

A Novel Sky-Subtraction Method Based on Non-negative Matrix Factorisation with Sparsity for Multi-object Fibre Spectroscopy

Bo Zhang¹, Long Zhang¹ and Zhongfu Ye^{1,2}

¹Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China

²Email: yezf@ustc.edu.cn

(RECEIVED March 17, 2016; ACCEPTED October 10, 2016)

Abstract

A novel sky-subtraction method based on non-negative matrix factorisation with sparsity is proposed in this paper. The proposed non-negative matrix factorisation with sparsity method is redesigned for sky-subtraction considering the characteristics of the skylights. It has two constraint terms, one for sparsity and the other for homogeneity. Different from the standard sky-subtraction techniques, such as the B-spline curve fitting methods and the Principal Components Analysis approaches, sky-subtraction based on non-negative matrix factorisation with sparsity method has higher accuracy and flexibility. The non-negative matrix factorisation with sparsity method has research value for the sky-subtraction on multi-object fibre spectroscopic telescope surveys. To demonstrate the effectiveness and superiority of the proposed algorithm, experiments are performed on Large Sky Area Multi-Object Fiber Spectroscopic Telescope data, as the mechanisms of the multi-object fibre spectroscopic telescopes are similar.

Keywords: instrumentation: spectrographs – methods: data analysis – techniques: spectroscopic – telescopes

1 INTRODUCTION

Sky-subtraction is one of the key links with vital importance in the data processing of the multi-object fibre spectroscopy. The precision will be affected by the mixed astigmatism in the telescope, the contamination of adjacent fibres, the transmission efficiency of the optical fibres, the accuracy of the wavelength calibration, etc. Except for reducing the influence of the aforementioned pretreatment processing by ramping up their degree of precision, it is particularly important to study the high precision methods for sky-subtraction to improve the accuracy of the sky-subtraction fundamentally. Therefore, the focus of this paper lies in researching the sky-subtraction methods that are exist and being used on the famous multi-object fibre spectroscopic telescopes and proposing a new sky-subtraction method with higher precision.

Relatively well-known sky-subtraction methods are listed as follows. The traditional sky-subtraction method is beam-switching method (Barden et al. 1993). The object fibres and the corresponding sampling sky fibres are observed in pairs, sacrificing half of the object fibres or half of the observing time. To solve the problem, the nod-shuffle method is proposed (Glazebrook & Bland-Hawthorn 2001). The object spectra and the corresponding sky spectra are obtained almost simultaneously and the accuracy of sky-subtraction under the

condition of faint light is improved. The nod-shuffle method is used on pipeline of the Wide Field Spectrograph (WiFeS) (Childress et al. 2014) in Australia. The sky background simulation methods are proposed to improve the efficiency of the optical observation, such as the B-spline curve fitting method (Kelson 2003). The B-spline curve fitting method has the ability of modelling the sky background flexibly. However, by B-spline curve fitting, all the sky fibres need to be arranged based on the wavelength and be combined into a complete sky fibre. In this procedure, all the sky fibres are thought to be the same. When the sky area is large, neglecting the differences amongst the sky fibres will bring superior error. Besides, due to the time consuming of the curve fitting modelling process, the method has low efficiency on processing the large-scale spectral data. The B-spline curve fitting method was used on SDSS on early stage (Gunn et al. 2006), and it is being used on Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) (Luo et al. 2015), the optical integral-field units (IFUs) of the Gemini GMOS-N spectrograph (Bolton & Burles 2007). The principal component analysis (PCA) is the common sky-subtraction method of the current frontier multi-object fibre spectroscopic data processing system, which was first put forward by Kurtz & Mink (2000). Then comes some improved and variant methods (Mink & Kurtz 2001; Wild & Hewett 2005; Sharp & Parkinson 2010). The PCA method

is being used on the APOGEE pipeline of SDSS (Nidever et al. 2015), the WiggleZ Dark Energy survey (Drinkwater et al. 2010), and the Galaxy And Mass Assembly (GAMA) of AAOmega (Liske et al. 2015).

Non-negative Matrix Factorisation (NMF) originates from PCA. The components in the vase vectors obtained from PCA have positive values and negative values, which is contradictory to the objective reality in many applications. NMF ensures the results to be non-negative, with the advantages of simple implementation and fast decomposition, it has been successfully applied in many fields of data analysis, especially for machine learning. NMF can automatically discover the hidden patterns or trends in the data. It is an effective tool to deal with large scale data processing and analysis. In view of the non-negativity of spectral energy, NMF method can be used to do sky-subtraction. A non-negative matrix factorisation with sparsity (NMF+S) method is proposed in this paper to achieve high-precision sky-subtraction.

