

Bilateral Joint-Sparse Regression for Hyperspectral Unmixing

Jie Huang, Wu-Chao Di, Jin-Ju Wang, Jie Lin and Ting-Zhu Huang

Abstract—Sparse hyperspectral unmixing has been a hot topic in recent years. Joint sparsity assumes that each pixel in a small neighborhood of hyperspectral images (HSIs) is composed of the same endmembers, which results in a few nonzero rows in the abundance matrix. Recall that a plethora of unmixing algorithms transform a 3-D HSI into a 2-D matrix with vertical priority. The transformation makes matrix computation easier. It is, however, hard to maintain the horizontal spatial information in HSIs in many cases. To make further use of the spatial information of HSIs, in this article, we propose a *bilateral joint-sparse structure* for hyperspectral unmixing in an attempt to exploit the local joint sparsity of the abundance matrix in both the vertical and horizontal directions. In particular, we introduce a permutation matrix to realize the bilateral joint-sparse representation and there is no need to construct the matrix explicitly. Moreover, we propose to simultaneously impose the bilateral joint-sparse structure and low rankness on the abundance and develop a new algorithm named *bilateral joint-sparse and low-rank unmixing*. The proposed algorithm is based on the alternating direction method of multipliers (ADMM) framework and employs a reweighting strategy. The convergence analysis of the proposed algorithm is investigated. Simulated and real experiments show the effectiveness of the proposed algorithm.

Index Terms—hyperspectral images, spectral unmixing, bilateral joint-sparse, low-rank matrix, alternating direction method of multipliers (ADMM), iterative reweighting.

I. INTRODUCTION

SPECTRAL unmixing of hyperspectral images (HSIs) has attracted much attention in different scientific fields [1]–[3]. It is the task of identifying the spectral signatures of distinct materials (*endmembers*) and estimating the fractions (*abundances*) of the materials for each pixel in HSI. The mixing process can be characterized as either linear [1] or nonlinear [4], [5]. In particular, bilinear mixture models have been proposed and used more commonly in practice [6]–[8]. Linear mixture model (LMM) assumes that the measurement spectrum of each pixel is a linear combination of the spectral signatures of the endmembers [3]. Due to its simplicity and tractability, we will adopt the LMM in spectral unmixing for the remainder of this article. Commonly, the abundance vector of a mixed pixel should satisfy the abundance nonnegative

This research is supported in part by NSFC (Grant No. 61772003), in part by Key Projects of Applied Basic Research in Sichuan Province (Grant No. 2020YJ0216), in part by National Key Research and Development Program of China (Grant No. 2020YFA0714001), in part by the Fundamental Research Funds for the Central Universities (Grant No. ZYGX2019J093). (Corresponding author: J. Huang.)

The authors are with School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, PR China (e-mail: hapjie07mo@163.com).

constraint (ANC) and the abundance sum-to-one constraint (ASC) for physical meaning [9].

Nonnegative matrix factorization for spectral unmixing aims to find two nonnegative matrices, one for an endmembers' dictionary and another for fractional abundances, so that their product is equal to the target HSI matrix [10]–[13]. Recently, nonnegative tensor factorization has been proposed for hyperspectral unmixing to further exploit spatial information of HSIs and improve the unmixing performance [14]–[16]. With the available of the spectral dictionaries, sparse unmixing provides an alternative way for spectral unmixing. It assumes that, compared with large-scale dictionaries, only a few of spectral signatures participate in the LMM of each pixel, leading a sparse structure in the abundance matrix [17]. Diverse sparse structures have been exploited in the literature, including the standard ℓ_1 -norm regularization [17], the $\ell_{2,1}$ -norm regularization [18], and their variants [10], [19]–[27].

Exploiting spatial information in HSIs significantly improves the accuracy of spectral unmixing. Typically, the total variation (TV) regularization imposes the piecewise smooth on each abundance map [28]–[31]. It suppresses the noise and generates smooth abundance maps with preserved edges. In many cases, however, it is overstrict to assume that neighboring pixels have both similar mixing endmembers and similar abundance fractions. The low-rank representation provides another perspective of spatial correlation for spectral unmixing [6], [32]–[34]. In this vein, the highly spatial correlation of mixed pixels is transferred into the linear dependence among their corresponding abundance vectors. Simultaneously imposing the low-rank characteristic with sparsity has offered stimulating results [32]. Moreover, it is shown that the unmixing performance can be further improved by adopting a joint-sparse-blocks structure in the low-rank representation based unmixing [26]. In addition, nonlocal spatial information has been utilized to improve the performance of abundance estimation [35]–[40]. Broadly speaking, simultaneously exploiting spatial information and sparsity provides an effective way for spectral unmixing. Notice that most above-mentioned sparse unmixing algorithms first transfer an HSI, a third-order tensor, into a matrix, and then estimate the abundance matrix under the LMM. Thus, spatial information of HSIs might be lost in the transferring process [14]–[16], [41], [42].

