A Semi Nonnegative Matrix Factorization Technique for Pattern Generalization in Single-Pixel Imaging — Source link ↗

Florian Rousset, Françoise Peyrin, Nicolas Ducros

Institutions: Claude Bernard University Lyon 1

Related papers:

- Fast Integral Image Estimation at 1% measurement rate

- A robust and computationally efficient motion detection algorithm based on sigma-delta background estimation

- Principled Design and Implementation of Steerable Detectors

- Fast Registration by Boundary Sampling and Linear Programming

- A New Algorithm For Difference Image Analysis

Share this paper: 🟦 🐦 in ✉

A semi nonnegative matrix factorization technique for pattern generalization in single-pixel imaging

Florian Rousset, Françoise Peyrin, Nicolas Ducros

HAL Id: hal-01635461

https://hal.archives-ouvertes.fr/hal-01635461v2

Submitted on 27 Feb 2018

A Semi Nonnegative Matrix Factorization Technique for Pattern Generalization in Single-Pixel Imaging

Florian Rousset, Françoise Peyrin, and Nicolas Ducros

Abstract—A single-pixel camera is a computational imaging device that only requires a single point detector to capture the image of a scene. This device measures the inner product of the scene and the spatial light modulator patterns. The image of the scene can be recovered through post-processing the measurements obtained for a set of different patterns. Independent of the strategy used for image recovery, real acquisitions require the spatial light modulator patterns to be positive. In addition, the dark current measured in the absence of modulation must be rejected. To date, both experimental issues have been addressed empirically. In this paper, we solve these from a general perspective. Indeed, we propose to seek positive patterns that are linear combinations of the desired patterns (with negative values), and the linear transformation matrices are chosen to reject the dark current. We refer to the problem of finding the positive patterns and the linear combinations as 'pattern generalization'. To the best of our knowledge, this is the first time that this problem has been introduced. In addition, we show that pattern generalization can be solved using a semi nonnegative matrix factorization algorithm. The data obtained from simulations demonstrate that our approach performs similarly to or better than conventional methods, while using fewer measurements.

Index Terms—Computational imaging, single-pixel camera, semi nonnegative matrix factorization, positivity constraint, wavelets, adaptive acquisition

I. INTRODUCTION

THE SINGLE-PIXEL CAMERA (SPC) design [1], [2] enables to build small, low-cost, and high-quality imaging devices suitable for a wide range of applications (e.g., remote imaging, hyperspectral imaging, video acquisition) [3]–[15]. A SPC is an optical setup composed of a *spatial light modulator* (SLM) and a single point detector, which measures the inner product of the scene under view with any SLM pattern. Exploiting a sequence of measurements acquired for different SLM patterns, the image of the scene can be recovered. Strategies for image recovery include compressed sensing, basis scan, and adaptive basis scan. The compressed sensing (CS) [16] theory provides an excellent theoretical framework for single-pixel imaging. It is based on the use of random SLM patterns

and ℓ_1-minimization for image recovery [1], [2]. Basis scan consists in acquiring the image in a predefined basis (e.g. Hadamard, Fourier, wavelets) and recovering the image using the inverse transformation [17]–[21]. *Adaptive* basis scan are similar to basis scan, but the most significant basis functions are predicted during the experiment, which allows to consider fewer SLM patterns [22]–[27]. Hybrid methods combining some of the previous techniques have also been investigated [28]–[31].

A. Problem statement

Pattern positivity and dark-current rejection are often addressed jointly by splitting each SLM pattern into its positive and absolute negative parts [24], [27]. Subtracting the two measurements leads to the desired measurement; i.e., the measurement that would have been acquired with the desired SLM pattern having both positive and negative values. While this approach is straightforward, it requires doubling of the number of measurements, which also doubles the total acquisition time. A second approach is pattern shifting. This consists of the addition of the same background value to all of the patterns, such that they all become positive [2]. The desired measurements are then obtained by subtracting a measurement acquired for the background value from all of the measurements. Only one additional measurement is required (i.e., for the background value), although pattern shifting is very sensitive to noise. While these two constraints are of particular importance for experimental data acquisition and image restoration, they have not been addressed from a general perspective yet.

B. Contributions

In this paper, we propose to tackle both of these experimental problems by seeking positive patterns that are linear combinations of the desired SLM patterns. This new problem is referred to as *pattern generalization*. The *ad-hoc* solutions based on pattern splitting or pattern shifting can be seen as two particular solutions of the pattern generalization problem. Pattern generalization is related to semi nonnegative matrix factorization (SNMF); i.e., the problem of factorizing a matrix into a product of two matrices with one having nonnegative entries [32]. By adapting an SNMF algorithm [33], we show that the number of measurements –hence, the acquisition time– can be halved compared to the standard pattern splitting approach. To summarize, our contribution is two-fold:

1) Formulation of the pattern generalization problem, which has never been addressed in single-pixel imaging, to the best of our knowledge;

2) Introduction of an SNMF-like algorithm to efficiently solve this problem.

Note that this paper does not address the image acquisition/ reconstruction problem of single-pixel imaging. Any scheme (e.g., compressive sensing, basis scans, adaptive basis scans) can be used in conjunction with the proposed approach.

A Matlab implementation of our SNMF pattern generalization algorithm is available online [34].

C. Outline

The paper is organized as follows. In Section II, we describe the acquisition by a SPC and the *ad-hoc* techniques used to deal with pattern positivity and dark-current rejection. Section III introduces the problem of pattern generalization, and Section IV proposes an algorithm to solve it. Section V describes our numerical experiments, and the associated data are reported in Section VI. Finally, we discuss our findings in Section VII, and report our conclusions in Section VIII.

II. CONVENTIONAL SPC ACQUISITION

A. SPC acquisition

A SPC consists of a spatial light modulator that is coupled with a single-point detector. A reflective SLM, such as a digital micromirror device, is commonly used. A lens is added to focus the light onto the detector. A SPC acquisition consists in experimentally measuring the inner product of an image and some SLM patterns, sequentially. Let $\mathbf{f} \in \mathbb{R}^{P \times 1}$ be the $N \times N$ image of the scene with $P = N^2$. The power emitted by the light source (e.g., laser, ambient light) is denoted as $N_0 \in \mathbb{R}_+$ where the units are photons per second (ph/s). We denote $\alpha \in \mathbb{R}_+$ (in ph/s) as the dark current; i.e., the signal read by the single detector when $N_0 = 0$ ph/s. The signal m_k (in ph) measured by the single detector during the integration time $\Delta t \in \mathbb{R}_+$ (in s) can be modeled as [1], [2]

$$m_k = \left(N_0 \mathbf{p}_k^\top \mathbf{f} + \alpha \right) \Delta t \tag{1}$$

where $\mathbf{p}_k \in \mathbb{R}_+^{P \times 1}$ is a SLM pattern. The values of $\mathbf{f}$ and $\mathbf{p}$ (no units) both range as $[0, 1]$; i.e., $\mathbf{f} \in [0, 1]^{P \times 1}$ and $\mathbf{p}_k \in [0, 1]^{P \times 1}$. Let $\mathbf{P} = (\mathbf{p}_1, \ldots, \mathbf{p}_K)^\top \in \mathbb{R}_+^{K \times P}$ be the matrix that contains the sequence of K SLM patterns $\mathbf{p}_k$. The measurement vector $\mathbf{m} = (m_1, \ldots, m_K)^\top \in \mathbb{R}_+^{K \times 1}$ that regroups the sequence of measurements m_k of (1) is hence given by the matrix equation

$$\mathbf{m} = (N_0 \mathbf{P} \mathbf{f} + \alpha \mathbf{v}) \Delta t \tag{2}$$

where $\mathbf{v} = (1, \ldots, 1)^\top \in \mathbb{R}^{K \times 1}$.

B. Ad-hoc *methods for experimental constraints*

Different approaches have been proposed to design a set of patterns $\mathbf{P}$ and to recover the image $\mathbf{f}$ from the measurements $\mathbf{m}$. While discussing these approaches is beyond the scope of this paper, it is important to note that most of these have considered patterns with negative values. Patterns can indeed be chosen in bases such as Fourier [19], [20], discrete cosine transform [21], wavelets [22]–[27], and Hadamard [29]. In addition, it is common to assume that there is no dark current,

so that the image formation model classically considered for image restoration is

$$\breve{m} = N_0 \breve{\mathbf{p}}^\top \mathbf{f} \, \Delta t \tag{3}$$

where $\breve{\mathbf{p}} \in \mathbb{R}^{P \times 1}$ is a SLM pattern with positive and negative values. From now on, we will refer to $\breve{\mathbf{p}}$ (resp. $\breve{m}$) as the desired SLM pattern (resp. measurement). Unfortunately such patterns cannot be physically implemented on a SLM, and the absence of dark current is unrealistic. However, two *ad-hoc* methods are (implicitly) used to correct for the two problems.

