arXiv:2303.13435v1 [cond-mat.mtrl-sci] 23 Mar 2023

“*Ultima Ratio*”：模拟宽范围X射线散射与衍射

Brian R. Pauw、Sofya Laskina、Aakash Naik、Glen J. Smales、Janine George、Ingo Breßler、Philipp Benner
德国联邦材料研究与测试研究所，Unter den Eichen 87，12205 柏林，德国

电子邮箱：brian.pauw@bam.de

2023年3月

摘要。我们展示了一种模拟宽范围X射线散射图样的策略，其涵盖小角和大角散射，也包括对分布函数（PDF）分析通常采用的散射角。这类模拟图样可用于测试整体分析模型；此外，由于衍射强度与散射强度处于同一尺度，它们还可能为测定结晶度提供一条新途径。

“Ultima Ratio”策略以一个 64 纳米的金属有机框架（MOF）颗粒为例进行演示，计算范围从 $Q < 0.01\mathrm{nm}^{-1}$ 到 $Q \approx 150\mathrm{nm}^{-1}$，分辨率为 0.16 $^{\circ}$A。计算采用了改进的三维快速傅里叶变换（3D-FFT），其改进使得能够对尺寸至少达到 8000^3 个体素的矩阵进行变换。将多个此类改进的三维快速傅里叶变换结合起来，以改善低-Q行为。所得曲线与在多分散金属有机框架粉末上测得的宽范围散射图样进行了比较。

尽管计算量很大，但该方法有望用于模拟各种真实复杂结构的散射，从（多）晶体颗粒到病毒和催化剂等分层多组分结构。

1. 引言

功能材料的性质源于层级结构之间复杂的相互作用，这些结构从原子尺度一直延伸至微米尺度，乃至更大尺度。这类层级结构有时适合采用现代 X 射线和中子散射方法进行研究。如今，这些方法默认可测量跨越多个数量级的散射矢量（Q），因此能够携带多个数量级结构长度尺度上的信息。如此宽广的测量范围促使人们构建能够解读这一扩展视野的分析模型。这些分析模型的输出应与基于充分了解的结构生成的模拟数据进行检验或比较，但在实践中，这类模拟往往难以构建。

模拟电子密度的傅里叶变换（FT）在数学上与物理散射过程高度相似，可用于准确计算 X 射线和中子散射图样 ‡。这些模拟图样既包含散射现象，也包含衍射现象，因为衍射本质上是来自规则结构的散射），并且其范围可延伸至对分布函数（PDF）区间。傅里叶变换的一大优势是，模拟可以根据需要尽可能贴近现实。其不受限制的特性使得模拟层级电子密度图成为可能，甚至还可实现仪器效应，例如有限的相干性、展宽效应和波长色散。

在采用这种方法之前，仍需解决若干挑战。尤其是，模拟网格分辨率稍有提高，所需计算资源便会急剧增长，乍看之下似乎难以克服。例如，电子密度（ED）盒子包含 512^3 个等距点时，普通硬件可在一分钟内完成傅里叶变换；但若将分辨率加倍，只为获得微不足道的额外 Q 值，就会耗尽大多数计算机系统的内存。反过来，增大 ED 采样点间距又会导致重要细节丢失，从而降低高 Q 值下的预测能力。虽然可以投入更多硬件来解决这一问题 [1]，但大多数实验室无力承担这类方案。因此，我们需要另寻解决之道。

这篇简短通讯探讨了两种方法，它们结合起来可以帮助小型实验室利用高分辨率傅里叶变换进行模拟。首先，我们展示了如何通过平铺和掩膜，将晶体晶胞的从头算电子密度模拟逐片生成晶体纳米颗粒的高分辨率电子密度（ED）图。随后，可逐片对所得 ED 进行傅里叶变换，以获得高分辨率的模拟散射信息。这在一定程度上规避了内存限制问题（转而受可用磁盘空间限制）。其次，我们通过合并模拟，免去了实空间零填充的需要：将初级颗粒的模拟（ED-1）与 ED-1 中单个颗粒较小、视野更广的密度模拟（ED-2 和 ED-3）相结合。最后，我们将这些方法整合成一套可行的策略，称为“Ultima Ratio”§，用于模拟更复杂的层级结构。我们通过一个软件演示程序（Jupyter 笔记本）展示了这一方法；该程序及所需的补充文件均可通过 Zenodo 获取，并遵循 CC-BY-4.0 许可 [2]。这篇简短通讯验证了实现路径，使得开发功能完备的解决方案成为可能。

