晶体结构预测的新方法和典型应用

吴福伦, 吴顺情*, 朱梓忠*

厦门大学物理科学与技术学院物理学系, 厦门 361005

*联系人, 吴顺情, E-mail: wsq@xmu.edu.cn; 朱梓忠, E-mail: zzhu@xmu.edu.cn

收稿日期: 2020-10-13; 接受日期: 2021-01-15; 网络出版日期: 2021-01-26

国家自然科学基金(编号: 11874307)和国家重点研发计划(编号: 2016YFA0202601, 2016YFB0901502)资助项目

摘要 物质的微观结构直接或间接决定了材料的几乎所有物理和化学性质. 这也是后续材料设计的重要基础. 近年来, 随着多种晶体结构预测方法的发展和改进, 仅根据化学组分就从理论上确定物质结构的研究已经取得了长足进步. 本文概述了当前主流的晶体结构理论预测方法, 并对基于自适应遗传算法(AGA)以及利用“结构单元(Motifs)”(XMsearch)的新型结构预测方法做了重点介绍. 近年来, 基于不同算法的AGA, XMsearch等程序包得到了开发与完善, 只需给定化学组分和压力等条件, 便能对材料的稳态、亚稳态结构进行预测研究. 目前, 这些新工具已在高压材料、电池材料、磁性材料、界面材料等凝聚态物质结构研究中得到了广泛应用.

关键词 结构预测, 自适应遗传算法, 结构单元

PACS: 61.50.-f, 61.66.Fn, 31.15.A-, 34.20.Cf

1 引言

材料结构决定了材料性质, 因而如何研究和预测晶体结构是物理上一个非常重要的课题. 1988年, Maddox (时任*Nature*杂志社主编)在*Nature*上发表了著名的评论: “仅仅根据给定的化学组分来确定物质结构, 即便是最简单的晶体结构, 迄今为止依然无法做到.”^[1] 理论上, 在给定化学组分和外界条件的前提下, 物质的结构组成可能有成千上万种, 而这些结构及其对应的能量构成了一个复杂的势能面. 以 $\text{H}_2\text{O}-\text{HO}_2$ 分子为例(图1), 不同键长、键角分别对应势能面中不同的点, 而稳定存在的分子结构通常是能量最低态, 对应

了势能面的最小值点. 所谓结构预测就是指在全局势能面中搜索最小值点, 并得到其对应的物质结构.

对势能面的探寻不是一件容易的事. 物质势能面高维且复杂, 对于含有 N 个粒子的晶体, 如果将质心固定在原点, 势能面通常有 $3N-3$ 个维度. 势能面中存在许多局域的能量低谷, 每个低谷都有各自的极小值点, 分别代表了具有不同特征的亚稳态结构. 因此, 对物质的结构预测分为了两部分: 在势能面的能量低谷中寻找极小值点, 以及根据极小值点确定最小值点. 然而, 势能面上极小值点的数目随着原子数的增加急剧增长. Tsai和Jordan指出^[2], 即使是最简单的Lennard-Jones体系, 13个原子的团簇, 其势能面也包含了超过一千个

引用格式: 吴福伦, 吴顺情, 朱梓忠. 晶体结构预测的新方法和典型应用. 中国科学: 物理学 力学 天文学, 2021, 51: 030006
Wu F L, Wu S Q, Zhu Z Z. New methods and typical applications of crystal structure prediction (in Chinese). Sci Sin-Phys Mech Astron, 2021, 51: 030006, doi: [10.1360/SSPMA-2020-0405](https://doi.org/10.1360/SSPMA-2020-0405)

图 1 (网络版彩图) $\text{H}_2\text{O}-\text{HO}_2$ 分子二维势能面示意图. 不同键长、键角分别对应势能面中不同点, 亚稳态结构对应势能面各个谷底. 摘自文献[3], 已获得授权

Figure 1 (Color online) Two-dimensional potential energy surface of $\text{H}_2\text{O}-\text{HO}_2$. Different bond lengths and bond angles correspond to different points in the potential energy surface. Metastable structures correspond to the valleys of the potential energy surface. Reprinted figure with permission from ref. [3]. Copyright (2006) by the AAAS.

能量极小值点; Wales和Doye的研究表明^[4], 当团簇的原子数增加到55个时, 能量极小值的数目将超过 10^{12} 个. 显然, 遍历所有极小值点来确定最小值是无法做到的. 人们转而寻求其他能够进行全局搜索的结构预测算法.

2 结构预测理论方法的发展

受限于计算资源的不足, 直至20世纪末, 结构预测的算法仍然没有明显进步. 到了21世纪, 随着计算机计算能力的大幅提升, 多种晶体结构预测方法得到了迅速发展和改进, 仅根据化学组分就从理论上确定物质结构的研究取得了长足的进步. 按照是否需要初始结构, 结构预测方法可分为两类: 一是盆地跳跃(Basin Hopping)、最小值跳跃(Minima Hopping)、模拟退火(Simulated Annealing)、巨动力学(Metadynamics)等依赖初始结构的方法; 二是随机搜索、粒子群优化、遗传算法等全局搜索算法. 基于以上算法也发展出了多种程序包, 解决了大量物理学、材料学领域的难题.

