

光学学报

基于改进的自适应迭代重加权惩罚最小二乘的空间外差拉曼光谱基线校正方法

白云飞^{1,2,3}, 罗海燕^{1,2,3}, 李志伟^{1,3}, 丁毅^{1,3}, 熊伟^{1,2,3*}¹中国科学院合肥物质科学研究院安徽光学精密机械研究所, 安徽 合肥 230031;²中国科学技术大学, 安徽 合肥 230026;³中国科学院通用光学定标与表征技术重点实验室, 安徽 合肥 230031

摘要 空间外差拉曼光谱技术因其非接触、无损、快速、高稳定性和高光谱分辨率等特点,已经广泛应用于各个物质探测领域。由于复原光谱中存在荧光背景干扰,对样品进行定性和定量分析时需要光谱进行基线校正。为了解决由拉曼峰引起的拟合基线抬升的问题,提出了一种改进的自适应迭代重加权惩罚最小二乘(airPLS)的基线拟合方法,即基于拉曼峰截断的 airPLS 基线拟合方法。该方法能够自动识别拉曼峰,并在对光谱进行拉曼峰截断后进行 airPLS 迭代拟合,以获得更准确的基线。使用仿真光谱和实测光谱进行验证,并与常见的基线拟合方法进行对比,结果显示,改进的 airPLS 基线拟合准确度显著提升,仿真光谱的基线拟合均方根误差优于 0.0052。实测拉曼光谱的校正谱特征峰清晰可见,荧光基线趋势被有效去除,满足拉曼光谱数据处理的需求。

关键词 拉曼光谱; 空间外差; 基线校正; 最小二乘法

中图分类号 O657.3

文献标志码 A

DOI: 10.3788/AOS231913

1 引言

拉曼光谱技术是一种基于物质对光的非弹性散射的无损分析技术^[1]。空间外差拉曼光谱技术将拉曼光谱技术和空间外差光谱技术相结合,同时具有拉曼光谱技术无损、快速和非接触的优势和空间外差光谱技术高稳定性、高光通量、高光谱分辨率和高探测灵敏度的优势^[2]。在诱导激发拉曼信号的同时,会不可避免地产生荧光信号。尽管荧光光谱技术也具有广泛的应用前景,但在拉曼光谱样品分析中,荧光背景直接影响了对样品拉曼特性的表征。常见的消荧光背景措施主要从软硬件两种渠道实现,硬件方法主要包括差频激发拉曼^[3]、深紫外拉曼和时间分辨拉曼^[4]等,这些方法虽然具有较好的效果,但仪器结构复杂且成本高。软件方法主要是指采用信号处理手段对拉曼光谱进行荧光背景扣除。拉曼光谱具有非连续和谱峰尖锐的特点,而荧光光谱大多连续且趋势平缓。针对二者光谱特性的差异,在进行拉曼光谱分析时,采用基线校正算法消除荧光干扰,可以确保拉曼光谱数据的可靠性和准确性。

常见的基线拟合方法包括平滑样条法、多项式拟

合、离散小波变换、形态学算法、变分模态分解、最小二乘法和神经网络等。平滑样条法是基于插值的原理通过估计荧光背景所属的拉曼光谱位置拟合基线,其易受到噪声的影响,无法处理具有强噪声的拉曼光谱^[5]。多项式拟合方法通过迭代切除信号峰或线性约束来适应基线^[6],但其依赖于多项式阶数,若多项式参数设置错误,就无法保证拟合结果。这意味着使用者必须手动检查每个光谱,限制了基线校正的全自动实现。基于离散小波变换的基线校正方法依赖于小波的频率分解能力^[7],这需要在基线信号中进行低频分解。由于基线随测试样品和激发条件的变化而变化,很难选择最佳的分解方法和相关的优化方法,且计算量巨大。形态学算法依赖于对信号和基线具有不同特征宽度的假设^[8],它们通过在图像处理中使用膨胀和腐蚀操作来实现基线拟合,然而没有可操作的标准来将光谱的宽尾部分与基线拐点分离开来。变分模态分解^[9]是一种信号分解估计方法,通过迭代搜寻变分模型最优解来确定每个分量的频率中心和带宽,从而自适应地实现信号的频域分离,并将低频信号作为光谱拟合基线。但其受到频率分解数量的影响,分解结果差异较大。其他方法包括使用遗传算法进行节点选择^[10]、切角^[11]

收稿日期: 2023-12-11; 修回日期: 2023-12-30; 录用日期: 2024-01-10; 网络首发日期: 2024-02-20

基金项目: 国家重点研发计划(2022YFB3901800, 2022YFB3901803)、国家自然科学基金(41975033, 61975212)、中国科学院重点部署项目(JCPYJJ-22010)、中国科学院合肥研究院院长基金(YZJJ202210-TS)

通信作者: *frank@aiofm.ac.cn

和导数传递累加^[12]。在实践中,这些方法的参数选择受到信号峰值的影响。这使得难以确保对各种信号形状进行有效的基线校正。虽然深度神经网络可以在基线校正中获得良好的结果^[13],但它需要数据库的支持、大量的训练集和复杂的计算。此外,其训练结果缺乏数学解释。

自适应迭代重加权惩罚最小二乘(airPLS)^[14]算法将最小二乘平滑与基线估计中的非平滑惩罚相结合,被认为是使用最广泛的基线校正算法。它具有自适应性和迭代加权惩罚特性,能够自适应光谱数据中的噪声和背景,减少人为干预。引入惩罚项迭代加权可提高基线拟合的准确性。当拟合的基线远高于信号时,权重增加以使拟合的基线快速接近真实信号。当拟合的基线远低于信号(通常出现在拉曼峰区域),权重增加以在强弱峰值处平衡拟合基线的平滑性,从而避免由局部过度平滑造成的基线估计漂移。但在拉曼峰区域,谱峰的存在将导致拟合基线抬升,并由此造成校正光谱的拉曼峰强度下降,给后续光谱分析和处理造成不利影响。

本文提出了一种拉曼峰截断的 airPLS 拟合方法,通过识别拉曼峰位置,截断拉曼峰区域,并采用 airPLS 算法进行基线拟合,降低拉曼峰区域强度突变造成的拟合基线抬升。仿真光谱实验表明,该方法优于其他先进的基线校正方法。此外,使用实测拉曼光谱进行的实验进一步证明了该方法的有效性。

