

数据论文 • 封面文章

纤锌矿铁电材料自发极化强度的本征影响因素*

康瑶¹⁾ 陈健¹⁾ 童祎²⁾ 王新朋²⁾ 段坤¹⁾ 王嘉琪¹⁾王旭东¹⁾ 周大雨^{1)†} 姚曼^{1)‡}

1) (大连理工大学材料科学与工程学院, 大连 116081)

2) (苏州实验室, 苏州 215123)

(2024 年 10 月 30 日收到; 2024 年 11 月 28 日收到修改稿)

自发极化强度是衡量铁电材料极化能力的关键指标. 新兴的纤锌矿铁电材料因较高的自发极化而受到广泛关注, 但目前对影响这一性质的关键因素的理解仍然不足. 本文旨在通过结合机器学习和第一性原理方法来解决这一问题. 首先, 计算了 40 种二元和 89 种简单三元纤锌矿材料的自发极化强度, 并从元素基本属性、晶体结构参数和电子性质中提取了多种特征. 随后, 采用 Boruta 算法和距离相关系数分析方法进行特征筛选, 提出了一个全面而精确的纤锌矿材料自发极化强度的机器学习预测模型. 进一步借助 SHapley Additive exPlanations 分析方法, 阐明了影响自发极化强度的关键因素是阳离子离子势的均值 IPi_Aave 和晶胞参数 a 等. 本研究弥补了目前对自发极化强度多因素的影响机制理解的缺乏, 为系统评估新兴纤锌矿材料的自发极化强度提供了帮助, 有助于加快性能优异的纤锌矿铁电材料的筛选. 本文数据集可在 <https://www.doi.org/10.57760/sciencedb.j00213.00073> 中访问获取.

关键词: 纤锌矿铁电材料, 自发极化强度, 机器学习, 第一性原理

PACS: 77.55.-g, 63.20.dk, 82.20.Wt, 42.25.Ja

DOI: 10.7498/aps.74.20241520

CSTR: 32037.14.aps.74.20241520

1 引言

铁电材料在面世后的一百年间, 凭借其优异的物理化学性能, 吸引了大量科研工作者的兴趣, 目前已广泛应用于铁电随机存取存储器 (ferroelectric random access memory, FeRAM)^[1,2]、铁电场效应晶体管 (ferroelectric field effect transistor, FeFET)^[3]、压电致动器^[4]、热电传感器^[5] 和光伏电池^[6] 等领域. 铁电体物理学研究的一个核心问题是自发极化, 描述这种极化能力的大小可以用自发极化强度 (spontaneous polarization magnitude, P_s)

这一关键指标来衡量. 自发极化的出现在晶体中造成了一个特殊方向, 每个晶胞中原子的构型使正负电荷重心沿该方向发生相对位移, 形成电偶极矩. 这一方向是在晶体所属点群的任何对称操作下都保持不动的方向, 32 个晶体学点群中仅有 10 个具有唯一的、在晶体所属点群的任何对称元素操作下都无法与其他方向重合极轴的极性点群可能存在具备自发极化性能的材料^[7]. 而铁电材料是指除了自发极化带来的压电和热电性能^[8,9], 还可以在电场作用下改变极化取向的一类特殊材料.

纤锌矿材料隶属于 6mm 极性点群, 自 2019 年 Fichtner 等^[10] 首次制备得到具有标准电滞回线的

* 国家自然科学基金 (批准号: 51974056, 51474047, 52472120)、中央高校基本科研业务费 (批准号: DUT24LAB117) 和苏州实验室科研项目 (批准号: SK-1202-2024-012) 资助的课题.

† 通信作者. E-mail: zhoudayu@dlut.edu.cn

‡ 通信作者. E-mail: yaoman@dlut.edu.cn

AlScN 材料后, 纤锌矿材料作为新兴的铁电材料, 获得了广泛的关注. 该研究发现, AlScN 材料的 P_s 比传统铁电材料高几倍^[11]. 后续, 对同属纤锌矿结构的 AlBN^[12], AlYN^[13], GaScN^[14], ZnMgO^[15] 等材料的进一步研究表明纤锌矿材料普遍具有较高的 P_s . 这种较高的 P_s 可以提高铁电材料极化的调节范围, 给 FeRAM 等器件带来更好的性能, 但过高的 P_s 又会在 FeFET 中导致显著的去极化场, 使得保留率较差^[16]. 因此需要选择合适的 P_s 来平衡铁电器件性能和稳定性的需求, 这要求我们对 P_s 有更强的调控能力. 近年来, 多项研究结果表明, AlScN 材料的 P_s 随着 Sc 含量的增大而降低^[10,17-19], 但除了考虑 Sc 元素的掺入会提高 AlScN 的结构平整程度进而降低 P_s 外, 少有文献进一步解释 Sc 含量影响 P_s 的原因, 目前仍缺乏对于影响纤锌矿铁电材料 P_s 关键因素的深入认知.

为了进一步得到影响纤锌矿铁电材料 P_s 的本征因素, 本文结合第一性原理计算和多种机器学习模型, 在 129 种简单的二元或三元纤锌矿材料中, 寻求纤锌矿材料组成元素基本属性、晶体结构、电子性质等因素与 P_s 的关系, 构建纤锌矿材料 P_s 数据集, 并提出精准预测纤锌矿铁电体系 P_s 的机器学习模型, 分析得到影响铁电性能 P_s 的多种关键因素, 以求在材料选择层面上, 进一步提高对 P_s 的调控能力.

2 计算方法

所有的计算均采用基于密度泛函理论 (density functional theory, DFT) 的第一性原理方法, 并使用维也纳从头算模拟包 (VASP)^[20,21] 进行. 电子的交换关联作用采用广义梯度近似 (GGA-PBE) 的 Perdew-Burke-Ernzerhof 形式来解释^[22]. 截断能被设置为 500 eV, 并且对于选定的二元纤锌矿材料单胞 ($N = 4$) 使用 $15 \times 15 \times 15$ k 点网格的 Γ 中心 Monkhorst-Pack 网格. 这里二元纤锌矿材料的初始结构模型是从 Materials Project 数据库^[23] 或通过类似的结构进行原子置换获得的, 在结构优化中晶格和原子位置得到了充分的弛豫. 总能量波动降低至 10^{-6} eV 以内, 并且残余力小于 10 meV/Å 的体系被认为是收敛的.

本文采用基于现代极化理论的 Berry 相方法计算 P_s , 该方法认为 P_s 是极性相与参考相之间形

式极化的差异, 参考相指的是宏观极化和形式极化均为零的对称结构. 早在 2016 年, Dreyer 等^[24] 就提出计算纤锌矿材料 P_s 的中间参考顺电相应应该选取层状六方结构, 而不是具有非零极化的闪锌矿结构. 上述假设也在 Fichtner 等^[10] 的工作中得到了验证, 使用层状六方结构作为参考相的计算结果和实验测量结果的量级相同. 并且 Fichtner 等^[25] 在 2024 年提出, 参考结构的选择并不依赖于特定的实验现象或原子运动路径, 而是具有一定的普遍性. 在计算纤锌矿材料的 P_s 时, 层状六方结构作为参考结构是普遍适用的, 即便在实际的铁电材料中, 极化翻转的原子路径和顺序可能不同^[26-29]. 但是 Berry 相存在 2π 的模糊性, 这导致了极化值存在极化量子模糊性. 极化量子可被计算为 $e\mathbf{R}/\Omega$, 其中 e 是元电荷, $\mathbf{R}$ 是晶格矢量, Ω 是晶胞体积^[30]. 为了避免极化量子对 P_s 结果的影响, 自发极化是通过识别全极化的连续演化作为归一化位移的函数严格确定. 并且对于中间参考结构存在的非绝缘现象, 使用正负极化相的极化强度差异的一半来代表体系的 P_s .

