



通用晶体结构预测软件：进化晶体学

用户手册

翻译：刘宁，李端，李坤，程翔，樊涛，姚茜，魏晓婷

统稿：曾庆丰，张瑾

USPEX

Universal Structure Predictor: Evolutionary Xtallography

编写：**A. R. Oganov, C. W. Glass, A. O. Lyakhov, G. -R. Qian, M. S. Rakitin,
H. T. Stokes, Q. Zhu**

版本 10.1, 修订日期 2015 年 7 月 28 日

石溪大学, 苏黎世联邦理工学院, 杨百翰大学, 莫斯科物理技术学院,
南开大学, LN 米塔尔学院, 奥维尔多大学,
巴黎大学, 西北工业大学



<http://uspex.stonybrook.edu>

原稿的其他参编者：

R. Agarwal, X. Dong, P. Perttierra, Z. Raza, M. A. Saivado, D. Dong, Q. Zeng

©A. R. Oganov, 部分章节由 Q. Zhu, M. S. Rakitin 和 G. -R. Qian 完成

<http://uspex.stonybrook.edu>

目 录

1 USPEX 的特点, 目标和历史.....	2
1.1 概述.....	2
1.2 USPEX 的特色.....	8
1.3 有关 USPEX 主要原理及应用的文章.....	9
1.4 版本的历史.....	10
2 初始化.....	16
2.1 如何获得 USPEX.....	16
2.2 版权引用.....	16
2.3 错误报告.....	16
2.4 运行 USPEX 的机型配置.....	16
2.5 与 USPEX 兼容使用的代码.....	17
2.6 如何安装 USPEX.....	17
2.7 如何运行 USPEX.....	18
2.8 运行 USPEX 示例.....	20
2.8.1 测试 USPEX Python 运行器.....	20
2.8.2 运行 EX13-3D special quasirandom structure TiCoO.....	21
2.8.3 运行一个与外部代码交互的示例.....	22
2.8.4 检查结果.....	24
3 输入与输出文件.....	25
3.1 输入文件.....	25
3.2 输出文件.....	26
3.3 SPECIFIC 文件夹.....	32
3.4 SPECIFIC 文件夹中的 INCAR 文件.....	33
4 输入文件 INPUT.TXT 的设置.....	37
4.1 计算类型的方法.....	37
4.2 种群.....	46
4.3 适者生存的选择.....	47
4.4 结构产生和变异操作.....	48
4.5 约束条件.....	52
4.6 晶胞.....	56
4.7 重新开始.....	58
4.8 从头算法的细节.....	58
4.9 指纹设置.....	62

4.10 反种子设置.....	63
4.11 判定空间群.....	65
4.12 统计学——针对开发者.....	66
4.13 不常使用的关键字.....	68
5 特殊情况的额外输入.....	73
5.1 MOL_1, MOL_2, ...等文件.....	73
5.1.1 分子晶体, calculatonType=310/311.....	73
5.1.2 聚合物晶体, calculationType=110 (“线性链状模型”)	74
5.1.3 经典力场的额外输入.....	74
5.1.4 如何准备 MOL 文件.....	76
5.2 表面.....	76
5.3 变成分代码.....	80
5.4 进化准动力学命令.....	83
5.5 粒子群算法 (PSO) 指令.....	87
6 相变路径预测.....	90
6.1 变胞微动弹性带 (VCNEB) 法.....	90
6.2 VCNEB 的 INPUT 选项.....	91
6.3 如何设置 VC-NEB 计算的初始路径.....	99
7 在线工具.....	102
7.1 结构表征.....	102
7.2 性质计算.....	102
7.3 分子晶体.....	103
7.4 表面.....	103
7.5 其它.....	103
8 常见问题.....	104
8.1 如何将结果可视化?	104
8.2 如何避免陷阱?.....	105
8.3 什么是单块计算?	105
8.4 如何使用种子技术?	105
8.5 如何巧妙考虑组成?	107
8.6 如何建立一个从本地机器到远程集群的密码连接?	109
8.7 如何从一个已坏了的 MAT-文件重新开始计算?	109
8.8 当 USPEX 在一段时间内不运行, 该如何操作?	109
8.9 如何建立一个计算使用的工作提交脚本?	110
8.9.1 第一步: 在提交文件夹中配置文件.....	110

8.9.2 第二步：定期运行 USPEX.....	113
8.9.3 Crontab.....	113
8.9.4 Shell 脚本.....	114
8.10 如何在带有对称性代码的 32 位机器上工作？	115
9 附录.....	116
9.1 例子目录.....	116
9.2 运行测试.....	118
9.3 INPUT.TXT 输入文件示例.....	119
9.4 空间群列表.....	120
9.5 平面群列表.....	121
9.6 点群列表.....	122
9.7 USPEX 中使用的共价原子半径表.....	123
9.8 USPEX 使用的默认化合价表.....	124
9.9 USPEX 使用的默认 GOODBONDS 表.....	125
10 参考书目.....	126

1 USPEX 的特点, 目标和历史

1.1 概述

USPEX 是通用结构预测器的简称：进化晶体结构。在俄语里，uspekh 代表“成功”，鉴于此方法的高效，以及通过这种方法产生了许多有用结构，这一命名是恰当的！USPEX 在计算材料发现方面具备许多独特功能。

最初，USPEX 项目的主要目的是不依赖经验知识预测晶体结构。这一目的已通过整合局部优化和实空间从头算法模拟，发展成为先进的进化算法。除了完全非经验式的搜索外，USPEX 同时能够预测大量亚稳结构，并运用不同程度的先验知识进行不同类型的模拟。

晶体结构预测是一个较早期的问题，实际上包括理论晶体化学的中心问题也很悠久。John Maddox 于 1988 年写道【2】：

“物理科学界长久以来的一个丑闻就是不能根据已知物质的化学组成进行结构预测，甚至是最简单的结晶固体…像结晶水（冰）这样简单固体的预测也被认为超出了凡人的知识范围”。

显然，目前的问题是全局优化，即在结构变动过程中发现晶体的最小自由能（每摩尔）。为了得到一些可能存在的结构，让我们考虑一种简化的情况，一个体积为 V ，有 N 个相同原子的固定立方晶胞。为了进一步简化，我们假设原子只能离散分布在一个分辨率为 δ 的网格节点上。这种数字离散化使得不同原子坐标组合的数目 C 有限化：

$$C = \frac{1}{(V/\delta^3)} \frac{(V/\delta^3)!}{[(V/\delta^3)-N]! N!} \quad (1)$$

如果将 δ 设为特征键长（例如， $\delta=1\text{\AA}$ ），由方程 1 给出的组合数量将是自由能局部极小的结构数量的合理估计。如果有不止一种原子类型，不同结构的数量将显著增加。假设一个典型体积 $\sim 10 \text{ \AA}^3$ 的原子，考虑到斯特灵的公式 ($n! \approx \sqrt{2\pi n} (n/e)^n$)，在一个有 10 原子的单胞内 A 元

素(化合物 AB)的可能结构数为 10^{11} (10^{14}) , 若在 20 个原子的单胞内为 10^{25} (10^{30}) , 30 个原子的单胞内为 10^{39} (10^{47}) 我们可以看到, 这些数目很大, 甚至是总原子数 $N \sim 10$ 的小体系实际处理起来也不可能实现。更糟的是, 随着 n 的增加其复杂性呈指数增加。很明显, 除了最简单的含有~1-5 个原子的单胞体系, 在自由能表面上逐点探索所有结构也不可行。

USPEX 【3, 4】采用的进化算法由 A. R. Oganov 和 C. W. Glass 设计, 还有随后做出很大贡献的 A. O. Lyakhov 和 Q. Zhu。它的高效来源于精心设计的变异运算, 其可靠性很大程度上取决于在进化算法中运用了目前最先进的从头算法模拟。进化算法的最大优势是不需要任何体系的特殊知识(除了化学成分之外), 同时还可以自我改进, 即随后代的不断增加, 越来越好的结构会被发现并用于产生新结构。下图 1 是一个能量面(或性质)的可能区域放大图。而且, 基于变异算法的灵活特性, 很容易把一些附加的特点融入到进化算法中。

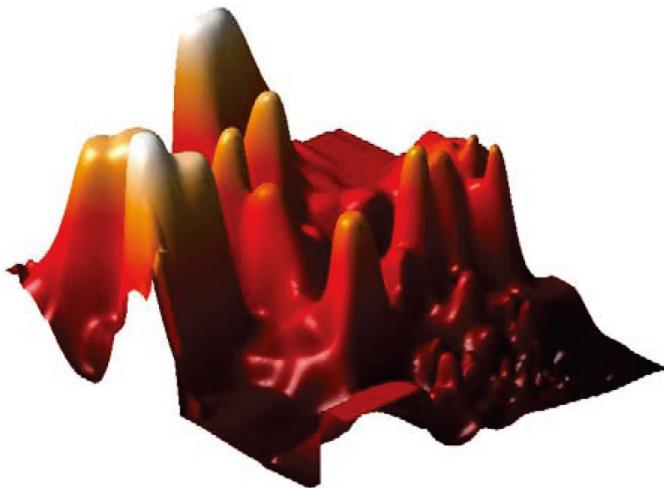


图 1: Au_8Pd_4 的简化面 2D 投影图, 显示了某一区域内具有低能量结构的聚集区。此图是由 Oganov & Valle(2009) 的方法获得。

我们开发 USPEX 的一个主要动力就是 2004 年 【5,6】发现了 MgSiO_3 的后钙钛矿结构(图 2), 这一发现显著改变了地球内核的结构模型。2010 年 9 月, 当 USPEX 公开发布时, 用户数近 200。用户增长速度很快——在 2012 年 5 月我们拥有 800 用户, 这一数字在 2014 年 12 月已超过 2100。

USPEX 的普及归功于它的高效性和可靠性, 而且在第一次对无机晶体结构预测的盲测【7】中体现出了 USPEX 优于其他方法测试, 因为它是针对性的测试(模拟退火和随机抽样)。随

机抽样（由 Freeman 和 Schmidt 分别于 1993 年和 1996 年提倡用于晶体结构预测，并于 2006 年被 Pickard 【8】以 AIRSS 的名字重新使用）是最简单的，也是成功率最低和计算成本最高的方法。由于在结构搜索中复杂的指数标度（公式 1），USPEX 的优势伴随体系大小的增加呈指数增长。然而，即使在小体系中，如有 8 个原子的 GaAs 单胞内，这些优势还是很大（在这个例子中，随机抽样需要平均 500 个结构弛豫来发现基态，而 USPEX 仅仅需要 30 个结构弛豫就能找到（图 3））。例如：

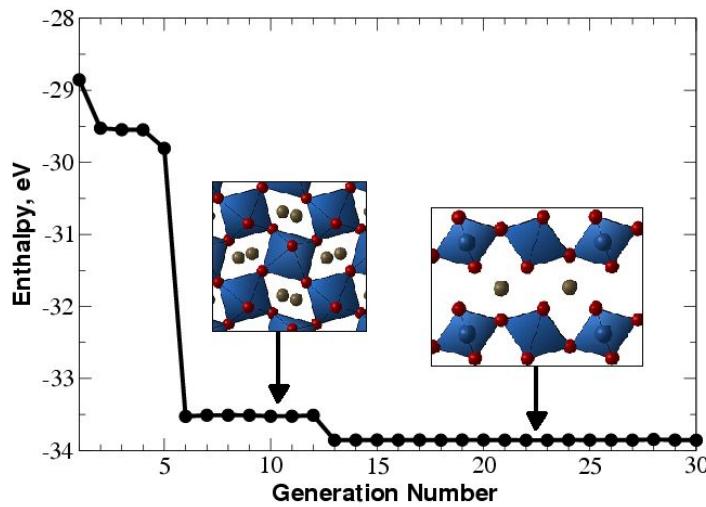


图 2：在 120 GPa 下的 MgSiO_3 晶体结构预测（20 原子/晶胞）。上图展示了随着代数的增加，最优结构的焓值变化。在第六代和第十二代间最优结构为钙钛矿，但在第十三代全局最低的（后钙钛矿）被发现。这次模拟计算在 2005 年结合了 USPEX 的第一个版本和从头算法完成。它没使用实验数据，证明了 USPEX 可以在一次模拟中同时找到稳定结构和低能量的亚稳结构，每一代都包含 30 个结构。这个图举例的是用最初的 USPEX 版本计算~10 次中最慢的一个，即使这样速度也相当快。

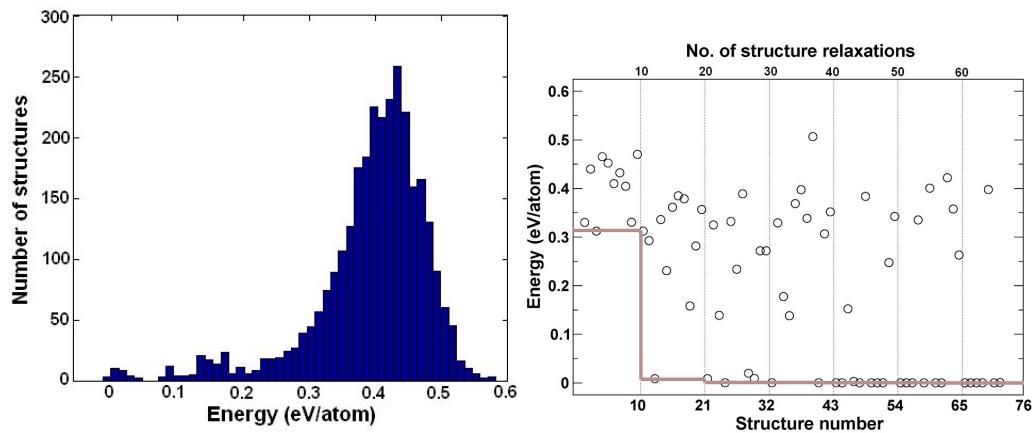


图 3: GaAs 结构预测。图 a)弛豫随机结构的能量分布图。图 b) 一个进化模拟进程图 (细垂直线表示结构代数, 而灰线则表现随代数变化最低能量的变化趋势。所有的能量都是基态结构能量的相对值, 进化模拟中每代会有 10 个结构。此外, 上一代具有最低能量的结构会延续到下一代。

由随机抽样预测得到的 SiH₄ 的稳定结构【8】中三分之二都被证明不稳定【9】; 同样, 使用随机抽样对 N【11】和 SnH₄ 的预测也被证实【10】不正确 (对 USOEX 和随机抽样【13】的预测结构进行比较【12】)。

对于较大的体系, 随机抽样倾向于产生几乎具有相同能量的完全无序结构, 这样就把成功率降低到几乎为零。从 MgSiO₃ 超胞中有 40 个原子的后钙钛矿的例子可以看出: 在 120,000 次的弛豫后随机抽样仍然没有找到正确的结构, 而 USPEX 在几百次的弛豫后就找到了 (图 4)。

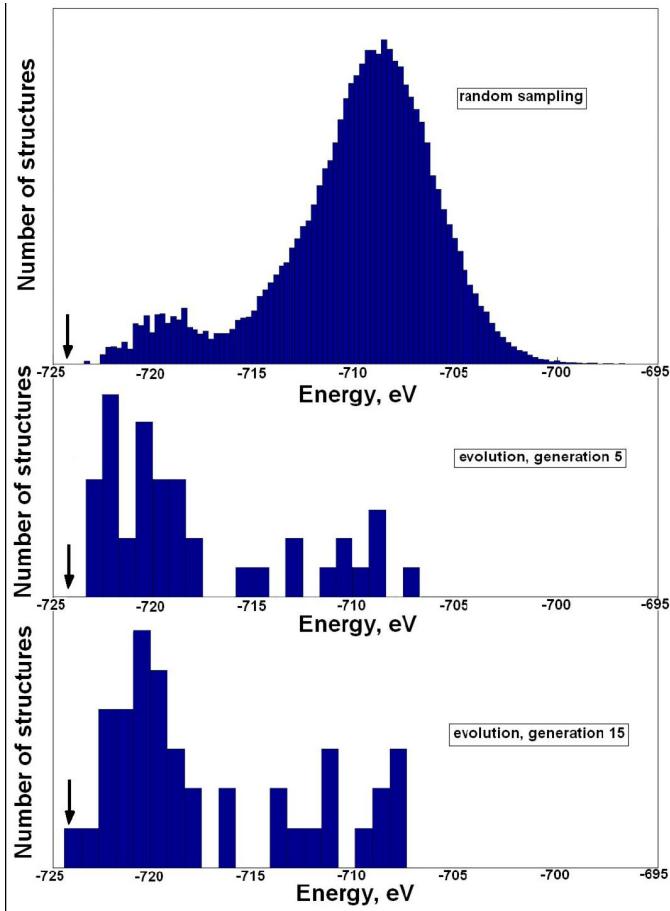


图 4: 能量表面取样: 分别采用随机抽样和 USPEX 对有 40 个原子的 MgSiO_3 后钙钛矿相晶格常数进行计算并对比。局域优化结构的能量如图所示。随机抽样共生成 1.2×10^5 个结构 (没有一种结构与基态一致)。USPEX 的搜索, 每一代包括 40 个结构, 在 15 代内发现基态结构。基态结构的能量为箭头标示处。这张图显示, 包含在进化搜索中的“学习”驱动模拟发现低能量结构。

在 USPEX 中实现随机抽样是非常容易的, 但大多数情况下只在测试的时候觉得这样很有用。同样, 粒子群优化 (PSO) 算法和晶体结构预测 (由 A. I. Boldyrev 开发, Wang, Lu, Zhu 和 Ma 改进实现) 已经被改造。它是在 USPEX 的基础上进行略微调整被作为修正的 PSO (croPSO) 算法, 它优于之前未修订版本的 PSO, 但实际上所有现存的 PSO 方法远不如 USPEX 那么高效可靠。如果有人想尝试, 我们认为 PSO 方法更适用于测试。结合了 Martonak 的准动力学和 Oganov-Glass 的进化方法的进化准动力学方法【14】非常强大, 弥补了进化算法的不足。对于全局优化和获得低能量的亚稳结构, 这个方法是足够的, 甚至可用于寻找可能的相转变路径。

为了研究对于相变机制的详细信息, 其他方法: 变胞 NEB 法【15】和转变路径方法【16】加入到该版本【17】。

在当前的版本, 代码只需要最少量的输入文件, 大多数的输入信息都有可靠的默认值。我们必须指定的是化学体系、压力、温度、硬约束【18】和用于产生下一代结构的结构数量。可以随意的固定晶格常数 (如果是实验已知的) 进行计算。只设置原子类型也可以进行变成分模拟, USPEX 可以找到稳定的组分和相应结构。它还可以预测低维物质的结构, 例如纳米团簇、二维晶体、表面和分子晶体。

1.2 USPEX 的特色

- 只根据化学组成就能预测其稳定和亚稳定结构。还能根据元素模拟搜索稳定的组成和结构。
- 可以结合结构部分已知信息
 - 可以根据实验获得的晶胞参数, 或者固定的晶胞形状, 又或者是固定的晶胞体积来对计算施加束缚条件 (详见 4.6) ;
 - 可以从已知或者猜测的结构开始搜索计算 (详见 8.4) ;
 - 可以从预先确定的分子 (包括结构灵活的分子) 进行晶体结构组装 (详见 5.1) 。
- 高效的束缚技术。可以消除在搜索空间累的非物理和无用的区域。晶胞缩小技术 (Oganov&Glass, 2008) 。
- 采用指纹函数的小生境技术。 (Oganov & Valle, 2009; Lyakhov, Oganov, Valle, 2010)。详见。
- 采用随机的方式, 或者使用空间群和晶胞分裂技术来进行初始化。
- 即使的结果分析—空间群的确定 (详见 4.11) , 硬度的计算, 有序参量等。
- 对纳米微粒结构和表面重构的结构预测 (详见 5.2) 。
- 重启设置。使得计算可以从任何终止的时刻从新恢复并接着计算 (详见 4.7) 。
- 镶嵌在 STM4 代码里强大的可视和分析技术 (M.Valle) 可以完全地与 USPEX 进行接口 (详见 8.1) 。
- USPEX 可以与 VASP, SIESTA, GULP, LAMMPS, DMACRYS, CP2K, Quantum Espresso, FHI-aims, ATK, CASTEP, Tinker, MOPAC 等代码进行接口 (详见 2.5) 。
- 可以从本地的工作站向远程集群和超算提交作业 (详见 8.9) 。

- 可以选择使用 USPEX 算法（默认算法），随机取样法，进化准动力学法（详见 5.4），修正粒子群优化法（详见 5.5），类最小跳跃阀来进行结构预测。可以使用进化准动力学法，变胞 NEB 法（详见 6.1）预测相变路径。
- 除了能量还可以选择以其他物理性质作为优化对象——例如硬度（Lyakhov&Oganov, 2011），密度（Zhu 等人, 2011），带隙和介电常数（Zeng 等人, 2014）和许多其它的性质。
- 为了方便开发者和用户，USPEX 是一个基于 MATLAB 的代码同时它也能在 Octave（一个类似于 MATLAB 环境的免费软件）——用户不需要进行任何编译，为了加强 MATLAB 版本的兼容性，只使用了基本的 MATLAB 命令。USPEX 代码已经在 MATLAB2012 版至 2015 版和 Octave 3.4（新的 Octave 版本将不被支持）进行了开发和测试。
- 从 9.4.1 版本开始，USPEX 具有了安装程序(install.sh 文件)和一个基于 Python 的 MATLAB 代码（USPEX 的 Python 模式），可以提供许多有用的命令行选项。

1.3 有关 USPEX 主要原理及应用的文章

1. Oganov A.R., Glass C.W. (2006). Crystal structure prediction using evolutionary algorithms: principles and applications. *J. Chem. Phys.*, 124, 244704.
2. Oganov A.R., Stokes H., Valle M. (2011). How evolutionary crystal structure prediction Works - and why. *Acc. Chem. Res.*, 44, 227-237.
3. Lyakhov A.O., Oganov A.R., Stokes H., Zhu Q. (2013). New developments in evolutionary structure prediction algorithm USPEX. *Comp. Phys. Comm.*, 184, 1172-1182.
4. Zhu Q., Oganov A.R., Glass C.W., Stokes H. (2012). Constrained evolutionary algorithm for structure prediction of molecular crystals: methodology and applications. *Acta Cryst. B*, 68, 215-226.
5. Zhu Q., Li L., Oganov A.R., Allen P.B. (2013). Evolutionary method for predicting surface reconstructions with variable stoichiometry. *Phys. Rev. B*, 87, 195317.
6. Zhu, Q., Sharma V., Oganov A.R., Ramprasad R. (2014). Predicting polymeric crystal structures by evolutionary algorithms. *J. Chem. Phys.*, 141, 154102.

7. Zhou X.-F., Dong X., Oganov A.R., Zhu Q., Tian Y., Wang H.-T. (2014). Semimetallic Two-Dimensional Boron Allotrope with Massless Dirac Fermions. *Phys. Rev. Lett.*, 112, 085502.
8. Lyakhov A.O., Oganov A.R., Valle M. (2010). Crystal structure prediction using evolutionary approach. In: *Modern methods of crystal structure prediction* (ed: A. R. Oganov), Berlin: Wiley-VCH.
9. Oganov A.R., Ma Y., Lyakhov A.O., Valle M., Gatti C. (2010). Evolutionary crystal structure prediction as a method for the discovery of minerals and materials. *Rev. Mineral. Geochem.*, 71, 271-298.
10. Duan, D., Liu, Y., Tian, F., Li, D., Huang, X., Zhao, Z., Yu, H., Liu, B., Tian, W., Cui, T. (2014). Pressure-induced metallization of dense $(\text{H}_2\text{S})_2\text{H}_2$ with high- T_c superconductivity. *Sci. Rep.*, 4.

