Quantitative Analysis of Powder Mixtures by Raman Spectrometry: the influence of particle size and its correction

Zeng-Ping Chen*[a], Li-Mei Li[a], Jing-Wen Jin[a], Alison Nordon[b], David Littlejohn[b], Jing Yang[a], Juan Zhang[a], Ru-Qin Yu[a]

a. State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, China

b. WestCHEM, Department of Pure and Applied Chemistry and Centre for Process Analytics and Control Technology, University of Strathclyde, Glasgow, G1 1XL, Scotland, UK

* *Corresponding author*

Tel.: (+86) 731 88821916; Fax: (+86) 731 88821916;

E-Mail Address: zpchen2002@hotmail.com (Z.P. Chen)

Abstract: Particle size distribution and compactness have significant confounding effects on Raman signals of powder mixtures, which cannot be effectively modeled or corrected by traditional multivariate linear calibration methods such as partial least squares (PLS), and therefore greatly deteriorate the predictive abilities of Raman calibration models for powder mixtures. The ability to obtain directly quantitative information from Raman signals of powder mixtures with varying particle size distribution and compactness is, therefore, of considerable interest. In this study, an advanced quantitative Raman calibration model was developed to explicitly account for the confounding effects of particle size distribution and compactness on Raman signals of powder mixtures. Under the theoretical guidance of the proposed Raman calibration model, an advanced dual calibration strategy was adopted to separate the Raman contributions caused by the changes in mass fractions of the constituents in powder mixtures from those induced by the variations in the physical properties of samples, and hence achieve accurate quantitative determination for powder mixture samples. The proposed Raman calibration model was applied to the quantitative analysis of backscatter Raman measurements of a proof-of-concept model system of powder mixtures consisting of barium nitrate and potassium chromate. The average relative prediction error of prediction obtained by the proposed Raman calibration model was less than one-third of the corresponding value of the best performing PLS model for mass fractions of barium nitrate in powder mixtures with variations in particle size distribution as well as compactness.

Keywords: Quantitative Raman Spectroscopic Analysis, Particle Size Distribution, Compactness, Multiplicative Confounding Effects, Powder Mixture, Dual Calibration Strategy

Introduction

Powder blending is an important process in the manufacture of many pharmaceutical products [1]. Raman spectroscopy has been increasingly applied to the qualitative analysis of powder mixtures [2-6], because of its flexibility of sampling (solids can be analyzed with little or no sample preparation), and exceptionally high chemical specificity and the use of fibre optics for convenient and remote analysis, which facilitate the non-invasive in-line and real time analysis of particulate systems [7-17]. However, some issues remain unresolved regarding the quantitative in-line monitoring of particulate systems by Raman spectroscopy.

One of the issues is that the Raman intensities of analyte peaks depend on not only the analyte concentration, but also on the intensity of the excitation source, the instrument's optical configuration and the sample alignment. Therefore, to gain quantitative information requires the use of internal or external standards [18-20]. Band ratios between the overall Raman intensities and that of an individual spectral peak arising from internal or external standards are calculated and used in quantitative analysis. But the use of internal or external standards can be difficult to apply accurately in many *in-situ* process analysis applications. Moreover, for samples involving solids such as powder mixtures, quantitative Raman analysis becomes even more difficult, because the Raman measurements from such samples depend on the particle size and compactness of the mixtures, which hinders the use of an internal or external standard. The application of multivariate calibration methods such as principal component regression (PCR) and partial least squares (PLS) has some advantages over univariate band ratio calibration models in the quantitative analysis of Raman measurements [20, 21]. However, when analyzing powder mixtures using Raman spectroscopy, the variations in the physical

properties such as particle size and compactness of the mixtures have confounding effects on the total Raman intensities. Such confounding effects cannot be effectively modeled by standard multivariate calibration methods, and will significantly affect the predictive accuracy of multivariate calibration models.

Although it has long been known that physical properties of powder samples can influence the intensity of the Raman spectrum, and several studies [22-26] have been conducted on the relationship between particle size and Raman intensity, relatively little research focuses on quantitative Raman spectroscopic analysis of powder mixtures. Some of the present authors conducted a preliminary investigation on quantitative Raman spectroscopic analysis of suspension samples [27]. However, due to the facility limitations at that time, we were unable to explicitly investigate the effects of particle size distribution and sample compactness on Raman signals of powder mixtures in that work. The objectives of this study are to 1) explicitly investigate the effects of particle size and compactness on Raman signals of powder mixtures, 2) develop an advanced quantitative Raman calibration model for powder mixtures, and 3) eventually achieve accurate quantitative analysis of powder mixtures using Raman spectrometry.

Theory

Raman intensities of powder mixtures

The intensity of Raman bands depends on a complex expression involving the polarisability tensor of a molecule [28]. For analytical purposes, the following less rigorous linear model

88 analogous to the Beer-Lambert law can be used.

$$I(v) = n \cdot r(v) \cdot I_o \quad (1)$$

89 Where $I(v)$ is the Raman intensity at Raman shift v, I_o is the intensity of the excitation
90 radiation, n is the number of molecules of the analyte illuminated by the source and viewed
91 by the spectrometer, and $r(v)$ is a composite term that represents the overall spectrometer
92 response, and the self absorption and molecular scattering properties of the analyte at Raman
93 shift v. For K powder samples comprising J constituents with amounts above their Raman
94 limits of detection, their overall Raman intensities can be expressed as the linear combination
95 of the contributions of all J constituents as well as other possible interference(s) such as
96 fluorescence.

$$I_k(v) = \sum_{j=1}^{J} [n_{k,j} \cdot r_j(v) \cdot I_{o,k}] + n_{k,interf} \cdot r_{interf}(v) \cdot I_{o,k}; \qquad k = 1, 2, ..., K \quad (2)$$

97 Where $n_{k,j}$ and $n_{k,interf}$ are the number of molecules of the j-th constituent and the
98 interference(s) in the k-th powder sample illuminated by the source and viewed by the
99 spectrometer, respectively; $r_{interf}(v)$ represents the molecular scattering/fluorescence
100 properties of the interference(s) at Raman shift v.

