

热电材料的第一性原理高通量研究

李鑫, 席丽丽, 杨炯

(上海大学 材料基因组工程研究院, 上海 200444)

摘要: 热电材料是一种新型能量转换材料, 在温差发电或通电制冷等领域具有广泛应用。热电优值 ZT 值是衡量热电材料能量转换效率的关键参数, ZT 值要求热电材料具有优异的电输运性能及较低的热导率。传统第一性原理热电材料研究往往关注少量样本下的电热输运性质理解与优化, 很难得到系统性的规律, 也不利于新体系的设计优化。材料基因组计划力求通过大数据、高通量手段去加速材料设计与发现, 具有广阔的发展前景。在热电材料研究领域, 第一性原理高通量计算也将在新材料预测与性能优化等方面起到越来越重要的作用。另一方面, 高通量研究也带来了新的挑战, 譬如电热输运性质的高通量算法发展、大数据分析手段等等, 这些方面的问题决定了高通量方法在材料应用中的效率与准确性。本文综述了热电材料中现有的电热输运性质高通量计算方法, 介绍了这些方法具体的应用案例, 并对高通量与热电材料结合的未来发展趋势进行了展望。

关键词: 高通量; 第一性原理; 热电材料; 电热输运; 综述

中图分类号: N34 文献标识码: A

First Principles High-throughput Research on Thermoelectric Materials: a Review

LI Xin, XI Li-Li, YANG Jiong

(Materials Genome Institute, Shanghai University, Shanghai 200444, China)

Abstract: Thermoelectric materials are a kind of energy conversion materials, which are extensively used in power generation or refrigeration. The key parameter that measure the performance of thermoelectric materials is the figure of merit ZT value, which requires material excellent electrical transport performance and low thermal conductivity. Standard first principles calculations on thermoelectric materials focus on small samples of materials, which is difficult to conclude general rules and propose new candidates. The Materials Genome Initiative speeds up the discovery and design of materials based on big data and high-throughput computational methods, which is promising in novel material screening. In thermoelectrics, first principles high-throughput calculations play an increasingly important role in the predicting and designing new materials. However, there are some drawbacks in the current high-throughput efforts for thermoelectric material screening, such as the demand of efficient high-throughput algorithms for transport properties, suitable tools for analyzing big data, etc. Solving these challenges strongly determines the efficiency and accuracy of high-throughput applications in thermoelectrics. This review summarizes several high-throughput theoretical methods and cases study on electrical and thermal transport properties in thermoelectric materials, and prospects the future trend of the combination of high-throughput and thermoelectric material research.

Key words: high-throughput; first principles; thermoelectric materials; electrical and thermal transport properties; review

收稿日期: 2018-07-16; 收到修改稿日期: 2018-10-02

基金项目: 国家重点研发计划(2017YFB0701600); 国家自然科学基金(51572167, 51632005, 11574333, 11674211)

National Key Research and Development Program of China (2017YFB0701600); National Natural Science Foundation of China (51572167, 51632005, 11574333, 11674211)

作者简介: 李鑫(1993-), 女, 博士研究生. E-mail: xinli@t.shu.edu.cn

通讯作者: 席丽丽, 副研究员. E-mail: lilyxi@t.shu.edu.cn; 杨炯, 教授. E-mail: jiongy@t.shu.edu.cn

热电材料通过泽贝克(Seebeck)效应和帕尔贴(Peltier)效应实现热能和电能之间的直接转换^[1-2],在热电发电和制冷方面具有重要应用。热电技术已经在工业废热发电利用、空间/深海特殊电源、半导体芯片降温/控温、高端制冷等领域取得应用,也可用于制备微型电源、微区冷却、光通信激光二极管和红外线传感器的调温系统^[3-7]。以热电材料为核心的热电发电器件和系统兼具全固态、无运动部件、无排放、结构简单、体积小、重量轻等诸多优点,已成为一种具有广泛应用前景的绿色能源技术。在环境污染和能源危机日益严重的今天,开展新型热电材料的研究具有很重要的现实意义。热电转换效率低是限制热电材料应用的一个瓶颈,如何有效提高热电材料转换效率是热电技术工业应用亟待解决的难题和热电科学家亟需研究的课题。因此,理解热电性能优化的关键因素,实现热电性能提升,发现高性能热电材料新体系尤为关键。

热电材料的转换效率与热电优值 ZT 值密切相关,提高热电材料转换效率的关键在于提高材料的 ZT 值。 ZT 值可以写成 $ZT = S^2 \sigma T / \kappa$, 其中 S 是泽贝克系数, σ 是电导率, $S^2 \sigma$ 是功率因子,是综合衡量材料电输运性能的参数; κ 是体系的热导率,包括电子热导率 κ_e 和晶格热导率 κ_L 两部分,是衡量材料热输运性能的关键参数; T 是绝对温度。通常泽贝克系数的提高伴随着电导率的降低,反之亦然。如何协同调控泽贝克系数与电导率以达到较高的功率因子是实现热电材料电性能优化的关键,这需要充分理解热电材料的电输运机理,包括电子能带结构和载流子散射机制。热导率包括载流子和晶格声子的贡献,载流子热导率与电导率直接相关,而晶格热导率决定于声子输运,与声子色散关系、电子-声子耦合及其声子输运散射机制等关联。所有这些物理特性相互耦合,使得发现热电材料的新体系和性能优化极具挑战性。

在传统的试错研究方法中,从一种材料的成功研发到商业应用往往需要几年甚至十年的时间。材料基因组计划(MGI)是美国 2011 年提出的,意在缩短材料研发流程的一项方案,主要从材料计算手段、实验手段、数字化数据库建立三个方面发展。其中,第一性原理计算从材料的微观角度出发,只需要材料的组分结构信息便可以直接获得材料的能量、电子结构以及与之相关的其它功能性质参数等,是一种行之有效的材料性能预测手段。MGI 与第一性原理结合能够有效加速材料的研发,预测材料的性质,创新大数据时代材料科学的研究方法,材料数据库与第一性原理计算平台也有了一定的发展。

目前国际上比较知名的数据库或者第一性原理计算平台有 Materials Project(MP: <https://www.materialsproject.org/>)、Automatic-FLOW for Materials Discovery(AFLOW: <http://www.aflowlib.org/>)、Novel Materials Discovery 的第一性原理计算数据库模块 CompES-x(http://compes-x.nims.go.jp/index_en.html)、中国科学院 MatCloud(<http://matcloud.cn/cn/>)、上海大学 Materials Informatics Platform(MIP: <http://mip.shu.edu.cn>)等。这些数据库或计算平台发展手段不尽相同,侧重点也不一样,各有优势(具体平台介绍见附录)。通过数据库或计算平台,可以对大量材料进行批量数据的生成、解析、性质预测等,为后续实验室制备、工业化应用提供充分的前期工作,加速新型材料从设计研发到应用的开发过程。

图 1 是第一性原理高通量材料研究的常规研究示意图,一般分为基础架构、大数据生成、大数据分析三步。(1)基础架构需要数据库、计算机集群、标准高通量流程的构建。数据库里的结构数据是基础研究样本,数据库的构建需要大量的计算机集群资源,而数据库与集群的管理则需要设计一套标准高通量流程,包括数据如何检索、交互、更新和生成等,三者的构建完成了第一性原理高通量材料研究基础。(2)大数据生成主要是基于材料基础数据库而得到材料性质信息,这个过程需要应用合适的第一性原理计算软件、简化的高通量性质算法。基于第一性原理计算软件设计的简化算法是第一性原理高通量研究的关键,如何平衡大数据计算的资源耗费与计算结果精确度是算法设计的重点。(3)大数据分析需要机器学习和理论推演等数据分析手段,以

