Reconstruction of Binary Shapes from Blurred Images via Hankel-structured Low-rank Matrix Recovery

Saeed Razavikia, Arash Amini, *Senior Member, IEEE*, and Sajad Daei

***Abstract*—With the dominance of digital imaging systems, we are often dealing with discrete-domain samples of an analog image. Due to physical limitations, all imaging devices apply a blurring kernel on the input image before taking samples to form the output pixels. In this paper, we focus on the reconstruction of binary shape images from few blurred samples. This problem has applications in medical imaging, shape processing, and image segmentation. Our method relies on representing the analog shape image in a discrete grid much finer than the sampling grid. We formulate the problem as the recovery of a rank r matrix that is formed by a Hankel structure on the pixels. We further propose efficient ADMM-based algorithms to recover the low-rank matrix in both noiseless and noisy settings. We also analytically investigate the number of required samples for successful recovery in the noiseless case. For this purpose, we study the problem in the random sampling framework, and show that with $\mathcal{O}(r\log^4(n_1n_2))$ random samples (where the size of the image is assumed to be $n_1 \times n_2$) we can guarantee the perfect reconstruction with high probability under mild conditions. We further prove the robustness of the proposed recovery in the noisy setting by showing that the reconstruction error in the noisy case is bounded when the input noise is bounded. Simulation results confirm that our proposed method outperform the conventional total variation minimization in the noiseless settings.**

***Index Terms*—Binary shape, Hankel structure, Low-rank matrix recovery.**

I. Introduction

The shape images are a subclass of general images that consist of two levels: inside and outside of the shape. Such binary structures appear in a number of applications such as medical imaging [1], millimeter-wave imaging [2], artworks (e.g., cutouts Figure 1) [3], quality monitoring in manufacturing, document analysis [4], astronomical imaging [5] and image segmentation [6]. Due to physical non-idealities of imaging devices, the digital measurements of binary shapes are no longer binary. More explicitly, the binary-valued points of the analog shape image are mixed in a weighted form within the acquisition device before creating the output digital image. Besides, due to the sampling effect of the digital acquisition device, the boundaries of the shape image are not exactly preserved. In summary, the digital image of a binary shape might be very different from its original form. In this paper, we aim at recovering a shape image from its blurred samples.

S. Razavikia and A. Amini are with the Advanced Communication Research Institute (ACRI), Department of Electrical Engineering, Sharif University of Technology (e-mail: saeedrazavikia@ee.sharif.edu, aamini@sharif.edu).

S. Daei is with the Department of Electrical Engineering, Iran University of Science and Technology (e-mail: sajaddaeiomshi@gmail.com).

(a) (b) (c) (d)

Fig. 1: (a) and (c) are two cut-outs made by Henri Matisse. Using black and white representation in (b) and (d), these artworks can be considered as shape images.

A related problem appears in super-resolution imaging. In this technique, a number of low-resolution images of the same scene are combined to achieve an image with higher resolution. Obviously, it is not generally possible to achieve a high-resolution image based on only a single low-resolution image. However, if the input image is known to have a specific structure, this might be possible; the example of blurred images of a cloud of separated points is considered in [7], [8].

The signals with finite rate of innovation (FRI) introduced in [9] for one-dimensional signals and extended to 2D signals in [10] provide a model for continuous-domain signals that can be exactly recovered from their generalized samples. Recently, the FRI structure has been of special interest for modeling the image structures, particularly, for shape images [11]–[17]. One popular model is to express the boundary of the shape image via a bi-variate polynomial or harmonic polynomial.

For instance, in [15], the shape boundaries are assumed to represent the zero-level set of a polynomial with finite degree; the recovery problem is then, formulated as estimating the polynomial coefficients via a set of annihilation equations. A similar approach is devised in [18] when the shape boundaries are expressed as the zero-level set of a trigonometric polynomial (a polynomial function of $e^{j2\pi x}$ and $e^{j2\pi y}$). The number of polynomial coefficients in [15], [18] roughly controls the degrees of freedom in repressing shapes. Specifically, to represent complicated shapes via the models in [15], [18], the degree of the polynomial shall be set quite high.

The annihilation equations are the most common tool in dealing with FRI signals. However, a matrix completion technique for 1D FRI signals has been recently proposed in [19]. This method (called EMAC) uses the Hankel transform (first proposed in [20]) to represent the FRI signal in a low-rank structure. With a slightly different matrix structure (block Toeplitz structure), [13] investigates the recovery of piece-wise constant images using a combination of rank minimization and annihilation equations. This could be considered as an extension of [19] to 2D FRI signals.

In a completely different setting, a blind deconvolution problem is studied in [21], where the discrete signals are subject to unknown blurring kernels. For both 1D and 2D signals, and under certain conditions, it is shown that a rank minimization problem can simultaneously estimate the blurring kernel and the original signal.

A. Contributions

The goal of this work is to recover a discrete binary image from blurred and sub-sampled measurements, where the blurring kernel is assumed to be known. Instead of recovering a continuous-domain image, we reconstruct the shape image in discrete form with an arbitrarily high sampling resolution. This enables us to avoid parametric continuous-domain models; indeed, the stability and robustness of recovering parametric models (by estimating the parameters) cannot be guaranteed in general, particularly, when studying FRI signals.

Our approach is based on the rank minimization of a Hankel structure. More precisely, we use the fact that the gradient of shape images is non-zero only at the boundaries, which implies that the gradient is sparse. Then, we recall a result that the Hankel transform of a signal which is sparse in the Fourier domain, is low-rank. Thus, if we apply the Hankel transform on the Fourier transform of the gradient image, we achieve a low-rank matrix. Hence, we base our image reconstruction on minimizing the rank of this latter matrix. The available measurements, which are linearly related to the shape image, form a set of constraints in the minimization problem. The rank r of the aforementioned matrix describes the complexity of the model, and plays a role similar to the degree of the polynomial in FRI models. However, unlike the FRI techniques, we do not need to know this parameter in advance. Indeed, our reconstruction technique (low-rank recovery) implicitly finds the simplest images that matches the constraints (i.e., measurements).

A simple block diagram of our reconstruction procedure is shown in Figure 2. After forming the final Hankel matrix, we devise a nuclear norm minimization as a convex relaxation of the rank minimization. Using an augmented Lagrangian form for the latter problem, we propose two algorithms based on the alternating direction method of multipliers (ADMM) for the noiseless and noisy cases.

As explained earlier, we assume that the desired high-resolution image generates the available measurements if the image is blurred via the known kernel and then sampled uniformly. This model differs from [13], [15], [18] as we do not take into account the continuous-domain model. While this model is somewhat close to the one used in [21], we should emphasize that we do not apply a random mask nor a random

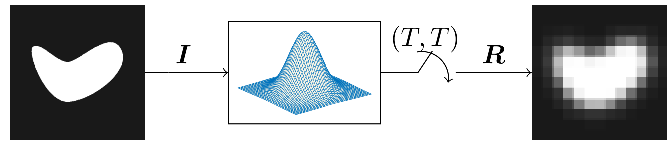


Fig. 2: In practical imaging devices, the physics of data acquisition frequently imply the existence of a 2-D blurring kernel before the sampling process.

sampler (which are used in [21]). It is worth mentioning that we assume to know the blurring kernel in this paper, while [21] estimates the kernel from multiple masked versions of the target image; in simple words, [21] assumes to have access to multiple images, while we are constrained to a single image accompanied with the blurring kernel.

To be able to analytically evaluate the performance of our method, we investigate the number of required random samples for recovering an $n_1 \times n_2$ image that results in a rank r matrix using the explained procedure. We prove that $\mathcal{O}(r \log^4(n_1 n_2))$ measurements are sufficient to guarantee perfect recovery with high probability. A preliminary result in this direction was previously presented in [22].

B. Notation

We use lower and upper-case bold letters (e.g., $\boldsymbol{x}$ and $\boldsymbol{X}$) to denote vectors and matrices, respectively. Calligraphic notations such as $\mathcal{X}$ represent linear operators. For a linear operator $\mathcal{X}$, $\mathcal{X}^*$ stands for the adjoint operator. We show the transpose and Hermitian of a matrix $\boldsymbol{X}$ by $\boldsymbol{X}^{\mathrm{T}}$ and $\boldsymbol{X}^{\mathrm{H}}$, respectively. With $\boldsymbol{X}$ a given matrix, $\|\boldsymbol{X}\|$, $\|\boldsymbol{X}\|_{\mathrm{F}}$ and $\|\boldsymbol{X}\|_*$ represent the spectral, Frobenius, and nuclear norms (sum of singular values), respectively. The norm of a linear operator $\mathcal{A} : \mathbb{C}^{n_1 \times n_2} \mapsto \mathbb{C}^{n_1 \times n_2}$ is defined as

$$\|\mathcal{A}\|_{\mathrm{op}} := \sup_{\boldsymbol{X} \in \mathbb{C}^{n_1 \times n_2}} \frac{\|\mathcal{A}(\boldsymbol{X})\|_{\mathrm{F}}}{\|\boldsymbol{X}\|_{\mathrm{F}}}.$$

Also $\|\boldsymbol{X}\|_1$ and $\|\boldsymbol{X}\|_\infty$ indicate the element-wise ℓ_1 and ℓ_∞ norms of $\boldsymbol{X}$ (treating $\boldsymbol{X}$ as a vector). We show the i th canonical basis vector by $\boldsymbol{e}_i$. For an integer n, $[n]$ represents $\{1, 2, \ldots, n\}$. We associate a 2D Hankel operator $\mathscr{H}_{d_1,d_2} : \mathbb{C}^{n_1 \times n_2} \mapsto \mathbb{C}^{d_1 d_2 \times (n_1-d_1+1)(n_2-d_2+1)}$ to the scalars $d_1 \in [n_1]$ and $d_2 \in [n_2]$ as

$$\mathscr{H}_{d_1,d_2}(\boldsymbol{X}) := \begin{bmatrix} \mathfrak{h}(\boldsymbol{x}_1) & \mathfrak{h}(\boldsymbol{x}_2) & \ldots & \mathfrak{h}(\boldsymbol{x}_{n_1-d_1+1}) \\ \mathfrak{h}(\boldsymbol{x}_2) & \mathfrak{h}(\boldsymbol{x}_3) & \ldots & \mathfrak{h}(\boldsymbol{x}_{n_1-d_1+2}) \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{h}(\boldsymbol{x}_{d_1}) & \mathfrak{h}(\boldsymbol{x}_{d_1+1}) & \ldots & \mathfrak{h}(\boldsymbol{x}_{n_1}) \end{bmatrix}, \tag{1}$$

where $\boldsymbol{x}_l$ is the lth row of the matrix $\boldsymbol{X}$, and $\mathfrak{h}(\boldsymbol{x}_{1 \times n_2})$ is given by

$$\mathfrak{h}(\boldsymbol{x}) := \begin{bmatrix} x_1 & x_2 & \ldots & x_{n_2-d_2+1} \\ x_2 & x_3 & \ldots & x_{n_2-d_2+2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{d_2} & x_{d_2+1} & \ldots & x_{n_2} \end{bmatrix}. \tag{2}$$

