BLUE NOISE AND OPTIMAL SAMPLING ON GRAPHS

by

Alejandro Parada-Mayorga

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Summer 2019

BLUE NOISE AND OPTIMAL SAMPLING ON GRAPHS

by

Alejandro Parada-Mayorga

Approved: _____
Kenneth E. Barner, Ph.D.
Chair of the Department of Electrical and Computer Engineering

Approved: _____
Levi T. Thompson, Ph.D.
Dean of the College of Engineering

Approved: _____
Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Gonzalo R. Arce, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Daniel Lau, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Javier García-Frías, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Dominique Guillot, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Sebastian Cioaba, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Austin Brockmeier, Ph.D.
Member of dissertation committee

ACKNOWLEDGEMENTS

I would like to express my gratitude to professor **Gonzalo Arce** for giving me the opportunity to study here at University of Delaware and providing me the funding to conduct my research. This journey has reshaped my life and my way of thinking for good.

My sincere thanks goes to professor **Daniel L. Lau** for his valuable insights and contributions to this dissertation.

I would like to express my deepest appreciation to professors **Dominique Guillot** and **Sebastian Cioaba**. Working with them was for me one of the best experiences I had at UD. It was nice to be in the middle of high level rational discussions about topics that I love with people that have an expertise in the field.

I thank the **National Science Foundation (NSF)** and the **University of Delaware** for providing me the funding and an exceptional environment and conditions to conduct my research in several projects.

I am deeply indebted to **my family** for their love and support during this process. In particular, I wanna thank my dear sister **Pilar** for her encouragement when I was thinking about coming to the united states, without her strength I would not be where I am now.

I would like to thank some of my colleagues here at UD for their support on different occasions during my time in Newark. Many thanks to: **Andre Rauh**, **Claudia Correa**, **Chen Fu**, **Li Li**, **Bohan Zhang**, and **Edgar Salazar**.

I thank those professors that also stimulated part of my intellectual growth teaching their courses in inspiring and stimulating ways: **Naya Banerjee** (Probability) and **Javier Garcia-Frias** (Information Theory). It was a pleasure to attend those classes.

Finally, I want to thank my collaborators here at UD for their enthusiasm and good disposition to work with me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

New data science tools are emerging to process signals on graph structures and concepts of algebraic and spectral graph theory are being merged with methods used in computational harmonic analysis to analyze these signals. A common problem in these networks is to determine which nodes play the most important role, assuming there is a quantity of interest defined on the network. Graph signal sampling thus becomes essential. In the first part of this dissertation, we explore a novel departure from prior work, inspired by sampling patterns in traditional dithering and halftoning. Specifically, we design graph signal sampling techniques that promote the maximization of the distance between sampling nodes on the vertex domain and that are characterized on some subclasses of graphs by a low frequency energy. Sampling patterns with these characteristics are referred to in the spatial dithering literature as blue-noise. The connection between existing theoretical results about sampling signals on graphs and blue noise sampling patterns on graphs is established, showing also how the spectral characteristics of these patterns are shaped by their vertex domain attributes. Additionally, for the generation of blue noise patterns a void and cluster algorithm on graphs is proposed exploiting the vertex-domain distribution of the sampling nodes. Numerical experiments show that the reconstruction error obtained with these patterns is similar to the one obtained by the state of the art approaches. Additionally, we explore the uniqueness sets for signals on cographs. Using the structure of the tree representation of a cograph, we proposed an algorithm that find its uniqueness sets from very simple small size graphs without any spectral decomposition or extensive searches on the vertex domain. The analysis performed on threshold graphs allowed us to calculate a closed form solution for the uniqueness sets.

In the second part of this dissertation we consider the problem of sampling on regular grids for compressed sensing applications. We design optimal sampling patterns in coded apertures for CASSI systems and compressive X-ray tomosynthesis architectures, providing closed form solutions that outperform the results achieved using designs obtained with previous approaches, at a very low computational cost. Additionally, a rigorous estimate of the spectral resolution in general colored CASSI systems is provided exploiting the structure of the non-ideal sampling patterns obtained when wide spectral filters are considered.

Chapter 1

GENERAL INTRODUCTION: UNIVERSAL PROPERTIES OF *GOOD* SAMPLING PATTERNS

The problem of representing complex, large scale or high-dimensional quantities with low size or low dimension representations has played a central role in numerical analysis, mathematics, physics, engineering and recently in data sciencel [4]. When these quantities were represented in Euclidean spaces, Shannon presented what would become one of the corner stone results in signal processing, *the Shannon-Nyquist theorem*. This result stated some minimum requirements over a sampling grid to provide a unique representation of a bandlimited signal [4]. The consequences and implications of such remarkable result are well known, as well as its scenario of applicability which assumes the sampling performed is on a regular grid. Naturally, there was the question about whether it was possible or not to obtain a generalization of this result to more general spaces, non-necessarily Euclidean, and to scenarios where the sampling grid was not necessarily regular.

Keeping the assumption that the signal of interest is defined on a Euclidean space, the works in irregular sampling were prolific [4] and different insights have been established in the context of several real life applications [4]. On the other hand, the problem of sampling on general spaces non necessarily Euclidean has been addressed more slowly with outstanding results obtained for the sampling of functions on manifolds and graphs [5]. In particular, the work of Pesenson in [5–7] stated formally for functions on manifolds a principle that lies behind almost every sampling strategy no matter the context, *uniformity*. Pesenson presented a criteria that allows one to compare the quality of sampling patterns on a manifold for the reconstruction of a

bandlimited signal. It was shown that the sampling pattern that provides the lowest error in the reconstruction is built as a *uniformly spread* pattern on the manifold, where the points are spread as far apart as possible from each other. This result was also connected with the results of irregular sampling if the manifold considered was $\mathbb{R}^n$. Pesenson also developed formal results to provide a measure of the quality of a sampling pattern for signals on graphs, however a connection between these results and that uniformity observed in manifolds was not established.

In the first part of this dissertation thesis we provide a connection between this uniformity principle and the quality of sampling sets in graphs introducing the concept of *blue-noise* sampling on graphs. Blue noise sampling has its roots in digital halftoning where the central interest lies on finding a binary representation of a gray scale image. We show that in subclasses of graphs with the same local isoperimetric dimension, a uniform spreading of the sampling nodes leads to high frequency patterns on the graph Fourier domain, and this result can be connected with parameters that are associated to the quality of the reconstructions. We develop algorithms for the generation of ideal blue noise sampling patterns and low complexity algorithms based on random walks and error diffusion to state the basis for the development of efficient algorithms of graph blue noise.

Additionally, we study *uniqueness sets* in cographs and threshold graphs. Exploiting the structure of the *cotree representation* of a cograph we provide an efficient algorithm for the calculation of the uniqueness sets without requiring the use of geodesic distances and/or spectral decompositions. In the case of threshold graphs, which is a subclass of cographs, a closed form solution for the uniqueness sets is derived.

In the second part of this dissertation the sampling problem is considered in the context of compressed sensing applications. In particular, we find optimal sampling patterns for signals that are projected on low dimensional spaces associated to colored CASSI systems and X-ray tomosynthesis architectures. The optimality of these sampling patterns is determined by the *coherence* of a sensing matrix and surprisingly an attribute of these patterns is the uniformity exhibited along the *dimension of shots*

or captures. The solutions obtained in my research are represented with closed form expressions that allow a fast design and computation that outperforms all the other approaches in the literature. Additionally, analyzing the properties of non-ideal sampling patterns in colored coded apertures, we provide a rigorous estimate of the spectral resolution that can be achieved in CASSI systems when a given distribution of spectral responses is used.

1.1 Dissertation Format

This dissertation contains eight chapters. In the first chapter a general introduction is provided, showing a panoramic view of the contributions achieved. In the second chapter the analysis of blue-noise sampling on graphs is considered, while in chapter three low complexity algorithms for the generation of blue noise are considered stating the basis for future work in the short term. In chapter four uniqueness sets for cographs are analyzed. Chapters five and six are devoted to the analysis of optimal sampling patterns in coded apertures for compressed sensing applications. In chapter seven a super-resolution analysis is performed on CASSI systems based on the sampling patterns of colored coded apertures. Finally in chapter 8 general conclusions and remarks about future work are presented.

1.2 My original contributions

Graph Signal Processing

1. Generalizing *blue-noise* sampling for signals on graphs.
2. Establishing a rigorous relationship between blue-noise sampling patterns on graphs and the quality of sampling sets measured by theoretical parameters stated in the literature.
3. Establishing a rigorous relationship between the vertex-domain and the spectral characteristics of blue noise sampling patterns.

4. Developing an algorithm for the generation of blue-noise sampling patterns on graphs.

5. Performing a complete numerical validation of blue-noise sampling patterns against sampling patterns generated by the state of the art approaches.

6. Establishing an estimate of the stability of blue-noise sampling patterns.

7. Developing an efficient algorithm for the calculation of uniqueness sets for the sampling of signals on cographs, without requiring spectral decompositions or the calculation of geodesic distances.

8. Performing a numerical validation of the effectiveness of the uniqueness sets of cographs, on graphs that are approximately cographs.

9. Obtaining a closed form solution for the uniqueness sets in Threshold graphs.

Compressed sensing applications

1. The Calculation of a closed form solution for a family of optimal colored coded apertures in CASSI systems. This solution provides up to the date the best reconstruction results in terms of PSNR.

2. A complete numerical validation of the optimal codes against codes obtained by the state of the art approaches.

3. The calculation of a closed form solution for a family of optimal coded apertures in a compressive X-ray tomosynthesis architecture. With these designs the results obtained were far superior to the ones obtained by other approaches that spent hundreds of hours for the computation of a solution.

4. The theoretical calculation of the spectral resolution limits in colored CASSI architectures.

5. Experimental validation of the spectral resolution limits in colored CASSI architectures.

1.3 My publications

Journal Papers

1. **A. Parada-Mayorga**, D. Lau, Jhony H. Giraldo and G.R. Arce, "Blue-Noise Sampling on Graphs" (**Accepted**: IEEE Transactions on Signal and Information Processing over Networks).

2. **A. Parada-Mayorga**, D. Guillot, S. Cioaba and G.R. Arce, "Uniqueness sets in the Paley-Wiener Space of Cographs" (To be submitted to IEEE Transactions on Signal and Information Processing over Networks).

3. **A. Parada-Mayorga** and G. R. Arce, "Colored Coded Aperture Design in Compressive Spectral Imaging via Minimum Coherence," in IEEE Transactions on Computational Imaging, vol. 3, no. 2, pp. 202-216, June 2017.

4. **A. Parada-Mayorga** and G.R. Arce, "Spectral Super-Resolution in Colored Coded Aperture Spectral Imaging," in IEEE Transactions on Computational Imaging, vol. 2, no. 4, pp. 440-455, Dec. 2016.

5. E. Salazar, **A. Parada-Mayorga** and G. Arce, "Spectral Zooming and Resolution Limits of Spatial Spectral Compressive Spectral Imagers," in IEEE Transactions on Computational Imaging. doi: 10.1109/TCI.2019.2893596

6. Juan F. Florez, **A. Parada-Mayorga**, Daniel L. Lau, Kenneth Barner, and Gonzalo R. Arce, "Signal Recovery via Concomitants of Order Statistics" (To be submitted to IEEE Transactions on Signal Processing).

Conference Papers

1. **A. Parada-Mayorga**, D. Lau, J. Giraldo, G. Arce, "Blue-Noise Sampling of Signals on Graphs", International Conference on Sampling Theory and Applications (SampTA), Bordeaux, 2019 (**Accepted**).

2. **A. Parada-Mayorga**, D. Lau, J. Giraldo, G. Arce, "Sampling of Graph Signals with Blue Noise dithering", IEEE data Science Workshop, Minneapolis, Minnesota, 2019 (**Accepted**).

3. D. Guillot, **A. Parada-Mayorga**, S. Cioaba, G. Arce, "Optimal Sampling Sets in Cographs", IEEE data Science Workshop, Minneapolis, Minnesota, 2019 (**Accepted**).

4. **A. Parada-Mayorga**, A. Cuadros and G.R. Arce, "Coded Aperture Design for Compressive X-ray Tomosynthesis via Coherence Analysis". IEEE International Symposium on Biomedical Imaging. Melbourne, Australia. April 2017 .

5. **A. Parada-Mayorga**, G.R. Arce, "Spectral Super-Resolution in Colored Coded Aperture Spectral Imaging," in Imaging and Applied Optics 2015, OSA Technical Digest (online) (Optical Society of America, 2015), paper CTh2E.2.

6. E. Salazar, **A. Parada-Mayorga** and G.R. Arce, "Spatial Super-resolution reconstruction via SSCSI Compressive Spectral Imagers" (OSA Imaging and Applied Optics Congress 2018).

7. E. Salazar, **A. Parada-Mayorga** and G.R. Arce, "Spectral zooming in SSCSI Compressive Spectral Imagers" (OSA Imaging and Applied Optics Congress 2018).

1.4 My collaboration with other students

During my Ph.D studies I had the pleasure of collaborating with some of my colleagues here at UD. In particular, I collaborated with:

- Ph.D student Edgar Salazar. Collaboration in compressive spectral imaging.
- Ph.D student Juan Florez. Collaboration in inverse problems.

- Ph.D student Jhony Giraldo. Collaboration in graph signal processing.

Additionally, I had the chance to play a role of mentorship with:

- Ph.D student Daniela Dapena. Collaboration in graph signal processing.

1.5 My collaboration work after my Ph.D studies

- Ph.D student Jhony Giraldo. I will be working after my Ph.D studies in collaboration with Jhony on the low complexity algorithms for graph blue-noise whose basis is stated in this dissertation.

- Ph.D student Daniela Dapena. I will be working in collaboration with Daniela in the analysis of optimal sampling patterns on directed graphs and its applications to problems in finance and analytics. Additionally, Daniela and I will be analyzing the role and effectiveness of graph blue-noise on classification and manifold learning.

Chapter 2

BLUE-NOISE SAMPLING ON GRAPHS

2.1 Introduction

Interesting phenomena in nature can often be captured by graphs since objects and data are invariably inter-related in some sense. Social [8], financial [9], ecological networks, and the human brain [10] are a few examples of such networks. Data in these networks reside on irregular or otherwise unordered structures [11]. New data science tools are thus emerging to process signals on graph structures where concepts of algebraic and spectral graph theory are being merged with methods used in computational harmonic analysis [12–14]. A common problem in these networks is to determine which nodes play the most important role, assuming there is a quantity of interest defined on the network.

Graph signal sampling thus becomes essential. Naturally, the mathematics of sampling theory and spectral graph theory have been combined leading to generalized Nyquist sampling principles for graphs [12, 15–19]. In general, these methods are based on the underlying graph spectral decompositions [19–22].

This work explores a somewhat radical departure from prior work, inspired by sampling patterns in traditional dithering and halftoning. Specifically, we intend to design graph signal sampling techniques that promote the maximization of the distance between sampling nodes on the vertex domain that are typically characterized by a low frequency energy. Sampling patterns with these characteristics are referred to in the spatial dithering literature as *blue-noise* [23, 24].

In this chapter, the connection between the properties of blue-noise sampling patterns and the results related with sampling sets in graphs is established, showing

that blue-noise like sampling patterns in graphs are connected with good sampling sets in terms of preserving the uniqueness of the representation of the sampled signal in a noise-free scenario. Additionally, it is shown how the inter-distance between the sampling nodes affects the redness in a given sampling pattern. We provide a measure of the bandwidth of the signals that can be uniquely represented from the vertex-domain characteristics of blue-noise sampling patterns. A numerical algorithm is proposed in order to compute these blue-noise patterns based on their vertex-domain distribution. In particular, trying to exploit the distribution of the sampling points on the nodes of the graph, a void and cluster algorithm on graphs is developed [25], allowing the generation of patterns that lead to reconstruction errors of bandlimited signals, similar to the ones obtained in the state-of-the-art literature.

2.2 Preliminaries

Sandryhaila [26] proposed a theoretical framework for the analysis and processing of signals on graphs based on the properties of the adjacency matrix. This approach is rooted in *algebraic signal processing*, whereas authors like Fuhr and Pesenson [15, 16, 27], Puy [28] and Shuman [12, 13, 29] based their analysis of signals on graphs, relying on the properties of the *Laplacian matrix*. In both approaches the Fourier transform of the signals on the graph is defined in terms of a spectral decomposition of the adjacency matrix and the Laplacian matrix respectively, using the set of eigenvectors as the Fourier basis for the representation of the signals.

The first approach offers a direct connection with the shift operator used in traditional signal processing, while the second resembles the main ideas of Fourier analysis in linear spaces in which the eigenfunctions of the Laplacian operator are used as the basis representation of the signal. The two approaches use a unitary operator, and problems like sampling and filtering can be successfully considered in both scenarios. In this work, the combinatorial Laplacian matrix is used as the building block, and the graphs considered are undirected, weighted, connected and simple. Consequently,

part of the developments proposed rely on the theoretical results obtained by Furh and Pesenson [15, 16, 27] in harmonic analysis on graphs.

2.2.1 Graph Signal Sampling

Let $G = (V(G), E(G))$ be an undirected, weighted, connected, simple graph with a set of nodes, $V(G)$, and a set of edges, $E(G)$. $\mathbf{W}$ is the adjacency matrix (symmetric), with $\mathbf{W}(u, v) \geq 0$ the weight connecting the nodes u and v and $u \sim v$ indicates that $\mathbf{W}(u, v) > 0$. The degree matrix, $\mathbf{D}$, is a diagonal matrix whose entries are given according to:

$$\mathbf{D}(u, u) = \sum_{v \in V(G)} \mathbf{W}(u, v). \tag{2.1}$$

For any graph G, its volume is defined as $vol(G) = \sum_{u \in V(G)} \mathbf{D}(u, u)$, and the volume of a subset $S \subset V(G)$ is defined as $vol(S) = \sum_{u \in S} \mathbf{D}(u, u)$. On the graph G, the combinatorial Laplacian operator is defined as the positive semi-definite operator:

$$\mathbf{L} = \mathbf{D} - \mathbf{W}, \tag{2.2}$$

whose eigenvalues are organized as $0 \leq \mu_1 \leq \mu_2 \leq \ldots \leq \mu_N$, $N = |V(G)|$ [30]. A real signal, $\boldsymbol{x}$, on the graph is then defined as the mapping $\boldsymbol{x} : V(G) \longrightarrow \mathbb{R}$ denoted by the vector $\boldsymbol{x} \in \mathbb{R}^N$ where $\boldsymbol{x}(v)$ is the value of the signal associated to $v \in V(G)$. The support of $\boldsymbol{x}$ is denoted by $\text{supp}(\boldsymbol{x})$, and the restriction of $\boldsymbol{x}$, to any subset $S \subset V(G)$, is denoted by $\boldsymbol{x}(S)$. It is worth noticing that:

$$(\mathbf{L}\boldsymbol{x})(v) = \sum_{u \in V(G)} (\boldsymbol{x}(v) - \boldsymbol{x}(u))\, \mathbf{W}(v, u). \tag{2.3}$$

If the spectral decomposition of the operator $\mathbf{L}$ is denoted by $\mathbf{L} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\mathsf{T}$, then the Graph Fourier Transform (GFT) of the signal $\boldsymbol{x}$ on G is given by $\hat{\boldsymbol{x}} = \mathbf{U}^\mathsf{T}\boldsymbol{x}$. There is a direct analogy between the concept of frequency in traditional Fourier Analysis and the behavior of the Graph Fourier Transform as is stated in [12]. Considering this analogy, the bandwidth of a signal $\boldsymbol{x}$ can be defined using the nonzero components of $\hat{\boldsymbol{x}}$. It is said that $\boldsymbol{x}$ has bandwidth $\omega \in \mathbb{R}_+$ on the spectral axis if $\hat{\boldsymbol{x}} \in PW_\omega(G) =$

span$\{\mathbf{U}_k : \mu_k \leq \omega\}$, where $PW_\omega(G)$ is the Paley-Wiener space of bandwidth ω [15] and $\mathbf{U}_k$ indicates the first k column vectors in $\mathbf{U}$. In some cases the bandwidth is also represented with the largest integer k such that $\mu_k \leq \omega$.

Given a notion of bandwidth, one invariably questions the notion of sampling rate and whether the number of samples or nodes of a graph can be reduced without loss of information to the signal. We, therefore, define sampling of a signal $\boldsymbol{x}$ on the graph G, by choosing the components of $\boldsymbol{x}$ on a subset of nodes, $S = \{s_1, \ldots, s_m\} \subset V(G)$. The sampled signal is given by $\boldsymbol{x}(S) = \mathbf{M}\boldsymbol{x}$ where $\mathbf{M}$ is a binary matrix whose entries are given by $\mathbf{M} = [\boldsymbol{\delta}_{s_1}, \ldots, \boldsymbol{\delta}_{s_m}]^\mathsf{T}$ and $\boldsymbol{\delta}_v$ is the $N-$ dimensional Kronecker column vector centered at v. Given $\boldsymbol{x}(S)$, it is possible to obtain a reconstructed version of $\boldsymbol{x}$ in different ways depending on whether the bandwidth of the signal is known. We assume that the bandwidth is known and that the reconstruction is given by:

$$\boldsymbol{x}_{rec} = \underset{\boldsymbol{z} \in span(\mathbf{U}_k)}{\operatorname{argmin}} \|\mathbf{M}\boldsymbol{z} - \boldsymbol{x}(S)\|_2^2 = \mathbf{U}_k \left(\mathbf{M}\mathbf{U}_k\right)^\dagger \boldsymbol{x}(S) \tag{2.4}$$

where $(\mathbf{M}\mathbf{U}_k)^\dagger$ is the Moore-Penrose pseudo-inverse of $\mathbf{M}\mathbf{U}_k$ [31, 32]. Alternatively, in [1] it is shown that a consistent reconstruction of the signal can be obtained from its samples using *interpolation splines*.

The problem of optimally sampling a signal on a graph can now be summarized as choosing S such that we maximize the available bandwidth of $\boldsymbol{x}(S)$. To this end, Pesenson defines [15,16] a Λ-removable set for $\Lambda > 0$ as the subset of nodes, $S \subset V(G)$, for which:

$$\|\boldsymbol{x}\|_2 \leq (1/\Lambda)\|\mathbf{L}\boldsymbol{x}\|_2 \qquad \forall\ \boldsymbol{x} \in L_2(S), \tag{2.5}$$

where $L_2(S)$ is the set of all signals, $\boldsymbol{x}$, with support in $S \subset V(G)$ (i.e. elements of $\boldsymbol{x}$ not included in S are equal to zero) and finite ℓ_2 norm. The largest value of Λ for which eqn. (2.5) holds is denoted by Λ_S. Notice that for any subset of nodes there exists a Λ-removable set with larger or smaller Λ_S. Therefore, Λ_S ultimately determines how much importance a given set has in the sampling process of a signal with a specific bandwidth. The relationship between properties of removable sets and the sampling problem was established by Pesenson in the following theorem:

Theorem 1 (Theorem 5.1 in [15]). *If for a set $S \subset V(G)$, its compliment $S^c = V(G)\backslash S$ is a $\Lambda_{S^c}-$removable set, then all signals in $PW_\omega(G)$ are completely determined by its values in S, whenever $0 < \omega < \Lambda_{S^c}$.*

In [27], another result related with sampling sets is established using a constant that can be calculated directly with the weights, $\mathbf{W}$, of the graph, G, stated in the following theorem:

Theorem 2 ([27]). *Every $S \subset V(G)$ is a uniqueness set for all functions in $PW_\omega(G)$ with any $\omega < K_S$, where*

$$K_S = \inf_{v \in S^c} w_S(v) \tag{2.6}$$

and $w_S(v) = \sum_{s \in S} \mathbf{W}(s, v)$.

Theorems 1 and 2 play a central role in the description of properties for different classes of sampling sets as it is possible to consider that a *good* sampling set, S, promotes the maximization of constants, Λ_{S^c} and K_S. In particular, it will be shown in the following sections that blue-noise sampling patterns indeed promote high values of these constants.

Recently Pesenson [33] introduced results that characterize the representation of a band limited signal in terms of induced subgraphs obtained from partitions of $V(G)$ that cover $V(G)$. This statement can be summarized in the following theorem.

Theorem 3 (5.1,6.1,6.2 [33]). *Let G be a connected finite or infinite and countable graph. Suppose that $\mathcal{P} = \{V(\Omega_j)\}_{j=1}^{j=|\mathcal{P}|}$ is a disjoint cover of $V(G)$ by **connected** and finite subgraphs Ω_j. Let $\mathbf{L}_{\Omega_j}$ be the Laplace operator of the **induced** graph Ω_j whose first nonzero eigenvalue is $\mu_{1,j}$. If $\Lambda_\mathcal{P} = \inf_j \mu_{1,j} > 0$ and $\Lambda_\mathcal{P} > \frac{1+\alpha}{\alpha}\omega$ with $\alpha > 0$, then every signal $\boldsymbol{x} \in PW_\omega(G)$ is uniquely determined by the values $\boldsymbol{x}^\mathsf{T}\boldsymbol{\xi}_j$, where $\boldsymbol{\xi}_j = \boldsymbol{\chi}_j/\sqrt{|V(\Omega_j)|}$ with $\boldsymbol{\chi}_j(V(\Omega_j)) = 1$ and $\boldsymbol{\chi}_j(V(\Omega_j)^c) = 0$. Additionally, $\boldsymbol{x}$ can be reconstructed from this set of values in a stable way.*

It is important to remark the meaning and implications of Theorem 3. This result shows that $V(G)$ can be divided into disjoint subsets that cover $V(G)$, and a given

band limited signal can be reconstructed from the average values of the signal in those regions. Additionally, the constant $\Lambda_{\mathcal{P}}$ associated to the partition provides a measure of the quality of the reconstruction obtained from the regions on $V(G)$ defined by $\mathcal{P}$. It is also worthy to point out that the size of the elements in the partition has a natural limit as $\mathbf{L}_{\Omega_j}$ is expected to have at least one nonzero eigenvalue, which would not be the case when Ω_j consist of one single vertex. This result will allow us to establish a connection between the spectral and vertex domain behavior of sampling patterns in some classes of graphs. Additionally, we will show that from a blue-noise sampling pattern $\boldsymbol{s}$, it is possible to build a partition that can be used to estimate the bandwidth of signals that are uniquely represented by their samples on the sampling nodes indicated by $\boldsymbol{s}$.

2.2.2 Optimal Graph Sampling

The problem of finding the best S is a combinatorial problem of calculating Λ_{S^c} for all sampling sets and choosing the set with the largest value of Λ_{S^c}, a prohibitively expensive process for large graphs. Allowing for some short cuts, a simple, greedy procedure for finding a good sampling set starts with an empty set of nodes and iteratively adds one node at a time, taking the best available node at each iteration according the value of a cost function. Several authors have formulated the problem of sampling and reconstruction in the presence of measurement noise, and in these works objective functions have been proposed that minimize the reconstruction error in terms of the worst case [21], where $S^{opt} = \arg\max_{|S|=m} \sigma_1^2$, the mean case [31], where $S^{opt} = \arg\max_{|S|=m} \sum_{i=1}^{\min\{m,k\}} \sigma_i^{-2}$, and the maximum volume case [32], where $S^{opt} = \arg\max_{|S|=m} \prod_{i=1}^{\min\{m,k\}} \sigma_i^2$; and σ_i represents the i^{th} singular value of the matrix $\mathbf{MU}_k$ consisting of the first k eigenvectors of $\mathbf{W}$ or $\mathbf{L}$ respectively, sampled on the rows indicated by S. In [34] the optimal sampling set is obtained considering the same cost function for the mean case, but using the singular values of $\mathbf{\Lambda U}^{\mathsf{T}}\text{diag}(S)$, where $\text{diag}(S)$ is the diagonal matrix whose entries are given by $\text{diag}(S)_{i,i} = 1 \Leftrightarrow i \in S$.

In order to reduce computational complexity, Anis et. al. [31] defines *graph*

spectral proxies of order q as estimates of the cutoff frequency of a given signal which can be used to define cutoff frequency estimates for a subset of nodes S according to:

$$\Omega_q(S) = \min_{\phi \in L_2(S^c)} \left(\frac{\|\mathbf{L}^q \phi\|_2}{\|\phi\|_2} \right)^{\frac{1}{q}}, \tag{2.7}$$

with $\mathbf{L}^q$ being the q^{th} power of $\mathbf{L}$ [20,31]. Anis et. al. further shows that, for any $q \in \mathbb{N}$ and S, it is possible to have perfect reconstruction when $\omega < \Omega_q(S)$. The value of $\Omega_q(S)$ can be calculated as $\Omega_q(S) = (\sigma_{1,q})^{\frac{1}{2q}}$, where $\sigma_{1,q}$ denotes the smallest eigenvalue of the reduced matrix $\mathbf{L}^{2q}_{S^c,S^c}$. The optimal sampling set can then be represented as the solution of the problem:

$$S_q^{opt} = \arg\max_{|S|=m} \Omega_q(S), \tag{2.8}$$

which is still combinatorial; however, Anis et. al. proposes a heuristic rule to solve eqn. (2.8) using the first eigenvector of $\mathbf{L}^q_{S^c,S^c}$. Basically, a node is added to the sampling set according to the index of the component with maximum absolute value for the first eigenvector of $\mathbf{L}^q_{S^c,S^c}$. The quality of the sampling set is also related to the value of q, which should be selected as large as possible at the expense of a higher computational cost. In [35], some performance theoretical bounds for these greedy sampling techniques are derived.

In some scenarios for sampling signals on graphs a spectral decomposition of the operators is not available, and therefore, there is a strong need for vertex domain sampling schemes that attempt to build good sampling patterns based entirely on the local graph structure around a node. In particular, for those cases where the graphs are too large for calculating the eigenvalues and eigenvectors of the GFT, several authors have looked at the problem of sampling using subsets of nodes that may not be optimal but are still very good at preserving band-limited signals. In the case of Puy et. al [28], the authors perform a random selection of nodes with a recovery algorithm that involves a probability distribution on a diagonal matrix, $\mathbf{P}$, in addition to the sampling matrix operator $\mathbf{M}$. The reconstructed signal $\boldsymbol{x}_{rec}$ can then be calculated as:

$$\boldsymbol{x}_{rec} = \arg\min_{\boldsymbol{z} \in \mathbb{R}^N} \left(\left\| \mathbf{P}^{-1/2} (\mathbf{M}\boldsymbol{z} - \boldsymbol{x}(S)) \right\|_2^2 + \tau \boldsymbol{z}^\intercal g(\mathbf{L}) \boldsymbol{z} \right), \tag{2.9}$$

where $\boldsymbol{x}(S) = \mathbf{M}\boldsymbol{x}$ is the sampled version of the signal $\boldsymbol{x}$, τ is a regularization parameter selected empirically and $g(\cdot)$ is a polynomial function selected also empirically.

Puy *et. al* [28] further show that an optimal $\mathbf{P}$ can be determined by the use of the *local graph coherence*, ν_k, on the nodes of the graph. The value of $\nu_k(i)$ at the node i can be calculated as $\nu_k(i) = \|\mathbf{U}_k\boldsymbol{\delta}_i\|_2$, where $\boldsymbol{\delta}_i$ is the Kronecker vector centered at node i, and it provides a measure about how important is the node i for the sampling of a signal with bandwidth k. If $\nu_k(i)$ is equal to 1 for a particular node i, then there exists k-bandlimited graph signals whose energy is solely concentrated in this i^{th} node. If $\nu_k(i)$ is equal to 0, then no k-bandlimited graph signal has any energy in this i^{th} node. Therefore, node i can be deleted with no repercussions.

Because the calculation of $\nu_k(i)$ requires the knowledge of the spectral decomposition, Puy *et. al* propose an approximate estimation of $\nu_k(i)$ that can be obtained without the calculation of any spectral decomposition, which allows the solution of eqn. (2.9). When the optimal $\mathbf{P}$ is used, Puy *et. al* show that the matrix $\mathbf{M}\mathbf{P}^{-1/2}$ satisfies a restricted isometry property when the number of samples is on the order of $O(k \log k)$, which provides a strong guarantee for the exact recovery of the signal. This represents an elegant result but with the drawback that $O(k \log k)$ is substantially higher than k, which is the optimal number of samples required to reconstruct a signal of bandwidth k.

Recently, Tremblay et. al. [32] proposed the use of *determinantal point processes (DPP)* in order to obtain the matrix $\mathbf{P}$ used in [28]. It is shown in [32] that an optimal $\mathbf{P}$ can be obtained using DPP when $\mathbf{U}_k$ is known. Additionally, when the spectral decomposition is not accessible, it is shown how a variant of the Wilson's Algorithm introduced in [36] can be used in order to obtain a sampling set that can be shown is related with a DPP that leads to an approximate version of the optimal $\mathbf{P}$. The reconstruction of the signal is obtained by as solution of eqn. (2.9); however, these results do not represent an improvement with respect to Anis et. al. [31] or Chen et. al. [21] and may lead to larger reconstruction errors when the graph considered does not have a strong community graph structure [32]. Wang et. al. [37] consider the

sampling and reconstruction of signals adapting concepts and ideas from frame theory, developing an iterative approach based on the concept of local sets.

Marques et. al. [38], proposed a different approach with respect to previous works, considering the sampling and reconstruction of the signal using its samples on a single node. The central idea is based on the information provided by the sequential application of the shift operator. The technique itself represents a novel alternative with potential applications in network analysis and its computational cost may be a drawback when large size graphs are considered.

2.3 Blue-noise Sampling on Graphs

This work proposes a different approach to graph signal sampling: the application of spatial dithering to the graph vertex domain where the spectral properties of well formed sampling patterns will equally benefit the graph vertex domain as they do the spatial. This approach is motivated by the well established research in digital halftoning, which is the process of converting a continuous tone image or photograph into a pattern of printed and not-printed dots for reproduction by inkjet or laser printers [23, 24, 39]. Halftoning algorithms based on error-diffusion are of particular importance because they produce random patterns of homogeneously distributed dots where minority pixels (black dots in highlights or white dots in shadows) are spaced as far apart as possible. These patterns have power spectra dominated by high frequency energy, earning the name, "blue-noise," since blue is the high frequency component of white light. Low frequency energy or red-noise contributes to halftone patterns looking fuzzy or noisy to the human visual system and are, therefore, to be avoided [23, 39].

In order to establish a blue-noise model for sampling signals on a graph, we first propose the idea of a binary dither pattern on a graph, $G = (V(G), E(G))$, as the binary graph signal, $\boldsymbol{s} \in \{0, 1\}^N$. We refer to the fraction of samples that we intend to preserve as the density $d = m/N$, where $\|\boldsymbol{s}\|_0 = m$. In the case of a white-noise dither pattern as illustrated in Fig. 2.1 (top) on a Sensor Network graph for $d = 0.1$, $\boldsymbol{s}$ is selected uniformly at random from the space of binary signals for which $\|\boldsymbol{s}\|_0 = dN$;

therefore each component of $\boldsymbol{s}$ can be modeled as a Bernoulli random variable with expected value $\mathbb{E}\{\boldsymbol{s}(\ell)\} = d$.

2.3.1 Vertex-domain Characteristics

We define blue-noise sampling on graphs in terms of its desired vertex domain characteristics which resemble the spatial characteristics of blue-noise sampling in traditional halftoning. As such, we need to define a measure of spacing between neighboring nodes on a graph by defining a path between the nodes v_a and v_b by the *sequence* $(v_a, u_1, u_2, \ldots, u_n, v_b)$ where each node in the sequence indicates the nodes visited when going from v_a to v_b, visiting between nodes with edge weights that are different from zero. Having a sequence of nodes defining a path, we define the length of this path according to:

$$|(v_a, u_1, u_2, \ldots, u_n, v_b)| \quad = \quad \mathbf{W}(v_a, u_1) \; + \; \mathbf{W}(u_1, u_2) \; + \; \cdots \; + \; \mathbf{W}(u_n, v_b), \quad (2.10)$$

where the shortest path between two nodes, v_a and v_b, is the path with minimum length and is represented by γ_{v_a,v_b}. For any $v \in V(G)$, the *open ball* of radius ρ and centered in v is defined as $B(v, \rho) = \{u \in V(G) : |\gamma_{v,u}| < \rho\}$. The symbol $\boldsymbol{\Gamma} \in \mathbb{R}^{N \times N}$ represents the matrix of geodesic distances in the graph, where $\boldsymbol{\Gamma}(u, v) = |\gamma_{u,v}|$. We will refer to a collection of subsets of $V(G)$ as a *cover* if the union of such subsets is equal to $V(G)$, and the cover will be called disjoint if the subsets are pairwise disjoint.

Having defined the notion of distance on the vertex domain of the graph, we can introduce blue-noise sampling taking into account its characteristics in traditional halftoning. Blue-noise halftoning is characterized on the spatial domain by a distribution of binary pixels where the minority pixels are spread as homogeneously as possible. Distributing pixels in this manner creates a pattern that is aperiodic, isotropic (radially symmetric), and does not contain any low-frequency spectral components. Halftoning a continuous-tone, discrete-space, monochrome image with blue-noise produces a pattern that, as Ulichney [23] describes, is visually "pleasant" and "does not clash with the structure of an image by adding one of its own, or degrade it by being too 'noisy' or

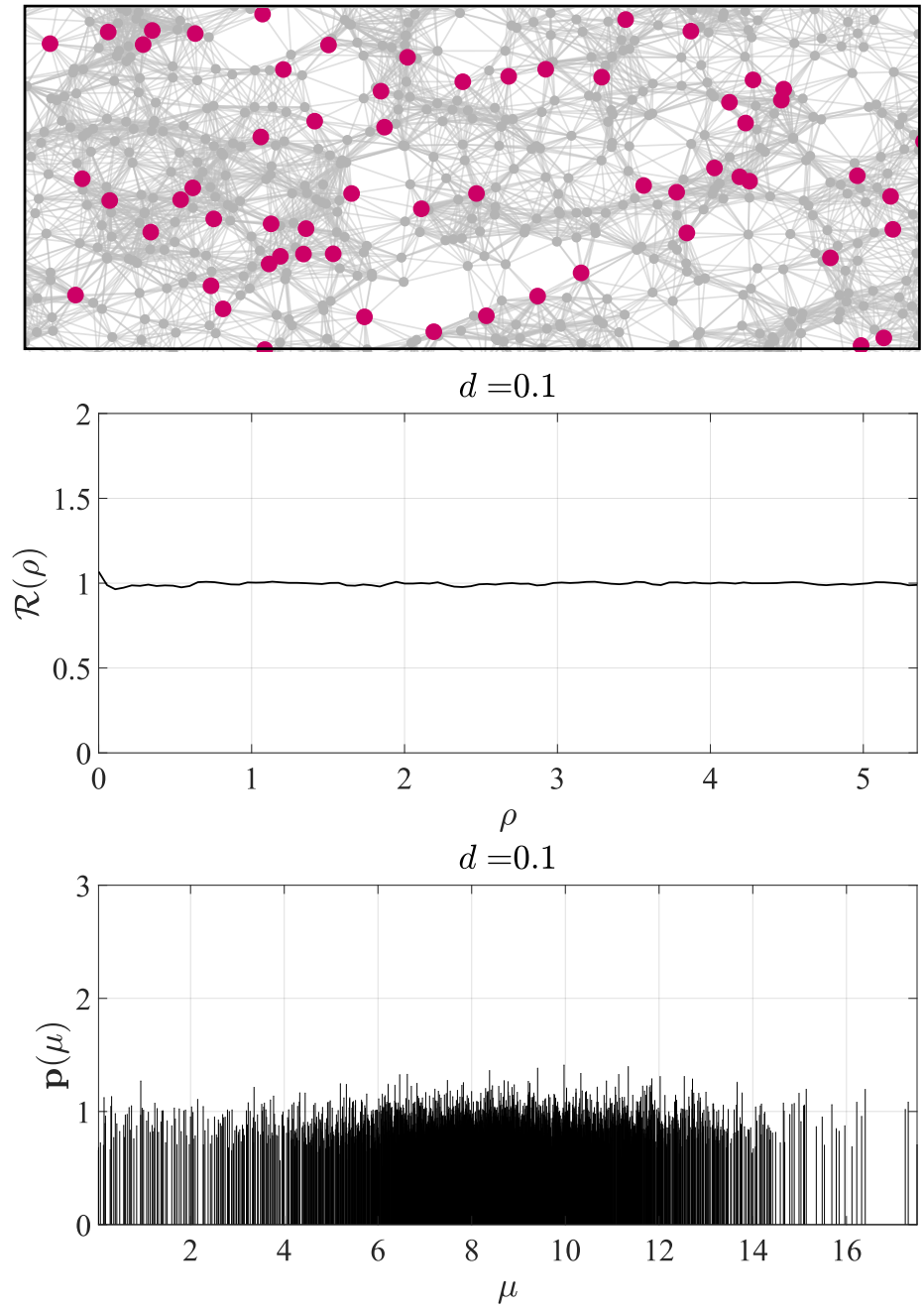


Figure 2.1: Illustration of the spatial and spectral properties of (top) a white-noise dither pattern on a Sensor Network graph with density, $d = 0.1$, with (center) a flat pair correlation approximately equal to 1.0 for all internode distances, ρ, and (bottom) an approximately flat power spectra for all frequencies, μ.

uncorrelated." Similarly on a graph, the minority nodes composing the binary signal are expected to be equally spaced apart when measuring distance as the sum of the weights forming the shortest path. With these ideas, we formally introduce blue-noise

in the following definition.

Definition 1 (Blue-Noise on Graphs)**.** *Let $S \subset V(G)$ be a subset of nodes in the graph G with $S = \{s_1, s_2, \dots s_m\}$. Then, it is said that S represents an ideal blue-noise sampling pattern, if the following conditions are satisfied:*

- *There is a collection of open balls $B(s_i, \lambda)$ that forms a cover of $V(G)$.*
- *The value of λ is the minimum possible for all the subsets of nodes of size m.*

Definition 1 implies that ideal blue-noise sampling patterns have their sampling nodes located as far as possible from each other, or in other words, there is a typical vertex domain spreading of the sampling nodes. Figure 2.5 (top) illustrates a typical blue-noise pattern on a sensor network. We use this attribute as the defining characteristic of a blue-noise sampling pattern; however, we will show in later sections that, in some classes of graphs this vertex domain spreading implies or is correlated with a high frequency behavior on the spectral domain.

2.3.1.1 Vertex-domain Metrics

For any $v \in V(G)$, the *annulus* of radius ρ, width θ, and center v is defined as $B_\theta(v, \rho) = \{u \in V(G) : \rho - \theta \leq |\gamma_{v,u}| < \rho + \theta\}$. Figure 2.2 illustrates an example of $B_\theta(v, \rho)$. With a notion of concentric rings in $B_\theta(v, \rho)$, we can now define the pair correlation on a graph. Specifically, let $S = \text{supp}(\boldsymbol{s}) = \{s_1, s_2, \dots, s_m\}$ be the support of the sampling pattern $\boldsymbol{s}$ and let $\|\boldsymbol{s}(B_\theta(s_i, \rho))\|_0$ be the number of 1s of $\boldsymbol{s}$ on $B_\theta(s_i, \rho)$, then the sample pair correlation function, $\mathcal{R}_{\boldsymbol{s}}(\rho)$, associated to $\boldsymbol{s}$ is defined by

$$\mathcal{R}_{\boldsymbol{s}}(\rho) = \frac{\frac{1}{m}\sum_{i=1}^{m} \|\boldsymbol{s}(B_\theta(s_i, \rho))\|_0}{\frac{1}{N}\sum_{v \in V(G)} \|\boldsymbol{s}(B_\theta(v, \rho))\|_0}. \tag{2.11}$$

Notice that the numerator in (2.11) indicates the average number of 1s in $\boldsymbol{s}$ on a ring of width θ that is centered on a 1 of $\boldsymbol{s}$, while the denominator indicates the average number of 1s on the ring of the same width when it is centered at any arbitrary node.

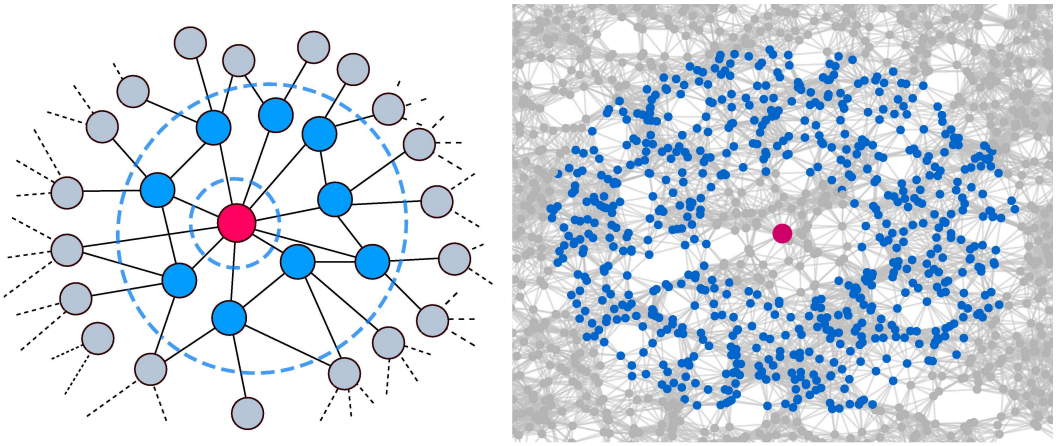


Figure 2.2: Illustration of $B_\theta(v,\rho)$ in a graph. Left: representation of $B_\theta(v,\rho)$ for small values of ρ and θ. Right: Illustration of $B_\theta(v,\rho)$ for large values of ρ and θ. The nodes in blue color are located in the annulus of radius ρ and width θ centered at the node v indicated in red color.

Now, the pair correlation for q realizations $\boldsymbol{s}_1,\ldots,\boldsymbol{s}_q$ of a random sampling pattern is defined as

$$\mathcal{R}(\rho) = \frac{1}{q}\sum_{r=1}^{q}\mathcal{R}_{\boldsymbol{s}_r}(\rho), \tag{2.12}$$

as the influence of a sampling point at node v on all other nodes in the geodesic annular region $B_\theta(v,\rho)$. Notice that for the computation of eqn. (2.11) several values of θ can be considered, in this work the value of θ is the average of nonzero edge weights.

Note that a maxima of $\mathcal{R}(\rho)$ can be considered as an indication of the frequent occurrence of the inter-node distance, ρ, between nodes set to 1 whereas minima indicate a reduced occurrence. Since for random patterns the expected number of 1s in any annular ring is proportional to the number of nodes within the ring, we expect a pair correlation equal to 1 for all $\rho > 0$ as illustrated in Fig. 2.1 (center).

Blue-noise, when applied to an image of constant gray-level g, spreads the minority pixels of the resulting binary image as homogeneously as possible such that the pixels are separated by an average distance, λ_b, referred to as the *principal wavelength* of blue-noise. These minority pixels are the pixels that are used to represent the properties of a region on a grayscale image, for instance in a dark region the minority pixels are labeled with 1 while in white regions the minority pixels are labeled with 0. Then,

the value of λ_b is defined as the radius of a round disc, surrounding a minority pixel, such that the ratio of the surface area of a minority pixel to the surface area of the disc is equal to the density of minority pixels, $d = g$ for $0 < g \leq 1/2$ and $d = 1 - g$ for $1/2 < g \leq 1$, which we can write as:

$$d = \frac{D_x D_y}{\lambda_b^2}, \tag{2.13}$$

where D_x and D_y are the sampling periods (distance between samples) of the digital image in the x and y directions, respectively.

In order to extend the notion of principal wavelength to graphs, we need a notion of surface area as the expected number of graph nodes, $\mathbb{E}\{\mathcal{N}(\lambda)\}$, within a distance or path length, λ, of a given minority node. We expect the ratio of our single, minority node to all nodes within a path length, λ_b, to equal the density level according to:

$$d = \frac{1}{\mathbb{E}\{\mathcal{N}(\lambda_b)\}}. \tag{2.14}$$

Being that $\mathbb{E}\{\mathcal{N}(\lambda_b)\}$ is graph dependent, the graph blue-noise wavelength, λ_b, is likewise graph dependent and its characterization is still an open area of research [40–42]. In general, one can derive λ_b versus d experimentally as we have in Fig. 2.3 where we show the principal wavelength versus the density sampling d for some commonly used graphs. We note that in the case of the sensor graph, λ varies smoothly with $d = 1/\mathbb{E}\{\mathcal{N}(\lambda_b)\}$ while, in the case of the community graph, it varies with a piecewise constant behavior with respect to d.

In light of the nature of graph blue-noise to isolate minority nodes, we can begin to characterize blue-noise graph signals in terms of the pair correlation, $\mathcal{R}(\rho)$, by noting that: (a) few or no neighboring minority nodes lie within a path length of $\rho < \lambda_b$; (b) for $\rho > \lambda_b$, the expected number of minority nodes per unit area tends to stabilize around a constant value; and (c) the average number of minority nodes within the path length, ρ, increases sharply nearly λ_b. The resulting pair correlation for blue-noise is, therefore, of the form in Fig. 2.4 (top), where $\mathcal{R}(\rho)$ shows: (a) a strong inhibition of minority nodes near $\rho = 0$, (b) a decreasing correlation of minority nodes

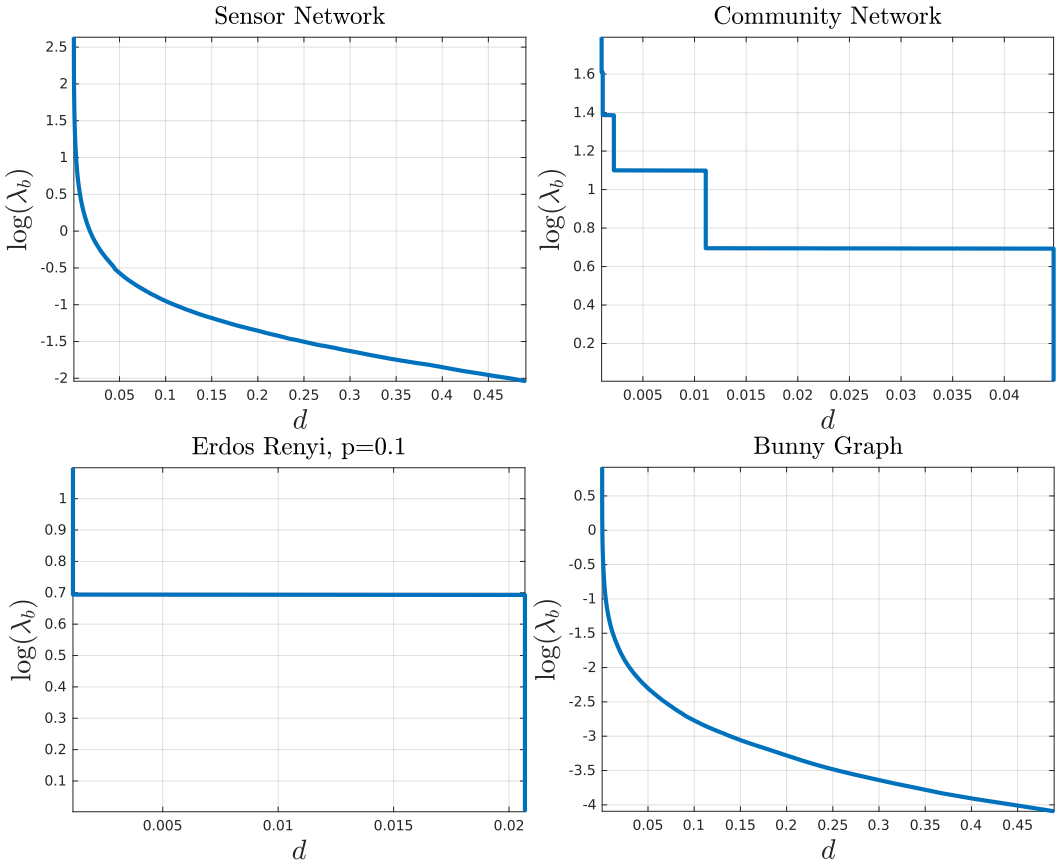


Figure 2.3: Principal wavelength λ_b versus the density of the sampling pattern for four different graphs traditionally considered in the literature.

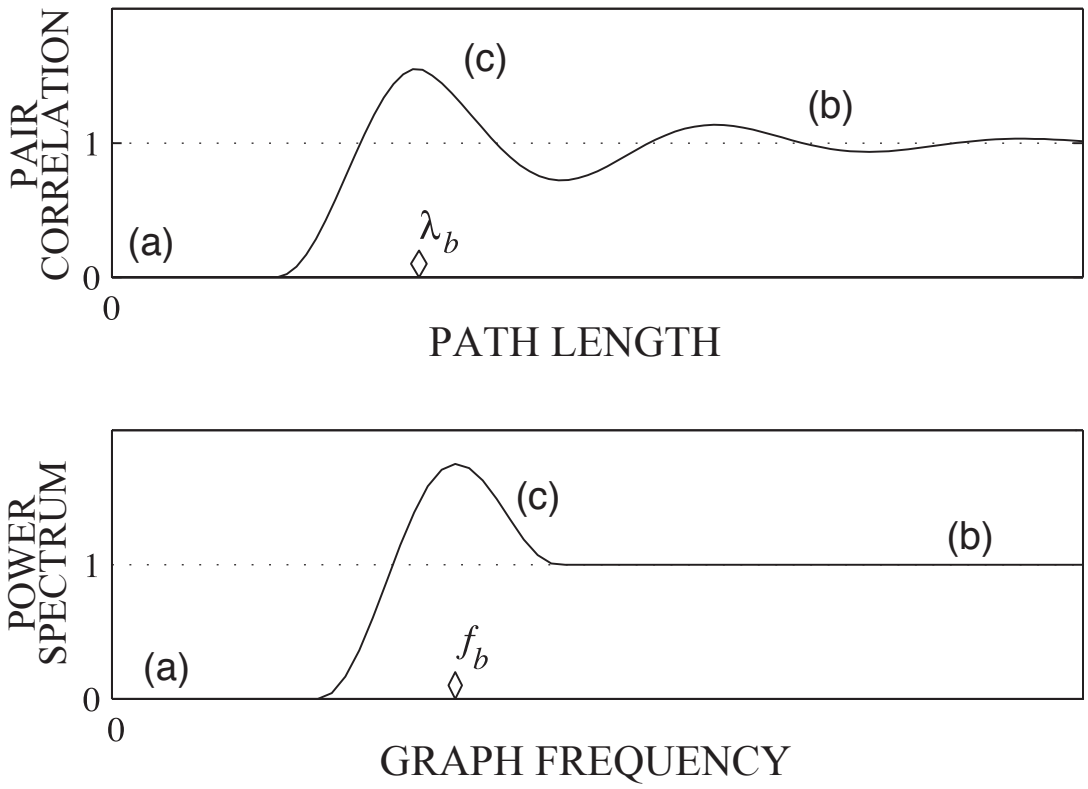


Figure 2.4: The ideal (top) pair correlation and (bottom) power spectra for blue-noise sampling patterns.

with increasing ρ ($\lim_{\rho\to\infty} \mathcal{R}(\rho) = 1$), and (c) a frequent occurrence of the inter-node distance λ_b, the principal wavelength, indicated by a series of peaks at integer multiples of λ_b. The principal wavelength is indicated in Fig. 2.4 (top) by a diamond located along the horizontal axis. Returning to the sample blue-noise signal of Fig. 2.5 (top), the resulting pair correlation of Fig. 2.5 (center) has a principal wavelength of $\lambda_b = 0.56$ with a clearly visible peak of 1.55, meaning that nodes equal to 1 are 55% more likely to occur at a distance of $\rho = 0.56$ from an existing 1 than for the unconstrained probability of a node being equal to 1.

2.3.2 Spectral Characteristics

Blue noise sampling patterns are characterized in traditional halftoning for a high frequency behavior [24]. In this section we state a connection between the spectral

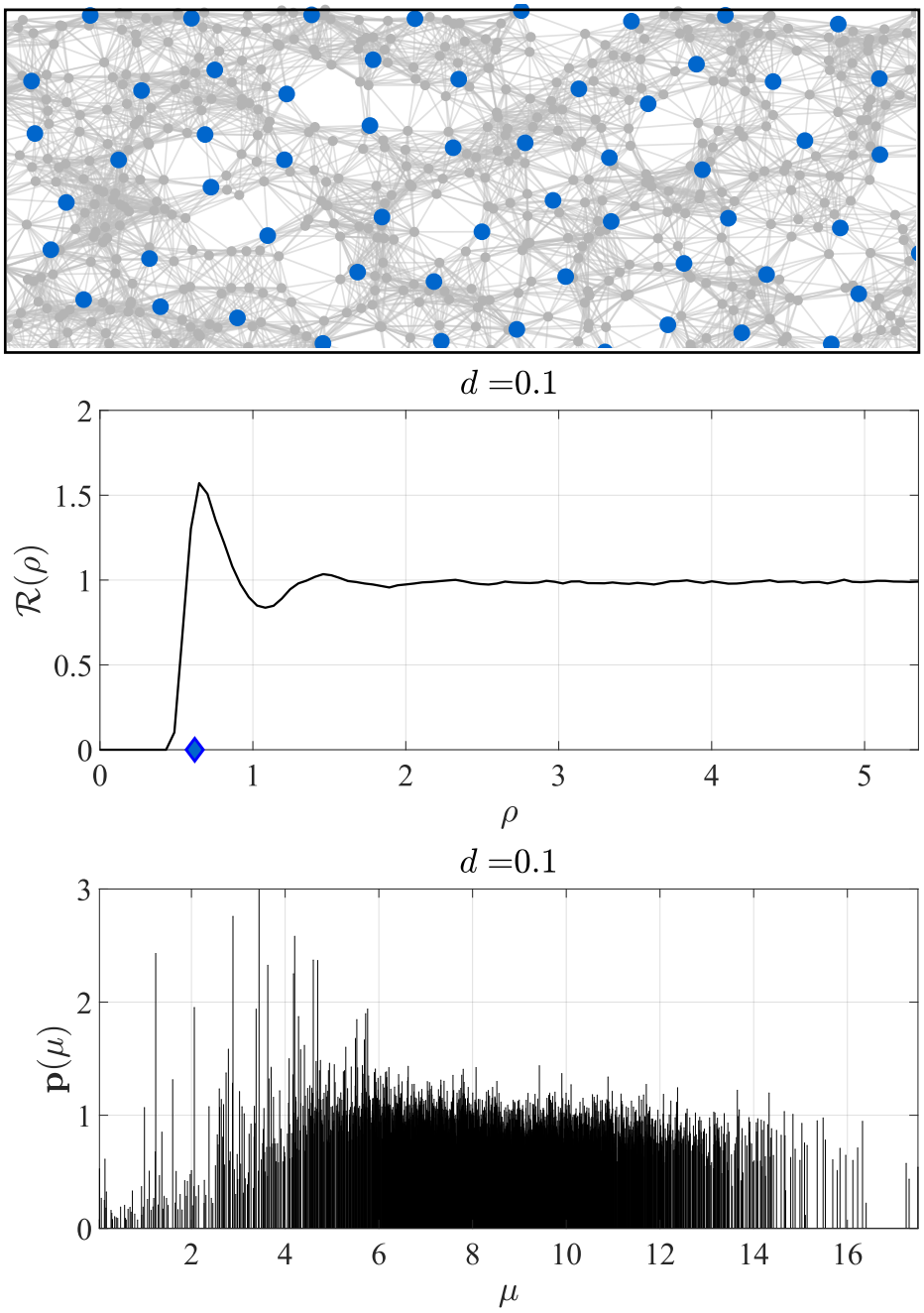


Figure 2.5: Illustration of the spatial and spectral properties of (top) a blue-noise dither pattern on a Sensor Network graph with density, $d = 0.1$, with (center) a pair correlation peak at the principal wavelength, λ_b, and (bottom) an approximately high frequency only power spectrum for frequencies, μ.

characteristics of a sampling pattern and its vertex domain characteristics, using the local properties of partitions of $V(G)$ that are measured by the isoperimetric constants

of local induced subgraphs. In order to characterize the frequency content of a sampling pattern, a cost function is proposed. In particular, we propose a scalar measure of low-frequency energy, $R_{\boldsymbol{s}}$, in the signal $\boldsymbol{s}$, as the weighted sum of all Fourier coefficients' energies:

$$R_{\boldsymbol{s}} = \frac{1}{\|\hat{\boldsymbol{s}}\|_2^2} \sum_{\ell=2}^{N} \frac{\hat{\boldsymbol{s}}^2(\ell)}{\mu_\ell} = \frac{1}{m} \sum_{\ell=2}^{N} \frac{\hat{\boldsymbol{s}}^2(\ell)}{\mu_\ell}, \tag{2.15}$$

where $\hat{\boldsymbol{s}}$ is the graph Fourier transform of $\boldsymbol{s}$. $R_{\boldsymbol{s}}$ is coined as the *redness* of $\boldsymbol{s}$ as it measures low frequency spectral content.

In order to establish a connection between $R_{\boldsymbol{s}}$ and the vertex domain characteristics of a sampling pattern, it is important to consider the following theorems.

Theorem 4. *For the graph $G = (V(G), E(G))$, let $\mathcal{P} = \{V(\Omega_1), V(\Omega_2), \ldots, V(\Omega_{|\mathcal{P}|})\}$ be a partition of $V(G)$, where Ω_j is the induced subgraph given by $V(\Omega_j)$. Let δ_j be the isoperimetric dimension of Ω_j. Then if*

$$\delta_1 = \delta_2 = \ldots = \delta_{|\mathcal{P}|} = \delta \tag{2.16}$$

it follows that

$$\Lambda_{\mathcal{P}} > \min \left\{ C_\delta \left(\frac{1}{vol(\Omega_1)} \right)^{\frac{2}{\delta}}, \ldots, C_\delta \left(\frac{1}{vol(\Omega_{|\mathcal{P}|})} \right)^{\frac{2}{\delta}} \right\} \tag{2.17}$$

where C_δ is a constant that depends on δ.

Proof: See Appendix 2.7

Theorem 4 indicates that when the graph has a local invariant isoperimetric dimension, the quality of a partition $\mathcal{P}$ for the representation of bandlimited signals, measured by $\Lambda_{\mathcal{P}}$, is defined by the set in $\mathcal{P}$ with the largest volume. The concept of isoperimetric dimension, originally defined on manifolds, provides a measure of how similar is the global behavior of a manifold with respect to a Euclidean space [43]. Similarly, in the case of graphs, the isoperimetric dimension indicates how close the behavior of a graph is with respect to regular grid-like graphs. For instance, the isoperimetric dimension of the $n-$dimensional regular grid is n [43]. In the following theorem, we indicate when the right-hand side of eqn. (2.17) is maximized.

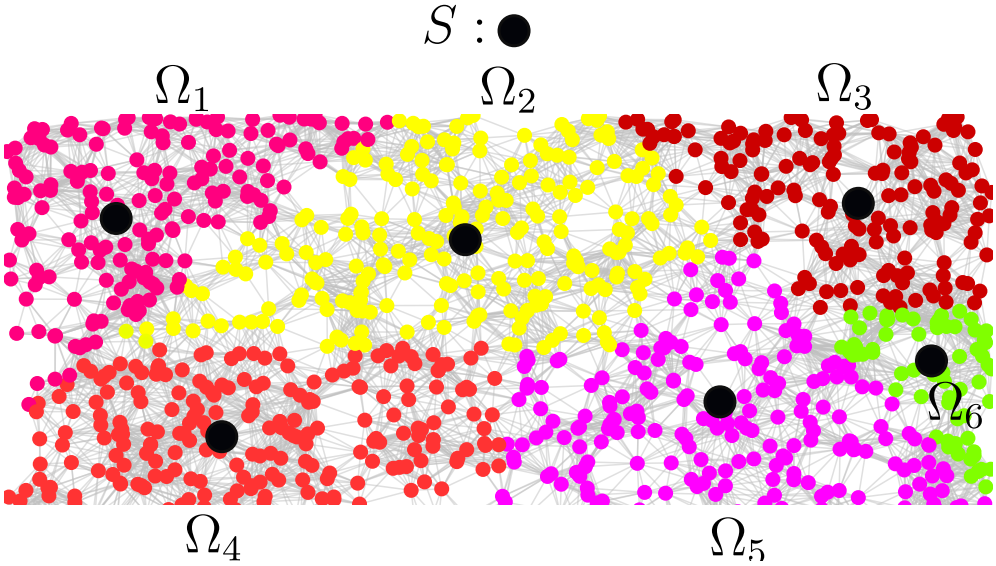


Figure 2.6: Partition of $V(G)$ for a graph G. Each color indicates the subgraph Ω_j induced by $V(\Omega_j)$. Illustration shows how a sampling pattern can be built from the partition selecting the sampling nodes on the set S, whose nodes are indicated in black color. Notice that the sampling pattern indicated satisfies eqn. (2.19) and eqn. (2.20).

Theorem 5. *Under the conditions stated in Theorem 4 and for a fixed value of $|\mathcal{P}|$, the partition that maximizes the right hand side of eqn. (2.17) satisfies that*

$$vol(\Omega_i) = vol(\Omega_j) \quad \forall i, j. \tag{2.18}$$

Proof: See Appendix 2.8

Under the conditions stated in Theorem 4, Theorem 5 provides the characteristics of the partition that will maximize the bandwidth of signals that can be represented in a unique way via their average values on the elements of the partition.

Now, it is important to notice that for any partition $\mathcal{P} = \{V(\Omega_1), V(\Omega_2), \ldots, V(\Omega_{|\mathcal{P}|})\}$, it is possible to build a sampling pattern, locating one sampling node per partition element (see Fig. 2.6). In the following theorem, we show that the spectral characteristics of such sampling patterns, measured by $R_{\boldsymbol{s}}$, are bounded by the local characteristics of the elements in $\mathcal{P}$.

Theorem 6. *Let* $\mathcal{P} = \left\{V(\Omega_1), V(\Omega_2), \ldots, V(\Omega_{|\mathcal{P}|})\right\}$ *a partition of* $V(G)$ *and let* $\boldsymbol{s} \in \{0,1\}^N$ *a sampling pattern chosen according to*

$$\|\boldsymbol{s}(V(\Omega_j))\|_0 = 1 \quad \forall j \tag{2.19}$$

$$\textit{If } \boldsymbol{s}(v) = 1, \textit{ then } \boldsymbol{s}(u) = 0 \quad \forall u \sim v \tag{2.20}$$

then

$$R_{\boldsymbol{s}} \leq \frac{(\mu_2 + \mu_N)^2(1 - |\mathcal{P}|/N)^2}{4\mu_2\mu_N \min\limits_j \left\{\frac{C_{\delta_j}}{vol(\Omega_j)^{2/\delta_j}}\right\}}. \tag{2.21}$$

If in addition, $\delta = \delta_1 = \ldots = \delta_{|\mathcal{P}|}$ *and* $vol(\Omega) = vol(\Omega_1) = \ldots = vol(\Omega_{|\mathcal{P}|})$*, then*

$$R_{\boldsymbol{s}} \leq \frac{(\mu_2 + \mu_N)^2(1 - |\mathcal{P}|/N)^2 vol(\Omega)^{\frac{2}{\delta}}}{4C_\delta\mu_2\mu_N}. \tag{2.22}$$

Proof: See Appendix 2.9

In order to discuss the meaning and implications of Theorem 6, it is important to mention that eqn. (2.19) and eqn. (2.20) imply that there is one sampling node per element of the partition with $\|\boldsymbol{s}\|_0 = |\mathcal{P}|$, and that there is a minimum interdistance between the sampling nodes in $\boldsymbol{s}$ (see Fig. 2.6). In particular, eqn. (2.20) assures that the sampling points in $\boldsymbol{s}$ are far from the boundaries of the elements of the partition.