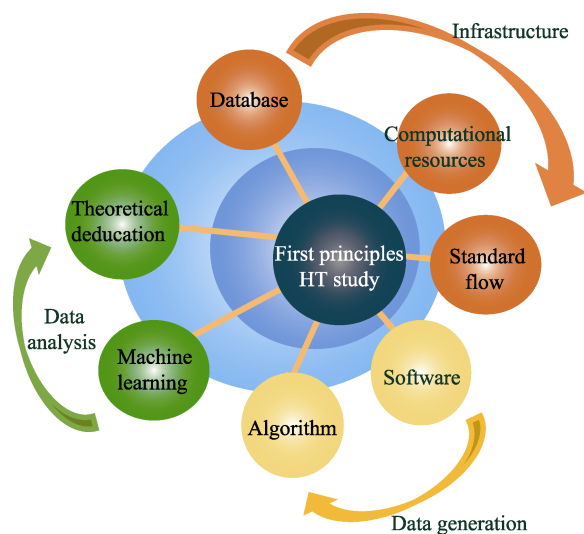


图 1 第一性原理高通量材料研究示意图
Fig. 1 Schematic diagram of first principles high-throughput (HT) study on materials

此对材料的功能性质信息进行分类、整合和预测等, 这些性质信息的反馈又可以充实数据库, 查漏补缺, 改进和完善高通量流程和算法, 并进行往复循环。三步工作既相互独立又相互耦合, 这种高效灵活的第一性原理高通量技术对于大数据时代的科学研究具有很高的应用价值。

截至目前, 高通量材料研发以其效率高、信息量大等特点, 被应用到许多功能材料开发领域中, 譬如电池的研发^[8]、弹性性质数据库^[9]、镁合金的研发^[10]、组分替换找寻新化合物^[11]等。无论是针对材料的某种性质进行高通量解析存储, 还是对某一类材料的应用进行性质预测与分析, 都能大大提高功能材料的研发速度。本文将重点介绍高通量在热电材料中的应用及研究方法。

1 高通量热电研究方法概述

热电材料的高通量研究通常从电、热输运两方面开展, 由于输运性能计算量大, 需要综合考虑计算成本及计算精度问题。图 2 给出了材料的计算成本与计算精确度之间的关系, 蓝色区域代表高精度, 红色区域代表高速, 绿色和黄色实心圆点选取了一些常见第一性原理热电材料计算方法进行标注^[12-20]。从图 2 可以看出, 无论是计算电输运性能, 还是计算热输运性能, 其准确度与计算成本成正相关, 而高通量热电材料研究则是力求达到图 2 中紫色区域所示范围, 既要在一定程度上保证计算的准确度,

又要节约计算成本。下面介绍几种常用的电热输运计算方法。

1.1 电输运性质理论研究方法及高通量处理

公式(1)、(2)给出了基于玻尔兹曼输运理论计算泽贝克系数与电导率的公式^[21], 其中 v_e 、 $N(\varepsilon)$ 分别是电子群速度与能态密度, 代表电子结构色散部分对于电输运性质的影响; τ_e 是电子弛豫时间, 代表散射部分对于电输运性质的影响; f_0 是费米狄拉克分布函数; Ω 是原胞体积; T 是温度; e 是电子电荷; ε_F 是费米能级。宏观电导率计算公式为 $\sigma = ne\mu$, 与公式(1)对比可知, 载流子浓度 n 与能态密度有关, 而迁移率 μ 跟电子群速度和弛豫时间都有关系。

$$\sigma = \frac{1}{\Omega} \int N(\varepsilon) v_e^2 \tau_e \left[-\frac{\partial f_0(T, \varepsilon)}{\partial \varepsilon} \right] d\varepsilon \quad (1)$$

$$S(T) = \frac{1}{eT\Omega\sigma} \int N(\varepsilon) v_e^2 \tau_e (\varepsilon - \varepsilon_F) \left[-\frac{\partial f_0(T, \varepsilon)}{\partial \varepsilon} \right] d\varepsilon \quad (2)$$

电子弛豫时间受到多种散射机制的影响, 其中最重要的电声相互作用散射的计算耗时较长, 无法直接应用于高通量研究, 通常需要对其进行简化。目前针对弛豫时间常用的处理方式常数弛豫时间近似或常数平均自由程近似。如何获得高效准确的弛豫时间算法是提高电输运高通量理论研究方法准确性的研究重点。

1.1.1 常数弛豫时间近似

对于某一类结构相近的材料体系, 其弛豫时间在较小范围内变化, 如果只是定性分析此类材料的电输运性质, 通常可以采用常数弛豫时间近似。常数弛豫时间近似认为材料的弛豫时间是一个固定值, 不受不同类型材料体系的影响。在常数弛豫时间近似下, 弛豫时间的大小不会影响泽贝克系数, 只对电导率有所修正。

基于常数弛豫时间近似的电输运性质研究, 对于泽贝克系数的预测比较好, 但受限于电导率预测的不准确性, 导致功率因子的预测结果偏差较大, 只能定性分析某一类结构类似体系的功率因子高低分布。虽然常数弛豫时间近似会影响预测结果的准确度, 但是由于计算速度比较快, 对于电输运计算模块的开发具有重要意义, 不仅能对特定的譬如有效质量^[22]、泽贝克系数等进行性质描述, 还能在更广范围、更大维度上对材料体系的各种性质进行预测、分类和评估, 标准化的数据类型对于高通量热电材料的研究也具有一定的参考价值。

1.1.2 常数平均自由程近似

当晶粒尺寸减小至纳米尺度范围, 晶粒表面积

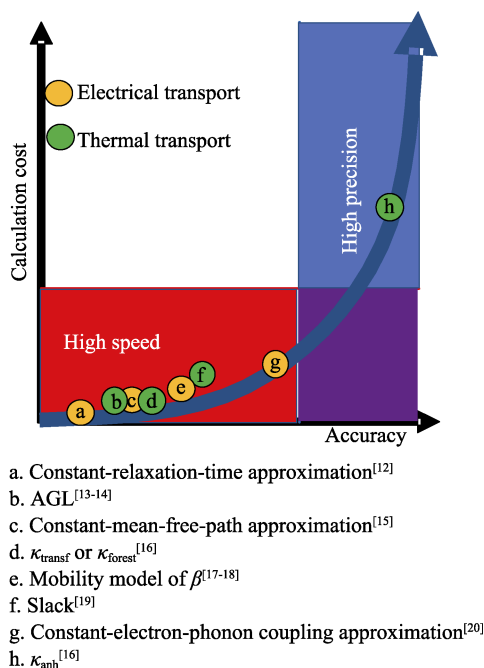


图 2 计算成本与计算精确度关系图

Fig. 2 Relationship between calculation cost and accuracy

增大,晶界散射比块体材料有很大的增强,有效降低了热电材料的热导率^[23]。此时,可以认为载流子的平均自由程 λ 与晶粒尺寸近似相等,即用晶粒尺寸代替载流子的平均自由程,从而对电输运性质进行简化计算。公式(3)和(4)给出了由平均自由程计算电导率与泽贝克系数的算法,其中 η 是化学势, k_B 是玻尔兹曼常数。

常数平均自由程近似将自由程设定为一个常数,可以避免复杂的弛豫时间计算,节约了计算成本,具有一定的可借鉴性。但是其研究对象被限定为纳米尺度材料,并且仅能在数量级上预测功率因子的大小趋势,不能准确地描述材料的电输运性质,普适性也有待考证。

$$\sigma = -\frac{e^2}{k_B T} \lambda \int f_0(1-f_0) \sum(\varepsilon) d\varepsilon \quad (3)$$

$$S = -\frac{e}{\sigma k_B T^2} \lambda \int f_0(1-f_0)(\varepsilon - \eta) \sum(\varepsilon) d\varepsilon \quad (4)$$

