

专题: 原子分子和材料物性数据

Au-Cu 合金的形成能数据集构建与团簇展开建模*

贾煜茹¹⁾ 刘倍雷^{2)†} 高兴誉^{2)‡} 陈华杰^{1)††} 宋海峰²⁾¹⁾ (北京师范大学数学科学学院, 北京 100875)²⁾ (北京应用物理与计算数学研究所, 计算物理全国重点实验室, 北京 100088)

(2026 年 3 月 17 日收到; 2026 年 4 月 18 日收到修改稿)

针对 Au-Cu 体系的团簇展开 (CE) 建模, 采用基于双屏蔽库仑修正 (DSCC) 方法自洽评估 U , J 值的 PBE+ U^{DSCC} 修正 Cu-3d 轨道与 Au-5d 轨道能级位置, 构建了实空间 CE 建模所需的训练数据集. 进一步计算了 Au 与 Cu 单组元在不同晶向下的外延应变能, 构建了 Au-Cu 体系组分应变能训练数据集. 在此基础上, 建立了包含长程修正项的完整 CE 模型, 有效分离了长程弹性应变效应与短程化学相互作用, 相比单纯的实空间 CE 模型, 拟合精度与凸包线的预测精度均得到提升. 本文数据集可在 <https://www.doi.org/10.57760/sciencedb.j00213.00207> 中访问获取.

关键词: 团簇展开, 双屏蔽库仑修正, Au-Cu 合金, 形成能**DOI:** 10.7498/aps.75.20260377**CSTR:** 32037.14.aps.75.20260377

1 引言

团簇展开 (cluster expansion, CE) 的基础框架是将多组元体系的能量表示为原子构型空间一组完备团簇基函数的线性组合^[1-3], 组合系数称为有效团簇相互作用参数 (effective cluster interaction, ECI). CE 建模的基本方法是通过构型空间采样结合第一性原理密度泛函理论 (density functional theory, DFT) 计算构建训练数据集, 在此基础上由线性回归确定 ECIs^[4,5]. CE 可视为构型空间中的 DFT 代理模型, 与蒙特卡罗 (Monte Carlo, MC) 方法相结合, 应用于合金体系的热力学性质、相稳定性及相图的计算^[6-15].

Au-Cu 的相稳定性一直是 CE 建模关注的焦点^[16-22], 在室温附近实验可观测到 $\text{Cu}_3\text{Au}(\text{L1}_2)$, $\text{CuAu}(\text{L1}_0)$ 的有序稳定相^[23]. 尽管富 Au 端有序相

的结构尚未完全确定, 已有研究表明稳定相为 $\text{CuAu}_3(\text{L1}_2)$ ^[24-27], 多数理论研究将其视为参考的稳定相^[17,19,28]. 另一方面, Au-Cu 的 CE 建模结果表现出明显的泛函依赖性^[17,22,28,29]. Perdew-Burke-Ernzerhof (PBE) 泛函^[30] 难以准确描述 Au-5d 轨道与 Cu-3d 轨道的相互作用, 对稳定相的形成能预测值约为实验值的一半, 且不能预测 CuAu_3 稳定相^[28]. HSE06 杂化泛函计算的形成能与实验测量结果更贴合^[28], 但理论计算复杂度相比 PBE 泛函高一个量级, 在 CE 建模需要大规模构型空间采样时可能遇到性能瓶颈. 为平衡精度与效率, PBE+ U 方法通过引入 Hubbard U 修正项优化 d 电子关联效应的描述^[17], 在保持与 PBE 计算代价相当的同时改善了形成能预测精度, 但 U 值选取一般依赖人工经验^[17,31,32]. 近期, Liu 等^[33] 发展了双屏蔽库仑修正 (doubly screened Coulomb correction, DSCC) 方法, 其基于介电函数模型模拟有效屏蔽

* 云南省重点研发计划 (批准号: 202403AA080016) 资助的课题.

† 通信作者. E-mail: beilei_liu@126.com

‡ 通信作者. E-mail: gao_xingyu@iapcm.ac.cn

†† 通信作者. E-mail: chen.huajie@bnu.edu.cn

下的电子间相互作用, 无需人工调参即可高效评估局域电子间在位库仑相互作用 U 值与 Hund 耦合 J 值, 显著提升了 PBE+ U 这类方法的预测能力^[34]. 所以, 值得尝试使用 DSCC 方法来构建 Au-Cu 体系 CE 建模的训练数据集.

Au 与 Cu 之间约 12% 的原子尺寸差异^[19,35]使得体系中不可避免地存在具有非局域特征的弹性应变效应^[21,36], 这类长程应变效应难以通过有限阶、有限范围的实空间团簇展开来描述, 导致实空间 CE 模型在对 Au-Cu 体系的训练数据拟合精度、相稳定性预测以及对新构型的可迁移性方面出现系统性偏差^[18]. Laks 等^[37]提出了引入组分应变能的倒易空间 CE 方法, Ozoliņš 等^[19]在此基础上发展了混合空间 CE 方法 (mixed-space cluster expansion, MSCE), 后续被应用于原子尺寸失配的合金体系^[38]. Blum 和 Zunger^[39]发展了混合基簇展开 (mixed-basis cluster expansion, MBCE) 方法. 上述基于倒易空间扩展的 CE 模型构建依赖于描述组分应变能的数据集. Yuge 和 Okawa^[18]使用连续自旋基变晶格团簇展开 (continuous-spin basis variable-lattice CE, CS-VLCE) 来完整描述任意给定合金中的应变效应, 建模时需输入描述晶胞应变量的数据. Behara 等^[40]发展了显式依赖于构型与应变的 CE 表达形式, 该方法需补充由应变分量表征的应变状态数据. 此外, Zhuravlev 等^[36]建立了构型依赖的晶格变形模型 (configuration-dependent lattice deformation model, CLDM), 其训练数据包括构型的 Kanzaki 力. 因此, 在 Au-Cu 的 CE 建模中有必要针对长程弹性应变效应修正构建并给出训练数据集, 从而支撑完整的 CE 建模.

本文研究了面向 CE 建模的 Au-Cu 体系形成能数据集构建方法, 建立了实空间 CE 建模所需的训练数据集及修正项建模所需的应变能训练数据集. 针对实空间 CE 建模所需的训练数据集, 采用 PBE+ U^{DSCC} 方法自洽修正 Cu-3d 轨道与 Au-5d 轨道能级位置, 有效地提高了形成能的预测精度, 正确预测了实验观测到的 Cu₃Au(L1₂), CuAu(L1₀), CuAu₃(L1₂) 三个稳定相. 进一步采用 PBE+ U^{DSCC} 方法, 计算了 Au 与 Cu 单组元外延应变能, 构建了引入组分应变能的倒易空间 CE 修正模型所需的应变能训练数据集. 在此基础上, 建立了包含长程修正项的完整 CE 模型, 相比单纯的实空间 CE 模型, 富 Cu 端的预测偏差有明显降低, 能够正确描述实验已知的 3 种稳定相.