为了获得纤锌矿材料的特征参数, 考虑了元素基本属性、晶体参数和电子性质这 3 个方面. 并且采用 Boruta 算法^[31] 进行特征选择, 解决特征的高维问题, 简化模型的复杂性, 增强模型的泛化能力. Boruta 算法是一种基于随机森林的特征选择方法, 通过使用影子特征和迭代过程来确定重要特征, 能够找到最优的特征组合, 而不只是单个特征的排序或删除. 并且为了得到对 P_s 存在关键影响的因素, 同时考虑特征对 P_s 以及特征之间的相关关系, 采用 Pearson 相关系数 (Pearson correlation coefficient, PCC) 和距离相关系数 (distance correlation coefficient, DCC) 来衡量两个变量之间的关系, 其计算公式如 (1a) 式和 (1b) 式所示:

$$\text{PCC}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}, \quad (1a)$$

$$\text{DCC}(X, Y) = \frac{\text{dCov}(X, Y)}{\sqrt{\text{dVar}(X) \cdot \text{dVar}(Y)}}, \quad (1b)$$

式中, $\text{Cov}(X, Y)$ 是特征之间的协方差, $\text{Var}(X)$, $\text{Var}(Y)$ 则分别代表特征的方差, $\text{dCov}(X, Y)$ 则表示距离协方差. Pearson 相关系数可以分析两个变量间的线性相关关系, 而距离相关系数可以分析两个变量间的独立性.

后续借助 6 种经典的机器学习方法, 核岭回归 (kernel ridge regression, KRR)、LASSO 回归 (least absolute shrinkage and selection operator regression, LASSO)、支持向量机回归 (support vector regression, SVR)、人工神经网络 (artificial neural networks, ANNs)、随机森林回归 (random forest regression, RFR) 和极端梯度提升回归 (extreme gradient boosting regression, XGBR), 通过五折交叉验证方法, 与统一的训练集、测试集的 4:1 划分, 以尽量避免数据划分的随机性, 并借助网格搜索方法寻求最佳的超参数组合, 提升模型的泛化能力, 对纤锌矿材料的 P_s 进行预测. 为了比较模型的预测效果, 采用与 P_s 单位相同, 易于解释且对异常点敏感的均方根误差 (root mean squared error, RMSE) 和表示模型解释效果的决定系数 R^2 指标来评估模型的能力, 其公式如 (2a) 式和 (2b) 式所示:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (2a)$$

$$R^2 = 1 - \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}. \quad (2b)$$

最后凭借模型评估指标进行模型的选择, 对预测效果最佳的模型, 利用 SHAP (SHapley Additive exPlanations) 分析方法 [32] 进行特征重要性分析. SHAP 分析方法将模型的预测效果看作是特征合作博弈的结果, 从而量化每个特征对模型预测的边际贡献, 为模型提供全局和局部的可解释性. 通过分析各特征 SHAP 值的分布和 SHAP 绝对值的均值等, 可以研究关键影响因素与 P_s 之间的复杂关系.

3 结果与讨论

3.1 数据集构建

纤锌矿材料是由阴阳离子分别组成两套沿 z 轴平行嵌套的六方密集堆积结构的材料体系, 其元素组成范围较广, 所以在进行大规模铁电性能计算之前, 需要确定哪些元素组成可以用作候选的纤锌矿铁电材料. 本文从纤锌矿材料或已知组成元素

的元素周期表相邻元素中提取了 17 种阳离子和 11 种阴离子, 如图 1(a) 所示, 并考虑了更为常见的阴阳离子外层电子数之和为 8 的离子配对关系 (这一关系已在多项工作 [33,34] 中得到证明), 组合成 64 个可能存在纤锌矿结构的初始体系. 后续数据集的筛选流程如图 1(b) 所示, 在具有纤锌矿结构的基础上, 进一步考虑了带隙、热力学稳定性等筛选条件后, 得到了 40 种二元纤锌矿材料. 我们用一条直线代表一组合理的纤锌矿材料阴阳离子组合, 汇总于图 1(c). 带隙和热力学稳定性的详细筛选过程分别示于补充材料表 S1 (online) 和图 S1 (online). 值得注意的是, 因 P_s 的计算对体系绝缘性的需求较高, 本文并不是以带隙等于 0 eV 为划分条件, 而是保留了其中带隙大于 0.03 eV 的材料体系, 即使 PBE 泛函计算低估了带隙, 但仍然可以定性地筛选出具有介电性质的材料.

当然, 仅考虑二元纤锌矿体系是不够的, 这没有考虑到其他元素掺入对 P_s 的影响, 为了提升数据集内体系的多样性, 以求得到更全面的 P_s 影响因素, 本文对上述二元体系进行元素替换, 得到了简单的三元纤锌矿结构材料. 三元纤锌矿结构的构建方法是在二元纤锌矿 AB 中 (为了叙述的清晰, 本文将阳离子标记为 A, 阴离子标记为 B, 原子结构示意图如图 1(a) 内插图所示), 替换其中一个阳离子 A, 构建 A_1A_2B 体系, 相应地替换一个阴离子 B, 构建 AB_1B_2 体系. 类似的简单三元纤锌矿模型也同样应用于 Sun 等 [35] 的工作中, 证实了此种规律掺杂的可靠性. 本文替换元素的选取原则是, 若要用 A_2 替换 A_1 , 则需保证 A_1B_1 , A_2B_1 均位于图 1(c) 展示的 40 种二元纤锌矿材料内. 经上述体系构建操作后, 共得到 60 种 A_1A_2B 和 51 种 AB_1B_2 三元纤锌矿体系. 需要注意的是, 这里考虑了常见的 Sc 元素作为额外的合金元素. 同样对上述三元纤锌矿体系进行带隙筛选, 共得到了 89 种候选的三元纤锌矿体系, 具体的带隙数值可见补充材料表 S2 (online). 这里的三元纤锌矿体系仅作为理论模型存在, 所以并未过多考虑热力学稳定性等实际应用价值.

