Hyperspectral Local Intrinsic Dimensionality

Lucas Drumetz, Miguel Angel Veganzones, Rubén Marrero, Guillaume Tochon, Mauro Dalla Mura, Giorgio Licciardi, Christian Jutten, Jocelyn Chanussot

Hyperspectral Local Intrinsic Dimensionality

Lucas Drumetz, Miguel A. Veganzones, *Member, IEEE,* Rubén Marrero, Guillaume Tochon, *Student Member, IEEE,* Mauro Dalla Mura, *Member, IEEE,* Giorgio Licciardi, *Member, IEEE,* Christian Jutten, *Fellow, IEEE,* and Jocelyn Chanussot, *Fellow, IEEE*

Abstract—The Intrinsic Dimensionality (ID) of multivariate data is a very important concept in spectral unmixing of hyperspectral images. A good estimation of the ID is crucial for a correct retrieval of the number of endmembers (the spectral signatures of macroscopic materials) in the image, for dimensionality reduction or for subspace learning, among others. Recently, some approaches to perform spectral unmixing and super-resolution locally have been proposed, which require a local estimation of the number of endmembers to use. However, the role of ID in local regions of hyperspectral images has not been properly addressed. Some important issues when dealing with small regions of hyperspectral data can seriously affect the performance of conventional hyperspectral ID estimators. We show that three factors mainly affect local ID estimation: the number of pixels in the local regions, which has to be high enough for the estimations to be relevant, the number of hyperspectral bands which complicates the estimations if the ambient space has a high dimensionality, and the noise, which can be misinterpreted as signal when its power is important. Here, we review the hyperspectral ID estimators on the literature for local ID estimation, we show how they behave in a local setting on synthetic and real datasets, and we provide some guidelines to make proper use of these estimators in local approaches.

Index Terms—Intrinsic dimensionality, virtual dimensionality, local spectral unmixing, hyperspectral imagery.

I. Introduction

HYPERSPECTRAL unmixing is one of the most important and widely used techniques in hyperspectral image analysis [9], [10]. It decomposes a hyperspectral image into a set of spectral signatures corresponding to macroscopically pure materials, named endmembers, and a set of cover proportions comprised in a fractional abundance matrix. The information provided by spectral unmixing enables a myriad of applications requiring fine identification of materials or estimation of physical parameters [9]. Many approaches exist in the literature to perform spectral unmixing [10], [19], [32],

L. Drumetz and C. Jutten are with Joseph Fourier University, Image and Signal Department (DIS), GIPSA-lab, F-38402 Saint Martin d'Heres Cedex (e-mail: {lucas.drumetz;christian.jutten}@gipsa-lab.grenoble-inp.fr).

M.A. Veganzones is with the CNRS, Department Image and Signal (DIS), GIPSA-lab, F-38402 Saint Martin d'Heres Cedex (e-mail: {miguel-angel.veganzones@gipsa-lab.grenoble-inp.fr})

R. Marrero is with the Institut de Planetologie et d'Astrophysique de Grenoble (IPAG), Université Joseph Fourier (UJF), Grenoble, France. (e-mail: ruben.marrero-gomez@obs.ujf-grenoble.fr)

G. Tochon, M. Dalla Mura, Giorgio Licciardi and J. Chanussot are with the Grenoble Institute of Technology, Image and Signal Department (DIS), GIPSA-lab, F-38402 Saint Martin d'Heres Cedex (e-mail: {guillaume.tochon;mauro.dalla-mura;giorgio-antonino.licciardi; jocelyn.chanussot}@gipsa-lab.fr)

This work has been partially supported by the European Research Council under the European Community's Seventh Framework Programme FP7/2007–2013, under Grant Agreements no.320684 (CHESS project) and no.320594 (DECODA project).

and almost all of them require an estimation of the intrinsic dimensionality.

Let us denote a hyperspectral dataset by $\mathbf{X} \in \mathbb{R}^{d \times n}$ where each column, $\mathbf{x}_i \in \mathbb{R}^d$, $i \in \{1, \ldots, n\}$, refers to a hyperspectral sample vector, *i.e.* a pixel with d spectral channels, and n denotes the number of samples. Usually, the dimensionality of hyperspectral vectors, d, is large, with hundreds or thousands of spectral bands. Assuming the observations may be decomposed into signal, $\mathbf{s}$, and noise, $\mathbf{n}$, that is, $\mathbf{x} = \mathbf{s} + \mathbf{n}$, Cawse-Nicholson *et al.* [14] introduce the following definition:

Definition 1. *The Intrinsic Dimension (ID) of a dataset, $\mathbf{x}_1, \ldots, \mathbf{x}_n$, is the dimension, K, of the vector subspace spanned by the signals, $\mathbf{s}_1, \ldots, \mathbf{s}_n$.*

Different authors have given alternative definitions of the intrinsic dimension or of similar terms. Chang and Du [16] define the "*virtual dimensionality*" as the the number of endmembers necessary to give accurate unmixing. Bajorski [5] defines the "*effective dimensionality*" as the dimensionality of the affine subspace giving an acceptable approximation to all pixels. Definition 1 is equivalent to the ones provided in [7], [40]. Besides conceptual aspects, all of them are used in spectral unmixing to estimate the actual number of endmembers or the dimensionality of the subspace spanned by these endmembers. Hereafter, for sake of clarity, we will make use of the ID term only, and follow Definition 1.

Recently, some local approaches have been proposed for spectral unmixing [12], [20], [23], [31], [42], in order to overcome some of the issues of global approaches, *i.e.* spectral variability [41], [45]. Furthermore, the local spectral unmixing approach proposed in [20] has proven to be useful to propose new unmixing-based segmentation techniques [44] or to improve unmixing-based hyperspectral super-resolution techniques using the local low rank property of hyperspectral data [30], [43]. In addition to the latter works, we envision incorporating local spectral unmixing to other hyperspectral applications such as unmixing-based anomaly/target detection, spectral-spatial classification or visualization, among others. Thus, there is an increasing need to better understand the role of ID in local neighborhoods of hyperspectral data, *i.e.* in patches or segmentation regions.

A. Related work

There exist many ID methods for hyperspectral data in the literature, as well as more general techniques in the signal processing community [11]. Nevertheless, the specificity of hyperspectral data reside in two aspects: the 2D spatial arrangement of the signal and the high dimensionality induced

by the numerous spectral bands. Most of the existing ID estimation algorithms are based on the eigen-decomposition of some data dependent statistical matrix, often second order statistics. The basic idea is that if some noiseless signals $\mathbf{s}$ span a k-dimensional vector space, then their covariance matrix $\mathbf{K_s}$ should have a rank which is equal to k. Then this covariance matrix should only have k nonzero eigenvalues. The main issue with this strategy is that noisy signals have more nonzero eigenvalues than their ID value, and the problem boils down to being able to sort the eigenvalues related to signal and the ones related to noise in the following eigenvalue decomposition:

$$\mathbf{K_x} = \mathbf{B}^\top \mathbf{D} \mathbf{B} \tag{1}$$

where $\mathbf{B}$ is a change of basis matrix, and $\mathbf{D}$ is a diagonal matrix containing the eigenvalues of $\mathbf{K_x}$ on its diagonal.

A simple baseline approach is to define the ID as the number of the largest eigenvalues that must be retained to represent a percentage of the total variance of the data [21], *i.e.*, 95% or 99%. Chang and Du [16] proposed the widely used *Harsanyi-Farrand-Chang (HFC)* method, based on the comparison of the eigenvalues obtained from the covariance and the correlation matrices. The validity of the HFC method has been questioned in [5], [6], and Bajorski proposed an alternative algorithm, called *Second Moment Linear dimensionality (SML),* based on similar concepts. Another popular algorithm to perform hyperspectral ID estimation is the *Hyperspectral Subspace Identification by Minimum Error (HySIME)* [8], which is an evolution of the *Signal Subspace Estimation (SSE)* algorithm presented in [7]. The HySIME algorithm works by identifying the signal subspace achieving a residual error comparable to the estimated noise power. A different approach has been proposed in [14], where new results in *Random Matrix Theory (RMT)* used to determine which eigenvalues are due to noise and which are due to signal have been adapted for the identification of the hyperspectral ID. The *Outlier Detection Method (ODM)* [4] is another eigen-based algorithm, although ODM focuses on modelling the noise and treats the signal as outliers to the noise distribution.

Three non eigen-based hyperspectral ID estimators have recently been proposed. The first one, introduced in [34] as part of a *Negative ABundance-Oriented (NABO)* unmixing algorithm, borrows the main idea from the HySIME algorithm. Basically, it decomposes the residual error from the unconstrained unmixing into two components, a first due to noise and a second due to ID. The algorithm works by starting from an underestimate of the ID, and then, iteratively increments the ID value until the unmixing error can be solely explained by the noise term. The second non eigen-based method, called *Hyperspectral Image Dimension Estimation through Nearest Neighbor distance ratios (HIDENN)* [26] is based on local geometrical properties of the data manifold. The technique is aimed at computing the correlation dimension of the dataset, which is itself closely related to the concept of fractal dimension. The basic idea is to count (in the neigborhood of one data point) the total number of pairs of points $g(\epsilon)$ which have a distance between them that is less than ϵ. Then it can be shown that if $n \to \infty$ and $\epsilon \to 0$, the so-called correlation

integral $C(\epsilon)$ has the following asymptotic behavior:

$$C(\epsilon) = \frac{g(\epsilon)}{n^2} \sim \epsilon^{k-1}, \tag{2}$$

where $k-1$ is here the dimension of the manifold (and k is the ID of the data). This behavior can be intuitively understood by the fact that in higher dimensions, there are more possible ways for one point to reach neighboring points. One can then recover the ID by computing:

$$k - 1 = \lim_{\epsilon \to 0} \frac{\ln(C(\epsilon))}{\ln(\epsilon)} \tag{3}$$

Note that since the ID is here estimated in each point of the data cloud, in the signal processing literature this category of ID estimation technique can be referred to as local ID estimation [13], [11]. However, the concept differs from the one we are interested in since we consider spatially local ID estimation.

In [28], Kyubeda *et al.* proposed the *Maximum Orthogonal Complement Algorithm (MOCA)*, which solves an optimiziation problem exploiting the sensitivity of the l_2^∞ norm to rare materials, so the signal subspace preserves them. In [1], Acito *et al.* proposed a version of MOCA, called *Robust Signal Subspace Estimator (RSSE)*, that improves the latter in terms of computational speed and lighter parametrization. The same authors summarized in [2] both approaches, MOCA and RSSE, using a common theoretical framework, and also proposed a more computationally efficient version of the MOCA algorithm named *Modified MOCA (MMOCA)*. They also derived from the RSSE algorithm a method to account for signal dependent noise [3]. Chang *et al.* [17] proposed a Neyman-Pearson detector version of MOCA linking the ideas behind MOCA with those of the HFC algorithm. Recently, Chang *et al.* [18] have proposed an extension of the latter work based on high-order statistics.

