无序合金高通量计算平台 HCPRA 的搭建及其应用

李婷婷^{1,2}, 黄孙超¹, 张小丽¹, 曾雉^{1,2*}

1. 中国科学院固体物理研究所, 中国科学院材料物理重点实验室, 合肥 230031;

2. 中国科学技术大学科学岛分院, 合肥 230026

* 联系人, E-mail: zzeng@theory.issp.ac.cn

2017-06-11 收稿, 2017-08-13 修回, 2017-08-14 接受, 2017-10-10 网络版发表

国家自然科学基金(11534012, U1230202)资助

摘要 介绍了无序合金高通量计算平台(high-throughput computation platform for random alloys, HCPRA)的设计难点、解决方案和基本框架. 详细地介绍了如何让平台自动化地创建大量的初始结构, 为各体系设置精度相当的计算参数以及如何提取和分析计算结果. 该平台采用Fortran和Shell语言搭建, 主要功能是实现大批量自动化地计算无序合金的力学和热力学性质. 为了测试其效率, 利用该平台计算了47种金属单质的弹性常数, 结果表明一台8核的计算机运行一天就能完成上述计算.

关键词 高通量平台, 无序合金, 力学性质, 弹性常数, 热力学性质

2013年美国科学家Jain等人^[1]发表了一篇关于材料基因组的评论文章. 文章指出新材料的研发对技术的进步有巨大的推动作用, 例如钢铁和铜基材料的研发对社会主要技术的发展产生了重要作用, 并且推动了人类文明的发展和传播. 如今人类社会面临气候变暖、水土流失和能源枯竭等众多危机, 亟需找到新材料和新技术来解决这些问题. 然而目前我们对新材料的研发仍然具有很大的不确定性和不可预测性, 从新材料的研发到生产往往需要几十年的时间, 这显然不利于我们解决目前的危机. 为了提高新材料研发效率, 2011年美国提出了材料基因组计划^[2]. 该计划的目标是将材料研发周期和研发成本同时减半, 它的核心部分是构建高通量计算平台和开放式的数据库. 高通量计算平台可以在很短的时间内产生大量的数据, 为材料筛选和设计提供原始的数据, 然后通过数据挖掘等手段对计算数据进行深度分析, 从而设计新材料. 因此这种方法是设计新材料, 解决社会问题的重要手段. 与材料基因组类

似的计划有欧盟的“冶金复兴”^[3]、日本的材料数据库计划^[4]和德国的工业4.0计划^[5].

高通量的思想萌芽于20世纪90年代, 并且在实验和理论领域都有一定的应用. 例如1995年实验科学家Xiang等人^[6]利用高通量实验方法制备了200多种薄膜样品, 他希望通过这种方法来加速电磁和光学材料的研究, 相应的科学成果发表在*Science*上. 理论科学家Chiang等人^[7]用高通量计算的思想研究了高密度锂电池, 他们认为用非过渡金属元素替换锂电池中的过渡金属元素可以提高电池的性能, 而这种替换存在很多可能性, 于是他们用高通量计算的思想研究了这些潜在对象. 最终发现锂氧电池不仅容量很高, 而且可以降低制备成本.

虽然高通量思想早已经在计算领域里萌芽, 但是并未得到广泛应用. 因为这种方法涉及大量的计算, 人工处理需要耗费大量的时间和精力, 而且容易出错, 例如在计算管理上出现重复、遗漏以及参数设置前后不一致等问题. 高通量计算平台可以高效地

引用格式: 李婷婷, 黄孙超, 张小丽, 等. 无序合金高通量计算平台 HCPRA 的搭建及其应用. 科学通报, 2017, 62: 3894–3901

Li T T, Huang S C, Zhang X L, et al. Construction of high-throughput computation platform for random alloys (HCPRA) and its applications (in Chinese). *Chin Sci Bull*, 2017, 62: 3894–3901, doi: 10.1360/N972017-00300

管理计算任务,从而极大地减少人力劳动,因此可以推广高通量计算方法在材料设计领域中的应用.为了所有已知无机化合物构造电子结构数据库,Setyawan等人^[8-10]建立了一个非常有名的高通量计算平台AFLOW.这个平台由15万余行C++语言组成,运行环境为Linux.原则上它可以调用众多的软件来进行第一性原理计算,但是目前只支持VASP(Vienna *ab-initio* simulation package).AFLOW的初始结构是从ICSD(the inorganic crystal structure database)和Pauling File等实验数据库中提取的.它本身也配备了一个数据库,用于存放所有的计算数据,以便后期的参考和调用.基于AFLOW至今已发表上百篇文献[9~13],例如Setyawan等人^[10]利用该平台从193456种化合物中筛选了7439种潜在研究对象,然后进行能带计算,最终找到了一系列有效质量低材料.但是AFLOW面向国内还未完全开放.

