

Progressive compressed sensing and reconstruction of multidimensional signals using hybrid transform/prediction sparsity model

Giulio Coluccia, *Member, IEEE*, Simeon Kamdem Kuiteing, Andrea Abrardo, *Senior Member, IEEE*, Mauro Barni, *Fellow, IEEE*, and Enrico Magli, *Senior Member, IEEE*

Abstract—Compressed sensing (CS) is an innovative technique allowing to represent signals through a small number of their linear projections. Hence, CS can be thought of as a natural candidate for acquisition of multidimensional signals, as the amount of data acquired and processed by conventional sensors could create problems in terms of computational complexity. In this paper we propose a framework for the acquisition and reconstruction of multidimensional correlated signals. The approach is general and can be applied to D dimensional signals, even if the algorithms we propose to practically implement such architectures apply to 2D and 3D signals. The proposed architectures employ iterative local signal reconstruction based on a hybrid transform/prediction correlation model, coupled with a proper initialization strategy.

Index Terms—Compressed Sensing, Multidimensional Signals, Linear Predictor, Image Scanning, Remote Sensing, Hyperspectral Imaging.

I. INTRODUCTION

Compressed Sensing (CS) [1], [2] has recently emerged as an efficient technique for sampling a signal with fewer coefficients than the number dictated by classical Shannon/Nyquist theory. The assumption underlying this approach is that the signal to be sampled is sparse or at least “compressible”, i.e., it must have a concise representation in a convenient basis. In CS, sampling is performed by taking a number of linear projections of the signal onto pseudorandom sequences. Therefore, the acquisition presents appealing properties such as low encoding complexity, since the basis in

which the signal is sparse does not need to be computed, and universality, since the sensing is blind to the source distribution. Reconstruction of a signal from its projections can be done e.g. using linear programming [2], with a complexity that is $O(N^3)$, with N the number of samples to be recovered.

Plenty of applications are possible, ranging from image and video to biomedical and spectral imaging, just to mention a few. A single-pixel camera has been demonstrated in [3], [4], which uses a single detector to sequentially acquire random linear measurements of a scene. This kind of design is very interesting for imaging at wavelength outside the visible light, where manufacturing detectors is very expensive. CS could be used to design cheaper sensors, or sensors providing better resolution for an equal number of detectors. E.g., in [5] an architecture is proposed based on Hadamard imaging, coupled with reconstruction techniques borrowed from CS.

The acquisition of multidimensional signals could benefit from CS due to its low-complexity sampling process and the reduction of the number of samples to be taken, processed and transmitted. In this case, a serious problem arises regarding the computational complexity of the reconstruction process. The conventional approach of measuring the signal along all dimensions at once leads to very large N , making the reconstruction computationally intractable. The simplest solution is to take separate sets of measurements grouping subsets of dimensions and to perform separate reconstructions. For example, an hyperspectral image could be acquired in the spatial or spectral dimensions. However, this “separate” approach does not yield satisfactory performance in terms of mean-squared error (MSE), as it neglects the overall correlation among every dimension. In the example above, the spatial CS approach completely neglects the spectral correlation, and the spectral approach neglects the spatial one.

The authors of [6] showed a way to recast a multidimensional CS problem to a one-dimensional one, by the

G. Coluccia and E. Magli are with Politecnico di Torino (Dipartimento di Elettronica e Telecomunicazioni), c.so Duca degli Abruzzi, 24 - 10129 Torino, Italy (email: giulio.coluccia@polito.it, enrico.magli@polito.it). Their research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement n 279848

S. Kamdem Kuiteing, A. Abrardo and M. Barni are with University of Siena (Dipartimento di Ingegneria dell'Informazione), via Roma, 56 - 53100 Siena, Italy (email: barni@dii.unisi.it, abrardo.dii.unisi.it, skamdemk@yahoo.fr).

means of Kronecker products of sensing and sparsity matrices. The problem of this approach is that the dimensionality of the CS problem to be solved rapidly grows as the product of the sizes of each dimension. The authors of [7] apply CS to blocks of images considering wavelets as sparsity basis. Reconstruction algorithms for multidimensional signals have also been proposed in [8], [9], [10] for hyperspectral images and multiview video.

In this paper, we propose a generic framework for CS acquisition and reconstruction of multidimensional signals. Then, we propose architectures for 2D and 3D signals. The framework is general since the principles behind the architectures we propose can be easily extended to D -dimensional signals, with arbitrary D . Moreover, the proposed architectures can be practically implemented following the algorithms we devise for each architecture. The architectures we propose can be applied to several scenarios we describe in the following paragraphs.

The first architecture describes a practical implementation for devices that acquire 2D visual information through progressive scanning [11]. These devices are equipped with a one-dimensional array of detectors, and a 2D image is obtained via the repeated use of the array over different slices of the 2D object to be imaged. This is a very important scenario, which encompasses many applications. Amongst others, it is worth mentioning at least two examples, which we will focus on in the respective sections of this paper. The first is given by *flatbed scanners*, where each line of the image is acquired by a 1D optical sensor moving in the orthogonal direction. The second one is represented by *airborne and spaceborne imagers* of the pushbroom type for remote sensing applications. In this case, the 1D sensor is carried on a flying platform such as an airplane or satellite; the sensor looks down at the Earth, and acquires a line-by-line scan of the underlying scene, while each line is oriented in the across-track direction, and the platform flight moves the sensor from one line to the next one. These applications, as well as several other ones, can clearly benefit from CS. Devices similar in principle to the Single Pixel Camera can be applied to progressive scanning, where a 1D micromirror array can be used to directly sense lines in the CS format. In the case of the remote sensing imaging system, CS can lead to a simpler and cheaper system, which uses a single detector and produces a reduced number of sampling. Detectors can be costly in the wavelengths outside the visible spectrum, and the reduced number of samples allows to implement simpler onboard processing systems. For the flatbed scanner, CS would be extremely useful in order to develop a scanner of small size, as the CS sensor needs

not be of the same physical size as the document being scanned. Moreover, in both cases, processing and data handling would be greatly reduced, which is important in order to reduce power consumption in the remote sensing case, and in order to enable application to small-sized low-power devices in the flatbed scanner case.

The second scenario is the acquisition of hyperspectral images. Satellite imaging is a highly effective tool in a variety of scientific and engineering contexts because of the information it provides about the nature of the materials being imaged. While traditional digital imaging techniques produce images with scalar values associated with each pixel location, in multi- and hyperspectral images these values are replaced with a vector containing the spectral information associated to that spatial location. The resulting image is therefore three-dimensional (two spatial and one spectral dimensions), and spectral resolution is very important for several applications, including classification, anomaly detection, and spectral unmixing. Despite the huge potential, however, many modern satellite imagers face a limiting trade-off between spatial and spectral resolution. In fact, the total number of samples that can be acquired is constrained by the size of the detector array. This limits the usefulness and cost-effectiveness of spectral imaging for many applications. This scenario intends to investigate the possibility of overcoming this limitation by means of a new imaging architecture based on CS, that is, an architecture in which the acquisition system does not detect single pixels of the scene, but rather a small number of measurements. Reconstruction of the image is going to be performed at the ground station, and all subsequent processing steps (radiometric and geometric calibration, orthorectification, and applications) would be performed on the reconstructed image. In this paper we consider several possible architectures that are based on the constraints imposed by real-world spectral imaging systems. A first architecture is derived from 2D imaging systems, and is based on the concept of a sensor that acquires the image as a whole, both spatially and spectrally. In this case, the single-pixel camera paradigm is still applicable, provided that this single pixel is actually a single-pixel spectral imaging device, i.e. one that separates the spectral components of the incoming light into several different wavelengths, in such a way that the integration of spatially modulated light can be done individually in each spectral channel, yielding a different set of linear measurements for each wavelength. The second architecture mimics spectral imagers of the pushbroom type. These scanners have a 2D sensor, but the two dimensions correspond to the spectral (wavelength) dimension and one spatial dimension, i.e. the

direction perpendicular to the flight path. At a given flight position, the scanner acquires at once one spatial-spectral slice, i.e. one line of the image with all its spectral components; as the satellite or airplane moves one, another spatial-spectral slice is acquired, and this process is repeated until the desired image length has been obtained. In this paper we also consider a CS imaging architecture based on this paradigm. We assume that the single-pixel concept is applied to spectral slices, i.e. there is a diffractive element before the single pixel, so that the spatial-spectral slice can be modulated before being sensed by the single detector. It is worth noticing that these architectures would yield significant benefits to hyperspectral imaging systems. They would allow to reduce the cost and size of the sensor itself, for the same spatial and spectral resolution, thanks to the reduced number of detectors. Moreover, they would allow to greatly simplify onboard data handling, since the compressed acquisition process would generate much less data than a conventional system, allowing to employ less memory and computing power, eliminating the need of onboard compression altogether, and eventually leading to reduced power consumption, which is a critical aspect of any remote sensing mission. Despite the very appealing advantages, however, these acquisition architectures entail a reconstruction problem of huge size. Indeed, optimal reconstruction must exploit signal sparsity (and hence correlation) in all dimensions, requiring to solve the reconstruction problem at once employing a three-dimensional transform as sparsity model. This problem has huge computational complexity, and becomes infeasible for rather small image sizes, highlighting the need of techniques that can achieve near-optimal performance with a reasonable computational complexity.