The LAMOST is designed with both large aperture (effective aperture of 3.6–4.9 m) and wide field of view (FOV) (5°) (Cui et al. 2012). It is equipped with 4 000 optical fibres and has 16 spectrographs with 32 CCDs. In each spectrograph, there are 250 fibres, 10 of them are sky fibres and the others are used for observing the targets, i.e the object fibres. The experiments are carried on simulated data from the LAMOST software system (Cui et al. 2012; Zhang 2007), which are designed based on the parameters and characters of LAMOST, and real data provided by the LAMOST astronomical observatory (Luo et al. 2015). Because of the similar mechanism of multi-object fibre spectroscopic telescopes, the method proposed can also be applied to other telescopes and has referential significance for other surveys.

2 THE NMF+S SKY-SUBTRACTION METHOD

2.1. Proposing a new NMF+S method

NMF is a method by which a $m \times n$ matrix \mathbf{X} , is approximated by the product of $m \times r$ matrix \mathbf{W} and $r \times n$ matrix \mathbf{H} , enforcing the constraint that all matrices are non-negative,

$$\mathbf{X} \approx \mathbf{WH} \text{ s.t. } \mathbf{W}, \mathbf{H} \geq 0. \quad (1)$$

To find an approximate factorisation of equation (1), a cost function can be constructed using some measure of distance between \mathbf{W} and \mathbf{H} (Lee & Seung 2001). Considering the feature of skylight, to effectively correspond to a Poisson process, the generalised Kullback–Leibler (KL) divergence is chosen here:

$$D(\mathbf{X} \parallel \mathbf{WH}) = \sum_{ij} \left(X_{ij} \log \frac{X_{ij}}{(\mathbf{WH})_{ij}} - X_{ij} + (\mathbf{WH})_{ij} \right). \quad (2)$$

Sparse NMF is an extension of NMF, an additional sparsity constraint is added to the cost function (Schmidt & Olsson 2006):

$$c_s(\mathbf{H}) = \|\mathbf{H}\|_1 = \sum_{k,j} H_{k,j}. \quad (3)$$

Considering the characteristic of the skylight, the flux of the sky fibres on each position in wavelength orientation has consistency, so an additional two-norm constraint is enforced on the matrix \mathbf{W} .

$$c_h(\mathbf{W}) = \sum_{i=1}^m \|\mathbf{w}^i\|_2 = \sum_{i=1}^m \sqrt{\sum_{k=1}^r W_{ik}^2}, \quad (4)$$

where \mathbf{w}^i is a row vector of \mathbf{W} . $\|\mathbf{w}^i\|_2$ stands for the norm of vector \mathbf{w}^i .

Thus, the cost function is defined as

$$c(\mathbf{W}, \mathbf{H}) = c_r(\mathbf{W}, \mathbf{H}) + \alpha c_s(\mathbf{H}) + \beta c_h(\mathbf{W}), \quad (5)$$

where α is a penalty term that enforces the sparsity and controls the trade-off between sparsity and accuracy of the approximation. β is the penalty term of the consistency constraint.

The multiplicative update rules are used to minimise equation (5) (Le Roux, Weninger, & Hershey 2015). The gradients of the cost function are calculated to gain the update criterions of \mathbf{H} and \mathbf{W} .

$$\begin{aligned} & \frac{\partial c(\mathbf{W}, \mathbf{H})}{\partial W_{ik}} \\ &= \frac{\partial}{\partial W_{ik}} \left(\sum_{ij} \left(X_{ij} \log \frac{X_{ij}}{W_{ik} \cdot H_{kj}} - X_{ij} + W_{ik} \cdot H_{kj} \right) \right. \\ & \quad \left. + \alpha \sum_{k,j} H_{k,j} + \beta \sum_{i=1}^m \sqrt{\sum_{k=1}^r W_{ik}^2} \right) \\ &= - \sum_j X_{ij} \cdot \frac{1}{W_{ik} \cdot H_{kj}} \cdot H_{kj} + \sum_j H_{kj} + \beta \cdot \frac{1}{\sqrt{\sum_{k=1}^r W_{ik}^2}} \cdot W_{ik} \end{aligned} \quad (6)$$

$$\begin{aligned} & \frac{\partial c(\mathbf{W}, \mathbf{H})}{\partial H_{kj}} \\ &= \frac{\partial}{\partial H_{kj}} \left(\sum_{ij} \left(X_{ij} \log \frac{X_{ij}}{W_{ik} \cdot H_{kj}} - X_{ij} + W_{ik} \cdot H_{kj} \right) \right. \\ & \quad \left. + \alpha \sum_{k,j} H_{k,j} + \beta \sum_{i=1}^m \sqrt{\sum_{k=1}^r W_{ik}^2} \right) \\ &= - \sum_i X_{ij} \cdot \frac{1}{W_{ik} \cdot H_{kj}} \cdot W_{ik} + \sum_i W_{ik} + \alpha. \end{aligned} \quad (7)$$