2 计算方法与参数设置

2.1 团簇展开模型

2.1.1 实空间团簇展开模型

团簇展开模型基于广义伊辛模型 (Ising model), 用于精确描述原子在晶格位点上的构型与材料总能量之间的映射关系^[1,2,41,42]. 对于二元合金 A_{1-x}B_x 体系, 定义每个格点 i 的自旋变量 $S_i = -1$ (若被 A 原子占据) 或 $+1$ (若被 B 原子占据). 任一原子构型 σ 的总能量可展开为一系列基于自旋乘积的基函数之和:

$$E(\sigma) = J_0 + \sum_i J_i S_i + \frac{1}{2} \sum_{i \neq j} J_{ij} S_i S_j + \frac{1}{3!} \sum_{i \neq j \neq k} J_{ijk} S_i S_j S_k + \cdots, \quad (1)$$

其中, $J_0, J_i, J_{ij}, J_{ijk}, \cdots$ 称为有效团簇相互作用参数. 若将晶格上的所有可能原子团簇记为 α , 则体系能量可进一步表达为更一般的形式:

$$E(\sigma) = \sum_{\alpha} J_{\alpha} \Pi_{\alpha}(\sigma), \quad (2)$$

其中 $\Pi_{\alpha}(\sigma)$ 为与构型 σ 相关的正交基函数, J_{α} 为团簇 α 的 ECI. 利用晶格对称性, 可将对称等价的团簇归为同一轨道 (orbit) β , 从而简化表达式为

$$E(\sigma) = \sum_{\beta} m_{\beta} J_{\beta} \langle \Pi_{\alpha}(\sigma) \rangle_{\beta}, \quad (3)$$

式中 m_{β} 为轨道 β 的多重度, $\langle \cdot \rangle_{\beta}$ 表示对该轨道内所有团簇的基函数取平均.

团簇展开模型中, ECIs 的强度随团簇阶数 (包含的格点数目) 及团簇内格点间距的增大而迅速衰减, 这一特性使得在实际应用中, 仅需对团簇展开进行有限项截断即可达到所需精度^[43]. 因此, 在构建模型前, 需通过系统的收敛性测试确定团簇截断方案, 选取既能保证精度又兼顾计算效率的截断参数, 避免因过度截断导致预测偏差, 或因保留过多团簇而引入冗余参数.

本文团簇展开模型的构建依赖于一组构型及其形成能数据. 形成能的定义为

$$\Delta H(\sigma) = E(\sigma) - (1-x)E_A - xE_B, \quad (4)$$

其中 $E(\sigma)$ 为构型 σ 的总能量; E_A 和 E_B 分别为纯组元 A, B 的能量; x 为组元 B 的摩尔分数. 通过正则化方法确定有限个关键团簇的 ECIs 后, 体系的

形成能可表示为

$$\Delta H_{\text{CE}}(\sigma) = J_0 + \sum_f^{N_f} D_f J_f \Pi_f(\sigma), \quad (5)$$

其中 J_0 为单点项能量, N_f 为所保留的非零 ECIs 数目, D_f 为每格点中 f 类团簇的个数, J_f 为对应团簇的 ECI, $\Pi_f(\sigma)$ 为构型 σ 中对应自旋乘积的格点平均. 基于该参数化模型, 可通过预测任意构型的形成能构建合金体系的基态凸包线, 判断低温下合金随组分变化的相稳定性. 进一步, CE 模型可作为 DFT 的代理模型, 结合 MC 系综采样, 给出构型熵对自由能的贡献, 支持高温下合金相图的计算^[4].

2.1.2 基于组分应变能修正的团簇展开模型

为描述长周期效应, 采用基于组分应变能项修正的 CE 模型^[37,44], 其体系构型 σ 的总能量表示为

$$E(\sigma) = E_{\text{CE}}(\sigma) + E_{\text{CS}}(\sigma) \\ = E_{\text{CE}}(\sigma) + \sum_{\mathbf{k}} \Delta E_{\text{CS}}(\hat{\mathbf{k}}, c) F(\mathbf{k}, \sigma), \quad (6)$$

其中, $E_{\text{CE}}(\sigma)$ 为传统实空间 CE 所描述的短程化学能量项, 而 $E_{\text{CS}}(\sigma)$ 则表示由组分不均匀性导致的长程弹性应变能. 组分应变能项被表示为对倒空间中浓度波分量的求和:

$$E_{\text{CS}}(\sigma) = \sum_{\mathbf{k}} \Delta E_{\text{CS}}(\hat{\mathbf{k}}, c) F(\mathbf{k}, \sigma), \quad (7)$$

其中, $\Delta E_{\text{CS}}(\hat{\mathbf{k}}, c)$ 是描述沿方向 $\hat{\mathbf{k}}$ 、总体浓度为 c 的超晶格形成所对应的单位原子外延应变能, 而 $F(\mathbf{k}, \sigma)$ 则衡量构型 σ 与波矢为 $\mathbf{k}$ 的浓度波之间的匹配程度. 具体而言, $\Delta E_{\text{CS}}(\hat{\mathbf{k}}, c)$ 可通过如下方式计算:

$$\Delta E_{\text{CS}}(\hat{\mathbf{k}}, c) = \min_{a_{\text{SL}}} [(1-c) \Delta E_{\text{epi}}^{\text{A}}(a_{\text{SL}}, \hat{\mathbf{k}}) \\ + c \Delta E_{\text{epi}}^{\text{B}}(a_{\text{SL}}, \hat{\mathbf{k}})], \quad (8)$$

其中, a_{SL} 为界面平面内的晶格参数, $\Delta E_{\text{epi}}^{\text{X}}(a_{\text{SL}}, \hat{\mathbf{k}})$ 表示相 X 在沿 $\hat{\mathbf{k}}$ 方向发生外延应变时相对于其平衡状态的能量差异. 该最小化过程反映了相干界面处两相通过共同晶格参数实现弹性匹配的物理过程. 构型匹配因子 $F(\mathbf{k}, \sigma)$ 通常写为

$$F(\mathbf{k}, \sigma) = \frac{|S(\mathbf{k}, \sigma)|^2}{4c(1-c)} \exp(-|\eta \mathbf{k}|^2), \quad (9)$$

其中 $S(\mathbf{k}, \sigma)$ 为构型的结构因子, 表征其在倒空间中对应变波矢 $\mathbf{k}$ 的浓度调制强度. 指数衰减项用于抑制高频波矢 (短波长) 浓度起伏对应的应变能贡献,

从而确保组分应变项主要作用于长周期构型.

在实际 CE 建模中, 为使实空间 CE 模型专注于描述短程化学相互作用, 组分应变能项在拟合 ECIs 之前需从第一性原理计算的训练数据中预先扣除.

2.2 构型空间采样与训练集生成方法

构型的生成采用结构枚举方法^[45], 即在给定母晶格与超胞原子数的条件下, 枚举所有不等价的原子占位方式. 该方法通过以下步骤实现构型空间系统采样: 1) 基于 Hermite 标准型 (HNF) 枚举超晶格指数 n (超胞中原子个数) 对应的所有整数变换矩阵, 结合母晶格对称性筛选出几何不等价的超晶格; 2) 将 Au, Cu 原子视为二元标记, 生成各超晶格下所有可能的原子占位构型 (总数为 2^n); 3) 利用 Smith 标准型 (SNF) 构建商群, 通过群论方法系统剔除平移等价、标记交换等价、超周期性及旋转等价的冗余构型. 该方法无需依赖几何比对, 仅通过完美哈希表与群论工具, 即可实现线性时间复杂度的等价性判断, 确保了采样的完整性、高效性与无冗余性.

