



## 混合气体测量中重叠吸收谱线交叉干扰的分离解析方法

王前进 孙鹏帅 张志荣 张乐文 杨曦 吴边 庞涛 夏滑 李启勇

### Separation and analysis method of overlapping absorption spectra with cross interference in gas mixture measurement

Wang Qian-Jin Sun Peng-Shuai Zhang Zhi-Rong Zhang Le-Wen Yang Xi Wu Bian Pang Tao  
Xia Hua Li Qi-Yong

引用信息 Citation: *Acta Physica Sinica*, 70, 144203 (2021) DOI: 10.7498/aps.70.20210286

在线阅读 View online: <https://doi.org/10.7498/aps.70.20210286>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

## 您可能感兴趣的其他文章

### Articles you may be interested in

基于可调谐激光吸收光谱技术的二硫化碳中红外光谱参数测量

Measurement of middle infrared spectroscopic parameters of carbon disulfide based on tunable diode laser absorption spectroscopy  
物理学报. 2019, 68(8): 084204 <https://doi.org/10.7498/aps.68.20182140>

Ar-O<sub>2</sub>混合气体电弧的数值模拟

Numerical simulation of mixture gas arc of Ar-O<sub>2</sub>

物理学报. 2019, 68(17): 178102 <https://doi.org/10.7498/aps.68.20190416>

可调谐掺铥光纤激光器线宽压缩及其超光谱吸收应用

Linewidth compression of tunable Tm-doped fiber laser and its hyperspectral absorption application  
物理学报. 2020, 69(3): 034205 <https://doi.org/10.7498/aps.69.20191515>

基于TDLAS技术的水汽低温吸收光谱参数测量

Measuring spectral parameters of water vapor at low temperature based on tunable diode laser absorption spectroscopy  
物理学报. 2017, 66(20): 204204 <https://doi.org/10.7498/aps.66.204204>

基于可调谐半导体激光器吸收光谱的高灵敏度甲烷浓度遥测技术

High sensitive scheme for methane remote sensor based on tunable diode laser absorption spectroscopy  
物理学报. 2017, 66(10): 100702 <https://doi.org/10.7498/aps.66.100702>

基于可调谐二极管激光技术利用小波去噪在2.008 μm波段对δ<sup>13</sup>C<sub>2</sub>O<sub>2</sub>的研究(已撤稿)

The research of δ<sup>13</sup>C<sub>2</sub>O<sub>2</sub> by use of wavelet de-noising at 2.008 μm based on tunable diode laser absorption spectroscopy  
物理学报. 2017, 66(2): 024202 <https://doi.org/10.7498/aps.66.024202>

# 混合气体测量中重叠吸收谱线交叉 干扰的分离解析方法\*

王前进<sup>1)2)</sup> 孙鹏帅<sup>2)</sup> 张志荣<sup>1)2)3)4)†</sup> 张乐文<sup>2)5)</sup> 杨曦<sup>1)2)</sup>  
吴边<sup>2)</sup> 庞涛<sup>2)</sup> 夏滑<sup>2)</sup> 李启勇<sup>6)</sup>

1) (中国科学技术大学环境科学与光电技术学院, 合肥 230026)

2) (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, 光子器件与材料安徽省重点实验室, 合肥 230031)

3) (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, 中国科学院环境光学与技术重点实验室, 合肥 230031)

4) (国防科技大学, 先进激光技术安徽省实验室, 合肥 230037)

5) (中国科学技术大学, 合肥 230022)

6) (安徽省生态环境监测中心, 合肥 230071)

(2021年2月7日收到; 2021年3月11日收到修改稿)

在基于可调谐二极管激光吸收光谱技术 (tunable diode laser absorption spectroscopy, TDLAS) 进行多种组分混合气体测量时, 经常会遇到吸收谱线之间存在相互干扰的现象, 这也是使用该技术测量过程中的主要“瓶颈”。比如在前期的应用中: 微量一氧化碳 (CO) 和甲烷气体 (CH<sub>4</sub>) 在同时检测时两者的吸收谱线存在严重的重叠干扰现象, 特别是在高浓度 CH<sub>4</sub> 存在的环境下, 微量 CO 气体吸收信号会被干扰甚至湮没, 无法实现有效解调, 这是通过谱线选取所不能解决的问题。因此, 针对此问题本文提出了基于支持向量回归模型, 以 CO 和 CH<sub>4</sub> 吸收谱线的严重重叠干扰问题为例, 通过选择线性核函数建立 CO 支持向量回归模型和 CH<sub>4</sub> 支持向量回归模型, 可对 CO 和 CH<sub>4</sub> 的混合气体吸收谱线进行解调, 最终获得两种气体浓度的准确测量结果。通过实验分别实现了四种不同浓度 CH<sub>4</sub> 环境下微量 CO 气体的检测, 得到的 CO 和 CH<sub>4</sub> 浓度 (气体的体积分数) 测量的绝对误差分别小于  $2 \times 10^{-6}$  和  $0.2 \times 10^{-2}$ ; 通过不同浓度的固定配比实验分析, CO 和 CH<sub>4</sub> 气体的测量值与实际值之间的相关系数分别达到了 0.998 和 0.9995, 且 CO 和 CH<sub>4</sub> 气体测量结果的绝对误差分别不超过  $2 \times 10^{-6}$  和  $0.1 \times 10^{-2}$ , 完全满足了混合气体的精确测量, 为实现多场合混合气体目标的实时监测提出了解决方案。

**关键词:** 可调谐二极管激光吸收光谱, 混合气体, 谱线重叠, 支持向量回归

**PACS:** 42.62.Fi, 87.55.kd, 42.55.Px, 33.20.Ea

**DOI:** 10.7498/aps.70.20210286

## 1 引言

可调谐二极管激光吸收光谱 (tunable diode laser absorption spectroscopy, TDLAS) 具有非接触实时测量、分辨率高、灵敏度高、响应速度快等

优点<sup>[1-6]</sup>, 能够实现对特定气体分子的高选择性精确测量。结合波长调制技术和长光程技术的 TDLAS 系统能够得到更高的检测灵敏度和更低的检测下限<sup>[7-9]</sup>, 已被广泛应用于环境监测、汽车尾气监测、工业气体检测、燃烧诊断、医学诊断等众多领域<sup>[10-14]</sup>。

\* 国家自然科学基金 (批准号: 11874364, 41877311, 42005107)、国家重点研发计划 (批准号: 2017YFC0805004)、安徽省重点研究与开发计划项目 (批准号: 201904c03020005) 和中国科学院合肥研究院“火花”基金项目 (批准号: YZJJ2020QN8) 资助的课题。