101 Suppose m_k and V_k are the overall mass and volume of the k-th powder sample,
102 respectively. $V_{spec,k}$ denotes the volume of the k-th powder sample illuminated by the source
103 and viewed by the spectrometer. $w_{k,j}$ ($\sum_{j=1}^{J} w_{k,j} = 1$) signifies the mass fraction of the j-th
104 constituent in the k-th sample. M_j is the molecular weight of the j-th constituent. The
105 multiplicative parameter, p_k, is introduced to account for the effects of the particle size

distribution and compactness of the *k*-th sample on the Raman intensities [24, 27]. Equation 2 then becomes:

$$I_k(v) = \sum_{j=1}^{J}[p_k \cdot \frac{m_k \cdot w_{k,j} \cdot V_{spec,k}}{M_j \cdot V_k} \cdot r_j(v) \cdot I_{o,k}] + n_{k,interf} \cdot r_{interf}(v) \cdot I_{o,k} \quad (3)$$

Define $q_k = p_k \cdot m_k \cdot V_{spec,k} \cdot I_{o,k} / V_k$ and $r_j^*(v) = r_j(v) / M_j$. Equation 3 can be simplified as follows.

$$I_k(v) = \sum_{j=1}^{J}[q_k \cdot w_{k,j} \cdot r_j^*(v)] + n_{k,interf} \cdot r_{interf}(v) \cdot I_{o,k} \quad (4)$$

In equation 4, q_k is a very important model parameter. It accounts for the variations in Raman intensities caused by the changes in variables other than the mass fractions of the constituents in the powder mixtures, such as the intensity of the excitation source, the sample's particle size distribution, sample compactness, the overall mass and volume of the powder sample as well as the volume illuminated by the source and viewed by the spectrometer.

Suppose the *j*-th constituent is the target component in the powder mixtures, and the Raman signals of *K* calibration samples have been measured over Raman shift range of v_1~v_m. As $\sum_{j=1}^{J} w_{k,j} = 1$, equation 4 can be rewritten as:

$$\begin{aligned} &\mathbf{x}_k = q_k \cdot w_{k,1} \cdot \Delta\mathbf{r}_1^* + q_k \cdot \mathbf{r}_2^* + \sum_{j=3}^{J}[q_k \cdot w_{k,j} \cdot \Delta\mathbf{r}_j^*] + n_{k,interf} \cdot \mathbf{r}_{interf} \cdot I_{o,k} \\ &\text{Where, } \mathbf{x}_k = [I_k(v_1), I_k(v_2), \cdots, I_k(v_m)]; \quad \mathbf{r}_j^* = [r_j^*(v_1), r_j^*(v_2), \cdots, r_j^*(v_m)], \ \mathrm{j} = 1, 2, \cdots, \mathrm{J} \\ &\Delta\mathbf{r}_j^* = \mathbf{r}_j^* - \mathbf{r}_2^*; \quad \mathbf{r}_{interf} = [r_{interf}(v_1), r_{interf}(v_2), \cdots, r_{interf}(v_m)] \end{aligned} \quad (5)$$

Assuming $\Delta\mathbf{r}_j^*$, $\mathbf{r}_2^*$, and $\mathbf{r}_{interf}$ are linearly independent of each other, it can be seen that a straightforward multivariate linear calibration model can be built only between $\mathbf{x}_k$ and $q_k \cdot w_{k,j}$ (or q_k). It is obvious that the multiplicative parameter, q_k, may be different for each

of the powder samples. Hence the relationship between Raman spectrum $\mathbf{x}_k$ and the mass fraction of the j-th constituent ($w_{k,j}$) is actually nonlinear; and the multiplicative parameter, q_k, has confounding effects on the estimation of $w_{k,j}$. In order to extract the quantitative information (mass fraction) of any constituent in powder samples from their Raman measurements, it is therefore imperative to estimate the multiplicative parameter, q_k, for each powder sample.

Dual Calibration Strategy (DCS) [27, 29-30]

For K training samples in which the mass fractions of the target constituent (say, the j-th constituent) are known, the multiplicative parameters, q_k (k = 1, 2, …, K), can be estimated by the modified Optical Path-Length Estimation and Correction method ($OPLEC_m$) [30] (the Matlab script for $OPLEC_m$ is provided in supporting information). After the estimation of q_k (k = 1, 2, …, K), the following two calibration models can be built by multivariate linear calibration methods such as PLS.

$$\begin{aligned} &diag(\mathbf{w}_j)\mathbf{q} = a_1\mathbf{1} + \mathbf{X}_{cal}\boldsymbol{\beta}_1; \quad \mathbf{q} = a_2\mathbf{1} + \mathbf{X}_{cal}\boldsymbol{\beta}_2 \\ &\mathbf{X}_{cal} = [\mathbf{x}_1;\ \mathbf{x}_2; ...; \mathbf{x}_K]; \quad \mathbf{w}_j = [w_{1,j};\ w_{2,j}; ...; w_{K,j}]; \quad \mathbf{q} = [q_1;\ q_2; ...; q_K] \end{aligned} \tag{6}$$

