

Technology Reports

Application of Band-Target Entropy Minimization to On-Line Raman Monitoring of an Organic Synthesis. An Example of New Technology for Process Analytical Technology

Effendi Widjaja,[†] Ying Yan Tan,[‡] and Marc Garland^{*,†,‡}

Process Science and Modeling, Institute of Chemical and Engineering Sciences, 1 Pesek Rd, Jurong Island, Singapore 627833, and Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576

Abstract:

The hydrolysis of acetic anhydride to acetic acid in water as solvent was monitored by Raman microscopy. Both static and flow-through configurations were used in the experiments, and various experimental designs, i.e., multiple-experimental runs and multiple-perturbation semibatch mode, were considered. Various spectral data preprocessing was performed and band-target entropy minimization (BTEM) was used in the spectral analysis to recover the pure-component spectra from the multicomponent data. Good and consistent spectral estimates of the solutes acetic anhydride and acetic acid were recovered. In addition, the pure-component spectrum of white-light interference was recovered. Together, these estimates permitted very good estimates of the individual time-dependent signal contributions. Taken together, the present results suggest that the combination of Raman spectroscopy and BTEM has considerable potential for organic syntheses and process analysis. The combination of Raman spectroscopy and BTEM represents another approach for reaction monitoring in process analytical technologies (PAT).

Introduction

Reaction monitoring has been one of the major interests of chemical engineers since it is closely associated with subsequent chemical process control and chemical process optimization. Recently, Raman spectroscopy has been one important process analytical technique that has experienced considerable instrumental improvement. From the discovery of Raman scattering in 1928, this technique has progressed tremendously, overcoming many impediments faced in the past. Earlier models of Raman spectrometers faced numerous problems including weak intensity, inefficient light collection and detection, and fluorescence interferences. These earlier designs have been superseded by newer models having more

sophisticated features such as a charged coupled device (CCD) array detectors, low-cost lasers, holographic optical elements, and fiber-optic probes. The renaissance of Raman spectroscopy for on-line process monitoring can be readily seen in some of the more recent process research, such as on-line monitoring of dispersion polymerization,¹ in situ monitoring of anhydride and acid conversion in supercritical water,² in situ monitoring of the acid-catalyzed esterification,³ kinetic investigation of imine formation in chloroform,⁴ the formation of melamine formaldehyde resins,⁵ on-line monitoring of the hydrolysis of acetonitrile in near-critical water,⁶ in situ investigation of polymorphic transformation of progesterone,⁷ on-line investigations of industrial reactions,⁸ and on-line characterization of hydrogenation reactions.⁹ Another factor that undoubtedly helped create a Raman renaissance was the development of multivariate spectral data processing techniques, which are frequently used and collectively termed “chemometrics”.

One popular generic chemometric technique is self-modeling curve resolution (SMCR). It is widely used to deconvolute the observable, pure-component spectra present in a set of multicomponent spectroscopic measurements. The major advantage of this approach is that it does not rely on any spectral libraries for resolving the pure-component spectra. Thus, it can be applied qualitatively to identify virtually any species present in the system including intermediates and other non-isolatable compounds. In addition, this approach can be used quantitatively to determine the concentration of compounds present in the system, since a

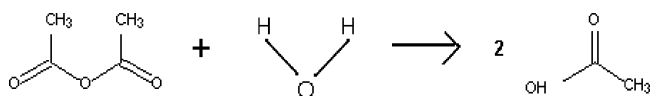
* To whom correspondence should be addressed. Present address: E-mail: marc_garland@ices.a-star.edu.sg. Dr. Marc Garland, Team Leader, Advanced Reaction Engineering, Process Analytics, and Chemometrics, Institute of Chemical and Engineering Sciences, 1 Pesek Rd, Jurong Island, Singapore 627833. Telephone: +65-67963947. Fax: +65-63166185.

[†] Institute of Chemical and Engineering Sciences.

[‡] National University of Singapore.