1.4 版本的历史

v.1—无局部优化进化算法。实空间表示, 与 VASP 接口。实验版。2004.10

v.2—CMA-ES 的启用 (CMA-ES 是一个由 N. Hansen 设计的功能强大的全局优化方法) 实验版。2005.1

v.3—有局部优化的进化算法

v3.1—工作版本, 序列。主要的基本改进。

3.1.4-3.1.5—第一产品版本。主要是片段移位和最小的亲代贡献的遗传 (硬编码为 0.25)。2005.5

3.1.8—适应 k 点网格。15 / 10 / 2005

3.1.11—从任意的一代开始重新计算。实验版。04 / 11 / 2005

3.1.12—基于 v.3.1.00 的产品版本, 变切片突变。11 / 11 / 2005

3.1.13—自适应缩放体积。29 / 11 / 2005

3.1.14—初步的种子技术。29 / 11 / 2005 (6/ 12/ 2005 调试)

v.3.2—大规模并行版本。

v.4—统一并行/串行版本。

4.1.1—晶格突变。20 /12 /2005 (10/ 01/ 2006 调试)

4.2.1—与 SIESTA 接口。初始群的大小可以不同于模拟运行中群的大小。24 /01 /2006 (20/ 04/ 2006 调试)

4.2.2—可以设定的保留最佳结构的数目。实验版。21 /04 /2006

4.2.3—最佳结构的弛豫可选择。完全调试了的大规模并行版本。25 /04 /2006

4.4.1—与 GULP 接口。08 /05 /2006

v.5—完全重写和调试版本, 代码模块化。

5.1.1—新平台: Blanc, Gonzales。暂时放弃序列模式。特定原子突变, 代码的互通性, 速读 INPUT_EA.txt 中的参数。20 /12/ 2006。

5.2.1—Z 矩阵的 SIESTA 与接口, 旋转变异运算, 只确保远程作业提交 (实验版)。01 /03 /2007。

v.6—产品版本, 与本地和远程作业都可提交。

6.1.3—可以高效地完成大体系的硬约束, USPEX 中可实现优化。07 /06 /2007。

6.2—开发版本

6.3.1-6.3.2—引入单元格对角线的角度约束。完全重写的远程提交。改进输入格式。进一步扩展标准测试。07 /12 /2007

6.3.3—X-com 网格界面 (有 S. Tikhonov 和 S. Sobolev 的共同参与)。在完善中。05 /03 /2008

6.4.1—指纹功能。07 /04 /2008

6.4.4—空间群识别 (调试过, 但仍有问题), 可以选择性的关掉。快速指纹 (参考表)。05 /05 /2008

6.5.1—针对大体系的分裂晶胞法。更容易的远程提交作业。可变数量的最佳结构（能量群）。

16 /07/2008

6.6.1—非常稳健的版本，与当地执行重新启用，改进了指纹和晶胞分裂技术。13/ 08/ 2008

6.6.3—实现多种亲代遗传。01 /10 /2008

6..6.4—增加亲代参与遗传的阈值。03 / 10 / 2008

6.6.6—第一次实现多元指纹功能。04/12/2008

6.6.7, 6.7.1 和 6.7.2—实现准熵测量种群的多样性，将 CEL 和 SPF 放入独立文件夹。10/12/2008

v.7—产品版本，含有变成分。

7.1.1-7.1.7—系列改进版本。版本 7.1.7 已分发给 200 用户。变成分一定程度上已编码，已知的漏洞已修复，基于能量面的技巧已改进。改进了晶胞分裂，实现亚晶胞。实现多组分指纹功能（跟单组分指纹功能相比对结构更为敏感）。28 /04 /2009 (28/ 05/ 2009 确定版本)。

7.2.5—第一个变成分计算全功能版本。引入了原子置换操作和组合熵。06 /09 /2009。

7.2.7—彻底调试，改进的重启功能，改进种子技术，与 M. Baskes 的 MD++ 代码接口，引入了结构弛豫中的扰动和变成分的偏置适应度函数。25 /09 /2009，在 7.2.8/9 版本中进一步改进。

7.3.0—指纹完全支持变成分的代码，包括 niching。“公平”算法用于产生组分的第一代结构。
22 /10 /2009

7.4.1—引入基于局部有序【22】坐标突变，遗传和原子置换也是基于局部有序。引入了硬度计算和基于硬度和密度的新型优化。04 / 01 / 2010

7.4.2—调试，实现基于局部有序多亲代遗传。15 / 01 / 2010

7.4.3—调试，实现新类型优化（最大化结构有序和种群多样性）。加入反种子技术，消除参数 voltimeconst, volbesthowmany。24 / 01 / 2010

v.8—将新型优化写入产品版本。

8.1.1-8.2.8—开发版本。局部有序与坐标变异算子, 软模变异运算, 硬度的计算与优化, 优化介质的敏感性。预测纳米粒子与表面结构, 实现点群, 软件整体表现提升。可选用 PSO 模拟, (由于 PSO 的低效率, 不建议在实际应用中使用 PSO 模拟——所以只作测试用途)。参数 goodbonds 转变为矩阵用以建立纳米粒子。22 / 09 / 2010

8.3.1—调试 PSO 算法, 优化介电常数, 清除输入。08 / 10 / 2010。

8.3.2—为集群引入连通性检查 (非常有用), 改进 dynamicalbesthm = 2 选项, 以及用于产生一代纯软模变异结构。改进集群的指纹功能。与 Quantum Espresso 和 CP2K 接口。11 / 10 / 2010

8.4—开发版本集合版, 包含改进的反种子功能和纳米团簇计算上做了一些改进, 建立表面重构、伪准动态和分子晶体分支程序。

8.5.1—工作界面的大量调试。对定晶胞, 变成分和亚晶胞实现空间群初始化。20 / 04 / 2011

8.6.0—H. Stokes.的工作为软件添加空间群判定程序。为分子晶体合并更新后代码 (包括空间群初始化)。修正了一个 SIESTA 中错误 (感谢 D. Skachkov)。06 / 05 / 2011

8.6.1-8.7.2—开发版本, 相当稳健。改进了定晶胞内的对称初始化。优化介电常数 (使用 GULP 和 VASP), 带隙 (使用 VASP) 和费米能级的 DOS 图 (VASP)。启用图形输出。改进的软模变异运算 (使用更好的标准模式与定向简并) 和遗传运算 (对实验物质使用能量规则相关性系数和余弦公式)。现在大多数变量有默认值, 它减少用户输入量。减短并改进格式日志。13 / 11 / 2011

8.7.5—修正变成分的代码中错误, 图形输出许多额外数据, 如用于可变成分计算的原子近似体积。增加将结构提取为凸包图的功能, 以便进行后期处理。21 / 03 / 2012

v.9—产品版本, 增加客户体验, 加入新类型的功能, 在该领域设置新标准。

9.0.0—将进化准动力学和 vc-NEB 代码添加到 USPEX 软件包, 增加准动力学的张量的版本, 添加额外的数据和后处理工具, 清理代码的输出。去除 INPUT 中的部分参数。改进软模变异。2012.4

9.1.0—发布版本。清理并整合。用户人数超过 800 人。发布于 28 / 05 / 2012

9.2.0—运行 GEM (广义进化准动力学)。持续开发 GEM 代码。空间群测定限度作为新的输入参数。改进大量置换变异参数。2012.7-2012.8

9.2.1-9.2.3—改进 GEM, 种群多样性和超胞尺寸, 改进选择模式。2012.9-2012.10

9.2.4-9.2.6—(9.2.4 发行版本)。多数输入参数采用智能默认值。改进的对称集群初始化能力。为纳米粒子增强有序性。新的参数用以调整空间群测定的限度。修正几个小错误。可以优化新的性质 (准熵)。完全集成的 vc-NEB 的代码。2012.11-2012.12

9.2.7—发行版本。在没有结构弛豫的情况下进行合金的有序优化, (在不依托于特定的准随机机制而是通常用的普通随机机制的情况下, 更易于产生准随机结构)。改进对称代 (这点对于定晶胞计算非常重要)。在固定晶胞计算过程中, 可以指定一个晶胞参数, 不一定非是一个 3×3 矩阵, 也可以是一排六个值 (三个以埃为单位的长度值和三个以度为单位的角度值)。为了置换数目的最大化 (参数 howManySwaps), 引入了智能默认值。添加新的测试, 清除当前参数并启用旧参数。与 CASTEP 接口 (感谢 Z. Raza, X. Dong 和 AL)。用户人数达到 1160 人。30 / 12 / 2012

9.3.0- 9.3.3—修正在一个在随机对称结构代中出现的错误 (这个错误出现于 9.2.7 版本)。显著简化输入和输出。创建 OUTPUT.TXT 用以存放重要信息。对分子晶体进行晶体分裂。通过允许用户设定初始成分来改进变成分计算与 LAMMPS 和 CASTEP 的接口。添加新的测试案例。20 / 03 / 2013

9.3.4—发行版本, 并做了清理。25 / 03 / 2013

9.3.5—修正了几个错误。添加并调试了预测二维晶体的代码。19 / 04 / 2013

9.3.6—集成和调试了 GEM。改进变成分计算。二维晶体包含平面种群。29 / 04 / 2013

9.3.8—一维聚合物晶体中添加平面群描述, 改进变化学计量比的表面结构预测。19 / 06 / 2013

9.3.9—发行版本。显著改进的版本, 加强用户体验, 新功能 (如二维晶体, GEM) 表现良好, 改进变成分算法 (同时支持单块计算, 即定组分但改变单胞内原子个数)、全功能表面计算, 新类型优化 (可优化介电常数、键长和新的介电材料的品质因数)。筹备在新的测试中与 LAMMPS 和 ATK 接口。通过微小调试不断更新 (最后调试于 10 / 02 / 2014)。

9.4.1—主要升级，极大地提高用户使用性便利性（自动预估体积和每种案例的变异运算操作的比例），新功能（弹性性质的优化和 Chen 氏硬度模型，预测聚合物的结构、反组分、自动统计分析和改进的种子技术），第一次发布 GEM，提供一系列 USPEX 的计算实例，测试案例，说明文件。用户人数超过 2100 人。发布于 30/12/2014

9.4.2—发行版本，兼容 Octave 3.4 中的错误。重新写入凸包图代码。与 MOPAC 接口。启用默认参数 goodBonds, valences, IonDistances。三元、四元和更复杂成分的变成分搜索表现良好。用户超过 2200 人。发行于 21 / 03 / 2015

2 初始话

2.1 如何获得 USPEX

USPEX 是一款开源软件，可以从以下网址下载：

<http://uspex.stonybrook.edu>

在下载页面，分别有 USPEX 源代码包，示例和手册文件。

2.2 版权引用

无论什么时候使用 USPEX，在所有出版物或者报告中必须在原文中指明，例如：“Crystal structure prediction was performed using the USPEX code[3,14,31]，based on an evolutionary algorithm developed by Oganov, Glass, Lyakhov and Zhu and featuring local optimization, real-space representation and flexible physically motivated variation operators”.

查阅 OUTPUT.txt 文件获得最重要的引用。

2.3 错误报告

像其他任何一个如此大型的代码，USPEX 在运行过程中可能存在错误。如果在你的模拟中发现异常的运行，请将 INPUT.TXT 和 log 日志发送给 USPEX 的管理员（目前可以发给朱强（alecfans@gmail.com））。

2.4 运行 USPEX 的机型配置

USPEX 可以运行于任何平台——所需要的只是一个能在 Linux 或者 Unix 系统环境下运行 MATLAB 或者 OCTAVE 软件的 CPU。通过使用该系统特殊的远程作业机制，USPEX 可以连接在任何远端机上进行计算。

2.5 与 USPEX 兼容使用的代码

为了便于编程和运行, USPEX 是以 MATLAB 语言编写的, 并且可以工作在 Octave (一种免费的类 MATLAB 环境) 下, 你不需要编译任何东西, 只需要接通并运行! 为了加强 MATLAB 版本的兼容性, 只有一些基本的 MATLAB 命令才能被使用。从 9.4.1 版本开始, USPEX 有一个安装器 (install.sh) 和一个基于 Python 语言的 MATLAB 代码运行器, 提供了许多有用的命令行选项。

通过 USPEX 产生的测试性结构是弛豫的, 然后通过外部与 USPEX 交互的代码进行评估。基于获得的弛豫结构的排列, USPEX 产生新的结构——它们又可以被用来弛豫和排列。我们的理念是使用现存的, 成熟的 ab initio 代码做结构弛豫和能量计算。目前, USPEX 可以和 VASP、SIESTA、Quantum Espresso、CASTEP、CP2K、DMACRYS、GULP、LAMMPS、ATK、Tinker 和 FHI-aims 这些软件交互作用。选择这些代码是基于 1) 它们结构弛豫的效率; 2) 稳健性; 3) 普遍性。当然还有一些其他代码也满足这些要求, 未来我们可以再加进去。

2.6 如何安装 USPEX

当你下载了 USPEX 的文件之后, 需要解压并运行以下命令安装 USPEX 到用户或者系统级的位置:

```
bash install.sh
```

安装者不要求具有 root 权限。你将会被询问是选择 MATLAB 还是 Octave, 安装目录以及确认创建或者使用这个目录。然后系统会提供给你关于环境变量的信息, 这项必须设置以确保 USPEX 在系统中可用, 例如:

For Bash shell system, add these lines in `~/.bashrc` or `~/.profile` or `/etc/profile`:

```
export PATH=/home/user/bin/USPEX:$PATH
```

```
export USPEXPATH=/home/user/bin/USPEX/src
```

For C shell system, add these lines in `~/.cshrc` or `~/.profile` or `/etc/profile`:

```
setenv PATH "/home/user/bin/USPEX:$PATH"  
  
setenv USPEXPATH "/home/user/bin/USPEX/src"
```

如果你想改变 MATLAB 或者 Octave 的路径，可以编辑 USPEX 安装目录里的 CODEPATH 文件。

2.7 如何运行 USPEX

为了运行 USPEX，你必须有 MATLAB 或者 Octave，USPEX 运行在 MATLAB（优先）或者 Octave 下，并且在电脑节点上有可运行的外部代码用于结构的弛豫和计算它们的能量。目前，USPEX 可以和 VASP、SIESTA、Quantum Espresso、CASTEP、CP2K、DMACRYS、GULP、LAMMPS、ATK、Tinker 和 FHI-aims 这些软件一起使用。开始你的计算时，根据你想要做什么体系来从示例中找到类似的例子（见附录：示例列表），然后从编辑 INPUT.txt 开始。对 INPUT 文件的详细使用说明键第四章节。然后，将外部代码做结构弛豫需要的文件放在 Specific 文件夹下，以 VASP 为例，这样的文件有 INCAR_1, INCAR_2, ...INCAR_N, 和 POTCAR_1, POTCAR_2, ...POTCAR_N。

有两种运行代码的方式——旧的和新的，两种都可以用。

(i) 在老的方式中，你的运行文件夹下需要有完整的 USPEX 代码（USPEX.m 文件和 FunctionFolder 目录等）。然后输入

```
nohup matlab < USPEX.m > log &
```

或者你用的是 Octave，输入

```
nohup octave < USPEX.m > log &
```

(ii)新的方式中，如果你使用了基于 Python 语言的运行器的 USPEX 安装器，你所需要做的是输入：

```
nohup USPEX -r > log &
```

或者，如果是 Octave，输入

```
nohup USPEX -r -o > log &
```

Log 文件中包含有模拟过程中信息的记录，如果有错误的话（如果你愿意报告软件的漏洞，请把 log 文件发送给我们）。

OUTPUT.txt 文件包含有每一代计算和分析的详细信息。

对于 USPEX 运行者，我们有许多人性化的选项：

-v, --version: 显示程序的版本号并退出

-h, --help: 显示帮助信息并退出

-p, --parameter: 指明需要获得帮助的参数。如果没有值或者是“all”，则整个 INPUT.txt 参数都会显示

-e, --example: 显示 USPEX 例子的细节。如果没有值或者是“all”，则所有的例子都将被显示。

-c, --NUM, --copy=NUM: 复制 INPUT 文件并指明例子所在文件夹。

-g, --generate: 产生准备 USPEX 计算的文件夹，包括 AntiSeeds、Seeds、Specific、Submission folders

-r, --run: 运行 USPEX 计算

--clean: 清理计算文件夹

当进行大规模并行计算时，用户需要做一点工作根据用户的电脑配置文件（因此，我们不敢保证支持解决大规模并行计算的问题）。

有两种提交作业的方式——（1）本地提交；（2）远程提交，取决于你是否将 ab initio 计算任务提交给你运行 USPEX 和 MATLAB 相同的机器，或者说你是否将你的工作发给远程的超级计算机。参见关键字 whichCluster 以及 8.4 和 8.5 章节。

2.8 运行 USPEX 示例

一旦下载了 USPEX 的软件包并安装了它，你就可以运行第一个 USPEX 例子。这个例子的详细信息列在附录：例子列表中。所需要的运行例子的外部代码显示如下：

GULP: EX02, EX03, EX08, EX12, EX15 (VC-NEB), EX16

VASP: EX01, EX07, EX09, EX14 (META)

LAMMPS: EX04

ATK: EX05

CASTEP: EX06

DMACRYS: EX10

TINKER: EX11

现在，开始我们的第一个 USPEX 体验：

2.8.1 测试 USPEX Python 运行器

为了得到版本信息，我们可以使用以下命令：

```
>> USPEX -v
```

如果 USPEX 运行正常，你应该得到以下信息。如果不工作，请检查你的安装步骤和系统环境设置，如上所述。

USPEX Version 9.4.2 (19/03/2015)

2.8.2 运行 EX13-3D special quasirandom structure TiCoO

例子 13 (EX13) 不需要任何外部代码，我们可以通过运行这个例子来熟悉 USPEX 的运行步骤。这个例子的运行大约会花费 30 分钟左右的时间。为了开始计算，我们首先创建一个测试文件夹，将示例文件复制过来，然后通过 USPEX Python 运行器计算，命令如下：

```
>> mkdir EX13
```

```
>> cd EX13
```

```
>> USPEX -c 13
```

```
>> USPEX -r
```

在例子运行的时候，我们现在有时间了解 EX13 更多的细节。在 EX13 中，我们通过进化算法优化结构阶数。所以，在 INPUT.txt 中，我们设置：

```
USPEX : calculationMethod
```

```
-4      : optType
```

我们使用以下参数：

```
300  : calculationType
```

```
% atomType
```

```
Co Ti O
```

```
% EndAtomType
```

```
% numSpecies
```

```
16 16 64
```

```
% EndNumSpecies
```

为举例说明 USPEX 研究 $\text{Co}_{16}\text{Ti}_{16}\text{O}_{64}$ 体系。为了减少 EX13 的计算时间，我们只是作为测试，我们可以减少 populationSize 和 numGenerations 这两个参数的值，如：

```
5 : populationSize
```

```
5 : numGenerations
```

然而，需要注意的是设置如此小的 populationSize 和 numGenerations，我们不能期望 USPEX 找到低能量的结构。

由于我们不需要外部代码，我们可以简单设置：

```
% abinitioCode
```

```
0
```

```
% ENDabinit
```

种子文件放置在 Seeds/POSCARs — 这是初始的 $\text{Co}_{16}\text{Ti}_{16}\text{O}_{64}$ 结构。当你发现 USPEX_IS_DONE 文件时，恭喜你，你已经成功完成了我们的第一个示例。接下来，我们将运行 USPEX 与外部代码交互的计算。

2.8.3 运行一个与外部代码交互的示例

这一步中，我们建议运行与 GULP 或者 VASP 交互的示例，从例子 02 (EX02) 或者例子 01 (EX01) 开始。我们使用 USPEX 运行器获得 EX02 的例子信息，创建一个单独的文件夹并复制文件，使用以下命令：

```
>> mkdir EX02
```

```
>> cd EX02
```

```
>> USPEX -c 2
```

因为在 EX02 中，我们使用的是 GULP，设置

```
% abinitioCode
```

```
3 3 3 3
```

```
% ENDabinit
```

为了运行一系列的计算而不是单独一个计算，我们应该改变 INPUT.txt 中的以下参数：

```
0 : whichCluster
```

```
1 : numParallelCalcs
```

在示例 INPUT.txt 文件中，whichCluster=QSH 是一个内部的机群名供我们自己的团队使用。我们也提供相同的方式帮助用户定义自己的机群。详细信息参见 FAQ.8.11。

在 INPUT.txt 中，我们没有指定怎么运行 GULP，因为我们假定所有用户使用相同的命令：

```
% commandExecutable
```

```
gulp < input > output
```

```
% EndExecutable
```

但是要确保这个命令是在你的机子上工作的。如果你想运行 EX01，使用的是 VASP，你应该这样设置：

```
% abinitioCode
```

```
1 1 1 1
```

```
% ENDabinit
```

```
% commandExecutable
```

```
mpirun -np 8 vasp
```

```
% EndExecutable
```

如果 commandExecutable 设置错误，启动 USPEX 计算就会失败。当每个参数都设置正确了，我们就可以通过 USPEX 运行器运行计算了。使用命令：

```
>> USPEX -r
```

2.8.4 检查结果

开始计算后，你可以检查 results1 文件夹中的 OUTPUT 文件。

现在，你有了使用 USPEX 运行简单计算的经验。要想了解更详细的关于 USPEX 和运行复杂计算的信息，我们可以从接下来章节的介绍中了解。

3 输入与输出文件

输入/输出文件依赖于进行结构弛豫的外部代码和作业提交的类型。

我们的原理中一个重要的技术要素是多步结构弛豫策略，该方法有坚实的理论基础。为了提供依据能量的正确的结构排列，最后一步的结构和能量必须是高质量的。大多数新产生的结构都距局部最小值很远（例如：所包含的键不是太短就是太长）而它们的高质量弛豫的代价极其昂贵。如果第一步的弛豫使用较为粗略的计算条件，只有最后一步的弛豫才需要高质量的计算，则这种高花费就可以避免。第一步的结构弛豫可以使用代价较低的方式（例如：原子间的势使用 GULP）。你可以改变计算条件（基本设置， k 点采样网格，赝势或者 PAW 势）或者近似的程度（原子间势、LDA 或 GGA）或者结构弛豫的代码（GULP、DMACRYS、SIESTA、VASP、CP2K、QE、CASTEP 或 ATK）对于每一个候选结构在结构弛豫期间。更进一步，为了保证结构弛豫的稳健性，也需要多步策略——如果初始力和应力太大，全弛豫可能会导致结构的“爆炸”。因此，我们强烈建议初始优化原子的位置时保证晶格参数为常数，或者优化晶胞的形状和原子位置时保证单位晶胞体积是常数。当体积固定进行优化时，你就不必担心在平面波计算中的 Pulay 应力。因此，使用小一点的基本设置也可以；然而，对于常压变晶胞弛豫，你需要一个高质量的基本设置。对于结构弛豫，你可以从一个小一点的 k 点设置开始，但是不要忘记在最后一步结构弛豫中适当增加，以获得精确的能量。

3.1 输入文件

假设计算所在文件目录是~/StructurePrediction，当按照序列运行 USPEX，这个目录下将会包含：

- INPUT.txt 文件
- Subdirectory ~/StructurePrediction/Specific 和 VASP、SIESTA 或 GULP 等软件的可执行文件，列出结构弛豫的输入文件——INCAR_1, INCAR_2, ... 和赝势文件。
- Subdirectory ~/StructurePrediction/Seeds 中包含种子文件，其中列出了成分/反成分。注意种子文件应该是 VASP5 POSCAR 格式来写入并以 POSCARS 或者 POSCARS_gen (gen 是代

数) 来对文件进行命名。成分和反成分文件用来控制在变成分或者单程序段计算中成分的变化。

- Subdirectory `~/StructurePrediction/AntiSeeds` — 你可以在这里放一些在计算中希望避免产生的特殊的结构, 或者不需要明确指定任何结构而使用反种子技术 (避免一些在运行中发现的结构)
- Subdirectory `~/StructurePrediction/results1` (如果这是一次新的计算) 和 `results2, results3,` (如果计算被中断后重启过或者运行了很多次, 对每一次计算都会有一个单独的结果文件夹)。

3.2 输出文件

输出文件一般在 `result` 文件夹里`~/StructurePrediction/results1` 是用于处处 USPEX 计算的结果。如果是一次新的计算, `results2, results3, ...` (如果计算被重新启动或运行了几次) 就会把每次计算的结果放在分开的 `results` 文件夹里。

The subdirectory `~/StructurePrediction/results1` 包含以下文件:

- `OUTPUT.txt` —— 总结输入变量, 由 USPEX 产生的文件和它们的特征值。
- `Parameters.txt`——`INPUT.txt` 文件的复制, 供参考
- `Individuals` —— 给出每个产生结构的详细信息 (能量, 单位晶胞体积, 空间群, 用于产生该结构所做的操作, 用于计算能量的 k 点网格, 有序度等)。文件 `BESTIndividuals` 给出每一代最好结构的信息。

Gen	ID	Origin	Composition	Enthalpy (eV)	Volume (A^3)	Density (g/cm^3)	Fitness	KPOINTS	SYMM	Q_entr	A_order	S_order
1	1	Random	[4 8 16]	-655.062	201.062	4.700	-655.062	[1 1 1]	1	0.140	1.209	2.632
1	2	Random	[4 8 16]	-650.378	206.675	4.572	-650.378	[1 1 1]	1	0.195	1.050	2.142
1	3	Random	[4 8 16]	-646.184	203.354	4.647	-646.184	[1 1 1]	1	0.229	0.922	1.746
1	4	Random	[4 8 16]	-649.459	198.097	4.770	-649.459	[1 1 1]	9	0.128	0.958	2.171
1	5	Random	[4 8 16]	-648.352	202.711	4.662	-648.352	[1 1 1]	2	0.154	1.014	2.148
1	6	Random	[4 8 16]	-643.161	206.442	4.577	-643.161	[1 1 1]	1	0.234	0.946	1.766
1	7	Random	[4 8 16]	-647.678	207.119	4.562	-647.678	[1 1 1]	1	0.224	1.108	2.106
1	8	Random	[4 8 16]	-644.482	203.844	4.636	-644.482	[1 1 1]	1	0.215	0.952	1.857
1	9	Random	[4 8 16]	-647.287	204.762	4.615	-647.287	[1 1 1]	40	0.136	1.142	2.563
1	10	Random	[4 8 16]	-649.459	198.097	4.770	-649.459	[1 1 1]	9	0.128	0.958	2.171

- Convex_hull ——只用于变成分计算，给出所有热力学稳定的成分，和它们的焓值（每原子）。

```
---- generation 1 -----
10 0 -8.5889
0 14 -8.5893
11 3 -8.7679
---- generation 2 -----

10 0 -8.5889
0 14 -8.5893
11 3 -8.8204
---- generation 3 -----
10 0 -8.5889
0 14 -8.5893
12 4 -8.9945
```

- gatheredPOSCARS ——弛豫的结构（以 VASP POSCAR 的格式）

```
EA1      9.346  8.002  2.688 90.000 90.000 90.000 Sym.group:    1
1.0000
         9.346156      0.000000      0.000000
         0.000000      8.002181      0.000000
         0.000000      0.000000      2.688367
Mg      Al      O
        4          8          16
```

```
Direct
0.487956      0.503856      0.516443
0.777565      0.007329      0.016443
0.987956      0.507329      0.016443
0.277565      0.003856      0.516443
0.016944      0.178753      0.016443
0.019294      0.833730      0.516443
0.746227      0.333730      0.516443
0.748577      0.678753      0.016443
0.516944      0.832431      0.516443
0.519294      0.177455      0.016443
0.246227      0.677455      0.016443
0.248577      0.332431      0.516443
0.416676      0.241774      0.516443
0.559871      0.674713      0.016443
0.205650      0.174713      0.016443
0.348845      0.741774      0.516443
0.613957      0.380343      0.016443
0.804054      0.542164      0.516443
0.113957      0.630842      0.516443
0.304054      0.469021      0.016443
0.848845      0.269411      0.016443
0.705650      0.836472      0.516443
0.059871      0.336472      0.516443
0.916676      0.769411      0.016443
0.651564      0.130842      0.516443
0.461467      0.969021      0.016443
0.151564      0.880343      0.016443
0.961467      0.042164      0.516443
```

```
EA2      9.487   4.757   4.580  90.243  90.188  89.349 Sym.group:      1
1.0000
  9.486893      0.000000      0.000000
  0.054041      4.756769      0.000000
 -0.014991     -0.019246      4.579857
Mg    Al    O
  4      8     16
Direct
  0.499837      0.633752      0.011361
  0.500082      0.131390      0.482012
  0.813573      0.257696      0.494520
  0.326111      0.625491      0.501746
  0.995267      0.254346      0.992293
  0.160822      0.689054      0.001270
  0.995907      0.760753      0.498354
  0.159742      0.192300      0.491958
  0.811206      0.761223      0.997857
  0.325692      0.125479      0.987935
  0.656355      0.695175      0.503322
  0.656596      0.199917      0.991605
  0.487990      0.763078      0.627771
  0.845518      0.645378      0.347890
  0.623474      0.895186      0.185946
  0.616379      0.395875      0.308861
  0.093745      0.991831      0.185467
  0.092669      0.494591      0.309957
  0.847697      0.118765      0.113434
  0.475636      0.251449      0.875207
  0.327510      0.787484      0.116764
  0.720411      0.975740      0.706398
  0.200804      0.880147      0.683027
  0.975416      0.612789      0.852917
  0.986131      0.108285      0.644081
  0.204805      0.364607      0.830780
  0.718464      0.496262      0.817031
  0.323904      0.257705      0.340590
```

- BESTgatheredPOSCAR——每一代中最好的结构
- non_optimized_structures ——给出弛豫之前产生的结构
- Enthalpies_complete.dat ——给出每一步弛豫所得结构的焓值
- Origin ——显示某结构是从哪个父代通过何种变化操作得到的

ID	Origin	Enthalpy	Parent-E	Parent-ID
1	Random	-23.395	-23.395	[0]
2	Random	-23.228	-23.228	[0]
3	Random	-23.078	-23.078	[0]
4	Random	-23.195	-23.195	[0]
5	Random	-23.155	-23.155	[0]
6	Random	-22.970	-22.970	[0]
7	Random	-23.131	-23.131	[0]
8	Random	-23.017	-23.017	[0]
9	Random	-23.117	-23.117	[0]
10	Random	-23.195	-23.195	[0]

