TOTAL VARIATION BASED PHASE RETRIEVAL FOR POISSON NOISE REMOVAL

HUIBIN CHANG*, YIFEI LOU†, AND YUPING DUAN‡

Abstract. Phase retrieval plays an important role in vast industrial and scientific applications. We consider a phase retrieval problem in which the magnitudes of the Fourier transform (or a general linear transform) of a underling object are corrupted by Poisson noise, since any optical sensors detect photons, and the number of detected photons follows the Poisson distribution. We propose a variational model for phase retrieval based on a total variation regularization as an image prior and maximum likelihood estimation of Poisson noise model, which is referred to as "TV-PoiPR". We also propose an efficient numerical algorithm based on alternating direction method of multipliers (ADMM) and establish its convergence. Extensive experiments are conducted using both real and complex valued images to demonstrate the effectiveness of our proposed methods.

Key words. Phase Retrieval; Poisson Noise; Total Variation; Kullback-Leibler (KL) divergence; Alternating direction method of multipliers.

AMS subject classifications. 46N10, 49N30, 49N45, 65F22, 65N21

1. Introduction.

Phase retrieval (PR) plays a very important role in vast industrial and scientific applications, such as in astronomical imaging [20, 35], crystallography [27, 41], and optics [50, 48], *etc.* The goal is to reconstruct an object where pointwise magnitudes of the Fourier transform (FT) of the object are available. Since the phase of the FT is missing, this procedure is referred to as "phase retrieval".

Throughout the paper, we consider PR in a discrete setting, i.e., an underlying object $u : \Omega = \{0, 1, \cdots, n-1\} \to \mathbb{C}$ is of size n with $n = n_1 \times n_2$, in which we represent a 2-dimensional (2D) object with resolution $n_1 \times n_2$ in terms of a vector of size n by a lexicographical order. The measured data are magnitudes of the Fourier transform of u, *i.e.*, $|\mathcal{F}u|^2$, where $|\cdot|^2$ denotes the pointwise square of the absolute value of a vector, $\mathcal{F} : \mathbb{C}^n \to \mathbb{C}^n$ denotes the discrete Fourier transform (DFT)

$$(\mathcal{F}u)(\omega_1 + \omega_2 n_1) := \frac{1}{\sqrt{n_1 n_2}} \sum_{\substack{0 \le t_j \le n_j - 1, \\ j=1,2}} u(t_1 + t_2 n_1) \exp(-2\pi \mathbf{i}(\omega_1 t_1/n_1 + \omega_2 t_2/n_2)),$$

$\forall\, 0 \le \omega_j \le n_j - 1,$ for $j = 1, 2$, where $\mathbf{i} = \sqrt{-1}$. In fact, DFT can be replaced with an arbitrary linear operator, thus leading to a general phase retrieval problem [2, 49, 29],

$$\text{To find } \; u \in \mathbb{C}^n, \quad s.t. \;\; |\mathcal{A}u|^2 = b, \tag{1.1}$$

where $\mathcal{A} : \mathbb{C}^n \to \mathbb{C}^m$ is a linear operator in the complex Euclidean space and $b : \Lambda = \{0, 1, \cdots, m-1\} \to \mathbb{R}_+$. In general, phase retrieval involves solving a quadratic inverse problem, and it is ill-posed and yet challenging, since PR does not have a unique solution without additional information.

The computational tools for phase retrieval can be classified into three categories. One is based on alternative projection since a pioneer work of error reduction (ER) by

*School of Mathematical Sciences, Tianjin Normal University, Tianjin, 300387, China,E-mail: changhuibin@gmail.com

†Department of Mathematical Sciences, University of Texas at Dallas, Dallas, TX 75080, USA, E-mail: yifei.lou@utdallas.edu

‡Center for Applied Mathematics, Tianjin University, 300072, China,E-mail: yuping.duan@tju.edu.cn

Gerchberg and Saxton [24], and its variants [22, 3, 4, 21, 34]; please refer to [39, 53] and the reference therein. However, these methods are lack of theoretical guarantees due to nonconvex constraint sets for alternative projection. Recently, a global convergence for Gaussian measurements was theoretically analyzed by Netrapalli *et al.* [42], while the convergence under general conditions was proved by Marchesini *et al.* [40]. Chen and Fannjiang [16] provided local and geometric convergence to a unique fixed point for a Douglas-Rachford (DR) algorithm. In addition, gradient-type methods [17] become popular. For example, a Wirtinger flow (WF) approach was proposed in [10] by comprising of a careful initialization by spectral method and adaptive steps, and was further improved by truncated Wirtinger flow (TWF) [17]. Gradient-based approaches are often with first-order convergence, while a higher-order method was proposed by Qian *et al.* [44] to accelerate the convergence in a ptychographic PR problem. The third category is convex methods based on semi-definite programming(SDP). For example, Candés *et al.* proposed PhaseLift [11] that formulates a convex trace (nuclear) norm minimization by the lift technique of SDP. PhaseCut by Waldspurger *et al.* [49] convexified the PR problem by separating phases and magnitudes. PhaseLiftOff, a nonconvex variant of PhaseLift by subtracting off Frobenius norm from the trace norm, was proposed by Yin and Xin [55] to retrieve the phase with less measurements than PhaseLift. The SDP-based methods were adopted to solve a sparse phase retrieval problem [43, 33] in the sense that the reconstructed object is sparse.

In addition to computational tools, researchers also devote to analyzing phase retrieval theoretically by studying the injectivity (or uniqueness) of the quadratic operator $|\mathcal{A}(\cdot)|^2$. Denote a nonlinear mapping $\mathcal{M} : \mathbb{C}^n \to \mathbb{R}^m_+$ as

$$\mathcal{M}(u) = |\mathcal{A}u|^2. \tag{1.2}$$

In general, there exist trivial ambiguities [48], such as global phase shift, conjugate inversion, and spatial shift, so that the injectivity of $\mathcal{M}$ does not hold for Fourier type of measurements without additional information. If u is further assumed to be a $d-$dimensional real-valued vector, then $2^d \times n$ Fourier measurements are required to ensure the injectivity of $\mathcal{M}$ [28]. Note that one often refers u as a 1D signal for $d = 1$ and as a 2D image for $d = 2$. Since we focus on 2D image in this paper, this translates to $4n$ measurements for uniquely recovering real valued images. The number of measurements for unique recovery is relaxed for PR in general (not limited to Fourier measurements). For example, the injectivity is guaranteed by collecting $m \geq 2n - 1$ [2] and $m \geq 4n - 4$ [18] measurements for real $u \in \mathbb{R}^n$ and complex $u \in \mathbb{C}^n$ signals respectively, provided that the transform $\mathcal{A}$ is generated by a generic frame[1]. In particular, Shechtman *et al.* [48] showed that the lower bound $2n - 1$ can be achieved with high probability by collecting full-spark random measurements. Following the idea of holography, Candés *et al.* [8] proved unique phase retrieval from $3n$ Fourier measurements, in which the linear operator $\mathcal{A}$ can be expressed as

$$\mathcal{A}u = \begin{bmatrix} \mathcal{F}u \\ \mathcal{F}(u + \mathscr{D}^{s_1,s_2}u) \\ \mathcal{F}(u - \mathbf{i}\mathscr{D}^{s_1,s_2}u) \end{bmatrix}, \tag{1.3}$$

[1]Generic frame means a $K-$element frame belongs to an open dense subset of the set of all $K-$element frames in $\mathbb{R}^n$ or $\mathbb{C}^n$ [2]

where

$$(\mathscr{D}^{s_1,s_2}u)(t_1+t_2n_1)=\exp\left(\frac{2\pi\mathbf{i}s_1t_1}{n_1}+\frac{2\pi\mathbf{i}s_2t_2}{n_2}\right)u(t_1+t_2n_1),\quad 0\le t_i\le n_i-1,$$

with $i=1,2$ and integers s_1,s_2 coprime to n_1,n_2 respectively. However, they found this amount of measurements is practically insufficient to recover u exactly and stably, and $7n$ measurements are suggested instead. In our preliminary work [15], we demonstrated that $3n$ measurements with $s_1=s_2=1/2$ can recover phase information both theoretically and empirically. Furthermore, additional prior information [30, 30, 11, 51, 47] is helpful to establish the uniqueness of $\mathcal{M}$ as well as to design efficient PR algorithms. For example, Jaganathan *et al.* [30] proved that signals of aperiodic support can be uniquely recovered with high probability if the DFT dimension is no less than $2n$, while Wang and Xu [51] focused on the minimal number of measurements required to deal with sparse signals in both the real and complex cases.

In this paper, we consider the phase retrieval problem from the measurements that are contaminated by the Poisson noise. It is very useful, since any optical sensor detects photons, and the number of measured photons varies following the Poisson distribution in the sense that the noise level depends on the ground-truth image, *i.e.*, stronger noise appears at lower intensity. Furthermore, when the intensity value is high enough, the Poisson noise at this pixel behaves like a "Gaussian" noise and therefore Poisson noise can be approximated by a Gaussian distribution via Anscombe transformation[1, 38]. In order to denoise the data from such measurements, *prior* information is important in the reconstruction procedure, and please refer to [32, 19, 56, 46, 25, 31] for various ways of imposing prior knowledge.

We formulate a variational model by introducing a total variation (TV) regularization to enforce sparsity, which is widely used in image processing since the seminal work of [45]. We extend our previous work [15] (focusing on holographic pattern for real-valued images) to more general PR setting (1.1) and deal with the Poisson noise in measurements for both real and complex valued images. We prove the existence and uniqueness of the minimizer to our proposed model, and an efficient alternative direction method of multipliers (ADMM) [26, 54, 5, 14] is designed with theoretical convergence guarantees. Numerical experiments are based on Fourier measurements generated by coded diffraction patterns (CDP) [9] and holographic pattern [11, 15] as expressed in (1.3). We show that satisfactory PR results can be obtained from $2n$ and $4n$ noisy measurements for real and complex images respectively, in consistent with the theoretical results in [2, 18]. Furthermore, the proposed method can deal with a large amount of downsampling; especially $0.4n$ measurements are shown to be sufficient for natural images.

The rest of this paper is organized as follows. In Section 2, a TV regularized model for Poisson noise removal in a phase retrieval problem, referred to as TV-PoiPR, is established, where the existence of the minimizer to the proposed model is obtained. Section 3 discusses an ADMM algorithm for TV-PoiPR with convergence analysis. Section 4 devotes to a special case, in which TV is not present, referred to as PoiPR. In this case, we can further prove the uniqueness of the solution. To the best of our knowledge, PoiPR, as a method to deal with noise-free PR measurements, is not systematically studied in the literature. Experiments are performed in Section 5 to demonstrate the effectiveness and robustness of the proposed methods for image recovery from noisy and incomplete magnitude data. Conclusions and future works are given in Section 6.