In this article, we propose a *bilateral joint-sparse* regression for hyperspectral unmixing to make further use of the spatial information of HSIs. On the one hand, we consider a local joint sparsity structure of the abundance matrix, which is obtained, as usual, by arranging the HSI to be a matrix along the *vertical* direction. On the other hand, we consider the local

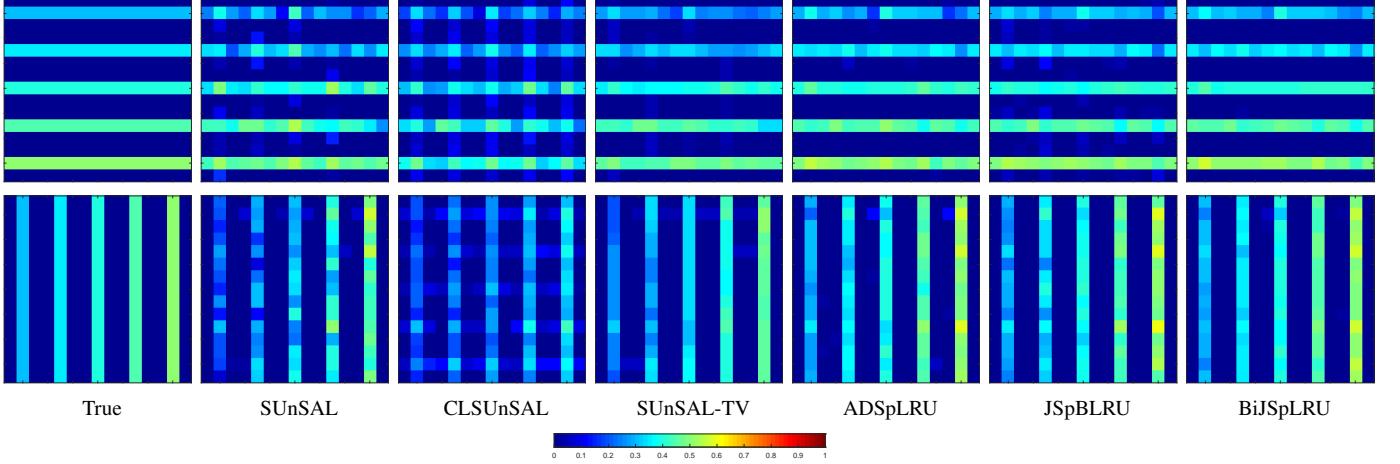


Fig. 4: True and estimated abundance maps for Example 1 with SNR = 30 dB for endmembers (top row) #1 and (bottom row) #2.

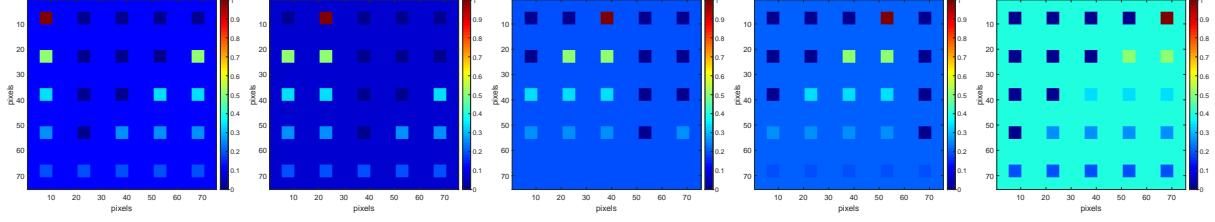


Fig. 5: True abundance maps of selected endmembers for Example 2. From left to right: Endmembers #1–#5.