- gatheredPOSCARS_order ——和 gatheredPOSCAR 的内容基本相同，额外给出了每个原子的有序度参量值。

```

EA1      9.346   8.002   2.688 90.000 90.000 90.000 Sym.group:      1
1.0000
  9.346156      0.000000      0.000000
  0.000000      8.002181      0.000000
  0.000000      0.000000      2.688367
Mg     Al    O
  4       8      16

Direct
  0.487956      0.503856      0.516443      1.1399
  0.777565      0.007329      0.016443      1.1399
  0.987956      0.507329      0.016443      1.1399
  0.277565      0.003856      0.516443      1.1399
  0.016944      0.178753      0.016443      1.1915
  0.019294      0.833730      0.516443      1.2474
  0.746227      0.333730      0.516443      1.2474
  0.748577      0.678753      0.016443      1.1915
  0.516944      0.832431      0.516443      1.1915
  0.519294      0.177455      0.016443      1.2474
  0.246227      0.677455      0.016443      1.2474
  0.248577      0.332431      0.516443      1.1915
  0.416676      0.241774      0.516443      1.2914
  0.559871      0.674713      0.016443      1.1408
  0.205650      0.174713      0.016443      1.1408
  0.348845      0.741774      0.516443      1.2914
  0.613957      0.380343      0.016443      1.2355
  0.804054      0.542164      0.516443      1.2161
  0.113957      0.630842      0.516443      1.2355
  0.304054      0.469021      0.016443      1.2161
  0.848845      0.269411      0.016443      1.2914
  0.705650      0.836472      0.516443      1.1408
  0.059871      0.336472      0.516443      1.1408
  0.916676      0.769411      0.016443      1.2914
  0.651564      0.130842      0.516443      1.2355
  0.461467      0.969021      0.016443      1.2161
  0.151564      0.880343      0.016443      1.2355
  0.961467      0.042164      0.516443      1.2161

```

```

EA2      9.487  4.757  4.580  90.243  90.188  89.349 Sym.group:   1
1.0000
  9.486893      0.000000      0.000000
  0.054041      4.756769      0.000000
 -0.014991     -0.019246      4.579857
Mg    Al    O
 4      8     16
Direct
  0.499837      0.633752      0.011361      0.9368
  0.500082      0.131390      0.482012      1.0250
  0.813573      0.257696      0.494520      0.9805
  0.326111      0.625491      0.501746      0.9437
  0.995267      0.254346      0.992293      0.9241
  0.160822      0.689054      0.001270      1.1731
  0.995907      0.760753      0.498354      0.9696
  0.159742      0.192300      0.491958      1.2666
  0.811206      0.761223      0.997857      1.0215
  0.325692      0.125479      0.987935      1.0353
  0.656355      0.695175      0.503322      1.2291
  0.656596      0.199917      0.991605      1.2547
  0.487990      0.763078      0.627771      1.1199
  0.845518      0.645378      0.347890      0.9725
  0.623474      0.895186      0.185946      0.9990
  0.616379      0.395875      0.308861      1.0141
  0.093745      0.991831      0.185467      1.1451
  0.092669      0.494591      0.309957      1.1164
  0.847697      0.118765      0.113434      0.8821
  0.475636      0.251449      0.875207      1.0609
  0.327510      0.787484      0.116764      1.0451
  0.720411      0.975740      0.706398      0.9539
  0.200804      0.880147      0.683027      0.9398
  0.975416      0.612789      0.852917      1.1191
  0.986131      0.108285      0.644081      0.9803
  0.204805      0.364607      0.830780      1.0915
  0.718464      0.496262      0.817031      1.0663
  0.323904      0.257705      0.340590      1.1471

```

- GoodStructures_POSCARs: 是一个在 v.9 版本 USPEX 可用的文件。这个文件非常便于分析，它以稳定性递减的顺序展示所有不同的结构，从最稳定的结构开始，到最不稳定的。
- compositionStatistic: 包含有该成分是由哪几种变化操作产生的信息

Comp/Ratio	Total	Random	Heredity	Mutation	Seeds	COPEX	Best/Convex
[0.0000 1.0000]	30(63)	21	8	1	0	0	33
0 8	4(4)	2	2	0	0	0	0
0 9	2(2)	2	0	0	0	0	0
0 10	5(5)	3	2	0	0	0	0
0 11	2(2)	0	1	1	0	0	0
0 12	6(35)	6	0	0	0	0	29
0 13	2(3)	1	1	0	0	0	1
0 14	0(0)	0	0	0	0	0	0
0 15	2(2)	2	0	0	0	0	0
0 16	5(8)	5	0	0	0	0	3
0 3	1(1)	0	1	0	0	0	0
0 4	1(1)	0	1	0	0	0	0
[0.1250 0.8750]	52(52)	32	9	11	0	0	0
1 7	20(20)	4	8	8	0	0	0
2 14	32(32)	28	1	3	0	0	0
[0.1111 0.8889]	25(25)	9	14	2	0	0	0
1 8	25(25)	9	14	2	0	0	0
[0.1000 0.9000]	16(16)	9	5	2	0	0	0
1 9	16(16)	9	5	2	0	0	0
[0.0909 0.9091]	11(11)	5	4	2	0	0	0
1 10	11(11)	5	4	2	0	0	0
[0.0833 0.9167]	17(17)	8	2	7	0	0	0
1 11	14(14)	8	2	4	0	0	0
2 22	3(3)	0	0	3	0	0	0

- graphical files (*.tif/*.pdf)——快速分析结果
- Folder AuxiliaryFiles 和许多其他文件比可能很少用到
- Energy_vs_N.pdf (Fitness vs N.pdf),——能量（适应度）与结构序号的函数;
- Energy_vs_Volume.pdf ——能量与体积的函数;
- BestEnthalpy.tif (BestFitness.tif)——最佳焓值（适应度）与代数的函数;
- quasiEntropy.tif ——准熵与代数的函数;
- Variation-Operators.pdf——子代与亲代的能量对比; 不同的颜色对应不同的运算方法 (此图可用于评估不同的变异运算的结果) ;
- E_series.pdf——弛豫的第 i 步和第 i+1 步的能量关联; 有助于发现问题和改进弛豫文件输入。
- 对于可变成分有另外的图 extendedconvexhull.tif, 用于显示形成焓与组分的函数。
- 为了给用户提供便利, USPEX 同时还在文件中输出所有独特的低能量结构和组分信息:
 1. goodStructures_POSCARS (and goodStructures_fitness.dat) ——定成分计算;
 2. extended_convex_hull_POSCARS and extended_convex_hull——变成分计算。

在这里，罗列出从最稳定到最不稳定的所有采样结构，它往往足以分析前几个结构。

3.3 Specific 文件夹

用于结构弛豫的可执行文件和列举的输入文件应该放在
subdirectory~/StructurePrediction/Specific 文件夹下，对于不同的交互代码都应如此。

- 对于 VASP，文件 INCAR_1, INCAR_2, ... 等定义了弛豫和能量计算在每一步的弛豫中应该怎样实施（我们推荐至少三步），相应的 POTCAR_1, POTCAR_2 文件是赝势文件。例如：INCAR_1, INCAR_2 对于原子位置和晶胞常数在体积固定的情况下进行了非常粗略的结构弛豫，INCAR_3 在恒定外压下以中等精度进行充分结构弛豫，INCAR_4 则进行精确的计算。每一个高一级的结构弛豫都是从低一级的优化结果开始并且改进它们。POTCAR 文件可以仅仅只是将相关元素文件放在 Specific 文件夹下而被选择性定义。
- 对于 SIESTA，你需要赝势文件和输入文件 input1.fdf, input.fdf
- 对于 GULP，文件 goptions_1, goptions_2, ... 和 gininput_1, gininput_2 必须有。前者指定哪种优化将被实施，后者指定细节（原子间势、压力、温度和优化循环数等）
- 对于 CASTEP，结构文件由 cell_1, cell_2, ... 给出，而计算参数由 param_1, param_2 给出，相应的赝势文件也必须有。
- 对于 CP2K，文件 cp2k_options_1, cp2k_options_2, ... 是必须准备好的。所有的文件都应该是含有所有除了原子坐标和晶胞参数（这些会被 USPEX 写在最后一行“&END FORCE_EVAL”）的所有参数的正常 CP2K 输入文件。因为程序会从 USPEX-1.cell 和 USPEX-pos-1.xyz 文件中读取输出文件，所以“项目名称”应该一直是 USPEX。我们建议是进行弛豫要至少三步（跟 VASP 类似）——第一次优化是固定晶格只优化原子位置，之后的都进行优化。
- 对于 Quantum Espresso，文件 QEspresso_options_1, QEspresso_options_2, ... 是必须准备的。所有的文件都应该是含有所有除了原子坐标、晶胞参数和 k 点（这些会被 USPEX 写在最后的文件里）的所有参数的正常 QE 输入文件。我们建议进行多步弛豫。例如，QEspresso_options_1 只是做了固定晶胞参数弛豫和原子位置的粗略计算，

QE espresso_options_2 是中等精度条件下的加常外压的全结构弛豫，QE espresso_options_3 就是高精度的计算。

3.4 Specific 文件夹中的 INCAR 文件

为了正确运行 USPEX，有一些对 Specific 文件夹中文件的提示，以便使用 USPEX 进行结构弛豫。我们以 VASP 为例子：

- 计算产生的最终结构必须被很好的弛豫，能量和精度收敛。要点是你设置的能量排名必须正确（要检查这个，查看输出中的 Energy-series 文件）。
- 计算的 POTCAR 文件：为了得到正确的结果，你的赝势的核（或者 PAW 势）不应该重叠超过 10-15%。
- 为了以一个低的代价获得精确的弛豫，对每一个结构设置至少三步弛豫。就是说至少有三个 INCAR 文件。通常我们会设 4-5 步弛豫。
- 你的初始结构通常远离局部最小点，在这种情况下，INCAR_1, 2 首先应该在保持体积固定的情况下 (ISIF=4) 弛豫原子和晶胞形状，然后在 INCAR_3, 4 中做完全弛豫 (ISIF=3)，在 INCAR_5 中完成非常精确的单点计算 (ISIF=2, NSW=0)。例外情况：当你做固定晶胞预测，和在进化分子动力学中（除了全弛豫），你必须在 INCAR_1, 2 里设 ISIF=2。
- 当你的体积不改变时，你可以使用默认的平面波截止能。当你再 INCAR 中使用 ISIF=3，你必须将其增加 30-40%，否则你会得到一个大的 Pulay 应力。你的收敛准则一开始比较松，但是在结束时应该紧一点，例如：EDIFF=1e-2 和 EDIFFG=1e-1 在 INCAR_1 中，逐渐的收紧至 EDIFF=1e-4 和 EDIFFG=1e-3 在 INCAR_4 中。timesteps(NSW)的最大值应该足够大保证好的弛豫，但也不能太大以避免在一些差的构型上浪费太多机时。你的系统越大，NSW 就越大。
- 选择一个有效的弛豫算法可以节省很多时间，在 VASP 中，我们推荐开始弛豫时使用共轭梯度 (IBRION=2, POTIM=0.02) 而当结构接近局部小值时，转换至 IBRION=1, POTIM=0.3。

- 即使你研究的是绝缘体系，你采样的许多构型将会是金属性的，所以为了得到收敛的结果，你必须使用“金属化”处理——这对于金属和绝缘体同样适用。我们推荐 Methfessel-Paxton 重叠方案（ISMEAR=1）。对于一个清楚的金属系统，使用 ISMEAR=1, SIGMA=0.1-0.2。对于一个清楚的绝缘系统，我们推荐 ISMEAR=1 和 SIGMA 从 0.1 开始降至 0.03-0.04。

这里，我们提供一个 INCAR 文件的例子，这是关于单元晶胞中有 16 个原子的 C，默认的 POTCAR 中 ENCUT=400ev:

INCAR_1:

PREC=LOW	NSW=55
EDIFF=1e-2	ISIF=4
EDIFFG=1e-1	IBRION=1
NSW=65	POTIM=0.30
ISIF=4	ISMEAR=1
IBRION=2	SIGMA=0.08

INCAR_3:

POTIM=0.02	PREC=NORMAL
ISMEAR=1	EDIFF=1e-3
SIGMA=0.10	EDIFFG=1e-2

INCAR_2:

PREC=NORMAL	ENCUT=520.0
EDIFF=1e-3	NSW=65
EDIFFG=1e-2	ISIF=3
	IBRION=2

POTIM=0.02

INCAR_5:

ISMEAR=1

PREC=NORMAL

SIGMA=0.07

EDIFF=1e-4

INCAR_4:

EDIFFG=1e-3

PREC=NORMAL

ENCUT=600.0

EDIFF=1e-4

NSW=0

EDIFFG=1e-3

ISIF=2

ENCUT=600.0

IBRION=2

NSW=55

POTIM=0.02

ISIF=3

ISMEAR=1

IBRION=1

SIGMA=0.05

POTIM=0.30

ISMEAR=1

SIGMA=0.06

不同于 USPEX, vc-NEB 方法并不需要来自外部代码的结构弛豫, 它通过 vc-NEB 自身并结合外部代码计算得到的力来进行结构弛豫。因此设置文件时会有一些不同。以 VASP 的 INCAR 文件为例, 我们需要设置 NSW=0 去避免结构弛豫, 但是 ISIF=2 或者 3 来提取原子上的力和晶格上的应力张量。我们也建议设置 PREC=Accurate 从而对力和应力有一个好的估计以加速 vc-NEB 的计算。一个针对 vc-NEB 的 INCAR 文件例子如下:

INCAR_1:

PREC=Accurate

EDIFF=1e-4

EDIFFG=1e-3

ENCUT=600.0

NSW=0

ISIF=2

IBRION=2

POTIM=0.02

ISMEAR=1

SIGMA=0.05

4 输入文件 INPUT.txt 的设置

典型的 INPUT.txt 文件在第 12 章的附录中给出。下面我们介绍输入中最重要的参数。许多参数都有可靠的默认值，当你在输入文件中遗漏掉它们时，这些默认值就变得有用（这就允许用户可以有非常短的输入文件！）。那些没有默认值选择的参数就需要用户根据具体的体系进行设置。可以在 <http://han.ess.sunysb.edu> 上查阅在线工具，这些小工具可以帮助用户更好的准备 INPUT.txt 文件，分子文件以及分析一些结果。手册的第 7 章简要地讨论了这些工具用法。

4.1 计算类型的方法

➤ 变量 calculationMethod

含义：明确计算方法

可能值（特性）：

- USPEX—晶体结构预测的发展算法
- META—进化准动态算法
- VCNEB—用变晶胞微动弹性带方法探索过渡态路径
- PSO—修正的 PSO 算法
- TPS—过渡路径抽样方法（迄今未发布）
- MINHOP—最小值跳跃方法（迄今未发布）
- COPEX—另一种新技术，不久将发布

默认：USPEX

格式：

USPEX: calculationMethod➤ 变量 **calculationType**

含义：明确计算类型，也就是不论是体晶、纳米团簇还是表面结构，都可以被预测。这种可变因素由三部分组成，包括维度、分子状态和成分的可变性，以及“s”或“S”表示的自旋选项。

● 维度：

“3”—体单晶

“2”—表面，“-2”—二维晶体

“1”—聚合物

“0”—纳米粒子

● 分子状态

“0”—无分子

“1”—分子计算

● 在计算中化学组成的变化

“0”—定成分

“1”—变成分

➤ 变量 **optType**

含义：明确以何种性质作为目标函数进行优化。

可能值（特性）

值	序号	说明
Enthalpy	1	为了发现稳定相
volume	2	体积最小化（找到最密结构）
hardness	3	硬度最大化（找到最硬相）
Struc_order	4	有序度最大化（发现最有序结构）
Aver_dist	5	在一代内结构差别的平均值的最大化
diel_sus	6	静态电介质极化率最大化（仅适用于 VASP 和 GULP）
gap	7	带隙最大化（仅适用于 VASP）
diel_gap	8	电能存储能量最大化（仅适用于 VASP）
Mag_moment	9	磁化强度最大化（仅适用于 VASP）
quasientropy	10	结构准熵值最大化

与弹性力学相关的性质 (“11**”)

值	序号	说明
K, 体积弹性模量	1101	体积弹性模量最大化
G, 剪切模量	1102	剪切模量最大化
E, 杨氏模量	1103	杨氏模量最大化
v, 泊松比	1104	泊松比最大化
G/K, Pugh 模量比	1105	Pugh 模量比最大化

Hv,维氏硬度	1106	维氏硬度最大化
Kg,断裂韧性	1107	断裂韧性最大化
D, 德拜温度	1108	德拜温度最大化
V _m , 声速	1109	声速最大化
S-波速	1110	S-波速最大化
P-波速	1111	P-波速最大化

注意：与弹性力学相关的性质仅仅适用于 VASP (从 VASP5.1 开始)和 GULP。对于 VASP 用户，需要在 Specific 文件夹中增加至少一个 INCAR_*文件，INCAR_*文件应该包括的参数有 IBRION=6, ISIF≥3 和 NFREE=4。体积弹性模量剪切模量和杨氏模量的估计值是 Voigh-Reuss-Hill (VRH) 的平均值。维氏硬度是用 Chen-Niu 模型【32】计算得到。断裂韧性的优化用理论断裂韧性的最低值作为适用值。

默认：焓值

格式：

Enthalpy: optType

注意：

(1) 如果想要反向优化，需添加负号 (-)。例如：要使静态介电系数最小化，赋上“-diel_sus”。

(2) 如果 optType = gap (带隙) 或 diel_gap (电介质带隙)，这里的带隙我们使用的是一个扩展函数，这个扩展函数对于金属表现出连续性。—也就是说， $\Delta E_g - g(E_F)/N$ ， ΔE_g 是带隙， $g(E_F)$ 是在费米能级处的态密度 (对金属而言)，而 N 是单胞内的原子数。由于这个扩展函数的连续性，与带隙相关量的全局最大化甚至可以用于解决金属的问题。对金属而言，这个值等于费米能级处的态密度。对于半导体和绝缘体而言这个值等于带隙。

图 6 是以硬度为目标优化函数搜索 TiO_2 硬度最大的结构（选择类型=硬度，显示了可能的最大硬度为 14GPa【33】，这是对 Dubrovinsky (2001) 关于 TiO_2 的超硬度的反驳【34】。这是一个很好的例子：一个简单的 USPEX 就可以解决长期存在的争论。

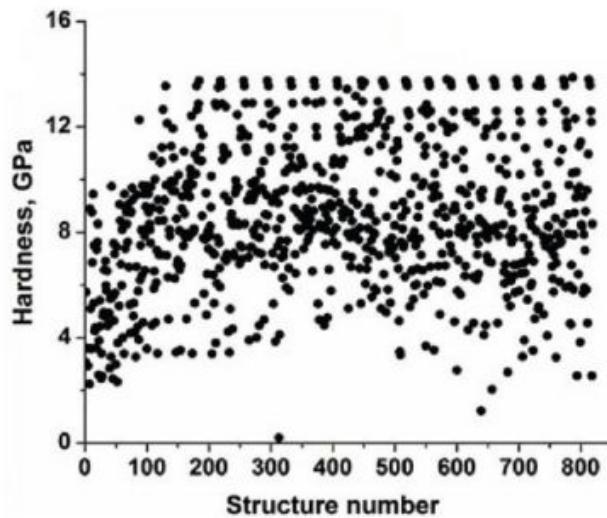


图 6: TiO_2 的最硬结构预测

现在，你需要明确所要计算体系的信息

➤ 变量 atomType

含义：描述每种原子类型的特性

默认：无，必须明确给出

格式：

如果你更喜欢用门捷列夫元素周期表的原子序数，

设置：

% atomType

12 14 8

% EndAtomType

或者，你更喜欢用元素名，设置

% atomType

Mg Si O

% EndAtomType

你也可以二者择一地定义元素的全名，例如：

% atomType

Magnesium Silicon Oxygen

% EndAtomType

➤ 变量 numSpecies

含义：描述每种类型原子的数目

默认：无，必须明确给出

格式：

% numSpecies

4 4 12

% EndNumSpecies

这表示第一种类型有 4 个原子，第二种类型有 4 个原子，第三种类型有 12 个原子

注意：对于变成分的计算，你需要列出如下所示的组成构建块

```
% nunSpecies
```

```
2 0 3
```

```
0 1 1
```

```
% EndNumSpecies
```

这表示第一个构建块的分子式是 A_2C_3 , 第二个构建块的分子式是 BC , A、B、C 在原子类型块中已经解释过。所有结构都满足公式 $xA_2C_3 + yBC$, 其中, $x, y = (0, 1, 2, \dots)$ 或者是 $A_{2x}B_yC_{3x+y}$ 。如果你想做 A-B-C 体系中所有可能组成的预测, 应该这样设置:

```
% numSpecies
```

```
1 0 0
```

```
0 1 0
```

```
0 0 1
```

```
% EndNumSpecies
```

你也可以用一个化学式的计量比数来做定成分计算, 这种情况下, 只需要设置一行 (计算类型=301) 如下为 A_2BC_4 化合物的例子:

```
% numSpecies
```

```
2 1 4
```

```
% EndNumSpecies
```

➤ 变量 valences

含义: 描述每类原子的化合价。这仅用于评估结合的硬度, 而结合硬度又用来计算近似的动力学矩阵 (对软模变异) 和晶体的硬度。

默认： USPEX 有一系列的系统默认化合价（见附录 17）。然而，需要注意的是，对一些元素（例如 N、S、W、Fe 和 Cr 等），有许多可能的化合价。除非你要计算硬度，否则这没有问题，你可以用系统默认化合价。如果你要做硬度计算，你必须明确给出化合价。

格式：

```
% valences
```

```
2 4 2
```

```
% endValences
```

➤ 变量 goodBonds

含义：在矩阵中指定两原子间最低的键价被认为是最重要的键价。就像离子距离矩阵（如下），这是一个上三角形式的方阵。这仅用来计算硬度和软模变异。我们可以用公式

$$\text{goodBonds} = \frac{\text{valence}}{\max_{\text{coordination number}}} \text{ 或者稍微更小。}$$

默认：USPEX 可以给出 goodBonds 合理的默认估值，你可以在输出文件中得到这些数值。对大多数情况，用默认值已经足够，但是对于硬度的计算，你必须要仔细地审查这些数值，或者手动设置。更多细节，请参考附录 18。

格式：

```
% goodBonds
```

```
10.0 10.0 0.2
```

```
0.0 10.0 0.5
```

```
0.0 0.0 10.0
```

```
% EndGoodBonds
```

注意：矩阵的维度必须与原子种类数或原子名称数相等。如果只有一种原子，那么矩阵就填充该数字。上面的矩阵解读如下：可被认为是一个键，Mg-Mg 间的距离要足够短使其键价为 10 或者更大，Mg-Mg 键的距离必须足够短。对 Mg-Si 键，Si-Si 键和 O-O 键也有同样要求（通过使用这种专一的标准，我们有效地排除了来自软模变异和硬度计算的交互作用），然而在用于硬度和软模变异计算的 Mg-O 键的键价取为 0.2 或者更大，Si-O 键的键价取为 0.5 或者更大。

➤ 变量 checkMolecules

含义：打开或关闭原始分子的再弛豫 (Mol_1, Mol_2, ... 文件) 是否完整。可用于分子晶体 (计算类型=310,311)

可能的数值（整数）：

- 0——未经核实，被认为是损坏了的结构或是合并的分子（我们强烈建议不使用这种类型）。
- 1——执行检查命令，所有损坏或是合并的分子均被遗弃。

默认：1

格式：

1: checkMolecules (核实分子)

➤ 变量 checkConnectivity

含义：打开或关闭硬度计算和在软模变异中的连通性相关标准。

可能的值（整数）：

- 0——连通性未经核实，没有硬度计算。
- 1——考虑连通性，计算硬度。

默认：0

格式:

1: checkConnectivity

4.2 种群

➤ 变量 populationSize

含义: 每一代中的结构数。如果需要的话, 第一代可以单独设置。

默认: 当 N 是晶胞中的原子数时 (或者是变成分结构中的 (maxAt) 最大原子数), $2*N$ 四舍五入最接近 10, 那么上限值是 60。通常可以依靠系统默认设置。

格式:

20: populationSize

➤ 变量 initialPopSize

含义: 第一代中的结构数。

默认: 等于 populationSize

格式:

20: initialPopSize

注意: 在许多情况下, 我们建议这两个参数值相等。有时 (尤其是在变成分计算中) 指定 initialPopSize 值大于 populationSize 值更为有用。设置一个很小的初始群也是可能的, 当你想要完全从几个种子结构产生第一代时这样设置就变得很有用。

➤ 变量 numGenerations

含义: 最大模拟代数。当 stopCrit 的最优结构保持不变时, 模拟可以较早地结束。

默认: 100

格式:

50: numGenerations

➤ 变量 stopCrit

含义: 当停止标准代数中的最优结构不再变化, 或者是代数已满的话——无论哪个先发生, 模拟就会停止。

默认: 定成分运行的总原子数, 变成分运行的最大原子数 maxAt。

格式:

20: stopCrit

4.3 适者生存的选择

➤ 变量 bestFrac

含义: 目前这一代中用于产生下一代的比例。

默认: 0.7

格式:

0.7: bestFrac

➤ 变量 keepBestHM

含义: 定义有多少个最优结构将会在下一代中存活。

默认: $0.5 * \text{populationSize}$ 。

格式:

3: keepBestHM

4.4 结构产生和变异操作

➤ 变量 symmetries

含义：晶体可能的空间群，二维晶体/表面的平面群组，或者是团簇的点群。在该列表中用随机选择的对称群，用随机生成的晶格常数和原子坐标产生确定结构数。在这个过程中，从普通位置上将生成特殊的威科夫（Wyckoff）位置。（图 7）。

默认：

- 对三维晶体：2-230
- 对二维晶体/表面：2-17
- 团簇：E C2 D2 C4 C3 C6 T S2 Ch1 Cv2 S4 S6 Ch3 Th Ch2 Dh2 Ch4 D3 Ch6 O D4 Cv3 D6 Td Cv4 Dd3 Cv6 Oh Dd2 Dh3 Dh4 Dh6 Oh C5 S5 S10 Cv5 Ch5 D5 Dd5 Dh5 I Ih

格式：

```
% symmetries
```

```
195-198 200 215-230
```

```
% endSymmetries
```

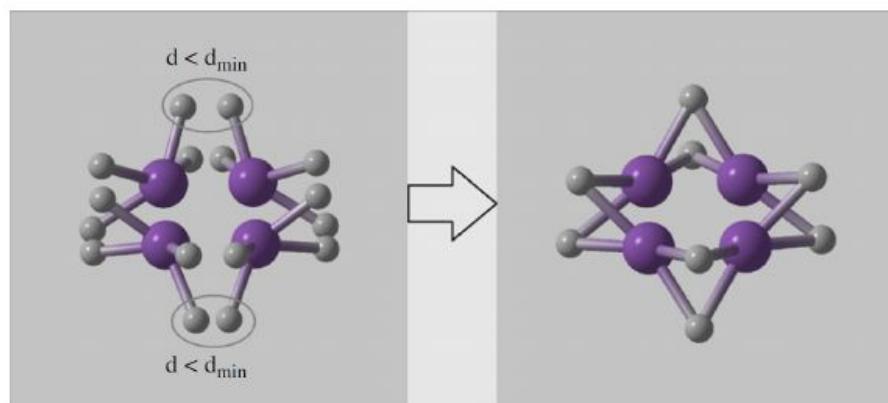


图 7：在特殊的 Wyckoff 位置上合并原子的例子(参考 Ref.【14】)

➤ 变量 fracGene

含义：通过遗传所产生的结构所占的百分比，0.1 代表 10%，等等。

默认：0.5

格式：

0.5: fracGene

➤ 变量 fracRand

含义：通过空间群随机产生的结构所占的百分比。

默认：0.2

格式：

0.2: fracRand

➤ 变量 fracPerm

含义：通过置换所产生的结构所占的百分比，0.1 代表 10%。等等。

默认：如果有超过一种类型的原子/分子，为 0.1，否则是 0。

格式：

0.1: fracPerm

➤ 变量 fracAtomsMut

含义：指定通过软模变异或交换所产生的机构所占的百分比。

默认：0.1

格式：

0.1: fracAtomsMut

注意: 你可以通过指定 softMutTill 用软模变异或协同变异。

➤ 变量 fracRotMut

含义: 通过分子取向变异所产生的结构所占的百分比。0.1 代表 10%，等等。

默认: 分子晶体为 1，其他为 0.