本文针对上述 40 种二元纤锌矿体系和 89 种三元纤锌矿体系, 借助 Berry 相方法计算各材料的 P_s , 得到 AlN 材料的 P_s 为 1.326 C/m², 该结果与 Ye 等 [18] (1.328 C/m²) 和 Dreyer 等 [24] (1.350 C/m²) 的计算较为吻合, 其余二元纤锌矿体系的 P_s 计算

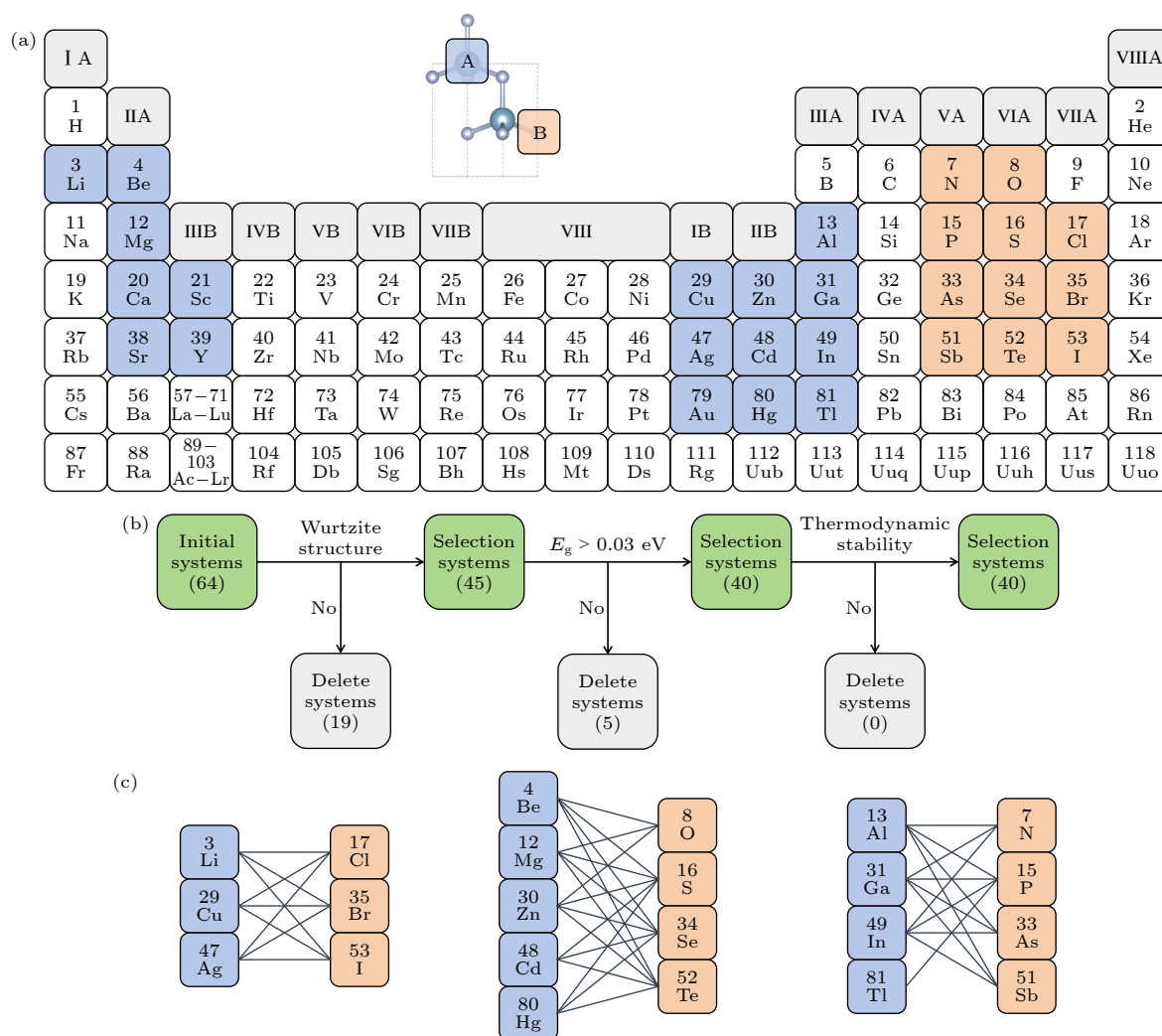


图1 (a) 初始筛选元素在元素周期表内的体现, 其中蓝色代表阳离子, 橙色代表阴离子; (b) 元素筛选的条件和具体流程; (c) 二元纤锌矿材料合理的元素组成

Fig. 1. (a) The representation of initial screening elements in the periodic table, where blue represents cations and orange represents anions; (b) the conditions and specific process for screening; (c) reasonable composition elements of binary wurtzite materials obtained through screening.

值与以往研究结果的对比汇总于补充材料表 S3 (online), 较高的一致性表明本文计算的数据集具备一定的合理性和准确性. 另外, BeO (1.217 C/m^2), GaN (1.311 C/m^2), InN (1.024 C/m^2) 等材料均具有和 AlN 相似的 P_s , 其中无毒性且具备较高带隙的 GaN 材料同样也是优秀的铁电候选材料.

为了更清晰地展示所有体系的 P_s 结果, 将所有体系分为 AB , $\text{A}_1\text{A}_2\text{B}$, AB_1B_2 三类, 分别归纳于图 2(a)–(c). 可以看出, 在阳离子相同的情况下, P_s 与体系阴离子半径呈负相关关系. 并且 P_s 与体系阳离子的价电子有关, 将数据集内体系按照阳离子价态的不同, 划分成单价、二价与三价这三类, 分别绘制不同价电子体系的箱线图, 如图 2(d) 所示. 可以看出, 随着体系阳离子价态的增加, P_s 呈

现上升趋势. 从整体看来, 129 种纤锌矿体系的 P_s 呈现近似正态的分布, 数据范围涵盖广泛, 集中在 $0.25\text{--}1.5 \text{ C/m}^2$ 区间内. 由此看来, 本文的 P_s 计算值可以作为研究数据集, 并为后续工作提供可靠的数据支持. 但很明显, 并不是只有阳离子价态与阴离子半径会对 P_s 存在影响, 这表明可能存在多种因素协同影响的复杂机制, 为此我们针对纤锌矿材料进行了全面的特征提取.

3.2 特征的提取与筛选

本文从组成元素的基本属性^[36]、纤锌矿材料的特殊晶体结构和电子性能这 3 个方面, 对可能影响 P_s 的纤锌矿材料特征进行了全面提取. 同时考虑了各类因素的最大值、最小值和平均值, 以确定