1.1.3 引入迁移率的半经验模型

基于热电材料领域基本理论^[17],Chasmar等^[18]采用半经验算符 β 来描述材料的输运性质。在这个模型下 $ZT = a\beta/(b\beta+1)$,而 β 是与能量无关的一个算符,具体定义如公式(5)所示,其中 \hbar 是约化普朗克常数, a 、 b 是与化学势相关的函数, μ_0 是载流子迁移率, m_{DOS}^* 是态密度有效质量。

$$\beta = \left(\frac{k_B}{e} \right)^2 \frac{2e(k_B T)^{3/2} T}{(2\pi)^{3/2} \hbar^3} \frac{\mu_0 m_{DOS}^{*3/2}}{\kappa_L} \quad (5)$$

β 定义式的最后一项中,分母 κ_L 部分代表晶格热导率,与热输运相关;分子 $\mu_0 m_{DOS}^{*3/2}$ 部分则与电输运相关。对于电输运部分, μ_0 可以用公式(6)表示^[24],其中 B 为体模量, m_b^* 为单能谷有效质量, N_b 为简并度。 A_0 、 s 、 t 等参数可以通过实验数据拟合得到,譬如在Yan等^[24]工作中 s 、 t 的取值分别为1、2.5。

$$\mu_0 = A_0 B^s (m_b^*)^{-t}, m_{DOS}^{*3/2} = N_b m_b^{*3/2} \quad (6)$$

引入半经验模型,用模量 B 、单能谷有效质量 m_b^* 来参数化迁移率的做法,相当于考虑了载流子弛豫时间中的散射通道部分。相比于常数弛豫时间以及常数平均自由程近似来讲,公式(6)的准确性更高,但由于该方法用到 N_b 、 m_b^* 等传统输运理论的概念,难以通过第一性原理计算自动化获得,并不适合高通量理论预测。

1.1.4 常数电声耦合近似

弛豫时间主要决定于电声耦合矩阵和散射通道,

Xi等^[20]通过常数电声耦合近似针对复杂的精确电声耦合弛豫时间计算做了一定的简化计算,即避免了常数弛豫时间近似带来的计算误差,又合理规避了精确电声耦合作用计算的高消耗。该方法精确计算了电子散射通道部分,并且认为能带变化对于电声耦合矩阵元的影响较小,将电声耦合部分写成常量,如公式(7),其中 $\tau_{n,k}$ 是弛豫时间, $C_{n,k}$ 是与电声耦合系数相关的常数。

$$\tau_{n,k}^{-1} = C_{n,k} \sum_{n',k'} \delta(\varepsilon_{n,k} - \varepsilon_{n',k'}) \quad (7)$$

常数电声耦合近似简化了弛豫时间计算的时间,并且其精度在一定程度上与精确的电声耦合计算结果可比,对热电材料高通量性能的准确预测具有重要的参考价值。在特定的工作中,还可以进一步参数化 $C_{n,k}$,比如用形变势常数和杨氏模量代替等。

1.2 热输运性质理论研究方法及高通量处理

热电材料一般为窄带隙半导体材料,相较于晶格热导率 κ_L 来说,材料的电子热导率 κ_e 几乎可以忽略,所以对于热电材料的热输运研究一般只关注晶格热导率。公式(8)是常规的晶格热导率计算公式^[21],其中 ω_q 、 v_q 分别是声子频率与声速,代表声子色散对于热输运方面的贡献; τ_q 是声子弛豫时间,代表各种散射对于热输运的贡献; N_q^0 是玻色爱因斯坦分布函数。声子色散方面通常需要计算声子谱,而散射方面则需要三阶力常数,而通常这种三阶非谐项跟格林艾森常数相关,一般格林艾森常数越大,晶格热导率越低。精确计算热导率相关量,但对于二阶的声子谱都较耗时,因此精确计算热导率很难运用于高通量研究。

$$\kappa_L = \frac{1}{3} \sum_q \hbar \omega_q v_q^2 \tau_q \frac{\partial N_q^0}{\partial T} \quad (8)$$

1.2.1 Slack模型计算热导率

晶格热导率常用简化的Slack模型去估算^[19]。考虑了高温极限的Slack模型,只有声学声子对热导率产生贡献,见公式(9)(其中的关键参量为德拜温度 θ_D 和格林艾森常数 γ 。 \bar{M} 是平均原子质量, δ^3 是原胞体积, n 是原胞内原子数)。德拜温度可以通过弹性模量求解声速得到,见公式(10)和(11)^[25](v_m 是平均声速, v_T 是横波声速, v_L 是纵波声速, \hbar 是普朗克常数, ρ 是材料的密度, N_A 是阿伏伽德罗常数, M 是原胞质量, B 是体变模量, G 是剪切模量)。格林艾森常数既可以通过弹性模量求解^[26],也可以通过吉布斯自由能求解^[13]。

$$\kappa_L = A \cdot \frac{\bar{M} \theta_D^3 \delta}{\gamma^2 n^3 T}, A_{(\gamma)} = \frac{5.720 \times 10^7 \times 0.849}{2 \times \left(1 - \frac{0.514}{\gamma} + \frac{0.228}{\gamma^2}\right)} \quad (9)$$

$$\theta_D = \frac{h}{k_B} \left[\frac{3n}{4\pi} \left(\frac{\rho N_A}{M} \right) \right]^{\frac{1}{3}} v_m, \frac{1}{v_m} = \frac{1}{3} \left(\frac{1}{v_L} + \frac{2}{v_T} \right) \quad (10)$$

$$v_L = \left(\frac{B + 4/3G}{\rho} \right)^{1/2}, v_T = \left(\frac{G}{\rho} \right)^{1/2} \quad (11)$$

1.2.1.1 通过弹性模量求解格林艾森常数

用公式(12)可以通过弹性模量计算声速,进一步得到泊松比,然后得到格林艾森常数^[26]。

$$\gamma = \frac{3}{2} \left(\frac{1 + \nu}{2 - 3\nu} \right), \nu = \frac{1 - 2(v_T/v_L)^2}{2 - 2(v_T/v_L)} \quad (12)$$

但是用公式(12)计算的格林艾森常数往往会有一定的误差,2017年Jia等^[27]发展了一种新的格林艾森常数的算法,考虑了体积变化下的弹性模量变化,通过公式(13)~(15),得到平均声学格林艾森常数 γ_e 。其中, γ_e 可以通过纵向格林艾森常数 γ_L^e 与剪切格林艾森常数 γ_S^e 求解, V 是原胞体积。

$$\gamma_e = \sqrt{\frac{\left[(\gamma_L^e)^2 + 2(\gamma_S^e)^2 \right]}{3}} \quad (13)$$

$$\gamma_L^e = -\frac{1}{2} \frac{V}{B + \frac{4G}{3}} \frac{\partial \left(B + \frac{4G}{3} \right)}{\partial V} - \frac{1}{6} \quad (14)$$

$$\gamma_S^e = -\frac{1}{2} \frac{V}{G} \frac{\partial G}{\partial V} - \frac{1}{6} \quad (15)$$

1.2.1.2 通过吉布斯自由能求解格林艾森常数

除了弹性模量求解法,也可以通过吉布斯自由能得到格林艾森常数。AUTOMATIC GIBBS LIBRARY (AGL)自动化准谐德拜模型是一种以吉布斯自由能准谐近似为理论基础发展的近似模型,对于一定温度压力下的平衡体系,可以通过求解格林艾森常数来获得热导率^[13-14]。平衡体系的吉布斯自由能可以用公式(16)表示,其中 E 是总能, A_{vib} 是振动亥姆霍兹自由能(可以通过德拜温度用公式(17)求解), p 是压力, V 是体积, D 是有关 $\frac{\theta_D}{T}$ 的积分函数。在AGL理论中,格林艾森常数可以用公式(18)、(19)求解^[13]。仍然可以利用弹性模量求解德拜温度。

$$G(V; p, T) = E(V) + pV + A_{\text{vib}}(\theta_D(V); T) \quad (16)$$

$$A_{\text{vib}}(\theta_D; T) = nk_B T \left[\frac{9\theta_D}{8T} + 3 \ln \left(1 - \frac{1}{e^{\frac{\theta_D}{T}}} \right) - D \right] \quad (17)$$

$$p - p_{T=0} = \gamma \frac{U_{\text{vib}}}{V} \quad (18)$$

$$U_{\text{vib}} = nk_B T \left[\frac{9}{8} \frac{\theta_D}{T} + 3D \left(\frac{\theta_D}{T} \right) \right] \quad (19)$$

