Full-Spectrum Denoising of High-SNR Hyperspectral Images

Miguel Colom[*], **Jean-Michel Morel**

CMLA, ENS Cachan, CNRS, Université Paris-Saclay, 94235 Cachan, France
[*]*colom@cmla.ens-cachan.fr*

Abstract: The high spectral redundancy of hyper/ultraspectral Earth-observation satellite imaging raises three challenges: a) to design accurate noise estimation methods, b) to denoise images with very high SNR, and c) to secure unbiased denoising. We solve (a) by a new noise estimation (b) by a novel Bayesian algorithm exploiting spectral redundancy and spectral clustering, and (c) by accurate measurements of the interchannel correlation after denoising. We demonstrate the effectiveness of our method on two ultraspectral Earth imagers, IASI and IASI-NG, one flying and the other in project, and sketch the major resolution gain of future instruments entailed by such unbiased denoising.

1. Introduction

Hyperspectral (HSI) and Ultraspectral Imaging (USI) have numerous applications, such as the characterization of the abundance of components found in planetary exploration [1], the measurement of the thermodynamic, chemical and climate variables of the Earth atmosphere through satellite sounders [2] but also, in the case of HSI, medical applications such as tumor detection [3], inspection of food and agricultural products [4], among many others. To make the reading easier, in the text we will refer from now to “HSIs" for both HSIs and USIs, without loss of generality. Hyperspectral imaging is generally characterized by a low resolution spatial sampling and high resolution spectral sampling, hence requiring unmixing algorithms to augment the spatial resolution and characterize the detected chemical or atmospheric events [5].

The data acquired by a hyperspectral instrument is affected by internal random perturbations such as the dark currents inside the sensor, readout errors, and absorption of electromagnetic energy by the semiconductors of the device, as well as by external sources, particularly photon noise due to the intrinsic quantum nature of light. All of these random perturbations of the measurement are known as *noise* in the acquired HSI.

With the exception of photon noise, the set of the other noise sources can be modeled as homoscedastic additive white Gaussian noise, whereas photon noise depends on the temperature of the scene and the frequency of the channel being observed. However, photon noise can be well characterized by Planck’s law. Thus one can formulate a complete and realistic noise model, which will be discussed in Sec. 2. Nevertheless, reestimating the noise components on board remains necessary.

Our main focus here is to reconsider and reformulate the noise estimation and the image denoising problem for high Signal-to-Noise Ratio (SNR) satellite hyperspectral images (where the noise is mainly governed by Planck’s law), by taking advantage of the high redundancy of the hyperspectrum. The latest advances in the design of the hyperpectral sensors allow one to measure very small variations of the signal in narrow spectral channels with a very low noise added by the electronics (i.e. dark, readout, shot, or electromagnetic interference noises). Therefore, modern instruments deliver high-SNR images. Yet, precisely, this high SNR and an even higher one are needed to evaluate (for example) the tiny variations of some diluted chemical components in the atmosphere.

We shall see that the observation of several thousands spectral channels on many data makes it possible to estimate the noise model from the data themselves. We shall also be in a position to check that the estimated noise matches the physical model.

The new method we present here is based on some of the ideas sketched in our conference paper [6]. We show Peak Signal-to-Noise Ratio (PSNR) and correlation results, before and after denoising. Actually we introduce a similar but more robust metric, the Median Signal to Noise Ratio (MSNR), Eq. (9). Our main goal is to reach a bias-free denoising method, which is defined as a method introducing no correlation in the signal. This is particularly important for applications to the detection of small variations in temperatures and densities of chemical components, where denoising can be followed by sophisticated detection and unmixing methods [7, 8].

The classic denoising tools that have been tried on HSIs in the literature comprise: wavelet thresholding, total variation, 3D patch-based methods, coupled with decorrelation tools. We now review the most effective ones.

The authors of [9] and [10] applied denoising to HSIs (70 to 200 channels), assuming low or very low SNR. In this low SNR context their algorithm could still be compared to classic patch-based algorithms such as Video Block Matching 3D (VBM3D) [11]. As we shall see for high SNR, HSIs with thousands of channels, patch-based methods are no longer competitive. They underperform when compared to algorithms exploiting the strong spectral redundancy.

The method proposed in [12] simulates uniform Gaussian noise and some "sparse noise" which is not related to Planck's law, and therefore not adapted to hyperspectral imaging. This method involving 2-D total variation along the spatial dimension and 1-D total variation along the spectral dimension may be used for multi-spectral images and low SNR. Again, this method is not relevant for high SNR hyperspectral images, because the spatial total variation forces local spatial correlations. Similarly the denoising method in [13] and [14] involves a minimization of the 2-D total variation along the spatial dimension and of the 1-D total variation along the spectral dimension. This is again adapted to HSIs with low SNR. The authors only consider a combination of Gaussian and impulse noise which does not correspond to the physics of hyperspectral images. A more up to date approach is presented in [15]. This paper builds spatial-spectral dictionaries of patches to exploit the sparsity across the spatial-spectral domain, the high correlation across spectra, and the non-local self-similarity over space. The denoising energy model which depends on three parameters is minimized by an iterative method. This model is clearly related to what we will propose, but again used in a different framework: hyperspectral rather than USIs, with low SNR. The proposed iterative method would be far too slow for our purposes. The dictionary approach can be accelerated as proposed in [16]. This paper describes an attempt at a fast greedy method based on orthogonal matching pursuit. This again is done in the context of multispectral images with low SNR, and is not adapted to the high SNR setup of ultraspectral imaging. A recent method based also on dictionary learning and decomposable nonlocal tensors (TDL) [17] is nowadays one of the methods that reach the best performance for low and moderate SNR HSIs. It exploits two well-known intrinsic characteristics of HSIs: the spatial nonlocal autosimilarity and the global correlation across the spectrum. We will also compare our results with TDL [17] and VBM3D [11] and show that even the best methods of the state of the art that are most performant in the literature are not adapted to the very high SNR images that are obtained by the next generation of ultraspectral sensors, including that of the future Infrared Atmospheric Sounding Interferometer Next Generation (IASI-NG) satellite.

We shall refer to a method based on 2-D non-subsampled shearlet transform (NSST) [18] and fully constrained least-squares unmmixing (FCLSU) [19]. This method splits the channels of the HSI between low and high noisy sets according to their spectral correlation, and then NSST is applied at each channel. Finally, the low-noise bands are denoised by thresholding shearlet coefficients and high-noise channels with FCLSU.

The method proposed by Chen [20] seems to be another state of the art method and it contains several arguments that we shall reuse here. Similarly to the present work, the authors consider hyperspectral images with signal-to-noise ratio (such as 600 : 1). They decorrelate the image information of hyperspectral data cubes from the noise by using principal component analysis (PCA). They argue that the first PCA channels contain a majority of the total energy of the HSI and that removing noise in the low-energy PCA channels will not harm the fine features of the data cubes. A 2-D bivariate wavelet thresholding method is therefore used to remove the noise for low-energy PCA channels, while instead a 1-D dual-tree complex wavelet transform denoising method is used to remove the noise of the spectrum of each pixel of the data cube.

The results of our denoising method, Dual Bispectral Bayesian Denoising (DBBD) outperform state-of-the art methods such as [19] and [20]. The DBBD results are presented in Sec. 4.

Our plan is as follows. In section 2 we describe the noise model. It is valid for IASI and the future satellite IASI-NG. In section 2.2 we describe the simulation data used for IASI-NG. Section 3 details our proposed method for hyperspectral images with high SNR and large spectrum. Section 4 details the evaluation of the method on simulated and real hyperspectral satellite images with high SNR and large spectrum. Section 5 is a conclusion on the perspectives open to satellite design by high performance unbiased denoising.

2. The Noise Model and its Estimation

Assuming a black body, the emission of photons is related only to its temperature T [21] and it can be modeled by a Poisson distribution whose variance σ^2 is equal to its mean μ. By definition of the Signal to Noise Ratio, we have SNR $= \frac{\mu}{\sigma} = \frac{\lambda}{\sqrt{\lambda}} = \sqrt{\sigma^2} = \sigma$. Since the Noise Equivalent Differential Temperature (NEDT) is defined as the change in temperature that yields an SNR ratio of unity, we can write that the standard deviation of the noise is equal to the Noise Equivalent Delta Radiance (NEDL), which is expressed as [22–24]:

$$\mathrm{NEDL}(\nu, T) = \mathrm{NEDT}(\nu, T) \times \left| \frac{\partial L(\nu, T)}{\partial T} \right|,$$

where $L(\nu, T)$ is the black body radiance and ν the corresponding wavenumber. The values of NEDT are an intrinsic characteristic of the acquisition sensor and they can be tabulated according to frequency and temperature values. In practice one assumes that the temperature of the black body is constant and it is fixed to a certain reference, typically $T = 280\ K$ (background-limited system).

