

LASP 培训讲义

网页

<http://www.lasphub.com>

<http://homepage.fudan.edu.cn/fdzpliu/lasp/>

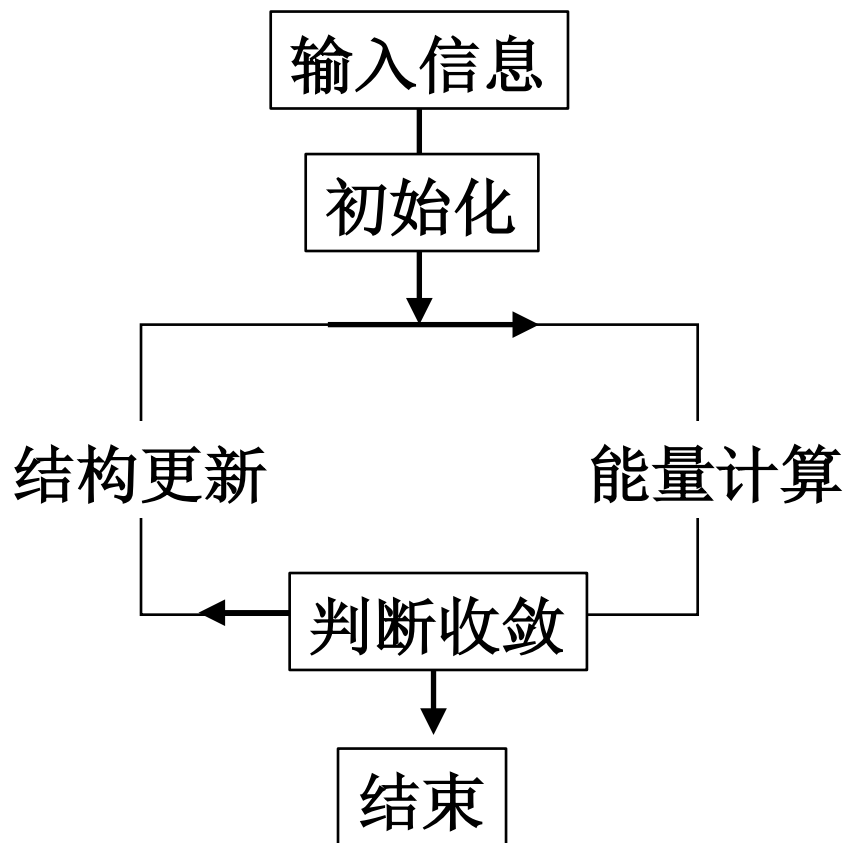
`ssh lasp_guest@10.158.134.250`

LASP 培训讲义 1 概述和程序接口

1. 软件基础功能和模块
2. 简单LASP计算：模块选择和输入输出
3. 程序实现和各部分模块接口
4. VASP等程序中的一些问题
5. 后续开发

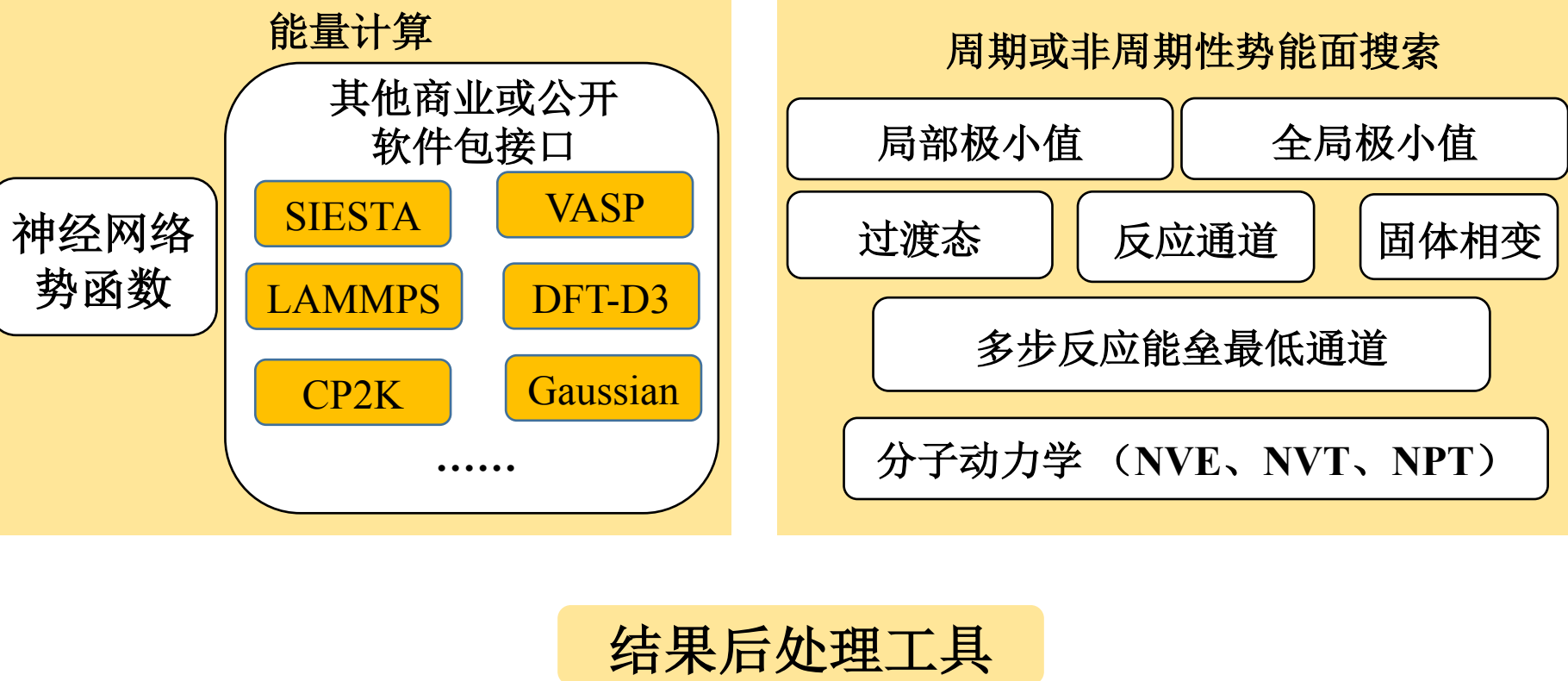
LASP 培训讲义 1 概述和程序接口

1. 软件基础功能和模块



LASP 培训讲义 1 概述和程序接口

1. 软件基础功能和模块



LASP 培训讲义 1 概述和程序接口

2. 简单LASP计算：模块选择和输入输出

- 在”lasp.in“文件中用 “**potential**” 关键词定义所使用的能量计算工具
(NN、 vasp、 siesta、 lammps、 d3、 gaussian、 cp2k)
(可以叠加: potential NN vasp siesta)
- 在”lasp.in“文件中用 “**explore_type**” 关键词定义所要进行的几何结构操作
(npt、 nvt、 nve、 ssw)
- 在SSW关键词中通过 “**RUN_TYPE**” 关键词选取SSW的功能
 - ◆ RUN_TYPE=1 固定晶胞单端法搜索过渡态
 - ◆ RUN_TYPE=2 固定晶胞双端法搜索反应路径
 - ◆ RUN_TYPE=5 固定晶胞全局结构搜索
 - ◆ RUN_TYPE=11 优化晶胞单端法搜索过渡态
 - ◆ RUN_TYPE=12 优化晶胞双端法搜索反应路径
 - ◆ RUN_TYPE=15 优化晶胞全局结构搜索

LASP 培训讲义 1 概述和程序接口

2. 简单LASP计算：模块选择和输入输出

必需文件：

- lasp.in （或input）：运行参数
- lasp.str （或input.arc）：初始结构

根据所选取的能量计算工具，准备相应的输入文件

NN： 势函数文件

VASP： INCAR POTCAR KPOINTS (POSCAR 会根据lasp.str文件生成)

Gaussian： gaussian.inp.pre, gaussian.inp.after

（源程序能否正常运行？）

输出文件： allkey.log, lasp.out, arc文件，外接程序的标准输出。

LASP 培训讲义 1 概述和程序接口

2. 简单LASP计算：模块选择和输入输出

输入example (lasp.in)

```
potential          cp2k
explore_type       ssw
MD.print_freq      100
MD.print_strfreq   100
Run_type           5
SSW.NG             9
DESW.ds            0.2000
DESW.ng            100
CBD.maxstep        20
CBD.maxcycle       40
CBD.NTS_repeat     1
%block             fixatommode
#                  empty
%endblock          fixatommode
```

自由格式

模块. 功能

参数

同一参数写多次，以第一遍为准

注释

LASP 培训讲义 1 概述和程序接口

2. 简单LASP计算：模块选择和输入输出

输入example (lasp.str)

```
!BIOSYM archive 3
PBC=ON
Materials Studio Generated CAR File
!DATE Sun Jul 22 14:41:11 2018
PBC 10 10 10 90.0000 90.0000 90.0000 (P1)
O 2.861839159 1.126221264 0.806771650 XXXX 1 xx O 0.000
H 3.788581204 1.221317879 0.483718885 XXXX 1 xx H 0.000
H 2.405079141 1.636760741 0.008699078 XXXX 1 xx H 0.000
end
end
```

➤ 可以通过Material Studio 生成

➤ 有效信息

- PBC a b c α β γ

- Element name x y z

LASP 培训讲义 1 概述和程序接口

2. 简单LASP计算：模块选择和输入输出

输出example(lasp.out,all.arc,allstr.arc,allfor.arc,allkey.log)

```
-----+
Welcome to Fantastic Atomic World in Neural Network
-----+
      L A S P
-----+
Large-scale Atomic Simulation with neural network Potential
Version 1.3   June 2018
      Shanghai, China
More on Website: www.lasphub.com

Compiled by IntelFortran version: ifort      12
CPU Parallel Run on   1      cores
-----+

+ The PES is calculated by
+ cp2k
+-----+

*****|
Welcome to SSW package
Stochastic Surface Walking Global Optimization 2.0
---- now as a part of LASP program
Fudan University Liu group 2013-2018
Developers: Cheng Shang, Xiaojie Zhang, Zhipan Liu
*****|

SSW Job Start at 00:41:15 22nd July      2018
System size: Number of atoms      8
Variable Cell Structure Search: Run_type=15
Key Parameters Utilized From Input:
Number of Gaussian for atomic move      9
*****|
```

LASP 培训讲义 1 概述和程序接口

2. 简单LASP计算：模块选择和输入输出

输出example(lasp.out,all.arc,allstr.arc,allfor.arc,allkey.log)

all.arc: 可以用MS读取

```
!BIOSYM archive 2
PBC=ON
Energy 0 0.0975 -1404.060589 C1
!DATE
PBC 10.00000000 10.00000000 10.00000000 90.00000000 90.00000000 90.00000000
O 2.890920223 1.173451973 0.856361342 CORE 1 O O 0.0000 1
H 3.826998565 1.171044212 0.547763441 CORE 2 H H 0.0000 2
H 2.424104792 1.637362086 0.083421092 CORE 3 H H 0.0000 3
end
end
```

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口

- **NN**、**FULL** 版本，可以直接运行，不需要编译
- **NSLV**版本，如使用VASP功能需自行编译vasp动态库
(xxx/Libraries/1.vasplib/genso.sh)
- **INTER**版本，自行编译所需软件的静态库，并编译主程序
(xxx/Libraries/1.vasplib/genlib.sh)
.....
(xxx/Src/makefile)
echo \$MKLROOT /home7/intel/compilers_and_libraries_2018.2.199/linux/mkl