结构预测需要在势能面的能量低谷中寻找极小值点, 以及根据极小值点确定最小值点. 其中, 寻找各个

极小值点的过程涉及势垒跃迁的问题, 以避免一直困在能量低谷中. 第一类方法主要着眼于势垒跃迁的能力. 盆地跳跃方法^[4]是局域极小化方法与蒙特卡罗算法的结合, 主要包含以下流程: (1) 对初始结构做局部极小化处理, 得到能量极小值; (2) 随机扰动后产生新的结构, 做同样处理后得到新的极小值; (3) 根据蒙特卡罗方法中的Metropolis准则来决定是否接受产生的新结构; (4) 不断循环以上过程以探索势能面未知区域. 此方法通过引入局域极小化方法把复杂的势能曲面投影成了阶梯形, 大大提升了蒙特卡罗算法的势垒跃迁能力. 最小值跳跃方法^[5]与此类似, 同样引入了局域极小化方法以得到能量极小值. 不同的是最小值跳跃法不依赖体系的热力学性质, 主要依靠分子动力学模拟进行新结构的演化迭代. 通过不断调整温度, 该算法能够改变势垒跃迁能力来进行势能面的探索. 此外, 通过将当前极小值与寻找过程中的历史轨迹相比较, 可避免重复探索已访问过的势能面. 模拟退火方法^[6]中的“退火”来源于冶金学, 意为将材料加热后再按照特定速率冷却的过程. 材料中的原子通过加热离开原先位置, 退火时冷却比较慢, 原子有更大的可能性落在比原先能量更低的位置上. 在结构预测中, 该方法可利用分子动力学模拟, 经由加热越过势垒, 再通过缓慢降温最终使体系达到能量最低的稳定态. 模拟退火算法具有渐近收敛性, 理论上已经证明该算法能够依概率收敛到全局最优解. 巨动力学方法^[7]是一种依赖于所定义反应坐标、能够有效重新构建体系势能面并加速采样的方法. 它通过在模拟过程中每隔一段时间给体系加入一定的高斯势, 使得势能面被逐渐“填平”, 从而使体系能够从各个局域极小值内“逃离”出来. 从而, 能够在较短时间尺度内实现对体系的全局采样, 并获得整个势能面的信息.

不依赖初始结构的结构预测算法主要有随机搜索、粒子群优化方法、遗传算法等三种. 随机搜索(Random Search)^[8]最为简单, 核心思想就是“纯粹的随机”. 该方法完全随机地产生大量结构, 通过局域极小化得到各自的能量极小值, 以期对比后获得全局的最小值. 随机搜索尽管原理十分简单, 但太过于依赖随机性, 较大的体系往往不会选择此算法. 粒子群优化算法(Particle Swarm Optimization)^[9]源于对简化社会模型的模拟, 最初用于模拟鸟群的捕食行为. 若鸟群在空间中搜索唯一的一块食物, 且所有的鸟仅知道

自身的位置离食物有多远, 则找到食物的最优策略就是整个鸟群都来搜索距食物最近的鸟的周围空间. 粒子群优化算法中的每只鸟都代表一个可能解, 称为一个“粒子”, 所有粒子跟随当前的最优粒子对势能面进行探索. 该算法收敛速度较快, 但可能会陷入局域极小值无法逃脱, 如何克服此缺陷一直是研究的热点. 遗传算法(Genetic Algorithm/Evolutionary Algorithm)^[10]是当前使用最广泛的全局优化算法之一. 此算法源于生物学中的进化论和遗传学说: 进化从随机产生或指定的种群开始, 一代一代继续下去, 直到获得最优的后代为止. 遗传的核心在于基因的继承和突变. 在结构预测中, 继承就是指下一代结构保留了上一代的部分特征, 而突变指对新结构的剪切、重塑、对称化等操作, 以增加整个体系的多样性. 通过将每次得到的最优后代投入下一次遗传, 就可以期望于 N 代后得到能量最低的稳定结构. 此外, 使用遗传算法进行结构预测需要处理大量的备选结构, 若对每一个备选结构都进行第一性原理计算会耗费大量的计算资源. 为加快运算速度, 我们一般只选取经验势来计算能量. 但由于晶体内部原子众多, 它们的势函数难以描述, 对整个晶体能量的准确评估就变得困难起来——而能量评估的准确性会直接影响预测结果. 因此, 为了克服结构预测中的经验势计算“快而不准”以及第一性原理计算“准而不快”的不足, 我们将遗传算法与第一性原理计算相结合, 提出并发展了自适应遗传算法(Adaptive Genetic Algorithm, AGA), 并开发了拥有自主知识产权的AGA结构搜索程序包^[11]. 该程序包仅需输入化学组分及比例, 就可进行二维表面与晶界、层状材料、三维晶体等的结构预测, 也取得了很多成果.

