

引用格式: 杨小渝, 郝德博, 舒城, 等. MatCloud-QE: 基于云原生理念的高通量第一性原理计算程序包[J]. 中国材料进展, 2024, 43(11): 1007-1015.

YANG X Y, HAO D B, SHU C, et al. MatCloud-QE: A Cloud-Native High-Throughput First-Principles Simulation Software Package[J]. Materials China, 2024, 43(11): 1007-1015.

# MatCloud-QE: 基于云原生理念的高通量 第一性原理计算程序包

杨小渝<sup>1, 2, 3</sup>, 郝德博<sup>3</sup>, 舒城<sup>3</sup>, 许立芳<sup>3</sup>, 梁亚男<sup>3</sup>, 马新杰<sup>3</sup>

(1. 中国科学院计算机网络信息中心, 北京 100190)

(2. 中国科学院大学, 北京 100049)

(3. 北京迈高材云科技有限公司, 北京 100190)

**摘要:** 国内开展第一性原理计算总是局限于某一两个商业软件, 不仅需要软件使用许可, 而且还需要下载、安装到本地, 再通过人工部署到高性能计算集群才能开展使用, 极大地限制了科研人员开展第一性原理计算。为了解决这个问题, 作者团队基于 Quantum ESPRESSO 第一性原理计算内核, 进一步开发了高通量第一性原理计算加速算法、参数智能推荐算法、输入结构统一算法、自动后处理引擎、组件化引擎、图形化引擎、更多物性计算功能以及 MatCloud+ 接口等算法程序, 形成了具有云原生特点的第一性原理计算程序包: MatCloud-QE; 进一步通过接入 MatCloud+ 材料云, 使用户仅通过浏览器就可在线开展第一性原理计算和数据自动化采集和管理, 极大地方便了用户开展第一性原理计算以及高通量计算筛选, 且无需软件使用许可。以面向太阳能电池应用的无机钙钛矿  $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  的材料特性的第一性原理计算为例, 说明通过 MatCloud+ 材料云平台使用 MatCloud-QE 第一性原理计算程序包可进行能隙、态密度和光学性质的计算, 计算结果不仅和文献报道结果相符, 且整个计算过程减少人工干预次数可达 90% 以上, 极大地提高了效率。目前, MatCloud-QE 第一性原理计算程序包深度集成在 MatCloud+ 材料云平台, 用户通过云端才能使用, 该模式为用户提供了丰富的计算资源和一站式第一性原理计算服务。

**关键词:** MatCloud-QE 第一性原理计算程序包; MatCloud+ 材料云; 第一性原理计算; 高通量材料计算; 云原生

**中图分类号:** TB30; TP319 **文献标识码:** A **文章编号:** 1674-3962(2024)11-1007-09

## MatCloud-QE: A Cloud-Native High-Throughput First-Principles Simulation Software Package

YANG Xiaoyu<sup>1, 2, 3</sup>, HAO Debo<sup>3</sup>, SHU Cheng<sup>3</sup>, XU Lifang<sup>3</sup>,  
LIANG Yanan<sup>3</sup>, MA Xinjie<sup>3</sup>

(1. Computer Network Information Centre, Chinese Academy of Sciences, Beijing 100190, China)

(2. University of Chinese Academy of Sciences, Beijing 100049, China)

(3. Beijing Maigao MatCloud Co., Ltd., Beijing 100190, China)

**Abstract:** In China, carrying out first-principles simulation is always limited to 1~2 commercial softwares, these softwares not only require software licenses, but also need to be downloaded and manually deployed to high-performance computing clusters, which greatly restricts the application of first-principles simulations. Based on fundamental Quantum ESPRESSO first-principles simulation package, we developed AI-driven high-throughput first-principles simulation acceleration

algorithm, parameter intelligent recommendation algorithm, input structure unifying algorithm, automatic post-processing engine, component engine, graphical engine, more simulation functions, and application programming interface (API) to MatCloud+ cloud platform, forming a Cloud-native first-principles simulation software package, namely MatCloud-QE. Through the MatCloud+ cloud platform, users can carry out QE first-principles simulation and data management on-line only through the Web browser

收稿日期: 2022-03-04 修回日期: 2022-08-24

基金项目: 云南省材料基因工程专项资助项目(2019CLJY06, 202002 AB080001-6); 国家自然科学基金资助项目(52142306); 信息功能材料国家重点实验室开放课题项目(SKL2021)

第一作者: 杨小渝, 男, 1968 年生, 研究员, 博士生导师,

Email: kxy@cnic.cn

DOI: 10.7502/j.issn.1674-3962.202202023

without downloading or installing any third-party softwares. This paper takes the first-principles simulation of material properties of inorganic perovskite  $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  as an example, showing that the MatCloud-QE first-principles simulation package can be used through the MatCloud+ cloud platform to calculate energy gap, state density and optical properties. The calculation results are not only consistent with the results from literatures, but also reduce the number of manual interventions by more than 90% during the entire computing process, which greatly improves efficiency.