B. Contributions

In [15], Cawse-Nicholson *et al.* studied the effect of correlated noise on ID estimation, and Hasanlou and Samadzadegan performed in [25] a comparative study of some ID estimation algorithms for classification. A recent survey of ID estimation algorithms compares five methods, three of which are also considered in this study, mostly in terms of ID estimation performance on the whole image, and in terms of the impact of the noise correlation and estimation [37]. Here, we are interested in the performance of hyperspectral ID estimation algorithms when going from global to local studies, that is, the capacity of the algorithms to correctly estimate the ID on small regions or subsamples of a hyperspectral image. In addition, the present study includes several algorithms not considered in [37].

Hyperspectral ID estimation algorithms can be grouped according to two main characteristics: i) whether they are based on eigen-decomposition or not, and ii) the requirement of a de-noising step or of a noise power estimation. When trying to identify the ID of local (often small) regions in hyperspectral images, eigenvalue-based methods can be severely affected by the so-called curse of dimensionality [29] and the high between-band correlation. The curse of dimensionality refers

to: i) the empty space phenomenon in high dimensions, which makes it necessary to use more and more data samples for estimation purposes when the dimension becomes higher, and to ii) the fact that high-dimensional data often show multicollinearity, which can hamper noise estimation regression (see the Appendix). The effects of the local de-noising and the local estimation of the noise power can also influence ID estimation. Usually, small regions present a relatively high spectral homogeneity, in the sense that the materials in the different pixels of small regions are likely to be the same, with slowly varying abundance coefficients. Then, noise can be sometimes misinterpreted as a signal, compromising the local de-noising and noise power estimation.

We describe and compare nine ID estimation algorithms when going from global to local studies of hyperspectral data. We catalogue the ID algorithms according to their base methodologies and we highlight their main drawbacks when working on local, often small, subsets of data. We also provide some guidelines for a better use of these algorithms in local studies which can be summarized as: (i) perform a global de-noising or estimation of the noise power, that is, avoid the use of local de-noising or local noise power estimation; (ii) subsets below a size threshold produce unreliable estimations, usually presenting an overestimation peak and an increase in the error variance.

The remainder of the paper is as follows: Sec. II is an overview of the compared ID estimation algorithms, Sec. III and Sec. IV present the experimental local ID studies using synthetic and real data, respectively; in Sec. V, we summarize the experimental observations and provide some guidelines to use the ID estimation algorithms in local experiments; and, finally, we give some conclusions in Sec. VI.

II. Intrinsic dimensionality estimation methods

In this section, some methods for the estimation of the ID of a hyperspectral image are listed and presented. These methods are the ones used for the experiments in Sections III and IV. Several algorithms in the following require a noise estimation step before computing the ID. The algorithm used in this paper to perform this noise estimation (originally suggested in [38]) is presented in the Appendix. In [22], several algorithms for noise estimation for hyperspectral images, based on linear regression are compared. The noise estimation method suggested in [38] was shown to be relatively robust in the simulations of that study. It is also the most widely used in the community. Anyway, by running similar experiments as the ones described below with known noise values (or equivalently a perfect noise estimation), we obtained comparable results to those obtained by estimating the noise globally on the whole image. This shows that the noise estimation provided by this method seems suited for local ID estimation. More details can be found in a supplementary material document provided by the authors. Next, we describe all the algorithms compared in this study. Some of the properties of those are listed in Table I.

A. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is an extremely popular technique for data analysis [27], which has been used extensively for dimension reduction, among other applications. The idea is to perform a Singular Value Decomposition on the sample covariance matrix of a given dataset. The resulting eigenvectors are then sorted by decreasing order of eigenvalues. The subspace spanned by the first k eigenvectors is the k-dimensional space whose explained variance percentage is the highest. This means that when the data cloud is projected onto this k-dimensional subspace, the relative difference between the variance of the data cloud and its projection is the lowest possible. To estimate the dimensionality of a dataset, one has to select a threshold on the percentage of the explained variance. However, for some applications, including hyperspectral imaging, the manual choice of a threshold is not an easy task, since explained variance is not directly linked to the number of sources, and also because variance can be very well explained in a very low-dimensional subspace while the intrinsic dimension of the data manifold might be higher. In the experiments, we selected a threshold of 95% of the explained variance to determine the ID values.

B. Harsanyi, Farrand, and Chang (HFC)

This dimensionality estimation method, termed HFC (for Harsanyi, Farrand, and Chang) is another rather simple and widely adopted technique to compute the ID of a hyperspectral dataset. The sample correlation and covariance matrices ($\mathbf{R_x}$ and $\mathbf{K_x}$, respectively) of the observations are both computed, and their eigenvalues are sorted in decreasing order. HFC assumes that the sources are deterministic and nonnegative, and that the noise is spectrally white (i.e. uncorrelated with constant variance) with zero mean. In this case, if the ID is k, then the k largest eigenvalues of $\mathbf{R_x}$ are supposed to be larger than those of $\mathbf{K_x}$ because in the corresponding components (coming from the transformation by the eigenvalue decomposition) an endmember contributes to the correlation eigenvalues in addition to the noise. Based on this, the algorithm performs a hypothesis test on each eigenvalue set to determine if the eigenvalues of the covariance and correlation matrices are statistically significantly different or not. Note that the algorithm's results depend on a user-tuned false alarm probability, set to $\alpha = 10^{-5}$ in the experiments. Every time the test fails in a component, the ID value is incremented. The ID finally corresponds to the number of times this test has failed. An alternative version of the algorithm, called Noise Whitened HFC (NWHFC), assumes that the noise is uncorrelated but with possibly non-constant variance. It includes a noise-whitening step before using the same methodology as HFC.

Bajorski has argued in [5], [6] that the HFC method can only measure the dependence of the difference between consecutive eigenvalues of the covariance to the average values of the bands, which is unrelated to the ID value. Therefore, the HFC method may be conceptually wrong. However, the method provides consistent results because the differences between consecutive covariance eigenvalues is in itself a useful indicator of the ID of the dataset, while relating this difference to eigenvalues of the correlation matrix is not relevant [6].

C. Hyperspectral Subspace Identification by Minimum Error (HySIME)

Bonjour En V

Another popular algorithm to perform hyperspectral intrinsic dimensionality estimation is Hyperspectral Subspace Identification by Minimum Error (HySIME) [8]. For this algorithm, the noise $\mathbf{n}$ is assumed to be zero-mean Gaussian distributed; the noise value $\mathbf{n}$ and the noise correlation matrix $\mathbf{R_n}$ are estimated using the band correlation method described in the Appendix. The sample observation correlation matrix $\mathbf{R_x}$ is computed, as well as the signal sample correlation matrix $\mathbf{R_{\hat{s}}}$, taking the signal values $\hat{s}$ by subtracting the estimated noise values $\hat{n}$ from the observations $\mathbf{x}$. The eigenvectors of the latter matrix are computed and sorted in descending order according to the corresponding eigenvalues. The subspace spanned by the first k eigenvectors corresponding to the k largest eigenvalues is the signal subspace, whereas the orthogonal complement is associated with the noise subspace. The separation between the two is found by looking for the value of k which minimizes the Mean Squared Error (MSE) between the signal and the projection of the observations on the subspace spanned by the first k eigenvectors, taking into account the projection error power (decreasing function of k) as well as the noise power (increasing with k). Note that in this case, using the correlation matrix is meaningful because its k first eigenvectors define the subspace minimizing the MSE between the projected data and the original data.

D. Random Matrix Theory (RMT)

This technique was recently introduced in [14] and makes use of the tools of Random Matrix Theory (RMT) to estimate the ID of a hyperspectral dataset. It requires a noise estimation step which, in [14], is performed by the method presented in [35]. The method extends an existing RMT-based method for dimensionality estimation [33] to the case of spectrally correlated Gaussian noise. The underlying mixing model is also assumed to be linear. The general idea is that, under the assumption that each column of the $d \times n$ noise image is distributed according to $\tilde{\mathbf{n}} \sim \mathcal{N}(0, \mathbf{\Phi})$, the random cross product matrix $\tilde{\mathbf{n}}\tilde{\mathbf{n}}^\top$ follows a Wishart distribution (which can be seen as a multivariate generalization of the χ^2 distribution) $\mathcal{W}_d(\mathbf{\Phi}, N)$, with d representing the degrees of freedom, and $\mathbf{\Phi}$ the $d \times d$ scale matrix. The probability density function of the largest eigenvalue of such matrices has been extensively studied in RMT. In the context of dimensionality estimation, a criterion has been found to test which is the largest sample covariance eigenvalue which is statistically consistent with the distribution of the largest eigenvalue of a Wishart matrix. In other words, this means that the eigenvalue of $\mathbf{K_y}$ found by this process is the largest noise eigenvalue, and that all the larger sample covariance eigenvalues are associated to a signal component. This criterion, originally derived for a number of samples $n \to \infty$ and a number of variables (bands in this application) $d \to \infty$, with their ratio constant: $\frac{d}{n} = c$ (usual conditions in RMT), has also shown to be reliable for large but finite n and d (see [14] and references therein). The computation of the eigenvalues of interest to be tested against those of a Wishart matrix, as well as the testing criterion, differ in the general case if the uncorrelated noise assumption has been dropped, but the basic principle remains the same.

E. Outlier Detection Method (ODM)

The algorithm introduced in [4] estimates the ID of a hyperspectral image by focusing on the noise and treating the signal data points as outliers to the noise distribution. It comprises three steps: the first is a whitening step performed by a Minimum Noise Fraction (MNF) transform [24], in which the noise estimation is performed using once again the band-regression method. The noise is then whitened by an eigenvalue decomposition of the noise covariance matrix $\mathbf{K_n}$ and scaled so as to get equal variances in each band, thus defining a noise hypersphere in the spectral space, and a principal component analysis is performed on the transformed data to obtain the final transformed components. The final step is the ID estimation through outlier detection, using Inter-Quartile Range (IQR) to define a boundary between the noise and the "outliers". The Euclidean distances between the standard deviation of each transformed band and the standard deviation of the previous one are computed, and the ID is incremented every time the value is above the IQR threshold. It is a nonparametric technique, which does not make any assumption on the noise distribution (even though the band-regression based noise estimation algorithm used will provide optimal performance when the noise is Gaussian, because of the least squares step), and hence the final step is supposed to be robust to a small number of samples used for the estimation.

F. Vertex Component Analysis/Negative ABundance Oriented algorithm (VCA/NABO)

This technique [34] performs spectral unmixing and dimensionality estimation at the same time. It is noteworthy that this method is not eigenvalue-based. The idea is to start from an underestimation of the dimensionality of the dataset, and an estimation of the noise. Then an endmember extraction (using any Endmember Extraction Algorithm (EEA)) is performed, and the abundances are computed through linear unconstrained least squares unmixing, dropping both the usual Abundance Sum-to-one Constraint (ASC) and the Abundance Nonnegativity Constraint (ANC). Then, the power of the Root Mean Square Error (RMSE) is compared to the estimated noise power. If the former is higher than the latter, the dimensionality is incremented until the error power becomes smaller than the estimated noise power. At this step, it should not be necessary to increase the dimensionality further since the potential gain in RMSE will not be meaningful, and so the number of endmembers has been found. It should be noted that the abundances are computed without using any constraints so that RMSE (in other words, the projection error) is not due to the projection of the data onto the feasible set of solutions but mainly to the fact that the subspace on which the data are projected has too small a dimension. In the experiments described below, the chosen EEA is Vertex Component Analysis (VCA) [36]. As this widely used EEA is stochastic by nature, the VCA/NABO algorithm is performed 20 times, and the final ID value is the mean of the results of each iteration.

