Spectral Super-Resolution in Colored Coded Aperture Spectral Imaging

Alejandro Parada-Mayorga and Gonzalo R. Arce

***Abstract*—Colored coded apertures have been recently introduced in compressive spectral imaging as a method to improve the quality of image reconstructions in terms of signal to noise ratio. This paper shows that colored coded apertures, in addition, can also provide a higher number of resolvable spectral bands. Colored coded apertures with real and ideal spectral responses are both considered. The maximum number of resolvable bands for the case of nonideal filters is estimated using the coherence of the sensing matrix which provides a condition that depends only on the characteristics of the optical filters involved in the colored coded aperture. Simulations and testbed experimental reconstructions with real data are presented.**

***Index Terms*—CASSI, coherence, colored coded aperture, optical filter, sensing matrix, spectral imaging.**

I. Introduction

SIGNIFICANT interest has emerged in compressed spectral imaging [1] and optical architectures like the coded aperture snapshot spectral imaging (CASSI) [14], [19]. The CASSI system, illustrated in Fig. 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 [1], [19]. The spectral resolution in CASSI is limited mainly by the pitch size of the detector and the spectral dispersion of the prism (Fig. 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. 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 [18], [2]. The use of this colored mask has been shown to improve the quality of the reconstructions [1], [18].

The colored coded aperture is a patterned multilayer optical coating, which physically allows the pass of specific bandwidths

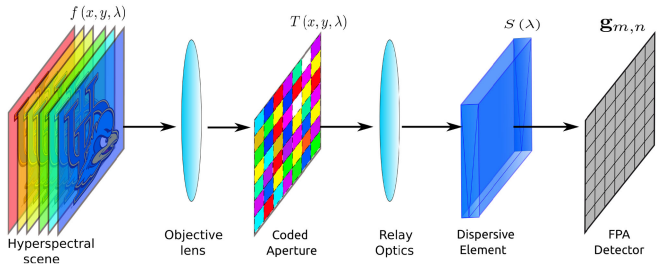

Fig. 1. Basic Components of the colored CASSI system. The spatial-spectral scene $f(x, y, \lambda)$ is modulated by a coded aperture $T(x, y, \lambda)$, after that the modulated field is dispersed by a prism with dispersion curve $S(\lambda)$ and finally the whole field is integrated in a detector array. Notice that in the traditional CASSI $T(x, y, \lambda) \in \{0, 1\}\ \forall\lambda$.

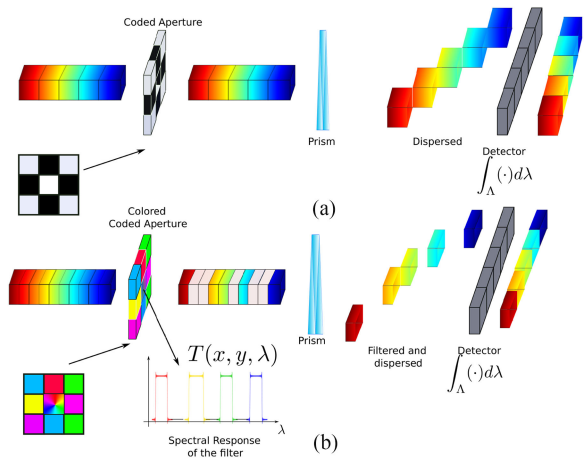

Fig. 2. (a): The spectral resolution in the traditional CASSI is determined by the dispersion of the prism, its spectral range and the detector pixel pitch. (b): The spectral resolution in the colored CASSI is determined by the spectral response of the optical filters in the colored coded aperture, by the dispersion of the prism, its spectral range, and the detector pixel pitch.

in different spatial locations [4], [18]. 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 [4].

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. Fig. 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

Manuscript received February 1, 2016; revised June 17, 2016 and August 29, 2016; accepted September 14, 2016. Date of publication September 22, 2016; date of current version November 4, 2016. This work was supported by the National Science Foundation under Grant VEC1538950 and by Colciencias. The associate editor coordinating the review of this manuscript and approving it for publication was Oliver Cossairt. *(Corresponding author: Alejandro Parada-Mayorga.)*

The authors are with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: alejopm@udel.edu; arce@udel.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCI.2016.2612943

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.

This paper is organized as follows. Section II presents the modeling of the CASSI system, and the discretization approach proposed with an analysis of the limits of the integral operators. It considers the existence of new bands inside the bands attained by the conventional CASSI. Section III shows how the model of the colored CASSI allows spectral super resolution and the super resolution factor for the case of ideal filters is derived. In Section IV, a matrix formulation of the colored CASSI is presented and used to analyze the coherence of the sensing matrix. An estimate of the super-resolution factor is presented when the optical filters involved in the color coded aperture are non ideal. This estimate is defined in terms of concentration inequalities of the coherence of the sensing matrices. In Section V, a set of simulation results are shown confirming the results obtained in previous sections. Section VI show the results of the reconstruction problem with real data. A summary and concluding remarks are showed in Section VII.

II. CASSI Modeling

The basic components in CASSI are shown in Fig. 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 (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 $f(x,y,\lambda)$ [1], [19]. 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 $q(x,y,\lambda)=f(x+S(\lambda),y,\lambda)T(x+S(\lambda),y,\lambda)$, 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 $g(x,y)=\int_\Lambda q(x,y,\lambda)d\lambda$.

A. Discretization of the Model

Consider the following terminology that is used hereafter.

Definition 1: 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 2: 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 [19].

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

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

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

$$\mathbf{g}_{m,n}=\iint_{\Omega_{m,n}} g(x,y)\text{rect}\left(\frac{x}{\Delta}-n,\frac{y}{\Delta}-m\right)dydx, \quad (1)$$

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

$$f(x,y,\lambda)T(x,y,\lambda)\approx \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(TF)_{m,n}(\lambda)\text{rect}\left(\frac{x}{\Delta}-n,\frac{y}{\Delta}-m\right), \quad (2)$$

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 $F_{m,n}(\lambda)=\iint_{\Omega_{m,n}}f(x,y,\lambda)dxdy$. A calculation of the limits involved in the spatial integral operators and its approximation using mid point rules is realized in the Appendix A, which shows that the value of the measurements at pixel (m,n) can be written as

$$\mathbf{g}_{m,n}= \int_\Lambda\Delta^2\left(\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor-\frac{S(\lambda)}{\Delta}+1\right)(TF)_{m,n+\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor}(\lambda)d\lambda - \int_\Lambda\Delta^2\left(\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor-\frac{S(\lambda)}{\Delta}\right)(TF)_{m,n+\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor+1}(\lambda)d\lambda, \quad (3)$$

where $\lfloor\cdot\rfloor$ is the floor function operator. Appendix A 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

$$\mathbf{g}_{m,n}=\sum_{k=0}^{L-1}\left(\mathbf{w}_{m,n,k}(\boldsymbol{TF})_{m,n+\left\lfloor\frac{k}{d}\right\rfloor,k}\right), \quad (4)$$

where $\mathbf{w}_{m,n,k,u}=\boldsymbol{\alpha}_{m,n,k}+\boldsymbol{\beta}_{m,n,k}$ for $u=1,\dots,c-2$, $\mathbf{w}_{m,n,k,0}=\boldsymbol{\alpha}_{m,n,k}$ 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 A 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 (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 $KN(N + \lfloor \frac{L}{d} \rceil - 1) \times N^2 L$ (see Fig. 4) with K being the number of shots and $\vec{\mathbf{F}}$ is the vectorized form of $\mathbf{F}$. Fig. 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

$$\vec{\mathbf{g}} = \mathbf{H}\boldsymbol{\psi}\mathbf{f}, \tag{5}$$

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 [15], [19]

$$\hat{\mathbf{f}} = \underset{\mathbf{z}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{A}\mathbf{z} - \vec{\mathbf{g}} \right\|_2^2 + \tau \left\| \mathbf{z} \right\|_1, \tag{6}$$

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

III. 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 [16]. The second case will be presented in the next section after showing the expression of the coherence of the sensing matrix.