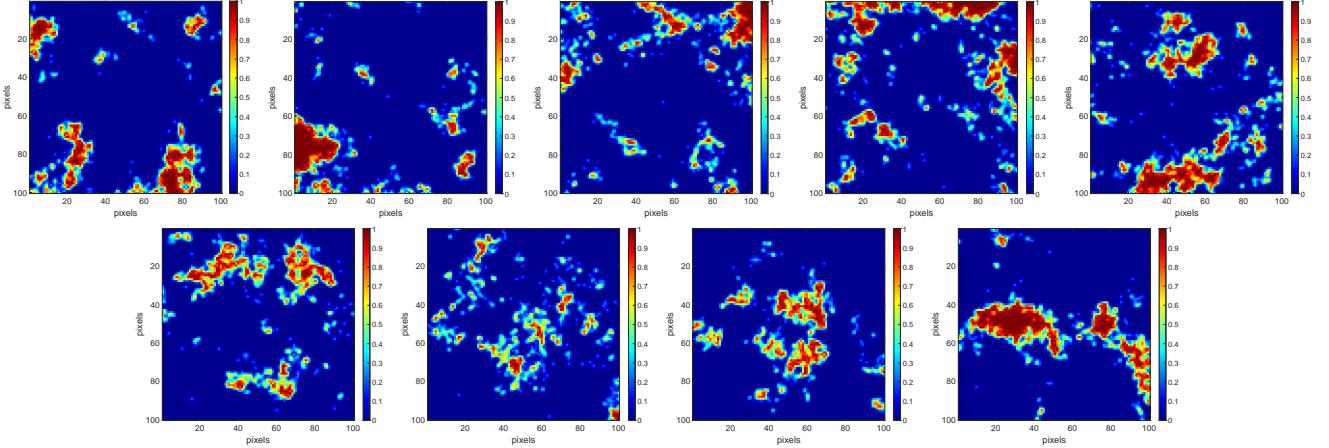


Fig. 6: True abundance maps for selected endmembers for Example 3. Top row: endmembers #1–#5. Bottom row: endmembers #6–#9.

TABLE I: Average of SRE (dB) and RMSE after 20 runs by different unmixing algorithms for Example 1 with SNR = 30 dB.

Criteria	SUN SAL	CLSUnSAL	SUnSAL-TV	ADSpLRU	JSpBLRU	BiJSpLRU
SRE	6.58	7.70	8.45	19.11	19.94	21.17
RMSE	0.0205	0.0180	0.0165	0.0048	0.0044	0.0038

