

· 研究论文(明日之星专栏) ·

## 基于区域二分法的结晶高分子纤维衍射指标化程序

胡文贤<sup>1,2</sup> 马天怡<sup>1,2</sup> 刘国明<sup>1,2\*</sup>

(<sup>1</sup>中国科学院化学研究所 中国科学院工程塑料院重点实验室 北京 100190)

(<sup>2</sup>中国科学院大学 北京 100049)

**摘要** 指标化是分析未知物衍射图案得出晶体晶胞参数和衍射峰对应 Miller 指数的步骤, 是晶体结构解析重要的组成部分. 高分子因其结晶的特殊性, 使晶体结构解析工作很难通过软件自动化完成, 其中也包括指标化. 为解决该问题, 本工作以二分递进法作为核心算法, 基于 Python 语言编写了可用于聚合物纤维衍射自动化指标化程序. 该程序可通过简单的参数输入, 实现高分子正交晶系与单斜晶系的快速指标化. 为验证程序可靠性和效率, 使用已知晶胞参数的 3 种聚合物对其进行了验证. 结果表明, 程序准确性高, 运算速度较快, 达到了应用水平.

**关键词** 聚合物晶体; 晶体结构; 指标化; 二分递进法; 纤维衍射

**引用:** 胡文贤, 马天怡, 刘国明. 基于区域二分法的结晶高分子纤维衍射指标化程序. 高分子学报, 2026, 57(3), 609–620.

**Citation:** Hu, W. X.; Ma, T. Y.; Liu, G. M. Indexing program for crystalline polymer fiber patterns based on dichotomy method. *Acta Polymerica Sinica* (in Chinese), 2026, 57(3), 609–620.

结晶性高分子材料是高分子领域的支柱品类, 在塑料、纤维等材料领域具有广泛应用. 对于可结晶聚合物, 其晶体结构的解析是研究其物理性质, 理解其性能, 规划其应用场景的首要环节. 高分子材料晶体结构解析主要使用 X 射线纤维衍射法进行<sup>[1]</sup>. 在单轴高倍拉伸的聚合物纤维或样条中, 拉伸方向与分子链投影方向重合. 在聚合物晶体学中, 通常将分子链轴方向人为定义为晶胞参数  $c$  轴方向, 因此, 纤维衍射法的衍射几何等效于沿  $c$  轴方向旋转的转晶衍射法, 相较于一维衍射更有利于晶体结构解析.

在聚合物晶体解析的过程中, 指标化是一大难点. 指标化是将晶体的衍射点的未知晶面指数全部解析的过程, 是晶体结构分析中生成晶胞参数的基础阶段. 晶胞参数的精确测定是确立晶系和识别可能对称元素的先决条件<sup>[2]</sup>. 对于小分子晶

体样品而言, 无论是单晶衍射晶体指标化方法, 还是粉末衍射指标化方法, 均已发展出诸多算法与软件. 在小分子单晶解析中, 常见的指标化程序有 HKL<sup>[3]</sup>, UNITCELL<sup>[4]</sup> 等, 其在单晶解析中都可达到很高的准确度. 在小分子、无机粉末多晶衍射领域, 科学家们已开发了多款基于不同算法的指标化程序, 如基于二分法的 DICVOL<sup>[5,6]</sup>、基于试错与非线性优化法的 TREOR<sup>[7]</sup>、基于倒易晶格几何解析法的 ITO<sup>[8]</sup>、基于遗传算法的 GAIN<sup>[9]</sup> 等. 尽管这些程序在小分子晶体指标化中成效显著, 但它们在聚合物晶体指标化中的应用仍存在局限. 由于聚合物晶体具有很大特殊性, 其分子链在一般情况下形成折叠链片晶, 极难制备大尺寸单晶样品, 因而无法直接使用单晶指标化程序进行解析. 此外, 聚合物材料固有的低结晶度、微小晶粒尺寸及较高无序性, 往往导致衍

2025-10-18 收稿, 2025-11-10 录用, 2026-01-14 网络出版.  
基金项目: 国家自然科学基金(基金号 22273113, 21922308).

\* 通信联系人, E-mail: gmliu@iccas.ac.cn

doi: 10.11777/j.issn1000-3304.2025.25264; CSTR: 32057.14.GFZXB.2025.7515

射峰宽化与重叠,小分子多晶解析软件难以适用.因此,高分子晶体结构解析对传统指标化软件提出了严峻挑战.

当前高分子晶体指标化方法主要采用人工操作分步进行,并以简单工具进行辅助.例如,在 高分子纤维晶衍射中,晶胞  $c$  轴长度可直接通过层线位置确定,可采用图解法拟合各层线内的衍射点.在图解法中,研究人员使用假定的倒易网格对不同层线上的点进行试配,与衍射点完全吻合的网格即可能为正确的晶胞参数对应的倒易点阵<sup>[10]</sup>.但这些方法本质上为人工试错,耗时较长,效率较低,且高度依赖解析者的经验与直觉.针对高分子晶体指标化的现状,开发一款可用于聚合物晶体指标化的程序对提高聚合物晶体解析准确度和效率具有重要意义.

在上述小分子粉末衍射指标化程序中,虽然直接应用于高分子解析中有困难,但部分程序的算法值得借鉴.其中, DICVOL 具有很高的成功率,其算法是基于穷举并裂分参数空间的二分递进法<sup>[11,12]</sup>.算法的基本原理就是通过划分未知晶胞参数空间,使用假想参数组合计算晶面间距后与实验数据进行匹配拟合,后对可能存在真实晶胞参数的参数空间进行二分递进,迭代计算,最终得到可能正确的晶胞参数组合.

本工作中,我们基于 Python 编程语言,开发了一种基于二分递进法的聚合物晶体指标化程序.该算法与之前本课题组提出的遗传算法不同<sup>[13]</sup>,理论上不会出现遗漏,且算法逻辑较为简单. Python 语言是当今应用最广泛的编程语言之一,与其他编程语言相比,科学计算、专业绘图与数据处理方面都具有编写简洁、实现简单、准确度高优势.该程序使用聚合物 X 射线纤维衍射数据,具有以下特点:(1)可自动指标化属于正交、单斜晶系的聚合物晶体,通过正交晶系也可拓展到四方晶系及高对称晶系.(2)计算速度较快,指标化速度可达秒级.(3)通过聚乙烯(PE)、等规聚丙烯(*i*PP)、聚(丁二酸丁二醇酯)(PBS)纤维衍射图案对程序的准确性进行了验证,其指标化准确度较高,对于上述聚合物都可得到正确的晶胞参数.该程序丰富了聚合物晶体结构解析的工具箱,提高了聚合物晶体解析准确度和效率,也为后续聚合物晶体指标化程序的继续完善与发展铺平了道路.

## 1 程序编写与实验部分

### 1.1 程序编写环境与测试条件

程序基于 Python 3.9 进行编写,主要基于 Numpy、Scipy、Matplotlib 等 Python 模块进行.使用 Numba 模块对程序进行加速,并使用 Spyder (Anaconda3)编译器进行编写.

