Hyperspectral and Multispectral Image Fusion Under Spectrally Varying Spatial Blurs – Application to High Dimensional Infrared Astronomical Imaging

Claire Guilloteau, Thomas Oberlin, Olivier Berné, Nicolas Dobigeon

Hyperspectral and multispectral image fusion under spectrally varying spatial blurs – Application to high dimensional infrared astronomical imaging

Claire Guilloteau, Thomas Oberlin, Olivier Berné and Nicolas Dobigeon

***Abstract*—Hyperspectral imaging has become a significant source of valuable data for astronomers over the past decades. Current instrumental and observing time constraints allow direct acquisition of multispectral images, with high spatial but low spectral resolution, and hyperspectral images, with low spatial but high spectral resolution. To enhance scientific interpretation of the data, we propose a data fusion method which combines the benefits of each image to recover a high spatio-spectral resolution datacube. The proposed inverse problem accounts for the specificities of astronomical instruments, such as spectrally variant blurs. We provide a fast implementation by solving the problem in the frequency domain and in a low-dimensional subspace to efficiently handle the convolution operators as well as the high dimensionality of the data. We conduct experiments on a realistic synthetic dataset of simulated observation of the upcoming James Webb Space Telescope, and we show that our fusion algorithm outperforms state-of-the-art methods commonly used in remote sensing for Earth observation.**

***Index Terms*—data fusion, hyperspectral imaging, high dimensional imaging, infrared astronomy, super-resolution, deconvolution**

I. Introduction

THE idea of combining spectroscopy and imaging has become very popular in the two past decades, leading to a new sensing paradigm referred to as *hyperspectral* or *spectral* imaging. Hyperspectral images can be thought as a whole cube of data which provides a full description of the acquired scene or sample both in space and wavelength, thus being suitable for numerous chemical or physical analyses in various applicative domains. Hyperspectral imaging finds applications in many different fields, including remote sensing for Earth observation [1], [2] or planetology [3], material science [4], [5], [6], dermatology [7] and food quality monitoring [8]. In this work, we will focus on astronomy in the visible and near-infrared range. Sensing the universe in this spectral range at high spatial and spectral resolution is indeed of particular interest to study key mechanisms in astrophysics and cosmology. More specifically, this concerns for instance the combined sensing of the morphology or spectral signatures of protoplanetary disks, the interstellar medium or galaxies in the near or distant universe. For those purposes, numerous astronomical instruments have, in the past couple decades, adopted observing modes or designs allowing to acquire hyperspectral datasets. A full review of these instruments is out of the scope of this paper, but, for instance, this concerns the instruments aboard a number of space missions such as ESA's Infrared Space Observatory, NASA's Spitzer Space Telescope, or the ESA's Herschel Space Observatory, and the upcoming James Webb Space Telescope.

However, instrumental constraints usually do not enable a direct acquisition of data-cubes combining full spatial *and* spectral resolutions simultaneously (when this is the case, it is at the price of much longer integration times). A common alternative for astronomers consists in acquiring two images of the same scene with complementary information, namely an hyperspectral (HS) image with high spectral resolution and a multispectral (MS) image with high spatial resolution. The HS and MS data fusion aims at combining these complementary observations to reconstruct a full data-cube at high spectral and spatial resolutions. This virtually allows to combine the performances of the data-sets at the post-processing step, without any modification of observing modes, instrumental designs, or integration times. From the astronomical point of view, the resulting fused product is expected to provide meaningful insights about the scene of interest. Thanks to its high spatial and spectral resolution, it opens the door to a finer mapping of physical quantities, such as gas temperature and density, radiation field, metalicity, chemical components, etc.

Image fusion has been extensively studied in the literature of Earth observation [9], [10], [11]. This task is undertaken to provide high spatial resolution multiband images from the measurements provided by a large majority of the optical sensors dedicated to Earth observation. Indeed, most of the airborne or spaceborne optical remote sensing platforms embed at least two sensors with complementary spatial and spectral resolutions. The first methods addressed the so called pansharpening problems, which consists in fusing a MS or HS image with a panchromatic (PAN) image, i.e., a grayscale image with a single spectral band. These heuristic approaches [12], [13] consisted in injecting spatial details extracted from the high spatial resolution image into an interpolated version

C. Guilloteau and N. Dobigeon are with University of Toulouse, IRIT/INP-ENSEEIHT, CNRS, 2 rue Charles Camichel, BP 7122, 31071 Toulouse Cedex 7, France (e-mail: {Claire.Guilloteau, Nicolas.Dobigeon}@enseeiht.fr).

N. Dobigeon is also with Institut Universitaire de France (IUF).

Th. Oberlin is with ISAE-SUPAERO, University of Toulouse, France (e-mail: Thomas.Oberlin@isae-supaero.fr).

C. Guilloteau and O. Berné are with Institut de Recherche en Astrophysique et Planétologie (IRAP), University of Toulouse, France (e-mail: {Claire.Guilloteau,Olivier.Berne}@irap.omp.eu).

Part of this work has been supported by the ANR-3IA Artificial and Natural Intelligence Toulouse Institute (ANITI) under grant agreement ANITI ANR-19-PI3A-0004, the French Programme Physique et Chimie du Milieu Interstellaire (PCMI) funded by the Conseil National de la Recherche Scientifique (CNRS) and Centre National d'Études Spatiales (CNES) Appel à Proposition de Recherche 4828.

of the low spatial resolution image. Those methods, in addition to be fast and easy to implement, are likely to recover spatial details with high accuracy, but they often produce significant spectral deformations [12]. Another class of data fusion methods is based on spectral unmixing and matrix factorization paradigms. One of the first methods was proposed in [14] for fusing infrared astronomical data. According to low-rank assumption on the spectral information contained in the HS image, the latter is decomposed into two factors, representing source spectra and spatial coefficients, following a non-negative matrix factorization (NMF) [15]. The source spectra matrix is then combined with a high spatial resolution coefficient matrix extracted with a non-negative least square algorithm from the MS image. The same idea has been pursued by Yokoya *et al.* for remote sensing images [16]. The so-called coupled-NMF (CNMF) method performs NMF alternatively on the HS and MS images to extract a high resolution source spectra matrix and a high resolution spatial coefficient matrix. The two methods assume linear spectral degradation and spectrally invariant spatial blur for the observations, that can be either known or estimated beforehand. The main drawback of these spectral unmixing-based fusion methods lies on their slow convergence to a local minimum, making the solution highly dependent on the initialization. More recently, capitalizing on the prior knowledge regarding the observation instruments, the data fusion task has been formulated as an inverse problem derived from explicit forward models and complemented by appropriate spatial and/or spectral regularizations. More precisely, the forward models rely on a spectral degradation operator associated with the MS filters and a spectrally invariant spatial blurring induced by the HS sensor. Most of these methods assume a low-rank structure for the spectral information provided by the HS image. They mainly differ by the adopted spatial regularization designed to promote particular behaviors of the spatial content. For instance, a convex regularization as a form of vector total variation has been used in [17], promoting sparsity in the distribution of the gradient of the reconstructed image. Therefore, this fused image is expected to be spatially smooth, except for a small number of areas, coinciding with sharp edges. Instead of promoting a smooth content, the regularization introduced in [18], represents the target image as a sparse combination of elements of a dictionary composed of spatial patches and learned from the MS image. The resulting optimization problems are solved iteratively thanks to particular instances of the alternating direction method of multipliers [19]. More recently, the authors in [20], [21] show that such fusion inverse problems can be formulated as a Sylvester equation and solved analytically, significantly decreasing the computational complexity of the aforementioned iterative methods.

However, all these techniques are not suitable to tackle the fusion of high dimensional astronomical data. The first challenge is to handle the high dimensionality of the data, considerably larger than the usual dimension encountered in remote sensing. Indeed, a high spatio-spectral fused image in Earth remote sensing is composed of at most a few hundreds of spectral bands while spatio-spectral astronomical data are typically composed of up to several thousands, or even tens of thousands of spectral measurements. Moreover, the spatial resolution of space- or airborne Earth observations is mainly limited by atmosphere turbulence [22]. Nevertheless, the spatial resolution of spaceborne astronomical observations is limited by diffraction. This limit is wavelength dependent and can be estimated by the Rayleigh criterion [23]. It defines the angular resolution $\theta = 1.220\frac{\lambda}{D}$, where λ is the wavelength of the light and D the diameter of the aperture. In practice, this physical property means that the operators associated with spatial blurs should be considered as spectrally varying while restoring astronomical MS and HS images [24], [25]. This crucial issue significantly increases the complexity of the forward models and make the fusion methods previously discussed inoperative. Indeed, as mentioned above, the forward models commonly used for Earth observation data fusion rely on a spectrally invariant spatial blur to describe the HS observation and a subsampling operator combined with a spectral degradation operator for the MS observation. The main contributions reported in this work tackle both challenges: we design a fusion method and its fast implementation suitable for fusing large-scale astronomical data while taking into account the specificities of astronomical imaging, in particular the spectrally variant blur underlying the MS and HS observations.

The paper is organized as follows. Section II describes the observational forward models and introduces the fusion inverse problem. Then, Section III presents our main contribution: a fast implementation to solve the inverse problem. To this end, the optimization problem is rewritten in the frequency domain, while an appropriate vectorization step enables to formulate the spatial degradations in a low-dimensional subspace. In Section IV, the performance of the proposed method is assessed using a realistic simulated astrophysical dataset and compared qualitatively and quantitatively with state of the art methods. Section V finally concludes the paper.

II. Problem formulation

A. Forward models

This section derives the mathematical models associated with two instruments providing images of complementary spatial and spectral resolutions. Contrary to models associated to Bayer-like or other color filter arrays mostly encountered in digital cameras [26, Chap. 10], they are formulated under the assumption of independent dual measurements encoding the information separately. The first instrument is an optical imager which acquires a MS image of high spatial resolution denoted $\mathbf{Y}_{\mathrm{m}} \in \mathbb{R}^{l_{\mathrm{m}} \times p_{\mathrm{m}}}$, where l_{m} and p_{m} denote the numbers of spectral bands and pixels, respectively. The second instrument is a spectrometer which acquires a full HS datacube $\mathbf{Y}_{\mathrm{h}} \in \mathbb{R}^{l_{\mathrm{h}} \times p_{\mathrm{h}}}$ of lower spatial resolution, with $l_{\mathrm{m}} < l_{\mathrm{h}}$ and $p_{\mathrm{h}} < p_{\mathrm{m}}$. From these measurements, the objective of the fusion process is to recover a HS image of high spatial resolution denoted $\mathbf{X} \in \mathbb{R}^{l_{\mathrm{h}} \times p_{\mathrm{m}}}$, which has the same spatial resolution as the MS image and the same spectral resolution of the HS one. The responses of the two sensors are modeled by a series of linear transformations that describe successive spatial and spectral degradations of light emerging from the scene of interest. With the adopted ordering of the elements in the

matrix $\mathbf{X}$, spectral and spatial degradations will be represented as left and right operators, respectively. More precisely, we assume that the MS and HS images result from the following forward models

$$\mathbf{Y}_{\mathrm{m}} \approx \mathbf{L}_{\mathrm{m}}\mathcal{M}(\mathbf{X}) \tag{1}$$

$$\mathbf{Y}_{\mathrm{h}} \approx \mathbf{L}_{\mathrm{h}}\mathcal{H}(\mathbf{X})\mathbf{S} \tag{2}$$