在AGL模型下,只要给定了温度、压力等条件,仅仅需要计算能量就能够预测材料的格林艾森常数,计算热导率,从而极大地减小了计算量,但是其精度相对较低。

1.2.2 最小热导率

在高温极限下,材料体系的热导率不得低于一个临界值 κ_{\min} ,根据公式(20)的Clarke模型^[28]和公式(21)的Cahill-Pohl模型^[29-30],均可以预测高温极限下的最小热导率 κ_{\min} ,其中, m 是原子数密度, E 是杨氏模量。这两种模型都是通过声速求解最小热导。Clarke模型中采取了杨氏模量求平均声速的方法,Cahill-Pohl模型中考虑了纵波声子与横波声子的贡献。Cahill-Pohl模型的声速可以用公式(11)求解,即最小热导率的两种预测方法均可以根据弹性模量得到。

$$\kappa_{\min}^{\text{Clarke}} = 0.87 k_B \left(\frac{N_A n \rho}{M} \right)^{2/3} \left(\frac{E}{\rho} \right)^{1/2} \quad (20)$$

$$\kappa_{\min}^{\text{Cahill}} = \frac{1}{2.48} k_B m^{2/3} (v_L + 2v_T) \quad (21)$$

1.2.3 结合机器学习研究热输运性质

热导率与材料的晶体结构、组成原子质量等密切相关。利用机器学习方法,从大量数据中学习得到热导率预测模型^[16],对目标体系进行快速分类、聚类、预测值回归估计。比如,常见的分类算法有逻辑回归算法(LR)、最近邻算法(KNN)和支持向量机算法(SVM)。回归算法有最小二乘回归(OLSR)、套索算法(LASSO)和梯度提升树(GBDT)。聚类算法有K均值聚类(Kmeans)、基于密度的DBSCAN聚类和层次聚类等。此外,可以使用皮尔逊相关性系数来分析特征描述符、特征属性和特征向量等特征变量之间以及特征变量与目标变量之间的关系^[31]。

一般来说,电子色散计算、声子色散计算、声子散射计算和电子散射计算的所用计算资源是依次升高的,目前高通量研究中唯一可以精确求解的只有电子色散关系,电输运研究的精度很大程度上取决于弛豫时间的处理方式。在热输运上,虽然无法精确求解声子谱与声子散射,但是体材料热输运的特点极大地取决于体系的声速^[32-34],而后者可以用简单的模量等参数代替。因此热输运高通量算法的准确性相对较高。下文将综述利用以上算法进行的第一性原理高通量热电材料研究工作。

2 高通量热电材料研究实例

2.1 用高通量评价热电材料的电输运性质

2.1.1 用常数弛豫时间近似研究电输运性质

基于常数弛豫时间近似的玻尔兹曼输运理论, Chen等^[12]以MP平台中48000多种材料体系为研究样本, 对其中约25000种半导体材料进行了电输运计算, 构成MP输运计算数据库。2017年Ricci等^[35]基于此电输运计算模块的数据, 结合MP数据库里面已有结构对晶体结构做了优化和能带计算的模块, 对其电学性质的总体分布做了比较细致的归纳。图3是温度为600 K, 载流子浓度为 10^{20} cm^{-3} 情况下, 所有研究的半导体材料的泽贝克系数、电导率、功率因子分布图^[35]。颜色越红, 代表功率因子越高, 电输运性质越好, 这种情况下一般泽贝克系数比较大。但是由于弛豫时间统一采用 10^{-14} s , 限制了电导率的预测准确性, Ricci等^[35]对于功率因子的预测结果误差较大。单独分析泽贝克系数预测结果, Ricci等^[35]预测的泽贝克系数的值与实验结果符合得比较好, 如果将计算所得的带隙调整到实验所得带隙数值上, 还可进一步提高预测精度。

根据MP常数弛豫时间近似计算电输运性质的结果, 2015年Zhu等^[36]发现了一类电性能比较优异的XYZ₂化合物, 比如从未有热电性能报道的TmAgTe₂, 预测其ZT值可以达到1.8。2016年Aydemir等^[37]对另一种XYZ₂类化合物YCuTe₂做了进一步的实验研究, 探究了Cu原子无序对于热导率的影响, 在780 K下, 实验所得Y_{0.96}Cu_{1.08}Te₂的ZT值达到0.75。

以上基于常数弛豫时间近似所做的工作, 可以定性地对材料功率因子高低进行评估, 指导发现电输运性质比较好的新体系。但是由于功率因子预测不准确以及实验工艺限制, 预测结果与实验结果会有误差。

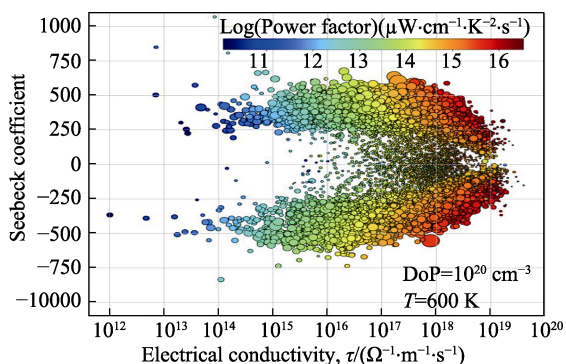


图3 MP常数弛豫时间近似电输运性质分布^[35]

Fig. 3 Electrical properties in MP with a constant relaxation time approximation^[35]

2.1.2 用高通量评价粉末烧结材料

2011年Wang等^[15]基于AFLOW平台, 着重研究了粉末烧结热电材料, 希望找到具有高热电转化效率的新型热电材料体系。粉末烧结热电材料具有较小的晶粒尺寸, 能够运用常数平均自由程近似方法, 得到载流子的平均自由程等于晶粒尺寸^[38]。通过常数平均自由程近似方法(见公式(3)和(4))计算粉末烧结材料电输运性质与晶粒尺寸的关系, 可以发掘具有较高功率因子的粉末烧结材料。还可以用于机器学习研究大量功率因子的预测数据, 分析功率因子与粉末烧结材料的一些内禀属性之间的关系, 通过内禀属性估测材料的电输运性质。

Wang等^[15]计算了2500种AFLOW数据库里的粉末烧结材料的功率因子与晶粒尺寸之比 P/L , 图4是20种性能比较好的n型与p型材料的平均 P/L 及

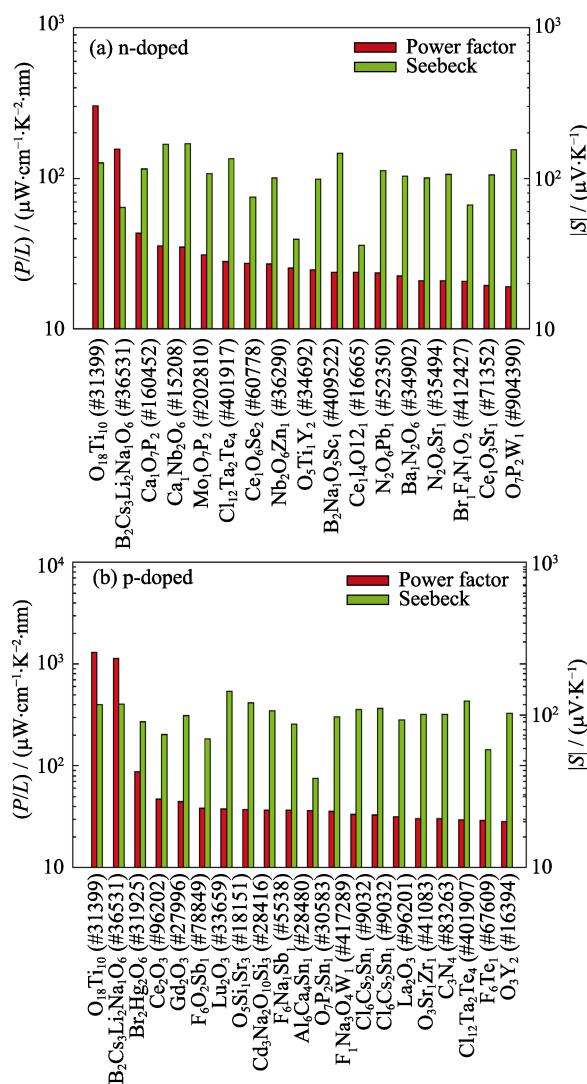


图4 n型掺杂(a)和p型掺杂(b)粉末烧结材料的平均 P/L 以及相应的泽贝克系数^[15]