Extending equations (6) and (7) to vector derivation, we have

$$\frac{\partial c(\mathbf{W}, \mathbf{H})}{\partial \mathbf{W}} = - \left(\frac{\mathbf{X}}{\mathbf{WH}} \right) \cdot \mathbf{H}^T + \mathbf{1} \cdot \mathbf{H}^T + \mathbf{D} \cdot \mathbf{W} \quad (8)$$

$$\begin{aligned} \mathbf{D} &= \beta \cdot \text{diag}(\mathbf{d}) \\ \mathbf{d} &= \left(\frac{1}{\|\mathbf{w}^1\|_2}, \dots, \frac{1}{\|\mathbf{w}^m\|_2}, \dots, \frac{1}{\|\mathbf{w}^m\|_2} \right)^T, \end{aligned}$$

where $\text{diag}(\mathbf{d})$ indicates a square diagonal matrix with the elements of vector \mathbf{d} on the main diagonal. $\mathbf{1}$ stands for a $m \times n$ matrix with all elements equal to 1.

$$\frac{\partial c(\mathbf{W}, \mathbf{H})}{\partial \mathbf{H}} = -\mathbf{W}^T \cdot \left(\frac{\mathbf{X}}{\mathbf{WH}} \right) + \mathbf{W}^T \cdot \mathbf{1} + \alpha \quad (9)$$

Then, the complete algorithm for the proposed NMF+S is given by the following rules:

- (1) Initialise \mathbf{H} and \mathbf{W} randomly.
- (2) Normalise the basis vectors according to $\mathbf{W} \leftarrow \bar{\mathbf{W}}$, where $\bar{\mathbf{W}}$ is the columnwise normalised basis matrix.
- (3) Calculate the reconstruction according to

$$\hat{\mathbf{X}} = \bar{\mathbf{W}}\mathbf{H}. \quad (10)$$

- (4) Update the activities according to

$$\mathbf{H} \leftarrow \mathbf{H} \bullet \frac{\bar{\mathbf{W}}^T \cdot (\frac{\mathbf{X}}{\bar{\mathbf{W}}\mathbf{H}})}{\bar{\mathbf{W}}^T \cdot \mathbf{1} + \alpha}. \quad (11)$$

The operator \bullet indicates pointwise multiplication and $\frac{\mathbf{X}}{\bar{\mathbf{W}}\mathbf{H}}$ indicates pointwise division.

- (5) Calculate the reconstruction with the new activities according to

$$\hat{\mathbf{X}} = \bar{\mathbf{W}}\mathbf{H}. \quad (12)$$

- (6) Update the basis vectors according to

$$\mathbf{W} \leftarrow \mathbf{W} \bullet \frac{(\frac{\mathbf{X}}{\bar{\mathbf{W}}\mathbf{H}}) \cdot \mathbf{H}^T}{\mathbf{1} \cdot \mathbf{H}^T + \mathbf{D} \cdot \bar{\mathbf{W}}}. \quad (13)$$

- (7) Normalise the basis vectors according to $\mathbf{W} \leftarrow \bar{\mathbf{W}}$.
- (8) Return to (3) until convergence.

2.2. Sky-subtraction using the proposed NMF+S method

In reducing the skylight, fibre efficiency corrections for the sky fibres and the object fibres are done first. Due to the different dispersion curves of each fibre, the wavelengths of each fibre are interpolated to a uniform wavelength coordinate.

The sky spectrum is composed of sky emission lines $\tilde{\mathbf{S}}$ and continuous spectrum $\tilde{\mathbf{S}}$. To separate them, median filtering is done on each sky spectrum. As the deviation of the continuous spectrum in each sky fibre is small, the mean value of the continuous spectrum of all the sky fibres is taken as the continuous spectrum flux. Subtracting the continuous spectrum flux from each sky spectrum, the flux of the sky emission lines $\tilde{\mathbf{S}}$ is obtained to do the NMF training. Do the median filtering in the same way on the object spectra \mathbf{O} , the flux of the sky emission lines $\hat{\mathbf{O}}$ of the object spectra is obtained to do the NMF testing.