‡ 要模拟中子散射，输入应使用核密度（以及适用时的磁结构），而非电子密度。§ “*Ultima Ratio Regum*”这一铭文历史上曾刻在路易十四下令铸造的大炮上，字面意思是“国王的最后论据”。与其同名者一样，这里的模拟软件只有在其他所有假设都不成立时才会派上用场，因此应将其视为解决这一特定问题的一种相当粗暴的工具。

2. 方法

2.1. 总体工作流程

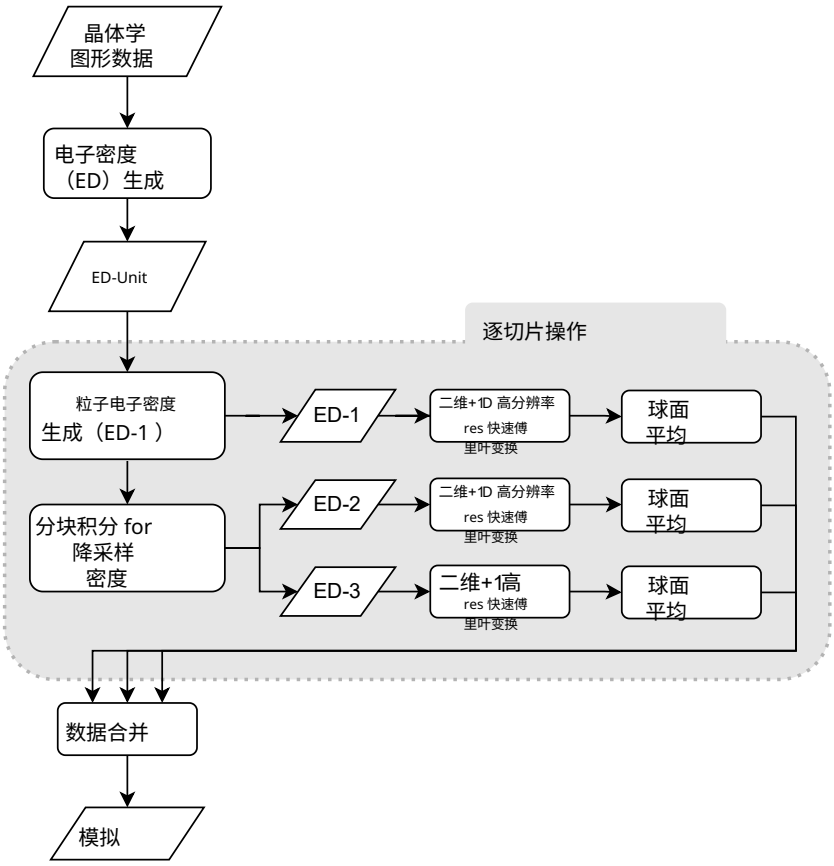


图1. *UltimaRatio*工作流程：用于获取宽范围 X射线散射模拟。结合多个尺度并分离二维+1维快速傅里叶变换步骤（无需占用三维快速傅里叶变换所需的内存），即可模拟真实密度和（层级）结构的散射。

为展示总体策略，我们针对一种具有沸石咪唑框架（ZIF-8）化学组成的代表性金属有机框架（MOF）颗粒，模拟了宽范围散射图样 [3]。实现这一策略所采用的总体工作流程以示意图的形式呈现在图1中，并将在下文段落中详细说明。简而言之，步骤如下：

(i) 生成晶胞的电子密度图（ED-Unit）。(ii) 随后将 ED-Unit 平铺在颗粒边界内，生成颗粒的大型高分辨率电子密度图（ED-1）。(iii) 对 ED-1 逐步进行块积分，以获得“缩小视图”结构的低分辨率 ED-2 和 ED-3 图。(iv) 对不同缩放级别下的颗粒（ED-1、ED-2 和 ED-3）进行零填充，并使用分离的三维快速傅里叶变换，按切片进行三维傅里叶变换。

(v) 对每个得到的快速傅里叶变换结果进行球面平均，以获得一条强度 I 与散射矢量 Q的 一维曲线。(vi) 通过合并步骤将这些曲线整合为一个数据集。

除第一步和最后一步外，所有步骤均以分片方式执行，并将完整矩阵临时存储在硬盘上，采用层次数据格式（HDF5），该格式支持数组的延迟加载和多个并发读取操作 [4]。

2.2. 密度泛函理论计算细节

使用密度泛函理论计算金属有机框架常规晶胞（ED-Unit）的全电子电荷密度图，晶胞尺寸为1.683^{3}纳米^{3} 。ZIF-8 的初始结构取自晶体学开放数据库（COD ID：4118891）[3, 5, 6]。该结构模型中的 C 和 H 原子具有部分占位率。为简化起见，我们使用 VESTA[7]，仅保留 C 和 H 占位率较高（0.630）的位点。

