RECONSTRUCTION ALGORITHM OPTIMIZATION AND SYSTEM MATRIX DESIGN IN COMPRESSIVE SPECTRAL IMAGING

YURI HERCILIA MEJIA MELGAREJO

Ingeniera Electrónica

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE INGENIERÍAS FISICOMECÁNICAS

ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE TELECOMUNICACIONES

DOCTORADO EN INGENIERÍA, ÁREA ELECTRÓNICA

BUCARAMANGA

2018

RECONSTRUCTION ALGORITHM OPTIMIZATION AND SYSTEM MATRIX DESIGN IN COMPRESSIVE SPECTRAL IMAGING

YURI HERCILIA MEJIA MELGAREJO

Ingeniera Electrónica

Informe final del trabajo de investigación para optar al título de

Doctora en Ingeniería, área Ingeniería Electrónica

Director:

Ph.D. HENRY ARGUELLO FUENTES

Doctor en Ingeniería Eléctrica

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE INGENIERÍAS FISICOMECÁNICAS

ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE TELECOMUNICACIONES

DOCTORADO EN INGENIERÍA, ÁREA ELECTRÓNICA

BUCARAMANGA

2018

CONTENTS

pag.

INTRODUCTION 19

1 COMPRESSIVE SPECTRAL IMAGING, CSI 25

1.1 SPECTRAL IMAGING 25

1.2 COMPRESSIVE SPECTRAL IMAGING AND SYSTEM MATRIX MODEL 26

1.3 PREVIOUS WORK ON MATRIX DESIGN 30

1.4 RECOVERY ALGORITHMS 30

1.5 PREVIOUS WORK ON RECOVERY ALGORITHMS 31

2 MATRIX DESIGN IN CSI BY EIGENVALUE ANALYSIS 33

2.1 STRUCTURED, SPARSE, AND BINARY MEASUREMENT MATRICES (SBIM) 33

2.2 MEASUREMENT MATRIX DESIGN 37

2.3 MEASUREMENT MATRIX OPTIMIZATION ALGORITHM 41

2.4 SIMULATION RESULTS 45

2.5 CHAPTER CONCLUSION 51

3 MATRIX DESIGN IN CSI WITH TWO SENSORS 55

3.1 CSI SYSTEMS MODELED FOR A FUSION FRAMEWORK 57

3.2 COMPRESSIVE SPECTRAL IMAGING FUSION 59

3.3 CODED APERTURE DESIGN METHODS 61
3.4 RECOVERY ALGORITHM 66
3.5 SIMULATIONS 68
3.6 CHAPTER CONCLUSION 70

4 CSI BASED ON MULTISPECTRAL FILTER ARRAY 72
4.1 TRADITIONAL DEMOSAICING 73
4.2 APPROACHES FOR RECONSTRUCTION WITH ADJUSTABLE SPATIAL RESOLUTION 73
4.3 MULTISPECTRAL FILTER DESIGN 76
4.4 SIMULATION AND RESULTS 76
4.5 CHAPTER CONCLUSION 79

5 OPTIMIZED COMPRESSIVE RECONSTRUCTION ALGORITHM BASED ON FILTERED GRADIENT 82
5.1 PROPOSED FILTERING STEP 83
5.2 PROPOSED ALGORITHM 84
5.3 SIMULATIONS AND RESULTS 86
5.4 CHAPTER CONCLUSION 93

6 BAYESIAN CSI RECONSTRUCTION BY USING A LOCAL STRUCTURED PRIOR 95
6.1 PROPOSED BAYESIAN RECONSTRUCTION 95
6.2 GIBBS SAMPLER 98
6.3 SIMULATION RESULTS 99
6.4 CHAPTER CONCLUSION 102

7 CONCLUSIONS, DISCUSSION, AND FUTURE WORK **103**

BIBLIOGRAPHIC REFERENCES **106**

BIBLIOGRAPHY **114**

LIST OF FIGURES

1 The matrix $\mathbf{\Phi}$ in (1.4) is shown for $Z = 3$, and $L = 4$, and 2 snapshots. Colored squares represent passing elements whereas the other elements are zero. 28

2 Sensing phenomena representation of the CASSI system. The qth slice of the data cube is coded by a row of the coded aperture and dispersed by the prism. The detector captures the intensity $\mathbf{y}$ by integrating the coded light. 29

3 Example of defined variables for a 5×7 matrix $\mathbf{\Phi}$, the dotted red squares represent the modifiable positions of the matrix, the white squares are ones, and the black squares are zeros. In this example, the desirable number of elements per columns is established at $\mu = 1$. 42

4 Singular values for the PFA system for (a) 50% and (b) 75% of compression ratio, with $Z = 32$ and $L = 4$. 45

5 Singular values for the C-CASSI system for (a) 55% and (b) 88% of compression ratio, with $Z = 32$ and $L = 4$. 46

6 Singular values for the compressive X-ray tomography with 50% of compression ratio, $M_1 = 552$ and $\hat{P} = 3$. 47

7 Histogram of the vector $\mathbf{p}$ for the PFA system in subsection II, A, 1) with $Z = 128$, $L = 16$, and compression ratio of 18% for (a) random realizations (b) designed with the CMD algorithm. Histogram of the vector $\mathbf{q}$ (c) random realizations (d) designed with the CMD algorithm. 48

8 Histogram of the vector $\mathbf{p}$ for the C-CASSI system in subsection II, A, 2) with $Z = 128$, $L = 16$, and a compression ratio of 21% with (a) random realizations (b) designed with the CDM algorithm. Histogram of the vector $\mathbf{q}$ (c) random realizations (d) designed with the CDM algorithm. 49

9 Histogram of the vector $\mathbf{p}$ for the compressive X-ray tomography system in subsection II, B, with $Z \times Z = 128 \times 128$, $\hat{P} = 3$, in a specific angle $m = 0$ for (a) random realizations (b) designed with the CMD algorithm. Histogram of the vector $\mathbf{q}$ (c) random realizations (d) designed with the CMD algorithm. 50

10 (a) Original. Differences between reconstructions on PFA system with 25% of compression ratio, (b) Original-Random and (c) Original-Proposed with SNR=20[dB], and (d) Original-Random and (e) Original-Proposed with SNR=10[dB]. 51

11 (a) Original. Differences between reconstructions on C-CASSI with 26% of compression ratio (b) Original-Random and (c) Original-Proposed with SNR=20[dB], and (d) Original-Random and (e) Original-Proposed with SNR=10[dB] 53

12 (a)Original. Differences between reconstructions on compressive X-ray tomography (a) Original-Random and (b) Original-Proposed noiseless and (c) Original-Random (d) Original-Proposed with SNR=20 dB 53

13 Reconstructions of the Pavia University scene for the proposed designs with 17% of compression ratio for the PFA system 69

14 Reconstructions of the butterfly scene for the proposed designs with 17% of compression ratio for the PFA system 69

15 Example of $q = 2$ that forms 4 subsets of measurements in a single snapshot of the MSFA-based sensor. 74

16 The downsampling matrix $\mathbf{D}^{\ell}$ for $q = 2$, $Z = 6$, and $\ell = 0, 1, 2, 3$. The white squares represent one-valued elements. 75

17 For the synthetic Balloons data cube, (first row) reconstruction results for band pass filters for (a) $q = 2$, (b) $q = 4$, and (c) $q = 8$, (second row) reconstruction results for dichroic filters for (d) $q = 2$, (e) $q = 4$, and (f) $q = 8$. 78

18 For the synthetic Beads data cube, (first row) reconstruction results for band pass filters for (a) $q = 2$, (b) $q = 4$, and (c) $q = 8$, (second row) reconstruction results for dichroic filters for (d) $q = 2$, (e) $q = 4$, and (f) $q = 8$. 78

19 For the Beads data cube, (first row) reconstruction results for band pass filters for (a) $q = 2$, (b) $q = 4$, and (c) $q = 8$, (second row) reconstruction results for dichroic filters for (d) $q = 2$, (e) $q = 4$, and (f) $q = 8$. 79

20 For the Beads data cube, (first row) reconstruction results for band pass filters for (a) $q = 2$, (b) $q = 4$, and (c) $q = 8$, (second row) reconstruction results for dichroic filters for (d) $q = 2$, (e) $q = 4$, and (f) $q = 8$. 80

21 Gray-scale versions of reconstructions for $L = 6$, and $q = 2$ neighborhood side size. For the Balloons database: (a) original, and (b) demosaicing reconstruction. For the CS approach reconstructions using (e) band pass filters, and (top) neighborhood CS, (bottom) complete CS; (f) dichroic filters, and (top) neighborhood CS, (bottom) complete CS. Similar results (c-d) and (g-h) for Beads datacube. The number in the images is the PSNR of the reconstruction. 81

22 Example of the products $\mathbf{x} = \mathbf{\Phi y}$ and $\mathbf{x}^{ref} = \mathbf{\Phi}^T\mathbf{y}$ for a spectral band in the CASSI system. 84

23 (a) A zero mean Gaussian kernel filter of size 3×3 and standard deviation of 0.6, (b) illustrates an example of a $\mathbf{W}$ for $N = 36$ created using the Gaussian kernel fitler in (a). 85

24 Mean PSNR of the Flowers reconstructed image with five different levels of standard deviation ($\sigma \in${0.2, 0.6, 0.9, 1.2, and 2}) and the traditional version of the GPSR as a function of the compression ratio, for a transmittance of the coded apertures of 40% in the CASSI system. 89

25 Evolution of the PSNR in each iteration for the filtered approach (blue) and the unfiltered approach (red), for the specific case of a transmittance of 40%, a compression ratio of 50%, and $\sigma = 0.6$, for (a) Feathers and (b) Flowers databases. 90

26 The 2^{nd}, 4^{th}, 6^{th}, and 8^{th} reconstructed spectral bands are presented in each row, respectively. The first and second columns show the original and their zoom versions. The third and fourth columns are the reconstructions from 20% compression and transmittance of 40% using unfiltered gradient algorithm and the fifth and sixth columns are the reconstructions using the filtered gradient algorithm. 91

27 The 2^{nd}, 4^{th}, 6^{th}, and 8^{th} reconstructed spectral bands are presented in each row, respectively. The first and second columns show the original and their zoom versions. The third and fourth columns are the reconstructions from 20% compression rate and transmittance of 40% using traditional gradient algorithm, and the fifth and sixth columns are the reconstructions with same parameters using the filtered gradient algorithm. 92

28 (a) Lego spectral data cube is shown as it would be viewed by a Stingray F-033C CCD Color Camera (b) PSNR of the reconstructed Spectral bands, comparison of the filtered approach, SALSA, AMP-3D-Wiener, and traditional GPSR, for the Lego image cube. Cube size is 256$\times$ 256 and 24 spectral bands. The measurements are captured with 2 shots. (c) PSNR comparison as a function of the number of shots. In both cases random Gaussian noise is added to the measurements such that the SNR is 20 dB. 93

29 Seventh spectral band of the image: (Left) Ground truth. Reconstruction results for: (top center) the proposed method, (bottom center) SpaRSA Smooth, (top right) Bayesian LASSO and (bottom right) SpaRSA LASSO. 100

30 Spectral signature for pixel $\#(20, 33)$. 101

LIST OF TABLES

1 Mean percentage of linearly dependent rows for 10 random realizations of spectral imagers matrices 38

2 Mean PSNR for different compression ratios [%] with SNR=20[dB], for PFA system 47

3 Mean PSNR as a function of noise (SNR) for the PFA system with 25% and 58% of compression ratio 51

4 Mean PSNR for different compression ratios [%], with SNR=20[dB] for C-CASSI system 52

5 Mean PSNR as a function of noise (SNR) for C-CASSI system with 26% and 61% of compression ratio 52

6 Mean PSNR as a function of compression [%] without noise and with SNR=20 dB for Compressive X-ray tomography 53

7 Mean PSNR for different noise levels for the Random, the Gradient, and the Proposed approaches 54

8 Mean PSNR in the PFA system, for different compression ratios with the $256 \times 256 \times 24$ Pavia image 68

9 Mean PSNR in the PFA system, for different compression ratios with the $128 \times 128 \times 24$ Butterfly image 70

10 Mean PSNR in the C-CASSI system, for different compression ratios with the $256 \times 256 \times 24$ Pavia image 70

11 Mean PSNR in the C-CASSI system, for different compression ratios with the $128 \times 128 \times 32$ Butterfly image 70

12 Mean PSNR of the Flowers reconstructed image with three different levels of standard deviation and filter sizes, and the traditional version of the GPSR for a transmittance of 20% in the Dual-coded system. . . 88
13 Mean PSNR of the Flowers reconstructed image with three different levels of standard deviation and filter sizes, and the traditional version of the GPSR for a transmittance of 50% in the in the Dual-coded system. 88
14 Mean PSNR of the Flowers reconstructed image with three different levels of standard deviation and filter sizes, and the traditional version of the GPSR for a transmittance of 60% in the in the Dual-coded system. 89
15 PSNR [dB] as function of filter size 90
16 Full conditionals $f(\delta_i^2|\theta_i, \sigma_n^2, a)$, $f(\boldsymbol{\theta}|\mathbf{y}, \sigma_n^2, \boldsymbol{\delta}^2, \lambda)$, $f(\lambda|\boldsymbol{\theta})$, $f(a|\boldsymbol{\delta}^2)$ and $f(\sigma_n^2|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\delta^2})$ associated with (6.11). 99
17 PSNRs [dB] obtained by the different algorithms. 100
18 Mean standard deviations of the estimations of $\boldsymbol{x}$. 101
19 Computational costs for a 53% compression ratio. 101

RESUMEN

TÍTULO: OPTIMIZACIÓN DEL ALGORITMO DE RECONSTRUCCIÓN Y DISEÑO DE LA MATRIZ DEL SISTEMA EN ADQUISICIÓN COMPRESIVA DE IMÁGENES ESPECTRALES *

AUTOR: Yuri Hercilia Mejia Melgarejo **

PALABRAS CLAVE: Muestreo compresivo, diseño de la matriz del sistema, algoritmo de reconstrucción, imágenes espectrales.

Los sensores compresivos de imágenes espectrales reducen el número de pixeles muestreados mediante la combinación de información espectral codificada de una escena en proyecciones bidimensionales. El diseño de la matriz de muestreo que modela el sensor y la optimización del algoritmo de reconstrucción son áreas importantes de investigación hoy en día. Este trabajo desarrolla algunos enfoques de diseño de la matriz de muestreo para sistemas compresivos de imágenes espectrales, además de propuestas para la mejorar el algoritmo de reconstrucción. El diseño de la matriz consiste en el estudio de la independencia lineal de las filas de la matriz de muestreo y su relación con los valores propios, dando lugar a restricciones físicas en los sensores. Estos diseños también se estudian en el caso en el que se reconstruye una imagen espectral a partir de la fusión de las medidas de dos sensores compresivos. Por otro lado, en la optimización del algoritmo de reconstrucción se explota el hecho de que las imágenes espectrales son suaves en el dominio espacial. Primero, se propone un filtrado pasa bajas dentro del proceso iterativo de los algoritmos de reconstrucción y segundo, se reformula un problema de minimización con un regularizador que promueve suavidad en la imagen y resolviendo por un enfoque Bayesiano.

*Trabajo de investigación

**Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director, Ph.D. Henry Arguello Fuentes.

ABSTRACT

TITLE: RECONSTRUCTION ALGORITHM OPTIMIZATION AND SYSTEM MATRIX DESIGN IN COMPRESSIVE SPECTRAL IMAGING *

AUTHOR: Yuri Hercilia Mejia Melgarejo **

KEYWORDS: Compressive sensing, system matrix design, reconstruction algorithm, spectral imaging.

Compressive spectral imagers take a reduced number of sampled pixels by coding and combining the spatio spectral information of a scene in two dimensional projections. The design of the matrix that represents the compressive spectral imager and the optimization of the reconstruction algorithm are important areas of research nowadays. This work develops some approaches to design the sampling matrix of compressive spectral imagers, as well as a proposal for improving the reconstruction algorithm. The design of the matrix consists of relating the linear independence of the sampling matrix rows and columns with the bounds of its eigenvalues, giving rise to physical restrictions in the sensors. These designs are also studied in the case where a spectral image is reconstructed by fusing the measurements of two different compressive spectral imagers, each one with different spatial and spectral resolutions. On the other hand, in the optimization of the reconstruction algorithm, we exploit the fact that the spectral images are soft in the spatial domain. First, we propose a low pass filter within the iterative process of the reconstruction algorithms and second, we reformulate the minimization problem with a regularizer that promotes smoothness in the image and solving it by a Bayesian approach.

*Research Work

**Faculty of Physical-Mechanical Engineering. Department of Electrical, Electronics, and Telecommunication Engineering. Advisor, Ph.D. Henry Arguello Fuentes.

List of Abbreviations

CS	**C**ompressive **S**ampling
HS	**H**yperspectral
MS	**M**ultispectral
CASSI	**C**oded **A**perture **S**napshot **S**pectral **I**mager
SSCSI	**S**patio-**S**pectral **E**ncoded **C**ompressive **S**pectral **I**mager
HR	**H**igh **R**esolution
ADMM	**A**lternating **D**irection**M**ethoid of **M**ultipliers
MSFA	**M**ulti-**S**pectral **F**ilter **A**rray

List of Symbols

Variable notatios

$\mathbf{x}$	Reference spectral image
$\mathbf{\Phi}_.$	Measurement matrix
$\mathbf{\Psi}$	Sparsity operator
$\boldsymbol{\theta}$	Sparse version of $\mathbf{x}$
$\mathbf{y}_.$	Compressed measurements
$._T$	Transpose operator

Size notatios

M	number of rows of $\mathbf{\Phi}$
N	number of columns$\mathbf{\Phi}$
L	number of spectral bands
Z	Spatial resolution

Mathematical symbols

$\ll$	Much lower
$\gg$	Much greater
$\otimes$	Kronecker product
$\|\|\cdot\|\|_1$	ℓ_1 norm
$\|\|\cdot\|\|_2$	ℓ_2 norm

INTRODUCTION

The advances in sensor technology allow high-resolution signals to be acquired with a large amount of digital data, which leads to high computational costs of acquisition, storage, transmission, and processing. These computational costs are a common issue in sensing systems dealing with radar signals, multimedia signals, medical, biomedical data, spectral remote sensing images, among others. On the other side, the compressive sensing theory has changed the traditional sampling paradigm, by reducing the amount of data and the acquisition costs [1], [2].

Compressive sensing (CS) is a mathematical framework that aims to recover a signal $\mathbf{x} \in \mathbb{R}^N$ from $M < N$ linear measurements, performing compression during signal acquisition [1], [3], [4]. Specifically, in CS a signal $\mathbf{x}$ is acquired by computing M inner products between each row of the measurement matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ and the underlying signal, that is, $\mathbf{y} = \mathbf{\Phi x}$. Although M is smaller than N, $\mathbf{y}$ contains plenty of useful information about the signal $\mathbf{x}$, when $\mathbf{x}$ is sparse or has a sparse representation in some domain $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ [5]. The signal $\mathbf{x}$ is called S-sparse in $\mathbf{\Psi}$ if $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$ with $\|\boldsymbol{\theta}\|_0 = S$, where $\|\boldsymbol{\theta}\|_0$ is used to count the non-zero entries in $\boldsymbol{\theta}$. Therefore, the measurements $\mathbf{y}$ of a compressible signal can be modeled as $\mathbf{y} = \mathbf{A}\boldsymbol{\theta} + \boldsymbol{\omega} = \mathbf{\Phi\Psi}\boldsymbol{\theta} + \boldsymbol{\omega}$, where $\boldsymbol{\theta} \in \mathbb{R}^N$ contains the coefficients of $\mathbf{x}$ in the sparsifying basis $\mathbf{\Psi}$, $\mathbf{A}$ is called the sensing matrix, and $\boldsymbol{\omega}$ is additive Gaussian noise. Typically, the underlying signal is reconstructed by solving an ℓ_2-ℓ_1 optimization problem as $\widehat{\mathbf{x}} = \mathbf{\Psi}\left(\arg\min_{\boldsymbol{\theta}} \|\mathbf{y} - \mathbf{\Phi\Psi}\boldsymbol{\theta}\|_2^2 + \tau \|\boldsymbol{\theta}\|_1\right)$, where τ is a regularization parameter.

An important set of applications such as image acquisition [6], [7], spectral imaging [8]–[12], and computed tomography [13] have successfully implemented compressive sensors, since they take a few linear projections of the signals to be measured, with a high probability of reconstruction. Those compressive sensors can be modeled as a system of linear equations, where the matrix representation of the linear equations is the measurement matrix and it is related to the real compressive sensing system.

One of the most important tasks in compressive sensing is the measurement matrix design. A wide range of developments have analyzed a popular family of theoretical matrices such as binary, sparse, Bernoulli, Fourier, Hadamard, among others [8], [14]–[17]. However, highly structured, sparse, and binary measurement matrices are still a design challenge.

The state-of-the-art optimized sensing matrices lead to dense real-valued matrices without a specific structure. However, these solutions are unfeasible in physical systems [18], [19]. In fact, very often, the sensing modality and the capabilities of the sensing devices limit the types of measurement matrices that can be implemented in a specific application. For instance, in physical systems, elements such as the digital micro-mirror devices (DMD) [20], the coded apertures [21], and the spatial light modulators (SLM) [22] perform the signal codification by allowing parts of the signal

to pass or not. This binary codification, the high dimensionality of the signal, and the array arrangement of the sensor lead to a highly structured and sparse compressive sensing matrix.

A compressive sensing matrix can be modeled by three characteristics: first, weight elements $w_{i,j}^k \in \mathbb{R}$ that involve the fixed intrinsic parameters of the system; second, binary variables $c_{i,j}^k \in \{0,1\}$, which represent the coding device; and third, a set of ordered pairs Ω, such that $|\Omega| \ll MN$, where the non-zero entries are positioned. In summary, each entry of the matrix is modeled as $[\mathbf{\Phi}]_{i,j} = \sum_{k=1}^{P} w_{i,j}^k c_{i,j}^k \ \forall \ (i,j) \in \Omega$, where P is determined by the compressive system. Usually, the measurement matrices have a few non-zero entries in a specific structure given by the system, and they can be represented as binary coding elements. Due to the structured, sparse, and binary nature of the measurement matrix, we will use SBiM to refer it.

The state-of-the-art approaches for designing the sensing matrices (the popular family of theoretical matrices) can not be applied directly to SBiM, due to the unavoidable dependence between the physical configuration and the matrix that models it, that carries strong limitations for any changes in the structure (element locations and weights). Instead it is possible to constrain the binary coding elements of a physical system for accomplishing basic requirements such as linear independence between rows when taking M linearly independent measurements. Achieving this property can be associated with minimizing the number of zero eigenvalues of the matrices $\mathbf{G}{=}\mathbf{\Phi}^T\mathbf{\Phi}$ and $\hat{\mathbf{G}}{=}\mathbf{\Phi}\mathbf{\Phi}^T$, which in turn is related to restrictions on the elements of $\mathbf{\Phi}$.

The first part of this dissertation focuses on the measurement matrix design. In chapter 3, we propose an algorithm to design the distribution of the binary coding elements in an SBiM by constraining the eigenvalues distribution of $\mathbf{G}$ and $\hat{\mathbf{G}}$. This design is tested in different compressive imaging systems.

In chapter 4, we propose three matrix designs for the compressive spectral imaging (CSI) fusion recovery problem. This problem aims to recover a high spatial and spectral resolution (HR) image from two images acquired with two CSI systems. The observed images are assumed to result from spectral or spatial degradations the target HR datacube.

After sensing the underlying spectral image, the reconstruction from the measurement vector can be implemented by solving an optimization problem, that is often solved with an iterative algorithm. Specifically, these algorithms establish a minimizing sequence given by $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \alpha\Omega_k(\mathbf{\Psi}^T\mathbf{\Phi}^T\mathbf{r}^{(k)})$, where $\boldsymbol{\theta}^{(k)}$ is the estimation of $\boldsymbol{\theta}$ at the iteration k, α is the step size, and $\Omega_k : \mathbb{R}^n \to \mathbb{R}^n$ is a function of $\mathbf{\Psi}^T\mathbf{\Phi}^T\mathbf{r}^{(k)}$ called the search function, where $\mathbf{r}^{(k)} = \mathbf{y} - \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\theta}^{(k)}$ is the k-th residual. Depending on the approach, the function Ω_k can be a thresholding operation [23], a denoising function [24], or the function gradient [25].

Since the matrix $\mathbf{\Phi}$ has a particular structure, the product by $\mathbf{\Phi}^T$ in the recovery algorithm can be exploited. The product by $\mathbf{\Phi}^T$ in each iteration returns rough approximations of the underlying signal $\mathbf{x}$, but the structure of the resulting product is given by the measurement system. Then, an interpolation action could be taken to

improve the approximation. An intuitive way to solve the problem is to apply filtering operations before the Φ^T products in each iteration.
When using a significant undersampling rate, there is an infinite number of possible images that can be associated with a particular set of measurements. Thus, it is common to promote realistic properties of the image to recover in order to regularize the problem and obtain a unique solution. One of the most used regularizations consists of enforcing the image of interest to have a sparse representation in a given basis [2]–[4], [26]. Although this technique has shown satisfactory results, the quality of the reconstruction can be further improved by exploiting additional properties of natural images such as their important degree of local structure similarity [27], [28].
The second part of this dissertation focuses on improving the recovery results by including additional knowledge of the spectral imaging systems. Chapter 5 considers the spectral image recovery from a specific compressive system that uses multispectral filter arrays (MSFA). We propose two models for adjustable spatial resolution reconstruction of multispectral images from the selection of MSFA-based sensor measurements neighborhoods. Each spectral filter encodes the data before it impinges onto the sensor using a random dichroic or a random bandpass filter. CS theory is then exploited to recover the underlying 3D spectral data cube from the compressed data captured in a single shot.
In chapter 6, we develop a gradient algorithm that includes a filtering step for solving the compressive problem. The optimization problem is solved by following a minimization sequence that is filtered using a matrix operation. A specific example in compressive spectral imaging is shown, where there are significant improvements in peak signal to noise ratio (PSNR) compared to the unfiltered approach and the message passing approach.
In chapter 7, we introduce a new hierarchical Bayesian model for the reconstruction of compressed hyperspectral images. This model promotes the solution to be sparse in a selected basis and smooth in the image domain. A Gibbs sampler is used to draw samples asymptotically distributed according to its posterior distribution. The generated samples are then used to build estimators of the unknown model parameters. It also allows the model's hyperparameters to be estimated from the observed measurements in an unsupervised framework. The samples generated by the Gibbs sampler can be used to determine measures of uncertainty for the estimates such as the estimation variance of confidence intervals.

GENERAL OBJECTIVE

To design the pseudorandom structure of a compressive spectral imaging system matrix in terms of its maximum eigenvalue concentration and to optimize the iterative reconstruction algorithm used to recover the spectral images in terms of reconstruction quality and computational complexity.

SPECIFIC OBJECTIVES

1. To determine the structure of the matrix model of the main compressive spectral imaging systems.
2. To design a strategy to optimize the pseudorandom structure of the compressive spectral imaging matrix in terms of its maximum eigenvalue concentration.
3. To establish the recovery algorithm that exploits the structure and sparse characteristics of the compressive spectral imaging matrices.
4. To optimize the iterative algorithm in compressive spectral imaging recovery and evaluate it in terms of the reconstruction quality, and computational complexity.
5. To evaluate the performance of the designed matrices and optimized algorithms in terms of quality of reconstruction and by comparison with state-of-the-art recovery algorithms.

IMPACT OF THE RESEARCH AND CONTRIBUTIONS

The contributions of the dissertation are mainly related to the optimal design of the measurement matrix and the improvement of the recovery algorithm in CSI. The innovation of the measurement matrix design is to model it as a highly structured, sparse, and binary-coded matrix (SBiM) and propose an algorithm to optimally select the location of the coded aperture elements in the matrix. The innovation of the recovery algorithm is to take advantage of the spatial smoothness characteristic of the spectral images by including a filtering step in a gradient descent recovery. On the other hand, exploiting the same smoothness characteristic, we propose another optimization problem with a penalty that promotes image smoothness and we solve it following a Bayesian approach.
Specifically, the main contributions are summarized as follows:

1. A measurement matrix design algorithm for the SBiM systems based on the optimal selection of non-zero entries positions. Instead of solving the associated problem of the traditional approaches, we proposed an algorithm that minimizes the associated objective function by optimal selecting the non-zero elements per rows and columns.
2. Three measurement matrix designs for the problem of recovering a spectral image from the measurements of two CSI systems. These methods attempt to impose constraints to the coded aperture elements such that the matrix $\mathbf{G} = \mathbf{\Phi}^T\mathbf{\Phi}$ be as close to the identity matrix as possible, where $\mathbf{\Phi}$ is the measurement matrix that takes into account the two CSI systems.

3. A reconstruction algorithm that recovers the spectral image with an adjustable spatial resolution. This algorithm is based on the assumption that $q \times q$ neighboring pixels, in a $Z \times Z \times L$ spectral image, have a similar spectral response. Then, $q \times q$ neighborhood measurement pixels of a multispectral filter array-based sensor are taken for the reconstruction of a single pixel spectral response in a decimated reconstruction.

4. A gradient algorithm that includes a filtering step for solving the compressive problem. This algorithm is based on the assumption that natural image patches can be represented as linear combinations of its neighbors in the spatial domain and that the operation $\mathbf{\Phi}^T\mathbf{y}$ produces an image approximation with elements near zero, producing gaps.

5. A hierarchical Bayesian model for the reconstruction of the compressed spectral images. This model promotes the solution to be sparse on a selected basis and smooth in the image domain. This solution allows the model hyperparameters to be estimated from the observed measurements in an unsupervised framework.

LIST OF PUBLICATIONS

Published journal papers:

1. Y. Mejia, F. Rojas, and H. Arguello, (2015). “Superpixel Compressive Sensing Recovery of Spectral Images Sensed by Multi-patterned Focal Plane.” Sensors & Transducers, 193 (10), 50–56.

2. Y. Mejia, O. Villarreal, and H. Arguello, (2016). “Adjustable spatial resolution of compressive spectral images sensed by multispectral filter array-based sensors.” Revista Facultad de Ingeniería, Universidad de Antioquia, (78), 89–98.

3. Y. Mejia, and H. Arguello, (2016). “Filtered gradient reconstruction algorithm for compressive spectral imaging.” Optical Engineering, 56(4), 041306.

4. Y. Mejia, and H. Arguello, (2018). “Binary Codification Design for Compressive Imaging by Uniform Sensing.” IEEE Transactions on Image Processing, 27(12), 5775-5786.

Conference papers:

1. Y. Mejia, and H. Arguello “Macropixel compressive sensing reconstruction of spectral images.” The Sixth International Conference on Sensor Device Technologies and Applications, Venice, Italy, (2015).

2. Y. Mejia, and H. Arguello, “Filtered gradient compressive sensing reconstruction algorithm for sparse and structured measurement matrices”, Proc. SPIE 9857, Compressive Sensing V: From Diverse Modalities to Big Data Analytics, 98570F (4 May 2016).

3. Y. Mejia, H. Arguello, F. Costa, J.-Y. Tourneret, and H. Batatia “Bayesian reconstruction of hyperspectral images by using compressed sensing measurements and a local structured prior.” The 42nd IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2017, 984023 (5 March 2017).

1. COMPRESSIVE SPECTRAL IMAGING, CSI

The purpose of this chapter is to present the main concepts that comprise this dissertation and to introduce the previous work on measurement matrix design and recovery algorithms for compressive spectral imaging.

1.1. SPECTRAL IMAGING

A gray-scale image is formed by projecting the light coming from a scene onto a sensor, which is usually a flat chip that detects electromagnetic energy. The sensor output is a 2D array of pixels, where the amount of them is the sensor spatial resolution. In color imaging, the process is the same, only this time the sensor pixels are covered with red, green, and blue filters, in some layout. Each pixel senses only one color, and spatial interpolation is used to obtain the red, green, and blue plane. Recently, imaging has evolved to include many bands encompassing the visible spectrum, the near-infrared, and shortwave infrared bands. The resulting spectral image can be modeled as a 3D datacube where each pixel is a vector of intensities representing a spectral signature. This evolution is aimed at exploiting the fact that the materials comprising the various objects in a scene reflect, scatter, absorb, and emit electromagnetic radiation in ways characteristic of their composition. If the radiation arriving at the sensor is measured at many wavelengths the resulting spectrum can be used to identify the materials in a scene [29].

To acquire hyperspectral data cubes, however, the traditional approaches are based on sequentially sensing 2D images, and composing them into a single data cube. In some spectral imagers, the scene is beam split into the desired wavelength components, for example, using a prism assembly, and each of these images is captured by a separate detector array. Although this method provides the highest spatial resolution, the sensing devices have significant size and weight disadvantages [30]. One of the most intuitive multispectral scanning techniques is the tunable filter, where a complete spectral image is produced after a sequence of exposures by capturing an image of one spectral band at a time. For instance, the spectral image can be sensed by using a filter wheel where some optical filters are installed in a rotatory mechanical structure [31]. Then the rotatory speed of such filter is limited by their mechanical nature. Another method is to change the spectrum of the illumination and take the snapshot in each band; it is used for fragile samples, because one may consider using different exposures under different light source [32]. Most of the methods are related to scanning operations where multiple exposures are used causing motion artifacts.

In all the traditional methods for hyperspectral data acquisition, there is a trade-off between acquisition time and SNR: the faster each band (or line) is captured, the fewer photons are acquired, decreasing SNR. The same problem exists in color

imaging, but in spectral imaging, there is lower energy per band and many more bands. Another difficulty is the huge amount of data that needs to be stored or transmitted.

1.2. COMPRESSIVE SPECTRAL IMAGING AND SYSTEM MATRIX MODEL

As a solution of the huge amount of data for sensing spectral images many architectures have been developed. The compressive sensing architectures are modeled by measurement matrices Φ. There are several approaches for compressing spectral data that can be divided in three remarkable classes, i) spatial coding based CS imagers, such as Coded Aperture Snapshot Spectral Imager (CASSI) [8], ii) spectral coding based spectral imagers, such as Spatio-spectral encoded Compressive Spectral Imager SSCSI [10], or Colored-CASSI [33], and iii) theoretical approaches for compressing spectral imaging data, such as Hyperspectral Coded Aperture (HYCA) [34] or Compressive-Projection Principal Component Analysis approach [35]. These systems are based on the fact that the light in the scene is projected on a prism or diffraction grid and scattered to its different spectral components by using light encoders (e.g. digital micromirror arrays or multispectral filter arrays). The scattered light is projected onto an imager. Traditionally, the scene is scanned line by line, creating a data cube, conversely, CS imagers acquire all data in a single projection by coding and integrating the spectral dimension. The following subsections describe some compressive sensing approaches, that for this work have a special interest, from the three classes mentioned above.

Multispectral filter arrays (spectral coding based imager) Some techniques of spectral imaging use multispectral filters and collect multiple wavelength spectra from a single detector array. Nowadays, the optimization of the optical coatings technologies allows manufacturers to create multipatterned arrays of different optical filters with traditional design and manufacturing methods [33]. The optical coatings production methodology combines modern optical thin film deposition techniques with micro-lithographic procedures. This process enables micron-scale precision patterning of optical thin film dichroic coatings on a single substrate. A dichroic filter is an accurate color filter used to selectively pass light of a small range of wavelengths while reflecting other wavelengths.
Let us model the multispectral filter array (MSFA) as a 3D array $T(x, y, \lambda)$ that codes a spatial-spectral datacube $X_0(x, y, \lambda)$, where (x, y) are the spatial coordinates, and λ is the wavelength. Then, the coded density impacts on the sensor. The coded density integrated into the detector can be expressed as

$$Y(x, y, \lambda) = \int\int T(x', y', \lambda) X_0(x', y', \lambda) h(x - x', y - y') dx' dy', \tag{1.1}$$

where $T(x, y, \lambda)$ is the transmission function representing the MSFA, and $h(x-x', y-y')$ is the optical response of the system.

Each pixel in the sensor is a discretized measurement. The source $X_0(x, y, \lambda)$ can be written in discrete form as $\mathbf{X}_{i,j,k}$ where i and j index the spatial coordinates, and k determines the kth spectral plane. Let us define $T_{i,j,k} \in \{0, 1\}$ as the discretization of the MSFA. Then, the discretized measurements of the MSFA-based sensor can be expressed as

$$Y_{i,j} = \sum_{k=0}^{L-1} X_{i,j,k} T_{i,j,k} + \omega_{i,j}, \tag{1.2}$$

where $Y_{i,j}$ is the intensity at the (i, j)th position of the detector, $i, j = 0, 1, ..., Z - 1$, and the dimensions of the detector are $Z \times Z$, $\mathbf{X}$ is a $Z \times Z \times L$ spectral data cube, and $\omega_{i,j}$ is the read noise of the sensing system, assumed as white Gaussian noise; for convenience and as traditionally, by its independence of the signal intensity.
The measurements $Y_{i,j}$ in (1.2) can be written in matrix notation as

$$\mathbf{y} = \mathbf{\Phi x} + \boldsymbol{\omega}, \tag{1.3}$$

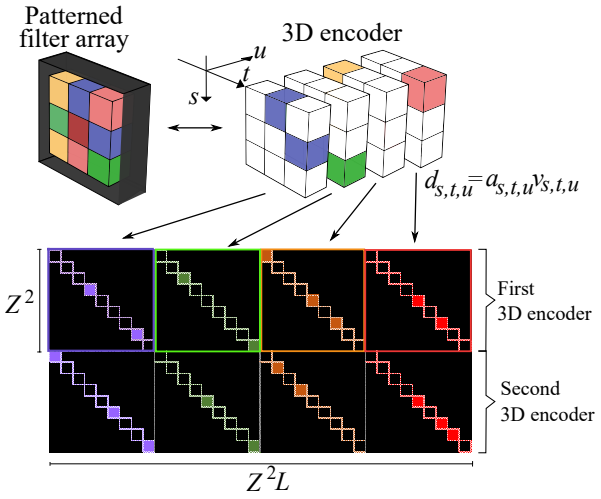
where $\mathbf{y}$ is an Z^2-long vector representation of $Y_{i,j}$, $\mathbf{x} = \text{vect}([\mathbf{x}_0, ..., \mathbf{x}_{L-1}])$ is the vector representation of the data cube $\mathbf{X}$ where $\mathbf{x}_k$ is the vectorization of the kth spectral band.
The output $\mathbf{y}$ in (1.3) can be extended as

$$\mathbf{y} = \underbrace{\begin{bmatrix} \mathbf{diag}(\mathbf{t}_0) & \mathbf{diag}(\mathbf{t}_1) & \cdots & \mathbf{diag}(\mathbf{t}_{L-1}) \end{bmatrix}}_{\mathbf{\Phi}} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{L-1} \end{bmatrix} + \boldsymbol{\omega}, \tag{1.4}$$

where $\mathbf{t}_k$ is the vectorization of the kth MSFA plane, more specifically $(\mathbf{t})_i = T_{(\lfloor i/Z \rfloor, i - \lfloor i/Z \rfloor, k)}$, for $i = 0, \ldots, Z^2 - 1$; $\mathbf{diag}(\mathbf{t}_0)$ is a $Z^2 \times Z^2$ diagonal matrix whose entries are the elements of $\mathbf{t}_k$. Figure 1 depicts the measurement matrix $\mathbf{\Phi}$ for the MSFA-based sensor for $Z = 3$, $L = 4$, and 2 independent acquisitions. In order to increase the number of measurements, it is common in compressive spectral imaging to take several acquisitions with different coding patterns and we will refer to every single acquisition as a snapshot.