格式:

0.1: fracRotMut

➤ 变量 fracLatMut

含义: 由晶格变异所获得的结构的百分比。0.1 代表 10%，等等。

默认: 固定晶胞预测为 0，其他为 0.1。

格式:

0.1: fracLatMut

注意: 如果所有分数的总和（遗传部分+随机部分+置换部分+...）不等于 1，它们将会被重新调节。

➤ 变量 howManySwaps

含义: 对置换来说，互换原子对的数量由在 1 和 howManySwaps 之间的统一分配所随机决定。

默认: $0.5 \times (\text{可能的互换原子对的最大值})$ ，如果原子 Na 和 Nb，原子 Nc 和 Nd 发生了互换，那么可能的互换原子对的总数是 $\min(\text{Na}, \text{Nb}) + \min(\text{Nc}, \text{Nd})$ ，并且 howManySwaps 的默认设置为 $0.5 \times [\min(\text{Na}, \text{Nb}) + \min(\text{Nc}, \text{Nd})]$ 。在许多情况下，采用默认值是一个不错的选择。

格式:

5: howManySwaps

➤ 变量 specificSwaps

含义：明确在置换中允许哪种原子类型发生互换。

默认：空行，这意味着没有明确的互换，所有的原子都发生了置换。

格式：

```
% specificSwaps
```

```
1 2
```

```
% EndSpecific
```

注意：在这种情况下，第一种类型的原子可以与第二种类型的原子发生互换。如果你想要尝试所有可能的互换，在该关键块里留一空行，或者删除这一部分。

➤ 变量 mutationDegree

含义：在软模变异的位移的最大值的单位是Å。设置了软模变异和协同变异的最大位移矢量使其最大位移等于 mutationDegree。

默认：3×平均原子半径

格式：

```
2.5: mutationDegree
```

➤ 变量 mutationRate

含义：晶格变异的应变矩阵的标准偏差的极小值。应变矩阵部分从高斯分布中随机选择，并且只允许取-1 到 1 之间的值。晶格变异的本质上是把准动力学的思想融入到了我们的方法中【7, 35】，新的结构是由一些异质结构的晶胞扭曲产生的。不像在准动力学中，在我们的方法中，变形是不累积的，所以弹性部分可以足够大去获得新结构。

默认: 0.5

格式:

0.5: mutationRate

将晶格变异与较弱的软模变异结合在一起时一个好选择。

➤ 变量 DisplaceInLatmutation

含义: 将软模变异指定为晶格变异的一部分, 并且设置最大位移的量级为Å。

默认: 1.0

格式:

1.0: DisplaceInLatmutation

➤ 变量 AutoFrac

含义: 程序根据计算的演变, 自动分配对变量算符所占的百分比 (参数控制)。这使得计算速度加快了两倍左右。当要选择用户自定义变异操作时, 设置 AutoFrac=0。

默认: 0

格式:

0: AutoFrac

4.5 约束条件

相同的结构可以再坐标系中以无限的方式展现出来 (模数不变性)。这些大量相同的选择将会导致非常扁平的晶胞单元, 从而为结构弛豫和能量计算带来问题 (例如需要很大力量的 k 点)。约束条件在结晶学中已经熟知, 那就是晶胞角度在 60° 到 120° 之间, 不会移出任何一个多余的和有问题的晶胞 (例如: 晶胞中允许的 $\alpha = \beta = \gamma \sim 120^\circ$ 实际上是扁平的)。因此我们发明了一种特殊的方法: 用最短的晶格矢量来获得特殊的晶胞形状【18, 36】。如果至少有一个晶

格矢量的投影在其他的晶格矢量上或者在对立面上的斜的晶格矢量的长度大于（通过模比较）该晶格矢量长度的一半，就会发生变形。也就是说，对于 a 和 b 电子对，或者是 c 和 $(a+b)$ 电子对，有以下标准：

例如，在标准 (2) 中，新的矢量 a^* 值等于

$$\left| \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|} \right| > \frac{|\mathbf{b}|}{2} \quad (2)$$

$$\left| \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|} \right| > \frac{|\mathbf{a}|}{2} \quad (3)$$

$$\left| \frac{\mathbf{c} \cdot (\mathbf{a} + \mathbf{b})}{|\mathbf{c}|} \right| > \frac{|\mathbf{c}|}{2} \quad (4)$$

$$\left| \frac{\mathbf{c} \cdot (\mathbf{a} + \mathbf{b})}{|\mathbf{a} + \mathbf{b}|} \right| > \frac{|\mathbf{a} + \mathbf{b}|}{2} \quad (5)$$

例如，在标准 (2) 中新的矢量值 a^* 等于

$$\mathbf{a}^* = \mathbf{a} - ceil\left(\frac{|\mathbf{a} \cdot \mathbf{b}|}{|\mathbf{b}|^2}\right) sign(\mathbf{a} \cdot \mathbf{b}) \mathbf{b} \quad (6)$$

这种变形迭代进行，完全避免了劣质的晶胞形状，从而解决了问题。在该变形中，原子坐标部分发生了改变，以保证初始结构和改变后的结构是完全相同的（在变形过程中，原子的笛卡尔坐标保持不变）。

➤ 变量 minVectorLength

含义：设置最新产生的结构的晶胞参数的最小长度。

默认：1.8 最大原子的共价直径。对于分子晶体（计算类型=310, 311），默认最大值是 1.8 (MolCenters)。

格式：

2.0: minVectorLength

当原子间距离足够小时，通常使用的计算方法（如赝势，PAW，LAPW和其他参数化的力场）将不会起作用。这种情况需要避免，你可以用上三角形式的离子距离方阵指定每对原子之间的最小距离。

➤ 变量 IonDistances

含义：在不同的原子类型之间设置最小原子距离矩阵。小于离子距离的数值在物理意义上是没有意义的，需要严格避免。

默认：A 和 B 原子之间的离子距离被估计为 $0.2(V_A^{1/3} + V_B^{1/3})$ ，但是不会超过 0.8Å，在分子类晶体计算中是 $0.45(V_A^{1/3} + V_B^{1/3})$ ，此处 V_A 和 V_B 是 USPEX 中所估计得到的 A 原子和 B 原子的默认体积值。

格式：

```
% IonDistances
```

```
1.0 1.0 0.8
```

```
0.0 1.0 0.8
```

```
0.0 0.0 1.0
```

```
% EndDistances
```

注意：矩阵的大小必须等于原子种类数。如果上面例子中的化合物时 MgSi₃，那么矩阵应该按如下解释：在一个新产生的结构中允许的 Mg-Mg 距离为 1.0 Å。Mg-Si 距离、Si-Si 距离和 O-O 距离也是 1.0 Å，而最小的 Mg-O 距离和 Si-O 距离是 0.8 Å。你可以把特定体系的更多信息放进这个键长矩阵。举例来说，如果你知道镁原子倾向于离得很远，在你的体系中最近不会小于 3Å，你可以列入该信息。然而需要注意的是最小距离越大，计算产生符合这些约束条件的结构就越困难(特别是大体系)。所以尽量折中，并且谨记离子距离必须远远小于实际的键长。

➤ 变量 constraint_enhancement

含义：可以用 IonDistances 矩阵（通过变为 constraint_enhancement 的倍数）来严格约束计算产生的对称随机结构（适用于所有变异操作，未增强的离子距离矩阵依然可以应用），。允许你应用离子距离矩阵中严格的约束条件（通过乘以增强约束条件）。确保你在知道自己正在用的是什么意义的时候使用。

默认：1

格式：

1: constraint_enhancement

对分子晶体，下面的关键块体非常重要。

➤ 变量 MolCenters

含义：分子中心之间的最小距离矩阵。任何小于那些明确给出的分子的大重叠的距离，在物理意义上都是不合理的，需要严格避免。

默认：零矩阵适用于非分子计算，分子晶体没有默认（必须明确给出）。

格式：

% MolCenters

5.5 7.7

0.0 9.7

% EndMol

注意：在上面的例子中，有两种分子类型。在所有产生的结构中，第一种类型的分子的几何中心之间的距离必须至少为 5.5 Å（A-A 距离），第一种类型和第二种类型分子几何中心之间的距离——7.7 Å（A-B 距离），第二种分子几何中心之间的距离——9.7 Å（B-B 距离）。

4.6 晶胞

在结构弛豫之前，指定在给定条件下的单胞体积对产生所有新结构是很有用的。这个指定的初始体积可以再 Latticevalues 这一项中输入。

➤ 变量 Latticevalues

含义：明确给出单胞或抑制晶格参数的初始体积。

默认：对于晶胞体积不需要明确给出值——USPEX 有一个强大的计算体系可以给出任何压力下的合理的估计值。

格式：

% Latticevalues

125.00

% Endvalues

注意：（1）该体积仅用于初始的猜测，并且只影响第一代，每个结构都经过充分优化，且采用与最低（自由）能相对应的体积。这个关键部分还有其他用途：当你知道晶格参数时（例如从实验中），你就可以在晶格参数中明确指定，而不是使用单胞体积，例如：

% Latticevalues

7.49 0.0 0.0

0.0 9.71 0.0

0.0 0.0 7.07

% Endvalues

你也可以通过列出 a、b、c、 α 、 β 和 γ 值来指定晶胞参数。

% Latticevalues

10.1 8.4 12.5 90.0 101.3 90.0

% Endvalues

注意：如果你要做定成分的单斜晶胞的计算，则要做专门的 β 角度设置（标准设置）。

(2) 对于变成分计算，需要明确给出计算搜索空间最后一个的体积。例如：

% Latticevalues

14.0 11.0

% Endvalues

(3) 用户不需要给出晶格常数关键子块中单胞或原子的体积——一个专门的算法已经被编写进 USPEX 中来对给定压力下体系的体积进行精确估计，因此不需要用户专门去设定它。这个选项很好地适用于需要在 INUPT 中设置体积的所有计算类型 (3**, 2D-crystals, 110, 000)，你也可以用在线程序 http://han.ess.sunysb.edu/volume_estimation 去估算体积。用户也可以手动输入体积。

(4) 如果你研究压力下的分子晶体，有时你可能需要稍微增加初始体积，为了能够用随机对称算子产生结构。

➤ 变量 splitInto

含义：定义了单胞中的子晶胞或伪子晶胞数。如果你不想使用分裂的方法，就用数值 1，或者删掉该块体。分裂选项仅适用于单胞中原子数大于 25-30 的情况。

默认： 1

格式：

% splitInto (该单胞分裂所形成的子晶胞数)

1 2 4

%EndSplitInto

子晶胞采用了其他的平移对称性（伪势）。除此之外，每个子晶胞可以使用一种特别的算法通过空间群构建，这个算法由 A.R. Oganov 和 H.T. Stokes 开发，且由 H.T. Stokes 执行的特殊算法（见参考文献【14】）。

4.7 重新开始

如果某些作业运行出错，你想要在它停止的点重新开始计算——或者是更早的点。如果你想做的就是从停止的那一点处重新开始运行，你不需要改变任何设置（所有信息都会被存储在*.mat 文件中），删掉 still_reading 文件，再次运行 USPEX 就足够了。

如果你想要在一个特殊结果文件夹中的特殊的一代重新开始，那么给出 pickUpYN=1, pickUpGen=你想要重新开始的代数。pickUpFolder=results 文件夹的编号（例如 resulsi1 为 1, results2 为 2, ...）也是从你想要重新开始的那里开始。如果 pickUpGen=0 就是要重新开始计算。三个参数的默认值都为 0。例如，想要从第 10 代开始在 results10 文件夹中重新启动运算，需设置：

1 : pickUpYN

10 : pickUpGen

5 : pickUpFolder

4.8 从头算法的细节

USPEX 倾向于强大的两级并行组合，这使得它的并行扩展性很难与许多其他计算方法匹配。并行第一级是在结构弛豫代码里执行的，使每个结构可以达到约 $10\text{-}10^2$ 个 CPUs 的高效率。并行第二级是把计算分配给了统一群中的个体（因为同一种群内的结构之间是彼此独立的）——也就是你通过设置 numParallelCals 模块同时进行结构弛豫的数目最大可以等于 populationSize（也可以是典型的 $10\text{-}10^2$ ）。

USPEX 的一个有趣的特征是具有结合不同水平理论的可能性——例如，在开始的时候，你通常并不需要执行高度精确的结构弛豫（仅仅最后几步的结构弛豫需要高度精确），在这种情况下，你可能想要在较低成本的情况下进行结构弛豫，例如刚开始用有原子间性质的（例如，用 GULP 代码），到最后转换成更精确的计算——例如，用 VASP。然而，使用这种设置时须注意，低成本的计算依然需要构建得有意义。该特点是由 abinitioCode 模块控制。例如：

➤ 变量 abinitioCode

含义：定义应用于每一步优化的准则。

默认：每一步优化都为 1 (VASP)。

格式：

```
% abinitioCode
```

```
3 2 2 1 1
```

```
% ENDabinit
```

注意：数字表示用于每一步结构弛豫所用的软件：1—VASP，2—SIESTA，3—GULP，4—LAMMPS，5—Neural Networks code，6—DNACRYS，7—CP2K，8—Quantum Espresso，9—FHI-aims，10—ATK，11—CASTEP，12—Tinker，13。

➤ 变量 KresolStart

含义：明确而给出 k 点网格倒易空间的分辨率（单位： $2\pi\text{\AA}^{-1}$ ）

默认：线性的从 0.2 到 0.08

格式：

```
% KresolStart
```

```
0.2 0.16 0.12 0.08
```

% Kresolend

注意：你可以输入几个数值(每一步结构弛豫对应一个数)，开始时粗略些（也就是值大一些），到最后，精确度高，这样就明显地加快了计算速度。特别是对金属而言，因为需要很多 k 点。如果你想用 VASP 或 QuantumEspresso，那么这个模块非常重要(GULP 完全没必要用它，而对于 SIESTA，你必须在 SIESTA 输入文件里定义 KresolStart)。

对于团簇，2D 晶体和表面，你必须要明确给定团簇（或表面）真空范围的厚度。

➤ 变量 vacuumSize

含义：定义在结构周围增加的空间的数量（在毗邻的单胞内邻近团簇之间的最近邻距离，Å）。仅适用于表面，2-D 晶体和纳米微粒。

默认：每一步弛豫设为 10Å.

格式：

```
% vacuumSize
```

```
10 10 15 20 20
```

```
% endVacuumSize
```

➤ 变量 numParallelCalcs

含义：指出在并行计算中结构弛豫的数目。

10: numParallelCalcs

你需要提供提交作业文件或者你正在使用的准则/模式下的可执行文件的名字。

➤ 变量 commandExecutable

含义：明确给定提交作业文件名或在给定准则下的可执行文件名。

默认：无默认值，需要用户自己设定。

格式：

```
% commandExecutable
```

```
gulp < input > output
```

```
mpirun -np 8 vasp > out
```

```
mpirun -np 8 vasp > out
```

```
mpirun -np 8 vasp > out
```

```
% EndExecutable
```

注意：每一行都对应于一个弛豫阶段——第一行描述了第一个弛豫阶段的执行，等等。例如：从头算准则等于“3111”意味着第一个弛豫阶段是执行 GULP，接下来的阶段会通过“mpirun -np 8 vasp > out 用 VASP”命令执行。如果在可执行命令中只显示出一行，那么所有的弛豫阶段都执行相同的命令。

实际上在你远程提交作业模式下几乎所有平台都可以使用 USPEX，你所需要的是在你的工作站有 MATLAB/Octave 运行。在这种情况下，你的工作站准备开始输入（包括 jobs），然后提交到远程计算节点上，计算完成时查看结果并分析，准备新的输入文件。传出和来往的数据不是很大，所以网络不需要非常快。当然，作业提交依赖于机器。

➤ 变量 whichCluster

含义：明确给定作业提交的类型

可能的值(整数)：

- 0——无作业脚本；
- 1——本地提交；
- 2——远程提交。

默认: 0

格式:

1: whichCluster

➤ 变量 remoteFolder

含义: 计算即将进行的远程超级计算机上的文件夹。只有当 whichCluster=2 时, 这个文件夹才会被激活。

默认: 没有

格式:

Blind test : remoteFolder

注意: 在远程提交文件夹中有一个相似的参数设置——homeFolder, 计算的真实路径实际上是*homeFolder*/*remoteFolder*/CalcFolderX, 在此处, X=1, 2, 3....

➤ 变量 PhaseDiagram

含义: 使计算类型=300、301 的相图的计算能够进行。这给出了一个想法 (比较粗劣, 仅仅是一个大致的结果): 给出新相可能会在较高或较低压力下, 同时粗劣的给出一个转变压力。

默认: 0

格式:

1: phaseDiagram

4.9 指纹设置

在改变下面这些参量之前, 建议你先读一下关于指纹概念和方法论的文章, 这样你才能知道你在设置的是什么 (Oganov & Valle, 2009 【22】)。

0.05 : sigmaFing

0.10 : deltaFing

10.0 : RmaxFing

0.015 : toleranceFing (如果距离小于公差, 结构是相同的)

0.015 : toleranceBestHM (如果距离小于公差, 结构是相同的)

SigmaFing 是原子间距的高斯扩展(0.05Å是可以的)。

deltaFing 是指纹函数的离散化 (单位为Å), (0.10Å是可以的)。

RmaxFing 是截断距离 (10Å是可以的, 但是有时你想将值取大或是取小一点, 依赖于体系或是你想要做什么)。

toleranceFing (默认值为 0.008) 和 toleranceBestHM (默认值为 0.04) 指定了结构之间的最小余弦距离, 证明它们是不相同的——分别用于产生子代和适者生存。它们依赖于结构弛豫的精度和体系的物理性质 (例如对于排序问题, 分别属于不同结构的指纹非常相似, 并且参数公差也设得非常小) 只有当 dynamicalBestHM=1 时 (我们已不用了), toleranceBestHM 才是有用的, 所以这个参数基本上被废弃。

4.10 反种子设置

反种子技术已经在 USPEX 中开发实施了, 所有基于对结构不利的想法都已被屏蔽以保证模拟没有陷入局部最小值。在这儿, 和时间有关的适应度是所有总的真实焓值 (或其他的适应性质) 和跟历史有关的项, 它是能量抽样面上叠加的高斯势能之和:

$$f = f_0 + \sum_a W_a \exp\left(-\frac{d_{ia}^2}{2\sigma_a^2}\right)$$

在这里, f 是适应度 (f_0 ——真实适应度, f ——跟历史有关的适应度), W_a 是高度, σ_a 是高斯展宽。在我们的方法中, 高斯参数的变化依赖于种群多样性, 能量在每一代中扩展。

有三种使用该技术的方法。第一种，将你认为不利的结构放置在反种子文件夹中。例如它可能是基态结构——在这种情况下，USPEX 将会尽力找到第二个最低焓值的结构。

在第二种和第三种方法中，你不需要明确给定反种子结构——USPEX 要么用所有的抽样结构作为反种子（经过很好的测试，是被推荐的方法），要么是每一代中的最优结构。你需要给定以下的一些设置：

➤ 变量 antiSeedsActivation

含义：明确从哪一代开始反种子模式开始。当 $\text{antiSeedsActivaton}=N>0$ 时，从第 N 代开始，高斯势加入到所有的结构中，当 $N<0$ 时，高斯势仅加入到每一代的最优结构中，从第 N 代开始。当 $n=0$ 时，高斯势仅加入到放在反种子文件夹中的结构中。如果你不想用反种子，那么给定非常大的 $\text{antiSeedsActivation}$ 值（例如 5000）和 $\text{antiSeedsMax}=0.0$ 。

默认：0

格式：

1: antiSeedsActivation

➤ 变量 antiSeedsMax

含义：明确高斯的高度，以每一代中的焓值的平均方差为单位（算出来的值仅在 bestFrac 之间，例如在准双亲之间）。我们推荐 $\text{antiSeedsMax}=0.01$ 。

默认：0.000

格式：

0.005: antiSeedsMax

➤ 变量 antiSeedsSigma

含义：明确高斯的宽度，以每一代中结构之间的平均距离为单位（算出来的值仅在 bestFrac 之间，例如在准双亲之间），我们建议 $\text{antiSeedsSigma}=0.005$ 。

默认: 0.001

格式:

0.005: antiSeedsSigma

图 8 是一个应用反种子技术的例子。

4.11 判定空间群

➤ 变量 doSpaceGroup

含义: 判定空间群, 还以结晶学的*.CIF 格式 (当你要准备发表的时候这会更简单一点, 但是需要注意的是: 如果弛豫不是非常精确, 或者是如果对于对称探测器设置了非常严格的公差, 那么有时候空间群可能是不确定的) 写在输出文件中。多亏了由 H.T. Stokes 提供的强大的对称准则, 这个选项才得以被使用。

默认: 如果计算类型-3** (300, 301, 310, 311-体单晶) 为 1, 否则为 0。

格式:

1 : doSpaceGroup (0——没有空间群, 1——判定空间群)

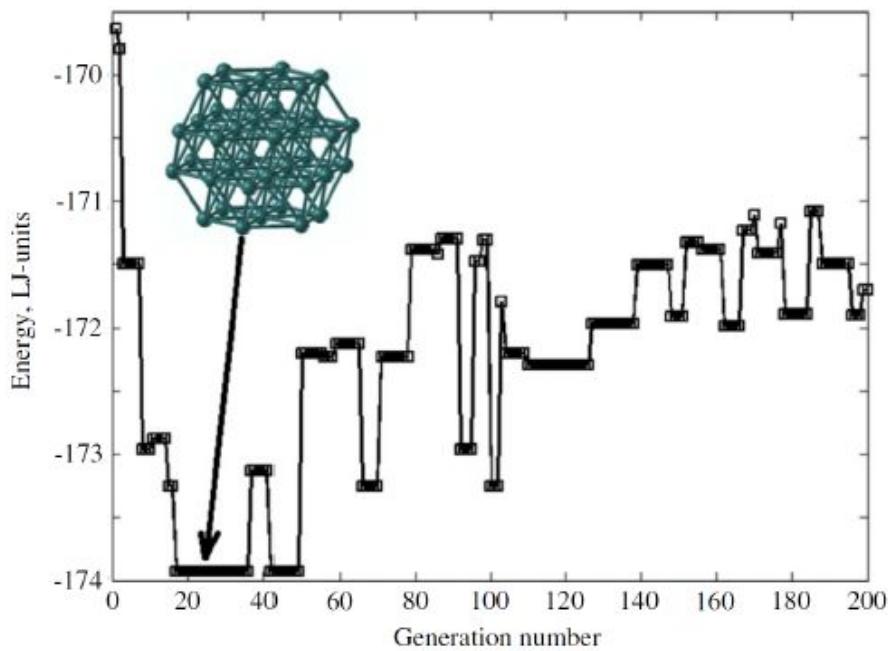


图 8：应用反种子的有 38 个原子的 Lennard-Jones 集群计算的例子。每一代中的最优结构的能量已被标绘出。我们可以清晰地看到，算法没有陷在任何一个候选最小值很长时间，并且很快发现了基态。在这儿我们用 antiSeedsActivation=1, antiSeedsMax=0.01, antiSeedsSigma=0.001。

➤ 变量 SymTolerance

含义：用 H.T. Stokes 的寻找对称性代码精确地确定对称性。可以设置为一个数字（单位为 Å），或者是 high/medium/low (=0.02/0.08/0.20)。

默认： medium

格式：

. medium : SymTolerance

4.12 统计学——针对开发者

➤ 变量 repeatForStatistics

含义： USPEX 自动运行的数量。 USPEX 模拟是随机的，用同样的输入参数再次模拟，得到的结果是必然不同的。然而最终的结果——基态值——是相同的（也是我们所期望的）。为了达到它所运行的步数和化学空间中的轨道数在每次运行时也是不同的。为了比较不同的算法，你必须收集一些统计资料——不要只依赖于单次运行（它可能是好运的，也可能不是好运的，USPEX 不能只依赖于好运）。这个选项仅是开发者感兴趣，而且仅对用简单的势函数（如 GULP）收集统计资料有意义。

默认： 1 （也就是说，没有统计资料将被收集）。

格式：

20 : repeatForStatistics

➤ 变量 stopFitness

含义： 设置适应度的值就可以当达到适应度值 \leq stopFitness 时，计算停止。

默认： 没有默认值，需要用户自己设置。

格式：

90.912 : stopFitness

注意： 当 stopFitness 值给定之后，统计资料才能够自动分析。为了收集 stopFitness 的可获取的统计资料建议 repeatForStatistics 模块部分设置的值 >1 。

