arXiv:2602.10064v1 [cond-mat.mtrl-sci] 10 Feb 2026

PySlice：利用通用原子间势进行常规振动电子能量损失谱预测

Harrison A. Walker[1, 2*], Thomas W. Pfeifer[3,4], Paul M. Zeiger[5,7],Jordan A. Hachtel[3], SokratesT. Pantelides[1,2,6], Eric R. Hoglund[3*]1范德堡大学跨学科材料科学项目，纳什维尔，田纳西州，美国。2范德堡大学物理与天文学系，纳什维尔，田纳西州，美国。3橡树岭国家实验室纳米相材料科学中心，橡树岭，田纳西州，美国。4弗吉尼亚大学机械与航空航天工程系，夏洛茨维尔，弗吉尼亚州，美国。5华盛顿大学材料科学与工程系，西雅图，华盛顿州，美国。6范德堡大学电气与计算机工程系，纳什维尔，田纳西州，美国。7乌普萨拉大学物理与天文学系，乌普萨拉，瑞典。

*通讯作者。电子邮箱：harrison.a.walker@vanderbilt.edu；hoglunder@ornl.gov；贡献者：pfeifertw@ornl.gov；pzeiger@uw.edu；hachtelja@ornl.gov；pantelides@vanderbilt.edu；

摘要

电子显微镜中的振动光谱学能够以纳米空间分辨率揭示声子激发，但由于工作流程零散且需要专业知识，常规预测仍难以实现。在此，我们介绍 PySlice，这是首个公开发布的辅助波函数时间自相关（TACAW）方法实现，可提供自动化框架，直接从原子结构生成动量分辨和能量分辨的振动电子能量

损失谱。通过将通用机器学习原子间势与 TACAW 相结合，PySlice 消除了针对每个系统开发势函数这一瓶颈。用户输入原子结构，即可通过一个统一工作流程获得声子色散、谱衍射图样和谱像；该流程涵盖分子动力学、GPU 加速的电子散射以及频域分析。我们概述了代码背后的理论形式，展示其在材料科学中典型系统上的应用，并探讨其在高级分析和材料探索中的用途。模块化的 Python 架构还支持常规电子显微镜模拟，为成像和衍射计算提供通用平台。PySlice 让振动光谱预测成为常规操作，而非专业领域的专门技术，从而能够开展用于实验设计的计算筛选、系统探索不同材料类别中的声子物理，以及高通量生成用于训练未来机器学习模型的模拟数据。

关键词：振动电子能量损失谱、声子光谱学、多层切片模拟、通用原子间势、TACAW、扫描透射电子显微镜

1 引言

透射电子显微镜中的振动电子能量损失谱（电子能量损失谱）在各种声子测量技术中占据独特地位，可灵活调节空间分辨率与动量分辨率[1]。拉曼光谱和红外光谱学等光学方法，包括纳米傅里叶变换红外光谱和扫描近场光学显微镜等近场变体，虽可实现纳米尺度的空间分辨率，却仍仅限于测量布里渊区中心声子（$\mathbf{q} \approx 0$），这是因为振动能量下光子的动量可以忽略不计。非弹性中子散射和非弹性硬X射线散射可提供完整的动量分辨声子色散，但需要毫米级样品，且缺乏空间分辨率。

振动电子能量损失谱弥补了这一差距：通过改变探针会聚角，该技术可在纳米空间分辨率（聚焦束）与高精度动量分辨率（平行光束）之间进行权衡，并根据实验需求对不确定性关系 $\Delta x \cdot \Delta p \gtrsim \hbar/2$进行采样。这种可调性使研究范围从动量分辨声子色散扩展到界面声子、局域应变场和缺陷诱导振动态的空间分辨成像[2–9]。这类实验往往十分复杂，因此需要借助预测模拟来准确解读结果并指导实验设计。

有多种计算方法可用于处理振动电子能量损失谱模拟的不同方面。通用多层切片模拟代码（如 abTEM[10] 和 Dr. Probe[11] ）可提供 GPU 加速电子散射计算，并对像差和探测器几何结构进行精细处理。这些代码支持冻结声子方法[12] ，通过从适当的分布中抽样静态原子位移来近似热无序。它们能够捕捉菊池带和热漫散射背景等漫散射现象。另一方面，基于谐声子模式[2, 13, 14] 的微扰处理方法则为.....提供理论模型

非弹性散射截面，从而能够根据特定声子激发预测电子能量损失谱强度。Dwyer 提出的另一种方法[15] 通过特定模式的扰动势传播电子，在考虑动力学衍射的同时，仍将各个谐声子模式分别处理。

除上述谐处理之外，分子动力学（MD）模拟能够捕捉原子运动的完整非谐动力学。Fransson等人与Pfeifer等人所发展的方法[16, 17] 可直接从分子动力学轨迹中提取动力学结构因子，从而高效获得非谐声子色散，并可与理论截面公式相结合，以预测电子能量损失谱强度。频率分辨冻结声子多层法（FRFPMS）[18]与辅助波函数时间自相关（TACAW）[19] 将电子束传播与分子动力学轨迹相结合，在纳入探针几何与动力学衍射效应的同时，也能获取非谐动力学。

FRFPMS对频率滤波轨迹执行多层切片计算，因此计算成本随频率区间数量和轨迹时长同时增加。TACAW将频率分解从输入端移至输出端，从而简化了这一流程：分子动力学生成同时采样所有热占据模式的原子构型时间序列，多层切片传播得到随时间变化的出射波函数，而对波函数数据做一次时间傅里叶变换即可恢复完整频谱。

机器学习原子间势（MLIPs）的最新进展拓展了原子尺度模拟的应用范围。MLIPs 是基于密度泛函理论（DFT）计算数据库训练的神经网络，能够以接近 DFT 的精度预测能量和力，速度却快几个数量级。ORB[20], 、MACE[21], 、CHGNet[22], 、Allegro-FM[23], 和 UMA[24],等通用机器学习原子间势（uMLIPs）在涵盖元素周期表大部分元素的多样化材料数据集（例如 Materials Project[25], 、Alexandria[26], 、OMat24[27], 、OMol25[28]）上训练，可在不同化学空间中迁移预测原子间力，而无需针对每个系统重新训练。用户只需提供任意晶体结构，即可获得符合物理规律的动力学行为，无需开发专用势。与此同时，PyTorch[29] 等现代张量库支持 CPU 和 GPU 后端的硬件抽象以及高效的批量运算，简化了数值算法的实现，使 MLIPs 和下游多层切片计算都能利用 GPU 并行能力。