In this paper we address these scenarios, and tackle the reconstruction problem for 2D and 3D signals. In particular, for 2D images, we propose a simple progressive acquisition algorithm, where rows are acquired independently of each other, but the reconstruction is performed jointly over all rows. Joint reconstruction is achieved through an iterative algorithm that correlates different rows through linear prediction filters, instead of taking a multidimensional transform as sparsity domain. Prediction filters allow to exploit correlation in both horizontal and vertical dimensions, even if the acquisition is performed in one direction only. The main concept is to exploit correlation along the vertical direction by iteratively predicting each line and reconstructing the prediction error only, which is more compressible than the line itself. Results show that few iterations of the proposed algorithm suffice to significantly improve the MSE of the reconstruction, allowing to obtain high-quality

reconstruction results with feasible complexity. On the other hand, for 3D images we have more degrees of freedom than in the 2D case. For example, in [12] it has been shown that 2D spatial CS (i.e., every spectral channel is measured independently) has better performance than spectral CS (in which every spectral vector is measured independently), just because the former approach models correlation in two dimensions, and the latter in only one. However, it should be noted that even spatial CS achieves an MSE that is not small enough for many hyperspectral applications, as the relative error is around $\pm 5\%$ for sensible values of the number of acquired samples. The key idea is that, in order to improve reconstruction quality, correlation must be exploited in all three dimensions of the spectral cube. To achieve this goal, we propose several approaches, combining an accurate modelling of the spatial-spectral correlations, with the low complexity of sequential, as opposed to fully joint, reconstruction. In particular, instead of modelling the correlation by means of a three-dimensional transform, and hence attempting to reconstruct the hyperspectral cube as a whole, we employ a linear correlation model of the hyperspectral image, and iteratively apply this model band by band, improving the quality of the reconstructed image. An alternative algorithm applies the model to the spectral rows of the image, iterating along rows.

Since the quality of the reconstructed signal depends on two factors: *i*) the initialization of the iterative procedure and *ii*) the accuracy of the linear prediction filters, we consider different initialization strategies based either on a 2D CS approach or on a simplified 3D strategy [13] and test different prediction filters looking for the one providing better performance.

This paper is organized as follows. Section II contains the background of this work, notations and definitions. In section III we describe in detail the algorithms we propose. In section IV we show some results obtained by simulations and we conclude our work in section V.

II. BACKGROUND

A. Notation and definitions

We denote (column-) vectors and matrices by lowercase and uppercase boldface characters, respectively. The (m, n) -th element of a matrix \mathbf{A} is $(\mathbf{A})_{m,n}$. The m -th row of matrix \mathbf{A} is $(\mathbf{A})_m$. The n -th element of a vector \mathbf{v} is $(\mathbf{v})_n$. The transpose of a matrix \mathbf{A} is \mathbf{A}^T .

The stack operator $\text{vec}\{\mathbf{A}\}$ denotes the column vector obtained by stacking the columns of \mathbf{A} on top of each other, from left to right.

We denote 3D variables by calligraphic letters, e.g. \mathcal{Q} . $\mathbf{X} = \mathcal{Q}_{i,:}$ is the matrix obtained by fixing index i

in the first dimension of \mathcal{Q} . $\mathbf{Y} = \mathcal{Q}_{:,j,:}$ is the matrix obtained by fixing index j in the second dimension of \mathcal{Q} . $\mathbf{Z} = \mathcal{Q}_{:,:,k}$ is the matrix obtained by fixing index k in the third dimension of \mathcal{Q} .

The notation $\|\mathbf{v}\|_{\ell_0}$ denotes the number of nonzero elements of vector \mathbf{v} . The notation $\|\mathbf{v}\|_{\ell_1}$ denotes the ℓ_1 -norm of the vector \mathbf{v} and is defined as $\|\mathbf{v}\|_{\ell_1} \triangleq \sum_i |(\mathbf{v})_i|$. The notation $\|\mathbf{v}\|_{\ell_2}$ denotes the Euclidean norm of the vector \mathbf{v} and is defined as $\|\mathbf{v}\|_{\ell_2} \triangleq \sqrt{\sum_i |(\mathbf{v})_i|^2}$. The notation $a \sim \mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian random variable a with mean μ and variance σ^2 .

The notation $\mathbf{A} \otimes \mathbf{B} = [(\mathbf{A})_{ij} \mathbf{B}]$ (written in block matrix form, where the pair i, j spans the range of indexes of \mathbf{A}) denotes the Kronecker product of \mathbf{A} times \mathbf{B} .

B. Compressed Sensing

In the standard CS framework, introduced in [14], a signal $\mathbf{x} \in \mathbb{R}^{N \times 1}$ which has a sparse representation in some basis $\Psi \in \mathbb{R}^{N \times N}$, i.e:

$$\mathbf{x} = \Psi \boldsymbol{\theta}, \quad \|\boldsymbol{\theta}\|_{\ell_0} = K, \quad K \ll N$$

can be recovered by a smaller vector $\mathbf{y} \in \mathbb{R}^{M \times 1}$, $K < M < N$, of linear measurements $\mathbf{y} = \Phi \mathbf{x}$, where $\Phi \in \mathbb{R}^{M \times N}$ is the *sensing matrix*. The optimum solution, requiring at least $M = K + 1$ measurements, would be

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_{\ell_0} \quad \text{s.t.} \quad \Phi \Psi \boldsymbol{\theta} = \mathbf{y}.$$

Since the ℓ_0 norm minimization is a NP-hard problem, one can resort to a linear programming reconstruction by minimizing the ℓ_1 norm

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_{\ell_1} \quad \text{s.t.} \quad \Phi \Psi \boldsymbol{\theta} = \mathbf{y}, \quad (1)$$

provided that M is large enough ($\sim K \log(N/K)$).

The same algorithm holds for signals which are not exactly sparse, but rather compressible, meaning that they (or their representation $\boldsymbol{\theta}$ in basis Ψ) can be expressed only by K significant coefficients, while the remaining ones are (close to) zero.

When the measurements are noisy the ℓ_1 minimization with relaxed constraints is used for reconstruction:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_{\ell_1} \quad \text{s.t.} \quad \|\Phi \Psi \boldsymbol{\theta} - \mathbf{y}\|_{\ell_2} < \varepsilon, \quad (2)$$

where ε bounds the amount of noise in the data. It has been shown in [15] that extracting the elements of Φ at random from a Gaussian or Rademacher distribution (i.e., ± 1 with the same probability), and, in general, from any Sub-Gaussian distribution, allows a correct reconstruction with overwhelming probability.

III. PROPOSED ARCHITECTURES

In this paper, we propose a general framework for multidimensional signals, allowing to exploit the low complexity and universality of CS in the acquisition process, with a manageable complexity of the reconstruction algorithm. Refer to Figure 1, which depicts the canonical approach of considering the multidimensional signal as a single signal (1(a)) and the architecture proposed in this paper (1(b)), based on progressive scanning and iterative reconstruction. Multidimensional signals are often captured in a progressive way, in a sequence of acquisitions corresponding to subsets of the coordinates. The principle is to acquire separately each signal dimension (or subsets of dimensions, considering them as single signals). Then, instead of reconstructing the whole set of measurements at once, as done with Kronecker Compressed Sensing [13], an iterative algorithm is applied to dimensions not involved in the measurement process. At each iteration, a linear prediction filter is used to aid the reconstruction process, measuring the prediction and reconstructing the prediction error only, which is supposed to be more compressible than the original signal. For example, if an hyperspectral image is acquired band by band, the iterative algorithm is applied on wavelength dimension. At each iteration, each band is predicted, acquired with the same sensing matrix used to acquire that band, the measurement of the predicted band is subtracted from the measurement of the band itself and the CS reconstruction is applied to this “measurement prediction error”, only.

In the following, we specialize the approach to 2D and 3D signals. We propose novel architectures for 2D and 3D acquisition and reconstruction, and the corresponding algorithms implementing the proposed architectures.

A. 2D signals

According to typical progressive scanning approaches, like the ones used by commercial flatbed scanners or by remote sensing systems acquiring environmental pictures, an image is acquired by sensing N_{COL} pixels of each row in a progressive fashion, until N_{ROW} rows are acquired. Hence, the acquired image will result as a matrix of pixels of size $N_{\text{ROW}} \times N_{\text{COL}}$, which will be compressed (and, accordingly, decoded) using a conventional technique. This process requires the acquisition (and processing) of $N_{\text{ROW}} N_{\text{COL}}$ pixels. When N_{ROW} and N_{COL} are large, processing of this huge amount of data may represent an issue, especially when dealing with low cost or low complexity devices.

For this reason, we propose a very simple acquisition scheme, based on CS linear measurements taken on

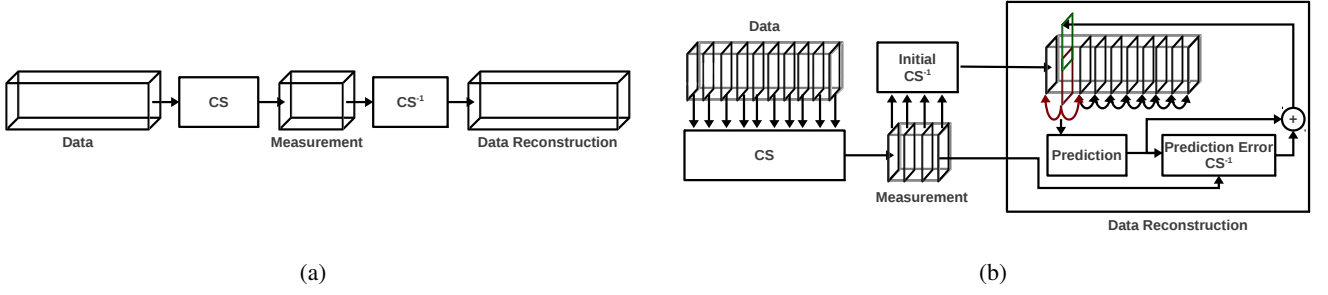


Fig. 1. Block Diagram of an architecture processing the whole signal (a) and of the novel iterative architecture (b).

