arXiv:2302.08991v1 [cs.CE] 17 Feb 2023

利用机器学习实现尺度桥接：从第一性原理统计力学到连续相场计算，研究 Li_xCoO_2 中的有序-无序转变

G.H. Teichert[a], S. Das[a], M. Faghih Shojaei[a], J. Holber[f], T. Mueller[g, b], L. Hung[b], V. Gavini[a, c, f, e], K. Garikipati[a, d, f, e, *]

[a]*Department of Mechanical Engineering, University of Michigan*[b]*Toyota Research Institute, Los Altos, CA*[c]*Department of Materials Science & Engineering, University of Michigan*
[d]*Department of Mathematics, University of Michigan*
[e]*Michigan Institute for Computational Discovery & Engineering, University of Michigan*
[f]*Applied Physics Program, University of Michigan*[g]*Materials Science & Engineering, Johns Hopkins University*

摘要

Li_xTMO_2 (TM=Ni、Co、Mn) 是锂离子电池正极材料的重要家族，其性能在很大程度上受锂组成依赖的晶体结构和相稳定性支配。在此，我们以 Li_xCoO_2 (LCO) 为模型系统，对一个用于材料物理学尺度桥接的机器学习驱动框架进行基准测试。我们聚焦于两个尺度：(a) 由密度泛函理论指导的统计力学描述的数千个原子组成的体系；以及 (b) 用于研究 LCO 中有序-无序转变动力学的连续相场模型。尺度桥接的核心，是通过对特定原子构型的形成能进行粗粒化，对该材料体系的自由能密度和化学势进行严格且定量准确的表征。我们开发了主动学习工作流，用于训练近期开发的可积深度神经网络，以表示这类高维自由能密度函数和化学势函数。由此得到的、受第一性原理指导的机器学习驱动相场计算，使我们能够从温度、形貌、充放电循环和颗粒尺寸等方面研究 LCO 正极的相演化。

1. 引言

Li_xTMO_2（$x \in [0, 1]$，TM=Ni、Mn、Co，因此称为NMC材料）型层状氧化物通常用作锂离子电池的正极，因为其化学组成可根据具体应用进行调节，同时具有较高的能量密度和充电速率。其中，$LiCoO_2$（LCO）一直是消费电子产品的首选，而富镍组分因其高能量密度而对高能量密度电池具有吸引力，这对于交通电气化至关重要。NMC材料在循环过程中会发生相变，在某些情况下可能影响其性能。LCO在低于 $x = 1/3$ [1]时会发生从O3相到H1-3结构的有害堆垛序列相变。要定量控制层状NMC氧化物的循环性能和结构稳定性，就需要基于第一性原理理解锂嵌入的热力学、其对晶体结构稳定性和长期电化学性质的影响，以及一个涵盖从原子到连续介质各个尺度的建模框架。然而，尺度桥接是一个难题；正如我们下文所述，先前的研究将不同长度尺度彼此孤立地进行了处理。在此，我们展示了进一步发展尺度桥接框架的关键进展 [2, 3]，即在一个考虑所有对称性的高维空间中，以严格且系统的方式将第一性原理能量学与描述相

*通讯作者

演化的自由能联系起来。我们以LCO的O3相为例，阐释锂嵌入过程中热力学相演化以及有序-无序转变背后的自由能基础。

LCO丰富的相行为已通过实验得到广泛研究。对于组分 $0.75 \leq x \leq 0.94$，已确定存在一阶相变，其原因是金属-绝缘体转变 [4, 5]。Reimers and Dahn [4] 报道了 $x = 1/2$处的有序化现象，随后确认其为行有序 [6, 7]。Shao-Horn et al.[6] 发现了 $x = 1/3$ 处在 $\sim$100 K时发生有序化的证据。此外，在175 K时观察到 $x = 1/2$ 和 $x = 2/3$ 处Co^{3+} 与Co^{4+} 原子发生电荷有序，而在室温下这些电荷态的占据是随机的 [8, 7]。对于 $x < 0.33$[9, 10]，还确定了从O3结构到H1-3结构的转变。

此外，还开展了大量关于LCO的计算研究。第一性原理 热力学方法结合密度泛函理论（DFT）和统计力学 [1, 11] ，已用于系统地预测相图、自由能数据和电压曲线。第一性原理方法与实验和CALPHAD模型 [12] 相结合，利用传统的Redlich-Kister多项式 [13, 10]定义吉布斯自由能函数。DFT预测还与实验数据相结合，例如用于对观测到的金属-绝缘体相变进行相场建模 [14]。然而，尽管第一性原理 计算与统计力学计算之间的联系已经确立，但第一性原理方法尚未纳入全面的尺度桥接框架中，该框架可连接到相场模型等连续介质处理，以研究这些体系中相变的热力学和动力学。

当前工作展示了一个严谨的框架，将先进的第一性原理方法与我们近期在机器学习方面的工作相结合，从而桥接基于DFT的统计力学与连续尺度模拟（图1a）。我们选用O3层状LCO对该方法进行基准测试，原因在于人们已对这一化学体系开展了广泛研究。据我们所知，此处展示的系统性尺度桥接尚未应用于这种正极材料。

我们进行DFT+U 计算 [15] ，以捕捉具有局域化 d 轨道的过渡金属氧化物中的电子关联。结合范德华泛函后，该方法通过与层状LCO [16]的实验电压高度一致，验证了DFT方法的有效性。利用DFT计算对簇展开哈密顿量进行参数化，并将其用于半巨正则蒙特卡罗采样，以计算锂嵌入热力学随温度和化学势的变化，并构建温度—组分相图 [17, 18]。

材料的晶体结构决定了在特定组分下，插入原子能够在晶格上采取的有序排列。自由能是组分、描述这些有序排列的参数以及应变和温度等其他变量的函数。对于本文所考察的LCO相稳定性范围，根据对称群的考虑确定的六个参数即可完整描述其有序性（§1.2.3、3.2.2）。统计力学模拟得到自由能密度相对于组分和序参量的导数，即化学势。作为本文核心的机器学习框架，通过自适应学习，根据自由能导数数据，将自由能密度表示为组分和六个序参量的七维函数。由于未纳入应变和振动熵，因此未考虑力学和热输运。作者此前提出的可积深度神经网络（IDNNs） [2, 3], 可以利用导数数据进行训练，例如使用通过蒙特卡罗采样获得的化学势，并通过解析积分恢复原函数，例如相场模型所需的自由能密度函数。这需要对非凸区域、极值、可行区域边界以及七维自由能密度中的高误差点进行自适应采样。

这些机器学习热力学表示具有可量化的精度，其精度由收敛容差控制。利用这些表示，我们确认了其与文献中广泛报道的LCO相图以及实验电压测量结果相吻合。我们还开展了大规模DFT计算，以计算不同相之间的反相界面能和界面能。最后，利用所得的自由能密度和反相畴界/界面能，我们对孤立颗粒中有序-无序转变的动力学，以及实际相关温度下单个LCO颗粒在充放电循环期间的动力学开展了相场研究。

1.1. 基于密度泛函理论的形成能和构型

由DFT+U预测的平均电压随锂组成 $x = \{0, 1\}$、$\{0, 0.5\}$、$\{0.5, 1\}$ 以及 U 值增加的变化绘于图1b中。我们发现，在计算LCO形成能时采用 $U = 2.5$ 电子伏特，可与参考文献 [16] 中报道的 $x = \{0, 1\}$ 和 $x = \{0.5, 1\}$的实验电压良好吻合。DFT+U 计算基于CASM（簇统计力学方法）软件 [22]识别出的333个构型进行。计算得到的形成能及其对应的凸包，以及预测的基态，绘于图1c中。$LiCoO_2$ 的晶胞体积计算为 32.502 $^{\circ}A^3$。

图1d展示了本研究中在 $x = 1/2$ 处出现的锂的三种有序构型。我们发现，之字形有序是基态，其形成能为 $E_{\mathrm{f}} = -0.170$ 电子伏特，而另外两种行构型的形成能分别为（$E_{\mathrm{f}} = -0.150$ 和 -0.164 电子伏特）。在我们的密度泛函理论结果中，只有图1d左侧的行构型出现了电荷分裂。

1.2. 统计力学

1.2.1. 针对 E_f 的簇展开

我们使用遗传算法在CASM代码中为形成能 E_{f}开发了一个簇展开模型，该算法从最初的221个基函数中选出了35个。该簇展开预测的均方根误差为每个原胞2.49电子伏特，这是通过10折交叉验证得到的。相应的有效簇相互作用（ECIs）如图1e所示。簇展开预测的 E_{f} 值与密度泛函理论得到的值+U 一起绘制于图1c中。请注意，相应的凸包在形成能和构型方面相匹配。

1.2.2. 有序/无序结构的蒙特卡罗模拟

通过半巨正则蒙特卡罗计算了温度范围为 200 K 至 400 K 时，锂含量随化学势和温度的变化。采用了无偏采样和有偏（伞式）采样 [24, 25, 26]。在无偏数据中，两相区域很容易表现为组分值的间隙（图 2a）。利用这些数据构建了 O3型LCO 的相图，其锂组分为 $x \geq 1/3$ ，如图 2c 所示。

根据蒙特卡罗结果， $x \approx 1/2$ 处有序排列的有序-无序转变温度略高于 330 K，与实验结果一致 [4]。在 300 K 下计算得到的电压曲线与实验测量结果也较为吻合 [23] （图 2b ）。在 $x \approx 1/2$ 附近的波动中可以观察到有序-无序转变。与以往针对 LCO 的第一性原理研究类似 [1, 18],，然而，这些计算未能体现导致在 $x \sim 0.7$-0.9 范围内观察到两相区域的金属-绝缘体转变。因此，模拟中未能捕捉到相应的电压平台。本文的蒙特卡罗结果及预测电压仅针对 O3结构，因此仅适用于锂组分为 $x \geq 1/3$的情况。

1.2.3. 对称性适配序参量

对称性适配序参量精确表示了之字形有序的不同变体。根据密度泛函理论，在组分为 $x = 1/2$ 时，在三角形锂子晶格的所有可能装饰中，之字形有序的能量最低。之字形有序具有 12 个变体，如图3a所示。这些变体源自属于三角晶格对称群的 3 个旋转 $\times$ 4 平移。随后计算了一个互相公度超胞，该超胞包含所有旋转和平移对应的超胞，并在连续的两个锂层上包含 32 个独特的子晶格位点（图3b-3c）。由此得到一个 32 维基，在该基上，12 个向量 $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(12)}\}$ 唯一描述这 12 个变体（图3c-3d）。

通过七个序参量 η_0、.....、η_6可以更高效地表示这12个变体；这些序参量对应于图3e中矩阵 $\boldsymbol{\eta}$ 的前七行，其余各行为零。其中， $\eta_0 = x$ 对应于对全部32个子晶格位点求平均得到的组分。它与正交矩阵 $\boldsymbol{Q} \in \mathbb{R}^{32\times32}$ 的一个独特特征值相关，该矩阵对于之字形有序的对称群保持不变（图3e和§3.2.2）。另外六个序参量与简并特征值之一的特征向量相关，并通过图3e中的 η_1、 $\ldots \eta_6 = \pm 1/2$ 表示这12个变体。

1.3. 大规 有序变体之间反相界面能的电子伏特尺度 *DFT*计算 s

不同 LCO 变体在有序组成$x = 1/2$ 下的共格反相界面能，是相场模拟所需的另一项关键输入。采用 $\eta_1 = 1/2$和 $\eta_3 = 1/2$ 定义的变体（图3f），我们进行了大规模 DFT计算，预测反相界面能为 $\gamma = 30.9$ mJ/m^2 （见 §3.1.1）。根据有序-无序界面能与反相界面能之间约为 $1/2$ 倍的估算关系 [27], ，我们将 ~ 15.45mJ/m^2 用作有序-无序界面能。

1.4. 主动学习 用于七维自由能曲面的可积深度神经网络（*IDNN*） e

图4a、b分别展示了示意图，以及针对按训练数据集编号索引的 IDNN 模型绘制的主动学习曲线。有关 IDNN 的详情，见 §3.3。每个 IDNN 模型的均方误差（MSE）均使用最后一个（25th）数据集计算。每个温度下均以实心圆标出均方误差最低的最佳模型；进行过超参数搜索的模型则以空心圆标出。图4c展示了每个温度下最佳模型在 $\eta_0 - \eta_1$ 和 $\eta_1 - \eta_2$ 切片上的自由能密度曲面。主动学习得到的 IDNN 在 260 K 时有两个隐藏层、每层 173 个神经元；在 300 K 时有三个隐藏层、每层 163 个神经元；在 340 K 时有两个隐藏层、每层 193 个神经元。

1.4.1. 保持对称性的函数作为 *IDNN* 特征

IDNN 表征自由能密度的特征是以下多项式，其关于 η_0，$\ldots \eta_6$ 的阶数最高为六阶。它们在三角形Li子晶格的变换下保持不变，这些变换可在有序变体之间进行映射。因此，以这些多项式表示的 IDNN 表征继承了自由能的对称性。这些保持对称性的 IDNN特征通过对 η_0、.....、 η_6 [28] 中的单项式应用 Reynolds算子确定：

$$
\begin{aligned}
h_1 &= \eta_0 \\
h_2 &= \frac{2}{3}\sum_{i=1}^{6}\eta_i^2 \\
h_3 &= \frac{8}{3}\sum_{i=1}^{6}\eta_i^4 \\
h_4 &= \frac{4}{3}\big[\left(\eta_1^2+\eta_2^2\right)\left(\eta_3^2+\eta_4^2+\eta_5^2+\eta_6^2\right)+ \\
&\qquad \left(\eta_3^2+\eta_6^2\right)\left(\eta_4^2+\eta_5^2\right)\big] \\
h_5 &= \frac{16}{3}\left(\eta_1^2\eta_2^2+\eta_3^2\eta_6^2+\eta_4^2\eta_5^2\right) \\
h_6 &= \frac{32}{3}\sum_{i=1}^{6}\eta_i^6 \\
h_7 &= \frac{8}{3}\big[\left(\eta_1^4+\eta_2^4\right)\left(\eta_3^2+\eta_4^2+\eta_5^2+\eta_6^2\right)+\left(\eta_3^4+\eta_6^4\right)\left(\eta_4^2+\eta_5^2\right)+ \\
&\qquad \left(\eta_1^2+\eta_2^2\right)\left(\eta_3^4+\eta_4^4+\eta_5^4+\eta_6^4\right)+\left(\eta_3^2+\eta_6^2\right)\left(\eta_4^4+\eta_5^4\right)\big] \\
h_8 &= \frac{16}{3}\big[\eta_1^2\eta_2^2(\eta_3^2+\eta_4^2+\eta_5^2+\eta_6^2)+\eta_3^2\eta_6^2(\eta_1^2+\eta_2^2+\eta_4^2+\eta_5^2)+ \\
&\qquad \eta_4^2\eta_5^2(\eta_1^2+\eta_2^2+\eta_3^2+\eta_6^2)\big] \\
h_9 &= \frac{32}{3}\left(\eta_1^4\eta_2^2+\eta_1^2\eta_2^4+\eta_3^4\eta_6^2+\eta_3^2\eta_6^4+\eta_4^4\eta_5^2+\eta_4^2\eta_5^4\right) \\
h_{10} &= 8(\eta_1^2+\eta_2^2)(\eta_3^2+\eta_6^2)(\eta_4^2+\eta_5^2) \\
h_{11} &= \frac{64}{5}\big[\eta_1\eta_2(\eta_3^2-\eta_6^2)(\eta_4^2-\eta_5^2)+\eta_3\eta_6(\eta_1^2-\eta_2^2)(\eta_4^2-\eta_5^2)+ \\
&\qquad \eta_4\eta_5(\eta_1^2-\eta_2^2)(\eta_3^2-\eta_6^2)\big] \\
h_{12} &= 64\sqrt{5}\eta_1\eta_2\eta_3\eta_4\eta_5\eta_6 \qquad (1)
\end{aligned}
$$

1.5. 相场模拟

IDNN 学习得到的表征表明，在 260 K 下，有序相发生异质成核时自由能密度降低了 1 meV/晶胞，而在 300 K 下降低了 0.5 meV/晶胞（图 S1）。利用 DFT 得到的反相界面能 $\gamma = 30.9$ mJ/m^2 和有序-无序界面能 ~ 15.45 mJ/m^2, ，我们得到的有序相在 260 K 下的成核速率远高于 300 K。图5a显示了260 K和300 K下成核速率相对于外加电压的变化，以及相应的无序基体组分（x_{mat}）（更多详情见 §S3）。在 340 K 下仅存在无序相，这一点可从相图（图2c）和学习得到的自由能密度曲面（图4c）中看出。

该尺度桥接框架应用于两项涉及约 $x = 1/2$ 发生有序-无序转变的 LCO正极相场研究。Cahn-Hilliard 和 Allen-Cahn 相

场方程（9-10）中的梯度参数分别根据反相畴界/界面能以及 260 K（300 K）温度下的自由能面 IDNN 表征（见 §3.5）确定为 $\chi_0 = 1 \times 10^{-4}$ mJ/m（1.88×10^{-4} mJ/m）和 $\chi_1, \ldots, \chi_6$ =2.12×10^{-8} mJ/m（4.91×10^{-8} mJ/m）。

第一项相场研究考察了无外部锂通量时二维颗粒中有序-无序转变的动力学。图5b显示，在直径为 $1\ \mu$ m、平均组分为 $x = 0.525$ 的颗粒中，260 K和300 K下的动力学过程分别为 $x = 0.53$ ，其形态与旋节分解后有序相畴发生奥斯特瓦尔德熟化的过程十分相似。这些组分位于相图的有序-无序区域内。在每个温度下，第一行显示组分， x 第二行显示有序状态 $\eta_1, \ldots, \eta_6$。对七维自由能面的研究表明，在每个有序区域内，十二种变体中只有一种形成（§S7和图S7）。由于所需计算时间较长，这些动力学模拟未运行至获得平衡微观结构（在260 K和300 K下）。然而，在考虑到IDNN学习得到的自由能密度所预测的有序-无序边界略有偏移后，由耦合相场方程在每个温度下预测的有序相与无序相的质量分数，与将杠杆定律应用于相图所得的结果非常一致（§S7、§S8和图S8）。这表明计算已达到近平衡状态。需要注意的是，与300 K相比，260 K下有序区域的数量更多，反相界面/界面也更窄，这是由于相应的界面能更低所致。还应注意，与300 K相比，260 K下的成核速率更高，但反相界面的运动更慢。成核速率的计算见§S3（参见图S1–S4）。

对应于 $\eta_i = \pm 1/2$ 的有序变体之间的界面宽度，大于由 $\eta_i = 1/2$ 或 $-1/2$ 以及 $\eta_j = 1/2$或 $-1/2$表示的一对变体之间的界面宽度，其中 $i \neq j$的实例如图5b第五行的插图所示。这是由 IDNN 学到的对称性所导致的；这些对称性是在上述通过 Reynolds算子（§1.4.1）施加的特征之外额外获得的，后文将对此作进一步讨论。

第二项相场研究同样针对260和300 K下的二维颗粒，重点研究通过指定随时间变化的外加锂通量 [29] 对钴酸锂颗粒充放电循环的影响。见（图 6）。在每个温度下，均将图5研究中无序基体内有序变体的最终构型作为初始条件。施加了初始电流密度0.6677 A m^{-2} 。在10、16、26和40 s时，施加的电流密度分别调整为-1.0683、0.6677、-0.3338和0 A m^{-2} 。该循环对应的 C倍率为3.125、-5、3.125、-1.5625和0 C，其中正号表示放电，负号表示充电。在组成范围内，放电过程（图 6中的a-c和e-g）注入锂，推动颗粒进一步进入无序区域并使有序畴缩小；相反，充电过程（c-e和g-i）提取锂，使有序畴长大。请注意，随着反相界面迁移且曲率减小，不同有序畴仍在持续长大或缩小。在260（300）K下，该循环导致净放电，平均锂组成从0.5425（0.5225）开始，在 ~ 4.15 V的电压平台上增加至平均最终组成为0.5455（0.5254）（见图2b）。

2. 讨论

我们的研究规模十分庞大，从数百次用于提供信息的 DFT计算开始，扩展到数万次半巨正则蒙特卡罗计算。主动学习框架引导蒙特卡罗采样，以获取用于训练连续体均匀自由能密度函数 IDNN 表征的数据。结合由其他大规模 DFT计算得到的反相界面能所提供信息的梯度自由能，这一连续体自由能描述实现了详细的第一性原理信息相场研究。据我们所知，此类系统性的尺度桥接框架尚未应用于任何材料体系。在机器学习的指导下，该框架使我们能够在每个阶段对预测结果的精度进行量化。为了将我们的工作置于适当背景下，我们将从 DFT计算入手，依次讨论第一性原理统计力学、机器学习和连续体计算，并阐述其研究范围、悬而未决的问题及未来发展。

我们已经开展了 了关于 O3层状结构的 DFT+U 计算，该结构是 LC
O 的稳定形式，这参数是通过与实验测得的平均 ge

电压进行匹配获得的，匹配区间为 $x = 0$ 和 $x = 1$、 $x = 0$ 和 $x = 1/2$，以及 $x = 1/2$ 和 $x = 1$（见图1b）。另一种方法是选择 U，通过迭代更新 E_{f} 的簇展开，并预测中间 x 处的电压，从而在图2a所示的整个组成范围内实现最佳拟合。利用自洽 Hubbard 方法 [30]，则可以完全避免通过经验方法确定 U 参数。

接下来讨论我们的 DFT 研究预测的构型。我们注意到，实验结果支持在 $x = 1/2$ 图1d [4, 6, 7]中观察到的行排序。Takahashi 等人 [7] 提出了左侧所示的行有序结构，而 Shao-Horn 等人 [6] 则提出了中间所示的另一种行有序结构。在他们的 DFT 研究中，Van der Ven 等人 [1] 发现中间的行有序结构形成能最低，右侧的之字形有序结构紧随其后。不过，他们指出，其研究中的能量差约为 1 毫电子伏特，小到仅凭 DFT计算难以确定基态。我们的结果与早期 DFT 工作在形成能上的差异，可能是由于我们采用了 DFT+U、vdW-DF 和广义梯度近似（GGA），而先前的研究使用的是局域密度近似（LDA）。尺度桥接框架同样适用于其他有序结构；本文以之字形有序结构为例进行展示。

鉴于这种之字形结构，η_0、……、η_6 对尺度桥接框架十分重要，因为与子晶格向量 $\boldsymbol{x} \in \mathbb{R}^{32}$相比，它们能够更高效地表示 $\mathbb{R}^7$ 中的 12 种变体；其中 $\eta_i = \pm 1/2$、 $i = 1$、……、6 代表完美有序结构（图3e）。然而，在 $\boldsymbol{\eta}$ 空间中对误差较高的区域、自由能井、有序-无序转变区域以及边界附近进行蒙特卡罗采样是一项挑战。对此，我们采用了主动学习工作流（图4a）；如图4b所示，该工作流改进了 IDNN 在各数据集上的训练效果，但并非单调改善。随着主动学习引导的利用与探索不断增加，蒙特卡罗采样数据逐渐增多，IDNN 均匀自由能密度模型所需的表示复杂度也越来越高。如果最近两次 IDNN 模型更新相对于最新数据集的均方误差持续增加，则会重新进行超参数搜索。将该 IDNN 模型绘制在 $\mathbb{R}^7$ 的二维切片上（图4c）可见，在 300 K 下，模型比 260 K（两个隐藏层，每层 173 个神经元）和 340 K（两个隐藏层，每层 193 个神经元）下更为复杂，包含三个隐藏层，每层 163 个神经元。