Notice that eqn. (2.21) presents a general upper bound for the redness of an arbitrary sampling pattern subject to eqn. (2.19) and eqn. (2.20). Meanwhile, eqn. (2.22) provides a tighter bound that is connected with blue-noise sampling patterns as a consequence of having the elements in the partition with the same volume and the same isoperimetric dimension. In this second case, we see that as the size of the partition, $\mathcal{P}$, increases (and therefore the number of sampling nodes in $\boldsymbol{s}$) $vol(\Omega)$ decreases and so it is the value $R_{\boldsymbol{s}}$, making clear the connection between a uniform vertex spreading of the sampling nodes in $\boldsymbol{s}$ and a low redness. As a consequence, a behavior like the one depicted in Fig. 2.5 (bottom) is expected. It is important to emphasize that this statement is connected to Theorems 4 and 5, where the characteristics of good

partitions for the representation of bandlimited signals is stated. In the case of the traditional halftoning scenario where the problem is modeled on a 2-dimensional grid, the conditions of Theorems 6, 5 and 4 hold and a typical response like the one shown in Fig. 2.4 (bottom) is obtained.

Theorem 6 also implies that there is an upper limit about the number of elements in the partition $\mathcal{P}$ that can be considered, and with that, it comes a limitation in the number of sampling nodes for which these inequalities hold. In particular, for very large values of $|\mathcal{P}|$, eqn. (2.19) and eqn. (2.20) cannot be satisfied. We point out that this does not diminish the quality of the sampling patterns, but instead points out that the relationship between the spectral domain and the vertex domain is not guaranteed to be governed by eqn. (2.21).

2.3.2.1 Spectral Metrics

It is also possible to characterize the spectral properties of binary dither patterns on a graph where we extend the idea of periodograms to graphs such that the GFTs of q realizations of $\boldsymbol{x}$, i.e. $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_q$, are averaged together to form the power spectrum:

$$\mathbf{p}(\ell) = \frac{N}{q} \sum_{i=1}^{q} \frac{\hat{\boldsymbol{x}}_i(\ell)^2}{\|\hat{\boldsymbol{x}}_i\|_2^2} \quad \ell = 2, \ldots, N. \tag{2.23}$$

Notice that the ℓ^{th} component of $\mathbf{p}$ is associated with the ℓ^{th} eigenvalue μ_ℓ. Like its binary halftone counterpart, the GFT of a white-noise sampling pattern is expected to be flat for all μ_ks, and to visualize this power spectra, Fig. 2.1 (bottom) shows an estimate of the power spectra for 100 unique white-noise dither patterns generated on the 2000-node Sensor Network graph with pattern density $d = 0.1$.

2.3.3 Blue-noise Sampling Sets

In the following corollary we state how from a blue-noise sampling pattern a good partition in the sense of Theorem 3 can be obtained.

Corollary 7. *Let $\boldsymbol{s}$ be a blue-noise sampling pattern obtained according to Definition 1, with $\|\boldsymbol{s}\|_0 = m$ and $\text{supp}(\boldsymbol{s}) = \{s_1, s_2, \ldots, s_m\}$. Let $B(s_j, \lambda)$ be as specified in Definition 1. Then, there exists a partition $\mathcal{P} = \{V(\Omega_1), V(\Omega_2), \ldots, V(\Omega_{|\mathcal{P}|})\}$ of $V(G)$ such that*

$$u \in V(\Omega_j) \Leftrightarrow u \in B(s_j, \lambda), u \notin B(s_i, \lambda) \ \forall \ i \neq j \tag{2.24}$$

and the elements in the intersection between the sets $B(s_i, \lambda)$ are distributed on the $V(\Omega_i)$ such that the quantity $\sum_{i \neq j} |vol(\Omega_i) - vol(\Omega_j)|$ is minimized. Additionally if $\Lambda_{\mathcal{P}} > (1 + 1/\alpha)\omega$, $\alpha > 0$ any $\boldsymbol{x} \in PW_\omega(G)$ can be uniquely determined from its values at $\{s_1, s_2, \ldots, s_m\}$ always that $\boldsymbol{x}(s_j) = \boldsymbol{x}^\mathsf{T}(\boldsymbol{\xi}_j \circ \boldsymbol{\xi}_j) \quad \forall j$.

Proof: See Appendix 2.11.

This corollary indicates that given a sampling pattern whose sampling points are located as far as possible from each other, it is possible to build a partition from which a unique representation of a set of bandlimited signals is possible. Additonally, if the conditions of Theorem 4 are satisfied, then the partitions obtained are the ones that maximize the value of $\Lambda_{\mathcal{P}}$.

Theorem 1 tells us that when a fixed value of the bandwidth ω is considered and a signal has to be sampled taking m samples, it is necessary to look for the set of nodes, S, such that S^c is a Λ_{S^c}-removable set with $\omega < \Lambda_{S^c}$. Finding the subset of m nodes with the maximum Λ_{S^c} would, therefore, give the *best sampling set*. On the other hand, if a signal has to be sampled taking a number of m samples, choosing a set of nodes S with the maximum value of Λ_{S^c} will extend the class of signals, $PW_\omega(G)$, that can be sampled and represented in a unique way with m samples.

Now if one can show that minimizing the redness in a sampling signal promotes high values of Λ_{S^c}, one could argue blue-noise was a desirable attribute for efficient sampling. The following theorem establishes this relationship:

Theorem 8. *Let $\boldsymbol{s} : V(G) \longrightarrow \{0, 1\}$ be a sampling pattern with $\boldsymbol{s}(S) = 1$, $\boldsymbol{s}(S^c) = 0$ for $S \subset V(G)$ and $|S| = \|\boldsymbol{s}\|_0 = m$, then the $\Lambda_{S^c}-$constant of the set S^c satisfies*

$$\Lambda_{S^c} > C_\delta \left(\frac{R_{\boldsymbol{s}}}{vol(G) R_{\boldsymbol{s}} - m \left(1 - \frac{m}{N}\right)^2} \right)^{\frac{2}{\delta}} \tag{2.25}$$

where $R_{\boldsymbol{s}}$ is the redness in $\boldsymbol{s}$ from eqn. (2.15); δ is the isoperimetric dimension of G [27, 44]; and C_δ a constant that depends only on δ.

Proof: See Appendix 2.12.

To summarize, Theorem 8 tells us that the best sampling set, S, is the one for which the value of Λ_{S^c} is a maximum; therefore while blue-noise sampling patterns (which minimize $R_{\boldsymbol{s}}$) are not necessarily the *best* sampling sets, they are *good* sampling sets. Notice that eqn. (2.25) is well defined as $vol(G)R_{\boldsymbol{s}} - m\left(1 - \frac{m}{N}\right)^2 > 0$, which is tight when $S^c \cup bS^c = V(G)$ where bS^c is the boundary of S^c. This criteria can be satisfied making the nodes in S as spread apart as possible in the graph, which is reasonable as a sampling set where all the nodes are too concentrated in one area could lead to poor reconstructions of signals that exhibit fast changes in the sparsely sampled areas left elsewhere.

As an approach to reinforce the benefits of blue-noise sampling sets, we can use the quantities introduced in Theorem 2 to show how blue-noise promotes those sampling sets that maximize the bandwidth of the signals that can be represented in a unique way on a given sampling set as indicated in the following theorem:

Theorem 9. *Let $\boldsymbol{s} : V(G) \longrightarrow \{0, 1\}$ with $\boldsymbol{s}(S) = 1$, $\boldsymbol{s}(S^c) = 0$, $S \subset V(G)$. If $K_S > 0$, then*

$$K_S \geq \left(\frac{m\left(1 - \frac{m}{N}\right)^2}{R_{\boldsymbol{s}}} - \gamma \right)^{1/2} \tag{2.26}$$

where $\gamma = \max_{S,v',v} \left(\sum_{v \in S^c \setminus v'} w_S(v)^2 \right)$, and $R_{\boldsymbol{s}}$ is, again, the redness in $\boldsymbol{s}$ from eqn. (2.15).

Proof: See Appendix 2.13.

Theorem 9 indicates that lowering the redness of the sampling pattern raises the minimum possible value of K_S and, therefore, extends the set of signals, $PW_\omega(G)$, that can be represented in a unique way on a given sampling set. Therefore, again the blue-noise sampling patterns that are characterized by small values of $R_{\boldsymbol{s}}$ represent a better option than arbitrary random sampling, which leads to large values of $R_{\boldsymbol{s}}$.

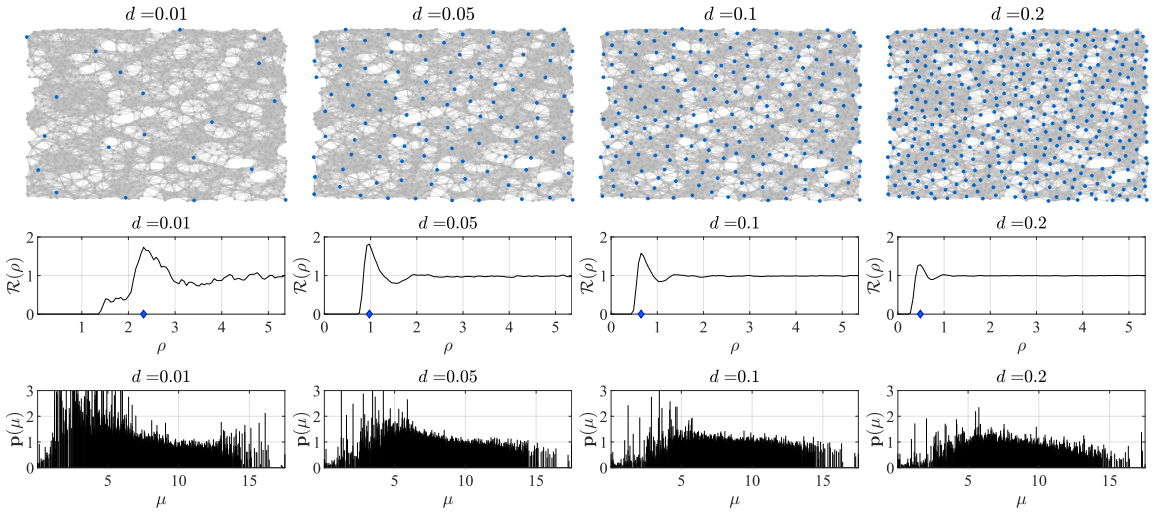


Figure 2.7: Void and cluster blue-noise sampling patterns for different intensities d for a sensor network graph. First row: Localization on the graph of the nodes selected in a blue-noise sampling pattern. Second row: The pair correlation function $\mathcal{R}(\rho)$ for the sampling patterns indicating with a diamond marker the value of λ_b. Third row: Power spectral density for the different blue-noise sampling patterns.

It is important to point out that, under the conditions stated in Theorem 4, Theorems 8 and 9 show that the reduction of the redness is a desirable attribute for any sampling pattern, which is something that can be considered with other sampling approaches. Additionally, the tightness of the inequalities depends on the graph structure which makes these results stronger in some families of graphs.

2.3.4 Stability and blue-noise sampling Sets

The selection of a sampling set, S, is not only associated to a possible unique representation of a signal but also to the stability of its reconstruction when the samples are corrupted by noise, or when the signal considered is not exactly bandlimited. This stability can be measured considering the condition of the matrix $\mathbf{U}_k(S,:)$, which is the matrix obtained by sampling $\mathbf{U}_k$ on the rows indicated by S [31]. Several cost functions can be formulated in terms of the condition of $\mathbf{U}_k(S,:)$, such that their minimum or maximum values are given by the sampling set that provides the best condition [31].

The measures of stability provided in [31] can be equivalently obtained from

a formal general definition of stability for sampling sets in arbitrary spaces [4]. In particular, recalling the definition of stability presented in [4] for general sampling schemes, we can say that $PW_\omega(G)$ posses a stable sampling expansion or reconstruction on $S \subset V(G)$ if there exists $C \geq 1$ such that $\|\boldsymbol{x}(S^c)\|_2^2 \leq (C-1)\|\boldsymbol{x}(S)\|_2^2 \quad \forall \boldsymbol{x} \in PW_\omega(G)$. The value of C provides a measure of stability associated to S; the larger the value of C the less stability we have. In the following theorem we provide an estimate of $C-1$ in terms of Λ_{S^c}.

Theorem 10. *Let $\boldsymbol{x} \in PW_\omega(G)$. Then, if $\Lambda_{S^c} > \tilde{\omega}$, it follows that*

$$\|\boldsymbol{x}(S^c)\|_2^2 \leq \frac{\left(\frac{\tilde{\omega}^r}{\Lambda_{S^c}}\right)^2}{1-\left(\frac{\tilde{\omega}^r}{\Lambda_{S^c}}\right)^2}\|\boldsymbol{x}(S)\|_2^2 \quad r \in \mathbb{R}_+ \tag{2.27}$$

where $\tilde{\omega}$ is the bandwidth of $\boldsymbol{x}_1$, $\boldsymbol{x}_1(S^c) = \boldsymbol{x}(S^c)$ and $\boldsymbol{x}_1(S) = 0$.

Proof: See Appendix 2.14

From this theorem, it is important to point out that finding the sampling set S for which Λ_{S^c} is maximum not only provides a unique representation of a bandlimited signal, but also provides the sampling set in which the highest stability is achieved. This result is consistent with the findings in [31].

As it was stated in Theorem 8, patterns with a low redness promote large values of Λ_{S^c}, therefore blue-noise sampling patterns not only promote uniqueness of the representation but also stability in the reconstruction.

2.3.5 Connection with other works

The implications and properties of spreading the sampling points as far as possible from each other on non Euclidean domains were formally established by Pesenson in [5] considering functions on compact Riemmanian manifolds. In [45–47] the concept of blue-noise was used for the sampling of surfaces embeded in $\mathbb{R}^3$ for applications in computer graphics. This last result can be considered an application of the results in [5] for two-dimensional manifolds embedded in $\mathbb{R}^3$. It is important to point out that

the results in [45–47] rely on the mapping that can be established between the surface and a subset of the 2-dimensional Euclidean domain, but they do not offer any insight of how to deal with the problem in higher dimensions. In [48] a method is proposed for the optimal location of sensors in Euclidean spaces. Exploiting the concepts of entropy, mutual information and Gaussian processes (GPs), the problem of selecting a subset of sensors among a predefined set of discrete positions, defined on a grid of an $n-$dimensional Euclidean space, is addressed. To deal with a large number of possible sensor locations some relaxations based on *lazy evaluations* and local structure of (GPs) are used.

In a different context, and before the emergence of graph signal processing, functions defined on the vertices of a graph have been considered under the concept of *fitness landscapes* [30, 49], which were introduced as a tool for the study of molecular evolution. In this context, the *length* of the autocorrelation function has been useful for the analysis of a landscape. In particular, the *correlation length* of a landscape, $\boldsymbol{x}$, on a K-regular graph is given by [49]

$$\ell_{\boldsymbol{x}} = \frac{K}{\|\hat{\boldsymbol{x}}\|_2^2} \sum_{\ell=2}^{N} \frac{\hat{\boldsymbol{x}}(\ell)^2}{\mu_\ell}. \tag{2.28}$$

The values of eqn. (2.28) provide an indication about how correlated are a given set of samples of the landscape obtained using a random walk. As can be observed in eqn. (2.28) this is proportional to the redness of $\boldsymbol{x}$. In this context, it is possible to conceive blue-noise sampling patterns on graphs as landscapes with a low length correlation.

2.4 Generating Blue-Noise Sampling Sets

Given that blue-noise graph signal sampling promotes the finding of good sampling sets, it is natural to ask how such sampling patterns can be generated. An algorithm that has been particularly successful in digital halftoning and that intuitively translates to graphs is the Void-And-Cluster (VAC) algorithm, introduced by

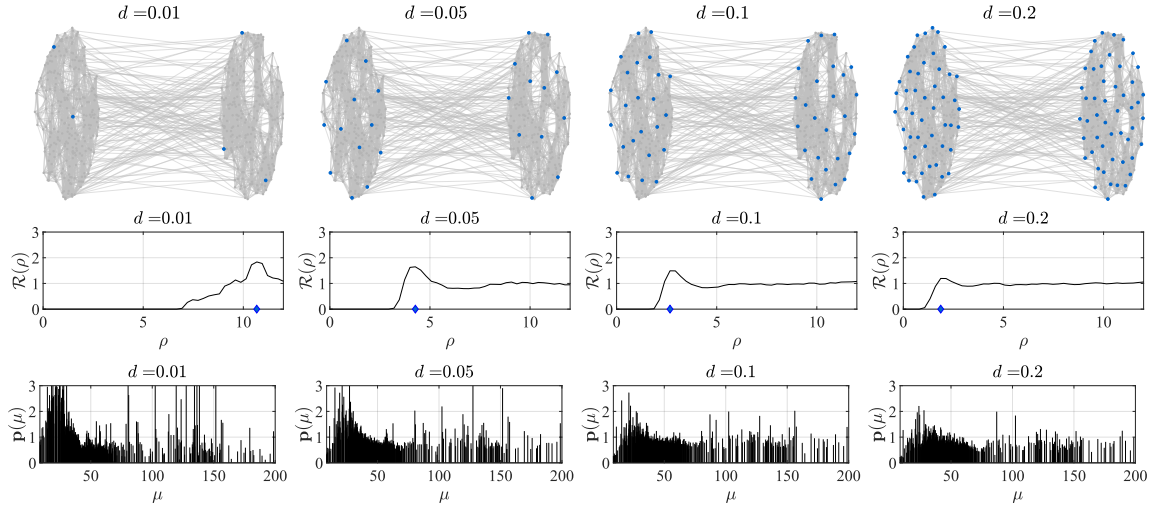


Figure 2.8: Void and cluster blue-noise sampling patterns for different intensities d for a community graph. First row: Localization on the graph of the nodes selected in a blue-noise sampling pattern. Second row: The pair correlation function $\mathcal{R}(\rho)$ for the sampling patterns indicating with a diamond marker the value of λ_b. Third row: Power spectral density for the different blue-noise sampling patterns.

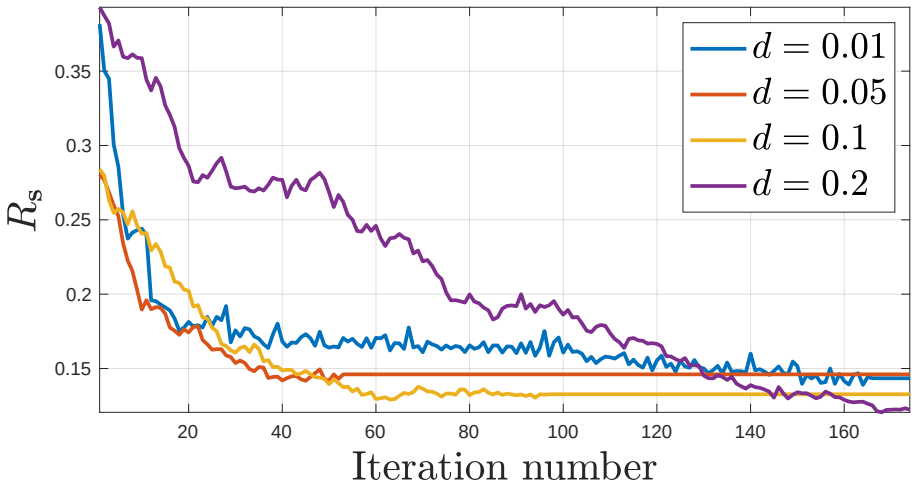


Figure 2.9: Illustration of the redness, $R_{\boldsymbol{s}} = \frac{1}{m}\sum_{\ell=2}^{N} \frac{\hat{\boldsymbol{s}}(\ell)^2}{\mu_\ell}$, of the void and cluster blue-noise sampling patterns on a sensor network with $N = 2000$ nodes, considering different densities.

Ulichney [25]. VAC allows for the construction of artifact-free homogeneous dithering patterns by iteratively measuring the concentration of minority pixels in a binary halftone image, using a gaussian low-pass filter, and swapping the minority pixels in the area of highest concentration with the non-minority pixel in the area of lowest concentration. The adaptation of this algorithm to sampling signals on graphs consists roughly speaking of the sequential computation of distances between sampling points in such a way that points with short geodesic distances between them are relocated trying to put them far from each other.

In order to exploit the above principle for the selection of sampling nodes on a graph, a Gaussian kernel $\mathbf{K}(u,v) = exp(-\mathbf{\Gamma}(u,v)^2/\sigma)$ is evaluated on the set of geodesic distances, $\mathbf{\Gamma}$. This provides a new set of distances that can be tuned according to the parameter, σ, where a small value of $\mathbf{\Gamma}(u,v)$ leads to a value of $\mathbf{K}(u,v)$ that is close to unity while a large value of $\mathbf{\Gamma}(u,v)$ leads to a value of $\mathbf{K}(u,v)$ close to zero. As a measure of how homogeneously distributed the sampling points are, the sum of all distances from one node to the others via the kernel $\mathbf{K}$ is calculated as $\mathbf{c} = \mathbf{K}\mathbf{1}_{N\times 1}$. With this, an initial sampling pattern is generated selecting the m components of $\mathbf{c}$ at random, where $m = dN$ is the number of 1's in the sampling pattern with density d.

The components of $\mathbf{c}$ whose index is given by the location of the 1's in $\boldsymbol{s}$, are then updated to be $\mathbf{c}(\text{supp}(\boldsymbol{s})) = \sum \mathbf{K}(\text{supp}(\boldsymbol{s}), \text{supp}(\boldsymbol{s}))$, where $\sum \mathbf{K}(A,B)$ is defined by

$$\sum \mathbf{K}(A,B) = \sum_{a_i,b_j} \mathbf{K}(a_i,b_j) \quad a_i \in A, b_j \subset B. \tag{2.29}$$

The remaining components of $\mathbf{c}$ are updated according to $\mathbf{c}(\text{supp}(\boldsymbol{s})^c) = \sum \mathbf{K}(\text{supp}(\boldsymbol{s}), \text{supp}(\boldsymbol{s})^c) - \tau$, where τ is selected as a large scalar value. With this update the distances between sampling points in the pattern are represented as positive quantities without adding the distances to other nodes. The distance between $\text{supp}(\boldsymbol{s})$ and $\text{supp}(\boldsymbol{s})^c$ is then represented with a negative value. Now the index of the component of $\mathbf{c}$ with the highest value will indicate the sampling point that is closest to the other sampling points, and then the value of $\boldsymbol{s}$ at that index is forced to be 0 whereas in the index

Algorithm 1 Void and cluster algorithm for graphs

Input: m: number of samples, σ, `NumIter`.
Output: $\boldsymbol{s}$: sampling pattern

Initialisation : $\boldsymbol{s} = \mathbf{0}$, `IndA=-1`, `IndB=-1`.
Calculate $\mathbf{K}(i,j) = e^{-\frac{\mathbf{\Gamma}(i,j)^2}{\sigma}}$ for all $1 \leq i, j \leq N$.
$\mathbf{c} = \mathbf{K}\mathbf{1}_{N\times 1}$.
Get $\mathcal{M}$ as m nodes selected at random.
$\boldsymbol{s}(\mathcal{M}) = 1$.
for $r = 1:1:$ `NumIter` **do**
  $\mathbf{c}(\text{supp}(\boldsymbol{s})) = \sum \mathbf{K}(\text{supp}(\boldsymbol{s}), \text{supp}(\boldsymbol{s}))$.
  $\mathbf{c}(\text{supp}(\boldsymbol{s})^c) = \sum \mathbf{K}(\text{supp}(\boldsymbol{s}), \text{supp}(\boldsymbol{s})^c) - \tau$.
  $\boldsymbol{s}\left(\arg\max_i\{\mathbf{c}(i)\}\right) = 0$.
  $\boldsymbol{s}\left(\arg\min_i\{\mathbf{c}(i)\}\right) = 1$.
  if `IndA`=$\arg\max_i\{\mathbf{c}(i)\}$ and `IndB`=$\arg\min_i\{\mathbf{c}(i)\}$ **then**
    break
  else
    `IndA`=$\arg\min_i\{\mathbf{c}(i)\}$.
    `IndB`=$\arg\max_i\{\mathbf{c}(i)\}$.
  end if
end for
return $\boldsymbol{s}$

where $\mathbf{c}$ is minimum, $\boldsymbol{s}$ is forced to be 1. Notice that the role of τ is to make sure that always $\mathbf{c}(\text{supp}(\boldsymbol{s})^c) < 0$ and a variety of values for τ would serve this purpose. Taking into account that $\sum \mathbf{K}(\text{supp}(\boldsymbol{s}), \text{supp}(\boldsymbol{s})^c) \leq N$, it is possible to select τ as any value such that $\tau > N$.

Repeating the above process iteratively, it is possible to achieve a sampling pattern with no clusters of 1s that exhibits a homogeneous distribution on $V(G)$. The details of the VAC algorithm can be appreciated in Algorithm 1 with example sampling patterns using VAC depicted in Fig. 2.7 for the Sensor Network graph and in Fig. 2.8 for a community graph. From observation, one can see a clear distinction with respect to random sampling when it comes to the nodes distribution of the sampling set. The spatial and spectral blue-noise-like behavior is obtained as a byproduct of the algorithm.

At this point, we note that the value of σ in the kernel $exp(-\mathbf{\Gamma}(u,v)^2/\sigma)$ plays a

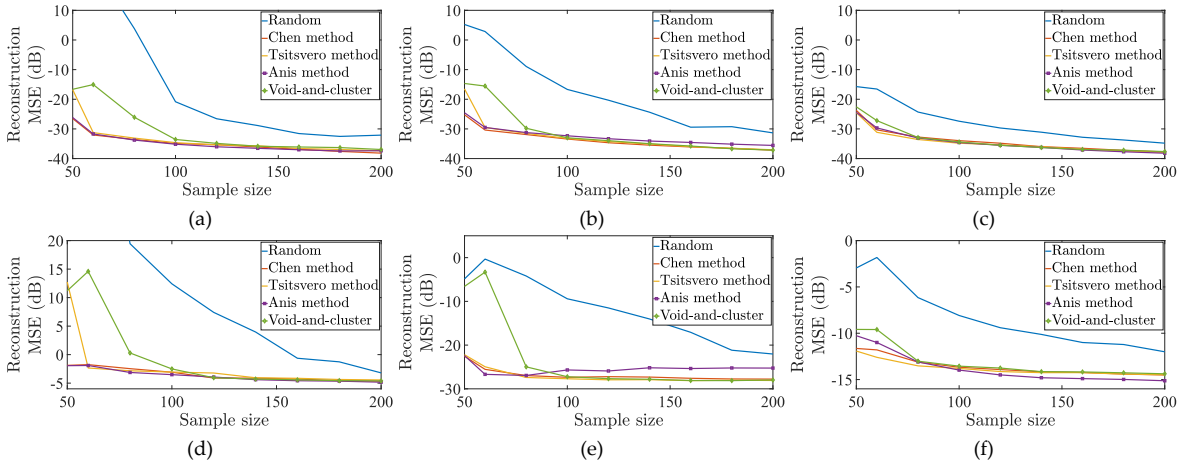


Figure 2.10: Averaged MSE using the reconstruction stated in (2.4) vs the sampling rate considering the reconstruction of 100 different signals from its samples using several sampling schemes and considering several graphs: (a) The graph G_1 and the signal model SM1. (b) The graph G_2 and the signal model SM1. (c) The graph G_3 and the signal model SM1. (d) The graph G_1 and the signal model SM2. (e) The graph G_2 and the signal model SM2. (f) The graph G_3 and the signal model SM2.

critical role in VAC as it defines which sampling nodes are close enough to another one in order to produce a relocation of the 1's in the sampling pattern. Taking into account the definition of λ_b presented in previous sections, it is possible to establish a natural connection between σ and λ_b. In order to do so, we note that if the blue-noise sampling pattern is ideally distributed on the set of nodes, $V(G)$, then when u and v are sampling nodes it follows that $exp(-\mathbf{\Gamma}(u,v)^2/\sigma) \approx 0$ if $\mathbf{\Gamma}(u,v) \geq \lambda_b$. This criteria is considered to be satisfied when $\sigma = \lambda_b^2/ln(10)$, i.e selecting σ in this way the exponential reaches a value of 0.1 when $\mathbf{\Gamma}(u,v) = \lambda_b$. The number of iterations `NumIter` is selected as a multiple of N. In the numerical experiments performed, we have found that choosing `NumIter` $= N$ is enough for the algorithm to reach a stationary behavior. As indicated in Fig. 2.9, there is a clear reduction of the redness of the patterns as they get better distributed on the nodes of the graph. It is important to mention that the number of iterations required for the redness to drop to its minimum value increases as the value of d increases. This is related with the fact that, as d is reduced, there are more possibilities for the relocation of the 1's in the sampling pattern.

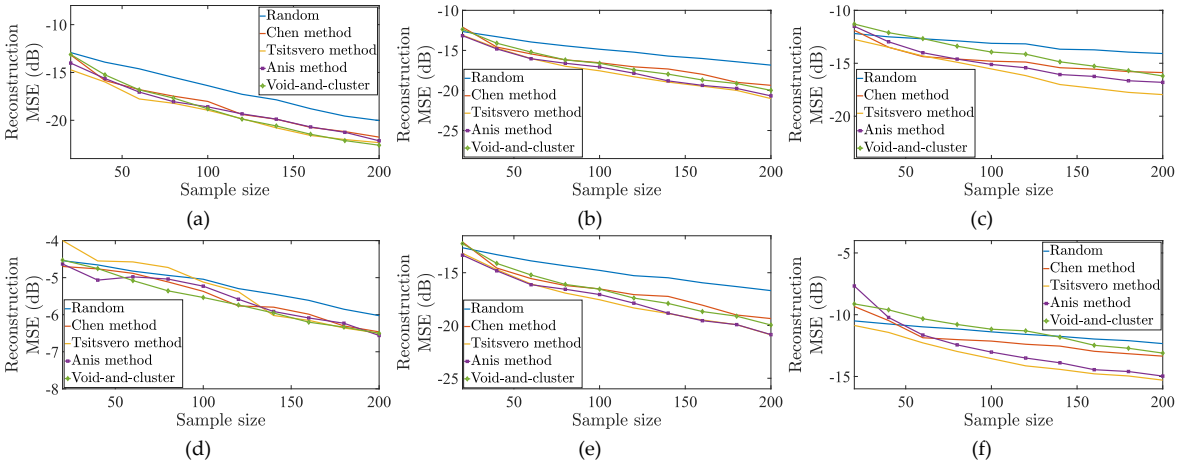


Figure 2.11: Averaged MSE using the reconstruction method proposed in [1] vs the sampling rate considering the reconstruction of 100 different signals from its samples using several sampling schemes and considering several graphs: (a) The graph G_1 and the signal model SM1. (b) The graph G_2 and the signal model SM1. (c) The graph G_3 and the signal model SM1. (d) The graph G_1 and the signal model SM2. (e) The graph G_2 and the signal model SM2. (f) The graph G_3 and the signal model SM2.

2.5 Experiments

In order to evaluate the benefits of blue-noise sampling, a set of numerical experiments is performed comparing the obtained results against state of the art techniques. The simulations are performed considering different graphs and signal models. The experiment is described by the following steps:

- For each graph model, a set of 100 signals is generated according to the specific signal models selected.
- Each signal is sampled by means of different sampling schemes.
- The signal reconstructed from the samples is compared to the original one, and its mean squared error (MSE) is calculated.
- The values of the MSE are averaged over 100.

The schemes of sampling considered for the experiment are the following:

- Blue noise sampling by void and cluster.

- Sampling scheme proposed by Chen et. al. [21].
- Sampling scheme proposed by Anis et. al. [31].
- Sampling scheme proposed by Tsitsvero et al. [34].

The signal models are:

- Signal model 1 (SM1): A random signal of bandwidth $k = 50$, where the Fourier coefficients are generated from the Gaussian distribution $\mathcal{N}(1, 0.5^2)$. The samples captured are contaminated with additive Gaussian noise such that the Signal to Noise Ratio is $SNR = 20$dB.
- Signal model 2 (SM2): A random signal with Fourier coefficients generated from the Gaussian distribution $\mathcal{N}(1, 0.5^2)$. This signal is modulated on the spectral axes by $h(\mu)$, where

$$h(\mu) = \begin{cases} 1 & \text{If} \quad \mu \leq \mu_{50} \\ e^{-4(\mu-\mu_{50})} & \text{If} \quad \mu > \mu_{50} \end{cases} \tag{2.30}$$

The graphs considered in the simulations are different from each other in their nature and represent typical graphs that can be found in different scenarios and applications. The graph models used are:

- Graph G_1: A random sensor network with $N = 1000$ nodes. The weights in the graph are given by the Euclidean distance between points. The maximum number of neighbors for each node is 6.
- Graph G_2: A community graph with $N = 1000$ nodes, 16 communities generated using the GSP toolbox [50].
- Graph G_3: A Barabási-Albert random network [11] with $N = 1000$ nodes.

The reconstructions are performed by means of eqn. (2.4) and by the interpolation splines proposed in [1] and implemented in [50]. In Figs. 2.10 and 2.11, the performance of different algorithms can be appreciated including VAC sampling. Notice that the

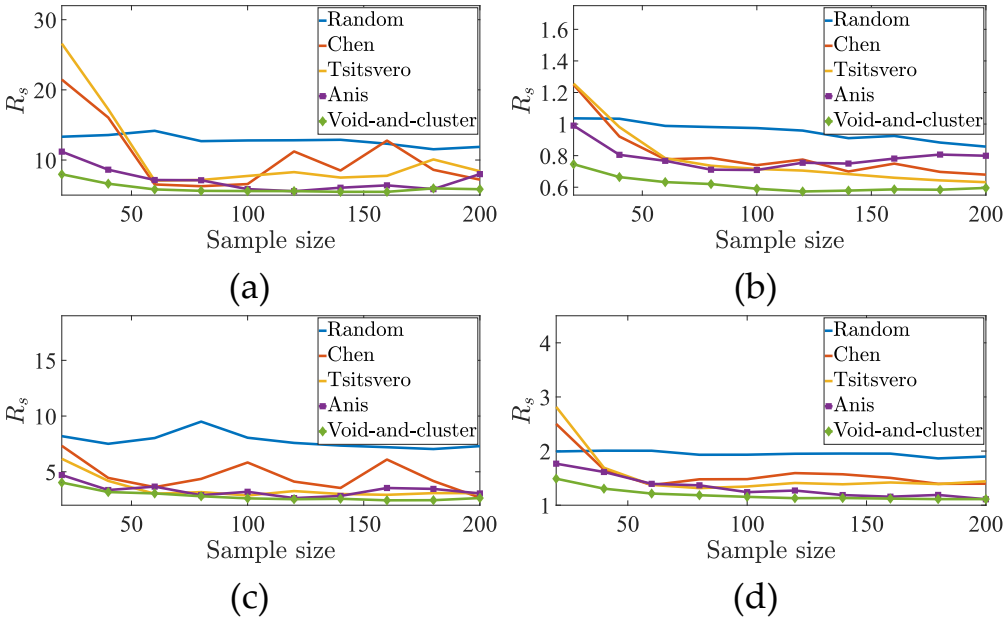


Figure 2.12: Illustration of the redness $R_{\boldsymbol{s}} = \frac{1}{m}\sum_{\ell=2}^{N} \frac{\hat{\boldsymbol{s}}(\ell)^2}{\mu_\ell}$ for the sampling patterns generated by different sampling approaches on different graphs. (a) Swiss roll graph; (b) Sensor network graph; (c) Sphere graph; (d) Bunny graph.

decay rate of the error curves show consistently the benefits of blue-noise sampling. The results obtained using VAC are close to the ones obtained in [31]. Additonally, in Fig. 2.12, the redness of the sampling patterns obtained by different techniques are presented considering different graphs. It is possible to see how low redness is a characteristic attribute of good sampling patterns.

2.6 Conclusion

Blue-noise sampling on graphs is defined based on the traditional blue-noise model associated with digital halftones. The properties and benefits of blue-noise sampling on graphs are linked with theoretical results related to uniqueness sets in sampling, showing why blue-noise patterns promote good sampling sets. We also extended void and cluster, a popular halftoning scheme to generating blue-noise sampling sets on graphs. Numerical tests on different graphs corroborate the good qualities of sampling with blue-noise. We specified conditions under which the traditional relationship between vertex-domain spreading and frequency behavior is preserved. We

further note that the results obtained in this work can be extended for specific families of graphs. The delimitation of the properties for the graphs under consideration could lead to sharper bounds and could allow the definition of other quantities extensively used in halftoning, like the principal frequency.

An overlooked benefit to blue-noise sampling is the wealth of computationally efficient algorithms used in digital halftones that can be extended to graph sampling without spectral estimation, namely error-diffusion where the produced halftone patterns conform to the local spectral content of the image to optimally preserve salient features like edges, gradients, flood fills, etc. Also, a very valuable attribute of error-diffusion that is not widely recognized outside the halftoning community is that error-diffusion can be trained to produce arbitrary spectral profiles (blue-noise, green-noise, etc) and even designed to match the dither patterns produced by other means, including ones with high computational complexity [24, 51]. For graph signal sampling, this opens up the possibility of error-diffusion algorithms trained to mimic sampling algorithms based on spectral estimation and vertex-domain characteristics. We consider that an interesting topic for future research would be the analysis and implications of blue-noise sampling on graphs for signals that are bandlimited but not necessarily low pass [52]. This could provide a generalization of the results that were stated in this work.

It is important to point out that blue-noise sampling promotes large values of Λ_{S^c}, but there is not a guarantee about reaching the maximum value of Λ_{S^c}. For this reason the stability is affected when the value of Λ_{S^c} is not large enough, which also happens when m is not large enough. This aspect is something that can be improved in future works adding additional constraints to the method used to generate blue-noise sampling patterns.

2.7 Proof of Theorem 4

Proof. As stated in [44] (page 168), by means of Sobolev inequalities, it is possible to state that

$$\mu_{1,j} \geq C_{\delta_j} \frac{1}{vol(\Omega_j)^{\frac{2}{\delta_j}}} \tag{2.31}$$

where δ_j is the isoperimetric dimension of Ω_j and C_{δ_j} is a constant that depends only on δ_j. Taking into account the definition of $\Lambda_{\mathcal{P}}$ in Theorem 3 we do have that

$$\Lambda_{\mathcal{P}} > \min \left\{ \frac{C_{\delta_1}}{vol(\Omega_1)^{\frac{2}{\delta_1}}}, \ldots, \frac{C_{\delta_{|\mathcal{P}|}}}{vol(\Omega_{|\mathcal{P}|})^{\frac{2}{\delta_{|\mathcal{P}|}}}} \right\}$$

if $\delta_1 = \delta_2 = \ldots = \delta_{|\mathcal{P}|} = \delta$, then it follows that

$$\Lambda_{\mathcal{P}} > \min \left\{ \frac{C_{\delta}}{vol(\Omega_1)^{\frac{2}{\delta}}}, \ldots, \frac{C_{\delta}}{vol(\Omega_{|\mathcal{P}|})^{\frac{2}{\delta}}} \right\}$$

□

2.8 Proof of Theorem 5

Proof. In order to simplify notation let us represent $x_i = vol(\Omega_i)^{2/\delta}/C_\delta$ and let us consider the optimization problem

$$\begin{aligned} &\underset{\{x_1,x_2,\ldots,x_{|\mathcal{P}|}\}}{\text{maximize}} \quad \min\{1/x_1, 1/x_2, \ldots, 1/x_{|\mathcal{P}|}\} \\ &\text{subject to} \quad \sum_{i=1}^{|\mathcal{P}|} x_i = c_1, \quad x_i > 0 \ \forall i \end{aligned} \tag{2.32}$$

where c_1 is a constant. Now, taking into account that

$$\min\{1/x_1, 1/x_2, \ldots, 1/x_{|\mathcal{P}|}\} \sum_{i=1}^{|\mathcal{P}|} x_i \leq \sum_{i=1}^{|\mathcal{P}|} x_i \frac{1}{x_i} = |\mathcal{P}|$$

$$\min\{1/x_1, 1/x_2, \ldots, 1/x_{|\mathcal{P}|}\} \leq |\mathcal{P}|/c_1.$$

Then, the maximum value of the objective function in eqn. (2.32) is $|\mathcal{P}|/c_1$. Let $(x_1^*, x_2^*, \ldots, x_{|\mathcal{P}|}^*)$ the optimal solution of (2.32), then it follows that $|\mathcal{P}|/c_1 \leq 1/x_i^*$.

Let us assume there exists a subset of indexes $\{j_1, j_2, \ldots, j_q\} \subset \{1, \ldots, |\mathcal{P}|\}$ such that $|\mathcal{P}|/c_1 < 1/x^*_{j_r}$ which implies $x^*_{j_r} < c_1/|\mathcal{P}|$. Then it follows that

$$\sum_{i=1}^{|\mathcal{P}|} x_i = (|\mathcal{P}| - q)\,\frac{c_1}{|\mathcal{P}|} + \sum_{j_r} x_{j_r} < c_1 \tag{2.33}$$

which is a contradiction. Therefore $x^*_i = c_1/|\mathcal{P}|$ which implies $x^*_1 = x^*_2 = \ldots = x^*_{|\mathcal{P}|}$. □

2.9 Proof of Theorem 6

Proof. Let $\boldsymbol{s} \in \{0,1\}^N$ selected according to (2.19) and (2.20), then it follows that $\boldsymbol{s}^\mathsf{T}\mathbf{L}\boldsymbol{s} = \sum_{j=1}^{|\mathcal{P}|} \boldsymbol{s}(V(\Omega_j))^\mathsf{T}\mathbf{L}_{\Omega_j}\boldsymbol{s}(V(\Omega_j))$. Additionally, directly from the definition of $\mu_{1,j}$ and using the Raleyigth coefficient we have $\mu_{1,j} \leq \boldsymbol{s}(V(\Omega_j))^\mathsf{T}\mathbf{L}_{\Omega_j}\boldsymbol{s}(V(\Omega_j))$. Therefore

$$\sum_{j=1}^{|\mathcal{P}|} \mu_{1,j} \leq \sum_{j=1}^{|\mathcal{P}|} \boldsymbol{s}(V(\Omega_j))^\mathsf{T}\mathbf{L}_{\Omega_j}\boldsymbol{s}(V(\Omega_j)) = \boldsymbol{s}^\mathsf{T}\mathbf{L}\boldsymbol{s}. \tag{2.34}$$

Now, taking into account eqn. (2.31) we have

$$|\mathcal{P}| \min_j \frac{C_{\delta_j}}{vol(\Omega_j)^{\frac{2}{\delta_j}}} \leq \sum_{j=1}^{|\mathcal{P}|} \mu_{1,j} \leq \boldsymbol{s}^\mathsf{T}\mathbf{L}\boldsymbol{s} = \sum_{\ell=2}^{N} \mu_\ell \hat{\boldsymbol{s}}(\ell)^2 \tag{2.35}$$

and using lemma 11, we obtain

$$R_{\boldsymbol{s}} \leq \frac{(\mu_2 + \mu_N)^2(1 - |\mathcal{P}|/N)^2}{4\mu_2\mu_N \min_j \left\{ \frac{C_{\delta_j}}{vol(\Omega_j)^{\frac{2}{\delta_j}}} \right\}}.$$

Now, when $\delta = \delta_1 = \ldots = \delta_{|\mathcal{P}|}$, it follows that

$$\frac{C_\delta |\mathcal{P}|}{vol(\Omega)^{\frac{2}{\delta}}} \leq \sum_{j=1}^{|\mathcal{P}|} \mu_{1,j} \leq \boldsymbol{s}^\mathsf{T}\mathbf{L}\boldsymbol{s} = \sum_{\ell=2}^{N} \mu_\ell \hat{\boldsymbol{s}}(\ell)^2 \tag{2.36}$$

and from lemma 11, it follows that

$$R_{\boldsymbol{s}} \leq \frac{(\mu_2 + \mu_N)^2(1 - |\mathcal{P}|/N)^2 vol(\Omega)^{\frac{2}{\delta}}}{4C_\delta\mu_2\mu_N}.$$

□

2.10 Redness inequality

In this section an important and useful lemma used in several proofs is stated.

Lemma 11. *For any sampling pattern $\boldsymbol{s} : V(G) \to \{0,1\}$, it follows that*

$$\frac{m\left(1-\frac{m}{N}\right)^2}{\sum_{\ell=2}\mu_\ell \hat{\boldsymbol{s}}(\ell)^2} \leq R_{\boldsymbol{s}} \leq \frac{m(\mu_2+\mu_N)^2\left(1-\frac{m}{N}\right)^2}{4\mu_2\mu_N\sum_{\ell=2}\mu_\ell \hat{\boldsymbol{s}}(\ell)^2} \tag{2.37}$$

Proof. By Cauchy inequality we know that

$$m^2\left(1-\frac{m}{N}\right)^2 = \left(\sum_{\ell=2}^{N}\sqrt{\mu_\ell}\hat{\boldsymbol{s}}(\ell)\frac{1}{\sqrt{\mu_\ell}}\hat{\boldsymbol{s}}(\ell)\right)^2 \leq \left(\sum_{\ell=2}^{N}\mu_\ell\hat{\boldsymbol{s}}(\ell)^2\right)\left(\sum_{\ell=2}^{N}\frac{1}{\mu_\ell}\hat{\boldsymbol{s}}(\ell)^2\right) \tag{2.38}$$

Now, as indicated in [53] when $\mu_\ell\hat{\boldsymbol{s}}(\ell) > 0$ for all ℓ we have that

$$\left(\sum_{\ell=2}^{N}\mu_\ell\hat{\boldsymbol{s}}(\ell)^2\right)\left(\sum_{\ell=2}^{N}\frac{1}{\mu_\ell}\hat{\boldsymbol{s}}(\ell)^2\right) \leq \left(\frac{\alpha+\beta}{2\sqrt{\alpha\beta}}\right)^2\left(\sum_{\ell=2}^{N}\sqrt{\mu_\ell}\hat{\boldsymbol{s}}(\ell)\frac{1}{\sqrt{\mu_\ell}}\hat{\boldsymbol{s}}(\ell)\right)^2$$
$$= \left(\frac{\alpha+\beta}{2\sqrt{\alpha\beta}}\right)^2 m^2\left(1-\frac{m}{N}\right)^2$$

with $0 < \alpha \leq \mu_\ell \leq \beta$. Then, with $\alpha = \mu_2$ and $\beta = \mu_N$ it follows that

$$\frac{4\mu_2\mu_N}{(\mu_2+\mu_N)^2}\left(\sum_{\ell=2}^{N}\mu_\ell\hat{\boldsymbol{s}}(\ell)^2\right)\left(\sum_{\ell=2}^{N}\frac{1}{\mu_\ell}\hat{\boldsymbol{s}}(\ell)^2\right) \leq m^2\left(1-\frac{m}{N}\right)^2 \tag{2.39}$$

combining eqn. (2.38) and eqn. (2.39) we obtain

$$\frac{m\left(1-\frac{m}{N}\right)^2}{\sum_{\ell=2}\mu_\ell\hat{\boldsymbol{s}}(\ell)^2} \leq \sum_{\ell=2}^{N}\frac{\hat{\boldsymbol{s}}(\ell)^2}{m\mu_\ell} \leq \frac{(\mu_2+\mu_N)^2\left(1-\frac{m}{N}\right)^2 m}{4\mu_2\mu_N\sum_{\ell=2}\mu_\ell\hat{\boldsymbol{s}}(\ell)^2} \tag{2.40}$$

□

2.11 Proof of Corollary 7

Proof. The first part of the proof follows directly from the Definition 1. Now, let us assume $\Lambda_{\mathcal{P}} > (1+1/\alpha)\omega$ with $\alpha > 0$, then according to Theorem 3 any signal $\boldsymbol{x} \in PW_\omega(G)$ is uniquely determined by the values $\boldsymbol{x}^\mathsf{T}\boldsymbol{\xi}_j = \sqrt{|V(\Omega_j)|}\boldsymbol{x}^\mathsf{T}(\boldsymbol{\xi}_j \circ \boldsymbol{\xi}_j)$, therefore if $\boldsymbol{x}(s_j) = \boldsymbol{x}^\mathsf{T}(\boldsymbol{\xi}_j \circ \boldsymbol{\xi}_j)$, $\boldsymbol{x}$ is uniquely determined from $\boldsymbol{x}(s_i)$. □

2.12 Proof of Theorem 8

In order to prove Theorem 8, some preliminary lemmas and theorems are discussed.

Lemma 12. *For any subset of nodes $S \subset V(G)$ and sampling pattern $\boldsymbol{s} \in \{0,1\}^N$ with* $\text{supp}(\boldsymbol{s}) = S$*, it follows that*

$$vol(S) \geq \frac{m^2 \left(1 - \frac{m}{N}\right)^2}{\sum_{\ell=2}^{N} \frac{1}{\mu_\ell} \hat{\boldsymbol{s}}(\ell)^2} \tag{2.41}$$

where $m = \|\boldsymbol{s}\|_0 = |S|$.

Proof. Let us consider the Laplacian matrix $\mathbf{L}$. Multiplying on the left by $\boldsymbol{s}^\mathsf{T}$ and on the right hand side by $\boldsymbol{s}$ it follows that $\boldsymbol{s}^\mathsf{T}\mathbf{L}\boldsymbol{s} = \boldsymbol{s}^\mathsf{T}\mathbf{D}\boldsymbol{s} - \boldsymbol{s}^\mathsf{T}\mathbf{W}\boldsymbol{s}$, which leads to $\sum_{\ell=2}^{N} \mu_\ell \hat{\boldsymbol{s}}(\ell)^2 = vol(S) - \boldsymbol{s}^\mathsf{T}\mathbf{W}\boldsymbol{s}$ and therefore $\sum_{\ell=2}^{N} \mu_\ell \hat{\boldsymbol{s}}(\ell)^2 \leq vol(S)$. Now, taking into account the Lemma 11 eqn. 2.41 is obtained. □

2.12.1 Proof of Theorem 8

Proof. Fuhr and Pesenson [27] show that if a subset of nodes $S \subset V(G)$ is removable with constant Λ_S, it follows that $\Lambda_S \geq \mu_D(S)$, where $\mu_D(S)$ is the Dirichlet eigenvalue of the induced subgraph[1] of S. This inequality is tight always that $S \cup bS = V(G)$, where bS is the vertex boundary of S.

As stated in [27], $\mu_D(S)$ satisfy the following inequality

$$\mu_D(S) > C_\delta \left(\frac{1}{vol(S)}\right)^{2/\delta} \tag{2.42}$$

where δ is the isoperimetric dimension of the graph, C_δ is a constant that depends only on δ and $vol(S) = \sum_{v \in S} D(v,v)$.

[1] Definitions and inequalities about induced subgraphs can be found in [44]

Now, taking into account that $vol(G) = vol(S) + vol(S^c)$ for any $S \subset V(G)$, and the lemma 12, we have that

$$vol(G) - vol(S) \leq vol(G) - \frac{m^2\left(1-\frac{m}{N}\right)^2}{\sum\limits_{\ell=2}^{N}\frac{1}{\mu_\ell}\hat{\boldsymbol{s}}(\ell)^2} \tag{2.43}$$

and then

$$C_\delta\left(\frac{1}{vol(G)-vol(S)}\right)^{\frac{2}{\delta}} \geq C_\delta\left(\frac{\sum_{\ell=2}\frac{1}{\mu_\ell}\hat{\boldsymbol{s}}(\ell)^2}{vol(G)\sum_{\ell=2}\frac{1}{\mu_\ell}\hat{\boldsymbol{s}}(\ell)^2 - m^2\left(1-\frac{m}{N}\right)^2}\right)^{\frac{2}{\delta}}. \tag{2.44}$$

Now, taking into account that $\Lambda_{S^c} \geq \mu_D(S^c)$, it follows that

$$\Lambda_{S^c} \geq C_\delta\left(\frac{\sum_{\ell=2}\frac{1}{\mu_\ell}\hat{\boldsymbol{s}}(\ell)^2}{vol(G)\sum_{\ell=2}\frac{1}{\mu_\ell}\hat{\boldsymbol{s}}(\ell)^2 - m^2\left(1-\frac{m}{N}\right)^2}\right)^{\frac{2}{\delta}}$$

□

2.13 Proof of Theorem 9

In this section the proof of Theorem 9 is provided. Before this proof is presented an important lemma is introduced.

Lemma 13. *Let $\boldsymbol{s} : V(G) \longmapsto \{0,1\}^N$ a binary signal defined on $V(G)$ and let $\bar{\boldsymbol{s}} = \mathbf{1} - \boldsymbol{s}$, then it follows that*

$$\sum_{\ell=2}^{N}\mu_\ell\hat{\bar{\boldsymbol{s}}}(\ell)^2 = \sum_{\ell=2}^{N}\mu_\ell\hat{\boldsymbol{s}}(\ell)^2 \tag{2.45}$$

Proof. Let us consider the Laplacian matrix $\mathbf{L}$ and multiply on the left by $\boldsymbol{s}^\mathsf{T}$ and on the right by $\boldsymbol{s}$, it follows that

$$\boldsymbol{s}^\mathsf{T}\mathbf{L}\boldsymbol{s} = (\mathbf{1}-\bar{\boldsymbol{s}})^\mathsf{T}\mathbf{L}(\mathbf{1}-\bar{\boldsymbol{s}}) = \bar{\boldsymbol{s}}^\mathsf{T}\mathbf{L}\bar{\boldsymbol{s}}. \tag{2.46}$$

Now, taking into account that $\boldsymbol{x}^\mathsf{T}\mathbf{L}\boldsymbol{x} = \sum_{\ell=1}^{N}\mu_\ell\hat{\boldsymbol{x}}(\ell)^2$, it follows that

$$\sum_{\ell=2}^{N}\mu_\ell\hat{\bar{\boldsymbol{s}}}(\ell)^2 = \sum_{\ell=2}^{N}\mu_\ell\hat{\boldsymbol{s}}(\ell)^2. \tag{2.47}$$

Notice that $\mu_1 = 0$ and consequently the sum can be computed for $\ell \geq 2$. □

2.13.1 Proof of Theorem 9

Proof. Taking into account that

$$\left(\mathbf{L}\boldsymbol{x}\right)(v) = \sum_{u \in V(G)} \left(\boldsymbol{x}(v) - \boldsymbol{x}(u)\right) \mathbf{W}(v,u) \tag{2.48}$$

and $w_S(v) = \sum_{u \in S} \mathbf{W}(u,v)$. It is possible to infer that

$$\left(\mathbf{L}\bar{\boldsymbol{s}}\right)(v) = \begin{cases} w_S(v) & \text{if} \quad v \in S^c \\ -w_{S^c}(v) & \text{if} \quad v \in S \end{cases} \tag{2.49}$$

where $\bar{\boldsymbol{s}} = \mathbf{1} - \boldsymbol{s}$. Now, taking into account eqn. (2.49) and Lemma 13 it follows that

$$\bar{\boldsymbol{s}}^{\mathsf{T}} \mathbf{L} \bar{\boldsymbol{s}} = \sum_{\ell=2}^{N} \mu_\ell \hat{\bar{\boldsymbol{s}}}(\ell)^2 = \sum_{v \in S^c} w_{S(v)}^2$$

$$\sum_{\ell=2}^{N} \mu_\ell \hat{\boldsymbol{s}}(\ell)^2 = K_S^2 + \sum_{v \in \{S^c \setminus v'\}} w_S(v)^2$$

$$K_S = \left(\sum_{\ell=2}^{N} \mu_\ell \hat{\boldsymbol{s}}(\ell)^2 - \sum_{v \in \{S^c \setminus v'\}} w_S(v)^2 \right)^{\frac{1}{2}}$$

which leads to

$$K_S \geq \left(\sum_{\ell=2}^{N} \mu_\ell \hat{\boldsymbol{s}}(\ell)^2 - \gamma \right)^{\frac{1}{2}}$$

where γ is given by $\gamma = \max_{S,v,v'} \sum_{v \in \{S^c \setminus v'\}} w_S(v)^2$ and taking into account Lemma 11, it follows that

$$K_S \geq \left(\frac{m^2 \left(1 - \frac{m}{N}\right)^2}{\sum_{\ell=2}^{N} \frac{1}{\mu_\ell} \hat{\boldsymbol{s}}(\ell)^2} - \gamma \right)^{\frac{1}{2}}$$

□

2.14 Proof of Theorem 10

Proof. Let us consider $\boldsymbol{x} \in PW_\omega(G)$ written as $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ with $\boldsymbol{x}_1(S^c) = \boldsymbol{x}(S^c)$, $\boldsymbol{x}_1(S) = 0$, $\boldsymbol{x}_2(S) = \boldsymbol{x}(S)$ and $\boldsymbol{x}_2(S^c) = 0$. From eqn. (2.5) we have that $\|\boldsymbol{x}_1\|_2^2 \leq (1/\Lambda_{S^c}^2)\|\mathbf{L}\boldsymbol{x}_1\|^2$. Now, using Bernstein's inequality [27,54] we get $\|\boldsymbol{x}_1\|_2^2 \leq (1/\Lambda_{S^c}^2)\tilde{\omega}^{2r}\|\boldsymbol{x}_1\|_2^2$. From this we have $\|\boldsymbol{x}_1\|_2^2 \leq (1/\Lambda_{S^c}^2)\tilde{\omega}^{2r}\left(\|\boldsymbol{x}_1\|_2^2 + \|\boldsymbol{x}_2\|_2^2\right)$ leading to eqn. (2.27). □

Chapter 3

LOW COMPLEXITY BLUE-NOISE ALGORITHMS

3.1 Introduction

As shown in the previous chapter, blue-noise sampling on graphs relies essentially on the simple principle of spreading up the sampling nodes of a sampling pattern. Having shown that there is a strong connection between the characteristics of these sampling patterns and the theoretical measures of the quality of sampling sets, we state the basis for the development of low computational cost algorithms that can follow these basic principles. The void and cluster algorithm proposed before provides sampling patterns close to the expected ideal blue noise sampling pattern relying on the use of geodesic distances. In this chapter we propose to exploit *random walks* on graphs and the concept of *error diffusion*, extensively used in digital halftoning, in order to generate blue-noise like sampling patterns. The notation and symbols used in this chapter are the same used in Chapter 2.

3.2 Random Walk Sampling

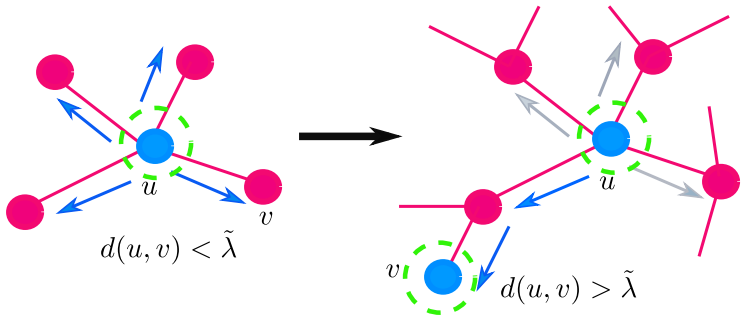


Figure 3.1: Illustration of how a random walk can be use to select sampling points that can be far away an approximate distance of $\tilde{\lambda}_b$.

Algorithm 2 Random Walk Sampling: RWS-1

Input: $\mathbf{W}$, λ_b, `numiter`.
Output: A blue-noise-like sampling pattern, $\boldsymbol{s}$.

```
    Initialisation : s = 0_N, count = 0
    Randomly select v ∈ V(G).
 2: s(v) = 1
    while count < numiter do
 4:    d_t = 0
       while d_t < λ_b do
 6:       Select u ∈ N(v) at random
          d_t = d(u, v) + d_t
 8:       v = u
       end while
10:    count = count + 1
       s(u) = 1
12:    v = u
    end while
14: return s.
```

The central idea in the definition of a blue-noise sampling pattern, is related with the minimum inter-distance between closest points. Ideally, this distance should be λ_b, which is graph dependent. Therefore, a blue-noise-like sampling pattern on the graph can be built jumping from one sampling node, u, to other nodes marking every node, v, that is at a distance greater than a given value λ_b, i.e., $d(v, u) > \lambda_b$. Figure 3.1 illustrates the intuition behind this idea. This approach offers a simple way to mark the sampling nodes at the expense of using a distance between nodes that is not necessarily geodesic, as it would be calculated as the length of a path followed by the random walk. These ideas are formalized in Algorithm 2 that we called *random walk sampling (RWS1)*, where $\mathcal{N}(v)$ indicates the set of nodes that are connected to a given node v.

It is important to remark that because the distance considered in the random walk is not geodesic, a rough estimate of λ_b is enough to run the algorithm and therefore no prior calculation before the sampling set search are needed. Indeed, as it will be shown in later sections, the guarantee of a minimum interdistance is going to be affected by the density of the sampling pattern as well as the size of the graph. However,

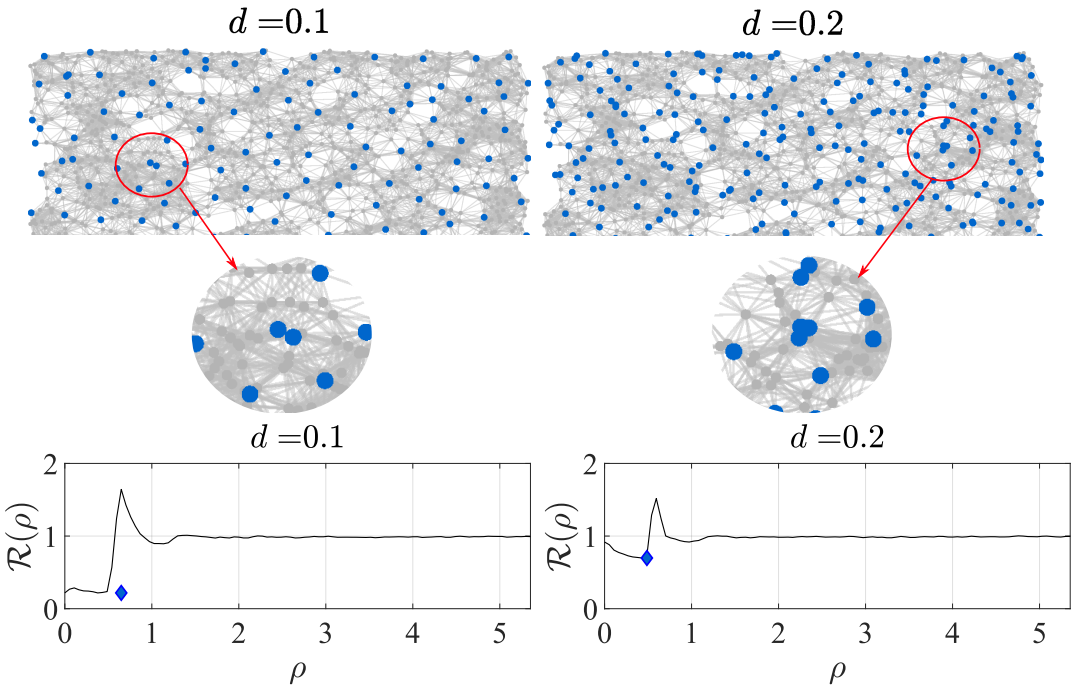


Figure 3.2: Vertex-domain distribution of a blue noise sampling pattern generated by means of Algorithm 2 considering different densities on a sensor network of 2000 nodes. Left: density $d = 0.1$. Right: density $d = 0.2$.

Algorithm 2 offers an extremely low computational cost, $\mathcal{O}(\texttt{numiter}\lceil\frac{\lambda_b}{\bar{w}}\rceil)$, where $\bar{w}$ is the average of nonzero weights in $\mathbf{W}$. Additionally, the results for some densities exhibit characteristics on the vertex domain close to the ideal blue-noise sampling patterns, as can be appreciated in Fig. 3.2 for a density of $d = 0.1$.

3.2.1 Ensuring a minimum distance and a minimum number of sampling nodes

Algorithm 2 exhibits a characteristic that is convenient when it comes to perform computations in large graphs, it requires only knowledge of the *local structure* around one node and no pre-calculations are needed for the search of the sampling set. Additionally, this local information is not stored as the random walk is performed on the graph. As a consequence, it is not possible to guarantee that two sampling nodes are not going to be connected by an edge. It is important to point out that the number

of sampling nodes cannot be controlled in advance. In this subsection we introduce a variant of Algorithm 2 in which we add the following constraints:

1. (**C1**) The sampling nodes cannot be connected by an edge.
2. (**C2**) The information of the nodes *visited* when marking the sampling nodes is stored.
3. (**C3**) The number of sampling nodes is fixed.

The first constraint is ensured as a post processing stage after having performed the random walk on the graph. Line 23 of Algorithm 3 shows how this condition is imposed. It is worth noticing that imposing this constraint will change the number of ones in the sampling pattern. Additionally, for high sampling densities this constraint is not well suited and may not be applied.

The second constraint is imposed taking into account not only the nodes visited in the random walk, but also the neighbors of those points. In this way, we try to maximize the distance between sampling points, and also promote a random walk that is not *concentrated* on one region of the graph. Algorithm 3 shows how this task can be performed by means of sets T, Q and R. The set Q is meant to store the neighbors of those nodes that have not been *visited* on the random walk. In that way, we can promote a random walk that is *spread* enough on the vertex domain. The set R keeps track of the nodes that are marked as sampling nodes, while set T has the cumulative storing of the neighbors locally stored in Q.

The imposition of the first and the second constraint affects the total number of sampling nodes generated by the algorithm. In some applications, the use of Algorithm 2 or Algorithm 3 would be satisfactory, but in some others the number of sampling points is fixed and the optimal distribution of those points has to be obtained.

If the number of sampling points, m, has to be fixed, challenges in two scenarios are faced. On one side we have the case where the number of sampling nodes is larger than the number of desired samples, and on the other side we have the case where

Algorithm 3 Random Walk Sampling: RWS-2

Input: $\mathbf{W}$, λ_b, `numiter`.
Output: A blue-noise-like sampling pattern, $\boldsymbol{s}$.

Initialisation : $\boldsymbol{s} = \mathbf{0}_N$, `count` $= 0$
Randomly select $v \in V(G)$.
$\boldsymbol{s}(v) = 1$
while `count` $<$ `numiter` **do**
  $d_t = 0$, $T = \emptyset$, $R = \emptyset$.
  while $d_t < \lambda_b$ **do**
    $Q = \mathcal{N}(v)$
    $Q = Q \setminus (T \cup R)$
    if $Q = \emptyset$ **then**
      $d_t = \lambda_b$
      $v \in \text{supp}(\boldsymbol{s})$
    else
      $u \in Q$
      $d_t = \mathbf{W}(v, u) + d_t$
      $R = \{v\} \cup R$
      $T = Q \cup T$
      $v = u$
    end if
    $\boldsymbol{s}(u) = 1$
  end while
  `count` $=$ `count` $+ 1$
  $\boldsymbol{s}(u) = 1$
end while
$\boldsymbol{x}(\mathcal{N}(v)) = 0 \quad \forall \boldsymbol{s}(v) = 1$
return $\boldsymbol{s}$.

Algorithm 4 Random Walk Sampling: RWS-3

Input: $\mathbf{W}$, λ_b, `numiter`, m: number of sampling nodes.
Output: A blue-noise-like sampling pattern, $\boldsymbol{s}$.

Initialisation : $\boldsymbol{s} = \mathbf{0}_N$, `count` $= 0$
Randomly select $v \in V(G)$.
$\boldsymbol{s}(v) = 1$
while `count` $<$ `numiter` **do**
  $d_t = 0$, $T = \emptyset$, $R = \emptyset$.
  while $d_t < \lambda_b$ **do**
    $Q = \mathcal{N}(v)$
    $Q = Q \setminus (T \cup R)$
    if $Q = \emptyset$ **then**
      $d_t = \lambda_b$
      $v \in \mathrm{supp}(\boldsymbol{s})$
    else
      $u \in Q$
      $d_t = \mathbf{W}(v, u) + d_t$
      $R = \{v\} \cup R$
      $T = Q \cup T$
      $v = u$
    end if
    $\boldsymbol{s}(u) = 1$
  end while
  `count` $=$ `count` $+ 1$
  $\boldsymbol{s}(u) = 1$
end while
$\boldsymbol{x}(\mathcal{N}(v)) = 0 \quad \forall \boldsymbol{s}(v) = 1$
if $\|\boldsymbol{s}\|_0 > m$ **then**
  Get $S \subset \mathrm{supp}(\boldsymbol{s})$, with $|S| = m$
  $\boldsymbol{s}(S) = 1$, $\boldsymbol{s}(S^c) = 0$
end if
if $\|\boldsymbol{s}\|_0 < m$ **then**
  Get $J = \{v \in \mathrm{supp}(\mathbf{1} - \boldsymbol{s}) : \mathcal{N}(v) \in \mathrm{supp}(\mathbf{1} - \boldsymbol{s})\}$
  $\boldsymbol{s}(\tilde{J}) = 1$, with $\tilde{J} \subset J$, $|\tilde{J}| = \min\{m - \|\boldsymbol{s}\|_0, |J|\}$
end if
while $\|\boldsymbol{s}\|_0 < m$ **do**
  $\boldsymbol{q}(v) = |\mathcal{N}(v) \cap \mathrm{supp}(\boldsymbol{s})^c|$ if $v \in \mathrm{supp}(\boldsymbol{s})$ and $\boldsymbol{q}(v) = 0$ if $v \in S^c$
  Get J, the $m - \|\boldsymbol{s}\|_0$ indexes where $\boldsymbol{q}$ has the largest value
  $U = \cup_{i=1}^{m-\|\boldsymbol{s}\|_0} U_i$, with $U_i = \{u_i\}$ and $u_i \in \mathcal{N}(v)$, $v \in J$
  $\boldsymbol{s}(U) = 1$
end while
return $\boldsymbol{s}$.

the number of selected sampling nodes is lower than m. The solutions we propose to deal with these issues differ substantially for each case. In particular we propose the following solutions:

- **When the number of sampling points is larger than the desired number of samples** m: The impositions of the first and second constraints considered above focus on the spreading of the nodes and try to ensure a minimum interdistance between sampling nodes. As a way to still preserve the overall distribution of the sampling nodes and reduce the number of sampling nodes, we discard the excess uniformly at random.