**Key words:** MatCloud-QE first-principles simulation software package; MatCloud+ cloud platform; first-principles simulation; high-throughput materials simulation; Cloud native

## 1 前言

第一性原理计算可以利用自洽场法求解薛定谔方程,进而通过计算模拟得到给定结构的各种性质<sup>[1,2]</sup>。在过去的 20 年中,第一性原理计算发展极为迅猛,到目前为止,用户认可并广泛使用的第一性原理计算程序包包括 CASTEP<sup>[3]</sup>、VASP<sup>[4,5]</sup>、Quantum ESPRESSO (QE)<sup>[6,7]</sup>、Siesta<sup>[8]</sup>、Elk<sup>[9]</sup>、WIEN2k<sup>[10]</sup>和 Abinit<sup>[11]</sup>等,其中每个程序包都有自己的特色。如 VASP 由于早期就具备了较为完整的赝势库,因而应用较为广泛,但是随着近些年赝势库的不断完善<sup>[12-15]</sup>,第一性原理计算软件呈现出百家争鸣的态势。然而,一直以来,国内用户开展第一性原理计算总是局限于某一两个商业程序包。这些软件不仅使用许可购买价格昂贵,而且还需要下载、本地安装,再通过人工部署到高性能计算集群才能开展使用。对于不熟悉 Linux,不熟悉编程,而又想开展第一性原理计算的用户来说,始终是个极大的困扰。

与其它第一性原理计算程序包相比,开源的 Quantum ESPRESSO 是一款非常出色的计算软件,ESPRESSO 的全名为 opEn-Source Package for Research in Electronic Structure, Simulation, and Optimization,这表明 QE 是一款用于电子结构、模拟、弛豫的量化计算且开源免费的软件。除了具备常规的第一性原理计算功能外, QE 还内置了一些主流第一性原理计算软件所没有的功能模块,如超导计算、电子能量损失谱、弹道输运等。

正是由于无需版权、功能全面的特点, QE 获得了大量用户的认可,目前,通过 QE 进行第一性原理计算的相关论文已经超过 2 万篇,稳居该技术领域内的第二梯队。自 2009 年 QE 问世以来,它被文章引用的次数逐年增加,表明使用 QE 的用户数逐年增加, QE 软件包使用用户数始终和 VASP 接近(图 1)。截至 2021 年,对 Google Scholar 的第一性原理计算软件用户数据进行统计,结果显示, QE 的用户增长率和其它主流软件几乎持平,呈现出良好的发展趋势。

但是 QE 本身也存在很多不足,给用户带来很多不便,比如以下几点:① QE 涉及参数众多,参数解释晦涩难懂,熟悉了解并挑选出合适的参数较为困难,且格式要求多,容易由于人为失误造成任务失败;② 输入

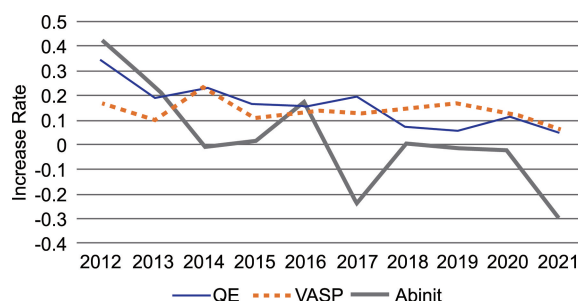


图 1 几个主要的的第一性原理计算软件使用人数增长率比较

Fig. 1 Increase trends of users number of major first-principles simulation softwares

结构层面,结构建模、结构调控需要利用某些商业软件进行模型搭建(通常需要版权),再导入到 QE;③在赝势方面,不同类型的赝势需要从网站采集,没有整合成库,不同赝势推荐截断能不一样,需要从各个渠道寻找;④计算提交方面,首先计算集群需要配置,安装编译 QE 还需要依赖一些底层的软件库,作业提交还需用户自行编写提交脚本,这需要用户有编程基础;此外,作业计算过程中,无实时任务监控及各任务间的衔接,需要用户人工干预,非常浪费时间。⑤后处理时,用户需要学习编程及熟悉后处理代码,数据难复用,易丢失。⑥基于 QE 的高通量计算筛选应用的加速,也有待解决。

为了让更多用户通过浏览器使用免费开源的第一原理计算软件,在线开展模型搭建和第一性原理计算,且无需下载、安装任何软件,作者团队基于 QE 基础程序内核,进一步开发了高通量第一原理计算加速算法、参数智能推荐算法、输入结构统一算法、自动前处理后处理引擎、组件化引擎、图形化引擎、MatCloud+材料云接口等算法程序等,形成了国内首个基于云原生(Cloud Native)理念的第一原理计算程序包:MatCloud-QE。云原生理念最早由 Matt Stine 提出,微软将其定义为“云原生体系结构和技术是一种设计、构造和操作在云中构建并充分利用云计算模型的工作负载方法”<sup>[16]</sup>。除云原生特点外, MatCloud-QE 不仅继承了 QE 的所有特点和计算功能,还动态拓展了 QE 程序所不支持的物理化学性质计算功能(如发射率的计算、断裂强度的计算等)。通过 MatCloud-QE 第一性原理计算程序包与 MatCloud+高通量

多尺度材料集成计算材料云接口, 用户使用浏览器通过 MatCloud+材料云就可使用 MatCloud-QE 的高通量计算筛选加速、参数智能推荐、自动前处理和后处理、功能组件化、图形化、云端拖拽式流程设计等功能, 在线开展第一性原理计算和数据的自动化采集和管理, 极大地方便用户开展第一性原理计算以及高通量计算筛选。

MatCloud-QE 创新之处在于: ① 通过 AI 实现了对大批量高通量第一原理计算作业处理的加速; ② 基于“软件定义材料计算模拟”的理念, 将计算模拟中的模型搭建/高通量建模、各任务处理间数据流动(如几何优化、静态计算)、参数设置、赝势处理、计算数据后处理、计算数据持久化以及机器学习等关键环节, 通过图形化和组件化使用户通过鼠标“拖拽”方式可实现高通量筛选逻辑的“自组装”; ③ 通过“建模→计算→数据→AI”的云端自动化流程, 解决了材料计算参数设置复杂、赝势处理繁琐、数据后处理易出错、计算数据易丢失等问题, 帮助实现材料自动化发现, 计算模拟一旦结束即可自动形成材料计算数据库。MatCloud-QE 是国内首个实现云端 Quantum ESPRESSO 高通量材料智能计算的云平台。