除此之外, 作为应用潜力最明显的工具之一, 机器学习在结构预测领域也变得日益重要. 我们知道, 机器学习算法可以在已知数据中寻找规律, 然后利用这些规律进行预测. 在材料结构预测中, 可以通过对材料数据库的大量学习, 预测出推荐的成分和结构以得到候选材料, 最后通过第一性原理计算对其稳定性进行检验. 比如Hautier等人^[12]通过对ICSD数据库(无机晶体结构数据库)中的183种常见的氧化物进行学习, 获得结构中原子分布规律的模型, 并以此预测得到了209种新的三元氧化物. 当然, 机器学习技术还可以与各种经典的结构搜索算法结合, 如模拟退火法、遗传算法、粒子群优化算法等, 通过构造基于神经网络的

原子间相互作用势函数, 以提高材料预测的效率. 前文提到, 对晶体能量评估的准确度会直接影响结构预测结果, 良好的机器学习势函数能够使经典结构预测算法的效果得到很大提升.

总体上, 本文对过去三十年来结构预测理论的研究进展做了简要回顾, 主要侧重于相关算法的发展, 特别是厦门大学计算凝聚态物理研究团队在结构预测方面的相关研究工作.

3 基于遗传算法的AGA结构预测方法

AGA始于经验势函数和遗传算法, 通过后者不断循环产生一定数量的较优结构. 随后对部分较优结构进行第一性原理计算, 并基于产生的力与能量等数据拟合出新的经验势函数. 其中, 拟合过程可利用pot-fit程序包^[13]的力场拟合算法. 最后将拟合后的新经验势传递给遗传算法, 进行下一次循环. 如此通过不断循环AGA最终实现对复杂势能面的探索以及新结构的不断演化.

AGA方法的流程如图2所示, 左侧就是遗传算法循环. 经过结构的继承、突变和选择, 遗传算法不断产生新结构, 并选取其中最优的子代结构进入循环. 在传统的遗传算法中, 最耗费计算资源的部分在于对子代结构的局部优化过程, 此时需要进行非常耗时的第一性原理计算. 一些复杂体系的收敛甚至要超过

图 2 (网络版彩图)自适应遗传算法流程图^[11]

Figure 2 (Color online) Flowchart of the adaptive genetic algorithm [11].

200次循环. 而在AGA方法中, 第一性原理计算被基于经验势的单点能计算所取代, 减少了五至六个数量级的计算量. 当然, 经验势计算的准确度不能与第一性原理计算相比, 因而在下一步需要通过potfit程序包对经验势进行拟合修正. 图2右侧即为势函数拟合过程, 选取第一步得到的少量较优结构进行第一性原理计算, 产生了一系列力、能量、压强等数据. 使用这些数据和基于模拟退火的potfit力场拟合算法, 经验势函数得到了拟合修正, 从而更加贴合势能面. 生成的新经验势可用于下一次遗传算法循环, 不断重复该过程就完成了AGA的迭代. 最后, 通过筛选每一代循环收集的亚稳态结构, 就能够得到最终的低能量稳态结构. 一般来说, AGA结构搜索需要30–50次迭代才能得到最终结构. 与传统的遗传算法相比, AGA具有明显优势. 整个AGA算法的净计算时间同比可减少三个数量级以上, 经验势的不断拟合修正也使其能够更有效地探索未知势能面. 也就是说, AGA结构搜索可同时具有经验势计算的效率和第一性原理计算的准确度.

将经验势函数与第一性原理计算相结合以进行结构优化的思想并非没有先例. 传统思路是利用一个较好的经验势从全势能面中筛选出所有候选结构, 再通过第一性原理计算进行细化处理. 但是该方法要求使用的经验势对整个势能面描述都比较准确, 这是非常困难的. 尤其是对于三元和多元系统, 想要得到一个能够准确描述各种结合环境的经验势基本是不可能实现的. 而在AGA方法中, 经过第一性原理计算修正后的经验势仅要求对一部分局域能谷的描述比较准确. 由于每次迭代时势函数都会得到修正, 从而AGA方法能够以不同的势函数对势能面不同的能谷进行采样. 尽管每个修正后的势函数仅能准确描述一部分能谷, 但通过不断迭代, AGA就实现了对整个势能面能谷的探索. 关于这一点的一个很好例证就是利用嵌入原子势(EAM)进行TiO₂晶体结构搜索的过程^[11]. AGA并不期望拟合的EAM势函数能够很好描述各种TiO₂晶体结构的能量, 但通过25次迭代, AGA最终能够找到TiO₂的两个低能量基态结构——金红石和锐钛矿结构.

AGA方法的有效性已通过多种实验结构的搜索测试得到了验证. 自2014年AGA程序包发布以来, 国内外研究人员利用它解决了很多物理、化学、地球物理、材料科学等领域的难题.

