

• 研究论文 •

## 二维相关分析中参考谱的选择对分离重叠峰的影响

王梦吟 赖祖亮 武培怡\*

(复旦大学高分子科学系和先进材料实验室 聚合物分子工程教育部重点实验室 上海 200433)

**摘要** 二维相关光谱是一项将光谱强度看作两个独立的光谱变量的函数的技术, 它是由动态光谱经过数学转化后得到的. 在扰动过程中, 动态光谱等于实际测得的光谱减去参考谱, 参考谱的选择是任意的, 甚至可以为 0, 但是在实际应用时, 人们逐渐发现参考谱的选择会对二维相关光谱产生一定的影响. 本篇文章采用模拟的方法, 建立光谱模型, 光谱强度按 e 指数形式单调变化, 比较以平均谱为参考谱和不设参考谱得到的二维相关光谱图, 分析它们的区别, 在不同参考谱条件下, 利用二维相关光谱分离重叠峰, 得到的结果也不相同, 将两种条件综合利用可以得到更多更正确的信息.

**关键词** 二维相关光谱; 参考谱; 重叠峰

## Effect of Selection of Reference Spectrum on the Separation of Overlapped Peaks

Wang, Mengyin Lai, Zuliang Wu, Peiyi\*

(Department of Macromolecular Science and Advanced Materials Laboratory, Key Laboratory of Polymer Engineering Science (Ministry of Education), Fudan University, Shanghai 200433)

**Abstract** 2D correlation spectroscopy is a technique where the spectral intensity is defined as a function of two independent spectral variables. With an external perturbation, the selection of the reference spectrum may affect the results of 2D correlation analysis. In many cases, 2D correlation spectroscopy has been used to separate highly overlapped peaks. In this report, the effects of separation have been tested in the generalized situations using simulated spectra with different reference spectra and assuming the intensity changes of bands take the exponential forms. It has been found that it provides more useful information by using time-averaged reference and none reference at the same time.

**Keywords** 2D correlation spectroscopy; reference spectrum; overlapped band

1993年, Noda<sup>[1]</sup>归纳出了广义二维相关方法的规律, 此后二维相关光谱得到了广泛的应用, 在应用过程中人们逐渐发现数据的预处理、参考谱的选择均可能对分析结果产生影响, 为了更好地运用二维相关光谱, 掌握其特点与性质, 科学家对此展开了一系列系统的研究, Huang<sup>[2]</sup>发现 Noda 的顺序判定规则存在漏洞, 在某些情况下会得到错误的结论, 而 Czarniecki<sup>[3]</sup>则提出运用适当的参考峰对原始谱归一化, 可以有效地减少线性非选择

性变化对二维相关分析的干扰.

目前为止, 针对参考谱进行的研究还不多, Czarniecki 发现尽管大多数情况下以平均谱作为参考谱, 但其实不选择参考谱同样可以提供有用的信息, 尤其是对于受噪音和基线漂移影响的体系<sup>[4,5]</sup>, 而对于峰有位移的情况, 选择不同的参考谱可能会得到不同的结果<sup>[6]</sup>. 本篇文章主要讨论参考谱的选择对重叠峰分离的影响, 比较在平均谱和无参考谱条件下得到的分析结果, 将两

\* E-mail: peiyiwu@fudan.edu.cn; Tel: 0086-021-65643255

Received October 23, 2008; revised February 16, 2009; accepted April 28, 2009.

国家自然科学基金(20774022)、上海市优秀学科带头人计划(No. 07XD14002)资助项目.

种条件综合利用以提高分离重叠峰的效果。

## 1 光谱模型

本篇文章通篇采用了模拟的方法, 每组数据包括 11 条模拟谱线, 分辨率为  $1 \text{ cm}^{-1}$ , 峰形为洛伦兹函数形式, 函数表达式如公式(1)所示:

$$L(x) = \frac{a}{1+(x-b)^2/c^2} \quad (1)$$

特征峰强度以  $e$  指数形式单调变化, 峰的起始值为  $A_0$ 、终止值为  $A_t$ 、半高宽为  $W_{1/2}$ , 半高宽是吸收谱带高度为峰值一半时谱带的全宽,  $A_0$ ,  $A_t$  和  $W_{1/2}$  的数值由具体情况决定, 模拟峰形如图 1 所示。