A. Optical Filters With Ideal Transition Bands

The prism in CASSI is used to separate the spectral information into bands [1], [19]. 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. 3, in which two cases of optical filtering are considered. In the first case (Fig. 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. 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.

1) When the cutoff frequencies of the filters do not match the boundaries of the basic bands: Let $B = \{B_0, B_1, \ldots, B_{L'-1}\}$ 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

$$|\hat{B}| = |B| + \sigma - 1, \tag{7}$$

where $\hat{B} = \{b_0, b_1, \ldots, b_{L-1}\}$ is the set of new bands.

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

$$|\hat{B}| = \begin{cases} |B| & \text{If} \quad \sigma \leq |B| \\ \sigma & \text{If} \quad \sigma > |B|. \end{cases} \tag{8}$$

IV. 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 [13], [15]. The former is in general difficult to calculate for large size matrices [15]. 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 [10]. Additionally, the coherence can be related with the RIP [7], [10], [13] 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 [15], [7]. 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.

A. Matrix Formulation

The structure of the sensing matrix $\mathbf{H}$ for the super-resolution model can be seen in Fig. 4. The elements that lie on the indicated diagonals represent the spectral response of the coded aperture in each band. In Fig. 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. 4 can be represented in precise mathematical terms. The m-th position in the j-th column of the $\mathbf{H}$ matrix can

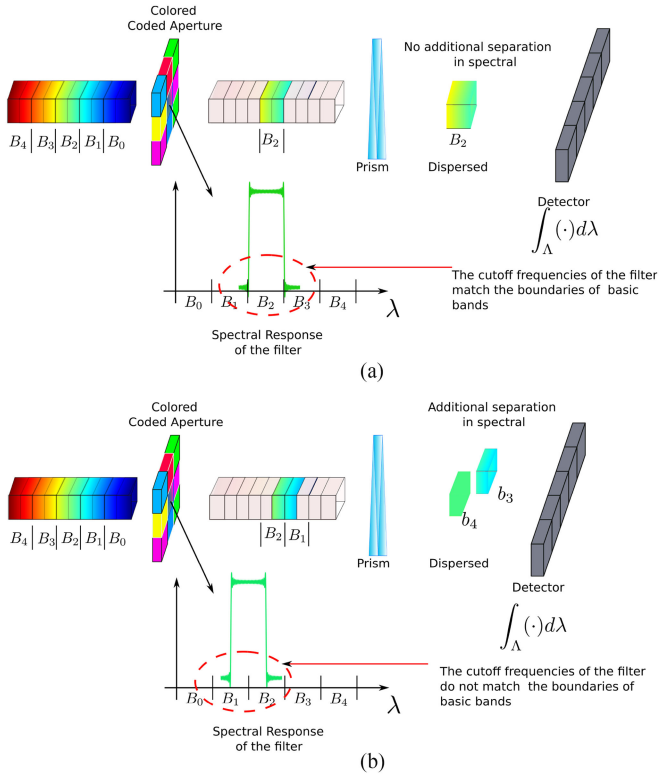

Fig. 3. (a) A pixel of the scene is filtered by one band pass filter for which the cutoff frequencies coincide exactly with the boundaries of the basic bands. (b) A pixel of the scene is filtered by one band pass filter for which the cutoff frequencies do not coincide with the boundaries of the basic bands. In this case super resolution is achieved.

be written as

$$\mathbf{h}_j(m) = \begin{cases} \left(\vec{\boldsymbol{T}}_\ell^{(s)} \vec{\mathbf{w}}_\ell\right)_{m-N(N+L-1)s-N\left\lfloor \frac{\ell}{d} \right\rfloor} \\ \text{If} \quad j = m + \ell N^2 - N\left\lfloor \frac{\ell}{d} \right\rfloor \ldots \\ \ldots - N(N+L-1)s \\ 0 \qquad \text{otherwise,} \end{cases} \tag{9}$$

with $s = \lfloor \frac{m}{N(N+L-1)} \rfloor, \ell = \lfloor \frac{j}{N^2} \rfloor$, 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 $\mathbf{w}(\cdot,\cdot,\ell)$.

B. Analysis of the Coherence

Let $\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \ldots \mathbf{h}_{N^2L}]$, where each $\mathbf{h}_i$ is the i^{th} column vector of $\mathbf{H}$, of dimension $KN(N + \lfloor \frac{L}{d} \rceil - 1) \times 1$. The basis matrix $\boldsymbol{\Psi}$ can be written as $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1^T, \quad \boldsymbol{\psi}_2^T, \cdots, \boldsymbol{\psi}_{N^2L}^T]^T$, 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 $\mathbf{A}(m,n) = \sum_{i=1}^{N^2L} \mathbf{h}_i(m)\boldsymbol{\psi}_i(n)$.

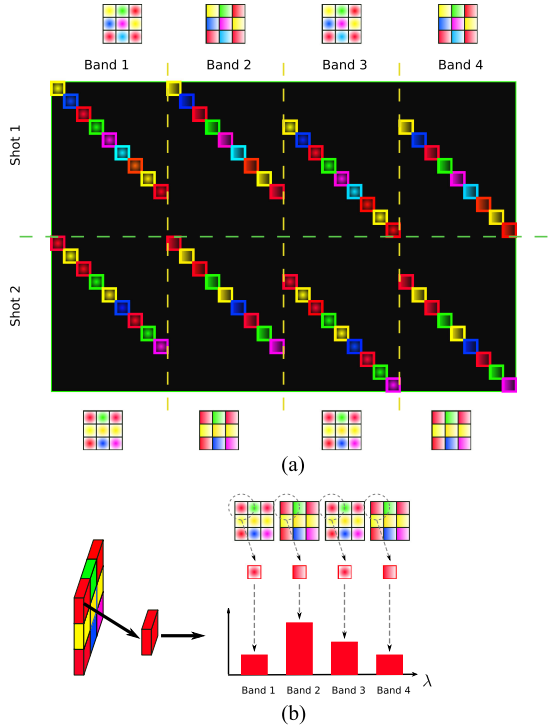

Fig. 4. a) The structure of the $\mathbf{H}$ matrix of the colored CASSI when a super resolution factor of $d = 2$ is considered. b) The spectral response one pixel in the coded aperture is detailed and how its values are distributed in the $\mathbf{H}$ matrix.

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

$$\begin{aligned} \mu(\mathbf{A}) &= \max_{\substack{m,n \\ m \neq n}} \left| \left\langle \frac{\mathbf{A}(\cdot,m)}{\|\mathbf{A}(\cdot,m)\|}, \frac{\mathbf{A}(\cdot,n)}{\|\mathbf{A}(\cdot,n)\|} \right\rangle \right| \\ &= \max_{\substack{m,n \\ m \neq n}} \mu_{mn}(\mathbf{A}), \end{aligned}$$

where

$$\mu_{mn}(\mathbf{A}) = \frac{\left| \sum_{i,j}^{N^2L} \varphi_{i,j} R_{i,j}^{(m,n)} \right|}{\left(\sum_{i,j}^{N^2L} \varphi_{i,j} R_{i,j}^{(m,m)} \right)^{\frac{1}{2}} \left(\sum_{i,j}^{N^2L} \varphi_{i,j} R_{i,j}^{(n,n)} \right)^{\frac{1}{2}}}, \tag{10}$$

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 $I = \{(i,j) | i \neq j, \varphi_{i,j} \neq 0\}$, which is the set of all possible inner products that are different from zero. Considering I, equation (10) can be written as

$$\mu_{mn}(\mathbf{A}) = \frac{|\varrho(m,n)|}{\varrho(m,m)^{\frac{1}{2}} \varrho(n,n)^{\frac{1}{2}}}, \tag{11}$$

where

$$\varrho\left(m,n\right)=\sum_{i=1}^{N^2L}\varphi_{i,i}R_{i,i}^{(m,n)}+\sum_{(i,j)\in I}\varphi_{i,j}R_{i,j}^{(m,n)}. \tag{12}$$

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.