3.1 高压物理

在天体物理领域, 天王星、海王星等冰巨星以及柯伯依带天体和大量系外行星上均贮有丰富的固态水, 引起了人们的注意. 作为行星的重要组成部分, 固态水的结构与性质一直是研究的热门. 已有的研究预测固态水在0 K、压强增大时的相变顺序为*Pbcm*→*Pbca*→*Cmcm*相^[14], 而通过AGA预测得到的*Pmc2₁*, *P2₁*和*P2₁/c*相在对应压强下均要比*Cmcm*相稳定^[15]. 亦即, H₂O的高压相变顺序应为*Pbcm*→*Pbca*→*Pmc2₁*→*P2₁*→*P2₁/c*相(图3). 后续研究发现, 这些预测得到的低温高压相并没有显著证据存在于天王星、海王星上. 研究人员猜测, 这些相可能与系外行星有关.

作为行星内部的重要组成部分, SiO₂也是人们研究的重点. 已有研究表明, 在10 Mbar条件下, 后钙钛矿MgSiO₃将会分解为体心立方结构的MgO和钙铝石型SiO₂^[16]. 而AGA结构预测得到了一种低温下比钙铝石型SiO₂更为稳定的Fe₂P型结构^[11], 对后钙钛矿MgSiO₃的重组与分解有了新的启发. 对于(MgO)_m(SiO₂)_n (简称M_mS_n)在高压下的重组与分解, 此前的研究给出了机理^[17]:

$$2\text{MS(后钙钛矿)} \rightarrow \text{M(体心立方)} + \text{MS}_2(\text{P2}_1/\text{c型}) \rightarrow 2\text{M(体心立方)} + 2\text{S(Fe}_2\text{P型)}.$$

通过AGA结构搜索, 研究人员得到了一个新的稳定相 $\sqrt{4}2d$ 型Mg₂SiO₄ (M₂S). 从而, 我们认为, 后钙钛矿的分解可分为三步^[18,19]:

图3 (网络版彩图)冰的超高压相结构. 蓝色和红色大球表示氧原子, 白色小球表示氢原子. 摘自文献[15], 已获得授权
Figure 3 (Color online) Crystal structures of ultrahigh-pressure phases of ice. Blue and red large spheres denote oxygen atoms. White small spheres denote hydrogen atoms. Reprinted figure with permission from ref. [15]. Copyright (2011) by the American Physical Society.

$2\text{MS}(\text{后钙钛矿}) \rightarrow \text{M}_2\text{S}(\text{新相}) \rightarrow \text{M}(\text{体心立方})$
 $+ \text{MS}_2(P2_1/c \text{型}) \rightarrow 2\text{M}(\text{体心立方}) + 2\text{S}(\text{Fe}_2\text{P型}).$

3.2 电池材料

随着资源和环境问题的日益突出, 发展新能源已成为人类社会发展的迫切需求. 钠元素作为地壳中含量最高的元素, 也是第二轻的碱元素, 引起了人们对钠离子电池的关注. 结合AGA算法和第一性原理计算, Wu等人^[20]对 $\text{Na}_2\text{FeSiO}_4$ 进行了全面的结构搜索, 最终在与基态相比0.5 eV/f.u. (分子式)的能量窗口内获得了84种不同的 $\text{Na}_2\text{FeSiO}_4$ 低能量结构. 对于电池正极材料而言, 充放电过程中保持不变的结构框架至关重要. 基于搜索得到的多种低能量结构, 他们进一步分析了其中的不同Fe-Si网络类型, 以及去除钠离子后对稳定性的影响.

3.3 磁性材料

近年来, 各产业对永磁体材料的需求不断增长. 目前应用最广泛的永磁体材料为 $\text{Nd}_2\text{Fe}_{14}\text{B}$ 和 SmCo_5 , 两种材料都包含了稀土元素. 然而稀土矿产资源十分有限, 人们转而寻找能够替代稀土元素永磁体的材料. 除了取决于微观结构, 永磁体的性能还取决于晶体结构的固有磁性, 例如饱和磁化强度、居里温度和磁晶各向异性等. 此时, 使用不同的加工路线制备的 $\text{Zr}_2\text{Co}_{11}$ 及其相关相引起了人们的关注. 研究表明, 由快速淬火而生成的 $\text{Zr}_2\text{Co}_{11}$ 亚稳态相在居里温度达到500°C时表现出很强的磁晶各向异性^[21,22]. 但是实验样品中多种相的存在以及晶粒尺寸的不同使得很难通过实验确定该相的结构, 极大阻碍了该材料在实际应用中的进一步开发. 为了解决这一问题, 研究人员利用AGA方法对 ZrCo_{5+x} ($x=0.0, 0.1, 0.2, 0.25, 0.5$)等组分材料进行了全面的结构搜索^[23], 最终确认了多种Zr-Co体系化合物的成分和晶体结构. 相似的研究还有对 Co_3Si 的分析, 实验研究表明以纳米颗粒形式制备的 Co_3Si 材料具有很高的矫顽力和相应的高饱和磁化强度^[24], 但这种纳米粒子相的结构及形成机理一直存在争议. 针对这一问题, 研究人员利用AGA方法对 Co_3Si 进行了结构搜索, 最终获得了一个新的 Co_3Si 结构($o\text{-Co}_3\text{Si}$)^[25]. 该结构具有正交对称性, 且第一性原理计算证明该相在0 K下比原有的 $h\text{-Co}_3\text{Si}$ 相更加稳定, 从而更新了人们对 Co_3Si 形成机理的认识.