Coded Aperture Snapshot Spectral Imager (CASSI) (spectral coding based imager) The concept for CASSI was developed from a generalization of coded aperture spectrometry. Coded aperture spectrometers replace the entrance slit of a dispersive spectrometer with a much wider field stop, inside which is inserted a binary-coded mask. This mask attempts to create a transmission pattern at each column within the slit such that each column transmission code is orthogonal to that of every other column. The encoded light, transmitted by the coded mask within the field stop, is then passed through a standard spectrometer back-end (i.e., collimating lens, disperser, reimaging lens, and detector array). Because the columns of the coded

Figure 1: The matrix $\mathbf{\Phi}$ in (1.4) is shown for $Z = 3$, and $L = 4$, and 2 snapshots. Colored squares represent passing elements whereas the other elements are zero.



mask are orthogonalizable, when they are smeared together by the disperser and multiplexed on the detector array, they can be demultiplexed during postprocessing. The resulting setup allows the system to collect light over a wide aperture without sacrificing the spectral resolution that one would lose by opening wide the slit of a standard slit spectrometer.

Let $\mathbf{X} \in \mathbb{R}^{Z \times Z \times L}$ be a spectral image with spatial resolution $Z \times Z$ and L spectral bands, the CASSI system captures only $K < L$ snapshots of $Z(Z+L-1)$ modulated information of the spectral image for recovering the complete data cube. The main characteristic of the coded apertures in CASSI is the percentage of transmissive elements or transmittance. The ℓth discretized measurement using a different coded aperture T^ℓ can be expressed as

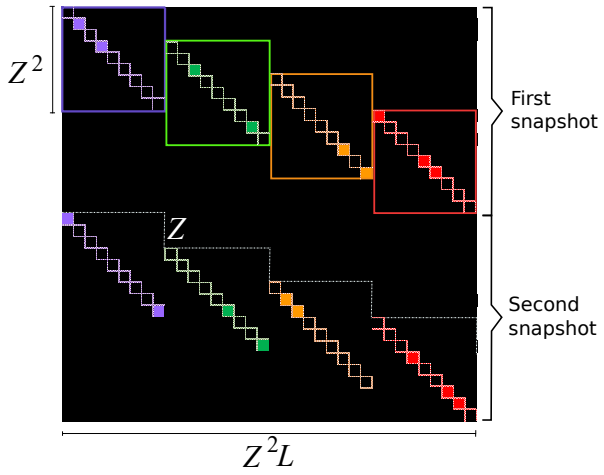
$$Y_{i,j}^{\ell} = \sum_{k=0}^{L-1} X_{i,(j-k),k} T_{i,(j-k)}^{\ell} + \nu_{i,j,k}, \tag{1.5}$$

where $\mathbf{X}_{i,j}^{\ell}$ is the intensity at the (i,j)th position of the detector whose dimensions are $Z \times (Z+L-1)$, and $\nu_{i,j,k}$ is the white Gaussian noise of the sensing system.

The output $Y_{i,j}^{\ell}$ in (1.5) can be written in matrix notation as $\mathbf{y}^{\ell} = \mathbf{\Phi}^{\ell}\mathbf{x} + \boldsymbol{\nu}$, where $\mathbf{y}^{\ell}$ is an $Z(Z+L-1)$-long vector representing Y_{ij}^{ℓ}, $\mathbf{x}$ is the vector representation of the data cube $\mathbf{X}$, and $\mathbf{\Phi}^{\ell}$ is the measurement matrix of the ℓth snapshot. The set of measurements taken by K snapshots can be written as $\mathbf{y} = [(\mathbf{\Phi}^0)^T, (\mathbf{\Phi}^1)^T, \cdots, (\mathbf{\Phi}^{K-1})^T]^T\mathbf{x} = \mathbf{\Phi x}$. Figure 2 shows the matrix $\mathbf{\Phi} \in \mathbb{R}^{Z(Z+L-1)K \times Z^2 L}$ for 2 snapshots. Note that its structure consists of a set of diagonal patterns, that repeat along the horizontal direction, such that one spatial dimension is shifted downward, as many times as the number of spectral bands. Each diagonal element is due to the sampling pattern, that has been vectorized column-wise. Another patterns of diagonal elements are

vertically stacked when several snapshots are considered by the system.

Figure 2: Sensing phenomena representation of the CASSI system. The qth slice of the data cube is coded by a row of the coded aperture and dispersed by the prism. The detector captures the intensity $\mathbf{y}$ by integrating the coded light.



The Colored CASSI (C-CASSI) is a modification of the traditional system, where the coded aperture is a multispectral filter array [33]. Therefore, the sensing matrix structure in C-CASSI allows the system to have more design possibilities.

Spatio-spectral encoded Compressive Spectral Imager SSCSI [10] (spectral coding based imager) The SSCSI compresses high-resolution HS images by combining optical spatial-spectral modulation allowing their reconstruction from a single sensor image. SSCSI employs a diffraction grating to disperse light into the spectral plane and adopts a coded attenuation mask mounted at a slight offset in front of the sensor. The mask modulates the target HS images in both spatial and spectral dimensions before projection into a sensor image. Thus, the coded projection operator is expressed as a sparse modulation matrix with Boolean entries.

Hyperspectral Coded Aperture (HYCA) [34] (theoretical imager) The HYCA system is a compressive sensing framework that computes for each spectral vector a few random projections using Gaussian independent identically distributed (i.i.d.) vectors based on the approaches of physical spatial/spectral coding. HYCA decouples the spectral and spatial domains, which greatly simplifies the reconstruction algorithms; and uses Gaussian i.i.d. vectors to compute the inner products yields incoherent CS matrices, a desirable property for the success of data recovery. Thus, the coded projection operator is expressed as a sparse modulation matrix with Gaussian entries.

1.3. **PREVIOUS WORK ON MATRIX DESIGN**

Recently, many works have focused on the construction and design of measurement matrices based on theoretical constraints. A common strategy to improve the measurement matrix is to solve the following minimization problem $\mathbf{\Phi} = \arg\min_{\hat{\mathbf{\Phi}} \in \mathbb{R}^{M \times N}} \|\mathbf{I} - \mathbf{\Psi}^T \hat{\mathbf{\Phi}}^T \hat{\mathbf{\Phi}} \mathbf{\Psi}\|_F^2$, which is related to the minimization of the mutual coherence of the sensing matrix $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}$. For instance, the work of Elad [36] proposes an iterative algorithm for reducing an average mutual coherence related to a threshold. Li et al. [37] consider the problem of designing the matrix $\mathbf{\Phi}$ given the representation basis $\mathbf{\Psi}$. They obtained a class of closed-form solutions for the mutual coherence problem. The works in [38], [39] proposed finding a robust projection matrix when the sparse representation error is considered by solving the minimization problem $\mathbf{\Phi} = \arg\min_{\hat{\mathbf{\Phi}} \in \mathbb{R}^{M \times N}} \|\mathbf{I} - \mathbf{\Psi}^T \hat{\mathbf{\Phi}}^T \hat{\mathbf{\Phi}} \mathbf{\Psi}\|_F^2 + \lambda \|\hat{\mathbf{\Phi}} \mathbf{E}\|_F^2$, where $\mathbf{E}$ represents the sparse representation error and has to be computed with a sufficient number of training signals. Hong and Zhu [40] propose to design a robust measurement matrix without the requirement of the training dataset by minimizing the problem $\mathbf{\Phi} = \arg\min_{\hat{\mathbf{\Phi}} \in \mathbb{R}^{M \times N}} \|\mathbf{I} - \mathbf{\Psi}^T \hat{\mathbf{\Phi}}^T \hat{\mathbf{\Phi}} \mathbf{\Psi}\|_F^2 + \lambda \|\hat{\mathbf{\Phi}}\|_F^2$.

State-of-the-art optimized sensing matrices are dense real-valued matrices without a specific structure. However, these matrices cannot be used in physical systems [19]. In fact, the sensing modality and the capabilities of the sensing devices limit the types of measurement matrices that can be implemented in a specific application. For instance, in physical systems, elements such as spatial light modulators (SLM) [22] perform the signal codification by blocking or letting pass part of the signal. This binary codification, the high dimensionality of the signal, and the array arrangement of the sensor lead to a highly structured and sparse compressive measurement matrix. Thus, measurement matrices of high-dimensional signals usually have a few non-zero entries with a specific structure given by the physical system that can be modified by the binary coding elements.

1.4. **RECOVERY ALGORITHMS**

In order to recover the spectral image $\mathbf{x}$ from the measurements $\mathbf{y} = \mathbf{\Phi}\mathbf{x}$, a compressive sensing problem can be posed and solved as a convex unconstrained optimization problem [8], [12], [18], [33]. This problem consists of minimizing an objective function that includes a quadratic error term combined with a sparsity-inducing regularization term [1], [3], [5]. When the solution of this convex problem is known in advance to be non-negative, it can be rewritten as [4], [25]

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\text{argmin}} (\tau \mathbf{1}_n - \mathbf{A}^T \mathbf{y})^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x}, \quad \text{s.t} \quad \mathbf{x} \geq 0. \tag{1.6}$$

Equation (1.6) can be seen as a standard bound constrained quadratic problem (BCQP) given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\text{argmin}}\ \mathbf{c}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{B}\mathbf{x} \equiv \underset{\mathbf{x}}{\text{argmin}} F(\mathbf{x}), \quad \text{s.t} \quad \mathbf{x} \geq 0, \tag{1.7}$$

where $\mathbf{c} = (\tau\mathbf{1}_n - \mathbf{A}^T\mathbf{y})$, $\mathbf{1}_n$ is a n-long vector of ones, and $\mathbf{B} = \mathbf{A}^T\mathbf{A}$.
There are several algorithms that solve the compressive sensing problem by Eq. (1.7) [4], [41]–[44].

1.5. **PREVIOUS WORK ON RECOVERY ALGORITHMS**

Related to the denoising approach, the work of Tan, et al. [45] focuses on the approximate message passing (AMP) algorithm for solving the imaging CS problem. The AMP algorithm reconstructs a signal iteratively by performing a scalar denoising within each iteration. In AMP, the denoising function is needed to be carefully chosen to obtain better reconstruction quality. One challenge in applying image denoisers within AMP is that it demands high computational resources to determine the so-called Onsager reaction term in the AMP iteration steps. The Onsager reaction term involves the derivative of the image denoising function, and thus, if an image function does not have a convenient closed form, then the Onsager reaction term can be extremely complex to compute. One complementary work [24] employs AMP with scalar denoisers that are better adapted to the statistics of natural images. Recently, the AMP approach has been applied to the CASSI system where they modified a Wiener filter to fit the particular sensing architecture [46].
The work of Figuereido *et al.* [25] is a well-known algorithm that uses a convex optimization approach. This is a gradient projection (GP) algorithm applied to a quadratic programming formulation of the CS problem, in which the search path from each iteration is obtained by projecting the negative-gradient direction onto the feasible set. The formulation of the problem as a bound constrained quadratic problem (BCQP) allows to identify the product for the transpose of the sensing matrix $\mathbf{\Phi}$ in the iterative process.
There are several gradient descent algorithms that solve the compressive sensing problem, such as the Stochastic Gradient Approach Based on Adaptive Filtering Framework [41], the Nonmonotone adaptive Barzilai-Borwein gradient algorithm [42], the Iterative gradient projection [43], the Fast and accurate first-order method for sparse recovery [44], the Sparse reconstruction by separable approximation [4].
For instance, the approaches of Figueiredo *et al.* [47], [48] for solving the CS problem based on a variable splitting to obtain an equivalent constrained optimization formulation, which is then addressed with an augmented Lagrangian method. The proposed algorithm is an instance of the so-called alternating direction method of multipliers (ADMM), for which convergence has been proved.
Some authors [49]–[51] consider the inversion of CS from a Bayesian perspective. Specifically, from this standpoint they have a prior belief that the θ should be sparse

in a basis and the objective is to provide a posterior belief (density function) for the values of the $\boldsymbol{\theta}$ given that $\mathbf{y}$ are observed from compressive measurements. Rather than providing a point (single) estimate for the weights $\boldsymbol{\theta}$, the Bayesian approach provides provides a full posterior density function, which yields "error bars" on the estimated $\mathbf{x}$. These error bars may be used to give a sense of confidence in the approximation to $\mathbf{x}$, and they may also be used to guide the optimal design of additional CS measurements, implemented with the goal of reducing the uncertainty in $\mathbf{x}$. In addition, the Bayesian framework provides an estimate for the posterior density function of additive noise encountered when implementing the compressive measurements.

2. MATRIX DESIGN IN CSI BY EIGENVALUE ANALYSIS

Recently, an important set of high dimensional signals (HDS) applications has successfully implemented compressive sensing (CS) sensors in which their efficiency depends on physical elements that perform a binary codification over the HDS. The structure of the binary codification is crucial as it determines the HDS sensing matrices. For a correct reconstruction, this class of matrices drastically differs from the dense or i.i.d. assumptions usually made in CS. Therefore, current CS matrix design algorithms are impractical.
It is possible to constrain the binary coding elements of a physical system to better fulfill basic requirements such as linear independence between rows, that allows the system to consider M linearly independent measurements. This property can be satisfied by minimizing the number of zero eigenvalues of the matrices $\mathbf{G}{=}\mathbf{\Phi}^T\mathbf{\Phi}$ and $\hat{\mathbf{G}}{=}\mathbf{\Phi}\mathbf{\Phi}^T$, which in turn is related to restrictions on the elements of $\mathbf{\Phi}$ according to the Gershgorin theorem [52], [53].
In this chapter, we introduce an algorithm for designing the distribution of the binary coding elements in an SBiM by constraining the eigenvalue distribution of $\mathbf{G}$ and $\hat{\mathbf{G}}$. The main contributions of this work are summarized as follows.

1. A practical methodology for designing SBiM that promotes the linear independence of its rows by relating the positions of its binary coding elements to the eigenvalues of $\mathbf{G}$ and $\hat{\mathbf{G}}$. The design constraints lead to a uniform choice of the non-zero elements per row and per column.

2. The proposition of an efficient algorithm for the design of a sensing matrix satisfying these constraints.

3. The proof that this design reduces the zero eigenvalues of the matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$, which increases the number of independent rows of the measurement matrix.

Simulations will show that the proposed optimization strategy improves the performance of image reconstruction in several real CS systems, with a gain in PSNR up to 8 dB.

2.1. STRUCTURED, SPARSE, AND BINARY MEASUREMENT MATRICES (SBIM)

The set of available SBiM entries can be expressed as a sum of products between weight elements that involve the intrinsic parameters of the system, and binary variables, which represent the coding devices. The other entries of the measurement matrix are zero. Specifically, the (i, j) non-zero entry of the measurement matrix $\mathbf{\Phi}$ is expressed as:

$$[\mathbf{\Phi}]_{i,j} = \sum_{k=1}^{P} w_{i,j}^{k} c_{i,j}^{k} \ \forall \ (i,j) \in \Omega \tag{2.1}$$

where $w_{i,j}^{k} \in \mathbb{R}$ is the weights related to the position (i,j) of $\mathbf{\Phi}$, $c_{i,j}^{k} \in \{0,1\}$ is a modifiable binary elements that depends on the coding device, defined as $c_{i,j}^{k} = T(\mathbf{v}_0^k, ..., \mathbf{v}_{D-1}^k)$, where T is a function of D coding vectors that maps them to a single binary element, and Ω is a set of ordered pairs that contains the non-zero positions of the measurement matrix $\mathbf{\Phi}$. Note that the set of positions Ω is fixed depending on the hardware of the studied system.
In order to explain how (2.1) can be constructed for real systems, the next subsections study specific examples of SBiM resulting from compressive spectral imaging and compressive X-ray tomography.

2.1.1. **CSI Systems modeled as SBiM** CSI systems capture spatio-spectral data simultaneously in multiple snapshots. For instance, the Patterned Filter Array (PFA) and the Colored Coded Aperture Snapshot Spectral Imaging (C-CASSI) approaches are CSI systems that have been widely used and have provided outstanding image reconstruction results [54], [55]. In the next subsections, we will propose the measuremente matrix model from the PFA and C-CASSI and how it is related to the matrix model (2.1).

Patterned Filter Array (PFA)

Let $\mathbf{X} \in \mathbb{R}^{Z \times Z \times L}$ be a spectral image with spatial resolution $Z \times Z$ and L spectral bands. For acquiring spectral images, the coding devices in the PFA approach are optical elements that perform spectral filtering and collect multiple wavelength spectra for each pixel of the detector array. These devices can be modeled as 3D encoders whose elements are given by $d_{s,t,u} = a_{s,t,u} v_{s,t,u}$, where $a_{s,t,u} \in [0,1]$ is an attenuation coefficient and $v_{s,t,u} \in \{0,1\}$ is an on-off element that represents the wavelength codification of the light. Let $X_{s,t,u}$ be an element of the spatio-spectral data cube, where $s,t = 0, ..., Z-1$ are used for the spatial coordinates and $u = 0, ..., L-1$ determines the uth spectral plane. The discretized measurements of the PFA approach for a single snapshot obtained by a particular encoder can be expressed as

$$[\mathbf{Y}]_{s,t} = \sum_{u=0}^{L-1} X_{s,t,u} d_{s,t,u} + \eta_{s,t}, \tag{2.2}$$

where $[\mathbf{Y}]_{s,t}$ is the intensity at the position (s,t) of the detector and $\eta_{s,t}$ is the white noise. The measurements $\mathbf{Y}$ in (2.2) can be written as $\mathbf{y} = \mathbf{R}\mathbf{x} + \boldsymbol{\eta}$, where $\mathbf{y} \in \mathbb{R}^{Z^2}$ is a vector representation of $\mathbf{Y}$, $\mathbf{x}$ is the vector representation of the data cube $\mathbf{X}$, and $\mathbf{R}$ represents the measurement matrix of a particular PFA. Each element of the

matrix $\mathbf{R}$ is related to the 3D encoder elements as

$$[\mathbf{R}]_{g,h} = d_{g',h',h''} \tag{2.3}$$

where $g' = g - \lfloor g/Z \rfloor Z$, $h' = \lfloor h/Z \rfloor$, and $h'' = \lfloor h/Z^2 \rfloor$ with $g = 0, ..., Z^2 - 1$, $h = 0, ..., N - 1$ and $N = Z^2 L$. Eq. (2.3) can be interpreted as the mapping of the 3D encoder array to the 2D matrix $\mathbf{R}$.
The sensing process can be generalized to D snapshot projections having different PFAs. The set of measurements obtained with D snapshot projections is defined as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}^0 \\ \vdots \\ \mathbf{y}^{D-1} \end{bmatrix} = \begin{bmatrix} \mathbf{R}^0 \\ \vdots \\ \mathbf{R}^{D-1} \end{bmatrix} \mathbf{x} = \mathbf{\Phi}\mathbf{x}. \tag{2.4}$$

Therefore, each element of the measurement matrix $\mathbf{\Phi}$ in (2.4) can be expressed as

$$\begin{aligned} [\mathbf{\Phi}]_{i,j} &= [\mathbf{R}^{\gamma_i}]_{i',j} \\ &= d^{\gamma_i}_{i'',j',j''} \\ &= a^{\gamma_i}_{i'',j',j''} v^{\gamma_i}_{i'',j',j''} \quad \forall\, (i,j) \in \Omega \end{aligned} \tag{2.5}$$

where $\gamma_i = \lfloor i/Z^2 \rfloor$, $i' = i - \lfloor i/Z \rfloor Z^2$, $i'' = i' - \lfloor i'/Z \rfloor Z$, $j' = \lfloor j/Z \rfloor$, and $j'' = \lfloor j/Z^2 \rfloor$ with $i = 0, ..., DZ^2$, $j = 0, ..., N - 1$, and $d^{\ell}_{s,t,u}$ refers to the (s,t,u) element of the encoder for the ℓ snapshot. Comparing (2.5) with (2.1), the PFA approach can be seen as a SBiM where $P = 1$, $w_{i,j} = a^{\gamma_i}_{i'',j',j''}$, and $c_{i,j} = v^{\gamma_i}_{i'',j',j''}$. The set Ω forms a set of block diagonal matrices related to the codification per spectral band and different PFAs.

Colored Coded Aperture Snapshot Spectral Imaging (C-CASSI) System

In the C-CASSI system, the coded information is spectrally shifted by a dispersive element, and a 2D focal plane array detects the coded and shifted projections. Specifically, a discretized snapshot from C-CASSI can be expressed as

$$[\mathbf{Y}]_{s,t} = \sum_{u=0}^{L-1} X_{s,(t-u),u} d_{s,(t-u),u} + \eta_{s,t}, \tag{2.6}$$

where $[\mathbf{Y}]_{s,t}$ is the intensity at the position (s,t) of the detector whose dimensions are $Z \times (Z + L - 1)$. Notice that (2.6) can be written in matrix notation as $\mathbf{y} = \mathbf{R}\mathbf{x} + \boldsymbol{\eta}$, where $\mathbf{y} \in \mathbb{R}^E$, with $E = Z(Z + L - 1)$, is the vector representation of $\mathbf{Y}$ and $\mathbf{R}$ is the measurement matrix of a single snapshot.
As in the previous subsection, each element of the matrix $\mathbf{R}$ is related to the 3D encoder, by taking into account the dispersion effect as follows

$$[\mathbf{R}]_{g,h} = d_{g',\hat{h},h''} \tag{2.7}$$

where $g' = g - \lfloor g/Z \rfloor Z$, $\hat{h} = \lfloor h/Z \rfloor - \lfloor h/Z^2 \rfloor Z$, and $h'' = \lfloor h/Z^2 \rfloor$, with $g = 0, ..., E-1$, $h = 0, ..., N-1$, and $N = Z^2 L$. Defining the set of measurements associated with the D different snapshots as in (2.4), each element of the measurement matrix $\mathbf{\Phi}$ can be expressed as

$$\begin{aligned} [\mathbf{\Phi}]_{i,j} &= [\mathbf{R}^{\rho_i}]_{\hat{\imath}',j} \\ &= d^{\rho_i}_{\hat{\imath}'',j',j''} \\ &= a^{\rho_i}_{\hat{\imath}'',j',j''} v^{\rho_i}_{\hat{\imath}'',j',j''} \quad \forall\, (i,j) \in \Omega \end{aligned} \tag{2.8}$$

where $\rho_i = \lfloor i/E \rfloor$, $\hat{\imath}' = i - \lfloor i/E \rfloor E$, $i'' = \hat{\imath}' - \lfloor \hat{\imath}'/Z \rfloor Z$, $j' = \lfloor j/Z \rfloor$, and $j'' = \lfloor j/Z^2 \rfloor$, with $i = 0, ..., DE-1$, and $j = 0, ..., N-1$. Comparing (2.8) with (2.1), the C-CASSI approach can be seen as a SBiM with $P = 1$, $w_{i,j} = a^{\rho_i}_{\hat{\imath}'',j',j''}$, and $c_{i,j} = v^{\rho_i}_{\hat{\imath}'',j',j''}$. The set Ω forms a set of stacked block diagonal matrices taking into account the dispersion of one spatial dimension.

2.1.2. **Compressive X-ray tomography** The X-ray transmission imaging model for a single source is established by the Beer-Lambert law as $I = I_0 e^{-\int_0^\infty a(l)dl}$, where I_0 is the intensity of an X-ray generated by the X-ray source passing through the object, I is the measured intensity, and $a(l)$ is the linear attenuation coefficient varying in the l location [56].

A commonly used physical configuration is the fan-beam geometry, with lines arranged into fans emanating from X-ray sources distributed over a circle around the origin [57]. If an X-ray source is located at position $\mathbf{s}$ and illuminates an object in direction φ, the data function for the imaging model is given by $\varpi(\mathbf{s}, \varphi) = -\ln(I/I_0)$. Therefore, the Beer-Lambert law can be rewritten as $\varpi(\mathbf{s}, \varphi) = \int_0^\infty x(\mathbf{s} + l\varphi)dl$, where x corresponds to the two-dimensional object function, i.e., the X-ray linear attenuation map.

The imaging model is discretized since only a discrete number of measurements can be acquired. Hence, the two-dimensional data array is represented by a discrete number of unknowns $[\mathbf{x}]_j$, where $j = 0, 1, ..., N-1$ corresponds to the jth pixel attenuation coefficient forming the object $\mathbf{x} \in \mathbb{R}^N$, where $N = Z_1 Z_2$ is the number of object pixels. The detector is a linear array of $M_1 \ll N$ elements located in an arc under the object.

The measurements are integrated at each detector element leading to $[\mathbf{y}]_i$, for the ith detector measurement, with $i = 0, 1, ..., M_1 - 1$. An X-ray fan beam with a single source k can be written as a system of linear equations $\mathbf{y}_k = \mathbf{Q}_k \mathbf{x}$, where the matrix $\mathbf{Q}_k \in \mathbb{R}^{M_1 \times N}$ is the system matrix of a single source. The entries of the matrix $\mathbf{Q}_k$ perform the mapping of the fan beam energy radiating from the X-ray source onto the detector, i.e., each entry $[\mathbf{Q}_k]_{i,j}$ represents the portion of area of the jth pixel that is irradiated by the X-ray measured at the ith detector element.

Compressive X-ray tomography fan beams multiplex measurements from several sources onto the detector. 1D coded apertures are arranged in front of the fan-beam sources to code the energy of each X-ray source, giving a distinct coded projection

onto the detector array. The size of the coded aperture elements is fixed to obtain a one-to-one correspondence between the detector elements.
Each source has a projection $\mathbf{y}_k$ and a system matrix $\mathbf{Q}_k$. The measurements obtained after irradiating several sources at a time instant in a specific angle m are given by

$$\mathbf{y}^m = \mathbf{R}^m \mathbf{x} \tag{2.9}$$

where the entries of the matrix $\mathbf{R}^m$ are given by

$$[\mathbf{R}^m]_{g,h} = \sum_{k=0}^{\hat{P}-1} [\mathbf{Q}_k^m]_{g,h} [\mathbf{t}_k^m]_g \tag{2.10}$$

and where $\mathbf{t}_k^m$ refers to the coded aperture of the kth source at the mth angle, and $\mathbf{Q}_k^m$ is the system matrix of the kth source at the mth angle.
The complete set of measurements for Λ different angles is denoted as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}^0 \\ \mathbf{y}^1 \\ \vdots \\ \mathbf{y}^{\Lambda-1} \end{bmatrix} = \begin{bmatrix} \mathbf{R}^0 \\ \mathbf{R}^1 \\ \vdots \\ \mathbf{R}^{\Lambda-1} \end{bmatrix} \mathbf{x} = \mathbf{\Phi}\mathbf{x}. \tag{2.11}$$

Each entry of the matrix $\mathbf{\Phi}$ is related to the $\mathbf{R}^m$ and $\mathbf{Q}_k^m$ entries as

$$[\mathbf{\Phi}]_{i,j} = [\mathbf{R}^{\xi_i}]_{\zeta_i,j} = \sum_{k=0}^{\hat{P}-1} [\mathbf{Q}_k^{\xi_i}]_{\zeta_i,j} [\mathbf{t}_k^{\xi_i}]_{\zeta_i} \quad \forall \ (i,j) \in \Omega \tag{2.12}$$

where $\xi_i = \lfloor i/M_1 \rfloor$, $\zeta_i = i - \lfloor i/M_1 \rfloor M_1$, and Ω is the set of modifiable positions obtained by specifying the hardware setting. Comparing (2.12) with (2.1), the compressive X-ray tomography approach can be seen as a SBiM where $P = \hat{P}$, $w_{i,j}^k = [\mathbf{Q}_k^{\xi_i}]_{\zeta_i,j}$, and $c_{i,j}^k = [\mathbf{t}_k^{\xi_i}]_{\zeta_i}$.

2.2. **MEASUREMENT MATRIX DESIGN**

The characteristics of highly structured, sparse, and binary-coded (SBiM) matrices lead to many constraints preventing state-of-the-art designs based on coherence to be applied directly. Instead, it is possible to constrain the binary coding elements of a physical system to better fulfill the linear independence requirement between rows. For instance, Table 1 shows the percentage of linearly dependent rows for 10 random realizations of 2048×4096 PFA measurement matrices and 2240×4096 C-CASSI measurement matrices. We can note that there is a percentage between 10% to 30% of linearly dependent rows in these matrices. Maximizing the number of linearly independent rows (columns) of $\mathbf{\Phi}$ or equivalently maximizing its rank leads a better-conditioned matrix and hence a better sensing system.

Table 1: Mean percentage of linearly dependent rows for 10 random realizations of spectral imagers matrices

SBiM	% Linearly dependent rows
PFA	10
C-CASSI	30

For maximizing the rank of $\mathbf{\Phi}$, it is possible to constrain the binary coding elements of a physical system. This property can be satisfied by minimizing the number of zero eigenvalues of the matrices $\mathbf{G}=\mathbf{\Phi}^T\mathbf{\Phi}$ and $\hat{\mathbf{G}}=\mathbf{\Phi}\mathbf{\Phi}^T$, which in turn is related to restrictions on the elements of $\mathbf{\Phi}$ according to the Gershgorin theorem [52], [53].
Let define the matrices $\mathbf{G}=\mathbf{\Phi}^T\mathbf{\Phi}$ and $\hat{\mathbf{G}}=\mathbf{\Phi}\mathbf{\Phi}^T$. There exist a relationship between the eigenvalues of the matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$ and the number of linearly independent rows (columns) of $\mathbf{\Phi}$. Specifically, the number of non-zero eigenvalues of $\mathbf{G}$ and $\hat{\mathbf{G}}$ is equal to the number of linearly independent rows (columns) of $\mathbf{\Phi}$ [52]. As a consequence, we design the matrix $\mathbf{\Phi}$ by minimizing the number of zero-valued eigenvalues of the matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$. To obtain this minimization, the Gershgorin theorem is used to find lower bounds for the eigenvalues of the $\mathbf{G}$ and $\hat{\mathbf{G}}$ matrices. The zero eigenvalues can be reduced by maximizing these lower bounds. The lower bounds are expressed as functions of the entries of the matrix $\mathbf{\Phi}$. Surprisingly, it is established that constraining the sum of the entries of each row in the matrix $\mathbf{\Phi}$ leads to the desired optimization.
Specifically, the Gershgorin theorem for the matrix $\mathbf{G}$ is given by

Theorem 2.1 (Gershgorin Theorem [53]) *Every eigenvalue λ of a matrix $\mathbf{G} \in \mathbb{R}^{N\times N}$ satisfies:*

$$|\lambda - [\mathbf{G}]_{i,i}| \leq \sum_{\substack{j=0 \\ j\neq i}}^{N-1} |[\mathbf{G}]_{i,j}| \qquad i \in \{0,1,...,N-1\}. \tag{2.13}$$

Denoting as $H_i = \sum_{j=0,j\neq i}^{N-1} |[\mathbf{G}]_{i,j}|$, the Gershgorin Theorem allows the eigenvalues of $\mathbf{G}$ to be related to its entries as follows

$$[\mathbf{G}]_{i,i} - H_i \leq \lambda \leq [\mathbf{G}]_{i,i} + H_i. \tag{2.14}$$

Notice that given $\mathbf{G} = \mathbf{\Phi}^T\mathbf{\Phi}$ all the eigenvalues of $\mathbf{G}$ satisfies $\lambda_i \geq 0$. Also, notice that Eq. (2.14) indicates that if the terms $\{H_i\}_{i=1}^N$ are minimized, the eigenvalues of $\mathbf{G}$ are close to $[\mathbf{G}]_{i,i}$. If $[\mathbf{G}]_{i,i}$ are constrained to be constants, the lower bounds for the eigenvalues of $\mathbf{G}$ are maximized by minimizing the terms H_i. This constraint can be satisfied for most physical systems by limiting the light transmittance.

Notice that the matrix $\mathbf{\Phi}$ represents a physical system such that the entries $[\mathbf{\Phi}]_{r,k} \geq 0$. Also, the element (k,l) of $\mathbf{G} = \mathbf{\Phi}^T\mathbf{\Phi}$ is given by

$$[\mathbf{G}]_{k,l} = \sum_{r=0}^{M-1} [\mathbf{\Phi}]_{r,k}[\mathbf{\Phi}]_{r,l} \tag{2.15}$$

then $[\mathbf{G}]_{k,l} \geq 0$ is the sum of positive numbers. The maximization of the lower bounds or equivalently the minimization of $H_i = \sum_{k=0,k\neq l}^{N-1} |[\mathbf{G}]_{k,l}|$ can be obtained by minimizing the entries $[\mathbf{G}]_{k,l}$ or equivalently by selecting a codification such that the products $[\mathbf{\Phi}]_{r,k}[\mathbf{\Phi}]_{r,l}$ are minimized.
Given the matrix $\mathbf{\Phi}$ is highly structured and sparse, it is convenient to express it in terms of its non-zero elements. Let $\hat{\chi}_j$ represents the set of non-zero indices in the jth column, $(\hat{\chi}_j)_r$ extracts the rth element of the set $\hat{\chi}_j$, and $\hat{\chi}_{kl} = \hat{\chi}_k \cap \hat{\chi}_l$ is the set of indices in common of non-zero elements in columns k and l. Therefore, Eq. (2.15) can be rewritten as:

$$[\mathbf{G}]_{k,l} = \sum_{r=0}^{|\hat{\chi}_{kl}|-1} [\mathbf{\Phi}]_{(\hat{\chi}_{kl})_r,k}[\mathbf{\Phi}]_{(\hat{\chi}_{kl})_r,l} \tag{2.16}$$

where $|\hat{\chi}_{kl}|$ is the cardinality of the set $\hat{\chi}_{kl}$.
Let us define the binary matrix $\overline{\mathbf{\Phi}}$ related to $\mathbf{\Phi}$ as

$$[\overline{\mathbf{\Phi}}]_{i,j} = \begin{cases} 1 & \text{if } [\mathbf{\Phi}]_{i,j} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.17}$$

whose elements are ones in the positions of the $\mathbf{\Phi}$ non-zero elements and zeros otherwise. For simplicity, define the number of non-zero indices in the jth column of $\mathbf{\Phi}$ as

$$[\mathbf{p}]_j = |\hat{\chi}_j| = \sum_{i=0}^{M-1} [\overline{\mathbf{\Phi}}]_{i,j}. \tag{2.18}$$

Notice that the cardinality of the set $\hat{\chi}_{kl}$ is at most the minimum between the cardinalities of the sets $\hat{\chi}_k$ and $\hat{\chi}_l$, additionally, let w_{max} represents the maximum value of the entries of $\mathbf{\Phi}$. Therefore, the values of the elements in $\mathbf{G}$, $[\mathbf{G}]_{k,l}$, in (2.16) are bounded by

$$[\mathbf{G}]_{k,l} \leq \sum_{r=0}^{\arg\min\{[\mathbf{p}]_k-1,[\mathbf{p}]_l-1\}} w_{max}^2 = w_{max}^2 \min\{[\mathbf{p}]_k, [\mathbf{p}]_l\}. \tag{2.19}$$

Equation (2.19) shows that the bound of the non-diagonal elements $[\mathbf{G}]_{k,l}$ is proportional to the number of non-zero elements by column. Since the number of non-zero elements can be modified by the coding variables, the elements $[\mathbf{G}]_{k,l}$ can be reduced by minimizing the number of non-zero elements in every column of $\mathbf{\Phi}$
Consider that the total number of non-zero elements is constrained by the measurement system to be a constant U, more specifically $\sum_{j=0}^{N-1}[\mathbf{p}]_j = U$. This constraint

is related to the quantity of light that passes through the compressive acquisition system

The optimization of the eigenvalues of $\mathbf{G}$ is equivalent to reduce the terms $[\mathbf{p}]_j$ in (2.19) which is equivalent to the following optimization problem

$$\begin{aligned} &\underset{\mathbf{p},\kappa}{\arg\min} \quad \kappa \\ &\textit{subject to} \quad \begin{array}{c} \kappa \geq [\mathbf{p}]_j, \\ \sum_{j=0}^{N-1}[\mathbf{p}]_j = U, \end{array} \end{aligned} \tag{2.20}$$

where κ is an auxiliary variable that works as a threshold for every element of $\mathbf{p}$. The solution of (2.20) promotes that all the elements of the vector $\mathbf{p}$ to be equal, then $[\mathbf{p}]_0 = [\mathbf{p}]_1 = \ldots = [\mathbf{p}]_{(N-1)} = U/N$.

Computing the entries of $\hat{\mathbf{G}} = \mathbf{\Phi}\mathbf{\Phi}^T$, defining the vector $\mathbf{q}$ as the number of non-zero elements in the rows of $\mathbf{\Phi}$, such that $[\mathbf{q}]_i = \sum_{j=0}^{N-1}[\overline{\mathbf{\Phi}}]_{i,j}$, and determining a similar bound as in (2.19) yields to the optimization of the eigenvalues of $\hat{\mathbf{G}}$ which is equivalent to the following optimization problem

$$\begin{aligned} &\underset{\mathbf{q},\hat{\kappa}}{\arg\min} \quad \hat{\kappa} \\ &\textit{subject to} \quad \begin{array}{c} \hat{\kappa} \geq [\mathbf{q}]_j, \\ \sum_{j=0}^{M-1}[\mathbf{q}]_j = U, \end{array} \end{aligned} \tag{2.21}$$

where $\hat{\kappa}$ is an auxiliary variable. The solution of (2.21) promotes that $[\mathbf{q}]_0 = [\mathbf{q}]_1 = \ldots = [\mathbf{q}]_{(M-1)} = U/M$.

To accomplish the solutions of the optimization problems (2.20) and (2.21), that is, the number of elements per rows is equal to U/N and the number of elements per columns is equal to U/M is not straightforward, because of the structure of the SBiM, since adjusting the non-zero entries by columns modifies at the same time the non-zero entries by rows, and vice versa.

Therefore, to approximate the solutions of (2.20) and (2.21) we formulate the following optimization problem

$$\underset{\mathbf{p},\mathbf{q}}{\arg\min} \quad \|\mathbf{p}-(U/N)\mathbf{1}_N\|_2^2+\|\mathbf{q}-(U/M)\mathbf{1}_M\|_2^2, \tag{2.22}$$

where $\mathbf{1}_N$ is defined as an N-long vector of ones. In order to associate (2.22) with the measurement matrix design, we use the binary matrix $\overline{\mathbf{\Phi}}$ defined in (2.17). Notice that $[\mathbf{p}]_j=\sum_{i=0}^{M-1}[\overline{\mathbf{\Phi}}]_{i,j}=[\mathbf{1}_M^T\overline{\mathbf{\Phi}}]_j$, and $[\mathbf{q}]_i=\sum_{j=0}^{N-1}[\overline{\mathbf{\Phi}}]_{i,j}=[\overline{\mathbf{\Phi}}\mathbf{1}_N^T]_i$. Therefore, (2.22) can be expressed in terms of $\overline{\mathbf{\Phi}}$ as

$$\begin{aligned} \underset{\overline{\mathbf{\Phi}}}{\arg\min} f(\overline{\mathbf{\Phi}}) = &\left\|\mathbf{1}_M^T\overline{\mathbf{\Phi}}-(U/N)\mathbf{1}_N^T\right\|_2^2+\left\|\overline{\mathbf{\Phi}}\mathbf{1}_N-(U/M)\mathbf{1}_M\right\|_2^2 \\ &\textit{subject to} \quad [\overline{\mathbf{\Phi}}]_{i,j} = 0 \ \ \forall \ \ (i,j) \in \Omega^C \end{aligned} \tag{2.23}$$

where Ω^C is the complement of the set Ω and imposes a restriction on the modifiable positions of $\overline{\mathbf{\Phi}}$.