Fig. 4 n- (a) and p-doped (b) average normalized power factor P/L and Seebeck coefficient^[15]

其对应的泽贝克系数,在各向异性材料中,平均 P/L 在最大 P/L 与最小 P/L 之间。对于这二十种材料体系, p 型材料的功率因子普遍比 n 型更高,热电性能更好。除此之外,通过数据回归分析, Wang 等^[15]还得到一些与热电性质有关的特征参量,譬如单胞内原子数越多、带隙越大、载流子有效质量越大,粉末烧结材料的功率因子越高。

由于纳米材料制备工艺对性质影响比较大,不同的制备工艺得到的纳米材料之间的性质存在差异,实验结果本身就具有不确定性,所以平均自由程近似的预测结果与实验结果也存在误差,但是其功率因子在 10^6 数量级,远大于功率因子误差的范围,可以用于定性预测电输运性质并进行机器学习分析。

2.1.3 基于常数电声耦合近似的高通量研究

2018 年 Xi 等^[20]基于上海大学 MIP 平台,开展了硫族类金刚石结构的热电材料预测工作,弛豫时间采用公式(7)的常数电声耦合近似法计算,电输运

参数与实验值可比。其中 $C_{n,k} = \frac{2\pi k_B T E_{\text{def}}^2}{V h E}$, E_{def} 是

带边的形变势常数, V 是原胞体积, E 是杨氏模量。通过带隙输运性能筛选, Xi 等^[20]预测了一类可能具有高性能的新组分化合物。该工作根据结构筛选条件,从 MIP 平台的 82412 个材料体系中选择具有 214 个硫族类金刚石热电材料进行电子结构计算,得到可以进行后续电输运性质分析的 161 个半导体材料。如图 5(a)是计算所得功率因子,空心点是带有空位的硫族类金刚石体系,圆圈内高功率因子体系大多数都含有空位,典型的带有空位(Vac, 符号[])的晶格结构见图 5(b)^[20]。结合大的原子量和适合的带隙范围等筛选条件, Xi 等^[20]发现一些譬如 $\text{Cd}_2\text{Cu}_3\text{In}_3\text{Te}_8$ 的新热电材料,实验验证其 ZT 值达到 1.0。

根据高通量结果的大数据分析, Xi 等^[20]发现在这一类材料中对电输运性质起主要作用的是硫族元素的 p 电子,即在这一类材料中存在着导电通道。他们定性分析了导电通道存在的原因,并对硫族类金刚石体系中 S、Se、Te 三类导电通道的电输运特性进行了比较分析,揭示了小样本无法得到的共性规律^[20]。

2.2 通过高通量评价热电材料的热输运性质

2.2.1 用 Slack 模型计算热导率实例

Slack 模型中,格林艾森常数可以用弹性矩阵、吉布斯自由能求解,AGL 理论属于后一种。AGL 以吉布斯准谐近似为理论基础^[14],引入吉布斯自由能,计算格林艾森常数。AGL 理论所需计算比较简单,适用于高通量。

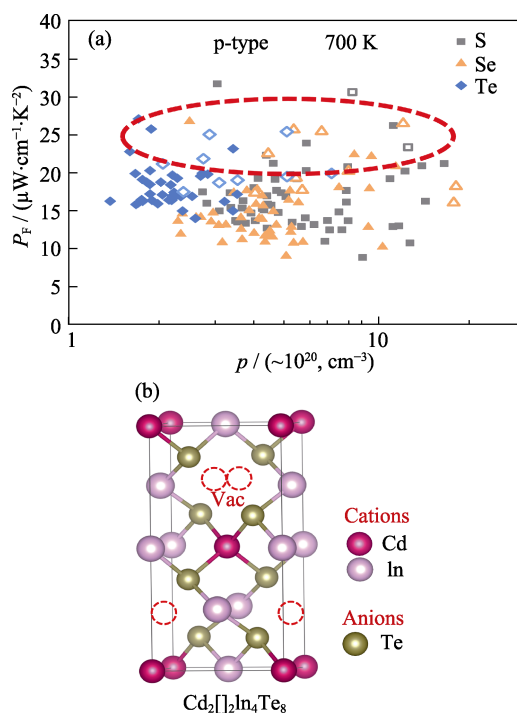


图 5 硫族类金刚石化合物的功率因子分布图(a)和带空位的三元硫族类金刚石结构图(b)^[20]

Fig. 5 Power factor of chalcogenides with diamond-like structures (a) and vacancy-containing ternary chalcogenides (b)^[20]

运用 AGL 理论, Toher 等^[13]基于 AFLOW 测试了 75 个包含金刚石结构、闪锌矿结构、盐岩结构、纤锌矿结构的材料体系以及 107 个半哈斯勒合金体系。图 6 是 AGL 预测热导率 κ_L^{AGL} 分别与第一性原理全非谐声子模拟计算热导率 κ_L^{anh} ^[16]、实验值热导率 κ_L^{exp} 的部分比较,灰色虚线是过原点斜率为 1 的直线。

除了四方体系, AGL 理论预测热导率与实验测量热导率之间的皮尔逊相关系数都比较高,主要是由于利用模量估算了德拜温度。传统晶格热导决定于热容、声速、平均自由程或声子弛豫时间。块体

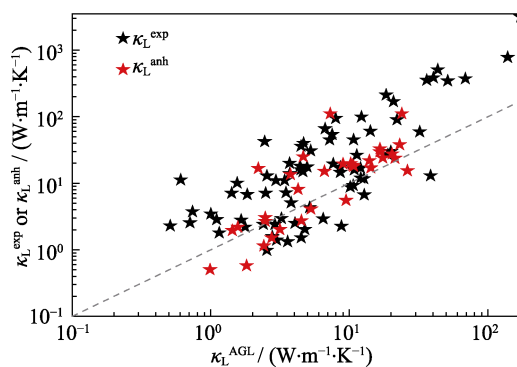


图 6 AGL 理论预测热导率与实验热导率以及全非谐声子热导率关系图(数据来源文献[13])

Fig. 6 Relationship between κ_L^{AGL} and $\kappa_L^{\text{anh}}/\kappa_L^{\text{exp}}$ (Data from Ref.[13])

材料对于声速有贡献的主要是声学声子,声学声子的声速可以用模量来表示,所以晶格热导的主要部分可以用模量估算德拜温度得到。Toher 等^[13]还通过预测材料体系的热导率,提出了几种可以用于热电材料的低热导率材料,譬如 AgI、CuI 等等。

