High-Resolution Multi-Spectral Imaging With Diffractive Lenses and Learned Reconstruction

Figen S. Oktem , *Member, IEEE*, Oğuzhan Fatih Kar, Can Deniz Bezek, and Farzad Kamalabadi, *Member, IEEE*

Abstract—Spectral imaging is a fundamental diagnostic technique with widespread application. Conventional spectral imaging approaches have intrinsic limitations on spatial and spectral resolutions due to the physical components they rely on. To overcome these physical limitations, in this paper, we develop a novel multi-spectral imaging modality that enables higher spatial and spectral resolutions. In the developed computational imaging modality, we exploit a diffractive lens, such as a photon sieve, for both dispersing and focusing the optical field, and achieve measurement diversity by changing the focusing behavior of this lens. Because the focal length of a diffractive lens is wavelength-dependent, each measurement is a superposition of differently blurred spectral components. To reconstruct the individual spectral images from these superimposed and blurred measurements, model-based fast reconstruction algorithms are developed with deep and analytical priors using alternating minimization and unrolling. Finally, the effectiveness and performance of the developed technique is illustrated for an application in astrophysical imaging under various observation scenarios in the extreme ultraviolet (EUV) regime. The results demonstrate that the technique provides not only diffraction-limited high spatial resolution, as enabled by diffractive lenses, but also the capability of resolving close-by spectral sources that would not otherwise be possible with the existing techniques. This work enables high resolution multi-spectral imaging with low cost designs for a variety of applications and spectral regimes.

Index Terms—Spectral imaging, diffractive lenses, photon sieves, inverse problems, learned reconstruction.

I. INTRODUCTION

SPECTRAL imaging, also known as multispectral or hyper-spectral imaging, is a fundamental diagnostic technique in the physical sciences with widespread application in physics,
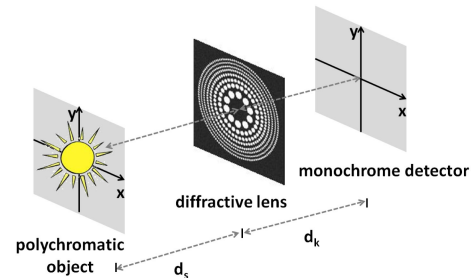
Fig. 1. Schematic view of the proposed diffractive lens based spectral imaging system.

biology, chemistry, medicine, astronomy, and remote sensing. In spectral imaging, images of a scene are formed as a function of wavelength. Capturing this three-dimensional (3D) spectral data cube enables unique identification of the biological, chemical, and physical properties of the scene. As a result, spectral imaging is a useful diagnostic tool in diverse applications including remote sensing of astrophysical plasmas, environmental monitoring, surveillance, biomedical diagnostics, and industrial inspection, among many others.

The inherent challenge in spectral imaging is to capture the 3D spectral cube with the 2D detectors. Conventional approaches perform scanning to form the spectral data cube from a series of 2D measurements. Generally this data cube is measured using an imager with a series of wavelength filters scanning the scene spectrally, or by using a spectrometer with a slit scanning the scene spatially. Since these techniques purely rely on physical components, an important drawback is the intrinsic physical limitations on their spatial, spectral, and temporal resolutions, as well as the inherent trade-off between spatial and spectral resolutions [1]–[3]. For example, for spectral imagers employing wavelength filters, the spectral resolution is strictly limited by the bandwidth of the producible wavelength filters in the desired spectral range, and the spatial resolution is limited by the quality and cost of the used reflective/refractive imaging optics [2], [4], [5].

In this paper, we develop a computational multi-spectral imaging technique with the goal of enabling diffraction-limited high spatial resolution and better spectral resolution than the conventional spectral imagers with wavelength filters. Our imaging technique exploits a diffractive lens for both dispersing and focusing the optical field, allowing us to overcome these limitations in the spatial and spectral resolutions, while not creating any physical trade-off between them. Fig. 1 provides

a schematic of the imaging system involved. This system takes advantage of the wavelength-dependent behavior of a diffractive lens, i.e. the fact that each wavelength is focused differently by a diffractive lens. The main idea is to differently focus each spectral component and acquire measurements that are superpositions of these spectral images (with each spectral component exposed to different type of blur). Hence each measurement is the superposition of differently blurred spectral components. To obtain the needed information for decomposing this multiplexed data, multiple measurements are acquired and jointly used to reconstruct the individual spectral images by solving an inverse problem.

To solve the image reconstruction problem, we first formulate the continuous image-formation model and then obtain its discrete version. The resulting inverse problem can be viewed as a multi-frame deconvolution problem where multiple images need to be deblurred from their superimposed and blurred versions. To solve this inverse problem, we develop two fast reconstruction algorithms. The first algorithm exploits analysis priors (such as total variation) using the alternating direction method of multipliers (ADMM). As the analytical priors have recently been outperformed by deep priors [6]–[8], the second algorithm learns the prior through a deep network that is end-to-end trained with an unrolling approach. Thanks to the shift-invariance behavior of the system and the resulting convolutional forward model, each step in the developed algorithms has fast computation using FFTs.

We illustrate the high spatial and spectral resolutions enabled by the developed spectral imaging technique for a potential application in astrophysical imaging of space plasmas in the extreme ultraviolet (EUV) regime. A detailed conditioning-based analysis is also performed for the spatial resolution of the technique. Although the presented results in this manuscript are for the UV regime, the developed imaging concept is equally applicable to other spectral regimes.

To achieve measurement diversity, multiple measurements can be acquired either with a moving monochrome detector (so measurements are taken from different distances) or with multiple diffractive lens designs (so spectral bands focused differently onto a fixed detector). The latter may require additional optical components such as beam splitters and a larger detector. It can also be performed using a compound diffractive lens design that focus multiple wavelengths onto different locations on the fixed detector or with a variable diffractive lens through a programmable spatial light modulator (SLM) or a digital micromirror device (DMD).

Earlier approaches in computational spectral imaging have been developed with the goal of providing faster acquisition and/or reduced hardware complexity than the conventional methods. These can be regarded as *compressive or snapshot* methods, and each use a different encoding scheme to acquire reduced number of measurements. Examples are coded aperture snapshot spectral imaging (CASSI) [9] and its variants [10]–[12], and coded aperture spectral imaging with diffractive lenses (CSID) [13], [14]. However, in none of these compressive/snapshot approaches, the main goal was to improve the

spatial and spectral resolutions of the conventional scanning-based methods beyond their physical limitations. To the best of our knowledge, there have been no reports of higher spatial and spectral resolutions than the conventional systems, and ours is the first work that demonstrates this. Moreover, our system has no trade-off between spatial and spectral resolutions unlike many of the earlier systems; that is, the spatial and spectral resolutions can be increased independently.

Note that our technique is *not a compressive imaging technique*; it has indeed a different objective than the compressive/snapshot methods. Different than these methods, we do not also solve an under-determined inverse problem for image reconstruction. Moreover, our technique is designed so as to be also applicable for scenes whose spectra consist of discrete spectral lines, or equivalently lack strong spectral correlation (i.e. has frequent peaks and fast transitions). Spectral sensing of such scenes are of interest in various applications including astrophysical imaging of space plasmas, atmospheric physics, and remote sensing [15], [16].

In the earlier diffractive imaging works, diffractive lenses have been mostly used with monochromatic sources [5], [17]–[19] because of their wavelength-dependent focal length. When working with monochromatic illumination (i.e. single wavelength), diffractive lenses have many attractive properties such as their low weight, low cost, and flat structure. Another important advantage of diffractive imaging elements is that they can provide diffraction-limited high spatial resolution for a wide spectral range including UV and x-rays [2], [5], [18], [20], [21]. This is not true for the refractive or reflective imaging elements. For example, at short wavelengths such as UV and x-rays, refractive lenses do not exist due to the strong absorption of available materials, and mirrors are costly to fabricate to achieve diffraction-limited resolution.

To reduce chromatic aberration and enable *broadband (full-spectrum)* imaging, different techniques have also been developed with diffractive lenses to focus different wavelengths onto a common plane either physically or computationally [22]–[29]. In contrast to these broadband imaging techniques, our approach *exploits chromatic aberration* to enable high-resolution multi-spectral imaging. Some preliminary results of this research were presented in [1], [30], [31]. Other *spectral imaging* works that take advantage of chromatic aberration focus on compressive or snapshot spectral imaging of spectrally correlated scenes. Different techniques and system architectures have been proposed for this purpose such as coded aperture spectral imaging with diffractive lenses (CSID) [13], [14], the use of a diffractive lens equipped with a light-field detector [32], a random diffractive filter in close proximity to an image sensor [33], a diffractive optical array after a pinhole array [34], and a specifically designed diffractive lens attached to a DSLR camera [35]. However, as mentioned before, in none of these compressive/snapshot imaging approaches, the main goal was to improve the spatial and spectral resolutions of the conventional systems beyond their physical limitations as in this work. Our approach can enable high spatial and spectral resolutions with low cost designs for a variety of applications and with scenes that can possibly lack spectral correlation. Moreover, the novel reconstruction

algorithms developed in this paper can be useful to improve the performance of the other diffractive imaging modalities.

The paper is organized as follows. We discuss the related work in spectral imaging and imaging with diffractive lenses in Section II. We introduce the forward model of the developed imaging technique in Section III. The inverse problem is discussed in Section IV. Section V presents the fast ADMM-based image reconstruction algorithm that exploits analysis priors (such as total variation). Section VI develops another model-based reconstruction algorithm that learns the prior through a deep network that is end-to-end trained with an unrolling approach. The performance of the developed imaging technique and the reconstruction algorithms are illustrated for EUV solar imaging in Section VII through computer simulations. A detailed analysis is also performed to demonstrate the resolution of the technique. In Section VIII, conclusions and final remarks are provided.

II. RELATED WORK

A. Multi-Spectral Imaging

Over the past decades, different techniques have been used for multi-spectral imaging. These techniques can be classified into three general groups: conventional methods, spectral filter array methods, and compressive or snapshot methods.

Conventional (scanning-based) methods obtain the spectral data cube from a series of 2D measurements that are acquired simultaneously, or sequentially. Generally this data cube is measured using an imager with a series of spectral filters (obtained through multilayer coating) scanning the scene spectrally, or by using a spectrometer that involves a dispersion element with a slit scanning the scene spatially. One of the most important drawbacks of these techniques is the intrinsic physical limitations on their spatial, spectral, and temporal resolutions, as well as the inherent trade-off between spatial and spectral resolutions [2], [3].

In fact, for multi-spectral imagers employing spectral filters, the spectral resolution is strictly limited by the bandwidth of the producible wavelength filters in that spectral range, and the spatial resolution is limited by the quality and cost of the used imaging optics [2], [4], [5]. For example, existing EUV imagers offer spatial resolutions worse than the diffraction-limit and provide spectral resolutions of 10% of the central wavelength at best [2], [4], [5], which preclude resolving important spatial and spectral features. At larger wavelengths such as visible and infrared, diffraction-limited spatial resolution is easier to attain while spectral resolution - achieved with the current multi-layer coating technology - is around 2-20% of the central wavelength [36], [37]. Our imaging technique enables to overcome these limitations on the spatial and spectral resolutions (without any physical trade-off between them), by using a single diffractive imaging element for both dispersing and focusing the optical field.

In the earlier works, super-resolution methods have been proposed to improve the resolutions of the conventional systems through post-processing [38]. However, these methods aim to guess information that is not captured by the observation system and rely on spectral correlations; as a result, their performance is limited. On the other hand, our approach improves the spatial and spectral resolutions by proposing a novel system architecture to capture the observation data and then reconstruct the high-resolution spectral cube without requiring spectral correlations.

Spectral filter array methods rely on recent technology developments that extend the Bayer filters used in color imaging to multi-spectral sensors [39]–[43]. These methods use tiled filter arrays and enable to measure different spectral bands with a single sensor. They have generally been realized at small spatial scales and particular spectral ranges, and have an inherent trade-off between spatial and spectral resolutions.

Compressive or snapshot methods use different encoding schemes to acquire reduced number of measurements than the conventional methods [3], [11], [44]. The first approach in this respect is the computed tomography imaging spectrometry (CTIS) [45]–[47], which uses a tomographic reconstruction approach with an optical system similar to a slit spectrometer, where the main difference is to remove the slit (to allow an instantaneous 2D field-of-view). An optical system similar to CTIS is also used with a parametric reconstruction approach that requires a parametric prior to hold along the spectral dimension [48]. In the compressive spectral imaging techniques - inspired by compressed sensing, the spectral data cube is reconstructed from compressive measurements by exploiting the correlations along the spatial and spectral dimensions. A commonly known compressive spectral imaging technique is coded aperture snapshot spectral imaging (CASSI) [9], with many variants [10]–[12]. The CASSI system is similar to the CTIS system where the major difference is the addition of a coded aperture, and the use of a sparse recovery algorithm for reconstruction. To provide a better alternative to CASSI, many other compressive or snapshot spectral imaging systems have been developed such as those exploiting single-pixel compressive camera architecture [49], diffractive lenses [13], [14], [32], [35], [50], random diffractive filter [33], diffractive optical array with a pinhole array [34], prism with a DSLR camera [51] and diffuser with a multi-spectral sensor [52].