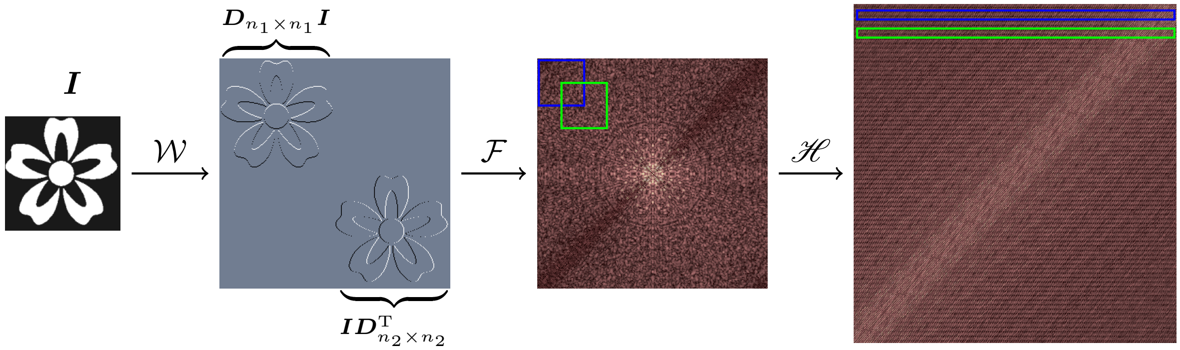


Fig. 3: The overview of the rank minimization of the Hankel format of a shape image. We first take row-wise and column-wise derivatives of the shape image $\boldsymbol{I}$ and form a block-diagonal matrix of double the size of $\boldsymbol{I}$. Next, we apply the 2D Fourier transform followed by the block-Hankel transform. Each block in the Fourier domain is thus, mapped to a row in the Hankel structure.

We also define the pseudo-inverse Hankel mapping $\mathscr{H}^{\dagger}_{d_1,d_2} : \mathbb{C}^{d_1d_2\times(n_1-d_1+1)(n_2-d_2+1)} \mapsto \mathbb{C}^{n_1\times n_2}$ by averaging those elements of the input $d_1d_2\times(n_1-d_1+1)(n_2-d_2+1)$ matrix that are supposed to be equal in a Hankel transform of an arbitrary $n_1\times n_2$ matrix, and then, reordering them to form an $n_1\times n_2$ matrix[1]. To simplify the notations, we often omit the subscript d_1, d_2 from $\mathscr{H}$ and $\mathscr{H}^{\dagger}$.

C. Organization of the paper

The rest of the paper is organized as follows. In Section II, we explain the image model and study the Hankel structure in details. Then, we explicitly define the sampling problem. Next, we present reconstruction algorithms for both the noiseless and noisy scenarios. The implementation of the algorithms is achieved via augmented Lagrangian method in Section III. In Section IV, we present numerical simulations and experimental results for both the noiseless and noisy cases. Finally, in Section V we provide theoretical guarantees for the recovery of binary shape using a random sampling strategy.

II. Shape Images and measurement procedure

A. Shape Images

In this work, we consider a discrete bi-level (black and white) image I of size $n_1\times n_2$ (or $\boldsymbol{I}\in\{0,1\}^{n_1\times n_2}$ using matrix representation) with the interior index set $S\subset[n_1]\times[n_2]$ defined as

$$\underset{(a,b)\in[n_1]\times[n_2]}{I[a,b]} = \begin{cases}1, & (a,b)\in S,\\ 0, & (a,b)\notin S,\end{cases} \tag{3}$$

where $\Theta=[n_1]\times[n_2]$ is the image plane. We assume to know a linear transform $\mathcal{W}:\mathbb{C}^{n_1\times n_2}\to\mathbb{C}^{\widetilde{n}_1\times\widetilde{n}_2}$ such that $\mathcal{W}(\boldsymbol{I})$ is sparse ($\widetilde{n}_i\geq n_i$). The transforms related to image derivatives are among important examples. To better clarify, let us define the finite difference matrix

$$\boldsymbol{D}_{m\times m} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0\\ 0 & 1 & -1 & \dots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \dots & 1 & -1\\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}_{m\times m}.$$

It is now easy to check that $\boldsymbol{D}_{n_1\times n_1}\boldsymbol{I}$ and $\boldsymbol{I}\boldsymbol{D}^{\mathrm{T}}_{n_2\times n_2}$ represent the vertical and horizontal derivatives of I, respectively. For common shape images, the non-zero elements of I are expected to be clustered such that $\boldsymbol{D}_{n_1\times n_1}\boldsymbol{I}$ and $\boldsymbol{I}\boldsymbol{D}^{\mathrm{T}}_{n_2\times n_2}$ are sparse (small number of non-zero elements compared to $n_1\times n_2$). Therefore, one can choose $\mathcal{W}(\boldsymbol{I})=\boldsymbol{D}_{n_1\times n_1}\boldsymbol{I}$ or $\mathcal{W}(\boldsymbol{I})=\boldsymbol{I}\boldsymbol{D}^{\mathrm{T}}_{n_2\times n_2}$, or a combination of both such as

$$\mathcal{W}(\boldsymbol{X}_{n_1\times n_2}) = \begin{bmatrix} \boldsymbol{D}_{n_1\times n_1}\boldsymbol{X} & \boldsymbol{0}_{n_1\times n_2}\\ \boldsymbol{0}_{n_1\times n_2} & \boldsymbol{X}\boldsymbol{D}^{\mathrm{T}}_{n_2\times n_2}\end{bmatrix}. \tag{4}$$

However, our model in this paper is general and we can use any invertible $\mathcal{W}$ (such as wavelets). Our goal in this paper is to recover I by using the sparse nature of $\mathcal{W}(\boldsymbol{I})$.

In our problem, the available data are a number of samples from a blurred version of I. Mathematically, a $(2L_1+1)\times(2L_2+1)$ smoothing kernel φ acts on I, and results in the blurred image $\widetilde{I}$ with the same size $(n_1\times n_2)$:

$$\widetilde{I}[i-L_1\,,\,j-L_2] = \sum_{(a,b)\in\Theta} I[a,b]\varphi[i-a,j-b] \qquad \forall i=L_1+[n_1], j=L_2+[n_2]. \tag{5}$$

The blurred image is also represented via $\widetilde{\boldsymbol{I}}\in\mathbb{R}^{n_1\times n_2}$. Finally, we assume to observe a subset $\Omega\in\Theta$ of the elements in $\widetilde{\boldsymbol{I}}$, which amounts to the sampling process. In other words, the available data can be modeled as a masked version $\boldsymbol{R}$ of $\widetilde{\boldsymbol{I}}$:

$$\boldsymbol{R}=\mathcal{P}_{\Omega}(\widetilde{\boldsymbol{I}}) = \begin{cases} \widetilde{I}[a,b], & (a,b)\in\Omega,\\ 0, & (a,b)\notin\Omega.\end{cases} \tag{6}$$

A practical case is when Ω stands for a uniform downsampling of Θ. In practice, due to the model mismatch and some non-

[1]One can verify that $\mathscr{H}_{d_1,d_2}\mathscr{H}^{\dagger}_{d_1,d_2}$ is the orthogonal projection onto the subspace of all Hankel matrices.

idealities such as quantization, the observed data is oftentimes contaminated with additive noise

$$\boldsymbol{R}_n = \mathcal{P}_\Omega\big(\widetilde{\boldsymbol{I}} + \boldsymbol{E}\big), \tag{7}$$

where $\boldsymbol{E} \in \mathbb{R}^{n_1 \times n_2}$ stands for the noise matrix; we assume the availability of an upper bound η on the Frobenius norm of the noise matrix, i.e. $\|\mathcal{P}_\Omega(\boldsymbol{E})\|_{\mathrm{F}} \leq \eta$. In summary, we assume the observations are given either by $\boldsymbol{R}$ or $\boldsymbol{R}_n$, where the smoothing kernel φ and the sampling subset Ω are available beforehand (η is also known in the noisy case).

For the sake of simplicity, we assume a separable smoothing kernel φ in this paper:

$$\underset{\substack{1\leq a\leq L_1\\ 1\leq b\leq L_2}}{\varphi[a,b]} = \phi(b)\cdot\psi(a). \tag{8}$$

With this assumption, the 2D convolution in (5) can be restated as

$$\widetilde{\boldsymbol{I}} = \boldsymbol{\Phi}\boldsymbol{I}\boldsymbol{\Psi}^{\mathrm{T}}, \tag{9}$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{n_1\times n_1}$ and $\boldsymbol{\Psi} \in \mathbb{R}^{n_2\times n_2}$ are Toeplitz matrices linked with (truncated) 1D convolution operators via $\phi[\cdot]$ and $\psi[\cdot]$ kernels, respectively. It should be noted that if the blurring kernel is replaced with the Fourier transform (i.e. $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are the DFT matrices), then, our measurement model (6) coincides with the MRI samples in [13], [18], [23]). Also, with $\boldsymbol{\Phi} = \mathbf{id}_{n_1\times n_1}$ and $\boldsymbol{\Psi} = \mathbf{id}_{n_2\times n_2}$, our measurements simplify to the exact samples of the image.