除了均匀自由能密度之外，相场计算还需要任意两个12种变体之间反相界面的能量。大规模DFT计算得出了 $\gamma = 30.9$ mJm^{-2} 之间 $\eta_1 = 1/2$ 和 $\eta_3 = 1/2$ 变体的反相界面能（图3f）。然而，由于随机采样的无序构型会随尺寸变化，有序-无序界面能的计算给系统收敛性研究带来了困难。因此，我们采用了文献 [27] 中提出的方法，即有序-无序界面能约为有序-有序反相边界能的一半。尽管有序/无序状态会引起局部原子弛豫，但晶格仍保持相干。因此，反相界面和有序-无序界面处的弹性应力预计不会很大；如实验文献 [31, 32, 33]所述，这些转变虽然对相动力学很重要，但可能不会导致退化。

这种从第一性原理统计力学到连续体自由能函数的机器学习驱动的尺度桥接，实现了定量严谨的相场计算。这些计算揭示了可与较低尺度关联的复杂性：由同一序参量的极值表示的（A）有序变体之间形成的反相界面，比由 $\eta_i = 1/2$ 或$-1/2$ 以及 $\eta_j = 1/2$ 或 $-1/2$、 $i \neq j$ 表示的（B）型变体之间的反相界面更厚，$\eta_i = \pm 1/2$ （§1.5，图5b底行）。A型必须包含一个无序区域（η_1，……，$\eta_6 = 0$），而B型中不存在该区域。图4c中的二维均匀自由能密度切片解释了这一点。考虑一个具有 $x = \eta_0 = 1/2$和 η_3，……，$\eta_6 = 0$的B型示例。均匀自由能密度曲面上存在连接具有 $\eta_1 = 1/2$ 或 $-1/2$ 与 $\eta_2 = 1/2$或 $-1/2$ 的畴的路径，而这些路径无需经过 η_1、$\eta_2 = 0$。因此，B型具有有序结构的区域之间的转变不包含完全无序界面，并表现为尖锐转变（图5b，底部插图）。相比之下，如图4c所示，对于 $x = \eta_0 = 1/2$和 η_3，……，$\eta_6 = 0$，A型变体之间的路径：(η_1，η_2) =(−1/2, 0)与(η_1，η_2) =(1/2, 0)必须经过点 $x = \eta_0 = 1/2$和(η_1，……，η_6) =(0，……，0)—一个完全无序区域。该反相界面呈现出更大的宽度。Allen-Cahn方程（10）表明，在处于平衡态的实际颗粒中，序参量分布

沿垂直于A型反相边界的路径可由位置的双曲正切函数很好地近似。对于B型变体，IDNN学习到了由尺度桥接所提供信息、受物理约束的对称性，其形式为 $\mu_i = \mu_j$ ，沿着 $\boldsymbol{\eta} \in \mathbb{R}^7$中的路径 $\eta_i = \pm\eta_j \pm 1/2$ 。这些对称性不同于§1.4.1、公式（1）中施加于IDNN特征的对称性。

除平衡结构外，有序-无序转变的相动力学也值得关注。这需要在组分为 x_{nuc} 的无序基体中，于组分为 x_{mat}处形成有序变体；这一过程取决于相应自由能密度的降低。对于均匀成核，我们的跨尺度研究结果见 $\Delta g_{\mathrm{hom}}^{260} \approx 2\Delta g_{\mathrm{hom}}^{300}$。因此，260 K 下的成核速率显著高于 300 K。在外场 V 下，异质成核的临界自由能降低量为 $\Delta g_{\mathrm{het}}^{*} = f(\theta)(16\pi\gamma^2/3(\Delta g_{\mathrm{hom}}^{3} \pm V(x_{\mathrm{mat}} - x_{\mathrm{nuc}}))$，具体取决于 V的符号，其中 $f(\theta)$是晶核接触角的函数。我们发现 $x_{\mathrm{mat}} - x_{\mathrm{nuc}} \sim 0.01$。参见 §S3 和图 S1b、S1c [34, 35, 36, 37]。在$x_{\mathrm{nuc}} = 1/2$下，图 5a 展示了成核速率和概率随电压及 x_{mat}的变化。由零跃升至指数级高成核速率、概率由 0 升至 1 的极陡、近乎不连续的转变，确保了一旦电压与 x_{mat} 的组合有利，无论在 260 K 还是 300 K，有序相的形成都会近乎瞬时完成。

通过 IDNN 均匀自由能密度函数模型本身，可以进一步深入了解 $x = 1/2$ 处的有序-无序转变。考察在这一七维流形上遵循的最低自由能路径，并结合 η_0 的 Cahn-Hilliard方程和 η_1的 Allen-Cahn方程所描述的相场动力学，可以深入了解这些转变。若忽略有序化，则可得到一个仅以组分 $x = \eta_0$为参数的一维均匀自由能密度。在 $x = 1/2$ 附近会出现非凸性，这表明旋节分解可能是一种相分离机制（§S1.5、§S4、图 S4 和图 S8a）。然而，对七维 IDNN 自由能密度模型的分析表明，在有序区域内，六个序参量中有五个保持为零，而另一个则相对于 η_0几乎不连续地变化至 $\sim 1/2$ （见 §S7、图 S9）。在 $g - \eta_0$ 平面中观察到的二维切片（图 4c 和图 S7）显示斜率存在不连续性，即化学势 μ_0。IDNN 仅以 η_0 为输入学习得到的一维自由能密度，会将这一 μ_0不连续性平滑为 $g - \eta_0$ 曲线的非凸性（§S7 和图 S8a），由此得出上述旋节分解的推断。然而，完整的七维 IDNN 模型揭示了一条更为复杂的轨迹：当 η_0 进入有序化区域时，在某些 $i \in \{1, \ldots, 6\}$条件下， $g - \eta_0 - \eta_i$ 空间中会出现非凸性。图 S7 利用关于$\eta_1, \ldots, \eta_6$的对称性，展示了 $g - \eta_0 - \eta_1$ 空间中的这种情况。然而，这并非仅仅是 g相对于 η_0 的非凸性，而是沿着 $(\eta_2, \ldots, \eta_6) = (0, \ldots, 0)$ 对应的 Hessian $\partial^2 g/\partial\eta_0\partial\eta_1$ 矩阵的特征向量方向排列的非凸性。这是发生在 $g - (\eta_0, \ldots, \eta_6)$ 空间中的旋节分解，无法仅通过考察 $g - \eta_0$ 空间来完整描述。七维视角揭示了决定 LCO 嵌锂极限处化学势、进而决定电压的相变。因此，这一视角有助于全面理解电池性能，并最终实现对其的控制。图 S8b 在相图中标出了与上述分析相关的点。在机器学习的助力下，跨尺度研究以可量化的精度揭示了动力学细节，并将其直接关联到第一性原理统计力学和 DFT计算。

如图5b所示，在 $260 - 300$ K（最高可达330 K）的温度范围内，对处于实际电池相关温度窗口中的正极颗粒进行快速充放电循环时，其有序-无序形貌可能出现显著变化。有序-无序转变不会产生晶格失配，因此预计不会导致应力诱发的损伤 [38, 33]。对于所研究的组成范围，在340 K以上不存在有序-无序转变（图2c和图4c），因此即使是原本可能在反相界面形成的低应力也得以消除。这或许可以推动采用能够稳定无序结构的元素进行掺杂。最近的研究观察到，Al和La掺杂能够稳定有序-无序转变，这一机制具有相关性 [33]。

较小颗粒具有更均匀的形貌。这一点已得到充分理解：模拟中瞬态微结构的长度尺度由梯度系数决定，而梯度系数与反相边界和界面能有关，并且独立于颗粒尺寸。当颗粒尺寸从 1 μm 减小到 50 nm 时，所得微结构会变得不那么复杂。参见 §S4 图 S6 和文献 [39]。在本文所考察的倍率下，对 50 nm 颗粒进行循环充放电时未观察到温度差异。

颗粒足够小，因此迁移率、颗粒尺寸和微观结构长度尺度的相对数值会使锂原子均匀且极其迅速地扩散；与模拟的时间尺度相比，这一扩散速度如此之快，以至于观察不到任何温度效应（图 S6）。

我们注意到，对于组分 $x < 1/3$，需要完成针对 H1-3结构的 DFT+U 研究。H1-3结构是在较低组成范围内稳定的形式，并且需要将其纳入我们的尺度桥接框架，从而将该框架扩展至整个组成范围。系统地纳入其他影响因素仍有待完成。这些因素包括弹性、晶格结构变化和振动熵，它们可以扩展至连续体热力学。在活性颗粒-电解质尺度上扩展电化学方程，将能够研究纳米颗粒电池 [40]中观察到的马赛克不稳定性等现象。我们的预测依赖于 DFT 方法的可靠性，而这些方法也远非第一性原理。因此，我们并不声称能够提出定量上精确的计算预测。然而，我们已经建立了一个尺度桥接框架，能够系统地将第一性原理统计力学中的定量信息注入连续相场计算中，并且其精度还可以进一步提高。这将使我们能够研究形貌如何在各种特定于微结构的、环境、动力学和循环条件下演化，并以此前在相关长度尺度上无法达到的计算保真度开展研究。通过这些模拟，我们可以将结果与 LCO [38, 31, 33, 32, 41]实验文献中的观测建立联系。

3. 方法

3.1. 用于计算能量的*DFT*计算

具有锂组成 x 的构型的形成能 E 可根据以下方程 [1]，利用通过密度泛函理论（DFT）计算得到的总能量值求得：

$$E_{\mathrm{Li}_x\mathrm{CoO}_2} = E^{\mathrm{tot}} - xE^{\mathrm{tot}}_{\mathrm{LiCoO}_2} - (1-x)E^{\mathrm{tot}}_{\mathrm{CoO}_2} \tag{2}$$

其中， E^{tot}、 $E^{\mathrm{tot}}_{\mathrm{LiCoO}_2}$和 $E^{\mathrm{tot}}_{\mathrm{CoO}_2}$ 分别为给定构型、LiCoO_2, 和CoO_2,的总能量。

我们采用简化的DFT旋转不变公式+U [15]和vdW-DF交换-相关泛函，即optB88交换-相关泛函 [42, 43, 44], , 使用Quantum Espresso [45, 46]计算所选LCO构型子集的形成能。CASM（簇统计力学方法）软件套件用于确定具有O3晶体结构的一组构型，以参数化簇展开。

我们校准Hubbard U 参数，使其在不同锂组成范围内与实验平均嵌锂电压相匹配。 x_1 至 x_2 的平均电压根据以下方程[47, 48, 16]计算：

$$V = -\frac{E_{\mathrm{Li}_{x_2}\mathrm{CoO}_2} - E_{\mathrm{Li}_{x_1}\mathrm{CoO}_2} - (x_2 - x_1)E_{\mathrm{Li}}}{(x_2 - x_1)e} \tag{3}$$

其中， E 是通过DFT计算得到的总能量， e 是电子所带的电荷。我们计算锂组成的平均电压：$\{x_1, x_2\} = \{0, 1\}$， $\{0, 1/2\}$ 和 $\{1/2, 1\}$。通过与实验电压（图1b）进行比较，我们选取了合适的 $U = 2.5$ 电子伏特值。关于形成能的DFT计算的更多细节见§S1.1。

3.1.1. 界面能

我们对 $\eta_1 = 1/2$ 和 $\eta_3 = 1/2$ 有序 LCO 旋转变体之间反相边界的共格界面能进行了大规模 DFT计算。这些计算使用 DFT-FE 软件 [49, 50, 51, 52], 完成；DFT-FE 是一种近期开发的开源代码，支持大规模并行，基于有限元离散开展大规模实空间 Kohn-Sham 密度泛函理论研究。我们采用 PBE 交换-相关泛函和优化范德堡范数守恒赝势（ONCV） [53] ，赝势来自 Pseudo Dojo 库 [54]。

表 S1 展示了用于确定反相界面能 γ的能量和系统尺寸；该界面能估计为 30.9 mJ/m^2。关于周期性模拟晶胞、反相边界的处理、弹性失配应变能的计算以及收敛容差的详细信息，见 §S1.1.1。

3.2. 统计力学

3.2.1. 用于形成能的簇展开

我们采用簇展开来处理统计力学研究中所需的大量构型。形成能写作 $E_{\mathrm{f}}(\boldsymbol{\sigma})$，其中 $\boldsymbol{\sigma}$ 是构型向量；当锂占据位点 i 时，占位变量 $\sigma_i = 1$ 取值为 $\sigma_i = 0$ ，而当该位点包含空位时则取值为 $\sigma_i = 0$ 。我们通过使用密度泛函理论计算部分构型的形成能来定义 $E_{\mathrm{f}}(\boldsymbol{\sigma})$，并利用这些数值对簇展开 $[55, 56]$ 进行参数化，将其作为形成能的快速查询替代模型。我们使用 CASM 软件[22], ；该软件有助于构建和参数化簇展开哈密顿量，并将其用于蒙特卡罗模拟，从而为本文的 DFT计算选择构型并执行统计力学计算 $[57, 58, 59]$。关于簇基函数、有效簇相互作用系数、回归技术以及在候选基函数中进行选择的算法的定义细节，见 §S1.2.1。

3.2.2. 对称性适配序参量

我们遵循 Natarajan 等人的方法 $[60]$ ，识别表示图3a中各变体的对称性适配序参量。之字形有序的对称群 P 由 384 个唯一的线性变换组成，这些变换作用于 12 个变体之间，每个变体均由 $\mathbb{R}^{32\times32}$ 中的一个矩阵表示。按照参考文献 $[61]$,中的算法，我们构造了一个 P不变矩阵，并对其进行了特征值分解，得到 8 个非零特征值：2 个互不相同的特征值、1 个重复 3 次的特征值，以及 1 个重复 6 次的特征值；同时得到与之对应的 8 组特征向量，共计 32 个特征向量，它们构成了正交变换矩阵$\boldsymbol{Q} \in \mathbb{R}^{32\times32}$的各行。该矩阵将亚晶格组成向量 $\boldsymbol{x} \in \mathbb{R}^{32}$ 映射到序参量空间 $\boldsymbol{\eta} \in \mathbb{R}^{32}$。

采用从 0 开始的索引时，对于所有变体 $i = 1, \ldots, 12$， $\boldsymbol{\eta} \in \mathbb{R}^{32}$ 的分量 $\eta_7^{(i)}$ 至 $\eta_{31}^{(i)}$ 均为零，与描述之字形有序无关（图3e）。在 7 个相关序参量中，η_0 与一个互不相同的特征值相关，对应于对全部 32 个子晶格位点的组分取平均；即 $\eta_0 = x$。另外 6 个序参量 $\eta_1, \ldots, \eta_6$ 与一个简并特征值相关，该特征值对应 6 个特征向量。

由于自由能密度在三角形Li子晶格的变换下保持不变，而这些变换可将有序变体相互映射，因此，IDNN 表示采用了在这些变换下关于 η_0， $\ldots \eta_6$ 对称的函数作为特征。我们选取了最高六阶的单项式，并通过对其求和施加 Reynolds算子：

$$h(\boldsymbol{\eta}) = \sum_{\boldsymbol{M}^{(\eta)} \in \mathcal{M}} f(\boldsymbol{M}^{(\eta)} \boldsymbol{\eta}), \tag{4}$$

其中，对于每个 $\boldsymbol{M} \in P$，我们有 $\boldsymbol{M}^{\eta} = \boldsymbol{Q}\boldsymbol{M}$。该操作得到公式（1）中的 P-不变多项式函数，并将其作为 IDNN 特征。关于 $\boldsymbol{\eta}$ 空间中序参量的保持对称性的函数，详见 §S5

3.2.3. 蒙特卡罗采样

给定 $E_{\mathrm{f}}(\boldsymbol{\sigma})$，我们在半巨正则系综中进行采样，其中化学势是指定的，并通过系综平均确定相应的组分和/或序参量。然而，这种方法无法在与相分离相关的不稳定区域内产生数据。另一方面，相场模拟需要这些两相区域内的自由能信息，才能一致地解析相界面。我们在 260 K、300 K 和 340 K 温度下，使用偏置势进行伞形采样，在有序-无序转变的不稳定

区域内额外进行了蒙特卡罗计算 [60, 24, 62, 25, 26]。配分函数为 :

$$\Theta = \sum_{\boldsymbol{\sigma}} \exp\left(-\frac{E(\boldsymbol{\sigma}) + \sum_{i=0}^{6} \phi_i(\eta_i(\boldsymbol{\sigma}) - \kappa_i)^2}{k_B T}\right) \tag{5}$$

其中， ϕ_i 和 κ_i 分别决定偏置势的曲率和中心，内层求和遍历组分和六个序参量。组分 $\langle \eta_0 \rangle$ 以及每个序参量 $\langle \eta_i \rangle$、 $i = 1$、.....、6 的系综平均值通过偏置参数与其相应的化学势相关:

$$\frac{1}{M}\mu_i\Big|_{\langle \boldsymbol{\eta} \rangle} = -2\phi_i(\langle \eta_i \rangle - \kappa_i), \qquad i = 0, \ldots, 6 \tag{6}$$

蒙特卡罗计算采用系综平均值 $\langle \eta_i \rangle = 3 \times 10^{-4}$ 的允许方差作为收敛判据，并据此计算序参量的精度 $\mathrm{Var}(\langle \eta_i \rangle - \kappa_i)$。 $\langle \mu_i \rangle$ 的精度通过式（6）获得。有关蒙特卡罗模拟的其他详细信息见 §S1.2.3。

3.3. 用于自由能表示的可积深度神经网络（*IDNNs*）

通过使用导数数据进行训练，可获得自由能密度函数的 IDNN 表示：以化学势作为标签，以组分或序参量相应的对称不变量函数作为特征（§3.2.2） [2, 3]。IDNN 的可积性源于微积分基本定理，因为它是全连接深度神经网络（DNN）的导数。DNN 是一个函数 $Y(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{b})$，表示系综平均化学势 $\langle \mu_i \rangle$；其自变量或输入 $\boldsymbol{X}$ 表示组分和序参量的对称不变量函数 $\langle \eta_i \rangle$、 $i = 0$、.....6、权重 $\boldsymbol{W}$ 以及偏置 $\boldsymbol{b}$。在给定 $\widehat{\ }$ $\widehat{\ }$ $\widehat{\ }$ $\widehat{\ }$数据集 $\{(\boldsymbol{X}_\theta, Y_\theta)\}$的情况下，DNN 训练是一个关于权重和偏置的优化问题。然而，在此处，数据集 $\{(\boldsymbol{X}_\theta, Y_\theta)\}$ 不可用。相反，我们拥有导数数据集 $\{(\widehat{\boldsymbol{X}}_\theta, \widehat{y}_{\theta_k})\}$，其中 $\widehat{y}_{\theta_k}$ 对应于 $\widehat{Y}_\theta$ 关于 k^{th} 分量$\widehat{\ }$的偏导数， $\boldsymbol{X}_\theta$。将 IDNN 定义为 Y 关于其输入 $\boldsymbol{X}$ 的梯度，即 $\partial Y(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{b})/\partial X_k$，训练/优化问题为:

$$\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{b}} = \underset{\boldsymbol{W}, \boldsymbol{b}}{\arg\min} \sum_{k=1}^{n} \mathrm{MSE}\left(\frac{\partial Y(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{b})}{\partial X_k}\Big|_{\widehat{\boldsymbol{x}}_\theta}, \widehat{y}_{\theta_k}\right) \tag{7}$$

$\widehat{\ }$ $\widehat{\ }$ $\widehat{\ }$ $\widehat{\ }$优化后的权重 $\widehat{\boldsymbol{W}}$和偏置 $\widehat{\boldsymbol{b}}$与 IDNN 函数 $\partial Y(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{b})/\partial X_k$ 结合使用，以预测$\widehat{\ }$ $\widehat{\ }$化学势。相同的权重和偏置还用于其原函数 DNN 函数 $Y(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{b})$，以预测均匀自由能密度。

3.4. 采样与主动学习工作流

用于表示之字形有序化学势的 IDNN 训练需要在七维 $\boldsymbol{\eta}$ 空间中采样数据。在该空间中进行均匀且密集的采样将需要数量多得难以承受的蒙特卡罗模拟。因此，我们聚焦于具有重要物理意义且难以捕捉特征的区域。这些区域包括与之字形有序变体相关的能量井，以及序参量空间边界处化学势的发散行为，其中包括位于 $\eta_0 = x = \{0, 1\}$ 的组分端元。此外，还对序参量空间进行了一些一般性的、无引导的探索采样，以捕捉整体趋势。

我们通过将利用 采样与探索相结合，在逐点误差较高的区域改进部分训练的 IDNN。主动学习工作流会反复执行探索采样、IDNN 训练和利用采样，直至满足停止准则。该采样必须在 $\boldsymbol{\eta}$-空间的可容许域边界内进行。我们在该空间中使用 Billiard Walk [63] 随机采样算法。更多详情请参见 §S1.4。

该工作流程（图4a、b）在以下情况下会强制重新搜索 IDNN超参数：(a) 第二次工作流程迭代时；以及 (b) 使用最新数据集计算的前两个 IDNN 模型的均方误差（MSE）从一个模型到另一个模型有所增加时。如果 MSE 降低，则该工作流程允许使用之前的 IDNN 在最新数据上继续训练（图4b）。

3.5. 相场理论及相关计算框架

忽略弹性效应时，系统的总自由能为：

$$\Pi[x,\widehat{\boldsymbol{\eta}}]=\int_{\Omega}\left(f(x,\widehat{\boldsymbol{\eta}})+\frac{1}{2}\chi_0|\nabla x|^2+\sum_{i=1}^{6}\frac{1}{2}\chi_i|\nabla\widehat{\eta}_i|^2\right)\mathrm{d}V \tag{8}$$

其中，χ_i 为梯度参数，$f(x,\widehat{\boldsymbol{\eta}})$为自由能密度，在本研究中由解析积分的深度神经网络表示。相场方程中使用的化学势 $\widetilde{\mu}_i$ 是总自由能的变分导数：$\widetilde{\mu}_0:=\delta\Pi/\delta x$ 和 $\widetilde{\mu}_i:=\delta\Pi/\delta\widehat{\eta}_i$, $i=1,\dots,6$（为清晰起见，使用 $\widehat{\bullet}$ 记号；然而，$\widehat{\eta}_i=\eta_i$ 来自前文）：

$$\widetilde{\mu}_0=\frac{\partial f}{\partial x}-\chi_0\nabla^2 x,\qquad \widetilde{\mu}_i=\frac{\partial f}{\partial\widehat{\eta}_i}-\chi_i\nabla^2\widehat{\eta}_i,\ i=1,\dots,6 \tag{9}$$

Cahn-Hilliard相场方程 [64] 描述守恒组分的动力学，而Allen-Cahn方程 [65] 则描述非守恒序参量场：

$$\frac{\partial x}{\partial t}=\nabla\cdot(\widetilde{M}\nabla\widetilde{\mu}_0),\quad\text{and}\quad\frac{\partial\widehat{\eta}_i}{\partial t}=-L\widetilde{\mu}_i,\qquad i=1,\dots,6 \tag{10}$$

$\widetilde{\ }$其中，M 为迁移率，L 为动力学系数。我们代入化学势，并将控制方程写成弱形式，以便采用混合有限元方法求解：

$$0=\int_{\Omega}\left(w_x\frac{\partial x}{\partial t}+\widetilde{M}\nabla w_x\cdot\nabla\tilde{\mu}_0\right)\mathrm{d}V-\int_{\partial\Omega}w_x j_n\mathrm{d}S \tag{11}$$

$$0=\int_{\Omega}\left[w_{\tilde{\mu}_0}\left(\tilde{\mu}_0-\frac{\partial f}{\partial x}\right)-\chi_0\nabla w_{\tilde{\mu}_0}\cdot\nabla x\right]\mathrm{d}V \tag{12}$$

$$0=\int_{\Omega}\left[w_i\frac{\partial\widehat{\eta}_i}{\partial t}+L\left(w_i\frac{\partial f}{\partial\widehat{\eta}_i}+\chi_i\nabla w_i\cdot\nabla\widehat{\eta}_i\right)\right]\mathrm{d}V,\qquad i=1,\dots,6 \tag{13}$$