在无稀土元素的永磁体中, 由于Fe的成本低且 $\alpha''\text{-Fe}_{16}\text{N}_2$ 薄膜具有较高的磁化强度^[26], Fe, N化合物中的四方亚稳相 $\alpha''\text{-Fe}_{16}\text{N}_2$ 材料吸引了大量的实验和理论关注. Ke等人^[27]的研究表明, 在 $\alpha''\text{-Fe}_{16}\text{N}_2$ 相中用Co或Ti代替少量Fe元素(不超过15%)可显著改善材料的磁晶各向异性. 这种掺杂是否会对晶体结构和相稳定性产生影响? 为了研究这一问题, 研究人员利用AGA方法对 $\text{Fe}_{16-x}\text{Co}_x\text{N}_2$ ($x=0, 1, \dots, 16$)等组分的材料进行了全面的结构搜索^[28], 最终确认了随着Co浓度的增加 $\text{Fe}_{16-x}\text{Co}_x\text{N}_2$ 会发生结构转变. 随着立方 Co_{16}N_2 结构转变为四方 Fe_{16}N_2 结构, 系统的磁矩不断增加, 但是磁晶各向异性在 $\text{Fe}_{12}\text{Co}_4\text{N}_2$ 成分处达到最大值. 类似的材料还有Fe-Co合金, 其自身具有较高的磁化强度, 并且可通过使晶胞应变(改变 c/a 比)的方式得到非常高的磁晶各向异性^[29]. 但是迄今为止, 四方畸变的Fe-Co合金仅能稳定存在于少数的薄层材料中. 为了得到更加稳定的四方畸变晶相结构, 有研究提出可以通过掺入B, C, N等轻元素来增加其稳定性^[30]. 针对这个思路, 研究人员通过AGA晶体结构预测以及第一性原理计算, 系统研究了 $\text{Fe}_{3-x}\text{Co}_x\text{C}$ 等组分的结构、磁矩和磁晶各向异性^[31]. 除了验证已知的 Fe_3C ($Pnma$)结构之外, AGA搜索还得到了几个室温下的新亚稳相, 其中贝氏体($P6_322$)结构在所有相中表现出了最大的磁矩, 并且每原子的能量仅比 $Pnma$ 相高了4 meV.

3.4 界面结构

了解固-固界面原子结构的信息对理解和预测材料性能至关重要. 由于界面的复杂性, 通过实验或计算机模拟来解析其原子结构十分困难. 为了解决这一问题, 研究人员开发了基于AGA方法的界面结构预测算法. 这种算法在保留第一性原理计算准确性的同时, 显著降低了界面结构搜索的计算成本. 一个典型的应用是对化学计量及非化学计量比 $\text{SrTiO}_3\sqrt{3}(112)[\bar{1}10]$ 晶界的研究^[32], 该体系包含了超过200个原子. 利用该算法, 研究人员发现了几种新颖的非化学计量比的边界基态结构(图4), 对 SrTiO_3 晶界稳定性有了新的认识. 研究表明, 在-0.72~-0.15 eV的化学势范围内, 化学计量比的边界是稳定的; 在该范围之外, 非化学计量比的晶界结构能量更低.

另一典型应用是对氧化物表面缺陷的研究. 氧化物表面的缺陷很大程度上影响了表面活性, 引起了人

图 4 (网络版彩图)利用AGA进行界面结构搜索示意图. 右图为预测得到的SrTiO₃晶界. 摘自文献[32], 已获得授权

Figure 4 (Color online) Interface structure prediction by adaptive genetic algorithm. Picture on the right shows the predicted SrTiO₃ grain boundary. Reprinted figure with permission from ref. [32]. Copyright (2011) by the American Chemical Society.

们对其原子构型和电子结构的重视. 锐钛矿型TiO₂ (001)-(1×4)表面缺陷就具有很高的活性, 但由于该体系表面发生了重构, 其缺陷在不同实验中表现出了不同的特征, 从而原子结构的确定异常困难. 针对这一问题, AGA结构预测算法成为了很好的工具. 基于该算法, 研究人员对锐钛矿型的TiO₂ (001)-(1×4)表面缺陷进行了系统的第一性原理分析^[33], 针对不同的Ti-O比例确定了低能量的缺陷构型, 并绘制了其在实际条件下能够稳定存在的区域. 该研究对洞悉表面缺陷的几何构型和电子结构提供了新的思路. 此外, 他们还通过AGA方法搜索得到了新的氧空位及钛间隙结构, 在局域电荷分布、磁性以及扫描隧道显微镜的明暗图像特征等方面都与常规结构不同.