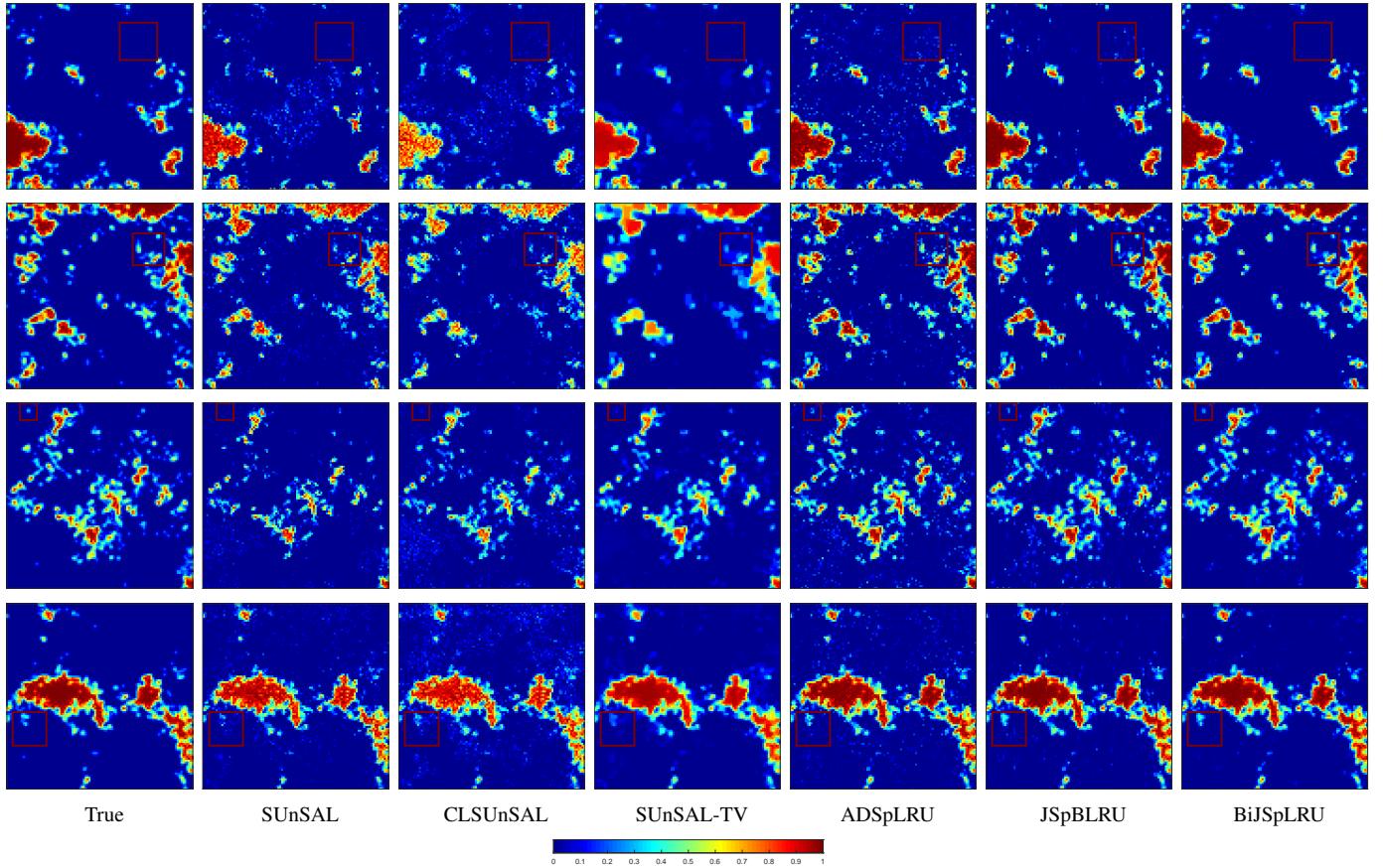


Fig. 8: True and estimated abundance maps by different unmixing algorithms for Example 3 with SNR = 30 dB for (from top to bottom) endmembers #2, #4, #7, and #9.

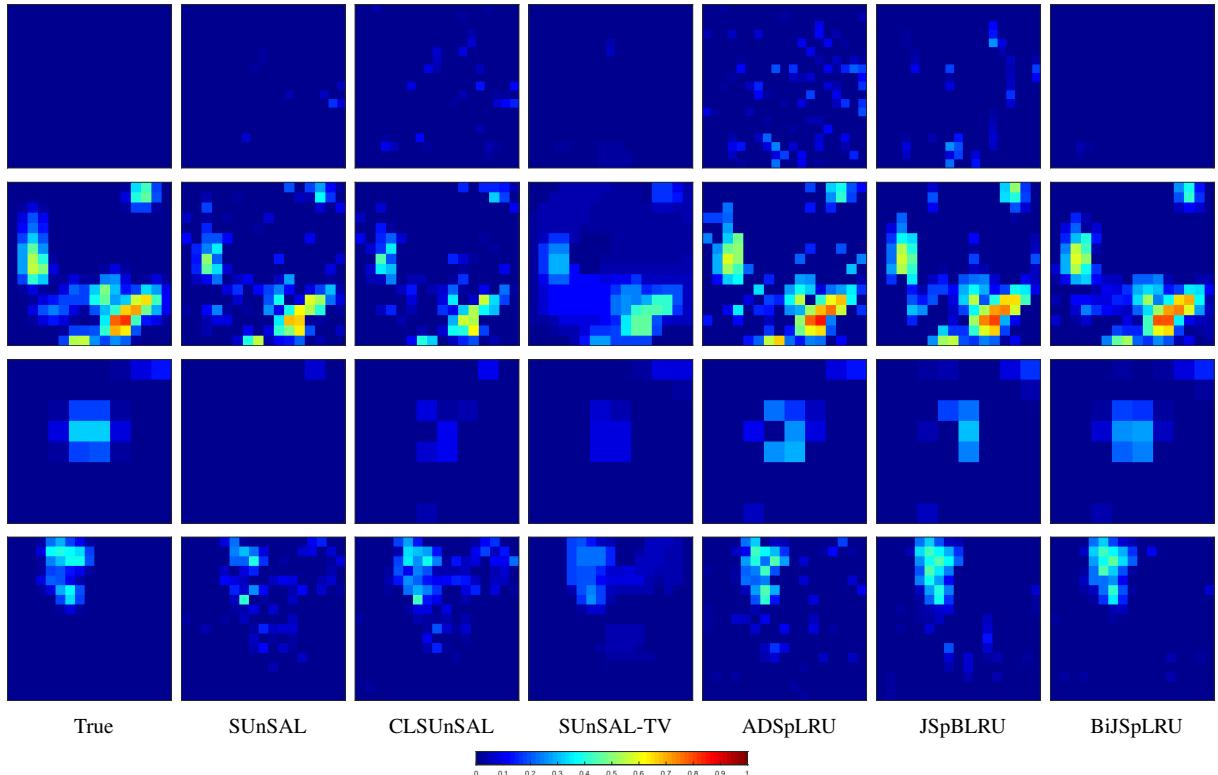


Fig. 9: Detailed regions in red boxes from Fig. 8 for (from top to bottom) endmembers #2, #4, #7, and #9.