2. Proposed Model.

2.1. Maximum likelihood Estimation. Poisson noise is one of the most common types of noise that occurs for photon-counting. Its name is stemmed from Poisson distribution, defined as follows

$$\mathrm{Pr}_\mu(n) = \frac{e^{-\mu}\mu^n}{n!}, \qquad n \geq 0,$$

where μ is mean and standard deviation. The number of photons measured at each pixel, denoted as $f(i)$, follows i.i.d. Poisson distributions with μ being the ground-truth value, $g(i)$, for $i \in \Omega$, denoted as

$$f(i) \overset{\text{ind.}}{\sim} \mathrm{Poisson}(g(i)), \ \forall i \in \Omega. \tag{2.1}$$

Given the measured data f, the denoising problem is then formulated via maximum likelihood estimation (MLE) of a clean image g, which can be expressed as $\max \mathrm{Pr}(g(i)|f(i))$. By Bayes' Law, we have

$$\mathrm{Pr}(g(i)|f(i)) = \frac{\mathrm{Pr}(f(i)|g(i))\mathrm{Pr}(g(i))}{\mathrm{Pr}(f(i))}. \tag{2.2}$$

Therefore, $\max \mathrm{Pr}(g(i)|f(i))$ is equivalent to $\max \mathrm{Pr}(f(i)|g(i))\mathrm{Pr}(g(i))$. It follows from the definition of Poisson distribution that

$$\mathrm{Pr}(f(i)|g(i)) = \mathrm{Pr}_{g(i)}(f(i)) = \frac{e^{-g(i)}g(i)^{f(i)}}{(f(i))!}, \tag{2.3}$$

which suggests to minimize the logarithm of the $\mathrm{Pr}(f(i)|g(i))\mathrm{Pr}(g(i))$ instead, *i.e.*

$$\begin{aligned} &\min_{g\geq 0} \sum_{i\in\Omega} -\log \mathrm{Pr}(f(i)|g(i)) - \log \mathrm{Pr}(g(i)) \\ =&\min_{g\geq 0} \sum_{i\in\Omega} (g(i) - f(i)\log g(i)) - \log \mathrm{Pr}(g(i)), \end{aligned} \tag{2.4}$$

where the first term is related with the famous Kullback-Leibler (KL) divergence, and define $0\log 0 = 0$ and $\log 0 = -\infty$.

Regularization often plays an important role in noise removal. Following the celebrated TV regularization, Le *et al.* considered a TV regularized model [32],

$$\min_{g\geq 0} \lambda \mathrm{TV}(g) + \sum_{i\in\Omega}(g(i) - f(i)\log g(i)), \tag{2.5}$$

where $\mathrm{TV}(g) = \|\nabla g\|_1 = \sum_j \sqrt{|(\nabla_x g)(j)|^2 + |(\nabla_y g)(j)|^2}$, and ∇_x and ∇_y define the $x-$direction and $y-$direction forward difference operators respectively. The model (2.5) can be regarded as incorporating a prior distribution, $\mathrm{Pr}(g) = \exp\big(-\lambda \mathrm{TV}(g)\big)$, into the MLE for Poisson noise removal. Efficient methods were proposed to solve the above total variation regularized model, such as gradient descent [32], expectation-maximization (EM) algorithm [6], mutigrid method [13], and fast splitting methods [23, 7, 12, 52].

2.2. TV-PoiPR. In this paper, we consider a phase retrieval problem in which the measurements $b = |\mathcal{A}u|^2$ in (1.1) are corrupted by Poisson noise, that is,

$$f(i) \overset{\text{ind.}}{\sim} \text{Poisson}(b(i)),\ \forall i \in D,$$

with an undersampling set $D \subseteq \Lambda$. In the light of (2.5), we establish a minimization problem, referred to as "TV-PoiPR",

$$\min_{u\in\mathbb{C}^n} \mathcal{G}(u) := \lambda \mathrm{TV}(u) + \frac{1}{2}\sum_{i\in D}(|(\mathcal{A}u)(i)|^2 - f(i)\log|(\mathcal{A}u)(i)|^2). \tag{2.6}$$

where $u \in \mathbb{C}^n$ is an underlying image that we want to reconstruct from magnitude data.

The rest of the section devotes to theoretical analysis of TV-PoiPR. In particular, we prove existence of solutions to (2.6) under mild conditions. The uniqueness of solutions to (2.6) with $\lambda = 0$ is studied in Section 4. Note that we only focus on the discrete setting, while it is straightforward to extend the analysis to a continuous setting using the compactness property of bounded variation (BV) space and lower semi-continuity of the objective functional $\mathcal{G}$.

Theorem 1. *Assume that there exists a positive number β for the operator $\mathcal{A}$, such that*

$$\beta\|u\|_2 \le \|\mathcal{A}u\|_{2,D},\ \forall u \in \mathbb{C}^n, \tag{2.7}$$

with $\|z\|_{2,D} = \sqrt{\sum_{i\in D} z(i)^2},\ z \in \mathbb{C}^m$, then there exists a minimizer u^* for (2.6), i.e.

$$u^* = \arg\min_{u\in\mathbb{C}^n} \mathcal{G}(u).$$

Proof. Define a data fidelity for Poisson noise,

$$J(u) = \sum_{i\in D}(|(\mathcal{A}u)(i)|^2 - f_i \log|(\mathcal{A}u)(i)|^2). \tag{2.8}$$

Since

$$x - f(i)\log x \ge f(i) - f(i)\log f(i), \forall x \ge 0, \text{and } i \in \{i \in D:\ f(i) > 0\},$$

we have

$$J(u) \ge \sum_{\{i\in D:\ f(i)>0\}} (f(i) - f(i)\log|f(i)|^2),\ \forall u \in \mathbb{C}^n, \tag{2.9}$$

such that $\mathcal{G}(u)$ is bounded below. Therefore, we can choose a minimizing sequence $\{u_n\}$, such that

$$\mathcal{G}(u_0) \ge \mathcal{G}(u_1) \ge \cdots \ge \mathcal{G}(u_j) \ge \cdots.$$

$$\begin{aligned}
\mathcal{G}(u_0) \ge & \lambda\mathrm{TV}(u_j) + \sum_{i\in D}(|z_j(i)|^2 - f(i)\log|z_j(i)|^2) \\
\ge & \sum_{\{i\in D:\ |z_j(i)|\ge 1\}} |z_j(i)|^2 - \|f\|_\infty \sum_{\{i\in D:\ |z_j(i)|\ge 1\}} \log|z_j(i)|^2 \\
= & \sum_{\{i\in D:\ |z_j(i)|\ge 1\}} |z_j(i)|^2 - \|f\|_\infty \log \prod_{\{i\in D:\ |z_j(i)|\ge 1\}} |z_j(i)|^2 \\
\ge & \sum_{\{i\in D:\ |z_j(i)|\ge 1\}} |z_j(i)|^2 - m_0\|f\|_\infty \log\Big(\frac{1}{m_0} \sum_{\{i\in D:\ |z_j(i)|\ge 1\}} |z_j(i)|^2\Big)
\end{aligned} \tag{2.10}$$

where $z_j = \mathcal{A}u_j$, and $m_0 = \#\{i \in D : |z_j(i)| \geq 1\} \leq m$. Therefore, one readily obtains that $\{\|\mathcal{A}u_j\|_{2,D}\}$ is a bounded sequence such that $\{\|u_j\|\}$ is also bounded by (2.7). By the compactness of discrete L^2 space, one can readily select a convergent subsequence $\{u_{j_k}\} \subseteq \{u_j\}$, such that $\lim_{k\to+\infty} u_{j_k} = u^*$. By the continuity of the objective functional $\mathcal{G}$, we have

$$\lim_{k\to+\infty} \mathcal{G}(u_{j_k}) = \mathcal{G}(u^*).$$

That concludes the proof. ▯

Furthermore, we can relax the assumption (2.7) as

$$\|\mathcal{A}\mathbf{1}\|_{2,D} \neq 0, \tag{2.11}$$

where $\mathbf{1} \in \mathbb{C}^m$ whose elements are all equal to one, and obtain a similar results.

Proposition 2.1. *If the assumption* (2.11) *holds, there exists a minimizer* u^* *for* (2.6)*,* i.e.

$$u^* = \arg\min_{u\in\mathbb{C}^n} \mathcal{G}(u).$$

Proof. Since $\mathcal{G}(u)$ is bounded below by (2.9), we can choose a minimizing sequence $\{u_n\}$, such that

$$\mathcal{G}(u_0) \geq \mathcal{G}(u_1) \geq \cdots \geq \mathcal{G}(u_j) \geq \cdots.$$

We will show the minimizing sequence is bounded. Rewrite

$$u_j = \hat{u}_j + c_j\mathbf{1}, \tag{2.12}$$

where the constant $c_j = \sum_{i\in\Lambda} u_j(i)/m$. Since it exists a positive constant C, such that

$$\|z - \frac{1}{m}\sum_{i\in\Lambda} z(i)\mathbf{1}\| \leq C\mathrm{TV}(z),\ \forall\ z \in \mathbb{C}^m,$$

we have the boundedness of $\{\hat{u}_j\}$. By (2.10), we obtain the boundedness of $\|\mathcal{A}u_j\|_{2,D}$. Since

$$\begin{aligned} |c_j|\|\mathcal{A}\mathbf{1}\|_{2,D} &= \|\mathcal{A}\hat{u}_j - \mathcal{A}u_j\|_{2,D} \\ &\leq \|\mathcal{A}\hat{u}_j\|_{2,D} + \|\mathcal{A}u_j\|_{2,D} \\ &\leq \|\mathcal{A}\hat{u}_j\| + \|\mathcal{A}u_j\|_{2,D} \\ &\leq \|\mathcal{A}\|\|\hat{u}_j\| + \|\mathcal{A}u_j\|_{2,D}, \end{aligned} \tag{2.13}$$

with the help of assumption (2.11), one can see that $\{c_j\}$ is bouneded. It follows from (2.12) that $\{u_j\}$ is bounded. Due to the lower semi-continuity of $\mathcal{G}$, we can complete the proof by similar analysis as in Theorem 1. ▯

Remark 2.1. *For any regularization term $\mathcal{R}(u)$ that is lower semi-continuous, we can prove the existence of minimizer of*

$$\min_{u\in\mathbb{C}^n} \mathcal{R}(u) + J(u),$$

where $J(u)$ is a data fidelity term, defined in (2.8)*.*

Remark 2.2. *The assumptions of* (2.7) *and* (2.11) *are satisfied for some special patterns as CDP and holographic pattern without undersampling, or, when $D = \Lambda$.*

In general, it is difficult to show that a nonconvex minimization problem has a unique solution. We will prove the uniqueness of the minimizer (2.6) without regularization term *i.e.* $\lambda = 0$ in Section 4.

3. Numerical Algorithms. We apply the ADMM algorithm [5] to solve the proposed model (2.6), which is equivalent to

$$\min_{u\in\mathbb{C}^n} \lambda\|\boldsymbol{p}\|_1 + \frac{1}{2}\sum_{i\in D}(|z(i)|^2 - f(i)\log|z(i)|^2), \quad s.t. \;\; z = \mathcal{A}u, \;\boldsymbol{p} = \nabla u, \tag{3.1}$$

where $\nabla u = (\nabla_x u, \nabla_y u)$ denotes the gradient operator and we assume zero boundary conditions for these two gradient operators. One can readily construct the augmented Lagrangian as

$$\begin{aligned}\mathcal{L}_{r_1,r_2}(u,z,\boldsymbol{p};v,\boldsymbol{q}) :=& \max_{\boldsymbol{q},v}\min_{u,z,\boldsymbol{p}} \lambda\|\boldsymbol{p}\|_1 + \frac{1}{2}\sum_{i\in D}(|z(i)|^2 - f(i)\log|z(i)|^2)\\ &+\mathrm{Re}\langle z-\mathcal{A}u, v\rangle + \mathrm{Re}\langle \boldsymbol{p}-\nabla u,\boldsymbol{q}\rangle + \frac{r_1}{2}\|z-\mathcal{A}u\|^2 + \frac{r_2}{2}\|\boldsymbol{p}-\nabla u\|^2,\end{aligned} \tag{3.2}$$

where $\langle\cdot,\cdot\rangle$ denotes the complex inner product of two vectors, and $v, \boldsymbol{q}$ are called Lagrange multipliers or dual variables. Alternative minimization for the above Lagrangian consists of solving three subproblems with respect to $u, z, \boldsymbol{p}$, followed by updating dual variables. Below we elaborate on how to solve for $u, z, \boldsymbol{p}$.