B. Block Hankel Structure

Let $\boldsymbol{X} = [X_{a,b}]_{a,b}$ be an $n_1 \times n_2$ matrix with elements of the form

$$X_{a,b} = \sum_{k=1}^{r} \exp\big(-\mathrm{j}([b,a]\cdot\boldsymbol{\omega}_k)\big), \tag{10}$$

where $\boldsymbol{\omega}_1,\dots,\boldsymbol{\omega}_r$ are arbitrary real-valued 2D vectors (2D frequencies). Indeed, $\boldsymbol{X}$ represents the Fourier transform of a sum of r discrete delta functions. It is shown in [20], that the rank of the Hankel transform $\mathscr{H}(\boldsymbol{X})$ is at most r. It is further shown in [24] that the Hankel transform of a matrix is low-rank if and only if the inverse 2D Fourier transform of the matrix is sparse. Hence, the sparse structure of a 2D matrix can be interpreted as the low-rank structure of the Hankel transform of its Fourier transform. In this work, we apply this result for the recovery of binary images for which the gradient has a sparse structure (when the shape boundary is small compared to the whole image). To simplify the notations, we define the operator $\mathcal{HF}_{\mathcal{W}}(\cdot) : \mathbb{C}^{n_1\times n_2} \mapsto \mathbb{C}^{d_1 d_2\times(\widetilde{n}_1-d_1+1)(\widetilde{n}_2-d_2+1)}$ as

$$\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}) := \mathscr{H}\Big(\mathcal{F}\big(\mathcal{W}(\boldsymbol{X})\big)\Big), \tag{11}$$

where $\mathcal{W} : \mathbb{C}^{n_1\times n_2} \mapsto \mathbb{C}^{\widetilde{n}_1\times\widetilde{n}_2}$ is the sparsifying transform with $\widetilde{n}_1 \geq n_1$ and $\widetilde{n}_2 \geq n_2$, and $\mathcal{F} : \mathbb{C}^{\widetilde{n}_1\times\widetilde{n}_2} \mapsto \mathbb{C}^{\widetilde{n}_1\times\widetilde{n}_2}$ stands for the 2D Fourier transform. The partial derivative operator $\mathcal{W} : \mathbb{C}^{n_1\times n_2} \mapsto \mathbb{C}^{2n_1\times 2n_2}$ of the form

$$\mathcal{W}(\boldsymbol{X}_{n_1\times n_2}) = \begin{bmatrix} \boldsymbol{D}_{n_1\times n_1}\boldsymbol{X} & \mathbf{0}_{n_1\times n_2} \\ \mathbf{0}_{n_1\times n_2} & \boldsymbol{X}\boldsymbol{D}^{\mathrm{T}}_{n_2\times n_2} \end{bmatrix} \tag{12}$$

is of particular interest in this paper. For now, however, we proceed with a general invertible $\mathcal{W}$ operator. The pseudo-inverse of $\mathcal{HF}_{\mathcal{W}}$ operator is then, defined by

$$\mathcal{HF}^{\dagger}_{\mathcal{W}}(\boldsymbol{Y}) := \mathcal{W}^{\dagger}\Big(\mathcal{F}^{-1}\big(\mathscr{H}^{\dagger}(\boldsymbol{Y})\big)\Big), \tag{13}$$

where $\boldsymbol{Y}$ is an arbitrary matrix of size $d_1 d_2 \times (\widetilde{n}_1 - d_1 + 1)(\widetilde{n}_2 - d_2 + 1)$, $\mathcal{F}^{-1} : \mathbb{C}^{\widetilde{n}_1\times\widetilde{n}_2} \mapsto \mathbb{C}^{\widetilde{n}_1\times\widetilde{n}_2}$ is the inverse 2D Fourier transform, and $\mathcal{W}^{\dagger} : \mathbb{C}^{\widetilde{n}_1\times\widetilde{n}_2} \mapsto \mathbb{C}^{n_1\times n_2}$ is the pseudo-inverse of $\mathcal{W}$. For the particular case in (12), we have that

$$\mathcal{W}^{\dagger}\Bigg(\begin{bmatrix} \boldsymbol{P}_{n_1\times n_2} & \boldsymbol{Q}_{n_1\times n_2} \\ \boldsymbol{R}_{n_1\times n_2} & \boldsymbol{S}_{n_1\times n_2} \end{bmatrix}\Bigg) = \frac{\boldsymbol{D}^{-1}_{n_1\times n_1}\boldsymbol{P} + \boldsymbol{S}(\boldsymbol{D}^{-1}_{n_2\times n_2})^{\mathrm{T}}}{2}. \tag{14}$$

Figure 3 illustrates the $\mathcal{HF}_{\mathcal{W}}$ operator with the specified $\mathcal{W}$ in (12).

III. Recovery From Blurred Samples

Let $\boldsymbol{I} \in \{0,1\}^{n_1\times n_2}$ be a binary shape image from which we have the noiseless observations $\boldsymbol{R} = \mathcal{P}_\Omega(\boldsymbol{\Phi}\boldsymbol{I}\boldsymbol{\Psi}^{\mathrm{T}})$. To take advantage of the low-rank structure in $\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X})$, one can recover $\boldsymbol{I}$ via

$$\begin{aligned} \widehat{\boldsymbol{I}} = \underset{\boldsymbol{X}\in\{0,1\}^{n_1\times n_2}}{\operatorname{argmin}} \quad & \operatorname{rank}\big(\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X})\big) \\ \text{s.t.} \quad & \mathcal{P}_\Omega(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}) = \boldsymbol{R}. \end{aligned} \tag{15}$$

As the feasible set $\boldsymbol{X} \in \{0,1\}^{n_1\times n_2}$ is discrete, and the $\operatorname{rank}(\cdot)$ function is non-convex and non-smooth, the optimization in (15) is NP-hard in general. A common alternative is to replace the $\operatorname{rank}(\cdot)$ with its convex relaxation, i.e., the nuclear norm [25]–[28], and enlarge the feasible set to $\boldsymbol{X} \in [0,1]^{n_1\times n_2}$:

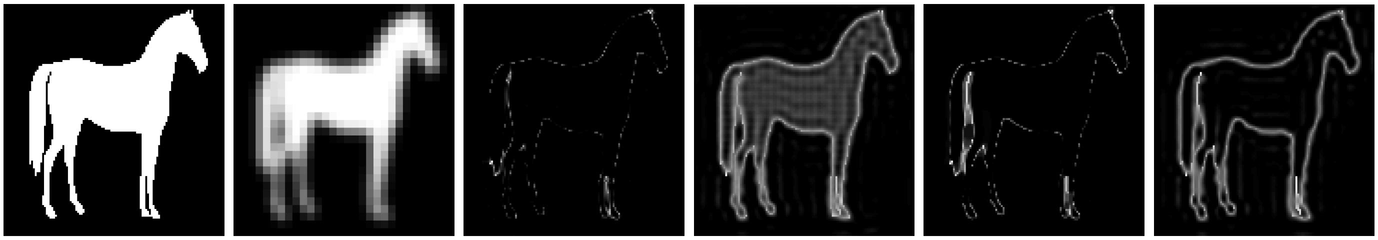
$$\begin{aligned} \widehat{\boldsymbol{I}} = \underset{\boldsymbol{X}\in[0,1]^{n_1\times n_2}}{\operatorname{argmin}} \quad & \big\|\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X})\big\|_* \\ \text{s.t.} \quad & \mathcal{P}_\Omega(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}) = \boldsymbol{R}. \end{aligned} \tag{16}$$

The problem (16) is convex. In the noisy case, the latter optimization can be rewritten as

$$\begin{aligned} \widehat{\boldsymbol{I}} = \underset{\boldsymbol{X}\in[0,1]^{n_1\times n_2}}{\operatorname{argmin}} \quad & \big\|\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X})\big\|_* \\ \text{s.t.} \quad & \big\|\mathcal{P}_\Omega(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}) - \boldsymbol{R}_n\big\|_{\mathrm{F}} \leq \eta, \end{aligned} \tag{17}$$

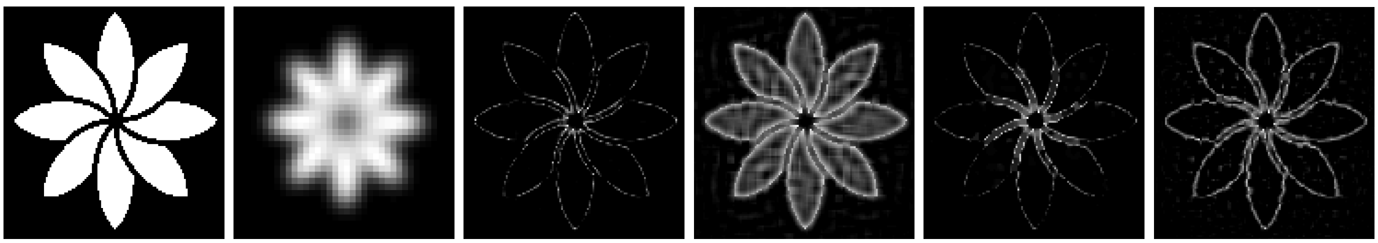
where $\eta > 0$ is an upper-bound on the standard deviation of noise. Next, we propose practical implementations for solving (16) and (17).

The Hankel transform inherently increases matrix dimensions, which in turn increases the computational complexity and memory requirements for solving (16) and (17). To partially alleviate this issue, we solve these problems using an ADMM implementation technique which is rather fast. It is worth noting that there are also other scalable algorithms to solve Hankel-based problems such as the techniques used in [19], [29]–[31].



(a) (b) (c) (d) (e) (f)

Fig. 4: Recovery of shape images. (a) is the original shape image. (b) is the blurred image (of size 30×30) when the sampling kernel is the 2D B-spline of order 2. The absolute error images for the proposed method and the least squares method are shown in (c) and (f), receptively. We have also included the result of TV-minimization (introduced in [3]) and the blind method of [21] (with no random mask) in (e) and (d), respectively. The reconstruction PSNR values for (c), (d), (e) and, (f) are 26.66dB, 15.07dB, 22.68dB, and 17.16dB respectively. The SSIM values are (the same order) 0.98, 0.57, 0.96, 0.63.



(a) (b) (c) (d) (e) (f)

Fig. 5: Recovery of a shape image. (a) is the original shape image. (b) is the blurred image (of size 38×38) when the blurring kernel is a 31×31 Gaussian filter. The absolute difference between the original shape and the recovered image using the proposed method is shown in (c) which corresponds to $\text{PSNR} = 23.15$dB and $\text{SSIM} = 0.97$. The error of the recovered image using the blind method of [21] is shown in (d) which achieves $\text{PSNR} = 13.64$dB and $\text{SSIM} = 0.49$. The results of TV-minimization (introduced in [3]) is depicted in (e) that achieves $\text{PSNR} = 20.54$dB and $\text{SSIM} = 0.94$. The image in (f) represents the error of the least squares solution with $\text{PSNR} = 17.77$dB and $\text{SSIM} = 0.70$.