- (1) McCaffery, T. R.; Durant, Y. G. *J. Appl. Polym. Sci.* **2002**, *86*, 1507–1515.
- (2) Bell, W. C.; Booksh, K. S.; Myrick, M. L. *Anal. Chem.* **1998**, *70*, 332–339.
- (3) Ampiah-Bonney, R. J.; Walmsley, A. D. *Analyst* **1999**, *124*, 1817–1821.
- (4) Lee, M.; Kim, H.; Rhee, H.; Choo, J. *Bull. Korean Chem. Soc.* **2003**, *24*, 205–208.
- (5) Kip, B. J.; Berghmans, T.; Palmén, P.; Van der Pol, A.; Huys, M.; Hartwig, H.; Scheepers, M.; Wienke, D. *Vib. Spectrosc.* **2000**, *24*, 75–92.
- (6) Venardou, E.; Garcia-Verdugo, E.; Barlow, S. J.; Gorbaty, Y. E.; Poliakov, M. *Vib. Spectrosc.* **2004**, *35*, 103–109.
- (7) Wang, F.; Wachter, J. A.; Antosz, F. J.; Berglund, K. A. *Org. Process Res. Dev.* **2000**, *4*, 391–395.
- (8) Wiss, J.; Zilian, A. *Org. Process Res. Dev.* **2003**, *7*, 1059–1066.
- (9) Tumuluri, V. S.; Kemper, M. S.; Sheri, A.; Choi S.-R.; Lewis, I. R.; Avery, M. A.; Avery, B. A. *Org. Process Res. Dev.* **2006**, *10*, 927–933.

Scheme 1. Hydrolysis of acetic anhydride to 2 M acetic acid



mixed-component spectrum is just a linear superposition of pure-component spectra and their concentrations.

Among the currently available SMCR techniques, band-target entropy minimization (BTEM) is one of the newer and robust methods,¹⁰ which was developed on the basis of Shannon's information entropy concept.¹¹ Thus far, BTEM has been successfully applied to resolve pure-component spectra from various one-dimensional (1D) spectra from multicomponent systems, measured by FT-IR,^{12–18} Raman,^{19,20} XRD,²¹ and MS.²² In addition, recently it has also been applied to two-dimensional (2D) spectra, such as COSY and HSQC 2D NMR data.²³ The main strengths of BTEM, which make it unique and different from other SMCR methods, are the following: (1) its ability to recover pure-component spectra of species at sub-ppm levels, (2) considerably enhanced signal-to-noise ratio of recovered minor compounds, (3) no requirement for any a priori estimate of the number of species present in a system, and (4) goal-oriented approach, whereby the user targets a single spectral feature of interest, and the algorithm yields the full-range, reconstructed, pure spectrum associated with the targeted feature.

The use of BTEM to recover pure-component spectra from Raman data of a mixture has been thus far limited to nonreactive and off-line systems. In the present contribution, we combine the use of Raman spectroscopic measurements and the novel multivariate data analysis method, BTEM, to monitor the hydrolysis of acetic anhydride to acetic acid (Scheme 1). Raman spectroscopy is a better choice than FT-IR for this reaction since water has an intrinsically weak Raman signal, and thus the Raman signals of other reactants and product can be more easily observed and measured. Two different monitoring setups, consisting of static and flow-through modes were employed. This contribution demonstrates the applicability of combined BTEM and Raman

measurements for robust multivariate process analytical technology (PAT) and suggests the extension to real-time analysis and control of manufacturing processes.

Theoretical Method

Band-Target Entropy Minimization (BTEM). As previously mentioned, the primary use of BTEM is to extract the pure-component spectra out of a set of mixture spectra. Subsequently, the individual concentration profiles are obtained, which contain significant information on the dynamical changes of each compound involved in the system.

Let $I_{k \times \nu}$ represent the Raman intensity in the consolidated spectroscopic data matrix where k denotes the number of spectra taken and ν denotes the number of data channels associated with the spectroscopic range. The experimental intensities $I_{k \times \nu}$ have a bilinear data structure and can be described as a product of two submatrices, namely, the concentration matrix $C_{k \times s}$ and the Raman pure-component scattering coefficient matrix $J_{s \times \nu}$ (where s denotes number of observable species in the chemical mixture). Accordingly, the associated error matrix $\epsilon_{k \times \nu}$ contains both experimental error and model error (nonlinearities).²⁴

$$I_{k \times \nu} = C_{k \times s} J_{s \times \nu} + \epsilon_{k \times \nu} \quad (1)$$