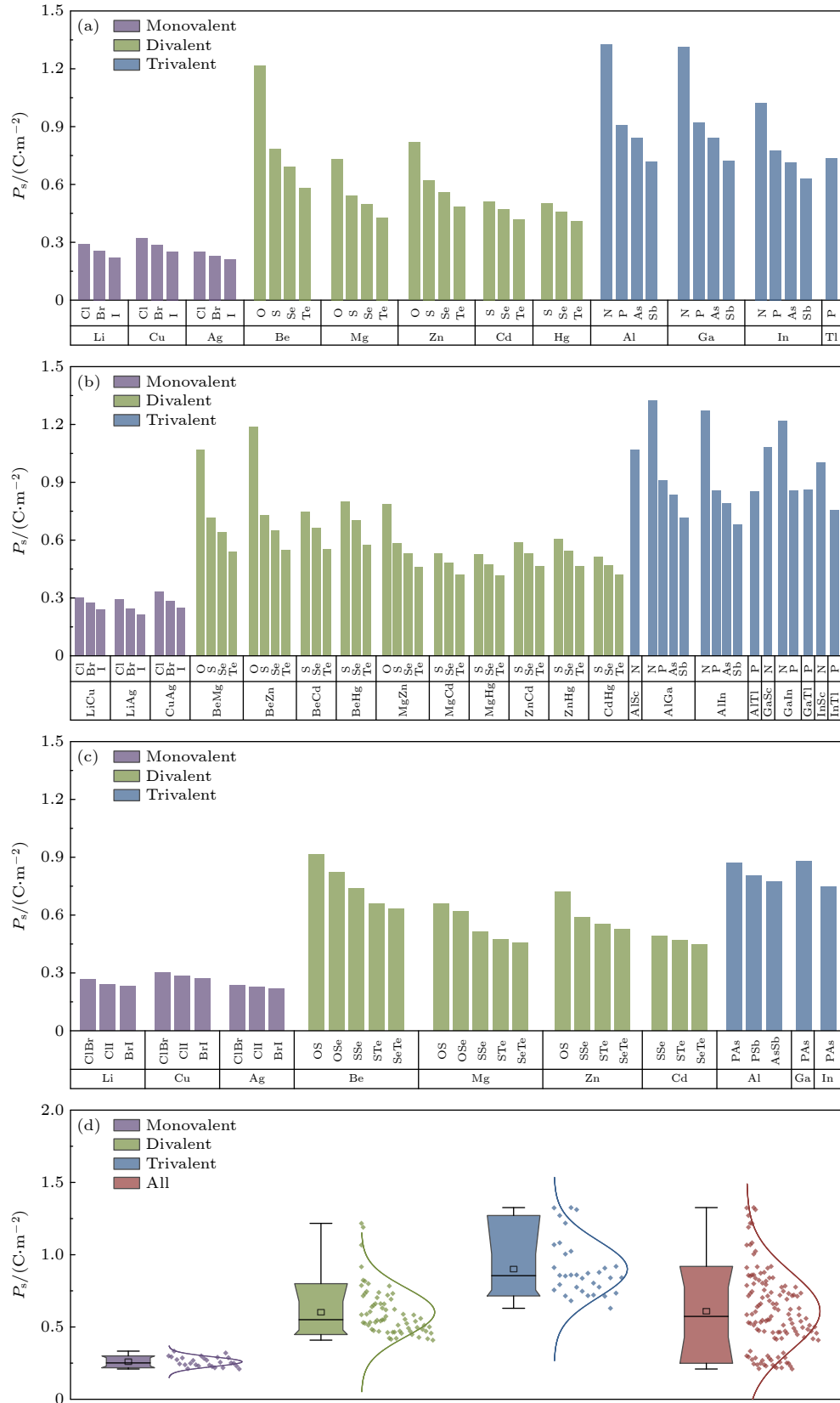


图 2 (a) AB, (b) A₁A₂B, (c) AB₁B₂ 体系的 P_s 数值; (d) 单价、二价、三价阳离子和整个数据集的箱线图
 Fig. 2. P_s of (a) AB, (b) A₁A₂B, and (c) AB₁B₂ systems; (d) box plots of P_s for monovalent, divalent, and trivalent cations as well as all systems.

更有物理意义的影响因素. 相较于传统钙钛矿铁电材料主要考虑元素单一属性和部分晶体结构的特征工程^[37], 本文创新提出了更适合纤锌矿材料的离子势、四面体内部曲折度等综合特征, 并对电子结构性质进行了定量分析, 有助于全面提高对纤锌矿材料的描述能力. 特征提取方法已在先前工作^[34]中详细介绍, 具体的特征介绍与统计处理方法归纳于补充材料, 总结如表1—表3所示, 其中结构示意图见图3和图4.

总之, 共提取得到了 505 种特征. 本文所构建的数据集, 不仅填补了基于 Berry 相方法的纤锌矿材料 P_s 计算值的空白, 更是为纤锌矿材料提供了详细全面的特征描述, 为进一步开发性能优异的纤

矿铁电材料奠定了理论基础. 但是当数据集具有大量特征时, 模型的训练和预测时间会大幅延长, 容易出现过拟合现象. 并且当特征之间存在强相关性时, 模型会受到多重共线性影响, 无法判断特征的重要性. 为了获得更好的 P_s 预测效果, 寻求 P_s 的关键影响因素, 本文进一步开展了特征的筛选工作.

首先本文借助 Boruta 算法, 寻求模型预测效果更好的特征组合, 得到了 105 个特征的复杂特征集. 但这些特征之间难免存在较强的自相关性, 这会影响对特征重要性的判断. 为此我们考虑了特征之间的距离相关性, 在距离自相关系数大于 0.7 的特征之中, 仅保留了物理意义明确, 特征提取方便, 且与目标性能 P_s 相关性较高的一个特征.

表 1 用于描述纤锌矿材料的特征类型——元素基本属性

Table 1. Parameters describing the characteristic types of wurtzite materials—Basic element attributes.

特征	定义	特征	定义
aN	原子序数	aM	相对原子质量
ICV	ICSD数据库内晶体体积	aD	原子密度
X	电负性	VE	价电子
R_c	共价半径	R_i	离子半径
R_a	原子半径	IPi	离子势(VE/Ri)
FIP	第一电离势	SIP	第二电离势
TIP	第三电离势	CP	化学势
FAE	自由原子能量	Men	门捷列夫数

表 2 用于描述纤锌矿材料的特征类型——晶体结构参数

Table 2. Parameters describing the characteristic types of wurtzite materials—Crystal structure parameters.

特征	定义	特征	定义
c, a	晶胞参数	c/a	晶胞参数比值
V	晶胞体积	S	晶胞底面积
H_{tetra}	四面体高度	V_{tetra}	四面体体积
d	阳离子到四面体底部的距离	μ	四面体内部曲折度(d/H_{tetra})
L_a, L_e	轴向键与平向键键长	$L_a - L_e$	轴向键与平向键键长的差值
L_a/L_e	轴向键与平向键键长的比值	L	四面体内平均键长
A_a, A_e	轴向键与平向键键角	$A_a - A_e$	轴向键与平向键键角的差值
A_a/A_e	轴向键与平向键键角的比值	A	四面体内平均键角

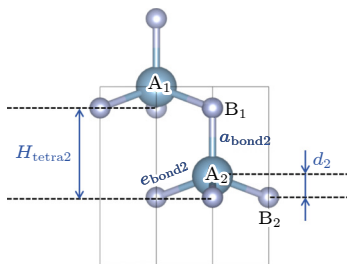


图 3 四面体结构参数示意图

Fig. 3. Schematic diagram of tetrahedral structure parameters.

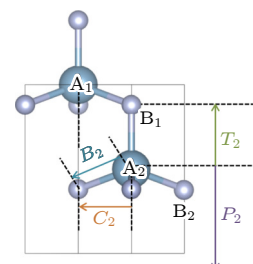


图 4 曲线路径示意图

Fig. 4. Schematic diagram of curve path.