尽管这些方法各自解决了振动电子能量损失谱预测问题的一部分，但目前还没有任何框架能将它们整合为统一的自动化工作流程。冻结声子方法缺乏振动光谱学所需的频率分辨率。微扰理论方法无法捕捉非谐效应。现有多层切片代码能够提供高质量的散射模拟，但缺乏振动电子能量损失谱的集成流程。TACAW 提供了所需的物理理论，但实际应用需要协调多个专用软件包——一个分子动力学引擎、一套多层切片代码以及自定义分析脚本——而且没有公开可用的实现。这种碎片化使振动电子能量损失谱预测一直是一个需要大量定制开发的专业任务。通用机器学习原子间势、GPU 加速的多层切片法与 TACAW 形式体系的融合，如今

创造了将这些元素整合为通用预测流程的契机。

在此，我们介绍 PySlice，这是首个公开发布的 TACAW 方法实现。这个模块化的 Python框架整合了三项关键功能：(1) 为包括通用机器学习原子间势在内的现代原子间势提供统一接口，并通过 ASE[30]和 OVITO[31]实现自动化平衡检测和分子动力学轨迹管理；(2) 使用 PyTorch[29], 进行 GPU加速的多层切片计算，并通过持久化缓存支持模拟重启和后处理分析；(3) 为 TACAW 分析提供结构化数据模型，处理时频变换，并提供提取声子色散、动量分辨谱和谱像的方法。其面向对象架构清晰区分分子动力学、电子散射和光谱分析，同时支持灵活组合这些要素，以适应多种模拟工作流程。

我们通过将 PySlice 系统地应用于过渡金属二硫属化物（TMDCs），展示其功能：只需运行单个脚本，即可生成特定于材料的声子色散和谱衍射图样，涵盖 MoS_2, WS_2, WSe_2, 和 $MoSe_2$。我们还展示了体相 Si/Ge 异质结构界面和石墨烯中的硅取代缺陷所对应的振动谱像。除了 TACAW 模拟，PySlice 的模块化设计还支持常规多层切片成像和衍射模拟，为透射电子显微镜/扫描透射电子显微镜模拟任务提供统一平台。通过消除振动电子能量损失谱预测的实际障碍并建立一致的工作流程，本研究旨在加速复杂材料声子光谱学的计算探索和实验验证，同时促进学界的研究可重复性。

2 方法

2.1 工作流架构

模拟振动电子能量损失谱需要衔接三个不同的计算领域：原子动力学、电子散射和光谱分析。PySlice通过模块化流程（图1）解决这一问题，在清晰分离这些关注点的同时保持各阶段之间高效的数据流。

工作流首先通过分子动力学模拟生成热布居的原子构型。PySlice的MDCalculator并不再实现另一个分子动力学引擎，而是通过ASE的分子动力学引擎充当通向现代原子间势的统一接口，并特别侧重通用机器学习势ORB。MDCalculator通过温度范围和运行平均值等收敛判据自动处理平衡检测，确保轨迹在无需人工干预的情况下捕获热声子布居。对于静态成像模拟，分子动力学是可选的，轨迹将仅支持单帧。

这些随时间变化的原子构型会输入 MultisliceCalculator，后者必须为每个分子动力学时间步计算一次完整的多层切片模拟。对于每个探针位置都要模拟数百个时间步的典型振动电子能量损失谱模拟，

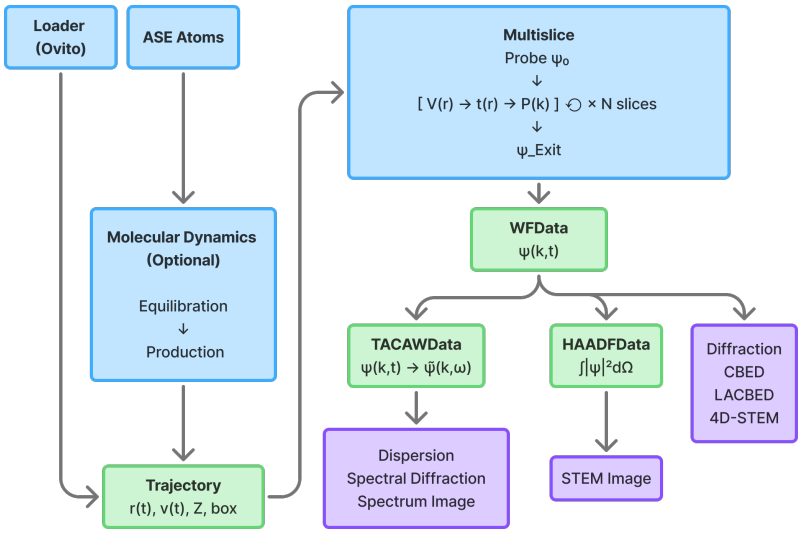


图 1 PySlice 自动化预测振动电子能量损失谱的工作流程。该流程包含四个阶段：(1) 输入结构，来源包括 CIF 文件、XYZ 文件、LAMMPS 轨迹或 ASE Atoms 对象；(2) 可选地使用通用机器学习原子间势（ORB）进行分子动力学模拟，并自动完成平衡（NVT）和生产阶段（NVE）；(3) 通过 GPU 加速的多层切片模拟，计算每个分子动力学时间步的出射波函数；(4) 通过 TACAW 分析将时域波函数转换为频域谱强度，或计算传统 TEM/STEM 信号，例如冻结声子高角环形暗场、透射电子显微镜衍射或四维扫描透射电子显微镜。数据容器（`Trajectory`、 `WFData`、 `TACAWData`、 `HAADFData`）负责管理各阶段之间的信息流。

这意味着需要投入大量计算资源。每个时间步都需要使电子波函数逐层切片地穿过整个样品厚度传播，而该快照中的原子位置则决定了散射势的分布。

该实现采用 PyTorch 进行 GPU 加速；在硬件支持的情况下，所有计算密集型操作均在 GPU 上执行。代码会根据可用硬件自动选择 CUDA、Apple Metal（MPS）或 CPU 后端。多层切片计算以帧为单位进行缓存：每个计算得到的出射波都会作为 NumPy 数组写入磁盘。这样既可在中断后恢复模拟，也可让后处理分析直接使用预先计算的波函数数据，而无需重新运行多层切片计算。虽然 GPU 加速是 PySlice 的核心功能，但即使未安装 PyTorch，代码仍保留完整功能，并可选择使用 NumPy [32] 作为后端。

随后，TACAWData类将时域波函数数据转换为频域谱。对于时间步长以 Δt为间隔的轨迹 $\{t_i\}$ ，在每个动量点 k 处对出射波函数 $\psi(\mathbf{k}\ t_i)$进行变换： $\psi(\mathbf{k}\ \omega) = \mathcal{F}_t[\psi(\mathbf{k}\ t) - \langle\psi(\mathbf{k}\ t)\rangle_t]$。减去平均值可分离出散射中随时间变化的部分，其中包含声子信息。随后，振动电子能量损失谱的强度 $I(\mathbf{k}\ \omega) = |\tilde{\psi}(\mathbf{k}\ \omega)|^2$ 会显示能量为 $\hbar\omega$ 的峰，对应声子激发