All of these compressive/snapshot methods have been developed with the goal of providing faster acquisition and/or reduced hardware complexity/cost than the conventional methods. However, in none of these approaches, the main goal was to improve the spatial and spectral resolutions of the conventional scanning-based methods beyond their physical limitations. To the best of our knowledge, there have been no reports of higher spatial and spectral resolutions than the conventional scanning-based systems, and ours is the first work that demonstrates this. Moreover, our system has no trade-off between spatial and spectral resolutions unlike many of the earlier systems; that is, the spatial and spectral resolutions can be increased independently of each other.

A related work in this respect aims to overcome the physical limitation on the spectral resolution of CASSI-type architectures [53]. The spectral resolution in CASSI is limited mainly by the pitch size of the detector and the spectral dispersion of the used disperser. Spectral resolution improvement by a factor of two has been demonstrated in this work by replacing the binary coded apertures in the original CASSI system with color coded ones; however there are some technical challenges in the

fabrication of such apertures that may affect the resulting spatial resolution [53]. There are also superresolution works that aims to improve the spatial resolution of a CASSI system with a low resolution detector by taking additional measurements [54]. A similar superresolution problem that takes into account the decimation effect caused by a low-resolution detector has been studied in [55] for the spectral imaging system developed in this paper.

B. Imaging With Diffractive Lenses

Diffractive lenses such as Fresnel lenses and photon sieves have been used for various imaging techniques. These techniques can be classified into three general groups: monochromatic imaging, broadband imaging, and spectral imaging.

Monochromatic imaging has been the main area of use for diffractive lenses because of their wavelength-dependent focal length. When working with monochromatic illumination (i.e. single wavelength), diffractive lenses have many attractive properties such as their low weight, low cost, and flat structure. Another important advantage of diffractive imaging elements is that they can provide diffraction-limited high spatial resolution for a wide spectral range including UV and x-rays [2], [5], [18], [20], [21]. This is not true for the refractive or reflective imaging elements. For example, at short wavelengths such as UV and x-rays, refractive lenses do not exist due to the strong absorption of available materials, and mirrors are costly to fabricate to achieve diffraction-limited resolution. In fact, for reflective optics, surface roughness and figure errors cause substantially worse spatial resolution than the diffraction limit [17]. On the other hand, diffractive lenses can provide diffraction-limited high spatial resolution with relaxed fabrication tolerances. Many monochromatic imaging systems have been developed with diffractive lenses at visible, UV, and x-ray wavelengths to approach diffraction-limited imaging performance [5], [17]–[19].

Broadband (full-spectrum) imaging techniques are based on the development of new diffractive lens designs with reduced chromatic aberration than the conventional designs [22]–[28]. These techniques aim to focus different wavelengths onto a common plane so as to operate with broad-band illumination. Recent approaches [29] also exploit the computational imaging framework and neural networks to improve the broadband imaging performance.

In contrast to these broadband imaging techniques, our approach exploits chromatic aberration to enable high-resolution multi-spectral imaging. Some preliminary results of this research were presented in [1], [30], [31], [56]. Other **spectral imaging** works exploiting chromatic aberration focus on compressive or snapshot spectral imaging of spectrally correlated scenes. Different techniques and system architectures have been proposed for this purpose such as coded aperture spectral imaging with diffractive lenses (CSID) [13], [14], the use of a diffractive lens equipped with a light-field detector [32], a random diffractive filter in close proximity to an image sensor [33], a diffractive optical array after a pinhole array [34], and a specifically designed diffractive lens attached to a DSLR camera [29], [35]. However, in none of these compressive/snapshot

approaches, the main goal was to improve the spatial and spectral resolutions of the conventional methods beyond their physical limitations as in this work. Our approach can enable unprecedented spatial and spectral resolutions with low cost designs for a variety of applications and for scenes that can possibly lack spectral correlation. Moreover, the efficient reconstruction algorithms developed in this paper can be useful to improve the performance of other diffractive imaging modalities as well.

III. FORWARD PROBLEM

A. Proposed Multi-Spectral Imaging System

Fig. 1 depicts a schematic of the proposed spectral imaging system, which contains a diffractive lens (such as a photon sieve) and a monochrome detector. Here the optical field is first passed through a diffractive lens before recording with a monochrome detector. The diffractive lens performs both focusing and dispersion of the field. Because the focal length of the diffractive lens varies with the wavelength, each spectral component is focused at a different distance. As a result of this, when one of the spectral components is in focus, on the same plane there also exists defocused images of the other spectral components. Hence, if a measurement is taken from a plane where one spectral component is focused, the focused image of this spectral component overlaps with the defocused images of the remaining components. A total of K such measurements are taken by the spectral imaging system. For the kth measurement, d_s and d_k respectively denote the distances from the object and measurement planes to the plane where the diffractive lens lies, where $k = 1, \ldots, K$.

The measurement diversity can be achieved in different ways. One possible way is to use a moving detector along the axial direction in order to take successive measurements at different planes. In this case, a single diffractive lens can be used by changing the distance d_k between the diffractive lens and the measurement plane. Alternatively, one can fix the measurement plane, i.e. the distance d_k, and use different diffractive lens designs with a different focusing behavior to take multiple measurements at the fixed held detector. Such measurements can be acquired in many different ways depending on the wavelength regime. For example, one possible way is to generate a variable diffractive lens through a programmable spatial light modulator (SLM) or a digital micromirror device (DMD). This is a feasible approach in visible and infrared regimes at the expense of an additional optical element. More cost-effective approaches that also enable *snapshot imaging* are i) to design a single compound diffractive lens that focus multiple wavelengths onto different locations on the fixed detector [23], [57]–[59], or ii) to use multiple optical paths with different diffractive lenses by using beam-splitters. These require a larger detector (or multiple detectors), but are feasible for a wide spectral range.

From the input object, we consider a polychromatic incoherent illumination consisting of P different wavelengths λ_p for $p = 1, \ldots, P$. In the general case, these wavelengths are uniformly sampled from the spectral range observed by the system. The spectral sampling interval should be chosen in

accordance with the spectral bandwidth of the designed diffractive lens (which determines the length of the spectral range that the response of the lens starts to change moderately with wavelength [17][Chap.9].) However, it is not necessary that the sampled wavelengths are contiguous or uniform. In various multi-spectral imaging applications such as astrophysical imaging, remote sensing, and tissue spectroscopy, the wavelengths that will be observed are known a priori from spectroscopic observations, and these wavelengths do not need to be contiguous, uniformly sampled, or correlated [3]. Unlike earlier approaches, our approach is general enough to work for such cases. This will be illustrated in Section VII.

In general, the diffractive lens can form images either with spatially coherent or incoherent illumination [21]. In this work, we consider the incoherent case where the diffractive lens produces images in intensity, but similar concepts can also be applied to the coherent or partially coherent case.

B. Continuous Image-Formation Model

The first task is to mathematically relate the input spectral images to the multiple measurements obtained. Each measurement is a superposition of differently blurred spectral images. Mathematically, this can be formulated as follows:

$$t_k(u,v) = \sum_{p=1}^{P} s_p(u,v) * g_{k,p}(u,v). \tag{1}$$

Here $t_k(u,v)$ is the intensity of the kth measurement, $*$ denotes convolution, and $s_p(u,v) = \frac{d_s^2}{d_k^2} f_p(-\frac{d_s}{d_k}u, -\frac{d_s}{d_k}v)$ is a scaled version of the pth spectral image $f_p(u,v)$ at the wavelength λ_p. Hence each measurement contains P terms resulting from P different spectral components.

At distance d_k the intensity of the spectral component at wavelength λ_p is convolved with the point-spread function (PSF), $g_{k,p}(u,v)$, of the diffractive lens, given by [21], [60]

$$g_{k,p}(u,v) = \left| i\frac{\lambda_p}{\Delta_k} e^{-i\pi \frac{u^2+v^2}{\Delta_k \lambda_p d_k^2}} * A_k\left(\frac{u}{\lambda_p d_k}, \frac{v}{\lambda_p d_k}\right) \right|^2. \tag{2}$$

Here $\Delta_k = 1/d_s + 1/d_k$, and $A_k(f_u, f_v)$ is the Fourier transform of the aperture (transmittance) function of the diffractive lens used in the kth measurement. The aperture function of a diffractive lens is defined as the ratio of the transmitted field amplitude to the incident field amplitude at every point on the lens. For the commonly used diffractive lens designs with circular holes [18], $A_k(f_u, f_v)$ is sum of jinc functions resulting from the Fourier transform of the circle functions corresponding to each hole. Note that this incoherent PSF formula is derived under Fresnel approximation [21], which is valid in almost all practical imaging scenarios of interest, and also has a fast computation algorithm [60].

An approximate, but a simpler model can also be used for the PSF in Eq. (2) when the number of zones in the diffractive lens design is large and the diffractive lens behaves like a conventional (refractive) lens in its first diffraction order. This approximate PSF is given as follows [21]:

$$g_{k,p}(u,v) = \begin{cases} \frac{1}{\pi^2}(\lambda_p d_k)^4 D_k^4 \, \text{jinc}^2(\frac{D_k}{\lambda_p d_k}u, \frac{D_k}{\lambda_p d_k}v), & \epsilon_k = 0 \\ \left| \frac{1}{\pi}(\lambda_p d_k)^2 D_k^2 \, \text{jinc}(\frac{D_k}{\lambda_p d_k}u, \frac{D_k}{\lambda_p d_k}v) * \right. \\ \left. i\frac{\lambda_p}{\epsilon_k} \exp(-i\pi \frac{u^2+v^2}{\epsilon_k \lambda_p d_k^2}) \right|^2, & \epsilon_k \neq 0, \end{cases} \tag{3}$$

where D_k is the outer diameter of the diffractive lens in the kth measurement, $f_1 = \frac{D_k w_k}{\lambda_p}$ is its first-order focal distance, w_k is the outer zone width of the lens, and $\epsilon_k = 1/d_k + 1/d_s - 1/f_1$ is a parameter related to the amount of defocusing. This approximate model provides easier computation as well as simpler analysis of the imaging system.

C. Discrete Model

In practice, only a finite number of discrete measurements is available through a digital sensor such as a CCD array. Since image reconstruction will be performed computationally on a computer, a discrete representation of the spectral images is also necessary. Now, our goal is to obtain such a discrete-to-discrete model between the unknown spectral images and measurements.

For this, we exploit the band-limitedness of the continuous functions involved. First note that each PSF $g_{k,p}$ is band-limited to a circle of diameter $2D_k/(\lambda_p d_k)$. This is because the argument inside the magnitude sign in Eq.(2) has a circular frequency support with diameter $D_k/(\lambda_p d_k)$. The incoherent PSF is the magnitude square of this function, and hence the frequency support of this PSF can be found by convolving this circular support with itself, which results in a circular support with twice diameter [61]. Due to the band-limitedness of these PSFs $g_{k,p}$ in Eq. (1), each measurement t_k is bandlimited.

Secondly, note that high frequencies of the spectral images, s_p, that are outside the frequency support of the PSF, $g_{k,p}$, are not captured in the measurements, which is known as the inherent *diffraction-limit* [17]. As a result, the forward operator involving convolutions with these PSFs has a non-trivial nullspace. For discretization, we restrict our attention to the band-limited version of each spectral image and aim for recovering these band-limited versions, which are simply given by

$$x_p(u,v) \equiv s_p(u,v) * \text{jinc}\left(\frac{2D_k}{\lambda_p d_k}u, \frac{2D_k}{\lambda_p d_k}v\right). \tag{4}$$

The jinc function $\text{jinc}(u,v) = \frac{J_1(\pi\sqrt{u^2+v^2})}{2\sqrt{u^2+v^2}}$ where $J_1(u)$ is the first-order Bessel function of the first kind. The forward model in Eq. (1) is still valid when the unknown source intensities, s_p, are replaced with their band-limited versions, x_p. Hence, we can treat all functions in the continuous forward model as band-limited functions and represent them using sinc basis [61].