其中，w_x、$w_{\tilde{\mu}_0}$和 w_i 为权函数。对于以这种混合形式写出的方程，已对 x、$\tilde{\mu}_0$和 $\widehat{\eta}_i$、$i=1$、.....、n在 $\partial\Omega$上的边界条件进行如下设定，其中 $\boldsymbol{n}$ 为向外的$\widetilde{\ }$单位法向量，j_n 为流入通量：$\nabla x\cdot\boldsymbol{n}=0$，$M\nabla\tilde{\mu}_0\cdot\boldsymbol{n}=j_n$，$\nabla\widehat{\eta}_i\cdot\boldsymbol{n}=0$，$i=1$、.....、6

我们使用参考文献 [66]中的预测值，为300 K下的扩散率 D 定义一个组分依赖型代理函数：$D=0.01\exp(-274(1.05-x)(0.47-x)(1.-x))$，并按照§S1.5中概述的方法近似其他温度下的 D。扩散率代理函数和预测数据见图S5，其中有效振动频率 ν^* 据报道约为 10^{13} s^{-1} [66]。迁移率

$\sim$ $\sim M$ 通过方程 $M=Dx/(k_BT)$ [29]与 D 相关。我们通过约束相场计算得到的界面能与§3.1.1中的DFT结果一致，求解方程（8）中的梯度参数 χ_i 逆问题，详见§S1.5

模拟采用有限元法，并使用 mechanoChemFEM 代码（可在 github.com/mechanoChem/mechanoChemFEM 获取）；该代码基于 deal.II [67] 库。模拟中使用了自适应网格划分（含悬挂节点）和自适应时间步进。相场方法的更多细节见 §S1.5。

致谢

我们感谢 Brian Puchala 和 Anirudh Natarajan 就 CASM 软件及相关方法提供的见解，也感谢 Chirranjeevi Gopal 和 Muratahan Aykol 就层状氧化物的密度泛函理论方法提出的建议。我们衷心感谢丰田研究院提供的资助，资助编号为 #849910，项目名称为“用于电池材料数据驱动、预测性、多尺度和多物理场建模的计算框架”。本研究还部分获得了美国国家科学基金会 DMREF 项目资助，编号为 #1729166，项目名称为“合金-氧化物结构设计的集成框架”。此外，美国国防高级研究计划局（DARPA）通过协议编号 HR0011199002 提供了支持，项目名称为“用于发现复杂涌现材料现象的人工智能引导多尺度多物理场框架”。我们感谢美国陆军研究办公室通过 DURIP 资助项目 W911NF1810242 提供的支持，该项目为本研究提供了计算资源。本研究中的模拟在密歇根大学的 Great Lakes 高性能计算集群上运行。另有部分模拟使用了极端科学与工程发现环境（XSEDE）提供的计算资源，包括圣迭戈超级计算机中心的 Comet 和德克萨斯高级计算中心的 Stampede2，资源分配编号为 TG-DMR180072 和 TG-MCH200011。XSEDE 获得了美国国家科学基金会编号为 ACI-1548562 的资助。部分模拟使用了美国国家科学基金会通过编号为 1531752 的资助项目提供的资源，该项目为“Conflux 采购：一种用于数据驱动计算物理的新型平台”（技术监测员：Ed Walker）；密歇根大学还提供了额外支持。

作者贡献

该研究的尺度桥接方法由 KG 和 GHT 开发；电子结构研究由 GHT、SD 和 VG 规划；电子结构计算由 GHT 和 SD 完成；统计力学和相场研究由 GHT、JH 和 MFS 开展；所有作者共同分析了研究结果；论文由所有作者共同撰写。

利益冲突

作者声明不存在利益冲突。

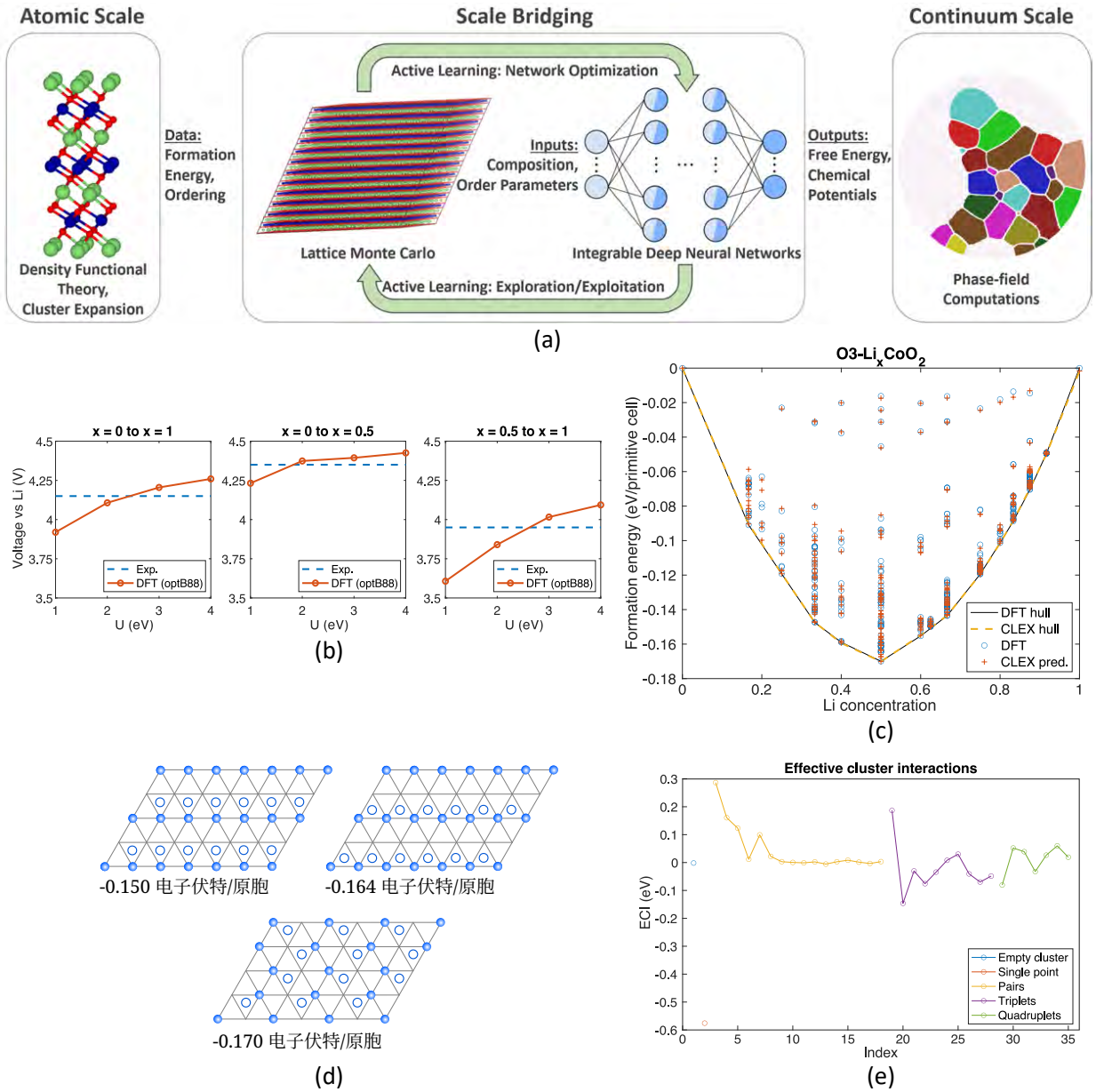


图1：（a）概述数据、计算方法以及机器学习支持的关联关系的流程图，展示了如何将原子尺度与连续体尺度连接起来。锂原子为绿色，Co为蓝色，O为红色（使用VESTA生成的可视化图像[19]）。（b）不同组分下的计算电压随U的变化，并与实验电压进行比较[20, 21]。（b）将簇展开（CLEX）预测的形成能和凸包与DFT计算得到的结果进行比较。簇展开预测位于凸包上的构型与DFT凸包上的构型相同。（d）三个锂浓度为 $x = 0.5$ 的构型及其对应的计算形成能。实心圆表示锂层内的锂原子，空心圆表示相邻锂层中的锂原子。（e）簇展开中与各基函数对应的有效簇相互作用值。

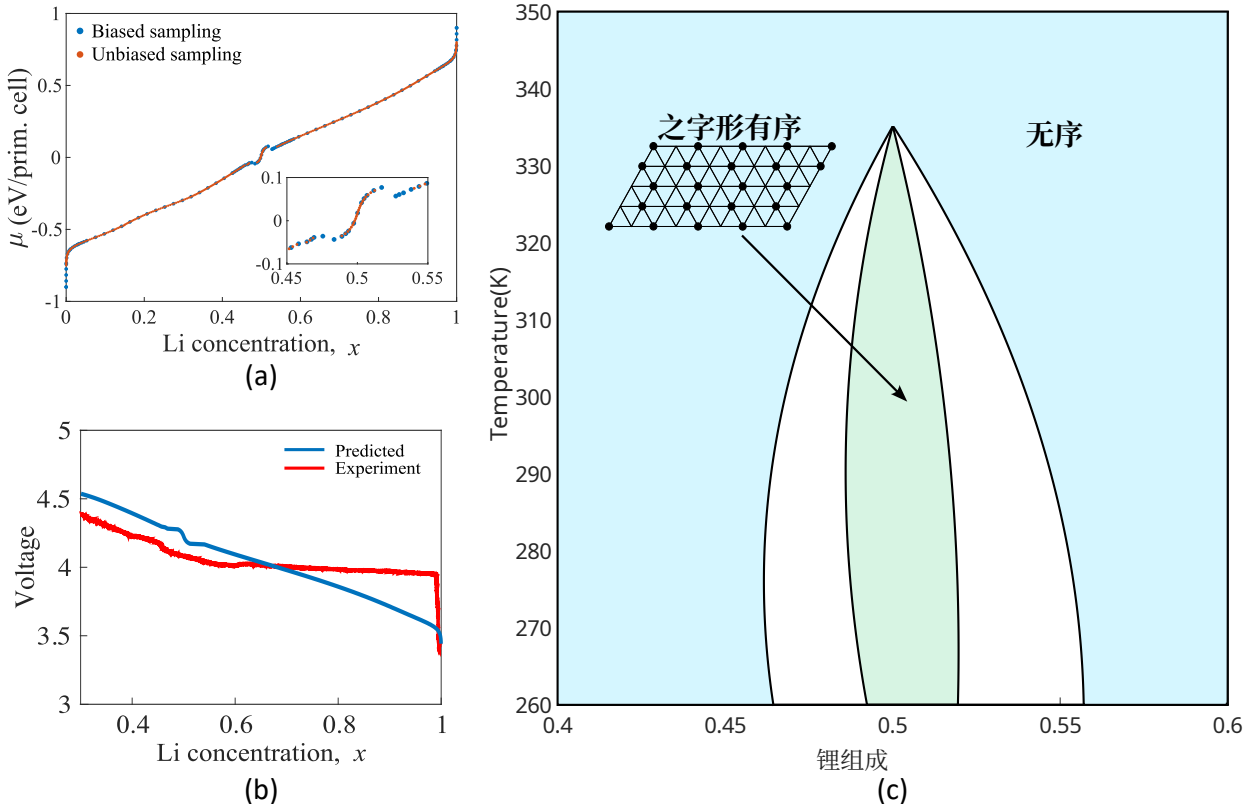


图 2：（a）使用有偏置势和无偏置势的蒙特卡罗模拟预测的 300 K 下的化学势值。（b）将 300 K 下预测的电压（蓝色）与文献 [23] 中的实验电压（红色）进行比较。（c）基于蒙特卡罗结果，展示 LCO 的 O3结构在 $x = 1/2$ 下具有所关注有序排列的相图。（a）中无偏置样本点的间隔对应于 300 K 下的有序-无序相变（白色区域）。

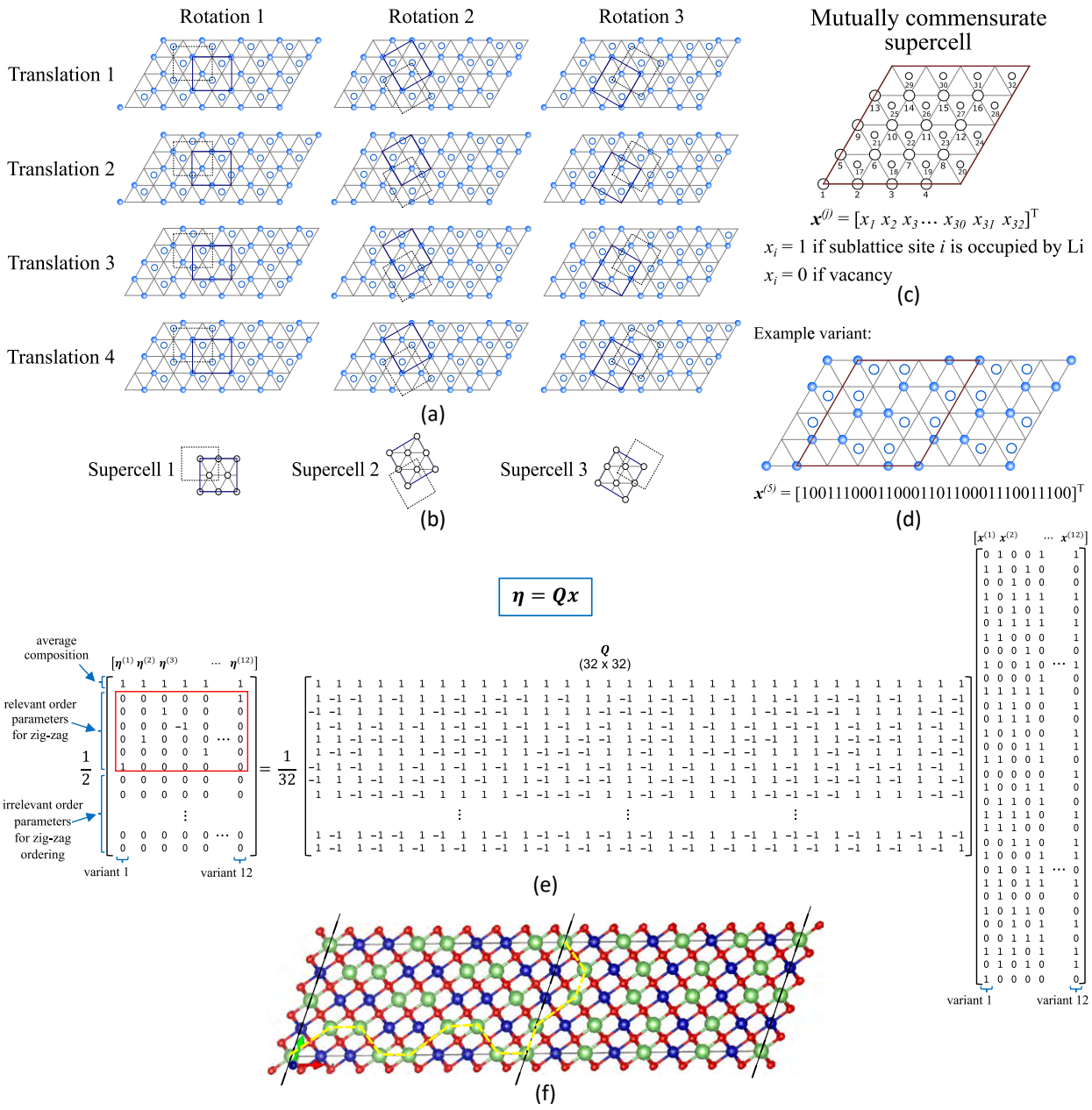


图3：（a）之字形有序有12种变体，由3种旋转和平移4次的组合产生。（b）还展示了与每种变体对应的超胞。请注意，这些超胞在平移下可以保持不变，但在旋转下不能保持不变。实心圆表示占据同一层的锂原子；空心圆表示位于下层（或上层）的锂原子。（c）还展示了包含所有超胞的互相公度超胞及其32个子晶格位点。（d）展示了12种之字形有序变体中的一个示例及其子晶格表示。（e）从子晶格到序参量空间的变换 $\boldsymbol{\eta} = \boldsymbol{Q}\boldsymbol{x}$ 。 $\boldsymbol{\eta}$ 空间中的12种之字形变体在第8—32行的值为零。因此，需要七个序参量来描述组分和有序性。（f）用于计算有序变体 $\eta_1 = 1/2$ （左）与 $\eta_3 = 1/2$ （右）之间反相边界（虚线）能量的原子构型。黄色路径描绘了之字形有序的变化。

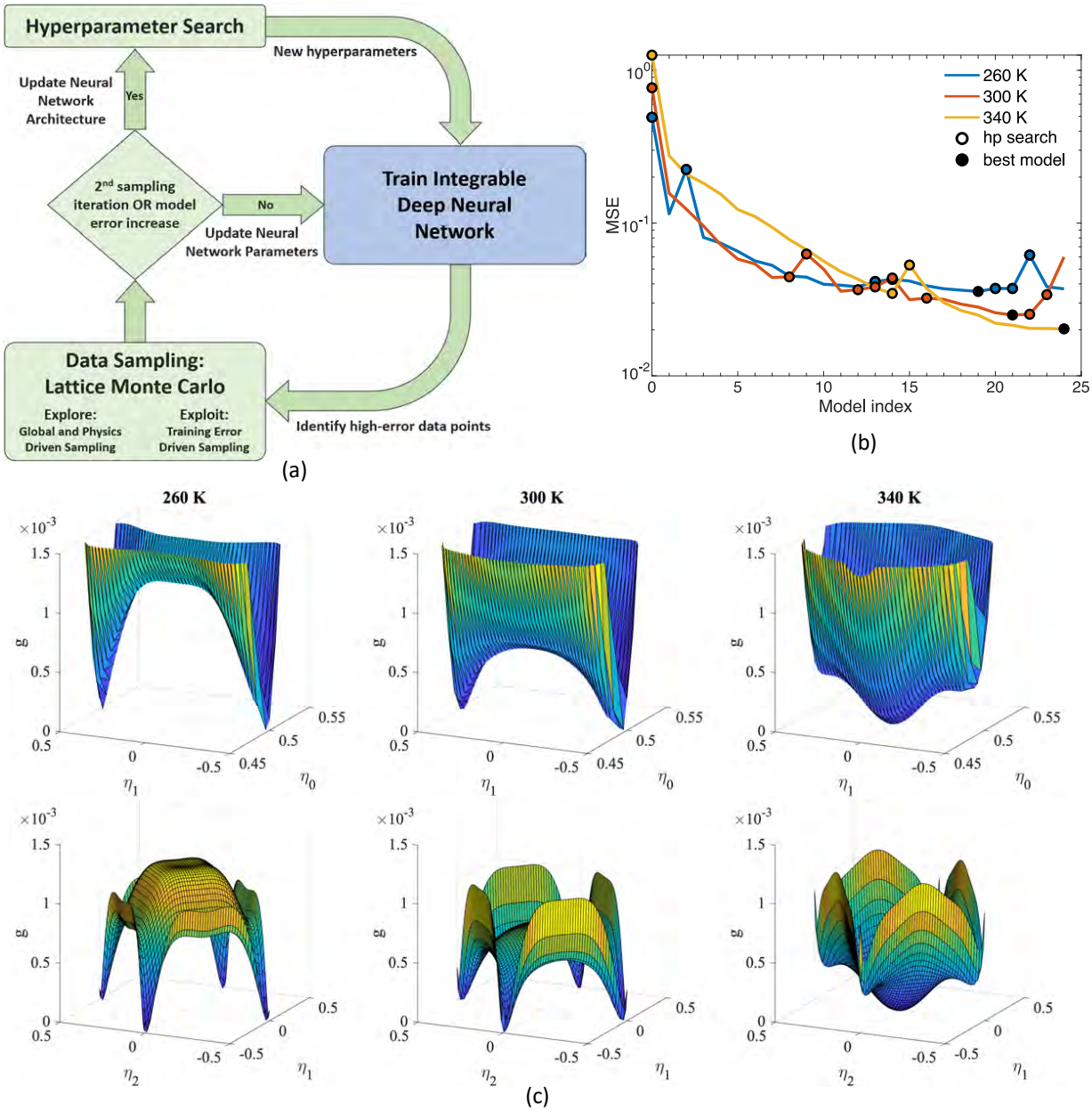
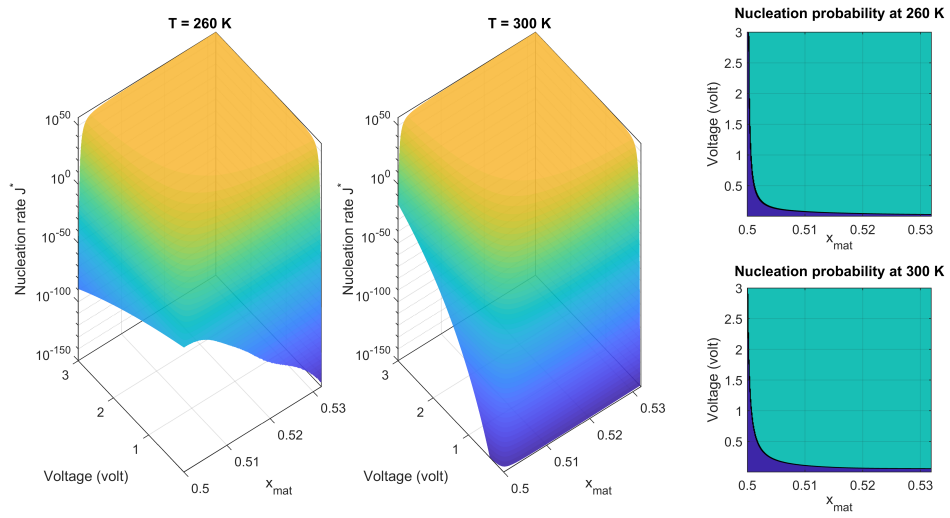
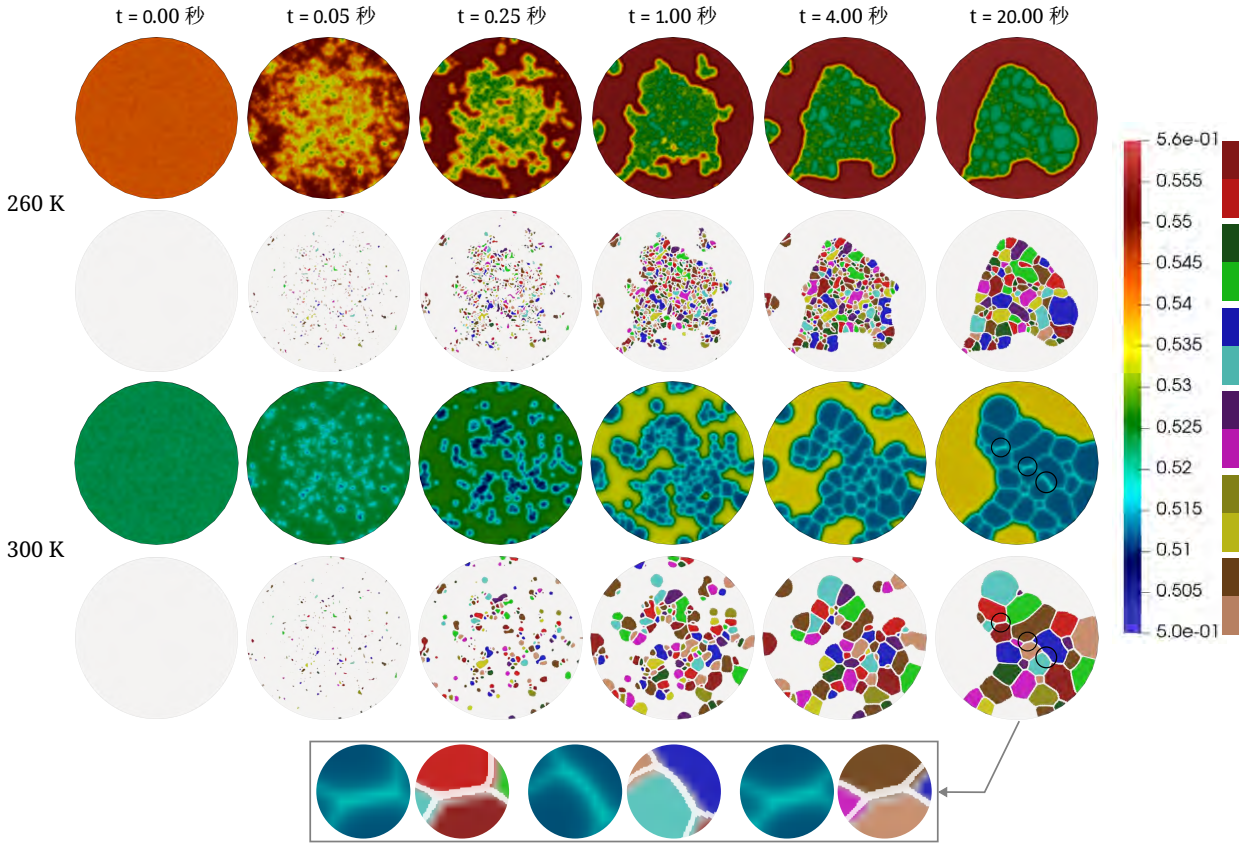


