Journal of the Optical Society of America A
OPTICS, IMAGE SCIENCE, AND VISION

Spatiotemporal blue noise coded aperture design for multi-shot compressive spectral imaging

CLAUDIA V. CORREA,[1] HENRY ARGUELLO,[2] AND GONZALO R. ARCE[1,*]

[1]Department of Electrical and Computer Engineering, University of Delaware, Newark, Delaware 19716, USA
[2]Universidad Industrial de Santander, Bucaramanga 680002, Colombia
*Corresponding author: arce@udel.edu

Received 28 June 2016; revised 12 September 2016; accepted 12 October 2016; posted 13 October 2016 (Doc. ID 269176); published 9 November 2016

Multi-shot coded aperture snapshot spectral imaging (CASSI) systems capture the spectral information of a scene using a small set of coded focal plane array (FPA) compressive measurements. Compressed sensing (CS) reconstruction algorithms are then used to reconstruct the underlying spectral 3D data cube from an underdetermined system of linear equations. Multiple snapshots result in a less ill-posed inverse problem and improved reconstructions. The only varying components in CASSI are the coded apertures, whose structure is crucial inasmuch as they determine the minimum number of FPA measurements needed for correct image reconstruction and the corresponding attainable quality. Traditionally, the spatial structures of the coded aperture entries are selected at random, leading to suboptimal reconstruction solutions. This work presents an optimal structure design of a set of coded apertures by optimizing the concentration of measure of the multi-shot CASSI sensing matrix and its incoherence with respect to the sparse representation basis. First, the CASSI matrix system representation in terms of the ensemble of random projections is established. Then, the restricted isometry property (RIP) of the CASSI projections is determined as a function of the coded aperture entries. The optimal coded aperture structures are designed under the criterion of satisfying the RIP with high probability, coined spatiotemporal blue noise (BN) coded apertures. Furthermore, an algorithm that implements the BN ensembles is presented. Extensive simulations and a testbed implementation are developed to illustrate the improvements of the BN coded apertures over the traditionally used coded aperture structures, in terms of spectral image reconstruction PSNR and SSIM.

OCIS codes: (170.1630) Coded aperture imaging; (110.1758) Computational imaging; (110.4234) Multispectral and hyperspectral imaging.

http://dx.doi.org/10.1364/JOSAA.33.002312

1. INTRODUCTION

Spectral imaging captures the spatial information of a scene across a range of wavelengths. Such three-dimensional (3D) information is arranged as a data cube and has been gaining relevance in several applications, including remote sensing in agriculture [1,2], defense [3], and medicine [4]. Traditionally, spectral image acquisition methods scan adjacent zones of the scene to construct the data cube. These methods are particularly expensive due to the overwhelming amount of required data. Recently, compressed sensing principles have been used to acquire the spatial and spectral information of a scene with a reduced amount of coded random projections, which is known as compressive spectral imaging [5]. This acquisition approach alleviates the cost of sensing, storing, and transmitting hyperspectral images. The coded-aperture snapshot spectral imager (CASSI) is a compressive spectral imaging (CSI) architecture that effectively captures the 3D information using a single two-dimensional (2D) projection measurement. Specifically, the image source density $f_0(x, y, \lambda)$ is encoded by means of a coded aperture $T(x, y)$ as depicted in Fig. 1, where (x, y) are the spatial coordinates and λ is the wavelength [6–8]. The resulting coded field $f_1(x, y, \lambda)$ is subsequently sheared by a dispersive element before it impinges on the FPA detector. The compressive measurements across the FPA are realized by the integration of the field $f_2(x, y, \lambda)$ over the spectral range sensitivity of the detector. Assuming that the detector and the coded aperture have a pixel pitch Δ_d and a bandpass filter of the instrument limits the spectral components between λ_1 and λ_2, the number of resolvable bands L is limited by $L = \beta\frac{\lambda_2-\lambda_1}{\Delta_d}$, where $\beta\lambda$ is the dispersion induced by the prism when a linear dispersion function is considered. Similarly, the spectral resolution is limited by $\frac{\Delta_d}{\beta}$. Thus, the horizontal and

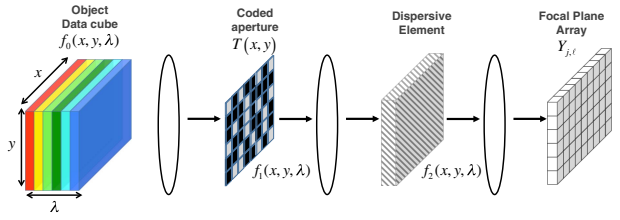

Fig. 1. Optical elements present in CASSI. The input scene $f_0(x,y,\lambda)$ is first spatially encoded by the coded aperture $T(x,y)$, and then, the encoded source is dispersed by a prism. Finally, the coded and dispersed scene is captured by the focal plane array.

vertical spatial resolutions are limited by Δ_d, and the number of spatially resolvable pixels of the underlying scene is $N \times N$.

It has been shown that for spectrally rich scenes or very detailed spatial scenes, a single-shot CASSI measurement may not provide a sufficient number of compressive measurements [9]. Increasing the number of measurement snapshots, each using a different coded aperture that remains fixed during the integration time of the detector, will rapidly increase the quality of the image reconstruction. Given that each CASSI shot simultaneously adds $N(N+L-1)$ compressive measurements, the total number of available measurements when K shots are taken is $m = KN(N+L-1)$. The time-varying coded apertures can be realized by moving a large photomask using a piezo-electric system [9]. A more versatile system was developed in [10], in which a digital micromirror device (DMD) was used to vary the random patterns on each snapshot.

Given a set of FPA compressive measurements, the compressive sensing (CS) theory is then exploited to recover the underlying 3D spectral data cube by finding the sparsest approximation with the minimum Euclidean distance to the 2D random projection measurements. More specifically, let $\mathbf{F}$ be the $N \times N \times L$ spectral data cube, or its vectorized representation $\mathbf{f}$. Then, the ith FPA measurement is obtained as $\mathbf{y}^i = \mathbf{H}^i\mathbf{f}$, where $\mathbf{H}^i$ is an $N(N+L-1) \times N^2L$ projection matrix representing the effects of the coded aperture and the dispersive element. The set of K FPA measurements, each with a different coded aperture, is then represented as $\mathbf{y} = [(\mathbf{y}^0)^T, \ldots, (\mathbf{y}^{K-1})^T]^T$. The underlying data cube is estimated as $\tilde{\mathbf{f}} = \mathbf{\Psi}(\text{argmin}_{\boldsymbol{\theta}}\|\mathbf{y} - \mathbf{H}\mathbf{\Psi}\boldsymbol{\theta}\|_2 + \tau\|\boldsymbol{\theta}\|_1)$, where $\mathbf{H} = [(\mathbf{H}^0)^T, \ldots, (\mathbf{H}^{K-1})^T]^T$, $\boldsymbol{\theta}$ is an S-sparse representation of $\mathbf{f}$ on the basis $\mathbf{\Psi}$, and τ is a regularization constant.

A critical component in the inverse problem that estimates $\mathbf{f}$ from $\mathbf{y}$ is the structure of the sensing matrix $\mathbf{A} \triangleq \mathbf{H}\mathbf{\Psi}$, as it ultimately determines the attainable quality of reconstruction. While the optical architecture in CASSI partially imposes a well-defined sparse structure onto the sensing matrix $\mathbf{A}$, the coded apertures used in each measurement shot determine the structure of $\mathbf{A}$. The objective in CASSI is thus to optimally design the set of coded apertures so as to forge a structure on $\mathbf{A}$ that minimizes the number of FPA snapshots while attaining the highest-quality reconstruction. To this end, this paper explores the interplay between the restricted isometry property (RIP) and the set of coded apertures used in CASSI.

The RIP gives a characterization of sensing matrices in CS. It establishes the necessary conditions for $\mathbf{A}$ such that the ℓ_2 norm of the underlying 3D spectral image is approximately preserved under the transformation $\mathbf{A}\boldsymbol{\theta}$. More precisely, the restricted isometry constant δ_s of the matrix $\mathbf{A}$ is the smallest constant such that $(1-\delta_s)\|\boldsymbol{\theta}\|_2 \le \|\mathbf{A}\boldsymbol{\theta}\|_2 \le \|\boldsymbol{\theta}\|_2(1+\delta_s)$ [11]. The RIP requires that all $m \times |\mathcal{T}|$ column submatrices $\mathbf{A}_{|\mathcal{T}|}$ of $\mathbf{A}$ are well conditioned for all $|\mathcal{T}| \le S$. Indeed, the RIP imposes that all the eigenvalues of the matrices $\mathbf{A}_{|\mathcal{T}|}^T\mathbf{A}_{|\mathcal{T}|}$ are in the interval $[1-\delta_s, 1+\delta_s]$. The probability of satisfying this condition is calculated by estimating the statistical distribution of the maximum eigenvalue $\lambda_{\max}$ of the matrices $\mathbf{A}_{|\mathcal{T}|}^T\mathbf{A}_{|\mathcal{T}|} - \mathbf{I}$, where $\mathbf{I}$ is an identity matrix. The distribution of the maximum eigenvalue $\lambda_{\max}$ is estimated using the concentration of measure for random matrices developed in [12]. The RIP condition also implies a stable recovery of the signal $\boldsymbol{\theta}$ from the projections $\mathbf{A}\boldsymbol{\theta}$ using a ℓ_1 optimization algorithm [11].

