# Joint Reconstruction and Spectral Unmixing from Single-Pixel Acquisitions

Séréna Hariga, Jérémy E Cohen, Nicolas Ducros

# Joint Reconstruction and Spectral Unmixing from Single-Pixel Acquisitions

Séréna Hariga*, Jérémy E. Cohen*, Nicolas Ducros*†
*Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne,
CNRS, Inserm, CREATIS UMR 5220, U1294, F-69621, LYON, France
†Institut universitaire de France (IUF), France

***Abstract*—Single-pixel imaging enables the acquisition of hyperspectral data across thousands of spectral channels at a low cost. After acquisition of the spectrum of scalar products of the scene with several light patterns, the hypercube of the scene is reconstructed by solving an inverse problem. Then, the hypercube can be decomposed into a few material maps and their corresponding spectral signatures. In this paper, we propose and discuss alternatives to the straightforward two-step approach. In particular, we propose a joint approach that recovers the material maps and spectral signatures directly from the measurements. The joint problem can be solved by tailoring existing nonnegative matrix factorization tools to single-pixel imaging. We demonstrate that the proposed joint method provides more accurate estimation of the spectra and material maps in simulations at different noise levels and sparsity levels.**

## I. Introduction

A single-pixel camera measures with a single point detector a sequence of scalar products between a scene and some light patterns uploaded onto a digital micro-mirror device (DMD) [1]. The image of the scene can be reconstructed from the measurements by solving an inverse problem (e.g., based on $\mathcal{L}_1$-minimization [2], [3] or deep learning techniques [4], [5]). In the case of hyperspectral imaging, the point detector is a spectrometer that acquires up to a few thousand wavelengths that can be reconstructed independently [6], [7].

The hyperspectral image often originates from the mixture of the response of a few materials that are present in the scene. Several unmixing methods are suggested in the literature [8]. A simple and widely-used method is the linear mixing model. When the spectral signatures are unknown, nonnegative matrix factorization (NMF) algorithms are required to estimate both the spectral signatures and material maps at each pixel [9], [10].

In this article, we present a family of strategies for estimating the material maps and spectral signatures. Among them, two-step methods are considered the most straightforward, leveraging the availability of standard methodologies in the field of spectral unmixing and single-pixel image reconstruction. However, these methods face limitations, which we aim to pinpoint in this article. In this context, we propose a joint approach that recovers the material maps and spectral signatures directly from the measurements. This kind of method has been proposed in other imaging domains, such as in Spectral CT [11], but not for single-pixel imaging yet, to the best of our knowledge. Our method relies on existing nonnegative matrix factorization tools tailored for single-pixel imaging. In particular, we consider alternating multiplicative updates that ensures nonnegativity of the solutions. In Section II, we model hyperspectral single-pixel imaging. In Section III, we discuss different strategies for the estimation of the material maps and spectral signatures. In Section IV, we introduce a joint method that we evaluate in the simulations described in Section V. Section VI reports our results.

## II. Hyperspectral single-pixel imaging model

Let $Y \in \mathbb{R}^{\Lambda \times M}$ be the raw single pixel measurements, where $M$ represents the number of patterns uploaded onto the DMD and $\Lambda$ the number of spectral channels provided by the spectrometer. We model single-pixel hyperspectral acquisition by

$$Y \sim \mathcal{P}(\alpha X A^T), \tag{1}$$

where $\mathcal{P}$ represents the Poisson distribution, $A \in \mathbb{R}^{M \times N}$ is the measurement matrix where each row represents a DMD pattern with $N$ pixels, $X \in \mathbb{R}_+^{\Lambda \times N}$ is the flattened hypercube of the scene and $\alpha$ is its intensity (in photons). Note that $\alpha$ sets the signal-to-noise ratio of the measurements and a higher value indicates better quality. Assuming further that the hypercube results from $R$ distinct components, we have

$$X = WH, \tag{2}$$

where $W \in \mathbb{R}_+^{\Lambda \times R}$ represents the spectral signature of each component and $H \in \mathbb{R}_+^{R \times N}$ represents the corresponding material maps. Finding matrices $W$ and $H$ from Eq. 2 corresponds to the NMF problem with rank $R$. NMF being an NP-hard problem [12], it is challenging to find the best solution efficiently. Moreover, NMF does not have generally a unique solution, making it an ill-posed problem [10]. Another challenge in NMF is the choice of the rank $R$, which remains an open problem [13], [9]. In the following, we assume that the rank is known, and we set $R$ such that $R < \min(\Lambda, N)$.