where the symbol $\approx$ accounts for random noises and model mismodeling, and the other operators are detailed hereafter. First, $\mathbf{L}_{\mathrm{m}} \in \mathbb{R}^{l_{\mathrm{m}} \times l_{\mathrm{h}}}$ and $\mathbf{L}_{\mathrm{h}} \in \mathbb{R}^{l_{\mathrm{h}} \times l_{\mathrm{h}}}$ are spectral degradation operators, respectively associated with MS and HS images. The MS observation instrument integrates the spectral bands of the initial scene $\mathbf{X}$ over the spectral dimension to provide each MS band. The rows of the matrix $\mathbf{L}_{\mathrm{m}}$ in (1) are thus made of the transmission functions of the l_{m} corresponding filters [16], [17], [18]. On the other hand, the spectral information of the initial scene $\mathbf{X}$ is attenuated by the optical system of the HS instrument. Therefore, $\mathbf{L}_{\mathrm{h}}$ is a diagonal matrix made of the spectral transmission function of the instrument. It is worth noting that these two spectral degradations are spatially invariant, which allows them to be formulated as the matrix products in (1) and (2). Second, $\mathcal{M} : \mathbb{R}^{l_{\mathrm{h}} \times p_{\mathrm{m}}} \to \mathbb{R}^{l_{\mathrm{h}} \times p_{\mathrm{m}}}$ in (1) and $\mathcal{H} : \mathbb{R}^{l_{\mathrm{h}} \times p_{\mathrm{m}}} \to \mathbb{R}^{l_{\mathrm{h}} \times p_{\mathrm{m}}}$ in (2) are spatial degradation operators which model the blurs caused by the optical system of both instruments. In the context of astronomical imaging addressed in this work, we can reasonably assume the associated point spread functions (PSFs) to be space-invariant, but they strongly depend on the wavelength, following a Rayleigh criterion [23]. Therefore, $\mathcal{M}(\cdot)$ and $\mathcal{H}(\cdot)$ are 2D spatial convolution operators with spectrally variant blurring kernels specific to each instrument [24], [25]. Finally, the spatial resolution of the HS image is impaired by a subsampling operator $\mathbf{S} \in \mathbb{R}^{p_{\mathrm{m}} \times p_{\mathrm{h}}}$ with an integer decimation factor d such that $p_{\mathrm{h}} = \frac{p_{\mathrm{m}}}{d^2}$. In other words, right-multiplying by $\mathbf{S}$ amounts to keeping one pixel over d^2. In this work, we assume that all the operators are known.

B. Inverse problem

To recover $\mathbf{X}$ from the two noisy observations, we adopt the general framework of (variational) inverse problem, trying to fit the observations while adding regularization terms to promote prior knowledge on the sought solution. Similar approaches have been widely advocated to address the problem of multiband image fusion in the remote sensing and Earth observation literature. By denoting $(\cdot)^H$ the Hermitian transpose and $\|\cdot\|_{\mathrm{F}}^2 = \mathrm{Tr}\left((\cdot)(\cdot)^H\right)$ the Frobenius norm, this amounts to solving the generic problem

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \left(\frac{1}{2\sigma_{\mathrm{m}}^2}\|\mathbf{Y}_{\mathrm{m}} - \mathbf{L}_{\mathrm{m}}\mathcal{M}(\mathbf{X})\|_{\mathrm{F}}^2 + \frac{1}{2\sigma_{\mathrm{h}}^2}\|\mathbf{Y}_{\mathrm{h}} - \mathbf{L}_{\mathrm{h}}\mathcal{H}(\mathbf{X})\mathbf{S}\|_{\mathrm{F}}^2 + \varphi_{\mathrm{spec}}(\mathbf{X}) + \varphi_{\mathrm{spat}}(\mathbf{X}) \right) \tag{3}$$

where the two first terms are data fidelity terms related respectively to the MS and the HS images. Minimizing these data fidelity terms is equivalent to maximize the log-likelihood associated to a white Gaussian noise model in the data, i.e., the symbols $\approx$ in (1) and (2) stand for additive corruptions $\mathbf{N}_{\mathrm{m}}$ and $\mathbf{N}_{\mathrm{h}}$ assumed to be independent white Gaussian noise with variance σ_{m}^2 and σ_{h}^2, respectively. Although this hypothesis may be not realistic for astronomical images as they are known to be rather corrupted by a mixed Poisson-Gaussian noise [27], the least-square loss is chosen for the sake of computational efficiency. It is worth noting that the experimental results reported in Section IV will show that this simplifying assumption does not significantly impair the relevance of the proposed method.

Besides, the terms $\varphi_{\mathrm{spec}}(\cdot)$ and $\varphi_{\mathrm{spat}}(\cdot)$ in (3) stand for spectral and spatial regularizations, respectively. Regarding $\varphi_{\mathrm{spec}}(\cdot)$, HS image bands are known to be highly correlated. Thus the pixels of the full scene $\mathbf{X}$ can be reasonably assumed to live in a subspace whose dimension l_{sub} is much smaller than its spectral dimension l_{h}. This property can be formulated by imposing a low-rank structure on the scene $\mathbf{X}$ to be recovered, i.e., $\mathbf{X} = \mathbf{V}\mathbf{Z}$ where the columns of $\mathbf{V} \in \mathbb{R}^{l_{\mathrm{h}} \times l_{\mathrm{sub}}}$ (with $l_{\mathrm{sub}} \leq l_{\mathrm{h}}$) spans the signal subspace and $\mathbf{Z} \in \mathbb{R}^{l_{\mathrm{sub}} \times p_{\mathrm{m}}}$ gathers the corresponding representation coefficients. This decomposition implicitly imposes a spectral regularization, since the spectra of the fused image are assumed to be linear combinations of the reference spectra defining $\mathbf{V}$. The columns of $\mathbf{V}$ spanning the signal subspace can be fixed beforehand thanks to prior knowledge regarding the composition of the scene of interest, for instance by stacking l_{sub} spectral signatures characterizing the components of the scene. In absence of such a knowledge, they can be directly estimated from the HS measurements, e.g., by conducting a principal component analysis (PCA) or resorting to a method specifically dedicated to HS data such as the maximum noise fraction (MNT) transform [28] or the hyperspectral signal subspace identification by minimum error (HySime) method [29]. This strategy has been widely adopted in numerous works of the literature dedicated to hyperspectral image enhancement [30], [17], [20]. Another asset of this change of variable lies in a significant reduction of the complexity of the optimization problem since *i)* estimating the decomposition coefficients $\hat{\mathbf{Z}}$ is sufficient to recover the fused image $\hat{\mathbf{X}} = \mathbf{V}\hat{\mathbf{Z}}$ and *ii)* this decomposition allows the forward models to be rewritten in the subspace spanned by $\mathbf{V}$, which leads to a scalable algorithm (see Section III-B for details).

Concerning the spatial regularization term $\varphi_{\mathrm{spat}}(\cdot)$, it is based on the assumption that the sought image is *a priori* spatially smooth, in agreement with typical scenes encountered in astrophysical observations. We thus propose to minimize the energy of the spatial discrete gradient of the image, also known as Sobolev regularization [31]. This writes

$$\varphi_{\mathrm{spat}}(\mathbf{Z}) = \mu\|\mathbf{Z}\mathbf{D}\|_{\mathrm{F}}^2$$

where the matrix $\mathbf{D}$ stands for a 1st order 2-D finite difference operator and the regularization parameter $\mu \geq 0$ controls the strength of the regularization. Note that, provided that the matrix $\mathbf{V}$ is orthonormal (i.e., $\mathbf{V}^T\mathbf{V} = \mathbf{I}_{l_{\mathrm{sub}}}$), this regularization formulated in the signal subspace is equivalent to the one that would be formulated in the original image

domain. The problem now becomes

$$\hat{\mathbf{Z}} = \underset{\mathbf{Z}}{\operatorname{argmin}} \left(\frac{1}{2\sigma_{\mathrm{m}}^2} \|\mathbf{Y}_{\mathrm{m}} - \mathbf{L}_{\mathrm{m}}\mathcal{M}(\mathbf{V}\mathbf{Z})\|_{\mathrm{F}}^2 + \frac{1}{2\sigma_{\mathrm{h}}^2} \|\mathbf{Y}_{\mathrm{h}} - \mathbf{L}_{\mathrm{h}}\mathcal{H}(\mathbf{V}\mathbf{Z})\mathbf{S}\|_{\mathrm{F}}^2 + \mu \|\mathbf{Z}\mathbf{D}\|_{\mathrm{F}}^2 \right). \quad (4)$$

The next section presents an efficient algorithmic scheme designed to solve the minimization problem (4).

III. Fast implementation

Although quadratic, the problem stated in (4) cannot be easily solved by conventional methods such as fast gradient descent [32] or conjugate gradient [33] because of the spectrally variant blurs in $\mathcal{M}(\cdot)$ and $\mathcal{H}(\cdot)$. Indeed, resorting to such algorithms needs to evaluate the gradient at each iteration, requiring the application of operators $\mathcal{M}(\cdot)$ and $\mathcal{H}(\cdot)$ and their respective adjoints, i.e., applying a set of $4l_{\mathrm{h}}$ distinct 2D spatial convolutions. Storing and processing thousands of distinct PSFs would annihilate the benefit of the dimension reduction induced by the low-rank decomposition associated with the matrix $\mathbf{V}$. This section details our main contribution which consists of a fast implementation tailored to this fusion task under spectrally variant blurring. Firstly, we show that the considered problem can be fully formulated in the frequency domain to handle the heavy convolution operators $\mathcal{M}(\cdot)$ and $\mathcal{H}(\cdot)$. Yet sporadically adopted in a few previous works [20], [21], this strategy departs from most of the fusion techniques already proposed in the literature. Indeed, most of these techniques usually solve the problem entirely in the image domain [34], [35], [36] or operate multiple and expensive back and forth between the image domain and the frequency domain along the iterations of the algorithm [17], [18], [37]. Secondly, we combine these convolution operators and vectorize the whole problem expressed in the low-dimensional subspace spanned by the columns of $\mathbf{V}$. Up to our knowledge, no similar approach has been proposed yet. Although the computational gain expected from a naive vectorization may appear limited, we show that it offers two benefits: *i)* it brings out appealing quantities that are computed only once in a pre-processing step and *ii)* the matrix associated with the resulting linear system to be solved is shown to be highly sparse with a particular block structure, which opens the door to the use of dedicated efficient solvers. Thus rather counterintuitively, this vectorization trick allows one to significantly reduce the computational cost of the subsequent iterative minimization algorithm. To highlight the benefits associated with each of these two steps, namely the formulation in the Fourier domain and the vectorization, we provide and discuss the computational complexity of the proposed implementation at the end of this section.

A. Resolution in the frequency domain

It is widely admitted that computing convolutions in the frequency domain using fast Fourier transform (FFT) and its inverse (iFFT) [38] can be faster than directly convolving in the spatial domain. Here, we propose to reformulate the whole problem in the Fourier domain to benefit from this computational advantage. Indeed, every spatial degradation operator

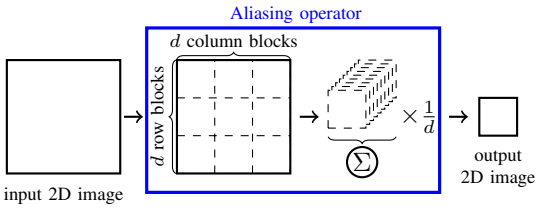

Fig. 1. Illustration of the aliasing operation $\dot{\mathbf{S}}$.