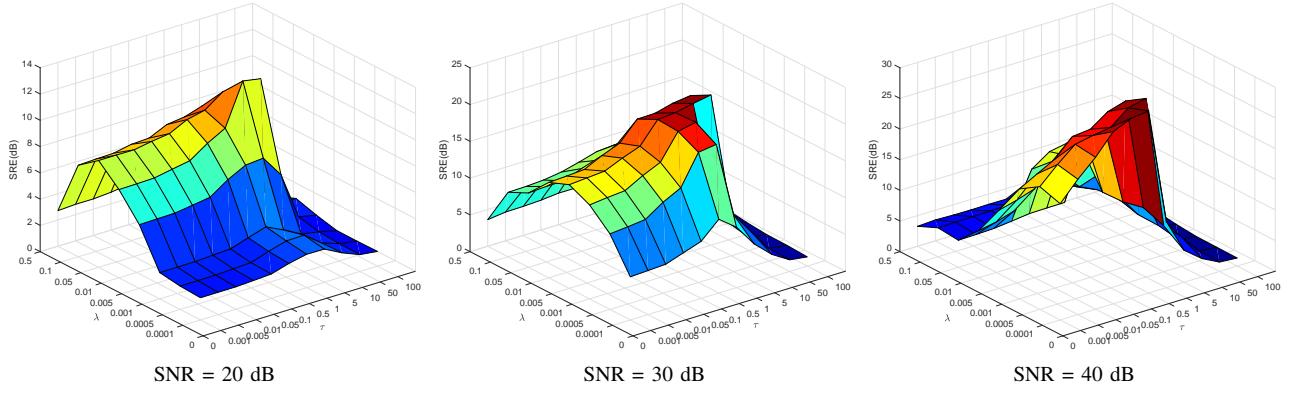


Fig. 10: SRE (dB) as a function of regularization parameters λ and τ in BiJSpLRU for Example 3.

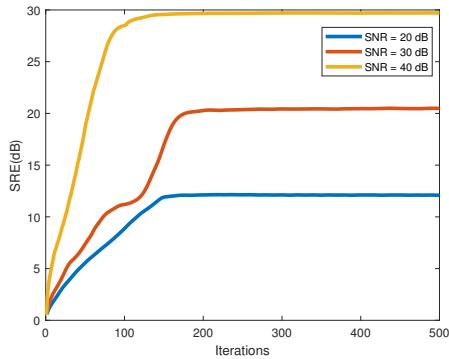


Fig. 11: Plot of SRE (dB) against iteration by BiJSpLRU for Example 3.

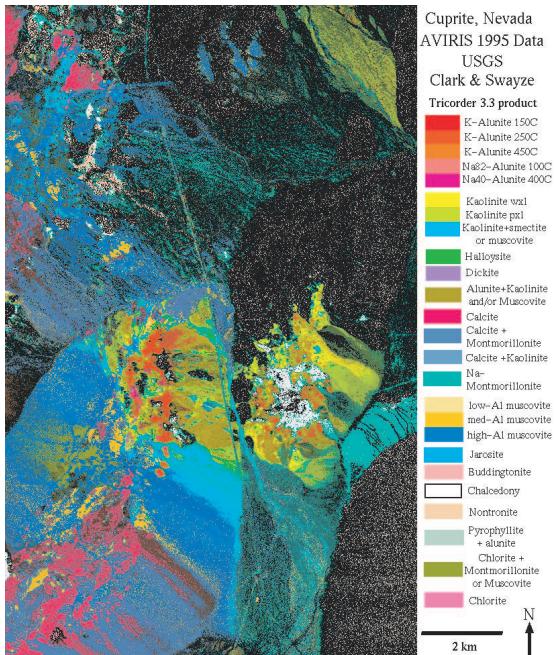


Fig. 12: The USGS map showing the location of different minerals in the Cuprite mining district in Nevada.

a similar conclusion, so we omit here for space considerations. Compared with SUoSAL and CLSUoSAL, SUoSAL-TV maintains the piecewise smooth behavior in the backgrounds, especially for endmember #5. But the abundance map of SUoSAL-TV is over-smooth for endmember #1. Low-rank unmixing algorithms delineate more square regions for endmember #1. Among them, BiJSpLRU produces most square blocks. It can be expected since BiJSpLRU utilizes more spatial information than ADSpLRU and JSpBLRU, especially for pixels in the boundaries.

