Blind Multiple-Input Multiple-Output Image Phase Retrieval

Yina Guo, *Member, IEEE*, Xiangning Zhao, Jianyu Li, Anhong Wang, and Wenwu Wang, *Senior Member, IEEE*

***Abstract*—In this paper, we consider the problem of recovering the phase information of the multiple images from the multiple mixed phaseless short-time Fourier transform image measurements, which is called the blind multiple-input multiple-output image phase retrieval (BMIPR) problem. It is an inherently ill-posed problem due to the lack of the phase and mixing information, and the existing phase retrieval algorithms are not explicitly designed for this case. To address the BMIPR phase retrieval problem, an integrated algorithm is presented, which combines a gradient descent algorithm by minimizing a nonconvex loss function with an independent component analysis algorithm and a nonlocal means algorithm. Experimental evaluation has been conducted to show that under appropriate conditions, the proposed algorithms can explicitly recover the images, the phases of the images, and the mixing matrix. In addition, the algorithm is robust to noise.**

***Index Terms*—Blind multiple-input multiple-output image phase retrieval (BMIPR), independent component analysis (ICA), nonconvex optimization, nonlocal means (NM), short-time Fourier transform (STFT).**

I. Introduction

THE PROBLEM of recovering a one-dimensional (1-D) signal from its Fourier transform magnitude, known as phase retrieval, is of paramount importance in various engineering and scientific applications, such as X-ray crystallography [1], [2], optics [3], [4], astronomy [5], [6], blind channel estimation [7], [8], and blind image deblurring [9], [10]. This problem has a long history and has been studied by many researchers [3], [11]–[14].

Manuscript received July 29, 2018; revised January 7, 2019; accepted February 7, 2019. Date of publication March 6, 2019; date of current version October 31, 2019. This work was supported by the National Natural Science Foundation of China under Grant 61301250, in part by the Key Research and Development Project of Shanxi Province under Grant 201803D421035, in part by the Outstanding Young Academic Leaders of Higher Learning Institutions of Shanxi Province under Grant [2015]3, in part by the Collaborative Innovation Center of Internet+3D Printing in Shanxi Province, and in part by the Scientific and Technological Innovation Team of Shanxi Province under Grant 201705D131025. *(Corresponding authors: Yina Guo and Wenwu Wang.)*

Y. Guo, X. Zhao, J. Li, and A. Wang are with the Department of Electronics and Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China (e-mail: zulibest@163.com; zxn@tyust.edu.cn; lijianyu@tyust.edu.cn; wah_ty@tyust.edu.cn).

W. Wang is with the Centre for Vision, Speech and Signal Processing, University of Surrey, Surrey GU2 7XH, U.K. (e-mail: w.wang@surrey.ac.uk).

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

Digital Object Identifier 10.1109/TIE.2019.2901661

The phase retrieval problem originally arises from detectors that can sometimes only record the magnitude square of the Fourier transform of a signal. Due to the lack of Fourier phase information, some forms of additional information is required to identify the underlying signal efficiently. In this respect, the phase retrieval methods can be mainly classified into two categories based either on additional prior information, such as sparsity [3], [11], [12] or additional magnitude-only measurements, including structured illuminations and masks [13], [15], [16], and short-time Fourier transform (STFT) magnitude-square measurements [14], [17], [18]. The key idea of using additional STFT magnitude-square measurements is to introduce redundancy in the magnitude-only measurements by maintaining a substantial overlap between adjacent short-time windows [17].

These phase retrieval methods have focused on recovering a single source from its Fourier transform magnitude. However, in certain cases, the problem of recovering multiple underlying images from multiple mixed Fourier transform magnitudes of images, called blind multiple-input multiple-output image phase retrieval (BMIPR), is ever-present in charge-coupled device (CCD) cameras and photosensitive films [19], [20], such as astronomy [11] or light field images [10], [21]. This problem is ill-posed due to the lack of the phase property and the mixing information. Recently, Guo *et al.* proposed a method for recovering the multiple one-dimensional (1-D) signals from the multiple mixed phaseless STFT measurements [22]. Bendory *et al.* consider the problem of recovering a pair of signals from their blind STFT [14]. Although these methods extend the research to a two-source scenario for 1-D signals, the existing phase retrieval methods cannot provide a solution to the problem of recovering the multiple underlying images from the multiple mixed phaseless STFT image measurements. Therefore, it is necessary to investigate the BMIPR problem.

Extending the former research of Guo *et al.* [22], a closely related problem of recovering the multiple underlying images from the multiple mixed phaseless STFT image measurements is considered. In this paper, our contribution is three-fold, which is as follows.

1) BMIPR Model: A new model of the BMIPR problem is proposed in order to recover multiple underlying images from multiple mixed STFT image magnitude-square measurements, corrupted by noise.
2) BMIPR Algorithms: Due to the absence of Fourier phase information and mixing information, we explore

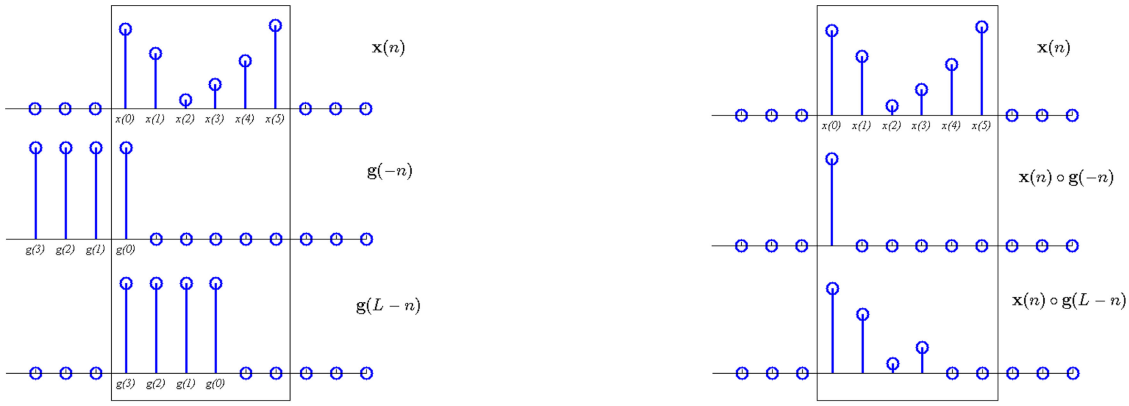


Fig. 1. Two shifting examples of the sliding window overlapped with the signal ($R = \lceil N/L \rceil = 2$) for $N = 6, W = 4$, and $L = 3$. (a) A signal $x(n)$ and two sliding windows $g(-n)$ and $g(L-n)$, shifted by 0 and L time units. (b) A signal $x(n)$, and the examples of $x(n)$ modulated by the sliding windows.

hybrid methods by introducing additional STFT magnitude-square measurements as well as estimating the mixing information. In view of the BMIPR model, an integrated algorithm is proposed, which combines a gradient descent (GD) algorithm by minimizing a nonconvex loss function with an improved complex maximization of non-Gaussianity (CMN) algorithm and a nonlocal means (NM) algorithm. At first, the mixed images can be recovered by the GD algorithm by minimizing a nonconvex loss function. Then, we use a composite algorithm that combined an improved CMN algorithm and an NM algorithm to estimate the mixing information and the underlying images from the mixed images.

3) Initialization of the GD Algorithm: It is shown in [22] that the initialization of the GD algorithm can be obtained by minimizing a nonconvex loss function and equivalently posed as a constrained least squares (LS) solution with a penalty term (ℓ_2). However, this method tends to limit the value range of model parameters and produce biases. To address this issue, we propose to use the principle eigenvector of a designed correlation matrix to initialize the GD algorithm that minimizes an LS solution with a penalty term ($\frac{\ell_1}{\ell_2}$). The new loss function may provide significant benefits in three aspects. First, it is more likely to get a sparser solution than the use of ℓ_p ($p \in (0,1)$) norm. Second, it has a better analytical structure than ℓ_q ($q \in (1,2)$). Third, it prevents over-fitting and improves generalization performance and relaxes the rank restriction of the regression variables.

This paper is organized as follows. Section II formulates a mathematical model and gives the assumptions for the BMRP problem from the multiple mixed STFT image magnitude-square measurements. Section III discusses the uniqueness of the BMRP problem and presents the conditions under which it has a solution by combining a GD algorithm, an independent component analysis (ICA) algorithm, and an NM algorithm. This section also explores the initialization method for the GD algorithm. Section IV shows numerical experimental results. Section V concludes this paper.

II. Mathematical Model and Assumptions

Consider a collaborative assessment task that is induced by blind multi-image phase retrieval from multiple mixed phaseless STFT image measurements in a noisy environment, we present a mathematical model and the assumptions for this task in this section.

A. Model for the BMIPR Problem

The multiple underlying image signals are denoted as $\mathbf{I} = \{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_M\}$, where $\mathbf{I}_i \in \mathbb{C}^{N\times N}$, $i = \{1, 2, \ldots, M\}$. The mixtures of multiple underlying images are defined as $\mathbf{X} = \mathbf{AI}$, with the mixtures $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_M\}, \mathbf{X}_i \in \mathbb{C}^{N\times N}$, $i = \{1, 2, \ldots, M\}$, and the mixing matrix $\mathbf{A} \in \mathbb{R}^{M\times M}$.

