Supporting Information

Title of the primary article:

Quantitative Spectroscopic Analysis of Heterogeneous Mixtures: the Correction of Multiplicative Effects Caused by Variations in Physical Properties of Samples

Authors' names:

Jing-Wen Jin[a], Zeng-Ping Chen[a], Li-Mei Li[a], Raimundas Steponavicius[b], Suresh N. Thennadil[c], Jing Yang[a] and Ru-Qin Yu*[a]*

Affiliations:

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

b. School of Chemical Engineering and Advanced Materials, Newcastle University, Merz Court, Newcastle upon Tyne, NE1 7RU, United Kingdom

c. Chemical and Process Engineering, University of Strathclyde, 75 Montrose Street, Glasgow, G1 1XJ, United Kingdom

Table of content:

1) The MATLAB code for the modified OPLEC method

2) Figure S-1: The plots of $\mathbf{p}$ vs $\mathbf{U}_s\mathbf{U}_s^{\mathrm{T}}\mathbf{p}$ (a) and $diag(\mathbf{c}_1)\mathbf{p}$ vs $\mathbf{U}_s\mathbf{U}_s^{\mathrm{T}}diag(\mathbf{c}_1)\mathbf{p}$ (b) for the four component suspension data.

3) Figure S-2: The 129 raw calibration spectra of the tecator data.

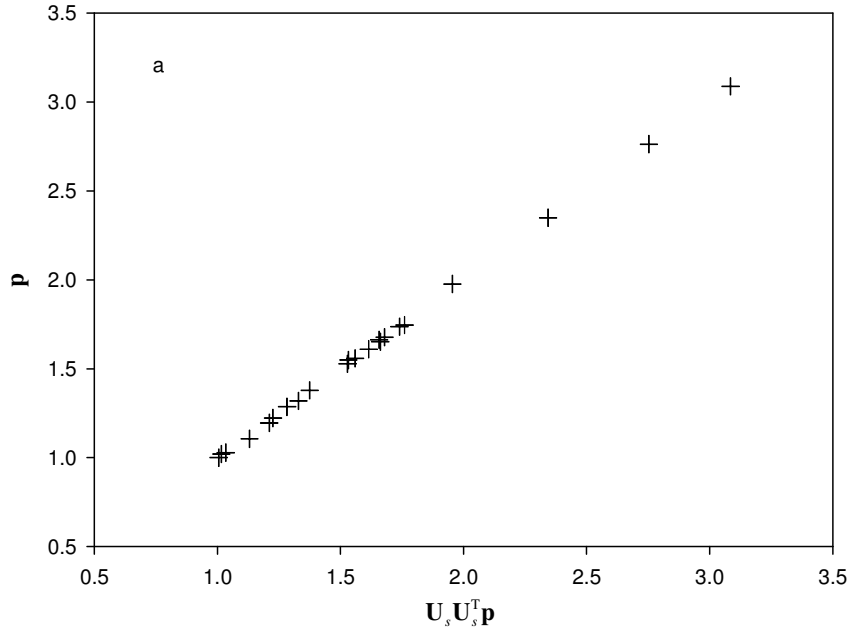
The MATLAB code for the modified OPLEC method

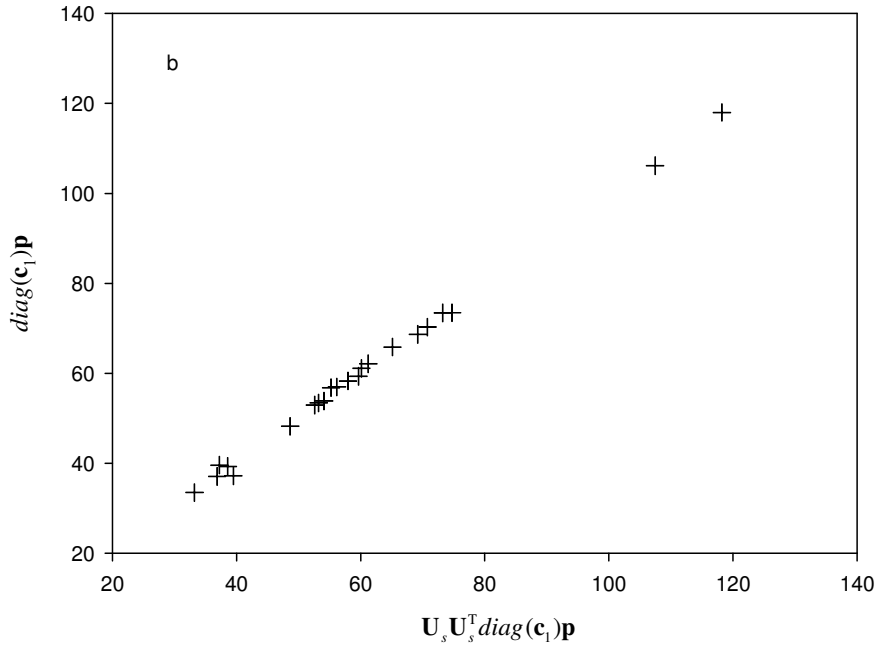
```
% [p, fval] = OPLEC_m(X, c, CompNumb);
% This is an m-file for the estimation of the multiplicative effect vector p for calibration samples;
% X contains x_i in its rows; x_i (i = 1,2,...,I) are the spectra of I calibration samples.
% c is the concentration vector of the target chemical component in the calibration samples;
% CompNumb is the number of spectroscopically active chemical components in mixture samples;
% p is a vector containing the multiplicative scattering parameters for the calibration samples;
% fval is the value of objective function at p;

function [p, fval]=OPLECm(X, c, CompNumb);
[U,S,V]=svd(X);
Us= U(:,1:CompNumb);
n=length(c);
w=max(c);
H1=eye(n, n)- Us* Us';
H2= diag(c./w)*H1* diag(c./w);
```

```
H=H1+H2;    % matrix H in min(0.5*p'*H*p+f'*p);
f=zeros(n,1);    % vector f in min(0.5*p'*H*p+f'*p);
A=-eye(n,n);    % matrix A in A*p<=b;
b=-ones(n,1);    % vector b in A*p<=b;
StartingVect=ones(n,1);
options=optimset('quadprog');
options=optimset(options,'LargeScale','off','Display','off');
[p,fval]=quadprog(H,f,A,b,[],[],[],[],StartingVect,options);
```

% After obtaining the model parameter vector **p** for calibration samples, two calibration models are built using the standard PLS toolbox. One is between the concentration vector ($\mathbf{c}$) of the target chemical component and the spectral data **X;** the other is between $diag(\mathbf{c})\mathbf{p}$ and **X**. The multiplicative effect on the test sample can then be corrected through dividing the prediction of the second calibration model by the prediction of the first calibration model.

1) **Figure S-1: The plots of p vs $\mathbf{U}_s\mathbf{U}_s^T\mathbf{p}$ (a) and $diag(\mathbf{c}_1)\mathbf{p}$ vs $\mathbf{U}_s\mathbf{U}_s^T diag(\mathbf{c}_1)\mathbf{p}$ (b) for the four component suspension data. The number of columns in $\mathbf{U}_s$ is three.**





2) Figure S-2: The 129 raw calibration spectra of the tecator data.