G. Hyperspectral Intrinsic Dimensionality Estimator through Nearest Neighbor distance ratios (HIDENN)

The dimensionality estimation method described here was presented in [26] and is called Hyperspectral Image Dimension Estimation through Nearest Neighbor distance ratios (HI-DENN). As VCA/NABO (though the methods are completely different in nature), it differs from most of the other methods mentioned in this paper in the sense that it is not based on any eigenvalue decomposition whatsoever. The data is assumed to come from samples of a manifold (it does not require any particular mixing model, so long as the abundances are subject to the ANC and the ASC), whose dimension is equal to the number of endmembers in the image minus one. A particular case of this is the $(k-1)$-simplex defined by a linear mixture of k materials. The algorithm estimates the dimension of the manifold (locally isomorphic to $\mathbb{R}^{k-1}$) using geometrical properties and then provides the number of endmembers. In that case, the distance between each data sample and its l-nearest neighbor is computed for two well chosen values of l, and using a variant of Eq. (3), an estimator of the correlation dimension is built to estimate the ID at this location in the spectral space. The choice of these values is critical since they need to be small enough to reduce the influence of the noise, but also large enough to be statistically robust. The individual pixel values are then averaged to give the global ID of the dataset, requiring a sufficient number of samples for the estimation to be meaningful. As the estimation of the dimension of such a manifold in the spectral space is highly sensitive to noise, a denoising may be performed beforehand in order to allow a more robust estimation of the ID. The algorithm becomes D-HIDENN (for Denoised-HIDENN) and makes use once again of the band correlation noise estimation technique described in [38].

H. Modified Maximum Orthogonal Complement Algorithm (MMOCA)

This non eigenvalue-based ID estimation technique [2], MMOCA (for Modified Maximum Orthogonal Complement Analysis), is actually a combination of the NWHFC algorithm described above and the MOCA algorithm [28]. The former is used to provide an underestimation of the ID of the dataset, so that the latter can iterate on the ID values from this starting point. More precisely, for a given candidate ID value k, MOCA aims at finding a suboptimal solution to the following optimization problem:

$$\hat{\mathbf{M}} = \arg \min_{\tilde{\mathbf{M}}} l_2^\infty \{\mathbf{P}_{\tilde{\mathbf{M}}}^\perp \mathbf{Y}\} \qquad (4)$$

where $\mathbf{Y}$ is the whitened data matrix, $l_2^\infty(\mathbf{X}) = \sup_i \|\mathbf{x}_i\|_2$ ($\mathbf{x}_i$ is the i$^{\text{th}}$ column of $\mathbf{X}$), and $\tilde{\mathbf{M}}$ is taken from the set of all possible bases of a k-dimensional subspace of $\mathbb{R}^d$. $\mathbf{P}_{\tilde{\mathbf{M}}}^\perp$ is the projection matrix on the orthogonal complement of the subspace spanned by $\mathbf{M}$, such that $\mathbf{P}_{\tilde{\mathbf{M}}}^\perp \mathbf{Y}$ is the error of the projection of the whitened data on the signal subspace. The l_2^∞ norm is used for its sensitivity to rare materials, since a rare material not accounted for by the $\tilde{\mathbf{M}}$ matrix will result in a high error on the concerned pixels, even if they are very few. The stopping criterion for this iterative process is
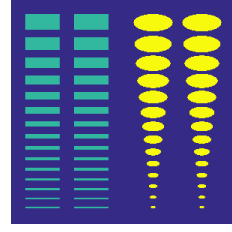

Fig. 1. The spatial pattern used for the creation of the synthetic datasets.

based on a hypothesis test using a Maximum A Posteriori (MAP) criterion. The idea is to determine whether $l_2^\infty \{\mathbf{P}_{\tilde{\mathbf{M}}}^\perp \mathbf{Y}\}$ depends only on the noise distribution or also on the residual signal.

III. Experiments on synthetic datasets

A. Datasets

The synthetic datasets built for this study were designed to evaluate how the previous algorithms behave from a local to a global scale, and to assess the effects of the SNR as well as the number of bands of the hyperspectral data in the ID estimation. A spatial pattern of 300×300 pixels comprising two kinds of aligned geometrical shapes (rectangles and ellipses) of various sizes was synthesized (see Fig. 1). Different variants of the dataset were created with different numbers of bands (480, 240, 120, 60 and 30 bands) and a spectrally and spatially white Gaussian noise was added so as to reach different values of SNR (20, 25, 30, 35, 40 dB), yielding a total of 25 synthetic images.

Here, we have considered a spectrally and spatially white noise. However as shown in [37], coloration of the noise (different variances in each bands, but still a diagonal noise covariance matrix) and correlation between bands for the noise can be significant in real scenarios. We have performed experiments on synthetic datasets accounting for these two properties of the noise, in order to see the impact of non white noise on local ID estimation. However, the conclusions are very similar to those of the experiments with white noise. Hence, these results are not shown here but gathered in a supplementary document file provided by the authors. Furthermore, since we want to focus on the capability of the different algorithms for local ID estimation, we only consider Roger's method [38] for noise estimation. We tested the impact of this choice by comparing the results of the local ID estimation using this noise estimation strategy on the whole image, to the use of the actual noise values. The results are similar in both cases, which shows that Roger's noise estimation strategy has little impact on the results, at least when the noise is estimated globally.

From the spatial pattern of Fig 1, three distinct mixtures were created: two mixtures of five endmembers and one of three endmembers. A mixture of three endmembers was employed to define the background, while two other mixtures of five endmembers were situated in the rectangles and the ellipses, respectively. The endmembers were randomly chosen from a mineral sublibrary of the U.S. Geological Survey

Property \ Algorithm	HySIME	RMT	ODM	VCA/NABO	HIDENN	D-HIDENN	PCA	HFC	MMOCA
Eigenvalue based	✓	✓	✓				✓	✓	
Nearest neighbor distance ratios					✓	✓			
Subspace estimation	✓			✓					✓
Noise estimation step	✓	✓	✓	✓		✓			
Underlying Mixing Model	Free	LMM	Free	LMM	Free	Free	Free	Free	Free

TABLE I
PROPERTIES OF THE ALGORITHMS USED.

(USGS) spectral library[1], with the constraint that the Spectral Angle Mapper between two signatures should not be less than 10 degrees or more than 30 degrees. This library contains the spectral signatures of various minerals acquired on the ground with a field spectrometer. The original endmembers were downsampled by a factor of 2, 4, 8 and 16 to provide datasets with the selected range of spectral bands. Note that some of the endmembers can be common to the different mixtures. In the end, there are 9 distinct endmembers in the image: 2 endmembers are common between the background and the ellipses, another is common between the background and the rectangles, and a last one is common between the ellipses and the rectangles. Thus we can deduce that there are 4 endmembers which are repeated among the different patterns in the image, leading to a total of 9 distinct endmembers. The abundances of each pixel are sampled from a uniform distribution over the probability simplex of the corresponding dimension (depending on the considered mixture), so that the ASC and ANC are enforced. The mixed pixels are finally generated using the linear mixing model [10]:

$$\mathbf{x}_i = \sum_{j=1}^{p_i} \mathbf{a}_{ij}\mathbf{e}_{ij} + \mathbf{n}_i \quad \forall i = 1, \cdots, n \tag{5}$$

with $\mathbf{x}_i, \mathbf{n}_i \in \mathbb{R}^q$, $\mathbf{e}_{ij} \in \mathbb{R}^q$ is the jth endmember in pixel i, $\mathbf{a}_i \in \mathbb{R}^{p_i}$ is the abundance vector for pixel i, and p_i is the number of endmembers in the considered pixel.

B. Experimental setup

Here we present the experimental methodology we followed to assess how the different algorithms behave to estimate the ID locally. Each of the 25 synthetic datasets was divided into non-overlapping square tiles of various sizes, ranging from 5×5 to 100×100 pixels with steps of 5×5 pixels, and from 100×100 to 300×300 pixels with steps of 10×10 pixels. Therefore, we can study the performance of ID estimation algorithms from a very small local subset (25 pixels) to a global scenario (90000 pixels). The actual ID of each tile depends on which region of the image it falls into (see Fig. 1). The possible actual ID values plotted against the tile length size are shown in Fig. 2: 5 if the tile falls into a rectangle or an ellipse only, 3 if the tile falls into the background only, 6 if the tile falls into the background and one or multiple ellipses, 7 if the tile falls into the background and one or multiple rectangles, and 9 if the tile falls into the background, one or multiple ellipses and one or multiple rectangles. A summary of these considerations is presented in Fig. 3, in which a stacked histogram of the tiles is shown. The first two tile sizes (the bars corresponding to 5×5

[1]http://speclab.cr.usgs.gov/spectral-lib.html

and 10×10 pixels are truncated for the sake of visibility, since 3600 and 900 tiles of this size can be fitted into the image, respectively).

For all the 25 configurations of SNR and number of bands, and for all tile sizes, each ID estimation algorithm is independently run on each tile. Since the noise is here spectrally white, we used the HFC algorithm rather than its noise whitened counterpart, which has minimal impact on the results. The ID estimation is performed in two different cases, depending on the way the noise is estimated: locally or globally. The *local noise estimation* makes use of the pixel values of the local subset only, while the *global noise estimation* makes use of the whole image. In both cases, we employed a fast implementation of Roger's method [38], due to [8], and presented in Appendix.