According to Planck's law,

$$L(\nu, T) = \frac{c_1 \nu^3}{e^{c_2 \nu / T} - 1},$$

where c_1 and c_2 are the first and the second radiation constants. Thus,

$$\frac{\partial L(\nu, T)}{\partial T} = \frac{c_1 c_2 \nu^4}{T^2} \frac{e^{c_2 \nu / T}}{\left(e^{c_2 \nu / T} - 1\right)^2}.$$

Finally, the standard deviation of the noise can be obtained by the following closed expression [25]:

$$\sigma_{\mathrm{noise}}(\nu, T) = \mathrm{NEDT}(\nu, T) \times \frac{c_1 c_2 \nu^4}{T^2} \frac{e^{c_2 \nu / T}}{\left(e^{c_2 \nu / T} - 1\right)^2}. \tag{1}$$

The physical details behind these well-known formulations are out of the scope of this article, and we refer the reader to [25], as well as to [26, 27] for more information.

Assuming a (non-available) ideal noise-free HSI or USI $\mathbf{G}$, the noisy HSI $\tilde{\mathbf{G}}$ is obtained as

$$\tilde{\mathbf{G}}_\nu = \mathbf{G}_\nu + \mathcal{N}_\nu, \tag{2}$$

where $\mathcal{N}_\nu$ is a Poisson random variable whose parameter is the NEDL.

2.1. Noise Estimation

Sec. 2 discussed how to obtain the standard deviation of the noise in a particular image in the hyperspectral cube corresponding to a wavenumber ν. While this physical model is valid for a black body, there are situations where it might be useful to estimate the noise from the data instead of assuming directly the ideal black body scenario. First, the radiometric characterization is not quite realistic. An Earth-observing scene in space-flight might not correspond to the one obtained in pre-flight calibration, where the instrument is calibrated using a sphere under constant lighting simulating a black body. The scene observed during space calibration is likely to have spectral radiance characteristics which do not correspond exactly to the black body assumption [28].

Second, the signal in the sensor might be censored or even get saturated if it overflows its dynamic range, causing a nonlinearity of the dependence of the variance of the noise with respect to the intensity when the signal approaches saturation points. This leads to censoring and clipping effects [29].

Third, the instrument mounted on the satellite might be made of several sensors with spectral overlapping. For example, in the case of the IASI-NG, the instrument is made of four hyperspectral sensors (four bands) with spectral overlapping of $\pm 40\text{cm}^{-1}$. Thus, the variance of the noise at the intersections is halved by averaging.

For these reasons, instead of assuming the black body model directly, it may be recommended to estimate the noise model from the noisy hyperspectral image itself. We propose in Algo. 1 a simple procedure to estimate the noise from HSIs. Its adequacy will be verified on both simulated and real data.

Algorithm 1 Estimate the noise level function of an HSI

1: **HSI noise estimation**
  Input $\tilde{\mathbf{G}}$: acquired noisy HSI
  Input $w = 100$: window size
  Output $\hat{\sigma}$: noise estimation
2: $R \leftarrow \text{corrcoef}(\tilde{\mathbf{G}})$ ▷ Pearson cross-correlation coeffs.
3: $\hat{\sigma} \leftarrow \text{zeros}(R.\text{shape}[0])$ ▷ First dimension: frequencies
4: # Raw noise estimation
5: **for** $i \in \text{range}(R.\text{shape}[0])$ **do**
6: $R[i,i] \leftarrow -\infty$ ▷ Exclude trivial diagonal maxs
7: $j \leftarrow \text{argmax } R[i,:]$ ▷ Get most correlated frequency
8: $k \leftarrow \text{mean}(\tilde{\mathbf{G}}_i)/\text{mean}(\tilde{\mathbf{G}}_j)$
9: $\tilde{\mathbf{G}}_j \leftarrow \tilde{\mathbf{G}}_j \times k$ ▷ Mean equalization
10: $\hat{\sigma}[i] \leftarrow ||\tilde{\mathbf{G}}_i - \tilde{\mathbf{G}}_j||/\sqrt{2}$ ▷ Raw noise estimation
11: **end for**
12: # Final filtered estimation
13: **for** $i \in \text{range}(0, \text{len}(\hat{\sigma}), w)$ **do** ▷ Non-overlapping window
14: $j \leftarrow i + w - 1$
15: **if** $j \geq \text{len}(\hat{\sigma})$ **then** ▷ Take care of array boundary
16: $j \leftarrow \text{len}(\hat{\sigma}) - 1$
17: **end if**
18: $\hat{\sigma}[i : j+1] \leftarrow \min(\hat{\sigma}[i : j+1])$ ▷ Assign min to window
19: **end for**
20: **return** $\hat{\sigma}$

The algorithm has two steps. First (lines 4–12) it finds for every channel i the channel $j \neq i$ that gives the maximum correlation, equalizes the data in both channels to have the same mean, and computes a raw estimation of the noise variance from their difference. This empirical variance (computed on line 11) is likely to contain outliers, since the most correlated channel with respect to a given one can have a larger difference than the one caused by the small noise level.

The second step (lines 13–19) removes the outliers. The assumption is that the noise levels inside a short frequential window should not be very different, since they are linked to similar frequencies. Thus, the minimum value of the noise empirical standard deviation should be close to the real noise value. The obvious drawback of using the min operator is that if most of the values are close to the real noise level, then the obtained result might get a small underestimation bias, as can be observed in the wavenumbers from 150000 to 200000 in Fig. 1. This figure shows the noise estimation in a simulated IASI-NG with simulated added noise with the physical model, Eq. (1). This means that after the NEDT of the sensor had been measured in the laboratory and tabulated, we applied Planck's law to obtain the NEDL and therefore the frequency-dependent standard deviation (STD) of the noise. In red, one can see the standard deviation of the noise according to the wavenumber given by the physical model. In green, our estimation. Both curves match, and their accuracy is more than sufficient for denoising purposes.

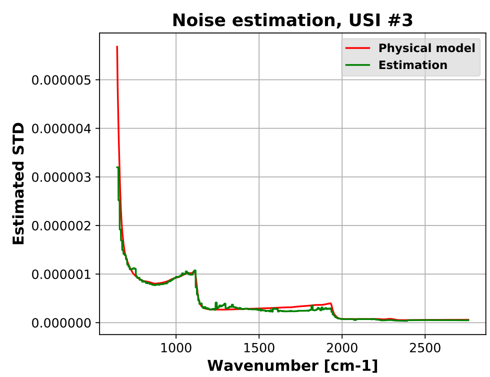


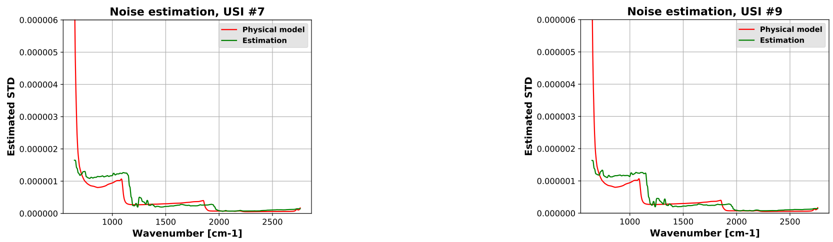
Fig. 1. Noise estimation with simulated data and simulated added noise according the physical model of the IASI-NG USI #3. In red, the standard deviation of the noise according to the wavenumber given by the physical model. In green, our estimation. Both curves match, and their accuracy is more than sufficient for denoising purposes. The noise estimation was obtained step-wise for each wavenumber.

Fig. 2 shows a noise estimation of two real IASI USIs. In that case the noise was not simulated and added: It is the real noise present in the acquired USI. Here again the estimations match well the noise predicted by the physical model.

This procedure is different and actually simpler than the noise estimator proposed in [30], which is, as the authors claim, valid for extremely noisy hyperspectral applications. Here we are instead working on high SNR hyperspectral data.

In order to confirm that the noise estimation procedure was accurate, we performed two kinds of tests: one with simulated data (see Sec. 2.2 for the details on the data model) and added noise according to the physical model presented in Sec. 2, as well estimating the noise in real HSIs and comparing the obtained result with the physical noise model.

The noise estimation algorithm is based on the strong redundancy of HSIs, in particular on the large cross-correlations along the different spectral channels. Fig. 3 shows an example of the band B2 of an USI of IASI-NG. Note that we use the correlation and not the L2 or similar norms to compare the data along different frequencies since the gain (say, the atmospheric absorption) at different channels can be different. However, the Pearson cross-correlation of the data is robust to these global changes.



USI #7 USI #9

Fig. 2. Noise estimation in the real IASI USIs #7 and #9. The noise was estimated from the original USI data. Similar results were obtained with the other real IASI USIs. In red, the standard deviation according to the wavenumber given by the physical model, Eq. (1). In green, our estimation. The noise estimation was obtained step-wise for each wavenumber.

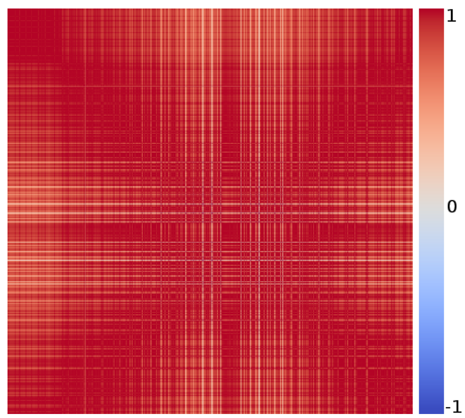


Fig. 3. The normalized Pearson cross-correlation matrix for band B2 (from 115000.0 m^{-1} to 195987.5 m^{-1}) in USI #1. The red color represents a high correlation, while white means low (*coolwarm* color map). Most of the channels are correlated with other many. Only a few frequencies are uncorrelated because of the presence of gases in the atmosphere at those particular frequencies.