(convolution, subsampling and finite differences operators) can be expressed or approximated in the Fourier domain by simple operators, thus reducing the computation burden. First, under periodic boundary assumptions, the set of 2D spatial convolutions in $\mathcal{H}(\cdot)$ and $\mathcal{M}(\cdot)$ can be achieved by cyclic convolutions acting on $\mathbf{V}\mathbf{Z}$. We denote by $\odot$ the element-wise matrix multiplication and $\mathbf{F}$ the 2D-discrete Fourier transform (DFT) matrix ($\mathbf{F}\mathbf{F}^H = \mathbf{F}^H\mathbf{F} = \mathbf{I}$) such that $\dot{\mathbf{Z}} = \mathbf{Z}\mathbf{F}$. Thus, the convolution at a specific spectral band l can be rewritten

$$[\mathcal{M}(\mathbf{V}\mathbf{Z})]^l = \left(\dot{\mathbf{M}}^l \odot [\mathbf{V}\dot{\mathbf{Z}}]^l\right)\mathbf{F}^H$$

$$[\mathcal{H}(\mathbf{V}\mathbf{Z})]^l = \left(\dot{\mathbf{H}}^l \odot [\mathbf{V}\dot{\mathbf{Z}}]^l\right)\mathbf{F}^H$$

where $\dot{\mathbf{M}}^l$ and $\dot{\mathbf{H}}^l$ denote the 2D-DFTs of the lth PSFs related to, respectively, the multi- and the hyperspectral observation instrument [39]. The down-sampling operator $\mathbf{S}$ of factor d can be written in the Fourier domain as an aliasing operator $\dot{\mathbf{S}} = \mathbf{S}\mathbf{F}$ of factor d [40], which sums d^2 blocks of an input matrix, as illustrated in Fig. 1. Similarly to the downsampling operator, the aliasing operator acts independently on every spectral band. Each 2D spatial image with p_{m} pixels is partitioned into d^2 blocks and these blocks are summed up to produce a 2D spatial image with $p_{\mathrm{h}} = \frac{p_{\mathrm{m}}}{d^2}$ pixels.

Regarding the spatial regularization, the 1st order 2D finite differences operator $\mathbf{D}$ can be seen as a 2D convolution operator with kernels $\begin{pmatrix} 1 & -1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. This operator needs to be applied to the low-dimensional representation maps $\mathbf{Z}$, whose spectral dimension l_{sub} is much smaller than $\mathbf{V}\mathbf{Z}$. Thus, the computational gain reached by computing this regularization in the Fourier domain remains negligible. However, for practical reasons and to simplify the implementation, we decide to adopt this strategy. More precisely, again, under cyclic boundary conditions, this regularization term can be expressed in the Fourier domain as a term-wise multiplication [39] such that

$$\mathbf{Z}\mathbf{D} = \left(\dot{\mathbf{Z}} \odot \dot{\mathbf{D}}\right)\mathbf{F}^H.$$

Finally, following Parseval's identity, the problem (4) is fully rewritten in the Fourier domain

$$\hat{\dot{\mathbf{Z}}} = \underset{\dot{\mathbf{Z}}}{\operatorname{argmin}} \left(\frac{1}{2\sigma_{\mathrm{m}}^2} \left\|\dot{\mathbf{Y}}_{\mathrm{m}} - \mathbf{L}_{\mathrm{m}}((\mathbf{V}\dot{\mathbf{Z}}) \odot \dot{\mathbf{M}})\right\|_{\mathrm{F}}^2 + \frac{1}{2\sigma_{\mathrm{h}}^2} \left\|\dot{\mathbf{Y}}_{\mathrm{h}} - \mathbf{L}_{\mathrm{h}}((\mathbf{V}\dot{\mathbf{Z}}) \odot \dot{\mathbf{H}})\dot{\mathbf{S}}\right\|_{\mathrm{F}}^2 + \mu \left\|\dot{\mathbf{Z}} \odot \dot{\mathbf{D}}\right\|_{\mathrm{F}}^2 \right) \quad (5)$$

where $\dot{\mathbf{Y}}_{\mathrm{m}} = \mathbf{Y}_{\mathrm{m}}\mathbf{F}$ and $\dot{\mathbf{Y}}_{\mathrm{h}} = \mathbf{Y}_{\mathrm{h}}\mathbf{F}$. Finally the fused image can be obtained as $\hat{\mathbf{X}} = \mathbf{V}\hat{\dot{\mathbf{Z}}}\mathbf{F}^H$.

B. Vectorization

The second step consists in computing the sequence of operators in the subspace spanned by the columns of $\mathbf{V}$ instead of being applied to the full image $\mathbf{V}\dot{\mathbf{Z}}$. To do so, we introduce the lexicographically ordered counterparts $\dot{\mathbf{y}}_{\mathrm{m}}$, $\dot{\mathbf{y}}_{\mathrm{h}}$, $\underline{\mathbf{V}}$ and $\dot{\mathbf{z}}$ of $\dot{\mathbf{Y}}_{\mathrm{m}}$, $\dot{\mathbf{Y}}_{\mathrm{h}}$, $\mathbf{V}$ and $\dot{\mathbf{Z}}$, respectively, such that

$$\dot{\mathbf{y}}_{\mathrm{m}} = \left[\dot{\mathbf{Y}}_{\mathrm{m}}^{1}, \cdots, \dot{\mathbf{Y}}_{\mathrm{m}}^{l_{\mathrm{m}}}\right]^{T} \qquad \dot{\mathbf{y}}_{\mathrm{h}} = \left[\dot{\mathbf{Y}}_{\mathrm{h}}^{1}, \cdots, \dot{\mathbf{Y}}_{\mathrm{h}}^{l_{\mathrm{h}}}\right]^{T}$$
$$\underline{\mathbf{V}} = \mathbf{V} \otimes \mathbf{I}_{p_{\mathrm{m}}} \qquad \dot{\mathbf{z}} = \left[\dot{\mathbf{Z}}^{1}, \cdots, \dot{\mathbf{Z}}^{l_{\mathrm{sub}}}\right]^{T}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{I}_p$ is the $p \times p$ identity matrix. With these notations, the problem (5) is equivalent to

$$\widehat{\dot{\mathbf{z}}} = \underset{\dot{\mathbf{z}}}{\operatorname{argmin}} \left(\frac{1}{2\sigma_{\mathrm{m}}^2} \|\dot{\mathbf{y}}_{\mathrm{m}} - \underline{\mathbf{L}}_{\mathrm{m}} \underline{\dot{\mathbf{M}}\mathbf{V}} \dot{\mathbf{z}}\|_2^2 + \frac{1}{2\sigma_{\mathrm{h}}^2} \|\dot{\mathbf{y}}_{\mathrm{h}} - \underline{\dot{\mathbf{S}}\mathbf{L}}_{\mathrm{h}} \underline{\dot{\mathbf{H}}\mathbf{V}} \dot{\mathbf{z}}\|_2^2 + \mu \|\underline{\dot{\mathbf{D}}} \dot{\mathbf{z}}\|_2^2 \right)$$

where $\underline{\mathbf{L}}_{\mathrm{m}}$, $\underline{\mathbf{L}}_{\mathrm{h}}$, $\underline{\dot{\mathbf{M}}}$, $\underline{\dot{\mathbf{H}}}$, $\underline{\dot{\mathbf{S}}}$ and $\underline{\dot{\mathbf{D}}}$ are vectorized forms of $\mathbf{L}_{\mathrm{m}}$, $\mathbf{L}_{\mathrm{h}}$, $\dot{\mathbf{M}}$, $\dot{\mathbf{H}}$, $\dot{\mathbf{S}}$ and $\dot{\mathbf{D}}$, respectively, such that

$$\underline{\mathbf{L}}_{\mathrm{m}} \underline{\dot{\mathbf{M}}\mathbf{V}} \dot{\mathbf{z}} = \begin{pmatrix} \left[\mathbf{L}_{\mathrm{m}}((\mathbf{V}\dot{\mathbf{Z}}) \odot \dot{\mathbf{M}})\right]^{1} \\ \vdots \\ \left[\mathbf{L}_{\mathrm{m}}((\mathbf{V}\dot{\mathbf{Z}}) \odot \dot{\mathbf{M}})\right]^{l_{\mathrm{m}}} \end{pmatrix}$$

and

$$\underline{\dot{\mathbf{S}}\mathbf{L}}_{\mathrm{h}} \underline{\dot{\mathbf{H}}\mathbf{V}} \dot{\mathbf{z}} = \begin{pmatrix} \left[\mathbf{L}_{\mathrm{h}}((\mathbf{V}\dot{\mathbf{Z}}) \odot \dot{\mathbf{H}})\dot{\mathbf{S}}\right]^{1} \\ \vdots \\ \left[\mathbf{L}_{\mathrm{h}}((\mathbf{V}\dot{\mathbf{Z}}) \odot \dot{\mathbf{H}})\dot{\mathbf{S}}\right]^{l_{\mathrm{h}}} \end{pmatrix}.$$

The structures and expressions of all vectorized spatial and spectral operators are detailed in Appendix A. Finally, the fusion task boils down to solving the linear system

$$\mathbf{A}\dot{\mathbf{z}} = \mathbf{b} \tag{6}$$

where $\mathbf{A} \in \mathbb{R}^{l_{\mathrm{sub}} p_{\mathrm{m}} \times l_{\mathrm{sub}} p_{\mathrm{m}}}$ and $\mathbf{b} \in \mathbb{R}^{l_{\mathrm{sub}} p_{\mathrm{m}}}$ are defined by

$$\mathbf{A} = \frac{1}{\sigma_{\mathrm{m}}^2} \underline{\mathbf{V}}^H \underline{\dot{\mathbf{M}}}^H \underline{\mathbf{L}}_{\mathrm{m}}^H \underline{\mathbf{L}}_{\mathrm{m}} \underline{\dot{\mathbf{M}}\mathbf{V}} + \frac{1}{\sigma_{\mathrm{h}}^2} \underline{\mathbf{V}}^H \underline{\dot{\mathbf{H}}}^H \underline{\mathbf{L}}_{\mathrm{h}}^H \underline{\dot{\mathbf{S}}}^H \underline{\dot{\mathbf{S}}\mathbf{L}}_{\mathrm{h}} \underline{\dot{\mathbf{H}}\mathbf{V}} + \mu \underline{\dot{\mathbf{D}}}^H \underline{\dot{\mathbf{D}}}, \tag{7}$$

$$\mathbf{b} = -\frac{1}{\sigma_{\mathrm{m}}^2} \underline{\mathbf{V}}^H \underline{\dot{\mathbf{M}}}^H \underline{\mathbf{L}}_{\mathrm{m}}^H \dot{\mathbf{y}}_{\mathrm{m}} - \frac{1}{\sigma_{\mathrm{h}}^2} \underline{\mathbf{V}}^H \underline{\dot{\mathbf{H}}}^H \underline{\mathbf{L}}_{\mathrm{h}}^H \underline{\dot{\mathbf{S}}}^H \dot{\mathbf{y}}_{\mathrm{h}}. \tag{8}$$

Interestingly, as suggested by (6) and explicitly expressed by (7) and (8), the quantities $\mathbf{A}$ and $\mathbf{b}$ resort to all spatial and spectral operators. In particular, they combine the individual wavelength-dependent PSFs defining $\mathcal{H}(\cdot)$ and $\mathcal{M}(\cdot)$ to be jointly expressed in the low-dimensional subspace through the left-composition by the projection operator $\underline{\mathbf{V}}^H$. Moreover, the symmetric matrix $\mathbf{A}$ is sparse and composed of, at most, $d^2 l_{\mathrm{sub}}^2 p_{\mathrm{m}}$ non-zero entries, i.e., only a d^2/p_{m}-th proportion of the matrix coefficients is non-zero, arranged according to a very particular structure detailed in Appendix B. As a consequence, its high level of sparsity, combined with its block structure, allows the matrix $\mathbf{A}$ to be computed only once as a pre-processing step and cheaply stored in memory (see Appendix B for a detailed description of its computation). Finally, this matrix can be easily called out along the iterations of a gradient-based descent algorithm implemented to solve (6). It is also worth noting that this matrix only depends on the forward models defined by the observation instruments, the adopted spatial regularization and the matrix $\mathbf{V}$ spanning the signal subspace. Thus, once this subspace does not change, this matrix does not need to be recomputed to fuse multiple sets of MS and HS measurements.