图4：（a）主动学习工作流示意图，展示了如何通过探索、训练和利用来指导额外采样。（b）IDNN 在 260、300 和 340 K 下相对于模型索引的主动学习曲线。每个模型对应一个规模逐步增大的数据集。请注意 y 轴采用对数刻度。空心圆表示为更新 IDNN 而进行的新一轮超参数搜索。损失为最终（25th）数据集的均方误差。实心圆标出各温度下表现最佳的 IDNN 模型。（c）以 $\eta_0 - \eta_1$ 切片（$\eta_2, \ldots, \eta_6 = 0$）和 $\eta_1 - \eta_2$ 切片（$\eta_0 = 1/2, \eta_3, \ldots \eta_6 = 0$）绘制的自由能密度曲面。有序变体对应于位于 $\eta_0 = x \sim 1/2$ 和 $\eta_1, \ldots, \eta_6 \sim \pm 1/2$ 处的势阱，分别对应 260 和 300 K。在 340 K 下，对于 η_0 的所有取值，唯一的势阱位于 $\eta_1, \ldots, \eta_6 = 0$；也就是说，仅存在无序形式。



（a）成核速率和成核概率



（b）有序区域生长的相场模拟

图 5：（a）260 K 和 300 K 下，无序区域中潜在成核位点的电压和组分的函数关系所对应的成核速率 x_{mat}，以及在 $\Delta t = 1\mathrm{e}-4$ sec 内的成核概率 P_n。深色和浅色分别表示 $P_n = 0$ 和 $P_n = 1$。（b）260 K（300 K）下的二维相场模拟，显示直径为 1 μm 的颗粒中的锂组成。初始锂组成围绕 $x = 0.5425$（$x = 0.5225$）进行随机扰动，且不存在边界通量。插图显示了在 300 K 下形成的部分最宽有序-有序界面，分别由 $\eta_1 = \pm\frac{1}{2}$（左）、 $\eta_3 = \pm\frac{1}{2}$（中）和 $\eta_6 = \pm\frac{1}{2}$（右）形成。

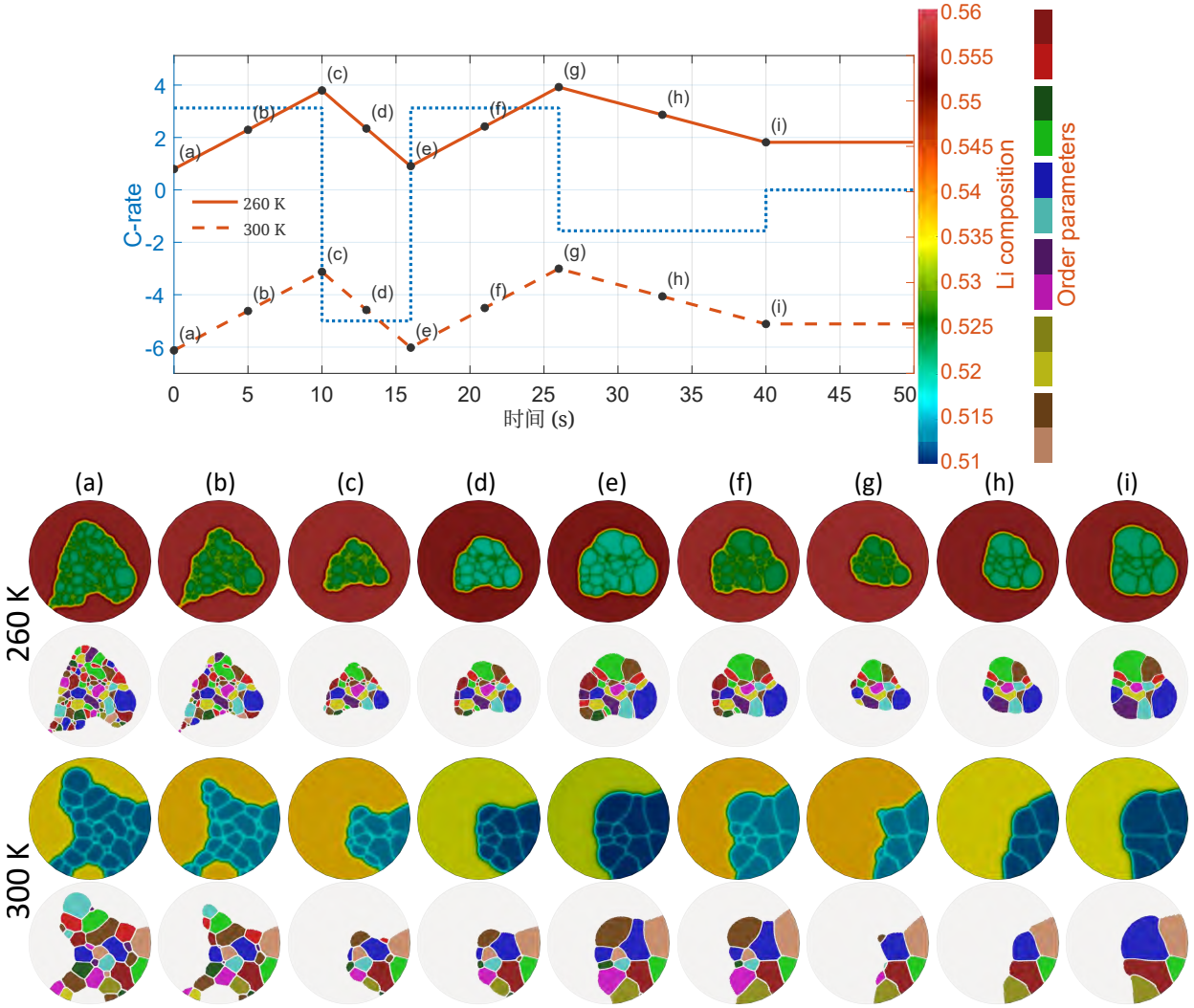


图 6：施加循环电流密度后，在 260 K 和 300 K 下直径为 1 μm 的颗粒所产生的锂组成场相场结果。蓝色虚线表示施加的 C倍率，正值表示放电，负值表示充电；红线表示颗粒对应的平均锂组成。

参考文献

[1] A. Van der Ven、M. K.. Aydinol、G. Ceder、G. Kresse 和 J. Hafner。Li_xCoO_2中相稳定性的第一性原理研究。《Physical Review B》，58：2975–2987，1998年。[2] G.H. Teichert、A.R. Natarajan、A. Van der Ven 和 K. Garikipati。机器学习材料物理学：可积深度神经网络通过学习自由能函数实现尺度桥接。《Computer Methods in Applied Mechanics and Engineering》，353：201–216，2019年。[3] G.H. Teichert、A.R. Natarajan、A. Van der Ven 和 K. Garikipati。材料物理学中的尺度桥接：用于合金自由能函数表示的主动学习工作流和可积深度神经网络。《Computer Methods in Applied Mechanics and Engineering》，371：113281，2020年。[4] J. N. Reimers 和 J. R. Dahn。锂在 Li_xCoO_2中的电化学和原位X射线衍射嵌入研究。《电化学学会志》，139(8)：2091–2097，1992年8月。[5] M. M´en´etrier、I. Saadoune、S. Levasseur 和 C. Delmas。$LiCoO_2$脱锂过程中的绝缘体-金属转变：电子性质和7Li核磁共振研究。《Journal of Materials Chemistry》，9(5)：1135–1140，1999年。被引用337次。[6] Y. Shao-Horn、S. Levasseur、F. Weill 和 C. Delmas。探究O3层状Li_xCoO_2 (x≈0.5)中的锂和空位有序化。《电化学学会志》，150(3)：A366，2003年。[7] Y Takahashi、N Kijima、K Tokiwa、T Watanabe 和 J Akimoto。$Li_{0.5}CoO_2$的单晶合成、结构精修和电学性质。《凝聚态物理学杂志》，19(43)：436202，2007年9月。[8] T. Motohashi、T. Ono、Y. Sugimoto、Y. Masubuchi、S. Kikkawa、R. Kanno、M. Karppinen 和 H. Ya- mauchi。层状钴氧化物体系Li_xCoO_2 ($0.0 \leq x \leq 1.0$)的电子相图。《Physical Review B》，80：165114，2009年10月。[9] Z. Chen 和 J.R. Dahn。获得$LiCoO_2$ 循环至4.5 v时优异容量保持率的方法。《Electrochimica Acta》，49(7)：1079–1090，2004年。[10] K. Chang、B. Hallstedt、D. Music、J. Fischer、C. Ziebert、Ulrich S 和 H. J. Seifert。层状O3和O2结构$LiCoO_2$–CoO_2 假二元体系的热力学描述。《Calphad》，41：6–15，2013年。[11] C. Wolverton 和 A. Zunger。Li_xCoO_2中空位有序-无序转变和嵌入电池电压的第一性原理预测。《Physical Review Letters》，81：606–609，1998年7月。[12] L. Kaufman 和 H. Bernstein。《相图的计算机计算》。Academic Press，纽约，1970年。[13] T. Abe 和 T. Koyama。$LiCoO_2$–CoO_2 假二元体系的热力学建模。《Calphad》，35(2)：209–218，2011年。[14] N. Nadkarni、T. Zhou、D. Fraggedakis、T. Gao 和 M. Z. Bazant。Li xCoO_2 中用于能量与信息存储的金属-绝缘体相变建模。《Advanced Functional Materials》，29(40)：1902821，2019年。[15] M. Cococcioni 和 S. de Gironcoli。采用LDA + U方法计算有效相互作用参数的线性响应方法。《Physical Review B》，71：035105，2005年1月。[16] M. Aykol、S. Kim 和 C. Wolverton。层状钴酸锂中的范德华相互作用。《The Journal of Physical Chemistry C》，119(33)：19053–19058，2015年。

[17] C. B. Gopal 和 A. van de Walle。氧化铈中本征氧空位的从头算热力学。《Physical Review B》，86：134117，2012年10月。[18] ME Arroyo y de Dompablo、A Van der Ven 和 G Ceder。Li x NiO 2 中锂有序化和相稳定性的第一性原理计算。《Physical Review B》，66(6)：064112，2002年。[19] K. Momma 和 F. Izumi。VESTA 3：用于晶体、体积和形貌数据的三维可视化。《J. Appl. Crystallogr.》，44：1272–1276，2011年。[20] GG Amatucci、JM Tarascon 和 LC Klein。Coo2，li x coo2固溶体的端元。《Journal of The Electrochemical Society》，143(3)：1114，1996年。[21] H Xia、L Lu、Y Sh Meng 和 G Ceder。采用脉冲激光沉积生长的LiCoO2薄膜中的相变和高电压电化学行为。《Journal of The Electrochemical Society》，154(4)：A337，2007年。[22] https://github.com/prisms-center/CASMcode。CASM：团簇统计力学方法，v0.3.dev，2018年。[23] G. G. Amatucci、J. M. Tarascon 和 L. C. Klein。CoO_2, , Li_xCoO_2 固溶体的端元。《Journal of The Electrochemical Society》，143(3)：1114–1123，1996年3月。[24] G.M. Torrie 和 J.P. Valleau。用于蒙特卡罗自由能估计的非物理采样分布：伞形采样。《Journal of Computational Physics》，23(2)：187–199，1977年。[25] B. Sadigh、P. Erhart、A. Stukowski、A. Caro、E. Martinez 和 L. Zepeda-Ruiz。用于合金析出原子尺度模拟的可扩展并行蒙特卡罗算法。《Physical Review B》，85：184203，2012年5月。[26] B. Sadigh 和 P. Erhart。通过跨越相界面的直接热力学积分计算析出物的过剩自由能。《Physical Review B》，86：134204，2012年10月。[27] Y. Wang、D. Banerjee、C.C. Su 和 A.G. Khachaturyan。面心立方固溶体中L12有序金属间化合物析出的场动力学模型与计算机模拟。《Acta Materialia》，46(9)：2983–3001，1998年。[28] M. S. Dresselhaus、G. Dresselhaus 和 A. Jorio。群论：在凝聚态物理学中的应用。Springer Science & Business Media，2007年。[29] T. Jiang、S. Rudraraju、A. Roy、A. Van der Ven、K. Garikipati 和 M. L. Falk。$Li_{1+x}Ti_2O_4$ 电极颗粒中嵌锂诱导应力的多物理场模拟。《The Journal of Physical Chemistry C》，120(49)：27871–27881，2016年。[30] I. Timrov、N. Marzari 和 M. Cococcioni。基于密度泛函微扰理论、超软和投影缀加波形式的自洽Hubbard参数。《Physical Review B》，103：045141，2021年1月。[31] S. H. Choi、J-W. Son、Y. S. Yoon 和 J. Kim。锂电池中$LiCoO_2$ 温度依赖性能的颗粒尺寸效应。《Journal of power sources》，158(2)：1419–1424，2006年。[32] F. Leng、C. M. Tan 和 M. Pecht。温度对高于室温运行的锂离子电池老化速率的影响。《Scientific reports》，5(1)：1–12，2015年。[33] J. P. Pender、G. Jha、D. H. Youn、J. M Ziegler、I. Andoni、E. J. Choi、A. Heller、B. S. Dunn、P. S. Weiss、R. M. Penner 等。锂离子电池中的电极退化。《ACS nano》，14(2)：1243–1295，2020年。[34] FK LeGoues 和 HI Aaronson。晶体学对均匀面心立方-面心立方成核的临界晶核形状和动力学的影响——IV。CuCo合金中理论与实验的比较。《冶金学报》，32(10)：1855–1864，1984年。

[35] J.P Simmons、C Shen 和 Y Wang。有序与无序转变中通过显式纳入成核事件对同时成核与生长进行相场建模。《材料学报》，43(10):935–942，2000年。[36] J.P. Simmons、Youhai Wen、C. Shen 和 Y.Z. Wang。使用相场法模拟涉及成核与长大现象的微观结构演化。《材料科学与工程：*A*辑》，365(1):136–143，2004年。多尺度材料建模。[37] R. P. Sear。成核：蛋白质溶液和胶体悬浮液的理论与应用。《凝聚态物理学杂志》，19(3):033101，2007年1月。[38] H. 王、Y-I. Jang、B. 黄、D. R. Sadoway 和 Y-M. Chiang。可充电锂电池中 $LiCoO_2$ 正极电化学循环诱导损伤和无序的透射电子显微镜研究。《电化学学会志》，146(2):473，1999年。[39] Gregory H Teichert、Sambit Das、Muratahan Aykol、Chirranjeevi Gopal、Vikram Gavini 和 Krishna Garikipati。通过机器学习驱动的电子结构、统计力学与相场理论之间的尺度桥接研究 Li_xCoO_2 相稳定性。*arXiv* 预印本 *arXiv:2104.08318*，2021年。[40] B. Orvananos、T. R. Ferguson、H-C. Yu、M. Z. Bazant 和 K. Thornton。纳米颗粒相分离锂离子电池电极充放电行为的颗粒尺度建模。《电化学学会志》，161(4):A535，2014年。[41] A. J. Merryweather、C. Schnedermann、Q. Jacquet、C. P. Grey 和 A. Rao。电池中单颗粒离子动力学的操作状态光学跟踪。《自然》，594(7864):522–528，2021年。[42] T. Thonhauser、V. R. Cooper、S. Li、A. Puzder、P. Hyldgaard 和 D. C. Langreth。范德华密度泛函：自洽势与范德华键的本质。*Physical Review B*，76:125112，2007年9月。[43] J. Klime˘s、D. R. Bowler 和 A. Michaelides。范德华密度泛函的化学精度。《凝聚态物理学杂志》，22(2):022201，2009年12月。[44] D C Langreth、B I Lundqvist、S D Chakarova-K¨ack、V R Cooper、M Dion、P Hyldgaard、A Kelkkanen、J Kleis、Lingzhu Kong、Shen Li、P G Moses、E Murray、A Puzder、H Rydberg、E Schr¨oder 和 T Thon-hauser。稀疏物质的密度泛函。《凝聚态物理学杂志》，21(8):084203，2009年1月。[45] P. Giannozzi、S. Baroni、N. Bonini、M. Calandra、R. Car、C. Cavazzoni、D. Ceresoli、G. L. Chiarotti、M. Cococcioni、I. Dabo、A. Dal Corso、S. de Gironcoli、S. Fabris、G. Fratesi、R. Gebauer、U. Gerstmann、C. Gougoussis、A. Kokalj、M. Lazzeri、L. Martin-Samos、N. Marzari、F. Mauri、R. Mazzarello、S. Paolini、A. Pasquarello、L. Paulatto、C. Sbraccia、S. Scandolo、G. Sclauzero、A. P. Seitsonen、A. Smogunov、P. Umari 和 R. M. Wentzcovitch。Quantum Espresso：用于材料量子模拟的模块化开源软件项目。《凝聚态物理学杂志》，21(39):395502（19页），2009年。[46] P Giannozzi、O Andreussi、T Brumme、O Bunau、M Buongiorno Nardelli、M Calandra、R Car、C Cavazzoni、D Ceresoli、M Cococcioni、N Colonna、I Carnimeo、A Dal Corso、S de Gironcoli、P Delugas、R. A. DiStasio Jr.、A. Ferretti、A. Floris、G. Fratesi、G. Fugallo、R. Gebauer、U. Gerstmann、F. Giustino、T. Gorni、J. Jia、M. Kawamura、H-Y. Ko、A. Kokalj、E. K¨u¸c¨ukbenli、M. Lazzeri、M. Marsili、N. Marzari、F. Mauri、N. L. Nguyen、H-V. Nguyen、A. Otero de-la Roza、L. Paulatto、S. Ponc´e、D. Rocca、R. Sabatini、B. Santra、M. Schlipf、A. P. Seitsonen、A. Smogunov、I. Timrov、T. Thonhauser、P. Umari、N. Vast、X. Wu 和 S Baroni。Quantum Espresso 的高级功能：用于材料建模的量子模拟。《凝聚态物理学杂志》，29(46):465901，2017年。[47] M. K. Aydinol、A. F. Kohan、G. Ceder、K. Cho 和 J. Joannopoulos。金属氧化物和金属二硫属化物中锂嵌入的第一性原理研究。*Physical Review B*，56:1354–1365，1997年7月。

[48] B. Meredig、A. Thompson、H. A. Hansen、C. Wolverton 和 A. van de Walle。在密度泛函理论 $+u$中定位低能解的方法。*Physical Review B*，82:195128，2010年11月。[49] S. Das、P. Motamarri、V. Subramanian、D. M. Rogers 和 V. Gavini。DFT-FE 1.0：采用有限元离散的大规模并行混合 CPU-GPU 密度泛函理论代码。《*Computer Physics Communications*》，280:108473，2022年。[50] P. Motamarri、S. Das、S. Rudraraju、K. Ghosh、D. Davydov 和 V. Gavini。DFT-FE——用于大规模密度泛函理论计算的大规模并行自适应有限元代码。《*Computer Physics Communications*》，246:106853，2020年。[51] P. Motamarri 和 V. Gavini。使用 Kohn-Sham 密度泛函理论进行电子结构计算中的构型力。*Physical Review B*，97:165132，2018年4月。[52] P. Motamarri、M.R. Nowak、K. Leiter、J. Knap 和 V. Gavini。用于 Kohn-Sham 密度泛函理论的高阶自适应有限元方法。《*Journal of Computational Physics*》，253:308–343，2013年。[53] D. R. Hamann。优化的范德堡范数守恒赝势。*Physical Review B*，88:239906，1995年。[54] M. J. van Setten、M. Giantomassi、E. Bousquet、M. J. Verstraete、D. R. Hamann、X. Gonze 和 G-M. Rignanese。the pseudodojo：85种元素优化守范赝势表的训练与分级。《*Computer Physics Communications*》，226:39–54，2018年。[55] J. M. Sanchez、F. Ducastelle 和 D. Gratias。多组分系统的广义团簇描述。$Physica\ A$，128:334–350，1984年。[56] D. De Fontaine。合金有序-无序转变的团簇方法。第47卷，《*Solid State Physics*》，第33–176页。Academic Press，1994年。[57] A. Van der Ven、J. C. Thomas、Q. Xu 和 J. Bhattacharya。将固体电子结构与其热力学和动力学性质联系起来。《*Mathematics and Computers in Simulation*》，80(7):1393–1410，2010年。[58] J. C. Thomas 和 A. Van der Ven。基于第一性原理研究强非谐性及力学不稳定晶相的有限温度性质。《*Physical Review B*》，88(21):214111–214111，2013年12月。[59] B. Puchala 和 A. Van der Ven。基于第一性原理计算的 Zr-O 系统热力学。*Physical Review B*，88:094108，2013年9月。[60] A. R. Natarajan、J. C. Thomas、B. Puchala 和 A. Van der Ven。经历有序-无序相变固体的对称性适配序参量与自由能。*Physical Review B*，96:134204，2017年10月。[61] J. C. Thomas 和 A. Van der Ven。非线性弹性的探索及其对晶体材料的高效参数化。《*Journal of the Mechanics and Physics of Solids*》，107:76–95，2017年10月。[62] Y. Mishin。Ni–Al体系中 γ 相和 γ'相的原子尺度建模。《*Acta Materialia*》，52(6):1451–1467，2004年。[63] B.T. Polyak 和 E.N. Gryazina。Billiard Walk——一种用于控制与优化的新型采样算法。《*IFAC Proceedings Volumes*》，47(3):6123–6128，2014年。第19届 IFAC 世界大会。

[64] J. W. Cahn 和 J. E. Hilliard。非均匀系统的自由能。I 界面能。《化学物理学杂志》，28:258–267，1958 年。[65] S. M. Allen 和 J. W. Cahn。反相畴界运动的微观理论及其在反相畴界粗化中的应用。《冶金学报》，27:1085–1091，1979 年。[66] A. Van der Ven 和 G Ceder。层状 Li_xCoO_2 中的锂扩散。《电化学与固态快报》，3(7):301，2000 年。[67] D. Arndt、W. Bangerth、D. Davydov、T. Heister、L. Heltai、M. Kronbichler、M. Maier、J-P. Pelteret、B. Turcksin 和 D. Wells。deal.II 库，8.5 版。《数值数学杂志》，25(3)，2017 年 4 月。

arXiv:2302.08991v1 [cs.CE] 17 Feb 2023

补充信息：机器学习驱动的尺度桥接——在第一性原理统计力学与连续相场计算之间，用于研究 Li_xCoO_2 的相稳定性

G.H. Teichert[a]、S. Das[a]、M. Faghih Shojaei[a]、J. Holber[f]、T. Mueller[g、b]、L. Hung[b]、V. Gavini[a、c、f、e]、K. Garikipati[a、d、f、e、*]