The design criteria in (2.23) can be referred as a uniform sensing. A non-zero entry $[\overline{\mathbf{\Phi}}]_{i,j}$ means that the jth element of the signal is sensed by the ith detector element. Therefore, each column of $\overline{\mathbf{\Phi}}$ is related to a particular signal element, hence, the sum per column indicates the number of times a particular signal element is measured. The minimization of the first term of (2.23) is related to uniform sensing for each signal element, that is, each element should be sensed equal times. Additionally, since each row of $\overline{\mathbf{\Phi}}$ is related to the linear combination of the signal elements that impinge onto a certain detector, the sum per row indicates the number of signal elements measured by the aforementioned detector. Consequently, minimizing the second term of (2.23) is equivalent to sense uniformly the signal. More specifically, we refer to the uniform sensing in the sense that $E\{\sum_i [\overline{\mathbf{\Phi}}]_{i,j}\} = U/N \ \forall \ j$ and $E\{\sum_j [\overline{\mathbf{\Phi}}]_{i,j}\} = U/M \ \forall \ i$, where $E\{.\}$ is the expected value.

2.3. MEASUREMENT MATRIX OPTIMIZATION ALGORITHM

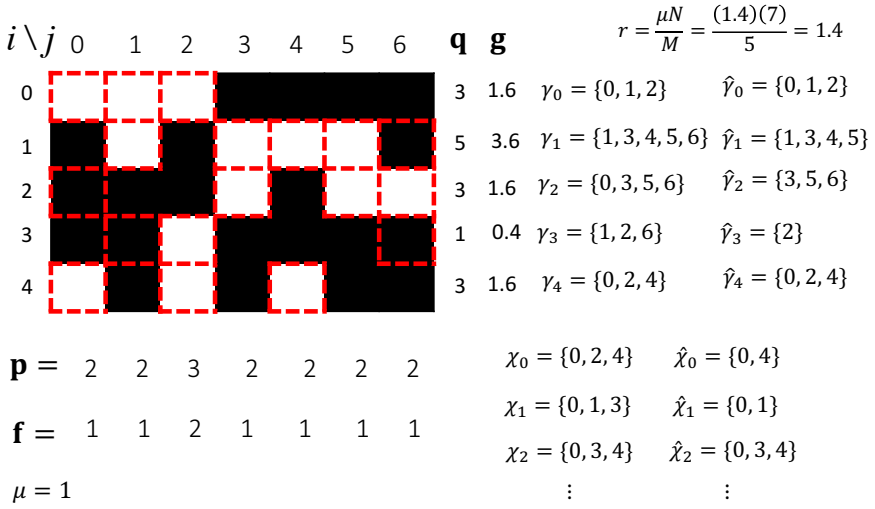
This section presents the algorithm for designing the measurement matrix $\mathbf{\Phi}$ such that the locations of the non-zero elements aim to accomplish the constraints for uniform sensing that were proposed in the last section. We propose to derive two processes of rearranging the entries of $\overline{\mathbf{\Phi}}$ per rows and per columns to satisfy the optimization problem in (2.23). Define χ as the set of sets with the modifiable positions of the $\overline{\mathbf{\Phi}}$ matrix by columns, such that χ_j extracts the set of modifiable indices of the jth column. Similarly, define Υ as the set of sets with the modifiable positions per rows, such as Υ_i is the set of modifiable indices of the ith row. Let us define the desired number of non-zero elements per column as $\mu = U/N$. Notice that minimizing the first term of (2.23) is equivalent to the elements of $\mathbf{p}$ to satisfy $|[\mathbf{p}]_i - \mu| < \sigma$, where σ is a small tolerance value.

The rearranging processes mentioned before consist of selecting randomly within the modifiable positions per columns until completing the non-zero elements average μ in $\hat{N}$ columns, then selecting randomly within the modifiable positions per rows until completing the non-zero elements average r in $\hat{M}$ rows, these processes are performed as many times as indicated by the input parameter $iter$. Algorithm 1, named Coding Matrix Design (CMD), performs these rearranging processes by following probability mass functions (pmf) related to the current number of ones and by using the sets of modifiable positions χ and Υ. In line 2, the algorithm randomly initializes the sets of non-zero positions $\hat{\chi}$ and $\hat{\Upsilon}$. In lines 3 and 6, the vectors $\mathbf{p}$ and $\mathbf{q}$ are formed by computing the cardinality of the sets $\hat{\chi}_j$ and $\hat{\Upsilon}_i$, respectively, also the vectors $\mathbf{f}$ and $\mathbf{g}$ that work as pmf for selecting columns and rows positions. Notice that line 8 performs the rearranging by columns and line 9 performs the process by rows. After updating the current non-zero positions, $\overline{\mathbf{\Phi}}$ is set in line 10.

Figure 3 shows an specific example of the defined variables for a 5×7 matrix $\mathbf{\Phi}$, where the dotted red squares represent the modifiable positions, the white squares are ones, and the black squares are zeros. In this example, the desirable number of elements per columns is established at $\mu = 1$. The calculations of the number of

elements per rows r, the definition of the sets χ_j, Υ_i, $\hat{\chi}_j$, $\hat{\Upsilon}_i$ and the vectors $\mathbf{p}$, $\mathbf{f}$, $\mathbf{q}$, $\mathbf{g}$ are also included in the example.

Figure 3: Example of defined variables for a 5×7 matrix $\mathbf{\Phi}$, the dotted red squares represent the modifiable positions of the matrix, the white squares are ones, and the black squares are zeros. In this example, the desirable number of elements per columns is established at $\mu = 1$.



The function Rearrange, detailed in Algorithm 2, is used for selecting the columns or rows, depending on the case, that least satisfy the optimization function (2.23), and rearranging their non-zero elements to better satisfy it. Specifically, for the selection part, line 4 estimates a pmf, $\hat{\mathbf{f}}$, that gives the highest probability to the positions with greater differences between the $\mathbf{p}$ elements and the desired μ. In line 6, the position j is selected by following the $\hat{\mathbf{f}}$ and the **while** loop in line 7 guarantees that the element to improve is out of the interval $[\mu - \sigma, \mu + \sigma]$. For the rearranging stage, the difference d between μ and the current $[\mathbf{p}]_b$ is computed in line 11. Depending on the sign of d, the next lines will put this number of elements or eliminate them in the b position.

In lines 12-15 of Algorithm 2 other input parameters for the function that put or eliminates non-zero elements are defined. In line 12, an indicator variable, I, is assigned related to the sign of d. A positive d means that the number of elements in the b column (row) is less than μ, therefore, the next steps will assign d non-zero elements to the b column (row). For that, the set $\mathit{\Pi}$ is established as the difference between the sets χ_b and $\hat{\chi}_b$, that is, the modifiable elements that are still zero. In line 15, the function Putd assigns d new elements to the non-zero sets. A negative d means that the number of elements in the b column (row) is larger than μ, therefore, the next steps will eliminate d non-zero elements to the specific sets, here, the set $\mathit{\Pi}$ is established as the set of non-zero elements $\hat{\chi}_b$.

Algorithm 3 shows the function of putting or eliminating $|d|$ ones in the bth column (row), called Putd. In lines 1-2, the pmf $\hat{\mathbf{g}}$ is defined in order to select the positions

Algorithm 1: Coding Matrix Design,*CMD*

input : $\mu, \sigma, M, N, \hat{M}, \hat{N}, \chi, \Upsilon, iter$
output: χ, Υ

% Assign ones randomly to some modifiable positions;
$\{\hat{\chi}, \hat{\Upsilon}\} \leftarrow$initialize($\chi, \Upsilon$);
for $j \leftarrow 1$ **to** N **do** $[\mathbf{p}]_j \leftarrow |\hat{\chi}_j|; \quad [\mathbf{f}]_j \leftarrow |[\mathbf{p}]_j - \mu|;$
% Estimates the number of ones per rows;
$r \leftarrow (\mu N)/M;$
for $i \leftarrow 1$ **to** M **do** $[\mathbf{q}]_i \leftarrow |\hat{\Upsilon}_i|; \quad [\mathbf{g}]_i \leftarrow |[\mathbf{q}]_i - r|;$
for $k \leftarrow 1$ **to** $iter$ **do**
    % Optimize by columns (see Algorithm 2) $\{\mathbf{p}, \alpha, \hat{\chi}, \chi, \mathbf{q}, \hat{\Upsilon}, \Upsilon\} =$ Rearrange($\hat{N}, \mathbf{p}, \mu, \mathbf{f}, \sigma, \hat{\chi}, \chi, \mathbf{q}, \hat{\Upsilon}, r$);
    % Optimize by rows (see Algorithm 2) $\{\mathbf{q}, \beta, \hat{\Upsilon}, \Upsilon, \mathbf{p}, \hat{\chi}, \chi\} =$ Rearrange($\hat{M}, \mathbf{q}, r, \mathbf{g}, \sigma, \hat{\Upsilon}, \Upsilon, \mathbf{p}, \hat{\chi}, \mu$);
$\overline{\mathbf{\Phi}} \leftarrow$setPhi($\hat{\chi}, \hat{\Upsilon}, M, N$);

Algorithm 2: Rearrange

input : $\hat{N}, \mathbf{p}, \mu, \mathbf{f}, \sigma, \hat{\chi}, \chi, \mathbf{q}, \hat{\Upsilon}, r$
output: $\mathbf{p}, \alpha, \hat{\chi}, \chi, \mathbf{q}, \hat{\Upsilon}, \Upsilon$

% Compute the lenght of $\mathbf{p}$;
$N \leftarrow \dim(\mathbf{p})$;
$\alpha \leftarrow \sum_{j=1}^{N} |[\mathbf{p}]_j - \mu|;$
$\hat{\mathbf{f}} \leftarrow \mathbf{f}/\alpha;$
for $s \leftarrow 1$ **to** $\hat{N}$ **do**
    $j \sim \hat{\mathbf{f}};$
    while $|[\mathbf{p}]_j - \mu| < \sigma$ **do** $j \sim \hat{\mathbf{f}};$
    % Assign the b position and update the pmf;
    $b \leftarrow j; \quad [\mathbf{f}]_b \leftarrow [\mathbf{f}]_b - 1; \quad \alpha \leftarrow \alpha - 1;$
    $\hat{\mathbf{f}} \leftarrow \mathbf{f}/\alpha;$
    $d \leftarrow \mu - [\mathbf{p}]_b;$
    $I \leftarrow$sgn(d);
    if $I = 1$ **then** $\Pi \leftarrow \chi_b \setminus \hat{\chi}_b;$
    else $\Pi \leftarrow \hat{\chi}_b;$
    $\{\hat{\chi}, \hat{\Upsilon}, \mathbf{g}, \mathbf{p}, \mathbf{q}\} =$ Putd($d, \Pi, \hat{\chi}, \hat{\Upsilon}, r, \sigma, \mathbf{g}, \mathbf{p}, \mathbf{q}, I, b$);

in $\mathit{\Pi}$ whose elements are farther from the desired r (μ). Inside the while loop of line 6 there are two different cases. Lines 7-14 show the case where the difference d is positive, lines 8-11 assign the selected position (ℓ, b) to the respective sets $\hat{\chi}_b$ and $\hat{\Upsilon}_\ell$. If the value $[\mathbf{q}]_\ell$ is larger than the bound $r-\sigma$, lines 12-14 select another possible position ℓ and update $\mathbf{g}$. Lines 15-22 shows the case where d is negative, lines 16-19 eliminate the selected position (ℓ, b) to the respective sets $\hat{\chi}_b$ and $\hat{\Upsilon}_\ell$. If the value $[\mathbf{q}]_\ell$ is lower than $r+\sigma$, lines 20-22 select another possible position ℓ and update $\mathbf{g}$.
The complexity of Algorithm 1 is in the order of $(\hat{N}|d||\mathit{\Pi}| + \hat{M}|d||\mathit{\Pi}|)$ per iteration, in the worse case. However, given that the selections of positions j in the 6th line of Algorithm 2 and ℓ in the 4th line of Algorithm 3 are defined by the pmf $\hat{\mathbf{f}}$ and $\hat{\mathbf{g}}$, respectively, it is expected to reduce the number of times in accomplishing the while loop in line 7 of Algorithm 2 and the while loop in line 6 of Algorithm 3.

Algorithm 3: Put or eliminate $|d|$ ones according to the case, Putd

input : $d, \mathit{\Pi}, \hat{\chi}, \hat{\Upsilon}, r, \sigma, \mathbf{g}, \mathbf{p}, \mathbf{q}, I, b$
output: $\hat{\chi}, \hat{\Upsilon}, \mathbf{g}, \mathbf{p}, \mathbf{q}$

1 $\beta \leftarrow \sum_{j \in \mathit{\Pi}} |[\mathbf{g}]_j - r|$;
2 $\hat{\mathbf{g}} \leftarrow [\mathbf{g}]_{\mathit{\Pi}}/\beta$;
3 **for** $i \leftarrow 1$ **to** $|d|$ **do**
4 $\quad \ell \sim \hat{\mathbf{g}}; \quad [\mathbf{g}]_\ell \leftarrow [\mathbf{g}]_\ell - 1; \quad \beta \leftarrow \beta - 1;$
5 $\quad flag \leftarrow 0; \quad t \leftarrow 0;$
6 $\quad$ **while** $flag = 0$ *and* $t < |\mathit{\Pi}|$ **do**
7 $\quad\quad$ **if** $I = 1$ **then**
8 $\quad\quad\quad$ **if** $[\mathbf{q}]_\ell < (r - \sigma)$ **then**
9 $\quad\quad\quad\quad \hat{\chi}_b \leftarrow \hat{\chi}_b \cup \ell; \quad \hat{\Upsilon}_\ell \leftarrow \hat{\Upsilon}_\ell \cup b;$
10 $\quad\quad\quad\quad [\mathbf{p}]_b \leftarrow [\mathbf{p}]_b + 1; \quad [\mathbf{q}]_\ell \leftarrow [\mathbf{q}]_\ell + 1;$
11 $\quad\quad\quad\quad flag \leftarrow 1;$
12 $\quad\quad\quad$ **else**
13 $\quad\quad\quad\quad \hat{\mathbf{g}} \leftarrow [\mathbf{g}]_{\mathit{\Pi}}/\beta; \quad \ell \sim \hat{\mathbf{g}};$
14 $\quad\quad\quad\quad [\mathbf{g}]_\ell \leftarrow [\mathbf{g}]_\ell - 1; \quad \beta \leftarrow \beta - 1;$
15 $\quad\quad$ **else**
16 $\quad\quad\quad$ **if** $[\mathbf{q}]_\ell > (r + \sigma)$ **then**
17 $\quad\quad\quad\quad \hat{\chi}_b \leftarrow \hat{\chi}_b \setminus \ell; \quad \hat{\Upsilon}_\ell \leftarrow \hat{\Upsilon}_\ell \setminus b;$
18 $\quad\quad\quad\quad [\mathbf{p}]_b \leftarrow [\mathbf{p}]_b - 1; \quad [\mathbf{q}]_\ell \leftarrow [\mathbf{q}]_\ell - 1;$
19 $\quad\quad\quad\quad flag \leftarrow 1;$
20 $\quad\quad\quad$ **else**
21 $\quad\quad\quad\quad \hat{\mathbf{g}} \leftarrow [\mathbf{g}]_{\mathit{\Pi}}/\beta; \quad \ell \sim \hat{\mathbf{g}};$
22 $\quad\quad\quad\quad [\mathbf{g}]_\ell \leftarrow [\mathbf{g}]_\ell - 1; \quad \beta \leftarrow \beta - 1;$
23 $\quad\quad t \leftarrow t + 1;$

2.4. SIMULATION RESULTS

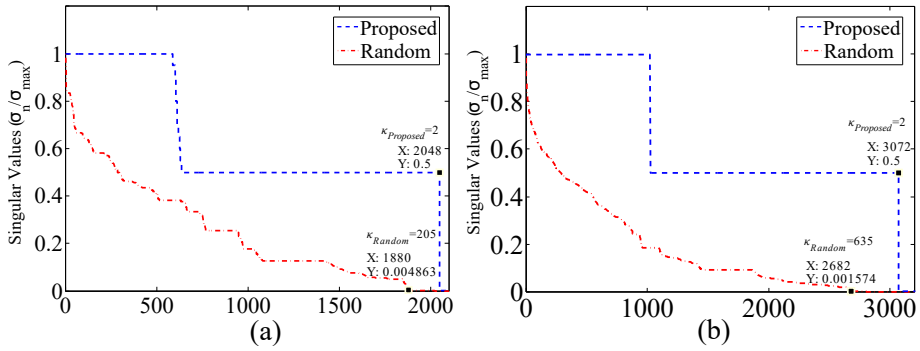
In this section, we present some simulations to illustrate the performance of Algorithm 1 for designing the SBiM, denoted as $Proposed$, and compare it with the CSI systems using a random coding element, denoted as $Random$.

2.4.1. MATRIX ANALYSIS

Singular value analysis

In this part, the eigenvalues and the condition number of the matrix $\mathbf{G}$ for the three SBiM are analyzed. The condition number (ratio between the greatest singular value and the least non-zero singular value) measures how ill-conditioned the measurement matrix is.
The eigenvalues for the PFA approach are showed in Fig. 4 for two scenarios (a) 50% and (b) 75% of compression ratio (percentage of the signal that is measured given by $(M/N)100\%$). The experiment computes the mean of 10 repetitions of singular values for PFA measurement matrices with $Z = 32$ and $L = 4$. Notice that for both compression ratios, the number of non-zero singular values in the *Proposed* case is greater than the *Random* one. Specifically, for a compression of 50% the *Proposed* has 168 non-zero eigenvalues more than the *Random*, and for a compression of 75% the *Proposed* has 390 non-zero eigenvalues more than the *Random*. The condition number for the above cases shows that *Proposed* measurement matrices are less ill-conditioned ($k_{Proposed} = 2$) compared with the *Random* ($k_{Random} = 205$ and 635).

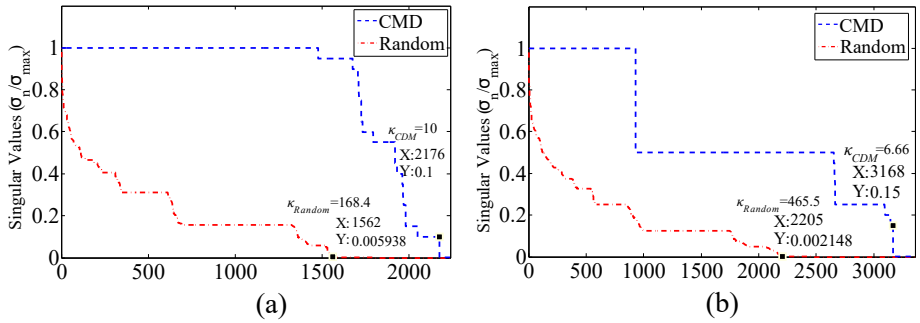
Figure 4: Singular values for the PFA system for (a) 50% and (b) 75% of compression ratio, with $Z = 32$ and $L = 4$.



The eigenvalues of the C-CASSI system are showed in Fig. 5 for (a) 55% and (b) 80% of compression ratio. The experiment computes the singular values for C-CASSI measurement matrices with $Z = 32$ and $L = 4$. Specifically, for a compression of 55% the *Proposed* has 614 non-zero eigenvalues more than the *Random*, and for a compression of 80% the *Proposed* has 390 non-zero eigenvalues more than the *Random*. The condition number of the *Proposed* measurement matrices shows that

they are less ill-conditioned ($k_{Proposed} = 10$ and 6.66) than the *Random* ($k_{Random} = 168.4$ and 465.5).

Figure 5: Singular values for the C-CASSI system for (a) 55% and (b) 88% of compression ratio, with $Z = 32$ and $L = 4$.



The singular values of the compressive X-ray tomography system with $M_1 = 552$ detector elements, $\hat{P} = 3$ fan-beam sources, and a compression ratio of 50% are shown in Fig. 6. The *Proposed* has 1357 non-zero singular values more than the *Random* case. The condition number of the *Proposed* measurement matrix shows that it is less ill-conditioned ($k_{Proposed} = 8.3 \times 10^4$) than the *Random* ($k_{Random} = 9.1 \times 10^4$).

Row and column sum histograms

In this part, the histograms corresponding to the vectors $\mathbf{q}$ and $\mathbf{p}$ are analyzed. Figure 7 shows the histograms for the PFA approach with $Z = 128$, $L = 16$, and a compression ratio of 18%. Figure 7 (a) and (b) show the histogram for the vector $\mathbf{p}$ before (*Random*) and after applying the CMD algorithm (*Proposed*), respectively. Figure 7 (b) shows that the distribution of the entries of $\mathbf{p}$ becomes concentrated around $\mu = 1$. Figure 7 (c) and (d) show the histograms for the vector $\mathbf{q}$. Figure 7 (d) shows that after the optimization, the distribution tends to $r = 5$.
Figure 8 shows the histograms for the C-CASSI approach with $Z = 128$, $L = 16$, and a compression ratio of 21%. Figure 8 (a) and (b) show the histograms for the vector $\mathbf{p}$. Figure 8 (b) shows that the entries of $\mathbf{p}$ are concentrated around $\mu = 1$ after the optimization. Figure 8 (c) and (d) show the histograms for the vector $\mathbf{q}$. Figure 8 (d) shows that the entries of $\mathbf{q}$ are concentrated around $r = 5$ after the optimization.
Fig. 9 shows the histograms for the compressive X-ray tomography measurement matrix in an angle $m = 0$ for $Z = 128$, $\hat{P} = 3$, $\mu = 200$, and $r = 5$. Fig. 9 (a) and (b) show the histograms for the vector $\mathbf{p}$. Fig. 9 (b) shows that the distribution is concentrated around $\mu = 200$. Fig. 9 (c) and (d) show the histogram for the vector $\mathbf{q}$. Fig. 9 (d) shows that the distribution is concentrated around $r = 5$.

2.4.2. **Recovery Results** In order to verify the proposed design algorithm, a set of compressive measurements is simulated using the forward models in (3.5) and

Figure 6: Singular values for the compressive X-ray tomography with 50% of compression ratio, $M_1 = 552$ and $\hat{P} = 3$.

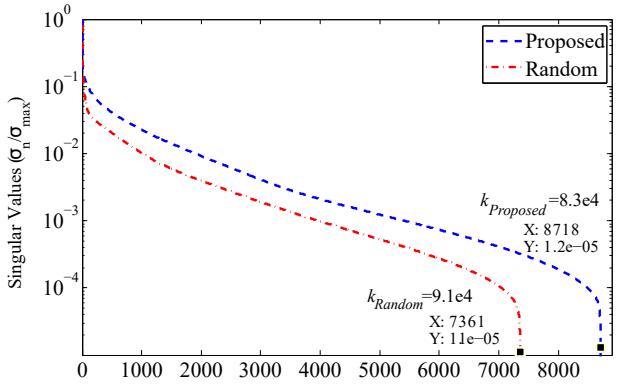


Table 2: Mean PSNR for different compression ratios [%] with SNR=20[dB], for PFA system

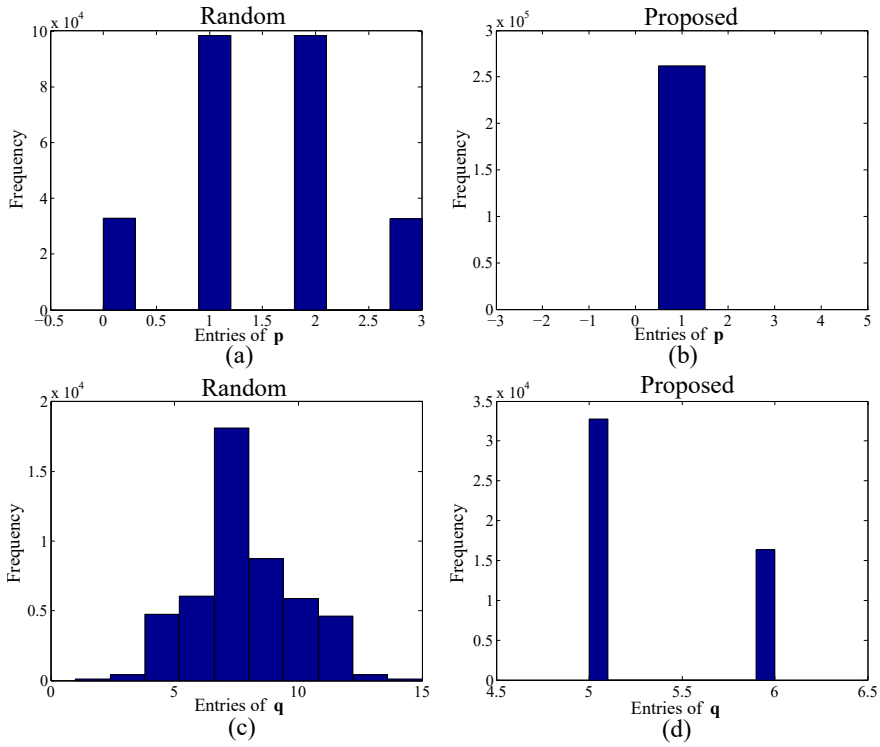
Compression ratio [%]	17	25	33	42	50	58	67
Random [dB]	32.2	33.3	33.8	34.0	34.1	34.2	41.7
Proposed [dB]	33.8	36.0	37.5	37.9	38.9	39.5	45.0

(2.11), for each compressive system, respectively. For the spectral data, the measurements are constructed by using the test spectral data cube C_1_5_1 taken from the database [58] which was acquired using Liquid Crystal Tunable Filters. The resulting test data cube has 256×256 pixels of spatial resolution and $L = 12$ spectral bands. For the compressive X-Ray tomography data, a cross-section of a human thorax of spatial resolution 128×128 was used.
For the reconstruction process, in the spectral case, the compressive sensing GPSR reconstruction algorithm [25] is used to recover the underlying data cube, with $\mathbf{\Psi} = \mathbf{W} \otimes \mathbf{\Psi}_{2D}$, where $\mathbf{\Psi}_{2D}$ is a 2D-Wavelet Symmlet and $\mathbf{W}$ is a 1D cosine transform. In the compressive X-Ray case, the recovery process was performed by the C-SALSA algorithm [48] with a Total Variation norm. The simulations are performed in a desktop architecture with an Intel Core i7 3.6GHz processor, 32GB RAM, and using Matlab R2014a.

PFA approach

The first experiment tests the design for different compression ratio levels, with a low noise addition of SNR=20[dB]. Table 2 shows the results where the PSNR values obtained with the Proposed method are higher than those obtained with Random.
Table 3 shows the results for the second experiment, where the compression ratios are fixed to 25% and 58% for different SNR levels. Despite the loss in reconstruction

Figure 7: Histogram of the vector $\mathbf{p}$ for the PFA system in subsection II, A, 1) with $Z = 128$, $L = 16$, and compression ratio of 18% for (a) random realizations (b) designed with the CMD algorithm. Histogram of the vector $\mathbf{q}$ (c) random realizations (d) designed with the CMD algorithm.



quality, the optimized measurement matrices conserve better performance than the Random.

Figure 10 shows the visual results of the differences between reconstructions and original data cube for the PFA approach with 25% of compression ratio and two SNR levels. There, it is possible to notice the better performance of the optimized PFA measurement matrix where a PSNR improvement of up to 5.2 [dB] is reached with SNR=20[dB] and to 2 [dB] with SNR=10.

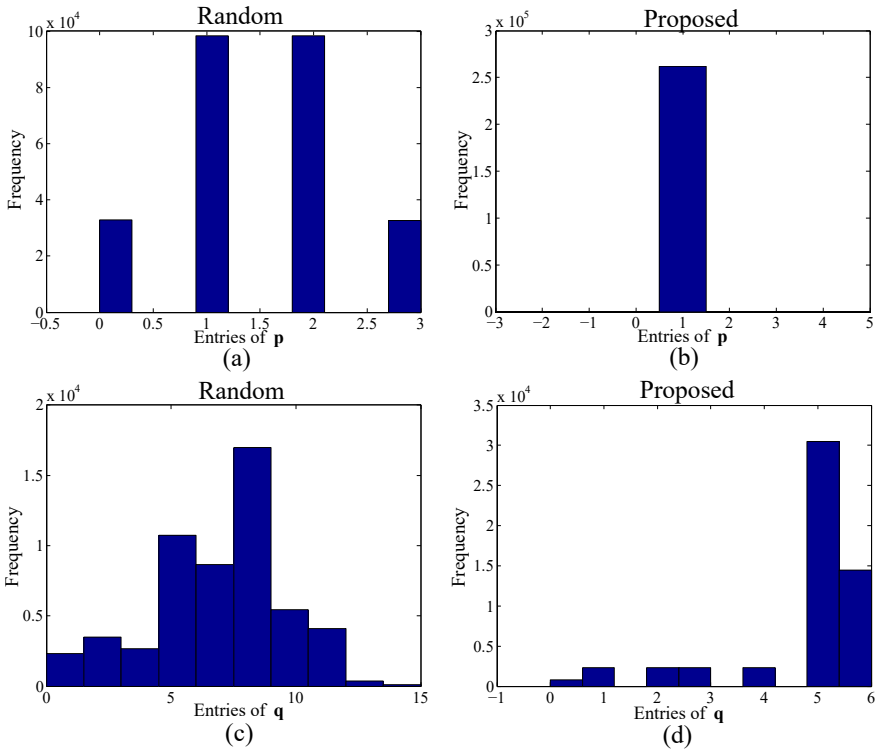
C-CASSI

For the C-CASSI system, Table 4 shows the results of mean PSNR as a function of the compression ratio for a SNR=20[dB].

The second experiment fixes the compression ratio to 26% and 61% for different SNR levels. Similarly to the PFA approach, despite the loss in PSNR, the designed matrices outperform the Random.

Figure 11 shows the visual results of the differences between reconstructions and original data cube for C-CASSI system with 26% of compression ratio. There, it is possible to notice the better performance of the optimized C-CASSI measurement

Figure 8: Histogram of the vector $\mathbf{p}$ for the C-CASSI system in subsection II, A, 2) with $Z = 128$, $L = 16$, and a compression ratio of 21% with (a) random realizations (b) designed with the CDM algorithm. Histogram of the vector $\mathbf{q}$ (c) random realizations (d) designed with the CDM algorithm.

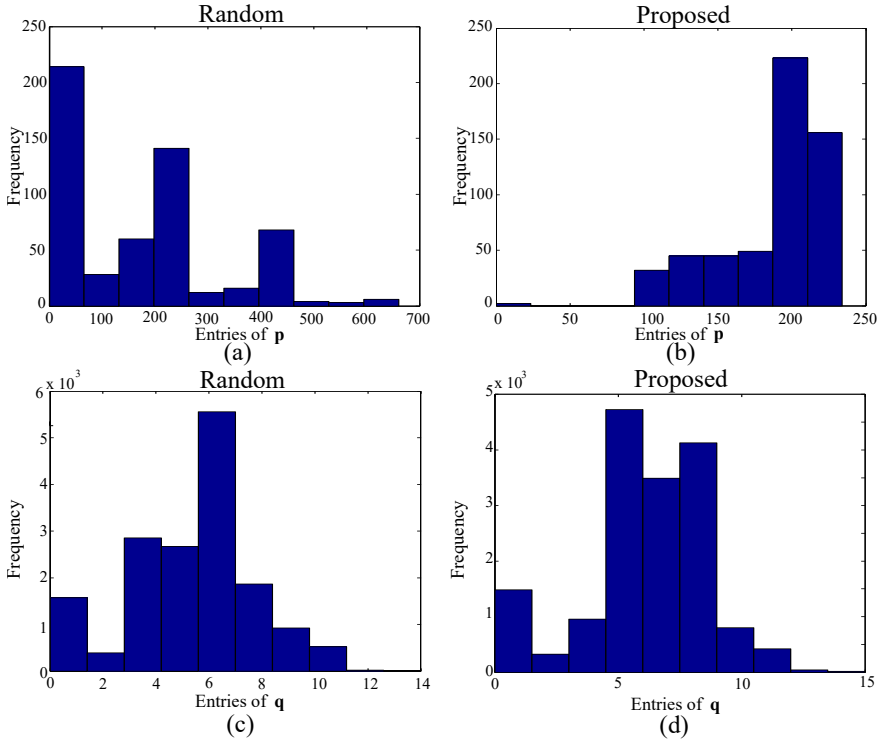


matrix where a PSNR improvement of up to 5.9 [dB] is attained with SNR=20[dB] and to 2.7 [dB] with SNR=10.

Compressive X-ray tomography

A scenario with a 1D line detector composed by $M_1 = 552$ elements and $\hat{P} = 3$ fan-beam X-ray sources placed uniformly in a 1×3 geometry, each coded aperture element corresponds to a particular detector element. Therefore, the coded apertures are placed in front of each of the sources composed by 552 elements. The ASTRA Tomography Toolbox [59] was used to obtain the system matrices $\mathbf{R}^m$.
The first experiment uses 3 configurations each one with different number of angles, 16 and 23 that correspond to 50% and 73% of compression ratios and without compression for 32 angles. Table 6 shows the results without noise and with a SNR level of 20 dB. In all the scenarios, the results of the designed matrices outperform the reconstruction quality of the Random ones.
Fig. 12 shows the visual results of the differences between reconstructions and original image with 50% of compression ratio without noise (Fig. 12 (b) and (c)) and with SRN=20 dB (Fig. 12 (d) and (e)). It is possible to notice the better performance

Figure 9: Histogram of the vector $\mathbf{p}$ for the compressive X-ray tomography system in subsection II, B, with $Z \times Z = 128 \times 128$, $\hat{P} = 3$, in a specific angle $m = 0$ for (a) random realizations (b) designed with the CMD algorithm. Histogram of the vector $\mathbf{q}$ (c) random realizations (d) designed with the CMD algorithm.



of the optimized compressive X-ray tomography measurement matrix where a PSNR improvement of up to 1.8 dB is attained without noise and to 0.5 dB with SNR=20 dB.

2.4.3. **Comparison with a coherence approach** In order to compare our results with a state-of-the-art solution we approximate the solution of the minimization problem $\mathbf{\Phi} = \arg\min_{\hat{\mathbf{\Phi}} \in \mathbb{R}^{M \times N}} \left\| \mathbf{I} - \mathbf{\Psi}^T \hat{\mathbf{\Phi}}^T \hat{\mathbf{\Phi}} \mathbf{\Psi} \right\|_F^2$, by a gradient descent approach followed by a thresholding and a structured operation, i.e. the iterative process consists of the next steps:

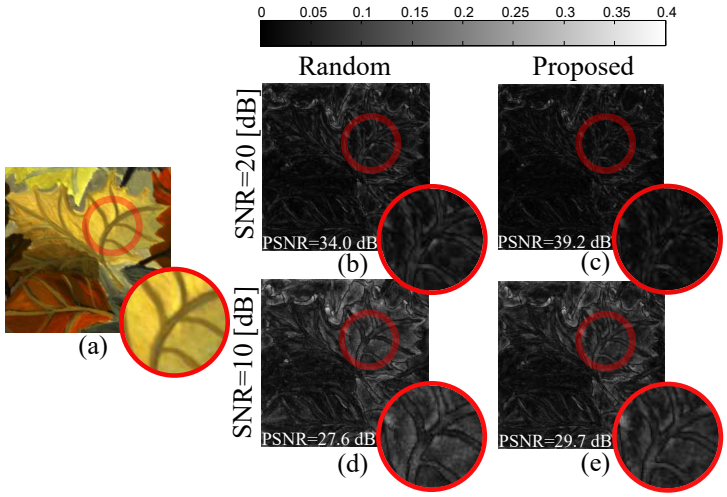
1. $\hat{\mathbf{\Phi}}^k = \hat{\mathbf{\Phi}}^{k-1} + \vartheta \hat{\mathbf{\Phi}}(\mathbf{I} - \hat{\mathbf{\Phi}}^T \hat{\mathbf{\Phi}})$, where the step $\vartheta = 0.125$ provides good results
2. *Thresholding:* $\bar{\mathbf{\Phi}}^k = \hat{\mathbf{\Phi}}^k > T$, with a threshold enables pass to 50% of the information
3. *Structure:* $\mathbf{\Phi}_{i,j} = \bar{\mathbf{\Phi}}^k_{i,j} \ \forall \ (i,j) \in \Omega$

This experiment was performed with the Ω for the PFA approach, with $Z = 32$, $L = 4$, and a compression ratio of 80%. Table 7 shows the PSNR results for different

Table 3: Mean PSNR as a function of noise (SNR) for the PFA system with 25% and 58% of compression ratio

		SNR [dB]			
Compression ratio [%]	Method	5	10	15	20
2*25	Random	22.6	27.6	32.8	34.0
	Proposed	25.5	29.7	33.4	39.2
2*58	Random	24.6	30.0	34.5	40.5
	Proposed	27.1	32.6	38.1	42.1

Figure 10: (a) Original. Differences between reconstructions on PFA system with 25% of compression ratio, (b) Original-Random and (c) Original-Proposed with SNR=20[dB], and (d) Original-Random and (e) Original-Proposed with SNR=10[dB].



SNR levels. Notice that the Proposed approach has a better performance than the Gradient and Random approaches.

2.5. CHAPTER CONCLUSION

This chapter proposed a new algorithm based on the optimal selection of non-zero entries positions for a structured, sparse, and binary measurement matrix (SBiM) design. Instead of solving the associated problem by traditional methods, an algorithm that minimizes the associated objective function by optimal selection of the non-zero elements per rows and columns has been designed. In this algorithm, the positions of the non-zero entries by rows and columns were updated alternatively, both using probability mass functions related to the sum of non-zero elements. This algorithm enables to solve an optimization problem without performing the product of large matrices, instead it takes advantage of the sparsity and structure of the SBiM providing a better performance than the traditional coherence problem. The proposed method

Table 4: Mean PSNR for different compression ratios [%], with SNR=20[dB] for C-CASSI system

Compression ratio [%]	17	26	35	43	52	61	70
Random [dB]	28.7	29.6	30.0	30.3	30.4	32.2	35.4
Proposed [dB]	32.1	35.8	38.1	39.4	40.2	40.6	47.8

Table 5: Mean PSNR as a function of noise (SNR) for C-CASSI system with 26% and 61% of compression ratio

		SNR [dB]			
Compression ratio [%]	Method	5	10	15	20
26	Random	21.7	25.2	28.8	29.6
	Proposed	23.6	27.9	32.1	35.5
61	Random	23.1	27.2	30.1	30.2
	Proposed	26.9	32.6	38.4	39.1

was tested on three different cases. Specifically, we developed the SBiM model for two compressive spectral sensing approaches and one compressive X-ray tomography. Numerical experiments showed that the proposed CMD algorithm compared competitively with Random and mutual coherence approaches, with the advantage of improving the reconstruction quality.