## 2 MatCloud+材料云框架

### 2.1 MatCloud+材料云

MatCloud+材料云是一个基于微服务架构的高通量多尺度材料智能计算大数据云平台(简称 MatCloud+材料云, 软件著作权号 2020SR0753167)(<http://mat-cloudplus.com>), 是在原中国科学院 MatCloud<sup>[17, 18]</sup>的基础上进行了底层的全部代码重构和二次开发。MatCloud+基于云计算理念, 将常用的材料计算方法、模型、计算流程、材料数据、人工智能算法以及计算集群一体化集成置于云端, 用户仅需一个浏览器, 通过 PC 端或手机端等便可登陆进行材料模型构建、模拟计算、数据存储与分析、数据挖掘等工作, 是用于在线开展“材料计算、数据、AI”的“新基建”(美国和韩国均有, MatCloud+是国内首家), 极大地方便老师和学生开展材料计算、快速获取数据和数据分析、通过人工智能构建 QSPR(quantitative structure-property relationship)模型以预测材料关键物性, 也便于各高校开展“计算材料学”的教学工作<sup>[1]</sup>。目前 MatCloud+主要支持量子力学和分子动力学高通量多尺度材料计算和数据管理。MatCloud+材料云的工作流程如图 2 所示: 用户在线进行模型搭建或导入, “拖拽式”设计一个计算流程; 流程启动后, 数据读取、几何优化计算、静态计算、能带计算、态密度计算将会自动化协同有序进行, 无需人工干预。尤其是, 上述的计算作业

将会被自动提交到绑定的国家超算中心, 或自有队列, 或集群。计算结束后, 系统自动获取数据, 并自动存入数据库中。

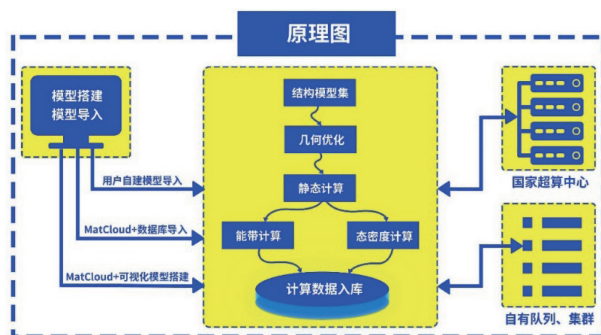


图 2 MatCloud+材料云工作流程图

Fig. 2 Procedure schematic of MatCloud+ for high-throughput materials simulations in Cloud

### 2.2 MatCloud+材料云的主要特点

MatCloud+材料云与国外一些具有类似功能的材料计算软件或计算数据库相比, 如 Materials Studio、MedeA、Material Project、AFLOW 和 AiiDA 等, 其主要特点是“高通量、高并发、网络化、图形化、集成化、流程化、自动化、多尺度”。它的主要创新点在于: ① 通过云计算的模式开展高通量材料计算和数据管理, 通过浏览器不仅可进行计算, 还能进行自动流程化的高通量材料筛选, 进而设计新的材料。整个使用过程无需安装和下载任何软件(商业软件用户自带版权)。② 通过电商模式开展材料计算。使用 MatCloud+材料云开展计算需要购买一定的机时, 借鉴电商模式, 支持“机时选择-在线支付-账户充值-电子发票”的“一站式”商务模式, 便于用户购买机时和在线开展材料计算。③ 支持 PC 端和移动端。传统的模拟计算依赖于本地电脑的配置, 而 MatCloud+将所有功能移植到云端实现, 不仅对电脑配置要求极低(只要能使用浏览器即可), 而且可以在不同的终端上实现模拟计算, 方便随时随地管理作业和数据。④ 支持材料计算数据库的快速构建, 以及在线机器学习和智能算法推荐, 并且能和多个超算中心对接, 支持 LSF、PBS 以及 SLURM 等作业管理系统。

### 2.3 前后端分离

MatCloud+材料云进行了前后端分离以及动态页面和静态页面的分离。后端仅通过 Restful API 返回前端所需的数据, 不再渲染 HTML 页面, 不再控制前端的效果; 前端的呈现效果由前端的程序专门开发。MatCloud+系统采用的是 VUE 前端处理框架, 无论是手机端或移动端, 所需的数据基本相同, 后端仅需开发一套逻辑, 基于定义好的 API 对前端提供数据即可。



## 2.4 微服务架构

MatCloud+材料云主要基于微服务框架 SpringCloud 进行开发,对材料计算和数据管理业务进行了有效的拆分应用,实现敏捷开发和部署。由图 3 微服务架构结构图可见,系统的主要微服务模块包括:用户管理、充值、用户认证、数据查询、机器学习、作业调度和管理、高

通量筛选和后台管理(如机时统计、任务统计、埋点分析等),上述服务均独立运行。数据库采用 MySQL 和 MongoDB 的混搭架构。通过上述各服务模块有机、协同地工作,完成高通量材料计算和数据管理工作。各模块之间的关系包含控制流、数据流和交互流,它们之间全面解耦,又能协同工作。