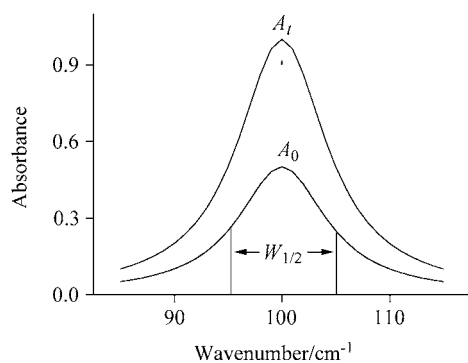


图 1 模拟峰形示意图  
Figure 1 Simulated spectra

## 2 二维相关光谱的计算

进行二维相关分析前, 对模拟光谱不需要进行任何预处理. 动态光谱  $\tilde{y}(v, t)$  的定义式如公式(2)所示<sup>[7]</sup>, 在外扰过程中等于原始谱减去参考谱, 而在非外扰过程中等于 0.

$$\tilde{y}(v, t) = \begin{cases} y(v, t) - \bar{y}(v) & T_{\min} < t < T_{\max} \\ 0 & \text{其它} \end{cases} \quad (2)$$

同步光谱强度  $\Phi(v_1, v_2)$  等于不同波数  $(v_1, v_2)$  的动态光谱强度的矢量积, 而异步光谱强度  $\Psi(v_1, v_2)$  则等于  $v_1$  处的动态光谱强度与  $v_2$  处的动态光谱强度的 Hilbert-Noda 转换式的矢量积<sup>[8]</sup>:

$$\Phi(v_1, v_2) = \frac{1}{m-1} \tilde{y}(v_1)^T \tilde{y}(v_2) \quad (3)$$

$$\Psi(v_1, v_2) = \frac{1}{m-1} \tilde{y}(v_1)^T N \tilde{y}(v_2) \quad (4)$$

公式(4)中的  $N$  代表 Hilbert-Noda 转化式,  $m$  是光谱的个数, 在本次实验中等于 11. 同步图中, 若  $(v_1, v_2)$  处的峰是正的, 则  $v_1, v_2$  处的光谱强度同向变化, 反之则是异向; 当  $v_1 > v_2$  时, 若同步图中  $(v_1, v_2)$  处和异步图中  $(v_1, v_2)$  处的峰值乘积为正, 则  $v_1$  先于  $v_2$  变化, 若为负, 则是  $v_1$  后于  $v_2$  变化<sup>[1]</sup>.

二维相关分析的结果利用 shige 软件得到, 该软件是由日本关西大学的 Shigeaki Morita 和 Yukihiro Ozaki 教授联合编写, 图像中的阴影部分代表负峰。

## 3 结果和讨论

### 3.1 不同参考谱的二维相关分析结果

为了研究不同参考谱的二维相关分析结果, 建立了第一组模拟数据, 包含 11 条谱线, 每条谱线含有 5 个特征峰, 半高宽均为  $10 \text{ cm}^{-1}$ , 各峰的起始值、终止值、变化幅度以及变化速度如表 1 所示, 其中  $1700 \text{ cm}^{-1}$  处特征峰的峰值不变, 得到的模拟谱图如图 2a 所示, 扰动过程中的平均谱如图 2b 所示。

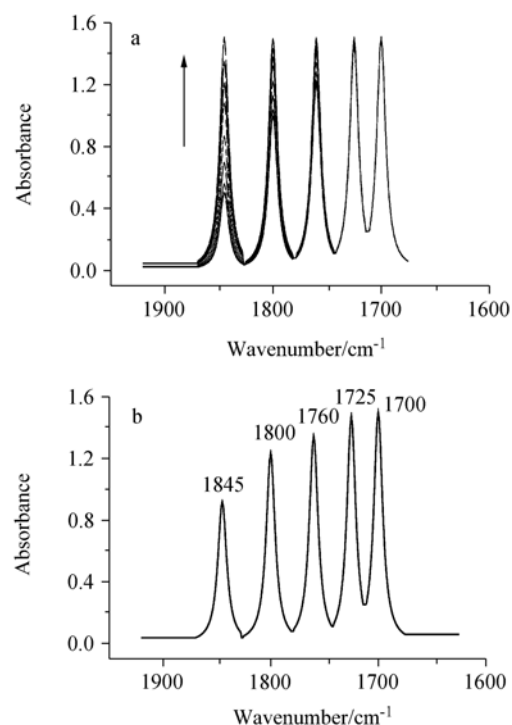


图 2 模型 1 的模拟谱(a)和平均谱(b)