中国科学家在此领域也做了许多工作并发表了相应的科学论文,例如王海舟等人^[14]很好地阐述了高通量实验方法的相关问题.文章系统地介绍了高通量实验方法的起源、发展、样品制备和样品表征.同时文章也提到了中国科学技术大学和中国科学院大连物理化学研究所等单位在这方面所做的工作.王卓等人^[15]系统地介绍了他们对高通量计算平台的看法和认知,以及他们课题组在这方面所做的工作.杨小瑜、曾雉、张平和林海清等材料计算模拟研究小组^[16,17]相互合作,以MatCloud为基础结合无序合金的应用需求,构建了用于设计二元合金的高通量计算模拟平台(MatCloud-SQS),实现了高通量锆基二元合金结构筛选、物性计算和数据库产生等功能.此外该平台配备了一个数据库系统^[18],用于自动保存平台所产生的数据. MatCloud-SQS是一种在线使用的、并以图形为交互界面的平台,在整个使用过程中不需要安装任何软件,此外图形界面易于初学者使用.但是该平台需要在联网的条件下使用,而且在计算前后需要通过互联网传输大量的数据,对网络资源方面有很高的要求.

为了解决这一问题,我们开发了一个离线版本的无序合金高通量计算平台HCPRA(high-throughput computation platform for random alloys)^[19].该平台是采用Fortran和Shell语言搭建的,能够在Linux平台上运行.虽然使用该平台需要掌握一定的Linux知识,但是它的使用是非常简单和灵活的,例如用户可以修改

其中的任何一段程序来使平台适合自己的研究工作.

1 无序合金高通量计算平台HCPRA

1.1 平台设计难点和相应解决方案

搭建一个高通量的计算平台,主要有以下3个方面的难点.首先是如何创建大量的初始结构并实现程序的自动调用,初始结构是计算的第一步,所以要先根据研究对象构建相应的初始结构.其次是如何为不同的体系和结构设置精度相当的计算参数,只有设置精度相当的计算参数才能使不同体系的计算结果具有可比性,这也是高通量计算区别于传统计算的主要特点之一.传统的计算是每一个课题组详细研究某一个或几个体系,虽然从全世界范围内来看,绝大多数的体系都研究过了,但是很难将这些数据收集到一起,就算将这些结果整合成一个数据库,不同数据也不具备可比性,因为不同组采用了不同的计算参数.而高通量计算是一个组采用精度相当的参数来研究不同的体系,所以利于收集计算结果且各计算结果之间具备可比性.最后一个难点是如何提取和呈现计算结果,高通量计算会一次性计算出上百个体系或结构,如何提取计算结果,并将这些结果分类呈现出来是一个很大的难题.

对于第一个难点我们的解决方案是采用特殊准随机结构模型为研究对象:二元无序合金构建初始结构(special quasi-random structures, SQSs^[20]).该方法采用一个有限大小的超胞来描述固溶体的无序现象,可以平衡计算量和精度之间的矛盾^[19,20].它的核心是构建一个超胞,然后通过调整原子的占位使该结构的关联函数尽可能地和标准无序合金的关联函数相匹配.二元标准无序合金的关联函数可以写成 $\Pi_{m,k} = (2x - 1)^k$,其中 x 是成分比, k 是表征该结构多体相互作用的指标;SQSs的关联函数是用Ising模型的语言定义,具体细节可参见文献[19,21~24].SQSs结构的选择标准是关联函数匹配度而不是通常所用的能量最低原则.因为研究体系是无序合金,所以在构建结构时,我们的第一准则是尽可能确保合金的无序性而不是能量最低.事实上,能量最低原则在描述无序合金时具有一定的局限性.无序合金具有明显的点阵结构,但是组分原子在格点上的排布是完全随机的.因此严格来讲,无序合金是没有周期性的,用有限大小的超胞来描述无序体系,一定会带来能