Frequently used CS matrices are unitary dense Gaussian random matrices whose entries are drawn from zero-mean Gaussian random variables with $1/m$ variance. Bernoulli random matrices with entries taking on values $\{-1/\sqrt{m}, 1/\sqrt{m}\}$ with equal probability are also common. These random matrices have well-established RIP properties. The RIP has been also studied for various structured random matrix topologies, including Toeplitz matrices with random Gaussian entries [13,14], random circulant matrices [11], and partial random Toeplitz matrices [15], which are often used in applications such as sparse channel estimation and multi-user detection. The RIP property for block diagonal random matrices (RBD) used in sensor networks and multiple-view imaging was studied in [16]. Also, a more general framework for structurally random matrices (SRM) has been established in [17]. SRM matrices are defined as the product $\mathbf{\Phi} = \sqrt{n/m}\,\mathbf{DFR}$, where $\mathbf{R}$ is a random permutation matrix, $\mathbf{F}$ is a fast computable matrix, $n = N^2L$ in the CSI problem, and $\mathbf{D}$ is a subsampling matrix. SRM can be designed with high flexibility using different combinations of fast computable matrices and block-based processing. The RIPs for other types of matrix structures were studied in [15,18].

The RIP can be used to determine bounds on the required number of measurements needed for successful CS reconstruction. These bounds depend on the structure of the underlying sensing matrix. The required number of measurements for Gaussian or Bernoulli ensembles is, for instance, $m \ge CS\log(n/S)$ for a positive constant C, whereas the required number for Toeplitz matrices is $m \ge CS^2\log(n)$ and, for random convolution, $m \ge CS\log(n)$. In the case of RBD and SRM matrices, $m \ge CS\log(n/S)$ is a bound that can be reached under certain conditions of the underlying signal $\mathbf{f}$ and by the proper selection of the matrices $\mathbf{R}$, $\mathbf{F}$, and $\mathbf{D}$.

Despite the importance of the RIP in CS, the RIP for CASSI sensing matrices has not been fully exploited. Given that the sensing matrices in CASSI are highly sparse and structured, the RIP characterization of dense sensing matrices is not applicable. The first approach to characterize the RIP in CASSI was given in [7], where it was beforehand assumed that if the RIP for the matrix $\mathbf{A}$ is satisfied for some constant δ_s, then conditions on the coded apertures were determined so that the RIP is better satisfied. These results, however, do not involve the RIP in the optimal design of the coded apertures. In fact, commonly used coded apertures in CASSI include Hadamard matrices $\mathbf{H}_N$, whose entries are $(H_N)_{ij} \in \{-1, 1\}^{N\times N}$,

Hadamard S matrices $\mathbf{S}_N$, where $\mathbf{S}_N = 1/2(1 - \mathbf{H}_N)$ [19], cyclic S-matrices consisting of cyclic permutations of a single master codeword [20], and Bernoulli random matrices [10,21]. The use of these coded apertures in CASSI has been principally motivated by the fact that they are well conditioned when used in least squares estimations [19,22]. However, these coded apertures do not fully exploit the rich theory of CS. In particular, they do not exploit the RIP condition or the concentration of measure of the respective random submatrices of $\mathbf{A}$ to define optimal coded-aperture sets. A more recent coded-aperture design, based on the concentration of measure, can be found in [23], where the correlation between measurement snapshots is taken into account to better satisfy the RIP. Despite the improved reconstructions that can be obtained with this design, its main weakness is that the spatial correlation of the one-valued entries is not considered. The design in [23] was called "Boolean coded apertures."

In this paper, the structure of the multi-shot CASSI sensing matrix is formulated, the RIP for CASSI sensing matrices is derived, and the corresponding RIP constants are expressed as a function of the structure of the random coded-aperture patterns. Instead of simply increasing the number of FPA projections to provide a better-quality image reconstruction, it will be shown that an optimal design of coded apertures can be used to better satisfy the RIP requirements in multi-shot CASSI. The resulting optimal set of coded apertures are called "blue noise (BN) coded apertures," since their distribution exhibits spatio-temporal characteristics of blue noise patterns that suppress low-frequency components of noise [24–26]. These coded apertures allow uniform sensing across spatial and spectral dimensions of the scene within different measurement snapshots. BN coded apertures are then compared with the traditional random coded apertures used in CASSI. Simulations are used to illustrate the benefits of this design. Finally, testbed reconstructions using BN coded apertures in the CASSI system are presented. The following notations should be used: scalars are denoted by italic lowercase roman, vectors by boldface lowercase roman, matrices by boldface uppercase roman, and multidimensional data, such as multispectral images, by uppercase calligraphic.

2. MULTI-SHOT CASSI MATRIX REPRESENTATION

Let $T^i_{j,\ell}$ be the (j,ℓ)th spatial index of a discretized representation of the coded aperture used to sense the ith FPA measurement, and let $F_{j,\ell,k}$ be the spatio-spectral source with k determining the kth spectral plane. The discrete output at the detector is given by [6,9,21]

$$Y^i_{j,\ell} = \sum_{k=0}^{L-1} F_{j,(\ell-k),(k)}\, T^i_{j,(\ell-k)} + \omega^i_{j,\ell} \quad i = 0, \ldots, K-1, \tag{1}$$

where i indexes the FPA snapshot, $Y^i_{j,\ell}$ is the intensity measured at the pixel (j,ℓ) of the FPA detector, and $\omega^i_{j,\ell}$ is the noise of the system. The coded apertures $T^i_{j,\ell}$ are time-varying and indexed by i. The output of the system Y^i in Eq. (1) is an $N \times V$ signal, with $V = N + L - 1$, yielding the compression ratio $NV/N^2L \approx 1/L$. The FPA measurement can be written in vector notation as

$$\mathbf{y}^i = \mathbf{H}^i\mathbf{f} + \boldsymbol{\omega}^i, \tag{2}$$

where $\mathbf{y}^i \in \mathbb{R}^{NV}$ is a NV-long vector representation of $\mathbf{Y}^i$, $\mathbf{H}^i \in \{0,1\}^{NV\times n}$ represents the effects of the coded aperture and the dispersive element, and $\mathbf{f} = \mathrm{vec}([\mathbf{f}_0, \ldots, \mathbf{f}_{L-1}])$ is the vector representation of the data cube $\mathbf{F}$, where $\mathbf{f}_k$ is the vector form of the kth spectral band. More specifically, the entries of $\mathbf{f}_k$ can be expressed as $(f_k)_\ell = F_{(\ell-rN),r,k}$ for $\ell = 0, \ldots, N^2 - 1$, $k = 0, \ldots, L-1$, where $r = \lfloor \frac{\ell}{N} \rfloor$. Using this definition of r, the vectorization of the coded aperture $T^i_{j,\ell}$ can be defined as $(t^i)_\ell = T^i_{(\ell-rN),r}$ for $\ell = 0, \ldots, N^2 - 1$, $i = 0, \ldots, K-1$. Similarly, the vector representation of the output $\mathbf{Y}^i$ is written as $(y^i)_\ell = Y^i_{(\ell-rN),r}$ for $\ell = 0, \ldots, NV - 1$, $i = 0, \ldots, K-1$. Based on the above matrix representation, the output $\mathbf{y}^i$ in Eq. (2) can be expressed as

$$\mathbf{y}^i = \overbrace{\begin{bmatrix} \mathrm{diag}(\mathbf{t}^i) & \mathbf{0}_{N(1)\times N^2} & \cdots & \mathbf{0}_{N(L-1)\times N^2} \\ & \mathrm{diag}(\mathbf{t}^i) & \cdots & \\ & & \ddots & \\ \mathbf{0}_{N(L-1)\times N^2} & \mathbf{0}_{N(L-2)\times N^2} & \cdots & \mathrm{diag}(\mathbf{t}^i) \end{bmatrix}}^{\mathbf{H}^i} \begin{bmatrix} \mathbf{f}_0 \\ \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_{L-1} \end{bmatrix}, \tag{3}$$

where $\mathrm{diag}(\mathbf{t}^i)$ is an $N^2 \times N^2$ diagonal matrix whose entries are the elements of the vectorized coded aperture $\mathbf{t}^i$, and $\mathbf{0}_{N(1)\times N^2}$ and $\mathbf{0}_{N(L-1)\times N^2}$ are $N(1) \times N^2$ and $N(L-1) \times N^2$ zero-valued matrices, respectively. The matrices for multiple snapshots $\mathbf{H}^i$ are assembled in the matrix $\mathbf{H}$ given by $\mathbf{H} = [(\mathbf{H}^0)^T, \ldots, (\mathbf{H}^{K-1})^T]^T$. Figure 2 illustrates the structure of the sensing matrix $\mathbf{H}$ for $N = 8$, $L = 4$, and $K = 2$, in which the entries $(t^i)_\ell$ are realizations of a Bernoulli random variable with parameter $p = 0.5$. In practice, the dimensions of the matrices $\mathbf{H}$ and $\mathbf{A}$ are much larger [9,21]. It can be observed in Fig. 2 that the jth row of $\mathbf{H}$ contains at most L non-zero elements. Further, the entries of the jth row of the sensing matrix $\mathbf{H}$ can be written as

$$(\mathbf{h}_j)_\ell = \begin{cases} t^i_{j-rN}, & \text{if } \ell = (j \bmod N') \text{ and } j - rN \geq 0 \\ 0, & \text{otherwise} \end{cases}, \tag{4}$$