3.1. u-subproblem. The u-subproblem is

$$\min_{u\in\mathbb{C}^n} \mathrm{Re}\langle z-\mathcal{A}u, v\rangle + \mathrm{Re}\langle \boldsymbol{p}-\nabla u,\boldsymbol{q}\rangle + \frac{r_1}{2}\|z-\mathcal{A}u\|^2 + \frac{r_2}{2}\|\boldsymbol{p}-\nabla u\|^2, \tag{3.3}$$

which can be simplified as

$$\min_{u\in\mathbb{C}^n} \frac{r_1}{2}\|z+v/r_1-\mathcal{A}u\|^2 + \frac{r_2}{2}\|\boldsymbol{p}+\boldsymbol{q}/r_2-\nabla u\|^2. \tag{3.4}$$

The operator $\mathcal{A}$ can be rewritten by the summation of two linear real operators $\mathcal{A} = \mathcal{A}_1 + \mathbf{i}\mathcal{A}_2$. In a similar way, $u = u_1 + \mathbf{i}u_2$ with $u_i \in \mathbb{R}^n$. One can obtain $\mathcal{A}u = \mathcal{A}_1u_1 - \mathcal{A}_2u_2 + \mathbf{i}(\mathcal{A}_1u_2 + \mathcal{A}_2u_1)$. Letting $v_0 = z + v/r_1, \boldsymbol{p}_0 = \boldsymbol{p} + \boldsymbol{q}/r_2$, (3.4) is equivalent to

$$\begin{aligned}&\frac{r_1}{2}\|v_0-\mathcal{A}u\|^2 + \frac{r_2}{2}\|\boldsymbol{p}_0-\nabla u\|^2\\ =&\frac{r_1}{2}\|\mathcal{A}_1u_1-\mathcal{A}_2u_2-\mathbf{Re}(v_0)\|^2 + \frac{r_1}{2}\|\mathcal{A}_1u_2+\mathcal{A}_2u_1-\mathbf{Im}(v_0)\|^2\\ &+\frac{r_2}{2}\|\mathbf{Re}(\boldsymbol{p}_0)-\nabla u_1\|^2 + \frac{r_2}{2}\|\mathbf{Im}(\boldsymbol{p}_0)-\nabla u_2\|^2.\end{aligned} \tag{3.5}$$

By computing the derivative with respect to u_i, one obtains

$$\begin{aligned}&\begin{bmatrix} r_1(A_1^TA_1 + A_2^TA_2) - r_2\Delta & -r_1(A_1^TA_2 - A_2^TA_1)\\ r_1(A_1^TA_2 - A_2^TA_1) & r_1(A_1^TA_1 + A_2^TA_2) - r_2\Delta\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}\\ =&\begin{bmatrix} r_1(A_1^T\mathbf{Re}(v_0) + A_2^T\mathbf{Im}(v_0)) - r_2(\nabla\cdot\mathbf{Re}(\boldsymbol{p}_0))\\ r_1(-A_2^T\mathbf{Re}(v_0) + A_1^T\mathbf{Im}(v_0)) - r_2(\nabla\cdot\mathbf{Im}(\boldsymbol{p}_0))\end{bmatrix},\end{aligned} \tag{3.6}$$

with suitable boundary condition for u_i, where $\Delta u = \nabla\cdot(\nabla u)$, the divergence operator $\nabla(\cdot)$ denotes the conjugate operator of gradient ∇. Readily we have

$$
\begin{aligned}
&\begin{bmatrix} r_1(\mathbf{Re}(\mathcal{A}^*\mathcal{A})) - r_2\Delta & -r_1(\mathbf{Im}(\mathcal{A}^*\mathcal{A})) \\ \\ r_1(\mathbf{Im}(\mathcal{A}^*\mathcal{A})) & r_1(\mathbf{Re}(\mathcal{A}^*\mathcal{A})) - r_2\Delta \end{bmatrix} \begin{bmatrix} u_1 \\ \\ u_2 \end{bmatrix} \\
&= \begin{bmatrix} r_1(\mathbf{Re}(\mathcal{A}^*v_0)) - r_2(\nabla\cdot\mathbf{Re}(\boldsymbol{p}_0)) \\ \\ r_1(\mathbf{Im}(\mathcal{A}^*v_0)) - r_2(\nabla\cdot\mathbf{Im}(\boldsymbol{p}_0)) \end{bmatrix},
\end{aligned} \tag{3.7}
$$

since

$$
A_1^T A_1 + A_2^T A_2 = \mathbf{Re}(\mathcal{A}^*\mathcal{A}), \qquad A_1^T A_2 - A_2^T A_1 = \mathbf{Im}(\mathcal{A}^*\mathcal{A}). \tag{3.8}
$$

Note that there is a unique solution to the $u-$subproblem, as characterized in Theorem 2.

Theorem 2. *The linear equations* (3.7) *admit a unique solution.*

Proof. Let

$$
\mathcal{B} := \begin{bmatrix} r_1(\mathbf{Re}(\mathcal{A}^*\mathcal{A})) - r_2\Delta & -r_1(\mathbf{Im}(\mathcal{A}^*\mathcal{A})) \\ \\ r_1(\mathbf{Im}(\mathcal{A}^*\mathcal{A})) & r_1(\mathbf{Re}(\mathcal{A}^*\mathcal{A})) - r_2\Delta \end{bmatrix}. \tag{3.9}
$$

We show that the linear operator $\mathcal{B}$ is strictly positive definite. For this purpose, we calculate

$$
\begin{aligned}
&\langle \mathcal{B}(u_1,u_2)^T, (u_1,u_2)^T\rangle = \langle r_1\mathbf{Re}(\mathcal{A}^*\mathcal{A})u_1 - r_2\Delta u_1, u_1\rangle - \langle r_1\mathbf{Im}(\mathcal{A}^*\mathcal{A})u_2, u_1\rangle \\
&\qquad\qquad + \langle r_1\mathbf{Im}(\mathcal{A}^*\mathcal{A})u_1, u_2\rangle + \langle r_1\mathbf{Re}(\mathcal{A}^*\mathcal{A})u_2 - r_2\Delta u_2, u_2\rangle \\
&= \langle r_1\mathbf{Re}(\mathcal{A}^*\mathcal{A})u_1 - r_2\Delta u_1, u_1\rangle + \langle r_1\mathbf{Re}(\mathcal{A}^*\mathcal{A})u_2 - r_2\Delta u_2, u_2\rangle \\
&= r_1\left(\langle \mathbf{Re}(\mathcal{A}^*\mathcal{A})u_1, u_1\rangle + \langle \mathbf{Re}(\mathcal{A}^*\mathcal{A})u_2, u_2\rangle\right) + r_2\left(\langle -\Delta u_1, u_1\rangle + \langle -\Delta u_2, u_2\rangle\right).
\end{aligned}
$$

Since the operator Δ is negative definite with the difference scheme of zero boundary condition and $\mathbf{Re}(\mathcal{A}^*\mathcal{A})$ is semi-positive by (3.8), we have

$$
\langle \mathcal{B}(u_1,u_2)^T, (u_1,u_2)^T\rangle > 0,
$$

for $(u_1,u_2) \neq 0$. Therefore the positivity of $\mathcal{B}$ guarantees the uniqueness of the solution to u-subproblem (3.7). □

Remark 3.1. *We can simplify the solution of u-subproblem if the matrix $\mathcal{A}$ involves Fourier measurements with masks $\{I_j\}_{j=1}^k$ as*

$$
\mathcal{A}u = \begin{bmatrix} \mathcal{F}(I_1\circ u) \\ \mathcal{F}(I_2\circ u) \\ \vdots \\ \mathcal{F}(I_k\circ u) \end{bmatrix}, \tag{3.10}
$$

where $\circ$ denotes the pointwise multiplication, I_j is a (mask) matrix indexed by j, each of which is represented by a vector in $\mathbb{C}^n$ in a lexicographical order. Therefore we have $\mathcal{A}^*\mathcal{A} = \sum_j I_j^*\circ I_j$, which is a real-valued matrix. For such pattern, we can obtain $u = u_1 + \mathbf{i}u_2$ as

$$
\begin{aligned}
u_1 &= \left(r_1(\mathbf{Re}(\mathcal{A}^*\mathcal{A})) - r_2\Delta\right)^{-1}\left(r_1\mathbf{Re}(\mathcal{A}^*v_0) - r_2\nabla\cdot\mathbf{Re}(\boldsymbol{p}_0)\right), \\
u_2 &= \left(r_1(\mathbf{Re}(\mathcal{A}^*\mathcal{A})) - r_2\Delta\right)^{-1}\left(r_1\mathbf{Im}(\mathcal{A}^*v_0) - r_2\nabla\cdot\mathbf{Im}(\boldsymbol{p}_0)\right),
\end{aligned} \tag{3.11}
$$

by solving the following block-diagonal equation

(3.12)
$$\begin{bmatrix} r_1\mathbf{Re}(\mathcal{A}^*\mathcal{A}) - r_2\Delta & 0 \\ 0 & r_1\mathbf{Re}(\mathcal{A}^*\mathcal{A}) - r_2\Delta \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} r_1\mathbf{Re}(\mathcal{A}^* v_0) - r_2\nabla\cdot\mathbf{Re}(\boldsymbol{p}_0) \\ r_1\mathbf{Im}(\mathcal{A}^* v_0) - r_2\nabla\cdot\mathbf{Im}(\boldsymbol{p}_0) \end{bmatrix}. \blacksquare$$

A simplified forms can be written as

$$u = (r_1\mathbf{Re}(\mathcal{A}^*\mathcal{A}) - r_2\Delta)^{-1}(r_1\mathcal{A}^* v_0 - r_2\nabla\cdot\boldsymbol{p}_0).$$

As the positivity of the coefficient matrix of the above equation in the proof of Theorem 2 is shown, one can use conjugate gradient methods to solve it efficiently and few iterations are needed in practice. Similar calculation reveals that the matrix $\mathcal{B}$ corresponding to the holographic pattern has a block-diagonal form as well.

Remark 3.2. *We consider the case of real signal $u \in \mathbb{R}^n$, and real auxiliary variables $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^{2n}$; in other words, we solve the minimization problem on a real-valued space. The Euler equation to the subproblem* (3.3) *with respect to real-valued u and $\boldsymbol{p}$ is obtained by*

$$(r_1\mathbf{Re}(\mathcal{A}^*\mathcal{A}) - r_2\Delta)\,u = r_1\mathbf{Re}(\mathcal{A}^* v_0) - r_2\nabla\cdot\boldsymbol{p}_0.$$

3.2. z-subproblem. We rewrite the z-subproblem as

$$\min_{z\in\mathbb{C}^m} \frac{1}{2}\sum_{i\in D}(|z(i)|^2 - f(i)\log|z(i)|^2) + \frac{r_1}{2}\|z - \mathcal{A}u + v/r_1\|^2. \tag{3.13}$$

It is straightforward that the minimization with respect to z is equivalent to minimizing with respect to each entry $z(i)$ independently and for $i \in \Lambda \setminus D$, an optimal solution is $z^*(i) = (\mathcal{A}u)(i) - v(i)/r_1$.