4 构建有效结构: Motif网络原理和应用

要进行晶体结构预测往往需要处理大量的备选结构, 这就对搜索效率提出了很高的要求. 多数结构预测算法都选择随机生成初始结构, 这种无约束的纯随机容易引入如原子过于聚集等的“低质量结构”, 造成了计算资源的浪费. 此外, 对于多元复杂体系, 由于体系的自由度随着原子数目迅速增加, 计算量也会大大上升. 能否根据材料体系的结构特征来提高搜索效率呢? 基于这个想法, Zhu等人^[34]开发了一种基于材料中“结构单元(Motifs)”的结构预测算法, 并由此开发了XMsearch结构搜索程序包.

该方法的核心思想是将材料体系中普遍存在的结

构特征, 如锂离子电池正极材料中普遍存在的四面体或八面体, 视为一个不变的“结构单元”. 这些“结构单元”可以在空间中平移或旋转, 但不会破坏其整体构型. 由此, 整个系统将由“结构单元”和余下原子两部分构成. 放入晶胞后的“结构单元”可以通过 (R, θ, ϕ) 五个初始自由度来描述(分别是描述重心位置所需的3个自由度和描述方位角所需的2个自由度). “结构单元”所含原子越多, 确定其初始位置所需参数减少也就越显著, 从而有效提升了随机结构搜索的效率. 我们将该方法应用于锂离子电池正极材料LiFePO₄的结构搜索^[34]中, 成功得到了Materials Project Database^[35]中的几种已有结构, 并预测得到了几种新型亚稳结构. 另一个典型的例子是对硅酸盐正极材料(A₂MSiO₄, A=Li, Na; M=Mn, Fe, Co)新结构的预测. 大量实验结果显示, 这种体系的材料普遍为四面体结构, 且拥有丰富的结构为四面体的同素异形体. 从而, 四面体可作为该材料体系中的“结构单元”. 基于这种方法, Zhao等人^[36]对A₂MSiO₄晶体结构进行了搜索. 他们首先确定了四面体“结构单元”中O原子的位置, 随后将其四个最近邻原子随机指定为两个A原子、一个M原子和一个Si原子; 继而, 将A, M, Si原子的最近邻指定为O原子. 以该形式不断迭代, 直至四面体结构中的所有位置都被合理占据(图5). 如此产生各原子分布后, 再通过第一性原理计算优化得到此结构. 研究结果显示, 该方法能够以很低的计算成本找到已有实验或理论报道的所有结构. 在此之外, 该方法同时搜索得到了一些能量相

图 5 (网络版彩图)四面体结构产生的示意图^[36]. 从中心氧原子出发, 四个最近邻原子随机指定为两个A原子、一个M原子及一个Si原子; 不断迭代直至四面体结构中的所有位置均被合理占据

Figure 5 (Color online) Schematic drawing of a tetrahedral structure^[36]. Starting from a central oxygen atom, the four nearest neighbors are randomly designated as two A atoms, one M atom, and one Si atom. Iterate until all positions in the tetrahedral structure are reasonably occupied.

近、甚至更加稳定的新结构, 为锂离子电池正极材料的合成与制备提供了新的信息. 当然, 除了电极材料, 该方法也能够应用于相似的具有结构特征的多元复杂体系, 并以此进行复杂晶体结构的广泛搜索. 值得一提的是, 该方法可作为AGA中生成初始结构方法的补充, 进一步提升搜索效率.

5 晶体结构相似性的实空间对比算法

由于晶体构造非常复杂、搜索空间巨大等因素的影响, 结构预测的效率与过程中的结构去重操作(去除重复的结构)有着密切的联系. Xtalopt结构预测软件^[37]单独开发了XtalComp重复性识别算法, 用于识别晶体的底层结构是否重复. 如今应用比较广泛的结构重复性识别算法有两种: 第一种是将两个结构的中心重合, 通过大量的随机旋转来确认相似性, 操作简单但是效率很低; 第二种通过列出及对比原子间距来分析相似结构, 效率虽高但是准确性不足. 尤其是对于晶体结构预测过程中生成的大量相似形变的备选结构, 该方法准确性偏低. 因此, 开发一种高效、准确的结构相似性分析算法可大大加快结构预测过程中的去重操作, 提升结构预测的效率. 与主流算法基于核心原子间距或基于晶体六个基本参数的想法不同, 我们以选取的中心原子和两个最近邻原子构成的面作为基准, 通过旋转将两个待对比结构的基准面重叠. 重叠后, 计算对应原子空间位置的偏差即可判断晶体结构的相似性, 在准确度和效率上均有较大提升. 目前, 该算法已实现对原子数目相同的结构对比, 正着力于开发处理原子缺失、换位、晶胞缩放等特殊情形的方法. 该方法作为提高搜索效率的模块加入到AGA程序中.