The abundance maps estimated by different algorithms for Example 3 are shown in Fig. 8. Particularly, the regions in red boxes are zoomed in and displayed in Fig. 9. SUoSAL and CLSUoSAL estimate abundance maps with low accuracy. Note that each of the two algorithms has only one regularization parameter so it is time-saving to select optimal parameter values compared with other algorithms. In addition, they are fast since there is no need to implement the TV or SVD at each iteration. SUoSAL-TV provides piecewise smooth estimations, in which many details are lost. Clearly, the backgrounds of the abundance maps estimated by JSpBLRU are more clear than those by ADSpLRU. However, when we zoom in the abundance maps of JSpBLRU, we can see much banded vertical noise in the backgrounds. The reason is, in part, that JSpBLRU utilizes the local joint-sparsity structure along the vertical direction of the HSI. BiJSpLRU addresses this problem by utilizing both vertical and horizontal spatial information. We can see that BiJSpLRU provides accurate abundance maps with clearer backgrounds. Finally, Table II shows the SRE and RMSE values by different unmixing algorithms for Examples 2 and 3. We see that BiJSpLRU provides the optimal results, which is consistent with the observations from Figs. 7-9.

C. Parameters Discussion

Now we discuss the influence of the regularization parameters λ and τ and the number of iterations of BiJSpLRU. Recall that in the simulated experiments, we test all possible combinations of parameters for BiJSpLRU and choose the optimal ones for best SREs. Fig. 10 plots the SRE values as a function of parameters λ and τ in the model (18) for Example 3. We can see that both optimal parameters decrease

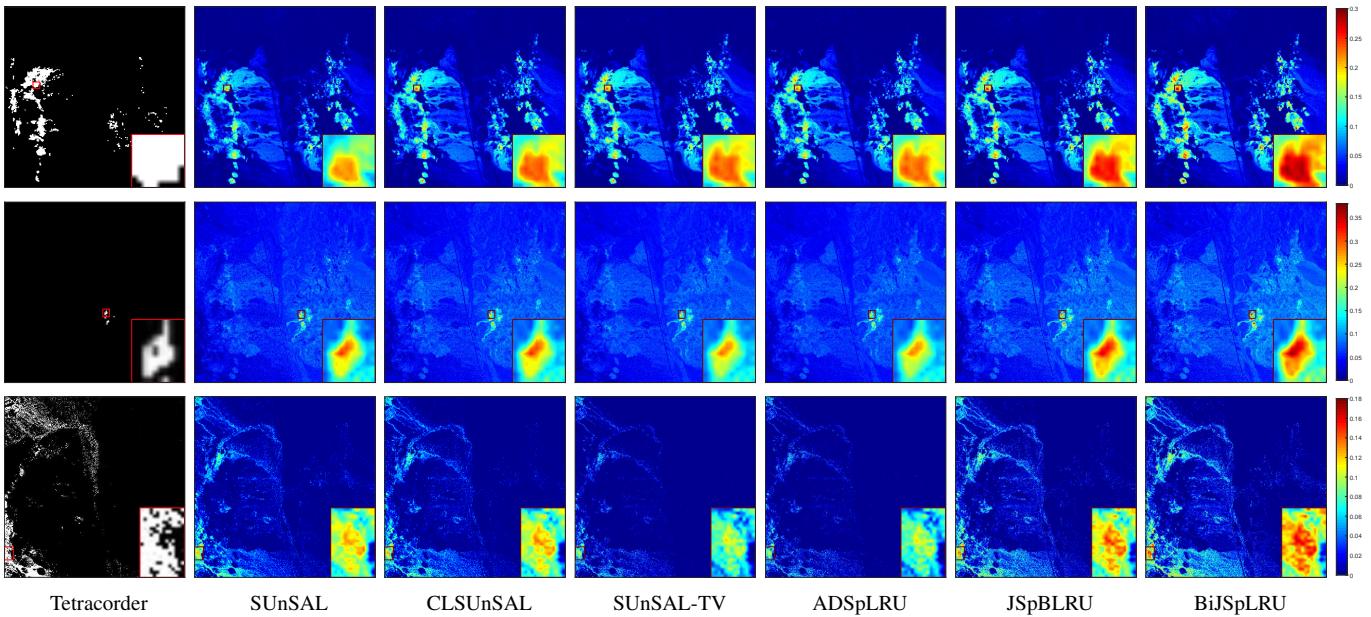


Fig. 13: Estimated abundance maps for (top row) Alunite, (middle row) Buddingtonite, and (bottom row) Muscovite by different unmixing algorithms. The demarcated area is enlarged in the right bottom corner for better visualization.