Figure 2 (a) Simulated spectra and (b) average spectrum of model 1

以平均谱为参考谱得到的二维相关光谱图(图 3)中, 变化速度为零的  $1700 \text{ cm}^{-1}$  没有对应的相关峰, 变化速度很小的  $1725 \text{ cm}^{-1}$  对应的相关峰均很微弱, 而不设参考谱得到的二维相关光谱图(图 4)中,  $1700$  和  $1725 \text{ cm}^{-1}$

表 1 光谱模型 1 的模拟数据

Table 1 Simulated data of spectra model 1

峰位置/ cm <sup>-1</sup>	起始峰值 A <sub>0</sub>	终止峰值 A <sub>11</sub>	峰值变化幅度 (A <sub>11</sub> -A <sub>0</sub> )	峰值变化速度 ln(A <sub>11</sub> /A <sub>0</sub> )/10
1700	1.5	1.5	0	0
1725	1.45	1.5	0.05	3.39×10 <sup>-3</sup>
1760	1.2	1.5	0.3	2.23×10 <sup>-2</sup>
1800	1	1.5	0.5	4.05×10 <sup>-2</sup>
1845	0.5	1.5	1.0	1.10×10 <sup>-1</sup>

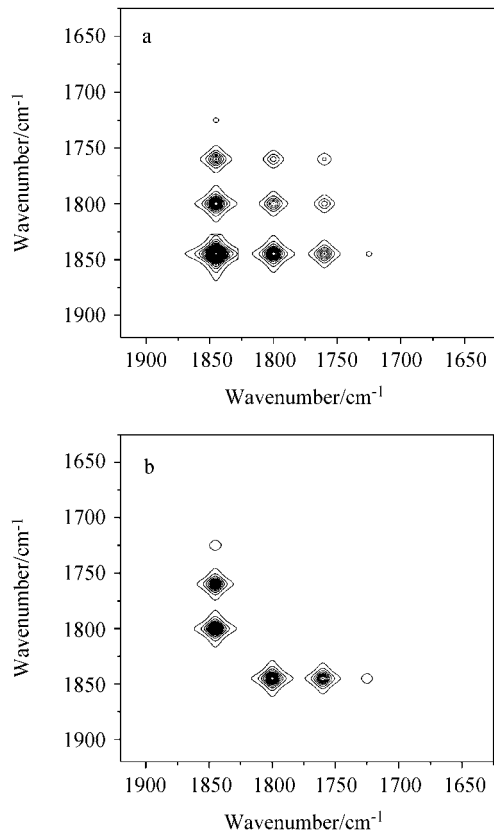


图 3 以平均谱为参考谱, 模型 1 的同步图(a)和异步图(b)  
Figure 3 (a) Synchronous and (b) asynchronous correlation spectra constructed from model 1 with average reference spectrum

均有十分明显的相关峰。

利用功率谱, 即同步图中自相关峰所在的对角线方向的切片谱, 可以更清晰地比较参考谱选择对二维相关光谱图的影响, 功率谱反映了同步图中自相关峰的峰高<sup>[5]</sup>。比较平均参考谱和不设参考谱时模型 1 的功率谱 (图 5), 发现图 5a 中, 变化速度为 0 的 1700 cm<sup>-1</sup> 没有峰, 而剩下四个峰峰高的排列顺序为: 1845 > 1800 > 1760 > 1725 cm<sup>-1</sup>, 图 5b 中, 变化速度为零的 1700 cm<sup>-1</sup> 峰值最大, 峰高排列顺序为: 1700 > 1725 > 1760 > 1800 > 1845 cm<sup>-1</sup>。