Where $diag(\mathbf{w}_j)$ denotes the diagonal matrix in which the corresponding diagonal elements are the elements of $\mathbf{w}_j$; $\mathbf{1}$ is a column vector with its elements equal to unity. After the estimation of model parameters a_1, a_2, $\boldsymbol{\beta}_1$, and $\boldsymbol{\beta}_2$ by multivariate calibration methods such as PLS, these two calibration models could then be used to predict the mass fraction of the target constituent in any test powder sample ($w_{test,j}$) from its Raman spectrum $\mathbf{x}_{test}$.

$$q_{test} \cdot w_{test,j} = a_1 + \mathbf{x}_{test}\boldsymbol{\beta}_1; \qquad q_{test} = a_2 + \mathbf{x}_{test}\boldsymbol{\beta}_2; \qquad w_{test,j} = \frac{a_1 + \mathbf{x}_{test}\boldsymbol{\beta}_1}{a_2 + \mathbf{x}_{test}\boldsymbol{\beta}_2} \tag{7}$$

The mass fraction of other constituents in the test sample can be obtained in a similar way.

Experimental

Materials

All chemicals were analytical grade, and were used as received without any further purification. Potassium chromate was obtained from Tianjin Windship Chemistry Technological Co., Ltd (Tianjin, China). Barium nitrate was purchased from Tianjin Kermel Chemical Reagent Co., Ltd (Tianjin, China).

Equipment

Raman spectra were acquired at room temperature on a LABRAM-0101 Laser Confocal Raman Spectrometer equipped with a 1024×256 pixels CCD detector. The microscope attachment was based on an Olympus BX41 system with a 10× objective. Radiation of 632.81 nm from a 17 mW He-Ne laser was used for excitation. The widths of the entrance slit and confocal pinhole were set to 100 μm and 1000 μm, respectively. Raman spectrum between 200 and 2000 cm^{-1} was collected with a 5 s exposure time and 3 accumulations for each spectrum.

Raman measurements of powder mixtures

The solids of both barium nitrate and potassium chromate were ground and sorted into

different particle sizes using standard sieves. The standard sieves were of mesh sizes 40, 60, 80, 100, 120, 140, 160 and 200 wires per inch. The hole sizes corresponding to the mesh sizes are 425, 250, 180, 150, 125, 109, 96 and 75 μm, respectively. A total of 72 powder mixtures of potassium chromate and barium nitrate powder with different weight ratios (1:0, 0.90:0.10, 0.75:0.25, 0.60:0.40, 0.50:0.50, 0.40:0.60, 0.25:0.75, 0.10:0.90 and 0:1) and different particle sizes (425, 250, 180, 150, 125, 109, 96 and 75 μm) were prepared by mixing appropriate amounts of the two constituents thoroughly (Table 1). For each of 72 powder mixtures, a sample was randomly taken and loosely packed into a cylindrical sample cup with a diameter of 6.9 mm and a height of 10.7 mm. The laser beam was focused at a point inside the sample so as to ensure the illumination of the whole upper surface of the sample by the laser beam, and then the Raman spectrum was acquired. Following this, each sample was packed more firmly, and a further Raman spectrum was recorded resulting in a total of 144 spectra. Seventy eight spectra (two outliers were removed) from the five mixtures with the ratios of potassium chromate to barium nitrate equal to 1:0, 0.75:0.25, 0.5:0.5, 0.25:0.75 and 0:1 formed the calibration data set. The test set comprised the remaining 64 spectra from the other four mixtures. Distinctive Raman peaks of potassium chromate (at around 351, 386.5, 396.8，853.4, 868.4, 877.8 and 906.8 cm^{-1}) and barium nitrate (at about 1047.5 cm^{-1}) can be readily observed between 292.8 and 1136.6 cm^{-1} (supporting information, Figure S-1). Therefore, Raman signals in this region were selected for the subsequent data analysis.

Table 1: Mass ratios and particle sizes of potassium chromate and barium nitrate in powder mixtures.

Sample No.	$K_2CrO_4/Ba(NO_3)_2$ （mass ratio）	Particle Size (μm)
1 – 8	1:0	425, 250, 180, 150, 125, 109, 96, 75
9 – 16	0.90:0.10	425, 250, 180, 150, 125, 109, 96, 75
17 – 24	0.75:0.25	425, 250, 180, 150, 125, 109, 96, 75
25 – 32	0.60:0.40	425, 250, 180, 150, 125, 109, 96, 75
33 – 40	0.50:0.50	425, 250, 180, 150, 125, 109, 96, 75
41 – 48	0.40:0.60	425, 250, 180, 150, 125, 109, 96, 75
49 – 56	0.25:0.75	425, 250, 180, 150, 125, 109, 96, 75
57 – 64	0.10:0.90	425, 250, 180, 150, 125, 109, 96, 75
65 – 72	0:1	425, 250, 180, 150, 125, 109, 96, 75

Data analysis

Due to the influence of particle size and compactness on Raman intensities, it is unlikely that univariate analysis will give accurate predictions of the mass fractions of barium nitrate in powder mixtures. Therefore, PLS and the dual calibration strategy (DCS) were adopted for the data analysis and their performance in terms of providing accurate predictions for the mass fractions of barium nitrate in powder mixtures were compared. The effectiveness of multiplicative signal correction (MSC) [31], standard normal variate (SNV) [32] and extended inverted signal correction (EISC) [33] in correcting the confounding effects of physical properties of powder samples on the Raman measurements and improving the predictive abilities of PLS calibration models were also investigated. For the convenience of presentation, PLS calibration models built on the mean-centred raw and preprocessed Raman spectra by MSC, EISC and SNV are denoted by PLS_raw, PLS_MSC, PLS_EISC and PLS_SNV, respectively. No pre-processing methods other than mean-centring were used when building DCS calibration models. The optimal calibration models were selected through a cross-validation procedure. During cross-validation, the Raman spectra of the calibration samples with the same mass ratio of potassium chromate to barium nitrate were left out in turn and the root mean square error of prediction from cross validation (RMSEPcv) values were calculated. The calibration models with the minimal RMSEPcv values were taken as the optimal models, and were then used to predict the mass fractions of barium nitrate in the test samples. The data analysis was performed on a Pentium class computer using Matlab version 6.5 (Mathworks, Inc). All the programmes including DSC, PLS, MSC, SNV, and EISC were written in house.