As for $i \in D$, we can decompose the minimization problem with respect to $z(i)$,

(3.14)
$$z^*(i) = \arg\min_{z(i)\in\mathbb{C}} \frac{1}{2}(|z(i)|^2 - f(i)\log|z(i)|^2) + \frac{r_1}{2}|z(i) - (\mathcal{A}u)(i) + v(i)/r_1|^2, \forall i \in D.$$

into two subproblems, i.e., $|z(i)|$ and $\mathrm{sign}(z(i))$ for $\mathrm{sign}(z(i)) = \frac{z(i)}{|z(i)|}$. One can readily obtain $\mathrm{sign}(z^*(i)) = \mathrm{sign}((\mathcal{A}u)(i) - v(i)/r_1)$. To minimize the subproblem with respect to $|z^*(i)|$, we have

$$|z^*(i)| = \arg\min_{\rho\in\mathbb{R}_+} \frac{1}{2}(\rho^2 - f(i)\log\rho^2) + \frac{r_1}{2}\big(\rho - |(\mathcal{A}u)(i) - v(i)/r_1|\big)^2,$$

which has a closed-form solution,

$$|z^*(i)| = \frac{r_1|(\mathcal{A}u)(i) - v(i)/r_1| + \sqrt{r_1^2|(\mathcal{A}u)(i) - v(i)/r_1|^2 + 4(1+r_1)f(i)}}{2(1+r_1)}.$$

Letting $w = \mathcal{A}u - v/r_1$, we arrive at a simplified expression of

$$z^*(i) = \frac{r_1|w(i)| + \sqrt{r_1^2|w(i)|^2 + 4(1+r_1)f(i)}}{2(1+r_1)}\mathrm{sign}(w(i)),\ \forall\, i \in D.$$

3.3. $\boldsymbol{p}$-subproblem and overall algorithm. At last, we consider the $\boldsymbol{p}$-subproblem

$$\min_{\boldsymbol{p}\in\mathbb{C}^{2n}} \lambda\|\boldsymbol{p}\|_1 + \frac{r_2}{2}\|\boldsymbol{p}-\nabla u+\boldsymbol{q}/r_2\|^2, \tag{3.15}$$

The solution is a soft shrinkage of variable $\nabla u - \boldsymbol{q}/r_2$ as

$$\boldsymbol{p}^* = \mathrm{Thresh}_{\lambda/r_2}(\nabla u - \boldsymbol{q}/r_2)$$

with

$$\mathrm{Thresh}_\eta(\boldsymbol{q}) = \max\{0, |\boldsymbol{q}|-\eta\}\,\mathrm{sign}(\boldsymbol{q}).$$

In summary, a pseudo code of ADMM for solving TV-PoiPR is provided in Algorithm I.

Algorithm I: ADMM for TV-PoiPR (2.6)

---

1. Initialization: Set $\boldsymbol{q}^0 = 0, v^0 = 0$, u^0 is randomly generated, $z^0 = \mathcal{A}u^0, \boldsymbol{p}^0 = \nabla u^0$. $j = 0$.
2. Solve $u^{j+1} = u_1^{j+1} + \mathbf{i}u_2^{j+1}$ by solving the following equations as
$$\mathcal{B}\begin{bmatrix} u_1^{j+1} \\ \\ u_2^{j+1}\end{bmatrix} = \begin{bmatrix} r_1\mathbf{Re}(\mathcal{A}^*v_0^j) - r_2\nabla\cdot\mathbf{Re}(\boldsymbol{p}_0^j) \\ \\ r_1\mathbf{Im}(\mathcal{A}^*v_0^j) - r_2\nabla\cdot\mathbf{Im}(\boldsymbol{p}_0^j)\end{bmatrix},$$
with $v_0^j = z^j + v^j/r_1, \boldsymbol{p}_0^j = \boldsymbol{p}^j + \boldsymbol{q}^j/r_2$, and the operator $\mathcal{B}$ defined in (3.9).
3. Solve z^{j+1} and $\boldsymbol{p}^{k+1}$ in parallel by
(3.16)
$$z^{j+1}(i) = \begin{cases} \dfrac{r_1|w^j(i)| + \sqrt{r_1^2|w^j(i)|^2 + 4(1+r_1)f(i)}}{2(1+r_1)}\mathrm{sign}(w^j(i)), \forall\ i\in D, \\ w^j(i), \qquad\qquad \forall i\in\Lambda\setminus D, \end{cases}$$
with $w^j = \mathcal{A}u^{j+1} - v^j/r_1$, and
$$\boldsymbol{p}^{j+1} = \mathrm{Thresh}_{\lambda/r_2}(\nabla u^{j+1} - \boldsymbol{q}^j/r_2). \tag{3.17}$$
4. Update multipliers as
$$\begin{aligned} v^{j+1} &= v^j + r_1(z^{j+1} - \mathcal{A}u^{j+1}), \\ \boldsymbol{q}^{j+1} &= \boldsymbol{q}^j + r_2(\boldsymbol{p}^{j+1} - \nabla u^{j+1}). \end{aligned} \tag{3.18}$$
5. If some stopping condition is satisfied, stop the iterations and output the iterative solution; else set $j = j+1$, and goto Step 2.

3.4. Convergence analysis. We then discuss a convergent behavior of the proposed algorithm. We show that the algorithm converges to a saddle point by satisfying Karush-Kuhn-Tucker (KKT) conditions, which is a typical situation for nonconvex problems. The KKT conditions for the Lagrangian $\mathcal{L}_{r_1,r_2}(u, z, \boldsymbol{p}; \boldsymbol{q}, v)$ in (3.2) are defined as follows,

$$\begin{cases} \partial_u \mathcal{L}_{r_1,r_2}(\tilde{u}, \tilde{z}, \tilde{\boldsymbol{p}}; \tilde{v}, \tilde{\boldsymbol{q}}) = 0, \\ \partial_z \mathcal{L}_{r_1,r_2}(\tilde{u}, \tilde{z}, \tilde{\boldsymbol{p}}; \tilde{v}, \tilde{\boldsymbol{q}}) = 0, \\ \partial_{\boldsymbol{p}} \mathcal{L}_{r_1,r_2}(\tilde{u}, \tilde{z}, \tilde{\boldsymbol{p}}; \tilde{v}, \tilde{\boldsymbol{q}}) \ni 0, \\ \partial_v \mathcal{L}_{r_1,r_2}(\tilde{u}, \tilde{z}, \tilde{\boldsymbol{p}}; \tilde{v}, \tilde{\boldsymbol{q}}) = 0, \\ \partial_{\boldsymbol{q}} \mathcal{L}_{r_1,r_2}(\tilde{u}, \tilde{z}, \tilde{\boldsymbol{p}}; \tilde{v}, \tilde{\boldsymbol{q}}) = 0. \end{cases} \tag{3.19}$$

for any saddle point $(\tilde{u}, \tilde{z}, \tilde{\boldsymbol{p}}, \tilde{v}, \tilde{\boldsymbol{q}})$. Since the Lagrangian $\mathcal{L}_{r_1,r_2}(u, z, \boldsymbol{p}; \boldsymbol{q}, v)$ is nonconvex with respect to $u, z, \boldsymbol{p}$, we detail the KKT conditions corresponding to these three variables:

$$\mathcal{B} \begin{bmatrix} \tilde{u}_1 \\ \\ \tilde{u}_2 \end{bmatrix} = \begin{bmatrix} r_1 \mathbf{Re}(\mathcal{A}^* \tilde{z} + \tilde{v}/r_1) - r_2 \nabla \cdot \mathbf{Re}(\tilde{\boldsymbol{p}} + \tilde{\boldsymbol{q}}/r_2) \\ \\ r_1 \mathbf{Im}(\mathcal{A}^* \tilde{z} + \tilde{v}/r_1) - r_2 \nabla \cdot \mathbf{Im}(\tilde{\boldsymbol{p}} + \tilde{\boldsymbol{q}}/r_2) \end{bmatrix}, \tag{3.20}$$

$$\tilde{z}(i) = \begin{cases} \dfrac{r_1 |\tilde{w}(i)| + \sqrt{r_1^2 |\tilde{w}(i)|^2 + 4(1+r_1) f(i)}}{2(1+r_1)} \mathrm{sign}(\tilde{w}(i)), \forall\ i \in D, \\ \tilde{w}(i), \forall\ i \in \Lambda \setminus D, \end{cases} \tag{3.21}$$

$$\text{with } \tilde{w} = \mathcal{A}\tilde{u} - \tilde{v}/r_1,$$

$$\begin{cases} 0 \ni \lambda \partial_{\boldsymbol{p}_1} \|\tilde{\boldsymbol{p}}_1\|_1 + r_2 \big(\tilde{\boldsymbol{p}}_1 - \nabla \tilde{u}_1 + \mathbf{Re}(\tilde{q})/r_2\big), \\ 0 \ni \lambda \partial_{\boldsymbol{p}_2} \|\tilde{\boldsymbol{p}}_2\|_1 + r_2 \big(\tilde{\boldsymbol{p}}_2 - \nabla \tilde{u}_2 + \mathbf{Im}(\tilde{q})/r_2\big), \end{cases} \tag{3.22}$$

$$\tilde{z} = \mathcal{A}\tilde{u}, \tag{3.23}$$

$$\tilde{\boldsymbol{p}} = \nabla \tilde{u}, \tag{3.24}$$

where $\tilde{u} = \tilde{u}_1 + \mathbf{i}\tilde{u}_2$ and $\tilde{\boldsymbol{p}} = \tilde{\boldsymbol{p}}_1 + \mathbf{i}\tilde{\boldsymbol{p}}_2$.

Theorem 3. *Assume that the successive differences of the two multiplies $\{v^j - v^{j-1}\}, \{\boldsymbol{q}^j - \boldsymbol{q}^{j-1}\}$ converge to zero and $\{u^j\}$ is bounded, then there exists an accumulation point of a subsequence of iterative sequences of Algorithm I that satisfies KKT conditions of the saddle point problem* (3.2).

Proof. We complete the proof in two steps. First, we show the boundedness of all the variables. Due to the update of two multipliers (3.18) and the assumption that their successive differences converge, one can derive that

$$\lim_{j \to +\infty} z^j - \mathcal{A}u^j = 0, \ \lim_{j \to +\infty} \boldsymbol{p}^j - \nabla u^j = 0, \tag{3.25}$$

which implies the boundedness of $\{z^j\}$ and $\{\boldsymbol{p}^j\}$. By (3.16), we have

$$|z^{j+1}(i)| = \begin{cases} \dfrac{r_1 |w^j(i)| + \sqrt{r_1^2 |w^j(i)|^2 + 4(1+r_1) f(i)}}{2(1+r_1)}, & \forall\ i \in D, \\ |w^j(i)|, & \forall\ i \in \Lambda \setminus D, \end{cases}$$

which demonstrates that $\{w^j\}$ is bounded and so is $\{v^j\}$ since $w^j = \mathcal{A}u^{j+1} - v^j/r_1$. By (3.17), we have

$$\begin{aligned} |\boldsymbol{p}^{j+1}| &= \max\big\{0, |\nabla u^{j+1} - \boldsymbol{q}^j/r_2| - \lambda/r_2\big\} \\ &\geq |\nabla u^{j+1} - \boldsymbol{q}^j/r_2| - \lambda/r_2 \\ &\geq |\boldsymbol{q}^j|/r_2 - |\nabla u^{j+1}| - \lambda/r_2, \end{aligned} \tag{3.26}$$

which gives the boundedness of $\{\boldsymbol{q}^j\}$ due to the boundedness of $\{u^j\}$ and $\{\boldsymbol{p}^j\}$.