密度泛函理论计算使用 VASP 5.4.4 [8, 9, 10] ，并采用投影缀加波（PAW） [11] 赝势（PP）。交换-相关能采用 Perdew–Burke–Ernzerhof（PBE）[12]泛函处理。进行两次不包含体积弛豫的结构优化。每次优化中，平面波截断能设为 520 eV，k点网格设为$1\times1\times1$，电子能量收敛判据设为 10^{-6} eV，离子收敛判据设为 10^{-5} eV。

接下来，对弛豫后的结构进行截断能和k点网格收敛测试。通过执行 Bader电荷分析[13, 14, 15]验证全电子电荷密度文件的可靠性。若基于 Bader分区得到的总电子数（包括芯电子）等于结构中的电子总数，且每个原子的 Bader 电荷均已收敛，则认为全电子电荷密度可靠。经测试，在 VASP静态计算中，采用 500 eV 的截断能、以 Γ 为中心的k点网格 $2\times2\times2$，以及包含 630 个点的快速傅里叶变换网格，是满足该标准的最佳设置。这些收敛测试的结果见以下数据存储库 [2]。对于 ED-Unit，总电子数误差为 0.023 %（$1200.227/1200$）。

2.3. 电子密度生成

上文生成的 ED-Unit 尺寸为 630^{3} 个体素，数据以 32 位浮点数形式提供。在用于模拟之前，我们通过 scikit-image 库的块积分功能降低其分辨率，即 `skimage.measure.block_reduce`[https://doi.org/10.7717/peerj.453]。选择的降采样因子需满足以下条件：1）VASP 晶胞的尺寸可被该因子整除；2）最大散射矢量 Q 仍处于模拟范围内。本次模拟中，降采样因子设为

6，生成尺寸为 105^3 个体素的降采样 ED-Unit。结合上文的晶胞晶格尺寸，降采样后的采样步长为 0.16 ° 埃¨。

随后，将缩减后的 ED-Unit 插值到旋转后的二维平面网格上（进行旋转是为了展示生成具有随机微晶取向的多晶材料的实际可能性），并在其边界之外重复 ED-Unit。此过程利用了 SciPy 在 ”`grid-wrap`” 模式下的 ”`map_coordinates`” 方法（`scipy.ndimage.map_coordinates`）。随后，每个二维平面网格都会乘以一个取决于平面坐标的圆形掩膜，该掩膜用于定义粒子外边界。该平面沿其法线方向每次移动一个体素长度，从而以逐片方式生成三维电子密度（ED-1），描述一个内部填充着晶胞晶格的球形粒子。

为了改进低 Q 模拟，需要在模拟盒中感兴趣对象周围进行额外的（零）填充。具体做法是将感兴趣对象“缩小显示”，从而生成 ED-2 和 ED-3：通过逐片块积分，并分别以 16 和 125 的因子对 ED-1 进行零填充来实现。此时已不再特别关注高分辨率，因此 ED-2 和 ED-3 的盒子体素分辨率也同时降低，以得到更快的小盒子。积分和分辨率降低均通过单个块积分步骤完成，但需要进行两个正交的积分步骤（第一个针对二维图像，第二个使用剩余方向上的一维核）。生成的电子密度以三维数据集的形式存储在 HDF5 文件中，采用 32 位浮点数。ED-1、ED-2 和 ED-3 如图 2 所示。

关于 ED-1、ED-2 和 ED-3 的尺寸，需要注意的是，快速傅里叶变换方法默认假设存在周期性边界条件（PBC）。对于包含一个或多个重复单元的结构模拟，可以利用这些周期性边界条件。然而，为了避免物体受到周期性边界相邻物体的干扰（孤立金属有机框架颗粒的情况正是如此），感兴趣的物体所占空间不应超过零填充盒子的一半。因此，虽然我们要计算装有 8000^3 个体素的盒子中 64 纳米颗粒的快速傅里叶变换，但颗粒本身只占 4000^3 个体素，因此无需写入其余密度，从而节省 8 倍的磁盘输入输出开销。其余体素会在快速傅里叶变换过程中添加。因此，ED-2 和 ED-3 的感兴趣密度所占空间按比例更小，其 250^3 和 32^3 的尺寸能够轻松放入总共包含 2000^3 个体素的快速傅里叶变换盒子中。