输出样本如下：

```

Number of files to be processed: 20
Target enthalpy: 90.912

Generation: 23 Number: 1326 Enthalpy: 90.9119 Mat-file: /home/USPEX/01/results1/USPEX.mat
Generation: 22 Number: 1224 Enthalpy: 90.9119 Mat-file: /home/USPEX/02/results1/USPEX.mat
Generation: 60 Number: 3451 Enthalpy: 90.9119 Mat-file: /home/USPEX/03/results1/USPEX.mat
Generation: 30 Number: 1739 Enthalpy: 90.9119 Mat-file: /home/USPEX/04/results1/USPEX.mat
Generation: 17 Number: 956 Enthalpy: 90.9119 Mat-file: /home/USPEX/05/results1/USPEX.mat
Generation: 36 Number: 2055 Enthalpy: 90.9119 Mat-file: /home/USPEX/06/results1/USPEX.mat
Generation: 35 Number: 1987 Enthalpy: 90.9119 Mat-file: /home/USPEX/07/results1/USPEX.mat
Generation: 22 Number: 1241 Enthalpy: 90.9119 Mat-file: /home/USPEX/08/results1/USPEX.mat
Generation: 18 Number: 1002 Enthalpy: 90.9119 Mat-file: /home/USPEX/09/results1/USPEX.mat
Generation: 29 Number: 1641 Enthalpy: 90.9119 Mat-file: /home/USPEX/10/results1/USPEX.mat
Generation: 21 Number: 1197 Enthalpy: 90.9119 Mat-file: /home/USPEX/11/results1/USPEX.mat
Generation: 27 Number: 1542 Enthalpy: 90.9119 Mat-file: /home/USPEX/12/results1/USPEX.mat
Generation: 44 Number: 2519 Enthalpy: 90.9119 Mat-file: /home/USPEX/13/results1/USPEX.mat
Generation: 32 Number: 1821 Enthalpy: 90.9119 Mat-file: /home/USPEX/14/results1/USPEX.mat
Generation: 15 Number: 835 Enthalpy: 90.9119 Mat-file: /home/USPEX/15/results1/USPEX.mat
Generation: 43 Number: 2477 Enthalpy: 90.9119 Mat-file: /home/USPEX/16/results1/USPEX.mat
Generation: 40 Number: 2278 Enthalpy: 90.9119 Mat-file: /home/USPEX/17/results1/USPEX.mat
Generation: 24 Number: 1358 Enthalpy: 90.9119 Mat-file: /home/USPEX/18/results1/USPEX.mat
Generation: 14 Number: 757 Enthalpy: 90.9119 Mat-file: /home/USPEX/19/results1/USPEX.mat
Generation: 27 Number: 1532 Enthalpy: 90.9119 Mat-file: /home/USPEX/20/results1/USPEX.mat

Found structures numbers : 1326 1224 3451 1739 956 2055 1987 1241 1002 1641 1197 1542 2519 1821 835 2477 2278 1358 757 1532
Found generations numbers: 23 22 60 30 17 36 35 22 18 29 21 27 44 32 15 43 40 24 14 27

Success rate: 100 percent
Average number of generations to get E=90.912: 29
Average number of structures to get E=90.912: 1647
Standard deviation: 670

```

4.13 不常使用的关键字

➤ 变量 ordering active

含义：通过局部有序参数打开有偏见的变量算符。

默认：1

格式：

1: ordering_active

➤ 变量 symmetrize

含义：把所有结构转换为可以适应对称性的结晶学设置

默认：0

格式：

1: symmetrize

➤ 变量 valenceElectr

含义：每种原子类型的价电子数。

默认：所有原子的值都是常量，我们已经将它们制成了表，不需要特别设置。

格式：

```
% valenceElectr
```

```
2 6
```

```
% endValenceElectr
```

➤ 变量 percSliceShift

含义：在所有维度内进行块移动（在遗传中用到）的概率。1.0 代表 100%。

默认：1.0

格式：

```
0.5 : percSliceShift
```

➤ 变量 dynamicalBestHM

含义：当 keepBestHM 为上限值时，设置在计算中存活的最优结构的数量是否不同。

可能的值（整数）：0=没有突变；1 和 2=看注解

默认：2

格式：

```
1 : dynamicalBestHM
```

注意：如果你设置 `dynamicalBestHM=1`，那么代码就会选择 `keepBestHM` 的最低能量结构（没有重复结构，这可以定义为有低于用户定义的 `toleranceBestHM` 的指纹距离）。如果 `dynamicalBestHM=2`（我们的优先选择），在与 `bestFrac` 相一致的全部能量间隔内，聚类算法恰好选择了 `keepBestHM` 的最大不同的结构，并且最佳 `toleranceBestHM` 值是自动决定的——在保留最优结构记忆的同时促进了多样性的发展。

➤ 变量 `softMutOnly`

含义：仅由软模变异可以产生多少代。

默认：0

格式：

```
% softMutOnly
```

1-5

```
% EndSoftOnly
```

注意：在上面的例子中，仅由软模变异产生的代数可以达到 5 代（当然，包括第一代）。注意根据软模变异，每个亲代产生两个软模变异突变体。你也可以在运行的过程中给定经过软模变异的特殊的几代，例如，每 10 代软模变异一次，你可以这样写：

```
% softMutOnly
```

2 12 22 32 42

```
% EndSoftOnly
```

➤ 变量 `maxDistHeredity`

含义：给定参与遗传的结构之间的最大余弦距离。它指定了与该结构能够匹配的能量面半径。

默认：0.5

格式:

0.5 : maxDistHeredity

➤ 变量 manyParents

含义: 明确是否应该要超过两片 (或超过两个亲代结构) 被用于遗传。这对大体系来说是有用的。

可能的值 (整数) :

0——仅用两个亲代, 每个一个片段;

1——许多结构被用作亲代, 每个一个片段;

2——两个结构被用于亲代, 许多片段 (用参数 minSlice 和 maxSlice 动态决定) 可以从每个亲代中独立选择;

3——两个结构被用于亲代, 许多片段 (用参数 minSlice 和 maxSlice 不断变化决定) 是从有特定设置的晶胞中切下来。这是适用于大体系的较好的选择。例如: 我们将两个结构切成近似相等厚度的薄片, 然后从亲代 1 中选择偶数片, 从亲代 2 中选择奇数片, 制成多层的“三明治”或是“斑马”。

默认: 0

格式:

3 : manyParents

minSlice, maxSlice: 决定了切片的最小和最大厚度, 单位为 Å, 他们将会从亲代中删掉来参加子代结构的产生。我们期望这些薄片足够薄, 能够携带一些亲代的信息 (但是不太薄, 以至于对遗传没有影响)。这些参数的合适值分别大约是 1 和 6 Å。

对于团簇, 你可以直接给定参与遗传的亲代数 (但是我们发现这几乎没有用)。

➤ 变量 numberparents

含义：定义了团簇中遗传的亲代数。

默认：2

格式：

2: numberparents

5 特殊情况的额外输入

5.1 MOL_1, MOL_2, ...等文件

5.1.1 分子晶体, calculationType=310/311

对于一个分子晶体, MOL_1 文件描述分子结构开始构建的结构部分。这个文件也定义如果分子可活动的哪个扭转角会突变。这个文件和它的格式跟 SIESTA's Z_Matrix 文件 (MOL_1 给出原子的笛卡尔坐标, 反之, Z_Matrix 文件从键长、键角和扭转角定义原子位置)不同。A_Matrix 文件是利用 MOL_1 文件给的信息创建的, 即: 键长和所有必要的角度都是用笛卡尔坐标计算的。重要的键长和角度用来创造Z_Matrix—这恰好是5-7列列举的。我们看一下苯C₆H₆的MOL_1文件:

Benzene							
Number of atoms: 12							
H	0.0000	0.0000	0.0000	0	0	0	1
C	1.0600	0.0000	0.0000	1	0	0	1
C	1.7550	1.2038	0.0000	2	1	0	1
C	3.1450	1.2038	0.0000	3	2	1	0
C	3.8400	-0.0000	0.0000	4	3	2	0
C	3.1450	-1.2038	0.0000	5	4	3	0
C	1.7550	-1.2038	0.0000	6	5	4	0
H	1.2250	2.1218	0.0000	3	2	1	0
H	3.6750	2.1218	0.0000	4	3	8	0
H	4.9000	0.0000	0.0000	5	4	9	0
H	3.6750	-2.1218	0.0000	6	5	10	0
H	1.2250	-2.1218	0.0000	7	6	11	0

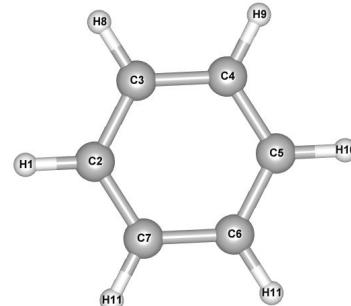


图 9: MOL_1 文件举例和相应的分子结构

第一个原子是 H, 它的坐标定义与其他原子无关 (“0 0 0”)。

第二个原子是 C, 它在 Z_matrix 的坐标 (在分子坐标系中) 仅根据它到第一个原子 (即, 上面描述的 H) 的距离来设置, 但没有角度-- (“1 0 0”)。

第三个原子是 C，它的坐标将根据它到第二个原子的距离来设置，键角 3-2-1，但不是扭转角—因此我们用“2 1 0”。

第四个原子是 C，它的坐标根据它到第三个原子的距离设置，键角 4-3-2，扭转角 4-3-2-1—因此，我们用“3 2 1”等等...知道最后，第 12 个原子，H，根据它到第七个原子（C）的距离定义，键角 12-7-6，扭转角 12-7-6-11—因此是“7-6-11”。

最后一列是扭转角灵活度的标示。例如，在 C4 中，扭转角由 4-3-2-1 定义。理想情况下，对于前 3 个原子来说这个标示应该是 1，其他原子的是 0，如果存在其他活动的扭转角，需在这一列说明是 1。

5.1.2 聚合物晶体，`calculationType=110`（“线性链状模型”）

对于聚合物，`MOL_1` 文件用来表示单体单元的几何结构，和分子晶体一样，除了要在最后一行列出活性原子，如下图列出的 PVDF（聚偏二氟乙烯）的 `MOL_1` 文件：

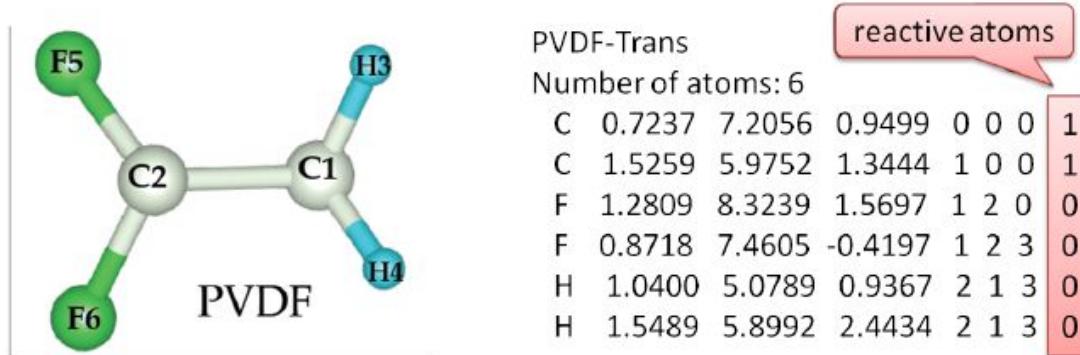


图 10：PVDF 的 `MOL_1` 文件及相应的单体结构图示

5.1.3 经典力场的额外输入

上面的 `MOL_1` 文件适用于 USPEX 的一般情况。然而，一些基于编码的经典力场需要额外的信息。例如，GULP 需要列出化学分类和电荷。阿司匹林的 `MOL_1` 文件的写法如下：

Aspirin_charge

Number of atoms: 21

H_1	0.2310	3.5173	4.8778	0	0	0	1	0.412884
O_R	0.7821	4.3219	4.9649	1	0	0	1	-0.676228
C_R	0.4427	5.0883	6.0081	2	1	0	1	0.558537
O_2	-0.5272	4.5691	6.6020	3	2	1	0	-0.658770
C_R	1.0228	6.3146	6.3896	3	2	4	0	0.116677
C_R	2.1330	6.8588	5.6931	5	3	2	0	0.311483
C_R	0.4810	7.0546	7.4740	5	3	6	0	-0.119320
O_R	2.8023	6.2292	4.6938	6	5	3	0	-0.574557
C_R	2.6211	8.1356	6.0277	6	5	8	0	-0.083091
C_R	0.9966	8.3146	7.8237	7	5	3	0	-0.103442
H_2	-0.3083	6.6848	8.0128	7	5	10	0	0.198534
C_R	3.6352	5.1872	4.9079	8	6	5	0	0.609295
C_R	2.0623	8.8613	7.0940	9	6	5	0	-0.119297
H_2	3.3963	8.5283	5.4906	9	6	13	0	0.174332
H_2	0.5866	8.8412	8.6013	10	7	13	0	0.205960
O_2	3.9094	4.7941	6.0632	12	8	6	0	-0.588433
C_3	4.2281	4.5327	3.7638	12	8	16	0	-0.271542
H_2	2.4227	9.7890	7.3367	13	9	10	0	0.196738
H_2	3.4269	4.1906	3.1183	17	12	8	0	0.151315
H_2	4.8283	3.6848	4.0792	17	12	19	0	0.131198
H_2	4.8498	5.2464	3.2337	17	12	19	0	0.127726

这里，标题的关键词电荷是告诉程序在最后一行读取电荷。

结合 Tinker 一起计算，附加行用来列出原子标签，如下所示：

Urea

Number of atoms: 8

C	0.000000	0.000000	0.000000	0	0	0	1	189
O	0.000000	0.000000	1.214915	1	0	0	1	190
N	1.137403	0.000000	-0.685090	1	2	0	1	191
N	-1.137403	0.000000	-0.685090	1	2	3	0	191
H	1.194247	0.000000	-1.683663	4	1	3	0	192
H	-1.194247	0.000000	-1.683663	4	1	3	0	192
H	1.998063	0.000000	-0.138116	2	1	3	0	192
H	-1.998063	0.000000	-0.138116	2	1	3	0	192

5.1.4 如何准备 MOL 文件

有许多程序可以生成 Zmatrix style 文件，如 Molden、Avogadro 等。有经验的用户或许有自己的方法来准备这些文件。为了用户的方便，我们创建了一个在线工具，允许用户仅用 XYZ 格式的文件来生成 USPEX-style MOL 文件。这个工具的网址为 <http://han.ess.sunysb.edu/zmatrix>。

5.2 表面

为了完成表面重构的预测，必须：

- 指定 200 或 201：计算类型
- 准备包含基底的 VASP5 POSCAR 格式的文件，如图 11 所示。
- 列举以下参数：

```
% symmetries
```

2-17

```
% endSymmetries
```

注意：如果对称标签 symmetries 出现，USPEX 将尝试使用平面组生成结构。

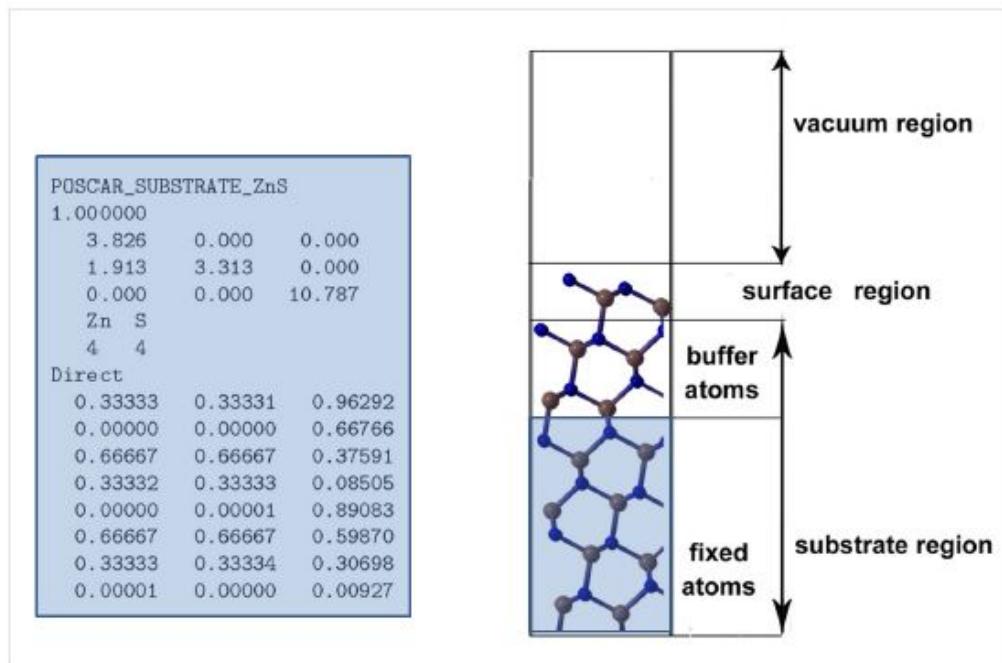


图 11：USPEX 使用的表面模型。注意 POSCAR_SUBSTRATE 应该准确代表没有真空的晶体的几何信息。如果输入中有大量的真空区域，程序将自动删除它并提供一个名为 POSCAR_SUBSTRATE NEW 的新文件，这个文件必须在计算中用到（重命名为 POSCAR_SUBSTRATE）。

➤ 变量 thicknessS

含义：表面区域的厚度。只有这个区域允许吸附原子。

默认：2.0 Å

格式：

3.5 : thicknessS

➤ 变量 thicknessB

含义：基底的缓冲区域的厚度。这个区域是 POSCAR_SUBSTRATE 的一部分，可以松弛。

默认：3.0 Å

格式:

3.0: thicknessB

➤ 变量 reconstruct

含义: 表面单元的最大增值, 允许复杂重建。

默认: 1

格式:

1: reconstruct

USPEX 也提供研究表面原子变成分搜索的可能 (在 USPEX 10.0 发布)。在这种情况下, 稳定的表面重建是由化学势【37】控制的。典型的输入设置如下:

* TYPE OF RUN AND SYSTEM *

USPEX : calculationMethod (USPEX, VCNEB, META)

201 : calculationType (dimension: 0-3; molecule: 0/1; varcomp: 0/1)

% atomType

Si O

% EndAtomType

% numSpecies

2 4

% EndNumSpecies

这里我们列出 1×1 晶胞表面原子的最大数。

* SURFACES *

% symmetries

2-17

% endSymmetries

% StoichiometryStart

1 2

% StoichiometryEnd

这定义晶体的化学计量数。

-23.7563 : E_AB (A_mB_n 化合物的 DFT 能量, eV/化学式)

-5.4254 : Mu_A (元素 A 的 DFT 能量, eV/原子)

-4.9300 : Mu_B (DFT energy of elemental B, in eV/atom)

3.5 : thicknessS (thickness of surface region, 2 Å by default)

3.0 : thicknessB (thickness of buffer region in substrate, 3 Å by default)

4 : reconstruct (maximum multiplications of cell)

现在, USPEX 支持如下的表面变成分计算:

- 基本的表面重建 (如 C 在 diamond(100) 表面)
- 二元化合物的表面重建 (如 GaN(0001) 表面)
- 包含不同物种的基本表面的重建 (如 PdO 在 Pd(100) 表面)

5.3 变成分代码

切换为变成分模式, 必须:

1. 设置 301、311 或 201: 计算类型
2. 在 numSpecies 下设置结构组成 (参照 numSpecies 的说明书)
3. 用 Latticevalues 设置每种类型原子 (或每个组成块) 粗略的原子体积
4. 设置下面的 varcomp-only 选项:

➤ 变量 firstGeneMax

含义: 第一代抽取多少不同的组分。如果是 0, 这个数值就等于 intialPopSize/4。对于二元的, 要求 firstGeneMax=11, 对于三元, 需要更高的值, 比如 30。

默认: 11

格式:

10 : firstGeneMax

➤ 变量 minAt

含义: 第一代单胞的最小原子数 (对于 calculationType=301/201/300) 或分子数 (对于 calculationType=311).

默认:无默认值

格式:

10: minAt

➤ 变量 maxAt

含义: 第一代单胞的最大原子数 (对于 calculationType=301/201/300 或 META 计算) 或分子数 (对于 calculationType=311)。

默认: 无默认值

格式:

20: maxAt

➤ 变量 fracTrans

含义: 通过原子替换获得的结构所占的百分比。这个过程中, 一个随机选取的原子被系统中其他化学成分的原子取代--新的化学组成由默认随机选取, 或者你可以在 specificTrans 下定义它, 就像明确的置换。

默认: 0.1

格式:

0.1: fracTrans

➤ 变量 howManyTrans

含义: 结构中被替换原子的最大百分比(0.1=10%)。要替换的原子的百分比从 0 到 howManyTrans 随机选取。

默认值: 0.2

格式:

0.2: howManyTrans

➤ 变量 specificTrans

含义：允许替换的种类

默认：虚线（没有特定的替换）

格式：

```
% specificTrans
```

```
1 2
```

```
% EndTransSpecific
```

注意：这个例子中，1类原子可换成第2类原子，反之亦然。如果你想尝试所有可能的置换，只需在这个参数下保留虚线。

变成分案例在运行时，参数 keepBestHM 有新含义--凸包（比如，多成分系统的热力学稳定状态）图上的所有结构保留，连同 convex hull 附近的一些亚稳结构--总量是 keepBestHM.

对于变成分的运行，明智地设置第一代尤其重要。选择适当大的初始代的大小 initialPopSize. 为不同组成的 firstGeneMax 选取合理大的数值作为第一代的样品。最后，minAt 和 maxAt 的差异不超过 2 倍，而且不同的体系大小你可能需要少量的计算：比如，4~8、8~16、16~30 原子，等等。

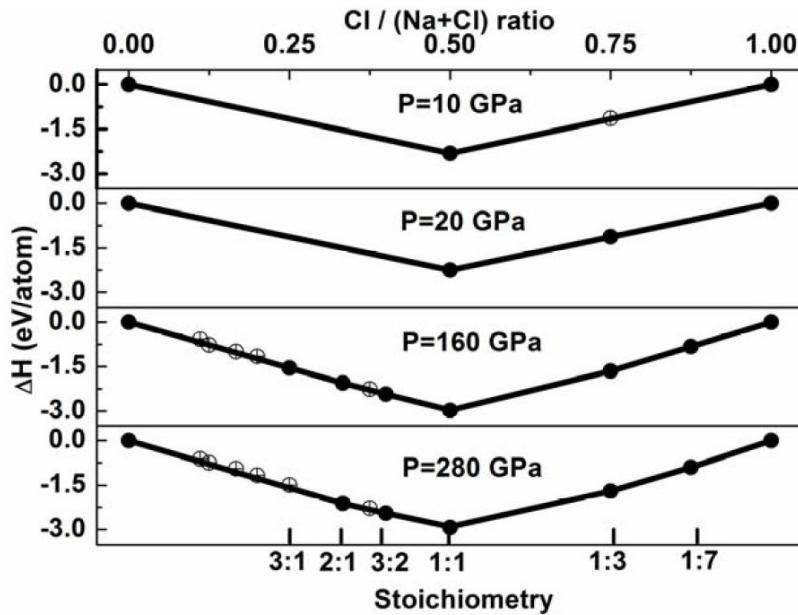


图 12: Na-Cl 系统在选定压力下的 convex hull 图。实心圆表示稳定结构, 空心圆表示亚稳结构。

给 VASP 用户一个附加意见--如果你想完成一个变成分的运行, 拿 Na-Cl 体系来说, 你需要确定 INPUT.txt 中给出了正确的原子类型, 并且将赝势文件 POTCAR_Na 和 POTCAR_Cl 放入文件夹 `~/StructurePrediction/Specific`。USPEX 将识别每个原子并且适当选取每个原子的 POTCAR 文件进行计算。图 12 展示了 USPEX 发现的且被实验证实【38】的热力学稳定的氯化钠。

5.4 进化准动力学命令

这对于找全局最小值是一个非常强大的方法, 以及开始结构中潜在的动力学接近的低能量亚稳结构。开始的结构必须是高质量的且在 POSCAR_1 中给出。进化准动力学只能在 VASP 和 GULP 中使用。

为了运行进化准动力学程序, 必须:

1. 设置

META: 计算方法

300: 计算类型

2. 在你的文件夹的 VASP5 版本中建立 POSCAR_1 文件（进化准动力学需要一个好的初始结构，在感兴趣的压力下弛豫）

3. 设定数量大小（这个例子中，这是每一步软模变异的数量）

30: populationSize

4. 设定压力

➤ 变量 ExternalPressure

含义：你想在哪个压力下执行计算，GPa。

默认：无默认值

格式：

10: ExternalPressure (GPa)

5. 设置准动力学选项：

➤ 变量 GaussianWidth

含义：每个加入到能量面以加速相变的高斯宽度。好的经验是从 0.1-0.15L 选一个值，这里 L 是单胞的最小长度，单位为 Å。

默认：0.15L (Å)

格式：

0.80: GaussianWidth

➤ 变量 GaussianHeight

含义：每个加入到能量面以加速相变的高斯高度。好的经验是选一个接近 $L(\delta h)2G$, L 是单胞的平均长度, 单位为Å, δh 是高斯宽度, G 是剪切模量, 单位为千巴。

默认: $1000 * (0.15L)^2 * L = 22.5L^3(\text{\AA}^3\text{kbar})$

格式:

2000 : GaussianHeight

➤ 变量 FullRelax

含义：这样的准动力学仅在固定晶胞中弛豫结构。为了分析，我们需要进行完整的结构弛豫（即弛豫晶胞）。

当 $\text{FullRelax}=0$, 没有完全弛豫要执行（非常快速的选择, 但对结果分析造成不便）

当 $\text{FullRelax}=1$, 仅一代中的最好结构要完全弛豫也很快, 有时充足)

当 $\text{FullRelax}=2$, 所有的等价结构完全弛豫（依旧很快, 只比=1时慢2倍, 但提供了更多的内涵信息。大多数案例都需要它）。

默认: 2

格式:

2: FullRelax

为了完全弛豫, 当执行进化准动力学计算时, `abinitioCode` 的格式略有不同, 比如:

`abinitioCode` (哪个指令可用于计算? 到目前为止: `vasp`、`siesta` 和 `gulp`)

3 3 3 (3 3)

`ENDabinit`

在上面的例子中, 一个固定晶胞的弛豫有4个阶段, 完全弛豫有2个(括号中)。记住, 在最后的固定晶胞的弛豫阶段, 压力张量必须精确--这才能驱动准动力学过程。

➤ 变量 maxVectorLength

含义：与 minVectorLength 一样，是基础晶胞在进化准动力学模拟时的长度边界值（注意在一般计算中 minVectorLength 含义不同，maxVectorLength 仅在进化准动力学中使用）。当任何基础晶胞的长度比 minVectorLength 或比 maxVectorLength 大，我们会在准动力学中添加一个夸张的修正力，这驱使晶胞进化朝好的方向发展。当所有基础晶胞的长度在正常范围时，这个修正力为零。