## III. Overview of the estimation strategies

Our goal is to estimate the material maps $H$ and spectral signatures $W$ from the single-pixel measurement $Y$, for which we identify and describe three main strategies.

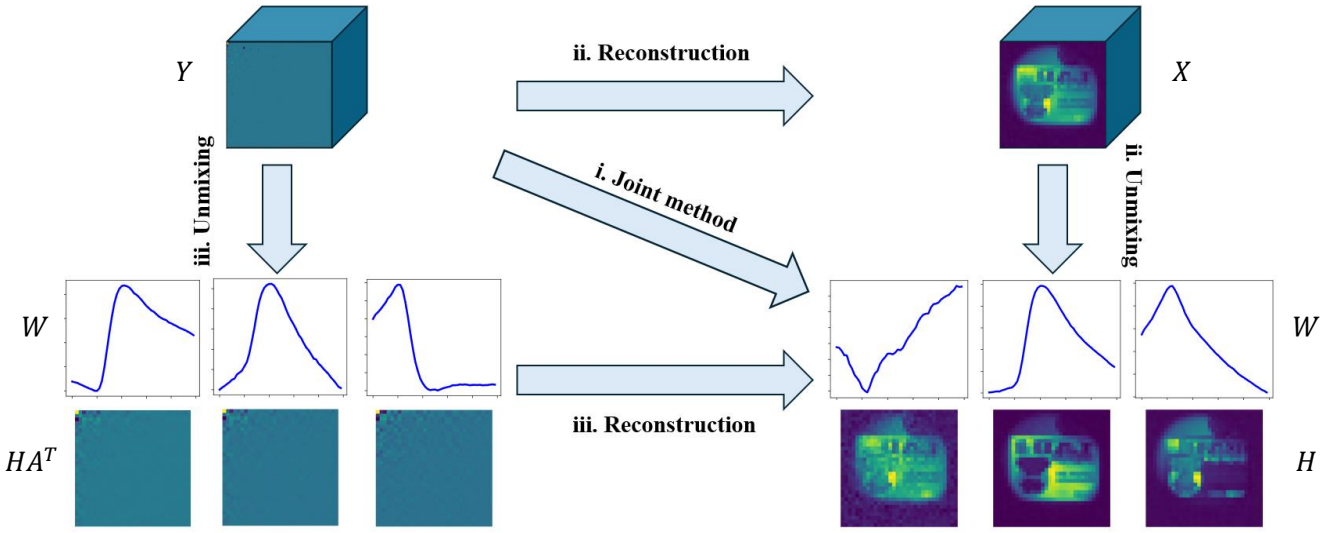



Fig. 1: Overview of the strategies for the estimation of $H$ and $W$. **(i)** Joint method. **(ii)** Reconstruct then unmix (RU). **(iii)** Unmix then reconstruct (UR).

### A. Reconstruct then unmix (RU)

A straightforward method is to reconstruct the hypercube from the raw data first, and then to unmix the obtained hypercube. This allows to reuse the standard tools in the field of spectral unmixing and single-pixel image reconstruction. For image reconstruction, a common approach involves maximizing the likelihood of a Poisson distribution, which is equivalent to minimizing the Kullback-Leibler divergence

$$X^* := \underset{X \in \mathbb{R}_+^{\Lambda \times N}}{\operatorname{argmin}} \; D_{KL}(Y, \alpha X A^T), \tag{3a}$$

where $D_{KL}$ is the Kullback-Leibler divergence applied element-wise. This is a well-known optimization problem for which various algorithms are available [14]. For spectral unmixing, one possibility is to consider NMF with the L2 norm

$$W^*, H^* := \underset{W \in \mathbb{R}_+^{\Lambda \times R},\; H \in \mathbb{R}_+^{R \times N}}{\operatorname{argmin}} \; \|X^* - WH\|_F^2. \tag{3b}$$

While other cost functions can be considered, choosing the L2 norm allows to benefit from efficient algorithms such as the Hierarchical Alternating Least Squares (HALS) [15]. However, it is not clear whether this choice is optimal, which constitutes the first limitation of the RU method. Another limitation is that the unmixing step is performed on $X^*$. Matrix $X^*$ is typically full rank, and computed to solve a different problem than the maximum likelihood estimator (3a), so the final cost value after spectral unmixing is likely to be suboptimal and not all the information is leveraged.