Results and discussion

The sensitivity of Raman intensities to the mass fraction of barium nitrate

Figure 1a shows the Raman spectra of powder mixtures samples with the same particle size (250 μm) and similar compactness but different mass ratios of potassium chromate to barium nitrate. The Raman peaks are relatively sensitive to the changes in the composition of the powder mixtures. The Raman peak intensity at 1047.5 cm^{-1} generally increases with mass fraction of barium nitrate in the powder samples. However, the relationship between Raman peak intensity and mass fraction of barium nitrate deviates from a perfect linear model even when samples have a similar particle size and degree of compactness (Figure 1b). Especially there is a discontinuity which might be caused by the variation in excitation intensity or packing density. This demonstrates the necessity to introduce the multiplicative parameter, q_k, in eq.4 to account for the variations in Raman intensities caused by the changes in variables other than the mass fractions of the constituents in the powder mixtures.

236

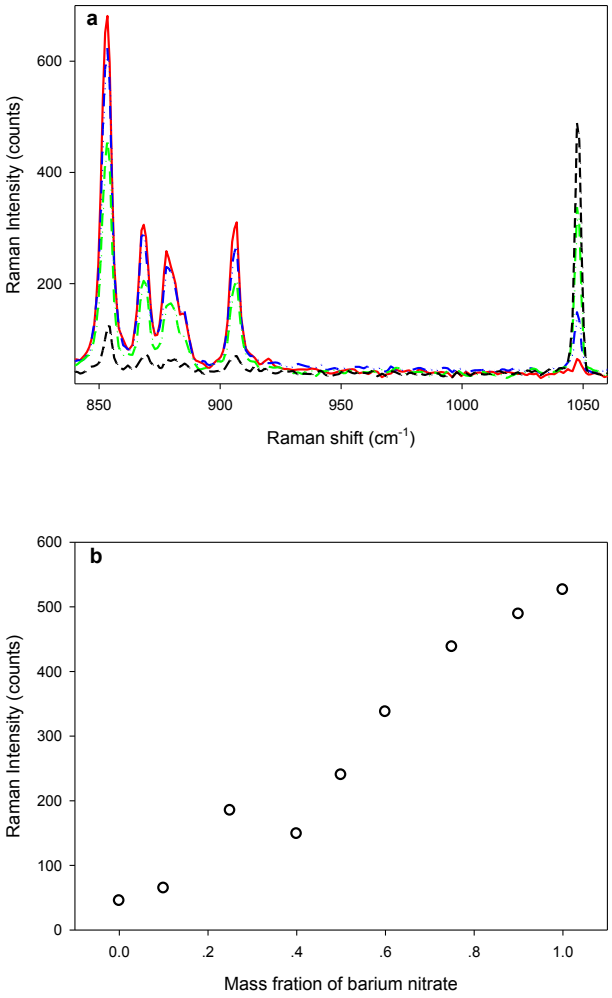


Figure 1: a) Raman spectra of loosely packed powder mixture samples (particle size: 250 μm) with different mass ratios of potassium chromate to barium nitrate (red solid line: 0.9:0.1; blue dash-dot-dot line: 0.60:0.40; green dash-dot line: 0.40:0.60; black dash line: 0.10:0.90); b) Raman peak intensity at 1047.5 cm^{-1} vs mass fraction of barium nitrate in loosely packed powder mixture samples (particle size: 250 μm).

The effects of particle size and compactness on Raman intensities

In addition to the mass ratio of potassium chromate to barium nitrate, the particle size and compactness of the powder mixture samples also have a significant influence on the Raman peak intensities. As shown in Figure 2a, a firmly packed sample has significantly more intense Raman peaks than those of a loosely packed sample with the same mass ratio and particle size. It has long been known that particle size differences make significant contributions to the spectral variations in Raman measurements of powders [22]. Our experimental results also show that variations in particle size of powder samples have significant effects on the Raman spectra (Figure 2b). For two samples with the same particle size (109 μm) but different mass ratios of potassium chromate to barium nitrate (e.g. 0.25:0.75 and 0.10:0.90), the difference between the peak intensities at 1047.5 cm^{-1} is 59.04. While for two samples with the same mass ratio of potassium chromate to barium nitrate (0.10:0.90) but different particle sizes (e.g. 109 and 75 μm), the difference between the corresponding peak intensities is 113.63, which is about 1.9 times that caused by a change in the mass ratio of potassium chromate to barium nitrate from 0.25:0.75 to 0.90:0.10. Moreover, variation in the particle size of a sample has the same effect on all Raman peaks in the spectrum. This makes it difficult to discriminate Raman intensity contributions caused by changes in a sample's particle size from those due to a variation in mass fractions of the chemical constituents using traditional univariate/multivariate calibration methods. If not properly modelled, this difference would significantly degrade the accuracy and reliability of calibration models built on Raman measurements contaminated by such confounding effects. The multiplicative parameter, q_k, in eq.4 accounts for the effects of particle size and

compactness on the Raman intensities and so their effects can be separated from those of the mass fractions of the chemical constituents in powder samples by the unique dual calibration strategy.