1) Pattern splitting: This natural method consists of separating each desired pattern $\breve{\mathbf{p}} \in \mathbb{R}^{P \times 1}$ into its positive $\mathbf{p}_+ \in \mathbb{R}_+^{P \times 1}$ and absolute negative $\mathbf{p}_- \in \mathbb{R}_+^{P \times 1}$ parts [24], [27], i.e.:

$$\breve{\mathbf{p}} = \mathbf{p}_+ - \mathbf{p}_- \quad \text{with} \quad \begin{cases} \mathbf{p}_+ = \max\left(\mathbf{0}_P, \breve{\mathbf{p}} \right) \\ \mathbf{p}_- = |\min\left(\mathbf{0}_P, \breve{\mathbf{p}} \right)| \end{cases} \tag{4}$$

where the $\max(.)$ and $\min(.)$ functions are applied to each entry of both vectors, with $\mathbf{0}_P$ as the null vector of size P. Then, the measurements m_+ and m_- acquired using the patterns $\mathbf{p}_+$ and $\mathbf{p}_-$, respectively, are subtracted to give $\breve{m}$. Indeed, we have

$$m_+ - m_- = \left(N_0 \mathbf{p}_+^\top \mathbf{f} + \alpha \right) \Delta t - \left(N_0 \mathbf{p}_-^\top \mathbf{f} + \alpha \right) \Delta t \tag{5}$$

$$= N_0 \left(\mathbf{p}_+ - \mathbf{p}_- \right)^\top \mathbf{f} \, \Delta t \tag{6}$$

$$= \breve{m}. \tag{7}$$

It is important to note that the dark current α cancels out. The drawback of this natural solution is that $K = 2I$ actual SPC measurements are required to get the I desired measurements, which leads to an increased (doubled) total acquisition time or to a lower Δt. This problem is further discussed in Section V.

2) Pattern shifting: A second approach consists of shifting the pattern $\breve{\mathbf{p}} \in \mathbb{R}^{P \times 1}$ toward the positive values [2]. Mathematically,

$$\breve{\mathbf{p}} = \mathbf{p}_{\text{shift}} - \mathbf{p}_{\text{back}} \quad \text{with} \quad \begin{cases} \mathbf{p}_{\text{shift}} = \breve{\mathbf{p}} + \mathbf{p}_{\text{back}} \in \mathbb{R}_+^{P \times 1} \\ \mathbf{p}_{\text{back}} = (b, \ldots, b)^\top \in \mathbb{R}_+^{P \times 1} \end{cases} \tag{8}$$

where b is a background value chosen so that $b \geq |\min\left(\breve{\mathbf{p}} \right)| \in \mathbb{R}_+$. As for the positive/negative separation, the subtraction of the two corresponding SPC measurements gives the desired measurement, while canceling out the dark current:

$$\breve{m} = m_{\text{shift}} - m_{\text{back}}. \tag{9}$$

Contrary to pattern splitting, it is *not* mandatory to double the number of measurements here. Indeed, choosing b large enough, the background value b can be the same for all of the desired measurements. Therefore, if I measurements are desired, only $K = I + 1$ measurement need to be acquired. However, as will be shown later, this method suffers dramatically from noise.

III. FRAMEWORK FOR PATTERN GENERALIZATION

A. Patterns as linear combinations

We denote $\check{\mathbf{P}} = (\check{\mathbf{p}}_1, \ldots, \check{\mathbf{p}}_I)^\top \in \mathbb{R}^{I \times P}$ as the set of I desired patterns and $\check{\mathbf{m}} = (\check{m}_1, \ldots, \check{m}_I)^\top$ as the vector that contains the corresponding set of desired measurements. Applying (3) to the collection of patterns $\check{\mathbf{P}}$ leads to

$$\check{\mathbf{m}} = N_0 \check{\mathbf{P}} \mathbf{f} \, \Delta t. \tag{10}$$

To acquire $\check{\mathbf{m}}$, which is not feasible in practice, our idea was to generalize the approaches described in Section II-B, to look for a collection of positive patterns $\mathbf{P} \in \mathbb{R}_+^{K \times P}$ such that

$$\check{\mathbf{P}} = \mathbf{TP} \tag{11}$$

where $\mathbf{T} \in \mathbb{R}^{I \times K}$ is a transformation matrix. Applying $\mathbf{T}$ to the measurements $\mathbf{m}$ acquired with the patterns $\mathbf{P}$ leads to

$$\mathbf{Tm} = \mathbf{T}(N_0 \mathbf{Pf} + \alpha \mathbf{v}) \, \Delta t \tag{12}$$
$$= (N_0 \check{\mathbf{P}} \mathbf{f} + \alpha \mathbf{Tv}) \, \Delta t \tag{13}$$

We note that the desired measurements $\check{\mathbf{m}}$ of (10) can be obtained by transforming the acquired measurements $\mathbf{m}$ as in (13), provided that

$$\mathbf{Tv} = \mathbf{0}_I \tag{14}$$

where $\mathbf{0}_I = (0, \ldots, 0)^\top \in \mathbb{R}^{I \times 1}$. If (14) is satisfied, then the desired measurements $\check{\mathbf{m}}$ of (10) are directly obtained as

$$\check{\mathbf{m}} = \mathbf{Tm} \tag{15}$$

In summary, for a given $\check{\mathbf{P}} \in \mathbb{R}^{I \times P}$, the pattern generalization problem we address is the following:

$$\text{Find } \mathbf{T} \in \mathbb{R}^{I \times K} \text{ and } \mathbf{P} \in \mathbb{R}^{K \times P} \text{ s.t. } \begin{cases} \check{\mathbf{P}} = \mathbf{TP} \\ \mathbf{P} \geq 0 \\ \mathbf{Tv} = \mathbf{0}_I \end{cases} \tag{16}$$

where $\mathbf{P} \geq 0$ is shorthand for $(\mathbf{P})_{k,n} \geq 0$, $\forall (k, n)$. The framework of the proposed pattern generalization method is presented in Fig. 1.

B. Link with ad-hoc methods

The *ad-hoc* methods described in Section II-B can easily be expressed in the proposed framework. The pattern splitting method described in (4) is equivalent to choosing $\mathbf{T}$ and $\mathbf{P}$ in (16) as

$$\mathbf{T} = \begin{pmatrix} 1 & -1 & 0 & \ldots & & 0 \\ 0 & 0 & 1 & -1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & & \ldots & 0 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{I \times 2I}$$

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1^\top \\ \mathbf{p}_2^\top \\ \vdots \\ \mathbf{p}_{2I-1}^\top \\ \mathbf{p}_{2I}^\top \end{pmatrix} = \begin{pmatrix} \max{(\mathbf{0}_P, \check{\mathbf{p}}_1)}^\top \\ |\min{(\mathbf{0}_P, \check{\mathbf{p}}_1)}|^\top \\ \vdots \\ \max{(\mathbf{0}_P, \check{\mathbf{p}}_I)}^\top \\ |\min{(\mathbf{0}_P, \check{\mathbf{p}}_I)}|^\top \end{pmatrix} \in \mathbb{R}_+^{2I \times P}. \tag{17}$$

The number of SLM patterns is $K_{\text{split}} = 2I$, where I is the number of desired patterns

For the pattern shifting method described in (8), $\mathbf{T}$ and $\mathbf{P}$ are given by the following formulae:

$$\mathbf{T} = \begin{pmatrix} 1 & \ldots & 0 & -\frac{1}{M} & \ldots & -\frac{1}{M} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 1 & -\frac{1}{M} & \ldots & -\frac{1}{M} \end{pmatrix} \in \mathbb{R}^{I \times (I+M)}$$

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1^\top \\ \vdots \\ \mathbf{p}_I^\top \\ \mathbf{p}_{I+1}^\top \\ \vdots \\ \mathbf{p}_{I+M}^\top \end{pmatrix} = \begin{pmatrix} (\check{\mathbf{p}}_1 + \mathbf{p}_{\text{back}})^\top \\ \vdots \\ (\check{\mathbf{p}}_I + \mathbf{p}_{\text{back}})^\top \\ \mathbf{p}_{\text{back}}^\top \\ \vdots \\ \mathbf{p}_{\text{back}}^\top \end{pmatrix} \in \mathbb{R}_+^{(I+M) \times P}. \tag{18}$$

The number of SLM patterns is $K_{\text{shift}} = I + M$ in this case. Setting M to 1 reduces the number of measurements. However, when there are noisy experimental measurements, it is better to average several measurements that are acquired with the same pattern $\mathbf{p}_{\text{back}}$; i.e., to choose $M > 1$.

C. Optimization problem

To find two matrices $\mathbf{T}$ and $\mathbf{P}$ that satisfy (16), a possibility is to seek two matrices that minimize the squared Frobenius norm of $\check{\mathbf{P}}$ minus $\mathbf{TP}$. With addition of the positivity constraint on $\mathbf{P}$ and the condition (14), this leads to the following optimization problem:

$$\min_{\mathbf{T}, \mathbf{P}} \left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\text{F}}^2 \quad \text{such that} \quad \mathbf{P} \geq 0 \text{ and } \mathbf{Tv} = \mathbf{0}_I \tag{19}$$

IV. PROPOSED SEMI NONNEGATIVE MATRIX FACTORIZATION ALGORITHM

A. Algorithm overview

The minimization problem (19) can be solved by adapting algorithms that are designed for SNMF, which usually solve (19) with the constraint $\mathbf{P} \geq 0$, but without the constraint on $\mathbf{T}$. Most SNMF methods are iterative, and they alternate between the minimization of $\left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\text{F}}^2$ for $\mathbf{P}$ and for $\mathbf{T}$ [35]–[37]. Herein, we also propose an alternating approach where each of the two minimization steps is solved with its respective constraint. An overview of our two-step iterative algorithm is given in Algorithm 1.