由于实验中测试了程序运行的速度,存在计算机性能配置的主观原因.本实验主要使用计算机中央处理器(CPU)进行计算.所用 CPU 为 AMD R7 5800H,程序只利用单核进行计算.

### 1.2 程序算法原理及步骤

程序的指标化功能所使用的数学原理是基于二分递进法修改而来,其本质是一种穷举与逼近法,原理和计算步骤如下:

首先,根据猜测的晶胞种类确定未知晶胞参数数量.由于聚合物纤维晶衍射可快速确定  $c$  轴长度,则相应晶系的未知晶胞参数都减少一个(如四方晶系为 1 个,正交晶系 2 个,单斜晶系 3 个,三斜晶系 5 个).

在确定未知参数个数后对参数空间进行区域划分,将多维参数空间以一定间隔分割为若干个子区域,在计算机中每个小空间对应的参数即为所有未知参数划分的遍历:

$$\Omega = \{a_i, b_i, c_i, \alpha_i, \beta_i, \gamma_i\} \quad (1)$$

对于区域的划分如图 1 所示(以正交晶解析为例):

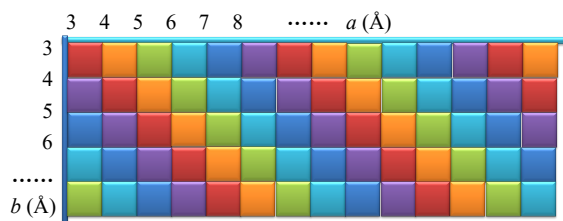


Fig. 1 Division of the subregions (orthorhombic).

假设真实的晶胞参数落在某个子区域中,那么该晶胞所有衍射晶面对应的晶面间距都应落在基于该子区域参数范围所计算的晶面间距范围内,则对于任意晶面指数,下式始终成立:

$$d_{hkl, \min} < d_{hkl} < d_{hkl, \max} \quad (2)$$

式中,  $h, k, l$  为晶面指数.由此,问题转化为如何快速求解出子区域中任意晶面间距的取值范围.

对于正交晶系, 其晶面间距计算方式相对简单, 其取值范围也很容易确定:

$$Q_{hkl, \min} = \frac{1}{d_{hkl, \max}^2} = \frac{h^2}{a_{\max}^2} + \frac{k^2}{b_{\max}^2} + \frac{l^2}{c^2} \quad (3)$$

$$Q_{hkl, \max} = \frac{1}{d_{hkl, \min}^2} = \frac{h^2}{a_{\min}^2} + \frac{k^2}{b_{\min}^2} + \frac{l^2}{c^2} \quad (4)$$

式中,  $d_{hkl, \min}$  和  $d_{hkl, \max}$  分别对应着在某子区域中晶面指数为  $(hkl)$  的晶面间距的最小值与最大值.  $Q_{hkl, \min}$  和  $Q_{hkl, \max}$  对应着晶面间距平方的倒数, 更便于进行计算. 而  $a_{\min}$ 、 $a_{\max}$ 、 $b_{\min}$ 、 $b_{\max}$  分别是正交晶未知参数  $a$  和  $b$  在该子区域内的最小值和最大值.

对于单斜晶系, 情况则复杂的多. 以  $\beta \neq 90^\circ$  单斜晶为例. 其  $Q$  值的表示式为(5)和(6):

$$Q_{hkl} = h^2 \times a^{*2} + k^2 \times b^{*2} + l^2 \times c^{*2} + 2hl \times a^* c^* \times \cos(\beta^*) \quad (5)$$

$$Q_{hkl} = \frac{1}{d_{hkl}^2} = \frac{h^2}{a^2 \times \sin^2(\beta)} + \frac{k^2}{b^2} + \frac{l^2}{c^2 \times \sin^2(\beta)} - \frac{2hl \times \cos(\beta)}{ac \times \sin^2(\beta)} \quad (6)$$

式中  $a^*$ 、 $b^*$ 、 $c^*$  代表着倒易矢量,  $\beta^*$  为倒易晶胞角参数, 数值为  $\beta$  的补角. 根据上述公式(5)和(6), 将  $Q$  的求解转化为一个多元函数求极值的问题. 式中,  $b$  项只有一个变量, 不含有交叉项, 因此,  $Q$  的最小值一定对应着  $b$  的最大值, 反之亦然. 因此我们可以不考虑这一项, 此时令  $a \times \sin(\beta)$  为独立变量  $A$ ,  $\beta$  为独立变量, 则上式可以变成:

$$Q_{hkl} = \frac{1}{d_{hkl}^2} = \frac{h^2}{A^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2 \times \sin^2(\beta)} - \frac{2hl \times \cos(\beta)}{Ac \times \sin(\beta)} \quad (7)$$

虽然  $A$  与  $a$  和  $\beta$  都有关, 但它可视为是独立变量, 可以被独立划分, 使用变量  $A$  代替  $a$  进行划分即可. 当  $hl$  乘积大于 0,  $\beta$  大于  $90^\circ$  时(初始参数取值范围可选  $90^\circ \sim 180^\circ$ ), 交叉项的值小于 0, 与前面的负号相乘后大于 0, 则很容易求出  $Q$  的极值:

$$Q_{hkl, \min} = f(A_{\max}, b_{\max}, c, \beta_{\min}) \quad (8)$$

$$Q_{hkl, \max} = f(A_{\min}, b_{\min}, c, \beta_{\max}) \quad (9)$$

公式形式用函数  $f$  代替.

当  $hl$  乘积小于 0 时, 情况就较为复杂. 式(7)中我们只考虑与  $A$ ,  $c$  相关的三项, 可将其简化为一个与  $A$  和  $\beta$  二元函数. 对  $A$  求偏导数为 0 时, 即为函数极值, 对应的  $A$  为:

$$A = \frac{ch \times \tan(\beta)}{l} \quad (10)$$

此时,  $Q$  对应最小值:

$$Q_{hkl, \min} = \frac{l^2}{c^2} + \frac{k^2}{b_{\max}^2} \quad (11)$$

根据多元函数的表达式, 可以绘制出前 3 项在  $A$  和  $\beta$  变化时对应的曲面:

图 2(a) 为三维图像, 图 2(b) 为高度图, 纵坐标为  $\beta$  (弧度制), 横坐标为  $A$ , 此时  $h=-1$ ,  $l=1$ , 该弧面为无鞍点凹面, 偏导数为 0 的极值即为最小值. 可以看到图中有一片深蓝色带, 为  $Q$  值较低的区域, 而最小值就落在  $A$  满足偏导数为 0 的谷线上. 即公式(10), 谷线形状如图 2(c) 所示.