本文以 Au 的面心立方 (FCC) 晶格作为母相, 限制原胞中原子数 $n = 1-7$ 个, 在 Au-Cu 合金完整的组分空间对构型采样, 共生成 241 个不等价构型. 由于构型枚举采样基于理想晶格, 而实际材料中原子往往会因弛豫而偏离理想格点位置, 且弛豫后的结构通常具有更低的能量和更高的热力学代表性, 因此在构建训练数据集时, 先通过 PBE+ U^{DSCC} 方法开展高通量第一性原理计算, 获得所有采样构型的弛豫结构及对应能量, 再将弛豫结构映射至最近似的理想晶格占位构型, 从而确保训练数据既保留理想晶格的构型特征, 又能真实反映实际材料的能量状态.

2.3 训练集数据计算方法

所有构型的结构弛豫与能量计算均基于 DFT, 通过 VASP (Vienna *ab-initio* simulation package) 软件包^[46-48] 完成. 计算采用投影缀加波 (PAW) 方法^[49], 平面波截断能设置为 600 eV, 布里渊区 k 点网格通过 KSPACING = 0.1 Å⁻¹ 自动生成. 电子交换关联能的处理采用广义梯度近似 (GGA), 选用 PBE 泛函^[30], 并利用 DSCC 方法^[33] 修正强关联效应, Au 和 Cu 元素对应的 U 值分别为 2.62 eV

和 6.24 eV, J 值分别为 0.75 eV 和 1.12 eV. 结构弛豫过程采用共轭梯度算法, 允许原子坐标、晶胞形状及体积同时优化 (构建单组元外延应变能数据集时仅弛豫原子坐标), 弛豫收敛判据为原子最大受力小于 0.01 eV/Å, 最大弛豫步数设置为 100 步.

3 结果与讨论

3.1 考虑 d 电子修正的训练数据集

为了获得 Au-Cu 合金更准确的形成能, 采用自主发展的 DSCC 方法对 Cu-3d 轨道和 Au-5d 轨道进行在位库仑修正, 自洽确定 Hubbard U 与 Hund 耦合 J 值. 对 Cu-3d 轨道评估获得的 U, J 值分别为 6.24, 1.12 eV, 对 Au-5d 轨道评估获得的 U, J 值分别为 2.62, 0.75 eV. 同时, 评估了不同组分下 Cu-3d 轨道和 Au-5d 轨道对应的 U 与 J 参数, 其值几乎不变, 如表 1 所列. 进一步分析 Au, Cu 的 d 分波电子数发现, 随着 Au 含量的增大, d 分波电子数仅有微弱的增大, 整体变化幅度很小 (见图 1), 这说明 Cu, Au 的 d 电子占据对组分变化并不敏感, 意味着没有发生显著的价态变化. 因此, 对所有的 Au-Cu 合金均采用自洽评估得到的固定 U, J 值来计算训练数据.

表 1 Au-Cu 合金中不同化合物的 Hubbard U 与 Hund 耦合 J 数值

Table 1. Hubbard U and Hund's coupling J parameters used for different compounds in Au-Cu alloys.

Compounds	Cu		Au	
	U/eV	J/eV	U/eV	J/eV
Cu ₃ Au	6.25	1.12	2.64	0.75
CuAu	6.24	1.12	2.62	0.75
CuAu ₃	6.23	1.12	2.62	0.75

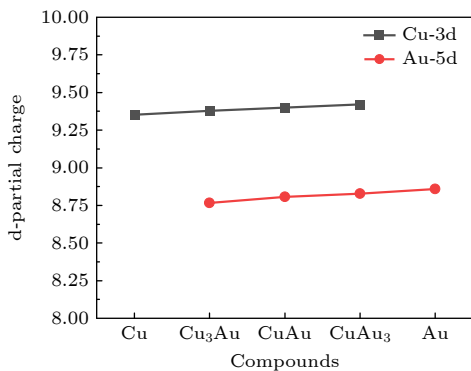


图 1 Cu 与 Au 的 d 分波电子数变化

Fig. 1. d-partial charge of Cu and Au.

应用 DSCC 修正后, 可以预测与实验一致的 Au-Cu 金属间化合物的形成能. 计算给出了基态凸包曲线 (如图 2), 这些曲线由热力学稳定结构的形成能构建, 位于包络线上的金属间化合物是稳定的. 可以发现, 基于 PBE 计算的 Au-Cu 形成能明显高于实验值, 并且实验合成的 CuAu₃ 不在凸包上. 然而, PBE+ U^{DSCC} 计算不仅能给出正确的形成能, 还能复现实验中报告的基态化合物 CuAu₃. 由于实验测量获得的 CuAu₃ 形成能来自一个不完全有序的结构 [19], 导致其在实验凸包曲线的包络线上方. 此外, 除实验已知的 3 个稳定相, PBE+ U^{DSCC} 计算还得到 2 个稳定相 CuAu₂(β), Cu₂Au₃(LPS). 这可能是由于实验中稳定相是自由能最小 (包括熵), 而静态能量计算没有考虑温度效应, 从而会显著影响相的热力学稳定性; 考虑温度后, 这些富 Au 相会变得不稳定或迁移到凸包以上, 与实验结果一致 [17].

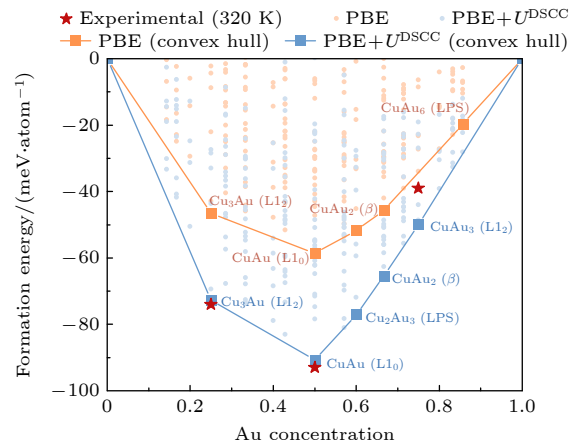


图 2 选用 PBE, PBE+ U^{DSCC} 的形成能凸包图 (★表示实验值 [23]), 其中 LPS 表示长周期结构

Fig. 2. Convex hulls of formation energies calculated using the PBE and PBE+ U^{DSCC} . The star symbols (★) denote experimental values [23]. LPS denotes the long-period structure.

已有文献 [17] 讨论了 d 带位置对 Au-Cu 等 Cu 合金形成能计算的重要影响. 在采用 PBE 泛函的 DFT 计算中, Cu-3d 和 Au-5d 间的 d 带中心距离被高估, 导致 3d-5d 杂化较弱, 形成能高于实验值. 然而, 需要说明的是, 在过去对 Au-Cu 体系进行 DFT+ U 计算中, U 值选取依赖人为经验, 而本文中采用完全自洽计算获得的 U, J 值.