- **When the number of sampling nodes is lower than the desired number of samples** m: This scenario exhibits more difficulties than the previous one, as the random selection of a set of nodes on the complement set can affect seriously the characteristics and benefits achieved by the imposition of the first (**C1**) and second (**C2**) constraints i.e., the uniform spreading and minimum distance between sampling nodes. To deal with these challenges, we propose a two-stage solution. In the first stage we *fill the voids* identifying nodes are not selected in the sampling pattern and whose neighbors are not selected either. On the second stage, we select at random the nodes that are needed to complete the pattern.

We formalize these ideas in Algorithm 4 that we call *fast blue-noise sampling*.

3.3 Error Diffusion on Graphs

Error diffusion in traditional halftoning represents an efficient way to generate dither patterns and has been extensively used for the generation of blue noise-like sampling patterns [39, 51]. The basic principle in error diffusion relies on the idea of using the error obtained as the difference between a signal at a given pixel and its halftone approximation, to influence the halftone approximation of the signal in other pixels. The way this error is *diffused* is known as the error diffusion algorithm. The application of these ideas to build sampling patterns on graphs is quite natural as

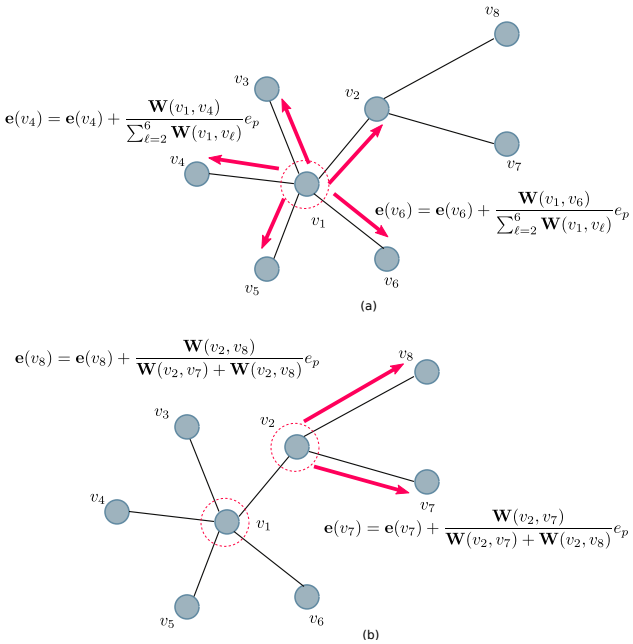


Figure 3.3: Pictorial illustration of how the error is diffused in the Error diffusion algorithm. In (a) the error calculated in v_1 is diffused to all the neighbors of v_1 and in (b) it is illustrated that the diffusion of the error from v_2 is diffused only to those nodes that have not been visited. Notice that the the contribution of the error is normalized by the weights of the edges that connect the actual node to the nodes to which the error is going to be diffused.

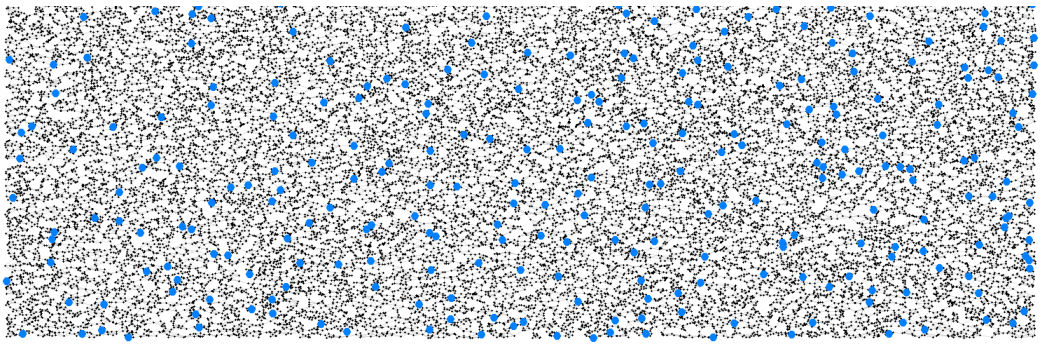


Figure 3.4: Illustration of a Blue-Noise sampling pattern generated on a sensor network with $N = 50000$ nodes.

nodes are the direct analogous of the pixels in an image and the edges define how the interaction between nodes is defined. The set of edges $E(G)$ in the graph offer the structure over which an error can be diffused. The main difference between the diffusion of the error in a regular rectangular grid and the graph will rely on the fact that there is not a standard way to label the nodes in order to *visit* them when it comes to the generation of the diffusion error. However, it will be possible to follow any given labeling of the nodes in the graph and diffuse the error according to that ordering.

Error diffusion on graphs proceeds as follows. Given a constant signal $\boldsymbol{x} = \left(\frac{m}{N}\right)\mathbf{1}_{N\times 1}$ with amplitude $d = m/N$ and following the ordering given by the labeling of the nodes in the graph, an error is generated in the first node as $e_p = \boldsymbol{s}(1) - \boldsymbol{x}(1)$, where $\boldsymbol{s}(1) = 1$ if $\boldsymbol{x}(1) > t_{th}$ or $\boldsymbol{s}(1) = 0$ if $\boldsymbol{x}(1) < t_{th}$. Then, the value of e_p is diffused to the neighbors of the first node, $\mathcal{N}(1)$, as $\boldsymbol{e}(i) = \frac{\mathbf{W}(1,i)e_p}{\sum_{i\in\mathcal{N}(1)}\mathbf{W}(v,i)}, \quad i \in \mathcal{N}(1)$. This process is repeated iteratively following the ordering given by the labels of the nodes in such a way that the error at the node v would be computed as $e_p = \boldsymbol{s}(v) - u$ with $\boldsymbol{s}(v) = 1$ if $u > t_{th}$ or $\boldsymbol{s}(v) = 0$ if $u < t_{th}$ and $u = \boldsymbol{x}(v) - \boldsymbol{e}(v)$, diffusing this error as $\boldsymbol{e}(i) = \boldsymbol{e}(i) + \frac{\mathbf{W}(v,i)e_p}{\sum_{i\in\mathcal{N}(v)}\mathbf{W}(v,i)}$. In this way, the error is cumulated according to the local connections of each node. Notice that this diffusion of the error is completely equivalent to the one performed in traditional halfoning [39]. The only difference is that the concept of directions on $V(G)$ is not defined and therefore the error is diffused as dictated by the neighbors of the nodes that are visited. This represents a substantial difference with its classical counterpart.

The steps and details of this approach are given in Algorithm 5. In Fig. 3.4 a sample of an error diffusion pattern is depicted indicating the distribution of the sampling points on $V(G)$ on a large sensor network graph.

3.4 Computational Complexity

We now compare the computational complexity of the developed algorithms with respect to the state of the art techniques and with respect to the void and cluster algorithm. This complexity is considered in two separated components. The first

Algorithm 5 Error diffusion Algorithm on Graphs

Input: $t_{th} = 0.5, m$

Output: $\boldsymbol{s}$: sampling pattern

Initialisation : $\boldsymbol{x} = \left(\frac{m}{N}\right)\mathbf{1}_{N\times 1}$, $\mathbb{V} = \emptyset$, $\mathbf{e} = \mathbf{0}_{N\times 1}$.

1: **for** $v = 1:1:|\mathcal{V}|$ **do**
2: $u = \boldsymbol{x}(v) - \boldsymbol{e}(v)$
3: **if** $u > t_{th}$ **then**
4: $\boldsymbol{s}(v) = 1$
5: **else**
6: $\boldsymbol{s}(v) = 0$
7: **end if**
8: $e_p = \boldsymbol{s}(v) - u$
9: **for** $i \in \mathcal{N}(v)$ **do**
10: $\boldsymbol{e}(i) = \boldsymbol{e}(i) + \frac{\mathbf{W}(v,i)e_p}{\sum_{i\in\mathcal{N}(v)}\mathbf{W}(v,i)}$
11: **end for**
12: $\mathbb{V} = \mathbb{V} \cup \{v\}$
13: **end for**
14: **return** $\boldsymbol{s}$

component states the calculations that are involved before the *sampling set search* (SS), while the second part states the cost of finding the sampling set assuming the knowledge of some quantities. Table 3.1 illustrates the computational cost of void and cluster (VC), the approach in [31], fast blue noise sampling (RWS-3) and error diffusion.

It is important to point out that approaches like [31] and void and cluster require previous calculations before performing the search of the sampling sets, while error diffusion and RWS-3 do not. This is a key component when it comes to the sampling scenario in huge graphs, where *pre-calculations* cannot be afforded.

Notice that in table 3.1 q is the power of the Laplacian considered in the application of [31], `n` corresponds to `numiter` in Algorithm 4, m is the number of sampling nodes, k is the bandwidth of the signals considered, $\bar{w}$ represents the average of the non-zero elements in $\mathbf{W}$, $\bar{d}_{deg}$ is the average number of neighbors of each node, α is a constant associated to the difference between the desired number of sampling nodes and its effective value and T_1 is the average number of iterations required for the

convergence of a single eigendecomposition pair.

Table 3.1: Computational Complexity.

Case	VC	Anis [31]	RWS-3	Error diffusion
pre-SS	$\mathcal{O}(N(\|E(G)\|+N)\log N)$	$\mathcal{O}(q\|E(G)\|kT_1)$	0	0
SS	$\mathcal{O}((N-1)(m+2))$	$\mathcal{O}(Nk)$	$\mathcal{O}(\mathtt{n}\lceil\frac{\lambda_b}{\bar{w}}\rceil+2N+m-\alpha)$	$\mathcal{O}(N(2+\bar{d}_{deg}))$

3.5 Numerical tests

In order to test the algorithms developed in this chapter in terms of the reconstruction error, we perform a set of numerical experiments considering two main scenarios: medium size graphs and large size graphs. In the first case, we consider graphs with one thousand nodes while in the second case we use graphs with fifty thousand nodes. In the first part of the experiments, we can use the approaches considered in the previous chapter. However, in the second case, it is impossible because of the high computational cost.

3.5.1 Medium Size Graphs

Considering three graphs and two types of signal models a set of sampling and reconstruction simulations is performed. The graphs considered are:

- Graph G_1: A random sensor network with $N = 1000$ nodes. The weights in the graph are given by the Euclidean distance between points. The maximum number of neighbors for each node is 6.

- Graph G_2: A community graph with $N = 1000$ nodes, 16 communities generated using the GSP toolbox [50].

- Graph G_3: A Barabasi-Albert random network [11] with $N = 1000$ nodes.

The signal models considered are

- Signal model 1 (SM1): A random signal of bandwidth $k = 2500$, where the Fourier coefficients are generated from the Gaussian distribution $\mathcal{N}(1, 0.5)$. The samples captured are contaminated with additive Gaussian noise, making the Signal to Noise Ratio to be $SNR = 20dB$.

- Signal model 2 (SM2): A random signal with Fourier coefficients generated from the Gaussian distribution $\mathcal{N}(1, 0.5)$. This signal is modulated on the spectral domain by the function $h(\mu)$, defined as

$$h(\mu) = \begin{cases} 1 & \text{If} \quad \mu \leq \mu_{2500} \\ e^{-4(\mu - \mu_{50})} & \text{If} \quad \mu > \mu_{50} \end{cases} \tag{3.1}$$

The details of these simulations are the following:

- We generate a set of 100 signals for each signal model.

- Each signal is sampled considering several approaches and then reconstructed by means of eqn. (2.4), calculating the mean squared error (MSE).

- The values of MSE are averaged.

The sampling approaches considered for this experiment are

- Uniform random sampling.

- Blue noise sampling by void and cluster.

- Blue noise by error diffusion.

- Blue noise sampling by RWS-3 (fast blue noise).

- Sampling scheme proposed by Chen et. al. [21].

- Sampling scheme proposed by Anis et. al. [31].

- Sampling scheme proposed by Tsitsvero et. al. [34].

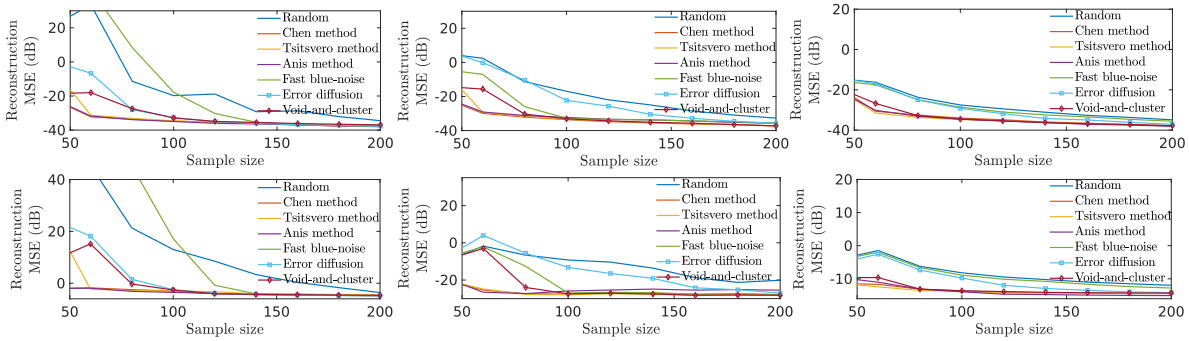


Figure 3.5: Averaged MSE considering different sampling approaches. From left to right we consider the graphs G_1, G_2 and G_3 respectively. First row is associated to the experiments considering the signal model $SM1$, while the second row is associated to $SM2$.

The results of these experiments are given in Fig. 3.5. Notice that the performance exhibited by error diffusion is close to the one obtained by the other approaches in graphs whose local isoperimetric dimension is homogeneous, like the sensor network for instance. However, in graphs like the community type its performance is still not competitive. In the case of the fast blue noise, we observe a poor performance on the sensor network, but a superior behavior on the community graph.

3.5.2 Large size graphs

The benefits of blue-noise sampling on graphs can be better appreciated when it comes to the analysis of signals on large graphs. The numerical evidence on medium size graphs shows that blue-noise can compete with the optimal state of the art approaches for the sampling of bandlimited signals, and the scenario in large graphs is substantially different as the applicability of approaches like the ones proposed in [21], [31], and [34] is not an option because of the computational complexity. The details of these simulations are follows:

- We generate a set of 100 signals for two signal models.
- Each signal is sampled considering several approaches and then reconstructed by means of eqn. (2.4), calculating the mean squared error (MSE).

- The values of MSE are averaged.

The sampling approaches considered for this experiment are

- Blue noise by Error Diffusion.
- Blue noise sampling by RWS-3 (fast blue noise).
- Uniform random sampling.

The signal models considered are

- Signal model 1 (SM1): A random signal of bandwidth $k = 2500$, where the Fourier coefficients are generated from the Gaussian distribution $\mathcal{N}(1, 0.5)$. The samples captured are contaminated with additive Gaussian noise, making the Signal to Noise Ratio to be $SNR = 20dB$.
- Signal model 2 (SM2): A random signal with Fourier coefficients generated from the Gaussian distribution $\mathcal{N}(1, 0.52)$. This signal is modulated on the spectral domain by the function $h(\mu)$, defined as

$$h(\mu) = \begin{cases} 1 & \text{If} \quad \mu \leq \mu_{2500} \\ e^{-4(\mu - \mu_{2500})} & \text{If} \quad \mu > \mu_{2500} \end{cases} \tag{3.2}$$

The graphs used in the simulations are:

- Graph G_1: A swiss-roll graph with $N = 50000$ nodes. The weights are given by the Euclidean distances between nodes.
- Graph G_2: A sphere graph with $N = 50000$ nodes. The weights are given by the Euclidean distances between nodes.
- Graph G_3: A random sensor network with $N = 50000$ nodes. The weights in the graph are given by the Euclidean distance between points. The maximum number of neighbors for each node is 6.

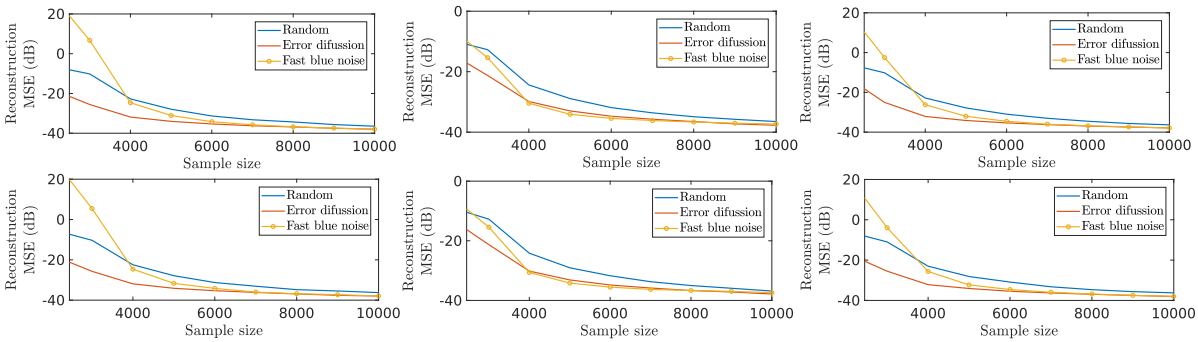


Figure 3.6: Averaged MSE for several sampling approaches on large graphs ($N = 50000$ nodes). From left to right we consider the graphs G_1, G_2 and G_3 respectively. First row is associated to the experiments considering the signal model $SM1$, while the second row is associated to $SM2$.

The results of these experiments are given in Fig. 3.6. We observe that error diffusion outperforms substantially random sampling, while fast blue noise sampling still exhibits some instability when the number of samples is close to the bandwidth of the signal.

3.6 Admisible partitions of $V(G)$ and future work

Now we discuss theoretical results that could represent a promising tool for the development of low complexity algorithms for sampling signals on graphs. This theoretical framework was developed by Fuhr and Pesenson [27] using a novel approach to quantify the quality of a sampling set, using partitions of $V(G)$ and quantities that can be calculated without requiring spectral decompositions. This represents the basis of the theoretical analysis that will be developed in the future to further advance the low complexity blue-noise algorithms discussed in previous sections. Moreover, it will serve as a tool to understand the performance of techniques like error diffusion.

3.6.1 Previous Theoretical Results

Let us consider a partition $\mathcal{P}$ of the set of nodes $V(G)$ in the graph G as $\mathcal{P} = \{\mathcal{S}_0, \mathcal{S}_1, \ldots, \mathcal{S}_n\}$, and let us define the following quantity for any subset $\mathcal{A} \subset V(G)$

$$w_{\mathcal{A}}(v) = \sum_{u \in \mathcal{A}} \mathbf{W}(u, v), \tag{3.3}$$

which is a measure of the connectivity between the node v and $\mathcal{A}$. Now, given the partition $\mathcal{P}$ we let

$$D_\ell = \max_{v \in \mathcal{S}_\ell} \left\{ w_{\mathcal{S}_{\ell+1}}(v) \right\} \tag{3.4}$$

and

$$K_\ell = \min_{v \in \mathcal{S}_{\ell+1}} \left\{ w_{\mathcal{S}_\ell}(v) \right\}. \tag{3.5}$$

The set $\mathcal{S}_0$ is called the *initial set* of the partition $\mathcal{P}$. The following theorem is presented in [27].

Theorem 14 ([27]). *If $\mathcal{P} = \{\mathcal{S}_0, \mathcal{S}_1, \ldots, \mathcal{S}_n\}$ is a partition of $V(G)$ with D_i, K_i finite and $K_i > 0$ for all $i = 0, \ldots n$, then for all $1 < p < \infty, 1 < q < \infty$ with $1/p + 1/q = 1$, and $\boldsymbol{x} \in \mathcal{D}^p(\nabla)$ with $\boldsymbol{x}|_{\mathcal{S}_0} \in L_p(\mathcal{S}_0)$, we have*

$$\|\boldsymbol{x}\|_p \leq \left(\sum_{m=0}^{n} \prod_{j=0}^{m-1} \frac{D_j}{K_j} \right)^{\frac{1}{p}} \|\boldsymbol{x}|_{\mathcal{S}_0}\|_p + \left(\sum_{m=1}^{n} \left(\sum_{k=1}^{m} \frac{1}{K_{k-1}^{q/p}} \left(\prod_{i=k}^{m-1} \frac{D_i}{K_i} \right) \right) \right)^{\frac{1}{p}} \|\nabla \boldsymbol{x}\|_p \tag{3.6}$$

where $\mathcal{D}^p(\nabla)$ is the space of functions $\boldsymbol{x} : V(G) \mapsto \mathbb{C}$ with $\|\nabla \boldsymbol{x}\|_p < \infty$ and

$$\|\nabla \boldsymbol{x}\|_p = \left(\sum_{u,v \in V(G)} \frac{1}{2} |\boldsymbol{x}(u) - \boldsymbol{x}(v)|^p \, \mathbf{W}(u,v) \right)^{\frac{1}{p}} \tag{3.7}$$

Additionally, in [27], it is established that $\left\| \mathbf{L}^{\frac{1}{2}} \boldsymbol{x} \right\|_2^2 = \|\nabla \boldsymbol{x}\|_2^2$ for all $\boldsymbol{x} \in \ell^2(G)$ contained in the domain of $\mathbf{L}$.

From these results it is possible to conclude that for any signal $\boldsymbol{x} \in PW_\omega(G)$ the initial set $\mathcal{S}_0$ of a given partition is a uniqueness set when $\omega < 1/\delta_{\mathcal{P},2}^2$, where $\delta_{\mathcal{P},2}^2$ is given by

$$\delta_{\mathcal{P},2}^2 = \sum_{m=1}^{n} \left(\sum_{k=1}^{m} \frac{1}{K_{k-1}} \left(\prod_{i=k}^{m-1} \frac{D_i}{K_i} \right) \right) \tag{3.8}$$

The problem of finding the best sampling set is then reformulated as the problem of building partitions of $V(G)$ such that the value of $\delta_{\mathcal{P},2}^2$ is minimized.

Notice that, as pointed out in [27], building the optimal partitions is a combinatorial problem. However, if this process has to start from an arbitrary set of nodes,

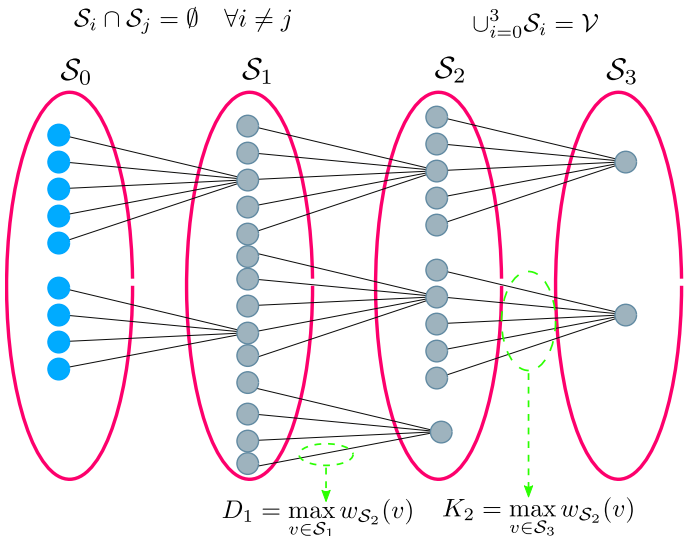


Figure 3.7: Pictorial illustration of a possible partition of $V(G)$, intended to maximize K_i and to minimize D_i.

it would be possible to eventually design the best possible partitions given the starting set of nodes given.

3.6.2 Minimizing $\delta^2_{\mathcal{P},2}$

Despite the fact that the minimization $\delta^2_{\mathcal{P},2}$ is a computationally expensive task, eqn (3.8) allows to establish some basic desirable properties of a given partition $\mathcal{P}$ of $V(G)$. The subsets $\mathcal{S}_\ell$ in $\mathcal{P}$ should be selected such that:

- The values of K_i are as large as possible.
- The values of D_i for $i \geq 1$ are as small as possible.

These basic principles can be used in order to build sequences step by step. In Fig. 3.3 a pictorial illustration of a partition of $V(G)$ is shown. It is intended to have large values for K_i and low values for D_i, always having $K_i > 0$.

3.7 Conclusions and Future work

In this chapter we presented two types of algorithms for a low complexity computation of blue-noise patterns on graphs. We exploited the properties of random

walks on graphs for the generation of patterns with a typical spreading of the sampling nodes, and we proposed an extension of the error diffusion algorithm extensively used in digital halftoning. The results obtained for large graphs are promising as they provide lower reconstruction error than random sampling. Additionally, at the end of the chapter, we proposed to explore the theoretical framework developed by Fuhr and Pesenson in [27] as a tool for the design of sampling patterns and also as a framework that would allow a better understanding of techniques like error diffusion.

Chapter 4

UNIQUENESS SETS IN THE PALEY-WIENER SPACE OF COGRAPHS

4.1 Introduction

The analysis and processing of signals on graphs has become a central topic in data science, and problems like sampling, filtering, and representation have been considered extensively in the literature [12, 21, 64–66]. These results are promoted for arbitrary graphs in general without considering the differences among several classes of graphs. As a consequence, some results that could be well suited for some graphs might not be the best option for others, and the limitations of a given approach are not completely highlighted.

In this work, we study the uniqueness sets for signals defined on *cographs* and provide an efficient algorithm for their closed form solution calculation. Studying the structure and tree representation of cographs, we show how the uniqueness set of bandlimited signals can be obtained by simple operations on small size graphs, without requiring spectral decompositions or the calculation of geodesic distances. The use of cographs is linked to crucial applications in orthology analysis [59, 60] and community detection [61]. Additionally, we provide a closed form solution for the sampling sets of *threshold graphs*. Taking into account that threshold graphs is a subfamily of cographs, we exploit the properties of a tree representation in order to calculate the uniqueness sets directly from the binary sequence that is traditionally used to represent threshold graphs.

Given the potential applications of cographs and threshold graphs for the representation of arbitrary graphs, we perform a set of numerical experiments in which we evaluate the effectiveness of the uniqueness sets of graphs that are obtained from

perturbed cographs. The numerical experiments show that these sampling sets are robust under changes of a considerable number of edges. This opens the door for the idea of considering the construction of a graph from data, forcing the resultant graph to be a cograph or threshold graph. We also point out that the cotree structure itself has potential applications on the analysis of power distribution grids [67].

4.2 Background and preliminaries

We will use the following notation. For any $n \in \mathbb{N}$, we let $[n] := \{1, \dots, n\}$. We identify vectors $\boldsymbol{f} = (f_1, \dots, f_n)^T \in \mathbb{R}^n$ with functions $\boldsymbol{f} : [n] \to \mathbb{R}$, where $\boldsymbol{f}(i) = f_i$. Given a $n \times n$ matrix $\mathbf{A}$ and subsets $S, T \subseteq \{1, \dots, n\}$, we will denote by $\mathbf{A}_{S,T}$ the submatrix of $\mathbf{A}$ with rows in S and columns in T. We also define $\mathbf{A}_S := \mathbf{A}_{S,S}$ when $S = T$ to simplify the notation. Similarly, for a vector $\boldsymbol{v} \in \mathbb{R}^n$, we will denote by $\boldsymbol{v}_S$ the restriction of the vector to its entries corresponding to indices in S. The matrices in $\mathbb{R}^{m \times n}$ with all entries equal to 0 and 1 are denoted by $\mathbf{0}_{m \times n}$ and $\mathbf{1}_{m \times n}$ respectively. Similarly, the vectors in $\mathbb{R}^n$ with all entries equal to 0 and 1 are denoted by $\mathbf{0}_n$ and $\mathbf{1}_n$.

Central to this paper is the notion of a *uniqueness set*.

Definition 2. *Let U be a subspace of $\mathbb{R}^n$. We will say that a subset $S \subseteq [n]$ is a uniqueness set for U if for all $\boldsymbol{g}, \boldsymbol{h} \in U$, the condition $\boldsymbol{g}(i) = \boldsymbol{h}(i)$ for all $i \in S$ implies $\boldsymbol{g} = \boldsymbol{h}$. A uniqueness set will be said to be minimal if it does not contain a uniqueness set as a proper subset.*

In other words, a set $S \subseteq [n]$ is a uniqueness set for U if the entries of any vector $\boldsymbol{f} \in U$ are entirely determined by its entries in S. Note that, equivalently, a subset $S \subseteq [n]$ is a uniqueness set for U if $\boldsymbol{f}(i) = 0 \ \forall i \in S$ implies $\boldsymbol{f} = \mathbf{0}_n$. The following simple result provides a useful way to test if a given set is a uniqueness set.

Proposition 15. *Let U be a k-dimensional subspace of $\mathbb{R}^n$ and let $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_k\} \subseteq \mathbb{R}^n$ be any basis of U. Let $\mathbf{W}$ be the $n \times k$ matrix with columns $\boldsymbol{w}_1, \dots, \boldsymbol{w}_k$. Then the following are equivalent for a subset $S \subseteq [n]$:*

1. *S is a minimal uniqueness set for U.*

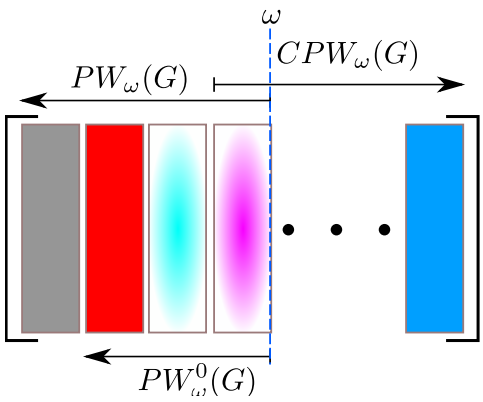


Figure 4.1: Representation of the spaces $PW_\omega(G)$, PW^0_ω and $CPW_\omega(G)$.

2. $|S| = k$ *and* $\det \mathbf{W}_{S,[k]} \neq 0$.

Proof: See Appendix 4.7.2

4.2.1 The Paley–Wiener space of a graph

Let $G = (V, E)$ be a simple graph with adjacency matrix $\mathbf{A}$ and graph Laplacian $\mathbf{L}_G := \mathbf{D} - \mathbf{A}$, where $\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_n)$ and d_i is the degree of the i-th vertex of G. Let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of $\mathbf{L}_G$, and let $\mathbf{1}_n = \boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \in \mathbb{R}^n$ be an associated orthogonal basis of eigenvectors. We identify functions $\boldsymbol{f} : V(G) \to \mathbb{R}$ to vectors in $\mathbb{R}^{V(G)}$.

Definition 3. *For $\omega \geq 0$, the Paley–Wiener space* $\mathrm{PW}_\omega(G)$ *is given by*

$$\mathrm{PW}_\omega(G) := \operatorname{span}\{\boldsymbol{w}_i : \lambda_i \leq \omega\}.$$

The Modified Paley-Wiener space $\mathrm{PW}^0_\omega(G)$ *is defined as*

$$\mathrm{PW}^0_\omega(G) := \operatorname{span}\{\boldsymbol{w}_i : i \geq 2 \text{ and } \lambda_i \leq \omega\}.$$

If $0 \leq \omega < \lambda_2$ with $\lambda_2 > 0$, we set $\mathrm{PW}^0_\omega(G) := \emptyset$. *Finally, we define the Complementary Paley-Wiener space $CPW_\omega(G)$ as*

$$CPW_\omega(G) = \operatorname{span}\{\boldsymbol{w}_i : \lambda_i \geq \omega\}.$$

Clearly, $\mathrm{PW}_\omega(G), \mathrm{PW}^0_\omega(G)$ and $CPW_\omega(G)$ are vector spaces over $\mathbb{R}$ of dimension $\dim \mathrm{PW}_\omega(G) = k$, $\dim \mathrm{PW}^0_\omega(G) = k-1$ and $\dim CPW_\omega(G) = n-k+1$, where k is the largest integer in $[n]$ such that $\lambda_k \leq \omega$. Notice that $\mathrm{PW}^0_\omega(G) = \mathrm{PW}_\omega(G) \cap \mathbf{1}^\perp_{|G|}$. If $\omega \geq \lambda_2$ and k is the largest integer in $[n]$ such that $\lambda_k \leq \omega$, then the modified Paley–Wiener space $\mathrm{PW}^0_\omega(G)$ is a subspace of $\mathbb{R}^{|V(G)|}$ of dimension $k-1$. For the spaces $PW_\omega(G), PW^0_\omega(G)$, and $CPW_\omega(G)$, the value of ω is going to be refered to as the *bandwidth* of the space.

Remark 16. *When some eigenvalues of $\mathbf{L}_G$ are repeated, there is some ambiguity in choosing a basis of eigenvectors of $\mathbf{L}_G$. When choosing such a basis below, we will always assume $\mathbf{1}_n$ is an eigenvector associated to $\lambda_1 = 0$, and that all the eigenvectors are orthogonal.*

Proposition 15 immediately implies the following characterization of uniqueness sets for $\mathrm{PW}_\omega(G)$, $PW^0_\omega(G)$, and $CPW_\omega(G)$.

Corollary 17. *Let G be a simple graph on n vertices with Laplacian $\mathbf{L}_G$. Let $\mathbf{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n] \in \mathbb{R}^{n\times n}$ be any matrix whose columns form a basis of eigenvectors associated to the eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ of $\mathbf{L}_G$. Also, $\omega \geq 0$ and let k be the largest integer in $[n]$ such that $\lambda_k \leq \omega$. Then the following are equivalent for a subset $S \subseteq V(G)$:*

1. *S is a minimal uniqueness set for $\mathrm{PW}_\omega(G)$.*
2. *We have $|S| = k = \dim \mathrm{PW}_\omega(G)$ and the matrix $\mathbf{W}_{S,[k]}$ is non-singular.*

Corollary 18. *Let G be a simple graph on n vertices with Laplacian $\mathbf{L}_G$. Let $\mathbf{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n] \in \mathbb{R}^{n\times n}$ be any matrix whose columns form a basis of eigenvectors associated to the eigenvalues $0 = \lambda_1 \leq \cdots \leq \lambda_n$ of $\mathbf{L}_G$. Also, let $\omega \geq \lambda_2$ and let k be the largest integer in $[n]$ such that $\lambda_k \leq \omega$. Then the following are equivalent for a subset $S \subseteq V(G)$:*

1. *S is a minimal uniqueness set for $\mathrm{PW}^0_\omega(G)$.*

2. *We have* $|S| = k - 1 = \dim \mathrm{PW}^0_\omega(G)$ *and the matrix* $\mathbf{W}_{S,\{2,\dots,k\}}$ *is non-singular.*

Corollary 19. *Let* G *be a simple graph on* n *vertices with Laplacian* $\mathbf{L}_G$*. Let* $\mathbf{W} = [\boldsymbol{w}_1, \dots, \boldsymbol{w}_n] \in \mathbb{R}^{n\times n}$ *be any matrix whose columns form a basis of eigenvectors associated to the eigenvalues* $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ *of* $\mathbf{L}_G$*. Also,* $\omega \geq 0$ *and let* k *be the largest integer in* $[n]$ *such that* $\lambda_k \leq \omega$*. Then the following are equivalent for a subset* $S \subseteq V(G)$*:*

1. *S is a minimal uniqueness set for $CPW_\omega(G)$.*

2. *We have* $|S| = n - k + 1 = \dim CPW_\omega(G)$ *and the matrix* $\mathbf{W}_{S,\{k,\dots,n\}}$ *is non-singular.*

Remark 20. *In the case where* $\mathbf{L}_G$ *has repeated eigenvalues, notice that Corollary 17 holds for any choice of a basis of eigenvectors of* $\mathbf{L}_G$*.*

In traditional Fourier analysis, a common practice relies on separating the analysis in frequency of the signal in two essential components. A first component, called the *dc value*, corresponds to the energy of the signal associated to the complex exponential of frequency zero, and a second component which contains the energy associated to all the remaining complex exponentials. This second component describes how fast the signal changes and determines where the signal can be sampled in order to have a unique representation. Informally speaking, this is way to state that the signal is essentially determined by the sampling set required to preserve its *variation* behavior. Now, it is important to notice that any signal in $PW^0_\omega(G)$ can be considered the representation of a signal in $PW_\omega(G)$ excluding its *dc* component. Therefore, there is a close relationship between the uniqueness sets of $PW^0_\omega(G)$ and of $PW_\omega(G)$. This is formalized in the following theorem.

Theorem 21. *Let* G *be a graph on* n *vertices and let* $\mathbf{U}$ *the matrix whose columns form an orthonormal basis for* $PW_\omega(G)$ *and let* S *a uniqueness set of* $PW^0_\omega(G)$*. Then,* T *is*

a uniqueness set of $PW_\omega(G)$ if and only if $T = S \cup \{v\}$ where $v \in S^c$. In particular, v is any node for which $\boldsymbol{x}(v) \neq 1$, where $\boldsymbol{x}$ is given by

$$\boldsymbol{x} = \mathbf{U}_{[n],[n]\setminus\{1\}}\boldsymbol{\alpha} \tag{4.1}$$

with

$$\boldsymbol{\alpha} = \mathbf{U}^{-1}_{S,[n]\setminus\{1\}}\mathbf{1}_{|S|\times 1}. \tag{4.2}$$

Proof: See Appendix 4.7.3

Theorem 21 is going to play a central role in the calculation of sampling sets of cographs as in some cases it is more convenient to work with $PW^0_\omega(G)$ than with $PW_\omega(G)$.

4.3 Cographs

Cographs are a subclass of graphs that are built from simple operations and have been considered as modeling tools in evolutionary biology and the study of evolution of diseases among different species. The fact that cographs are define in terms of only two basic operations endows these type of graphs with particular structures for the uniqueness sets that we study in this chapter.

4.3.1 Definition and characterizations

Definition 4. *Let $G = (V(G), E(G)), H = (V(H), E(H))$ be two graphs on disjoint sets of vertices. The join of G and H, denoted $G \vee H$, is the union of the two graphs G, H, together with all edges joining $V(G)$ and $V(H)$. Formally, $V(G \vee H) = V(G) \cup V(H)$ and $ab \in E(G \vee H)$ if and only if either*

- *$ab \in E(G)$ or $ab \in E(H)$; or*
- *$a \in V(G),\ b \in V(H)$ or $a \in V(H)$ and $b \in V(G)$.*

Recall that the graph complement of G is the graph $G^c := (V(G), E(G)^c)$. The union of two graphs G and H on disjoint sets of vertices is the graph on $V(G \cup H) :=$

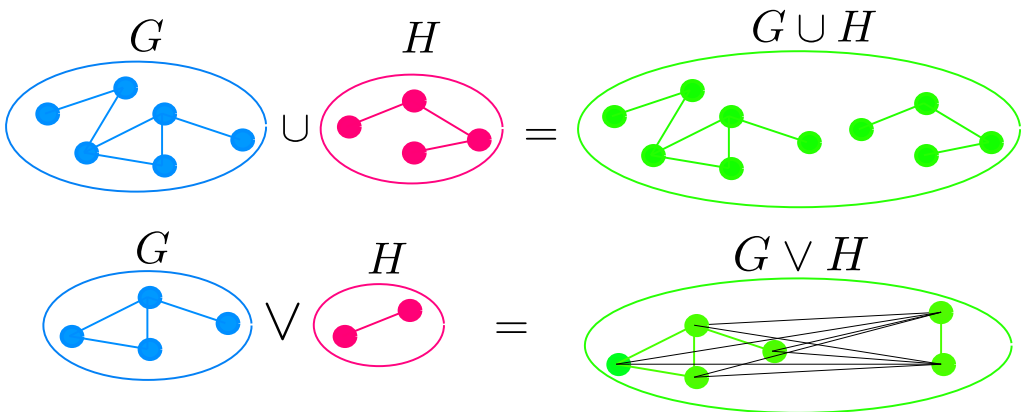


Figure 4.2: Top: Representation of the union operation between two graphs. Bottom: Representation of the join operation of two graphs.

$V(G) \cup V(H)$ where $ab \in E(G \cup H)$ if and only if either $ab \in E(G)$ or $ab \in E(H)$. Using these notions one can equivalently define $G \vee H = (G^c \cup H^c)^c$.

Definition 5. *A cograph (or a complement-reducible graph) is a graph defined recursively as follows:*

1. *Isolated vertices are cographs;*
2. *If G and H are cographs on disjoint vertex sets, then so is their join $G \vee H$;*
3. *If G and H are cographs on disjoint vertex sets, then so is their union $G \cup H$.*

In other words, cographs are graphs that can be constructed from isolated vertices by joints and unions.

Remark 22. *Alternatively, cographs are often defined as follows [58]:*

1. *An isolated vertex is a cograph.*
2. *If G is a cograph, then so is its complement G^c.*
3. *If G and H are cographs, then so is their union $G \cup H$.*

Clearly graphs that can be constructed from joins and unions can be constructed from complements and unions since $G \vee H = (G^c \cup H^c)^c$. Conversely, if G can be constructed

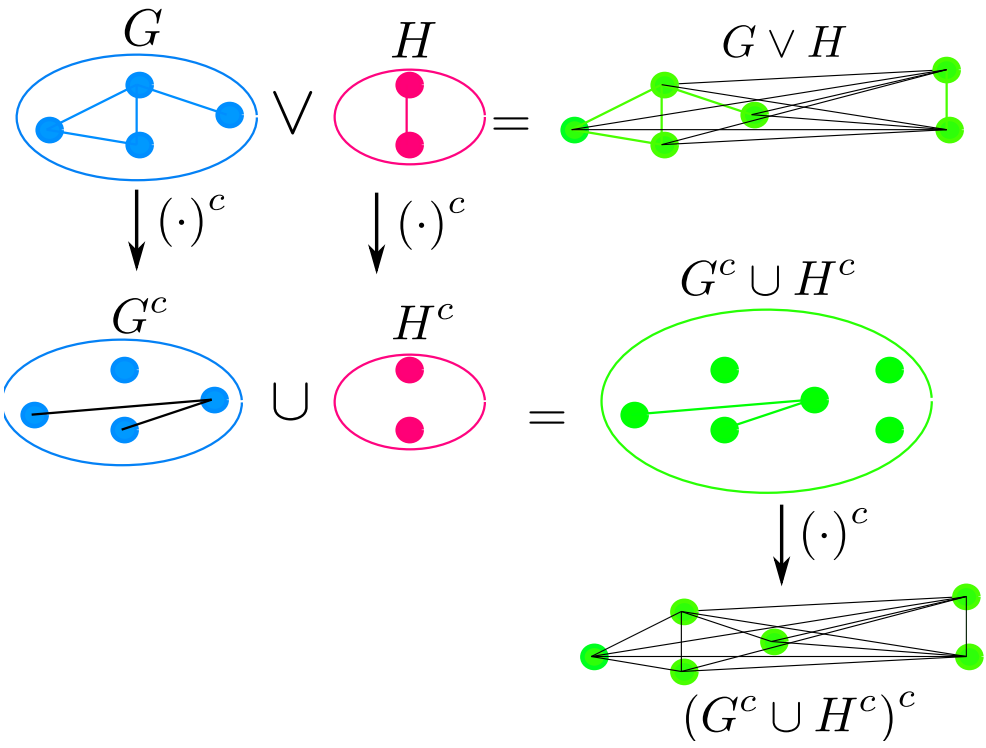


Figure 4.3: Top: Representation of the union operation between two graphs. Bottom: Representation of the join operation of two graphs.

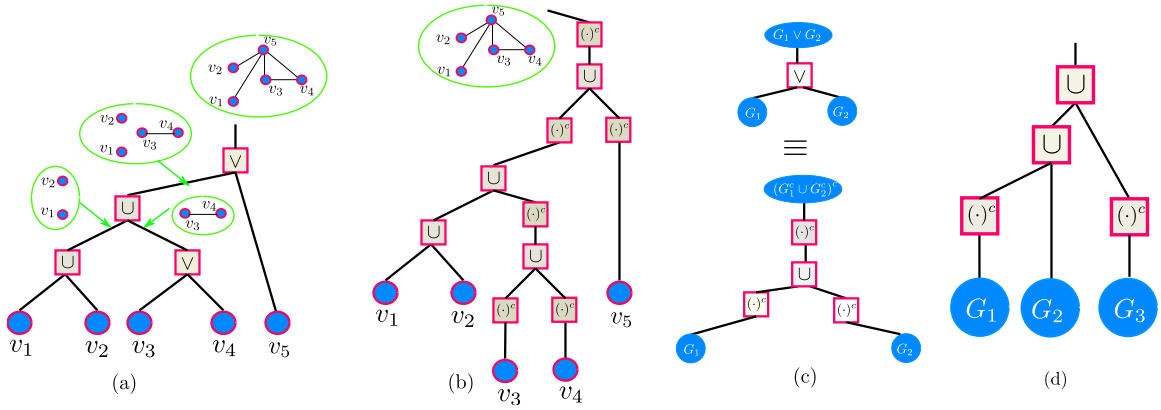


Figure 4.4: (a) The cotree representation of the cograph $G = ((\{v_1\} \cup \{v_2\}) \cup (\{v_3\} \vee \{v_4\})) \vee \{v_5\}$. (b) The cotree representation of the cograph G depicted in (a) using complements and unions. (c) A tree representation showing the equivalence $G_1 \vee G_2 = (G_1^c \cup G_2^c)^c$. (d) Representation of a cotree indicating the construction of a cograph from smaller size cographs G_1, G_2 and G_3.

from complements and unions, then every non-trivial connected induced subgraph of G has a disconnected complement (see [58, Theorem 2]). Thus, if $G_1, \ldots, G_k$ denote the connected components of G^c, then $G = G_1^c \vee \cdots \vee G_k^c$ and $G_1^c, \ldots, G_k^c$ are induced subgraphs of G. It follows easily that G can be constructed from joins and unions and so the two definitions are equivalent.

Cographs can be characterized in several interesting ways. The following result shows that they are precisely the graphs with no induced paths of length 3 or more. Denote by P_n the path graph on n vertices.

Theorem 23 ([58, Theorem 2])**.** *A graph is a cograph if and only if it does not contain P_4 as an induced subgraph.*

4.3.2 Representations of Cographs

Cographs are traditionally represented by a *rooted tree* or *cotree* in which the main nodes indicate the operations of join $\vee$, union $\cup$ and complement $(\cdot)^c$. For instance, consider Figure 4.4(a) in which the cograph $G = ((\{v_1\} \cup \{v_2\}) \cup (\{v_3\} \vee \{v_4\})) \vee \{v_5\}$ is represented. The nodes of this cotree indicate the operations of joint and union. In Figure 4.4(b) we show exactly the same cograph but using the operations of union and complement.

It is important to point out that in Figure 4.4(b) the nodes of the cotree are related to the union operation and the complements are considered operations realized on each branch. In this chapter we will describe our main results in terms of unions and complements, taking into account that any given cograph can be always represented in this way since $G_1 \vee G_2 = (G_1^c \cup G_2^c)^c$. As we will show in the following sections, the cotree representation provides a very useful tool for the calculation of the sampling sets.

4.4 Uniqueness sets of cographs

Before doing a formal calculation of the uniqueness sets in cographs it is natural to find the uniqueness sets for graphs that are obtain from unions and joins separately.

4.4.1 Uniqueness sets of unions and complements

We start this section by describing the minimal uniqueness sets for $\mathrm{PW}_\omega(G)$ and $\mathrm{PW}^0_\omega(G)$ when G is a union of two graphs.

Lemma 24. *Let $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$ be two simple graphs, let $G := G_1 \cup G_2$, and let $\omega \geq 0$. Then*

1. $\dim \mathrm{PW}_\omega(G) = \dim \mathrm{PW}_\omega(G_1) + \dim \mathrm{PW}_\omega(G_2)$.
2. *The following are equivalent for a subset $S \subseteq V(G_1) \cup V(G_2)$:*
   (a) *S is a minimal uniqueness set for $\mathrm{PW}_\omega(G)$.*
   (b) *$S = S_1 \cup S_2$ where S_1 is a minimal uniqueness set for $\mathrm{PW}_\omega(G_1)$ and S_2 is a minimal uniqueness set for $\mathrm{PW}_\omega(G_2)$.*

Proof. The result follows easily from the fact that $L_G = L_{G_1} \oplus L_{G_2}$ and from Corollary 17. □

The pictorial representation of Lemma 24 is presented in Fig. 4.5(top). Observe that any submatrix that is obtained by sampling the eigenvectors associated to $\lambda_i \leq \omega$ on the rows related to a given subset of nodes is a block diagonal matrix whose determinant is the product of the determinants of each block. Therefore, the determinant of the submatrix is only different from zero when the determinant of each block is different from zero.

The analogous statement for $\mathrm{PW}^0_\omega(G_1 \cup G_2)$ is more complicated.

Lemma 25. *Let $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$ be two simple graphs, let $G := G_1 \cup G_2$, and let $\omega \geq 0$. Then*

1. $\dim \mathrm{PW}^0_\omega(G) = \dim \mathrm{PW}^0_\omega(G_1) + \dim \mathrm{PW}^0_\omega(G_2) + 1$.
2. *The following are equivalent for a subset $S \subseteq V_1 \cup V_2$:*
   (a) *S is a minimal uniqueness set for $\mathrm{PW}^0_\omega(G)$.*

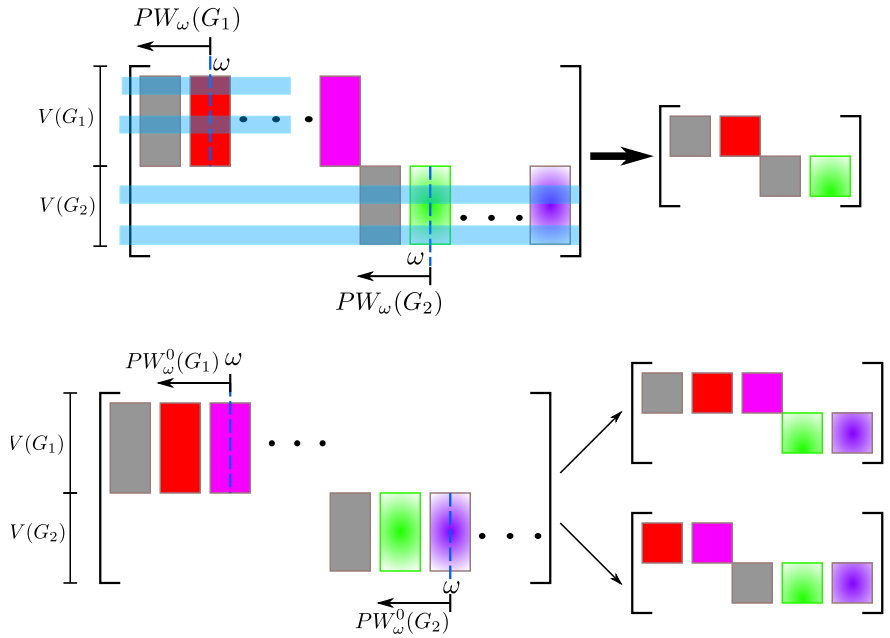


Figure 4.5: Top: Pictorial representation of the results indicated in Lemmas 24. Bottom: Pictorial representation of the results indicated in Lemma 25. The sampling on the rows considering a given sampling set is indicated by the indexes highlighted in blue color. The resultant submatrix is always a block diagonal matrix when unions are considered.

(b) $S = S_1 \cup S_2$ *where* S_1 *is a minimal uniqueness set for* $\mathrm{PW}_\omega(G_1)$ *and* S_2 *is a minimal uniqueness set for* $\mathrm{PW}^0_\omega(G_2)$*, or* S_1 *is a minimal uniqueness set for* $\mathrm{PW}^0_\omega(G_1)$ *and* S_2 *is a minimal uniqueness set for* $\mathrm{PW}_\omega(G_2)$*.*

Proof: See Appendix 4.7.4

A simple graphical representation of Lemma 25 is indicated in Fig. 4.5(bottom).

Remark 26. *It is important to point out that the pictorial representations indicated in Fig. 4.5 can be considered even when one of the Paley-Wiener spaces involved is empty. For instance, if Lemma 25 is applied in a scenario where* $PW^0_\omega(G_2) = \emptyset$*, the matrix basis representation,* $\mathbf{U}$*, for the calculation of the uniqueness sets of* $PW^0_\omega(G_1 \cup G_2)$ *would be given by*

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_{PW_\omega(G_1)} \\ \mathbf{0}_{|V(G_2)| \times \dim PW_\omega(G_1)} \end{bmatrix}. \tag{4.3}$$

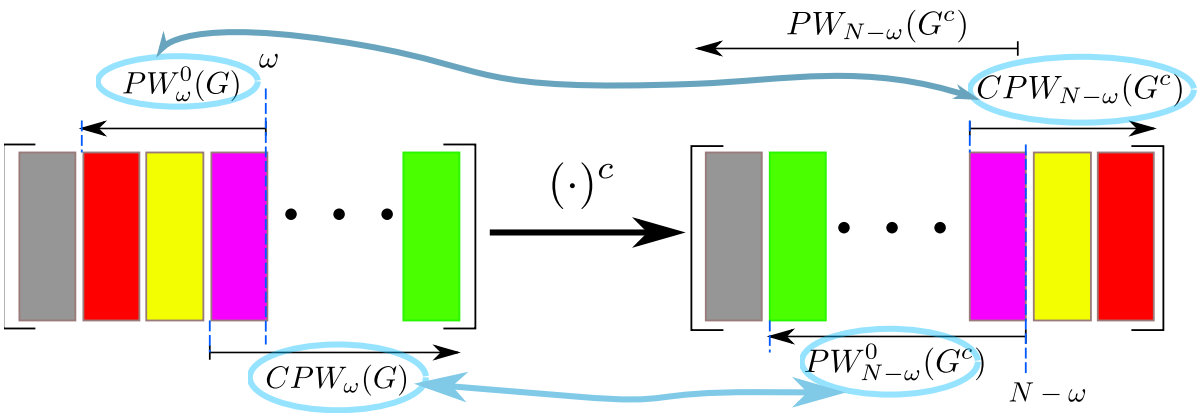


Figure 4.6: Pictorial representation of the results stated in Theorem 28.

From this equation we can see clearly that if we want a submatrix with a determinant different from zero, we cannot select nodes associated to G_2.

We now introduce the characterization of the Paley-Wiener spaces for the complement graph. In order to do so, let us consider the following lemma.

Lemma 27 (see e.g. [63, Section 1.3.2])**.** *Let G be a graph with Laplacian eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ and associated eigenvectors $\mathbf{1}_n = \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$. Then the complement graph G^c has Laplacian eigenvalues $\{0, n-\lambda_n, \ldots, n-\lambda_2\}$, and eigenvectors $\mathbf{1}_n = \boldsymbol{v}_1, \boldsymbol{v}_2 = \boldsymbol{u}_n, \ldots, \boldsymbol{v}_n = \boldsymbol{u}_2$.*

Using the above lemma, we can now describe the Paley-Wiener space of the complement of a graph.

Theorem 28. *Let $G = (V(G), E(G))$ be a graph with $|V(G)| = n$. Then for any $\omega \geq 0$, we have $PW^0_{n-\omega}(G^c) = span\left(\mathbf{U}_{CPW_\omega(G)} \setminus \{\mathbf{1}_n\}\right)$ and $span\left(\mathbf{U}_{CPW_{n-\omega}(G^c)} \setminus \{\mathbf{1}_n\}\right) = PW^0_\omega(G)$.*

Proof: See Appendix 4.7.5

Theorem 28 provides a connection between the complementary and modified Paley-Wiener spaces on a graph and its complement, and is going to play a central role for understanding the structure of eigenvectors in a cograph.

Remark 29. *Notice that in light of Theorem 21, Theorem 28 provides a complete characterization of the uniqueness sets of the Paley-Wiener spaces in the complement graph.*

Remark 30. *In the applications of Lemmas 24, 25 and Theorem 28 we consider that* $PW_\omega(G) = PW^0_\omega(G) = \emptyset$ *for any* $\omega < 0$.

4.4.2 Cotree representation and Paley-Wiener spaces

The results stated previously about $PW_\omega(G)$ when G is obtained from unions or joints of graphs, can be connected to a cotree representation of G. Specifically, given a cotree representation like the one depicted in Fig. 4.4(d), it is possible to determine what is the bandwidth of the Paley-Wiener spaces associated to the smaller cographs size that are being used to build a large size cograph. For instance, considering Fig. 4.4(d), we can see that if we want to calculate the uniqueness sets of $PW_\omega(G)$, we require the uniqueness sets of the two Paley-Wiener spaces associated to the cographs seen from those branches, and therefore we can consider the bandwidth ω suffers no changes when it *propagates trough the cotree*, while it changes after a complement node is met. If we travel on the cotree, applying systematically Lemma 25 and Theorem 28 we reach G_1, G_2 and G_3, knowing exactly what is the bandwidth of the Paley-Wiener spaces involved in the calculation of the uniqueness sets of the cograph.

Lemma 31. *Let* G *a cograph, then the uniqueness sets of* $PW_\omega(G)$, $PW^0_\omega(G)$ *and* $CPW_\omega(G)$ *can be calculated with systematic and sequential application of Lemma 25 and Theorem 28 on the cotree structure of the cograph.*
Proof: See Appendix 4.7.6.

As a consequence of the previous Lemma 31, the convention $PW_\omega(G) = PW^0_\omega(G) = \emptyset$ for any $\omega < 0$ and the fact that every cotree representation reaches its end on the leaves always on a vertex, we can determine which nodes can be selected as part of a uniqueness sets considering the sign of the bandwidth of the Paley-Wiener spaces associated to each vertex.

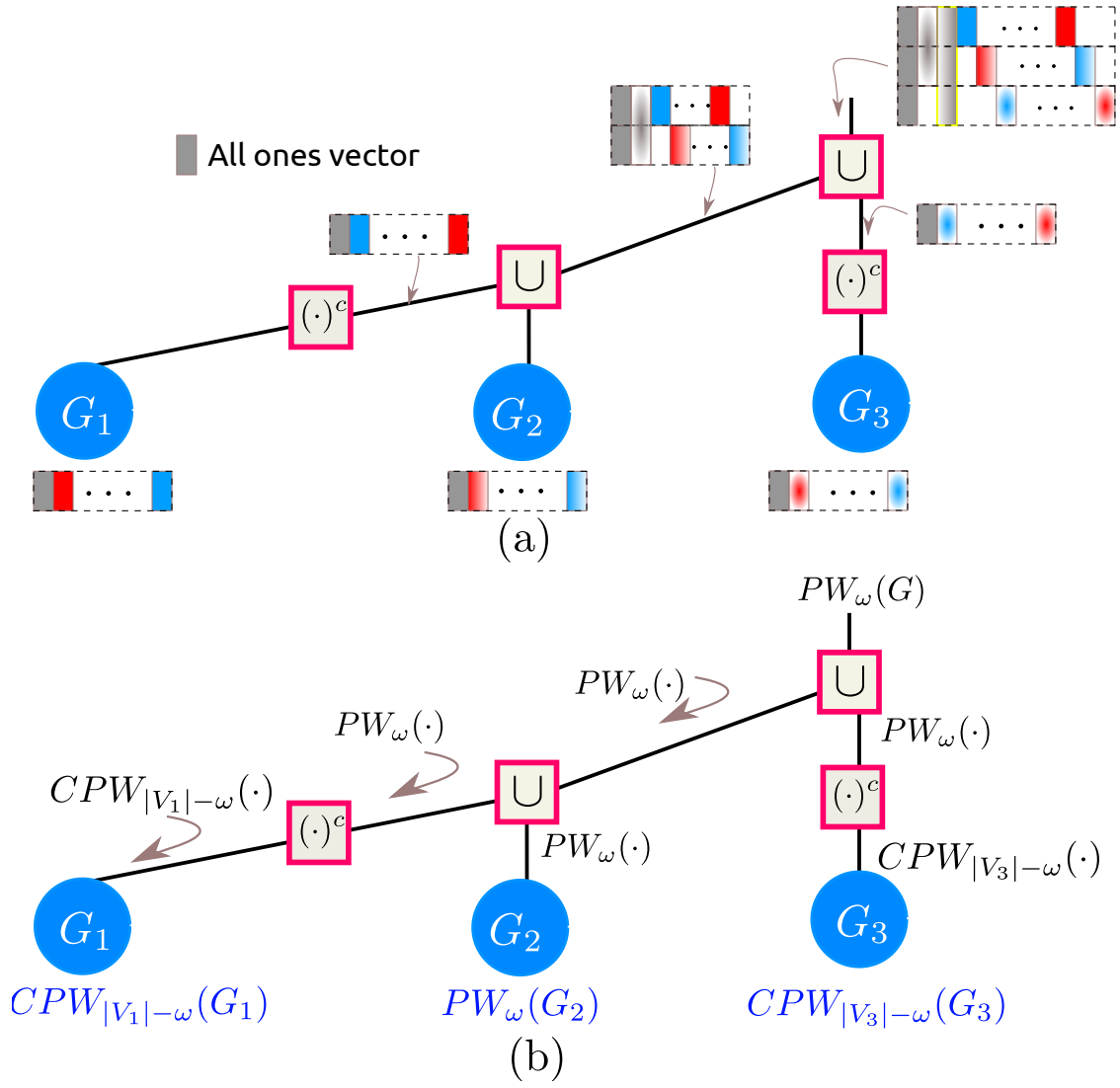


Figure 4.7: Cotree represention of the cograph $G = (G_1^c \cup G_2)^c \cup G_3^c$ that is built from more elementary cographs G_1, G_2 and G_3. (a) Indication of how the set of eigenvectors associated to G are obtained from G_1, G_2, G_3 and the transformations involved when *moving* on the cotree. (b) Indication on the cotree of the Paley-Weiner subspaces of G_1, G_2, G_3 involved in the calculation of the uniqueness set of $PW_\omega(G)$, and how the bandwidth changes when *moving* on the cotree.

Before stating these ideas formally, we introduce the following definition:

Definition 6. *Let us consider a cotree representation of a cograph G where the leaves end always on a vertex i.e., no knowledge of small cographs is used in the cotree representation. Let $\tilde{\omega}(v)$ be the bandwidth of the Paley-Wiener spaces associated to the node v that can be obtained applying Lemma 25 and Theorem 28 on the cotree. We define $V_{\tilde{\omega}\geq 0}(G)$ by*

$$V_{\tilde{\omega}\geq 0}(G) = \{v \in V(G) : \tilde{\omega}(v) \geq 0\}. \tag{4.4}$$

Definition 6 introduces a notation to represent the set of nodes in the cotree representation of a cograph, that have a nonnegative bandwidth associated to the their Paley-Wiener spaces once the cotree representation is considered for the calculation of the uniqueness sets.

Corollary 32. *Let G be a cograph and let $\tilde{\omega}(v)$ be the bandwidth of the Paley-Wiener spaces associated to the node v on the cotree structure built from **unions** and **joins**. Then, any subset $S \subset V_{\tilde{\omega}\geq 0}(G)$ with $|S| = \dim PW_{\omega}(G)$ nodes is a uniqueness set of $PW_{\omega}(G)$.*
Proof: See Appendix 4.7.7.

It is important to point out that this result is, as far as we know, the first broad family of graphs for which the uniqueness sets can be computed in closed form.

Corollary 32 provides a concrete characterization of the nodes in the uniqueness sets of cographs of a Paley-Wiener spaces with a given bandwidth. The application of this corollary can be considered in those cases where the *detailed cotree* is known, i.e. the leaves on the cotree structure reach their end on a vertex.

Remark 33. *When the size of the cograph is large, we do not necessarily assume we start building the cograph from a set of vertices, but instead we start with some small size cographs as it is indicated in Figures. 4.4(d) and 4.7.*

Algorithm 6 Uniqueness set for Cographs

Input: Cotree of cograph G, ω.
Output: A minimal uniqueness set $\mathcal{U}$ of $PW_\omega(G)$.

Initialisation : $\mathcal{U} = \emptyset$
$\mathcal{P} = \mathcal{PW}_\omega(\cdot)$.
for $i = 1 : 1 : c_T$ **do**
  Get the branches $b = b_\ell, b_r$ of $c(i)$. Then for each b:
  if b has no complement **then**
    $\mathcal{U}_b = \mathcal{P}(b)$.
  else
    $\mathcal{P} = \mathcal{CPW}_{|b|-\omega}(\cdot)$
    $\mathcal{U}_b = \mathcal{P}(b \rightarrow)$
  end if
  $\mathcal{U} = \mathcal{U}_{b_\ell} \cup \mathcal{U}_{b_r} \cup \mathcal{U}$
end for
return $\mathcal{U}$.

4.4.3 Complete multipartite graphs

An important example of cographs with extensive applications in communications and data sicence are complete multipartite graphs. In this section we derive the uniqueness sets for these graphs from the above results.

Let $K_i = (V(K_i), E(K_i))$ be a family of complete graphs $(i = 1, \ldots, q; q \in \mathbb{N})$ and consider an arbitrary complete multipartite graph $M = K_1^c \vee K_2^c \vee \cdots \vee K_q^c$. Now, taking into account the connections between the join operation and unions and complements we have that $M = (K_1 \cap K_2 \cap \cdots \cap K_q)^c$. Therefore, considering Lemma 24 and Theorem 28 and their pictorial representations in Fig. 4.7 we know that the eigenbasis matrix associated to M is as indicated in Fig. 4.8, where the support of the eigenvectors is indicated in colors. From Fig. 4.8 we can see clearly the dependency of the uniqueness sets in $PW_\omega(M)$ with respect to the uniqueness sets in the Paley-Wiener spaces for each K_i. The results stated in Fig. 4.8 are formally presented in the following theorem:

Theorem 34. *Let $M = K_1^c \vee K_2^c \vee \cdots \vee K_q^c$, with $|V_{K_1}| < |V_{K_2}| < \ldots < |V_{K_q}|$ the complete multipartite graph. Then the uniqueness sets for the Paley-Weiner space $PW_\omega(M)$ are given by:*

1. *If* $0 \leq \omega < n - |V_{K_q}|$*: Any node in* $\cup_{j=1}^{q} V_{K_j}$.

2. *If* $n - |V_{K_{q-r}}| \leq \omega < n - |V_{K_{q-(r+1)}}|$*:*

$$\left(\bigcup_{\ell=0}^{r} S_{K_{q-\ell}} \right) \bigcup \{v\} \tag{4.5}$$

where $S_{K_{q-\ell}} \subset V_{K_{q-\ell}}$, $|S_{K_{q-\ell}}| = |V_{K_{q-\ell}}| - 1$ *and*

$$v \in \left(\left(\bigcup_{i=1}^{q} V_{K_i} \right) \setminus \left(\bigcup_{\ell=0}^{r} S_{K_{q-\ell}} \right) \right) \tag{4.6}$$

with $r = 0, \ldots, q-1$.

3. *If* $n \leq \omega$*:* $\left(\bigcup_{\ell=1}^{q} V_{K_\ell} \right)$

Proof: See Appendix 4.7.8

In Appendix 4.7.8 an alternative calculation of the uniqueness sets of $PW_\omega(M)$ is presented as a confirmation of the results obtained.