图 3 中的 3 个红框代表我们在分割变量空间时所对应的 3 种情况, 作为子区域时它们对应的  $Q$  的最大最小值如下:

区域 1:

$$A_{\min}, A_{\max} < \frac{ch \times \tan(\beta_{\max})}{l} \quad (12)$$

$$Q_{hkl, \min} = f(A_{\max}, b_{\max}, c, \beta_{\max})$$

$$Q_{hkl, \max} = f(A_{\min}, b_{\min}, c, \beta_{\min})$$

区域 3:

$$A_{\min}, A_{\max} > \frac{ch \times \tan(\beta_{\min})}{l} \quad (13)$$

$$Q_{hkl, \min} = f(A_{\min}, b_{\max}, c, \beta_{\min})$$

$$Q_{hkl, \max} = f(A_{\max}, b_{\min}, c, \beta_{\max})$$

区域 2:

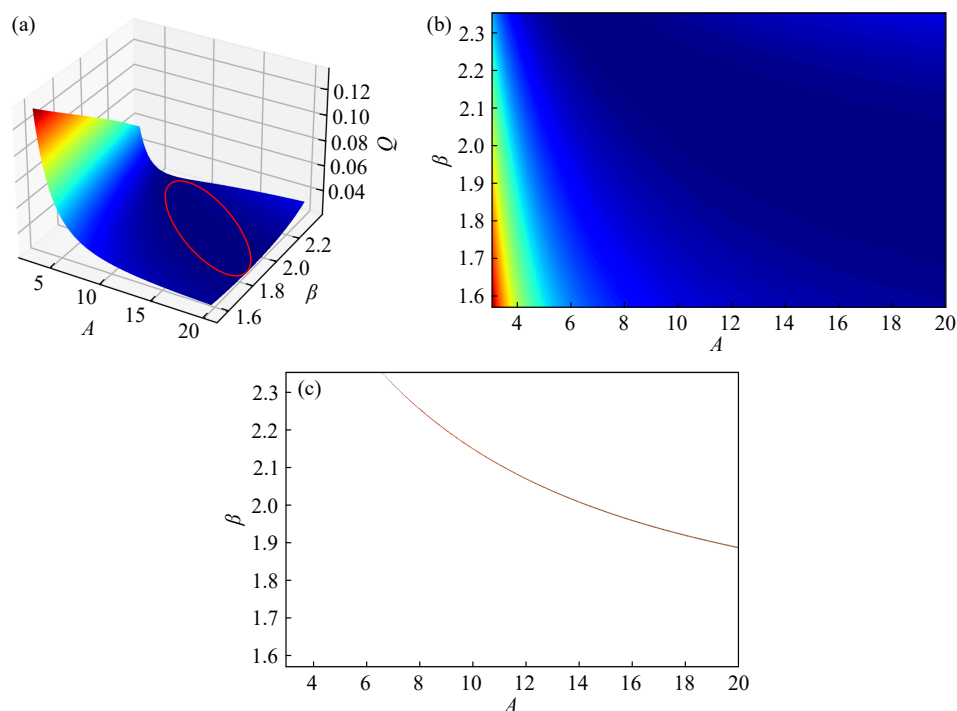
$$Q_{hkl, \min} = \frac{l^2}{c^2} + \frac{k^2}{b_{\max}^2} \quad (14)$$

$$Q_{hkl, \max} =$$

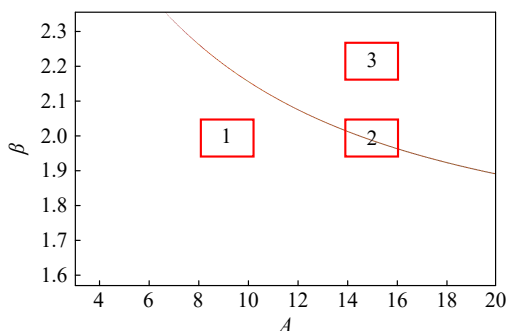
$$\max \left( f(A_{\max}, b_{\min}, c, \beta_{\max}), f(A_{\max}, b_{\min}, c, \beta_{\min}), f(A_{\min}, b_{\min}, c, \beta_{\min}), f(A_{\min}, b_{\min}, c, \beta_{\max}) \right)$$

至此, 我们完成了单斜晶系子区域晶面间距取值范围的计算.

在计算出某个子区域中任意晶面指数所对应的晶面间距最小值与最大值后, 便可在计算机中生成 2 个三维的数据矩阵, 其 3 个维度即为晶面指数  $h$ 、 $k$ 、 $l$ . 其数值分别由晶面间距最大值和最小值构成. 同理, 从纤维衍射图中提取的衍射点晶面间距同样可由层线划分而生成 1 个二维数组. 计算机通过算法, 将二维数组中的数值与三维矩阵所限定的范围进行逐一比对, 当所有实验数据都能落入固定  $h$ 、 $k$ 、 $l$  组合之间的限定范围

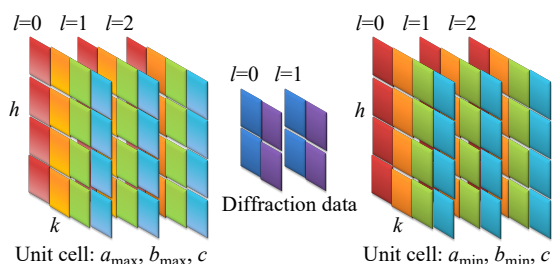


**Fig. 2** The visualization of the multivariate functions described in Eq. (7): (a) 3D surface (minima highlight in red circle), (b) the contour figure of (a), (c) the valley line of the surface.



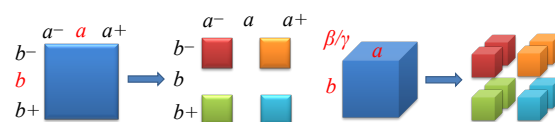
**Fig. 3** Handling extrema in different subregions of multivariate functions (monoclinic system).

内时, 便将此子区域列为可能包含正确晶胞参数的区域, 将其参数记录下来并进行下一步操作, 计算机匹配示意图如图4所示.



**Fig. 4** Schematic diagram of experimental data-subregion matching algorithm (left: orthorhombic system; right: monoclinic system).

将提取出的可能包含正确晶胞参数的子区域全部进行二分递进处理, 即将子区域的所有未知参数取值范围全部裂分, 形成更小的子区域, 每次裂分所产生的新子区域数量为原区域的 $2^n$ 倍, 其中 $n$ 为未知参数的数量, 如图5所示. 将这些新裂分产生的子区域再次进行以上运算, 求解其对应的晶面间距取值范围, 并与实验数据进行匹配, 筛选出可能包含真实晶胞参数的子区域.



**Fig. 5** Schematic diagram of the dichotomy (subregion bisection and progressive splitting) (orthorhombic system).

再次重复二分递进, 直到子区域的参数范围很小(本程序为 $0.0625 \text{ \AA}$ ). 当某些小区域依然满足匹配条件时, 再通过对小区域参数的多等分计算, 求解出可能正确的、精确的晶胞参数, 并将结果打印. 根据该组晶胞参数计算出每个衍射点对应晶面间距并与实验数据进行比对, 计算总误差. 通过总误差判断该组晶胞参数与真实参数的吻合程度. 总误差如公式(15)所示:

$$\Delta E = \sum |Q_{\text{obs}}^2 - Q_{\text{calc}}^2| \quad (15)$$

### 1.3 程序输入与输出

依据数学原理, 程序运行前需输入已解析出的 $c$ 轴方向长度, 以及衍射数据二维数组. 第一维度是层线的序号, 例如( $hk0$ )晶面的第一维度均为0; 第二维度为某条层线上的所有可提取的衍射点所对应的 $Q$ 值; 为保证第二维度上数据量相同, 当不同层线上衍射点数量不同时, 程序采用0补齐, 在运算时程序将自动跳过. 除此之外, X射线波长等固有信息也需输入.