A. Noiseless Recovery

It is well-known that the nuclear norm can be expressed as [32]:

$$\|\boldsymbol{A}\|_* = \min_{\substack{\boldsymbol{U},\boldsymbol{V} \\ \boldsymbol{A}=\boldsymbol{U}\boldsymbol{V}^{\mathrm{H}}}} \|\boldsymbol{U}\|_{\mathrm{F}}^2 + \|\boldsymbol{V}\|_{\mathrm{F}}^2. \tag{18}$$

Therefore, we can reformulate (16) as

$$\begin{aligned} \underset{\substack{\boldsymbol{U},\boldsymbol{V} \\ \boldsymbol{X}\in[0,1]^{n_1\times n_2}}}{\operatorname{argmin}} \quad & \|\boldsymbol{U}\|_{\mathrm{F}}^2 + \|\boldsymbol{V}\|_{\mathrm{F}}^2 \\ \text{s.t.} \quad & \mathcal{P}_\Omega\big(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}\big) = \boldsymbol{R}, \\ & \mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}) = \boldsymbol{U}\boldsymbol{V}^{\mathrm{H}}. \end{aligned} \tag{19}$$

We first transform the conditional minimization into an augmented Lagrangian form, and then, apply an ADMM technique to achieve the global minimizer. Note that (19) is bilinear in terms of $\boldsymbol{U}$ and $\boldsymbol{V}$, and not necessarily convex. Using a result in [33], we know that the ADMM converges in this case, when the penalty parameter of ADMM is set sufficiently large. It is not difficult to verify that except some constant terms (which do not affect the minimizer), the augmented Lagrangian form can be written as

$$\begin{aligned} L_{\mu_1,\mu_2}(\boldsymbol{U},\boldsymbol{V},\boldsymbol{X},\boldsymbol{\Lambda}_1,\boldsymbol{\Lambda}_2) = \|\boldsymbol{U}\|_{\mathrm{F}}^2 + \|\boldsymbol{V}\|_{\mathrm{F}}^2 \\ + \mu_1\|\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}) - \boldsymbol{U}\boldsymbol{V}^{\mathrm{H}} + \boldsymbol{\Lambda}_1\|_{\mathrm{F}}^2 \\ + \mu_2\|\mathcal{P}_\Omega(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}) - \boldsymbol{R} - \boldsymbol{\Lambda}_2\|_{\mathrm{F}}^2, \end{aligned} \tag{20}$$

where μ_1, μ_2 are arbitrary positive scalars (above the threshold required for convergence) and $\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2$ are the Lagrange multipliers in form of matrices that match the size of $\mathcal{HF}_{\mathcal{W}}(\boldsymbol{I})$ and $\boldsymbol{I}$, respectively. We minimize the unconditional cost L_{μ_1,μ_2} using an ADMM approach; i.e., we sequentially update $\boldsymbol{X},\boldsymbol{U},\boldsymbol{V},\boldsymbol{\Lambda}_1,\boldsymbol{\Lambda}_2$. More precisely, let $\boldsymbol{X}^{(k)}$,$\boldsymbol{U}^{(k)}$, $\boldsymbol{V}^{(k)}$, $\boldsymbol{\Lambda}_1^{(k)}$ and $\boldsymbol{\Lambda}_2^{(k)}$ be the achieved matrices until the kth iteration. The updated matrices at the end of the $(k+1)$th iteration are given by

$$\begin{aligned} \boldsymbol{X}^{(k+1)} = \underset{\boldsymbol{X}\in[0,1]^{n_1\times n_2}}{\operatorname{argmin}} \ & \mu_1\|\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}) - \boldsymbol{U}^{(k)}\boldsymbol{V}^{(k)\mathrm{H}} + \boldsymbol{\Lambda}_1^{(k)}\|_{\mathrm{F}}^2 \\ & \mu_2\|\boldsymbol{R} - \mathcal{P}_\Omega(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}) + \boldsymbol{\Lambda}_2^{(k)}\|_{\mathrm{F}}^2, \end{aligned} \tag{21}$$

$$\begin{aligned} \boldsymbol{U}^{(k+1)} =& \arg\min_{\boldsymbol{U}} \|\boldsymbol{U}\|_{\mathrm{F}}^2 + \\ & \mu_1\|\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}^{(k+1)}) - \boldsymbol{U}\boldsymbol{V}^{(k)\mathrm{H}} + \boldsymbol{\Lambda}_1^{(k)}\|_{\mathrm{F}}^2, \end{aligned} \tag{22}$$

$$\begin{aligned} \boldsymbol{V}^{(k+1)} =& \arg\min_{\boldsymbol{V}} \|\boldsymbol{V}\|_{\mathrm{F}}^2 + \\ & \|\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}^{(k+1)}) - \boldsymbol{U}^{(k+1)}\boldsymbol{V}^{\mathrm{H}} + \boldsymbol{\Lambda}_1^{(k)}\|_{\mathrm{F}}^2, \end{aligned} \tag{23}$$

$$\boldsymbol{\Lambda}_1^{(k+1)} = \boldsymbol{\Lambda}_1^{(k)} + \mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}^{(k+1)}) - \boldsymbol{U}^{(k+1)}\boldsymbol{V}^{(k+1)\mathrm{H}} \tag{24}$$

$$\boldsymbol{\Lambda}_2^{(k+1)} = \boldsymbol{\Lambda}_2^{(k)} + \boldsymbol{R} - \mathcal{P}_\Omega(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}). \tag{25}$$

Except for (24) and (25) that have simple update rules, we need to derive the minimizers. Fortunately, we are dealing with

TABLE I: Simplified operators when $\mathcal{W}$ is given by (12) and $\Omega = \{R, 2R, 3R, \ldots, \lfloor \frac{n_1}{R} \rfloor R\} \times \{C, 2C, 3C, \ldots, \lfloor \frac{n_2}{C} \rfloor C\}$ which corresponds to a uniform subsampling.

Operator	Matrix Representation	Details of Calculation
$\mathscr{H}^*(\boldsymbol{Y})$		$[\mathscr{H}^*(\boldsymbol{Y})]_{i,j} = \langle \boldsymbol{Y}, \mathscr{H}(\boldsymbol{e}_i \boldsymbol{e}_j^{\mathrm{T}}) \rangle, \quad \forall\ (i,j) \in [n_1] \times [n_2]$
$\mathscr{H}^*(\mathscr{H}(\boldsymbol{X}))$	$\boldsymbol{K}_{\mathrm{r}} \boldsymbol{X} \boldsymbol{K}_c^T$	$\boldsymbol{K}_{\mathrm{r}} = \mathrm{diag}(1, 2, \cdots, \underbrace{d_1, \cdots, d_1}_{(2n_1 - 2d_1 + 2)}, d_1 - 1, \cdots, 1),$ $\boldsymbol{K}_{\mathrm{c}} = \mathrm{diag}(1, 2, \cdots, \underbrace{d_2, \cdots, d_2}_{(2n_2 - 2d_2 + 2)}, d_2 - 1, \cdots, 1)$
$\mathcal{W}^*(\boldsymbol{X})$	$\boldsymbol{D}_{n_1 \times n_1}^{\mathrm{T}} \boldsymbol{P} + \boldsymbol{S} \boldsymbol{D}_{n_2 \times n_2}$	$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{P}_{n_1 \times n_2} & \boldsymbol{Q}_{n_1 \times n_2} \\ \boldsymbol{R}_{n_1 \times n_2} & \boldsymbol{S}_{n_1 \times n_2} \end{bmatrix} \in \mathbb{C}^{2n_1 \times 2n_2}$
$\mathcal{HF}_{\mathcal{W}}^* \mathcal{HF}_{\mathcal{W}}(\boldsymbol{X})$	$\boldsymbol{D}_{n_1 \times n_1}^{\mathrm{T}} (\boldsymbol{M}_{\mathrm{r}1} \boldsymbol{D}_{n_1 \times n_1} \boldsymbol{X} \boldsymbol{M}_{\mathrm{r}1}^{\mathrm{T}} + \boldsymbol{M}_2 \boldsymbol{X} \boldsymbol{D}_{n_2 \times n_2}^{\mathrm{T}} \boldsymbol{M}_2^{\mathrm{T}})$ $+ (\boldsymbol{M}_{\mathrm{c}3} \boldsymbol{D}_{n_1 \times n_1} \boldsymbol{X} \boldsymbol{M}_{\mathrm{c}3}^{\mathrm{T}} + \boldsymbol{M}_{\mathrm{c}4} \boldsymbol{X} \boldsymbol{D}_{n_2 \times n_2}^{\mathrm{T}} \boldsymbol{M}_{\mathrm{c}4}^{\mathrm{T}}) \boldsymbol{D}_{n_2 \times n_2}$	$\boldsymbol{M}_{\mathrm{r}} = \boldsymbol{F}_{\mathrm{r}}^{\mathrm{T}} \boldsymbol{K}_{\mathrm{r}} \boldsymbol{F}_{\mathrm{r}}, \qquad \boldsymbol{M}_{\mathrm{c}} = \boldsymbol{F}_{\mathrm{c}}^{\mathrm{T}} \boldsymbol{K}_{\mathrm{c}} \boldsymbol{F}_{\mathrm{c}},$ $\boldsymbol{F}_{\mathrm{r}} \in \mathbb{C}^{2n_1 \times 2n_1}$ and $\boldsymbol{F}_{\mathrm{c}} \in \mathbb{C}^{2n_2 \times 2n_2}$ are discrete Fourier transform matrices $\boldsymbol{M}_{\mathrm{r}} = \begin{bmatrix} \boldsymbol{M}_{\mathrm{r}1} & \boldsymbol{M}_{\mathrm{r}2} \\ \boldsymbol{M}_{\mathrm{r}3} & \boldsymbol{M}_{\mathrm{r}4} \end{bmatrix}$ *and* $\boldsymbol{M}_{\mathrm{c}} = \begin{bmatrix} \boldsymbol{M}_{\mathrm{c}1} & \boldsymbol{M}_{\mathrm{c}2} \\ \boldsymbol{M}_{\mathrm{c}3} & \boldsymbol{M}_{\mathrm{c}4} \end{bmatrix} \in \mathbb{C}^{2n_1 \times 2n_2}$
$\mathcal{J}(\boldsymbol{X})$	$\boldsymbol{J}_{\mathrm{r}} \boldsymbol{X} \boldsymbol{J}_c^{\mathrm{T}}$	$\boldsymbol{J}_{\mathrm{r}} = \mu_1 (\boldsymbol{D}_{n_1 \times n_1}^{\mathrm{T}} \boldsymbol{M}_{\mathrm{r}1} \boldsymbol{D}_{n_1 \times n_1} + \boldsymbol{D}_{n_1 \times n_1}^{\mathrm{T}} \boldsymbol{M}_{\mathrm{r}2} + \boldsymbol{M}_{\mathrm{c}3} \boldsymbol{D}_{n_1 \times n_1} + \boldsymbol{M}_{\mathrm{c}4}) + \mu_2 (\widetilde{\boldsymbol{\Phi}}^{\mathrm{T}} \widetilde{\boldsymbol{\Phi}})$ $\boldsymbol{J}_{\mathrm{c}} = \mu_1 (\boldsymbol{D}_{n_1 \times n_1}^{\mathrm{T}} \boldsymbol{M}_{\mathrm{c}4} \boldsymbol{D}_{n_1 \times n_1} + \boldsymbol{D}_{n_1 \times n_1}^{\mathrm{T}} \boldsymbol{M}_{\mathrm{c}3} + \boldsymbol{M}_{\mathrm{r}2} \boldsymbol{D}_{n_1 \times n_1}^{\mathrm{T}} + \boldsymbol{M}_{\mathrm{r}1}) + \mu_2 (\widetilde{\boldsymbol{\Psi}}^{\mathrm{T}} \widetilde{\boldsymbol{\Psi}})$ $(\widetilde{\boldsymbol{\Phi}}^{\mathrm{T}})_i = (\boldsymbol{\Phi}^{\mathrm{T}})_{Ri}, \quad \forall\ i = 1 : \lfloor n_1 / R \rfloor$ $(\widetilde{\boldsymbol{\Psi}}^{\mathrm{T}})_i = (\boldsymbol{\Psi}^{\mathrm{T}})_{Ci}, \quad \forall\ i = 1 : \lfloor n_2 / C \rfloor$