6 总结与展望

仅仅基于材料化学组分的晶体结构预测, 长期以来都是固体物理和材料科学等领域的一大挑战. 本文对两类预测算法做了系统介绍: 一是依赖初始结构的盆地跳跃、最小值跳跃、模拟退火、巨动力学等方法; 二是随机搜索、粒子群优化、遗传算法等全局搜索算法. 总结来说, 晶体结构预测方法的开发, 一方面由于搜索空间巨大, 需要提高探索相空间的效率, 另一方面也需要能够快速、准确计算结构能量的原子间相互作用势. 基于这个想法, 文章重点介绍了一种自适应遗传算法(AGA)的晶体结构预测方法, 并简述了其在高压物理、电池材料、磁性材料、界面结构等领域的应用. 此外, 体系材料中普遍存在的结构特征, 如锂离子电池正极材料中普遍存在的四面体或八面体, 可被视为一个不变的“结构单元”. 基于这个想法, 我们开发了一种基于材料中“结构单元”的结构预测算法, 并以此进行了复杂晶体结构的广泛搜索, 提升了搜索效率. 最后, 本文概述了一种基于基准面的结构重复性识别算法, 在准确度和效率上均有较大提升.

尽管各种结构预测算法已经广泛应用于物理、化学、地球物理、材料科学等多个领域, 并解决了很多科学难题, 但是这些方法的发展仍然面临着许多挑战. 例如, 现有结构预测算法大多只适用于不超过一百个原子的简单体系, 对于大尺寸多元复杂体系的预测仍然十分吃力; 鲜有算法能够进行超高温、超高压等极限环境下的复杂材料的结构预测等. 但是随着计算机计算能力的不断提升和全局优化算法的不断更新发展, 我们相信结构预测方法必将大有可为, 可在更多领域发挥重要作用.

参考文献

- 1 Maddox J. Crystals from first principles. *Nature*, 1988, 335: 201
- 2 Tsai C J, Jordan K D. Use of an eigenmode method to locate the stationary points on the potential energy surfaces of selected argon and water clusters. *J Phys Chem*, 1993, 97: 11227–11237
- 3 Suma K, Sumiyoshi Y, Endo Y. The rotational spectrum of the water-hydroperoxy radical ($\text{H}_2\text{O}-\text{HO}_2$) complex. *Science*, 2006, 311: 1278–1281
- 4 Wales D J, Doye J P K. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J Phys Chem A*, 1997, 101: 5111–5116
- 5 Goedecker S. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *J Chem Phys*, 2004, 120: 9911–9917
- 6 Pannetier J, Bassas-Alsina J, Rodriguez-Carvajal J, et al. Prediction of crystal structures from crystal chemistry rules by simulated annealing.