† 通信作者。E-mail: zhangzr@aiofm.ac.cn

通常情况下,采用 TDLAS 技术同时测量混合气体浓度时,需要根据每一种气体分子吸收谱线的“指纹识别区”设计一套特定波段的监测系统,也就是气体和激光光源是一一对应的关系,这也就导致了整套系统的复杂程度和成本的增加.虽然当两种气体的吸收谱线比较接近并且不会互相重叠时(如 CO 和 CO<sub>2</sub> 在 1579 nm 附近的吸收谱线),可以采用温度控制和驱动电流控制的方法使得一台分布反馈式激光器(distributed feedback laser, DFB laser)能够同时扫描过这两种气体的吸收波长,进而实现两种气体同时在线检测<sup>[15]</sup>.这样不仅能够降低检测系统的复杂程度以节省成本,也提高了系统的响应率和检测速度.然而像 CO 和 CO<sub>2</sub> 气体的这类相近而又独立的吸收谱线是特殊的,大多数情况下两种气体之间的吸收谱线足够接近且存在交叉干扰现象.为了对混合气体中多种成分进行测量,科研人员进行了多种相关研究<sup>[16,17]</sup>,如: Zhang 等<sup>[18]</sup>利用中心波长为 1578 和 1747 nm 的两个激光器对比了分时锯齿法、光开关法和多频正弦调制法同时检测 H<sub>2</sub>S 和 HCl 两种气体的优缺点; Shao 等<sup>[19]</sup>利用中心波长为 2.3 μm 的 DFB 激光器在低压条件下完成了 CO 和 CH<sub>4</sub> 气体的同时检测.以上方法虽然提高了混合气体实时检测的可靠性,但是同时也会增加系统电路、光路或者气路的复杂性.此外,在某些应用场合,上述处理方式仍不可避免地遇到谱线间的交叉干扰问题.例如常温常压下,CO 和 CH<sub>4</sub> 的吸收谱线存在严重的交叉干扰现象,特别是在含有高浓度 CH<sub>4</sub> 的特殊场合(例如矿井、天然气泄漏检测等),CO 的吸收信号会被湮没在不规则的多峰 CH<sub>4</sub> 吸收谱线之下,此时难以对这两种气体,特别是微量 CO 气体进行准确地解析.

为了解决上述谱线干扰的共性问题,进一步实现简化系统复杂程度、降低系统成本、增加系统可靠性的目的,本文提出了一种基于支持向量回归(support vector regression, SVR)模型的方法,来解决 TDLAS 系统遇到的多组分混合气体检测中的谱线干扰问题.支持向量机在气体检测方面已经得到了广泛应用. Qu 等<sup>[20]</sup>利用改进的网格搜索算法对支持向量机进行参数优化,建立了气体的红外光谱定量分析方法,并且在 1%—25% 浓度范围内对 CO<sub>2</sub> 气体进行定量分析,得到的绝对误差小于 4%; Laref 等<sup>[21]</sup>对比了支持向量回归和偏最小二乘回归在定量分析松树精油蒸气应用中的性能,得

出了支持向量回归具有更高测量精度的结论.本文以微量 CO 和较高浓度 CH<sub>4</sub> 为例,经过小样本数据集训练得到 CO 和 CH<sub>4</sub> 气体浓度回归模型,分别为 CO-SVRmodel 和 CH<sub>4</sub>-SVRmodel.经不同浓度的气体测量实验验证,该模型能够显著解决微量 CO 气体吸收信号在高浓度 CH<sub>4</sub> 吸收信号下被湮没的谱线重叠干扰问题,实现了微量 CO 和较高浓度 CH<sub>4</sub> 的同时检测,为实现多组分气体的 TDLAS 同时测量技术奠定了基础和可行性论证.

## 2 实验原理

### 2.1 TDLAS 原理

根据 Lambert-Beer 定律,特定波长的激光光束穿过待测气体会被吸收,透射光强变化可以表示为

$$I_t = I_0 \exp[-S(T)XPL\phi(v)] = I_0 \exp[-\alpha(v)L], \quad (1)$$

其中  $I_0$  是激光出射强度;  $I_t$  是透射激光强度;  $S(T)$  ( $\text{cm}^{-2} \cdot \text{atm}^{-1}$ ) 为气体温度为  $T$  时的吸收谱线强度;  $X$  为待测气体浓度;  $P(\text{atm})$  为气体压强;  $L(\text{cm})$  为光在待测气体中传播的距离即光程;  $\phi(v)$  是吸收线型函数,与气体温度和压力有关,描述了气体吸收谱线的形状;  $\alpha(v)$  ( $\text{cm}^{-1}$ ) 为待测气体对特定波数  $v$  的吸收率.

为了提高系统的抗干扰能力和检测灵敏度,引入波长调制技术.在激光器的扫描信号上叠加一个高频的正弦调制信号,则激光器在某时刻  $t$  的瞬时频率和出光强度可以表示为

$$v(t) = \bar{v} + a \cos(\omega t + \varphi), \quad (2)$$

$$I_0(t) = \bar{I}_0 + i_0 \cos(\omega t + \varphi), \quad (3)$$

其中  $\bar{v}$  和  $\bar{I}_0$  分别为激光器的中心波长和平均光强;  $a$  和  $i_0$  分别为频率调制幅度和光强调制幅度;  $\omega$  和  $\varphi$  分别为正弦调制的频率和初始相位.调制之后的激光,被待测气体吸收,经锁相放大器检测可以获得各个分量的谐波信号,并且一般情况下  $-\alpha(v)L \ll 1$ ,结合 Reid 和 Labrie<sup>[22]</sup>给出的二次谐波表达式,得到二次谐波信号峰值与气体浓度成正比,即

$$P_{2f} \propto \frac{I_0 S(T) P X L}{\pi \Delta v} \left\{ \frac{2}{m^2} \left[ \frac{2+m^2}{(1+m^2)^{1/2}} - 2 \right] \right\}, \quad (4)$$

其中  $P_{2f}$  为二次谐波的峰值;  $\Delta v$  为气体吸收谱线

的半高宽;  $m$  为无量纲变量,  $m = a/\Delta\nu$ . 二次谐波信号的峰值和初始光强  $I_0$  也成正比关系. 因此实际测量中, 常使用一次谐波进行光强修正, 消除光强波动引起的峰值变化.

## 2.2 支持向量回归 (SVR) 原理

SVR 原理可以分为线性回归和非线性回归 [23], 首先给定一组样本集:  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ . 其中,  $x_i (x_i \in \mathbb{R})$  是第  $i$  个样本的光谱数据;  $y_i (y_i \in \mathbb{R})$  为实际测量得到的第  $i$  个样本的气体浓度值.

当所测气体为单一气体时, 气体浓度与光谱数据之间符合线性关系, 构造线性回归模型如下:

$$f(x) = \mathbf{w}^T x + b, \quad (5)$$

其中  $\mathbf{w}$  为权重向量;  $b$  为偏置.