Algorithm 1 Noiseless Matrix Recovery with ADMM

Input:
  Measurement matrix $\boldsymbol{\Phi}, \boldsymbol{\Psi}$.
  Image samples $\boldsymbol{R} \in [0,1]^{n_1 \times n_2}$.
  Augmented Lagrange multiplier parameters μ_1, μ_2
Output:
  Reconstruct Image $\boldsymbol{X} \in [0,1]^{n_1 \times n_2}$.
procedure LOW-RANK RECOVERY($\boldsymbol{R}, \boldsymbol{X}$)
  $\boldsymbol{\Lambda}_1^{(0)} \leftarrow \emptyset$
  $\boldsymbol{\Lambda}_2^{(0)} \leftarrow \emptyset$
  $\boldsymbol{X}^{(0)} \leftarrow \boldsymbol{\Phi}^{\dagger} \boldsymbol{R} \boldsymbol{\Psi}^{\dagger \mathrm{T}}$
  $\boldsymbol{U}^{(0)}, \boldsymbol{V}^{(0)} \leftarrow PolarDecomposition(\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}^{(0)}))$
  for $k = 1 : \mathrm{iter}$ **do**
    Calculate $\boldsymbol{X}^{(k)}$ and project to set with (26)
    Update $\boldsymbol{U}^{(k)}$ within (28)
    Update $\boldsymbol{V}^{(k)}$ within (29)
    $\boldsymbol{\Lambda}_1^{(k)} \leftarrow \boldsymbol{\Lambda}_1^{(k-1)} + \mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}^{(k)}) - \boldsymbol{U}^{(\boldsymbol{k})} \boldsymbol{V}^{(k)\mathrm{H}}$,
    $\boldsymbol{\Lambda}_2^{(k)} \leftarrow \boldsymbol{\Lambda}_2^{(k-1)} + \boldsymbol{R} - \mathcal{P}_{\Omega}(\boldsymbol{\Phi} \boldsymbol{X} \boldsymbol{\Psi}^{\mathrm{T}})$
  end for
  $\widehat{\boldsymbol{I}} \leftarrow \boldsymbol{X}^{(\mathrm{iter})}$
  return $\widehat{\boldsymbol{I}}$
end procedure

convex quadratic forms in all the involved minimizations in (21)-(23), which have closed-form expressions. In particular,

$$\boldsymbol{X}^{(k+1)} = \mathcal{P}_{[0,1]} \Bigg(\mathcal{J}^{-1} \Big(\mu_1 \mathcal{HF}_{\mathcal{W}}^* \big(\boldsymbol{U}^{(k)} \boldsymbol{V}^{(k)\mathrm{H}} - \boldsymbol{\Lambda}_1^{(k)} \big) + \mu_2 \boldsymbol{\Phi} \big(\mathcal{P}_{\Omega}(\boldsymbol{R} + \boldsymbol{\Lambda}_2^{(k)}) \big) \boldsymbol{\Psi}^{\mathrm{T}} \Big) \Bigg), \tag{26}$$

where the linear operator $\mathcal{J} : \mathbb{C}^{n_1 \times n_2} \mapsto \mathbb{C}^{n_1 \times n_2}$ refers to

$$\mathcal{J}(\boldsymbol{X}) = \mu_1 \mathcal{HF}_{\mathcal{W}}^* \mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}) + \mu_2 \boldsymbol{\Phi}^{\mathrm{T}} \mathcal{P}_{\Omega}(\boldsymbol{\Phi} \boldsymbol{X} \boldsymbol{\Psi}^{\mathrm{T}}) \boldsymbol{\Psi}, \tag{27}$$

and the mapping $\mathcal{P}_{[0,1]}$ is given by

$$\mathcal{P}_{[0,1]}(\boldsymbol{Z}) = \begin{cases} 1, & Z_{i,j} > 1, \\ Z_{i,j}, & 1 \geq Z_{i,j} \geq 0, \\ 0, & Z_{i,j} < 0. \end{cases}$$

The update rules of $\boldsymbol{U}$ and $\boldsymbol{V}$ as suggested in (22) and (23) are

$$\boldsymbol{U}^{(k+1)} = \mu_1 \big(\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}^{(k+1)}) + \boldsymbol{\Lambda}_1^{(k)} \big) \boldsymbol{V}^{(k)} \cdot \big(\mathbf{id}_{n_1 \times n_1} + \mu_1 \boldsymbol{V}^{(k)\mathrm{H}} \boldsymbol{V}^{(k)} \big)^{-1}, \tag{28}$$

$$\boldsymbol{V}^{(k+1)} = \mu_1 \big(\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}^{(k+1)}) + \boldsymbol{\Lambda}_1^{(k)} \big) \boldsymbol{U}^{(k+1)} \cdot \big(\mathbf{id}_{n_2 \times n_2} + \mu_1 \boldsymbol{U}^{(k+1)\mathrm{H}} \boldsymbol{U}^{(k+1)} \big)^{-1}, \tag{29}$$

where $\mathbf{id}_{n_1 \times n_1}$ and $\mathbf{id}_{n_2 \times n_2}$ denote the identity matrices of size n_1 and n_2, respectively. The overall procedure of image reconstruction based on noiseless samples is summarized in Algorithm 1. We have simplified the required operators for the special case of $\mathcal{W}$ as in (12) in conjunction with uniform 2D sampling in Table I.

B. Reconstruction from Noisy Samples

When the measurements are noisy, $\mathcal{P}_{\Omega}(\boldsymbol{\Phi} \boldsymbol{X} \boldsymbol{\Psi}^{\mathrm{T}}) = \boldsymbol{R}$ is no longer valid and shall be replaced with $\|\mathcal{P}_{\Omega}(\boldsymbol{\Phi} \boldsymbol{X} \boldsymbol{\Psi}^{\mathrm{T}}) - \boldsymbol{R}_n\|_{\mathrm{F}}^2 \leq \eta^2$, where η^2 stands for an upper-bound on the noise variance. Similar to the noiseless case, we propose the following Lagrangian form for solving (17):

$$\widetilde{L}_{\mu_1,\mu_2}(\boldsymbol{U},\boldsymbol{V},\boldsymbol{X},\boldsymbol{\Lambda}_1) = \|\boldsymbol{U}\|_{\mathrm{F}}^2 + \|\boldsymbol{V}\|_{\mathrm{F}}^2 + \mu_1\|\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X}) - \boldsymbol{U}\boldsymbol{V}^{\mathrm{H}} + \boldsymbol{\Lambda}_1\|_{\mathrm{F}}^2 + \mu_2\|\mathcal{P}_{\Omega}\big(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}\big) - \boldsymbol{R}_n\|_{\mathrm{F}}^2. \tag{30}$$

Again μ_1 is an arbitrary positive real (above the threshold required for convergence), while μ_2 needs to be set such that $\|\mathcal{P}_{\Omega}\big(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}\big) - \boldsymbol{R}_n\|_{\mathrm{F}}^2 \leq \eta^2$ is satisfied. In our implementation, we initialize μ_2 by a large value, and then gradually decrease it until the minimizer of $\widetilde{L}_{\mu_1,\mu_2}$ fails to satisfy $\|\mathcal{P}_{\Omega}\big(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}\big) - \boldsymbol{R}_n\|_{\mathrm{F}}^2 \leq \eta^2$ for the first time. We should emphasize that due to the inequality constraint in the noisy case, we no longer have the Lagrange multiplier $\boldsymbol{\Lambda}_2$. Due to the similarity to the noiseless case, the update rules for $\boldsymbol{U}$ and $\boldsymbol{V}$ remain unchanged as (28) and (29). However, for $\boldsymbol{X}^{(n+1)}$, we have:

$$\boldsymbol{X}^{(k+1)} = \mathcal{P}_{[0,1]}\Bigg(\mathcal{J}^{-1}\Big(\mu_1\mathcal{HF}_{\mathcal{W}}^{*}\big(\boldsymbol{U}^{(k)}\boldsymbol{V}^{(k)^{\mathrm{H}}} - \boldsymbol{\Lambda}_1^{(k)}\big) + \mu_2\boldsymbol{\Phi}\,\mathcal{P}_{\Omega}\big(\boldsymbol{R}_n\big)\boldsymbol{\Psi}^{\mathrm{T}}\Big)\Bigg), \tag{31}$$

where $\mathcal{J}$ is defined in (27). The update rule for $\boldsymbol{\Lambda}_1$ is again the same as (24). The overall procedure for image recovery from noisy measurements is summarized in Algorithm 2.

Algorithm 2 Noisy Matrix Recovery with ADMM

```
Input:
    Measurement matrices Φ, Ψ.
    Image of Noisy samples  R_n ∈ [0,1]^{n1×n2}.
    Augmented Lagrange multiplier parameters μ1, μ2
Output:
    Reconstruct Image X ∈ [0,1]^{n1×n2}.
procedure LOW-RANK RECOVERY(R_n, X)
    Λ1^(0) ← ∅
    X^(0) ← Φ† R_n Ψ†T
    U^(0), V^(0) ← PolarDecomposition(HF_W(X^(0)))
    for k = 1 : iter do
        Calculate X^(k) and project to set with (31)
        Update U^(k) within (28)
        Update V^(k) within (29)
        Λ1^(k) ← Λ1^(k−1) + HF_W(X^(k)) − U^(k) V^(k)H,
    end for
    Î ← X^(iter)
    return Î
end procedure
```

The computational complexity of our ADMM method is mainly determined by the matrix inversions in (28) and (29). Each matrix inversion involves $\mathcal{O}((\widetilde{n}_1 - d_1 + 1)(\widetilde{n}_2 - d_2 + 1)d_1d_2r + r^3)$ multiplications. With a lesser computational impact, the update for the Lagrange multipliers in (24) requires $\mathcal{O}((\widetilde{n}_1 - d_1 + 1)(\widetilde{n}_2 - d_2 + 1)d_1d_2r)$ scalar multiplications. We further need to compute the inverse of two diagonal (block diagonal) matrices in (26) with the computational cost of $\mathcal{O}(\widetilde{n}_1^3 + \widetilde{n}_2^3)$ multiplications. Besides the computational cost, the variables $\boldsymbol{U}$, $\boldsymbol{V}$, $\mathcal{HF}_{\mathcal{W}}(\boldsymbol{X})$, and the Lagrangian multipliers $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ shall be stored throughout the iterations in our ADMM implementation. The associated memory requirement is at least $(\widetilde{n}_1 - d_1 + 1)(\widetilde{n}_2 - d_2 + 1)r + rd_1d_2 + 2(\widetilde{n}_1 - d_1 + 1)(\widetilde{n}_2 - d_2 + 1)d_1d_2 + n_1n_2$.