The BTEM algorithm proceeds by first decomposing $I_{k \times \nu}$ into its singular vectors using singular value decomposition (SVD).²⁵ This is then followed by the appropriate transformation of right singular vectors, V^T , into physically meaningful, pure-component spectra.

$$I_{k \times \nu} = U_{k \times k} \Sigma_{k \times \nu} V_{\nu \times \nu}^T \quad (2)$$

Different from other SMCR methods, BTEM is uniquely developed to resolve one pure spectrum at a time. The number of eigenvectors, z , taken for inclusion in the transformation is usually much larger than s due to the nonlinearities present (nonstationary spectral characteristics). The number z is usually determined by identifying the vectors which possess localized signals of clear physical significance and retaining these, while discarding the vectors that are more-or-less randomly distributed noise.

$$\hat{J}_{1 \times \nu} = T_{1 \times z} V_{z \times \nu}^T \quad (3)$$

To extract a pure spectrum $\hat{J}_{1 \times \nu}$, a selected band present in the first few V^T vectors is targeted. The BTEM algorithm then retains this feature and, at the same time, returns an entire spectrum which has a minimum entropy. This routine is repeated for all important observable physical features in the selected V^T vectors. A superset of reconstructed spectra is obtained, and this set is reduced to eliminate redundancies. This results in an enumeration of all observable pure-component spectra.

The targeted bands are retained during the reconstruction. As part of the process, the resulting pure spectral patterns

(10) Widjaja, E. *Development of Band-Target Entropy Minimization (BTEM) and Associated Software Tools*. Ph.D. Thesis, National University of Singapore, 2002.

(11) Shannon, C. E. *Bell Syst. Tech. J.* **1948**, *3*, 379–423.

(12) Chew, W.; Widjaja, E.; Garland, M. *Organometallics* **2002**, *21*, 1982–1990.

(13) Widjaja, E.; Li, C. Z.; Garland, M. *Organometallics* **2002**, *21*, 1991–1997.

(14) Li, C. Z.; Widjaja, E.; Chew, W.; Garland, M. *Angew. Chem., Int. Ed.* **2002**, *41*, 3785–3789.

(15) Li, C. Z.; Widjaja, E.; Garland, M. *J. Am. Chem. Soc.* **2003**, *125*, 5540–5548.

(16) Li, C. Z.; Widjaja, E.; Garland, M. *J. Catal.* **2003**, *213*, 126–134.

(17) Widjaja, E.; Li, C. Z.; Chew, W.; Garland, M. *Anal. Chem.* **2003**, *75*, 4499–4507.

(18) Widjaja, E.; Li, C. Z.; Garland, M. *J. Catal.* **2004**, *223*, 278–289.

(19) Ong, L. R.; Widjaja, E.; Stanforth, R.; Garland, M. *J. Raman Spectrosc.* **2003**, *34*, 282–289.

(20) Widjaja, E.; Crane, N. J.; Chen, T.-S.; Morris, M. D.; Igelzi, M. A., Jr.; McCreddie, B. R. *Appl. Spectrosc.* **2003**, *57*, 1353–1362.

(21) Guo, L. F.; Kooli, F.; Garland, M. *Anal. Chim. Acta* **2004**, *517*, 229–236.

(22) Zhang, H. J.; Garland, M.; Zeng, Y. Z.; Wu, P. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 1295–1305.

(23) Guo, L. F.; Wiesmath, A.; Sprenger, P.; Garland, M. *Anal. Chem.* **2005**, *77*, 1655–1662.

(24) Garland, M.; Visser, E.; Terwiesch, P.; Rippin, D. W. T. *Anal. Chim. Acta* **1997**, *351*, 337–358.

(25) Golub, G. H.; Van Loan, C. F. *Matrix Computations*; Johns Hopkins University Press: Baltimore, 1996.