2 原理和方法

2.1 airPLS 原理

airPLS 假设信号 \mathbf{y} 的长度为 N ,在等间隔采样下得到,同时 \mathbf{z} 为拟合基线,拟合的基线应该跟随 \mathbf{y} 具有相同的趋势,同时保持平滑性。假设 \mathbf{y} 和 \mathbf{z} 为列向量,可以通过求取一个带惩罚项的最小二乘函数 \mathbf{F} 的最小值来确定 \mathbf{z} 。

$$\mathbf{F} = \sum_{i=1}^N \mathbf{w}_i (\mathbf{y}_i - \mathbf{z}_i)^2 + \lambda \sum_{i=1}^N (\mathbf{z}_i'')^2, \quad (1)$$

式中: i 是信号采样位置的索引; \mathbf{y}_i 和 \mathbf{z}_i 分别是向量 \mathbf{y} 和 \mathbf{z} 中第 i 个采样点的强度; \mathbf{w}_i 是每个采样位置控制拟合基线和信号接近程度的权重。式(1)中的第一项和第二项分别代表光谱信号的保真度和 \mathbf{z} 的平滑程度, λ 是平衡这两项的参数。也可以用矩阵形式表示为

$$\mathbf{F} = (\mathbf{y} - \mathbf{z})^T \mathbf{W} (\mathbf{y} - \mathbf{z}) + \lambda \mathbf{z}^T \mathbf{D}^T \mathbf{D} \mathbf{z}, \quad (2)$$

式中: \mathbf{W} 是一个对角矩阵,权重向量在它的对角线上; \mathbf{D} 是一个二阶差分矩阵:

$$\mathbf{D} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix}. \quad (3)$$

通过求偏导数向量并令其值为 0,即 $\partial \mathbf{F} / \partial \mathbf{z}^T = 0$,得到式(2)的最小值解如下:

$$\frac{\partial \mathbf{F}}{\partial \mathbf{z}^T} = -2\mathbf{W}(\mathbf{y} - \mathbf{z}) + 2\lambda \mathbf{D}^T \mathbf{D} \mathbf{z} = 0, \quad (4)$$

$$\mathbf{z} = (\mathbf{W} + \lambda \mathbf{D}^T \mathbf{D})^{-1} \mathbf{W} \mathbf{y}. \quad (5)$$

自适应迭代重加权过程根据迭代次数使用不同的方法来分配权重并添加惩罚来控制拟合基线的平滑性。权重向量 \mathbf{w}_i 随迭代步骤 t 的变化由下式求得:

$$\mathbf{w}_i = \begin{cases} 0, & \mathbf{y}_i \geq \mathbf{z}_i^{t-1} \\ \frac{t(\mathbf{y}_i - \mathbf{z}_i^{t-1})}{e^{-|\mathbf{d}^-|}}, & \mathbf{y}_i < \mathbf{z}_i^{t-1} \end{cases}, \quad (6)$$

式中,向量 \mathbf{d}^- 由相减 $\mathbf{y} - \mathbf{z}$ 的负元素组成。在之前的 $t-1$ 次迭代中,拟合向量 \mathbf{z} 是基线的候选。如果一个信号 \mathbf{y}_i 大于基线的候选值 \mathbf{z}_i ,则认为它是峰值的一部分,令其权值设为 0。否则根据式(6)调整权值。当迭代次数达到最大值或满足下式终止条件时,迭代停止:

$$|\mathbf{d}^-| < \epsilon \times |\mathbf{y}_i|, \quad (7)$$

式中, ϵ 为阈值系数,通常取 0.001^[15]。

加权方案是 airPLS 的关键,它决定了估计的基线是否符合实际的拉曼光谱。但是,受到拉曼峰强度的影响,拟合基线会高于理论基线,这将对后续定性和定量分析的准确性造成影响。

2.2 拉曼峰截断的 airPLS

为进一步提升 airPLS 的基线拟合精度,提出了一种基于拉曼峰截断的 airPLS 基线拟合方法。具体步骤如下:

1) 利用 Savitzky-Golay (S-G) 滤波器对拟合光谱信号 s 进行去噪,得到 y_0 。

2) 使用一阶导数对 y_0 进行寻峰,得到拉曼峰对应的拉曼频移 Peak-positions。以每个拉曼峰横坐标为中心,分别向左右两侧寻得极小值点 L_{\min} 和 R_{\min} 。

3) 将两极小值点之间作为拉曼峰区域进行截断,其余区域信号保留得到 L_n 。对 L_n 进行 airPLS 拟合,得到新的基线 L_{n+1} 。若 L_{n+1} 与 L_n 相同,则将 L_{n+1} 视为原光谱信号的最优拟合基线,基线校正后的信号为 $s' = s - L_{n+1}$;若 L_{n+1} 与 L_n 不同,则将 L_{n+1} 和 L_n 相减后求绝对值,将绝对值大于阈值 P 的 L_{n+1} 截断,得到新的待拟合信号 L_n ,重复 airPLS 拟合直至两次拟合基线相同,输出 L_{n+1} 。则基线校正后的信号为 $s' = s - L_{n+1}$ 。其中 P 为 L_{n+1} 与 L_n 之差的绝对值的均值,两次基线拟合相同的判别标准为

$$\sum_{i=1}^N \left| \frac{L_n(i) - L_{n-1}(i)}{L_n(i)} \right| < \varphi. \quad (8)$$