2.4. 傅里叶变换

三维快速傅里叶变换（快速傅里叶变换）通过将傅里叶变换拆分为 n 2D 快速傅里叶变换，然后沿正交方向进行 n^2 1D 快速傅里叶变换来实现。电子密度的三维快速傅里叶变换定义如下：

$$F(u,v,w) = \frac{1}{MNL}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\sum_{z=0}^{L-1}\rho_{xyz}\cdot e^{-i2\pi(\frac{ux}{M}+\frac{vy}{N}+\frac{wz}{L})}, \tag{1}$$

其中，电子密度 ρ_{xyz} 现在由位于 $(x,\ y,\ z)$ 的采样点给出

三维空间中的坐标，u、v 和 w 分别描述测量信号所在空间中的 3 个域。

通过分离一个指数项，可将其分解为 n 2D 和 n^2 个正交的一维分量，从而得到我们采用的方法：

$$F(u,v,w) = \frac{1}{M}\sum_{x=0}^{M-1} e^{-i2\pi\frac{ux}{M}} \left(\frac{1}{NL}\sum_{y=0}^{N-1}\sum_{z=0}^{L-1} \rho_{xyz} e^{-i2\pi(\frac{vy}{N}+\frac{wz}{L})} \right) \tag{2}$$

这些部分可以分别计算，并在二维与一维快速傅里叶变换之间使用中间磁盘存储，从而将内存需求从 $\mathcal{O}(n^3)$ 降低到 $\mathcal{O}(n^2)$（不过，最终三维快速傅里叶变换所需的磁盘存储仍为 $\mathcal{O}(n^3)$。这种分离方式的另一项优势是，对于初始二维快速傅里叶变换，只需计算具有非零密度的切片，这分别可使中间存储和计算的空间与时间开销减少一半。

快速傅里叶变换使用 PyTorch 库（`torch.fft`），因此在有可用硬件时，可以将任务卸载到支持 cuFFT 的图形处理器上 [16]。在存入中间 HDF5 结构之前，会对二维快速傅里叶变换结果进行 fft-shift，使最低频率位于中心。存储二维快速傅里叶变换切片后，会沿正交方向执行一系列一维快速傅里叶变换，并将输出存储到新的 HDF5 文件中。中间快速傅里叶变换结果以 128 位复数存储，以避免精度损失；最终快速傅里叶变换的绝对振幅平方（即强度）则以 32 位浮点数存储。如前所述，为节省计算资源，首次二维快速傅里叶变换仅针对具有密度的切片执行，第二次快速傅里叶变换则会填充未计算的空切片。

*2.5. Q*值计算与球面平均

如果需要研究该样品的各向异性性质，可以通过模拟具有一定厚度的球面（即埃瓦尔德球）来提取反映实际实验的二维数据；球壳厚度由波长分布决定，并沿该球面提取强度。然而，由于实验金属有机框架数据具有各向同性，因此无需二维强度数据。因此，我们逐层对整个基体进行球面平均。对于每一层，根据模拟盒各边的长度信息、每边的体素数量以及该层到中心的距离，计算出一个 Q Q值图。

强度切片和 Q值图都会输入我们的 dataMerge 软件进行分箱（参见 [17]的附录 D）。此分箱过程会在每个对数间隔分箱边界内，同时对强度取平均，并计算均值标准误（以及其他一些统计量）。该软件的核心部分已用编译型 Python（Cython）重写，从而显著提升了运行速度。

这样得到的逐切片平均强度，随后还需通过单独的平均步骤进行合并。这一步同样由 dataMerge 完成，并传播此前确定的不确定性估计。因此，每个高分辨率模拟都会得到一条曲线（如图 2 下半部分所示）。最后，第三步 dataMerge 将多个尺度模拟所得的曲线合并起来。需要

注意的是，过去我们还会将散射图样与体素大小的 sinc 函数进行卷积，以改善高 Q 行为（参见 [18]）；但在这里，由于体素极小，这一步的影响微乎其微，因此我们省略了它。如果需要在接近上限 Q 边界处获得精确结果，可以考虑重新加入这一步。