In order to improve the operational speed of NMF, formulate a criteria to select the pixels containing components of skylights, such as calculating the variance or mean value of the flux on each wavelength position of each fibre. As the sky emission lines are considered to be positive, the non-negativeness of the flux is also included in the criteria. Mark the locations of these points on the sky emission lines $\tilde{\mathbf{S}}$. The selected pixels of sky fibres form a sky sample matrix $\mathbf{X}_{\tilde{\mathbf{S}}}$ for NMF training. Select the pixels in the same marked locations in the wavelength orientation on the object spectra, forming the object sampling matrix $\mathbf{X}_{\hat{\mathbf{O}}}$ for NMF testing.

For the 10 sky fibres on LAMOST, the selected sample matrix $\mathbf{X}_{\tilde{\mathbf{S}}}$ can be formulated as

$$\mathbf{X}_{\tilde{\mathbf{S}}} \approx \mathbf{W}\mathbf{H}.s.t.\mathbf{W}, \mathbf{H} \geq 0. \quad (14)$$

Utilising the NMF+S method in Section 2.1 for training, \mathbf{W} is obtained. Then, take it as the base vectors, keeping fixed, to do testing with the selected sample matrix of the object fibres, by the NMF+S method in Section 2.1.

$$\mathbf{X}_{\hat{\mathbf{O}}} \approx \mathbf{W}\mathbf{H}'.s.t.\mathbf{W}, \mathbf{H}' \geq 0. \quad (15)$$

Through the NMF+S method, the sky emission lines on the selected positions $\tilde{\mathbf{S}}_{\text{select}}$ are rebuilt by

$$\tilde{\mathbf{S}}_{\text{select}} = \mathbf{W}\mathbf{H}'. \quad (16)$$

Replacing the flux of $\tilde{\mathbf{S}}_{\text{select}}$ in $\tilde{\mathbf{S}}$ on the selected positions and remaining the flux of other positions unchanged, $\tilde{\mathbf{S}}$ is refreshed as $\tilde{\mathbf{S}}'$. Adding the continuous spectral part of the sky $\tilde{\mathbf{S}}$, the sky is reconstructed. The flux of the object spectra after sky-subtraction can be expressed as

$$\hat{\mathbf{O}} = \mathbf{O} - \tilde{\mathbf{S}}' - \tilde{\mathbf{S}}. \quad (17)$$

3 EXPERIMENTS AND DISCUSSION

3.1. Simulated data on LAMOST

The experiments in this part are simulated data from LAMOST software system (Cui et al. 2012). The data is designed strictly in accordance with the characteristics of LAMOST. The 250 fibres, 10 of which are sky fibres, are exactly simulated as the real data on LAMOST. Each of the fibres has 4 096 pixels in the wavelength orientation. The detailed simulation mechanisms and the relevant parameters can be found in a dissertation (Zhang 2007).

Taking all the errors of the instruments and the influences of noises into account, the simulated object spectra can be expressed as

$$O(i, \lambda) = \{ [O_r(i, \lambda) \times d_{\text{sky}}(\lambda) + \text{sky}_r(\lambda)] \times d_{\text{tel}}(\lambda) \times d_{\text{fib}}(i, \lambda) \times d_{\text{spm}}(i, \lambda) + \text{scatter}(i, \lambda) \} \times d_{\text{CCD}}(i, \lambda) + c_h(i, \lambda) + \text{Bias} \quad (18)$$

where, O_r is the flux of the object outside the atmosphere. d_{sky} is the atmospheric extinction function. sky_r is the flux of the sky. d_{tel} is the flux response function of the optical telescope system. d_{fib} is the optical fibre transmissivity. d_{spm} is the flux response function of the spectrograph, i.e. the influence of spectrometer dispersion on flux, containing the distortion and the vignetting of the spectra. scatter is the flux of the stray light, containing the uniform stray light and the contamination from adjacent fibres. d_{CCD} is the response function of CCDs. c_h is the flux of the cosmic ray. Bias is the background flux of the CCDs.

The sky can be expressed as

$$\text{sky}(i, \lambda) = [\text{sky}_r(\lambda) \times d_{\text{tel}}(\lambda) \times d_{\text{fib}}(i, \lambda) \times d_{\text{spm}}(i, \lambda) + \text{scatter}(i, \lambda)] \times d_{\text{CCD}}(i, \lambda) + c_h(i, \lambda) + \text{Bias}. \quad (19)$$

More details of the simulation refer to Zhang (2007).