Table 1. Varying initial concentration of acetic anhydride and water in simple hydrolysis reactions

reaction	volume of acetic anhydride (mL)	volume of water (mL)	molarity of acetic anhydride
1	1.5	5	3.18
2	0.5	4	1.32
3	0.2	4	0.53
4	0.8	4	2.12

are returned in a normalized form. When all normed, observable, pure-component spectra have been reconstructed, the relative concentration profiles are recovered by projection onto the original data set. For detailed descriptions of BTEM algorithm, readers may refer to refs 10, 12, 13, and 17.

Experimental Methods

Chemicals. Acetic anhydride (analytical grade, 98%) was purchased from Sigma Aldrich Chemical. Further purification of acetic anhydride by distillation was performed before it was used for reaction. Water was filtered and deionized in an ultrapure water system (Aquamax, Younglin Instrument, Korea).

Dispersive Raman Microscope. A dispersive LabRAM HR Raman microscope from JobinYvon Horiba was used in all experiments presented in this paper. This microscope was equipped with a confocal microscope, a liquid nitrogen-cooled charge-coupled device (CCD) multichannel detector (256 pixels \times 1024 pixels), and a visible laser excitation source. The laser used in all experiments was a 20-mW visible 514.5 nm argon ion laser. A 50 \times long working distance objective lens was used, and the Raman shift range acquired was in the range of 100–1800 cm^{-1} with spectral resolution 1.7–2 cm^{-1} .

Experimental Setup. Two reaction-monitoring setups were used. In the first setup, the hydrolysis reaction was performed in a closed quartz cell situated directly under the Raman microscope (in situ reaction). The Raman spectra were acquired every 2 min during the 1 h reaction time, and altogether, 30 spectra were collected during this period. The acquisition time for one spectrum was 60 s. A total of four reactions with varying initial concentrations of acetic anhydride and water was performed. Table 1 shows the quantities of reagents and acetic anhydride molarities used in the reactions.

In the second experimental setup, an on-line reaction-monitoring system was used. The reactants were placed in a round-bottomed flask and continuously stirred with a magnetic stirrer to ensure good mixing. The flask was immersed in a silicon oil bath which kept the temperature constant at 23 $^{\circ}\text{C}$. The liquid phase was then pumped continuously by a positive displacement pump through a flow-through cell placed under the Raman microscope and then recycled back to the flask. Raman spectra were continuously acquired every minute. From time to time, perturbations were performed by adding some reactants to the system using a syringe (Table 2). The purpose of these perturbations was to break the concentration collinearities present among the observed components. The loadings of water and acetic anhydride were

Table 2. Experimental design for reagent perturbations and the corresponding spectral numbers

step	volume of water (mL)	volume of acetic anhydride (mL)	numbers of spectra
1	10	0	1–10
2	10	3	11–140
3	10	5	41–155
4	12	5	56–170
5	14	5	71–180
6	14	7	81–125

changed according to the experimental design found in Table 2. A total of five perturbations was performed, and a total of 125 Raman spectra was obtained for spectral analysis.

Data Analysis. All data analyses were performed in MATLAB 6.5 (The Mathworks Inc., Natick, MA). A personal computer with a 3-GHz Pentium IV processor and 522 MB of RAM was used for computations.

Results and Discussion

First Experimental Setup. As mentioned in the experimental method, two setups were used. The first setup consisted of a quartz cell under a long distance 50 \times microscope objective. Acetic anhydride and water were first mixed externally, before being injected into the static cell. Since the reaction was relatively slow, mixing was not a crucial issue. The objective was focused on a small spot (approximately a few micrometers in size) inside the static cell. Four separate reaction runs with varying concentrations were carried out. These multiple reactions were needed, in this simple reaction-monitoring setup, in order to break concentration collinearities among observed components. This is a standard prerequisite for successful multivariate data analysis, especially when accurate, pure-component spectra and hence accurate concentration information is needed. The Raman spectra from the four data sets were collected into one data matrix, $\mathbf{A}_{120 \times 952}$, with 120 reaction spectra (rows) and 952 channels of data (columns). Figure 1 presents one of the raw Raman data sets before any spectral preprocessing was employed. As can be seen, typical spikes due to cosmic rays are present. Also, the signal contributions in the spectral range of 100–300 cm^{-1} show pronounced fluorescence and no noticeable Raman signal (consistent with