C. Complexity analysis

This section discusses the complexity imposed by one iteration for three different gradient descent algorithms that solve the fusion problem. More precisely, we compare a naive implementation minimizing (4), the so-called frequency algorithm minimizing the problem (5) formulated in the Fourier domain and the proposed algorithm solving the vectorized formulation yielding the linear system (6). The respective complexities are expressed as functions of the spatial and spectral dimensions of the data to be fused, namely p_{m}, l_{h} and l_{m}, and the intrinsic dimension l_{sub} of the subspace. They are reported in Table I for the general case, i.e., without assuming any particular prevalence of one of these quantities over the others. However, for typical scenarios arising in the applicative context of astronomical imaging that will be considered in the experiments (see Section IV), we have $l_{\mathrm{sub}} \leq l_{\mathrm{m}} \leq \log p_{\mathrm{m}}$. In this context, the following findings can be drawn.

When considering a naive implementation, the heaviest computational burden to solve (4) directly results from evaluating the gradient of the corresponding quadratic cost function, which amounts to $\mathcal{O}(l_{\mathrm{h}} p_{\mathrm{m}} \log p_{\mathrm{m}})$ operations. Note that this implementation relies on cyclic convolutions operated in the Fourier domain but requires back and forth in the image domain by FFT and inverse FFT at each iteration. When the problem is fully formulated in the Fourier domain (see Section III-A), the cost of computing the gradient associated to (5) reduces to $\mathcal{O}(l_{\mathrm{h}} p_{\mathrm{m}} l_{\mathrm{m}})$. By vectorizing the problem (see Section III-B), the gradient is directly given by the matrix $\mathbf{A}$ in (7). Thus the core steps of the iterative algorithm solving (6) consist in matrix-vector products, which requires only $\mathcal{O}(p_{\mathrm{m}} l_{\mathrm{sub}}^2)$ operations thanks to the high level of sparsity of $\mathbf{A}$. Consequently, one iteration of this vectorized implementation is significantly less complex than the naive and Fourier domain-based resolutions.

Besides, while the naive implementation does not require any pre-processing step, the two alternative schemes proposed in Sections III-A and III-B rely on quantities computed beforehand. More precisely, to solve (5) in the frequency domain, FFT of the MS and HS images and PSFs are required, for a overall complexity of $\mathcal{O}(l_{\mathrm{h}} p_{\mathrm{m}} \log p_{\mathrm{m}})$. In addition, solving (6) requires to compute the matrix $\mathbf{A}$ in (7) and the vector $\mathbf{b}$ in (8). Specifically, the most heavy step is computing the two first terms in the right-hand side of (7), for a overall complexity of $\mathcal{O}(l_{\mathrm{h}} p_{\mathrm{m}} l_{\mathrm{sub}}^2)$. Therefore, the pre-processing step involved in the vectorized implementation is more time-consuming but this step is performed only once before solving

TABLE I
ASYMPTOTIC COMPLEXITY (AS $\mathcal{O}(\cdot)$): ONE ITERATION OF THE GRADIENT-BASED ALGORITHM AND PRE-PROCESSING.

Method	Iteration	Pre-processing
Naive	$l_{\mathrm{h}} p_{\mathrm{m}} \max\{l_{\mathrm{sub}}, l_{\mathrm{m}}, \log p_{\mathrm{m}}\}$	–
Frequency	$l_{\mathrm{h}} p_{\mathrm{m}} \max\{l_{\mathrm{sub}}, l_{\mathrm{m}}\}$	$l_{\mathrm{h}} p_{\mathrm{m}} \log p_{\mathrm{m}}$
Vectorized	$p_{\mathrm{m}} l_{\mathrm{sub}}^2$	$l_{\mathrm{h}} p_{\mathrm{m}} l_{\mathrm{sub}}^2$

the problem iteratively. Moreover, as already highlighted, this pre-processing is significantly lightened when fusing several sets of HS and MS measurements since only $\mathbf{b}$ needs to be updated, provided the spatial regularization and the signal subspace remain unchanged.

Beyond the computational complexity, given the high dimensionality of the problem, issues raised by handling the data and instrument models should be also discussed. In particular, loading the entire fused product and all spectrally variant PSFs is impossible in high dimension when using conventional computing resources. As a consequence, when solving the fusion problem with the naive strategy or in the frequency domain (see Section III-A), computing $(\mathbf{V}\dot{\mathbf{Z}}) \odot \dot{\mathbf{M}}$ and $(\mathbf{V}\dot{\mathbf{Z}}) \odot \dot{\mathbf{H}}$ at each iteration of the descent algorithm requires on-the-fly loading of each PSF, which increases the computational times significantly. Conversely, the vectorized implementation requires to load these PSFs only once during the pre-processing.

IV. EXPERIMENTAL RESULTS

This section assesses the performance of the proposed fusion method when applied to a simulated yet realistic astronomical dataset. This dataset is discussed in the next paragraph. The considered figures-of-merit, compared methods and quantitative and qualitative results are reported subsequently.

A. Simulated dataset

The simulated dataset considered in the experiments was specifically designed to assess multi- and hyperspectral data fusion in the particular context of high dimensional astronomical observations performed by the James Webb Space Telescope (JWST). The generation process is accurately described in [41] and more briefly recalled hereafter. This dataset is composed of a high spatial and high spectral resolution synthetic scene of a photodissociation region (PDR) located in the Orion Bar. This scene is accompanied with a pair of corresponding simulated MS and HS observations. The resolution of the synthetic scene matches the spectral resolution of the HS instrument and the spatial resolution of the MS sensor, and its field of view and spectral range corresponds to plausible real acquisitions that will be performed by the JWST.

The synthetic scene has been generated under a low-rank assumption such that its constitutive spectra are linear mixtures of 4 synthetic elementary spectra spatially distributed according to 4 maps representing the spatial abundances of each elementary spectrum over the scene. To simulate the expected spatial and spectral content of the Orion bar, four real images acquired by different telescopes are combined to build the spatial maps and the spectral signatures of the elementary components were chosen to be those likely present in this region (see [41] for more details). This simulated scene will be denoted $\mathbf{X}$ in the following and will represent the reference (i.e., ground-truth) data-cube we aim to recover by fusing the HS and MS measurements. It is composed of 90×900 pixels and 4974 spectral bands ranging from 1 to 2.35 μm.

The corresponding MS and HS observed images were simulated from this reference synthetic image following the forward models introduced in Section II-A, where the spatial and spectral degradation operators are those of the JWST instrumentation documentation[1]. The MS image $\mathbf{Y}_{\mathrm{m}}$ simulates the output of the near-infrared camera (NIRCam) imager and is composed of 90×900 pixels and 11 spectral bands. The HS image $\mathbf{Y}_{\mathrm{h}}$ consists of 30×300 pixels and 4974 spectral bands with the specificities of the integral field unit (IFU) of the near-infrared spectrograph (NIRSpec). The spatial subsampling factor d is thus set to $d = 3$. The spectral degradation operators $\mathbf{L}_{\mathrm{m}}$ and $\mathbf{L}_{\mathrm{h}}$ are the spectral responses of those two instruments as specified by the documentation. The 2-D spatial convolution operators $\mathcal{M}(\cdot)$ and $\mathcal{H}(\cdot)$ are each composed of 4974 PSFs whose full width at half-maximum (FWHM) is linearly varying with wavelength. Therefore, the widest PSF is 2.35 times larger than the thinnest. For multi- and hyperspectral observation instruments, they are of size 161×161 pixels and 145×145 pixels, respectively and, because of the specific shape of JWST mirrors, these PSFs are strongly anisotropic, as illustrated in Fig. 3. This figure emphasizes the crucial need of accounting for spectrally variant spatial convolution operators in the two forward models.

Finally, the simulated images $\mathbf{Y}_{\mathrm{m}}$ and $\mathbf{Y}_{\mathrm{h}}$ include a realistic Poisson-Gaussian mixed noise which is expected to corrupt astronomical data. They are first corrupted with a Poisson noise approximated by a multiplicative Gaussian noise of mean and variance the photon count in each pixel. The instrumental so-called readout noise is subsquently modeled by an additive spatially correlated Gaussian noise, with mean and covariance matrix depending on instruments and readout patterns, assumed to be known. It is worth noting that this mixed noise significantly departs from the simplifying assumption of Gaussian noise underlying the two data-fitting terms that define the minimization problem (3).

Red-green-blue (RGB) color compositions (left) and spectra (right) of the reference synthetic image (top), the simulated MS observed image (middle) and simulated HS observed image (bottom) are shown in Fig. 2. Each color in the composite images is associated to a specific emission line chemically related to a particular region of the PDR to highlight the various structures of the scene. Spectra in the right-hand side of the figure coincide with a pixel in the dark-blue region. Those illustrations show how the signal is degraded by the instruments. For the MS observations, the RGB composition

[1]Instrumental documentation available on STScI website: https://jwst-docs.stsci.edu/

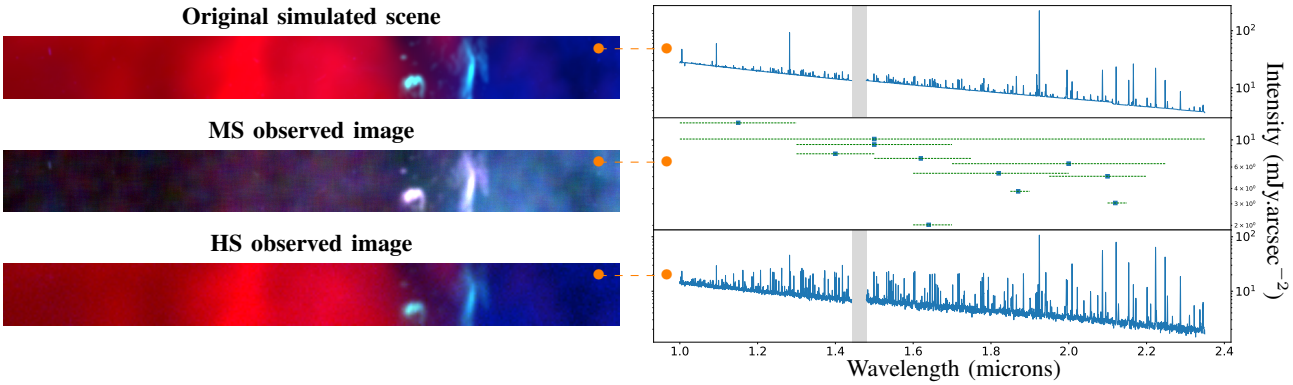

Fig. 2. Left: RGB compositions of the synthetic simulated scene (top), the NIRCam Imager MS image (middle) and the NIRSpec IFU HS image (bottom) [Red channel: H_2 emission line pic intensity at 2.122μm, Green channel: H recombination line pic intensity at 1.865μm, Blue channel: Fe^+ emission line pic intensity at 1.644μm]. Right: A spectrum from 1.0 to 2.35 microns related to a pixel of each image on their left. From top to bottom, the first two are original spectra from the synthetic scene with 4974 points, the following two are observed spectra from the multiband image provided by the NIRCam Imager forward model with 11 spectral points, the last two are calibrated observed spectra from the HS image provided by the NIRSpec IFU forward model with about 5000 spectral points.

