Hyperspectral Image Unmixing with Endmember Bundles and Group Sparsity Inducing Mixed Norms

Lucas Drumetz, *Member, IEEE*, Travis R. Meyer, Jocelyn Chanussot, *Fellow, IEEE* , Andrea L. Bertozzi, *Member, IEEE*, and Christian Jutten, *Fellow, IEEE*

***Abstract*—Hyperspectral images provide much more information than conventional imaging techniques, allowing a precise identification of the materials in the observed scene, but because of the limited spatial resolution, the observations are usually mixtures of the contributions of several materials. The spectral unmixing problem aims at recovering the spectra of the pure materials of the scene (endmembers), along with their proportions (abundances) in each pixel. In order to deal with the intra-class variability of the materials and the induced spectral variability of the endmembers, several spectra per material, constituting endmember bundles, can be considered. However, the usual abundance estimation techniques do not take advantage of the particular structure of these bundles, organized into groups of spectra. In this paper, we propose to use group sparsity by introducing mixed norms in the abundance estimation optimization problem. In particular, we propose a new penalty which simultaneously enforces group and within group sparsity, to the cost of being nonconvex. All the proposed penalties are compatible with the abundance sum-to-one constraint, which is not the case with traditional sparse regression. We show on simulated and real datasets that well chosen penalties can significantly improve the unmixing performance compared to classical sparse regression techniques or to the naive bundle approach.**

***Index Terms*—Hyperspectral imaging, remote sensing, spectral unmixing, endmember variability, group sparsity, convex optimization**

I. Introduction

THIS document gathers complementary results for the experiments on the synthetic dataset used in the main manuscript, and on the Cuprite dataset for the paper "Hyperspectral Image Unmixing with Endmember Bundles and Group Sparsity Inducing Mixed Norms". For the synthetic data, the obtained abundance matrices are displayed as images to reveal the group sparsity structure of the abundances enforced by the different algorithms. For Cuprite, the results include some additional within group abundance maps and their interpretation. The abundance maps are to be compared to a reference land cover map, which can be found for example in Fig. 10 of [1]. We reproduce it here for the reader's convenience (Fig. 2). We also make a comparison between the extracted endmembers and the USGS reference endmembers available for this scene. However, note that these endmembers and land cover maps, though useful for qualitative comparisons, cannot be considered as absolute ground truths, because the abundance maps were algorithmically generated and do not correspond to in situ ground truth data. As for the endmembers, one reference USGS signature is available only for each endmember, and thus they cannot be fully representative of the spectral variability occurring in the scene.

L. Drumetz is with IMT Atlantique, Lab-STICC, UBL, Technopôle Brest-Iroise CS 83818, 29238 Brest Cedex 3, France (e-mail: lucas.drumetz@imt-atlantique.fr).

T. R. Meyer is with SmartKYC Ltd., Los Angeles, CA, USA (e-mail: travis.meyer@smartkyc.com).

J. Chanussot and C. Jutten are Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, 38000 Grenoble, France. * Institute of Engineering Univ. Grenoble Alpes (e-mail: {jocelyn.chanussot,christian.jutten}@gipsa-lab.grenoble-inp.fr).

A. L. Bertozzi is with the Department of Mathematics at the University of California, Los Angeles, 90095 Los Angeles, CA, USA (e-mail: {tmeyer;bertozzi}@math.ucla.edu).

This work was supported by the European Research Council FP7/2007–2013, under Grant no.320684 (CHESS project), as well as projects ANR-16-ASTR-0027-01 APHYPIS, by CNRS PICS 263484, and by NSF grant DMS-1118971, NSF grant DMS-1417674, ONR grant N000141210838, UC Lab Fees Research grant 12-LR- 236660. L. Drumetz was also supported by a Campus France outgoing postdoctoral mobility grant, PRESTIGE-2016-4 0006.