- Nature, 1990, 346: 343–345
- 7 Martonák R, Laio A, Parrinello M. Predicting crystal structures: The Parrinello-Rahman method revisited. *Phys Rev Lett*, 2003, 90: 075503, arXiv: [cond-mat/0211551](#)
- 8 Pickard C J, Needs R J. *Ab initio* random structure searching. *J Phys-Condens Matter*, 2011, 23: 053201, arXiv: [1101.3987](#)
- 9 Wang Y, Lv J, Zhu L, et al. CALYPSO: A method for crystal structure prediction. *Comput Phys Commun*, 2012, 183: 2063–2070, arXiv: [1205.2264](#)
- 10 Deaven D M, Ho K M. Molecular geometry optimization with a genetic algorithm. *Phys Rev Lett*, 1995, 75: 288–291, arXiv: [mtrl-th/9506004](#)
- 11 Wu S Q, Ji M, Wang C Z, et al. An adaptive genetic algorithm for crystal structure prediction. *J Phys-Condens Matter*, 2014, 26: 035402
- 12 Hautier G, Fischer C C, Jain A, et al. Finding nature's missing ternary oxide compounds using machine learning and density functional theory. *Chem Mater*, 2010, 22: 3762–3767
- 13 Brommer P, Gähler F. Potfit: Effective potentials from *ab initio* data. *Model Simul Mater Sci Eng*, 2007, 15: 295–304, arXiv: [0704.0185](#)
- 14 Militzer B, Wilson H F. New phases of water ice predicted at megabar pressures. *Phys Rev Lett*, 2010, 105: 195701, arXiv: [1009.4722](#)
- 15 Ji M, Umemoto K, Wang C Z, et al. Ultrahigh-pressure phases of H₂O ice predicted using an adaptive genetic algorithm. *Phys Rev B*, 2011, 84: 220105, arXiv: [1108.4164](#)
- 16 Umemoto K, Wentzcovitch R M, Allen P B. Dissociation of MgSiO₃ in the cores of gas giants and terrestrial exoplanets. *Science*, 2006, 311: 983–986
- 17 Umemoto K, Wentzcovitch R M. Two-stage dissociation in MgSiO₃ post-perovskite. *Earth Planet Sci Lett*, 2011, 311: 225–229
- 18 Umemoto K, Wentzcovitch R M, Wu S, et al. Phase transitions in MgSiO₃ post-perovskite in super-Earth mantles. *Earth Planet Sci Lett*, 2017, 478: 40–45, arXiv: [1708.04767](#)
- 19 Wu S, Umemoto K, Ji M, et al. Identification of post-pyrite phase transitions in SiO₂ by a genetic algorithm. *Phys Rev B*, 2011, 83: 184102
- 20 Wu P, Wu S Q, Lv X, et al. Fe-Si networks in Na₂FeSiO₄ cathode materials. *Phys Chem Chem Phys*, 2016, 18: 23916–23922
- 21 Demczyk B G, Cheng S F. Structures of Zr₂Co₁₁ and HfCo₇ intermetallic compounds. *J Appl Crystlogr*, 1991, 24: 1023–1026
- 22 Gabay A M, Zhang Y, Hadjipanayis G C. Cobalt-rich magnetic phases in Zr-Co alloys. *J Magn Magn Mater*, 2001, 236: 37–41
- 23 Zhao X, Nguyen M C, Zhang W Y, et al. Exploring the structural complexity of intermetallic compounds by an adaptive genetic algorithm. *Phys Rev Lett*, 2014, 112: 045502
- 24 Balasubramanian B, Manchanda P, Skomski R, et al. High-coercivity magnetism in nanostructures with strong easy-plane anisotropy. *Appl Phys Lett*, 2016, 108: 152406
- 25 Zhao X, Yu S, Wu S, et al. Structures, phase transitions, and magnetic properties of Co₃Si from first-principles calculations. *Phys Rev B*, 2017, 96: 024422, arXiv: [1704.04514](#)
- 26 Kim T K, Takahashi M. New magnetic material having ultrahigh magnetic moment. *Appl Phys Lett*, 1972, 20: 492–494
- 27 Ke L, Belashchenko K D, van Schilfgaarde M, et al. Effects of alloying and strain on the magnetic properties of Fe₁₆N₂. *Phys Rev B*, 2013, 88: 024404, arXiv: [1304.2428](#)
- 28 Zhao X, Wang C Z, Yao Y, et al. Large magnetic anisotropy predicted for rare-earth-free Fe_{16-x}Co_xN₂ alloys. *Phys Rev B*, 2016, 94: 224424, arXiv: [1610.03544](#)
- 29 Burkert T, Nordström L, Eriksson O, et al. Giant magnetic anisotropy in tetragonal FeCo alloys. *Phys Rev Lett*, 2004, 93: 027203
- 30 Choe H, Terai T, Fukuda T, et al. Easy axis of magnetization of Fe₃C prepared by an electrolytic extraction method. *J Magn Magn Mater*, 2016, 417: 1–5
- 31 Wu S Q, Balamurugan B, Zhao X, et al. Exploring new phases of Fe_{3-x}Co_xC for rare-earth-free magnets. *J Phys D-Appl Phys*, 2017, 50: 215005
- 32 Zhao X, Shu Q, Nguyen M C, et al. Interface structure prediction from first-principles. *J Phys Chem C*, 2014, 118: 9524–9530
- 33 Shi Y, Sun H, Nguyen M C, et al. Structures of defects on anatase TiO₂ (001) surfaces. *Nanoscale*, 2017, 9: 11553–11565
- 34 Zhu Z, Wu P, Wu S, et al. An efficient scheme for crystal structure prediction based on structural motifs. *J Phys Chem C*, 2017, 121: 11891–11896
- 35 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
- 36 Zhao X, Wu S, Lv X, et al. Exploration of tetrahedral structures in silicate cathodes using a motif-network scheme. *Sci Rep*, 2015, 5: 15555, arXiv: [1504.02070](#)

- 37 Lonie D C, Zurek E. XtalOpt: An open-source evolutionary algorithm for crystal structure prediction. [Comput Phys Commun](#), 2011, 182: 372–387

New methods and typical applications of crystal structure prediction

WU FuLun, WU ShunQing^{*} & ZHU ZiZhong^{*}

Department of Physics, College of Physical Science and Technology, Xiamen University, Xiamen 361005, China

The microstructure of a material considerably influences its physical and chemical properties; this observation forms the basis of subsequent material design. Recently, the study of crystal structure prediction beginning from the chemical composition has progressed considerably with the development and improvement of many crystal structure prediction methods. Here, we present an overview of the currently used theoretical crystal structure prediction methods and new crystal structure prediction methods based on adaptive genetic algorithm (AGA) and motifs (XMsearch). Further, we introduce AGA and XMsearch along with their applications. These new tools have been widely used in the structural studies of condensed matter such as high-pressure materials, battery materials, magnetic materials, and interfacial systems.

structure prediction, adaptive genetic algorithm, motifs

PACS: 61.50.-f, 61.66.Fn, 31.15.A-, 34.20.Cf

doi: [10.1360/SSPMA-2020-0405](https://doi.org/10.1360/SSPMA-2020-0405)