Algorithm 1 Acquisition algorithm for 2D signals

Require: the image \mathbf{X} , M

Ensure: the measurement matrix \mathbf{Y}

- 1: **for** $i = 1$ to N_{ROW} **do**
 - 2: Draw Φ^i of size $M \times N_{\text{COL}}$ s.t. $(\Phi^i)_{kj} \sim \mathcal{N}(0, 1/M)$
 - 3: $(\mathbf{Y})_i^\top \leftarrow \Phi^i (\mathbf{X})_i^\top$
 - 4: **end for**
 - 5: **return** \mathbf{Y}
-

each row, without any further processing. This reduces the amount of data to be acquired and processed. The reconstruction algorithm relies on linear prediction filters in order to improve the quality of CS reconstruction, by correlating the measurements of adjacent rows in order to exploit their statistical dependencies during the reconstruction stage, largely improving over individual separate reconstruction.

1) *Acquisition:* The image acquisition algorithm we propose, labelled as Algorithm 1, is very simple and consists in taking linear measurements of each row of the image in a progressive fashion. To minimize the risks of failures in the reconstruction side, a different sensing matrix Φ is drawn for each row.

The image to be measured can be divided into N_{ROW} rows. For each row, M linear measurements are taken, where $M < N_{\text{COL}}$ and N_{COL} is the desired vertical resolution.

In summary, the scene we wish to acquire is represented by the matrix $\mathbf{X} \in \mathbb{R}^{N_{\text{ROW}} \times N_{\text{COL}}}$. For each row of \mathbf{X} , we draw a matrix $\Phi^i \in \mathbb{R}^{M \times N_{\text{COL}}}$ whose elements are Gaussian i.i.d. such that $(\Phi^i)_{kj} \sim \mathcal{N}(0, 1/M)$, with $k = 1, \dots, M$ and $j = 1, \dots, N_{\text{COL}}$. Then, we take M linear measurements of $(\mathbf{X})_i$ which will form the rows of the matrix of measurements $\mathbf{Y} \in \mathbb{R}^{N_{\text{ROW}} \times M}$, namely

$$(\mathbf{Y})_i^\top = \Phi^i (\mathbf{X})_i^\top$$

A more complex algorithm, based on Compressed Sensing and able to capture spatial correlation in both

directions (horizontal and vertical), could acquire in a single shot the whole image in a single measurement vector of length M' .

$$\mathbf{y}' = \Phi' \text{vec}\{\mathbf{X}\},$$

where $\text{vec}\{\mathbf{X}\} \in \mathbb{R}^{N_{\text{ROW}} N_{\text{COL}} \times 1}$, $\Phi' \in \mathbb{R}^{M' \times N_{\text{ROW}} N_{\text{COL}}}$, $\mathbf{y}' \in \mathbb{R}^{M' \times 1}$.

Even if this algorithm performed better than the one proposed here since the reconstruction would optimally exploit the correlation in 2 dimensions through a 2D transform matrix, it would require the solution of (1) for a vector of length $N = N_{\text{ROW}} N_{\text{COL}}$. For realistic values of N_{ROW} and N_{COL} , the solution of (1) would be impossible to perform in reasonable time. On the other hand, the proposed approach splits the problem into smaller (and hence tractable) subproblems. However, in doing so, it does not neglect the spatial correlation in vertical direction, which is modelled and employed in the reconstruction process through the use of linear prediction filters.

2) *Reconstruction:* A trivial reconstruction algorithm based on the acquisition scheme described in section III-A1 would simply apply the ℓ_1 reconstruction (1) to recover separately each line of \mathbf{X} given the corresponding Φ^i and $(\mathbf{Y})_i$.

Instead, we propose an algorithm which iteratively improves the current estimate of \mathbf{X} by modelling statistical dependencies between adjacent lines. We label this Algorithm 2. We count the iterations using the index n . The estimation of \mathbf{X} at iteration n is denoted with $\mathbf{X}^{(n)}$.

In particular, the algorithm evaluates a first image reconstruction using some initialization strategy $F(\mathbf{Y}, \Phi)$ (iteration $n = 0$). Then, the iterations start. The intuition is as follows. For each row, if we are able to reliably predict it using the reconstruction of the upper and lower lines at previous iteration with some linear prediction filter $P(\cdot, \cdot)$, obtaining \mathbf{x}_p , we can compute the “measurement” \mathbf{y}_p of this prediction by applying matrix Φ^i to \mathbf{x}_p . Then we calculate the prediction error in the linear measurement domain \mathbf{e}_y by subtracting this

Algorithm 2 Reconstruction algorithm for 2D signals**Require:** the measurement matrix \mathbf{Y} , the set of Φ^i **Ensure:** the estimation $\hat{\mathbf{X}}$

```

1:  $n \leftarrow 0$ 
2:  $\mathbf{X}^{(n)} \leftarrow F(\mathbf{Y}, \Phi)$ 
3: repeat
4:    $n \leftarrow n + 1$ 
5:   for  $i = 1$  to  $N_{\text{ROW}}$  do
6:     if  $i = 1$  or  $i = N_{\text{ROW}}$  then
7:        $\mathbf{x}_P \leftarrow (\mathbf{X}^{(n-1)})_i^T$ 
8:     else
9:        $\mathbf{x}_P \leftarrow P\left(\left((\mathbf{X}^{(n-1)})_{i-1}, (\mathbf{X}^{(n-1)})_{i+1}\right)^T\right)$ 
10:    end if
11:     $\mathbf{y}_P \leftarrow \Phi^i \mathbf{x}_P$ 
12:     $\mathbf{e}_Y \leftarrow (\mathbf{Y})_i^T - \mathbf{y}_P$ 
13:     $\mathbf{e}_\theta \leftarrow \arg \min_{\mathbf{e}} \|\mathbf{e}\|_{\ell_1} \quad \text{s.t.} \quad \Phi^i \Psi \mathbf{e} = \mathbf{e}_Y$ 
14:     $\mathbf{e}_X \leftarrow \Psi \mathbf{e}_\theta$ 
15:     $(\mathbf{X}^{(n)})_i^T \leftarrow (\mathbf{x}_P + \mathbf{e}_X)^T$ 
16:  end for
17: until Convergence is reached
18: return  $\mathbf{X}^{(n)}$ 

```

“predicted measurement” from the original measurement row $(\mathbf{Y})_i$. The error \mathbf{e}_Y will be then reconstructed using (1), leading to a prediction error on the signal samples equal to \mathbf{e}_X . Adding \mathbf{e}_X to \mathbf{x}_P provides a new estimate of \mathbf{x} . Since the new estimate is more accurate than the old one, the process can be repeated by estimating a new, more accurate prediction. If the prediction of the row is accurate enough, the prediction error is going to be more compressible than the original vector. As a consequence, for an equal number of measurements, the ℓ_1 reconstruction will yield lower MSE. We support this claim by numerical simulations, whose results are shown in Section IV-A2.

Different initialization strategies are described in section III-D, while prediction filters for 2D signals are described in section III-C1. Since (1) is a convex problem and the prediction filters we test are linear, the overall algorithm can be considered as a projection onto convex sets. This ensures the convergence of the algorithm to the intersection of the constraint sets (if any) [16].

B. 3D signals

As has been said, acquisition of 3D hyperspectral images can be performed in different ways. Common to the various approaches are the signal dimensions, which determine the spatial and spectral resolution of the imaging system. The spectral resolution is given by the number of individual wavelengths N_{BAND} that

Algorithm 3 Acquisition algorithm for 3D signals**Require:** the hyperspectral image \mathcal{F} , M **Ensure:** the measurement matrix \mathbf{Y}

```

1: for  $i = 1$  to  $N_{\text{BAND}}$  do
2:   Draw  $\Phi^i$  of size  $M \times N_{\text{ROW}}N_{\text{COL}}$  s.t.  $(\Phi^i)_{kj} \sim \mathcal{N}(0, 1/M)$ 
3:    $\mathbf{F}^i = \mathcal{F}_{:, :, i}$ 
4:    $(\mathbf{Y})_i^T \leftarrow \Phi^i \text{vec}\{\mathbf{F}^i\}$ 
5: end for
6: return  $\mathbf{Y}$ 

```

the system is able to discriminate. In one possible approach, each wavelength is sensed individually, leading to different measurements for each spectral channel. In another approach, the system measures spatial-spectral slices individually, and different wavelengths are separated during the CS reconstruction process. In both cases, we assume that the user intends to acquire or reconstruct each spectral channel with a resolution of $N_{\text{ROW}} \times N_{\text{COL}}$ pixels. In both cases, we can represent the original data as a cube with two spatial and one spectral dimension, and interpret them either as a collection of N_{BAND} spectral channels, or as a collection of N_{ROW} spatial-spectral slices. In the following we take the first approach, and consider an hyperspectral image $\mathcal{F} \in \mathbb{R}^{N_{\text{ROW}} \times N_{\text{COL}} \times N_{\text{BAND}}}$ as a collection of N_{BAND} spectral channels $\mathbf{F}^i = \mathcal{F}_{:, :, i}$, $i = 1, \dots, N_{\text{BAND}}$, each consisting of a $N_{\text{ROW}} \times N_{\text{COL}}$ frame, i.e.