表 3 用于描述纤锌矿材料的特征类型——电子结构参数
Table 3. Parameters describing the characteristic types of wurtzite materials—Electronic parameters.

分析	特征	定义	特征	定义
COHP分析	ICa, ICe	轴向键与平向键键强	ICa-ICe	轴向键与平向键键强的差值
	ICa/ICe	轴向键与平向键键强的比值	IC	四面体内平均键强
BC分析	AtV	原子体积	Bad	巴德电荷
DOS分析	Bc_tot	总能带中心	Bc_out	外层轨道的能带中心
	Bw_tot	总能带宽度	Bw_out	外层轨道的能带宽度
	Bs_tot	总能带偏度	Bs_out	外层轨道的能带偏度
	Bk_tot	总能带峰度	Bk_out	外层轨道的能带峰度
ELF分析	ELF_Cat	阳离子位置的ELF值	ELF_Ani	阴离子位置的ELF值
	ELF_dCA	阴阳离子位置ELF值的差值	ELF_Max	ELF曲线峰值
	ELF_MP	ELF曲线峰值的位置	ELF_F	ELF曲线半峰全宽
DC分析	DC_Cat	阳离子位置的DC值	DC_Ani	阴离子位置的DC值
	DC_dCA	阴阳离子位置DC值的差值	DC_Max	DC曲线峰值
	DC_MP	DC曲线峰值的位置	DC_F	DC曲线半峰全宽

我们把特征分为元素基本属性、晶体结构参数、COHP、Bader、DOS、ELF 和 DC 这 7 类. 因为组内特征存在各类统计分析, 且相似程度较高, 所以首先在各组内进行特征筛选. 这里需要注意的是, 因为元素基本属性组别的特征较多, 为了方便展示, 我们将该组分成了 3 组进行初步筛选, 其相关性热图绘制于补充材料图 S2 ([online](#)), 并对元素基本属性的初步保留特征进行了图 5(a) 所示的分析. 整理其余六类特征的组内相关性分析, 得到如图 5(b)—(e) 所示的相关系数热图, 图中红色框表明该特征被保留. 补充材料图 S3 ([online](#)) 中展示了各特征与 P_s 之间的 Pearson 相关系数和距离相关系数, Pearson 相关系数可以更清楚地表明特征的正负相关关系. 在进行组内特征筛选之后, 汇总所有保留特征, 进行组间的特征筛选, 得到如图 5(f) 所示的相关系数热图. 经过上述筛选, 最终生成一个包含 10 个特征的简化特征集, 这两个优化后的特征集在后续的分析中被用于机器学习模型拟合.

3.3 机器学习模型预测

本文分别针对特征筛选后得到的 105 个特征的复杂特征集, 以及 10 个特征的简化数据集, 进行 KRR, LASSO, SVR, ANNs, RF, XGBR 模型的机器学习训练, 分别得到如图 6(a), (b), (d), (e) 所示的训练集与测试集的预测效果. 可以看出, 简化后的特征集虽然在大部分情况下使模型的预测

精度降低, 但对于平均值为 0.609 C/m^2 的 P_s 数据集来说, 仍保持着良好的预测效果. 这表明上述因素的组合, 在很大程度上决定了体系的 P_s . 图 6(c), (f) 中分别展示了两种特征集的 XGBR 模型预测值与本文第一性原理计算值的对比散点图, 直观表明了 XGBR 模型优秀的预测能力. 各模型的具体超参数与其余模型的预测散点图分别归纳于补充材料表 S4 ([online](#)) 和表 S5 ([online](#)) 及图 S4 ([online](#)) 和图 S5 ([online](#)) 所示.

3.4 特征重要性分析

针对性能较好且物理意义清晰的含 10 个特征 XGBR 模型, 进行特征重要性分析. 本文借助 SHAP 方法, 得到特征重要性排序的 SHAP 散点图和 SHAP 绝对值均值, 分别如图 7(a), (b) 所示. 研究表明, 纤锌矿体系阳离子离子势的均值 IPi_Aave , 晶胞参数 a 对纤锌矿材料 P_s 存在较高的影响作用, 其中 IPi_Aave 更为突出. 这一结果表明, 相比于 Moriwake 等^[33] 仅考虑了价电子对 P_s 的单一影响, 本文通过价电子和离子半径比值的综合特征——离子势来评估对 P_s 的影响, 可以更好地解释不同体系 P_s 变化的原因. 剩余的影响因素中, 如体系阴离子离子势的均值 IPi_Bave , 阴离子第一电离势的均值 FIP_Bave , 阳离子门捷列夫数的均值 Men_Aave 等, 虽然对 P_s 也存在一定的影响, 但影响程度不如 IPi_Aave 和 a 强烈. 为此, 我们认为应尽量关注前两个更重要的影响因素.