2.2.2 用两种模型计算最小热导率

除了探讨电输运性质相关,Chen 等^[12]还基于 MP 用电输运预测了具有较高功率因子体系为热导率研究的样本,研究了热输运性质。采用 Clarke 模型^[28]和 Cahill-Pohl 模型^[29-30]来预测高温极限下的最小热导率 κ_{\min} ,这两种模型只需要比较少的计算量就能达到比较理想的效果。图 7 是两种模型最小热导率的预测结果与实验结果对比图^[12],蓝色实线是过原点斜率为 1 的直线。两种模型的预测结果相差不多,预测结果与实验结果线性关系比较好,大部分预测结果与实验结果的误差在合理范围。Chen 等^[12]还试图在多样化的数据中找到一些对于预测热电材料性质有用的特征描述符,例如带隙、原子半径、电负性等等,并采用密度聚类算法(DBSCAN)将 5431 种材料数据分成了六类,还分析了每一类中对热电性质影响比较大的关键描述符。

Chen 等^[12]的工作基于 MP 第一性原理计算平台,探索了热电材料电、热输运两方面的性质,集合了数据库技术、高通量计算手段和数据挖掘技术,是一项比较全面的高通量热电材料研究,研究手法比较成熟,研究思路也值得学习和借鉴。

2.2.3 结合机器学习研究半哈斯勒合金的热导率性质

目前关于半哈斯勒合金热电性质的高通量研究很多^[32,39-40],在结合高通量计算与机器学习的工作中,以 Carrete 等^[16]的研究最具代表性。他们对元素周期表中所有非放射性元素进行排列组合,得到约 79000 种半哈斯勒合金结构的材料。通过形成能计

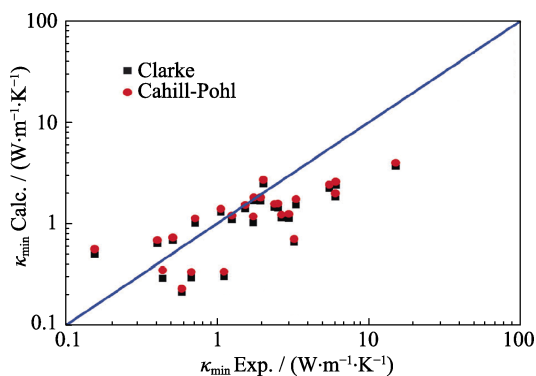


图 7 MP 弹性张量计算热导率与实验数据关系图^[12]

Fig. 7 Relationship between calculated κ_L and experimental κ_L ^[12]

算、声子色散计算以及热力学稳定性分析,得到 75 种稳定的半哈斯勒合金。结合机器学习,Carrete 等^[16]提出了三种方法来估算其热导率,分别是非谐力常数在晶格结构一样的体系(Mg_2Si)中具有可移植性方法,其预测热导率为 κ_{transf} ;随机回归森林模型方法,其预测热导率为 κ_{forest} ;高阶力常数机器学习模型方法,其预测热导率为 κ_{anh} 。

图 8(a)是 κ_{transf} 、 κ_{forest} 预测结果,横坐标是两个数量级跨度的热导率,纵坐标是预测体系出现在某热导率下的频率密度。 κ_{transf} 预测模型假设非谐力常数在结构相似的体系(Mg_2Si)中具有可移植性,计算的热导率主要在 $1\sim 100 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$ 范围内; κ_{forest} 预测模型采用第一性原理方法计算了 32 种体系的热导率作为训练集,采用随机森林回归模型^[41]来学习热导率与一些特征描述符之间的关系。这两种预测方法与非谐项没有相关性,影响了预测热导率的准确性,只能用于定性分析。为了更精准地描述热导率,Carrete 等^[16]又发展了第三种方法来探究原子间非谐力常数与热导率间的关系(图 8(b))。每种半哈斯勒合金中有 737 种独立的原子间非谐力常数,通过主成分分析法(PCA)^[42]分析,这 737 种非谐力常数仅有四个比较重要,其它非谐力常数可以用四个参

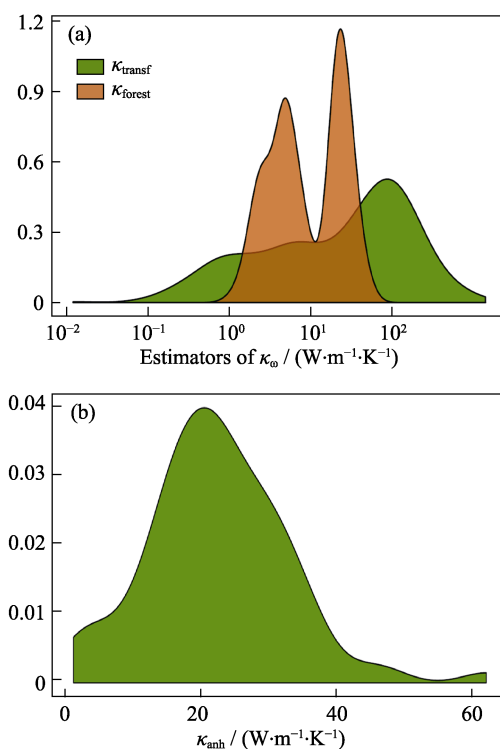


图 8 半哈斯勒合金热导率性质三种预测结果^[16]

Fig. 8 Three prediction models of the κ_L in half-Heusler compounds^[16]

(a) Frequency densities of the estimators of thermal conductivity at 300 K κ_{transf} and κ_{forest} ; and (b) distribution of κ_{anh} over the 75 thermodynamically stable half-Heuslers

量的线性模型来表示。这四个参量可以用 16 个密度泛函计算得到,极大地简化了三阶力常数的计算。在这 75 种热力学稳定体系中,可能存在三个热导率低于 $3 \text{ W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$ 的材料。

虽然机器学习难以解释材料性质的本征机理,但是该工作降低了最耗时间的三阶力常数计算成本,达到很高的计算精度。机器学习预测模型的简单化和多样化,拓展了研究方法,有利于挖掘特征描述符与目标变量之间的关系,促进了热电材料的第一性原理高通量研究。

3 结束语

本文综述了最近几年第一性原理高通量研究方法在热电材料中的应用,从电、热输运两方面介绍了其对热电材料快速发展的影响,并且介绍了机器学习与高通量结合的部分工作。常用的电输运高通量简化算法有常数弛豫时间近似、常数平均自由程近似、常数电声耦合近似。Chen、Ricci、Zhu 等^[12,35-36]均采用常数弛豫时间近似做了高通量搜索,Zhu 等^[36]还提出了 XYZ_2 类新型热电材料,并做了实验分析^[37]。Wang 等^[15]采用常数平均自由程近似对纳米材料做了高通量搜索,得到了 20 种电输运性质比较好的材料。这两种简化算法减小了电输运计算的计算量,但牺牲了计算精度,只能对电输运性质做定性分析。Xi 等^[20]采用常数电声耦合近似筛选了一类具有本征缺陷结构的新型硫族类金刚石化合物,并实验验证其具有较好的热电性能,实现了从理论预测到实验实现的高通量筛选新材料全过程。该方法兼顾了计算的精确性与效率,将对未来高性能新型热电材料筛选产生重大影响。常用的热输运高通量简化算法有 Slack 模型、Clarke 模型和 Cahill-Pohl 模型。Slack 模型可以计算有限温度下材料晶格热导率,针对 Slack 模型下的格林艾森常数可以通过弹性模量或者 AGL 理论得到,其预测精确度比电输运简化算法稍高。本文选取了 Toher 等^[13]的工作为例,指出了 AgI、CuI 材料具有比较好的热输运性质。用 Clarke 模型和 Cahill-Pohl 模型可以计算高温极限下的最小热导率,Chen 等^[12]采用此方法做了高通量研究,其预测结果较为准确。热导率计算的资源耗费较大,以利用模量来计算声速(或德拜温度)的高通量热输运简化方法有利于高速预测材料的晶格热导率;但是这些模型不适用于低温情况。大数据的产生便于引入机器学习方法来研究热电材料,Carrete 等^[16]发展了三种机器学习模型来预测半哈斯勒合金的热导率,找到了三种具有较低热