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口 1.vasplib

```
genso.sh libvasp.so makefile.lib makefile.vasp
```

```
patch.vasp.tar.gz README
```

```
rm -rf vasp.5.3 vasp.5.lib
tar -xzf vasp.5.3.5.tar.gz
tar -xzf vasp.5.lib.tar.gz
cd vasp.5.lib
cp ../makefile.lib makefile
make dynamic=T
cd ..
lvasp=`pwd`
sed -i '/for_lasp_use/d' ~/.bashrc
echo export LD_LIBRARY_PATH=$lvasp:'$LD_LIBRARY_PATH #for_lasp_use' >> ~/.bashrc
cd vasp.5.3
for i in `ls *.F`; do sed -i 's/USE radial/USE radial_vasp/g' $i; done
tar -xzf ../patch.vasp.tar.gz
for i in `cat patch.vasp/patch.list`; do patch $i patch.vasp/$i.patch ;done
cp ../makefile.vasp ./makefile
if [ "$MKLROOT" != "" ]; then sed -i 's/\#for_lasp //g' makefile ; fi
make dynamic=T
rm -rf patch.vasp
if [ -f libvasp.so ]; then rm -rf ../libvasp.so;cp libvasp.so ../;fi
cd ..
```

echo \$MKLROOT

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口 INTERFACE version

```
version.sh*  
version.pre  
version.f90  
README  
readkey.F90  
otherpot.F90  
otherpes_init.F90  
othermove.F90  
modules/  
module_pes.F90  
makefile
```

version.sh

version.pre : 版本控制文件

version.f90

Readkey.F90 : 添加关键词

otherpes_init.F90 : 添加势函数初始化

otherpot.F90 : 添加势函数

othermove.F90 : 添加结构操作

module_pes.F90 : 关键参数传递

Makefile : 编译相关

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口

INTERFACE version

Read_key.F90、 module_pes.F90

```
Subroutine readkey()

use mpiinfo
use getkeys
use readkeys
use module_pes
use module_move
use module_md
use module_system
implicit none
integer::uinput=754212312
character(130),allocatable::keyall(:)
integer::nkey
character(100),allocatable::lines(:)

if(cpuid==0) then
  open(uinput,file=adjustl(trim(infn)))
  call ioassign(iokeys)
  open(iokeys,file=outkeyfn,status='replace')
  close(iokeys)
endif
call readkeys_init(uinput)
if(cpuid==0) close(uinput)

call get_string_array("potential",potential_type,npotential)
call get_string_array("lasp_pot",potential_type,npotential)
```

get_int	读整型
get_real	读双精度浮点数
get_string	读字符串
get_logical	读逻辑变量
get_int_array	整型数组
get_logical_array	逻辑数组
get_real_array	双精度数组
get_string_array	字符串数组
get_block	Block结构

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口 INTERFACE version

Read_key.F90、 module_pes.F90

```
Module module_pes
#ifdef LMP
  use LAMMPS
  use, intrinsic :: ISO_C_binding, only : C_ptr
#endif
implicit none
#ifdef LMP
  type (C_ptr) :: lmp
  double precision :: lmpeunit, lmpfunit, lmpeunit, lmpfunit, lmpeunit, lmpfunit
#endif
character(100), allocatable :: potential_type(:)
double precision, allocatable :: pesfact(:), fcurr(:)
double precision :: ecurr, stresscur(3,3)
integer :: npotential
double precision :: r0=1.0d0, k0=2d0, sig=2.5d0, epsl
logical :: Lbravcell=.true.
integer :: tall(10)
character(10) :: d3_func="pbe"
integer :: enemodify=0
integer :: iymd=0
End Module module_pes
```

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口 INTERFACE version

Otherpes_init.F90

```
select case (trim(adjustl(tolowc_main(potential_type(potin
case("vasp")
#ifdef VASP
  if(cpuid==0) then
    call system("sed '/^IBRION/d' INCAR >1; echo 'IBRION
    call system("sed -i '/^ISYM/d' 1; echo 'ISYM = 0' >>
    call system("sed -i '/^NSW/d' 1; echo 'NSW = 2000000
    call system("sed -i '/^ISIF/d' 1; echo 'ISIF = 3' >>
  endif
#endif
  case("siesta")
#ifdef SIESTA
  if(cpuid==0) then
    call system("sed '/^MD.TypeOfRun/d' siesta.fdf >1; echo
    call system("sed -i '/^DivideAndConquer/d' 1; echo 'Divi
    ! call system("rm -fr ./siesta")
  endif
#endif
#ifdef MPI
  call siesta_launch("siesta", mpi_comm=MPI_COMM_WORLD)
#else
  call siesta_launch("siesta")
#endif
  call siesta_units('Ang', 'eV')
```

```
case("gaussian")
  call ioassign(uindex)
  open(uindex,file=trim(adjustl(infn))
  call readkeys_init(uindex)
  call get_int("gaussian.charge",charge)
  call get_int("gaussian.multi",multi)
  call get_string("gaussian.exec",exec)
```

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口 INTERFACE version

Other_pot.F90

```
! write(6,*) 'VASP....2..'
call VAMP(strt%nalocal,strt%xcart,strt%fcart,strt%lat,strt%ene,strt%stress33,CBD_Reset_DM,.false.,.false)
! write(6,*) 'VASP....3..'
if(cpuid==0) call write_poscar(strt%nalocal,strt%fcart,strt%lat,1)
if(cpuid==0) call write_poscar(strt%nalocal,strt%xcart,strt%lat,1)
#ifdef MPI
call MPI_Bcast(strt%fcart,3*strt%nalocal,MPI_DOUBLE_PRECISION,0,MPI_COMM_WORLD,mpier)
#endif
strt%stress33=-strt%stress33
#endif
#ifdef SIESTA
case("siesta")
call siesta_forces('siesta', strt%nalocal, strt%xcart, strt%lat, strt%ene, strt%fcart, strt%stress33)
#endif
```

坐标

能量

张力(3,3)

原子数

晶胞(3,3)

力(3,n)

输入：原子数，坐标，晶胞

输出：能量，力，张力

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口 INTERFACE version

Othermove.F90

```
case("test")  
  strt%xcart=strt%xcart+strt%fcart*0d0  
  if(istep==teststeps) converge=.true.
```

新结构=旧结构+力×步长

收敛 converge=.true.

LASP 培训讲义 1 概述和程序接口

3. 程序实现和各部分模块接口

INTERFACE version

Makefile

```
# ===== Change here for MKL library linkage =====  
#write MKLROOT here if is not in your environment variables, and change the mkl-lib correspondingly  
MKLROOT =  
mkl-lib = ${MKLROOT}/lib/intel64/libmkl_scalapack_lp64.a -Wl,--start-group ${MKLROOT}/lib/intel64/libmkl_intel_lp64.a  
64/libmkl_core.a ${MKLROOT}/lib/intel64/libmkl_blacs_intelmpi_lp64.a -Wl,--end-group -lpthread -lm -ldl  
# =====  
mkl-lib += -lstdc++ ${MKLROOT}/interfaces/fftw3xf/libfftw3xf_intel.a  
  
FC      = mpiifort  
FFLAGS  = $(INC) -O2 -xHost -g -traceback #-check all -fp-stack-check  
INC      = -I$(LSSWOOP) -I./modules -I${MKLROOT}/include  
lib      = modules/lasplib.a $(LSSWOOP)/libssw.a modules/libp.a modules/libold.a modules/lib.s.a modules/libr.a  
  
FuncObj = otherpes_init.o readkey.o module_pes.o  otherpot.o othermove.o\
```

主要修改:

MKLROOT

mkl-lib

FuncObj

LASP 培训讲义 1 概述和程序接口

4. VASP等程序中的一些问题

1、INCAR

IBRION = 22

ISYM = 0

NSW = 2000000

ISIF = 3

2、晶胞变化需要reset VASP

	N	E	dE	d eps	ncg	rms	rms (c)
DAV:	1	-0.364986314967E+03	0.51371E-02	-0.57170E+00	1088	0.714E+00	0.594E-01
DAV:	2	-0.364994439456E+03	-0.81245E-02	-0.10932E-01	1256	0.984E-01	0.366E-01
DAV:	3	-0.364993450656E+03	0.98880E-03	-0.25005E-03	1344	0.221E-01	0.211E-01
DAV:	4	-0.364993375711E+03	0.74945E-04	-0.24220E-03	1288	0.169E-01	0.599E-02
DAV:	5	-0.364993341505E+03	0.34206E-04	-0.39643E-04	1320	0.843E-02	0.249E-02
DAV:	6	-0.364993348058E+03	-0.65537E-05	-0.72021E-05	1232	0.252E-02	0.154E-02
DAV:	7	-0.364993346751E+03	0.13075E-05	-0.75481E-06	1320	0.116E-02	
162 F= -.36499335E+03 E0= -.36499335E+03 d E =-.364993E+03							
	N	E	dE	d eps	ncg	rms	rms (c)
DAV:	1	-0.364989097086E+03	0.42510E-02	-0.45153E+00	1088	0.624E+00	0.486E-01
DAV:	2	-0.364995539682E+03	-0.64426E-02	-0.77846E-02	1264	0.839E-01	0.302E-01
DAV:	3	-0.364994888227E+03	0.65146E-03	-0.16409E-03	1344	0.170E-01	0.175E-01
DAV:	4	-0.364994824367E+03	0.63860E-04	-0.14473E-03	1264	0.131E-01	0.452E-02
DAV:	5	-0.364994824475E+03	-0.10826E-06	-0.23517E-04	1312	0.622E-02	0.218E-02
DAV:	6	-0.364994827203E+03	-0.27276E-05	-0.34487E-05	1232	0.182E-02	0.145E-02
DAV:	7	-0.364994825907E+03	0.12959E-05	-0.79712E-06	1312	0.115E-02	
163 F= -.36499483E+03 E0= -.36499483E+03 d E =-.364995E+03							
	N	E	dE	d eps	ncg	rms	rms (c)
DAV:	1	0.126074675393E+04	0.16257E+04	0.00000E+00	816	0.000E+00	
DAV:	2	-0.271656103243E+03	-0.15324E+04	-0.14768E+04	1016	0.327E+02	
DAV:	3	-0.400703722863E+03	-0.12905E+03	-0.12803E+03	1008	0.995E+01	
DAV:	4	-0.403033158412E+03	-0.23294E+01	-0.23216E+01	992	0.144E+01	
DAV:	5	-0.403091741786E+03	-0.58583E-01	-0.58520E-01	1016	0.206E+00	0.382E+01
DAV:	6	-0.366883112033E+03	0.36209E+02	-0.10223E+02	976	0.271E+01	0.159E+01
DAV:	7	-0.365351912843E+03	0.15312E+01	-0.78617E+00	992	0.818E+00	0.102E+01
DAV:	8	-0.364955136530E+03	0.39678E+00	-0.22092E+00	984	0.450E+00	0.187E+00
DAV:	9	-0.364917986545E+03	0.37150E-01	-0.15191E-01	976	0.138E+00	0.656E-01
DAV:	10	-0.364914654545E+03	0.33320E-02	-0.49967E-02	920	0.777E-01	0.193E-01
DAV:	11	-0.364914701785E+03	-0.47239E-04	-0.29119E-03	1008	0.192E-01	0.970E-02
DAV:	12	-0.364914826571E+03	-0.12479E-03	-0.17461E-03	912	0.172E-01	0.289E-02
DAV:	13	-0.364914850062E+03	-0.23492E-04	-0.76817E-05	1016	0.390E-02	0.149E-02
DAV:	14	-0.364914857138E+03	-0.70762E-05	-0.32484E-05	920	0.221E-02	0.707E-03
DAV:	15	-0.364914857114E+03	0.23796E-07	-0.60052E-06	976	0.895E-03	
1 F= -.36491486E+03 E0= -.36491486E+03 d E =-.364915E+03							
	N	E	dE	d eps	ncg	rms	rms (c)
DAV:	1	-0.364186942647E+03	0.72791E+00	-0.10286E+02	864	0.314E+01	0.483E+00
DAV:	2	-0.364020048678E+03	0.16689E+00	-0.34136E+00	1016	0.568E+00	0.221E+00
DAV:	3	-0.363936361222E+03	0.83687E-01	-0.26116E-01	1000	0.169E+00	0.103E+00
DAV:	4	-0.363923708548E+03	0.12653E-01	-0.52160E-02	1000	0.785E-01	0.404E-01
DAV:	5	-0.363922201221E+03	0.15073E-02	-0.75234E-03	1000	0.336E-01	0.142E-01