算法流程如图 1 所示。

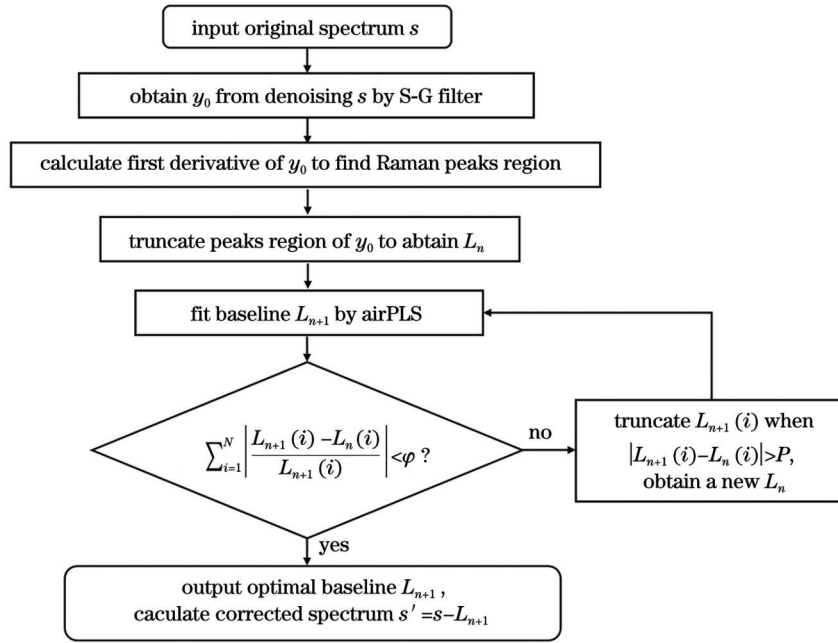


图1 算法流程图

Fig. 1 Flow chart of algorithm

3 仿真实验

3.1 仿真光谱基线拟合实验

由于拉曼探测中真实的荧光基线未知,因此无法对所提出的基线拟合方法进行定量评价。采用模拟拉曼光谱进行仿真实验。对于模拟拉曼光谱数据 s , 可以用数学方式描述如下:

$$s(i) = R(i) + l(i) + n(i), \quad (9)$$

式中: $R(i)$ 表示仿真理论拉曼光谱; $l(i)$ 表示仿真基线; $n(i)$ 表示随机噪声。一般来说, 仿真理论拉曼光谱由多个相互叠加的洛伦兹函数^[16]或者高斯函数^[17]组成, 或者使用二者之和表示^[18]。但在空间外差拉曼光谱中, 复原光谱由干涉图经傅里叶变换后得到。空间外差干涉光谱仪的仪器线型函数可以认为是对单色光干涉图进行复原后的光谱曲线, 该光谱曲线以 sinc 函数的形式^[19]表达为

$$R(i) = \sum_i I(i) \cdot \text{sinc} \left\{ \frac{\pi U_{\max} [R_s - P(i)]}{W(i)} \right\}^2, \quad (10)$$

式中: $I(i)$ 为拉曼峰强度; $P(i)$ 为拉曼峰位置; R_s 为拉曼峰对应的拉曼频移; $W(i)$ 为拉曼峰宽; U_{\max} 为最大光程差, 与光栅照明宽度 W_e 和 Littrow 波数 θ_L 有关:

$$U_{\max} = 2W_e \sin \theta_L. \quad (11)$$

实际拉曼探测中的荧光基线线型复杂, 不具备单一函数特征, 因此模拟基线数据采用四种混合类型函数: 正弦函数、指数函数、幂函数和多项式函数^[20], 如下式所示:

$$\begin{cases} l_1 = \alpha + \beta \sin(2\pi f \cdot R_s) \\ l_2 = \gamma e^{-\delta \cdot R_s} \\ l_3 = R_s^\rho \\ l_4 = aR_s^4 + bR_s^3 + cR_s^2 + dR_s + e \\ l(r) = l_1 + l_2 + l_3 + l_4 \end{cases}, \quad (12)$$

式中: a 和 β 分别表示正弦函数的偏置和幅值; f 为正弦函数频率; γ 和 δ 为指数函数的参数; ρ 为幂函数的指数项; $a \sim e$ 为多项式函数的各级次系数。仿真基线为四种类型函数之和。随机噪声使用随机函数生成, 通过单一变量设置噪声强度以获得不同信噪比的仿真拉曼光谱。

根据所搭建的空间外差拉曼探测实验装置, 为验证基线拟合方法在不同类型基线和不同信号质量下的适用性, 设置仿真参数如表 1 所示。使用式(10)中 sinc 函数生成仿真拉曼光谱: 6 个半峰全宽为 5.33 cm^{-1} 的

表1 仿真拉曼光谱参数设置

Table 1 Parameter settings of simulated Raman spectrum		
Parameter	Baseline-1	Baseline-2
a	1×10^4	1×10^4
β	4×10^6	8×10^6
f	1.5×10^{-4}	1
γ	9×10^5	1×10^6
δ	1×10^{-3}	1×10^{-3}
ρ	1.7	2
(a, b, c, d, e)	$(1.1 \times 10^{-7}, 5.6 \times 10^{-5}, -0.15, 110, 1 \times 10^6)$	$(1.1 \times 10^{-7}, 5.6 \times 10^{-5}, -0.15, 110, 1 \times 10^6)$

拉曼峰, 对应拉曼频移分别为 300、400、1000、1200、1400、1500 cm^{-1} , 相对强度分别为 1、2、5、3、4、2。分别生成具有单一下降趋势和复杂变化趋势的两种基线类型 (Baseline-1 和 Baseline-2)。引入不同强度的噪声, 生成信噪比为 5、10、50、100 的拉曼光谱共 8 个。仿真拉曼光谱如图 2 和图 3 所示。信噪比 S_{NR} 的计算方法如下式所示:

$$S_{\text{NR}} = \frac{s_{\text{peak}}}{s_{\text{id-noise}}}, \quad (13)$$

式中: s_{peak} 为无基线的拉曼峰高度; $s_{\text{id-noise}}$ 为无峰区域噪声标准差。以 300 cm^{-1} 处最弱拉曼峰的信噪比对光谱质量进行表征。

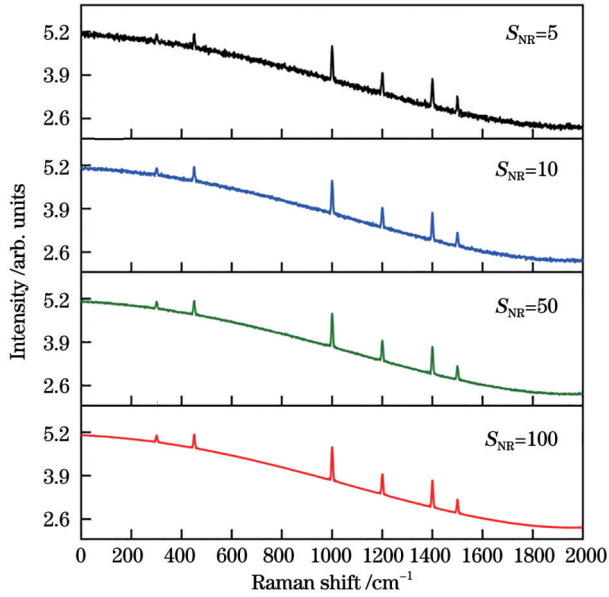


图 2 附带 Baseline-1 基线的仿真拉曼光谱
Fig. 2 Simulated Raman spectra with Baseline-1

将常用基线拟合方法与本文算法进行对比, 主要包括多项式拟合 (Polyfit)、离散小波变换 (DWT)、变分模态分解 (VMD) 和自适应迭代重加权惩罚最小二乘 (airPLS)。由于上述算法的基线拟合精度均受到参数设置的影响, 因此将拟合基线与理论基线的均方根误差 RMSE 作为评价标准, 对上述各算法中涉及的参数进行大范围、小步距的遍历, 确定最优参数配置