Algorithm 1 Solving the pattern generalization problem (16)

Initialization: $\mathbf{P} = \text{rand}(K, P)$
while $\left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\text{F}}^2 > \varepsilon$ **do**
   Step 1: Minimize $\left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\text{F}}^2$ w.r.t. $\mathbf{T}$ such that $\mathbf{Tv} = \mathbf{0}_I$
   Step 2: Minimize $\left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\text{F}}^2$ w.r.t. $\mathbf{P}$ such that $\mathbf{P} \geq 0$
end while

For each of the two steps in Algorithm 1, different resolution methods are available, which include alternating least-squares algorithms [38], multiplicative update algorithms [39], and gradient descent algorithm [40], [41]. We propose a new solution for the equality constrained problem of step 1 in Section IV-B, and describe a solution for the problem of step
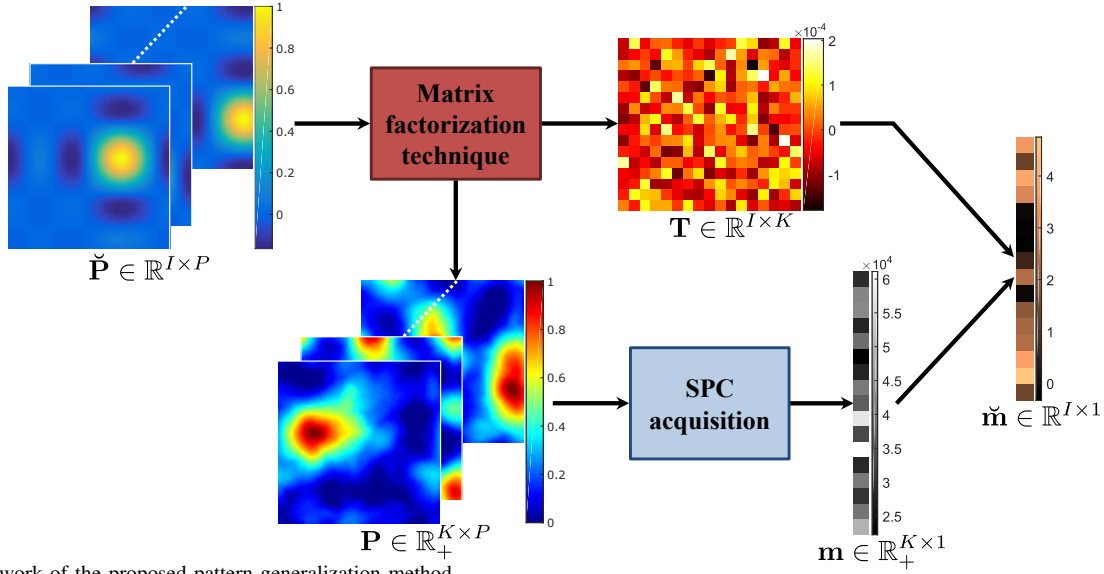
Fig. 1. Framework of the proposed pattern generalization method.

2 in Section IV-C. The solution for step 2 is taken from the literature and retained among the existing methods due to its low factorization error [33].

B. Solution for $\mathbf{T}$

Assuming $\mathbf{P}$ is fixed, the following subproblem must be solved:

$$\min_{\mathbf{T}} \left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\mathrm{F}}^2 \quad \text{such that} \quad \mathbf{Tv} = \mathbf{0}_I \tag{20}$$

In standard SNMF problems where no constraints act on $\mathbf{T}$, the subproblem (20) has a straightforward solution $\mathbf{T} = \check{\mathbf{P}}\mathbf{P}^\top \left(\mathbf{PP}^\top\right)^{-1}$ when $\mathbf{PP}^\top$ is invertible. In our case, the constraint (14) makes it an equality constrained optimization problem. Nevertheless, we show in Appendix A that this can still be solved analytically using the duality property. We have

$$\mathbf{T} = \check{\mathbf{P}}\mathbf{P}^\top \left(\mathbf{PP}^\top\right)^{-1} \left(\mathbf{I} - \frac{1}{\mathbf{v}^\top \left(\mathbf{PP}^\top\right)^{-1} \mathbf{v}} \mathbf{V} \left(\mathbf{PP}^\top\right)^{-1} \right) \tag{21}$$

where $\mathbf{I}$ is the identity matrix of size $K \times K$, and $\mathbf{V} = \mathbf{vv}^\top$ is the matrix of size $K \times K$, with all entries equal to one. In all of our experiments, $\mathbf{P}$ is observed to be full rank, and therefore $\mathbf{PP}^\top$ is invertible. If $\mathbf{PP}^\top$ is not invertible, it is possible to use the Moore-Penrose pseudoinverse instead, as suggested in [33], [37].

C. Solution for $\mathbf{P}$

Assuming $\mathbf{T}$ is fixed, the following subproblem must be solved:

$$\min_{\mathbf{P}} \left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\mathrm{F}}^2 \quad \text{such that} \quad \mathbf{P} \geq 0 \tag{22}$$

The smaller K, the faster the acquisition. There are many existing algorithms for solving this problem with $K < I$, but they lead to large factorization errors ($\sim 1 - 10$). This is unacceptable in our problem, as the approximation error propagates from the acquired measurements to the desired measurements from which image reconstruction is performed.

As this leads to low factorization error (i.e., error in the range $[10^{-10} - 10^{-3}]$) for $K = I + 1$, we selected the technique proposed by Gillis in [33]. This is a block coordinate descent that has a closed-form solution that can be computed rapidly [33]. To derive it, the cost function $\mathcal{F}(\mathbf{P}) = \left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\mathrm{F}}^2$ is rewritten as

$$\mathcal{F}(\mathbf{P}) = \mathcal{F}(\mathbf{P}_{-k}, \mathbf{p}_k) = \left\| \check{\mathbf{P}} - \mathbf{T}_{|k}\mathbf{P}_{-k} - \mathbf{t}_k \mathbf{p}_k^\top \right\|_{\mathrm{F}}^2 \tag{23}$$

where $\mathbf{T}_{|k}$ (resp. $\mathbf{P}_{-k}$) is the matrix $\mathbf{T}$ (resp. $\mathbf{P}$) deprived of its column (resp. row) k, and $\mathbf{t}_k \in \mathbb{R}^{I \times 1}$ (resp. $\mathbf{p}_k^\top \in \mathbb{R}^{1 \times P}$) is the k-th column (resp. row) of $\mathbf{T}$ (resp. $\mathbf{P}$). The problem (22) is solved by minimizing $\mathcal{F}(\mathbf{P})$ with respect to $\mathbf{p}_k$, and iterating over the K rows of $\mathbf{P}$ [33]. Hence, at each iteration, the following problem is solved

$$\min_{\mathbf{p}_k} \mathcal{F}(\mathbf{P}_{-k}, \mathbf{p}_k) \quad \text{such that} \quad \mathbf{p}_k \geq 0, \tag{24}$$

Interestingly, this allows an analytical solution:

$$\mathbf{p}_k = \max \left(\mathbf{0}_P, \frac{\left(\check{\mathbf{P}} - \mathbf{T}_{|k}\mathbf{P}_{-k}\right)^\top \mathbf{t}_k}{\left\| \mathbf{t}_k \right\|_2^2} \right) \tag{25}$$

where the max $(.)$ function is applied entrywise, and $\mathbf{0}_P$ is the null vector of size P. For consistency, a demonstration of (23) is provided in Appendix B. In the rest of this paper, we choose to set the number of actual patterns for our SNMF algorithm to $K_{\mathrm{snmf}} = I + 1$ as done in [33], which guarantees that the factorization error is low.

D. Proposed algorithm

Using the closed-formula (21) and (25) allows Algorithm 1 to be completely written as in Algorithm 2, to solve for $\check{\mathbf{P}} \approx \mathbf{TP}$ with only $K = I + 1$.