Now by representing each continuous band-limited function with sinc basis, the continuous convolution operations in Eq. (1) reduce to discrete convolutions of the form

$$t_k[m,n] = \sum_{p=1}^{P} x_p[m,n] * g_{k,p}[m,n], \tag{5}$$

where $m, n = [0, \ldots, N-1]$. Here $t_k[m,n]$, $x_p[m,n]$, and $g_{k,p}[m,n]$ are sampled versions of their continuous counterparts, e.g. $t_k[m,n] = t_k(m\Delta, n\Delta)$ for some Δ smaller than the Nyquist sampling interval (i.e. $\Delta < \frac{\lambda_p d_k}{2D_k}$ for all k). Here the uniformly sampled observations, $t_k[m,n]$, are assumed to be same as the detector measurements, i.e. the averaged intensity over detector pixels of width Δ.

Let the PSF $g_{k,p}[m,n]$ has $M \times M$ support, i.e. $g_{k,p}[m,n] \approx 0$ for $m,n \notin [0, M-1]$. We assume that the supports of the spectral images are slightly smaller than the detector range covered by N pixels, i.e. $x_p[m,n] = 0$ for $m,n \notin [0, N-M]$. Then the linear convolution in Eq. (5) can be replaced with a circular convolution of N points, which will be useful to develop fast image reconstruction methods.

Due to the linearity of the convolution operator, the discrete model can be formulated in the following matrix-vector form using appropriate lexicographic ordering:

$$\mathbf{t}_k = \sum_{p=1}^{P} \mathbf{H}_{k,p} \mathbf{x}_p. \tag{6}$$

Here $\mathbf{H}_{k,p}$ is an $N^2 \times N^2$ block circulant matrix with circular blocks corresponding to the circular convolution operation with the PSF $g_{k,p}[m,n]$. Vectors $\mathbf{t}_k$ and $\mathbf{x}_p$ correspond to lexicographically ordered versions of the observation, $t_k[m,n]$, and spectral image, $x_p[m,n]$, respectively. By combining all the observation vectors into a single vector, $\mathbf{t}$, and all the image vectors into the vector $\mathbf{x}$, we obtain

$$\mathbf{t} = \mathbf{Hx}, \tag{7}$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{1,1} & \ldots & \mathbf{H}_{1,P} \\ \vdots & & \vdots \\ \mathbf{H}_{K,1} & \ldots & \mathbf{H}_{K,P} \end{bmatrix}, \tag{8}$$

where $\mathbf{t} = [\mathbf{t}_1^T | \ldots | \mathbf{t}_K^T]^T$, $\mathbf{x} = [\mathbf{x}_1^T | \ldots | \mathbf{x}_P^T]^T$, and $\mathbf{H}$ is a $KN^2 \times PN^2$ matrix. The final observation model can be expressed with noise as follows:

$$\mathbf{y} = \mathbf{t} + \mathbf{w} = \mathbf{Hx} + \mathbf{w} \tag{9}$$

where $\mathbf{w} = [\mathbf{w}_1^T | \ldots | \mathbf{w}_K^T]^T$ is the additive noise vector. Here, white Gaussian noise is assumed, where $(\mathbf{w}_k)_i \sim N(0, \sigma_k^2)$ is uncorrelated over different pixels i and measurements k, with σ_k representing the noise standard deviation for the kth measurement. This noise assumption is generally valid for practical multi-spectral imaging scenarios of interest.

IV. INVERSE PROBLEM

In the inverse problem, the goal is to recover the unknown spectral images, $\mathbf{x}$, from the measurements, $\mathbf{y}$, obtained with the proposed spectral imaging system. This inverse problem can be considered as a multi-frame deconvolution problem where the measurements are in the form of superimposed blurred images. That is, each measurement is a superposition of focused or defocused versions of all spectral images. This problem is inherently ill-posed; as the PSFs, $g_{k,p}[m,n]$, of different wavelengths, p, and measurements, k, become more similar (for example, when

the difference between different wavelengths or measurement distances decreases), the conditioning of the problem gets worse due to increased dependency of the columns or rows of $\mathbf{H}$, respectively.

There are various approaches for solving ill-posed linear inverse problems. A systematic approach to regularization uses prior knowledge about the unknown solution and leads to the minimization of an appropriately formulated cost function [62], [63]. The prior information can be introduced in a deterministic or a statistical setting, with the latter related to the Bayes framework [64]. A general formulation of the inverse problem can be expressed as

$$\min_{\mathbf{x}} \ ||\mathbf{y} - \mathbf{Hx}||_2^2 + \lambda \Phi(\mathbf{Px}). \tag{10}$$

where the first term measures data fidelity, and the second term, $\Phi(\mathbf{Px})$, controls how well the solution matches the prior knowledge, with the parameter λ trading off between these two terms. Here $\mathbf{P}$ is a matrix representing an analysis operator.

Equivalently, we can reformulate this problem as a constrained problem:

$$\min_{\mathbf{x}} \ \Phi(\mathbf{Px}) \text{ subject to } ||\mathbf{y} - \mathbf{Hx}||_2 \leq \epsilon, \tag{11}$$

where $\epsilon \geq 0$ is a parameter that depends on noise level. Note that if the problem in Eq. (11) is feasible for some $\epsilon \geq 0$, then it is equivalent to Eq. (10) for some $\lambda \geq 0$. The formulation in Eq. (11) has advantage over Eq. (10) due to the fact that the parameter ϵ is directly proportional to the noise standard deviation, so it is easier to choose than the parameter λ [65].

There are popular and powerful choices for the regularizer $\Phi(.)$ [7], [8], [65], [66]. Because we want our approach to be applicable for scenes that can possibly lack spectral correlation, we exploit 2D priors for each spectral image (rather than a 3D prior for the entire spectral cube). Among 2D analytical priors, total-variation (isotropic or anisotropic) has significant popularity as it provides superior results for images with significant structure or piecewise-constant characteristics [63]. As the analytical priors have recently been shown to be outperformed by deep priors [6]–[8], there is also significant interest in learning the prior through a deep network that is trained offline or end-to-end with an unrolling approach. In this work, we develop two reconstruction algorithms to solve the inverse problem, one with analytical priors and another with deep priors, and comparatively evaluate their performance.

V. IMAGE RECONSTRUCTION WITH ANALYTICAL PRIORS

In this work, we first develop a fast reconstruction algorithm using the ADMM framework to solve the resulting optimization problem in Eq. (11) with an analytical prior (particularly, with an analysis prior). ADMM belongs to the family of augmented Lagrangian methods [67] and is used in many signal and image reconstruction problems [65], [68]. It provides a divide-and-conquer approach by splitting the minimization steps in an unconstrained multi-objective optimization problem. Its convergence is guaranteed under mild conditions.

Following the ADMM framework [65], we first convert the problem in Eq. (11) to an unconstrained problem by adding the

constraint to the cost function as a penalty:

$$\min_{\mathbf{x}} \ \Phi(\mathbf{Px}) + \iota_{(\|\mathbf{y}-\mathbf{Hx}\|_2 \leq \epsilon)}(\mathbf{x}), \qquad (12)$$

where the indicator function $\iota_{(\|\mathbf{y}-\mathbf{Hx}\|_2 \leq \epsilon)}(\mathbf{x})$ is defined as

$$\iota_{(\|\mathbf{y}-\mathbf{Hx}\|_2 \leq \epsilon)}(\mathbf{x}) = \begin{cases} 0, & \text{if } \|\mathbf{y} - \mathbf{Hx}\|_2 \leq \epsilon \\ +\infty, & \text{if } \|\mathbf{y} - \mathbf{Hx}\|_2 > \epsilon. \end{cases}$$

After variable-splitting, we arrive at the following problem:

$$\begin{aligned} \underset{\mathbf{x},\mathbf{u},\mathbf{v}}{\text{minimize}} \ & \Phi(\mathbf{u}) + \iota_{(\|\mathbf{y}-\mathbf{v}\|_2 \leq \epsilon)}(\mathbf{v}) \\ \text{subject to} \ & \mathbf{u} = \mathbf{Px}, \ \mathbf{v} = \mathbf{Hx} \end{aligned} \qquad (13)$$

where $\mathbf{u}$ and $\mathbf{v}$ are the auxiliary variables in the ADMM framework. After formulating the problem in (13) in augmented Lagrangian form [65], minimization over $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{v}$ is required. We perform each minimization in an alternating fashion. For the lth iteration, the required minimizations are

$$\mathbf{x}^{l+1} = \arg\min_{\mathbf{x}} \frac{\mu}{2} \left\| \begin{bmatrix} \mathbf{H} \\ \mathbf{P} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{v}^l \\ \mathbf{u}^l \end{bmatrix} - \begin{bmatrix} \mathbf{f}^l \\ \mathbf{d}^l \end{bmatrix} \right\|_2^2 \qquad (14)$$

$$\mathbf{u}^{l+1} = \arg\min_{\mathbf{u}} \Phi(\mathbf{u}) + \frac{\mu}{2} \|\mathbf{u} - (\mathbf{Px}^{l+1} - \mathbf{d}^l)\|_2^2 \qquad (15)$$

$$\mathbf{v}^{l+1} = \arg\min_{\mathbf{v}} \iota_{(\|\mathbf{y}-\mathbf{v}\|_2 \leq \epsilon)}(\mathbf{v}) + \frac{\mu}{2} \|\mathbf{v} - (\mathbf{Hx}^{l+1} - \mathbf{f}^l)\|_2^2 \qquad (16)$$

with $\mathbf{d}$ and $\mathbf{f}$ denoting the dual variables in the ADMM framework, and μ is a penalty parameter. We now explain how these update steps, referred to as $\mathbf{x}$-update, $\mathbf{u}$-update, and $\mathbf{v}$-update, are performed.

In the $\mathbf{x}$-update step, we face the least squares problem in Eq. (14) with the following closed-form solution:

$$\mathbf{x}^{l+1} = (\mathbf{I} + \mathbf{H}^H \mathbf{H})^{-1}(\mathbf{P}^H(\mathbf{u}^l + \mathbf{d}^l) + \mathbf{H}^H(\mathbf{v}^l + \mathbf{f}^l)). \quad (17)$$

Here $\mathbf{P}$ is assumed to be a unitary transform resulting in $\mathbf{P}^H \mathbf{P} = \mathbf{I}$. This solution can be efficiently obtained through computations in the frequency domain by exploiting the diagonalizability of the 2D circular convolution operations [69].

Note that each block of $\mathbf{H}$ is diagonalized by the discrete Fourier transform (DFT) matrix because $\mathbf{H}_{k,p}$ is block circulant with circular blocks (BCCB). Hence $\mathbf{H}_{k,p} = \mathbf{F}_{2D}^H \mathbf{\Lambda}_{k,p} \mathbf{F}_{2D}$ where $\mathbf{F}_{2D}$ is the unitary 2D DFT matrix and $\mathbf{\Lambda}_{k,p}$ is a diagonal matrix whose diagonal consists of the 2D DFT of the PSF $g_{k,p}[m,n]$, for $k = 1, \dots, K$ and $p = 1, \dots, P$. As a result, the overall matrix $\mathbf{H}$ can be written as $\mathbf{H} = \bar{\mathbf{F}}^H \mathbf{\Lambda} \tilde{\mathbf{F}}$ where $\bar{\mathbf{F}} = \mathbf{I}_K \otimes \mathbf{F}_{2D}$ and $\tilde{\mathbf{F}} = \mathbf{I}_P \otimes \mathbf{F}_{2D}$ with $\otimes$ denoting the Kronecker product and $\mathbf{I}_n$ denoting the identity matrix of size $n \times n$. Here $\mathbf{\Lambda}$ is a matrix of $K \times P$ blocks with each block given by $\mathbf{\Lambda}_{k,p}$. By inserting this expression of $\mathbf{H}$ in Eq. (17), the following form can be obtained for the efficient computation of the image update step:

$$\mathbf{x}^{l+1} = \tilde{\mathbf{F}}^H (\mathbf{I} + \mathbf{\Lambda}^H \mathbf{\Lambda})^{-1} (\tilde{\mathbf{F}} \mathbf{P}^H(\mathbf{u}^l + \mathbf{d}^l) + \mathbf{\Lambda}^H \bar{\mathbf{F}}(\mathbf{v}^l + \mathbf{f}^l)). \quad (18)$$

To compute the solution in Eq. (18), we do not need to form huge matrices, which provides significant savings for memory and computation time. In fact, multiplication by $\tilde{\mathbf{F}}$ or $\tilde{\mathbf{F}}^H$ corresponds to taking the DFT or inverse DFT of all

2D spectral bands for $p = 1, \dots, P$. Similarly, multiplication by $\bar{\mathbf{F}}$ corresponds to taking the DFT of the 2D measurement signals for $k = 1, \dots, K$. For example, $\bar{\mathbf{F}}(\mathbf{v}^l + \mathbf{f}^l) = [(\mathbf{F}_{2D}(\mathbf{v}_1^l + \mathbf{f}_1^l))^T | \dots | (\mathbf{F}_{2D}(\mathbf{v}_K^l + \mathbf{f}_K^l))^T]^T$, where each term can be computed via the 2D FFT. Moreover, because $\mathbf{\Lambda}$ is a block matrix consisting of diagonal matrices, multiplication by $\mathbf{\Lambda}^H$ corresponds to element-wise 2D multiplication with the conjugated DFTs of the underlying PSFs and then summation. Furthermore, for a unitary $\mathbf{P}$, multiplication by $\mathbf{P}^H$ corresponds to taking the inverse transform. Note that when the image data is correlated in all directions, this can be chosen as a 3D transform; otherwise, it can be a 2D transform applied on each spectral band separately.