Figure 11: (a) Original. Differences between reconstructions on C-CASSI with 26% of compression ratio (b) Original-Random and (c) Original-Proposed with SNR=20[dB], and (d) Original-Random and (e) Original-Proposed with SNR=10[dB]

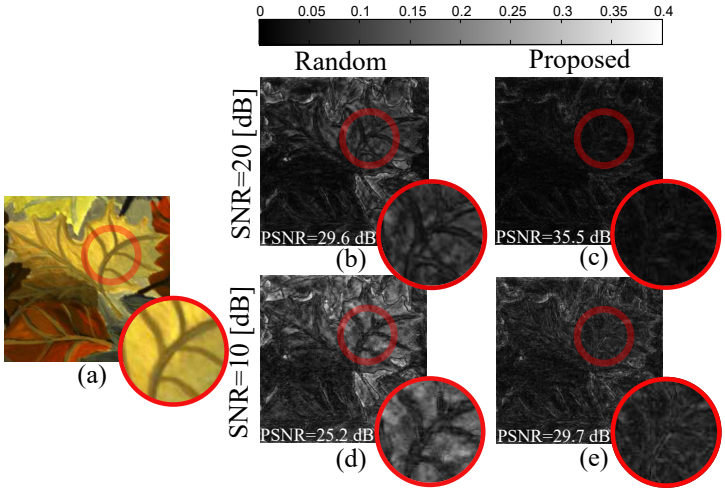


Table 6: Mean PSNR as a function of compression [%] without noise and with SNR=20 dB for Compressive X-ray tomography

		Compression ratio [%]		
SNR [dB]	Method	50	73	100
2*Noiseless	Random	27.6	33.4	42.4
	Proposed	29.1	39.0	49.0
2*20	Random	23.6	27.1	29.1
	Proposed	24.2	27.6	30.2

Figure 12: (a)Original. Differences between reconstructions on compressive X-ray tomography (a) Original-Random and (b) Original-Proposed noiseless and (c) Original-Random (d) Original-Proposed with SNR=20 dB

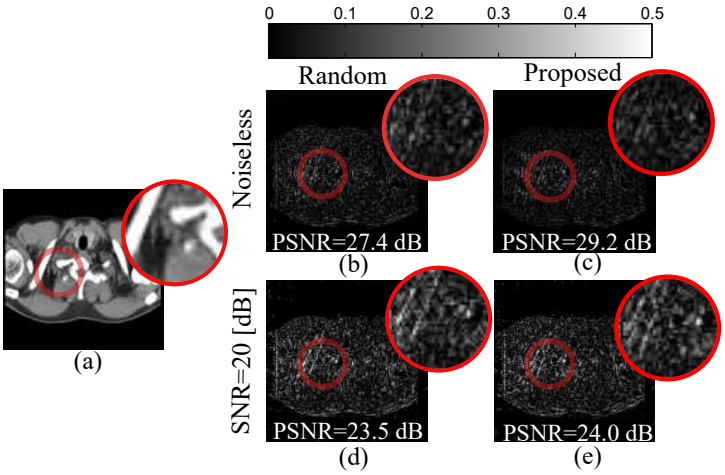


Table 7: Mean PSNR for different noise levels for the Random, the Gradient, and the Proposed approaches

SNR[dB]	5	10	15	20	40
Random	20.1	26.7	30.7	32.7	33.2
Gradient	20.2	26.8	31.4	34.3	35.2
Proposed	21.5	27.2	32.4	35.5	36.3

3. MATRIX DESIGN IN CSI WITH TWO SENSORS

Each image sensor is designed based on a particular signal-to-noise ratio (SNR). The reflected incoming light must be of sufficient energy to guarantee a sufficient SNR for a proper acquisition. To increase the energy level of the arriving signal, either the instantaneous field of view or the spectral window width must be increased. However, these solutions are mutually exclusive. That is, optical sensors suffer from an intrinsic energy tradeoff that limits the possibility of acquiring images of both high spatial and high spectral resolutions. This tradeoff prevents any simultaneous decrease in both the GSI and the spectral window width [60].

As a consequence, reconstructing a high-spatial and high-spectral multiband (HR) image from two degraded and complementary observed images is a relevant issue that has been addressed in different scenarios. In particular, fusing a high-spatial low-spectral resolution image and a low-spatial high-spectral image is a classic situation of multiband image reconstruction, such as pansharpening (MS+PAN) or hyperspectral pansharpening (HS+PAN). Generally, the linear degradations applied to the observed images with respect to the target high-spatial and high-spectral image reduce to spatial and spectral transformations. Thus, the multiband image fusion problem can be interpreted as restoring a 3D datacube from two degraded datacubes.

Due to the extremely large volumes of data collected by imaging spectrometers, hyperspectral data compression has received considerable interest in recent years. Compressive Spectral Imaging (CSI) is a technique to capture spectral images efficiently by encoding the spatio-spectral information. This is based on the fact that spectral images can be represented with a few coefficients in some basis. More formally, suppose a spectral image $\mathbf{X} \in \mathbb{R}^{Z \times Z \times L}$, or its vector representation $\mathbf{x} \in \mathbb{R}^N$ with $N = Z^2 L$, is S-sparse in a basis $\mathbf{\Psi}$, such that $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$ can be approximated by a linear combination of S vectors from $\mathbf{\Psi}$ with $S \ll N$. The number of multiplexed measurements in CSI is less than the required by the conventional spectral imaging devices.

To recover the HR image $\mathbf{X}$ from the measurements, the inverse sensing operator of the matrix that represents the CSI system must be found. However, this system is under-determined, therefore direct inversion is unfeasible. Instead, it is possible to force the sensing operator to exhibit a close to well-conditioning behavior in order to obtain a better inversion. Mainly, the performance of the compressive sensor is related to the coded apertures. A coded aperture can be modeled as a 2D matrix, where each pixel is an optical filter. The selection of the spectral characteristics of each pixel and their spatial distributions control the quality of the compressive measurements, which in turn determine the quality of the estimated spectral image. Therefore, there exist some research related to design the structure of the coded aperture in a CSI system.

Recently, some works have attended to recover an HR image from compressive

measurements of two CSI systems with different spatial and spectral resolution. There are a few works for designing coded apertures in a single CSI system, there are a handful of works that attempt to reconstruct an HR image from compressed measurements of two CSI systems.

In this work, we propose three coded aperture designs for the compressive spectral imaging fusion recovery problem. The first method attempts to perform uniform sampling on the spectral datacube along the snapshots taken by a CSI system. The second method includes another uniform constraint such as the expected value of the number of elements in a 2D spatial window be the same in the coded aperture for the multi-spectral system. The third method designs the coded aperture for spreading the pass elements in a 3D window that moves along the coded aperture.

There exist numerous approaches for fusing two low-resolution images of the same scene for recovering a better hyperspectral image. For instance, pan-sharpening [61] is an initial class of fusion methods that aims at fusing a panchromatic image (gray image) with a multispectral one to form an image with the high spatial resolution of the former and the high spectral resolution of the latter [62]–[64]. In spatial-spectral image fusion, a class of methods exploits the Bayesian framework to fuse an MSI with an HSI, that is, they determine a posterior distribution based on prior knowledge and on the observation model [65], [66]. Recently, the matrix factorization approach has been used by assuming that the HR image only contains a small number of significant elements. This approach first unfolds the HR image as a matrix and then represents it into a basis matrix and a coefficient matrix [61], [67].

All the above methods assume that the there are two complete datacubes of the same scene with different spatial and spectral resolutions. Instead, some works have attended to recover an HR image from compressive measurements of two CSI systems with different spatial and spectral resolutions. The work of [68] introduces a model allowing compressive data from high spatial/low spectral and low spatial/high spectral resolution sensors to be fused. The compressive fusion process is formulated as an inverse problem that minimizes an objective function defined as the sum of a quadratic data fidelity term and smoothness and sparsity regularization penalties. The work of [69] proposes an optimization algorithm based on the linear spectral unmixing model and using a block coordinate descent strategy. The main advantage of these approaches is using a significant reduced number of measurements.

The development of CSI systems is an active area of research, both in terms of hardware and the signal processing algorithms. Every CSI architecture can be modeled by a characteristic measurement matrix Φ, whose entries and structure depend on the different optical elements. For instance, the Patterned Filter Array (PFA) and the Colored Coded Aperture Snapshot Spectral Imaging (C-CASSI) approaches are CSI systems that have been widely used and have provided outstanding image reconstruction results [33], [54], [55].

Optimization of color coded apertures in the C-CASSI system has been demonstrated in the past. The work [33] proposes a row-per-row optimization approach based on concentration of measurement. Another approach exploits the spatial and

spectral correlation of the spectral images by modeled the CCA as a 3D black-and-white coded aperture, and thus seeks to spread the translucent elements (band-pass filters of the color coded aperture) along the 3D cube extent [70], [71].

3.1. CSI SYSTEMS MODELED FOR A FUSION FRAMEWORK

This section presents the sensing models of the CSI considered in this work: the Patterned Filter Array (PFA) and the Colored Coded Aperture Snapshot Spectral Imaging (C-CASSI).

Patterned Filter Array (PFA)

Let $\mathcal{X} \in \mathbb{R}^{Z\times Z\times L}$ be a spectral image with spatial resolution $Z \times Z$ and L spectral bands. For acquiring spectral images, the PFA approach uses optical elements that perform spectral filtering and collect multiple wavelength spectra on each pixel detector array in few snapshots. This colored coded aperture (CCA) can be modeled as a 4D encoder $\mathcal{D}$, where the element $\mathcal{D}_{s,t,u,k} \in \{0, 1\}$ represent the on-off wavelength codification of the light in the spatial position (s, t), the uth spectral band, and the kth snapshot. Let $\mathcal{X}_{s,t,u}$ be an element of the spectral image, where $s, t = 0, ..., Z - 1$ index the spatial coordinates and $u = 0, ..., L - 1$ determines the uth spectral band. The discretized measurements of the PFA approach from a single snapshot can be expressed as

$$\mathcal{Y}_{s,t,k} = \sum_{u=0}^{L-1} \mathcal{X}_{s,t,u}\mathcal{D}_{s,t,u,k} + \eta_{s,t,k}, \tag{3.1}$$

where $\mathcal{Y}_{s,t,k}$ is the intensity at the position (s, t) of the detector in the kth snapshot and $\eta_{s,t,k}$ is the white noise of the sensing system.
Let us define the set of measurements in (3.1) for the kth snapshot as $\mathbf{Y}_k \in \mathbb{R}^{Z\times Z}$. $\mathbf{Y}_k$ can be written in matrix notation as $\mathbf{y}_k = \mathbf{\Phi}_k\mathbf{x} + \boldsymbol{\eta}$, where $\mathbf{y}_k \in \mathbb{R}^{Z^2}$ is the vector representation of $\mathbf{Y}_k$, $\mathbf{x}$ is the vector representation of the spectral image $\mathcal{X}$, and $\mathbf{\Phi}_k$ represents the measurement matrix for the kth snapshot. More precisely, if we define the column vector $\mathbf{d}_{u,k} \in \{0, 1\}^{Z^2}$ as the vectorization of the CCA elements for the uth spectral band and the kth snapshot, such that $\mathbf{d}_{u,k} = \mathrm{vect}(\mathcal{D}_{:,:,u,k})$, the matrix $\mathbf{\Phi}_k$ is given by

$$\mathbf{\Phi}_k = \begin{bmatrix}\mathrm{diag}(\mathbf{d}_{0,k}) & \mathrm{diag}(\mathbf{d}_{1,k}) & \cdots & \mathrm{diag}(\mathbf{d}_{L-1,k})\end{bmatrix}. \tag{3.2}$$

The matrix $\mathbf{\Phi}_k$ in eq. (3.2) can be expressed as

$$\mathbf{\Phi}_k = \sum_{u=0}^{L-1} \mathrm{diag}(\mathbf{d}_{u,k})\mathbf{I}_C(\mathbf{P}_{Z^2L}^T)^{uN^2} \tag{3.3}$$

where $\mathbf{P}_n$ is the following $n \times n$ permutation matrix

$$\mathbf{P}_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \tag{3.4}$$

$\mathrm{diag}(\mathbf{v})$ is a diagonal matrix formed from the vector $\mathbf{v}$, $\mathbf{I}_C = [\mathbf{I}, \mathbf{0}_C]$, $\mathbf{I}$ is the $Z^2 \times Z^2$ identity matrix, and $\mathbf{0}_C$ is the $Z^2 \times Z^2(L-1)$ zero matrix.
The measurement process can be generalized to D snapshot projections where the set of measurements is defined as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \vdots \\ \mathbf{y}_{D-1} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_0 \\ \vdots \\ \mathbf{\Phi}_{D-1} \end{bmatrix} \mathbf{x} + \boldsymbol{\eta} = \mathbf{U}\mathbf{x} + \boldsymbol{\eta}. \tag{3.5}$$

where $\mathbf{U}$ is a $DZ^2 \times Z^2L$ matrix that accounts D snapshots each one with Z^2 measurements.

Colored Coded Aperture Snapshot Spectral Imaging (C-CASSI) System

In the C-CASSI system, the spectral image is coded, then spectrally shifted by a dispersive element, and a 2D Focal Plane Array detects the coded and shifted projections. Specifically, let $\mathcal{X}_{s,t,u}$ be an element of the spectral image and $\mathcal{D}_{s,t,u,k}$ an element of the encoder that represents the CCA, a discretized snapshot from C-CASSI can be expressed as

$$\mathcal{Y}_{s,t,k} = \sum_{u=0}^{L-1} \mathcal{X}_{s,(t-u),u} \mathcal{D}_{s,(t-u),u} + \eta_{s,t,k}, \tag{3.6}$$

where $\mathcal{Y}_{s,t,k}$ is the intensity at the position (s,t) of the detector that collects $Z \times (Z + L - 1)$ measurements in the kth snapshot. Let us define $\mathbf{Y}_k \in \mathbb{R}^{Z \times Z+L-1}$ as the set of measurements in the kth snapshot and $\mathbf{y}_k$ its vectorized version. Therefore, Eq. (3.6) can be written in matrix notation as $\mathbf{y}_k = \mathbf{\Phi}_k \mathbf{x} + \boldsymbol{\eta}$, where $\mathbf{y}_k \in \mathbb{R}^E$, with $E = Z(Z + L - 1)$, and $\mathbf{\Phi}_k$ is the measurement matrix of the kth snapshot.
The matrix $\mathbf{\Phi}_k$ can be related to the CCA, in this case, taking into account the dispersion effect as:

$$\mathbf{\Phi}_k = \begin{bmatrix} \mathrm{diag}(\mathbf{d}_{0,k}) & \mathbf{0}_{Z \times Z^2} & \cdots & \mathbf{0}_{Z(L-1) \times Z^2} \\ & \mathrm{diag}(\mathbf{d}_{1,k}) & \cdots & \\ & & \ddots & \\ \mathbf{0}_{Z(L-1) \times N^2} & \mathbf{0}_{Z(L-2) \times Z^2} & \cdots & \mathrm{diag}(\mathbf{d}_{L-1,k}) \end{bmatrix} \tag{3.7}$$

where $\mathbf{d}_{u,k} \in \mathbb{R}^{Z^2}$ is the vectorization of the CCA elements for the uth spectral band and the kth snapshot, and $\mathbf{0}_{Z \times Z^2}$ is the $Z \times Z^2$ zero matrix. As in the previous subsection, Eq. (3.7) can be constructed using permutation matrices as

$$\mathbf{\Phi}_k = \sum_{u=0}^{L-1} (\mathbf{P}_m)^{uZ} [\mathrm{diag}(\mathbf{d}_{u,k}), \mathbf{0}]^T \mathbf{I}_C (\mathbf{P}_{Z^2 L}^T)^{uZ^2} \tag{3.8}$$

where $m = Z^2 + Z(L-1)$, $\mathbf{0}$ is the $Z^2 \times Z(L-1)$ zero matrix, $\mathbf{I}_C = [\mathbf{I}, \mathbf{0}_C]$ with $\mathbf{I}$ is the $Z^2 \times Z^2$ identity matrix and $\mathbf{0}_C$ is the $Z^2 \times Z^2(L-1)$ zero matrix.
Here, also the set of measurements taken by D different snapshots can be defined as in (3.5), where $\mathbf{U}$ is a $DZ(Z+L-1) \times Z^2L$ matrix that takes D snapshot each one of $Z(Z+L-1)$ measurements.

3.2. **COMPRESSIVE SPECTRAL IMAGING FUSION**

The fusion of compressed measurements aims to recover a high spatial and high spectral resolution (HR) image from two images acquired with two CSI systems. The observed images are assumed to result from spectral or spatial degradations of an HR image acquired with two different CSI systems. Specifically, let $\mathbf{X}_M \in \mathbb{R}^{Z \times Z \times L_M}$ and $\mathbf{X}_H \in \mathbb{R}^{Z_H \times Z_H \times L}$, with $L_M \ll L$ and $Z_H \ll Z$, be the multispectral (MS) and the hyperspectral (HS) datacubes of a scene, respectively. We define the sensing matrix for the MS datacube as $\mathbf{U}_M$ and for the HS datacube as $\mathbf{U}_H$, therefore, the compressive measurements for both CSI systems are given by

$$\begin{aligned} \mathbf{y}_M &= \mathbf{U}_M \mathbf{x}_M + \boldsymbol{\eta}_M, \\ \mathbf{y}_H &= \mathbf{U}_H \mathbf{x}_H + \boldsymbol{\eta}_H, \end{aligned} \tag{3.9}$$

where $\mathbf{x}_M \in \mathbb{R}^{Z^2 L_M}$ and $\mathbf{x}_H \in \mathbb{R}^{Z_M^2 L}$ are the vectorization of $\mathbf{X}_M$ and $\mathbf{X}_H$, respectively, and $\boldsymbol{\eta}_M$ and $\boldsymbol{\eta}_H$ are additive noise terms. The vectors $\mathbf{x}_M$ and $\mathbf{x}_H$ can be related to the HR image by using an operator that represents the reduction of the number of bands and one that represents the spatial blurring and downsampling.
In summary, the measurements considered in this work are HS and MS compressed images given by

$$\begin{aligned} \mathbf{y}_M &= \mathbf{U}_M \mathbf{x}_M + \boldsymbol{\eta}_M = \mathbf{U}_M \mathbf{E}\mathbf{x} + \boldsymbol{\eta}_M = \boldsymbol{\Phi}_M \mathbf{x} + \boldsymbol{\eta}_M, \\ \mathbf{y}_H &= \mathbf{U}_H \mathbf{x}_H + \boldsymbol{\eta}_H = \mathbf{U}_H \mathbf{S}\mathbf{x} + \boldsymbol{\eta}_H = \boldsymbol{\Phi}_H \mathbf{x} + \boldsymbol{\eta}_H, \end{aligned} \tag{3.10}$$

where

- $\mathbf{E} \in \mathbb{R}^{Z^2 L_M \times Z^2 L}$ ($L_M \ll L$) represents the reduction of the number of spectral bands
- $\mathbf{S} \in \mathbb{R}^{Z_H^2 L \times Z^2 L}$ ($Z_H \ll Z$) represents the spatial blurring and downsampling
- $\mathbf{U}_M \in \mathbb{R}^{P \times Z^2 L_M}$ and $\mathbf{U}_H \in \mathbb{R}^{Q \times Z_H^2 L}$ are appropriate compressive sensing matrices, with $P, Q \ll Z^2 L$
- $\mathbf{y}_M \in \mathbb{R}^P$ and $\mathbf{y}_H \in \mathbb{R}^Q$ are the observed compressive measurements of the MS and HS images. P and Q are the number of measurements used to sense the HS and MS images, respectively.

Notice that a similar fusion model was posed in [72] except that the observed images were not compressed, i.e., without the matrices $\mathbf{U}_M$ and $\mathbf{U}_H$. The aim of this work is to design the entries of the matrices $\mathbf{U}_M$ and $\mathbf{U}_H$ in order to better estimate the HR image $\mathbf{x}$ from the observed measurements $\mathbf{y}_M$ and $\mathbf{y}_H$.

3.2.1. **Spectral Reduction Matrix $\mathbf{E}$** Let us define the matrix $\mathbf{E}$ in (3.10) for L spectral bands as

$$\mathbf{E} = \mathbf{I}_{L_M} \otimes \bar{\mathbf{E}} \tag{3.11}$$

where $\bar{\mathbf{E}} \in \{0,1\}^{Z^2 \times pZ^2}$ is a spectral reduction operator that performs a linear combination of p adjacent spectral bands given by $\bar{\mathbf{E}} = \mathbf{1}_p^T \otimes \mathbf{I}_{Z^2}$, with $\mathbf{1}_p$ is a ones column vector of length p and $\mathbf{I}_{Z^2}$ is the $Z^2 \times Z^2$ identity matrix, $\mathbf{I}_{L_M}$ is an $L_M \times L_M$ identity matrix with $L_M = L/p$, and $\otimes$ represents the Kronecker product.

3.2.2. **Measurement Matrices $\mathbf{\Phi}_M$ in a C-CASSI system** Let us define the matrix $\mathbf{U}_M$ in (3.10) as a C-CASSI measurement matrix that senses a $Z \times Z \times L_M$ spectral image. Given that the matrix $\mathbf{\Phi}_M$ in (3.10) is the product between the measurement matrix $\mathbf{U}_M$ and the spectral reduction matrix $\mathbf{E}$, $\mathbf{\Phi}_M$ can be decomposed into D_M submatrices, for D_M snapshots, each one related to a specific snapshot as follow

$$\begin{aligned} \mathbf{\Phi}_{M,k} &= \mathbf{\Phi}_k \mathbf{E} \\ &= \sum_{u=0}^{L_M-1} (\mathbf{P}_m)^{uZ} [(\mathrm{diag}(\mathbf{d}_{u,k}^M)\bar{\mathbf{E}})^T, \mathbf{0}]^T \mathbf{I}_C (\mathbf{P}_{Z^2L}^T)^{upZ^2} \end{aligned} \tag{3.12}$$

where $k = 0, 1, ..., D_M - 1$, $m = Z^2 + Z(L_M - 1)$, $\mathbf{d}_{u,k}^M \in \{0,1\}^{Z^2}$ is the vectorization of the CCA elements in the uth spectral band and the kth snapshot for the MS C-CASSI system like the one defined in Eq. (3.7), $\mathbf{0}$ is the $pZ^2 \times Z(L_M - 1)$ zero matrix, $\mathbf{I}_C = [\mathbf{I}, \mathbf{0}_C]$, $\mathbf{I}$ is the $pZ^2 \times pZ^2$ identity matrix, and $\mathbf{0}_C$ is the $pZ^2 \times pZ^2(L_M - 1)$ zero matrix. The matrix $\mathbf{\Phi}_M$ that accounts the D_M snapshots is given by

$$\mathbf{\Phi_M} = \begin{bmatrix} \mathbf{\Phi}_{M,0} \\ \vdots \\ \mathbf{\Phi}_{M,D_M-1} \end{bmatrix}. \tag{3.13}$$

3.2.3. **Spatial Blurring and Downsampling Matrix $\mathbf{S}$** Let us define the matrix $\mathbf{S}$ in (3.10) for L spectral bands as

$$\mathbf{S} = \mathbf{I}_L \otimes \bar{\mathbf{S}}\bar{\mathbf{B}} \tag{3.14}$$

where $\mathbf{I}_L$ is an $L \times L$ identity matrix, $\bar{\mathbf{B}} \in \mathbb{R}^{Z^2 \times Z^2}$ is a cyclic convolution operator operator acting on each band, and $\bar{\mathbf{S}}$ is a spatial downsampling operator that performs a linear combination of q^2 adjacent pixels in a single image, that is $\bar{\mathbf{S}}$ will degrade spatially each spectral band by a factor of q. Mathematically, $\bar{\mathbf{S}} \in \{0,1\}^{Z_H \times Z}$ can be expressed as

$$\bar{\mathbf{S}} = \mathbf{I}_{Z_H} \otimes [\mathbf{1}_q^T \otimes (\mathbf{I}_{N_H} \otimes \mathbf{1}_q^T)], \tag{3.15}$$

where q is the downsampling spatial factor, $Z_H = Z/q$ with N the size of the high spatial resolution spectral band, $\mathbf{I}_{Z_H}$ is an $Z_H \times Z_H$ identity matrix, and $\mathbf{1}_q$ is a ones column vector of length q.

3.2.4. **Measurement Matrices $\mathbf{\Phi}_H$ in a C-CASSI system** Let us define the matrix $\mathbf{U}_H$ in (3.10) as a C-CASSI measurement matrix that senses an $Z_H \times Z_H \times L$ spectral image. Given that the matrix $\mathbf{\Phi}_H$ in (3.10) is the product between the measurement matrix $\mathbf{U}_H$ and the spatial blurring and downsampling matrix $\mathbf{S}$, $\mathbf{\Phi}_H$ can be decomposed into D_H submatrices, for D_H snapshots, each one related to a specific snapshot as follow

$$\begin{aligned} \mathbf{\Phi}_{H,k} &= \mathbf{\Phi}_k \mathbf{S} \\ &= \sum_{u=0}^{L-1} (\mathbf{P}_m)^{uZ_H} [(\mathrm{diag}(\mathbf{d}_{u,k}^H)\bar{\mathbf{S}}\bar{\mathbf{B}})^T, \mathbf{0}]^T \mathbf{I}_C (\mathbf{P}_{Z^2L}^T)^{uZ^2} \end{aligned} \tag{3.16}$$

where $k = 0, 1, ..., D_H - 1$, $m = N_H^2 + N_H(L-1)$, $\mathbf{d}_{u,k}^H \in \{0,1\}^{N_H^2}$ is the vectorization of the CCA elements in the uth spectral band and the kth snapshot for the HS C-CASSI system like the one defined in Eq. (3.7), $\mathbf{0}$ is the $Z^2 \times Z_H(L-1)$ zero matrix, $\mathbf{I}_C = [\mathbf{I}, \mathbf{0}_C]$, $\mathbf{I}$ is the $Z^2 \times Z^2$ identity matrix, and $\mathbf{0}_C$ is the $Z^2 \times Z^2(L-1)$ zero matrix. The matrix $\mathbf{\Phi}_H$ that accounts the D_H snapshots is given by

$$\mathbf{\Phi_H} = \begin{bmatrix} \mathbf{\Phi}_{H,0} \\ \vdots \\ \mathbf{\Phi}_{H,D_H-1} \end{bmatrix}. \tag{3.17}$$

3.2.5. **Observations related to the CSI systems** The definitions of matrices $\mathbf{\Phi}_{M,k}$ in (3.12) and $\mathbf{\Phi}_{H,k}$ in (3.16) were related to the C-CASSI system described by (3.8). However, these definitions can be related to the patterned system (3.3) following the same procedure that in the last subsections.

3.3. CODED APERTURE DESIGN METHODS

The observation model in (3.10) has the same input $\mathbf{x}$ and two sets of measurements $\mathbf{y}_M$, $\mathbf{y}_H$ as outputs. If we stacked the measurements in a single vector, the sensing process can be represented as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_M \\ \mathbf{y}_H \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_M \\ \mathbf{\Phi}_H \end{bmatrix} \mathbf{x} = \mathbf{\Phi x}. \tag{3.18}$$

If we take into account that the spectral image $\mathbf{x}$ is sparse in a basis $\mathbf{\Psi}$, such that $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$, Eq. (3.18) can be seen as the CS problem $\mathbf{y} = \mathbf{\Phi\Psi}\boldsymbol{\theta} = \mathbf{A}\boldsymbol{\theta}$.

CS requires that the sensing matrix $\mathbf{A}$ accomplish certain orthonormality properties, for instance, the restricted isometry property (RIP) [73]. The matrix $\mathbf{A}$ accomplishes the RIP if any set of S of its columns is orthonormal. In general, if the sparsity of the signal is neglected, the objective is that the matrix $\mathbf{A}$ be as orthonormal as possible, that is $\mathbf{A}^T\mathbf{A} = \mathbf{I}$. In our case $\mathbf{A}^T\mathbf{A} = \mathbf{\Psi}^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{\Psi} = \mathbf{\Psi}^T\mathbf{G}\mathbf{\Psi}$, where we define $\mathbf{G} = \mathbf{\Phi}^T\mathbf{\Phi}$. If the sparsifying basis $\mathbf{\Psi}$ is orthonormal, selected such that it induces high sparsity, and fixed, the design of the matrix will focus on that $\mathbf{\Phi}$ be as orthonormal as possible, that is $\mathbf{G} = \mathbf{\Phi}^T\mathbf{\Phi} = \mathbf{I}$.
Given that the matrix $\mathbf{G}$ is related to the entries of the CCAs $\mathcal{D}^M$ and $\mathcal{D}^H$, our approach attempts to impose restrictions to the CCAs such that $\mathbf{G}$ tends to the identity. Specifically, we express the entries of $\mathbf{G}$ as a function of the CCA elements for the MS system $\mathcal{D}^M_{s,t,u,k}$ and the CCA elements of the HS system $\mathcal{D}^H_{s,t,u,k}$, then we constraint the $\{0,1\}$ values such that the diagonal elements of $\mathbf{G}$ tend to 1 and the off-diagonal elements be close to zero.
We will determine the entries of the matrix $\mathbf{G}$ in the fusion framework, as follows

$$\mathbf{G} = \begin{bmatrix}\mathbf{\Phi}_M \\ \mathbf{\Phi}_H\end{bmatrix}^T \begin{bmatrix}\mathbf{\Phi}_M \\ \mathbf{\Phi}_H\end{bmatrix} = \mathbf{\Phi}_M^T\mathbf{\Phi}_M + \mathbf{\Phi}_H^T\mathbf{\Phi}_H = \mathbf{G}_M + \mathbf{G}_H, \tag{3.19}$$

where the matrix related to the spectral reduction and compressive sensing $\mathbf{G_M}$ is given by

$$\begin{bmatrix}\mathbf{\Phi}_{M,0} \\ \vdots \\ \mathbf{\Phi}_{M,D_M-1}\end{bmatrix}^T \begin{bmatrix}\mathbf{\Phi}_{M,0} \\ \vdots \\ \mathbf{\Phi}_{M,D_M-1}\end{bmatrix} = \sum_{k=0}^{D_M-1} \mathbf{\Phi}_{M,k}^T\mathbf{\Phi}_{M,k} = \sum_{k=0}^{D_M-1} \mathbf{G}_{M,k}, \tag{3.20}$$

where each product $\mathbf{G}_{M,k} = \mathbf{\Phi}_{M,k}^T\mathbf{\Phi}_{M,k}$ can be modeled as

$$\begin{bmatrix} \mathbf{D}_0^M & \mathbf{F}_{0,1}^M & \mathbf{F}_{0,2}^M & \cdots & \mathbf{F}_{0,L_M-1}^M \\ (\mathbf{F}_{0,1}^M)^T & \mathbf{D}_1^M & \mathbf{F}_{1,2}^M & \cdots & \mathbf{F}_{1,L_M-1}^M \\ (\mathbf{F}_{0,2}^M)^T & (\mathbf{F}_{1,2}^M)^T & \mathbf{D}_2^M & \cdots & \mathbf{F}_{2,L_M-1}^M \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (\mathbf{F}_{0,L_M-1}^M)^T & (\mathbf{F}_{1,L_M-1}^M)^T & (\mathbf{F}_{2,L_M-1}^M)^T & \cdots & \mathbf{D}_{L_M-1}^M \end{bmatrix} \tag{3.21}$$

which is a symmetric matrix whose diagonal submatrices $\mathbf{D}_u^M \in \{0,1\}^{pZ^2\times pZ^2}$ are related to the coded aperture elements of the uth spectral band. Specifically, each diagonal matrix is given by

$$\mathbf{D}_u^M = \bar{\mathbf{E}}^T\mathbf{T}_u\mathbf{T}_u\bar{\mathbf{E}} = \bar{\mathbf{E}}^T\mathbf{T}_u^2\bar{\mathbf{E}} \tag{3.22}$$

where $\mathbf{T}_u = \mathrm{diag}(\mathbf{d}_u^M) \in \{0,1\}^{Z^2\times Z^2}$ and $\bar{\mathbf{E}} \in \{0,1\}^{Z^2\times pZ^2}$ has been defined in the subsection 4.2.1.
The off-diagonal submatrices in Eq. (3.21), $\mathbf{F}_{i,j}^M \in \{0,1\}^{pZ^2\times pZ^2}$, are related to the interactions between the elements of the coded aperture that affects the ith spectral band and the one that affects the jth spectral band. Specifically, in the C-CASSI system, these matrices are given by

$$\mathbf{F}_{i,j}^M = \bar{\mathbf{E}}^T\mathbf{T}_i(\mathbf{P}_{Z^2})^{(i-j)N}\mathbf{T}_j\bar{\mathbf{E}} \tag{3.23}$$

where $\mathbf{P}_{Z^2} \in \{0,1\}^{Z^2 \times Z^2}$ is a permutation matrix as the one defined in (3.4).
If we take into account the interactions of the L_M spectral bands, it is possible to express the entries of the matrix $\mathbf{G}_{M,k}$ as a function of the CCA elements as follows

$$[\mathbf{G}_{M,k}]_{i,j} = \begin{cases} \sum_{\ell_2=0}^{L_M-1} \sum_{\ell_1=0}^{L_M-1} d^M_{\hat{i},\hat{j}+\ell_2,\ell_1} d^M_{\hat{i},\hat{j}+\ell_1,\ell_2} & \text{if} \quad \hat{i}'-\ell_2 Z = \hat{j}'-\ell_2 Z \\ 0, & \text{otherwise}, \end{cases} \tag{3.24}$$

where $\hat{\imath} = \mathrm{mod}(i, Z)$, $\hat{\jmath} = \left\lfloor \frac{j}{Z} \right\rfloor - \left\lfloor \frac{j}{Z^2} \right\rfloor Z$, $\hat{i}' = \mathrm{mod}(i, Z^2)$, $\hat{j}' = \mathrm{mod}(j, Z^2)$.
Similarly, the matrix related to the spatial blurring, downsampling, and sensing process $\mathbf{G}_H$ has the same structure of (3.20) but with the D_H matrices $\mathbf{\Phi}_{H,k}$. The product $\mathbf{G}_{H,k} = \mathbf{\Phi}^T_{H,k}\mathbf{\Phi}_{H,k}$ is also a block symmetric matrix as (3.21) whose diagonal submatrices $\mathbf{D}^H_u \in \mathbb{R}^{Z^2 \times Z^2}$ are related to the coded apertures of the uth spectral band as

$$\mathbf{D}^H_u = \bar{\mathbf{B}}^T\bar{\mathbf{S}}^T\mathbf{T}_u\mathbf{T}_u\bar{\mathbf{S}}\bar{\mathbf{B}} = \bar{\mathbf{B}}^T\bar{\mathbf{S}}^T\mathbf{T}^2_u\bar{\mathbf{S}}\bar{\mathbf{B}} \tag{3.25}$$

where $\mathbf{T}_u = \mathrm{diag}(\mathbf{d}^H_u)$ and $\bar{\mathbf{S}}$ and $\bar{\mathbf{B}}$ has been defined in subsection 4.2.3.
The off-diagonal submatrices $\mathbf{F}^H_{i,j} \in \mathbb{R}^{Z \times Z}$ are related to the interactions between the elements of the coded aperture that affects the ith band with the one that affects the jth band. Specifically, these matrices are given by

$$\mathbf{F}^H_{i,j} = \bar{\mathbf{B}}^T\bar{\mathbf{S}}^T\mathbf{T}_i(\mathbf{P}_{Z^2_H})^{(i-j)Z_H}\mathbf{T}_j\bar{\mathbf{S}}\bar{\mathbf{B}} \tag{3.26}$$

where $\mathbf{P}_{Z^2_H} \in \{0,1\}^{Z^2_H \times Z^2_H}$ is a permutation matrix as the one defined in (3.4).
If we take into account the interactions of the L spectral bands without including the blurring, the entries of the matrix $\mathbf{G}_{H,k}$ can be expressed as

$$[\mathbf{G}_{H,k}]_{i,j} = \begin{cases} \sum_{k_2=0}^{L-1} \sum_{k_1=0}^{L-1} d^H_{i',j'+k_2,k_1} d^H_{i',j'+k_1,k_2} & \text{if} \quad i' = j''\delta\left(\hat{\jmath}''-\hat{\imath}''\right) \\ 0, & \text{otherwise}, \end{cases} \tag{3.27}$$

where $i' = \lfloor i/q \rfloor - \lfloor i/Z \rfloor Z_H$, $j' = \lfloor j/(qZ) \rfloor$, $j'' = \lfloor j/q \rfloor - \lfloor j/Z \rfloor Z_H$, $\hat{\jmath}'' = \lfloor (j + k_2 Z q)/(qZ) \rfloor$, $\hat{\imath}'' = \lfloor (i + k_1 Z q)/(qZ) \rfloor$, δ is the Dirac delta function.
The following observations can be used for the optimization of the CCAs:

1. The sparsifying base $\mathbf{\Psi}$ is considered properly selected such that it induces high sparsity. Thus, it is considered fixed.

2. The sum of the correlated products of the CCA in a single shot along all the spectral bands, as the equations (3.24) and (3.27) show, tends to increase the off-diagonal elements of $\mathbf{G}_M$ and $\mathbf{G}_H$. Therefore, impose a constraint that decreases the number of non-zero elements along the spectral bands is required.

3. The sum of the matrices $\mathbf{G}_{M,k}$ and $\mathbf{G}_{H,k}$ along the snapshots to form the matrices $\mathbf{G}_M$ and $\mathbf{G}_H$, respectively, increases the value of the elements, therefore, minimizing the quantity of light that passes through the CCA in each snapshot is desired in a complementary way. With complementary we mean that in each position s, t, u of the CCA could be a non-zero element only in one of the snapshots.

4. The possible intersections between the matrix $\mathbf{G}_M$ and $\mathbf{G}_H$ could occurs in neighborhoods related to the spatial downsampling, that is, each set of $q \times q$ pixels of the multispectral CCA is related to a single pixel of the hiperspectral CCA, therefore, a restriction that minimizing the quantity of light that passes through the CCA for the multispectral system in each set of $q \times q$ pixels is desired.

5. Given that the above observations attempt to decrease the quantity of light that passes through the systems and that the number of elements that we require to recover is Z^2L, we restrict the systems transmittance to collect this minimum amount of elements.

In order to interpret these observations in a physical scenario, the following can be analyzed:
The most faithful sampling process results when the system is modeled as the combination of the image with an identity matrix such that each voxel in a discrete image is measured at least once. However, the compressive system having fewer rows than columns and requiring randomness restricts the choice of an identity.
Therefore, the idea of measuring each voxel at least once must be held because it is analogous to the identity but fulfilling the restrictions of the compressive system, then the motivation for the CCA designs is to measure all the image voxels in random acquisitions.