V. NUMERICAL EXPERIMENTS

Our numerical simulations are based on the computation of the forward model (2). The Jaszczak target is chosen as the

---

Algorithm 2 Solve $\check{\mathbf{P}} \approx \mathbf{TP}$ as in (19)

Initialization: Set $K = I + 1$ and $\mathbf{P} = \text{rand}(K, P)$

while $\left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_{\mathrm{F}}^{2} > \varepsilon$ **do**

  1: Update $\mathbf{T}$ using (21)

  2: Update $\mathbf{P}$

    for $k = 1 : K$ **do**

      2.1: $\mathbf{p}_k \leftarrow \max \left(\mathbf{0}_P, \frac{(\check{\mathbf{P}} - \mathbf{T}_{|k}\mathbf{P}_{-k})^{\top} \mathbf{t}_k}{\|\mathbf{t}_k\|_2^2} \right)$

      2.2: Update the k-th row of $\mathbf{P}$ with $\mathbf{p}_k^{\top}$

    end for

end while

Fig. 2. Charge-coupled device image of the Jaszczak target used for the numerical experiments.

imaged object $\mathbf{f}$, which is classically considered to determine the resolution of an imaging device. The desired SLM patterns $\check{\mathbf{P}}$ are chosen adaptively on the wavelet basis. The strategy for choosing the patterns, as well as that for recovering the image of the scene from the acquired measurements $\check{\mathbf{m}}$, is described in [27]. The image reconstruction algorithm implementation is based on the open-source SPIRIT toolbox [34]. In particular, 64×64 ($P = 4096$) SPC images are obtained from $I = 1024$ wavelet patterns. Note that the wavelet patterns $\check{\mathbf{P}}$ typically have both positive and negative entries. In practice, SLM patterns are coded on b bits, and typically $b = 8$ bits. Therefore, the maximum of each of the obtained positive patterns of $\mathbf{P}$ is set to $2^b - 1$ using a rescaling factor. The matrix $\mathbf{T}$ is adequately compensated by dividing the elements of each column with the same corresponding rescaling factor, so that the equality $\check{\mathbf{P}} = \mathbf{TP}$ still holds. This operation is done during Algorithm 2, at each iteration after the update of $\mathbf{P}$. Once the algorithm is completed, $\mathbf{P}$ is rounded to the nearest integer so as to get integer values in the range $[0, 2^b - 1]$, and the patterns are therefore implementable on the SLM. Details concerning the influence of pattern quantization with respect to the image restoration quality are given in [27]. The measurements $\mathbf{m}$ are the numbers of photons collected at the single detector, and these are necessarily corrupted by Poisson noise[1]. Mathematically, the following noise model is used:

$$\mathbf{m} = \mathscr{P}\left((N_0 \mathbf{Pf} + \alpha \mathbf{v}) \, \Delta t \right) \tag{26}$$

where $\mathscr{P}(.)$ is the Poisson distribution applied to each entry of the vector. Changing N_0 and/or Δt in (26) allows several levels of noise to be simulated. The larger N_0 and Δt, the larger the number of photons collected, and hence, the better the signal-to-noise ratio of the measurements.

The proposed SNMF method is compared to the pattern splitting and pattern shifting methods presented in Section II-B. The same matrix $\check{\mathbf{P}}$ was considered both for the factorizations of Section III-B and the SNMF Algorithm 2. The stopping criterion for the SNMF method is set to $\varepsilon = 10^{-6}$. For pattern shifting, we choose $M = 10$ in (18) to reduce the influence of noise and to better estimate the background value.

Different sets of acquisitions are considered for the numerical experiments, which correspond to different values for the parameters of (2). In particular, varying light power N_0 is considered. A low N_0 indicates a low-light scenario, while a

---

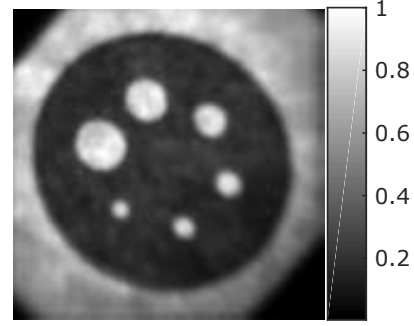[1] There are other sources of noise, but the predominant one is Poisson noise.

large N_0 simulates a very bright object. Varying dark current α is also considered, as each specific optical set-up has its own value that depends on the single-point detector, and the illumination conditions used, and also on other factors.

The integration time Δt is set to 1 in all of our experiments. As the total acquisition time is $\Delta T = K \Delta t$, where K is the number of measurements, different pattern generalization strategies lead to different acquisition times. In particular, with $K_{\text{split}} = 2I$ and $K_{\text{snmf}} = I + 1$, the total acquisition time for the splitting approach is almost twice the total acquisition time for our SNMF approach; i.e., $\Delta T_{\text{split}} \approx 2\Delta T_{\text{snmf}}$, for large values of I.

The charge-coupled device (CCD) image of the Jaszczak target shown in Fig. 2 serves as the reference image. Image acquisition is simulated for N_0 ranging from 100 ph/s to 3600 ph/s, and α ranging from 500 ph/s to 80,000 ph/s. Different values for N_0 account for different laser power or object brightness. The dark current α is related to the number of undesirable photons that reach the detector. This depends on both the imaging environment and the selectivity of the spatial light modulator.

VI. RESULTS

a) Convergence of the proposed algorithm: Figure 3 illustrates the typical positive patterns that are obtained using the proposed SNMF method and using the *ad-hoc* methods. The Le Gall wavelet (CDF 5/3 biorthogonal) patterns of size $P = N \times N = 64 \times 64 = 4096$ are considered. The top row of Fig. 3 shows the pattern generalization $\check{\mathbf{P}} = \mathbf{TP}$ that is obtained using our SNMF approach. In this example, the matrix $\check{\mathbf{P}}$ contains a sequence of $I = 4$ desired patterns, while the matrix $\mathbf{P}$ contains $K = I + 1 = 5$ (positive) patterns, which are also shown in the middle row of Fig. 3. The positive patterns obtained using the pattern-splitting and the pattern-shifting methods for one of the desired patterns of $\check{\mathbf{P}}$ are given in the bottom row of Fig. 3. As can be seen, the proposed SNMF method generates SLM patterns where their main structures and shapes are those of the desired patterns. The desired pattern Fig. 3-(i) indeed has a star shape in its center, which is also visible in the SLM patterns created in Fig. 3-(d)–(h).

The decrease in the pattern generalization error $\left\| \check{\mathbf{P}} - \mathbf{PT} \right\|_{\mathrm{F}}^{2}$ during the SNMF iterations is presented in Fig. 4. The SNMF method is also evaluated for several matrices $\check{\mathbf{P}}$ that contain
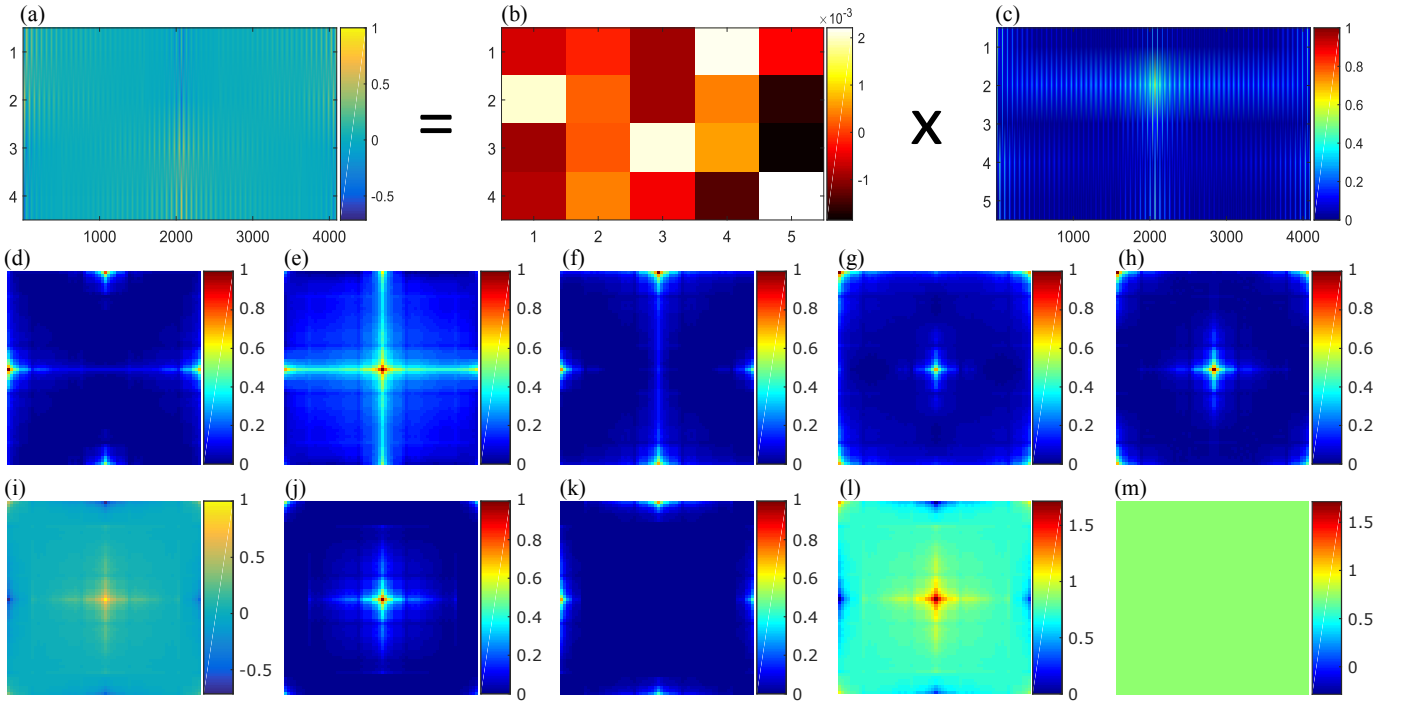
Fig. 3.   Examples of the patterns created with the proposed matrix factorization algorithm using Le Gall wavelet patterns. (a) Sequence matrix $\check{\mathbf{P}}$ of $I = 4$ desired SLM patterns of size $P = 64 \times 64 = 4096$. (b) Transformation matrix $\mathbf{T}$. (c) Sequence matrix $\mathbf{P}$ of $K = I + 1 = 5$ positive SLM patterns. (d-h) Five SNMF patterns of $\mathbf{P}$ (rows of image (c)). (i) Example of a desired pattern $\check{\mathbf{p}}$ (4th row of image (a)). (j) Positive part $\mathbf{p}_+ = \max\left(\mathbf{0}_P, \check{\mathbf{p}}\right)$. (k) Negative part $\mathbf{p}_- = |\min\left(\mathbf{0}_P, \check{\mathbf{p}}\right)|$. (l) Shifted pattern $\mathbf{p}_{\text{shift}} = \check{\mathbf{p}} + \mathbf{p}_{\text{back}}$. (m) Background pattern $\mathbf{p}_{\text{back}}$.
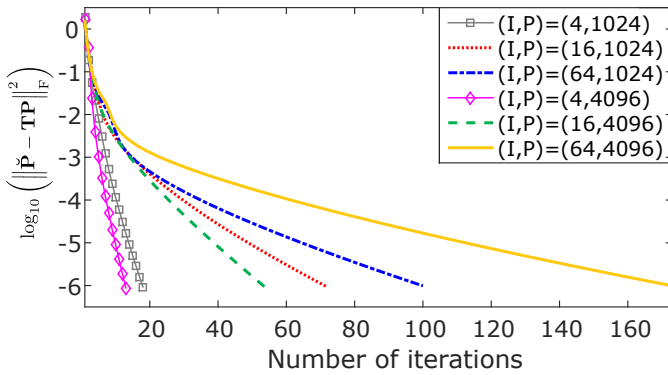