Lastly, the inverse of $\mathbf{\Sigma} \triangleq (\mathbf{I} + \mathbf{\Lambda}^H \mathbf{\Lambda})$ needs to be computed only once, and hence does not affect the computational cost of the iterations. Nevertheless, it is possible to reduce the required time and memory for this pre-computation through a recursive block matrix inversion approach [70]. Note that $\mathbf{\Sigma}$ is a block matrix of $P \times P$ blocks, where each block is a diagonal matrix given by $\mathbf{\Sigma}_{i,j} = \sum_{k=1}^{K} \mathbf{\Lambda}_{k,i}^H \mathbf{\Lambda}_{k,j} + \delta_{i,j} \mathbf{I}$ with $\delta_{i,j}$ denoting the Kronecker delta function and $i, j = 1, \dots, P$. Hence, for $P = 2$, the inverse can be computed as

$$\begin{bmatrix} \mathbf{\Sigma}_{1,1} & \mathbf{\Sigma}_{1,2} \\ \mathbf{\Sigma}_{2,1} & \mathbf{\Sigma}_{2,2} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{C} & \mathbf{\Sigma}_{1,1}^{-1} \mathbf{\Sigma}_{1,2} \mathbf{K} \\ \mathbf{K} \mathbf{\Sigma}_{2,1} \mathbf{\Sigma}_{1,1}^{-1} & -\mathbf{K} \end{bmatrix} \qquad (19)$$

where $\mathbf{K} = -(\mathbf{\Sigma}_{2,2} - \mathbf{\Sigma}_{2,1} \mathbf{\Sigma}_{1,1}^{-1} \mathbf{\Sigma}_{1,2})^{-1}$ and $\mathbf{C} = \mathbf{\Sigma}_{1,1}^{-1} - \mathbf{\Sigma}_{1,1}^{-1} \mathbf{\Sigma}_{1,2} \mathbf{K} \mathbf{\Sigma}_{2,1} \mathbf{\Sigma}_{1,1}^{-1}$. For $P > 2$ case, the overall matrix $\mathbf{\Sigma}$ is partitioned into 2×2 blocks and each block is inverted recursively using Eq. (19). Because all the matrices involved in these computations are diagonal, computing the inverse of $\mathbf{\Sigma}$ requires element-wise multiplication and division operations.

For the $\mathbf{u}$-update step, the problem in Eq. (15) needs to be solved, which corresponds to a denoising problem of the form:

$$\Psi_f(\mathbf{z}) = \arg\min_{\mathbf{u}} f(\mathbf{u}) + \frac{\mu}{2} \|\mathbf{u} - \mathbf{z}\|_2^2. \qquad (20)$$

Here f is a regularization functional and $\mathbf{z}$ is a noisy observation for $\mathbf{u}$. The solution is given by the Moreau proximal mapping of the regularization functional f/μ evaluated at $\mathbf{z}$. Hence, the solution of Eq. (15) is, by definition, the Moreau proximal mapping of $\Phi(.)/\mu$ evaluated at $\mathbf{Px}^{l+1} - \mathbf{d}^l$. We denote this proximal mapping as $\Psi_{\Phi/\mu}(\mathbf{Px}^{l+1} - \mathbf{d}^l)$:

$$\mathbf{u}^{l+1} = \Psi_{\Phi/\mu}(\mathbf{Px}^{l+1} - \mathbf{d}^l). \qquad (21)$$

There are efficient computations of $\Psi_\Phi(\mathbf{Px}^{l+1} - \mathbf{d}^l)$ for the different choices of the functional $\Phi(.)$. For example, if ℓ_1-norm is used, i.e. $\Phi(\mathbf{Px}) = \|\mathbf{Px}\|_1$, then $\Psi_{\Phi/\mu}(\mathbf{Px}^{l+1} - \mathbf{d}^l)$ is simply soft-thresholding, i.e. $soft(\mathbf{Px}^{l+1} - \mathbf{d}^l, 1/\mu)$ with $1/\mu$ denoting the threshold. Here, multiplication by $\mathbf{P}$ corresponds to either a single 3D transform or multiple 2D transforms along each spectral band. The soft-thresholding operation, $soft(\mathbf{n}, \tau)$, is component-wise computed as $\mathbf{n}_i \to sign(\mathbf{n}_i) \max(|\mathbf{n}_i| - \tau, 0)$ for each voxel i, with $sign(\mathbf{n}_i)$ taking value 1 if $\mathbf{n}_i > 0$ and -1 otherwise [65]. If $\Phi(.)$ is chosen as the 2D isotropic TV operator applied on each spectral band, $\mathbf{P}$ becomes an identity matrix, and the resulting proximal mapping, $\Psi_{\text{TV}}(\mathbf{x}^{l+1} - \mathbf{d}^l)$, has an efficient decoupled calculation for each spectral band

TABLE I
IMAGE RECONSTRUCTION ALGORITHM WITH ANALYTICAL PRIOR

Input: $\mathbf{y}$: measurements, $\mathbf{H}$: system matrix, $\mathbf{P}$: transform
Output: $\mathbf{x}$: reconstructed image
1. set: $l = 0$
choose: $\mu > 0$, $\mathbf{u}^0$, $\mathbf{v}^0$, $\mathbf{d}^0$, $\mathbf{f}^0$
2. **repeat**
3.   update $\mathbf{x}^{l+1}$ using Eq. (18).
4.   update $\mathbf{u}^{l+1}$ using Eq. (21).
5.   update $\mathbf{v}^{l+1}$ using Eq. (22).
6.   update $\mathbf{d}^{l+1}$ using Eq. (23).
7.   update $\mathbf{f}^{l+1}$ using Eq. (24).
9.   set $l \leftarrow l + 1$.
10. **until:** some stopping criterion is satisfied.

using Chambolle's algorithm [65], [71]. Specifically, we compute the proximal mapping for each band of $\mathbf{x}^{l+1} - \mathbf{d}^l$ using Chambolle's algorithm and then concatenate the updated bands along the spectral dimension.

For the $\mathbf{v}$-update step given in Eq. (16), similar to the $\mathbf{u}$-update, we have the proximal mapping of $\iota_{(\|\mathbf{y}-\mathbf{v}\|_2 \le \epsilon)/\mu}(.)$ evaluated at $\mathbf{H}\mathbf{x}^{l+1} - \mathbf{f}^l$, which we denote as $\Psi_{\iota_{(\|\mathbf{y}-\mathbf{v}\|_2 \le \epsilon)/\mu}}(\mathbf{H}\mathbf{x}^{l+1} - \mathbf{f}^l)$. Calculation of this proximal mapping requires a projection of $\mathbf{s} \triangleq (\mathbf{H}\mathbf{x}^{l+1} - \mathbf{f}^l)$ onto ϵ-radius ball centered at $\mathbf{y}$. The solution is given [65] by

$$\Psi_{\iota_{(\|\mathbf{y}-\mathbf{v}\|_2 \le \epsilon)}}(\mathbf{s}) = \begin{cases} \mathbf{y} + \epsilon \frac{\mathbf{s}-\mathbf{y}}{\|\mathbf{s}-\mathbf{y}\|_2}, & \text{if } \|\mathbf{s}-\mathbf{y}\|_2 > \epsilon \\ \mathbf{s}, & \text{if } \|\mathbf{s}-\mathbf{y}\|_2 \le \epsilon. \end{cases} \quad (22)$$

Finally, ADMM dual variables, $\mathbf{d}$ and $\mathbf{f}$, are updated as

$$\mathbf{d}^{l+1} = \mathbf{d}^l - (\mathbf{P}\mathbf{x}^{l+1} - \mathbf{u}^{l+1}), \quad (23)$$

$$\mathbf{f}^{l+1} = \mathbf{f}^l - (\mathbf{H}\mathbf{x}^{l+1} - \mathbf{v}^{l+1}). \quad (24)$$

The overall algorithm is summarized in Table I. In the numerical results, we choose $\Phi(.)$ as the 2D isotropic TV operator and Step 4 is solved efficiently using Chambolle's algorithm [71] for each spectral band, as explained before.

A. Computational Complexity

The computational complexity of the algorithm is dominated by the $\mathbf{x}$-update, given in Step 3 in Table I. This requires $2P$ FFT and P inverse FFT computations. Thus, its computational complexity is $O(PN^2 \log(N))$ where N^2 is the size of each spectral image and P is the number of spectral bands. Calculation of the proximal mapping in Step 4 and ADMM dual variable update in Step 6 have $O(N^2)$ complexity if $\mathbf{P}$ is a diagonal matrix, or $O(PN^2 \log(N))$ if $\mathbf{P}$ has a fast implementation such as with FFT. Steps 5 and 7 require multiplication with $\mathbf{H}$, which are computed using FFT with $O(PN^2 \log(N))$ complexity. Thus, the overall complexity of the algorithm is $O(PN^2 \log(N))$ as determined by the complexity of the computation of the forward operator and its adjoint. It is also worth noting that the complexity of the recursive inversion in Eq. (19) is $O(P^3 N^2)$. Since it is pre-calculated once, it does not affect the algorithm's overall complexity.

VI. LEARNED RECONSTRUCTION WITH DEEP PRIORS

We also develop a model-based learned reconstruction method with deep priors motivated by the idea presented in [7]. Model-based learned reconstruction methods can simultaneously incorporate physics-based forward model and the data, and have been successfully applied to various inverse problems in imaging [6]–[8], [35].

Our approach learns a 2D prior through a deep network that is end-to-end trained through unrolling. For this, we apply half-quadratic splitting to Eq. (10) with $\mathbf{P} = \mathbf{I}$, resulting in

$$\min_{\mathbf{x},\mathbf{z}} ||\mathbf{y} - \mathbf{H}\mathbf{x}||_2^2 + \lambda\Phi(\mathbf{z}) + \nu||\mathbf{x} - \mathbf{z}||_2^2. \quad (25)$$

Here ν is a learned penalty parameter. This problem can be alternatingly solved as follows:

$$\mathbf{z}^l = \mathbf{D}(\mathbf{x}^l), \quad (26a)$$

$$\mathbf{x}^{l+1} = \arg\min_{\mathbf{x}} ||\mathbf{y} - \mathbf{H}\mathbf{x}||_2^2 + \nu||\mathbf{x} - \mathbf{z}^l||_2^2, \quad (26b)$$

These can be regarded as denoising and data fidelity updates, respectively. Here $\mathbf{D}(\mathbf{x})$ is the denoised and artifact free version of $\mathbf{x}$ that is obtained by passing $\mathbf{x}$ through a convolutional neural network (CNN).

For the data-fidelity update, the least-squares problem in Eq. (26b) can be solved in closed-form as follows:

$$\mathbf{x}^{l+1} = (\nu\mathbf{I} + \mathbf{H}^H\mathbf{H})^{-1}(\mathbf{H}^H\mathbf{y} + \nu\mathbf{z}^l). \quad (27)$$

This solution can be efficiently obtained through computations in the frequency domain by exploiting the diagonalizability of the convolution operations involved. Remember that $\mathbf{H}$ can be written as $\mathbf{H} = \bar{\mathbf{F}}^H \mathbf{\Lambda} \tilde{\mathbf{F}}$. By inserting this expression in Eq. (27), we obtain the following form for the efficient data fidelity update step:

$$\mathbf{x}^{l+1} = \tilde{\mathbf{F}}^H (\nu\mathbf{I} + \mathbf{\Lambda}^H\mathbf{\Lambda})^{-1}(\mathbf{\Lambda}^H\bar{\mathbf{F}}\mathbf{y} + \nu\tilde{\mathbf{F}}\mathbf{z}^l). \quad (28)$$

Here, multiplication by $\tilde{\mathbf{F}}$ or $\tilde{\mathbf{F}}^H$ corresponds to taking the FFT or inverse FFT of all 2D spectral bands for $p = 1, \ldots, P$. Similarly, multiplication by $\bar{\mathbf{F}}$ corresponds to taking the FFT of the 2D measurements for $k = 1, \ldots, K$. Moreover, because $\mathbf{\Lambda}$ is a block matrix consisting of diagonal matrices, multiplication by $\mathbf{\Lambda}^H$ corresponds to element-wise 2D multiplication with the conjugated FFTs of the PSFs and then summation. The inverse of $\nu\mathbf{I} + \mathbf{\Lambda}^H\mathbf{\Lambda}$ can be computed through a recursive block matrix inversion approach [70], as before.