4.5 Numerical Experiments with Uniqueness sets of Cographs

In this section we perform a set of numerical experiments in which the benefits of the calculation of uniqueness sets in cographs show promising applications for the analysis of signals defined on more general graphs. The characteristics of the experiment are given as follows:

- Considering a cograph $\mathcal{G}$, a graph G is obtained from $\mathcal{G}$ by modifying a number of edges.

- A uniqueness set of $PW_\omega(\mathcal{G})$ is calculated and used as a sampling set for $PW_\omega(G)$ for several values of ω.

- A set of 100 signals are randomly generated from each $PW_\omega(G)$, then sampled and reconstructed. The mean squared error is calculated and averaged over the 100 signals.

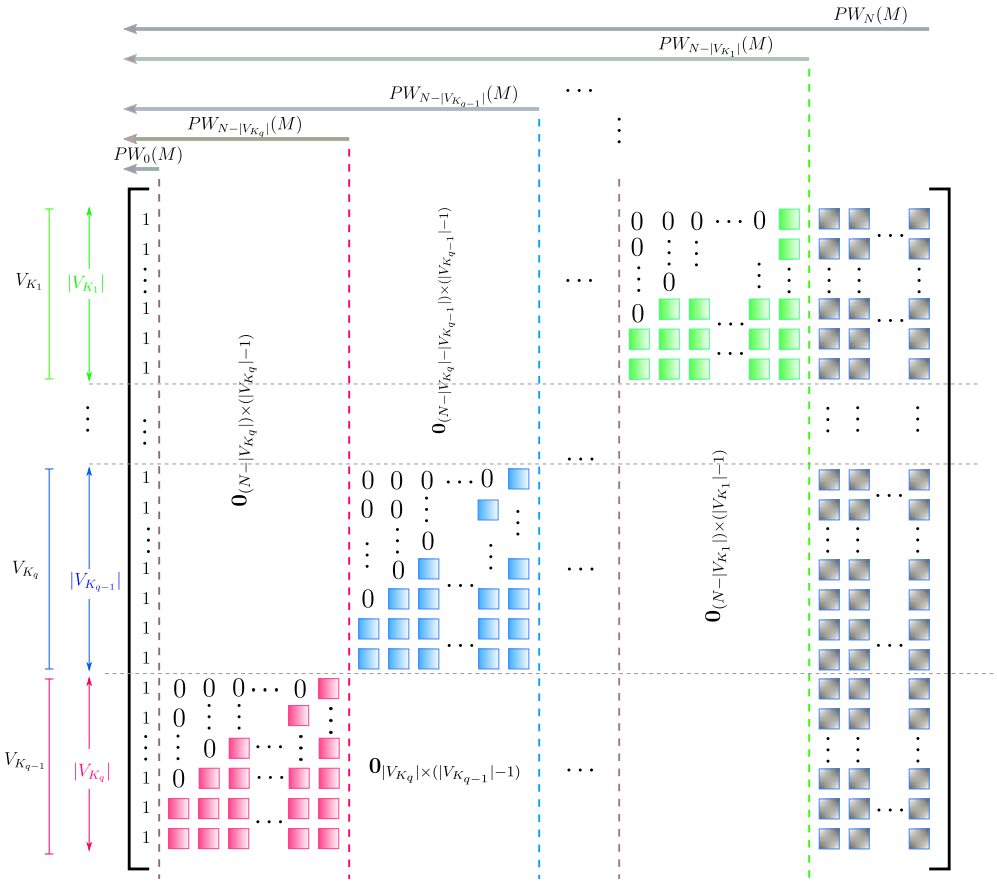


Figure 4.8: Illustration of the support of the eigenvectors of the combinatorial Laplacian for the complete multipartite graph.

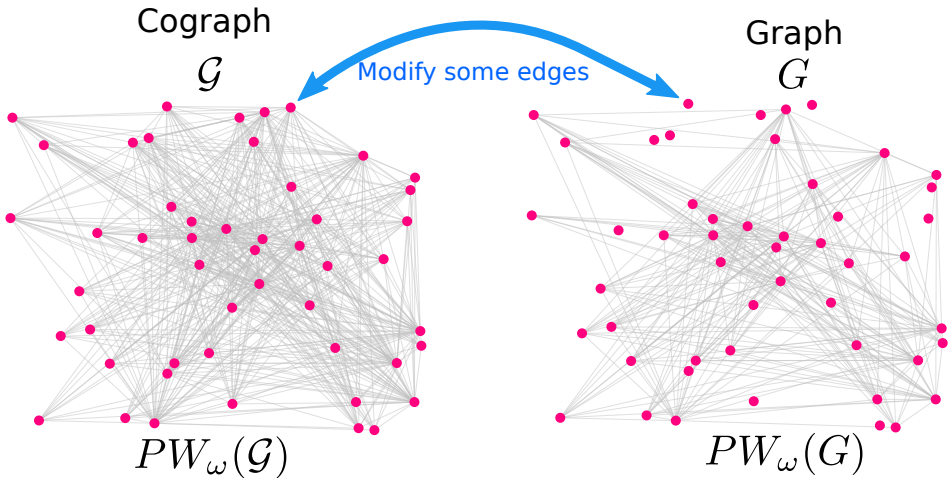


Figure 4.9: Pictorial representation of the numerical experiments performed. A cograph $\mathcal{G}$ is generated and then a subset of edges is modified to generate a graph G (not necessarily a cograph), then the uniqueness sets of $PW_\omega(\mathcal{G})$ are used on $PW_\omega(G)$. The numerical results of this numerical tests can be appreciated in Fig. 4.10.

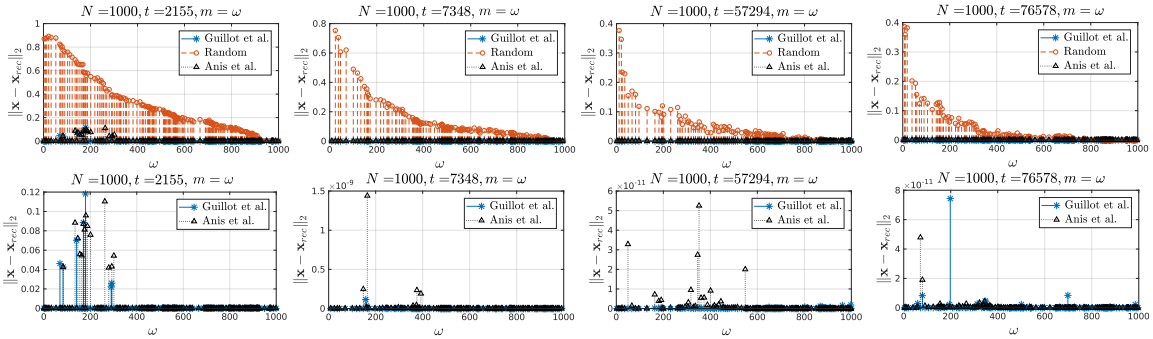


Figure 4.10: First row: reconstruction error for random bandlimited signals defined on a graph G with cograph approximation $\mathcal{G}$. The number t indicates the number of edges that differ between the graph G and $\mathcal{G}$, whereas $m = \omega$ indicates that the number of samples is equal to the bandwidth of the signal. Second row: same results without the random sampling approach.

- The approaches considered are: uniform random sampling, sampling approach proposed by Anis et al. [64] and the approach we proposed.

The results of this experiment are given in Fig. 4.10. It is possible to notice that the values of the error attained by the use of the uniqueness sets obtained from $PW_\omega(\mathcal{G})$ are similar to the ones obtained by using the approach in [64]. However the uniqueness sets calculated from $PW_\omega(\mathcal{G})$ are obtained at a very low computational complexity.

4.6 Threshold graphs

A particular family of cographs is the *threshold graphs*.

Definition 7. *A simple graph G is said to be a threshold graph if it can be constructed from a one-vertex graph by repeated application of one of the following operations:*

1. *Adding an isolated vertex to the graph.*
2. *Adding a dominating vertex to the graph, i.e., adding a vertex that is adjacent to all the other vertices.*

It is important to point out that, as threshold graphs are a particular case of a cograph. Figure 4.11 depicts the particular structure of the cotree for a threshold graph, observe that any operation, join or union, is always performed between a cograph and

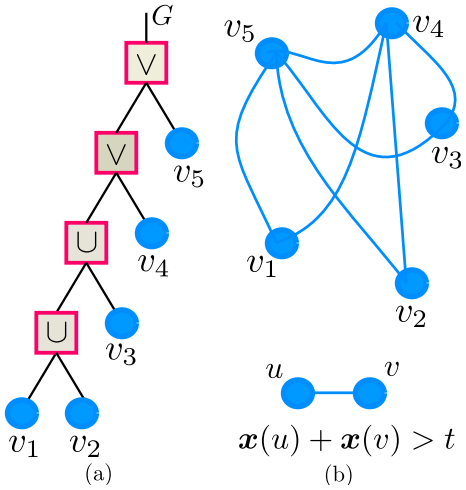


Figure 4.11: (a) The cotree structure of a threshold graph. (b) An equivalent representation of the threshold graph depicted in (a) using the characterization indicated in Theorem 35.

a vertex. As we will show later, some results about the uniqueness sets for the join between two graphs simplify considerably when threshold graphs are considered.

Threshold graphs admit several interesting characterizations. Recall that a subset of vertices is said to be *independent* if no two vertices in the set are adjacent. In what follows, we shall denote the neighborhood of a vertex $a \in V$ in a graph $G = (V, E)$ by

$$n(a) = \{x \in V : ax \in E\}.$$

Theorem 35. *Then following are equivalent for a simple graph $G = (V, E)$:*

1. *G is a threshold graph.*

2. *There exist an integer vertex labeling $c : V \to \mathbb{N}$ and an integer $t \in \mathbb{N}$ (called the threshold) such that for any distinct vertices $a, b \in V$,*

$$ab \in E \iff c(a) + c(b) > t.$$

3. *There exist an integer vertex labeling $c : V \to \mathbb{N}$ and an integer $t \in \mathbb{N}$ such that for any subset of vertices $S \subseteq V$,*

$$S \text{ is independent} \iff \sum_{a \in S} c(a) \leq t.$$

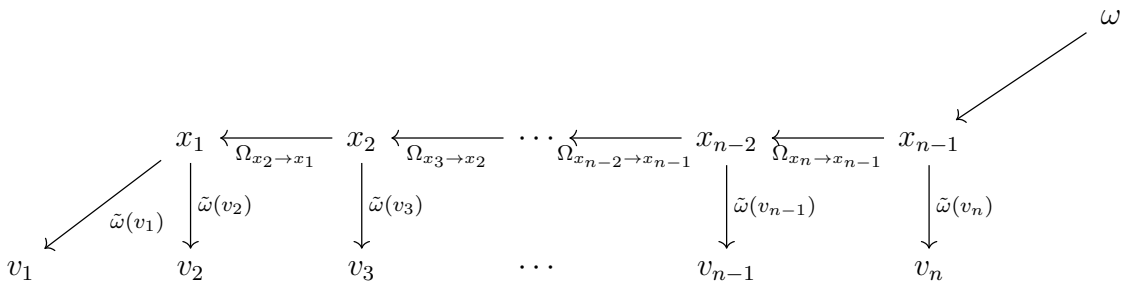


Figure 4.12: Cotree representation of a threshold graph, G, obtained from the binary sequence $\{x_1, x_2, \ldots, x_n\}$, showing the changes in the bandwidth in the Paley-Wiener spaces involved in the calculation of the uniqueness sets for $PW_\omega(G)$.

4. *The set of vertices V can be partitioned as $V = C \cup I$ where C induces a maximal clique in G, and $I := \{i_1, \ldots, i_k\}$ is an independent set such that*

$$n(i_1) \subseteq n(i_2) \subseteq \cdots \subseteq n(i_k).$$

5. *G does not contain P_4, C_4 and $2K_2$ as an induced subgraph, where P_4 and C_4 denote the path and cycle on 4 vertices respectively, and $2K_2$ denotes the union of two disjoint edges.*

Remark 36. *Note that a vertex labeling satisfying property (3) in Theorem 35 automatically satisfies property (2). The converse is false. However, if (2) is satisfied, the theorem guarantees the existence of a (possibly different) labeling satisfying (3).*

The following result provides a characterization of the eigenvectors of the Laplacian matrix of the join of two graphs.

Theorem 37 ([68, Theorem 2.1])**.** *Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be graphs on disjoint sets of $|V(G_1)|$ and $|V(G_2)|$ vertices, respectively. Then*

1. *If the eigenvalues of the Laplacian of G_1 are $0 = \mu_1 \leq \mu_2 \leq \cdots \leq \mu_{|V(G_1)|}$ with associated eigenvectors $\mathbf{1}_{|V(G_1)|} = u_1, \ldots, u_{|V(G_1)|}$, and those of G_2 are $0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_{|V(G_2)|}$ with eigenvectors $\mathbf{1}_{|V(G_2)|} = v_1, \ldots, v_{|V(G_2)|}$, then the eigenvalues of the Laplacian of $G_1 \vee G_2$ are*

- 0.
- $|V(G_1)| + |V(G_2)|$;
- $\mu_2 + |V(G_2)| \leq \cdots \leq \mu_{|V(G_1)|} + |V(G_2)|$;
- $\nu_2 + |V(G_1)| \leq \cdots \leq \nu_{|V(G_2)|} + |V(G_1)|$.

2. *Assume* $u_2, \ldots, u_{|V(G_1)|}$ *are orthogonal to* $\mathbf{1}_{|V(G_1)|}$ *and* $v_2, \ldots, v_{|V(G_2)|}$ *are orthogonal to* $\mathbf{1}_{|V(G_2)|}$. *The eigenvectors of* $G_1 \vee G_2$ *are as follows:*

   - $\mathbf{1}_{|V(G_1)|+|V(G_2)|}$ *is associated to* 0;
   - $(-|V(G_2)|\mathbf{1}_{V_1}) \oplus (|V(G_1)|\mathbf{1}_{V_2})$ *is associated to* $|V(G_1)| + |V(G_2)|$;
   - $u_2 \oplus \mathbf{0}_{V_2}, \ldots, u_{|V(G_1)|} \oplus \mathbf{0}_{V_2}$ *are associated to* $\mu_2 + |V(G_2)|, \ldots, \mu_{|V(G_1)|} + |V(G_2)|$;
   - $\mathbf{0}_{V_1} \oplus v_2, \ldots, \mathbf{0}_{V_1} \oplus v_{|V(G_2)|}$ *are associated to* $\nu_2 + |V(G_1)|, \ldots, \nu_{|V(G_2)|} + |V(G_1)|$.

We can now explain how to compute the uniqueness set of a join of two graphs.

Theorem 38. *Let* $|V(G_1)| \leq |V(G_2)|$ *and let* $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$ *be two simple graphs on* $|V(G_1)|$ *and* $|V(G_2)|$ *vertices, with spectrum* $0 = \mu_1 \leq \mu_2 \leq \cdots \leq \mu_{|V(G_1)|}$ *and* $0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_{|V(G_2)|}$ *respectively. Let* $G := G_1 \vee G_2$, *and let* $\omega \geq 0$. *Also, let* $S \subseteq V_1 \cup V_2$, *and let* $S_1 := S \cap V_1$ *and* $S_2 := S \cap V_2$. *Then the following are equivalent to* S *being a minimal uniqueness set for* $\mathrm{PW}_\omega(G)$*:*

a) If $|V(G_1)| > \omega$, *then* $S = \{a\}$ *for any* $a \in V_1 \cup V_2$.

b) If $|V(G_1)| \leq \omega < |V(G_2)|$, *then either* S_2 *is a minimal uniqueness set for* $\mathrm{PW}_{\omega-|V(G_1)|}(G_2)$ *and* $S_1 = \emptyset$, *or* S_2 *is a minimal uniqueness set for* $\mathrm{PW}^0_{\omega-|V(G_1)|}(G_2)$ *and* $S_1 = \{a\}$ *for some* $a \in V_1$.

c) If $|V(G_2)| \leq \omega < |V(G_1)| + |V(G_2)|$, *then either* S_1 *is minimal uniqueness set for* $\mathrm{PW}_{\omega-|V(G_2)|}(G_1)$ *and* S_2 *is a minimal uniqueness set for* $\mathrm{PW}^0_{\omega-|V(G_1)|}(G_2)$; *or* S_1 *is minimal uniqueness set for* $\mathrm{PW}^0_{\omega-|V(G_2)|}(G_1)$ *and* S_2 *is a minimal uniqueness set for* $\mathrm{PW}_{\omega-|V(G_1)|}(G_2)$.

d) If $\omega \geq |V(G_1)| + |V(G_2)|$, *then* $S = V_1 \cup V_2$.

Proof: See appendix 4.7.9.

Theorem 38 can be applied recursively to compute the minimal uniqueness sets of a cograph. However, in the process, ones needs to be able to compute minimal uniqueness sets for the modified Paley–Wiener space $\mathrm{PW}^0_\lambda(G_1 \vee G_2)$. We now prove an analog of Theorem 38 to address that case. Notice that when $\lambda < \min(|V_1|, |V_2|)$, the space $\mathrm{PW}^0_\lambda(G_1 \vee G_2)$ is empty and there is nothing to prove. We therefore restrict ourselves to the case where $\lambda \geq \min(|V_1|, |V_2|)$.

Theorem 39. *Let* $|V(G_1)| \leq |V(G_2)|$ *and let* $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$ *be two simple graphs on* $|V(G_1)|$ *and* $|V(G_2)|$ *vertices with spectrum* $0 = \mu_1 \leq \mu_2 \leq \cdots \leq \mu_m$ *and* $0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_n$ *respectively. Let* $G := G_1 \vee G_2$, *let* $S \subseteq V(G_1) \cup V(G_2)$, *and let* $S_1 := S \cap V(G_1)$ *and* $S_2 := S \cap V(G_2)$. *Then the following are equivalent to* S *being a minimal uniqueness set for* $\mathrm{PW}^0_\omega(G)$*:*

a) If $|V(G_1)| \leq \omega < |V(G_2)|$ *then* S_2 *is a minimal uniqueness set for* $\mathrm{PW}^0_{\omega-|V(G_1)|}(G_2)$ *and* $S_1 = \emptyset$.

b) If $|V(G_2)| \leq \omega < |V(G_1)| + |V(G_2)|$, *then* S_1 *is minimal uniqueness set for* $\mathrm{PW}^0_{\omega-|V(G_2)|}(G_1)$ *and* S_2 *is a minimal uniqueness set for* $\mathrm{PW}^0_{\omega-|V(G_1)|}(G_2)$.

c) If $\omega \geq |V(G_1)| + |V(G_2)|$, *then* S *is any subset of size* $\dim \mathrm{PW}^0_\omega(G) = |V(G_1)| + |V(G_2)| - 1$.

Proof: See Appendix 4.7.10

Remark 40. *It is important to point out that the items a), b) and c) in Theorem 39 can be unified and reformulated in one statement considering the Paley-Wiener spaces* $PW^0_{\omega-|V(G_1)|}(G_2)$ *and* $PW^0_{\omega-|V(G_2)|}(G_1)$. *In particular:*

- *If* $\omega < |V(G_1)|$ *then* $PW^0_{\omega-|V(G_1)|}(G_2) = \emptyset$ *and* $PW^0_{\omega-|V(G_2)|}(G_1) = \emptyset$ *and therefore* $\mathrm{PW}^0_\omega(G) = \emptyset$.

- *If* $|V(G_1)| \leq \omega < |V(G_2)|$, *then* $PW^0_{\omega-|V(G_2)|}(G_1) = \emptyset$ *and therefore the uniqueness sets are derived from* $PW^0_{\omega-|V(G_1)|}(G_2)$.

- *If* $|V(G_2)| \leq \omega < |V(G_1)|+|V(G_2)|$ *then both* $PW^0_{\omega-|V(G_1)|}(G_2)$ *and* $PW^0_{\omega-|V(G_2)|}(G_1)$ *are nonempty and the uniqueness sets are derived from them.*

4.6.1 Cotree structure of a Threshold graph

Any threshold graph is obtained as a sequence of joints and unions, therefore Theorems 38 and 39 play a central role for the calculation of uniqueness sets. Now, it is important to point out that in light of Theorem 21, we can focus on the systematic application of Theorem 39. The consequence of the systematic application of these results lead to closed form solution for the uniqueness sets in threshold graphs. In order to show this, let us consider the following notation:

- A threshold graph G with $|V(G)| = n$ nodes is represented by the ordered sequence $\{x_1, x_2, \ldots, x_{n-1}\}$, $x_i \in \{0, 1\} \forall i$.

- Starting with a single vertex unions and joins are computed according to x_i. If $x_i = 0$ a union is performed, and if $x_i = 1$ a joint is performed.

- The cotree of the threshold graph is depicted in Fig. 4.12, where the changes of the bandwidth trough the cotree are shown. In particular $\tilde{\omega}(v)$ represents the bandwidth of the Paley-Wiener spaces associated to the node v.

- In Fig. 4.12 the term $\Omega_{x_{k+1}\to x_k}$ indicates the bandwidth that *propagates* from the node x_{k+1} to the node x_k in the cotree.

- Remember from the remarks and results of previous subsections that if the bandwidth in one of the branches of the cotree is negative, the Paley-Wiener space associated to that branch is empty and therefore we do not select nodes associated to that branch.

- Notice that Theorem 39 simplifies substantially when we perform the join between a cograph and a vertex, as it is always obvious which graph is G_1 and which graph is G_2.

Then, the following result can be stated.

Theorem 41. *Let the ordered binary sequence* $\{x_1, x_2, \ldots, x_{n-1}\}$ *be the representation of a threshold graph* G*. Then, any uniqueness set of* $PW_\omega(G)$ *is given by any subset* $\{v_{i_1}, v_{i_2}, \ldots, v_{i_q}\} \subset V(G)$ *of* q *nodes where* $\omega - d(v_{i_r}) \geq 0$*, when* $d(v_{i_r})$ *denotes the degree of the node* v_{i_r} *and* $q = \dim PW_\omega(G)$*. Additionally, the uniqueness set with maximum stability is obtained by the* q *nodes with the lowest degree.*

Proof. From the cotree structure of the threshold graph depicted in Fig. 4.12, and considering Lemma 25 and Theorems 38, 39 we have that:

$$\tilde{\omega}(v_{k+1}) = \begin{cases} \Omega_{x_{k+1}\to x_k} & \text{If} \quad x_k = 0 \\ \Omega_{x_{k+1}\to x_k} - k & \text{If} \quad x_k = 1 \end{cases} \tag{4.7}$$

Additionally, we have

$$\Omega_{x_{k+1}\to x_k} = \begin{cases} \Omega_{x_{k+2}\to x_{k+1}} & \text{If} \quad x_{k+1} = 0 \\ \Omega_{x_{k+2}\to x_{k+1}} - 1 & \text{If} \quad x_{k+1} = 1 \end{cases} \tag{4.8}$$

Now, from eqn. 4.8 and the structure of the cotree, we can conclude that

$$\Omega_{x_{k+1}\to x_k} = \omega - \sum_{j=k+1}^{n} x_j. \tag{4.9}$$

Combining eqn. 4.9 and eqn. 4.7 we have

$$\tilde{\omega}(v_{k+1}) = \begin{cases} \omega - \sum_{j=k+1}^{n} x_j & \text{If} \quad x_k = 0 \\ \omega - \sum_{j=k+1}^{n} x_j - k & \text{If} \quad x_k = 1 \end{cases} \tag{4.10}$$

Now, taking into account that

$$d(v_{k+1}) = \begin{cases} \sum_{j=k+1}^{n} x_j & \text{If} \quad x_k = 0 \\ k + \sum_{j=k+1}^{n} x_j & \text{If} \quad x_k = 1 \end{cases} \tag{4.11}$$

we obtain

$$\tilde{\omega}(v_k) = \omega - d(v_k) \tag{4.12}$$

With this result and the previous results we got for cographs in the cotree, we emphasize that if $\tilde{\omega}(v_k) < 0$ the associated Paley-Wiener space is empty and therefore no nodes are selected from that space. □

4.7 Proofs

4.7.1 Schur Complement Lemma

Lemma 42. *Let*

$$M = \begin{pmatrix} A & \mathbf{0}_{n_1 \times n_2} \\ C & D \end{pmatrix}$$

be a matrix with blocks $A \in \mathbb{R}^{n_1 \times n_1}, C \in \mathbb{R}^{n_2 \times n_1}$, *and* $D \in \mathbb{R}^{n_2 \times n_2}$. *Then* M *is invertible if and only if* A *and* D *are invertible.*

Proof. ($\Rightarrow$) Suppose D is not invertible. Then the columns of D are linearly dependent. Hence the last n_2 columns of M are linearly dependent as well and so M is not invertible. A similar reasoning applies to the rows of A if A is not invertible. Therefore A and D are invertible if M is.

($\Leftarrow$). Now, suppose A and D are invertible. Using the theory of Schur complements, we have

$$\det M = (\det D) \times \det(A - \mathbf{0}_{n_1 \times n_2} D^{-1} C) = \det A \times \det D \neq 0.$$

Thus, M is invertible. □

4.7.2 Proof of Proposition 15

Proof. (2) $\implies$ (1). Suppose $|S| = k$ and $W_{S,\{1,\ldots,k\}}$ is non-singular. Let $f = \sum_{i=1}^{k} a_i w_i$ be such that $f_S = \sum_{i=1}^{k} a_i (w_i)_S = \mathbf{0}_{|S|}$. Since the vectors $\{(w_i)_S : i = 1, \ldots, k\}$ are linearly independent, we conclude that $a_i = 0$ for all $i = 1, \ldots, k$, and so $f = \mathbf{0}_n$ as desired.

(1) $\implies$ (2). now suppose (1) holds. If $|S| < k$, then the vectors $(w_1)_S, \ldots (w_k)_S$ are linearly dependent. Thus, there exist scalars $a_1, \ldots, a_k \in \mathbb{R}$ (not all zero) such that

$$\sum_{i=1}^{k} a_i (w_i)_S = \mathbf{0}_{|S|}. \tag{4.13}$$

Let $f = \sum_{i=1}^{k} a_i w_i$. Then $f_S = \mathbf{0}_{|S|}$. Since S is a uniqueness set, we conclude that $f = \mathbf{0}_n$. This contradicts the linear independence of $w_1, \ldots, w_k$. Hence $|S| \geq k$. now, assume $|S| > k$, and suppose Equation (4.13) holds for some $a_1, \ldots, a_k \in \mathbb{R}$. By assumption, it follows that $\sum_{i=1}^{k} a_i w_i = \mathbf{0}_n$, and so $a_1 = \cdots = a_k = 0$ by the linear independence of $w_1, \ldots, w_k$. Thus, the vectors $(w_1)_S, \ldots, (w_k)_S$ are linearly independent. Since the row rank equals the column rank of a matrix, there exists $S' \subseteq S$ of size k such that $(w_1)_{S'}, \ldots, (w_k)_{S'}$ are also linearly independent. By the (2) $\implies$ (1) implication above, the set S' is a uniqueness set, contradicting the minimality of S.

Finally, suppose $|S| = k$ and $W_{S,\{1,\ldots,k\}}$ is singular. Then there exist scalars $a_1, \ldots, a_k \in \mathbb{R}$ (not all zero) such that Equation (4.13) holds. Defining $f = \sum_{i=1}^{k} a_i w_i$ as above, we obtain that $f_S = \mathbf{0}_{|S|}$. By (1), we conclude that $f = \mathbf{0}_n$, contradicting the linear independence of the eigenvectors $w_1, \ldots, w_k$. □

4.7.3 Proof of Theorem 21

Proof. Let us consider the matrix $\mathbf{M}$ given by

$$\mathbf{M} = \begin{bmatrix} \mathbf{U}_{S,[n]\setminus\{1\}} & \mathbf{1}_{|S|\times 1} \\ \mathbf{U}_{\{v\},[n]\setminus\{1\}} & 1 \end{bmatrix} \tag{4.14}$$

with $v \in S^c$. Calculating the determinant of $\mathbf{M}$ using the Schur-complement theorem, we obtain

$$\det(\mathbf{M}) = (1 - \mathbf{U}_{\{v\},[n]\setminus\{1\}} \mathbf{U}^{-1}_{S,[n]\setminus\{1\}} \mathbf{1}_{|S|\times 1}) \det(\mathbf{U}_{S,[n]\setminus\{1\}}) \tag{4.15}$$

therefore $\det(\mathbf{M}) \neq 0$ if and only if

$$\mathbf{U}_{\{v\},[n]\setminus\{1\}} \mathbf{U}^{-1}_{S,[n]\setminus\{1\}} \mathbf{1}_{|S|\times 1} \neq 1 \tag{4.16}$$

We claim that there exists always $v \in S^c$ such that eqn. (4.16) holds. Indeed, suppose for a contradiction that there is no such v. Notice that the sampled version of a signal $\boldsymbol{x} \in PW_\omega^0(G)$ on a subset of nodes S is given by

$$\boldsymbol{x}(S) = \mathbf{U}_{S,[n]\setminus\{1\}}\boldsymbol{\alpha} \tag{4.17}$$

If S is a uniqueness set for $PW_\omega^0(G)$, then $\boldsymbol{\alpha}$ is unique and can be calculated as

$$\boldsymbol{\alpha} = \mathbf{U}_{S,[n]\setminus\{1\}}^{-1}\boldsymbol{x}(S) \tag{4.18}$$

If $\boldsymbol{x}(S) = 1$, then

$$\mathbf{U}_{\{v\},[n]\setminus\{1\}}\boldsymbol{\alpha} = \mathbf{U}_{\{v\},[n]\setminus\{1\}}\mathbf{U}_{S,[n]\setminus\{1\}}^{-1}\mathbf{1}_{|S|\times 1} \tag{4.19}$$

Now, if there is no $v \in S^c$ such that $\mathbf{U}_{\{v\},[n]\setminus\{1\}}\boldsymbol{\alpha} \neq 1$, this would imply that

$$\mathbf{1}_{|V(G)|\times 1} = \mathbf{U}_{[n],[n]\setminus\{1\}}\boldsymbol{\alpha} \tag{4.20}$$

which is a contradiction since the columns of $\mathbf{U}$ are assumed to be an orthonormal basis. □

4.7.4 Proof of Lemma 25

Proof. Let $m := |V_1|$ and $n := |V_2|$. Also, let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{k_1}$ and $0 = \mu_1 \leq \mu_2 \leq \cdots \leq \mu_{k_2}$ denote the eigenvalues of L_{G_1} and L_{G_2} that are less than λ, respectively. Denote the associated eigenvectors of L_{G_1} and L_{G_2} by $\mathbf{1}_m =: u_1, u_2, \ldots, u_{k_1}$ and $\mathbf{1}_n =: v_1, v_2, \ldots, v_{k_2}$. Then $\mathrm{PW}_\lambda^0(G) = \mathrm{span}(-n\mathbf{1}_m \oplus m\mathbf{1}_n, u_2 \oplus \mathbf{0}_n, \ldots, u_{k_1} \oplus \mathbf{0}_n, \mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{k_2})$. Hence $\dim \mathrm{PW}_\lambda^0(G) = \dim \mathrm{PW}_\lambda^0(G_1) + \dim \mathrm{PW}_\lambda^0(G_2) + 1$. Let W be the matrix with columns $-n\mathbf{1}_m \oplus m\mathbf{1}_n, u_2 \oplus \mathbf{0}_n, \ldots, u_{k_1} \oplus \mathbf{0}_n, \mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{k_2}$, and let S be a minimal uniqueness set for $\mathrm{PW}_\lambda^0(G)$. By Corollary 18, we have $|S| = \dim \mathrm{PW}_\lambda^0(G_1) + \mathrm{PW}_\lambda^0(G_2) + 1$. Proceeding as in the Proof of Theorem 38(d), we obtain that $|S_1| \geq \dim \mathrm{PW}_\lambda^0(G_1)$ and $|S_2| \geq \dim \mathrm{PW}_\lambda^0(G_2)$. Hence, either $|S_1| = \dim \mathrm{PW}_\lambda^0(G_1) + 1$ and $|S_2| = \dim \mathrm{PW}_\lambda^0(G_2)$, or $|S_1| = \dim \mathrm{PW}_\lambda^0(G_1)$ and $|S_2| = \dim \mathrm{PW}_\lambda^0(G_2) + 1$. The result now follows by writing W in block form as in the proof of Theorem 38(d) (see eqn. (4.36)). □

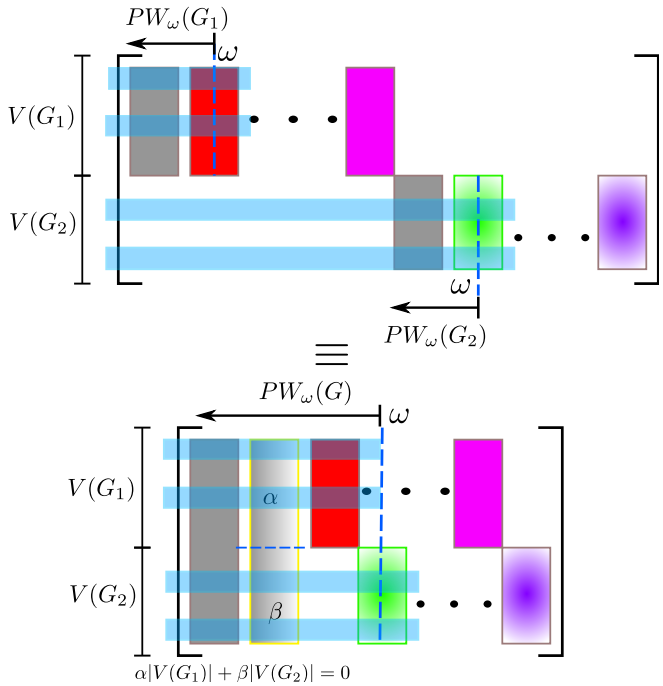


Figure 4.13: Top: Representation of the union operation between two graphs. Bottom: Representation of the join operation of two graphs.

4.7.5 Proof of Theorem 28

Proof. Let $0 = \mu_1, \mu_2, \ldots, \mu_n$ denote the Laplacian eigenvalues of G^c. We know that $PW_{n-\omega}(G^c)$ is generated by all eigenvectors associated to the eigenvalues $\mu_i \leq n - \omega$, and by Lemma 27 the μ_i's are given by $0, n - \lambda_n, \ldots n - \lambda_i$ with $i > 1$. This implies that $n - \lambda_i \leq n - \omega \Rightarrow \lambda_i \geq \omega$ then $PW_{n-\omega}(G^c) = \mathbf{1}_n \oplus CPW_\omega(G)$. Similarly, we have that $CPW_{n-\omega}(G^c)$ is generated by the eigenvectors associated to the eigenvalues $\mu_i \geq n - \omega$ which implies $\lambda_i \leq \omega$. □

4.7.6 Proof of Lemma 31

Proof. Any cograph has a cotree representation based on unions or joins. At any node in the cotree, the calculation of a uniqueness set for any modified or complementary Paley-Wiener space is based on the application of Lemma 25 and Theorem 28. Now, by means of Theorem 21 a uniqueness set of a Paley-Wiener space can be obtained as described. □

4.7.7 Proof of Corollary 32

Proof. The proof follows by the systematic application of Theorems 38 and 39 □

4.7.8 Proof of Theorem 34

Lemma 43. *The eigenvalues of* $K_q = (V_{K_q}, E_{K_q})$ *are given by* $\{0, |V_{K_q}|^{|V_{K_q}|-1}\}$ *with eigenvectors given by:* $u_1 = \mathbf{1}_{|V_{K_q}|\times 1}$ *and*

$$u_i(\ell) = \begin{cases} (i-1) & \ell = |V_{K_q}| - (i-1) \\ -1 & \ell \geq |V_{K_q}| - (i-2) \\ 0 & \textit{otherwise} \end{cases} \tag{4.21}$$

for $i = 2, \ldots, |V_{K_q}|$.

Proof. Given that the components of the j^{th} row, $\mathbf{r}_j$, of the Laplacian matrix for the complete graph can be written as

$$\mathbf{r}_j(\ell) = \begin{cases} |V_{K_q}| - 1 & \text{if} \quad \ell = j \\ -1 & \ell \neq j \end{cases} \tag{4.22}$$

we have that

$$(Lu_i)(j) = \left(|V_{K_q}| - 1\right) u_i(j) + \sum_{\ell \neq j} (-1) u_i(\ell). \tag{4.23}$$

Now, taking into account 4.21 for the resultant sum in equation 4.23, we have the following possibilities:

- If $j < |V_{K_q}| - (i-1)$ then $u_i(j) = 0$ and therefore

$$\left(|V_{K_q}| - 1\right) u_i(j) \quad + \quad \sum_{\ell \neq j} (-1) u_i(\ell) \qquad = \qquad 0 \qquad = \qquad |V_{K_q}| u_i(j) \tag{4.24}$$

- If $j = |V_{K_q}| - (i-1)$ then $u_i(j) = (i-1)$ and therefore

$$\left(|V_{K_q}| - 1\right) u_i(j) + \sum_{\ell \neq j} (-1) u_i(\ell) = (|V_{K_q}| - 1)(i-1) + (i-1) = |V_{K_q}| u_i(j) \tag{4.25}$$

- If $j \geq |V_{K_q}| - (i-2)$ then $u_i(j) = -1$ and therefore

$$\left(|V_{K_q}| - 1\right) u_i(j) + \sum_{\ell \neq j} (-1) u_i(\ell) = (|V_{K_q}| - 1)(-1) - (i-1) + (i-2) = -|V_{K_q}|$$

$$= |V_{K_q}| u_i(j) \quad (4.26)$$

  and therefore $Lu_i = |V_{K_q}| u_i$ for $i = 2, \ldots, |V_{K_q}|$.

The fact that $L\mathbf{1}_{|V_{K_q}| \times 1} = 0$ follows directly from the fact that $\mathbf{r}\mathbf{1}_{|V_{K_q}| \times 1} = 0$ for all j. □

Lemma 44 ([63]). *Let G be a graph with Laplacian eigenvalues $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$ and eigenvectors $u_1 = \mathbf{1}, u_2, \ldots, u_n$; then the complement graph G^c has Laplacian eigenvalues $\{0, n - \mu_n, \ldots, n - \mu_2\}$.*

Lemma 45. *Considering the eigenvalues of the complement graph indicated in Lemma 44, it follows that the set of eigenvectors associated to the eigenvalues $\{0, n-\mu_n, \ldots, n-\mu_\ell\}$ with $\mu_\ell \neq 0$, are given by $v_1 = \mathbf{1}, v_2 = u_n, \ldots, u_\ell$.*

Proof. Let us consider the Laplacian matrix of the complement graph $\bar{L} = nI - J - L$, where J is the all ones matrix,. Then, let us consider $\mu_\ell \neq 0$ the eigenvalue of L with eigenvector u_ℓ, then

$$\bar{L}u_\ell = nIu_\ell - Ju_\ell - Lu_\ell$$

$$\bar{L}u_\ell = nu_\ell - \mu_\ell u_\ell = (n - \mu_\ell) u_\ell$$

Additionally, considering the vector $\mathbf{1}$ we do have

$$\bar{L}\mathbf{1} = n\mathbf{1} - n\mathbf{1} - L\mathbf{1} = 0$$

□

Lemma 46. *Let us consider the complete graphs K_i with $i = 1, \ldots, q$ and $|V_{K_1}| < |V_{K_2}| < \ldots < |V_{K_q}|$. Then the set of eigenvalues of $G = K_1 \cup K_2 \cup \cdots \cup K_q$ is given by*

$$\left\{0^q, |V_{K_1}|^{|V_{K_1}|-1}, |V_{K_2}|^{|V_{K_2}|-1}, \ldots, |V_{K_q}|^{|V_{K_q}|-1}\right\} \quad (4.27)$$

with associated eigenvectors

$$\left\{\mathbf{1}_{n\times 1}, \tilde{u}_1^{K_1}, \ldots, \tilde{u}_1^{K_{q-1}}, \tilde{u}_2^{K_1}, \ldots, \tilde{u}_{|V_{K_1}|}^{K_1}, \tilde{u}_2^{K_2}, \ \ldots \tilde{u}_{|V_{K_2}|}^{K_2}, \ldots, \tilde{u}_2^{K_q}, \ldots \tilde{u}_{|V_{K_q}|}^{K_q}\right\}, \quad (4.28)$$

where $n = \sum_{i=1}^{q} |V_{K_i}|$, *and*

$$\tilde{u}_i^{K_j} = \begin{cases} 0 & \ell \notin V_{K_j} \\ u_i(\ell) & \ell \in V_{K_j}, \end{cases} \quad (4.29)$$

with $i \geq 2$, $j = 2, \ldots, q$ *and where* u_i *denotes* i^{th} *Laplalcian eigenvector of* K_j.

Proof. By lemma 43, it is immediate that the eigenvalues of $G = K_1 \cup K_2 \cup \cdots \cup K_q$ are given by

$$\left\{0^q, |V_{K_1}|^{|V_{K_1}|-1}, |V_{K_2}|^{|V_{K_2}|-1}, \ldots, |V_{K_q}|^{|V_{K_q}|-1}\right\} \quad (4.30)$$

Now, since

$$L_G = \begin{bmatrix} L_{K_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & L_{K_2} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & L_{K_q} \end{bmatrix}, \quad (4.31)$$

it follows that

$$L_G \tilde{u}_i^{K_j} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ L_{K_j} u_i \\ \vdots \\ \mathbf{0} \end{bmatrix} = \lambda_i^{K_j} \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ u_i \\ \vdots \\ \mathbf{0} \end{bmatrix} \quad (4.32)$$

where $\lambda_i^{K_j} \in \{0, |V_{K_j}|^{|V_{K_j}|-1}\}$ and u_i is the i^{th} eigenvector associated to $\lambda_i^{K_j}$. Additionally $L_G\mathbf{1} = 0$. □

Theorem 47. *Let* $M = K_1^c \vee K_2^c \vee \cdots \vee K_q^c$ *with* $|V_{K_1}| < |V_{K_2}| < \ldots < |V_{K_q}|$ *the complete multipartite graph. Then the set of Laplacian eigenvalues is given by*

$$\left\{0, (n - |V_{K_q}|)^{|V_{K_q}|-1}, \ldots, (n - |V_{K_2}|)^{|V_{K_2}|-1}, (n - |V_{K_1}|)^{|V_{K_1}|-1}, n^{q-1}\right\} \quad (4.33)$$

and the set of eigenvectors associated to

$$\left\{0, (n-|V_{K_q}|)^{|V_{K_q}|-1}, \dots, (n-|V_{K_2}|)^{|V_{K_2}|-1}, (n-|V_{K_1}|)^{|V_{K_1}|-1}\right\} \quad (4.34)$$

is given by

$$\left\{\mathbf{1}_{n\times 1}, \tilde{u}_2^{K_q}, \dots, \tilde{u}_{|V_{K_q}|}^{K_q}, \dots, \tilde{u}_2^{K_2}, \dots \tilde{u}_{|V_{K_2}|}^{K_2}, \tilde{u}_2^{K_1}, \dots, \tilde{u}_{K_1}^{K_1}\right\} \quad (4.35)$$

Proof. $M = (K_1 \cup K_2 \cup \cdots \cup K_q)^c$. The result follows by applying Lemmas 46, 43, and 44 to M. □

4.7.8.1 Proof of Multipartite uniqueness sets

Proof. Let us take into account the results of lemma 47, that describes the set of eigenvectors of the Laplacian matrix for all eigenvalues except n^q and its pictorial representation in Figure 4.8.

In Figure 4.8 It is possible to see the structure of the support of the set of vectors that span each PW_ω space. For example on left side of the gray line there is just the vector $\mathbf{1}_{n\times 1}$, which indicates that the Paley-Weiner space PW_ω for $0 \leq \omega < n - |V_{K_q}|$ is generated by $\mathbf{1}_{n\times 1}$. On the left side of the magenta line, it is possible to see the vectors that span PW_ω for $n-|V_{K_{q-r}}| \leq \omega < n-|V_{K_{q-(r+1)}}|$ which includes the vectors whose support is in magenta color and the vector $\mathbf{1}_{n\times 1}$. The use of this figure allows to determine how the rows of the submatrices must be selected in order to guarantee a full rank matrix.

1. When $0 \leq \omega < n - |V_{K_q}|$, we do have that $PW_\omega = span(\mathbf{1}_{n\times 1})$, therefore any node in M is a minimal uniqueness set for PW_ω.

2. When $n-|V_{K_{q-r}}| \leq \omega < n-|V_{K_{q-(r+1)}}|$ it is possible to see in Figure 4.8 that $PW_\omega = span(\mathbf{1}_{n\times 1}, \tilde{u}_2^{K_q}, \dots, \tilde{u}_{|V_{K_q}|}^{K_q}, \dots, \tilde{u}_2^{K_{q-r}}, \dots \tilde{u}_{|V_{K_2}|}^{K_{q-r}})$. Additionally, from the figure, notice that:

- The selection of rows from the matrix (nodes in the graph) should be done in such a way that no zero column is obtained in the resultant submatrix. This happens only when at least $|V_{K_i}| - 1$ nodes (rows) are selected in each V_{K_i} involved in the span of PW_ω.
- The size of the sampling sets is given by $1 + \sum_{\ell=0}^{r}(|V_{K_{q-\ell}}| - 1)$, and therefore it is possible to select one node v out of the V_{K_i} involved in the span of PW_ω.

3. If $n \leq \omega$ $PW_\omega(M)$ is generated by the complete set of eigenvectors, therefore the size of the sampling set must be n and all nodes in the graph are part of the minimal uniqueness set.

□

4.7.9 Proof of Theorem 38

Proof. Let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k$ denote the eigenvalues of G that are smaller or equal to λ. By Theorem 37, each of the eigenvalues $\lambda_2, \ldots, \lambda_k$ is of the form $\mu_i + n$ or $\nu_i + m$ for some eigenvalue μ_i of L_{G_1}, or some eigenvalue ν_i of L_{G_2}. Let $\{\lambda_2, \ldots, \lambda_k\} = \Lambda_1 \cup \Lambda_2$, with $\Lambda_1 = \{\mu_2 + n, \ldots, \mu_{k_1} + n\}$, and $\Lambda_2 = \{\nu_2 + m, \ldots, \nu_{k_2} + m\}$. Let $u_2 \oplus \mathbf{0}_n, \ldots, u_{k_1} \oplus \mathbf{0}_n$ and $\mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{k_2}$ denote the eigenvectors of L_G that correspond to the eigenvalues in Λ_1 and Λ_2 respectively. We consider each of the four case separately.

a) If $m > \lambda$, then $\dim \mathrm{PW}_\lambda(G) = 1$ and only contains constant signals. Thus S is a minimal uniqueness set if and only if $|S| = 1$.

b) If $m \leq \lambda < n$, then

$$\mathrm{PW}_\lambda(G) = \mathrm{span}(\mathbf{1}_{m+n}, \mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{k_2}).$$

Consider the matrix W with columns $\mathbf{1}_{m+n}, \mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{k_2}$. If $S_1 = \emptyset$, then it follows easily by Corollary 17 that S_2 is a minimal uniqueness set for $\mathrm{PW}_\lambda(G)$ if and only if it is a minimal uniqueness set for $\mathrm{PW}_{\lambda-m}(G_2)$. If $|S_1| \geq 2$, then two rows of $W_{S_1,\{1,\ldots,k_2\}}$ are equal and Corollary 17 implies that S is not a uniqueness set. If $|S_1| = 1$,

then it is not hard to see that the columns of $W_{S,\{1,\dots,k_2\}}$ are linearly independent if and only if the columns of $W_{S_2,\{2,\dots,k_2\}}$ are linearly independent. It follows that S is a minimal uniqueness set for $\mathrm{PW}_\lambda(G)$ if and only if S_2 is a minimal uniqueness set for $\mathrm{PW}^0_{\lambda-m}(G_2)$, as claimed.

c) Now, suppose that $n \leq \lambda < m+n$. Let S be a minimal uniqueness set for $\mathrm{PW}_\lambda(G)$. Note that $\dim \mathrm{PW}_\lambda(G) = k = k_1 + k_2 - 1$. Hence, $|S| = k_1 + k_2 - 1$ by Corollary 17. We first claim that $|S_1| \geq k_1 - 1$ and $|S_2| \geq k_2 - 1$. Indeed, suppose $|S_1| < k_1 - 1$. Then the vectors $(u_2)_{S_1}, \dots (u_{k_1})_{S_1}$ are linearly dependent, say $\sum_{i=2}^{k_1} t_i (u_i)_{S_1} = 0$ with not all t_i equal to zero. Consider the signal $f = \sum_{i=2}^{k_1} t_i u_i \oplus \mathbf{0}_n \in \mathrm{PW}_\lambda(G)$. Clearly, $f \equiv 0$ on S. Hence $f \equiv 0$ on $V_1 \cup V_2$ since S is a uniqueness set for $\mathrm{PW}_\lambda(G)$. This contradicts the linear independence of $u_2, \dots, u_{k_1}$. We therefore must have $|S_1| \geq k_1 - 1$. A similar argument shows that $|S_2| \geq k_2 - 1$. Hence, either $|S_1| = k_1$ and $|S_2| = k_2 - 1$, or $|S_1| = k_1 - 1$ and $|S_2| = k_2$. Assume $|S_1| = k_1$ and $|S_2| = k_2 - 1$. Consider the matrix W with columns $\mathbf{1}_{m+n}, u_2 \oplus \mathbf{0}_n, \dots, u_{k_1} \oplus \mathbf{0}_n, \mathbf{0}_m \oplus v_2, \dots, \mathbf{0}_m \oplus v_{k_2}$. Observe that W can be written in block form:

$$W = \begin{pmatrix} U_1 & \mathbf{0}_{m\times(k_2-1)} \\ X & U_2 \end{pmatrix} \tag{4.36}$$

where $U_1 \in \mathbb{R}^{m\times k_1}$, $U_2 \in \mathbb{R}^{n\times(k_2-1)}$, and $X \in \mathbb{R}^{n\times k_1}$. By Lemma 42, the matrix $W_{S,\{1,\dots,k\}}$ is invertible if and only if $(U_1)_{S_1,\{1,\dots,k_1\}}$ and $(U_2)_{S_2,\{1,\dots,k_2-1\}}$ are invertible. It follows by Corollary 17 that S_1 is a minimal uniqueness set for $\mathrm{PW}_{\lambda-n}(G_1)$ and S_2 a uniqueness set for $\mathrm{PW}^0_{\lambda-m}(G_2)$. If instead $|S_1| = k_1 - 1$ and $|S_2| = k_2$ then a similar argument shows that S_1 is a minimal uniqueness set for $\mathrm{PW}^0_{\lambda-n}(G_1)$ and S_2 a uniqueness set for $\mathrm{PW}_{\lambda-m}(G_2)$. Conversely, suppose without loss of generality that S_1 is a uniqueness set for $\mathrm{PW}_{\lambda-n}(G_1)$ and S_2 a uniqueness set for $\mathrm{PW}^0_{\lambda-m}(G_2)$. Then the matrices $(U_1)_{S_1,\{1,\dots,k_1\}}$ and $(U_2)_{S_2,\{1,\dots,k_2-1\}}$ in Equation (4.36) are invertible. Thus $W_{S,\{1,\dots,k\}}$ is invertible and S is a uniqueness set by Corollary 17.

d) Finally, if $\lambda \geq m + n = |V(G)|$, then Theorem 37 shows that $\mathrm{PW}_\lambda(G) = \mathbb{R}^{m+n}$. Thus $\dim \mathrm{PW}_\lambda(G) = m + n$ and it follows easily from Corollary 17 that $S = V_1 \cup V_2$ is the only uniqueness set for $\mathrm{PW}_\lambda(G)$. □

4.7.10 Proof of Theorem 39

Proof. The proof follows the proof of Theorem 38, with some minor adjustments. With the same notation as in that proof, we consider each of the three case separately.

a) If $m \leq \lambda < n$, then

$$\mathrm{PW}^0_\lambda(G) = \mathrm{span}(\mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{k_2}).$$

Consider the matrix W with columns $\mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{k_2}$. If $S_1 = \emptyset$, then it follows easily by Corollary 18 that S_2 is a minimal uniqueness set for $\mathrm{PW}^0_\lambda(G)$ if and only if it is a minimal uniqueness set for $\mathrm{PW}^0_{\lambda-m}(G_2)$. If $|S_1| \geq 1$, then $W_{S,\{1,\ldots,k_2-1\}}$ contains a row of zeros and S is not a uniqueness set.

b) Suppose $n \leq \lambda < m + n$. Let S be a minimal uniqueness set for $\mathrm{PW}^0_\lambda(G)$. Note that $\dim \mathrm{PW}_\lambda(G) = k_1 + k_2 - 2$. Hence, $|S| = k_1 + k_2 - 2$ by Corollary 18. We first claim that $|S_1| \geq k_1 - 1$ and $|S_2| \geq k_2 - 1$. Indeed, suppose $|S_1| < k_1 - 1$. Then the vectors $(u_2)_{S_1}, \ldots (u_{k_1})_{S_1}$ are linearly dependent, say $\sum_{i=2}^{k_1} t_i (u_i)_{S_1} = 0$ with not all t_i equal to zero. Consider the signal $f = \sum_{i=1}^{k_1} t_i u_i \oplus \mathbf{0}_n \in \mathrm{PW}^0_\lambda(G)$. Clearly, $f \equiv 0$ on S. Hence $f \equiv 0$ on $V_1 \cup V_2$ since S is a uniqueness set for $\mathrm{PW}^0_\lambda(G)$. This contradicts the linear independence of $u_2, \ldots, u_{k_1}$. We therefore must have $|S_1| \geq k_1 - 1$. A similar argument shows that $|S_2| \geq k_2 - 1$. We therefore have $|S_1| = k_1 - 1$ and $|S_2| = k_2 - 1$. Now, consider the matrix W with columns $u_2 \oplus \mathbf{0}_n, \ldots, u_{k_1} \oplus \mathbf{0}_n, \mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{k_2}$. Observe that W can be written in block form:

$$V = \begin{pmatrix} U_1 & \mathbf{0}_{m \times (k_2-1)} \\ \mathbf{0}_{n \times (k_1-1)} & U_2 \end{pmatrix} \tag{4.37}$$

where $U_1 \in \mathbb{R}^{m \times (k_1-1)}$, and $U_2 \in \mathbb{R}^{n \times (k_2-1)}$. Clearly, the matrix $W_{S,\{1,\ldots,k_1+k_2-2\}}$ is invertible if and only if $(U_1)_{S_1,\{1,\ldots,k_1-1\}}$ and $(U_2)_{S_2,\{1,\ldots,k_2-1\}}$ are invertible. It follows by Corollary 18 that S_1 is a minimal uniqueness set for $\mathrm{PW}^0_{\lambda-n}(G_1)$ and S_2 a uniqueness set for $\mathrm{PW}^0_{\lambda-m}(G_2)$. Conversely, suppose that S_1 is a uniqueness set for $\mathrm{PW}^0_{\lambda-n}(G_1)$ and S_2 a uniqueness set for $\mathrm{PW}^0_{\lambda-m}(G_2)$. Then the matrices $(U_1)_{S_1,\{1,\ldots,k_1-1\}}$ and $(U_2)_{S_2,\{1,\ldots,k_2-1\}}$ in Equation (4.37) are invertible. Thus $W_{S,\{1,\ldots,k_1+k_2-2\}}$ is invertible and S is a uniqueness set by Corollary 18.

c)Finally, assume $\lambda \geq m+n$. Then $\mathrm{PW}^0_\lambda(G) = \mathbf{1}^\perp_{m+n}$. Let $\mathbf{1}_{m+n}, u_2 \oplus \mathbf{0}_n, \ldots, u_{m-1} \oplus \mathbf{0}_n, \mathbf{0}_m \oplus v_2, \ldots, \mathbf{0}_m \oplus v_{n-1}, -n\mathbf{1}_m \oplus m\mathbf{1}_n$ be an orthogonal basis of eigenvectors of L_G. Denote these vectors by $\mathbf{1}_{m+n} = w_1, w_2, \ldots, w_{m+n} = -n\mathbf{1}_m \oplus m\mathbf{1}_n$ and let W be the matrix with columns $w_1, \ldots, w_{m+n}$. By Corollary 18, any minimal uniqueness set for $\mathrm{PW}^0_\lambda(G)$ had size $m+n-1$. Let S be any subset of $[m+n]$ of size $m+n-1$, say $S = [m+n] \setminus \{p\}$ for some $p \in [m+n]$. By Corollary 18, it suffices to show that the determinant of the matrix obtained from W by deleting its p-th row and its first column is nonzero. Let $X = W^{-1}$. Using properties of the adjunct matrix, this is equivalent to showing that $X_{1,p} \neq 0$. Let $x_1 = (x_{1,i})_{1 \leq i \leq m+n}$ denote the first row of X. Then x_1 is characterized by:

$$\sum_{i=1}^{m+n} x_{1,i} = 1 \tag{4.38}$$

$$(x_1)_{V_1} \perp u_i \qquad 2 \leq i \leq m \tag{4.39}$$

$$(x_1)_{V_2} \perp v_i \qquad 2 \leq i \leq n-1 \tag{4.40}$$

$$-n\sum_{i=1}^{m} x_{1,i} + m\sum_{i=m+1}^{m+n} x_{1,i} = 0, \tag{4.41}$$

where $(x_1)_{V_1}$ and $(x_1)_{V_2}$ denote the restriction of x_1 to its first m and its last n entries respectively. Now, since $u_2, \ldots, u_{m-1}$ form a basis of the orthogonal complement of $\mathbf{1}_m$ in $\mathbb{R}^m$, we conclude by Equation (4.39) that $(x_1)_{V_1} = \alpha\mathbf{1}_m$ for some $\alpha \in \mathbb{R}$. Similarly,using Equation (4.40), we obtain $(x_1)_{V_2} = \beta\mathbf{1}_n$ for some $\beta \in \mathbb{R}$. Now, by Equation (4.38) and (4.41), we have

$$m\alpha + n\beta = 1$$
$$-nm\alpha + mn\beta = 0.$$

Solving for α, β, we obtain

$$\alpha = \beta = \frac{1}{m+n} \neq 0.$$

Therefore, S is a minimal uniqueness set, as claimed. □

Chapter 5

COLORED CODED APERTURE DESIGN IN COMPRESSIVE SPECTRAL IMAGING VIA MINIMUM COHERENCE

5.1 Introduction

Spectral imaging is extensively used in remote sensing applications and thus the development of more efficient sensing architectures is of interest, including compressed sensing approaches such as coded aperture spectral imaging (CASSI). Multispectral imaging based on CASSI has received considerable interest in recent years. In particular, the design of coded apertures has been shown to be a key to increase the quality of the reconstructions [69]. Whereas initial designs of the CASSI considered the use of random binary coded apertures [70] [71], it has been proven that the use of colored coded apertures can significantly improve the quality of the reconstructions [72] [2]. Colored coded apertures are two dimensional arrays of pixels that have selective spectral response, i.e. each pixel allows or blocks specific parts of the light spectrum. It is fabricated as a patterned multilayer optical coating, that allows to have a compact two dimensional array of pixels, each one with an specific spectral response [73].

This paper addresses the design and optimization of compressive spectral imagers with this new coded aperture technology. The solution proposed is based on the analysis of the coherence of the sensing matrix. It is found that exploiting the highly structured transfer function matrix of the system, leads to a solution that can be obtained as the set of designs of smaller submatrices. It is shown that diverse families of codes exists that can be considered optimal in the sense that an upper bound of the coherence is minimized. The obtained codes are tested by simulation against traditional random binary codes and the codes obtained with other methods [2], showing that the results obtained with the approach proposed in the present work exhibit a

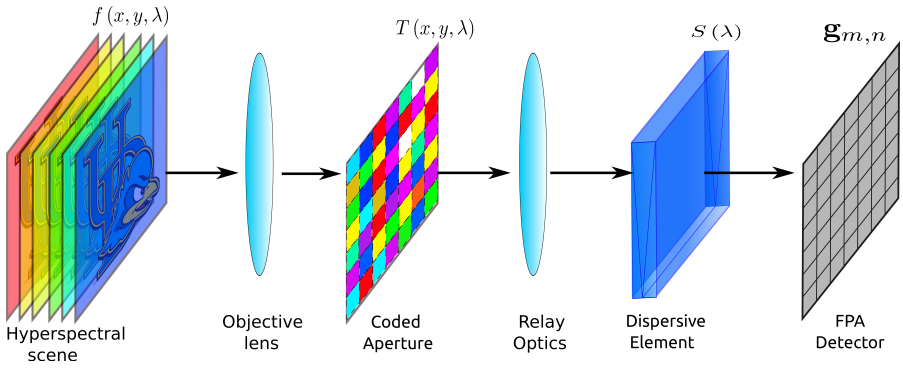


Figure 5.1: Components of the CASSI with the use of a general coded aperture $T(x, y, \lambda)$. The multispectral scene $f(x, y, \lambda)$ is modulated by $T(x, y, \lambda)$, and the resultant modulated field is dispersed by an Amici prism with dispersion curve $S(\lambda)$ to be finally integrated in the Focal Plane Array (FPA). When the classical binary coded aperture is used $T(x, y, \lambda) \in \{0, 1\}\ \forall\lambda$.

clear improvement in the quality of the reconstructions. Additionally, the methods introduced present a simple and fast way to generate optimized codes with high diversity in the kind of spectral responses in the colored coded aperture, reducing the large scale optimization problem to a subset of smaller problems, which is also an advantage with respect to previous designs that are obtained using the analysis of the RIP constant [2] using genetic type optimization algorithms, leading to limited specific constraints in the spectral responses of the colored coded apertures. The designed codes based on the coherence are tested against random codes in the presence of noise, showing that despite the fact that the coherence is considered less suited than the RIP constant to deal with the noise, the resultant codes lead to reconstructions of higher quality than the obtained with random binary codes.

5.2 CASSI Modeling

The components of CASSI are shown in Figure 7.1. The multispectral scene $f(x, y, \lambda)$ is coded by a coded aperture $T(x, y, \lambda)$, where (x, y) indicates the spatial coordinates and λ is the wavelength component. If traditional binary coded apertures are used, then $T(x, y, \lambda) \in \{0, 1\} \quad \forall\lambda \in \Lambda$, where Λ is the spectral range of $f(x, y, \lambda)$ [69] [71]. When $T(x, y, \lambda)$ is a non constant and non negative function with

respect to λ, $T(x, y, \lambda)$ represents colored coded apertures which modulates the incoming light field both in space and wavelength [74]. Once $f(x, y, \lambda)$ has been modulated by $T(x, y, \lambda)$, the resultant field is dispersed by a prism and $q(x, y, \lambda)$ is obtained as $q(x, y, \lambda) = f\left(x - S(\lambda), y, \lambda\right) T\left(x - S(\lambda), y, \lambda\right)$, where $S(\lambda)$ is the dispersion curve of the prism. The continuous representation of the compressed measurements $g(x, y)$ in the Focal Plane Array (FPA) are obtained by the integration of $q(x, y, \lambda)$ across the spectral axes as $g(x, y) = \int_{\Lambda} q(x, y, \lambda) d\lambda$ [69] [74].

5.2.1 Discretization of the Model

The focal plane array (FPA) measures a sampled set of points. In order to adapt the continuous model to a discrete sampled formulation, all operators are approximated [74]. Let $\Omega \in \mathbb{R}^2$ be the spatial domain of the FPA, represented as $\Omega = \cup_{m,n \in [N]} \Omega_{m,n}$, where $\Omega_{m,n}$ is given by

$$\Omega_{m,n} = \{ (x, y) | \, \Delta(n - 1/2) \leq x \leq \Delta(n + 1/2) \, ,$$

$$\Delta(m - 1/2) \leq y \leq \Delta(m + 1/2) \} \, ,$$

and $[N] = \{0, 1, \ldots N - 1\}$. The FPA measurement is then represented as an $N \times (N + L - 1)$ array conformed by the union of $N(N + L - 1)$ domains of size $\Delta \times \Delta$. The pixel (m, n) with pitch size Δ is associated with $\Omega_{m,n}$, and its measurement sample is

$$\mathbf{g}_{m,n} = \iint_{\Omega_{m,n}} g(x, y) \text{rect}\left(\frac{x}{\Delta} - n, \frac{y}{\Delta} - m\right) dy dx, \tag{5.1}$$

which represents the contribution of $g(x, y)$ to each $\Omega_{m,n}$. The multispectral scene $f(x, y, \lambda)$ is modulated by $T(x, y, \lambda)$ such that

$$f(x, y, \lambda) T(x, y, \lambda) \approx \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (TF)_{m,n}(\lambda) \text{rect}\left(\frac{x}{\Delta} - n, \frac{y}{\Delta} - m\right), \tag{5.2}$$

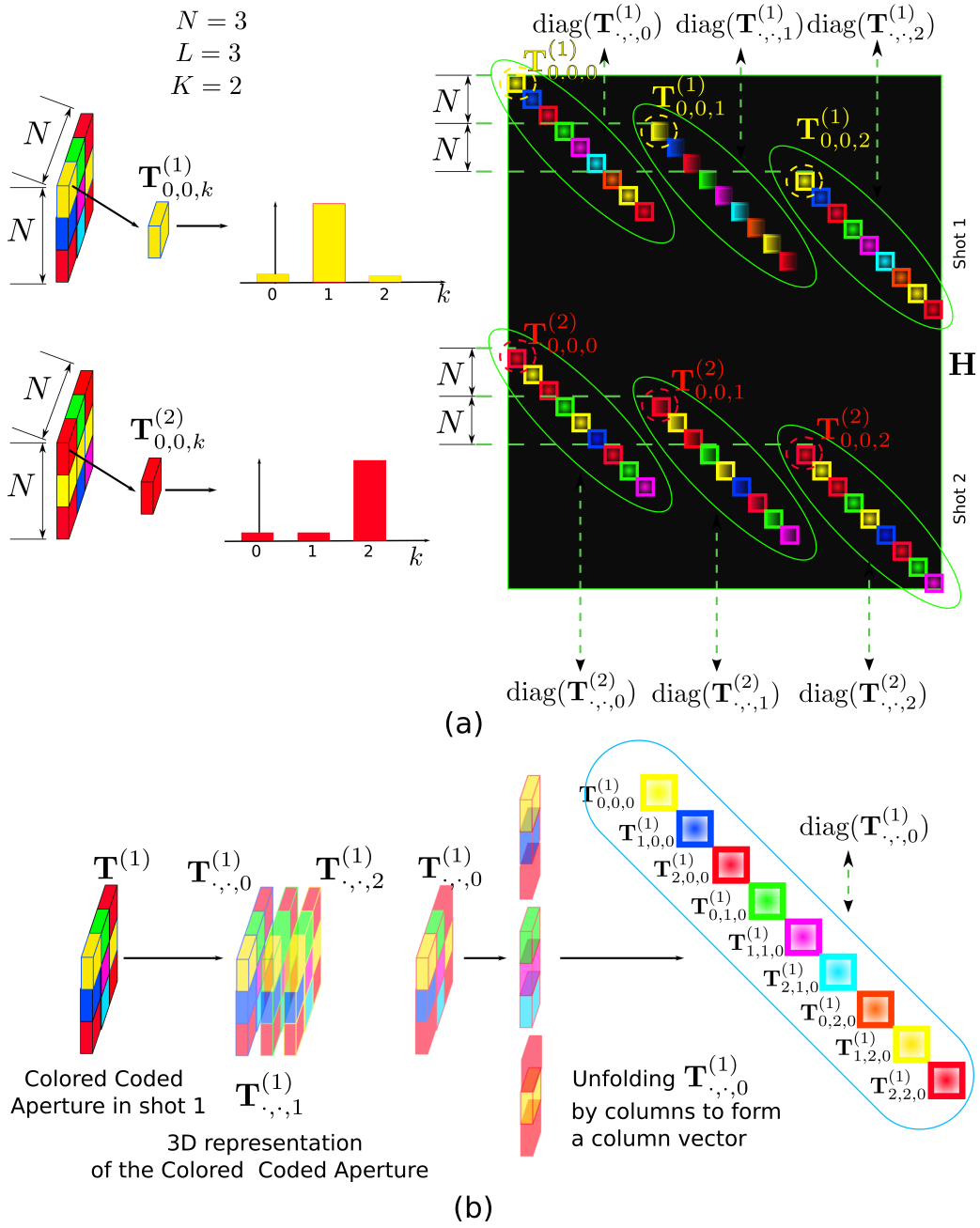


Figure 5.2: The graphic representation of the transfer function matrix $\mathbf{H}$ is depicted, when the capture of one multispectral scene of $L = 3$ bands and $N = 3$ is considered using $K = 2$ shots. On the indicated diagonals the spectral response of the pixels of the coded aperture are shown for each band respectively, i.e. on the diagonal of the band i, the lexicographic ordering of the elements of $\mathbf{T}_{\cdot,\cdot,i}$ is considered.

where $(TF)_{m,n}(\lambda) = T_{m,n}(\lambda)F_{m,n}(\lambda)$ and $T_{m,n}(\lambda)$ is the spectral response of the coded aperture at pixel (m,n), and where $F_{m,n}(\lambda)$ is given by $F_{m,n}(\lambda) = \iint\limits_{\Omega_{m,n}} f(x,y,\lambda)dxdy$. Taking into account these facts and using (1), $\mathbf{g}_{m,n}$ can be written as

$$\mathbf{g}_{m,n} = \sum_{k=0}^{\min\{n,L-1\}} (\mathbf{T} \circ \mathbf{F})_{m,n-k,k}, \tag{5.3}$$

Appendix 5.7 provides a detailed derivation of (7.4). $(\mathbf{T} \circ \mathbf{F})_{m,n,k} = \mathbf{T}_{m,n,k}\mathbf{F}_{m,n,k}$ where $\mathbf{T}_{m,n,k}$ and $\mathbf{F}_{m,n,k}$ are the (m,n,k) elements of the arrays $\mathbf{T}$ and $\mathbf{F}$, respectively, where the first two dimensions represent the spatial location and the third one indicates the spectral component. The term $\mathbf{T}_{m,n,k}$ is the spectral response of the coded aperture at pixel (m,n) at wavelength k. Equation (7.4) can be rewritten in matrix notation as $\vec{\mathbf{g}} = \mathbf{H}\vec{\mathbf{F}}$, where $\vec{\mathbf{g}}$ is the vectorized representation of $\mathbf{g}$. $\mathbf{H}$ is of dimension $KN\,(N+L-1)\times N^2L$ with K being the number of shots and $\vec{\mathbf{F}}$ is the vectorized form of $\mathbf{F}$ (see Fig. 5.2). Figure 5.2 shows the structure of $\mathbf{H}$ for CASSI with colored coded apertures. The vectorization of 2-dimensional arrays is performed by columns. For 3-dimensional arrays like $\mathbf{F} \in \mathbb{R}^{N\times N\times L}$, the vectorization is performed concatenating vertically the vectorizations of each $\mathbf{F}(:,:,i)$ with $i = 1,\ldots,L$.