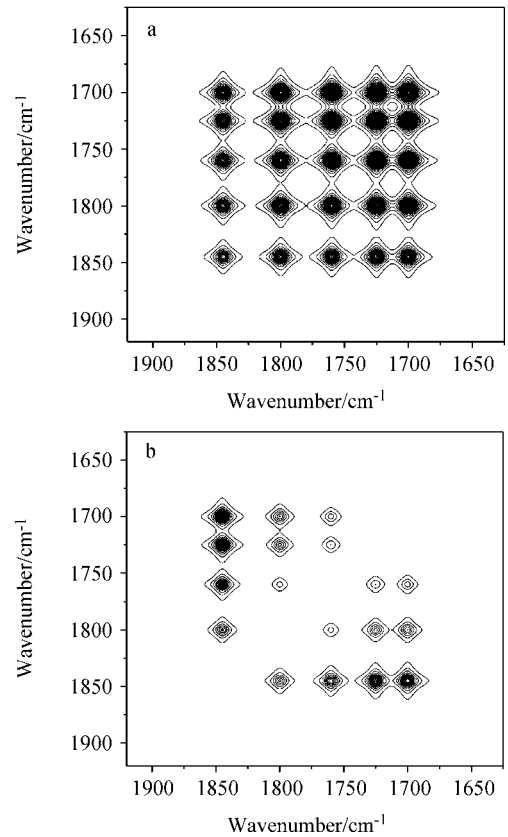


图 4 无参考谱, 模型 1 的同步图(a)和异步图(b)  
Figure 4 (a) Synchronous and (b) asynchronous correlation spectra constructed from model 1 without reference spectrum

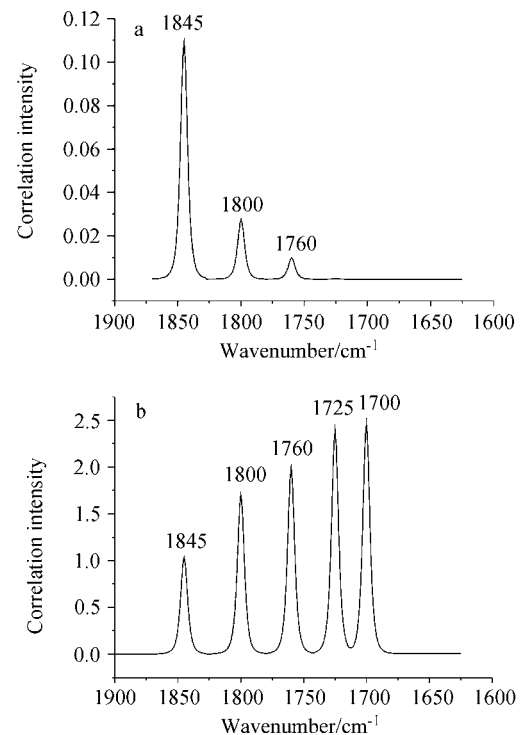


图 5 模型 1 的功率谱: (a)平均参考谱; (b)无参考谱  
Figure 5 Power spectrum of model 1 constructed (a) with average reference spectrum and (b) without reference spectrum

不设参考谱时,数据的动态光谱等于原始谱,峰高决定其信号强度,平均峰值可能会影响其二维相关峰的强度.为了验证这一推测,建立第二组模拟数据,同样含有11条谱线,特征峰的形态、分布以及平均峰值与模型1均相同,但变化速度均为0,得到的模拟谱图如图6a所示,扰动过程中的平均谱如图6b所示.

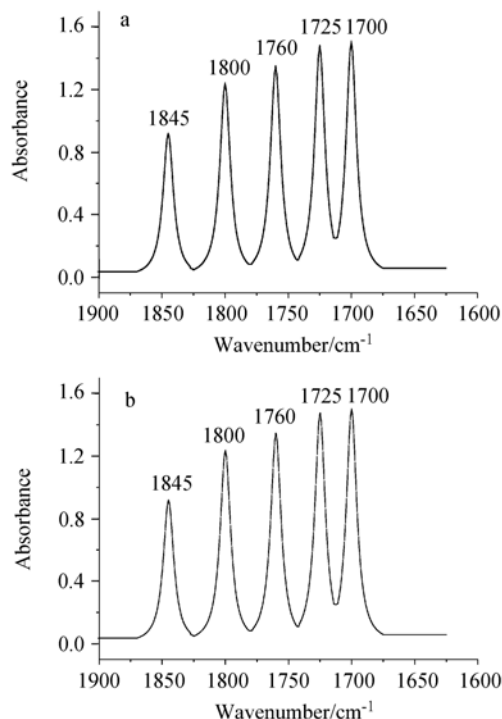


图6 模型2的模拟谱(a)和平均谱(b)