Fig. 4.   Logarithm (base 10) of the error $\left\|\check{\mathbf{P}} - \mathbf{TP}\right\|_{\text{F}}^2$ during the iterations of Algorithm 2 for different values of desired patterns I and pattern sizes P. The convergence criterion ε was set to 10^{-6}.

Value of (I,P)	Number of iterations	Computation time (s)
$(4, 1024)$	18	0.01
$(16, 1024)$	72	0.10
$(64, 1024)$	100	0.68
$(4, 4096)$	13	0.02
$(16, 4096)$	54	0.25
$(64, 4096)$	173	3.58

TABLE I
NUMBER OF ITERATIONS AND COMPUTATION TIME FOR THE PROPOSED
SEMI NONNEGATIVE MATRIX FACTORIZATION ALGORITHM 2 TO
CONVERGE FOR SEVERAL VALUES OF (I,P). THE CONVERGENCE
CRITERION ε WAS SET TO 10^{-6}.

different numbers of patterns I and pattern sizes P. The resulting computation times and number of iterations are reported in Table I.

b) Influence of the conditions for acquisition: The SPC-recovered images of size 64×64 for four different paired (N_0, α) are shown in Fig. 5. The peak signal-to-noise ratio (PSNR) of the restored images is shown as a function of N_0 for two values of α (i.e., 20000 ph/s, 80,000 ph/s) in Fig. 6-(a), and as a function of α for two values of N_0 (i.e., 600 ph/s, 1600 ph/s) in Fig. 6-(b). The larger the PSNR, the better the image quality. The PSNR of the restored images for all (N_0, α) pairs and pattern generalization methods are given in Table II.

All of these three pattern generalization methods perform better for increasing N_0 (see Fig. 6-(a)) and decreasing α (see Fig. 6-(b)). In all cases, pattern shifting provides the lowest PSNR. For low values of α, pattern splitting gives the best image quality. However, the pattern splitting image quality degrades dramatically in low-light scenarios for increasing α (see Fig. 6-(b)). Overall, the proposed SNMF algorithm performs better in most of the scenarios assessed.

c) Influence of the constraint on $\mathbf{T}$: To motivate the introduction of the transformation matrix constraint (14), the SNMF pattern generalization is performed for all of the previously described scenarios without the constraint $\mathbf{Tv} = \mathbf{0}_I$. Table II gives the PSNR of the images recovered with no constraint on $\mathbf{T}$. If α is large enough, the image quality dramatically improves when the equality constraint is added. For the case of Table II when α is small (i.e., 500, 5,000 ph/s), the results are close or even better without the constraint. This is because
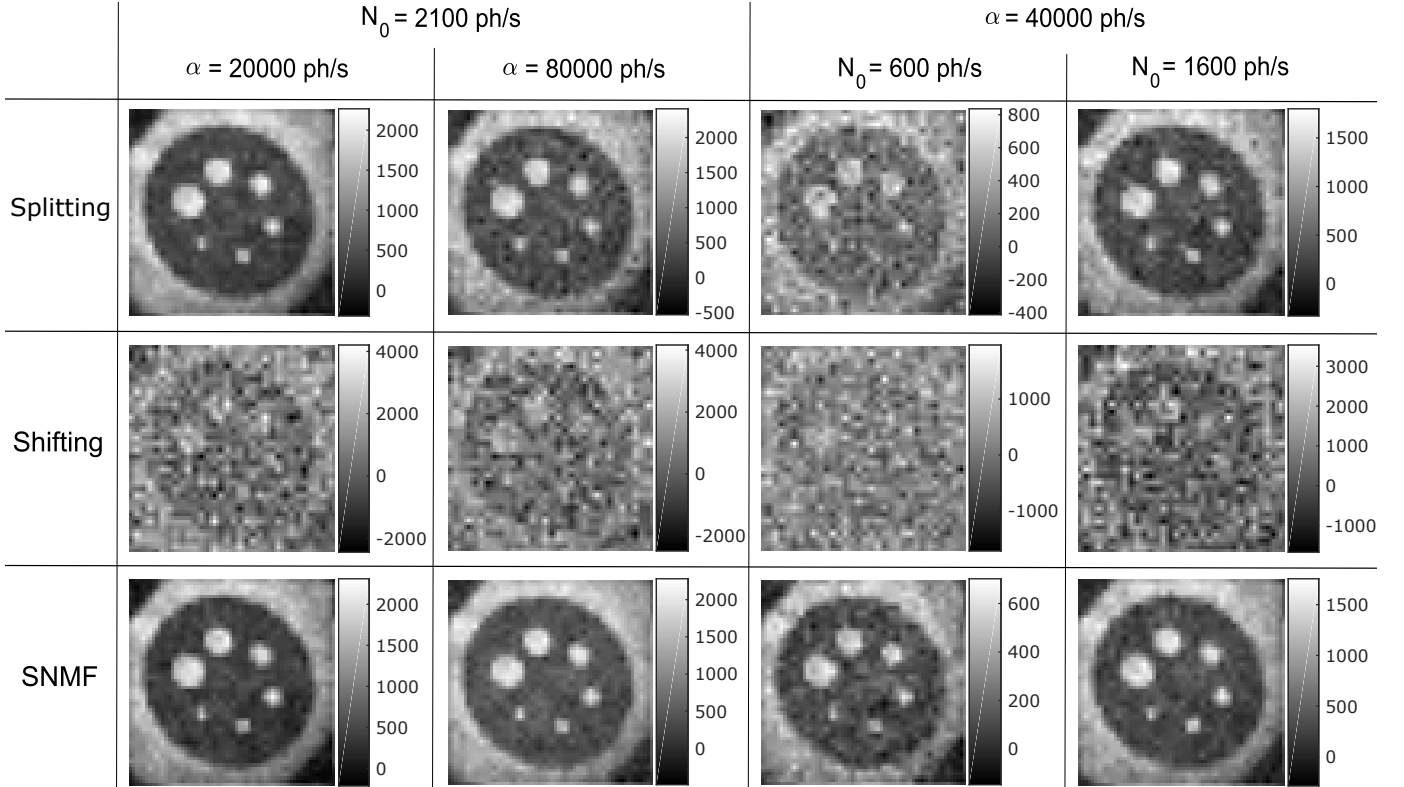
Fig. 5. Single-pixel camera restored images using adaptive basis scan by wavelet prediction for the three matrix factorization techniques for different values of N_0 and α on the test image of Fig. 6. The corresponding peak signal-to-noise ratio compared to the ground-truth image (shown in Fig. 6) can be read from Fig. 6 and Table II.

even if $\mathbf{T}$ is obtained as $\mathbf{T} = \check{\mathbf{P}}\mathbf{P}^\top \left(\mathbf{P}\mathbf{P}^\top\right)^{-1}$ (no constraint), the sum of the rows of $\mathbf{T}$ is around 10^{-4}, which is sufficient to cancel out the small dark-current values. With the constraint, this sum is lowered to about 10^{-18}, which cancels out any value of dark current.

VII. DISCUSSION

In this study, we have introduced the so-called pattern generalization problem (16) to allow for experimental limitations. This problem has analytical solutions for any kind of pattern (e.g., binary random pattern, wavelet, Fourier, discrete cosine transform) for the cases $K = 2I$ and $K = I + M$. These solutions are provided by the pattern splitting method (for $K = 2I$) and by the pattern shifting method (for $K = I + M$). For the case $K = 2I$, the solution of the pattern generalization problem is not unique, as both the pattern splitting and pattern shifting methods can be used (setting $M = I$). Our numerical SNMF approach provides a solution for the case $K = I + 1$. The solution obtained is seen to be different from that obtained using the analytical pattern shifting method, which suggests that the solution is not unique in this case too. From a general point of view, determining the existence and uniqueness of the solution of the pattern generalization problem remains an open question.