采用  $\epsilon$ -不敏感损失函数作为误差函数, 表达式如下:

$$l_\epsilon = \begin{cases} 0, & |z| \leq \epsilon, \\ |z| - \epsilon, & \text{otherwise.} \end{cases} \quad (6)$$

回归函数的求解问题可以描述为一个二次凸优化问题, 即:

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2, \quad \text{s.t.} \begin{cases} y_i - (\mathbf{w} \cdot x_i + b) \leq \epsilon, \\ (\mathbf{w} \cdot x_i + b) - y_i \leq \epsilon, \end{cases} \quad (i=1, 2, \dots, n). \quad (7)$$

根据 (6) 式和 (7) 式, 将 SVR 问题形式化定义为

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n l_\epsilon(f(x_i) - y_i) \quad (i=1, 2, \dots, n), \quad (8)$$

其中  $C$  为正则化常数,  $C$  越大表示对训练误差大于  $\epsilon$  的样本惩罚越大;  $l_\epsilon$  是 (6) 式对应的  $\epsilon$ -不敏感损失函数. 当我们允许 SVR 在某些样本上出错时, 即引入松弛变量  $\xi_i, \xi_i^*$  后, 目标函数变为

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*), \quad \text{s.t.} \begin{cases} y_i - (\mathbf{w} \cdot x_i + b) \leq \epsilon + \xi_i^*, \\ (\mathbf{w} \cdot x_i + b) - y_i \leq \epsilon + \xi_i, \quad (i=1, 2, \dots, n), \\ \xi_i, \xi_i^* \geq 0. \end{cases} \quad (9)$$

为了更加高效地求解目标函数, 可以对 (9) 式使

用拉格朗日乘子法. 引入拉格朗日乘子  $\mu_i \geq 0, \hat{\mu}_i \geq 0, \alpha_i \geq 0, \hat{\alpha}_i \geq 0$ , 最终得到目标优化函数的“对偶问题”如下:

$$\begin{aligned} & \max_{\alpha, \hat{\alpha}} \sum_{i=1}^n y_i (\hat{\alpha}_i - \alpha_i) - \epsilon (\hat{\alpha}_i + \alpha_i) \\ & - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\hat{\alpha}_i - \alpha_i) (\hat{\alpha}_j - \alpha_j) x_i^T x_j \\ & \text{s.t.} \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) = 0, \\ & 0 \leq \alpha_i, \hat{\alpha}_i \leq C. \end{aligned} \quad (10)$$

通过对偶问题的最优值求解, 可以得到线性 SVR 的解如下:

$$f(x) = \sum_{i=1}^n (\hat{\alpha}_i - \alpha_i) x_i^T x + b. \quad (11)$$

但是对于混合气体而言, 在原始样本空间是线性不可分的, 此时可以通过引入非线性映射, 将样本空间映射到一个高维的特征空间中, 然后在该特征空间应用 SVR 线性回归方法求解回归函数.

令  $\varphi(x)$  表示将样本  $x$  映射后的特征向量, 于是, 在特征空间中划分超平面所对应的模型可表示为

$$f(x) = \mathbf{w}^T \varphi(x) + b. \quad (12)$$

对于 (12) 式, 经过与线性回归中相同的一系列求解后, 可以得到特征空间中回归模型的解如下:

$$f(x) = \sum_{i=1}^n (\hat{\alpha}_i - \alpha_i) \varphi(x_i)^T \varphi(x) + b, \quad (13)$$

其中  $\varphi(x_i)^T \varphi(x)$  是样本映射到特征空间之后的内积, 而特征空间维数很高, 甚至可能是无穷维, 因此直接计算  $\varphi(x_i)^T \varphi(x)$  是困难的, 为了避开这个障碍, 引入了核函数:

$$\kappa(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle = \varphi(x_i)^T \varphi(x_j). \quad (14)$$

最终, 特征空间下 SVM 回归模型的解如下:

$$f(x) = \sum_{i=1}^n (\hat{\alpha}_i - \alpha_i) \kappa(x, x_i) + b. \quad (15)$$

常用的核函数有线性核函数、多项式核函数和高斯核函数等, 经过实验得出结论, 对于混合气体数据, 使用线性核函数得到的 SVR 模型效果最优.

### 3 实验装置

为了实现 CO 的低浓度测量, 该实验系统选择中心波长为 2.3  $\mu\text{m}$  的 DFB 激光器作为光源. 其波长随着温度和电流变化的调谐性能如图 1 所示. 共测试了 6 组不同温度下的电流调谐性能, 均能够获得较好的线性特征, 因此在调谐过程中不会出现激光输出的不稳定性因素. 在本实验中将激光驱动器的扫描电流调节到合适的范围, 使得 DFB 激光器从 2326.6 nm 扫描到 2327.7 nm 的波数范围, 此时中心发射波长为 2327.2 nm, 完全覆盖 CO 的中心吸收谱线.

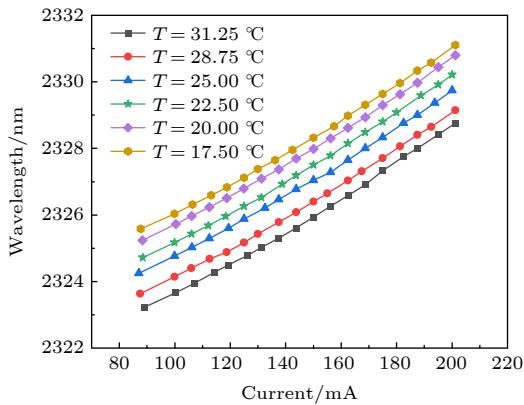


图 1 DFB 激光器电流-波长调谐性能曲线  
Fig. 1. Current-wavelength tuning curve of DFB laser.

实验装置原理图如图 2 所示, 分别由电路部分 (Part A)、光学部分 (Part B)、数据采集与分析部分 (Part C)、配气系统 (Part D) 四个功能模块组成. 首先, 电路部分包括函数发生器、激光驱动

器和双通道锁相放大器, 该部分完成 DFB 激光器驱动以及信号调制和解调的功能. 函数发生器产生低频扫描信号和高频正弦信号, 频率分别为 10 Hz 和 10 kHz. 低频扫描信号使得激光器扫描输出 2  $\text{cm}^{-1}$  的波长范围, 而高频正弦信号使激光器波长得到调制输出. 通过锁相放大器对接收到的信号进行解调而获得一次谐波信号 (1f) 和二次谐波信号 (2f). 其次, 光学部分包括蝶形封装的 DFB 激光器、带有光纤准直和光电探测器的光学吸收池, 其光程为 20 m. 最后, 在数据采集和分析部分中, 从锁相放大器获得的 1f 和 2f 分别包含 2000 个采样点, 采集触发信号为 10 Hz. 最后将 1f 和 2f 数据通过串口通信发送给计算机进行分析和计算. 配气系统主要用于不同浓度 CO 和  $\text{CH}_4$  混合气体的配比, 标准浓度的 CO 和  $\text{CH}_4$  气体分别通过质量流量计 1 和质量流量计 2 混合后通入吸收池, 通过已知的 CO 和  $\text{CH}_4$  气体浓度和对应质量流量计的读数, 可计算得到混合气体中 CO 和  $\text{CH}_4$  的浓度.