When the sparsity properties of the signal $\vec{\mathbf{F}}$ in a basis $\boldsymbol{\psi}$ are used, the problem can be written as

$$\vec{\mathbf{g}} = \mathbf{H}\boldsymbol{\psi}\mathbf{f}, \tag{5.4}$$

where $\vec{\mathbf{F}} = \boldsymbol{\psi}\mathbf{f}$ and $\mathbf{f}$ is a column vector whose entries are the coefficients representation in the basis. The recovery of $\mathbf{f}$ is obtained as the solution of the nonlinear optimization problem [75] [71]

$$\hat{\mathbf{f}} = \underset{\mathbf{z}}{\operatorname{argmin}} \frac{1}{2} \left\|\mathbf{A}\mathbf{z} - \vec{\mathbf{g}}\right\|_2^2 + \tau \left\|\mathbf{z}\right\|_1, \tag{5.5}$$

where $\mathbf{A} = \mathbf{H}\boldsymbol{\psi}$ is the sensing matrix of the problem and τ a regularization parameter.

5.3 Coherence of the Sensing Matrix

Two measures for the quality are often used to describe the effectiveness of compressed sensing projections, the coherence of the sensing matrix and the *Restricted Isometry Property* (RIP) [75] [76] [4]. The RIP is traditionally used in theoretical analysis of compressed sensing because of its elegance. However, even with simple representations of the sensing matrix, it is in general difficult to calculate [75] [4]. On the other hand, the coherence offers a measure of quality that is often more tractable. In particular it allows one to exploit the structure of the sensing matrix, and it also gives the degree of ill-posedness of the system. The coherence of the sensing matrix $\mathbf{A} = \mathbf{H}\boldsymbol{\psi}$ is defined as the maximum absolute value of the inner product between any two columns of $\mathbf{A}$, with each column normalized by its ℓ_2-norm. A good sensing matrix will have the coherence as small as possible to guarantee uniqueness of the solution [77] [76].

In the following, a detailed calculation of the coherence is presented showing how the structure of $\mathbf{H}$ can be exploited, in turn, leading to the proposed optimization framework.

5.3.1 Matrix Formulation and Analysis of the Coherence

The structure of $\mathbf{H}$ is depicted in Fig. 5.2. The nonzero elements lie on the indicated diagonals representing the spectral response of the coded aperture in each band.

Let $\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \ldots \mathbf{h}_{N^2L}]$, where each column $\mathbf{h}_i$ is of dimension $KN\,(N+L-1)\times 1$. The basis matrix $\boldsymbol{\Psi}$ can be written as $\boldsymbol{\Psi} = \left[\boldsymbol{\psi}_1^T, \quad \boldsymbol{\psi}_2^T, \cdots, \boldsymbol{\psi}_{N^2L}^T\right]^T$, where each $\boldsymbol{\psi}_i$ is of dimension $1 \times N^2L$. Then, the sensing matrix can be represented as $\mathbf{A} = \mathbf{H}\boldsymbol{\Psi} = \sum_{i=1}^{N^2L} \mathbf{h}_i\boldsymbol{\psi}_i$, and the element of $\mathbf{A}$ in the (m,n) position can be written as $\mathbf{A}(m,n) = \sum_{i=1}^{N^2L} \mathbf{h}_i(m)\boldsymbol{\psi}_i(n)$. The inner product between the m and n columns of $\mathbf{A}$ is given by

$$\langle \mathbf{A}(\cdot,m), \mathbf{A}(\cdot,n) \rangle = \sum_{i,j=1}^{N^2L} \langle \mathbf{h}_i, \mathbf{h}_j \rangle \, \boldsymbol{\psi}_i(m) \boldsymbol{\psi}_j^*(n).$$

The coherence of the sensing matrix $\mathbf{A}$ can be written as

$$\mu(\mathbf{A}) = \max_{\substack{m,n \\ m \neq n}} \left| \left\langle \frac{\mathbf{A}(\cdot,m)}{\|\mathbf{A}(\cdot,m)\|}, \frac{\mathbf{A}(\cdot,n)}{\|\mathbf{A}(\cdot,n)\|} \right\rangle \right| \tag{5.6}$$

$$= \max_{\substack{m,n \\ m \neq n}} \mu_{m,n}(\mathbf{A}), \tag{5.7}$$

where

$$\mu_{m,n}(\mathbf{A}) = \frac{\left| \sum_{i,j=1}^{N^2L} \varphi_{i,j} R_{i,j}^{(m,n)} \right|}{\left(\sum_{i,j=1}^{N^2L} \varphi_{i,j} R_{i,j}^{(m,m)} \right)^{\frac{1}{2}} \left(\sum_{i,j=1}^{N^2L} \varphi_{i,j} R_{i,j}^{(n,n)} \right)^{\frac{1}{2}}}, \tag{5.8}$$

and $R_{i,j}^{(m,n)} = \boldsymbol{\psi}_i(m) \boldsymbol{\psi}_j^*(n)$, $\varphi_{i,j} = \langle \mathbf{h}_i, \mathbf{h}_j \rangle$.

From the structure of $\mathbf{H}$ (see Figure 5.2), it is possible to identify a set I of pairs of columns of $\mathbf{H}$, that can be written as $I = \{(i,j) | \, i \neq j, \varphi_{i,j} \neq 0\}$, that is the set of all possible pairs of columns of $\mathbf{H}$ whose inner products are different from zero. Taking into account I, equation (7.10) can be written as

$$\mu_{m,n}(\mathbf{A}) = \frac{|\varrho(m,n)|}{\varrho(m,m)^{\frac{1}{2}} \, \varrho(n,n)^{\frac{1}{2}}}, \tag{5.9}$$

where

$$\varrho(m,n) = \sum_{i=1}^{N^2L} \varphi_{i,i} R_{i,i}^{(m,n)} + \sum_{(i,j) \in I} \varphi_{i,j} R_{i,j}^{(m,n)}. \tag{5.10}$$

The coherence $\mu(\mathbf{A})$ in (5.7) is therefore determined by the functions $\varphi_{i,j}$ and $\varphi_{i,i}$. In the next sections, it is established how the values of $\varphi_{i,j}$ and $\varphi_{i,i}$ can be used in order to get an upper bound on the values of $\mu_{m,n}(\mathbf{A})$ and consequently $\mu(\mathbf{A})$.

5.4 Optimization of the H matrix

The intended values of $\mu(\mathbf{A})$ should be as small as possible in order to improve the quality of the reconstructions [78] [79]. The problem of designing $\mathbf{H}$ such that the minimum value of $\mu(\mathbf{A})$ is attained, can be formulated as

$$\begin{aligned} \underset{\mathbf{H}}{\text{minimize}} \quad & \mu(\mathbf{A}) \\ \text{subject to} \quad & \mathbf{H} \in \mathcal{C}_{N,L,K} \end{aligned} \tag{5.11}$$

where $\mathcal{C}_{N,L,K}$ is the set of matrices for a CASSI system reconstructing multispectral images with image size N, L bands, using K shots and, the entries of $\mathbf{H}$ are binary nonnegative. The expression for $\mu(\mathbf{A})$ is nonconvex [78] [80], therefore a direct solution of the problem (5.11) is not workable. However, it is possible to take into account the relationship between $\mu_{m,n}(\mathbf{A})$ and $\varrho(m,n)$ to get an alternative formulation. This relation is presented in the following theorem.

Theorem 48. *Let $\mathbf{A} = \mathbf{H}\boldsymbol{\psi}$ be the sensing matrix for the CASSI system considering the reconstruction of multispectral scenes of size $N \times N \times L$ using K shots. Then, there is $\xi \in \mathbb{R}_+$ such that*

$$\mu_{m,n}(\mathbf{A}) \leq \xi \left[\left(\sum_{i=1}^{N^2 L} \varphi_{i,i} \right) \max_i \left| R_{i,i}^{(m,n)} \right| + \left(\sum_{(i,j)\in I} \varphi_{i,j} \right) \max_{(i,j)\in I} \left| R_{i,j}^{(m,n)} \right| \right] \tag{5.12}$$

Proof: See Appendix 5.8.

The quantities from (7.12) $\varphi_{i,i}$ and $\varphi_{i,j}$ play a key role in bounding the value of $\mu_{m,n}(\mathbf{A})$. In particular, note that minimizing $\varphi_{i,i}$ and $\varphi_{i,j}$ implies a minimization of the right side of equation (5.12). This reduces the range of values of $\mu(\mathbf{A})$, and therefore it represents an indirect minimization of $\mu(\mathbf{A})$ that allows the finding of local minimums.

The maximum values of $\varphi_{i,j}$ can be determined by the use of the Cauchy-Schwarz inequality as $\varphi_{i,j} = |\langle \mathbf{h}_i, \mathbf{h}_j \rangle| \leq \|\mathbf{h}_i\|_2 \|\mathbf{h}_j\|_2$ and for $\varphi_{i,i}$ it follows that $\varphi_{i,i} = \|\mathbf{h}_i\|_2^2$. Then, the values of the functions $\varphi_{i,i}$ and $\varphi_{i,j}$ are bounded by the ℓ_2-norm of the columns $\mathbf{h}_i$. Because all the vectors $\mathbf{h}_i$ have binary nonnegative components the minimum value of $\|\mathbf{h}_i\|_2 \neq 0$ for an arbitrary i is attained when $\|\mathbf{h}_i\|_2 = 1$.

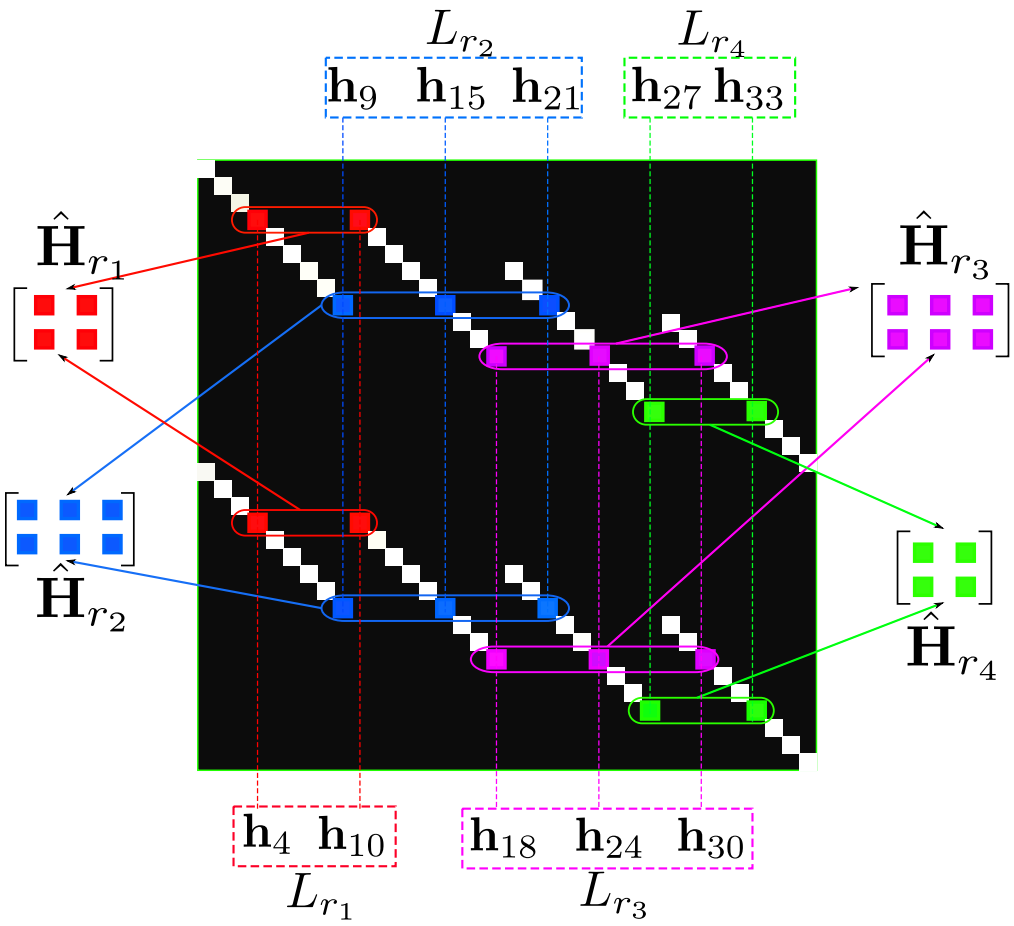


Figure 5.3: The structure of $\mathbf{H}$ is depicted showing how the different sets I_r are defined and how its structure is related with the submatrices $\hat{\mathbf{H}}_r$. The support of some subset of columns of $\mathbf{H}$ for which $\varphi_{i,j} \neq 0$ is indicated in different colors, and it is shown how is its relation with $\hat{\mathbf{H}}_r$.

Taking into account this, the term $\varphi_{i,i}$ in (5.12) must be such that $\varphi_{i,i} = 1$. With this value of $\varphi_{i,i}$ it is possible to write an alternative formulation of (5.11) as

$$\begin{aligned} \underset{\mathbf{H}}{\text{minimize}} \quad & \sum_{(i,j)\in I} \varphi_{i,j} \\ \text{subject to} \quad & \varphi_{i,j} = \langle \mathbf{h}_i, \mathbf{h}_j \rangle \\ & \mathbf{H} \in \mathcal{C}_{N,L,K} \\ & \|\mathbf{h}_i\|_1 = 1. \end{aligned} \tag{5.13}$$

5.4.1 About the structure of H

The set I indicates the set of inner products between columns of $\mathbf{H}$ that can be different from zero, with $|I| \ll N^2L(N^2L-1)$. The structure of I can be additionally described as a union of simpler and smaller sets as

$$I = \bigcup_{r=1}^{R} I_r \tag{5.14}$$

where each I_r represents a subset of pairs of I that indicate the subset of columns of $\mathbf{H}$ that have the same support (see Figure 5.3), and R is the total number of subsets. As shown in Fig. 5.3, these sets have the property that

$$I_r \bigcap I_\ell = \emptyset \quad \forall r \neq \ell. \tag{5.15}$$

Hence the possible values of the inner products between columns of $\mathbf{H}$ are described by the set of inner products of the submatrices $\hat{\mathbf{H}}_r$. Because the structure of $\mathbf{H}$ is explicitly known as indicated in Fig. (5.2) and Fig. (5.3), it is possible to determine the sets I_r and the indexes L_r of columns that define I_r in a systematic way. In Appendix 5.12 an auxiliary algorithm used to obtain the indexes of the columns L_r is indicated.

The design of $\mathbf{H}$ can then be formulated as the design of the submatrices $\hat{\mathbf{H}}_r$. This idea is summarized in Algorithm 1. It is pointed out here that the structure of $\mathbf{H}$ indicates where the elements of $\mathbf{T}$ are located in $\mathbf{H}$, this has been stated in previous paragraphs and Figures (5.2) and (5.3).

Algorithm 7 Design of $\mathbf{H}$ based on the design of submatrices $\hat{\mathbf{H}}_r$

Input: structure of $\mathbf{H}$
Output: $\mathbf{H}$

1: Identify the sets L_r and the dimensions of the matrices $\hat{\mathbf{H}}_r$ (see Figure 5.3 and Appendix 5.12);
2: Design the submatrices $\hat{\mathbf{H}}_r$ for each r;
3: Assemble values of $\hat{\mathbf{H}}_r$ in $\mathbf{H}$ (see Figure 5.3 and Appendix 5.12);
4: **return** $\mathbf{H}$

The assembling of the values of $\hat{\mathbf{H}}_r$ in $\mathbf{H}$ can be obtained given the knowledge of the structure of $\mathbf{H}$ (see Fig. (5.2) and Fig. (5.3)). Additionally an auxiliary algorithm is presented in Appendix 5.12 showing how this task is done.

Different criteria can be applied in the design of the submatrices $\hat{\mathbf{H}}_r$ in order to get direct or relaxed solutions of (5.13). In the following subsections a direct solution of (5.13) is presented and a relaxed solution of this problem based on the coherence of $\mathbf{H}$ is also presented.

5.4.2 Minimum Coherence design of $\hat{\mathbf{H}}_r$

Before a direct minimization of $\sum_{(i,j)\in I}\varphi_{i,j}$ is obtained, a relaxed formulation of the problem is presented based on the minimization of the coherence of $\mathbf{H}$.

Despite the fact that the relationship between $\mu(\mathbf{H})$ and $\mu_{m,n}(\mathbf{A})$ is nonlinear, a minimized value of $\mu(\mathbf{H})$ is related to the minimization of $\sum_{(i,j)\in I}\varphi_{i,j}$ as $\sum_{(i,j)\in I}\varphi_{i,j} \leq |I|K\mu(\mathbf{H})$. Then as the value of $\mu(\mathbf{H})$ is reduced, the upper bound on $\sum_{(i,j)\in I}\varphi_{i,j}$ is also reduced, which represents an indirect way of minimization of $\sum_{(i,j)\in I}\varphi_{i,j}$. This fact makes any formulation in which $\mu(\mathbf{H})$ is minimized, a relaxed formulation of (5.13). Taking into account that $\mu(\mathbf{H}) = \max_r \mu(\hat{\mathbf{H}}_r)$ because the sets I_r are disjoint, the formulation using the minimum coherence criteria is presented as

$$\begin{aligned} \underset{\hat{\mathbf{H}}_r}{\text{minimize}} \quad & \mu(\hat{\mathbf{H}}_r) \quad \forall I_r \\ \text{subject to} \quad & \hat{\mathbf{h}}_i^{(r)} \in \{0,1\}^K \end{aligned} \tag{5.16}$$

where the I_r are specified according to (5.14) and $\hat{\mathbf{h}}_i^{(r)}$ is the i^{th} column of $\hat{\mathbf{H}}_r$. This formulation, which constitutes a set of disjoint formulations for each $\hat{\mathbf{H}}_r$, allows one to get $\mathbf{H}$ with the minimum coherence $\mu(\mathbf{H})$.

The algorithm proposed to solve this minimization problems in an efficient way, without extensively evaluating all possibilities in the feasible set, is presented below. It is shown how we can exploit the structure that the vectors on the unit hypercube have.

5.4.2.1 Minimum coherence matrix construction with nonnegative binary entries

The general procedure to build nonnegative binary matrices is presented here in order to be used for the design of the submatrices $\hat{\mathbf{H}}_r$. Let Δ_θ be defined as $\Delta_\theta =$

$\{\mathbf{x} \in Q_K \,|\|\mathbf{x}\|_{\ell_1} = \theta\}$ where Q_K represents the hypercube in the $K-$dimensional space described as $Q_K = \left\{\mathbf{x} = [x_1, \ldots, x_K]^T \in \mathbb{R}^K \,|x_i \in \{0,1\} \quad \forall i\right\}$. Then Δ_θ represents the set of binary vectors that have exactly θ of its components equal to 1 and the remaining components equal to 0. In particular it is important to take into account that

$$Q_K \setminus \{\mathbf{0}\} = \bigcup_{\theta=1}^{K} \Delta_\theta. \tag{5.17}$$

This means that the set of possible vectors that can be used as column vectors in the construction of a matrix can be represented as a union of the sets Δ_θ. Then, the representation of Q_K in (5.17) is exploited in order to calculate the coherence of any matrix whose columns are in Q_K. The following theorem exploits this last representation.

Theorem 49. *Consider* $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \bigcup_{i=1}^{m} \Delta_{\theta_i}$ *with* $\mathbf{x}_i \neq \mathbf{x}_j \quad \forall i \neq j$ *and* $\theta_1 < \ldots < \theta_m$. *Then*

$$\mu\left([\mathbf{x}_1, \ldots, \mathbf{x}_n]\right) \leq \max\left\{\max_{\ell} \sqrt{\frac{\theta_\ell}{\theta_{\ell+1}}}, 1 - \frac{1}{\theta_m}\right\} \tag{5.18}$$

with $\ell = 1, \ldots, m-1$. *The bound is satisfied with equality when*

$$n = \sum_{i=1}^{m} |\Delta_{\theta_i}| \tag{5.19}$$

Proof: See Appendix 5.9

Theorem (49) establishes an upper bound for the values of the coherence of a set of vectors given the knowledge of the sets Δ_θ from which these vectors are taken. It can be interpreted as a worst case value of the coherence for those vectors. Additionally, this upper bound is in the set of possible values of the coherence. From the computational point of view, it is important to note that this worst case value can be calculated basically considering the largest value of θ and the value of $\max_\ell \sqrt{\frac{\theta_\ell}{\theta_{\ell+1}}}$.

On the other hand it is possible to see that given any value of the term $\max\left\{\max_\ell \sqrt{\frac{\theta_\ell}{\theta_{\ell+1}}}, 1 - \frac{1}{\theta_m}\right\}$, it is always possible to choose a set of vectors from a collection of sets $\Delta_{\theta_i}, i = 1 \ldots m$ that have exactly this value of the coherence. This fact

can be used in order to determine the minimum coherence as the minimum between those worst case values. This idea can be described by the following theorem.

Theorem 50. *The minimum coherence that can be achieved for a matrix of u columns, whose column vectors are distinct elements of the set $Q_K \setminus \{\mathbf{0}\}$ with $2 \leq u \leq 2^K - 1$, is given by*

$$\mu_{min}(u) = \min_{\tilde{\mu}} \{f_{\tilde{\mu},s}(u)\} \tag{5.20}$$

where the function $f_{\tilde{\mu},s}(t)$ is defined as

$$f_{\tilde{\mu},s}(t) = \begin{cases} \tilde{\mu} & for \quad t = 2, \ldots, s \\ \infty & otherwise \end{cases} \tag{5.21}$$

where $\tilde{\mu}$ is in the set of values given by the terms $\max\left\{\max_\ell \sqrt{\frac{\theta_\ell}{\theta_{\ell+1}}}, 1 - \frac{1}{\theta_m}\right\}$, i.e. $\tilde{\mu}$ is the highest possible value of the coherence estimated by the upper bound in theorem 49. The value of s is given by

$$s = \sum_{i=1}^{m} |\Delta_{\theta_i}| \tag{5.22}$$

considering the collection of sets $\{\Delta_{\theta_i}\}_{i=1}^m$ with $\theta_1 < \ldots < \theta_m$.

Proof: In Appendix 5.10

The function $f_{\tilde{\mu},s}(t)$ establishes all the possible number of vectors that can have the value of the coherence $\tilde{\mu}$. Given a number of vectors it is possible to achieve different values of the coherence, and then between those possible values the minimum is chosen in order to determine the minimum coherence value.

These ideas are considered in Algorithm 2, used in order to calculate the minimum coherence given a number of vectors u and the dimension of the embedding space K. Additionally, it is also possible to obtain the classes Δ_θ from which the vectors should be selected in order to achieve the minimum coherence[1]. The set

[1] The number of possible combinations of vectors that achieve this minimum value is not of interest.

Algorithm 8 Minimum Coherence Calculation for Binary Matrices

Input: K
Output: $\mu_{min}(u)$ and $C_{min}(u)$
1: **for** $j = 2$ to K **do**
2: $V = comb\{K, j\}$.
3: **for** $i = 1$ to $|V|$ **do**
4: $\tilde{\mu} = \max\left\{\max_\ell \sqrt{\frac{\theta_\ell}{\theta_{\ell+1}}}, 1 - \frac{1}{\theta_j}\right\}$ with $\ell = 1, \ldots, j-1$.
5: $s = \sum_{r=1}^{|V^{(i)}|} |\Delta_{\theta_r}|$
6: $C^{(\tilde{\mu},s)} = \{\theta_1, \ldots, \theta_j\}$
7: **end for**
8: **end for**
9: **for** $u = 2$ to $2^K - 1$ **do**
10: $\mu_{min}(u) = \min_{\tilde{\mu}} \{f_{\tilde{\mu},s}(u)\}$
11: $C_{min}(u) = C^{(\mu_{min}(u), u \leq s)}$
12: **end for**
13: **return** $\mu_{min}(u)$ and $C_{min}(u)$

$V = comb\{K, j\}$ is the set of combinations of j numbers taken from the set $\{1, \ldots K\}$ and indicate the sets Δ_θ that are considered. The i^{th} combination in the set V is given by $V^{(i)} = \{\theta_1, \ldots, \theta_j\}$. Once the combination of sets Δ_θ is selected, they are ordered as $\theta_1 < \ldots < \theta_j$. The function $\mu_{min}(u)$ represents the minimum coherence that can be achieved with a matrix whose number of columns is u, and the term $C_{min}(u)$ indicates a collection of sets Δ_θ from which the set of vectors achieving $\mu_{min}(u)$ can be selected. Then any submatrix $\hat{\mathbf{H}}_r$ with K rows and u columns can be designed taking a set of u vectors from the classes Δ_θ indicated by $C_{min}(u)$. The construction of the submatrices $\hat{\mathbf{H}}_r$ based on the results of Algorithm 8 are presented in Algorithm 9.

Algorithm 9 Building the matrices $\hat{\mathbf{H}}_r$ considering the results of Algorithm 8

Input: K, L_r (the L_r can be obtained using Algorithm 12)
Output: All submatrices $\hat{\mathbf{H}}_r$
1: Considering K use Algorithm 8 to get $C_{min}(u)$;
2: **for** $r = 1$ to R **do**
3: Choose $u_r = |L_r|$ different columns from the classes indicated in $C_{min}(u_r)$;
4: Put those columns in a matrix $\hat{\mathbf{H}}_r$;
5: **end for**
6: **return** All submatrices $\hat{\mathbf{H}}_r$

In the minimum coherence designs of $\hat{\mathbf{H}}_r$, it is important to remark that the maximum number of columns that these submatrices can have is L (See Figure 5.3). Then, taking into account that in K shots the total number of vertices of the hypercube Q_K without the zero vector is $2^K - 1$, it is necessary to have $L \leq 2^K - 1$. If that is not the case, then it will be necessary to use again one of the vectors already used in the construction of the $\hat{\mathbf{H}}_r$, which implies $\mu(\hat{\mathbf{H}}_r) = 1$ for some of the $\hat{\mathbf{H}}_r$. Therefore, this condition could be represented as

$$K \geq \lfloor \log_2(L+1) \rfloor + 1. \tag{5.23}$$

5.4.3 Minimizing $\sum_{(i,j)\in I} \varphi_{i,j}$

Taking into account that

$$\sum_{(i,j)\in I} \varphi_{i,j} = \sum_{(i,j)\in I_1} \varphi_{i,j} + \ldots + \sum_{(i,j)\in I_R} \varphi_{i,j} \tag{5.24}$$

with $\sum_{(i,j)\in I_r} \varphi_{i,j} \geq 0 \quad \forall r$ and the sets I_r are disjoint. A direct minimization of the term $\sum_{(i,j)\in I} \varphi_{i,j}$ in (5.13) can be formulated as

$$\begin{aligned} &\underset{\hat{\mathbf{H}}_r}{\text{minimize}} \quad \sum_{(i,j)\in I_r} \varphi_{i,j} \quad \forall I_r \\ &\text{subject to} \quad \varphi_{i,j} = \left\langle \hat{\mathbf{h}}_i^{(r)}, \hat{\mathbf{h}}_j^{(r)} \right\rangle, \quad \left\| \hat{\mathbf{h}}_i^{(r)} \right\|_1 = 1. \end{aligned} \tag{5.25}$$

Then, the problem is again decomposed in a set of smaller problems in which the sum of all possible inner products are minimized but considering the submatrices $\hat{\mathbf{H}}_r$, which would lead to a direct minimization of $\sum_{(i,j)\in I} \varphi_{i,j}$. For (5.25) it is possible to establish closed form solutions as it is shown next.

Theorem 51. *The solution submatrices $\hat{\mathbf{H}}_r$ to the problem (5.25) are given by*

$$\hat{\mathbf{H}}_r = \boldsymbol{\pi} \left\{ [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n_r}] \right\}, \tag{5.26}$$

where n_r is the number of columns required in the submatrix $\hat{\mathbf{H}}_r$ and the operator $\boldsymbol{\pi}$ represents the random permutation operator on columns. The vectors $\boldsymbol{u}_i$ are obtained as the first n_r columns of $\mathbf{U}$, which is given by

$$\mathbf{U} = \mathbf{1}_{1\times L} \otimes \mathbf{I}_{K\times K}. \tag{5.27}$$

Proof: See Appendix 5.11

Algorithm 10 Building the matrices $\hat{\mathbf{H}}_r$ based on the minimization of $\sum_{(i,j)\in I}\varphi_{i,j}$

Input: K and L_r
Output: All submatrices $\hat{\mathbf{H}}_r$
1: **for** $r = 1$ to R **do**
2: $\quad \mathbf{U} = \mathbf{1}_{1\times L} \otimes \mathbf{I}_{K\times K}$.
3: $\quad \hat{\mathbf{H}}_r = \boldsymbol{\pi}\{\mathbf{U}(:, 1 : |L_r|)\}$
4: **end for**
5: **return** All submatrices $\hat{\mathbf{H}}_r$

The construction of $\hat{\mathbf{H}}_r$ based on the minimization of $\sum_{(i,j)\in I}\varphi_{i,j}$ and the closed form solution of (5.25) is presented in Algorithm 10.

5.4.4 About additional restrictions on H

Since $\mathbf{H}$ models a physical device, it is important to consider some physical constraints related with colored coded apertures. Coded apertures with large number of color filters lead to costly implementation since its cost increases directly with the number of colors [73]. Thus, a constrained optimization procedure restricting the number of colors is of interest.

In this work this restriction is added in a post-optimization stage considering the relation that different pixels exhibit in the structure of the $\mathbf{H}$ matrix. In Figure 5.4 it is possible to see, how the spectral responses of some pixels are related. In particular, it is possible to appreciate that following the lexicographic order on the pixels of the colored coded aperture, the pixels $i, i+N, i+2N, \ldots, i+(L-1)N$ are related, in the sense that their spectral responses are involved in at least one of the $\hat{\mathbf{H}}_r$ matrices. As it is indicated in Fig. 5.4, this related pixels are identified looking at the support in the rows of $\mathbf{H}$. The idea is to keep those spectral responses in the designed $\mathbf{H}$, such that the values of the inner products between columns of $\mathbf{H}$ are preserved as much as possible. For this reason the value of i should be chosen such that the spectral responses selected

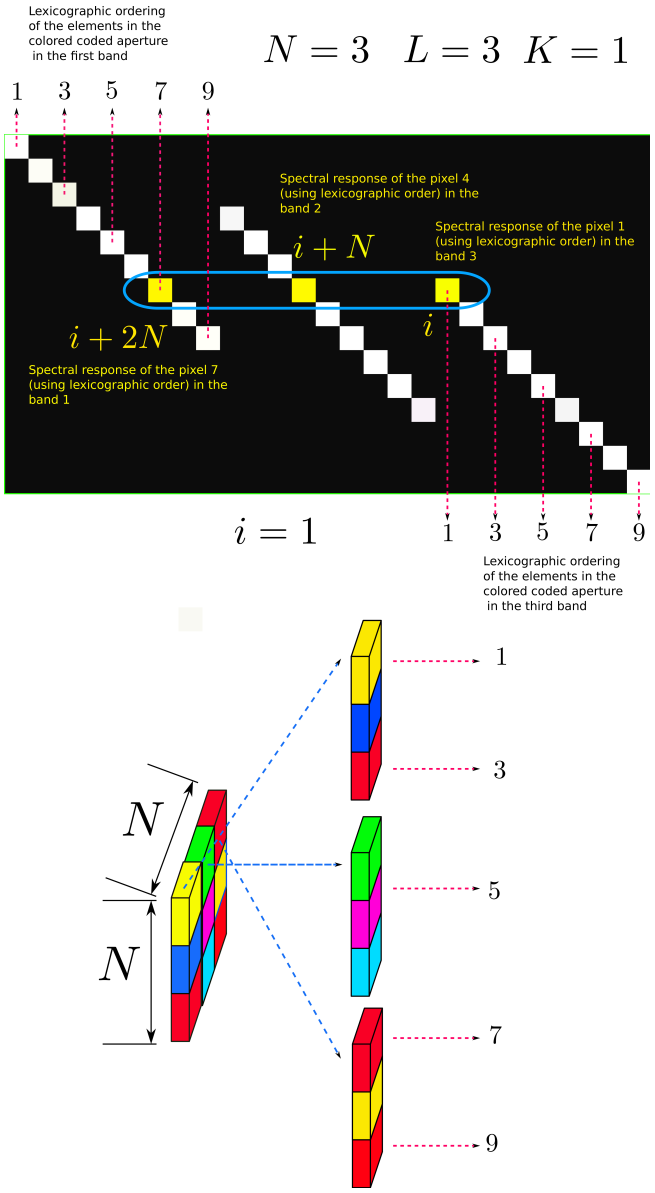


Figure 5.4: Illustration of $\mathbf{H}$ for 1 shot considering a multispectral scene of 3 bands and $N = 3$. As it is indicated in yellow color, three entries in the support of one of the rows of $\mathbf{H}$ are involved in one of the submatrices $\hat{\mathbf{H}}_r$. The lexicographic order in each diagonal indicates which pixels have their spectral responses related by $\hat{\mathbf{H}}_r$.

are related with the submatrices $\hat{\mathbf{H}}_r$ with the largest number of columns, in particular $i = 1$ satisfies this requirement.

The proposed strategy then consists on using the optimal designs of $\mathbf{H}$ obtained without restrictions on the number of spectral responses and after that, the maximum number of spectral responses per shot ζ is taken into account such that the new colored coded aperture uses some of the spectral responses of the pixels $i, i+N, i+2N, \ldots, i+(L-1)N$ (using lexicographic order). This procedure is summarized in Algorithm 11.

Algorithm 11 $\mathbf{H}$ design with restrictions on the number of colors

Input: ζ (maximum number of spectral responses per shot)
Output: $\mathbf{H}$

Design $\mathbf{H}$ using Algorithm 1, with criteria 1 (see Algorithm 9) or criteria 2 (see Algorithm 10).
for $k = 1$ to K **do**
  Get $\mathbf{T}^{(k)}$ from $\mathbf{H}$ (see Fig. (5.2) and Appendix 5.12)
  for $n = 1$ to N **do**
    $V = \mathbf{T}^{(k)}(1, 1 : \zeta, :)$
    $Q = []$;
    for $q = 1$ to $\lceil N/\zeta \rceil$ **do**
      $Q = [Q, \boldsymbol{\pi}_c(V, 2)]$;
    end for
    $Q = Q(1, 1 : N, :)$;
    $\mathbf{T}^{(k)}_{res}(n, :, :) = Q$;
  end for
end for
Put the values of $\mathbf{T}^{(k)}_{res}$ in $\mathbf{H}$ (see Fig. (5.2) and Appendix 5.12)
return $\mathbf{H}$

The variable $\mathbf{T}^{(k)}$ is the three-dimensional array representation of the colored coded aperture in the shot k, and $\mathbf{T}^{(k)}_{res}$ the resultant colored coded aperture after applying the restrictions on the number of colors. The values of $\mathbf{T}^{(k)}_{res}$ can be located in the final designed $\mathbf{H}$ matrix considering the structure of $\mathbf{H}$ showed in Fig. 5.2 and Fig. 5.3. In Appendix 5.12 an auxiliary algorithm designed to do this task is presented. The operator $\boldsymbol{\pi}_c(\cdot, 2)$ is the random circular shifting operator acting through the second dimension. In Fig. 5.5 it is shown how the restriction on the maximum number of spectral responses in the coded aperture is applied. The statement $V =$

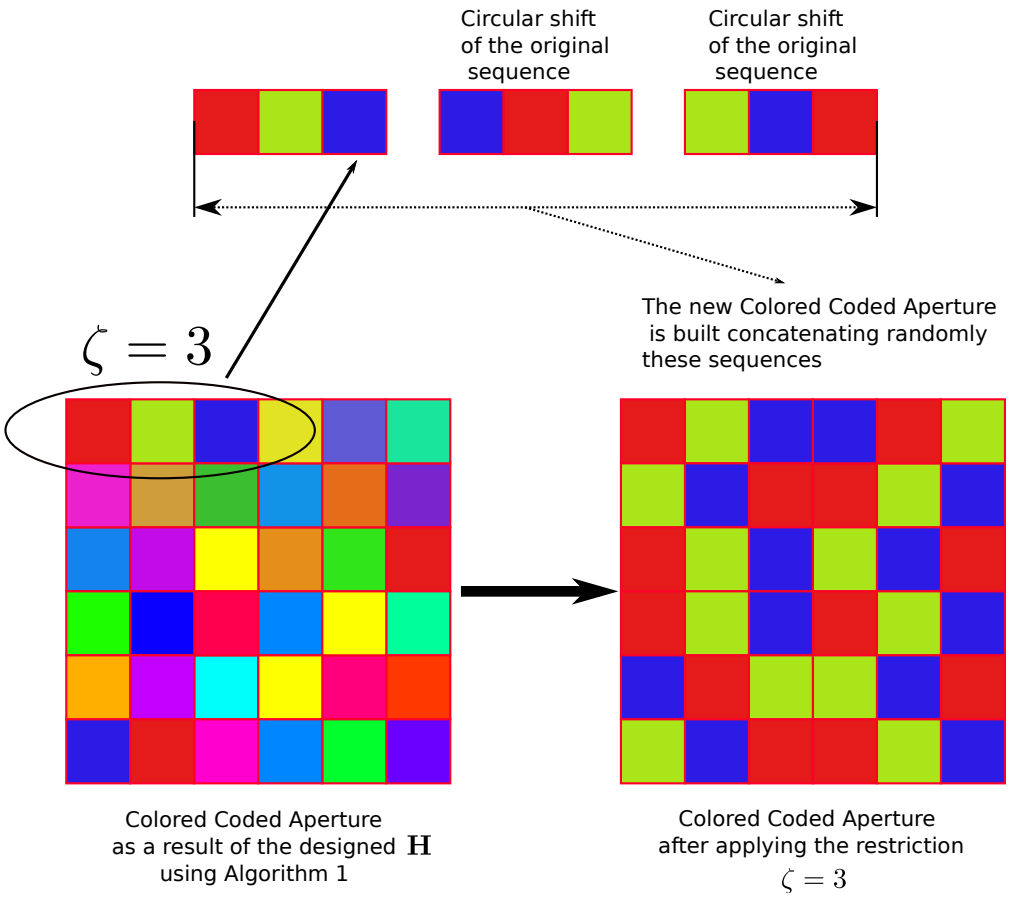


Figure 5.5: Illustration of how the designed coded aperture with restrictions on the number of colors is obtained. In the original coded aperture there is a large number of spectral responses whereas the coded aperture obtained with the restriction $\zeta = 3$ can have at most 3 different spectral responses.

$\mathbf{T}^{(k)}(1, 1 : \zeta, :)$ indicates that in the variable V are stored the spectral responses of the coded aperture in the row 1 from the column 1 up to the column ζ (see Fig. 5.5), which considering the lexicographic ordering of the pixels of $\mathbf{T}^{(k)}$ makes reference to the pixels $1 + N, 1 + 2N, \ldots, 1 + \zeta N$. The maximum possible number of these spectral responses is given by $K\zeta$. In line 8 of Algorithm 5, the arrays Q and $\boldsymbol{\pi}_c(V, 2)$ are concatenated along the second dimension, i.e. if $\dim(Q) = a \times b \times c$ and $\dim(\boldsymbol{\pi}_c(V, 2)) = a \times d \times c$, then $\dim([Q, \boldsymbol{\pi}_c(V, 2)]) = a \times (b + d) \times c$.

5.5 Simulations

In order to evaluate the performance in CASSI of the designed matrices $\mathbf{H}$ using the strategies and algorithms proposed, a set of simulations is realized considering also the effects of the noise in the measurements.

5.5.1 Parameters of the Simulations

5.5.1.1 Multispectral scene

A datacube of dimension $64 \times 64 \times 12$ is considered. The representation of this datacube decomposed by bands can be appreciated in Figure 5.7.

5.5.1.2 Reconstruction Algorithm and Basis Representation

The multispectral scene is represented in the basis $\mathbf{\Psi} = \mathbf{\Psi}_{DCT} \otimes \mathbf{\Psi}_{W}$ where $\mathbf{\Psi}_{DCT}$ is the Discrete Cosine Transform (DCT) basis for the spectral domain and $\mathbf{\Psi}_{W}$ is the wavelet basis for the spatial domain. The GPSR algorithm is used in order to perform the reconstructions, the regularization parameter is tunned empirically following the guidelines in [81].

5.5.1.3 Measure of the quality of the reconstructions

The Peak Signal to Noise Ratio (PSNR) is used as a measure of the quality of the reconstructions.

5.5.2 Optimal H designs

The results of the simulations are presented in Figures 5.6, 5.7 and 5.8 using the two different optimization criteria discussed before

- Criteria 1: Minimizing the Coherence of the submatrices $\hat{\mathbf{H}}_r$. Here the designed $\mathbf{H}$ is obtained as the result of the methods and algorithms presented in section IV(B).

- Criteria 2: Minimizing $\sum_{(i,j)} \varphi_{i,j}$ using the solution presented for (5.25). Here the designed $\mathbf{H}$ is obtained as the result of the methods presented in section IV(C).

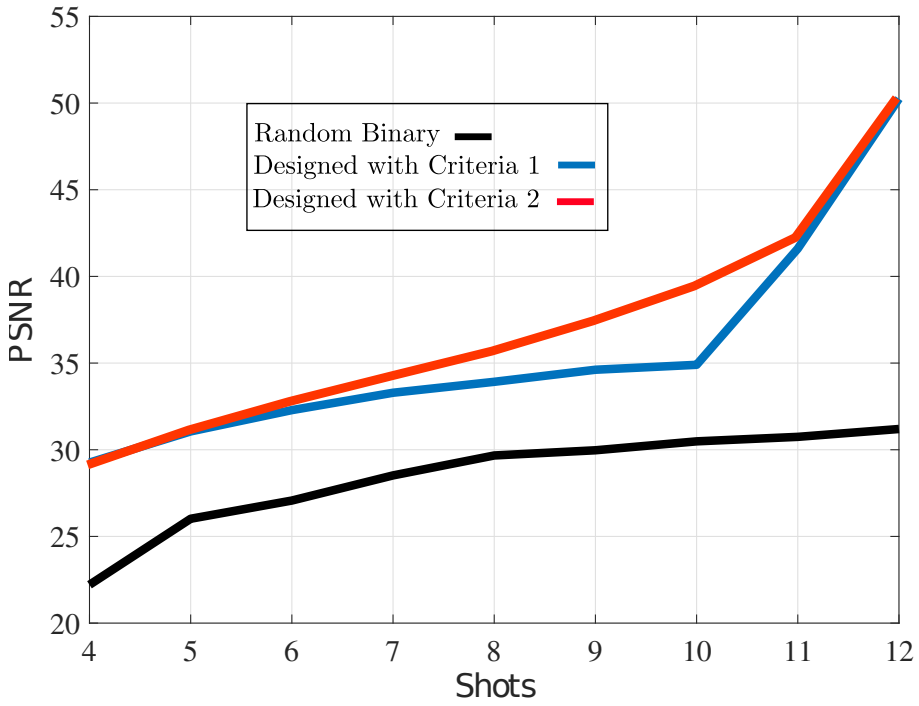


Figure 5.6: Simulations results comparing the performance of the different designs of $\mathbf{H}$ using two different optimization criteria. Several values for the number of shots K are considered. In blue color, the results using optimization criteria 1, based on the minimum coherence of $\mathbf{H}$, are presented. In red color, the results using optimization criteria 2, based on minimizing $\sum_{(i,j)\in I} \varphi_{i,j}$, are presented. It is possible to see that the results obtained with the designed $\mathbf{H}$ are consistently superior than the results obtained with random binary coded apertures.

As can be appreciated in Figure 5.6, the highest performance is obtained when criteria 2 is used, i.e. when $\sum_{(i,j)} \varphi_{i,j}$ is minimized. This superiority with respect the use of random black and white coded apertures is consistent through the different number of shots used. Additionally, despite the fact that the reconstructions using criteria 1 are not as good as the ones for criteria 2, the result is still superior to the results that can be obtained with the classical random black and white coded apertures.

In Figures 5.7 and 5.8 the reconstructed multispectral scene is shown using the designs of $\mathbf{H}$ based on the criteria 2 and with the configuration of $\mathbf{H}$ using a black and white coded aperture. The results show clearly how the bands reconstructed with the designed codes are superior. The differences are evident in the bands $519[nm], 539[nm], 559[nm], 619[nm]$ for instance. In Figure 5.8 a zoomed version of the band $559[nm]$ can be observed, in particular it is possible to appreciate the substantial improvement in the reconstructed bands when the designed codes are used.

Figures 5.9 and 5.10 show the details of the resultant coded aperture from the designed $\mathbf{H}$ using the criteria 2. In Figure 5.9 the two dimensional representation of the coded aperture is depicted, and the spectral responses of selected pixels are shown in Figure 5.10 as a function of the wavelength. In Figure 5.11 it is possible to see the results obtained using criteria 2 versus the results obtained with the codes designed in [2]. The results obtained with the approach present in this paper allow to obtain higher values of PSNR than the values obtained with the RIP based designs. The result that the RIP methods allow, do not have arbitrary spectral responses [2], whereas the proposed solution here allow the use of a more diverse variety of spectral responses, which gives a more flexible design.

5.5.3 Optimal H designs with restrictions

In Figure 5.12, the results of the reconstructions using optimal designs with restrictions on the number of spectral responses in the coded aperture are presented. Different cases are considered, going from a maximum number of spectral responses per shot $\zeta = 3$ up to $\zeta = 8$. As the value of ζ is increased, the PSNR of the reconstructions

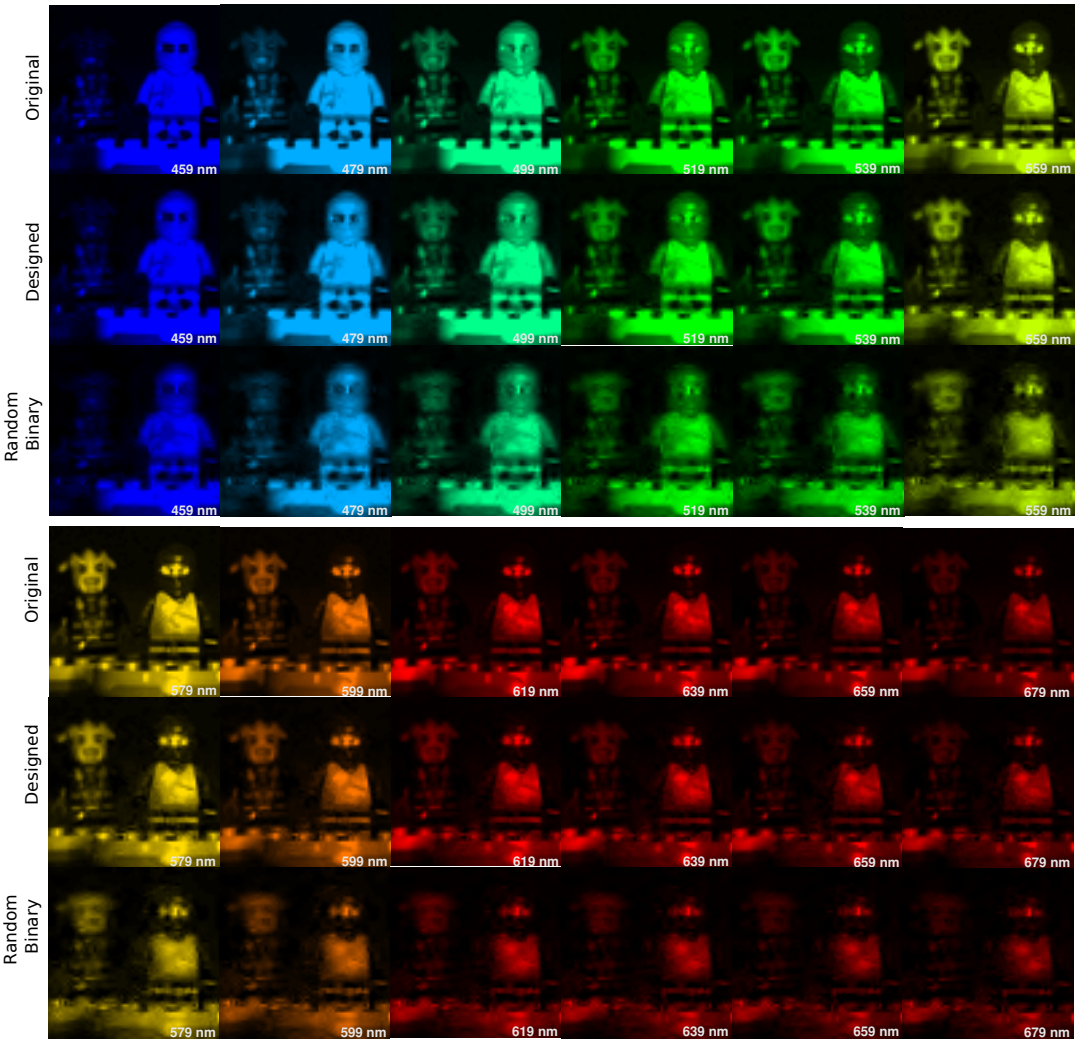


Figure 5.7: Simulation results showing the reconstructed bands of the multispectral scene are considered. In the first row of each group of images, the original bands of the multispectral scene are shown. In the second row the reconstructed bands obtained with the designed **H** are presented, and in the third row the reconstructed bands using the traditional black and white coded aperture are depicted.

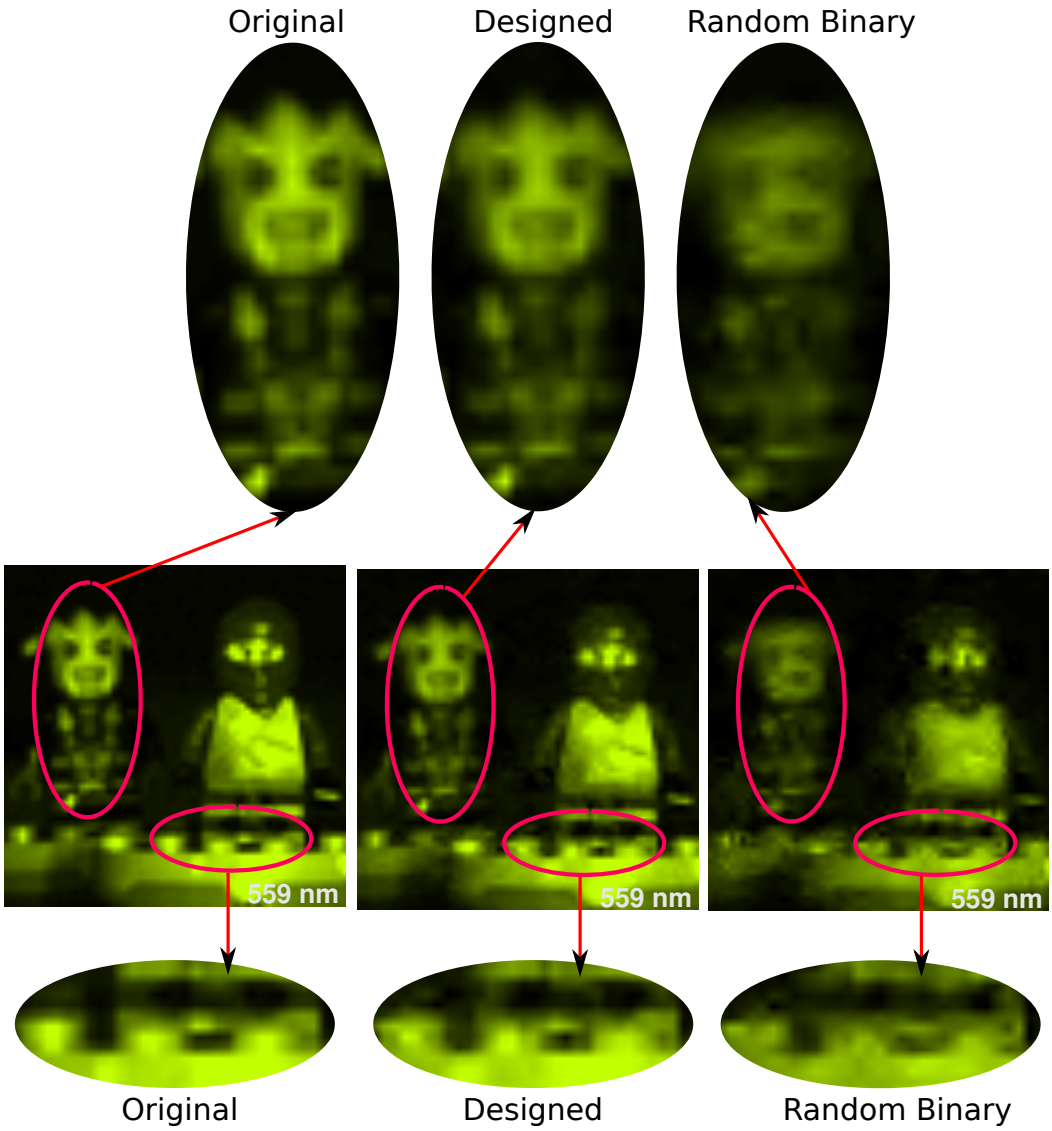


Figure 5.8: A zoomed version of the band at 559[nm]. On the left, the original band of the multispectral scene is shown. In the center, the reconstructed band using the designed **H** with the criteria 2 is presented. On the right, the reconstructed band using the random black and white coded aperture.

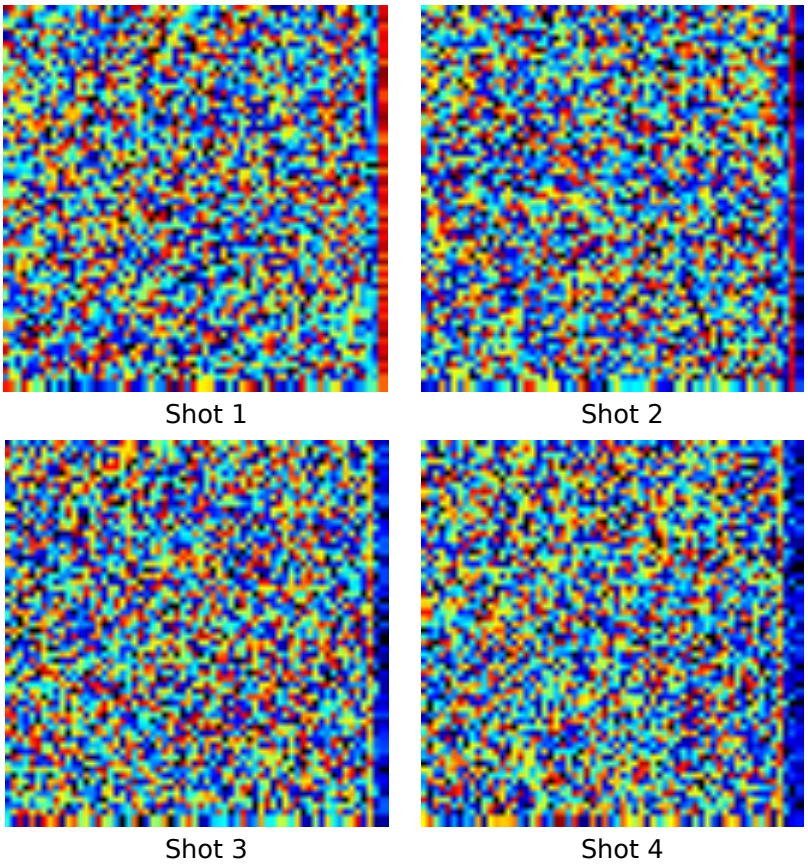


Figure 5.9: A representation of the spectral response of the resultant coded aperture from the designed **H** used in the reconstructions of Figures 5.6, 5.7, 5.8 is presented. For each shot, each spectral pattern in the coded aperture is shown. The representation of the spectral response of some pixels as functions of the wavelength is depicted in Figure 5.10.

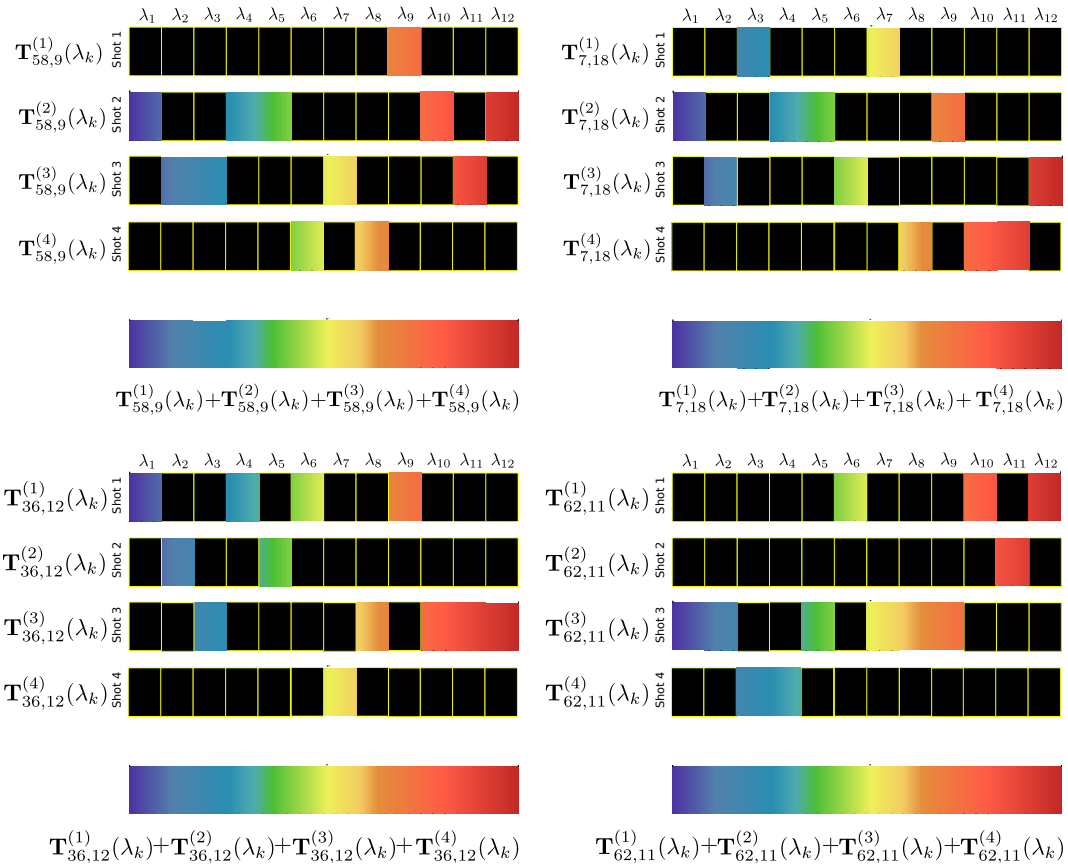


Figure 5.10: Spectral responses of selected pixels in the resultant coded aperture used in the reconstructions of Figures 5.6, 5.7, 5.8 (See also Figure 5.9) as functions of the wavelength. The units of horizontal axes are in nanometers $[nm]$. The term $\mathbf{T}^{(\ell)}_{m,n}(\lambda_k)$ indicates the spectral response of the coded aperture at position (m, n) in the shot ℓ.

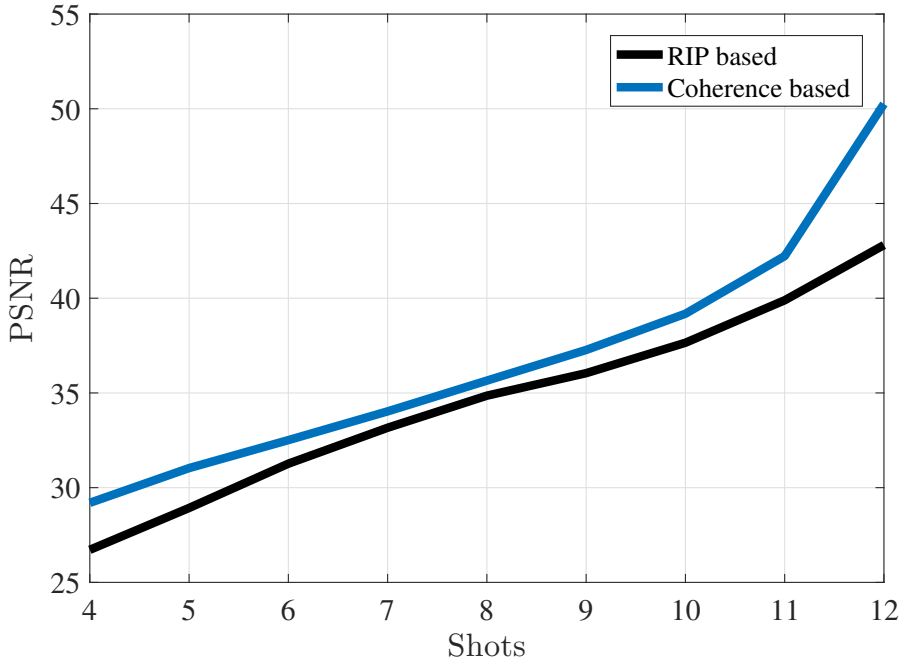


Figure 5.11: Simulation results comparing the designed codes proposed in this work using criteria 2 versus the results obtained in [2].

with restrictions gets closer to the curve that is obtained with the designed $\mathbf{H}$ with criteria 2. It is also possible to see, how the final number of spectral responses in the resultant coded aperture is smaller than its maximum limit which is $K\zeta$.

5.5.4 Considering the effects of the noise

The designed codes are tested against the classical random binary codes when the measurements $\mathbf{g}_{m,n}$ are contaminated with white Gaussian noise. Different values of SNR on $\mathbf{g}_{m,n}$ are considered.

In Figure 5.13 it is possible to appreciate the performance of the designed codes with criteria 2 versus the random binary codes for several levels of noise. Clearly the designed codes allow to get higher PSNR than the obtained with random black and white codes, and these results are consistent for different values of SNR in the measurements.

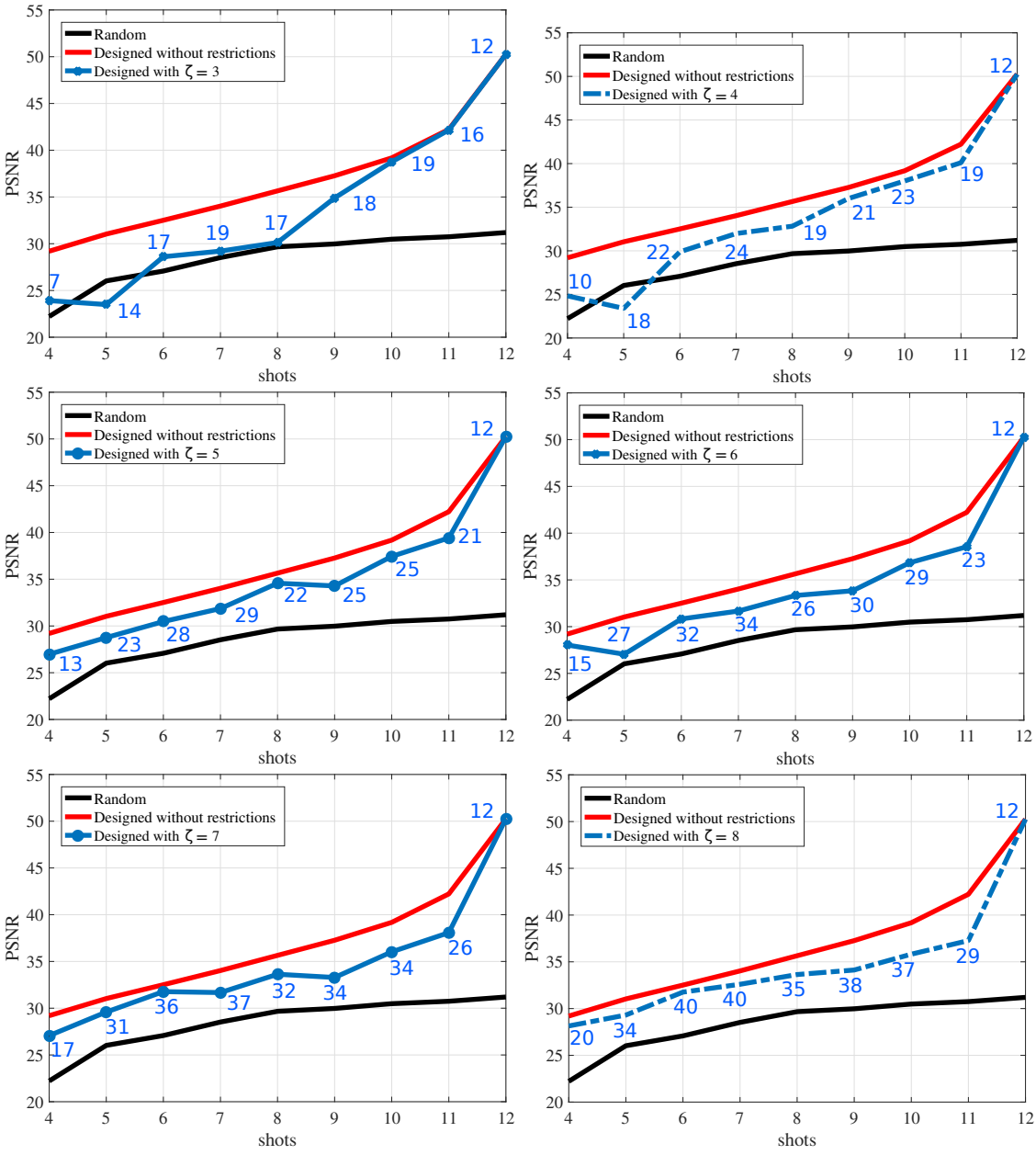


Figure 5.12: Results of the reconstructions using the designed **H** with restrictions on the number of spectral responses in the coded aperture. The resultant PSNR is shown for different values of ζ ranging from $\zeta = 3$ up to $\zeta = 8$. The number of spectral responses in the resultant coded aperture is indicated on each marker of the curve.