281

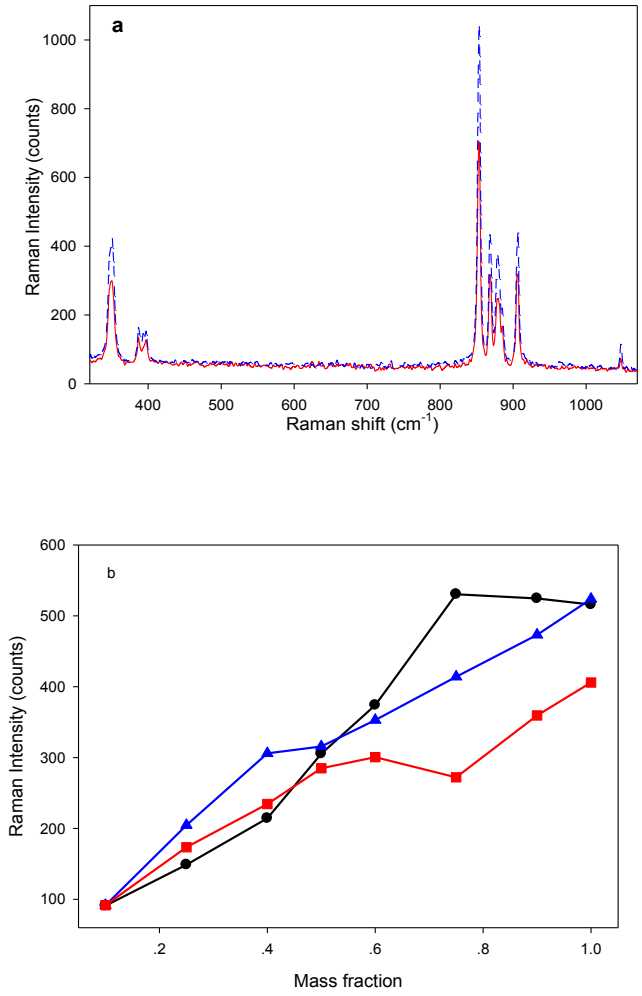


Figure 2: a) Raman spectra of a binary powder mixture sample (potassium chromate:barium nitrate: 0.90:0.10, particle size: 425 μm) with different compactness (blue dash line: firmly packed; red solid line: loosely packed); b) peak intensity at 1047.5 cm^{-1} vs mass fraction of barium nitrate with different particle sizes (black circle: 180 μm ; blue triangle: 109 μm ; red square: 75 μm).

Correction of the multiplicative effects of particle size and compactness on Raman intensities

As shown in the preceding section, the presence of significant multiplicative confounding effects (arising from differences in particle size and compactness) caused deviations in the linear relationship between the Raman intensities and the mass fraction of solid powder samples. With a view to mitigate the influence of the multiplicative confounding effects present in the Raman spectral data, the dual calibration strategy (DCS) was employed to correct such confounding effects. For the purpose of comparison, PLS models with and without the use of pre-processing methods MSC, EISC and SNV were also applied to the same Raman spectral data. DCS involves the estimation of the multiplicative parameter, q_k, for each calibration sample by $OPLEC_m$ [30]. The implementation of $OPLEC_m$ requires the determination of the number of spectral variation sources including chemical components and possible interference(s). For the powder mixture samples studied in this paper, the number of spectral variation sources is two, i.e. potassium chromate and barium nitrate. The results of $OPLEC_m$ are shown in Figure 3. It is evident that different calibration samples generally have different multiplicative parameter values (q_k) and the multiplicative parameter, q_k, of the calibration samples varies in the range of 1 – 2.23. These results demonstrate that the presence of significant multiplicative confounding effects in the Raman spectral data and the introduction of the multiplicative parameter, q_k, in eq.4 is theoretically sound and also practically relevant. Otherwise, the multiplicative parameter values (q_k) calculated by $OPLEC_m$ for the calibration samples would vary within a narrow range, and would also be quite close to 1.

303

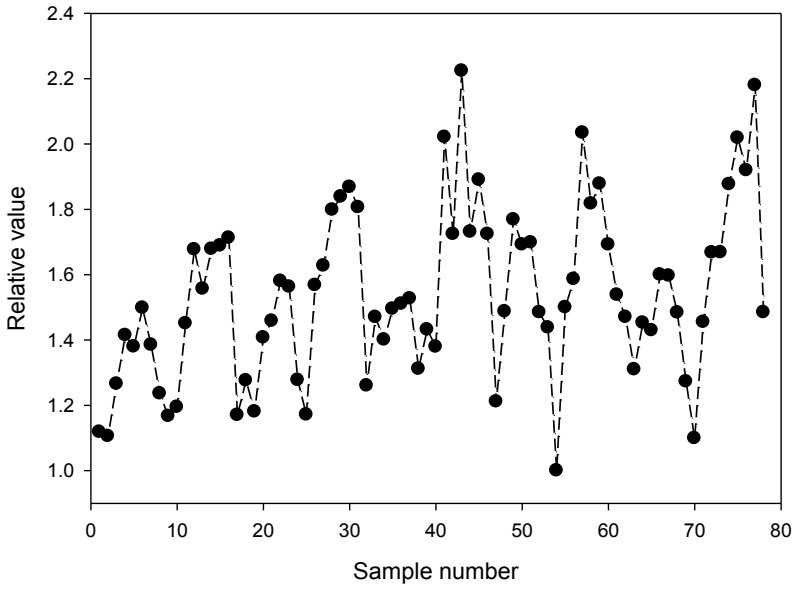


Figure 3: the multiplicative parameter q_k for the calibration samples estimated by $OPLEC_m$.