Next, we describe the quality metrics defined to evaluate the performance of each algorithm. Given the set $\mathcal{S} = \{5, 10, 15, \cdots, 100, 110, 120, \cdots, 300\}$ of window sizes, let $s = card(\mathcal{S})$ be the number of possible lengths. N_i denotes the number of windows of size $\mathcal{S}_i$, $1 \leq i \leq s$. Let d_{ij} and $\hat{d}_{ij}$ respectively denote the actual and estimated ID values of the jth window of size $\mathcal{S}_i$. We define μ_i as the average of the relative absolute errors committed on all windows of size $\mathcal{S}_i$:

$$\mu_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{|d_{ij} - \hat{d}_{ij}|}{d_{ij}}. \tag{6}$$

We also define μ as the average of all the μ_i values for all possible window lengths. This provides a single number to assess the overall performance of the algorithms from the most local (*i.e.* smallest window size) to global ID estimation:

$$\mu = \frac{1}{s} \sum_{i=1}^{s} \mu_i. \tag{7}$$

Finally, σ_i^2 is an estimator of the variance of the absolute relative error committed on all tiles of size $\mathcal{S}_i$:

$$\sigma_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} \left(\frac{|d_{ij} - \hat{d}_{ij}|}{d_{ij}} - \mu_i \right)^2. \tag{8}$$

C. Results

The results of the ID estimations on the 25 synthetic datasets are presented for all algorithms in Figs. 4 to 14. In Fig. 4, the value of μ (see Eq. (7)) is displayed as an image, for all noise powers and numbers of bands, in the case of a local noise estimation. From this figure, we see that the algorithms of the bottom row (PCA, HFC and MMOCA) are nearly insensitive to the number of bands or the noise power. This is

because these algorithms do not require any noise estimation. The results of PCA are highly dependent on the chosen threshold for the explained variance, which is not directly related to the ID value. MMOCA and HFC seem to perform relatively well in all cases. The case of HIDENN is different since estimating the dimension of a manifold is an operation which is highly sensitive to noise, and also dependent of the dimension of the ambient space. We can see that if any of the two tested parameters here are tuned to a more favorable value (higher SNR or lower number of bands), the overall results get better, while in unfavourable configurations, outliers in the estimated values severely decrease the performance. The denoised version of the algorithm, D-HIDENN, helps to reduce the impact of this phenomenon, although it is still present. This algorithm is still sensitive to the noise power, because the noise is not only estimated through its covariance matrix, but also subtracted from the observations. The last four algorithms, Hysime, RMT, ODM, and VCA/NABO present a more similar behavior. They all require a noise estimation, whose performance greatly impacts the ID estimation. We can notice immediately that the ID estimation for these algorithms is much more sensitive to the number of bands than to the noise power, which can be explained by the fact that the algorithm used for the noise estimation is based on a regression of each band on the others, an operation becoming less precise when the number of bands increases. This is due to the multicollinearity effect: when there are more bands, they are more correlated since adjacent wavelengths become closer and closer, and there are multiple good candidates for the regression coefficients. Hence a small change in the data can induce a large change in the regression coefficients (see the Appendix).

Fig. 5 shows the same metric μ in the case of a global noise estimation. For MMOCA, HFC, PCA and HIDDEN, the results are very similar to the ones obtained for the local noise estimation since these algorithms do not estimate the noise (they are not exactly equivalent since for both experiments a different noise realization was used). However, for the other algorithms, notable differences are visible: the algorithms perform much better in the least favorable cases. As we will see in the following, this is due to the fact that global noise estimation allows a much better ID estimation in small windows (provided the noise distribution is the

same everywhere), where a precise local noise estimation is impossible because of the too low number of samples. The ODM algorithm does not seem very affected by the change in the noise estimation. This probably comes from the paradigm used in this algorithm: the objective of ODM is to identify the signal as an outlier in a noise distribution.

Fig. 6 sums up these considerations by showing the difference between the μ values estimated using local and global noise estimations, $\mu_{\text{local}} - \mu_{\text{global}}$. Thus, a positive value means that local noise estimation performed worse than global noise estimation, and vice versa. From the figure, it is clear than in almost all cases, global noise estimation performs better for algorithms sensitive to the way the noise is estimated. We see that when the configuration becomes more favorable, the results of local noise estimation become closer to the ones with global noise estimation. It happens in some cases that estimating the noise locally performs slightly better than doing it globally, but in most cases the results show that global noise estimation is much more robust.

In Figs. 7 and 8 we show in detail the results of the ID estimation for all algorithms and all window lengths in one representative noise and band number configuration, respectively 30 dB and 120 bands, corresponding to the central pixels of the images of Figs. 4 and 5. This configuration was chosen because it is representative of many real scenarios. These figures are to be compared to the actual ID values in Fig. 2. Two patterns in the ID estimations can be found for most algorithms: (i) a window size range where the ID estimate has a peak, which is too large, and (ii) a set of window sizes for which there is a slow stabilization of the results, until the support of the global image is reached.

Fig. 7 shows, as expected, that for local noise estimation and for most algorithms, the ID estimation provides erroneous values for the smallest windows. For HySIME, RMT and VCA/NABO we can observe an important peak in the estimated ID values for a certain window size. This peak means that nearly all the values between zero and the maximum of the peaks were attained for the different windows of this size, confirming the instability of the algorithms, and more specifically of the noise estimation for small windows. The height and position of the peak depends on the noise and band configuration, as we will discuss in the following. The peak is also present, to a lesser extent for HIDENN/D-HIDDEN
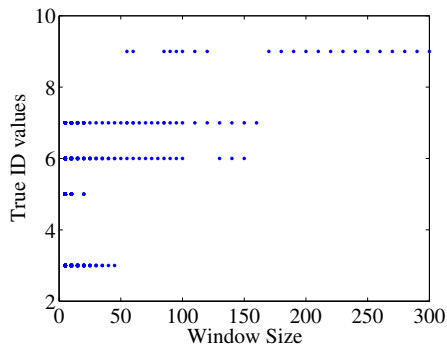

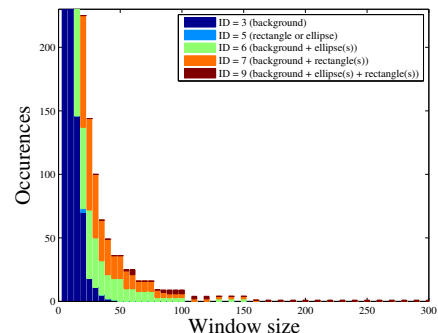Fig. 2. True ID values plotted against size of the subset.

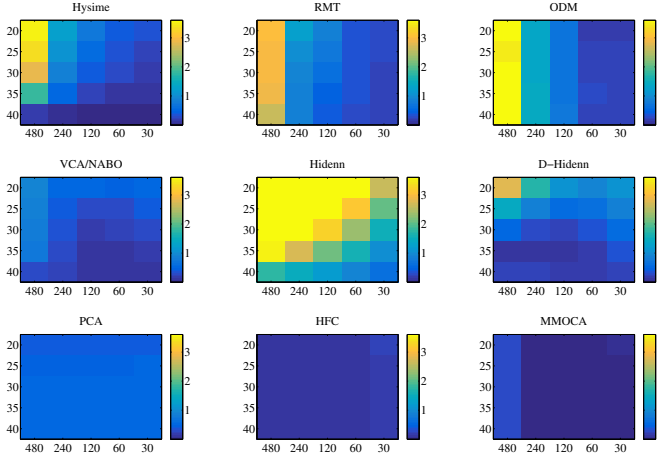Fig. 3. Stacked Histogram of the tiles for each size, depending on the ID value.

Fig. 4. μ_{local} as a function of SNR in dB (y axis) and number of bands (x axis). The color scale ranges from blue (0.0002) to yellow (3.6 or higher).
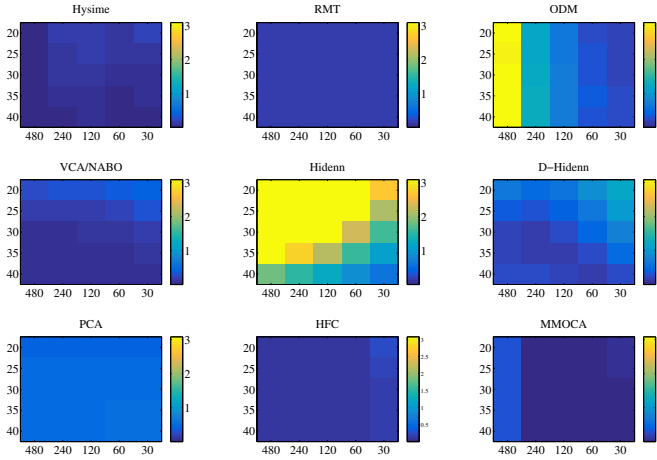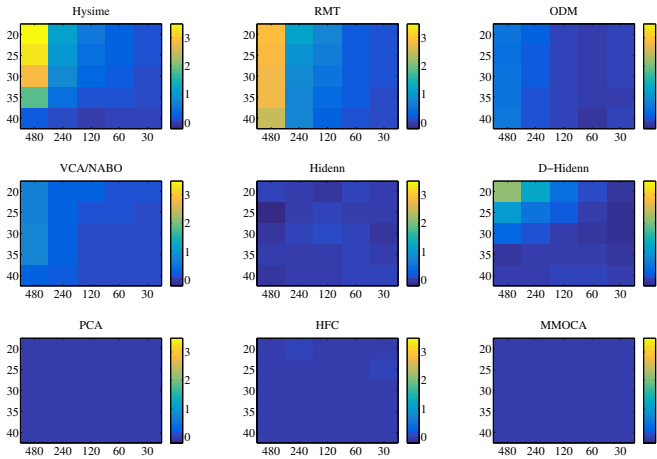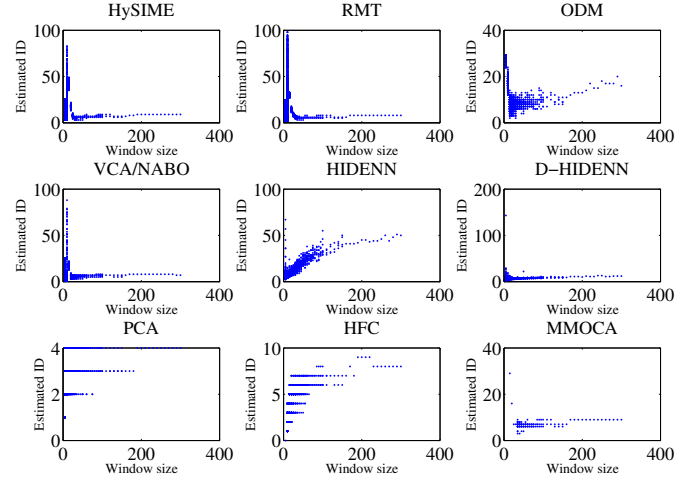

Fig. 7. Estimated ID in the case of local noise estimation plotted against window size for all algorithms, for SNR = 30dB and 120 bands.

Fig. 5. μ_{global} as a function of SNR in dB (y axis) and number of bands (x axis). The color scale ranges from blue (0.0002) to yellow (3.1 or higher).

Fig. 6. $\mu_{local} - \mu_{global}$ as a function of SNR in dB (y axis) and number of bands (x axis). The color scale ranges from blue (-0.48) to yellow (3.49 or higher).

algorithms because they estimate the dimension of a manifold with too few samples, in which case noise is mistaken for signal, especially for small regions which are likely to have a low rank. ODM also shows this peak because of the noise estimation, although its importance is mitigated by the outlier in noise paradigm. For MMOCA, the peak has another origin since the low dimensional subspace is estimated by resorting to an optimization problem. In this case, for too small windows, this problem is very ill-conditioned, which entails erroneous estimations. Below a certain size, singular matrices appear during the estimation and the algorithm fails to produce an estimated value. Finally, PCA seems affected inasmuch as the (overall small) variance seems harder to capture with only a few dimensions. Finally, HFC seems to be less affected by the number of pixels in the local regions, since the estimation does not show a peak in the ID values but more a linear increase with the window size.

In Fig. 8, the same plots are presented, but in this case for global noise estimation. As before, HIDDEN, PCA, HFC and MMOCA are not affected since they do not require a noise estimation step. ODM does not seem very affected either, probably because of its particular signal and noise model. For HySIME the peak is also present, because while the noise correlation matrix estimation is much more precise, the signal correlation matrix still has to be estimated in a small dataset. However, the peak decreases faster and is less important in amplitude than in the local case. However, RMT and VCA/NABO, seem very affected by the change in noise estimation. The corresponding plots are now quite similar in shape to the actual ID values in Fig. 2 (VCA/NABO does not require the estimation of the signal covariance matrix). Finally, for D-HIDENN, global noise estimation allows the suppression of the most aberrant outliers from the estimated ID values. Overall, it seems that global noise estimation is very beneficial to ID estimation, but it relies on the assumption that the noise is spatially i.i.d. in all the image.