LASP 培训讲义 1 概述和程序接口

4. VASP等程序中的一些问题

3、siesta.fdf

输入文件名字不可更改

MD.TypeOfRun Forces

DivideAndConquer False

INPUT_DEBUG中的元素顺序要与lasp.str 保持一致

4、Gaussian

串行运行lasp，内部并行

LASP 培训讲义 1 概述和程序接口

5. 后续功能开发

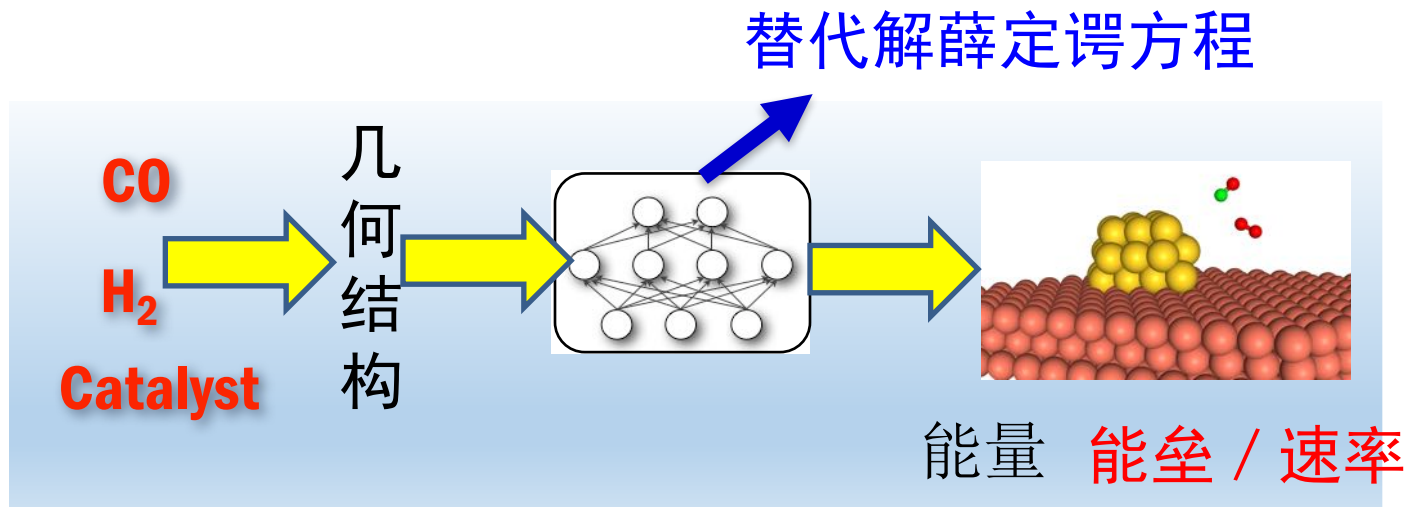
- 其他程序包接口
- 统一输入结构文件
- 通用程序接口

LASP 培训讲义 2 神经网络势函数

1. 理论基础
2. 势函数文件格式
3. LASP中的NN计算
4. NN模拟注意事项
5. 积累数据和NN拟合

<http://homepage.fudan.edu.cn/fdzpliu/publication/>

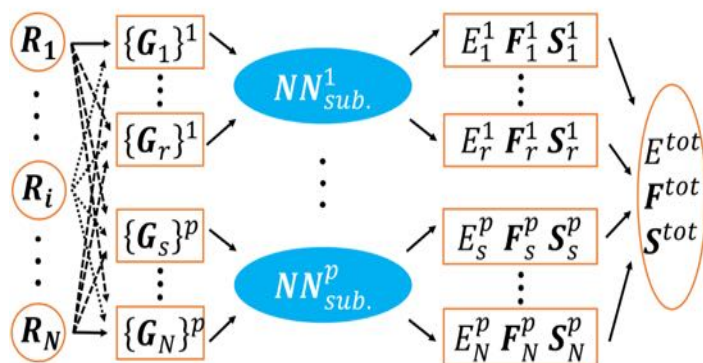
LASP 培训讲义 2 神经网络势函数



关键技术问题

- 实现高效PES预测
- (反应) 大数据 学习
 - 结构特征提取 学习什么
 - 势函数导数连续性 学习结果

LASP 培训讲义 2 神经网络势函数



High-Dimensional NN (HDNN)

1. 总能是原子能量加和
2. 原子的化学环境由原子间长度和角度的函数确定
3. 原子的化学环境可以区分元素
4. 标准的前馈神经网络模型

$$E^{tot} = \sum_i E_i$$

J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**, 146401.

LASP 培训讲义 2 神经网络势函数

Behler 提出的高斯类型的结构描述函数

Rotation-invariant Symm func

Cutoff function--- $\tan^3 h$ 函数

Radial function---Gauss 函数

Angular function--- 三角函数

两体，三体 函数

$$f_c(R_{ij}) = \begin{cases} 0.5 \times \tanh^3 \left[1 - \frac{R_{ij}}{R_c} \right] & \text{for } R_{ij} \leq R_c \\ 0 & \text{for } R_{ij} > R_c \end{cases} \quad (2)$$

$$G_i^{1,p} = \sum_j f_c(R_{ij}), \quad (3)$$

$$G_i^{2,p} = \sum_j e^{-\eta(R_{ij}-R_s)^2} \cdot f_c(R_{ij}), \quad (4)$$

$$G_i^{3,p} = \sum_j \cos(\kappa R_{ij}) \cdot f_c(R_{ij}). \quad (5)$$

$$G_i^{4,p} = 2^{1-\zeta} \sum_{j,k \neq i}^{all} (1 + \lambda \cos \theta_{ijk})^\zeta \cdot e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \cdot f_c(R_{jk}), \quad (6)$$

$$G_i^{5,p} = 2^{1-\zeta} \sum_{j,k \neq i}^{all} (1 + \lambda \cos \theta_{ijk})^\zeta \cdot e^{-\eta(R_{ij}^2 + R_{ik}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}), \quad (7)$$

LASP 培训讲义 2 神经网络势函数

LASP 采用的指数函数类型的结构描述函数

$$f_c(R_{ij}) = \begin{cases} 0.5 \times \tanh^3 \left[1 - \frac{r_{ij}}{r_c} \right], & \text{for } r_{ij} \leq r_c \\ 0 & \text{for } r_{ij} > r_c \end{cases} \quad (1)$$

$$R^n(r_{ij}) = r_{ij}^n \cdot f_c(r_{ij}), \quad (2)$$

$$s_i^1 = \sum_{j \neq i} \text{GU}_1 = \sum_{j \neq i} R^n(r_{ij}), \quad (3)$$

$$s_i^2 = \left[\sum_{m=-L}^L \left| \sum_{j \neq i} \text{GU}_2 \right|^2 \right]^{\frac{1}{2}}$$

$$= \left[\sum_{m=-L}^L \left| \sum_{j \neq i} R^n(r_{ij}) Y_{Lm}(r_{ij}) \right|^2 \right]^{\frac{1}{2}} \quad (4)$$

$$s_i^3 = 2^{1-\zeta} \sum_{j,k \neq i} \text{GU}_3 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos \theta_{ijk})^\zeta$$

$$\cdot R^n(r_{ij}) \cdot R^m(r_{ik}) \cdot R^p(r_{jk}), \quad (5)$$

$$s_i^4 = 2^{1-\zeta} \sum_{j,k \neq i} \text{GU}_4 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos \theta_{ijk})^\zeta$$

$$\cdot R^n(r_{ij}) \cdot R^m(r_{ik}), \quad (6)$$

Cutoff function--- $\tan^3 h$ 函数

Radial function--- Power函数 和球谐函数

Angular function--- 三角函数

$$s_i^5 = \left[\sum_{m=-L}^L \left| \sum_{j,k \neq i} \text{GU}_5 \right|^2 \right]^{\frac{1}{2}}$$

$$= \left[\sum_{m=-L}^L \left| \sum_{j,k \neq i} R^n(r_{ij}) \cdot R^m(r_{ik}) \cdot R^p(r_{jk}) \right. \right. \\ \left. \left. \cdot (Y_{Lm}(\mathbf{r}_{ij}) + Y_{Lm}(\mathbf{r}_{ik})) \right|^2 \right]^{\frac{1}{2}}, \quad (7)$$

两体，三体，四体函数

$$s_i^6 = 2^{1-\zeta} \sum_{j,k,l \neq i} \text{GU}_6 = 2^{1-\zeta} \sum_{j,k,l \neq i} (1 + \lambda \cos \delta_{ijkl})^\zeta$$