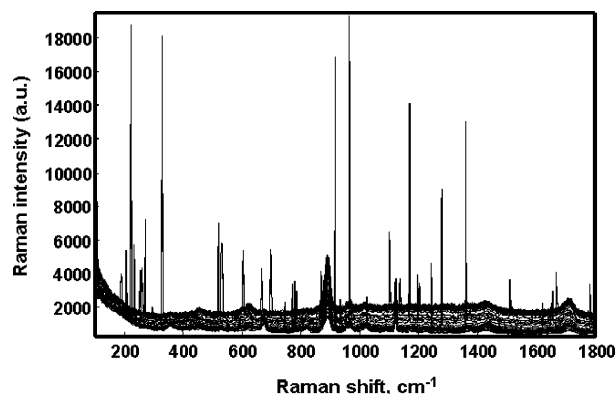


Figure 1. Raw Raman spectra from one data set taken in the static quartz cell before any spectral preprocessing.

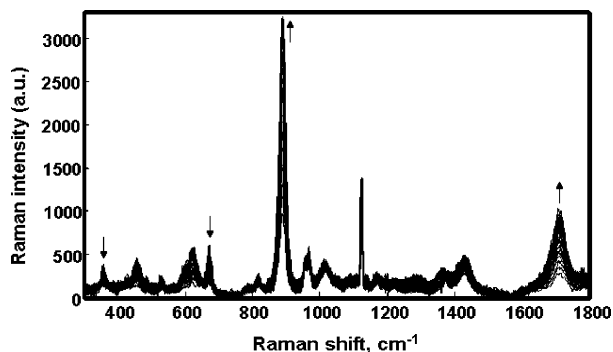


Figure 2. Raman spectra after spectral preprocessing taken from one of the data sets in first monitoring setup.

the present chemistry and hence lack of low wavenumber vibrations). Accordingly, this low wavenumber region was truncated. The remaining spectral region 300–1800 cm^{-1} has a relatively low signal-to-noise ratio, and a distorted baseline which increases in intensity with reaction time.

To remove most of the fluorescence and other nonchemical contributions, some spectral preprocessing was performed. Spikes were removed in the first step followed by adjacent 5-point averaging smoothing. The last step was baseline correction using third-order modified polynomial fitting.²⁶ The preprocessed Raman spectra from one of the data sets are shown in Figure 2.

After spectral preprocessing, the dynamic changes of the Raman bands intensities became much more obvious. Bands which clearly increased or decreased due to chemical reaction included 357, 675, 891, and 1710 cm^{-1} (Figure 2). However, the peak at 1124 cm^{-1} exhibits almost no changes in intensity. This peak is associated with the white light signal.

The preprocessed spectral data $\mathbf{A}_{120 \times 952}^{\text{Pre}}$ was then subjected to singular value decomposition yielding two orthonormal matrices $\mathbf{U}_{120 \times 120}$ and $\mathbf{V}_{952 \times 952}^{\text{T}}$ and the diagonal singular value matrix $\Sigma_{120 \times 952}$. The row vectors of \mathbf{V}^{T} matrix (normally called right singular vectors) are independent basis vectors that contain abstract information concerning the pure-component spectra of the observed components present in the system. These basis vectors are ordered according to their contribution to the total variance in the observations. Hence, the first few vectors are associated primarily with the chemically significant information in the system, and the remaining vectors are associated primarily with the random instrumental and experimental error. In addition, it should be noted that, although there are a total of 952 basis vectors of \mathbf{V}^{T} after decomposition, only 120 vectors are physically meaningful (either chemical signals or random error signals). This occurs since there are only 120 spectra in the original data set. Furthermore, not all of these 120 \mathbf{V}^{T} vectors were used in the BTEM spectral reconstructions. Figure 3 clearly shows that only the first six right singular vectors have significant and important signals related to this hydrolysis reaction and white light interference and that the last row of \mathbf{V}^{T} vectors (vectors 13–16) are primarily associated with the random noise. Therefore, 12 \mathbf{V}^{T} vectors were later used for all BTEM spectral reconstructions.