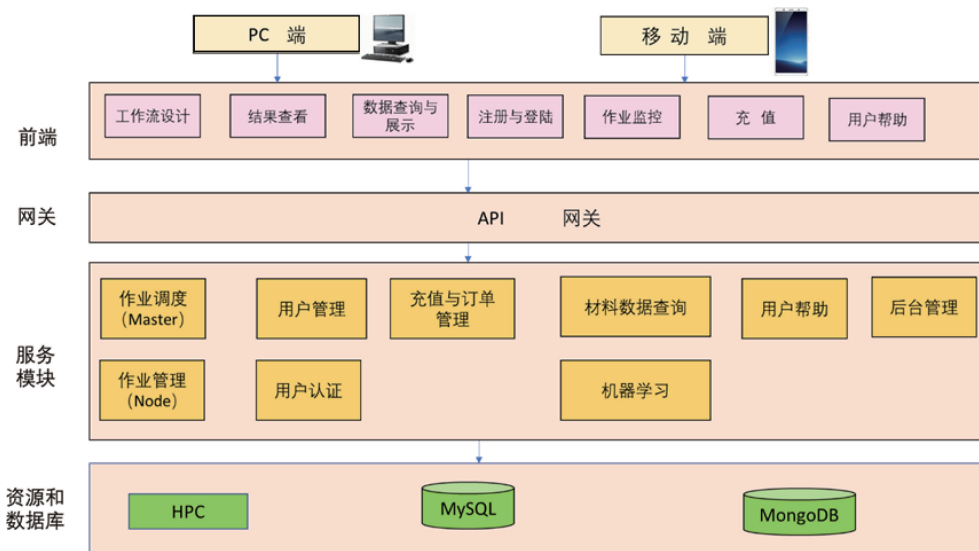


图 3 MatCloud+材料云微服务架构

Fig. 3 MatCloud+ platform architecture diagram

## 3 MatCloud-QE 第一性原理计算程序包

### 3.1 基于 AI 的高通量第一性原理计算加速算法

随着超级计算的普及和机时成本的下降,开展第一性原理计算的模式不再停留在以往单一的第一性原理计算,高通量计算筛选已成为一种新的模式和手段。高通量材料计算筛选是指对大量的晶体或分子结构通过理论计算计算出预定义的筛选描述符值,然后按这些筛选描述符值进行筛选,得到满足预定义筛选描述符的目标结构,也称高通量虚拟筛选。高通量计算筛选一般有如下 3 个准则:① 时间准则。从大量的化学空间或候选结构空间中,要在尽量短的时间内,找出“命中”(hit),时间是一个很重要的考虑因素。② 自动化流程。人工进行高通量筛选不仅极具挑战性,而且也是根本不可能的。高通量计算筛选涉及大量数据的生成、存储和查询,因此需要一定程度的自动化才能有效。③ 筛选漏斗。通常情况下,对所有潜在结构进行描述符筛选的大规模计算太过昂贵,因此采用计算漏斗是一个较好的办法。漏斗的每一层都代表一个有明确误差界限的计算,在每一层中,根据误差界限定义选择标准,不符合的结构将被排除。由于每一级的计算强度逐渐增加,只对最可能感兴趣的分子用最昂贵的方法进行计算。实际上,如何从大量候

选结构空间中通过“筛选漏斗”快速地筛选出目标结构,时效性和准确性这对矛盾体一直是高通量材料计算筛选一个最核心问题。

为了解决这个问题,作者团队基于 QE 基础量子力学程序开发了基于 AI 的高通量第一原理计算加速算法,采用与第一性原理计算获取物性具有等效作用的代理模型机制来加快高通量计算筛选。也就是说,“筛选漏斗”中的描述符值获取可由代理模型帮助获取。由于通过代理模型获取相关描述符值比直接通过 QE 计算在速度上有着量级的提升,因而可提高筛选效率。通过高通量计算筛选生成结构候选空间,一般按照指定的规则,依据组合的思路生成给定分子或晶体的部分或所有可能的组合,一般有替代策略、连接策略、融合策略等,往往会形成千量级、万量级的第一性原理计算作业(如掺杂)等。这些作业往往具有一定的“同构”特点,也就意味着可用机器学习模型,作为量子力学计算的代理模型或等效模型进行预测。因此针对这类高通量计算筛选,在 QE 标准量子力学程序基础上,作者团队开发了相关的 AI 算法:它能够按照一定的分布,从候选空间中选取部分结构进行第一性原理计算,得到标注的数据,形成基础训练集。基于此基础训练集,自动调用不同机器学习算法进行迭代优化训练,并自动选择最优机器学习



算法, 输出能达到设定计算精度误差范围的最优代理模型。如果该代理模型不能达到计算精度的误差范围(误差范围由用户通过图形化界面输入), 则进一步通过第一性原理计算补充训练集, 并将此次模型预测值与计算值的误差值反馈给下一次模型训练迭代, 进一步训练代理模型和修正误差。一旦代理模型预测值达到设定的计算精度误差范围, 描述符值的获取就可切换为通过直接调用该代理模型获取描述符值。基于该方法, 可实现在一定计算精度误差范围内 10 倍以上的加速。

### 3.2 参数智能推荐算法

不同元素种类、计算精度要求、K 点的选取, 往往会有着不同的第一原理计算参数设置, 或每个作业所需计算机核数、内存要求不同。不正确选择它们, 往往会导致计算作业的失败。对于初次开展第一性原理计算的用户, 如何选取这些参数给用户带来很大困扰。因此, 作者团队参照大量文献, 开发了 QE 参数智能推荐算法, 能够针对常见的元素种类、计算精度、K 点的选取等, 自动推荐出一些最优参数, 免去用户设置参数时的困扰。

### 3.3 结构模型统一算法

QE 输入文件中结构数据的获取需要涉及到模型的搭建、结构文件转换和文件拆分 3 个步骤, 模型的搭建和结构转换通常需要借助第三方软件来实现, 对于用户的使用非常不便捷。基于 QE 基础程序内核, 作者团队开发了统一的结构导入算法, 对结构进行了逻辑上的统一。用户不需要考虑文件格式, 同时支持批量结构导入, 支持导入多种晶体结构文件作为 QE 的输入结构。同时, 该结构模型统一算法还提供了与 MatCloud+材料云的接口, 使之与 MatCloud+材料云的建模型件协同, 实现了大部分结构的创建或调控功能, 可以让用户完全不用借助第三方建模软件, 仅通过浏览器就可实现结构“模型搭建—计算—数据—AI”的端到端一体化。