II. Complementary results on the synthetic data

We show in Fig. 1 a part of the abundance matrices obtained for the synthetic dataset as images (only for some pixels of the image), to show the structure that each penalty induces on the coefficients. We confirm that FCLSU and the group penalty obtain similar performance. Collaborative sparsity deteriorates the abundance estimation because it forces some endmembers to be zero on the whole support of the image. The elitist penalty tends to incorporate spurious endmembers because of the between group density of the coefficients. The fractional penalty obtains a more accurate support of the endmembers than the other two, but at the price of sometimes overestimating the abundances because of a too violent sparse behavior, which is necessary to get rid of the noise but too harsh for a high precision abundance estimation.

III. Complementary Results on the Cuprite dataset

The second dataset we consider is a $200 \times 200 \times 186$ subset of the Cuprite dataset. The image was acquired by NASA's AVIRIS sensor and covers the Cuprite mining district in western Nevada, USA. We extracted 14 bundles according to the intrinsic dimensionality estimated by the Hyperspectral Subspace Identification by Minimum Error (HySIME) [2] on our subset. We used VCA three times on a third of the dataset, giving a total of $Q = 42$ signatures, clustered into 14 groups as was done for the Houston data. We compare the same algorithms as before and show in Fig. 3 the estimated abundance maps. The results are shown for only for 6 of the 14 extracted endmembers. The materials have been identified by visual comparison between the estimated abundance maps and endmember signatures with those recovered in [3]. For

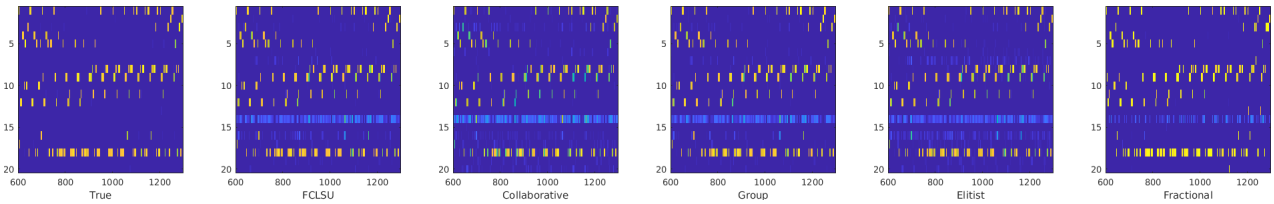


Fig. 1. Abundance matrices (pixels are on the x-axis, and candidate endmembers are on the y-axis. Candidates 1-5 belong to group 1, candidates 5-10 to group 2, and so on.

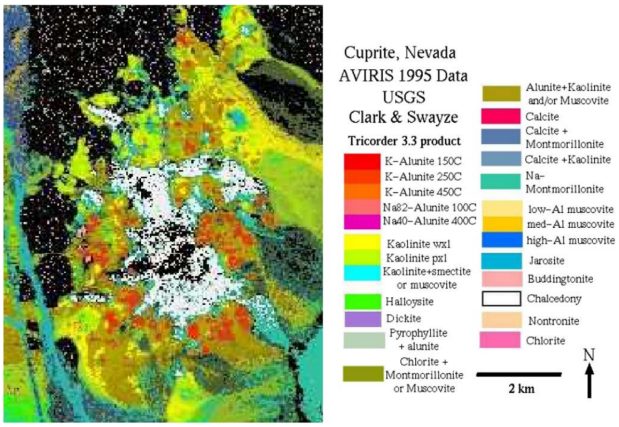


Fig. 2. Reference map showing the location of different minerals in the Cuprite mining district in Nevada (taken from [1]).

four of these materials, i.e. sphene, alunite, buddingtonite and chalcedony, we also show the abundances of the different instances of the bundles, in Fig. 5. The visual results (to be compared to a reference land cover map, which can be found e.g. in Fig. 10 of [1]) are in accordance with the conclusions drawn from the Houston data, i.e. that the collaborative and group approaches obtain abundances which are sparser and smoother at the global level than FCLSU, because they tend to distribute abundance equally between active signatures in each bundle. The elitist penalty leads to denser mixtures over the groups, and the fractional penalty obtains sparser abundances than other algorithms because it promotes sparsity more aggressively. Also, it obtains within group sparsity, which helps identifying the endmember candidates which are locally predominant, revealing spectral variations for each endmember.