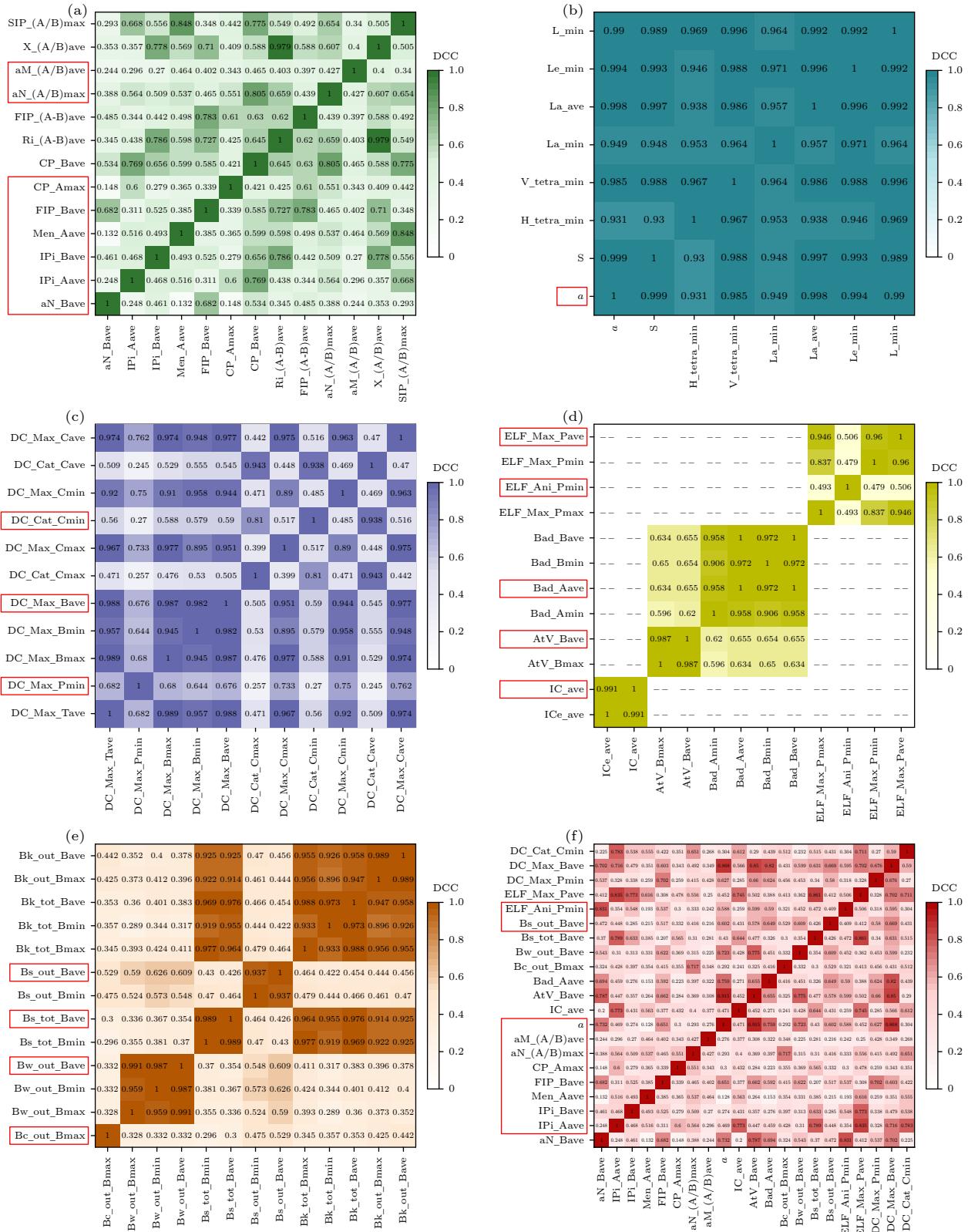


图5 组内距离自相关系数热图 (a) 元素基本属性特征; (b) 晶体结构特征; (c) DC 特征; (d) COHP, BC, ELF 特征; (e) DOS 特征. (f) 保留特征的组内距离自相关系数热图

Fig. 5. Heatmap of distance correlation coefficients: (a) element properties; (b) crystal parameters; (c) DC properties; (d) COHP, BC, ELF properties; (e) DOS properties. (f) Feature preserving inter group distance autocorrelation coefficient heat map.

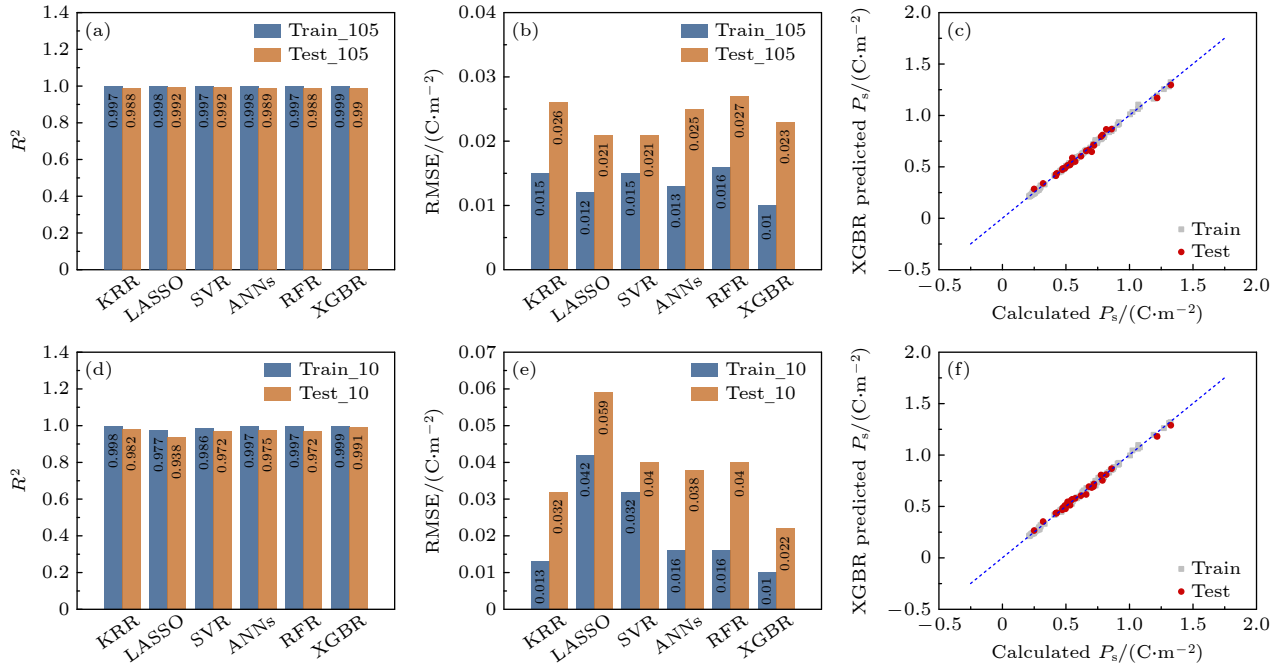


图6 6种含105特征机器学习模型的(a) R^2 , (b) RMSE, (c) XGBR模型的预测散点图; 6种含10特征机器学习模型的(d) R^2 , (e) RMSE, (f) XGBR模型的预测散点图

Fig. 6. (a) R^2 and (b) RMSE values of six machine-learning models and (c) scatter plot of the XGBR model with 105 features; (d) R^2 and (e) RMSE values of six machine-learning models and (f) scatter plot of the XGBR model with 10 features.