量上的截断误差. 此外, 不同超胞的原子弛豫和晶格畸变都有所不同, 这些因素都会造成能量误差, 从而导致能量最低的结构不一定能描述真实体系. 相反, 无序度判断准则不仅完全不受原子弛豫和晶格畸变的影响, 而且与具体的原子没有关系. 这就使得我们可以用一组SQSs模板来模拟所有的二元无序合金. 本平台采用一组包含7种组分0%(A_2), 6.25%($A_{15}B$), 25%($A_{12}B_4$), 50%(A_8B_8), 75%(A_4B_{12}), 93.75%(AB_{15})和100%(B_2)的SQSs模板来描述每一个无序合金体系 $A_{1-x}B_x$, 其中 $A_{12}B_4$, A_8B_8 和 A_4B_{12} 为SQSs结构, $A_{15}B$ 和 AB_{15} 为计算条件下的固溶极限结构, 即超胞中溶质原子的个数为1, 他们是 A_8B_8 的一个变体, 即超胞形状和格点位置保持一致. 图1为密排六方体系的SQSs结构模板示意图, 其中图1(c)的基本框架是由图1(d)沿 x , y , z 方向等比扩大两倍所形成的, 但是成分原子在格点上的占位是无序的.

SQSs模板可以显著降低构建大量初始结构的工作量, 但是SQSs模板并不等同于初始结构. 当用SQSs模板来描述某一具体体系时, 还需要将体系的原子信息和晶格常数信息添加到结构模板. 因为不同的体系具有不同的晶格常数, 同一体系的不同组分也具有不同的晶格常数信息. 本平台的研究对象是无序合金, 所以几乎找不到实验或理论晶格常数. 我们的解决方案是采用Neumann-Kopp定理:

$$y = a(1-x) + bx, \quad (1)$$

来确定初始结构的晶格常数, 这里的 y , a 和 b 分别为合金、单质 A 和单质 B 的晶格常数, x 为 B 的组分. 合理的初始结构可以有效地降低第一性原理计算量并增加得到平衡态的几率, 因此这一定理对开展高通量计算非常重要. 以锆钛体系为例, 由Neumann-Kopp定理得到的晶格常数(虚线)和第一性原理计算得到的

晶格常数(实线)非常接近, 如图2所示. 同时图2给出了锆和钛单质的实验晶格常数, 和第一性原理计算结果之间的误差仅为0.5%左右, 说明了结构文件的构建是非常高效和合理的.

如何为不同的体系和结构设置精度相当的计算参数, 是实现高通量计算的核心所在, 也是搭建高通量计算平台的一大难点. 本平台是调用VASP做第一性原理计算, 因此需要设置的参数文件有两个, 分别为KPOINTS和INCAR文件. 其中KPOINTS文件决定的是倒空间的网格是如何划分, 为了使各体系的计算精度一致, 需要保证网格点的划分密度一致, 所以只需要设置自动划分网格点即可. INCAR是决定第一性原理计算收敛条件的文件, 为了保证各体系的计算精度一致, 需要保证各系统的收敛条件相当. 然而INCAR文件中有一个特殊的参数: 能量截断值(ENCUT), 它对计算精度有着绝对性地影响, 而且对于不同体系它的大小不一样, 因此不能在INCAR中显式地给出一个具体的数值. 我们的处理方案是分别从各目标体系的势函数(POTCAR)当中读取最大的默认截断能, 然后再统一乘以 n 倍作为各体系的ENCUT的值, n 的值由参数测试决定, 通常应为1.3. ISMEAR也是INCAR当中一个非常重要的参数, 它决定了价电子对费米面附近能态的占据情况. 对于无杂质的半导体或者绝缘体, 价电子恰好填充完所有的价带而导带全空, 对每个态的占据概率为1或者0; 对于金属或者合金, 波函数的相互交叠会使得价电子以分数的形式占据费米面附近的能态. 在VASP当中, 这种占据可以通过不同的分布函数来控制. 在计算当中, 我们采用官方推荐, 也就是使用ISMEAR=-5(带Blochl修正的四面体方法, 占据概率只能为1或者0)来计算所有体系的总能和能态密度,

图1 (网络版彩色)密排六方体系的不同组分结构模板. (a) $A_{15}B$ 和 AB_{15} ; (b) $A_{12}B_4$ 和 A_4B_{12} ; (c) A_8B_8 ; (d) A_2 和 B_2