Some materials, however, are not perfectly separated at the group level, but can be distinguished at the intra group level. For example, column 3 of Fig 5 (a) corresponds to a patch of kaolinite. Also, the bottom right part of the buddingtonite abundance maps should correspond to a Montmorillonite patch, according to the reference map. Since all the methods seem affected by these two issues, we can assume that the bundle extraction step was not carried out perfectly. This could be due to the absence of pure pixels for some of these materials, or an incorrect clustering step. In theory, if VCA extracted one instance for each material, the number of instances per material should be the same (3), which is not the case here. However, enabling group and within group sparsity allows the abundance estimation algorithm to separate the incorrectly clustered instances at the intra group level. When the clustering has been correctly carried out, local maps corresponding to different instances can be defined, as show in Fig. 5, e.g. for alunite and buddingtonite.

We show in Fig. 4 some of the signatures extracted by the fractional approach in randomly chosen pixels of the images for the same endmembers as in Fig. 3, along with the signatures present in the bundles (dashed). In black is shown a reference signature from the USGS spectral library. These references signatures were recalibrated using a scaling factor ψ to match the extracted signatures using a least squares regression:

$$\hat{\psi} = \arg\min_{\psi} \frac{1}{2} \sum_{i=1}^{n} ||\psi \mathbf{s}_0 - \mathbf{s}_i||_2^2 \tag{1}$$

where $\mathbf{s}_0$ is the reference USGS signature for a given material, and the $\mathbf{s}_i$ gather the extracted signatures for the n selected pixels. We see that overal the USGS signatures and the extracted ones by the fractional approach show a good agreement, except for Chalcedony and the lower part of the spectrum for alunite. Again, this could be caused by the fact that VCA was not able to extract pure pixels from the data, or because of nonlinear variabilities which could affect the image spectra. Nevertheless, even though the pixelwise endmembers' quality are inherently conditioned on the quality of the extracted bundles, the group information allows to interpolate for each pixel between the bundles signature to refine the variability estimation locally.

We also show the RMSE and SAM values associated to the reconstruction of the data, as well as the running times of all algorithms in Table I. The numbers follow the trends identified with the Houston data. FCLSU obtains the best quantitative reconstruction results since there is no sparsity involved. The fractional penalty obtains the second best results. The running times are almost equal for the group and elitist penalties, while the fractional algorithm is slightly more computationally intensive.

References

[1] J. Li, J. M. Bioucas-Dias, A. J. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing.," *IEEE Trans. Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6076–6090, 2016.

Metric \ Algorithm	FCLSU	Collaborative	Group	Elitist	Fractional
RMSE	0.0034	0.0048	0.0075	0.0053	0.0047
SAM (degrees)	0.5233	0.7391	1.1514	0.8288	0.7282
Running time (s)	54	102	175	174	194
λ	$\times$	1	0.05	0.1	0.125

TABLE I

VALUES OF THE RMSE AND SAM (COMPUTED ON THE RECONSTRUCTION) AND RUNNING TIMES FOR EACH TESTED ALGORITHM ON THE CUPRITE DATA. THE BEST VALUES ARE IN RED, AND THE SECOND BEST ARE IN BLUE. THE VALUES OF THE REGULARIZATION PARAMETERS ARE ALSO REPORTED, WHEN APPLICABLE.

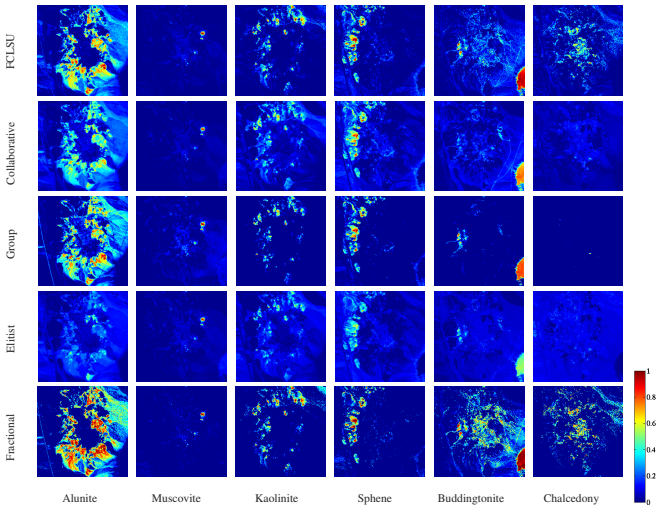


Fig. 3. Abundances for all the tested algorithms on the Cuprite dataset.