以及满足声子色散关系的动量 $\hbar$k。其背后的物理原理很直接：原子以各自特征声子频率振动时，会在电子散射势上产生周期性调制（同时涉及时间和空间）。这种调制会使透射电子波函数呈现相同的时间特征，并观测到向相应波矢的弹性散射。对时间进行傅里叶变换可提取这些频率分量，从而在动量和能量分辨强度 I(k ω)中显示声子模。PySlice将这类分析整合到一种数据结构中，并提供执行常见声子光谱学任务的方法：沿倒易空间路径提取色散、生成动量分辨谱，以及计算谱像。

2.2 关键技术要素

多层切片传播。多层切片算法通过将样品分解为垂直于电子束方向的薄片，求解电子散射问题。当切片厚度相对于横向相干长度足够小时，该近似成立；它使完整的三维散射问题能够作为一系列二维相移（透射）与自由空间传播交替进行的过程来处理。在每个厚度为 Δz 的切片内，波函数会因投影势而产生相移： t(r) $=$ exp($i\sigma V$(r))，其中 $\sigma = 2\pi m_e \lambda / h^2$ 是相互作用参数， m_e 包含相对论修正。在切片之间，倒易空间中的菲涅耳传播应用算符 P(k) $=$ exp($-i\pi\lambda\Delta z$|k|^2)。通过 N 个切片的完整传播过程如下： $\psi_{\text{exit}} = P_{N-1} t_N \cdots P_1 t_2 P_0 t_1 \psi_{\text{probe}}$。

原子散射势。准确的原子势对于定量电子能量损失谱模拟至关重要。PySlice 使用 Kirkland 参数化的原子结构因子[33], 这些因子经过数十年的改进，并与量子力学计算和实验测量进行了比对。对于每种元素 Z，其动量空间形式因子为 $f(\mathbf{q})$ $= \sum_{i=1}^{3} \left[\frac{a_i}{|\mathbf{q}|^2 + b_i} + c_i \exp(-d_i |\mathbf{q}|^2) \right]$，其中系数 $\{a_i, b_i, c_i, d_i\}$ 已列表给出。对于典型的透射电子显微镜条件，该参数化能够以足够的精度同时描述高角散射（第一项）和低角散射（第二项）。每个切片的投影势通过在倒易空间中对原子位置对应的结构因子乘以适当的相位因子后求和而构建： V(k) $= \sum_j f_{Z_j}$(k) exp($-2\pi i$k $\cdot$ r$_j$)，其中求和范围为切片 z 方向边界内的原子。随后进行逆快速傅里叶变换，得到实空间投影势。

分子动力学。TACAW方法通过对含时电子波函数进行傅里叶变换来提取声子频率，该波函数的时间结构继承自原子运动。这对分子动力学轨迹提出两项要求：构型必须在目标温度下采样热声子布居，且时间演化必须反映确定性的声子动力学，而非恒温器驱动的涨落。PySlice的MDCalculator通过两阶段方案满足这些要求：朗之万动力学（恒温恒容系综）将系统驱动至热平衡，随后以速度-Verlet积分（NVE）生成生产轨迹。或者，生产阶段也可采用摩擦系数极低的朗之万动力学，这在某些机器学习原子间势下能更好地保持能量守恒，同时

对声子动力学的扰动最小。平衡不能仅凭瞬时温度达到目标值来推断，因为系统可能在仍向热平衡弛豫的过程中短暂经过该值。MDCalculator转而依据在滑动窗口上评估的三项判据来判断收敛。首先，瞬时温度必须至少达到过一次目标值，以确认热激发充分。其次，近期步数内的平均温度必须保持在目标值的容差范围内，且其标准差低于稳定性阈值，从而确保系统围绕正确的热状态涨落，而非发生漂移。第三，每原子势能的相对标准差必须低于某一阈值，表明系统已进入构型空间中的稳定区域，而非仍在进行结构弛豫。只有当这三项条件同时满足时，模拟才会进入生产阶段，从而将传统上需要人工检查热化曲线的过程自动化。虽然默认参数适用于大多数材料，但用户可完全控制收敛判据、时间步长和恒温器摩擦系数，并可根据具体任务进行调整。这些参数的选取指南见补充材料。分子动力学运行所得的构型与速度会自动打包为PySlice的Trajectory格式，并附带正确的时间步长元数据，从而无缝交接至多层切片计算阶段。

通过对多个独立轨迹或轨迹片段的结果取平均，可以进一步改善谱统计，但这会使多层切片法计算的计算成本成倍增加；补充材料中提供了有关此方法的指导。

数据结构设计。 该软件架构通过四个核心类将数据存储与分析分离开来。`Trajectory`类封装了原子构型，并提供访问位置、速度和派生量（平均位置、位移）的方法，而不会暴露底层存储格式。 `WFData`类将多层切片法结果存储为复数波函数，并按照探针位置、时间和倒易空间坐标建立索引。这种组织方式与TACAW计算的自然坐标轴相匹配，并支持高效的切片操作以进行分析。 `TACAWData` 类继承自 `WFData`，但将时间轴替换为频率，并提供用于声子分析的专用方法：`dispersion()` 用于提取沿k空间路径的色散关系，`spectraldiffraction()` 用于获取特定频率下按动量分辨的强度，以及 `spectrum image()`用于在扫描透射电子显微镜几何中进行位置分辨的振动映射。使用 `WFData` 类还可以执行常规透射电子显微镜和四维扫描透射电子显微镜分析，例如用于成像和衍射的实空间与倒易空间绘图。还提供了 `HAADFData` 类，以简化对 x、y、k_x、k_y 中所含数据集的后处理，这些数据集存储于 `WFData`中，例如用于常规明场或环形暗场扫描探针成像。