[a]*Department of Mechanical Engineering、 University of Michigan*[b]*Toyota Research Institute、 Los Altos、 CA*[c]*Department of Materials Science & Engineering、 University of Michigan*
[d]*Department of Mathematics、 University of Michigan*
[e]*Michigan Institute for Computational Discovery & Engineering、 University of Michigan*
[f]*Applied Physics Program、 University of Michigan*[g]*Materials Science & Engineering、 Johns Hopkins University*

S1. 详细方法

S1.1. 密度泛函理论方法

组分为 Li E 的构型的形成能 x 可根据使用密度泛函理论（DFT）计算得到的总能量值，按照以下方程 [1] 计算：

$$E_{\mathrm{Li}_x\mathrm{CoO}_2} = E^{\mathrm{tot}} - xE^{\mathrm{tot}}_{\mathrm{LiCoO}_2} - (1-x)E^{\mathrm{tot}}_{\mathrm{CoO}_2} \tag{S1}$$

其中， E^{tot}、 $E^{\mathrm{tot}}_{\mathrm{LiCoO}_2}$和 $E^{\mathrm{tot}}_{\mathrm{CoO}_2}$ 分别为给定构型、$LiCoO_2$, 和 CoO_2,的总能量。在对 Li_xCoO_2，进行DFT计算时，Aykol 及其合作者发现，不仅需要加入 Hubbard修正（DFT+U），以减少自相互作用误差 [2, 3, 4]），还需要加入 van der Waals（vdW）相互作用 [5]。加入 vdW 相互作用后，在使用经过适当调节的 U值时，DFT 预测的电压能够与实验电压相匹配。尽管 vdW修正和 vdW密度泛函（vdW-DF）都能改善 DFT 结果，但使用 vdW-DF 得到的预测结果最为一致且准确。

在此工作流程中，我们采用简化的DFT旋转不变公式+U [6]和一种vdW-DF交换-相关泛函，即optB88交换-相关泛函 [7, 8, 9, 10, 5, 11, 12],，计算选定LCO构型子集的形成能。CASM（簇展开统计力学方法）软件套件用于确定一组具有O3晶体结构的构型，以便参数化簇展开。我们对其中333个构型完成了完全弛豫的DFT计算，CASM根据原子构型自动调整k点网格的大小。我们在Quantum Espresso中 [13, 14], 使用采用Perdew-Burke-Ernzerhof（PBE）泛函计算的PAW赝势，进行带几何优化的基态DFT计算；这些赝势来自PSlibrary 1.0.0 [15, 16]。波函数截断能和电荷密度截断能的取值分两步确定。首先，以Co赝势文件中建议的截断能值为起点，仅提高电荷密度截断能，直到总能量收敛至 $<$1电子伏特/原子。随后，在保持两者比值不变的情况下，同时提高波函数截断能和电荷密度截断能，直到总能量再次收敛至 <1 电子伏特/原子，最终波函数截断能为55 Ha，电荷密度截断能为301.5 Ha。此外还采用 $6\times6\times3$ 的k点网格，以确保总能量在 <1 电子伏特/原子以内收敛。结构优化持续进行，直到晶胞应力和

*通讯作者

离子力分别低于0.5 kbar和0.00005 Ha/Bohr。O3-$LiCoO_2$ 的晶体结构如图1a所示。

在本研究中，我们校准 Hubbard U 参数，使其在各种锂组成范围内与实验平均嵌锂电压相匹配。为此，对于给定的 x_1 和 x_2 [17]组分值，我们使用先前由实验报道的晶体结构，计算逐渐增加 U 时的电压。根据以下方程 [18, 5]计算从 x_1 到 x_2 的平均电压：

$$V = -\frac{E_{\mathrm{Li}_{x_2}\mathrm{CoO}_2} - E_{\mathrm{Li}_{x_1}\mathrm{CoO}_2} - (x_2 - x_1)E_{\mathrm{Li}}}{(x_2 - x_1)e} \tag{S2}$$

其中， E 是根据密度泛函理论计算得到的总能量， e 是电子所带的电荷。我们计算了以下锂组成的平均电压：$\{x_1, x_2\} = \{0, 1\}$， $\{0, 1/2\}$ 和 $\{1/2, 1\}$。通过与实验电压（图1b）进行比较，我们选取了 $U = 2.5$ 电子伏特这一适当数值。

通过密度泛函理论计算中的磁矩，监测了电荷有序的发展 [5]。在本研究的密度泛函理论结果中，图1d左侧的行有序显示出电荷分裂的证据，但另一种行有序和之字形有序中没有观察到。实验上已观察到 Co^{3+}和Co^{4+} 原子的电荷有序，其出现于低温以及锂组分为 $x = 1/2$ 和$x = 2/3$ [19]时。

S1.1.1. 界面能的 *DFT*计算 $nergy$

我们对 $\eta_1 = 1/2$ 与 $\eta_3 = 1/2$ 有序 LCO旋转变体之间反相边界的共格界面能进行了大规模 DFT计算。这些计算采用 DFT-FE 软件 [20, 21, 22, 23], ——一种新近开发的、用于基于有限元离散开展大规模实空间 Kohn-Sham 密度泛函理论研究的大规模并行开源代码。我们采用 PBE 交换 - 相关泛函和优化范德堡范数守恒赝势（ONCV） [24] ，其来自 Pseudo Dojo 库 [25]。DFT-FE 中的所有数值参数均经过选择，使基态能量的收敛精度达到 1.5 电子伏特/原子。此外，离子力和晶胞应力分别松弛至低于 5 meV/ $^{\circ}$A 和 0.5 Kbar，并且所有模拟均采用温度为 500 K 的费米-狄拉克展宽。

为了计算界面能，我们考虑了两类模拟晶胞。第一类是表示界面系统的周期性模拟晶胞，其中为每个有序变体构建了数量相等且彼此相容（MC）的超胞（见§S1.2.2）；考虑到模拟区域的周期性，这样会形成两个界面。如图3f所示，在包含两个MC超胞的情况下，MC超胞沿着不属于界面所在平面的晶格矢量排列。此外，我们还考虑了另一种用于表示体相的周期性模拟晶胞，该晶胞由对应于任一有序变体的一个MC超胞构成。由于相应的超胞可通过旋转相互关联，因此两个有序变体的体相能量相等。随后，我们对界面系统和体相系统进行了基态DFT计算，并对离子力和晶胞应力进行完全结构弛豫。离子力和晶胞应力以自洽方式弛豫，其中每次晶胞应力弛豫更新都包括在固定晶胞矢量的条件下对离子力进行完全弛豫。单位界面面积的总界面能（定义为两个界面能的平均值，并记为 γ'）由下列能量差给出

$$\gamma' = \frac{E_{(\mathrm{Var1+Var2})} - 2N_{\mathrm{cell}}E_{\mathrm{bulk}}}{2A}\,, \tag{S3}$$

其中， N_{cell} 表示每种变体对应的 MC 超胞数量， A 表示结构弛豫后界面的面积，分母中的因子 2 则用于对周期性体系中的两个界面取平均。界面能 γ' 由两部分组成。第一部分是短程项，源于界面处原子排列的局部偏离，记为 γ。第二部分是长程贡献，源于维持界面相干性所需的弹性失配应变。尽管失配弹性应变会随远离界面而衰减，但通过周期性超胞进行显式 DFT计算来解析这种应变，在计算上并不可行

这种应变。由于界面计算采用了周期性边界条件，失配弹性应变会对 γ' 贡献一个随 MC 超胞数量线性增长的能量：

$$\gamma' = \gamma + (k_1 + k_2) \times N_{\text{cell}}\,, \tag{S4}$$

其中， k_1 和 k_2 表示两个有序变体每个 MC 超胞的失配弹性能。接下来，通过求方程 (S4) 在 $N_{\text{cell}} = 0$ 处的截距，消除对 γ' 呈线性标度的弹性失配应变能贡献，从而得到 γ。此前的原子尺度计算也采用了类似方法来计算共格界面能 [26, 27]。表 S1 展示了用于确定 LCO体系中 γ 的能量和体系尺寸；其中，我们考虑 N_{cell} 在 1–3 之间变化，并得到 γ' 与 N_{cell}之间预期的近似线性关系。最后，估算得到 γ 为 30.9 mJ/m^2。此外需要指出的是，上述计算要求模拟域包含 $\sim$ 500–1400 个原子（4,000–12,000 个电子），并且每种模拟域尺寸约进行 200 次几何更新，这表明这些计算具有大规模特征。

S1.2. 统计力学

在本研究中，我们遵循 Van der Ven 等人 [1] 在对 Li$_x$CoO$_2$ (LCO)开展第一性原理研究时所提出的统计力学方法，以及 Natarajan 等人 [28], 等人的方法。

S1.2.1. 形成能的簇展开

鉴于 DFT计算的高昂成本，我们采用簇展开，以处理统计力学研究所需的大量构型。形成能表示为 $E_{\text{f}}(\boldsymbol{\sigma})$，其中$\boldsymbol{\sigma}$ 是构型向量；当 Li 占据位点 i 时，占位变量 $\sigma_i = 1$ 为 $\sigma_i = 1$ ，而当该位点包含空位时为 $\sigma_i = 0$ 。按照惯例，我们通过使用 DFT 计算一部分构型的形成能来定义 $E_{\text{f}}(\boldsymbol{\sigma})$，并利用这些数值对簇展开 [29, 30] 进行参数化，将其作为可快速查询的形成能替代模型。我们使用 CASM 软件[31], ，该软件有助于构建簇展开哈密顿量并对其进行参数化，以及将其用于蒙特卡罗模拟；在本研究中，我们利用它选择用于 DFT 计算的构型并开展统计力学计算 [32, 33, 34]。候选构型选自 O3-Li$_x$CoO$_2$ ，其锂组成范围为 $x = 0$ 至 $x = 1$。

簇是锂子晶格上若干位点的集合。给定一个位点簇 $\alpha = \{i, j, \ldots, k\}$，可以将多项式 ϕ_α 定义为这些位点的占位变量之积，即 $\phi_\alpha = \sigma_i \sigma_j \cdots \sigma_k$。簇展开是多项式 ϕ_α的线性组合，因此形成能可表示为以下形式：

$$E_{\text{f}}(\boldsymbol{\sigma}) = V_0 + \sum_{\alpha} V_\alpha \phi_\alpha(\boldsymbol{\sigma}), \tag{S5}$$

其中，系数 V_0 和 V_α 经过优化以获得“最佳拟合”，称为有效簇相互作用（ECI）。

我们使用 CASM代码，对 O3宿主结构的 333 个构型通过密度泛函理论计算得到的形成能进行簇展开拟合。拟合采用一种将遗传算法与加权线性回归相结合的稀疏回归技术。在这种方法中，利用彼此距离在给定范围内的晶格位点的单点簇、二元簇、三元簇和四元簇，构建多个候选基函数。遗传算法用于从这些基函数中选择一个子集，将其纳入簇展开。随后使用线性回归计算每个基函数子集的系数。我们采用最高至四元簇的簇类型；二元簇、三元簇和四元簇的最大晶格位点距离分别为 24 $\mathring{A}$ 、 8 $\mathring{A}$ 和 6 $\mathring{A}$ ，总计得到 221 个候选基函数。在线性回归过程中，我们采用权重 w_i ，通过数据点 i 与该构型处凸包之间的能量差 ΔE_i^{hull}，使拟合偏向于提高对凸包上或其附近构型的准确性：

$$w_i = 15 \exp\left(-\frac{\Delta E_i^{\text{hull}}}{0.005}\right) + 0.5 \tag{S6}$$

此外，遗传算法受到约束，必须选择这样的基函数：由簇展开对全部 333 个构型的预测结果构建的凸包，与密度泛函理论凸包包含相同的构型。

S1.2.2. 对称性适配序参量

除组分外，对称性适配序参量对于观察和追踪有序-无序转变，以及识别给定有序结构的各种平移和旋转变体也很有用，例如图3a所示组分为 $x \approx 1/2$ 的之字形有序结构中的变体。识别对称性适配序参量的方法已在其他文献中介绍；例如参见 Natarajan 等人的文献 [28]。在这一过程中，会针对每个变体确定相应的超胞（图3b）。随后确定一个互相公度超胞，该超胞涵盖每个不同的超胞，因此足以表示全部12个变体。对于之字形有序结构，该超胞包含32个锂子晶格位点（图3c-3d）。该超胞的亚晶格组分构成一个基，可将每个变体描述为向量 $\boldsymbol{x} \in \mathbb{R}^{32}$；如果相应的子晶格位点完全被锂占据，则对应分量为1，否则为0。

之字形有序的对称群 P 由12种变体之间的384个不同线性变换组成，每种变体均由 $\mathbb{R}^{32\times32}$ 中的一个矩阵表示。根据 Thomas和Van der Ven [35]，提出的算法，我们构造了一个 P-不变矩阵，并对其进行了特征值分解。结果得到8个非零特征值：其中两个各不相同，两个分别重复三次，四个分别重复六次；相应地得到8组特征向量，共计32个特征向量。这些特征向量构成了正交变换矩阵 $\boldsymbol{Q} \in \mathbb{R}^{32\times32}$ 的行，该矩阵将亚晶格组成向量 $\boldsymbol{x}$ 映射为包含32个序参量的向量 $\boldsymbol{\eta}$。

通过对描述该有序结构12种变体的向量 $\boldsymbol{x}^{(i)}$、 $i = 1$、……、12 施加 $\boldsymbol{Q}$ ，确定了与之字形有序相关的序参量子集。采用从零开始的索引， $\boldsymbol{\eta} \in \mathbb{R}^{32}$ 的分量 $\eta_7^{(i)}$ 至 $\eta_{31}^{(i)}$ 对所有变体 $i = 1$、……、12 均为零，因此与描述之字形有序无关（图3c）。由此确定了7个相关序参量，即 η_0、……、η_6。其中第一个序参量 η_0 与两个不同特征值中的一个相关，对应于在全部 32 个子晶格位点上取平均得到的组分；即 $\eta_0 = x$。另外六个序参量 η_1、……、η_6 与具有6个相应特征向量的简并特征值中的一个相关。

由于自由能密度在将三角形Li子晶格的有序变体相互映射的变换下保持不变，因此，IDNN 表示采用了在这些变换下关于 η_0，$\ldots\eta_6$ 对称的特征。我们选取了最高六阶的单项式，并通过求和作用Reynolds算子：

$$h(\boldsymbol{\eta}) = \sum_{\boldsymbol{M}^{(\eta)}\in\mathcal{M}} f(\boldsymbol{M}^{(\eta)}\boldsymbol{\eta}), \tag{S7}$$

其中，对于每个 $\boldsymbol{M} \in P$，我们有 $\boldsymbol{M}^{\eta} = \boldsymbol{Q}\boldsymbol{M}$。此操作生成了式（1）中的P-不变多项式函数，并将其作为IDNN特征。

S1.2.3. 蒙特卡罗采样

给定 E_{f}（$\boldsymbol{\sigma}$），我们在半巨正则系综中进行采样。在该系综中，化学势是指定的，而相应的组分和/或序参量则通过系综平均确定。半巨正则系综的配分函数如下：

$$\Theta = \sum_{\boldsymbol{\sigma}} \exp\left(-\frac{E(\boldsymbol{\sigma}) - M\widehat{\boldsymbol{\eta}}(\boldsymbol{\sigma})\cdot\widehat{\boldsymbol{\mu}}}{k_B T}\right) \tag{S8}$$

其中， $\widehat{\boldsymbol{\eta}} = \langle\eta_0,\ \ldots\ldots,\ \eta_6\rangle^{\mathrm{T}}$ 表示序参量的约化向量， $\eta_0 = x$ 和 $\widehat{\boldsymbol{\mu}} = \langle\mu_0,\ \ldots\ldots,\ \mu_6\rangle$表示相应的化学势向量，$M$ 表示铺满该构型的参考超胞数量， k_B 表示玻尔兹曼常数， T 表示温度。每个化学势分量都是自由能相对于对应组分或序参量分量的导数。

与锂组成相关的化学势与电压之间的关系可表示为组分的函数 $V(x)$:

$$V(x) = -\frac{\mu_0^{\mathrm{cathode}}(x) - \mu_0^{\mathrm{anode}}}{e}, \tag{S9}$$

其中， μ_0^{cathode}是正极（LCO）相对于锂的化学势， μ_0^{anode}是负极的化学势。将负极设为金属锂后， μ_0^{anode}是一个常数，等于锂 [18]的吉布斯自由能。不过，必须为这两个化学势采用一致的参考系。我们用通过密度泛函理论计算得到的总能量近似锂的吉布斯自由能，但蒙特卡罗采样报告的 LCO 化学势基于形成能，而这些形成能根据方程（S1）中的线性关系重新以端元为参考。由于化学势是自由能对组分的导数，这种重新参考使 LCO 的化学势发生了一个常数偏移 k ，其值等于方程（S1）对 x的导数，即 k : $= -E^{\mathrm{tot}}_{\mathrm{LiCoO_2}} + E^{\mathrm{tot}}_{\mathrm{CoO_2}}$。因此，在根据方程（S9）计算电压之前，我们通过从 $\mu_0^{\mathrm{cathode}}(x)$中减去 k 来抵消化学势的这一偏移。这样便为 μ_0^{cathode}和 μ_0^{anode}施加了一致的参考系。

半巨正则系综中的蒙特卡罗采样存在一个局限：它无法生成与相分离相关的不稳定区域内的数据。虽然这足以绘制相图并预测电压，但相场模拟需要两相区域内的自由能信息，才能一致地解析相界面。针对温度为 260 K、300 K 和 340 K 的情况，进行了额外的蒙特卡罗计算；这一次采用伞形采样，通过使用偏置势，对有序-无序转变的不稳定区域进行采样 [28, 36, 37, 38, 39]。由于化学势在 $\eta_0 = x = 1/2$ 附近的涨落，以及在$\eta_1, \ldots, \eta_6 \approx \pm 1/2$ 处势阱位置的涨落，对于表示低于 ∼330 K 温度下相场模型中的有序状态十分重要，因此在 260 K 和 300 K 下，对这些区域的化学势数据进行了密集采样。此外，还在 $\eta_0 = x = \{0, 1\}$ 的极端组分和序参量 $\eta_1, \ldots, \eta_6 = \pm 1/2$ 附近进行了额外采样，以帮助捕捉化学势的发散特性以及不稳定区域。此时所使用的配分函数如下:

$$\Theta = \sum_{\boldsymbol{\sigma}} \exp\left(-\frac{E(\boldsymbol{\sigma}) + \sum_{i=0}^{6} \phi_i (\eta_i(\boldsymbol{\sigma}) - \kappa_i)^2}{k_B T}\right) \tag{S10}$$

其中， ϕ_i 和 κ_i 分别决定偏置势的曲率和中心，内层求和则遍历组分和六个序参量。通常，可将 ϕ_i 固定为适当的值，同时改变 κ_i ，以便在蒙特卡罗采样过程中对所需的组分和序参量空间进行采样。组分 $\langle \eta_0 \rangle$ 和各个序参量 $\langle \eta_i \rangle$、 $i = 1$、.....、6的系综平均值通过偏置参数与其对应的化学势相关:

$$\frac{1}{M}\mu_i\Big|_{\langle \boldsymbol{\eta} \rangle} = -2\phi_i(\langle \eta_i \rangle - \kappa_i), \qquad i = 0, \ldots, 6 \tag{S11}$$

具有有序结构的区域可通过连接化学势数据点所描绘曲线的陡峭程度和正斜率来识别。这表明自由能势阱具有强凸性，并且在相应的组成区间内，热力学上更偏好这些结构。

在 $x = 1/3$处也预测到一种有序结构，其出现温度低于约235 K。实验上尚未在室温下观察到这种有序结构，但已有研究报告了低温下的部分证据[40], ，这与我们的结果一致。与此前的第一性原理工作 [1] 相比，所预测的有序-无序温度准确性有所提高，这可能归因于采用了包含范德华相互作用的先进密度泛函理论方法+U （§S1.1）。

蒙特卡罗计算在系综平均值中以方差 $\langle \eta_i \rangle = 3 \times 10^{-4}$ 运行，据此将序参量的精度计算为 Var($\langle \eta_i \rangle - \kappa_i$)。$\langle \mu_i \rangle$ 中的精度可通过（S11）得到。对于标准采样和伞形采样，最小、平均和最大计算的计算运行时间分别为（17 秒、97 秒、64941 秒）和（1 秒、166 秒、13700 秒）。标准采样的每次蒙特卡罗模拟平均有 5487 个点，而伞形采样有 52789 个点。

S1.3. 用于自由能表示的可积深度神经网络（*IDNNs*）

通过使用导数数据训练自由能密度函数，可以得到 IDNN 表示；导数数据的形式为化学势（标签）以及相应的（对称不变量函数）组分或序参量（特征）[41, 42]。IDNN 的可积性内置于其结构中，具体方法是将其构造为标准全连接深度神经网络（DNN）的导数形式。虽然对描述标准 DNN 的方程进行求导以推导 IDNN 方程十分直接，但现代深度学习库使这一步变得不必要。用户只需构建一个标准 DNN，并应用梯度运算来创建 IDNN，随后即可用于训练。在数学上，DNN 可表示为函数 $Y(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{b})$，表示系综平均化学势 $\langle \mu_i \rangle$；其参数或输入 $\boldsymbol{X} \in \mathbb{R}^7$ 表示组分 $\langle x \rangle = \langle \eta_0 \rangle$ 以及其余序参量$\langle \eta_i \rangle$, $i = 1, \ldots 6$、权重 $\boldsymbol{W}$和偏置 $\boldsymbol{b}$。DNN 的训练涉及针对给定数据集 $\{(\boldsymbol{X}_\theta, Y_\theta)\}$，求解关于权重和偏置的优化问题：

$$\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{b}} = \underset{\boldsymbol{W}, \boldsymbol{b}}{\arg\min}\, \mathrm{MSE}\left(Y(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{b})\Big|_{\widehat{\boldsymbol{X}}_\theta}, \widehat{Y}_\theta \right) \tag{S12}$$

这里关注的情况是，数据集 $\{(\boldsymbol{X}_\theta,\ Y_\theta)\}$ 不可用。相反，我们拥有导数数据集 $\{(\widehat{\boldsymbol{X}}_\theta,\ \widehat{y}_{\theta_k})\}$，其中 $\widehat{y}_{\theta_k}$对应于 $\widehat{Y}_\theta$相对于 $\boldsymbol{X}_\theta$的 k^{th}分量的偏导数。为了使用这些数据，IDNN 被定义为 Y 相对于其输入 $\boldsymbol{X}$的梯度，即$\partial Y(\boldsymbol{X},\ \boldsymbol{W},\ \boldsymbol{b})/\partial X_k$。训练定义如下：

$$\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{b}} = \underset{\boldsymbol{W}, \boldsymbol{b}}{\arg\min} \sum_{k=1}^{n} \mathrm{MSE}\left(\frac{\partial Y(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{b})}{\partial X_k}\Big|_{\widehat{\boldsymbol{X}}_\theta}, \widehat{y}_{\theta_k} \right) \tag{S13}$$

所得的优化后的权重 $\boldsymbol{W}$和偏置 $\boldsymbol{b}$可与函数 $\partial Y(\boldsymbol{X},\ \boldsymbol{W},\ \boldsymbol{b})/\partial X_k$结合使用，以返回化学势的预测值。使用函数 $Y(\boldsymbol{X},\ \boldsymbol{W},\ \boldsymbol{b})$中的相同权重和偏置，可以精确表示其原函数。对于当前工作，$\partial Y(\boldsymbol{X},\ \boldsymbol{W},\ \boldsymbol{b})/\partial X_k$给出了化学势的 IDNN 表示，而 $Y(\boldsymbol{X},\ \boldsymbol{W},\ \boldsymbol{b})$ 是自由能的 DNN 表示。