图 3 为 HITRAN 数据库 [24] 仿真得到的在上述 DFB 激光器的扫描波数范围内浓度分别为  $20 \times 10^{-6}$  和  $400 \times 10^{-6}$  的 CO 气体、浓度为  $1 \times 10^{-2}$  的  $\text{CH}_4$  气体吸收谱线, 该处及下文 (包括表 1) 中气体浓度均为气体对应的体积分数. 可以发现浓度为  $400 \times 10^{-6}$  的 CO 气体, 其吸收谱线的峰值信号强度仍然小于浓度为  $1 \times 10^{-2}$  的  $\text{CH}_4$  气体在该波长位置的吸收信号强度. 所以利用传统的浓度反演方法, 在高浓度  $\text{CH}_4$  气体的大信号背景下, 要实现精确地定量检测微量 CO 气体是非常困难的, 因此本文采用基于支持向量回归模型解决激光吸收光谱遇到的类似谱线重叠的共性问题.

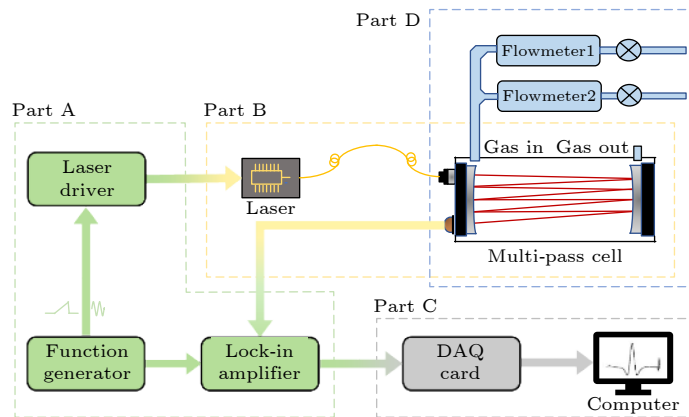


图 2 实验系统原理图  
Fig. 2. Schematic diagram of the experimental system.

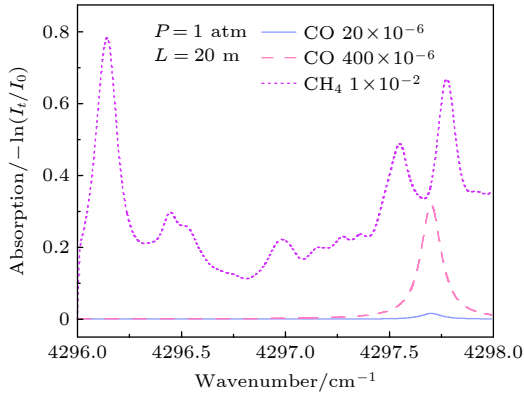


图 3 CO ( $20 \times 10^{-6}$ ,  $400 \times 10^{-6}$ ) 和  $\text{CH}_4$  ( $1 \times 10^{-2}$ ) 的模拟吸收光谱图

Fig. 3. Simulated absorption spectra of CO ( $20 \times 10^{-6}$ ,  $400 \times 10^{-6}$ ) and  $\text{CH}_4$  ( $1 \times 10^{-2}$ ).

表 1 训练数据集  
Table 1. Training data set.

Group	Standard gas category and concentration		Ratio	Gas category and concentration in multi-pass cell	
	$\text{CO}/10^{-6}$	$\text{CH}_4/10^{-2}$		$\text{CO}/10^{-6}$	$\text{CH}_4/10^{-2}$
1	19.3	0	—	19.3	0
2	54.0	0	—	54.0	0
3	102.0	0	—	102.0	0
4	0	0.50	—	0	0.50
5	0	1.04	—	0	1.04
6	0	2.02	—	0	2.02
7	0	5.02	—	0	5.02
8	54.0	1.04	1:1	27.0	0.52
9	102.0	1.04	1:1	51.0	0.52
10	19.3	1.04	1:1	9.6	0.52
11	19.3	0.50	1:1	9.6	0.25
12	19.3	5.02	1:1	9.6	2.51
13	19.3	2.02	1:1	9.6	1.01
14	19.3	8.50	1:1	9.6	4.25

## 4 实验结果与分析

### 4.1 模型训练

基于支持向量回归原理和 CO,  $\text{CH}_4$  两种气体在 2327 nm 附近的重叠吸收光谱特征, 选择了 14 组经过 1f 归一化处理之后的 2f 数据 (2f/1f)<sup>[19]</sup> 作为训练集构建 SVR 模型. 训练数据集的气体种类及其对应浓度如表 1 所列. 其中包含 3 组 CO 标气、4 组  $\text{CH}_4$  标气、7 组 CO 和  $\text{CH}_4$  混合气体 (标气按照 1:1 比例配比而成). 经过计算, 光学多通池

中气体种类及实际浓度如表 1 中最后两列所示.

选择表 1 中三组气体, 分别为 Group 1 (CO 标气), Group 4 ( $\text{CH}_4$  标气) 和 Group 10 (混合气体), 它们的信号如图 4 所示, 其对应浓度如图 4 中图例所示. 对比图 4 中曲线和划线的差别可知,  $\text{CH}_4$  气体的吸收信号 (Group 4) 和包含微量 CO 的混合气体信号 (Group 10) 在图 4 黑框中 (对应采样点范围为 100 到 500 之间) 基本一致, 采用传统的计算方法无法将微弱的 CO 吸收信号解调出来.

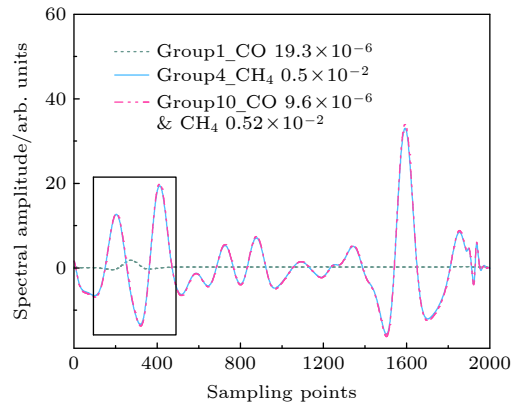


图 4 三组气体 (Group 1, Group 4, Group 10) 的 2f/1f 信号  
Fig. 4. 2f/1f signals of three groups of gases (Group 1, Group 4, Group 10).