3 结果

3.1 过渡金属二硫属化物的自动化工作流程

为验证自动化工作流，我们将 PySlice 应用于四种单层过渡金属二硫属化物：MoS_2，WS_2，$MoSe_2$，和 WSe_2 （图2）。这四种材料

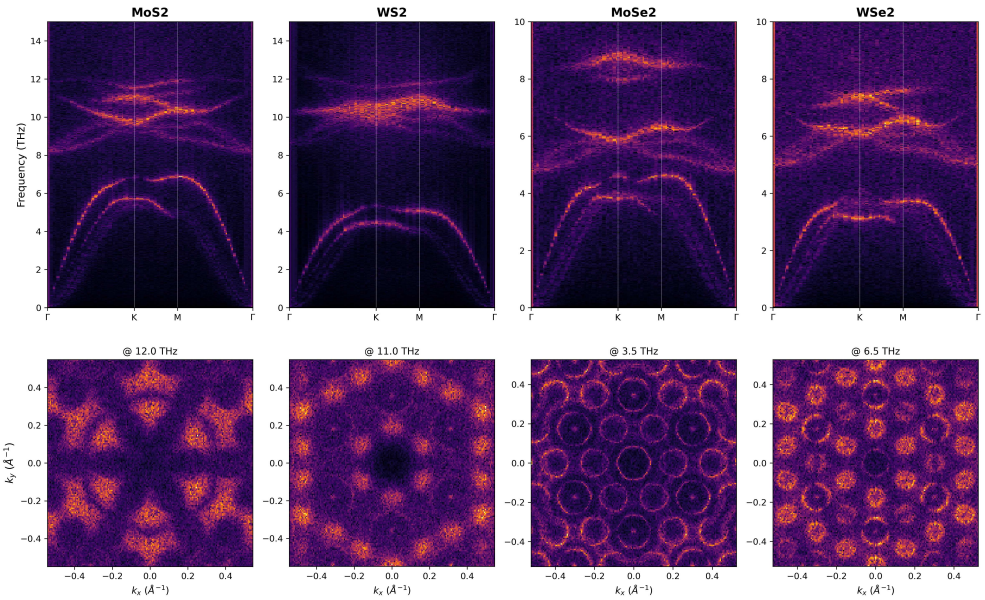


图 2 过渡金属二硫属化物家族的端到端自动化 TACAW 光谱。所有模拟均由单个 PySlice 脚本生成，该脚本负责完整工作流程：使用 ASE 生成晶体结构、通过 ORB 驱动分子动力学并自动进行平衡、通过 GPU 加速的多层切片传播运行数百个时间步长，以及进行 TACAW 频域分析。在不同材料之间，无需操作中间文件、转换格式或手动调整参数。所得的声子色散和光谱衍射图样展示了 MoS_2, WS_2,WSe_2 and $MoSe_2$ 中由原子质量和成键差异所引起的材料特异性振动特征。

均通过同一个脚本处理，并采用相同的模拟参数，逐一遍历各化学式。对于每种材料，工作流都从晶体结构生成开始，自动完成 ORB 驱动的分子动力学、跨分子动力学帧的多层切片传播，以及 TACAW 频域分析。

所得的声子色散（图2，上排）呈现出预期的材料特征：声学支从 Γ 点向上延伸，较高频率处出现光学支，并在布里渊区边界附近呈现典型的平坦化。硫化物（MoS_2, WS_2）的光学模延伸至 ∼14 THz，而较重的硒化物（$MoSe_2$,WSe_2）则呈现相应较低的频率，最高达到 ∼10THz。代表性频率下的光谱衍射图样（图2，下排）显示出这些 2H 相材料所具有的预期六方对称性。这些结果表明，单一自动化流程能够为化学组成相关的一系列材料生成符合物理规律的特定材料声子谱。这些预测与实验测量结果的一致程度取决于所采用的原子间势，我们将在第4节讨论这一点。

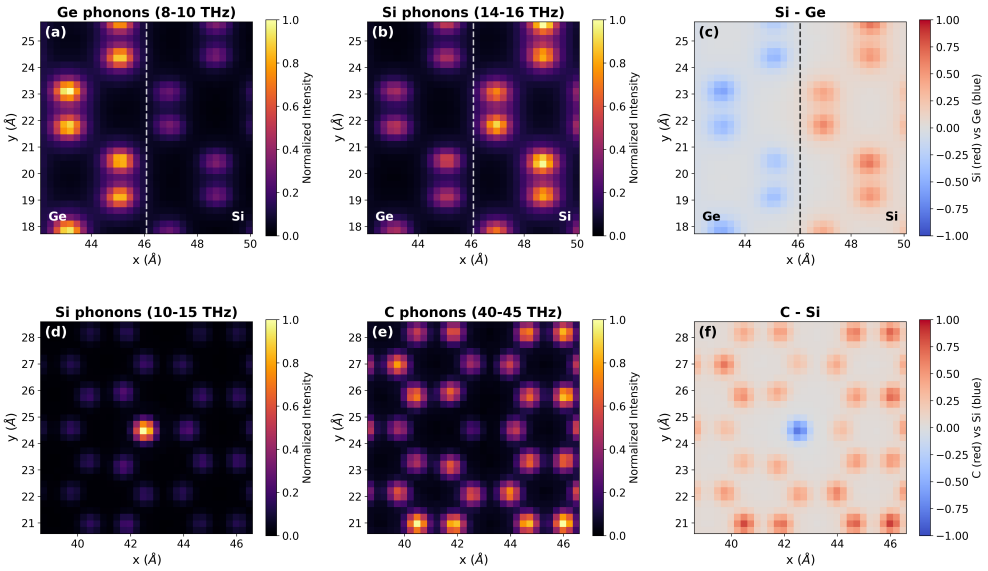


图3 局域振动谱像。(a-c) 展示光学声子约束的体相 Si/Ge 异质结构界面：(a) 局限于 Ge 区域的 Ge 光学模（∼8-10 THz），(b) 局限于 Si 区域的 Si 光学模（∼14-16 THz），以及 (c) 两个能量区域的归一化差值。(d-f) 展示石墨烯中的硅取代缺陷及其局域缺陷模：(d) 硅振动（10-15 THz）局限于孤立原子，而 (e) 碳振动（40-45 THz）则遍布整个结构。(f) 中的差值图突出显示了硅杂质。这些结果展现了 PySlice 预测界面和点缺陷处空间分辨振动信号的能力。

3.2 空间分辨声子模

PySlice 的 STEM 扫描功能支持空间分辨的声子光谱学，可揭示局域振动模。我们以两个典型体系为例进行展示：Si/Ge 半导体界面，以及石墨烯中的硅取代（图3）。

Si/Ge 界面体现了异质结构边界处的声子约束。由于 Si（28 amu）和 Ge（73 amu）之间存在质量差异，这两种材料的光学声子频率有显著差异。TACAW 频谱成像显示，Ge 光学模（图3a）局限于 Ge 区域，而 Si 光学模（图3b）则局限于界面的 Si 侧。归一化差值（图3c）突显了这种空间分离。声学模的波长远大于界面宽度，因此能够在两种材料中传播；在谱像中，它们不会局限于任一区域。

石墨烯中的硅取代展示了点缺陷振动的检测。硅（28 amu）与碳（12 amu）之间显著的质量差异产生了鲜明的振动特征：石墨烯晶格中的碳原子以约 40–45 THz 的特征频率振动，而较重的硅杂质则呈现 10–15 THz 范围内的低频振动模式。在这些频率下进行 TACAW 频谱成像，可在均匀的

石墨烯背景中显示出硅原子这一局域亮斑，证明 PySlice 对孤立的原子尺度缺陷具有灵敏的检测能力。

4 讨论

4.1 对材料探索的启示

通用机器学习原子间势使基于分子动力学的振动电子能量损失谱预测变得广泛可及。过去，研究人员需要针对特定材料的原子间势；如今，通用机器学习原子间势可在整个化学空间中提供可迁移的力。PySlice 借此实现自动化光谱预测。