as SNR increases. Optimal τ is far greater than optimal λ , which shows the effectiveness of the low-rank regularizer.

Fig. 11 plots the SRE values against iteration for Example 3. Clearly for BiJSpLRU, it is enough to set the maximum number of iterations to be 300. In addition, we see from Fig. 11 that for different noise levels, the convergence behavior of BiJSpLRU becomes stable as the iteration increases.

IV. EXPERIMENTS ON REAL DATA

In this test, we show the performance of BiJSpLRU on the well-known Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Cuprite data set. Fig. 12 shows a mineral map produced in 1995 by USGS, in which different minerals were mapped by a Tetracorder software product [55]. We use a subscene of the Cuprite data with 350×350 pixels and 188 spectral bands. The spectral dictionary A of size 188×240 is generated from the USGS spectral library. Similarly as in [17], [18], [28], we set regularization parameters $\lambda = 0.001$ for SUNSAL, $\lambda = 0.01$ for CLSUnSAL, and $\lambda = \lambda_{TV} = 0.001$ for SUNSAL-TV. Similarly as in [26] and [32], we set $\lambda = \tau = 0.001$ for ADSpLRU, JSpBLRU, and BiJSpLRU.

Fig. 13 shows the estimated abundance maps obtained by different unmixing algorithms. Due to the low noise of the real data, all the algorithms produce overall similar results. From the enlarged area in the right bottom corner in Fig. 13, however, we can see that BiJSpLRU gives comparable or higher estimates in the regions considered as respective materials. It shows again the effectiveness of BiJSpLRU by simultaneously utilizing the bilateral joint sparsity structure and the low rankness of the abundance matrix.

V. CONCLUSION

In this paper, we have proposed a bilateral joint-sparse structure to utilize the vertical and horizontal spatial information

of HSIs. In particular, we realize the bilateral joint-sparse structure by introducing a permutation matrix, which is no need to be constructed explicitly. Then we have proposed an unmixing model by imposing the bilateral joint-sparse structure and the low-rank property on the abundance matrix, via the weighted $\ell_{2,1}$ -norm and the weighted nuclear norm, respectively. We solve the proposed model under the ADMM framework with an iterative reweighting strategy, leading to an algorithm called BiJSpLRU. The residual and objective convergence results of BiJSpLRU have been provided. The simulated and real-data experiments show the effectiveness of the proposed algorithm, compared with other state-of-the-art unmixing algorithms.

The proposed bilateral joint-sparse structure has the potential to contribute other image processing problems via sparse representation, such as (hyperspectral) image restoration, super-resolution, blind hyperspectral unmixing, etc. In addition, the article provides a way to obtain theoretical convergence results of algorithms which are based on the ADMM framework but with a reweighted strategy.

APPENDIX

In order to prove Theorem 1 about the convergence of the sequence $(X^k, V^k, W^k, \Lambda^k)$ by (26), we first give three lemmas.

Lemma 1. Let $(X^k, V^k, W^k, \Lambda^k)$ be the sequence generated by (26). Then under Assumption 1, we have

$$p^* - p^{k+1} \leq \langle \Lambda^*, r^{k+1} \rangle. \quad (47)$$

Proof. Recall from (43), we have

$$\mathcal{L}_0(X^*, V^*, W^*, \Lambda^*) \leq \mathcal{L}_0(X^{k+1}, V^{k+1}, W^{k+1}, \Lambda^*). \quad (48)$$