Figure 1 (Color online) Structure templates for hexagonal close-packed (HCP) systems. (a) $A_{15}B$ and AB_{15} ; (b) $A_{12}B_4$ and A_4B_{12} ; (c) A_8B_8 ; (d) A_2 and B_2

图2 (网络版彩色)ZrTi合金晶格常数(单位为Å)随成分的变化关系
Figure 2 (Color online) Variation of ZrTi lattice parameter vs Ti concentration

该方法适合所有的大带隙半导体和绝缘体体系以及块体材料的总能和态密度的计算;采用ISMEAR=1(Methfessel-Paxton方法,允许部分占据,适合金属体系)和默认的SIGMA值来优化合金的几何结构和计算他们的力学性质。

高通量计算的一个特点是产生大数据输出,有效地提取和分析数据至关重要。在我们的平台中,数据处理特点是纵向提取数据,横向分析数据。纵向提取数据是指尽可能多地从第一性原理计算结果中提取力学性质数据,然后利用各种理论模型和经验公式从这些数据中得到其他力学量。纵向提取的核心是得到具有结构对称性的弹性常数矩阵。平台采用SQSs技术构建结构,而SQSs结构的尺寸是有限的,使得等价晶向[1000]和[0100]的平均原子分布出现偏差,造成 C_{11} 和 C_{22} 的大小不再严格相等;此外在计算弹性常数时,第一性原理方法本身也会引入一些误差,导致计算得到的弹性常数会偏离完美的晶格对称性,这部分误差可以认为是计算误差。通常来讲,这些因素造成的弹性常数矩阵对称性偏离度各不一样,为了后续的对比和研究,必须先对弹性常数矩阵进行重整化。采用的方法是Moakher等人^[25]提出的弹性常数投影技术。从重整化后的弹性常数矩阵中,可以提取弹性常数、体弹模量、剪切模量和泊松比等一系列力学参数。横向分析是指按合金体系整理数据。虽然同一体系各组分间的计算是相互独立的,但是在数据分析时,可以将他们放在一起。这样可以更好地使用数据挖掘技术,从大量的计算结果中提取有用的信息和未被发现的规律。例如在研究Zr-Ti、Zr-Hf和Zr-Sc体系时,我们发现合金的力学性质随成分

的变化关系可以用一个单参数的特殊二次函数来描述^[26]。

$$y = M_{Zr}(1-x) + M_x x + x(1-x)F. \quad (2)$$

图3为ZrSc体系的拟合结果,从中可以看出拟合效果非常理想,即第一性原理计算结果基本落在拟合曲线上。有了这些拟合曲线,我们可以快捷地得到其他成分点的力学性质,同时可以得到力学性质随成分变化的最大值或最小值。此外拟合参数 F 也具有一定的物理意义,它表示合金成分对力学性能的贡献。

我们将这个规律集成到计算平台中,并利用它来分析所有二元体系的力学性质,希望它能够帮助我们进一步了解合金力学性质的变化规律。当有新规律被发现后,这些新规律又会被集成到计算平台中使它初步具备深度自我学习的数据挖掘能力。

高通量计算会产生大量的原始数据,虽然目标性质已经在数据分析阶段提取完毕,这些原始数据仍然具有许多潜在价值,例如验证已有数据的可靠性和提取其他性质。所以在高通量计算的过程中,所有的原始数据都分类保存在数据库中,以便后期访问。

错误信息提示和处理是一项非常重要的功能,也是平台设计难点之一。高通量计算会提交大量的任务,这就不可避免地出现各种各样的错误。该平台在任务管理上的一大特点是,各任务之间是相互独立的,当一个任务出现错误时,该任务终止运行,并给出相应的错误信息,其他任务继续运行。根据错误信息可以修改相应参数重新提交任务,直至得到目标性能。平台暂时不提供自动纠错功能,这并不是因为存在技术上或理论上的困难,而是为了保证数据可靠性。高通量计算和材料高通量筛选不同于传统的计算,这里不希望通过调整各种各样的计算参数来获得与实验一致的结果,而是更加注重各数据间的横向可比性。为了达到这个目的,在整个计算过程中都尽可能使用相当的计算精度和计算参数。例如对所有的计算都采用相同的 K 点密度,能量截断半径都为默认值的 n 倍。因此当计算出现错误时,更倾向于手动调节参数,以便了解各数据间的可比性。

