

## Robust Calibration Model Transfer

Yi Mou<sup>\*a</sup>, Long Zhou<sup>a</sup>, Shujian Yu<sup>c</sup>, WeiZhen Chen<sup>a</sup>, Xu Zhao<sup>a</sup>, Xinge You<sup>b</sup>,

<sup>a</sup>School of Electrical and Electronics Engineering, Wuhan Polytechnic University, Wuhan 430074, China

<sup>b</sup>Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

<sup>c</sup>Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA

### Abstract

Calibration model transfer is an important issue in infrared spectra analysis. For identical sample, spectra collected with master and slave spectrometers share same components. In the sense of mathematics, they share same basis. If the basis and corresponding coefficient matrices can be obtained, the model transfer can be efficiently realized. On the other hand, the performance of calibration model transfer method will degraded if there are outliers and noise in samples. In this paper, a robust calibration transfer model is proposed. Cauchy estimator are employed to learn same basis shared by master and slave spectra robustly. Transformation matrix can be calculated with the two corresponding coefficients matrices. Slave testing spectra are represented with the common basis and corresponding coefficients are then transferred using the transformation matrix. The slave testing spectra can be transferred using common basis and the corrected coefficients. The convergence property

---

\*Corresponding author. Tel & Fax: +86-27-8550-4729  
Email address: mouyi@whpu.edu.cn (Yi Mou\*)

and bound of proposed model are also discussed. Extensive experiments are conducted, experimental results demonstrate that our robust calibration transfer model can generally outperform the existing methods.

*Keywords:* Infrared spectra; Model transfer; Subspace learning

---

## 1. Introduction

Multivariate calibration model transfer is the key problem in infrared spectral quantitative analysis [1–3]. Multivariate calibration techniques are commonly employed but in practical the model is invalid if an existing model is applied to spectra measured under different circumstance (temperature and humidity) or on a separate instrument[4]. Recalibration can be utilized to break the limitation, but it is expensive and time consuming. This motivates calibration transfer method which refers to the transfer of quantitative analysis model between different instruments or conditions[5, 6]. Calibration transfer plays an important role, because of the possibility of using an existing model to analyze new samples obtained in new conditions or with a new instrument without the need to build the calibration model again[7–10].

Various calibration model transfer methods have been proposed, which have been comprehensively discussed[11, 12]. These methods fall into three categories. First is the pre-processing methods which eliminate or decrease the differences between master and slave spectra, including baseline elimination, derivative techniques, multiplicative scatter correction (MSC) [8], FIR filtering [9], orthogonal signal correction (OSC) [10], and generalized least squares (GLS). The second is to find a transformation matrix that maps the response of the slave instrument onto the master instrument, including Direct

1  
2  
3  
4  
5  
6  
7  
8  
9 standardization (DS) and piecewise direct standardization (PDS). The third  
10 is based on subspace learning. The transforation matrix is built in subspace.  
11  
12

13 Direct standardization (DS) and piecewise direct standardization (PDS)  
14 method are the representative approaches of second category[13, 14]. DS uses  
15 the whole spectrum on the slave instrument to fit each spectral point on the  
16 master instrument. While in PDS, a small window from the slave spectrum  
17 is used instead of the entire spectral range.  
18  
19

20 CCA is a widely employed subspace learning tool in machine learning and  
21 pattern recognition, which is successfully applied to correct the differences  
22 between spectra measured on different instruments because of its ability to  
23 reveal the correlations between them[15]. CCA is first employ to reduce the  
24 dimensionality of master and slave spectra. The transformation matrix is  
25 calculated in lower subspace. Peng pointed out that the CCA is a linear  
26 subspace learning method. There are many subspace learning techniques,  
27 such as PCA, locality preserving projections (LPP)[16, 17], and neighborhood  
28 preserving embedding (NPE) [18]. However, the dense matrices eigenvalue  
29 decomposition in these algorithms is expensive in both time and memory, and  
30 the solution of the optimization problem is usually unstable when the number  
31 of features is larger than the number of samples. Thus spectral regression  
32 based model transfer method is proposed by Peng[19] to cast the problem  
33 of learning an embedding function into a regression framework, which avoids  
34 eigenvalue decomposition. The difference between the CCA based model  
35 transfer and Peng’s method is that the direction vectors for master and slave  
36 spectra are calculated with spectral regression.  
37  
38

39 The development of multivariate calibration models involves several stages.  
40  
41

To obtain high accuracy, spectra acquisition is one of the most important stage in quantitative analysis. In particular, it would be desirable to eliminate or minimize the sources of data variability that are not related to the analytical property of interest. In fact, the spectra collected with infrared spectrometer are often affected by noise and outliers. Thus, the methods should be robust against the noise and outliers to obtain higher accuracy.

In machine learning and pattern recognition, Cauchy estimator is more robust than least square estimator[20, 21]. The underlying assumption of calibration model transfer is that the spectra collected with different instruments or in different environmental conditions share components. In the sense of mathematics the share same basis. It's the basis of calibration model transfer. Without this assumption, it's chemically and mathematically meaningless. On the other hand, the performance of calibration model transfer method will degraded if there are outliers and noise in samples. At this point, we proposed a new robust method based on Cauchy estimator to correct the spectral data.

The rest of the paper is organized as follows. We review general principles of calibration model transfer methods in Section 2. We formulate robust calibration model transfer and provide an efficient algorithm for solving the proposed model in Section 3, where analysis to convergence analysis and bound discussion are also conducted. Experiments and result analysis are provided in Sections 4 and conclusions are drawn in Section 5.

## 2. Related work

Infrared spectroscopy is an extensively employed analytical technique in many industrial applications because of its rapidness and the fact that it is non-destructive to the samples. Multivariate calibration techniques are commonly used method to build quantitative analysis model, such as partial least squares (PLS) regression [1] and principal component regression (PCR) [2]. However, a problem occurs when an existing model is applied to spectra that were measured under new environmental conditions or on another instrument. New spectra contain variation which can lead to erroneous predictions. A possible solution to this calibration transfer problem is to measure every sample in the new instrument and construct a new model for it. However, this process would be both costly and time consuming. A more acceptable way is to apply chemometrics techniques to correct the difference of spectra measured on two instruments.

### *2.1. Classical Calibration Model*

Let  $\mathbf{X}_s$  and  $\mathbf{X}_m$  be the spectral matrices obtained from master and slave instruments respectively.  $\bar{\mathbf{X}}_s$  and  $\bar{\mathbf{X}}_m$  are subsets of  $\mathbf{X}_s$  and  $\mathbf{X}_m$  respectively,  $C$  and  $\bar{C}$  are corresponding concentration matrices.  $K_s$  and  $K_m$  are the matrices of sensitivities on both instruments, each row is the pure components spectra. The relationships between the concentration matrices and observed matrices  $\mathbf{X}_s$  and  $\mathbf{X}_m$  are:

$$\mathbf{X}_m = \mathbf{C}\mathbf{K}_m \quad (1)$$

$$\mathbf{X}_s = \mathbf{C}\mathbf{X}_s = \mathbf{C}(\mathbf{K}_m + \Delta\mathbf{K}) \quad (2)$$

where  $\Delta\mathbf{K}$  is the difference matrix between  $\mathbf{K}_s$  and  $\mathbf{K}_m$ . The same relationship is hold for the subsets.

$$\bar{\mathbf{X}}_m = \bar{\mathbf{C}}\mathbf{K}_m \quad (3)$$

$$\bar{\mathbf{X}}_s = \bar{\mathbf{C}}\mathbf{X}_s = \bar{\mathbf{C}}(\mathbf{K}_m + \Delta\mathbf{K}) \quad (4)$$

The difference matrix  $\Delta K$  can be calculated by:

$$\Delta \mathbf{K} = \bar{\mathbf{C}}^+ (\bar{\mathbf{X}}_s - \bar{\mathbf{X}}_m) \quad (5)$$

Then  $\mathbf{X}_s$  is estimated as

$$\hat{\bar{\mathbf{X}}}_s = \bar{\mathbf{X}}_m + \mathbf{C}\bar{\mathbf{C}}^+(\bar{\mathbf{X}}_s - \bar{\mathbf{X}}_m) \quad (6)$$

with  $\bar{\mathbf{X}}_s$  and  $\mathbf{C}$ , a new calibration model can be built for prediction on the slave instrument. Two assumptions are implied in this method: the linear relationship should hold on both instruments and the concentrations for all analytes contributing to the response must be known. For complex compound, where the concentrations are not known, this method is invalid.

## 2.2. Direct and Piecewise Direct Standardization

In direct standardization, response matrices on both instruments are related to each other by a transformation matrix  $\mathbf{F}$ :

$$\bar{\mathbf{X}}_m = \bar{\mathbf{X}}_s \mathbf{F} \quad (7)$$

where  $\mathbf{F}$  is calculated as:

$$\mathbf{F} = \bar{\mathbf{X}}_s^+ \bar{\mathbf{X}}_m \quad (8)$$

1  
2  
3  
4  
5  
6  
7  
8  
9 J. Nie studied calibration model transfer between different temperatures.  
10  
11 The method is based on direction standardization.

12  
13 Piecewise direct standardization builds several local regression models to  
14 fit the every wavelength of the master spectra using a range of wavelengths of  
15 the slave spectra. the  $i$ th wavelength of master spectra  $\mathbf{X}_m(:, i)$  is regressed  
16 on the slave piece  $[\mathbf{X}_s(:, i - j), \dots, \mathbf{X}_s(:, j + k)]$ . According to the PLS or  
17 PCR regression model,  $\mathbf{X}_m(:, i) = [\mathbf{X}_s(:, i - j), \dots, \mathbf{X}_s(:, j + k)]b_i$ . We obtain  
18 the matrix  $\mathbf{F} = \text{diag}(b_1^T, \dots, b_p^T)$ . The linear relation between master and  
19 slave instruments is  $\mathbf{X}_m = \mathbf{X}_s\mathbf{F}$ .  
20  
21  
22  
23  
24  
25  
26  
27  
28

### 2.3. CCA based Calibration Model Transfer

29 CCA based model transfer is proposed by W. Fan[15]. Canonical correlation  
30 analysis is employed to analyzed the mean centered standardization sets  
31  $\mathbf{T}_m$  and  $\mathbf{T}_s$ . Direction matrices  $\mathbf{W}_m$  and  $\mathbf{W}_s$  are obtained and the canonical  
32 variables  $\mathbf{L}_m$  and  $\mathbf{L}_s$  are obtained by projecting the standardization sets  
33 on to the direction matrices. The transform matrix  $\mathbf{F}_1$  and  $\mathbf{F}_2$  are deter-  
34 mined by  $\mathbf{F}_1 = \mathbf{L}_s^+\mathbf{L}_m$  and  $\mathbf{F}_2 = \mathbf{L}_m^+\mathbf{T}_m$ . The transferred prediction set  $\mathbf{Z}_s$  is  
35  
36  $\mathbf{Z} = \mathbf{K}_s\mathbf{F}_1\mathbf{F}_2$  where  $\mathbf{K}_s = \mathbf{P}_s\mathbf{W}_s$ .

### 43 2.4. Spectral regression based Calibration Model Transfer

44 The samples are divided into two parts: calibration set and prediction set.  
45 The standardization set  $\mathbf{X}_{ms}$  is selected among the calibration set of master  
46 instrument using subset selection method, and the same subset on the slave  
47 instrument gives rise to standardization set  $\mathbf{X}_{ss}$ . Spectral regression is carried  
48 out to  $\mathbf{X}_{ms}$  and  $\mathbf{X}_{ss}$  respectively. The transformation matrix  $\mathbf{U}$  and  $\mathbf{V}$  map  
49 the standardization sets to two low-dimensional subspaces, and two subspace  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

variables  $\mathbf{L}_m$  and  $\mathbf{L}_s$  are obtained. Two transformation matrices  $\mathbf{F}_1$  and  $\mathbf{F}_2$  are obtained  $\mathbf{F}_1 = \mathbf{L}_s^+ \mathbf{L}_m$  and  $\mathbf{F}_2 = \mathbf{L}_m^+ \mathbf{X}_{ms}$ . The transferred prediction set  $\mathbf{Z}_s$  is  $\mathbf{Z} = \mathbf{K}_s \mathbf{F}_1 \mathbf{F}_2$  where  $\mathbf{K}_s = \mathbf{P}_s \mathbf{W}_s$ . It's clear that the difference between CCA based calibration model transfer and spectral regression based method is the way of dimension reduction.

### 3. Calibration model transfer based on subspace learning

For identical sample, spectra collected with master and slave spectrometers share same components. In the sense of mathematics, they share same basis. The differences between them are due to the representation coefficients, as shown in Fig.1. If the same basis and corresponding coefficients are obtained, the model transfer can be efficiently realized. On the other hand, the performance of calibration model transfer method will degraded if there are outliers and noise in samples. In this paper, a robust calibration transfer model is proposed.

#### 3.1. Notation

Samples collected on master and slave instruments are defined as  $\mathbf{X}_m \in R^{n \times p}$  and  $\mathbf{X}_s \in R^{n \times p}$  respectively, where  $n$  is the number of samples and  $p$  is the dimension of data.

#### 3.2. Robust estimator

M-estimator is popular in robust statistics. Let  $r_i$  be the residual of the  $i$ -th data point, i.e., the difference between the  $i$ -th observation and its fitted values. The standard least squares method tries to minimize  $\sum_i r_i^2$ , which is unstable if outliers are present, and which has a strong effect to distort the

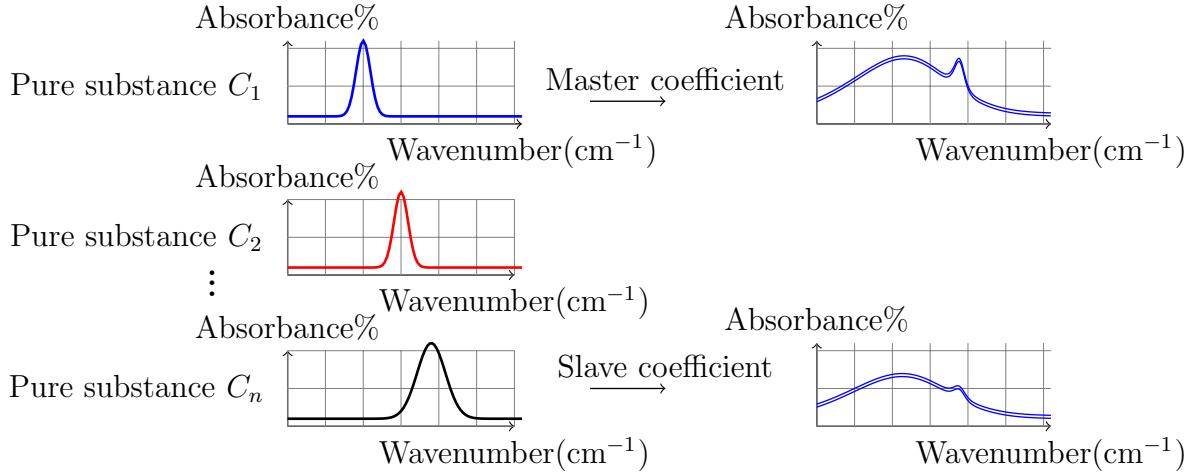


Figure 1: Difference between master and slave spectra

estimated parameters. M-estimators try to reduce the effect of outliers by replacing the squared residuals  $r_i^2$  with another function of residuals

$$\min \sum_i \rho(r_i) \quad (9)$$

where  $\rho$  is a symmetric, positive-definite function with a unique minimum at zero and chosen to be less increasing than the square function. The corresponding influence function is defined as:

$$\psi(x) = \frac{\partial \rho(x)}{x} \quad (10)$$

which measures the influence of a random data point on the value of the parameter estimate.

As shown in Fig.2, for the  $\ell_2$  estimator (least-squares) with  $\rho(x) = \frac{x^2}{2}$ , the influence function is  $\psi(x) = x$ ; that is, the influence of a data point on the estimate increases linearly with the size of its error. This confirms the non-robustness of the least-squares estimate. Although the  $\ell_1$  (absolute value)

estimator with  $\rho(x) = |x|$  reduces the influence of large errors, its influence function has no cut-off. When an estimator is robust, it is inferred that the influence of any single observation is insufficient to yield a significant offset. Cauchy estimator has been shown to own this valuable property

$$\rho(x) = \log\left(1 + \left(\frac{x}{c}\right)^2\right) \quad (11)$$

along with the upper bounded influence function:

$$\psi(x) = \frac{2x}{c^2 + x^2} \quad (12)$$

Therefore, we deploy the Cauchy estimator in PLS.