For the denoising update in Eq. (26a), the CNN architecture proposed in [7] is used. This CNN consists of N layers, with each layer having 64 filters of size 3×3. First $N - 1$ layer has convolution, batch normalization, and non-linear activation function ReLU units. The last layer, i.e., N^{th} layer, does not have ReLU so that the negative part of the learned noise is not truncated. Residual learning strategy is used to speed up the learning. After fixing the number of iterations to L for the alternating updates, the steps in Eq. (26a) and (26b) can be unrolled to an end-to-end trainable network as shown in Fig. 2. Sharing weights are used between different iterations to decrease the number of trainable parameters. After training the network,
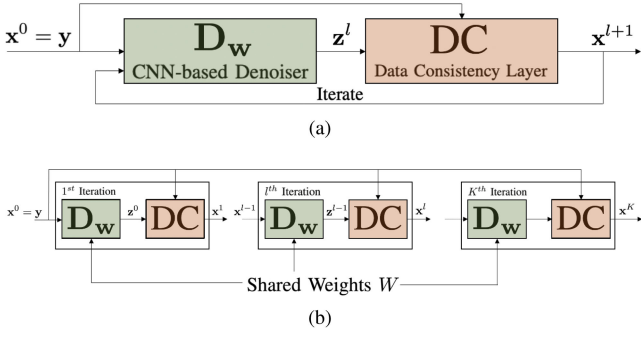
(a)

(b)

Fig. 2. Model-based learned reconstruction framework. (a) The iterative algorithm described by (26a) and (26b), (b) the resulting architecture after unrolling.

the images are reconstructed by simply feeding the network with the available measurements.

VII. NUMERICAL RESULTS

In this section, we demonstrate the high spatial and spectral resolutions enabled by the developed multi-spectral imaging technique in solar imaging [2]. For this, we consider three nearby solar emission lines in the EUV regime, with central wavelengths of $\lambda_1 = 33.28$ nm, $\lambda_2 = 33.42$ nm and $\lambda_3 = 33.54$ nm (i.e., $P = 3$) [72]. Our goal is to obtain the spectral images for these three nearby emission lines with diffraction-limited spatial resolution. Such high spatial and spectral resolutions are not possible to achieve with the conventional spectral imagers involving wavelength filters [1], [2], [4], [5]. Obtaining these spectral images will provide information about the important physical parameters of the Sun's extended atmosphere, and hence will enable the investigation of the dynamic plasma behavior [72], [73].

For the diffractive lens, a sample photon sieve (PS) design developed in [2] for the considered application is used, where the outer diameter is chosen as $D = 25$ mm, and the diameter of the smallest hole as $\Delta = 5\mu$m. The resulting photon sieve has first-order focal lengths of $f_1 = 3.756$ m, $f_2 = 3.740$ m, and $f_3 = 3.727$ m at the corresponding wavelengths (i.e. $f = D\Delta/\lambda$) [2]. As determined by the smallest hole diameter Δ, the resulting (Abbe's) diffraction-limited spatial resolution is 5μm [17]. The detector pixel size is chosen as 2.5μm to match the diffraction-limited resolution with two pixels on the detector (i.e. the Nyquist rate sampling). Moreover, the expected spectral resolution is $4\Delta\lambda/D \approx 0.03$ nm near a wavelength of 33 nm, as given by the spectral bandwidth of the diffractive lens [17][Chap.9]. This bandwidth is slightly larger than the typical width of the considered emission lines and hence the entire spectral line will be in good focus at each focal plane [2]. Moreover, the distance between the wavelength sources is ~ 0.1 nm, which is larger than the spectral resolution supported by the system. Note that the spatial and spectral resolutions of the system can be increased independently of each other, as the former is determined by the minimum hole diameter Δ while the latter is by $4\Delta\lambda/D$.

For the imaging system, two different measurement settings are considered. In the *moving detector* (MD) case, the system measures the intensities at the three focal planes, f_1, f_2 and f_3

(i.e., $K = 3$). Hence at the first focal plane, the measurement contains a focused image of the first source and the defocused images of the second and third sources, and vice versa at the other focal planes. Note that the spectral bandwidth of the PS design is equal to $4\Delta\lambda/D \approx 0.03$ nm near a wavelength of 33 nm, which is slightly larger than the typical width of an emission line in this application. Hence, in each measurement, one of the spectral (emission) lines will be entirely in focus, while the other two will be defocused.

Alternatively, in the *fixed detector* (FD) case, we fix the measurement plane at the distance f_2 and obtain the measurements by changing the PS design at each shot. For each measurement, same design parameters are used as before by only changing the diameter D to focus λ_1 or λ_3 onto the distance f_2. Because the measurement plane is fixed to $f_2 = 3.740$ m, the outer diameter is decreased by 104.6 μm for λ_1 and is increased by 89.9 μm for λ_3. These modifications on the PS design are well within the limits of modern lithography techniques [19]. FD measurement setting obtains similar measurements with the MD setting without moving the detector in the axial direction, but at the expense of additional optical components (SLM or beamsplitters), larger detector array, and/or lower SNR due to the reduced diffraction efficiency of the compound lens designs. Replacing the diffractive lens may also be considered for the FD setting using a mechanical component, but this may be more challenging in practice.

A. Performance Analysis

We present numerical simulations to demonstrate the performance of the developed imaging technique and algorithms under different imaging scenarios with MD and FD settings. For this, solar EUV images of size 256×256 pixels are used as the input images. However, because the spatial resolution of the existing solar imagers are worse than the diffraction-limited resolution enabled by the proposed technique, realistic (high-resolution) solar images are not available for the simulations. Here we use these images as if they were images of some other sun-like object, and demonstrate the diffraction-limited resolution for this. Hence this experiment will illustrate that objects with similar characteristics can be observed with diffraction-limited high spatial resolution.

Using the forward model in Eq. (9), we first simulate the measurements $\mathbf{y}$ at the signal-to-noise ratio (SNR) of 25 dB (where SNR is defined as $10\log_{10}\frac{\sigma_{\bar{x}}^2}{\sigma_w^2}$). Fig. 3 illustrates the simulated measurements at the three focal planes together with the contribution from each spectral source. To also demonstrate the wavelength-dependent behavior of the system, the acting PSFs of the diffractive lens at the three source wavelengths are illustrated in Fig. 3(e), (j), (o) when the measurement is taken at the 1st focal plane. These PSFs illustrate that each spectral source is exposed to a different amount of blur. As clearly seen from the shown contributions and PSFs, the first source is focused and the other two sources are defocused when the measurement is taken at the 1st focal plane. We have similar behavior for the measurements at the 2nd and 3 rd focal planes. Hence the measurements involve not only the superposition of
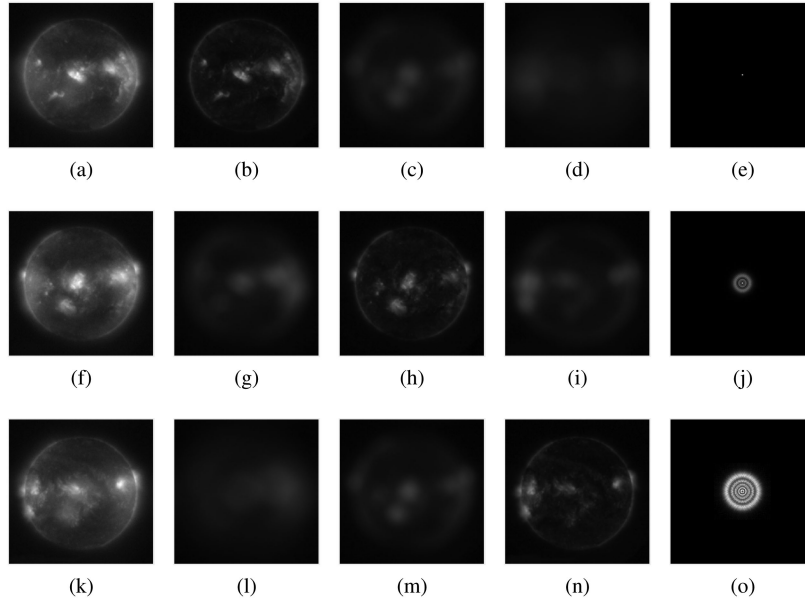
Fig. 3. Measured intensities for SNR = 25 dB at the first focal plane (a), at the second focal plane (f), and at the third focal plane (k); the contribution of the 1st, 2nd, and 3 rd sources to the measurement at the first focal plane (b)-(d), at the second focal plane (g)-(i), at the third focal plane (l)-(n); sampled and zoomed PSFs of the system at the first focal plane for the 1st, 2nd, and 3 rd sources, (e), (j), and (o), respectively.

all spectral images but also substantial and varying amount of blur and noise.

We obtain the reconstructions using the algorithm in Table I with isotropic TV prior and the learned reconstruction method with deep prior. For TV prior, one reconstruction takes a couple of minutes while for learned reconstruction it takes a couple of seconds on a computer with 32 GB of RAM, Intel i $9 - 9900$ K 3.60 GHz CPU, and Nvidia GeForce RTX 2080 Ti GPU.

For the model-based learned reconstruction method, we train our model using solar EUV images from the NASA SDO database [4]. We use SDO images from the 1st, 5th, 10th, 15th, 25th, and 30th days of every month for years between 2012 and 2020, to have training data diversity. After removing the images with artifacts (such as out-of-focus), we randomly pick 6 images from each day. These images are then converted to grayscale, resized to 256×256, and normalized between [0,1]. Finally, we combine randomly chosen three images to form 725 training datasets (for $P = 3$ case). In a similar way, we have generated separate 105 test datasets.

We train four different models for signal-to-noise ratios (SNR) of 15 dB, 20 dB, 25 dB, and 30 dB. For this, we simulate the measurements $\mathbf{y}$ at these SNRs using the forward model in Eq. (9). We then train our models by minimizing the l_2 loss between the ground truth and reconstructed images by using Adam solver [74]. We initialize the model parameters with Xavier initialization. We choose $N = 5$, $L = 20$, and the learning rate as 0.001 after some empirical efforts. Codes are implemented with Tensorflow on Nvidia GeForce RTX 2080 Ti GPU. Training takes approximately 30 hours.

To illustrate the performance of the developed imaging modality and the two reconstruction methods, we use the randomly constructed 105 test datasets described above. For each dataset, we perform 5 Monte Carlo runs and compute the average PSNR

and SSIM values between the diffraction-limited ground truth images and the reconstructed spectral images using TV and learned reconstruction methods. Note that here the comparison is performed with the diffraction-limited versions of the original images to check the imaging system's capability for achieving diffraction-limited spatial resolution. Table II presents the resulting PSNR and SSIM values. The first row demonstrates the average performance over the entire test dataset as SNR changes from 15 dB to 30 dB. To illustrate the performance at different levels of difficulty, 5 spectral datasets are also selected from this large test dataset and their PSNR/SSIM values are also separately given. Numbered as 1 to 5, these selected datasets have increasing levels of abrupt changes along both spatial and spectral dimensions, corresponding to observations with different sizes of active and quiet regions on the sun.