而本文基于 Matlab R2018a 平台, 采用 5 折交叉验证经过 15 次迭代构建 SVR 模型, 并且采用序列最小最优化 (sequential minimal optimization, SMO) 算法进行参数优化, 得到两个 SVR 模型, 分别命名为 CO-SVRmodel 和  $\text{CH}_4$ -SVRmodel, 这两个模型的主要参数如表 2 中所列. 其中, Box-Constraint( $C$ ) 为惩罚系数, 主要影响模型的泛化能力; KernelScale 是归一化常数;  $\epsilon$  是不敏感损失函数设置的损失阈值.

表 2 SVR 模型主要参数  
Table 2. Optimal parameters of SVR model.

	CO-SVRmodel	$\text{CH}_4$ -SVRmodel
BoxConstraint( $C$ )	0.3443	0.0180
KernelScale	0.0617	0.9205
$\epsilon$	0.1710	65.1810
Total function evaluations	15	15
Total elapsed time in seconds	26.2545	12.5730

鉴于混合气体中 CO 和 CH<sub>4</sub> 气体的浓度差距在 10<sup>3</sup>—10<sup>4</sup> 量级, 且两种气体的吸收峰会产生明显的交叠, 所以为了使两种气体的回归模型都能够得到精准的结果, 两个模型采用的训练数据集不完全相同. CH<sub>4</sub>-SVRmodel 的训练采用整个 2f/1f 光谱数据进行训练, 能够充分反映 CH<sub>4</sub> 气体的吸收谱线特征; CO-SVRmodel 则放弃了光谱冗余数据, 选择只包含 CO 吸收信息的光谱数据 (光谱中的第 100—500 个点之间的范围) 进行训练, 如图 4 中黑色方框所示.

为了初步验证两种气体 SVR 模型的可用性, 首先对表 1 中的 14 组气体分别进行 60 组测试, 然后根据回归模型 (CO-SVRmodel, CH<sub>4</sub>-SVRmodel) 分别对 CO 和 CH<sub>4</sub> 气体进行浓度计算, 得到 CO 和 CH<sub>4</sub> 气体浓度的计算结果分别如图 5(a) 和图 5(b) 所示. 从图 5 中可以看出, 不管是 CO 气体还是 CH<sub>4</sub> 气体, 基于 SVR 模型计算得到的结果和真实值相差都不大, 其中 CO 气体的误差最大值为  $2 \times 10^{-6}$ , CH<sub>4</sub> 气体的误差最大值为  $0.16 \times 10^{-2}$ , 该处及下文中误差均是指气体浓度的绝对误差. 并且可以发现 CO 和 CH<sub>4</sub> 气体发生最大误差具有统一性, 即都发生在 CO 气体浓度为 0, CH<sub>4</sub> 气体浓度为  $5.02 \times 10^{-2}$  时的测量结果. 出现这种结果的原因是 CO 和 CH<sub>4</sub> 气体的浓度相差越大, 就会导致 CO 信号被湮没地越彻底, 并且在实验中也发现了浓度为  $5.02 \times 10^{-2}$  的 CH<sub>4</sub> 气体在信号稳定 (观察信号

稳定) 后仍然会存在相对大的波动, 也导致了 CO 和 CH<sub>4</sub> 气体的测量结果在此处产生了最大的误差. 经过验证, 采用线性核函数, 并且基于以上两个数据集分别训练得到的两个浓度回归模型比采用整个光谱数据集得到的回归模型更加精确.

## 4.2 CO 和 CH<sub>4</sub> 配比实验

4.1 节训练集数据的验证从一定程度上反映了 SVR 模型对 CO 和 CH<sub>4</sub> 两种气体浓度计算有效. 但是其中训练集和测试集的浓度设置完全一致, 所以不能很好地说明系统具有良好的泛化能力 (泛化能力是指对于训练数据集之外的数据仍然能够得到准确的计算结果, 标志着模型是否具有通用性和普适性). 为了更进一步探讨该模型的泛化能力, 进行了气体配比实验.

将浓度为  $102 \times 10^{-6}$  的 CO 气体与浓度分别为  $2.02 \times 10^{-2}$ ,  $4.5 \times 10^{-2}$  的 CH<sub>4</sub> 气体进行 1:4, 2:3, 3:2, 4:1 比例的混合, 并利用该实验系统对以上配比的混合气体进行检测, 待信号稳定后记录每组混合气体 30 组的数据, 并对其求平均值, 然后根据两个气体浓度回归模型 (CO-SVRmodel 和 CH<sub>4</sub>-SVRmodel) 对得到的数据进行回归计算. CO 和 CH<sub>4</sub> 气体浓度的真实值和测量值之间的相关性, 及其测量误差如图 6 所示.

图 6(c), (d) 可以说明: 基于上文的浓度回归模型分别对 CO 和 CH<sub>4</sub> 气体浓度进行浓度测量, CO

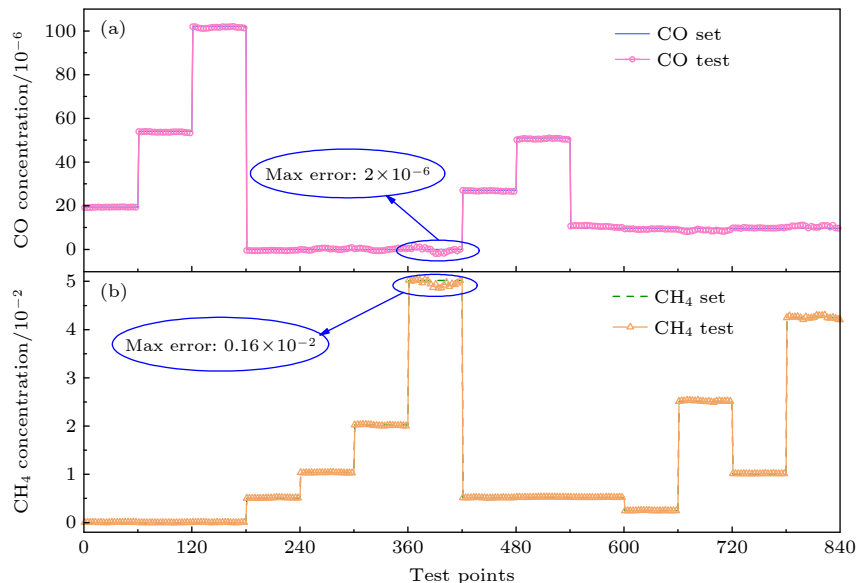


图 5 CO 和 CH<sub>4</sub> 浓度设置值和测试值对比 (a) CO 浓度值; (b) CH<sub>4</sub> 浓度值

Fig. 5. Comparisons between set and test values: (a) CO concentration; (b) CH<sub>4</sub> concentration.