### B. Unmix then reconstruct (UR)

By noticing that Eq. 1 and Eq. 2 are equivalent to $Y \sim \mathcal{P}(\alpha WG)$ with $G = HA^T$, one can first unmixing the spectra $W$ and materials maps in the transformed domain $G$ from the raw measurements $Y$, before reconstructing $H$ from $G$. Again, standard tool can be considered for both the unmixing and reconstruction problems.

In the absence of noise (i.e., $\alpha = \infty$), the unmixing step consists in finding $W \in \mathbb{R}_+^{\Lambda \times R}$ and $G \in \mathbb{R}_+^{R \times M}$ such that $Y = WG$. This problem can be geometrically expressed as finding $R$ vectors in the nonnegative orthant, such that the cone generated by these vectors contains the columns of $Y$. To be able to identify a solution, it is preferable that the hyperspectral data are sufficiently scattered over the unit simplex [10]. The so-called Sufficiently Scattered Conditions (SSC) are illustrated in Fig. 2 in the simple case $R = N = 3$ when $A$ corresponds to the Hadamard patterns defined in Section V-A. It show that even if the hypercube satisfies the SSC in the image domain, it may no longer satisfies the SSC in the Hadamard domain. Therefore, one cannot expect NMF to be unique in the Hadamard domain and the UR strategy is not suitable.

### C. Joint method

To overcome the limitations associated with the two-step methods, the material maps and spectral signatures can be recovered directly from the measurements by solving

$$\min_{W \in \mathbb{R}_+^{\Lambda \times R},\; H \in \mathbb{R}_+^{R \times N}} D_{KL}(Y, \alpha WHA^T). \tag{4}$$

This problem can be seen as a generalization of NMF. Note that the RU strategy given by Eq. 3 differs from the joint strategy of Eq. 4. In particular, the latter formulation guarantees that the likelihood of the measurements resulting from the decomposed $W$ and $H$ is maximised. As far as the RU strategy is concerned, the reconstructed hypercube $X$ maximises the likelihood of the measurements but in general it is not low-rank and therefore does not minimize Eq. 4.

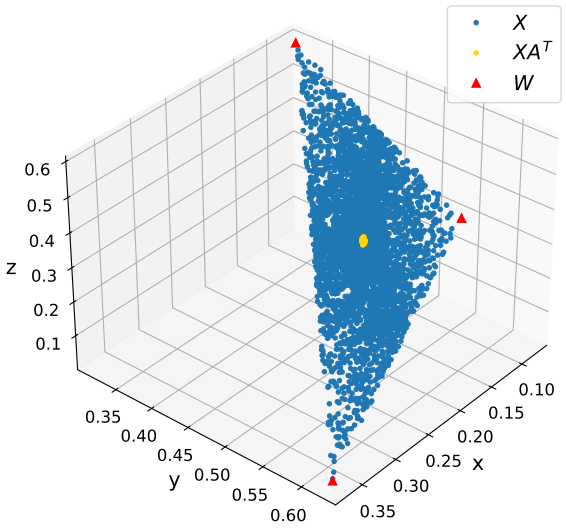



Fig. 2: Sufficiently scattered conditions (SSC). In blue: hypercube $X$ in the image domain which is simulated as the product of two matrices $W$ and $H$, following a uniform distribution and normalized by the L1 norm. In yellow: hypercube after Hadamard transformation $XA^T$. In red: columns of $W$. This shows that data in the Hadamard domain is concentrated, resulting in non-unique NMF decomposition as there exists an infinity of cones containing the yellow data and included in the nonnegative orthant. Even if the hypercube satisfies the SSC in the image domain, it may no longer satisfies the SSC in the Hadamard domain.