As seen from Table II, the learned reconstruction method provides the best performance in this large dataset with the average PSNR changing from 29.7 dB to 35.8 dB and SSIM changing from 0.79 to 0.92 as SNR increases from 15 to 30 dB. The obtained PSNR and SSIM values between the reconstructions and diffraction-limited original images illustrate the imaging system's diffraction-limited spatial resolution capability over a wide range of SNRs. As expected, the learned reconstruction method provides significantly better performance than the TV reconstruction on the average both in terms of PSNR and SSIM. This illustrates the better generalization capability of the learned reconstruction method for a particular application. Moreover, as expected, the performance of the TV reconstruction decreases with the increased structure in the images (i.e. from dataset 1 to dataset 5). The performance improvement achieved with the learned reconstruction method is more significant for the datasets with significant structure (for example, for datasets 4 and 5). Note that for the case that all three spectral images

TABLE II
PSNR (dB) AND SSIM VALUES FOR TV AND LEARNED RECONSTRUCTION METHODS AT DIFFERENT INPUT SNRS AND FOR DIFFERENT SPECTRAL DATASETS

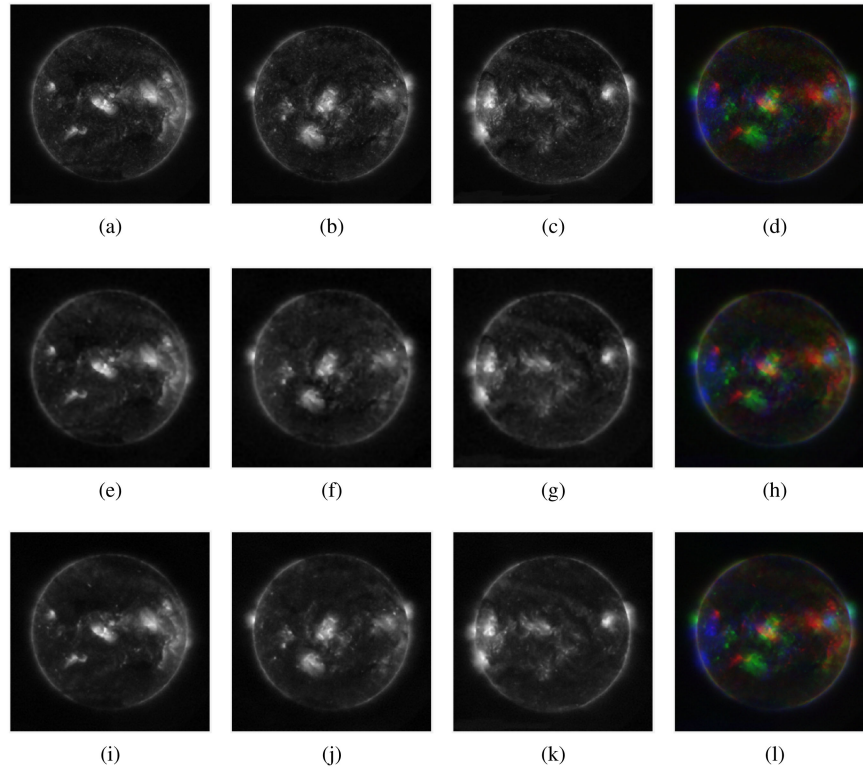| SNR	15dB		20dB		25dB		30dB	
Method	TV Recon.	Learned Recon.	TV Recon.	Learned Recon.	TV Recon.	Learned Recon.	TV Recon.	Learned Recon.
Average over 105 images	27.56/0.78	29.70/0.79	28.61/0.83	31.57/0.84	29.73/0.87	33.55/0.89	30.80/0.90	35.77/0.92
Sample Dataset 1	38.56/0.93	37.12/0.91	39.99/0.95	37.70/0.92	41.82/0.97	40.85/0.96	43.54/0.98	43.06/0.97
Sample Dataset 2	34.05/0.89	35.54/0.87	35.49/0.92	36.40/0.89	37.09/0.94	39.04/0.94	38.97/0.96	41.32/0.96
Sample Dataset 3	32.38/0.83	31.20/0.82	33.80/0.86	32.62/0.86	35.16/0.89	32.26/0.90	36.80/0.93	35.88/0.93
Sample Dataset 4	29.99/0.86	33.04/0.86	30.32/0.88	35.22/0.90	33.06/0.92	37.25/0.94	34.60/0.94	35.56/0.95
Sample Dataset 5	26.44/0.83	29.09/0.84	27.48/0.86	31.25/0.88	28.83/0.90	35.72/0.92	29.98/0.93	36.72/0.94

Fig. 4. Diffraction-limited original images (top), reconstructed spectral images with isotropic TV regularization (middle), and reconstructed spectral images with learned-reconstruction method (bottom) for dataset 2 at 25 dB SNR. (Courtesy of NASA/SDO and the AIA, EVE, and HMI science teams.) For each case, false-color representations are also provided on the right.

are mostly smooth and have few edges (such as dataset 1) TV reconstruction performs better. This may also result from using limited number of such examples in the training phase of the learning-based method.

To visually evaluate the results for images with different amount of structure, we provide the reconstructed spectral images for the datasets 2 and 5 in Fig. 4 and 5, respectively. For comparison, the diffraction-limited versions of the original images are also given. Here SNR is taken as 25 dB. These also illustrate that the developed imaging system provides near diffraction-limited resolution. Visual inspection shows that important characteristic features, such as solar flares, are successfully recovered. The learned reconstruction method is more successful in this respect than the TV reconstruction, but the TV reconstruction performs closely to the learned reconstruction for the dataset 2. However, even for these mostly smooth spectral images, TV reconstruction introduces some blur for the rapidly varying spatial structures.

TABLE III
PSNR (dB) AND SSIM VALUES FOR DIFFERENT INPUT SNRS WHEN *MOVING DETECTOR* (MD) AND *FIXED DETECTOR* (FD) MEASUREMENT SETTINGS ARE USED. SNR AND PSNR VALUES ARE REPORTED IN DECIBEL (DB)

Input SNR	MD PSNR	FD PSNR	MD SSIM	FD SSIM
15	31.98	32.00	0.88	0.88
20	33.43	33.38	0.91	0.91
25	34.75	34.73	0.93	0.93
30	35.82	35.75	0.94	0.94

Another important observation is given in Table III, which illustrates that MD and FD settings provide similar reconstruction quality. Here the solar images in dataset 2 with a larger size of 512×512 are used. As seen, for a given input SNR level, maximum differences in the reconstructions for the two measurement settings are less than 0.07 dB in PSNR and 0.01 in
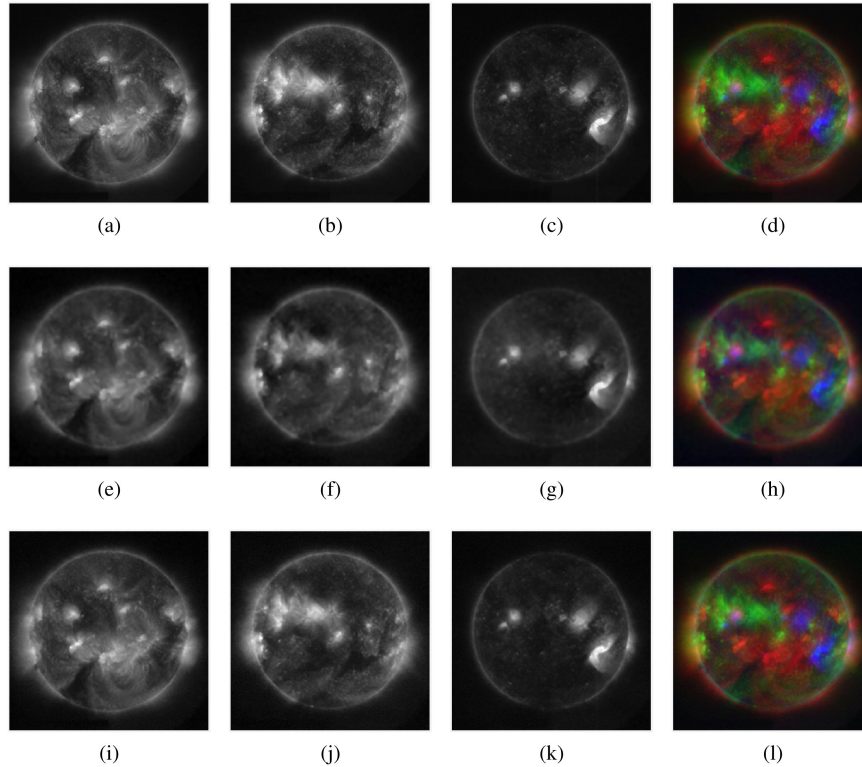
Fig. 5. Diffraction-limited original images (top), reconstructed spectral images with isotropic TV regularization (middle), and reconstructed spectral images with learned-reconstruction method (bottom) for dataset 5 at 25 dB SNR. (Courtesy of NASA/SDO and the AIA, EVE, and HMI science teams.) For each case, false-color representations are also provided on the right.

SSIM. This suggests that one setting can be chosen over the other based on the design requirements of a particular application.

In addition to the measurement noise, other practical challenges should also be considered in real implementation such as the dynamic range of the detector, calibration of the system, or model mismatch. Dynamic range of the chosen monochromatic detector will actually contribute to the noise level as another source of error, but this error can be kept at a desired level to achieve a particular SNR. Likewise, the calibration is also an easy task as the imaging system is shift-invariant (unlike many of the earlier spectral imaging systems). In particular, measuring the PSFs for different wavelengths is sufficient for calibration. Hence there is no need to measure the system response for each voxel in the spectral data cube.

Model mismatch can be, however, more critical in practice. For example, for the moving detector (MD) case, there can be misplacement errors due to the mechanically moving component, resulting in model errors for the assumed PSFs of different spectral bands. Other non-ideal factors during operation (such as in-flight distortions) can also cause similar widening in the PSFs. Hence errors in the assumed PSFs will constitute the main sources of model mismatch. To illustrate the effect of such model errors on the performance, we have performed a sensitivity analysis with varying misplacement errors for the MD setting. Note that in the considered MD setting, the moving detector takes measurements from the distances varying between $d_k = 3.727$ m (f_3) to 3.756 m (f_1), covering a range of 29 mm. In the sensitivity analysis, maximum placement error, $\Delta d_{\max}$, is considered
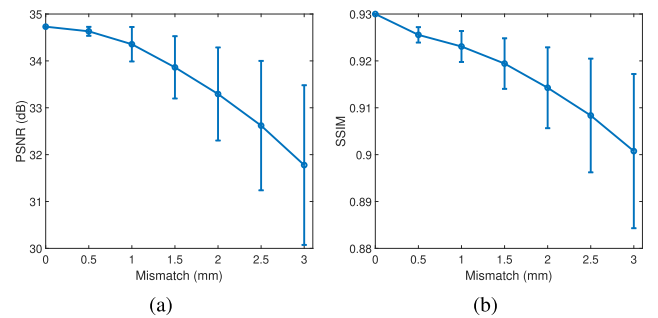

Fig. 6. The average PSNR (a) and SSIM (b) values for the reconstructed spectral images as a function of maximum placement error $\Delta d_{\max}$ in the measurement planes.

between 0 and 3 mm. Then the error in the measurement distance d_k, denoted as Δd, is modeled as a uniformly distributed random variable between $[-\Delta d_{\max}, \Delta d_{\max}]$ for each measurement taken. After performing a Monte-Carlo simulation at 25 dB SNR, the average PSNR and SSIM values for the reconstructed spectral images are obtained in the presence of these measurement errors. The results are given in Fig. 6 as a function of $\Delta d_{\max}$. As seen from this figure, the random misplacement errors up to 3 mm degrades the performance gracefully; hence the reconstructions are mostly robust to such errors. But for larger misplacement errors, it could be better to modify the reconstruction methods to take into account such model mismatch errors (for example, through semi-blind multi-channel deconvolution).

TABLE IV
PSNR (dB) AND SSIM VALUES FOR DIFFERENT NUMBER OF SOURCES (P)
AND MEASUREMENTS (K) FOR SNR = 25 Db

Setting	Average PSNR (dB)	Average SSIM
K=2, P=2	37.55	0.94
K=3, P=3	34.75	0.93
K=4, P=4	31.91	0.91

Another type of model error can arise when the spectral sampling interval is not chosen to be smaller than the spectral bandwidth of the designed diffractive lens. To illustrate this, we have considered Eq. (5) with a continuous spectrum consisting of three Gaussian spectral lines, each with a width slightly larger than the spectral bandwidth. Although the reconstruction was still successful for the three spectral images reconstructed at the central wavelengths, the performance was slightly decreased (36.57 dB PSNR and 0.92 SSIM for dataset 2 with 25 dB SNR). To avoid such issues, one can consider to revise the design of the diffractive lens to decrease its spectral bandwidth. Other options are to use finer sampling for the spectral discretization (i.e. more terms in Eq. (5)) and take more measurements, or train the network with the measurements obtained through a more realistic simulation.