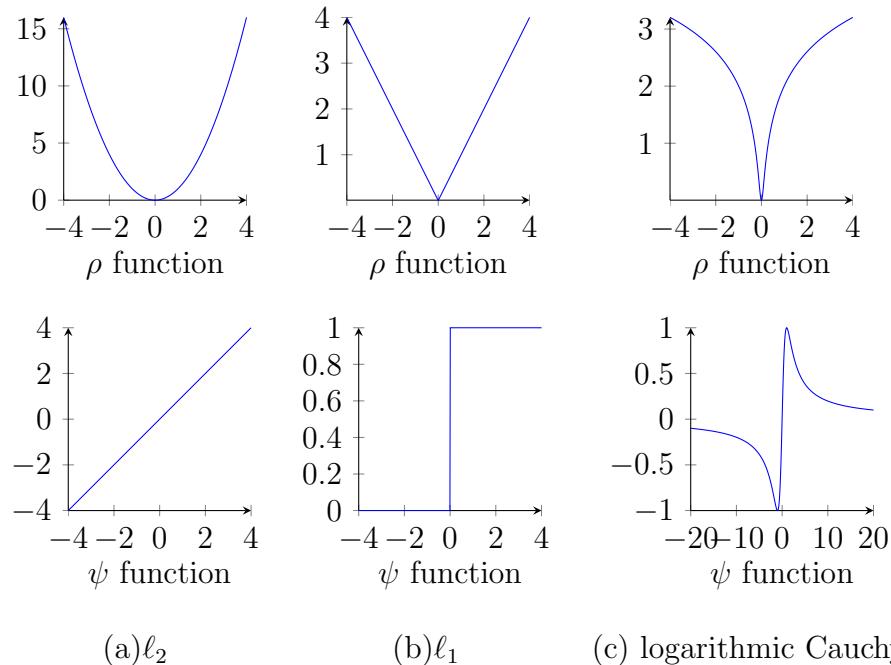
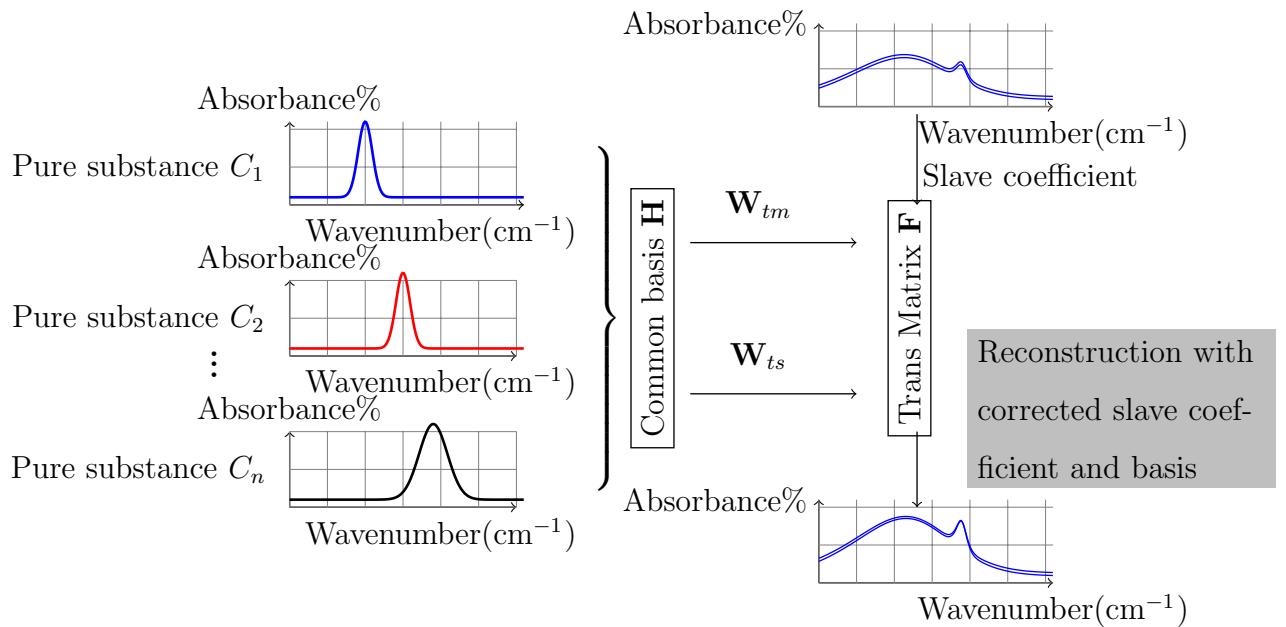


Figure 2:  $\rho$  and  $\psi$  function

1  
2  
3  
4  
5  
6  
7  
8  
9     3.3. Robust model transfer  
10

11 Samples collected from different instruments own same basis. Linear rep-  
12 resentation coefficient leads to the difference between  $\mathbf{X}_m$  and  $\mathbf{X}_s$ . We can  
13 obtain their common basis  $\mathbf{H}$ , and corresponding representation coefficients  $\mathbf{W}_m$   
14 and  $\mathbf{W}_s$ . Transformation matrix  $\mathbf{F}$  can be calculated as  $\mathbf{F} = \mathbf{W}_s^+ \mathbf{W}_m$ . The  
15 basis  $\mathbf{H}$  is then utilized to represent the slave test sets  $\mathbf{X}_{ts}$ , and  $\mathbf{W}_{ts}$  is  
16 corresponding coefficient. The slave testing sample can be corrected with  
17 transformation matrix and basis. The flow chart of proposed algorithm is  
18



47     Figure 3: Flow of proposed model transfer  
48  
49  
50

51     shown in Fig.3.  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

In learning procedure, the basis  $\mathbf{H}$  satisfies the model:

$$\begin{aligned} & \min \log \left(1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1}\right) + \log \left(1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2}\right) \quad (13) \\ & s.t. \mathbf{H}^T \mathbf{H} = \mathbf{I} \end{aligned}$$

where  $\mathbf{W}_m$  and  $\mathbf{W}_s \in R^{m \times p}, \mathbf{H} \in R^{n \times m}$ . The corresponding Lagrange function:

$$\mathcal{L} = \log \left(1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1}\right) + \log \left(1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2}\right) + \frac{1}{2}\lambda(\mathbf{I} - \mathbf{H}^T \mathbf{H}) \quad (14)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{H}} = \frac{(\mathbf{X}_m - \mathbf{HW}_m)\mathbf{W}_m^T}{c_1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2} + \frac{(\mathbf{X}_s - \mathbf{HW}_s)\mathbf{W}_s^T}{c_2 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2} - \lambda \mathbf{H} = 0 \quad (15)$$

Define:

$$\begin{aligned} \alpha &= \frac{1}{c_1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2} \\ \beta &= \frac{1}{c_2 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2} \end{aligned}$$

Then

$$\frac{\partial \mathcal{L}}{\partial \mathbf{H}} = \alpha(\mathbf{X}_m - \mathbf{HW}_m)\mathbf{W}_m^T + \beta(\mathbf{X}_s - \mathbf{HW}_s)\mathbf{W}_s^T - \lambda \mathbf{H} = 0$$

$$\alpha \mathbf{X}_m \mathbf{W}_m^T + \beta \mathbf{X}_s \mathbf{W}_s^T = \mathbf{H}(\alpha \mathbf{W}_m \mathbf{W}_m^T + \beta \mathbf{W}_s \mathbf{W}_s^T + \lambda)$$

$$(\alpha \mathbf{X}_m \mathbf{W}_m^T + \beta \mathbf{X}_s \mathbf{W}_s^T)(\alpha \mathbf{W}_m \mathbf{W}_m^T + \beta \mathbf{W}_s \mathbf{W}_s^T + \lambda)^{-1} = \mathbf{H} \quad (16)$$

Transformation matrix  $\mathbf{F}$  is:

$$\mathbf{F} = \mathbf{W}_m \mathbf{W}_s^+ \quad (17)$$

---

**Algorithm 1** Robust model transfer algorithm

---

**Input:** Master and slave spectral matrices  $\mathbf{X}_m$  and  $\mathbf{X}_s$  for training,  $\mathbf{X}_{tm}$  and  $\mathbf{X}_{ts}$  for testing, initial  $\lambda$ ,  $c_1 = c_2 = 1$ ,  $\alpha = \alpha^0$ ,  $\beta = \beta^0$ ,  $\mathbf{W}_m = \mathbf{W}_m^0$ ,  $\mathbf{W}_s = \mathbf{W}_s^0$ , Set the iteration times  $k$

- 1:  $\mathbf{H}^k = (\alpha^{k-1}\mathbf{X}_m(\mathbf{W}_m^{k-1})^T + \beta^{k-1}\mathbf{X}_s(\mathbf{W}_s^{k-1})^T)(\alpha^{k-1}\mathbf{W}_m^{k-1}(\mathbf{W}_m^{k-1})^T + \beta^{k-1}\mathbf{W}_s^{k-1}(\mathbf{W}_s^{k-1})^T + \lambda)^{-1}$ ;
- 2:  $\mathbf{W}_m^k = (\mathbf{H}^k)^+\mathbf{X}_m$ ,  $\mathbf{W}_s^k = (\mathbf{H}^k)^+\mathbf{X}_s$ ;
- 3:  $\alpha^k = \frac{1}{c_1 + \|\mathbf{X}_m - \mathbf{H}^k \mathbf{W}_m^k\|_F^2}$ ,  $\beta^k = \frac{1}{c_2 + \|\mathbf{X}_s - \mathbf{H}^k \mathbf{W}_s^k\|_F^2}$ ;
- 4:  $k = k + 1$ ;
- 5: Transformation matrix  $\mathbf{F}$  is:  $\mathbf{F} = \mathbf{W}_m \mathbf{W}_s^+$ .
- 6: Represent  $\mathbf{X}_{ts}$  with obtained  $\mathbf{H}$  and the corresponding coefficient is  $\mathbf{W}_{ts}$
- 7: Transferred slave testing spectra matrix is:  $\mathbf{X}_{ts}^* = \mathbf{H} \mathbf{F} \mathbf{W}_{ts}$

---

### 3.4. Convergence analysis

For simplicity, we consider  $\mathbf{h} \in R^{n \times 1}$ ,  $\mathbf{w} \in R^{1 \times p}$ . After having found  $\mathbf{h}^{(k)}$ , we can construct a quadratic function  $\psi(\mathbf{h}, \mathbf{h}^{(k)})$  to upper bound  $\mathcal{L}(\mathbf{h})$  such that the following conditions hold[22–24]:

$$\psi(\mathbf{h}^{(k)}, \mathbf{h}^{(k)}) = \mathcal{L}(\mathbf{h}^{(k)}) \quad (18)$$

$$\psi'(\mathbf{h}^{(k)}, \mathbf{h}^{(k)}) = \mathcal{L}'(\mathbf{h}^{(k)}) \quad (19)$$

The function  $\psi(\mathbf{h}^{(k)}, \mathbf{h}^{(k)})$  is:

$$\begin{aligned} \psi(\mathbf{h}, \mathbf{h}^{(k)}) &= \mathcal{L}(\mathbf{h}^{(k)}) + (\mathbf{h} - \mathbf{h}^{(k)})^T \mathcal{L}'(\mathbf{h}^{(k)}) \\ &+ (\mathbf{h} - \mathbf{h}^{(k)})^T (\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda) (\mathbf{h} - \mathbf{h}^{(k)}) \end{aligned} \quad (20)$$

Assume that  $\psi(\mathbf{h}, \mathbf{h}^{(k)})$  is locally convex with respect to  $\mathbf{h}$  and has a local minimizer. Setting  $\mathbf{h}^{k+1}$  as this minimizer, then

$$\psi'(\mathbf{h}^{(k+1)}, \mathbf{h}^{(k)}) = \mathcal{L}'(\mathbf{h}^{(k)}) + 2(\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda)(\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}) = 0 \quad (21)$$

By appropriately choosing  $\mathbf{h}^{k+1}$  near  $\mathbf{h}$ , we have  $\mathcal{L}(\mathbf{h}) \leq \psi(\mathbf{h}, \mathbf{h}^{(k)})$ , which means that

$$\begin{aligned} \mathcal{L}(\mathbf{h}^{(k+1)}) &\leq \psi(\mathbf{h}^{(k+1)}, \mathbf{h}^{(k)}) \\ &= \mathcal{L}(\mathbf{h}^{(k)}) + (\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)})^T \mathcal{L}'(\mathbf{h}^{(k)}) \\ &+ (\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)})^T (\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda)(\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}) \end{aligned}$$

By combining Eqs. (21) and (22), we have:

$$\begin{aligned} \mathcal{L}(\mathbf{h}^{(k+1)}) - \mathcal{L}(\mathbf{h}^{(k)}) &\leq -(\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda)(\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)})^T (\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}) \\ &\leq -(\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda) \|\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}\|_F^2 \leq 0 \end{aligned} \quad (23)$$

### 3.5. Bound discuss

The proposed model is represented as (24). We noticed that function  $\log(1 + x^2)(x > 0)$  is monotonously increased function, shown as Fig.4:

The object function can be expressed as:

$$\begin{aligned} \min \log &\left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2 \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_1 c_2} \right) \\ s.t. &\mathbf{H}^T \mathbf{H} = \mathbf{I} \end{aligned} \quad (24)$$

The inequality holds:

$$\frac{\left( \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_t - \mathbf{HW}_t\|_F^2}{c_2} \right)^2}{4} \geq \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2 \|\mathbf{X}_t - \mathbf{HW}_t\|_F^2}{c_1 c_2} \quad (25)$$

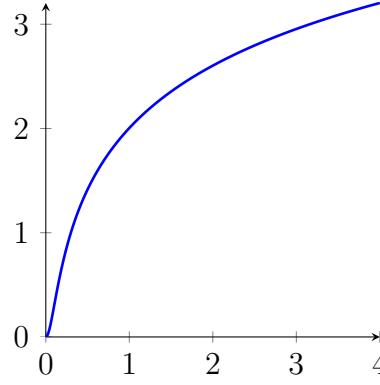


Figure 4:  $\log(1 + x^2)$

Then we have:

$$\begin{aligned}
 & \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2 \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_1 c_2} \\
 & \leq \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\left(\frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_t - \mathbf{HW}_t\|_F^2}{c_2}\right)^2}{4} \quad (26)
 \end{aligned}$$

We notice that:

$$\begin{aligned}
 \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} &= \left\| \begin{pmatrix} \frac{1}{\sqrt{c_1}} \mathbf{X}_m \\ \frac{1}{\sqrt{c_2}} \mathbf{X}_s \end{pmatrix} - (\mathbf{H}, \mathbf{H}) \begin{pmatrix} \frac{1}{\sqrt{c_1}} \mathbf{W}_m \\ \frac{1}{\sqrt{c_2}} \mathbf{W}_s \end{pmatrix} \right\|_F^2 \\
 &= \|\mathbf{X} - \mathbf{HW}\|_F^2 \quad (27)
 \end{aligned}$$

where  $\mathbf{X} = \begin{pmatrix} \frac{1}{\sqrt{c_1}} \mathbf{X}_m \\ \frac{1}{\sqrt{c_2}} \mathbf{X}_s \end{pmatrix}$ ,  $\mathbf{H} = (\mathbf{H}, \mathbf{H})$  and  $\mathbf{W} = \begin{pmatrix} \frac{1}{\sqrt{c_1}} \mathbf{W}_m \\ \frac{1}{\sqrt{c_2}} \mathbf{W}_s \end{pmatrix}$ .

$$\begin{aligned}
 & \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2 \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_1 c_2} \\
 & \leq \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\left(\frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_t - \mathbf{HW}_t\|_F^2}{c_2}\right)^2}{4} \\
 & \leq \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \quad (28)
 \end{aligned}$$

Then we have:

$$\begin{aligned} 0 &\leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) \\ &\leq \log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) \end{aligned} \quad (29)$$

where  $\sigma$  is the singular value of the matrix  $\mathbf{X}$ ,  $c_1 = c_2 = 1$ .

#### 4. Experiments and discussing

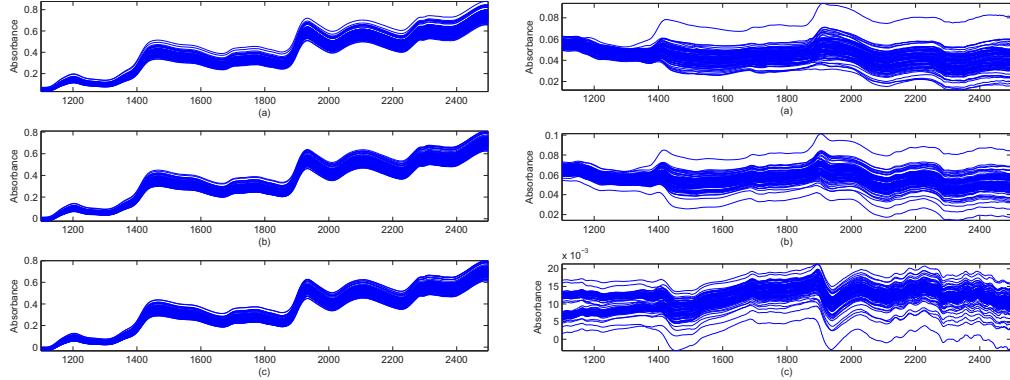
In order to evaluate the performance of proposed method, three data sets are used and five benchmark methods are re-implemented,i.e., PDS, DS, CCA, KCCA, SR Square Error of Prediction(RMSEP) to measure the prediction precision which is defined as:

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (30)$$

All experiments are implemented using Matlab and run in a personal computer with a 2.80GHz Intel Core i5-2300 CPU, 4GB RAM, and Windows 7 operation system. We set the upper limit of number of latent components to be 10. For each latent variable, the experiment is repeated 50 times independently and average RMSEP is calculated.