IV. Numerical Results

In this section, we present experimental results for both the noiseless and noisy cases. In addition, we consider different blurring kernels such as Gaussian and B-spline filters. We also compare the results against the TV-minimization technique in [3], the blind method in [21] and the least squares solution with which we initialize our ADMM algorithms ($\boldsymbol{X}^{(0)}$ in Algorithms 1 and 2). We should highlight that the method of [21] estimates the blurring kernel from multiple blurred measurements of the same image subject to random binary masks. Here, we restrict the input to all methods (including the blind method of [21]) to a single blurred image. Therefore, the method of [21] needs to struggle with estimating the kernel and the original image with only a single blurred image. We also ignore the random binary mask in our simulations to avoid missing information in the single blurred image. We use both the PSNR and SSIM metrics to check the quality. The binary shape images used in this section are all of size 120×120 pixels (from which we take samples).

As we described earlier, the employed ADMM methods are guaranteed to converge when the penalty parameters are sufficiently large. For this purpose, we use $\mu_1 = 10^5$ and $\mu_2 = 10^{10}$ in our noiseless experiments. In the noisy setting, however, μ_2 also depends on the standard deviation of noise (η). In this case, we initialize μ_2 with 10^{10} and gradually decrease μ_2 until $\left\|\mathcal{P}_{\Omega}(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}) - \boldsymbol{R}_n\right\|_F^2$ exceeds η^2 for the first time (μ_2 shall be the smallest value which guarantees $\left\|\mathcal{P}_{\Omega}(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{T}}) - \boldsymbol{R}_n\right\|_F^2 < \eta^2$).

A. Noiseless Recovery

In our first experiment, we consider a 2nd order B-spline kernel that is scaled to fit in an area of 16×16 pixels. After applying this kernel on the original 120×120 image in Figure 4-(a), we take 30×30 uniform noiseless samples to reach Figure 4-(b) and proceed with the noiseless reconstruction methods. In Figures 4-(c) to 4-(f) we plot the reconstruction error for our method (with $\mathcal{W}$ as in (12)), the method in [21], TV-minimization and the least-squares approach. Both the PSNR and SSIM metrics reveal that our proposed method has the best performance; the superiority of our reconstruction is particularly visible in places where small areas of one color is surrounded by the other color.

In Figure 5, we consider Gaussian blurring kernel. We use 31×31 Gaussian kernel for Figure 5, while we take 38×38 noiseless samples. Again, the results (with $\mathcal{W}$ as in (12)) confirm superiority of the proposed method compared to the TV technique in reconstruction the details.

To study the role of $\mathcal{W}$ in our experiments, we use a 7-layer Haar wavelet transform in Figure 6. The blurring kernel here is a 2nd order B-spline of size 19×19 and the size of the blurred image after sampling is 24×24. The results indicate

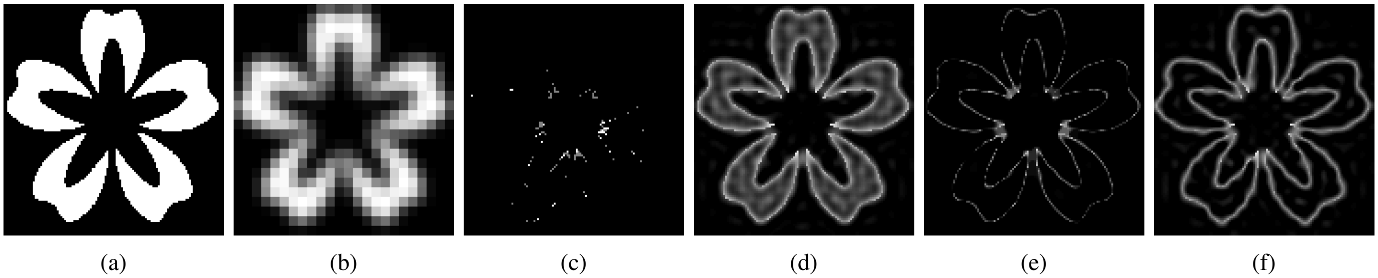


Fig. 6: Recovery of a shape image with Haar wavelet as the sparsifying transform. (a) is a binary image with size 120×120. (b) is the blurred image (of size 24×24) when the blurring kernel is the 2nd order B-spline of size 19×19. The absolute error of the recovered image using our method with Haar wavelet is shown in (c) which has $\mathrm{PSNR} = 24.36\mathrm{dB}$ and $\mathrm{SSIM} = 0.97$. The result of TV minimization is located in (e) and has $\mathrm{PSNR} = 21.46\mathrm{dB}$ and $\mathrm{SSIM} = 0.94$. The images in (d) and (f) represent the absolute error of the blind method of [21] and the least squares solution, respectively. Their PSNR values are $14.37\mathrm{dB}$ and $15.66\mathrm{dB}$, and SSIM values are 0.58 and 0.54, respectively.

that our method still performances very well as the wavelet transform properly sparsifies the image.

Next, we investigate the performance of our method in recovering a non-binary piece-wise constant image in Figure 7. The TV minimization is a well-studied approach for recovering such images. Interestingly, the partial derivatives of piece-wise constant images are sparse and our theory still holds even though the images are non-binary. We have applied a Gaussian blurring kernel of size 31×31 in Figure 7 to obtain a noiseless image of size 38×38 after sampling. As expected, our method performs suitably in recovering the original image with a performance superior to the TV-minimization.

Our final noiseless experiment is dedicated to a part of a Matisse artwork in Figure 8 from Figure 1. Note that in this case the white area is not fully contained within the plane boundaries. In other words, the boundary itself is composed of black and white parts. One can check that the proposed rank-minimization technique is not affected by this issue and still performs suitably.

For a more user-friendly presentation of the results in Figures 4 to 8, the PSNR and SSIM values of each method is reported in Table II.

Remark 1. The performance of our proposed method is directly linked with the nuclear norm (or rank) of the Hankel transform $\mathcal{HF}\big(\mathcal{W}(\mathbf{I})\big)$. While this value could be upper-bounded by the sparsity level of $\mathcal{W}(\mathbf{I})$, the support pattern of $\mathcal{W}(\mathbf{I})$ is also important. For instance, it is shown in [13] that if the edges of the shape could be described by low degree trigonometric polynomials, the rank of $\mathcal{HF}\big(\mathcal{W}(\mathbf{I})\big)$ is better approximated by the degree of the polynomial rather than the length of the edge (sparsity level).

B. Noisy Recovery

In this part, we investigate the effect of additive noise on the performance of the reconstruction method. We present the results for two experiments in Figure 9. More precisely, Figure 9-(a) shows a noiseless 120×120 binary image from which we have a set of 54×54 blurred and noisy samples as in Figure 9-(b). The blurring kernel is a 41×41 Gaussian filter and the noise level is such that $\mathrm{PSNR} = 20.25\mathrm{dB}$ (the PSNR is measured against the blurred but noise-free samples). We use the recovery method in (31), where μ_2 is set with the assumption of $\mathrm{PSNR} = 19\mathrm{dB}$; i.e., we assume an upper-bound on the noise variance in our recovery. The reconstruction error shown in Figure 9-(c) reveals that the proposed method has a descent performance; only the edges of the shape are slightly miss-located. The second experiment on the shape image in Figure 9-(d) also confirms this observation. For this experiment, we applied a 2nd order B-spline blurring kernel of size 12×12 to obtain the 30×30 noisy samples depicted in 9-(e); the level of the noise is similar to the first experiment ($\mathrm{PSNR} = 20.8\mathrm{dB}$).

V. Theoretical Guarantee

Our image reconstruction method relies on a rank minimizing problem. Therefore, it is important to check the conditions under which we can guarantee that the solution to the minimization problem corresponds to the ground truth image. Obviously, such conditions depend on the shape of the image, the smoothing kernel and the number of samples. In this section, we state a bound for the number of noiseless samples that guarantees perfect recovery. Unlike the practical scenario which consists of uniform image sampling, we assume a random sampling strategy to take advantage of the probabilistic recovery guarantees. In other words, we derive a lower-bound on the number of random noiseless samples (for the given shape image and the smoothing kernel) which results in perfect image recovery with high probability.

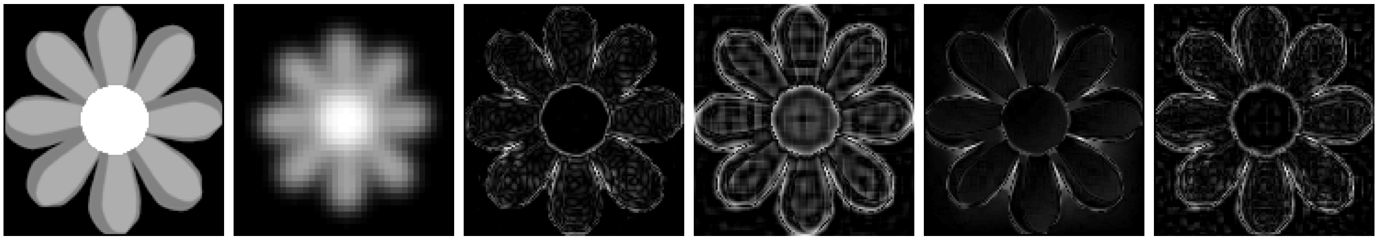
Before we state the bound, we need to introduce a few definitions. First, we define the Dirichlet kernel corresponding to $\mathscr{H}_{d_1,d_2}$ for an input of size $\widetilde{n}_1 \times \widetilde{n}_2$ as

$$\mathcal{D}(d_1, d_2, \boldsymbol{\kappa}) := \frac{1}{d_1 d_2} \frac{\sin(\pi d_1 \kappa_1)}{\sin(\pi \kappa_1)} \frac{\sin(\pi d_2 \kappa_2)}{\sin(\pi \kappa_2)}, \qquad (32)$$

where $\boldsymbol{\kappa} = (\kappa_1, \kappa_2) \in [\widetilde{n}_1] \times [\widetilde{n}_2]$. Let $\boldsymbol{\Xi} \in \mathbb{C}^{\widetilde{n}_1 \times \widetilde{n}_2}$ be a matrix with r nonzero elements located at $\{\boldsymbol{\kappa}_i\}_{i=1}^r$. We define $r \times r$ Gram matrices $\boldsymbol{G}_L(\boldsymbol{\Xi})$ and $\boldsymbol{G}_R(\boldsymbol{\Xi})$ as

$$\big(\boldsymbol{G}_L(\boldsymbol{\Xi})\big)_{i,j} = \mathcal{D}(d_1, d_2, \boldsymbol{\kappa}_i - \boldsymbol{\kappa}_j), \qquad (33)$$

$$\big(\boldsymbol{G}_R(\boldsymbol{\Xi})\big)_{i,j} = \mathcal{D}(n_1 - d_1 + 1, n_2 - d_2 + 1, \boldsymbol{\kappa}_i - \boldsymbol{\kappa}_j). \qquad (34)$$