The boundedness of all variables guarantees that there exists a subsequence $\{(u^{j_l},\boldsymbol{p}^{j_l},z^{j_l},v^{j_l},\boldsymbol{q}^{j_l})\}\subset\{(u^j,\boldsymbol{p}^j,z^j,v^j,\boldsymbol{q}^j)\}$ and $(\tilde u,\tilde{\boldsymbol{p}},\tilde z,\tilde v,\tilde{\boldsymbol{q}})$, such that

$$\lim_{l\to+\infty}(u^{j_l},\boldsymbol{p}^{j_l},z^{j_l},v^{j_l},\boldsymbol{q}^{j_l})=(\tilde u,\tilde{\boldsymbol{p}},\tilde z,\tilde v,\tilde{\boldsymbol{q}}).$$

We then prove that the point $(\tilde u,\tilde{\boldsymbol{p}},\tilde z,\tilde v,\tilde{\boldsymbol{q}})$ satisfies the KKT conditions. It follows from (3.25) that the KKT conditions with respect to z and $\boldsymbol{p}$, i.e., (3.23) and (3.24), are satisfied. Since $\mathcal{B}$ is a linear operator in a finite dimensional space, it is straightforward to prove (3.20). By the continuity of (3.16), (3.21) is obtained. Finally one can obtain that

$$\tilde{\boldsymbol{p}}=\mathrm{Thresh}_{\lambda/r_2}(\nabla\tilde u-\tilde{\boldsymbol{q}}/r_2),$$

which implies (3.22). Hence the proof is completed. □

4. A special case. We consider a special case of TV-PoiPR by setting $\lambda=0$ in (2.6), referred to as "PoiPR",

$$\min_{u\in\mathbb{C}^n}\mathcal{H}(u):=\frac{1}{2}\sum_{i\in D}(|(\mathcal{A}u)(i)|^2-f(i)\log|(\mathcal{A}u)(i)|^2). \tag{4.1}$$

It reduces to a phase retrieval problem from noise-free measurements. In this case, we can prove that there exists a unique solution to (4.1) under some conditions. Particularly, the existence of the solution is given in Theorem 1, while the uniqueness is based on the work of Conca *et al.* [18]. To make our paper self-contained, we include an important result from their paper. Define a $m-$element complex *frame* $\Phi=\{\phi_0,\phi_1,\cdots,\phi_{m-1}\}$, and a linear operator $\mathcal{A}_\Phi$ generated by the frame Φ in the sense of $(\mathcal{A}_\Phi u)_j=\langle\phi_j,u\rangle$. The following lemma gives the injectivity of phase retrieval for a generic frame Φ.

Lemma 4.1. *[Conca* et al. *[18]] If $m\geq 4n-4$, the mapping $\mathcal{M}(u)=|\mathcal{A}_\Phi u|^2$ is injective for a generic frame Φ.*

Using this lemma, we can prove the uniqueness of the solution to (4.1).

Theorem 4. *Under the same assumptions in Lemma 4.1 and $D=\Lambda, f\in$* Range$(\mathcal{M})$*, then the minimizer to* (4.1) *is unique.*

Proof. We consider a minimization problem with respect to w as follows,

$$w^*=\min_{w(i)\in\mathbb{R}_+}\frac{1}{2}\sum_{i\in\Lambda}(w(i)-f(i)\log w(i)), \tag{4.2}$$

which can be solved pointwisely. Particularly for each entry $w(i)$, the minimization problem has a unique solution, i.e.,

$$w^*(i)=\min_{\rho\in\mathbb{R}_+}\Big\{\rho-f(i)\log\rho\Big\}=f(i).$$

Therefore f is the unique minimizer to (4.2). Furthermore, since $f\in$ Range$(\mathcal{M})$, there exists a vector $u\in\mathbb{C}^n$ such that $f=\mathcal{M}(u)$. The uniqueness of such u is guaranteed by Lemma 4.1, which completes the proof. □

Computationally, one can construct the augmented Lagrangian for PoiPR, similarly to (3.2),

$$\max_v\min_{u,z}\frac{1}{2}\sum_{i\in D}(|z(i)|^2-f(i)\log|z(i)|^2)+\mathrm{Re}\langle z-\mathcal{A}u,v\rangle+\frac{r}{2}\|z-\mathcal{A}u\|^2. \tag{4.3}$$

Following a similar procedure of solving TV-PoiPR, we obtain an ADMM-based algorithm for PoiPR. The pseudo-code is summarized in Algorithm II, which is much simpler compared to Algorithm I (for solving TV-PoiPR).

Algorithm II: ADMM for PoiPR (4.1)

---

1. Initialization: Set $v^0 = 0$, u^0 is randomly generated, $z^0 = \mathcal{A}u^0$ and $j = 0$.
2. Solve $u^{j+1} = u_1^{j+1} + \mathbf{i}u_2^{j+1}$ by solving the following equations as

$$\begin{bmatrix} \mathbf{Re}(\mathcal{A}^*\mathcal{A}) & -\mathbf{Im}(\mathcal{A}^*\mathcal{A}) \\ \mathbf{Im}(\mathcal{A}^*\mathcal{A}) & \mathbf{Re}(\mathcal{A}^*\mathcal{A}) \end{bmatrix} \begin{bmatrix} u_1^{j+1} \\ u_2^{j+1} \end{bmatrix} = \begin{bmatrix} \mathbf{Re}(\mathcal{A}^* v_0^j) \\ \mathbf{Im}(\mathcal{A}^* v_0^j) \end{bmatrix}, \tag{4.4}$$

   with $v_0^j = z^j + v^j/r$.
3. Solve z^{j+1} by (3.16).
4. Update the multiplier as

$$v^{j+1} = v^j + r(z^{j+1} - \mathcal{A}u^{j+1}). \tag{4.5}$$

5. If the some stopping condition is satisfied, stop the iterations and output the iterative solution; else set $j = j + 1$, and goto Step 2.

Similar theoretical results of PoiPR can be obtained from the analysis of TV-PoiPR. In particular, we can prove that there exists a unique solution of $u-$subproblem (4.4) and Algorithm II converges to a stationary point, as characterized in Theorems 5-6 respectively. The proofs are omitted here.

Theorem 5. *Assume that the operator* $\mathbf{Re}(\mathcal{A}^*\mathcal{A})$ *is positive, then the linear equations* (4.4) *admit a unique solution vector.*

REMARK 4.1. *Note that the assumption in Theorem 5 holds for the Fourier measurements* (3.10) *without undersampling, since each entry of* $\sum_j I_j^* \circ I_j$ *is nonzero. Note that it also holds for holographic pattern* (1.3).

Theorem 6. *Assume that the successive differences of the multiplier* $\{v^j - v^{j-1}\}$ *converge to zero and* $\{u^j\}$ *is bounded, then there exists an accumulation point of a subsequence of iterative sequences of Algorithm II satisfies KKT condition of the saddle point problem* (4.3).

5. Numerical experiments. Although the proposed approaches are suitable for phase retrieval in general with arbitrary linear operator $\mathcal{A}$, we only focus on Fourier type measurements in the experimental section. In particular, we consider two different types of linear operators $\mathcal{A}$: coded diffraction pattern (CDP) with random masks and holographic pattern [11] with deterministic masks. For coded diffraction pattern, we use octanary CDPs; specifically each element of I_j in (3.10) takes a value randomly among the eight candidates, *i.e.*, $\{\pm\sqrt{2}/2, \pm\sqrt{2}\mathbf{i}/2, \pm\sqrt{3}, \pm\sqrt{3}\mathbf{i}\}$. For holographic pattern, the linear operator $\mathcal{A}$ is given in (1.3). Following our previous work [15], we choose $s_1 = s_2 = 1/2$, which is shown to give better phase retrieval with less measurements than s_1, s_2 taking integer values suggested by Candés [8].

The amount of Poisson noise at each pixel depends on its intensity value, as

η	r	r_1	r_2	λ
0.01	5	5	5×10^4	200
0.02	2	2	5×10^4	200
0.05	2	2	5×10^4	200
0.1	2	2	2×10^4	100

Table 1
Parameters of the proposed algorithms for recovering real-valued images from CDP measurements and results are shown in Figure 2.

discussed in the introduction section. Therefore, we introduce a scale factor $\eta \in (0, 1]$ to control the scale of the image intensities (or the number of photons), which is inverse proportional to the amount of noise added to the data. Let $u_\eta = \eta u$, and the measured data is expressed as

$$f(i) \overset{\text{ind.}}{\sim} \text{Poisson}(|(\mathcal{A}u_\eta)(i)|^2),\ \forall i \in D.$$

Signal-Noise-Ratio (SNR) is used to measure the reconstruction quality from noisy measurements

$$\text{SNR}(u, u_g) = -10 \log_{10} \frac{\sum\limits_{j\in\Omega} |u(j) - u_g(j)|^2}{\sum\limits_{j\in\Omega} |u_g(j)|^2},$$

where u_g is the ground truth image of size $n_1 \times n_2$ and u is the reconstructed image.

5.1. Effectiveness of TV regularization. We first show the effectiveness of TV regularization when measurements from CDP or holographic patterns are corrupted by Poisson noise. We conduct experiments to compare PR results with and without TV.

We start with phase retrieval of real-valued images from CDP measurements with $k = 2$ in (3.10) such that two groups of data are measured, i.e.,

$$\mathcal{A}u = \begin{bmatrix} \mathcal{F}(I_1 \circ u) \\ \mathcal{F}(I_2 \circ u) \end{bmatrix}. \tag{5.1}$$

In this setting, the operator $\mathcal{M}$ takes $2n$ measurements and hence its uniqueness to reconstruct a real signal is guaranteed [2]. We then consider $\eta = 0.01, 0.02, 0.05, 0.1$ to generate four different scales ground truth images in the sense that the smaller η is, the more noisy the image exhibits. The testing images are “Flower” with resolution 256×256, “Leaf” with resolution 363×378, and “Cameraman” with resolution 256×256, as shown in Figure 1, and the PR results using Algorithm I (with TV) and Algorithm II (without TV) are present in Figure 2. Readily one can infer TV’s effectiveness in Poisson noise removal in that the SNR values are increased at least 3dB and the reconstructed images have sharper edges, cleaner background, and higher contrast than the non-TV version. Improvement of TV over non-TV is more obvious for a smaller η, which corresponds to larger amount of noise in the data or lower amount of photon counts. The parameters for this set of examples are recorded in Table 1, and the choices of parameters are discussed in Section 5.3.

We then consider phase retrieval of a complex-valued image, called “Goldballs”, which is used in [53, 8]. We set $k = 4$ in (3.10) to guarantee the uniqueness of phase

(a) Flower (b) Maple (c) Camerman

Fig. 1. *Ground truth real-valued image (a) Flower with resolution* 256×256*, (b) Maple with resolution* 363×378*, and (c) Cameraman with resolution* 256×256.

retrieval for complex data [18], and we choose the parameters as $\lambda = 200, r_1 = r = 5, r_2 = 5.0 \times 10^4$. Figure 3 shows the phase retrieval results, which again illustrates the improvement of using TV regularization over the non-TV version by increasing at least 5dB in SNR values.