## IV. Proposed joint method

In this section, we present an algorithm to compute solutions to the joint estimation problem (4). This problem is not convex in $W$ and $H$ simultaneously but it is with respect to $W$ when $H$ is fixed and reciprocally. A workhorse class of solvers for such bi-convex problems is alternating minimization algorithms, that estimate alternately each matrix $W$ and $H$ until convergence, while ensuring the nonnegativity constraints are satisfied. When $H$ is fixed, a widely-used update for matrix $W$ is given by the Multiplicative Update (MU) rule [16], which is easily adapted to handle the Hadamard measurement matrix,

$$W \longleftarrow \max\left(W \circ \frac{\frac{Y}{WHA^T}AH^T}{\alpha \mathbb{1}_M AH^T},\ \epsilon\right), \tag{5}$$

where $\circ$ denotes the element-wise product and $\mathbb{1}_M$ is a vector of size $M$ composed of 1. When $W$ is fixed, the following MU updates for $H$ have been proposed in the literature [17],

$$H \longleftarrow \max\left(H \circ \frac{W^T\frac{Y}{WHA^T}A}{\alpha W^T \mathbb{1}_{\Lambda\times M} A},\ \epsilon\right), \tag{6}$$

where $\mathbb{1}_{\Lambda\times M}$ is a matrix of size $\Lambda \times M$ with all entries set to 1. Note that a projection with a small $\epsilon$ is performed to ensure convergence of the MU updates [10], [18], [19]. In practice, we iterate the update of each factor several times before switching to the update of the other factor.

## V. Experiments

This section presents a series of experiments aimed at evaluating the performance of the following algorithms :

**RU :** This algorithm follows the RU method explained in Section III-A. It consists of one reconstruction step computed by MU and one unmixing step performed by HALS.

**Joint :** This algorithm is based on the MU outlined in Section IV.

The purpose of these experiments is to study the influence of noise level and sparsity on the estimation of $W$ and $H$. To this end, the simulation is conducted for different noise levels and varying sparsity percentages on $H$. We conduct a comparison between the RU and joint methods to assess whether the proposed approach provides more accurate estimations as expected.

### A. Choice of patterns

An efficient choice for $A$ is to consider the Hadamard matrix of order $N$ denoted as $\mathcal{H}_N$, which satisfies $\mathcal{H}_N\mathcal{H}_N^T = NI_N$ where $I_N$ is the identity matrix of size $N$. Opting for the Hadamard matrix allows reducing the variance of an image by a factor of $\sqrt{N}$, compared to the case where the matrix $A$ is the identity matrix, making it an optimal choice [20]. Each pattern loaded onto the DMD corresponds to a row of this matrix. The Hadamard matrix has negative entries, but a DMD can only load nonnegative patterns. To overcome this challenge, one possibility is to split the patterns into two sets of matrices with nonnegative entries, representing the nonnegative and negative components of the Hadamard basis [21]. Matrix containing the patterns is denoted $A \in \mathbb{R}^{2N-1\times N}$ such that

$$A = \begin{pmatrix} \mathcal{H}_N^+ \\ \mathcal{H}_N^- \end{pmatrix}, \tag{7}$$

where $\mathcal{H}_N^+$ and $\mathcal{H}_N^-$ are respectively the positive and negative parts of the Hadamard matrix. In practice, the first row of the negative part is removed because it is entirely composed of zeros.

### B. Simulation description

We propose to simulate a hypercube $X$ of rank $R = 3$ that is the product of two matrices $W$ and $H$ containing respectively the spectra and the material maps. We simulate $R$ spectra of size $\Lambda = 50$ as Gaussian distributions. To avoid spectra being too correlated, we sample the Gaussian means uniformly between five and forty-five with a minimum separation of seven. The standard deviations are sampled from a uniform distribution ranging between zero and four. To simulate material maps, we choose $R$ different images of size $N = 16 \times 16$ from the STL-10 database. Additionally, we impose a sparsity constraint on $H$ to enhance the likelihood of having a unique NMF [22], [23]. We set $\beta\%$ of the smallest pixels of each image to zero. We introduce Poisson noise to the simulated hypercube as described in Eq. 1. We adjust the parameters $\alpha$ and $\beta$ to capture various levels of noise and sparsity. This simulation is conducted on one hundred hypercubes for each parameter combination of $\alpha$ and $\beta$.