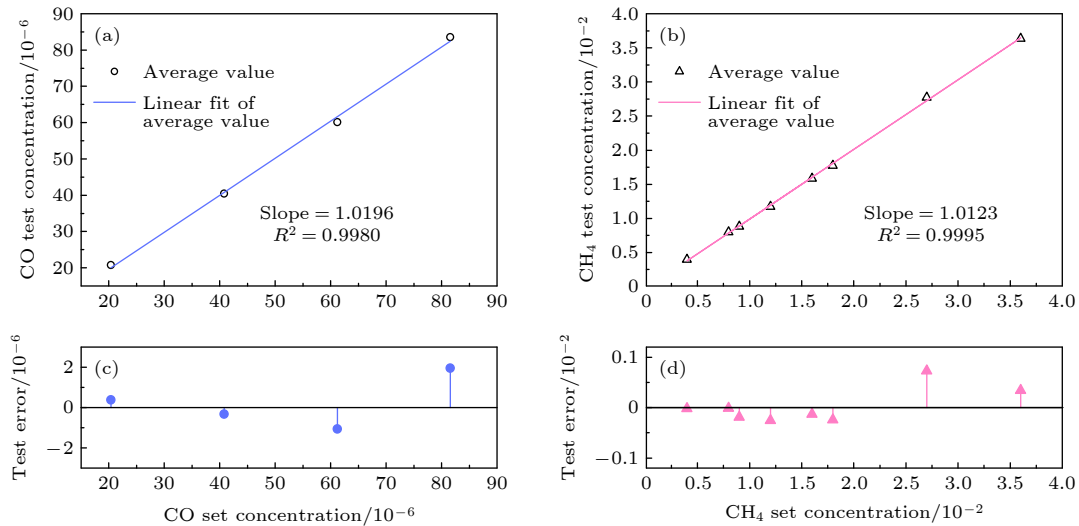


图 6 设置值和测量值之间相关性及其测量误差 (a) CO 相关性; (b) CH<sub>4</sub> 相关性; (c) CO 测量误差; (d) CH<sub>4</sub> 测量误差  
 Fig. 6. Correlation and test errors between set and average values: (a) CO correlation; (b) CH<sub>4</sub> correlation; (c) CO errors; (d) CH<sub>4</sub> errors.

和 CH<sub>4</sub> 气体浓度误差分别不超过  $2 \times 10^{-6}$  和  $0.1 \times 10^{-2}$ ; 且由图 6(a),(b) 可知, CO 和 CH<sub>4</sub> 的真实值与预测值之间的相关系数分别为 0.998 和 0.9995. 由此可得 CO-SVRmodel 和 CH<sub>4</sub>-SVRmodel 回归模型在固定配比数据集上能够取得良好的浓度回归测量效果, 证明上述两个模型良好的泛化能力.

### 4.3 混合气体浓度定量解析

为了验证该算法在不同浓度的 CH<sub>4</sub> 环境下仍能够对微量 CO 进行准确测量. 将  $18 \times 10^{-6}$  的 CO 气体和四种浓度的 CH<sub>4</sub> 气体 (浓度分别为  $0.5 \times 10^{-2}$ ,  $1.04 \times 10^{-2}$ ,  $2.02 \times 10^{-2}$ ,  $5.02 \times 10^{-2}$ ) 进行 1:3 的配比, 待信号稳定后进行浓度结果回归计算, 每组配比的气体进行 120 组数据测量, 其中 CO 和 CH<sub>4</sub> 的浓度真实值如图 7 中黑色数字标注. CO 浓度回归结果和 CH<sub>4</sub> 回归结果如图 7 所示. 从图 7 中可以看出, 该 TDLAS 系统结合 SVR 模型计算得到的 CO 和 CH<sub>4</sub> 气体浓度在较小的范围之内波动. 微量 CO 气体在受高浓度 CH<sub>4</sub> 气体干扰的条件下, 其最大误差小于  $2 \times 10^{-6}$ ; 而 CH<sub>4</sub> 气体得到的测量结果最大误差小于  $0.2 \times 10^{-2}$ , 证明了系统的抗干扰能力, 说明 CO-SVRmodel 和 CH<sub>4</sub>-SVRmodel 能够克服吸收谱线数据中的特征干扰问题并得到精确的浓度回归结果.

为了更进一步验证 CO-SVRmodel 和 CH<sub>4</sub>-SVRmodel 两个浓度回归模型的动态可靠性, 利用上述实验系统进行了两组充气过程实验, 分别如

下: 在两组过程中均利用 CO-SVRmodel 和 CH<sub>4</sub>-SVRmodel 得到 CO 和 CH<sub>4</sub> 的浓度结果.

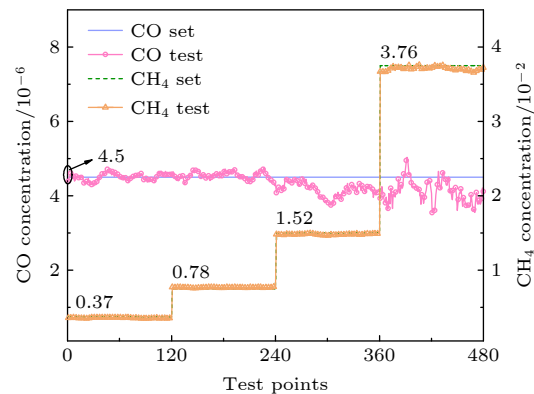


图 7 不同浓度 CH<sub>4</sub> 对 CO 测量结果的干扰  
 Fig. 7. Interference on CO measurement results by different concentrations of CH<sub>4</sub>.

第一组充气过程中 CO 和 CH<sub>4</sub> 气体的浓度变化趋势如图 8 所示. 初始状态气体池中底气为  $1.04 \times 10^{-2}$  的 CH<sub>4</sub> 气体, 在某一时刻, 将浓度为  $20 \times 10^{-6}$  的 CO 气体和浓度为  $1.04 \times 10^{-2}$  的 CH<sub>4</sub> 气体按照 1:1 的配比同时打开, 可以观察到 CO 气体浓度逐渐上升, CH<sub>4</sub> 气体浓度逐渐下降, 且两者浓度变化过程具有时间上的统一性, 也会随着时间的推移接近按照 1:1 比例配比的浓度 (CO 浓度接近  $10 \times 10^{-6}$ , CH<sub>4</sub> 浓度接近  $0.52 \times 10^{-2}$ ); 当 CO 和 CH<sub>4</sub> 混合气体吸收信号无明显变化, 待信号稳定一段时间后, 关闭 CH<sub>4</sub> 气体气阀, 保持 CO 气体

气阀打开,可以观察到 CO 气体浓度逐渐上升并趋近  $20 \times 10^{-6}$ ,  $\text{CH}_4$  气体浓度继续逐渐下降并趋近 0,且两者浓度变化过程具有时间上的统一性.