Another aspect of local ID estimation shown in Figs. 9 to 12, is the transition between erroneous ID estimations for small
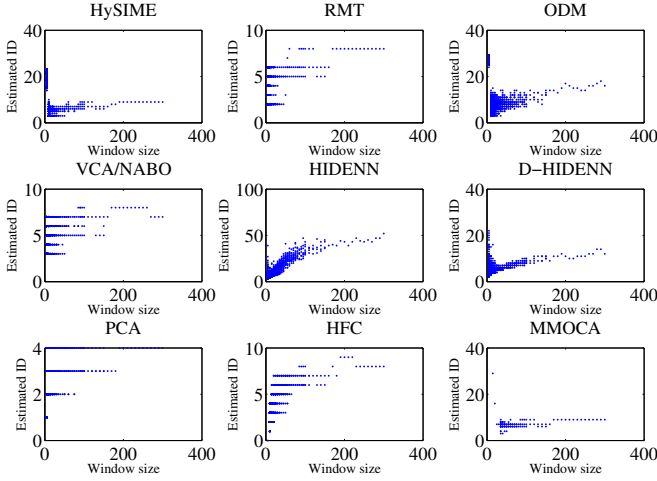
Fig. 8. Estimated ID in the case of global noise estimation plotted against window size for all algorithms, for SNR = 30dB and 120 bands.

Fig. 9. μ_i as a function of the window size $i \in S$ in the case of local noise estimation, for SNR = 30dB and 120 bands. The standard deviation of the estimated values is represented by the red curves.

windows to correct ID estimations when the window sizes get sufficiently high, until the size of the whole image is reached. The quality metric μ_i (see Eq. (6)) is plotted in blue against the window length $\mathcal{S}_i$, while the dashed red curves correspond to the quality metric plus and minus one standard deviation (the standard deviation is defined as the square root of Eq. (8)) are shown in dashed red. Fig. 9 shows the value of μ_i for the local noise estimation. Fig. 10 is simply a zoomed version of Fig. 9 to show what happens after the peak in the estimations. From these two figures, we clearly see that for the algorithms concerned, the peak is accompanied by a large variance in the estimations, which quickly decreases as the number of samples get higher. Note that for large window sizes, this phenomenon is also due to the fact that there are fewer windows of this size that we can fit into the image. For global noise estimation (Figs. 11 and 12), we see that apart from HySIME and ODM, the estimations in small windows are less subject to a high variance, and the estimation for each window size in small windows is much more precise, which confirms the results of the previous figures. The observations drawn from these figures allow one to define empirically a size threshold above which the noise estimation will be reliable.

Finally, Figs. 13 and 14 depict a last but nonetheless important aspect of the noise estimation: for which window size does the peak appear? We discuss this particular point, very linked to the definition of a reliability threshold for the estimation, considering this time several band number configurations at fixed SNR, and vice versa, but only for the algorithms concerned (i.e. HySIME, RMT, ODM, VCA/NABO and MMOCA). For local noise estimation (Fig. 13), we immediately see that the size at which the peak appears for all algorithms is much more related to the number of bands considered in the estimation than it is to the noise level, which more influences its height. The higher the number of bands, the later the peak appears, which means that larger windows will be necessary for a correct ID estimation. In the case of global noise estimation (Fig. 14), many cases are favorable enough for the algorithms not to present a peak, since the noise
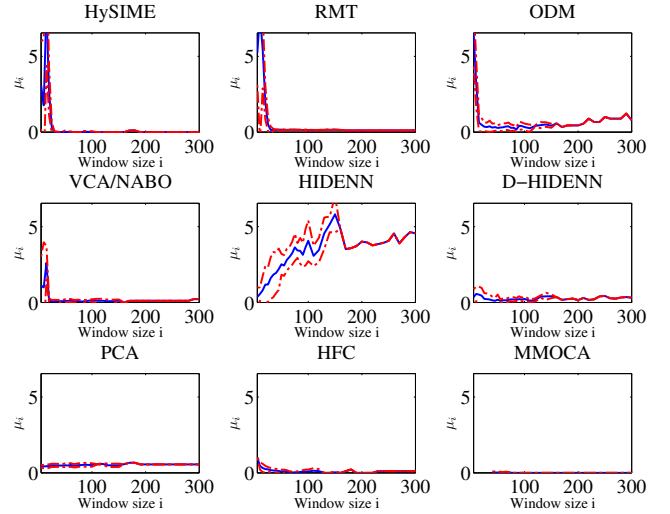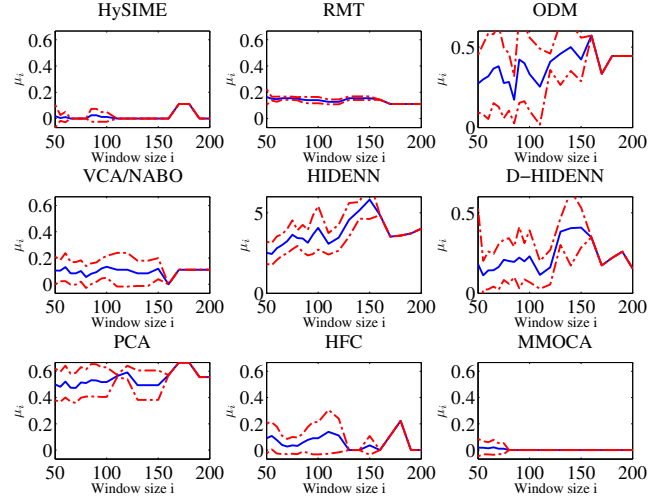

Fig. 10. μ_i as a function of the window size $i \in S \cap [50, 200]$ in the case of local noise estimation, for SNR = 30dB and 120 bands. The results for Hidenn are shown on a different scale than the other algorithms.

is correctly estimated (except for MMOCA which does not require a noise estimation), and its position is less influenced by the number of bands.

IV. EXPERIMENTS ON REAL DATASETS

In this section we present the experiments we performed on two real datasets in order to validate the observations made on the synthetic datasets.

A. Datasets

The first dataset we used is an image acquired by NASA's AVIRIS sensor over the Cuprite mining district in Nevada, USA. It is a 350×350 image comprising 188 spectral bands, which has been often used to validate ID estimation algorithms. We estimated the SNR of each band of this image using the algorithm presented in the Appendix and obtained an
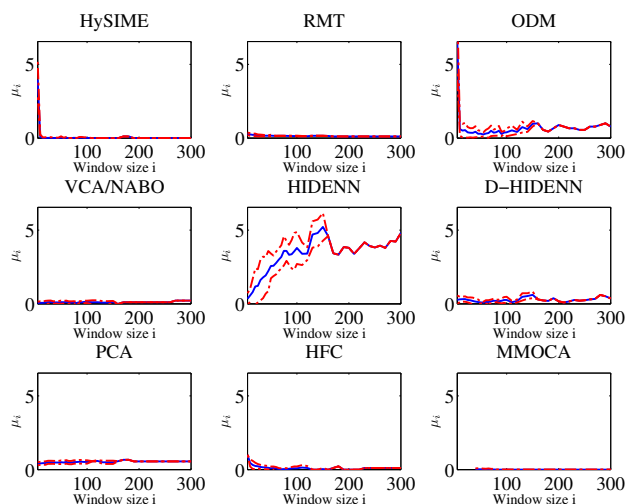
Fig. 11. μ_i as a function of the window size $i \in S$ in the case of global noise estimation, for SNR = 30dB and 120 bands.
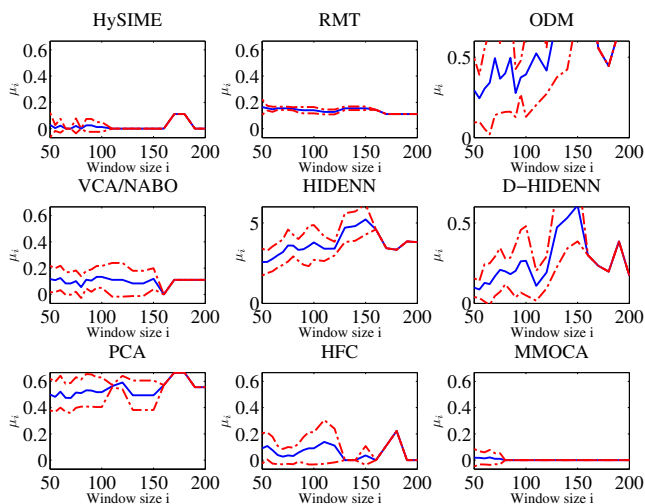

Fig. 12. μ_i as a function of the window size $i \in S \cap [50, 200]$ in the case of global noise estimation, for SNR = 30dB and 120 bands. The results for Hidenn are shown on a different scale than the other algorithms.
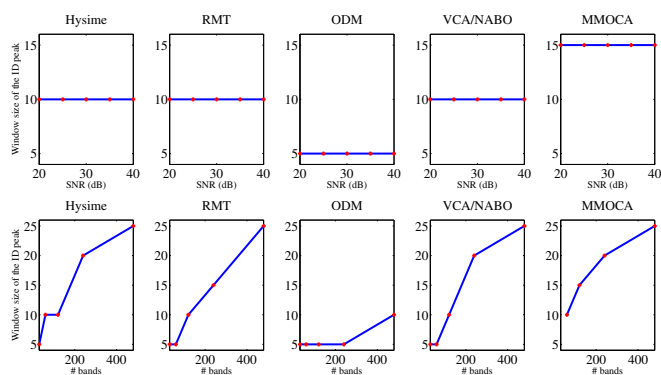

Fig. 13. Window size of the ID peak (corresponding to spurious estimation) plotted against SNR for 120 bands (top row) or against number of bands at fixed SNR = 30dB (bottom row), for local noise estimation. Blank values indicate that no peak is present in the corresponding configuration.
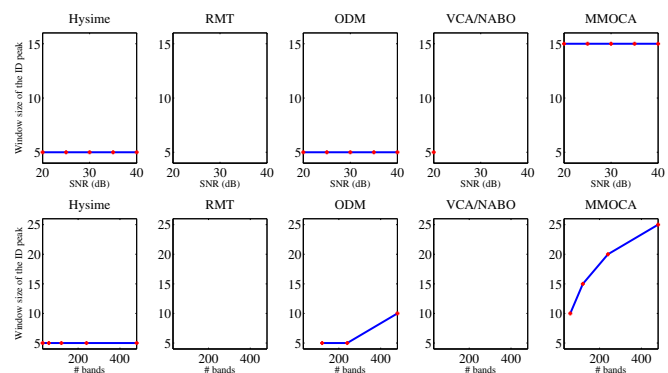

Fig. 14. Window size of the ID peak (corresponding to spurious estimation) plotted against SNR for 120 bands (top row) or against number of bands at fixed SNR = 30dB (bottom row), for global noise estimation. Blank values indicate that no peak is present in the corresponding configuration.
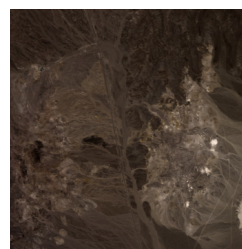

Fig. 15. RGB representation of the Cuprite dataset.

average SNR (over all bands) of 27dB. It is usually considered that there are at least 17 different materials (mostly minerals) in this image, based on ground observations and mineral maps of the site[2]. In addition, according to experiments performed in [15], the noise in this image in not very spectrally correlated. An RGB representation of this image is shown in Fig 15, using bands 40,30 and 20 of the image.