304

305

306

307

308

309

310

311

312

313

314

After the calculation of the multiplicative parameters, q_k, for each calibration sample by $OPLEC_m$, DCS models with different underlying components were built. Values of the root mean square error of prediction from cross validation (RMSEPcv) obtained by DCS and the various PLS calibration models (i.e. PLS_raw, PLS_MSC, PLS_EISC and PLS_SNV) with different number of latent variables are given in supporting information, Figure S-2. Both the PLS_raw and PLS_MSC models attained minimum RMSEPcv values of 0.08 and 0.12, respectively, when two latent variables were used. For PLS_EISC and PLS_SNV, only one latent variable was suggested by cross validation; however, the RMSEPcv values of 0.26 and 0.23, respectively, were significantly larger than that for PLS_raw, which to some extent indicates the inappropriateness of applying EISC and SNV to this particular Raman spectral data set. In contrast with the above PLS calibration models, a DCS model with three latent variables had a minimum RMSEPcv value of 0.03, which is less than half that of the corresponding value obtained with the PLS_raw model.

For a more convincing comparison, the performance of the optimal DCS and various PLS calibration models for the independent test samples was investigated. As shown in Figure 4 and Figure 5, the RMSEP value of the optimal PLS_raw model obtained for the independent test samples was 0.08 (equivalent to a mean relative prediction error of 30.8%), which clearly demonstrates the presence of detrimental multiplicative confounding effects caused by variations in the particle size and compactness of powder samples. The application of MSC, EISC and SNV resulted in a deterioration of the predictive ability of the PLS calibration models. This confirms that the pre-processing methods MSC, EISC and SNV cannot effectively correct the multiplicative confounding effects of particle size and

337 compactness on Raman intensities. In contrast, the optimal DCS model with 3 latent variables
338 achieved a RMSEP value of 0.04 for the independent test samples, which is equivalent to a
339 mean relative prediction error of 9.6%, less than one third of the corresponding value for the
340 optimal PLS_raw model. Even more interestingly, though the construction of the DCS model
341 requires no extra information or data compared to the PLS models, it consistently
342 outperformed the various PLS models built on the raw and pre-processed Raman spectra, no
343 matter how many latent variables were used (Figure S-3 in supporting information). The
344 significant reduction in the RMSEP value achieved with the optimal DCS model results
345 solely from the introduction of the multiplicative parameter, q_k, in eq. 4 to account for the
346 variations in Raman intensities caused by the changes in variables other than the mass
347 fractions of the chemical constituents in powder mixtures, in this case particle size and
348 compactness.
349
350
351
352
353
354
355
356
357
358

359

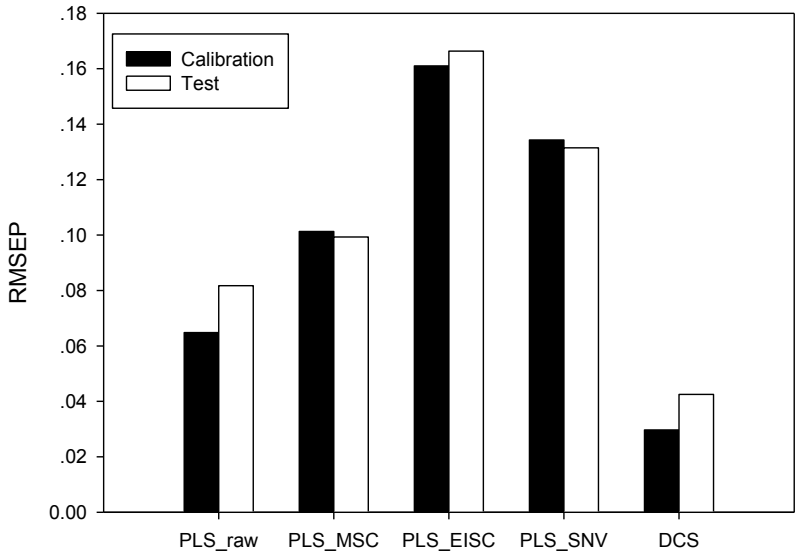


Figure 4: the RMSEP values for both the calibration and independent test samples obtained by different calibration methods.