程序输出为候选的晶胞参数组合, 以及所有输入衍射点在该晶胞参数下对应的晶面指数, 实验得到的晶面间距, 计算得到的理论晶面间距, 以及计算误差.

### 1.4 程序优化

为了进一步提高程序的运行速度, 使用Numba模块对程序进行优化提速. Numba是一个基于LLVM的即时编译器(Just-In-Time, JIT), 专门用于加速Python中数值计算和科学计算相关的代码. 其核心原理是通过将Python代码(尤其是涉及循环和NumPy操作的代码)动态编译为高效的机器码, 从而绕过Python解释器的性能瓶颈. 使用该模块可大幅提高程序的运算效率, 极大减少程序计算所消耗的时间.

### 1.5 聚合物拉伸样品制备与纤维衍射实验

为验证程序对聚合物纤维X射线衍射图案指标化的准确度与稳定性, 我们制备了几种常见的聚合物(聚乙烯、聚丙烯、聚(丁二酸丁二醇酯))的高取向拉伸样品, 并进行了X射线纤维衍射实验. 使用纤维衍射图案对程序进行了测试.

#### 1.5.1 样品信息

高密度聚乙烯(HDPE)和等规聚丙烯(*i*PP)均从Sigma-Aldrich公司购买; PBS样品由西班牙巴斯克大学Müller教授提供.

#### 1.5.2 聚合物样品拉伸实验

HDPE样条通过热压膜制备, 在50 MPa, 160 °C下排气1 min, 后热压3 min, 最后冷压3 min结晶定型(冷压即是在一定温度热压后将聚合物熔体转移到非加热的常温冷压台上再次压膜成型. 在此过程中, 聚合物完全冷却并结晶, 形成可以裁剪的样片), 使用裁刀裁成样条形状. 将样条置于拉伸机上(Linkam TST 350), 在常温下以25  $\mu\text{m/s}$ 的速率拉伸至工程应变850%, 并保持一段时间, 使分子链充分松弛. 在实验前将样品放

置在真空腔中120 °C热处理8 h, 后进行衍射实验.

*i*PP样条同样通过热压膜制备, 在50 MPa, 210 °C下排气1 min, 后热压3 min, 最后冷压3 min结晶定型, 使用裁刀裁成样条形状. 将样条置于拉伸机上(Linkam TST 350), 在145 °C下以25  $\mu\text{m/s}$ 的速率拉伸至工程应变700%, 并保持一段时间, 使分子链充分松弛. 在实验前将样品放置在真空腔中以150 °C热处理36 h, 后进行衍射实验.

PBS样条处理方法与上述方法类似, 在50 MPa, 150 °C下排气1 min, 后热压3 min, 最后冷压3 min结晶定型, 使用裁刀裁成样条形状. 将样条置于拉伸机上(Linkam TST 350), 在100 °C下以25  $\mu\text{m/s}$ 的速率拉伸至工程应变700%, 并保持一段时间, 使分子链充分松弛. 在实验前将样品放置在真空腔中以100 °C热处理12 h, 后进行衍射实验.

#### 1.5.3 聚合物拉伸样条X射线衍射实验

使用Xenocs Xeuss 2.0 X射线散射系统测试高度取向的三种聚合物样条以获得纤维衍射图像. X射线源为铜靶, 波长为1.54189 Å, 探测器为Dectris公司Pilatus 300K半导体探测器, 分辨率为 $619 \times 487$ , 像素尺寸 $172 \mu\text{m} \times 172 \mu\text{m}$ . 探测器经校正已完全垂直于入射光斑. 测试温度为常温. 为校准纤维样品拉伸方向垂直于入射光, 先使用掠入射法校正样品台基座, 后将待测试样品固定在精密加工的自制样品台垂直的固定机构上. 同时, 自制纤维样品台将样品到探测器距离(SD值)缩短至85 mm左右, 以获得更多高角度信息. 为了进一步获得更多高角度信息并提高信噪比, 在实验中根据纤维衍射图案调整探测器位置进行曝光, 并在数据处理过程中拼接图像.

#### 1.5.4 纤维图像衍射点提取与程序数据输入

当获得上述三种聚合物纤维衍射图后, 使用Fit2D对图像中的衍射点进行切片分析, 结合自编数据处理程序, 将纤维衍射图上可提取的衍射点的衍射角全部提取. 使用Fraser校正法将纤维衍射图转化为倒易空间坐标<sup>[4]</sup>, 并使用自编程序对转化后的图像的 $Z$ 方向进行积分. 通过层线间隔计算出精确的 $c$ 轴长度, 用于输入程序进行计算.

## 2 结果与讨论

### 2.1 PE衍射图像分析与程序指标化

通过拼接探测器接收到的信号, 可得PE纤

维衍射图像,如图6(a)所示:通过拓展信号收集范围,可观测到2个衍射层线,其层线间间隔较大,表明PE的 $c$ 轴长度较短.对图像应用Fraser校正后得到图6(b),其中横坐标 $R$ 为倒易空间圆

柱坐标中平面极坐标长度,纵坐标 $Z$ 为倒易空间圆柱坐标系的纵坐标.将图6(b)沿其 $Z$ 方向积分(垂直方向)可得 $c$ 轴长度约为2.57 Å.此数值可用于输入指标化程序.

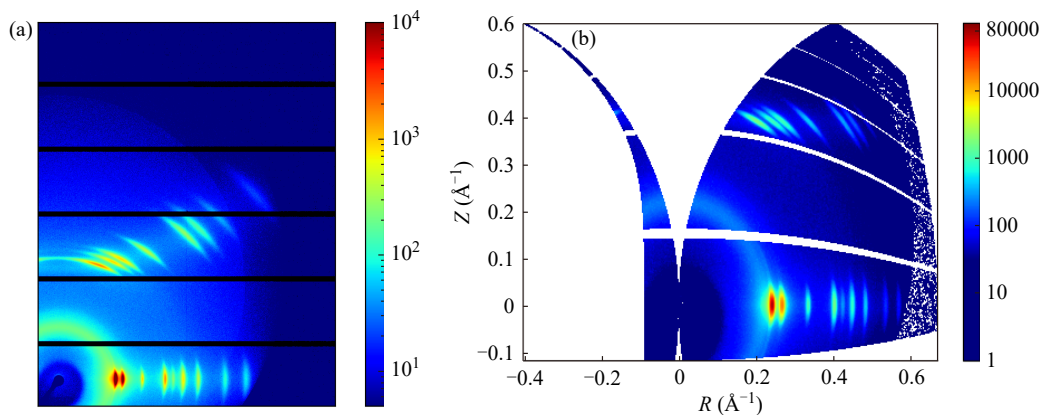


Fig. 6 PE fiber diffraction: (a) stitched diffraction pattern, (b) diffraction pattern treated by Fraser correction.