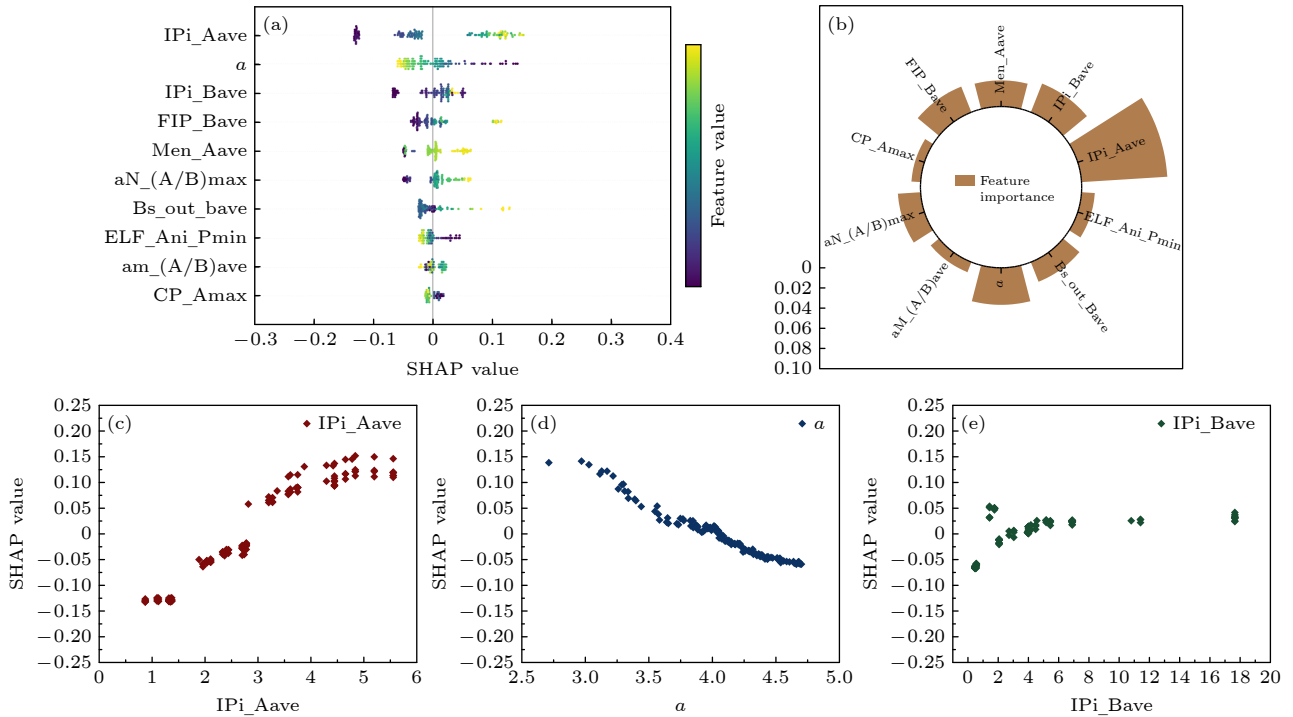


图7 含10个特征XGBR模型的(a) SHAP散点图和(b) SHAP特征重要性; (c) IPi_Aave, (d) a , (e) IPi_Bave特征的部分依赖图

Fig. 7. (a) SHAP summary plot, (b) SHAP feature importance of XGBR model with 10 features; partial dependence plots for the top three important features of (c) IPi_Aave, (d) a , (e) IPi_Bave.

此外, 为了更清楚地了解重要特征对 P_s 的具体影响, 我们对排名前三的特征, 绘制了影响 P_s 的部分依赖图, 分别如图 7(c)—(e) 所示. 可以明显看

出, 在最关键的影响因素 IPi_Aave 较高的情况下, IPi_Aave 对 P_s 存在显著的正向影响, 且 IPi_Aave 较低时存在显著负向影响. 表明较高的 IPi_Aave

代表阳离子对其周围电子云产生的吸引力越强,使得相邻电荷分布发生较大的位移,从而更容易发生极化现象. 而晶胞参数 a 则具有相反的影响趋势,并且在体系的 $a = 3.75 \text{ \AA}$ 左右时,对 P_s 的影响较为平缓. 而 IPi_Bave 整体对 P_s 的影响均处于 ± 0.05 之间,影响效果并不显著. IPi_Aave 和 a 这两个关键影响因素的提出,合理解释了 AlScN 体系随着 Sc 浓度的增加使体系 P_s 降低的现象^[18],因为 Sc 原子的掺入导致体系 IPi_Aave 降低,并且令 a 升高. 后续对纤锌矿材料铁电性能的优化,可优先考虑通过 IPi_Aave 和 a 这两个因素进行 P_s 性能的调控.

4 结 论

本文为 40 种二元和 89 种三元纤锌矿体系提供了精确的自发极化强度 P_s 的第一性原理计算值,构建了包含 505 个特征在内的可靠数据集. 并且采用机器学习和统计分析方法,不仅提出了预测 P_s 的机器学习模型,更深入研究了影响 P_s 的关键因素. 我们发现纤锌矿铁电材料的 P_s 在很大程度上由体系阳离子离子势的均值 IPi_Aave 和晶胞参数 a 等因素综合影响. 我们的研究提出了纤锌矿材料 P_s 的多因素综合影响机制,找到了更为本质的、易于获得的影响因素,有助于调控纤锌矿材料的 P_s , 加快初步筛选出铁电综合性能优异的候选体系.

数据可用性声明

本研究成果的数据集可在科学数据银行 <https://www.doi.org/10.57760/sciencedb.j00213.00073> 中访问获取.

参考文献

- [1] Ishiwara H 2012 *J. Nanosci. Nanotechnol.* **12** 7619
- [2] Arimoto Y, Ishiwara H 2004 *Mrs. Bull.* **29** 823
- [3] Song S, Kim K H, Chakravarthi S, Han Z, Kim G, Ma K Y, Shin H S, Olsson R H, Jariwala D 2023 *Appl. Phys. Lett.* **123** 183501
- [4] Muralt P 2000 *J. Micromech. Microeng.* **10** 136
- [5] Takayama R, Tomita Y, Iijima K, Ueda I 1991 *Ferroelectrics* **118** 325
- [6] Han X, Ji Y, Yang Y 2022 *Adv. Funct. Mater.* **32** 2109625
- [7] Zhong W L 1996 Physics of Ferroelectric Materials (Beijing: Science Press) p1 (in Chinese) [钟维烈 1996 (北京: 科学出版社) 第 1 页]

- [8] Li F, Zhang S J, Xu Z 2020 *Acta Phys. Sin.* **69** 217703 (in Chinese) [李飞, 张树君, 徐卓 2020 物理学报 **69** 217703]
- [9] Fan Y T, Tan G J 2022 *Mater. Lab* **1** 220008
- [10] Fichtner S, Wolff N, Lofink F, Kienle L, Wagner B 2019 *J. Appl. Phys.* **125** 114103
- [11] Zhao Z, Chen Y R, Wang J F, Chen Y W, Zou J R, Lin Y X, Xing Y F, Liu C W, Hu C M 2022 *IEEE Electr. Device L* **43** 553
- [12] Hayden J, Hossain M D, Xiong Y H, Ferri K, Zhu W L, Imperatore M V, Giebink N, Trolrier-McKinstry S, Dabo I, Maria J P 2021 *Phys. Rev. Mater.* **5** 044412
- [13] Wang D, Mondal S, Liu J N, Hu M T, Wang P, Yang S, Wang D H, Xiao Y X, Wu Y P, Ma T, Mi Z 2023 *Appl. Phys. Lett.* **123** 033504
- [14] Wang D, Wang P, Wang B Y, Mi Z 2021 *Appl. Phys. Lett.* **119** 111902
- [15] Ferri K, Bachu S, Zhu W L, Imperatore M, Hayden J, Alem N, Giebink N, Trolrier-McKinstry S, Maria J P 2021 *J. Appl. Phys.* **130** 044101
- [16] Zhang Y L, Zhu Q X, Tian B B, Duan C G 2024 *Nano-Micro Lett.* **16** 227
- [17] Yasuoka S, Shimizu T, Tateyama A, Uehara M, Yamada H, Akiyama M, Hiranaga Y, Yasuo C, Funakubo H 2020 *J. Appl. Phys.* **128** 114103
- [18] Ye K H, Han G, Yeu I W, Hwang C S, Choi J H 2021 *Phys. Status Solidi R* **15** 2100009
- [19] Wang P, Wang D, Vu N M, Chiang T, Heron J T, Mi Z 2021 *Appl. Phys. Lett.* **118** 223504
- [20] Kresse G, Furthmüller J 1996 *Phys. Rev. B* **54** 11169
- [21] Kresse G, Joubert D 1999 *Phys. Rev. B* **59** 1758
- [22] Perdew J P, Burke K, Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
- [23] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G, Persson K A 2013 *APL Mater.* **1** 011002
- [24] Dreyer C E, Janotti A, Van de Walle C G, Vanderbilt D 2016 *Phys. Rev. X* **6** 021038
- [25] Fichtner S, Yassine M, Van de Walle C G, Ambacher O 2024 *Appl. Phys. Lett.* **125** 040501
- [26] Liu Z J, Wang X Y, Ma X Y, Yang Y R, Wu D 2023 *Appl. Phys. Lett.* **122** 122901
- [27] Calderon S, Hayden J, Baksa S M, Tzou W, McKinstry S T, Dabo I, Maria J P, Dickey E C 2023 *Science* **380** 1034
- [28] Lee C W, Yazawa K, Zakutayev A, Brennecke G L, Gorai P 2024 *Sci. Adv.* **10** ead10848
- [29] Lee C W, Din N U, Yazawa K, Brennecke G L, Zakutayev A, Gorai P 2024 *Matter* **7** 1644
- [30] Spaldin N A 2012 *J. Solid State Chem.* **195** 2
- [31] Kursa M B, Rudnicki W R 2010 *J. Stat. Softw.* **36** 1
- [32] Kim Y, Kim Y 2022 *Sustain. Cities Soc.* **79** 103677
- [33] Moriwake H, Yokoi R, Taguchi A, Ogawa T, Fisher C A J, Kuwabara A, Sato Y, Shimizu T, Hamasaki Y, Takashima H, Itoh M 2020 *APL Mater.* **8** 121102
- [34] Kang Y, Chen J, Sui J Y, Wang X D, Zhou D Y, Yao M 2024 *ACS Appl. Mater. Interfaces* **16** 49484
- [35] Sun J A, Zhang J F, Yan J Y, Wang Y L, Lou J Z, Yan X B, Guo J X 2022 *Phys. Status. Solidi. B* **259** 2200079
- [36] Kirklin S, Saal J E, Meredig B, Thompson A, Doak J W, Aykol M, Rühl S, Wolverton C 2015 *NPJ Comput. Mater.* **1** 15010
- [37] Yang Z X, Gao Z R, Sun X F, Cai H L, Zhang F M, Wu X S 2019 *Acta Phys. Sin.* **68** 210502 (in Chinese) [杨自欣, 高章然, 孙晓帆, 蔡宏灵, 张凤鸣, 吴小山 2019 物理学报 **68** 210502]