表 2 各个基线拟合算法中使用的最优参数配置

Table 2 Optimal parameter configuration used in each baseline fitting algorithm

Parameter	Baseline-1	Baseline-2
Polyfit ^a	4	7
DWT ^b	(db4, 9)	(db4, 7)
VMD ^c	(1×10^5 , 0, 10, 1, 1×10^{-6})	(1×10^5 , 0, 10, 1, 1×10^{-5})
airPLS ^d	(7.2×10^7 , 2, 0, 0, 1)	(2.68×10^7 , 2, 0, 0, 1)
Proposed ^e	(0.18, 5, 0.01, 7.2×10^7 , 2, 0, 0, 1)	(0.18, 5, 0.01, 2.68×10^7 , 2, 0, 0, 1)

Note: a: order; b: (wavelet name, level); c: (alpha, tau, K , init, tolerance); d: (lambda, order, wep, p , itermax); e: (min-height, min-distance, threshold, lambda, order, wep, p , itermax)

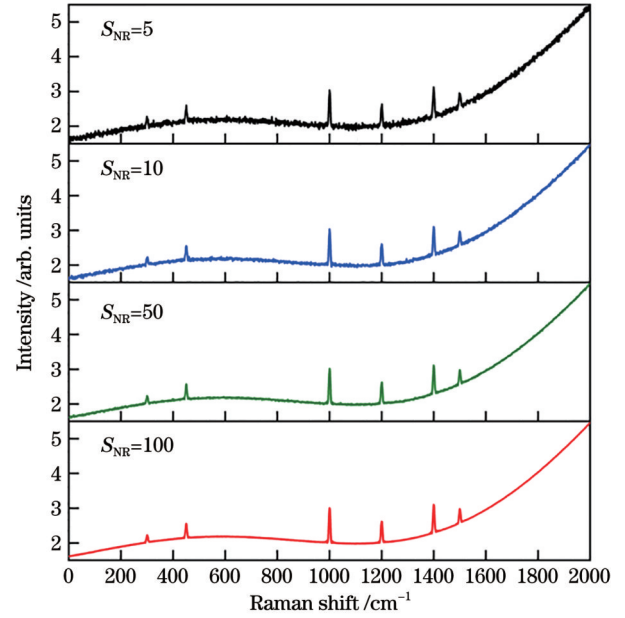


图 3 附带 Baseline-2 基线的仿真拉曼光谱
Fig. 3 Simulated Raman spectra with Baseline-2

如表 2 所示。

$$R_{\text{MSE}} = \sqrt{\sum_{i=1}^N [L_n(i) - l(i)]^2 / N}, \quad (14)$$

式中, 多项式拟合仅需要输入多项式阶数 order。DWT 需要输入函数名称 wavelet name 和小波变换级数 level。VMD 需要确定数据保真度 alpha、上升时间步长 tau、模态数量 K 、频率初始化参数 init 和收敛容差 tolerance。airPLS 中 lambda 为平滑参数, 较大的 lambda 值会增加平滑程度, 导致拟合的基线更加平滑, 但可能会忽略一些细节和小尺度的变化。较小的 lambda 值会减小平滑程度, 拟合的基线可能更接近原始数据, 但可能会受到噪声的干扰; order 为构建差分矩阵的阶数, 其值越大, 拟合基线越能够适应光谱数据的变化; wep 为权重控制参数, 调整其值能够合理分配信号两端和中部区域所占权重; p 为权重修正系数, 控制拟合的平滑程度; itermax 为最大迭代次数。本文采用拉曼峰截断的 airPLS 在寻峰过程中设置识别最小峰高为 0.18, 可识别峰的最小距离为 5 cm^{-1} 。

3.2 实测拉曼光谱基线拟合实验

使用自行研制的便携式空间外差拉曼光谱探测实验装置对高荧光样品口红(购自电商)进行拉曼探测,验证基线拟合算法在实测拉曼光谱中的拟合效果。实验布局如图 4 所示,激发波长为 785 nm,可探测拉曼频移范围为 300~2000 cm^{-1} ,光谱间隔为 5.33 cm^{-1} 。实验中设置积分时间为 5 s,激光功率为 500 mW,单次采集 10 幅干涉数据,平均后得到原始干涉图。经盲元校正^[21]、相位校正^[22]、三角窗切趾、充零、傅里叶变换和光谱平均后,得到复原光谱。