导率的材料。

随着理论与技术的发展,各种高通量计算平台应运而生,高通量热电研究也会逐渐成熟。为了使高通量研究与热电材料更紧密地结合,未来需要提出准确性更高的电热输运简化算法以得到更加准确的计算结果。在数据分析方面,针对大数据结果,加深机器学习与传统热电材料性质分析手段的结合,以便更好地得到热电材料输运性质的构效关系。

4 附录

Materials Project (MP)是由 Ceder, Persson 等人牵头,美国能源部(DOE)、劳伦斯伯克利国家实验室(LBNL)、储能研究联合中心(JCESR)、美国国家自然科学基金会(NSF)支持,多家单位,多位合作者共同开发的以 Vienna *Ab Initio* Simulation Package(VASP)第一性原理计算软件包为主要研究手段^[43],结合材料应用程序接口(MAPI)^[44]和 Open-source Python Materials Genomics (pymatgen)^[45]等材料数据分析手段的高通量平台^[46]。它的结构数据存储量几乎涵盖了无机晶体结构数据库的所有数据,并且计算了能量、磁性、能带、态密度、声子、X 射线衍射、相图、弹性、表面能等性质,数据量比较全,计算内容比较丰富。结合数据挖掘手段,MP 也做了关于电池、透明导体氧化物和热电材料等的研究^[35,47-50]。

Automatic-FLOW for Materials Discovery (AFLOW)是由美国杜克大学 Curtarolo 教授团队、Nardelli 教授团队和美国杨百翰大学 Hart 教授团队等联合开发的以 VASP^[43]和 Quantum Espresso (QE)^[51]为主要研究手段的高通量计算数据库平台^[52]。除了基础的数据结构以及基本性质计算信息,AFLOW 还包含了各种模块。譬如结合机器学习算法来预测材料性质的模块 AFLOW-ML、二元合金库 Binary alloy library、超级合金搜索模块 Superalloys search、与计算相关模块 PAOFLOW 和 AFLOW π ^[53]等等。AFLOW 还支持在线提交计算任务,计算模式采用高通量统一标准,保证了计算快速性的同时还具有一定的准确性。

除了美国的上述两个数据库平台外,欧洲、亚洲等也相继开发了各自的数据库平台。NOMAD 是欧洲以数据共享为目的开发的计算数据库平台,内容涵盖了各种第一性原理计算软件支持的数据类型,支持用户上传下载各种计算软件过程中的输入输出数据文件,这与材料基因理念中的数据由内部生成的模式不太一样,但是庞大的数据量也使其成为一个较好的数据库平台。

MatNavi 是由日本国家材料研究所(NIMS)徐一斌团队开发的包含聚合物数据库、无机材料数据库、计算相图数据库、计算电子结构数据库、放射性物质净化吸附剂数据库、中子嬗变数据库、界面热导率数据库、扩散数据库、超导材料数据库等数据库合集的平台,内容涵盖了实验数据与计算数据。MatNavi 的第一性原理计算数据库模块(CompES-x)采用的计算软件是 VASP,也具有固定的计算参数标准。同上述所有数据库一样,MatNavi 也是免费的。

MatCloud 是由中国科学院杨小渝教授团队开发的以提供第一性原理计算资源为主的收费式国产高通量计算平台,目前仅支持 VASP 计算。其运用网页提交计算作业模式,简化了操作过程,并且采用图形交互将计算结果直接显示在网页上,支持用户上传下载数据。MatCloud 的主要功能包括晶体结构建模、图形化界面的流程设计、电子结构以及力学等多种物理化学性质预测等。其计算过程中使用的所有数据同样支持上传与下载。

上述数据库或者计算平台虽然有少数涉及了热电材料的研究,但是并没有出现一个针对热电材料设计的数据库或者计算平台。由上海大学材料基因组工程研究院、上海大学计算机工程与科学学院等开发的MIP是针对热电材料研究的第一性原理高通量计算平台。MIP以大量基础晶体数据库为原始数据,以 VASP 为主要计算软件,对材料进行性质计算和预测。除了最基础的电子结构、磁性、费米面、弹性等计算数据外,MIP还采用 Transoptic 软件包(<http://www.mgi.shu.edu.cn/Portals/675/transoptic.zip>)计算了材料的泽贝克系数和功率因子等电输运性质。该程序采用跃迁矩阵元方法有效避免能带交叉问题,并利用常数电声耦合近似处理弛豫时间,比常数弛豫时间近似精确度更高。

参考文献:

- [1] SEEBECK T J. On the magnetic polarization of metals and minerals by temperature differences. *Annals of Physics*, 1826, **82(3)**: 253–286.
- [2] PELTIER J C A. New experiments on the heat effects of electric currents. *Annals of Chemistry and Physics*, 1834, **56**: 371–386.
- [3] ZHANG Q, LIAO J, TANG Y, *et al.* Realizing a thermoelectric conversion efficiency of 12% in bismuth telluride/skutterudite segmented modules through full-parameter optimization and energy-loss minimized integration. *Energy & Environmental Science*, 2017, **10(4)**: 956–963.
- [4] BULMAN G E, SIIVOLA E, SHEN B, *et al.* Large external delta t and cooling power densities in thin-film Bi_2Te_3 -superlattice thermoelectric cooling devices. *Applied Physics Letters*, 2006, **89(12)**: 122117–1–3.
- [5] SHAKOURI A, ZHANG Y. On-chip solid-state cooling for integrated circuits using thin-film microrefrigerators. *IEEE Transactions on Components and Packaging Technologies*, 2005, **28(1)**: 65–69.
- [6] WANG W, JIA F, HUANG Q, *et al.* A new type of low power thermoelectric micro-generator fabricated by nanowire array thermoelectric material. *Microelectronic Engineering*, 2005, **77(3/4)**: 223–229.
- [7] LI JING-FENG. Macrofabrication technology of three-dimensional microdevices and their MEMS applications. *Journal of Inorganic Materials*, 2002, **17(4)**: 657–664.
- [8] HAUTIE G, JAIN A, CHEN H, *et al.* Novel mixed polyanions lithium-ion battery cathode materials predicted by high-throughput *ab initio* computations. *Journal of Materials Chemistry*, 2011, **21(43)**: 17147–17153.
- [9] DE JONG M, CHEN W, ANGSTEN T, *et al.* Charting the complete elastic properties of inorganic crystalline compounds. *Sci. Data*, 2015, **2**: 150009–1–13.
- [10] TAYLOR R H, CURTAROLO S, HART G L W. Guiding the experimental discovery of magnesium alloys. *Physical Review B*, 2011, **84(8)**: 084101–1–17.
- [11] HAUTIER G, FISCHER C, EHRLACHER V, *et al.* Data mined ionic substitutions for the discovery of new compounds. *Inorg. Chem.*, 2011, **50(2)**: 656–663.
- [12] CHEN W, POHLS JAN-HENDRIK, HAUTIER G, *et al.* Understanding thermoelectric properties from high-throughput calculations: trends, insights, and comparisons with experiment. *Journal of Materials Chemistry C*, 2016, **4(20)**: 4414–4426.
- [13] TOHER C, PLATA J J, LEVY O, *et al.* High-throughput computational screening of thermal conductivity, debye temperature, and gruneisen parameter using a quasiharmonic debye model. *Physical Review B*, 2014, **90(17)**: 174107–1–14.
- [14] BLANCO M, FRANCISCO E, LUANA V. Gibbs: isothermal-isobaric thermodynamics of solids from energy curves using a quasi-harmonic debye model. *Computer Physics Communications*, 2004, **158(1)**: 57–72.
- [15] WANG S, WANG Z, SETYAWAN W, *et al.* Assessing the thermoelectric properties of sintered compounds *via* high-throughput *ab-initio* calculations. *Physical Review X*, 2011, **1(2)**: 021012–1–8.
- [16] CARRETE J, LI W, MINGO N, *et al.* Finding unprecedentedly low-thermal-conductivity half-Heusler semiconductors *via* high-throughput materials modeling. *Physical Review X*, 2014, **4(1)**: 011019–1–9.
- [17] GOLDSMID H, DOUGLAS R. The use of semiconductors in thermoelectric refrigeration. *British Journal of Applied Physics*, 1954, **5(11)**: 386–390.
- [18] CHASMAR R, STRATTON R. The thermoelectric figure of merit and its relation to thermoelectric generators. *International Journal of Electronics*, 1959, **7(1)**: 52–72.
- [19] SLACK G A. Nonmetallic crystals with high thermal conductivity. *Journal of Physics & Chemistry of Solids*, 1973, **34(2)**: 321–335.
- [20] XI L, PAN S, LI X, *et al.* Discovery of high performance thermoelectric chalcogenides through reliable high throughput material screening. *Journal of the American Chemical Society*, 2018, **140(34)**: 10785–10793.
- [21] YANG J, XI L, QIU W, *et al.* On the tuning of electrical and thermal transport in thermoelectrics: an integrated theory–experiment perspective. *npj Computational Materials*, 2016, **2**: 15015–1–17.
- [22] GIBBS Z M, RICCI F, LI G, *et al.* Effective mass and Fermi surface complexity factor from *ab initio* band structure calculations. *npj Computational Materials*, 2017, **3(1)**: 8–1–7.
- [23] CHEN LI-DONG, XIONG ZHEN, BAI SHENG-QIANG. Recent