2.2. Our Real and Simulated Data

We performed experiments on both simulated and real data. The simulated data model was obtained by the α-IASI-β radiative transfer model, originally developed for IASI by the European Space Agency EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites). Yet this model is realistic for any nadir viewing satellite, aerial sensor, or ground observation [31].

Fig. 4 shows the simulated scene radiance at two different frequencies of the first band of USI #0 in our dataset. Fig. 5 corresponds however to real data obtained from an actual IASI instrument.

3. The proposed denoising method

This section describes our proposed *Dual Bispectral Bayesian Denoising* (DBBD) method. Its pseudo-code is given in Algo. 2. Most of the frequencies of an HSI are highly correlated with the rest (see Fig. 3) with the exception of a few of them, and therefore also the acquired images (Fig. 4). A fundamental denoising principle is to exploit redundancy of the data, and particularly

(a) 88362.5 m^{-1} (b) 96500.0 m^{-1}

Fig. 4. Simulation of the images acquired by the IASI-NG instrument at different frequencies (USI #0, band B1). The intensity of each pixel represents the energy measured in $\mathrm{W/m^2/sr/m^{-1}}$ units.

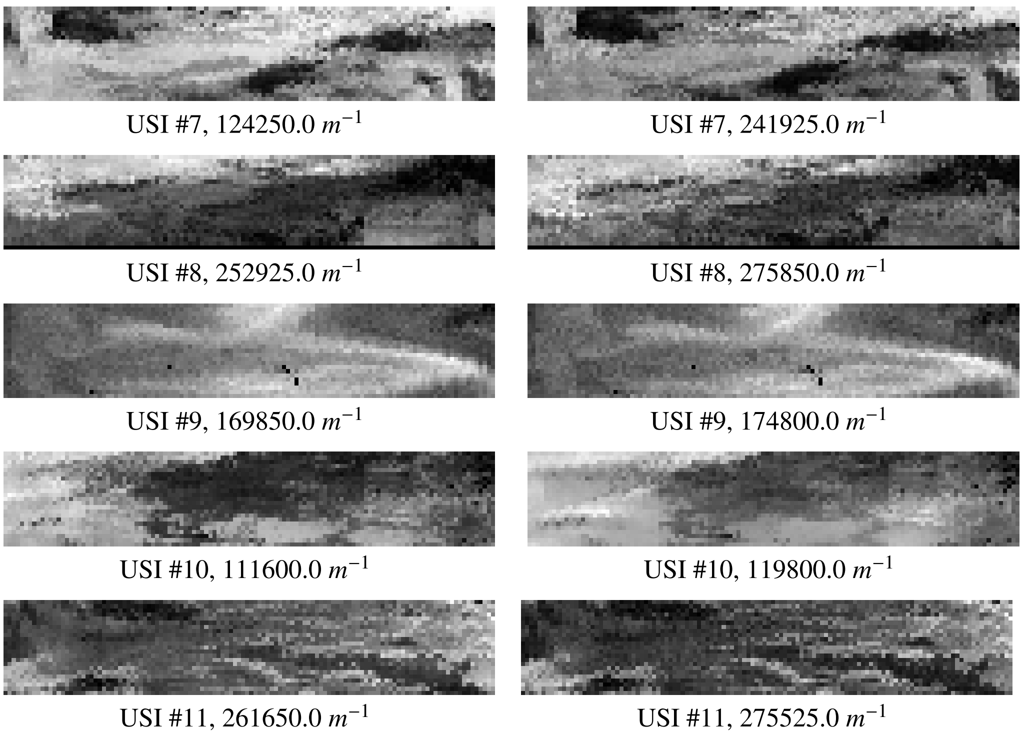


Fig. 5. Two different frequencies of the USIs #7–#11 in our database. This corresponds to real IASI acquisitions, each of them of 120×23 pixels.

its self-similarity. The general principle invented by Lee [32] is to group pixels having a similar model. Groups of similar pixels are detected by comparing their position in the image domain and their values. In a recent image denoising method, the same principle was expanded by grouping pixels which have similar neighborhoods [33], before denoising them together. The reason for comparing patches instead of pixels is to ensure that the grouped pixels really obey the same model. In the case of hyperspectral images, the high number of values available at each pixel makes it unnecessary and actually counter-productive to use patches. Each pixel is indeed extremely well characterized by its spectral signature. We even found that the self similarity principle should be applied first in the frequency domain. This means that we treat frequencies as dual pixels and group them according to their similarities, before going back to the image domain and grouping the pixels in clusters by their similarity. Thus, we build frequency clusters for each HSI by grouping highly correlated frequencies. We perform this using the K-means clustering algorithm.

Each frequency cluster Λ_i therefore contains highly-correlated frequencies. The data at each of the frequencies is divided by the STD of the noise to equalize all frequency STDs to 1. The next step is to reduce the frequency domain dimensionality to $N = 20$ dimensions by PCA. We verified empirically that 20 PCs are enough to represent reliably the IASI-NG USIs (see Fig. 6). This is done in conformity with the literature which recommends a PCA reduction for an efficient representation of observations from high-resolution infrared sounders [34]. As pointed out in [35], one of the main advantages of the spectral domain over the spatial domain is the computation of low-noise basis vectors even from very noisy USIs. As these authors remark,

> Spectral distributions can be very compactly represented and there is an overwhelmingly large number of spectra in a typical USI. As we also expect noise to occur randomly, the impact of noise on the greatest directions of variance in the large set of spectra in an USI should be minimal. Thus PCA can be used to obtain low-noise top basis vectors even when noise is high.

We denote by $\tilde{\mathbf{A}}$ the pixels in Λ_i after dimensionality reduction.

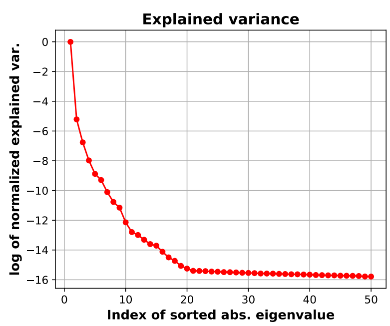


Fig. 6. Decay of the normalized explained variance according to the sorted eigenvalues, in log scale. There is no significant improvement if more than 20 PCs are used. This plot was obtained with data from USI #0. Similar results are obtained with the others.

For each cluster Λ_i we apply optimal Bayesian denoising [36] to each of the pixels of the HSI after PCA dimensionality reduction (only the first N principal components -PCs- are kept), where $\tilde{P}$ is the noisy pixel and P the denoised pixel:

$$P = \overline{\tilde{P}} + \left[\mathbf{C}_{\tilde{P}} - \mathbf{C}_n\right] \mathbf{C}_{\tilde{P}}^{-1} \left(\tilde{P} - \overline{\tilde{P}}\right). \tag{3}$$

In this equation, $\mathbf{C}_{\tilde{P}}$ is the empirical covariance matrix of the pixels most similar to the reference pixel in the cluster. We use the Pearson cross-correlation coefficient as the similarity criterion. The empirical covariance matrix $\mathbf{C}_n$ of the noise after PCA transformation is required too. It needs to be computed only for the wavenumbers in Λ_i. Sec. 3.1 gives the details on how to compute $\mathbf{C}_n$. For each cluster Λ_i, we apply the optimal Bayesian estimator. This estimator is optimal under two assumptions:

a) The frequency-dependent noise model is Gaussian and entirely characterized by its covariance matrix $\mathbf{C}_n$;

b) The pixel model is equally frequency-dependent Gaussian and characterized by its covariance matrix $\mathbf{C}_{\tilde{P}}$.

Assumption a) is realistic for spectrometers [22, 23]. Indeed, it is a classic result in imaging that the Poisson distribution can be well approximated by a Gaussian when the number of photons is large enough. Assumption b) is all the more true if our double grouping procedure, clustering first similar frequencies and second similar pixels is efficient. Very similar such pixels are only different by small variations caused by sampling and random local variations of all chemical components that can be accounted for by a Gaussian model.

At this point, the pixel has been partially denoised for the wavenumbers contained in the cluster, and only for the first N PCs. The next step is to denoise the rest of the PCs, which contains a very noisy signal but still some useful information. Thus, the rest of the spectrum encoded at those last PCs can be denoised, leading to a *full-spectrum* technique.

Here we apply the dual method of Chen [20] to the data in the cluster, which accounts for applying a dual-tree complex wavelet transform (DT-CWT) to the rest of the PCs and denoising them with Donoho's universal threshold, defined as

$$t = \sigma\sqrt{2\log(N_x N_y)}, \tag{4}$$

where $\sigma = 1$ is the STD of the noise (it is one because we first divided the noisy signal by the STD of the noise), and (N_x, N_y) is the size of the 2D image.