这一能力为材料探索开辟了多种途径。高通量筛选因而成为可能：研究人员可以系统地计算不同材料类别、合金成分或缺陷构型的振动光谱，从而找出具有目标声子性质的体系。自动化工作流程还支持预测驱动的实验设计，即在占用显微镜机时之前，先利用模拟光谱指导材料和测量条件的选择。此外，生成大量模拟光谱数据集的能力，将为未来用于分析实验电子能量损失谱数据的机器学习模型提供训练数据。

预测的声子频率和强度的准确性取决于底层原子间势；PySlice 会忠实传递通用机器学习原子间势所编码的物理规律，但无法弥补其不足。近期针对通用机器学习原子间势开展的系统基准评测，将其与 DFT声子数据库及实验非弹性中子散射光谱进行比较，结果表明，ORB 等模型在多种材料类别中都能实现接近 DFT 的振动性质预测精度[34]。PySlice 利用这一经验证的准确性，实现自动化振动电子能量损失谱预测。随着通用机器学习原子间势不断改进，该工作流程的预测结果也将相应提升，而无需更改模拟基础设施。

计算参数也会限制可访问的光谱范围。超胞尺寸决定可分辨的最小波矢，而分子动力学时间步则通过奈奎斯特准则决定最高频率。记录帧的总数决定频率分辨率。这些是任何时域声子模拟中常见的权衡，用户可以调整超胞尺寸和轨迹长度，以匹配感兴趣的光谱特征。

4.2 超越振动光谱学

PySlice 基于 TACAW 架构构建，但同一套多层切片法机制也可自然地用于常规扫描透射电子显微镜模拟（图4）。将两者整合在同一框架中十分便利，用户无需切换工具，即可验证探针配置、生成对比图像并进行完整的振动分析。

对于扫描透射电子显微镜成像，HAADFData类会在可配置的收集角范围内对散射强度进行积分，直接从多层切片出射波函数生成环形暗场图像。高角度配置（内角为 50–80 mrad）可生成对原子序数敏感的 Z衬度图像，而调整探测器几何结构则可探测不同的散射机制。底层四维数据集，即衍射

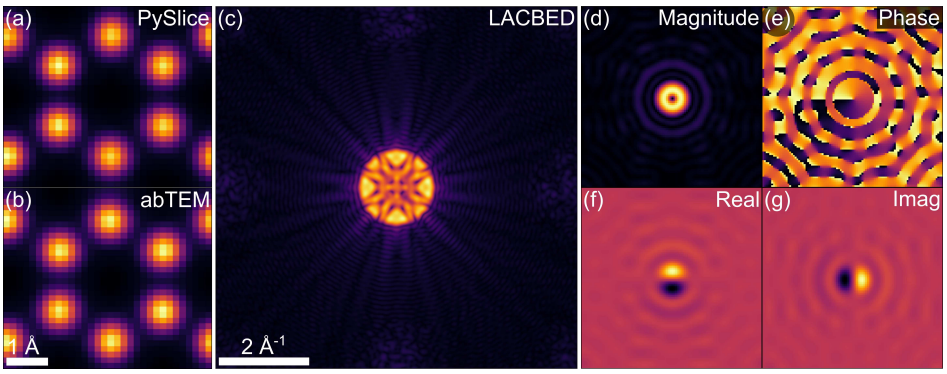


图 4 振动光谱学之外的扩展模拟能力。(a) 环形暗场（ADF）成像与成熟的多层切片代码进行验证。(b) 大角度会聚束电子衍射（LACBED）图样，展示了与厚度相关的动力学衍射效应。(c) 具有轨道角动量的涡旋束探针，展示了 PySlice 对任意波前分布的支持，包括手性敏感照明模式。

图样仍可供更复杂的四维扫描透射电子显微镜分析使用，这些图样是在每个探针位置记录的。

真实的模拟通常需要对透镜像差进行建模。PySlice 实现了标准的像差函数，支持离焦、像散、彗差、球差及更高阶项。像差既可在探针形成过程中施加，也可在事后施加到已存储的波函数上，从而无需重复进行多层切片计算，即可高效探索穿焦序列或像差校正策略。

多层切片传播可以选择保留每个切片处的波函数，从而提供与会聚束电子衍射分析相关的深度分辨散射信息。Probe 类也接受任意复波前数组，因此用户可以构建自定义相位分布，例如用于涡旋束的螺旋相位，并通过同一套基础设施传播这些波前。这种灵活性使 PySlice 适用于振动光谱学这一主要应用之外的一般 TEM/STEM 模拟任务。

5 结论

我们介绍了 PySlice，这是一个可根据原子结构预测振动电子能量损失谱的自动化框架。通过整合通用机器学习原子间势、GPU 加速的多层切片模拟和 TACAW 频域分析，PySlice 将振动光谱预测的门槛从数周的专业开发工作降至数小时的计算时间。我们在过渡金属二硫属化物家族中验证了这一方法，证明只需运行一个自动化脚本，即可获得具有材料特异性的声子色散和谱衍射图样。我们还展示了异质结构界面和点缺陷处的振动谱像，这些功能可用于复杂材料中的纳米尺度声子映射。PySlice 为高通量光谱预测、训练数据生成以及预测驱动的振动电子能量损失谱实验奠定了基础

设计。补充材料和代码仓库中提供了示例脚本和完整工作流程。

致谢。H.A.W. 和 S.T.P. 获得了美国能源部基础能源科学办公室（DOE-BES）材料科学与工程部项目（编号 DE-FG02-09ER46554）、美国能源部科学办公室用户设施以及麦克明捐赠基金的支持。

H.A.W. 和 S.T.P. 感谢国家能源研究科学计算中心（NERSC）提供计算资源。NERSC 是美国能源部科学办公室的用户设施，位于劳伦斯伯克利国家实验室，由该实验室根据合同编号 DE-AC02-05CH11231 运营。

本研究得到美国国家科学基金会研究生研究奖学金计划的支持（资助编号 2444112）。本文所表达的任何观点、发现、结论或建议均为作者个人观点，不一定反映美国国家科学基金会的看法。

T.W.P. 和 E.R.H. 感谢橡树岭国家实验室材料表征核心设施提供支持。该设施由 UT-Battelle有限责任公司根据与美国能源部签订的合同（编号 DE-AC05-00OR22725）运营，并由橡树岭国家实验室实验室定向研究与开发计划资助；该计划由 UT-Battelle有限责任公司代表美国能源部管理。