默认：无默认值

格式：

12.0: minVectorLength

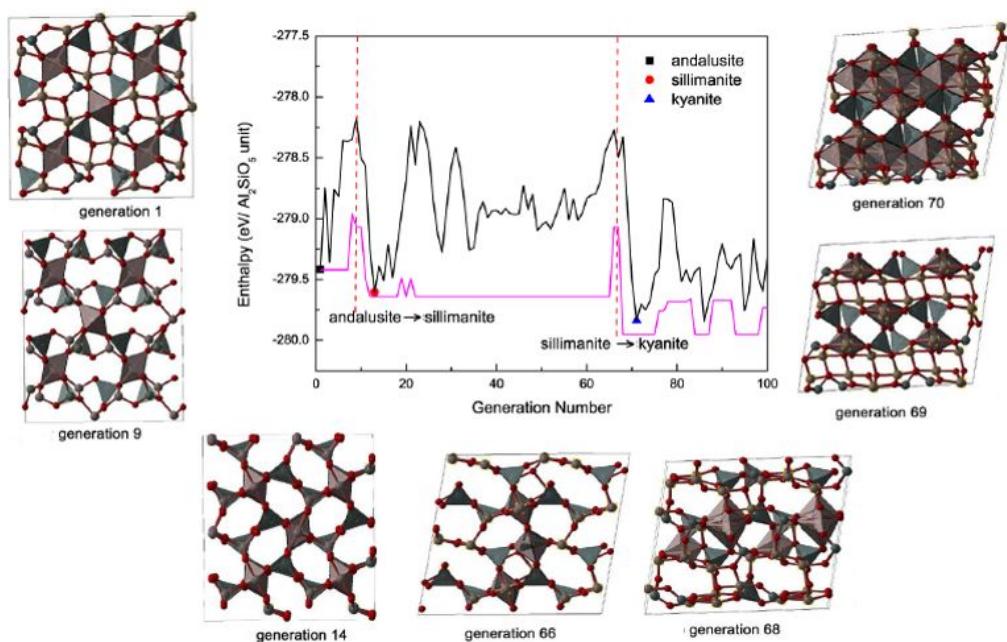


图 13: 10GPa 时压缩红柱石 (Al_2SiO_5) 过程中的焓演变 (黑线: 常量 h 下最好结构的焓; 红线: 最好的结构完全弛豫后的焓)。得到结构的顺序为: 第 1 代 (红柱石) -> 第 9 代 (硅线石) -> 第 14 代-> 第 66 代-> 第 68 代-> 第 69 代-> 第 70 代 (蓝晶石)。

当运行准动力学时，附加文件可在 results1 文件夹中找到，更重要的是：

Force.dat --分析晶胞上的力，内部的(f_c)和来自高斯的 (f_g)；

Presten --压力张量

Lattice.dat--模拟过程中晶胞形状的改变

Enthalpies and enthalpies_relaxed --固定晶胞每一步结构的焓和完全弛豫后的焓；

GatheredPOSCARS and gatheredPOSCAR_relaxed--固定晶胞的结构和完全弛豫后的结构。

图 13 展示了进化准动力学应用的例子：从一个 Al_2SiO_5 同质多形体（红柱石）开始，我们得到另外两个已知的同质多形体（蓝晶石和硅线石）和重要的相变机理。

5.5 粒子群算法 (PSO) 指令

在晶体和簇结构预测领域，一些算法对于小体系非常成功。粒子群算法，在这个领域由 Boldyrev 【39】所倡导，是进化算法的特殊情形，大量的（集群）候选值（称为“微粒”）根据一些简单的公式移动到搜索空间。微粒的运动由它们在搜索空间中自己最显著的位置决定，和全体集群最显著的位置一样。首先，微粒的坐标 x 和速度 v 随机产生。随后每一步，位置和速度都根据以下公式更新：

$$\begin{aligned} v'_i &= \omega \cdot v_i + \varphi_p \cdot r_p \cdot (p_i - \chi_i) + \varphi_g \cdot r_g \cdot (g - x_i) \\ \chi'_i &= \chi_i + v'_i \end{aligned} \quad (7)$$

这里 ω , φ_p 和 φ_g 是加权因子，控制 PSO 算法的行为和效率； r_p 和 r_g 是 $[0, 1]$ 范围内的随机数，分别在每一步为每一个微粒产生； p_i 是微粒 i 最显著的位置， g 是整个集群最显著的位置。

这种算法，尽管它简单但可以工作 【39】。改进之前安装运行 【39, 40】的关键：（1）搜索空间的度量（单独在坐标系中绘制各个晶体结构不重要）；（2）用 PSO 进化结构的方法，也就是变异操作。

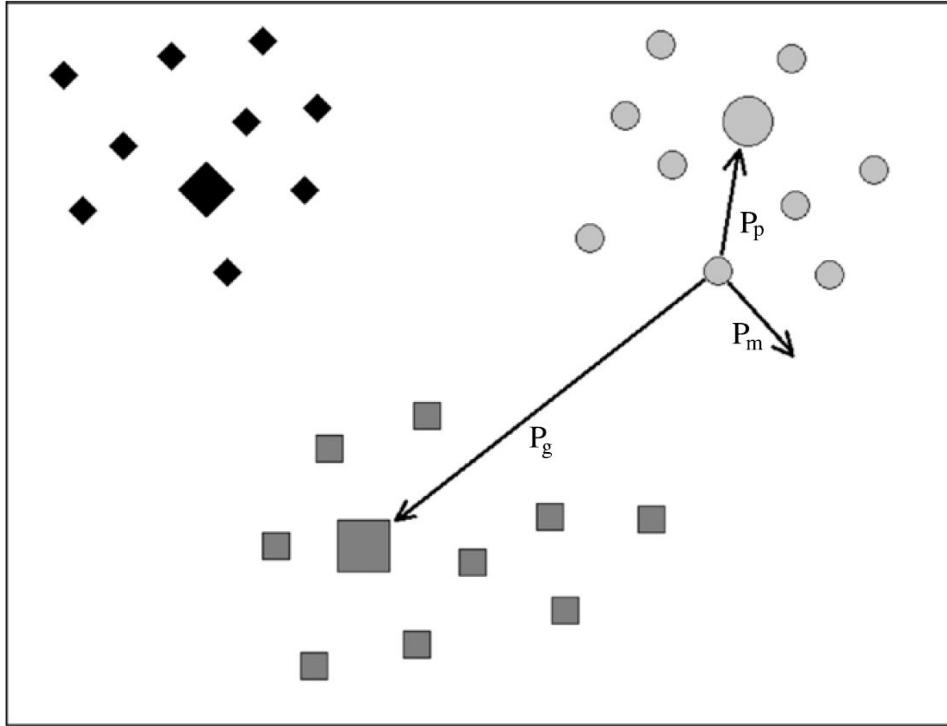


图 14: 三个个体数量 10 代之后的的 PSO-USPEX 混合算法图例（用菱形、正方形和圆形标记）。每个微粒的最好位置都用放大的符号标记。最好的结构是大的正方形。圆形标记的结构可突变，用其历史中最好的位置（大圆）产生子代或通过遗传操作以 P_m , P_p 和 P_g 的概率产生全部种群（大正方形）的最好位置。

通过直接由原子坐标和两个结构的参数决定的速度 v_i 来进化微粒是不可行的。我们的问题是用指纹距离作为能谱最原始的坐标，且 USPEX 用来进化 PSO 微粒的变量作为进化种群结构最有效公正的方法。即，微粒要么突变（模拟自由移动），要么是以最佳位置参与遗传或最佳的种群位置（模拟 PSO 朝这些方向移动）参与遗传。代替了在所有的移动的每一步都加上一些权重，我们通过以下描述的可能性一次性加上：

$$\begin{aligned}
 P_m &= \frac{\omega}{\Sigma}; & P_p &= \frac{\varphi_p \cdot r_p \cdot D_p}{\Sigma}; & P_g &= \frac{\varphi_g \cdot r_g \cdot D_g}{\Sigma}; \\
 \Sigma &= \omega + \varphi_p \cdot r_p \cdot D_p + \varphi_g \cdot r_g \cdot D_g,
 \end{aligned} \tag{8}$$

其中 D_p 是一个微粒现在和最佳位置间的指纹间距, D_g 是微粒现在位置和全部种群最佳位置的指纹间距。我们的测试在几个不同的系统进行, 表明这种方法(我们称之为“cor-PSO”, 即校正 PSO)是比较成功的。而且比 PSO 之前的版本好, 但在成功率和效率方面仍不能和 USPEX 算法【3,41】相比。

下面的变量对 calculationMethod=PSO 是唯一的:

➤ 变量 PSO_softMut

含义: 软模变异的权重(公式 8 中的 ω)

默认: 1

格式:

1: PSO_softMut

➤ 变量 PSO_BestStruc

含义: 用同样 PSO 微粒的最好位置遗传的权重

默认: 1

格式:

1: PSO_BestStruc

➤ 变量 PSO_BestEver

含义: 由全局最好 PSO 遗传的微粒的权重

默认: 1

格式:

1 : PSO_BestEver

6 相变路径预测

相变决定材料许多方面的行为。因此，很有必要揭示结构相变可能存在的机理。

6.1 变胞微动弹性带（VCNEB）法

预测相变机制可以认为是一个两端的问题，算法必须定位在中间状态。变胞微动弹性带（NEB）【42,43,44】方法是解决两端问题广泛运用的技术，一个有效而稳健的方法来寻找反应路径和鞍点，在最小能量路径（MEP）上的两个端点之间的势能面。该 NEB 方法已成功应用于分子化学反应和缺陷迁移，尤其它可以提供相变过程给定的初始和最终状态间的能量势垒。不幸的是，大多数用 NEB 方法解决的问题都是在单胞的限定下，这排除了它用于相变的可能（这涉及沿相变路径变异的单胞）。

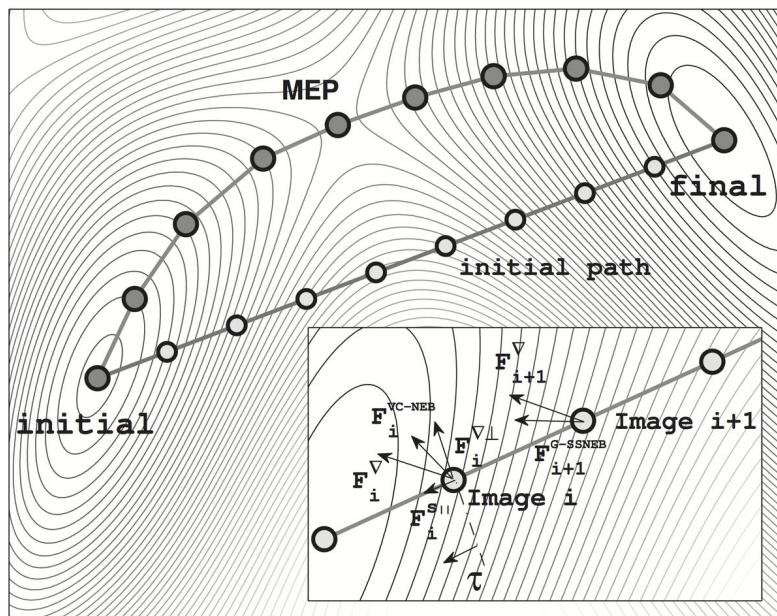


图 15：2D 焰表面模型的最小能量路径（有灰色圆的线）和初始路径。影像 i 用 VC-NEB 方法的力示于插图。 F_i^V 是梯度方向的潜在力。 $F_i^{V\perp}$ 和 $F_i^{s||}$ 分别是 $F_i^{V\perp}$ 的横向分量和弹性力。

变胞 NEB (VC-NEB)方法【15】，我们用不同的公式来完善，平等对待单胞和原子坐标，在恒定压力下在膨胀的结构空间运行。以第一性原理为框架的 VC-NEB 方法已经被添加到 USPEX 代码中。VC-NEB 方法是探索更大结构空间相变过程两个端点激活路径的常用工具。VCNEB 方法中路径上的所有结构都被示为“Image”。

6.2 VCNEB 的 Input 选项

VCNEB 方法目前只能用 VASP, GULP 和 Quantum Espresso 指令。

为了切换 VCNEB 模式，必须：

1. 设定

VCNEB: calculationMethod

2. 在 VASP4 版本你的文件夹中创建 Images 文件（VCNEB 需要至少两个结构，初始和最终相，来进行相变路径预测）。

3. 设置以下 VCNEB 选项：

➤ 变量 vcnebType

含义：说明 VCNEB 计算的类型。这个变量包含 3 个目录：calculation option, Image number variability, 和 constant variability:

● Calculation option:

“1”--VCNEB 方法

“2”--没有 VCNEB 计算的结构弛豫模型

● Variable-Image-Number method:

“0”--VCNEB 的 Images 数量固定

“1”--VCNEB 的 Images 数量可变

- Variability of spring constant:

“0”--固定弹性常数

“1”--弹性常数可变

默认: 110

格式:

111: vcnebType

注意: 如果 vcnebType=111, 即, VCNEB 计算的 Images 值可变, 弹性常数可变。我们强烈建议用户用 VCNEB 方法研究重建相变时使用可变的 Images 值。

➤ 变量 numImages

含义: 执行计算初始的 Images 值

默认: 9

格式:

13: numImages

➤ 变量 numSteps

含义: 执行 VCNEB 计算的最大步数

默认: 200

格式:

500: numSteps

注意: (1) 当 numSteps=-1, 初始路径仅在不执行能量计算时产生。 (2) VCNEB 路径通常收敛很慢。我们要求 numSteps 最少设为 500.

➤ 变量 optReadImages

含义： Images 文件的参考选项：

- Images 文件参考选项：

“0”--所有 images 都需要尤其是 Images 文件

“1”--只需要初始和最终的 images 且可以在 Images 中看的

“2”--初始、最终和任何特定的中间 Images 都能在 Images 文件中看的

默认： 2

格式：

1: optReadImages

注意：所有选项中，初始和最终的 images 必须指定。自动线性插值法被应用于在选项 1 和 2 中产生初始 Images

➤ 变量 optimizerType

含义：结构弛豫的优化算法选项

- 优化算法；

“1”--急速下降（SD）

“2”--快速惯性弛豫算法（FIRE，Fast Inertial Relaxation Engine）算法 **【45】**

默认： 1（SD）对于 VCNEB 计算； 2（FIRE）对于结构弛豫

格式：

1: optimizerType

➤ 变量 optRelaxType

含义：结构弛豫类型

- 结构弛豫类型：

“1”--仅弛豫原子位置（晶胞固定），跟传统 NEB 方法一样

“2”--仅弛豫晶格（仅用于测试）

“3”--原子位置和晶格完全弛豫

默认：3

格式：

3: optRelaxType

➤ 变量 dt

含义：结构弛豫的时间步数

默认：0.05

格式：

0.1: dt

注意：如果 dt 很小，计算就会很慢。如果 dt 很大，计算将不稳定且产生没有意义的路径。

➤ 变量 ConvThreshold

含义：images 中均方根力（RMS, Root Mean Square forces）弛豫的标准条件

默认：0.003eV/Å

格式：

0.005: ConvThreshold

➤ 变量 VarPathLength

含义：变 Images 方法的标准路径长度。当两个相邻 images 的长度比 VarPathLength 的 1.5 倍还大，新的 image 将用线性插入法加入两个 images 间；当小于 0.5 倍，第二个 image 将会移动。

默认：初始 Images 间的平均路径长度

格式：

0.3: VarPathLength

➤ 变量 K_min

含义：最小弹性常数，仅用于变弹性常数 VCNEB，eV/Å²

默认：5

格式：

3: K_min(eV/Å²)

➤ 变量 K_max

含义：最大弹性常数，仅用于变弹性常数 VCNEB，eV/Å²

默认：5

格式：

6: K_max(eV/Å²)

➤ 变量 Kconstant

含义：弹性常数，仅用于固定弹性常数 VCNEB，eV/Å²

默认: 5

格式:

4: Kconstant(eV/Å²)

➤ 变量 optFreezing

含义: 冻结 Images 结构的选项。当 ConvThreshold 被激活时 Image 结构将被冷冻。

● Images 结构冻结选项:

“0”--任何时候都不冷冻 Images

“1”--当 ConvThreshold 被激活时冷冻 Image

默认: 0

格式:

1: optFreezing

➤ 变量 optMethodCIDI

含义: Climbing-Image (CI) 和 Downing-Image (DI) 方法的选项。这个方法只建议在你有一个合理且收敛的路径时使用。

● CI/DI-Image 方法选项:

“0” --CI/DI 方法不用

“1” --只用 CI 方法, 只有最高能量或用户提供相变 (TS) 才用 CI

“2” -混合多个-CI/DI 方法, 需提供 FS 和 LM 的连续数值

默认: 0

格式:

1: optMethodCIDI

➤ 变量 startCIDIStep

含义: CI/DI 法开始的步数, 只有 optMethodCIDI=1 时可用

默认: 100

格式:

200: startCIDIStep

➤ 变量 pickupImages

含义: Images ID 为 CI/DI-Image 法挑选

默认: 过渡态 Image ID 和局域最小态 Images

格式:

% pickupImages

9 11 17

% EndPickupImages

注意: 在这个例子中, 第 9、11、17 个 Images 将被挑选应用 CI/DI-Image 法。过渡态的 Image 将应用 CI-Image 法, 局域最小态 Image 将自动应用 DI-Image 法。

➤ 变量 FormatType

含义: 路径 output 文件中结构的格式, 位于 results1/PATH/.

● 路径结构 output 格式:

“1” --XCRYSDEN 格式 (.xsf 文件)

“2” --VASP POSCAR

“3”--晶格的 XYZ 格式

默认： 2

格式：

1: FormatType

➤ 变量 PrintStep

含义：保存在经过时间间隔后的每一步后位于 result1/STEP 的 VNCEB 重新开始的文件。

默认： 1

格式：

10: PrintStep

注意：对于经验指令，如 GULP，我们建议用户设 PrintStep=10 以减少保存新文件耗费的时间。

图 16 展示了用 VCNEB 方法的例子：相变机制和能量势垒在 BH 体系在 160GPa 发生 $Ibam \longrightarrow P6/mmm$ 的转变开始，我们得到了 $Pbcm$ 中间相。

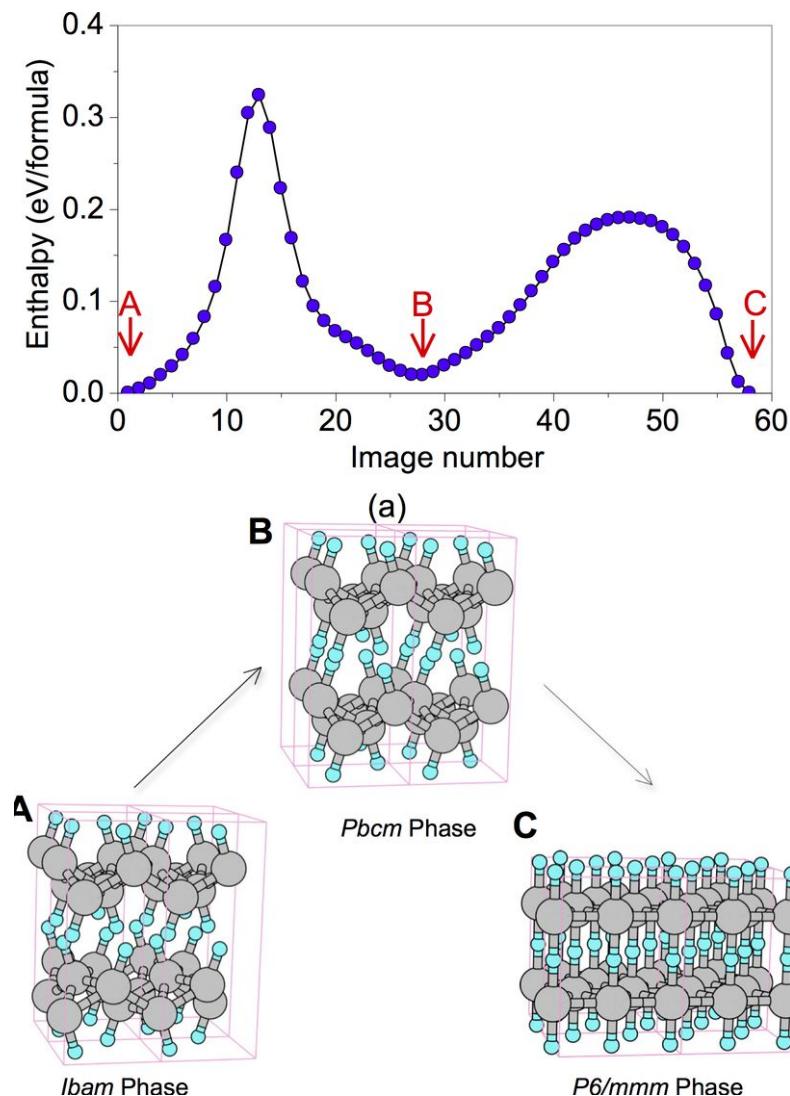


图 16：BH 在 160GPa 时的 $Ibam \rightarrow P6/mmm$ 的转变。出现 $Pbcm$ 中间相。 $Ibam \longrightarrow Pbcm$ 和 $Pbcm \longrightarrow P6/mmm$ 片段上的鞍点的势垒分别为 0.32、0.19 eV/f.u.

6.3 如何设置 VC-NEB 计算的初始路径

VC-NEB 方法是寻找相变路径非常有效的方法，但我们必须小心地准备初始路径。在 VC-NEB 计算中，晶胞旋转发生在初始和最后的结构附近，这里初始和最终的图像包含许多相

同结构。为了排除这些无用的旋转，改进的 Variable-Image-Number 方法将自动出现晶胞旋转，这会节约很多时间。

作为一种选择，当产生初始 image 设置时，可以在应用 VC-NEB 方法前采用拒绝循环技术。一般的 3×3 矩阵，欧拉角 $R(\phi, \theta, \psi)$ 和晶格反射 $M(x, y, z)$ 都被定义。执行 VC-NEB 计算前，全部空间的欧拉角和反射操作数值搜索都用来寻找最小晶胞转变距离 Δh ：

$$\Delta h = |h_{initial} - R(\phi, \theta, \psi)M(x, y, z)h_{final}| \quad (9)$$

最后的 \tilde{h}_{final} 拒绝循环晶格适量被指定为终点的 image:

$$\tilde{h}_{final} = R(\phi, \theta, \psi)M(x, y, z)h_{final} \quad (10)$$

更重要的是，我们要预防任意分配初始和最终的 images 的原子部分坐标 r_v （修改初始和最终结构的绘制原子）。否则，计算将会很困难或者一些相同的路径会出现在一个计算中，如图 17 所示。对于更复杂的系统，如果你没有好的初始路径，将会得到一些不合理的或散乱的路径。在两个终点 images 上的原子最小距离的全局搜索帮助 VC-NEB 方法再分配原子序列。在 VC-NEB 计算前自动创造模型路径的能力对相转变的稳定和收敛非常重要，而且是研究大的复杂系统的先决条件。

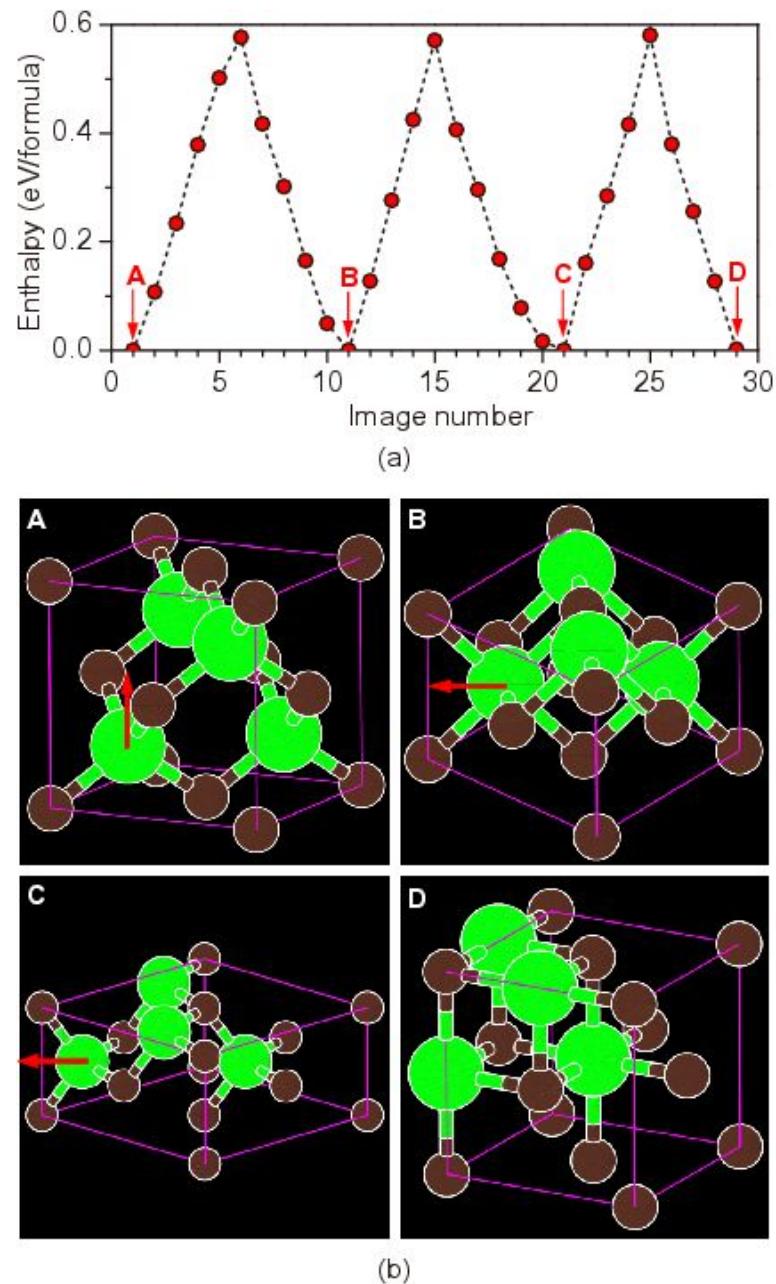


图 17：当设置不好的初始 Image 文件会发现相同的路径。路径是 GaN 在 45 GPa 静压下 B->B1 的相变。在 11 和 21 Images，单斜晶胞中的 B1 和 B3 结构在 MEP 搜索时被发现。Ga 原子在相变过程中沿箭头方向移动。

7 在线工具

我们开发许多有用的在线工具，它们被用于准备 USPEX 的输入和后处理。查询这些工具请链接：

<http://han.ess.sunysb.edu>

下面你可以找到有关每个在线工具的信息。

7.1 结构表征

在这里我们有 4 个工具：

- 指纹指令——计算和绘制指纹功能。这是一种晶体结构描述符。这一描述符是一种与对相关函数和衍射图案相关的一维函数。它不依赖于绝对的原子坐标，而只依赖于原子间的距离。原子位置的小偏差会轻微地影响指纹，即它们是数值稳健的。
- 多触点指令——计算平均准熵，以及关于一系列结构的一阶和 S 阶。它通过余弦距离差 ≥ 0.003 这一条件过滤出独特的结构，识别出这些结构的对称性，最后在 `uniq gatheredPOSCARS` 文件中将它们一一列举出来。
- POSCAR2CIF——确定了空间群和从 POSCAR 文件中准备一个 CIF 文件。
- CIF2POSCAR——从 CIF 文件中准备一个 POSCAR 文件。

7.2 性质计算

这里我们有 2 个工具：

- 硬度——这一工具基于 Lyakhov-Oganov 模型计算硬度。
- EELS — 这一工具用于计算电子能量损失谱，EELS 是 Priya Johari 编写的。

7.3 分子晶体

在这里我们有 2 个工具：

- MOL 预检验——该工具允许你在运行含有 calculationType=310/311/110 的 USPEX 计算器之前查阅 MOL_1 文件。
- Z 矩阵——该工具可将 XYZ 文件转换为 USPEX MOL 1 文件。

7.4 表面

基底——一个从 POSCAR / CIF 文件和指定的密勒指数，层厚度和转变中准备基底的程序。由此产生的 POSCAR 文件可用于 calculationType = 200/201 作为基底的表面计算。

7.5 其它

这里我们有 2 个功能和 1 个链接：

- 输入文件生成器——USPEXINPUT.txt 生成器。该实用程序可以帮助初学者建立正确的 USPEX 输入文件。
- 体积估计——对于 USPEX（或对于 INPUT.txt 文件）该功能可以估计非分子晶体与分子晶体的体积。
- USPEX 手册——一本手册的在线版本。

8 常见问题

8.1 如何将结果可视化?