1.2 平台的基本框架

当平台的设计难点被扫除后,就可以开始平台构建了。该无序合金高通量计算平台(HCPRA)是采用Fortran90和Shell编写的程序代码,运行环境为Linux,主要功能是实现无序合金热力学性质和力学性质的

图3 (网络版彩色)ZrSc合金的力学性质. 圆点为第一性原理计算结果, 实线为特殊二次函数拟合结果

Figure 3 (Color online) Mechanical properties of ZrSc alloy system. Spots are the first principles calculation results, lines are the special quadratic fitting results

高通量计算和相应数据库的构建. 目前能够计算的力学性质有弹性常数、弹性模量和泊松比等. 平台采用特殊准无序模型(SQSs)建模^[20], 利用VASP进行第一性原理计算, 共包含5个部分, 如图4所示: 分别为结构文件、性质计算、数据库、数据分析和数据应用.

1.3 性质计算的理论依据

(i) 热力学性质. 热力学稳定性是判断结构是否合理的一个重要依据. 一般来讲, 如果合金能够稳定存在, 那么它的混合能 E 应该小于0, 其定义如下式所示:

$$E = (E(A_x B_y) - xE(A) - yE(B)) / (x + y), \quad (3)$$

其中 E 代表混合能, $E(A)$ 、 $E(B)$ 和 $E(A_x B_y)$ 分别代表

组分单质A、组分单质B和合金的自由能. 自由能通常是由第一性原理计算直接得到, 然而第一性原理计算只能得到0 K下的能量, 因此需要借用德拜模型来考虑物质在有限温度下的自由能. 在有限温度下, 物质的自由能包含3部分, 分别为0 K下的能量, 振动能和振动熵, 其中振动能和振动熵由德拜模型来估算. 在德拜模型中, 振动能和振动熵分别由下列两式给出:

$$E_{\text{vib}} = \frac{9}{8} k_B \Theta_D + 3k_B T D(\Theta_D/T), \quad (4)$$

$$S_{\text{vib}} = 3k_B \left(\frac{4}{3} D(\Theta_D/T) - \ln(1 - \exp(-\Theta_D/T)) \right), \quad (5)$$

其中 k_B 为玻尔兹曼常数, $D(\Theta_D/T)$ 为德拜函数, Θ_D 为德拜温度, 德拜温度可以利用德拜频率和弹性模量计算得到, 具体计算公式如下所示:

$$\Theta_D = \frac{h}{2\pi k_B} \left(\frac{6\pi^2}{V} \right)^{\frac{1}{3}} v_D, \quad (6)$$

$$\frac{1}{v_D^3} = \frac{1}{3} \left(\left(\frac{\rho}{B} \right)^{\frac{3}{2}} + 2 \left(\frac{\rho}{G} \right)^{\frac{3}{2}} \right), \quad (7)$$

$$D(\Theta_D/T) = \frac{3}{y^3} \int_0^y \frac{x^3}{e^x - 1} dx. \quad (8)$$

图4 (网络版彩色)HCPRA 平台的基本构架

Figure 4 (Color online) Fundamental frame of HCPRA

此处 V 为体积, ρ 为密度, v_D 为平均声速, B 和 G 分别为体弹模量和剪切模量. 考虑温度效应后, 晶体的自由能变为

$$F = E + E_{\text{vib}} - TS_{\text{vib}}. \quad (9)$$

事实上在求无序合金 $A_{1-x}B_x$ 的混合能时, 除了考虑温度效应还需要考虑组态熵的贡献, 我们平台采用玻尔兹曼关系式来考虑组态熵的贡献, 如下式所示:

$$S_m = -Nk_B(x \ln x + (1-x) \ln(1-x)). \quad (10)$$

其中 N 和 x 分别是阿伏伽德罗常数和合金组分. 当考虑这一项贡献后, 晶体的自由能改写为

$$F = E + E_{\text{vib}} - TS_{\text{vib}} - TS_m. \quad (11)$$

通过这种方法, 我们成功解释了在描述锆钛二元无序合金时, 长期存在的有关混合能的理论计算和实验结果不相符的问题^[27].