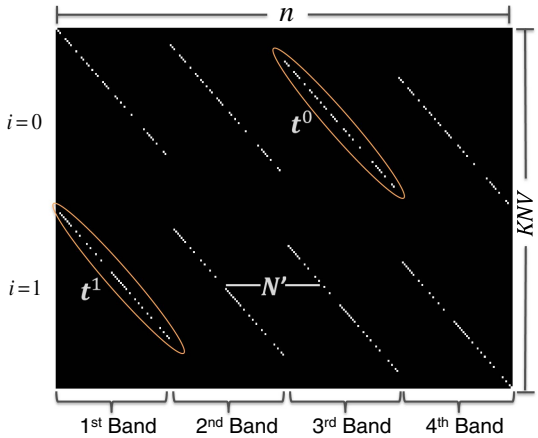

Fig. 2. Structure of the matrix $\mathbf{H}$ is illustrated for $K = 2$, $N = 8$, and $L = 4$. For illustration, the entries of the coded apertures are $\mathbf{t}^i_j \sim$ Bernoulli(0.5). The dark elements represent zero values and the white elements correspond to the one-valued entries of the coded aperture.

for $\ell = 0, \ldots, N^2L - 1$, where $r = \lfloor \frac{\ell}{N^2} \rfloor$, $N' = N^2 - N$ and $i = \lfloor j/NV \rfloor$. The ensemble of CASSI outputs $\mathbf{y} = [(\mathbf{y}^0)^T, \ldots, (\mathbf{y}^{K-1})^T]^T$ can be rewritten as

$$\mathbf{y} = \mathbf{A}\boldsymbol{\theta} = \mathbf{H}\boldsymbol{\Psi}\boldsymbol{\theta} + \boldsymbol{\omega}, \tag{5}$$

where $\mathbf{A} = \mathbf{H}\boldsymbol{\Psi}$ is the CASSI sensing matrix, $\boldsymbol{\theta}$ is a sparse representation of $\mathbf{f}$ on the basis $\boldsymbol{\Psi}$, and $\boldsymbol{\omega}$ represents the noise of the system. Notice that $\mathbf{A} \in \mathbb{R}^{m \times n}$, where $m = KNV$, $n = N^2L$, and $m \leq n$. For instance, the basis matrix $\boldsymbol{\Psi}$ can be a Kronecker representation basis given by $\boldsymbol{\Psi} = \boldsymbol{\Psi}^{2D} \otimes \boldsymbol{\Psi}^{C}$, where $\boldsymbol{\Psi}^{C}$ is the one-dimensional cosine matrix transform, $\boldsymbol{\Psi}^{2D}$ is the 2D symlet wavelet transform, which treats each spectral band independently, and $\otimes$ is the Kronecker product. In this case, the entries of $\boldsymbol{\theta} = \boldsymbol{\Psi}^T\mathbf{f}$ account for the correlation among all the elements in the data cube $\mathbf{f}$. In summary, given that $\boldsymbol{\Psi}$ is fixed, the only variable elements in $\mathbf{A}$ correspond to the coded aperture entries $T^i_{j,\ell}$, which, in turn, determine $\mathbf{H}$. Thus, the goal of this work is to design $T^i_{j,\ell}$ such that the sensing matrix $\mathbf{A}$ is better conditioned in order to improve the attainable reconstruction quality.

3. RESTRICTED ISOMETRY PROPERTY IN MULTI-SHOT CASSI

The restricted isometry property provides guidelines to determine the minimum number of measurements needed for signal reconstruction in CS. In CASSI, the RIP is critical to designing optimal coded aperture ensembles in order to maximize the quality of the reconstructions. In particular, the RIP of the CASSI sensing matrix $\mathbf{A}$ of order S is defined as the smallest constant δ_S such that $(1 - \delta_S)\|\boldsymbol{\theta}\|_2^2 \leq \|\mathbf{A}\boldsymbol{\theta}\|_2^2 \leq (1 + \delta_S)\|\boldsymbol{\theta}\|_2^2$ holds for all S-sparse vectors $\boldsymbol{\theta}$. The RIP constant δ_S is given by

$$\delta_S = \max_{\mathcal{T} \subset [n], |\mathcal{T}| \leq S} \|\mathbf{A}_{|\mathcal{T}|}^T\mathbf{A}_{|\mathcal{T}|} - \mathbf{I}\|_2^2, \tag{6}$$

where the operator $\|\cdot\|_2^2$ is the squared norm from ℓ_2 into ℓ_2, $\mathbf{A}_{|\mathcal{T}|}$ is a $m \times |\mathcal{T}|$ matrix whose columns are $|\mathcal{T}|$ columns of the CASSI matrix $\mathbf{A}$ indexed by the set Ω, and $\mathbf{I}$ is an identity matrix [11,27]. In other words, the RIP determines whether all column submatrices $\mathbf{A}_{|\mathcal{T}|}$ of $\mathbf{A}$ are well conditioned. The RIP constant in Eq. (6) can be rewritten in terms of the eigenvalues of $\mathbf{A}_{|\mathcal{T}||\mathcal{T}|} = \mathbf{A}_{|\mathcal{T}|}^T\mathbf{A}_{|\mathcal{T}|}$ as

$$\delta_S = \max_{|\mathcal{T}| \subset [n], |\mathcal{T}| \leq S} \sqrt{\lambda_{\max}(\mathbf{A}_{|\mathcal{T}||\mathcal{T}|} - \mathbf{I})}, \tag{7}$$

where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue [11]. Given that the sensing matrix $\mathbf{A}$ is determined by the matrix $\mathbf{H}$ and the basis representation matrix $\boldsymbol{\Psi}$, the entries of $\mathbf{A}_{|\mathcal{T}|}$ can be expressed as the product of the entries of $\boldsymbol{\Psi}$ and $\mathbf{H}$. Let the columns of $\boldsymbol{\Psi}$ be written as $[\boldsymbol{\psi}_0, \ldots, \boldsymbol{\psi}_{n-1}]$. Using the structure of $\mathbf{H}$ in Eq. (4), the entries of $\mathbf{A}_{|\mathcal{T}|}$ can be expressed as the product of the rows of $\mathbf{H}$ and the columns of $\boldsymbol{\Psi}$ indexed by the set Ω:

$$(\mathbf{A}_{|\mathcal{T}|})_{j,k} = \mathbf{h}_j\boldsymbol{\psi}_{\Omega_k} = \sum_{r=0}^{L-1} T^i_{\lfloor \frac{j-rN}{N} \rfloor, j-rN-\lfloor \frac{j-rN}{N} \rfloor N}\psi_{j+rN',\Omega_k}, \tag{8}$$

for $j = 0, \ldots, m - 1$, $k = 0, \ldots, |\mathcal{T}| - 1$, where $i = \lfloor j/NV \rfloor$, $N' = N^2 - N$, and $\Omega_k \in \{0, \ldots, n - 1\}$.

Given that the entries of $\mathbf{A}_{|\mathcal{T}||\mathcal{T}|}$ are given by the product of two columns of $\mathbf{A}_{|\mathcal{T}|}$, Eq. (8) can be used to express the elements of $\mathbf{A}_{|\mathcal{T}||\mathcal{T}|}$ as

$$(\mathbf{A}_{|\mathcal{T}||\mathcal{T}|})_{j,k} = \sum_{i=0}^{K-1}\sum_{\ell_1=0}^{N-1}\sum_{\ell_2=0}^{V-1}\sum_{r=0}^{L-1}\sum_{u=0}^{L-1} T^i_{\ell_1,\ell_2-r}T^i_{\ell_1,\ell_2-u} \times \psi_{\ell_2 N+\ell_1+rN',\Omega_j}\psi_{\ell_2 N+\ell_1+uN',\Omega_k}, \tag{9}$$

for $j, k = 0, \ldots, |\mathcal{T}| - 1$.