Figure 6 (a) Simulated spectra and (b) average spectra of model 2

比较不设参考谱时二者的功率谱(图7),发现它们十分相似,利用t检验,证明在检验水准 $\alpha=0.05$ ,即置信度为95%时,两组数据没有明显区别,说明特征峰的平均峰值对功率谱的峰高有较大的影响,平均峰值高的特征峰对应的信号较强.

从上面结果可以看出,对于同一个体系,选择不同的参考谱,二维相关峰的强度有较大区别,不设参考谱时,平均峰值大的特征峰对应的二维相关峰较强,功率谱中的峰值也较高,对于受噪音干扰的体系,特征峰的平均峰值远高于噪音,因此使用无参考谱的条件可以有效减少噪音的影响;而以平均谱为参考谱时,变化速度大的特征峰对应的相关峰的强度较大,这是由于选择不同参考谱得到的动态光谱不同所造成的,如图8所示,以平均谱为参考谱和不设参考谱时,模型1的动态光谱图有明显的差别.而二维同步图的自相关峰强度是由动态光谱计算得到的,与之平方成正比,所以以平均谱为参考谱时,功率谱中变化速度大的特征峰信号强,而不

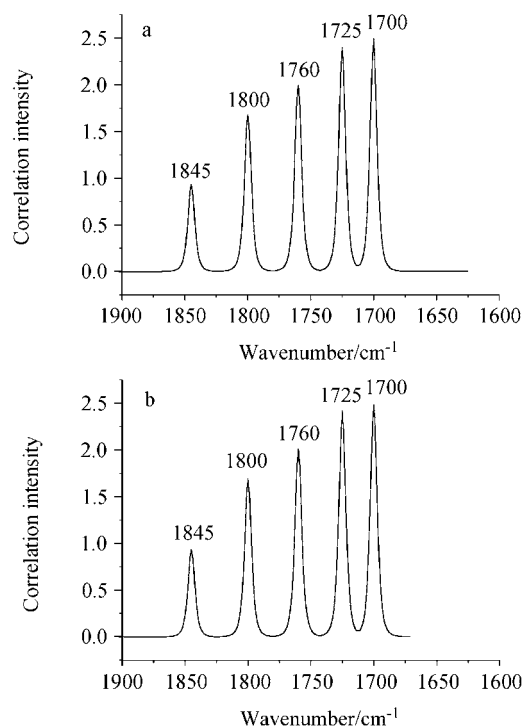


图7 不设参考谱时,数据1(a)和数据2(b)的功率谱  
Figure 7 Power spectrum of (a) model 1 and (b) model 2 constructed without reference spectrum

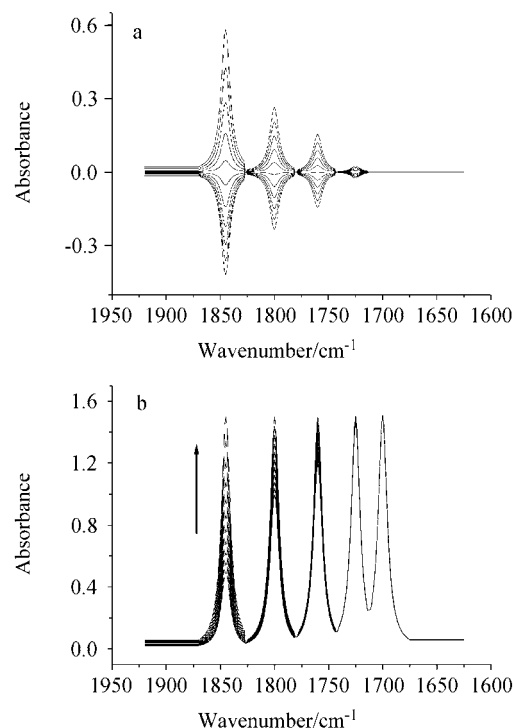


图8 模型1的动态光谱:(a)平均参考谱;(b)无参考谱  
Figure 8 Dynamic spectrum of model 1 constructed (a) with average reference spectrum and (b) without reference spectrum

设参考谱时,功率谱中平均峰值大的特征峰信号强.