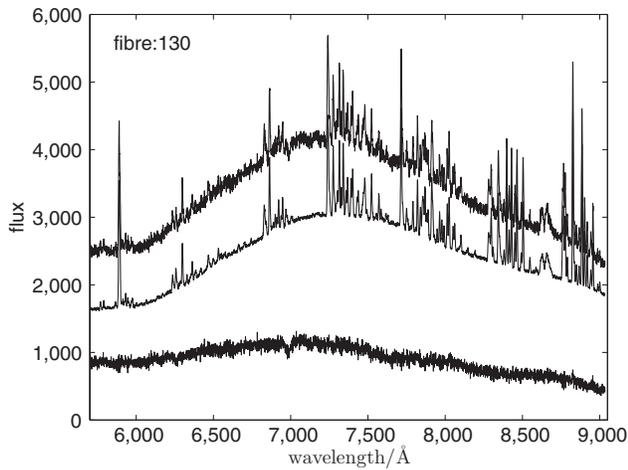


Figure 1. Sky-subtraction results on fibre 130 in the red channel by the proposed NMF+S method, from top to bottom: The original object spectrum, the reconstructed sky spectrum, the object spectrum after sky-subtraction.

The sky spectrum contains emission lines and continuous spectrum. The emission lines are simulated by the existing empirical data. The continuous spectrum is derived from the empirical formula. Since the sky emission lines assume highly localised distributions in the red channel, and the flux of them is generally higher than the blue channel, the experiments are carried on the red channel as the condition described above is more difficult for sky-subtraction. The sky-subtraction is carried out after the bias subtraction, flat field correction, centroid tracing, spectrum extraction, and wavelength calibration (Luo et al. 2015).

The two B-spline curve fitting methods in Zhu & Ye (2012) (The 2-D B-spline method has been proposed in it. The 1-D B-spline method in it is being used on LAMOST (Luo et al. 2015)) and the PCA method (Sharp & Parkinson 2010) being used on the APOGEE pipeline of SDSS (Nidever et al. 2015) are contrast methods to demonstrate the effectiveness and superiority of the proposed NMF+S sky-subtraction method.

In the experiments, considering the width of the sky emission lines, a median filter with window length 111 is done on the flux of the sky to calculate the continuous spectrum of skylight \tilde{S} . The criterion to select the sample pixels in the NMF+S method is selecting the pixels whose flux is bigger than 10% of the max flux. By debugging the parameters, we found that the parameters chosen in the training and testing of the proposed NMF method are the same. The parameters in the proposed method are chosen as $\alpha = 1.5$, $\beta = 6$. The number of basis vectors is 15, which is a little bigger than the number of sky fibres, i.e. 10, to retain some redundancy.

The sky-subtraction result on fibre 130 by the proposed method is shown in Figure 1.

The experimental results on contrast methods are shown in Figures 2 and 3.

The experimental result figures show that the PCA method and the NMF+S method are more accurate than the B-spline methods, based on the facts that the object spectrum after sky-

Table 1. The variances e of the object spectra after sky-subtraction by contrast methods.

Fibre no.	B-spline	2-D B-spline	PCA	NMF+S
15	229.10	225.54	217.06	214.92
25	384.46	380.70	371.72	368.90
60	341.85	340.76	330.78	328.14
100	249.19	247.53	236.30	234.01
130	196.03	192.55	181.89	180.15
145	162.66	159.46	149.18	147.78
165	580.62	579.43	566.95	564.87
190	191.93	186.87	176.55	174.67
205	304.34	300.54	292.81	290.02
240	178.59	177.26	168.04	166.21

subtraction is more smooth with less burrs, illustrating that the sky emission lines are subtracted more completely. Since the difference between the PCA method and the NMF+S method is not obvious, and moreover, to validate the performance of the proposed algorithm, a target of variance is defined to measure the accuracy of the methods objectively and mathematically. The variance of the object spectrum after sky-subtraction is defined as

$$e = \frac{1}{M} \sum_{k=1}^M \left[\hat{O}(k) - \frac{1}{M} \sum_{k=1}^M \hat{O}(k) \right]^2, \quad (20)$$

where $\hat{O}(k)$ stands for the flux of the object spectrum after sky-subtraction, M is the number of pixels on the wavelength of the full object spectrum. The smaller the value of e is, the less the sky emission lines residuals of the object spectrum are, i.e. the more complete the sky-subtraction is.

Ten object fibre spectra are selected randomly and the variances of them after sky-subtraction are shown in Table 1.