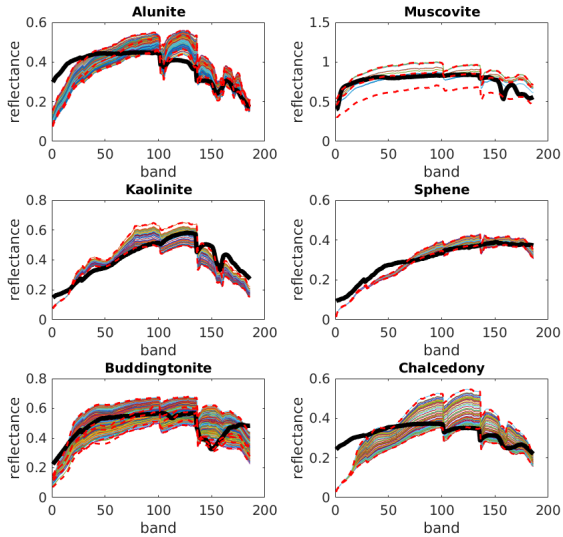
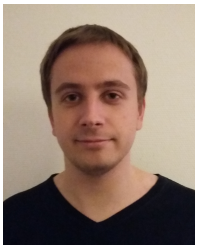


Fig. 4. Endmembers recovered from the bundles by the fractional algorithm, and the bundles' signatures (dashed), for 6 endmembers of the Cuprite data. In black is shown a reference signature from the USGS spectral library.

[2] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 2435–2445, Aug 2008.

[3] J. Nascimento and J. Bioucas Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 898–910, April 2005.



Lucas Drumetz (M'17) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 2013, and the Ph.D degree for the Université de Grenoble Alpes, Grenoble, France, in 2016. He received the 2017 best PhD award from the Université de Grenoble Alpes for this work. In 2017, he was a visiting assistant professor at the University of California, Los Angeles (UCLA), USA. The same year, he was a visiting researcher at the University of Tokyo, Tokyo, Japan, for three months. Since 2018, he has been an Associate Professor in the department of signal and communications at IMT-Atlantique, Brest, France. His research interests include signal and image processing, inverse problems, optimization techniques and machine learning for remote sensing data.

Travis R. Meyer received the B.Sc. degree in physics and the B.A. degree in applied mathematics from the University of California, Los Angeles (UCLA), CA, USA, in 2011. He received his Ph.D. degree in applied mathematics from UCLA in 2017 with focused work on machine learning techniques for text and image analysis. While completing his Ph.D., he helped develop and teach courses on machine learning for his department and mentored many advanced undergraduate students during their first research projects. Additionally, his publications have contributed to areas including atomic force microscopy and hyperspectral imaging. Currently he is the head machine learning developer for SmartKYC where he pursues his interest in variational machine learning techniques by developing cutting-edge knowledge aggregation and analysis systems to help banks understand potential client risks.