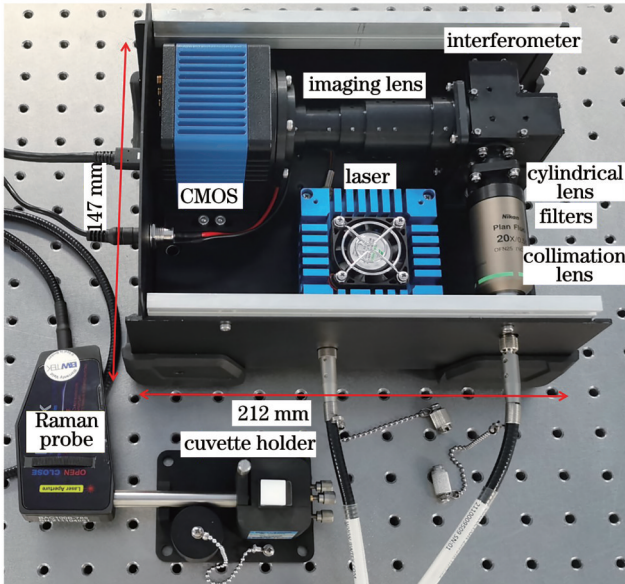


图 4 空间外差拉曼光谱探测实验现场

Fig. 4 Detection experimental site of spatial heterodyne Raman spectroscopy

4 结果与讨论

4.1 仿真拉曼光谱基线拟合结果

拉曼峰截断的 airPLS 基线拟合过程如图 5 所示,

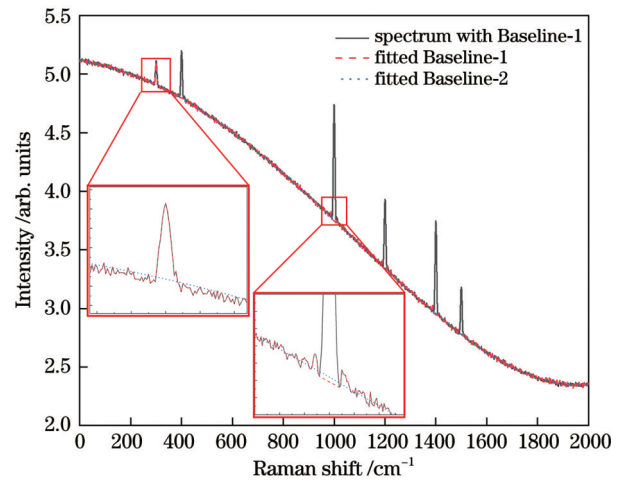


图 5 拉曼峰截断的 airPLS 基线拟合过程

Fig. 5 Baseline fitting procedure of airPLS based on Raman peak truncation

300 cm^{-1} 处的拉曼峰由于强度较弱,寻峰阶段未将其识别为拉曼峰,因此在第一次基线拟合过程中,将其作为基线保留了下来。但拟合迭代至第二次时,拉曼峰趋势被完全去除,第三次迭代满足基线输出条件。由此可见,改进方法的基线拟合结果对寻峰参数的设置并不敏感。

下面分析拟合算法在不同基线类型和不同光谱信噪比下的拟合效果。在图 6 和图 7 中给出了 300 cm^{-1} 的拉曼峰在信噪比为 5 时,1 类型基线和 2 类型基线的各种算法拟合结果。通过计算上述 5 种拟合方法的 RMSE 来评价其基线拟合准确度,其值越小,表明拟合基线与理论基线越接近,50 次仿真计算结果的平均值如表 3 所示。

从图 6(a)中可以发现,在低信噪比情况下,300 cm^{-1} 拉曼峰强度较弱,因此 5 种算法对于拉曼峰附近基线的拟合效果接近,本文算法虽更加平坦并接近理论基线,但优势并不明显。而在图 6(b)和 6(c)中可以明显

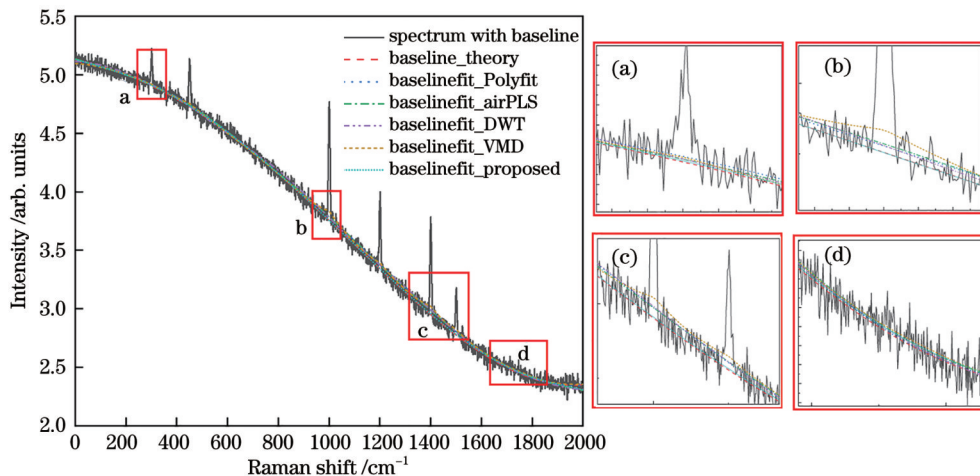


图 6 各算法在 Baseline-1 中的拟合结果。(a)~(d)为左图中对应 a~d 的局部放大图

Fig. 6 Fitting results of each algorithm in Baseline-1. (a)~(d) are locally enlarged images of corresponding a~d in left graph

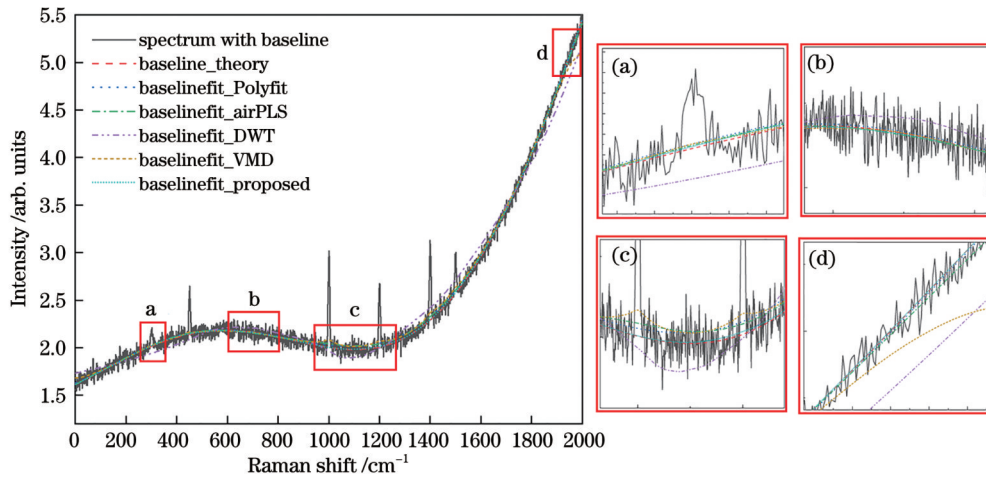


图 7 各算法在 Baseline-2 中的拟合结果。(a)~(d) 为左图中对应 a~d 的局部放大图

Fig. 7 Fitting results of each algorithm in Baseline-2. (a)~(d) are locally enlarged images of corresponding a~d in left graph

表 3 各算法拟合基线的 RMSE

Table 3 RMSE of each algorithm fitting baseline

	RMSE	Polyfit	airPLS	DWT	VMD	Proposed
Baseline-1	SNR is 100	0.01762	0.01976	0.02010	0.02182	0.00216
	SNR is 50	0.01763	0.01976	0.02011	0.02316	0.00224
	SNR is 20	0.01765	0.01976	0.02014	0.02290	0.00254
	SNR is 5	0.01762	0.01992	0.02063	0.02839	0.00413
Baseline-2	SNR is 100	0.01941	0.02159	0.02630	0.05096	0.00214
	SNR is 50	0.01941	0.02159	0.02630	0.05087	0.00219
	SNR is 20	0.01941	0.02158	0.02626	0.05078	0.00340
	SNR is 5	0.01954	0.02194	0.02687	0.05122	0.00514