A necessary condition for satisfying the RIP is that the expected value of the diagonal terms of $\mathbf{A}_{|\mathcal{T}||\mathcal{T}|}$ should be equal to one, so $E((\mathbf{A}_{|\mathcal{T}||\mathcal{T}|})_{s,s}) = 1$ for $s = 0, \cdots, |\mathcal{T}| - 1$. This condition holds when the coded-aperture entries are constrained to satisfy $\Sigma_{i=0}^{K-1}(T^i_{\ell_1,\ell_2-r})^2 = C$, with C a selectable constant. Hence, $\mathbf{A}_{|\mathcal{T}||\mathcal{T}|}$ can be normalized by defining the matrix $\mathbf{B}_{|\mathcal{T}||\mathcal{T}|} = \mathbf{A}_{|\mathcal{T}||\mathcal{T}|}/C$. Using Eq. (9), the normalized matrix can be written as

$$(\mathbf{B}_{|\mathcal{T}||\mathcal{T}|})_{j,k} = \sum_{\ell_1=0}^{N-1}\sum_{\ell_2=0}^{V-1}\sum_{r=0}^{L-1}\sum_{u=0}^{L-1} \gamma_{\ell_1\ell_2 ru}\phi_{\ell_1\ell_2 ru}/C, \tag{10}$$

where $\gamma_{\ell_1\ell_2 ru} = \Sigma_{i=0}^{K-1} T^i_{\ell_1,\ell_2-r}T^i_{\ell_1,\ell_2-u}$ and $\phi_{\ell_1\ell_2 ru} = \psi_{\ell_2 N+\ell_1+rN',\Omega_j}\psi_{\ell_2 N+\ell_1+uN',\Omega_k}$.

Equation (10) can be analyzed from two perspectives: the first one assumes that $\phi_{\ell_1\ell_2 ru}$ is a constant in order to study the effect of the term $\gamma_{\ell_1\ell_2 ru}$; the second perspective studies the effect of the coded apertures on the coherence between $\mathbf{H}$ and $\boldsymbol{\Psi}$. Given that the elements of $\boldsymbol{\Psi}$ are fixed and bounded, the sparse representation basis term $\phi_{\ell_1\ell_2 ru}$ is also bounded. More specifically, $|\phi_{\ell_1\ell_2 ru}| < C_1$, for all ℓ_1, ℓ_2, r, and u. Thus, Eq. (10) can be rewritten as $(\mathbf{B}_{|\mathcal{T}||\mathcal{T}|})_{j,k} \leq C_1/C\Sigma_{\ell_1=0}^{N-1}\Sigma_{\ell_2=0}^{V-1}\Sigma_{r=0}^{L-1}\Sigma_{u=0}^{L-1}\gamma_{\ell_1\ell_2 ru}$. In consequence, the elements $(\mathbf{B}_{|\mathcal{T}||\mathcal{T}|})_{j,k}$ are the sum of bounded random variables and they can be modeled as a sub-Gaussian random variable $(\mathbf{B}_{|\mathcal{T}||\mathcal{T}|})_{j,k} \sim \mathrm{Sub}(\alpha^2)$, with parameter given by

$$\alpha = \max_{j,k} \frac{C_1}{C}\sum_{\ell_1=0}^{N-1}\sum_{\ell_2=0}^{V-1}\sum_{r=0}^{L-1}\sum_{u=0}^{L-1}\gamma_{\ell_1\ell_2 ru}. \tag{11}$$

Previous works on the RIP for sub-Gaussian random variables have established that $\mathbb{P}(\|\mathbf{B}_{|\mathcal{T}||\mathcal{T}|} - \mathbf{I}\| \leq \delta_S) \geq 1 - \varepsilon$, where $\varepsilon = 2(1 + 2/\rho)^2 e^{-\delta_S(2-(1+\rho)^2)^2 KVc_2/\alpha}$, with $\rho = 2/(e^3 - 1)$ and c_2 a constant independent of K and V [28,29]. Therefore, the error probability ε is minimized either when the number of measurement shots K is increased, or when the coded-aperture ensemble is designed such that the sub-Gaussian parameter α is minimized. It can be noted in Eq. (11) that the minimization of α, based on the design of the coded apertures, is directly related to the minimization of the variable $\gamma_{\ell_1\ell_2 ru} = \Sigma_{i=0}^{K-1} T^i_{\ell_1,\ell_2-r}T^i_{\ell_1,\ell_2-u}$, which depends on the horizontal separation of the one-valued entries of $\mathbf{T}^i$ and their correlation across shots.

Equivalent to the RIP, the notion of coherence provides an alternative approach for evaluating the properties of $\mathbf{A}$ that guarantee the correct recovery of the underlying signal. In particular, CS requires the acquired measurements to be uncorrelated, and the coherence is used to measure the correlation between $\mathbf{H}$ and $\boldsymbol{\Psi}$, such that the CS conditions hold. Recently, the coherence has been defined as the largest absolute

inner product between any two columns of the sensing matrix [30]. More specifically, the coherence of the CASSI sensing matrix can be calculated as

$$\mu \triangleq \max_{i \neq j} |\mathrm{TPSF}(i,j)| / \sqrt{\mathrm{TPSF}(i,i)\mathrm{TPSF}(j,j)}, \quad (12)$$

where $\mathrm{TPSF}(i,j) = (\mathbf{\Psi}^T \mathbf{H}^T \mathbf{H} \mathbf{\Psi})_{i,j}$ is a transform point-spread function, as defined in [31,32]. In other words, the coherence is calculated as the maximum off-diagonal entry of the TPSF matrix. Given that low correlation values are desired, one can define the incoherence as $\tilde{\mu} = 1 - \mu$ for simplicity in evaluating low correlation values between $\mathbf{H}$ and $\mathbf{\Psi}$.

4. CODED-APERTURE DESIGN BASED ON RIP

The results from the previous section provide useful guidelines for the design of coded-aperture ensembles that help satisfy the RIP and incoherence conditions so better sensing matrices can be designed in order to improve image reconstructions. In particular, the following design criteria are taken into account to generate optimized coded-aperture ensembles:

(a) *Horizontal separation:* The minimization of the sub-Gaussian parameter α in Eq. (11) is determined by the term $\gamma_{\ell_1 \ell_2 r u}$. More specifically, given that $r, u \in \{0, \ldots, L-1\}$, the products $T^i_{\ell_1,\ell_2-r} T^i_{\ell_1,\ell_2-u}$ should be minimized within a horizontal neighborhood of size L of the coded aperture $\mathbf{T}^i$. It can be noted that this minimization is achieved when the one-valued entries of $\mathbf{T}^i$ within the same row are maximally separated. This condition, in essence, requires the codes to satisfy spatial blue noise pixel distribution.

To illustrate the effect of this condition, different coded-aperture ensembles are employed to determine $\mathbf{H}$ and, α from Eq. (11) is evaluated for the corresponding matrices with a fixed $\mathbf{\Psi}$. Given that Kronecker product bases have shown good sparse representations of spectral images, a Kronecker product basis $\mathbf{\Psi} = \mathbf{\Psi}^{2D} \otimes \mathbf{\Psi}^{C}$ is used, where $\mathbf{\Psi}^{2D}$ is the $N^2 \times N^2$ matrix of the 2D wavelet transform that is applied to each spectral band, and $\mathbf{\Psi}^{C}$ is the $L \times L$ matrix of the cosine transform that looks for the sparse representation across the spectral coordinates. The value of α is evaluated for the traditionally employed random coded apertures with 50% transmittance, and it is compared with coded-aperture ensembles intentionally designed to exhibit horizontal separation of one-valued entries for $K = 4$. Figure 3 shows one realization of these coded apertures and their corresponding average values of α. It can be noted that random coded apertures result in larger values of α.

(b) *Vertical separation:* The presence of clusters of one-valued elements in the vertical direction of the coded-aperture ensemble is related to an increase of the incoherence related to Eq. (12). This fact is illustrated in Fig. 4, where a random coded aperture is compared with a coded aperture intentionally designed to exhibit vertical separation of one-valued entries. It can be noted that the random coded aperture results in a lower incoherence value. Furthermore, given that clusters of one-valued pixels prevent uniform sensing of the scene, one of the design criteria of an optimal coded-aperture ensemble is to reduce these vertical cluster occurrences.

(c) *Temporal correlation:* Given that each measurement snapshot employs a different coded aperture, low correlation across shots is desired. An approach to reduce this correlation consists of constraining the set of codes to be complementary. In practical terms, this is equivalent to having just one coded aperture pixel set to one at each particular spatial position of the ensemble. This guarantees that a sensed voxel of the scene is captured just once. This property corresponds to temporal blue noise characteristics.