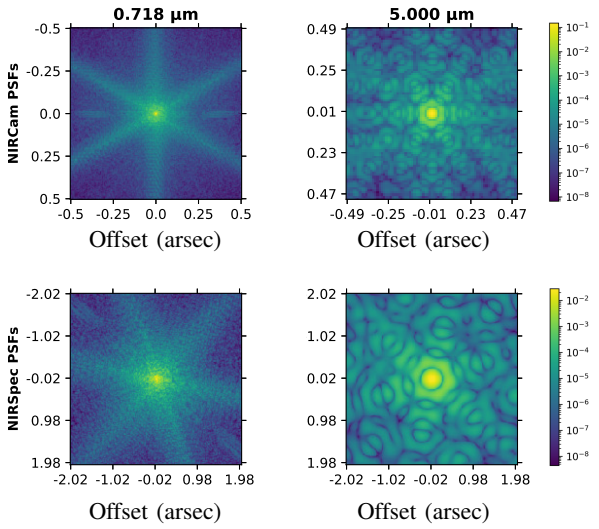

Fig. 3. PSFs of the NIRCam Imager (top) and NIRSpec IFU (bottom) calculated with *webbpsf* [42] for two particular wavelengths (logarithmic scale).

shows less contrast, due to the loss of spectral information induced by the filters. On the other hand, the hyperspectral data is clearly less spatially resolved, and the spectrum exhibits a lower signal-to-noise ratio (SNR).

B. Quality metrics

The performances of the compared data fusion algorithms are assessed according to three reconstruction quality measures. We propose to evaluate the spectral distortion between reconstructed and target spectra through the average spectral angle mapper (SAM), defined by

$$\text{aSAM}(\hat{\mathbf{X}}, \mathbf{X}) = \frac{1}{p_\text{m}} \sum_{p=1}^{p_\text{m}} \arccos\left(\frac{\langle \mathbf{X}_p, \hat{\mathbf{X}}_p \rangle}{\|\mathbf{X}_p\|_2 \|\hat{\mathbf{X}}_p\|_2} \right)$$

where $\hat{\mathbf{X}}_p$ is a reconstructed spectrum and $\mathbf{X}_p$ is the corresponding reference spectrum. The structural similarity (SSIM) index is then used to estimate the degradation of spatial structural information. The SSIM index is defined as

$$\text{SSIM}(\hat{\mathbf{X}}^l, \mathbf{X}^l) = \frac{\left(2\mu_{\hat{\mathbf{X}}^l}\mu_{\mathbf{X}^l} + C_1\right)\left(2\sigma_{\hat{\mathbf{X}}^1\mathbf{X}^l} + C_2\right)}{\left(\mu^2_{\hat{\mathbf{X}}^l} + \mu^2_{\mathbf{X}^l} + C_1\right)\left(\sigma^2_{\hat{\mathbf{X}}^1} + \sigma^2_{\mathbf{X}^1} + C_2\right)}$$

where $\hat{\mathbf{X}}^l$ is a lth reconstructed spectral band, $\mathbf{X}^l$ is the corresponding reference spectral band and $\mu_{\hat{\mathbf{X}}^l}$, $\mu_{\mathbf{X}^l}$, $\sigma^2_{\hat{\mathbf{X}}^1}$, $\sigma^2_{\mathbf{X}^1}$, $\sigma_{\hat{\mathbf{X}}^1\mathbf{X}^l}$ are empirical statistics defined in [43] and $C_j \propto L^2$ $(j = 1, 2)$ is the dynamic range of $\mathbf{X}^l$. In this paper, we rather consider the average complementary SSIM (acSSIM) across all bands defined by

$$\text{acSSIM}(\hat{\mathbf{X}}, \mathbf{X}) = 1 - \frac{1}{l_\text{m}} \sum_{l=1}^{l_\text{m}} \text{SSIM}(\hat{\mathbf{X}}^l, \mathbf{X}^l).$$

Finally, the overall peak SNR (PSNR) measures the overall reconstruction quality in the least-square sense:

$$\text{PSNR}(\hat{\mathbf{X}}, \mathbf{X}) = 10 \log_{10}\left(\frac{\max(\mathbf{X})}{\|\mathbf{X} - \hat{\mathbf{X}}\|^2_\text{F}} \right)$$

where $\hat{\mathbf{X}}$ is the reconstructed image and $\mathbf{X}$ is the reference. Note that a good performance is achieved when both the aSAM and acSSIM are low while the PSNR is large. All these quantities have been averaged over 20 Monte-Carlo runs.

C. Compared methods

We first consider a naive super-resolution method relying on a low-rank assumption, referred to afterwards as the baseline method. This approach consists in spatially upsampling the projection of the HS image onto the subspace spanned by the columns of $\mathbf{V}$ with a bi-cubic spline interpolation to reach the spatial resolution of the MS image.

We also compare our fusion algorithm, designated as "Proposed", with two methods widely known for fusing MS and HS or MS and PAN remote sensing data: the Brovey method

[44] and the robust fast fusion using a Sylvester equation (R-FUSE) [21]. The first one is a component substitution approach originally designed to fuse MS and PAN images. It interpolates the projection of the HS image over the spectral subspace to the spatial resolution of the MS image and injects extracted details from the MS image. It only requires the prior knowledge of the spectral blur operator $\mathbf{L}_{\mathrm{m}}$. The second method formulates the fusion task as an inverse problem derived from forward models of observation instruments complemented with a Gaussian prior. This problem uses a spectral degradation operator $\mathbf{L}_{\mathrm{m}}$ related to the multispectral instrument and a spectrally invariant PSF related to the hyperspectral instrument. In these experiments, this unique PSF is chosen as the PSF corresponding to the mean-energy wavelength. The problem is written as a Sylvester equation and solved analytically, substantially decreasing the computational complexity.

Finally to evaluate the relevance of the fusion task, we also compare our fusion algorithm with its two non-symmetric versions, where one of the data-fit term is removed. The first version, called MS-only, solves the following spectral deconvolution problem

$$\hat{\mathbf{Z}} = \underset{\mathbf{Z}}{\operatorname{argmin}} \left(\frac{1}{2\sigma_{\mathrm{m}}^2} \|\mathbf{Y}_{\mathrm{m}} - \mathbf{L}_{\mathrm{m}} \mathcal{M}(\mathbf{VZ})\|_{\mathrm{F}}^2 + \mu_{\mathrm{m}} \|\mathbf{ZD}\|_{\mathrm{F}}^2 \right)$$

where only the data fitting term related to the MS image is considered. Similarly, the second version, called HS-only and similar to [45], solves the following HS super-resolution problem including only the data fitting term related to the HS image

$$\hat{\mathbf{Z}} = \underset{\mathbf{Z}}{\operatorname{argmin}} \left(\frac{1}{2\sigma_{\mathrm{h}}^2} \|\mathbf{Y}_{\mathrm{h}} - \mathbf{L}_{\mathrm{h}} \mathcal{H}(\mathbf{VZ})\mathbf{S}\|_{\mathrm{F}}^2 + \mu_{\mathrm{h}} \|\mathbf{ZD}\|_{\mathrm{F}}^2 \right).$$

All the aforementioned methods require a subspace identification to find the basis matrix $\mathbf{V}$. This step is performed by PCA conducted on the HS image, as it is expected to contain all the relevant spectral information. These methods also require an hyperparameter setting. In this paper, the hyperparameter is set such that it leads to the highest PSNR value and the lowest aSAM and acSSIM.

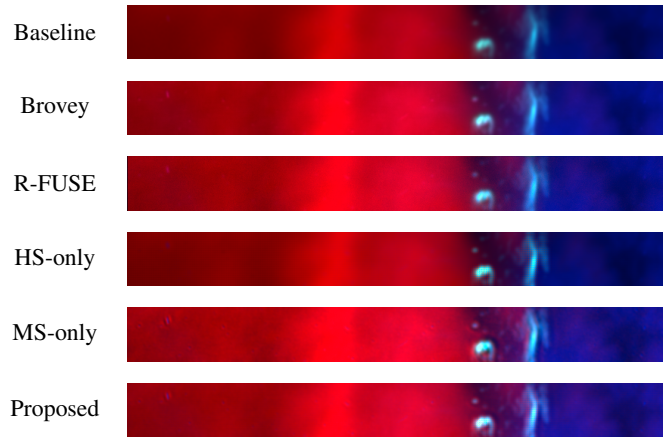

Fig. 4. From top to bottom: RGB compositions of fused images reconstructed by the baseline, Brovey, R-FUSE, HS-only, MS-only and proposed method. The color composition is the same as for Fig. 2 (left).

D. Results

The fusion results obtained by the six compared methods are depicted in Fig. 4 as RGB images using the same color composition as in Fig. 2. Zooms on sharp structures in the scene are shown in Fig. 5. Qualitatively, the reconstruction appears to be excellent. Denoising seems to be efficient for most methods, with a slightly noisier fused product obtained with R-FUSE. The baseline and HS-only methods are not able to restore the energy in the signal while the MS-only and proposed methods appear to better recover even very high intensities, especially on sharp edges, as shown in Fig. 5. The gain in spectral and spatial resolution of reconstructed images of the proposed algorithm with respect to MS and HS images respectively is clearly noticeable. The contrast between color components in the MS observed image is restored, as well as spatial details blurred in the HS observed image.

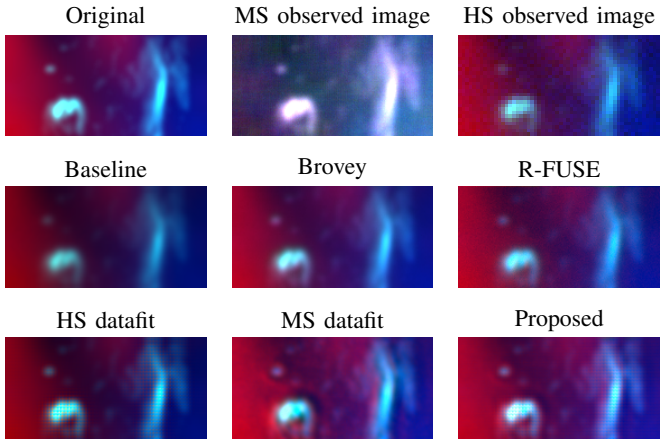

Fig. 5. Zooms on strong structures excerpt from Fig. 2 (simulated, MS observed and HS observed images) and from Fig. 4 (fused images by the compared methods).

To better assess method performances, quantitative results are reported in Table II. The two best results for each measure are highlighted in bold. As expected, the HS-only method shows a very low aSAM, i.e., an excellent spectral reconstruction but a poor spatial reconstruction with the second worst cSSIM index. On the other side, the MS-only method provides the best spatial reconstruction but the worst spectral reconstruction. Our method provides, as a trade-off between HS-only and MS-only, the second best spatial and spectral reconstructions. The best overall PSNR values are reached by MS-only, our proposed method and HS-only, improving the baseline performance up to 8dB. State-of-the-art methods give similar quantitative results, with a slightly better PSNR value for the Brovey method. Pre-processing time aside, all compared methods in their optimal implementation are very fast and perform data fusion in less than 30 seconds. To emphasize the interest of the formulation in the Fourier domain and the vectorization underlying the proposed algorithm, this table also reports the computational times required by depreciated counterparts of the proposed algorithm, namely the naive implementation and the frequency-based implementation (i.e., without vectorization). Note again that the three algorithms exactly solve the same minimization problem and thus are

expected to provide the same fusion performance. These times show the respective gains reached when conducting each of the two steps described in Section III. In particular, one can observe that the proposed vectorized resolution (including the pre-processing step) saves up to 8 and 16 times of the computational time when compared to the frequency-based and the naive implementations, respectively. It is also worth noting that the pre-processing step required by the proposed method after vectorization consists in pre-computing quantities that do not depend on the images to be fused. Thus, if the instrumental specifications defining the forward model and the signal subspace defined by $\mathbf{V}$ do not change, this pre-processing step needs to be conducted only once to fuse multiple data sets.