360

361

362

363

364

365

366

367

368

369

370

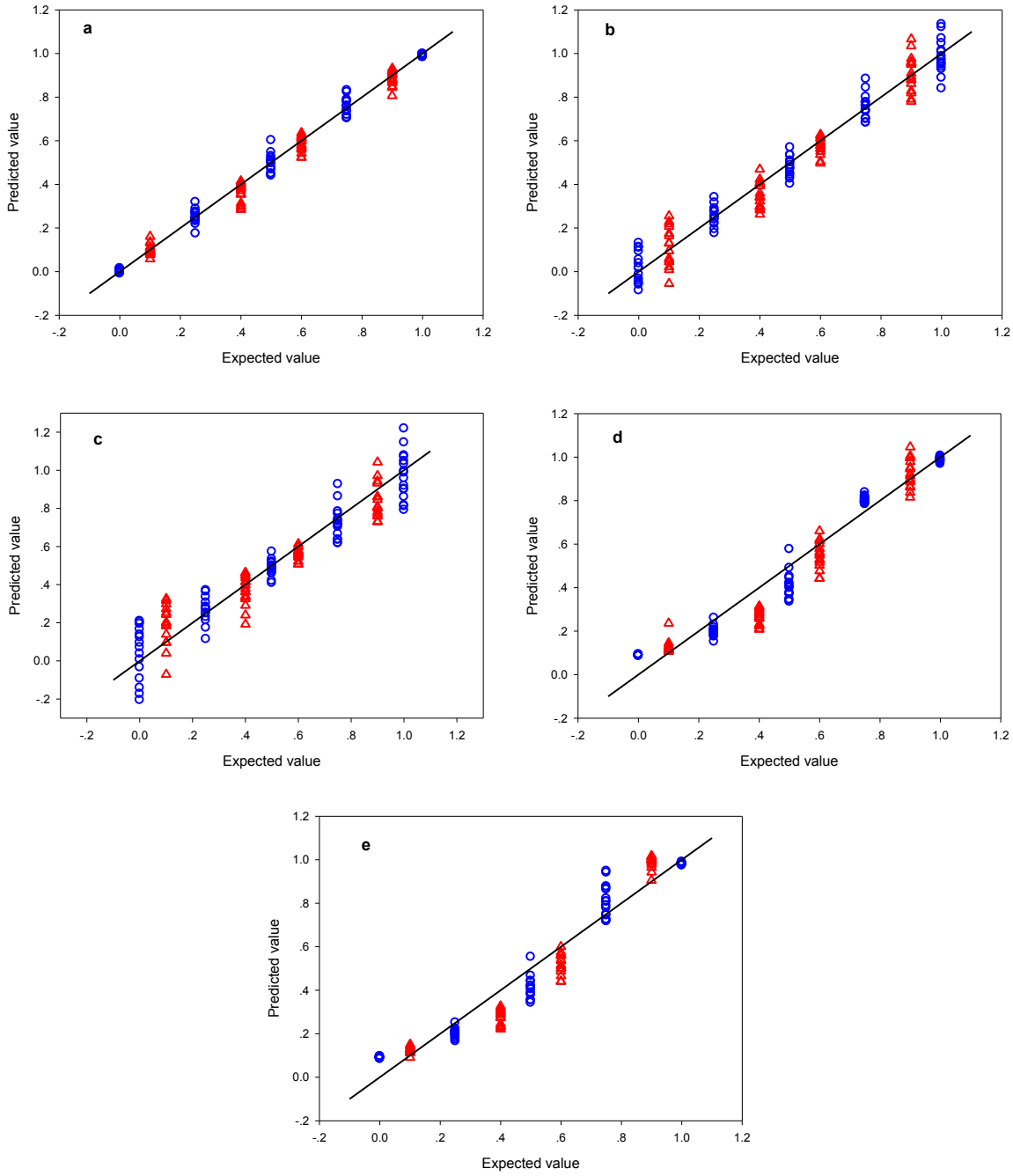


Figure 5: the mass fractions of $Ba(NO_3)_2$ in the calibration (blue circle) and independent test (red triangle) samples predicted by various calibration models (a: DCS; b: PLS_raw; c: PLS_MSC; d: PLS_SNV; e: PLS_EISC)

371

Conclusions

The Raman intensities of powder mixture samples depend on not only the mass fractions of the chemical constituents but also the physical properties of samples such as particle size distribution and compactness. The experimental results on a model system of powder mixtures consisting of barium nitrate and potassium chromate showed that the effects of particle size distribution and compactness on Raman measurements are multiplicative, which cannot be effectively modelled by multivariate linear calibration methods such as PLS. Pre-processing the Raman measurements with multiplicative confounding effects of particle size and compactness by MSC, SNV or EISC could not improve but rather deteriorated the predictive performance of Raman calibration models. In this work, we introduced a multiplicative parameter in the quantitative Raman calibration model to explicitly account for the confounding effects of particle size and compactness on Raman signals of powder mixtures, and then eliminated the confounding effects through a unique dual calibration strategy. The average relative prediction error of predictions obtained by the dual calibration strategy for the independent test samples was less than one-third of the corresponding value of the optimal PLS calibration models built using the raw Raman spectra and considerably better than the results of PLS models based on spectra pre-processed by application of MSC, EISC or SNV. These results demonstrate that the dual calibration strategy can effectively mitigate the confounding effects of samples' physical properties and so improve the accuracy of quantitative analysis of powders using Raman spectrometry. Hence, the dual calibration strategy will be of major benefit for quantitative measurement of particulate samples such as powder blends and pharmaceutical dosage forms.

Acknowledgements

The authors acknowledge the financial support of the National Natural Science Foundation of China (grant no. 21075034 and grant no. 21035001) and the National Instrumentation Program of China (grant no. 2011YQ0301240102). AN acknowledges the award of a University Research Fellowship by the Royal Society, UK.

Supporting Information Available

MATLAB code for the modified OPLEC method, Raman spectra of K_2CrO_4 and $Ba(NO_3)_2$; The RMSEP values from cross validation obtained by DCS and various PLS calibration models with different number of latent variables built on the raw and pre-processed Raman spectra by MSC, SNV, and EISC; The RMSEP values for the test powder mixtures obtained by DCS and various PLS calibration models with different number of latent variables built on the raw and pre-processed Raman spectra by MSC, SNV, and EISC. This material is available free of charge via the Internet at http://pubs.acs.org.