The second dataset was acquired by the CASI 1500 sensor over the Barrax region, in the south of Spain, in 2005[3]. The 97×847 image comprises 144 bands in the VNIR region (370-1050 nm) and the estimated average SNR is 43dB. A RGB representation of this scene is shown in Fig. 16, using bands 52,35 and 25.

B. Experimental setup

For both datasets, as for the synthetic data, we perform local ID estimation on non-overlapping square tiles of different sizes, from 5×5 to 100×100 pixels size with steps of 5×5 pixels, and from then on, from 100×100 pixels size to the maximum possible with steps of 10×10 pixels. For the Barrax dataset, we considered only the tiles in which no unobserved values were present. For both datasets, ID

[2]http://speclab.cr.usgs.gov/cuprite95.tgif.2.2um_map.gif.
[3]http://www.uv.es/~leo/sen2flex/

Fig. 16. RGB representation of the Barrax dataset.

estimation was carried out for all algorithms, for local and global noise estimation. In the absence of ground truth, we cannot compute the metrics used for the synthetic datasets, but we can compare qualitatively the shapes of the local ID plots to the observations made for the synthetic data.

C. Results

First, we compared the results of the ID estimations for both datasets in the case of local noise estimation (see Figs. 17 and 18). The results show that in both cases, the general behavior of the algorithms is similar to that for the synthetic datasets. We can see that HySIME, RMT, ODM and VCA/NABO show a clear peak in the ID estimations for small windows, which is clearer for the Cuprite dataset, probably because it is noisier than the Barrax data. The peaks appear roughly for the same window sizes as in the simulated data, which is logical since the number of bands is comparable in both datasets. Then the peaks quickly decrease and seem to stabilize around different values for each algorithm when we approach global ID estimation. Note that the zero values which can appear for very small windows and some algorithms are due to a very poor noise estimation. For example, in the case of HySIME, the estimated noise values were so high that the curves showing the projection error and the noise power never crossed, hence the zero estimated value. For HIDENN and D-HIDENN, the results are consistent with the synthetic data: large outliers appear for very small windows, and then, the algorithm quickly stabilizes with smaller outlier values if the data has been de-noised beforehand. The performance of PCA still depends heavily on the arbitrary choice of the variance percentage (still 95% here). For a percentage lower than 95%, the estimated ID is rarely above 3 for the global images, showing that the ID is not linked to the variance of the data cloud. The performance improves for larger thresholds, but the tuning is empirical and data-dependent. HFC still obtains a more or less linear behavior with the increase in window size. Finally, the MMOCA algorithm fails to produce a value for a large range of window sizes because of the ill-conditioning of the subspace estimation problem (which explains why only windows bigger than 50×50 pixels appear for the Cuprite dataset and windows over 20×20 pixels for the Barrax dataset).

For global noise estimation, and for the algorithms requiring noise estimation, the results are still consistent with the ones obtained on the synthetic data (see Figs. 19 and 20). The peak in the estimated values is still present for the HySIME and ODM algorithms with the Cuprite data, but very attenuated with respect to the case of local noise estimation. For RMT and VCA/NABO, as for the synthetic datasets in such noise and band configuration, the peak has vanished. We can see that when the windows get larger, both noise estimation strategies perform in an increasingly similar way, as expected. From the figures above, we can define an empirical threshold above which the ID estimation would be reliable: for instance, for the Cuprite dataset, we can set the window size threshold to 30×30 pixels for the case of local noise estimation, and a window size threshold of 15×15 pixels for global noise estimation, for all algorithms. Note that the algorithms can differ a lot in
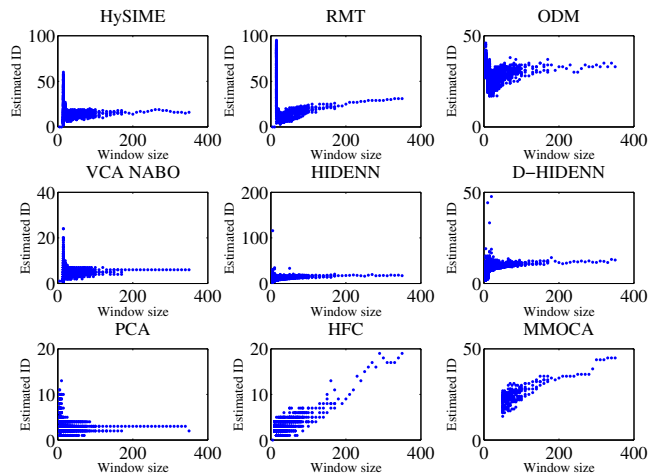

Fig. 17. Estimated ID of the Cuprite dataset in the case of local noise estimation plotted against window size for all algorithms.
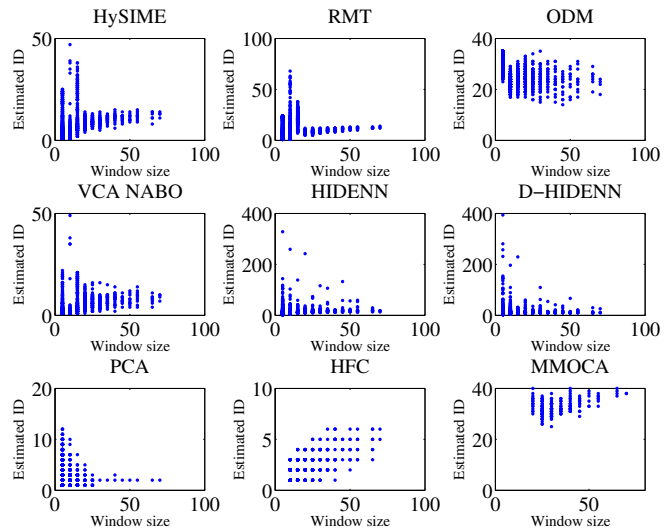

Fig. 18. Estimated ID of the Barrax dataset in the case of local noise estimation plotted against window size for all algorithms. Some outliers for Hidenn and D-Hidenn are not displayed.

their estimated global ID values (in the Cuprite case, around 20 for HySIME, HFC and HIDENN, and less for VCA/NABO and PCA, and much higher for RMT, ODM and MMOCA).

Notice that for both datasets, the estimated global noise ID values for large window sizes match the ones with the same windows, but obtained in a local noise estimation context. This tends to confirm that the spatial i.i.d. assumption for the noise holds in these datasets.

V. DISCUSSION

In this section, we summarize the observations made for the synthetic and real datasets, and we provide some indications on how to use the ID estimation algorithms in a local setting. From the results, we observed that there are three main parameters influencing local ID estimation:

1) The number of pixels in the local region.
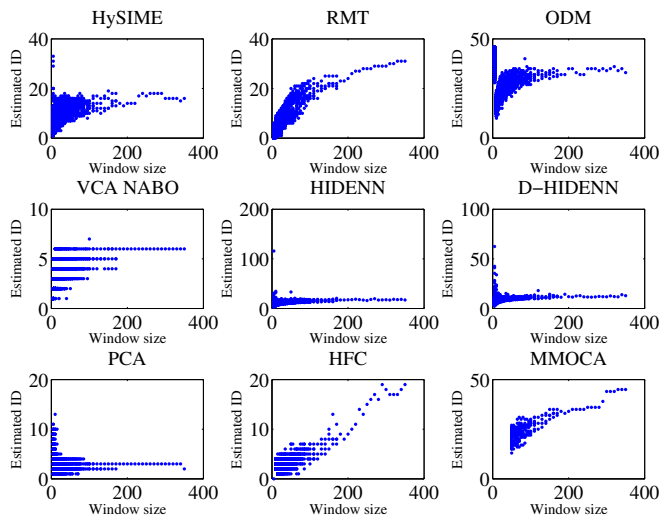2) The number of spectral bands.
3) The noise level.

Fig. 19. Estimated ID of the Cuprite dataset in the case of global noise estimation plotted against window size for all algorithms
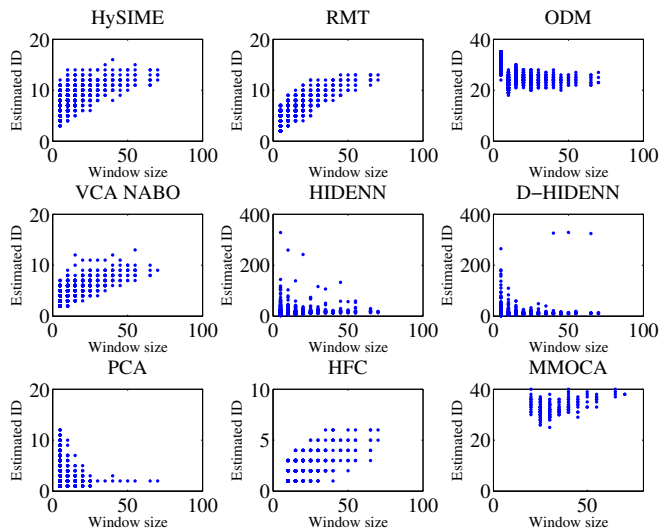

Fig. 20. Estimated ID of the Barrax dataset in the case of global noise estimation plotted against window size for all algorithms. Some outliers for Hidenn and D-Hidenn are not displayed.

A majority of the tested algorithms require a noise estimation step. For these algorithms, a clear pattern can be seen when estimating ID in regions of the datasets at different scales, often comprising a peak in the overestimation of the ID in unfavorable cases. This pattern is especially present when the noise is estimated locally, in each tile of the image. When such a peak appears, its amplitude increases with the noise power, while its position is especially determined by the number of bands in the image: the more bands, the larger the window size where the peak appears, which means that the ID estimation will be unreliable for larger windows than if there were fewer spectral bands. This phenonenon is linked to the curse of dimensionality: since higher dimensional spaces are sparser, more samples are required in order that estimation algorithms obtain reliable results. In addition, the multicollinearity phenomenon in high dimensions can also

hamper noise estimation strategies exploiting the between band correlations.

These considerations raise the question of how to choose a minimum value for the window size, below which the ID estimation is unreliable. One has to take into account the position of the peak, but also the speed of the decrease after it. A threshold can be roughly defined visually from plots similar to those in Fig. 7. The most favorable configuration for local ID estimation is then a low number of bands and a good SNR. In any case, for those algorithms, a global noise estimation is preferable, since it largely reduces the uncertainty due to the noise estimation. The only case when a local noise estimation is preferable is in the case of a spatially non-i.i.d. noise. MMOCA does not require a noise estimation, but fails to produce a result when the underlying optimization problem is too ill-conditioned. For the other algorithms, HIDDEN and D-HIDDEN have a tendency to produce large outliers when the number of samples is too few. HFC seems to behave more naturally for small windows, since small estimated ID values come out in this case.