3.3.1. **Uniform Sensing** Let us define $\mathbf{d}_{u,k}$ as the vectorization of the CCA elements for the spectral band u and the kth snapshot. In this case, we constraint each CSI system to measure a voxel at least once, therefore we want to collect $U = Z^2L$ elements. The uniform sensing approach [74] attempts to locate these elements in the CCA uniformly random constraining the sum of passing elements across all snapshots to be one. Specifically, this can be mathematically expressed as:

$$\sum_{k=0}^{D-1} \mathbf{d}_{u,k} = \mathbf{1}_{Z^2} \quad \forall \ u \in \{0, 1, ..., L-1\} \tag{3.28}$$

where $\mathbf{1}_{Z^2}$ is a Z^2-long one-valued vector. This constraint agrees with the third observation of the last part allowing only one passing element along all the snapshots. Additionally, in order to uniformly distribute the measurements along the spectral bands and following the second observation, we want to accomplish that the expected value of the sum of elements along the spectral bands in each snapshot be

L/K, that is:

$$E\left\{\sum_{u=0}^{L-1}\mathbf{d}_{u,k}\right\}=(L/K)\mathbf{1}_{Z^2}\quad\forall\ k\in\{0,1,...,D-1\}. \tag{3.29}$$

The constraints of equations (3.28) and (3.29) are equivalent to the uniform restrictions proposed in Chapter 3, therefore, it is possible to use the proposed algorithm from chapter 3 to design the matrix and therefore the distributions of the CCA of each CSI system independently. The benefits of this design is that each CSI system will have suitable properties of linearly independent measurements. The drawback of this design is that it does not have into account the spatial restrictions related to the fusion of the two CSI systems measurements.

3.3.2. **Uniform with local spatial constraints** In order to follow the fourth observation, this approach uses the uniform constraints of the subsection 3.3.1 and restricts the high spatial resolution CCA, $\mathbf{d}_u^M$, to have approximately the same number of passing elements in spatial regions of neighborhoods $p\times p$, such that the high spatial resolution CCA has the following additional restriction:

$$\begin{aligned}E\left\{\textstyle\sum_{i\in\Omega_j}[\mathbf{d}_{u,k}^M]_i\right\}=\tfrac{Z^2}{Kp^2}\mathbf{1}_{p^2}\quad&\forall\ k\in\{0,1,...,D_M-1\}\\ &\text{and}\ \ u\in\{0,1,...,L_M-1\},\end{aligned} \tag{3.30}$$

where Ω_j is the set of positions in the jth $p\times p$ neighborhood along the spatial dimension of the CCA $\mathbf{d}_{u,k}^M$. In order to accomplish the constraint (3.30), after applying the uniform algorithm to both CSI systems, the spatial distribution of the elements in the CCA for the MS system is rearranged. This can be done by an algorithm that computes the sum of passing elements in each neighborhood of $p\times p$ pixels and moves randomly through the shots and bands the ones that exceed the desired average ($Z^2/(Kp^2)$), keeping the uniform constraints. The benefit of this approach is that spreading the passing elements in the spatial neighborhoods of the CCA for the MS system decreases the interactions with the elements iof the CCA for the HS system. The drawback of this design is that it requires an additional step that moves one step per neighborhood through $(Z/q)^2L$ spatial neighborhoods.

3.3.3. **3D Spread** In order to minimizing the possible interactions, one general approach requires separating the non-zero elements of the CCA in a 3D sub-window. The aim of this approach is to spread the passing elements of each CCA in every sub-window of size $2L\times 2L\times L$ along all its rows, columns, and slices [70], [71]. Here, the CCA entries have to satisfy the minimum light throughput constraint of Eq. (3.28). Additionally, in order to spread the elements of the CCA, the design attempts to better accomplish the following optimization problem

$$\begin{aligned}&\arg\min_{\{\mathbf{d}_{0,k},...,\mathbf{d}_{L-1,k}\}_{k=0}^{D-1}}\ \beta\\ &\text{s.t.}\ \textstyle\sum_{k=0}^{D-1}\mathbf{d}_{u,k}=\mathbf{1}_{Z^2}\quad\forall\ u\in\{0,1,...,L-1\}\end{aligned} \tag{3.31}$$

where β is defined as

$$\beta = \sum_{k=0}^{D-1} \sum_{s=0}^{Z-1} \sum_{t=0}^{Z-1} \sum_{u=0}^{L-1} \sum_{\hat{u}=0}^{L-1} d_{s,t,u}^{k} d_{s,t,\hat{u}}^{k}. \tag{3.32}$$

where $d_{s,t,u}^{k}$ is the 3D representation of the CCA with k denoting the snapshot. An iterative algorithm for accomplishing this design is presented in [70], [71]. This randomly walks along all the pixels of a of randomly generated CCA satisfying the complementary (sum-to-one) constraint. In every step, the algorithm filters the sub-window of size $2L \times 2L \times L$ centered at the current pixel with a 3D-Euclidean filter and evaluates the objective function of the problem for every snapshot. The snapshot pixel that gives the lowest objective value is set to 1 (passing element) and the rest of the snapshots at the same positions are set to 0 (blocking element). The algorithm iterates on the updated CCA until the objective function stops decreasing, or until the objective function lies within a certain threshold. The 3D-Euclidean filter is designed such that the weights are calculated inversely proportional to the Euclidean distance between the points in a 3D mesh, with respect to the central pixel. The advantage of this design is that spreading the passing elements along 3D sub windows of each CSI system, the correlations between both CSI systems will be reduced. The drawback of this design is its high complexity.

3.4. **RECOVERY ALGORITHM**

To recover the high spatial and spectral resolution spectral image, the following optimization problem is established

$$\arg\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{U}_M \mathbf{E}\mathbf{x} - \mathbf{y}_M\|_2^2 + \frac{1}{2} \|\mathbf{U}_H \mathbf{S}\mathbf{x} - \mathbf{y}_H\|_2^2 + \phi(\mathbf{x}), \tag{3.33}$$

where the two first terms are the data fidelity terms associated with the HS and MS observations, $\|\cdot\|_2$ is the ℓ_2 norm and the last term corresponds to regularization. In this work, we consider that the vectorized image $\mathbf{x}$ has a sparse representation on a basis $\mathbf{\Psi} \in \mathbb{R}^{Z^2L \times Z^2L}$. Also, we consider a TV regularization to preserve sharp edges or object boundaries in the spatial domain.
The minimization problem in (3.33) can be solved by using the ADMM algorithm, where the objective function and the constraints are split by using auxiliary variables as follows

$$\begin{aligned} \underset{\mathbf{x},\mathbf{v}_i}{\text{argmin}} \quad & \tfrac{1}{2}\|\mathbf{U}_M \mathbf{E}\mathbf{v}_1 - \mathbf{y}_M\|_2^2 + \tfrac{1}{2}\|\mathbf{U}_H \mathbf{S}_d \mathbf{v}_2 - \mathbf{y}_H\|_2^2 \\ & + \lambda \|\mathbf{v}_3\|_1 + \lambda_{TV} \|\mathbf{v}_4\|_1 \\ \text{subject to} \quad & \mathbf{v}_1 = \mathbf{x} \\ & \mathbf{v}_2 = \mathbf{B}_s \mathbf{x} \\ & \mathbf{v}_3 = \mathbf{\Psi}\mathbf{x} \\ & \mathbf{v}_4 = \mathbf{D}\mathbf{x} \end{aligned} \tag{3.34}$$

where $i = 1, ..., 4$, $\mathbf{S}_d = \mathbf{I}_L \otimes \bar{\mathbf{S}}$, $\mathbf{B}_s = \mathbf{I}_L \otimes \bar{\mathbf{B}}$, $\mathbf{\Psi}$ is a basis of the wavelet domain, $\mathbf{D} = \mathbf{I}_L \otimes \mathbf{D}_{im}$, with $\mathbf{D}_{im}$ refers as the 2-dimensional TV operator in [75], therefore $\mathbf{D}$ evaluates the total variation band per band, and λ and λ_{TV} are two positive regularization parameters for adjusting the importance of the sparsity term and the TV term, respectively. For convenience, we introduce the following notations

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}1 \\ \mathbf{v}2 \\ \mathbf{v}3 \\ \mathbf{v}4 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} \mathbf{I} \\ \mathbf{B}s \\ \mathbf{\Psi} \\ \mathbf{D} \end{bmatrix},$$

and the cost function

$$\begin{aligned} h(\mathbf{v}) = \; & \tfrac{1}{2}\|\mathbf{U}_M\mathbf{E}\mathbf{v}_1 - \mathbf{y}_M\|_2^2 + \tfrac{1}{2}\|\mathbf{U}_H\mathbf{S}_d\mathbf{v}_2 - \mathbf{y}_H\|_2^2 \\ & + \lambda\|\mathbf{v}_3\|_1 + \lambda_{TV}\|\mathbf{v}_4\|_1 \end{aligned}$$

with $\mathbf{v} = \mathbf{C}\mathbf{x}$. Thus, (3.34) can be expressed as

$$\begin{aligned} \underset{\mathbf{x},\mathbf{v}}{\text{argmin}} \quad & h(\mathbf{v}) \\ \text{subject to} \quad & \mathbf{v} = \mathbf{C}\mathbf{x}. \end{aligned} \tag{3.35}$$

The augmented Lagrangian associated with (3.35) is

$$\mathcal{L}(\mathbf{x}, \mathbf{v}, \mathbf{g}) = h(\mathbf{v}) + \tfrac{\rho}{2}\|\mathbf{v} - \mathbf{C}\mathbf{x} + \mathbf{g}\|_2^2 \tag{3.36}$$

where $\mathbf{g}$ is the scaled dual variable and $\rho \geq 0$ is weighting the augmented Lagrangian term. The exact procedure used for estimating $\mathbf{x}$ is summarized in Algorithm 1.

Algorithm 4: ADMM algorithm to estimate $\mathbf{x}$

Input : $\mathbf{y}_M$, $\mathbf{y}_H$, $\mathbf{U}_M$, $\mathbf{U}_H$, $\mathbf{E}$, $\mathbf{S}_d$, $\mathbf{B}_s$, $\mathbf{G}$, $\mathbf{D}$, $\rho \geq 0$

Output: $\mathbf{x}^{(k+1)}$

$\mathbf{v}^{(0)}, \mathbf{g}^{(0)}$;

for $k \leftarrow 1$ **to** *stopping rule* **do**

  $\mathbf{x}^{(k+1)} = \underset{\mathbf{x}}{\text{argmin}}\ \mathcal{L}(\mathbf{x}, \mathbf{v}^{(k)}, \mathbf{g}^{(k)})$;

  $\mathbf{v}^{(k+1)} = \underset{\mathbf{v}}{\text{argmin}}\ \mathcal{L}(\mathbf{x}^{(k+1)}, \mathbf{v}, \mathbf{g}^{(k)})$;

  $\mathbf{g}^{(k+1)} = \mathbf{g}^{(k)} + \mathbf{v}^{(k+1)} - \mathbf{C}\mathbf{x}^{(k+1)}$;

Table 8: Mean PSNR in the PFA system, for different compression ratios with the $256 \times 256 \times 24$ Pavia image

$CR = 100(P+Q)/(Z^2L)[\%]$	9	12	17	21
Uniform	35.94	37.96	37.69	37.85
Unif+ Spatial Restriction	34.97	33.98	38.56	**38.88**
3D Spread	**37.60**	**38.20**	**38.63**	**38.88**
Random	32.03	35.46	31.76	31.68

3.5. **SIMULATIONS**

In this section, we present some simulations to illustrate the performance of the approaches for designing the coded apertures in a fusion framework showed in section 3.3, and compare them with random coded apertures. We test the designs in two different datasets, with the real high resolution image as reference image and the degraded images (MS and HS) obtained by simulation. For all the datasets, the MS image was generated by a spectral degradation of the reference image resulting in $L_M = L/p$ bands MS image. The HS was generated by applying a spatial blur $\bar{\mathbf{B}}$ to the reference image followed by a downsampling $\bar{\mathbf{S}}$ with a factor of q in each direction. The MS and HS measurements were generated by compressing the MS and HS images with sensing matrices whose entries were generated by different coded apertures design. The compression ratio is defined as $CR = 100(P+Q)/Z^2L$, where $(P+Q)$ is the total number of measurements in the two CSI systems for fusing. Additionally, the MS and HS measurements were both contaminated by additive Gaussian noise, with $SNR = 15$ [dB].

3.5.1. **Effect of the Compression Ratio** In this experiment we use two different databases, the hyperspectral image acquired over the University of Pavia (Italy) by the ROSIS-03 (Reflective Optics Systems Imaging Spectrometer) hyperspectral sensor, where we use a section of $256 \times 256 \times 24$, and the Butterfly HR image of size $128 \times 128 \times 24$ taken in our laboratory the HDSP research group optical laboratory.
In this case, the reduction of bands is $p = 4$, the blurring matrix is imposed to be the identity, and the downsampling factor is $q = 4$ resulting in an MS image of $256 \times 256 \times 6$ and an HS image of size $64 \times 64 \times 24$ in the Pavia case, and an MS image of $128 \times 128 \times 6$ and an HS image of $32 \times 32 \times 24$ for the Butterfly image.
Table 8 shows the results for the Patterned Filter Array (PFA) system for the Pavia University image. Notice that the larger the number of measurement, the greater the quality of the recovery. Similar quality results are shown for the Butterfly dataset in table 9. A similar behavior is shown in the tables 10 and 11 in the C-CASSI system. Figure 13 shows the reconstructed images for the Pavia University scene and figure 14 shows the reconstructed images for the Butterfly using the PFA system with 17% of compression ratio. Notice that in both cases the results for the designed CCA outperform those for random CCA.

Figure 13: Reconstructions of the Pavia University scene for the proposed designs with 17% of compression ratio for the PFA system

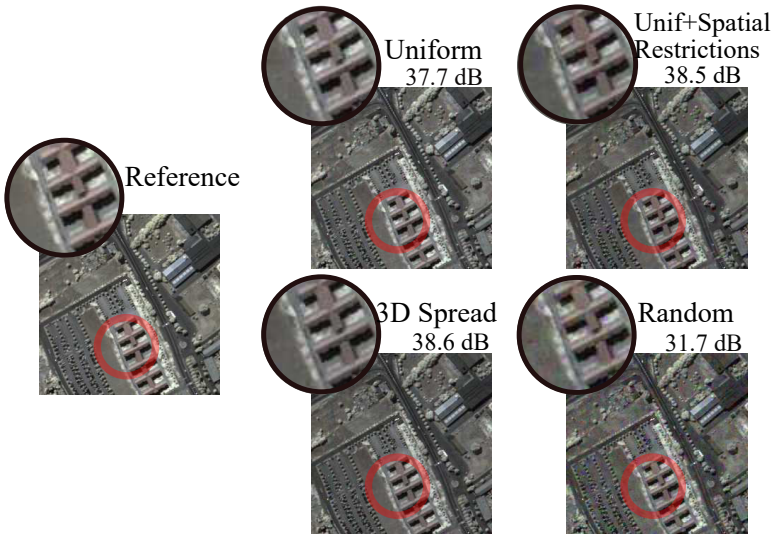


Figure 14: Reconstructions of the butterfly scene for the proposed designs with 17% of compression ratio for the PFA system

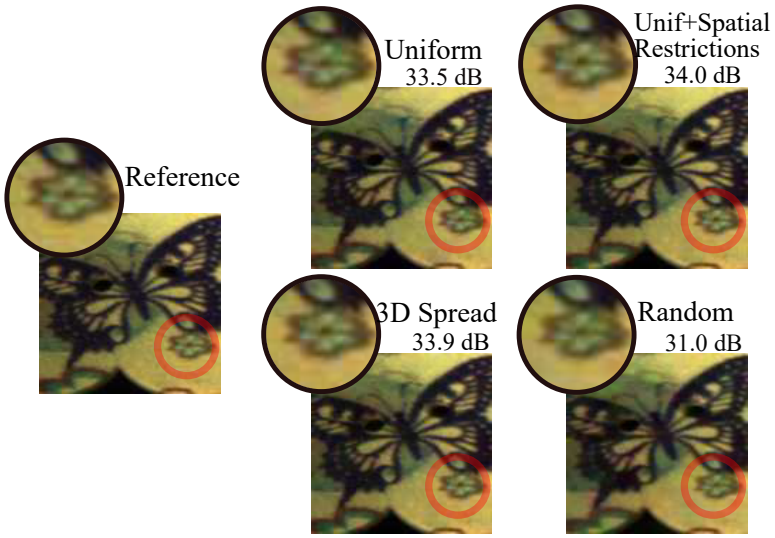


Table 9: Mean PSNR in the PFA system, for different compression ratios with the $128 \times 128 \times 24$ Butterfly image

$CR = 100(P+Q)/(Z^2L)[\%]$	9	12	17	21
Uniform	31.84	32.68	33.48	33.51
Unif+Spatial Restriction	32.20	32.87	**34.00**	33.65
3D Spread	**32.43**	**33.67**	33.99	**34.37**
Random	30.04	30.64	31.02	31.71

Table 10: Mean PSNR in the C-CASSI system, for different compression ratios with the $256 \times 256 \times 24$ Pavia image

$CR = 100(P+Q)/(Z^2L)[\%]$	9	13	18	22
Uniform	29.11	31.67	32.35	33.56
Unif+Spatial Restriction	28.96	31.25	32.58	**33.69**
3D Spread	**30.19**	**31.69**	**32.84**	33.68
Random	29.11	29.88	29.92	29.83

Notice that in all the cases the designed CCA outperforms the Random CCA. The biggest differences between the designed CCA and the Random ones are for the PFA system of 3 dB for the Pavia image, with the 3D spread design and 12% of CR of 7.2 dB for the Butterfly image, with the 3D spread design and 21% of CR for the C-CASSI system of 3.86 dB for the Pavia image, with the Unif+Spatial Restriction design and 22% of CR and of 1.58 dB for the butterfly image, with the Uniform and 17% of CR. In comparison with the C-CASSI the PFA the system obtains better results with similar compressive ratios, up to 7 dB.

3.6. **CHAPTER CONCLUSION**

This chapter exposes three CCA designs for improving the image reconstruction in a framework of fusing measurements of two CSI systems. For the designs, we analyzed and modeled the structure of the matrices related to spectral reduction, spatial downsampling, and compression.

Table 11: Mean PSNR in the C-CASSI system, for different compression ratios with the $128 \times 128 \times 32$ Butterfly image

$CR = 100(P+Q)/(Z^2L)[\%]$	9	13	17	21
Uniform	31.28	32.23	**33.61**	33.4
Unif+ Spatial Restriction	31.20	32.62	32.85	33.25
3D Spread	**31.55**	**32.75**	33.25	**33.97**
Random	30.68	31.31	32.03	32.91

As a result, we first proposed a Uniform Sensing strategy that attempts to locate the pass elements in the CCA uniformly random such that the sum of pass elements along the snapshots be one and that the sum of pass elements along the spectral bands in each snapshot be L/K. The benefit of this design is that each CSI system will have suitable properties of linearly independent measures. The drawback of this design is that it ignores the spatial constraints related to the fusion of the two CSI systems measurements.

For including the spatial restrictions in the MS CSI system, a second design is proposed such that, additional to the uniform sensing constraints, it tries to keep the same average of pass elements in spatial windows of $p \times p$ pixels. The benefit of this approach is that spreading the pass elements in the spatial neighborhoods of the CCA for the MS system decreases the interactions with the elements of the CCA for the HS system. The drawback of this design is that requires an additional step that moves through $(Z/q)^2 L$ spatial neighborhoods.

The third design is a generalization of the required spreading of the pass elements along the 3D CCA. The aim of this approach is to spread the pass elements of the CCA in 3D sub-windows by minimizing an objective function. A possible algorithm to perform the optimization is to generate random realizations of the coded aperture, to move a 3D-Euclidean filter that promotes the spreading of the pass elements, to evaluate the objective function, and to preserve the one that minimizes the objective function. The advantage of this design is that the correlations between both CSI systems will be reduced by spreading the pass elements along 3D subwindows of each compressive system. The drawback of this design is the increasing of the complexity in the design.

The proposed designs were tested on two different CSI systems and two datacubes. Numerical experiments showed that the proposed designs compared competitively with the Random distribution of the pass elements, with the advantage of improving the reconstruction quality up to 5dB. Between the proposed designs the Uniform with local spatial constraints and the 3D spread outperform the Uniform sensing, in the majority of cases, up to 2 dB.

4. CSI BASED ON MULTISPECTRAL FILTER ARRAY

Assuming that the MSFA-based sensor performs a linear measurement process that computes $M \ll N$ internal products between the vectorized spectral image $\mathbf{x}$ and a collection of vectors $\{\mathbf{\Phi}_j\}_{j=1}^{M}$, as $[\mathbf{y}]_i = \langle \mathbf{x}, \mathbf{\Phi}_j \rangle$ then $\mathbf{y} = \mathbf{\Phi x}$, where the set of $[\mathbf{y}]_i$ projections form the vector $\mathbf{y}$ of m elements, $\mathbf{\Phi}$ is the measurement matrix with dimensions $M \times N$, with $\mathbf{\Phi}_j^T$ rows, and $\mathbf{x}$ is the original signal of size N.
Some works have proposed the demosaicing process by CS theory. Sadeghipoor et al. [76] use an MSFA of 4 spectral bands (Red-Green-Blue-NIR), similar to the traditional RGB. The difference is that the two green filters have different transmittances. Here, transmittance means the amount of light the filter transmits in the given wavelength range. Also, a mixture of one color channel and the NIR is captured at each spatial position on the sensor. For demosaicing task, they use a CS approach for separating the NIR of the color bands, after that a traditional RGB demosaicing for the complete reconstruction. Aggarwal and Majumdar [77] explore the problem of CS reconstruction of multi-spectral images acquired with a single sensor architecture. They propose random and uniform filter array designs. For reconstruction, they use a group-sparse optimization and the Kronecker product between Fourier and Wavelet basis.
These methods reconstruct fixed resolution data cubes, there is not an introduced approach to adjust the spatial resolution of the reconstructed images. Probably the selection of neighborhoods measurements to estimate the spectral response of a single pixel optimizes CS reconstruction from the MSFA-based sensor measurements. The neighborhoods size leads to an adjustable spatial resolution in the reconstruction preserving the filters spectral resolution, reconstructing a spatial decimated data cube. This information can be used in applications requiring higher spectral than spatial image quality, also for a quick view of the scene, for instance, in transmission and communication applications.
This chapter considers the spectral images recovery from the measurements of the multispectral filter array-based sensor by using CS. We propose two models for adjustable spatial resolution reconstruction of multispectral images from the selection of MSFA-based sensor measurements neighborhoods. They are based on measurements taken in an architecture that includes an MSFA-based sensor. Each spectral filter modulates the data before it impinges onto the sensor using a random dichroic or random bandpass filter. CS theory is then exploited to recover the underlying 3D spectral data cube from the compressed data captured in a single shot. The quality of the reconstructions is analyzed based on the number of spectral bands, and the size of the neighborhoods.

4.1. TRADITIONAL DEMOSAICING

Given the set of measurements $\mathbf{y}$, a traditional demosaicing algorithm estimates for each reconstructed pixel the intensities for all wavelength components. In traditional cases, measurements are taken for a mosaic of multi-spectral filters, where a pattern is formed subject to the number of spectral bands to sense. The most commonly used configuration is the Bayer filter for RGB reconstructions [78]. For reconstruction, the common approach minimizes the linear mean square error between the measurements and the vector estimation multiplied by the sensing matrix. More specifically, the estimated signal is given by

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 \tag{4.1}$$

A closed-form solution of (4.1) is given by

$$\tilde{\mathbf{x}} = (\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\mathbf{y} = \mathbf{\Phi}^\dagger\mathbf{y}, \tag{4.2}$$

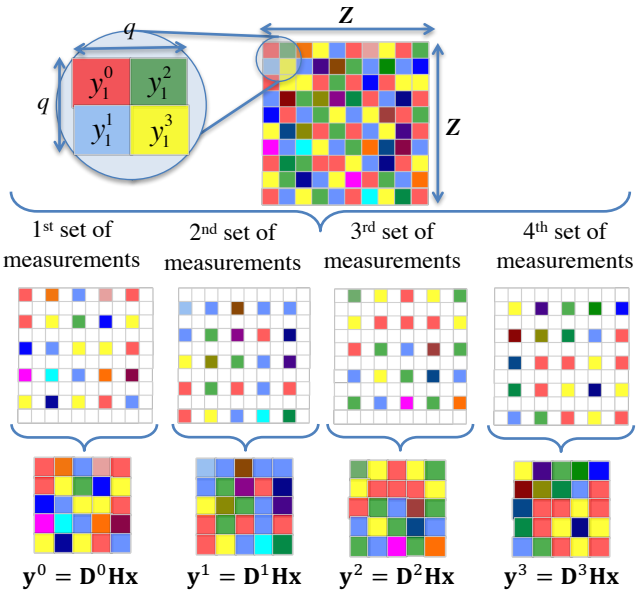
where $\mathbf{\Phi}^\dagger$ is known as the pseudo-inverse of $\mathbf{\Phi}$, and $\mathbf{\Phi}^T$ is its transpose.

4.2. APPROACHES FOR RECONSTRUCTION WITH ADJUSTABLE SPATIAL RESOLUTION

Notice that the methods explained in this section are carried out after the sensing process. Thus, the resolution of the sensor remains fixed, however, the methods explained in this section provide spectral images reconstructions with different resolutions. For exploiting CS theory and the measurements of the MSFA-based sensor, two approaches are proposed for reconstructing a datacube with an adjustable spatial resolution. Both are related to selecting a neighborhood of $q \times q$ pixels for reconstructing a single pixel. The parameter q is selected at the time to reconstruct the data cube, q refers to the neighborhood side size. In the first model, all the $Z \times Z$ pixels of the sensor (measurements) are taken for reconstructing a $Z \times Z \times L$ datacube following of a spatial average decimation for the given size of the neighborhood. The second model selects neighborhoods of measurements and reformulates the CS reconstruction problem for reconstructing a $Z/q \times Z/q \times L$ datacube.

4.2.1. **Decimation of a complete data cube reconstruction** The underlying datacube projection in this case is reconstructed solving an $\ell_1 - \ell_2$ minimization problem [14]. In this case, for using the neighborhood approach the data cube spatial resolution is adjusted following an average decimation matrix operation. This operation is applied to the reconstructed data cube as $\tilde{\mathbf{x}}_q = \mathbf{B}\tilde{\mathbf{x}}$, where $\mathbf{B}$ does a block averaging of size $q \times q$ in each spectral band reducing the size of the vectorized reconstructed

Figure 15: Example of $q = 2$ that forms 4 subsets of measurements in a single snapshot of the MSFA-based sensor.



data cube to Z^2L/q^2, and q is the side size of the neighborhood. The spatial decimation matrix element-by-element can be expressed as:

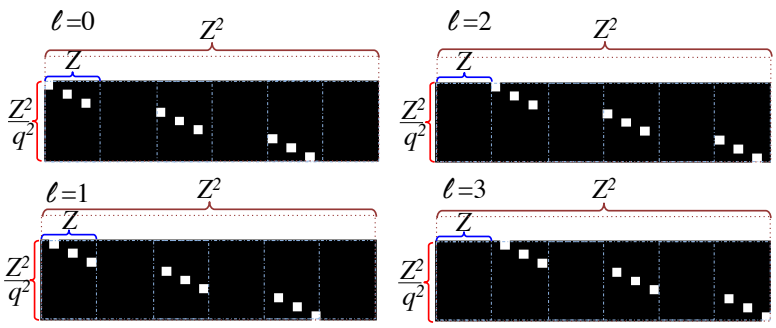
$$B_{i,j} = \begin{cases} 1/q^2, & \text{if } i = \left\lfloor \frac{j}{q} \right\rfloor - \frac{Z}{q}\left\lfloor \frac{j}{Z} \right\rfloor + \frac{Z}{q}\left\lfloor \frac{j}{qZ} \right\rfloor, \\ 0, & \text{otherwise,} \end{cases} \tag{4.3}$$

where $i = 0, 1, ..., \frac{Z^2L}{q^2} - 1$.

4.2.2. **Reconstruction based on neighborhood measurements selection** This model uses the assumption that $q \times q$ neighboring pixels, in a $Z \times Z \times L$ spectral image, have the same spectral response. Then $q \times q$ neighborhood measurement pixels in the sensor are taken for the reconstruction of a single pixel spectral response in a decimated reconstruction. For instance, Fig. 16 shows the set of measurements taken for $q = 2$, in total q^2 subsets of measurements are grouped from a single snapshot of the MSFA-based sensor.
In the reconstruction model, each subset of measurements is selected from the total measurements by performing a matrix product between a downsampling matrix that accomplishes the function of selecting the specific set of measurements, and the vectorized version of the measurements. More specifically, each subset is given by

$$\mathbf{y}^\ell = \mathbf{D}^\ell\mathbf{H}\mathbf{x}, \tag{4.4}$$

Figure 16: The downsampling matrix $\mathbf{D}^\ell$ for $q = 2$, $Z = 6$, and $\ell = 0, 1, 2, 3$. The white squares represent one-valued elements.



where $\mathbf{D}^\ell$ performs the downsampling in each $q \times q$ neighborhood of pixels for taking q^2 different subsets of the total measurements and $\mathbf{y}^\ell$ is the ℓth subset of measurements where $\ell \in \{0, ..., q^2 - 1\}$. Notice that q is a parameter that can be selected by the user. Precisely, the function of the downsampling matrix $\mathbf{D}^\ell$ is selecting the ℓth element in each $q \times q$ neighborhood for forming the $\ell-$th subset of measurements. The (i, j)th element of the ℓth downsampling matrix can be expressed as:

$$D^\ell_{i,j} = \begin{cases} 1, & \text{if } \; j = iq + (q-1)Z\left\lfloor \frac{iq}{Z} \right\rfloor + \ell + \left\lfloor \frac{\ell}{q} \right\rfloor (Z - q), \\ 0, & \text{otherwise}, \end{cases} \tag{4.5}$$

where $i = 0, 1, ..., \left\lfloor \frac{Z^2}{q^2} \right\rfloor - 1$, $j = 0, 1, ..., Z^2 - 1$, and $\ell = 0, 1, ..., q^2 - 1$. Figure 7 depicts the downsampling matrix $\mathbf{D}^\ell$ for $q = 2$, $Z = 6$, and $\ell = 0, 1, 2, 3$. The white squares represent one-valued elements. In this case, the complete set of measurements is rearranged as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}^0 \\ \mathbf{y}^1 \\ \vdots \\ \mathbf{y}^{q^2-1} \end{bmatrix} = \begin{bmatrix} \mathbf{D}^0 \\ \mathbf{D}^1 \\ \vdots \\ \mathbf{D}^{q^2-1} \end{bmatrix} \mathbf{H}\mathbf{f} = \mathbf{H}_q \mathbf{f}, \tag{4.6}$$

where the underlying datacube projection is also reconstructed solving an $\ell_1 - \ell_2$ minimization problem. However, in this case the decimation process is taken into account. More specifically, the optimization problem is given by

$$\tilde{\mathbf{x}} = \mathbf{\Psi} \left(\arg\min_{\boldsymbol{\theta}} \|\mathbf{y} - \mathbf{H}_q \mathbf{\Psi} \boldsymbol{\theta}\|_2^2 + \tau \|\boldsymbol{\theta}\|_1 \right) \tag{4.7}$$

where $\mathbf{y}$ is given by (4.6), and $\boldsymbol{\theta}$ is an S-sparse representation of a $Z^2 L/q^2$ version of $\mathbf{x}$ on the basis $\mathbf{\Psi}$.

4.3. MULTISPECTRAL FILTER DESIGN

The quality of the reconstructed datacubes depends on the selection of the multispectral filter design. For developing this work, given the possible flexibility in the bandpass of each MSFA element, two MSFA were selected. First, a spectral response for each MSFA element that can be selected randomly from a set of bandpass filters; and second, dichroic filters, that is, a special case of random bandpass whose spectral response lets only one spectral band to pass for each sensor pixel. These designs still follow a random selection of elements and the motivation of differentiating them is to evaluate the performance of combining more than one spectral band, for the bandpass case, in comparison with the dichroic filters.
The spectral response of a $\lambda_i^{\mathcal{L}}, \lambda_i^{\mathcal{H}}$ band-pass filter can be defined as

$$(\mathbf{t}_k)_i = \begin{cases} 1, & \text{if } \lambda_i^{\mathcal{L}} \leq k \leq \lambda_i^{\mathcal{H}}, \\ 0, & \text{otherwise}, \end{cases} \tag{4.8}$$

for $k = 0, 1, ..., L-1$, $i = 0, 1, ..., Z^2 - 1$, and $\lambda_i^{\mathcal{L}} \leq \lambda_i^{\mathcal{H}} \in \{0, 1, ..., L-1\}$. For instance, $L = 4$, $\lambda_1^{\mathcal{L}} = 2$, and $\lambda_1^{\mathcal{H}} = 3$ define the spectral response of the spatial position $i = 1$ as $(\mathbf{t}_k)_1 = [0\ 0\ 1\ 1]$. Dichroic filters are a special case of band-pass filters that let pass only one spectral band. Then the spectral response of a $(\lambda_i^{\mathcal{D}})$ dichroic filter pixel can be defined as

$$(\mathbf{t}_k)_i = \begin{cases} 1, & \text{if } \lambda_i^{\mathcal{D}} = k, \\ 0, & \text{otherwise}, \end{cases} \tag{4.9}$$

for $\lambda_i^{\mathcal{D}} \in \{0, .., L-1\}$ and $i = 0, ..., Z^2 - 1$. For example, if $L = 4$, and $\lambda_5^{\mathcal{D}} = 3$, then the spectral response of the spatial position $i = 5$ in the vectorized MSFA is $(\mathbf{t}_k)_5 = [0\ 0\ 0\ 1]$.

4.4. SIMULATION AND RESULTS

To verify the MSFA-based sensor reconstructions, a set of compressive measurements is simulated using the model of eq. (2). These measurements are constructed with two spectral images captured with a CCD camera Apogee Alta U260 and a VariSpec liquid crystal tunable filter, in the range of wavelength 400nm-560nm, with steps of 10nm [15]. The resulting test data cubes have 512×512 pixels of spatial resolution and $L = 16$ spectral bands. The grayscale images mapped versions of the selected data cubes are shown in Fig. 9. The experiments were carried out using the images Balloons and Beads with a decimation processing for creating synthetic ones and in their real form. Compressive sensing reconstruction is performed by using the GPSR algorithm [16]. Simulations results are analyzed in terms of PSNR (Peak-Signal-to-Noise-Ratio) of the reconstructed images. The representation basis $\mathbf{\Psi}$ is a Kronecker product $\mathbf{\Psi} = \mathbf{\Psi}_1 \otimes \mathbf{\Psi}_2$, where $\mathbf{\Psi}_1$ is the two-dimensional-wavelet Symmlet-8 basis and $\mathbf{\Psi}_2$ is the cosine basis. The simulations are performed in a desktop architecture with an Intel Core i7 3.6GHz processor, 32GB RAM, and using

Matlab R2014a. Each experiment is repeated ten times and the respective results are averaged.

4.4.1. **Synthetic multispectral data cubes** When the input data cube fits the neighborhood measurements selection model the reconstructions outperform the traditional demosaicing approach. To illustrate this, synthetic multispectral datacubes based on Balloons and Beads images are constructed. For that, an average decimation of factor q followed of a duplication of q pixels is applied in each spectral band. Then the synthetic data cubes satisfy the assumption that some neighboring pixels attain the same spectral response. Ensembles of dichroic and band pass filters based on a random selection of spectral bands are used for obtaining measurements. To simulate the measurements, the first L spectral bands of synthetic data cubes are used. Then selecting a value of q the two CS approaches reconstruct a decimated data cube. For comparison, a demosaicing traditional process using dichroic filters mosaics is implemented, after performing the reconstruction an average decimation matrix is applied to the reconstructed data cube. Figure 17 shows a comparison of the average PSNR reconstruction, for the synthetic Balloons data cubes, as a function of the number of sensed and reconstructed spectral bands. The first row is related to the reconstructions of measurements sensed with an ensemble of band pass filters. The frequency response of the bandpass filter is selected at random. The second row depicts the results from random dichroic filters measurements. The columns are associated to the reconstruction varying the neighborhood size to $q = 2, 4, 8$. For example, Fig. 18(a) shows the results for band pass filters with a neighborhood of $q = 2$, and Fig. 18(f) the results for dichroic filters with a neighborhood of $q = 8$. The PSNR evaluation is calculated between the reconstructed image and a spatial decimated version (of size $Z/q \times Z/q \times L$) of the test data cube. Figure 18 shows similar results for synthetic Beads. In all the cases, with increasing the number of spectral bands decreases the PSNR. Also, it is possible to observe the improvement of the reconstruction based on neighborhood measurements selection method (neighborhood CS) compared to the decimation of a complete data cube reconstruction method (complete CS) and the traditional demosaicing method, when the data cube has the property of having similar spectral signature in a neighborhood $q \times q$.

4.4.2. **Experiment with real multispectral data cubes** In this case, the measurements are simulated using L spectral bands of real data cubes. Ensembles of dichroic and band pass filters based on a random selection of spectral bands are used to obtain measurements. Figure 19 shows a comparison of the average PSNR reconstruction for the Balloons data cube. Figure 20 shows similar results for PSNR reconstructions in the Beads data cube. It can be observed in Fig. 19 and Fig. 20 that the performance of the CS proposed methods improves when the size of the neighborhood is increased. Additionally, the reconstruction based on neighborhood

Figure 17: For the synthetic Balloons data cube, (first row) reconstruction results for band pass filters for (a) $q = 2$, (b) $q = 4$, and (c) $q = 8$, (second row) reconstruction results for dichroic filters for (d) $q = 2$, (e) $q = 4$, and (f) $q = 8$.

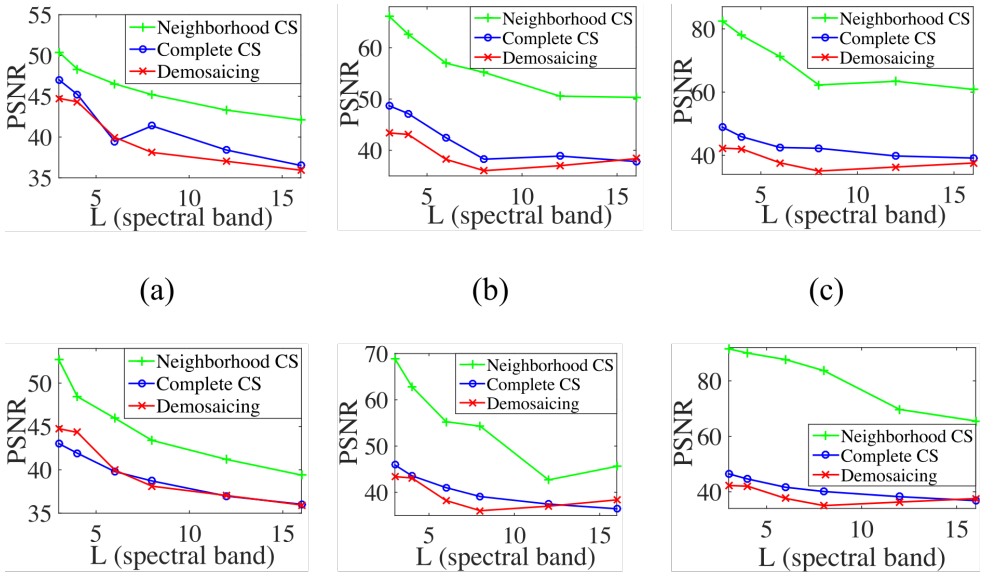


Figure 18: For the synthetic Beads data cube, (first row) reconstruction results for band pass filters for (a) $q = 2$, (b) $q = 4$, and (c) $q = 8$, (second row) reconstruction results for dichroic filters for (d) $q = 2$, (e) $q = 4$, and (f) $q = 8$.