### 3.2 参考谱对重叠峰分离的影响

二维相关光谱的主要优点之一就是极大提高了谱图的分辨率, 常用于分析包含重叠峰的复杂体系, 继曲线拟合、二阶导、傅立叶去卷积<sup>[9]</sup>后, 成为第四种常用分离重叠峰的方法. 通常采用平均参考谱, 但通过上面的分析发现不设参考谱也可以为分离重叠峰提供一定的信息, 该条件下得到的二维相关分析结果对平均峰值大的特征峰更敏感, 不设参考谱的条件有利于分离平均峰值大的特征峰.

数据 3 模拟的就是一个由两个峰重叠而成的体系, 峰形为洛伦兹函数形式, 1695 和 1705  $\text{cm}^{-1}$  处的两个峰, 前者变化速度较大, 后者平均峰值较高, 半高宽均为 10  $\text{cm}^{-1}$ , 详细数据见表 2, 模拟得到的谱图如图 9 所示.

表 2 光谱模型 3 的模拟数据  
Table 2 Simulated data of spectra model 3

峰位置/ $\text{cm}^{-1}$	起始峰值 $A_0$	终止峰值 $A_{11}$	峰值变化幅度 $(A_{11}-A_0)$	峰值变化速度 $\ln(A_{11}/A_0)/10$
1695	0.5	0.75	0.25	$4.05 \times 10^{-2}$
1705	1.5	1.55	0.05	$3.28 \times 10^{-3}$

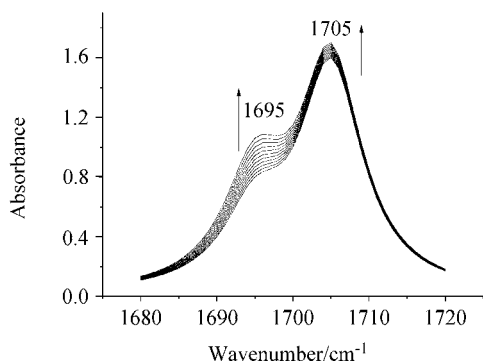


图 9 模型 3 的模拟谱图  
Figure 9 Simulated spectra of model 3

比较不同参考谱时模型 3 的功率谱(图 10), 发现它们的分离效果有明显区别. 利用 Origin7.5 软件确定峰位置, 不设参考谱时, 平均峰值大的 1705  $\text{cm}^{-1}$  可被分离出来, 而使用平均参考谱时, 变化速度大的 1695  $\text{cm}^{-1}$  可被分离出来, 而 1705  $\text{cm}^{-1}$  的峰却很微弱.

通过上述模拟, 发现利用二维相关光谱分离重叠峰时, 选择不同的参考谱得到的结果有较大区别, 要想能更加完整地认识这个重叠峰, 可以同时使用两种参考谱条件结合得到的结果.

尝试用上述方法研究更为复杂的情况, 比如三个峰重叠的体系. 聚  $\epsilon$ -己内酯和多酚儿茶酚的混合体系包含了三种类型的羰基, 分别是无定型(1736  $\text{cm}^{-1}$ )、结晶型(1724  $\text{cm}^{-1}$ )、氢键型(1706  $\text{cm}^{-1}$ ), 纯的聚  $\epsilon$ -己内酯中含有前两种羰基, 加入儿茶酚后出现第三种类型的羰基,