Denote $\mathbf{F} = \{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_M\}$ as the STFT matrices of $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_M\}$, where $\mathbf{X}_i = \{\mathbf{x}_{i1}, \mathbf{x}_{i2}, \ldots, \mathbf{x}_{iN}\}$, $\mathbf{x}_{ij} \in \mathbb{C}^N$, $\mathbf{F}_i = \{\mathbf{F}_{i1}, \mathbf{F}_{i2}, \ldots, \mathbf{F}_{iN}\}$, $\mathbf{F}_{ij} \in \mathbb{C}^{R\times N}$, $i = \{1, 2, \ldots, M\}$, and $j = \{1, 2, \ldots, N\}$.

The elements of $\mathbf{F}_i$ can be defined as

$$\begin{cases} F_{i1}(\tau, k) = \sum_{n=0}^{N-1} x_{i1}(n) g(\tau L - n) e^{-2j\pi kn/N} \\ \qquad\qquad \vdots \\ F_{iN}(\tau, k) = \sum_{n=0}^{N-1} x_{iN}(n) g(\tau L - n) e^{-2j\pi kn/N} \end{cases} \tag{1}$$

for $\tau = 0, \ldots, R-1$ and $k = 0, \ldots, N-1$, where W is the window length, L depicts the separation in time between adjacent short-time windows, $R = \lceil N/L \rceil$ denotes the number of short-time windows considered, and $\lceil\ \rceil$ rounds the argument to the smallest integer that is not less than the argument. Fig. 1 shows an example $\mathbf{g}_{\tau L} = \{g(\tau L - n)\}_{n=0}^{N-1}$ and applying it to a signal by shifting the sliding window $\mathbf{g}$ by τL time units. $\mathbf{x}$ and $\mathbf{g}$ are zero padded over the boundaries of (1), where $i = \{1, 2, \ldots, M\}$, and $j = \{1, 2, \ldots, N\}$. The τth row of $\mathbf{F}_{ij}$ corresponds to the N-point discrete Fourier transform (DFT) of $\mathbf{x}_{ij} \circ \mathbf{g}_{\tau L}$.

Denote the STFT magnitude-squared measurements as $|\mathbf{F}|^2 = \{|\mathbf{F}_1|^2, \ldots, |\mathbf{F}_M|^2\}$. We have the following signal

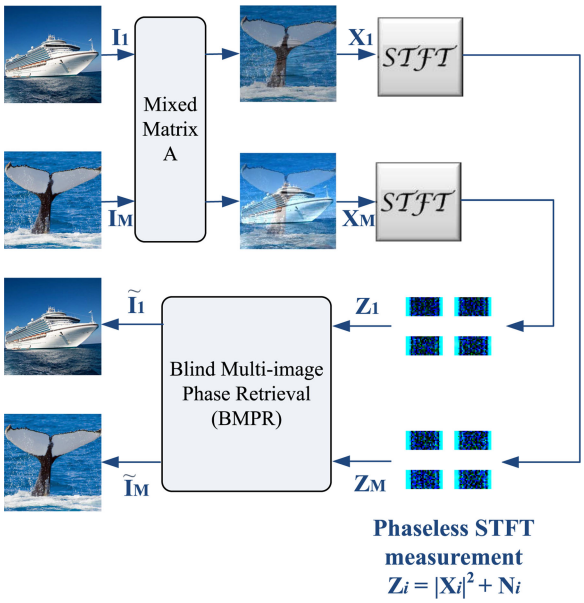


Fig. 2. Proposed BMIPR model.

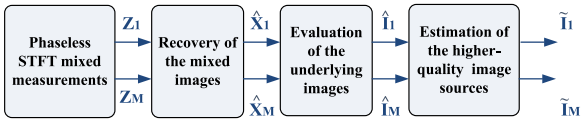


Fig. 3. Schematics diagram of the proposed BMIPR algorithm.

model:

$$\mathbf{Z}_i = |\mathbf{F}_i|^2 + \mathbf{N}_i \tag{2}$$

where $i = 1, \ldots, M$ and $\mathbf{N}_i$ is a random $N \times R \times N$ tensor that represents noise. Thus, $\mathbf{Z} = \{\mathbf{Z}_1, \ldots, \mathbf{Z}_M\}$ and $\mathbf{N} = \{\mathbf{N}_1, \ldots, \mathbf{N}_M\}$.

The aim of multisource phase retrieval is to recover the phases of the underlying sources $\mathbf{I}$ from the phaseless STFT measurements $\mathbf{Z}$ corrupted by noise $\mathbf{N}$. The model of the BMIPR problem is illustrated in Fig. 2.

To address this problem, two assumptions and a two-step algorithm are proposed as discussed in the following.

B. Assumptions for the BMIPR Problem

To address the BMIPR problem, following two assumptions are utilized for constructing the BMIPR model.

1) The mixed images $\mathbf{X}$ are linear mixtures of the multiple underlying image sources $\mathbf{I}$ multiplied by the mixing matrix $\mathbf{A}$.
2) The multiple underlying image sources $\mathbf{I}$ are independent from each other.

III. BMIPR Algorithm

In this section, we find the conditions for solving the BMIPR problem and under which we provide a three-step solution (see Fig. 3). The first step is to recover the mixed image signals $\hat{\mathbf{X}}$ from its mixed phaseless STFT measurements $\mathbf{Z}$ corrupted by noises $\mathbf{N}$. The second step is to estimate multiple image sources $\mathbf{I}$ from the recovered mixed image signals $\hat{\mathbf{X}}$. The third step is to reduce the noises of $\hat{\mathbf{I}}$ and obtain multiple higher quality underlying image sources $\tilde{\mathbf{I}}$.

The fundamental question in BMIPR is whether the underlying image sources $\tilde{\mathbf{I}}$ can be determined uniquely from $\mathbf{Z}$. Based on the research of Guo *et al.* [22], we propose a BMIPR algorithm, which is effective to solve this problem.

A. Recovery of the Phases of the Mixed Image Signals $\hat{\mathbf{X}}$

For the purpose of recovering the phases of the mixed image signals $\mathbf{X}$, we first take DFT of the phaseless STFT measurements $\mathbf{Z}$ to simplify the quadratic system of equations and obtain required correlation data [23]. We take $\mathbf{Z}_{ij}$ as an example, where $\mathbf{Z}_{ij} \in \mathbb{C}^{R\times N}$, $i = \{1, 2, \ldots, M\}$, and $j = \{1, 2, \ldots, N\}$. The DFT of the measurement $Y_{ij}(\tau, l)$ can be described by

$$\begin{aligned} Y_{ij}(\tau, l) &= \frac{1}{N}\sum_{k=0}^{N-1} Z_{ij}(\tau, k) e^{-2j\pi kl/N} \\ &= \sum_{n=0}^{N-1} x_{ij}(n) x_{ij}^*(n+l) g(\tau L - n) g(\tau L - n - l) \end{aligned} \tag{3}$$

where $Y_{ij}(\tau, l)$ is equal to zero for all τ when $W \le l \le (N - W)$ and can be interpreted as a "W bandlimited" function. For fixed τ, $\mathbf{Y}_{ij}$ can be seen as the autocorrelation of $\mathbf{x}_{ij} \circ \mathbf{g}_{\tau L}$, where $\mathbf{g}_{\tau L}$ is defined as in (1). The DFT is normalized by $1/N$. Note that the τth row of $\mathbf{Z}_{ij}$ and the τth row of $\mathbf{Y}_{ij}$ are Fourier pairs. Hence, for a particular τ, if $Z_{ij}(\tau, l)$ for $0 \le l \le N - 1$ is available, then $Y_{ij}(\tau, l)$ for $0 \le l \le N - 1$ can be calculated by taking an inverse Fourier transform [17].

Theorem 3.1: $Z(\tau, k)$ for $0 \le k \le 2W - 2$ is sufficient to calculate $Y(\tau, l)$ for $0 \le l \le N - 1$.

Proof: See Appendix A. ■

Let $\mathbf{D}_{\tau L} \in \mathbb{R}^{N\times N}$ be a diagonal matrix composed of the entries of $\mathbf{g}_{\tau L}$, the problem of recovering $\mathbf{x}_{ij}$ from the measurement $\mathbf{Z}_{ij}$ can therefore be equivalently posed as a nonconvex loss function derived from

$$f(\mathbf{x}_{ij}) = \frac{1}{2}\sum_{\tau=0}^{R-1}\sum_{l=-(W-1)}^{W-1} (\mathbf{x}_{ij}^* \mathbf{H}_{\tau,l} \mathbf{x}_{ij} - Y_{ij}(\tau, l))^2 \tag{4}$$

where $\mathbf{H}_{\tau,l} = \mathrm{P}_{-l}\mathbf{D}_{\tau L}\mathbf{D}_{\tau L - l}$, $\mathbf{x}_{ij}^* \mathbf{H}_{\tau,l} \mathbf{x}_{ij} = \mathrm{tr}(\mathbf{X}_{ij}\mathbf{H}_{\tau,l})$, $\mathrm{P}_{-l} = \mathrm{P}_l^T$, and $(\mathrm{P}_l x)(n) = x(n + l)$.