TABLE II
PERFORMANCE OF FUSION METHODS: ASAM (RAD), ACSSIM, PSNR (DB), AND TIME (PRE-PROCESSING + FUSION, SECONDS).

Methods		aSAM	acSSIM	PSNR	Time
Baseline		0.0296	0.0428	66.88	/
Brovey		0.0304	0.0040	71.04	17
R-FUSE		0.0360	0.0036	69.68	26
HS-only		**0.0118**	0.0239	72.90	1600 + 20
MS-only		0.0389	**0.0018**	**75.00**	600 + 15
Proposed	Naive				$\approx 36 \times 10^3$
	Frequency	**0.0247**	**0.0029**	**74.90**	$\approx 18 \times 10^3$
	Vectorized				2200 + 20

Fig. 6 presents SAM errors maps. These spectral errors have been calculated between reference and reconstructed spectra without averaging over the pixels. This figure highlights that, for each method, spectra located around a sharp structure of the scene (i.e., characterized by a high gradient region) show bad reconstructions (yellow pixels). This bad reconstruction is even worse for baseline, Brovey, R-FUSE and MS-only methods while HS-only and the proposed method provide the lowest SAM maxima. For most methods, this can be explained by the adopted regularizations, which promote spatially smooth content and therefore distribute the flux over neighboring pixels, leading to higher SAM values. On the contrary, in spatially smooth regions, all the methods present a very low spectral error. Fig. 7 represents cSSIM errors as function of the wavelength, i.e., without averaging over the spectral bands. Baseline and HS-only methods show very large spatial errors whereas MS-only and the proposed method provide the best cSSIM values. In between, Brovey and R-FUSE present intermediate and slightly increasing with wavelength cSSIM values. This may be explained by the fact that R-FUSE exploits a unique PSF, i.e., neglecting the spectrally spatial blur affecting the data. Indeed, larger wavelength bands are blurrier than short wavelength bands and therefore spatially worse reconstructed with an inappropriate model. The reference spectrum displayed in the bottom of the graph emphasizes that variations in cSSIM w.r.t. wavelength are correlated with high intensity emission lines in the scene. This is also likely due to the regularization term which tends to favor smooth images especially for high intensity spectral bands.

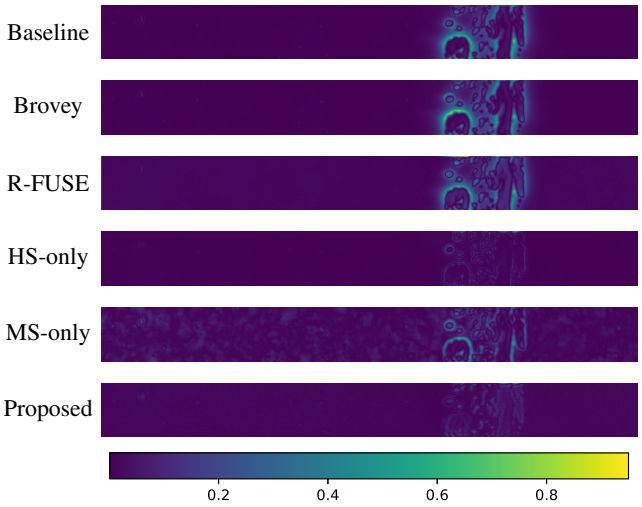

Fig. 6. Spatial maps of the SAM obtained by, from top to bottom, the baseline, Brovey, R-FUSE, HS-only, MS-only and proposed method. The smaller SAM, the better the reconstruction.

Figs. 8 and 9 show cumulative histograms of SAM and cSSIM errors respectively. According to Fig. 8, R-FUSE appears to provide a high systematic error and a large number of pixels with a SAM value larger than 10^{-1}rad. On the contrary MS-only and the proposed method also show a high systematic error but a small number of pixels with a large SAM value. On the other hand, the baseline and Brovey present a low systematic error but a large number of pixels with a large SAM value. HS-only shows the best cumulative histogram, with a low systematic error and a small number of pixels with a large SAM value. However in Fig. 9, this method, as well as the baseline, provide a very high (larger than 10^{-2}) cSSIM value for all spectral bands, while all the other methods show a much lower systematic error. Brovey and MS-only seem to have a very low number of spectral bands with a large cSSIM, but Brovey shows a much larger

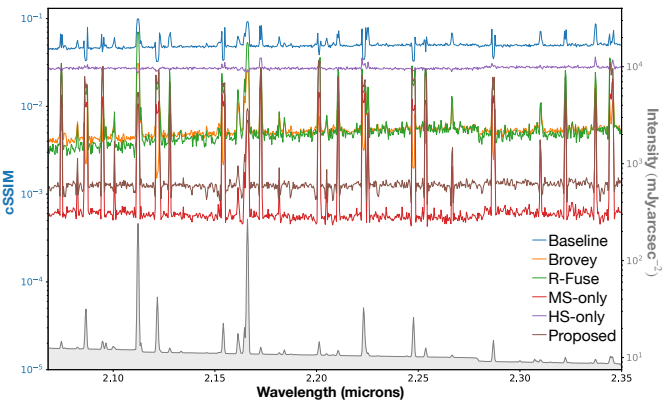

Fig. 7. Color lines: cSSIM as a function of the wavelength obtained by the compared methods. The smaller cSSIM, the better the reconstruction. Gray line and shaded area: a spectrum located in the reference scene around a sharp structure.

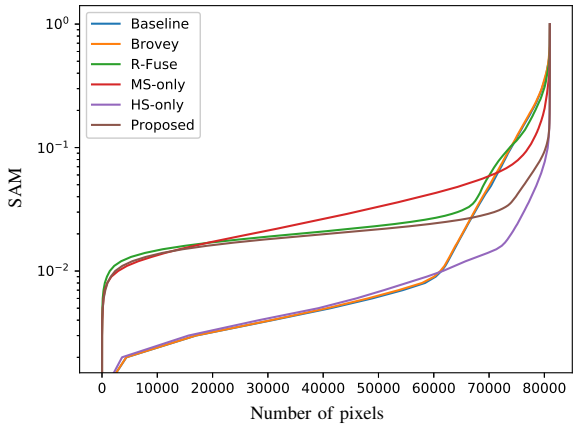

Fig. 8. SAM cumulative histograms obtained by the compared methods.

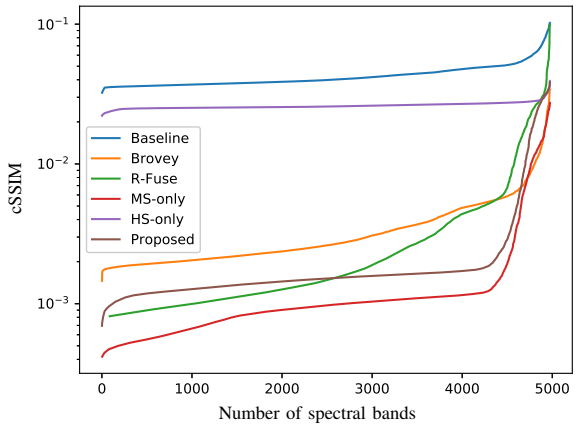

Fig. 9. cSSIM cumulative histograms obtained by the compared methods.

systematic error. R-FUSE and the proposed method present an intermediate cSSIM cumulative histogram, with a larger number of spectral bands with a high cSSIM value for the R-FUSE method. Considering these two figures, our proposed method emerges once more as a trade-off between good spatial and spectral reconstructions.

E. Selecting the regularization parameter

To choose an appropriate value for the regularization parameter μ in (4), we evaluated performances of the proposed fusion algorithm by monitoring the obtained aSAM and PSNR as functions of μ. Results are displayed in Fig. 10. In the simulations, we selected $\mu = 2.10^{-5}$ as a trade-off between the values providing the best PSNR and aSAM. We see that, for a wide range of μ values (light green, typically between 5.10^{-6} and $2.5.10^{-4}$), the proposed algorithm still outperforms state-of-the-art algorithms.

In a real-world scenario, i.e., when no ground truth is available and thus quantitative performance measures cannot be computed, we propose to adjust the regularization parameter μ automatically thanks to a dichotomous approach. More

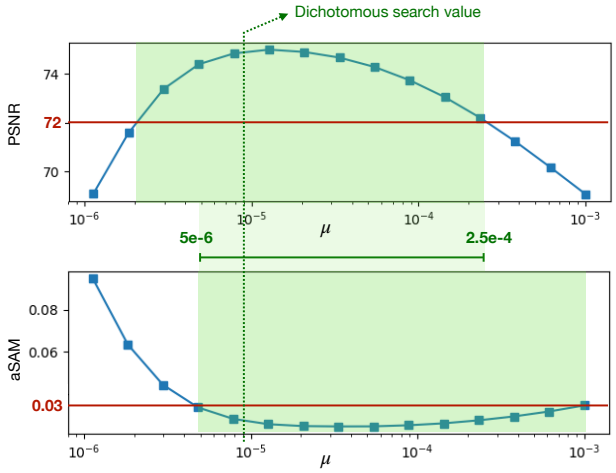

Fig. 10. Performance (in terms of PSNR and aSAM) of the proposed fusion algorithm as a function of the regularization parameter μ. Shaded green areas indicate the ranges of values for which the proposed algorithm outperforms state-of-the-art methods. The value of the parameter obtained by the proposed dichotomous search is highlighted with a vertical dotted line.

precisely, the optimal value of the parameter is assumed to provide a fused product $\hat{\mathbf{X}}$ such that the residuals defined by the forward models are of magnitude of the noise levels, i.e.,

$$\|\mathbf{Y}_{\mathrm{m}} - \mathbf{L}_{\mathrm{m}}\mathcal{M}(\hat{\mathbf{X}})\|_{\mathrm{F}}^2 \approx \sigma_{\mathrm{m}}^2 \quad (9)$$

$$\|\mathbf{Y}_{\mathrm{h}} - \mathbf{L}_{\mathrm{h}}\mathcal{H}(\hat{\mathbf{X}})\mathbf{S}\|_{\mathrm{F}}^2 \approx \sigma_{\mathrm{h}}^2. \quad (10)$$

Therefore, if the residuals are higher (resp. lower) than the noise levels, we increase (resp. decrease) the value of μ. As illustrated in Fig 10, the final value obtained by this iterative procedure is shown to belong to the range of acceptable values.

F. Robustness with respect to model mismatch

All experiments reported above have been conducted under a perfect knowledge of the forward models (1) and (2), i.e., an ideal design of the spectral filters $\mathbf{L}_{\mathrm{m}}$ and $\mathbf{L}_{\mathrm{h}}$ and an ideal calibration of the PSFs defining $\mathcal{M}(\cdot)$ and $\mathcal{H}(\cdot)$. Indeed, the performance of the proposed fusion algorithm has been assessed on pairs of MS and HS images generated following the simulation framework detailed in Section IV-A. Hence, both the data generation and fusion procedure rely on exactly the same forward models, a testing strategy usually referred to as "*inverse crime*" [46]. To depart from this questionable experimental protocol, new sets of MS and HS images are generated after perturbing the nominal spatial and spectral responses of the forward models (1) and (2) while, at the same time, these nominal responses are still used in the forward models underlying the fusion process. This additional experiment detailed below will allow ones to evaluate the impact of a forward model mismatch on the expected fusion performance.