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 $f_{X_rY_r}(n)=p^{2n}(1-p^2)^{1-n} \quad n=0,1 \quad r=1,\ldots,K$ since the sequence of products X_rY_r are i.i.d Bernoulli random variables with parameter p^2. Using Chebyshev's inequality, it follows that

$$\mathbb{P}\left(\left|\varphi_{i,j}-\mathbb{E}\left\{\varphi_{i,j}\right\}\right|\geq\varepsilon\right)\leq\frac{1}{\varepsilon^2}\sum_{i=1}^{K}\mathrm{Var}\left(X_rY_r\right) \quad (i,j)\in I,$$

which implies

$$\mathbb{P}\left(\left|\varphi_{i,j}-\left(Kp\right)p\right|\geq\varepsilon\right)\leq\frac{Kp}{\varepsilon^2}p(1-p^2) \quad (i,j)\in I. \tag{13}$$

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

$$\mathbb{P}\left(\left|\varphi_{i,i}-\left(Kp\right)\right|\geq\varepsilon\right)\leq\frac{Kp}{\varepsilon^2}(1-p) \quad \forall i. \tag{14}$$

In the particular case $p=1/2$, the expression is

$$\mathbb{P}\left(\left|\varphi_{i,i}-\left(\frac{K}{2}\right)\right|\geq\varepsilon\right)\leq\left(\frac{K}{2}\right)\frac{1}{2\varepsilon^2}. \tag{15}$$

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

$$\mathbb{P}\left(\left|\varphi_{i,j}-\left(\frac{K}{\sigma}\right)\frac{1}{\sigma}\right|\geq\varepsilon\right)\leq\left(\frac{K}{\sigma}\right)\frac{1}{\sigma\varepsilon^2}\left(1-\frac{1}{\sigma^2}\right), \tag{16}$$

[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

$$\mathbb{P}\left(\left|\varphi_{i,i}-\left(\frac{K}{\sigma}\right)\right|\geq\varepsilon\right)\leq\left(\frac{K}{\sigma}\right)\frac{1}{\varepsilon^2}\left(1-\frac{1}{\sigma}\right). \tag{17}$$

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. 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

$$\mathbb{P}\left(\left|\varphi_{i,j}-\left(\frac{K}{\sigma}\right)\frac{q_{11}}{\sigma}\right|\geq\varepsilon\right)\leq\left(\frac{K}{\sigma}\right)\frac{1}{\sigma\varepsilon^2}\left(q_{22}-\frac{q_{11}^2}{\sigma^2}\right), \tag{18}$$

where

$$q_{11}=\sum_{u,v=1}^{\sigma}T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)T_v\left(\lambda_{\lfloor\frac{j}{N^2}\rfloor}\right), \tag{19}$$

$$q_{22}=\sum_{u,v=1}^{\sigma}T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)^2T_v\left(\lambda_{\lfloor\frac{j}{N^2}\rfloor}\right)^2. \tag{20}$$

For the norm of each column the expression is

$$\mathbb{P}\left(\left|\varphi_{i,i}-\left(\frac{K}{\sigma}\right)q_2\right|\geq\varepsilon\right)\leq\left(\frac{K}{\sigma}\right)\frac{1}{\varepsilon^2}\left(q_4-\frac{q_2^2}{\sigma}\right), \tag{21}$$

where

$$q_2=\sum_{u=1}^{\sigma}T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)^2, \tag{22}$$

$$q_4=\sum_{u=1}^{\sigma}T_u\left(\lambda_{\lfloor\frac{i}{N^2}\rfloor}\right)^4. \tag{23}$$

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 (18), (19) and (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

$$\mathbb{P} \left(\left| \varphi_{i,j} - \left(\frac{K}{\sigma} \right) \hat{q}_1 \right| \geq \varepsilon \right) \leq \left(\frac{K}{\sigma} \right) \frac{1}{\varepsilon^2} \left(\hat{q}_2 - \frac{\hat{q}_1^2}{\sigma} \right), \quad (24)$$

where

$$\hat{q}_1 = \sum_{u=1}^{\sigma} T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \right) T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \pm \ell \frac{\Lambda_{\lfloor \frac{i}{dN^2} \rfloor}}{d} \right), \quad (25)$$

and

$$\hat{q}_2 = \sum_{u=1}^{\sigma} T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \right)^2 T_u \left(\lambda_{\lfloor \frac{i}{N^2} \rfloor} \pm \ell \frac{\Lambda_{\lfloor \frac{i}{dN^2} \rfloor}}{d} \right)^2. \quad (26)$$

The concentration equations for the norm of the columns are still the same as when $d = 1$.

C. 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 (11) and (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 [5]

$$\mathbb{P} \left(\mu(\mathbf{A}) > \varepsilon \right) \leq \frac{\mathbb{E}(\mu(\mathbf{A}))}{\varepsilon}. \quad (27)$$

Then, given a fixed value of ε, it is possible to compare two configurations of the CASSI using the equation (27), in order to get a measure of the performance of the system [9], [8]. 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 1: 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

$$\mathbb{P} \left\{ \mu_{mn}(\mathbf{A})^2 \geq \varepsilon \right\} \leq \frac{\mathbb{E} \left\{ \varrho(m,n)^2 \right\}}{\varepsilon} \xi \quad (28)$$

$$\mathbb{P} \left\{ \mu_{mn}(\mathbf{A}_\sigma)^2 \geq \varepsilon \right\} \leq \frac{\mathbb{E} \left\{ \varrho_\sigma(m,n)^2 \right\}}{\varepsilon} \xi \quad (29)$$

$\forall m \neq n$.

Proof: See Appendix C.

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 3: 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

$$t = \left(\sum_{i=1}^{N^2} \int_{\Lambda} T_i(\lambda) d\lambda \right) \Big/ \Lambda N^2, \quad (30)$$

and $T_i(\lambda)$ is the spectral response of the i^{th} pixel.

Definition 4: 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 2: Let $\varphi_{i,j}$ be as specified in equation (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 (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 $\mathbb{E}(\varphi_{i,j}^{(\sigma)}) \leq \mathbb{E}(\varphi_{i,j})$ and $\text{Var}(\varphi_{i,j}^{(\sigma)}) \leq \text{Var}(\varphi_{i,j})$, then

$$\mathbb{E} \left\{ \varrho_\sigma(m,n)^2 \right\} \leq \mathbb{E} \left\{ \varrho(m,n)^2 \right\} \quad \forall m \neq n. \quad (31)$$

Proof: See Appendix D.

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 3: 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

$$\mathbb{E} \left\{ \varrho_\sigma(m,n)^2 \right\} \leq \mathbb{E} \left\{ \varrho(m,n)^2 \right\}. \quad (32)$$

Proof: See Appendix D.

This corollary and Theorem 1 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.

[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 $\mathbf{H}$ matrix.

D. 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 [17], 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 [17], 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 $K_\sigma = dK$.

This estimation of d is represented by the following definition.

Definition 5: If d represents the super resolution factor described in equations (4) and (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 (4) is the highest value such that

$$d \leq \min\left\{\hat{d}_1, \hat{d}_2\right\} \tag{33}$$

with

$$\hat{d}_1 = \frac{\sigma}{4\hat{q}_1}, \quad \hat{d}_2 = \frac{3\sigma}{16\left(\hat{q}_2 - \frac{\hat{q}_1^2}{\sigma}\right)} \tag{34}$$

where the values of $\hat{q}_1(d,\sigma), \hat{q}_2(d,\sigma)$ are specified by equations (25) and (26).