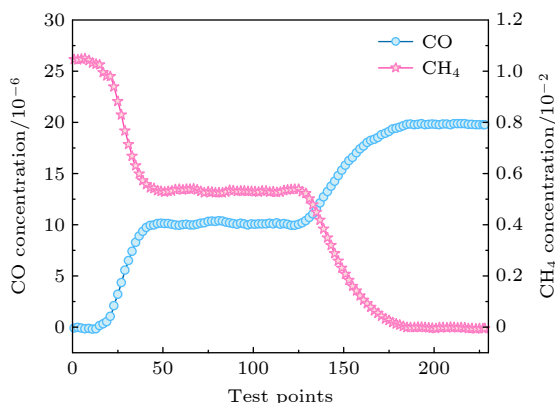


图 8 第一组充气过程中 CO 和  $\text{CH}_4$  浓度变化

Fig. 8. Concentration measured of CO and  $\text{CH}_4$  during the first set of inflation.

第二组充气过程中 CO 和  $\text{CH}_4$  气体的浓度变化趋势如图 9 所示. 气体池中底气为浓度  $8.5 \times 10^{-2}$  的  $\text{CH}_4$  气体和浓度  $20 \times 10^{-6}$  的 CO 气体按照 1:1 比例混合的气体,从图 9 中可以看出 CO 和  $\text{CH}_4$  气体均能够获得比较准确且稳定的结果,在某一时刻,关闭  $\text{CH}_4$  气体气阀,保持 CO 气体气阀打开,此时会发现  $\text{CH}_4$  气体浓度会微微增大,再缓慢减小,直至信号稳定,此时  $\text{CH}_4$  气体浓度已经减小到 0 附近并保持;在关闭  $\text{CH}_4$  气阀的同时,观察到 CO 气体浓度会下降约  $3 \times 10^{-6}$ ,然后再逐渐上升至  $20 \times 10^{-6}$  左右并保持. 出现上述现象的原因是由于此时混合气体中  $\text{CH}_4$  含量较高(为  $4.25 \times 10^{-2}$ ),关闭  $\text{CH}_4$  气阀瞬间,信号会产生波动,而该波动对多峰的  $\text{CH}_4$  浓度回归计算影响较小,但是对特征量较少的 CO 气体浓度回归计算会产生较大的影响.

通过上述两组充气过程的验证, TDLAS 系统结合 SVR 模型能够对包含微量 CO 和浓度不超过  $4.25 \times 10^{-2}$  的  $\text{CH}_4$  的混合气体进行动态实时监测,且在整个动态过程中能够将 CO 和  $\text{CH}_4$  的误差分别控制在  $3 \times 10^{-6}$  和  $0.5 \times 10^{-2}$ ;待信号稳定后,利用该回归模型得到的 CO 和  $\text{CH}_4$  浓度误差仍然不超过  $2 \times 10^{-6}$  和  $0.2 \times 10^{-2}$ . 充气动态过程实验充分验证了基于 SVR 模型解调 CO 和  $\text{CH}_4$  混合气体重叠干扰吸收谱线的可靠性和完备性.

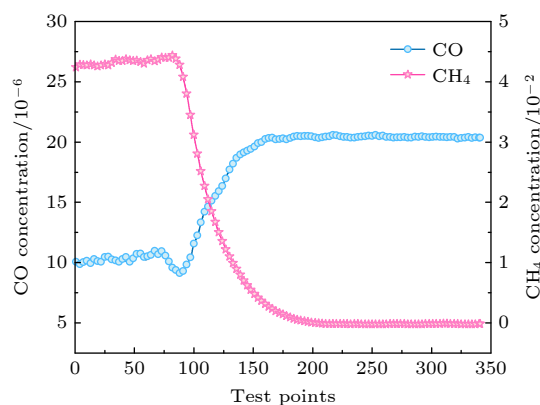


图 9 第二组充气过程中 CO 和  $\text{CH}_4$  浓度变化

Fig. 9. Concentration measured of CO and  $\text{CH}_4$  during the second set of inflation.

## 5 结 论

本文基于  $2.3 \mu\text{m}$  的 DFB 激光器以 CO、 $\text{CH}_4$  和氮气混合气体浓度测量时的谱线交叉干扰共性问题为例,提出了基于支持向量回归模型的混合气体浓度解析方法,实现了  $\text{CH}_4$  和微量 CO 的混合气体浓度的同时测量. 为了确保浓度回归模型的准确性和高效性,根据吸收谱线的特征对  $2f/1f$  数据进行了合理切割和利用,分别得到了两种气体的回归预测模型. 基于这两个模型,进行了固定配比的验证实验、混合气体浓度定量解析及组分变化过程实验. 在固定配比实验中,可以得到 CO 和  $\text{CH}_4$  气体的测量值与真实值的相关性分别为 0.998, 0.9995,证明了该系统具有良好的泛化能力. 混合气体浓度同时测量实验说明在高浓度  $\text{CH}_4$  与微量 CO 气体同时共存的气体环境下,CO 气体回归计算结果绝对误差不超过  $2 \times 10^{-6}$ ,  $\text{CH}_4$  回归结果的绝对误差不超过  $0.2 \times 10^{-2}$ . 上述结果证明,基于 SVR 原理构建的回归预测模型具有自动提取吸收谱线特征、抗干扰性强、回归预测结果准确等优点,可有效地解决激光光谱技术在混合气体测量过程中遇到的吸收谱线重叠干扰问题,为该技术面临的共性技术难题及其应用提供了新的解决方法.

## 参考文献

- [1] Zhang Z R, Pang T, Yang Y, Xia H, Cui X J, Sun P S, Wu B, Wang Y, Sigrist M W, Dong F Z 2016 *Opt. Express* **24** A943
- [2] Wang F P, Chang J, Wang Q, Wei W, Qin Z G 2017 *Sens. Actuators A* **259** 152
- [3] Avetisov V, Bjoroey O, Wang J, Geiser P, Paulsen K G 2019