More precisely, in this new experiment, the nominal spectral filters $\mathbf{L}_{\mathrm{m}}$ and $\mathbf{L}_{\mathrm{h}}$ of models (1) have been corrupted by an additive zero-mean Gaussian noise. The standard deviation of this noise has been adjusted such that the corresponding SNR is equal to 50dB. This corruption is assumed to be of low

TABLE III
PERFORMANCE OF THE PROPOSED FUSION METHOD WITH RESPECT TO SIMULATED TELESCOPE JITTER.

Jitter	aSAM	acSSIM	PSNR
$\times$ 1.00	0.0250	0.0029	74.90
$\times$ 1.20	0.0250	0.0029	74.84
$\times$ 1.95	0.0250	0.0029	74.99
$\times$ 3.15	0.0250	0.0029	74.97
$\times$ 5.15	0.0249	0.0029	74.93
$\times$ 8.35	0.0250	0.0029	74.62

energy since the design of the spectral filters embedded on spaceborne astronomical telescopes does not depend on the observed scene, is known with a high accuracy and does not vary with time. Conversely, the main source of errors on the calibration of the PSFs $\mathcal{M}(\cdot)$ and $\mathcal{H}(\cdot)$ is possible telescope instabilities, which are likely to occur for long exposure times. To mimic this telescope jitter, the nominal PSFs have been blurred with a Gaussian kernel following the strategy used in *webbpsf* [42]. The standard deviations of these kernels have been chosen to be from 1 to more than 8 times the nominal jitter value wich is an angle of 7.0×10^{-3}arcsec. (2.0×10^{-6} degrees) provided by *webbpsf*.

The new resulting data set has been fused by the proposed method when assuming nominal spatial and spectral filters. Table III reports the performance of the fusion method for noisy spectral responses and increasing levels of telescope jitter. Note that a jitter level of $\times 1$ corresponds to nominal spatial responses and, thus, this case reflects a sole spectral model mismatch. Spectral (aSAM) and spatial (acSSIM) reconstructions remain stable and of equivalent accuracy to performance reported in Table II without considering model mismatch. Although a slow decay of the overall reconstruction (PSNR) performance is observed when the jitter increases, the proposed fusion method still outperforms state-of-the-art methods.

As for the experiment reported in Section IV-D, these performances have been obtained for an optimal value of the regularization parameter μ, i.e., providing the highest PSNR. The automatic procedure for adjusting this hyperparameter (see Section IV-E) has been also performed and has shown to provide consistent results (not reported here for brevity). This can be explained by the impact of the model mismatch on the achieved reconstruction performance as a function of this parameter. To illustrate, similarly to the results presented in Fig. 10, Fig. 11 plots the PSNR as a function of μ for various levels of telescope jitter and a constant degradation of the spectral response. These curves show that a model mismatch only shifts vertically the achieved performance. As a consequence, in case of model mismatch, the proposed strategy to automatically adjust the regularization parameter is able to provide consistent results, without any dramatic impact on the expected fusion performance.

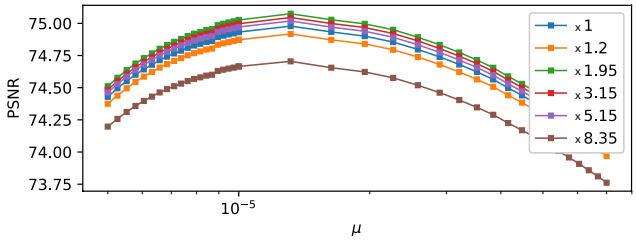

Fig. 11. Performance (in term of PSNR) of the proposed fusion algorithm as a function of the regularization parameter μ for various levels of telescope jitter and noisy spectral responses.

V. DISCUSSION AND CONCLUSION

In this paper, we proposed a novel hyperspectral and multispectral image fusion method when the observed images were affected by spectrally variant blurs. To computationally handle this particularity, we designed a fast algorithm to minimize the objective function associated with the fusion problem. Operating in the Fourier domain, this algorithm exploited the frequency properties of cyclic convolution operators and capitalized on a low-rank decomposition of the fused image. This implicit spectral regularization allowed the problem to be solved in a subspace of significantly lower dimension. These two computational advantages made the proposed algorithm able to handle large data sets since it solved the fusion problem with reasonable processing times. The relevance of the proposed method was evaluated in the specific context of astronomical imaging. We applied this method to a realistic simulated scene of the Orion Bar and compared the results with fused products obtained with state-of-the-art methods and non-symmetric versions of our approach. We showed that the proposed method appeared as an excellent trade-off with the best spectral, spatial *and* overall reconstruction results.

Improvements of the fusion method are however required. Further work will be dedicated to design a tailored regularization term, that could be more suitable than the currently chosen one to our kind of data. In a wider perspective, we also would like to include a realistic noise model in our fusion method.

APPENDIX A
VECTORIZED OPERATORS

To handle the vectorized counterpart $\dot{\mathbf{z}}$ of the DFT of the representation coefficients, the subspace basis matrix $\mathbf{V}$ shoud be rewritten $\underline{\mathbf{V}} = \mathbf{V} \otimes \mathbf{I}_{p_\mathrm{m} \times p_\mathrm{m}}$ such that

$$\underline{\mathbf{V}}\dot{\mathbf{z}} = \begin{pmatrix} [\mathbf{V}\dot{\mathbf{Z}}]^1 \\ \vdots \\ [\mathbf{V}\dot{\mathbf{Z}}]^{l_\mathrm{h}} \end{pmatrix}.$$

Similarly, within this vectorized formulation, the spectral degradation operators $\mathbf{L}_\mathrm{m}$ and $\mathbf{L}_\mathrm{h}$ should be rewritten as

$$\underline{\mathbf{L}_\mathrm{m}} = \mathbf{L}_\mathrm{m} \otimes \mathbf{I}_{p_\mathrm{m}}$$
$$\underline{\mathbf{L}_\mathrm{h}} = \mathbf{L}_\mathrm{h} \otimes \mathbf{I}_{p_\mathrm{h}}.$$

Corresponding convolution operators $\underline{\dot{\mathbf{M}}}$ and $\underline{\dot{\mathbf{H}}}$ are two block-diagonal matrices

$$\underline{\dot{\mathbf{M}}} = \mathrm{diag}\left\{\dot{\mathbf{M}}^1, \cdots, \dot{\mathbf{M}}^{l_\mathrm{h}}\right\}$$
$$\underline{\dot{\mathbf{H}}} = \mathrm{diag}\left\{\dot{\mathbf{H}}^1, \cdots, \dot{\mathbf{H}}^{l_\mathrm{h}}\right\}$$

defined by the DFTs of the MS and HS PSFs along the spectral bands. Finally, the spatial operators $\underline{\dot{\mathbf{S}}}$ and $\underline{\dot{\mathbf{D}}}$ are written as

$$\underline{\dot{\mathbf{S}}} = \mathbf{I}_{l_\mathrm{h}} \otimes \dot{\mathbf{S}}^H$$
$$\underline{\dot{\mathbf{D}}} = \mathbf{I}_{l_\mathrm{sub}} \otimes \dot{\mathbf{D}}^H.$$

APPENDIX B
STRUCTURE AND EFFICIENT COMPUTATION OF THE LINEAR SYSTEM MATRIX $\mathbf{A}$

Capitalizing on the vectorized formulation of the objective function (5), the matrix $\mathbf{A}$ defining the linear system to solve exhibits a particular structure. More precisely, as stated by (7), $\mathbf{A}$ can be written as a weighted sum of 3 matrices denoted here as $\mathbf{A}_\mathrm{m}$, $\mathbf{A}_\mathrm{h}$ and $\mathbf{A}_\mathrm{r}$ whose computations are discussed in what follows. First, the matrix $\mathbf{A}_\mathrm{m} \triangleq \underline{\mathbf{V}}^H \underline{\dot{\mathbf{M}}}^H \underline{\mathbf{L}}_\mathrm{m}^H \underline{\mathbf{L}}_\mathrm{m} \underline{\dot{\mathbf{M}}\mathbf{V}}$ associated with the MS forward model can be decomposed into $l_\mathrm{sub} \times l_\mathrm{sub}$ elementary blocks such that

$$\mathbf{A}_\mathrm{m} = \begin{bmatrix} [\mathbf{A}_\mathrm{m}]_1^1 & \dots & [\mathbf{A}_\mathrm{m}]_{l_\mathrm{sub}}^1 \\ \vdots & \ddots & \vdots \\ [\mathbf{A}_\mathrm{m}]_1^{l_\mathrm{sub}} & \dots & [\mathbf{A}_\mathrm{m}]_{l_\mathrm{sub}}^{l_\mathrm{sub}} \end{bmatrix}$$

where each block $[\mathbf{A}_\mathrm{m}]_i^j \in \mathbb{R}^{p_\mathrm{m} \times p_\mathrm{m}}$ is a diagonal matrix defined by

$$[\mathbf{A}_\mathrm{m}]_i^j = \mathrm{diag}\left\{\sum_{l=1}^{l_\mathrm{m}} \boldsymbol{\alpha}_i^l \odot \bar{\boldsymbol{\alpha}}_j^l\right\}$$

with $\boldsymbol{\alpha}_j^l = \sum_{b=1}^{l_\mathrm{h}} [\mathbf{L}_\mathrm{m}]_b^l \dot{\mathbf{M}}^b \mathbf{V}_j^b$. This computation is detailed in Algo. 1. Since the matrix $\mathbf{A}_\mathrm{m}$ is symmetric, note that only its upper (or lower) triangular part needs to be calculated.

Similarly, the matrix $\mathbf{A}_\mathrm{h} \triangleq \underline{\mathbf{V}}^H \underline{\dot{\mathbf{H}}}^H \underline{\mathbf{L}}_\mathrm{h}^H \underline{\dot{\mathbf{S}}}^H \underline{\dot{\mathbf{S}}\mathbf{L}_\mathrm{h}\dot{\mathbf{H}}\mathbf{V}}$ defined by the HS forward model can be decomposed into $l_\mathrm{sub} \times l_\mathrm{sub}$ elementary blocks such that

$$\mathbf{A}_\mathrm{h} = \begin{bmatrix} [\mathbf{A}_\mathrm{h}]_1^1 & \dots & [\mathbf{A}_\mathrm{h}]_{l_\mathrm{sub}}^1 \\ \vdots & \ddots & \vdots \\ [\mathbf{A}_\mathrm{h}]_1^{l_\mathrm{sub}} & \dots & [\mathbf{A}_\mathrm{h}]_{l_\mathrm{sub}}^{l_\mathrm{sub}} \end{bmatrix}$$

where each block $[\mathbf{A}_\mathrm{h}]_i^j \in \mathbb{R}^{p_\mathrm{m} \times p_\mathrm{m}}$ is also decomposed into $d^2 \times d^2$ diagonal matrices of size ${}^{p_\mathrm{m}}/_{d^2} \times {}^{p_\mathrm{m}}/_{d^2}$, i.e.,

$$[\mathbf{A}_\mathrm{h}]_i^j = \boldsymbol{\beta}_i^j \odot \boldsymbol{\Upsilon}_{p_\mathrm{m}, d^2}$$

with $\boldsymbol{\beta}_i^j = \frac{1}{d^2} \sum_{l=1}^{l_\mathrm{h}} \left([\mathbf{L}_\mathrm{h}]_l^l\right)^2 \mathbf{V}_j^l \mathbf{V}_i^l (\dot{\mathbf{H}}^l)(\dot{\mathbf{H}}^l)^H$ and

$$\boldsymbol{\Upsilon}_{p_\mathrm{m}, d^2} = \begin{bmatrix} \mathbf{I}_{p_\mathrm{m}/d^2} & \dots & \mathbf{I}_{p_\mathrm{m}/d^2} \\ \vdots & \ddots & \vdots \\ \mathbf{I}_{p_\mathrm{m}/d^2} & \dots & \mathbf{I}_{p_\mathrm{m}/d^2} \end{bmatrix} \tag{11}$$

Note that a large number of coefficients in $\mathbf{A}_\mathrm{h}$ are zeros, which avoids to compute all the entries in the matrices $\boldsymbol{\beta}_i^j$ but only its non-zero coefficients whose positions correspond to the non-zero values in $\boldsymbol{\Upsilon}_{p_\mathrm{m}, d^2}$. This is summarized in Algo. 2 which also benefits from the Hermitian symmetry of $\mathbf{A}_\mathrm{h}$. It is also worth noting that the $2l_\mathrm{sub}^2$ blocks defining the matrices $\mathbf{A}_\mathrm{m}$ and $\mathbf{A}_\mathrm{h}$ of the system matrix $\mathbf{A}$ can be computed independently, which is highly amenable to parallelization to benefit from multi-core processing.