The definition of d in equation (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 (33) indicates how large d could be in model (4), with a solution of equation (6) with better quality results than in the traditional CASSI with no super resolution.

E. 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 (5) which is considered as independent of the sensing matrix $\mathbf{A}$. The measurements are given by

$$\vec{\mathbf{g}} = \mathbf{A}\mathbf{f} + \mathbf{e}. \tag{35}$$

For any reconstruction algorithm, represented by the operator Δ, it follows that [11], [15]

$$\|\Delta(\mathbf{A}\mathbf{f} + \mathbf{e}) - \mathbf{f}\|_2 \leq C_{\mathbf{A}} \|\mathbf{e}\|_2, \tag{36}$$

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 $C_{\mathbf{A}} = 1/\sqrt{1-\delta_{2k}(\mathbf{A})}$ [11], [15]. 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

$$\frac{\sqrt{1-\delta_{2k}(\mathbf{A})}}{\sqrt{1-\delta_{2k}(\mathbf{A}_\sigma)}} \leq 1. \tag{37}$$

When the matrix $\mathbf{A}_\sigma$ has columns with equal norm, (37) can be written as[3]

$$\frac{\sqrt{1-\delta_{2k}(\mathbf{A})}}{\sqrt{\alpha\left(1-(2k-1)\mu(\mathbf{A}_\sigma)\right)}} \leq 1, \tag{38}$$

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 (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 (38) is satisfied in different ways for each case. Additionally, when $\alpha < 1$ there is one point α_0 at which inequality (38) is not satisfied for all $\alpha < \alpha_0$.

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

[3]The fact that in a unit norm column matrix, the coherence and the RIP are related is used [11], [15].

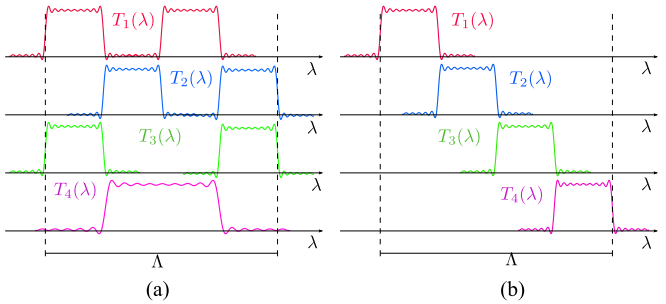

Fig. 5. Two different choices of the set of filters σ in the case that $\sigma = 4$. (a) The set of filters cover the spectral range of interest Λ, each filter allows the transmission of the same amount of energy which corresponds to 50% of the energy contained in Λ. (b) The set of filters cover the 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.

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 (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 Fig. 5 it is possible to see two possible choices for the set of filters in a colored coded aperture where $\sigma = 4$. In Fig. 5(a) the filters used allow the pass of 50% of the energy in the spectral range of interest Λ, whereas in Fig. 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.

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.

V. Simulations

A. Parameters of the Simulations

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 Fig. 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.

2) Spectral responses of the filters: In order to represent the spectral responses of the filters, combinations of Butterworth

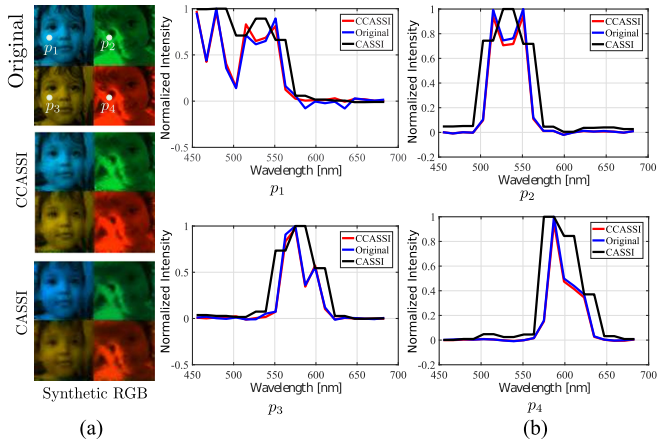

Fig. 6. Results of the simulations considering a super resolution factor $d = 2$. (a) First row: the original target as an RGB representation of the hyperspectral scene. Second row: The RGB representation of the hyperspectral scene reconstructed by the colored CASSI. Third row: The RGB representation 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.

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 $[450\ \text{nm} - 689\ \text{nm}]$. The number of these filters is changed in order to appreciate its effect on the simulations.

3) Prism Curve: The prism curve used for simulations is a realistic prism curve adapted from [3], in order to emulate as close as possible the real nonlinear behavior of the prism.

4) Reconstruction algorithm and basis used in simulations: In order to represent the multispectral scene in terms of a basis, a DCT basis $\mathbf{\Psi}_{\text{DCT}}$ is selected for the spectral domain, whereas a wavelet $\mathbf{\Psi}_W$ for the spatial domain, such that the whole basis is represented as $\mathbf{\Psi}_{\text{DCT}} \otimes \mathbf{\Psi}_W$. The GPSR algorithm [12] 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.

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.

B. Reconstructions

In order to illustrate the performance of the super-resolution, the case in which $d = 2$ is considered. In Figs. 6 and 7 the results of simulations are presented.

In Fig. 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. Fig. 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

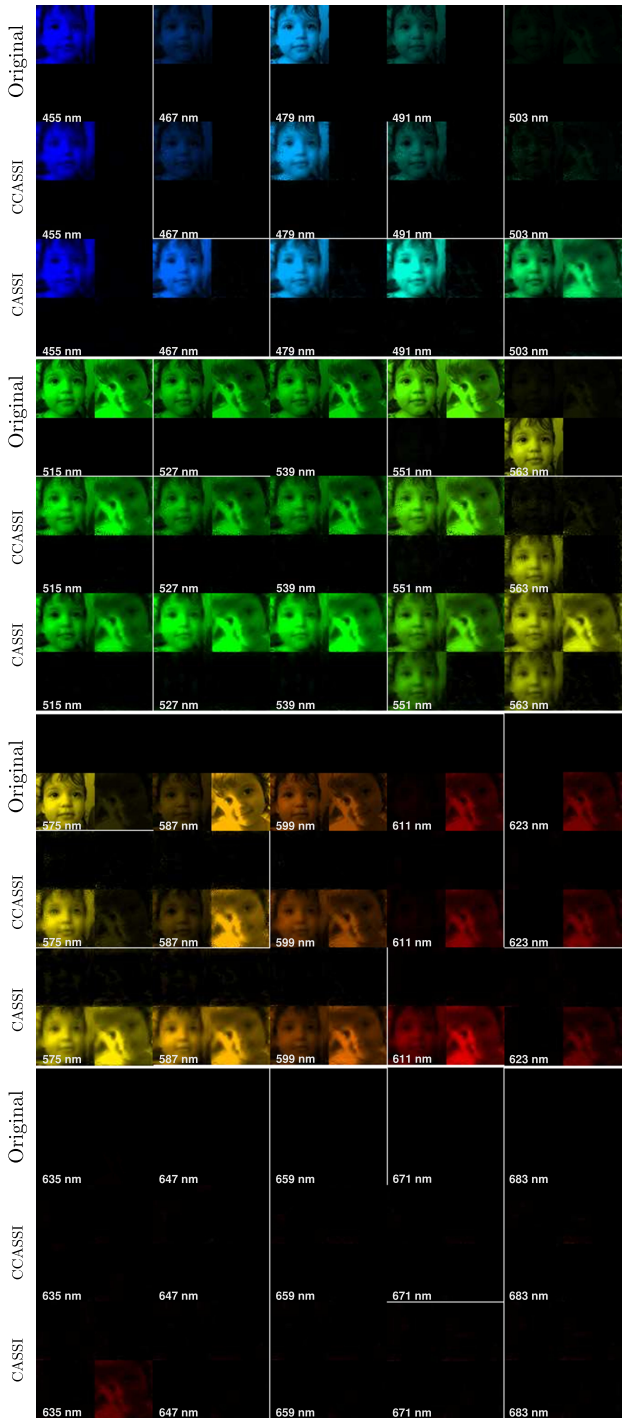

Fig. 7. Results of the simulations considering a super resolution factor $d = 2$. The number of shots used for the traditional CASSI is 5 whereas the number of shots used for the colored CASSI is 10, such that both architectures have the same $\mathcal{C} = 0.5$. The first row of each group depict the 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.