$$\mathcal{F} = [\mathbf{F}^1, \mathbf{F}^2, \dots, \mathbf{F}^{N_{\text{BAND}}}] .$$

1) *Acquisition:* According to Algorithm 3, for each spectral channel a collection $\mathbf{y}_i \in \mathbb{R}^{M \times 1}$ of M measurements is acquired as

$$\mathbf{y}_i = \Phi^i \text{vec}\{\mathbf{F}^i\} ,$$

where each sensing matrix $\Phi^i \in \mathbb{R}^{M \times N_{\text{ROW}}N_{\text{COL}}}$ is taken as Gaussian i.i.d. and $M < N_{\text{ROW}}N_{\text{COL}}$. For simplicity, M is taken as the same value for all spectral channels. The measurements of all channels are then collected in the matrix \mathbf{Y} . This setting is amenable to separate spatial reconstruction of each spectral channel using a two-dimensional transform as sparsity domain. However, we expect that separate spatial reconstruction does not yield a sufficiently accurate estimate of the original image, since it lacks modelling of spectral correlation, which is very strong for hyperspectral images.

2) *Reconstruction:* The idea behind the iterative reconstruction is that, as in the 2D case, if we can obtain a prediction of a spectral channel \mathbf{F}^i , e.g. applying the operator $P(\cdot, \cdot)$ to channels \mathbf{F}^{i-1} and \mathbf{F}^{i+1} of some initial reconstruction, then we can cancel out the contribution of

this predictor from the measurements of \mathbf{F}^i , and reconstruct only the prediction error instead of the full spectral channel. If the prediction filter is accurate, the prediction error is expected to be more compressible than the full signal, and the reconstruction will yield better results as shown in Section IV-A2. In particular, the iterative procedure starts from the initial reconstruction $\mathcal{F}^{(0)}$ of all spectral channels. At this stage, we do not specify how we generate such initial reconstruction, which is generically denoted by $F(\mathbf{Y}, \Phi)$ to indicate that it is computed from random projections \mathbf{Y} and measurement matrices Φ . Then, for every channel we first obtain $\mathbf{F}_P = P(\mathcal{F}_{:,i-1}^{(n-1)}, \mathcal{F}_{:,i+1}^{(n-1)})$. After that, we compute prediction error measurements as $\mathbf{e}_y = \mathbf{y}_i - \Phi^i \text{vec}\{\mathbf{F}_P\}$, and we use \mathbf{e}_y to reconstruct the i -th channel summing the CS reconstruction of \mathbf{e}_y to \mathbf{F}_P . Ψ_{2D} is the 2D sparsity transform matrix. If $\theta = \Psi_{\text{ROW}} \mathbf{X} \Psi_{\text{COL}}^T$ is the sparse representation of the 2D signal \mathbf{X} for some Ψ_{ROW} and Ψ_{COL} , then $\text{vec}\{\mathbf{X}\} = \Psi_{2D} \text{vec}\{\theta\}$, where

$$\Psi_{2D} = (\Psi_{\text{COL}}^T \otimes \Psi_{\text{ROW}}^T). \quad (3)$$

This process is performed on all bands, and is iterated until convergence. Again, it is worth noting that since (1) is a convex problem and the prediction filter is linear, this algorithm can be cast in terms of projections onto convex sets [16], guaranteeing convergence to the intersection of the constraint sets (if not empty). The proposed iterative reconstruction scheme is shown in Algorithm 4.

Remark 3.1: The proposed acquisition and reconstruction architecture for 3D signals can be applied in any of the 3 dimensions. For example, the acquisition can be performed for each spectral row (instead of spectral channels), and the reconstruction can be performed iterating over rows. The respective algorithms can be obtained by Algorithms 3 and 4 by properly rotating indexes and dimensions.

C. Linear Prediction Filters

1) *Row prediction filters for 2D signals:* We consider here several linear prediction filters $P((\mathbf{X})_{i-1}, (\mathbf{X})_{i+1})$ for 2D signals, looking for the one providing fastest convergence and best MSE performance. We denote as \mathbf{x}_P the result of the prediction.

Prediction filter labelled as *P1* estimates the current¹ line to be predicted as the average of the upper and lower lines:

$$\mathbf{x}_P = \frac{1}{2} ((\mathbf{X})_{i-1} + (\mathbf{X})_{i+1})^T$$

¹Here and in the following equations, we omit the index (n) denoting current iteration

Algorithm 4 Reconstruction algorithm for 3D signals

Require: the measurement matrix \mathbf{Y} , the set of Φ^i

Ensure: the estimation $\hat{\mathcal{F}}$

```

1:  $n \leftarrow 0$ 
2:  $\mathcal{F}^{(n)} = F(\mathbf{Y}, \Phi)$ 
3: repeat
4:    $n \leftarrow n + 1$ 
5:   for  $i = 1$  to  $N_{\text{BAND}}$  do
6:     if  $i = 1$  then
7:        $\mathbf{F}_P \leftarrow P(\cdot, \mathcal{F}_{:,i+1}^{(n-1)})$ 
8:     else if  $i = N_{\text{BAND}}$  then
9:        $\mathbf{F}_P \leftarrow P(\mathcal{F}_{:,i-1}^{(n-1)}, \cdot)$ 
10:    else
11:       $\mathbf{F}_P \leftarrow P(\mathcal{F}_{:,i-1}^{(n-1)}, \mathcal{F}_{:,i+1}^{(n-1)})$ 
12:    end if
13:     $\mathbf{y}_P \leftarrow \Phi^i \text{vec}\{\mathbf{F}_P\}$ 
14:     $\mathbf{e}_y \leftarrow (\mathbf{Y})_i^T - \mathbf{y}_P$ 
15:     $\mathbf{e}_\theta \leftarrow \arg \min_{\mathbf{e}} \|\mathbf{e}\|_{\ell_1} \quad \text{s.t.} \quad \Phi^i \Psi_{2D} \mathbf{e} = \mathbf{e}_y$ 
16:     $\mathbf{e}_F \leftarrow \Psi_{2D} \mathbf{e}_\theta$ 
17:     $\text{vec}\{\mathcal{F}_{:,i}^{(n)}\} \leftarrow (\text{vec}\{\mathbf{F}_P\} + \mathbf{e}_F)$ 
18:  end for
19: until Convergence is reached
20: return  $\mathcal{F}^{(n)}$ 

```

Prediction filter labelled as *P2* predicts each pixel of current line as the average of adjacent pixels of upper and lower lines

$$(\mathbf{x}_P)_j = \frac{1}{6} [(\mathbf{X})_{i-1,j-1} + (\mathbf{X})_{i-1,j} + (\mathbf{X})_{i-1,j+1} + (\mathbf{X})_{i+1,j-1} + (\mathbf{X})_{i+1,j} + (\mathbf{X})_{i+1,j+1}]$$

Finally, prediction filter labelled as *P3* predicts each pixel of current line as the *weighted* average of adjacent pixels of upper and lower lines. Weights depend on the distance from the pixel to be predicted, namely

$$(\mathbf{x}_P)_j = [a(\mathbf{X})_{i-1,j-1} + b(\mathbf{X})_{i-1,j} + a(\mathbf{X})_{i-1,j+1} + a(\mathbf{X})_{i+1,j-1} + b(\mathbf{X})_{i+1,j} + a(\mathbf{X})_{i+1,j+1}]$$

with $a = \frac{2-\sqrt{2}}{4}$ and $b = \frac{\sqrt{2}-1}{2}$.

In section IV-A, we test the performance of the linear prediction filters $P(\cdot, \cdot)$ and of the overall algorithm.

2) *Band prediction filter for 3D signals:* In the following we describe the linear prediction stage $P(\cdot, \cdot)$ employed in Algorithm 4. The prediction filter operates in a blockwise fashion. Prediction of spectral channel i is performed dividing the channel into non-overlapping spatial blocks of size 16×16 pixels. Each block is predicted from the spatially co-located block in a reference spectral channel l (typically the previous or next band). Focusing on a single 16×16 block, we denote

by $f_{m,n,i}$ the pixel of an hyperspectral image in m -th line, n -th pixel, and i -th band, with $m, n = 0, \dots, 15$, and $i = 0, \dots, N_{\text{BAND}} - 1$.

Samples $f_{m,n,i}$ belonging to the block are predicted from the samples $\hat{f}_{m,n,l}$ of the *reconstructed* reference band. In particular, a least-squares estimator [17] is computed over the block. First, a gain factor is calculated as $\alpha = \frac{\alpha_N}{\alpha_D}$, with $\alpha_N = \sum_{m,n} [(\hat{f}_{m,n,l} - \mu_l)(\hat{f}_{m,n,i} - \mu_i)]$ and $\alpha_D = \sum_{m,n} [(\hat{f}_{m,n,l} - \mu_l)^2]$. μ_i and μ_l are the average values of the co-located reconstructed blocks in bands \mathbf{F}^i and \mathbf{F}^l . Then the predicted values within the block are computed for all $m, n = 0, \dots, 15$ as $\tilde{f}_{m,n,i}^{(l)} = \mu_i + \alpha(\hat{f}_{m,n,l} - \mu_l)$.