The main advantage of our SNMF approach is that it guarantees low factorization error for a low number of positive patterns $K_{\text{snmf}} = I + 1$. While a low number of patterns is necessary to limit the number of measurements, and hence the acquisition time, low factorization error is crucial to limit any model deviation. Most SNMF methods propose the choice of $K << I$, but the factorization error is often important. Here, only $I + 1$ measurements with positive patterns are required to obtain the I desired measurements, which gives an effective measurement ratio $\frac{K}{I}$ of almost 1 for large values of I.

The conventional pattern splitting approach requires $K_{\text{split}} = 2I$ actual measurements. Compared to this, the proposed SNMF method requires only $K_{\text{snmf}} = I + 1$ patterns, which is a reduction by almost a factor 2. As for the SNMF method, the pattern shifting method allows to choose $K_{\text{shift}} = I + 1$ (setting $M = 1$). However, the pattern shifting method led to degraded image quality in all of the experiments conducted here, and even for larger M (e.g., $M = 10$).

We chose to compare the three approaches for the same capture time Δt. Alternatively, the *total* capture time $K\Delta t$ can be kept constant by reducing either the capture time Δt or the number of patterns K. Compared to the SNMF approach, the pattern splitting approach will either suffer from noisier measurements (having set $\Delta t_{\text{split}} = \Delta_{\text{snmf}}/2$) or from higher compression rate (having set $K_{\text{split}} = K_{\text{snmf}}/2$), which will result in a degraded image quality. The patterns shifting method will be affected the same way.

The computational complexity of Algorithm 2 is $\mathcal{O}(IPK^2) \approx \mathcal{O}(PK^3)$ flops per iteration of the while loop. This is dominated by the update of $\mathbf{p}_k$ in the *for* loop in step 2. The number of iterations and the computation time of the SNMF algorithm depends on the size of the problem. As the SNMF algorithm
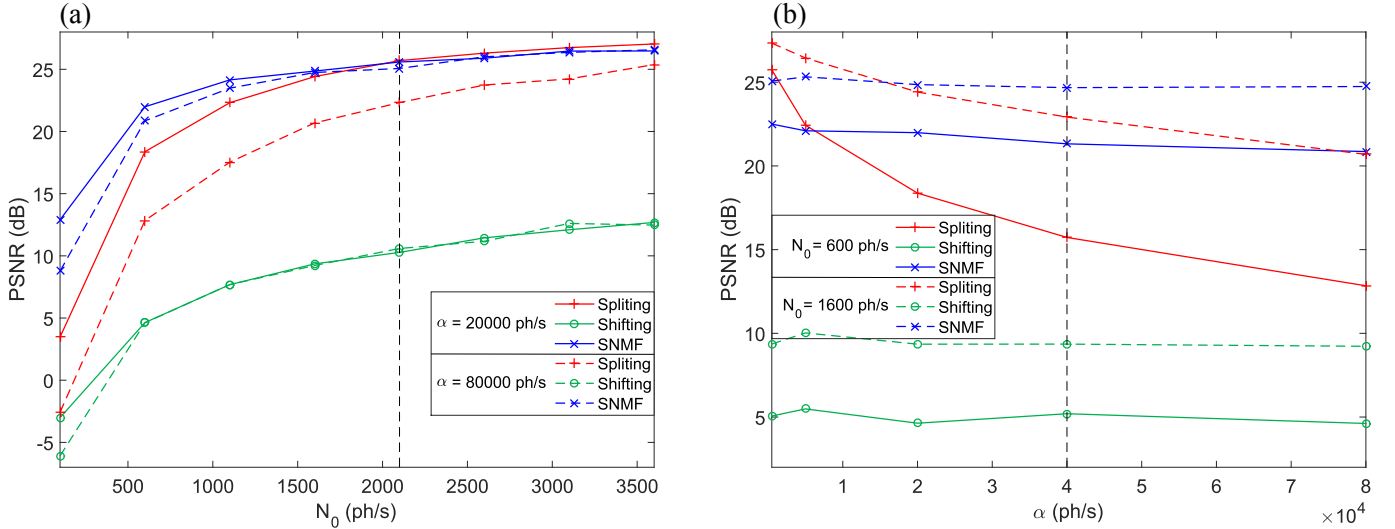
Fig. 6. Jaszczak target and peak signal-to-noise ratio curves of single-pixel camera restored images for (a) (resp. (b)) two fixed values of N_0 (resp. α) and increasing values of N_0 (resp. α) for $\Delta t = 1$ s. The dotted black bars correspond to the results shown in Fig. 5.

Value of α	Technique	PSNR (dB)							
		$N_0 = 100$	$N_0 = 600$	$N_0 = 1100$	$N_0 = 1600$	$N_0 = 2100$	$N_0 = 2600$	$N_0 = 3100$	$N_0 = 3600$
500	Pattern splitting	**17.56**	**25.74**	**26.95**	**27.33**	**27.63**	**27.77**	**27.83**	**27.94**
	Pattern shifting	-2.41	5.06	7.85	9.36	10.49	11.17	12.94	13.22
	SNMF with $\mathbf{Tv} \neq \mathbf{0}_I$	16.00	22.63	24.70	25.32	25.90	26.29	26.42	26.81
	SNMF with $\mathbf{Tv} = \mathbf{0}_I$	15.22	22.50	24.16	25.07	25.68	26.10	26.63	26.68
5000	Pattern splitting	9.34	**22.43**	**25.34**	**26.45**	**26.98**	**27.40**	**27.55**	**27.73**
	Pattern shifting	-2.63	5.49	7.92	10.03	10.18	11.33	12.34	13.36
	SNMF with $\mathbf{Tv} \neq \mathbf{0}_I$	6.91	19.73	22.93	24.53	25.17	25.87	26.32	26.76
	SNMF with $\mathbf{Tv} = \mathbf{0}_I$	**14.93**	22.10	24.29	25.32	25.52	26.14	26.40	26.73
20000	Pattern splitting	3.47	18.37	22.31	22.41	**25.70**	**26.30**	**26.75**	**27.04**
	Pattern shifting	-3.01	4.63	7.68	9.35	10.28	11.45	12.10	12.69
	SNMF with $\mathbf{Tv} \neq \mathbf{0}_I$	-4.58	10.72	15.75	18.74	20.39	21.98	22.83	23.68
	SNMF with $\mathbf{Tv} = \mathbf{0}_I$	**12.90**	**21.99**	**24.15**	**24.86**	25.59	25.89	26.47	26.47
40000	Pattern splitting	0.91	15.73	19.97	22.92	24.58	25.15	26.03	26.44
	Pattern shifting	-4.91	5.20	7.72	9.36	10.42	11.50	12.30	13.01
	SNMF with $\mathbf{Tv} \neq \mathbf{0}_I$	-10.51	4.88	10.17	13.26	15.46	17.27	18.68	19.68
	SNMF with $\mathbf{Tv} = \mathbf{0}_I$	**10.61**	**21.32**	**23.68**	**24.68**	**25.82**	**25.81**	**26.43**	**26.48**
80000	Pattern splitting	-2.57	12.84	17.53	20.69	22.32	23.74	24.23	25.38
	Pattern shifting	-6.08	4.62	7.69	9.23	10.59	11.17	12.61	12.47
	SNMF with $\mathbf{Tv} \neq \mathbf{0}_I$	-16.62	-1.07	4.18	7.42	9.79	11.54	13.17	14.36
	SNMF with $\mathbf{Tv} = \mathbf{0}_I$	**8.78**	**20.86**	**23.48**	**24.74**	**25.07**	**26.00**	**26.36**	**26.57**

TABLE II

PEAK SIGNAL-TO-NOISE RATIOS OF THE SINGLE-PIXEL CAMERA RESTORED IMAGES FOR DIFFERENT VALUES OF N_0 AND α FOR $\Delta t = 1$. SOME CASES CORRESPOND TO THE IMAGES SHOWN IN FIG. 5 AND THE CURVES OF FIG. 6.

updates the rows of $\mathbf{P}$, the increasing of $I \approx K$ (the number of rows of $\mathbf{P}$) has a bigger impact than the increasing of P (the number of columns of $\mathbf{P}$). This can be seen by comparing the computation times of Table I for the cases of $(I,P) = (16,4096)$ and $(I,P) = (64,1024)$, for which the matrix $\check{\mathbf{P}}$ has the same number of entries ($I \times P = 65536$). The SNMF algorithm converges more rapidly for the smaller value of I.

For high resolution images (i.e., large P and I), the SNMF computation can take up to several minutes. For nonadaptive SPC acquisition (e.g., based on compressed sensing), SNMF can be performed prior to the acquisition. However, long computation times can be a limitation for adaptive acquisition strategies where the set of patterns $\check{\mathbf{P}}$ is determined during acquisition from the knowledge of the previous measurements. Therefore, it is necessary to wait for the SNMF algorithm

outputs $\mathbf{P}$ and $\mathbf{T}$ before proceeding with the acquisition. In this study, a computation time of about 4 minutes was required for $I = 1024$ and $P = 4096$, when running Matlab code on a single-core 2.10 GHz CPU. A $C++$ implementation running on a GPU would significantly lower these computation times and make them compatible with adaptive acquisition strategies.