### 3.4 后处理引擎

标准的 QE 计算任务正常结束之后, 传统情况下, 用户需要利用后处理软件或脚本自行处理数据(如 Origin), 处理得到的结果大多保存在本地电脑或者各大云盘中, 数据相互之间无关联, 也不能直接进行计算结果的可视化(如能带、态密度、声子谱)。另外, 由于团队人员的流进流出, 数据往往容易丢失, 后期想要复用数据难度极大。因此, 在 QE 基础程序的基础上, 开发了后处理引擎, 使得在标准 QE 计算程序完成后, 不仅可以自动提取输出文件中的关键数据, 还可以将得到的关键数据自动实时保存在云端数据库, 在保证数据安全的基础上, 还可以利用数据库对数据进行二次筛选和复用。

### 3.5 组件化引擎

QE 拥有十几个功能模块, 而仅 PWscf 一个就包含有 200 多个参数, 对于初学者来说晦涩难懂, 在 QE 基础程序的基础上, 开发了组件化引擎。该组件化引擎包括组件库以及组件参数设置与调用算法。根据 QE 功能的不同, 从 QE 的各个模块中将独立的功能进行抽取, 形成 QE 基础程序所不具备的组件库。同时开发了与 MatCloud+材料云的量子力学程序接口, 使组件库能呈现为一个一个图形化组件。同时结合上述 3.3 节的参数智能推荐算法, 为每个组件推荐关键参数。用户使用时, 仅通过点选方式即可选择组件, 设置参数, 操作灵活简便。另外, 为了方便中文用户, 组件化引擎还开发了中文参数库, 给每个参数都提供了中文使用帮助, 极大减少了用户调研参数花费的时间。

针对赝势及其内部参数获取繁琐的问题, 组件化引擎将全种类赝势整合到了每个组件中, 用户仅需点选想要选择的赝势种类, 就可实现为每个元素自动分配此类型的赝势, 并结合元素和赝势种类自动给出推荐的截断能。

LibXC 是目前最全面、最强大的交换关联泛函库, 一直在业界开源社区得到维护, LibXC 泛函库包括各种泛函, 如 LDA、GGA、杂化泛函以及 metaGGA 等。为了让用户更加便捷地选择交换关联泛函, 团队给组件库开发了专门接口, 使其对 LibXC 全面支持。

### 3.6 图形化引擎

在 QE 基础量子力学程序的基础上, 开发了图形化引擎算法, 它的核心理念在于: 基于软件定义的理念, 重新定义了第一性原理计算的操作, 将模型搭建、高通量建模、各处理间数据流动(如几何优化、静态计算)、参数设置、赝势处理、计算数据后处理、计算数据持久化等关键环节进行图形化处理。此外, 开发了图形化引擎与 MatCloud+材料云的专门接口, 使得用户通过灵活的“拖拽”方式实现自组装的各个功能集中整合。同时实现了 QE 高通量材料计算“建模—计算—数据—AI”的“云端”一体自动化流程, 直接连接千核/万核集群, 解决了材料计算参数设置复杂、赝势处理繁琐、数据后处理易出错、计算数据易丢失等问题。

### 3.7 支持更多物理化学性质的计算

MatCloud-QE 除了支持基础 QE 第一性原理计算所支持的物理化学性质计算外, 还支持其他一些物理化学性质计算(如发射率、断裂强度等), 且所支持的物理化学性质计算还在不断增加。

### 3.8 与 MatCloud+材料云的 API 接口

使用 QE 基础程序, 用户不仅需要购买计算集群,

还需要编译安装 QE、准备提交脚本等等,这就需要用户具备一定的 Linux 语言基础。为了更好地利用 MatCloud+材料云框架所提供的功能, MatCloud-QE 第一原理计算程序包专门提供了与 MatCloud+材料云的接口,使得通过 MatCloud+材料云可以便捷地使用 MatCloud-QE 第一原理计算程序包。由于 MatCloud+材料云本身直接对接国家超算中心,用户不需要考虑集群和软件配置环境的问题,直接通过浏览器进行结构建模、设计工作流程、选择超算、点击提交按钮就可开展第一性原理计算。计算一旦结束,数据直接可视化和入库,无需使用任何第三方软件。另外,通过该接口程序,可通过 MatCloud+材料云拖拽式流程实现批量 QE 任务的提交,一次性可以高通量计算多个结构的多个性质,整个工作流程无需人为操作任何文件,各计算任务有效协同,不仅最大化利用了碎片化时间,也杜绝了由于误操作导致的各种问题。

## 4 应用案例

本文以面向太阳能电池应用的无机钙钛矿  $\text{Rb}_x\text{Cs}_{1-x}\text{PbI}_3$  材料特性的第一性原理计算研究为例<sup>[19]</sup>,来说明通过 MatCloud+材料云,如何利用 MatCloud-QE 第一原理计算程序包开展第一性原理计算。

### 4.1 案例阐述

研究表明,用无机 Cs 或 Rb 阳离子代替或混合杂化卤化物钙钛矿中的有机分子可以提高材料稳定性,同时提高太阳能电池性能<sup>[19]</sup>。 $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  的结构如图 4 所示。在文献[19]中,作者采用了元素替代法,掺杂不同

含量的 Rb 到  $\text{CsPbI}_3$  中,生成多个结构,使用第一性原理计算系统地研究了全无机碱金属碘化物钙钛矿  $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  的电子和光学性质,因此该计算属于典型的多结构、多性质计算(如图 5)。