A GD algorithm is adopted to recover the ijth mixed image signal by minimizing (4). The kth iteration is

$$\mathbf{x}_{ijk} = \mathbf{x}_{ij(k-1)} - \mu \nabla f(\mathbf{x}_{ij(k-1)})$$

$$\begin{aligned} &\nabla f(\mathbf{b}) \\ &= \sum_{\tau=0}^{R-1}\sum_{l=-(W-1)}^{W-1} (\mathbf{b}^T \mathbf{H}_{\tau,l} \mathbf{b} - Y_{i(j-1)}(\tau, l))[(\mathbf{H}_{\tau,l} + \mathbf{H}_{\tau,l}^T)\mathbf{b}] \end{aligned} \tag{5}$$

where $\mathbf{b} = \mathbf{x}_{ij(k-1)}$, μ is the gradient step size, and R is defined as in (1).

According to the above-mentioned procedures, $\mathbf{x}_{ij}$ is recovered from $\mathbf{Z}_{ij}$, then we can get $\mathbf{X}_i$. For each $\mathbf{Z}_{ij}$, the above-mentioned procedures from (3) to (5) are repeated until all the mixed signals $\mathbf{X}$ have been recovered. In order to remove the magnitude effect of the mixing matrix, the recovered mixed image signals are normalized as $\hat{\mathbf{X}} = \{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \ldots, \hat{\mathbf{X}}_M\}$.

Improved loss function for initialization: For $L = 1$, the research of Guo *et al.* [22] show that the geometry of the loss function for the initialization heavily affects the properties of the GD algorithm and the initialization $\mathbf{x}_0$ can be determined by a constrained LS solution with a penalty term (ℓ_2). However, it may limit the value range of the model parameters and produce biases. To address this issue, a penalty term is introduced as follows, for $L = 1$:

$$\min_{\mathbf{x}_{ij} \in \mathbb{C}^N} \sum_{l=-(W-1)}^{W-1} \|\mathbf{y}_{ijl} - \mathbf{G}_l \mathrm{diag}(\mathbf{F}_{ij}, l)\|^2 + \lambda \frac{\|\mathrm{diag}(\mathbf{F}_{ij}, l)\|_1}{\|\mathrm{diag}(\mathbf{F}_{ij}, l)\|_2} \quad \text{subject to } \mathbf{F}_{ij} = \mathbf{x}_{ij}\mathbf{x}_{ij}^* \tag{6}$$

where $\mathbf{y}_{ijl} = \{Y_{ij}(\tau, l)\}_{\tau=0}^{R-1}$, λ is a regularization coefficient, and the (τ, n)th entry of the matrix $\mathbf{G}_l \in \mathbb{R}^{R \times N}$ is given by $g(\tau L - n)g(\tau L - n - l)$. Then, the first column of $\mathbf{G}_l$ can be given by the nonvanishing matrix $\mathbf{g} \circ (\mathbf{P}_{-l}\mathbf{g})$, where $\mathbf{g} = \{g(n)\}_{n=0}^{N-1}$, $\mathrm{P}_{-l} = \mathrm{P}_l^T$, and $(\mathrm{P}_l x)(n) = x(n + l)$. $\mathbf{G}_l$ as a circulant matrix can be factored as $\mathbf{G}_l = \mathbf{F}^* \mathbf{\Sigma}_l \mathbf{F}$, where $\mathbf{F}$ is the DFT matrix and $\mathbf{\Sigma}_l$ is a diagonal matrix. The new loss function may provide significant benefits in three aspects. First, it is more likely to get a sparser solution than ℓ_p ($p \in (0, 1)$). Second, it has better analytical structure than ℓ_q ($q \in (1, 2)$). Third, it prevents over-fitting and improves generalization performance and relaxes the rank restriction of the regression variables [24].

Thus, we construct a matrix $\mathbf{F}_{ij0}$ from (6) as follows:

$$\min_{\mathbf{x}_{ij} \in \mathbb{C}^N} \sum_{l=-(W-1)}^{W-1} \|\mathbf{y}_{ijl} - \mathbf{G}_l \mathrm{diag}(\mathbf{F}_{ij0}, l)\|^2 + \lambda \frac{\|\mathrm{diag}(\mathbf{F}_{ij0}, l)\|_1}{\|\mathrm{diag}(\mathbf{F}_{ij0}, l)\|_2} \quad \text{subject to } \mathbf{F}_{ij0} = \mathbf{x}_{ij0}\mathbf{x}_{ij0}^* \tag{7}$$

where $\mathbf{G}_l$ and $\mathbf{y}_{ijl}$ are represented in (6), and $\mathbf{x}_{ij0}$ is a principle eigenvector of $\mathbf{F}_{ij0}$. Then, the initialization $\mathbf{x}_{ij0}$ of the proposed GD algorithm can be constructed by $\mathbf{F}_{ij0}$.

Theorem 3.2: Denote $\mathbf{z} = \mathrm{diag}(\mathbf{F}_{ij0}, l)$, and $\mathcal{S}(\mathbf{z})$ as the support of $\mathbf{z}$. $\mathbf{z} \in \mathcal{F}$ is called locally sparse if $\nexists \mathbf{y} \in \mathcal{F} \setminus \{\mathbf{z}\}$ such that $\mathcal{S}(\mathbf{y}) \subseteq \mathcal{S}(\mathbf{z})$. Denote by $\mathcal{F} = \{\mathbf{z} \in \mathbb{C}^N : \mathbf{z}$ is locally sparse, $i = \{1, 2, \ldots, M\}, j = \{1, 2, \ldots, N\}\}$ as the set of feasible solutions.

Proof: See Appendix B. ■

In the case that $L > 1$, we need to expand $\mathbf{y}_{ijl} = \{Y_{ij}(\tau, l)\}_{\tau=0}^{R-1}$ to an up-sampled version $\tilde{\mathbf{y}}_{ijl}$ by expansion and interpolation as

$$Y_{ij}(n, l) = \begin{cases} Y_{ij}(\tau, l), & n = \tau L \\ 0, & \text{otherwise} \end{cases}.$$

$$\text{Let } \hat{\mathbf{y}}_{ijl} = \{Y_{ij}(n, l)\}_{n=0}^{N-1} \text{ for fixed } l$$

$$\tilde{\mathbf{y}}_{ijl} = (\mathbf{F}_p^* \mathbf{F}_p)\hat{\mathbf{y}}_{ijl} \tag{8}$$

where $\mathbf{F}_p$ is a partial Fourier matrix consisting of the first R rows of the DFT matrix $\mathbf{F}$ defined as in (9). Then, the initialization $\mathbf{x}_{ij0}$ can be obtained by (8).

B. Evaluation of Multiple Image Sources $\hat{\mathbf{I}}$

All the normalized mixed image signals $\mathbf{X}$ are used as the input for the recovery of the underlying image sources. On the basis of the ICA idea [25]–[27], we extend the CMN algorithm presented in [28] and [29] for blind image separation, which is an effective algorithm for both circular and noncircular sources using complex functions.

We take $\hat{\mathbf{X}}_i$ as an example, where $\hat{\mathbf{X}}_i$ is converted from a matrix to a vector $\hat{\mathbf{x}}_i$ row by row. The whitening mixed image signals $\tilde{\mathbf{x}}_i = \mathbf{U}\hat{\mathbf{x}}_i$ are transformed by a whitening matrix $\mathbf{U}$ [28]. Here, $E\{\tilde{\mathbf{x}}_i \tilde{\mathbf{x}}_i^*\} = \mathbf{I}$.

The use of whitening allows us to search for an orthogonal matrix $\mathbf{W}$ as $E\{\mathbf{W}\tilde{\mathbf{x}}_i(\mathbf{W}\tilde{\mathbf{x}}_i)^*\} = \mathbf{W}E\{\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_i^*\}\mathbf{W}^* = \mathbf{I}$.

Each source s_k is estimated by finding a vector $\mathbf{w}$ such that

$$s_{ik} = \mathbf{w}_k^* \tilde{\mathbf{x}}_i \tag{9}$$

where $\mathbf{w}_k$ is a column of $\mathbf{W}^*$. Constraining the source to $E\{s_{ik}s_{ik}^*\} = 1$, the weights to $\|\mathbf{w}\|^2 = 1$, and $\mathbf{W}$ unitary due to the whitening transform [29].

The optimal weights $\mathbf{w}_{\text{opt}}$ are determined by maximizing the cost under the unit norm constraint where $\tilde{\mathbf{x}}_i$ has been whitened and G is any complex analytic function $\mathbb{C} \mapsto \mathbb{C}$, such as polynomials or transcendental functions, which is given by

$$\mathbf{w}_{\text{opt}} = \arg \max_{\|\mathbf{w}\|^2 = 1} E\{|G(\mathbf{w}^* \tilde{\mathbf{x}}_i)|^2\}. \tag{10}$$

To calculate the optimal weights, a gradient optimization algorithm is used, followed by a normalization step

$$\mathbf{w} \leftarrow \mathbf{w} + \mu\nu \frac{\partial J(s_i)}{\partial \mathbf{w}}$$

$$\mathbf{w} \leftarrow \frac{\mathbf{w}}{\|\mathbf{w}\|} \tag{11}$$

where $J(s_i) = E\{|G(s_i)|^2\}$, $s_i = \mathbf{w}^* \tilde{\mathbf{x}}_i$, μ is the learning rate, and $\nu \in \{-1, 1\}$ is the parameter that determines whether we are maximizing or minimizing the cost function.