Next, we need to determine which algorithm to choose to estimate the ID locally. To guide the reader in his choice, we summarize below and in Table II the strengths and weaknesses of each tested algorithm:

- HySIME: relatively robust for local ID estimation, provided the noise is estimated globally, but still subject to overestimation when the window size is too small because it requires the estimation of the signal correlation matrix. It is also relatively fast and produced good results on synthetic datasets.
- RMT: comparable to HySIME, with good performance on the synthetic datasets. It does not show a peak in the ID values when the noise is estimated globally (at least for reasonable band and noise configurations). Relatively fast.
- ODM: Relatively fast, but less precise and more sensitive to the number of bands than the previous two algorithms. Less sensitive to local/global noise estimation.
- VCA/NABO: same advantages as the previous ones, which fall in the same category (although NABO is not eigenvalue-based), but quite computationally intensive since it requires a spectral unmixing step. Slightly more sensitive to noise than most algorithms.
- HIDENN / D-HIDENN: Not eigenvalue based, but very sensitive to noise, even though its effect can be mitigated but not suppressed when a de-noising step is performed. Poor precision in low SNR cases. Relatively slow.
- HFC: Practically insensitive to noise and band number. Provides underestimated ID values independently of the scale, although they are overall relatively accurate. Depends on a user-defined threshold. It can be argued that it is theoretically wrong and that the results depend on the average values of the bands and not directly on the ID of the data. Fast.
- PCA: Definitely not a good candidate: the performance is conditioned to the arbitrary choice of the threshold.
- MMOCA: Does not require a noise estimation, good performance. Computationally rather intensive, especially

for small windows. Does not work for too small windows because of ill conditioning.

VI. CONCLUSIONS

In this paper, we presented a study of several Intrinsic Dimensionality estimation algorithms for hyperspectral imaging in the context of local ID estimation. The results on both synthetic and real data show that in general, when trying to use these algorithms on local subsets of a large image, one has to be careful with: (i) the number of samples in the subsets, which have to be sufficiently numerous for estimation processes to be reliable; and, (ii) when noise estimation or denoising is required, a local approach will yield a decrease in performance, although this problem can be highly mitigated by estimating the noise on the whole image. Two other important factors also have consequences on the results: the noise level and the number of spectral bands. A low SNR and a high number of bands will increase the chance of mistaking noise for signal and make the estimation more prone to fail in higher dimensional settings, respectively. We summed up the properties of nine ID estimation algorithms and showed how they behaved in local areas of the image, and evidenced their respective strengths and weaknesses for local ID estimation. Future work will includeconsiderations developed in this paper in the pipelines of algorithms designed for other applications on hyperspectral imaging which resort to local subsets of the image, such as local spectral unmixing or local super-resolution techniques.

APPENDIX
NOISE ESTIMATION

The noise estimation algorithm used in the experiments is based on the use of the high correlation between adjacent bands and was first brought to the hyperspectral imaging community in [38]. The idea behind this strategy is to perform a linear regression of each spectral band on all the other bands, that is to say to express all the pixels from one spectral band (stacked into a $n \times 1$ vector) as a linear combination of the pixel vectors of all the other bands. If we denote by $\mathbf{X}_{\neq k}$ the data matrix $\mathbf{X}$ with the kth row $\mathbf{x}_k$ (one entire band) removed, we can estimate the optimal regression parameters $\mathbf{b}_k \in \mathbb{R}^{q-1}$ of $\mathbf{x}_k$ on $\mathbf{X}_{\neq k}$ in a least square sense by:

$$\mathbf{b}_k = \mathbf{x}_k \mathbf{X}_{\neq k}^\top (\mathbf{X}_{\neq k} \mathbf{X}_{\neq k}^\top)^{-1} \qquad (9)$$

and we can finally estimate the noise vector $\mathbf{w}_k \in \mathbb{R}^n$ in band k by:

$$\mathbf{w}_k = \mathbf{x}_k - \mathbf{b}_k \mathbf{X}_{\neq k} \qquad (10)$$

This difference between the observations in the considered spectral band and the result of the regression is assumed to be due to noise, providing the estimated noise values and allowing the estimation of the noise sample correlation matrix, which is assumed to be diagonal (and hence does not consider spectrally correlated noise) with difference variances in each spectral band. Other methods exist to perform hyperspectral noise estimation, such as the so-called shift difference method [24] for instance, which assumes that the differences between adjacent pixels are mainly due to different realizations of i.i.d. noise, the signal component being practically the same. Two other noise estimation strategies [35] [39] which have been

used in hyperspectral data analysis, have been evaluated and discussed (especially for their behavior in case of correlated noise) in [15].

REFERENCES

[1] N. Acito, M. Diani, and G. Corsini. A new algorithm for robust estimation of the signal subspace in hyperspectral images in the presence of rare signal components. *Geosc. and Rem. Sens., IEEE Trans. on*, 47(11):3844–3856, Nov 2009.
[2] N. Acito, M. Diani, and G. Corsini. Hyperspectral signal subspace identification in the presence of rare signal components. *Geosc. and Rem. Sens., IEEE Trans. on*, 48(4):1940–1954, April 2010.
[3] N. Acito, M. Diani, and G. Corsini. Hyperspectral signal subspace identification in the presence of rare vectors and signal-dependent noise. *Geosc. and Rem. Sens., IEEE Trans. on*, 51(1):283–299, Jan 2013.
[4] C. Andreou and V. Karathanassi. Estimation of the number of endmembers using robust outlier detection method. *Sel. Topics in App. Earth Obs. and Rem. Sens., IEEE Journal of*, 7(1):247–256, Jan 2014.
[5] P. Bajorski. Does virtual dimensionality work in hyperspectral images? In *Proc. SPIE*, volume 7334, pages 73341J–73341J–11, 2009.
[6] P. Bajorski. Second moment linear dimensionality as an alternative to virtual dimensionality. *Geosc. and Rem. Sens., IEEE Trans. on*, 49(2):672–678, Feb 2011.
[7] J. M. Bioucas-Dias and José M. P. Nascimento. Estimation of signal subspace on hyperspectral data. In *Proc. SPIE*, volume 5982, pages 59820L–59820L–8, 2005.
[8] J.M. Bioucas-Dias and J.M.P. Nascimento. Hyperspectral subspace identification. *Geos. and Rem. Sens. IEEE Trans. on*, 46(8):2435–2445, Aug 2008.
[9] J.M. Bioucas-Dias, A Plaza, G. Camps-Valls, P. Scheunders, N.M. Nasrabadi, and J. Chanussot. Hyperspectral remote sensing data analysis and future challenges. *Geosc. and Rem. Sens. Mag., IEEE*, 1(2):6–36, June 2013.
[10] J.M. Bioucas-Dias, A Plaza, N. Dobigeon, M. Parente, Qian Du, P. Gader, and J. Chanussot. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *Sel. Topics in App. Earth Obs. and Rem. Sens., IEEE Journal of*, 5(2):354–379, April 2012.
[11] F. Camastra. Data dimensionality estimation methods: a survey. *Pattern Recognition*, 36(12):2945 – 2954, 2003.
[12] K. Canham, A Schlamm, A Ziemann, B. Basener, and D. Messinger. Spatially adaptive hyperspectral unmixing. *Geosc. and Rem. Sens., IEEE Trans. on*, 49(11):4248–4262, Nov 2011.
[13] K. M. Carter, R. Raich, and A.O. Hero. On local intrinsic dimension estimation and its applications. *Sig. Proc., IEEE Trans. on*, 58(2):650–663, Feb 2010.
[14] K. Cawse-Nicholson, S.B. Damelin, A Robin, and M. Sears. Determining the intrinsic dimension of a hyperspectral image using random matrix theory. *Im. Proc., IEEE Trans. on*, 22(4):1301–1310, April 2013.
[15] K. Cawse-Nicholson, A. Robin, and M. Sears. The effect of correlation on determining the intrinsic dimension of a hyperspectral image. *Sel. Topics in App. Earth Obs. and Rem. Sens., IEEE Journal of*, 6(2):482–487, April 2013.
[16] C.-I. Chang and Q. Du. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *Geosc. and Rem. Sens., IEEE Trans. on*, 42(3):608–619, March 2004.
[17] C.-I. Chang, W. Xiong, H.-M. Chen, and J.-W. Chai. Maximum orthogonal subspace projection approach to estimating the number of spectral signal sources in hyperspectral imagery. *Sel. Topics in Sig. Proc., IEEE Journal of*, 5(3):504–520, June 2011.

Algorithm / Property	HySIME	RMT	ODM	HIDENN	D-HIDENN	VCA/NABO	PCA	HFC	MMOCA
Noise level sensitivity	+	+	+	++	++	+	-	- -	- -
Sample size sensitivity	++	+	+	++	++	+	+	-	+
Band number sensitivity	+	+	++	-	-	+	- -	- -	
Computational burden	-	- -	- -	+	+	++	- -	- -	+
Overall estimation error on synthetic datasets	-	- -	+	++	+	-	+	-	- -

TABLE II

STRENGTHS AND WEAKNESSES OF THE TESTED ALGORITHMS FOR LOCAL ID ESTIMATION. ONE OR TWO + SIGNS MEANS HIGH OR VERY HIGH, AND ONE OR TWO - SIGNS MEANS LOW OR VERY LOW.

[18] C.-I. Chang, W. Xiong, and C.-H. Wen. A theory of high-order statistics-based virtual dimensionality for hyperspectral imagery. *Geosc. and Rem. Sens., IEEE Trans. on*, 52(1):188–208, Jan 2014.

[19] N. Dobigeon, J.-Y. Tourneret, C. Richard, J.C.M. Bermudez, S. McLaughlin, and AO. Hero. Nonlinear unmixing of hyperspectral images: Models and algorithms. *Sig. Proc. Mag., IEEE*, 31(1):82–94, Jan 2014.

[20] L. Drumetz, M. A. Veganzones, R. Marrero, G. Tochon, M. Dalla Mura, A. Plaza, and J. Chanussot. Binary partition tree-based local spectral unmixing. In *Proc. IEEE WHISPERS*, 2014.

[21] K. Fukunaga. *Introduction to Statistical Pattern Recognition, Second Edition*. Academic Press, 2 edition, October 1990.

[22] L. Gao, Q. Du, B. Zhang, W. Yang, and Y. Wu. A comparative study on linear regression-based noise estimation for hyperspectral imagery. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 6(2):488–498, April 2013.

[23] M.A. Goenaga, M.C. Torres-Madronero, M. Velez-Reyes, S.J. Van Bloem, and J.D. Chinea. Unmixing analysis of a time series of hyperion images over the guanica dry forest in puerto rico. *Sel. Topics in App. Earth Obs. and Rem. Sens., IEEE Journal of*, 6(2):329–338, April 2013.

[24] A.A. Green, M. Berman, P. Switzer, and M.D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *Geosc. and Rem. Sens., IEEE Trans. on*, 26(1):65–74, Jan 1988.