本研究使用了橡树岭国家实验室的科学计算与数据环境（CADES）资源。该环境由美国能源部科学办公室根据合同编号 DE-AC05-00OR22725 提供支持。

P.M.Z. 获得瑞典研究委员会 2024-06617 号项目资助。

声明：本手稿由 UT-Battelle有限责任公司根据与美国能源部（DOE）签订的合同 DE-AC05-00OR22725 撰写。美国政府保留相关权利；出版商接受本文发表即承认，美国政府拥有非独占、已付清、不可撤销、全球范围的许可，可为美国政府目的发表或复制本手稿的已发表版本，或授权他人这样做。DOE 将依据美国能源部公共获取计划（https://www.energy.gov/doe-public-access-plan）为公众提供这些联邦资助研究成果的获取途径。

声明

- 资助：参见致谢。
- 利益冲突：作者声明不存在任何竞争性利益。
- 数据可用性：本文所展示的所有数据均使用 PySlice 生成，可通过代码仓库中提供的脚本复现。

- 代码可用性：PySlice 可在以下地址获取：https://github.com/h-walk/PySlice。
- 作者贡献：H.A.W. 撰写了手稿。H.A.W.、T.W.P. 和 P.M.Z. 开发了 PySlice 软件。J.A.H.、S.T.P. 和 E.R.H. 提供了指导。所有作者均审阅并编辑了手稿。

参考文献

[1] Krivanek, O. L. 等. 电子显微镜中的振动光谱学。《自然》514, 209–212 (2014).[2] Senga, R. 等. 石墨烯纳米结构中振动的位置与动量映射。《自然》573, 247–250 (2019).[3] Hage, F. S., Radtke, G., Kepaptsoglou, D. M., Lazzeri, M. & Ramasse, Q. M. 扫描透射电子显微镜中的单原子振动光谱学。《科学》367, 1124–1127 (2020).[4]Hoglund, E. R. 等. 氧化物超晶格的涌现界面振动结构。《自然》601, 556–561 (2022).[5] Haas, B. 等. 晶界处局域声子模式的原子分辨率映射。《纳米快报》23, 5975–5980 (2023).[6] Hoglund, E. R. 等. 复杂晶界处局域振动的直接可视化。《先进材料》35, 2208920 (2023).[7]Xu, M. 等. 具有化学键敏感性的单原子振动光谱学。《自然·材料》22, 612–618 (2023).[8] Hoglund, E. R. 等. 极性超晶格界面处的非等价原子振动。《先进材料》36, 2402925 (2024).[9] Yan, J. 等. Al_2O_3 晶界处的纳米尺度局域声子。《纳米快报》24, 3323–3330 (2024).[10]Madsen, J. & Susi, T. abTEM：基于第一性原理的透射电子显微术。Open Research Europe 1, 24 (2021).[11] Barthel, J. Dr. Probe：用于高分辨率扫描透射电子显微镜图像模拟的软件。《超微显微学》193, 1–11 (2018).[12] Loane, R., Xu, P. & Silcox, J. 汇聚束电子衍射中的热振动。《晶体学报》A辑47, 267–278 (1991).[13] Nicholls, R. 等. 电子显微镜中动量分辨声子光谱学的理论。《物理评论B》99, 094105 (2019).[14]Rossi, A. W., Bourgeois, M. R., Walton, C. & Masiello, D. J. 利用动量分辨电子能量损失谱学, 探测二维材料从原子晶体到纳米光子阵列中的低能激发极化。《纳米快报》24, 7748–7756 (2024).

[15] Dwyer, C. 原子振动导致的高能电子散射局域化。《物理评论B》89, 054103 (2014).[16] Fransson, E., Slabanja, M., Erhart, P. & Wahnstr¨om, G. dynasor——一种从分子动力学模拟中提取动力学结构因子和电流关联函数的工具。《先进理论与模拟》4, 2000240 (2021).[17]Pfeifer, T. W. et al. 理解动量分辨电子能量损失谱学中的声子选择与干涉。arXiv 预印本 arXiv:2503.09792 (2025).[18] Zeiger, P. M. & Rusz, J. 用于振动扫描透射电子显微镜-电子能量损失谱的高效且通用模型。《物理评论快报》124, 025501 (2020).[19] Castellanos-Reyes, J. A., Zeiger, P. M. & Rusz, J. 透射电子显微术中声子和磁振子的角分辨电子能量损失与增益谱学的动力学理论（纳入多重散射效应）。《物理评论快报》134, 036402 (2025).[20] Neumann, M. et al. Orb：一种快速、可扩展的神经网络势 (2024). 2410. 22570.[21] Batatia, I., Kov´acs, D. P., Simm, G. N., Ortner, C. & Cs´anyi, G. MACE：用于快速、精确力场的高阶等变消息传递神经网络。《神经信息处理系统进展》35, 11423–11436 (2022).[22] Deng, B. et al. CHGNet：一种用于电荷信息驱动的原子建模的预训练通用神经网络势。《自然·机器智能》5, 1031–1041 (2023).[23] Nomura, K.-i. et al. Allegro-FM：面向百亿亿次级分子动力学模拟的等变基础模型。《物理化学快报》16, 6637–6644 (2025).[24] Wood, B. M. et al. UMA：一组原子通用模型 (2025). 2506. 23971.[25] Jain, A. et al. 材料项目：加速材料创新的材料基因组方法。《APL材料》1, 011002 (2013).[26] Schmidt, J., Wang, H.-C., Cerqueira, T. F. T., Botti, S. & Marques, M. A. L. 使用 PBEsol 和 SCAN 泛函计算的 17.5 万种稳定与亚稳材料数据集。《科学数据》9, 64 (2022).

[27]Barroso-Luque, L. et al. 开放材料 2024（OMat24）无机材料数据集与模型 (2024). 2410.12771.[28] Levine, D. S. et al. 开放分子 2025（OMol25）数据集、评估与模型 (2025). 2505.08762.[29] Paszke, A. et al. PyTorch：一种命令式风格、高性能深度学习库。《神经信息处理系统进展》32, 8024–8035 (2019).[30] Larsen, A. H. et al. 原子模拟环境——一个用于处理原子的 Python 库。《物理学杂志：凝聚态》29, 273002 (2017).[31] Stukowski, A. 使用 OVITO——开放可视化工具对原子模拟数据进行可视化与分析。《材料科学与工程中的建模与模拟》18, 015012 (2010).[32] Harris, C. R. et al. 使用 NumPy 进行数组编程。《自然》585, 357–362 (2020).[33] Kirkland, E. J. 《电子显微镜中的高级计算》第3版（施普林格，2020）。[34]Han, B. & Cheng, Y. 对通用机器学习原子间势进行基准测试，以快速分析非弹性中子散射数据。《机器学习：科学与技术》6, 030504 (2025).