(ii) 力学性质计算. 力学性质计算的核心是通过第一性原理方法得到弹性常数矩阵. 弹性常数在结构材料的设计和表征中占据非常重要的位置, 它是连接应力与应变的桥梁, 而且许多力学量都可以从中提取. 目前计算弹性常数的方法有两种, 一种是能量体积法, 即通过改变体积, 求取相应的能量, 然后根据一定的规则和公式求得弹性常数. 这种方法对于求解高对称体系是非常有效的, 例如对于BCC体系, 弹性常数只有3个独立变量, 因此只需3组能量和体积的变化关系就可以求出弹性常数, 但是对于低对称的体系, 这种方法存在明显的不足, 例如对于单斜晶体, 弹性常数有21个独立变量, 这时采用能量和体积法就非常复杂和繁琐. 另一种方法是应力应变法, 即采用广义胡克定律来计算弹性常数, 如下式所示:

$$T = SC, \quad (12)$$

其中 T 和 S 分别是应力矩阵和应变矩阵, C 为弹性常数矩阵. 不管目标结构的对称性有多低, 最多只需要6组应力应变就可以求得弹性常数矩阵(36个分量). 因为平台采用SQSs技术来研究无序合金, 所以体系的弹性常数不会像有序相那样具有严格的对称性,

所以我们采用应力应变法来计算弹性常数^[28].

1.4 平台应用实例

平台的主要功能是高通量计算无序二元合金的力学性质, 并建立相应的数据库. 为了快速验证平台的合理性和高效性, 我们利用该平台计算了47种金属单质在密排六方(HCP)结构下的弹性常数. 对于每一种单质, 平台单独给出计算报告. 此处以金属镁为例, 给出平台计算报告(表1), 报告的第一行给出体系信息, 即我们的研究对象是什么; 随后给出VASP计算中所使用的重要参数, 分别为截断能(ENCUT)和布里渊区网格点划分密度(KPOINTS); 我们的平台以研究力学性能为切入点, 目前能详细地给出各种力学性质如弹性常数、弹性模量和泊松比等. 这些数据与实验结果符合很好, 如 $C_{11}=64.5$ GPa(实验结果为59.5 GPa), $C_{12}=24.1$ GPa(实验结果为26.1 GPa), $C_{13}=29.3$ GPa(实验结果为21.8 GPa), $C_{33}=64.4$ GPa(实验结果为61.5 GPa)和 $C_{44}=12.9$ GPa (实验结果为16.4 GPa), 理论和实验结果的一致性验证了我们计算平台的可靠性. 该应用一方面说明我们的平台可以用来研究单质体系的力学性质, 并得到与实验相吻合的结果; 另一方面也说明我们的平台实现了高通量计算, 为批量计算提供了一种有效的工具.

2 结论

本文设计了一种可以实现二元合金力学性质高通量计算的方法, 明确地指出了构建相应高通量计算平台的难点所在, 并给出了解决方案. 平台由结构文件、性质计算、数据分析、数据应用和数据库5部分组成. 它能够自动为二元无序合金体系创建结构, 调用第一性原理计算软件VASP进行物性计算, 分析和收集数据并将其保存在数据库中, 为大批量计算提供一种有效的工具. 该高通量计算平台已被国家版权局登记, 登记号为: 2017SR012458. 该工作的主要贡献是为高通量计算和材料基因组计划的实施提供了一种参考方案.

表1 金属镁的平台计算报告

Table 1 Calculation report of Mg

体系	计算参数		力学特性											
	截断能 (eV)	KPOINTS	C_{11} (GPa)	C_{33} (GPa)	C_{12} (GPa)	C_{13} (GPa)	C_{44} (GPa)	B (GPa)	G (GPa)	E (GPa)	B/G	ν	$P(\%)$	A
Mg	273.016	18×18×10	64.487	64.393	24.134	20.290	12.868	35.819	17.275	44.643	2.074	0.292	1.192	0.582

致谢 相关计算在中国科学院合肥物质科学院超算中心、中国科学院计算机网络工作信息中心网格ScGrid以及北京计算科学研究中心(CSRC)天河二号上完成.