If we call $c_{i,j,l}$ the wavelet coefficient at pixel $[i, j]$ and level l, the denoised coefficient is obtained by soft thresholding as

$$\hat{c}_{i,j,l} = c_{i,j,l}\left(1 - t^2/|d_{i,j,l}|^2\right)^+, \tag{5}$$

where $|d_{i,j,l}|^2 = (|c_{i-1,j,l}|^2 + |c_{i,j,l}|^2 + |c_{i+1,j,l}|^2)/3$ is the average of the absolute squared coefficients in a small neighborhood.

Each cluster contains only a part of the total set of wavenumbers. Thus each pixel is denoised partly at each iteration. After all iterations, the pixel is completely denoised. Finally, the denoised pixels are projected back (the data is expressed in the same axes as before the PCA rotation) and the complete denoised USI is obtained. At the end, all the PCs (the first N with optimal Bayesian denoising and the rest with DT-CWT thresholding) and all the wavenumbers have been denoised.

3.1. Estimation of $\mathbf{C}_n$

Eq. (3) requires the empirical covariance matrix of the noise $\mathbf{C}_n$ on the PCA axes. This matrix is not available directly and must be estimated from the noisy data itself.

The rationale behind the estimation procedure presented here is that, after PCA, the first N PCs (in decreasing absolute eigenvalue) are dominated by the underlying signal, whereas the N-th and following PCs are mainly the contribution of the noise. We shall call M the number of wavenumbers in the current noisy cluster and $\mathbf{C}_n$ the empirical covariance matrix of the noise being estimated. For the PCs from N to M the signal is dominated by the noise and therefore the variance computed at these PCs corresponds mainly to the variance of the noise after PCA rotation [37].

For the first N PCs the underlying signal dominates and it is not possible to directly measure the variance of the noise. In order to estimate the noise, we make the following assumptions:

1. The PCA concentrated most of the variance of the noisy cluster Λ_i in the first PCs
2. Given that Λ_i was normalized in order that the noise had STD=1 along any wavenumber therefore (as a consequence of the previous point) $\mathbf{C}_n[0, 0] \approx 1$.

The variance of the noise from the N-th PC is well estimated and therefore one can approximate the variance of the noise from the first PC to the N-th. We chose to use a simplest estimator (a linear function) given that we do not assume any priors, but it could be improved if additional information is available. For example, from the statistics of many observations of real IASI-NG images, but for the moment these data is not available. However, the simple linear estimator had a good performance, as shown in Sec. 4.

Finally, we obtain the estimated empirical covariance matrix of the noise as:

$$\mathbf{C}_n[i,j] = \begin{cases} 1 + \frac{i \times \mathrm{var}(\Lambda_i[N,:]) - 1}{N} & \text{if i = j and i < N,} \\ \mathrm{var}(\Lambda_i[i,:]) & \text{if i = j and i} \geq \text{N,} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

DBBD performs a denoising which is adapted to the data itself and groups similar pixels in clusters. This covariance matrix is obtained for each of the clusters, given that the PCA rotation is performed at each of them.

Fig. 7 shows our estimated noise variance after PCA rotation, in one of the clusters. Both the estimation and the ground-truth curves match sufficiently to estimate an empirical covariance matrix of the noise, $\mathbf{C}_n$.

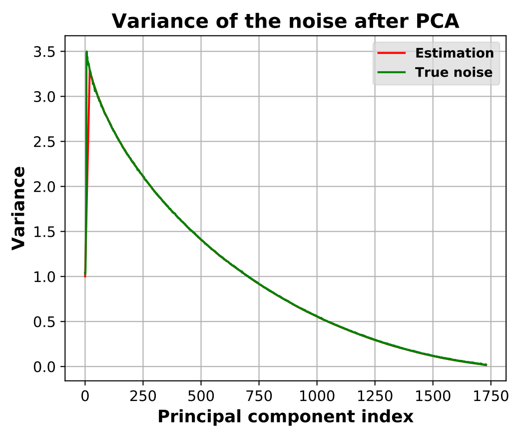


Fig. 7. Estimated noise variance after PCA rotation for USI #6, band #4, cluster #2 (similar results are obtained with the other HSIs, bands, and clusters). In red: our estimation. In green: the true value. The curves match sufficiently to estimate an empirical covariance matrix of the noise, $\mathbf{C}_n$.

3.2. The conservative Lee correction

Eq. (3) performs totally blind denoising of each pixel of the input USI, for the first N PCs. It assumes that the K most similar hyperspectral pixels used to compute the covariance matrix $\mathbf{C}_{\tilde{P}}$ are samples of a Gaussian vector, their differences being due to noise and to a combination of feeble independent factors. The other PCs (assumed to contain mainly noise but also minimal detail information) are denoised by soft thresholding.

If the denoising process is perfect, the variance of the removed signal should match the variance of the noise. If its variance is larger than the variance given by the physical model, this means that the denoising is not only removing noise, but also removing details. Thus, in order to improve the denoised signal we compute a weighted sum of the denoised and of the noisy USI to put back details from the noisy image when the denoising has been detected as too *aggressive*. This technique is inspired by Lee's local filtering [32, 38], which estimates the true value of a pixel with a weighted sum of the noisy pixel and the average of other noisy pixels in a small neighborhood.

The weighted sum for a given wavenumber j is given by:

$$\hat{\mathbf{G}}'_j = \alpha\hat{\mathbf{G}}_j + (1-\alpha)\tilde{\mathbf{G}}_j. \tag{7}$$

The weighting factor is obtained by solving a simple equation which sets the variance of the removed signal to the variance given by the model. The unique solution is

$$\alpha_j^2 = \frac{\sigma_{N,j}^2}{\mathrm{var}(\tilde{\mathbf{G}}_j)} + \mathrm{var}(\hat{\mathbf{G}}_j) - 2\,\mathrm{cov}(\tilde{\mathbf{G}}_j, \hat{\mathbf{G}}_j). \tag{8}$$

Our final denoising algorithm, DBBD, is summarized in Algo. 2.

3.3. Discussion on the choice of the hyperparameters

DBBD has three hyperparameters: the number of spectral clusters Q, the number of similar pixels K, and N the number of PCA principal components kept.

These hyperparameters are fixed and depend on the application. For IASI-NG we found that the optimal values are $Q = 3$, $K = 400$, and $N = 20$. The actual values depend on the characteristics of the hyperspectral image, including the SNR, the size of the image, and the number of channels per pixel. The number of PCA components which is kept (N) can be obtained from many observations of different HSIs and choosing the value for which adding more PCs has no significant influence on the explained variance. We chose $N = 20$ from the experiment shown in Fig. 6. The number of spectral clusters depends on the number of channels. A low Q will create clusters whose pixels are not similar enough, and conversely a large Q will produce clusters with not enough samples. The exact value for Q depends again on the application and the number of channels of the HSI. We fixed $Q = 3$ which produces 1410 channels/cluster. The number of similar pixels N that will be used to compute the covariance matrix of the noisy pixels depends directly on the number of pixels of the sensor and the characteristics of the image. It should be large enough to allow computing the inverse of $\mathbf{C}_{\tilde{P}}$. If K is too large $\mathbf{C}_{\tilde{P}}$ will be computed with pixels which have a different spectral signature and the algorithm will fail to obtain a reliable model of the noisy. For IASI-NG we found that $K = 400$ is a good compromise. Table 1 shows the MSNRs obtained when varying the hyperparameters of the algorithm for HSI #2. *Optimal* is the default case of IASI-NG with $Q, N, K = \{3, 20, 400\}$, case #1 with $Q, N, K = \{8, 20, 400\}$, case #2 with $Q, N, K = \{3, 7, 400\}$, and case #3 with $Q, N, K = \{3, 20, 150\}$. Deviations from the optimal hyperparameters cause a slight decrease of the performance. For other type of sensors the values for Q, K, and N need to be re-calibrated from the observations.

Optimal	**Case #1**	**Case #2**	**Case #3**
63.44	62.39	62.08	62.84

Table 1. MSNRs obtained when varying the hyperparameters of the algorithm for HSI #2. *Optimal* is the default case of IASI-NG with $Q, N, K = \{3, 20, 400\}$, case #1 with $Q, N, K = \{8, 20, 400\}$, case #2 with $Q, N, K = \{3, 7, 400\}$, and case #3 with $Q, N, K = \{3, 20, 150\}$. Deviations from the optimal hyperparameters cause a slight decrease of the performance.

4. Experimental validation of DBBD

Our denoising method was tested and validated on both simulated and real USIs.