Based on the previous design criteria, the coded-aperture optimization problem is mathematically described below. First, it is easy to see that small regions of the ensemble can be designed such that the optimization criteria are satisfied; hence, the whole ensemble will fulfill them as well. Thus, let U^i_p be a $\Delta \times \Delta$ window of $\mathbf{T}^i$ centered at a specific point $P = (j, \ell)$, defined as $U^i_P = \{T^i_{\ell_1,\ell_2} | \ell_1 \in [x - \lfloor \Delta/2 \rfloor, x + \lfloor \Delta/2 \rfloor], \ell_2 \in [y - \lfloor \Delta/2 \rfloor, y + \lfloor \Delta/2 \rfloor]\}$, for $i = 0, \ldots, K-1$. Given that $j = 0, \ldots, N-1$, and $\ell = 0, \ldots, N-1$, $P = 0, \ldots, N^2 - 1$. In order to reduce the concentration of one-valued terms within U^i_p, let

$$d^i_{P1} = \{T^i_{\ell_1,\ell_2} | \ell_1 \in [x - \lfloor \Delta/2 \rfloor, x + \lfloor \Delta/2 \rfloor], \ell_2 = y\} \quad (13)$$

be the subset of horizontal neighboring pixels within U^i_p around P. Similarly define

$$d^i_{P2} = \{T^i_{\ell_1,\ell_2} | \ell_1 = x, \ell_2 \in [y - \lfloor \Delta/2 \rfloor, y + \lfloor \Delta/2 \rfloor]\} \quad (14)$$

as the subset of vertical neighboring pixels in U^i_p around P. Finally, let the subsets of diagonal neighbors of P in U^i_p be written as

$$d^i_{P3} = \{T^i_{\ell_1,\ell_2} | \ell_1 \in [x - \lfloor \Delta/2 \rfloor, x + \lfloor \Delta/2 \rfloor], \ell_2 \in [y - \lfloor \Delta/2 \rfloor, y + \lfloor \Delta/2 \rfloor], \ell_1 + \ell_2 = \Delta + 1\}, \quad (15)$$

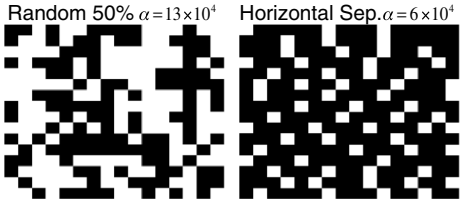

Fig. 3. Comparison for different coded aperture ensembles with $K = 4$ in terms of parameter α. (Left) Random coded aperture entries are realizations of a Bernoulli random variable with $p = 50\%$. (Right) Coded aperture intentionally designed to exhibit horizontal separation of one-valued features.

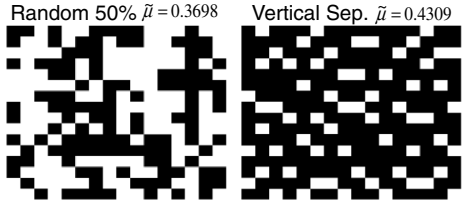

Fig. 4. Incoherence comparison for different coded aperture ensembles with $K = 4$. (Left) Random coded aperture entries are realizations of a Bernoulli random variable with $p = 50\%$. (Right) Coded aperture intentionally designed to exhibit vertical separation of one-valued features.

$$d^i_{P4} = \{T^i_{\ell_1,\ell_2} | \ell_1, \ell_2 \in [x - \lfloor \Delta/2 \rfloor, x + \lfloor \Delta/2 \rfloor], \ell_1 = \ell_2\}. \tag{16}$$

Figure 5 shows a graphic representation of the definitions from Eqs. (13)–(16). These subset definitions are then used to define a weighted local metric S^i_P given by

$$S^i_P = \sum_{k=1}^{4} w_k \left\| d^i_{Pk} \right\|_1, \tag{17}$$

which measures the concentration of one-valued elements in U^i_P for $i = 0, \ldots, K-1$. More specifically, this metric penalizes the appearance of clusters of ones according to the weights w_k. Thus, if horizontal and vertical clusters are not desirable, w_1 and w_2 should be assigned greater values. Equivalently, w_3 and w_4 should be greater than w_1, w_2 if the distance between one-valued elements in the diagonals is to be increased.

Hence, an optimal coded-aperture ensemble can be obtained by minimizing the variance of S^i_P, given that all K coded apertures in the ensemble are intended to exhibit the same transmittance value. The optimal set of coded apertures for compressive spectral imaging can thus be expressed as the solution of the problem

$$\underset{\{T^0, T^1, \ldots, T^{K-1}\}}{\operatorname{argmin}} \operatorname{var}(S^i_P) \quad \text{subject to} \quad \sum_{i=0}^{K-1} T^i_{\ell_1,\ell_2} = 1, \tag{18}$$

where $P = 0, \ldots, N^2 - 1$. To generate the set of optimal coded apertures, an algorithm that finds a solution to Eq. (18) has been developed and is illustrated in Fig. 6. The required inputs are: the size of the coded-aperture ensemble N, the number of measurement shots K, the window size Δ, the set of weights $\mathbf{w} = \{w_k\}_{k=1}^4$, and a set of randomly ordered spatial coordinates $\Omega = (\Omega_r, \Omega_c)$. Notice that Ω represents a permuted version of the spatial positions indexed by P, such that the procedure is performed in a random raster fashion. Also, an initial guess of the coded aperture is required.

The proposed algorithm moves over the pairs (Ω_r, Ω_c) for $r, c = 0, \ldots, N-1$, and for each coded aperture of the ensemble, i.e., $i = 0, \ldots, K-1$, creates a $\Delta \times \Delta$ window, U^i_P, centered at $T^i_{\Omega_r,\Omega_c}$, as illustrated in Fig. 5. On each window,

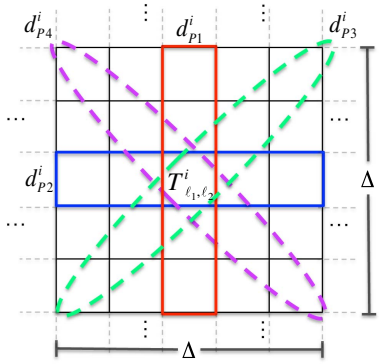

Fig. 5. Graphic representation of the metric used to determine the concentration of one-valued entries in a window U^i_P of the code. This metric helps with deciding the best i for setting the one-valued element.

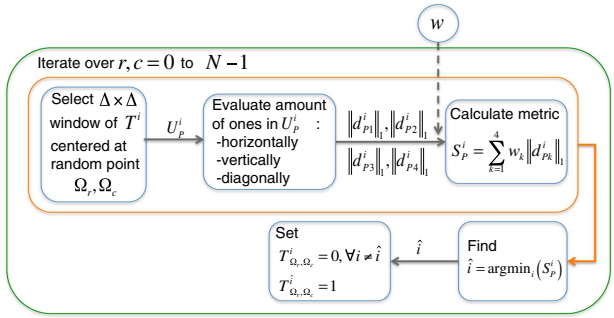

Fig. 6. Block diagram of the algorithm that generates the designed coded apertures. This procedure iterates over all (Ω_r, Ω_c) points and determines the best coded aperture of the ensemble in which the (Ω_r, Ω_c)th pixel is set to one.

the algorithm calculates the concentration of one-valued entries as $\left\| d^i_{P1} \right\|_1$, $\left\| d^i_{P2} \right\|_1$, $\left\| d^i_{P3} \right\|_1$, and $\left\| d^i_{P4} \right\|_1$, which are then used to calculate S^i_P using Eq. (17). Hence, the criterion to decide the best coded apertures of the ensemble, i.e., the value of i, to insert the one-valued element for the specific spatial position (Ω_r, Ω_c) given by $\hat{i} = \arg \min_i (S^i_P)$, which in other words refers to the coded aperture with the minimum concentration of ones in the window around (Ω_r, Ω_c). Finally, the algorithm sets the corresponding values of the coded aperture ensemble in the particular point $P = (\Omega_r, \Omega_c)$ as

$$T^i_{\Omega_r,\Omega_c} = 0 \quad \forall\, i \neq \hat{i}; \;\; T^{\hat{i}}_{\Omega_r,\Omega_c} = 1. \tag{19}$$

Figure 7 illustrates an example of the operations performed by the proposed algorithm. More specifically, an 8×8 coded aperture ensemble for $K = 4$ snapshots is used for this example. It can be noted that the initial guess has clusters of one-valued elements and, at the end of the execution of the algorithm, the clusters of ones in the resulting coded-aperture ensemble are considerably minimized.

The resulting coded-aperture ensemble is coined spatiotemporal blue noise. A single realization of the BN coded apertures obtained with the proposed algorithm for $K = 4$ is illustrated in Fig. 8 (top) along with the average incoherence and RIP parameters, $\tilde{\mu}$ and α. Another approach to obtain BN coded

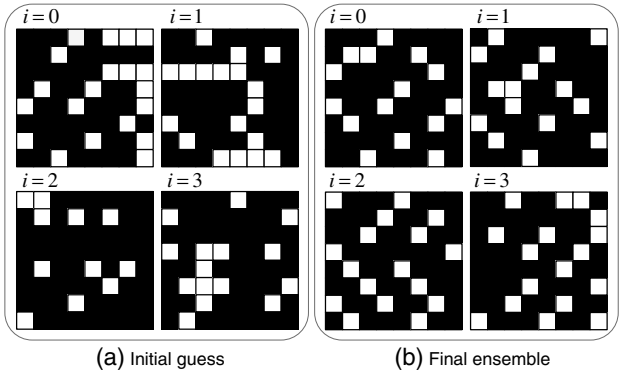

Fig. 7. Example of the operations performed by the proposed algorithm. (a) Initial guess of an 8×8 coded aperture ensemble for $K = 4$ snapshots. (b) Resulting ensemble after applying the proposed algorithm.