S1.4. 采样和主动学习工作流

为了训练一个能够表示组分加上与之字形有序相关的六个序参量的化学势数据的 IDNN，我们必须在七维空间中进行采样。我们不在整个空间内均匀且密集地采样，因为这可能需要数量多到难以承受的蒙特卡洛模拟，而是重点关注那些在物理上具有重要意义且其特征可能难以捕捉的区域。这些区域包括与之字形有序变体相关的能量井，以及序参量空间边界处化学势的发散行为，其中包括 $\eta_0 = x = \{0, 1\}$ 处的组分端元。我们还对序参量空间进行一些一般性的无引导采样，以捕捉总体趋势。因此，这种第一种采样方法涉及对探索已知先验感兴趣区域的采样。

由于采样的目的是开发一个自由能代理模型，我们将探索采样与第二种采样方法相结合，利用利用部分训练的代理模型来识别其他可能包含有助于改进代理模型的数据的区域。对于当前工作流程，这种利用采样仅包括识别指定数量的数据点，在这些点处代理模型表现出较高的逐点误差。较高的误差表明该区域的能量景观难以捕捉，因此我们在此处采集更多数据。探索/利用采样相结合的方法构成了主动学习工作流，其中依次进行探索采样、IDNN 训练和利用采样，直至满足停止准则。

在对该空间进行采样之前，有必要先定义空间的边界。子晶格参数空间是32维单位超立方体，而将 $\boldsymbol{Q}$（图3e中）限制为其前七行后得到的 $\boldsymbol{Q}$ 约化矩阵，则实现了从亚晶格空间到序参量空间的变换。

$\widehat{\ }$通过 $\boldsymbol{Q}$变换后的亚晶格空间均匀采样，并不会产生序参量的均匀采样，部分原因在于降维。相反，我们在序参量空间中定义边界平面。首先，我们将空间限制为 η_7 至 η_{31} 均等于零，即任何有序结构都只能是之字形有序的一种变体。利用这一限制条件、关系式 $\boldsymbol{\eta} = \boldsymbol{Q}\boldsymbol{x}$，并应用亚晶格组分满足 $0 \leq x_i \leq 1$、 $i = 0$、.....、31这一物理约束，我们可以使用完整 $\boldsymbol{Q}$ 矩阵的逆来定义以下上边界平面：

$$\sum_{j=0}^{6} Q_{ij}^{-1}\eta_j \leq 1, \qquad i = 0, \ldots, 31 \tag{S14}$$

以及下界平面：

$$\sum_{j=0}^{6} Q_{ij}^{-1}\eta_j \geq 0, \qquad i = 0, \ldots, 31 \tag{S15}$$

其中， $\mathbb{R}^{32}$ 和 $\boldsymbol{Q} \in \mathbb{R}^{32\times 32}$中的向量采用从零开始的索引。

我们使用 Billiard Walk [43] 随机采样算法，在边界平面内进行采样。该算法概述如下：给定边界平面内的一个初始点，随机选择一条轨迹及其长度。沿轨迹前进，直至到达某个边界平面；此时根据轨迹在该平面上的反射情况更新轨迹。反射过程持续进行，直至走完轨迹的全长，由此确定 Billiard Walk 中的下一个点。重复此过程，直到获得所需数量的全局采样内部点。

Billiard Walk 的一个副产物是一组准随机边界点。在工作流程的早期迭代中，我们随机选取其中一部分边界点，以帮助捕捉边界处化学势的发散行为。在端元和能量井附近的区域内，我们在边长为 0.15 的超立方体中进行随机采样。此外，我们还明确地沿着并在序参量空间中的有序-无序转变路径附近进行采样，采样位置为 $\eta_0 = x = 0.5$。

主动学习工作流（图4a、b）通过引导采样，在降阶 $\widehat{\boldsymbol{\eta}} \in \mathbb{R}^7$ 子空间的边界、 高误差点、势阱和不稳定区域附近进行探索、训练和利用。该工作流在（a）第二次工作流迭代时，或（b）使用最新数据集计算的两个先前IDNN模型的均方误差（MSE）从一个模型到另一个模型增加时，强制重新搜索IDNN超参数。如果MSE降低，则该工作流允许使用最新数据继续训练先前的IDNN（图4b）。

S1.5. 相场理论及相关计算框架

微结构的演化和相变可以使用相场方程进行建模。Cahn-Hilliard方程 [44] 描述守恒量（例如组分）的动力学，而非守恒序参量场则使用Allen-Cahn方程 [45]进行建模。在忽略弹性效应时，具有组分和 $n = 6$ 序参量的系统总自由能可描述如下：

$$\Pi[x, \widehat{\boldsymbol{\eta}}] = \int_{\Omega} \left(f(x, \widehat{\boldsymbol{\eta}}) + \frac{1}{2}\chi_0 |\nabla x|^2 + \sum_{i=1}^{6} \frac{1}{2}\chi_i |\nabla \widehat{\eta}_i|^2 \right) \mathrm{d}V \tag{S16}$$

其中， χ_i 为梯度参数， $f(x,\ \widehat{\boldsymbol{\eta}})$为自由能密度，在本研究中由解析积分的深度神经网络表示。

相场方程中使用的化学势 $\widetilde{\mu}_i$ 由总自由能的变分导数给出，即 $\widetilde{\mu}_0 := \delta\Pi/\delta x$ 和 $\widetilde{\mu}_i := \delta\Pi/\delta\widehat{\eta}_i$, $i = 1,\ .....n$（为清晰起见，使用 $\widehat{\bullet}$记号；

不过， $\widehat{\eta}_i = \eta_i$ 沿用前几节中的记号）：

$$\widetilde{\mu}_0 = \frac{\partial f}{\partial x} - \chi_0 \nabla^2 x \tag{S17}$$

$$\widetilde{\mu}_i = \frac{\partial f}{\partial \widehat{\eta}_i} - \chi_i \nabla^2 \widehat{\eta}_i, \qquad i = 1, \ldots, n \tag{S18}$$

Cahn-Hilliard方程和Allen-Cahn方程分别如下：

$$\frac{\partial x}{\partial t} = \nabla \cdot (\widetilde{M} \nabla \widetilde{\mu}_0) \tag{S19}$$

$$\frac{\partial \widehat{\eta}_i}{\partial t} = -L \widetilde{\mu}_i, \qquad i = 1, \ldots, n \tag{S20}$$

$\widetilde{\ }$其中， M 为迁移率， L 为动力学系数。我们将化学势方程代入，并以弱形式写出控制方程，以便使用混合有限元方法求解：

$$0 = \int_\Omega \left(w_x \frac{\partial x}{\partial t} + \widetilde{M} \nabla w_x \cdot \nabla \tilde{\mu}_0 \right) \mathrm{d}V - \int_{\partial \Omega} w_x j_n \mathrm{d}S \tag{S21}$$

$$0 = \int_\Omega \left[w_{\tilde{\mu}_0} \left(\tilde{\mu}_0 - \frac{\partial f}{\partial x} \right) - \chi_0 \nabla w_{\tilde{\mu}_0} \cdot \nabla x \right] \mathrm{d}V \tag{S22}$$

$$0 = \int_\Omega \left[w_i \frac{\partial \widehat{\eta}_i}{\partial t} + L \left(w_i \frac{\partial f}{\partial \widehat{\eta}_i} + \chi_i \nabla w_i \cdot \nabla \widehat{\eta}_i \right) \right] \mathrm{d}V, \qquad i = 1, \ldots, n \tag{S23}$$

其中， w_x、 $w_{\tilde{\mu}_0}$和 w_i 为权函数。对于以这种混合形式写出的方程，在 x、 $\tilde{\mu}_0$和 $\widehat{\eta}_i$、 $i = 1$、.....、 n上施加了以下诺伊曼边界条件，其中 ∂为Ω， $\boldsymbol{n}$ 为外法向单位向量， j_n 为流入通量：

$$\nabla x \cdot \boldsymbol{n} = 0 \tag{S24}$$

$$\widetilde{M} \nabla \tilde{\mu}_0 \cdot \boldsymbol{n} = j_n \tag{S25}$$

$$\nabla \widehat{\eta}_i \cdot \boldsymbol{n} = 0, \qquad i = 1, \ldots, n \tag{S26}$$

为比较LCO在不同温度下的行为，我们在260 K和300 K下进行相场模拟，并采用这些温度下蒙特卡罗计算得到的数据。使用各温度对应的解析积分自由能深度神经网络来表示式（S18）中的 $f(x、\ \widehat{\boldsymbol{\eta}})$。利用文献 [46]中的预测值，我们为300 K下的扩散率 D 定义如下组分依赖型代理函数：

$$D = 0.01 \exp(-274(1.05 - x)(0.47 - x)(1. - x)) \tag{S27}$$

扩散率代理函数和预测数据见图 S5，其中有效振动$\widetilde{\ }$频率 ν^* 据报告约为 10^{13} s^{-1} [46]量级。迁移率 M 与 D 的关系由方程$\sim M = Dx/(k_B T)$ [47]给出。将 D 乘以 4，以近似计算 340 K 下的扩散率；对于 260 K，则除以 4。该近似源于相关研究：研究人员结合密度泛函理论和动力学蒙特卡罗计算了 LCO 中锂的扩散，并考虑了迁移能垒 [46]随组分变化的情况。这些计算给出了 300 K 和 400 K 下随组分变化的扩散率，其中 400 K 下的 D 约为 300 K 下扩散率的 30 倍；据此，我们推得温度每变化 40 K，扩散率因子约变化 4 倍。

我们获得了方程（S16）中梯度参数 χ_i 的取值，使界面能的相场计算结果与 S1.1.1 节给出的 DFT 结果一致。设 γ 和 β 分别为通过 DFT 和相场$\widehat{\ }$理论计算得到的、组分为 $x = 0.5$ 时两个有序 LCO 变体的反相界面能。同样地，考虑有序变体与以下组分的无序基体之间的界面能：

组分为 $x = 0.5$ 的有序变体与组分为 $x > 0.5$ 和 $\eta_i = 0$、$i = 1$、.....、6 的无序基体之间的界面能，分别通过 DFT 和相场理论计算得到。由于无法以任何程度的唯一性定义无序 LCO 原子结构，因此一致地计算 $\widehat{\gamma}$ 具有挑战性。因此，我们采用如下估计：$\widehat{\gamma} = \frac{1}{2}\gamma$ [48]。

随后，我们定义一个逆相场问题：给定 γ，求解 χ_0 和 $\chi = \chi_1 = \ldots = \chi_6$，使得 $\sqrt{(\beta - \gamma)^2 + (\widehat{\beta} - \frac{\gamma}{2})^2} < \varepsilon$，其中我们将阈值设为 $\varepsilon = 0.1\,\mathrm{mJ/m}$。我们在长度和高度分别为 0.2 m 和 0.005 m 的薄矩形区域上计算 β 和 $\widehat{\beta}$。通过执行穷举二维网格搜索，我们获得最优的 χ_0 和 χ。推断得到的数值为：260 K 时 $\chi_0 = 1 \times 10^{-4}\,\mathrm{mJ/m}$，300 K 时 $\chi_0 = 1.88 \times 10^{-4}\,\mathrm{mJ/m}$；260 K 时 $\chi_1, \ldots, \chi_6 = 2.12 \times 10^{-8}\,\mathrm{mJ/m}$，300 K 时 $\chi_1, \ldots, \chi_6 = 4.91 \times 10^{-8}\,\mathrm{mJ/m}$。在这些数值下，最小化后的函数在 260 K 时为 0.05 mJ/m，在 300 K 时为 0.08 mJ/m。

模拟采用有限元法，通过基于 deal.II [49] 库的 mechanoChemFEM 代码[1], 执行，并在密歇根大学的 Great Lakes 高性能计算集群上运行。模拟中采用了带悬挂节点的自适应网格划分和自适应时间步进。

S2. 反相界面能的 DFT计算数据

表 S1：使用大规模 DFT计算，并对离子力和晶胞应力进行弛豫，计算得到的有序变体 $\eta_1 = \frac{1}{2}$ 与 $\eta_3 = \frac{1}{2}$ 之间的反相畴界界面能（见图3a）。

MC 数量 超晶胞（$2N_{\mathrm{cell}}$）	数量 原子（电子）	γ' （毫焦耳/米²）
2	448（3904）	45.9
4	896（7808）	57.9
6	1344（11712）	74.4
γ		30.9

S3. 显式成核

利用经典成核速率在相场模拟中定义显式成核的想法由 Simmons 等人提出 [50, 51]。晶核在一个单元内于一个时间步 P_n 中形成的概率 Δt 为：

$$P_n = 1 - \exp(-J^* \Delta t) \tag{S28}$$

单元内的成核速率 J^* 为

$$J^* = ZN\beta^* \exp\left(-\frac{\Delta G^*}{k_B T}\right) \exp\left(\frac{\tau}{t}\right) \tag{S29}$$

其中， Z 为 Zeldovich 因子， N 为该微元中的原子数（在本例中为锂位点数）， β^* 为频率因子， ΔG^* 为形成稳定晶核所需的活化能， k_B 为玻尔兹曼常数， T为温度。我们在 τ下忽略孕育项。 Z 和 β^* 的表达式由 [52]给出：

$$Z = \frac{3v_{\mathrm{o}}}{4\pi^{3/2}} \left[\frac{\Delta G^*}{k_B T}\right]^{1/2} \frac{1}{R^{*3}} \tag{S30}$$

$$\beta^* = \frac{4\pi R^{*2} D x_0}{a^4} \tag{S31}$$

[1]代码可在 github.com/mechanoChem/mechanoChemFEM 获取

其中， v_{o} 为有序相中每个（锂）原子的平均体积（约为晶胞体积的两倍）， R^{*}为临界半径， D 为扩散率（见公式（S42））， x_0 为平均成分， a 为晶核相和基体相的平均晶格参数（近似取晶胞体积的立方根）。我们可以合并前因子中的三个项并略作简化，并令 v_{e} 为$= Nv_o$，即该微元的体积：

$$ZN\beta^{*} = \frac{3v_{\mathrm{e}}Dx_0}{R^{*}a^4}\left[\frac{\Delta G^{*}}{\pi k_B T}\right]^{1/2} \tag{S32}$$

Sear 对经典成核理论的综述中给出了类似的成核速率表达式，并解释了如何推导 ΔG^{*} [53]。形成晶核的自由能变化等于体能变化加上引入的界面能。对于球形析出物，总能量变化 ΔG 如下

$$\Delta G = -\frac{4}{3}\pi R^3 \Delta G_v + 4\pi R^2 \gamma \tag{S33}$$

其中， R 是析出物的半径， ΔG_v 是单位体积自由能的降低量（即，正的ΔG_v 表示自由能密度下降；稍后将进一步介绍这一术语）， γ 是界面能。ΔG^{*} 使该能量达到最大值，因为它就是能垒，因此我们对 R 求导并令其等于零，从而求得临界半径 R^{*}:

$$0 = -4\pi {R^{*}}^2 \Delta G_v + 8\pi R^{*}\gamma \tag{S34}$$

$$\implies R^{*} = \frac{2\gamma}{\Delta G_v} \tag{S35}$$

代回公式（S33）即可得到均匀成核的数值

$$\Delta G^{*}_{\mathrm{hom}} = -\frac{4}{3}\pi\left(\frac{2\gamma}{\Delta g_v}\right)^3 \Delta G_v + 4\pi\left(\frac{2\gamma}{\Delta g_v}\right)^2 \gamma \tag{S36}$$

$$= \frac{16\pi\gamma^2}{3\Delta G_v^3} \tag{S37}$$

严格来说，这适用于均匀成核。对于异质成核，还需乘以另一个项 f（θ），该项是界面与（平坦）边界之间夹角 θ 的函数:

$$\Delta G^{*}_{\mathrm{het}} = f(\theta)\Delta G^{*}_{\mathrm{hom}} \tag{S38}$$

$$f(\theta) = \frac{1}{2} - \frac{3}{4}\cos\theta + \frac{1}{4}\cos^3\theta \tag{S39}$$

请注意， $f(\pi/2) = 0.5$，这就是此处采用的数值。

Poduri 和 Chen [54] 建议通过在潜在成核位点（当前处于无序状态，在图 S1a 中标记为 x_{mat} ）的组分处，针对无序自由能密度绘制相对于 x 的切线，来计算驱动力 ΔG_v。 ΔG_v 是这条无序切线与有序自由能密度之间的最大差值。这发生在组分（在图 S1 中标记为 x_{nuc} ）处，此时有序切线与无序切线平行。当无序相的组分使其与有序相具有重合的公切线时， ΔG_v 为零（见图 S1b）。

自由能密度 DNN 的单位为电子伏特/晶胞。晶胞体积为 32.502 $\mathring{\mathrm{A}}^3$, 时，转换为 J/m^3 的换算因子为 4.926×10^9 J/m^3 电子伏特/晶胞。图 S2 显示了~~300 K~~下无序相内给定组分处的驱动力 ΔG_v。在电池的工作范围内（即 $x \geq 0.5$），ΔG_v 的最大值为 3.563×10^6 J/m^3，对应于 $x = 0.5$，并在约 $x = 0.537$处变为零。

为探究界面能的影响，图 S3 展示了 ΔG 和 R 随界面能的变化，其中采用的最大驱动力为 $\Delta G_v = 3.563 \times 10^6$ J/m^3。

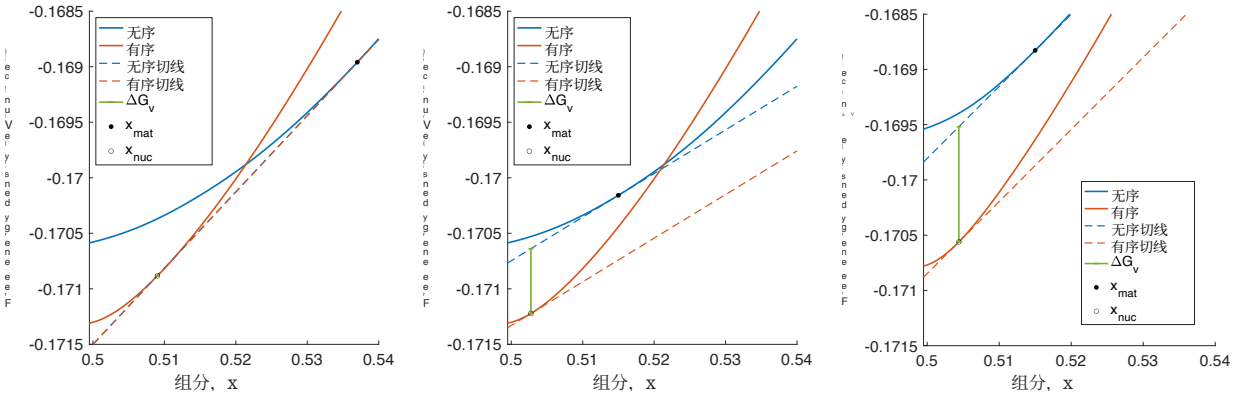
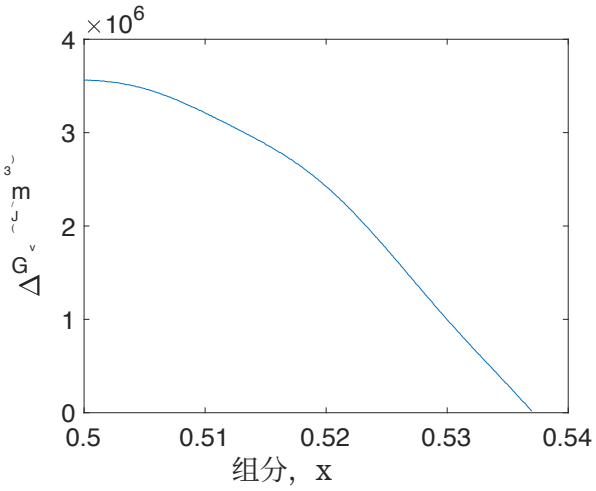


图 S1：（a）展示如何计算成核驱动力 ΔG_v：在组分为 x_{mat} 和 x_{nuc} 时，它们分别对应无序基体和有序晶核。驱动力 ΔG_v 在形成公切线的组分下为零。（b） ΔG_v 是有序和无序自由能密度平行切线之间的垂直距离，如图所示为 300 K 时的情况。（c）260 K 时的 ΔG_v 是 300 K 时数值的 $\sim 2\times$ 。



Fi图 S2：300 K 下无序相内给定组分处的驱动力 ΔG_v 图。

在外加电压下， V，异质成核的临界自由能为

$$\Delta G^*_{\mathrm{het}} = \frac{16\pi\gamma^2}{3(\Delta G_v - V\Delta x)^3} \tag{S40}$$

对于 $\Delta x = x_{\mathrm{nuc}} - x_{\mathrm{mat}}$。相应的异质成核速率为

$$J^* = ZN\beta^* \exp\left(-\frac{\Delta G^*_{\mathrm{het}}}{k_B T}\right) \exp\left(\frac{\tau}{t}\right) \tag{S41}$$

正文中的图5b表示异质成核速率与 V 和 x_{mat}的关系，其中考虑了 x_{nuc} 对 x_{mat} 的依赖性，如图 S1所示。

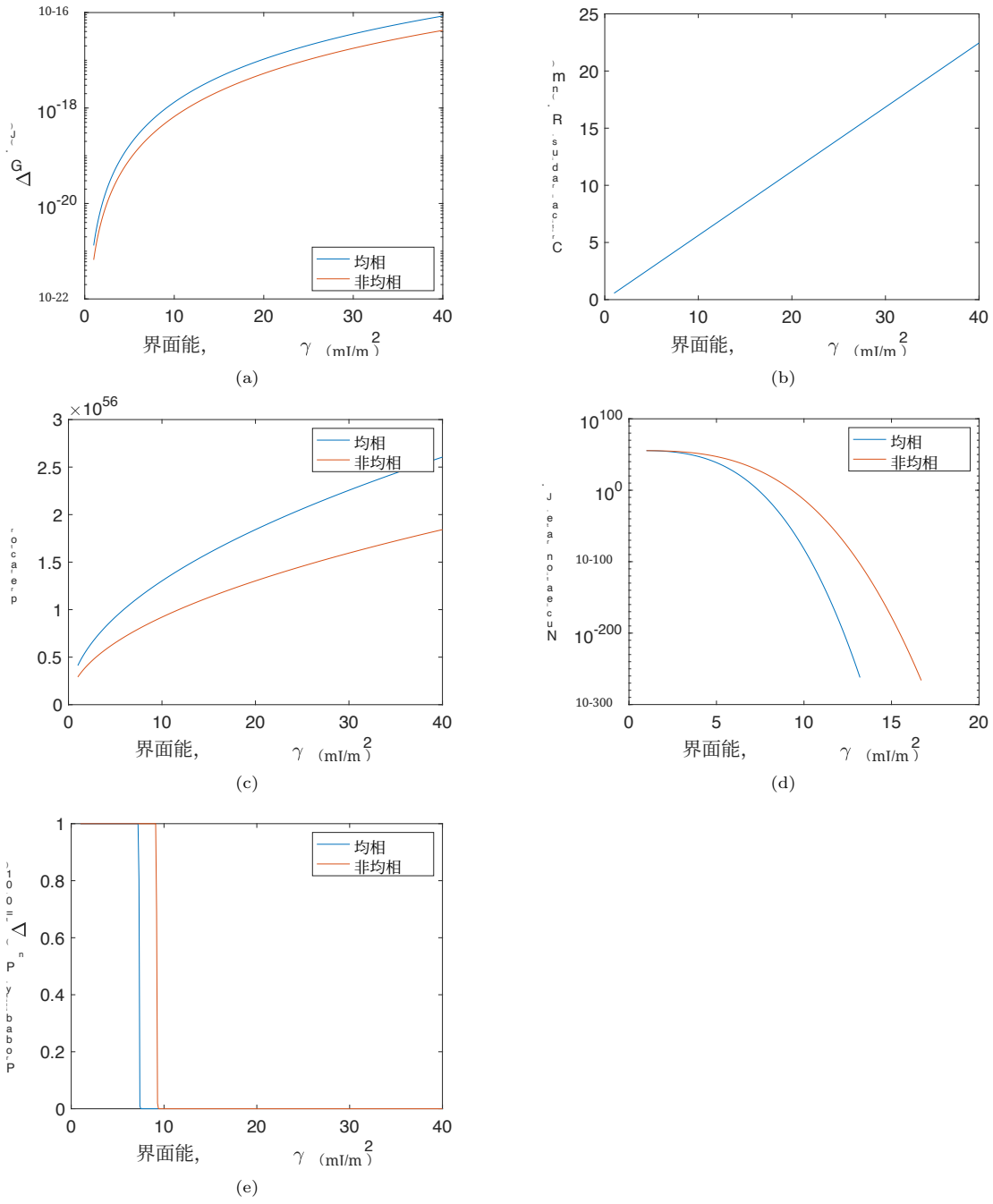


图 S3：（a）成核临界自由能， ΔG^*；（b）临界半径， R^*；（c）前因子， $ZN\beta^*$；（d）成核速率， J^*；以及（e）在 0.01 s 的时间步长内，一个单元中发生成核的概率， P_n；以上各项均为界面能， γ 的函数，并采用 ΔG_v 的最大值。