Algorithm 2 DBBD denoising of an HSI

Input/output:
  Input $\tilde{\mathbf{G}}$: acquired noisy HSI, of size (N_x, N_y, N_z)
  Input σ_n: standard deviations of the noise (wavenumber dep.)
  Input $Q = 3$: number of spectral clusters
  Input $K = 400$: number of similar pixels
  Input $N = 20$: number of PCA PCs kept
  Output $\hat{\mathbf{G}}'$: denoised HSI

$\hat{\mathbf{G}} \leftarrow$ zeros($\tilde{\mathbf{G}}$.shape) ▹ Placeholder
Clusterize in Q spectral groups $\Lambda_i, i \in [1, Q]$ ▹ K-means for cosine distance
for each cluster $\Lambda_i, i \in [1, Q]$ **do**
  Noise normalization: divide each coordinate of each sample in Λ_i by its STD taken from σ_n
  $\mathbf{Z} \leftarrow \mathrm{PCA}(\Lambda_i)$ ▹ Matrix of eigenvectors of the PCA of Λ_i, in decreasing abs. eigenvalue order
  sigmas_pca $\leftarrow \mathrm{STD}(\Lambda_i[:])$ ▹ STD of each wavenumber after PCA: Eq. (6).
  # Process the first PC's (from 0 to $N-1$):
  $\tilde{\mathbf{A}} \leftarrow \tilde{\mathbf{A}}[0:N]$ ▹ First N PCs
  $\mathbf{C}_n[i, j] \leftarrow$ sigmas_pca$[i]^2$ if $i = j$ (0 otherwise) ▹ Get noise covariance matrix
  for each noisy pixel $\tilde{P}$ **do**
    $\mathbf{V} \leftarrow$ set of the K most correlated pixels with $\tilde{P}$
    $\overline{\tilde{P}} \leftarrow \mathrm{mean}(\mathbf{V})$ ▹ Compute their empirical mean and
    $\mathbf{C}_{\tilde{P}} \leftarrow \mathrm{cov}(\mathbf{V})$ ▹ covariance matrix
    $P \leftarrow \overline{\tilde{P}} + \left[\mathbf{C}_{\tilde{P}} - \mathbf{C}_n\right] \mathbf{C}_{\tilde{P}}^{-1} \left(\tilde{P} - \overline{\tilde{P}}\right)$ ▹ Eq. (3)
    $\hat{\mathbf{A}} \leftarrow P$ ▹ Append denoised P
  end for
  $\mathbf{Z}[0:N] \Leftarrow \tilde{\mathbf{A}}$ ▹ Store Bayesian denoised
  # Process the remaining PC's (from N to end):
  $\mathbf{D} \leftarrow \mathrm{DTCWT}(\tilde{\mathbf{A}}[N:])$ ▹ DT-CWT of the remaining PCs
  $\mathbf{M}^2 \leftarrow (|\mathbf{D}^2_{i-1,j,l}| + |\mathbf{D}^2_{i,j,l}| + |\mathbf{D}^2_{i+1,j,l}|)/3$
  $\mathbf{T}[:] \leftarrow$ sigmas_pca$[:]\sqrt{2 \log N_x N_y}$ ▹ Universal threshold
  $\mathbf{Z}[N:] \leftarrow \mathbf{Z}[N:](1 - \mathbf{T}[:]^2/\mathbf{M}^2[:])$ ▹ Pointwise soft thresholding
  $\hat{\mathbf{G}}(\Lambda_i) \leftarrow \mathrm{PCA}^{-1}(\mathbf{Z})$ ▹ Invert PCA rotation
  $\hat{\mathbf{G}}(\Lambda_i) \leftarrow \hat{\mathbf{G}}(\Lambda_i) \times \sigma_n[\mathrm{idx}(\Lambda_i)]$ ▹ Invert STD normalization
end for
# Apply conservative variance correction:
for $j \in [0, \mathrm{shape}(\hat{\mathbf{G}})[0] - 1$ **do** ▹ Loop over wavenumbers
  $\alpha^2 \leftarrow \frac{\sigma_n^2[j]}{\mathrm{var}(\tilde{\mathbf{G}}[j,:])} + \mathrm{var}(\hat{\mathbf{G}}[j,:]) - 2\,\mathrm{cov}(\tilde{\mathbf{G}}[j,:], \hat{\mathbf{G}}[j,:])$ ▹ Eq. (8)
  $\hat{\mathbf{G}}'[j,:] \leftarrow \alpha\hat{\mathbf{G}}[j,:] + (1-\alpha)\tilde{\mathbf{G}}[j,:]$ ▹ Weighted sum, Eq. (7)
end for
return $\hat{\mathbf{G}}'$ ▹ Final denoised HSI

4.1. Evaluation with Simulated USIs

The reason to evaluate the method in simulations is that they allow to create an ideal USI where the physically correct noise model is respected. Thus, a reliable ground-truth (GT) is available. This yields a quantitative exact measurement of the performance of the method. To this aim we used simulated data of the IASI-NG satellite project [2, 39, 40]. The location of the seven simulated IASI-NG USIs is given in Fig. 8

The first performance test measures the PSNR of the denoised signal with respect to the noise-free USI. In order to avoid the effect of outlier radiance values, the median of the signal should be preferred to its maximum for the PSNR definition. This leads to using the MSNR between an USI $\mathbf{T}$ and the reference noise-free USI $\mathbf{G}$ at wavenumber j defined by

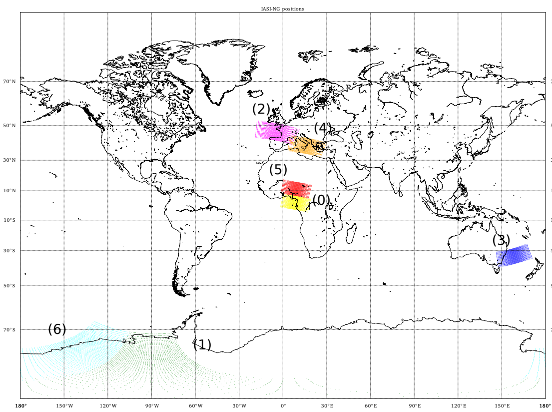


Fig. 8. Location of the seven IASI-NG USIs.

$$\mathrm{MSNR}_j(\mathbf{T}, \mathbf{G}) = 10 \log_{10} \left[\frac{\mathrm{median}(\mathbf{T}_j)^2}{\mathrm{MSE}(\mathbf{T}_j, \mathbf{G}_j)} \right], \tag{9}$$

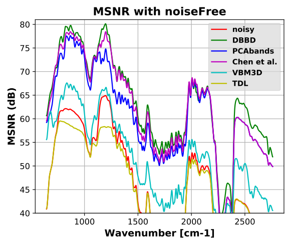
where MSE is the Mean Squared Error.

Fig. 9 plots the respective MSNRs of the noisy USI, the DBBD result, dimensionality reduction by PCA band by band, and Chen's method [20]. We display results for USIs #2, #3, #4, and #5. Similar results were obtained for all simulated USIs. We used the following parameters in the comparison: $N = 20$ in PCAbands, $N = 20$ PCs and 4 DT-CWT levels in Chen. For VBM3D, the parameters are T_2D=5, T_3D=7, N=8, Nf=4, Ns=7, Npr=5, Nb=2, k=8, kt=1, p=6, d=0.000754, lambda3D=2.700000, and tau=0.046136 (first step of the algorithm) and T_2D=4, T_3D=7, N=8, Nf=4, Ns=7, Npr=5, Nb=2, k=7, kt=1, p=4, d=0.000138, lambda3D=0, tau=0.023068 (second step). We refer the reader to [11, 41] for more details on the parameters. For TDL, the original source code sets the optimal values according to the level of noise. The IASI-NG USIs have 16920 channels and 2240 pixels/channel. VBM3D [11] improves slightly the noisy HSI, Chen's method [20] performs significantly well and beats the PCA band-by-band. DBBD (ours) outperforms all methods. TDL [17] has the worst performance, with a performance below the noisy HSI, confirming that TDL is not adapted to very-high SNR frequency-dependent HSIs.

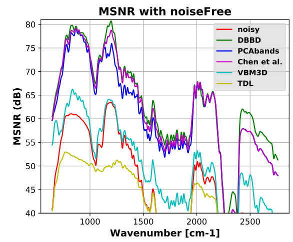
Given that neither TDL nor VBM3D can deal with frequency-dependent noise, we multiplied each channel by the proper factor in order to equalize the STD of the noise along all the spectrum, as required by these methods. Since exact noise level is known for each wavenumber (estimated noise (Sec. 2.1), simulated (Sec. 4.1), or tabulated empirically from the NEDT of the sensor (Eq. (1)), it is straighforward to perform this STD equalization.

The observed MSNR increase of the PCA result with respect to the noisy USI shows that PCA separates efficiently noise from the signal. Since the USIs were simulated, we dispose of an exact GT. It is therefore possible to examine the portion of signal that the denoising procedure actually removed. We express this in terms of *removed signal* and not *removed noise* because unless the denoising algorithm is *perfect*, it will remove some details along with the noise, as well as introducing some bias in the result.