- Sensors* **19** 5313
- [4] Nwaboh J A, Werhahn O, Ortwein P, Schiel D, Ebert V 2013 *Meas. Sci. Technol.* **24** 015202
- [5] Vallon R, Soutade J, Verant J L, Meyers J, Paris S, Mohamed A 2010 *Sensors* **10** 6081
- [6] Xia H, Wu B, Zhang Z R, Pang T, Dong F Z, Wang Y 2013 *Acta Phys. Sin.* **62** 214208 (in Chinese) [夏滑, 吴边, 张志荣, 庞涛, 董凤忠, 王煜 2013 物理学报 **62** 214208]
- [7] Zhang L W, Zhang Z R, Sun P S, Pang T, Xia H, Cui X J, Guo Q, Sigrist M W, Shu C M, Shu Z F 2020 *Spectrochim. Acta, Part A* **239** 118495
- [8] He Q X, Dang P P, Liu Z W, Zheng C T, Wang Y D 2017 *Opt. Quantum Electron.* **49** 115
- [9] Kluczynski P, Gustafsson J, Lindberg A M, Axner O 2001 *Spectrochim. Acta, Part B* **56** 1277
- [10] Wang F, Wu Q, Huang Q X, Zhang H D, Yan J H, Cen K F 2015 *Opt. Commun.* **346** 53
- [11] Lin X, Yu X L, Li F, Zhang S H, Xin J G, Chang X Y 2013 *Appl. Phys. B: Lasers Opt.* **110** 401
- [12] Köhring M, Huang S, Jahjah M, Jiang W, Ren W, Willer U, Caneba C, Yang L, Nagrath D, Schade W, Tittel F K 2014 *Appl. Phys. B* **117** 445
- [13] Zhang T, Kang J, Meng D, Wang H, Mu Z, Zhou M, Zhang X, Chen C 2018 *Sensors* **18** 4295
- [14] Zhang Z R, Wu B, Xia H, Pang T, Wang G X, Sun P S, Dong F Z, Wang Y 2013 *Acta Phys. Sin.* **62** 234204 (in Chinese) [张志荣, 吴边, 夏滑, 庞涛, 王高旋, 孙鹏帅, 董凤忠, 王煜 2013 物理学报 **62** 234204]
- [15] Gabrysch M, Corsi C, Pavone F S, Inguscio M 1997 *Appl. Phys. B* **65** 75
- [16] Cai T D, Gao G Z, Wang M R 2016 *Opt. Express* **24** 859
- [17] Malegori C, Marques E J N, Freitas S T d, Pimentel M F, Pasquini C, Casiraghi E 2017 *Talanta* **165** 112
- [18] Zhang Z R, Xia H, Dong F Z, Pang T, Wu B 2013 *Opt. Precis. Eng.* **21** 2771 (in Chinese) [张志荣, 夏滑, 董凤忠, 庞涛, 吴边 2013 光学精密工程 **21** 2771]
- [19] Shao L G, Fang B, Zheng F, Qiu X B, He Q S, Wei J L, Li C L, Zhao W X 2019 *Spectrochim. Acta, Part A* **222** 117118
- [20] Qu J, Chen H Y, Liu W Z, Zhang B, Li Z B 2015 *Proceedings of the 2015 International Conference on Intelligent Systems Research and Mechatronics Engineering* Zhengzhou, PRC, April 11–13, 2015 p2111
- [21] Laref R, Losson E, Sava A, Adjallah K, Siadat M 2018 *2018 Ieee International Conference on Industrial Technology (Icit)* Lyon, France, February 19–22, 2018 p1335
- [22] Reid J, Labrie D 1981 *Appl. Phys. B* **26** 203
- [23] Smola A J, Scholkopf B 2004 *Stat. Comput.* **14** 199
- [24] Gordon I E, Rothman L S, Hill C, et al. 2017 *J. Quant. Spectrosc. Radiat. Transfer* **203** 3

## Separation and analysis method of overlapping absorption spectra with cross interference in gas mixture measurement\*

Wang Qian-Jin<sup>1)2)</sup> Sun Peng-Shuai<sup>2)</sup> Zhang Zhi-Rong<sup>1)2)3)4)†</sup>  
 Zhang Le-Wen<sup>2)5)</sup> Yang Xi<sup>1)2)</sup> Wu Bian<sup>2)</sup>  
 Pang Tao<sup>2)</sup> Xia Hua<sup>2)</sup> Li Qi-Yong<sup>6)</sup>

1) (School of Environment Science and Optoelectronic Technology, University of Science and Technology of China, Hefei 230026, China)

2) (Anhui Provincial Key Laboratory of Photonic Devices and Materials, Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)

3) (Key Laboratory of Environmental Optics & Technology, Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)

4) (Advanced Laser Technology Laboratory of Anhui Province, National University of Defense Technology, Hefei 230037, China)

5) (University of Science and Technology of China, Hefei 230022, China)

6) (Monitoring Center of Ecology and Environment of Anhui Province, Hefei 230071, China)

(Received 7 February 2021; revised manuscript received 11 March 2021)

### Abstract

The interference between overlapping gas absorption lines often occurs in the measurement of multi-component gas mixture with using tunable diode laser absorption spectroscopy (TDLAS). This is also the main

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 11874364, 41877311, 42005107), the National Key Research and Development Program of China (Grant No. 2017YFC0805004), the Key Research and Development Projects of Anhui Province, China (Grant No. 201904c03020005), and the "Spark" Fund Project of Hefei Institutes of Physics Science, Chinese Academy Sciences, China (Grant No. YZJJ2020QN8).

† Corresponding author. E-mail: zhangzr@aiofm.ac.cn

problem of the technology in some applications. For instance, in the early application of multi-component gas mixture measurement in coal mines, we found that the absorption lines of carbon monoxide (CO) and methane (CH<sub>4</sub>) seriously overlapped. The absorption signal of trace CO gas was annihilated and could not be effectively demodulated, especially in the presence of high concentration of CH<sub>4</sub>. This problem could not be solved just by accurately selecting the spectral lines due to the band absorption of CH<sub>4</sub>. Therefore, in this paper, we introduce the support vector regression (SVR) model to deal with the interference between CO and CH<sub>4</sub> absorption lines. The spectral signals of 14 groups of mixed gases with different concentrations of CO and CH<sub>4</sub> are used as the training sets, and the five-fold cross-validation is adopted to prevent the model from overfitting. After 15 iterations in 30 seconds, the optimal regression model of CO and CH<sub>4</sub> can be obtained respectively. Furthermore, it is worth noting that based on the experimental data, the linear kernel function is selected to construct the two gas SVR models, and the parameters of the SVR models are optimized by the sequential minimal optimization(SMO) algorithm. With the assistance of the SVR models, the absorption spectra of the two gases can be demodulated effectively, and finally the accurate measurement results are obtained. The measurement results show that the absolute error of trace CO and CH<sub>4</sub> concentration(volume fraction of gas) are less than  $2 \times 10^{-6}$  and  $0.2 \times 10^{-2}$  respectively. Meanwhile, the correlation coefficient between the measured values and the actual values of CO and CH<sub>4</sub> are 0.998 and 0.9995, respectively. In addition, the dynamic stability for each of the two regression models is fully verified by the experiment of the inflation process. Consequently, this method can eliminate the interference between the overlapping spectra, and can fully meet the requirements for accurately measuring the gas mixture. We hope that the SVR model can provide an effective solution for the real-time monitoring of multi-component gas mixture, and thus greatly improving the adaptability of TDLAS technology in the future.

**Keywords:** tunable diode laser absorption spectroscopy (TDLAS), gas mixture, overlapping spectral lines, support vector regression

**PACS:** 42.62.Fi, 87.55.kd, 42.55.Px, 33.20.Ea

**DOI:** [10.7498/aps.70.20210286](https://doi.org/10.7498/aps.70.20210286)