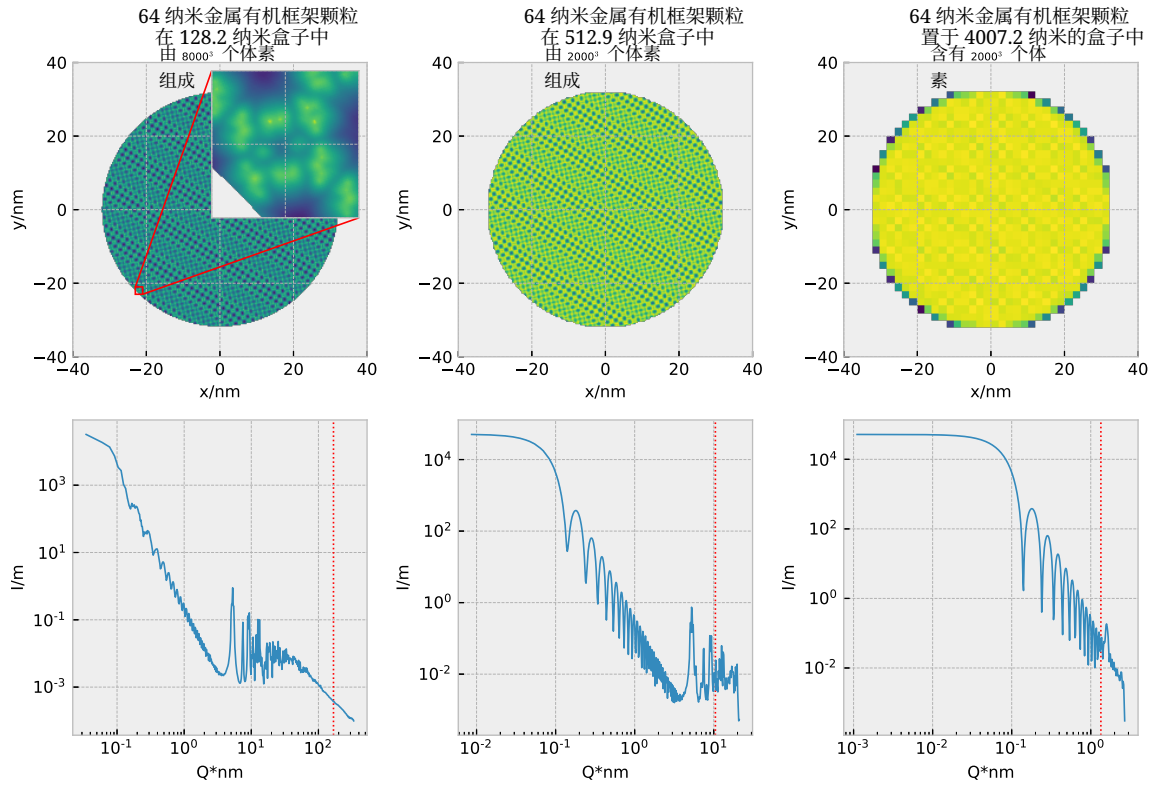


图 2. 上排：电子密度图 ED-1（左）、ED-2（中）和 ED-3（右）的中心切片。图中仅显示感兴趣的部分，填充的零值区域未显示。对于 ED-1，插图显示了电子密度图中一个 2 × 2 纳米区域的放大图。在此色彩映射中，蓝色表示低密度，黄色表示高密度。下排：根据这些结构计算得出的散射图样。红色虚线表示可识别出伪影的界限。该界限出现在 $Q > Q_{max}/2$。

2.6. 计算硬件详情及性能说明

这些模拟在一台运行 Ubuntu Linux 22.04.2 LTS 的双 CPU 惠普 Z8 G4 工作站上进行。该工作站配备两颗标称频率为 1.8 GHz 的 Intel Xeon Silver 4108 处理器。除 20 TB 的临时存储空间外，还配有 384 GB 内存；临时存储空间分布在三块采用 RAID-0 条带配置的机械硬盘上。计算通过 JupyterLab 中的 Python 3.10 内核完成。上述计算需要约 6 TB 的硬盘空间。

虽然 ED-1、ED-2 和 ED-3 的生成速度足够快，但对 ED-1 进行快速傅里叶变换需要三天 8000^3 （相比之下，ED-2 和 ED-3 各需约 20 分钟）。处理时间如此之长，主要是因为存在输入/输出瓶颈

原因是使用了旋转磁盘，而一次 4000^3 快速傅里叶变换仅需约四小时。对于不超过 4000^3 体素的情况，CPU 核心始终处于满负荷状态；但在 8000^3 时，它们会明显空闲，同时输入/输出传输已达到上限。因此，如果预算允许且需要执行大型快速傅里叶变换，建议改用速度更快的固态硬盘存储。改进代码也可能解决其中一些问题。

2.7. 合成与表征

ZIF-8（锌咪唑骨架-8）由两种储备溶液合成：第一种为六水合硝酸锌溶于甲醇（MeOH）所得，第二种为2-甲基咪唑（2-MeIm）溶于MeOH所得。将每种储备溶液各取10 ml，以15 ml/min的速率注入离心管中，并在环境实验室温度 28°C下以200 rpm搅拌21.5(5)分钟。最终合成体系中Zn：2-MeIm：MeOH的摩尔比为1：7.5：290。反应达到规定时间后，将反应混合物以6000 rpm离心20分钟，随后在 60°C下干燥22小时。之后使用小角/广角X射线散射对合成所得粉末进行表征。