$$\cdot R^n(r_{ij}) R^m(r_{ik}) R^p(r_{il}), \quad (8)$$

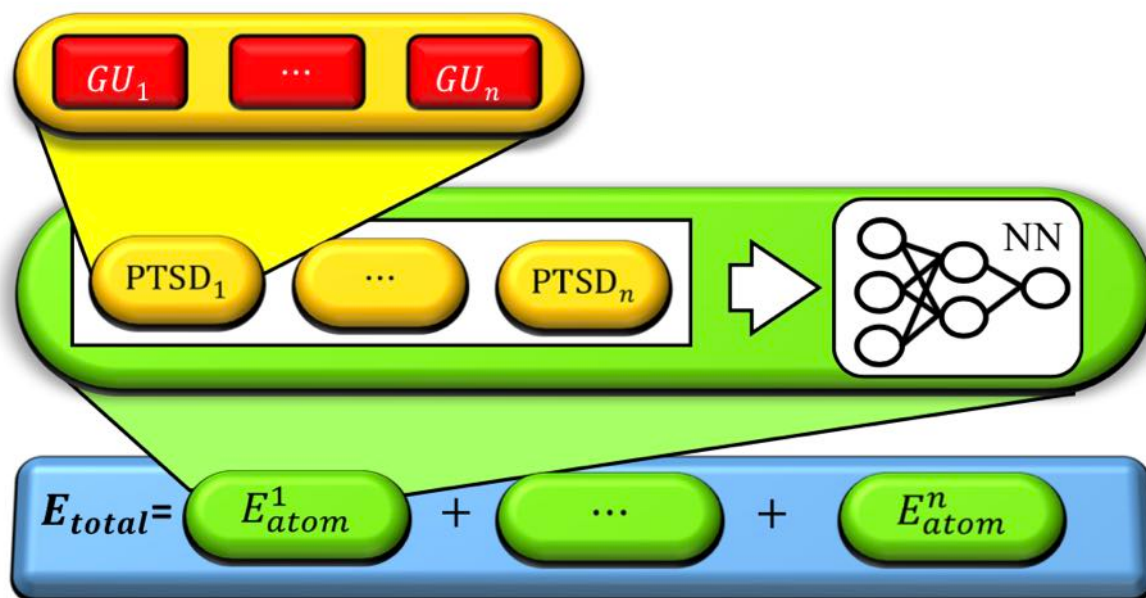
LASP 培训讲义 2 神经网络势函数

LASP NN计算的并行架构

原子并行

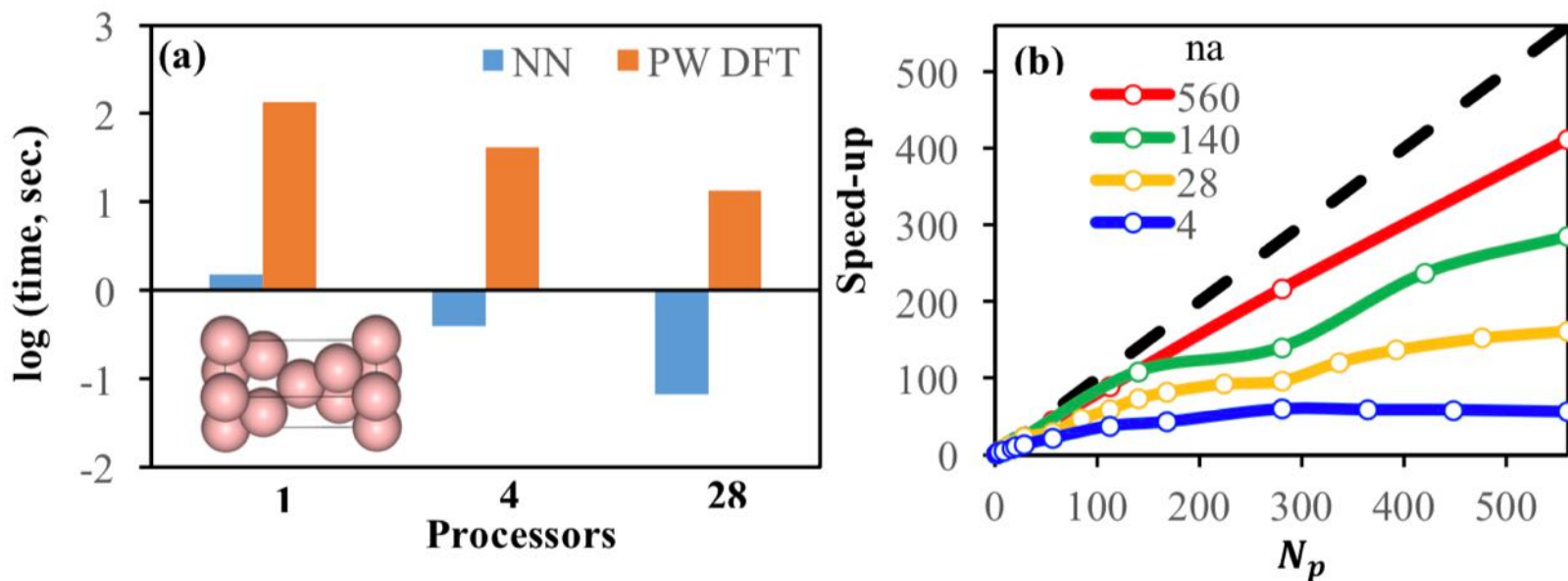
结构描述函数并行

多体函数并行



LASP 培训讲义 2 神经网络势函数

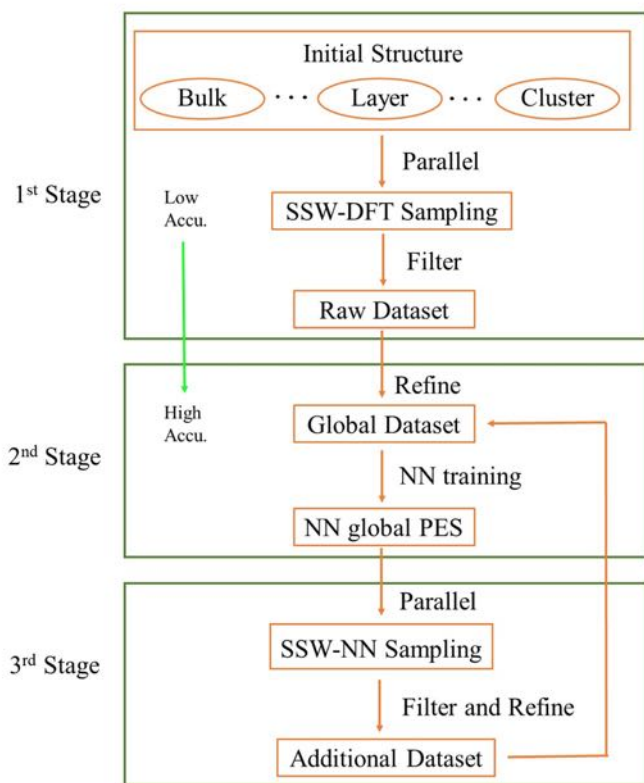
LASP NN计算的并行效率



(a) Speed-up comparison of a 28-atom Boron crystal between NN and DFT using Plane wave basis set (PW DFT). The x axis is the logarithm of time in the unit of seconds. The inset shows the structure of the lattice. (b) Speed-up of Boron crystals on 560 cores. The x axis (N_p) is the number of processors and the y axis is the speed-up.

LASP 培训讲义 2 神经网络势函数

SSW 第一性原理势能面大数据 + NN 拟合 == NN 势函数



数据集产生

1. **VASP**: 平面波 DFT + PBE (+U) 泛函

2. **SSW 全局势能面**数据
一般没有高温MD势能面数据

3. **自动训练**, 直到SSW能顺利进行

- 结构描述函数不越界
- DFT验证两次能量误差在势函数RMS误差范围

LASP 培训讲义 2 神经网络势函数

1. 理论基础
2. 势函数文件格式
3. LASP中的NN计算
4. NN模拟注意事项
5. 积累数据和NN拟合

LASP 培训讲义 2 神经网络势函数

```
release_Sn0.pot
release_Pd0.pot
release_Si0H.pot
release_Ag0.pot
release_Rh0.pot
release_Ti0H.pot
release_Cu0.pot
release_Co0.pot
release_Au0.pot
release_PtNi.pot
release_Pt0_1st.pot
release_Rh0_1st.pot
release_Co0_1st.pot
release_B.pot
release_Ti0.pot
release_Pd0_1st.pot
release_Si.pot
release_H2O.pot
release_Ru0.pot
release_B_solid.pot
release_C.pot
release_Ir0.pot
release_GST.pot
release_Zr0.pot
release_P.pot
README
NNpot_lib.tgz
```

文件命名规则

release_AB(_x).pot

A 和B: 元素名

x: 前几次的拟合 (不建议)

原则上

1. 可以研究不同配比 如 TiO_2 , Ti_2O_3 , Ti
2. 可以研究固体, 表面和团簇

实际使用:

1. 关键数据需要通过DFT再验证, 比如小体系先测试
2. 团簇预测难度高, 一般需要针对具体体系再训练

LASP 培训讲义 2 神经网络势函数

Neural Network Potential For LASP

Reference : www.lasphub.com

Name : TiOx (Ti metal to TiOx autotrained)

Species : 0 Ti

Version : 20180616232752

Dataset size : 30022

NN 产生时间日期

数据集大小

Fitting Accuracy

	Energy/(meV/atom)	Force/(eV/A)	Stress/(GPa)
RMS	8.624	0.154	2.130
MAX	317.526	3.476	81.790

RMS 精度

Element 0

Neural Network Architecture

%block 0_netinfo

# input	147	2	float	newrun
full	50	2	float	newrun
full	50	2	float	newrun
full	1	0	float	newrun

%endblock 0_netinfo

前馈网络

网络大小 147-50-50-1

网络越大计算越慢

一般来说, NN计算主要代价是结构描述函数计算

LASP 培训讲义 2 神经网络势函数

前馈网络

结构描述子 个数列表

PowerType Structure Descriptors						
	S1	S2	S3	S4	S5	S6
SDType*	two	two	three	three	three	four
Number	52	32	15	24	18	6

*: Structural Descriptors type, being two-body, three-body and four-body.

结构描述函数具体的参数

%block	C_S1				
1.20	1	8	0.00000000000000e+00	2.66162625000000e-02	# 1
1.40	1	2	0.00000000000000e+00	1.28365840909091e-01	# 2
1.80	1	4	0.00000000000000e+00	4.21399522727273e-01	# 3
1.20	6	8	0.00000000000000e+00	1.39950204545455e-02	# 4
1.40	6	2	0.00000000000000e+00	6.86706488636364e-02	# 5
1.80	6	4	0.00000000000000e+00	2.40084659090909e-01	# 6
1.20	8	8	0.00000000000000e+00	9.97279102272727e-03	# 7
1.40	8	2	0.00000000000000e+00	5.31754943181818e-02	# 8
1.80	8	4	0.00000000000000e+00	2.44338636363636e-01	# 9
1.20	7	8	0.00000000000000e+00	1.39493272727273e-02	# 10
1.40	7	2	0.00000000000000e+00	7.06525000000000e-02	# 11

结构描述函数的函数范围
(min Scale ---- max Scale)

LASP 培训讲义 2 神经网络势函数

1. 理论基础
2. 势函数文件格式
3. **LASP中的NN计算**
4. NN模拟注意事项
5. 积累数据和NN拟合

LASP 培训讲义 2 神经网络势函数

```
21:41:04:[zpliu@storage3 Examples]$ ll
total 24K
drwxr-xr-x  2 root root 4.0K Jun  5 15:14 allexamples/
drwxr-xr-x  9 root root 4.0K Jun  5 15:14 VASP/
drwxr-xr-x  4 root root 4.0K Jun  5 15:14 LAMMPS/
-rw-r--r--  1 root root 3.9K Jun  5 15:14 README
drwxr-xr-x 12 root root 4.0K Jun  5 15:14 NN/
drwxr-xr-x  6 root root 4.0K Jun  5 15:14 SIESTA/
```

LASP Examples
www.lasphub.com 下载

LASP输入文件 lasp.in (或input)

NN计算中的关键词

potential NN

```
%block netinfo
  O CHON.pot
  H CHON.pot
  C CHON.pot
%endblock netinfo
```

元素名和文件名

PrintChg 0

0/1 debug, 输出原子能量

ExSymR 5

结构描述函数允许的越界范围 (缺省5倍)

LASP 培训讲义

LASP主输出文件 `lasp.out`

NN计算 应该出现在lasp.out头部

```
+-----+  
+  The PES is calculated by  
+  nn  
+-----+
```

LASP输出文件

`Badstr.arc`

输出NN计算中可能遇到的“预测出错”结构（arc 格式）

1. 结构描述函数范围越界 `ExSymR` 5
2. 势能面上的频率超过范围 `SSW.Safe_hardcurv` 150.0000
3. 势能面上的优化未收敛的结构（不一定是NN的问题）

LASP 培训讲义 2 神经网络势函数

1. 理论基础
2. 势函数文件格式
3. LASP中的NN计算
4. NN模拟注意事项
5. 积累数据和NN拟合

LASP 培训讲义

1. NN效率

目前原子并行效率最高：原子数=核数

2. NN基于VASP DFT/PBE的精度

后加 D3 等 vdW修正

数据集一般是30个原子左右的第一性原理数据集，对于大体系，绝对能量会有偏差，需要检查相对能量

可以用于寻找过渡态等，但是一般过渡态精度低于稳态

原则上可以用于MD等高温动力学研究

高温结构（高于熔点的液体）预测精度可能会降低

LASP 培训讲义

3. 利用NN势函数学习/发展PES方法，熟练各种参数

全局优化

过渡态寻找

常见分子动力学 (MD)