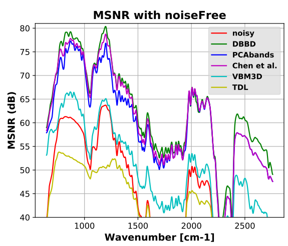
Since we know in advance that the noise is not correlated across different spectral channels, we can measure the Pearson auto-correlation of the removed signal ($\hat{\mathbf{G}} - \mathbf{G}$), and compute histograms of correlation values (with the exception of the diagonal of the correlation matrix). Once the



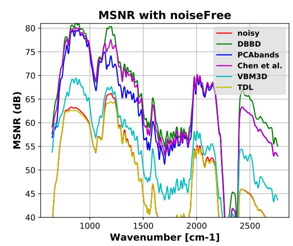
USI #2 (IASI-NG), MSNR



USI #3 (IASI-NG), MSNR



USI #4 (IASI-NG), MSNR



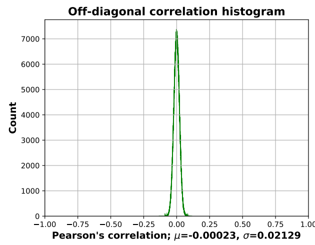
USI #5 (IASI-NG), MSNR

Fig. 9. MSNR denoising performances of simulated USIs #2, #3, #4, and #5. In red: the MSNR corresponding to the noisy USI. In green: the MSNR of our DBBD method. In blue: the denoising result obtained by simply applying PCA dimensionality reduction with $N = 20$ principal components band by band. Chen's method [20] in magenta, [11] in cyan, and TDL in olive color. Since the curves oscillate much, they have been filtered with a Gaussian of standard deviation $\sigma = 50$ for visualization and comparison purposes only. VBM3D [11] improves slightly the noisy HSI, Chen's method [20] performs significantly well and beats the PCA band-by-band. DBBD (ours) outperforms all methods. TDL [17] has the worst performance, with a performance below the noisy HSI, confirming that TDL is not adapted to very-high SNR frequency-dependent HSIs.

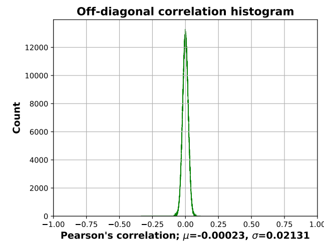
trivial ones at the diagonal have been excluded, two parameters of the histogram distribution need to be analyzed: its mean and its sharpness. Ideally, the distribution of correlation values should be a very sharp peak centered at zero (implying that there is no cross-correlation with other frequencies). Fig. 10 shows the histograms of the three USIs analyzed. As can be observed on the left, the cross-correlations of the signal removed by DBBD are centered at zero and their distribution is sharp and symmetric, which means that the method is not introducing artificial correlation in the denoised signal.

Table 2 compares the mean MSNR for the seven simulated IASI-NG USIs with respect to VBM3D [11], DBBD, Chen's method [20], PCA band by band, and TDL [17] and it shows that DBBD outperforms all other methods. Table 3 shows the evaluation with the SSIM [42] metric, where again DBBD outperforms the compared methods.

Table 4 gives our denoising factors. It is the factor by which the standard deviation of the noise has been divided for each of the simulated USIs. Reducing the standard deviation of the noise by the factor in (a) is theoretically equivalent to reducing the area of the sensor by the square of this factor while keeping the same SNR, or increasing the resolution of the sensor while keeping the same area. Column (c) gives the equivalent reduction in the area of the sensor.



USI #2 (IASI-NG) autocor.



USI #3 (IASI-NG) autocor.

Fig. 10. Histograms of the Pearson auto-correlation along different frequencies of the removed signal in USIs #2 and #3 (IASI-NG), band #3. As can be observed, the auto-correlations of the signal removed by DBBD are centered at zero and their distribution is sharp and symmetric. This gives strong evidence that the part of the signal that was removed by DBBD is indistinguishable from noise.

USI #	**Noisy**	**VBM3D [11]**	**DBBD**	**Chen [20]**	**PCA bands**	**TDL [17]**
#0	41.76	44.96	**59.39**	58.13	57.91	40.46
#1	33.64	41.00	**52.27**	50.38	50.25	33.31
#2	45.89	51.55	**63.44**	62.04	60.65	44.91
#3	44.83	49.88	**63.22**	61.84	60.82	41.92
#4	44.51	50.25	**62.46**	61.09	60.12	39.91
#5	47.63	53.69	**65.08**	63.69	62.35	47.47
#6	40.06	49.39	**58.50**	56.91	56.65	39.88
Mean	42.62	48.39	**60.62**	59.15	58.30	41.12

Table 2. Mean MSNR for the seven simulated IASI-NG USIs for VMB3D, DBBD, Chen's method, PCA band by band, and TDL. The column labeled as *Noisy* corresponds to the MSNR when no denoising is applied. The fact that TDL shows a MSNR lower than the noise itself must be understood as this method not being adapted to these kind of high-SNR frequency-dependent noise HSIs, despite of being one of the best methods available.

Fig. 11 shows the noise reduction before and after DBBD denoising in terms of the NEDT parameter. It can be observed that the amount of noise decreases significantly after denoising. We show results for only one of the seven IASI-NG USIs, but similar results are obtained with the others.

Our proposed method DBBD is specially adapted to hyperspectral images where the SNR is high and the number of channels is very large (several thousands), as in the case of the sensor of future satellites such as IASI-NG. To complete our study we have included a hypothetical scenario with a worse SNR and a reduced number of channels. We multiplied by 10 the STD of the noise and took one out of 70 wavenumbers (a total of 242). The aim of this is to have a quick review of these state-of-the-art methods in a scenario which is not adapted to DBBD. As can be observed in Table 5, in this scenario the difference between Chen and DBBD is not significant. The method by Chen should be preferred in this scenario, since it performed better in a majority of USIs. TDL shows competitive results, as expected. Indeed, when the noise is larger, it is easier to separate the noise from the signal, and if the number of channels is reduced TDL and other

USI #	Noisy	VBM3D [11]	DBBD	Chen [20]	PCA bands	TDL [17]
#0	22.73	23.47	**26.27**	26.12	25.75	22.48
#1	22.73	24.08	**26.19**	26.16	26.02	22.73
#2	22.73	23.85	**25.98**	25.81	24.92	22.27
#3	22.73	23.54	**26.06**	25.91	25.18	20.97
#4	22.73	23.75	**25.95**	25.83	24.93	20.94
#5	22.73	23.95	**26.04**	25.83	24.87	22.65
#6	22.73	24.14	**26.12**	26.11	25.89	22.72
Mean	22.73	23.80	**26.08**	25.96	25.26	21.80

Table 3. $|\log(1 - \mathrm{SSIM})|$ for the seven simulated IASI-NG USIs for VBM3D, DBBD, Chen, PCA band by band, and TDL. *Noisy* refers to the case when no denoising is applied. The evaluation of TDL shows again that certain methods (including VBM3D) are not adapted to high-SNR frequency-dependent HSIs, even if they can be considered the best methods for low and medium SNR HSIs.

USI #	(a)	(b)	(c)	(d)
#0	7.492	56.124	-5.513-e5	0.021
#1	7.331	53.740	-5.346-e5	0.021
#2	6.663	44.400	-5.365-e5	0.021
#3	6.966	48.530	-5.442-e5	0.021
#4	6.406	41.037	-5.429-e5	0.021
#5	6.778	45.943	-5.207-e5	0.021
#6	7.125	50.767	-5.318-e5	0.021

Table 4. (a) Standard deviation reduction factor after denoising with DBBD, (b) the associated surface reduction for the IASI-NG simulated USIs, (c) mean and (d) STD of the Pearson auto-correlation of $\tilde{\mathbf{G}} - \hat{\mathbf{G}}'_{\nu}$ (the removed signal) excluding the diagonal. The noise is simulated according to the tabulated NEDT of the sensor and applying Eq. (1) to obtain its realistic frequency-dependent STD.

dictionary-based methods start to be more adapted to denoise such signals. Nevertheless, notice that DBBD learns a basis for each group of similar pixels, and that the decay of eigenvalues for each basis is very sharp. Hence, DBBD can be assimilated to a local dictionary learning.

4.2. Evaluation with Real USIs

The evaluation of the denoising performance with real USIs can only be done through indirect methods such as the analysis of the correlations introduced (Sec. 4.1).

We used the five real USIs shown in Fig. 5 acquired by the IASI instrument [43]. We will refer to these real USIs as USI #7, . . . , #11. Each of the acquired images is 120×23 pixels with 8461 spectral channels. Since these are real USI images, there is no GT to compare to, but the autocorrelation of the removed signal can be analyzed. As explained in Sec. 4.1, we expect that the distribution of the Pearson auto-correlation coefficients out of the diagonal is

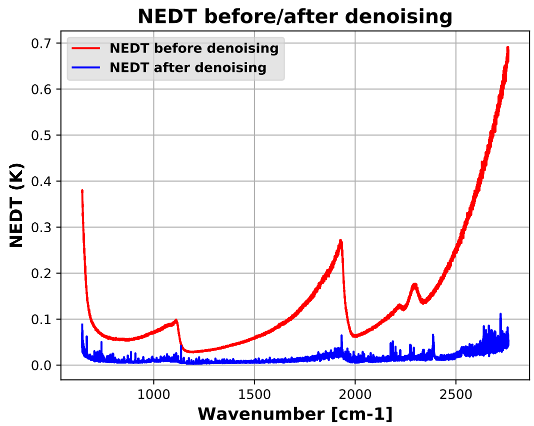


Fig. 11. Noise reduction before and after denoising with DBBD with IASI-NG USI #6. In red: the model's theoretical NEDT of the noise according to the physical model. In blue: the NEDT of the noise after DBBD denoising. The amount of noise has decreased significantly.