[25] M. Hasanlou and F. Samadzadegan. Comparative study of intrinsic dimensionality estimation and dimension reduction techniques on hyperspectral images using k-nn classifier. *Geosc. and Rem. Sens. Letters, IEEE*, 9(6):1046–1050, Nov 2012.

[26] R. Heylen and P. Scheunders. Hyperspectral intrinsic dimensionality estimation with nearest-neighbor distance ratios. *Sel. Topics in App. Earth Obs. and Rem. Sens., IEEE Journal of*, 6(2):570–579, April 2013.

[27] I.T. Jolliffe. *Principal Component Analysis*. Springer Verlag, 1986.

[28] O. Kuybeda, D. Malah, and M. Barzohar. Rank estimation and redundancy reduction of high-dimensional noisy signals with preservation of rare vectors. *Sig. Proc., IEEE Trans. on*, 55(12):5579–5592, Dec 2007.

[29] D. Landgrebe. Hyperspectral image data analysis. *Sig. Proc. Mag., IEEE*, 19(1):17–28, Jan 2002.

[30] G. Licciardi, M. A. Veganzones, M. Simoes, J. Bioucas-Dias, and J. Chanussot. Super-resolution of hyperspectral images using local spectral unmixing. In *Proc. IEEE WHISPERS*, 2014.

[31] J. Liu, J. Zhang, Y. Gao, C. Zhang, and Z. Li. Enhancing spectral unmixing by local neighborhood weights. *Sel. Topics in App. Earth Obs. and Rem. Sens., IEEE Journal of*, 5(5):1545–1552, Oct 2012.

[32] W.-K. Ma, J.M. Bioucas-Dias, Tsung-Han Chan, N. Gillis, P. Gader, AJ. Plaza, A Ambikapathi, and Chong-Yung Chi. A signal processing perspective on hyperspectral unmixing: Insights from remote sensing. *Sig. Proc. Mag., IEEE*, 31(1):67–81, Jan 2014.

[33] Z. Ma. Accuracy of the Tracy–Widom limits for the extreme eigenvalues in white Wishart matrices. *Bernoulli*, 18(1):322–359, 02 2012.

[34] R. Marrero, S. Lopez, G.M. Callico, M.A. Veganzones, A. Plaza, J. Chanussot, and R. Sarmiento. A novel negative abundance-oriented hyperspectral unmixing algorithm. *Geoscience and Remote Sensing, IEEE Transactions on*, 53(7):3772–3790, July 2015.

[35] P. Meer, J. Jolion, and A. Rosenfeld. A fast parallel algorithm for blind estimation of noise variance. *Pattern Analysis and Mach. Intel., IEEE Trans. on*, 12(2):216–223, Feb 1990.

[36] J.M.P. Nascimento and J.M. Bioucas Dias. Vertex component analysis: a fast algorithm to unmix hyperspectral data. *Geosc. and Rem. Sens., IEEE Trans. on*, 43(4):898–910, April 2005.

[37] A. Robin, K. Cawse-Nicholson, A. Mahmood, and M. Sears. Estimation of the intrinsic dimension of hyperspectral images: Comparison of current methods. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 8(6):2854–2861, June 2015.

[38] R.E. Roger. Principal components transform with simple, automatic noise adjustment. *Int. Journal of Rem. Sens.*, 17(14):2719–2727, 1996.

[39] R.E. Roger and J.F. Arnold. Reliably estimating the noise in AVIRIS hyperspectral images. *Int. Journal of Rem. Sens.*, 17(10):1951–1962, 1996.

[40] A. Schlamm, D. Messinger, and W. Basener. Geometric estimation of the inherent dimensionality of single and multi-material clusters in hyperspectral imagery. *Journal of Applied Rem. Sens.*, 3(1):033527–033527–16, 2009.

[41] B. Somers, G. P. Asner, L. Tits, and P. Coppin. Endmember variability in spectral mixture analysis: A review. *Rem. Sens. of Environment*, 115(7):1603 – 1616, 2011.

[42] B. Somers, M. Zortea, A Plaza, and G.P. Asner. Automated extraction of image-based endmember bundles for improved spectral unmixing. *Sel. Topics in App. Earth Obs. and Rem. Sens., IEEE Journal of*, 5(2):396–408, April 2012.

[43] M. A. Veganzones, M. Simoes, G. Licciardi, J. Bioucas-Dias, and J. Chanussot. Hyperspectral super-resolution of locally low rank images from complementary multisource data. In *Proc. IEEE Int. Conf. on Im. Proc. (ICIP)*, 2014.

[44] M.A Veganzones, G. Tochon, M. Dalla-Mura, AJ. Plaza, and J. Chanussot. Hyperspectral image segmentation using a new spectral unmixing-based binary partition tree representation. *Im. Proc., IEEE Trans. on*, 23(8):3574–3589, Aug 2014.

[45] A. Zare and K.C. Ho. Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing. *Sig. Proc. Mag., IEEE*, 31(1):95–104, Jan 2014.

Lucas Drumetz received the M.Sc degree in Electical Engineering from the Grenoble Institute of Technology (Grenoble-INP), France, in 2013. Since then, he has been working towards the PhD degree in signal and image processing at the University of Grenoble-Alpes. His research interests focus on spectral unmixing of hyperspectral data, as well as the applications of optimization techniques to hyperspectral images and remote sensing. He is a reviewer for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

Miguel A. Veganzones (M'12) received the M.Sc. and Ph.D. degrees in computer science from the Basque Country University (UPV/EHU), Donostia, Spain, in 2005 and 2012 respectively. Since 2012, he is a postdoctoral researcher at the Images-Signal department in the GIPSA-lab, Grenoble, France. His current main research activities focus in the analysis of hyperspectral images by means of computational intelligence, statistical techniques and tensor data analysis. He is a Reviewer of IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING and the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.

Rubén Marrero Gómez received the Engineer and the M.Sc. in Telecommunication Technologies from Las Palmas de Gran Canaria University (ULPGC), Spain, in 2011 and 2012 respectively. He has also received the Electronics Engineer from the same University in 2013. In 2012 he focused his research on hyperspectral imaging working at the Institute for Applied Microelectronics (IUMA), ULPGC and at the Institut de Planetologie et d'Astrophysique de Grenoble (IPAG), Université Joseph Fourier (UJF), CNRS, France.

Giorgio A. Licciardi (M'11) received the M.S. degree in telecommunication engineering and the Ph.D. degree in "geoinformation" from the Tor Vergata University, Rome, Italy, in 2005 and 2010, respectively. In 2010 he joined the Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique (GIPSA-Lab) as a Postdoctoral Fellow. His main research is focused on hyperspectral image processing, including feature extraction techniques, spectral unmixing, super-resolution and Pansharpening. From 2014 he is Ingénieur de recherché at the Grenoble Institute of Technology. He is also a European Space Agency Category-1 Principal Investigator for Earth observation data. Dr. Licciardi serves as a Referee for several scientific journals such as the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, THE IEEE GEOSCIENCE AND REMOTE SENSING LETTERS AND IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

Guillaume Tochon (S'12) received the M.Sc degree in Electrical Engineering from the Grenoble Institute of Technology (Grenoble-INP), France, in 2012, and the PhD degree in signal and image processing from the University of Grenoble Alpes, France, in 2015. He is currently a postdoctoral researcher at the Images-Signal department in the GIPSA-lab, Grenoble, France. His research activities focus on mathematical morphology and data fusion, with applications in remote sensing. He also serves as a reviewer for the IEEE Transactions on Geoscience and Remote Sensing, the IEEE Journal of Selected Topics in Earth Observations and Remote Sensing and the IEEE Transactions on Image Processing journals. He is a member of the IEEE Geoscience and Remote Sensing society and the IEEE Signal Processing society.

Mauro Dalla Mura (S'08 – M'11) received the laurea (B.E.) and laurea specialistica (M.E.) degrees in Telecommunication Engineering from the University of Trento, Italy, in 2005 and 2007, respectively. He obtained in 2011 a joint Ph.D. degree in Information and Communication Technologies (Telecommunications Area) from the University of Trento, Italy and in Electrical and Computer Engineering from the University of Iceland, Iceland. In 2011 he was a Research fellow at Fondazione Bruno Kessler, Trento, Italy, conducting research on computer vision. He is currently an Assistant Professor at Grenoble Institute of Technology (Grenoble INP), France. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His main research activities are in the fields of remote sensing, image processing and pattern recognition. In particular, his interests include mathematical morphology, classification and multivariate data analysis. Dr. Dalla Mura was the recipient of the IEEE GRSS Second Prize in the Student Paper Competition of the 2011 IEEE IGARSS 2011 and co-recipient of the Best Paper Award of the International Journal of Image and Data Fusion for the year 2012-2013 and the Symposium Paper Award for IEEE IGARSS 2014. He is a Reviewer of IEEE Transactions on Geoscience and Remote Sensing, IEEE Geoscience and Remote Sensing Letters, IEEE Journal of Selected Topics in Earth Observations and Remote Sensing, IEEE Journal of Selected Topics in Signal Processing, Pattern Recognition Letters, ISPRS Journal of Photogrammetry and Remote Sensing, Photogrammetric Engineering and Remote Sensing (PE&RS). He is a member of the Geoscience and Remote Sensing Society (GRSS) and IEEE GRSS Data Fusion Technical Committee (DFTC) and Secretary of the IEEE GRSS French Chapter (2013-2016). He was a lecturer at the RSSS12 - Remote Sensing Summer School 2012 (organized by the IEEE GRSS), Munich, Germany.

Christian Jutten (AM'92-M'03-SM'06-F'08) received Ph.D. and Doctor ès 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 University Joseph Fourier of Grenoble, in 1989. For 35 years, his research interests have been machine learning and source separation, including theory (separability, source separation in nonlinear mixtures, sparsity, multimodality) and applications (brain and hyperspectral imaging, chemical sensor array, speech). He is author or coauthor of more than 90 papers in international journals, 4 books, 25 keynote plenary talks and about 200 communications in international conferences.

He has been visiting professor at Swiss Federal Polytechnic Institute (Lausanne, Switzerland, 1989), at Riken labs (Japan, 1996) and at Campinas University (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 for the French National Research Center (2003–2006). From May 2012 to September 2014, he was deputy director at the Institute for Information Sciences (INS2I) at French National Center of Research (CNRS) in charge of signal and image processing.

Christian Jutten was organizer or program chair of many international conferences, especially of the 1st International Conference on Blind Signal Separation and Independent Component Analysis in 1999. He has been a member of a few IEEE Technical Committees, and currently in " SP Theory and Methods" of the IEEE Signal Processing society. He received best paper awards of EURASIP (1992) and of IEEE GRSS (2012), and Medal Blondel (1997) from the French Electrical Engineering society for his contributions in source separation and independent component analysis. He is IEEE fellow (2008) and EURASIP fellow (2013). He is a Senior Member of the Institut Universitaire de France since 2008, with renewal in 2013 for 5 years. He is the recipient of a 2012 ERC Advanced Grant for a project on challenges in extraction and separation of sources (CHESS).

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 Savoie University, 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 was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and 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 2014-2015, he is 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).