From the first four \mathbf{V}^{T} vectors, three significant spectral features were selected for the spectral reconstructions. The three targeted band regions were selected as 672–677, 888–894, and 1121–1126 cm^{-1} , and ultimately three pure-component spectra were resolved. These three pure-component spectral estimates corresponded to reactant acetic anhydride, the product acetic acid, and the white light signals. After obtaining the estimates of the pure-component spectra, it was possible to calculate the relative contributions of these observed components by solving eq 1 using multilinear regression. The resulting $\mathbf{C}_{k \times s}$ are relative contributions since the BTEM-resolved, pure-component spectra are normalized. Figure 4 shows the BTEM-resolved, pure-component spectral estimates and their relative contributions. As expected, there is a monotonic decrease of the acetic anhydride relative contributions and monotonic increase of the acetic acid relative contributions. In addition, it can also be seen that the initial acetic acid concentrations for the first, second, and fourth runs are not zero. This is due to some reaction which has occurred during external mixing before the liquid mixture was injected into the static cell for the Raman measurements. The white light contribution, however, remained relatively constant during each independent run. The variation of white light contributions between runs could be caused by slight variations in the shielding of the experimental setup from the laboratory white light.

The data corresponding to the four different reaction runs, with varying initial concentrations of acetic anhydride and water, can be clearly differentiated in the relative contributions profiles.

As a comparison for the spectral results, Table 3 shows the Raman peak assignments of prominent spectral features of acetic anhydride and acetic acid taken from the present BTEM estimates and the literature.^{27,28} The differences in the observed wavenumbers can be attributed to the different concentrations used and hence different contributions to hydrogen bonding. The very strong solvation effects for water, acetic anhydride, and acetic acid have been noted before in the literature. Hydrogen bonding with water leads to a weakening of the C–C bond and a strengthening of the C=O bond in acetic acid.²

Second Experimental Setup. In the second experimental setup, the hydrolysis reaction was performed with a simple continuous stirred tank (CSTR) reactor system with a recycle loop. Thus the reacting solution was pumped continuously through a 1-mm path length flow-through cell which was positioned under the Raman microscope objective. The aims for this flow-through Raman monitoring experiment were to create better mixing and homogeneity for the liquid phase and, at the same time, reduce the exposure time of the reacting liquid phase to the laser light. With the flow-through system, the fluid has only a short residence time in the cell and, hence, only a brief exposure time. This type of approach appears useful for potentially laser-light-sensitive reactions and might avoid or at least significantly reduce the problem of compound degradation.

(26) Lieber, C. A.; Mahadevan-Jansen, A. *Appl. Spectrosc.* **2003**, *57*, 1363–1367.

(27) Mirone, P.; Fortunato, B.; Canziani, P. *J. Mol. Struct.* **1970**, *5*, 283–295.

(28) Bertie, J. E.; Michaelian, K. H. *J. Chem. Phys.* **1982**, *77*, 5267–5271.