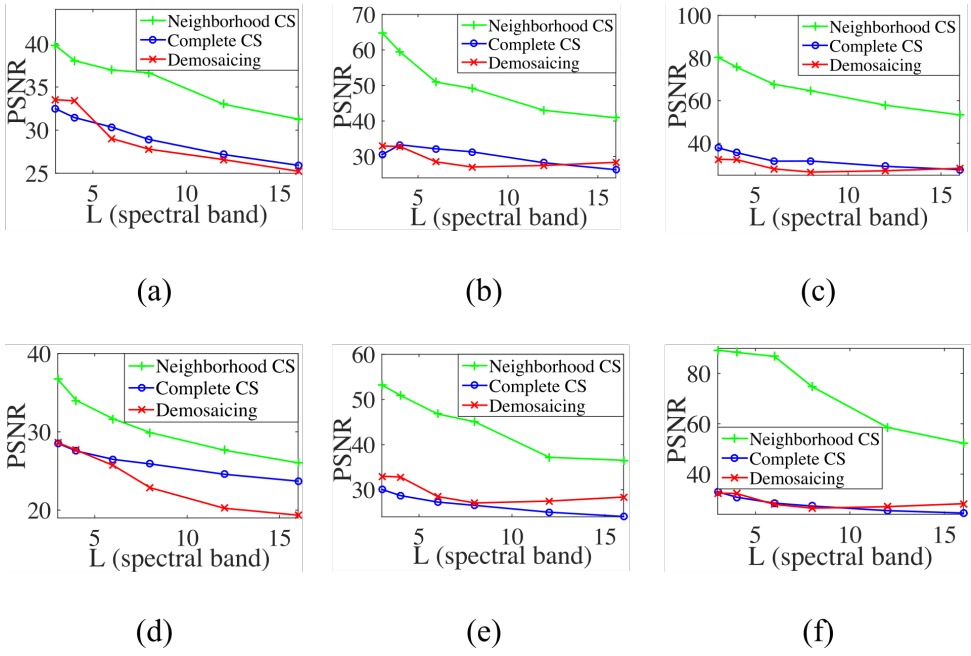
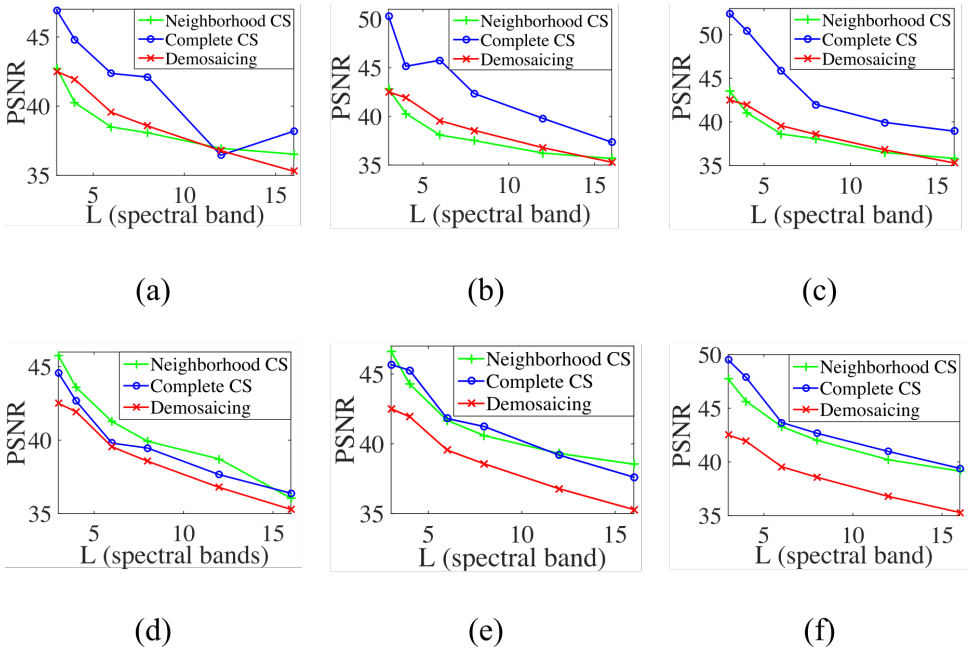


Figure 19: For the Beads data cube, (first row) reconstruction results for band pass filters for (a) $q = 2$, (b) $q = 4$, and (c) $q = 8$, (second row) reconstruction results for dichroic filters for (d) $q = 2$, (e) $q = 4$, and (f) $q = 8$.



measurements selection method (neighborhood CS) approach has a better performance than the traditional demosaicing approach with the measurements captured with dichroic filters. On the contrary, the decimation of a complete data cube reconstruction (complete CS) takes advantage of band pass filters.

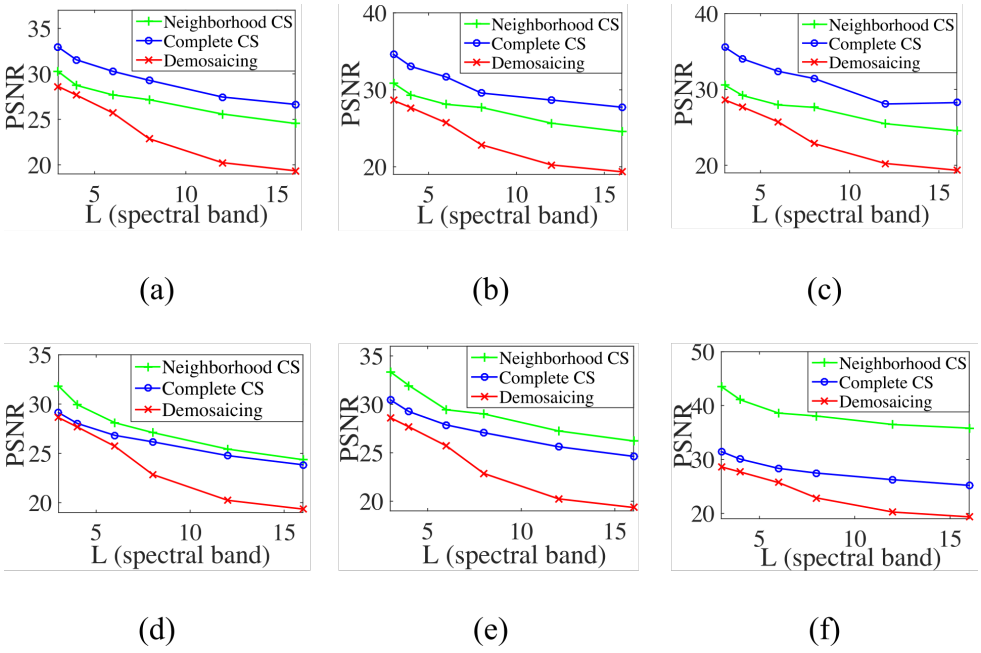
Figure 21 illustrates the whole reconstructed data cubes mapped in gray-scale, for $L = 6$, and $q = 2$ neighborhood side size. For the Balloons database, it can be observed that the reconstruction of measurements taken for random dichroic filter ensemble using the neighborhood CS approach provides an improvement of up 1.72dB in PSNR over the demosaicing approach.

Furthermore, the PSNR reached by the complete CS reconstruction using band pass filters is 2.81dB higher than the demosaicing reconstructions. For the Beads database, the reconstruction of neighborhood CS using dichroic filters improves up 2.4dB over the demosaicing, and the complete CS using band pass filter up 4.58dB.

4.5. CHAPTER CONCLUSION

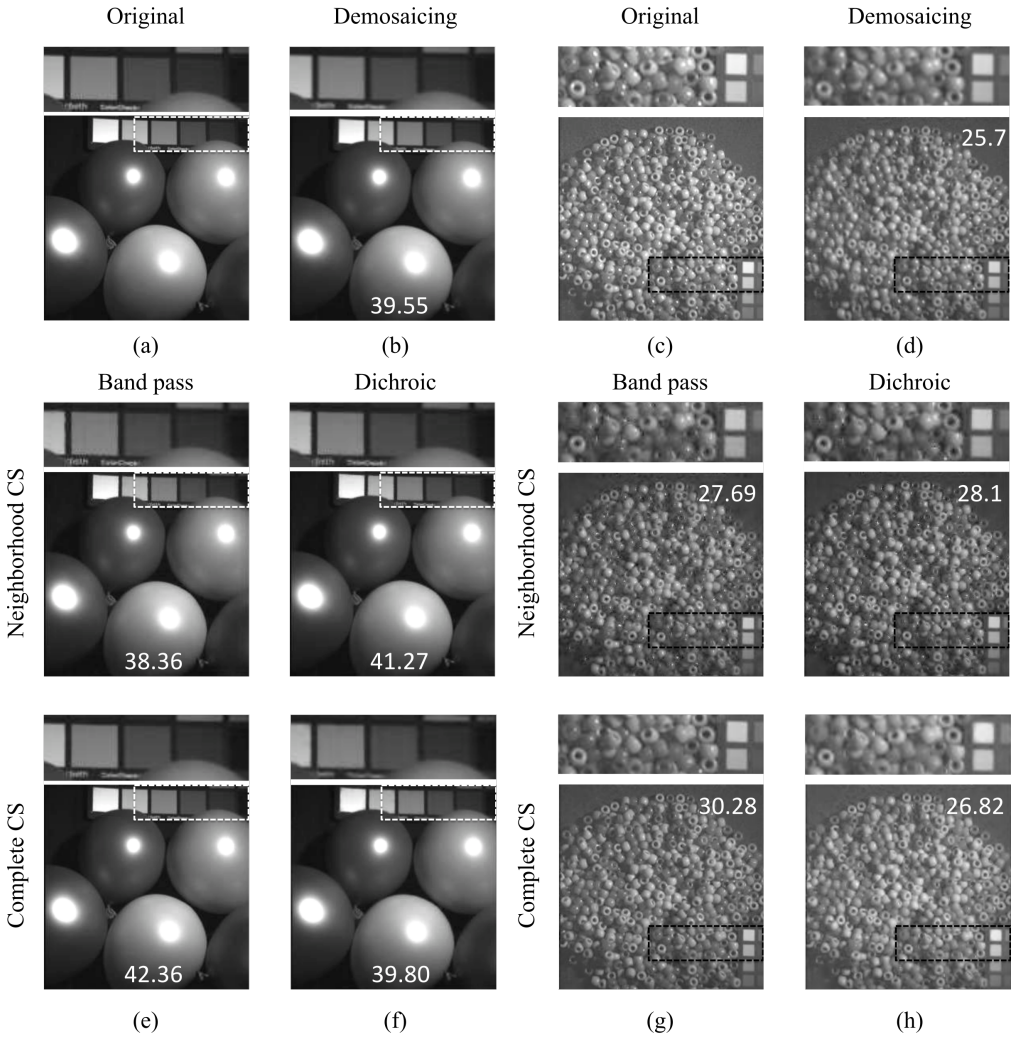
In this chapter we presented two models for CS reconstruction of spectral images sensed by MSFA-based sensors using a neighborhood approach. The first model reconstructs a complete data cube and applies neighborhood decimation. The second model performs a selection of measurements subsets to form neighborhoods that have spectral information of a single reconstructed pixel. The two CS reconstruction approaches are compared with a traditional demosaicing reconstruction

Figure 20: For the Beads data cube, (first row) reconstruction results for band pass filters for (a) $q = 2$, (b) $q = 4$, and (c) $q = 8$, (second row) reconstruction results for dichroic filters for (d) $q = 2$, (e) $q = 4$, and (f) $q = 8$.



method. For the CS reconstruction, the PSNR increases with the neighborhood side size. The improvements range from 0.5 dB to 7 dB with respect to the traditional approach in real data cubes. Results show that increasing the number of spectral bands decreases the PSNR for all reconstruction methods.

Figure 21: Gray-scale versions of reconstructions for $L = 6$, and $q = 2$ neighborhood side size. For the Balloons database: (a) original, and (b) demosaicing reconstruction. For the CS approach reconstructions using (e) band pass filters, and (top) neighborhood CS, (bottom) complete CS; (f) dichroic filters, and (top) neighborhood CS, (bottom) complete CS. Similar results (c-d) and (g-h) for Beads datacube. The number in the images is the PSNR of the reconstruction.



5. OPTIMIZED COMPRESSIVE RECONSTRUCTION ALGORITHM BASED ON FILTERED GRADIENT

The problem of recovering the spectral image $\mathbf{x}$ from the compressive measurements $\mathbf{y} = \mathbf{\Phi x}$ can be posed and solved as a convex unconstrained optimization problem [8], [12], [18], [33]. It consists of minimizing an objective function that includes a quadratic error term combined with a sparsity-inducing regularization term, given by

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \frac{1}{2} \left\| \mathbf{y} - \mathbf{A}\boldsymbol{\theta} \right\|_2^2 + \tau \left\| \boldsymbol{\theta} \right\|_1, \tag{5.1}$$

where $\boldsymbol{\theta}$ is the sparse representation of $\mathbf{x}$ in a basis $\mathbf{\Psi}$, such that $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$, $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}$, and τ is a regularization parameter for the sparsity prior [1], [3], [5]. When the solution of Eq. (5.1) is known in advance to be non-negative, the problem can be rewritten as [4], [25]

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} (\tau \mathbf{1}_n - \mathbf{A}^T\mathbf{y})^T\boldsymbol{\theta} + \frac{1}{2}\boldsymbol{\theta}^T\mathbf{A}^T\mathbf{A}\boldsymbol{\theta}, \quad \text{s.t} \quad \boldsymbol{\theta} \geq 0. \tag{5.2}$$

Eq. (5.2) can be seen as a standard bound constrained quadratic problem (BCQP) given by

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}}\, \mathbf{c}^T\boldsymbol{\theta} + \frac{1}{2}\boldsymbol{\theta}^T\mathbf{B}\boldsymbol{\theta} \equiv \underset{\boldsymbol{\theta}}{\text{argmin}} \quad F(\boldsymbol{\theta}), \quad \text{s.t} \quad \boldsymbol{\theta} \geq 0, \tag{5.3}$$

where $\mathbf{c} = (\tau \mathbf{1}_n - \mathbf{A}^T\mathbf{y})$, $\mathbf{1}_n$ is a n-long vector of ones, and $\mathbf{B} = \mathbf{A}^T\mathbf{A}$.
There are several algorithms that solve the compressive sensing problem (5.3) [4], [41]–[44]. A well studied method for solving (5.3) is the gradient descent algorithm [79]–[82]. This method computes a sequence of points following the opposite direction of the gradient, given by

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \alpha \nabla F(\boldsymbol{\theta}^{(k)}), \tag{5.4}$$

where $\nabla F(\boldsymbol{\theta}^{(k)})$ is the gradient of the function F, α is the step size, and $k = 0, 1, ...$ is the iteration. The gradient of $F(\boldsymbol{\theta}^{(k)})$ in (5.3) is calculated as

$$\nabla F(\boldsymbol{\theta}^{(k)}) = \begin{cases} \mathbf{c} + \mathbf{B}\boldsymbol{\theta}^{(k)} = \tau \mathbf{1}_n - \mathbf{A}^T(\mathbf{y} - \mathbf{A}\boldsymbol{\theta}^{(k)}), & \text{if} \quad \tau \mathbf{1}_n < \mathbf{A}^T(\mathbf{y} - \mathbf{A}\boldsymbol{\theta}^{(k)}) \\ 0, & \text{otherwise}, \end{cases} \tag{5.5}$$

to ensure the restriction that the reconstructed signal is positive. Therefore, replacing (5.5) in (5.4), the gradient sequence of (5.3) is given by

$$\begin{aligned} \boldsymbol{\theta}^{(k+1)} &= \boldsymbol{\theta}^{(k)} - \alpha(\tau \mathbf{1}_n - \mathbf{A}^T(\mathbf{y} - \mathbf{A}\boldsymbol{\theta}^{(k)})) \\ &= \boldsymbol{\theta}^{(k)} - \alpha(\mathbf{A}^T\mathbf{A}\boldsymbol{\theta}^{(k)} - \mathbf{A}^T\mathbf{y} + \tau \mathbf{1}_n) \\ &= \boldsymbol{\theta}^{(k)} - \alpha\mathbf{\Psi}^T(\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\theta}^{(k)} - \mathbf{\Phi}^T\mathbf{y}) - \alpha\tau \mathbf{1}_n, \end{aligned} \tag{5.6}$$

where $\mathbf{A}$ has been substituted for $\mathbf{A} = \mathbf{\Phi\Psi}$. The gradient descent method adopts $\left\|\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)}\right\|_2^2 \leq \delta$ or equivalent $\left\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\right\|_2^2 \leq \delta_2$ as the stop criterion, where $\mathbf{x}^{(k)} = \mathbf{\Psi}\boldsymbol{\theta}^{(k)}$. From Eq. (5.6) the difference between iterations $\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)}$ is given by

$$\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)} = \alpha\mathbf{\Psi}^T\kappa(\boldsymbol{\theta}^{(k)}) - \alpha\tau\mathbf{1}_n, \tag{5.7}$$

such that the variable term $\kappa(\boldsymbol{\theta}^{(k)})$ is calculated by

$$\kappa(\boldsymbol{\theta}^{(k)}) = \mathbf{x}^{ref} - \mathbf{x}^{ref(k)}, \tag{5.8}$$

where $\mathbf{x}^{ref} = \mathbf{\Phi}^T\mathbf{y}$ and $\mathbf{x}^{ref(k)} = \mathbf{\Phi}^T\mathbf{\Phi\Psi}\boldsymbol{\theta}^{(k)}$. The difference in (5.7) stops changing when $\kappa(\boldsymbol{\theta}^{(k)}) = 0$, thus when

$$\mathbf{x}^{ref(k)} = \mathbf{x}^{ref}. \tag{5.9}$$

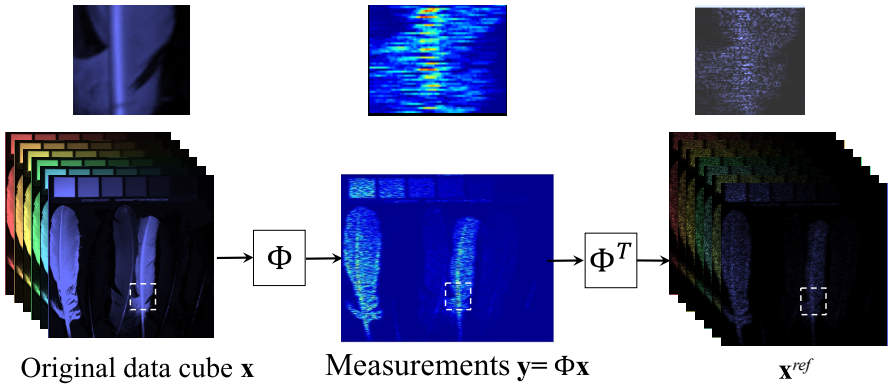
Since the matrix $\mathbf{\Phi}$ has a specific structure, its operation $\mathbf{\Phi}^T$ in the recovery algorithm can be exploited. The product by $\mathbf{\Phi}^T$ in each iteration returns rough approximations of the underlying signal $\mathbf{x}$, but this structure is given by the measurement system. Then, an interpolation action could be taken for improving the approximation. An intuitive way to solve the problem is using filtering operations before the $\mathbf{\Phi}^T$ products in each iteration. In this chapter, we propose a filtering step for solving the compressive problem into the gradient algorithm. For that, the optimization problem is solved using a minimizing sequence that is filtered using a matrix operation, a specific example in compressive spectral imaging of the approach is shown, where there are significant improvements in peak signal to noise ratio (PSNR) compared to the traditional approach.

5.1. **PROPOSED FILTERING STEP**

Notice that because of the sparse structure of the CSI sensing matrices the quantity $\mathbf{x}^{ref}$ can be seen as a rough approximation of the original data cube $\mathbf{x}$. To visualize this approximation, note that $\mathbf{y} = \mathbf{\Phi x}$ and $\mathbf{x}^{ref} = \mathbf{\Phi}^T\mathbf{y}$, such that $\mathbf{x}^{ref}$ can be seen as an estimation of $\mathbf{x}$ when the inverse of $\mathbf{\Phi}$ is approximated as $\mathbf{\Phi}^{-1} \cong \mathbf{\Phi}^T$. In a real scenario, $\mathbf{\Phi}^T\mathbf{\Phi}$ differs from the identity matrix and $\mathbf{\Phi}^T$ is an imprecise approximation of $\mathbf{\Phi}^{-1}$ such that the error $\left\|\mathbf{x} - \mathbf{x}^{ref}\right\|_2^2$ is considerably high. However, as it will be illustrated in short, $\mathbf{x}^{ref}$ preserves the smoothness characteristic of the signal $\mathbf{x}$. Thus from Equations (5.7) and (5.8) $\mathbf{x}^{ref}$ is an objective point for the algorithm in (5.6), such that if $\mathbf{x}^{ref}$ can be properly modified, the algorithm converges faster and can possibly obtain a better error approximation $\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2$, where $\hat{\mathbf{x}}$ is the result of the final iteration in (5.6).

On the other hand, natural spectral images $\mathbf{x}$ are smooth in their primary domain. For example, natural images can be represented as a linear combination of its neighbors in the spatial domain, such as $\mathbf{x} = \mathbf{Wx}$, where $\mathbf{W}$ is a weighted matrix [28], [83]. Specifically, to exploit this characteristic of natural signals we are going to modify the

Figure 22: Example of the products $\mathbf{x} = \mathbf{\Phi y}$ and $\mathbf{x}^{ref} = \mathbf{\Phi}^T\mathbf{y}$ for a spectral band in the CASSI system.



Eq. (5.6) by including a filtering operation to the terms $\mathbf{x}^{ref}$ and $\mathbf{x}^{ref(k)}$ to get a closer solution to a natural spectral image.
Given the structured and sparse characteristic of $\mathbf{\Phi}$ in the CSI systems, the spectral image approximation $\mathbf{x}^{ref}$ loses its smoothness and presents elements close to zero, like gaps. This characteristic disappears when sampling is performed by a dense matrix. Figure 22 shows an example of the measurements $\mathbf{y}$ and the reference $\mathbf{x}^{ref}$ for the CASSI sensing of a spectral image. The zoom versions in Figure 22 show that the $\mathbf{x}^{ref}$ is smooth and is highly correlated with the original data cube. Also, note in Figure 22 that $\mathbf{x}^{ref}$ can be filtered such that its smoothness characteristic increases, which leads to the decreasing of this error the error $\left\|\mathbf{x} - \mathbf{x}^{ref}\right\|_2^2$.
If we filter $\mathbf{x}^{ref}$ and $\mathbf{x}^{ref(k)}$ in each iteration of Eq. (5.6) we argue that the iterative process including the filtering operation converges to a better solution. The new stop criterion by including the filtered approach is modified as

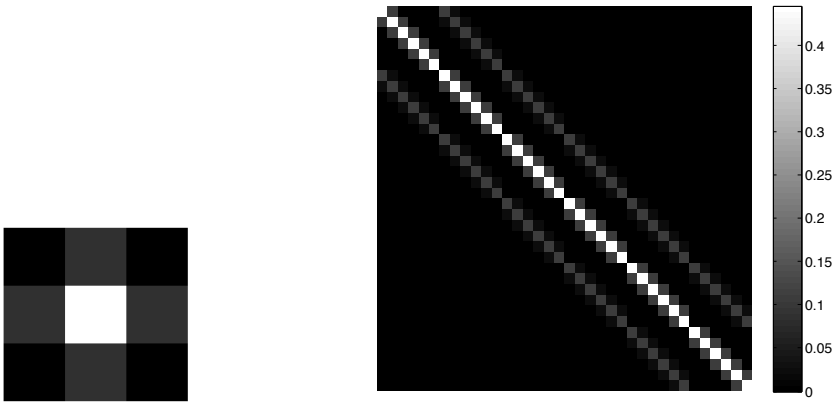
$$\left\|\mathbf{W}\mathbf{x}^{ref} - \mathbf{W}\mathbf{x}^{ref(k)}\right\|_2^2 = \left\|\mathbf{x}_{\mathbf{W}}^{ref} - \mathbf{x}_{\mathbf{W}}^{(k)}\right\|_2^2 < \delta, \tag{5.10}$$

where $\mathbf{x}_{\mathbf{W}}^{ref} = \mathbf{W}\mathbf{x}^{ref}$ and $\mathbf{x}_{\mathbf{W}}^{(k)} = \mathbf{W}\mathbf{x}^{ref(k)}$, and $\mathbf{W} \in \mathbb{R}^{N\times N}$, with elements $w_{i,j}$, is a convolution matrix that represents the filtering operation. Then, the sequence in Eq. (5.6) converges to a filtered version in each iteration. Figure 23 (a) shows a zero mean gaussian kernel filter of size 3×3 and a standard deviation of 0.6, Figure 23 (b) illustrates the respective matrix $\mathbf{W}$ for $N = 36$ created using the Gaussian kernel filter in Figure 23 (a) [27].

5.2. **PROPOSED ALGORITHM**

Algorithm 4 shows the filtered gradient approach for recovering a spectral image from its compressed measurements $\mathbf{y}$. In line 2, we fix the filtered version of the reference as $\mathbf{x}_{\mathbf{W}}^{ref} = \mathbf{W}\mathbf{x}^{ref}$, and iteratively filter the product $\mathbf{x}^{ref(k)} = \mathbf{\Phi}^T\mathbf{\Phi}\mathbf{x}^{(k)}$ in line

Figure 23: (a) A zero mean Gaussian kernel filter of size 3×3 and standard deviation of 0.6, (b) illustrates an example of a $\mathbf{W}$ for $N = 36$ created using the Gaussian kernel fitler in (a).



(a) (b)

7 to get smoothness on each estimation. The update is calculated in line 9, which corresponds to the gradient descent iteration with an additional term $\mathbf{W}$ including the filtering step.

We use to calculate the steps $\alpha^{(k)}$ the Barzilian-Borwein approach [84]. This approach calculates each step such that $\alpha^{(k)}\nabla F(\boldsymbol{\theta}^{(k)})$ approximates the Hessian of F at $\boldsymbol{\theta}^{(k)}$. Barzilai and Borwein propose that the approximation of the Hessian be a multiple of the identity $\eta^{(k)}\mathbf{I}$, where $\eta^{(k)}$ is chosen so that this approximation has similar behavior to the true Hessian over the most recent step, that is

$$\nabla F(\boldsymbol{\theta}^{(k)}) - \nabla F(\boldsymbol{\theta}^{(k-1)}) \approx \eta^{(k)}[\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)}] \tag{5.11}$$

with $\eta^{(k)}$ chosen to satisfy this in the least-squares sense. In order to use this approach, we follow the extension to BCQPs in [85] that is also descripted in [25] as follows:

Step 0: (initialization): Given $\theta^{(0)}$, choose parameters α_{min}, α_{max}, with $0 < \alpha_{min} < \alpha_{max}$, $\alpha^{(0)} \in [\alpha_{min}, \alpha_{max}]$, and set $k = 0$.

Step 1 Compute step:

$$\boldsymbol{\delta}^{(k)} = \left(\boldsymbol{\theta}^{(k)} - \alpha^{(k)}\nabla F(\boldsymbol{\theta}^{(k)})\right)_+ - \boldsymbol{\theta}^{(k)}. \tag{5.12}$$

Step 2 (line search): Find the scalar $\lambda^{(k)}$ that minimizes $F(\boldsymbol{\theta}^{(k)} + \lambda^{(k)}\boldsymbol{\delta}^{(k)})$ on the interval $\lambda^{(k)} \in [0, 1]$, and set $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \lambda^{(k)}\boldsymbol{\delta}^{(k)}$. Since F is quadratic, the line

search parameter $\lambda^{(k)}$ can be calculated using the following closed-form expression:

$$\lambda^{(k)} = \text{mid}\left\{0, \frac{\left(\boldsymbol{\delta}^{(k)}\right)^T \nabla F\left(\boldsymbol{\theta}^{(k)}\right)}{\left(\boldsymbol{\delta}^{(k)}\right)^T \mathbf{B}\boldsymbol{\delta}^{(k)}}, 1\right\} \tag{5.13}$$

where the operator mid$\{a, b, c\}$ is the middle value of the three scalar arguments.
Step 3 (update) α **):** Compute

$$\gamma^{(k)} = \left(\boldsymbol{\delta}^{(k)}\right)^T \mathbf{B}\boldsymbol{\delta}^{(k)}; \tag{5.14}$$

if $\gamma^{(k)} = 0$, let $\alpha^{(k+1)} = \alpha_{max}$, otherwise

$$\lambda^{(k)} = \text{mid}\left\{\alpha_{min}, \frac{\left\|\boldsymbol{\delta}^{(k)}\right\|_2^2}{\gamma^{(k)}}, \alpha_{max}\right\}. \tag{5.15}$$

Step 4: Perform convergence test and terminate with approximate solution $\boldsymbol{\theta}^{(k+1)}$ if it is satisfied; otherwise set $k \leftarrow k + 1$ and return to **Step 1**.
Table 1 summarizes the variables in the algorithm. It is important to clarify that the filtered gradient proposed approach here can be easily incorporated in many elaborated gradient based algorithms such as the Stochastic Gradient Approach Based on Adaptive Filtering Framework [41], the Non-monotone adaptive Barzilai-Borwein gradient algorithm [42], the Iterative gradient projection [43], the Fast and accurate first-order method for sparse recovery [44], the Sparse reconstruction by separable approximation [4], among others. Particularly, in this work, we implement the filtered version of the Gradient projection for sparse reconstruction (GPSR) algorithm in the CSI problem.

5.3. **SIMULATIONS AND RESULTS**

In order to verify the filtered gradient approach, a set of compressive measurements is simulated using the forward models in (1.5) and (1.4). For reconstruction, the basis representation $\boldsymbol{\Psi}$ is set to be the Kronecker product of three basis $\boldsymbol{\Psi} = \boldsymbol{\Psi}_1 \otimes \boldsymbol{\Psi}_2 \otimes \boldsymbol{\Psi}_3$, where the combination $\boldsymbol{\Psi}_1 \otimes \boldsymbol{\Psi}_2$ is the 2D-Wavelet Symmlet 8 basis and $\boldsymbol{\Psi}_3$ is the discrete cosine basis. Simulation results are the average of 10 experiments conducted for each case. The tests were carried out on a personal computer with an Intel(R) Core(TM) i7-4790 CPU 3.60 GHz processor and 32 GB of memory, using MATLAB. To test the filtered gradient approach we take a gradient based reconstruction algorithm and modify it to include the filtered gradient descent method presented in Algorithm 1. We have noted that the benefits of the filtered approach are preserved for the different gradient based approaches.

Algorithm 5: Filtered gradient descent method

input : $\mathbf{y}, \mathbf{\Phi}, \mathbf{\Psi}, \mathbf{W}, \tau, \delta, \boldsymbol{\theta}^{(0)}$
output: $\boldsymbol{\theta}^{(k+1)}$

$\mathbf{x}^{ref} \leftarrow \mathbf{\Phi}^T\mathbf{y}$;
$\%$ (`Precompute the filtered` $\mathbf{x}^{ref}$) $\mathbf{x}_{\mathbf{W}}^{ref} \leftarrow \mathbf{W}\mathbf{x}^{ref}$;
$k \leftarrow 0$;
$\mathbf{x}_{\mathbf{W}}^{0} \leftarrow 0$;
while $\left\|\mathbf{x}_{\mathbf{W}}^{ref} - \mathbf{x}_{\mathbf{W}}^{(k)}\right\| > \delta$ **do**
  $\mathbf{x}^{(k)} \leftarrow \mathbf{\Psi}\boldsymbol{\theta}^{(k)}$;
  $\%$ (`Filtering` $\mathbf{x}^{(k)}$) $\mathbf{x}_{\mathbf{W}}^{(k)} \leftarrow \mathbf{W}\mathbf{\Phi}^T\mathbf{\Psi}\mathbf{x}^{(k)}$;
  $\%$ (`Line Search`) Choose a step size $\alpha_k > 0$;
  $\%$(`Update`) $\boldsymbol{\theta}^{(k+1)} \leftarrow \boldsymbol{\theta}^{(k)} - \alpha^{(k)}(\tau\mathbf{1}_n - \mathbf{\Psi}^T(\mathbf{x}_{\mathbf{W}}^{ref} - \mathbf{x}_{\mathbf{W}}^{(k)}))$;
  $k \leftarrow k + 1$;

5.3.1. **PSNR as a function of the kernel filter parameters** The filtered approach is evaluated by the reconstruction of the Feathers and Flowers spectral images captured with a CCD camera Apogee Alta U260 and a VariSpec liquid crystal tunable filter, in the wavelength range 400nm-560nm, with steps of 10nm [86]. For simulations, a spatial section of 512×512 pixels, and 8 spectral bands were used. The matrix $\mathbf{W}$ was selected to be a convolution matrix of symmetric Gaussian lowpass filter. In each experiment, the value of τ for the GPSR and the Filtered approaches was selected by cross-validation where the interval [1e-4,1e-5] provides the best results for the estimated image. For the first experiment, the spectral image Flowers was used. The first experiment consists of comparing the average PSNR of the reconstructed image with three different levels of standard deviation of the filter kernel ($\sigma \in \{0.2, 0.6,$ and $1.2\}$) and three filter kernel sizes (3×3, 5×5, and 7×7) and the GPSR as a function of the transmittance of the Dual-coded approach. Table 12 shows the results for a transmittance of 20%, Table 13 for a transmittance of 50%, and Table 14 for a transmittance of 60%. These results show that the PSNR values obtained with the filtered approach are higher than those achieved with the traditional version of the GPSR algorithm for a standard deviation of 0.6, and overcomes up to 4 dB the GSPR algorithm, however, when the standard deviation is 0.2 the performance of the filtered approach is worse than the GPSR algorithm for the three transmittance values, and when the standard deviation is 1.2 it is worse for 20% and 60% of transmittance. On the other hand, note that from a filter size of 3×3 the PSNR improves significantly, nevertheless, larger sizes of filters have similar performance. Then, the standard deviation is a decisive parameter of the filter and can be tuned for getting better results than those of the GPSR algorithm.
For the next experiment, five levels of compression ratio $\left(\frac{M}{N} \cdot 100\%\right)$ were selected (10%, 20%, 50%, 60%, and 90%) for the CASSI system. This experiment consists of comparing the average PSNR of the reconstructed images with five different levels

Table 12: Mean PSNR of the Flowers reconstructed image with three different levels of standard deviation and filter sizes, and the traditional version of the GPSR for a transmittance of 20% in the Dual-coded system.

GPSR		34.86		
		σ		
		0.2	0.6	1.2
3*Filter size	3×3	34.68	**38.67**	**35.16**
	5×5	34.42	**38.80**	**36.44**
	7×7	34.76	**38.78**	**36.35**

Table 13: Mean PSNR of the Flowers reconstructed image with three different levels of standard deviation and filter sizes, and the traditional version of the GPSR for a transmittance of 50% in the in the Dual-coded system.

GPSR		39.94		
		σ		
		0.2	0.6	1.2
3*Filter size	3×3	39.81	**40.66**	35.22
	5×5	39.89	**40.65**	37.37
	7×7	39.90	**40.66**	37.48

of standard deviation of the filter kernel ($\sigma \in$0.2, 0.6, 0.9, 1.2, and 2}) and the GPSR as a function of the compression ratio, for a transmittance of the coded apertures of 40% in the CASSI system. Figure 24 shows that the PSNR values obtained with the filtered approach are higher than those achieved with the traditional version of the GPSR algorithm for a standard deviation between 0.6 and 0.9, and overcomes up to 6 dB the GSPR algorithm, however when the standard deviation is larger than 1.2 the performance of the filtered approach is worse than the GPSR algorithm. Then, also for the CASSI architecture the standard deviation is a decisive parameter of the filter.

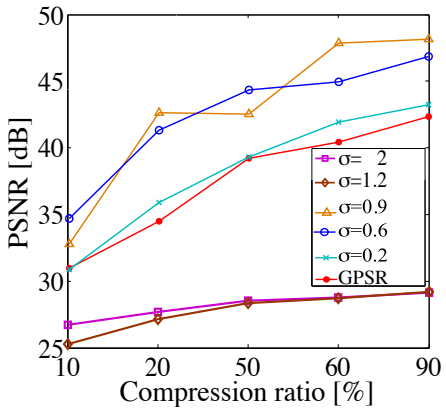
Table 15 shows the results with a $\sigma = 0.6$ and different filter size and compression ratio. Note that from a filter size of 3×3 the PSNR improves significantly, however larger sizes of filters have similar performance. Therefore, the filter size is not as crucial as the selected standard deviation. On the other hand, a larger filter size to 3×3 increase the computing time in each iteration.

Another important issue related to the filtered algorithm is the convergence. Despite there is no way to directly compute the convergence of new algorithm, in practice the following experiments show the improvement of the algorithm convergence when the filtering step is used. For the following experiments a filter size of 3×3 and $\sigma = 0.6$ was selected for the CASSI system. Figure 25 shows the evolution of the PSNR in

Table 14: Mean PSNR of the Flowers reconstructed image with three different levels of standard deviation and filter sizes, and the traditional version of the GPSR for a transmittance of 60% in the in the Dual-coded system.

GPSR		39.86		
		σ		
		0.2	0.6	1.2
3*Filter size	3×3	39.94	**40.33**	34.52
	5×5	39.90	**40.31**	37.51
	7×7	39.85	**40.31**	37.55

Figure 24: Mean PSNR of the Flowers reconstructed image with five different levels of standard deviation ($\sigma \in${0.2, 0.6, 0.9, 1.2, and 2}) and the traditional version of the GPSR as a function of the compression ratio, for a transmittance of the coded apertures of 40% in the CASSI system.