补充材料：PySlice——利用通用原子间势进行常规振动电子能量损失谱预测

Harrison A. Walker[1, 2*], Thomas W. Pfeifer[3,4], Paul M. Zeiger[5,7],Jordan A. Hachtel[3], SokratesT. Pantelides[1,2,6], Eric R. Hoglund[3*]1范德堡大学跨学科材料科学项目，美国田纳西州纳什维尔。[2]范德堡大学物理与天文学系，美国田纳西州纳什维尔。[3]橡树岭国家实验室纳米相材料科学中心，美国田纳西州橡树岭。[4]弗吉尼亚大学机械与航空航天工程系，美国弗吉尼亚州夏洛茨维尔。[5]华盛顿大学材料科学与工程系，美国华盛顿州西雅图。[6]范德堡大学电气与计算机工程系，美国田纳西州纳什维尔。[7]乌普萨拉大学物理与天文学系，瑞典乌普萨拉。

*通讯作者。电子邮箱：harrison.a.walker@vanderbilt.edu；hoglunder@ornl.gov；贡献作者：pfeifertw@ornl.gov；pzeiger@uw.edu；hachtelja@ornl.gov；pantelides@vanderbilt.edu；

1 补充材料

自动分子动力学平衡

PySlice 实现了一种自动化的两阶段分子动力学方案：先进行 NVT 平衡，再进行 NVE（或低摩擦NVT）生产阶段。平衡阶段

使用具有适度摩擦系数的朗之万恒温器，使系统快速达到热平衡。检测到会聚后，模拟便切换至生产阶段动力学，并将恒温器干涉降至最低，从而确保声子动力学清晰，适用于 TACAW 分析。

平衡过程中的会聚情况通过在滑动窗口上评估多个标准来判断。算法要求以下所有条件同时满足：

1. 温度稳定性：最近 N_{temp} 步的温度标准差必须低于阈值：

$$\sigma_T = \sqrt{\langle T^2 \rangle - \langle T \rangle^2} < \sigma_T^{\max} \tag{1}$$

2. 温度达标：平均温度必须处于目标温度的容差范围内：

$$|\langle T \rangle - T_{\mathrm{target}}| < \Delta T_{\mathrm{tol}} \tag{2}$$

3. 能量稳定性：势能的相对标准差必须低于阈值：

$$\frac{\sigma_E}{|\langle E \rangle|} < \epsilon_E \tag{3}$$

该算法的一个关键特性是已达到目标温度标志：在宣布会聚之前，系统必须至少有一次达到目标温度。这可以防止在初始加热阶段出现虚假会聚，此时温度和能量可能看似稳定，但数值并不正确。最小步数要求可确保在开始会聚检查之前累积足够的统计数据。

表S1列出了默认会聚参数及其说明。

表 S1 分子动力学平衡收敛检查的默认参数。

参数	默认值	说明
温度 阈值	5.0 K	最大允许 σ_T
温度容差	20.0 K	最大 $\lvert\langle T\rangle - T_{\mathrm{target}}\rvert$
能量阈值	0.01	最大相对 $\sigma_E/\lvert\langle E\rangle\rvert$
温度窗口	50	温度平均窗口
能量窗口	50	能量平均窗口
最小平衡步数	100	检查前的最小步数
最大平衡步数	20000	强制转变

这些默认设置有意保持宽松，以确保不同材料均能稳健收敛，无需用户干预。对于需要精确温度控制的定量研究，用户应考虑收紧 温度容差（例如设为 10 K），并增加平均窗口，以更好地捕捉慢弛豫动力学。

图 S1展示了 MoS_2 在 300 K 下具有代表性的平衡轨迹。系统从初始构型开始平衡时，温度最初会出现波动，随后稳定在目标值附近。一旦满足所有收敛判据，模拟便会自动进入生产阶段。

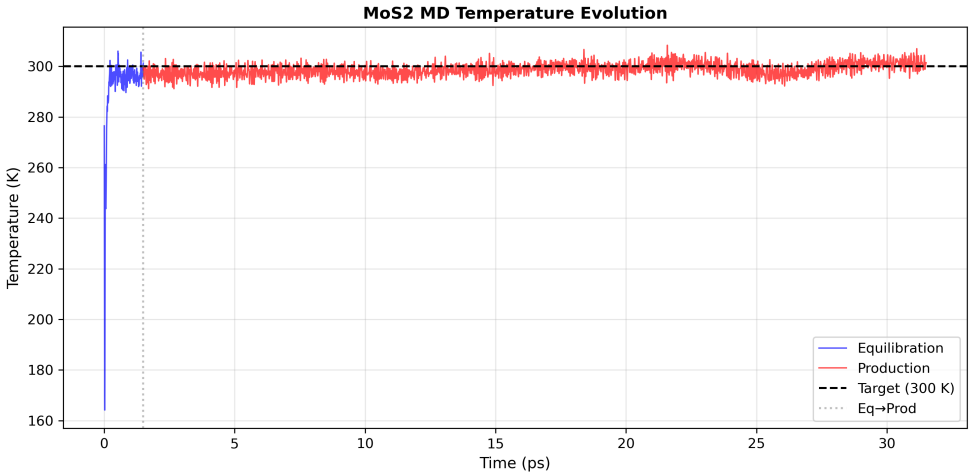


图 S1 使用 ORB 势对 MoS_2 进行分子动力学平衡。平衡期间的温度变化表明，系统达到热平衡后，满足收敛判据，并开始生产动力学。

采样注意事项

若干采样参数决定了 TACAW 分析中可获取声子频率的分辨率和范围：

超胞尺寸决定了可分辨的最小波矢。对于沿给定方向尺寸为 L 的模拟晶胞：

$$\Delta k = \frac{1}{L} \tag{4}$$

较大的超胞可对布里渊区进行更精细的采样，但会增加计算成本。

轨迹长度决定频率分辨率。对于间隔时间为 Δt 的 N 帧：

$$\Delta f = \frac{1}{N \cdot \Delta t} \tag{5}$$

更长的轨迹可提供更高的频率分辨率，但需要更多存储空间和计算资源。

采样时间步长决定了可分辨最高频率的奈奎斯特极限：

$$f_{\max} = \frac{1}{2\Delta t_{\text{sample}}} \tag{6}$$

请注意，TACAW 分析所用的采样间隔时间步长可以不同于分子动力学积分时间步长——为保证数值稳定性，积分可采用 1 fs，同时以 10 fs 的间隔保存帧。

轨迹平均可以提高信噪比。以不同随机速度初始化的多条独立分子动力学轨迹会采样相空间的不同区域；对其 TACAW 光谱取平均可降低统计噪声，同时保留底层的声子结构。

图S2展示了采样不足所带来的后果。由单条短轨迹计算得到的石墨烯中硅取代的谱像，在等效碳位点处呈现强度变化，反映出对声子相空间的遍历采样不完整。红圈标出硅原子的位置。延长轨迹或对多次独立运行取平均，可改善谱统计的收敛。