看出,当拉曼峰强度较大时,各种拟合算法出现差异。其中 VMD 的拟合方法受到拉曼峰的严重影响,拟合基线在拉曼峰根部出现严重凸起。而 airPLS、DWT 和 Polyfit 的拟合效果接近,虽然基线平坦,但受到拉曼峰的影响,相对于理论基线发生抬升。而改进的 airPLS 拟合方法由于将拉曼峰区域信号截断,大幅消除了因拉曼峰的凸起对拟合基线造成的影响。因此相较于另外 4 种方法,本文方法的基线拟合结果更加接近理论值。而在无拉曼峰区域,各方法拟合效果接近。

在第二种基线类型下的拟合结果中,改进的 airPLS 拟合方法的优势同样明显。在位于 300 cm^{-1} 处弱拉曼峰附近和 $600\sim 800\text{ cm}^{-1}$ 区域无拉曼峰的拟合结果中,如图 7(a) 和 7(b) 所示,各种方法体现出了相似的拟合精度。但在拉曼峰强度较高的 1000 cm^{-1} 和 1200 cm^{-1} 区域附近,如图 7(c) 所示,受到拉曼峰趋势的影响,拟合基线出现不同程度的向上偏移。其中 VMD 偏移最为严重,在拉曼峰根部出现了局部凸起。这表明 VMD 对光谱频率进行分离时,低频信号中仍然包含部分高频信息。而 airPLS、DWT 和 Polyfit 的拟合效果接近,虽然基线平坦,但受到拉曼峰的影响,

拟合基线出现抬升。观察发现,受到 1000 cm^{-1} 和 1500 cm^{-1} 处强拉曼峰的影响,拟合基线的抬升效果发生自 840 cm^{-1} ,结束自 1600 cm^{-1} ,影响了拉曼峰附近较大区域的基线拟合精度。而在整个光谱的两端区域,由于光谱截断,DWT 和 VMD 均出现了不同程度的拟合失真。而本文中采用拉曼峰截断的 airPLS 方法,在 airPLS 的基础上消除了拉曼峰趋势影响,大幅提高了拉曼峰附近的基线拟合精度。且无论是在光谱两端区域还是无信号区域,对基线的拟合精度均为最高。由表 3 的对比发现,在不同信噪比光谱的仿真实验中,本文方法的拟合精度均为最高,其 RMSE 相较于其余 4 种拟合方法约降低一个数量级。

4.2 实测光谱基线拟合结果

实验中采集了口红的拉曼光谱,使用本文提出的改进 airPLS 和上述几种常见基线拟合方法得到的拟合基线如图 8 所示。口红样品主要由油(40%~70%)、蜡(约 15%)、着色剂(约 8%)、抗氧化剂和香料等成分构成,成分复杂且具有强烈的荧光背景^[23]。即使采用 785 nm 近红外激发,荧光强度仍远高于拉曼强度。在 $1100\sim 1650\text{ cm}^{-1}$ 频移范围内,存在多个拉曼峰,其中

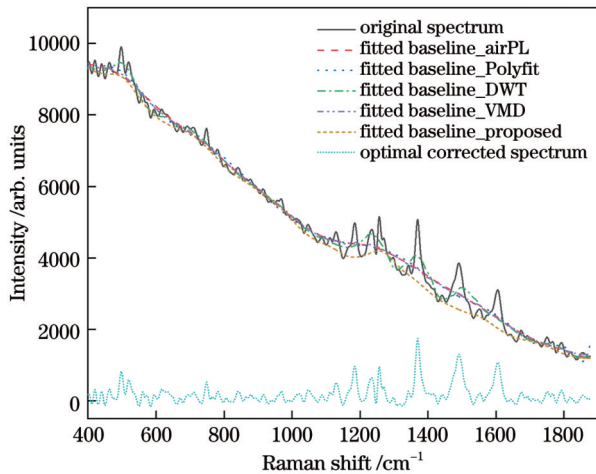


图8 口红拉曼光谱的基线校正结果

Fig. 8 Baseline correction results of lipstick Raman spectra

强度较大的 1492 cm^{-1} 和 1606 cm^{-1} 分别是由甘油三酯 C—C 键以及氨基 ($-\text{NH}_2$) 的伸缩振动产生的, 1370 cm^{-1} 由口红中油类成分产生, 1650 cm^{-1} 归结于口红蜡质成分中脂肪酸 $\text{C}=\text{C}$ 和 $-\text{CH}$ 键的振动。对比发现, 在弱拉曼峰区域如 $800\sim 1000\text{ cm}^{-1}$ 范围内, 各算法拟合基线接近。而在拉曼峰密集的区域如 $1100\sim 1650\text{ cm}^{-1}$ 范围, 各算法拟合的基线均不同程度地受到谱峰影响出现基线抬升现象, 其中 DWT 受到拉曼峰影响最为严重。本文所提出的基线拟合方法由于进行了谱峰截断, 在拟合基线时通过迭代的方法极大地降低谱峰对拟合结果的影响, 最大限度地保留了拉曼峰强度, 为几种方法中最优的。

5 结 论

拉曼光谱的产生伴随着荧光背景, 直接影响样品的定性和定量分析。因此, 对拉曼光谱的分析进行基线校正以消除荧光干扰是必要的。本文提出了拉曼峰截断的 airPLS 基线拟合方法。使用拉曼峰截断的方法降低了尖锐的拉曼峰对基线拟合的影响。既继承了 airPLS 在无峰区域的良好基线拟合效果, 又解决了 airPLS 基线拟合中因拉曼峰强度造成的基线拟合抬升的问题, 提高了基线拟合的精度。在对模拟光谱进行的比较实验中, 拉曼峰截断的 airPLS 在单一下降趋势荧光背景和复杂变化趋势的荧光背景中展示了优越的基线拟合性能。在不同的光谱信噪比下, 1 类型光谱拟合基线的 RMSE 优于 0.0042 , 2 类型光谱拟合基线的 RMSE 优于 0.0052 , 在各种方法中均为最低。在实测口红拉曼光谱的基线拟合实验中, 本文基线拟合方法的结果相较几种常用算法为最优, 在不引起校正光谱失真的情况下, 拉曼峰强度被准确恢复, 荧光基线趋势被有效去除, 满足拉曼光谱数据处理的需求。