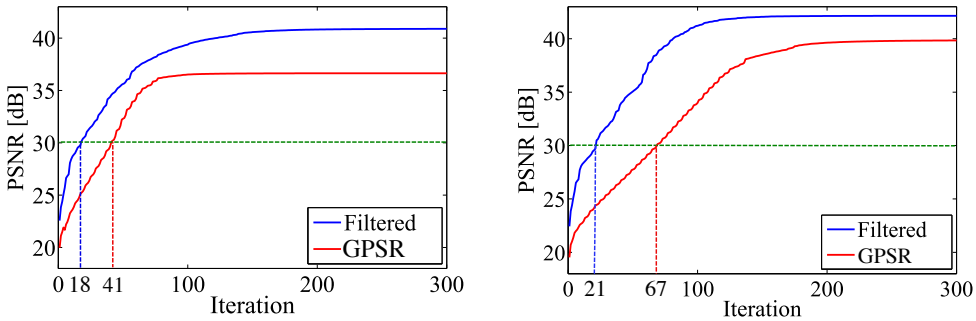


each iteration for the modified filtered approach (blue) and the traditional approach (red), for the specific case of a transmittance of 40%, and a compression ratio of 50%. Ideally, a transmittance of 100% is desired, however, high transmittance decreases the coding randomness. Figure 25 (a) corresponds to the results for the Feathers database, and the Fig. 25(b) for the Flowers database. Figure 25 shows that if a threshold of 30 [dB] in PSNR is imposed for a good quality reconstruction (dashed green line) the proposed algorithm reaches the threshold in fewer iterations than the traditional approach since filtered step allows to reduce the error of each estimation of the traditional gradient method. Specifically, in the case of the Feathers database Fig. 25(a) shows that the filtered takes 18 iterations to reach the 30 [dB] level while the unfiltered takes 41. Fig. 25(b) shows that for the Flowers database the filtered takes 21 iterations to reach the threshold and the unfiltered takes 67. It can be noted that the number of iterations to get an acceptable estimation of the spectral image is

Table 15: PSNR [dB] as function of filter size

Compression ratio [%]	10	20	40	50	60	80	90
GPSR	28.91	31.09	32.08	34.35	35.09	35.09	38.38
2x2	31.74	34.74	36.46	37.93	39.21	39.93	40.57
3x3	33.82	36.09	38.04	39.21	40.68	42.19	43.69
4x4	32.55	35.25	36.75	38.24	39.95	41.36	42.16
5x5	33.66	36.05	37.17	37.77	39.39	41.74	42.61

Figure 25: Evolution of the PSNR in each iteration for the filtered approach (blue) and the unfiltered approach (red), for the specific case of a transmittance of 40%, a compression ratio of 50%, and $\sigma = 0.6$, for (a) Feathers and (b) Flowers databases.

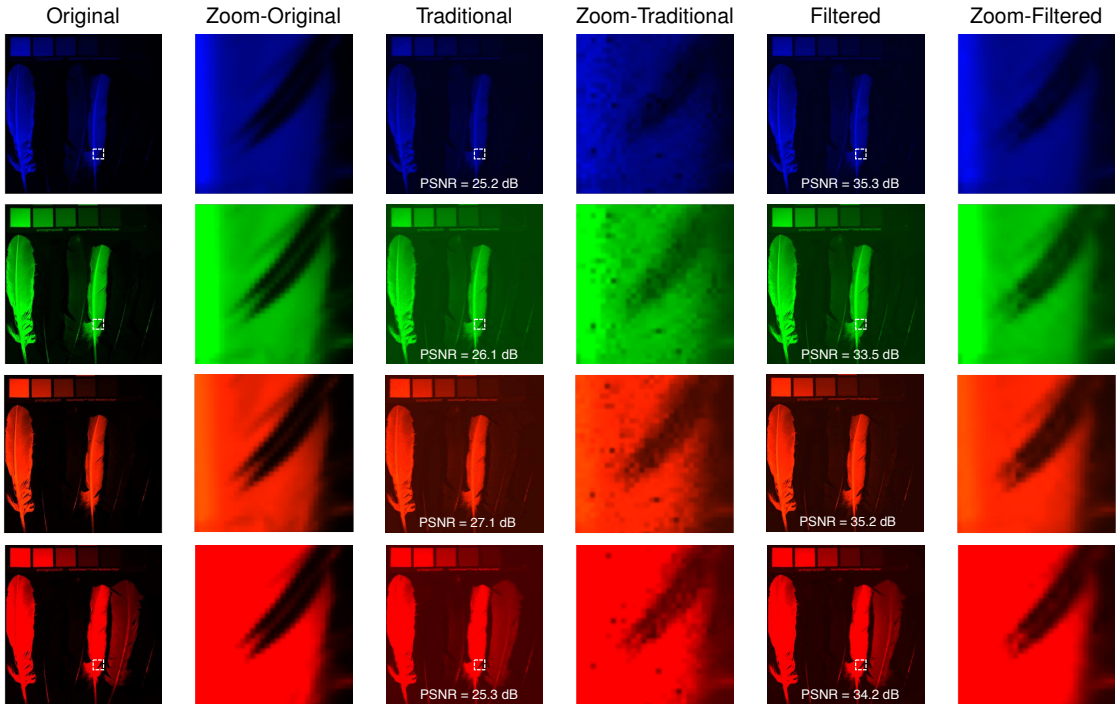


(a) (b)

obtained in fewer iterations than without the filtering step.
To better visualize the differences, the reconstructed second, fourth, sixth, and eighth spectral bands and their zoom versions are presented in Fig. 26 for the feather database and in Fig. 27 for the flowers database, using a compression ratio of 20% and transmittance of 40%. There, it is possible to note the better performance of the filtered approach.

5.3.2. **Comparison with the Approximate Message Passing (AMP-3D-Wiener) and the Split Augmented Lagrangian Shrinkage (SALSA) algorithms** For the purpose of comparison, we select the approximate message passing (AMP-3D-Wiener) algorithm of Tan, et al. [45], which is a sparse recovery approach based on iterative filtering; and the Split Augmented Lagrangian Shrinkage algorithm of Afonso et al. [47] that is an instance of the so-called alternating direction method of multipliers (ADMM), both present an ideal simulation scenario to test the filtered proposed method. The experiments are performed using the same settings 1 and 2 of the work of Tan et al. [46] using the scene shown in Figure 28 (a). This data

Figure 26: The 2^{nd}, 4^{th}, 6^{th}, and 8^{th} reconstructed spectral bands are presented in each row, respectively. The first and second columns show the original and their zoom versions. The third and fourth columns are the reconstructions from 20% compression and transmittance of 40% using unfiltered gradient algorithm and the fifth and sixth columns are the reconstructions using the filtered gradient algorithm.



cube was acquired using a wide-band Xenon lamp as the illumination source, modulated by a visible monochromator spanning the spectral range between $448nm$ and $664nm$, and each spectral band has $9nm$ width. The image intensity was captured using a grayscale CCD camera, with pixel size $9,9\mu m$, and 8 bits of intensity levels. The resulting test data cube has 256×256 pixels of spatial resolution and 24 spectral bands.

In the setting 1, the CASSI measurements are captured with 2 shots. In the case of AMP-3D-Wiener algorithm the coded apertures required to be complementary, in this particular example the first shot is generated randomly with 50% of the aperture being opaque, and the coded aperture in the second shot is the complement of the aperture in the first shot. The compression ratio $\left(\frac{M}{N}\right)$ with two shots is approximately of 10%. Moreover, we add Gaussian noise with zero mean to the measurements such that the SNR is 20 dB. Figure 28 (b) shows the PSNR of each 2D slice in the reconstructed cube separately for setting 1 using the proposed filtered approach, the SALSA algorithm, the AMP-3D-Wiener approach, and the traditional version of the GPSR. The results of the AMP-3D-Wiener recovery were taken of the work of

Figure 27: The 2^{nd}, 4^{th}, 6^{th}, and 8^{th} reconstructed spectral bands are presented in each row, respectively. The first and second columns show the original and their zoom versions. The third and fourth columns are the reconstructions from 20% compression rate and transmittance of 40% using traditional gradient algorithm, and the fifth and sixth columns are the reconstructions with same parameters using the filtered gradient algorithm.

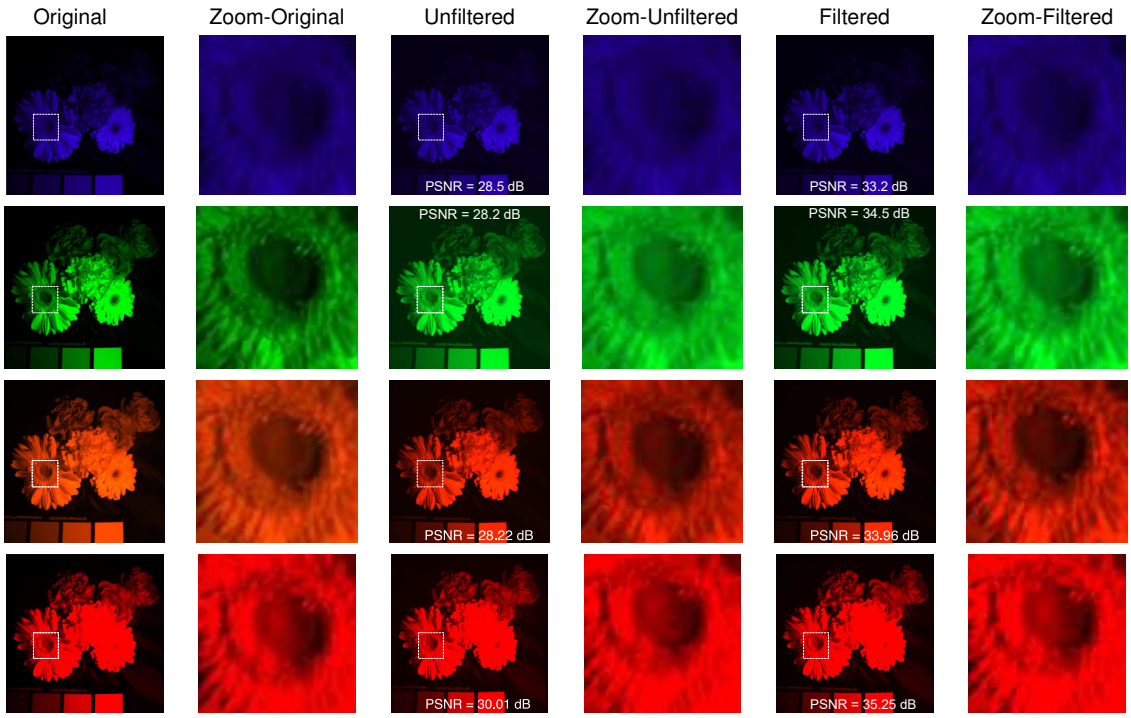
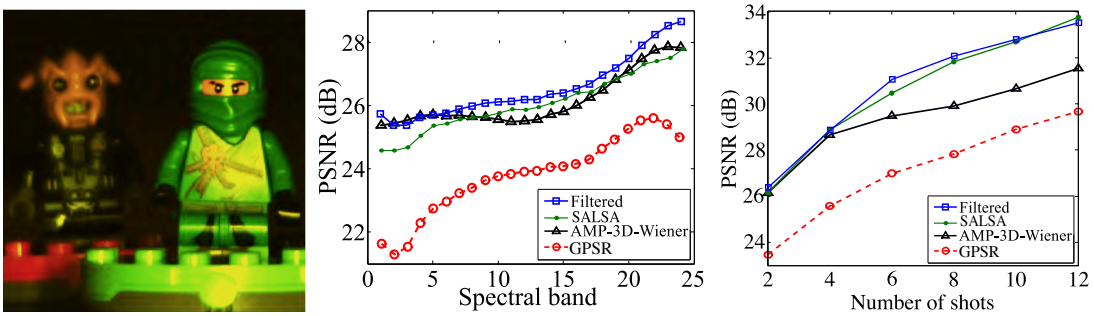


Figure 28: (a) Lego spectral data cube is shown as it would be viewed by a Stingray F-033C CCD Color Camera (b) PSNR of the reconstructed Spectral bands, comparison of the filtered approach, SALSA, AMP-3D-Wiener, and traditional GPSR, for the Lego image cube. Cube size is 256$\times$ 256 and 24 spectral bands. The measurements are captured with 2 shots. (c) PSNR comparison as a function of the number of shots. In both cases random Gaussian noise is added to the measurements such that the SNR is 20 dB.



(a) (b) (c)

Tan et al.[46], and a version of the SALSA algorithm was implemented for reconstructing the data cube of simulated measurements of setting 1 using random coded apertures. It is shown that the cube reconstructed by the Filtered approach has a comparable or higher PSNR than the cubes reconstructed by AMP-3D-Wiener and SALSA algorithms, and 2-5 dB higher than those of traditional GPSR for all 24 slices. It is important to note that the AMP-3D-Wiener is an approach that was implemented specifically for the CASSI system, requiring complementary coded apertures and greater computing capacity at each iteration, unlike the approach proposed in this article that is not specific to the application or demands high computing capacity.
In the second experiment we set the number of shots in $K = \{2, 4, 6, 8, 10, 12\}$, and also we add Gaussian noise with zero mean to the measurements such that the SNR is 20 dB. Figure 28 (b) shows a comparison of the reconstructed cubes PSNR as a function of the number of shots. Note that the proposed filtered approach has a similar performance than the SALSA algorithm, has an improvement up to 4dB over the traditional GPSR, and after 4 shots beats the AMP-3D-Wiener up to 3dB.

5.4. CHAPTER CONCLUSION

A gradient algorithm that includes a filtering step for solving the compressive sensing problem has been proposed. The methodology to include this filtering step in a gradient descent algorithm is exposed. Simulations show an improvement of up to

7 dB of PSNR for reconstructions obtained from the modified filtered GPSR algorithm compared to its traditional version, in the compressive spectral imaging CASSI system. The experiments show the better visual quality of the reconstructed images when the filtered algorithm was used. This general approach was compared with the state-of-the-art Approximate Message Passing (AMP-3D-Wiener), which is a specific implementation for the CASSI system, and Split Augmented Lagrangian Shrinkage (SALSA) algorithms and the results show better performance.

Additionally, experiments show that the filtered approach accelerates the convergence of the algo-rithm, this could be because the smoothness process accelerates the estimation of each iteration tobe closer to a tolerable one. On the other hand, in order to improve the conditions to convergence,it is possible to adapt the effect of the filter in each iteration such that it is stronger for the firstiterations and decreases along them.

6. BAYESIAN CSI RECONSTRUCTION BY USING A LOCAL STRUCTURED PRIOR

This chapter introduces a hierarchical Bayesian model for the reconstruction of hyperspectral images using compressed sensing measurements. This model exploits known properties of natural images, promoting the recovered image to be sparse on a selected basis and smooth in the image domain. The posterior distribution of this model is too complex to derive closed form expressions for the estimators of its parameters. Therefore, an MCMC method is investigated to sample this posterior distribution. The resulting samples are used to estimate the unknown model parameters and hyperparameters in an unsupervised framework. The results obtained on real data illustrate the improvement in reconstruction quality when compared to some existing techniques.

When using a significant undersampling rate, there is an infinite number of possible images that can be associated with a particular set of measurements. Thus, it is usual to promote realistic properties of the image to recover in order to regularize the problem and obtain a unique solution. One of the most common regularizations consists of enforcing the image of interest to have a sparse representation in a given basis [2]–[4], [26]. Although this technique has shown satisfactory results, the quality of the reconstruction can be further improved by exploiting additional properties of natural images such as their important degree of local structure similarity [27], [28].

In this work, we introduce a new hierarchical Bayesian model for the reconstruction of compressed hyperspectral images. This model promotes the solution to be sparse in a selected basis and smooth in the image domain. A Gibbs sampler is used to draw samples asymptotically distributed according to its posterior distribution. The generated samples are then used to build estimators of the unknown model parameters.

The proposed model provides improved reconstruction quality due to the exploitation of the local structure similarity of natural images. It also allows the model hyperparameters to be estimated from the observed measurements in an unsupervised framework (instead of fixing them a-priori as required by most convex optimization methods [42]–[44]). Finally, the samples generated by the Gibbs sampler can be used to determine measures of uncertainty for the estimates such as the estimation variance of confidence intervals.

6.1. PROPOSED BAYESIAN RECONSTRUCTION

The CSI inverse problem addressed in this paper can be formulated as follows

$$\mathbf{y} = \mathbf{\Phi x} + \boldsymbol{e} \tag{6.1}$$

where $\mathbf{x} \in \mathbb{R}^{Z^2 L}$ is the lexicographically ordered vectorization of an image of size $Z \times Z$ with L spectral bands, $\mathbf{\Phi} \in \mathbb{R}^{P \times Z^2 L}$ is the sparse measurement matrix that depends on the compressive sensing imager (see [9]–[11], [33] for more details), $\boldsymbol{e} \in \mathbb{R}^P$ is an additive white Gaussian noise and $\mathbf{y} \in \mathbb{R}^P$ is the observed measurement vector. In this paper, we concentrate on the CASSI sampler [8], [18], [87]. We propose to recover $\mathbf{x}$ from $\mathbf{y}$ by solving the following problem

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left(\frac{1}{2\sigma_n^2} ||\mathbf{A}\boldsymbol{\theta} - \mathbf{y}||^2 + \tau ||\boldsymbol{\theta}||_1 + \lambda ||(\mathbf{B} - \mathbf{I})\mathbf{\Psi}\boldsymbol{\theta}||^2 \right) \tag{6.2}$$

where $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}$, $\boldsymbol{\theta} \in \mathbb{R}^{Z^2 L}$ contains the coefficients of $\mathbf{x}$ in a given basis $\mathbf{\Psi} \in \mathbb{R}^{Z^2 L \times Z^2 L}$ (i.e., $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$) and $\mathbf{B} \in \mathbb{R}^{Z^2 L \times Z^2 L}$ is a weighting matrix (as the one considered in [28], [83]) associated with a low-pass filter. Note that the first term in (6.2) is the data-fidelity term, that the second term is a penalty enforcing sparsity of $\boldsymbol{\theta}$, and the third term promotes image smoothness by encouraging each pixel of $\mathbf{x}$ to be close to a linear combination of its neighbors. The model hyperparameters τ and λ adjust the relative importance of the three terms. In the following section, we show that the inverse problem (6.2) is equivalent to the determination of the maximum a posteriori (MAP) estimator of a Bayesian model whose likelihood and priors are provided.

6.1.1. **Likelihood** Assuming that the additive noise in (6.1) is white Gaussian with variance σ_n^2, the distribution of the observation vector is a Gaussian distribution with mean vector $\mathbf{A}\boldsymbol{\theta}$ and covariance matrix $\sigma_n^2 \mathbf{I}_P$, i.e.,

$$f(\mathbf{y}|\boldsymbol{\theta}, \sigma_n^2) = \mathcal{N}\left(\mathbf{y} \middle| \mathbf{A}\boldsymbol{\theta}, \sigma_n^2 \mathbf{I}_P\right) \tag{6.3}$$

where $\mathbf{I}_P$ is the $P \times P$ identity matrix and $\mathcal{N}$ denotes the Gaussian distribution. This likelihood is the Bayesian equivalent of the data-fidelity term previously shown in (6.2).

6.1.2. **Prior distributions**

Image coefficients θ

To promote sparsity in the $\mathbf{\Psi}$ domain and smoothness in the image domain, the following prior distribution is introduced for the image basis coefficients $\boldsymbol{\theta}$

$$f(\boldsymbol{\theta}|\lambda, \tau) \propto \mathcal{N}\left(\boldsymbol{\theta} \middle| 0, \boldsymbol{C}/\lambda\right) \prod_{i=1}^{Z^2 L} \exp\left(-\tau|\theta_i|\right) \tag{6.4}$$

with $\boldsymbol{C}^{-1} = \mathbf{\Psi}^T (\mathbf{B} - \mathbf{I})^T (\mathbf{B} - \mathbf{I}) \mathbf{\Psi}$, and where λ and τ are two hyperparameters. This prior can be shown to be the Bayesian equivalent of the regularization terms in (6.2). To simplify the analysis, it is convenient to define the hyperparameter $a = \sigma_n^2 \tau^2$ which transforms the prior to

$$f(\boldsymbol{\theta}|\lambda, a, \sigma_n^2) \propto \mathcal{N}\left(\boldsymbol{\theta}\middle|0, \boldsymbol{C}/\lambda\right) \prod_{i=1}^{Z^2L} \exp\left(-\sqrt{\frac{a}{\sigma_n^2}}|\theta_i|\right). \tag{6.5}$$

It is then easy to show that (6.5) is the marginal distribution of the data-augmented density $f(\boldsymbol{\theta}, \boldsymbol{\delta}^2|\lambda, \tau)$, where $\boldsymbol{\delta}^2 \in (\mathbb{R}^+)^{Z^2L}$ is a vector of independent latent variables distributed according to gamma distributions such that (see [88] for details)

$$f(\boldsymbol{\delta}^2|a) = \prod_{i=1}^{Z^2L} \mathcal{G}\left(\delta_i^2\middle|1, \frac{a}{2}\right). \tag{6.6}$$

The advantage of using this data augmentation scheme is that the conditional distribution of $\boldsymbol{\theta}|\lambda, \boldsymbol{\delta}^2, \sigma_n^2$ is much easier to sample than the conditional distribution of $\boldsymbol{\theta}|\lambda, a, \sigma_n^2$ since

$$f(\boldsymbol{\theta}|\lambda, \boldsymbol{\delta}^2, \sigma_n^2) \propto \mathcal{N}\left(\boldsymbol{\theta}\middle|0, \boldsymbol{C}/\lambda\right) \mathcal{N}\left(\boldsymbol{\theta}\middle|0, \sigma_n^2\boldsymbol{\Delta}\right) \tag{6.7}$$

where $\boldsymbol{\Delta} \in \mathbb{R}^{Z^2L \times Z^2L}$ is a diagonal matrix whose ith diagonal element is δ_i^2.

Noise variance σ_n^2

The noise variance is assigned a Jeffrey's prior

$$f(\sigma_n^2) \propto \frac{1}{\sigma_n^2} 1_{\mathbb{R}^+}(\sigma_n^2) \tag{6.8}$$

where $1_{\mathbb{R}^+}(\xi) = 1$ if $\xi \in \mathbb{R}^+$ and 0 otherwise (indicator function on $\mathbb{R}^+$). This choice is very classical when no information about a scale parameter is available (see [89]).

6.1.3. **Hyperparameter priors** To be able to estimate the hyperparameters a and λ of a Bayesian model, one can define a hierarchical Bayesian model defined by hyperparameter priors. The hyperpriors considered in this work are summarized in this section.

Hyperprior of a

A Jeffrey's prior is assigned to a to keep the system as non-informative as possible

$$f(a) \propto \frac{1}{a} 1_{\mathbb{R}^+}(a). \tag{6.9}$$

Hyperprior of λ

A conjugate gamma hyperprior is assigned to λ

$$f(\lambda) = \mathcal{G}\left(\lambda\middle|\alpha_\lambda, \beta_\lambda\right). \tag{6.10}$$

The values of α_λ and β_λ are chosen to make the hyperprior non-informative. The values we used for our experiments are specified in Section 6.3.

6.1.4. **Posterior distribution** Using the likelihood as well as the parameter and hyperparameter priors defined in the previous sections, the posterior distribution associated with the proposed hierarchical Bayesian model is

$$f(\sigma_n^2, \boldsymbol{\theta}, a, \lambda, \boldsymbol{\delta}^2|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\theta}, \sigma_n^2) f(\boldsymbol{\theta}|\lambda, \boldsymbol{\delta}^2, \sigma_n^2) \\ f(\boldsymbol{\delta}^2|a) f(\sigma_n^2) f(a) f(\lambda). \tag{6.11}$$

6.2. GIBBS SAMPLER

The posterior distribution (6.11) is intractable, in the sense that it does not allow closed-form expressions for the Bayesian estimators of the parameters and hyperparameters to be derived. Thus we propose to draw samples distributed according to (6.11) using a Markov chain Monte Carlo (MCMC) method. The generated samples are then used to estimate the unknown image jointly with the other model parameters and hyperparameters. More precisely, after an appropriate burn-in period, the samples associated with a given parameter generated by the MCMC method are averaged in order to approximate the minimum mean square error estimator of this parameter. The MCMC method considered in this paper is a Gibbs sampler, which generates the unknown variables sequentially according to their conditional distributions, as shown in Algorithm 6 (see [89]). Note that the full vector $\boldsymbol{\theta}$ can be sampled in a single step since it is more efficient than sampling each element of $\boldsymbol{\theta}$ separately (it allows convergence to be reached with a smaller number of iterations). The corresponding conditional distributions are shown in Table 16 where $\mathcal{GIG}$, $\mathcal{N}$, $\mathcal{G}$, and $\mathcal{IG}$ are the generalized inverse Gaussian, normal, gamma, and inverse gamma distributions (see [88] for details about the $\mathcal{GIG}$ distribution).

Algorithm 6: Gibbs sampler

```
Initialize(a, σ_n^2, λ)
Sample(θ from its prior distribution)
while No convergence do
    for i ← 1 to N do
        Sample(δ_i^2 from f(δ_i^2|θ_i, σ_n^2, a));
    Sample(λ from f(λ|θ));
    Sample(a from f(a|δ^2));
    Sample(σ_n^2 from f(σ_n^2|y, θ, δ^2));
```

6.2.1. **Sampling considerations** The variables δ_i^2, λ, a and σ_n^2 can be easily sampled using standard generators of random variables. However, sampling all the elements of $\boldsymbol{\theta}$ jointly in a direct manner is not straightforward, since the inversion of the precision matrix $\boldsymbol{\Sigma}^{-1} \in \mathbb{R}^{Z^2L \times Z^2L}$ becomes intractable even for small hyperspectral images. In order to solve this problem, we propose to use the sampling technique introduced by Orieux *et. al.* [90] that was proposed to sample from high dimensional

Table 16: Full conditionals $f(\delta_i^2|\theta_i,\sigma_n^2,a)$, $f(\boldsymbol{\theta}|\mathbf{y},\sigma_n^2,\boldsymbol{\delta}^2,\lambda)$, $f(\lambda|\boldsymbol{\theta})$, $f(a|\boldsymbol{\delta}^2)$ and $f(\sigma_n^2|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\delta^2})$ associated with (6.11).

δ_i^2	$\mathcal{GIG}\left(\frac{1}{2}, a, \frac{\theta_i^2}{\sigma_n^2}\right)$
$\boldsymbol{\theta}$	$\mathcal{N}\left(\frac{\boldsymbol{\Sigma}\mathbf{A}^T\mathbf{y}}{\sigma_n^2}, \boldsymbol{\Sigma}\right), \boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_n^2}(\mathbf{A}^T\mathbf{A} + \boldsymbol{\Delta}^{-1}) + \lambda\boldsymbol{C}^{-1}$
λ	$\mathcal{G}\left(\frac{Z^2L}{2} + \alpha_\lambda, \frac{\|(\mathbf{B}-\mathbf{I})\boldsymbol{\Psi}\boldsymbol{\theta}\|^2}{2} + \beta_\lambda\right)$
a	$\mathcal{G}\left(Z^2L, \frac{\|\boldsymbol{\delta}\|^2}{2}\right)$
σ_n^2	$\mathcal{IG}\left(\frac{Z^2L+P}{2}, \frac{1}{2}\left[\|\mathbf{y}-\mathbf{A}\boldsymbol{\theta}\|^2 + \sum\frac{\theta_i^2}{\delta_i^2}\right]\right)$

multivariate Gaussian distributions. Note that the method bypasses the problem of inverting the precision matrix $\boldsymbol{\Sigma}^{-1}$ by using a perturbation-optimization algorithm. A conjugate gradient method has been used in this paper to solve the perturbation-optimization problem.

6.3. **SIMULATION RESULTS**

In order to evaluate the performance of the proposed algorithm, experiments were performed on a real hyperspectral image. The image was acquired using a monochromator with wavelengths separated by 1nm, more precisely with a CCD camera AVT Marlin F033B, with 656×492 pixels, 24 spectral bands and a pixel pitch of 9.9μm. The experiments were conducted on a section of 128×128 pixels and 8 spectral bands (from 461nm to 596nm) that was extracted from the acquired image. Five matrices $\boldsymbol{\Phi}$ were calculated, each corresponding to a random realization of a CASSI sensing matrix with a different compression ratio $\frac{P}{Z^2L}$, more precisely $13\%, 26\%, 40\%, 53\%$, and 66%. Compressed images were then generated by multiplying the hypercube by each of the matrices $\boldsymbol{\Phi}$ and adding Gaussian noise to obtain SNR = 25dB.
Four different algorithms were used to recover the hypercube from each set of measurements: i) The SpaRSA optimization algorithm for the LASSO problem (SpaRSA LASSO) [4], ii) the SpaRSA algorithm for solving the proposed problem (6.2) (SpaRSA smooth), iii) the Bayesian LASSO [91] and iv) the proposed method. Note that algorithms ii and iv estimate the image by solving (6.2) whereas algorithms i and iii do not use the smoothing term. Note also that algorithms i and ii require to set a priori values of σ_n^2, τ and λ, e.g., by using cross-validation, whereas algorithms iii and iv are implemented in a Bayesian framework, that estimates is own hyperparameters from the observed data.
The basis representation $\boldsymbol{\Psi}$ used in all experiments is defined as the Kronecker product of three bases $\boldsymbol{\Psi} = \boldsymbol{\Psi}_1 \otimes \boldsymbol{\Psi}_2 \otimes \boldsymbol{\Psi}_3$, where the combination $\boldsymbol{\Psi}_1 \otimes \boldsymbol{\Psi}_2$ is the 2D-Wavelet Symmlet 8 basis and $\boldsymbol{\Psi}_3$ is the cosine basis. For algorithms ii and iv, $\mathbf{B}$ was chosen as a low-pass Gaussian filter with radius 3 and standard deviation 0.6. Finally, the hyperprior parameters were fixed by cross validation to $\alpha_\lambda = 1 \times 10^{-3}$ and $\beta_\lambda = 10 \times 10^{-6}$.

Table 17: PSNRs [dB] obtained by the different algorithms.

Compression ratio	13%	26%	40%	53%	66%
Proposed method	24.4	27.1	28.6	29.6	30.4
Bayesian LASSO	22.9	26.0	27.5	28.4	28.4
SpaRSA smooth	25.2	27.1	28.8	29.7	30.6
SpaRSA LASSO	23.5	26.8	28.5	29.4	30.4

Figure 29: Seventh spectral band of the image: (Left) Ground truth. Reconstruction results for: (top center) the proposed method, (bottom center) SpaRSA Smooth, (top right) Bayesian LASSO and (bottom right) SpaRSA LASSO.

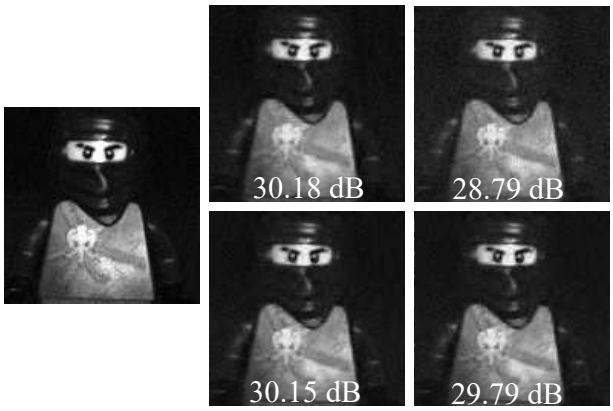


Table 17 shows the PSNRs of the reconstructed images obtained for different compression ratios of Φ and for the four algorithms. We can observe that the algorithms using the smoothness term have a performance that is up to 2dB higher than the ones that do not. Note that the optimization algorithms yield slightly better results than their Bayesian counterparts. We believe that this improved performance is due to the fact that the hyperparameters in the optimization algorithms have been manually set to optimize the PSNR whereas these hyperparameters are estimated from the data by the Bayesian algorithms. Fig. 29 illustrates the seventh spectral band recovered by the algorithms when applied to the measurements corresponding to a 53% compression ratio. The algorithms that include the smoothness term clearly provide results that are visually less noisy.

Examples of reconstructions along the spectral axis for spatial pixel $\#(20, 33)$ (selected randomly) are compared in Fig. 30. The methods that use the smoothing term are closer to the ground truth. In addition, the Bayesian methods are able to calculate the standard deviation (SD) of their estimations, which is included in the graph showing that the reference is within their 2 SD error margin. Table 18 shows the values of SD for both Bayesian approaches. The proposed method presents significantly lower values of SD, providing more confident estimates than the Bayesian LASSO.

Finally, Table 19 shows the computational costs obtained with the algorithms for

Figure 30: Spectral signature for pixel $\#(20, 33)$.

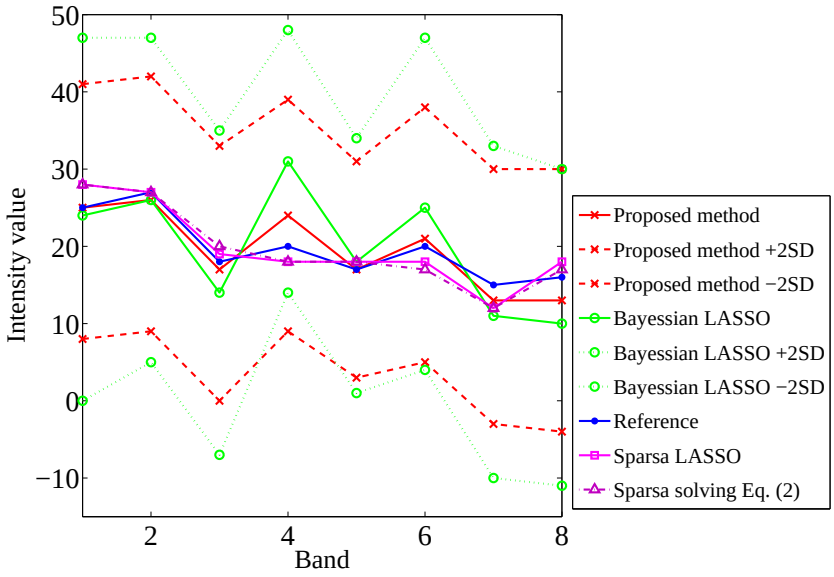


Table 18: Mean standard deviations of the estimations of $\boldsymbol{x}$.

Compression ratio	13%	26%	40%	53%	66%
Proposed method	10.85	10.51	9.79	9.21	9.05
Bayesian LASSO	25.90	18.21	15.06	13.10	13.10

Table 19: Computational costs for a 53% compression ratio.

Computational cost	Iterations	Seconds
Proposed method	500	20×10^3
Bayesian LASSO	750	18×10^3
SpaRSA smooth	300	316
SpaRSA LASSO	300	42

measurements associated with a 53% compression ratio. All algorithms were ran on a personal computer with an Intel Core i7-4790 CPU 3.60GHz processor and 32GB of memory. Algorithms were implemented using MATLAB with MEX files written in C. The algorithms based on MCMC methods are significantly slower than the ones based on optimization, as usual. However, it is important to note that optimization algorithms require to have their hyperparameters manually adjusted, and thus have to be run several times to find the optimal solution. Conversely, Bayesian methods can estimate their hyperparameters jointly with the recovered image coefficients using a single run.

6.4. **CHAPTER CONCLUSION**

This chapter introduced a hierarchical Bayesian model to solve the compressive spectral imaging problem by promoting the image to be sparse in a given basis and smooth in the spatial domain. A Gibbs sampler was developed to draw samples asymptotically distributed according to the corresponding posterior, sampling the full image in a single step to accelerate the convergence speed. The generated samples were used to calculate the Bayesian estimators of the unknown image. The resulting algorithm was compared to other reconstruction methods for a hyperspectral image compressed with different compression ratios. Our experiments showed that including a spatial smoothing term can improve the PSNR of the recovered image up to 2dB.

7. CONCLUSIONS, DISCUSSION, AND FUTURE WORK

The hypothesis of this dissertation was that it is feasible to design a strategy for optimizing the pseudorandom structure of the main compressive spectral imagers in terms of its maximum eigenvalue concentration and to optimize the computational efficiency of a recovery algorithm that exploits the structure characteristics of the compressive spectral imaging matrices where the computational optimization is quantified in terms of reconstruction quality and speed of convergence.
In this dissertation, a set of works demonstrates that it is possible to design the structure of the compressive spectral imaging matrix and to improve the recovery reconstruction algorithm for increasing the quality of the recovered images.
The main difference between the state-of-the-art matrix design minimizing problems and the proposed algorithm of Chapter 3 is that the gradient step in the state-of-the-art algorithms requires matrix products with complexity dependent on the size of the signal. Alternatively, the proposed algorithm takes advantage of the structure of the measurement matrix Φ and only operates in the modifiable elements which result a reduced computational complexity. Additionally, traditional algorithms assume that matrix normalization is feasible, which is not the case in optical systems with highly structured matrices.
In the work of Chapter 3, we aim to increase the rank of the measurement matrix by uniformly distributing the non-zero elements across rows and columns. In particular, we present a novel algorithm for achieving a uniform distribution of the non-zero elements when the locations of the modifiable elements are known in advance. The proposed algorithm performs well when the number of non-zero entries is significantly smaller than the total number of entries. It is possible to determine the locations of the non-zero elements that match the established expected values by performing an exhaustive search over the rows and columns. We randomly selected the locations by sampling a probability distribution that is related to the number of ones.
Several lines of research stem from this work. A future line of research is to design a proper representation basis Ψ for the proposed uniform binary sensing. Again, this design needs to take into account the physical limitations of the measurement matrix Φ and requires the spacing between the eigenvalues of the matrix $\mathbf{A} = \Phi\Psi$ being reduced while preserving the sparsification properties of Ψ.
Another line of research is to design the codification in a local neighborhood manner, different to the optimization presented there that aims to sense uniformly at a pixel by pixel or a voxel by voxel level. More specifically, the proposed design can be extrapolated to a multiscale approach where the optimization aims to design uniformly the codification at a pixel (voxel) level, but also at different groups of pixels (voxels) levels. The selection of these groups depends on the application and can consider physical characteristics of the sensing process as 2D or 3D local structures of the codification patterns.

The current developments in tensor imaging can also benefit from our designs, for instance, in the case of compressive sensing of sparse tensors, the projections acquired through the different dimensions of the tensorial data can be optimized following our method [92], [93].
It is important to emphasize that the codification design methods presented in Chapters 3 and 4 of this dissertation can be extended to several compressive architectures with binary codification. Some examples of these architectures include several compressive spectral imaging sensors [11], [94], [95], seismic data acquisition [96], and compressive spectrum sensing [97]. Another area of application includes the design of the binary codification applied in phase retrieval imaging [98], [99].
An approach for recovering an image with a given spatial resolution is presented in Chapter 5 where it is used the assumption that in a spatial neighborhood the spectral image has a similar spectral behavior. This model lets the image to be recovered with different spatial resolutions, this characteristic can be used in applications where it is important to create a quick preview of the scene. The models presented in Chapter 5 have better results in comparison with a traditional demosaicing recovering, obtaining an improvement range from 0.5 dB to 7 dB with respect to the traditional approach.
On the other hand, the proposed filtered gradient reconstruction algorithm of Chapter 6 achieves better performance, in terms of PSNR, than a traditional gradient-based reconstruction algorithm and similar results to the SALSA and AMP approaches, by taking advantage of a spatial filter operation. In a spatial filter operation, each pixel value is changed by a function of the intensities of pixels in a neighborhood, here, the spatial filter operation is used for increasing the smoothness characteristic of the spectral images that it is lost following the traditional reconstruction steps. This recovery algorithm can be seen as a variant of the gradient descent, where the main improvement attained is the better quality results achieved in fewer iterations than the traditional recovery algorithm.
Although that approach was implemented for the CASSI system, it can be extended to several compressive spectral imaging architectures [9]–[11], [33]. Additionally, due to the generality of the algorithm, it can be useful in several compressive sensing architectures modeled by sparse and structured sensing matrices where the signal for recovering is smooth, for instance, compressive computed tomography [13], single pixel imaging [6], and compressive video imaging [100].
We proposed a Gaussian low pass filter tuning the variance to 0.6 and a neighborhood of 5×5 pixels, this filter achieved better performance in the CASSI reconstructions. However, a drawback of the filtered recovery algorithm is that the parameters of the low pass filter, such as the size of the neighborhood and the variance, have to be tuned for the specific application. Once the appropriate values for the filter have been found, a general spatial filter operation is applied in each iteration, which is an advantage of the approach. The complexity of the algorithm increases only in one additional product in each iteration. On the other hand, this approach can be included in several gradient descent algorithms that solve the compressive sensing

problem [4], [41]–[44].