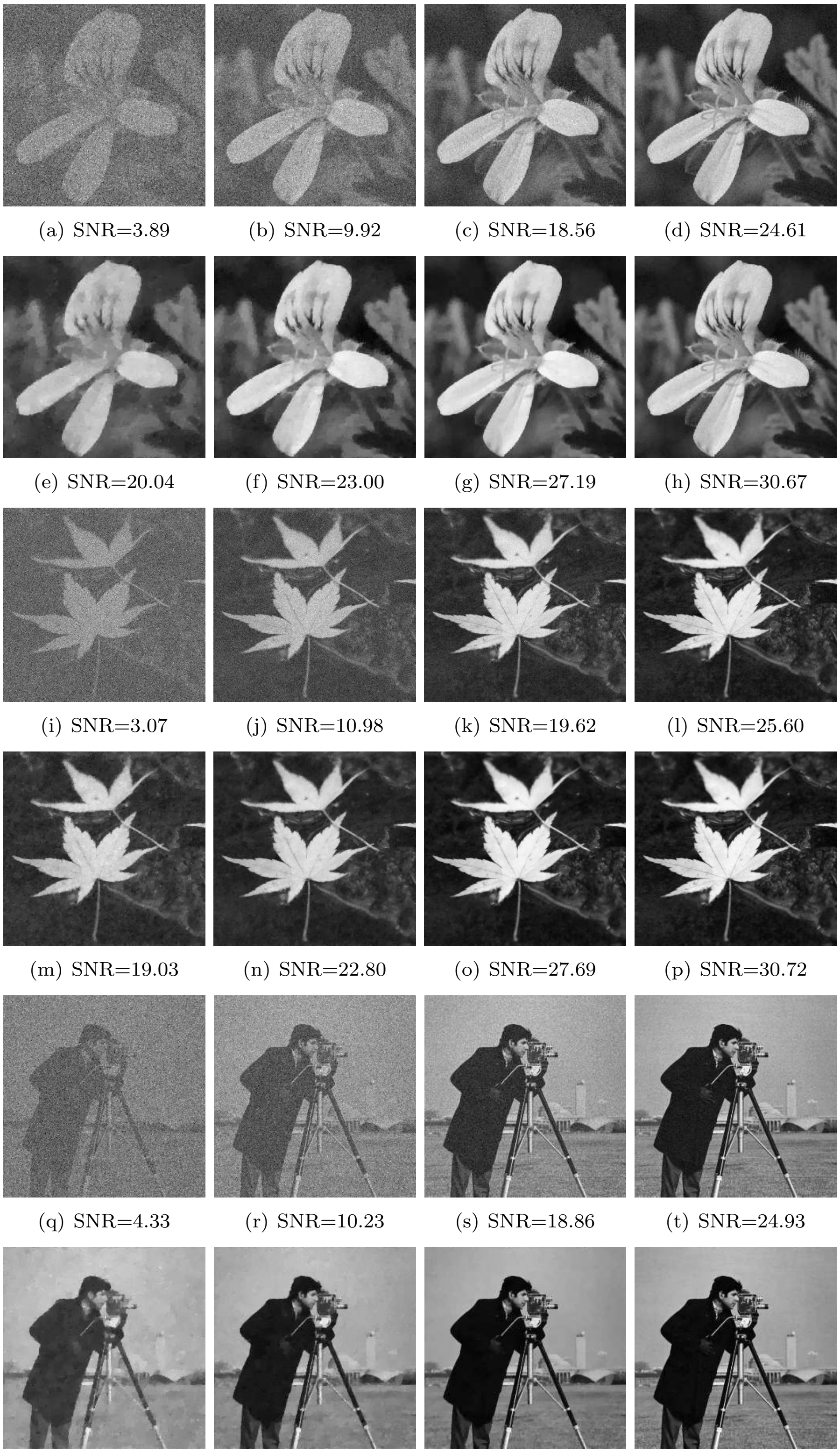
Finally we look at holographic patterns and set $s_1 = s_2 = 1/2$ in (1.3), in which the uniqueness is given in our preliminary work [15]. The parameter setting for this example is listed in Table 2 and we stop Algorithm I and Algorithm II after 500 iterations since it seems to require more iterations for such kind of deterministic masks than the ones for CDP. The results are given in Figure 4, which shows that Algorithm II produces a lot of artifacts, such as ringing artifacts and distortions, unlike the case of CDP as in Figure 2 and Figure 3. These severe artifacts inherited from noisy measurements are removed by the use of TV regularization, shown on the second row of Figure 4.

η	r	r_1	r_2	λ
0.01	5	5	2×10^5	200
0.02	2	3	1×10^5	200
0.05	1	1	8×10^4	200
0.1	0.5	1	1×10^4	200

Table 2
Parameters of the proposed algorithms for recovering real-valued images from holographic pattern measurements and results are shown in Figure 4.

5.2. Oversampling and undersampling. We examine the performance of the proposed algorithms with respect to oversampling in the sense that one collects noisy data with different number of masks, *i.e.* $k = 1, 3, 5$ and 7, which acts as an oversampling factor. Figure 5 shows that the more data we have (larger k), the better the phase retrieval results by either Algorithm I or Algorithm II. When comparing TV-PoiPR results for different k values, there is a diminishing gain in terms of SNR when the oversampling factor is increasing.

The amount of undersampling is controlled by the subset D. In particular, we define the undersampling ratio as $r_s = |D|/m$. The sampling subsets D are randomly generated, and please refer to [37] for details about various choices of the sampling masks. In this paper, we consider a random sampling mask D by using probability density function for 2D random sampling pattern with polynomial variable density sampling [36]. For different undersampling ratios $r_s = 0.1, 0.2, 0.4, 0.8$, we show one



(a) SNR=3.89 (b) SNR=9.92 (c) SNR=18.56 (d) SNR=24.61

(e) SNR=20.04 (f) SNR=23.00 (g) SNR=27.19 (h) SNR=30.67

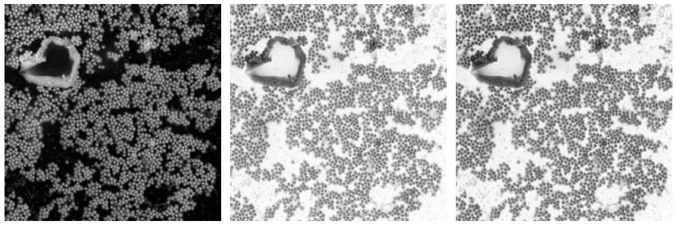
(i) SNR=3.07 (j) SNR=10.98 (k) SNR=19.62 (l) SNR=25.60

(m) SNR=19.03 (n) SNR=22.80 (o) SNR=27.69 (p) SNR=30.72

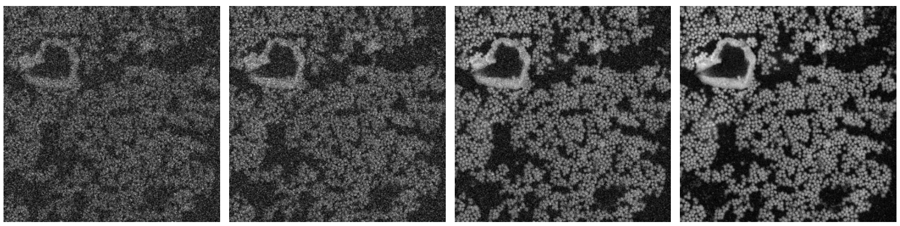
(q) SNR=4.33 (r) SNR=10.23 (s) SNR=18.86 (t) SNR=24.93

(u) SNR=18.11 (v) SNR=20.80 (w) SNR=24.99 (x) SNR=29.75

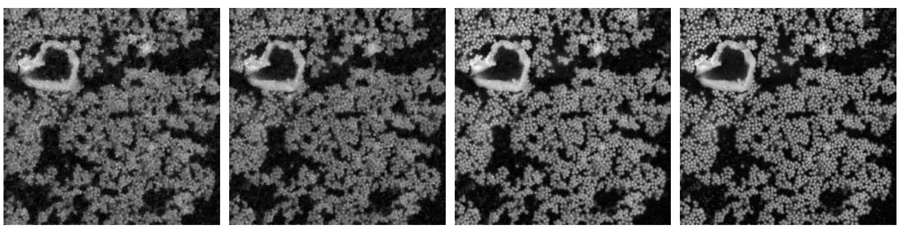
FIG. 2. *Performances of Algorithm I for TV-PoiPR (with TV) and Algorithm II for PoiPR (without TV) to recover three real-valued images in Figure 1 from CDP measurements. 1st, 3rd and 5th rows: results by Algorithm II for PoiPR; 2nd, 4th and 6th rows: results by Algorithm I for TV-PoiPR. From left to right:* $\eta = 0.01, 0.02, 0.05, 0.1$ *(the smaller* η *is, the more noisy the image exhibits.)*



(a) Magnitude (b) Real Part (c) Imaginary Part



(d) SNR=2.71 (e) SNR=4.20 (f) SNR=9.41 (g) SNR=15.73



(h) SNR=11.00 (i) SNR=12.09 (j) SNR=15.78 (k) SNR=21.51

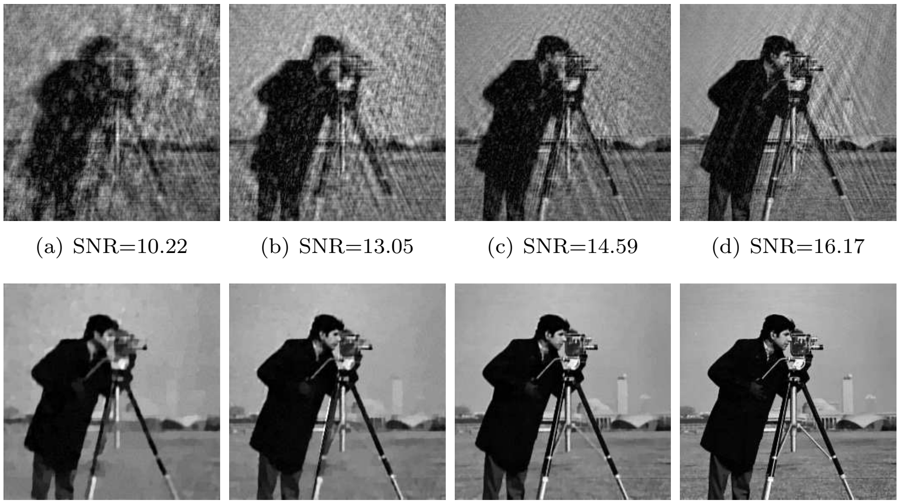
FIG. 3. *Performances of Algorithm I for TV-PoiPR and Algorithm II for PoiPR with ground truth complex-valued image "Goldballs" of resolution* 256×256. *First row: magnitude of the ground-truth image (a), real part (b) and imaginary parts (c). Second row: Resulted magnitudes by Algorithm II for PoiPR; Third row: Resulted magnitudes by Algorithm I for TV-PoiPR. For the second and third rows:* $\eta = 0.08, 0.1, 0.2, 1$ *(from the left to right).*

realization of each corresponding mask in Figure 6, and the PR results are given in Figure 7 with $k = 2, \eta = 0.05$ and the algorithmic parameters being $r = r_1 = 2, r_2 = 2.0 \times 10^4, \lambda = 200$. Large improvements of Algorithm I over Algorithm II are observed both visually and in terms of SNR. Figure 7 also demonstrates that TV can give satisfactory results at 20% undersampling ratio. As $k = 2$, the number of measurements at $r_s = 20\%$ is $0.4n$, which is below the theoretical limit capped at $2n$ for real signal recovery.

5.3. Discussion on the proposed algorithms. We will illustrate some properties of the proposed algorithms based on numerical simulation studies.

Impact by parameters. The parameters λ, r_1, r_2 in Algorithm I and r in Algorithm II are chosen by hand in order to obtain visually satisfactory results; and heuristically we stop the iterations of the proposed Algorithm I and II after 50 iterations as a default stopping condition.

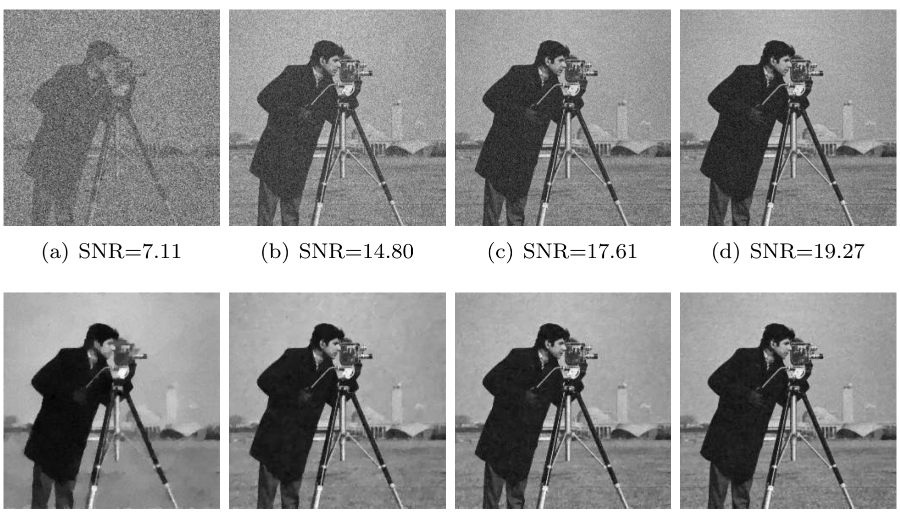
We now discuss the impact of r for Algorithm II and r_1, r_2 for Algorithm I when the amount of Poisson noise is at $\eta = 0.02$. Particularly, we choose r from $\{r^0 \times 2^{-l}, r^0 \times 2^{-l+1}, \cdots, r^0 \times 2^{l-1}, r^0 \times 2^l\}$ with $l = 7$, $r_0 = 2$ and plot the corresponding



(a) SNR=10.22 (b) SNR=13.05 (c) SNR=14.59 (d) SNR=16.17

(e) SNR=17.54 (f) SNR=19.10 (g) SNR=22.76 (h) SNR=27.10

FIG. 4. *Performances of Algorithm I for TV-PoiPR and Algorithm II for PoiPR with ground truth image "Cameraman" in Figure 1 (a) from measurements of holographic patterns. First row: results by Algorithm II for PoiPR; Second row: results by Algorithm I for TV-PoiPR. From left to right:* $\eta = 0.01, 0.02, 0.05, 0.1$.