如图 3 所示, 对 Cu-3d 轨道和 Au-5d 轨道开展修正后, Cu-3d 带的位置向下移动约 0.9 eV, 而 Au-5d 的位置向下移动约 0.3 eV, Cu-3d 带与 Au-

5d 带能级接近. 还可以观察到对 CuAu 合金开展修正后计算获得的 Au—Cu 键长相较未进行修正的 DFT 计算缩短 0.004 Å, 这暗示了 Cu 和 Au 间更强的键合作用. 进一步进行了晶体轨道哈密顿布

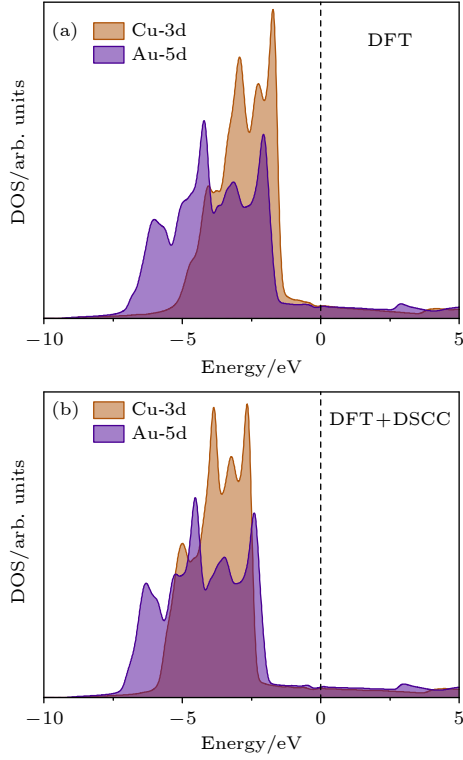


图 3 采用 DFT (a) 和 DFT+DSCC (b) 计算的 Cu 单质的 3d 轨道, Au 单质的 5d 轨道态密度, 图中竖直虚线表示费米能级

Fig. 3. Density of states of the Cu-3d orbital and Au-5d orbital for elemental Cu and Au calculated using DFT (a) and DFT+DSCC (b). The vertical dashed line indicates the Fermi level.

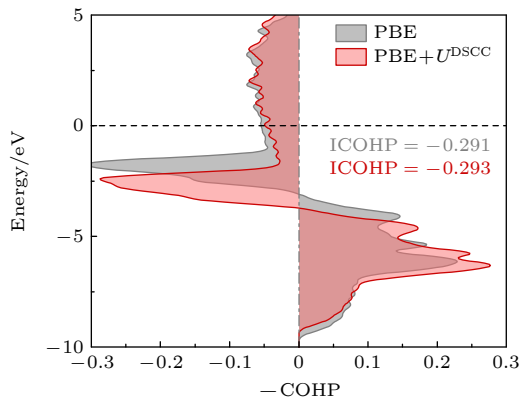


图 4 基于 PBE 与 $PBE+U^{DSCC}$ 计算的 CuAu 键合相互作用晶体轨道哈密顿布居 (COHP) 分析

Fig. 4. Crystal orbital Hamilton population (COHP) analysis of the Cu-Au bonding interaction in CuAu calculated using PBE and $PBE+U^{DSCC}$.

居 (COHP) 分析, 以研究键强度. CuAu 中 Cu 和 Au 之间的相互作用在图 4 中展示. -COHP 为正值时, 表示存在成键作用, 而为负值时, 表明存在反键作用. 可以清晰地看出, 相较于 PBE 计算, DSCC 修正下在 Fermi 面附近的反键峰值更弱, 而在 -7 eV 附近的成键峰值更强. 这表明, DSCC 修正强化了成键态, 同时抑制了反键态, 导致 Au-Cu 间具有更强的键合作用, 这也是导致 Au-Cu 具有更低形成能的微观机制.

3.2 组分应变能训练数据集

为实现基于组分应变能项修正的 CE 建模, 从单组元外延应变能出发, 构建了 Au-Cu 体系的组分应变能训练数据集. 该应变能数据集可直接用于基于组分应变能的倒易空间 CE 构建, 同时也适用于 MSCE, MBCE 模型. 首先选取 6 个晶向 [001], [011], [111], [201], [221], [311], 通过第一性原理计算获得每个晶向下单组元的能量-应变关系, 这些晶向对组分应变能的贡献占主导地位, 由它们拟合的结果仅对组分应变能的整体计算结果产生边际效应 [44,50].

对纯 Cu 和纯 Au 的弛豫计算显示, 二者的平衡晶格常数分别为 $a_{Cu} = 3.618 \text{ Å}$, $a_{Au} = 4.151 \text{ Å}$. 分别在 6 个晶向下, 沿晶向方向施加法向应变, 而在垂直于晶向的两个正交方向施加等效的面内应变, 从而再现实验中外延薄膜受基底约束、法向自由弛豫的物理情形. 对于 Au 元素, 其面内晶格参数从 $a_{Cu}/a_{Au} \approx 0.87$ 连续扫描至 1.12, 对于 Cu 元素则从 0.86 扫描至 $a_{Au}/a_{Cu} \approx 1.15$, 从而完整覆盖了由外延条件可能引入的主要拉伸与压缩应变状态. 在每个晶向和元素中, 面内应变区间被均匀离散为 11 个取样点, 在保证应变能曲面分辨率的同时控制第一性原理计算的总体规模. 该取样密度经测试能够稳定重构应变能随晶格参数变化的平滑趋势. 在固定面内应变取值 a 的条件下, 法向缩放因子在 $[a^{-0.8} - 0.12, a^{-0.8} + 0.12]$ 区间内均匀采样, 并离散为 11 个点. 随后, 对每一组固定面内应变条件下得到的法向应变能与晶格参数关系采用三次多项式进行拟合, 并通过数值最小化精确确定对应的最低应变能及其平衡法向晶格参数, 该处理等价于在外延约束条件下对法向自由度进行充分弛豫. 进一步, 对不同晶向下的离散曲线采用五次多项式进行整体拟合, 并将能量零点统一归一化至最低能量值, 获得连续、光滑且物理一致的单组元

外延应变能函数. 在此基础上, 合金应变能表示为 Au 与 Cu 单组元外延应变能的线性组合, 并通过对合金晶格参数进行能量最小化确定平衡组分应变能. 在最小化过程中, 初始晶格参数采用 Vegard 定律给出, 组分空间内选取 101 个等间隔浓度点, 对每一晶向分别求解合金外延应变能. 图 5 展示了基于 PBE+ U^{DSCC} 得到的 Au-Cu 体系平衡组分应变能随成分变化的结果.

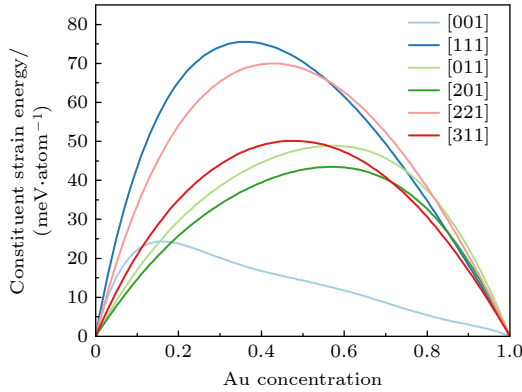


图 5 6 个低指数晶向下体系平衡组分应变能随 Au 浓度的变化

Fig. 5. Equilibrium constituent strain energies as a function of Au concentration along six low-index crystallographic directions.