S4. 忽略有序性的热力学表示

图S4展示了通过仅使用与该组分对应的自由能/化学势数据进行训练，并将组分作为唯一特征而获得的I DNN 表示。这些是一维的

未考虑有序性的表征 [55]。相应的相微结构也一并展示。

图 S5 显示 了根据 [46] 的数据拟合得到的扩散率对组分的依赖关系。拟合函数为 s:

$$D(x) = 0.01\exp(-274(1.05 - x)(0.47 - x)(1. - x)) \tag{S42}$$

以围绕 x =0.525 随机扰动的锂组成初始化、且无边界通量条件下进行的相场模拟结果，对于 50 n m 颗粒展示于图 S6 中。与较大颗粒一样，在 260 K 和 300 K 下会发生旋节分解，但在 340 K 下不会发生。然而，根据温度不同，较小颗粒达到平衡态的速度比大颗粒快 100 到 10,000 倍。此外，尽管两种尺寸颗粒的平衡微观结构相近，但较小颗粒的瞬态微结构要简单得多。

图 S6 中所示 50 nm 颗粒循环模拟的结果表明，与较大颗粒不同，三种温度下的锂组成场基本相同。这再次与 Choi等人 [56] 的实验结果一致，后者显示在 258 至 333 K 的温度范围内性能保持一致。

S5. 对称函数

这12个对称变体在能量上是等价的，因此将这种对称性构建到IDNN中非常重要。我们的做法是使IDNN成为对称函数的函数，而这些对称函数封装了系统的对称性。对称函数本身是序参量的函数。例如，考虑一个更简单的情况（例如B2系统），其中有序化可以通过单个序参量来定义：正值表示一个变体，负值表示第二个变体。无论序参量为正还是为负，能量都应相同，即能量函数关于$\eta = 0$对称。为了实现这一点，我们可以令能量成为 η^2的函数，因为 η^2 封装了适当的对称性。对于更多序参量，对称函数会变得更加复杂，但基本思想相同。

Thomas和Van der Ven在其论文的附录C（以及第5节）中描述了定义对称函数的算法 [35]。其思想是从由序参量构成的给定阶数的完整单项式集合开始，并对每个单项式应用Reynolds算子。请注意，由于现在我们作用于序参量向量 $\boldsymbol{\eta}$ ，而不是子晶格向量 $\boldsymbol{x}$，因此首先将变换矩阵 $\boldsymbol{M}_i^{(x)}$ 转换为 $\boldsymbol{M}_{(\eta)i} = \boldsymbol{Q}\boldsymbol{M}_{(i}^{x}\boldsymbol{Q}^{-1})$。其中许多函数会归约为零函数，而所得的非零函数之间也很可能存在重复，但那些保留下来的函数都是对称不变量。此外，在大多数情况下，最终得到的函数不再是单项式。该过程从一阶单项式开始，然后逐步提高单项式的阶数，直到获得足够数量的对称不变量函数。

并非所有单独的对称函数都体现相同的对称性。例如，函数

$$f(\boldsymbol{\eta}) = \eta_1\eta_2\eta_3\eta_4\eta_5\eta_6 \tag{S43}$$

表明，如果改变单个序参量的符号，自由能可能会有所不同，但这一特征在例如下面这个低阶对称函数中就无法体现：

$$f(\boldsymbol{\eta}) = \eta_1^2 + \eta_2^2 + \eta_3^2 + \eta_4^2 + \eta_5^2 + \eta_6^2 \tag{S44}$$

Reynold 算子通过依次将所有变换应用于单项式的输入并对结果求和来实现，如下所示：

$$h(\boldsymbol{\eta}) = \sum_{\boldsymbol{M}^{(\eta)}\in\mathcal{M}} f(\boldsymbol{M}^{(\eta)}\boldsymbol{\eta}) \tag{S45}$$

其中， $\mathcal{M}$ 是作用于序参量空间的所有变换矩阵 $\boldsymbol{M}^{(\eta)}$ 组成的群， $h(\boldsymbol{\eta})$ 是一个对称不变量函数，而 $f(\boldsymbol{\eta})$ 是初始单项式。

我们可以表示指数向量为 $\boldsymbol{n}$ 、相关系数为 $a^{(}$ 的单项式 $f(\boldsymbol{\eta})$ $^{)}$:

$$f(\boldsymbol{\eta}) = a^{(\boldsymbol{n})} \prod_{m=1}^{\mathrm{len}(\boldsymbol{n})} \eta_m^{n_m} \tag{S46}$$

此处使用的最终对称函数包括六阶以内的函数，具体如下:

$$h_1 = \eta_0 \tag{S47}$$

$$h_2 = \frac{2}{3}\sum_{i=1}^{6}\eta_i^2 \tag{S48}$$

$$h_3 = \frac{8}{3}\sum_{i=1}^{6}\eta_i^4 \tag{S49}$$

$$h_4 = \frac{4}{3}\left[\left(\eta_1^2+\eta_2^2\right)\left(\eta_3^2+\eta_4^2+\eta_5^2+\eta_6^2\right)+\left(\eta_3^2+\eta_6^2\right)\left(\eta_4^2+\eta_5^2\right)\right] \tag{S50}$$

$$h_5 = \frac{16}{3}\left(\eta_1^2\eta_2^2+\eta_3^2\eta_6^2+\eta_4^2\eta_5^2\right) \tag{S51}$$

$$h_6 = \frac{32}{3}\sum_{i=1}^{6}\eta_i^6 \tag{S52}$$

$$\begin{aligned} h_7 = \frac{8}{3}\big[&\left(\eta_1^4+\eta_2^4\right)\left(\eta_3^2+\eta_4^2+\eta_5^2+\eta_6^2\right)+\left(\eta_3^4+\eta_6^4\right)\left(\eta_4^2+\eta_5^2\right)+ \\ &\left(\eta_1^2+\eta_2^2\right)\left(\eta_3^4+\eta_4^4+\eta_5^4+\eta_6^4\right)+\left(\eta_3^2+\eta_6^2\right)\left(\eta_4^4+\eta_5^4\right)\big] \end{aligned} \tag{S53}$$

$$\begin{aligned} h_8 = \frac{16}{3}\big[&\eta_1^2\eta_2^2(\eta_3^2+\eta_4^2+\eta_5^2+\eta_6^2)+\eta_3^2\eta_6^2(\eta_1^2+\eta_2^2+\eta_4^2+\eta_5^2)+ \\ &\eta_4^2\eta_5^2(\eta_1^2+\eta_2^2+\eta_3^2+\eta_6^2)\big] \end{aligned} \tag{S54}$$

$$h_9 = \frac{32}{3}\left(\eta_1^4\eta_2^2+\eta_1^2\eta_2^4+\eta_3^4\eta_6^2+\eta_3^2\eta_6^4+\eta_4^4\eta_5^2+\eta_4^2\eta_5^4\right) \tag{S55}$$

$$h_{10} = 8(\eta_1^2+\eta_2^2)(\eta_3^2+\eta_6^2)(\eta_4^2+\eta_5^2) \tag{S56}$$

$$\begin{aligned} h_{11} = \frac{64}{5}\big[&\eta_1\eta_2(\eta_3^2-\eta_6^2)(\eta_4^2-\eta_5^2)+ \\ &\eta_3\eta_6(\eta_1^2-\eta_2^2)(\eta_4^2-\eta_5^2)+ \\ &\eta_4\eta_5(\eta_1^2-\eta_2^2)(\eta_3^2-\eta_6^2)\big] \end{aligned} \tag{S57}$$

$$h_{12} = 64\sqrt{5}\eta_1\eta_2\eta_3\eta_4\eta_5\eta_6 \tag{S58}$$

S6. 蒙特卡洛精度

每次蒙特卡罗运行都会针对给定的 ϕ、 κ和温度，求得 $\boldsymbol{\eta}$ 和 $\boldsymbol{\mu}$ 的系综平均。用于七维 IDNN 的蒙特卡罗计算中，序参量的精度为 0.0003，置信度为 0.95。蒙特卡洛模拟收敛后，随后使用系综平均计算 $\boldsymbol{\mu}$ 。我们发现，序参量的精度不会影响所确定的$\boldsymbol{\eta}$ 与 $\boldsymbol{\mu}$ 之间的关系，因此不会影响用于训练 IDNN 的数据的准确性。不过，它确实会

影响用于训练 IDNN 的数据，因为根据输入确定的 $\boldsymbol{\eta}$ 准确性较低。我们发现，在 300 K 和 340 K 时，需要更高的精度才能使 $\boldsymbol{\kappa}$ 与 $\boldsymbol{\eta}$之间的关系保持一致。需要精确采样时，就有必要提高序参量的精度。对于我们当前的主动学习方法，全局采样和局部采样都具有随机性，因此，在用于求得 $\boldsymbol{\mu}$ 的 $\boldsymbol{\eta}$ 中引入额外的随机性并不会造成问题。

S7. 最低自由能面

自由能面绘制于 $\eta_0-\eta_1$ 空间中，其中每个 η_0（η_2-η_6 保持为 0）对应的最低自由能曲线以红色绘制。通过使用七维 IDNN 预测 η_0 和正 η_1 空间中每个点的自由能，并确定每个 η_0处的最低自由能，得到该自由能曲线。由于序参量具有对称不变性，这一结论同样适用于负 η_1 和所有其他序参量。对于 260 K 和 300 K（ η_2-η_6 保持为 0），最低自由能的导数存在不连续性，这对应于 η_1 从 0 跳变至（.33-.39）。在 340 K 时，最低能量状态始终为无序状态。

七维空间中的最低自由能曲线绘制于图 S8a 中。这些曲线通过在完整的 $\boldsymbol{\eta}$空间中评估数据点得到，条件是 $\eta_1-\eta_6$ 为正且按递减顺序排列（因为每个序参量都具有对称不变性），并且满足正文方程 15 和 16 中列出的条件。对于每个 η_0，随后确定对应最低自由能的 $\eta's$ 。对于 300 K，沿最低自由能曲线仅有一个序参量非零。该曲线显示于图 S8a 和图 S7 中。对于 260 K，存在两个明显的非无序区域。在其中一个区域，η_1 和 η_2 均为非零，并从 0 跳变至 $\sim .18$。在第二个区域中，η_2 变为 0，而 η_1变为 $\sim .37$。这些不连续性发生的点以图 S8 中的红点表示。此外，图 S7 中 260 K 的 $\eta_0-\eta_1$ 曲线作为绿色曲线显示于图 S8a 中，不连续性则以 S8 中的绿色点表示。

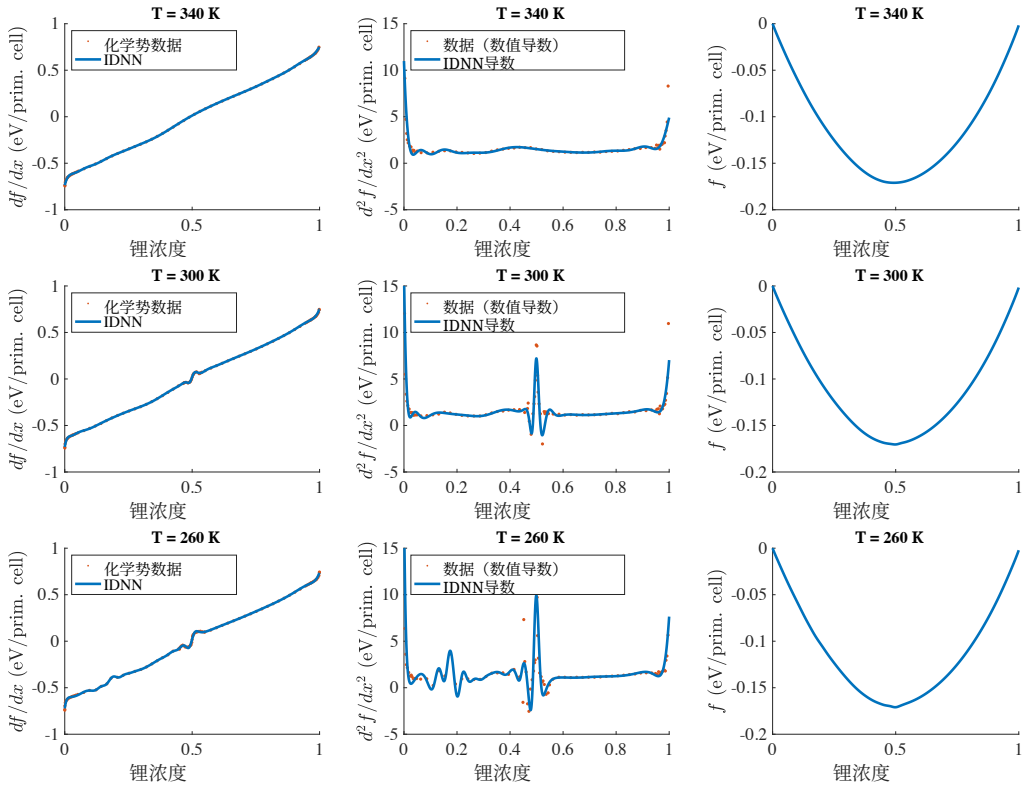
S8. 相图

我们的相图（图 S8b）使用一维 IDNN 基于一维蒙特卡罗伞形采样结果训练后预测的自由能构建。我们在自由能曲线上构造了一组切线，分别对应锂含量大于 0.5 和小于 0.5 的组分，用于确定相不稳定性区域的边界。相图中使用的点在图 S8a 中以蓝色菱形表示。

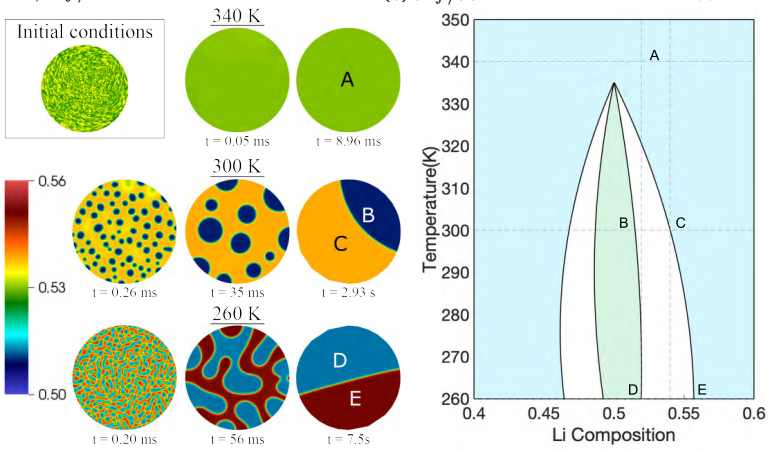
为了确定旋节分解区域，我们还可以使用自由能的二阶导数来确定边界。该区域的一个特征是自由能曲线呈凸形，而在其他情况下，自由能曲线呈凹形。对于一维蒙特卡罗，在两相区内自由能的二阶导数为负，当 $\frac{\partial^2 g}{\partial x^2}=0$ 时即可找到该区域的边界。这些点在图 S8a 中以蓝色圆点表示。

对于七维 IDNN，在所有特征值均为负值的区域中发现自由能具有凸性。然而，我们发现这种情况仅存在于边界区域之外，即正文第 4.4 节（公式 15 和 16）中针对 $\eta_1-\eta_6$ 所描述的区域。我们认为，这解释了 1D 曲线与七维曲线之间的差异。这意味着我们无法使用特征值确定旋节分解区域的边界。相反，我们考察与最低自由能曲线相对应的 $\eta's$ 。

μ_0 空间中存在不连续性，其中 η_1 发生跳变。这些跳变显示在图 S9 中。S8 中的绿色点和红色点表示了这些不连续点。由于我们没有观察到零值或负值的 $\frac{\partial^2 g}{\partial x^2}$，因此使用这些不连续点来近似无序态与两相态之间的转变。从图 S8 可以看出，该不连续点发生在与 1D 无序态—旋节区域边界相近的 η_0 处。在七维中，有序态到旋节分解边界更难确定，可能需要额外采样才能在 IDNN 中正确捕捉该边界。我们确实预计，1D 结果与七维 IDNN 得到的最低自由能将呈现相同的趋势。这种差异可能是由蒙特卡罗计算的不准确性、次优采样，或未使用最优的机器学习模型造成的。



(a) 化学势，$\mathrm{d}f/\mathrm{d}c$ (b) $\mathrm{d}^2 f/\mathrm{d}c^2$ (c) 自由能， $f(c)$



（d）Cahn-Hilliard模拟。

图S4：260 K、300 K 和 340 K 下 IDNN拟合与数据的比较：（a）用于训练的化学势数据（通过偏置势采样）以及 IDNN预测；（b）数值微分数据和化学势 IDNN 的导数；（c）解析积分自由能DNN。（d）基于仅由组分参数化的自由能密度函数，对有序-无序转变进行 Cahn-Hilliard相场计算。其中未包含序参量，因此未表示有序变体。图中标出了无序区域 A、C、E 以及有序区域 C、D。

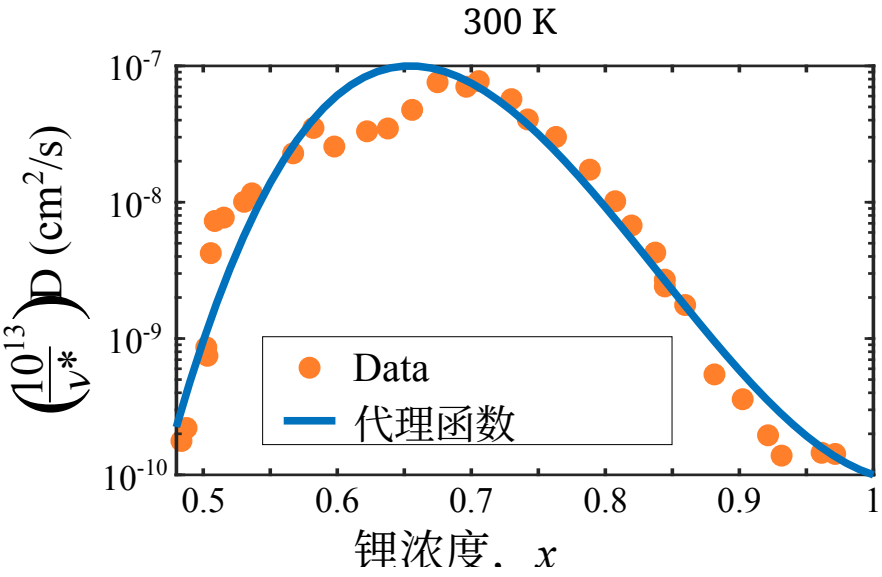


图 S5：根据数据拟合锂扩散率。

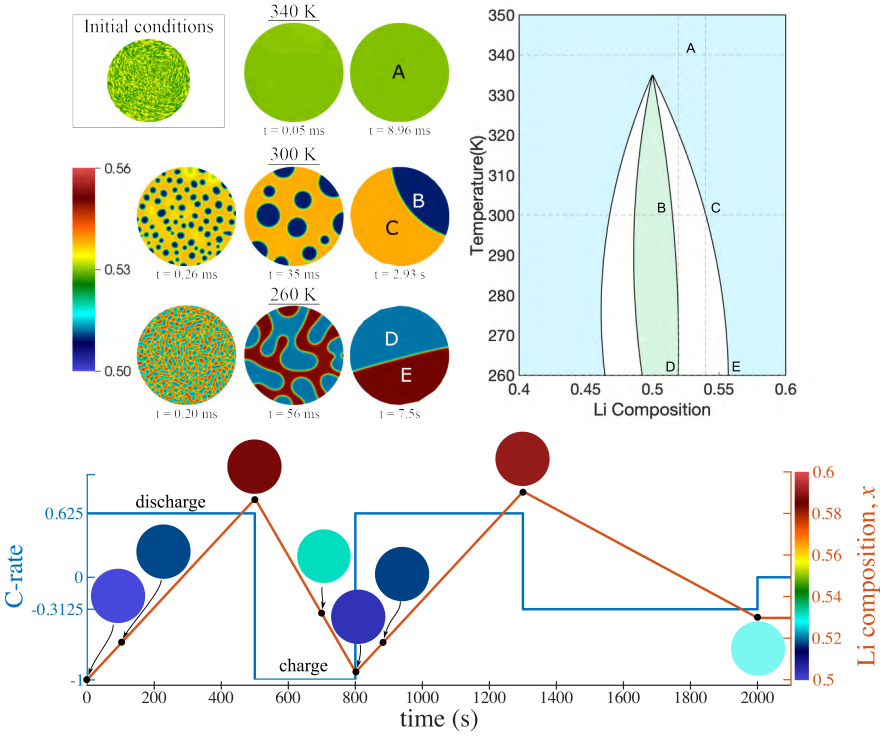


图 S6：260、300 和 340 K 下直径为 50 nm 的颗粒的二维相场模拟结果，显示锂组成；初始锂组成在 $x = 0.53$ 附近随机扰动，且无边界通量。与图 5 中使用 1 μm 颗粒的结果进行比较。相场结果显示了对直径为 50 nm 的颗粒施加循环电流密度后得到的锂组成场。260、300 和 340 K 下的结果在视觉上无法区分，因此这里只列出一组结果。与 1 μm 颗粒相比，施加的电流密度缩放了 1/20 倍，以保持两种颗粒尺寸下相同的 C倍率（见图 5）。蓝线表示施加的 C倍率，红线表示颗粒相应的平均锂组成。