##### 4.1. Corn spectra analysis

This data set consists of 80 samples of corn measured on 3 different NIR spectrometers. The wavelength range is 1100-2498nm at 2 nm intervals (700 channels), as shown in Fig.5. The moisture, oil, protein and starch values for each of the samples is also included.



(a) Corn spectra. (a) m5; (b) mp5; (c) mp6 (b) Spectral differences. (a) m5-mp5; (b) m5-mp6; (c) mp5-mp6

Figure 5: Corn spectra and differences

Two experiments are conducted with this set. M5 is selected as master instrument, while MP5 and MP6 are selected as slave instruments, respectively. The number of training samples are 15,20,25,30,35,40. The default number of the basis is fixed as 10. Fig.6 and 7 are the convergence analysis results of corn spectral model transfer analysis. During the iteration process of the proposed algorithm, the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$ ,  $\mathbf{H}$ ,  $\alpha$  and  $\beta$  finally converged as shown in Fig. 6 and 7. It's clear that during the iteration process,  $\mathbf{HW}_m$  and  $\mathbf{HW}_s$  are approaching  $\mathbf{X}_m$  and  $\mathbf{X}_s$ . Thus the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  is decreased and converged.

According to Eq.(29), in corn analysis, the upper bound of objection is :

$$m5 \text{ is master and } mp5 \text{ is slave: } \log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) = 0.0050$$

$$m5 \text{ is master and } mp6 \text{ is slave: } \log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) = 0.0054$$

In our experiments, the inequalities hold:

$$0 \leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) < 0.0050$$

$$0 \leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) < 0.0054$$

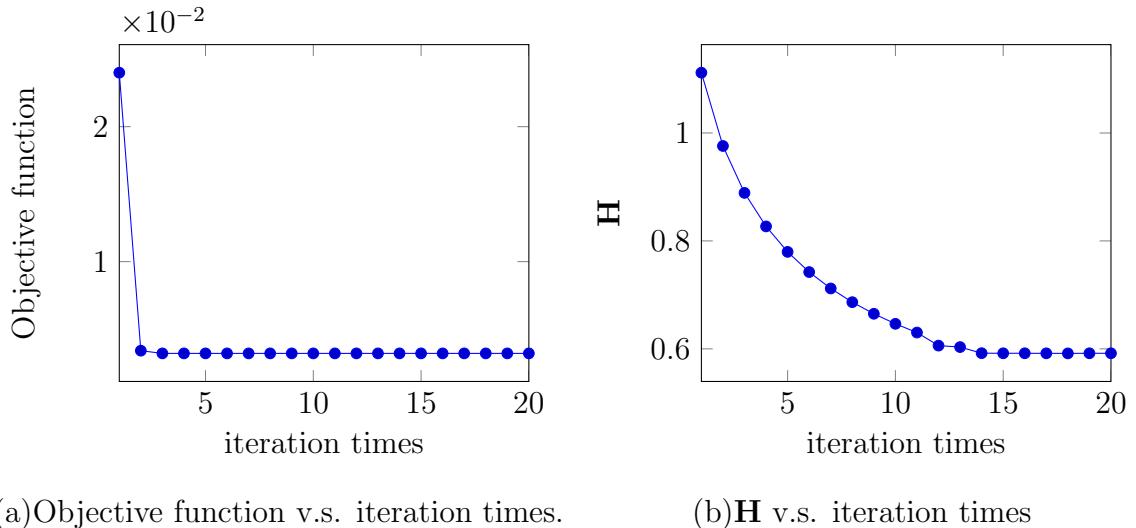


Figure 6:  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  and  $\mathbf{H}$  v.s. iteration times in corn spectral analysis

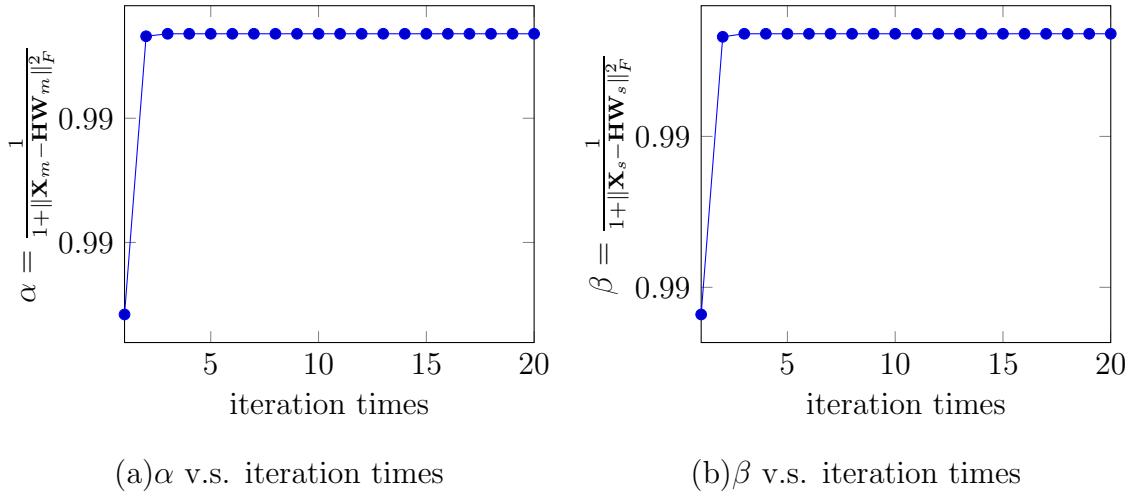


Figure 7:  $\alpha$  and  $\beta$  v.s. iteration times in corn spectral analysis

The results are listed in Table 1 and Table 2. The figures in bold represent the best results. As shown in Table 1 and 2, the proposed method have higher accuracy than other method on the whole.

#### 4.2. Tablet spectra analysis

Data set 2 is a public source data for calibration transfer from the IDRC Shootout 2002 [18,19]. Spectra of pharmaceutical tablets from two Multi-tab spectrometers (FossNIRsystems, Silverspring, MD) are measured in the transmittance mode. Tablets data from two instruments have been split into two calibration sets (155 tablets, Calibrate 1 and Calibrate 2) and two test sets (460 tablets, Test 1 and Test 2). All log (1/T) spectra cover the region from 600 to 1898 nm in 2 nm increments. The data set includes tablets with a wide ASSAY range, 152-239 mg. The spectra and the difference are shown as Fig.8.

Calibrate1 is selected as master instrument, while calibrate2 is selected

Table 1: Result: m5 is master, mp5 is slave

Response	Method	Samples	15	20	25	30	35	40
Water	NO	2.4218	2.8003	1.6277	1.5132	0.5088	0.3726	
	PDS	15.1419	6.1223	3.7922	3.7577	2.8557	1.7235	
	DS	0.4387	0.4456	0.5635	0.5534	0.4898	0.4235	
	CCA	0.8909	0.7772	0.8308	0.7272	0.5253	0.4213	
	KCCA	0.8308	0.8209	1.3517	0.9640	0.3831	0.7299	
	SR	1.3156	1.1146	0.5313	0.4255	0.3768	0.3799	
	Proposed	1.2998	0.8905	0.5308	0.3904	0.3371	<b>0.3297</b>	
Oil	NO	2.3112	0.2118	0.3478	0.3567	0.3236	0.7346	
	PDS	6.6020	3.2630	2.2290	0.3112	0.2134	0.5998	
	DS	0.5324	0.4434	0.3678	0.2346	0.1975	0.2389	
	CCA	0.3390	0.2701	0.2370	0.2670	0.1790	0.2601	
	KCCA	0.3599	0.2834	0.1911	0.2698	0.1690	0.2793	
	SR	0.3410	0.3561	0.2710	0.1998	0.2197	<b>0.1621</b>	
	Proposed	0.3591	0.3361	0.2498	0.1997	0.1727	0.1658	
Protein	NO	1.8174	3.3435	6.4569	2.5678	1.2348	1.3654	
	PDS	2.7890	9.5678	4.3467	7.3257	5.4556	4.6012	
	DS	0.4315	0.5790	0.5362	0.5137	0.5314	0.4290	
	CCA	0.7457	0.9156	0.6457	0.6465	0.6343	0.4148	
	KCCA	0.7855	0.9545	0.6143	0.6458	0.5557	0.4665	
	SR	0.7789	0.7783	0.6836	0.6012	0.5024	0.4454	
	Proposed	1.0345	0.8904	0.8723	0.5244	0.4903	<b>0.4012</b>	
Starch	No	3.2123	2.3432	3.4356	1.2123	1.8904	1.5235	
	PDS	16.1278	16.0324	12.1244	4.4636	5.23524	4.4078	
	DS	0.9823	1.3256	1.0349	0.9568	1.1890	0.9513	
	CCA	1.7601	2.2223	1.5219	1.1908	1.4023	1.0123	
	KCCA	1.6690	2.2794	1.6001	1.2101	1.4123	0.8901	
	SR	1.6698	2.2526	1.6572	1.0501	0.9810	0.8553	
	Proposed	1.4978 <sub>20</sub>	2.1126	1.6123	0.9546	0.9256	<b>0.8443</b>	

Table 2: Results: m5 is master, mp6 is slave

Response	Samples method	15	20	25	30	35	40
Water	No	6.1435	5.1906	3.2563	1.6789	1.8345	1.9101
	PDS	8.5698	7.3290	12.0116	5.6571	5.3909	3.9562
	DS	0.5437	0.3048	0.4547	0.4325	0.4821	0.3790
	CCA	1.1412	0.4987	0.5909	0.6019	0.4890	0.4014
	KCCA	1.1909	0.7802	0.7279	0.4015	0.3801	0.3210
	SR	0.7012	0.3669	0.4998	0.3992	0.3621	0.3225
	Proposed	1.2120	0.3612	0.5812	0.4230	0.3561	<b>0.3105</b>
Oil	NO	2.2563	1.9235	1.6312	1.1568	0.6378	0.7187
	PDS	5.0231	0.9021	2.7352	2.2134	1.7234	0.7123
	DS	0.2410	0.2211	0.1998	0.1940	0.1825	0.1730
	CCA	0.2912	0.3543	0.2271	0.2478	0.2123	0.2012
	KCCA	0.5234	0.2812	0.2524	0.2378	0.2344	0.1956
	SR	0.2623	0.2680	0.1923	0.1846	0.1804	0.1725
	Proposed	0.2742	0.2612	0.1712	0.1856	<b>0.1520</b>	0.1629
Protein	NO	9.2131	7.3245	6.3490	1.6324	3.1096	4.3245
	PDS	9.2352	15.9851	7.3271	2.5487	1.5878	1.8901
	DS	0.7201	0.5989	0.5870	0.5023	0.5324	0.5339
	CCA	0.8781	0.9245	0.5675	0.5123	0.6436	0.5501
	KCCA	0.8790	0.8330	0.5810	0.4636	0.6309	0.5871
	SR	0.9310	1.0789	0.8234	0.5660	0.5478	0.5340
	Proposed	1.5012	1.0787	0.8234	0.4279	<b>0.4014</b>	0.5001
Starch	NO	10.4366	6.1341	3.2145	2.4367	6.6578	7.54745
	PDS	20.6576	14.8798	14.3453	10.3243	8.4353	7.7234
	DS	0.8767	1.1237	0.9126	0.8787	1.0553	0.9673
	CCA	1.3998	1.8790	1.3289	1.2016	1.1203	0.9494
	KCCA	1.4490	2.1023	1.4343	1.1590	1.1413	0.8445
	SR	1.4567	1.4312	1.4023	1.1870	1.058	0.9340
	Proposed	2.6787	1.5235	1.2789	0.9413	0.8713	<b>0.8234</b>

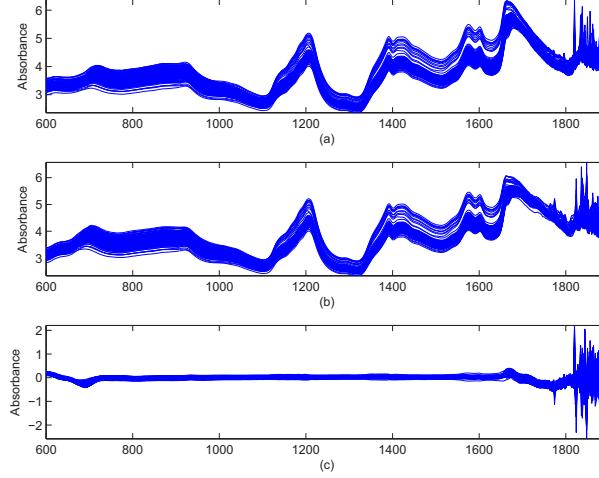


Figure 8: Tablets spectra. Spectra of 50 pharmaceutical tablets on instrument 1; (b) spectra of 50 pharmaceutical tablets on instrument 2; (c) spectral differences between the two instruments.

as slave instruments. The number of training samples are 15,20,25,30,35,40. The default number of the basis is fixed as 10. Fig.9 and 10 are the convergence analysis results of tablet spectral model transfer analysis. During the iteration process of the proposed algorithm, the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$ ,  $\mathbf{H}$ ,  $\alpha$  and  $\beta$  finally converged as shown in Fig. 9 and 10. It's clear that during the iteration process,  $\mathbf{HW}_m$  and  $\mathbf{HW}_s$  are approaching  $\mathbf{X}_m$  and  $\mathbf{X}_s$ . Thus the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  is decreased and converged. According to (29), in tablet analysis, the upper bound of objection is :

$$\log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) = 8.9455$$

In our experiments, the inequalities hold:

$$0 \leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) < 8.9455$$

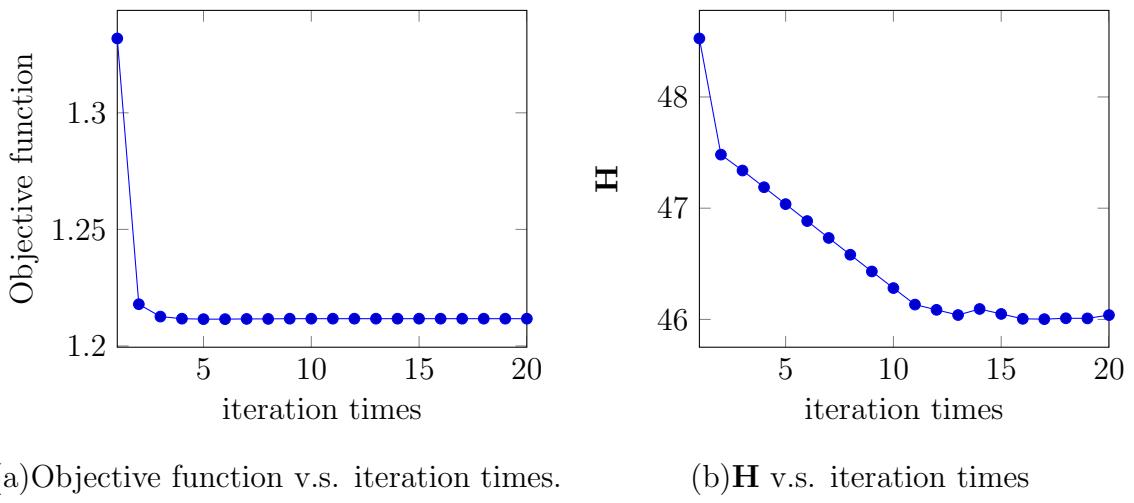


Figure 9:  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  and  $\mathbf{H}$  v.s. iteration times in tablet analysis

The results are listed in Table.3. The figures in bold represent the best results. As shown in Table.3, the proposed method have higher accuracy than other method on the whole.