Lastly, we demonstrate the system performance for increased number of spectral bands, P. For this, we also perform simulations for $P = 2$ and $P = 4$ cases by taking same number of measurements as the number of spectral bands (i.e. $K = P$). In the first case, we consider two spectral bands at $\lambda_2 = 33.42$ nm, $\lambda_3 = 33.54$ nm with the measurements obtained at f_2 and f_3 for the MD setting. In the latter case, we consider four spectral bands, namely $\lambda_0 = 33.16$ nm, $\lambda_1 = 33.28$ nm, $\lambda_2 = 33.42$ nm, $\lambda_3 = 33.54$ nm, and obtain four measurements at the corresponding focal planes, f_0, f_1, f_2, and f_3. Table IV shows PSNR and SSIM values for these cases together with $P = 3$ case described before. The results demonstrate that the reconstruction quality gets worse with increasing P. This is expected since the ill-posedness of the inverse problem increases with the number of spectral bands, P. Nevertheless, the system provides high quality reconstructions even for $P = 4$ case, with a PSNR of 31.91 dB and an SSIM of 0.91 at 25 dB input SNR.

B. Resolution Analysis

An important advantage of the developed technique is the unprecedented spectral resolution it enables. The technique offers the capability of separating close-by spectral components that would not otherwise be possible using conventional spectral imagers employing wavelength filters. In the experiments, the sources of interest have wavelengths 33.28 nm, 33.42 nm, and 33.54 nm; hence the developed spectral imager can achieve a spectral resolution of approximately 0.1 nm, which is nearly 0.3% of the central wavelengths. On the other hand, the state-of-the-art EUV wavelength filters at best provide a spectral resolution of 10% of the central wavelength [4]; hence such a high spectral resolution cannot be achieved with the conventional filter-based imagers. This becomes critical especially

when more than one spectral line lie in this 10% range as in the considered EUV spectral imaging application because, using feasible wavelength filters, it is not possible to separately resolve each line.

The obtained PSNR and SSIM values between the reconstructions and diffraction-limited original images also illustrate that the technique offers diffraction-limited high spatial resolution as enabled by the diffractive lenses. Here we further analyze the spatial resolution of the system to quantify its diffraction-limited imaging performance better. For this purpose, we first perform a conditioning-based analysis for the resolution. We also reconstruct point targets separated by various distances to verify that the resolvability of the point targets in these results agree with the diffraction-limited resolution. Each of these analysis provides additional support for the imaging system's diffraction-limited spatial resolution capability. Note that existing EUV spectral imagers cannot achieve diffraction-limited resolution due to surface roughness and figure errors of the used reflective optics [2], [4].

In the conditioning-based resolution analysis, we investigate the conditioning of the forward model in Eq. (9) when the scene consists of point targets. For this, we consider different number of point sources with varying distances between them. If we assume that the locations of the point targets are known, the reconstruction quality will only depend on the columns of $\mathbf{H}$ associated with the locations of the point targets. Thus, by examining the conditioning of the relevant submatrices of $\mathbf{H}$, we will gain an understanding of the resolving capability of the system [75]. In this analysis, we suppose an oracle tells us the exact locations of the point sources in the 3D spectral data cube, which effectively corresponds to knowing the support of the cube a priori. Then, the reconstruction task is to determine the values of these nonzero components. If this problem cannot be solved, the original problem of finding both the locations and values of the point targets will also fail. As a result, the conditioning of the resulting submatrices of $\mathbf{H}$ provides information about the best possible capability of the system for resolving point targets.

Since this is not a conventional camera but a computational imaging system, two-point resolution may not reveal the system performance for more complex scenes. For this reason, we also consider a higher number of point targets than two in our analysis. In particular, we consider 2, 4, 16, 32, and 64 point sources placed in a square grid. We choose the pixel size on the detector as 1 μm for fine analysis of the resolution. We change the spacing between the point sources from 1 μm to 20 μm with 1 μm steps. Then, we calculate the conditioning of the submatrix of $\mathbf{H}$ that contains the columns of $\mathbf{H}$ associated with the locations of these point sources. The results are plotted in Fig. 7 when all point sources are placed in the first, second or third spectral bands for $P = 3$ and MD setting case.

As seen, conditioning is similar for each spectral band and gets worse as the number of point sources increases, or the separation distance between them decreases, as expected. An important observation is the rapid decrease in the condition number around 5 μm separation distance, which indeed corresponds to the theoretical diffraction-limited resolution of the imaging system for the monochromatic case. As the distance gets larger than 5 μm, the change in conditioning becomes small. This observation

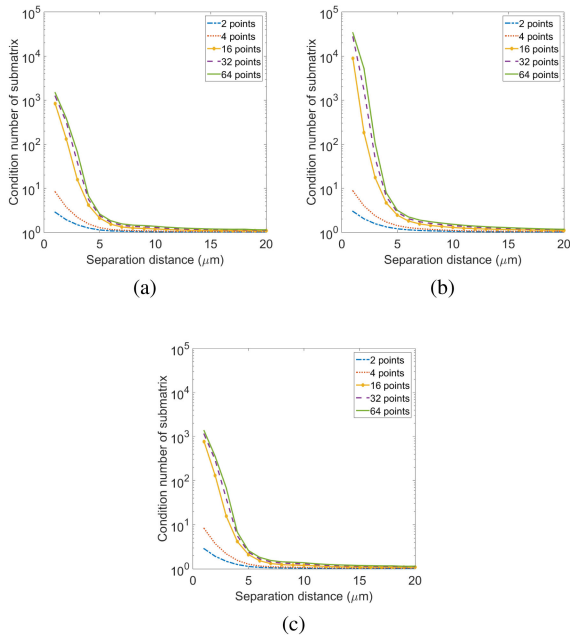Fig. 7. Conditioning of the relevant submatrices of **H** for different number of point sources and separation distances. Conditioning results when the point sources are placed at (a) the first, (b) the second, and (c) the third spectral bands.

Fig. 8. Demonstration of 5 μm resolution using point targets for SNR=25 dB. (a)-(c) Noisy measurements with 25 dB SNR at the focal planes f_1, f_2, and f_3, (d)-(f) reconstructed images from the noisy measurements, (g)-(i) diffraction-limited images, (j)-(l) ground truth images.

agrees with the expected diffraction-limited resolution of the system.

In the reconstruction-based resolution analysis, we simultaneously place 2, 4, and 16 point sources of size 1 μm in the first, second, and third spectral bands, respectively. The separation distance between point sources is chosen as the expected spatial resolution of 5 μm. This data cube is reconstructed from the measurements generated using the forward model in Eq. (9) with 25 dB SNR. Fig. 8 shows the reconstructed and diffraction-limited images for each band, together with the ground truth and measurements. As seen, the imaging system can successfully resolve point sources with 5 μm separation. In fact, the reconstructed images are even sharper than their diffraction-limited versions at this SNR level. This observation is certainly related to our regularization choice (isotropic TV) as well.

By repeating this analysis for different SNRs and separation distances, we also observe that point sources of 5 μm separation can be resolved for SNRs as low as 3 dB. This suggests that the expected theoretical resolution can be achieved even for highly noisy cases. Moreover, resolving point sources with 4 μm spacing is also possible for SNRs as low as 5 dB. That is, the imaging system can even provide spatial resolution beyond the diffraction limit for a wide range of SNR values. However, the system can not resolve point sources even at high SNR levels when the separation distance becomes 3 μm. That is, the system fails to resolve point sources shortly after 5 μm separation distance, as the conditioning starts to degrade significantly. Hence, the reconstruction-based resolution analysis is in agreement with the conditioning-based analysis and also supports the high spatial resolution enabled by the developed multi-spectral imaging technique.
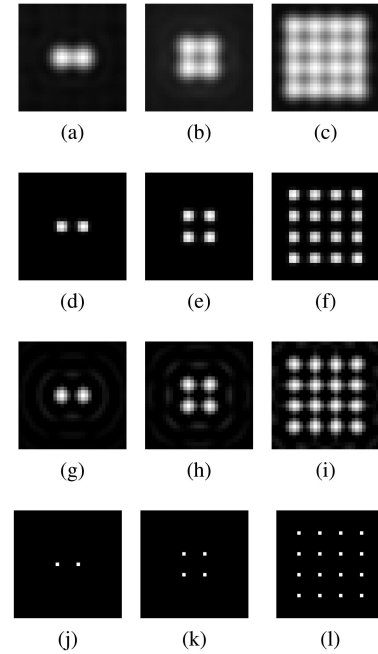
VIII. CONCLUSION

In this paper, we have developed a novel multi-spectral imaging modality that enables higher spatial and spectral resolutions than the conventional scanning-based techniques. Our approach exploited the use of a diffractive lens and powerful image reconstruction to overcome the resolution limits of conventional techniques. The proposed lightweight and low-cost imaging system enables not only unprecedented spectral resolution but also diffraction-limited high spatial resolution for a wide spectral range including UV and x-rays. Spectral imaging with these capabilities has profound impact for a variety of applications. Moreover, the developed learning-based method is capable of adapting the reconstruction (i.e. deep prior) to a particular application. To illustrate these capabilities, the developed computational imaging technique was applied to a potential solar imaging application in the EUV regime. The presented results for EUV spectral imaging illustrate the diffraction-limited spatial resolution and unprecedented spectral resolution enabled with this technique.

The developed imaging technique opens up new possibilities for high-resolution multi-spectral imaging with low-cost and simpler designs. One important application area is the observation of the spectra of space plasmas with new capabilities, as demonstrated in the paper. Lightweight and low-cost diffractive lenses provide diffraction-limited high spatial resolution for a wide spectral range including UV and x-rays. This is not possible with the imaging/collimating optics (such as mirrors and refractive lenses) prevalently used in the other spectral imaging techniques. Another important advantage of the developed technique is the unprecedented spectral resolution it enables. The technique

provides the capability of resolving nearby spectral components that would not otherwise be possible with the wavelength filters used in conventional spectral imagers. For the considered EUV application (which was the initial targeted domain), missions are currently under development. The details of a mission concept based on a distributed solar imager that exploits diffractive lenses and two small satellites flying in formation can be found elsewhere [30], [76].

Our technique is designed to provide high resolution for general scenes including those that lack strong spectral correlation (i.e. has frequent peaks and fast transitions) or have spectra consisting of discrete spectral lines. Spectral sensing of such scenes are of interest in various applications including astrophysical imaging of space plasmas, atmospheric physics, and remote sensing [15], [16]. Note that our approach will naturally work better for spectrally correlated scenes (such as those encountered in the visible regime) using 3D priors to exploit spectral correlation. Sample results in visible regime can be found elsewhere [55], [77].

Lastly we note that the performance of the developed technique can be further improved by taking more measurements (such as at the intermediate planes) and optimizing the measurement configuration or the diffractive lens design [78]. In the presented results, a conventional photon sieve has been used, but the performance may be improved with a diffractive lens specifically designed for this imaging setting. The performance of the learned reconstruction method can also be improved by using different loss functions and priors.

ACKNOWLEDGMENT

REFERENCES

[1] F. S. Oktem, F. Kamalabadi, and J. M. Davila, "High-resolution computational spectral imaging with photon sieves," in *Proc. IEEE Int. Conf. Image Process.*, Paris, France: IEEE, 2014, pp. 5122–5126.

[2] J. M. Davila, "High-resolution solar imaging with a photon sieve," in *Proc. SPIE vol. 8148, Solar Physics and Space Weather Instrumentation IV*, San Diego, CA, USA: 2011, pp. 81480O: 1–11.

[3] G. Arce, D. Brady, L. Carin, H. Arguello, and D. Kittle, "Compressive coded aperture spectral imaging: An introduction," *Signal Process. Mag.*, vol. 31, no. 1, pp. 105–115, 2014.

[4] J. R. Lemen *et al.*, "The atmospheric imaging assembly (AIA) on the solar dynamics observatory (SDO)," in *The Solar Dynamics Observatory*, P. C. Chamberlin, W. D. Pesnell, and B. J. Thompson, Eds., New York, NY, USA, Springer, 2011, pp. 17–40.

[5] G. Andersen, "Large optical photon sieve," *Opt. Lett.*, vol. 30, no. 22, pp. 2976–2978, 2005.

[6] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, "Using deep neural networks for inverse problems in imaging: Beyond analytical methods," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 20–36, Jan. 2018.

[7] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, Feb. 2019.

[8] G. Ongie, A. Jalal, C. A. M. R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," *IEEE J. Sel. Areas Inf. Theory*, vol. 1, no. 1, pp. 39–56, May 2020.

[9] 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. B 44-B51, 2008.

[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.*, vol. 36, no. 14, pp. 2692–2694, 2011.

[11] X. Cao *et al.*, "Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world," *IEEE Signal Process. Mag.*, vol. 33, no. 5, pp. 95–108, Sep. 2016.