尝试编写自己的模块

(biased MD, metadynamics, MC, 遗传算法)

4. 验证NN的计算结果；针对目标体系用NN快速构建数据集，后续NN再训练

最低的结构，最优的反应路径：对比实验

频率计算：ZPE等的合理性

LASP 培训讲义 2 神经网络势函数

1. 理论基础
2. 势函数文件格式
3. LASP中的NN计算
4. NN模拟注意事项
5. 积累数据和NN拟合

LASP 培训讲义 2 神经网络势函数

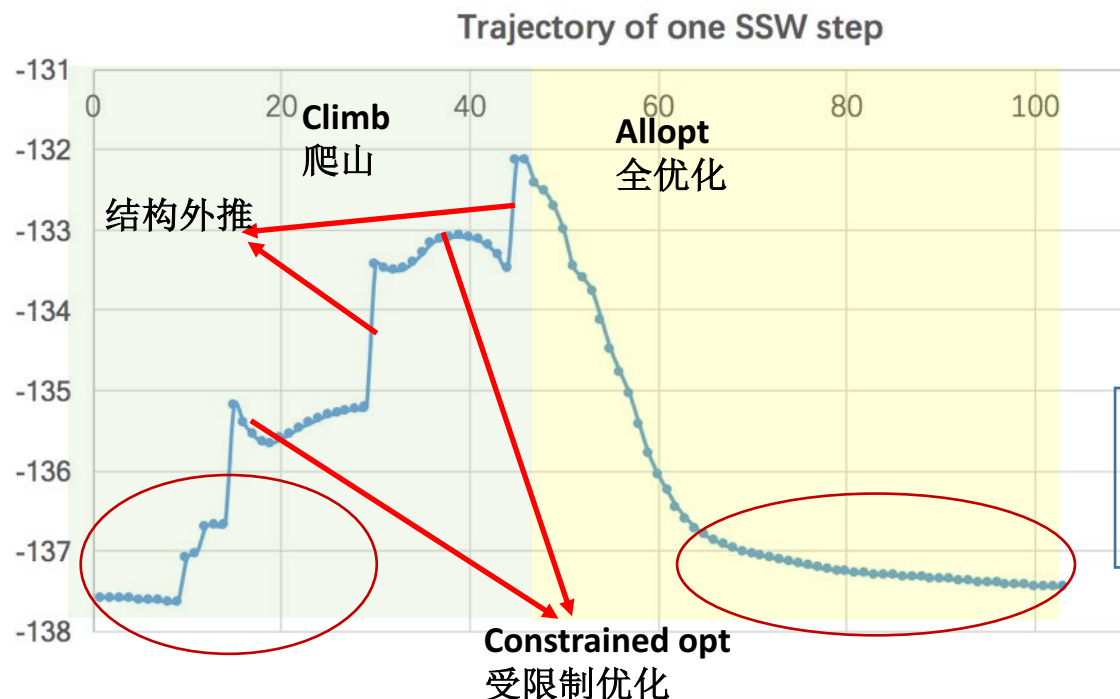
参见LASP使用入门 5

关键1: 利用NN进行 SSW计算时, 打开结构/能量/力输出选项 **SSW.printevery T**



allstr.arc 结构/能量

allfor.arc 力/应力

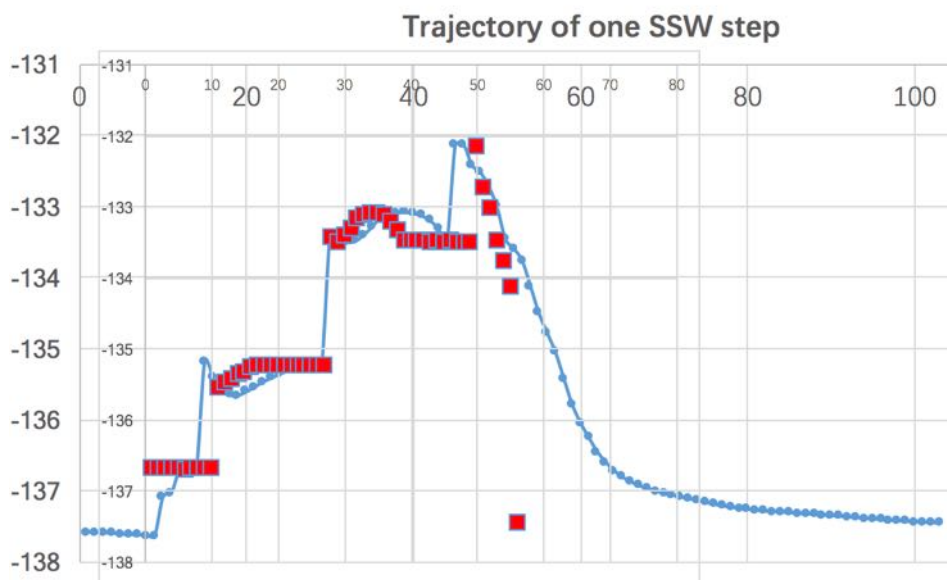


从allstr.arc中,
grep Energy allstr.arc
简单画出SSW轨迹 (左图)

```
SSW.printevery T  
SSW.printselect 1  
SSW.printdelay -1
```

LASP 培训讲义 2 神经网络势函数

只有对势能面**有效均匀采样**, 势函数才能够具有更好**可移植性**和**反应预测性**, 建议选择 **SSW.printselect 6**



```
SSW.printevery T  
SSW.printselect 6  
SSW.printdelay -1
```

改变SSW.printselect, 会大大减少**SSW**轨迹取点

如图, 红色点, 相比蓝色点

1. 去掉能量低区间点
2. 结构外推处插值取点 (由于是插值的, 能量没有真正计算, 看起来还是有能量跳跃)

LASP 培训讲义 2 神经网络势函数

关键2：针对目标体系构建数据集

- 控制原子数（每晶包），但是需要有相似的局部结构特征
- 可以从Badstr.arc文件寻找低能量的结构 作为SSW的初始结构
- 可以从DFT验证结果中，预测最差的结构作为SSW的初始结构

LASP 培训讲义 2 神经网络势函数

关键3: 拟合用数据集大小 ---- 尽量小（高度代表性）

从巨量的SSW轨迹结构中取十分之一到百分之一数据

按照20个SSW目录，每个目录100个SSWsteps
(SSW.SSWsteps 100)

预期产生 $20 \times 100 \times 200 \times 0.4 = 16$ 万数据
随机取其中5000-10000数据，

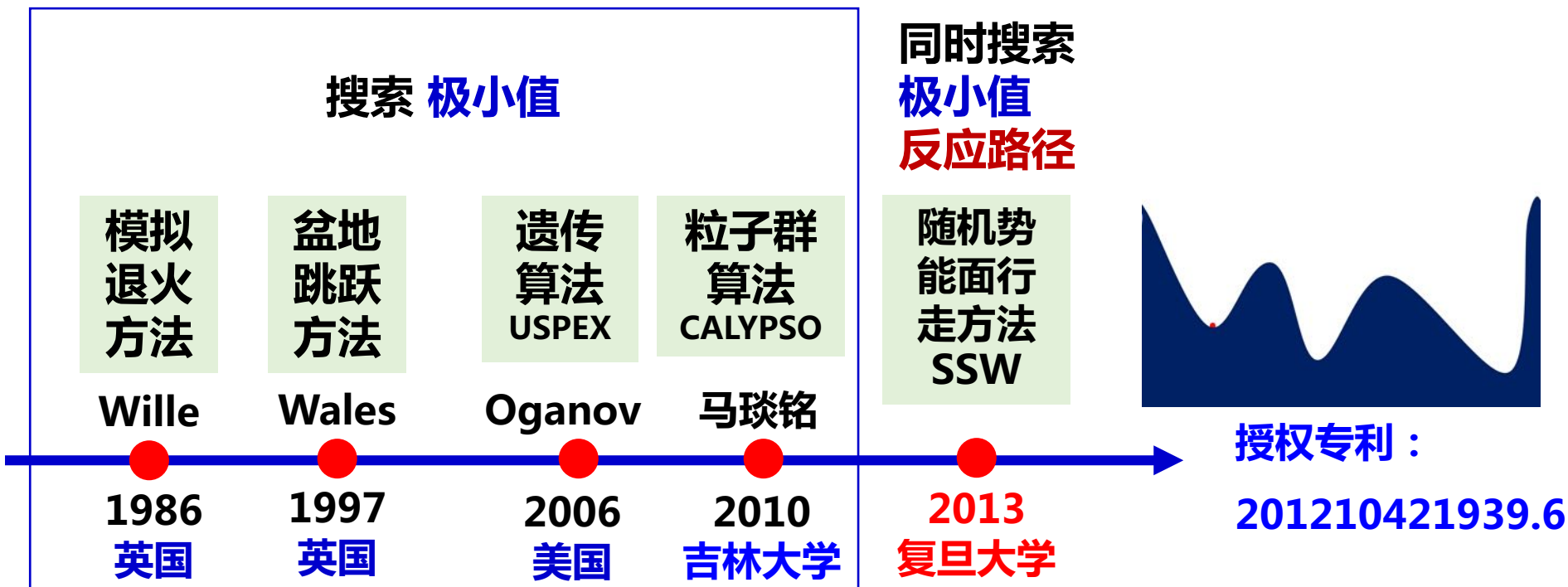
进行DFT高精度单点计算（需要联系LASP团队取得原始NN拟合采用VASP的setup 文件，INCAR，KPOINTS，POTCAR等