为实现对未显式计算组分以及非选定晶向的高效插值, 进一步采用 Redlich-Kister 多项式对合金外延应变能的组分依赖关系进行参数化^[44,50]. 对每一个晶向, 应变能被表示为 Au 浓度 x 的函数:

$$E_{\text{CS}}(\hat{k}, x) = \sum_n L_n(\hat{k}) x(1-x)(1-2x)^n,$$

其中 $L_n(\hat{k})$ 为待拟合的 RK 系数. 结果表明, 对于所有研究晶向, 五阶 Redlich-Kister 多项式即可在整个组分区间内准确拟合合金外延应变能曲线, 未出现非物理振荡或异常极值点. 最终得到的 RK 系数见表 2.

表 2 6 个晶向的 RK 系数

Table 2. RK coefficients along six crystallographic directions.

$\hat{k}$	L_0	L_1	L_2	L_3	L_4
[001]	0.0573	0.0469	0.0504	0.1345	0.1221
[111]	0.2815	0.1394	0.0750	0.0547	0.0309
[011]	0.1922	-0.0441	0.0547	0.0088	-0.0184
[201]	0.1709	-0.0406	0.0395	0.0140	-0.0205
[221]	0.2745	0.0740	0.0331	0.0237	0.0133
[311]	0.2002	0.0168	0.0119	0.0171	0.0045

3.3 团簇展开模型训练结果

3.3.1 实空间团簇展开模型

本文的团簇展开模型构建均基于 ICET 软件完成^[51]. 在最优模型参数下, 采用十倍交叉验证方法得到团簇展开模型, 其验证集上的均方根误差 (root mean square error, RMSE) 为 4.85 meV/atom, 对应的决定系数 R^2 为 0.958. 从图 6 可以清楚地看到, 大多数 CE 预测值接近 DFT 计算值, 表明模型能够基本再现第一性原理计算所得的构型形成能.

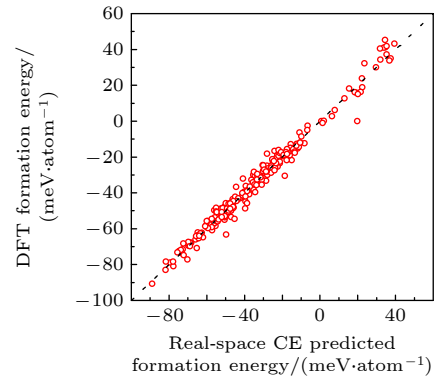


图 6 构型形成能的第一性原理计算结果与实空间 CE 模型预测值的对比

Fig. 6. Comparison between first-principles calculated formation energies and real-space CE model predictions.

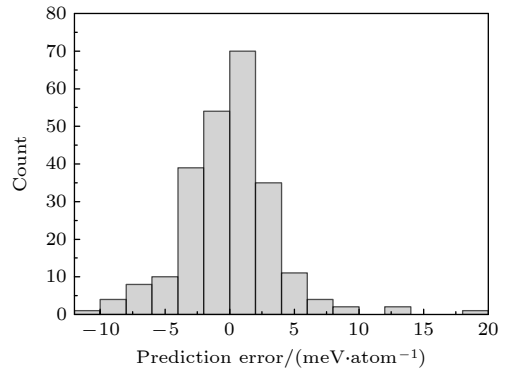


图 7 实空间 CE 模型的预测误差分布

Fig. 7. Prediction error distribution of the real-space CE model.

此外, 进一步考察了模型预测误差的分布 (图 7), 最大绝对误差 (maximum absolute error, MAE) 为 19.80 meV/atom. 结合图 8 中 DFT 计算与 CE 预测的形成能凸包对比结果可看出, 实空间 CE 能预测 Cu_3Au , CuAu 这 2 种实验已知稳定相, 不能预测 CuAu_3 稳定相, 且富 Cu 端存在明显预测偏

差. 增加富 Cu 端的训练数据后, 预测偏差仍未得到有效改善, 表明该问题并非源于训练数据规模不足.

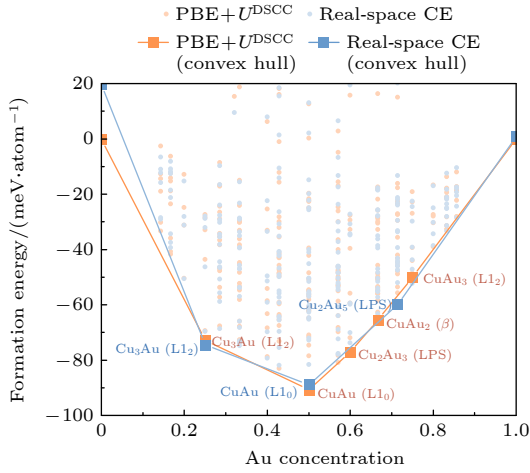


图 8 基于第一性原理计算 (橙) 与实空间 CE 模型预测结果 (蓝) 构建的 Au-Cu 合金形成能凸包对比, 其中 LPS 表示长周期结构

Fig. 8. Comparison of the formation energy convex hulls of the Au-Cu system constructed from first-principles calculations (orange) and real-space CE model predictions (blue). LPS denotes the long-period structure.

3.3.2 基于组分应变能修正的团簇展开模型

选择自动相关确定回归 (automatic relevance determination regression, ARDR) 作为正则化拟合方法^[43], 测试选取截断距离 (cutoff distance) 和团簇阶数 (cluster order). 以步长 0.5 Å 展开截断扫描, 遍历范围为两原子团簇 8—20 Å, 三原子团簇 5—12 Å, 四原子团簇 4—9 Å, 五原子团簇 3—5 Å, 过程中剔除基函数矩阵条件数大于 1×10^{10} 的截断距离. 采用十倍交叉验证法, 以验证集 RMSE 作为拟合精度的评判标准, 将使验证集 RMSE 最小的参数组合定义为最优截断距离. 测试结果如图 9 所示, 可知最优截断距离为 13.5, 6.0, 4.5 Å.

同时, 图 10 中对团簇阶数测试的 ECI 结果表明, ECI 贡献随阶数增加而衰减, 模型展开到五阶团簇时已经收敛. 值得注意的是, 高阶团簇的存

在受截断距离的严格限制, 五原子及以上团簇可能因原子间距约束而无法形成, 但即使在高阶团簇存在的前提下, 其引入也并未带来模型精度的提升. 因此最终 CE 模型仅考虑展开到四原子团簇.

表 3 展示了引入组分应变修正前后模型的拟合效果. 引入应变修正后, 验证集 RMSE 降低, 同时 R^2 略微升高, 表明模型对第一性原理能量的整体拟合精度有所提升. 为进一步评估基于组分应变能修正的 CE 模型的物理合理性, 本文同时对未修