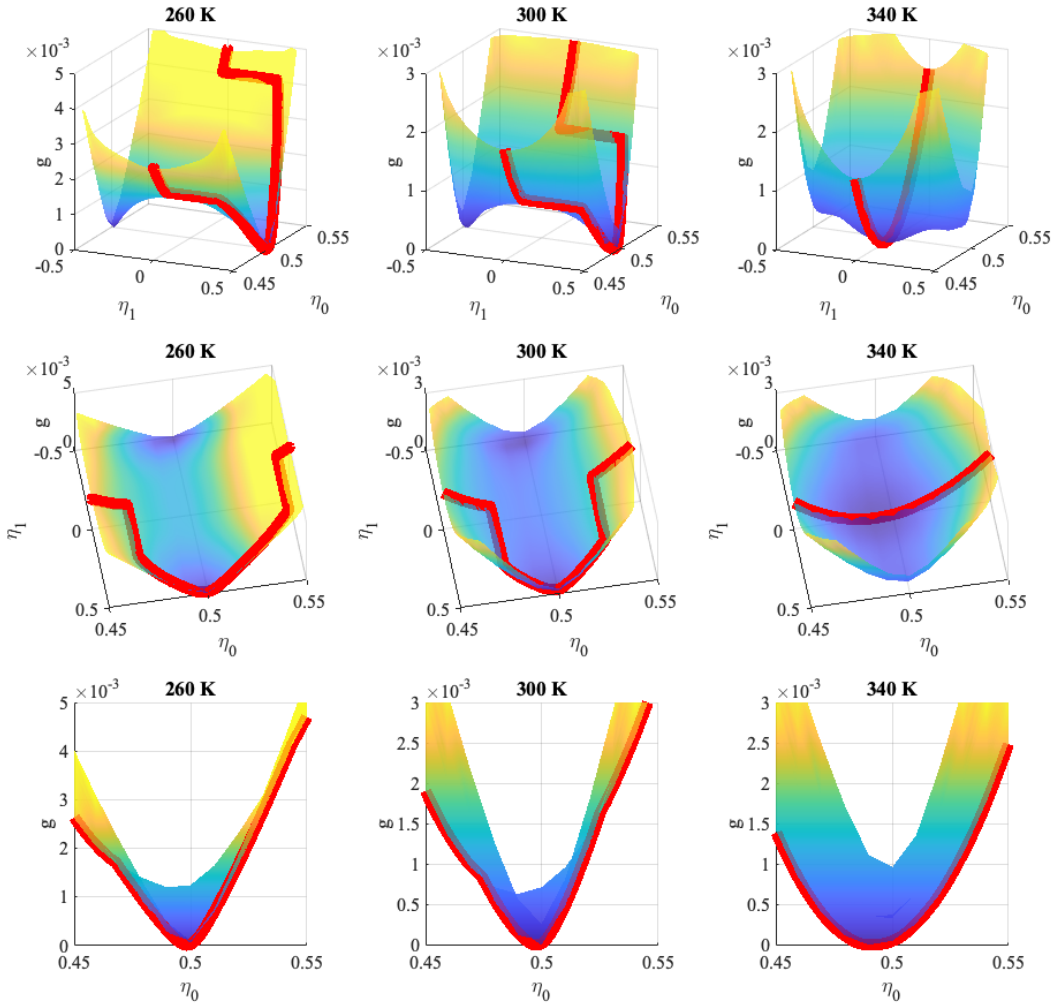
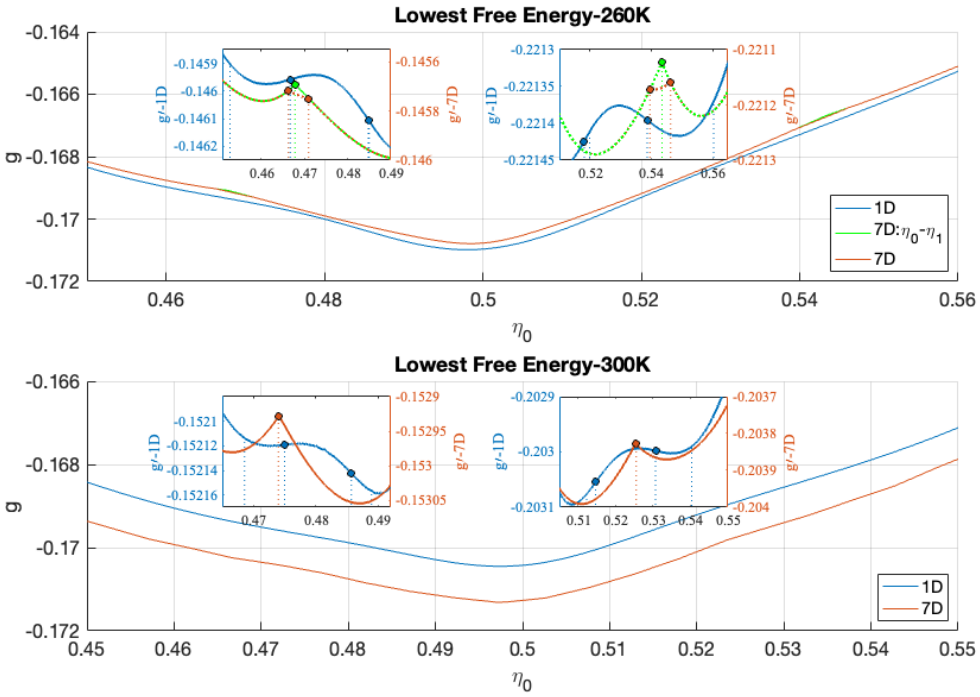
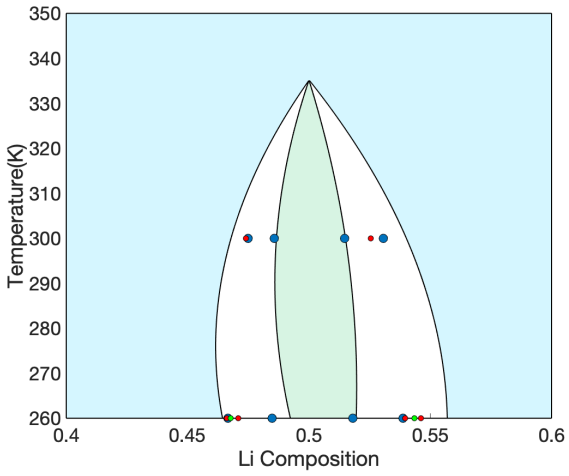


图 S7：在不同取向下，于 $\eta_0 - \eta_1$ 子空间中绘制的自由能密度曲面。由七维 IDNN 确定的 η_1 空间中每个 η_0 对应的最低自由能以红色绘制。



(a)



(b)

图 S8：（a）由 IDNN 确定的最低自由能曲线。一维 IDNN 的曲线以蓝色表示。完整 η 空间中的最低自由能由红色七维曲线表示。对于 260 K，我们还绘制了 $\eta_0 - \eta_1$ 空间中的最低自由能曲线，以绿色表示。（b）中用于绘制相图的点以蓝色菱形表示。这些点在 T260 时位于（0.453 0.485 0.520 0.560），在 T300 时位于（0.469 0.489 0.512 0.540）。对于（a）和（b），1 4667 0 0 4848 0 5181 0 5387 260 0 475 0 4857 0 5147 0 5307 300 7 D IDNN 的二阶导数为零的点以蓝色圆圈表示。这些点在 T 时位于（. . . .），在 T 时位于（. . . .）。D IDNN 自由能中的不连续点以绿色和红色点表示。对于 T260，完整 η 空间中的不连续点（红点）位于（0.466 0.471 0.540 0.546）。在 $\eta_0 - \eta_1$ 空间中，绿色点位于（0.468 0.544）。对于 T300，不连续点位于（0.474 0.525）。（a）中标示两相区凸性的方框中加入了化学势的线性项，以夸大凸区，使得 $g' = g + l\eta_0$。对于 T260：l 分别为 0.05 和 -0.1002，对应左侧和右侧方框。类似地，对于 T300：l 分别为 0.037 和 -0.065。

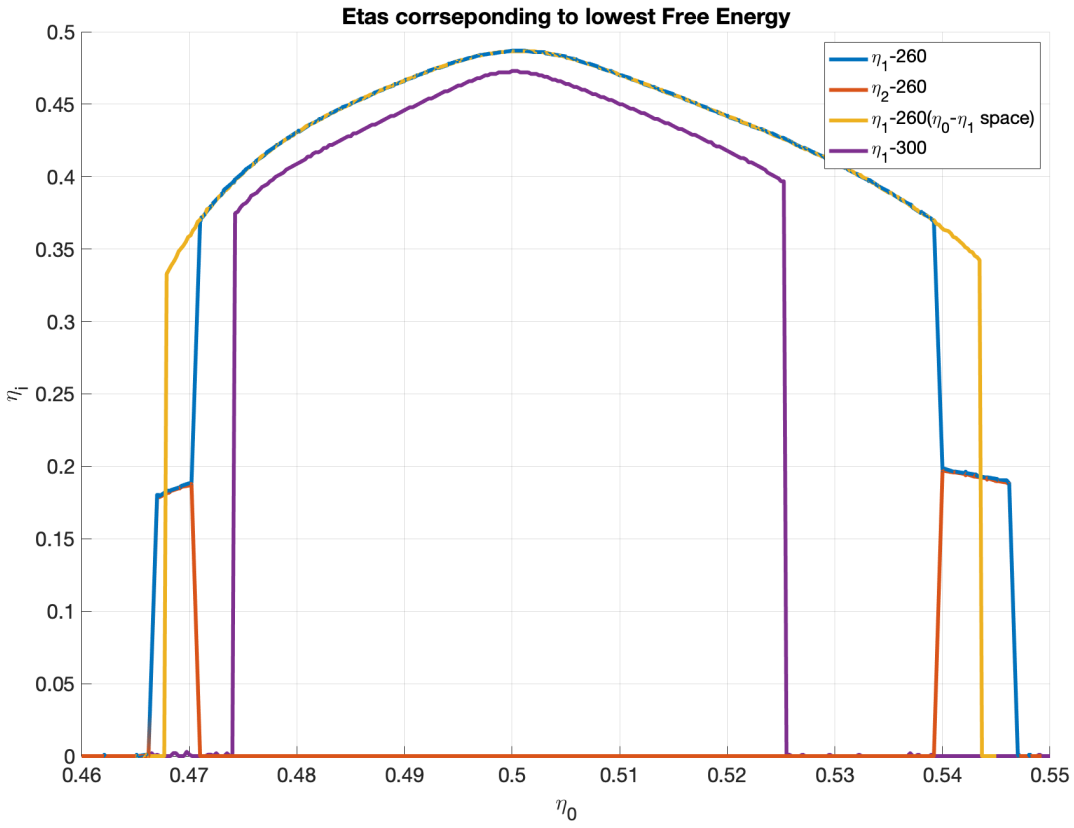


图 S9：该图展示了与最低自由能曲线相对应的 η_0。红色和蓝色曲线展示了在 260 K 下，与最低自由能相对应的 η_1 和 η_2 如何变化。黄色曲线展示了在 260 K 下给定 $\eta_2 - \eta_6$ 为零时， η_1 如何变化。紫色曲线展示了在 $\eta_2 - \eta_6$ 为零的情况下， η_1 在 300 K 下如何变化。

参考文献

[1] A. Van der Ven、M. K.. Aydinol、G. Ceder、G. Kresse 和 J. Hafner。锂$_x$钴氧化物$_2$相稳定性的第一性原理研究。《Physical Review B》，58:2975–2987，1998年。[2] V. I. Anisimov、J. Zaanen 和 O. K. Andersen。能带理论与莫特绝缘体：用 Hubbard U 代替 Stoner I。《Physical Review B》，44:943–954，1991年7月。[3] F. Zhou、M. Cococcioni、C. A. Marianetti、D. Morgan 和 G. Ceder。采用 LDA ＋U第一性原理预测过渡金属化合物中的氧化还原电势。《Physical Review B》，70:235121，2004年12月。[4] M. Aykol 和 C. Wolverton。用于准确计算过渡金属化合物热化学性质的局域环境依赖 GGA＋u 方法。《Physical Review B》，90:115105，2014年9月。[5] M. Aykol、S. Kim 和 C. Wolverton。层状钴酸锂中的范德华相互作用。《The Journal of Physical Chemistry C》，119(33):19053–19058，2015年。[6] M. Cococcioni 和 S. de Gironcoli。LDA＋U 方法中有效相互作用参数计算的线性响应方法。《Physical Review B》，71:035105，2005年1月。[7] T. Thonhauser、V. R. Cooper、S. Li、A. Puzder、P. Hyldgaard 和 D. C. Langreth。范德华密度泛函：自洽势与范德华键的本质。《Physical Review B》，76:125112，2007年9月。[8] J. Klime˘s、D. R. Bowler 和 A. Michaelides。范德华密度泛函的化学精度。《凝聚态物理学杂志》，22(2):022201，2009年12月。[9] D C Langreth、B I Lundqvist、S D Chakarova-K¨ack、V R Cooper、M Dion、P Hyldgaard、A Kelkkanen、J Kleis、Lingzhu Kong、Shen Li、P G Moses、E Murray、A Puzder、H Rydberg、E Schr¨oder 和 T Thon-hauser。稀疏物质的密度泛函。《凝聚态物理学杂志》，21(8):084203，2009年1月。[10] R. Sabatini、E. K¨u.c¨ukbenli、B. Kolb、T Thonhauser 和 S. de Gironcoli。非经验密度泛函描述的应力作用下氨基酸晶体结构演化。《凝聚态物理学杂志》，24(42):424209，2012年10月。[11] T. Thonhauser、S. Zuluaga、C. A. Arter、K. Berland、E. Schr¨oder 和 P. Hyldgaard。金属有机框架中非局域关联结合的自旋特征。《Physical Review Letters》，115:136402，2015年9月。[12] K. Berland、V. R. Cooper、K. Lee、E. Schr¨oder、T Thonhauser、P. Hyldgaard 和 B. I. Lundqvist。密度泛函理论中的范德华力：vdW-DF方法综述。《Reports on Progress in Physics》，78(6):066501，2015年5月。[13] P. Giannozzi、S. Baroni、N. Bonini、M. Calandra、R. Car、C. Cavazzoni、D. Ceresoli、G. L. Chiarotti、M. Cococcioni、I. Dabo、A. Dal Corso、S. de Gironcoli、S. Fabris、G. Fratesi、R. Gebauer、U. Gerstmann、C. Gougoussis、A. Kokalj、M. Lazzeri、L. Martin-Samos、N. Marzari、F. Mauri、R. Mazzarello、S. Paolini、A. Pasquarello、L. Paulatto、C. Sbraccia、S. Scandolo、G. Sclauzero、A. P. Seitsonen、A. Smogunov、P. Umari 和 R. M. Wentzcovitch。Quantum Espresso：用于材料量子模拟的模块化开源软件项目。《凝聚态物理学杂志》，21(39):395502（19页），2009年。[14] P Giannozzi、O Andreussi、T Brumme、O Bunau、M Buongiorno Nardelli、M Calandra、R Car、C Cavaz-zoni、D Ceresoli、M Cococcioni、N Colonna、I Carnimeo、A Dal Corso、S de Gironcoli、P Delugas、R A DiStasio Jr、A Ferretti、A Floris、G Fratesi、G Fugallo、R Gebauer、U Gerstmann、F Giustino、T Gorni、J Jia、M Kawamura、H-Y Ko、A Kokalj、E K¨u.c¨ukbenli、M Lazzeri、M Marsili、N Marzari、F Mauri、N L Nguyen、H-V Nguyen、A Otero de-la Roza、L Paulatto、S Ponc´e、D Rocca、R Sabatini、B Santra、M Schlipf、A P Seitsonen、A Smogunov、I Timrov、T Thonhauser、P Umari、N Vast、X Wu,

以及 S Baroni。使用 Quantum Espresso 进行材料建模的高级功能。《凝聚态物理学杂志》，29(46):465901，2017年。[15] A. Dal Corso。赝势周期表：从 H 到 Pu。《计算材料科学》，95:337 – 350，2014年。[16] PSlibrary 1.0.0。https://github.com/dalcorso/pslibrary。[17] B. Meredig、A. Thompson、H. A. Hansen、C. Wolverton 和 A. van de Walle。在 DFT + u中定位低能解的方法。*Physical Review B*，82:195128，2010年11月。[18] M. K. Aydinol、A. F. Kohan、G. Ceder、K. Cho 和 J. Joannopoulos。锂在金属氧化物和金属二硫属化物中嵌入的第一性原理研究。*Physical Review B*，56:1354–1365，1997年7月。[19] T. Motohashi、T. Ono、Y. Sugimoto、Y. Masubuchi、S. Kikkawa、R. Kanno、M. Karppinen 和 H. Ya-mauchi。层状钴氧化物体系 Li_xCoO_2 ($0.0 \leq x \leq 1.0$) 的电子相图。*Physical Review B*，80:165114，2009年10月。[20] S. Das、P. Motamarri、V. Subramanian、D. M. Rogers 和 V. Gavini。DFT-FE 1.0：使用有限元离散的大规模并行混合 CPU-GPU 密度泛函理论代码。《*Computer Physics Communications*》，280:108473，2022年。[21] P. Motamarri、S. Das、S. Rudraraju、K. Ghosh、D. Davydov 和 V. Gavini。DFT-FE——用于大规模密度泛函理论计算的大规模并行自适应有限元代码。《*Computer Physics Communications*》，246:106853，2020年。[22] P. Motamarri 和 V. Gavini。使用 Kohn-Sham 密度泛函理论进行电子结构计算中的构型力。*Physical Review B*，97:165132，2018年4月。[23] P. Motamarri、M.R. Nowak、K. Leiter、J. Knap 和 V. Gavini。Kohn–Sham 密度泛函理论的高阶自适应有限元方法。《*Journal of Computational Physics*》，253:308–343，2013年。[24] D. R. Hamann。优化的范德堡范数守恒赝势。*Physical Review B*，88:239906，1995年。[25] M. J. van Setten、M. Giantomassi、E. Bousquet、M. J. Verstraete、D. R. Hamann、X. Gonze 和 G-M. Rignanese。the pseudodojo：训练和分级 85 种元素的优化守范赝势表。《*Comput. Phys. Commun.*》，226:39–54，2018年。[26] D. Farkas、M.F. de Campos、R.M. de Souza 和 H. Goldenstein。共格 Ni/Ni3Al 界面的原子尺度结构。*Scripta Metallurgica et Materialia*，30(3):367–371，1994年。[27] Y. Mishin。Ni–Al体系中 γ 和 γ'相的原子尺度建模。《*Acta Materialia*》，52(6):1451–1467，2004年。[28] A. R. Natarajan、J. C. Thomas、B. Puchala 和 A. Van der Ven。经历有序-无序相变的固体的对称性适配序参量与自由能。*Physical Review B*，96:134204，2017年10月。[29] J. M. Sanchez、F. Ducastelle 和 D. Gratias。多组分体系的广义团簇描述。*Physica A*，128:334–350，1984年。[30] D. De Fontaine。合金有序-无序转变的团簇方法。第47卷，《*Solid State Physics*》，第33–176页。Academic Press，1994年。

[31] https://github.com/prisms-center/CASMcode。CASM：统计力学的簇方法，v0.3.dev，2018年。[32] A. Van der Ven、J. C. Thomas、Q. Xu 和 J. Bhattacharya。将固体的电子结构与其热力学和动力学性质联系起来。《Mathematics and Computers in Simulation》，80(7):1393–1410，2010年。[33] J. C. Thomas 和 A. Van der Ven。基于第一性原理研究强非谐性和力学不稳定晶相的有限温度性质。《Physical Review B》，88(21):214111–214111，2013年12月。[34] B. Puchala 和 A. Van der Ven。基于第一性原理计算的 Zr-O 系统热力学。《Physical Review B》，88:094108，2013年9月。[35] J. C. Thomas 和 A. Van der Ven。非线性弹性的探索及其在晶体材料中的高效参数化。《Journal of the Mechanics and Physics of Solids》，107:76–95，2017年10月。[36] G.M. Torrie 和 J.P. Valleau。自由能估计中的非物理采样分布：伞形采样。《Journal of Computational Physics》，23(2):187–199，1977年。[37] Y. Mishin。Ni–Al 体系中 γ 和 γ'相的原子尺度建模。《Acta Materialia》，52(6):1451–1467，2004年。[38] B. Sadigh、P. Erhart、A. Stukowski、A. Caro、E. Martinez 和 L. Zepeda-Ruiz。用于合金中析出物原子尺度模拟的可扩展并行蒙特卡罗算法。《Physical Review B》，85:184203，2012年5月。[39] B. Sadigh 和 P. Erhart。通过跨越相界面的直接热力学积分计算析出物的过剩自由能。《Physical Review B》，86:134204，2012年10月。[40] Y. Shao-Horn、S. Levasseur、F. Weill 和 C. Delmas。探究 O3 层状 Li_xCoO_2 (x≈0.5) 中的锂和空位有序化。《电化学学会志》，150(3):A366，2003年。[41] G.H. Teichert、A.R. Natarajan、A. Van der Ven 和 K. Garikipati。机器学习材料物理学：可积深度神经网络通过学习自由能函数实现尺度桥接。《Computer Methods in Applied Mechanics and Engineering》，353:201–216，2019年。[42] G.H. Teichert、A.R. Natarajan、A. Van der Ven 和 K. Garikipati。材料物理学的尺度桥接：合金自由能函数表示的主动学习工作流与可积深度神经网络。《Computer Methods in Applied Mechanics and Engineering》，371:113281，2020年。[43] B.T. Polyak 和 E.N. Gryazina。Billiard Walk——一种用于控制与优化的新采样算法。《IFAC Proceedings Volumes》，47(3):6123–6128，2014年。第19届 IFAC 世界大会。[44] J. W. Cahn 和 J. E. Hilliard。非均匀系统的自由能。I 界面能。《化学物理学杂志》，28:258–267，1958年。[45] S. M. Allen 和 J. W. Cahn。反相畴界运动的微观理论及其在反相畴界粗化中的应用。《冶金学报》，27:1085–1091，1979年。[46] A. Van der Ven 和 G Ceder。层状 Li_xCoO_2中的锂扩散。《电化学与固态快报》，3(7):301，2000年。[47] T. Jiang、S. Rudraraju、A. Roy、A. Van der Ven、K. Garikipati 和 M. L. Falk。$Li_{1+x}Ti_2O_4$ 电极颗粒中嵌锂诱导应力的多物理场模拟。《The Journal of Physical Chemistry C》，120(49):27871–27881，2016年。

[48] Y. Wang、D. Banerjee、C.C. Su 和 A.G. Khachaturyan。从面心立方固溶体中析出 L12有序金属间化合物的场动力学模型与计算机模拟。《*Acta Materialia*》，46(9):2983–3001，1998年。[49] D. Arndt、W. Bangerth、D. Davydov、T. Heister、L. Heltai、M. Kronbichler、M. Maier、J-P. Pelteret、B. Turcksin 和 D. Wells。deal.II 库，8.5版。《数值数学杂志》，25(3)，2017年4月。[50] J.P Simmons、C Shen 和 Y Wang。通过明确纳入成核事件，对同时成核与生长进行相场建模。《材料学报》，43(10):935–942，2000年。[51] J.P. Simmons、Youhai Wen、C. Shen 和 Y.Z. Wang。采用相场法模拟涉及成核与长大现象的微观结构演化。《材料科学与工程：*A*辑》，365(1):136–143，2004年。多尺度材料建模。[52] FK LeGoues 和 HI Aaronson。晶体学对均匀面心立方-面心立方成核的临界晶核形状和动力学的影响——IV。CuCo合金中理论与实验的比较。《冶金学报》，32(10):1855–1864，1984年。[53] R. P. Sear。成核：蛋白质溶液和胶体悬浮液中的理论与应用。《凝聚态物理学杂志》，19(3):033101，2007年1月。[54]R Poduri 和 L-Q Chen。有序金属间析出物的非经典成核理论——应用于 AlLi合金。《*Acta materialia*》，44(10):4253–4259，1996年。[55] Gregory H Teichert、Sambit Das、Muratahan Aykol、Chirranjeevi Gopal、Vikram Gavini 和 Krishna Garikipati。通过电子结构、统计力学与相场理论之间的机器学习驱动尺度桥接研究 Li_xCoO_2 相稳定性。*arXiv* 预印本 *arXiv:2104.08318*，2021年。[56] S. H. Choi、J-W. Son、Y. S. Yoon 和 J. Kim。锂电池中 $LiCoO_2$ 的颗粒尺寸效应与温度相关性能。《电源学报》，158(2):1419–1424，2006年。