The better results of the SNMF method compared to most cases of pattern splitting can be understood by looking at the patterns in Fig. 3. The patterns (j) and (k) were obtained from pattern splitting, and they have many zeros. When α is large with respect to N_0, the useful part of the pattern (i.e., nonzeros) has only a small contribution to the SPC measurement. The SNMF patterns have fewer zeros, and therefore this effect is reduced (see Fig. 3, patterns (d)–(h)). This effect is also reduced for pattern shifting but this method leads to the lowest image quality; only the shape of the target can be

recovered (see Fig. 5). An explanation here is that this method is very sensitive to noise, as the useful information of the shifted pattern is overwhelmed by the useless information from the background (see Fig. 3-(m)). Looking at the shifted pattern (see Fig. 3-(l)), it can be seen that all of the pixels contribute to the measurement, which leads to the collection of many undesired photons in the presence of noise. Overall, the SNMF technique works well when $N_0 \ll \alpha$; i.e., when the dark current cannot be neglected, which is always the case in real experiments.

The pattern splitting and SNMF methods give similar results, with a slight improvement for pattern splitting for small values of α. A possible explanation here concerns the construction of the measurement vector. For the proposed SNMF technique, the measurements $\check{\mathbf{m}}$ that are used for image restoration are obtained from the measurement transformation $\mathbf{Tm}$ (see (15)). In the presence of noise, the variance of $\check{\mathbf{m}}$ depends on the sum of the variances of $\mathbf{m}$. While many measurements are combined when the SNMF method is used, only two measurements are combined when pattern splitting is used. Hence pattern splitting is less affected by noise than SNMF. However, compared to pattern splitting, the proposed SNMF method an excellent compromise that gives an image quality that is similar to that obtained using pattern splitting, while halving the number of measurements.

VIII. CONCLUSION

In this paper, we introduce a new problem in single-pixel imaging that we refer to as pattern generalization. This consists of determining a set of positive patterns that can be actually loaded onto a SLM. We also show that a second issue that has to be addressed in pattern generalization is dark-current removal. Our basis idea is to see the set of desired patterns (with negative entries) as a linear transform of positive patterns.

We report a SNMF technique that is shown to solve the problem. To our knowledge, this is the first time that a SNMF algorithm is used in the context of single-pixel imaging. This provides an elegant way to dispose of both the positivity constraint induced by the use of a SLM, and the dark-current rejection. At the same time, it allows the number of measurements to be reduced compared to the conventional pattern splitting method.

In the future, we will investigate algorithms for which the number of measurements is further reduced; i.e., $K \leq I$. The main challenge is to limit the factorization error that results in model deviations.

APPENDIX

For the matrix and vector derivation presented below, we refer the reader to the Matrix Cookbook [42].

A. Dual problem to solve for *T*

To solve the dual problem, we first write the Lagrange function $\mathcal{L}$:

$$\mathcal{L}(\mathbf{T}, \boldsymbol{\lambda}) = \left\| \check{\mathbf{P}} - \mathbf{TP} \right\|_F^2 + \boldsymbol{\lambda}^\top \mathbf{Tv}$$

$$\mathcal{L}(\mathbf{T}, \boldsymbol{\lambda}) = \mathrm{tr}\left(\left(\check{\mathbf{P}} - \mathbf{TP}\right)\left(\check{\mathbf{P}} - \mathbf{TP}\right)^\top \right) + \boldsymbol{\lambda}^\top \mathbf{Tv}$$

$$= \mathrm{tr}\left(\mathbf{TPP}^\top \mathbf{T}^\top - \mathbf{TP}\check{\mathbf{P}}^\top - \check{\mathbf{P}}\mathbf{P}^\top \mathbf{T}^\top + \check{\mathbf{P}}\check{\mathbf{P}}^\top \right) + \boldsymbol{\lambda}^\top \mathbf{Tv} \quad (27)$$

where $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_I)^\top \in \mathbb{R}^{I \times 1}$ are the Lagrange multipliers. We now write the dual function $\mathcal{D}$:

$$\mathcal{D}(\boldsymbol{\lambda}) = \min_{\mathbf{T}}(\mathbf{T}, \boldsymbol{\lambda}) = \mathcal{L}(\mathbf{T}(\boldsymbol{\lambda}), \boldsymbol{\lambda})$$

with $\mathbf{T}(\boldsymbol{\lambda}) = \arg\min \mathcal{L}(\mathbf{T}, \boldsymbol{\lambda})$ for $\boldsymbol{\lambda}$ fixed. If the dual function $\mathcal{D}$ is differentiable, and if $\boldsymbol{\lambda}^* = \arg\max \mathcal{D}(\boldsymbol{\lambda})$, then $\mathbf{T}(\boldsymbol{\lambda}^*) = \mathbf{T}$ is the solution of the primal problem (20).

First, we search $\mathbf{T}(\boldsymbol{\lambda}) = \arg\min \mathcal{L}(\mathbf{T}, \boldsymbol{\lambda})$ by expressing the derivative of $\mathcal{L}$, using (27):

$$\frac{\partial \mathcal{L}(\mathbf{T}, \boldsymbol{\lambda})}{\partial \mathbf{T}} = 2\mathbf{TPP}^\top - 2\check{\mathbf{P}}\mathbf{P}^\top + \boldsymbol{\lambda}\mathbf{v}^\top \quad (28)$$

$\mathbf{T}(\boldsymbol{\lambda})$ is found when (28) is null. This gives

$$\mathbf{T}(\boldsymbol{\lambda}) = \left(\check{\mathbf{P}}\mathbf{P}^\top - \frac{1}{2}\boldsymbol{\lambda}\mathbf{v}^\top \right) \left(\mathbf{PP}^\top \right)^{-1} \quad (29)$$

We now use (29) in the equality constraint (14), to get the optimal Lagrange multipliers $\boldsymbol{\lambda}^*$:

$$\mathbf{T}(\boldsymbol{\lambda}^*)\mathbf{v} = \mathbf{0} \Rightarrow \left(\check{\mathbf{P}}\mathbf{P}^\top - \frac{1}{2}\boldsymbol{\lambda}^*\mathbf{v}^\top \right) \left(\mathbf{PP}^\top \right)^{-1} \mathbf{v} = \mathbf{0}$$

which leads to

$$\boldsymbol{\lambda}^* = \frac{2}{\mathbf{v}^\top \left(\mathbf{PP}^\top \right)^{-1} \mathbf{v}} \check{\mathbf{P}}\mathbf{P}^\top \left(\mathbf{PP}^\top \right)^{-1} \mathbf{v} \quad (30)$$

Finally, the final expression of $\mathbf{T}$ can be obtained by solving the problem (20), by replacing (30) in (29):

$$\mathbf{T} = \check{\mathbf{P}}\mathbf{P}^\top \left(\mathbf{PP}^\top \right)^{-1} \left(\mathbf{I} - \frac{1}{\mathbf{v}^\top \left(\mathbf{PP}^\top \right)^{-1} \mathbf{v}} \mathbf{V} \left(\mathbf{PP}^\top \right)^{-1} \right) \quad (31)$$

where $\mathbf{I}$ is the identity matrix of size $K \times K$, and $\mathbf{V} = \mathbf{vv}^\top$ is the matrix of size $K \times K$, with all entries equal to one.

B. Block coordinate descent for *P*

This time, $\mathbf{T}$ is considered fixed and the problem to solve is

$$\min_{\mathbf{p}_k} \mathcal{F}(\mathbf{P}_{-k}, \mathbf{p}_k) \quad \text{such that} \quad \mathbf{p}_k \geq 0 \quad (32)$$

with

$$\mathcal{F}(\mathbf{P}_{-k}, \mathbf{p}_k) = \left\| \check{\mathbf{P}} - \mathbf{T}_{|k}\mathbf{P}_{-k} - \mathbf{t}_k \mathbf{p}_k^\top \right\|_F^2 = \left\| \mathbf{S} - \mathbf{t}_k \mathbf{p}_k^\top \right\|_F^2 \quad (33)$$

where $\mathbf{S} = \check{\mathbf{P}} - \mathbf{T}_{|k}\mathbf{P}_{-k}$, $\mathbf{T}_{|k}$ (resp. $\mathbf{P}_{-k}$) is the matrix $\mathbf{T}$ (resp. $\mathbf{P}$) deprived of its column (resp. row) k, and $\mathbf{t}_k \in \mathbb{R}^{I \times 1}$ (resp. $\mathbf{p}_k^\top \in \mathbb{R}^{1 \times P}$) is the k-th column (resp. row) of $\mathbf{T}$ (resp. $\mathbf{P}$). Let us rewrite $\mathcal{F}(\mathbf{P}_{-k}, \mathbf{p}_k)$ as

$$\left\| \mathbf{S} - \mathbf{t}_k \mathbf{p}_k^\top \right\|_F^2 = \mathrm{tr}\left(\left(\mathbf{S} - \mathbf{t}_k \mathbf{p}_k^\top \right) \left(\mathbf{S} - \mathbf{t}_k \mathbf{p}_k^\top \right)^\top \right)$$

for which we now take its derivative with respect to $\mathbf{p}_k$:

$$\frac{\partial \mathcal{F}(\mathbf{P}_{-k}, \mathbf{p}_k)}{\partial \mathbf{p}_k} = 2\mathbf{p}_k \mathbf{t}_k^\top \mathbf{t}_k - 2\mathbf{S}^\top \mathbf{t}_k = 2\left(\mathbf{p}_k \|\mathbf{t}_k\|_2^2 - \mathbf{S}^\top \mathbf{t}_k \right). \quad (34)$$

The minimum of (32) is found when (34) is null and adding the projection operation onto the feasible set $\mathbb{R}_+^{P \times 1}$ leads to the solution

$$\mathbf{p}_k = \max \left(\mathbf{0}_P, \frac{\mathbf{S}^\top \mathbf{t}_k}{\|\mathbf{t}_k\|_2^2} \right) \tag{35}$$

with the maximum function applied to each entry of both vectors.