参考文献

- 1 Jain A, Ong S P, Hautier G, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *Apl Mater*, 2013, 1: 011002
- 2 Holdren J P. Materials Genome Initiative for Global Competitiveness. Technical Report, NSTC. 2011
- 3 Jarvis D, Raabe D, Singer R, et al. Metallurgy Europe—A Renaissance Programme for 2012–2022. Technical Report, MatSEEC. 2012
- 4 Yamazaki M, Xu Y B. Current status of NIMS structural materials database. *Pres Ves P*, 2010, 6: 1561–1568
- 5 Heng S. Industry 4.0: Upgrading of Germany's Industrial Capabilities on the Horizon. Technical Report, SSRN. 2014
- 6 Xiang X D, Briceno G, Lou Y, et al. A combinatorial approach to materials discovery. *Science*, 1995, 268: 1738–1740
- 7 Chiang Y M, Aydinol M K, Jang Y I, et al. Identification of cathode materials for lithium batteries guided by first-principles calculations. *Nature*, 1998, 392: 694–696
- 8 Curtarolo S, Setyawan W, Hart G L W, et al. An automatic framework for high-throughput materials discovery. *Comp Mater Sci*, 2013, 58: 218–226
- 9 Toher C, Plata J J, Levy O, et al. High-throughput computational screening of thermal conductivity, Debye temperature, and Grüneisen parameter using a quasiharmonic Debye model. *Phys Rev B*, 2014, 90: 174107
- 10 Setyawan W, Gaume R M, Lam S, et al. High-throughput combinatorial database of electronic band structures for inorganic scintillator materials. *ACS Comb Sci*, 2011, 13: 382–390
- 11 Setyawan W, Curtarolo S. High-throughput electronic band structure calculations: Challenges and tools. *Comp Mater Sci*, 2010, 49: 299–312
- 12 Taylor R H, Curtarolo S, Hart G L W. Guiding the experimental discovery of magnesium alloys. *Phys Rev B*, 2011, 84: 1787–1794
- 13 Wang S D, Wang Z, Setyawan W, et al. Assessing the thermoelectric properties of sintered compounds via high-throughput *ab-initio* calculations. *Phys Rev X*, 2011, 1: 1160–1164
- 14 Wang H Z, Wang H, Ding H, et al. Progress in high-throughput materials synthesis and characterization (in Chinese). *Sci Technol Rev*, 2015, 10: 31–49 [王海舟, 汪洪, 丁洪, 等. 材料的高通量制备与表征技术. 科技导报, 2015, 10: 31–49]
- 15 Wang Z, Yang X Y, Zheng Y F, et al. Integrated design of materials and the information platform under the framework of materials genome initiative (in Chinese). *Chin Sci Bull*, 2013, 58: 3733–3744 [王卓, 杨小渝, 郑宇飞, 等. 材料基因组框架下的材料集成设计及信息平台初探. 科学通报, 2013, 58: 3733–3744]
- 16 Yang X, Wang J, Zou F, et al. Towards the integration of data and high-throughput computation for in-silico material design. *SciData-Con*, 2014
- 17 Yang X Y, Lin H Q, Zeng Z, et al. MatCloud-SQS High Throughput Software for Binary Alloys First-principles Calculations Simulation and Design (in Chinese). Software Copyright, 2015SR074470, 2015-05-05 [杨小渝, 林海青, 曾雉, 等. MatCloud-SQS 高通量二元合金第一性原理模拟与设计软件. 软件著作权, 2015SR074470, 2015-05-05]
- 18 Wang X Y, Yang X Y, Chi X B. A computational material database system (in Chinese). *e-Sci Technol Appl*, 2015, 6: 12–17 [王幸阳, 杨小渝, 迟学斌. 一种计算材料数据库系统. 科研信息化技术与应用, 2015, 6: 12–17]
- 19 Wei S H, Ferreira L G, Bernard J E, et al. Electronic properties of random alloys: Special quasirandom structures. *Phys Rev B*, 1990, 42: 9622–9649
- 20 Zunger A, Wei S, Ferreira L, et al. Special quasirandom structures. *Phys Rev Lett*, 1990, 65: 353–356
- 21 Lieser A C, Zacherl C L, Saengdeejing A, et al. First-principles calculations and thermodynamic re-modeling of the Hf-W system. *Calphad*, 2012, 38: 92–99
- 22 Von P J, Dick A, Friák M, et al. Generation and performance of special quasirandom structures for studying the elastic properties of random alloys: Application to Al-Ti. *Phys Rev B*, 2010, 81: 094203
- 23 Shin D, Arróyave R, Liu Z K, et al. Thermodynamic properties of binary hcp solution phases from special quasirandom structures. *Phys Rev B*, 2006, 74: 024204
- 24 Zhou S H, Wang Y, Chen L, et al. Solution-based thermodynamic modeling of the Ni-Ta and Ni-Mo-Ta systems using first-principle calculations. *Calphad*, 2009, 33: 631–641
- 25 Moakher M, Norris A N. The closest elastic tensor of arbitrary symmetry to an elasticity tensor of lower symmetry. *J Elasticity*, 2006, 85: 215–263
- 26 Huang S C, Zhang Y S, Zeng Z, et al. Mechanical properties of zirconium-based random alloys: Alloying elements and composition dependencies. *Comp Mater Sci*, 2017, 127: 60–66
- 27 Huang S C, Zhang X L, Zhi Z, et al. Investigations of the mechanical properties of the Zr_8Ti_8 random alloy. *Int J Mod Phys C*, 2016, 27: 1650076
- 28 Page L Y, Saxe P. Symmetry-general least-squares extraction of elastic data for strained materials from *ab-initio* calculations of stress. *Phys Rev B*, 2002, 65: 104104