#### 4.3. Spirit analysis

We have collected spirit spectrum with two Nicolet spectrometers, and the 17 components reference concentrations values are tested with Agilent HLPC. In our experiments, 4 components are analyzed. 100 spectra are collected with instruments A and B, respectively. A is selected as master instrument, while B is selected as slave instruments. The number of training samples are 15,20,25,30,35,40. The default number of the basis is fixed as 50. Fig.11 and 12 are the convergence analysis results of corn spectral model transfer analysis. During the iteration process of the proposed algorithm,

Table 3: Results of tablet analysis

Response	Method	Samples	20	25	30	35	40	45
Weight	NO	20.4543	38.4351	8.2132	43635.2314	19.2147	19.2343	
	DS	9.4355	10.2345	5.4567	9.8341	8.2344	8.5463	
	PDS	15.4564	17.7967	6.5745	15.3413	11.1212	11.6343	
	CCA	7.7896	7.8341	6.2345	10.5357	6.7896	6.7695	
	KCCA	7.3141	7.2356	6.0124	8.5674	7.2354	6.7854	
	SR	6.4563	6.4643	6.1032	7.3256	5.6623	5.6467	
	Proposed	7.4363	5.3140	5.6573	6.0196	4.6425	<b>4.6123</b>	
Hardness	NO	5.2131	5.7858	2.3141	1127.7867	4.0231	4.0213	
	DS	5.2735	2.9243	2.3579	3.9425	2.1241	1.9945	
	PDS	7.2132	4.2342	2.1234	4.6878	2.6342	2.7653	
	CCA	3.8962	2.2143	2.1231	3.7689	1.6546	1.7234	
	KCCA	3.8578	2.5742	2.1234	3.5778	1.5246	1.2231	
	SR	3.8906	1.7454	1.8986	2.1790	1.4012	1.3890	
	Proposed	4.2133	1.5747	1.9786	2.2133	1.2134	<b>1.1903</b>	
ASSAY	NO	48.3423	120.2131	65.2143	203579.01321	58.2132	70.56765	
	DS	74.2112	32.7686	20.2131	34.2143	30.2141	31.0213	
	PDS	88.2413	63.1231	17.5464	39.7868	41.9886	42.5464	
	CCA	38.3413	23.9869	22.4354	34.1232	24.6879	23.8797	
	KCCA	38.1231	24.2132	22.3243	34.1243	18.9869	20.2434	
	SR	50.2131	19.9989	22.4543	28.5363	19.9121	19.8986	
	Proposed	39.3032	18.2131	22.2131	23.5655	17.9769	<b>17.1231</b>	

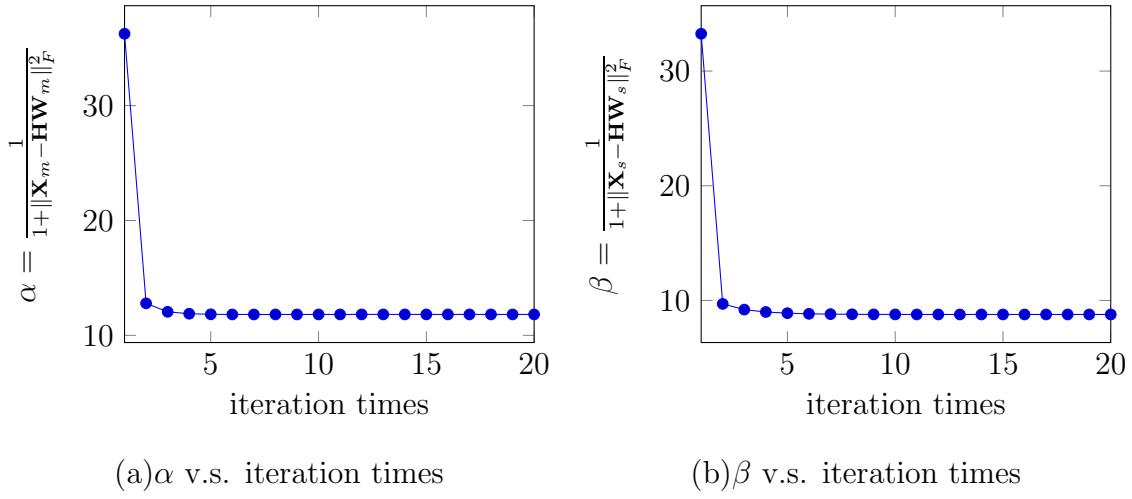


Figure 10:  $\alpha$  and  $\beta$  v.s. iteration times in tablet analysis

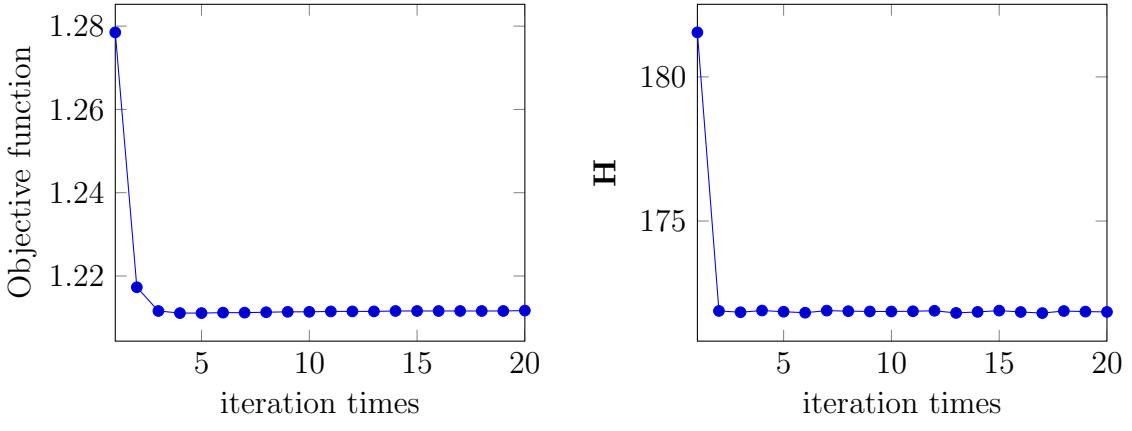
the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$ ,  $\mathbf{H}$ ,  $\alpha$  and  $\beta$  finally converged as shown in Fig. 11 and 12. It's clear that during the iteration process,  $\mathbf{HW}_m$  and  $\mathbf{HW}_s$  are approaching  $\mathbf{X}_m$  and  $\mathbf{X}_s$ . Thus the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  is decreased and converged. According to (29), in tablet analysis, the upper bound of objection is :

$$\log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) = 6.345$$

In our experiments, the inequalities hold:

$$0 \leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{H}\mathbf{W}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{H}\mathbf{W}_s\|_F^2}{c_2} \right) < 6.345$$

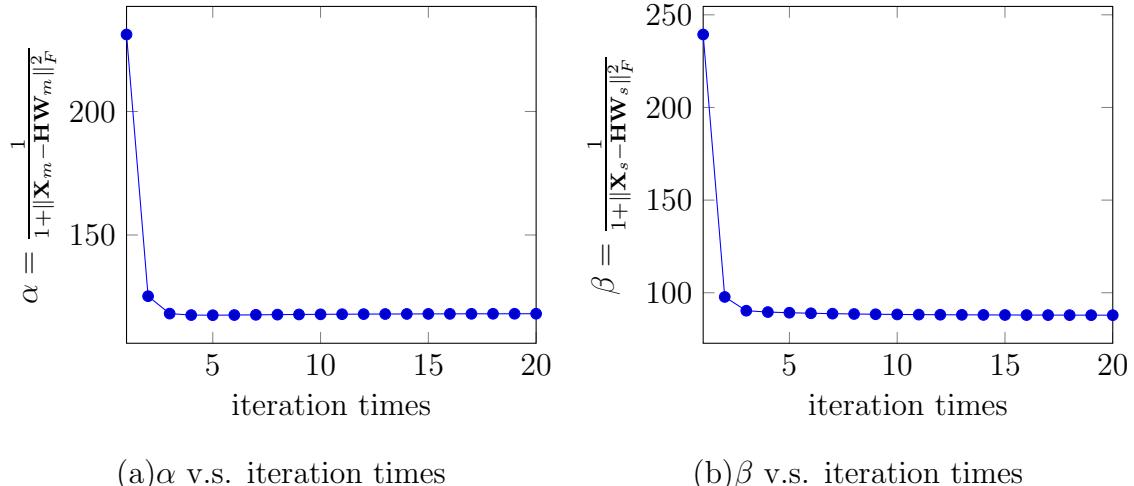
The results are listed in Table.4. The figures in bold represent the best results. As shown in Table.4, the proposed method have higher accuracy than other method on the whole.



(a) Objective function v.s. iteration times.

(b)  $\mathbf{H}$  v.s. iteration times

Figure 11:  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  and  $\mathbf{H}$  v.s. iteration times  
in sprite analysis



(a)  $\alpha$  v.s. iteration times

(b)  $\beta$  v.s. iteration times

Figure 12:  $\alpha$  and  $\beta$  v.s. iteration times in sprite analysis

Table 4: Results of spirit analysis

Response	Method	Samples	20	25	30	35	40	45
Ethanol	NO	8.6575	10.2131	5.9807	10.565	6.0311	5.5756	
	DS	6.4654	8.2311	4.6786	9.1231	7.8797	6.3242	
	PDS	6.4354	7.6766	4.5657	5.5756	7.1231	5.5657	
	CCA	4.7686	5.3435	5.2423	4.3543	3.8980	4.2342	
	KCCA	4.2334	3.9879	5.1230	5.6356	4.2423	3.9897	
	SR	5.6546	4.2131	5.3545	5.2132	3.231	4.1231	
	Proposed	4.2131	3.2131	4.7877	3.5656	3.2131	<b>3.1231</b>	
acetic ether	NO	7.3242	6.7342	4.4353	6.2342	4.0243	4.0232	
	DS	6.2342	3.1231	4.5654	3.9789	6.2423	4.2434	
	PDS	5.2423	2.2344	3.2311	2.3454	4.3242	2.7864	
	CCA	2.3242	2.3242	2.3454	3.2141	1.3432	1.6897	
	KCCA	4.4564	2.4645	1.8987	3.5657	<b>1.3242</b>	1.4325	
	SR	4.0242	1.5213	1.8342	2.2310	1.4353	1.4234	
	Proposed	6.2423	4.2131	3.3214	2.343	1.5645	2.2141	
ethyl butyrate	NO	43.5464	60.4534	34.5464	20.3235	23.3242	18.5465	
	DS	9.2423	8.4543	8.3242	7.3423	8.3534	9.3234	
	PDS	8.3423	7.2423	9.3242	6.2342	8.3443	8.2342	
	CCA	6.2342	5.2342	7.3242	7.2342	6.2342	5.3243	
	KCCA	5.3533	5.2311	4.8987	4.8987	4.5657	5.3252	
	SR	5.2423	4.8978	5.4654	5.4545	4.4344	4.3543	
	Proposed	4.3253	5.4354	4.2432	3.9778	<b>3.8977</b>	5.5457	
Ethyl lactate	NO	37.5654	36.5645	30.4456	28.8978	36.4646	30.4546	
	DS	10.2432	16.3423	15.3242	12.2432	10.2423	9.3242	
	PDS	36.3254	35.2352	30.3543	28.3543	25.6765	24.35435	
	CCA	7.4234	8.3534	7.4534	7.1657	6.32434	6.2131	
	KCCA	8.2343	8.2132	7.2343	6.8797	6.3242	5.8932	
	SR	8.3423	6.2343	6.2122	6.0123	5.8978	5.5343	
	Proposed	27.4353	8.8978	7.6241	7.0131	6.3432	<b>5.4181</b>	

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
**5. Conclusion**

12 In this paper, a robust calibration transfer model for infrared spectra  
13 is proposed. Robust objective function is employed to obtain same basis  
14 shared by master and slave spectra. Transformation matrix can be calcu-  
15 lated with the two corresponding coefficients matrices. Slave testing spectra  
16 are represented with the common basis and corresponding coefficients are  
17 then transferred using the transformation matrix. The slave testing spectra  
18 can be transferred using common basis and the corrected coefficients. The  
19 convergence and bound are also discussed. Extensive experiments are con-  
20 ducted, experimental results demonstrate that our robust calibration transfer  
21 model can generally outperform the existing methods.

22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
**Conflict of interest**

33  
34 The authors declare that they have no conflict of interest.  
35  
36

37  
38  
39  
40  
**Acknowledgment**

41 This work is supported partially by the the national natural science foun-  
42 dation of Hubei Province (No.2014CFC1131), Research and Innovation Ini-  
43 tiatives of WHPU(No.2016J06).  
44  
45

46  
47  
48  
49  
**References**

- 50  
51 [1] Harald Martens and Tormod Naes. *Multivariate calibration*. John Wiley  
52 & Sons, 1992.  
53  
54

- 1  
2  
3  
4  
5  
6  
7  
8  
9 [2] Yongdong Wang, Michael J Lysaght, and Bruce R Kowalski. Improvement  
10 of multivariate calibration through instrument standardization.  
11 *Analytical Chemistry*, 64(5):562–564, 1992.  
12  
13  
14 [3] Kenneth R Beebe and Bruce R Kowalski. An introduction to multi-  
15 variate calibration and analysis. *Analytical Chemistry*, 59(17):1007A–  
16 1017A, 1987.  
17  
18 [4] A Herrero and MC Ortiz. Multivariate calibration transfer applied to  
19 the routine polarographic determination of copper, lead, cadmium and  
20 zinc. *Analytica chimica acta*, 348(1):51–59, 1997.  
21  
22 [5] Robert N Feudale, Nathaniel A Woody, Huwei Tan, Anthony J Myles,  
23 Steven D Brown, and Joan Ferré. Transfer of multivariate calibration  
24 models: a review. *Chemometrics and Intelligent Laboratory Systems*,  
25 64(2):181–192, 2002.  
26  
27 [6] Fernanda Araújo Honorato, Roberto Kawakami Harrop Galvão,  
28 Maria Fernanda Pimentel, Benício de Barros Neto, Mário César Ugulino  
29 Araújo, and Florival Rodrigues de Carvalho. Robust modeling for mul-  
30 tivariate calibration transfer by the successive projections algorithm.  
31 *Chemometrics and intelligent laboratory systems*, 76(1):65–72, 2005.  
32  
33 [7] Graciela M Escandar, Patricia C Damiani, Héctor C Goicoechea, and  
34 Alejandro C Olivieri. A review of multivariate calibration methods ap-  
35 plied to biomedical analysis. *Microchemical journal*, 82(1):29–42, 2006.  
36  
37 [8] Yusuf Sulub, Rosario LoBrutto, Richard Vivilecchia, and Busolo Wa  
38 Wabuyele. Content uniformity determination of pharmaceutical tablets  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9 using five near-infrared reflectance spectrometers: a process analytical  
10 technology (pat) approach using robust multivariate calibration transfer  
11 algorithms. *Analytica chimica acta*, 611(2):143–150, 2008.  
12  
13  
14

- 15 [9] Claudete Fernandes Pereira, Maria Fernanda Pimentel, Roberto  
16 Kawakami Harrop Galvão, Fernanda Araújo Honorato, Luiz Stragevitch,  
17 and Marcelo Nascimento Martins. A comparative study of calibration  
18 transfer methods for determination of gasoline quality parameters in  
19 three different near infrared spectrometers. *Analytica chimica acta*,  
20 611(1):41–47, 2008.  
21  
22 [10] Anne Andrew and Tom Fearn. Transfer by orthogonal projection: making  
23 near-infrared calibrations robust to between-instrument variation.  
24 *Chemometrics and Intelligent Laboratory Systems*, 72(1):51–56, 2004.  
25  
26 [11] E Bouveresse and DL Massart. Standardisation of near-infrared spec-  
27 trometric instruments: a review. *Vibrational Spectroscopy*, 11(1):3–15,  
28 1996.  
29  
30 [12] Thomas Dean and Tomas Isaksson. Standardisation: What is it and  
31 how is it done? part 2. *NIR news*, 4(4):14–15, 1993.  
32  
33 [13] Yongdong Wang, David J Veltkamp, and Bruce R Kowalski. Multivari-  
34 ate instrument standardization. *Analytical chemistry*, 63(23):2750–2756,  
35 1991.  
36  
37 [14] Ana Herrero and M Cruz Ortiz. Piecewise direct standardization method  
38 applied to the simultaneous determination of pb (ii), sn (iv) and cd (ii)  
39 by differential pulse polarography. *Electroanalysis*, 10(10):717–721, 1998.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 1  
2  
3  
4  
5  
6  
7  
8  
9 [15] Wei Fan, Yizeng Liang, Dalin Yuan, and Jiajun Wang. Calibration  
10 model transfer for near-infrared spectra based on canonical correlation  
11 analysis. *Analytica chimica acta*, 623(1):22–29, 2008.  
12  
13  
14  
15 [16] Yong Xu, Aini Zhong, Jian Yang, and David Zhang. Lpp solution  
16 schemes for use with face recognition. *Pattern Recognition*, 43(12):4165–  
17 4176, 2010.  
18  
19  
20 [17] Sibao Chen, Haifeng Zhao, Min Kong, and Bin Luo. 2d-lpp: a two-  
21 dimensional extension of locality preserving projections. *Neurocomputing*,  
22 70(4):912–921, 2007.  
23  
24  
25 [18] Xiaofei He, Deng Cai, Shuicheng Yan, and Hong-Jiang Zhang. Neighbor-  
26 hood preserving embedding. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1208–1213.  
27 IEEE, 2005.  
28  
29  
30 [19] Jiangtao Peng, Silong Peng, An Jiang, and Jie Tan. Near-infrared cali-  
31 bration transfer based on spectral regression. *Spectrochimica Acta Part*  
32 *A: Molecular and Biomolecular Spectroscopy*, 78(4):1315–1320, 2011.  
33  
34  
35 [20] Moshe Idan and Jason L Speyer. Cauchy estimation for linear scalar  
36 systems. *Automatic Control, IEEE Transactions on*, 55(6):1329–1342,  
37 2010.  
38  
39  
40 [21] Moshe Idan and Jason L Speyer. Multivariate cauchy estimator with  
41 scalar measurement and process noises. *SIAM Journal on Control and*  
42 *Optimization*, 52(2):1108–1141, 2014.