- J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5009–5022, 2019.
- [28] D. Iordache, J. Bioucas-Dias, and A. Plaza, “Total variation spatial regularization for sparse hyperspectral unmixing,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, 2012.
- [29] X.-L. Zhao, F. Wang, T.-Z. Huang, M. Ng, and R. Plemmons, “Deblurring and sparse unmixing for hyperspectral images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4045–4058, 2013.
- [30] X.-L. Zhao, W. Wang, T.-Y. Zeng, T.-Z. Huang, and M. Ng, “Total variation structured total least squares method for image restoration,” *SIAM J. Sci. Comput.*, vol. 35, pp. 1304–1320, 2013.
- [31] X. Li, J. Huang, L.-J. Deng, and T.-Z. Huang, “Bilateral filter based total variation regularization for sparse hyperspectral image unmixing,” *Inf. Sci.*, vol. 504, pp. 334–353, 2019.
- [32] P. Giampouras, K. Themelis, A. Rontogiannis, and K. Kourtoumbas, “Simultaneously sparse and low-rank abundance matrix estimation for hyperspectral image unmixing,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4775–4789, 2016.
- [33] R. Wang, W. Liao, H. Li, H. Zhang, and A. Pižurica, “Hyperspectral unmixing by reweighted low rank and total variation,” in *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Los Angeles, California, USA, Aug. 21–24, 2016.
- [34] J. Yang, Y.-Q. Zhao, J. Chan, and S. Kong, “Coupled sparse denoising and unmixing with low-rank constraint for hyperspectral image,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1818–1833, 2016.
- [35] Y. Zhong, R. Feng, and L. Zhang, “Non-local sparse unmixing for hyperspectral remote sensing imagery,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1889–1909, 2014.
- [36] R. Feng, Y. Zhong, and L. Zhang, “An improved nonlocal sparse unmixing algorithm for hyperspectral imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 4, pp. 915–919, 2015.
- [37] R. Feng, Y. Zhong, and L. Zhang, “Adaptive non-local Euclidean medians sparse unmixing for hyperspectral imagery,” *ISPRS J. Photogram. Remote Sens.*, vol. 97, pp. 9–24, 2014.
- [38] R. Wang, H.-C. Li, W. Liao, X. Huang, and W. Philips, “Centralized collaborative sparse unmixing for hyperspectral images,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 1949–1962, 2017.
- [39] L. Sun, F. Wu, T. Zhan, W. Liu, J. Wang, and B. Jeon, “Weighted nonlocal low-rank tensor decomposition method for sparse unmixing of hyperspectral images,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1174–1188, 2020.
- [40] J. Huang, T.-Z. Huang, X.-L. Zhao, and L.-J. Deng, “Nonlocal tensor-based sparse hyperspectral unmixing,” *IEEE Trans. Geosci. Remote Sens.*, pp. 1–15, 2020.
- [41] G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [42] H. A. L Kiers, “Towards a standardized notation and terminology in multiway analysis,” *J. Chemometr.*, vol. 14, no. 3, pp. 105–122, 2000.
- [43] E. Candès, M. Wakin, and S. Boyd, “Enhancing sparsity by reweighted ℓ_1 minimization,” *J. Fourier Anal. Appl.*, vol. 14, no. 5–6, pp. 877–905, 2008.
- [44] X. Fu, W.-K. Ma, J. Bioucas-Dias, and T.-H. Chan, “Semiblind hyperspectral unmixing in the presence of spectral library mismatches,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5171–5184, 2016.
- [45] S. Gu, L. Zhang, W. Zuo, and X. Feng, “Weighted nuclear norm minimization with application to image denoising,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2014, pp. 2862–2869.
- [46] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd Edition, Cambridge University Press, New York, USA, 2013.
- [47] J. Bioucas-Dias and M. A. T. Figueiredo, “Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing,” in *Proc. 2nd Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, pp. 1–4, 2010.
- [48] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, “Weighted nuclear norm minimization and its applications to low level vision,” *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, 2017.
- [49] S. Wright, R. Nowak, and M. Figueiredo, “Sparse reconstruction by separable approximation,” *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, 2009.
- [50] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, Elsevier, Amsterdam, North-Holland, 1983.
- [51] P. Tseng, “Applications of a splitting algorithm to decomposition in convex programming and variational inequalities,” *SIAM J. Control Optim.*, vol. 29, pp. 119–138, 1991.
- [52] J. Eckstein and D. P. Bertsekas, “On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Math. Program.*, vol. 55, pp. 293–318, 1992.
- [53] M. Fukushima, “Application of the alternating direction method of multipliers to separable convex programming problems,” *Comput. Optim. Appl.*, vol. 1, pp. 93–111, 1992.
- [54] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
- [55] R. Clark *et al.*, “Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems,” *J. Geophys. Res.*, vol. 108, no. E12, pp. 5131–5135, 2003.