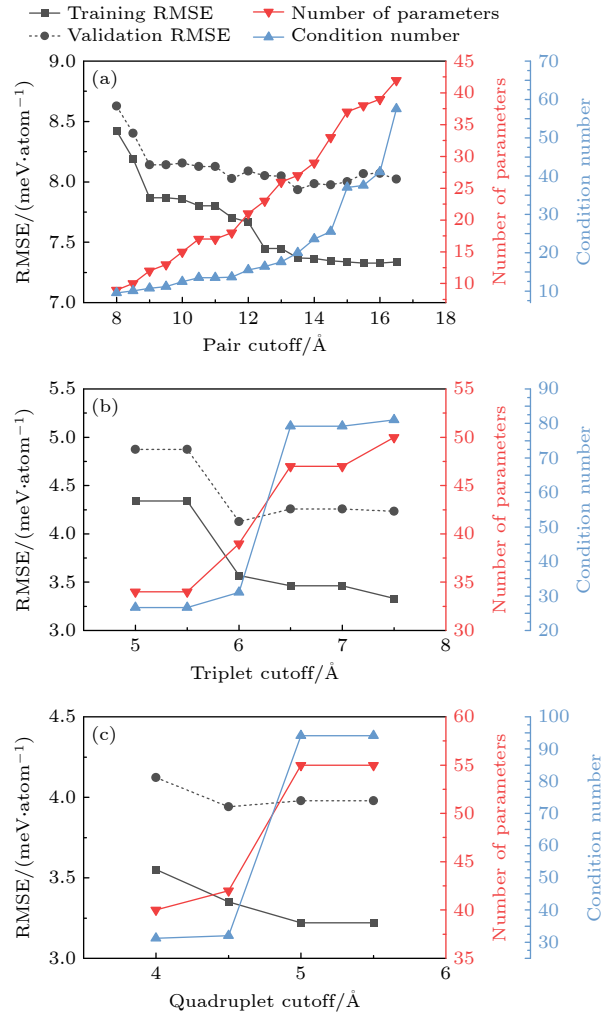


图 9 截断距离测试结果 (a) 二阶团簇; (b) 三阶团簇; (c) 四阶团簇

Fig. 9. Cutoff distance test results: (a) Second-order clusters; (b) third-order clusters; (c) fourth-order clusters.

表 3 组分应变能修正前后模型拟合精度比较

Table 3. Comparison of model fitting accuracy for the Au-Cu system before and after constituent strain energy correction.

CE type	Number of nonzero ECIs	Optimal cutoff distance/Å			Training-set RMSE/(meV·atom ⁻¹)	Validation-set RMSE/(meV·atom ⁻¹)	R^2 (validation)
		Pair	Triplet	Quadruplet			
Real-space CE	40	15.0	7.5	5.0	3.65	4.92	0.958
Strain-corrected CE	30	13.5	6.0	4.5	3.37	3.97	0.968

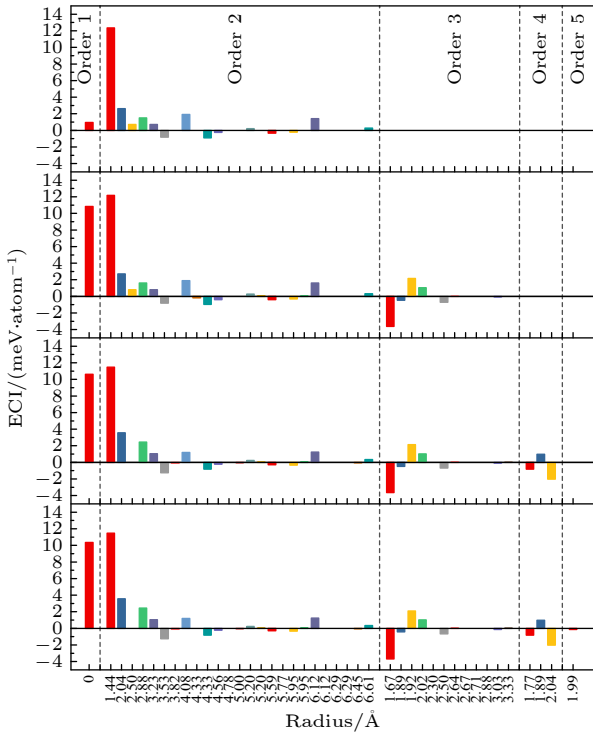


图 10 团簇阶数测试结果, 其中相同颜色代表不同阶数下的同一近邻

Fig. 10. Cluster order test results. Curves with the same color represent the same neighbor shell at different cluster orders.

正的实空间非零 ECIs 和应变修正后的实空间非零 ECIs 进行比较, 见图 11. 可以观察到, 在两种模型中, ECIs 的主导贡献均来源于低阶、短程团簇, 这表明 Au-Cu 体系的主要化学相互作用具有明显的局域性特征. 在实空间 CE 模型中, 部分远程团簇的 ECIs 会出现非物理放大的趋势; 而在扣除组分应变能并重新拟合后, 这类 ECIs 明显减少, 使得 ECIs 随团簇半径呈现出更为清晰的衰减行为. 这表明引入组分应变修正项, 能够从实空间剥离部分长程作用, 使 CE 能更合理地描述原子尺寸失配引起的长程弹性相互作用. 表 4 汇总了本文应变修正 CE 模型的各项关键参数.

表 4 应变修正 CE 模型的关键参数
Table 4. Key parameters for the strain-corrected CE.

Parameters	Value		
Primitive structure	Au (FCC)		
Cluster order	2	3	4
Cutoff distance/Å	13.5	6.0	4.5
Complete ECIs	Detailed in Supplementary Material 3, Science Data Bank		
Strain energy parameters	RK coefficients, see Table 2		

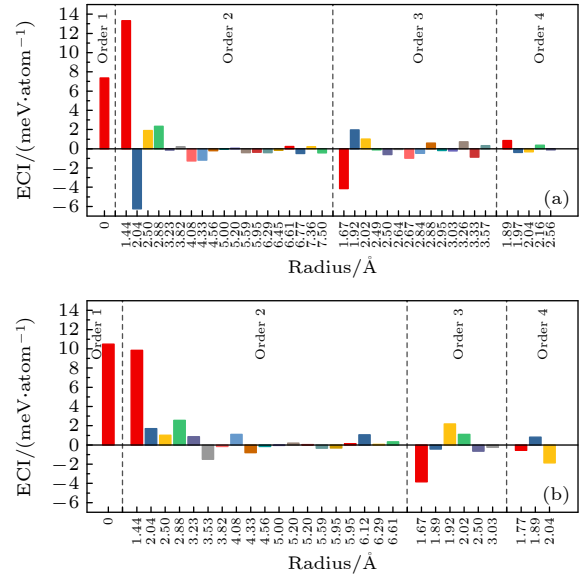


图 11 组分应变能修正前后实空间非零 ECIs 比较 (a) 未修正的实空间非零 ECIs; (b) 应变修正后的实空间非零 ECIs
Fig. 11. Comparison of nonzero real-space ECIs before and after constituent strain energy correction: (a) Uncorrected real-space ECIs; (b) strain-corrected real-space ECIs.

最后, 基于引入组分应变能修正的 CE 模型, 对 Au-Cu 体系稳定化合物进行预测, 其形成能凸包如图 12 所示. 可以看出, 相较于未修正模型, 富 Cu 端的预测偏差明显降低 (由 19.80 meV/atom 减小为 9.98 meV/atom). 同时, 该模型能够正确预测实验已知的 Cu_3Au , CuAu 及 CuAu_3 三种稳定相, 这表明 CE 模型在稳定相判定与能量预测方面的可靠性得

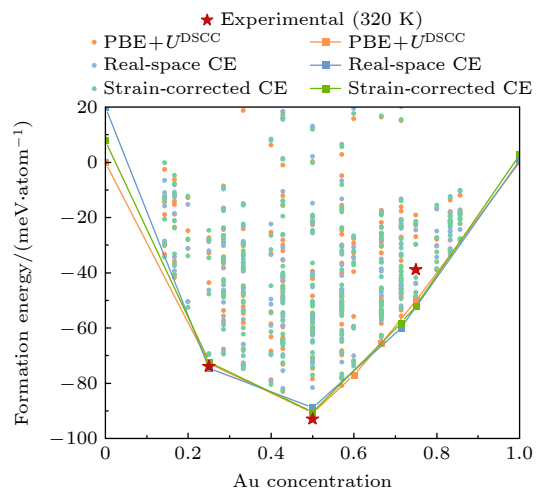


图 12 PBE+ U^{DSCC} 与 CE 预测凸包对比: PBE+ U^{DSCC} (橙)、实空间 CE 预测 (蓝)、基于组分应变能修正的 CE 预测 (绿)
Fig. 12. Comparison of the formation energy convex hulls of the Au-Cu system predicted by PBE+ U^{DSCC} (blue), the real-space CE model (orange), and the strain-corrected CE model (green).