Summary for “无序合金高通量计算平台 HCPRA 的搭建及其应用”

Construction of high-throughput computation platform for random alloys (HCPRA) and its applications

LI TingTing^{1,2}, HUANG SunChao¹, ZHANG XiaoLi¹ & ZENG Zhi^{1,2*}

¹ Key Laboratory of Materials Physics, Institute of Solid State Physics, Chinese Academy of Sciences, Hefei 230031, China;

² Science Island Branch of Graduate School, University of Science and Technology of China, Hefei 230026, China

* Corresponding author, E-mail: zzeng@theory.issp.ac.cn

Materials play not only an important role in academic research but also a rather important role in daily life. The discovery of a new material can make a significant contribution to both scientific research and society development. For example, the discovery of graphene in 2004 has attracted enormous attention in academic circle and graphene has shown promising applications in electronic and optical devices. However, the period of discovering a new material is too long to satisfy the demand of society. To speed up the process, Advanced Manufacturing Partnership was proposed in 2011 by US President Obama. This plan has a key part: Materials Genome Initiative (MGI). MGI is made up with three part, (1) construction of high-throughput computation platforms, (2) construction of high-throughput experimental platforms and (3) construction of databases. The high-throughput computation and experimental platforms have promising applications in investigating known materials as well as discovering new materials with high efficiency since they can produce a large amount of data in a short term. Databases can provide convenience for data analysis and machine learning that is a prevailing way to design new materials. In this paper, we present our new designed High-throughput Computational Platform for Random Alloys. The platform is built by using Fortran and Shell languages and its main function is to calculate mechanical and thermodynamic properties of random alloys in large quantities automatically by employing the Vienna *ab-initio* Simulation Package (VASP). There are two main problems that are needed to be solved in the construction of High-throughput Computational Platform for Random Alloys. The first one is the creation of structures for random binary alloys. Another is to set the comparable computational parameters for all calculations automatically. To create structures for binary random alloys, we employ Alloy Theoretic Automated Toolkit to obtain a set of special quasi-random structure templates. Then, we collect all structure information for elementary substance from literature or by performing the first principles calculations. Finally, we design the platform to create special quasi-random structures (POSCAR) for all kinds of binary random alloys by using the templates and structure information of elementary substance combine with the Neumann-Kopp rule. In order to set comparable computational parameters for the VASP calculations, we write a code to read the structure information from POSCAR and use the information to set a comparable k-mesh for all calculations. The INCAR file is created by using the information from POSCAR and POTCAR. After all input files are generated, the platform employs VASP to perform the first-principles calculations. The platform is also designed to extract and analyze calculation results as well as print a calculation report and save all the calculation data into a database. The platform also has some useful functions such as error information-outputting system and error-skipping system. The function is rather important since convergence problem is inevitable in the first-principle calculations especially in high throughput calculations. To test the efficiency of the platform, elastic constants of 47 pure metals are investigated which indicates a computer with eight cores can finish all calculations in one day. And all calculated results match those in the literature well. By using the platform, the mechanical properties of zirconium based binary alloys are investigated and we find the elastic constants and elastic modulus can be well described by a parabolic function.

high-throughput computation platform, random alloys, mechanical properties, elastic constant, thermodynamic properties

doi: 10.1360/N972017-00300