使用MOUSE仪器进行了小角/广角X射线散射测量，采用微焦点单色X射线源（$\lambda_{\mathrm{Cu}} = 0.154$纳米） [17]。使用置于真空中的Dectris Eiger R 1M探测器采集数据，探测器与样品之间的距离设为多个值，约为57–2507 mm。粉末样品夹在两片Scotch Magic Tape™透明胶带之间，置于光束中。随后使用DAWN软件包按照标准化的数据校正流程处理数据[19, 20, 21]。校正后，采用一种合并流程整合不同距离下的数据集，以优化重叠区域的不确定度，从而得到一条连续曲线 [22]。

3. 结果与讨论

所得模拟X射线散射曲线涵盖从小角到广角再到对分布函数的散射矢量范围，如图3所示。值得注意的是，采用标准方法无法模拟如此大的盒子尺寸，但采用分离方法则可以实现，而且由于中间数据和最终数据的存储效率更高，所需空间也更少。文中还加入了一条对实验中一颗名义尺寸为65纳米的金属有机框架颗粒进行的MOUSE宽范围X射线散射测量结果，以作比较。整个范围内的吻合度非常出色，但仍需作几点说明。

在小角范围内，单个颗粒与多分散颗粒集合之间的差异显而易见。此外，测量期间材料所处的状态（粉末）还会在最低 Q 值处产生额外散射。隆起峰值的位置基本吻合，但如何处理多分散性仍是一个问题。根据我们开发 McSAS软件 [23], 的经验，我们知道，仅用 200 个按体积加权的贡献，就能够准确描述广泛的多分散性。这意味着在这里要进行 200 次模拟，虽然次数不少，但

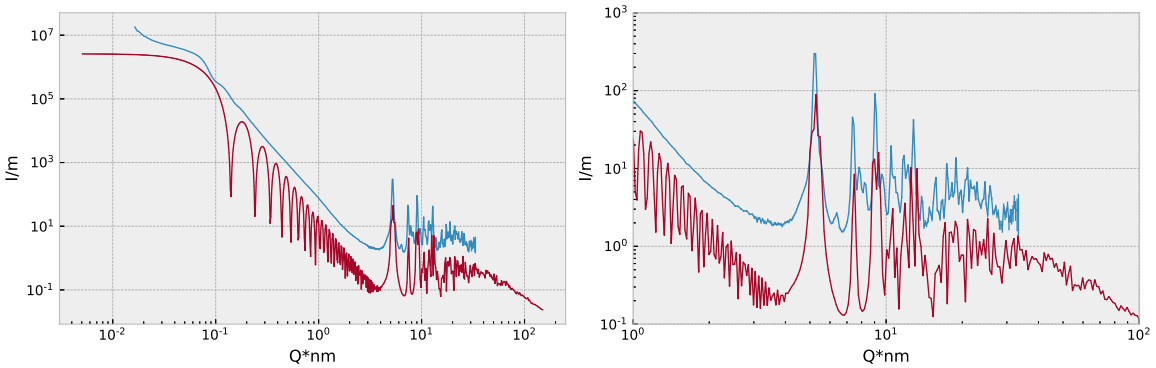


图 3. 合并后的模拟曲线（红色）与我们实验室对 ZIF-8 型金属有机框架颗粒的测量结果（蓝色）的比较。左图显示完整范围，右图仅显示广角和对分布函数范围。

在合理的时间范围内是可行的，特别是改进计算硬件和代码效率之后。这样做还有一个好处，就是能够精确描述广角范围内的峰展宽效应。

即使采用单颗粒模拟，广角范围的吻合度也非常高，无论是峰展宽还是峰位均是如此。虽然一两个峰存在差异，但这并不令人意外，因为：1）样品可能含有微量污染物；2）模拟仅限于由理想化晶格生成的密度，因此可能不包含来自不完美晶体的衍射反射。再看更宽的角度范围，对分布函数范围内的强度逐渐下降，这与迄今采集到的实验数据相符。

如今，结合小角信号所获得的衍射信息，或许开辟了一种确定结晶度的新途径。由于模拟假定结晶度完美，模拟衍射峰下的积分强度应对应于颗粒结晶度为 100%，并可用绝对强度单位评估，或相对于小角散射信号进行评估。有了这一认识，便可将实际测量中的积分峰强度与模拟的纯晶态衍射图谱进行比较，从而提取实际结晶度。