- 1  
2  
3  
4  
5  
6  
7  
8  
9 [22] H Voss and Ulrich Eckhardt. Linear convergence of generalized  
10 weiszfeld's method. *Computing*, 25(3):243–251, 1980.  
11  
12  
13  
14 [23] Ivan Singer. *Duality for nonconvex approximation and optimization*.  
15 Springer Science & Business Media, 2007.  
16  
17  
18 [24] Chang Xu, Dacheng Tao, and Chao Xu. Multi-view intact space learn-  
19 ing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*,  
20 37(12):2531–2544, 2015.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

# Robust Calibration Model Transfer

Yi Mou<sup>\*a</sup>, Long Zhou<sup>a</sup>, Shujian Yu<sup>c</sup>, WeiZhen Chen<sup>a</sup>, Xu Zhao<sup>a</sup>, Xinge You<sup>b</sup>,

<sup>a</sup>*School of Electrical and Electronics Engineering, Wuhan Polytechnic University, Wuhan 430074, China*

<sup>b</sup>*Department of Electronics and Information Engineering, HuaZhong University of Science and Technology, Wuhan 430074, China*

<sup>c</sup>*Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA*

---

## Abstract

Calibration model transfer is an important issue in infrared spectra analysis. For identical sample, spectra collected with master and slave spectrometers share same components. In the sense of mathematics, they share same basis. If the basis and corresponding coefficient matrices can be obtained, the model transfer can be efficiently realized. On the other hand, the performance of calibration model transfer method will degrade if there are outliers and noise in samples. In this paper, a robust calibration transfer model is proposed. Cauchy estimator are employed to learn same basis shared by master and slave spectra robustly. Transformation matrix can be calculated with the two corresponding coefficients matrices. Slave testing spectra are represented with the common basis and corresponding coefficients are then transferred using the transformation matrix. The slave testing spectra can be transferred using common basis and the corrected coefficients. The convergence property

---

\*Corresponding author. Tel & Fax: +86-27-8550-4729  
Email address: mouyi@whpu.edu.cn (Yi Mou\*)

and bound of proposed model are also discussed. Extensive experiments are conducted, experimental results demonstrate that our robust calibration transfer model can generally outperform the existing methods.

*Keywords:* Infrared spectra; Model transfer; Subspace learning

---

## 1. Introduction

Multivariate calibration model transfer is the key problem in infrared spectral quantitative analysis [1–3]. Multivariate calibration techniques are commonly employed but in practice the model is invalid if an existing model is applied to spectra measured under different circumstances (temperature and humidity) or on a separate instrument[4]. Recalibration can be utilized to break the limitation, but it is expensive and time consuming. This motivates calibration transfer method which refers to the transfer of quantitative analysis model between different instruments or conditions[5, 6]. Calibration transfer plays an important role, because of the possibility of using an existing model to analyze new samples obtained in new conditions or with a new instrument without the need to build the calibration model again[7–10].

Various calibration model transfer methods have been proposed, which have been comprehensively discussed[11, 12]. These methods fall into three categories. The first category is the pre-processing methods which eliminate or decrease the differences between master and slave spectra, including baseline elimination, derivative techniques, multiplicative scatter correction (MSC) [8], FIR filtering [9], orthogonal signal correction (OSC) [10], and generalized least squares (GLS). The second is to find a transformation matrix that maps the response of the slave instrument onto the master instrument,

including Direct standardization (DS) and piecewise direct standardization (PDS). The third is based on subspace learning. The transformation matrix is built in subspace.

Direct standardization (DS) and piecewise direct standardization (PDS) method are the representative approaches of second category[13, 14]. DS uses the whole spectrum on the slave instrument to fit each spectral point on the master instrument. While in PDS, a small window from the slave spectrum is used instead of the entire spectral range.

CCA is a widely employed subspace learning tool in machine learning and pattern recognition, which is successfully applied to correct the differences between spectra measured on different instruments because of its ability to reveal the correlations between them[15]. CCA is first employed to reduce the dimensionality of master and slave spectra. The transformation matrix is calculated in lower subspace. Peng pointed out that the CCA is a linear subspace learning method. There are many subspace learning techniques, such as PCA, locality preserving projections (LPP)[16, 17], and neighborhood preserving embedding (NPE) [18]. However, the dense matrices eigenvalue decomposition in these algorithms is expensive in both time and memory, and the solution of the optimization problem is usually unstable when the number of features is larger than the number of samples. Thus spectral regression based model transfer method is proposed by Peng[19] to cast the problem of learning an embedding function into a regression framework, which avoids eigenvalue decomposition. The difference between the CCA based model transfer and Peng’s method is that the direction vectors for master and slave spectra are calculated with spectral regression.

Calibration model transfer is to find the relationship between master and slave spectra. For simplicity, master and slave spectra are obtained using same sample. According to Lambert-Beer’s Law, mixture spectrum is weighted sum of pure substance spectra. Thus, if the pure substance spectra are obtained, calibration model transfer can be implemented efficiently. Unfortunately, to obtain pure substance spectra of complex mixture is impossible. In the sense of mathematics, master and slave spectra share same basis. If the basis and corresponding coefficient matrices can be obtained, the model transfer can be efficiently realized.

To obtain high accuracy, spectra acquisition is one of the most important stage in quantitative analysis. It would be desirable to eliminate or minimize the sources of data variability that are not related to the analytical property of interest. In fact, the spectra collected with infrared spectrometer are often affected by noise and outliers. Thus, the proposed methods should be robust against the noise and outliers to obtain higher accuracy.

In machine learning and pattern recognition, Cauchy estimator is more robust than least square estimator and  $\ell_1$  estimator[20, 21]. Compared with least square estimator and  $\ell_1$  estimator, Cauchy estimator can heavily reduce the influence of large errors. Meanwhile, when an estimator is robust, it is inferred that the influence of any single observation is insufficient to yield a significant offset. Cauchy estimator has been shown to own this property. At this point, we proposed a new robust method based on Cauchy estimator to correct the spectral data.

The rest of the paper is organized as follows. We review general principles of calibration model transfer methods in Section 2. We formulate robust

calibration model transfer and provide an efficient algorithm for solving the proposed model in Section 3, where analysis to convergence analysis and bound discussion are also conducted. Experiments and result analysis are provided in Sections 4 and conclusions are drawn in Section 5.

## 2. Related work

Infrared spectroscopy is an extensively employed analytical technique in many industrial applications because of its rapidness and the fact that it is non-destructive to the samples. Multivariate calibration techniques are commonly used method to build quantitative analysis model, such as partial least squares (PLS) regression [1] and principal component regression (PCR) [2]. However, a problem occurs when an existing model is applied to spectra that were measured under new environmental conditions or on another instrument. New spectra contain variation which can lead to erroneous predictions. A possible solution to this calibration transfer problem is to measure every sample in the new instrument and construct a new model for it. However, this process would be both costly and time consuming. A more acceptable way is to apply chemometrics techniques to correct the difference of spectra measured on two instruments.

### 2.1. Classical Calibration Model

Let  $\mathbf{X}_s$  and  $\mathbf{X}_m$  be the spectral matrices obtained from master and slave instruments respectively.  $\bar{\mathbf{X}}_s$  and  $\bar{\mathbf{X}}_m$  are subsets of  $\mathbf{X}_s$  and  $\mathbf{X}_m$  respectively,  $\mathbf{C}$  and  $\bar{\mathbf{C}}$  are corresponding concentration matrices.  $\mathbf{K}_s$  and  $\mathbf{K}_m$  are the matrices of sensitivities on both instruments, each row is the pure components

spectra. The relationships between the concentration matrices and observed matrices  $\mathbf{X}_s$  and  $\mathbf{X}_m$  are:

$$\mathbf{X}_m = \mathbf{C}\mathbf{K}_m \quad (1)$$

$$\mathbf{X}_s = \mathbf{C}\mathbf{K}_s = \mathbf{C}(\mathbf{K}_m + \Delta\mathbf{K}) \quad (2)$$

where  $\Delta\mathbf{K}$  is the difference matrix between  $\mathbf{K}_s$  and  $\mathbf{K}_m$ . The same relationship is hold for the subsets.

$$\bar{\mathbf{X}}_m = \bar{\mathbf{C}}\mathbf{K}_m \quad (3)$$

$$\bar{\mathbf{X}}_s = \bar{\mathbf{C}}\mathbf{K}_s = \bar{\mathbf{C}}(\mathbf{K}_m + \Delta\mathbf{K}) \quad (4)$$

The difference matrix  $\Delta\mathbf{K}$  can be calculated by:

$$\Delta\mathbf{K} = \bar{\mathbf{C}}^+(\bar{\mathbf{X}}_s - \bar{\mathbf{X}}_m) \quad (5)$$

Then  $\mathbf{X}_s$  is estimated as

$$\hat{\mathbf{X}}_s = \mathbf{X}_m + \mathbf{C}\bar{\mathbf{C}}^+(\bar{\mathbf{X}}_s - \bar{\mathbf{X}}_m) \quad (6)$$

with  $\hat{\mathbf{X}}_s$  and  $\mathbf{C}$ , a new calibration model can be built for prediction on the slave instrument. Two assumptions are implied in this method: the linear relationship should hold on both instruments and the concentrations for all analytes contributing to the response must be known. For complex compound, where the concentrations are not known, this method is invalid.

## 2.2. Direct and Piecewise Direct Standardization

In direct standardization, response matrices on both instruments are related to each other by a transformation matrix  $\mathbf{F}$ :

$$\bar{\mathbf{X}}_m = \bar{\mathbf{X}}_s\mathbf{F} \quad (7)$$

where  $\mathbf{F}$  is calculated as:

$$\mathbf{F} = \bar{\mathbf{X}}_s^+ \bar{\mathbf{X}}_m \quad (8)$$

Lin[22] studied calibration model transfer between different temperatures. The method is based on direction standardization.

Piecewise direct standardization builds several local regression models to fit the every wavelength of the master spectra using a range of wavelengths of the slave spectra. The  $i$ th wavelength of master spectra  $\mathbf{X}_m(:, i)$  is regressed on the slave piece  $[\mathbf{X}_s(:, i - j), \dots, \mathbf{X}_s(:, j + k)]$ . According to the PLS or PCR regression model,  $\mathbf{X}_m(:, i) = [\mathbf{X}_s(:, i - j), \dots, \mathbf{X}_s(:, j + k)]b_i$ . We obtain the matrix  $\mathbf{F} = \text{diag}(b_1^T, \dots, b_p^T)$ . The linear relation between master and slave instruments is  $\mathbf{X}_m = \mathbf{X}_s \mathbf{F}$ .

### 2.3. CCA based Calibration Model Transfer

CCA based model transfer is proposed by W. Fan[15]. Canonical correlation analysis is employed to analyzed the mean centered standardization sets  $\mathbf{T}_m$  and  $\mathbf{T}_s$ . Direction matrices  $\mathbf{W}_m$  and  $\mathbf{W}_s$  are obtained and the canonical variables  $\mathbf{L}_m$  and  $\mathbf{L}_s$  are obtained by projecting the standardization sets on to the direction matrices. The transform matrix  $\mathbf{F}_1$  and  $\mathbf{F}_2$  are determined by  $\mathbf{F}_1 = \mathbf{L}_s^+ \mathbf{L}_m$  and  $\mathbf{F}_2 = \mathbf{L}_m^+ \mathbf{T}_m$ . The transferred prediction set  $\mathbf{Z}_s$  is  $\mathbf{Z}_s = \mathbf{K}_s \mathbf{F}_1 \mathbf{F}_2$  where  $\mathbf{K}_s = \mathbf{P}_s \mathbf{W}_s$ .

### 2.4. Spectral regression based Calibration Model Transfer

The samples are divided into two parts: calibration and prediction set. The standardization set  $\mathbf{X}_{ms}$  is selected among the calibration set of master instrument using subset selection method, and the same subset on the slave

instrument gives rise to standardization set  $\mathbf{X}_{ss}$ . Spectral regression is carried out to  $\mathbf{X}_{ms}$  and  $\mathbf{X}_{ss}$  respectively. The transformation matrix  $\mathbf{U}$  and  $\mathbf{V}$  map the standardization sets to two low-dimensional subspaces, and two subspace variables  $\mathbf{L}_m$  and  $\mathbf{L}_s$  are obtained. Two transformation matrices  $\mathbf{F}_1$  and  $\mathbf{F}_2$  are obtained  $\mathbf{F}_1 = \mathbf{L}_s^+ \mathbf{L}_m$  and  $\mathbf{F}_2 = \mathbf{L}_m^+ \mathbf{X}_{ms}$ . The transferred prediction set  $\mathbf{Z}_s$  is  $\mathbf{Z} = \mathbf{K}_s \mathbf{F}_1 \mathbf{F}_2$  where  $\mathbf{K}_s = \mathbf{P}_s \mathbf{W}_s$ . It's clear that the difference between CCA based calibration model transfer and spectral regression based method is the way of dimension reduction.

### 3. Calibration model transfer based on subspace learning

For identical sample, spectra collected with master and slave spectrometers share same components. In the sense of mathematics, they share same basis. The differences between them are due to the representation coefficients, as shown in Fig.1. If the same basis and corresponding coefficients are obtained, the model transfer can be efficiently realized. On the other hand, the performance of calibration model transfer method will degraded if there are outliers and noise in samples. In this paper, a robust calibration transfer model is proposed.

#### 3.1. Notation

Samples collected on master and slave instruments are defined as  $\mathbf{X}_m \in R^{n \times p}$  and  $\mathbf{X}_s \in R^{n \times p}$  respectively, where  $n$  is the number of samples and  $p$  is the dimension of data.

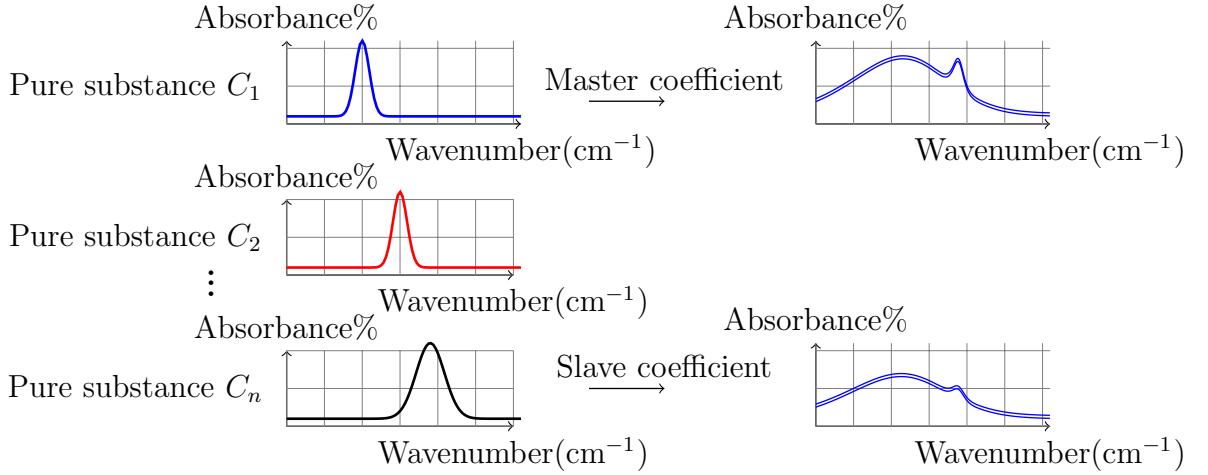


Figure 1: Difference between master and slave spectra

### 3.2. Robust estimator

M-estimator is popular in robust statistics. Let  $r_i$  be the residual of the  $i$ -th data point, i.e., the difference between the  $i$ -th observation and its fitted values. The standard least squares method tries to minimize  $\sum_i r_i^2$ , which is unstable if outliers are present, and which has a strong effect to distort the estimated parameters. M-estimators try to reduce the effect of outliers by replacing the squared residuals  $r_i^2$  with another function of residuals

$$\min \sum_i \rho(r_i) \quad (9)$$

where  $\rho$  is a symmetric, positive-definite function with a unique minimum at zero and chosen to be less increasing than the square function. The corresponding influence function is defined as:

$$\psi(x) = \frac{\partial \rho(x)}{x} \quad (10)$$

which measures the influence of a random data point on the value of the parameter estimate.

As shown in Fig.2, for the  $\ell_2$  estimator (least-squares) with  $\rho(x) = \frac{x^2}{2}$ , the influence function is  $\psi(x) = x$ ; that is, the influence of a data point on the estimate increases linearly with the size of its error. This confirms the non-robustness of the least-squares estimate. Although the  $\ell_1$  (absolute value) estimator with  $\rho(x) = |x|$  reduces the influence of large errors, its influence function has no cut-off. When an estimator is robust, it is inferred that the influence of any single observation is insufficient to yield a significant offset. Cauchy estimator has been shown to own this valuable property

$$\rho(x) = \log\left(1 + \left(\frac{x}{c}\right)^2\right) \quad (11)$$

along with the upper bounded influence function:

$$\psi(x) = \frac{2x}{c^2 + x^2} \quad (12)$$

Therefore, we deploy the Cauchy estimator in PLS.