The statistical data in the Table 1 shows that the proposed NMF+S method on sky reduction has the minimum variance, indicating that it is the most thorough sky-subtraction method to some extent. Furthermore, as the NMF methods have many variant, with appropriate improvements, the algorithm proposed has great room for improvement.

3.2. Real data on LAMOST

To verify the accuracy and practicability of the proposed algorithm, we repeat the experiments on LAMOST real data. The data are provided by LAMOST astronomical observatory (Luo et al. 2015). In view of the performances of the methods in section 3.1, the PCA method (Sharp & Parkinson 2010) is taken as the contrast method.

As there are cosmic rays in real data, to avoid the influence of them in sky background modelling, the cosmic rays in the sky should be eliminated when choosing the sky sample matrix $\mathbf{X}_{\tilde{S}}$. The criteria to select the sample pixels in the NMF+S method and the PCA method are both based on the flux and the variance of the spectra. For the NMF+S method,

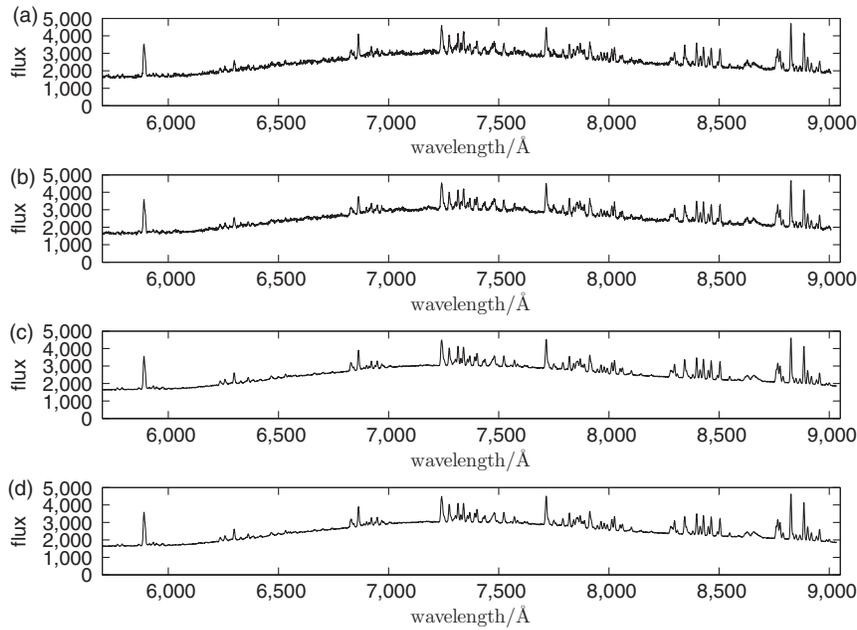


Figure 2. Comparison results of the reconstructed sky spectrum on fibre 130 in the red channel. (a) B-spline method in LAMOST reduction. (b) 2-D B-spline method in Zhu & Ye (2012). (c) PCA method (d) NMF+S method.

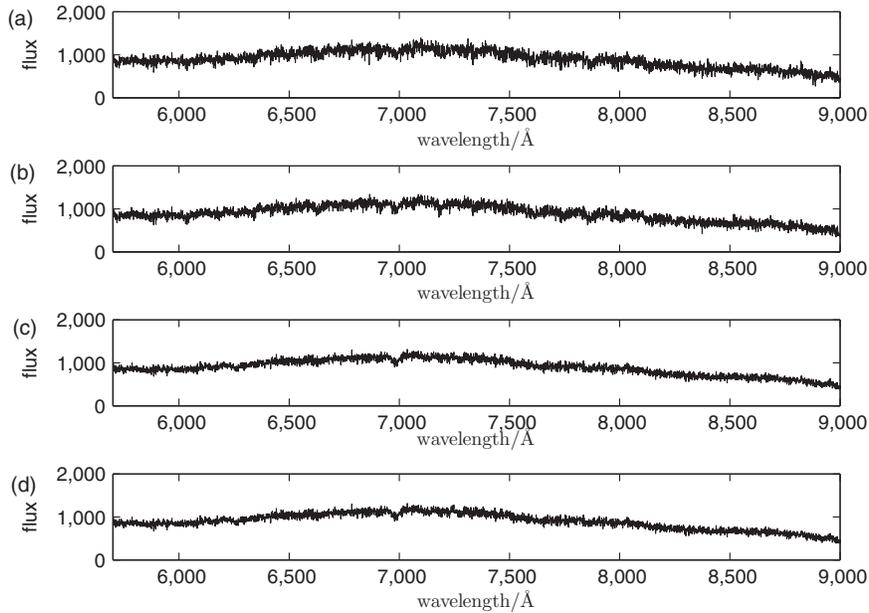


Figure 3. Comparison results of the object spectrum after sky-subtraction on fibre 130 in the red channel. (a) B-spline method in LAMOST reduction. (b) 2-D B-spline method in Zhu & Ye (2012). (c) PCA method (d) NMF+S method.