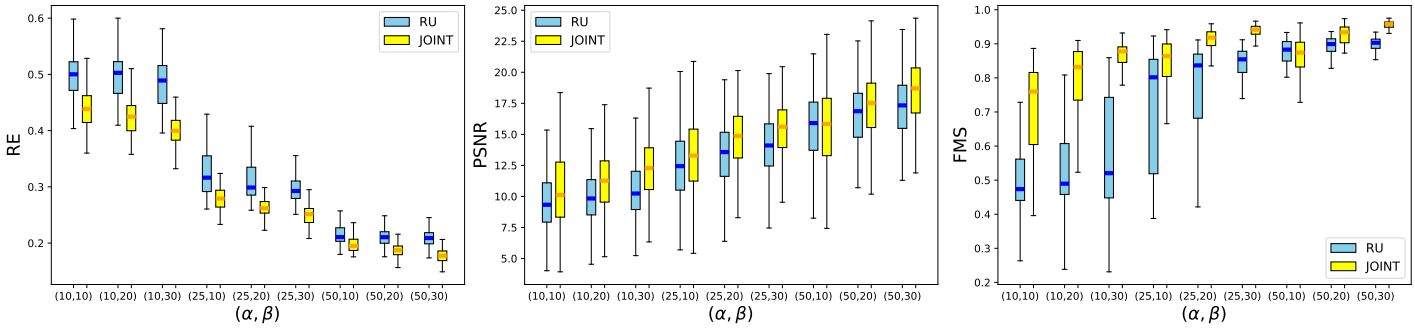



Fig. 3: Boxplot of the distribution of different metrics on the estimation of spectra and material maps using the RU and joint methods with different combination of values $(\alpha, \beta)$. On the left: the Relative Error on the estimation of the hypercube $X$. On the middle: the PSNR on the estimation of the material maps $H$. On the right: the Factor Match Score for the estimation of the spectra and material maps pair $(W, H)$.

### C. Choice of hyperparameters

To initialize the algorithms, we use the least squares solution $\tilde{X} = YA^{\dagger}$, where $A^{\dagger}$ represents the pseudo-inverse of $A$. To acquire the initializations of $W$ and $H$, we compute a truncated Singular Value Decomposition (SVD) with a rank of $R$ on $\tilde{X}$ and apply element-wise absolute values to ensure nonnegativity.

**RU:** For the reconstruction, we perform 2000 iterations of the MU algorithm until convergence. To address the unmixing, the maximal number of iterations for HALS is set to 50 and the number of inner iterations is fixed to 20, after which convergence is typically observed in our setting.

**Joint:** We fix 2000 iterations and 10 inner iterations.

In both cases the parameter $\epsilon$, employed to guarantee updates convergence, is expected to be small. Therefore, we set $\epsilon$ to a value of $10^{-9}$.

### D. Evaluation metrics

The Relative Error (RE) quantifies the accuracy of the hypercube $X$ estimation and can be defined as

$$\mathrm{RE}(X, \widehat{X}) = \frac{\|X - \widehat{X}\|_2}{\|X\|_2},$$

where $\widehat{X}$ is the estimated hypercube.

The reconstruction quality of the material maps $H$ is measured, for each row of $H$, by the Peak Signal-to-Noise Ratio (PSNR). The PSNR between the original image $h$ and the estimated image $\hat{h}$ is defined by

$$\mathrm{PSNR}(h, \hat{h}) = 20 \log_{10} \left( \frac{1}{\|h - \hat{h}\|_2} \right).$$

The Factor Match Score (FMS) is used to measure the similarity between the couples of matrices $(W, H)$ and $(\widehat{W}, \widehat{H})$,

$$\mathrm{FMS}\big((W, H), (\widehat{W}, \widehat{H})\big) = \frac{1}{R} \sum_{r=1}^{R} \frac{w_r^T \hat{w}_r}{\|w_r^T\| \|\hat{w}_r\|} \frac{h_r^T \hat{h}_r}{\|h_r^T\| \|\hat{h}_r\|},$$

where $\widehat{W}$ and $\widehat{H}$ are respectively $W$ and $H$ estimates. Note that the ordering may differ between the true and the estimated components. To resolve this indeterminacy, we use a linear sum assignment solver to compute the optimal permutation.