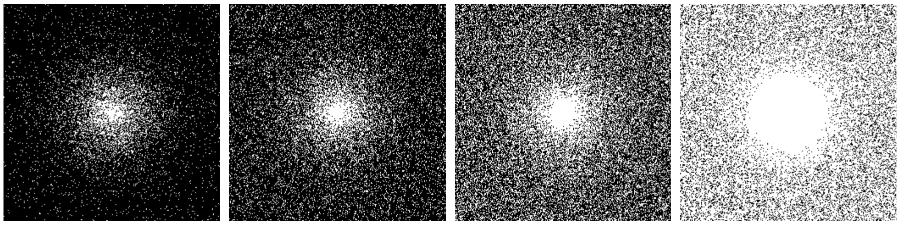
(a) SNR=7.11 (b) SNR=14.80 (c) SNR=17.61 (d) SNR=19.27

(e) SNR=19.00 (f) SNR=22.52 (g) SNR=23.80 (h) SNR=24.46

FIG. 5. *Performances with respect to oversampling factor* $k = 1, 3, 5$ *and* 7 *(from left to right). First row: results by Algorithm II for PoiPR; Second row: results by Algorithm I for TV-PoiPR*

SNRs of the reconstructed images in Figure 8, where we include the results of running Algorithm II 1000 iterations (red plus) versus default 50 iterations (blue dots). It seems that the parameter r only affects the convergence rate of proposed Algorithm II, and Algorithm II for solving a non-convex PR problem is rather insensitive to the parameter r in the range of $[1, 100]$.

As for Algorithm I, the impact of r_1, r_2, and λ are illustrated in Figure 9, in



(a) $r_s = 10\%$ (b) $r_s = 20\%$ (c) $r_s = 40\%$ (d) $r_s = 80\%$

FIG. 6. *Sampling masks D with respect to undersampling ratios r_s.*

which we fix $\lambda \in \{100, 200, 400\}$ and vary the parameters $(r_1, r_2) \in \{r_1^0 \times 2^{-l_1}, r_1^0 \times 2^{-l_1+1}, \cdots, r_1^0 \times 2^{l_1-1}, r_1^0 \times 2^{l_1}\} \times \{r_2^0 \times 2^{-l_2}, r_2^0 \times 2^{-l_2+1}, \cdots, r_2^0 \times 2^{l_2-1}, r_2^0 \times 2^{l_2}\}$ with $r_1^0 = 2, r_2^0 = 5 \times 10^4$, and $l_1 = l_2 = 7$. For the sake of better visualization, we raise the negative SNR values to zeros in Figure 9, which shows that Algorithm I is less sensitive to λ than r_1, r_2. Two groups tests are performed to study the impact of parameters r_1, r_2 by two different stopping condition as 50 and 1000 iterations in the first row and second row in Figure 9 respectively. One can see that the parameters r_1, r_2 affect the convergence rates of Algorithm I, and we should prevent selecting the values, in which Algorithm I diverges. For each $\lambda = 100, 200$, or 400, we show the best result among various combinations of r_1, r_2 in Figure 9. One can see that large λ leads to over-smoothed image recovery, and hence a moderate λ should be chosen for the best results. In addition, we observe heuristically that Algorithm II is less sensitive to parameters than Algorithm I, as the TV regularization in Algorithm I introduces a non-differential term, which mysteriously interacts with the non-convex fidelity term. It is helpful to determine optimal parameters automatically, which will be left as a future work.

Convergence. We empirically demonstrate the convergence of the proposed algorithms by plotting objective functional values $\mathcal{G}(u^j)$ and $\mathcal{H}(u^j)$ as well as relative errors $\frac{\|u^j - u^{j-1}\|}{\|u^j\|}$ with respect to the iteration number j in Figure 10. The test image in this example is "Cameraman" as shown in Figure 1 (c) and noise level is at $\eta = 0.1$. We consider both CDP and holographic measurements, each with default parameters in Table 1 and Table 2 respectively. As illustrated in Figure 10 and Figure 11, all the curves are monotonically decreasing, which validate the convergence. In addition, we observe that 30∼50 iterations are sufficient to obtain a satisfactory recovery result from noisy CDP measurements, while more iterations are required for holographic case to converge. Note that the convergence of Algorithm II for noise free measurements is further examined in Figure 14, and please refer more details in Section 5.4.

5.4. Comparison to the state-of-the-art. We compare our proposed algorithms with three PR methods, error reduction (ER) algorithm [24], Wirtinger flow (WF) [10] and truncated Wirtinger flow (TWF) [17]. The Matlab implementation of WF and TWF can be found on authors' website[2], while we implement the ER by ourselves, which consists of the following three steps [15]:

Step 1. Initialize $u^0, z^0 := \mathcal{A}u^0$, which satisfies $|z^0|^2 = b$, and set $k := 0$.

[2]WF: `http://www-bcf.usc.edu/~soltanol/WFcode.html` and TWF: `http://web.stanford.edu/~yxchen/TWF/code.html`

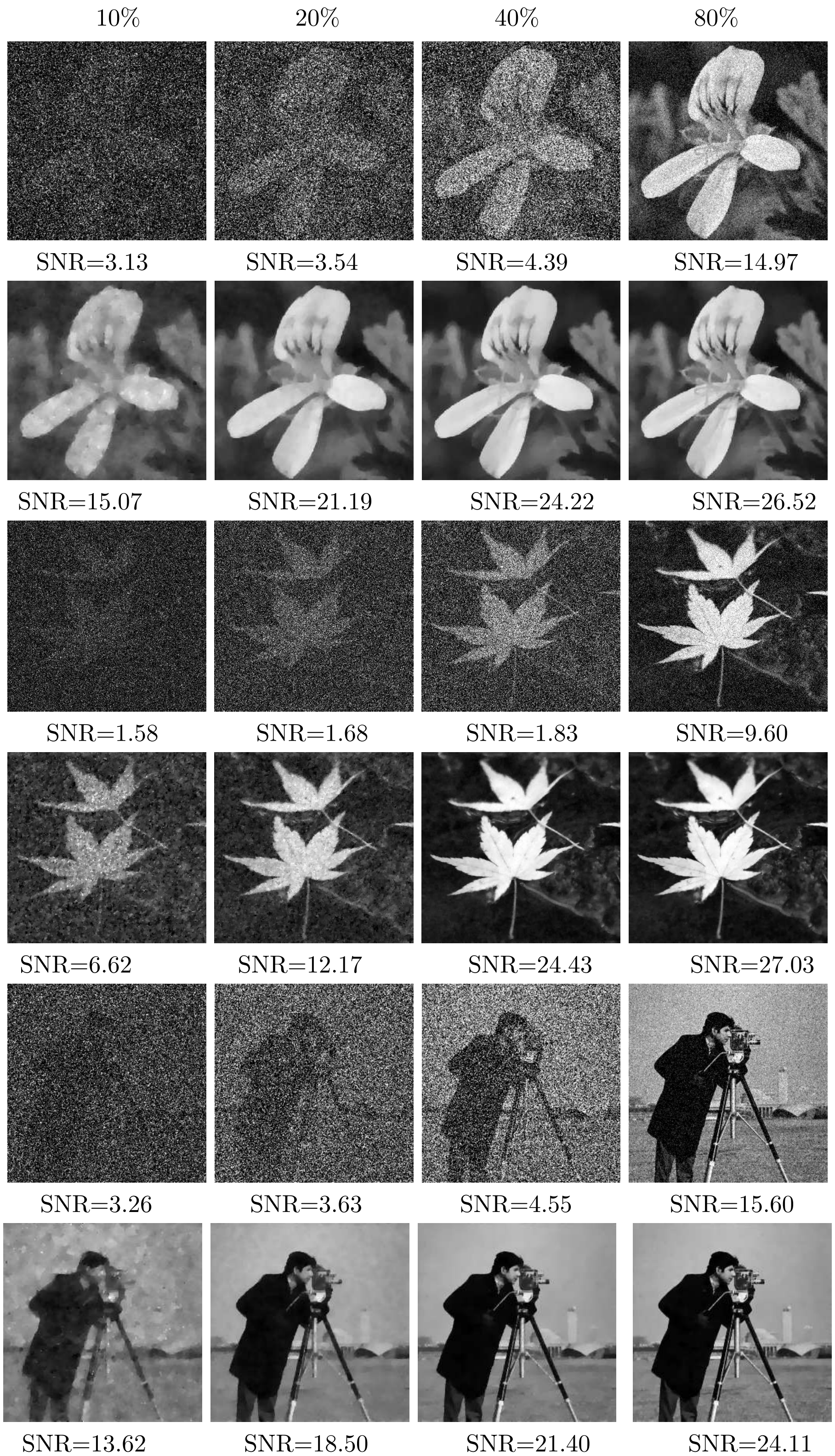


FIG. 7. *Performances with respect to different sampling masks D shown in Figure 6 (using the same examples in Figure 1 and $\eta = 0.05$). 1st, 3rd and 5th rows: results by Algorithm II for PoiPR; 2nd, 4th,and 6th rows: results by Algorithm I for TV-PoiPR. From the left to right: $r_s = 10\%, 20\%, 40\%, 80\%$.*

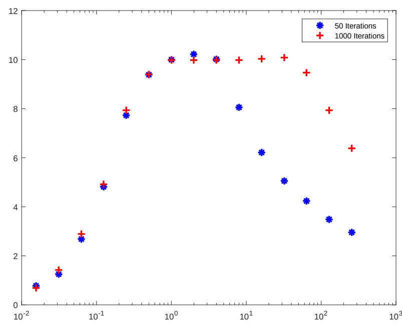


FIG. 8. *The performance of Algorithm II with respect to r stopped after 50 iterations in blue dots and after 1000 iterations in red plus. The y-axis gives the corresponding SNR values.*