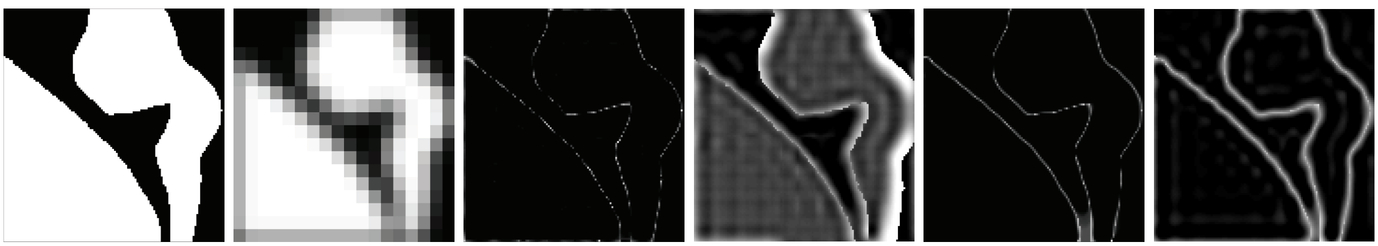


(a) (b) (c) (d) (e) (f)

Fig. 7: Recovery of a piece-wise constant shape image. (a) is a multi-level image with size 120×120. (b) is the blurred image (of size 38×38) when the blurring kernel is the Gaussian filter of size 31×31. The absolute difference between the original image and the recovered image using the proposed method is shown in (d) which corresponds to $\mathrm{PSNR} = 27.27\mathrm{dB}$ and $\mathrm{SSIM} = 0.86$. The result of TV-minimization, the blind method of [21], and the least squares method are depicted in (e), (d), and (f) that achieve PSNR values of $24.14\mathrm{dB}$, $21.61\mathrm{dB}$, $24.37\mathrm{dB}$ and SSIM values of 0.75, 0.63, 0.65, respectively.

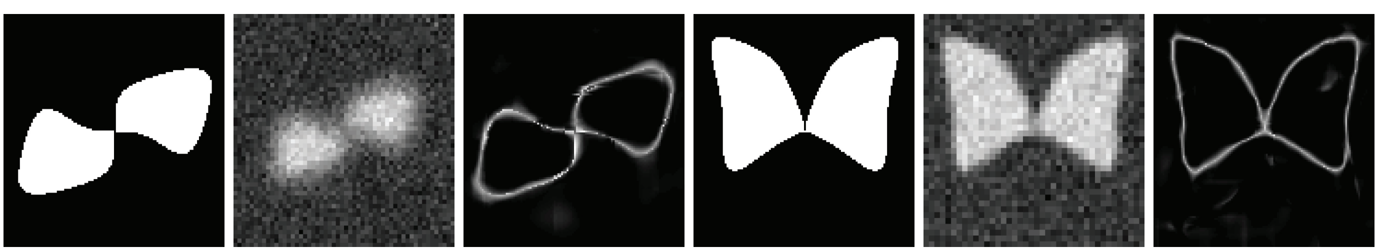
TABLE II: Reconstruction quality of the methods in terms of PSNR and SSIM metrics (noiseless samples).

Figure	PSNR Metric (dB)				SSIM Metric			
	Proposed method	**TV-minimization [3]**	**Blind** [21]	**Least Square**	**Proposed method**	**TV-minimization [3]**	**Blind** [21]	**Least Square**
4	**26.66**	22.68	15.07	17.16	**0.98**	0.96	0.57	0.63
5	**23.15**	20.54	13.64	17.77	**0.97**	0.94	0.49	0.70
6	**24.36**	21.46	14.37	15.66	**0.97**	0.94	0.58	0.54
7	**27.27**	24.14	21.61	24.37	**0.86**	0.75	0.63	0.65
8	**25.38**	23.06	13.12	17.29	**0.98**	0.96	0.49	0.57



(a) (b) (c) (d) (e) (f)

Fig. 8: Recovery of a shape image. (a) is a part of Matisse artwork depicted in Figure 1 with size 120×120. Unlike the previous experiments, the white region is not fully surrounded by the black region. (b) is the blurred image (of size 18×18) when the blurring kernel is the 2nd order B-spline of size 27×27. The Absolute difference between the original shape and the recovered images of the proposed method, TV-minimization [3], the blind method of [21], and the least squares method are shown in (c), (e), (d), and (f), respectively. Their PSNR values are $25.38\mathrm{dB}$, $23.06\mathrm{dB}$, $13.12\mathrm{dB}$, $17.29\mathrm{dB}$, respectively, and SSIM values are 0.98, 0.96, 0.49, 0.57 (the same order).



(a) (b) (c) (d) (e) (f)

Fig. 9: Recovery from noisy samples. (a) and (d) are the original shape images with size 120×120. (b) represents 54×54 noisy and blurred measurements of (a) by applying a Gaussian blurring kernel of size 41×41 and including additive noise level of $\mathrm{PSNR} = 20.25\mathrm{dB}$. (e) represents 30×30 noisy and blurred measurements of (d) by applying the 2nd order B-spline blurring kernel of size 12×12 and including additive noise level of $\mathrm{PSNR} = 20.8\mathrm{dB}$. (c) and (f) depict the absolute differences between the original shapes and the recovered ones using the proposed method in (31); for setting μ_2 (and in turn η), we have assumed an upper-bound for the noise variance corresponding to $\mathrm{PSNR} = 19\mathrm{dB}$. The recovered images based on (b) and (e) correspond to $\mathrm{PSNR} = 18.64\mathrm{dB}$ and $\mathrm{PSNR} = 20.17\mathrm{dB}$, and $\mathrm{SSIM} = 0.76$ and $\mathrm{SSIM} = 0.76$, respectively.

Definition 1. (*Incoherence measure*) Let $\boldsymbol{I} \in \{0,1\}^{n_1 \times n_2}$ be a discrete shape image for which $\mathcal{W}(\boldsymbol{I})$ is an r-sparse matrix. The incoherence measure of $\boldsymbol{I}$ in correspondence with $\mathscr{H}_{d_1,d_2}$ is defined as

$$\rho_1(\boldsymbol{I}; \mathcal{W}, d_1, d_2) = \max\left\{\tfrac{1}{\sigma_{\min}(\boldsymbol{G}_L(\mathcal{W}(\boldsymbol{I})))}, \tfrac{1}{\sigma_{\min}(\boldsymbol{G}_R(\mathcal{W}(\boldsymbol{I})))}\right\}. \tag{35}$$

Definition 2. (Kernel parameter) For a smoothing kernel associated with $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$, we define a kernel parameter in correspondence with $\mathcal{HF}_{\mathcal{W}}$ as

$$\begin{aligned}\rho_2(\boldsymbol{\Phi}, \boldsymbol{\Psi}; \mathcal{W}, d_1, d_2) = \\ \max_{(i,j)\in[n_1]\times[n_2]} \Big\{ \big\|\mathcal{HF}_{\mathcal{W}}^{\dagger *}(\Upsilon_1^{(i,j)})\big\|^2 \big\|\mathcal{HF}_{\mathcal{W}}(\Upsilon_2^{(i,j)})\big\|_1^2, \\ \big\|\mathcal{HF}_{\mathcal{W}}^{\dagger *}(\Upsilon_1^{(i,j)})\big\|_{\mathrm{F}}^2 \big\|\mathcal{HF}_{\mathcal{W}}(\Upsilon_2^{(i,j)})\big\|_{\mathrm{F}}^2 \Big\}, \end{aligned} \tag{36}$$

where

$$\Upsilon_1^{(i,j)} = \left(\boldsymbol{\Phi}^{\mathrm{T}}\right)_i \left(\left(\boldsymbol{\Psi}^{\mathrm{T}}\right)_j\right)^{\mathrm{T}}, \quad \Upsilon_2^{(i,j)} = \left(\boldsymbol{\Phi}^{-1}\right)_i \left(\left(\boldsymbol{\Psi}^{-1}\right)_j\right)^{\mathrm{T}}, \tag{37}$$

with $(\cdot)_i$ representing the ith columns of the matrix.

The proof of the following results are provided in a separate file as supplementary material.

Theorem 1. *Let $\boldsymbol{I} \in \{0,1\}^{n_1 \times n_2}$ be a shape image such that $\mathcal{W}(\boldsymbol{I})$ is r-sparse. We observe samples from this image according to* (6) *where the elements in Ω are drawn uniformly at random from $[n_1] \times [n_2]$. If*

$$|\Omega| > c\rho_1\rho_2 r \frac{n_1\, n_2}{d_1\, d_2} \log^4(n_1 n_2), \tag{38}$$

where ρ_1 and ρ_2 stand for $\rho_1(\boldsymbol{I}; \mathcal{W}, d_1, d_2)$ and $\rho_2(\boldsymbol{\Phi}, \boldsymbol{\Psi}; \mathcal{W}, d_1, d_2)$, respectively, d_1, d_2 are the dimensions used in the Hankel operator, and

$$c = \max\left\{\tfrac{112}{3}(b_1+1),\, 3b_2^2\right\}, \tag{39}$$

($b_1 \geq 2$ and $b_2 \geq 4$), then, $\boldsymbol{I}$ is the unique solution to (16) *with probability no less than $1-(n_1n_2)^{\max\{2-b_1, 4-b_2\}}$.*