通过程序处理和手工提取衍射峰的方法相结合,由PE第0层线上提取出9个衍射峰,从第1层线上共提取出8个衍射峰.将数据转化成 $2 \times 10$ 的二维数组后输入程序(未满10个数据点的使用0补齐).程序运行部分结果如表1所示:

程序运行时间为1.73 s,完全满足应用需求.在参数空间中,可能存在多种参数组合以满足不同的指标化组合.程序会根据总误差的大小对计算得到的可能晶胞进行排序.计算出可能的晶胞参数如表2所示:

Table 1 Program-generated PE unit cell parameters and diffraction plane indexing results.

$a=4.969 \text{ \AA}$ $b=7.469 \text{ \AA}$ $c=2.569 \text{ \AA}$ Total error: 0.000031				
Equator				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[1. 1. 0.]	0.058538	0.058432	4.133	4.137
[0. 2. 0.]	0.072127	0.071707	3.724	3.734
[1. 2. 0.]	0.112422	0.112212	2.982	2.985
[2. 0. 0.]	0.163155	0.162019	2.476	2.484
[2. 1. 0.]	0.180817	0.179946	2.352	2.357
[1. 3. 0.]	0.202590	0.201846	2.222	2.226
[2. 2. 0.]	0.234740	0.233726	2.064	2.068
[0. 4. 0.]	0.287676	0.286830	1.864	1.867
[2. 3. 0.]	0.325046	0.323361	1.754	1.759
1 <sup>st</sup> Layer				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[1. 0. 1.]	0.193419	0.192041	2.274	2.282
[1. 1. 1.]	0.211382	0.209968	2.175	2.182
[0. 2. 1.]	0.224504	0.223244	2.111	2.116
[1. 2. 1.]	0.265462	0.263748	1.941	1.947
[2. 1. 1.]	0.333517	0.331482	1.732	1.737
[1. 3. 1.]	0.354756	0.353383	1.679	1.682
[2. 2. 1.]	0.387149	0.385263	1.607	1.611
[0. 4. 1.]	0.440761	0.438366	1.506	1.510

**Table 2** Lattice indices, observed/calculated  $d$ -spacings, and structure factors of PE.

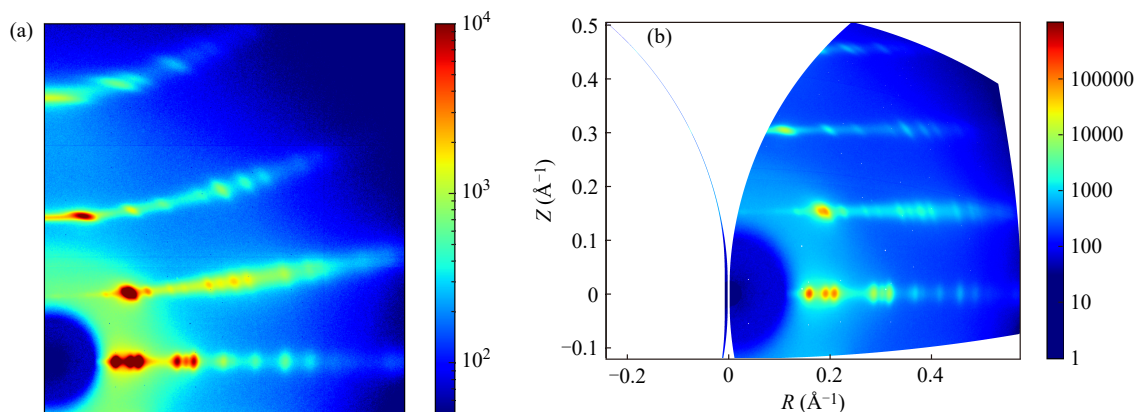
Unit cell parameters ( $\alpha=\beta=\gamma=90^\circ$ )			Total error
$a=4.969 \text{ \AA}$	$b=7.469 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000031
$a=4.156 \text{ \AA}$	$b=8.844 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000474
$a=4.594 \text{ \AA}$	$b=11.344 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000624
$a=4.344 \text{ \AA}$	$b=14.531 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000550
$a=5.031 \text{ \AA}$	$b=7.406 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000136
$a=8.219 \text{ \AA}$	$b=8.969 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000184
$a=8.969 \text{ \AA}$	$b=8.219 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000184
$a=8.156 \text{ \AA}$	$b=9.281 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000163
$a=8.344 \text{ \AA}$	$b=12.906 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000159
$a=8.469 \text{ \AA}$	$b=14.844 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000249
$a=8.031 \text{ \AA}$	$b=19.344 \text{ \AA}$	$c=2.569 \text{ \AA}$	0.000152

表2中, 总误差最小的晶胞即对应了真实的晶胞参数(第一行). 结合文献, PE的晶胞参数

为 $a=4.93 \text{ \AA}$ ,  $b=7.40 \text{ \AA}$ ,  $c=2.54 \text{ \AA}$ ,  $\alpha=\beta=\gamma=90^\circ$  [15], 与指标化程序得到的晶胞参数误差很小. 提取出的衍射峰全部成功指标化, 且与文献报道的晶面指数一致. 此处, 除了正确的晶胞参数外, 程序也运行得到了多组“错误的晶胞参数”. 但这些额外的结果是有必要的, 因为晶体结构解析的过程存在较大不确定性, 随时可能由于晶胞不正确而推翻重来, 因此, 程序给出其他结果也体现了其穷举算法对所有可能结果筛选的完备性. 总之, PE的指标化获得成功, 表明程序可对正交晶系晶体进行快速有效的指标化处理.

## 2.2 PP衍射图像分析和程序指标化

等规聚丙烯在常规降温结晶过程中产生 $\alpha$ 晶, 其晶胞属单斜晶胞,  $\beta \neq 90^\circ$ . 其衍射图像与Fraser校正图像如图7所示.



**Fig. 7** *i*PP fiber diffraction: (a) stitched diffraction pattern, (b) diffraction pattern treated by Fraser correction.

*i*PP分子链存在 $3_1$ 螺旋, 因此其 $c$ 轴长度较PE长. 图中可见4个层线, 但第三层线的衍射点较为模糊不清, 因此提取第0,1,2层线上共25个衍射点, 用于指标化处理. 根据对图7(b)的积分, 可计算出 $c$ 轴长度约为 $6.61 \text{ \AA}$ . 将其代入程序进行计算. 程序运行时长约 $8.18 \text{ s}$ , 体现出在大运算

量下的优化效果. 其部分运行结果如下(筛选出的可能的晶胞参数组合):

*i*PP  $\alpha$ 晶晶胞参数文献报道值为 $a=6.65 \text{ \AA}$ ,  $b=20.96 \text{ \AA}$ ,  $c=6.50 \text{ \AA}$ ,  $\alpha=\gamma=90^\circ$ ,  $\beta=99.10^\circ$  [16]. 通过程序运行得到的晶胞参数与文献值基本一致. 从表3中可看出, 第一组晶胞参数组合和第二组组合非