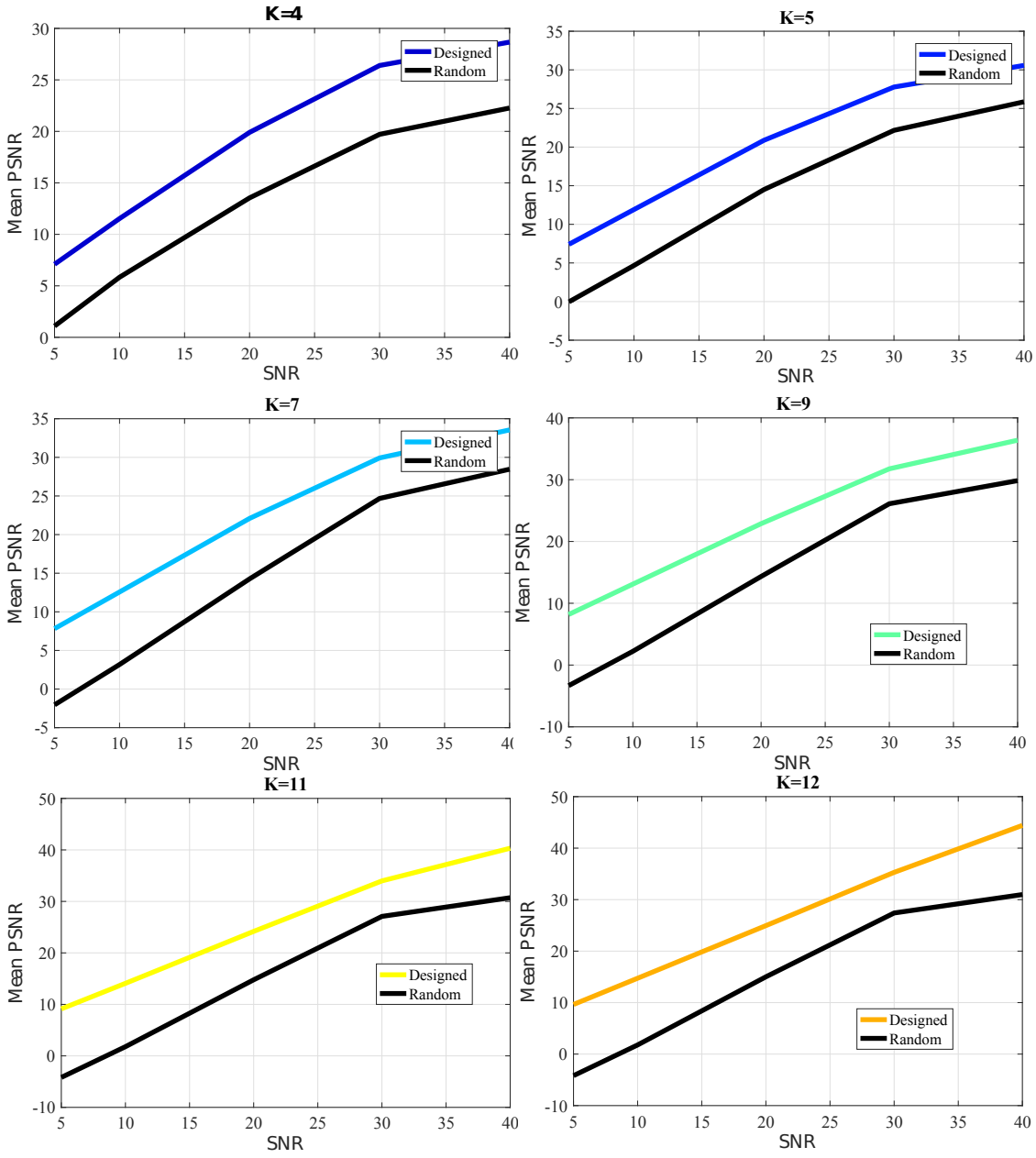


Figure 5.13: Simulation results showing the quality of the reconstructions in the presence of noise. The designed codes allow to obtain higher quality reconstructions than the random binary codes in the presence of noise, considering different values of the SNR in the measurements.

5.6 Conclusions and Future work

The coherence analysis of the sensing matrix of the CASSI system proposed in this work, allows the design of families of matrices $\mathbf{H}$ which lead to reconstructions with higher PSNR than the obtained using random black and white coded apertures and coded apertures designs obtained with the RIP approaches. It is also shown that the proposed method allows to obtain the designs of $\mathbf{H}$ in a fast and simple way.

The approach used in this work, showed that there are numerous designs of $\mathbf{H}$ that are obtained as a result of the minimization of an upper bound of the coherence. It was also shown that given those designs, it is possible to restrict the number of spectral responses exploiting again the structure of $\mathbf{H}$. This allows more flexible designs with higher diversity of spectral responses than the proposed in [2].

The results obtained with the designed $\mathbf{H}$ are more robust than the configurations obtained when the random black and whited coded apertures are used, if noise in the measurements is considered. This behavior is consistent for different values of SNR.

In future works the approach presented in this paper can be improved considering more specific knowledge on the basis $\mathbf{\Psi}$.

5.7 Discretization of CASSI Model

The *rect* function used in equation (5.1), can be represented as

$$\text{rect}(x,y) = \begin{cases} 1 & \text{if} \quad -1/2 \leq x \leq 1/2, \quad -1/2 \leq y \leq 1/2 \\ 0 & \text{otherwise} \end{cases}$$

This function is separable and can be equivalently written as

$$\text{rect}(x,y) = \text{rect}(x)\text{rect}(y)$$

with

$$\text{rect}(x) = \begin{cases} 1 & \text{if} \quad -1/2 \leq x \leq 1/2 \\ 0 & \text{otherwise} \end{cases}$$

Considering this facts, it is possible to see that there are concrete cases in which the product between *rect* functions is different from zero. With this, it is possible to define the limits of the integral operators in (5.1); in particular it is considered that[2]

$$\text{rect}\left(\frac{y}{\Delta}-m^{'}\right)\text{rect}\left(\frac{y}{\Delta}-m\right)\neq 0 \Longleftrightarrow m^{'}=m,$$

whereas $\text{rect}\left(\frac{x}{\Delta}-n\right)\text{rect}\left(\frac{x-S(\lambda)}{\Delta}-n^{'}\right)$ is different from zero if $\Delta\left(n^{'}+\frac{1}{2}\right)+S(\lambda)>\Delta\left(n-\frac{1}{2}\right)$ and $\Delta\left(n^{'}+\frac{1}{2}\right)+S(\lambda)<\Delta\left(n+\frac{1}{2}\right)$. Therefore, it follows that $n^{'}=-\left\lceil\frac{S(\lambda)}{\Delta}\right\rceil+n$.

Then, in the x variable the integration limits are

$$\left\{\Delta\left(n-\frac{1}{2}\right),\quad \Delta\left(n+\frac{1}{2}\right)-\Delta\left\lceil\frac{S(\lambda)}{\Delta}\right\rceil+S(\lambda)\right\}.$$

Additionally, if $\Delta\left(n^{'}-\frac{1}{2}\right)+S(\lambda)>\Delta\left(n-\frac{1}{2}\right)$ and $\Delta\left(n^{'}-\frac{1}{2}\right)+S(\lambda)<\Delta\left(n+\frac{1}{2}\right)$, it follows that $n^{'}=-\left\lceil\frac{S(\lambda)}{\Delta}\right\rceil+n+1$. In this case for the x axes the integration limits are

$$\left\{\Delta\left(n+\frac{1}{2}\right)-\Delta\left\lceil\frac{S(\lambda)}{\Delta}\right\rceil+S(\lambda),\quad \Delta\left(n+\frac{1}{2}\right)\right\}.$$

Using these facts, the measurements at pixel (m,n) can be written as

$$\begin{aligned}\mathbf{g}_{m,n}=&\int_{\Lambda}\int_{\Delta\left(n-\frac{1}{2}\right)}^{\Delta\left(n+\frac{1}{2}\right)-\Delta\left\lceil\frac{S(\lambda)}{\Delta}\right\rceil+S(\lambda)}\int_{\Delta\left(m-\frac{1}{2}\right)}^{\Delta\left(m+\frac{1}{2}\right)}\\&(TF)_{m,n-\left\lceil\frac{S(\lambda)}{\Delta}\right\rceil}(\lambda)dydxd\lambda\\&+\int_{\Lambda}\int_{-\Delta\left\lceil\frac{S(\lambda)}{\Delta}\right\rceil+\Delta\left(n+\frac{1}{2}\right)+S(\lambda)}^{\Delta\left(n+\frac{1}{2}\right)}\int_{\Delta\left(m-\frac{1}{2}\right)}^{\Delta\left(m+\frac{1}{2}\right)}\\&(TF)_{m,n-\left\lceil\frac{S(\lambda)}{\Delta}\right\rceil+1}(\lambda)dydxd\lambda,\end{aligned}$$

[2] In the analysis considered in this section the variables $m, m^{'}, n, n^{'}$ represent integer numbers.

by the mid point rule approximation[3] for the spatial operators, the measurements are given by

$$\mathbf{g}_{m,n} \approx \int_{\Lambda} \Delta^2 \left(- \left\lceil \frac{S(\lambda)}{\Delta} \right\rceil + \frac{S(\lambda)}{\Delta} + 1 \right) (TF)_{m,n-\left\lceil \frac{S(\lambda)}{\Delta} \right\rceil}(\lambda) d\lambda \\ - \int_{\Lambda} \Delta^2 \left(- \left\lceil \frac{S(\lambda)}{\Delta} \right\rceil + \frac{S(\lambda)}{\Delta} \right) (TF)_{m,n-\left\lceil \frac{S(\lambda)}{\Delta} \right\rceil+1}(\lambda) d\lambda. \quad (5.28)$$

When discretization in λ axes is considered, the bands in the super resolution model define the limit points of the intervals of integration. Using again the mid-point rule for the approximation of the operator in λ axes, it follows that

$$\mathbf{g}_{m,n} = \sum_{k=0}^{L-1} \Delta_{\lambda(k)} \Delta^2 \left(- \left\lceil \frac{S(\hat{\lambda}_k)}{\Delta} \right\rceil + \frac{S(\hat{\lambda}_k)}{\Delta} + 1 \right) \\ \times (TF)_{m,n-\left\lceil \frac{S(\hat{\lambda}_k)}{\Delta} \right\rceil}(\hat{\lambda}_k) \\ + \sum_{k=0}^{L-1} \Delta_{\lambda(k)} \Delta^2 \left(- \left\lceil \frac{S(\hat{\lambda}_k)}{\Delta} \right\rceil + \frac{S(\hat{\lambda}_k)}{\Delta} \right) \times (TF)_{m,n-\left\lceil \frac{S(\hat{\lambda}_k)}{\Delta} \right\rceil+1}(\hat{\lambda}_k),$$

where $\hat{\lambda}_k = (\lambda_{k+1} + \lambda_k)/2$, $\Delta_{\lambda(k)} = \lambda_{k+1} - \lambda_k$. Considering the properties of $\left\lfloor \frac{-S(\lambda)}{\Delta} \right\rfloor$ presented with all detail in [74] (also discussed at the end of this appendix) and the fact that $\lfloor -u \rfloor = -\lceil u \rceil$, the representation of $\mathbf{g}_{m,n}$ takes the form

$$\mathbf{g}_{m,n} = \sum_{k=0}^{L-1} \sum_{u'=0}^{c-1} \\ \left[\Delta_{\lambda(k)} \Delta^2 \left(- \left\lceil \frac{S(\hat{\lambda}_k)}{\Delta} \right\rceil + \frac{S(\hat{\lambda}_k)}{\Delta} + 1 \right) (TF)_{m,n-k+u'}(\hat{\lambda}_k) \right. \\ \left. + \Delta_{\lambda(k)} \Delta^2 \left(- \left\lceil \frac{S(\hat{\lambda}_k)}{\Delta} \right\rceil + \frac{S(\hat{\lambda}_k)}{\Delta} \right) (TF)_{m,n-k+u'+1}(\hat{\lambda}_k) \right],$$

[3] For any integrable function $f(x)$ in the interval $[a, b]$, the midpoint rule approximation can be represented as $\int_a^b f(x)dx \approx (b-a) f((a+b)/2)$.

with $c \in \mathbb{N}$. The value of c represents the number of pixels of the detector affected by one voxel of the datacube model [82] [74]. If

$$\boldsymbol{\alpha}_{m,n,k} = \Delta_{\lambda(k)}\Delta^2 \left(- \left\lceil \frac{S(\hat{\lambda}_k)}{\Delta} \right\rceil + \frac{S(\hat{\lambda}_k)}{\Delta} + 1 \right),$$

and

$$\boldsymbol{\beta}_{m,n,k} = -\Delta_{\lambda(k)}\Delta^2 \left(- \left\lceil \frac{S(\hat{\lambda}_k)}{\Delta} \right\rceil + \frac{S(\hat{\lambda}_k)}{\Delta} \right),$$

it follows that

$$\mathbf{g}_{m,n} = \sum_{k=0}^{L-1} \sum_{u=0}^{c} \left(\mathbf{w}_{m,n,k,u} (\mathbf{T} \circ \mathbf{F})_{m,n-k+u,k} \right),$$

where

$$\mathbf{w}_{m,n,k,u} = \begin{cases} \boldsymbol{\alpha}_{m,n,k} & \text{If} \quad u = 0 \\ \boldsymbol{\alpha}_{m,n,k} + \boldsymbol{\beta}_{m,n,k} & \text{If} \quad 0 < u < c \\ \boldsymbol{\beta}_{m,n,k} & \text{If} \quad u = c \end{cases}$$

for $c \geq 2$. When $c = 1$ the traditional model of the CASSI[4] can be obtained, with an abuse of notation, redefining $\mathbf{F}_{m,n-k,k} = \boldsymbol{\alpha}_{m,n,k}\mathbf{F}_{m,n-k,k} + \boldsymbol{\beta}_{m,n,k}\mathbf{F}_{m,n-k+1,k}$ such that

$$\mathbf{g}_{m,n} = \sum_{k=0}^{\min\{n,L-1\}} (\mathbf{T} \circ \mathbf{F})_{m,n-k,k},$$

For simplicity $c = 1$ is used for the presented analysis.

5.7.1 About the meaning of $\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor$

Taking into account the Weierstrass approximation theorem it is possible to obtain a good representation of $S(\lambda)$ as a polynomial [74]. In particular this representation can be stated as

$$S(\lambda) \approx \sum_{r=0}^{Q} \alpha_r \lambda^r \quad Q \in \mathbb{N}, \alpha_r \in \mathbb{R}.$$

[4] This model was used and proposed in [70] [71].

With this representation it is possible to separate the linear and nonlinear components of $S(\lambda)$. Renaming $q(\lambda) = \sum_{r=2}^{Q} \alpha_r \lambda^r$ and using the properties of the floor function [74], it follows that

$$\left\lfloor \frac{\alpha_0}{\Delta} + \frac{q(\lambda)}{\Delta} \right\rfloor + \left\lfloor \alpha_1 \frac{\lambda}{\Delta} \right\rfloor \leq \left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor \leq \left\lfloor \frac{\alpha_0}{\Delta} + \frac{q(\lambda)}{\Delta} \right\rfloor + \left\lfloor \alpha_1 \frac{\lambda}{\Delta} \right\rfloor + 1.$$

The term α_0 is associated with the mismatching in the dispersion axes of the multispectral image on the focal plane array [74], and it is considered that the nonlinearities in $S(\lambda)$ are contained in $q(\lambda)$, whereas the term $\left\lfloor \alpha_1 \frac{\lambda}{\Delta} \right\rfloor$ is directly related with the changes in the dispersion axes with respect to λ. It is possible then to associate a variable index to the term $\left\lfloor \alpha_1 \frac{\lambda}{\Delta} \right\rfloor$ indicating the changes through the spectral axes and another variable index can be associated to the term $\left\lfloor \frac{\alpha_0}{\Delta} + \frac{q(\lambda)}{\Delta} \right\rfloor$. This leads to the indexes involved in the representation of $\mathbf{g}_{m,n}$.

5.8 Proof of the Theorem 48

Taking into account equations (7.11) and (7.12) it follows that

$$\mu_{m,n}(\mathbf{A}) \leq |\varrho(m,n)|\, \xi$$

with

$$\xi = \frac{1}{\min\limits_{\mathbf{H}, \|\mathbf{h}_i\|>0 \forall i} \varrho(m,m)^{\frac{1}{2}} \varrho(n,n)^{\frac{1}{2}}}.$$

Using the triangle inequality on $|\varrho(m,n)|$ it is possible to get

$$|\varrho(m,n)| \leq \sum_{i=1}^{N^2 L} \varphi_{i,i} \left| R_{i,i}^{(m,n)} \right| + \sum_{(i,j)\in I} \varphi_{i,j} \left| R_{i,j}^{(m,n)} \right|.$$

Now, applying Holder's inequality to the right-hand side of the previous equation, it follows that

$$|\varrho(m,n)| \leq \left[\left(\sum_{i=1}^{N^2 L} \varphi_{i,i} \right) \max_i \left| R_{i,i}^{(m,n)} \right| + \left(\sum_{(i,j)\in I} \varphi_{i,j} \right) \max_{(i,j)\in I} \left| R_{i,j}^{(m,n)} \right| \right].$$

Therefore it finally follows that

$$\mu_{m,n}(\mathbf{A}) \leq \xi \left[\left(\sum_{i=1}^{N^2 L} \varphi_{i,i} \right) \max_i \left| R_{i,i}^{(m,n)} \right| + \left(\sum_{(i,j)\in I} \varphi_{i,j} \right) \max_{(i,j)\in I} \left| R_{i,j}^{(m,n)} \right| \right].$$

5.9 Proof of the Theorem 49

The inner product between any two different vectors in the set $Q_K \setminus \{\mathbf{0}\}$ can be written as

$$\frac{\langle \mathbf{x}_i, \mathbf{x}_j \rangle}{\|\mathbf{x}_i\|_2 \|\mathbf{x}_j\|_2} \leq \begin{cases} 1 - \frac{1}{\theta} & \text{If} \quad \mathbf{x}_i, \mathbf{x}_j \in \Delta_\theta \quad \forall \theta \\ \sqrt{\frac{\theta'}{\theta''}} & \text{If} \quad \mathbf{x}_i \in \Delta_{\theta'}, \mathbf{x}_j \in \Delta_{\theta''}, \theta' < \theta'' \end{cases} \tag{5.29}$$

Because the function $1 - 1/\theta$ is monotonic increasing in θ, it follows that

$$\frac{\langle \mathbf{x}_i, \mathbf{x}_j \rangle}{\|\mathbf{x}_i\|_2 \|\mathbf{x}_j\|_2} \leq \begin{cases} 1 - \frac{1}{\theta_m} & \text{If} \quad \mathbf{x}_i, \mathbf{x}_j \in \Delta_\theta, \quad \theta \in \{\theta_1, \ldots, \theta_m\} \\ \sqrt{\frac{\theta'}{\theta''}} & \text{If} \quad \mathbf{x}_i \in \Delta_{\theta'}, \mathbf{x}_j \in \Delta_{\theta''}, \theta' < \theta'' \end{cases} \tag{5.30}$$

where the sequence of values $\{\theta_1, \ldots, \theta_m\}$ satisfies $\theta_1 < \ldots < \theta_m$. Consider the term $\sqrt{\theta_i/\theta_j}$ with $i < j$. It is possible to see that

$$\sqrt{1 - \frac{1}{\theta_j}} > \sqrt{1 - \frac{r}{\theta_j}} \quad \forall r = 2, \ldots, \theta_j \tag{5.31}$$

then the maximum value of the term $\sqrt{\theta_i/\theta_j}$ is given when θ_i and θ_j are successive elements in the sequence $\{\theta_1, \ldots, \theta_m\}$. Therefore, it follows that

$$\frac{\langle \mathbf{x}_i, \mathbf{x}_j \rangle}{\|\mathbf{x}_i\|_2 \|\mathbf{x}_j\|_2} \leq \begin{cases} 1 - \frac{1}{\theta_m} & \text{If} \quad \mathbf{x}_i, \mathbf{x}_j \in \Delta_\theta, \quad \theta \in \{\theta_1, \ldots, \theta_m\} \\ \max\limits_{r<\ell} \sqrt{\frac{\theta_r}{\theta_\ell}} \end{cases} \tag{5.32}$$

with θ_r, θ_ℓ successive elements in the sequence $\theta_1, \ldots, \theta_m$, and therefore from this it is obtained that

$$\mu\left([\mathbf{x}_1, \ldots, \mathbf{x}_n]\right) \leq \max\left\{ \max_{r<\ell} \sqrt{\frac{\theta_r}{\theta_\ell}}, 1 - \frac{1}{\theta_m} \right\}$$

5.10 Proof of theorem 50

Let us consider $u \in \mathbb{N}$ such that $2 \leq u \leq 2^K - 1$, then the minimum coherence for this number of vectors in $Q_K \setminus \{\mathbf{0}\}$ is calculated as

$$\mu_{min}(u) = \min_{\tilde{\mu}} \{f_{\tilde{\mu},s}(u)\}$$

as indicated in theorem (50). Let us assume by contradiction that there exists a set of u vectors in $Q_K \setminus \{\mathbf{0}\}$, with coherence equal to $\gamma < \mu_{min}(u)$, i.e. there exists $u \in \mathbb{N}$ vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_u$ in $Q_K \setminus \{\mathbf{0}\}$ such that

$$\mu([\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_u]) = \gamma$$

The value of γ is in the finite set of the possible values of the coherence for any set of vectors taken from $Q_K \setminus \{\mathbf{0}\}$. Therefore, γ can be written as

$$\gamma = \max\left\{\max_{\ell} \sqrt{\frac{\theta_\ell}{\theta_{\ell+1}}}, 1 - \frac{1}{\theta_m}\right\}$$

for a collection of classes $\{\Delta_{\theta_i}\}_{i=1,\ldots,m}$. In particular, taking into account Theorem 49, we have that for any set of different vectors $\mathbf{y}_1, \ldots, \mathbf{y}_v \in \{\Delta_{\theta_i}\}_{i=1,\ldots,m}$

$$\mu\left([\mathbf{y}_1, \ldots, \mathbf{y}_v]\right) \leq \max\left\{\max_{\ell} \sqrt{\frac{\theta_\ell}{\theta_{\ell+1}}}, 1 - \frac{1}{\theta_m}\right\}.$$

When $v = \sum_{i=1}^{m} |\Delta_{\theta_i}|$, equality is achieved and the maximum number of vectors that can have this coherence value of γ is v. Therefore $u \leq v$, which implies that $\exists f_{\gamma,v}(u) < \infty$. This is a contradiction because $\mu_{min}(u) = \min_{\tilde{\mu}}\{f_{\tilde{\mu},s}(u)\}$.

5.11 Proof of theorem 51

In this proof, two cases are considered separately. In the first case, the number of columns n_r of the submatrices $\hat{\mathbf{H}}_r$ is such that $n_r \leq K$. In the second case, $n_r > K$.

It is important to take into account that the constraint $\|\hat{\mathbf{h}}_i\|_1 = 1$ in (5.25) and the fact that the entries of $\mathbf{H}$ belong to the set $\{0, 1\}$, indicate that the vectors $\hat{\mathbf{h}}_i$ must be taken from the set $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_K\}$, where $\mathbf{e}_i$ is the vector with 1 in the i^{th} component and 0 in the remaining components.

1. When $n_r \leq K$: In this case the columns of $\hat{\mathbf{H}}_r$ can be selected as any subset of the set E, and as a consequence of this $\varphi_{i,j} = 0$ for any choice of these subsets. Therefore

$$\hat{\mathbf{H}}_r = \boldsymbol{\pi}\{\mathbf{U}\}$$

where $\mathbf{U}$ is the matrix whose columns are selected as any subset of the set E. This is a particular case of (5.26).

2. When $n_r > K$: Let u be the number of columns of the matrix $\hat{\mathbf{H}}_r$. This number u is going to be represented as $u = j + mK$ where $m \in \mathbb{N}$ and $1 \leq j \leq K$. If $\ell_i > 1$ is the number of times that the element i of the set E appears in the choosing of the columns of the matrix $\hat{\mathbf{H}}_r$, then the number of ones in the term $\varphi_{i,j}$ (with $j > i$) in (5.25) as a consequence of this choice, is given by $\binom{\ell_i}{2} = \frac{\ell_i(\ell_i-1)}{2}$. This expression can be also used to describe the number of ones in $\varphi_{i,j}$ for the case $\ell_i = 1$.

   Taking into account this, the objective function in (5.25) can be rewritten as $G(\ell_1, \ldots, \ell_K) = \sum_{i=1}^{K} \phi(\ell_i)$ where $\phi(x) = x(x-1)/2$ and then, the original problem can be rewritten as

$$\begin{aligned} &\underset{(\ell_1,\ldots,\ell_K)}{\text{minimize}} \quad \sum_{i=1}^{K} \phi(\ell_i) \\ &\text{subject to} \quad \sum_{i=1}^{K} \ell_i = u, \quad \ell_i \in \mathbb{N} \end{aligned} \tag{5.33}$$

   The solution of the problem (5.25) presented in the equation (5.26) corresponds in the rewritten version of the problem (5.33) to the point

$$\boldsymbol{\ell}_O = (\ell_1, \ldots, \ell_K) = \boldsymbol{\pi}\left(m, m, \ldots, m, \ , m+1, m+1 \ldots, m+1\right)$$

   where there are j components with value m and $K - j$ components with value $m + 1$.

   It is going to be shown that the solution to (5.25), which is the same solution of (5.33) is given by $\boldsymbol{\ell}_O$.

   In order to get a better knowledge of the objective function we are dealing with in the specific domain (constraint), it is possible to use Lagrange multipliers

optimization, relaxing the domain of the ℓ_i to be $\ell_i \in \mathbf{R}_+$. Then, the Lagrangian can be written as

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{K} \ell_i(\ell_i - 1) + \lambda \left[\sum_{i=1}^{K} \ell_i - u\right].$$

Calculating the gradient and equating to zero all the components, it is obtained that

$$\ell_i - \frac{1}{2} + \lambda = 0 \quad \forall i = 1, \ldots, K; \quad \sum_{i=1}^{K} \ell_i = u,$$

which gives the optimal point $\tilde{\boldsymbol{\ell}} = (u/K, u/K, \ldots, u/K)$. This solution would be acceptable for (5.33) always that $u = rK$, where $r \in \mathbb{N}$, however that is not always the case.

Now, the function G restricted on the set of points $S = \{(\ell_1, \ldots, \ell_K) | \sum_{i=1}^{K} \ell_i = u\}$ is given by

$$G|_S = \frac{1}{2}\left[\sum_{i=1}^{K} \ell_i^2 - u\right] \tag{5.34}$$

which is a convex function on S. Then, the solution to (5.33) can be found solving the problem

$$\begin{aligned} \underset{\mathbf{v}}{\text{minimize}} \quad & \left\| \left(\frac{u}{K}, \frac{u}{K}, \ldots, \frac{u}{K}\right) - \mathbf{v} \right\|_2^2 \\ \text{subject to} \quad & \sum_{i=1}^{K} v_i = u, \quad v_i \in \mathbb{N} \end{aligned} \tag{5.35}$$

Because $v_i \in \mathbb{N}$, the solution vector of (5.35) must be obtained such that its components are obtained either as $v_i = \lfloor u/K \rfloor$ or $v_i = \lceil u/K \rceil$. It is clear that in general[5] not all the components can be obtained using just one type of operator because the condition $\sum_{i=1}^{K} v_i = u$ could be violated. Therefore, some of this terms are approximated using the floor function and the others the ceiling function, case in which some components of $\mathbf{v}$ are given by $\left\lfloor \frac{j+mK}{K} \right\rfloor = m$ while the others by $\left\lceil \frac{j+mK}{K} \right\rceil = m + 1$. Now, let q the number of components obtained

[5] The only situation in which both operators can be used for all the components is when u/K is an integer.

using the floor function operator for the rounding process and p the number of components obtained using the ceiling function operator, then it must happen that

$$qm + p(m+1) = u = j + mK, \quad p, q \in \mathbb{N}.$$

Because, $p + q = K$ then $j = p$, therefore when $u = j + mK$ with $1 \leq j \leq K$ the optimal solution for (5.35) which is the same solution for (5.33), is given by

$$\boldsymbol{\pi}(m, \ldots, m, m+1, \ldots, m+1) = \boldsymbol{\ell}_O \tag{5.36}$$

where there are j components with value $m+1$ and $K-j$ components with value m.

5.12 Auxiliary Algorithms

5.12.1 Obtaining the L_r and the dimension of the matrices $\hat{\mathbf{H}}_r$ (Algorithm 12)

With the purpose of determining the sets L_r in a systematic way, it is convenient to consider a matrix $\mathbf{H}$ obtained when all the entries of the coded aperture have a value of 1. This matrix is denoted by $\boldsymbol{\mathcal{H}}$. The set $L_r = \text{supp}(\boldsymbol{\mathcal{H}}(i,:))$ indicates the indexes of the columns of the matrix $\mathbf{H}$ in the support of the row i. J_r represents the row indexes of the support of the columns of $\mathbf{H}$ indicated by the set L_r in the first shot. The information contained in J_r is used to assemble the values of $\hat{\mathbf{H}}_r$ in $\mathbf{H}$.

Algorithm 12 Obtaining the sets L_r

Input: $\boldsymbol{\mathcal{H}}$

Output: L_r, J_r

Initialisation : $r = 0$

1: **for** $i = 1$ to $N(N + L - 1)$ **do**
2: $r = r + 1$;
3: $L_r = \text{supp}(\boldsymbol{\mathcal{H}}(i, :))$;
4: $J_r = i$;
5: **end for**
6: **return** L_r, J_r

5.12.2 Assembling the values of $\hat{\mathbf{H}}_r$ into $\mathbf{H}$ (Algorithm 13)

The designed values of $\hat{\mathbf{H}}_r$ can be assembled in $\mathbf{H}$ using Algorithm 13.

Algorithm 13 Assembling the values of $\hat{\mathbf{H}}_r$ into $\mathbf{H}$

Input: I_r, J_r and $\hat{\mathbf{H}}_r$
Output: $\mathbf{H}$
1: **for** $r = 1$ to R **do**
2: **for** $k = 1$ to K **do**
3: $\mathbf{H}(J_r + (k-1)N(N+L-1); L_r) = \hat{\mathbf{H}}_r(k, :)$;
4: **end for**
5: **end for**
6: **return** $\mathbf{H}$

5.12.3 Obtaining the values of the coded aperture from H (Algorithm 14)

Algorithm 14 Obtaining the values of the coded aperture from **H**

Input: **H**

Output: $\mathbf{T}^{(k)}$ for $k = 1, \ldots, K$

1: **for** $k = 1$ to K **do**

2: **for** $\ell = 1$ to N^2 **do**

3: $\mathbf{T}^{(k)}_{lex}(\ell, :) = [\mathbf{H}(\ell + (k-1)N(N+L-1), \ell), \mathbf{H}(\ell + N + (k-1)N(N+L-1), \ell + N^2), \ldots, \mathbf{H}(\ell + (L-1)N + (k-1)N(N+L-1), \ell + (L-1)N^2)]$;

4: **end for**

5: **end for**

6: **for** $\ell = 1$ to N **do**

7: $\mathbf{T}^{(k)}(:, \ell, :) = \mathbf{T}^{(k)}_{lex}((\ell-1)N + 1 : \ell N, :)$;

8: **end for**

9: **return** $\mathbf{T}^{(k)}$ for $k = 1, \ldots, K$

The term

$$\mathbf{T}^{(k)}_{lex}(\ell, :) = [\mathbf{H}(\ell + (k-1)N(N+L-1), \ell),$$
$$\mathbf{H}(\ell + N + (k-1)N(N+L-1), \ell + N^2), \ldots$$
$$, \mathbf{H}(\ell + (L-1)N + (k-1)N(N+L-1), \ell + (L-1)N^2)]$$

indicates that the spectral response of the pixel ℓ (using lexicographic ordering) in the colored coded aperture is saved in the $\ell-$row of $\mathbf{T}^{(k)}_{lex}$. The assignment $\mathbf{T}^{(k)}(:, \ell, :) = \mathbf{T}^{(k)}_{lex}((\ell-1)N + 1 : \ell N, :)$ indicates that the spectral responses in the rows from $(\ell-1)N + 1$ up to ℓN of $\mathbf{T}^{(k)}_{lex}$ are located along the second dimension in the position ℓ on the 3D-dimensional array $\mathbf{T}^{(k)}$. Then, the spectral values are contained through the third dimension of $\mathbf{T}^{(k)}$.

5.12.4 Locating the values of $\mathbf{T}^{(k)}$ in H (Algorithm 15)

The term $\text{diag}(\mathbf{T}^{(k)}(:, :, \ell))$ in line 3 of Algorithm 9, represents a diagonal matrix where the diagonal is the vectorized version of $\mathbf{T}^{(k)}(:, :, \ell)$.

Algorithm 15 Locating the values of $\mathbf{T}^{(k)}$ in $\mathbf{H}$

Input: $\mathbf{T}^{(k)}$ for $k = 1, \ldots, K$
Output: $\mathbf{H}$

Initialisation : $\mathbf{H} = []$;
for $k = 1$ to K **do**
  for $\ell = 1$ to L **do**
    $\mathbf{H}^{(k)}(1 + (\ell - 1)N : N^2 + (\ell - 1)N, 1 + (\ell - 1)N^2 : \ell N^2) = \mathrm{diag}(\mathbf{T}^{(k)}(:, :, \ell))$;
  end for

$$\mathbf{H} = \begin{bmatrix} \mathbf{H} \\ \mathbf{H}^{(k)} \end{bmatrix}$$

end for
return $\mathbf{H}$

Chapter 6

CODED APERTURE DESIGN FOR COMPRESSIVE X-RAY TOMOSYNTHESIS VIA COHERENCE ANALYSIS

6.1 Introduction

X-ray tomosynthesis imaging plays an important role in biomedical imaging applications like mammography and angiography [83]. As in most X-ray computational tomography imaging systems, radiation exposure can significantly increase the risk of adverse radiation effects, producing damages in body cells [84]. To reduce the damage that radiation can cause, different approaches have been proposed by lowering the number of angles at which projections are taken [85]. However, the consequent reduction of measurements leads to an ill-posed problem, highly sensitive to modeling and measurement errors. Moreover, the reconstructions based on filtered backprojection (FBP) with ill-posed systems of equations produces artifacts and noise which makes the reconstructions useless for medical diagnosis [86] [87].

In order to acquire measurements in parallel, coded aperture X-ray tomosynthesis was introduced in [88]. The substantial differentiation in this approach is the use of a coded aperture between the sources of radiation and the objects. This coded aperture codes the radiation signal that impinges on the object allowing a differentiation between the projections on the detector. As a consequence, multiple projections can be captured at the same time instead of capturing sequential measurements as it is done in conventional systems [88]. The projections used in [88], however, used totally random coded apertures. No coded aperture optimization was considered. The optimized design of coded apertures for the compressive X-ray tomosynthesis system was considered first in [3], the results obtained in [3] were superior to the ones obtained with totally random patterns. However, the computational complexity required for the

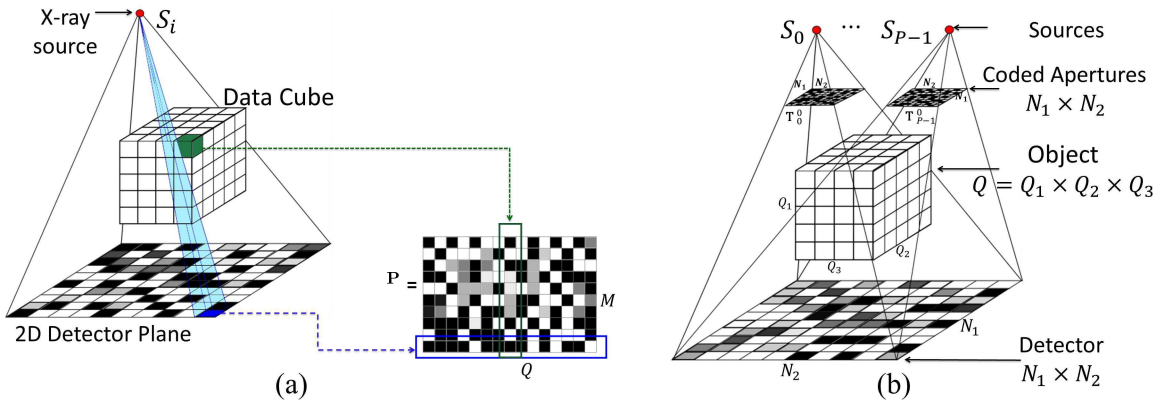


Figure 6.1: (a)The matrix **P** determines the mapping of the X-ray sources to the detector. (b) Coded aperture compressive X-ray tomosynthesis. The radiation of each source is modulated by the coded aperture.

computation of the optimized codes in [3] is high and the objective function used for that purpose is only indirectly linked with the parameters that are commonly used in compressed sensing. In this work, the coded aperture design is addressed considering the analysis of the coherence of the sensing matrix. It exploits the highly structured sensing matrix that represents the X-ray tomosynthesis architecture. The idea is to minimize the inner products between columns of the sensing matrix considering a general basis representation of the signal of interest. It is shown that, families of codes can be obtained which provide better results than the ones obtained by the use of totally random patterns, and the results can be also comparable to the ones obtained in [3]. The reduction in the computation of the solution is dramatic, because the solution is obtained in seconds whereas in [3] the time is in the order of hours.

6.2 Forward Projection Model

Let us consider an X-ray source that is located at position $\vec{s}$ and radiates an object in direction $\widehat{\theta}$, the continuous X-ray model is given by: $y(\vec{s},\widehat{\theta}) = \int_0^\infty f(\vec{s}+x\widehat{\theta})dx$, where the function f corresponds to the three-dimensional object of interest. This imaging model is known as the X-ray transform [3] [85].

Because only a discrete number of radon measurements can be acquired, the continuous model is discretized. Let $\mathbf{F} \in \mathbb{R}^{Q_1 \times Q_2 \times Q_3}$ be the three dimensional array that represents the object. The value of Q_1 indicates the number of slices of dimensions $Q_2 \times Q_3$. The detector considered is of dimension $N_1 \times N_2$ and is placed under the considered object as indicated in Fig 6.1(a).

The projection measurements are represented by the vector $\mathbf{y}$. Then, the traditional forward model in tomosynthesis can be written as $\mathbf{y} = \mathbf{P}\vec{\mathbf{F}}$ where $\vec{\mathbf{F}}$ is a vectorized version of $\mathbf{F}$, and the matrix $\mathbf{P}$ is the system matrix obtained by specifying the hardware settings. The entries of $\mathbf{P}$ correspond to the mapping of the cone-beam energy radiating from the X-ray source onto the detector [3]. As it is shown in Fig. 6.1(a), each entry of $\mathbf{P}$ represents the portion of the volume of a given voxel that is irradiated by the X-ray associated with one detector element. In particular, each row of $\mathbf{P}$ indicates the information gathered by one detector and each column corresponds to the information gathered from a single voxel [3].

In compressive X-ray tomosynthesis the measurements are multiplexed from multiple sources onto the detector. Coded apertures are located between the sources and the object to modulate the radiation of each source producing a coded projection onto the detector plane [88]. The size of the elements of the coded apertures is fixed to obtain a one to one correspondence with the detector elements [3]. Let $\mathbf{T}_i^{(k)}$ be the coded aperture related with the source i in the shot k, then the measurements $\mathbf{y}$ can be represented as

$$\mathbf{y} = \mathbf{C}\mathbf{P}\vec{\mathbf{F}} \tag{6.1}$$

where the matrix $\mathbf{C}$ is given by

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1^{(1)} & \mathbf{C}_2^{(1)} & \dots \mathbf{C}_S^{(1)} \\ \mathbf{C}_1^{(2)} & \mathbf{C}_2^{(2)} & \dots \mathbf{C}_S^{(2)} \\ & \vdots & \\ \mathbf{C}_1^{(K)} & \mathbf{C}_2^{(K)} & \dots \mathbf{C}_S^{(K)} \end{bmatrix} \tag{6.2}$$

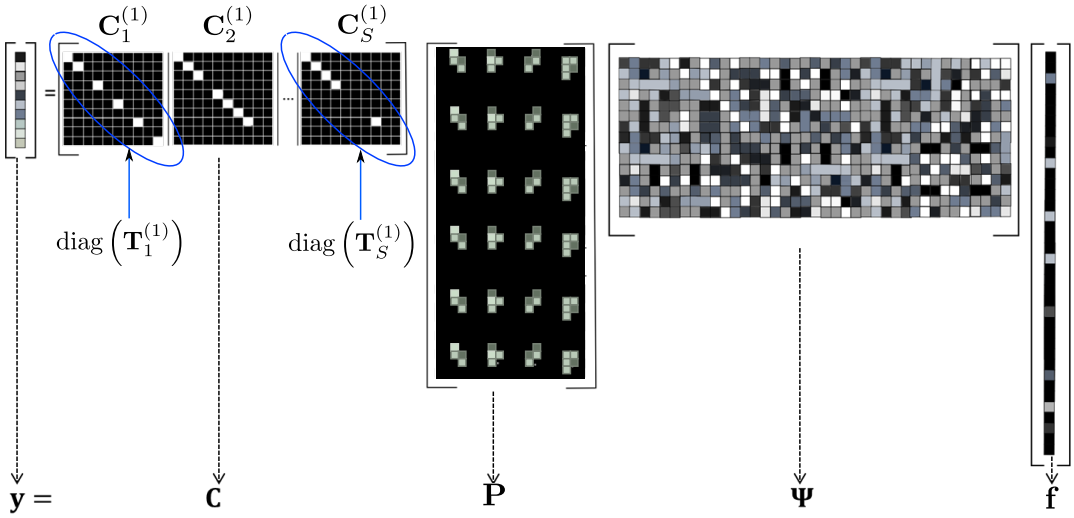


Figure 6.2: A graphical representation of the sensing matrix is depicted when $K = 1$ shots and S sources are considered. The matrix $\mathbf{C}$ is composed by the diagonalized version of the coded apertures related to each source.

where $\mathbf{C}_i^{(k)} = \text{diag}\left(\mathbf{T}_i^{(k)}\right)$. The problem (6.1) is ill conditioned and cannot be solved using traditional approaches. Compressed sensing allows the solution of this problem considering an sparse representation of $\vec{\mathbf{F}}$ in one basis $\mathbf{\Psi}$. The quality of the reconstructed solution is directly related with the coherence of the matrix $\mathbf{CP\Psi}$ [75] [77].

Let us consider the representation of $\vec{\mathbf{F}}$ in the basis $\mathbf{\Psi}$ as $\vec{\mathbf{F}} = \mathbf{\Psi f}$, where $\mathbf{f}$ is the sparse coefficients vector representation. Then, the problem (6.1) can be equivalently written as $\mathbf{y} = \mathbf{Af}$ where $\mathbf{A} = \mathbf{CP\Psi}$ is the sensing matrix. The solution of this problem via compressed sensing can be obtained as the solution of

$$\hat{\mathbf{f}} = \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{y} - \mathbf{Af}\|_2^2 + \tau \|\mathbf{f}\|_1 \tag{6.3}$$

where τ is a regularization parameter.

6.3 Analysis of the sensing matrix and Coded aperture optimization

A measure of the quality of the solutions of (6.3) in compressed sensing is given by the coherence of the sensing matrix, which is the maximum absolute value for the normalized inner products between any two columns of the sensing matrix [75] [77]. The value of this parameter is desired to be as small as possible, to guaranty unique

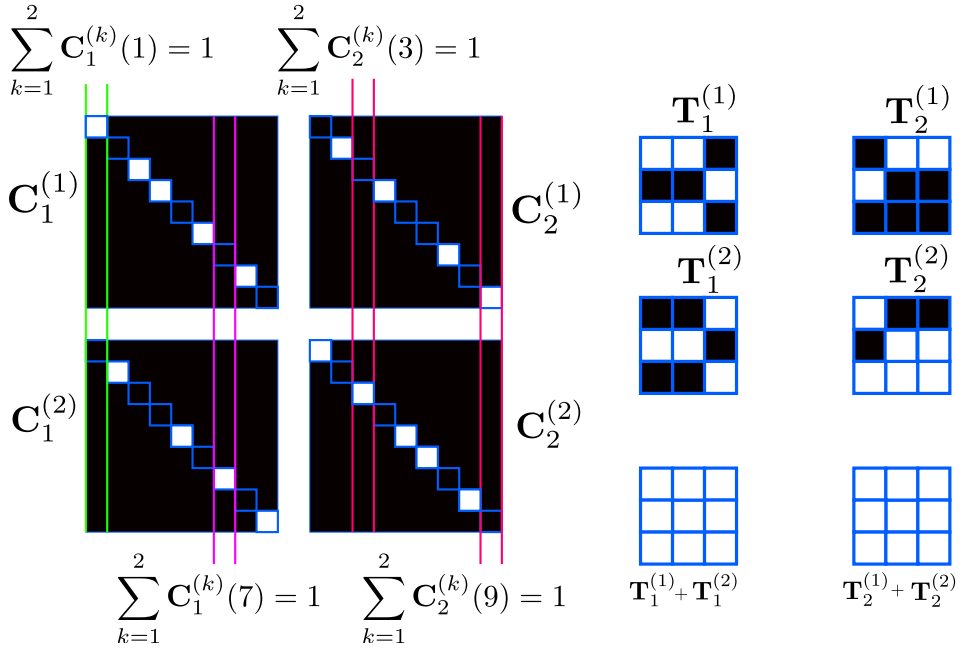


Figure 6.3: A graphical representation of the matrix $\mathbf{C}$ is depicted, showing the effects of the condition $\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u) = 1$ when the number of shots is $K = 3$ and the number of sources is $S = 2$.

recovery and low error in the numerical solution of the problem as the quality of the solution is directly related to the coherence [4].

In [89] an approach based on the analysis of the coherence was developed to exploit the structure of the sensing matrix in compressive spectral imaging, in order to increase the quality of the reconstructions. The approach proposed in [89] shows how an upper bound of the coherence can be minimized, analyzing the structure of the inner products of the transfer function matrix of the system. It is shown, that when the set of measurements is given by $\mathbf{y} = \mathbf{H}\boldsymbol{\Psi}\mathbf{f}$, it is possible to achieve a minimum for an upper bound of the coherence when the inner products for the columns of the matrix $\mathbf{H}$ are minimized. In this work, this approach is used to consider the computed tomography problem such that a designed sensing matrix is obtained minimizing the inner products in the matrix $\mathbf{H} = \mathbf{CP}$, considering arbitrary values on the entries of $\mathbf{P}$.

The inner product of the columns m and n of the matrix $\mathbf{CP}$ is given by

$$\varphi_{m,n} = \sum_{u=1}^{N^2} \sum_{k=1}^{K} \sum_{i=1}^{S} \sum_{j=1}^{S} \mathbf{C}_i^{(k)}(u) \mathbf{C}_j^{(k)}(u) Q_{u,u}^{(i,j)}(m,n) \tag{6.4}$$

where $Q_{u,u}^{(i,j)}(m,n) = \mathbf{P}_u^{(i)}(m)\mathbf{P}_u^{(j)}(n)$ with $\mathbf{P}_u^{(i)}(m)$ representing the $m^{th}-$component of the row u of the projection submatrix i related with the source i. Additionally the convention $\mathbf{C}_i^{(k)}(u) \equiv \mathbf{C}_i^{(k)}(u,u)$ is used to simplify the notation.

Equation (6.4) can be equivalently written as

$$\begin{aligned}\varphi_{m,n} = &\sum_{u=1}^{N^2} \sum_{i=1}^{S} \left(\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u) \right) Q_{u,u}^{(i,i)}(m,n) + \cdots \\ &+ \sum_{u=1}^{N^2} \sum_{i \neq j} \left(\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u) \mathbf{C}_j^{(k)}(u) \right) Q_{u,u}^{(i,j)}(m,n)\end{aligned} \tag{6.5}$$

Taking into account that $Q_{u,u}^{(i,j)}(m,n) \geq 0$, it follows that the minimum of $\varphi_{m,n}$ is achieved when the terms $\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u)$ and $\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u)\mathbf{C}_j^{(k)}(u)$ are minimized. Additionally, using the Cauchy-Schwartz inequality the following relation is obtained

$$\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u) \mathbf{C}_j^{(k)}(u) \leq \sqrt{\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u)} \sqrt{\sum_{k=1}^{K} \mathbf{C}_j^{(k)}(u)}. \tag{6.6}$$

From this relation it is possible to see that the minimization of $\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u)$ and $\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u)\mathbf{C}_j^{(k)}(u)$ is achieved by the minimization of $\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u)$. Because the entries of the coded apertures are represented by binary nonnegative entries, this minimization is achieved when[1]

$$\sum_{k=1}^{K} \mathbf{C}_i^{(k)}(u) = 1. \tag{6.7}$$

Then, the family of solutions for the coded apertures that satisfy this equality can be written as $\left[\mathbf{C}_i^{(1)}(u), \mathbf{C}_i^{(2)}(u), \ldots, \mathbf{C}_i^{(K)}(u)\right]^T = \boldsymbol{\pi}\{\text{diag}(\mathbf{1}_{K\times 1})\}_1 \quad \forall u$, where $\boldsymbol{\pi}$ represents the random permutation operator of the columns, and the subindex 1 the first column of the matrix after applying $\boldsymbol{\pi}$.

[1] The zero solution is not considered because using it would imply that there would be a voxel that is not sensed in any of the shots used in the measurement process

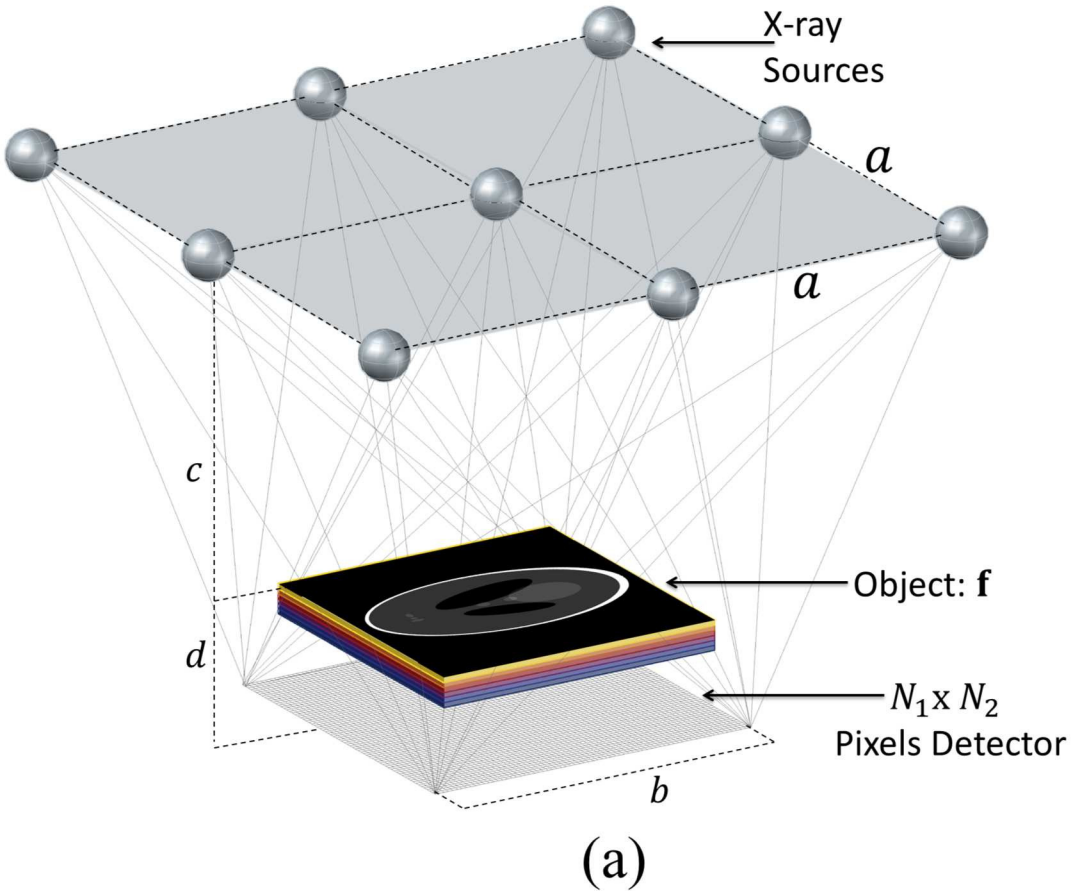


Figure 6.4: (a) Configuration for X-ray tomosynthesis simulation. The 9 sources are placed uniformly over a 128×128 phantom with 16 slices. For the simulation scenario that was studied here $a = 128, b = 128, c = 675, d = 60, e = 150$

6.4 Simulations

In order to have a precise comparison with the approach presented in [3], the same simulation scenario is considered. Then, to simulate the compressive X-ray tomosynthesis the configuration of a flat 2D detector plane composed by $N_1 \times N_2 = 150 \times 150$ elements, $S = 9$ cone-beam X-ray sources placed uniformly in a 3×3 geometry and an object of interest $\mathbf{F}$ of dimensions $Q_2 \times Q_3 \times Q_1 = 128 \times 128 \times 16$ are used. Each pixel in the coded aperture corresponds to a particular detector element as detailed in Fig. 6.4(a). Therefore, the coded apertures placed in front of each of the sources are also composed by 150×150 elements.

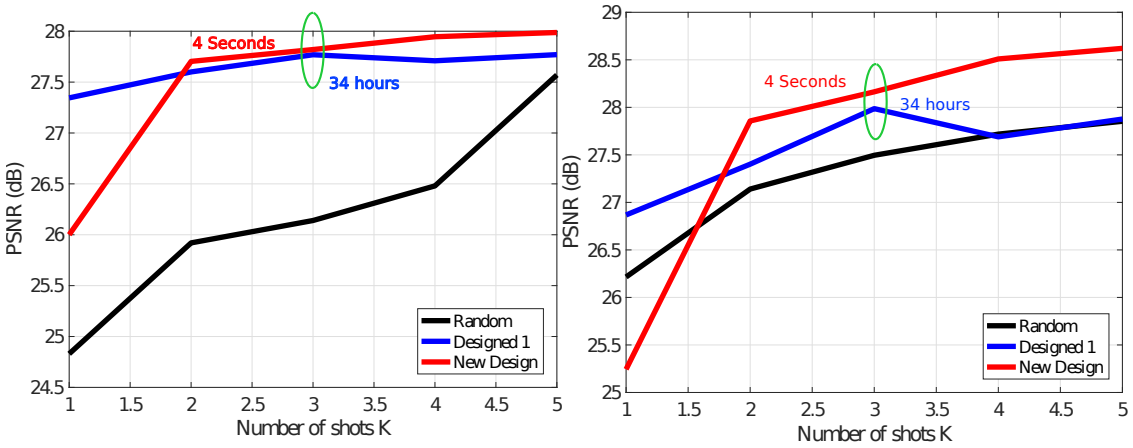


Figure 6.5: Left: PSNR of the 13^{th} slice in the reconstructed datacube. Right: The mean PSNR of the reconstructed datacube. $K = 3$ shots are considered. The results obtained in [3] are depicted in blue color whereas the new designs in red color.

The ASTRA Tomography Toolbox (All Scale Tomographic Reconstruction Antwerp) [90] is used to obtain the system matrix $\mathbf{P}$ as well as the projection measurements for each of the X-ray cone beam sources. The codes developed according to the ideas presented in Section III are generated and compared with the totally random codes and the codes generated in [3]. In the last case the algorithm developed in [3] is used to obtain a set of codes for $K = 1, 2, 3, 4, 5$ shots. The GPSR algorithm [81] is used for the reconstructions, doing an experimental tunning of the regularization parameter.

6.4.1 Results

Figure 6.5 plots the PSNR of all the reconstructions obtained. It is possible to appreciate that for a number of shots from $K = 2$ up to $K = 5$ the PSNR of the reconstructions with the designed codes is superior. Also appealing is that the time necessary to generate those codes is in the order of seconds. Figure 6.5 also shows that the time necessary to obtain the designed codes according to [3] for $K = 3$ shots is 34 hours, whereas the designs obtained with the presented approach is 4 seconds. Additionally in Figure 6.6 the slices 4^{th} and 13^{th} are showed when random codes are used against the results obtained with the new designs. In the zoomed regions it is clear that with the new designs more details in the reconstructed object can be obtained.

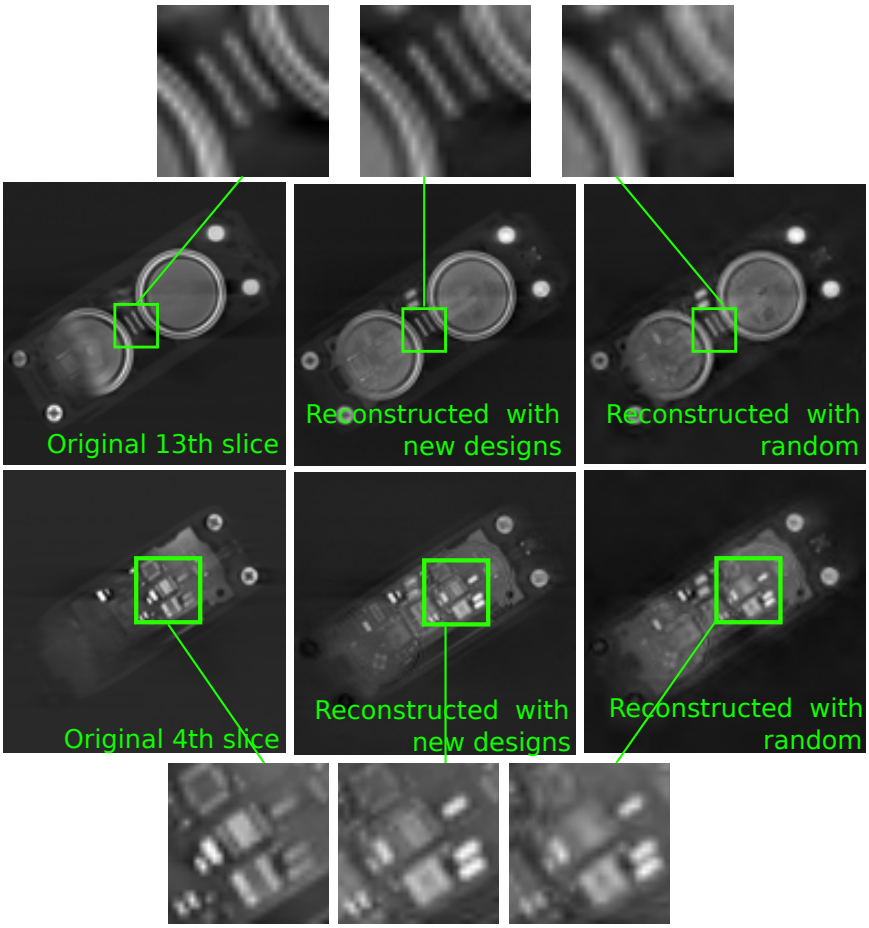


Figure 6.6: The slices 4 and 13 of the reconstructed datacube are depicted, comparing the results of using random codes versus the new design approach.

PSNR on 13^{th}-slice			
K	New approach	Approach of [3]	Random
1	**26**	27.34	24.83
2	**27.70**	27.6	25.92
3	**27.82**	27.76	26.14
4	**27.94**	27.70	26.48
5	**27.98**	27.76	27.57
Mean PSNR			
1	**25.24**	26.86	26.21
2	**27.85**	27.40	27.14
3	**28.16**	27.98	27.49
4	**28.50**	27.68	27.71
5	**28.62**	27.87	27.85

Table 6.1: The PSNR on the 13^{th} slice is indicated for the different codes used and also the mean PSNR is indicated.

K	1	2	3	4	5
Random	0.21ms	0.3ms	0.22ms	0.23ms	0.39ms
New Design	18.65ms	22.66ms	27.39ms	34.50ms	38.41ms
Designed 1	87.54s	1395s	1411s	3725s	4601s

Table 6.2: Time spent in the generation of the coded apertures for each approach and different values of K when the scene considered is of size 32x32x4.

In Table 6.1 the values of the PSNR obtained in the simulations and used in Fig. 6.5 are showed. In Figure 6.7 a sample of the coded apertures obtained by the use of the new approach can be appreciated. Additionally the time spent in the generation of the coded apertures in each approach for different number of shots is presented in Table 6.2 for an object of size 32x32x4.

6.5 Conclusions

A new strategy for the design of coded apertures in compressed X-ray tomosynthesis has been presented. The proposed approach relies on the analysis of the coherence of the sensing matrix and allows to obtain a family of designs that has a closed

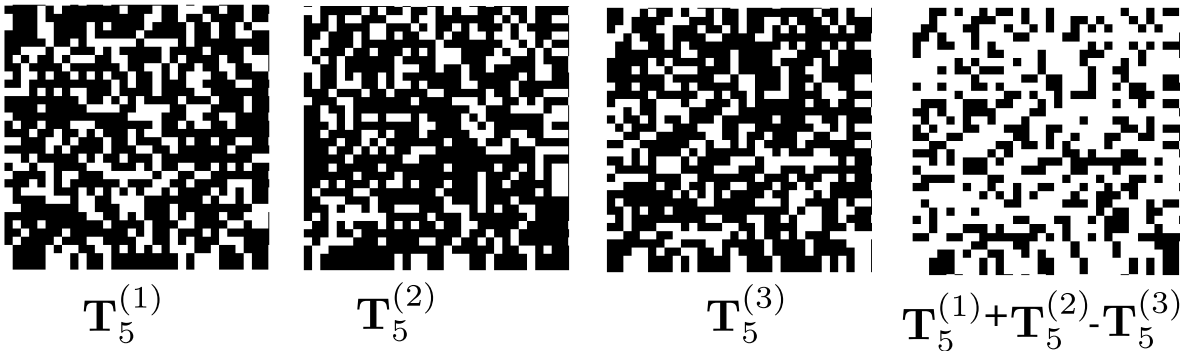


Figure 6.7: A 64×64 window of the designed coded aperture with the new design approach is depicted for the source number 5 when $K = 3$ shots are considered.

form solution. The results of the simulations, show that the designs obtained with the presented approach are better than the results obtained with totally random patterns and with the designs generated in [3]. Additionally, the time necessary to generate the designed codes in the presented approach is in the order of seconds while in [3] is in the order of hours.

Chapter 7

SPECTRAL SUPER-RESOLUTION IN COLORED CODED APERTURE SPECTRAL IMAGING

7.1 Introduction

Significant interest has emerged in compressed spectral imaging [69] and optical architectures like the coded aperture snapshot spectral imaging (CASSI) [70] [71]. The CASSI system, illustrated in Fig. 7.1, encodes the spectral-spatial information of a hyperspectral scene into 2D compressed projections. The projections are attained by spatially coding the optical field using a coded aperture. The coded optical signal is then dispersed by a prism and integrated by a focal plane array (FPA) with wide spectral response. The spectral scene is then reconstructed by solving a ℓ_1-minimization problem [69] [71]. The spectral resolution in CASSI is limited mainly by the pitch size of the detector and the spectral dispersion of the prism (Fig. 7.2(a)).

In the traditional CASSI, the coded aperture is a binary valued mask which blocks or allows the light in a pixel to pass. A modified version of CASSI known as colored CASSI (Fig. 7.1) has been introduced recently. In the colored CASSI, the coded aperture is replaced by a colored mask in which each pixel has a band pass spectral response [72] [2]. The use of this colored mask has been shown to improve the quality of the reconstructions [69] [72].

The colored coded aperture is a patterned multilayer optical coating, which physically allows the pass of specific bandwidths in different spatial locations [73] [72]. The basic idea is to have a compact two dimensional array of pixels, each one with a different spectral response. In the fabrication process of these devices many technical challenges need to be addressed such as preserving the spatial resolution that could be affected in the coating process [73].

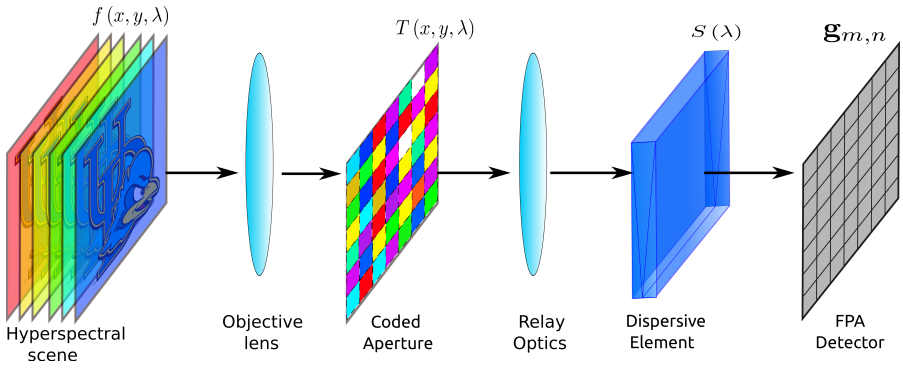


Figure 7.1: Basic Components of the colored CASSI system. The spatial-spectral scene $f(x, y, \lambda)$ is modulated by a coded aperture $T(x, y, \lambda)$, after that the modulated field is dispersed by a prism with dispersion curve $S(\lambda)$ and finally the whole field is integrated in a detector array. Notice that in the traditional CASSI $T(x, y, \lambda) \in \{0, 1\}\ \forall\lambda$.

This work shows that colored coded apertures in CASSI not only improve the quality of image reconstruction but, in addition, they allow the reconstruction of a larger number of bands, which represents an improvement in the spectral resolution. Figure 7.2(b) illustrates how colored coded apertures allow the sensing of smaller regions of the spectrum that cannot be sensed with the traditional CASSI architecture. This property is proved in two steps. First, a model of the colored CASSI is used to write the reconstruction problem as an inverse problem. In the second step, a matrix formulation of the system is used to calculate the coherence of the sensing matrix, which is used to estimate the value of the super resolution factor. This relation provides an estimate of the number of additional resolvable bands based on the wavelengths provided by the colored coded aperture. The estimate is valid even when non ideal filters are used in the colored mask. Numerical simulations confirm these facts and reconstructions with real data shows the accuracy of this approach in a real implementation of the colored CASSI.