CASSI (black line). Because the CASSI can reconstruct just 10 bands, interpolation is used in order 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

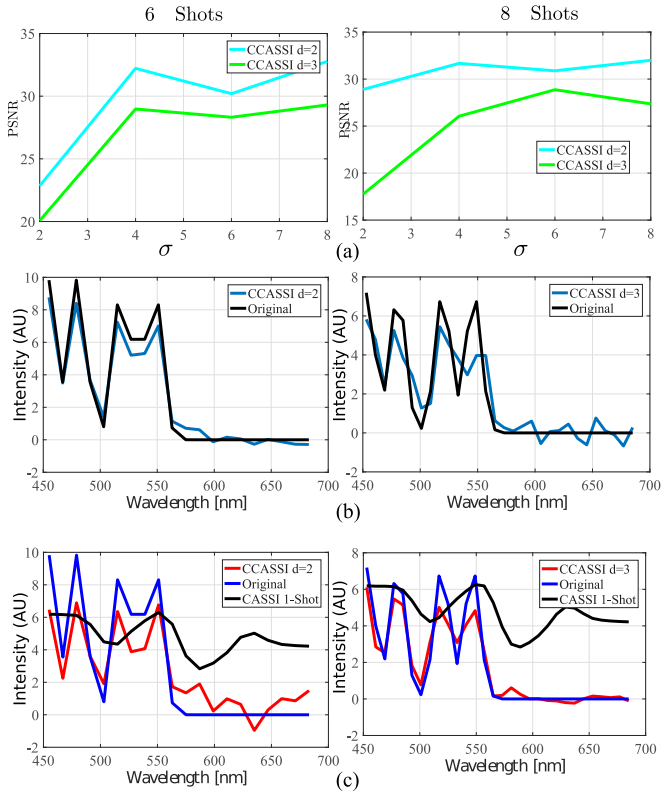

Fig. 8. Simulation results showing the performance of the colored CASSI with super-resolution for different values of σ and d. (a) The PSNR of the reconstructions as a function of σ for $d = 2, 3$ and different number of shots. (b) The spectral response at point p_1 on the target is reconstructed (blue 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.

CASSI are closer to the ground truth than the results obtained by the use of the basic CASSI.

In Fig. 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 $551[nm], 563[nm], 575[nm], 587[nm], 611[nm], 635[nm]$.

C. Simulations: Mean PSNR vs σ

In Fig. 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 $\mathcal{C}$. Fig. 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

TABLE I
THE VALUES OF $\mu(\mathbf{A})$ RELATED WITH THE FIG. 8. K AND σ INDICATE THE NUMBER OF SHOTS AND THE NUMBER FILTERS USED RESPECTIVELY

$K\backslash\sigma$	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
$d = 3$				

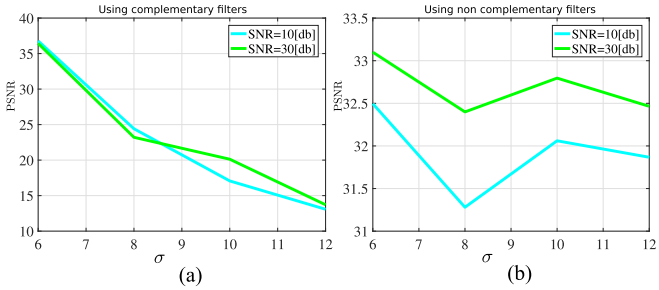

Fig. 9. PSNR of the reconstruction results for different values of σ when the measurements are polluted with additive Gaussian noise. (a) The filters used in the colored coded aperture are selected in a complementary way (see Fig. 5). (b) The filters used in the coded aperture are not complementary but cover the whole spectral range of interest.

TABLE II
THE VALUE OF $\hat{d}_1, \hat{d}_2$

$(\cdot)\backslash\sigma$	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							

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 I. It can be seen how for a given number of shots K, the values of the coherence decrease as σ is increased.

In Fig. 9 the effect of additive Gaussian noise in the measurements is considered using the model defined by equations

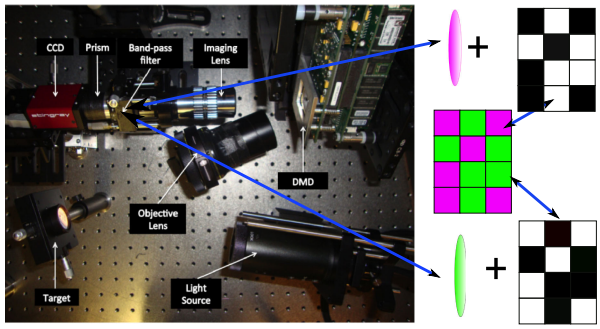

Fig. 10. Picture of the testbed used in the implementation of the colored CASSI. In the picture it is also indicated how the colored coded aperture are implemented for one example of a color coded aperture of 2 filters (See also Fig. 11).

(5) and (35). In Fig. 9(a) the filters used in the coded aperture are chosen as complementary filters (see Fig. 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 Fig. 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.

D. Estimates of the Super-Resolution Factor d

The data presented in the Table II, 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 (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.

VI. EXPERIMENTAL RESULTS

Experimental results were obtained considering the super resolution model proposed. The testbed, shown in Fig. 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 (25 mm, VIS 400-694 nm 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

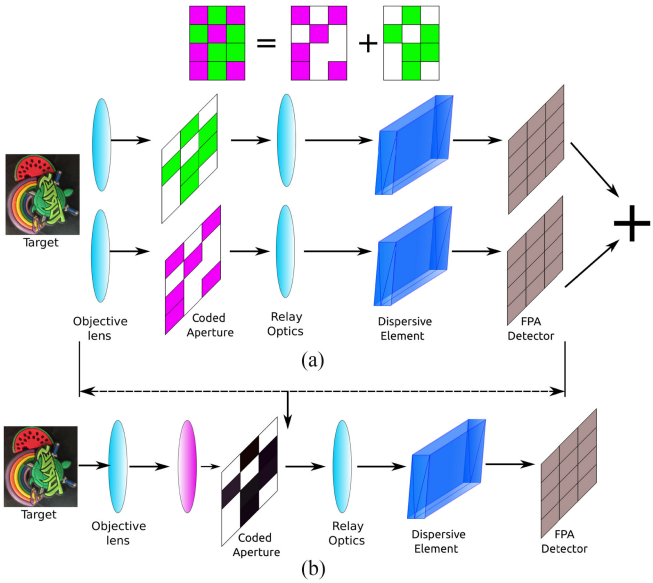

Fig. 11. Details on the implementation of the colored coded aperture. (a) The capture of one shot with a colored coded aperture is the sum of the captures using coded apertures with one single spectral response. (b) In order to get the capture with each of the patterns with one single spectral response a filter is located as indicated in the picture.

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 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 [3] 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.