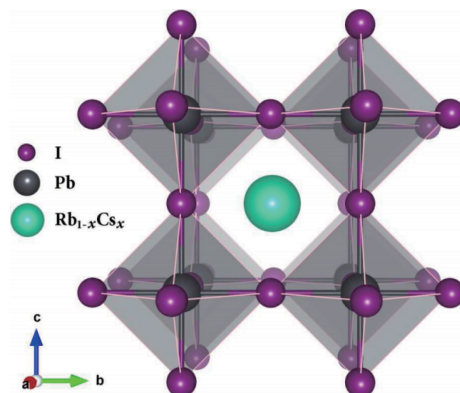


图 4  $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  结构示意图

Fig. 4 Illustration of  $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  structure

### 4.2 利用 MatCloud-QE 开展第一性原理计算的过程

通过 MatCloud+材料云,使用 MatCloud-QE 第一性原理计算程序包重复上述文献案例,具体的操作步骤如下。

#### (1) 高通量建模: $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$

通过“手动”方式生成大量掺杂结构,通常既耗费时间又无法保证结果的严谨性。MatCloud+材料云的【随机取代】功能模块,通过调用 MatCloud-QE 程序包的结构统一算法,可以一次性建模,生成大量的可供筛选的结构,形成结构候选空间(图 6)。

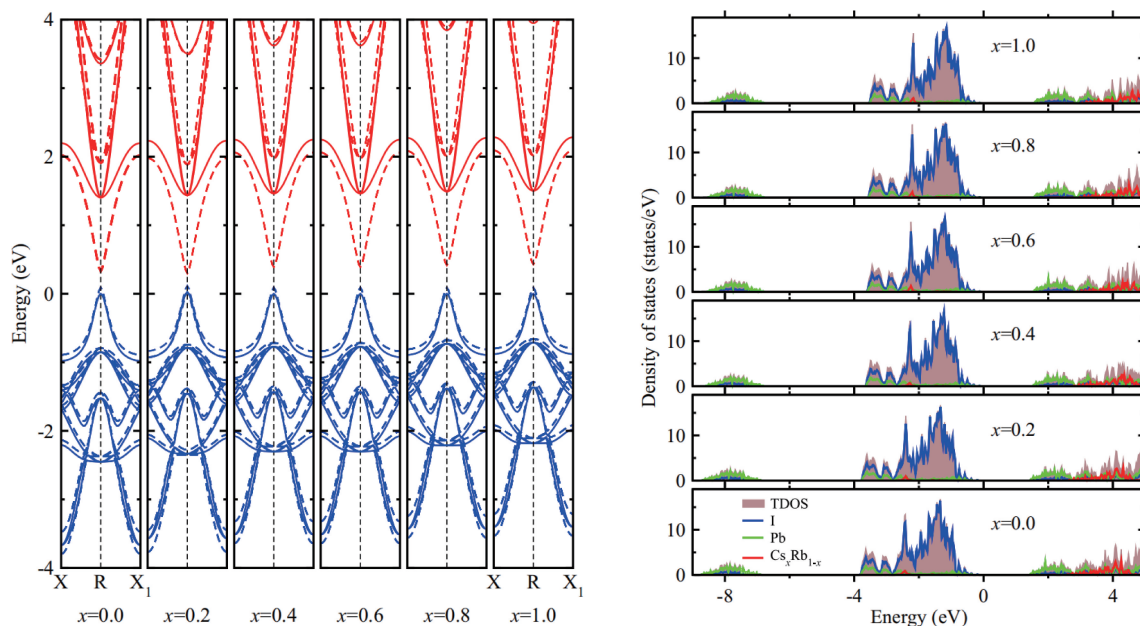


图 5 无机碘化钙钛矿  $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  的电子性质随着 Cs 含量的变化<sup>[19]</sup>

Fig. 5 Electronic properties changes of  $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  with Cs content increasing<sup>[19]</sup>

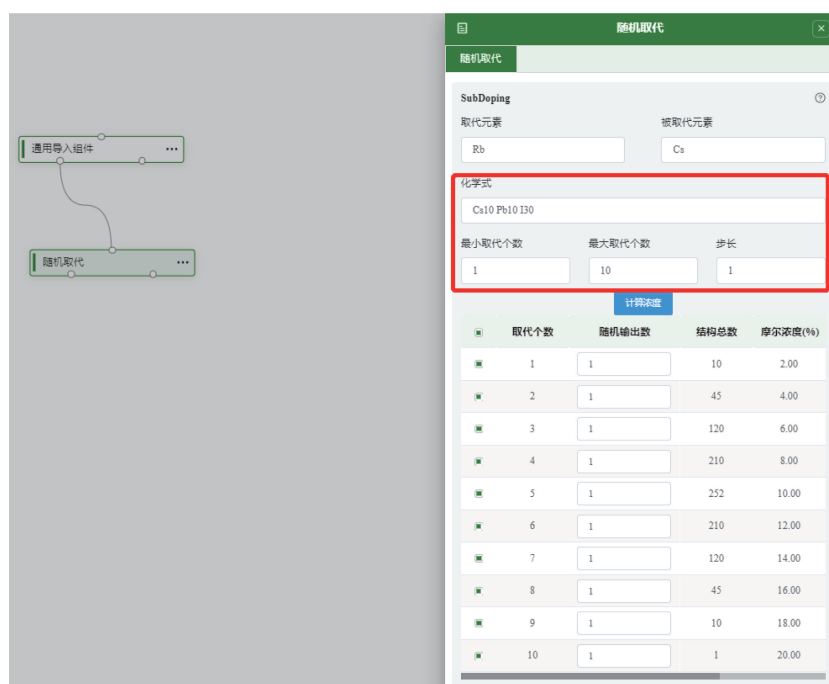


图 6 MatCloud+材料云的【随机取代】功能模块调用 MatCloud-QE 程序包的结构统一算法生成结构候选空间

Fig. 6 “Random Substitution” functional component of MatCloud+ invoking “Structure Unify Algorithm” of MatCloud-QE first-principles simulation package generating large number of candidate structures