反馈数据给LASP团队，再次拟合（一般比较容易）

LASP 培训讲义 3 全局搜索和反应路径

1. 理论基础
2. 固定晶包和变化晶包
3. 全局搜索
4. 反应路径搜索
5. 固体相变与异相结构

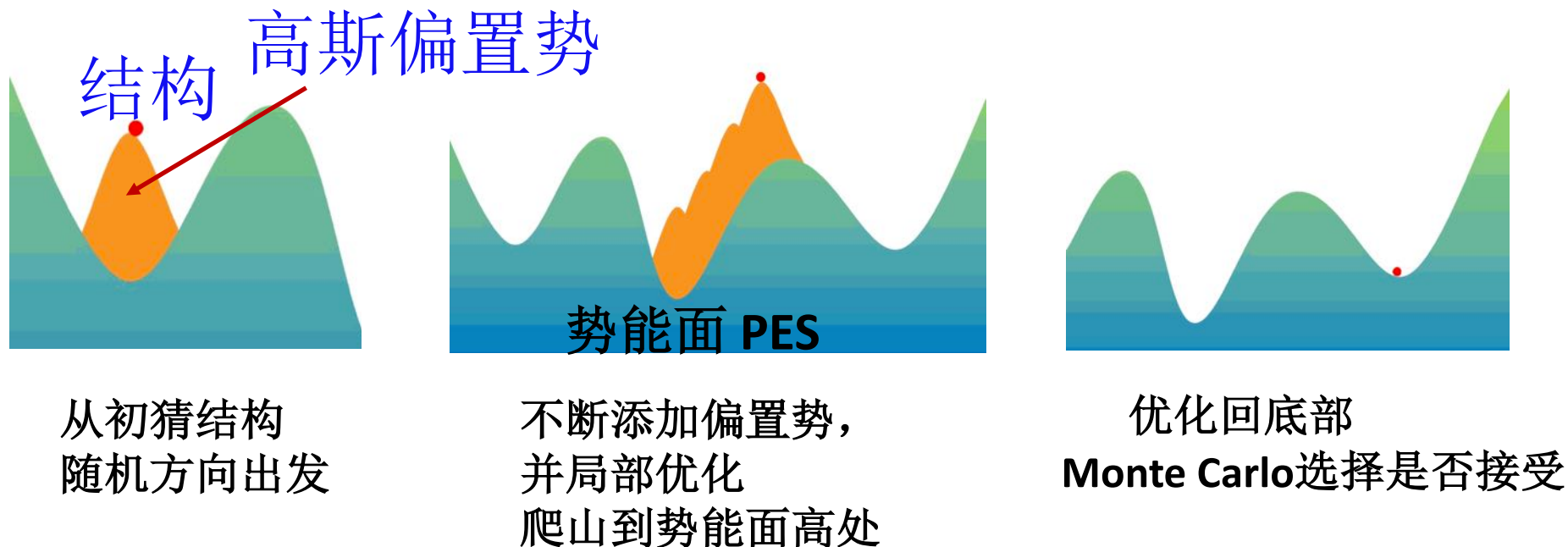
LASP 培训讲义 3 全局搜索和反应路径



SSW的特点 引入二阶梯度, 轨迹连续, 采样更均匀

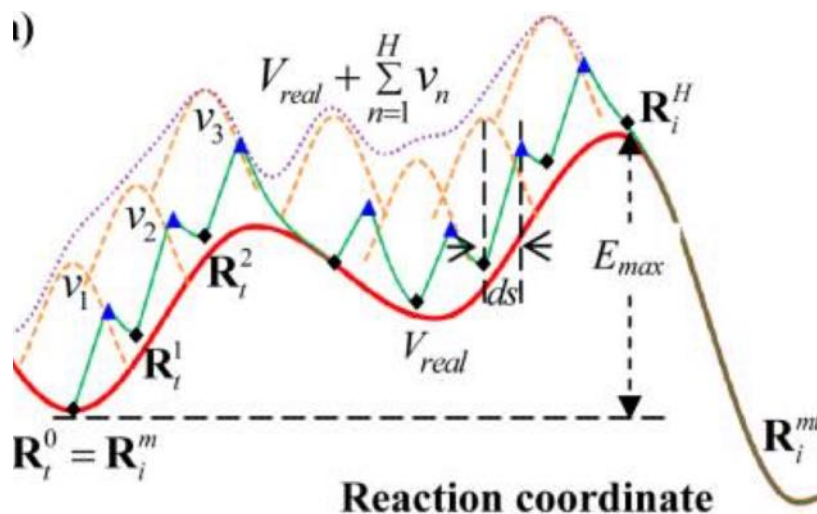
LASP 培训讲义 3 全局搜索和反应路径

随机表面行走 (SSW) 工作原理



J. Chem. Theory Comput, 2013, 9, 1838

LASP 培训讲义 3 全局搜索和反应路径



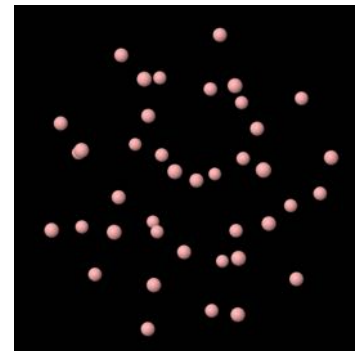
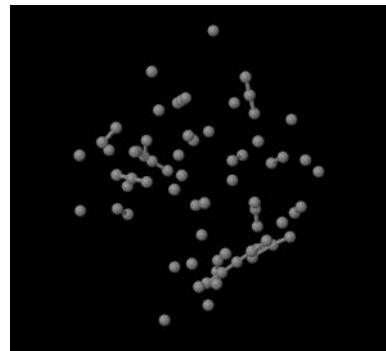
$$\begin{aligned}
 V_{m-to-H} &= V_{\text{real}} + \sum_{n=1}^H v_n \\
 &= V_{\text{real}} + \sum_{n=1}^H w_n \times \exp\left[-\frac{((\mathbf{R}^t - \mathbf{R}_t^{n-1}) \cdot \mathbf{N}_i^n)^2}{2 \times ds^2}\right]
 \end{aligned}$$

$$\begin{aligned}
 F_{\text{tot}} &= F_{\text{real}} + \sum_n w_n \cdot \exp\left[-\frac{((\mathbf{R}^t - \mathbf{R}_t^{n-1}) \cdot \mathbf{N}_i^n)^2}{2 \times ds^2}\right] \\
 &\quad \frac{(\mathbf{R}^t - \mathbf{R}_t^{n-1}) \cdot \mathbf{N}_i^n}{ds^2} \cdot \mathbf{N}_i^n
 \end{aligned}$$

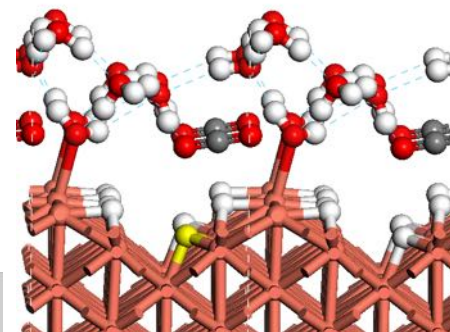
LASP 培训讲义

特色

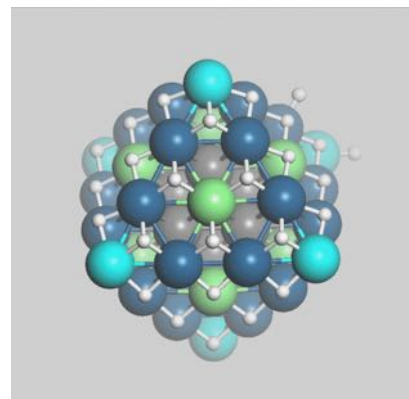
共价成键体系



分子，表面） 化学反应

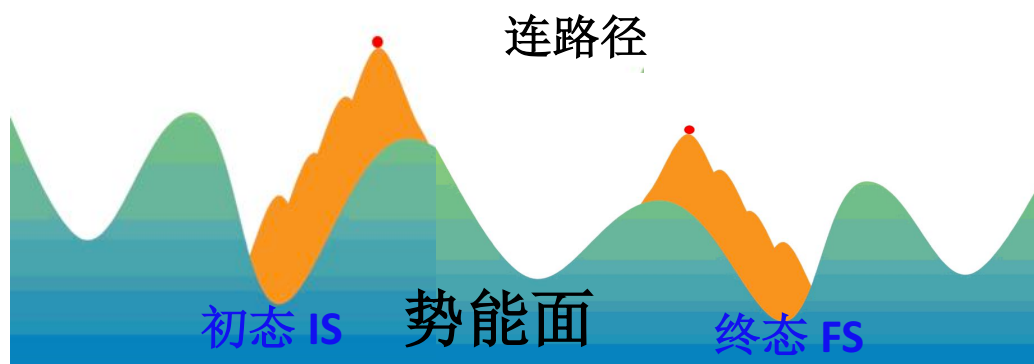


固体相变，
缺陷， 重构

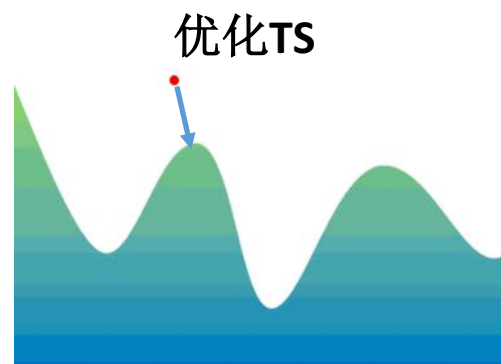


LASP 培训讲义 3 全局搜索和反应路径

两点行走法（DESW）工作原理



分别从初态和终态添加偏置势，
并局部优化，直到两个结构相遇

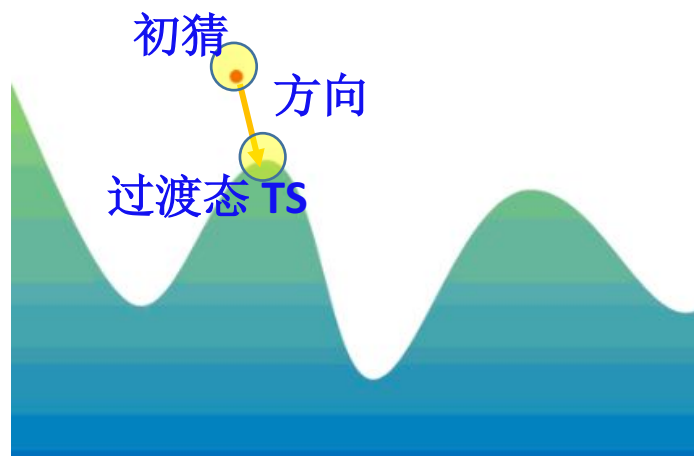


从行走轨迹中取最高点，
优化到过渡态

J. Chem. Theory Comput., 2013, 9, 5745

LASP 培训讲义 3 全局搜索和反应路径

单点行走法（CBD）工作原理



J. Chem. Theory Comput, 2010, 6, 1136

J. Chem. Theory Comput, 2012, 8, 2215

LASP 培训讲义 3 全局搜索和反应路径

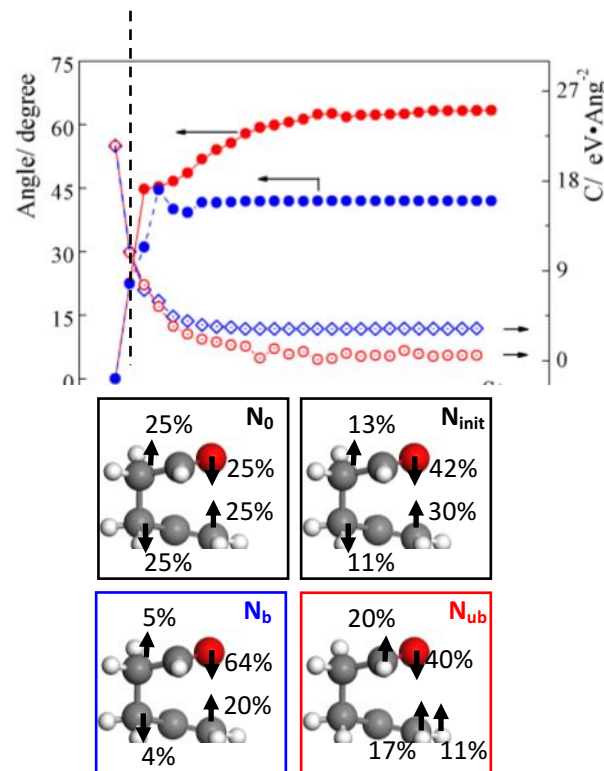
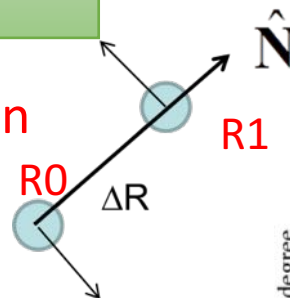
随机模式（频率）优化工作原理