This one-step prediction filter is employed in such a way as to take full advantage of the correlation between bands. In particular, the current band is very correlated with its two adjacent bands, while the correlation tends to decrease moving further away. Eventually, we define a predictor for a block in the current band \mathbf{F}^i as the average of two predictors obtained from the previous and the next band. $\tilde{f}_{m,n,i} = (\tilde{f}_{m,n,i}^{(i-1)} + \tilde{f}_{m,n,i}^{(i+1)})/2$. Hence, the prediction filter $P(\cdot, \cdot)$ applies this predictor to the two adjacent reconstructed spectral channels in a blockwise manner as described above, yielding a predicted spectral channel \mathbf{F}_p . Exceptions are made for the first and last band, where only the available previous/next band is used for the prediction.

D. Initialization Strategies

1) *Separate reconstruction*: Separate reconstruction represents the trivial way for reconstructing signal acquired with progressive algorithms like Algorithm 1 or Algorithm 3. It simply consists in applying the ℓ_1 reconstruction (1) to recover each separately acquired portion of the original signal (i.e. a line $(\mathbf{X})_i$ of an image or the frame \mathbf{F}^i of an hyperspectral image) independently from each other, given the corresponding Φ^i and $(\mathbf{Y})_i$.

2) *Kronecker Compressed Sensing*: Given the above, we have investigated the possibility of implementing a more sophisticated reconstruction algorithm which allows the proposed scheme to achieve good performance even for low M , i.e., for high compression ratios. To this aim, we considered the simplified 3D reconstruction scheme proposed in [13], [6], where it is shown that Kronecker product matrices are a natural way to generate sparsifying and measurement matrices for the application of CS to multidimensional signals, resulting in a formulation that is denoted by Kronecker Compressive Sensing (KCS). In KCS, starting from the assumption that the signal structure along each dimension can be

expressed via sparsity, Kronecker product sparsity bases combine the structures for each signal dimension into a single matrix and representation. This allows to obtain separable transforms matrices, thus maintaining the computational complexity to an acceptable level. Similarly, Kronecker random product measurement matrices for multidimensional signals can be implemented by performing a sequence of separate random measurements obtained along each dimension. Given the above, the application of KCS to the problem at hand is straightforward. We describe the initialization based on KCS for the more general 3D signal case, which can be trivially specialized to the 2D case: the separate (band by band) random projections $(\mathbf{Y})_i^T = \Phi^i \text{vec}\{\mathbf{F}^i\}$ can be used to get a reconstruction scheme which profitably exploits correlation in all dimensions by using a separable 3D Kronecker product sparsity domain. More specifically, we consider DCT transforms for both spatial and spectral domains since DCT transform is better than other typical transforms used in CS (e.g. Wavelet transform) on small spatial crops, while a wavelet transform would arguably provide better performance over a larger image. Accordingly, denoting by Ψ_{2D} and Ψ_λ the DCT sparsifying operator for the spatial and spectral domain, respectively, reconstruction is given by

$$\text{vec}\{\mathcal{F}^{(0)}\} = \Psi_{3D} \hat{\theta},$$

where $\Psi_{3D} = \Psi_{2D} \otimes \Psi_\lambda$ and $\hat{\theta}$ can be obtained by means of linear program reconstruction

$$\hat{\theta} = \arg \min_{\theta} \|\theta\|_{\ell_1} \quad \text{s.t.} \quad \Phi \Psi_{3D} \theta = \mathbf{y},$$

where $\mathbf{y} = \text{vec}\{\mathbf{Y}^T\}$ and Φ is the *block-diagonal* sensing matrix obtained as

$$\Phi = \text{diag}(\Phi^1, \dots, \Phi^{N_{\text{BAND}}}).$$

The reconstructed set of images $\mathcal{F}^{(0)}$ can then be used as starting point for the iterative algorithm proposed in Algorithm 4.

E. Complexity

We explain here the complexity reduction obtained using Algorithms 2 and 4 instead of the standard CS reconstruction algorithm, processing the 2D/3D signal as a whole. We specialize the discussion to 2D and 3D signals, but it can be easily extended to any multidimensional signal. For an $N_{\text{ROW}} \times N_{\text{COL}}$ image, the standard CS reconstruction algorithm has an $O(N_{\text{ROW}}^3 N_{\text{COL}}^3)$ complexity. Our algorithm performing N_{ITER} iterations has an $O(N_{\text{ITER}} N_{\text{ROW}} N_{\text{COL}}^3)$ complexity, with, usually, $N_{\text{ITER}} \ll N_{\text{ROW}}, N_{\text{COL}}$. Hence, the complexity gain that

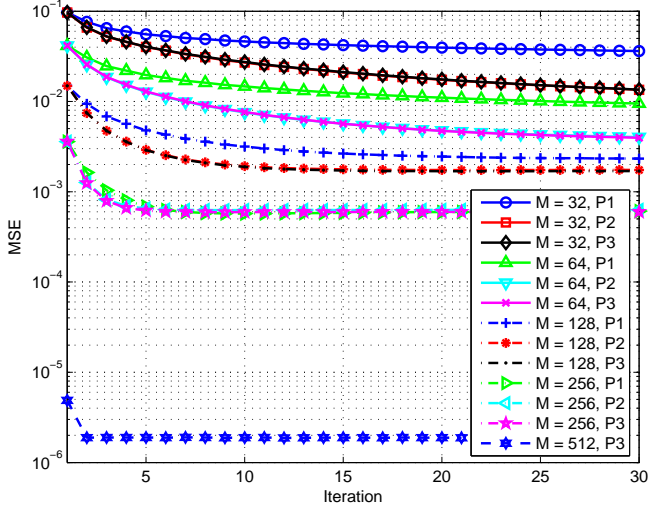


Fig. 2. Test of different prediction filters on *lena* image.

can be obtained is $\sim O(N_{\text{ROW}}^2)$. For 3D signals the same considerations hold. In this case, the gain will be $\sim O(N_{\text{BAND}}^2)$.

IV. NUMERICAL RESULTS

A. 2D images

1) *Choice of the Prediction Filter*: First, we start by seeking the linear prediction filter $P((\mathbf{X})_{i-1}, (\mathbf{X})_{i+1})$ providing fastest convergence and best MSE performance. For this test, we use the standard *lena* black and white image of size 512×512 . M takes the values of 32, 64, 128, 256 and the transform matrix Ψ is the DCT matrix. We denote as \mathbf{x}_P the result of the prediction.

Fig. 2 shows the MSE performance of the overall system for different values of M and using the prediction filters described above. Results show that the convergence is reached for each value of M . The bigger M , the faster the convergence and the smaller is the MSE at convergence. In any case, it can be noticed that the best performance is obtained for each value of M using prediction filter labelled as *P3*, i.e. the weighted average. Hence, we will use this prediction filter in our further tests, omitting to mention it from now on.

For $M = 64$, the MSE obtained with separately recovered lines is $4.16 \cdot 10^{-2}$. After 30 iterations, an MSE of $3.96 \cdot 10^{-3}$ is obtained, with a gain of 10.2 dB. The convergence in this case is quite slow, but the MSE is decreased as much as one order of magnitude. Faster convergence is obtained with $M = 128$, as after 15 MSE is decreased from $1.49 \cdot 10^{-2}$ to $1.72 \cdot 10^{-3}$, with a gain of 9.38 dB. Finally, with $M = 256$ the MSE decreases from $3.59 \cdot 10^{-3}$ to $6.18 \cdot 10^{-4}$ in 5 iterations only, with a gain of 7.64 dB.

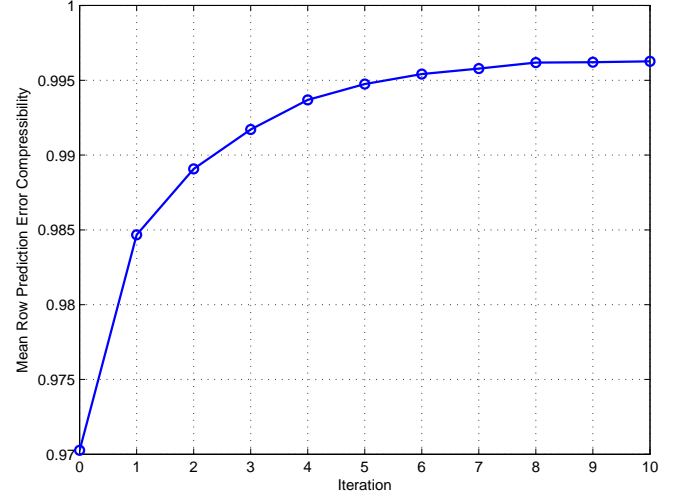


Fig. 3. Mean row compressibility of *lena* image. $M = 128$.

Finally, Fig. 2 shows as a reference the performance of a “non compressive” system. It can be noticed that the convergence is reached in only 1 iteration, leaving a residual MSE of $2 \cdot 10^{-6}$ due to unpredictable signal components.