REFERENCES

[1] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," in *in Proc. of Computational Imaging IV at SPIE Electronic Imaging*, 2006, pp. 43–52.

[2] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-pixel imaging via compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 83–91, March 2008.

[3] J. Ma, "Single-pixel remote sensing," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 2, pp. 199–203, April 2009.

[4] J. H. Shapiro, "Computational ghost imaging," *Phys. Rev. A*, vol. 78, p. 061802, Dec 2008.

[5] S. S. Welsh, M. P. Edgar, R. Bowman, P. Jonathan, B. Sun, and M. J. Padgett, "Fast full-color computational imaging with single-pixel detectors," *Opt. Express*, vol. 21, no. 20, pp. 23 068–23 074, Oct 2013.

[6] Q. Pian, R. Yao, and X. Intes, "Time-resolved hyperspectral single-pixel camera implementation for compressive wide-field fluorescence lifetime imaging," in *Proc. SPIE*, vol. 9701, 2016, pp. 970 115–970 115–6.

[7] V. Studer, J. Bobin, M. Chahid, H. Shams Mousavi, E. Candes, and M. Dahan, "Compressive fluorescence microscopy for biological and hyperspectral imaging," in *Proceedings of the National Academy of Sciences of the USA*, vol. 109 (26), 2012, pp. E1679–E1687.

[8] F. Magalhaes, M. Abolbashari, F. M. Araujo, M. V. Correia, and F. Farahi, "High-resolution hyperspectral single-pixel imaging system based on compressive sensing," *Optical Engineering*, vol. 51, no. 7, pp. 071 406–1–071 406–6, 2012.

[9] H. Dai, G. Gu, W. He, L. Ye, T. Mao, and Q. Chen, "Adaptive compressed photon counting 3d imaging based on wavelet trees and depth map sparse representation," *Opt. Express*, vol. 24, no. 23, pp. 26 080–26 096, Nov 2016.

[10] W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Applied Physics Letters*, vol. 93, no. 12, 2008.

[11] C. M. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, S. Krishna, D. R. Smith, and W. J. Padilla, "Terahertz compressive imaging with metamaterial spatial light modulators," *Nature Photonics*, vol. 8, no. 8, pp. 605–609, 2014.

[12] F. Magalhaes, F. M. Araujo, M. V. Correia, M. Abolbashari, and F. Farahi, "Active illumination single-pixel camera based on compressive sensing," *Appl. Opt.*, vol. 50, no. 4, pp. 405–414, Feb 2011.

[13] E. Tajahuerce, V. Durán, P. Clemente, E. Irles, F. Soldevila, P. Andrés, and J. Lancis, "Image transmission through dynamic scattering media by single-pixel photodetection," *Opt. Express*, vol. 22, no. 14, pp. 16 945–16 955, Jul 2014.

[14] Y. Zhang, M. P. Edgar, B. Sun, N. Radwell, G. M. Gibson, and M. J. Padgett, "3d single-pixel video," *Journal of Optics*, vol. 18, no. 3, p. 035203, 2016.

[15] A. Rodriguez, P. Clemente, E. Tajahuerce, and J. Lancis, "Dual-mode optical microscope based on single-pixel imaging," *Optics and Lasers in Engineering*, vol. 82, pp. 87 – 94, 2016.

[16] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, 2006.

[17] R. A. DeVerse, R. R. Coifman, A. C. Coppi, W. G. Fateley, F. Geshwind, R. M. Hammaker, S. Valenti, F. J. Warner, and G. L. Davis, "Application of spatial light modulators for new modalities in spectrometry and imaging," in *Spectral Imaging: Instrumentation, Applications, and Analysis II*, vol. 4959, Jul. 2003, pp. 12–22.

[18] S. S. Welsh, M. P. Edgar, R. Bowman, B. Sun, and M. J. Padgett, "Near video-rate linear stokes imaging with single-pixel detectors," *Journal of Optics*, vol. 17, no. 2, p. 025705, 2015.

[19] Z. Zhang, X. Ma, and J. Zhong, "Single-pixel imaging by means of fourier spectrum acquisition," *Nature communications*, vol. 6, 2015.

[20] Z. Zhang and J. Zhong, "Three-dimensional single-pixel imaging with far fewer measurements than effective image pixels," *Opt. Lett.*, vol. 41, no. 11, pp. 2497–2500, Jun 2016.

[21] B.-L. Liu, Z.-H. Yang, X. Liu, and L.-A. Wu, "Coloured computational imaging with single-pixel detectors based on a 2d discrete cosine transform," *Journal of Modern Optics*, vol. 64, no. 3, pp. 259–264, 2017.

[22] S. Deutsch, A. Averbush, and S. Dekel, "Adaptive compressed image sensing based on wavelet modeling and direct sampling," in *SAMPTA'09*, Laurent Fesquet and Bruno Torrésani, Ed., Marseille, France, May 2009, p. General session.

[23] A. Averbuch, S. Dekel, and S. Deutsch, "Adaptive compressed image sensing using dictionaries," *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 57–89, 2012.

[24] H. Dai, G. Gu, W. He, F. Liao, J. Zhuang, X. Liu, and Q. Chen, "Adaptive compressed sampling based on extended wavelet trees," *Appl. Opt.*, vol. 53, no. 29, pp. 6619–6628, Oct 2014.

[25] J. Hahn, C. Debes, M. Leigsnering, and A. M. Zoubir, "Compressive sensing and adaptive direct sampling in hyperspectral imaging," *Digital Signal Processing*, vol. 26, pp. 113 – 126, 2014.

[26] Y.-R. Huo, H.-J. He, F. Chen, and H.-M. Tai, "Adaptive single-pixel imaging based on guided coefficients," *J. Opt. Soc. Am. A*, vol. 34, no. 1, pp. 39–51, Jan 2017.

[27] F. Rousset, N. Ducros, A. Farina, G. Valentini, C. D'Andrea, and F. Peyrin, "Adaptive basis scan by wavelet prediction for single-pixel imaging," *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 36–46, March 2017.

[28] W.-K. Yu, M.-F. Li, X.-R. Yao, X.-F. Liu, L.-A. Wu, and G.-J. Zhai, "Adaptive compressive ghost imaging based on wavelet trees and sparse representation," *Opt. Express*, vol. 22, no. 6, pp. 7133–7144, Mar 2014.

[29] N. Radwell, K. J. Mitchell, G. M. Gibson, M. P. Edgar, R. Bowman, and M. J. Padgett, "Single-pixel infrared and visible microscope," *Optica*, vol. 1, no. 5, pp. 285–289, Nov 2014.

[30] F. Soldevila, E. Salvador-Balaguer, P. Clemente, E. Tajahuerce, and J. Lancis, "High-resolution adaptive imaging with a single photodiode," *Scientific Reports*, no. 5, 09 2015.

[31] A. C. Sankaranarayanan, M. A. Herman, P. Turaga, and K. F. Kelly, "Enhanced compressive imaging using model-based acquisition: Smarter sampling by incorporating domain knowledge," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 81–94, Sept 2016.

[32] C. Ding, T. Li, and M. Jordan, "Convex and semi-nonnegative matrix factorizations," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 1, pp. 45–55, Jan 2010.

[33] N. Gillis and A. Kumar, "Exact and heuristic algorithms for semi-nonnegative matrix factorization," *SIAM Journal on Matrix Analysis and Applications*, vol. 36, no. 4, pp. 1404–1424, 2015.

[34] S. S. pixel image reconstruction toolbox. (version 1.0, 2017). [Online]. Available: https://www.creatis.insa-lyon.fr/~ducros/WebPage/spirit.html

[35] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," in *Computational Statistics and Data Analysis*, 2006, pp. 155–173.

[36] Y. X. Wang and Y. J. Zhang, "Nonnegative matrix factorization: A comprehensive review," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 6, pp. 1336–1353, June 2013.

[37] N. Gillis, "The why and how of nonnegative matrix factorization," in *Regularization, Optimization, Kernels, and Support Vector Machines*. Chapman and Hall/CRC, 2014, pp. 257–291.

[38] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, 1994.

[39] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.

[40] ——, "Algorithms for non-negative matrix factorization," in *NIPS*. MIT Press, 2001, pp. 556–562.

[41] C.-J. Lin, "Projected gradient methods for nonnegative matrix factorization," *Neural Comput.*, vol. 19, no. 10, pp. 2756–2779, Oct. 2007.

[42] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, p. 15, 2008.