**Table 3** Lattice indices, observed/calculated  $d$ -spacings, and structure factors of *i*PP.

Unit cell parameters ( $\alpha=\gamma=90^\circ$ )				Total error
$a=6.721 \text{ \AA}$	$b=20.969 \text{ \AA}$	$c=6.613 \text{ \AA}$	$\beta=97.969^\circ$	0.000018
$a=6.668 \text{ \AA}$	$b=21.156 \text{ \AA}$	$c=6.613 \text{ \AA}$	$\beta=98.594^\circ$	0.000012
$a=10.454 \text{ \AA}$	$b=12.719 \text{ \AA}$	$c=6.613 \text{ \AA}$	$\beta=95.469^\circ$	0.000049
$a=11.817 \text{ \AA}$	$b=12.719 \text{ \AA}$	$c=6.613 \text{ \AA}$	$\beta=118.281^\circ$	0.000050
$a=12.930 \text{ \AA}$	$b=12.656 \text{ \AA}$	$c=6.613 \text{ \AA}$	$\beta=126.406^\circ$	0.000050
$a=12.626 \text{ \AA}$	$b=19.219 \text{ \AA}$	$c=6.613 \text{ \AA}$	$\beta=97.031^\circ$	0.000050
$a=12.802 \text{ \AA}$	$b=20.969 \text{ \AA}$	$c=6.613 \text{ \AA}$	$\beta=93.281^\circ$	0.000023

常接近, 总误差也很接近, 出现这种情况的原因是由于不同晶面的间距非常接近. 在这种情况下需要对晶体的对称性进行研究, 通过系统消光性来再次进行筛选. 表4中我们筛选了正确的晶胞参数组合并列出了其对应的晶面指标结果. 除正

确的晶胞参数组合外, 程序还发现了其他可满足指标化的晶胞参数组合, 这与晶体解析过程中的反复试错和验证原理相同, 即晶体结构解析必须对某些未知数进行假设, 才能通过有限的方程解出可能的正确值.

**Table 4** Program-generated iPP unit cell parameters and diffraction plane indexing results.

$a=6.721 \text{ \AA}$ $b=20.969 \text{ \AA}$ $c=6.613 \text{ \AA}$ $\beta=97.969^\circ$ Total error: 0.000018				
Equator				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[1. 1. 0.]	0.025043	0.024845	6.319	6.344
[0. 4. 0.]	0.036435	0.036389	5.239	5.242
[1. 3. 0.]	0.043069	0.04304	4.819	4.820
[0. 6. 0.]	0.081168	0.081876	3.510	3.495
[2. 1. 0.]	0.091689	0.092556	3.302	3.287
[2. 2. 0.]	0.099596	0.099379	3.169	3.172
[1. 7. 0.]	0.133998	0.134013	2.732	2.732
[2. 6. 0.]	0.172006	0.172158	2.411	2.410
[3. 1. 0.]	0.206805	0.205409	2.199	2.206
[2. 8. 0.]	0.233982	0.23584	2.067	2.059
1 <sup>st</sup> Layer				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[0. 4. 1.]	0.05873	0.059704	4.126	4.093
[1. 3. 1.]	0.073325	0.072715	3.693	3.708
[-1. 5. 1.]	0.096189	0.096383	3.224	3.221
[-2. 2. 1.]	0.110564	0.109973	3.007	3.015
[-2. 4. 1.]	0.137836	0.137265	2.694	2.699
[-1. 7. 1.]	0.15098	0.150967	2.574	2.574
[0. 8. 1.]	0.166627	0.168872	2.450	2.433
[-2. 6. 1.]	0.181906	0.182752	2.345	2.339
2 <sup>nd</sup> Layer				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[-1. 1. 2.]	0.10435	0.105383	3.096	3.080
[1. 1. 2.]	0.131941	0.130825	2.753	2.765
[-2. 0. 2.]	0.157644	0.158099	2.519	2.515
[-2. 4. 2.]	0.193997	0.194489	2.270	2.268
[-2. 5. 2.]	0.215501	0.214958	2.154	2.157
[-2. 6. 2.]	0.239953	0.239976	2.041	2.041
[-3. 0. 2.]	0.258275	0.258231	1.968	1.968

通过指标化 iPP 也验证了程序在单斜晶系指标化中的高准确度.

### 2.3 PBS 衍射图像分析与指标化

由于单斜晶体相对复杂, 因此还需更多案例对其进行测试. PBS 在简单的熔体降温结晶过程

中形成  $\alpha$  晶, 属于单斜晶,  $\beta \neq 90^\circ$ . 相关衍射图像与 Fraser 校正图像如图 8 所示.

PBS 衍射图中可见 6 个层线, 但第五层线上衍射点清晰度欠佳, 因此只提取了第 0 层线至 4 层线上的衍射点, 共 25 个衍射峰. 对 Fraser 校正

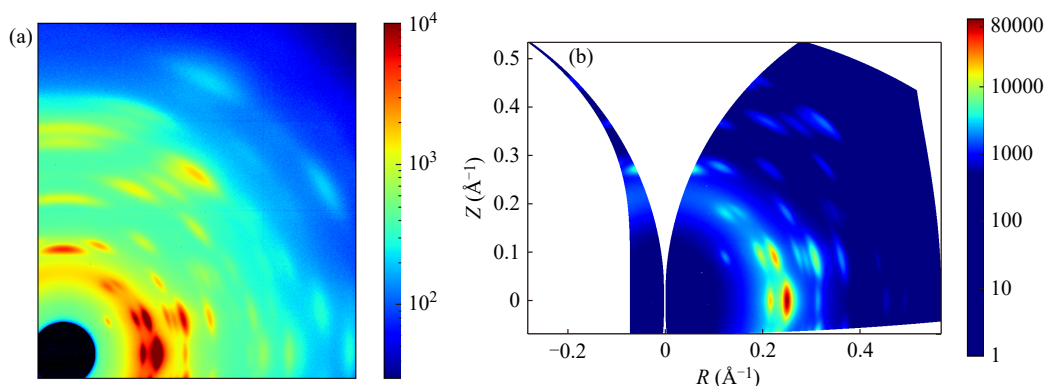


Fig. 8 PBS fiber diffraction: (a) stitched diffraction pattern, (b) diffraction pattern treated by Fraser correction.

图像进行积分, 得到其  $c$  轴长度约为  $10.94 \text{ \AA}$ . 通过提取这些峰的衍射角度并计算出其晶面间距平方的倒数, 输入程序进行计算. 计算得到的可能晶胞参数和接近正确的晶胞参数对应的衍射指标如表 5 和表 6 所示:

PBS  $\alpha$  晶胞参数文献报道值为  $a=5.23 \text{ \AA}$ ,  $b=9.15 \text{ \AA}$ ,  $c=10.90 \text{ \AA}$ ,  $\alpha=\gamma=90^\circ$ ,  $\beta=123.90^\circ$  [17].

第一个晶胞参数组合与 PBS 真实晶胞相似. 我们注意到, 较大的晶胞可能反而误差更小. 但从衍射峰指标化的经验来看, 某一层线较为靠前的衍射峰通常指数也不会太高, 所以在选择可能的晶胞参数时, 会优先考虑较小的晶胞. 但相较于 PE 和 PP 的指标化, PBS 的计算结果误差较大, 效果差一些.