2) *Prediction error compressibility*: In this section, we show results supporting the claim that accurate prediction leads to prediction errors which are more compressible than the original signal. Fig. 3 shows the mean row compressibility of prediction error of the same image as previous section, measured at each of the first 10 iterations of the reconstruction algorithm. The data at 0-th iteration corresponds to the compressibility of the original image. With the term *compressibility* here we mean the fraction of coefficients of the DCT of a row of the prediction error (or of the original image) below a certain threshold θ_{TH} , averaged over the rows. The threshold, evaluated for each row and at each iteration in order to take into account norm fluctuations, is computed as $\theta_{\text{TH}} = 5 \frac{\|\theta_i\|_{\ell_1}}{N_{\text{COL}}}$, where θ_i is the DCT of the current row.

Fig. 3 supports the claim that the prediction error is more compressible than the original signal and that it gets more and more compressible along iterations.

3) *Flatbed Scanner*: In this section, we apply our algorithm to images suitable to a *flatbed scanner* scenario. These are black and white graphics and text, and are depicted in Fig. 4. M takes the values of 8, 16, 32, 64, 128, 256 and (where possible) 512. Since they all have a completely white background (representing paper), they can be considered sparse in the pixel domain. Hence, the matrix Ψ is the identity matrix of size N_{COL} , namely $\mathbf{I}_{N_{\text{COL}}}$.

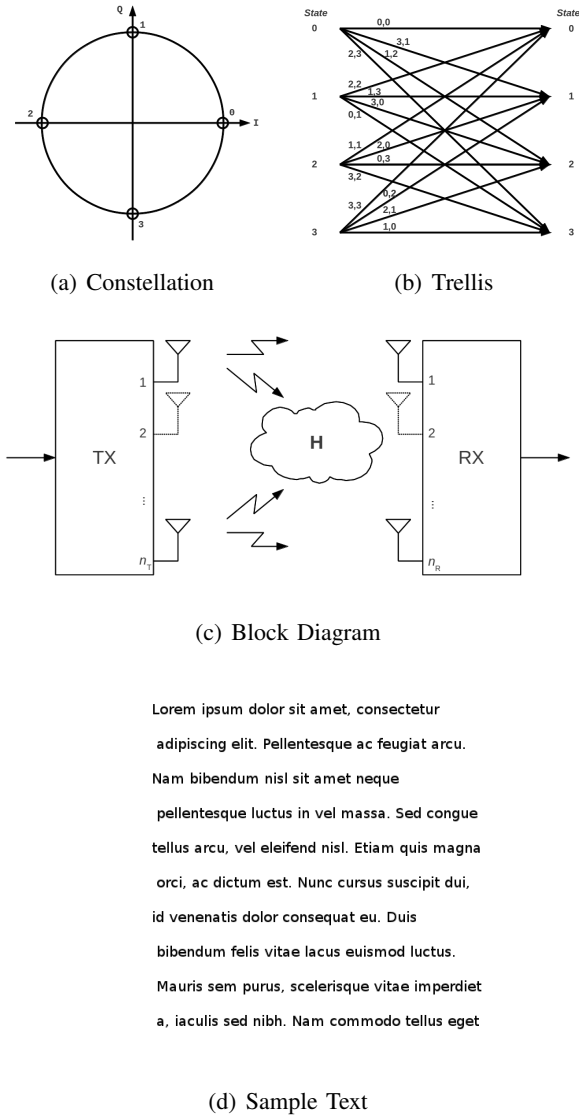
Fig. 4. The graphics used as test image for *flatbed scanner* scenario.

Fig. 4(a) is the simplest graphic, representing a QPSK constellation. Fig. 4(b) represents a slightly more complicated (hence, less sparse) graphic, the trellis of a convolutional code. Fig. 4(c) is a larger figure representing a generic block diagram. Finally, fig. 4(d) depicts a sample of generic text.

Table I reports the results obtained using the proposed algorithm. The table shows, for each image, the initial MSE (obtained using separate CS reconstruction of each line), the MSE the algorithm converges to, the performance gain, and the number of iterations necessary to reach convergence. Figures confirm the results obtained in the previous section. The more measurements are taken, the faster is the convergence and the lower is the MSE that can be obtained when the algorithm has converged. When the picture is very sparse, it is

TABLE I
MSE AND CONVERGENCE RESULTS ON SAMPLE GRAPHICS

M	init. MSE	conv. MSE	gain (dB)	steps
<i>Constellation</i> (680×576)				
64	$2.08 \cdot 10^{-2}$	$7.86 \cdot 10^{-3}$	4.23	18
128	$9.99 \cdot 10^{-3}$	$3.05 \cdot 10^{-3}$	5.15	14
256	$4.72 \cdot 10^{-3}$	$7.25 \cdot 10^{-4}$	8.14	10
<i>Trellis</i> (680×576)				
64	$8.56 \cdot 10^{-2}$	$3.98 \cdot 10^{-2}$	3.33	18
128	$7.45 \cdot 10^{-2}$	$2.00 \cdot 10^{-2}$	5.71	13
256	$3.39 \cdot 10^{-2}$	$6.51 \cdot 10^{-3}$	7.17	8
<i>Block Diagram</i> (529×1123)				
64	$8.38 \cdot 10^{-3}$	$7.02 \cdot 10^{-3}$	0.77	7
128	$5.79 \cdot 10^{-3}$	$3.97 \cdot 10^{-3}$	1.64	6
256	$2.79 \cdot 10^{-3}$	$1.66 \cdot 10^{-3}$	2.25	5
512	$1.23 \cdot 10^{-3}$	$4.71 \cdot 10^{-4}$	4.17	5
<i>Sample Text</i> (512×512)				
64	$6.40 \cdot 10^{-2}$	$4.59 \cdot 10^{-2}$	1.44	10
128	$5.46 \cdot 10^{-2}$	$2.99 \cdot 10^{-2}$	2.62	7
256	$3.39 \cdot 10^{-2}$	$1.41 \cdot 10^{-2}$	3.81	4

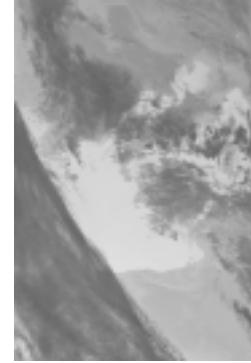


Fig. 5. The AIRS sensor gran 9 hyperspectral image, 600-th band.

possible to obtain a reduction of one order of magnitude, while when the picture is less sparse the contribution of Compressed Sensing is weaker, but still a reduction of about 50% in MSE can be obtained.

4) *Remote Sensing*: To test the performance of the proposed scheme in a *remote sensing* scenario, we use a spectral band extracted from hyperspectral image “granule 9” of the AIRS sensor. AIRS is an ultraspectral sounder with 2378 spectral channels, used to create 3D maps of air and surface temperature. The spatial size is $N_{\text{COL}} = 90$ pixels and $N_{\text{ROW}} = 135$ lines. The dataset consists in the raw output of the detector, without any processing, calibration or denoising applied. We choose the 600-th band, which is depicted in Fig. 5, but very similar results have been obtained with other bands and are omitted for brevity. M takes the values of 8, 16, 32 e 64. The sparsity basis Ψ is the DCT.

Table II (Basic Algorithm) summarizes the results obtained applying the proposed algorithm to the 600-th band of the test image. Results show that with $M = 8$

TABLE II
MSE AND CONVERGENCE RESULTS ON AIRS SENSOR IMAGE

M	init. MSE	conv. MSE	gain (dB)	steps
Basic algorithm				
8	$2.40 \cdot 10^{-1}$	$2.63 \cdot 10^{-2}$ (20-th it.)	9.6	20+
16	$9.57 \cdot 10^{-2}$	$5.21 \cdot 10^{-3}$ (20-th it.)	12.6	20+
32	$2.17 \cdot 10^{-2}$	$1.56 \cdot 10^{-3}$	11.4	10
64	$2.89 \cdot 10^{-3}$	$3.40 \cdot 10^{-4}$	9.29	4
Kronecker improved algorithm				
8	$4.60 \cdot 10^{-3}$	$3.80 \cdot 10^{-3}$	0.83	7
16	$2.62 \cdot 10^{-3}$	$2.02 \cdot 10^{-3}$	1.13	5
32	$1.22 \cdot 10^{-3}$	$9.73 \cdot 10^{-4}$	0.98	3
64	$2.95 \cdot 10^{-4}$	$2.64 \cdot 10^{-4}$	0.48	1

and $M = 16$ the convergence is very slow and is not reached after 20 iterations. On the other hand, when $M = 32$ the convergence is obtained after 10 iterations (reducing from $2.17 \cdot 10^{-2}$ to $1.56 \cdot 10^{-3}$, with a gain of 11.4 dB), while taking $M = 64$ measurements per row implies the convergence after 4 steps only (with MSE reduction from $2.89 \cdot 10^{-3}$ to $3.40 \cdot 10^{-4}$ and a gain of 9.29 dB).

As a term of comparison, we report here the MSE performance of a simple reconstruction algorithm named *Orthogonal Matching Pursuit* [18], whose complexity is linear in the number of samples of the original signal ($N_{\text{ROW}}N_{\text{COL}}$ in this case). We acquire and reconstruct the entire image as a whole using $M = 32 \cdot N_{\text{ROW}}$ and $M = 64 \cdot N_{\text{ROW}}$ measurement, to be compared with the performance of our algorithm with $M = 32$ and $M = 64$, respectively. For $M = 32 \cdot N_{\text{ROW}}$, we obtain an MSE of $1.9 \cdot 10^{-3}$, while for $M = 64 \cdot N_{\text{ROW}}$ we obtain an MSE of $1.7 \cdot 10^{-3}$. Hence, our algorithm with $M = 32$ performs 3 dB better than OMP with the same total amount of measurements, while with $M = 64$ the gain is 8 dB.