Remark 2. The signal model in [20] ignores the blurring kernel and assumes that the available samples are in the Fourier domain of a sparse signal. We show that our result in Theorem 1 matches [20, Theorem 1] in this special case. Indeed, ρ_2 is the distinguishing factor between our bound and the one in [20]. First note that $\mathcal{W}$ is equal to $\mathcal{F}^{-1}$ in this special case, as the signal is sparse and the available samples are in the Fourier domain (i.e., $\mathcal{F}^{-1}$ is the sparsifying transform). This implies that $\mathcal{HF}_{\mathcal{W}} \equiv \mathscr{H}$. Besides, the lack of a blurring kernel can be modeled via $\boldsymbol{\Phi} = \mathbf{id}_{n_1\times n_1}$ $\boldsymbol{\Psi} = \mathbf{id}_{n_2\times n_2}$. With these choices, $\Upsilon_1^{(i,j)}$ and $\Upsilon_2^{(i,j)}$ in (37) both simplify to a $n_1 \times n_2$ matrix filled with 0 except for the (i,j) element which is 1; let us denote this matrix with $\Upsilon^{(i,j)}$. We can now check that $\mathcal{HF}_{\mathcal{W}}(\Upsilon^{(i,j)})$ is also formed by 0s and 1s. The number of non-zero elements in the latter matrix depends on i, j; we represent this number with $N_{i,j}$. In addition, each row and column of $\mathcal{HF}_{\mathcal{W}}(\Upsilon^{(i,j)})$ contains at most one non-zero element. We can verify that $\mathcal{HF}_{\mathcal{W}}^{\dagger *}(\Upsilon^{(i,j)}) = \frac{1}{N_{i,j}}\mathcal{HF}_{\mathcal{W}}(\Upsilon^{(i,j)})$. It is now easy to check that

$$\begin{aligned} \big\|\mathcal{HF}_{\mathcal{W}}^{\dagger *}(\Upsilon^{(i,j)})\big\| &= \frac{1}{N_{i,j}}, \quad \big\|\mathcal{HF}_{\mathcal{W}}(\Upsilon^{(i,j)})\big\|_1 = N_{i,j}, \\ \big\|\mathcal{HF}_{\mathcal{W}}^{\dagger *}(\Upsilon^{(i,j)})\big\|_{\mathrm{F}} &= \frac{1}{\sqrt{N_{i,j}}}, \quad \big\|\mathcal{HF}_{\mathcal{W}}(\Upsilon^{(i,j)})\big\|_{\mathrm{F}} = \sqrt{N_{i,j}}. \end{aligned} \tag{40}$$

Thus, ρ_2 in (36) simplifies to 1.

To evaluate the bound in Theorem 1, we consider the shape image in Figure 9-(d) with size 10000×8000 blurred with a 3211×3211 B-spline kernel of order 2. The values of ρ_1, ρ_2 and r in this case are given as 8.03×10^{-6}, 57.77 and 13540, respectively. Using $d_1 = \frac{10000}{2}$ and $d_2 = \frac{8000}{2}$, our bound is equivalent to 27.79 percent of the total number of pixels for a successful recovery with probability at least 0.9894 ($b_1 = 2.25$ and $b_2 = 4.25$).

For the noisy case, we can no longer expect perfect recovery. Instead, we bound the reconstruction error in Theorem 2 linearly in terms of the input noise level.

Theorem 2. *Let $\boldsymbol{I} \in \{0,1\}^{n_1 \times n_2}$ be a shape image such that $\mathcal{W}(\boldsymbol{I})$ is r-sparse. We observe samples from this image according to* (7)*, where the noise term $\boldsymbol{E}$ satisfies $\|\mathcal{P}_\Omega(\boldsymbol{E})\|_{\mathrm{F}} \leq \eta$, and the elements in Ω are drawn uniformly at random from $[n_1] \times [n_2]$. If* (38) *holds, then, any solution $\widehat{\boldsymbol{I}}$ of* (17) *satisfies*

$$\|\mathcal{HF}_{\mathcal{W}}(\boldsymbol{I}) - \mathcal{HF}_{\mathcal{W}}(\widehat{\boldsymbol{I}})\|_{\mathrm{F}} \leq \frac{17 n_1 n_2 \sqrt{\widetilde{n}_1 \widetilde{n}_2}\,\eta}{\sigma_{\min}(\boldsymbol{\Phi})\sigma_{\min}(\boldsymbol{\Psi})} \|\mathcal{HF}_{\mathcal{W}}\|_{\mathrm{op}} \tag{41}$$

with probability at least $1-(n_1n_2)^{\max\{2-b_1, 4-b_2\}}$ ($b_1 \geq 2$ and $b_2 \geq 4$). Here, $\sigma_{\min}(\boldsymbol{\Phi})$ and $\sigma_{\min}(\boldsymbol{\Psi})$ are the minimum singular values of $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$, respectively.

Remark 3. Similar to the arguments in Remark 2, when there is no blurring effect we have that $\boldsymbol{\Phi} = \mathbf{id}_{n_1\times n_1}$, $\boldsymbol{\Psi} = \mathbf{id}_{n_2\times n_2}$. Hence, $\sigma_{\min}(\boldsymbol{\Phi}) = \sigma_{\min}(\boldsymbol{\Psi}) = 1$, $\widetilde{n}_1 = n_1$, $n_2 = \widetilde{n}_2$, and $\|\mathcal{HF}_{\mathcal{W}}\|_{\mathrm{op}} \leq \sqrt{n_1 n_2}$. Thus, (41) simplifies to the same bound derived in [19, Theorem 3].

VI. Conclusion

In this paper, we presented a method to recover binary shape images from sub-sampled and blurred measurements. These images are commonly sparsified by applying wavelets or gradient-related operators. We linked the sparsity level of the image to the rank of a matrix generated by the Hankel transform. Then, we formulated the recovery procedure as a rank-minimization problem. Besides the simulation results, we provided some theoretical guarantees to ensure suitable recovery performances in both noiseless and noisy setups.

References

[1] T. Zhang, J. M. Pauly, and I. R. Levesque, "Accelerating parameter mapping with a locally low rank constraint," *Magnetic resonance in medicine*, vol. 73, no. 2, pp. 655–661, 2015.

[2] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 9, pp. 1581–1592, Sep. 2001.

[3] M. Fatemi, A. Amini, L. Baboulaz, and M. Vetterli, "Shapes from pixels," *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1193–1206, Mar. 2016.

[4] J. Nemeth and P. Balazs, "Restoration of blurred binary images using discrete tomography," in *Advanced Concepts for Intelligent Vision Systems*. Cham: Springer International Publishing, 2013, pp. 80–90.

[5] H. Pan, T. Blu, and M. Vetterli, "Towards generalized fri sampling with an application to source resolution in radioastronomy," *IEEE Transactions on Signal Processing*, vol. 65, no. 4, pp. 821–835, 2017.

[6] N. Alajlan, M. S. Kamel, and G. H Freeman, "Geometry-based image retrieval in binary image databases," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, pp. 1003 – 1013, 07 2008.

[7] V. I. Morgenshtern and E. J. Candes, "Super-resolution of positive sources: The discrete setup," *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, pp. 412–444, 2016.

[8] G. Schiebinger, E. Robeva, and B. Recht, "Superresolution without separation," *Information and Inference: A Journal of the IMA*, vol. 7, no. 1, pp. 1–30, 2017.

[9] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, 2002.

[10] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets strang–fix," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1741–1757, 2007.

[11] S. Poddar, Y. Mohsin, D. Ansah, B. Thattaliyath, R. Ashwath, and M. Jacob, "Manifold recovery using kernel low-rank regularization: application to dynamic imaging," *IEEE Transactions on Computational Imaging*, pp. 1–1, 2019.

[12] Y. Hu, X. Liu, and M. Jacob, "A generalized structured low-rank matrix completion algorithm for mr image recovery," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2018.

[13] G. Ongie, S. Biswas, and M. Jacob, "Convex recovery of continuous domain piecewise constant images from nonuniform fourier samples," *IEEE Transactions on Signal Processing*, vol. 66, no. 1, pp. 236–250, 2017.

[14] H. Pan, T. Blu, and P. L. Dragotti, "Sampling curves with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 458–471, Jan. 2014.

[15] M. Fatemi, A. Amini, and M. Vetterli, "Sampling and reconstruction of shapes with algebraic boundaries," *IEEE Transactions on Signal Processing*, vol. 64, no. 22, pp. 5807–5818, 2016.

[16] H. Mohaghegh Dolatabadi and A. Amini, "A sampling theorem for convex shapes with algebraic boundaries," in *2017 International Conference on Sampling Theory and Applications (SampTA)*, Jul. 2017, pp. 499–503.

[17] S. Razavi-Kia, H. Zamani, and A. Amini, ""ecovery of a class of binary images from fourier samples," in *Sampling Theory and Applications 2019 (SampTA 2019)*, Bordeaux, France, Jul. 2019.

[18] G. Ongie and M. Jacob, "Off-the-grid recovery of piecewise constant images from few fourier samples," *SIAM Journal on Imaging Sciences*, vol. 9, no. 3, pp. 1004–1041, 2016.

[19] J. C. Ye, J. M. Kim, K. H. Jin, and K. Lee, "Compressive sampling using annihilating filter-based low-rank interpolation," *IEEE Transactions on Information Theory*, vol. 63, no. 2, pp. 777–801, 2017.

[20] Y. Chen and Y. Chi, "Robust spectral compressed sensing via structured matrix completion," *IEEE Transactions on Information Theory*, vol. 60, no. 10, 2014.

[21] S. Bahmani and J. Romberg, "Lifting for blind deconvolution in random mask imaging: Identifiability and convex relaxation," *SIAM Journal on Imaging Sciences*, vol. 8, no. 4, pp. 2203–2238, 2015.

[22] S. Razavi-Kia, H. Zamani, and A. Amini, "Sampling and recovery of binary shapes via low-rank structures," in *Sampling Theory and Applications 2019 (SampTA 2019)*, Bordeaux, France, Jul. 2019.

[23] A. Balachandrasekaran, V. Magnotta, and M. Jacob, "Recovery of damped exponentials using structured low rank matrix completion," *IEEE Transactions on Medical Imaging*, vol. 36, no. 10, pp. 2087–2098, Oct. 2017.

[24] K. H. Jin, D. Lee, and J. C. Ye, "A general framework for compressed sensing and parallel mri using annihilating filter based low-rank hankel matrix," *IEEE Transactions on Computational Imaging*, vol. 2, no. 4, pp. 480–495, 2016.

[25] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, p. 717, 2009.

[26] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.

[27] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[28] D. Gross, "Recovering low-rank matrices from few coefficients in any basis," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1548–1566, 2011.

[29] G. Ongie and M. Jacob, "A fast algorithm for convolutional structured low-rank matrix recovery," *IEEE transactions on computational imaging*, vol. 3, no. 4, pp. 535–550, 2017.

[30] J. Min, K. H. Jin, M. Unser, and J. C. Ye, "Grid-free localization algorithm using low-rank hankel matrix for super-resolution microscopy," *IEEE Transactions on Image Processing*, vol. 27, no. 10, pp. 4771–4786, Oct. 2018.

[31] J. P. Haldar, "Low-rank modeling of local k-space neighborhoods (loraks) for constrained mri," *IEEE Transactions on Medical Imaging*, vol. 33, no. 3, pp. 668–681, Mar. 2014.

[32] N. Srebro, "Learning with matrix factorizations [ph. d. thesis]," *Massachusetts Institute of Technology*, 2004.

[33] Z.-Q. L. M. Hong and M. Razaviyayn, "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 337–364, 2016.