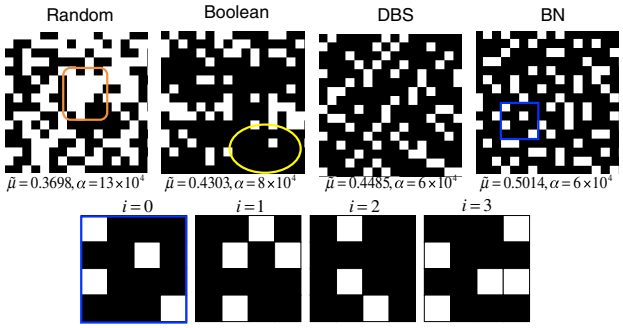

Fig. 8. (Top) Comparison of one of the $K = 4$ realizations of the BN coded apertures with the traditional random, Boolean, and DBS coded apertures. (Bottom) Zoomed-in portion of a BN coded aperture ensemble for $K = 4$ measurement snapshots to illustrate the low correlation across shots.

apertures is the direct binary search (DBS) algorithm, which has been originally used in lithography to optimally represent continuous-tone images [33,24]. In particular, blue noise patterns suppress the low-frequency components of noise [25,26], and the characteristics of the patterns generated with the DBS algorithm correspond to the desired design criteria of the proposed BN coded apertures, as shown in Fig. 8 (top). For comparison purposes, Fig. 8 (top) presents one realization of the traditional random coded apertures and Boolean coded apertures, which exhibit a spatially random distribution but exploit the temporal correlation restriction of BN patterns. Notice that the clusters of one-valued entries of the random coded aperture are highlighted in the orange rounded square. In addition, given that the spatial distribution of the Boolean coded apertures is random, clusters of zero-valued elements are likely to appear, as shown in the yellow circle, as well as clusters of ones. The DBS and the BN coded apertures, on the other hand, have a very similar structure. The corresponding average values of the incoherence and RIP parameters, $\bar{\mu}$ and α, respectively, are also indicated for each case. As expected, BN coded apertures exhibit larger incoherence and lower RIP constant values. Furthermore, Fig. 8 (bottom) illustrates the 4×4 blue highlighted portion of the BN coded aperture ensemble for $K = 4$ snapshots, where the temporal correlation restriction is satisfied.

5. SIMULATION AND EXPERIMENTAL RESULTS

The performance of the proposed coded aperture design was tested through simulations and experiments. First, the CASSI acquisition process is simulated. In addition, a testbed implementation of the CASSI system is used for capturing experimental measurements. Reconstruction results and an analysis are also presented for each case.

A. Simulations

The compressive spectral imaging acquisition was simulated using Eq. (5) for three different spectral data cubes with 256×256 pixels of spatial resolution and 8 spectral bands with central wavelengths of 450, 456, 484, 503, 524, 549, 580, and 621 nm. The first data cube was acquired by illuminating the scene with a monochromatic light source. The other data cubes are real scenes illustrating a set of colorful balloons and a color palette, available in [34] and [35], respectively. The proposed blue noise coded apertures are compared against the traditional random coded apertures with 50% transmittance and the Boolean coded apertures [23]. In addition, the simulations are conducted using blue noise coded apertures obtained with the DBS algorithm [24]. Recall that the Boolean and the DBS coded aperture ensembles are complementary as the BN, resulting in low correlation across snapshots. The number of snapshots is varied in the simulations from $K = 2$ to $K = 8$, which matches the number of spectral bands, that is, the compression ratios from 1:4 to 1:1.

Figure 9 shows the average behavior of the reconstruction quality when different types of coded apertures are used and the number of snapshots is varied. The quality of the reconstructions is measured using two widely known error metrics: the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM), averaged across spectral bands. The PSNR measures the log scale of the inverse of the mean squared error and is defined as $20 \log_{10}(\max_I / \mathrm{MSE}^{1/2})$, where $\max_I$ is the maximum possible value of the image and MSE is the mean squared error with respect to the original image. On the other hand, the SSIM, introduced in [36], compares local patterns of pixel intensities that have been normalized for luminance and contrast. It can be noted in Fig. 9 that the random coded apertures exhibit the poorest reconstructions, while the coded apertures with low correlation across snapshots, i.e., the Boolean, DBS, and BN, yield improved reconstruction quality. In addition, the DBS and BN coded apertures provide the best PSNR and SSIM results. More specifically, the proposed BN coded apertures outperform the performance of the random

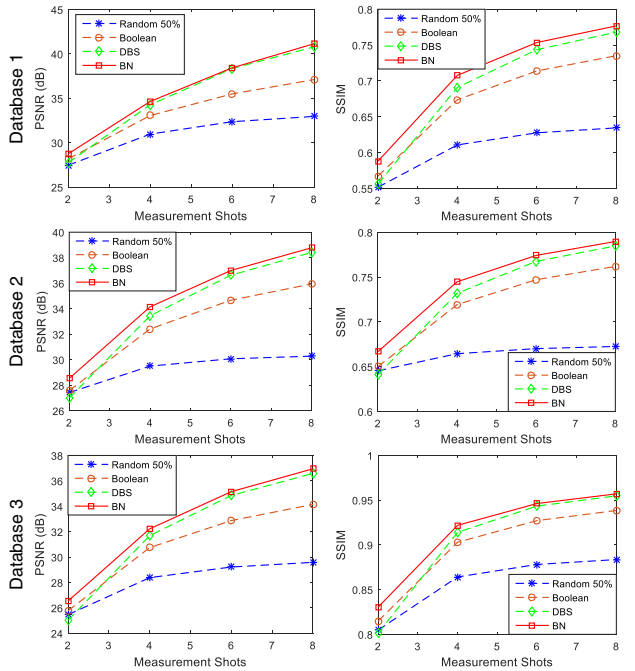

Fig. 9. Average reconstruction quality measured by the PSNR and SSIM as a function of the number of captured snapshots. Results for three different data bases are illustrated using the random, Boolean, DBS, and BN coded apertures.

coded apertures in up to 9 dB of PSNR, and the DBS coded apertures in up to 1 dB. Similarly, for the SSIM measure, the BN coded apertures present improvements of up to 0.14 and 0.04 with respect to the random and DBS coded apertures, respectively. In general, these results indicate that the restrictions on the horizontal, vertical, and diagonal separations of the one-valued entries are determinant to obtaining improved reconstructions.

Figure 10 shows an RGB mapping of the recovered spectral data cubes for the different types of coded apertures, compared to the original scene. In particular, these reconstructions are obtained from $K = 4$ measurement snapshots, which corresponds to 50% of the data. In addition, Fig. 11 illustrates a zoomed-in portion of the reconstructions in Fig. 10, in which the differences between the recovered images are noticeable. Specifically, in spite of the good results obtained with the random, Boolean, and DBS coded apertures, the proposed BN ensemble provides more detailed and accurate reconstructions. Moreover, to illustrate the spectral accuracy of the BN coded apertures, the spectral signatures for two different points of each scene are illustrated in Fig. 12. Additional simulations with a fourth spectral data set captured with an AVIRIS sensor [37] were performed to test the BN coded apertures in these types of remotely sensed images. This data set is a 256 × 256

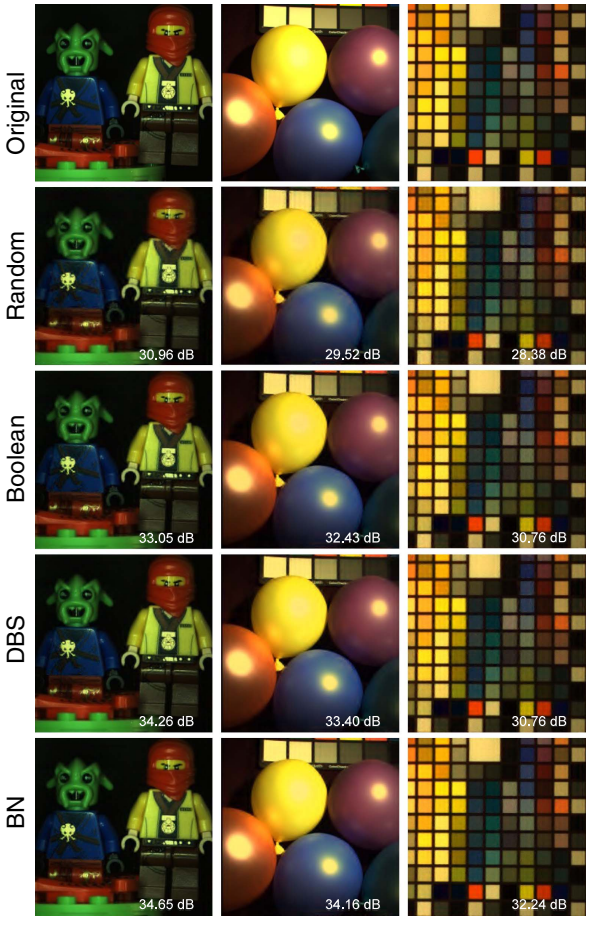

Fig. 10. RGB profiles of the reconstructed images obtained using the random, Boolean, DBS, and BN coded aperture ensembles, and $K = 4$ measurement snapshots, compared to the original data cubes.