USI #	Noisy	VBM3D [11]	DBBD	Chen [20]	PCA bands	TDL [17]
#0	21.74	27.45	35.75	**35.90**	28.91	35.26
#1	13.83	24.71	**30.85**	29.53	25.31	29.47
#2	25.82	34.50	39.41	**39.85**	30.95	38.92
#3	24.79	32.51	39.58	**40.13**	29.86	38.26
#4	24.47	33.57	39.24	**39.35**	30.22	37.92
#5	27.67	37.18	40.53	**40.86**	30.72	40.81
#6	20.02	30.96	**36.56**	35.83	29.41	35.24
Mean	22.62	31.55	**37.42**	37.35	29.34	36.55

Table 5. Mean MSNR for the seven simulated IASI-NG USIs for VMB3D, DBBD, Chen's method, PCA band by band, and TDL in the case of low SNR and a reduced number of channels. We multiplied by 10 the STD of the noise and took one out of 70 wavenumbers (a total of 242). Now the difference between Chen and DBBD is not significant. The method by Chen should be preferred in this scenario, since it performed better in a majority of USIs. TDL shows competitive results, as expected.

centered at zero and with small variance (it should be zero in the case of an ideal perfect denoiser). The introduced correlation remains small even if groups of channels have similar underlying information. This is particularly important for remote sensing applications which can benefit from increased SNR, for example for retrieval of atmospheric and surface temperatures or ozone atmospheric profiles only if no bias has been introduced [22, 44].

Table 6 shows the statistics of the histograms of the Pearson cross-correlations coefficients along all frequencies of the removed signal in real USIs #7 . . . #11. For the real IASI data (8461 channels) Chen's method has values very similar to PCA-bands because with a reduced number of channels its thresholding simply set to zero the less significant PCs (as PCA does). The autocorrelation is better in DBBD. Note however that when the number of available channels is larger (as in the IASI-NG case, see Table 5, and also Tables 2 and 3), Chen clearly outperforms PCA. The goal of this experiment is to show that indeed these methods are valid to denoise ultraspectral images.

Statistics of the histograms of the Pearson cross-correlations coefficients along all frequencies of the removed signal in real USIs #7 . . . #11. For the real IASI data (8461 channels) Chen's method has values very similar to PCA-bands because with a reduced number of channels its thresholding simply set to zero the less significant PCs (as PCA does). The autocorrelation is better in DBBD. Note however that when the number of available channels is larger (as in the IASI-NG case, see Table 5, and also Tables 2 and 3), Chen clearly outperforms PCA. The goal of this experiment is to show that indeed these methods are valid to denoise ultraspectral images.

	Mean		STD	
USI #	**(a)**	**(b)**	**(c)**	**(d)**
#7	-2.21e-5	6.88e-5	3338.96e-5	3446.04e-5
#8	-4.80e-5	63.16e-5	3452.47e-5	5675.18e-5
#9	-5.77e-5	11.81e-5	3327.54e-5	3722.74e-5
#10	-3.77e-5	32.29e-5	3516.86e-5	6539.90e-5
#11	9.91e-5	-9.23e-5	3297.69e-5	2510.47e-5
Mean	**-1.33e-5**	20.98e-5	**3386.71e-5**	3386.71e-5

Table 6. Statistics of the histograms of the Pearson cross-correlations coefficients along all frequencies of the removed signal in real USIs #7 . . . #11. (a) means in DBBD, (b) means in PCA-bands and Chen, (c) STDs in DBBD, and (d) STDs in PCA-bands and Chen. For the real IASI data (8461 channels) Chen's method has values very similar to PCA-bands because with a reduced number of channels its thresholding simply set to zero the less significant PCs (as PCA does). The autocorrelation is better in DBBD. Note however that when the number of available channels is larger (as in the IASI-NG case, see Table 5, and also Tables 2 and 3), Chen clearly outperforms PCA. The goal of this experiment is to show that indeed these methods are valid to denoise ultraspectral images.

IASI measures wavenumbers from 645 cm^{-1} to 2760 cm^{-1} along 8461 channels. Therefore it has a spectral resolution of 0.25 cm^{-1}/ch. For example, the NO_2 molecule can be detected between 1290 and 1306.19 cm^{-1} (a bandwidth of 15.52 cm^{-1} or equivalently 51 IASI channels). If a denoising method happened to correlate these 51 wavenumbers, it could prevent the detection of NO_2 or produce false detection of other molecules.

5. Consequences for the Instrument's Resolution

In this paper we presented a new denoising method for hyperspectral and ultraspectral imaging based on the physical noise model given by Planck's law. We designed a noise estimation procedure adapted to real USIs and verified on simulated and real data that it retrieves the physical noise model. The noise estimation can favorably be used instead of the physical noise model as discussed in Sec. 2.1. It is a simple remedy for the insufficiency of the black body assumption, the differences between pre and post flight radiometric characterization, the signal saturation in the sensor, and finally the band overlapping. It is well known that a bad noise model leads to bad unmixing results [45]. Our experiments substantiate the usefulness of the physical model and, still better, of its real estimation. They not only allow to remove the exact amount of noise, but also help keep low the correlation introduced in the denoised signal. Finally, the method is *full-spectrum* in the sense that it denoises the totality of the spectrum [20] and not only high-energy PCA channels.

We did three types of evaluations. The MSNR and Structural Similarity (SSIM) metrics are important indicators of success when a GT is available. But we saw that the autocorrelation of the removed signal is perhaps a still more faithful indicator, actually more reliable than the MSNR or SSIM. While the MSNR and SSIM give crucial information on the denoising performance (for simulated data), a high MSNR does not guarantee that the removed noise is uncorrelated.

We showed that the performance of both TDL and VBM3D was poor in terms of the MSNR and SSIM metric. This must be understood in the context of the problem of high-SNR frequency-dependent noise in the case of ultraspectral imaging, for which these methods are not adapted. They can be considered the state of the art and clearly they are the best methods for data with a lower SNR and with a noise model which does not depend on the frequency, but they are simply not adapted to this new scenario. These methods do not take into account the fact that the variance of the noise depends on the frequency in their model, while in our method it is deeply integrated. Of course, one can always try to stabilize the variance with the Anscombe transform [46, 47] or other specific variance stabilizing transforms (VST), but the performance of the stabilization depends greatly on the size of the data and the introduced errors might be too large for the high-SNR case. In general, it is preferable to avoid the VST and integrate the fact that the noise depends on the frequency in the model itself. Other methods, such as Chen's and DBBD are more adapted to this kind of high-SNR USI data.

Yet one of the most important goals of hyperspectral imaging is the characterization of materials and their abundances. The performance of unmixing algorithms might be altered by any unduly introduced correlation in the restored signal. A significant advantage of the correlation measurement is that it can be measured in absence of GT. This has led us to demonstrate that the correlation introduced by our denoising method is very low on IASI real data.

The fact that DBBD does not introduce a significant correlation in the denoised signal and that MSNR and classification results after denoising is largely improved raises an important question in satellite design: is its resolution adequate to its goals? Indeed, in Table 4 we showed that for IASI-NG the standard deviation of the noise can be reduced by a factor of the order of 7. This means that it might be possible to reduce the surface of the sensor by a factor of the order of 50 while keeping the same SNR.

Thus, we propose to improve the current hyperspectral sensors by means of image denoising algorithms, with a special focus on the future IASI-NG to be launched in 2021. It is typical for hyperspectral sensors to have a large spectral resolution, but a very small one spatially. Getting a large enough SNR is the reason for the scarce spatial resolution. An increase of the SNR would allow to increase the spatial resolution delivered by the satellite, assuming the same sensor, or to design a new sensor whose surface is adapted to the image processing algorithms that will be applied. The resulting optical apparatus (and satellite) would therefore be much cheaper for an equal final resolution and SNR after processing. Alternatively, a much higher resolution apparatus would be obtained for the same overall cost. These comments apply retrospectively to the flying IASI and to future projects like IASI-NG. In short, we propose to move towards hybrid systems combining the optics of the instrument with the potential of image denoising algorithms.

In a future work, we seek to get *a-priori* information on the natural correlations introduced by the instrument itself in the spectral channels, by analyzing the data acquired at the same zones along different satellite passes. This information would be added to the noise model and help measure the correlation introduced, thus further improving the results.

Acknowledgement: Centre National d'Etudes Spatiales (CNES, R&T action) for the access to IASI and IASI-NG data and advice, EUMETSAT and NOVELTIS for the simulated data, and the European Research Council (Advanced Grant *Twelve Labours*), Office of Naval Research (Grant N00014-97-1-0839), and Direction Générale de l'Armement (DGA) for their funding of the project.