总的来说，这项演示充分证明，X射线散射的高分辨率、大范围模拟是可行的。只要有足够的磁盘空间和计算能力，就能克服此前的分辨率限制。瓶颈如今已转移到实空间，限制在于能否生成足够准确地反映真实结构的电子密度图。已有可用于解决这些问题的工具集和软件库，但可能需要加以改造，才能分块生成高分辨率矩阵。我们预计，近期我们自己以及其他人都会在各种散射相关研究中采用这一方法。

4. 后续步骤

虽然这一概念已得到验证，但若能由专门从事高性能计算（HPC）编程的专家改进底层代码，仍有很大的提升空间。通过以下方式，速度可能会大幅提升：1. 巧妙利用内存和磁盘输入输出，例如使用“zarr”软件库取代 HDF5，因为“zarr”支持同时写入；2. 改进并行化，并更好地划分多个线程和多个进程，还可加入自动调优；3. 通过更完善的性能分析识别并缓解处理瓶颈；4. 可选地将任务卸载至图形处理器硬件，以加快傅里叶变换。

毋庸置疑，经验丰富的程序员能够找出更多改进之处；而且可以预期，只要投入足够的时间，就能发现至少 10 倍的速度提升空间。

其次，可以改进电子密度生成，以提高所得颗粒密度的真实感。这可以包括多晶和非晶颗粒晶粒，以及对颗粒边缘进行更好的截断。SciPy 的“ndimage”库或许是构建此类结构的良好起点，例如对颗粒边缘进行偏置侵蚀，以改善其边界 [https://docs.scipy.org/doc/scipy/tutorial/ndimage.html]。或者，也可以利用其他软件包（如 `PoreSpy` [24] 或 `XRD_simulator` [25] ）生成这类逼真的结构，并通过以晶态电子密度图填充不同相来进一步完善。此外，还可以将周围基体的密度设定为更真实的值，例如水的密度，并加入局部密度涨落（可通过等温压缩率求得）。

最后，可以利用融合多个不同长度尺度的模拟所带来的优势，以更少的资源更快地计算相同或更宽的范围。此外，还可以通过组合不同尺度下的各个组成部分，对多相系统进行模拟，例如由多种不同类型的颗粒负载于多孔碳基体上所构成的系统。

5. 总结

克服内存限制，并结合多尺度模拟后，傅里叶变换方法为模拟宽范围散射图样提供了一条可行途径。这些模拟数据集可直接与真实散射数据进行比较，可用于测试解析模型，也可通过比较信号强度与背景水平来检验实验的可行性。除时间和磁盘空间外，该方法剩下的限制在于电子密度模型是否可实现，以及研究人员的想象力。

6. 致谢

BRP 谨此感谢 Steen L. Hansen 博士多年前随口提到可以将快速傅里叶变换拆分处理。感谢 Thilo Muth 博士提供 Sofya Laskina，让她能够探索这一问题。