417 **References:**

418 [1]. Berman, J. *PDA J. Pharm. Sci. Tech.* **2001**, *55*, 209-222

419 [2]. De Beer, T.R.M.; Bodson, C.; Dejaegher, B.; Walczak, B.; Vercruysse, P.; Burggraeve,

420 A.; Lemos, A.; Delattre, L.; Varder Herden, Y.; Remon, J.P.; Vervaet, C.; Baeyens,

421 W.R.G. *J. Pharm. Biomed. Anal.* **2008**, *48*, 772-779

422 [3]. Campbell Roberts, S.N.; Williams, A.C.; Grimsey, I.M.; Booth, S.W. *J. Pharm.*

423 *Biomed. Anal.* **2002**, *28*, 1135-1147

424 [4]. Al-Zoubi, N.; Koundourellis, J.E.; Malamataris, S. *J. Pharm. Biomed. Anal.* **2002**, *29*,

425 459-467

426 [5]. Pratiwi, D.; Fawcett, J.P.; Gordon, K.C.; Rades, T. *Eur. J. Pharm. Biopharm.* **2002**, *54*,

427 337-341

428 [6]. Strachan, C.J.; Pratiwi, D.; Gordon, K.C.; Rades, T. *J. Raman Spectrosc.* **2004**, *35*,

429 347-352

430 [7]. De Beer, T.R.M.; Baeyens, W.R.G.; Ouyang, J.; Vervaet, C.; Remon, J. P. *Analyst*,

431 **2006**, *131*, 1137-1144

432 [8]. Kogermann, K.; Aaltonen, J.; Strachan, C.J.; Pöllänen, K.; Heinämäki, J.; Yliruusi, J.;

433 Rantanen, J. *J. Pharm. Sci.* **2008**, *97*, 4983-4999

434 [9]. Févotte, G. *Chem. Eng. Res. Des.* **2007**, *85(A7)*, 906-920

435 [10]. Caillet, A.; Puel, F.; Fevotte, G. *Int. J. Pharm.* **2006**, *307*, 201-208

436 [11]. Hu, Y.R.; Liang, J.K.; Myerson, A.S.; Taylor, L.S. *Ind. Eng. Chem. Res.* **2005**, *44*,

437 1233-1240

438 [12]. Ono, T.; Kramer, H.J.M.; ter Horst, J.H.; Jansens, P.J. *Cryst. Growth Des.* **2004**, *4*,

1161-1167

[13]. O'Sullivan, B.; Barrett, P.; Hsiao, G.; Carr, A.; Glennon, B. *Org. Proc. Res. Dev.* **2003**, *7*, 977-982

[14]. Schöll, J.; Bonalumi, D.; Vicum, L.; Mazzotti, M.; Muller, M. *Cryst. Growth Des.* **2006**, *6*, 881-891

[15]. Starbuck, C.; Spartalis, A.; Wai, L; Wang, J.; Fernandez, P.; Lindemann, C.M.; Zhou, G.X.; Ge, Z. *Cryst. Growth Des.* **2002**, *2*, 515-522

[16]. Wang, F.; Wachter, J.A.; Antosz, F.J.; Berglund, K.A. *Org. Proc. Res. Dev.* **2000**, *4*, 391-395

[17]. Agarwal, P.; Berglund, K.A. *Cryst. Growth Des.* **2003**, *3*, 941-946

[18]. Scheweinsberg, D.P.; West Y.D. *Spectrochimica Acta Part A*, **1997**, *53*, 25-34

[19]. Aarnoutse, P.J.; Westerhuis, J.A. *Anal. Chem.* **2005**, *77*, 1228-1236

[20]. Pelletier, M.J. *Appl. Spectrosc.* **2003**, *57*, 20A-42A

[21]. Szostak, R.; Mazurek, S. *Journal of Molecular Structure*, **2004**, *704*, 235-245

[22]. Wang, H.; Mann, C. K.; Vickers, T. J. *Appl. Spectrosc.* **2002**, 56, 1538-1544

[23]. Hu, Y.; Wikström, H.; Byrn, S. R.; Taylor, L. S. *Appl. Spectrosc.* **2006**, 60, 977-984

[24]. Pellow-Jarman, M.V.; Hendra, P.J.; Lehnert, R.J. *Vibrational Spectroscopy,* **1996**, *12*, 257-261

[25]. Reis, M. M.; Araujo, P. H. H.; Sayer, C.; Giudici, R. *Polymer,* **2003**, *44*, 6123-6128

[26]. Hamilton, P.; Littlejohn, D.; Nordon, A.; Sefcik, J.; Slavin, P.; Dallin, P.; Andrews, J. *Analyst,* **2011**, *136*, 2168-2174

[27]. Chen, Z.P.; Fevotte, G.; Caillet, A.; Littlejohn, D.; Morris, J. *Anal. Chem.* **2008**, *80*, 6658-6665

[28]. Anderson, A. *The Raman effect, volume 2: Applications*, Mercel Dekker, Inc: New York, **1973**

[29]. Chen, Z.P.; Zhong, L.-J.; Nordon, A.; Littlejohn, D.; Holden, M.; Fazenda, M.; Harvey, L.; McNeil, B.; Faulkner, J.; Morris, J. *Anal. Chem.* **2011**, *83*, 2655-2659

[30]. Jin, J.W.; Chen, Z.P.; Li, L.M.; Steponavicius, R.; Thennadil, S.N.; Yang, J.; Yu, R.Q. *Anal. Chem.* **2012**, *84*, 320–326

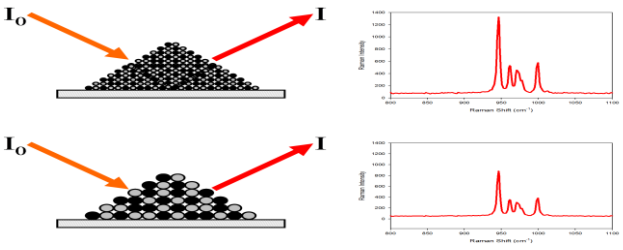
[31]. Geladi, P.; MacDougall, D.; Martens, H. *Appl. Spectrosc.* **1985**, *39 (3)*, 491-500

[32]. Barnes, R.J.; Dhanoa, M.S.; Lister, S.J. *Appl. Spectrosc.,* **1989**, *43*, 772-777

[33]. Pedersen, D.; Martens, H.; Nielsen, J.; Engelsen, S. *Appl. Spectrosc.* **2002**, *56 (9)*, 1206-1214

482

For TOC only



483