Algorithm 1 Computing $\mathbf{A}_\mathrm{m}$

Input: $\mathbf{L}_\mathrm{m}$, $\mathbf{V}$, $\dot{\mathbf{M}}$
 # Compute all $\boldsymbol{\alpha}_j^l$ ($\forall j, l$)
for $l = 1$ to l_m **do**
  for $j = 1$ to l_sub **do**
    $\boldsymbol{\alpha}_j^l = \sum_{b=1}^{l_\mathrm{h}} [\mathbf{L}_\mathrm{m}]_b^l \dot{\mathbf{M}}^b \mathbf{V}_j^b$
  end for
end for
 # Fill-in $\mathbf{A}_\mathrm{m}$ block-by-block
for $i = 1$ to l_sub **do**
  for $j = 1$ to $(l_\mathrm{sub} - i)$ **do**
    if $(j \neq 0)$ **then**
      $[\mathbf{A}_\mathrm{m}]_i^{j+i} = \sum_{l=1}^{l_\mathrm{m}} \boldsymbol{\alpha}_i^l \odot \bar{\boldsymbol{\alpha}}_j^l$
      $[\mathbf{A}_\mathrm{m}]_{j+i}^i = \overline{[\mathbf{A}_\mathrm{m}]_i^{j+i}}$
    else
      $[\mathbf{A}_\mathrm{m}]_i^i = \sum_{l=1}^{l_\mathrm{m}} \boldsymbol{\alpha}_i^l \odot \bar{\boldsymbol{\alpha}}_i^l$
    end if
  end for
end for
Output: $\mathbf{A}_\mathrm{m}$

Algorithm 2 Computing $\mathbf{A}_\mathrm{h}$

Input: $\mathbf{L}_\mathrm{h}$, $\mathbf{V}$, $\dot{\mathbf{H}}$, d
 # Fill-in $\mathbf{A}_\mathrm{h}$ block-by-block
for $i = 1$ to l_sub **do**
  for $j = 1$ to $(l_\mathrm{sub} - i)$ **do**
    if $(j \neq 0)$ **then**
      # Identify non-zero elements in the block $\binom{j}{i}$
      for (m, n) s.t. $[\boldsymbol{\Upsilon}_{p_\mathrm{m}, d^2}]_m^n \neq 0$ **do**
        $[\mathbf{A}_\mathrm{h}]_{i,m}^{i+j,n} = \frac{1}{d^2} \sum_{l=1}^{l_\mathrm{h}} \left([\mathbf{L}_\mathrm{h}]_l^l\right)^2 \mathbf{V}_j^l \mathbf{V}_i^l \dot{\mathbf{H}}_m^l \dot{\mathbf{H}}_n^l$
        $[\mathbf{A}_\mathrm{h}]_{i+j}^i = \overline{(\mathbf{A}_\mathrm{h})_i^{i+j}}$
      end for
    else
      # Identify non-zero elements in the block $\binom{i}{i}$
      for (m, n) s.t. $[\boldsymbol{\Upsilon}_{p_\mathrm{m}, d^2}]_m^n \neq 0$ **do**
        $[\mathbf{A}_\mathrm{h}]_{i,m}^{i,n} = \frac{1}{d^2} \sum_{l=1}^{l_\mathrm{h}} \left([\mathbf{L}_\mathrm{h}]_l^l\right)^2 \mathbf{V}_i^l \mathbf{V}_i^l \dot{\mathbf{H}}_m^l \dot{\mathbf{H}}_n^l$
      end for
    end if
  end for
end for
Output: $\mathbf{A}_\mathrm{h}$

Finally, the last matrix involved in the definition of $\mathbf{A}$ is $\mathbf{A}_\mathrm{r} \triangleq \underline{\dot{\mathbf{D}}}^H \underline{\dot{\mathbf{D}}}$, which is a diagonal matrix and easily computable.

REFERENCES

[1] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral detection*

and classification. New York: Kluwer, 2003.
[2] ——, *Hyperspectral Data Exploitation: theory and applications*. Hoboken, NJ: John Wiley & Sons, 2007.
[3] S. Douté, B. Schmitt, Y. Langevin, J.-P. Bibring, F. Altieri, G. Bellucci, B. Gondet, F. Poulet, and the MEX OMEGA team, "South Pole of Mars: Nature and composition of the icy terrains from Mars Express OMEGA observations," *Planetary and Space Science*, vol. 55, pp. 113–133, Jan. 2007.
[4] C. Colliex, M. Tencé, E. Lefèvre, C. Mory, H. Gu, D. Bouchet, and C. Jeanguillaume, "Electron energy loss spectrometry mapping," *Microchimica Acta*, vol. 114, pp. 71–87, 1994.
[5] N. Dobigeon and N. Brun, "Spectral mixture analysis of EELS spectrum-images," *Ultramicroscopy*, vol. 120, pp. 25–34, Sept. 2012.
[6] E. Monier, T. Oberlin, N. Brun, M. Tenc, M. de Frutos, and N. Dobigeon, "Reconstruction of partially sampled multiband images – application to stem-eels imaging," *IEEE Trans. Comput. Imag.*, vol. 4, no. 4, pp. 585–598, Dec 2018.
[7] P. A. A. De Beule, C. Dunsby, N. P. Galletly, G. W. Stamp, A. C. Chu, U. Anand, P. Anand, C. D. Benham, A. Naylor, and P. M. W. French, "A hyperspectral fluorescence lifetime probe for skin cancer diagnosis," *Review of Scientific Instruments*, vol. 78, no. 12, 2007.
[8] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, and J. Frias, "Hyperspectral imaging : an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, no. 12, pp. 590–598, 2007.
[9] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, M. Selva, and C. Chen, "25 years of pansharpening: a critical review and new developments," *Signal and Image Processing for Remote Sensing*, pp. 533–548, 2011.
[10] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simes, J. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya, "Hyperspectral pansharpening: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.
[11] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion : A comparative review of the recent literature," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 29–56, Jun 2017.
[12] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.
[13] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images. II. Channel ratio and "chromaticity" transformation techniques," *Remote Sens. Environment*, vol. 22, no. 3, pp. 343–365, Aug 1987.
[14] O. Berné, A. Helens, P. Pilleri, and C. Joblin, "Non-negative matrix factorization pansharpening of hyperspectral data: An application to mid-infrared astronomy," in *Proc. IEEE GRSS Workshop Hyperspectral Image SIgnal Process.: Evolution in Remote Sens. (WHISPERS)*, 06 2010, pp. 1–4.
[15] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct 1999.
[16] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb 2012.
[17] M. Simoes, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, Jun 2015.
[18] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul 2015.
[19] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep 2010.
[20] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Fast fusion of multi-band images based on solving a Sylvester equation," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4109–4121, Nov 2015.
[21] Q. Wei, N. Dobigeon, J. Tourneret, J. Bioucas-Dias, and S. Godsill, "R-fuse: Robust fast fusion of multiband images based on solving a sylvester equation," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1632–1636, Nov 2016.
[22] J. E. Pearson, "Atmospheric turbulence compensation using coherent optical adaptive techniques," *Appl. Opt.*, vol. 15, no. 3, pp. 622–631, Mar 1976.
[23] L. Rayleigh, "Investigations in optics, with special reference to the spectroscope," *Monthly Notices of the Royal Astronomical Society : Letters*, vol. 40, p. 254, Feb 1880.
[24] F. Soulez, E. Thiébaut, and L. Denis, "Restoration of hyperspectral astronomical data with spectrally varying blur," in *EAS Publications Series*, ser. EAS Publications Series, D. Mary, C. Theys, and C. Aime, Eds., vol. 59, Mar 2013, pp. 403–416.
[25] M. A. Hadj-Youcef, F. Orieux, A. Fraysse, and A. Abergel, "Restoration from multispectral blurred data with non-stationary instrument response," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Aug 2017, pp. 503–507.
[26] R. Szeliski, *Computer vision: algorithms and applications*, ser. Texts in Computer Science. London, U.K.: Springer-Verlag, 2011.
[27] J.-L. Starck and F. Murtagh, *Astronomical Image and Data Analysis*, 2nd ed., ser. Astronomy and Astrophysics Library. Heidelberg, Germany: Springer-Verlag, 2006.
[28] 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," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1988.
[29] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
[30] E. Wycoff, T.-H. Chan, K. Jia, W.-K. Ma, and Y. Ma, "Non-negative sparse promoting algorithm for high resolution hyperspectral imaging," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, Vancouver, Canada, 2013.
[31] W. C. Karl, "Regularization in image restoration and reconstruction," in *Handbook of Image and Video Processing*. Elsevier, 2005, pp. 183–V.
[32] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, pp. 183–202, 01 2009.
[33] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," Carnegie Mellon University, Pittsburgh, PA, USA, Tech. Rep., 1994.
[34] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *European Conf. Computer Vision (ECCV)*. Springer, 2014, pp. 63–78.
[35] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of multi-band images," *IEEE J. Sel. Topics Signal Processing*, vol. 9, no. 6, pp. 1117–1127, Sept. 2015.
[36] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2015, pp. 3586–3594.
[37] Q. Wei, J. M. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, M. Chen, and S. Godsill, "Multi-band image fusion based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7236–7249, Dec. 2016.
[38] J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," *Math. Comp.*, vol. 19, pp. 297–301, 1965.
[39] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed. USA: Prentice Hall Press, 2009.
[40] J. O. Smith, *Spectral Audio Signal Processing*. W3K-publishing, 2011, online book, 2011 edition.
[41] C. Guilloteau, T. Oberlin, O. Berné, É. Habart, and N. Dobigeon, "Simulated JWST datasets for multispectral and hyperspectral image fusion," *The Astronomical Journal*, vol. 160, no. 1, p. 28, Jun. 2020.
[42] M. D. Perrin, R. Soummer, E. M. Elliott, M. D. Lallo, and A. Sivaramakrishnan, "Simulating point spread functions for the James Webb Space Telescope with WebbPSF," in *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 8442, Sep 2012, p. 84423D.
[43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr 2004.
[44] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at ihs-like image fusion methods," *Information Fusion*, vol. 2, no. 3, pp. 177 – 186, 2001.
[45] M. e. A. Hadj-Youcef, F. Orieux, A. Fraysse, and A. Abergel, "Spatio-spectral multichannel reconstruction from few low-resolution multispectral data," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Rome, Italy, Sept. 2018.
[46] J. Kaipio and E. Somersalo, "Statistical inverse problems: discretization, model reduction and inverse crimes," *J. Computation. Appl. Math.*, vol. 198, no. 2, pp. 493–504, 2007.