A. *Implementation of the Colored Coded Aperture*

Fig. 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 Fig. 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 lo-

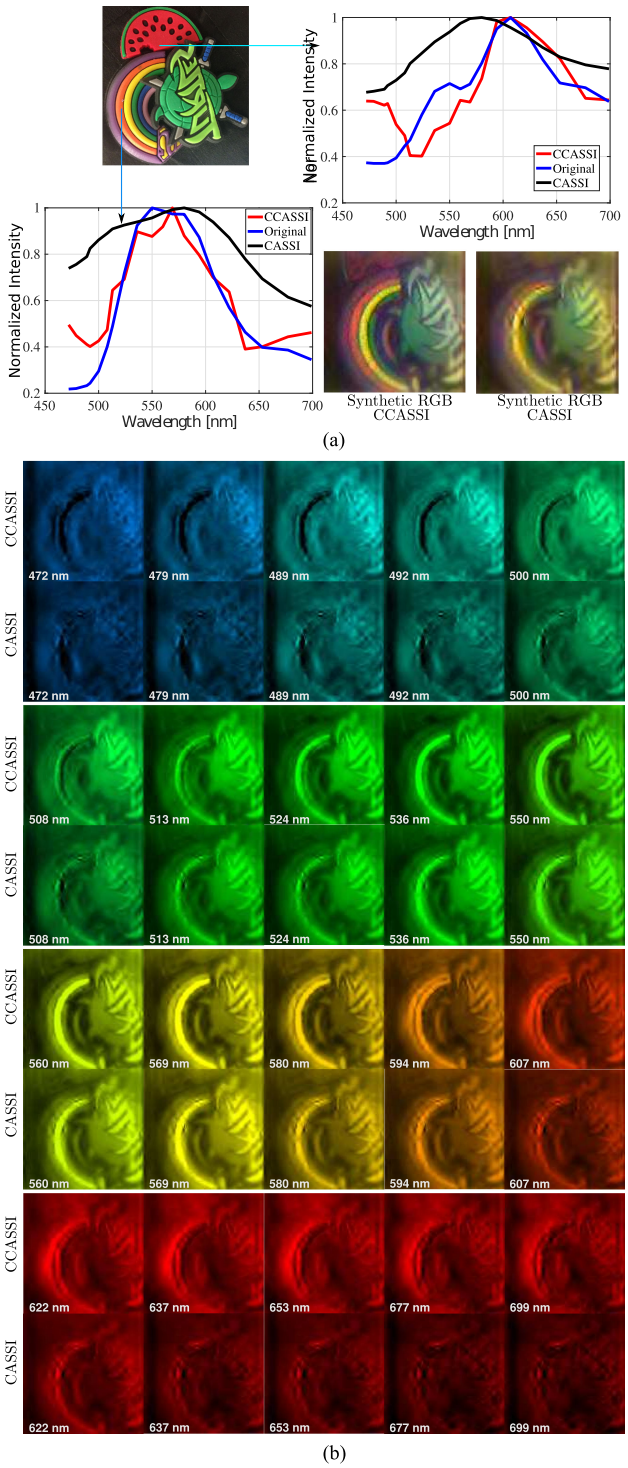

Fig. 12. Results considering a super-resolution factor of $d = 2$ and using $\sigma = 6$ filters. (a): The spectral responses of two different points are reconstructed. The original spectrum (blue line) measured with an spectrometer is compared with the reconstructed spectrum using the colored CASSI (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.

cated. 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.

B. 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 (33) is 6. For the reconstructions the GPSR algorithm [12] 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 Fig. 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 Fig. 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.

Fig. 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.

VII. Conclusion and Future work

This paper 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.

Appendix A
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 $\text{rect}\,(\frac{x}{\Delta} - n)\,\text{rect}\,(\frac{x+S(\lambda)}{\Delta} - n^{'})$ is different from zero when $\Delta(n^{'} + \frac{1}{2}) - S(\lambda) > \Delta(n - \frac{1}{2})$ and $\Delta(n^{'} + \frac{1}{2}) - S(\lambda) < \Delta(n + \frac{1}{2})$, it follows that $n^{'} = \lfloor \frac{S(\lambda)}{\Delta} \rfloor + n$. 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 $\Delta(n^{'} - \frac{1}{2}) - S(\lambda) > \Delta(n - \frac{1}{2})$ and $\Delta(n^{'} - \frac{1}{2}) - S(\lambda) < \Delta(n + \frac{1}{2})$, it follows that $n^{'} = \lfloor \frac{S(\lambda)}{\Delta} \rfloor + n + 1$. 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

$$\begin{aligned}\mathbf{g}_{m,n} = &\int_{\Lambda}\int_{\Delta\left(n-\frac{1}{2}\right)}^{\Delta\left(n+\frac{1}{2}\right)+\Delta\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor-S(\lambda)}\int_{\Delta\left(m-\frac{1}{2}\right)}^{\Delta\left(m+\frac{1}{2}\right)} \\ &(TF)_{m,n+\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor}(\lambda)dydxd\lambda \\ &+\int_{\Lambda}\int_{\Delta\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor+\Delta\left(n+\frac{1}{2}\right)-S(\lambda)}^{\Delta\left(n+\frac{1}{2}\right)}\int_{\Delta\left(m-\frac{1}{2}\right)}^{\Delta\left(m+\frac{1}{2}\right)} \\ &(TF)_{m,n+\left\lfloor\frac{S(\lambda)}{\Delta}\right\rfloor+1}(\lambda)dydxd\lambda,\end{aligned}$$

and using the mid point rule approximation for the spatial integral operators, the measurements can be written as

$$\mathbf{g}_{m,n} = \int_{\Lambda} \Delta^2 \left(\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor - \frac{S(\lambda)}{\Delta} + 1 \right) (TF)_{m,n+\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor}(\lambda) d\lambda - \int_{\Lambda} \Delta^2 \left(\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor - \frac{S(\lambda)}{\Delta} \right) (TF)_{m,n+\left\lfloor \frac{S(\lambda)}{\Delta} \right\rfloor+1}(\lambda) d\lambda. \tag{39}$$

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 $\lfloor \frac{S(\lambda)}{\Delta} \rfloor$, which are explained in Appendix B, it is possible to write equation (3) as

$$\mathbf{g}_{m,n} = \sum_{k=0}^{L-1} \sum_{u'=0}^{c-1} \left[\Delta_{\lambda(k)} \Delta^2 \left(\left\lfloor \frac{S(\hat{\lambda}_k)}{\Delta} \right\rfloor - \frac{S(\hat{\lambda}_k)}{\Delta} + 1 \right) (TF)_{m,n+\left\lfloor \frac{k}{d} \right\rfloor+u'}(\hat{\lambda}_k) - \Delta_{\lambda(k)} \Delta^2 \left(\left\lfloor \frac{S(\hat{\lambda}_k)}{\Delta} \right\rfloor - \frac{S(\hat{\lambda}_k)}{\Delta} \right) (TF)_{m,n+\left\lfloor \frac{k}{d} \right\rfloor+u'+1}(\hat{\lambda}_k) \right],$$

where $c \in \mathbb{N}$. Here c represents the number of pixels of the detector affected by one voxel of the datacube model [3]. 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.

Appendix B
The Meaning of $\lfloor \frac{S(\lambda)}{\Delta} \rfloor$

According to the Weierstrass approximation theorem [6], 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 $\lfloor \frac{\alpha_0}{\Delta} + \frac{q(\lambda)}{\Delta} \rfloor$, 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 $\lfloor \frac{S(\lambda)}{\Delta} \rfloor$ from one integer to another [1], [19]. 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 $\lfloor \frac{S(\lambda)}{\Delta} \rfloor$, varying λ from $\inf\{\Lambda\}$ to $\sup\{\Lambda\}$ [1], [19].

In the term $\lfloor \alpha_1 \frac{\lambda}{\Delta} \rfloor$, 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 $\lfloor \frac{k}{d} \rfloor$, where k is the index that represent the number of the band and d the super resolution factor.