Biased dimer rotation

$$V_{R1} = V_{real} + V_N$$

$$V_N = -\frac{a}{2} \cdot [(\mathbf{R}_1 - \mathbf{R}_0) \cdot \mathbf{N}_{init}]^2 = -\frac{a}{2} \cdot (\Delta \mathbf{R} \cdot \mathbf{N}_t \cdot \mathbf{N}_{init})^2$$

$$\begin{aligned} C_{rot} &= C_e + C_N \\ &= \frac{(\mathbf{F}_0 - \mathbf{F}_N - \mathbf{F}_1) \cdot \mathbf{N}_t}{\Delta R} \\ &= \frac{(\mathbf{F}_0 - \mathbf{F}_1) \cdot \mathbf{N}_t}{\Delta R} - a \cdot (\mathbf{N}_t \cdot \mathbf{N}_{init})^2 \end{aligned}$$



J. Chem. Theory Comput, 2010, 6, 1136

LASP 培训讲义 3 全局搜索和反应路径

1. 理论基础
2. 固定晶包和变化晶包
3. 全局搜索
4. 反应路径搜索
5. 固体相变与异相结构

下面通过
NN-1...NN-12
例子讲解

LASP 培训讲义 4 程序的后处理和分析

1. 结构后处理：动画和反应模式
2. 结构区分和表征：对称性和有序参量
3. 全局势能面图
4. 反应路径图

LASP 培训讲义 4 程序的后处理和分析

1. 结构后处理：动画和反应模式

LASP_Example/NN/solid-TS-VCDESW

➤ TSmode.arc : 过渡态对应的振动模式

➤ SSWPath.arc DESW的pseudo-pathway 大概反应通道

LASP_Example/NN/molecule-ConfirmPathway-ExtrapolatefromTS

➤ allstr.arc、all.arc: SSW或Opt的轨迹

LASP 培训讲义 4 程序的后处理和分析

2. 结构区分和表征：对称性和有序参量

Steinhardt-type:

$$OP_l = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^l |\overline{Y_{lm}(\mathbf{n})}|^2 \right)^{1/2}$$

Used in SSW:

$$OP_l = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^l \left| \frac{1}{N_{bonds}} \sum_{i \neq j} e^{-\frac{1}{2} \frac{r_{ij} - r_c}{r_c}} Y_{lm}(\mathbf{n}) \right|^2 \right)^{1/2}$$

LASP 培训讲义 4 程序的后处理和分析

2. 结构区分和表征：对称性和有序参量

LASP_Example/LAMMPS/varcell-ssw-eam/lasp.out

```
Stru symm and Q      981      0 longQ T      0.02088      0.06169      0.30327      65 Cmmn
Stru symm and Q      981      0 longQ T      0.02142      0.04654      0.10513      63 Cmmn
minimum found        981      -162.27976      -161.87299      100.0 K      63 Cmmn      8.4459      9.27 F      0.09      38      195
Stru symm and Q      982      0 longQ T      0.02043      0.06165      0.30336      65 Cmmn
Stru symm and Q      982      0 longQ T      0.04778      0.05826      0.18477      1 P1
minimum found        982      -162.27976      -158.64177      100.0 K      1 P1      5.7600      5.24 F      0.10      26      181
Stru symm and Q      983      0 longQ T      0.02087      0.06167      0.30327      65 Cmmn
Stru symm and Q      983      0 longQ T      0.01059      0.05675      0.25888      65 Cmmn
minimum found        983      -162.27976      -162.06630      100.0 K      65 Cmmn      8.3799      8.34 T      0.59      57      216
Stru symm and Q      984      0 longQ T      0.01059      0.05675      0.25888      65 Cmmn
```

\$grep found lasp.out

```
minimum found        976      -162.09844      -159.28570      100.0 K      11 P2_1/m      5.7600      4.21 F      0.10      8      159
minimum found        977      -162.09844      -160.06992      100.0 K      12 C2/m      8.4151      6.18 F      0.10      8      160
minimum found        978      -162.09844      -159.51798      100.0 K      1 P1      5.7600      5.23 F      0.09      16      173
minimum found        979      -162.09844      -162.27976      100.0 K      65 Cmmn      8.3989      7.71 T      0.10      90      257
minimum found        980      -162.27976      -158.04448      100.0 K      1 P1      5.7600      8.32 F      0.09      15      204
minimum found        981      -162.27976      -161.87299      100.0 K      63 Cmmn      8.4459      9.27 F      0.09      38      195
minimum found        982      -162.27976      -158.64177      100.0 K      1 P1      5.7600      5.24 F      0.10      26      181
minimum found        983      -162.27976      -162.06630      100.0 K      65 Cmmn      8.3799      8.34 T      0.59      57      216
minimum found        984      -162.06630      -158.07858      100.0 K      1 P1      5.7600      4.68 F      0.10      14      175
minimum found        985      -162.06630      -162.31723      100.0 K      221 Pm-3m      8.3419      7.73 T      0.48      29      163
minimum found        986      -162.31723      -162.36348      100.0 K      221 Pm-3m      5.7600      3.88 T      0.17      20      177
```