到提升. 值得注意的是, $\text{Cu}_2\text{Au}_5(\text{LPS})$ 相形成能的 DFT 计算结果为 -47.87 meV/atom , 组分应变能修正的 CE 模型计算结果为 -49.09 meV/atom . 若将 CE 拟合的 RMSE(3.97 meV/atom) 视为标准不确定度 ($k = 1$), 则 DFT 计算结果位于 CE 预测结果的不确定度范围内. CE 拟合的不确定度源于训练集构型的有限性、团簇展开的截断近似等因素^[52]. 本文所有数据集 (包括所有构型的形成能数据、组分应变能数据、组分应变能修正 CE 模型的完整 ECI 列表) 均可在科学数据银行 <https://www.doi.org/10.57760/sciencedb.j00213.00207> 中访问获取.

4 结 论

本文围绕 Au-Cu 合金体系的 CE 建模, 构建了实空间 CE 建模所需的训练数据集及修正项建模所需的应变能训练数据集, 建立了包含长程修正项的完整 CE 模型, 主要结论如下:

1) 基于 PBE+ U^{DSCC} 方法, 构建了包含 241 个 Au-Cu 不等价构型的训练数据集, 计算得到的 Cu_3Au , CuAu , CuAu_3 化合物形成能与实验值的偏差明显低于 PBE 泛函结果. DSCC 方法通过自洽修正 d 轨道能级位置, 有效提高了 Au-Cu 体系形成能的预测精度, 在保持与 PBE 相当计算代价的同时, 避免了人工调参, 更利于开展高通量计算.

2) 基于单组元的外延应变能, 构建了包含 1452 条数据的 Au-Cu 体系组分应变能训练数据集, 定量刻画了 Au-Cu 体系中由显著原子尺寸失配引起的组分应变能贡献. 该数据集为考虑组分应变能的倒易空间 CE 建模提供了数据, 可进一步支撑 MSCE, MBCE 模型的构建. 相比于 CS-VLCE 和 CLDM 等修正建模所需的超胞应变构型数据集, 组分应变能训练数据集的计算开销较小, 在开展应变修正 CE 建模时可优先考虑.

3) 基于组分应变能修正的 CE 模型在富 Cu 端的能量预测精度明显提高, 能够再现实验观测到的 Cu_3Au , CuAu 与 CuAu_3 三种稳定化合物, 且整体拟合误差进一步降低. 在实空间 CE 模型中显式引入组分应变能项后, 实现了短程化学相互作用与长程弹性贡献的有效分离. 此类修正 CE 模型可以为描述短程化学作用与长程弹性应变效应耦合提供有效的理论工具.

4) 目前 CE 模型对稳定相 $\text{Cu}_2\text{Au}_5(\text{LPS})$ 的预

测结果在其不确定度 ($k = 1$) 范围内与实验相符, 后续需要进一步开展 CE 不确定度量化的研究.

数据可用性声明

支撑本研究成果的数据集可在科学数据银行 <https://www.doi.org/10.57760/sciencedb.j00213.00207> 访问获取.

感谢北京应用物理与计算数学研究所王越超副研究员、陈欣副研究员的讨论.

参考文献

- [1] Sanchez J M, Ducastelle F, Gratias D 1984 *Physica A* **128** 334
- [2] Sanchez J M 2010 *Phys. Rev. B* **81** 224202
- [3] Sanchez J M 2017 *J. Phase Equilib. Diffus.* **38** 238
- [4] Walle A, Ceder G 2002 *JPE* **23** 348
- [5] Connolly J W D, Williams A R 1983 *Phys. Rev. B* **27** 5169
- [6] Hua M, Tian X, Li S, Zhang X, Shao A, Song L, Lin X 2022 *Phys. Chem. Chem. Phys.* **24** 2251
- [7] Wu Q, He B, Song T, Gao J, Shi S 2016 *Comput. Mater. Sci.* **125** 243
- [8] Ke J H, Jokisaari A M 2023 *JOM* **75** 3208
- [9] Abe Y, Kubo A, Ukai S, Tsuru T 2026 *J. Nucl. Mater.* **618** 156221
- [10] Doležal T D, Valverde N A, Yuwono J A, Kemnitz R A 2024 *Mater. Design* **244** 113226
- [11] Jindal V, Lele S 2025 *Calphad* **89** 102825
- [12] Liu S, Martínez E, LLorca J 2020 *Acta Mater.* **195** 317
- [13] Fan W, Gong T, Hao W, Chen Y, Chen X Q 2025 *Comput. Mater. Sci.* **252** 113798
- [14] Shi S, Gao J, Liu Y, Zhao Y, Wu Q, Ju W, Ouyang C, Xiao R 2016 *Chin. Phys. B* **25** 018212
- [15] Zhang F Z, Xue H T, Tang F L, Li X K, Lu W J, Feng Y D 2016 *Chin. Phys. B* **25** 013103
- [16] Oates W A, Spencer P J, Fries S G 1996 *Calphad* **20** 481
- [17] Zhao J, Xia W, Zeng Z, Wang X 2024 *npj Comput. Mater.* **10** 71
- [18] Yuge K, Okawa R 2014 *Intermetallics* **44** 60
- [19] Ozoliņš V, Wolverton C, Zunger A 1998 *Phys. Rev. B* **57** 6427
- [20] Tian L Y, Levämäki H, Kuusma M, Kokko K, Nagy Á, Vitos L 2019 *Phys. Rev. B* **99** 064202
- [21] Mohri T, Chen Y, Horiuchi T 2024 *Comput. Mater. Sci.* **235** 112772
- [22] Pandey S, Koch R J, Li G, Misture S T, Wang H, Phillpot S R 2019 *J. Alloys Compd.* **809** 151615
- [23] Orr R L 1960 *Acta Metall.* **8** 489
- [24] Ogawa S, Watanabe D 1954 *J. Phys. Soc. Jpn.* **9** 475
- [25] Jaumot F E, Charles H S 1954 *Acta Metall.* **2** 63
- [26] Oriani R A 1954 *Acta Metallurgica* **2** 608
- [27] Kuczynski G C, Doyama M, Fine M E 1956 *J. Appl. Phys.* **27** 651
- [28] Zhang Y, Kresse G, Wolverton C 2014 *Phys. Rev. Lett.* **112** 075502
- [29] Nepal N K, Adhikari S, Bates J E, Ruzsinszky A 2019 *Phys. Rev. B* **100** 045135
- [30] Perdew J P, Burke K, Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
- [31] Himmetoglu B, Floris A, De Gironcoli S, Cococcioni M 2014 *Int. J. Quantum Chem.* **114** 14
- [32] Cococcioni M, De Gironcoli S 2005 *Phys. Rev. B* **71** 035105