DATA PAPERS • COVER ARTICLE

Key factors of spontaneous polarization
magnitude in wurtzite materials^{*}KANG Yao¹⁾ CHEN Jian¹⁾ TONG Yi²⁾ WANG Xinpeng²⁾ DUAN Kun¹⁾
WANG Jiaqi¹⁾ WANG Xudong¹⁾ ZHOU Dayu^{1)†} YAO Man^{1)‡}¹⁾ (School of Materials Science and Engineering, Dalian University of Technology, Dalian 116081, China)²⁾ (Suzhou Laboratory, Suzhou 215123, China)

(Received 30 October 2024; revised manuscript received 28 November 2024)

Abstract

Emerging wurtzite ferroelectric materials have aroused significant interest due to their high spontaneous polarization magnitude (P_s). However, there is a limited understanding of the key factors that influence P_s . Herein, a machine-learning regression model is developed to predict the P_s using a dataset consisting of 40 binary and 89 simple ternary wurtzite materials. Features are extracted based on elemental properties, crystal parameters and electronic properties. Feature selection is carried out using the Boruta algorithm and distance correlation analysis, resulting in a comprehensive machine learning model. Furthermore, SHapley Additive exPlanations analysis identifies the average cation-ion potential (IPi_Aave) and the lattice parameter (a) as significant determinants of P_s , with IPi_Aave having the most prominent effect. A lower IPi_Aave corresponds to a lower P_s in the material. Additionally, a exhibits an approximately negative correlation with P_s .

This multifactorial analysis fills the existing gap in understanding the determinants of P_s , and makes a foundational contribution to the evaluating emerging wurtzite materials and expediting the discovery of high-performance ferroelectric materials.

The dataset in this work can be accessed in the Scientific Data Bank <https://www.doi.org/10.57760/sciencedb.j00213.00073>.

Keywords: wurtzite ferroelectric materials, spontaneous polarization magnitude, machine learning, first principles methods

PACS: 77.55.-g, 63.20.dk, 82.20.Wt, 42.25.Ja

DOI: [10.7498/aps.74.20241520](https://doi.org/10.7498/aps.74.20241520)

CSTR: [32037.14.aps.74.20241520](https://cstr.cn/32037.14.aps.74.20241520)

* Project supported by the National Natural Science Foundation of China (Grant Nos. 51974056, 51474047, 52472120), the Fundamental Research Funds for the Central Universities of China (Grant No. DUT24LAB117), and the Suzhou Laboratory Project (Grant No. SK-1202-2024-012).

† Corresponding author. E-mail: zhoudayu@dlut.edu.cn

‡ Corresponding author. E-mail: yaoman@dlut.edu.cn

纤锌矿铁电材料自发极化强度的本征影响因素

康瑶 陈健 童祎 王新朋 段坤 王嘉琪 王旭东 周大雨 姚曼

Key factors of spontaneous polarization magnitude in wurtzite materials

KANG Yao CHEN Jian TONG Yi WANG Xinpeng DUAN Kun WANG Jiaqi WANG Xudong
ZHOU Dayu YAO Man

引用信息 Citation: *Acta Physica Sinica*, 74, 027701 (2025) DOI: 10.7498/aps.74.20241520

在线阅读 View online: <https://doi.org/10.7498/aps.74.20241520>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于机器学习和第一性原理计算的Janus材料预测

Prediction of magnetic Janus materials based on machine learning and first-principles calculations

物理学报. 2024, 73(23): 230201 <https://doi.org/10.7498/aps.73.20241278>

外加电磁场下周周期性体系的第一性原理计算方法

First-principles calculation method for periodic system under external electromagnetic field

物理学报. 2023, 72(23): 237102 <https://doi.org/10.7498/aps.72.20231313>

基于机器学习的无机磁性材料磁性基态分类与磁矩预测

Classification of magnetic ground states and prediction of magnetic moments of inorganic magnetic materials based on machine learning

物理学报. 2022, 71(6): 060202 <https://doi.org/10.7498/aps.71.20211625>

机器学习结合固溶强化模型预测高熵合金硬度

Machine learning combined with solid solution strengthening model for predicting hardness of high entropy alloys

物理学报. 2023, 72(18): 180701 <https://doi.org/10.7498/aps.72.20230646>

机器学习在光电子能谱中的应用及展望

Application and prospect of machine learning in photoelectron spectroscopy

物理学报. 2024, 73(21): 210701 <https://doi.org/10.7498/aps.73.20240957>

取向相关的 $\text{Pb}(\text{Zr}_{0.52}\text{Ti}_{0.48})\text{O}_3$ 外延薄膜的相图和介电性能

Phase diagram and dielectric properties of orientation-dependent $\text{PbZr}_{0.52}\text{Ti}_{0.48}\text{O}_3$ epitaxial films

物理学报. 2021, 70(12): 127701 <https://doi.org/10.7498/aps.70.20202164>