In the same line of reconstruction algorithms, Chapter 7 presents a hierarchical Bayesian model to solve the compressive spectral imaging problem by promoting the image to be sparse in a given basis and smooth in the spatial domain. The resulting algorithm was compared to other reconstruction methods for a hyperspectral image compressed with different compression ratios. The experiments show that including a spatial smoothing term can improve the PSNR of the recovered image up to 2dB. Future work includes the introduction of different regularization terms to promote smoothness of the recovered image, such as total variation. It would also be interesting to analyze the effect of small uncertainties affecting the sensing matrix on the performance of the image reconstruction algorithm.

BIBLIOGRAPHIC REFERENCES

[1] E. J. Candes and M. B. Wakin, “An introduction to compressive sampling”, *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[2] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, “Compressive sensing: From theory to applications, a survey”, *Journal of Communications and Networks*, vol. 15, no. 5, pp. 443–456, 2013.

[3] D. L. Donoho, “Compressed sensing”, *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[4] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, “Sparse reconstruction by separable approximation”, *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.

[5] E. J. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information”, English, *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[6] M. F. Duarte, M. A. Davenport, D. Takbar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, “Single-pixel imaging via compressive sampling”, *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, 2008, ISSN: 1053-5888.

[7] J. Ma, “Single-pixel remote sensing”, *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 199–203, 2009, ISSN: 1545-598X.

[8] A. Wagadarikar, R. John, R. Willett, and D. Brady, “Single disperser design for coded aperture snapshot spectral imaging”, *Applied Optics*, vol. 47, no. 10, B44, 2008, ISSN: 0003-6935.

[9] C. V. Correa, H. Arguello, and G. R. Arce, “Snapshot colored compressive spectral imager”, *Journal of the Optical Society of America A*, vol. 32, no. 10, pp. 1754–1763, 2015, ISSN: 1084-7529.

[10] X. Lin, Y. Liu, J. Wu, and Q. Dai, “Spatial-spectral encoded compressive hyperspectral imaging”, *ACM Trans. Graph.*, vol. 33, no. 6, 233:1–233:11, 2014, ISSN: 0730-0301.

[11] L. Wang, Z. Xiong, D. Gao, G. Shi, and F. Wu, “Dual-camera design for coded aperture snapshot spectral imaging.”, *Applied optics*, vol. 54, no. 4, pp. 848–58, 2015, ISSN: 1539-4522.

[12] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, “Compressive coded aperture spectral imaging: An introduction”, *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 105–115, 2014.

[13] D. J. Brady, A. Mrozack, K. MacCabe, and P. Llull, “Compressive tomography”, *Adv. Opt. Photon.*, vol. 7, no. 4, pp. 756–813, 2015.

[14] T. T. Do, L. Gan, N. H. Nguyen, and T. D. Tran, “Fast and Efficient Compressive Sensing Using Structurally Random Matrices”, *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 139–154, 2012, ISSN: 1053-587X.

[15] V. Tiwari, P. Bansod, and A. Kumar, “Designing sparse sensing matrix for compressive sensing to reconstruct high resolution medical images”, *Cogent Engineering*, vol. 2, no. 1, Z. Jin, Ed., 2015, ISSN: 2331-1916.

[16] N. Y. Yu and Y. Li, “Deterministic construction of Fourier-based compressed sensing matrices using an almost difference set”, *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, p. 155, 2013, ISSN: 1687-6180.

[17] P. C. Petrantonakis and P. Poirazi, “A compressed sensing perspective of hippocampal function”, *Frontiers in Systems Neuroscience*, vol. 8, p. 141, 2014, ISSN: 1662-5137.

[18] R. M. Willett, M. F. Duarte, M. A. Davenport, and R. G. Baraniuk, “Sparsity and structure in hyperspectral imaging : Sensing, reconstruction, and target detection”, *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 116–126, 2014.

[19] M. F. Duarte and Y. Eldar, “Structured compressed sensing: From theory to applications”, *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.

[20] S. A-qian, Z. Ding-fu, Y. Sheng, H. You-jun, Z. Peng, Y. Jian-ming, and Z. xin, “Optical scanning holography based on compressive sensing using a digital micro-mirror device”, *Opt. Commun.*, vol. 385, pp. 19 –24, 2017, ISSN: 0030-4018.

[21] M. L. Don, C. Fu, and G. R. Arce, “Compressive imaging via a rotating coded aperture”, *Appl. Opt.*, vol. 56, no. 3, B142–B153, 2017.

[22] Z. Li, J. Suo, X. Hu, C. Deng, J. Fan, and Q. Dai, “Efficient single-pixel multispectral imaging via non-mechanical spatio-spectral modulation”, *Scientific Reports*, vol. 7, pp. 1 –7, 2017.

[23] T. Blumensath and M. E. Davies, “Iterative hard thresholding for compressed sensing”, *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[24] C. A. Metzler, A. Maleki, and R. G. Baraniuk, “From denoising to compressed sensing”, *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 5117–5144, 2016, ISSN: 0018-9448.

[25] M. A. Figueiredo, R. D. Nowak, and S. J. Wright, “Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems.”, *Selected Topics in Signal Processing, IEEE Journal*, vol. 1, no. 4, pp. 586–597, 2007.

[26] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, ser. Applied and Numerical Harmonic Analysis. Springer New York, 2013.

[27] J. Fessler, “Image reconstruction”, in *Image Reconstruction*, 2016, ch. 1, pp. 1–18.

[28] U. Feige, “Why are images smooth?”, in *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, ser. ITCS ’15, Rehovot, Israel: ACM, 2015, pp. 229–236, ISBN: 978-1-4503-3333-7.

[29] G. A. Shaw and H.-h. K. Burke, “Spectral Imaging for Remote Sensing”, *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 3–28, 2013.

[30] D. Wu and D. W. Sun, “Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review - part i: Fundamentals”, *Innovative Food Science and Emerging Technologies*, vol. 19, pp. 1 –14, 2013.

[31] J. M. Eichenholz, N. Barnett, and D. Fish, “Sequential filter wheel multispectral imaging systems”, in *Imaging and Applied Optics Congress*, Optical Society of America, 2010, ATuB2.

[32] R. Jolivot, P. Vabres, and F. Marzani, “Reconstruction of hyperspectral cutaneous data from an artificial neural network-based multispectral imaging system”, *Computerized Medical Imaging and Graphics*, vol. 35, no. 2, pp. 85 –88, 2011, Advances in Skin Cancer Image Analysis, ISSN: 0895-6111.

[33] H. Arguello and G. R. Arce, “Colored coded aperture design by concentration of measure in compressive spectral imaging”, *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1896–1908, 2014.

[34] G. Martín, J. M. Bioucas-Dias, and A. Plaza, “Hyca: A new technique for hyperspectral compressive sensing”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2819–2831, 2015.

[35] J. E. Fowler, “Compressive-projection principal component analysis”, *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2230–2242, 2009.

[36] M. Elad, “Optimized projections for compressed sensing”, *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, 2007, ISSN: 1053-587X.

[37] G. Li, Z. Zhu, D. Yang, L. Chang, and H. Bai, “On projection matrix optimization for compressive sensing systems”, *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2887–2898, 2013, ISSN: 1053-587X.

[38] G. Li, X. Li, S. Li, H. Bai, Q. Jiang, and X. He, “Designing robust sensing matrix for image compression”, *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5389–5400, 2015, ISSN: 1057-7149.

[39] T. Hong, H. Bai, S. Li, and Z. Zhu, “An efficient algorithm for designing projection matrix in compressive sensing based on alternating optimization”, *Signal Process.*, vol. 125, pp. 9 –20, 2016, ISSN: 0165-1684.

[40] T. Hong and Z. Zhu, “An efficient method for robust projection matrix design”, *Signal Process.*, vol. 143, pp. 200 –210, 2018, ISSN: 0165-1684.

[41] J. Jin, Y. Gu, and S. Mei, “A Stochastic Gradient Approach on Compressive Sensing Signal Reconstruction Based on Adaptive Filtering Framework”, *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 409–420, 2010.

[42] Y. Qiu, J. Yan, and F. Xu, “Nonmonotone adaptive Barzilai-Borwein gradient algorithm for compressed sensing”, *Abstract and Applied Analysis*, vol. 2014, 2014.

[43] G. Chen, D. Li, and J. Zhang, “Iterative gradient projection algorithm for two-dimensional compressive sensing sparse image reconstruction”, *Signal Processing*, vol. 104, pp. 15–26, 2014, ISSN: 01651684.

[44] S. Becker, J. Bobin, and E. J. Candes, “Nesta: A fast and accurate first-order method for sparse recovery”, *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.

[45] J. Tan, Y. Ma, and D. Baron, “Compressive imaging via approximate message passing with wavelet-based image denoising”, *2014 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2014*, vol. 63, no. 8, pp. 424–428, 2014.

[46] J. Tan, Y. Ma, H. Rueda, D. Baron, and G. Arce, “Compressive Hyperspectral Imaging via Approximate Message Passing”, pp. 1–20, 2015.

[47] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “Fast image recovery using variable splitting and constrained optimization”, *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2345–2356, 2010, ISSN: 10577149. arXiv: 0910.4887.

[48] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems”, *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, 2011, ISSN: 1057-7149.

[49] S. Ji, Y. Xue, and L. Carin, “Bayesian compressive sensing”, *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008, ISSN: 1053-587X.

[50] L. He and L. Carin, “Exploiting structure in wavelet-based bayesian compressive sensing”, *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3488–3497, 2009, ISSN: 1053-587X.

[51] M. Wang, X. He, L. Qing, and S. Xiong, “Tree-structured bayesian compressive sensing via generalised inverse gaussian distribution”, *IET Signal Processing*, vol. 11, no. 3, pp. 250–257, 2017, ISSN: 1751-9675.

[52] S. Banerje and A. Roy, *Linear Algebra and Matrix Anal. for Statist.* Baltimore, MA: Chapman & Hall/CRC, 2014, ch. 5, p. 129.

[53] H. E. Bell, “Gershgorin’s theorem and the zeros of polynomials”, *The American Math. Monthly*, vol. 72, no. 3, pp. 292–295, 1965, ISSN: 00029890, 19300972.

[54] H. Rueda, H. Arguello, and G. R. Arce, “Compressive spectral testbed imaging system based on thin-film color-patterned filter arrays”, *Appl. Opt.*, vol. 55, no. 33, pp. 9584–9593, 2016.

[55] C. V. Correa, H. Arguello, and G. R. Arce, “Snapshot colored compressive spectral imager”, *J. Opt. Soc. Am. A*, vol. 32, no. 10, pp. 1754–1763, 2015.

[56] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. Soc. of Ind. and Appl. Math., 2001, ch. 4, pp. 114–116.

[57] Y. Kaganovsky, D. Li, A. Holmgren, H. Jeon, K. P. MacCabe, D. G. Politte, J. A. O’Sullivan, L. Carin, and D. J. Brady, “Compressed sampling strategies for tomography”, *J. Opt. Soc. Am. A*, vol. 31, no. 7, pp. 1369–1394, 2014.

[58] Munsell Color Science Laboratory, *Studio for Scientific Imaging and Archiving of Cultural Heritage: Spectral Image Database*. [Online]. Available: https://www.rit.edu/cos/colorscience/mellon/im/spectralDatabase.php (visited on 05/15/2017).

[59] “The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography”, *Ultramicroscopy*, vol. 157, pp. 35–47, 2015, ISSN: 18792723.

[60] V. Ferraris, N. Dobigeon, Q. Wei, and M. Chabert, “Detecting changes between optical images of different spatial and spectral resolutions: A fusion-based approach”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1566–1578, 2018, ISSN: 0196-2892.

[61] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, J. Y. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya, “Hyperspectral pansharpening: A review”, *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 27–46, 2015, ISSN: 2473-2397.

[62] X. He, L. Condat, J. M. Bioucas-Dias, J. Chanussot, and J. Xia, “A new pansharpening method based on spatial and spectral sparsity priors”, *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 4160–4174, 2014, ISSN: 1057-7149.

[63] R. Restaino, G. Vivone, M. D. Mura, and J. Chanussot, “Fusion of multispectral and panchromatic images based on morphological operators”, *IEEE Transactions on Image Processing*, vol. 25, no. 6, pp. 2882–2895, 2016, ISSN: 1057-7149.

[64] Z. Chen, H. Pu, B. Wang, and G. M. Jiang, “Fusion of hyperspectral and multispectral images: A novel framework based on generalization of pan-sharpening methods”, *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 8, pp. 1418–1422, 2014, ISSN: 1545-598X.

[65] Q. Wei, N. Dobigeon, and J. Y. Tourneret, “Fast fusion of multi-band images based on solving a sylvester equation”, *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4109–4121, 2015, ISSN: 1057-7149.

[66] N. Akhtar, F. Shafait, and A. Mian, “Bayesian sparse representation for hyperspectral image super resolution”, in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3631–3640.

[67] C. Lanaras, E. Baltsavias, and K. Schindler, “Hyperspectral super-resolution by coupled spectral unmixing”, in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 3586–3594.

[68] O. Espitia, H. Arguello, and J. Y. Tourneret, “High-resolution spectral image reconstruction based on compressed data fusion”, in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 1970–1974.

[69] E. Vargas, H. Arguello, and J. Y. Tourneret, “Spectral image fusion from compressive measurements using spectral unmixing”, in *2017 IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2017, pp. 1–5.

[70] H. Rueda, H. Arguello, and G. R. Arce, “High-dimensional optimization of color coded apertures for compressive spectral cameras”, in *25th European Signal Processing Conference (EUSIPCO)*, Kos island, Greece: IEEE, 2017, pp. 493–497.

[71] H. Rueda, H. Arguello, and G. R. Arce, “Optimal Colored Coded Apertures for Compressive Spectral Imaging Systems”, in *Imaging and Applied Optics 2017 (3D, AIO, COSI, IS, MATH, pcAOP)*, Washington, D.C.: OSA, 2017, JTu5A.25, ISBN: 978-1-943580-29-3.

[72] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J. Y. Tourneret, M. Chen, and S. Godsill, “Multiband image fusion based on spectral unmixing”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7236–7249, 2016, ISSN: 0196-2892.

[73] E. Candès and T. Tao, “Decoding by Linear Programming”, *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[74] Y. Mejia and H. Arguello, “Binary codification design for compressive imaging by uniform sensing”, *IEEE Transactions on Image Processing*, vol. 27, no. 12, pp. 5775–5786, 2018, ISSN: 1057-7149.

[75] A. Barbero and S. Sra, “Fast algorithms for total-variation based optimization”, MPI fur biologische Kybernetik, Germany, Tech. Rep., 2010, pp. 1–25.

[76] Z. Sadeghipoor, Y. M. Lu, and S. Süsstrunk, “A novel compressive sensing approach to simultaneously acquire color and near-infrared images on a single sensor”, in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 1646–1650.

[77] H. K. Aggarwal and A. Majumdar, “Compressive sensing multi-spectral demosaicing from single sensor architecture”, in *2014 IEEE China Summit International Conference on Signal and Information Processing (ChinaSIP)*, 2014, pp. 334–338.

[78] J. Brauers and T. Aach, “A color filter array based multispectral camera”, in *12. Workshop Farbbildverarbeitung*, G. C. Group, Ed., Ilmenau, 2006.

[79] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2010, vol. 25, pp. 487–487, ISBN: 9780521833783. arXiv: 1111.6189v1. [Online]. Available: https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.

[80] M. A. Branch, T. F. Coleman, and Y. Li, “A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems”, *SIAM Journal on Scientific Computing*, vol. 21, no. 1, pp. 1–23, 1999.

[81] J. J. Moré and G. Toraldo, “Algorithms for bound constrained quadratic programming problems”, *Numerische Mathematik*, vol. 55, no. 4, pp. 377–400, 1989, ISSN: 0945-3245.

[82] P. H. Calamai and J. J. Moré, “Projected gradient methods for linearly constrained problems”, *Mathematical Programming*, vol. 39, no. 1, pp. 93–116, 1987, ISSN: 1436-4646.

[83] K. Zhang, X. Gao, D. Tao, and X. Li, “Single Image Super-Resolution With Non-Local Means and Steering Kernel Regression”, English, *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4544–4556, 2012, ISSN: 1057-7149.

[84] J. Barzilai and J. M. Borwein, “Two-Point Step Size Gradient Methods”, *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.

[85] T. Serafini, G. Zanghirati, and L. Zanni, “Gradient projection methods for quadratic programs and applications in training support vector machines”, *Optimization Methods and Software*, vol. 20, no. 2-3, pp. 353–378, 2005.

[86] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, *CAVE | Projects: Multispectral Image Database*, www1.cs.columbia.edu/CAVE/databases/multispectral/, 2008. [Online]. Available: http://www1.cs.columbia.edu/CAVE/databases/multispectral/ (visited on 01/30/2016).

[87] H. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce, “Higher-order computational model for coded aperture spectral imaging”, *Appl. Opt.*, vol. 52, no. 10, pp. D12–D21, 2013.

[88] S. Raman, T. J Fuchs, P. J Wild, E. Dahl, and V. Roth, “The Bayesian group-lasso for analyzing contingency tables”, in *Proc. 26th ACM Annu. Int. Conf. Mach. Learn. (ICML)*, Montreal, Quebec, 2009.

[89] G. Casella and C. P Robert, *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.

[90] F. Orieux, O. Féron, and J.-F. Giovannelli, “Sampling high-dimensional Gaussian distributions for general linear inverse problems”, *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 251–254, 2012.

[91] T. Park and G. Casella, “The Bayesian Lasso”, *J. Amer. Stat. Soc.*, vol. 103, no. 482, pp. 681–686, 2008.

[92] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, “Tensor decomposition for signal processing and machine learning”, *Trans. Sig. Proc.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017, ISSN: 1053-587X.

[93] S. Friedland, Q. Li, and D. Schonfeld, “Compressive sensing of sparse tensors”, *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4438–4447, 2014, ISSN: 1057-7149.

[94] G. Martín and J. M. Bioucas-Dias, “Spatial-spectral hyperspectral image compressive sensing”, in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 3988–3991.

[95] Y. August, C. Vachman, Y. Rivenson, and A. Stern, “Compressive hyperspectral imaging by random separable projections in both the spatial and the spectral domains”, *Appl. Opt.*, vol. 52, no. 10, pp. D46–D54, 2013.

[96] C. C. Mosher, E. Keskula, S. T. Kaplan, R. G. Keys, C. Li, E. Z. Ata, L. C. Morley, J. D. Brewer, F. D. Janiszewski, P. M. Eick, R. A. Olson, and S. Sood, “Compressive seismic imaging, seg technical program expanded abstracts”, in. 2012, pp. 1–5.

[97] J. Marín Alfonso, J. I. Martínez Torre, H. Arguello Fuentes, and L. B. Agudelo, “Compressive multispectral spectrum sensing for spectrum cartography”, *Sensors*, vol. 18, no. 2, 2018.

[98] K. Jaganathan, Y. Eldar, and B. Hassibi, “Phase retrieval with masks using convex optimization”, in *2015 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2015, pp. 1655–1659, ISBN: 978-1-4673-7704-1.

[99] S. Pinilla, H. García, L. Díaz, J. Poveda, and H. Arguello, “Coded aperture design for solving the phase retrieval problem in x-ray crystallography”, *Journal of Computational and Applied Mathematics*, vol. 338, pp. 111 –128, 2018, ISSN: 0377-0427.

[100] A. Sankaranarayanan, L. Xu, C. Studer, Y. Li, K. Kelly, and R. Baraniuk, “Video compressive sensing for spatial multiplexing cameras using motion-flow models”, *SIAM Journal on Imaging Sciences*, vol. 8, no. 3, pp. 1489–1518, 2015.

BIBLIOGRAPHY

A-QIAN, Sun et al. Optical scanning holography based on compressive sensing using a digital micro-mirror device. In: Opt. Commun, 2017, vol. 385, pp. 19 –24.

AFONSO, Manya V., BIOUCAS-DIAS, José M., and FIGUEIREDO, Mário A. An Augmented Lagrangian Approach to the Constrained Optimization Formulation of Imaging Inverse Problems. In: IEEE Trans. Image Process, 2011, vol. 20, no. 3, pp. 681–695.

AFONSO, Manya V., BIOUCAS-DIAS, José M., and FIGUEIREDO, Mário A. Fast image recovery using variable splitting and constrained optimization". In: IEEE Transactions on Image Processing, 2011, vol. 19, no. 9, pp. 2345–2356.

AGGARWAL, H. K. and MAJUMDAR, A. Compressive sensing multi-spectral demosaicing from single sensor architecture. In: 2014 IEEE China Summit International Conference on Signal and Information Processing (ChinaSIP), 2014, pp. 334–338.

AKHTAR, N., SHAFAIT, F., and MIAN, A. Bayesian sparse representation for hyperspectral image super resolution. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3631–3640.

ARCE, G. R. et al. Compressive Coded Aperture Spectral Imaging: An Introduction. In: IEEE Signal Processing Magazine, 2014, vol. 31, no. 1, pp. 105–115.

ARGUELLO, H. and ARCE, G. R. Colored Coded Aperture Design by Concentration of Measure in Compressive Spectral Imaging. In: IEEE Trans. Image Process, 2014, vol. 23, no.4, pp. 1896–1908.

ARGUELLO, H. et al. Higher-order computational model for coded aperture spectral imaging. In: Appl. Opt., 2013, vol. 52, no. 10, pp. D12–D21.

AUGUST, Yitzhak et al. Compressive hyperspectral imaging by random separable projections in both the spatial and the spectral domains. In: Appl. Opt., 2013, vol. 52, no. 10, pp. D46–D54.

BANERJE, Sudipto and ROY, Anindya. Linear Algebra and Matrix Anal. for Statist. Baltimore, MA: Chapman & Hall/CRC, 2014. Chap. 5, p. 129.

BARBERO, Alvaro and SUVRIT. A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems". In: SIAM Journal on Scientific Computing, 1999, vol. 21, no. 1, pp. 1–23.

BRAUERS, Johannes and AACH, Til. A Color Filter Array Based Multispectral Camera. In: 12. Workshop Farbbildverarbeitung, 2006, Ed. by German Color Group. Ilmenau.

CALAMAI, Paul H. and MORÉ, Jorge J. Projected gradient methods for linearly constrained problems. In: Mathematical Programming, 1987, vol. 39, no. 1, pp. 93–116.

CANDES, E. J., ROMBERG, J., and TAO, T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. In: IEEE Transactions on Information Theory, 2006, vol. 52, no. 2, pp. 489–509.

CANDES, E. J. and WAKIN, M. B. An Introduction to Compressive Sampling. In: IEEE Signal Processing Magazine, 2008, vol. 25, no. 2, pp. 21–30.

CANDÈS, Emmanuel and TAO, Terence. Decoding by Linear Programming. In: IEEE Trans. Inf. Theory, 2005, vol. 51, no. 12, pp. 4203–4215.

CASELLA, G. and ROBERT, C. P. Monte Carlo Statistical Methods. New York: Springer-Verlag, 1999.

CHEN, G., LI, D., and ZHANG, J. Iterative gradient projection algorithm for two dimensional compressive sensing sparse image reconstruction. In: Signal Processing, 2014, vol. 104, pp. 15–26.

CHEN, Z. et al. Fusion of Hyperspectral and Multispectral Images: A Novel Framework Based on Generalization of Pan-Sharpening Methods. In: IEEE Geoscience and Remote Sensing Letters, 2014, vol. 11, no. 8, pp. 1418–1422.

CORREA, Claudia V., ARGUELLO, Henry, and ARCE, Gonzalo R. Snapshot colored compressive spectral imager. In: J. Opt. Soc. Am., 2015, vol. 32, no. 10, pp. 1754–1763.

DO, Thong T. et al. Fast and Efficient Compressive Sensing Using Structurally Random Matrices. In: IEEE Transactions on Signal Processing, 2012, vol. 60, no. 1, pp. 139–154.

DON, Michael L., FU, Chen and ARCE, Gonzalo R. Compressive imaging via a rotating coded aperture. In: Appl. Opt., 2017, vol. 56, no. 3, B142–B153.

DONOHO, D. L. Compressed sensing. In: IEEE Transactions on Information Theory, 2006, vol. 52, no. 4, pp. 1289–1306.

DUARTE, M. F. and ELDAR, Y.C. Structured Compressed Sensing: From Theory to Applications. In: IEEE Trans. Signal Processing, 2011, vol. 59, no. 9, pp. 4053 4085.

DUARTE, M. F. et al. Single-Pixel Imaging via Compressive Sampling. In: IEEE Signal Process. Mag., 2008, vol. 25, no. 2, pp. 83–91.

EICHENHOLZ, Jason M., BARNETT, Nick, and FISH, Dave. Sequential Filter Wheel
Multispectral Imaging Systems. In: Imaging and Applied Optics Congress. Optical Society of America, 2010, ATuB2.

ELAD, M. Optimized Projections for Compressed Sensing. In: IEEE Trans. Signal Process, 2007, vol. 55, no. 12, pp. 5695–5702.

ESPITIA, O., ARGUELLO, H., and TOURNERET, J. Y. High-resolution spectral image reconstruction based on compressed data fusion. In: 2017 IEEE International
Conference on Image Processing (ICIP), 2017, pp. 1970–1974.

FEIGE, U. Why Are Images Smooth? In: Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science. ITCS ’15. Rehovot, Israel: ACM, 2015, pp. 229–236.

FERRARIS, V. et al. Detecting Changes between Optical Images of Different Spatial and Spectral Resolutions: A Fusion-Based Approach. In: IEEE Transactions on Geoscience and Remote Sensing, 2018, vol. 56, no. 3, pp. 1566–1578.

FESSLER. Image reconstruction. In: Image Reconstruction. Chap. 1, 2016, pp. 1–18.

FIGUEIREDO, Mário, NOWAK, Robert D, and WRIGHT, Stephen J. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. In: Selected Topics in Signal Processing, IEEE, 2007, vol. 1, no. 4, pp. 586–597.

FOUCART, S. and RAUHUT, H. A Mathematical Introduction to Compressive Sensing. Applied and Numerical Harmonic Analysis. Springer New York, 2013.

FOWLER, J. E. Compressive-Projection Principal Component Analysis. In: IEEE Transactions on Image Processing, 2009, vol. 18, no. 10, pp. 2230–2242.

FRIEDLAND, S., LI, Q., and D. SCHONFELD. Compressive Sensing of Sparse Tensors. In: IEEE Transactions on Image Processing, 2014, vol. 23, no. 10, pp. 4438–4447.

HE, L. and CARIN, L. Exploiting Structure in Wavelet-Based Bayesian Compressive Sensing. In: IEEE Transactions on Signal Processing, 2009, vol. 57, no. 9, pp. 3488–
3497.

HE, X. et al. A New Pansharpening Method Based on Spatial and Spectral Sparsity Priors. In: IEEE Transactions on Image Processing, 2014, vol. 23, no. 9, pp. 4160–
4174.

HONG, Tao and ZHU, Zhihui. An efficient method for robust projection matrix design. In: Signal Process., 2018, vol. 143, pp. 200 –210.

HONG, Tao et al. An efficient algorithm for designing projection matrix in compressive sensing based on alternating optimization. In: Signal Process., 2016, vol. 125, pp. 9 –20.

JAGANATHAN, Kishore, ELDAR, Yonina, and HASSIBI, Babak. Phase retrieval with masks using convex optimization. In: 2015 IEEE International Symposium on Information Theory (ISIT). IEEE, 2015, pp. 1655–1659.

JI, S., XUE, Y., and CARIN, L. Bayesian Compressive Sensing. In: IEEE Transactions on Signal Processing, 2008, vol. 56, no. 6, pp. 2346–2356.

JIN, Jian, GU, Yuantao, and MEI, Shunliang. A Stochastic Gradient Approach on Compressive Sensing Signal Reconstruction Based on Adaptive Filtering Framework. In: IEEE Journal of Selected Topics in Signal Processing, 2010, vol. 4, no. 2, pp. 409–420.

JOLIVOT, Romuald, VABRES, Pierre, and MARZANI, Franck. Reconstruction of hyperspectral cutaneous data from an artificial neural network-based multispectral imaging system. In: Computerized Medical Imaging and Graphics, Advances in Skin Cancer Image Analysis, 2011, vol. 35, no. 2. pp. 85 –88. ISSN: 0895-6111.

KAGANOVSKY, Yan et al. Compressed sampling strategies for tomography. In: J. Opt. Soc. Am., 2014, vol. 31, no. 7, pp. 1369–1394.

KAK, Avinash C and SLANEY, Malcolm. Principles of computerized tomographic imaging. Soc. of Ind. and Appl. Math. Chap. 4, 2001, pp. 114–116.

LANARAS, C., BALTSAVIAS, E., and SCHINDLER, K. Hyperspectral Super-Resolution by Coupled Spectral Unmixing. In: 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 3586–3594.

LI, G. et al. On Projection Matrix Optimization for Compressive Sensing Systems. In: IEEE Trans. Signal Process, 2013, vol. 61, no. 11, pp. 2887–2898.

LI, G. et al. Designing robust sensing matrix for image compression. In: IEEE Trans. Image Process, 2015, vol. 24, no. 12, pp. 5389–5400.

LI, Ziwei et al. Efficient single-pixel multispectral imaging via non-mechanical spatio-spectral modulation". In: Scientific Reports 7, 2017, pp. 1 –7.

LIN, Xing et al. Spatial-spectral encoded compressive hyperspectral imaging. In: ACM Trans. Graph, 2014, vol. 33, no. 6, pp. 233:1–233:11.

LONCAN, L. et al. Hyperspectral Pansharpening: A Review. In: IEEE Geoscience and Remote Sensing Magazine, 2015, vol. 3, no. 3, pp. 27–46.

MA, J. Single-Pixel Remote Sensing. In: IEEE Geosci. Remote Sens. Lett., 2009, vol. 6, no. 2, pp. 199–203.

MARÍN ALFONSO, Jeison et al. Compressive Multispectral Spectrum Sensing for Spectrum Cartography. In: Sensors, 2018, vol. 18, no. 2.

MARTÍN, G. and BIOUCAS-DIAS , J. M. Spatial-spectral hyperspectral image compressive sensing. In: 2017 IEEE International Geoscience and Remote Sensing
Symposium (IGARSS), 2017, pp. 3988–3991.

MARTÍN, G., BIOUCAS-DIAS, J. M. and PLAZA, A. HYCA: A New Technique for Hyperspectral Compressive Sensing. In: IEEE Transactions on Geoscience and Remote Sensing, 2015, vol. 53, no. 5, pp. 2819–2831.

MEJIA, Y. and ARGUELLO, H. Binary Codification Design for Compressive Imaging
by Uniform Sensing. In: IEEE Transactions on Image Processing, 2018, vol. 27, no. 12, pp. 5775–5786.

METZLER, C. A., MALEKI, A., and BARANIUK, R. G. From Denoising to Compressed Sensing. In: IEEE Transactions on Information Theory, 2016, vol. 62, no. 9, pp. 5117–5144.

MORÉ, Jorge J. and TORALDO, Gerardo. Algorithms for bound constrained quadratic programming problems. In: Numerische Mathematik, 1989, vol. 55, no. 4, pp. 377–400.

MOSHER, Charles C. et al. Compressive Seismic Imaging. In: SEG Technical Program Expanded Abstracts, 2012, pp. 1–5.

Munsell Color Science Laboratory. Studio for Scientific Imaging and Archiving of Cultural Heritage: Spectral Image Database. URL: https://www.rit.edu/ cos / colorscience / mellon / im / spectralDatabase . php (visited on 05/15/2017).

ORIEUX, F., FÉRON, O., and GIOVANNELLI, J-F. Sampling high-dimensional Gaussian distributions for general linear inverse problems. In: IEEE Signal Process. Lett., 2012, vol. 19, no. 5, pp. 251–254.

PARK, T. and CASELLA, G. The Bayesian Lasso. In: J. Amer. Stat. Soc., 2008, vol. 103, no. 482, pp. 681–686.

PETRANTONAKIS, Panagiotis C. and POIRAZI, Panayiota. A compressed sensing perspective of hippocampal function. In: Frontiers in Systems Neuroscience, 2014, vol. 8, pp. 141.

PINILLA, Samuel et al. Coded aperture design for solving the phase retrieval problem in X-ray crystallography. In: Journal of Computational and Applied Mathematics, 2018, vol. 338, pp. 111 –128. ISSN: 0377-0427.

QAISAR, S. et al. Compressive sensing: From theory to applications, a survey. In: Journal of Communications and Networks, 2013, vol. 15, no. 5, pp. 443–456.

QIU, Yuanying, YAN, Jianlei, and XU, Fanyong. Nonmonotone adaptive Barzilai-Borwein gradient algorithm for compressed sensing. In: Abstract and Applied Analysis, 2014.

RAMAN, S. et al. The Bayesian group-lasso for analyzing contingency tables. In: Proc. 26th ACM Annu. Int. Conf. Mach. Learn. (ICML). Montreal, Quebec, 2009.

RESTAINO, R. et al. Fusion of Multispectral and Panchromatic Images Based on Morphological Operators. In: IEEE Transactions on Image Processing, 2016, vol. 25, no. 6, pp. 2882–2895. ISSN: 1057-7149.

RUEDA, Hoover, ARGUELLO, Henry, and ARCE, Gonzalo R. Compressive spectral testbed imaging system based on thin-film color-patterned filter arrays. In: Appl. Opt., 2016, vol. 55, no. 33, pp. 9584–9593.

RUEDA, Hoover, ARGUELLO, Henry, and ARCE, Gonzalo R. High-dimensional optimization of color coded apertures for compressive spectral cameras. In: 25th European Signal Processing Conference (EUSIPCO). Kos island, Greece: IEEE, 2017, pp. 493–497.

RUEDA, Hoover, ARGUELLO, Henry, and ARCE, Gonzalo R. Optimal Colored Coded Apertures for Compressive Spectral Imaging Systems. In: Imaging and Applied Optics 2017 (3D, AIO, COSI, IS, MATH, pcAOP).Washington, D.C.: OSA, 2017, JTu5A.25.

SADEGHIPOOR, Z., LU, Y. M., and SÜSSTRUNK, S. A novel compressive sensing approach to simultaneously acquire color and near-infrared images on a single sensor. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 1646–1650.

SANKARANARAYANAN, A. et al. Video Compressive Sensing for Spatial Multiplexing Cameras Using Motion-Flow Models. In: SIAM Journal on Imaging Sciences, 2015, vol. 8, no. 3, pp. 1489–1518.

SERAFINI, Thomas, ZANGHIRATI, Gaetano, and ZANNI, Luca. Gradient projection methods for quadratic programs and applications in training support vector machines. In: Optimization Methods and Software, 2005, vol. 20, no. 2-3, pp. 353–378.

SHAW, Gary A. and K. BURKE, Hsiao-hua. Spectral Imaging for Remote Sensing. In: Lincoln Laboratory Journal, 2013, vol. 14, no. 1, pp. 3–28.

SIDIROPOULOS, Nicholas D. et al. Tensor Decomposition for Signal Processing and Machine Learning. In: Trans. Sig. Proc., 2017, vol. 65, no. 13, pp. 3551–3582.

TAN, Jin, MA, Yanting, and BARON, Dror. Compressive imaging via approximate message passing with wavelet-based image denoising. In: 2014 IEEE Global Conference on Signal and Information Processing, GlobalSIP, 2014, vol. 63, no. 8, pp. 424–428.

TAN, Jin et al. Compressive Hyperspectral Imaging via Approximate Message Passing. In: IEEE Journal of Selected Topics in Signal Processing, 2016, pp. 1–20.

PALENSTIJN, W J., BATENBURG, K J., and SIJBERS, J. The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography. In: Ultramicroscopy, 2015, vol. 157, pp. 35–47.

TIWARI, Vibha, BANSOD, P.P., and KUMAR, Abhay. Designing sparse sensing matrix for compressive sensing to reconstruct high resolution medical images. In: Cogent Engineering, 2015, vol. 2, no. 1. Ed. by Zhongmin Jin.

VARGAS, E., ARGUELLO, H., and TOURNERET, J. Y. Spectral image fusion from compressive measurements using spectral unmixing. In: 2017 IEEE 7th

International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2017, pp. 1–5.

WAGADARIKAR, Ashwin et al. Single disperser design for coded aperture snapshot spectral imaging. In: Applied Optics, 2008, vol. 47, no. 10, B44.

WANG, L. et al. Dual-camera design for coded aperture snapshot spectral imaging. In: Applied optics, 2015, vol. 54, no. 4, pp. 848–58.

WANG, M. et al. Tree-structured Bayesian compressive sensing via generalized inverse Gaussian distribution. In: IET Signal Processing, 2017, vol 11, no. 3, pp. 250–257.

WEI, Q., DOBIGEON, N., and TOURNERET, J. Y. Fast Fusion of Multi-Band Images Based on Solving a Sylvester Equation. In: IEEE Transactions on Image Processing, 2015, vol. 24, no. 11, pp. 4109–4121.

WEI, Q. et al. Multiband Image Fusion Based on Spectral Unmixing. In: IEEE Transactions on Geoscience and Remote Sensing, 2016, vol. 54, no. 12, pp. 7236–7249.

WILLETT, R. M. et al. Sparsity and Structure in Hyperspectral Imaging: Sensing, Reconstruction, and Target Detection. In: IEEE Signal Processing Magazine, 2014, vol. 31, no. 1, pp. 116–126.

WRIGHT, S. J., NOWAK, R. D., and FIGUEIREDO, M. A. T. Sparse Reconstruction by Separable Approximation. In: IEEE Transactions on Signal Processing, 2009, vol. 57, no. 7, pp. 2479–2493.

WU, Di and WEN SUN, Da. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review - Part I: Fundamentals. In: Innovative Food Science and Emerging Technologies, 2013, vol. 19, pp. 1 –14.

YASUMA, F. et al. (www1.cs.columbia.edu/CAVE/databases/multispectral/, 2008). CAVE | Projects: Multispectral Image Database. URL: http://www1.cs.columbia.edu/CAVE/databases/multispectral/ (visited on 01/30/2016).

YU, Nam Yul and LI, Ying. Deterministic construction of Fourier-based compressed sensing matrices using an almost difference set. In: EURASIP Journal on Advances in Signal Processing, 2013, vol. 1, p. 155.

ZHANG, K. et al. Single Image Super-Resolution with Non-Local Means and Steering Kernel Regression. In: IEEE Transactions on Image Processing, 2012, vol. 21, no. 11, pp. 4544–4556.