通过指定最小取代个数、最大取代个数、步长,便可搭建一定取代浓度范围内的多个模型,并得到可视化的所有浓度百分比饼图,同时可将多个取代掺杂模型一键导入数据库、结构集或本地(图 7)。

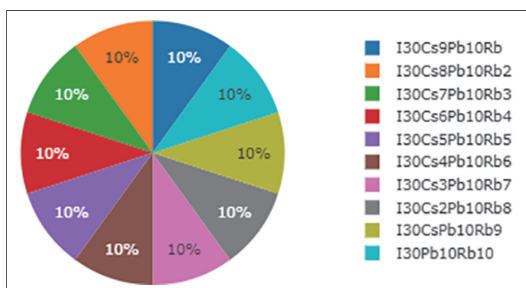


图 7 调用 MatCloud-QE 程序包的结构统一算法得到的随机取代建模结果

Fig. 7 Results of “Random Substitution” functional component of MatCloud+ invoking “Structure Unify Algorithm” of MatCloud-QE first-principles simulation package

## (2) 性质计算: $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$

采用 MatCloud-QE 第一原理计算程序包,通过与 MatCloud+材料云接口,仅需如下 4 步即可完成文献中全无机碱金属碘化物钙钛矿  $\text{Rb}_{1-x}\text{Cs}_x\text{PbI}_3$  的电子和光学性质计算:上传结构模型—创建计算工作流—设置计算参数并提交计算—计算结果查看。

文献[19]中计算了 10 个结构的电子和光学性质,共

需要提交 50 次计算(既结构优化、静态计算、能带结构、电子态密度、光学性质),通过 MatCloud-QE 第一原理计算组件化引擎算法,用户只需要拖拽 MatCloud-QE 组件形成计算工作流(图 8),就可完成传统的 50 次计算,而所需人工干预只需 3 次(工作流设计、参数设置、作业提交等),减少的人工干预次数可达 90% 以上,且一旦计算结束,可快速地形成计算物性数据库。对于千量级、万量级的第一性原理计算作业,可进一步使用 MatCloud-QE 的高通量第一原理计算加速算法(图 8)。

## (3) 与文献计算结果的对比

MatCloud-QE 后处理引擎提供了可视化的计算结果,且支持不同形式的结果下载,见表 1 和图 9。这里我们仅比较  $\text{CsPbI}_3$  掺杂前的能隙和态密度计算结果。从表 1 可以看出,通过 MatCloud+材料云,使用 MatCloud-QE 第一原理计算程序包计算的  $\text{CsPbI}_3$  的能隙值与文献是一致的。

## 5 结 语

作者团队在 Quantum ESPRESSO (QE) 第一性原理计算基础程序基础上,进一步开发了高通量第一性原理计算加速算法、参数智能推荐算法、输入结构统一算法、自动前处理后处理引擎、组件化引擎、图形化引擎、工作流接口等算法程序等,形成了具备云原生特点的第一性原理计算程序包: MatCloud-QE。进一步通过与 MatCloud+



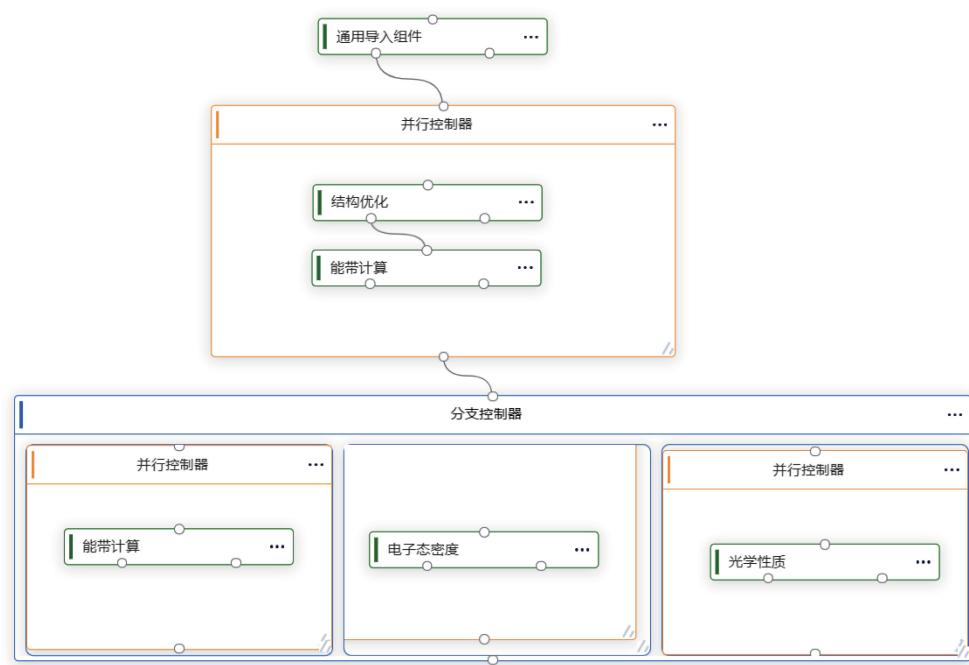


图 8 通过 MatCloud+材料云，使用 MatCloud-QE 程序包开展高通量第一性原理计算的工作流  
Fig. 8 A workflow of MatCloud+ invoking MatCloud-QE first-principles simulation package for high-throughput simulations