7.2 CASSI Modeling

The basic components in CASSI are shown in Fig.7.1. The coding of the scene $f(x, y, \lambda)$ is realized by a coded aperture $T(x, y, \lambda)$ at the image plane, where

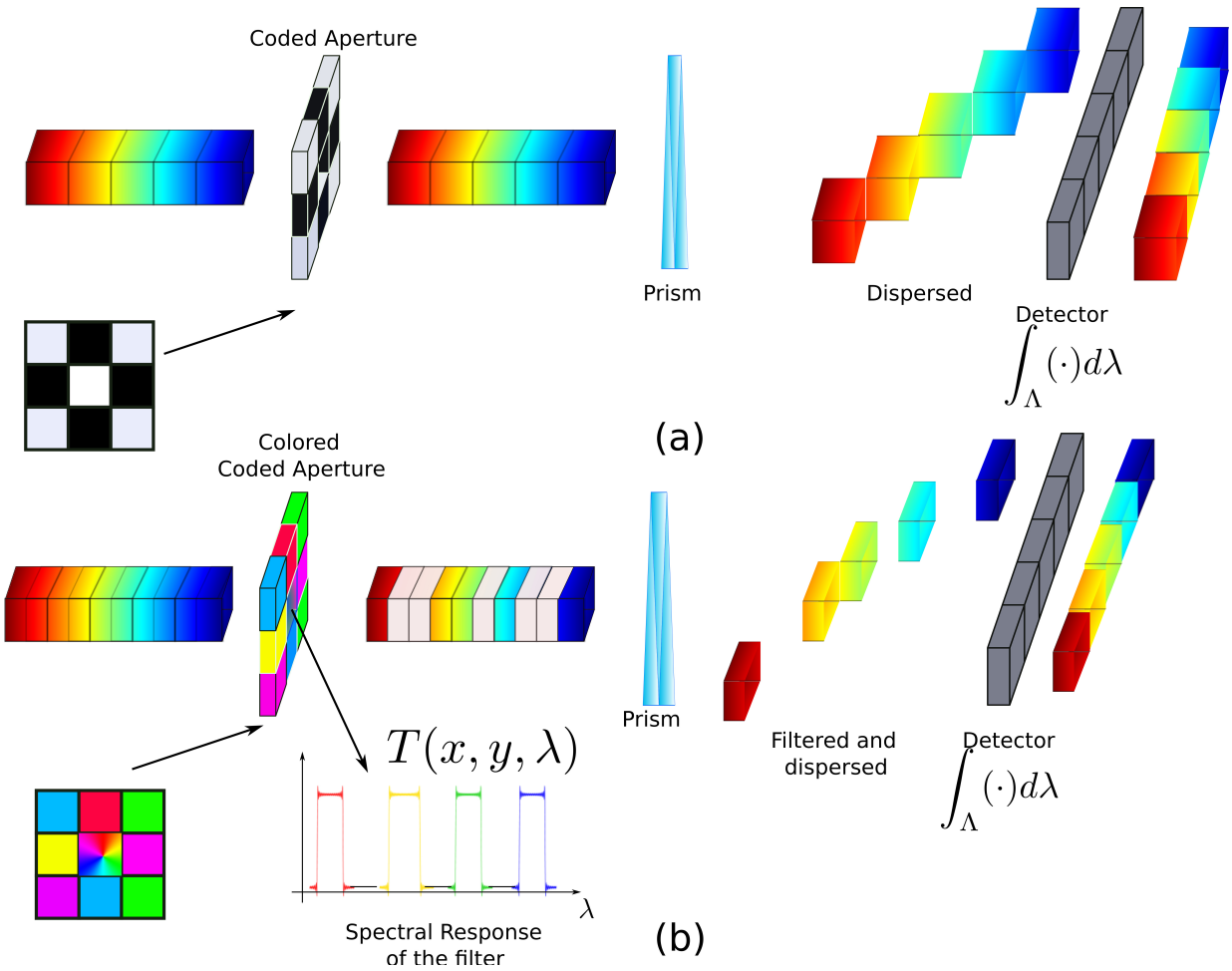


Figure 7.2: (a): The spectral resolution in the traditional CASSI is determined by the dispersion of the prism, its spectral range and the detector pixel pitch. (b): The spectral resolution in the colored CASSI is determined by the spectral response of the optical filters in the colored coded aperture, by the dispersion of the prism, its spectral range, and the detector pixel pitch.

(x, y) are the spatial coordinates and λ represents the wavelength components. When block-unblock coded apertures are used in the CASSI system, $T(x, y, \lambda)$ is such that $T(x, y, \lambda) \in \{0, 1\} \quad \forall \lambda \in \Lambda$, where Λ is the spectral range of $f(x, y, \lambda)$ [69] [71]. If $T(x, y, \lambda)$ is non constant with respect to λ, there is a representation of the CASSI with a color coded aperture. Once the spectral signal $f(x, y, \lambda)$ has been modulated by $T(x, y, \lambda)$, the optical field is dispersed by the prism and $q(x, y, \lambda)$ is obtained as $q(x, y, \lambda) = f\left(x + S(\lambda), y, \lambda\right) T\left(x + S(\lambda), y, \lambda\right)$, where $S(\lambda)$ is the dispersion curve of the prism. The compressed measurements $g(x, y)$ in the Focal Plane Array (FPA) are realized by the integration of $q(x, y, \lambda)$ across the spectral axes as $g(x, y) = \int_{\Lambda} q(x, y, \lambda) d\lambda$.

7.2.1 Discretization of the Model

Consider the following terminology that is used hereafter.

Definition 8. *Let $L^{'}$ be the number of bands that can be reconstructed by the traditional CASSI (block-unblock coded aperture). These are referred as the basic bands. Its number is given by the number of discrete detector pixels subtended by a pixel of the scene that is dispersed throughout all its spectral components.*

Definition 9. *Let L be the number of bands that can be reconstructed using the colored CASSI, the super-resolution factor d is defined as $d = L/L^{'}$.*

Since the FPA is a finite array of pixel sensors, there is just a finite number of spatial points that can be sensed. To obtain a discretization, the integral operators and the coded aperture are approximated [71].

Let $\Omega \in \mathbb{R}^2$ be the spatial domain of the FPA, which can be written as $\Omega = \cup_{m,n \in [N]} \Omega_{m,n}$, where $\Omega_{m,n}$ is defined as

$$\Omega_{m,n} = \{ (x, y) | \, \Delta(n - 1/2) \leq x \leq \Delta(n + 1/2) \, ,$$

$$\Delta(m - 1/2) \leq y \leq \Delta(m + 1/2) \} \, ,$$

and $[N] = \{1, 2, \dots N\}$. The representation of the FPA is then an array of size $N \times N$ conformed by the disjoint union of N^2 smaller domains of size $\Delta \times \Delta$, and the pixel (m, n) is associated with the corresponding domain $\Omega_{m,n}$. The value Δ represents the size of each pixel in the FPA. Let $\mathbf{g}_{m,n}$ be the value of the measurements at the (m, n) pixel, then

$$\mathbf{g}_{m,n} = \iint\limits_{\Omega_{m,n}} g(x, y)\text{rect}\left(\frac{x}{\Delta} - n, \frac{y}{\Delta} - m\right) dydx, \tag{7.1}$$

represents the contribution of $g(x, y)$ to each of the disjoint domains, $g(x, y)$ is the optical field at the image plane summed over the spectral dimension. Now, the continuous data cube $f(x, y, \lambda)$ is modulated by $T(x, y, \lambda)$ such that

$$f(x, y, \lambda)T(x, y, \lambda) \approx$$

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(TF)_{m,n}(\lambda)\text{rect}\left(\frac{x}{\Delta} - n, \frac{y}{\Delta} - m\right), \tag{7.2}$$

where $(TF)_{m,n}(\lambda) = T_{m,n}(\lambda)F_{m,n}(\lambda)$ and $T_{m,n}(\lambda)$ is the spectral response of the colored mask at pixel (m, n), and where $F_{m,n}(\lambda)$ is defined by $F_{m,n}(\lambda) = \iint\limits_{\Omega_{m,n}} f(x, y, \lambda)dxdy$. A calculation of the limits involved in the spatial integral operators and its approximation using mid point rules is realized in the Appendix 7.8, which shows that the value of the measurements at pixel (m, n) can be written as

$$\mathbf{g}_{m,n} =$$

$$\int_\Lambda \Delta^2 \left(\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor - \frac{S(\lambda)}{\Delta} + 1\right)(TF)_{m,n+\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor}(\lambda)d\lambda$$

$$-\int_\Lambda \Delta^2 \left(\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor - \frac{S(\lambda)}{\Delta}\right)(TF)_{m,n+\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor+1}(\lambda)d\lambda, \tag{7.3}$$

where $\lfloor\cdot\rfloor$ is the floor function operator. Appendix 7.8 also shows the discretization in λ, where the limits of the new bands in the super resolution model are used as the limit points of the intervals and the mid-point rule is used again for the approximation

of the integral operator in Λ, such that the value of the measurements at pixel (m, n) can be written as

$$\mathbf{g}_{m,n} = \sum_{k=0}^{L-1} \left(\mathbf{w}_{m,n,k} (\boldsymbol{T}\mathbf{F})_{m,n+\left\lfloor \frac{k}{d} \right\rfloor,k} \right), \tag{7.4}$$

where $\mathbf{w}_{m,n,k,u} = \boldsymbol{\alpha}_{m,n,k} + \boldsymbol{\beta}_{m,n,k}$ for $u = 1, \ldots, c-2$, $\mathbf{w}_{m,n,k,0} = \boldsymbol{\alpha}_{m,n,k}$ and $\mathbf{w}_{m,n,k,c-1} = \boldsymbol{\beta}_{m,n,k}$. The terms $\boldsymbol{\alpha}_{m,n,k}, \boldsymbol{\beta}_{m,n,k}$ are weights whose values are

$$\boldsymbol{\alpha}_{m,n,k} = \Delta_{\lambda(k)} \Delta^2 \left(\left\lfloor \frac{S(\hat{\lambda}_k)}{\Delta} \right\rfloor - \frac{S(\hat{\lambda}_k)}{\Delta} + 1 \right),$$

$$\boldsymbol{\beta}_{m,n,k} = -\Delta_{\lambda(k)} \Delta^2 \left(\left\lfloor \frac{S(\hat{\lambda}_k)}{\Delta} \right\rfloor - \frac{S(\hat{\lambda}_k)}{\Delta} \right),$$

Appendix 7.8 provides the definition and calculation of this expressions. The $\boldsymbol{T}_{m,n,k}$ and $\mathbf{F}_{m,n,k}$ terms are the (m, n, k) elements of the 3-dimensional arrays $\boldsymbol{T}$ and $\mathbf{F}$, respectively, whose first two dimensions indicate the spatial location and the third one indicates the spectral position. The vector $\boldsymbol{T}_{m,n,\cdot}$ is the discretized version of the spectral response of the color coded aperture at pixel (m, n). Equation (7.4) can be represented as the matrix equation $\vec{\mathbf{g}} = \mathbf{H}\vec{\mathbf{F}}$, where $\vec{\mathbf{g}}$ is the vectorized representation of $\mathbf{g}$. The matrix $\mathbf{H}$ is of dimension $KN\left(N + \left\lfloor \frac{L}{d} \right\rceil - 1\right) \times N^2 L$ (see Fig. 7.4) with K being the number of shots and $\vec{\mathbf{F}}$ is the vectorized form of $\mathbf{F}$. Figure 7.4 show the structure of $\mathbf{H}$ for the colored CASSI with a super resolution factor $d = 2$.

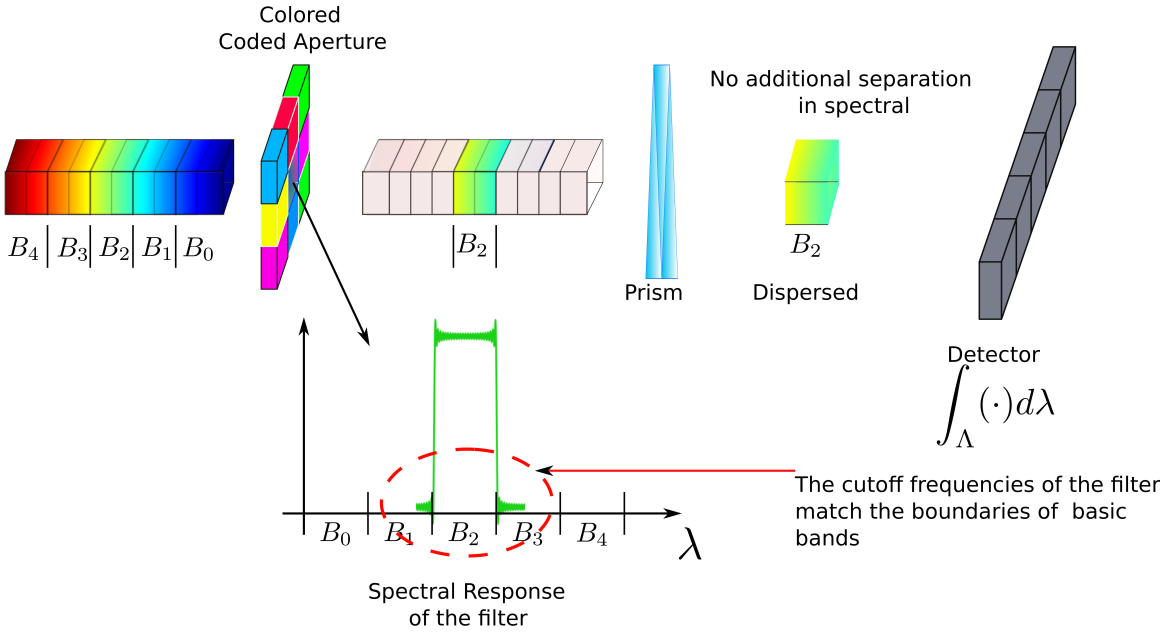
In order to use the sparsity properties of the signal $\vec{\mathbf{F}}$ in a basis $\boldsymbol{\psi}$, it is possible to write the problem as

$$\vec{\mathbf{g}} = \mathbf{H}\boldsymbol{\psi}\mathbf{f}, \tag{7.5}$$

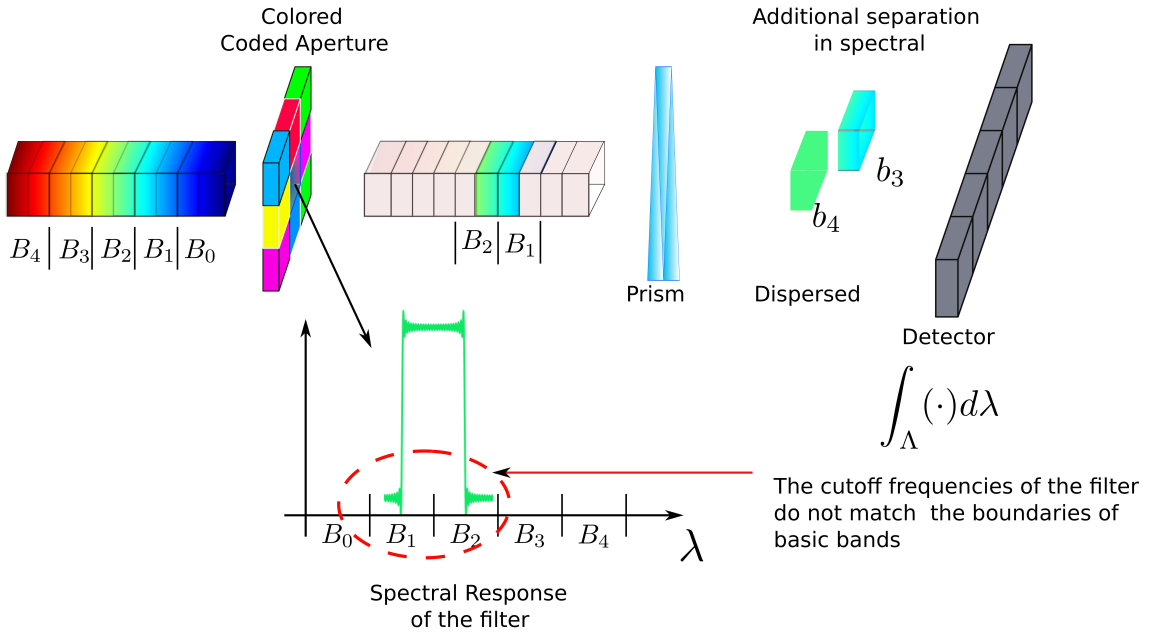
where $\vec{\mathbf{F}} = \boldsymbol{\psi}\mathbf{f}$ and $\mathbf{f}$ is a column vector whose entries are the coefficients representation in the basis. The signal recovery of $\mathbf{f}$ can be obtained as the solution of the nonlinear optimization problem [75] [71]

$$\hat{\mathbf{f}} = \underset{\mathbf{z}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{A}\mathbf{z} - \vec{\mathbf{g}} \right\|_2^2 + \tau \left\| \mathbf{z} \right\|_1, \tag{7.6}$$

where $\mathbf{A} = \mathbf{H}\boldsymbol{\psi}$ is the sensing matrix of the problem and τ a regularization parameter.



(a)



(b)

Figure 7.3: (a) A pixel of the scene is filtered by one band pass filter for which the cutoff frequencies coincide exactly with the boundaries of the basic bands. (b) A pixel of the scene is filtered by one band pass filter for which the cutoff frequencies do not coincide with the boundaries of the basic bands. In this case super resolution is achieved.

7.3 Super-Resolution Analysis

In this section, an estimation of the value of the super resolution factor d is presented. Two cases are separately analyzed. In the first case, the optical filters in the colored coded aperture have short transition bands and they are also selected as complimentary on the spectral range of interest, i.e. the supports in their spectral responses are disjoint from one filter to another. In the second case, the filters have broad transition bands and are not necessarily complementary in their spectral response. For the first case, a simple analysis is used to obtain the number of attainable new bands, whereas in the second case it is necessary to define a criteria to evaluate the quality of the reconstructions [91]. The second case will be presented in the next section after showing the expression of the coherence of the sensing matrix.

7.3.1 Optical filters with ideal transition bands

The prism in CASSI is used to separate the spectral information into bands [69] [71]. If optical filters are added to separate the spectrum into smaller sections, there is a spectral resolution increase. There are, however, different cases that depend on the number of bands and the number of filters in the colored mask. For instance, consider Fig. 7.3, in which two cases of optical filtering are considered. In the first case (Fig. 7.3 (a)) the cutoff frequencies of the band pass filter coincide exactly with the boundaries of the basic bands and therefore there is no additional separation of the spectral information. In the second case (Fig. 7.3 (b)), the cutoff frequencies of the filter do not match the boundaries of the basic bands and thus there is separation of the spectral information as a combined effect of the prism and the filter.

7.3.1.1 When the cutoff frequencies of the filters do not match the boundaries of the basic bands

Let $B = \left\{B_0, B_1, \ldots, B_{L'-1}\right\}$ represent the set of basic bands that can be recovered with the classical CASSI system, and σ the number of optical filters used in

the colored coded aperture. Then, the new number of bands that can be reconstructed $|\hat{B}|$ is given by

$$|\hat{B}| = |B| + \sigma - 1, \tag{7.7}$$

where $\hat{B} = \{b_0, b_1, \ldots, b_{L-1}\}$ is the set of new bands.

7.3.1.2 When the cutoff frequencies of the filters match the boundaries of the basic bands

In the case all the boundaries of the basic bands meet with the boundaries of the bandpass of some of the filters, the new number of attainable bands is directly related to the number of filter as

$$|\hat{B}| = \begin{cases} |B| & \text{If} \quad \sigma \leq |B| \\ \sigma & \text{If} \quad \sigma > |B|. \end{cases} \tag{7.8}$$

7.4 Coherence of the Sensing Matrix

In compressed sensing, two effective measures to predict the quality of signal reconstruction are the *restricted isometry property* (RIP), and the coherence [75] [76]. The former is in general difficult to calculate for large size matrices [75]. On the other hand, the coherence offers a measure of the ill-posedness of the system, and it must be as small as possible to guarantee uniqueness of the solution [77]. Additionally, the coherence can be related with the RIP [76] [77] [92] and, therefore, it is possible to get an analysis of the problem based on the coherence which implies an specific behavior of the RIP constant [75] [92]. The use of the coherence also allows to exploit the structure of the CASSI and colored CASSI, and to quantify the effects of super-resolution factor d as function of the spectral response of the pixels in the coded aperture.

7.4.1 Matrix Formulation

The structure of the sensing matrix $\mathbf{H}$ for the super-resolution model can be seen in Fig. 7.4. The elements that lie on the indicated diagonals represent the spectral response of the coded aperture in each band. In Fig. 7.4 the structure of $\mathbf{H}$ indicates

that there is no shifting in the transitions Band 1-Band 2 and Band 3-Band 4. The reason for this, is that in those transitions the separation of the spectral information is due to the filters in the coded aperture and not the prism. The qualitative behavior that can be appreciated in Fig. 7.4 can be represented in precise mathematical terms. The m-th position in the j-th column of the $\mathbf{H}$ matrix can be written as

$$\mathbf{h}_j(m) = \begin{cases} \left(\vec{\boldsymbol{T}}_\ell^{(s)} \vec{\mathbf{w}}_\ell\right)_{m-N(N+L-1)s-N\left\lfloor \frac{\ell}{d} \right\rfloor} \\ \text{If} \quad j = m + \ell N^2 - N\left\lfloor \frac{\ell}{d} \right\rfloor \dots \\ \qquad \dots - N(N+L-1)s \\ \qquad\quad 0 \qquad \text{otherwise,} \end{cases} \tag{7.9}$$

with $s = \left\lfloor \frac{m}{N(N+L-1)} \right\rfloor$, $\ell = \left\lfloor \frac{j}{N^2} \right\rfloor$, and $\vec{\boldsymbol{T}}_\ell^{(s)}$ represents the vectorized version of $\boldsymbol{T}(\cdot,\cdot,\ell)$ in the shot s, which is the coded aperture in band ℓ whereas $\vec{\mathbf{w}}_\ell$ is the vectorized version of $\mathbf{w}(\cdot,\cdot,\ell)$.

7.4.2 Analysis of the Coherence

Let $\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \dots \mathbf{h}_{N^2L}]$, where each $\mathbf{h}_i$ is the i^{th} column vector of $\mathbf{H}$, of dimension $KN\left(N + \left\lfloor \frac{L}{d} \right\rceil - 1\right) \times 1$. The basis matrix $\boldsymbol{\Psi}$ can be written as $\boldsymbol{\Psi} = \left[\boldsymbol{\psi}_1^T, \quad \boldsymbol{\psi}_2^T, \cdots, \boldsymbol{\psi}_{N^2L}^T\right]^T$, where each $\boldsymbol{\psi}_i$ is of dimension $1 \times N^2L$. Then, the sensing matrix can be represented as $\mathbf{A} = \mathbf{H}\boldsymbol{\Psi} = \sum_{i=1}^{N^2L} \mathbf{h}_i \boldsymbol{\psi}_i$, and the element of $\mathbf{A}$ in the (m,n) position can be written as $\mathbf{A}(m,n) = \sum_{i=1}^{N^2L} \mathbf{h}_i(m)\boldsymbol{\psi}_i(n)$.

The inner product between the columns m and n of $\mathbf{A}$ is

$$\langle \mathbf{A}(\cdot,m), \mathbf{A}(\cdot,n) \rangle = \sum_{i,j}^{N^2L} \langle \mathbf{h}_i, \mathbf{h}_j \rangle \, \boldsymbol{\psi}_i(m)\boldsymbol{\psi}_j(n).$$

Then, the coherence of the sensing matrix $\mathbf{A}$ can be written as

$$\begin{aligned} \mu(\mathbf{A}) &= \max_{\substack{m,n \\ m \neq n}} \left| \left\langle \frac{\mathbf{A}(\cdot,m)}{\|\mathbf{A}(\cdot,m)\|}, \frac{\mathbf{A}(\cdot,n)}{\|\mathbf{A}(\cdot,n)\|} \right\rangle \right| \\ &= \max_{\substack{m,n \\ m \neq n}} \mu_{mn}(\mathbf{A}), \end{aligned}$$

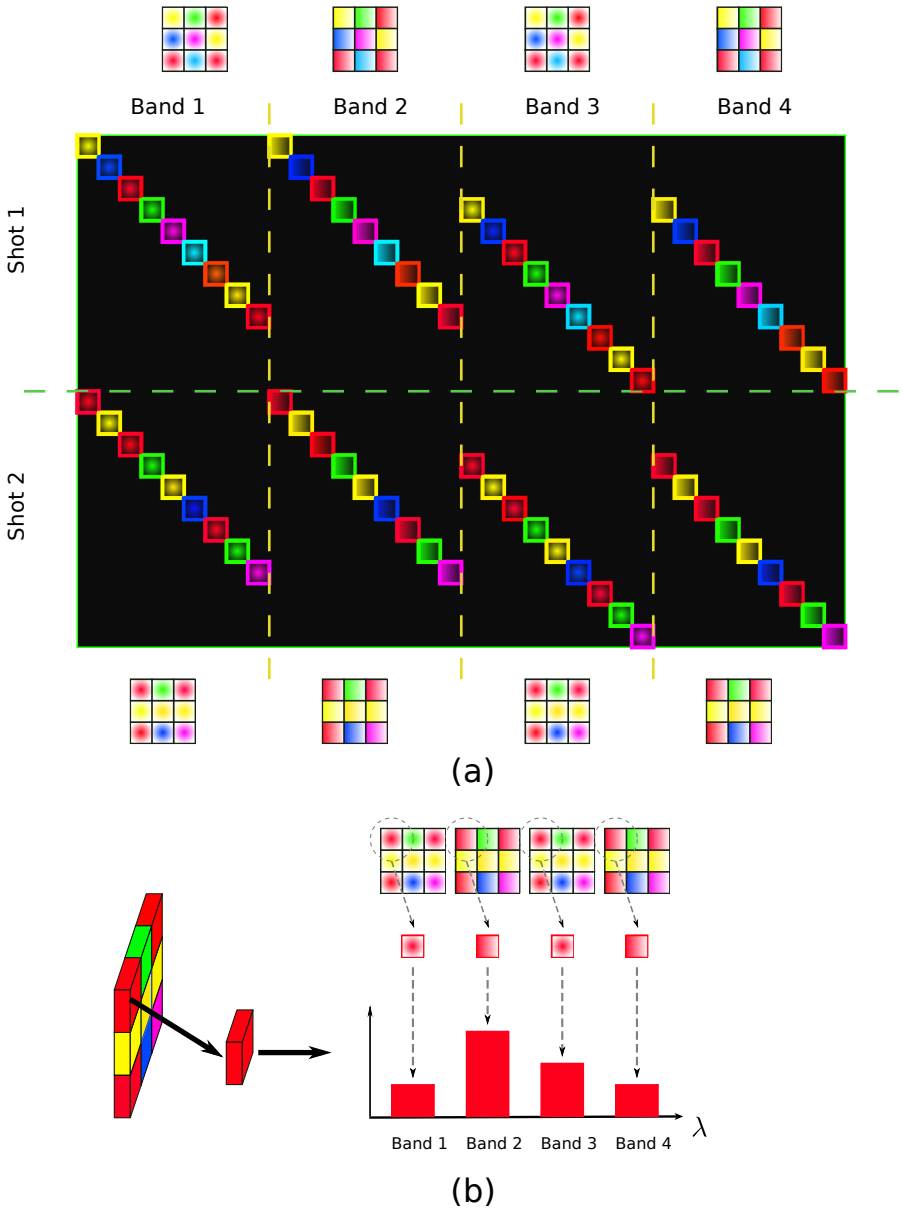


Figure 7.4: a) The structure of the **H** matrix of the colored CASSI when a super resolution factor of $d = 2$ is considered. b) The spectral response one pixel in the coded aperture is detailed and how its values are distributed in the **H** matrix.

where

$$\mu_{mn}\left(\mathbf{A}\right)=\frac{\left|\sum\limits_{i,j}^{N^2L}\varphi_{i,j}R_{i,j}^{(m,n)}\right|}{\left(\sum\limits_{i,j}^{N^2L}\varphi_{i,j}R_{i,j}^{(m,m)}\right)^{\frac{1}{2}}\left(\sum\limits_{i,j}^{N^2L}\varphi_{i,j}R_{i,j}^{(n,n)}\right)^{\frac{1}{2}}}, \tag{7.10}$$

and $R_{i,j}^{(m,n)}=\boldsymbol{\psi}_i(m)\boldsymbol{\psi}_j^*(n)$, $\varphi_{i,j}=\langle\mathbf{h}_i,\mathbf{h}_j\rangle$.

Taking into account the structure of the matrix $\mathbf{H}$ presented previously, it is possible to see that there is a set I contained in the set of all possible inner products of the columns of the $\mathbf{H}$ matrix, that can be written as $I=\left\{(i,j)|\, i\neq j,\varphi_{i,j}\neq 0\right\}$, which is the set of all possible inner products that are different from zero. Considering I, equation (7.10) can be written as

$$\mu_{mn}\left(\mathbf{A}\right)=\frac{\left|\varrho\left(m,n\right)\right|}{\varrho\left(m,m\right)^{\frac{1}{2}}\varrho\left(n,n\right)^{\frac{1}{2}}}, \tag{7.11}$$

where

$$\varrho\left(m,n\right)=\sum_{i=1}^{N^2L}\varphi_{i,i}R_{i,i}^{(m,n)}+\sum_{(i,j)\in I}\varphi_{i,j}R_{i,j}^{(m,n)}. \tag{7.12}$$

The coherence is then completely determined by the random quantities $\varphi_{i,j}$ and $\varphi_{i,i}$. In the following subsections a detailed description and analysis of these two quantities is presented. The entries of the coded aperture of the traditional CASSI are modeled using Bernoulli random variables whereas the entries of the color coded aperture are represented by a uniform random distribution over the number of filters σ in the colored coded aperture.

7.4.2.1 Block-Unblock Coded Aperture with $d=1$

Consider the coded aperture modeled as an array of i.i.d random variables. Then, let $\varphi_{i,j}=\sum_{r=1}^{K}X_rY_r \qquad (i,j)\in I$, where X_r,Y_r are the i.i.d random variables that represents the value of the pixels involved in the inner product of the columns i and j. Notice that $X_r,Y_r\in\{0,1\}$ are Bernoulli random variables with parameter p. Therefore, the probability mass function of X_rY_r is given by $f_{X_rY_r}(n)=p^{2n}\left(1-p^2\right)^{1-n} \quad n=$

$0, 1 \quad r = 1, \ldots, K$ since the sequence of products $X_r Y_r$ are i.i.d Bernoulli random variables with parameter p^2. Using Chebyshev's inequality, it follows that

$$\mathbb{P}\left(\left|\varphi_{i,j} - \mathbb{E}\left\{\varphi_{i,j}\right\}\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sum_{i=1}^{K} \mathrm{Var}\left(X_r Y_r\right) \qquad (i,j) \in I,$$

which implies

$$\mathbb{P}\left(\left|\varphi_{i,j} - (Kp)\, p\right| \geq \varepsilon\right) \leq \frac{Kp}{\varepsilon^2} p(1-p^2) \qquad (i,j) \in I. \tag{7.13}$$

In the particular case of $p = 1/2$, the following expression is obtained

$$\mathbb{P}\left(\left|\varphi_{i,j} - \left(\frac{K}{2}\right)\frac{1}{2}\right| \geq \varepsilon\right) \leq \left(\frac{K}{2}\right)\frac{3}{8\varepsilon^2} \qquad (i,j) \in I.$$

This inequality describes the behavior of $\varphi_{i,j}$ in terms of its mean in K shots, and indicates the degree of concentration of the inner product of two columns of the matrix $\mathbf{H}$. For the norm of the columns of $\mathbf{H}$, the expression is

$$\mathbb{P}\left(\left|\varphi_{i,i} - (Kp)\right| \geq \varepsilon\right) \leq \frac{Kp}{\varepsilon^2}(1-p) \quad \forall i. \tag{7.14}$$

In the particular case $p = 1/2$, the expression is

$$\mathbb{P}\left(\left|\varphi_{i,i} - \left(\frac{K}{2}\right)\right| \geq \varepsilon\right) \leq \left(\frac{K}{2}\right)\frac{1}{2\varepsilon^2}. \tag{7.15}$$

7.4.2.2 Color Coded Aperture with $d = 1$

Considering first short transition bands (ideal filters). In this case, the value of $\varphi_{i,j}$ is given by a Binomial distribution with parameters $(K, 1/\sigma^2)$, if a uniform distribution[1] is used to choose the filters. Using again the Chebyshev's inequality

$$\mathbb{P}\left(\left|\varphi_{i,j} - \left(\frac{K}{\sigma}\right)\frac{1}{\sigma}\right| \geq \varepsilon\right) \leq \left(\frac{K}{\sigma}\right)\frac{1}{\sigma\varepsilon^2}\left(1 - \frac{1}{\sigma^2}\right), \tag{7.16}$$

[1] The purpose of choosing a uniform distribution for the selection of the filters is to make a fair comparison with the traditional CASSI in which the entries of the coded aperture are modeled as two point uniform distribution (i.e. Bernoulli with parameter $p = 1/2$)

and for the norm

$$\mathbb{P}\left(\left|\varphi_{i,i}-\left(\frac{K}{\sigma}\right)\right|\geq\varepsilon\right)\leq\left(\frac{K}{\sigma}\right)\frac{1}{\varepsilon^2}\left(1-\frac{1}{\sigma}\right). \tag{7.17}$$

When wide transition band filters are used, the spectral response of one filter $T_r(\lambda)$ can take $L^{'}$ different values in the interval $[0,1]$. These values are given for a particular filter $T_r(\lambda)$ as indicated in Fig. 7.4, in which one pixel has the sequence of values $T_r(\lambda_1), T_r(\lambda_2), \ldots, T_r(\lambda_{L'})$ in each band, respectively. Then, there is a random variable that represents the entries of the $\mathbf{h}_i$ column of $\mathbf{H}$ and take its values on the set $\{T_1(\lambda_k), T_2(\lambda_k), \ldots T_\sigma(\lambda_k)\}$, $\quad k=1,\ldots,L^{'}$, where k represents the band that is associated with the $\mathbf{h}_i$ column. Therefore, the distribution function for X_rY_r is given by

$$f_{X_rY_r}(n)=\sum_{u,v=1}^{\sigma}\frac{1}{\sigma^2}\delta\left(n-T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)T_v\left(\lambda_{\lfloor\frac{j}{N^2}\rfloor}\right)\right),$$

where X_r and Y_r represent the random variables involved in the inner product of the vectors $\mathbf{h}_i$ and $\mathbf{h}_j$, and $\delta(n)$ is the delta Kronecker function. Then, again using the Chebyshev's inequality it is possible to get

$$\mathbb{P}\left(\left|\varphi_{i,j}-\left(\frac{K}{\sigma}\right)\frac{q_{11}}{\sigma}\right|\geq\varepsilon\right)\leq\left(\frac{K}{\sigma}\right)\frac{1}{\sigma\varepsilon^2}\left(q_{22}-\frac{q_{11}^2}{\sigma^2}\right), \tag{7.18}$$

where

$$q_{11}=\sum_{u,v=1}^{\sigma}T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)T_v\left(\lambda_{\lfloor\frac{j}{N^2}\rfloor}\right), \tag{7.19}$$

$$q_{22}=\sum_{u,v=1}^{\sigma}T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)^2T_v\left(\lambda_{\lfloor\frac{j}{N^2}\rfloor}\right)^2. \tag{7.20}$$

For the norm of each column the expression is

$$\mathbb{P}\left(\left|\varphi_{i,i}-\left(\frac{K}{\sigma}\right)q_2\right|\geq\varepsilon\right)\leq\left(\frac{K}{\sigma}\right)\frac{1}{\varepsilon^2}\left(q_4-\frac{q_2^2}{\sigma}\right), \tag{7.21}$$

where

$$q_2=\sum_{u=1}^{\sigma}T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)^2, \tag{7.22}$$

$$q_4=\sum_{u=1}^{\sigma}T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)^4. \tag{7.23}$$

7.4.2.3 Color Coded Aperture with $d > 1$

In this case, two situations in which $\varphi_{i,j}$ could be different from zero are considered. In the first situation, the inner products between columns are related to different *basic bands*, therefore, they can be represented with the equations (7.18), (7.19) and (7.20). In the second situation, the inner products for columns inside the basic bands are considered.

If σ ideal filters with $\sigma > L^{'}$ are used, and their cutoff frequencies match exactly with the boundaries of the basic bands, it is possible to use equations defined for the analysis with $d = 1$ and ideal filters. If real filters are used, the inner product between two columns in the same basic band is described by the probability mass function[2]

$$\sum_{u=1}^{\sigma} \frac{1}{\sigma} \delta \left(n - T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \right) T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \pm \ell \frac{\Lambda_{\lfloor \frac{i}{N^2} \rfloor}}{d} \right) \right),$$

where $\Lambda_{\lfloor \frac{i}{N^2} \rfloor}$ represents the support of the $\lfloor \frac{i}{N^2} \rfloor$ basic band and $\ell = 1, \ldots, d-1$. Therefore, taking into account the Chebyshev's inequality it is possible to get

$$\mathbb{P} \left(\left| \varphi_{i,j} - \left(\frac{K}{\sigma} \right) \hat{q}_1 \right| \geq \varepsilon \right) \leq \left(\frac{K}{\sigma} \right) \frac{1}{\varepsilon^2} \left(\hat{q}_2 - \frac{\hat{q}_1^2}{\sigma} \right), \tag{7.24}$$

where

$$\hat{q}_1 = \sum_{u=1}^{\sigma} T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \right) T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \pm \ell \frac{\Lambda_{\lfloor \frac{i}{dN^2} \rfloor}}{d} \right), \tag{7.25}$$

and

$$\hat{q}_2 = \sum_{u=1}^{\sigma} T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \right)^2 T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \pm \ell \frac{\Lambda_{\lfloor \frac{i}{dN^2} \rfloor}}{d} \right)^2. \tag{7.26}$$

The concentration equations for the norm of the columns are still the same as when $d = 1$.

---

[2] This inner product for columns inside the basic bands is the most relevant because it the represents the worst case scenario in terms of the possible values of the inner product between any two columns of the **H** matrix

7.4.3 Quality of reconstructions

Given that the coherence $\mu(\mathbf{A})$ is a random variable when the elements of the coded aperture are chosen in a random way, it is possible to use concentration inequalities in equations (7.11) and (7.12) to compare the coherence of CASSI and colored CASSI. Since the probability space of $\mu(\mathbf{A})$ is contained in $\mathbb{R}_+$, it follows that [93]

$$\mathbb{P}\left(\mu(\mathbf{A}) > \varepsilon\right) \leq \frac{\mathbb{E}(\mu(\mathbf{A}))}{\varepsilon}. \tag{7.27}$$

Then, given a fixed value of ε, it is possible to compare two configurations of the CASSI using the equation (7.27), in order to get a measure of the performance of the system [80] [78]. In this work, it is taken into account the fact that $\mu(\mathbf{A}) = \max_{\substack{m,n \\ m \neq n}} \mu_{mn}(\mathbf{A})$, where $\mu_{mn}(\mathbf{A})$ is specified as before. Thus, is possible to establish a comparison between two configurations of the CASSI represented by matrices $\mathbf{A}$ and $\mathbf{A}_s$ through a comparison between $\mu_{mn}(\mathbf{A})$ and $\mu_{mn}(\mathbf{A}_s)$. This comparison is done considering the variables $\mu_{mn}(\mathbf{A})^2$ and $\mu_{mn}(\mathbf{A}_s)^2$, and thus, it implies a direct comparison between the values of the coherence for the sensing matrix of both architectures. In this way, it is possible to know which value of the coherence parameter is concentrated around an small value with higher probability.

To this end, the relationship between the coherence for the traditional CASSI and the colored CASSI is determined first.

Theorem 52. *Let* $\mathbf{A}$ *be the sensing matrix in the traditional CASSI system and* $\mathbf{A}_\sigma$ *be the sensing matrix in the colored CASSI, considering reconstructions of scenes of dimensions* $N \times N \times L'$. *There exists a constant* $\xi \in \mathbb{R}_+$ *such that*

$$\mathbb{P}\left\{\mu_{mn}(\mathbf{A})^2 \geq \varepsilon\right\} \leq \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\xi \tag{7.28}$$

$$\mathbb{P}\left\{\mu_{mn}(\mathbf{A}_\sigma)^2 \geq \varepsilon\right\} \leq \frac{\mathbb{E}\left\{\varrho_\sigma(m,n)^2\right\}}{\varepsilon}\xi \tag{7.29}$$

$\forall m \neq n$.

Proof: See Appendix 7.10.

This result establishes a method to compare the quality of the reconstructions considering the concentration of the coherence as a consequence of the concentration of $\mu_{mn}(\mathbf{A})^2$. In order to make this comparison, it is necessary to introduce the following definitions

Definition 10. *Let K be the number of shots in a configuration of the CASSI system for the reconstruction of scenes of dimensions $N \times N \times L^{'}$, and let t be the transmittance per shot. The $\mathcal{V}$ factor is defined as $\mathcal{V} = Kt$, where the t is calculated as*

$$t = \left(\sum_{i=1}^{N^2} \int_{\Lambda} T_i(\lambda) d\lambda \right) \Big/ \Lambda N^2, \tag{7.30}$$

and $T_i(\lambda)$ is the spectral response of the i^{th}pixel.

Definition 11. *Let K be the number of shots used in a CASSI system for the reconstruction of scenes of dimensions $N \times N \times L^{'}$. The compression factor is defined as $\mathcal{C} = K/L^{'}$.*

Considering these definitions it is possible to establish the following result.

Theorem 53. *Let $\varphi_{i,j}$ be as specified in equation (7.18) for the sensing matrix in the CASSI system and consider the reconstruction of a scene of dimensions $N \times N \times L^{'}$ and factor $\mathcal{V}$. Let $\varphi_{i,j}^{(\sigma)}$ be as specified in equation (7.18) for the sensing matrix of the colored CASSI and consider the reconstruction of scenes of dimensions $N \times N \times L^{'}$ and factor $\mathcal{V}$. If $\mathbb{E}\left(\varphi_{i,j}^{(\sigma)}\right) \leq \mathbb{E}\left(\varphi_{i,j}\right)$ and $Var\left(\varphi_{i,j}^{(\sigma)}\right) \leq Var\left(\varphi_{i,j}\right)$, then*

$$\mathbb{E}\left\{\varrho_{\sigma}(m,n)^2\right\} \leq \mathbb{E}\left\{\varrho(m,n)^2\right\} \quad \forall m \neq n. \tag{7.31}$$

Proof: See Appendix 7.11.

It is important to remark that if the factor $\mathcal{V}$ is equal for both systems, this indicates that the same number of voxels are sensed in both architectures.

Taking into account this result, it is possible to establish the relation of the colored CASSI and the CASSI in the following corollary.

Corollary 54. *Let* $\mathbf{A}$ *be the sensing matrix for the CASSI system, and consider the reconstruction of a scene of dimensions* $N \times N \times L'$. *Let* $\mathbf{A}_\sigma$ *be the sensing matrix of the colored CASSI and take the same* $\mathcal{V}$ *factor for both systems. If* $\sigma > 2$ *then*

$$\mathbb{E}\left\{\varrho_\sigma(m,n)^2\right\} \leq \mathbb{E}\left\{\varrho(m,n)^2\right\}. \tag{7.32}$$

Proof: See Appendix 7.11.

This corollary and Theorem 52 imply that the quality of the reconstruction in the colored CASSI is better than in the traditional CASSI, when the same $\mathcal{V}$ factor is considered for both systems.

7.4.4 On the super-resolution factor d

As shown in the previous theorems, the functions $\varphi_{i,j}$ and $\varphi_{i,i}$, allow to establish a way to compare the performance of two configurations of the CASSI in terms of the coherence of the sensing matrix. This fact is used to compare the performance of the colored CASSI when super resolution is required, with the traditional CASSI with no super resolution comparing the functions $\varphi_{i,j}$ and $\varphi_{i,i}$ of these two configurations. In this manner, the estimate of the maximum value of d is obtained, according to the number and characteristics of the spectral responses of the filters involved in the coded aperture.

In [94], it was proved how the resolution of filter array based spectrometers, could be recalculated when DSP techniques are used to process the collected data. The approach used in [94], show that even in the case of non ideal spectral filter responses, it is possible to get more resolution. The way the authors of that work estimate a super-resolution factor, is by brute force search and do not exploit the properties of the transfer function matrix of the system. In this Section, an estimate of the super resolution factor d is proposed analyzing the behavior of the functions $\varphi_{i,j}$ and $\varphi_{i,i}$, always making a comparison with the traditional CASSI architecture considering $t = 1/2$, which represents a transmittance of 50% in the black and white coded aperture.

Before doing this comparison, it is necessary to establish a relationship between the parameters $\mathcal{C}$ and $\mathcal{V}$ for CASSI and colored CASSI with super resolution factor d. The parameter $\mathcal{V}$ gives the proportion of the set of voxels sensed with the total number of shots used in the measurement process. Therefore, it is natural to consider the same value of this factor for both architectures, which implies $K_\sigma t_\sigma = Kt$, where K, K_σ represents the number of shots and t, t_σ the transmittance per shot in each architecture, respectively. On the other hand, the compression factor is also considered as the same for both architectures, which implies $K_\sigma/dL = K/L$ and therefore $K_\sigma = dK$.

The value of d is estimated as follows.

Definition 12. *If d represents the super resolution factor described in equations (7.4) and (7.6), the value of d when real optical filters are used in the color coded aperture is estimated as follows: $d \in \mathbb{N}$ in the model (7.4) is the highest value such that*

$$d \leq \min\left\{\hat{d}_1, \hat{d}_2\right\} \tag{7.33}$$

with

$$\hat{d}_1 = \frac{\sigma}{4\hat{q}_1}, \quad \hat{d}_2 = \frac{3\sigma}{16\left(\hat{q}_2 - \frac{\hat{q}_1^2}{\sigma}\right)} \tag{7.34}$$

where the values of $\hat{q}_1(d,\sigma), \hat{q}_2(d,\sigma)$ are specified by equations (7.25) and (7.26).

The definition of d in equation (7.33) indicates that it is possible to get a new number of bands $L = dL^{'}$, always that the concentration of the random quantities $\varphi_{i,j}$ and $\varphi_{i,i}$ that determine the condition of the sensing matrix is tighter than in the traditional CASSI with no super resolution. The idea is to make a comparison of the expected value of this quantities taking into account the variance.

Equation (7.33) indicates how large d could be in model (7.4), with a solution of equation (7.6) with better quality results than in the traditional CASSI with no super resolution.

7.4.5 Impact of measurement noise on d

Previous analysis did not consider the effect of the noise on the attained super-resolution. Ideally, increasing σ with the appropriate set of filters, and the appropriate $\mathcal{C}$ and $\mathcal{V}$ would imply an increased spectral resolution. However, in real implementations with noise, distortions and other non ideal characteristics of the hardware are always present. In this situation, increasing σ could imply a reduction on the bandwidth of the filters used, which reduces the light throughput and the Signal to Noise Ratio (SNR) in the captured measurements.

The effect of the noise in super-resolution can introduce a natural limitation on the super-resolution factor as discussed next.

Let $\mathbf{e}$ be the noise added to the model in (7.5) which is considered as independent of the sensing matrix $\mathbf{A}$. The measurements are given by

$$\vec{\mathbf{g}} = \mathbf{A}\mathbf{f} + \mathbf{e}. \tag{7.35}$$

For any reconstruction algorithm, represented by the operator Δ, it follows that [4] [75]

$$\|\Delta\left(\mathbf{A}\mathbf{f} + \mathbf{e}\right) - \mathbf{f}\|_2 \leq C_{\mathbf{A}}\|\mathbf{e}\|_2, \tag{7.36}$$

for all $\mathbf{f} \in \Sigma_{2k}$, where Σ_{2k} is the set of all $2k-$sparse signals and $C_{\mathbf{A}} \in \mathbb{R}_+$. The constant $C_{\mathbf{A}}$ can be related with the RIP constant as $C_{\mathbf{A}} = 1/\sqrt{1-\delta_{2k}(\mathbf{A})}$ [4] [75]. Let $\mathbf{A}$ be the sensing matrix of the black and white CASSI, and $\mathbf{A}_\sigma$ be the sensing matrix of the colored CASSI. A limit of the performance with $\mathbf{A}_\sigma$ in the presence of $\mathbf{e}$ is given by the comparison of the upper bounds $C_{\mathbf{A}}\|\mathbf{e}\|_2$ and $C_{\mathbf{A}_\sigma}\|\mathbf{e}\|_2$. Super-resolution is then achieved whenever $C_{\mathbf{A}_\sigma} \leq C_{\mathbf{A}}$, which can be equivalently formulated as

$$\frac{\sqrt{1-\delta_{2k}(\mathbf{A})}}{\sqrt{1-\delta_{2k}(\mathbf{A}_\sigma)}} \leq 1. \tag{7.37}$$

When the matrix $\mathbf{A}_\sigma$ has columns with equal norm, (7.37) can be written as[3]

$$\frac{\sqrt{1-\delta_{2k}(\mathbf{A})}}{\sqrt{\alpha\left(1-(2k-1)\mu(\mathbf{A}_\sigma)\right)}} \leq 1, \tag{7.38}$$

[3] The fact that in a unit norm column matrix, the coherence and the RIP are related is used [4], [75]

where α is the value of the norm of the columns of $\mathbf{A}_\sigma$. Notice that $\mathbf{A}_\sigma = \alpha\tilde{\mathbf{A}}_\sigma$ where $\tilde{\mathbf{A}}_\sigma$ is a normalized column version of $\mathbf{A}_\sigma$.

In (7.38), it is possible to see that two sensing matrices $\tilde{\mathbf{A}}_\sigma$ and $\mathbf{A}_\sigma$ with equal coherence $\mu(\tilde{\mathbf{A}}_\sigma) = \mu(\mathbf{A}_\sigma)$, lead to different results in the presence of noise, as (7.38) is satisfied in different ways for each case. Additionally, when $\alpha < 1$ there is one point α_0 at which inequality (7.38) is not satisfied for all $\alpha < \alpha_0$.

7.4.5.1 About the light throughput

The total energy related with an specific filter in the coded aperture can be estimated approximately as $|\operatorname{supp}(T(\lambda))|t_\sigma$, where $|\operatorname{supp}(T(\lambda))|$ is the support of the filter and t_σ the gain (maximum amplitude) of the filter. Therefore, an scaling of a sensing matrix $\mathbf{A}_\sigma$ can be considered as an scaling on the product $|\operatorname{supp}(T(\lambda))|t_\sigma$. In this way it is possible to see that a reduction in the support of the filters can be interpreted as an scaling of the sensing matrix with a factor $\alpha < 1$. This means that when the number of filters is increased such that the support of each filter is reduced, the presence of noise establishes one point α_0 at which equation (7.38) is not satisfied for all $\alpha < \alpha_0$. This also indicates that if a given σ is required, the filters selected should have a support large enough to deal with the presence of noise.

In Figure 7.5 it is possible to see two possible choices for the set of filters in a colored coded aperture where $\sigma = 4$. In Figure 7.5(a) the filters used allow the pass of 50% of the energy in the spectral range of interest Λ, whereas in Figure 7.5(b) the filters selected are complementary on Λ and allow the pass of just 25% of the energy in Λ.

It is also important to remark that an scaling of the sensing matrix $\mathbf{A}_\sigma$, as an scaling on the product $|\operatorname{supp}(T(\lambda))|t_\sigma$ can be also considered as an scaling on t_σ which means a change in the gain of the filters. This implies that two different configurations of the CASSI with filters with identical support but with different gains could lead to the same values of $\mu(\mathbf{A})$ but the quality of their reconstructions in the presence of noise is different.

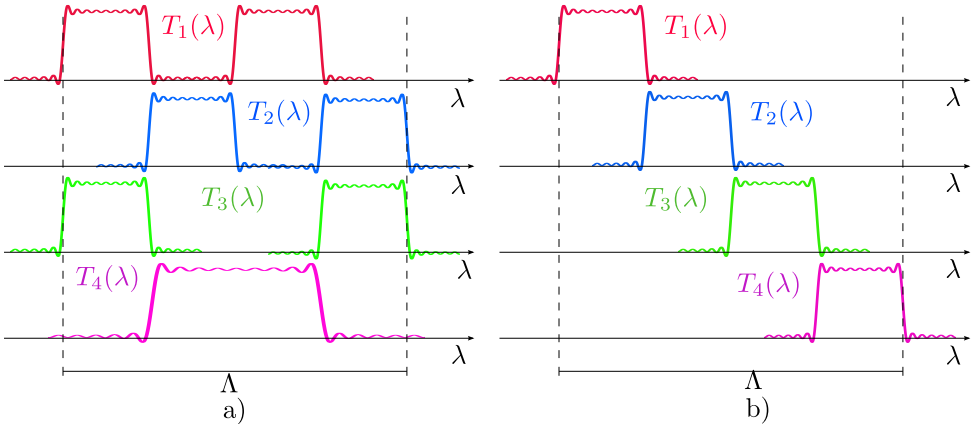


Figure 7.5: Two different choices of the set of filters σ in the case that $\sigma = 4$. (a) The set of filters cover the spectral range of interest Λ, each filter allows the transmission of the same amount of energy which corresponds to 50% of the energy contained in Λ. (b) The set of filters cover the spectral range of interest Λ, each filter allows the transmission of the same amount of energy which corresponds to 25% of the energy contained in Λ.

In the next section, simulation results showing the effect on the quality of the reconstructions when complementary filters are used in presence of noise, with respect to non complementary filters are presented.

7.5 Simulations

7.5.1 Parameters of the simulations

7.5.1.1 Multispectral scene

A datacube of dimensions $128 \times 128 \times 10d$ is generated for $d = 2, 3$. The RGB representation of this target can be appreciated in the Figure 7.6 (a). This target is artificially crated in order to have spectral responses with peaks and fast transitions, this is done with the purpose to appreciate the performance of the colored CASSI with respect to the classical CASSI in the super-resolution problem.

7.5.1.2 Spectral responses of the filters

In order to represent the spectral responses of the filters, combinations of Butterworth transfer functions with order 3 are used. This is done with the purpose of modeling realistic optical filters. The set of filters selected in each simulation covers

the spectral range of interest $[450nm - 689nm]$. The number of these filters is changed in order to appreciate its effect on the simulations.

7.5.1.3 Prism Curve

The prism curve used for simulations is a realistic prism curve adapted from [82], in order to emulate as close as possible the real nonlinear behavior of the prism.

7.5.1.4 Reconstruction algorithm and basis used in simulations

In order to represent the multispectral scene in terms of a basis, a DCT basis $\mathbf{\Psi}_{DCT}$ is selected for the spectral domain, whereas a wavelet $\mathbf{\Psi}_W$ for the spatial domain, such that the whole basis is represented as $\mathbf{\Psi}_{DCT} \otimes \mathbf{\Psi}_W$. The GPSR algorithm [81] is used for the reconstructions. The regularization parameter τ is chosen in an empirical way, so reconstructions are performed for different values of τ and the final result is selected as the one in which higher PSNR is obtained.

7.5.1.5 The measure of the quality

In order to measure the quality of the reconstructions, the *Peak Signal to Noise Ratio* (PSNR) is used. The comparison is made between the reconstructed hyperspectral scene and the *ground truth*, which is given by the original datacube generated for the simulations. Additionally, sample points of the image are selected in order to check the quality of the reconstructions in spectral.

7.5.2 Reconstructions

In order to illustrate the performance of the super-resolution, the case in which $d = 2$ is considered. In Figures 7.6 and 7.7 the results of simulations are presented.

In Figure 7.6(a) an RGB representation of the target used in simulations is depicted, indicating four points p_1, p_2, p_3, p_4 for which the spectral response is considered. Figure 7.6(b) shows the reconstructed spectrum at points p_1, p_2, p_3, p_4, considering the ground truth (blue line), the colored CASSI (red line), and the classical CASSI (black line). Because the CASSI can reconstruct just 10 bands, interpolation is used in order

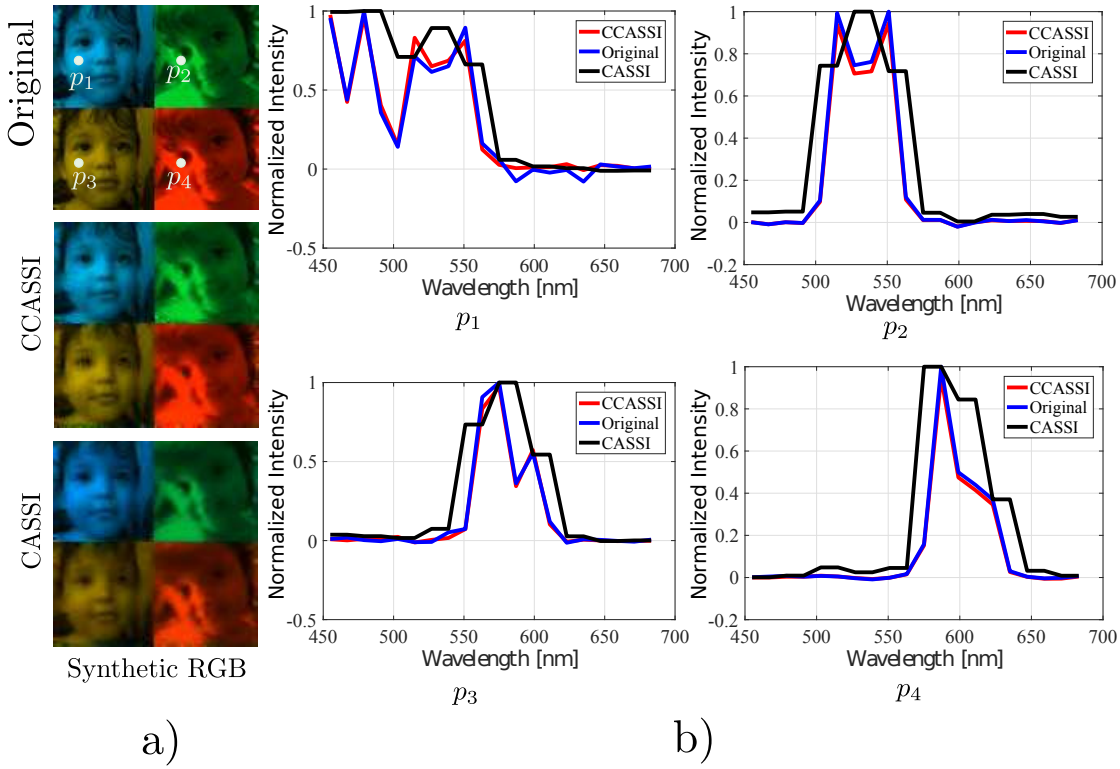


Figure 7.6: Results of the simulations considering a super resolution factor $d = 2$. (a) First row: the original target as an RGB representation of the hyperspectral scene. Second row: The RGB representation of the hyperspectral scene reconstructed by the colored CASSI. Third row: The RGB representation of the hyperspectral scene reconstructed by the traditional CASSI. (b) Comparison of the reconstructed spectral responses at points p_1, p_2, p_3, p_4 indicated in the target, using the colored CASSI and the traditional CASSI.

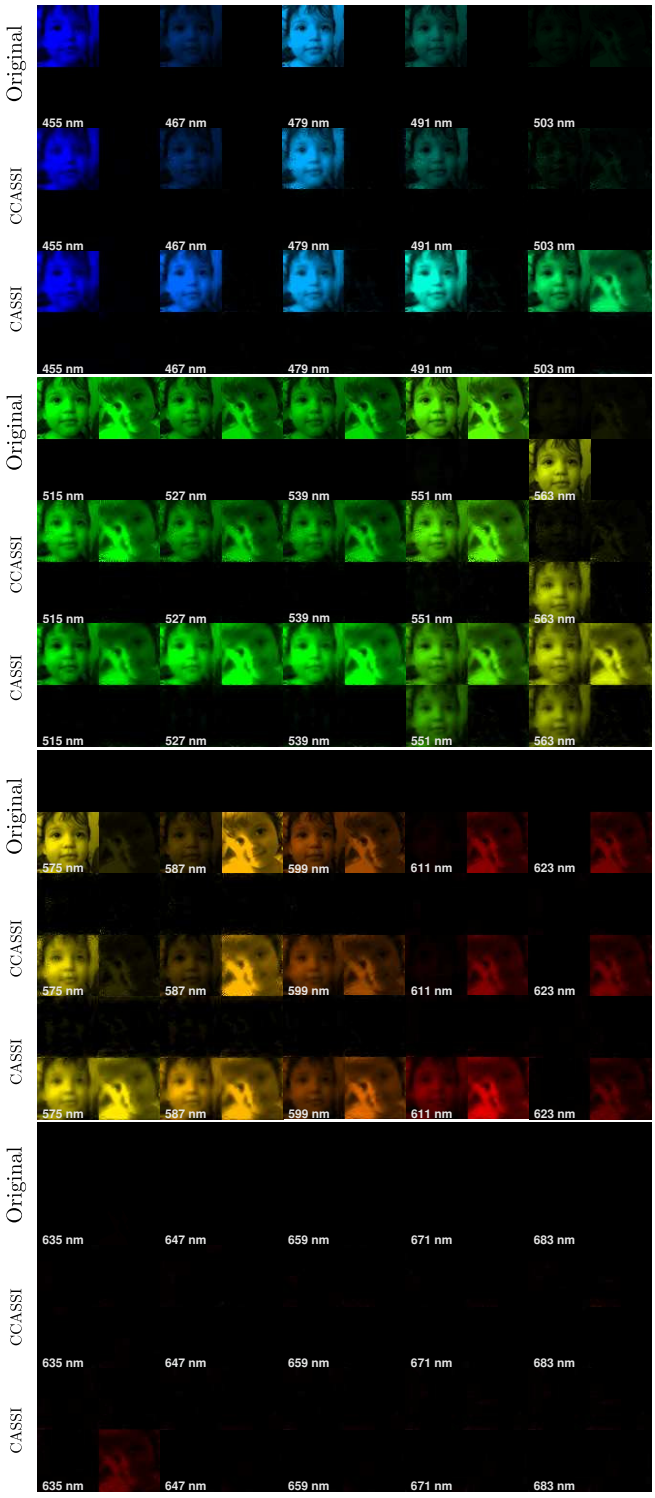


Figure 7.7: Results of the simulations considering a super resolution factor $d = 2$. The number of shots used for the traditional CASSI is 5 whereas the number of shots used for the colored CASSI is 10, such that both architectures have the same $\mathcal{C} = 0.5$. The first row of each group depict the original bands (ground truth) of the target. The second row of each group depict the reconstructed bands using the colored CASSI (CCASSI) considering $d = 2$. The third row of each group depict the bands obtained with the traditional CASSI reconstructing 10 bands and interpolating in order to get a new set of bands.

$K\backslash\sigma$	2	4	6	8
4	1	0.8571	0.7599	0.6470
6	1	0.8517	0.7517	0.6460
8	1	0.8529	0.7073	0.6044
10	1	0.8366	0.6664	0.5779
$d = 2$				
4	1	0.9997	0.7902	0.7109
6	1	0.9996	0.7532	0.6580
8	1	0.9996	0.7089	0.6060
10	1	0.9995	0.6810	0.6079
$d = 3$				

Table 7.1: The values of $\mu(\mathbf{A})$ related with the Figure 7.8. K and σ indicate the number of shots and the number filters used respectively.

to display 20 bands for the comparison with the colored CASSI with super-resolution. It is clear from the pictures that the results obtained with the colored CASSI are closer to the ground truth than the results obtained by the use of the basic CASSI.

In Figure 7.7 it is possible to see all the reconstructed bands using the colored CASSI and the basic CASSI in comparison with the ground truth. The performance of the colored CASSI with super resolution is consistently better with respect the basic CASSI. At the $503[nm]$ band for instance the colored CASSI reconstructs the band accurately. In the reconstructed band of the basic CASSI, false information is present, which is not present in the ground truth. This same behavior occurs in the results in bands such as $551[nm], 563[nm], 575[nm], 587[nm], 611[nm], 635[nm]$.

7.5.3 Simulations: Mean PSNR vs σ

In Figure 7.8(a) it is possible to see how the PSNR of the reconstructions changes according to the values of σ. In this simulations no noise is added to the measurements. It is clear that under this context, the quality of reconstructions improve with higher values of σ with the appropriate factors $\mathcal{V}$ and $\mathcal{C}$. Figure 7.8(b) shows the spectral response at p_1 on the target of the reconstructions, as d changes from $d = 2$ to $d = 3$ when $\sigma = 6$. Here it is possible to see how as d is increased with a fixed value of σ the

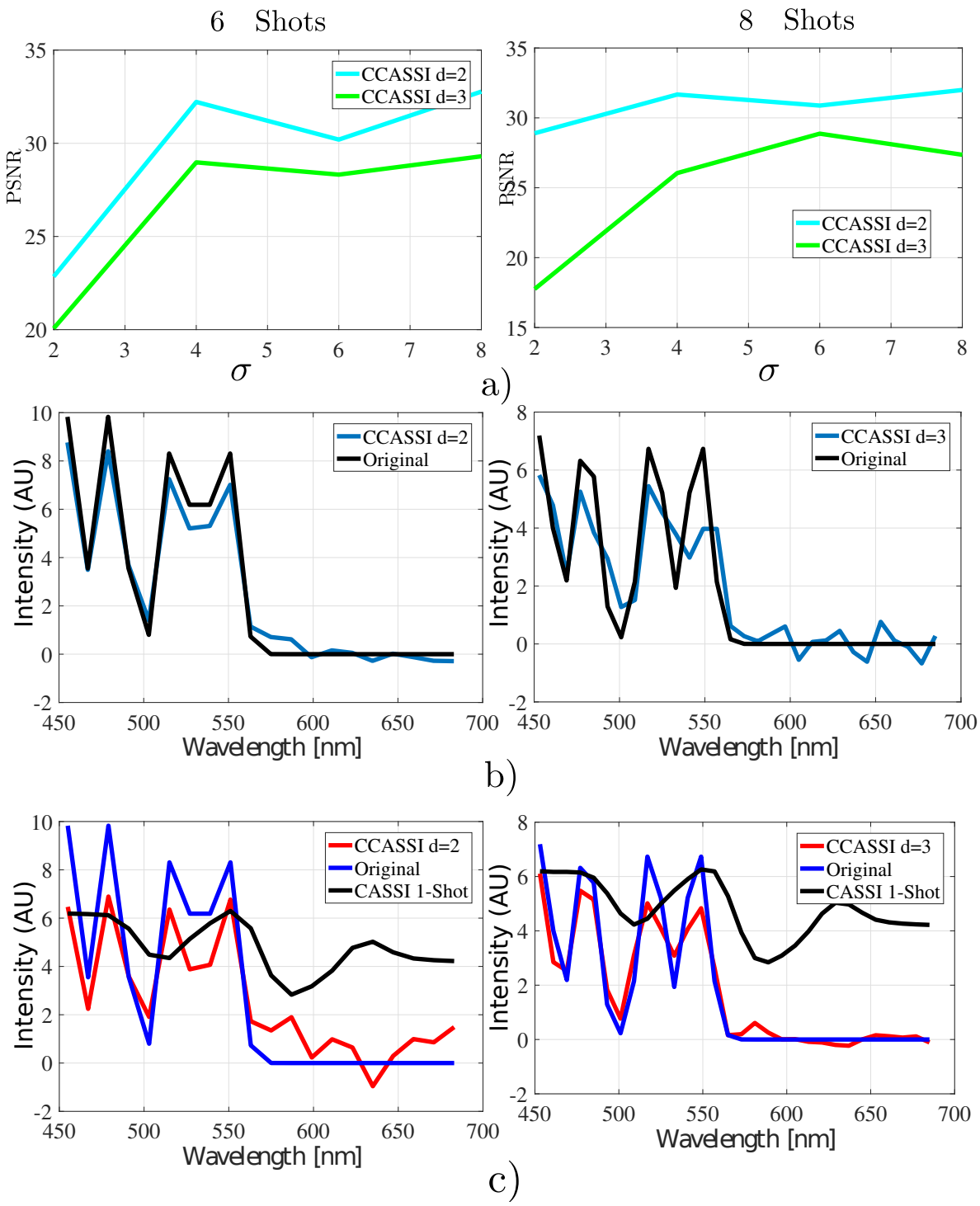


Figure 7.8: Simulation results showing the performance of the colored CASSI with super-resolution for different values of σ and d. (a) The PSNR of the reconstructions as a function of σ for $d = 2, 3$ and different number of shots. (b) The spectral response at point p_1 on the target is reconstructed (blue line) and compared with the original (black line) for different values of d, while the number of filters is $\sigma = 6$. (c) A comparison of the reconstructions obtained with the colored CASSI (red line), the traditional CASSI (black line) using the single shot modality and the real spectral response (blue line) at p_1.

quality of the reconstructions is decreased. This is consistent with the nature of the problem, because increasing d rises the ill posedness of the problem as the number of bands required is higher, given a fixed value of σ. The values of the coherence of the sensing matrix $\mu(\mathbf{A})$ are presented in Table 7.1. It can be seen how for a given number of shots K, the values of the coherence decrease as σ is increased.

In Figure 7.9 the effect of additive Gaussian noise in the measurements is considered using the model defined by equations (7.5) and (7.35). In Figure 7.9 (a) the filters used in the coded aperture are chosen as complementary filters (see Figure 7.5) that cover the entire spectral range of interest, and the noise measurements are such that the SNR takes the values $SNR = 10[db], 20[db]$. It is clear from the simulations, that in the presence of noise and with the use of complementary filters, as σ is increased the light throughput is reduced and consequently the results are affected. On the other hand in Figure 7.9(b) noisy measurements are considered but the filters used in the colored coded aperture are not complementary, and they cover the spectral range of interest as well. It is possible to see that in this second case, the reconstruction results are more robust in the presence of noise when σ is increased than in the case when complementary filters are used.