7. 参考文献

[1] Cambier L, Pan D, Ligowski L. 多节点多GPU：大规模使用 NVIDIA cuFFTMp 快速傅里叶变换；2022年。最后查阅日期：2023年3月15日。获取地址：https://developer.nvidia.com/blog/multinode-multi-gpu-using-nvidia-cufftmp-ffts-at-scale/.[2] Pauw BR, Laskina S, Naik A, Smales GJ, George J, Breßler I, 等。与论文《Ultima Ratio：模拟宽范围X射线散射与衍射》配套的 Jupyter 笔记本和 VASP 计算详情。Zenodo。2023年3月。获取地址：https://dx.doi.org/10.5281/zenodo.7764045.[3]Karagiaridi O, Lalonde MB, Bury W, Sarjeant AA, Farha OK, Hupp JT. 开启 ZIF-8：一种具有方钠石拓扑结构、连接体未经取代且具催化活性的沸石咪唑酯骨架。美国化学会志。2012年；134(45)：18790-6。获取地址：https://doi.org/10.1021/ja308786r.[4] Folk M, Pourmal E. HDF5 的性能与保存平衡：经验教训。载于：2010年数字保存互操作框架路线图研讨会论文集——US-DPIF ’10。马里兰州盖瑟斯堡：ACM出版社；2010年。第1-8页。获取地址：http://dl.acm.org/citation.cfm?doid=2039274.2039285.[5] Graˇzulis S, Daˇskeviˇc A, Merkys A, Chateigner D, Lutterotti L, Quiros M, 等。晶体学开放数据库（COD）：开放获取的晶体结构集合与全球协作平台。《核酸研究》。2012年；40卷（D1期）：D420-7。获取地址：https: //doi.org/10.1093/nar/gkr900.[6] Downs RT, Hall-Wallace M. 《美国矿物学家》晶体结构数据库。《美国矿物学家》。2003年；88(1)：247-50.[7] Momma K, Izumi F. VESTA 3：用于晶体、体积和形貌数据三维可视化的软件。《应用晶体学杂志》。2011年；44(6)：1272-6。获取地址：https://doi.org/10.1107/S0021889811038970.[8] Kresse G, Hafner J. 液态金属的从头算分子动力学。《物理评论B》。1993年1月；47：558-61。获取地址：https://link.aps.org/doi/10.1103/PhysRevB.47.558.[9]Kresse G, Furthm¨uller J. 使用平面波基组进行从头算总能量计算的高效迭代方案。《物理评论B》。1996年10月；54：11169-86。获取地址：https: //link.aps.org/doi/10.1103/PhysRevB.54.11169.[10] Kresse G, Joubert D. 从超软赝势到投影缀加波方法。《物理评论B》。1999年；59(3)：1758。获取地址：https://doi.org/10.1103/PhysRevB.59. 1758.[11] Bl¨ochl PE. 投影缀加波方法。《物理评论B》。1994年；50(24)：17953。获取地址：https://doi.org/10.1103/PhysRevB.50.17953.[12] Perdew J P, Burke K, Ernzerhof M. 简明易用的广义梯度近似。《物理评论快报》。1996年；77卷（18期）：3865。获取地址：https://doi.org/10.1103/PhysRevLett. 77.3865.[13] Bader R. 《分子中的原子：量子理论》。牛津大学出版社：牛津。1990年。[14]Tang W, Sanville E, Henkelman G. 一种无晶格偏差的基于网格的Bader分析算法。《物理学杂志：凝聚态》。2009年；21(8)：084204。获取地址：https: //doi.org/10.1088/0953-8984/21/8/084204.[15]Sanville E, Kenny SD, Smith R, Henkelman G. 改进的基于网格的Bader算法

电荷分配。《计算化学杂志》。2007年；28(5)：899-908。获取地址：https://doi.org/10.1002/jcc.20575.[16]Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, 等。Py- Torch：一种命令式风格、高性能的深度学习库。载于：《神经信息处理系统进展》第32卷。Curran Associates, Inc.；2019年。第8024-35页。获取地址：http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.[17] Smales GJ, Pauw BR. MOUSE项目：一种严谨获取可追溯宽范围X射线散射信息的方法。《仪器学杂志》。2021年；16：P06034。[18] Pauw BR, Vigild ME, Mortensen K, Andreasen JW, Klop EA. 芳纶纤维纳米多孔结构分析。《应用晶体学杂志》。2010年；43(4)：837-49。获取地址：https://doi.org/10.1107/S0021889810017061.[19] Basham M, Filik J, Wharmby MT, Chang PCY, El Kassaby B, Gerring M, 等。数据分析工作台（DAWN） [期刊文章]。《同步辐射杂志》。2015年；22：853-8。[20] Filik J, Ashton AW, Chang PCY, Chater PA, Day SJ, Drakopoulos M, 等。在 DAWN 中处理二维X射线衍射和小角散射数据 2 [期刊文章]。《应用晶体学杂志》。2017年；50(3)：959-66。获取地址：https://doi.org/10. 1107/S1600576717004708.[21]Pauw BR, Smith AJ, Snow T, Terrill NJ, Th¨unemann AF. 模块化小角X射线散射数据校正流程。《应用晶体学杂志》。2017年；50：1800-11。获取地址：https://doi.org/10.1107/S1600576717015096.[22] Pauw BR, Laskina S. dataMerge——一种具备统计感知能力的分箱与合并后端。FigShare。2022年11月。获取地址：https://figshare.com/articles/software/dataMerge_-_A_具备统计感知能力的_分箱_与_合并_后端/21591360.[23] Bressler I, Pauw BR, Th¨unemann AF. *McSAS*：一种用于从散射图样中提取模型参数分布的软件。《应用晶体学杂志》。2015年；48(3)：962-9。获取地址：https://doi.org/10.1107/S1600576715007347.[24] Gostick JT, Khan ZA, Tranter TG, r Kok MD, Agnaou M, Sadeghi M, 等。PoreSpy：用于定量分析多孔介质图像的 Python 工具包。《开源软件杂志》。2019年；4(37)：1296。获取地址：https://doi.org/10.21105/joss.01296.[25] Henningsson A, Hall SA. *xrd*模拟器：支持三维多晶微观结构形貌描述的三维X射线衍射模拟软件。《应用晶体学杂志》。2023年；56(1)：282-92。获取地址：https://doi.org/10.1107/S1600576722011001。