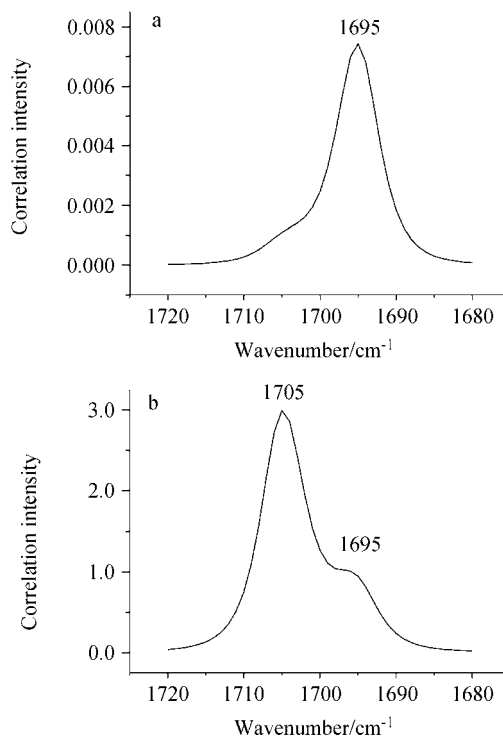


图 10 模型 3 的功率谱: (a) 平均参考谱; (b) 无参考谱  
Figure 10 Power spectrum of model 3 constructed (a) with average reference spectrum and (b) without reference spectrum

三者位置较接近, 容易形成难分离的重叠峰, 随着儿茶酚质量分数的增加, 三个特征峰的峰值随之改变, 变化的速度各不相同<sup>[10]</sup>. 用模拟的方法构建一个类似的体系, 即模型 4, 峰形为洛伦兹函数的形式, 数据见表 3, 起始时, 体系中只含有纯的聚  $\epsilon$ -己内酯, 故 1706  $\text{cm}^{-1}$  处的吸收强度为 0, 从第二条谱线开始有峰值, 模拟的谱图见图 11.

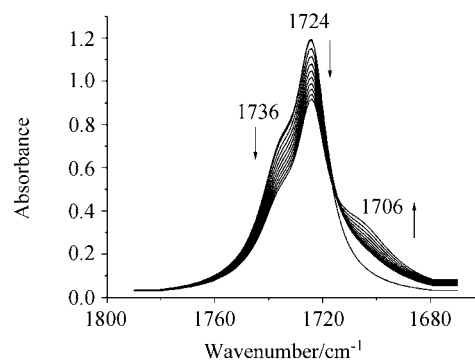


图 11 模型 4 的模拟光谱图  
Figure 11 Simulated spectra of model 4

比较不同参考谱条件下模型 4 的功率谱(图 12), 以平均谱为参考谱时, 1706 和 1724  $\text{cm}^{-1}$  有明显的峰, 位置均有偏差, 而不设参考谱时, 平均峰值最高的 1724  $\text{cm}^{-1}$  有明显的峰, 位置不变.

表3 光谱模型4的模拟数据  
Table 3 Simulated data of spectra model 4

峰位置/ $\text{cm}^{-1}$	半高宽/ $\text{cm}^{-1}$	起始峰值 $A_0$	终止峰值 $A_{11}$	峰值变化速度 $\ln(A_{11}/A_0)/10$
1736	19	0.5	0.3	$-5.11 \times 10^{-2}$
1724	13	1	0.7	$-3.57 \times 10^{-2}$
1706	30	0.1	0.25	$9.16 \times 10^{-2}$

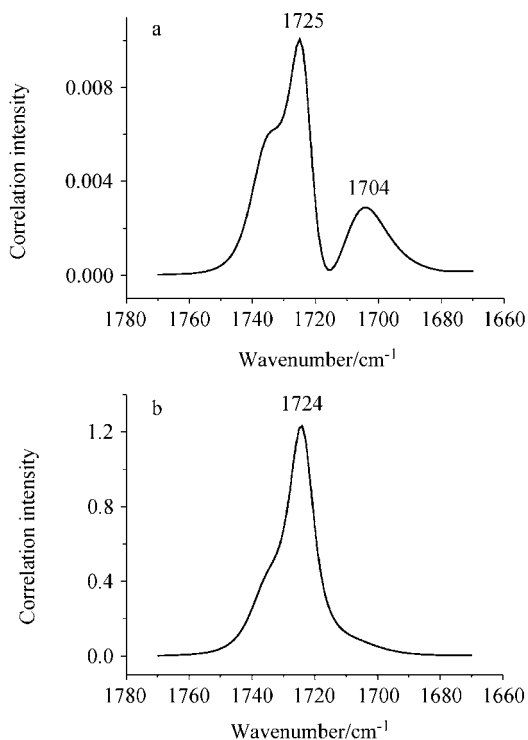


图12 模型4的功率谱: (a)平均参考谱(b)无参考谱  
Figure 12 Power spectrum of model 4 constructed (a) with average reference spectrum and (b) without reference spectrum