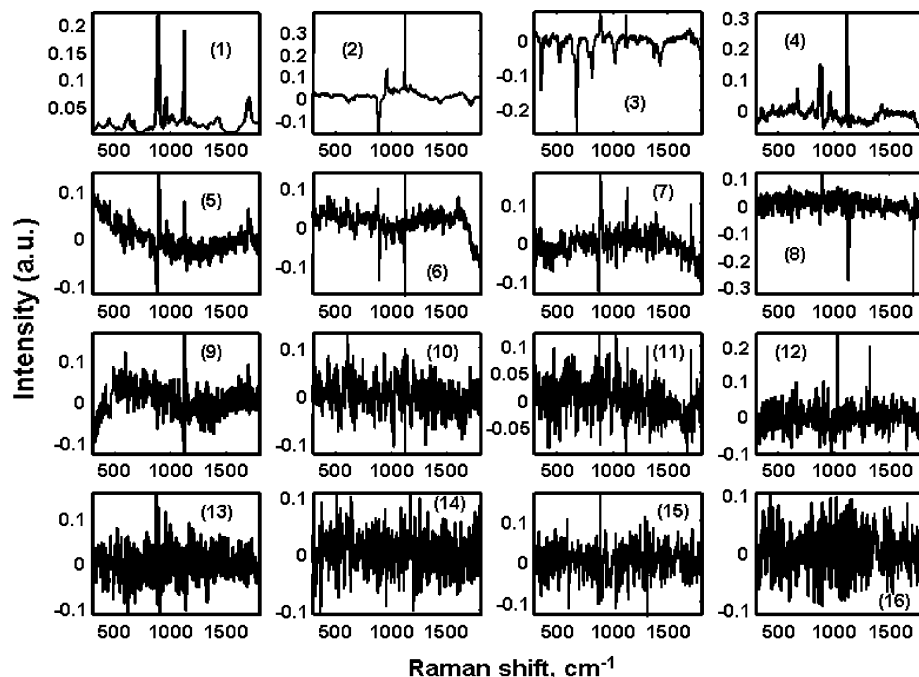


Figure 3. First 16 right singular vectors of the V^T matrix: (1) first vector, (2) second vector, (3) third vector, ..., (16) 16th vector.

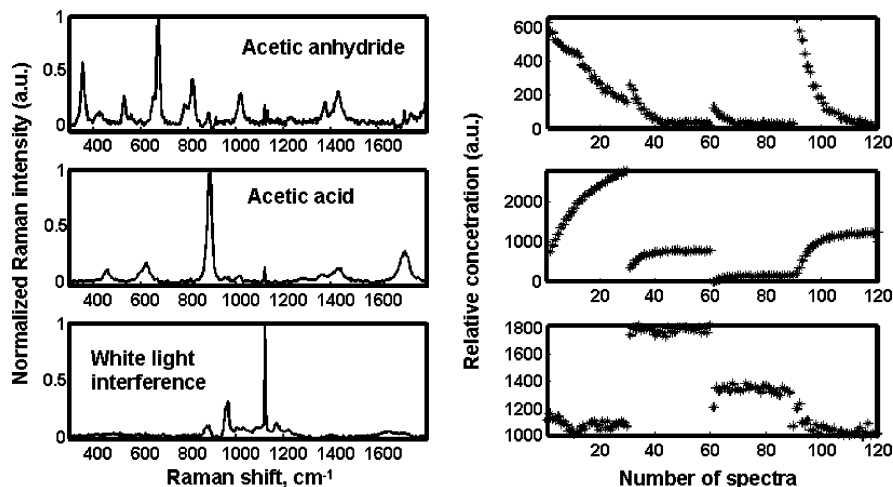


Figure 4. Reconstructed, pure-component spectral estimates of solutes and interference and their relative contributions in the static quartz cell experiments. A total of 120 reaction spectra was collected from the four reaction runs.

Table 3. Peak assignments of acetic anhydride and acetic acid Raman spectra

wavenumber, cm^{-1}		
literature ^{27,28}	BTEM estimates	assignment
Acetic Anhydride		
677	675	γ C=O + δ C=O
798	818	ν C–C
1009	1021	ρ CH ₃
1441	1432	γ CH ₃
Acetic Acid		
894	891	ν C–C
1432	1429	ν CH ₃
1690	1712	ν C=O

For the flow-through experiment, five reaction perturbations were performed (by adding either reactant or solvent into the reaction system), and a total 125 reaction spectra were collected. The raw Raman spectra from this flow-

through experiment before any spectral preprocessing are presented in Figure 5a. As can be seen, the signal-to-noise ratio in these spectra is quite low, and there are a lot of intense spikes due to cosmic rays. Spectral preprocessing consisting of spike removal, smoothing using adjacent 5-point averaging, and baseline correction using third-order modified polynomial fitting were again performed in order to improve the Raman reaction spectra. Figure 5b presents the preprocessed Raman spectra.