Appendix C
Proof of Theorem 1

Lets 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 $\mathbb{P}\{\varrho(m,n)^2 > x_\ell\varepsilon\} \leq \frac{\mathbb{E}\{\varrho(m,n)^2\}}{x_\ell\varepsilon}$, it follows that

$$\sum_{\ell} \mathbb{P}\left\{\varrho(m,n)^2 > x_\ell\varepsilon\right\}\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}$$
$$\leq \sum_{\ell} \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{x_\ell\varepsilon}\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}$$
$$= \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\sum_{\ell}\frac{\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}}{x_\ell}$$
$$\leq \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\sum_{\ell}\frac{1}{\min_\ell(x_\ell)}\mathbb{P}\left\{\varrho(m,m)\varrho(n,n) = x_\ell\right\}$$
$$= \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\frac{1}{\min_\ell(x_\ell)}.$$

Therefore

$$\mathbb{P}\left\{\mu_{mn}(\mathbf{A})^2 > \varepsilon\right\} \leq \frac{\mathbb{E}\left\{\varrho(m,n)^2\right\}}{\varepsilon}\frac{1}{\min_\ell(x_\ell)}.$$

Following the same steps with $\mathbf{A}_\sigma$ it is possible to get

$$\mathbb{P}\left\{\mu_{mn}(\mathbf{A}_\sigma)^2 > \varepsilon\right\} \leq \frac{\mathbb{E}\left\{\varrho_\sigma(m,n)^2\right\}}{\varepsilon}\frac{1}{\min_\ell(y_\ell)},$$

where y_ℓ represents the values of the random variable $\varrho_\sigma(m,m)\varrho_\sigma(n,n)$. Then, taking

$$\xi = \max\left\{\frac{1}{\min_\ell(x_\ell)}, \frac{1}{\min_\ell(y_\ell)}\right\}$$

the desired result is obtained.

Appendix D
Proof of Theorem 2

In order to simplify the notation for the proof, consider the following convention $\varphi_i \equiv \varphi_{i,i}$, $R_i^{(m,n)} \equiv R_{i,i}^{(m,n)}$ and $\hat{\varphi}_\ell \equiv \varphi_{i,j}$, $\hat{R}_\ell^{(m,n)} \equiv R_{i,j}^{(m,n)}$ $(i,j) \in I$; 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, lets consider

$$\mathbb{E}\left\{\varrho_\sigma(m,n)^2\right\} - \mathbb{E}\left\{\varrho(m,n)^2\right\}$$
$$= \underbrace{\sum_{i=1}^{N^2L}\left(\text{Var}(\varphi_i^{(\sigma)}) - \text{Var}(\varphi_i) + \mathbb{E}\left\{\varphi_i^{(\sigma)}\right\}^2 - \mathbb{E}\{\varphi_i\}^2\right)\left[R_i^{(m,n)}\right]^2}_{①}$$
$$+ \underbrace{\sum_{i\neq j}^{N^2L}\left(\mathbb{E}\left\{\varphi_i^{(\sigma)}\varphi_j^{(\sigma)}\right\} - \mathbb{E}\left\{\varphi_i\varphi_j\right\}\right)R_i^{(m,n)}R_j^{(m,n)}}_{②}$$
$$+ \underbrace{2\left(\mathbb{E}\{\varphi_i^{(\sigma)}\}\mathbb{E}\{\hat{\varphi}_\ell^{(\sigma)}\} - \mathbb{E}\{\varphi_i\}\mathbb{E}\{\hat{\varphi}_\ell\}\right)\sum_{i=1}^{N^2L}\sum_{\ell}R_i^{(m,n)}\hat{R}_\ell^{(m,n)}}_{③}$$
$$+ \underbrace{\sum_{\ell}\left(\text{Var}(\hat{\varphi}_\ell^{(\sigma)}) - \text{Var}(\hat{\varphi}_\ell) + \mathbb{E}\left\{\hat{\varphi}_\ell^{(\sigma)}\right\}^2 - \mathbb{E}\{\hat{\varphi}_\ell\}^2\right)\left[\hat{R}_\ell^{(m,n)}\right]^2}_{④}$$
$$+ \underbrace{\sum_{\ell\neq\ell'}\left(\mathbb{E}\left\{\hat{\varphi}_\ell^{(\sigma)}\hat{\varphi}_{\ell'}^{(\sigma)}\right\} - \mathbb{E}\left\{\hat{\varphi}_\ell\hat{\varphi}_{\ell'}\right\}\right)\hat{R}_\ell^{(m,n)}\hat{R}_{\ell'}^{(m,n)}}_{⑤}.$$

In the following, the analysis of each term of the previous equation es presented separately.

A. About the Term ①

Taking into account that $\mathbb{E}\{\varphi_i^{(\sigma)}\} = \mathbb{E}\{\varphi_i\} = \mathcal{V}$, $\text{Var}(\varphi_i^{(\sigma)}) = \mathcal{V}(1-\frac{1}{\sigma})$ and $\text{Var}(\varphi_i) = \frac{\mathcal{V}}{2}$ the original expression ② can be represented as

$$= \mathcal{V}\left(\frac{1}{2} - \frac{1}{\sigma}\right)\sum_{i=1}^{N^2L}\left[R_i^{(m,n)}\right]^2.$$

B. About the Term ②

For the non independent terms $i \neq j, i \sim j$ it follows that

$$\mathbb{E}\left\{\varphi_i\varphi_j\right\} = \mathbb{E}\left\{\left(\sum_{i=1}^{K}T_{X_i}(\lambda_k)^2\right)\left(\sum_{i=1}^{K}T_{X_i}(\lambda_{k'})^2\right)\right\}$$
$$= \sum_{i=1}^{K}\mathbb{E}\left\{T_{X_i}(\lambda_k)^2T_{X_i}(\lambda_{k'})^2\right\}$$
$$+ \sum_{i\neq j}^{K}\mathbb{E}\left\{T_{X_i}(\lambda_k)^2\right\}\mathbb{E}\left\{T_{X_j}(\lambda_{k'})^2\right\},$$

where $k \neq k^{'} \in \{1,\dots L^{'}\}$, 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: $\frac{K}{2} + \sum_{i\neq j}^{K}\frac{1}{4} = \mathcal{V}(\mathcal{V}+\frac{1}{2})$ when the block unblock coded aperture with transmittance $t = 1/2$ is used, and $\sum_{i\neq j}^{K}\frac{1}{\sigma^2} = \mathcal{V}(\mathcal{V}-\frac{1}{\sigma})$ when the colored coded aperture is used. Now, calculating the difference between the colored case and the binary case, and taking into account that $\mathbb{E}\{\varphi_i^{(\sigma)}\} = \mathbb{E}\{\varphi_i\} = \mathcal{V}$, it

follows that

$$\sum_{i\neq j}^{N^2L}\left(\mathbb{E}\left\{\varphi_i^{(\sigma)}\varphi_j^{(\sigma)}\right\}-\mathbb{E}\left\{\varphi_i\varphi_j\right\}\right)R_i^{(m,n)}R_j^{(m,n)}$$
$$=-\mathcal{V}\left(\frac{1}{\sigma}+\frac{1}{2}\right)\sum_{\substack{i\neq j\\ i\sim j}}R_i^{(m,n)}R_j^{(m,n)}.$$

C. About the Term ③

Notice that

$$2\sum_{i=1}^{N^2L}\sum_{\ell}\mathbb{E}\left\{\varphi_i\hat{\varphi}_\ell\right\}R_i^{(m,n)}\hat{R}_\ell^{(m,n)}$$
$$=2\sum_{\substack{i,\ell\\ i\sim\ell}}\mathbb{E}\left\{\varphi_i\hat{\varphi}_\ell\right\}R_i^{(m,n)}\hat{R}_\ell^{(m,n)}$$
$$+2\sum_{\substack{i,\ell\\ i\nsim\ell}}\mathbb{E}\left\{\varphi_i\right\}\mathbb{E}\left\{\hat{\varphi}_\ell\right\}R_i^{(m,n)}\hat{R}_\ell^{(m,n)}.$$

Now, analyzing the term $\mathbb{E}\{\varphi_i\hat{\varphi}_\ell\}$ 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 $\mathbb{E}\{\varphi_i\hat{\varphi}_\ell\}=\frac{1}{2}\mathcal{V}^2+\frac{1}{4}\mathcal{V}$ and for the colored case $\mathbb{E}\{\varphi_i\hat{\varphi}_\ell\}=\frac{1}{\sigma}\mathcal{V}^2+(\frac{1}{\sigma}-\frac{1}{\sigma^2})\mathcal{V}$.