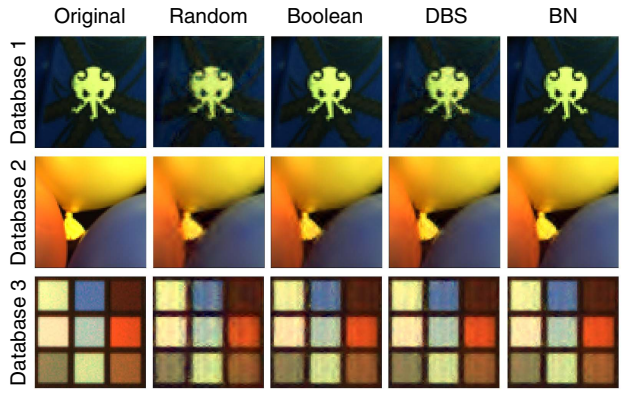

Fig. 11. Comparison of zoomed-in portions of the RGB reconstructions in Fig. 10.

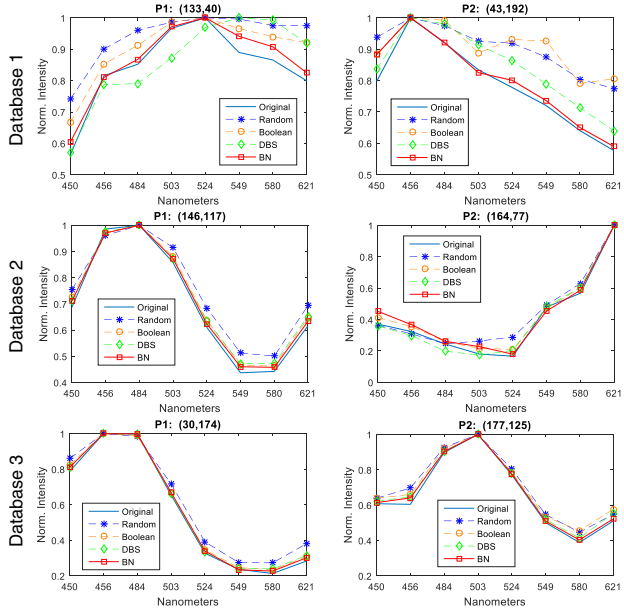

Fig. 12. Spectral signature comparison for two spatial points on each data base. Reconstructions obtained from $K = 4$ snapshots.

portion of the aerial view of Moffett Field with $L = 32$ spectral bands. Figure 13 illustrates the average reconstruction PSNR and SSIM for this data base as a function of the number of snapshots. It can be noted that these results exhibit a similar behavior to those in Fig. 9.

Even though in most applications, the coded apertures are usually generated offline, the computation time required to generate a coded aperture ensemble is an additional comparison factor that can be taken into account. The performed simulations show that the DBS algorithm is up to 5 times slower than the proposed algorithm for generating coded apertures with blue noise characteristics. Another metric to compare different types of coded apertures is the radially average power spectrum density (RAPSD) [24], which is used to identify spectral characteristics of lithographic patterns. Given that the designed coded apertures have blue noise characteristics, RAPSD can be used to verify this behavior. In essence, RAPSD relies on

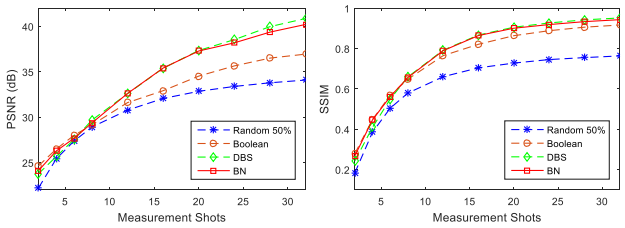

Fig. 13. Average reconstruction PSNR and SSIM for the Moffett Field data set as a function of the number of snapshots.

estimating the magnitude square of the Fourier transform of the output ensemble to produce a spectral estimate with Bartlett's method of averaging periodograms. Thus, the RAPSD is obtained by partitioning the spectral domain into a series of annular rings, and it is calculated as the average power in the annular ring for each center radius [25,26]. Figure 14 illustrates the RAPSD for random, Boolean, DBS, and BN coded aperture ensembles, in which the random and Boolean designs present a flat behavior. Meanwhile, the high-frequency characteristics of DBS and BN coded apertures are clearly noticeable.

B. Testbed Implementation

The optical setup of the CASSI system illustrated in Fig. 15 was constructed to experimentally verify the performance of the proposed BN coded apertures. This setup comprises an objective lens, a custom Amici prism, a relay lens, a monochrome CCD, a photomask wafer, and an x–y linear stage. In particular, the wafer is a $17 \times 17 \times 0.015''$ chrome-on-quartz photomask that consists of several 128×128 coded aperture realizations

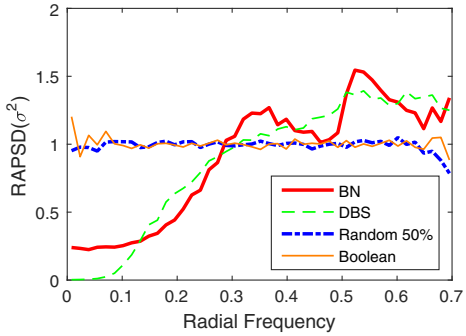

Fig. 14. Comparison of the radially averaged power spectral density for the random, Boolean, DBS, and BN coded apertures.

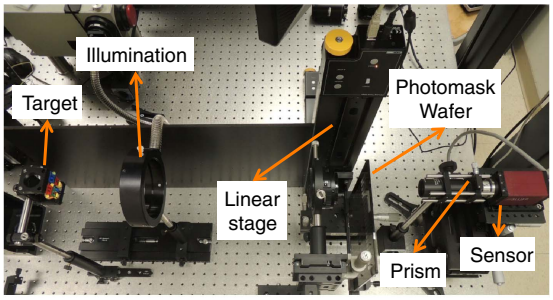

Fig. 15. Testbed implementation of the CASSI system.

with a pixel pitch of 19.35 μm. The linear stage moves the photomask accordingly in order to locate and align the correspondent pattern to be used on each snapshot. The nonlinear Amici prism disperses the visible incoming radiation into 10 spectral bands with the following central wavelengths: 448, 460, 473, 489, 507, 526, 550, 579, 611, and 653 nm. Compressive measurements for $K = 2, 4, 6$ measurement snapshots were captured using random, Boolean, DBS, and BN coded apertures. These sets of measurements and the corresponding calibration data cubes are used to estimate the original scene.

Figure 16 presents a comparison of an RGB map of the recovered scenes using the random, Boolean, DBS, and BN coded apertures obtained from $K = 2, 4$, and 6 measurement snapshots. It can be noted that increasing the number of snapshots results in improved and more detailed reconstructions.

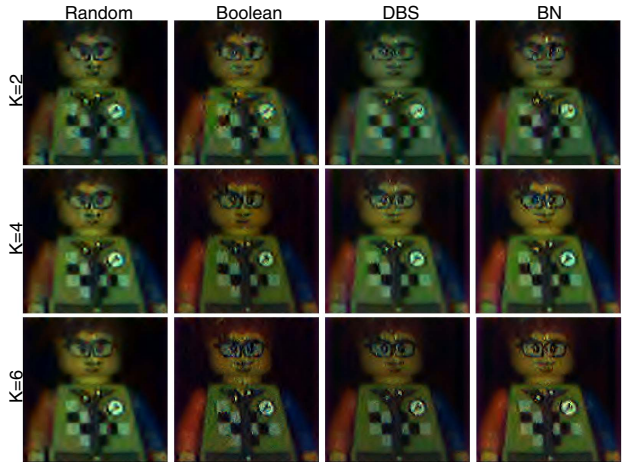

Fig. 16. RGB representations of the experimental reconstructions from $K = 2, 4, 6$ for different types of coded apertures.

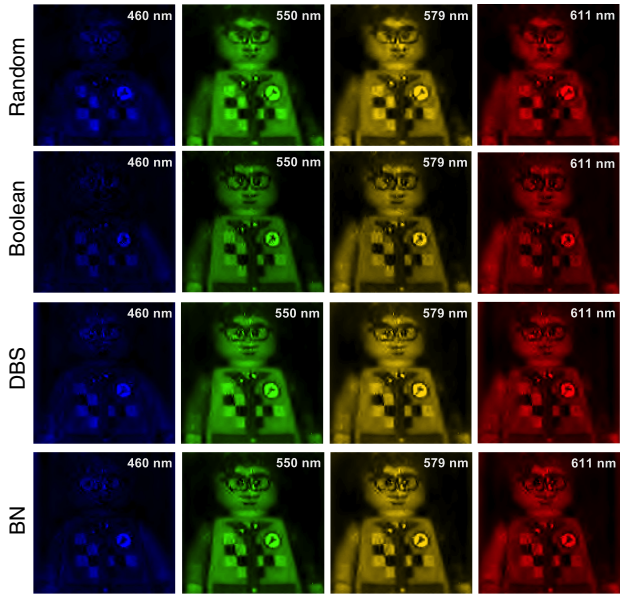

Fig. 17. Selected reconstructed spectral bands for the random, Boolean, DBS, and BN coded apertures with $K = 4$.