### 3.3. Robust model transfer

Samples collected from different instruments own same basis. Linear representation coefficient leads to the difference between  $\mathbf{X}_m$  and  $\mathbf{X}_s$ . We can obtain their common basis  $\mathbf{H}$ , and corresponding represent coefficients  $\mathbf{W}_m$  and  $\mathbf{W}_s$ . Transformation matrix  $\mathbf{F}$  can be calculated as  $\mathbf{F} = \mathbf{W}_s^+ \mathbf{W}_m$ . The basis  $\mathbf{H}$  is then utilized to represent the slave test sets  $\mathbf{X}_{ts}$ , and  $\mathbf{W}_{ts}$  is corresponding coefficient. The slave testing sample can be corrected with transformation matrix and basis. The flow chart of proposed algorithm is shown in Fig.3.

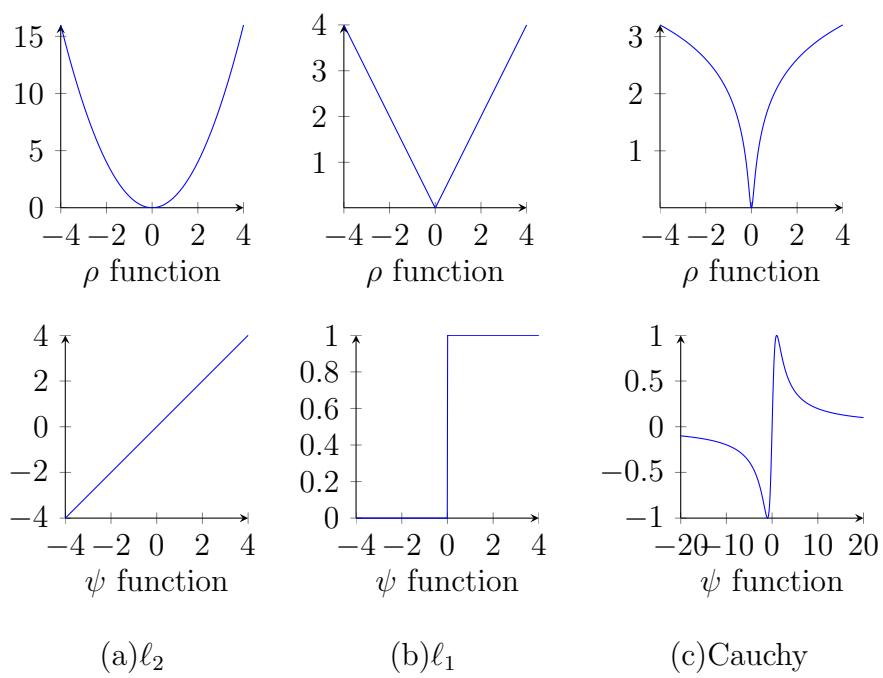


Figure 2:  $\rho$  and  $\psi$  function

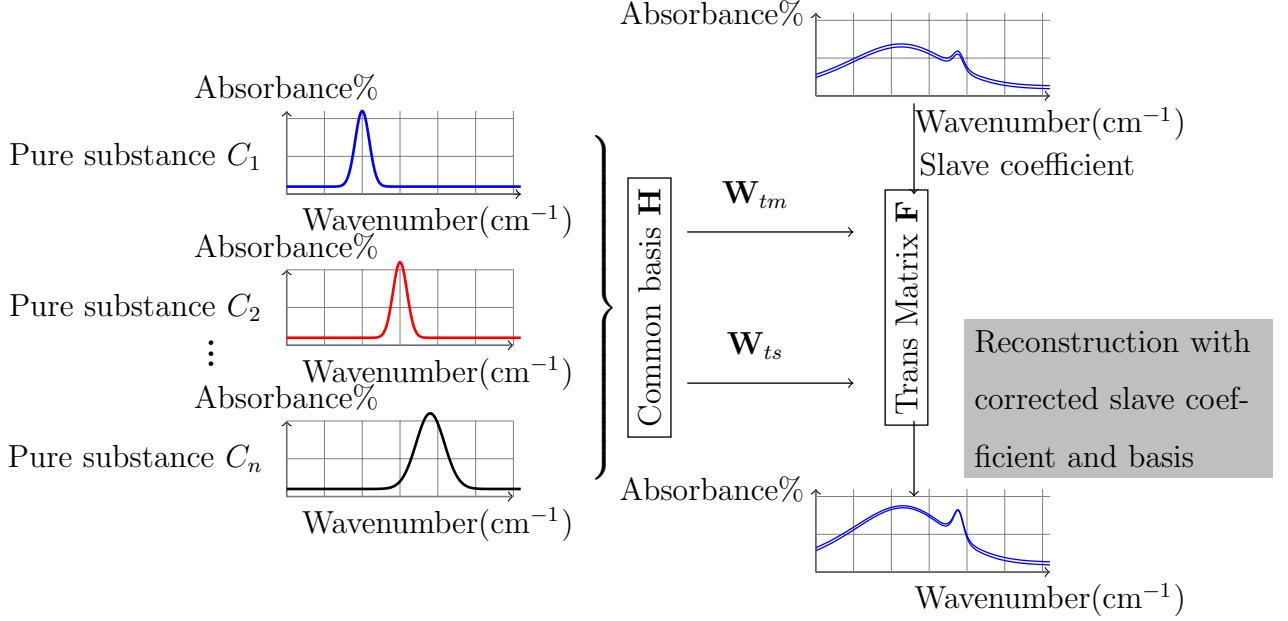


Figure 3: Flow of proposed method

In learning procedure, the basis  $\mathbf{H}$  satisfies the model:

$$\begin{aligned} \min_{\mathbf{H}, \mathbf{W}_m, \mathbf{W}_s} & \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) \\ s.t. & \mathbf{H}^T \mathbf{H} = \mathbf{I} \end{aligned} \quad (13)$$

where  $\mathbf{H} \in R^{n \times m}$ ,  $\mathbf{W}_m \in R^{m \times p}$  and  $\mathbf{W}_s \in R^{m \times p}$  represent the shared basis and corresponding coefficients. The term  $\log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right)$  and  $\log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right)$  measure the reconstruction error,  $c_1 > 0$  and  $c_2 > 0$  are two scale parameters.

The Lagrange function is:

$$\mathcal{L} = \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) + \frac{1}{2} \lambda (\mathbf{I} - \mathbf{H}^T \mathbf{H}) \quad (14)$$

(1) If  $\mathbf{W}_m$ ,  $\mathbf{W}_s$  are fixed,

$$\frac{\partial \mathcal{L}}{\partial \mathbf{H}} = \frac{(\mathbf{X}_m - \mathbf{HW}_m)\mathbf{W}_m^T}{c_1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2} + \frac{(\mathbf{X}_s - \mathbf{HW}_s)\mathbf{W}_s^T}{c_2 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2} - \lambda \mathbf{H} = 0 \quad (15)$$

Define:

$$\begin{aligned}\alpha &= \frac{1}{c_1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2} \\ \beta &= \frac{1}{c_2 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}\end{aligned}$$

Then

$$\frac{\partial \mathcal{L}}{\partial \mathbf{H}} = \alpha(\mathbf{X}_m - \mathbf{HW}_m)\mathbf{W}_m^T + \beta(\mathbf{X}_s - \mathbf{HW}_s)\mathbf{W}_s^T - \lambda\mathbf{H} = 0$$

$$\alpha\mathbf{X}_m\mathbf{W}_m^T + \beta\mathbf{X}_s\mathbf{W}_s^T = \mathbf{H}(\alpha\mathbf{W}_m\mathbf{W}_m^T + \beta\mathbf{W}_s\mathbf{W}_s^T + \lambda)$$

$$(\alpha\mathbf{X}_m\mathbf{W}_m^T + \beta\mathbf{X}_s\mathbf{W}_s^T)(\alpha\mathbf{W}_m\mathbf{W}_m^T + \beta\mathbf{W}_s\mathbf{W}_s^T + \lambda)^{-1} = \mathbf{H} \quad (16)$$

(2) If the shared basis  $\mathbf{H}$  is obtained,

$$\mathbf{W}_m = \mathbf{H}^+\mathbf{X}_m \quad (17)$$

$$\mathbf{W}_s = \mathbf{H}^+\mathbf{X}_s \quad (18)$$

Transformation matrix  $\mathbf{F}$  is:

$$\mathbf{F} = \mathbf{W}_m\mathbf{W}_s^+ \quad (19)$$

### 3.4. Convergence analysis

For simplicity, we consider  $\mathbf{h} \in R^{n \times 1}, \mathbf{w} \in R^{1 \times p}$ . After having found  $\mathbf{h}^{(k)}$ , we can construct a quadratic function  $\psi(\mathbf{h}, \mathbf{h}^{(k)})$  to upper bound  $\mathcal{L}(\mathbf{h})$  such that the following conditions hold[23–25]:

$$\psi(\mathbf{h}^{(k)}, \mathbf{h}^{(k)}) = \mathcal{L}(\mathbf{h}^{(k)}) \quad (20)$$

$$\psi'(\mathbf{h}^{(k)}, \mathbf{h}^{(k)}) = \mathcal{L}'(\mathbf{h}^{(k)}) \quad (21)$$

---

**Algorithm 1** Robust model transfer algorithm

---

**Input:** Master and slave spectral matrices  $\mathbf{X}_m$  and  $\mathbf{X}_s$  for training,  $\mathbf{X}_{tm}$  and  $\mathbf{X}_{ts}$  for testing, initial  $\lambda$ ,  $c_1 = c_2 = 1$ ,  $\alpha = \alpha^0$ ,  $\beta = \beta^0$ ,  $\mathbf{W}_m = \mathbf{W}_m^0$ ,  $\mathbf{W}_s = \mathbf{W}_s^0$ , Set the iteration times  $k$

- 1:  $\mathbf{H}^k = (\alpha^{k-1}\mathbf{X}_m(\mathbf{W}_m^{k-1})^T + \beta^{k-1}\mathbf{X}_s(\mathbf{W}_s^{k-1})^T)(\alpha^{k-1}\mathbf{W}_m^{k-1}(\mathbf{W}_m^{k-1})^T + \beta^{k-1}\mathbf{W}_s^{k-1}(\mathbf{W}_s^{k-1})^T + \lambda)^{-1};$
  - 2:  $\mathbf{W}_m^k = (\mathbf{H}^k)^+\mathbf{X}_m$ ,  $\mathbf{W}_s^k = (\mathbf{H}^k)^+\mathbf{X}_s$ ;
  - 3:  $\alpha^k = \frac{1}{c_1 + \|\mathbf{X}_m - \mathbf{H}^k \mathbf{W}_m^k\|_F^2}$ ,  $\beta^k = \frac{1}{c_2 + \|\mathbf{X}_s - \mathbf{H}^k \mathbf{W}_s^k\|_F^2}$ ;
  - 4:  $k = k + 1$ ;
  - 5: Transformation matrix  $\mathbf{F}$  is:  $\mathbf{F} = \mathbf{W}_m \mathbf{W}_s^+$ .
  - 6: Represent  $\mathbf{X}_{ts}$  with obtained  $\mathbf{H}$  and the corresponding coefficient is  $\mathbf{W}_{ts}$
  - 7: Transferred slave testing spectra matrix is:  $\mathbf{X}_{ts}^* = \mathbf{H} \mathbf{F} \mathbf{W}_{ts}$
- 

The function  $\psi(\mathbf{h}^{(k)}, \mathbf{h}^{(k)})$  is:

$$\begin{aligned} \psi(\mathbf{h}, \mathbf{h}^{(k)}) &= \mathcal{L}(\mathbf{h}^{(k)}) + (\mathbf{h} - \mathbf{h}^{(k)})^T \mathcal{L}'(\mathbf{h}^{(k)}) \\ &+ (\mathbf{h} - \mathbf{h}^{(k)})^T (\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda) (\mathbf{h} - \mathbf{h}^{(k)}) \end{aligned} \quad (22)$$

Assume that  $\psi(\mathbf{h}, \mathbf{h}^{(k)})$  is locally convex with respect to  $\mathbf{h}$  and has a local minimizer. Setting  $\mathbf{h}^{k+1}$  as this minimizer, then

$$\psi'(\mathbf{h}^{(k+1)}, \mathbf{h}^{(k)}) = \mathcal{L}'(\mathbf{h}^{(k)}) + 2(\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda)(\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}) = 0 \quad (23)$$

By appropriately choosing  $\mathbf{h}^{k+1}$  near  $\mathbf{h}$ , we have  $\mathcal{L}(\mathbf{h}) \leq \psi(\mathbf{h}, \mathbf{h}^{(k)})$ , which means that

$$\begin{aligned} \mathcal{L}(\mathbf{h}^{(k+1)}) &\leq \psi(\mathbf{h}^{(k+1)}, \mathbf{h}^{(k)}) \\ &= \mathcal{L}(\mathbf{h}^{(k)}) + (\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)})^T \mathcal{L}'(\mathbf{h}^{(k)}) \\ &+ (\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)})^T (\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda) (\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}) \end{aligned} \quad (24)$$

By combining Eqs. (21) and (22), we have:

$$\begin{aligned}\mathcal{L}(\mathbf{h}^{(k+1)}) - \mathcal{L}(\mathbf{h}^{(k)}) &\leq -(\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda)(\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)})^T (\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}) \\ &\leq -(\alpha^k \mathbf{w}_m^k (\mathbf{w}_m^k)^T + \beta^k \mathbf{w}_s^k (\mathbf{w}_s^k)^T + \lambda) \|\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}\|_F^2 \leq 0\end{aligned}\quad (25)$$

### 3.5. Bound discuss

The proposed model is represented as (26). We noticed that function  $\log(1 + x^2)(x > 0)$  is monotonously increased function, shown as Fig.4:

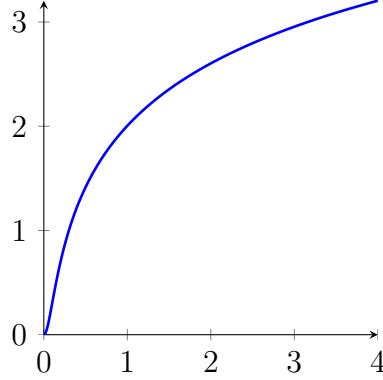


Figure 4:  $\log(1 + x^2)$

The object function can be expressed as:

$$\begin{aligned}\min \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2 \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_1 c_2} \right) \\ s.t. \mathbf{H}^T \mathbf{H} = \mathbf{I}\end{aligned}\quad (26)$$

The inequality holds:

$$\frac{\left( \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right)^2}{4} \geq \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2 \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_1 c_2} \quad (27)$$

Then we have:

$$\begin{aligned}
& \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2 \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_1 c_2} \\
\leq & \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\left( \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right)^2}{4} \quad (28)
\end{aligned}$$

We notice that:

$$\begin{aligned}
\frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} &= \left\| \begin{pmatrix} \frac{1}{\sqrt{c_1}} \mathbf{X}_m \\ \frac{1}{\sqrt{c_2}} \mathbf{X}_s \end{pmatrix} - (\mathbf{H}, \mathbf{H}) \begin{pmatrix} \frac{1}{\sqrt{c_1}} \mathbf{W}_m \\ \frac{1}{\sqrt{c_2}} \mathbf{W}_s \end{pmatrix} \right\|_F^2 \\
&= \|\mathbf{X} - \mathbf{HW}\|_F^2 \quad (29)
\end{aligned}$$

where  $\mathbf{X} = \begin{pmatrix} \frac{1}{\sqrt{c_1}} \mathbf{X}_m \\ \frac{1}{\sqrt{c_2}} \mathbf{X}_s \end{pmatrix}$ ,  $\mathbf{H} = (\mathbf{H}, \mathbf{H})$  and  $\mathbf{W} = \begin{pmatrix} \frac{1}{\sqrt{c_1}} \mathbf{W}_m \\ \frac{1}{\sqrt{c_2}} \mathbf{W}_s \end{pmatrix}$ .

$$\begin{aligned}
& \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2 \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_1 c_2} \\
\leq & \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} + \frac{\left( \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right)^2}{4} \\
\leq & \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \quad (30)
\end{aligned}$$

Then we have:

$$\begin{aligned}
0 &\leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) \\
&\leq \log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) \quad (31)
\end{aligned}$$

where  $\sigma$  is the singular value of the matrix  $\mathbf{X}$ ,  $c_1 = c_2 = 1$ .

## 4. Experiments and discussion

In order to evaluate the performance of proposed method, three data sets are used and five benchmark methods are re-implemented,i.e., PDS, DS, CCA, KCCA, SR Square Error of Prediction(RMSEP) to measure the prediction precision which is defined as:

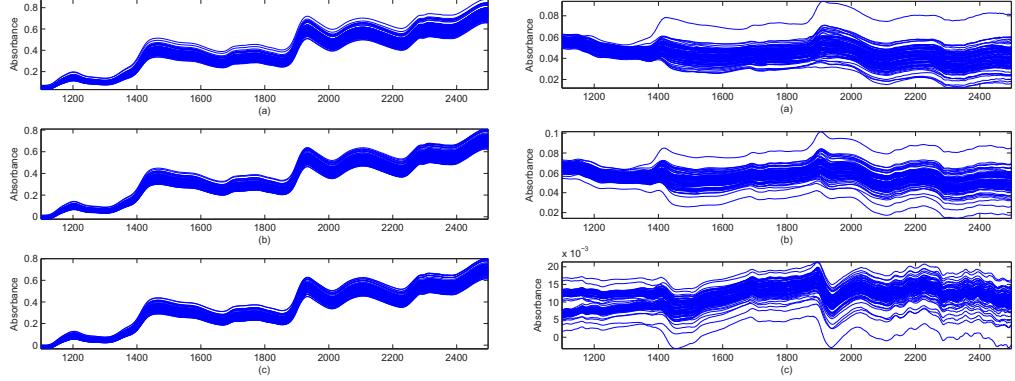
$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (32)$$

All experiments are implemented using Matlab and run in a personal computer with a 2.80GHz Intel Core i5-2300 CPU, 4GB RAM, and Windows 7 operation system. We set the upper limit of number of latent components to be 10. For each latent variable, the experiment is repeated 50 times independently and average RMSEP is calculated.