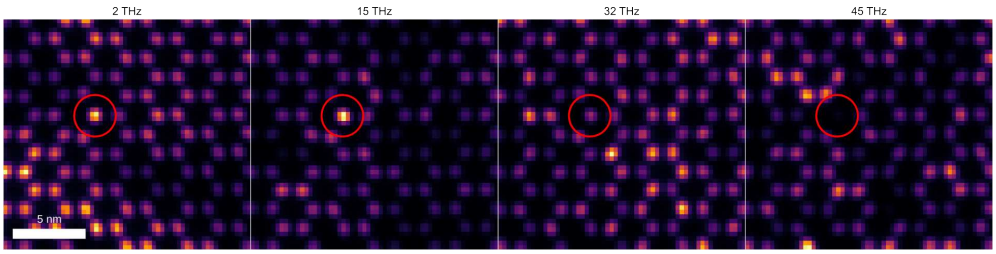


图 S2 由单条轨迹计算得到的石墨烯中硅取代缺陷在代表性频率（2、15、32 和 45 THz）下的谱像。红圈标出硅原子。等效碳位点之间的强度变化表明，轨迹长度不足会导致非遍历采样伪影。对多条独立轨迹取平均或延长模拟时长可改善收敛。

物理近似与局限性

PySlice 中实现的多层切片算法用于计算电子在随时间变化的样品势中的弹性散射。该方法本身包含若干近似：

弹性散射形式。每个切片的透射函数都会根据瞬时静电势施加相移：

$$T(\mathbf{r}_\perp) = \exp[i\sigma V(\mathbf{r}_\perp)] \tag{7}$$

不会计算电子与声子之间的显式能量转移。相反，会从衍射强度的时间调制中提取声子频率

随着原子沿分子动力学轨迹运动。由于声子能量（毫电子伏特量级）相对于电子动能（100 千电子伏特）可以忽略不计，因此这种方法是有效的，能量转移实际上不会改变电子的波长和散射几何。

电子-样品动力学解耦。原子轨迹的计算独立于电子束。模拟假设：

- 无电子诱导加热或辐射损伤
- 探测电子不会扰动原子运动
- 电子对热声子布居进行采样，但不会影响它

这种近似适用于振动电子能量损失谱测量中典型的低剂量条件，在这些条件下，实验上会尽量减少束流诱导效应。

扩展至其他原子尺度动力学

PySlice 接受任何提供原子位置随时间变化的轨迹。该框架并不局限于热分子动力学——任何能够产生时间分辨位置的原子尺度动力学方法均可与 TACAW 分析流程配合使用。潜在应用包括：

原子尺度自旋动力学：用于研究自旋和晶格自由度相互耦合的材料中的磁振子-声子耦合与磁激发。

路径积分分子动力学：用于在低温下纳入量子核效应，此时零点运动和隧穿变得显著。

驱动动力学：激光激发或外加场等非平衡模拟可揭示瞬态声子群。

关键要求是轨迹需提供：(i) 形状为 $(N_{\text{frames}}, N_{\text{atoms}}, 3)$ 的位置数组，(ii) 原子种类识别，以及 (iii) 模拟晶胞几何。

最小工作示例

以下 Python 脚本展示了完整的 TACAW 工作流程：生成晶体结构、运行分子动力学、计算多层切片法的出射波函数，并提取声子色散。本示例计算体硅沿 $\Gamma \to X$ 方向的声子色散。

```
from ase.build import bulk
from pyslice import MDCalculator, \
MultisliceCalculator, TACAWData
import numpy as np

# 1. Create structure: 20x20x5 Si supercell
atoms = bulk("Si", "diamond", a=5.431, cubic=True)
atoms = atoms * (20, 20, 5)

# 2. Run MD with automatic equilibration
md = MDCalculator(
model_name="orb-v3-conservative-inf-omat")
md.setup(atoms=atoms,temperature=300,
           # Ktimestep=2.0,              # fs
```

```
    production_steps=1000,       # 2 ps trajectory
save_interval=5)
trajectory = md.run()        # Save every 10 fs)
# 200 frames at 0.01 ps spacing -> Delta_f = 0.5 THz


# f_max = 50 THz (Nyquist limit)
# 3. Run multislice (parallel beam for TACAW)

calc = MultisliceCalculator()calc.setup(trajectory,
aperture=0,                   # Parallel beam
voltage_eV=100e3             # 100 keV , ,
sampling=0.1                 # Angstrom/pixel
slice_thickness=0.5          # Angstrom       , )


wf_data = calc.run()

# 4. Compute TACAW spectrum
tacaw = TACAWData(wf_data)

# 5. Extract dispersion along Gamma -> X
a = 5.431
kx = np.linspace(0, 1/(2*a), 100)  # Gamma to X
ky = np.zeros_like(kx)
dispersion = tacaw.dispersion(kx, ky)

# 6. Plot spectral diffraction at 15 THz
import matplotlib.pyplot as plt
Z = tacaw.spectral_diffraction(15.0)
plt.imshow(Z**0.25, cmap="inferno")
plt.savefig("spectral_15THz.png")
```

此示例的分辨率分析： $20 \times 20 \times 5$ 超胞的尺寸为 a =5.431 °A，实空间范围为 $L = 108.6$ °A，倒易空间分辨率为 Δk =$1/L \approx 0.009$ °A^{-1}。采样参数（0.1 °A）决定最大波矢：$k_{\max} = 5$ °A^{-1}，远高于硅布里渊区边界的 $k_X \approx 0.092$ °A^{-1}。约有 10 个 k 点对 $\Gamma \to X$ 路径进行采样。包含 200 帧、帧间隔为 0.01 ps 的轨迹提供了 $\Delta f = 0.5$ THz 的频率分辨率和 $f_{\max} = 50$ THz 的奈奎斯特极限，足以涵盖硅光学声子（∼15T Hz）。

对于扫描透射电子显微镜谱成像，请将探针位置添加到 MultisliceCalculator 中：

```
# Define probe scan region
xs = np.linspace(0, 20, 32)
ys = np.linspace(0, 20, 32)

calc.setup(trajectory,aperture=30,
               # 30 mrad convergencevoltage_eV=100e3,
sampling=0.1     ,probe_xs=xs,
               # Scan gridprobe_ys=ys,)
wf_data = calc.run()tacaw = TACAWData(wf_data)




# Extract spectrum image at specific frequency
spectrum_image = tacaw.spectrum_image(15.0)
```

加载已有轨迹（LAMMPS、ASE、XYZ）时：

```
from pyslice import Loader

# LAMMPS dump file
```

```
traj = Loader(
    "dump.lammpstrj",
    timestep=0.01,              # ps
    atom_mapping={1: "Si", 2: "Ge"}
).load()

# ASE trajectory file
traj = Loader("production.traj", timestep=0.01).load()
```

更多示例涵盖 HAADF成像、像差、LACBED花样和涡旋束，可在代码库的 tests/ 目录中找到。

代码可用性

所有代码，包括示例和测试脚本，均可在 https://github.com/h-walk/PySlice。