## VI. Results

### A. Statistical results

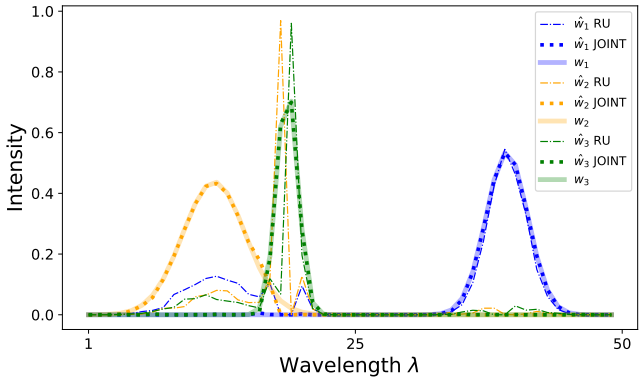
Simulation results are presented in Figure 3, illustrating different metrics (RE, FMS, PSNR) for different pairs of $\alpha$ and $\beta$ values.

We observe that as the noise level increases (smaller $\alpha$), the discrepancy in the results becomes significant, favoring the joint method. This is especially noticeable when examining the FMS for $\alpha = 10$, where the RU method yields very low FMS values (around 0.5), while the joint method provides an improvement of approximately 0.3 to 0.4 in FMS. The difference in terms of PSNR and RE is less significant, but the joint method consistently provides improvements in most cases.

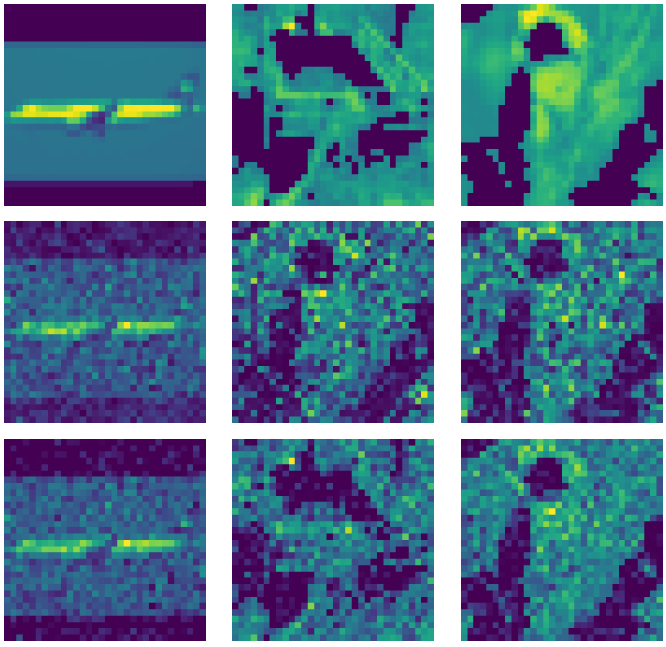
Sparsity does not seem to have a substantial impact on the results obtained with RU. Indeed, the PSNR, RE, and FMS results are similar regardless of the $\beta$ value for the RU method. Sparsity has a noticeable effect on the joint method, especially when examining the FMS. For example, with $\alpha = 10$, there is an increase of approximately 0.15 in FMS, when comparing $\beta = 10$ and $\beta = 30$.

### B. Qualitative results

The simulation results are illustrated in Fig. 4 in the case where $\alpha = 25$ and $\beta = 30$. Fig. 4a displays spectra decomposition using the RU and joint algorithms. We observe that the spectra estimated by the joint method better fit the original spectra compared to the RU method, especially where spectra overlap. In this example, a component has not been found by the RU method, which may explain the significant variability in FMS and poor values in certain cases. Fig. 4b illustrates the material maps reconstructed by the two methods. The difference is less significant in the images than in the spectra, but it can still be observed that the joint method slightly improves the reconstruction of material maps compared to the RU method.




(a) Decomposed spectra using the RU and joint methods. The RU method fails to detect the yellow component.



(b) Decomposed material maps. Top row: ground truth material map. Middle row: decomposition using the RU method. Bottom row: decomposition using the joint method.

Fig. 4: Spectral decomposition for $R = 3$ components. We consider $N = 32 \times 32$ pixels, $\alpha = 25$ photons and $\beta = 30\%$ of the pixels of each material map is zero.

## VII. Conclusion

This paper provides an overview of the strategies for the estimation of material maps and spectral signatures in the context of hyperspectral single-pixel imaging. We demonstrate through simulations that the joint method generally provides a better estimation of material maps and spectral signatures compared to the two-step method, especially in scenarios with high noise and sparsity. One limitation of this work is that it does not investigate the case where the spectra are highly correlated, making the decomposition problem more challenging. In future work, we plan to incorporate regularization in the optimization problem to achieve finer control over both material maps and spectral signatures.