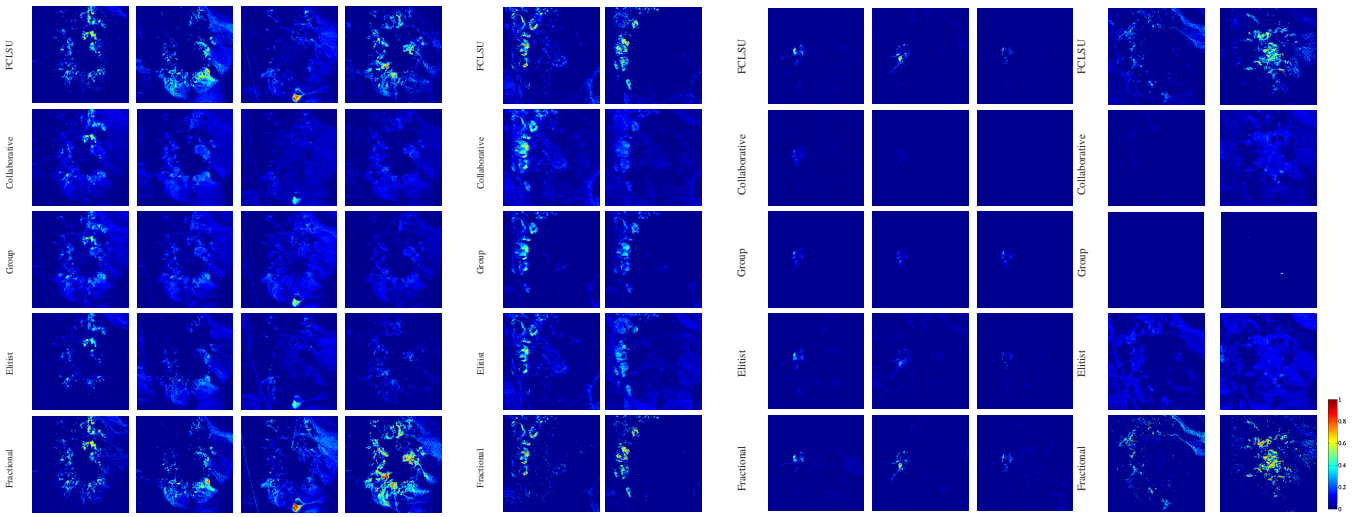


Fig. 5. Abundances for some of the (a) alunite, (b) sphene (c) buddingtonite, and (d) chalcedony endmember candidates, for all the tested algorithms.

Jocelyn Chanussot (M'04–SM'04–F'12) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from the Université de Savoie, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA - French National Defense Department). Since 1999, he has been with Grenoble INP, where he is currently a Professor of signal and image processing. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing. He has been a visiting scholar at Stanford University (USA), KTH (Sweden) and NUS (Singapore). Since 2013, he is an Adjunct Professor of the University of Iceland. In 2015-2017, he was a visiting professor at the University of California, Los Angeles (UCLA). Dr. Chanussot is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007-2010) which received the 2010 IEEE GRS-S Chapter Excellence Award. He was the co-recipient of the NORSIG 2006 Best Student Paper Award, the IEEE GRSS 2011 and 2015 Symposium Best Paper Award, the IEEE GRSS 2012 Transactions Prize Paper Award and the IEEE GRSS 2013 Highest Impact Paper Award. He was a member of the IEEE Geoscience and Remote Sensing Society AdCom (2009-2010), in charge of membership development. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009-2011) and Cochair of the GRS Data Fusion Technical Committee (2005-2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006-2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing, (2009). He was an Associate Editor for the IEEE Geoscience and Remote Sensing Letters (2005-2007) and for Pattern Recognition (2006-2008). Since 2007, he is an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing. He was the Editor-in-Chief of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2011-2015). In 2013, he was a Guest Editor for the Proceedings of the IEEE and in 2014 a Guest Editor for the IEEE Signal Processing Magazine. He is a Fellow of the IEEE and a member of the Institut Universitaire de France (2012-2017).