The 125 preprocessed Raman spectra were then subjected to singular value decomposition in order to reduce the data set and at the same time to generate independent orthonormal basis vectors, V^T , which were used to recover the observed pure-component spectra. Again, three significant spectral features were observed in the first few V^T vectors. These were targeted in the BTEM spectral reconstructions. The three targeted band regions were 672–677, 888–894, and

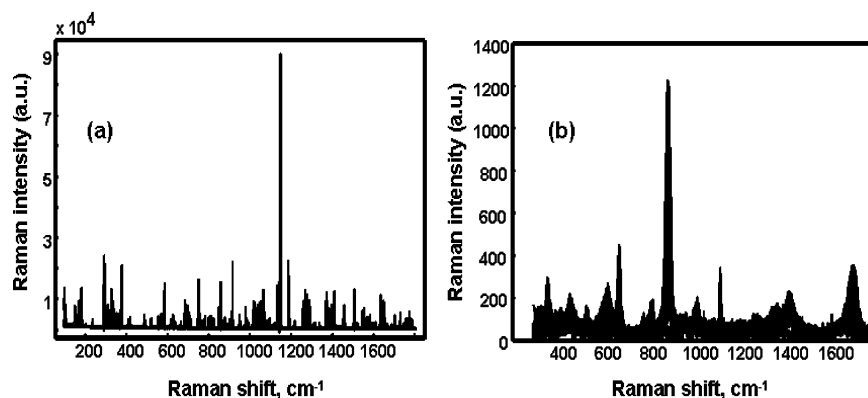


Figure 5. (a) Raw Raman spectra before any spectral preprocessing and (b) Raman spectra after spectral preprocessing for the flow-through experiments.

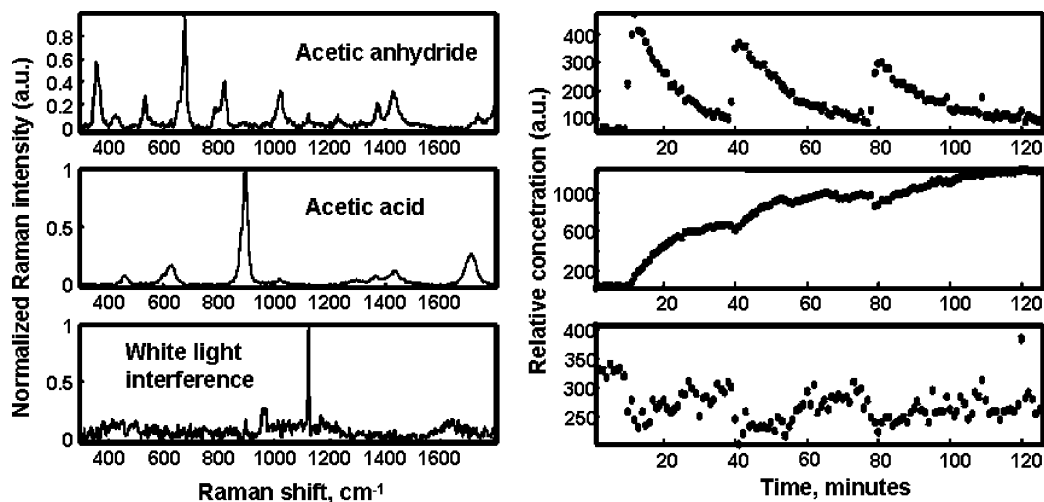


Figure 6. Reconstructed, pure-component spectral estimates and their relative contributions in flow-through experimental setup. A total of 125 reaction spectra was collected from the five-perturbation semibatch reaction.

1121–1126 cm^{-1} . Subsequently, BTEM resolved three pure-component spectra, which corresponded to the reactant acetic anhydride, the product acetic acid, and the white light signals. It should be noted that these spectral estimates are in good agreement with those obtained in the previous static experiments. The relative contributions of the observed components were then readily calculated. The pure-component spectral estimates via BTEM and their relative contributions in this flow-through experiment are shown in Figure 6.

Conclusions

The present work addressed the combined use of in situ Raman spectroscopy coupled with the novel multivariate data analysis algorithm, BTEM, for monitoring an aqueous-phase organic synthesis. The results clearly indicate the usefulness

of this approach, since good solute pure-component spectra as well as interference signals can be resolved. The high quality of the spectral estimates then allowed accurate evaluation of the individual signal contributions. At this point, and in the general case, subsequent reaction modeling can readily proceed. It appears that the combined use of Raman spectroscopy and BTEM represents a future tool in PAT for reaction monitoring.

Acknowledgment

We thank Dr. Selvasothi Selvaratnam and Teoh Siang Ping for purified acetic anhydride.

Received for review October 5, 2006.

OP0602066