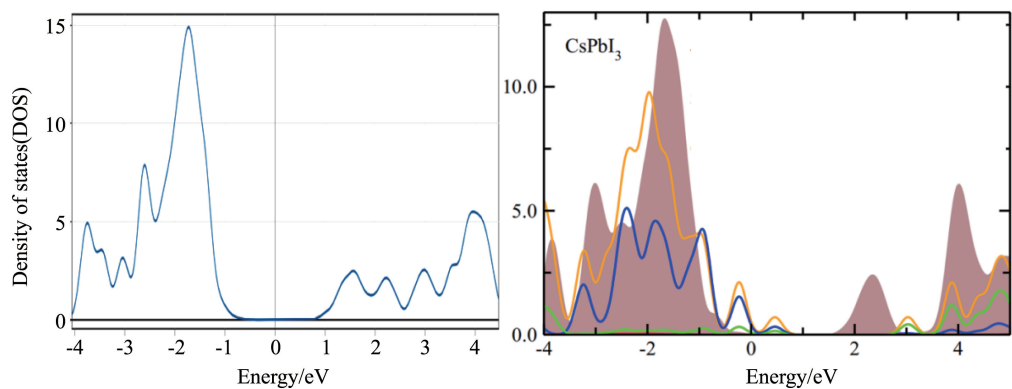


图 9 MatCloud-QE 后处理引擎处理得到的 CsPbI<sub>3</sub> 电子态密度和文献 [19] 报道值(褐色)的对比  
Fig. 9 Totaol DOS of CsPbI<sub>3</sub> by “Post-processing Engine” of MatCloud-QE first-principles simulation package, compared with result from literature [19] (in brown color)

表 1 采用 MatCloud-QE 第一性原理计算程序包计算的 CsPbI<sub>3</sub> 的带隙值与文献 [19] 结果的对比

Table 1 Band gap value of CsPbI<sub>3</sub> by MatCloud-QE, compared with value from literature [19] (eV)

MatCloud-QE	Calculation value in ref. [19]
1.47	1.45

高通量多尺度材料集成计算(材料云)接口，实现用户通过浏览器就可在线开展第一性原理计算和数据的自动化采集和管理，极大地方使用户开展第一性原理计算以及高通

量计算筛选，且无需软件使用许可。MatCloud-QE 云原生第一原理计算程序包深度集成在 MatCloud+材料云平台，为用户提供高效便捷的云端计算服务。该程序包专注于通过云端运行，为用户提供灵活的计算资源和一站式的第一原理计算功能。尽管 MatCloud-QE 第一原理计算程序包使用免费，但应用案例表明，结合 MatCloud+材料云使用 MatCloud-QE 比直接使用 MatCloud-QE，有更好的用户体验和效果：不仅免去用户安装和下载 MatCloud-QE，还能体验到 QE 基础程序所不支持的高通量第一原理计算加速、输入结构统一、全程图形化展示、智能参数推

荐、数据自动提取并实时入库、更多物理化学性质计算等功能, 从而进一步降低了第一性原理计算的使用门槛, 并大幅度提升其计算效率。

## 参考文献 References

- [1] MARTIN R M. Electronic Structure: Basic Theory and Practical Methods[M]. Cambridge: Cambridge University Press, 2004.
- [2] 李正中. 固体理论[M]. 北京: 高等教育出版社, 2002.  
LI Z Z. Solid State Theory[M]. Beijing: Higher Education Press, 2002.
- [3] CLARK S J, SEGALL M D, PICKARD C J, *et al.* Journal of Zeitschrift für Kristallographie-Crystalline Materials[J], 2005, 22(5/6): 567-570.
- [4] KRESSE G, FURTHMÜLLER J. Computational Material Science[J], 1996, 6(1): 15-50.
- [5] KRESSE G, FURTHMÜLLER J. Physical Review B[J], 1996, 54: 11169-11186.
- [6] GIANNOZZI P, BARONI S, BONINI N, *et al.* Journal of Physics: Condensed Matter[J], 2009, 21(39): 395502.
- [7] GIANNOZZI P, ANDREUSSI O, BRUMME T, *et al.* Journal of Physics: Condensed Matters[J], 2017, 29(46): 465901.
- [8] SOLER J M, ARTACHO E, GALE J D, *et al.* Journal of Physics: Condensed Matter[J], 2002, 14(11): 2745-2779.
- [9] DEWHURST K, SHARMA S, NORDSTROM L, *et al.* The Elk Fp-Lapw Code[EB/OL]. (2020-04-04) [2022-04-19]. [http: //](http://elk.sourceforge.net/)
- elk.sourceforge.net/.
- [10] BLAHA P, SCHWARZ K, TRAN F, *et al.* Journal of Chemical Physics[J], 2020, 152(7): 074101.
- [11] GONZE X, JOLLET F, ARAUJO F A, *et al.* Computer Physics Communication[J], 2016, 205: 106-131.
- [12] PRANDINI G, MARRAZZO A, CASTELLI I E, *et al.* npj Computational Materials[J], 2018, 4: 72.
- [13] LEJAEGERE K, BIHLMAYER G, BJORKMAN T, *et al.* Science[J], 2016, 351(6280): 1415.
- [14] GARRITY K F, BENNETT J W, RABE K M, *et al.* Computational Material Science[J], 2014, 81: 446.
- [15] SCHLIPF M, GYGI F. Computer Physics Communication[J], 2015, 196: 36.
- [16] Microsoft Corporation. What is Cloud Native? [M/OL]// Architecting Cloud Native .NET Applications for Azure. Washington, 2023. [2024-11-15] [https: //learn.microsoft.com/en-us/dotnet/architecture/cloud-native/definition](https://learn.microsoft.com/en-us/dotnet/architecture/cloud-native/definition)
- [17] YANG X, WANG Z, SONG J, *et al.* Computational Material Science[J], 2018, 146: 319-333.
- [18] YANG X, WANG Z, SONG J, *et al.* Chinese Physics[J], 2018, 27(11): 110301.
- [19] JONG U, YU C, KIM Y, *et al.* Physical Review B[J], 2018, 98: 125116.

(编辑 惠 琼)