Andrea L. Bertozzi is an applied mathematician with expertise in nonlinear partial differential equations and fluid dynamics. She also works in the areas of geometric methods for image processing, crime modeling and analysis, and swarming/cooperative dynamics. Bertozzi completed all her degrees in Mathematics at Princeton. She was an L. E. Dickson Instructor and NSF Postdoctoral Fellow at the University of Chicago from 1991-1995. She was the Maria Geoppert-Mayer Distinguished Scholar at Argonne National Laboratory from 1995-6. She was on the faculty at Duke University from 1995-2004 first as Associate Professor of Mathematics and then as Professor of Mathematics and Physics. She has served as the Director of the Center for Nonlinear and Complex Systems while at Duke. Bertozzi moved to UCLA in 2003 as a Professor of Mathematics. Since 2005 she has served as Director of Applied Mathematics, overseeing the graduate and undergraduate research training programs at UCLA. In 2012 she was appointed the Betsy Wood Knapp Chair for Innovation and Creativity. Bertozzi's honors include the Sloan Research Fellowship in 1995, the Presidential Early Career Award for Scientists and Engineers in 1996, and SIAM's Kovalevsky Prize in 2009. She was elected to the American Academy of Arts and Sciences in 2010 and to the Fellows of the Society of Industrial and Applied Mathematics (SIAM) in 2010. She became a Fellow of the American Mathematical Society in 2013 and a Fellow of the American Physical Society in 2016. She won a SIAM outstanding paper prize in 2014 with Arjuna Flenner, for her work on geometric graph-based algorithms for machine learning. Bertozzi is a Thomson-Reuters/Clarivate Analytics 'highly cited' Researcher in mathematics for both 2015 and 2016, one of about 100 worldwide in her field. She was awarded a Simons Math + X Investigator Award in 2017, joint with UCLA's California NanoSystems Institute (CNSI). Bertozzi was appointed Professor of Mechanical and Aerospace Engineering at UCLA in 2018, in addition to her primary position in the Mathematics Department. In May 2018 Bertozzi was elected to the US National Academy of Sciences. Bertozzi has served on the editorial boards of fourteen journals: SIAM Review, SIAM J. Math. Anal., SIAM's Multiscale Modeling and Simulation, Interfaces and Free Boundaries, Applied Mathematics Research Express (Oxford Press), Applied Mathematics Letters, Mathematical Models and Methods in the Applied Sciences (M3AS), Communications in Mathematical Sciences, Nonlinearity, and Advances in Differential Equations, Journal of Nonlinear Science, Journal of Statistical Physics, Nonlinear Analysis Real World Applications; and the J. of the American Mathematical Society. She served as Chair of the Science Board of the NSF Institute for Computational and Experimental Research in Mathematics at Brown University from 2010-2014 and previously on the board of the Banff International Research Station. She served on the Science Advisory Committee of the Mathematical Sciences Research Institute at Berkeley from 2012-2016. To date she has graduated 35 PhD students and has mentored over 40 postdoctoral scholars.

Christian Jutten (AM'92-M'03-SM'06-F'08) received Ph.D. and Doctor es Sciences degrees in signal processing from Grenoble Institute of Technology (GIT), France, in 1981 and 1987, respectively. From 1982, he was an Associate Professor at GIT, before being Full Professor at Univ. Grenoble Alpes, in 1989. Since 80's, his research interests have been machine learning and source separation, including theory (separability, nonlinear mixtures, sparsity, multimodality) and applications (brain and hyperspectral imaging, chemical sensor array, speech). He coauthored 105+ papers in international journals, 4 books, 27 keynote plenary talks and 230+ communications in international conferences. He was a visiting professor at EPFL (Lausanne, Switzerland, 1989), Riken labs (Japan, 1996) and Campinas Univ. (Brazil, 2010). He was director or deputy director of his lab from 1993 to 2010, especially head of the signal processing department (120 people) and deputy director of GIPSA-lab (300 people) from 2007 to 2010. He was a scientific advisor for signal and images processing at the French Ministry of Research (1996–1998) and CNRS (2003–2006 and since 2012). He was organizer or program chair of many international conferences, especially of the 1st Int. Conf. on Blind Signal Separation and Independent Component Analysis in 1999 (ICA'99). He has been a member of a few IEEE Technical Committees. He received many awards, e.g. best paper awards of EURASIP (1992) and IEEE GRSS (2012), Medal Blondel (1997) from the French Electrical Engineering society, and one Grand Prix of the French Académie des Sciences (2016). He was elevated as IEEE fellow (2008), EURASIP fellow (2013) and as a Senior Member of Institut Universitaire de France since 2008. He is the recipient of a 2012 ERC Advanced Grant for the project Challenges in Extraction and Separation of Sources (CHESS).