以平均谱为参考谱时, 模型4的动态光谱如图13

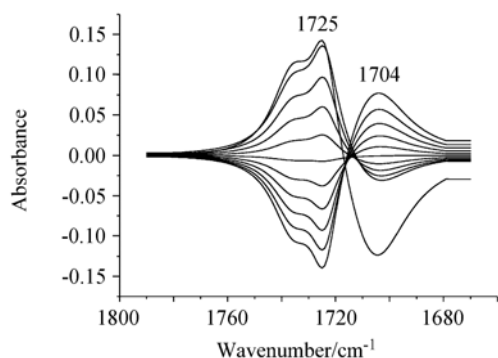


图13 平均参考谱时, 模型4的动态光谱  
Figure 13 Dynamic spectrum of model 4 constructed with average reference spectrum

所示, 其中变化速度最快的  $1706 \text{ cm}^{-1}$  附近信号较强, 但峰的位置偏移至  $1704 \text{ cm}^{-1}$ , 同样平均峰值最大的  $1724 \text{ cm}^{-1}$  附近信号也较强, 峰的位置偏移至  $1725 \text{ cm}^{-1}$ , 这可能是由于重叠程度较大, 致使峰形发生改变所造成的. 动态光谱进一步转化得到二维相关光谱, 对应相关峰的位置也随之改变.

利用二维相关光谱分离重叠峰, 通常选用平均谱为参考谱, 得到较好的分离效果, 但是当峰的重叠程度较大时, 判定得到的峰位置容易产生误差. 而不设参考谱时, 可以分离出平均峰值较大的峰, 位置不变. 所以我们可以同时利用两种参考谱条件, 将所得结果综合分析, 得到更多更准确的信息.

#### 4 结论

参考谱的选择会影响二维相关分析的结果, 以平均谱为参考谱或不设参考谱, 得到的二维相关峰的强度有明显区别. 以平均谱为参考谱时, 变化速度大的峰对应的相关峰信号强, 功率谱中峰值也较高, 有利于找出在外界扰动下产生感应的特征峰; 不设参考谱时, 平均峰值大的峰对应的相关峰信号强, 功率谱中峰值也较高, 实际体系中特征峰的平均峰值往往远大于噪音, 所以利用不设参考谱的条件可以有效降低噪音对体系的干扰. 利用二维相关光谱分离重叠峰, 通常使用平均参考谱, 但是当峰的重叠程度较大时, 判定得到的峰位置可能会有误差, 而不使用参考谱时, 可以分离出平均峰值较高的峰, 且位置不变, 因此将两种参考谱条件综合使用, 可以有效提高重叠峰的分离效果, 与单纯使用平均参考谱相比, 能够得到更多更正确的信息.

#### References

- 1 Noda, I. *Appl. Spectrosc.* **1993**, *47*, 1329.
- 2 Huang, H. *Anal. Chem.* **2007**, *79*, 8281.
- 3 Czarnecki, M. A. *Appl. Spectrosc.* **1999**, *53*, 1392.
- 4 Czarnecki, M. A. *Appl. Spectrosc.* **1998**, *52*, 1583.
- 5 Czarnecki, M. A. *Appl. Spectrosc.* **2003**, *57*, 991.
- 6 Czarnecki, M. A. *Appl. Spectrosc.* **2000**, *54*, 986.
- 7 Noda, I.; Dowrey, A. E.; Marcott, C.; Story, G. M.; Ozaki, Y. *Appl. Spectrosc.* **2000**, *54*, 236A.
- 8 Noda, I. *Appl. Spectrosc.* **2000**, *54*, 994.
- 9 Czarnecki, M. A.; Ozaki, Y. *Spectrochim. Acta, Part A* **1996**, *52*, 1593.
- 10 Zhu, B.; Li, J. C.; He, Y.; Yoshie, N.; Inoue, Y. *Macromol. Biosci.* **2003**, *3*, 684.

(A0810234 Lu, Y.; Lu, Z.)