USPEX 会产生大量数据（晶体结构和能量等等）。通过手动分析这些数据将会非常繁琐和耗时。通过采用高效可视化，分析可以更迅速地进行，可以产生有价值的新见解。Mario Valle 使用他的 STM4 可视化工具包【46】已经制定了具体的读取功能和可视化 USPEX 输出文件，在这里我们强烈建议使用与 USPEX 相匹配的工具包。使用 STM4，需要在计算机上安装 AVS / Express。AVS / Express 不是公开发行的，所以你需要一个许可证。或者，你可以使用其他软件，例如 OpenDX 或 VESTA 进行 USPEX 结果的可视化。VESTA 已经包含了直接阅读 USPEX 结果文件的功能。图 5 显示了由 STM4 产生的典型图谱，此外 STM4 还可以制作影像，用以分析电荷密度图，你可在 <http://mariovalle.name/STM4> 里找寻应用。

关于具体怎么后续处理或数据分析。Mario Valle 在他的 STM4 的代码(图 5)中添加了 USPEX 界面，因而 USPEX 在这方面也具有独特优势。包括在几分钟内分析成千上万的结构，结构性能的相关性测定，算法性能分析，能量面的量化，最先进的结构可视化，空间群的确定等等，甚至可以展示模拟过程的动画！对于结构的可视化，我们也建议用 VESTA，它直接导入 USPEX 的结构文件。USPEX 本身也会产生一些图片文件，用来帮助查看和分析计算结果：

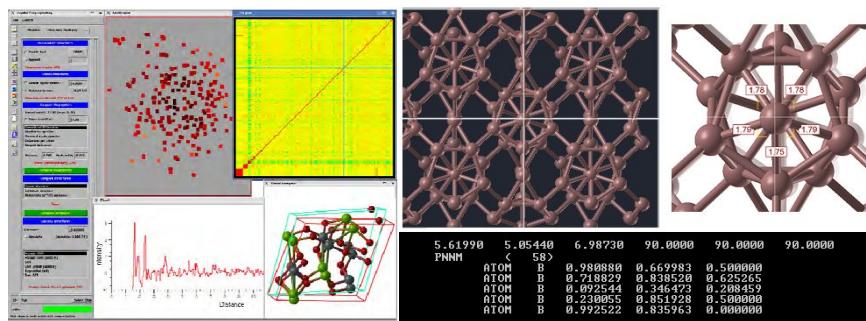


图 5：STM4 的 USPEX 界面

8.2 如何避免陷阱?

第一, USPEX 使用一个足够大的种群规模。第二, USPEX 通过默认采用一种强大的指纹识别方法。任何增加种群多样性的因素会减少受困于局部最小值的几率。为了确保模拟不陷入困境, 运行一个带有不同参数的二次模拟是有必要的。反种子技术是避免陷入困境的一个强大的方式。

8.3 什么是单块计算?

在 USPEX 9.3.9 中对单块特征进行了介绍, 单块特征使 USPEX 用户能够利用相同混合物可变数目的计算单位进行结构预测。例如:

```
% atomType  
Si O  
% EndAtomType  
% numSpecies  
1 2  
% EndNumSpecies  
12: minAt  
  
24: maxAt
```

这意味着我们取样的带有可变数目计算单位的二氧化硅化合物(按 1:2 的比例)具有 12–24 个原子。

从 USPEX 9.4.1 开始, 单块功能已经从 calculationType = 301/311 变为 300/310。仍然是相同的设置, 用户只需要设置 minAt、maxAt 和 numSpecies 关键词。

8.4 如何使用种子技术?

这种技术是有用的, 如果不是从随机结构开始, 你会更愿意输入一些你已经知道的化合物或相关材料的结构。

在 VASP5 格式中的级联 POSCAR 文件的格式如下，为下一代结构在 Seeds 中创建 POSCARS 文件，或为一次 USPEX 计算的特殊代创建 POSCARS gen(gen 指代号)。在此文件名中不要遗忘字母“S”。例如：

EA33 2.69006 5.50602 4.82874 55.2408 73.8275 60.7535 no SG

1.0

2.6901 0 0

2.6901 4.8041 0

1.3449 2.4021 3.9671

Mg Al O

1 2 4

Direct

0.79919 0.56784 0.85959

0.79352 0.23095 0.54475

0.79354 0.91609 0.17445

0.050972 0.81606 0.85961

0.17223 0.19481 0.8596

0.43825 0.65517 0.40688

0.43823 0.20244 0.31233

EA34 7.61073 2.85726 2.85725 60.0001 79.1809 79.1805 no SG

1.0

7.6107 0 0

0.53635 2.8065 0

0.53633 1.352 2.4593

Mg Al O

1 2 4

Direct

0.70891 0.50744 0.068339

0.37405 0.28573 0.84663

0.023663 0.069185 0.63009

0.88956 0.78056 0.34146

```
0.35047  0.62692  0.18782  
0.59729  0.21131  0.77221  
0.11644  0.37159  0.9325
```

在计算中的任何时候都可以加入种子。每一代开始时 USPEX 将尝试读取种子文件的两个类型。相应的信息将被记录在 results1/Seeds_history 中，种子文件 (POSCARS 或 POSCAR_gen) 将在种子中的 POSCARS_gen 文件内保存。

当种子加入时，我们建议用户检查 the results1/Seeds history 和警告文件。如果你的种子有问题时，可能会有一个警告信息 “Meet a problem when reading Seeds-…”。当一个错误出现在种子文件，比如是缺失了一行，错误出现后的结构将不会被添加进去。

注意：确保你指定的所有原子符号在每一结构的 6 号线位置。例如，添加 P63 / MC H2 结构到 H-O 变成分计算，你要编辑的文件：

```
H_I-P63/mc  
1  
4.754726 -2.74514  0.000000  
-0.00000  5.490285  0.000000  
0.000000  0.000000  4.508715  
H O  
16 0  
Direct ...  
(原子位置信息在这就省略了)
```

即使在空间群为 $P6_3/mc$ 的 H_2 原子结构中没有氧原子，你写下氧的名字标签也是很重要的。否则 USPEX 将会跳过此结构。

8.5 如何巧妙考虑组成？

对于可变组分和单块计算，当计算开始时，它便会生成一个带有所有可能组成的种子文件 compositions，其中遗传代码从随机结构发生器中随机选取组成。例如，C-O 体系

8 0

0 8

2 4

意味着你会对随机产生 C₈、O₈ 和 C₂O₄ 结构怀有兴趣。由于遗传变异和原子置换，其他的成分也将被采样。当想要产生含有指定组成的结构时，可以使用 `anticompositions` 的特点——将所有不想要的组成输入至 Seeds/Anti-compositions 文件中。主要有 3 种可行方式：

1. 对于所有同一比例不需要的成分，可以写下某一成分比以禁止某些成分。例如，可以使用“1 2 1”禁止所有的以相同的比例组成的成分，如“1 2 1”、“2 4 2”和“3 6 3”等。
2. 只为特定的成分，而不是其他具有同一成分比的组成。可以用一个负号标识组成。例如，可以使用“-3 2 0”或“3 -2 0”禁止“3 2 0”的成分，但不禁止“6 4 0”或“9 6 0”组成。（注意：“3 2 -0”这种情况下无意义）。
3. 对于所有的单一/二元/三元化合物，如果你不想对所有的单一/二元/三元化合物进行采样，请在 Anti-compositions 文件中写下关键词 `single/binary/ternary`。

例如：

```
single  
binary  
1 1 2  
-2 2 1
```

如果你不太确定你正在做什么，就让 Anti-compositions 文件空着吧。为了得到你不想要的组分的信息，你可以查看下 `results1/compositionStatistic`。

注意：

- 即使在计算开始之前 `compositions` 或 `Anti-compositions` 文件已存在，他们会被忽略。`Anti-compositions` 文件会被重命名为 `Anti-compositions-back` 备份文件。因此，在计算开始之后请编辑 `compositions` 或 `Anti-compositions` 文件。
- 也请注意，在带有组合块的 USPEX 计算中，组合物通常意味着这些块的数量。因此，为了确保 `Anti-compositions` 文件的格式正确，请首先检查组成。

8.6 如何建立一个从本地机器到远程集群的密码连接？

你需要复制从本地机器到远程集群的公共密匙（目录./ SSH 或./ SSH2）。这里是需要执行的命令列表：

```
local # ssh-keygen -t dsa  
local # scp ~/.ssh2/id_dsa.pub oganov@palu.cscs.ch:~/ssh/tmp.pub  
remote # cd ~/ssh/  
remote # ssh-keygen -f tmp.pub -i >> authorized_keys  
remote # rm tmp.pub
```

8.7 如何从一个已坏了的 MAT-文件重新开始计算？

当在文件系统中存在一个问题，例如磁盘已满，文件系统超载，USPEX 难以准确写下这些 MAT-文件，在计算过程中便会出现如下信息：

```
???Error using ==> load  
Unable to read MAT file /home/USPEX/Current_POP.mat  
File may be corrupt.
```

当遇到一个坏了的 MAT-文件引起的错误时，请利用备份恢复损毁的 mat-文件，删除 matfilelocker 文件，然后重新启动计算。不幸地，若是当备份 mat-文件是空的或还是错误时，你必须重新在这一代中重新启动计算。

8.8 当 USPEX 在一段时间内不运行，该如何操作？

当发现 USPEX 在一段时间里不运行，而文件 still_running 仍存在了较长的时间（通常超过 30 分钟，请使用命令‘ls -l’以检查仍在运行文件的时间戳），你应该考虑到该计算存在某些错误。在这种情况下，你需要做的是遵循以下的检查程序：

- 使用命令‘top’检查确保您的 MATLAB 计算没有运行，有时，USPEX 在结构的生成和软模变异上会花费较长时间。一旦你确定 MATLAB 停止，你可以开始下一步。

- 停止 crontab 或作业运行脚本以避免在检查程序过程中 USPEX 运行。这是非常重要的，否则你会搞砸了你的 USPEX 计算。
 - 删除 still_running 文件。
 - 输入命令‘USEPX -r’ 或 ‘matlab <USPEX.m’运行 USPEX 看会发生什么。如果你遇到错误或缺陷，你可以尝试修复它们，或是向我们的谷歌 USPEX 论坛寻求帮助，或让我们知道，并报告缺陷。
- 如果一切都很好，只需要重新启动 crontab 或作业运行脚本继续计算。

8.9 如何建立一个计算使用的工作提交脚本？

为了建立一个工作提交脚本，我们希望用户了解关于 MATLAB 编程和你的作业提交系统的一些基本知识，至少要明白如何处理 MATLAB 中的字符串和如何获得工作信息的基本思路。

有两种工作脚本提交模式：本地提交或远程提交，这取决于你是将从头计算提交到你运行 USPEX 和 MATLAB 的本地机器还是远程超级计算机。

8.9.1 第一步：在提交文件夹中配置文件

案例一：本地提交

请在 INPUT.txt 文件中编辑下列标签 1

1: whichCluster (0: no-job-script, 1: local submission, 2: remote submission)

然后到 Submission 文件夹下，这里你需要编辑 submitJob local.m 和 checkStatus local.m.这两个文件。

你可以在这些文件中找到详细的说明。在一般情况下，你只需告诉 USPEX 如何提交工作和检查工作是否已经完成。

在 submitJob_local.m 里：

```

1 function jobNumber = submitJob_local()
%
2 %This routine is to check if the submitted job is complete or not
3 %One needs to do a little edit based on your own situation.
4 %l : whichCluster (default 0, 1: local submission , 2: remote submission)
%
5
6 %Step 1: to prepare the job script that is required by your supercomputer
7 fp = fopen('myrun', 'w');
8 fprintf(fp, '#!/bin/sh\n');
9 fprintf(fp, '#PBS -l nodes=1:ppn=8,walltime=1:30:00 -q cfn_short\n');
10 fprintf(fp, '#PBS -N USPEX\n');
11 fprintf(fp, '#PBS -j oe\n');
12 fprintf(fp, '#PBS -V \n');
13 fprintf(fp, 'cd ${PBS_O_WORKDIR}\n');
14 fprintf(fp, 'mpirun -np 4 vaspl > vasp.out\n');
15 fclose(fp);
%
16 %Step 2: to submit the job with a command like qsub, bsub, llssubmit , etc.
17 [a,b]=unix(['qsub myrun'])
%
18 %Step 3: to get the jobID from the screen message
19 %It will output some message on the screen like '2350873.nano.cfn.bn1.local'
20
21 end_marker = findstr(b,'.');
22 jobNumber = b(1:end_marker(1)-1);

```

在 checkStatus_local.m 里：

```

1 function doneOr = checkStatus_local(jobID)
%
2 %This routine is to check if the submitted job is complete or not
3 %One needs to do a little edit based on your own case.
4 %l : whichCluster (0: no-job-script , 1: local submission , 2: remote submission)
%
5
6 %Step1: the command to check job by ID.
7 [a,b] = unix(['qstat ' jobID ''])
%
8 %Step2: to find the keywords from the screen message to determine if the job is complete
9 %Below is just a sample:
10
11 %Job id          Name        User        Time Use S Queue
12 %-----          -----        -----      -----
13 %2455453.nano   USPEX      qzhu       02:28:42 R cfn_gen04
14 %
15 %If the job is still running, it will show as above.
16 %If there are no key words like 'Q/R Cfn_gen04' , it indicates the job is complete.
17 %Therefore, we can use a small MATLAB function findstr to apply this argument.
18 if isempty(findstr(b,'R cfn_')) & isempty(findstr(b,'Q cfn_'))
19     doneOr = 1
20     unix('rm USPEX*');    % to remove the log file
21 end

```

案例二：远程提交

请在 INPUT.txt 文件中编辑下列标签：

2: whichCluster (默认是 0, 1: 本地提交; 2: 远程提交)

C-20GPa: 远程文件夹

最后到 Submission 文件夹下，此处你需要编辑两个文件：submitJob_remote.m 和 checkStatus_remote.m

```

function jobNumber = submitJob_remote(USPEX, Index)
%
%This routine is to check if the submitted job is complete or not
%2 : whichCluster (default 0, 1: local submission; 2: remote submission)
%C=20GPa : remoteFolder
%
%
%Step1: To prepare the job script, runvasp.sh
fp = fopen('runvasp.sh', 'w');
fprintf(fp, '#!/bin/sh\n');
fprintf(fp, '#PBS -l nodes=2:ppn=2,walltime=1:30:00\n');
fprintf(fp, '#PBS -N USPEX\n');
fprintf(fp, '#PBS -j oe\n');
fprintf(fp, '#PBS -V \n');
fprintf(fp, 'cd ${PBS_O_WORKDIR}\n');
fprintf(fp, '/usr/local/pkg/openmpi-1.4.5/bin/mpirun -np 4 vaspl > vasp.out\n');
fclose(fp);
%
%Step 2: Copy the files to the remote machine
%
%Step2-1: Specify the PATH to put your calculation folder
Home = ['/nfs/user08/qinzhu']; %'pwd' of your home directory on remote machine
Address = 'qinzhu@seawulf.stonybrook.edu'; %your target server: username@address
Path = [Home '/' USPEX '/CalcFold' num2str(Index)]; %Just keep it
%
%Step2-2: Create the remote directory
% Please change the ssh/scp command if necessary!
% Sometimes you don't need the -i option
try
[a,b]=unix(['ssh -i ~/.ssh/seawulf ' Address ' mkdir ' USPEX ]);
catch
end
try
[a,b]=unix(['ssh -i ~/.ssh/seawulf ' Address ' mkdir ' Path ]);
catch
end
%
%Step2-3: Copy the necessary files (for VASP calculations, we need POSCAR, INCAR, POTCAR,
% KPOINTS and job script)
unix(['scp -i ~/.ssh/seawulf POSCAR ' Address ':' Path]);
unix(['scp -i ~/.ssh/seawulf INCAR ' Address ':' Path]);
unix(['scp -i ~/.ssh/seawulf POTCAR ' Address ':' Path]);
unix(['scp -i ~/.ssh/seawulf KPOINTS ' Address ':' Path]);
unix(['scp -i ~/.ssh/seawulf runvasp.sh ' Address ':' Path]);
%
%Step 3: to submit the job and get JobID, i.e., the exact command to submit the job.
[a,v]=unix(['ssh -i ~/.ssh/seawulf ' Address ' /usr/local/pkg/torque/bin/qsub '
Path '/runvasp.sh']);
%
% format: Job 1587349.nagling is submitted to default queue <monoc>
end_marker = findstr(v,'.');

```

```

56 if strfind(v,'error')
57 jobNumber=0;
58 else
59   jobNumber = v(1:end_marker(1)-1);
60 end

```

在 checkStatus_remote.m 里：

```

1 function doneOr = checkStatus_remote(jobID, USPEX, Folder)
2 %
3 %This routine is to check if the submitted job is complete or not
4 %One needs to do a little edit based on your own situation.
5 %
6 %Step1: Specify the PATH to put your calculation folder
7 Home = ['/nfs/user08/qiazhu']; %'pwd' of your home directory of your remote machine
8 Address = 'qiazhu@seawulf.stonybrook.edu'; %Your target: username@address.
9 Path = [Home '/' USPEX '/CalcFold' num2str(Folder)]; %just keep it
10 %Step2: Check JobID, the exact command to check job by jobID
11 [a,b]=unix(['ssh -i ~/.ssh/seawulf ' Address ' /path/to/qstat ' num2str(jobID)])
12 tempOr1 = strfind(b, 'R batch');
13 tempOr2 = strfind(b, 'Q batch');
14 if isempty(tempOr1) & isempty(tempOr2)
15     doneOr = 1;
16 % for vasp, we usually need OSZICAR for reading energy and CONTCAR for reading
17 %structure OUTCAR, EIGENVAL, DOSCAR might be needed for reading other properties.
18 % unix(['sep -i ~/.ssh/seawulf ' Address ': ' Path '/OUTCAR ./'])
19 %OUTCAR is not necessary by default
20 unix(['sep -i ~/.ssh/seawulf ' Address ': ' Path '/OSZICAR ./'])
21 %For reading enthalpy/energy
22 unix(['sep -i ~/.ssh/seawulf ' Address ': ' Path '/CONTCAR ./'])
23 %For reading structural info
24 end

```

这可能需要一些时间来正确配置这些文件。为了测试它是否可以工作，你可以输入两次“USPEX -r”指令，然后监测屏幕信息。第一次的尝试是检查工作是否提交，而第二次是检查USPEX 是否可以正确检测提交工作的状态。所有的相关信息，可以在屏幕上输出信息中找到。如果 MATLAB 输出没有任何错误，你可以准备好离开了。

8.9.2 第二步：定期运行 USPEX

实际计算以‘USPEX -r > log’命令开始。每次 MATLAB 程序都会检查从头计算的运行状态。如果工作完成，MATLAB 就会前往计算文件夹读取结果，随后提交新的计算。之后，MATLAB 将退出。因此，需要定期调用命令（例如，每 5 分钟）。周期性的脚本可以通过使用 crontab 或 shell 脚本被执行。

8.9.3 Crontab

这可以在你的 Linux 机器上执行一个 crontab 后台程序。在你的用户主目录中，现在应该有这些文件：

```

~/call_job
~/CronTab

```

下面是来自我们的集群的 CronTab 文件例子的一行:

```
* /5 * * * * sh call job
```

它表示作业提交之间的间隔为 5 分钟，并指向文件 call_job，它应该包含 USPEX 将被执行目录的地址，call_job 大概看起来是这样：

```
#!/bin/tcsh
source /etc/csh.login
source ${HOME}/.cshrc
cd /ExecutionDirectory
date >> log
matlab < USPEX.m >> log
```

为了激活 Crontab，键入

```
crontab ~/CronTab
```

如果你想终止运行，或编辑 call_job 文件或通过键入 crontab -r 删除 crontab

检查 crontab 是否工作正常，你应该在计算开始时追踪日志文件的更新

8.9.4 Shell 脚本

可以在 Linux shell 中使用睡眠命令准备脚本。下面是一个非常简单的脚本 run-uspex.sh:

```
#!/bin/sh
while true
do
    date >> log
    matlab < USPEX.m >> log
    sleep 300
done
```

注意：记住这个计算可以通过终止该脚本的进程 ID 来停止。

8.10 如何在带有对称性代码的 32 位机器上工作？

在 FunctionFolder/Tool/32bit 下执行“./install-32bit.sh”能够确保 32 位文件替换默认的 64 位二进制对称性代码。

9 附录

9.1 例子目录

- **EX01-3D_Si_vasp:** 零压力下的硅（每个单胞中含 8 个原子）。使用 VASP，PBE96 泛函进行变胞 DFT 计算。感谢 G. Kresse 和他的 PAW 文件（POTCAR）对我们的贡献。
- **EX02-3D_MgAl₂O₄_gulp:** 在 100 GPa 下的 MgAl₂O₄（每一个单胞中有 28 个原子）。使用 Buckingham 势，GULP 代码进行变胞计算。如果需要很可靠的结果，最好用从头算法。
- **EX03-3D-const_cell_MgSiO₃_gulp:** 这个例子向我们展示在知道单胞参数时如何进行结构预测。含有 Buckingham 势，GULP 代码的 MgSiO₃（每个单胞有 20 个原子）。晶胞参数与后钙钛矿一致。后钙钛矿的发现（Oganov & Ono, Nature 2004; Murakami *et al.*, Science 2004）在地球科学领域是一个重大突破。
- **EX04-3D_C_lammps:** 这个例子展示了如何使用结合了 LAMMPS 码的 USPEX 进行晶体结构预测。在这个简单的例子中：有 8 个碳原子，采用 Tersoff 势。
- **EX05-3D_Si_atk:** 这是一个使用密度泛函紧束缚近似和 ATK 代码对每一单胞含 8 个原子的硅进行晶体结构预测的例子。
- **EX06-3D_C_castep:** 使用 CASTEP 代码对每个单胞含 8 个原子的碳的晶体结构进行基于 DFT 的预测。请注意，两个不同的版本的 CASTEP 在输入文件格式上会有些许不同，我们给出了两种变体。
- **EX07-2D_Si_vasp:** 使用 DFT 和 VASP 预测硅的二维晶体。简单而有效。
- **EX08-0D_LJ_gulp:** 纳米结构预测。含有 30 个原子的 Lenard-Jones 纳米颗粒，GULP 代码。
- **EX09-3D-molecules_CH4_vasp:** 甲烷分子晶体预测，压力 20GPa，DFT，VASP。分子在文件 MOL1 中描述。
- **EX10-3D-molecules_CH4_dmacrys:** 常压，带有 8 个分子的单胞，采用力场和 DMACRYS 代码。在文件 MOL1 中描述分子，但要注意其 DMACRYS 计算略有不同的格式。请把可执行文件 dmacrys、neighcrys-PP 和 neighcrys-VV 放在特定的文件夹。

- **EX11-3D-molecules_urea_tinker:** 常压，2个分子的单胞，采用力场和TINKER代码。在文件MOL1中描述分子。
- **EX12-3D_varcomp_LJ_gulp:** 含有伪“MO”和“B”原子，GULP和变组成USPEX(Lyakhov和Oganov, 2010)的Lenard-Jones二元体系。
- **EX13-3D_special_quasirandom_structure_TiCoO:** USPEX可以很容易地找到最无序(或最有序)合金结构。这里表示为 $Ti_xCo_{(1-x)}O$ 。你需要在seed/POSCARS指定初始结构和仅使用置换操作。在这种情况下，不需要使用任何外部代码。在这个例子中，我们在无弛豫的情况下优化(最小化)结构有序度(Oganov and Valle (2009); Lyakhov, Oganov, Valle (2010))abinitioCode=0)。种子结构(Ti-Co-O结构的超晶胞)被置换找到结构的最小/最大化有序度。在这种情况下最小化有序度，可以得到“特殊准随机结构”的通用版本。
- **EX14-GeneralizedMetadynamics_Si_vasp:** 关于开始用一个简单的种子结构搜索复杂低能量结构的简单例子(Zhu *et al.*, 2013)。硅，最多为16个原子的单胞，DFT，VASP。要特别注意到INCAR文件。最重要的是，只要保留你在这里看到的文件，只改变ENCUT，SIGMA，进化准动力学不仅预测低能量的结构，同时也给出了晶体结构之间的转换机制。
- **EX15-VCNEB_Ar_gulp:** 氩气，在0GPa压力模型系统中采用变胞微动弹性带(VCNEB: Qian *et al.*, 2013)计算fcc-hcp转变。Lenard-Jones势，GULP代码。
- **EX16-USPEX-performance_SrTiO₃_gulp:** SrTiO₃(50原子/单胞)在零压力下。使用Buckingham势，GULP代码进行可变单胞计算。从这个例子中可以看到，即使是这样一个比较大的系统，USPEX代码仍有>90%的成功率和显著的效率。与此相反，通过Zurek和Lonie使用相同的体系，相同的势能报道的成功率只有7-12%。显然，USPEX优于我们通过Zurek和Lonie的实现方法。我们已经见证了我们的代码的性能在更大的系统也很优异。
- **EX17-3D_DebyeTemp_C_vasp:** 对弹性相关性能优化的例子(体积和剪切模量，泊松比，Chen-Niu硬度，德拜温度)在这个例子中，我们使用VASP代码最大限度地提高碳的德拜温度。
- **EX18-3D_varcomp_ZnOH_gulp:** 正如你所知道的，USPEX对于变成分的搜索具有独特的能力。这个例子为我们展现了一个非常具有挑战性的情况——对于三元体系Zn-O-H的变成分计算。这个计算在GULP代码中使用ReaxFF力场。USPEX可以为任何成分组成进行计算——例如，在其它范围内可实现的四元和五元体系。当然，成分越多，计算越昂贵(和越有风险)。目前没有任何参考结果。