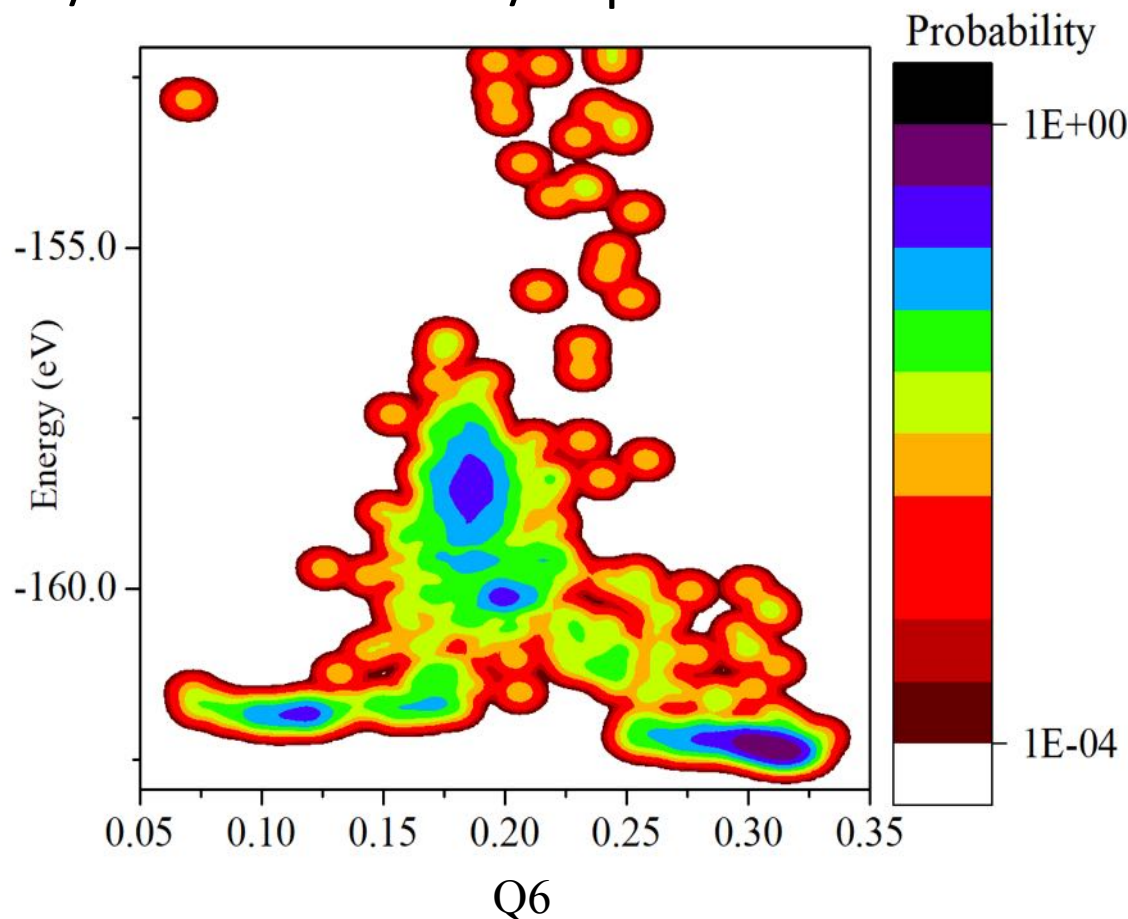
LASP 培训讲义 4 程序的后处理和分析

3. 全局势能面图

LASP_Example/LAMMPS/varcell-ssw-eam/lasp.out

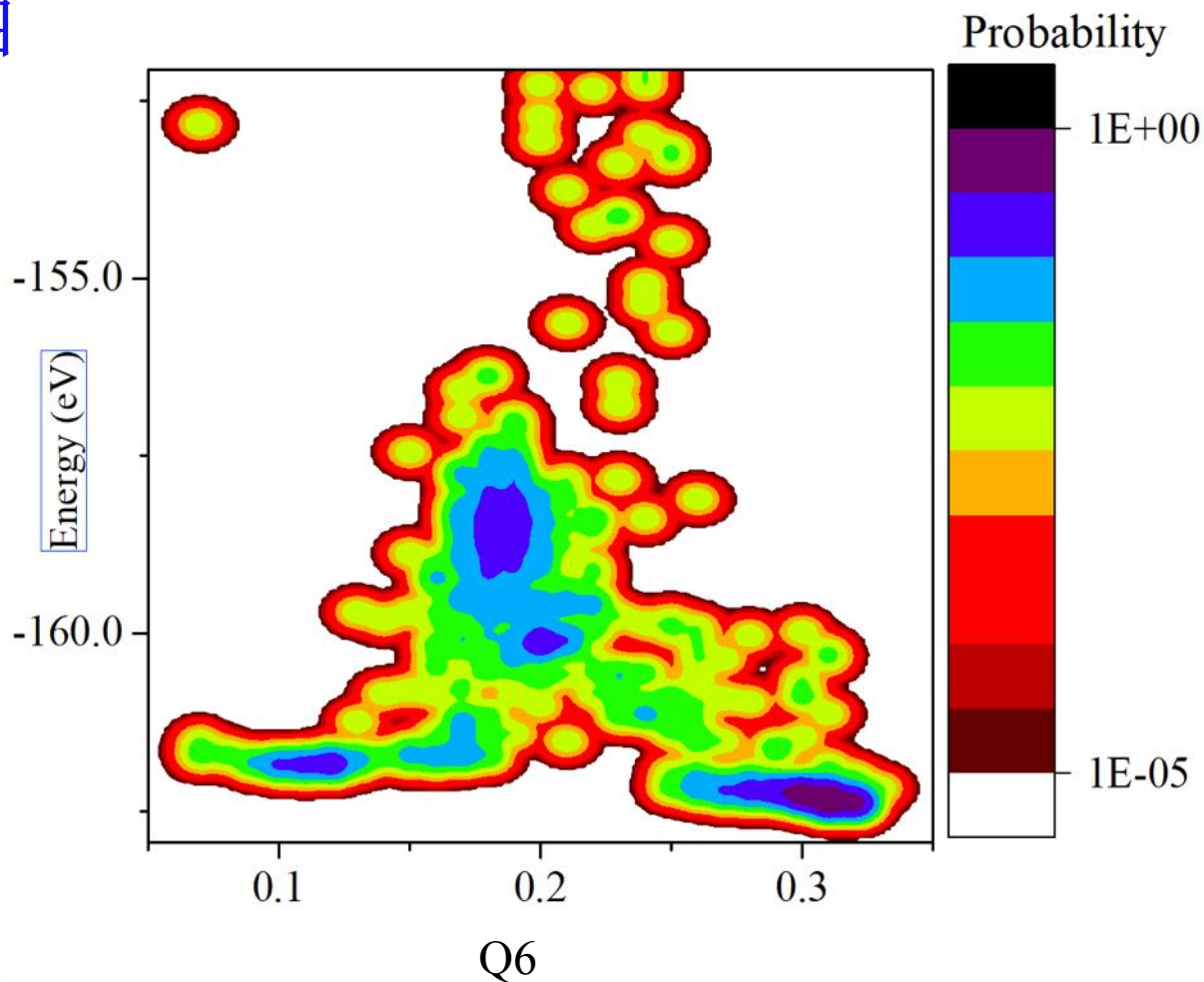
➤ QE.sh

➤ dos2D.f90



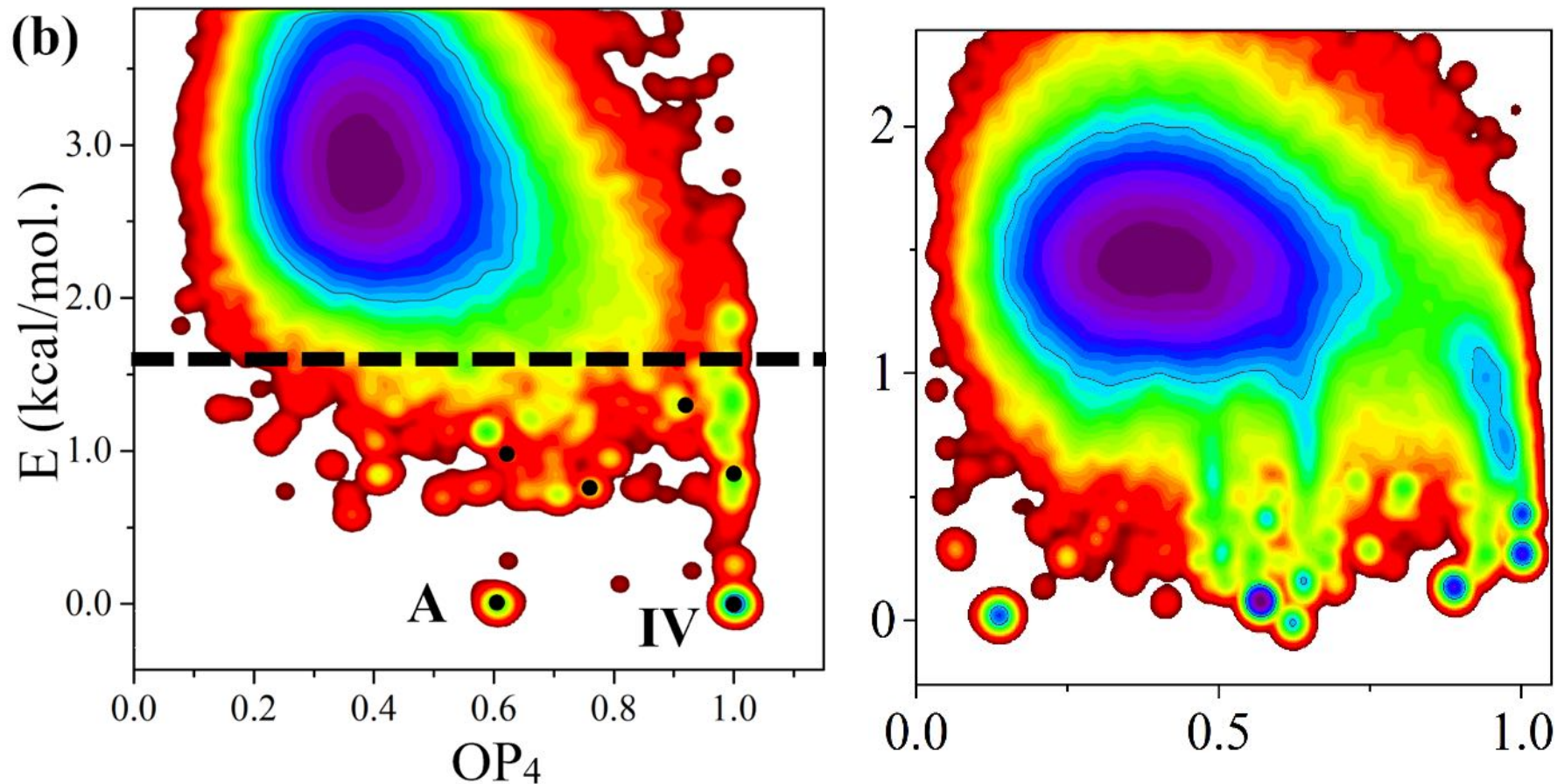
LASP 培训讲义 4 程序的后处理和分析

3. 全局势能面图



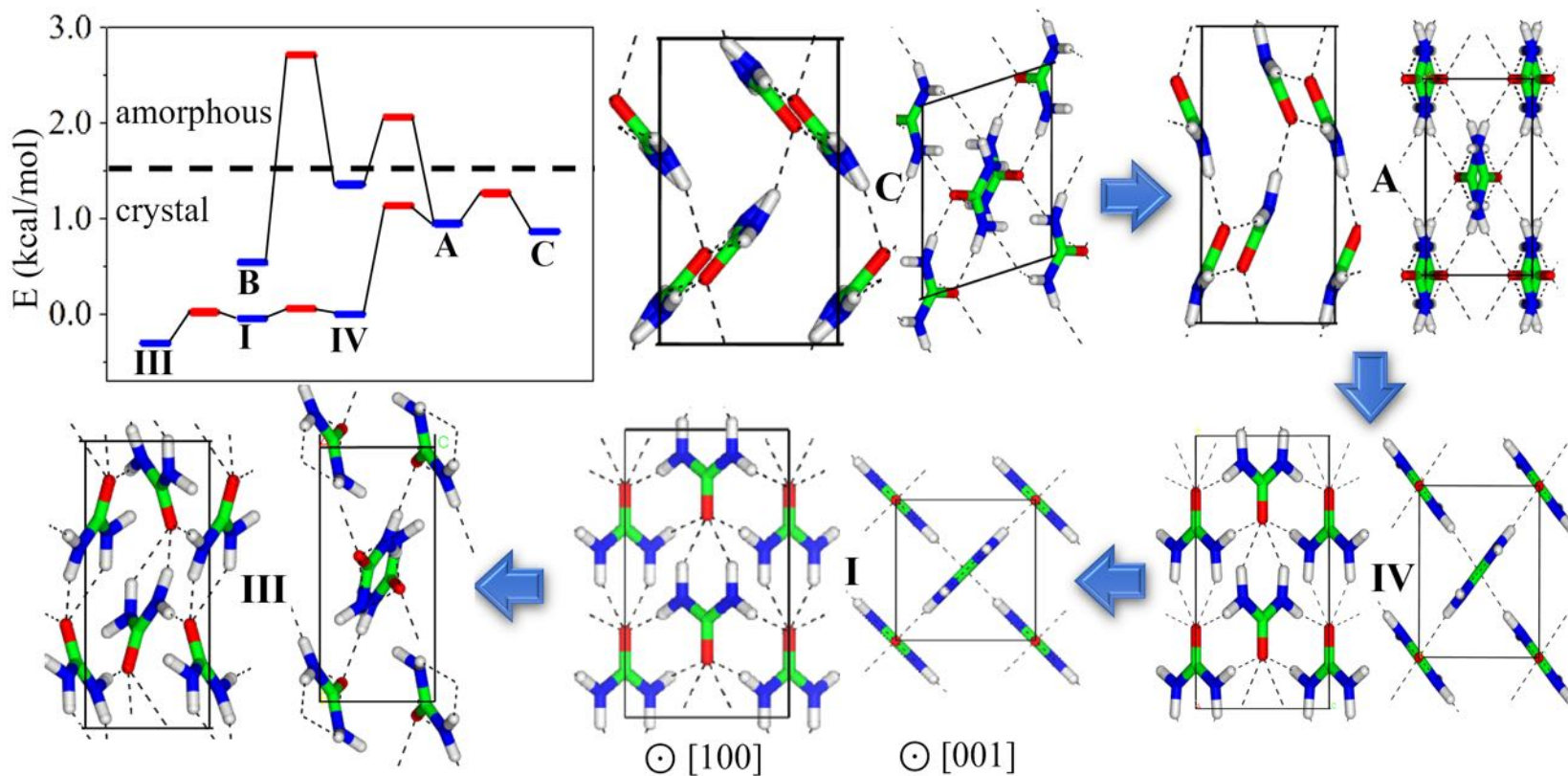
LASP 培训讲义 4 程序的后处理和分析

3. 全局势能面图



LASP 培训讲义 4 程序的后处理和分析

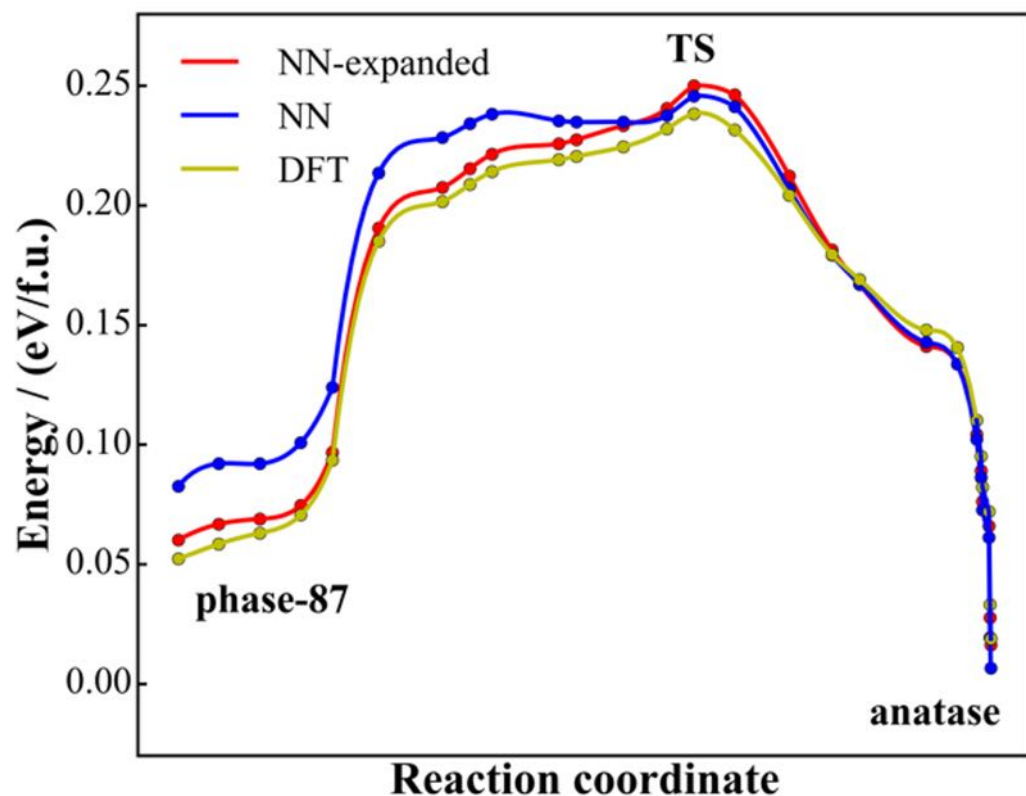
4. 反应路径图



Phys. Chem. Chem. Phys., 2017, **19**, 32125-32131

LASP 培训讲义

4. 反应路径图



反应坐标: $\overrightarrow{FS}-\overrightarrow{IS}$

Chem. Sci., 2017, **8**, 6327-6337