5) *Improving performance with Kronecker Compressed Sensing*: An improvement to the performance of the algorithm is obtained using *Kronecker Compressed Sensing* (KCS), described in section III-D2, into the algorithm we showcase. Hence, we use KCS to initialize the iterative algorithm we propose in this paper (instead of separate linewise reconstruction) and apply it to the *remote sensing* scenario. The performance of this modified version of the algorithm are reported in Table. II (Kronecker improved algorithm). The figures show two effects. First, the initial MSE is much lower than in the separate reconstruction case. This gain can be noticed in particular when M is small and is due to the better performance of KCS reconstruction with respect to separate reconstruction; second, the iterative algorithm slightly improves the overall performance and

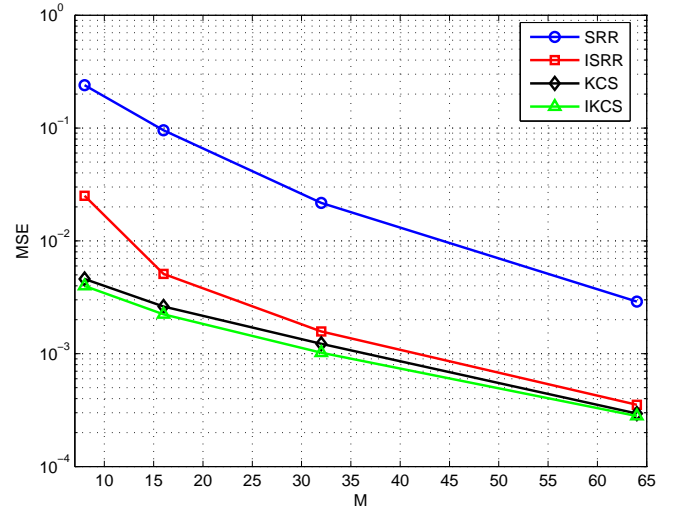


Fig. 6. Performance comparison of proposed algorithms vs. M .

converges in very few steps. This is due to the fact that KCS captures also correlation in vertical direction, making the contribution of each iteration less effective.

Figure 6 summarizes the best MSE performance obtained by Separate Row Reconstruction (SRR), our Iterative algorithm initialized with Separate Row Reconstruction (ISRR), the Kronecker Compressed Sensing (KCS) and our Iterative algorithm with KCS initialization (IKCS) vs. the number of measurements M . Best performing algorithms are the ones implementing KCS. Plain KCS shows a gain of 7.37 dB over ISSR when $M = 8$, and 1.10 dB when $M = 32$. When using IKCS, roughly 1 dB of additional gain can be obtained with very few iterations.

B. 3D signals

We report reconstruction results on a few scenes that are used as reference for onboard lossy compression in the “multispectral and hyperspectral data compression” working group of the Consultative Committee for Space Data Systems (CCSDS), namely scene *sc0* of AVIRIS (Yellowstone) and granule 9 (*gran9*) of AIRS. AVIRIS is a spectrometer with 224 bands, and the size of this image is 512 lines and 680 pixels. AIRS has already been described in previous sections. Because of the complexity of the reconstruction process and the large amount of data, we do not use the complete images, but rather a 32x32 spatial crop with all spectral channels. Both are *raw* images, i.e. they are the output of the detector, with no processing, calibration or denoising applied. These images are noisier than the corresponding processed images, but more realistic for application to onboard sensors.

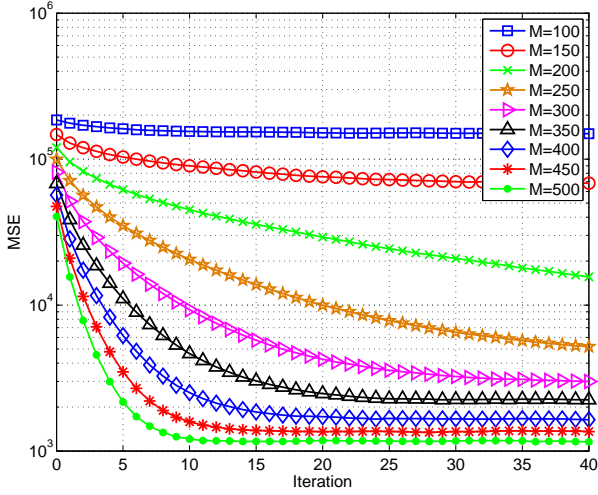


Fig. 7. AVIRIS: Reconstruction MSE of Algorithm 4 as a function of the number of iterations with initial separate 2D reconstruction.

1) *Preliminary experimental analysis with initial separate 2D reconstruction:* We have carried out some experiments to preliminarily assess the validity of the proposed algorithm when the initial reconstruction images are computed using separate 2D DCT transforms band by band. In particular, Figure 7 shows the MSE behaviour experienced on AVIRIS images as a function of the number of iterations for different values of the number of projections M . A similar behaviour is observed for AIRS images. Note that for medium to high M , iterations are effective in reducing MSE, e.g., for $M > 400$ the proposed algorithm improves the MSE up to a factor of 35 with respect to the initial reconstruction. Moreover, convergence to the minimum attainable MSE is obtained in a relatively small number of iterations. For lower M , convergence is slower and MSE reduction is less effective. In particular, for very low M , e.g. for $M = 100$ convergence is very slow and MSE reduction is negligible. In essence, the algorithm shows a threshold behaviour with respect to the initial reconstructed images: a poor initial reconstruction prevents the iterative algorithm to improve the MSE while if the initial reconstruction's MSE falls below a minimum threshold, the improvement is remarkable and convergence very fast.

2) *Improving initial reconstruction by means of Kronecker CS:* To assess the effectiveness of such an approach, in Figure 8 we show the MSE behaviour experienced on AVIRIS images as a function of the number of iterations, for different M , when the starting point of the iterative scheme proposed in Algorithm 4 is obtained through Kronecker 3D reconstruction. Comparing with

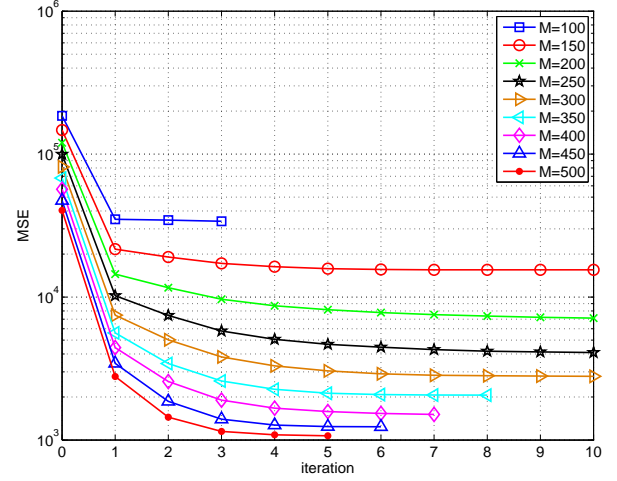


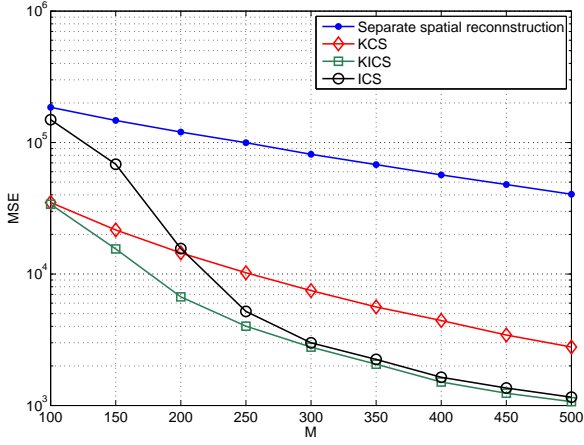
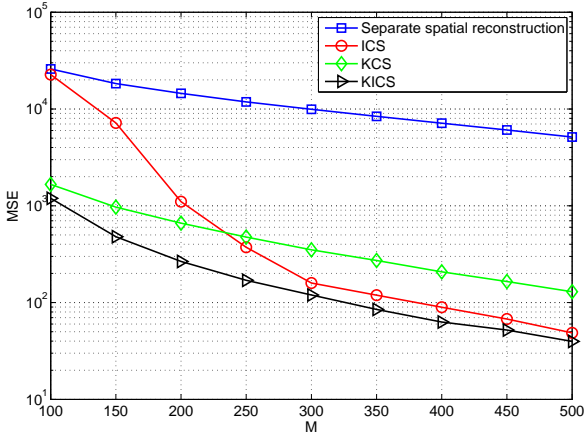
Fig. 8. AVIRIS: Reconstruction MSE of Algorithm 4 with 3D Kronecker starting point, as a function of the number of iterations.

Figure 7 it can be observed that, as expected, the MSE starting point is much lower and convergence is achieved in few iterations. Moreover, despite 3D Kronecker reconstruction already exploits correlation in the spectral domain, the proposed iterative algorithm still allows to improve the MSE up to a factor of 3 with respect to the initial reconstruction. In the next section, we describe in more details the experiments we conducted to evaluate the performance of the two proposed reconstruction schemes, namely iterative compressed sampling (ICS) and Kronecker-iterative compressed sampling (KICS), which are both based on the iterative procedure described in Algorithm 4, with the initial point computed by means of separate reconstruction and KCS, respectively.