参 考 文 献

- [1] 刘凤翔, 张礼豪, 黄霞. 拉曼光谱技术在肿瘤诊断中的应用[J]. 激光与光电子学进展, 2022, 59(6): 0617016.
Liu F X, Zhang L H, Huang X. Application of Raman spectroscopy in cancer diagnosis[J]. Laser & Optoelectronics Progress, 2022, 59(6): 0617016.
- [2] Gojani A B, Palásti D J, Paul A, et al. Analysis and classification of liquid samples using spatial heterodyne Raman spectroscopy[J]. Applied Spectroscopy, 2019, 73(12): 1409-1419.
- [3] 刘迎丽, 牟涛涛, 陈少华. 移频激发差分拉曼光谱技术的研究进展综述[J/OL]. 激光与光电子学进展: 1-22 [2023-11-24]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20230714.1054.198.html>.
Liu Y L, Mu T T, Chen S H. A review of the research progress of shifted excitation Raman difference spectroscopy[J/OL]. Laser & Optoelectronics Progress: 1-22 [2023-11-24]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20230714.1054.198.html>.
- [4] Sandford M W, Misra A K, Acosta-Maeda T E, et al. Detecting minerals and organics relevant to planetary exploration using a compact portable remote Raman system at 122 meters[J]. Applied Spectroscopy, 2021, 75(3): 299-306.
- [5] 姚泽楷, 蔡耀仪, 李诗文, 等. 基于平滑样条曲线结合离散状态转移算法的拉曼光谱基线校正方法[J]. 中国激光, 2022, 49(18): 1811001.
Yao Z K, Cai Y Y, Li S W, et al. Baseline correction for Raman spectroscopy using cubic spline smoothing combined with discrete state transformation algorithm[J]. Chinese Journal of Lasers, 2022, 49(18): 1811001.
- [6] Baek S J, Park A, Shen A G, et al. A background elimination method based on linear programming for Raman spectra[J]. Journal of Raman Spectroscopy, 2011, 42(11): 1987-1993.
- [7] 叶松, 甘永莹, 熊伟, 等. 采用小波变换的空间外差光谱仪基线校正[J]. 红外与激光工程, 2016, 45(11): 1117009.
Ye S, Gan Y Y, Xiong W, et al. Baseline correction of spatial heterodyne spectrometer using wavelet transform[J]. Infrared and Laser Engineering, 2016, 45(11): 1117009.
- [8] Koch M, Suhr C, Roth B, et al. Iterative morphological and mollifier-based baseline correction for Raman spectra[J]. Journal of Raman Spectroscopy, 2017, 48(2): 336-342.
- [9] Dragomiretskiy K, Zosso D. Variational mode decomposition[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544.
- [10] He S X, Fang S X, Liu X H, et al. Investigation of a genetic algorithm based cubic spline smoothing for baseline correction of Raman spectra[J]. Chemometrics and Intelligent Laboratory Systems, 2016, 152: 1-9.
- [11] Liu Y J, Zhou X G, Yu Y D. A concise iterative method using the Bezier technique for baseline construction[J]. Analyst, 2015, 140(23): 7984-7996.
- [12] Liu Y J, Lin J H. A general-purpose signal processing algorithm for biological profiles using only first-order derivative information [J]. BMC Bioinformatics, 2019, 20(1): 611.
- [13] Gao C, Zhao P, Fan Q, et al. Deep neural network: As the novel pipelines in multiple preprocessing for Raman spectroscopy [J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2023, 302: 123086.
- [14] Zhang Z M, Chen S, Liang Y Z. Baseline correction using adaptive iteratively reweighted penalized least squares[J]. The Analyst, 2010, 135(5): 1138-1146.
- [15] Guo Y X, Jin W Q, Wang W L, et al. Baseline correction for Raman spectra using a spectral estimation-based asymmetrically reweighted penalized least squares method[J]. Applied Optics, 2023, 62(18): 4766-4776.
- [16] Chen K, Zhang H Y, Wei H Y, et al. Improved Savitzky-

- Golay-method-based fluorescence subtraction algorithm for rapid recovery of Raman spectra[J]. *Applied Optics*, 2014, 53(24): 5559-5569.
- [17] Xu D G, Liu S, Cai Y Y, et al. Baseline correction method based on doubly reweighted penalized least squares[J]. *Applied Optics*, 2019, 58(14): 3913-3920.
- [18] Ye J F, Tian Z Y, Wei H Y, et al. Baseline correction method based on improved asymmetrically reweighted penalized least squares for the Raman spectrum[J]. *Applied Optics*, 2020, 59(34): 10933-10943.
- [19] 熊伟, 施海亮, 俞能海. 空间外差干涉光谱仪仪器线型函数测量新方法研究[J]. *光谱学与光谱分析*, 2015, 35(1): 267-271.
Xiong W, Shi H L, Yu N H. Study on a new method for instrumental line shape measurement of spatial heterodyne interference spectrometer[J]. *Spectroscopy and Spectral Analysis*, 2015, 35(1): 267-271.
- [20] 赵恒, 陈娱欣, 续小丁, 等. 基于局部对称重加权惩罚最小二乘的拉曼基线校正[J]. *中国激光*, 2018, 45(12): 1211001.
Zhao H, Chen Y X, Xu X D, et al. Baseline correction for Raman spectra based on locally symmetric reweighted penalized least squares[J]. *Chinese Journal of Lasers*, 2018, 45(12): 1211001.
- [21] Ding Y, Luo H Y, Shi H L, et al. Correction of invalid data based on spatial dimension information of a temporally and spatially modulated spatial heterodyne interference imaging spectrometer[J]. *Applied Optics*, 2021, 60(22): 6614-6622.
- [22] Wang Q S, Luo H Y, Li Z W, et al. Greenhouse gases monitoring instrument on a GF-5 satellite- II : correction of spatial and frequency-dependent phase distortion[J]. *Optics Express*, 2023, 31(2): 3028-3045.
- [23] Gładysz M, Król M, Kościelniak P. Differentiation of red lipsticks using the attenuated total reflection technique supported by two chemometric methods[J]. *Forensic Science International*, 2017, 280: 130-138.