## Acknowledgment

This work is supported financially by ANR JCJC LoRAiA ANR-20-CE23-0010 and ANR ULHYB ANR-22-CE19-0030-01.

## References

[1] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, *Single-pixel imaging via compressive sampling*. IEEE signal processing magazine, 2021, vol. 25, no. 2.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, p. 1289–1306, 2006.

[3] M. P. Edgar, G. M. Gibson, and M. Padgett, "Principles and prospects for single-pixel imaging," *Nature photonics*, vol. 13, no. 1, pp. 13–20, 2019.

[4] C. Higham, R. Murray-Smith, M. Padgett, and M. Edgar, "Deep learning for real-time single-pixel video," *Scientific Reports*, 2018.

[5] A. Lorente Mur, P. Leclerc, F. Peyrin, and N. Ducros, "Single-pixel image reconstruction from experimental data using neural networks," *Optics Express*, vol. 29, no. 11, pp. 17 097–17 110, 2021.

[6] T. Sun, , and K. Kelly, "Compressive Sensing Hyperspectral Imager," *Optics Express*, vol. 26, no. 6, 2009.

[7] G. B. Martins, L. Mahieu-Williame, T. Baudier, and N. Ducros, "OpenSpyrit: an ecosystem for open single-pixel hyperspectral imaging," *Optics Express*, vol. 31, no. 10, p. 15599, 2023.

[8] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches," *IEEE Journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, 2012.

[9] N. Gillis, "The Why and How of Nonnegative Matrix Factorization," *Machine Learning and Pattern Recognition Series*, 2014.

[10] ——, *Nonnegative Matrix Factorization*. Society for Industrial and Applied Mathematics, 2020.

[11] M. Zeegers, A. Kadu, and K. J. Batenburg, "ADJUST: A Dictionary-based Joint reconstruction and Unmixing method for Spectral Tomography," *Inverse Problems 38.12*, p. 125002, 2022.

[12] S. A. Vavasis, "On the complexity of Nonnegative Matrix Factorization," *SIAM Journal on Optimization*, 2008.

[13] J. Bioucas-Dias and J. Nascimento, "Estimation of signal subspace on hyperspectral data," *Remote Sensing*, 2005.

[14] M. Voichita, Y. Feng, H. Banjak, and E. Bretin, "Tomographic reconstruction from Poisson distributed data: a fast and convergent EM-TV dual approach," *Springer Nature*, p. 15599, 2023.

[15] A. Cichocki, R. Zdunek, and S. Amari, "Hierarchical ALS algorithms for Nonnegative Matrix and 3d Tensor Factorization," *Lecture Notes in Computer Science*, p. 169–176, 2007.

[16] D. D. Lee and H. S. Seung, "Algorithms for Nonnegative Matrix Factorization," *Advances in Neural InformationProcessing*, vol. 13, 2001.

[17] C. Févotte and J. Idier, "Algorithms for Nonnegative Matrix Factorization with the $\beta$-divergence," *Neural Computation*, 2011.

[18] N. Gillis and F. Glineur, "Nonnegative factorization and the maximum edge biclique problem," *Center for Operations Research and Econometrics (CORE)*, 2008.

[19] N. Takahashi and R. Hibi, "Global convergence of modified multiplicative updates for nonnegative matrix factorization," *Computational Optimization and Applications*, p. 246, 2014.

[20] N. Ducros, "Une introduction à l'imagerie computationnelle monodétecteur," *ISTE Editions*, 2019.

[21] A. Lorente Mur, M. Ochoab, J. E. Cohen, X. Intesb, and N. Ducros, "Handling negative patterns for fast single-pixel lifetime imaging," *SIAM Journal on Optimization*, 2008.

[22] D. Donoho and V. Stodden, "When Does Non-Negative Matrix Factorization Give a Correct Decomposition into Parts?" *Neural Information Processing Systems*, 2003.

[23] N. Gillis, "Sparse and Unique Nonnegative Matrix Factorization through data preprocessing," *Journal of Machine Learning Research*, vol. 13, 2012.