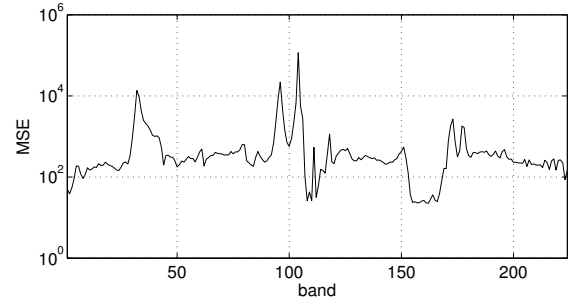
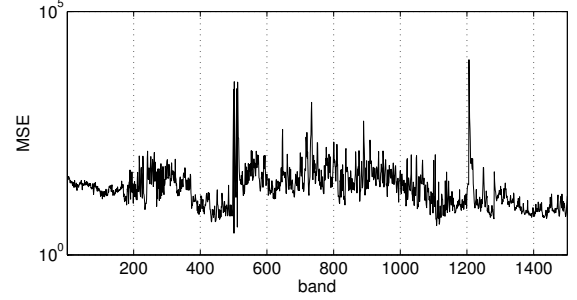
3) *Spectral channel analysis:* We compare results of the proposed ICS and KICS with those obtained through separate spatial reconstruction (S2D) of each spectral channel and through 3D KCS. The reconstruction algorithm for the iterative schemes is run for 40 iterations, with several values of M .

Results in terms of MSE versus M are shown in Figures 9 and 10 for the AVIRIS and AIRS scenes, respectively. As can be seen, S2D spatial reconstruction yields very large mean-squared error (MSE), typically in excess of $5 \cdot 10^4$ for AVIRIS and of $7 \cdot 10^3$ for AIRS. Considering that the average signal energy for this crop is equal to $2.76 \cdot 10^7$ for AVIRIS and $4.85 \cdot 10^6$ for AIRS, spatial reconstruction yields an average percentage error of nearly $\pm 4\%$ both test images, which is inadequate for most applications.

As anticipated in Figure 7, the proposed ICS reconstruction algorithm allows to improve the MSE up to a

Fig. 9. AVIRIS: Reconstruction MSE vs. M for different schemes.Fig. 10. AIRS: Reconstruction MSE vs. M for different schemes.

factor of 35 for high M , but it is not effective for low M . On the other hand, the 3D KCS reconstruction without iterative predictions performs quite well for low M but its performance are not so good for high M , e.g., it is even worse than ICS for $M > 200 - 250$. Eventually, KICS gives the best performance over the whole range of considered M . In other words, combining 3D KCS with predictive CS allows to accurately reconstruct original images requiring a number of linear measurements much smaller than the original samples. On the other hand, average results provide a somewhat biased picture though. In Fig. 11 and 12, the individual MSE per band and for $M = 450$ obtained through KICS algorithm on AVIRIS and AIRS scene is shown, respectively. As can be seen, in most bands the MSE is very small, between 100 and 400. The average MSE is biased by a relatively small number of bands which are reconstructed with large error. Visual inspection shows that e.g. band 104

Fig. 11. AVIRIS: Reconstruction MSE for each band, $M = 450$.Fig. 12. AIRS: Reconstruction MSE for each band, $M = 450$.

is extremely noisy (hence not at all sparse) and contains almost no information, while band 32 is misregistered with respect to band 31, yielding poor prediction. This shows that, on average, a much lower relative error can be achieved in most bands, except for noisy bands, which are not very important altogether, or misregistered bands, where improved prediction models can be employed to improve the reconstruction.

4) *Separate acquisition of spectral rows*: Fig. 13 refers to a $32 \times 32 \times 32$ crop of AIRS. Instead of taking separate measurements of spectral channels, here we separately acquire horizontal spectral rows $\mathbf{F}^i \in \mathbb{R}^{N_{\text{COL}} \times N_{\text{BAND}}}$, where $\mathbf{F}^i = \mathcal{F}_{i,:}$ and $i = 1, \dots, N_{\text{ROWS}}$. The cube is then reconstructed using algorithm 4, iterating over rows instead of wavelength. In this case, to predict spectral rows we use prediction filter $P1$ described in section III-C1. Results show that separate reconstruction (the MSE of the initial step) applied to spectral rows leads to better performance than the same algorithm applied to spectral channels. Nevertheless, the iterative algorithm is less effective when iterating over rows than when iterating over wavelength. Results show that only a slight MSE gain can be obtained from iterative algorithm. This effect is due to stronger correlation along wavelength direction than between rows. Stronger correlation is better exploited by Compressed Sensing, leading to better separate reconstruction performance,

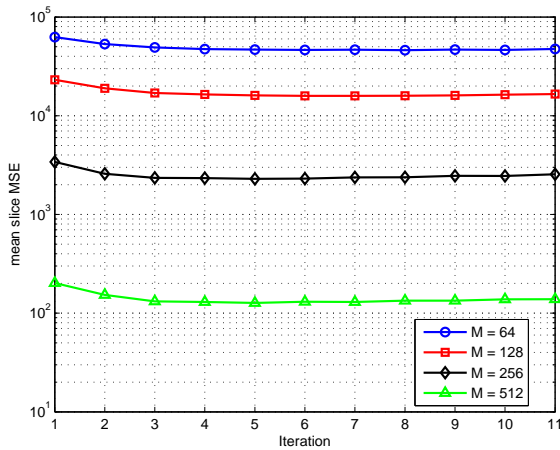


Fig. 13. MSE performance of ICS applied to spectral rows.

while weaker correlation between rows yields only a minor contribution of the iterative algorithm.

V. CONCLUSIONS

In this paper, we proposed a general architecture for the acquisition and reconstruction of multidimensional correlated signals with manageable computational complexity. The acquisition is based on Compressed Sensing and consists in taking a sequence of separate random linear measurements of the signal, grouping subsets of the coordinates, in a progressive fashion. The reconstruction process implements an iterative architecture relying on linear prediction filters and the Compressed Sensing reconstruction of the prediction error, which is supposed to be more compressible than the original signal. Then, we specialize this framework to 2D signals and 3D signals, proposing practical applications for these scenarios. For 2D signals, we envisage applications to *flatbed scanners* and *remote sensing*, which already perform progressive scanning. For 3D signals, the straightforward application is *hyperspectral imaging*, even if applications related to *video* could be also be imagined. For both scenarios, we show that the performance in terms of MSE and speed of convergence depend on two factors. On one hand, the initial MSE of the algorithm depends on the initialization strategy. We show that the performance obtained by trivially initializing the algorithm with separate measurement reconstruction can be improved by using the so-called *Kronecker Compressed Sensing*, which is able to capture the correlation in all dimensions at the cost of an increased computational complexity. On the other hand, the effectiveness of the iterative algorithm in terms of MSE gain strongly depends on the choice of the linear prediction filter and on the amount of signal correlation along the iteration dimension.

REFERENCES

- [1] E. Candes, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, vol. 1, no. 3, 2006.
- [2] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [3] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 83, pp. 83–91, Mar. 2008.
- [4] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. Baraniuk, "An architecture for compressive imaging," in *International Conference on Image Processing*, 2006, pp. 1273–1276.
- [5] R. Willett, M. Gehm, and D. Brady, "Multiscale reconstruction for computational spectral imaging," in *Proc. of SPIE Electronic Imaging*, 2007.
- [6] M. Duarte and R. Baraniuk, "Kronecker product matrices for compressive sensing," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 3650–3653.
- [7] J. Fowler, S. Mun, and E. Tramel, "Multiscale block compressed sensing with smoothed projected landweber reconstruction," in *19th European Signal Processing Conference (EUSIPCO 2011)*, 2011, pp. 564–568.
- [8] A. Abrardo, M. Barni, C. Carretti, S. Kamdem, and E. Magli, "A compressive sampling scheme for iterative hyperspectral image reconstruction," in *19th European Signal Processing Conference (EUSIPCO 2011)*, 2011, pp. 1120–1124.
- [9] M. Trocan, T. Maugey, E. Tramel, J. Fowler, and B. Pesquet-Popescu, "Multistage compressed-sensing reconstruction of multiview images," in *Multimedia Signal Processing (MMSp), 2010 IEEE International Workshop on*. IEEE, 2010, pp. 111–115.
- [10] M. Wakin, "A manifold lifting algorithm for multi-view compressive imaging," in *Picture Coding Symposium, 2009. PCS 2009*. IEEE, 2009, pp. 1–4.
- [11] G. Coluccia and E. Magli, "A Novel Progressive Image Scanning and Reconstruction Scheme based on Compressed Sensing and Linear Prediction," in *IEEE International Conference on Multimedia and Expo 2012*, Melbourne, Australia, July 2012.
- [12] M. Duarte and R. Baraniuk, "Kronecker product matrices for compressive sensing," in *Proc. of IEEE ICASSP*, 2010.
- [13] —, "Kronecker compressive sensing," *Preprint*, <http://dsp.rice.edu/publications/kroneckercompressive-sensing-0>, Submitted to *IEEE Transactions on Image Processing*, 2009.
- [14] E. Candes and T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [15] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.
- [16] P. Combettes, "The foundations of set theoretic estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, 1993.
- [17] M. Slyz and L. Zhang, "A block-based inter-band lossless hyperspectral image compressor," in *Data Compression Conference, 2005. Proceedings. DCC 2005*. IEEE, 2005, pp. 427–436.
- [18] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.