Spatial Heterodyne Raman Spectral Baseline Correction Based on Improved Adaptive Iterative Re-Weighted Penalized Least Square Method

Bai Yunfei^{1,2,3}, Luo Haiyan^{1,2,3}, Li Zhiwei^{1,3}, Ding Yi^{1,3}, Xiong Wei^{1,2,3*}

¹Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, Anhui, China;

²University of Science and Technology of China, Hefei 230026, Anhui, China;

³Key Laboratory of Optical Calibration and Characterization of Chinese Academy of Sciences, Hefei 230031, Anhui, China

Abstract

Objective Raman spectroscopy is a non-destructive analytical technique based on the inelastic scattering of light by matter. While inducing Raman signals, the fluorescence background directly affects the characterization of sample Raman properties. The common approaches to reducing fluorescence background are primarily implemented through hardware and software methods. Hardware methods mainly involve techniques such as shifted excitation Raman difference spectroscopy, time-resolved Raman spectroscopy, and deep ultraviolet Raman spectroscopy. Although these methods exhibit effective outcomes, they often entail complex instrument setups and high costs. Software methods, on the other hand, refer to utilizing signal processing techniques to subtract fluorescence background from Raman spectra. Raman spectra are characterized by typically discontinuous with sharp peaks, contrasted with the continuous and smooth trends often present in fluorescence spectra. Given the difference in their spectral characteristics, when Raman spectroscopy analysis is carried out, employing baseline correction algorithms to eliminate fluorescence interference helps ensure the reliability and accuracy of Raman spectroscopy data. Common methods for mitigating fluorescence background include polynomial fitting, discrete wavelet transform, morphological algorithms, variational mode decomposition, least squares methods, and neural networks. However, due to the presence of Raman peaks, these methods typically result in varying degrees of baseline rise in the fitting outcomes. In the present study, we report an adaptive iterative re-weighted penalized least square (airPLS) method based on Raman peaks truncation. By identifying the positions of Raman peaks, truncating the corresponding regions, and employing the airPLS algorithm for baseline fitting, the method reduces the rise in the fitted baseline caused by abrupt changes in intensity within the Raman peak regions, making the fitted baseline approach closer to the true baseline. We hope that this improved method will further enhance the accuracy of Raman spectroscopy baseline fitting.

Methods Baseline fitting is conducted with airPLS based on Raman peaks truncation. Initially, the spectral signal is denoised by the Savitzky-Golay filter. Subsequently, we employ a peak-finding algorithm to identify Raman peaks within the denoised spectrum and use the first derivative of the spectrum to determine the left and right boundaries for each Raman

peak. Following this, we truncate the Raman peaks within these defined boundaries to obtain the original baseline. An airPLS fitting is performed on this original baseline to derive a new baseline. At this stage, we compute the difference between the new baseline and the original baseline and truncate the regions where the absolute difference exceeds a threshold. We iterate this process by subjecting the truncated signal to successive airPLS fitting until the absolute difference between the baselines from two consecutive fittings is less than the threshold, concluding the iteration. The resulting fitted baseline is output. Here, the threshold represents the average of the absolute differences between the baselines from two consecutive fittings. Subtracting the fitted baseline from the original Raman spectrum yields the corrected Raman spectrum.

Results and Discussions The airPLS based on Raman peaks truncation has demonstrated outstanding performance in baseline fitting. Simulated Raman spectra and measured Raman spectra from lipstick are utilized to validate the proposed baseline fitting method, respectively. Comparative analyses are conducted against commonly used baseline fitting methods, including polynomial fitting, discrete wavelet transform, variational mode decomposition, and airPLS. As depicted in Figs. 6(a), 7(a), and 7(b), although the algorithm proposed in this article achieves a baseline fit closer to the theoretical baseline in regions with weak Raman peaks or without Raman peaks, the improvement compared to the aforementioned algorithms is not notably conspicuous. This indicates that the fitting capabilities of these algorithms are similar in spectra exhibiting gradual trends. However, near the Raman peak regions, as shown in Figs. 6(b), 6(c), and 7(c), these methods experience baseline elevation due to abrupt changes in spectral peak intensity. In contrast, the proposed baseline fitting method, which incorporates Raman peaks truncation, minimizes the influence of Raman peaks on the fitting results, resulting in the closest fit to the theoretical values. Table 3 presents the root mean square errors (RMSE) between the fitted baseline and the theoretical baseline for our method and the aforementioned commonly used methods, evaluating the performance of these methods in spectra exhibiting both single and complex trends. Comparative analysis indicates that under various signal-to-noise ratios of spectra, our method yields the lowest RMSE, showcasing its superior performance. The baseline fitting results of the lipstick Raman spectrum shown in Fig. 8 are consistent with the simulated analysis outcomes. In the region devoid of Raman peaks ($800\text{--}1000\text{ cm}^{-1}$), the baseline fitting capabilities of each algorithm are similar. However, within the region with numerous Raman peaks ($1100\text{ to }1650\text{ cm}^{-1}$), the baseline fits of all algorithms are affected to varying degrees by the spectral peaks, resulting in baseline rises. Our baseline fitting method, employing spectral peak truncation and an iterative approach, significantly mitigates the influence of spectral peaks on the fitting outcomes. This method maximizes the preservation of Raman peak intensities and stands out as the optimal choice among the various methods evaluated.

Conclusions We introduce the Raman peak-truncated airPLS baseline fitting method. The utilization of Raman peak truncation mitigates the influence of abrupt changes in Raman peak intensity on baseline fitting. The method not only inherits the effective baseline fitting performance of airPLS in peak-free regions but also resolves the issue of baseline elevation caused by Raman peak intensity, thereby enhancing the accuracy of baseline fitting. Comparative experiments conducted on simulated spectra demonstrate the superior baseline fitting performance of Raman peak-truncated airPLS. Under different spectral signal-to-noise ratios, the RMSE for Type 1 spectra fitted by our method is below 0.0042, and for Type 2 spectra, it is below 0.0052, which is the lowest among various methods. In experiments fitting Raman spectra from lipstick samples, airPLS based on Raman peaks truncation outperforms several commonly used algorithms. It accurately restores Raman peak intensities without distorting the corrected spectra, effectively removing fluorescence baseline trends and meeting the requirements of Raman spectroscopy data processing.

Key words Raman spectroscopy; spatial heterodyne; baseline correction; least squares method