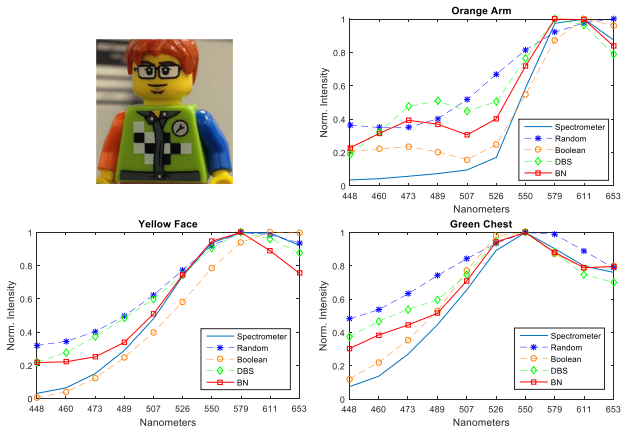

Fig. 18. Spectral signature reconstructions of three points of the scene, using $K = 4$ snapshot measurements. (Top right) Orange arm, (bottom left) yellow face, and (bottom right) green chest.

Furthermore, the BN and DBS coded apertures yield better reconstructions than the random and Boolean coded apertures.

To better visualize the improvements in the reconstructions, Fig. 17 presents four selected recovered spectral bands, from $K = 4$ measurement snapshots, for the random, Boolean, DBS, and BN coded apertures. Finally, Fig. 18 compares the spectral signatures of three different points of the scene with the original spectrum measured with a point spectrometer. This comparison shows that in spite of the good approximation of the spectra obtained with the random and Boolean coded apertures, the DBS and BN designs provide a better approximation of the bands of interest.

6. CONCLUSIONS

Spatiotemporal blue noise coded apertures for compressive spectral imaging have been introduced. This design is based on the restricted isometry property of the CASSI sensing matrix, as well as the incoherence with a given sparsifying basis. The mathematical development of the proposed design and a computational algorithm to generate BN coded apertures have been also presented. The simulations and experiments show that the proposed design results in improved reconstructions with respect to the traditional random coded apertures, while similar quality can be attained with blue noise coded apertures generated with the DBS algorithm. In particular, gains of up to 9 dB of PSNR and 0.15 in SSIM are obtained with respect to the random ensembles. However, the proposed algorithm generates BN coded apertures up to 5 times faster than the DBS algorithm. It is worth noting that although BN coded apertures were tested for the CASSI system, the concept can be extended to other CSI architectures, such as those in [5], and applications like remote sensing.

Funding. National Science Foundation (NSF)/Intel Partnership on Visual and Experiential Computing (VEC) (1538950); Vicerrectoría de Investigación y Extensión of the Universidad Industrial de Santander, Colombia (1890); Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS).

Acknowledgment. Claudia V. Correa is supported by a Colciencias scholarship.

REFERENCES

1. P. S. Thenkabail, J. G. Lyon, and A. Huete, *Hyperspectral Remote Sensing of Vegetation* (CRC Press, 2016).
2. D. Liu, X. Zeng, and D. Sun, "Recent developments and applications of hyperspectral imaging for quality evaluation of agricultural products: a review," CRC Crit. Rev. Food Sci. Nutr. **55**, 1744–1757 (2015).
3. C. Chang, "Hyperspectral target detection," in *Real-Time Progressive Hyperspectral Image Processing: Endmember Finding and Anomaly Detection*, C. DeCusatis, D. Clement, E. Maass, and R. Lasky, eds. (Springer, 2016), pp. 131–172.
4. G. Lu and B. Fei, "Medical hyperspectral imaging: a review," J. Biomed. Opt. **19**, 010901 (2014).
5. X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational snapshot multispectral cameras: toward dynamic capture of the spectral world," IEEE Signal Process. Mag. **33**(5), 95–108 (2016).
6. A. A. Wagadarikar, R. John, R. Willett, and D. J. Brady, "Single disperser design for coded aperture snapshot spectral imaging," Appl. Opt. **47**, B44–B51 (2008).
7. A. A. Wagadarikar, "Compressive spectral and coherence imaging," Ph.D. thesis (Duke University, 2010).
8. A. Wagadarikar, N. Pitsianis, X. Sun, and D. Brady, "Video rate spectral imaging using a coded aperture snapshot spectral imager," Opt. Express **17**, 6368–6388 (2009).
9. D. Kittle, K. Choi, A. A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," Appl. Opt. **49**, 6824–6833 (2010).
10. Y. Wu, I. O. Mirza, G. R. Arce, and D. W. Prather, "Development of a digital-micromirror-device-based multishot snapshot spectral imaging system," Opt. Lett. **36**, 2692–2694 (2011).
11. H. Rauhut, "Compressive sensing and structured random matrices," in *Theoretical Foundations and Numerical Methods for Sparse Recovery*, Radon Series on Computational and Applied Mathematics (deGruyter, 2010), Vol. **9**, pp. 1–92.
12. M. Ledoux, *The Concentration of Measure Phenomenon* (American Mathematical Society, 2001), Vol. **89**.
13. J. Haupt, W. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," IEEE Trans. Inf. Theory **56**, 5862–5875 (2010).
14. J. L. Paredes, G. R. Arce, and Z. Wang, "Ultra-wideband compressed sensing: channel estimation," IEEE J. Sel. Top. Signal Process. **1**, 383–395 (2007).
15. M. F. Duarte and Y. Eldar, "Structured compressed sensing: from theory to applications," IEEE Trans. Signal Process. **59**, 4053–4085 (2011).
16. J. Y. Park, H. L. Yap, C. J. Rozell, and M. B. Wakin, "Concentration of measure for block diagonal matrices with applications to compressive signal processing," IEEE Trans. Signal Process. **59**, 5859–5875 (2011).
17. T. T. Do, L. Gan, N. Nguyen, and T. Tran, "Fast and efficient compressive sensing using structurally random matrices," IEEE Trans. Signal Process. **60**, 139–154 (2012).
18. J. Romberg, "Compressive sensing by random convolution," SIAM J. Imag. Sci. **2**, 1098–1128 (2009).
19. D. J. Brady, *Optical Imaging and Spectroscopy* (Wiley, 2009).
20. R. M. Willett, M. E. Gehm, and D. J. Brady, "Multiscale reconstruction for computational spectral imaging," Proc. SPIE **6498**, 64980L (2007).
21. H. Arguello and G. R. Arce, "Code aperture optimization for spectrally agile compressive imaging," J. Opt. Soc. Am. A **28**, 2400–2413 (2011).
22. N. J. A. Sloane and M. Harwit, "Mask for Hadamard transform optics, and weighing designs," Appl. Opt. **15**, 107–114 (1976).
23. 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).
24. D. Lau and G. R. Arce, *Modern Digital Halftoning* (CRC Press, 2008).

25. D. L. Lau, G. R. Arce, and N. C. Gallagher, "Green-noise digital halftoning," Proc. IEEE **86**, 2424–2444 (1998).
26. D. L. Lau, R. Ulichney, and G. R. Arce, "Blue and green noise halftoning models," IEEE Signal Process. Mag. **20**(4), 28–38 (2003).
27. E. Candes and T. Tao, "Decoding by linear programming," IEEE Trans. Inf. Theory **51**, 4203–4215 (2005).
28. U. Ayaz and H. Rauhut, "Nonuniform sparse recovery with sub-Gaussian matrices," Electron. Trans. Numer. Anal. **32**, 242–254 (2012).
29. H. Arguello and G. R. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging," IEEE Trans. Image Process. **23**, 1896–1908 (2014).
30. Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications* (Cambridge University, 2012).
31. Y. Kaganovsky, D. Li, A. Holmgren, H. Jeon, K. P. MacCabe, D. G. Politte, J. A. O'sullivan, L. Carin, and D. J. Brady, "Compressed sampling strategies for tomography," J. Opt. Soc. Am. A **31**, 1369–1394 (2014).
32. D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization," Proc. Natl. Acad. Sci. USA **100**, 2197–2202 (2003).
33. F. A. Baqai and J. P. Allebach, "Halftoning via direct binary search using analytical and stochastic printer models," IEEE Trans. Image Process. **12**, 1–15 (2003).
34. "Multispectral image database," http://www.cs.columbia.edu/CAVE/databases/multispectral/.
35. G. Finlayson, S. Hordley, and P. Morovic, "Using the SpectraCube to build a multispectral image database," in *European Conference on Graphics, Imaging and Vision*, Germany, 2004.
36. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).
37. "Airborne visible infrared imaging spectrometer," http://aviris.jpl.nasa.gov/data/free_data.html.