Table 5 Lattice indices, observed/calculated  $d$ -spacings, and structure factors of PBS.

Unit cell parameters ( $\alpha=\gamma=90^\circ$ )				Total Error
$a=4.990 \text{ \AA}$	$b=9.156 \text{ \AA}$	$c=10.938 \text{ \AA}$	$\beta=123.594^\circ$	0.001022
$a=9.836 \text{ \AA}$	$b=9.156 \text{ \AA}$	$c=10.938 \text{ \AA}$	$\beta=119.843^\circ$	0.000408
$a=10.279 \text{ \AA}$	$b=9.156 \text{ \AA}$	$c=10.938 \text{ \AA}$	$\beta=123.906^\circ$	0.000258
$a=9.281 \text{ \AA}$	$b=17.406 \text{ \AA}$	$c=10.938 \text{ \AA}$	$\beta=92.969^\circ$	0.000036
$a=9.281 \text{ \AA}$	$b=20.031 \text{ \AA}$	$c=10.938 \text{ \AA}$	$\beta=98.593^\circ$	0.000201
$a=9.218 \text{ \AA}$	$b=21.781 \text{ \AA}$	$c=10.938 \text{ \AA}$	$\beta=99.218^\circ$	0.000065
$a=9.227 \text{ \AA}$	$b=21.906 \text{ \AA}$	$c=10.938 \text{ \AA}$	$\beta=105.156^\circ$	0.000358
.....				

Table 6 Program-generated PBS unit cell parameters and diffraction plane indexing results.

$a=4.990 \text{ \AA}$ $b=9.156 \text{ \AA}$ $c=10.938 \text{ \AA}$ $\beta=123.594^\circ$ Total error: 0.001022				
Equator				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[0. 2. 0.]	0.047647	0.047712	4.581	4.578
[1. 1. 0.]	0.063374	0.069817	3.972	3.785
[1. 2. 0.]	0.099199	0.105601	3.175	3.077
[2. 0. 0.]	0.208262	0.231556	2.191	2.078
[1. 4. 0.]	0.243266	0.248736	2.027	2.005
1 <sup>st</sup> Layer				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[0. 1. 1.]	0.023711	0.023973	6.494	6.459
[-1. 1. 1.]	0.047864	0.052641	4.571	4.359
[0. 2. 1.]	0.059125	0.059757	4.113	4.091
[-1. 2. 1.]	0.083557	0.088425	3.459	3.363

Table 6 Continued.

$a=4.990 \text{ \AA}$ $b=9.156 \text{ \AA}$ $c=10.938 \text{ \AA}$ $\beta=123.594^\circ$ Total error: 0.001022				
1 <sup>st</sup> Layer				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[1. 0. 1.]	0.091144	0.099156	3.312	3.176
[1. 1. 1.]	0.103207	0.111084	3.113	3.000
[1. 2. 1.]	0.13919	0.146867	2.680	2.609
2 <sup>nd</sup> Layer				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[-1. 1. 2.]	0.055045	0.055837	4.262	4.232
[1. 0. 2.]	0.154323	0.153898	2.546	2.549
3 <sup>rd</sup> Layer				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[-1. 0. 3.]	0.076361	0.078633	3.619	3.566
[-1. 1. 3.]	0.089544	0.090561	3.342	3.323
[0. 1. 3.]	0.120079	0.120336	2.886	2.883
[-1. 3. 3.]	0.184161	0.185985	2.330	2.319
4 <sup>th</sup> Layer				
Index	$Q_{\text{obs}} (\text{\AA}^{-2})$	$Q_{\text{calc}} (\text{\AA}^{-2})$	$d_{\text{obs}} (\text{\AA})$	$d_{\text{calc}} (\text{\AA})$
[-1. 2. 4.]	0.180655	0.181441	2.353	2.348
[0. 1. 4.]	0.204502	0.204653	2.211	2.211
[0. 2. 4.]	0.240041	0.240437	2.041	2.039

上表中部分晶面指标不正确, 例如(-112)和(102)晶面, 应为(002)和(-122). 这可能与衍射峰位置过于接近或重合, 以及晶粒尺寸过小而造成衍射峰半峰宽, 其精确位置的测量出现较大误差有关. 这些衍射峰的精确指标, 将在后续空间群确定后通过对称性与消光进行修正. 尽管如此, 程序依然给出了基本正确的晶胞参数和晶面指标, 且只需 10.15 s 即完成计算过程, 表现出较高的效率与准确度.

虽然程序可以得到正确的晶胞参数, 对于未知体系, 未知晶胞参数时, 还需要结合对称性, 空间群与系统消光规律进行判定, 从中筛选出正确的晶胞参数. 但程序的引入无疑大幅减少了指标化所需要的时间与人力, 加快了晶体解析的全流程工作速度.

### 3 结论

主要基于 Python 语言及其数据运算模块成

功编写了可用于聚合物纤维衍射指标化的自动化程序. 程序采用二分递进法作为算法核心, 并通过高分子纤维衍射特有的衍射几何对其数学原理进行了重新推导, 优化了其多元函数极值判定条件. 程序通过简单的参数输入便可实现正交晶系与单斜晶系的快速指标化. 同时我们使用已知晶胞参数的 3 种聚合物(PE, *i*PP, PBS)等对程序的运算效率和可靠性进行了验证. 验证结果表明, 程序对 3 种聚合物的晶胞预测均表现出很高的准确率, 其给出的可能正确的晶胞中都包含实际正确的晶胞参数, 衍射峰所对应的晶面指数也都可由程序计算得出. 此外, 程序运算效率高, 在运行正交晶系解析时, 其运行时间在普通便携式计算机上即可达到秒级, 在运行单斜晶系解析时, 其运行时间可达十秒级, 具有很高的应用价值.

致谢 感谢 Müller 教授提供 PBS 样品.



**作者简介:** 刘国明, 男, 1983年生, 中国科学院化学研究所研究员. 2006年本科毕业于四川大学, 2011年博士毕业于中国科学院化学研究所, 曾在德国 Leibniz 聚合物研究所和 Max-Planck 高分子研究所访问学习, 2016~2018在牛顿学者项目资助下在剑桥大学卡文迪许实验室开展博士后研究工作. 主要从事高分子物理研究, 发表论文 100 余篇, 包括 *PRL*、*JACS*、*Macromolecules*、*Adv. Mater.* 等. 曾获冯新德高分子奖提名奖、中国化学会高分子科学创新论文奖、北京市自然科学二等奖等.