Theorem 3.3: Let $J = J(s_i)$. Suppose $\mathbf{w} = \mathbf{w}^R + j\mathbf{w}^I$, and $\mathbf{w}^R$ and $\mathbf{w}^I$ are two real variables. The partial derivative of the cost function J with respect to the conjugate of the weight vector $\mathbf{w}$, referred to as the conjugate gradient, is obtained as

$$\frac{\partial J}{\partial \mathbf{w}} = \frac{\partial J}{\partial \mathbf{w}^R} + j\frac{\partial J}{\partial \mathbf{w}^I} = E\{\tilde{\mathbf{x}}_i G^*(s_i) g(s_i)\}$$

where g is the derivation of G.

Proof: See Appendix C. ■

After each source is estimated, the vectors $\mathbf{w}$ are orthogonalized to prevent multiple solutions from converging to the same maximum since $\mathbf{W}$ is unitary due to the prewhitening step.

The mixing matrix is estimated as $\mathbf{A} = \mathbf{U}^{-1}\mathbf{W}^*$. We recover the underlying signals by $\mathbf{S} = \mathbf{A}^{-1}\tilde{\mathbf{X}}$. Given $\mathbf{S} = \mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M$, each recovered source $\mathbf{s}_i$ needs to be converted from a vector to a matrix $\bar{\mathbf{I}}_i$, where $i = 1, 2, \ldots, M$. The recovered underlying image sources are normalized as $\hat{\mathbf{I}} = \{\hat{\mathbf{I}}_1, \hat{\mathbf{I}}_2, \ldots, \hat{\mathbf{I}}_M\}$.

C. Estimation of Higher Quality Image Sources $\tilde{\mathbf{I}}$

For purpose of removing the interferences of the underlying image sources, we use the Stein's unbiased risk estimate NM algorithm proposed in [30].

Let the ith recovered image $\hat{\mathbf{I}}_i = \{I_i(l) \mid l \in \hat{\mathbf{I}}\}$ and $i = 1, 2, \ldots, M$, the pixel-based NM algorithm maps the evaluated image source $\hat{\mathbf{I}}_i$ into $\tilde{\mathbf{I}}_i$ as follows:

$$\tilde{I}_i(l) = \frac{\Sigma_{k \in \mathcal{S}_l} w_i(k,l)\hat{I}_i(k)}{\Sigma_{k \in \mathcal{S}_l} w_i(k,l)} \tag{12}$$

where $\mathcal{S}_l$ is the search region around the pixel l and $w_i(k,l)$ are the weights that compare the neighborhoods around pixels l and k, respectively. The weights are defined as

$$w_i(k,l) = e^{-\frac{\Sigma_{b \in \mathcal{B}}(\hat{I}_i(k+b) - \hat{I}_i(l+b))^2}{h^2}} \tag{13}$$

where $\mathcal{B}$ defines the neighborhood, B is the total size of $\mathcal{B}$, and the parameter h acts as a degree of filtering.

Finally, we estimate the higher quality underlying image sources $\tilde{\mathbf{I}}_i = \{\tilde{\mathbf{I}}_1, \tilde{\mathbf{I}}_2, \ldots, \tilde{\mathbf{I}}_M\}$.

IV. Numerical Experiments

In this section, we carry out numerical simulations on the astronomy images captured by the NASA and the oceanographic images to demonstrate the performance of the proposed BMIPR algorithm depends on the length of the window and the maximal overlapping between adjacent windows, and how the algorithm is affected by noise for solving the BMIPR problem.

The relative root-mean-squared error (RRMSE), the correlation coefficient, and signal-to-noise ratio (SNR) are used to evaluate the performance of the proposed algorithms.

For $i = 1, 2$, RRMSE_i is defined as follows:

$$\text{RRMSE}_i = \frac{\text{RMS}(\mathbf{I}_i - \tilde{\mathbf{I}}_i)}{\text{RMS}(\mathbf{I}_i)} \tag{14}$$

where $\tilde{\mathbf{I}}_i$ is the recovered underlying image and $\mathbf{I}_i$ is the original image source.

The correlation coefficient is similar in nature to the convolution of two functions. For $i = 1, 2$, r_i can be defined as

$$r_i = \frac{n\sum_i \mathbf{I}_i\tilde{\mathbf{I}}_i - \sum_i \mathbf{I}_i \sum_i \tilde{\mathbf{I}}_i}{\sqrt{n\sum_i \mathbf{I}_i^2 - (\sum_i \mathbf{I}_i)^2}\sqrt{n\sum_i \tilde{\mathbf{I}}_i^2 - (\sum_i \tilde{\mathbf{I}}_i)^2}}. \tag{15}$$

Algorithm 1: BMIPR Algorithm.

Input: The phaseless STFT measurements $\mathbf{Z}$ as given in (2), a low-pass interpolation filter with bandwidth R as depicted in (8).

Output: Initialization of $\mathbf{x}_0$, recovery of $\hat{\mathbf{X}}$, $\bar{\mathbf{I}}$, and $\tilde{\mathbf{I}}$.

1: *DFT*. Compute DFT $\mathbf{Y}_{ij}$ of the STFT $\mathbf{Z}_{ij}$ as shown in (3).

2: *Up-sampling*. For $W \le l \le (N - W)$,
if $L = 1$, omit this step,
else $L > 1$, $\mathbf{y}_{ijl}$ is expanded and interpolated to the up-sampled version $\tilde{\mathbf{y}}_{ijl}$ by (8).

3: *Initialization*. Construct an initial matrix $\mathbf{F}_{ij0}$.

$$\min_{\mathbf{x}_{ij} \in \mathbb{C}^N} \sum_{l=-a}^{a} \|\mathbf{y}_{ijl} - \mathbf{G}_l \text{diag}(\mathbf{F}_{ij0}, l)\|^2 + \lambda \frac{\|\text{diag}(\mathbf{F}_{ij0}, l)\|_1}{\|\text{diag}(\mathbf{F}_{ij0}, l)\|_2}$$

where $a = W - 1$, $\mathbf{G}_l$, $\mathbf{y}_{ijl}$, and λ are defined as in (6) and (7). Find the initialization $\mathbf{x}_{ij0}$ by the eigenvector decomposition of $\mathbf{F}_{ij0}$.

4: *Recovery* of $\hat{\mathbf{X}}$. Recover the mixed image source $\mathbf{X}_i$ by a GD algorithm by minimizing a nonconvex function as given in (5). For each $\mathbf{Z}_i$, repeat steps 1–4 until all the mixed image sources have been recovered, and then normalized as $\hat{\mathbf{X}} = \{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \ldots, \hat{\mathbf{X}}_M\}$.

5: *Whitening*. Estimate a whitening matrix $\mathbf{U}$ and the whitening mixed signal $\tilde{\mathbf{x}}_i = \mathbf{U}\mathbf{x}_i$.

6: *Orthogonalization*. Search for an orthogonal matrix $\mathbf{W}$. The optimal weights are determined by $\mathbf{w}_{\text{opt}} = \arg \max_{\|\mathbf{w}\|^2=1} E\{|G(\mathbf{w}^*\tilde{\mathbf{x}}_i)|^2\}$.

7: *Normalization*. To calculate the optimal weights, a gradient optimization algorithm is used and followed by a normalization step as in (11).

8: *Evaluation* of $\hat{\mathbf{I}}$. Estimate the mixing matrix $\mathbf{A} = \mathbf{U}^{-1}\mathbf{W}^*$ and recover the underlying image sources $\hat{\mathbf{I}} = \{\hat{\mathbf{I}}_1, \hat{\mathbf{I}}_2, \ldots, \hat{\mathbf{I}}_M\}$ by $\mathbf{S} = \mathbf{A}^{-1}\tilde{\mathbf{X}}$ and a conversion from a vector to a matrix.

9: *Estimation* of $\tilde{\mathbf{I}}$. Estimate the higher-quality underlying image sources $\tilde{\mathbf{I}}_i = \{\tilde{\mathbf{I}}_1, \tilde{\mathbf{I}}_2, \ldots, \tilde{\mathbf{I}}_M\}$ by (12).

A. BMIPR for the Astronomy Phaseless STFT Measurements

The astronomy images are collected normally by the far-distance measurement methods and rich in phase information. Sometimes the phases of the astronomy images may be corrupted by some noises and interferences.

The original astronomy images considered for the first simulation are captured by the NASA and can be found online.[1] The mixed phaseless STFT measurements are corrupted by additive Gaussian noises with zero mean and unit variance with the level of noise from 5 to 25 dB.

[1][Online] Available: https://www.nasa.gov/multimedia/imagegallery/index.html).

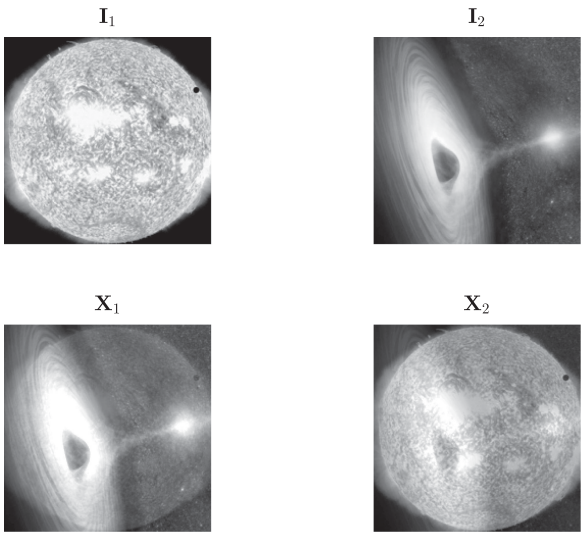


Fig. 4. Image sources **I** (size 203×203) and the mixed images **X**.

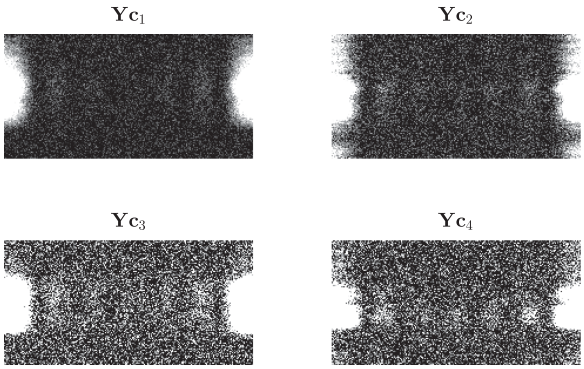


Fig. 5. Four phaseless STFT images.

In the first set of simulations, we evaluate the estimation performance of the proposed algorithm described in Algorithm 1. We choose the maximal overlapping between adjacent windows ($L = 2$), the window length ($W = 7$), the image size (203×203), and the noise at SNR $= 25$ dB. The number of short-time windows is denoted by $R = \lceil N/L \rceil$. The gradient step size μ is set to be 0.005, the regularization coefficient λ is set to be 0.01, the maximal number of iterations for stopping Algorithm 1 is 5000.

Fig. 4 shows the image sources $\mathbf{I} = \{\mathbf{I}_1, \mathbf{I}_2\}$ and the mixed images $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2\}$.

We give four example images of the phaseless STFT images converted from $\mathbf{X}$ in Fig. 5.

Fig. 6 demonstrates the recovered mixed images $\hat{\mathbf{X}} = \{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2\}$ and the mixed images $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2\}$. It shows that the capability of the proposed algorithm in recovering the phases of the mixed images.

In Fig. 7, we can see that the evaluated underlying images $\hat{\mathbf{I}} = \{\hat{\mathbf{I}}_1, \hat{\mathbf{I}}_2\}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}} = \{\tilde{\mathbf{I}}_1, \tilde{\mathbf{I}}_2\}$ are similar to the image sources.

As shown in Fig. 8, the final RRMSEs of the recovered underlying images $\tilde{\mathbf{I}} = \{\tilde{\mathbf{I}}_1, \tilde{\mathbf{I}}_2\}$ are less than 0.13. The correlation coefficients of the image sources and the underlying images are

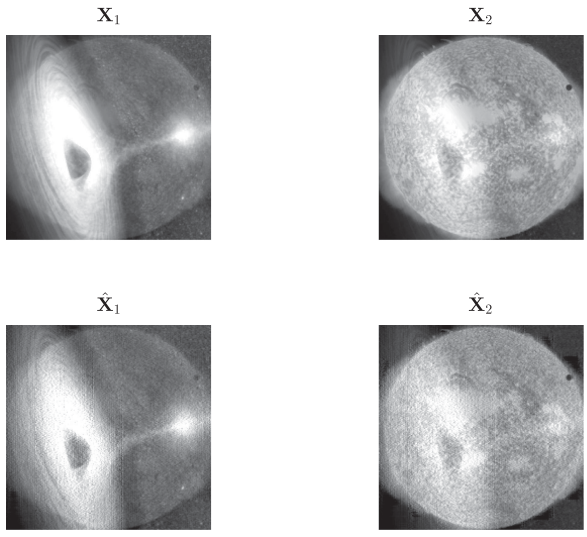


Fig. 6. Recovered mixed images $\hat{\mathbf{X}}$ and the mixed images **X**.

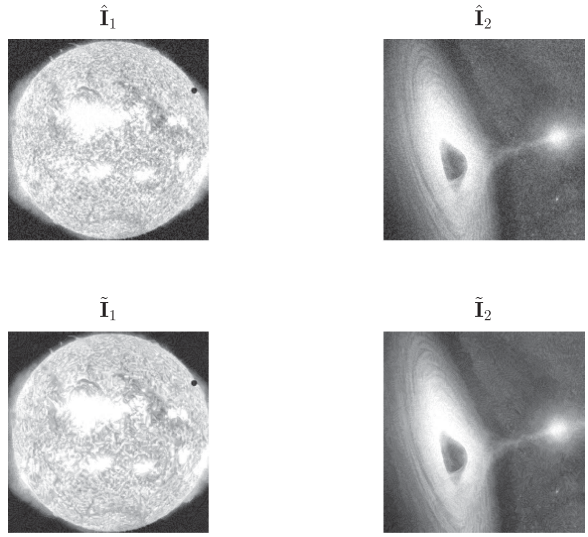


Fig. 7. Evaluated underlying images $\hat{\mathbf{I}}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}}$.

0.7874 and 0.8102, whereas those of the image sources and the higher quality underlying images are 0.9203 and 0.9422, respectively. These demonstrate the effectiveness of the proposed algorithm in recovering the phases of multiple underlying astronomy images.

We also estimated the performance of the proposed BMIPR algorithm with respect to different types of noise and levels of noise (in terms of SNRs). Fig. 9 demonstrates the average RRMSE for 50 experiments of the evaluated underlying images $\hat{\mathbf{I}}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}}$ for Gaussian noise, Laplacian noise, and Poisson noise at different SNR when $L = 2$ and $W = 7$. With the increase in SNR, the RRMSEs of $\hat{\mathbf{I}}$ and $\tilde{\mathbf{I}}$ drop rapidly. For SNR $= 20, 25$ dB, the RRMSEs of $\tilde{\mathbf{I}}$ are relatively low and less than 0.15. The results mean that the proposed algorithm has better anti-noise performance in $\tilde{\mathbf{I}}$ than $\hat{\mathbf{I}}$ especially for high values of SNR.

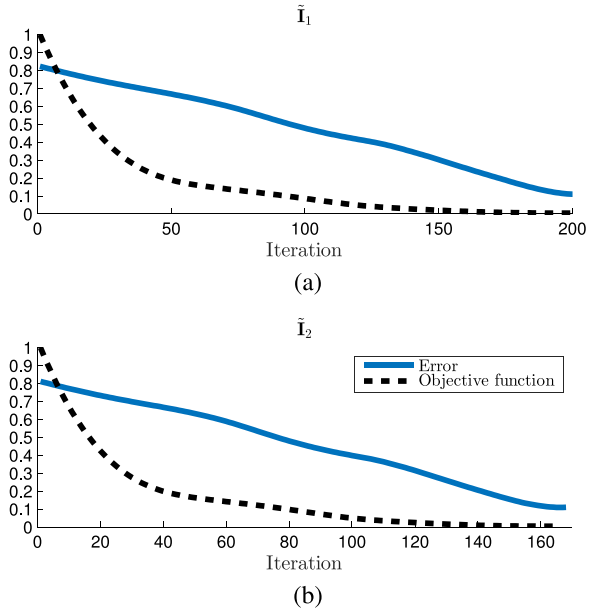


Fig. 8. RRMSE and the normalized objective function values of the recovered underlying images $\tilde{\mathbf{I}}_1$ and $\tilde{\mathbf{I}}_2$.

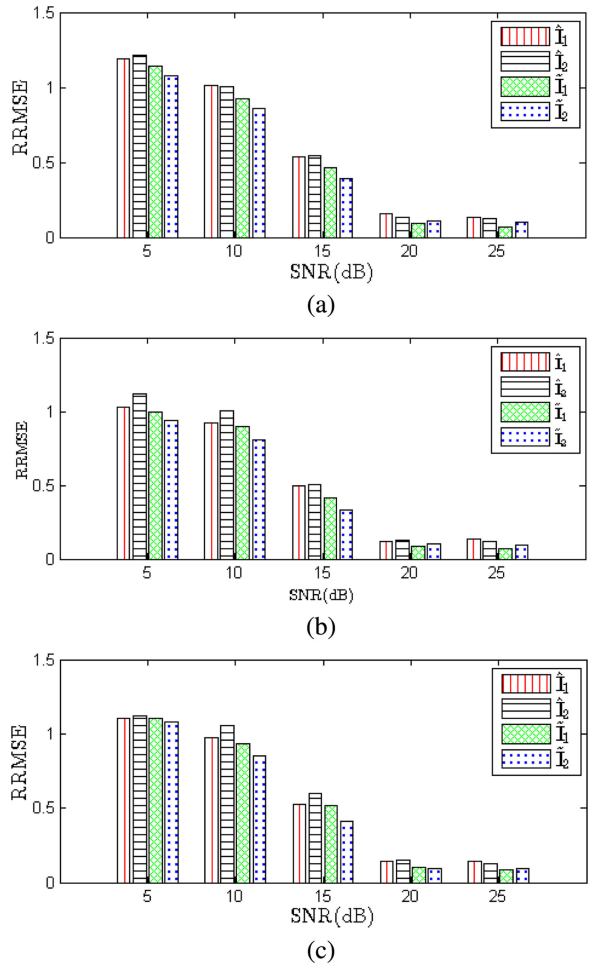


Fig. 9. Average final RRMSE for 50 experiments of the evaluated underlying images $\hat{\mathbf{I}}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}}$ for different types of noise at different SNRs. (a) Gaussian noise. (b) Laplacian noise. (c) Poisson noise.

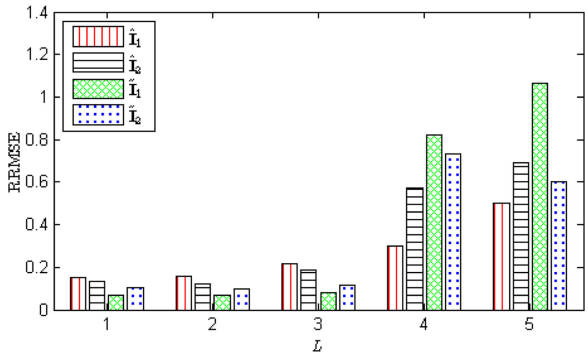


Fig. 10. Average final RRMSE for 50 experiments of the evaluated underlying images $\hat{\mathbf{I}}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}}$ for different L.

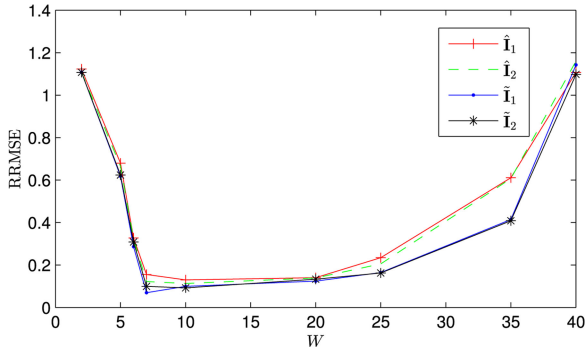


Fig. 11. Average final RRMSE for 50 experiments of the evaluated underlying images $\hat{\mathbf{I}}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}}$ for different W.

Fig. 10 presents the average RRMSE for 50 experiments of the evaluated underlying images $\hat{\mathbf{I}}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}}$ for different L when SNR $= 25$ dB and $W = 7$. For low values of $1 \leq L \leq 3$, the RRMSEs of $\hat{\mathbf{I}}$ and $\tilde{\mathbf{I}}$ are relatively low (at around 0.2 and 0.15, respectively), and in this case, the underlying images are well-recovered. For high values of $L = 4, 5$, the RRMSE of $\hat{\mathbf{I}}$ is smaller than that of $\tilde{\mathbf{I}}$ and the values are mostly above 0.5. The proposed algorithm has better performance in recovering $\tilde{\mathbf{I}}$ than $\hat{\mathbf{I}}$ only for low values of L.

Fig. 11 presents the average RRMSE for 50 experiments of the evaluated underlying images $\hat{\mathbf{I}}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}}$ for different W when $L = 2$, and SNR $=$ 25 dB. For $W \leq 6$ and $W \geq 25$, the RRMSEs of $\hat{\mathbf{I}}$ and $\tilde{\mathbf{I}}$ are relatively high and the values are mostly above 0.16. For middle values of $6 < W < 25$, the RRMSE of $\hat{\mathbf{I}}$ and $\tilde{\mathbf{I}}$ are about 0.1. The proposed algorithm has better performance in recovering $\tilde{\mathbf{I}}$ than $\hat{\mathbf{I}}$ for $6 < W < 25$.

B. RGB BMIPR for the Oceanographic Phaseless STFT Measurements

In the marine photogrammetric survey, many sea animals, boats, and rocks are filmed by the aerial photoing or other photography from a great distance. The oceanographic images

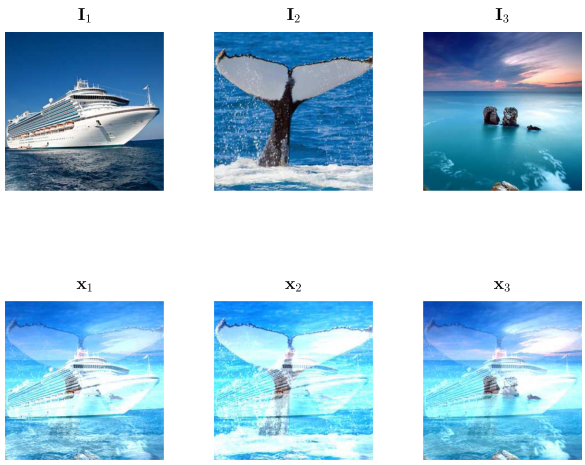


Fig. 12. Oceanographic image sources **I** and the mixed images **X**.

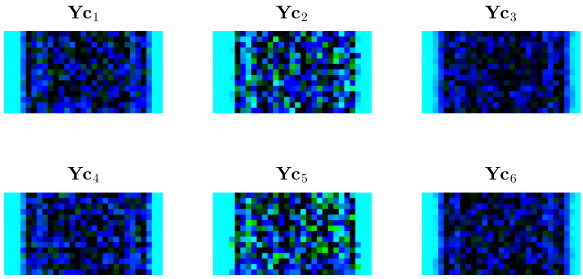


Fig. 13. Six phaseless STFT images.

should contain abundant phase information. However, the phases of these images may be corrupted by some interferences.

In the second set of simulations, we choose the image size ($290 \times 290 \times 3$), the maximal overlapping between adjacent windows ($L = 2$), the window length ($W = 7$), and the noise at SNR $= 25$ dB. The number of short-time windows is denoted by $R = \lceil N/L \rceil$. The gradient step size μ is set to be 0.005, the regularization coefficient λ is set to be 0.01, and the maximal number of iterations for stopping Algorithm 1 is 5000.

Fig. 12 shows the image sources $\mathbf{I} = \{\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3\}$ and the mixed images $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$.

Six example images of the phaseless STFT images converted from **X** are shown in Fig. 13.

Fig. 14 demonstrates the phase retrieval capability of the proposed algorithm.

As shown in Figs. 12 and 15, the evaluated underlying images $\hat{\mathbf{I}} = \{\hat{\mathbf{I}}_1, \hat{\mathbf{I}}_2, \hat{\mathbf{I}}_3\}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}} = \{\tilde{\mathbf{I}}_1, \tilde{\mathbf{I}}_2, \tilde{\mathbf{I}}_3\}$ resemble the oceanographic image sources $\mathbf{I} = \{\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3\}$.

The final RRMSEs of $\tilde{\mathbf{I}}_1$, $\tilde{\mathbf{I}}_2$, and $\tilde{\mathbf{I}}_3$ are 0.1502, 0.1791, and 0.1377, respectively. The correlation coefficients of the image sources and the higher quality underlying images are 0.9251, 0.9076, and 0.9679. This demonstrates the efficiency of the proposed BMIPR algorithm in recovering the phases of the oceanographic images.

In Table I, the phase retrieval performance of the proposed BMIPR algorithm, PhaseLamp algorithm [16], and PhaseLift algorithm [31] are compared for SNR $= 25$ dB by three different images. The codes of the PhaseLamp algorithm and PhaseLift algorithm are obtained from a phase retrieval library PhasePack [32]. It shows clearly that the phase retrieval performance of the proposed BMIPR algorithm outperforms the other two

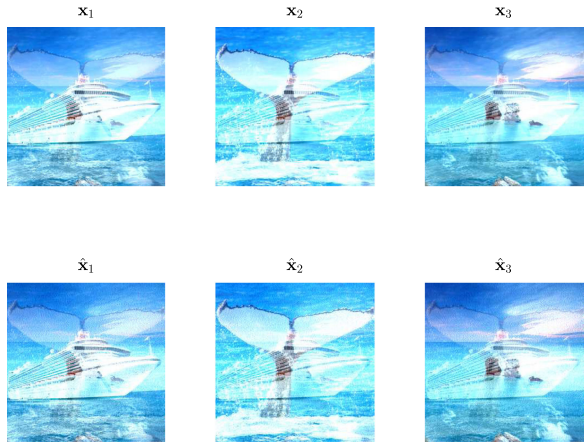


Fig. 14. Recovered mixed images $\hat{\mathbf{X}} = \{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \hat{\mathbf{X}}_3\}$ and the mixed images $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$.

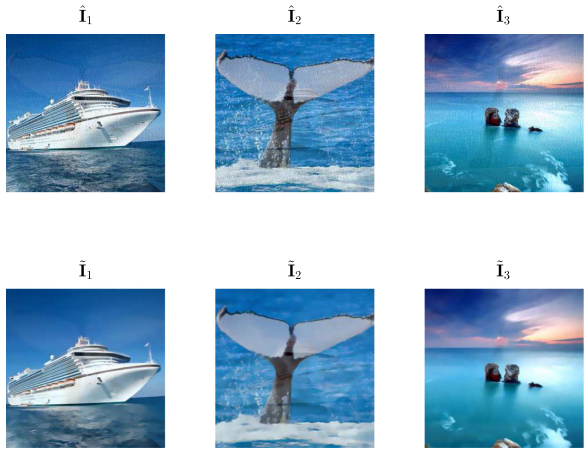


Fig. 15. Evaluated underlying images $\hat{\mathbf{I}}$ and the estimated higher quality underlying images $\tilde{\mathbf{I}}$.

TABLE I
COMPARISON OF THE PERFORMANCE FOR THE BMIPR ALGORITHM, PHASELAMP ALGORITHM, AND PHASELIFT ALGORITHM FOR SNR $= 25$ DB

Image	Algorithm	RRMSE	PSNR(dB)	Correlation	Time(s)
Boat	BMIPR	0.1502	35.69	0.9251	619
	PhaseLamp	0.1747	22.97	0.8554	622
	PhaseLift	0.2334	21.01	0.7926	947
Whale	BMIPR	0.1791	25.03	0.9076	679
	PhaseLamp	0.1811	25.04	0.9027	771
	PhaseLift	0.2232	21.53	0.8111	1211
Stone	BMIPR	0.1377	35.74	0.9679	563
	PhaseLamp	0.1536	28.88	0.921	576
	PhaseLift	0.2082	22.52	0.8263	875

TABLE II
COMPARISON OF THE PERFORMANCE FOR THE BMIPR ALGORITHM, PHASELAMP ALGORITHM, AND PHASELIFT ALGORITHM FOR SNR = 20 DB

Image	Algorithm	RRMSE	PSNR(dB)	Correlation	Time(s)
Boat	BMIPR	0.205	22.43	0.8789	741
	PhaseLamp	0.3273	15.86	0.6618	744
	PhaseLift	0.3812	11.52	0.6152	1053
Whale	BMIPR	0.2917	18.43	0.8553	875
	PhaseLamp	0.4136	10.41	0.5264	901
	PhaseLift	0.4088	11.28	0.582	1342
Stone	BMIPR	0.1822	24.93	0.9012	815
	PhaseLamp	0.2966	18.21	0.6827	818
	PhaseLift	0.3284	15.34	0.6513	952

algorithms for SNR = 25 dB in terms of RRMSE, PSNR, and correlation.

Table II gives a phase retrieval performance comparison of the proposed BMIPR algorithm, PhaseLamp algorithm [16], and PhaseLift algorithm [31] for SNR = 20 dB by three different images. Again, we observe the similar advantages using the proposed BMIPR algorithm over the two baseline algorithms. In terms of computational complexity, the proposed BMIPR algorithm is more efficient than the PhaseLift and PhaseLamp algorithms.

V. CONCLUSION

The model and the algorithm for the problem of blind multi-image phase retrieval from multiple mixed phaseless STFT image measurements were investigated in this paper. Our contributions to this challenging problem are as follows.

Model: We formed a new model for the problem of blind multi-image phase retrieval from multiple mixed phaseless STFT image measurements.

Algorithm: Due to the absence of Fourier phase information and mixed information, a BMIPR algorithm, which combined a GD algorithm by minimizing a nonconvex loss function with an improved CMN algorithm and an NM algorithm, was presented as a solution to BMIPR problem.

Initialization: We showed the significance of the initialization method to the GD-ICA algorithm and demonstrated that the initialization method could be constructed by an improved LS loss function with a penalty term ($\frac{\ell_1}{\ell_2}$).

Numerical experiments showed that the proposed algorithms performed well in estimating the phases of multiple image sources and the mixing information. In terms of future research directions, it was interesting to investigate how to incorporate conditions, such as window length, additional magnitude-only measurement, mixing mode, or maximal overlapping between adjacent windows into the BMIPR algorithm. It was also tempting to consider different mixing mode and extend a fast algorithm for recovering multiple underlying image sources from multiple mixed phaseless STFT image measurements.

APPENDIX

A. Proof of Theorem 3.1

The problem of STFT phase retrieval is equivalent to the short-time autocorrelation $\mathbf{Y}$ by taking an N-point DFT of the phaseless STFT measurement $\mathbf{Z}$

$$Y(\tau,l)=\frac{1}{N}\sum_{k=0}^{N-1}Z(\tau,k)e^{-2j\pi kl/N}$$
$$=\sum_{n=0}^{N-1}x(n)x^*(n+l)g(\tau L-n)g(\tau L-n-l)$$

where $0\le l\le N-1$ and $0\le\tau\le R-1$.

If the window length is W, the values of $\mathbf{Y}$ are nonzero only in the interval $0\le l\le W-1$ and $N-W+1\le l\le N-1$. By circularly shifting $\mathbf{Y}$ by $W-1$ rows, a signal $\mathbf{H}$ that has nonzero values only in the interval $0\le l\le 2(W-1)$ is obtained.

Let $\mathbf{b}$ be the sub-matrix of the DFT matrix $\mathbf{Y}$ acquired by considering the first $2W-1$ rows and $2W-1$ columns (the Vandermonde structure is retained). Since $\mathbf{b}$ is invertible, $Z(\tau,k)$ for $0\le k\le 2(W-1)$ and $H(\tau,l)$ for $0\le l\le 2(W-1)$ are related by an invertible matrix. Note that $Z(\tau,k)$ for $0\le k\le N-1$ can be trivially calculated from $H(\tau,l)$ for $0\le l\le 2(W-1)$.

This proof gives evidence that $\mathbf{Z}$ is sufficient to calculate $\mathbf{Y}$ [17].

B. Proof of Theorem 3.2

Lemma 1 means that any locally sparse solution is the sparsest solution in essence locally.

Lemma 1: $\forall\mathbf{z}\in\mathcal{F}_L,\exists\delta>0$ such that $\forall\mathbf{y}\in\mathcal{F}$, if $0<\|\mathbf{y}-\mathbf{z}\|_2<\delta$, we have $\mathcal{S}(\mathbf{z})\subset\mathcal{S}(\mathbf{y})$.

Let $\mathbf{y}=\mathbf{z}+\mathbf{v}$ and select $\delta=\min_{i\in\mathcal{S}(\mathbf{z})}\{\mathbf{z}_i\}$, then

$$\|v\|_\infty\le\|v\|_2<\min_{i\in\mathcal{S}(\mathbf{z})}\{\mathbf{z}_i\}.$$

Hence

$$\mathbf{y}_i\ge\mathbf{z}_i-\|v\|_\infty>\mathbf{z}_i-\min_{i\in\mathcal{S}(\mathbf{z})}\{\mathbf{z}_i\ge 0\quad\forall i\in\mathcal{S}(\mathbf{z}).$$

The above-mentioned equations mean

$$\mathcal{S}(\mathbf{z})\subseteq\mathcal{S}(\mathbf{y}).$$

However, $\mathcal{S}(\mathbf{z})\ne\mathcal{S}(\mathbf{y})$ for $\mathbf{z}\in\mathcal{F}_L$. Then, $\min_{\mathbf{z}\ge 0}\frac{\|\mathbf{z}\|_1}{\|\mathbf{z}\|_2}$ that subjects to $\mathbf{y}_{ijl}=\mathbf{G}_l\mathbf{z}$ must be locally sparse, thereby being at least locally the sparsest feasible solution.

C. Proof of Theorem 3.3

The derivative of the function $J=J(s_i)$ is calculated based on real-valued functions because J is not analytic. Let $s_i=\mathbf{w}^*\tilde{\mathbf{x}}_i$, $G(s_i)$ is expanded in terms of two real-valued functions $u(y)$ and $v(y)$

$$J=|G(s_i)|^2=|G(\mathbf{w}^*\tilde{\mathbf{x}}_i)|^2=|u(\mathbf{w}^*\tilde{\mathbf{x}}_i)+jv(\mathbf{w}^*\tilde{\mathbf{x}}_i)|^2$$
$$\equiv u^2(a,b)+v^2(a,b)$$

where a and b are the real part and imaginary part of $\mathbf{w}^*\tilde{\mathbf{x}}_i$, respectively. The partial derivative of J with respect to the real weight w_i^R is obtained by use of the chain rule and results in

$$\frac{\partial J}{\partial w_i^R} = 2u\left(\frac{\partial u(a,b)}{\partial a}\frac{\partial a}{\partial w_i^R} + \frac{\partial u(a,b)}{\partial b}\frac{\partial b}{\partial w_i^R}\right) + 2v\left(\frac{\partial u(a,b)}{\partial a}\frac{\partial a}{\partial w_i^R} + \frac{\partial u(a,b)}{\partial b}\frac{\partial b}{\partial w_i^R}\right).$$

The above-mentioned expression is rearranged as follows:

$$\frac{\partial J}{\partial w_i^R} = 2u(u_a\tilde{x}_{ij}^R + u_b\tilde{x}_{ij}^I) + 2v(v_a\tilde{x}_{ij}^R + v_b\tilde{x}_{ij}^I) = 2[\tilde{x}_{ij}^I(uu_a + vv_a) + \tilde{x}_{ij}^R(uu_b + vv_b)]$$

where $u_a \equiv \frac{\partial u(a,b)}{\partial a}$, $u_b \equiv \frac{\partial u(a,b)}{\partial b}$, $v_a \equiv \frac{\partial v(a,b)}{\partial a}$, and $v_b \equiv \frac{\partial v(a,b)}{\partial b}$.

The derivative of J with respect to the imaginary weight w_i^I is calculated as

$$\frac{\partial J}{\partial w_i^I} = 2[\tilde{x}_{ij}^I(uu_a + vv_a) - \tilde{x}_{ij}^R(uu_b + vv_b)].$$

It is advantageous to utilize complex operators for a more compact notation. Noticing that

$$(uu_a + vv_a) + j(uu_b + vv_b) = g^*(\mathbf{w}^*\tilde{\mathbf{x}}_i)G(\mathbf{w}^*\tilde{\mathbf{x}}_i)$$

where g is the derivative of G. According to the Cauchy–Riemann equations: $g^R = u_a = v_b$ and $g^I = v_a = -u_b$, the derivative of J with respect to the complex weight vector $\mathbf{w}$ is shown as

$$\frac{\partial J}{\partial \mathbf{w}} = 2(g^*(\mathbf{w}^*\tilde{\mathbf{x}}_i)G(\mathbf{w}^*\tilde{\mathbf{x}}_i))^*\tilde{\mathbf{x}}_i = 2\tilde{\mathbf{x}}_i G^*(\mathbf{w}^*\tilde{\mathbf{x}}_i)g(\mathbf{w}^*\tilde{\mathbf{x}}_i)$$

which is equivalent to

$$\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} = \frac{1}{2}\left(\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}^R} + j\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}^I}\right).$$

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their constructive comments for improving this paper.

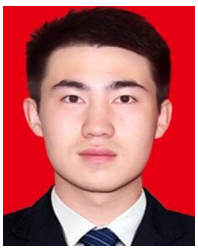
REFERENCES

[1] B. Yu *et al.*, "Evaluation of phase retrieval approaches in magnified x-ray phase nano computerized tomography applied to bone tissue," *Opt. Express*, vol. 26, no. 9, pp. 11110–11124, 2018.
[2] P. Baran *et al.*, "High-resolution x-ray phase-contrast 3-D imaging of breast tissue specimens as a possible adjunct to histopathology," *IEEE Trans. Med. Imag.*, vol. 37, no. 12, pp. 2642–2650, Dec. 2018.
[3] J. R. Fienup, "Phase-retrieval algorithms for a complicated optical system," *Appl. Opt.*, vol. 32, no. 10, pp. 1737–1746, 1993.
[4] L. J. Wright, M. Karpinski, C. Soller, and B. J. Smith, "Spectral shearing of quantum light pulses by electro-optic phase modulation," *Phys. Rev. Lett.*, vol. 118, no. 2, 2017, Art. no. 023601.
[5] N. Arnaud *et al.*, "Controlling kilometre-scale interferometric detectors for gravitational wave astronomy: Active phase noise cancellation using EOMs," *Nucl. Instrum. Methods Phys. Res. Sect. A*, vol. 845, pp. 347–349, 2017.
[6] R. Abuter *et al.*, "First light for GRAVITY: Phase referencing optical interferometry for the very large telescope interferometer," *Astron. Astrophys.*, vol. 602, 2017, Art. no. A94.
[7] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1711–1732, Mar. 2014.
[8] T. J. Lee and Y. C. Ko, "Channel estimation and data detection in the presence of phase noise in MIMO-OFDM systems with independent oscillators," *IEEE Access*, vol. 5, pp. 9647–9662, 2017.
[9] A. Ahmed, A. Aghasi, and P. Hand, "Blind deconvolutional phase retrieval via convex programming," in *Proc. Ann. Conf. Adv. Neural Inform. Process. Syst.*, 2018, pp. 10 051–10 061.
[10] Q. Qu, Y. Q. Zhang, Y. C. Eldar, and J. Wright, "Convolutional phase retrieval via gradient descent," in *Proc. Ann. Conf. Adv. Neural Inform. Process. Syst.*, 2017.
[11] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Uniqueness guarantees and recovery algorithms," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2402–2410, May 2017.
[12] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, no. 2, pp. 237–246, 1972.
[13] E. J. Candes and P. Sur, "The phase transition for the existence of the maximum likelihood estimate in high-dimensional logistic regression," 2018, arXiv:1804.09753.
[14] T. Bendory, D. Edidin, and Y. C. Eldar, "Blind phaseless short-time Fourier transform recovery," 2018, arXiv:1808.07414.
[15] T. Goldstein and C. Studer, "PhaseMax: Convex phase retrieval via basis pursuit," *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 2675–2689, Apr. 2018.
[16] O. Dhifallah, C. Thrampoulidis, and Y. M. Lu, "Phase retrieval via linear programming: Fundamental limits and algorithmic improvements," in *Proc. IEEE 55th Ann. Allerton Conf. Commun., Cont., Comput. (Allerton)*, 2017, pp. 1071–1077.
[17] K. Jaganathan, Y. Eldar, and B. Hassibi, "STFT phase retrieval: Uniqueness guarantees and recovery algorithms," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 4, pp. 770–781, Jun. 2016.
[18] T. Bendory, Y. C. Eldar, and N. Boumal, "Nonconvex phase retrieval from STFT measurements," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 467–484, Jan. 2018.
[19] K. Jaganathan, Y. C. Eldar, and B. Hassibi, "Phase retrieval: An overview of recent developments," 2015, arXiv:1510.07713.
[20] B. Rajaei, E. W. Tramel, S. Gigan, F. Krzakala, and L. Daudet, "Intensity-only optical compressive imaging using a multiply scattering material and a double phase retrieval approach," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 4054–4058.
[21] G. Cheng, S. Cheng, J. Tan, X. Bao, S. Liu, and Z. Liu, "A robust multi-image phase retrieval," *Opt. Lasers Eng.*, vol. 101, pp. 16–22, 2018.
[22] Y. Guo, A. Wang, and W. Wang, "Multisource phase retrieval from multichannel phaseless STFT measurements," *Signal Process.*, vol. 144, pp. 36–40, 2018.
[23] D. Lyon, "The discrete Fourier transform, Part 6: Cross-correlation," *J. Object Technol.*, vol. 9, no. 2, pp. 17–22, 2010.
[24] P. Yin, E. Esser, and J. Xin, "Ratio and difference of l1 and l2 norms and sparse representation with coherent dictionaries," *Commun. Inf. Syst.*, vol. 14, no. 2, pp. 87–109, 2014.
[25] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York, NY, USA: Wiley, 2004.
[26] Y. Guo, S. Huang, and Y. Li, "Single-mixture source separation using dimensionality reduction of ensemble empirical mode decomposition and independent component analysis," *Circuits, Syst., Signal Process.*, vol. 31, no. 6, pp. 2047–2060, 2012.
[27] Y. Guo, S. Huang, Y. Li, and G. R. Naik, "Edge effect elimination in single-mixture blind source separation," *Circuits, Syst., Signal Process.*, vol. 32, no. 5, pp. 2317–2334, 2013.
[28] M. Novey and T. Adali, "ICA by maximization of nonGaussianity using complex functions," in *Proc. IEEE Workshop Mach. Learn. Signal Process.*, 2005, pp. 21–26.
[29] M. Novey and T. Adali, "Complex ICA by negentropy maximization," *IEEE Trans. Neural Netw.*, vol. 19, no. 4, pp. 596–609, Apr. 2008.
[30] D. Van De Ville and M. Kocher, "Sure-based nonlocal means," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 973–976, Nov. 2009.
[31] E. J. Candes, T. Strohmer, and V. Voroninski, "PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. Pure Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, 2013.
[32] R. Chandra, Z. Zhong, J. Hontz, V. McCulloch, C. Studer, and T. Goldstein, "PhasePack: A phase retrieval library," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, 2017, pp. 1617–1621.

Yina Guo (M'16) received the B.Sc. degree in electronics and information technology from the China University of Mining and Technology, Xuzhou, China, in 2002, and the M.E. and Ph.D. degrees in communication engineering from the Taiyuan University of Science and Technology, Taiyuan, China, in 2007 and 2014, respectively.

She is currently a Professor in Signal and Information Processing with the Taiyuan University of Science and Technology. She has authored/coauthored more than 30 papers and two books, and was granted six patents, and four software copyrights in China. Her research interests include blind source separation, biosignal processing, and phase retrieval.

Prof. Guo was the recipient of several science and technology awards from Shanxi Province and holds grants from the National Natural Science Foundation of China and Shanxi Province. She is a Senior Member of the China Institute of Communications and a member of the IEEE Signal Processing Society.



Xiangning Zhao received the B.E. degree in electronic engineering from Shanxi University, Taiyuan, China, in 2016. He is currently working toward the master's degree in electronics and communication engineering with the Taiyuan University of Science and Technology, Taiyuan.

His research interests include image signal processing and phase retrieval.

Jianyu Li received the B.E. degree in electronic engineering in 2016 from the Taiyuan University of Science and Technology, Taiyuan, China, where he is currently working toward the master's degree in electronics and communication engineering.

His research interests include blind source separation and phase retrieval.

Anhong Wang received the B.Sc. and M.E. degrees in electronic information engineering from the Taiyuan University of Science and Technology, Taiyuan, China, in 1994 and 2002, respectively, and the Ph.D. degree in information science from Beijing Jiaotong University, Beijing, China, in 2009.

She is currently a Professor in Signal and Information Processing and the Director of the Institute of Digital Media and Communication, Taiyuan University of Science and Technology. She has authored/coauthored more than 90 papers in international journals and conferences. She is leading several research projects, including two funded by the National Science Foundations of China. Her research interests include image and video coding and transmission, compressed sensing, and secret image sharing.

Wenwu Wang (M'02–SM'11) received the B.Sc. degree in automatic control, the M.E. degree in control science and engineering, and the Ph.D. degree in navigation, guidance, and control from Harbin Engineering University, Harbin, China, in 1997, 2000, and 2002, respectively.

He is currently a Reader in Signal Processing with the University of Surrey, Surrey, U.K., and a Co-Director of the Machine Audition Laboratory within the Centre for Vision, Speech and Signal Processing. He has been a Guest Professor with Qingdao University of Science and Technology, Qingdao, China, since 2018. He has coauthored more than 250 publications. His research interests include signal processing and machine learning.

Prof. Wang has been a Senior Area Editor since 2019 and an Associate Editor from 2014 to 2018 for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is a Publication Co-Chair for the International Conference on Acoustics, Speech, and Signal Processing 2019, Brighton, U.K.