- [33] Liu B L, Wang Y C, Liu Y, Liu H F, Song H F 2023 *J. Phys. Chem. Lett.* **14** 8930
- [34] Liu B L, Wang Y C, Gao X, Liu H F, Song H F 2025 *Phys. Rev. B* **111** 115139
- [35] Zhang C, Meng Y, Yan C, Tang X, Wang Y L, Zhang Q Y 2007 *Acta Phys. Sin.* **56** 452 (in Chinese) [张超, 孟旻, 颜超, 唐鑫, 王永亮, 张庆瑜 2007 物理学报 **56** 452]
- [36] Zhuravlev I A, An J M, Belashchenko K D 2014 *Phys. Rev. B* **90** 214108
- [37] Laks D B, Ferreira L G, Froyen S, Zunger A 1992 *Phys. Rev. B* **46** 12587
- [38] Wang K, Cheng D, Zhou B C 2023 *npj Comput. Mater.* **9** 75
- [39] Blum V, Zunger A 2004 *Phys. Rev. B* **70** 155108
- [40] Behara S S, Thomas J C, Puchala B, Van Der Ven A 2024 *Phys. Rev. Mater.* **8** 033801
- [41] Zunger A, Wang L G, Hart G L W, Sanati M 2002 *Modelling Simul. Mater. Sci. Eng.* **10** 685
- [42] Fontaine D D 1994 *Solid State Physics* (Academic Press) pp33–176
- [43] Eklborg-Tanner P, Rosander P, Fransson E, Erhart P 2024 *PRX Energy* **3** 042001
- [44] Rahm J M, Löfgren J, Erhart P 2022 *Acta Mater.* **227** 117697
- [45] Hart G L W, Forcade R W 2008 *Phys. Rev. B* **77** 224115
- [46] Kresse G, Furthmüller J 1996 *Comput. Mater. Sci.* **6** 15
- [47] Kresse G, Furthmüller J 1996 *Phys. Rev. B* **54** 11169
- [48] Kresse G, Hafner J 1993 *Phys. Rev. B* **47** 558
- [49] Blöchl P E 1994 *Phys. Rev. B* **50** 17953
- [50] Ozoliņš V, Wolverton C, Zunger A 1998 *Phys. Rev. B* **57** 4816
- [51] Ångqvist M, Muñoz W A, Rahm J M, et al. 2019 *Adv. Theory Sims* **2** 1900015
- [52] Aldegunde M, Zabarar N, Kristensen J 2016 *J. Comput. Phys.* **323** 17

SPECIAL TOPIC — Atomic, molecular and materials properties data

Construction of formation energy dataset and cluster expansion for Au-Cu alloys*

JIA Yuru¹⁾ LIU Beilei^{2)†} GAO Xingyu^{2)‡} CHEN Huajie^{1)††} SONG Haifeng²⁾

¹⁾ (School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China)

²⁾ (National Key Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, Beijing 100088, China)

(Received 17 March 2026; revised manuscript received 18 April 2026)

Abstract

This study presents a dataset construction method for the cluster expansion (CE) of Au-Cu alloys. To construct the formation energy training dataset, the PBE+ U^{DSCC} method is employed to correct the energy levels of the Cu-3d and Au-5d orbitals, where the U and J parameters are self-consistently evaluated using the doubly screened Coulomb correction (DSCC) method. This improves the calculation accuracy of the formation energy for the Au-Cu alloys to allow the prediction of three experimentally observed stable compounds, namely Cu₃Au, CuAu, and CuAu₃. To construct the constituent strain energy training dataset, we use the PBE+ U^{DSCC} method to calculate the epitaxial strain energies of pure Au and Cu along different crystallographic orientations, thereby fitting the constituent strain energy of Au-Cu alloys. Based on these two datasets, a CE model corrected by constituent strain energy is established, in which the long-range elastic strain effect arising from the significant atomic size mismatch is separated from the short-range chemical interactions. Compared with the pure real-space CE model, both the fitting accuracy and the prediction accuracy of the convex hull are improved. The datasets presented in this paper are openly available at <https://www.doi.org/10.57760/sciencedb.j00213.00207>.

Keywords: cluster expansion, doubly screened Coulomb correction, Au-Cu alloys, formation energy

DOI: 10.7498/aps.75.20260377

CSTR: 32037.14.aps.75.20260377

* Project supported by the Key Research and Development Program of Yunnan Province, China (Grant No. 202403AA080016).

† Corresponding author. E-mail: beilei_liu@126.com

‡ Corresponding author. E-mail: gao_xingyu@iapcm.ac.cn

†† Corresponding author. E-mail: chen.huajie@bnu.edu.cn

Au-Cu合金的形成能数据集构建与团簇展开建模

贾煜茹 刘倍雷 高兴誉 陈华杰 宋海峰

Construction of formation energy dataset and cluster expansion for Au-Cu alloys

JIA Yuru LIU Beilei GAO Xingyu CHEN Huajie SONG Haifeng

引用信息 Citation: *Acta Physica Sinica*, 75, 120807 (2026) DOI: 10.7498/aps.75.20260377

CSTR: 32037.14.aps.75.20260377

在线阅读 View online: <https://doi.org/10.7498/aps.75.20260377>

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

锂离子电池富锂锰基三元材料中氧空位簇的形成: 第一原理计算

Formation of oxygen vacancy clusters in Li-rich Mn-based cathode Materials of lithium-ion batteries: First-principles calculations

物理学报. 2023, 72(7): 078201 <https://doi.org/10.7498/aps.72.20222300>

H离子团簇高次谐波平台展宽与团簇膨胀

High-order harmonic platform extension and cluster expansion of H ion cluster

物理学报. 2023, 72(21): 214203 <https://doi.org/10.7498/aps.72.20230534>

Ni-Al-Cr合金中团簇加连接原子模型的第一性原理计算

First-principles calculations of Ni-Al-Cr alloys using cluster-plus-glue-atom model

物理学报. 2022, 71(20): 207101 <https://doi.org/10.7498/aps.71.20221036>

采用薄靶方法测量低能电子致Al, Ti, Cu, Ag, Au元素K壳层电离截面与L壳层特征X射线产生截面

Measurements of K-shell ionization cross sections and L-shell X-ray production cross sections of Al, Ti, Cu, Ag, and Au thin films by low-energy electron impact

物理学报. 2022, 71(17): 173402 <https://doi.org/10.7498/aps.71.20220162>

相场法研究 $\text{Fe}_{84}\text{Cu}_{15}\text{Mn}_1$ 合金富Cu相析出的内磁能作用机理

Mechanism of internal magnetic energy of Cu-rich phase precipitation in $\text{Fe}_{84}\text{Cu}_{15}\text{Mn}_1$ alloy by phase field method

物理学报. 2022, 71(8): 080201 <https://doi.org/10.7498/aps.71.20212087>

Ti-V-Ta多主元合金辐照位错环形成的级联重叠模拟

Cascade overlap simulation of formation of dislocation loops in Ti-V-Ta multi-principal element alloy

物理学报. 2024, 73(22): 226102 <https://doi.org/10.7498/aps.73.20241074>