## REFERENCES

- 1 扈健, 王梦梵, 吴婧华. X 射线晶体结构解析技术在高分子表征研究中的应用. *高分子学报*, **2021**, 52(10), 1390-1405.
- 2 Christopher, J.; Gilmore, James, A. Kaduk; Henk Schenk. *International Tables for Crystallography, Volume H, Powder Diffraction*. Hoboken: Wiley Press, **2019**.
- 3 Otwinowski, Z.; Minor, W. I. *Methods in Enzymology, Vol. 276, Macromolecular Crystallography, Part A*. New York: Academic Press, **1997**.
- 4 Shi, H. L.; Li, Z. A. Unitcell tools, a package to determine unit-cell parameters from a single electron diffraction pattern. *IUCrJ*, **2021**, 8(5), 805-813.
- 5 Louër, D.; Boultif, A. Powder pattern indexing and the dichotomy algorithm. *Z. Kristallogr. Suppl.*, **2007**, 2007(suppl\_26), 191-196.
- 6 Louër, D.; Louër, M. Méthode d'essais et erreurs pour l'indexation automatique des diagrammes de poudre. *J. Appl. Crystallogr.*, **1972**, 5(4), 271-275.
- 7 Altomare, A.; Giacovazzo, C.; Guagliardi, A.; Moliterni, A. G. G.; Rizzi, R.; Werner, P. E. New techniques for indexing: N-treor in expo. *J. Appl. Crystallogr.*, **2000**, 33(4), 1180-1186.
- 8 Visser, J. W. A fully automatic program for finding the unit cell from powder data. *J. Appl. Crystallogr.*, **1969**, 2(3), 89-95.
- 9 Kariuki, B. M.; Belmonte, S. A.; McMahon, M. I.; Johnston, R. L.; Harris, K. D. M.; Nelmes, R. J. A new approach for indexing powder diffraction data based on whole-profile fitting and global optimization using a genetic algorithm. *J. Synchrotron Radiat.*, **1999**, 6(2), 87-92.
- 10 Tashiro, K.; Asanaga, H.; Ishino, K.; Tazaki, R.; Kobayashi, M. Development of a new software for the X-ray structural analysis of polymer crystals by utilizing the X-ray imaging plate system. *J. Polym. Sci. B Polym. Phys.*, **1997**, 35(11), 1677-1700.
- 11 Boultif, A.; Louër, D. Indexing of powder diffraction patterns for low-symmetry lattices by the successive dichotomy method. *J. Appl. Crystallogr.*, **1991**, 24(6), 987-993.
- 12 Boultif, A.; Louër, D. Powder pattern indexing with the dichotomy method. *J. Appl. Crystallogr.*, **2004**, 37(5), 724-731.
- 13 Ma, T. Y.; Hu, W. X.; Wang, D. J.; Liu, G. M. A global optimization approach to automated indexing of fiber diffraction patterns. *J. Appl. Crystallogr.*, **2025**, 58(3), 759-767.
- 14 Fraser, R. D. B.; MacRae, T. P.; Miller, A.; Rowlands, R. J. Digital processing of fiber diffraction patterns. *J. Appl. Crystallogr.*, **1976**, 9(2), 81-94.
- 15 Bunn, C. W. The crystal structure of long-chain normal paraffin hydrocarbons. The "shape" of the CH<sub>2</sub> group. *Trans. Faraday Soc.*, **1939**, 35, 482-491.
- 16 Natta, G.; Corradini, P. *Stereoregular Polymers and Stereospecific Polymerizations, vol. 2, Structure and Properties of Isotactic Polypropylene*. Oxford: Pergamon Press **1967**.
- 17 Ichikawa, Y.; Kondo, H.; Igarashi, Y.; Noguchi, K.; Okuyama, K.; Washiyama, J. Crystal structures of  $\alpha$  and  $\beta$  forms of poly(tetramethylene succinate). *Polymer*, **2000**, 41(12), 4719-4727.

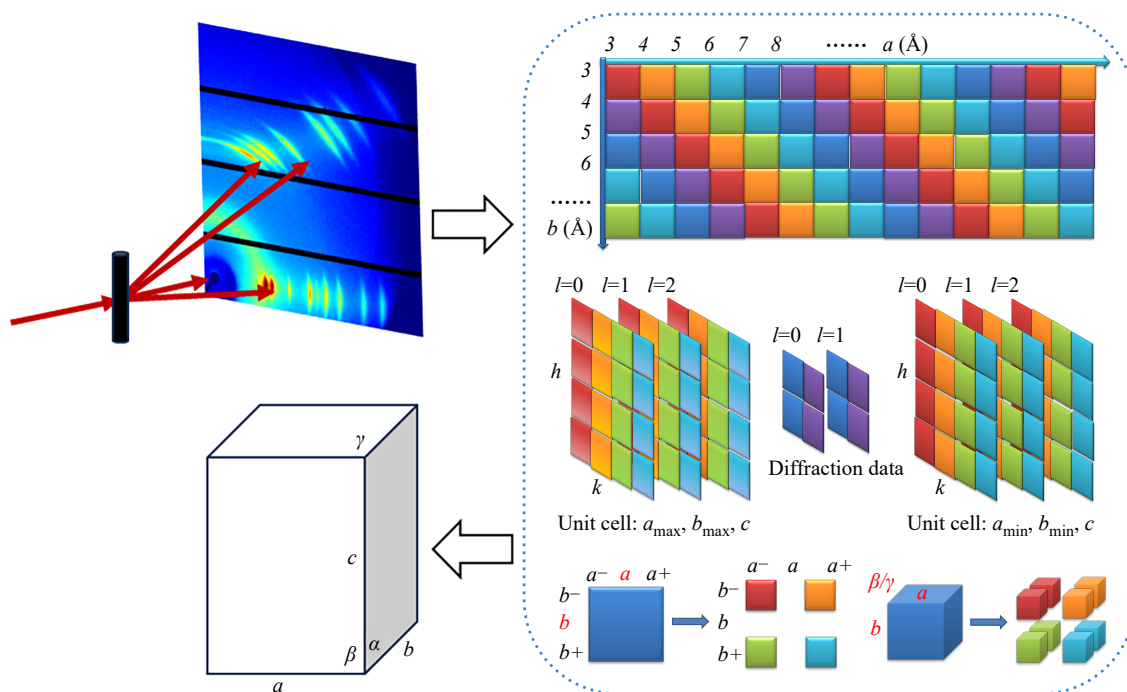
## Indexing Program for Crystalline Polymer Fiber Patterns Based on Dichotomy Method

Wen-xian Hu<sup>1,2</sup>, Tian-yi Ma<sup>1,2</sup>, Guo-ming Liu<sup>1,2\*</sup>

(<sup>1</sup>CAS Key Laboratory of Engineering Plastics, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100190)

(<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049)

**Abstract** Indexing, the process of deriving unit cell parameters from unknown diffraction patterns, is a crucial step in crystal structure determination. However, the crystallization behavior of polymers makes it difficult to accomplish crystal structure analysis, including indexing, using fully automated software. To address this issue, we developed an automated Python-based program for indexing polymer fiber diffraction patterns using the dichotomy method as the core algorithm. With simple parameter inputs, the program enables the rapid indexing of orthorhombic and monoclinic polymer systems. To validate its reliability and efficiency, the program was tested on several polymers with known unit-cell parameters. The results demonstrate excellent accuracy and rapid computational performance, indicating their suitability for practical applications.



**Keywords** Polymer crystal; Crystal structure; Indexing; Dichotomy; Fiber pattern

\* Corresponding author: Guo-ming Liu, E-mail: gmliu@iccas.ac.cn