[12] E. Salazar, A. Parada-Mayorga, and G. R. Arce, "Spectral zooming and resolution limits of spatial spectral compressive spectral imagers," *IEEE Trans. Comput. Imag.*, vol. 5, no. 2, pp. 165–179, Jun. 2019.

[13] O. F. Kar, U. Kamaci, F. C. Akyon, and F. S. Oktem, "Compressive photon-sieve spectral imaging," in *Proc. OSA Comput. Opt. Sens. Imag.*, Orlando, FL, USA: OSA, 2018, p. CTu 5D.

[14] O. F. Kar and F. S. Oktem, "Compressive spectral imaging with diffractive lenses," *Opt. Lett.*, vol. 44, no. 18, pp. 4582–4585, 2019.

[15] G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics*, Weinheim, Germany: Wiley, 2008.

[16] J. M. Hollas, *Modern Spectroscopy*, Chichester, West Sussex, England: Wiley, 2004.

[17] D. Attwood, *Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Applications*. Cambridge Univ. Press, Cambridge, 2000.

[18] L. Kipp *et al.*, "Sharper images by focusing soft X-rays with photon sieves," *Nature*, vol. 414, no. 6860, pp. 184–188, 2001.

[19] R. Menon, D. Gil, G. Barbastathis, and H. I. Smith, "Photon-sieve lithography," *J. Opt. Soc. Amer. A*, vol. 22, no. 2, pp. 342–345, 2005.

[20] G. Kim, J. A. Domínguez-Caballero, and R. Menon, "Design and analysis of multi-wavelength diffractive optics," *Opt. Exp.*, vol. 20, no. 3, pp. 2814–2823, 2012.

[21] F. S. Oktem, F. Kamalabadi, and J. M. Davila, "Analytical fresnel imaging models for photon sieves," *Opt. Exp.*, vol. 26, no. 24, pp. 32259–32279, 2018.

[22] G. Andersen and D. Tullson, "Broadband antihole photon sieve telescope," *Appl. Opt.*, vol. 46, no. 18, pp. 3706–3708, 2007.

[23] C. Zhou, X. Dong, L. Shi, C. Wang, and C. Du, "Experimental study of a multiwavelength photon sieve designed by random-area-divided approach," *Appl. Opt.*, vol. 48, no. 8, pp. 1619–1623, 2009.

[24] Y. Wang, W. Yun, and C. Jacobsen, "Achromatic fresnel optics for wideband extreme-ultraviolet and X-ray imaging," *Nature*, vol. 424, no. 6944, pp. 50–53, 2003.

[25] X. Zhao, F. Xu, J. Hu, and C. Wang, "Broadband photon sieves imaging with wavefront coding," *Opt. Exp.*, vol. 23, no. 13, pp. 16812–16822, 2015.

[26] P. Wang, N. Mohammad, and R. Menon, "Chromatic-aberration-corrected diffractive lenses for ultra-broadband focusing," *Sci. Rep.*, vol. 6, 2016, Art. no. 21545.

[27] N. Mohammad, M. Meem, B. Shen, P. Wang, and R. Menon, "Broadband imaging with one planar diffractive lens," *Sci. Rep.*, vol. 8, no. 1, pp. 1–6, 2018.

[28] Y. Li, C. Wang, X. Zhao, F. Xu, and C. Wang, "Multispectral and large bandwidth achromatic imaging with a single diffractive photon sieve," *Opt. Exp.*, vol. 26, no. 16, pp. 21141–21152, 2018.

[29] X. Dun, H. Ikoma, G. Wetzstein, Z. Wang, X. Cheng, and Y. Peng, "Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging," *Optica*, vol. 7, no. 8, pp. 913–922, 2020.

[30] J. M. Davila *et al.*, "Milli-arcsecond (MAS) imaging of the solar corona," in *AAS/Solar Phys. Division Abstr. #47, ser. AAS/Solar Physics Division Meeting*, vol. 47, May 2016, pp. 3–10.

[31] O. F. Kar and F. S. Oktem, "Fast computational spectral imaging with a programmable diffractive lens," in *Proc. OSA Comput. Opt. Sens. Imag..*, Munich, Germany: OSA, 2019, pp. CTh 3A-3.

[32] F. D. Hallada, A. L. Franz, and M. R. Hawks, "Fresnel zone plate light field spectral imaging," *Opt. Eng.*, vol. 56, no. 8, pp. 081811: 1–11, 2017.

[33] P. Wang and R. Menon, "Computational multispectral video imaging," *JOSA A*, vol. 35, no. 1, pp. 189–199, 2018.

[34] M. Nimmer, G. Steidl, R. Riesenberg, and A. Wuttig, "Spectral imaging based on 2D diffraction patterns and a regularization model," *Opt. Exp.*, vol. 26, no. 22, pp. 28335–28348, 2018.

[35] D. S. Jeon *et al.*, "Compact snapshot hyperspectral imaging with diffracted rotation," *ACM Trans. Graph.*, vol. 38, no. 4, pp. 117: 1–13, 2019.

[36] R. Le Goff *et al.*, "Multispectral filters assemblies for earth remote sensing imagers," in *Proc. Int. Conf. Space Opt.*2014, vol. 10563. International Society for Optics and Photonics, 2017, Art. no. 1056305.

[37] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, Burlington, MA, USA: Elsevier, 2006.

[38] C. Yi, Y.-Q. Zhao, and J. C.-W. Chan, "Spectral super-resolution for multispectral image based on spectral improvement strategy and spatial preservation strategy," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9010–9024, Nov. 2019.

[39] B. Geelen, N. Tack, and A. Lambrechts, "A snapshot multispectral imager with integrated tiled filters and optical duplication," in *Adv. Fabr. Technol. Micro/Nano Opt. Photon. VI*, vol. 8613. International Society for Optics and Photonics, 2013, p. 861314.

[40] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," *Sensors*, vol. 14, no. 11, pp. 21626–21659, 2014.

[41] Z. Wang and Z. Yu, "Spectral analysis based on compressive sensing in nanophotonic structures," *Opt. Exp.*, vol. 22, no. 21, pp. 25608–25614, 2014.

[42] Z. Wang *et al.*, "Single-shot on-chip spectral sensors based on photonic crystal slabs," *Nat. Commun.*, vol. 10, no. 1, pp. 1–6, 2019.

[43] R. Hahn *et al.*, "Detailed characterization of a mosaic based hyperspectral snapshot imager," *Opt. Eng.*, vol. 59, no. 12, 2020, Art. no. 125102.

[44] F. S. Oktem, L. Gao, and F. Kamalabadi, "Computational spectral and ultrafast imaging via convex optimization," in *Handbook Convex Optim. Methods Imag. Sci.*, Cham, Switzerland: Springer, 2017, pp. 105–127.

[45] T. Okamoto and I. Yamaguchi, "Simultaneous acquisition of spectral image information," *Opt. Lett.*, vol. 16, no. 16, pp. 1277–1279, 1991.

[46] M. Descour and E. Dereniak, "Computed-tomography imaging spectrometer: Experimental calibration and reconstruction results," *Appl. Opt.*, vol. 34, no. 22, pp. 4817–4826, 1995.

[47] B. K. Ford, C. E. Volin, S. M. Murphy, R. M. Lynch, and M. R. Descour, "Computed tomography-based spectral imaging for fluorescence microscopy," *Biophysical J.*, vol. 80, no. 2, pp. 986–993, 2001.

[48] F. S. Oktem, F. Kamalabadi, and J. M. Davila, "A parametric estimation approach to instantaneous spectral imaging," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5707–5721, Dec. 2014.

[49] Y. August, C. Vachman, Y. Rivenson, and A. Stern, "Compressive hyperspectral imaging by random separable projections in both the spatial and the spectral domains," *Appl. Opt.*, vol. 52, no. 10, pp. D 46--D54, 2013.

[50] S.-H. Baek *et al.*, "End-to-End hyperspectral-depth imaging with learned diffractive optics," 2020, *arXiv:2009.00463v2*.

[51] S.-H. Baek, I. Kim, D. Gutierrez, and M. H. Kim, "Compact single-shot hyperspectral imaging using a prism," *ACM Trans. Graph.*, vol. 36, no. 6, pp. 1–12, 2017.

[52] K. Monakhova, K. Yanny, N. Aggarwal, and L. Waller, "Spectral diffusercam: Lensless snapshot hyperspectral imaging with a spectral filter array," *Optica*, vol. 7, no. 10, pp. 1298–1307, 2020.

[53] A. Parada-Mayorga and G. R. Arce, "Spectral super-resolution in colored coded aperture spectral imaging," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 440–455, Dec. 2016.

[54] H. Arguello, H. F. Rueda, and G. R. Arce, "Spatial super-resolution in code aperture spectral imaging," in *Compressive Sens.*, vol. 8365. International Society for Optics and Photonics, 2012, Art. no. 83650A.

[55] O. F. Kar, "Computational spectral imaging techniques using diffractive lenses and compressive sensing," M.S. thesis, *Dept. Elect. Electron. Eng.*, METU, Turkey, 2019.

[56] F. S. Oktem, F. Kamalabadi, and J. M. Davila, "High-resolution solar imaging with photon sieves," in *AGU Fall Meeting Abstracts*, vol. 2014, Dec. 2014, pp. SH 53B-4219.

[57] H.-H. Chung, N. Bradman, M. R. Davidson, and P. H. Holloway, "Dual wavelength photon sieves," *Opt. Eng.*, vol. 47, no. 11, 2008, Art. no. 118001.

[58] A. Ö. Yöntem, J. Li, and D. Chu, "Imaging through a projection screen using bi-stable switchable diffusive photon sieves," *Opt. Exp.*, vol. 26, no. 8, pp. 10162–10170, 2018.

[59] S. Ayazgok and F. S. Oktem, "Snapshot spectral imaging with generalized photon sieves," in *Proc. OSA Imag. Syst. Appl.*, Munich, Germany: OSA, 2019, pp. JW 2A-8.

[60] S. Ayazgok and F. S. Oktem, "Efficient computation of 2D point-spread functions for diffractive lenses," *Appl. Opt.*, vol. 59, no. 2, pp. 445–451, 2020.

[61] R. E. Blahut, *Theory of Remote Image Formation*. Cambridge University Press, Cambridge, 2004.

[62] P. C. Hansen, *Discrete Inverse Problems: Insight and Algorithms*. Philadelphia, PA, USA: SIAM, 2010, vol. 7.

[63] C. R. Vogel, *Computational methods for inverse problems*, Philadelphia, PA, USA: SIAM, 2002, vol. 23.

[64] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.

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

[66] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[67] J. Nocedal and S. Wright, *Numerical Optimization*, New York, NY, USA: Springer, 2006.

[68] M. K. Ng, P. Weiss, and X. Yuan, "Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods," *SIAM J. Sci. Comput.*, vol. 32, no. 5, pp. 2710–2736, 2010.

[69] U. Kamaci, F. C. Akyon, T. Alkanat, and F. S. Oktem, "Efficient sparsity-based inversion for photon-sieve spectral imagers with transform learning," in *Proc. IEEE Glob. Conf. Signal Inf. Process.*, Montreal, Quebec, Canadag, 2017, pp. 1225–1229.

[70] B. Noble and J. W. Daniel, *Applied Linear Algebra*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1988.

[71] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imag. Vis.*, vol. 20, no. 1-2, pp. 89–97, 2004.

[72] P. Brekke, W. Thompson, T. Woods, and F. Eparvier, "The extreme-ultraviolet solar irradiance spectrum observed with the coronal diagnostic spectrometer (CDS) on SOHO," *Astrophysical J.*, vol. 536, no. 2, p. 959, 2000.

[73] P. R. Young *et al.*, "EUV emission lines and diagnostics observed with Hinode/EIS," *Pub. Astrono. Soc. Jpn.*, vol. 59, no. sp3, pp. S 857–S864, 2007.

[74] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations.*, 2015, pp. 1–15.

[75] N. Antipa *et al.*, "Diffusercam: Lensless single-exposure 3D imaging," *Optica*, vol. 5, no. 1, pp. 1–9, 2018.

[76] D. Rabin, A. N. Daw, P. Calhoun *et al.*, "A mission to achieve ultrahigh-resolution imaging of the solar corona," *AGUFM*, vol. 2018, pp. A 44G–07, 2018.

[77] D. Dogan and F. S. Oktem, "Efficient algorithms for convolutional inverse problems in multidimensional imaging," 2020, *arXiv:2006.08303*.

[78] E. Widloski, U. Kamaci, and F. Kamalabadi, "Optimal measurement configuration in computational diffractive imaging," in *Proc. IEEE Int. Conf. Image Process.*, Abu Dhabi, United Arab Emirates, 2020, pp. 281–285.