### 4.1. Corn spectra analysis

This data set consists of 80 samples of corn measured on 3 different NIR spectrometers. The wavelength range is 1100-2498nm at 2 nm intervals (700 channels), as shown in Fig.5. The moisture, oil, protein and starch values for each of the samples is also included.

Two experiments are conducted with this set. M5 is selected as master instrument, while MP5 and MP6 are selected as slave instruments, respectively. The number of training samples are 15,20,25,30,35,40. The default number of the basis is fixed as 10. Fig.6 and 7 are the convergence analysis results of corn spectral model transfer analysis. During the iteration process of the proposed algorithm, the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$ ,  $\mathbf{H}$ ,  $\alpha$  and  $\beta$  finally converged



(a) Corn spectra. (a) m5; (b) mp5; (c) mp6 (b) Spectral differences. (a) m5-mp5; (b) mp5-mp6; (c) mp5-mp6

Figure 5: Corn spectra and differences

as shown in Fig. 6 and 7. It's clear that during the iteration process,  $\mathbf{HW}_m$  and  $\mathbf{HW}_s$  are approaching  $\mathbf{X}_m$  and  $\mathbf{X}_s$ . Thus the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  is decreased and converged. According to Eq.(31), in corn analysis, the upper bound of objection is :

$$\begin{aligned} \text{m5 is master and mp5 is slave: } & \log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) = 0.0050 \\ \text{m5 is master and mp6 is slave: } & \log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) = 0.0054 \end{aligned}$$

In our experiments, the inequalities hold:

$$\begin{aligned} 0 \leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) < 0.0050 \\ 0 \leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) < 0.0054 \end{aligned}$$

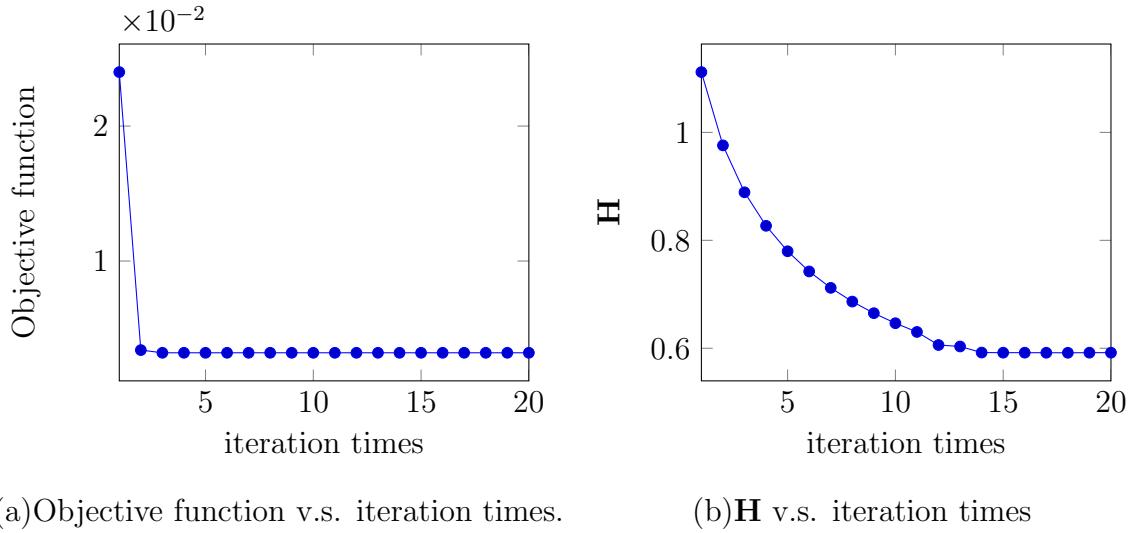


Figure 6:  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  and  $\mathbf{H}$  v.s. iteration times in corn spectral analysis

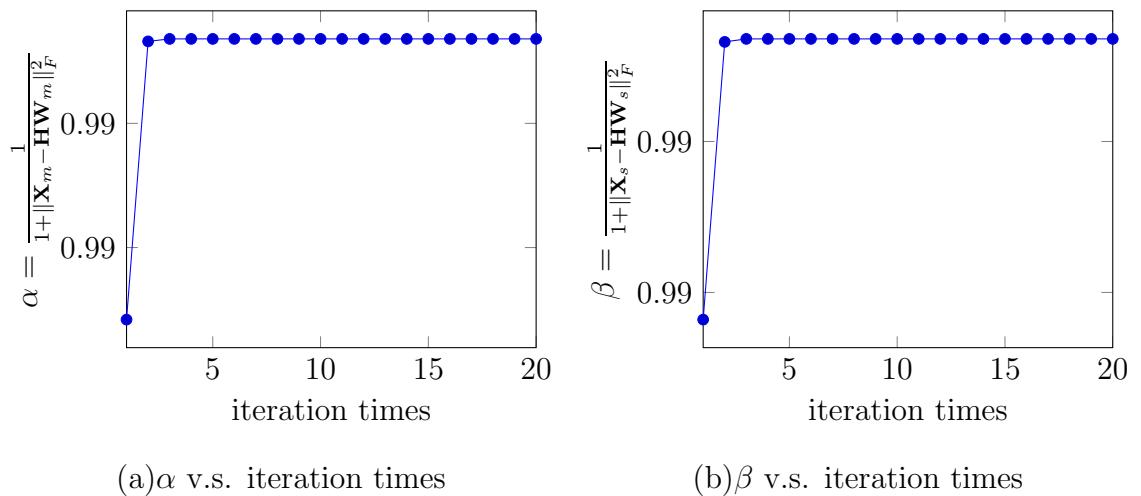


Figure 7:  $\alpha$  and  $\beta$  v.s. iteration times in corn spectral analysis

The results are listed in Table 1 and Table 2. The figures in bold represent the best results. As shown in Table 1 and 2, the proposed method have higher accuracy than other method on the whole.

#### 4.2. Tablet spectra analysis

Data set 2 is a public source data for calibration transfer from the IDRC Shootout 2002 [18,19]. Spectra of pharmaceutical tablets from two Multi-tab spectrometers (FossNIRsystems, Silverspring, MD) are measured in the transmittance mode. Tablets data from two instruments have been split into two calibration sets (155 tablets, Calibrate 1 and Calibrate 2) and two test sets (460 tablets, Test 1 and Test 2). All log (1/T) spectra cover the region from 600 to 1898 nm in 2 nm increments. The data set includes tablets with a wide ASSAY range, 152-239 mg. The spectra and the difference are shown as Fig.8.

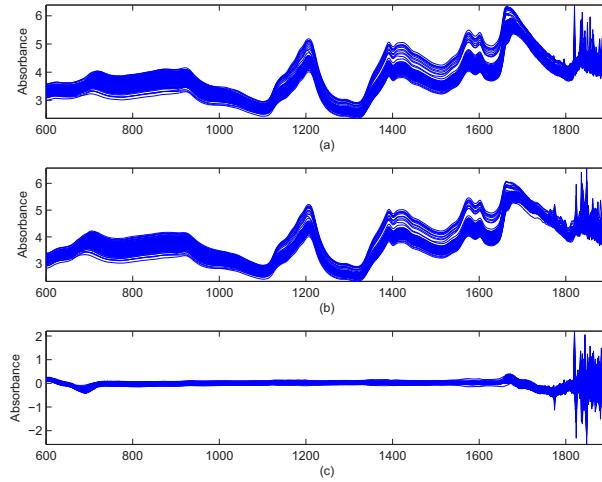


Figure 8: Tablets spectra. Spectra of 50 pharmaceutical tablets on instrument 1; (b) spectra of 50 pharmaceutical tablets on instrument 2; (c) spectral differences between the two instruments.

Table 1: Result: m5 is master, mp5 is slave

Response	Method \ Samples	15	20	25	30	35	40
Water	NO	2.4218	2.8003	1.6277	1.5132	0.5088	0.3726
	PDS	15.1419	6.1223	3.7922	3.7577	2.8557	1.7235
	DS	0.4387	0.4456	0.5635	0.5534	0.4898	0.4235
	CCA	0.8909	0.7772	0.8308	0.7272	0.5253	0.4213
	KCCA	0.8308	0.8209	1.3517	0.9640	0.3831	0.7299
	SR	1.3156	1.1146	0.5313	0.4255	0.3768	0.3799
	Proposed	1.2998	0.8905	0.5308	0.3904	0.3371	<b>0.3297</b>
Oil	NO	2.3112	0.2118	0.3478	0.3567	0.3236	0.7346
	PDS	6.6020	3.2630	2.2290	0.3112	0.2134	0.5998
	DS	0.5324	0.4434	0.3678	0.2346	0.1975	0.2389
	CCA	0.3390	0.2701	0.2370	0.2670	0.1790	0.2601
	KCCA	0.3599	0.2834	0.1911	0.2698	0.1690	0.2793
	SR	0.3410	0.3561	0.2710	0.1998	0.2197	<b>0.1621</b>
	Proposed	0.3591	0.3361	0.2498	0.1997	0.1727	0.1658
Protein	NO	1.8174	3.3435	6.4569	2.5678	1.2348	1.3654
	PDS	2.7890	9.5678	4.3467	7.3257	5.4556	4.6012
	DS	0.4315	0.5790	0.5362	0.5137	0.5314	0.4290
	CCA	0.7457	0.9156	0.6457	0.6465	0.6343	0.4148
	KCCA	0.7855	0.9545	0.6143	0.6458	0.5557	0.4665
	SR	0.7789	0.7783	0.6836	0.6012	0.5024	0.4454
	Proposed	1.0345	0.8904	0.8723	0.5244	0.4903	<b>0.4012</b>
Starch	No	3.2123	2.3432	3.4356	1.2123	1.8904	1.5235
	PDS	16.1278	16.0324	12.1244	4.4636	5.23524	4.4078
	DS	0.9823	1.3256	1.0349	0.9568	1.1890	0.9513
	CCA	1.7601	2.2223	1.5219	1.1908	1.4023	1.0123
	KCCA	1.6690	2.2794	1.6001	1.2101	1.4123	0.8901
	SR	1.6698	2.2526	1.6572	1.0501	0.9810	0.8553
	Proposed	1.4978	2.1126	1.6123	0.9546	0.9256	<b>0.8443</b>

Table 2: Results: m5 is master, mp6 is slave

Response	Samples method	15	20	25	30	35	40
Water	No	6.1435	5.1906	3.2563	1.6789	1.8345	1.9101
	PDS	8.5698	7.3290	12.0116	5.6571	5.3909	3.9562
	DS	0.5437	0.3048	0.4547	0.4325	0.4821	0.3790
	CCA	1.1412	0.4987	0.5909	0.6019	0.4890	0.4014
	KCCA	1.1909	0.7802	0.7279	0.4015	0.3801	0.3210
	SR	0.7012	0.3669	0.4998	0.3992	0.3621	0.3225
	Proposed	1.2120	0.3612	0.5812	0.4230	0.3561	<b>0.3105</b>
Oil	NO	2.2563	1.9235	1.6312	1.1568	0.6378	0.7187
	PDS	5.0231	0.9021	2.7352	2.2134	1.7234	0.7123
	DS	0.2410	0.2211	0.1998	0.1940	0.1825	0.1730
	CCA	0.2912	0.3543	0.2271	0.2478	0.2123	0.2012
	KCCA	0.5234	0.2812	0.2524	0.2378	0.2344	0.1956
	SR	0.2623	0.2680	0.1923	0.1846	0.1804	0.1725
	Proposed	0.2742	0.2612	0.1712	0.1856	<b>0.1520</b>	0.1629
Protein	NO	9.2131	7.3245	6.3490	1.6324	3.1096	4.3245
	PDS	9.2352	15.9851	7.3271	2.5487	1.5878	1.8901
	DS	0.7201	0.5989	0.5870	0.5023	0.5324	0.5339
	CCA	0.8781	0.9245	0.5675	0.5123	0.6436	0.5501
	KCCA	0.8790	0.8330	0.5810	0.4636	0.6309	0.5871
	SR	0.9310	1.0789	0.8234	0.5660	0.5478	0.5340
	Proposed	1.5012	1.0787	0.8234	0.4279	<b>0.4014</b>	0.5001
Starch	NO	10.4366	6.1341	3.2145	2.4367	6.6578	7.54745
	PDS	20.6576	14.8798	14.3453	10.3243	8.4353	7.7234
	DS	0.8767	1.1237	0.9126	0.8787	1.0553	0.9673
	CCA	1.3998	1.8790	1.3289	1.2016	1.1203	0.9494
	KCCA	1.4490	2.1023	1.4343	1.1590	1.1413	0.8445
	SR	1.4567	1.4312	1.4023	1.1870	1.058	0.9340
	Proposed	2.6787	1.5235	1.2789	0.9413	0.8713	<b>0.8234</b>

Calibrate1 is selected as master instrument, while calibrate2 is selected as slave instruments. The number of training samples are 15,20,25,30,35,40. The default number of the basis is fixed as 10. Fig.9 and 10 are the convergence analysis results of tablet spectral model transfer analysis. During the iteration process of the proposed algorithm, the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$ ,  $\mathbf{H}$ ,  $\alpha$  and  $\beta$  finally converged as shown in Fig. 9 and 10. It's clear that during the iteration process,  $\mathbf{HW}_m$  and  $\mathbf{HW}_s$  are approaching  $\mathbf{X}_m$  and  $\mathbf{X}_s$ . Thus the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  is decreased and converged. According to (31), in tablet analysis, the upper bound of objection is :

$$\log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) = 8.9455$$

In our experiments, the inequalities hold:

$$0 \leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) < 8.9455$$

The results are listed in Table.3. The figures in bold represent the best results. As shown in Table.3, the proposed method have higher accuracy than other method on the whole.

#### 4.3. Spirit analysis

We have collected spirit spectrum with two Nicolet spectrometers, and the 17 components reference concentrations values are tested with Agilent HPLC. In our experiments, 4 components are analysized. 100 spectra are collected with instruments A and B, respectively. A is selected as master

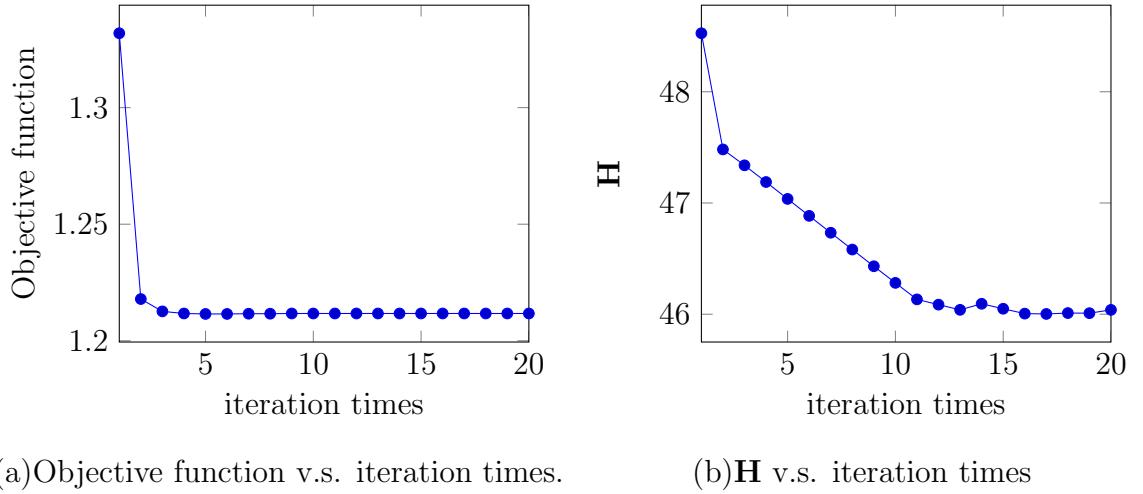


Figure 9:  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  and  $\mathbf{H}$  v.s. iteration times in tablet analysis