9.2 运行测试

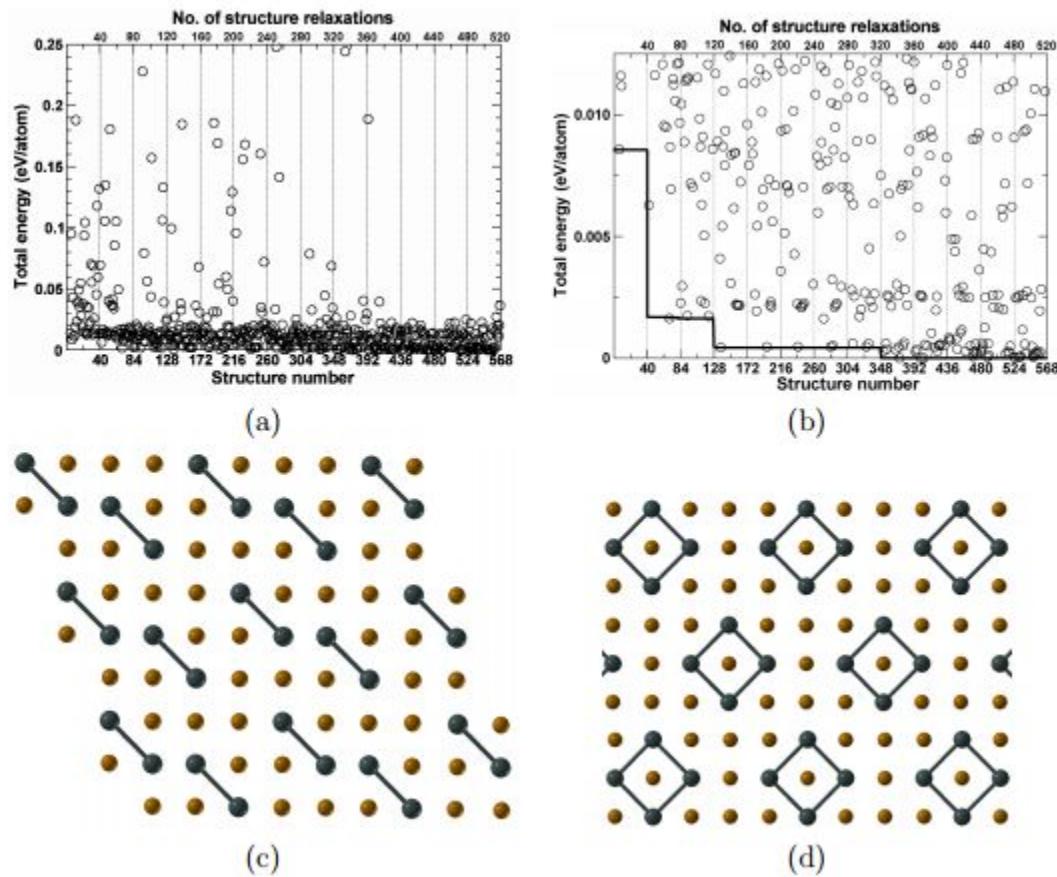


图 18：对 Au_8Pd_4 进化结构搜索。(a, b)--总能量进化（为了清晰，(b) 可以在同一数据集能量最低的区域急剧上升），(c)--在我们的进化模拟中发现的能量最低结构，d---利用准噶尔集群扩展的计算发现的能量最低结构。请注意我们的结构 (c) 是这种化合物已知的最低能量结构。这建立了我们方法的权威性（即使是旧的 2007 版）。

9.3 INPUT.txt 输入文件示例

```
1 PARAMETERS EVOLUTIONARY ALGORITHM
2 % Example of the short input, using most options as defaults
3
4 % atomType
5 Mg Al O
6 % EndAtomType
7
8 % numSpecies
9 2 4 8
10 % EndNumSpecies
11
12 50 : numGenerations
13
14 % abinitioCode
15 3 3 3 3 3
16 % ENDabinit
17
18 % commandExecutable
19 gulp < input > output
20 % EndExecutable
```

9.4 空间群列表

1	P1	2	P-1	3	P2	4	P21
5	C2 (A2)*	6	Pm	7	Pc (Pa)*	8	Cm (Am)*
9	Cc (Aa)*	10	P2/m	11	P21/m	12	C2/m (A2/m)*
13	P2/c (P2/a)*	14	P21/c (P21/a)*	15	C2/c (A2/a)*	16	P222
17	P222 ₁	18	P2 ₁ 2 ₁ 2	19	P2 ₁ 2 ₁ 2 ₁	20	C222 ₁
21	C222	22	F222	23	I222	24	I2 ₁ 2 ₁ 2 ₁
25	Pmm2	26	Pmc2 ₁	27	Pcc2	28	Pma2
29	Pca2 ₁	30	Pnc2	31	Pmn2 ₁	32	Pba2
33	Pna2 ₁	34	Pnn2	35	Cmm2	36	Cmc2 ₁
37	Ccc2	38	Amm2 (C2mm)*	39	Aem2 (C2mb)*	40	Ama2 (C2cm)*
41	Aea2 (C2cb)*	42	Fmm2	43	Fdd2	44	Imm2
45	Iba2	46	Ima2	47	Pmmm	48	Pnnn
49	Pccm	50	Pban	51	Pmma	52	Pnna
53	Pmna	54	Pcca	55	Pbam	56	Pcen
57	Pbcm	58	Pnmm	59	Pmmn	60	Pbcn
61	Pbca	62	Pnma	63	Cmcn	64	Cmce (Cmcn)*
65	Cmmm	66	Ccem	67	Cmme (Cmma)*	68	Ccce (Ccea)*
69	Fmmm	70	Fddd	71	Immm	72	Ibam
73	Ibca	74	Imma	75	P4	76	P4 ₁
77	P4 ₂	78	P4 ₃	79	I4	80	I4 ₁
81	P-4	82	I-4	83	P4/m	84	P4 ₂ /m
85	P4/n	86	P4 ₂ /n	87	I4/m	88	I4 ₁ /a
89	P422	90	P4212	91	P4 ₁ 22	92	P4 ₁ 2 ₁ 2
93	P4 ₂ 22	94	P4 ₂ 2 ₁ 2	95	P4 ₃ 22	96	P4 ₃ 2 ₁ 2
97	I422	98	I4 ₁ 22	99	P4mm	100	P4bm
101	P4 ₂ cm	102	P4 ₂ nm	103	P4cc	104	P4nc
105	P4 ₂ mc	106	P4 ₂ bc	107	I4mm	108	I4cm
109	I4 ₁ md	110	I4 ₁ cd	111	P-42m	112	P-42c
113	P-42 ₁ m	114	P-42 ₁ c	115	P-4m2	116	P-4c2
117	P-4b2	118	P-4n2	119	I-4m2	120	I-4c2
121	I-42m	122	I-42d	123	P4/mmm	124	P4/mcc
125	P4/nbm	126	P4/ncc	127	P4/mbm	128	P4/mnc
129	P4/nmm	130	P4/ncc	131	P4 ₂ /mmc	132	P4 ₂ /mcm
133	P4 ₂ /nbc	134	P4 ₂ /nmn	135	P4 ₂ /mbe	136	P4 ₂ /mnmm
137	P4 ₂ /nmc	138	P4 ₂ /ncm	139	I4/mmm	140	I4/mem
141	I4 ₁ /amd	142	I4 ₁ /acd	143	P3	144	P3 ₁
145	P3 ₂	146	R3	147	P-3	148	R-3
149	P312	150	P321	151	P3 ₁ 12	152	P3 ₁ 21
153	P3 ₂ 12	154	P3 ₂ 21	155	R32	156	P3m1
157	P31m	158	P3c1	159	P31c	160	R3m
161	R3c	162	P-31m	163	P-31c	164	P-3m1
165	P-3cl	166	R-3m	167	R-3c	168	P6
169	P6 ₁	170	P6 ₅	171	P6 ₂	172	P6 ₄
173	P6 ₃	174	P-6	175	P6/m	176	P6 ₃ /m
177	P622	178	P6 ₁ 22	179	P6 ₅ 22	180	P6 ₂ 22
181	P6 ₄ 22	182	P6 ₃ 22	183	P6mm	184	P6cc
185	P6 ₃ cm	186	P6 ₃ mc	187	P-6m2	188	P-6c2
189	P-62m	190	P-62c	191	P6/mmm	192	P6/mcc
193	P6 ₃ /mcmm	194	P6 ₃ /mmc	195	P23	196	F23
197	I23	198	P2 ₁ 3	199	I2 ₁ 3	200	Pm-3
201	Pn-3	202	Fm-3	203	Fd-3	204	Im-3
205	Pa-3	206	Ia-3	207	P432	208	P4 ₂ 32
209	F432	210	F4 ₁ 32	211	I432	212	P4 ₃ 32
213	P4 ₁ 32	214	I4 ₁ 32	215	P-43m	216	F-43m
217	I-43m	218	P-43n	219	F-43c	220	I-43d
221	Pm-3m	222	Pn-3n	223	Pm-3n	224	Pn-3m
225	Fm-3m	226	Fm-3c	227	Fd-3m	228	Fd-3c
229	Im-3m	230	Ia-3d				

*括号里是在 USPEX 代码里使用的非标准空间群标号。

9.5 平面群列表

Number	Group
1	Group p1
2	Group p2
3	Group pm
4	Group pg
5	Group cm
6	Group pmm
7	Group pmg
8	Group pgg
9	Group cmm
10	Group p4
11	Group p4m
12	Group p4g
13	Group p3
14	Group p3m1
15	Group p31m
16	Group p6
17	Group p6m

9.6 点群列表

罗列出了在 Schönflies 和 Hermann-Maugin (国际) 符号中的所有晶体和最重要非晶体点群。

<i>Crystallographic point groups:</i>		
Hermann-Maugin	Schönflies	In USPEX
1	C ₁	C1 or E
2	C ₂	C2
222	D ₂	D2
4	C ₄	C4
3	C ₃	C3
6	C ₆	C6
23	T	T
1	S ₂	S2
M	C _{1h}	Ch1
mm2	C _{2v}	Cv2
2	S ₄	S4
3	S ₆	S6
6	C _{3h}	Ch3
m̄3	T _h	Th
2/m	C _{2h}	Ch2
mmm	D _{2h}	Dh2
4/m	C _{4h}	Ch4
32	D ₃	D3
6/m	C _{6h}	Ch6
432	O	O
422	D ₄	D4
3m	C _{3v}	Cv3
622	D ₆	D6
43m	T _d	Td
4mm	C _{4v}	Cv4
3m	D _{3d}	Dd3
6mm	C _{6v}	Cv6
m̄3m	O _h	Oh
42m	D _{2d}	Dd2
62m	D _{3h}	Dh3
4/mmm	D _{4h}	Dh4
6/mmm	D _{6h}	Dh6
m̄3m	O _h	Oh

Important non-crystallographic point groups

Hermann-Maugin	Schönflies	In USPEX
5	C ₅	C5
5/m	S ₅	S5
5̄	S ₁₀	S10
5m	Cv _{5v}	Cv5
10	Ch _{5h}	Ch5
52	D ₅	D5
5̄m	D _{5d}	Dd5
102m	D _{5h}	Dh5
532	I	I
5̄3m	I _h	Ih

9.7 USPEX 中使用的共价原子半径表

USPEX 中所使用的原子半径（主要用于计算硬度）：

Z	Element	r	Z	Element	r	Z	Element	r
1	H	0.31	30	Zn	1.22	63	Eu	1.98
2	He	0.28	31	Ga	1.22	64	Gd	1.96
3	Li	1.28	32	Ge	1.20	65	Tb	1.94
4	Be	0.96	33	As	1.19	66	Dy	1.92
5	B	0.84	34	Se	1.20	67	Ho	1.92
6	Csp ³	0.76	35	Br	1.20	68	Er	1.89
	Csp ²	0.73	36	Kr	1.16	69	Tm	1.90
	Csp	0.69	37	Rb	2.20	70	Yb	1.87
7	N	0.71	38	Sr	1.95	71	Lu	1.87
8	O	0.66	39	Y	1.90	72	Hf	1.75
9	F	0.57	40	Zr	1.75	73	Ta	1.70
10	Ne	0.58	41	Nb	1.64	74	W	1.62
11	Na	1.66	42	Mo	1.54	75	Re	1.51
12	Mg	1.41	43	Tc	1.47	76	Os	1.44
13	Al	1.21	44	Ru	1.46	77	Ir	1.41
14	Si	1.11	45	Rh	1.42	78	Pt	1.36
15	P	1.07	46	Pd	1.39	79	Au	1.36
16	S	1.05	47	Ag	1.45	80	Hg	1.32
17	Cl	1.02	48	Cd	1.44	81	Tl	1.45
18	Ar	1.06	49	In	1.42	82	Pb	1.46
19	K	2.03	50	Sn	1.39	83	Bi	1.48
20	Ca	1.76	51	Sb	1.39	84	Po	1.40
21	Sc	1.70	52	Te	1.38	85	At	1.50
22	Ti	1.60	53	I	1.39	86	Rn	1.50
23	V	1.53	54	Xe	1.40	87	Fr	2.60
24	Cr	1.39	55	Cs	2.44	88	Ra	2.21
25	Mn l.s.	1.39	56	Ba	2.15	89	Ac	2.15
	h.s.	1.61	57	La	2.07	90	Th	2.06
26	Fe l.s.	1.32	58	Ce	2.04	91	Pa	2.00
	h.s.	1.52	59	Pr	2.03	92	U	1.96
27	Co l.s.	1.26	60	Nd	2.01	93	Np	1.90
	h.s.	1.50	61	Pm	1.99	94	Pu	1.87
28	Ni	1.24	62	Sm	1.98	95	Am	1.80
29	Cu	1.32				96	Cm	1.69

Source: Cordero *et al.*, Dalton Trans. 2832-2838, 2008.

9.8 USPEX 使用的默认化合价表

USPEX 中使用的化合价表（主要用于硬度计算）：

Z	Element	valence	Z	Element	valence	Z	Element	valence
1	H	1	35	Br	1	69	Tm	3
2	He	0.5	36	Kr	0.5	70	Yb	3
3	Li	1	37	Rb	1	71	Lu	3
4	Be	2	38	Sr	2	72	Hf	4
5	B	3	39	Y	3	73	Ta	5
6	C	4	40	Zr	4	74	W	4
7	N	3	41	Nb	5	75	Re	4
8	O	2	42	Mo	4	76	Os	4
9	F	1	43	Tc	4	77	Ir	4
10	Ne	0.5	44	Ru	4	78	Pt	4
11	Na	1	45	Rh	4	79	Au	1
12	Mg	2	46	Pd	4	80	Hg	2
13	Al	3	47	Ag	1	81	Tl	3
14	Si	4	48	Cd	2	82	Pb	4
15	P	3	49	In	3	83	Bi	3
16	S	2	50	Sn	4	84	Po	2
17	Cl	1	51	Sb	3	85	At	1
18	Ar	0.5	52	Te	2	86	Rn	0.5
19	K	1	53	I	1	87	Fr	1
20	Ca	2	54	Xe	0.5	88	Ra	2
21	Sc	3	55	Cs	1	89	Ac	3
22	Ti	4	56	Ba	2	90	Th	4
23	V	4	57	La	3	91	Pa	4
24	Cr	3	58	Ce	4	92	U	4
25	Mn	4	59	Pr	3	93	Np	4
26	Fe	3	60	Nd	3	94	Pu	4
27	Co	3	61	Pm	3	95	Am	4
28	Ni	2	62	Sm	3	96	Cm	4
29	Cu	2	63	Eu	3	97	Bk	4
30	Zn	2	64	Gd	3	98	Cf	4
31	Ga	3	65	Tb	3	99	Es	4
32	Ge	4	66	Dy	3	100	FM	4
33	As	3	67	Ho	3	101	Md	4
34	Se	2	68	Er	3	102	No	4

9.9 USPEX 使用的默认 goodBonds 表

USPEX 中使用的 goodBonds 表（主要用于硬度计算）：

Z	Element		Z	Element		Z	Element	
1	H	0.20	35	Br	0.10	69	Tm	0.20
2	He	0.05	36	Kr	0.05	70	Yb	0.20
3	Li	0.10	37	Rb	0.05	71	Lu	0.20
4	Be	0.20	38	Sr	0.10	72	Hf	0.30
5	B	0.30	39	Y	0.20	73	Ta	0.40
6	C	0.50	40	Zr	0.30	74	W	0.30
7	N	0.50	41	Nb	0.35	75	Re	0.30
8	O	0.30	42	Mo	0.30	76	Os	0.30
9	F	0.10	43	Tc	0.30	77	Ir	0.30
10	Ne	0.05	44	Ru	0.30	78	Pt	0.30
11	Na	0.05	45	Rh	0.30	79	Au	0.05
12	Mg	0.10	46	Pd	0.30	80	Hg	0.10
13	Al	0.20	47	Ag	0.05	81	Tl	0.20
14	Si	0.30	48	Cd	0.10	82	Pb	0.30
15	P	0.30	49	In	0.20	83	Bi	0.20
16	S	0.20	50	Sn	0.30	84	Po	0.20
17	Cl	0.10	51	Sb	0.20	85	At	0.10
18	Ar	0.05	52	Te	0.20	86	Rn	0.05
19	K	0.05	53	I	0.10	87	Fr	0.05
20	Ca	0.10	54	Xe	0.05	88	Ra	0.10
21	Sc	0.20	55	Cs	0.05	89	Ac	0.20
22	Ti	0.30	56	Ba	0.10	90	Th	0.30
23	V	0.30	57	La	0.20	91	Pa	0.30
24	Cr	0.25	58	Ce	0.30	92	U	0.30
25	Mn	0.30	59	Pr	0.20	93	Np	0.30
26	Fe	0.25	60	Nd	0.20	94	Pu	0.30
27	Co	0.25	61	Pm	0.20	95	Am	0.30
28	Ni	0.15	62	Sm	0.20	96	Cm	0.30
29	Cu	0.10	63	Eu	0.20	97	Bk	0.30
30	Zn	0.10	64	Gd	0.20	98	Cf	0.30
31	Ga	0.25	65	Tb	0.20	99	Es	0.30
32	Ge	0.50	66	Dy	0.20	100	FM	0.30
33	As	0.35	67	Ho	0.20	101	Md	0.30
34	Se	0.20	68	Er	0.20	102	No	0.30

10 参考书目

- [1] A.R. Oganov. Modern Methods of Crystal Structure Prediction. Wiley-VCH, 1 edition, 2011.
- [2] J. Maddox. Crystals from first principles. *Nature*, 335:201, 1988.
- [3] A.R. Oganov and C.W. Glass. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *J. Chem. Phys.*, 124:244704, 2006.
- [4] C.W. Glass, A.R. Oganov, and N. Hansen. USPEX-evolutionary crystal structure prediction. *Comp. Phys. Comm.*, 175:713-720, 2006.
- [5] A.R. Oganov and C.W. Glass. Theoretical and experimental evidence for a post-perovskite phase of MgSiO₃ in Earth's D" layer. *Nature*, 430:445-448, 2004.
- [6] M. Murakami, K. Hirose, K. Kawamura, N. Sata, and Y. Ohishi. Post-perovskite phase transition in MgSiO₃. *Science*, 304(5672):855-858, 2004.
- [7] A.R. Oganov, J.C. Schon, M. Jansen, S.M. Woodley, W.W. Tipton, and R.G. Hennig. First blind test of inorganic crystal structure prediction. In A.R. Oganov, editor, *Modern Methods of Crystal Structure Prediction*. Wiley-VCH, 2010.
- [8] C.J. Pickard and R.J. Needs. High-pressure phases of silane. *Phys. Rev. Lett.*, 97:045504, 2006.
- [9] M. Martinez-Canales, A.R. Oganov, A. Lyakhov, Y. Ma, and A. Bergara. Novel structures of silane under pressure. *Phys. Rev. Lett.*, 102:087005, 2009.
- [10] Y. Ma, A.R. Oganov, Y. Xie, Z. Li, and J. Kotakoski. Novel high pressure structures of polymeric nitrogen. *Phys. Rev. Lett.*, 102:065501, 2009.
- [11] C.J. Pickard and R.J. Needs. High-pressure phases of nitrogen. *Phys. Rev. Lett.*, 102:125702, 2009.
- [12] G. Gao, A.R. Oganov, Z. Li, P. Li, T. Cui, A. Bergara, Y. Ma, T. Iitaka, and G. Zou. Crystal structures and superconductivity of stannane under high pressure. *Proc. Natl. Acad. Sci.*, 107:1317-1320, 2010.
- [13] C.J. Pickard and R.J. Needs. Structures at high pressure from random searching. *Phys. Status Solidi*, 246:536-540, 2009.

- [14] A.O. Lyakhov, A.R. Oganov, H.T. Stokes, and Q. Zhu. New developments in evolutionary structure prediction algorithm USPEX. *Comp. Phys. Comm.*, 184:1172-1182, 2013.
- [15] G.R. Qian, X. Dong, X.-F. Zhou, Y. Tian, A.R. Oganov, and H.-T. Wang. Variable cell nudged elastic band method for studying solid-solid structural phase transitions. *Comput. Phys. Comm.*, 184:2111-2118, 2013.
- [16] C. Dellago, P.G. Bolhuis, F.S. Csajka, and D. Chandler. Transition path sampling and the calculation of rate constants. *J. Chem. Phys.*, 108(5):1964-1977, 1998.
- [17] S.E. Boulfelfel, A.R. Oganov, and S. Leoni. Understanding the nature of "superhard graphite". *Scientific Reports*, 2(471), 2012.
- [18] A.R. Oganov and C.W. Glass. Evolutionary crystal structure prediction as a tool in materials design. *J. Phys.: Cond. Matter*, 20:064210, 2008.
- [19] G. Kresse and J. Furthmuller. Efficient iterative schemes for ab initio total-energy calculations using plane wave basis set. *Phys. Rev.*, B54:11169-11186, 1996.
- [20] J.M. Soler, E. Artacho, J.D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal. The SIESTA method for ab initio order-N materials simulation. *J. Phys.: Cond. Matter*, 14:2745-2779, 2002.
- [21] J.D. Gale. GULP: Capabilities and prospects. *Z. Krist.*, 220:552-554, 2005.
- [22] A.R. Oganov and M. Valle. How to quantify energy landscapes of solids. *J. Chem. Phys.*, 130:104504, 2009.
- [23] G.H. Johannesson, T. Bligaard, A.V. Ruban, H.L. Skriver, K.W. Jacobsen, and J.K.Nørskov. Combined electronic structure and evolutionary search approach to materials design. *Phys. Rev. Lett.*, 88:255506, 2002.
- [24] A.D. Becke. Density functional thermochemistry. III. the role of exact exchange. *J. Chem. Phys.*, 98(7):5648-5652, 1993.
- [25] J.M. Tao, J.P. Perdew, V.N. Staroverov, and G.E. Scuseria. Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.*, 91:146401, 2003.
- [26] M. Fuchs and X. Gonze. Accurate density functionals: approaches using the adiabaticconnection

- uctuation-dissipational theorem. Phys. Rev., B65:235109, 2002.
- [27] A.I. Liechtenstein, V.I. Anisimov, and J. Zaanen. Density functional theory and strong interactions - orbital ordering in mott-hubbard insulators. Phys. Rev., B52:R5467-R5470, 1995.
- [28] P. Strange, A. Svane, W.M. Temmerman, Z. Sszotek, and H. Winter. Understanding the valency of rare earths from first-principles theory. Nature, 399:756-758, 1999.
- [29] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. Rev. Mod. Phys., 68:13-125, 1996.
- [30] W.M.C. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal. Quantum monte carlo simulation of solids. Rev. Mod. Phys., 73:33-83, 2001.
- [31] A.R. Oganov, A.O. Lyakhov, and M. Valle. How evolutionary crystal structure prediction works -and why. Accounts of Chemical Research, 44(3):227-237, 2011.
- [32] X.-Q. Chen, H. Niu, D. Li, and Y. Li. Modeling hardness of polycrystalline materials and bulk metallic glasses. Intermetallics, 19(9):1275-1281, 2011.
- [33] A.R. Oganov and A.O. Lyakhov. Towards the theory of hardness of materials. Journal of Superhard Materials, 32(3):143-147, 2010.
- [34] L.S. Dubrovinsky, N.A. Dubrovinskaia, V. Swamy, J. Muscat, N.M. Harrison, R. Ahuja, B. Holm, and B. Johansson. Materials science: The hardest known oxide. Nature, 410(6829):653-654, 2001.
- [35] R. Martonak, A. Laio, M. Bernasconi, C. Ceriani, P. Raiteri, F. Zipoli, and M. Parrinello. Simulation of structural phase transitions by metadynamics. Z. Krist., 220:489-498, 2005.
- [36] A.O. Lyakhov, A.R. Oganov, and M. Valle. Crystal structure prediction using evolutionary approach. In A.R. Oganov, editor, Modern Methods of Crystal Structure Prediction, pages 147-180. Berlin: Wiley-VCH, 2010.
- [37] Q. Zhu, L. Li, A.R. Oganov, and P.B. Allen. Evolutionary method for predicting surface reconstructions with variable stoichiometry. Phys. Rev. B, 87:195317, May 2013.
- [38] W. Zhang, A.R. Oganov, A.F. Goncharov, Q. Zhu, S.E. Boulfelfel, A.O. Lyakhov, E. Stavrou, M. Somayazulu, V.B. Prakapenka, and Z. Konopkova. Unexpected stable stoichiometries of sodium chlorides. Science, 342(6165):1502-1505, 2013.

- [39] S.T. Call, D.Yu. Zubarev, and A.I. Boldyrev. Global minimum structure searches via particle swarm optimization. *Journal of Computational Chemistry*, 28(7):1177-1186, 2007.
- [40] Y. Wang, J. Lv, L. Zhu, and Y. Ma. Crystal structure prediction via particle-swarm optimization. *Phys. Rev. B*, 82:094116, Sep 2010.
- [41] A.O. Lyakhov, A.R. Oganov, and M. Valle. How to predict very large and complex crystal structures. *Comp. Phys. Comm.*, 181:1623-1632, 2010.
- [42] G. Mills, H. Jónsson, and G.K. Schenter. Reversible work transition state theory: application to dissociative adsorption of hydrogen. *Surf. Sci.*, 324(2):305-337, 1995.
- [43] G. Henkelman, B.P. Uberuaga, and H. Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.*, 113(22):9901-9904, 2000.
- [44] G. Henkelman and H. Jonsson. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.*, 113(22):9978-9985, 2000.
- [45] E. Bitzek, P. Koskinen, F. Gähler,, M. Moseler, and P. Gumbsch. Structural relaxation made simple. *Phys. Rev. Lett.*, 97(17):170201, 2006.
- [46] M. Valle. STM3: a chemistry visualization platform. *Z. Krist.*, 220:585-588, 2005.