Making the difference between binary and colored case, the expression ③ is finally represented as

$$=2\mathcal{V}\left(\frac{1}{\sigma}-\frac{1}{\sigma^2}-\frac{1}{4}\right)\sum_{\substack{i,\ell\\ i\sim\ell}}R_i^{(m,n)}\hat{R}_\ell^{(m,n)}.$$

D. 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

$$=\left(\mathcal{V}\left(\frac{1}{\sigma}-\frac{1}{\sigma^3}-\frac{3}{8}\right)+\mathcal{V}^2\left(\frac{1}{\sigma^2}-\frac{1}{4}\right)\right)\sum_{\ell}\left[\hat{R}_\ell^{(m,n)}\right]^2.$$

E. 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 $\frac{\mathcal{V}}{4}(\mathcal{V}+\frac{1}{2})$ for the binary case, and $\frac{\mathcal{V}}{\sigma^2}(\mathcal{V}+1-\frac{1}{\sigma})$ 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

$$a_2=\left(\frac{1}{\sigma^2}-\frac{1}{4}\right)\left(\sum_{\ell}\left[\hat{R}_\ell^{(m,n)}\right]^2+\sum_{\ell\neq\ell'}\hat{R}_\ell^{(m,n)}\hat{R}_{\ell'}^{(m,n)}\right),$$

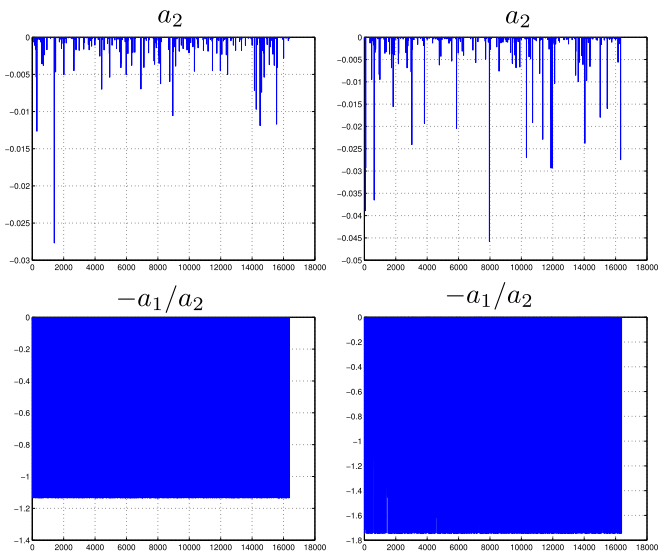

Fig. 13. Samples of the behavior of the coefficients a_2 and a_1. (a): The behavior of the coefficient a_2 and the quotient $-a_1/a_2$ for $\sigma = 3$. (b): The behavior of the coefficient a_2 and the quotient $-a_1/a_2$ for $\sigma = 5$.

$$a_1 = \left(\frac{1}{2} - \frac{1}{\sigma}\right) \sum_{i=1}^{N^2 L} \left[R_i^{(m,n)}\right]^2 - \left(\frac{1}{2} + \frac{1}{\sigma}\right) \sum_{\substack{i \neq j \\ i \sim j}} R_i^{(m,n)} R_j^{(m,n)} + 2\left(\frac{1}{\sigma} - \frac{1}{\sigma^2} - \frac{1}{4}\right) \sum_{\substack{i,\ell \\ i \nsim \ell}} R_i^{(m,n)} \hat{R}_\ell^{(m,n)} + \left(\frac{1}{\sigma} - \frac{1}{\sigma^3} - \frac{3}{8}\right) \sum_{\ell} \left[\hat{R}_\ell^{(m,n)}\right]^2 + \left(\frac{1}{\sigma^2} - \frac{1}{\sigma^3} - \frac{1}{8}\right) \sum_{\substack{\ell \neq \ell' \\ \ell \sim \ell'}} \hat{R}_\ell^{(m,n)} \hat{R}_{\ell'}^{(m,n)}.$$

Then, the behavior of the coefficients a_2 and a_1 completely determine if the inequality (32) is satisfied or not. In Fig. 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 .

REFERENCES

[1] 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.*, vol. 31, no. 1, pp. 105–115, Jan. 2014.

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

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

[4] 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*, vol. 270, no. 12, pp. 6–9, 1995.

[5] D. Bertsimas and I. Popescu, "Optimal inequalities in probability theory: A convex optimization approach," *SIAM J. Optim.*, vol. 15, pp. 780–804, 1998.

[6] P. Borwein and T. Erdelyi, "Polynomials and Polynomial Inequalities," in *Graduate Texts in Mathematics*. New York, NY, USA: Springer, 1995.

[7] T. Tony Cai, Guangwu Xu, and Jun Zhang, "On recovery of sparse signals via l1 minimization," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3388–3397, 2009.

[8] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *Trans. Imgage Process.*, vol. 18, no. 7, pp. 1395–1408, Jul. 2009.

[9] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.

[10] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, 1st ed. New York, NY, USA: Springer, 2010.

[11] Y.C. Eldar *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[12] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.

[13] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Basel, Switzerland: irkhäuser , 2013.

[14] D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," *Appl. Opt.*, vol. 49, no. 36, pp. 6824–6833, Dec. 2010.

[15] Y. C. Eldar, G. Kutyniok, M. A Davenport, and M. F. Duarte, "Introduction to compressed sensing," in *Compressed Sensing*, G. Kutyniok, and Y. C. Eldar, Ed. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[16] J. Oliver, Woong-Bi Lee, and Heung-No Lee, "Filters with random transmittance for improving resolution in filter-array-based spectrometers," *Opt. Express*, vol. 21, no. 4, pp. 3969–3989, Feb. 2013.

[17] J. Oliver, Woongbi Lee, Sangjun Park, and Heung-No Lee, "Improving resolution of miniature spectrometers by exploiting sparse nature of signals," *Opt. Express*, vol. 20, no. 3, pp. 2613–2625, Jan. 2012.

[18] H. Rueda, H. Arguello, and G. R. Arce, "Experimental demonstration of a colored coded aperture-based compressive spectral imaging system," in *Proc. Classical Opt.*, 2014, Paper no. CTu2C.6.

[19] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, no. 10, pp. B44–B51, Apr. 2008.

Alejandro Parada-Mayorga received the B.Sc. and the Master's degrees in electrical engineering from the Industrial University of Santander, Bucaramanga, Colombia. He is currently working toward the Ph.D. degree in electrical engineering at the University of Delaware, Newark, DE, USA, under the supervision of Dr. G. R. Arce. His research interests include analysis and processing of high-dimensional data, compressed sensing, high-dimensional geometry, computational harmonic analysis, optimization, and compressive spectral imaging.

Gonzalo R. Arce is the Charles Black Evans Professor in the Electrical and Computer Engineering Department, University of Delaware, Newark, DE, USA. He held the 2011 Nokia-Fulbright Distinguished Chair in Information and Communications Technologies, Aalto University, Helsinki, Finland. From 1999 to 2009, he served as the Chair of the Electrical and Computer Engineering Department. He holds 15 U.S. patents. His research interests include computational imaging and spectroscopy, signal processing, and the analysis and processing of high-dimensional data. He received the JPMorgan-Chase Faculty Award from the Institute of Financial Services Analytics, University of Delaware. He is the author of four textbooks in the areas of computational imaging and statistical signal processing. He is a Fellow of the Center for Advanced Studies at the University of Delaware.