References

1. R. N. Greenberger, J. F. Mustard, B. L. Ehlmann, D. L. Blaney, E. A. Cloutis, J. H. Wilson, R. O. Green, and A. A. Fraeman, "Imaging spectroscopy of geological samples and outcrops: Novel insights from microns to meters," GSA Today **25**, 4–10 (2015).
2. C. Crevoisier, C. Clerbaux, V. Guidard, T. Phulpin, R. Armante, B. Barret, C.-P. C., J. P. Chaboureau, P. F. Coheur, L. Crépeau *et al.*, "Towards IASI-new generation (IASI-NG): impact of improved spectral resolution and radiometric noise on the retrieval of thermodynamic, chemistry and climate variables," Atmospheric Meas. Tech. **7**, 4367–4385 (2014).
3. B. S. Sorg, B. J. Moeller, O. Donovan, Y. Cao, and M. W. Dewhirst, "Hyperspectral imaging of hemoglobin saturation in tumor microvasculature and tumor hypoxia development," J. biomedical optics **10**, 044004–044004 (2005).
4. R. Lu and Y.-R. Chen, "Hyperspectral imaging for safety inspection of food and agricultural products," in *Photonics East (ISAM, VVDC, IEMB),* (International Society for Optics and Photonics, 1999), pp. 121–133.
5. N. Keshava, "A survey of spectral unmixing algorithms," Linc. Lab. J. **14**, 55–78 (2003).
6. M. Colom, G. Blanchet, A. Klonecki, O. Lezeaux, E. Pequignot, F. Poustomis, C. Thiebaut, S. Ythier, and J.-M. Morel, "BBD: a new bayesian bi-clustering denoising algorithm for IASI-NG hyperspectral images," in *8th workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing,* (IEEE, Los Angeles, USA, 2016), pp. –.
7. A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri *et al.*, "Recent advances in techniques for hyperspectral image processing," Remote. sensing environment **113**, S110–S122 (2009).
8. A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gaultieri *et al.*, "Advanced processing of hyperspectral images," in *2006 IEEE International Symposium on Geoscience and Remote Sensing,* (IEEE, 2006), pp. 1974–1978.
9. Y.-Q. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," Geosci. Remote. Sensing, IEEE Transactions on **53**, 296–308 (2015).
10. Y. Zhao and J. Yang, "Hyperspectral image denoising via sparsity and low rank," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International,* (IEEE, 2013), pp. 1091–1094.
11. D. Kostadin, F. Alessandro, and E. Karen, "Video denoising by sparse 3d transform-domain collaborative filtering," in *European signal processing conference,* vol. 149 (Tampere, Finland, 2007).
12. B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using 3d wavelets," in *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International,* (IEEE, 2012), pp. 1349–1352.
13. H. K. Aggarwal and A. Majumdar, "Mixed gaussian and impulse denoising of hyperspectral images," in *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International,* (IEEE, 2015), pp. 429–432.
14. H. K. Aggarwal and A. Majumdar, "Hyperspectral image denoising using spatio-spectral total variation," IEEE Geosci. Remote. Sens. Lett. **13**, 442–446 (2016).
15. Y. Fu, A. Lam, I. Sato, and Y. Sato, "Adaptive spatial-spectral dictionary learning for hyperspectral image denoising," in *Proceedings of the IEEE International Conference on Computer Vision,* (2015), pp. 343–351.
16. M. Ye, Y. Qian, and Q. Wang, "Mt-omp for hyperspectral imagery denoising with model parameter estimation," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International,* (IEEE, 2013), pp. 1079–1082.
17. Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* (2014), pp. 2949–2956.
18. A. Karami, R. Heylen, and P. Scheunders, "Band-specific shearlet-based hyperspectral image noise reduction," IEEE Transactions on Geosci. Remote. Sens. **53**, 5054–5066 (2015).
19. A. Karami, R. Heylen, and P. Scheunders, "Denoising of hyperspectral images using shearlet transform and fully constrained least squares unmixing," in *8th workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing,* (Los Angeles, USA, 2016).
20. G. Chen and S.-E. Qian, "Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage," Geosci. Remote. Sensing, IEEE Transactions on **49**, 973–980 (2011).
21. M. Planck, "Ueber das gesetz der energieverteilung im normalspectrum," Annalen der Physik **309**, 553–563 (1901).
22. F. Aires, A. Chedin, N. A. Scott, and W. B. Rossow, "A regularized neural net approach for retrieval of atmospheric and surface temperatures with the IASI instrument," J. Appl. Meteorol. **41**, 144–159 (2002).
23. F. Cayla, B. Tournier, and P. Hebert, "Performance budgets of iasi options," Tech. rep., Tech. Rep. IA-TN-0B-5476-CNE, 21 pp.[Available from Centre National d'Études Spatiales, 2 place Maurice Quentin, 75039 Paris Cedex 01, France.] (1995).
24. M. Jiméneza, R. Díaz-Delgadob, P. Vaughana, A. De Santisc, A. Fernández-Renaua, E. Pradoa, and O. G. de la Cámaraa, "Airborne hyperspectral scanner (ahs) mapping capacity simulation for the doñana biological reserve scrublands," in *Proceedings of the 10th International Symposium on Physical Measurements and Signatures in Remote Sensing, Davos, Switzerland,* (2007), pp. 12–14.
25. C. Cao and A. K. Heidinger, "Inter-comparison of the longwave infrared channels of modis and avhrr/noaa-16 using simultaneous nadir observations at orbit intersections," in *Earth Observing Systems VII,* vol. 4814 (International Society for Optics and Photonics, 2002), pp. 306–317.
26. C. Cao, M. Weinreb, and J. Sullivan, "Solar contamination effects on the infrared channels of the advanced very high

resolution radiometer (avhrr)," J. Geophys. Res. Atmospheres **106**, 33463–33469 (2001).

27. S. B. Campana, J. S. Accetta, and D. L. Shumaker, *The infrared and electro-optical systems handbook(Volume 5, Passive electro-optical systems)* (Infrared Information Analysis Center, 1993).
28. J. Rice, S. W. Brown, B. Johnson, and J. Neira, "Hyperspectral image projectors for radiometric applications," Metrologia **43**, S61 (2006).
29. A. Foi, "Clipped noisy images: Heteroskedastic modeling and practical denoising," Signal Process. **89**, 2609–2629 (2009).
30. M. Farzam, S. Beheshti, and M. Hashemi, "Adaptive noise variance estimation and intrinsic order selection for low snr hyperspectral signals," Electr. Comput. Eng. Can. J. **36**, 4–10 (2013).
31. G. Grieco, C. Serio, and G. Masiello, "σ-IASI-β: a hyperfast radiative transfer code to retrieve surface and atmospheric geophysical parameters," EARSeL eProceedings **12**, 149 (2013).
32. J.-S. Lee, "Digital image smoothing and the sigma filter," Comput. vision, graphics, image processing **24**, 255–269 (1983).
33. M. Lebrun, M. Colom, A. Buades, and J. M. Morel, "Secrets of image denoising cuisine," Acta Numer. **21**, 475–576 (2012).
34. A. D. Collard, A. P. McNally, F. I. Hilton, S. B. Healy, and N. C. Atkinson, "The use of principal component analysis for the assimilation of high-resolution infrared sounder observations for numerical weather prediction," Q. J. Royal Meteorol. Soc. **136**, 2038–2050 (2010).
35. A. Lam, I. Sato, and Y. Sato, "Denoising hyperspectral images using spectral domain statistics," in *Pattern Recognition (ICPR), 2012 21st International Conference on,* (IEEE, 2012), pp. 477–480.
36. M. Lebrun, A. Buades, and J.-M. Morel, "A nonlocal Bayesian image denoising algorithm," SIAM J. on Imaging Sci. **6**, 1665–1688 (2013).
37. S. Pyatykh, J. Hesser, and L. Zheng, "Image noise level estimation by principal component analysis," IEEE transactions on image processing **22**, 687–699 (2013).
38. J.-S. Lee, "Refined filtering of image noise using local statistics," Comput. graphics image processing **15**, 380–389 (1981).
39. F. Bernard, B. Calvel, F. Pasternak, R. Davancens, C. Buil, E. Baldit, C. Luitot, and A. Penquer, "Overview of IASI-NG the new generation of infrared atmospheric sounder," in *International Conference on Space Optics,* vol. 7 (2014), p. 10.
40. C. Crevoisier, C. Clerbaux, V. Guidard, E. Pequignot, and F. Pasternak, "IASI-NG: un concentré d'innovations technologiques pour l'étude de l'atmosphère terrestre," La Météorologie **2014**, 3–5 (2014).
41. M. Lebrun, "An Analysis and Implementation of the BM3D Image Denoising Method," Image Processing On Line **2**, 175–213 (2012).
42. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE transactions on image processing **13**, 600–612 (2004).
43. EUMETSAT, "IASI level 1: Product guide," http://www.eumetsat.int/website/home/Data/Products/Level1Data/index.html (2014). [Online; accessed 30-Mai-2016].
44. F. Aires, W. Rossow, N. Scott, and A. Chédin, "Remote sensing from the infrared atmospheric sounding interferometer instrument 2. simultaneous retrieval of temperature, water vapor, and ozone atmospheric profiles," J. Geophys. Res. Atmospheres **107** (2002).
45. K. Cawse-Nicholson, A. Robin, and M. Sears, "The effect of spectrally correlated noise on noise estimation methods for hyperspectral images," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2012 4th Workshop on,* (IEEE, 2012), pp. 1–4.
46. F. J. Anscombe, "The transformation of poisson, binomial and negative-binomial data," Biometrika **35**, 246–254 (1948).
47. M. Makitalo and A. Foi, "Optimal inversion of the anscombe transformation in low-count poisson image denoising," IEEE transactions on Image Process. **20**, 99–109 (2011).