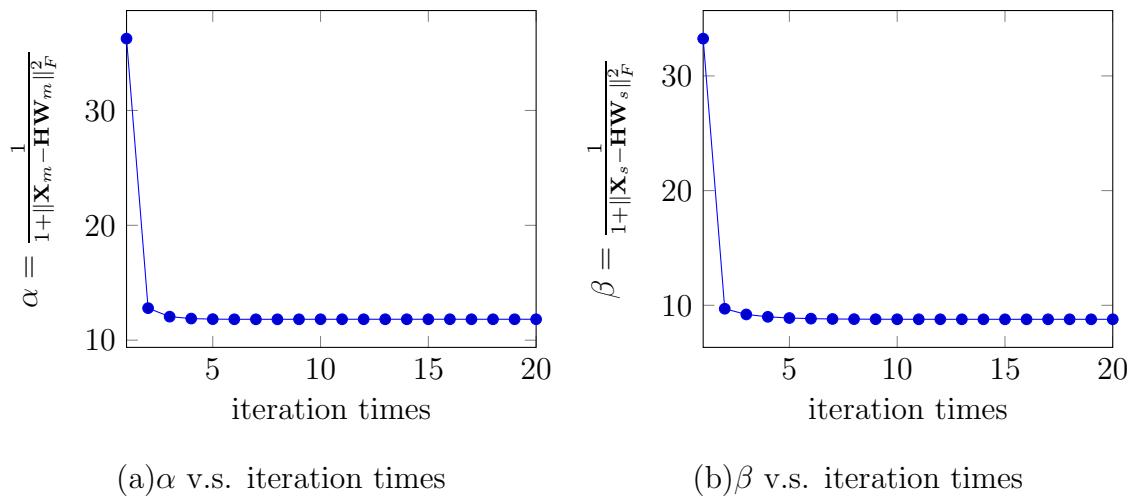


Figure 10:  $\alpha$  and  $\beta$  v.s. iteration times in tablet analysis

Table 3: Results of tablet analysis

Response	Method \ Samples	20	25	30	35	40	45
Weight	NO	20.4543	38.4351	8.2132	43635.2314	19.2147	19.2343
	DS	9.4355	10.2345	5.4567	9.8341	8.2344	8.5463
	PDS	15.4564	17.7967	6.5745	15.3413	11.1212	11.6343
	CCA	7.7896	7.8341	6.2345	10.5357	6.7896	6.7695
	KCCA	7.3141	7.2356	6.0124	8.5674	7.2354	6.7854
	SR	6.4563	6.4643	6.1032	7.3256	5.6623	5.6467
	Proposed	7.4363	5.3140	5.6573	6.0196	4.6425	<b>4.6123</b>
Hardness	NO	5.2131	5.7858	2.3141	1127.7867	4.0231	4.0213
	DS	5.2735	2.9243	2.3579	3.9425	2.1241	1.9945
	PDS	7.2132	4.2342	2.1234	4.6878	2.6342	2.7653
	CCA	3.8962	2.2143	2.1231	3.7689	1.6546	1.7234
	KCCA	3.8578	2.5742	2.1234	3.5778	1.5246	1.2231
	SR	3.8906	1.7454	1.8986	2.1790	1.4012	1.3890
	Proposed	4.2133	1.5747	1.9786	2.2133	1.2134	<b>1.1903</b>
ASSAY	NO	48.3423	120.2131	65.2143	203579.01321	58.2132	70.56765
	DS	74.2112	32.7686	20.2131	34.2143	30.2141	31.0213
	PDS	88.2413	63.1231	17.5464	39.7868	41.9886	42.5464
	CCA	38.3413	23.9869	22.4354	34.1232	24.6879	23.8797
	KCCA	38.1231	24.2132	22.3243	34.1243	18.9869	20.2434
	SR	50.2131	19.9989	22.4543	28.5363	19.9121	19.8986
	Proposed	39.3032	18.2131	22.2131	23.5655	17.9769	<b>17.1231</b>

instrument, while B is selected as slave instruments. The number of training samples are 15,20,25,30,35,40. The default number of the basis is fixed as 50. Fig.11 and 12 are the convergence analysis results of corn spectral model transfer analysis. During the iteration process of the proposed algorithm, the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$ ,  $\mathbf{H}$ ,  $\alpha$  and  $\beta$  finally converged as shown in Fig. 11 and 12. It's clear that during the iteration process,  $\mathbf{HW}_m$  and  $\mathbf{HW}_s$  are approaching  $\mathbf{X}_m$  and  $\mathbf{X}_s$ . Thus the objective function  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  is decreased and converged. According to (31), in tablet analysis, the upper bound of objection is :

$$\log \left( 1 + \sum_{i=m+1}^{2n} \sigma_i^2 + 0.25 \left( \sum_{i=m+1}^{2n} \sigma_i^2 \right)^2 \right) = 6.345$$

In our experiments, the inequalities hold:

$$0 \leq \log \left( 1 + \frac{\|\mathbf{X}_m - \mathbf{HW}_m\|_F^2}{c_1} \right) + \log \left( 1 + \frac{\|\mathbf{X}_s - \mathbf{HW}_s\|_F^2}{c_2} \right) < 6.345$$

The results are listed in Table.4. The figures in bold represent the best results. As shown in Table.4, the proposed method have higher accuracy than other method on the whole.

## 5. Conclusion

In this paper, a robust calibration transfer model for infrared spectra is proposed. Robust objective function is employed to obtain same basis shared by master and slave spectra. Transformation matrix can be calculated with the two corresponding coefficients matrices. Slave testing spectra

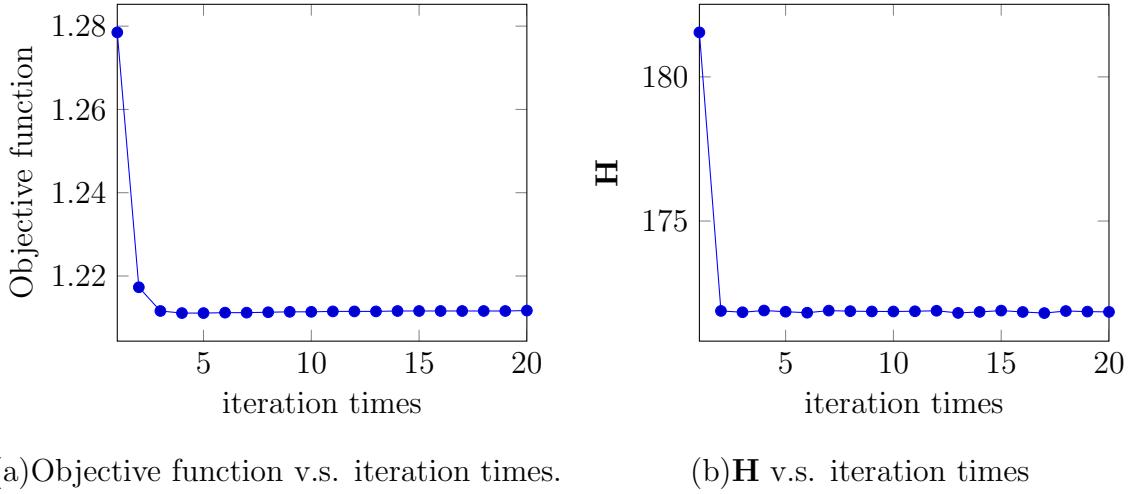


Figure 11:  $\log(1 + \|\mathbf{X}_m - \mathbf{HW}_m\|_F^2) + \log(1 + \|\mathbf{X}_s - \mathbf{HW}_s\|_F^2)$  and  $\mathbf{H}$  v.s. iteration times in sprite analysis

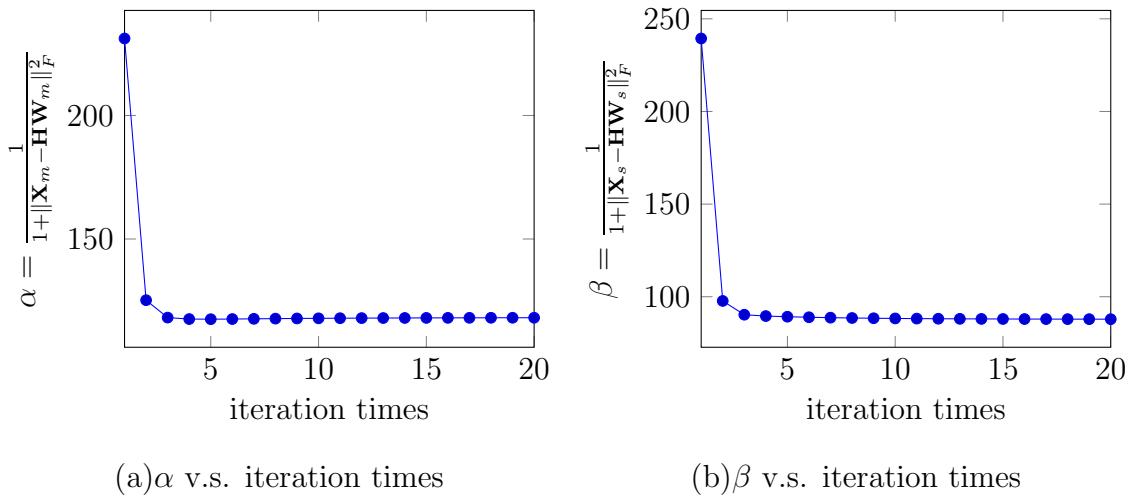


Figure 12:  $\alpha$  and  $\beta$  v.s. iteration times in sprite analysis

Table 4: Results of spirit analysis

Response	Method \ Samples						
		20	25	30	35	40	45
Ethanol	NO	8.6575	10.2131	5.9807	10.565	6.0311	5.5756
	DS	6.4654	8.2311	4.6786	9.1231	7.8797	6.3242
	PDS	6.4354	7.6766	4.5657	5.5756	7.1231	5.5657
	CCA	4.7686	5.3435	5.2423	4.3543	3.8980	4.2342
	KCCA	4.2334	3.9879	5.1230	5.6356	4.2423	3.9897
	SR	5.6546	4.2131	5.3545	5.2132	3.231	4.1231
	Proposed	4.2131	3.2131	4.7877	3.5656	3.2131	<b>3.1231</b>
acetic ether	NO	7.3242	6.7342	4.4353	6.2342	4.0243	4.0232
	DS	6.2342	3.1231	4.5654	3.9789	6.2423	4.2434
	PDS	5.2423	2.2344	3.2311	2.3454	4.3242	2.7864
	CCA	2.3242	2.3242	2.3454	3.2141	1.3432	1.6897
	KCCA	4.4564	2.4645	1.8987	3.5657	<b>1.3242</b>	1.4325
	SR	4.0242	1.5213	1.8342	2.2310	1.4353	1.4234
	Proposed	6.2423	4.2131	3.3214	2.343	1.5645	2.2141
ethyl butyrate	NO	43.5464	60.4534	34.5464	20.3235	23.3242	18.5465
	DS	9.2423	8.4543	8.3242	7.3423	8.3534	9.3234
	PDS	8.3423	7.2423	9.3242	6.2342	8.3443	8.2342
	CCA	6.2342	5.2342	7.3242	7.2342	6.2342	5.3243
	KCCA	5.3533	5.2311	4.8987	4.8987	4.5657	5.3252
	SR	5.2423	4.8978	5.4654	5.4545	4.4344	4.3543
	Proposed	4.3253	5.4354	4.2432	3.9778	<b>3.8977</b>	5.5457
Ethyl lactate	NO	37.5654	36.5645	30.4456	28.8978	36.4646	30.4546
	DS	10.2432	16.3423	15.3242	12.2432	10.2423	9.3242
	PDS	36.3254	35.2352	30.3543	28.3543	25.6765	24.35435
	CCA	7.4234	8.3534	7.4534	7.1657	6.32434	6.2131
	KCCA	8.2343	8.2132	7.2343	6.8797	6.3242	5.8932
	SR	8.3423	6.2343	6.2122	6.0123	5.8978	5.5343
	Proposed	28.4353	8.8978	7.6241	7.0131	6.3432	<b>5.4181</b>

are represented with the common basis and corresponding coefficients are then transferred using the transformation matrix. The slave testing spectra can be transferred using common basis and the corrected coefficients. The convergence and bound are also discussed. Extensive experiments are conducted, experimental results demonstrate that our robust calibration transfer model can generally outperform the existing methods.

### **Conflict of interest**

The authors declare that they have no conflict of interest.

### **Acknowledgment**

This work is supported partially by the the national natural science foundation of Hubei Province (No.2014CFC1131), Research and Innovation Initiatives of WHPU(No.2016J06).

### **References**

- [1] Harald Martens and Tormod Naes. *Multivariate calibration*. John Wiley & Sons, 1992.
- [2] Yongdong Wang, Michael J Lysaght, and Bruce R Kowalski. Improvement of multivariate calibration through instrument standardization. *Analytical Chemistry*, 64(5):562–564, 1992.
- [3] Kenneth R Beebe and Bruce R Kowalski. An introduction to multivariate calibration and analysis. *Analytical Chemistry*, 59(17):1007A–1017A, 1987.

- [4] A Herrero and MC Ortiz. Multivariate calibration transfer applied to the routine polarographic determination of copper, lead, cadmium and zinc. *Analytica chimica acta*, 348(1):51–59, 1997.
- [5] Robert N Feudale, Nathaniel A Woody, Huawei Tan, Anthony J Myles, Steven D Brown, and Joan Ferré. Transfer of multivariate calibration models: a review. *Chemometrics and Intelligent Laboratory Systems*, 64(2):181–192, 2002.
- [6] Fernanda Araújo Honorato, Roberto Kawakami Harrop Galvão, Maria Fernanda Pimentel, Benício de Barros Neto, Mário César Ugulino Araújo, and Florival Rodrigues de Carvalho. Robust modeling for multivariate calibration transfer by the successive projections algorithm. *Chemometrics and intelligent laboratory systems*, 76(1):65–72, 2005.
- [7] Graciela M Escandar, Patricia C Damiani, Héctor C Goicoechea, and Alejandro C Olivieri. A review of multivariate calibration methods applied to biomedical analysis. *Microchemical journal*, 82(1):29–42, 2006.
- [8] Yusuf Sulub, Rosario LoBrutto, Richard Vivilecchia, and Busolo Wa Wabuyele. Content uniformity determination of pharmaceutical tablets using five near-infrared reflectance spectrometers: a process analytical technology (pat) approach using robust multivariate calibration transfer algorithms. *Analytica chimica acta*, 611(2):143–150, 2008.
- [9] Claudete Fernandes Pereira, Maria Fernanda Pimentel, Roberto Kawakami Harrop Galvão, Fernanda Araújo Honorato, Luiz Stragevitch, and Marcelo Nascimento Martins. A comparative study of calibration

transfer methods for determination of gasoline quality parameters in three different near infrared spectrometers. *Analytica chimica acta*, 611(1):41–47, 2008.

- [10] Anne Andrew and Tom Fearn. Transfer by orthogonal projection: making near-infrared calibrations robust to between-instrument variation. *Chemometrics and Intelligent Laboratory Systems*, 72(1):51–56, 2004.
- [11] E Bouveresse and DL Massart. Standardisation of near-infrared spectrometric instruments: a review. *Vibrational Spectroscopy*, 11(1):3–15, 1996.
- [12] Thomas Dean and Tomas Isaksson. Standardisation: What is it and how is it done? part 2. *NIR news*, 4(4):14–15, 1993.
- [13] Yongdong Wang, David J Veltkamp, and Bruce R Kowalski. Multivariate instrument standardization. *Analytical chemistry*, 63(23):2750–2756, 1991.
- [14] Ana Herrero and M Cruz Ortiz. Piecewise direct standardization method applied to the simultaneous determination of pb (ii), sn (iv) and cd (ii) by differential pulse polarography. *Electroanalysis*, 10(10):717–721, 1998.
- [15] Wei Fan, Yizeng Liang, Dalin Yuan, and Jiajun Wang. Calibration model transfer for near-infrared spectra based on canonical correlation analysis. *Analytica chimica acta*, 623(1):22–29, 2008.
- [16] Yong Xu, Aini Zhong, Jian Yang, and David Zhang. Lpp solution schemes for use with face recognition. *Pattern Recognition*, 43(12):4165–4176, 2010.

- [17] Sibao Chen, Haifeng Zhao, Min Kong, and Bin Luo. 2d-lpp: a two-dimensional extension of locality preserving projections. *Neurocomputing*, 70(4):912–921, 2007.
- [18] Xiaofei He, Deng Cai, Shuicheng Yan, and Hong-Jiang Zhang. Neighborhood preserving embedding. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1208–1213. IEEE, 2005.
- [19] Jiangtao Peng, Silong Peng, An Jiang, and Jie Tan. Near-infrared calibration transfer based on spectral regression. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 78(4):1315–1320, 2011.
- [20] Moshe Idan and Jason L Speyer. Cauchy estimation for linear scalar systems. *Automatic Control, IEEE Transactions on*, 55(6):1329–1342, 2010.
- [21] Moshe Idan and Jason L Speyer. Multivariate cauchy estimator with scalar measurement and process noises. *SIAM Journal on Control and Optimization*, 52(2):1108–1141, 2014.
- [22] Jie Lin. Near-ir calibration transfer between different temperatures. *Applied spectroscopy*, 52(12):1591–1596, 1998.
- [23] H Voss and Ulrich Eckhardt. Linear convergence of generalized weiszfeld’s method. *Computing*, 25(3):243–251, 1980.
- [24] Ivan Singer. *Duality for nonconvex approximation and optimization*. Springer Science & Business Media, 2007.

- [25] Chang Xu, Dacheng Tao, and Chao Xu. Multi-view intact space learning. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 37(12):2531–2544, 2015.