7.5.4 Estimates of the super-resolution factor d

The data presented in the Table 7.2, shows the numerical values involved in the terms $\hat{d}_1, \hat{d}_2$. In order to check the estimate of the super-resolution factor d in equation (7.33), these values were calculated for different number of filters and three different cases of partition of the basic bands. When the basic bands are broken in two parts, the super-resolution factor is achieved when 7 or more filters are used, and higher values of d are not possible with less than 8 filters.

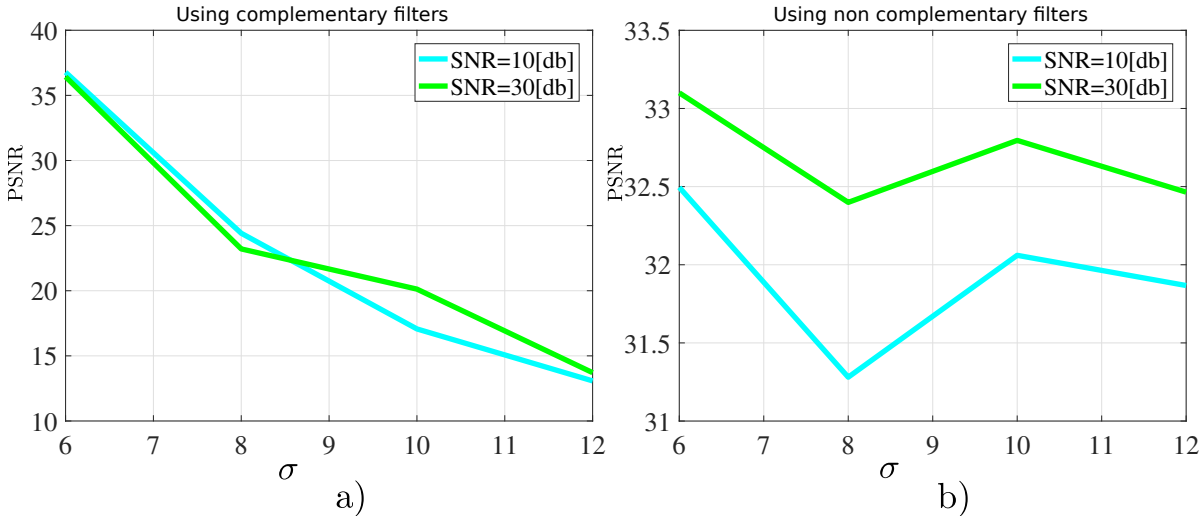


Figure 7.9: PSNR of the reconstruction results for different values of σ when the measurements are polluted with additive Gaussian noise. (a) The filters used in the colored coded aperture are selected in a complementary way (see Figure 7.5). (b) The filters used in the coded aperture are not complementary but cover the whole spectral range of interest.

$(\cdot)\backslash\sigma$	2	3	4	5	6	7	8
$\hat{d}_1$	0.54	0.83	1.27	1.40	1.73	2.03	2.65
$\hat{d}_2$	1.86	1.67	6.05	2.10	2.89	4.40	8.96
Partitions in 2 parts							
$d\backslash\sigma$	2	3	4	5	6	7	8
$\hat{d}_1$	0.55	0.74	1.05	1.33	1.60	2.06	2.31
$\hat{d}_2$	0.96	1.18	1.23	1.61	1.85	2.27	2.7
Partitions in 3 parts							
$d\backslash\sigma$	2	3	4	5	6	7	8
$\hat{d}_1$	0.50	0.73	1.06	1.22	1.69	1.74	1.96
$\hat{d}_2$	1.02	1.06	1.34	1.40	1.96	1.98	2.14
Partition in 4 parts							

Table 7.2: The value of $\hat{d}_1, \hat{d}_2$

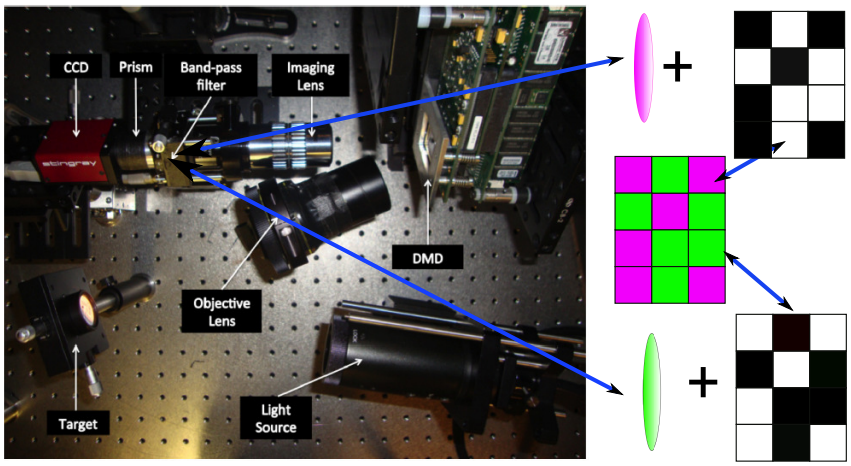


Figure 7.10: Picture of the testbed used in the implementation of the colored CASSI. In the picture it is also indicated how the colored coded aperture are implemented for one example of a color coded aperture of 2 filters (See also Fig. 7.11)

7.6 Experimental Results

Experimental results were obtained considering the super resolution model proposed. The testbed, shown in Fig. 7.10, is used to implement the colored CASSI system and to verify the simulation results. It is composed of a light source, the target, the objective lens, the Digital micro-mirror device (DMD) which plays the role of the coded aperture, imaging lenses, a bandpass filter in which the filters of the colored coded aperture are contained, the dispersive element and the CCD camera.

The target is illuminated with the source light and the reflected light on the target is filtered by the bandpass filter (25mm, VIS 400-694nm CWL Mounted Diameter Filter Kit of Edmund optics), and then redirected trough the objective lenses on the DMD. Then, the light reflected on the DMD (Texas instruments DMD) is focused into the prism (Amici prism) imaging plane that disperses the light onto the CCD camera (Stingray F-033C CCD camera), which integrates all information in a 2-dimensional array of data.

The system is characterized in order to reduce non uniform conditions and external noise. For that purpose, the light source intensity distribution and the CCD spectral

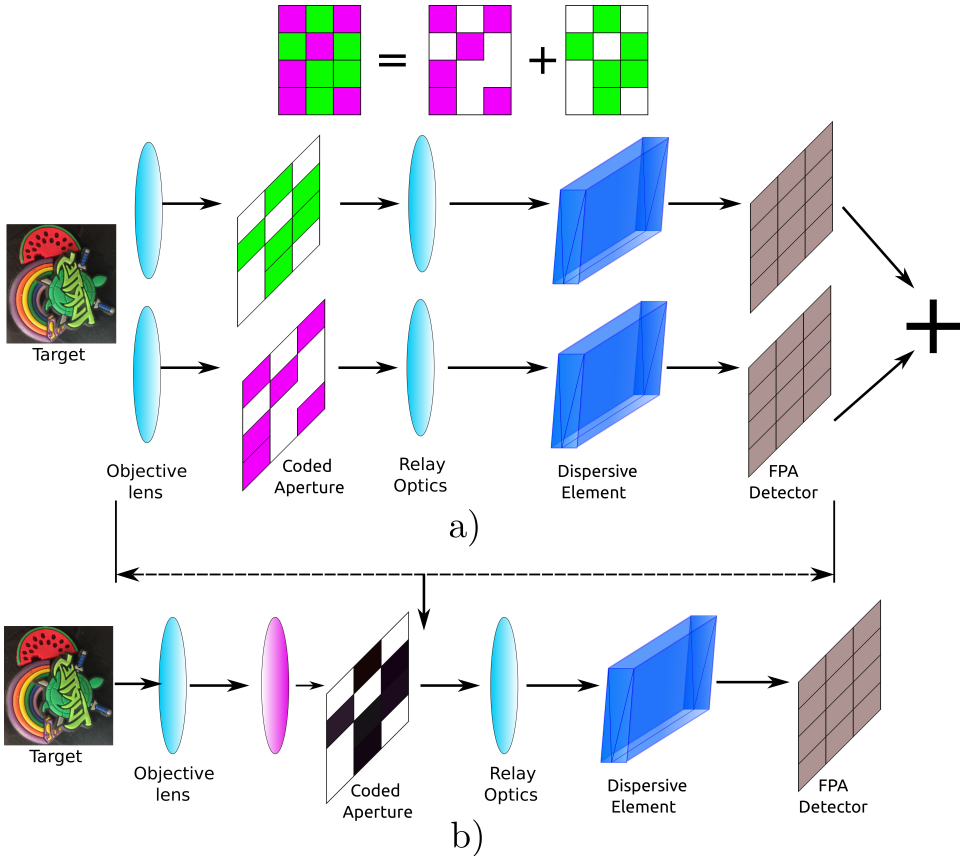


Figure 7.11: Details on the implementation of the colored coded aperture. (a) The capture of one shot with a colored coded aperture is the sum of the captures using coded apertures with one single spectral response. (b) In order to get the capture with each of the patterns with one single spectral response a filter is located as indicated in the picture

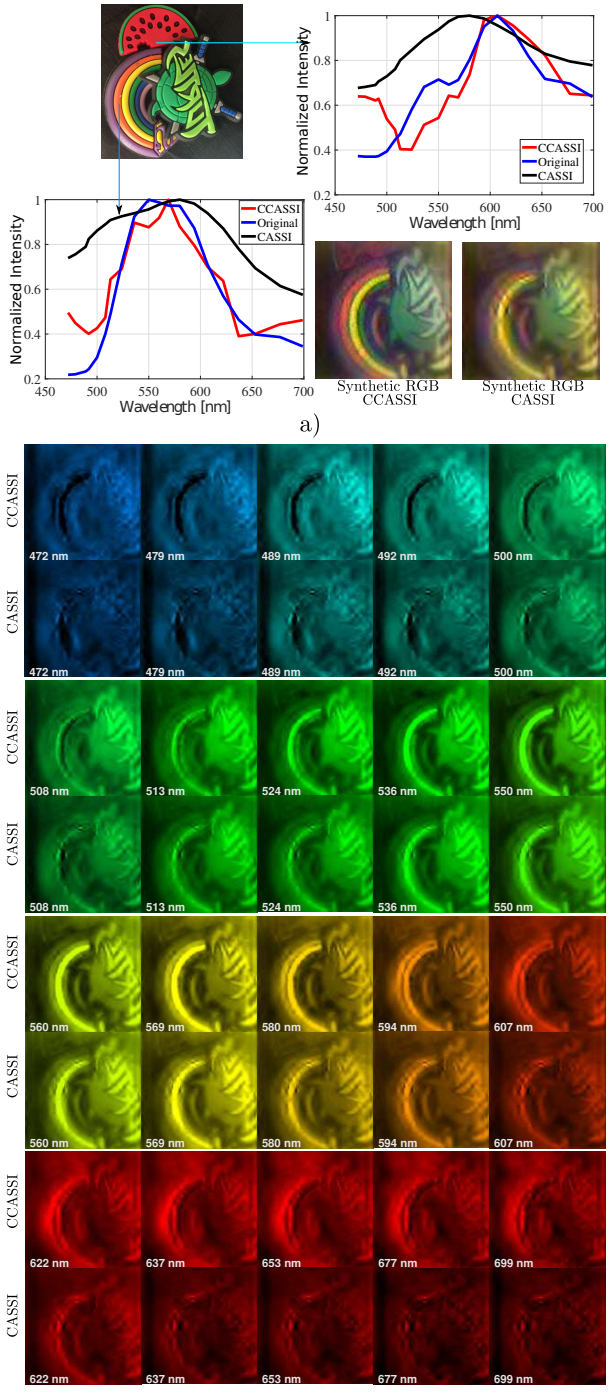


Figure 7.12: Results considering a super-resolution factor of $d = 2$ and using $\sigma = 6$ filters. (a): The spectral responses of two different points are reconstructed. The original spectrum (blue line) measured with an spectrometer is compared with the reconstructed spectrum using the colored CASSI (red line), and the reconstructed spectrum using the traditional CASSI (black line). The number of shots used with the traditional CASSI is 5 whereas for the colored CASSI is 10, such that the same value of $\mathcal{C} = 0.5$ is considered in both cases. It can be appreciated that the curve obtained with the Colored CASSI with super resolution is closer to the real spectrum than the curve obtained with the basic CASSI. (b) The reconstructed bands using the traditional CASSI and the colored CASSI are presented. The first row of each group depict the reconstructed bands using the colored CASSI (CCASSI). The second row of each group depict the reconstructed bands using the traditional CASSI.

sensitivity are characterized analyzing their spectral responses using a USB2000+VIS-NIR Ocean Optics spectrometer with a known spectral response. These non-uniform spectral response curves are taken into account to quantify the whole spectral responses in the coded aperture. The non flat spectral nature of the light and the spectral response of the camera, causes a final effect in the spectral response of each color $T_T(\lambda)$. It is represented as $T_T(\lambda) = T_{\text{light}}(\lambda)T_{\text{camera}}(\lambda)T(\lambda)$, where $T_{\text{light}}(\lambda)$ is the spectral response of the light, $T_{\text{camera}}(\lambda)$ is the spectral response of the camera and $T(\lambda)$ the spectral response of the filter. The CCD exposure time is 100 microseconds. The prism is characterized in order to take into account its non-linear response curve, and the resultant bandwidth of each spectral basic band. In order to get an estimate of the values for the weights $w_{i,j,k}$, the procedure presented in [82] is followed, in which some measurement shots are captured using monochromatic light, allowing the estimation of the effect of a single voxel impinging onto the CCD.

7.6.1 Implementation of the colored coded aperture

Figure 7.11 shows the physical implementation of one colored coded aperture. It is based on the decomposition of a measurement shot in the sum of different captures, each of them involving just one of the spectral responses of the colored coded aperture, and complementary patterns on the DMD. In Figure 7.10 it is indicated that, there is a fixed position in the testbed in which the color filters involved in the colored coded aperture are located. Additionally, a set of complementary binary patterns are associated (one pattern per filter) with the captures according to the desired filter. The mathematical description of this situation, taking into account that if $\mathbf{T}^{(k)}$ is the colored coded aperture in shot k, can be represented as $\mathbf{T}^{(k)} = \sum_{i=1}^{\sigma} T_i(\lambda)\mathbf{T}_i$, where $T_i(\lambda)$ represents the spectral response of the optical filter i in the colored coded aperture, and the $\mathbf{T}_i$ are binary patterns such that $\sum_i^{\sigma} \mathbf{T}_i = \mathbf{1}_{n\times n}$. Therefore, for each shot, σ captures are done putting the DMD with the binary pattern defined by each $\mathbf{T}_i$ and locating the $T_i(\lambda)$ in the bandpass position respectively.

7.6.2 Experimental results for $d = 2$

In order to show the experimental results of super-resolution, a factor $d = 2$ is considered. According to the characteristics of the filters, the number of filters required considering equation (7.33) is 6. For the reconstructions the GPSR algorithm [81] was used. The value of the scalar parameter for regularization was chosen in an empiric way. The wavelet transform was used as a basis for the spatial domain, and the discrete cosine transform (DCT) was used for the spectral domain. In Figure 7.12, it is possible to appreciate the reconstruction of the spectrum for some specific points in the target, considering 20 bands. This reconstruction is compared with the interpolation of a 10 bands CASSI. The expected coherence calculated for the sensing matrix using the basic CASSI is $\mu(\mathbf{A}) = 0.8366$ whereas the coherence for the colored CASSI is $\mu(\mathbf{A}_\sigma) = 0.6564$.

In Figure 7.12(a) the real target used in the experiments is presented with the spectral responses of some specific points on the target, and an RGB representation of the reconstructions as well. It is possible to appreciate how the reconstructions obtained with the colored CASSI with super-resolution (red line) match better with the real spectral response (blue line), than the reconstructions obtained by interpolation with the traditional CASSI (black line). The RGB mapping of the reconstructed multispectral images shows also that the results with the colored CASSI with super-resolution are closer to the real target, than the reconstructed scene using basic CASSI with interpolation.

Figure 7.12(b) shows the reconstructed bands in both configurations of the CASSI. As can be seen, the bands reconstructed with the colored CASSI show much more clear details than the reconstructed bands using basic CASSI with interpolation. Consider for instance the band at $508[nm]$, where the details on the letters on the target are much more clear when the colored CASSI is used that when basic CASSI is used.

7.7 Conclusions and Future work

This chapter demonstrates that color coded apertures in CASSI systems can be used to obtain higher spectral resolution than that achieved by CASSI systems using binary coded apertures. The increased resolution is related to the number of different colored filters used in the coded apertures, and on their spectral responses. An estimate of the super-resolution factor d is obtained using the coherence of the sensing matrix. This, in turn, provide concentration inequalities on the projection matrices that involve the characteristics of the set of filters used.

The presence of noise in the measurements is analyzed. It is shown that filters with complementary spectral responses, in multi-shot measurements, are more affected by noise than colored coded apertures whose filters' spectral responses are not complementary in the spectral range of interest. It is shown that two sensing matrices having the same set of filters, but with different gains, do not guarantee the same performance. Thus, the spectral shape, the central wavelengths, and their gain are all important in the design of the sensing matrix.

The distribution of the filters on the colored coded aperture was defined as uniformly distributed, in order to have a fair comparison with the traditional CASSI which uses Bernoulli random variables on the entries of the black and white coded aperture. A significant improvements can be achieved, however, when the filters are selected according to some optimal pattern design. This strategy is left for future work.

7.8 Calculation of the integration limits

Taking into account that the *rect* function is separable, it is possible to determine when the product between the rect functions involved is different from zero, in order to define the limits of the integral operators. Because

$$\text{rect}\left(\frac{y}{\Delta}-m^{'}\right)\text{rect}\left(\frac{y}{\Delta}-m\right)\neq 0 \Longleftrightarrow m^{'}=m,$$

and the product $\text{rect}\left(\frac{x}{\Delta}-n\right)\text{rect}\left(\frac{x+S(\lambda)}{\Delta}-n^{'}\right)$ is different from zero when $\Delta\left(n^{'}+\frac{1}{2}\right)-S(\lambda)>\Delta\left(n-\frac{1}{2}\right)$ and $\Delta\left(n^{'}+\frac{1}{2}\right)-S(\lambda)<\Delta\left(n+\frac{1}{2}\right)$, it follows that $n^{'}=\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor+n$.

Then, the integration limits in the x variable are

$$\left\{\Delta\left(n-\frac{1}{2}\right),\quad \Delta\left(n+\frac{1}{2}\right)+\Delta\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor-S(\lambda)\right\}.$$

On the other hand, if $\Delta\left(n^{'}-\frac{1}{2}\right)-S(\lambda)>\Delta\left(n-\frac{1}{2}\right)$ and $\Delta\left(n^{'}-\frac{1}{2}\right)-S(\lambda)<\Delta\left(n+\frac{1}{2}\right)$, it follows that $n^{'}=\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor+n+1$. Then, the integration limits in the x variable are

$$\left\{\Delta\left(n+\frac{1}{2}\right)+\Delta\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor-S(\lambda),\quad \Delta\left(n+\frac{1}{2}\right)\right\}.$$

Putting all this together, it follows that the value of the measurements at pixel (m,n) is

$$\mathbf{g}_{m,n}=$$
$$\int_{\Lambda}\int_{\Delta\left(n-\frac{1}{2}\right)}^{\Delta\left(n+\frac{1}{2}\right)+\Delta\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor-S(\lambda)}\int_{\Delta\left(m-\frac{1}{2}\right)}^{\Delta\left(m+\frac{1}{2}\right)}(TF)_{m,n+\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor}(\lambda)dydxd\lambda$$
$$+\int_{\Lambda}\int_{\Delta\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor+\Delta\left(n+\frac{1}{2}\right)-S(\lambda)}^{\Delta\left(n+\frac{1}{2}\right)}\int_{\Delta\left(m-\frac{1}{2}\right)}^{\Delta\left(m+\frac{1}{2}\right)}(TF)_{m,n+\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor+1}(\lambda)dydxd\lambda,$$

and using the mid point rule approximation for the spatial integral operators, the measurements can be written as

$$\mathbf{g}_{m,n}=$$
$$\int_{\Lambda}\Delta^2\left(\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor-\frac{S(\lambda)}{\Delta}+1\right)(TF)_{m,n+\lfloor\frac{S(\lambda)}{\Delta}\rfloor}(\lambda)d\lambda$$

$$-\int_{\Lambda}\Delta^2\left(\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor-\frac{S(\lambda)}{\Delta}\right)(TF)_{m,n+\lfloor\frac{S(\lambda)}{\Delta}\rfloor+1}(\lambda)d\lambda. \tag{7.39}$$

For the discretization in λ, the limits of the new bands in the super resolution model are used as the limit points of the intervals, and the mid-point rule is used again for the approximation of the integral operator in the λ variable as

$$\mathbf{g}_{m,n}=$$
$$\sum_{k=0}^{L-1}\Delta_{\lambda(k)}\Delta^2\left(\left\lfloor\frac{S(\hat{\lambda}_k)}{\Delta}\right\rfloor-\frac{S(\hat{\lambda}_k)}{\Delta}+1\right)(TF)_{m,n+\left\lfloor\frac{S(\hat{\lambda}_k)}{\Delta}\right\rfloor}(\hat{\lambda}_k)$$
$$-\sum_{k=0}^{L-1}\Delta_{\lambda(k)}\Delta^2\left(\left\lfloor\frac{S(\hat{\lambda}_k)}{\Delta}\right\rfloor-\frac{S(\hat{\lambda}_k)}{\Delta}\right)(TF)_{m,n+\left\lfloor\frac{S(\hat{\lambda}_k)}{\Delta}\right\rfloor+1}(\hat{\lambda}_k),$$

where $\hat{\lambda}_k = (\lambda_{k+1} + \lambda_k)/2$, $\Delta_{\lambda(k)} = \lambda_{k+1} - \lambda_k$. Taking into account the properties of the term $\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor$, which are explained in Appendix 7.9, it is possible to write equation (7.3) as

$$\mathbf{g}_{m,n} = \sum_{k=0}^{L-1} \sum_{u'=0}^{c-1}$$
$$\left[\Delta_{\lambda(k)} \Delta^2 \left(\left\lfloor \frac{S(\hat{\lambda}_k)}{\Delta} \right\rfloor - \frac{S(\hat{\lambda}_k)}{\Delta} + 1 \right) (TF)_{m,n+\left\lfloor \frac{k}{d} \right\rfloor + u'} (\hat{\lambda}_k) \right.$$
$$\left. - \Delta_{\lambda(k)} \Delta^2 \left(\left\lfloor \frac{S(\hat{\lambda}_k)}{\Delta} \right\rfloor - \frac{S(\hat{\lambda}_k)}{\Delta} \right) (TF)_{m,n+\left\lfloor \frac{k}{d} \right\rfloor + u' + 1} (\hat{\lambda}_k) \right],$$

where $c \in \mathbb{N}$. Here c represents the number of pixels of the detector affected by one voxel of the datacube model [82]. Letting

$$\boldsymbol{\alpha}_{m,n,k} = \Delta_{\lambda(k)} \Delta^2 \left(\left\lfloor \frac{S(\hat{\lambda}_k)}{\Delta} \right\rfloor - \frac{S(\hat{\lambda}_k)}{\Delta} + 1 \right),$$

and

$$\boldsymbol{\beta}_{m,n,k} = -\Delta_{\lambda(k)} \Delta^2 \left(\left\lfloor \frac{S(\hat{\lambda}_k)}{\Delta} \right\rfloor - \frac{S(\hat{\lambda}_k)}{\Delta} \right),$$

it follows that

$$\mathbf{g}_{m,n} = \sum_{k=0}^{L-1} \sum_{u=0}^{c-1} \left(\mathbf{w}_{m,n,k,u} (\boldsymbol{TF})_{m,n+\left\lfloor \frac{k}{d} \right\rfloor + u,k} \right),$$

where

$$\mathbf{w}_{m,n,k,u} = \begin{cases} \boldsymbol{\alpha}_{m,n,k} & \text{If} \quad u = 0 \\ \boldsymbol{\alpha}_{m,n,k} + \boldsymbol{\beta}_{m,n,k} & \text{If} \quad u = 1, \ldots, c-2 \\ \boldsymbol{\beta}_{m,n,k} & \text{If} \quad u = c-1 \end{cases}$$

For the sake of simplicity $c = 1$ is used for other analysis, therefore the index u would not be longer necessary.

7.9 The meaning of $\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor$

According to the Weierstrass approximation theorem [95], if $S(\lambda)$ is continuous in Λ, then $\forall \varepsilon > 0$ exists a polynomial $p(\lambda)$ such that $|p(\lambda) - S(\lambda)| < \varepsilon \quad \forall \lambda \in \Lambda$. Letting $\varepsilon > 0$ be an infinitesimal fixed value, it is possible to write $S(\lambda)$ in an approximate way as $S(\lambda) \approx \sum_{r=0}^{Q} \alpha_r \lambda^r \quad Q \in \mathbb{N}, \alpha_r \in \mathbb{R}$. This representation allows the

separation of the linear and nonlinear components in the dispersion phenomena. Now, with $q(\lambda) = \sum_{r=2}^{Q} \alpha_r \lambda^r$, it follows that

$$\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor = \left\lfloor \frac{\alpha_0}{\Delta} + \alpha_1 \frac{\lambda}{\Delta} + \frac{q(\lambda)}{\Delta} \right\rfloor.$$

Using the basic properties of the floor function it is possible to get

$$\left\lfloor \frac{\alpha_0}{\Delta} + \frac{q(\lambda)}{\Delta} \right\rfloor + \left\lfloor \alpha_1 \frac{\lambda}{\Delta} \right\rfloor \leq \left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor \leq \left\lfloor \frac{\alpha_0}{\Delta} + \frac{q(\lambda)}{\Delta} \right\rfloor + \left\lfloor \alpha_1 \frac{\lambda}{\Delta} \right\rfloor + 1.$$

Each term in this last equation has a different meaning. In the term $\left\lfloor \frac{\alpha_0}{\Delta} + \frac{q(\lambda)}{\Delta} \right\rfloor$, the coefficient α_0 includes the effect of the mismatching in the x axes of the dispersed and modulated hyperspectral image on the FPA, whereas $q(\lambda)$ represents the nonlinearities in the prism curve. The values of this term are represented by $u^{'}$.

In the classical CASSI system the boundaries of the support of each basic band Λ_k are defined by the changes of $\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor$ from one integer to another [69] [71]. Now, this concept of bands is generalized taking into account that the basic bands can be broken such that a new set of bands is obtained. Because the basic bands are conformed by d successive new bands, the number of bands L to reconstruct can be written as $L = dL^{'}$, where d is the super-resolution factor, which indicates the number of parts in which the basic bands are separated, and $L^{'}$ represents the number of basic bands, defined by the relation $\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor$, varying λ from $\inf\{\Lambda\}$ to $\sup\{\Lambda\}$ [69] [71].

In the term $\left\lfloor \alpha_1 \frac{\lambda}{\Delta} \right\rfloor$, once the value of α_1 is fixed, it changes from one integer to another for the values of lambda that are multiples integers of Δ/α_1, therefore, this term defines a delay that it is directly related with the changes from one basic band to another, but remains invariant when λ changes inside the support of the basic bands. This value is represented by $\left\lfloor \frac{k}{d} \right\rfloor$, where k is the index that represent the number of the band and d the super resolution factor.

7.10 Proof of Theorem 52

Let us consider

$$\mu_{mn}(\mathbf{A})^2 = \frac{\varrho(m,n)^2}{\varrho(m,m)\varrho(n,n)},$$

then

$$\mathbb{P}\left\{\mu_{mn}(\mathbf{A})^2 > \varepsilon\right\} = \mathbb{P}\left\{\frac{\varrho(m,n)^2}{\varrho(m,m)\varrho(n,n)} > \varepsilon\right\}$$
$$= \mathbb{P}\left\{\varrho(m,n)^2 > \varepsilon\varrho(m,m)\varrho(n,n)\right\}$$
$$= \sum_{\ell} \mathbb{P}\left\{\varrho(m,n)^2 > x_\ell\varepsilon\right\}\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}.$$

Taking into account that $\mathbb{P}\left\{\varrho(m,n)^2 > x_\ell\varepsilon\right\} \leq \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{x_\ell\varepsilon}$, it follows that

$$\sum_{\ell} \mathbb{P}\left\{\varrho(m,n)^2 > x_\ell\varepsilon\right\}\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}$$
$$\leq \sum_{\ell} \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{x_\ell\varepsilon}\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}$$
$$= \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\sum_{\ell} \frac{\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}}{x_\ell}$$
$$\leq \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\sum_{\ell} \frac{1}{\min\limits_{\ell}(x_\ell)}\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}$$
$$= \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\frac{1}{\min\limits_{\ell}(x_\ell)}.$$

Therefore

$$\mathbb{P}\left\{\mu_{mn}(\mathbf{A})^2 > \varepsilon\right\} \leq \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\frac{1}{\min\limits_{\ell}(x_\ell)}.$$

Following the same steps with $\mathbf{A}_\sigma$ it is possible to get

$$\mathbb{P}\left\{\mu_{mn}(\mathbf{A}_\sigma)^2 > \varepsilon\right\} \leq \frac{\mathbb{E}\left\{\varrho_\sigma(m,n)^2\right\}}{\varepsilon}\frac{1}{\min\limits_{\ell}(y_\ell)},$$

where y_ℓ represents the values of the random variable $\varrho_\sigma(m,m)\varrho_\sigma(n,n)$. Then, taking

$$\xi = \max\left\{\frac{1}{\min\limits_{\ell}(x_\ell)}, \frac{1}{\min\limits_{\ell}(y_\ell)}\right\}$$

the desired result is obtained.

7.11 Proof of Theorem 53

In order to simplify the notation for the proof, consider the following convention $\varphi_i \equiv \varphi_{i,i}, \quad R_i^{(m,n)} \equiv R_{i,i}^{(m,n)}$ and $\hat{\varphi}_\ell \equiv \varphi_{i,j}, \quad \hat{R}_\ell^{(m,n)} \equiv R_{i,j}^{(m,n)} \quad (i,j) \in I$; where the index ℓ indicates a numeration of the set I. The quantities $\varphi_i, \hat{\varphi}_\ell$ are related with the traditional CASSI and $\varphi_i^{(\sigma)}, \hat{\varphi}_\ell^{(\sigma)}$ with the colored CASSI. The symbols $\sim$ and $\nsim$ are to be used to represent dependency and non-dependency between two random variables, respectively. For instance, in one expression in which the variables φ_i and φ_j are involved, $i \sim j$ represents that φ_i and φ_j are dependent random variables.

Then, let us consider

$$
\begin{aligned}
&\mathbb{E}\left\{\varrho_\sigma(m,n)^2\right\} - \mathbb{E}\left\{\varrho(m,n)^2\right\} \\
&= \underbrace{\sum_{i=1}^{N^2L} \left(\mathrm{Var}(\varphi_i^{(\sigma)}) - \mathrm{Var}(\varphi_i) + \mathbb{E}\left\{\varphi_i^{(\sigma)}\right\}^2 - \mathbb{E}\{\varphi_i\}^2 \right) \left[R_i^{(m,n)}\right]^2}_{①} \\
&\quad + \underbrace{\sum_{i \neq j}^{N^2L} \left(\mathbb{E}\left\{\varphi_i^{(\sigma)} \varphi_j^{(\sigma)}\right\} - \mathbb{E}\left\{\varphi_i \varphi_j\right\} \right) R_i^{(m,n)} R_j^{(m,n)}}_{②} \\
&\quad + \underbrace{2\left(\mathbb{E}\{\varphi_i^{(\sigma)}\}\mathbb{E}\{\hat{\varphi}_\ell^{(\sigma)}\} - \mathbb{E}\{\varphi_i\}\mathbb{E}\{\hat{\varphi}_\ell\} \right) \sum_{i=1}^{N^2L} \sum_{\ell} R_i^{(m,n)} \hat{R}_\ell^{(m,n)}}_{③} \\
&\quad + \underbrace{\sum_{\ell} \left(\mathrm{Var}(\hat{\varphi}_\ell^{(\sigma)}) - \mathrm{Var}(\hat{\varphi}_\ell) + \mathbb{E}\left\{\hat{\varphi}_\ell^{(\sigma)}\right\}^2 - \mathbb{E}\{\hat{\varphi}_\ell\}^2 \right) \left[\hat{R}_\ell^{(m,n)}\right]^2}_{④} \\
&\quad + \underbrace{\sum_{\ell \neq \ell'} \left(\mathbb{E}\left\{\hat{\varphi}_\ell^{(\sigma)} \hat{\varphi}_{\ell'}^{(\sigma)}\right\} - \mathbb{E}\left\{\hat{\varphi}_\ell \hat{\varphi}_{\ell'}\right\} \right) \hat{R}_\ell^{(m,n)} \hat{R}_{\ell'}^{(m,n)}}_{⑤} .
\end{aligned}
$$

In the following, the analysis of each term of the previous equation es presented separately.

7.11.1 About the term ①

Taking into account that $\mathbb{E}\left\{\varphi_i^{(\sigma)}\right\} = \mathbb{E}\left\{\varphi_i\right\} = \mathcal{V}$, $\mathrm{Var}\left(\varphi_i^{(\sigma)}\right) = \mathcal{V}\left(1 - \frac{1}{\sigma}\right)$ and $\mathrm{Var}\left(\varphi_i\right) = \frac{\mathcal{V}}{2}$ the original expression ① can be represented as

$$= \mathcal{V}\left(\frac{1}{2} - \frac{1}{\sigma}\right) \sum_{i=1}^{N^2 L} \left[R_i^{(m,n)}\right]^2.$$

7.11.2 About the term ②

For the non independent terms $i \neq j, i \sim j$ it follows that

$$\mathbb{E}\left\{\varphi_i \varphi_j\right\} = \mathbb{E}\left\{\left(\sum_{i=1}^{K} T_{X_i}(\lambda_k)^2\right)\left(\sum_{i=1}^{K} T_{X_i}(\lambda_{k'})^2\right)\right\}$$

$$= \sum_{i=1}^{K} \mathbb{E}\left\{T_{X_i}(\lambda_k)^2 T_{X_i}(\lambda_{k'})^2\right\} + \sum_{i \neq j}^{K} \mathbb{E}\left\{T_{X_i}(\lambda_k)^2\right\} \mathbb{E}\left\{T_{X_j}(\lambda_{k'})^2\right\},$$

where $k \neq k^{'} \in \{1, \ldots L^{'}\}$, and $T_{X_i}(\lambda_k)$ represents the spectral response of the filter X_i in the band λ_k. The values of this last expression according to the characteristics of the coded aperture are: $\frac{K}{2} + \sum_{i \neq j}^{K} \frac{1}{4} = \mathcal{V}\left(\mathcal{V} + \frac{1}{2}\right)$ when the block unblock coded aperture with transmittance $t = 1/2$ is used, and $\sum_{i \neq j}^{K} \frac{1}{\sigma^2} = \mathcal{V}\left(\mathcal{V} - \frac{1}{\sigma}\right)$ when the colored coded aperture is used. Now, calculating the difference between the colored case and the binary case, and taking into account that $\mathbb{E}\left\{\varphi_i^{(\sigma)}\right\} = \mathbb{E}\left\{\varphi_i\right\} = \mathcal{V}$, it follows that

$$\sum_{i \neq j}^{N^2 L} \left(\mathbb{E}\left\{\varphi_i^{(\sigma)} \varphi_j^{(\sigma)}\right\} - \mathbb{E}\left\{\varphi_i \varphi_j\right\}\right) R_i^{(m,n)} R_j^{(m,n)}$$

$$= -\mathcal{V}\left(\frac{1}{\sigma} + \frac{1}{2}\right) \sum_{\substack{i \neq j \\ i \sim j}} R_i^{(m,n)} R_j^{(m,n)}.$$

7.11.3 About the term ③

Notice that

$$2 \sum_{i=1}^{N^2 L} \sum_{\ell} \mathbb{E}\left\{\varphi_i \hat{\varphi}_\ell\right\} R_i^{(m,n)} \hat{R}_\ell^{(m,n)} = 2 \sum_{\substack{i,\ell \\ i \sim \ell}} \mathbb{E}\left\{\varphi_i \hat{\varphi}_\ell\right\} R_i^{(m,n)} \hat{R}_\ell^{(m,n)}$$

$$+2\sum_{\substack{i,\ell\\ i\nsim\ell}}\mathbb{E}\left\{\varphi_i\right\}\mathbb{E}\left\{\hat{\varphi}_\ell\right\}R_i^{(m,n)}\hat{R}_\ell^{(m,n)}.$$

Now, analyzing the term $\mathbb{E}\left\{\varphi_i\hat{\varphi}_\ell\right\}$ when $i\sim\ell$ and taking into account the fact that $\varphi_i\equiv\varphi_{i,i}=\langle\mathbf{h}_i,\mathbf{h}_i\rangle=\sum_{u=1}^{K}T_{X_u}(\lambda_k)^2$, and $\hat{\varphi}_\ell\equiv\varphi_{i,j}=\langle\mathbf{h}_i,\mathbf{h}_j\rangle=\sum_{u=1}^{K}T_{X_u}(\lambda_k)T_{Y_u}(\lambda_{k'})$. It follows that the product $\varphi_i\hat{\varphi}_\ell$ can be written as

$$\varphi_i\hat{\varphi}_\ell=\sum_{u=1}^{K}T_{X_u}(\lambda_k)^3T_{Y_u}(\lambda_{k'})+\sum_{u\neq v}^{K}T_{X_u}(\lambda_k)^2T_{X_v}(\lambda_k)T_{Y_v}(\lambda_{k'}).$$

Using the expectation operator on this expression, it is possible to get

$$\mathbb{E}\left\{\varphi_i\hat{\varphi}_\ell\right\}=\sum_{u=1}^{K}\mathbb{E}\left\{T_{X_u}(\lambda_k)^3\right\}\mathbb{E}\left\{T_{Y_u}(\lambda_{k'})\right\}$$

$$+\sum_{u\neq v}^{K}\mathbb{E}\left\{T_{X_u}(\lambda_k)^2\right\}\mathbb{E}\left\{T_{X_v}(\lambda_k)\right\}\mathbb{E}\left\{T_{Y_u}(\lambda_{k'})\right\}.$$

For the binary case $\mathbb{E}\left\{\varphi_i\hat{\varphi}_\ell\right\}=\frac{1}{2}\mathcal{V}^2+\frac{1}{4}\mathcal{V}$ and for the colored case $\mathbb{E}\left\{\varphi_i\hat{\varphi}_\ell\right\}=\frac{1}{\sigma}\mathcal{V}^2+\left(\frac{1}{\sigma}-\frac{1}{\sigma^2}\right)\mathcal{V}$.

Making the difference between binary and colored case, the expression ③ is finally represented as

$$=2\mathcal{V}\left(\frac{1}{\sigma}-\frac{1}{\sigma^2}-\frac{1}{4}\right)\sum_{\substack{i,\ell\\ i\sim\ell}}R_i^{(m,n)}\hat{R}_\ell^{(m,n)}.$$

7.11.4 About the term ④

Taking into account that

$$\sum_{\ell}\mathbb{E}\left\{[\hat{\varphi}_\ell]^2\right\}\left[\hat{R}_\ell^{(m,n)}\right]^2=\sum_{\ell}\left(\mathrm{Var}(\hat{\varphi}_\ell)+\mathbb{E}\left\{\hat{\varphi}_\ell\right\}^2\right)\left[\hat{R}_\ell^{(m,n)}\right]^2,$$

it is possible to make the difference between the colored case and the binary case, getting a representation of expression ④ as

$$=\left(\mathcal{V}\left(\frac{1}{\sigma}-\frac{1}{\sigma^3}-\frac{3}{8}\right)+\mathcal{V}^2\left(\frac{1}{\sigma^2}-\frac{1}{4}\right)\right)\sum_{\ell}\left[\hat{R}_\ell^{(m,n)}\right]^2.$$

7.11.5 About the term ⑤

Notice that

$$\mathbb{E}\left\{\hat{\varphi}_\ell \hat{\varphi}_{\ell'}\right\} = \begin{cases} \mathbb{E}\left\{\hat{\varphi}_\ell\right\}\mathbb{E}\left\{\hat{\varphi}_{\ell'}\right\} & \ell \neq \ell', \ell \nsim \ell' \\ \mathbb{E}\left\{\hat{\varphi}_\ell \hat{\varphi}_{\ell'}\right\} & \ell \neq \ell', \ell \sim \ell'. \end{cases}$$

Because of the symmetry of the problem, for $\ell \neq \ell', \ell \sim \ell'$ it follows that

$$\mathbb{E}\left\{\hat{\varphi}_\ell \hat{\varphi}_{\ell'}\right\} =$$

$$\mathbb{E}\left\{\left(\sum_{i=1}^{K} T_{X_i}(\lambda_k) T_{Y_i}(\lambda_{k'})\right)\left(\sum_{i=1}^{K} T_{X_i}(\lambda_k) T_{Z_i}(\lambda_{k''})\right)\right\}$$

$$= \sum_{i=1}^{K} \mathbb{E}\left\{T_{X_i}(\lambda_k)^2\right\}\mathbb{E}\left\{T_{Y_i}(\lambda_{k'})\right\}\mathbb{E}\left\{T_{Z_i}(\lambda_{k''})\right\}$$

$$+ \sum_{i\neq j}^{K} \mathbb{E}\left\{T_{X_i}(\lambda_k)\right\}\mathbb{E}\left\{T_{X_j}(\lambda_k)\right\}\mathbb{E}\left\{T_{Y_i}(\lambda_{k'})\right\}\mathbb{E}\left\{T_{Z_i}(\lambda_{k''})\right\},$$

which is equivalent to $\frac{\mathcal{V}}{4}\left(\mathcal{V}+\frac{1}{2}\right)$ for the binary case, and $\frac{\mathcal{V}}{\sigma^2}\left(\mathcal{V}+1-\frac{1}{\sigma}\right)$ for the colored case. Therefore, it follows that

$$\sum_{\ell\neq\ell'}\left(\mathbb{E}\left\{\hat{\varphi}_\ell^{(\sigma)}\hat{\varphi}_{\ell'}^{(\sigma)}\right\} - \mathbb{E}\left\{\hat{\varphi}_\ell\hat{\varphi}_{\ell'}\right\}\right)\hat{R}_\ell^{(m,n)}\hat{R}_{\ell'}^{(m,n)} =$$

$$2\sum_{\ell>\ell'}\mathcal{V}^2\left(\frac{1}{\sigma^2}-\frac{1}{4}\right)\hat{R}_\ell^{(m,n)}\hat{R}_{\ell'}^{(m,n)}$$

$$+2\sum_{\substack{\ell>\ell' \\ \ell\sim\ell'}}\mathcal{V}\left(\frac{1}{\sigma^2}-\frac{1}{\sigma^3}-\frac{1}{8}\right)\hat{R}_\ell^{(m,n)}\hat{R}_{\ell'}^{(m,n)}.$$

Putting all this together, it follows that

$$\mathbb{E}\left\{\varrho_\sigma(m,n)^2\right\} - \mathbb{E}\left\{\varrho(m,n)^2\right\} = a_2\mathcal{V}^2 + a_1\mathcal{V},$$

where

$$a_2 = \left(\frac{1}{\sigma^2}-\frac{1}{4}\right)\left(\sum_\ell\left[\hat{R}_\ell^{(m,n)}\right]^2 + \sum_{\ell\neq\ell'}\hat{R}_\ell^{(m,n)}\hat{R}_{\ell'}^{(m,n)}\right),$$

$$a_1 = \left(\frac{1}{2} - \frac{1}{\sigma}\right) \sum_{i=1}^{N^2 L} \left[R_i^{(m,n)}\right]^2 - \left(\frac{1}{2} + \frac{1}{\sigma}\right) \sum_{\substack{i \neq j \\ i \sim j}} R_i^{(m,n)} R_j^{(m,n)}$$
$$+2\left(\frac{1}{\sigma} - \frac{1}{\sigma^2} - \frac{1}{4}\right) \sum_{\substack{i,\ell \\ i \nsim \ell}} R_i^{(m,n)} \hat{R}_\ell^{(m,n)}$$
$$+\left(\frac{1}{\sigma} - \frac{1}{\sigma^3} - \frac{3}{8}\right) \sum_{\ell} \left[\hat{R}_\ell^{(m,n)}\right]^2$$
$$+\left(\frac{1}{\sigma^2} - \frac{1}{\sigma^3} - \frac{1}{8}\right) \sum_{\substack{\ell \neq \ell' \\ \ell \sim \ell'}} \hat{R}_\ell^{(m,n)} \hat{R}_{\ell'}^{(m,n)}.$$

Then, the behavior of the coefficients a_2 and a_1 completely determine if the inequality (7.32) is satisfied or not. In Fig. 7.13 it is possible to appreciate the most representative values of the whole set of values for the coefficients a_2 and $-a_1/a_2$ considering a hyperspectral image of dimensions $64 \times 64 \times 4$, and the same basis functions used in the simulations and real reconstructions for different values for σ. It is clear that the value of a_2 is always non positive, which implies that the polynomial $a_2\mathcal{V}^2 + a_1\mathcal{V}$ is a parabola that opens downwards. One of the roots of this polynomial is on $\mathcal{V} = 0$ and the other one is on $-a_1/a_2$ which means that $a_2\mathcal{V}^2 + a_1\mathcal{V} \leq 0 \quad \forall \mathcal{V} \geq 0$ because $-a_1/a_2$ is always non positive .

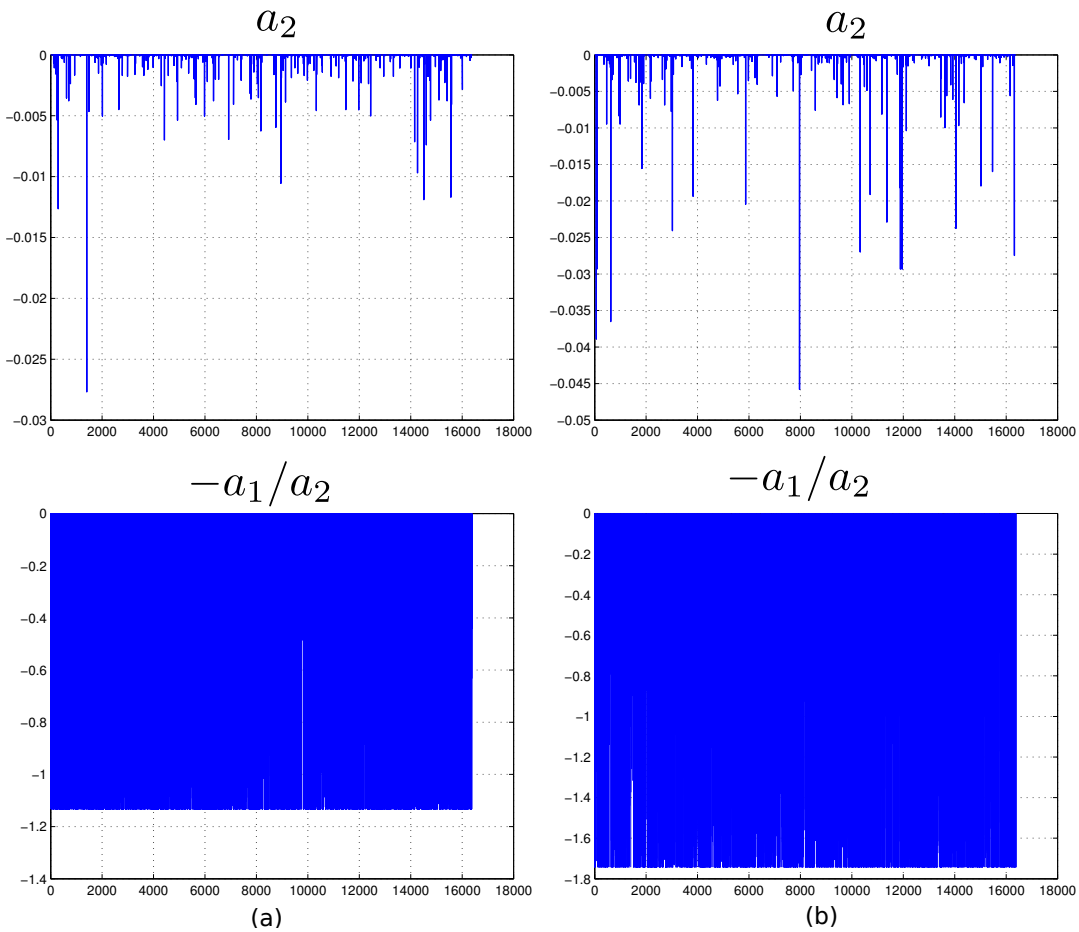


Figure 7.13: Samples of the behavior of the coefficients a_2 and a_1. (a): The behavior of the coefficient a_2 and the quotient $-a_1/a_2$ for $\sigma = 3$. (b): The behavior of the coefficient a_2 and the quotient $-a_1/a_2$ for $\sigma = 5$

Chapter 8

CONCLUSIONS AND FUTURE WORK

In this dissertation, the study of simple and universal principles for the generation of good sampling patterns was considered for signals on graphs and for the analysis of projected measurements in compressed sensing applications. In particular, the concept of blue-noise sampling was extended to graphs and its connection to previous theoretical results was explored. In graphs with a locally homogeneous isoperimetric dimension, blue-noise sampling patterns are characterized by a low redness that is associated to good values of the constants that measure the quality of a sampling set. The analysis performed was validated by a set of numerical experiments that showed the effectiveness of blue noise for the sampling and reconstruction of bandlimited signals on graphs. Additionally, the basis for the development of low complexity algorithms was given, exploiting simple principles that can be derived from vertex-domain characteristics of a blue noise sampling pattern.

As suggested by professor Austin J. Brockmeier, a relationship between blue-noise sampling on graphs and clustering techniques like K-means is worth exploring. In particular, a possible connection between clusters and the open balls used in the definition of ideal blue-noise sampling can be established, opening the door for new applications and theoretical insights of graph sampling.

Performing an analysis of bandlimited signals on cographs, it was shown that by taking advantage of the structure of the cotree, it is possible to calculate with a low complexity the uniqueness sets for bandlimited signals. Additionally, a closed form solution was obtained for a subclass of cographs called threshold graphs. These findings open the door to promising applications where datasets can be modeled by cographs

and threshold graphs, besides the applications in evolutionary biology and scheduling problems already existing in the literature.

In the context of compressed sensing applications, optimal sampling patterns on colored coded apertures are obtained for CASSI systems and compressed X-ray tomosynthesis architectures. In particular, a family of optimal codes is determined in closed form via a rigorous analysis of the coherence of the sensing matrix. This contribution provided a substantial improvement with respect to all the approaches existing in the literature, where iterative procedures were required for the generation of optimal patterns. Additionally, analyzing the sampling patterns on a not necessarily ideal colored coded aperture, a rigorous estimate of the spectral resolution achievable in colored CASSI systems is provided. These conditions are derived from concentration inequalities that are calculated by exploiting the structure of the sensing matrix.

BIBLIOGRAPHY

[1] I. Pesenson. Variational splines and paley–wiener spaces oncombinatorial graphs. *Constructive Approximation*, 29(1):1–21, Feb 2009.

[2] H. Arguello and G. R. Arce. Colored coded aperture design by concentration of measure in compressive spectral imaging. *IEEE Transactions on Image Processing*, 23(4):1896–1908, April 2014.

[3] Angela P. Cuadros, Christopher Peitsch, Henry Arguello, and Gonzalo R. Arce. Coded aperture optimization for compressive x-ray tomosynthesis. *Opt. Express*, 23(25):32788–32802, Dec 2015.

[4] Y.C. Eldar. *Sampling Theory: Beyond Bandlimited Systems*. Cambridge University Press, 2015.

[5] I. Z. Pesenson. A sampling theorem on homogeneous manifolds. *Trans. Amer. Math.*, 352:4257–4269, April 2000.

[6] I. Pesenson. Poincaré-type inequalities and reconstruction of paley-wiener functions on manifolds. *The Journal of Geometric Analysis*, 14(1):101–121, Mar 2004.

[7] H. Feichtinger and I. Pesenson. A reconstruction method for band-limited signals on the hyperbolic plane. 4, 05 2011.

[8] F. S. Roberts. *Graph Theory and Its Applications to Problems of Society*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1978.

[9] A. A. Keller. Graph theory and economic models: from small to large size applications. *Electronic Notes in Discrete Mathematics*, 28(Supplement C):469 – 476, 2007. 6th Czech-Slovak International Symposium on Combinatorics, Graph Theory, Algorithms and Applications.

[10] A. Fornito, A. Zalesky, and E. Bullmore. *Fundamentals of Brain Network Analysis*. Elsevier Science, 2016.

[11] A.L. Barabási and M. Pósfai. *Network Science*. Cambridge University Press, 2016.

[12] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30(3):83–98, May 2013.

[13] D. I Shuman, B. Ricaud, and P. Vandergheynst. Vertex-frequency analysis on graphs. *Applied and Computational Harmonic Analysis*, 40(2):260 – 291, 2016.

[14] D. I Shuman, M. J Faraji, and P. Vandergheynst. A multiscale pyramid transform for graph signals. *IEEE Transactions on Signal Processing*, 64(8):2119–2134, 2016.

[15] I. Z. Pesenson. Sampling solutions of schrodinger equations on combinatorial graphs. In *2015 International Conference on Sampling Theory and Applications (SampTA)*, pages 82–85, May 2015.

[16] I. Z. Pesenson and Meyer Z. Pesenson. Sampling filtering and sparse approximations on combinatorial graphs. *Journal of Fourier Analysis and Applications*, 16(6):921–942, Dec 2010.

[17] A. Ortega, P. Frossard, J. Kovaevi, J. M. F. Moura, and P. Vandergheynst. Graph signal processing: Overview, challenges, and applications. *Proceedings of the IEEE*, 106(5):808–828, May 2018.

[18] S. K. Narang, A. Gadde, and A. Ortega. Signal processing techniques for interpolation in graph structured data. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5445–5449, May 2013.

[19] S. Chen, R. Varma, A. Singh, and J. Kovaevi. Signal recovery on graphs: Fundamental limits of sampling strategies. *IEEE Transactions on Signal and Information Processing over Networks*, 2(4):539–554, Dec 2016.

[20] A. Anis, A. Gadde, and A. Ortega. Towards a sampling theorem for signals on arbitrary graphs. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3864–3868, May 2014.

[21] S. Chen, R. Varma, A. Sandryhaila, and J. Kovaevi. Discrete signal processing on graphs: Sampling theory. *IEEE Transactions on Signal Processing*, 63(24):6510–6523, Dec 2015.

[22] S. Barbarossa P. Di Lorenzo and P. Banelli. Sampling and recovery of graph signals. In P. Djuric and C.Richard, editors, *Cooperative and Graph Signal Processing*. Elsevier, 2018.

[23] R. A. Ulichney. Dithering with blue noise. *Proceedings of the IEEE*, 76(1):56–79, 1988.

[24] D. L. Lau, R. Ulichney, and G. R. Arce. Blue and green noise halftoning models. *IEEE Signal Processing Magazine*, 20(4):28–38, July 2003.

[25] Robert A. Ulichney. Void-and-cluster method for dither array generation, 1993.

[26] A. Sandryhaila and J. M. F. Moura. Discrete signal processing on graphs. *IEEE Transactions on Signal Processing*, 61(7):1644–1656, April 2013.

[27] H. Fuhr and I. Z. Pesenson. Poincar and plancherel polya inequalities in harmonic analysis on weighted combinatorial graphs. *SIAM Journal on Discrete Mathematics*, 27(4):2007–2028, 2013.

[28] G. Puy, N. Tremblay, R. Gribonval, and P. Vandergheynst. Random sampling of bandlimited signals on graphs. *Applied and Computational Harmonic Analysis*, 44(2):446 – 475, 2018.

[29] D. I. Shuman, M. J. Faraji, and P. Vandergheynst. A multiscale pyramid transform for graph signals. *IEEE Transactions on Signal Processing*, 64(8):2119–2134, April 2016.

[30] T. Biyikoglu, J. Leydold, and P. F. Stadler. *Laplacian Eigenvectors of Graphs: Perron-Frobenius and Faber-Krahn Type Theorems*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2007.

[31] A. Anis, A. Gadde, and A. Ortega. Efficient sampling set selection for bandlimited graph signals using graph spectral proxies. *IEEE Transactions on Signal Processing*, 64(14):3775–3789, July 2016.

[32] N. Tremblay, P. O. Amblard, and S. Barthelmé. Graph sampling with determinantal processes. *CoRR*, abs/1703.01594, 2017.

[33] I Z. Pesenson. Average sampling and average splines on combinatorial graphs. *arXiv e-prints*, page arXiv:1901.08726, Jan 2019.

[34] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo. Signals on graphs: Uncertainty principle and sampling. *IEEE Transactions on Signal Processing*, 64(18):4845–4860, Sep. 2016.

[35] L. F. O. Chamon and A. Ribeiro. Greedy sampling of graph signals. *IEEE Transactions on Signal Processing*, 66(1):34–47, Jan 2018.

[36] L. Avena and A. Gaudillire. On some random forests with determinantal roots. 10 2013.

[37] X. Wang, P. Liu, and Y. Gu. Local-set-based graph signal reconstruction. *IEEE Transactions on Signal Processing*, 63(9):2432–2444, May 2015.

[38] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro. Sampling of graph signals with successive local aggregations. *IEEE Transactions on Signal Processing*, 64(7):1832–1843, April 2016.

[39] D. Lau and G. R. Arce. *Modern digital halftoning*. CRC Press, 2 edition, 2008.

[40] V. de Silva and R. Ghrist. Coordinate-free coverage in sensor networks with controlled boundaries via homology. *The International Journal of Robotics Research*, 25(12):1205–1222, 2006.

[41] V. de Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Algebr. Geom. Topol.*, 7(1):339–358, 2007.

[42] V. De Silva and R. Ghrist. Homological sensor networks. *Notices Amer. Math. Soc*, pages 10–17, 2007.

[43] F. Chung. Discrete isoperimetric inequalities. In *Surveys in Differential Geometry IX, International Press, 5382*, pages 53–82. Press, 2004.

[44] F.R.K. Chung. *Spectral Graph Theory*. Number no. 92 in CBMS Regional Conference Series. American Mathematical Society, 1997.

[45] H. Li, Li-Yi Wei, P. V. Sander, and Chi-Wing Fu. Anisotropic blue noise sampling. *ACM Trans. Graph.*, 29(6):167:1–167:12, December 2010.

[46] Y. Xu, R. Hu, C. Gotsman, and L. Liu. Blue noise sampling of surfaces. *Computers and Graphics*, 36(4):232 – 240, 2012. Applications of Geometry Processing.

[47] S. Zhang, J. Guo, H. Zhang, X. Jia, D. Yan, J. Yong, and P. Wonka. Capacity constrained blue-noise sampling on surfaces. *Computers and Graphics*, 55:44 – 54, 2016.

[48] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.*, 9:235–284, June 2008.

[49] P. F. Stadler. Landscapes and their correlation functions. *Journal of Mathematical Chemistry*, 20(1):1–45, Mar 1996.

[50] N. Perraudin, J. Paratte, D. I. Shuman, V. Kalofolias, P. Vandergheynst, and D. K. Hammond. GSPBOX: A toolbox for signal processing on graphs. *CoRR*, abs/1408.5781, 2014.

[51] D. L. Lau, G. R. Arce, and N. C. Gallagher. Digital color halftoning with generalized error diffusion and multichannel green-noise masks. *IEEE Transactions on Image Processing*, 9(5):923–935, May 2000.

[52] Daniel L. Lau, Gonzalo R. Arce, and Neal C. Gallagher. Digital halftoning by means of green-noise masks. *J. Opt. Soc. Am. A*, 16(7):1575–1586, Jul 1999.

[53] J.M. Steele and Mathematical Association of America. *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. MAA problem books series. Cambridge University Press, 2004.

[54] Isaac Pesenson. Sampling in paley-wiener spaces on combinatorial graphs. *Transactions of The American Mathematical Society - TRANS AMER MATH SOC*, 361, 11 2011.

[55] P. Djuric and C. Richard. *Cooperative and Graph Signal Processing: Principles and Applications*. Elsevier Science, 2018.

[56] A. Parada-Mayorga, D.L. Lau, J. Giraldo, and G. . Arce. Blue-Noise Sampling on Graphs. *arXiv:1811.12542*, December 2018.

[57] Ravindra B Bapat. *Graphs and matrices*, volume 27. Springer, 2010.

[58] Derek G Corneil, Helmut Lerchs, and L Stewart Burlingham. Complement reducible graphs. *Discrete Applied Mathematics*, 3(3):163–174, 1981.

[59] F. Kühnl and M. Lechner. orthoDeprime: A tool for heuristic cograph editing on estimated orthology graphs. 2014.

[60] M. Hellmuth, M. Hernandez-Rosales, K.T. Huber, V. Moulton, P. Stadler, and N. Wieseke. Orthology relations, symbolic ultrametrics, and cographs. *Journal of mathematical biology*, 66:399–420, 03 2013.

[61] J. Songwei, G. Lin, G. Yong, J. Nastos, Y. Wang, X. Zhang, and H. Wang. Defining and identifying cograph communities in complex networks. *New Journal of Physics*, 17, 01 2015.

[62] M. Hellmuth, A. Fritz, N. Wieseke, and P. Stadler. Techniques for the cograph editing problem: Module merge is equivalent to editing P4's, 09 2015.

[63] A.E. Brouwer and W.H. Haemers. *Spectra of Graphs*. Universitext. Springer New York, 2011.

[64] Aamir Anis, Akshay Gadde, and Antonio Ortega. Efficient sampling set selection for bandlimited graph signals using graph spectral proxies. *IEEE Transactions on Signal Processing*, 64(14):3775–3789, 2016.

[65] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega. Eigendecomposition-free sampling set selection for graph signals. *IEEE Transactions on Signal Processing*, 67(10):2679–2692, May 2019.

[66] A. Parada-Mayorga, D. L. Lau, J. H. Giraldo, and G. R. Arce. Blue-noise sampling on graphs. *IEEE Transactions on Signal and Information Processing over Networks*, pages 1–1, 2019.

[67] D. Deka, S. Backhaus, and M. Chertkov. Learning topology of the power distribution grid with and without missing data. In *2016 European Control Conference (ECC)*, pages 313–320, June 2016.

[68] Russell Merris. Laplacian graph eigenvectors. *Linear algebra and its applications*, 278(1-3):221–236, 1998.

[69] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Process. Mag.*, 31(1):105–115, 2014.

[70] D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady. Multiframe image estimation for coded aperture snapshot spectral imagers. *Appl. Opt.*, 49(36):6824–6833, Dec 2010.

[71] A. Wagadarikar, R. John, R. Willett, and D. Brady. Single disperser design for coded aperture snapshot spectral imaging. *Appl. Opt.*, 47(10):B44–B51, Apr 2008.

[72] H. Rueda, H. Arguello, and G. R. Arce. Experimental demonstration of a colored coded aperture-based compressive spectral imaging system. In *Classical Optics 2014*, page CTu2C.6. Optical Society of America, 2014.

[73] J.D. Barrie, K.A. Aitchison, G.S. Rossano, and M.H. Abraham. Patterning of multilayer dielectric optical coatings for multispectral ccds. *Thin Solid Films*, 270(12):6 – 9, 1995. 22nd International Conference on Metallurgical Coatings and Thin Films.

[74] A. Parada-Mayorga and G. Arce. Spectral super-resolution in colored coded aperture spectral imaging. *IEEE Transactions on Computational Imaging*, PP(99):1–1, 2016.

[75] Y. C. Eldar G. Kutyniok M. A Davenport, M. F. Duarte. Introduction to compressed sensing. In Gitta Kutyniok Yonina C. Eldar, editor, *Compressed Sensing*. Cambridge University Press, 2012.

[76] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Birkhäuser Basel, 2013.

[77] M. Elad. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer Publishing Company, Incorporated, 1st edition, 2010.

[78] J. M. Duarte-Carvajalino and G. Sapiro. Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization. *Trans. Img. Proc.*, 18(7):1395–1408, July 2009.

[79] H. Bai, G. Li, S. Li, Q. Li, Q. Jiang, and L. Chang. Alternating optimization of sensing matrix and sparsifying dictionary for compressed sensing. *IEEE Transactions on Signal Processing*, 63(6):1581–1594, March 2015.

[80] M. Elad. Optimized projections for compressed sensing. *IEEE Transactions on Signal Processing*, 55(12):5695–5702, 2007.

[81] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. Technical report, IEEE Journal of Selected Topics in Signal Processing, 2007.

[82] H. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce. Higher-order computational model for coded aperture spectral imaging. *Appl. Opt.*, 52(10):D12–D21, Apr 2013.

[83] J. T. Dobbins III and D. J. Godfrey. Digital x-ray tomosynthesis: current state of the art and clinical potential. *Physics in Medicine and Biology*, 48(19):p.R65, 2003.

[84] R. Marcus R. Smith-Bindman, J. Lipson and et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. *Archives of Internal Medicine*, 169(22):2078–2086, 2009.

[85] I. Reiser and S. Glick. *Tomosynthesis Imaging*. CRC press, 2014.

[86] F. Natterer. *The Mathematics of Computerized Tomography*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1986.

[87] V. Kolehmainen M. Lassas K. Hmlinen, A. Kallonen. Sparse tomography. *Computational Methods in Science and Engineering, SIAM*, 35:B644–B665, February 2013.

[88] Kerkil Choi and David J. Brady. Coded aperture computed tomography. *Proc. SPIE*, 7468:74680B–74680B–10, 2009.

[89] A. Parada-Mayorga and G. Arce. Colored coded aperture design in compressive spectral imaging via minimum coherence. *IEEE Transactions on Computational Imaging-Special Issue in Earth sciences (Submitted)*, 2016.

[90] Wei Xu, Fang Xu, Mel Jones, Bettina Keszthelyi, John Sedat, David Agard, and Klaus Mueller. High-performance iterative electron tomography reconstruction with long-object compensation using graphics processing units (gpus). *Journal of Structural Biology*, 171(2):142 – 153, 2010.

[91] J. Oliver, Woong-Bi Lee, and Heung-No Lee. Filters with random transmittance for improving resolution in filter-array-based spectrometers. *Opt. Express*, 21(4):3969–3989, Feb 2013.

[92] T. Tony Cai, Guangwu Xu, and Jun Zhang. On recovery of sparse signals via l1 minimization. *IEEE Transactions on Information Theory*, 55(7):3388–3397, 2009.

[93] D. Bertsimas and I. Popescu. Optimal inequalities in probability theory: A convex optimization approach, 1998.

[94] J. Oliver, Woongbi Lee, Sangjun Park, and Heung-No Lee. Improving resolution of miniature spectrometers by exploiting sparse nature of signals. *Opt. Express*, 20(3):2613–2625, Jan 2012.

[95] P. Borwein and T. Erdelyi. *Polynomials and Polynomial Inequalities*. Graduate Texts in Mathematics. Springer New York, 1995.

Appendix

COPYRIGHT NOTICE

This appendix contains the copyright notices for the contents of this dissertation.

1. Chapter 2 was published in the IEEE Transactions on Signal and Information Processing over Networks ©2019 IEEE. Reprinted, with permission, from Alejandro Parada-Mayorga, Daniel Lau, Jhony Giraldo and Gonzalo Arce.

2. Part of the material in Chapter 4 was published in the IEEE data Science Workshop, Minneapolis, Minnesota, ©2019 IEEE. Reprinted, with permission, from Dominique Guillot, Alejandro Parada-Mayorga, Sebastian Cioaba and Gonzalo Arce.

3. Chapter 5 was published in the IEEE transactions on Computational Imaging, ©2017 IEEE. Reprinted, with permission, from Alejandro Parada-Mayorga and Gonzalo Arce.

4. Chapter 6 was published in the IEEE International Symposium on Biomedical Imaging. Melbourne, Australia, ©2017 IEEE. Reprinted, with permission, from Alejandro Parada-Mayorga, Angela Cuadros and Gonzalo Arce.

5. Chapter 7 was published in the IEEE transactions on Computational Imaging, ©2016 IEEE. Reprinted, with permission, from Alejandro Parada-Mayorga and Gonzalo Arce.

In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of University of Delawares products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing