Constrained phase retrieval: When alternating projection meets regularization

4 authors, including:

Shi Baoshun
Yan Shan University
21 PUBLICATIONS   **57** CITATIONS

Journal of the
Optical Society
of America **B**

OPTICAL PHYSICS

Constrained phase retrieval: when alternating projection meets regularization

Baoshun Shi,[1,2] **Qiusheng Lian,**[1,2,*] **Xin Huang,**[3,4] **and Ni An**[5]

[1]*School of Information Science and Engineering, Yanshan University, Qinhuang Dao, Hebei Province 066004, China*
[2]*Hebei Key Laboratory of Information Transmission and Signal Processing, Yanshan University, Qinhuang Dao, Hebei Province 066004, China*
[3]*School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei Province 430079, China*
[4]*State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei Province 430079, China*
[5]*Department of Electrical and Computer Engineering, Drexel University, Philadelphia, Pennsylvania 19104, USA*
Corresponding author: lianqs@ysu.edu.cn

Simplicity and few finely tuned parameters are the main advantages of alternating projection (AP) methods, a fundamental class of phase retrieval (PR) methods in the optical imaging field. However, AP methods often suffer from low-quality imaging when few diffraction patterns are recorded. Regularized PR methods avoid this deficiency by using some proper regularization models, but many finely tuned parameters are needed. In this work, we propose a novel unified framework called constrained PR (ConPR), which brings the AP method and the regularization approach together. The proposed ConPR framework not only can recover high-quality images from few diffraction patterns, but also does not need fine-tuning of the parameters. Our proposed generalized constrained PR optimization model consists of a relation function term, a regularization term, and a measurement constraint. The measurement constraint ensures that the recovered image matches with the measurement, and the regularization term can impose some desirable properties on the recovered image. The relation function promotes the approximation of the two underlying variables. The sparsity under the block-matching and 3D filtering frame is incorporated into the proposed ConPR framework. The problem formulation consists of an image updating sub-problem and a constrained optimization sub-problem. The epigraph set of the data fidelity function is defined, and the constrained optimization sub-problem is solved via the epigraph concept. Diffraction imaging from one noisy coded diffraction pattern demonstrates the effectiveness of the proposed algorithm.   © 2018 Optical Society of America

OCIS codes: (100.5070) Phase retrieval; (110.1758) Computational imaging; (110.3010) Image reconstruction techniques; (100.3200) Inverse scattering; (100.3190) Inverse problems.

https://doi.org/10.1364/JOSAB.35.001271

1. INTRODUCTION

Phase retrieval (PR) aims to recover the object or image of interest from the diffraction pattern without phase. PR arises in various imaging fields, such as optical imaging [1], coherent diffractive imaging [2], and ultrafast non-linear optics [3], to name just a few. The PR problem is an old but challenging problem for its non-convex and non-linear property. In the past decades, various efforts of exploiting priors to address this problem were developed.

A longstanding of PR algorithms is the alternating projection (AP) method [4,5]. This class of PR algorithms often starts from an initial estimation, and projects the estimation onto two constraint sets (the measurement constraint set and the spatial domain constraint set) alternatively. The measurement

constraint set is defined based on the sampling model. For instance, the measurement constraint set is the Fourier magnitude constraint set in the far-field PR regime. The spatial domain constraint set can exploit the inherent prior information of the underlying image. Particularly, the seminal AP method Gerchberg and Saxton (GS) [4] exploits the magnitude constraint of the complex-valued image, and the popular alternating projection methods error reduction (ER) and hybrid input output (HIO) [5] exploit the support constraint. These three algorithms are simple, but the inherent priors of the underlying image are ignored in these algorithms. Recently, inspired by compressed sensing [6], a sparse constraint was incorporated into the spatial domain constraint set. Experimental simulations demonstrate that these sparsity-based AP methods

can improve the reconstruction performance compared to their original AP methods without utilizing sparsity [7,8]. AP methods mainly have an advantage that few finely tuned parameters are needed. However, this class of algorithms often suffers from low-quality imaging when the recorded diffraction patterns are few [9,10].

A more recent PR approach is to formulate a non-convex but smooth least-squares estimation problem. This type of algorithm often exploits the gradient descent method to solve the formulated problem, examples like Wirtinger flow (WF) [11] and truncated Wirtinger flow [12]. The strict theoretical guarantee of the convergence to stationary points leads these two algorithms to popular ones. Unfortunately, this class of algorithm also has a drawback of low-quality imaging in the case of few diffraction patterns [13,14]. The neglect of the priors about the underlying image should account for this deficiency.

To exploit the inherent prior information for PR, regularized PR algorithms were proposed [9,10,13–15]. The regularized PR algorithms often formulate a non-convex problem that combines a data fidelity term and an elaborate regularization term. The regularization term can enforce some desirable properties on the recovered image. The regularized PR algorithms that exploit proper regularization models can obtain better reconstruction accuracy, compared to the PR algorithms without regularization models [13,14]. Moreover, the data fidelity term can be designed based on the distribution type of the noise. Therefore, the regularized PR algorithms are often robust to noise. Since regularized PR algorithms can perform high-quality imaging in the case of few noisy diffraction patterns, they have become popular in recent years. Many finely tuned parameters (such as regular parameters or thresholding parameters) in regularized methods are needed. Moreover, the optimal parameters often vary with the sampling ratios and noise levels [13]. Therefore, fine-tuning of the parameters in regularized PR algorithms is a difficult, but not a trivial task.

Keeping the advantages and disadvantages of these two classes of PR algorithms (AP methods and regularized methods) in mind, we propose a unified framework called constrained phase retrieval (ConPR). The ConPR framework not only can exploit priors for diffraction imaging via an image denoising operator, but also does not need fine-tuning of the parameters. The AP methods usually solve a PR feasible problem that contains two constraint sets. From the optimization perspective, this feasible problem can be rewritten as a bivariate optimization problem. Based on this fact, we develop a generalized ConPR model, and it is a bivariate optimization model. We should mention a similar work by Pauwels *et al.* in a preprinted literature [16]. Pauwels *et al.* [16] showed that AP methods could lead to smoothly converging sequence of estimates theoretically when used with Fourier transform and convex priors. On the practical side, they derived a fast gradient descent method based on the l_1 regularization. The main differences between their work and our work are twofold, as follows. (1) We consider the generalized PR problem, but the authors in [16] only consider the Fourier PR problem; and (2) our generalized ConPR framework can exploit any proper regularization model, but the authors in [16] only exploit the l_1 regularization.

The proposed generalized ConPR model combines a relation function term, a regularization term, and a measurement constraint. The relation function is utilized to promote the approximation of the two variables, i.e., the estimated image and the auxiliary variable. The regularization term enforces some desirable properties on the recovered image, while the measurement constraint promotes the data fidelity. To present a concrete PR algorithm, we exploit the sparsity under a certain frame for diffraction imaging. Motivated from the block-matching and 3D filtering (BM3D) algorithm [17,18], an advanced image denoising algorithm in the last decade, we select the BM3D frame as the representation frame. The proposed BM3D-based problem formulation consists of two sub-problems, i.e., the image updating sub-problem and the auxiliary variable updating sub-problem (constrained optimization sub-problem). The image updating sub-problem can be solved by the BM3D image denoising algorithm. To solve the constrained optimization sub-problem effectively, we define the epigraph set of the data fidelity function and translate the sub-problem into a projection problem. By doing so, the regular parameter is unnecessary. In fact, our algorithm consists of two steps: one is the BM3D image denoising step, and the other one is the projection step. For BM3D, the default parameters except for the noise standard deviation are utilized in our algorithm, and the noise standard deviation is estimated from the recovered image at each iteration.

The advantages of our proposed ConPR framework are summarized as follows:

- **Flexibility:** our proposed framework is flexible, and it is twofold, as follows: *i.* plug-and-play: any effective image denoising operator can be incorporated into this model via adjusting the regularization term. In other words, various prior information can be utilized in ConPR for diffraction imaging; *ii.* feasible for the generalized PR problem: any non-linear sampling operator whose gradient is computable can be incorporated into the proposed generalized PR framework.
- **Effectiveness:** coded diffraction imaging of utilizing the ConPR algorithm with BM3D is simulated. With one noisy diffraction pattern, the proposed ConPR algorithm can produce both higher imaging quality and faster imaging speed, compared with the existing sparsity-based PR algorithms.

In what follows, in Section 2, we review the AP methods and the regularized PR algorithms. In Section 3, we introduce the proposed constrained PR framework. In Section 4, we present our experimental results. Finally, concluding remarks and directions for future research are discussed in Section 5.

2. AP METHODS VERSUS REGULARIZED PR METHODS

A light illuminates an image or object of interest $\mathbf{x} \in \mathbb{R}^N$ (we consider the recovery problem of real images in this paper). Through the propagation of the light in the space, the information of the arrived light is recorded by sensors. Due to the limitations of modern sensors, only the intensity or magnitude of the complex-valued light field can be recorded. Therefore, the sampling process is non-linear. Mathematically, the sampling model can be described as

$$\mathbf{y} = G[\Phi(\mathbf{x})], \tag{1}$$

where $\Phi(\bullet)$ is a non-linear operator. The non-linear sampling operator $\Phi(\bullet)$ describes the theoretical sampling process. In particular, the non-linear sampling operator $\Phi(\mathbf{x}) = |\mathbf{Fx}|$ (here, $\mathbf{F}$ represents the Fourier transform matrix, and $|\bullet|$ is the magnitude operator) for the far-field scenario, and $\Phi(\mathbf{x}) = |\mathbf{F}(\mathbf{I} \odot \mathbf{x})|^2$ (here, $\mathbf{I}$ represents the random mask, and $\odot$ is the element-wise product) for the coded diffraction pattern (CDP) model [19]. $\mathbf{y}$ is the non-linear measurement, and $G(\bullet)$ represents an operator that adds some noise onto the clear measurement. Recovering the image $\mathbf{x}$ from the measurement $\mathbf{y}$ is the goal of the PR algorithm. Since the non-linear sampling operator exists in the PR problem, this problem is ill-posed. To address this challenge, researchers usually oversample the signal or exploit the priors for phase retrieval [7–10,13–15].

A. AP Methods

The AP method of solving the PR problem usually formulates a feasible problem:

$$\text{find} \quad \mathbf{x} \in M \cap S, \tag{2}$$

where M and S represent the measurement constraint set and the spatial domain constraint set, respectively. For a general PR problem, the measurement constraint set can be written as $\mathrm{M} = \{\mathbf{x} \in \mathbb{R}^N | \|\mathbf{y} - \Phi(\mathbf{x})\|_2^2 \leq \epsilon\}$ (here, $\epsilon > 0$, and we assume that Gaussian noise exists in the measurement device). The set M can promote the data fidelity. The spatial domain constraint set S can be designed based on the priors of the underlying image. The most commonly utilized priors are support, smoothness, and sparsity [5,7,8]. Given an initialization for the specific PR feasible problem, the AP method optimizes this problem by projecting the iterated estimation onto these constraint sets alternatively. Define the indicator function $\mathbb{I}_C(\mathbf{x}) = 0$ if $\mathbf{x} \in \mathrm{C}$; $\mathbb{I}_C(\mathbf{x}) = +\infty$ otherwise. Here, C represents a certain constraint set. Equation (2) can be rewritten as a non-convex optimization problem:

$$\min_{\mathbf{x}} \mathbb{I}_M(\mathbf{x}) + \mathbb{I}_S(\mathbf{x}). \tag{3}$$

In fact, the AP method, such as the GS method or ER method, solves the above problem via incorporating an auxiliary variable, namely,

$$\min_{\mathbf{x}, \mathbf{z}} \mathbb{I}_M(\mathbf{z}) + \mathbb{I}_S(\mathbf{x}) + \|\mathbf{x} - \mathbf{z}\|_2^2, \tag{4}$$

where $\mathbf{z}$ is the auxiliary variable. Essentially, the GS method or ER method solves the above problem via alternating minimization, namely, updating one variable while keeping the other one fixed. To solve Eq. (4), alternating minimization for each variable is essentially projecting the estimation onto the two sets alternatively. For Eq. (4) itself, there is not a parameter, which is the main advantage of the GS or ER method.

B. PR-Based Regularization Methods

AP methods often suffer from low-quality imaging in the case of noisy and few measurements [9,10,14]. Regularized PR algorithms have attacked this deficiency. They often combine a data fidelity term and a regularization term to formulate an optimization problem:

$$\min_{\mathbf{x}} \|\mathbf{y} - \Phi(\mathbf{x})\|_2^2 + \lambda R(\mathbf{x}), \tag{5}$$

where λ is the regular parameter. Here, the first term is called the data fidelity term, and the second term is the regularization term. The data fidelity term ensures that the recovered image agrees with the measurement, whereas the regularization term enforces some desirable properties onto the unknown image.

Recent trends [9,10,13–15,20,21] are concentrated on exploiting an elaborate regularization term for PR, and developing high-quality imaging algorithms in the case of noisy measurements. For the far-field PR field, l_1 regularization term [20], sparse representation regularization under the translation invariant Haar pyramid tight frame [9], and the regularization term based on transfer orthogonal sparsifying transform learning [10] were proposed for Fourier phase retrieval. For a CDP model, the sparse representation model based on synthesis dictionary learning [13], the sparse models under the BM3D frames [14], the total variation regularization term [15], and the regularization model based on orthogonal dictionary learning [21] were proposed for coded diffraction imaging. Although the regularized PR algorithms have performed better reconstruction performance compared to AP methods, they have a drawback that fine-tuning of many parameters is needed. In the literature [13], the authors discussed the importance of the regularization parameter to the final reconstruction result, and showed that this parameter varied with sampling ratios and noise levels. To avoid this difficult process of tuning the regular parameter, we propose a novel constrained PR framework. The proposed framework is simple as the AP framework, and it can be fused with any regularization term that characterizes the proper prior of the underlying image.

3. CONSTRAINED PHASE RETRIEVAL

A. Generalized ConPR Framework

For the optimization model [Eq. (3)], the prior, such as sparsity, is utilized via the spatial domain constraint set S, while for the optimization model [Eq. (5)], the prior is utilized via the defined regularization term $R(\mathbf{x})$. We attempt to construct a unified framework that fuses the regularization term and the measurement constraint. Since the AP method can be regarded as solving Eq. (4) via the alternating minimization method, one can incorporate the regularization term into Eq. (4):

$$\min_{\mathbf{x}, \mathbf{z}} \mathbb{I}_M(\mathbf{z}) + \|\mathbf{x} - \mathbf{z}\|_2^2 + \lambda R(\mathbf{x}). \tag{6}$$

We rewrite the above optimization model as a constrained optimization model, and propose the generalized ConPR optimization model as follows:

$$\min_{\mathbf{x}, \mathbf{z}} \|\mathbf{x} - \mathbf{z}\|_2^2 + \lambda R(\mathbf{x}) \quad \text{s.t.} \ f(\mathbf{z}) \leq \epsilon. \tag{7}$$

Here, f represents the data fidelity function, which is defined based on the noise distribution type. In this paper, we consider that Gaussian noise exists in the optical measurement device. Therefore, the data fidelity function is defined as $f(\mathbf{z}) = \|\mathbf{y} - \Phi(\mathbf{z})\|_2^2$, which promotes the robustness of the algorithm to Gaussian noise. The first term in Eq. (7) is a smooth function that characterizes the relationship between the underlying image and the auxiliary variable. We denote this function as the

relation function. The second term is the regularization term that can incorporate the priors into the image recovery model. The constraint guarantees that the recovered auxiliary variable matches with the measurement. Through the relation function, the recovered image can be constrained by the measured data. The above optimization model is a generalized model, which is suitable for any non-linear sampling operator. Moreover, any effective regularization term can be utilized in this model theoretically. The proposed ConPR model [Eq. (7)] is suitable for any effective regularization term and any non-linear sampling operator; therefore, we term the model [Eq. (7)] as the generalized ConPR model. Indeed, Eq. (7) can be solved via dividing it into the following two sub-problems (to the tth iteration):

$$\mathbf{x}^{(t)} = \arg \min_{\mathbf{x}} \{\|\mathbf{x} - \mathbf{z}^{(t-1)}\|_2^2 + \lambda R(\mathbf{x})\}, \qquad \textbf{(8)}$$

$$\mathbf{z}^{(t)} = \arg \min_{\mathbf{z}} \{\|\mathbf{x}^{(t)} - \mathbf{z}\|_2^2\} \quad \text{s.t.} \ f(\mathbf{z}) \le \epsilon. \qquad \textbf{(9)}$$

Equation (8) is essentially an image denoising problem whose regularization term can be written explicitly. Theoretically, any proper image denoising algorithm can be utilized to solve Eq. (8) with a certain regularization term $R(\mathbf{x})$. In general, the image denoising algorithms with their original parameters can produce good denoising performance. Therefore, Eq. (8) can be solved by using an effective denoising algorithm with default parameters. Equation (9) is a constrained optimization problem, and we solve it by using the epigraph concept [22].

B. ConPR Problem Formulation Based on BM3D Sparse Model

The BM3D algorithm is an advanced image denoising algorithm in the last decade. Motivated from the successful denoising performance of the BM3D algorithm, we exploit the sparsity under the BM3D frame for the proposed generalized ConPR model. The core idea of BM3D is exploiting the non-local similarity and the 3D transform sparsity for image denoising. Based on the frame interpretation of the BM3D algorithm [18], the algorithm can be split into three steps: *analysis*, *processing*, and *synthesis*. In the *analysis* step, the image is partitioned into small overlapping patches, and then these patches are grouped to form the 3D groups. The formed groups are decorrelated by using an invertible 3D transform. The 3D transform coefficients are called 3D group spectra. Mathematically, the *analysis* process can be modeled as

$$\boldsymbol{\alpha} = \Psi_{\text{BM3D}}\mathbf{x}. \qquad \textbf{(10)}$$

Here, $\boldsymbol{\alpha}$ is the 3D group spectra. The analysis operation Ψ_{BM3D} models the group and the 3D transform process. The *processing* step is filtering the obtained 3D group spectra by using the hard thresholding operator: $\hat{\boldsymbol{\alpha}} = T_\tau(\Psi_{\text{BM3D}}\mathbf{x})$ (here, $T_\tau(\bullet) = \max(\bullet, \tau)$ represents the hard thresholding operator, and τ is the thresholding value). In the *synthesis* step, the estimates for each group are provided via inverting the filtered spectra. These patch-wise estimates are aggregated to construct the final image reconstruction. Mathematically, the synthesis sparse model of utilizing the BM3D frame can be described as [18]

$$\hat{\mathbf{x}} = \Phi_{\text{BM3D}}\hat{\boldsymbol{\alpha}}, \qquad \textbf{(11)}$$

where Φ_{BM3D} is the synthesis operation. The explicit representations of the BM3D synthesis operation Φ_{BM3D} and the

analysis operation Ψ_{BM3D} are defined in [18]. With the synthesis operation and the analysis operation, the BM3D image denoising process can be described as $\mathbf{x} = \Phi_{\text{BM3D}}T_\tau(\Psi_{\text{BM3D}}\mathbf{x})$. The frame interpretation of BM3D [18] is beneficial for constructing BM3D-based image recovery optimization models.

Our problem formulation based on the BM3D frame in the case of Gaussian noise is as follows:

$$\min_{\mathbf{x}, \mathbf{z}} \|\mathbf{x} - \mathbf{z}\|_2^2 + \lambda \|\Psi_{\text{BM3D}}\mathbf{x}\|_0 \quad \text{s.t.} \ \|\mathbf{y} - \Phi(\mathbf{z})\|_2^2 \le \epsilon. \qquad \textbf{(12)}$$

Since the non-linear sampling operator exists in the constraint of Eq. (12), it is a non-convex optimization problem. Solving this challenging problem effectively is also one of our main contributions, and the next sub-section will introduce the effective numerical method that has few finely tuned parameters to solve this problem. It should be noted that finely tuned parameters in other image denoising algorithms utilized in the proposed framework may be needed. Nevertheless, for BM3D, the default parameters in the BM3D image denoising algorithm of the proposed PR algorithm are suitable for most noise levels. This is one of the reasons why finely tuned parameters are not needed in the proposed PR algorithm of utilizing BM3D.

C. BM3D-Based ConPR Algorithm without Extra Hand-Tuned Parameter

We solve Eq. (12) by dividing it into the following two subproblems (to the tth iteration):

$$\mathbf{x}^{(t)} = \arg \min_{\mathbf{x}} \{\|\mathbf{x} - \mathbf{z}^{(t-1)}\|_2^2 + \lambda \|\Psi_{\text{BM3D}}\mathbf{x}\|_0\}, \qquad \textbf{(13)}$$

$$\mathbf{z}^{(t)} = \arg \min_{\mathbf{z}} \{\|\mathbf{x}^{(t)} - \mathbf{z}\|_2^2\} \quad \text{s.t.} \ \|\mathbf{y} - \Phi(\mathbf{z})\|_2^2 \le \epsilon. \qquad \textbf{(14)}$$

Through solving the above two sub-problems alternatively, the final recovered image can be obtained. The above two subproblems are called the image $\mathbf{x}$ updating sub-problem and the auxiliary variable $\mathbf{z}$ updating sub-problem, respectively.

1. x Sub-Problem

Equation (13) can be considered as a denoising problem with a regularization term defined over the BM3D analysis frame Ψ_{BM3D}. In our algorithm, we attempt to use the following equation as an approximate solution for Eq. (13):

$$\mathbf{x}^{(t)} = \Phi_{\text{BM3D}}T_\tau(\Psi_{\text{BM3D}}\mathbf{z}^{(t-1)}). \qquad \textbf{(15)}$$

The above equation is the BM3D image denoising process where the auxiliary variable $\mathbf{z}^{(t-1)}$ is the noise image and $\mathbf{x}^{(t)}$ is its denoised version [18]. For BM3D image denoising, the thresholding value τ is often set as $\tau = C\sigma$, where C is a constant and σ is the noise standard deviation. For realistic applications of our algorithm, the default parameters (such as the constant C, patch size, group size, and so on) in the original BM3D algorithm are utilized. The input noise standard deviation for BM3D is evaluated based on the noise image $\mathbf{z}^{(t-1)}$ via the robust median operator [23]. Therefore, keeping the default parameters of the original BM3D image denoising method fixed, the image updating process has not a finely tuned parameter.

2. z Sub-Problem

Equation (14) is a constrained optimization problem where the constraint is the measurement constraint. With $\mathbf{x}^{(t)}$ fixed, this problem can be regarded as a projection problem, namely, projecting the estimated image $\mathbf{x}^{(t)}$ onto the measurement constraint set $\{\mathbf{z} \in \mathbb{R}^N \| \|\mathbf{y} - \Phi(\mathbf{z})\|_2^2 \le \epsilon\}$. For the far-field regime, we can set the magnitude of the Fourier spectrum of the denoised image $\mathbf{x}^{(t)}$ to the recorded noiseless measurement, and inverse the updated spectrum to obtain the projected solution. In this work, we consider the general PR problem. Under this scenario, the sampling operator is a general non-linear sampling operator.

Different from the regularized PR problem [Eq. (5)], the regularization parameter in the constrained problem [Eq. (14)] is unnecessary. However, fine-tuning of the parameter ϵ that is an upper bound on the error in the measurement domain is needed. Moreover, a fixed ϵ is not always optimal for each noise level or each sampling ratio. Therefore, fine-tuning of the parameter ϵ is a difficult, but not a trivial task. To avoid the tedious tuning process, we solve Eq. (14) via the epigraph concept and select this parameter adaptively. We extend the epigraph set concept of convex cost functions in [22] to the non-convex cost functions. We attempt to solve Eq. (14) via projecting an estimated point onto the defined epigraph set. Through the successive projections, the value of the error function $f(\mathbf{z})$ is decreased. We plan to use the final projected solution as the approximated solution of Eq. (14).

Specifically, we define an epigraph set of the function $f(\mathbf{z}) = \|\mathbf{y} - \Phi(\mathbf{z})\|_2^2$, and select the parameter ϵ based on the estimated auxiliary variable $\mathbf{z}^{(t)}$ (for the tth iteration). Define the epigraph set of the error function $f(\mathbf{z})$ via increasing the dimension by one:

$$C_{\text{error}} = \{\underline{\mathbf{z}} = [\mathbf{z}^T q]^T : q \ge f(\mathbf{z})\}. \tag{16}$$

Here, q is a positive constant, which is determined based on the estimation $\mathbf{z}$. The set C_{error} is the set of column vectors with $N + 1$ dimensional whose $(N + 1)$th components are greater than $f(\mathbf{z})$. Keeping the epigraph set concept in mind, Eq. (14) can be solved via attacking the following projection equation:

$$\underline{\mathbf{z}}^{(t)} = \text{proj}(\underline{\mathbf{u}}^{(t)}) = \arg \min_{\underline{\mathbf{z}} \in C_{\text{error}}}\{\|\underline{\mathbf{u}}^{(t)} - \underline{\mathbf{z}}\|_2^2\}, \tag{17}$$

where $\underline{\mathbf{u}}^{(t)} = [\mathbf{x}^{(t)T} 0]^T$. Equation (17) is essentially the orthogonal projection of $\underline{\mathbf{u}}^{(t)}$ onto the set C_{error}. In other words, the goal of the above optimization model is to search for the nearest vector $\underline{\mathbf{z}}$ to $\underline{\mathbf{u}}^{(t)}$. Through the orthogonal projection onto the supporting hyperplanes of the epigraph set C_{error} iteratively, we can obtain the solution to Eq. (17). To present a concrete PR algorithm, in this work, we consider the coded diffraction imaging problem, and focus on recovering the image of interest from one coded diffraction pattern corrupted with Gaussian noise. Under this scenario, the non-linear sampling operator is $\Phi(\mathbf{x}) = |\mathbf{Ax}|^2$, where $\mathbf{A}$ describes the propagation process of the illuminate light from the object plane to the sensor plane [19]. In the case of one diffraction pattern, the matrix $\mathbf{A}$ and its conjugate transpose matrix are defined as [13]

$$\mathbf{Ax} = \mathbf{F}(\mathbf{I} \odot \mathbf{x}), \qquad \mathbf{A}^H \mathbf{c} = \frac{1}{M} \bar{\mathbf{I}} \odot (\mathbf{F}^H \mathbf{c}). \tag{18}$$

Here, $\mathbf{I}$ is the illumination mask, and $\bar{\mathbf{I}}$ represents its dual filter. M is the number of the measurements, and c is a vector with size M. $\mathbf{F}^H$ is the conjugate transpose matrix of the Fourier transform matrix. $\odot$ represents the element-wise product. For the CDP model, the error function is $f(\mathbf{z}) = \|\mathbf{y} - |\mathbf{Az}|^2\|_2^2$. To solve Eq. (17), we project $\underline{\mathbf{u}}^{(t)}$ onto the supporting hyperplanes of the epigraph set C_{error} successively. For the jth projection, the projected point on the supporting hyperplane is [22]

$$[\mathbf{v}_j^T f(\mathbf{v}_j)]^T, \qquad \mathbf{v}_j = \mathbf{v}_{j-1} - \frac{f(\mathbf{v}_{j-1})}{\|\nabla f(\mathbf{v}_{j-1})\|_2^2 + 1} \nabla f(\mathbf{v}_{j-1}). \tag{19}$$

The initial value is $\mathbf{v}_0 = \mathbf{u}^{(t)}$. The gradient $\nabla f(\mathbf{v}_{j-1})$ in Eq. (19) is defined as

$$\nabla f(\mathbf{v}_{j-1}) = \text{real}\{\mathbf{A}^H[\mathbf{Av}_{j-1} \odot (|\mathbf{Av}_{j-1}|^2 - \mathbf{y})]\}, \tag{20}$$

where real($\bullet$) is the operator that extracts the real part of a complex-valued vector element-wise. For the total J times successive projections, the final orthogonal projection onto the supporting hyperplane of the epigraph set can be represented as $[\mathbf{v}_J^T f(\mathbf{v}_J)]^T$. Therefore, an approximate solution of Eq. (17) is achieved:

$$\text{proj}(\underline{\mathbf{u}}^{(t)}) = [\mathbf{v}_J^T f(\mathbf{v}_J)]^T, \tag{21}$$

where $\mathbf{v}_J = \mathbf{v}_{J-1} - \frac{f(\mathbf{v}_{J-1})}{\|\nabla f(\mathbf{v}_{J-1})\|_2^2 + 1} \nabla f(\mathbf{v}_{J-1})$. Although the Jth projected point is an approximated solution to Eq. (17), the effectiveness of this method is validated in the experimental simulation section. Obtain the projection of $\mathbf{u}^{(t)}$ onto the supporting hyperplane of the epigraph set C_{error}, then the solution to Eq. (14) is $\mathbf{z}^{(t)} = \mathbf{v}_J$. Essentially, Eq. (14) can be regarded as a projection problem. Based on the aforementioned epigraph set of the data fidelity function, we can solve Eq. (14) via attacking Eq. (17). We solve Eq. (17) by using the orthogonal projection onto the epigraph set, and let $\underline{\mathbf{u}}^{(t)} = [\mathbf{x}^{(t)T} 0]^T$, $\underline{\mathbf{z}}^{(t)} = [\mathbf{z}^{(t)T} f(\mathbf{z}^t)]^T$. By doing this, we can obtain the solution to Eq. (17) as well as Eq. (14).

So far, all issues to solve the $\mathbf{x}$ sub-problem and the $\mathbf{z}$ sub-problem are attacked. The method of solving these two sub-problems consists of two steps, i.e., the image denoising step and projection step. We can obtain a satisfied solution to the proposed formulation Eq. (12) via performing these two steps iteratively. We summarize the proposed ConPR algorithm of utilizing the BM3D frame in algorithm 1.

Algorithm 1: Constrained Phase Retrieval (ConPR) Algorithm

1: **Input:** non-linear measurement $\mathbf{y}$.
2: **Initialization:** initial image $\mathbf{z}^{(0)} = \mathbf{x}^{(0)}$, maximum iteration t_{max}, J.
3: **Outer Loop:** for $t = 1, 2, \ldots, t_{max}$ do
4:   Obtain the denoised image $\mathbf{x}^{(t)}$ via Eq. (15). ▷ BM3D image denoising step
5:   Set the initial point $\mathbf{v}_0 = \mathbf{x}^{(t)}$.
6:   **Inner Loop:** for $j = 1, 2, \ldots, J$ do
7:     Compute the gradient at $\mathbf{v}_{j-1}$ via Eq. (20).
8:     Obtain the projected point via Eq. (19). ▷ Projection step
9:   **End Inner Loop**

10:    Let $\mathbf{z}^{(t)} = \mathbf{v}_J$, and output the recovered image if $t = t_{max}$ or $\frac{\|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|_2^2}{\|\mathbf{x}^{(t)}\|_2^2} \leq 10^{-4}$.

11:    **End Outer Loop**

12:    **Output:** image $\mathbf{x}^{(t)}$.

4. EXPERIMENTAL SIMULATIONS

In this section, we discuss various numerical experiments in the case of few measurements to study the effectiveness of the proposed ConPR algorithm. To that end, we consider recovering the real image of interest from one noisy coded diffraction pattern at several noise levels. The details of the experimental setup and our concrete implementation are given in the first sub-section. The next sub-section will present the results achieved by the proposed ConPR algorithm and the existing PR algorithms. The final sub-section presents the convergence behavior of the ConPR algorithm and the comparison of the running time.

A. Experimental Setup and Our Concrete Implementation

We consider sampling the testing images via the CDP model. The size of the testing image is 512×512, and each image is scaled to [0,1]. The eight original gray-scale images are presented in Fig. 1. The first six natural images can be downloaded from http://www.cs.tut.fi/~foi/GCF-BM3D/. The original cell images can be downloaded from http://www.cellimagelibrary. org/pages/project_20269, and we translate the two cell images to the gray-scale versions with size 512×512.

An illumination mask is placed behind the image of interest in the coded diffraction pattern realistic setup, and the sensor records the coded diffraction pattern. Concretely, we use the quaternary illumination mask, namely, each element of the illumination mask is selected randomly from the set $\{\pm 1, \pm i\}$. Mathematically, we sample the image via the following model:

$$\mathbf{y} = |\mathbf{F}(\mathbf{I} \odot \mathbf{x})|^2 + \mathbf{n}. \tag{22}$$

Here, $\mathbf{n}$ represents the Gaussian noise vector, and the noise level is controlled by signal-to-noise ratio (SNR) defined as [14]

$$\text{SNR} = 10 \log_{10}[\|\bar{\mathbf{y}}\|_2^2 / \|\bar{\mathbf{y}} - \mathbf{y}\|_2^2]\text{dB}, \tag{23}$$

where $\bar{\mathbf{y}} = |\mathbf{F}(\mathbf{I} \odot \mathbf{x})|^2$ ($\mathbf{x}$ is the original image) represents the intensity of the theoretical diffraction pattern.

Four recent publicly released PR algorithms are selected as the benchmark algorithms to demonstrate the effectiveness of

Fig. 1. Original images. Top row (left to right): Lena, Barbara, hill, boat; bottom row (left to right): couple, fingerprint, acinar cell, chromaffin cell.

the proposed ConPR algorithm. They are WF [11], DOLPHIn (DictiOnary Learning for PHase retrIeval) [13], BM3D-prGAMP (BM3D denoiser within the generalized approximate message passing framework for phase retrieval) [24], and sparse phase amplitude reconstruction (SPAR) [14]. The codes of WF and DOLPHIn can be found in the DOLPHIn package (https://www.graphics.rwth-aachen.de/media/resource_files/ DOLPHIn.zip). Moreover, we use the BM3D-prGAMP algorithm from the D-AMP_Toolbox package (https://github.com/ ricedsp/D-AMP_Toolbox/tree/master/Algorithms). The code of the SPAR algorithm can be downloaded from the SPARSE project homepage (http://www.cs.tut.fi/sgn/imaging/sparse/). The concrete implementations of the proposed ConPR algorithm and the four benchmark PR algorithms are as follows.

The WF algorithm only exploits the non-linear measurements for diffraction imaging, and its maximum iteration is set as 100. The DOLPHIn algorithm exploits the sparsity under an adaptive synthesis dictionary for diffraction imaging, and optimizes the dictionary and the image simultaneously from the non-linear measurements. The default parameters of the original DOLPHIn algorithm were utilized. The BM3D-prGAMP algorithm incorporates the BM3D image denoising method into the iteration of the GAMP (generalized approximate message passing) algorithm. The SPAR algorithm exploits the BM3D frame to design the sparse model, and a variational approach is utilized to formulate the optimization model. The maximum iteration for BM3D-prGAMP is 70 and for SPAR is 50. For the proposed ConPR algorithm, we set the maximum iteration of the outer loop $t_{\max} = 20$, and the inner loop $J = 2$. For the BM3D image denoising process in our algorithm, the default parameters are utilized. Essentially, only the maximum iteration is needed to be specified in ConPR, and finely tuned parameters are not needed.

All the tests are performed on a desktop computer with a Core i7-7700 CPU, 3.6 GHz, with 8 GB of RAM, and running the Windows 10 64-bit operating system and MATLAB 2017a software (MathWorks, USA). The code of the proposed ConPR algorithm can be downloaded from https://raw.githubusercontent.com/shibaoshun/Sparsity-based-Phase-Retrieval/phase-retrieval/ConPR.zip.
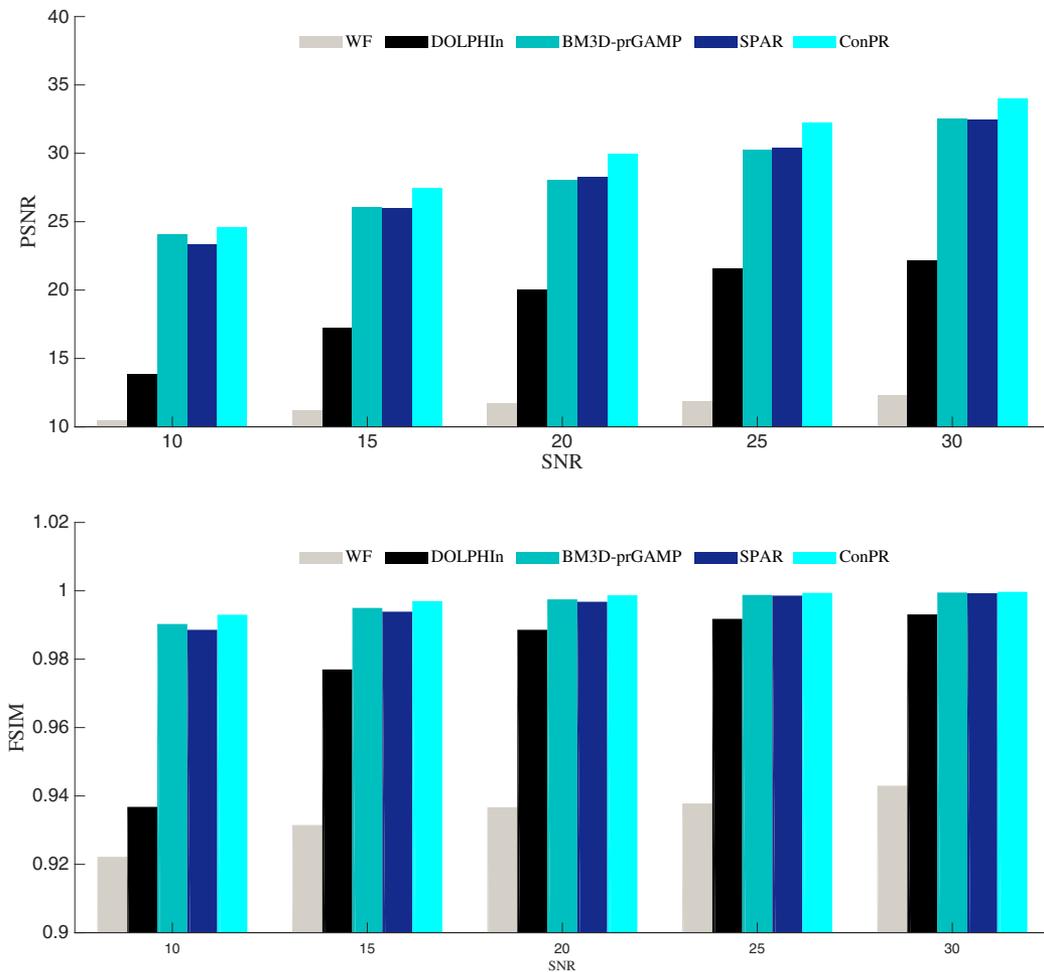
B. Comparison with the Previous PR Algorithms

Diffraction imaging from one noisy diffraction pattern was simulated, and we conducted the four previous PR algorithms and the proposed ConPR algorithm at various noise levels. All PR algorithms start from the same random initial guess. To assess the quality of the recovered image, we consider two measure metrics: the peak signal-to-noise ratio (PSNR) of a reconstruction and its feature similarity (FSIM) value [25]. The PSNR value of a reconstruction $\mathbf{w}$ with size 512×512 is defined as

$$\text{PSNR} = 20 \log_{10}[512 / \|\mathbf{w} - \mathbf{x}\|_2]\text{dB}, \tag{24}$$

where $\mathbf{x}$ represents the original image. For PSNR, larger values are better, and a FSIM value (ranging from 0 to 1) closer to 1 indicates better visual quality.

Figure 2 presents the recovered PSNR and FSIM values achieved by the five PR algorithms in the case of one coded diffraction pattern. From Fig. 2, one can see that the proposed

Fig. 2. Comparison with the previous algorithms in the case of one diffraction pattern. Top: the plot of PSNR versus SNR; bottom: the plot of FSIM versus SNR. In each case, results are the average PSNR or FSIM values of the eight testing images.

ConPR algorithm can provide the highest average PSNR and FSIM value at each noise level. As can be seen from the figure, the WF algorithm without any prior is the worst, and the DOLPHIn algorithm of exploiting sparsity is better than the WF algorithm. These two algorithms are extremely worse than the other BM3D-based PR algorithms. The BM3D-prGAMP algorithm and the SPAR algorithm can obtain relatively good reconstructions, but still worse than the proposed ConPR algorithm. For the most part, the proposed ConPR algorithm can achieve the best reconstructions among the five algorithms. For the average PSNR and FSIM values, the proposed ConPR algorithm is the highest among the five algorithms.

To further demonstrate the effectiveness of the proposed ConPR algorithm, some recovered images by the five algorithms are presented in Figs. 3–5. For comparing clearly, Fig. 3 gives the parts of the recovered Lena images at SNR = 20 dB. As can be seen from Figs. 3–5, the reconstructions by the WF algorithm and the DOLPHIn algorithm are extremely worse, and many noise-like components exist in their reconstructions. The BM3D-prGAMP algorithm and the SPAR algorithm suppress these components and provide better reconstructions than the WF algorithm and the DOLPHIn

algorithm. However, from Fig. 3, the reconstruction by BM3D-prGAMP still contains some noise-like components, and much detail information is lost in their reconstructions. The SPAR algorithm is slightly better than the BM3D-prGAMP algorithm, but its reconstructions still lose some detail information. The proposed ConPR algorithm not only suppresses many noise-like components but also preserves much detail information, such as the texture information on the hat (see Fig. 3). Moreover, from Figs. 4 and 5, one can see that the proposed ConPR algorithm suppresses the noise-like components, and its reconstructions preserve both the edge information and the texture information. Therefore, from the visual perspective view, the proposed ConPR algorithm can produce the best reconstructions among the five PR algorithms.

C. Convergence Behavior and Running Time

The proposed ConPR algorithm solves Eq. (12) via alternating between updating the image $\mathbf{x}$ and updating the auxiliary variable $\mathbf{z}$. Indeed, the image $\mathbf{x}$ is updated via the BM3D image denoising algorithm. The auxiliary variable $\mathbf{z}$ is updated via solving a constrained non-convex optimization problem. We solve this problem based on the epigraph concept, and give

Fig. 3. Lena images recovered by the five PR algorithms at SNR = 20 dB. For comparing clearly, the parts of the images are presented. From left to right and top to bottom: the original image, the image recovered by WF (PSNR = 10.93 dB, FSIM = 0.8967), DOLPHIn (PSNR = 21.36 dB, FSIM = 0.9819), BM3D-prGAMP (PSNR = 31.94 dB, FSIM = 0.9972), SPAR (PSNR = 31.82 dB, FSIM = 0.9966), and ConPR (PSNR = 33.51 dB, FSIM = 0.9985).

Fig. 4. Hill images recovered by the five PR algorithms at SNR = 25 dB. From left to right and top to bottom: the original image, the image recovered by WF (PSNR = 10.94 dB, FSIM = 0.9129), DOLPHIn (PSNR = 24.04 dB, FSIM = 0.9921), BM3D-prGAMP (PSNR = 31.97 dB, FSIM = 0.9984), SPAR (PSNR = 31.89 dB, FSIM = 0.9980), and ConPR (PSNR = 33.73 dB, FSIM = 0.9992).
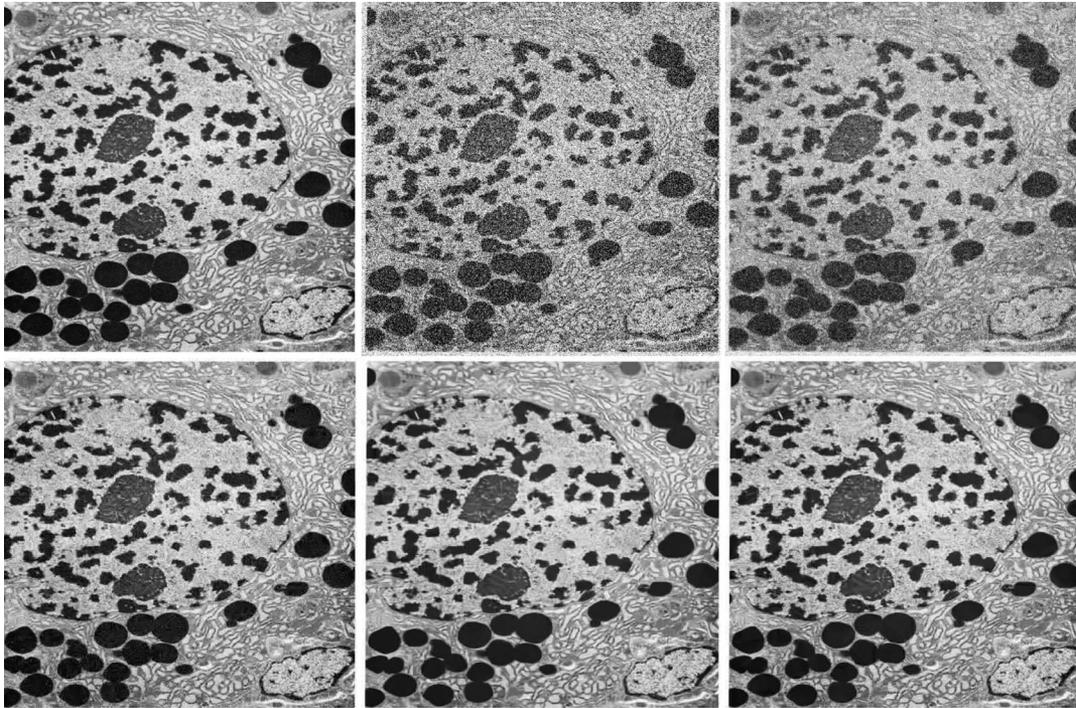
Fig. 5. Acinar cell images recovered by the five PR algorithms at SNR = 25 dB. From left to right and top to bottom: the original image, the image recovered by WF (PSNR = 13.16 dB, FSIM = 0.9891), DOLPHIn (PSNR = 17.19 dB, FSIM = 0.9960), BM3D-prGAMP (PSNR = 24.39 dB, FSIM = 0.9994), SPAR (PSNR = 25.80 dB, FSIM = 0.9995), and ConPR (PSNR = 27.36 dB, FSIM = 0.9997).

an approximated solution to the problem. Updating these two variables alternatively can obtain a satisfied solution to the formulated problem. However, the convergence is yet to be proved rigorously. Since the PR problem is non-convex, the strict convergence proof is difficult to provide. Empirically, PSNR converges to a stable point as the iteration increases. The curves of PSNR versus iteration for the testing images at SNR = 15 dB and SNR = 20 dB are presented in Fig. 6. For observing clearly, the convergence curve of the seven testing images

are presented. As can be seen from Fig. 6, with the growth of the iteration number, the proposed ConPR algorithm becomes stable through about 10 iterations. The fast convergence speed demonstrates good convergence behavior of the proposed ConPR algorithm. Nonetheless, one can see some slight perturbations on the curves for some images. The non-convexity of the optimization problem should account for this behavior.

To compare the imaging speed between the proposed ConPR algorithm and the four benchmark PR algorithms,
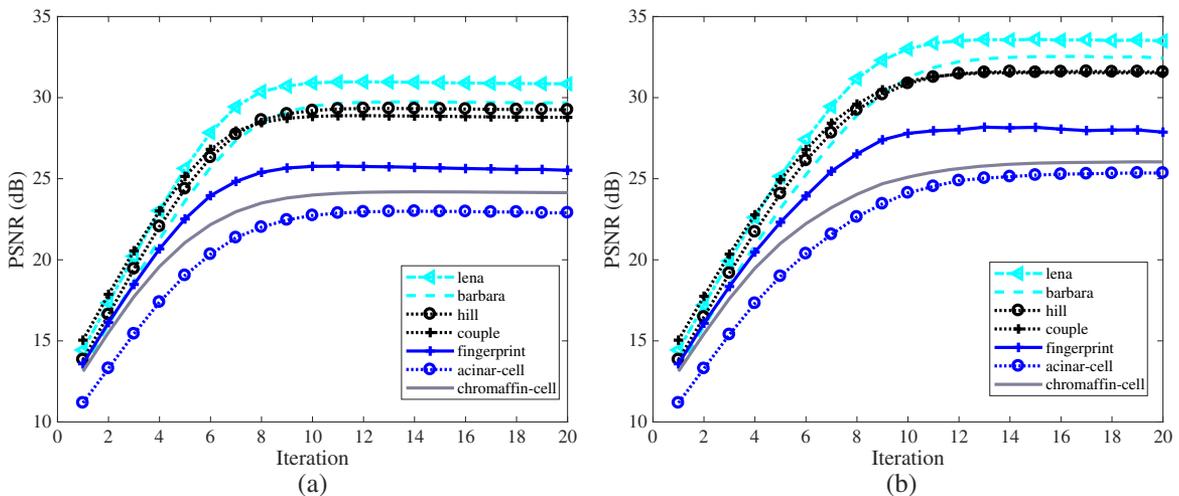
Fig. 6. Convergence curves of the proposed ConPR algorithm for recovering the testing images at (a) SNR = 15 dB and (b) SNR = 20 dB.

Table 1. Comparison of the Imaging Time (s) Consumed by the PR Algorithms[a]

WF	DOLPHIn	BM3D-prGAMP	SPAR	ConPR
7.40	16.24	188.85	130.95	13.28

[a]We report the average running time of processing the eight testing images. The first row is the testing algorithms, and the second row is the running time (s).

we give the average running time of recovering the eight testing images in Table 1. As can be seen from the table, the proposed ConPR algorithm is faster than the sparsity-based algorithms including DOLPHIn, BM3D-prGAMP, and SPAR, but slower than WF. The BM3D image denoising process should account for the slower imaging speed of the proposed algorithm, compared to the WF algorithm. However, the WF algorithm is extremely worse than the proposed algorithm in terms of PSNR and FSIM. Therefore, compared with the WF algorithm, the proposed ConPR algorithm is more suitable for applications. One can see from the table that the ConPR algorithm is nearly 1.2 times, 14.2 times, and 10 times faster than the DOLPHIn algorithm, the BM3D-prGAMP algorithm, and the SPAR algorithm, respectively. For both the reconstruction quality and the imaging speed, the proposed ConPR algorithm beats the three sparsity-based benchmark PR algorithms.

5. CONCLUSION

In this paper, we introduced a novel framework called ConPR. The ConPR framework is a unified framework that brings the AP method and the regularization approach together. We formulated the generalized ConPR optimization model that consists of a relation function term, a regularization term, and a constraint. The generalized ConPR optimization model is plug-and-play, and any image denoising method can be incorporated into this model. In particular, the BM3D frame was utilized for constructing the regularization term. We formulated the BM3D-based ConPR optimization model and solved it via updating the image and the auxiliary variable alternately. The two sub-problems were solved by using the BM3D image denoising method and the epigraph concept, respectively.

The proposed ConPR algorithm can exploit the non-local similarity and the 3D transform sparsity of the underlying image implicitly for PR. Coded diffraction imaging was simulated, and the experimental results showed that the ConPR algorithm could perform higher-quality imaging compared with previous PR algorithms. Moreover, our proposed algorithm is faster than the sparsity-based PR algorithms, such as DOLPHIn, BM3D-prGAMP, and SPAR. Most importantly, with the default parameters of the image denoising algorithm fixed, the finely tuned parameter in our algorithm is unnecessary, and it is suitable for realistic applications. Future work will focus on applying the deep learning method to our proposed ConPR framework. Moreover, applying the ConPR framework to other imaging fields is also our main research direction in the future.

REFERENCES

1. Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," IEEE Signal Process. Mag. **32**(3), 87–109 (2015).
2. B. Leshem, R. Xu, Y. Dallal, J. Miao, B. Nadler, D. Oron, N. Dudovich, and O. Raz, "Direct single-shot phase retrieval from the diffraction pattern of separated objects," Nat. Commun. **7**, 10820 (2016).
3. S. Birkholz, G. Steinmeyer, S. Koke, D. Gerth, S. Bürger, and B. Hofmann, "Phase retrieval via regularization in self-diffraction-based spectral interferometry," J. Opt. Soc. Am. B **32**, 983–992 (2015).
4. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," Optik **35**, 237–246 (1972).
5. J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," Opt. Lett. **3**, 27–29 (1978).
6. D. L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory **52**, 5406–5425 (2006).
7. S. Mukherjee and C. S. Seelamantula, "Fienup algorithm with sparsity constraints: application to frequency-domain optical-coherence tomography," IEEE Trans. Signal Process. **62**, 4659–4672 (2014).
8. S. Qin, X. Hu, and Q. Qin, "Compressed sensing phase retrieval with phase diversity," Opt. Commun. **310**, 193–198 (2014).
9. B. S. Shi, Q. S. Lian, and S. Z. Chen, "Sparse representation utilizing tight frame for phase retrieval," EURASIP J. Adv. Signal Process. **2015**, 1–11 (2015).
10. Q. S. Lian, B. S. Shi, and S. Z. Chen, "Transfer orthogonal sparsifying transform learning for phase retrieval," Digital Signal Process. **62**, 11–25 (2017).
11. E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: theory and algorithms," IEEE Trans. Inf. Theory **61**, 1985–2007 (2015).
12. Y. X. Chen and E. J. Candes, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," Commun. Pure Appl. Math. **70**, 822–883 (2017).
13. A. M. Tillmann, Y. C. Eldar, and J. M. Mairal, "DOLPHIn: dictionary learning for phase retrieval," IEEE Trans. Signal Process. **64**, 6485–6500 (2016).
14. V. Katkovnik, "Phase retrieval from noisy data based on sparse approximation of object phase and amplitude," arXiv: 1709.01071 (2017).
15. H. Chang, Y. Lou, M. K. Ng, and T. Zeng, "Phase retrieval from incomplete magnitude information via total variation regularization," SIAM J. Sci. Comput. **38**, A3672–A3695 (2016).
16. E. Pauwels, A. Beck, Y. C. Eldar, and S. Sabach, "On Fienup methods for regularized phase retrieval," arXiv: 1702.08339 (2017).
17. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Trans. Image Process. **16**, 2080–2095 (2007).
18. A. Danielyan, V. Katkovnik, and K. Egiazarian, "BM3D frames and variational image deblurring," IEEE Trans. Image Process. **21**, 1715–1728 (2012).

19. E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," Appl. Comput. Harmon. Anal. **39**, 277–299 (2015).
20. M. L. Moravec, J. K. Romberg, and R. G. Baraniuk, "Compressive phase retrieval," Proc. SPIE **6701**, 670120 (2007).
21. H. Chang and S. Marchesini, "Denoising Poisson phaseless measurements via orthogonal dictionary learning," arXiv: 1612.08656 (2016).
22. M. Tofighi, K. Kose, and A. E. Cetin, "Denoising using projections onto the epigraph set of convex cost functions," in *IEEE International Conference on Image Processing* (2015), pp. 2709–2713.

23. D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," Biometrika **81**, 425–455 (1994).
24. C. A. Metzler, A. Maleki, and R. G. Baraniuk, "BM3D-PRGAMP: compressive phase retrieval based on BM3D denoising," in *IEEE International Conference on Image Processing (ICIP)* (IEEE, 2016), pp. 2504–2508.
25. L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment," IEEE Trans. Image Process. **20**, 2378–2386 (2011).