the non-negativeness of the sample pixels is also considered as an additional criterion. The numbers of the sample pixels in both the NMF+S method and the PCA method are about the same. The parameters in the proposed method are cho-

sen as $\alpha=0.1$, $\beta=3$. Other parameters used are the same as section 3.1.

The sky-subtraction result on fibre 130 by the proposed method is shown in Figure 4.

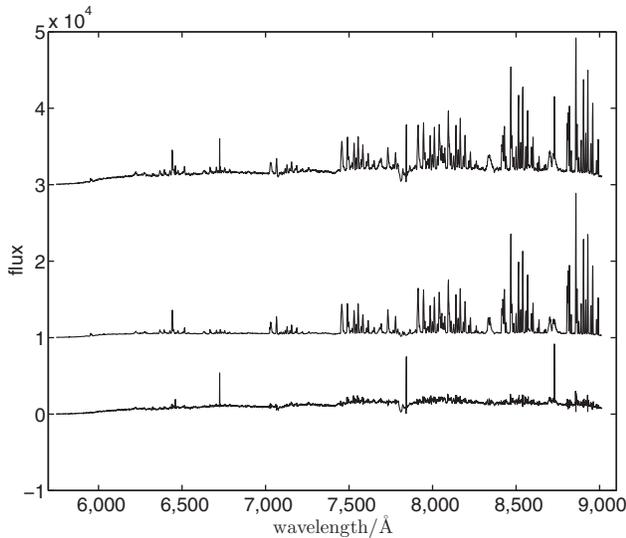


Figure 4. Sky-subtraction results on fibre 40 in the red channel by the proposed NMF+S method, from top to bottom: The original object spectrum, the reconstructed sky spectrum, the object spectrum after sky-subtraction. To make it seem more clear, the original object spectrum is moved up to 30 000, the reconstructed sky spectrum is moved up to 10 000.

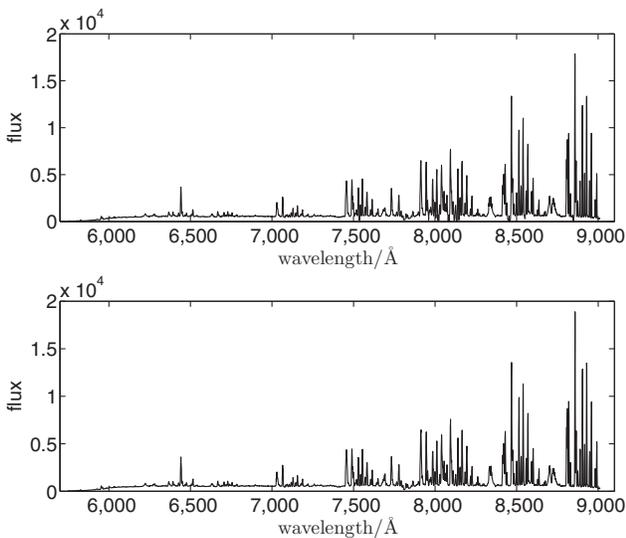


Figure 5. Comparison results of the reconstructed sky spectrum on fibre 40 in the red channel. top: PCA method; bottom: NMF+S method.

The experimental results on contrast method are shown in Figures 5 and 6.

The high flux emission lines of in Figures 5 and 6 are the cosmic rays in the object spectrum fibre 40. The cosmic rays in the object spectra will be removed in the subsequent processing steps.

Ten object fibre spectra are selected randomly and the variances of them after sky-subtraction are shown in Table 2. The cosmic rays in the object spectra is included when calculating the variances.