- progress of thermoelectric nano-composites. *Journal of Inorganic Materials*, 2010, **25**(6): 561–568.
- [24] YAN J, GORAI P, ORTIZ B, *et al.* Material descriptors for predicting thermoelectric performance. *Energy & Environmental Science*, 2015, **8**(3): 983–994.
- [25] ANDERSON ORSON L. A simplified method for calculating the debye temperature from elastic constants. *Journal of Physics and Chemistry of Solids*, 1963, **24**(7): 909–917.
- [26] HILL RICHARD. The elastic behaviour of a crystalline aggregate. *Proceedings of the Physical Society. Section A*, 1952, **65**(5): 349–354.
- [27] JIA T, CHEN G, ZHANG Y. Lattice thermal conductivity evaluated using elastic properties. *Physical Review B*, 2017, **95**(15): 155206–1–6.
- [28] CLARKE D R. Materials selection guidelines for low thermal conductivity thermal barrier coatings. *Surface and Coatings Technology*, 2003, **163**: 67–74.
- [29] CAHILL D G, POHL R. Lattice vibrations and heat transport in crystals and glasses. *Annual Review of Physical Chemistry*, 1988, **39**(1): 93–121.
- [30] CAHILL D G, BRAUN P V, CHEN G, *et al.* Nanoscale thermal transport. II. 2003–2012. *Applied Physics Reviews*, 2014, **1**(1): 011305–1–45.
- [31] HAUKE J, KOSSOWSKI T. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae*, 2011, **30**(2): 87–93.
- [32] YANG J, LI H, WU T, *et al.* Evaluation of half-Heusler compounds as thermoelectric materials based on the calculated electrical transport properties. *Advanced Functional Materials*, 2008, **18**(19): 2880–2888.
- [33] YING P, LI X, WANG Y, *et al.* Hierarchical chemical bonds contributing to the intrinsically low thermal conductivity in α -MgAgSb thermoelectric materials. *Advanced Functional Materials*, 2017, **27**(1): 1604145–1–8.
- [34] LI W, LIN S, GE B, *et al.* Low sound velocity contributing to the high thermoelectric performance of Ag_8SnSe_6 . *Advanced Science*, 2016, **3**(11): 1600196–1–7.
- [35] RICCI F, CHEN W, AYDEMIR U, *et al.* An *ab initio* electronic transport database for inorganic materials. *Sci. Data*, 2017, **4**: 170085–1–13.
- [36] ZHU H, HAUTIER G, AYDEMIR U, *et al.* Computational and experimental investigation of TmAgTe_2 and XYZ_2 compounds, a new group of thermoelectric materials identified by first-principles high-throughput screening. *Journal of Materials Chemistry C*, 2015, **3**(40): 10554–10565.
- [37] AYDEMIR U, P HLS J, ZHU H, *et al.* YCuTe_2 : a member of a new class of thermoelectric materials with CuTe_2 -based layered structure. *Journal of Materials Chemistry A*, 2016, **4**(7): 2461–2472.
- [38] BERA C, SOULIER M, NAVONE C, *et al.* Thermoelectric properties of nanostructured $\text{Si}_{1-x}\text{Ge}_x$ and potential for further improvement. *Journal of Applied Physics*, 2010, **108**(12): 124306–1–8.
- [39] ZIOLKOWSKI P, WAMBACH M, LUDWIG A, *et al.* Application of high-throughput seebeck microprobe measurements on thermoelectric half-Heusler thin film combinatorial material libraries. *ACS Combinatorial Science*, 2018, **20**(1): 1–18.
- [40] CARRETE J, MINGO N, WANG S D, *et al.* Nanograined half-Heusler semiconductors as advanced thermoelectrics: an *ab initio* high-throughput statistical study. *Advanced Functional Materials*, 2014, **24**(47): 7427–7432.
- [41] LIAW A, WIENER M. Classification and regression by random forest. *R News*, 2002, **23**(23): 18–22.
- [42] JOLLIFFE I T. Principal component analysis. Berlin, Heidelberg: Springer, 2011: 1094–1096.
- [43] KRESSE G, FURTHMULLER J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Physical Review B*, 1996, **54**(16): 11169–11186.
- [44] ONG S P, CHOLIA S, JAIN A, *et al.* The materials application programming interface (API): a simple, flexible and efficient API for materials data based on representational state transfer (rest) principles. *Computational Materials Science*, 2015, **97**: 209–215.
- [45] ONG S P, RICHARDS W D, JAIN A, *et al.* Python materials genomics (pymatgen): a robust, open-source python library for materials analysis. *Computational Materials Science*, 2013, **68**: 314–319.
- [46] JAIN A, ONG S P, HAUTIER G, *et al.* The materials project: a materials genome approach to accelerating materials innovation. *APL Materials*, 2013, **1**(1): 011002–1–11.
- [47] ZHOU F, COCCIONI M, MARIANETTI C A, *et al.* First-principles prediction of redox potentials in transition-metal compounds with LDA + *U*. *Physical Review B*, 2004, **70**(23): 235021–1–8.
- [48] WANG L, MAXISCH T, CEDER G. A first-principles approach to studying the thermal stability of oxide cathode materials. *Chemistry of Materials*, 2007, **19**(3): 543–552.
- [49] ONG S P, JAIN A, HAUTIER G, *et al.* Thermal stabilities of delithiated olivine MPO_4 ($\text{M} = \text{Fe}, \text{Mn}$) cathodes investigated using first principles calculations. *Electrochemistry Communications*, 2010, **12**(3): 427–430.
- [50] ADAMS S, RAO R P. High power lithium ion battery materials by computational design. *Physica Status Solidi a-Applications and Materials Science*, 2011, **208**(8): 1746–1753.
- [51] GIANNOZZI P, BARONI S, BONINI N, *et al.* Quantum espresso: a modular and open-source software project for quantum simulations of materials. *Journal of Physics-Condensed Matter*, 2009, **21**(39): 395502–1–19.
- [52] ISAYEV O, OSES C, TOHER C, *et al.* Universal fragment descriptors for predicting properties of inorganic crystals. *Nat. Commun.*, 2017, **8**: 15679–1–12.
- [53] SUPKA A R, LYONS T E, LIYANAGE L, *et al.* AFLOW $_{\pi}$: a minimalist approach to high-throughput *ab initio* calculations including the generation of tight-binding hamiltonians. *Computational Materials Science*, 2017, **136**: 76–84.