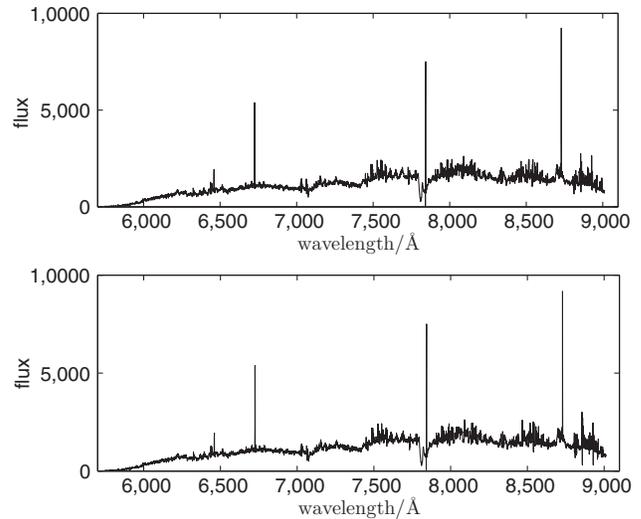


Figure 6. Comparison results of the object spectrum after sky-subtraction on fibre 40 in the red channel. top: PCA method; bottom: NMF+S method.

Table 2. The variances ϵ of the object spectra after sky-subtraction by contrast methods.

Fibre no.	PCA	NMF+S
40	542.67	536.77
50	224.72	223.90
80	330.78	328.14
100	250.60	251.05
120	111.26	110.88
140	413.47	412.84
170	893.64	893.35
190	114.08	113.30
210	759.55	758.66
230	121.51	121.79

The results of the experiments on real data show that in most instances, the accuracy of the NMF+S method is higher than the PCA method. As the real data is more unstable, the performance of the proposed method is close to but better than the PCA method.

4 CONCLUSION

A novel sky-subtraction method based on an improved NMF+S method has been proposed in this paper. The method has been designed for sky-subtraction step of the surveys on multi-object fibre spectroscopy. Taking the characteristics of the skylights into account, the proposed NMF+S method has a sparseness constraint term and a homogeneity constraint term. By experiments on LAMOST data, the proposed method has higher accuracy and flexibility, contrasted to the universal sky-subtraction methods, such as the B-spline curve fitting methods and the PCA method. Since the mechanisms of the multi-object fibre spectroscopic telescopes are similar, the proposed NMF+S method has research value for

sky-subtraction in the surveys on all multi-object fibre spectroscopic telescopes.

ACKNOWLEDGEMENTS

The authors acknowledge support from the National Natural Science Foundation of China (NSFC 11078016).

REFERENCES

- Barden, S. C., Elston, R., Armandroff, T., & Pryor, C. P. 1993, in ASP Conf. Ser., Vol. 37, *Fiber Optics in Astronomy II*, ed. P. M. Gray (San Francisco: ASP), 223
- Bolton, A. S., & Burles, S. 2007, *NJPh*, 9, 443
- Childress, M. J., Vogt, F. P., Nielsen, J., & Sharp, R. G. 2014, *Ap&SS*, 349, 617
- Cui, X.-Q., et al. 2012, *RAA*, 12, 1197
- Drinkwater, M. J., et al. 2010, *MNRAS*, 401, 1429
- Glazebrook, K., & Bland-Hawthorn, J. 2001, *PASP*, 113, 197
- Gunn, J. E., et al. 2006, *AJ*, 131, 2332
- Kelson, D. D. 2003, *PASP*, 115, 688
- Kurtz, M. J., & Mink, D. J. 2000, *ApJL*, 533, L183
- Le Roux, J., Weninger, F., & Hershey, J. 2015, Mitsubishi Electric Research Labs (MERL), Cambridge, MA, USA, Tech. Rep., no. TR2015-023
- Lee, D. D., & Seung, H. S. 2001, in *Advances in neural information processing systems*, vol. 13 (Cambridge, MA: MIT Press), 556–562
- Liske, J., et al. 2015, *MNRAS*, 452, 2087
- Luo, A.-L., et al. 2015, *RAA*, 15, 1095
- Mink, D., & Kurtz, M. 2001, in ASP Conf. Proc., Vol. 238, *Astronomical Data Analysis Software and Systems X*, ed. F. R. Harneden, Jr., F. A. Primini, and H. E. Payne (San Francisco: ASP), 491
- Nidever, D. L., et al. 2015, *AJ*, 150, 173
- Schmidt, M. N., & Olsson, R. K. 2006, “Single-channel speech separation using sparse non-negative matrix factorization,” *International Conference on Spoken Language Processing (INTER-SPEECH)*
- Sharp, R., & Parkinson, H. 2010, *MNRAS*, 408, 2495
- Wild, V., & Hewett, P. C. 2005, *MNRAS*, 358, 1083
- Zhang, L. 2007, PhD thesis, University of Science and Technology of China
- Zhu, J., & Ye, Z. 2012, *PASA*, 29, 78