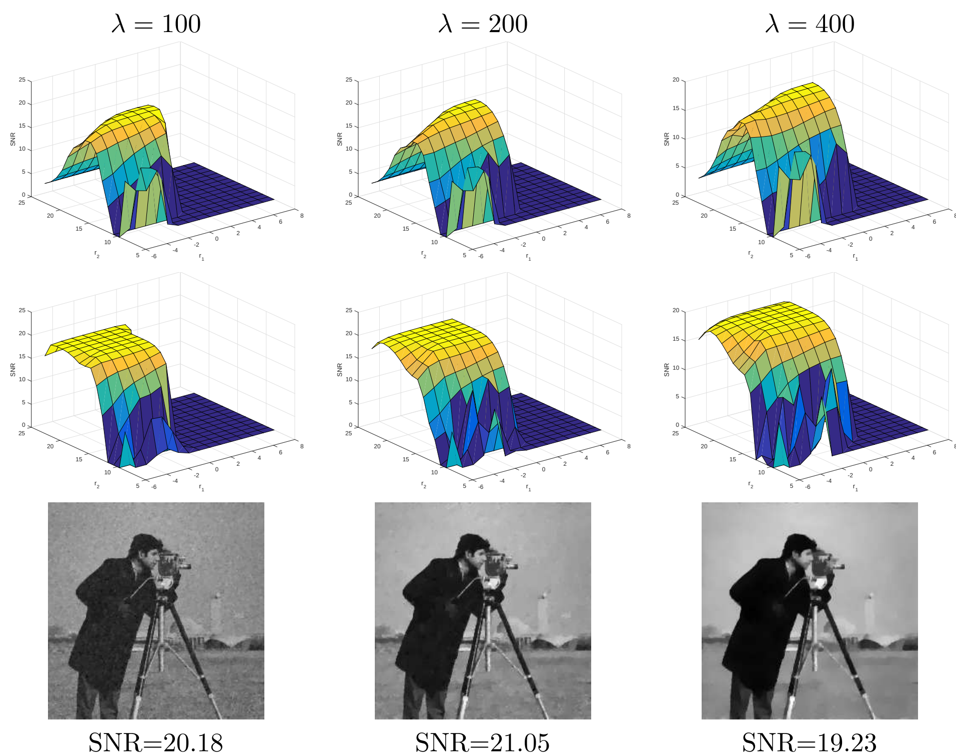


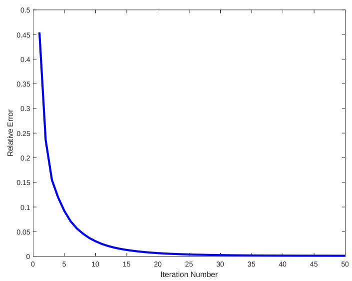
FIG. 9. *The performance of Algorithm I with respect to r_1, r_2 for $\lambda = 100, 200,$ and 400 (from left to right). 1st-2nd row: the SNR values with respect to r_1, r_2 for each λ stopped after 50 iterations in 1st row and 1000 iterations in 2nd row and bottom row: the image recovery results corresponding to the highest SNR.*

Step 2. Update u by solving the following least square problem

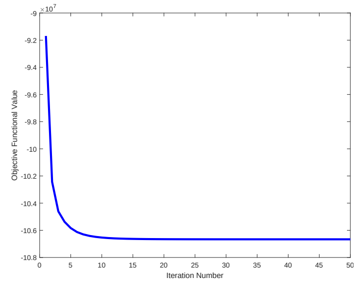
$$u^{k+1} = \arg\min_{u \in \mathbb{R}^n} \|\mathcal{A}u - \tilde{z}^k\|^2,$$

where $\tilde{z}^k(i) = \sqrt{b(i)} z^k(i)/|z^k(i)|, \ 0 \le i \le m-1.$

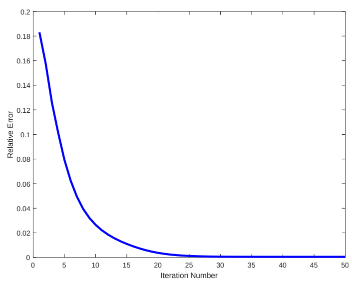
Step 3. Update z as $z^{k+1} = \mathcal{A}u^{k+1}$. If some stopping condition is satisfied, stop and output u^{k+1} as the final result; Otherwise, set $k := k+1$ and goto Step 2.



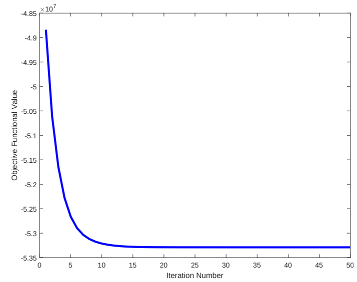
(a) Error of Algorithm II



(b) Objective of Algorithm II

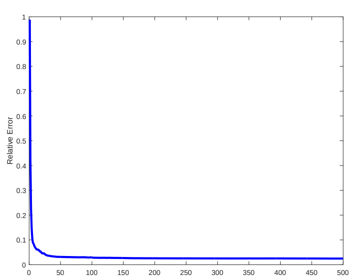


(c) Error of Algorithm I

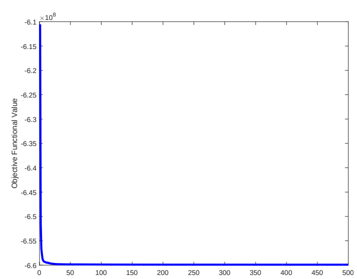


(d) Objective of Algorithm I

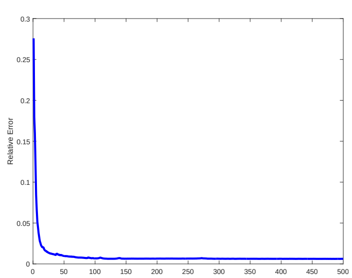
FIG. 10. *Convergence analysis of Algorithm I and Algorithm II for CDP by plotting objective functional values and relative errors to the ground-truth image with respect to the iteration number.*



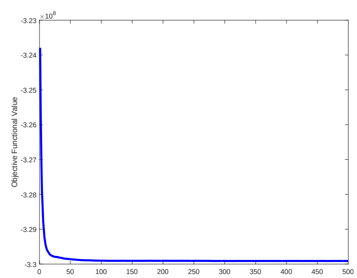
(a) Error of Algorithm II



(b) Objective of Algorithm II

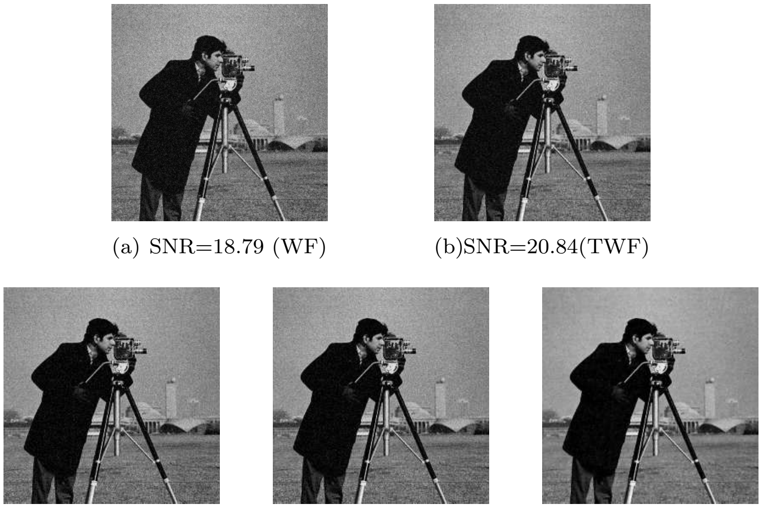


(c) Error of Algorithm I



(d) Objective of Algorithm I

FIG. 11. *Convergence analysis of Algorithm I and Algorithm II for holographic pattern by plotting objective functional values and relative errors to the ground-truth image with respect to the iteration number.*



(a) SNR=18.79 (WF) (b)SNR=20.84(TWF)

(c) SNR=21.08 (ER) (d) SNR=21.41 (II) (e) SNR=25.63 (I)

Fig. 12. *Comparison of noisy phase retrieval from CDP measurements* (3.10) *with more masks* $k = 12$ *with WF in (a), TWF in (b), ER in (c). The proposed methods are labeled by "I" and "II", short for Algorithm I (with TV) and Algorithm II (without TV) respectively.*

(a) SNR=10.48 (ER) (b) SNR=10.73 (II) (c) SNR=21.05 (I)

Fig. 13. *Comparison of noisy phase retrieval from CDP measurements* (3.10) *with less masks* $k = 2$ *with ER only, as neither WF nor TWF gives satisfactory results. The proposed methods are labeled by "I" and "II", short for Algorithm I (with TV) and Algorithm II (without TV) respectively.*

We first test on noisy measurements of CDP (3.10) with the number of masks $k = 12$ and $\eta = 0.02$. As shown in Figure 12, ER, WF and TWF are even worse than the proposed method without TV; and with the help of TV regularization, the improvement is about 4dB in SNR values over the non-TV version. The results of similar setting but less number of masks $k = 2$ are shown in Figure 13, in which WF and TWF are not present as they can not give satisfactory results. In the case of less amount of measurements (2n) and same amount of noise ($\eta = 0.02$), TV shows its effectiveness in noise removal; specifically the SNR value of the TV reconstructed image is almost doubled compared to the ones given by ER and non-TV version. In summary, our proposed methods outperform the state-of-the-art phase retrieval methods.

It is true that ER, WF and TWF are not designed to deal with Poisson noise. So we consider the noise free data of "Cameraman" and only compare the non-TV version (PoiPR and Algorithm II) with ER, WF and TWF methods. When the noise is not present, all the methods are able to find the ground-truth image. We

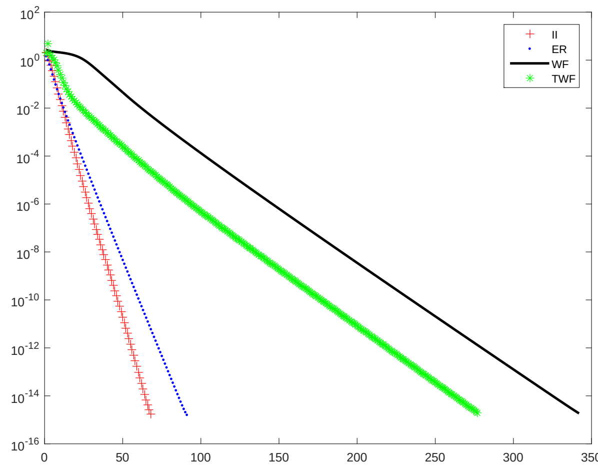


Fig. 14. *Convergence curves of Algorithm II (II in short) in comparison to ER, WF, and TWF for noise free data "Cameraman" with $k = 12$ masks.*

plot the convergence curves in Figure 14, which shows that the proposed Algorithm II converges the fastest among all. In addition, a linear convergence rate is observed practically, while theoretical analysis on convergence rate will be studied in the future. As all these algorithms are of "first-order" convergence, higher-order algorithms [44] will be exploited in the future.

6. Conclusion. In this paper, we proposed a total variation regularization model "TV-PoiPR" to recover an image (taking real or complex values) from its partial and noisy magnitude measurements. We proved that there exists one and only one solution for proposed models. Numerically, an efficient ADMM was designed with guaranteed convergence, which was also validated by numerical experiments. Experiments further demonstrated the effectiveness of our proposed methods over the the state-of-the-art methods. One future direction will be patch based sparse modeling [19, 46, 25] in order to further improve image recovery from magnitude data.

Acknowledgment. The authors would like to thank Prof. Z. Wen in the International Center for Mathematical Research (BICMR), Perking University to provide the complex-valued images "Goldballs" for experiments. Dr. H. Chang is partially supported by National Natural Science Foundation of China (NSFC 11426165 and 11501413) and Tianjin 131 Talent Project. Dr. Y. Lou is partially supported by NSF DMS-1522786.

REFERENCES

[1] Francis J Anscombe, *The transformation of poisson, binomial and negative-binomial data*, Biometrika, 35 (1948), pp. 246–254.

[2] R. Balan, P. Casazza, and D. Edidin, *On signal reconstruction without phase*, Applied and Computational Harmonic Analysis, 20 (2006), pp. 345–356.

[3] H. H. Bauschke, P. L. Combettes, and D. R. Luke, *Phase retrieval, error reduction algorithm, and fienup variants: a view from convex optimization*, J. Opt. Soc. Amer. A, 19 (2002), pp. 1334–1345.

[4] ———, *Hybrid projectioncreflection method for phase retrieval*, J. Opt. Soc. Amer. A, 20 (2003), pp. 1025–1034.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Foundations and Trends® in Machine Learning, 3 (2011), pp. 1–122.
[6] C. Brune, A. Sawatzky, and M. Burger, *Bregman-em-tv methods with application to optical nanoscopy*, in Scale Space and Variational Methods in Computer Vision, Springer, 2009, pp. 235–246.
[7] ———, *Primal and dual bregman methods with application to optical nanoscopy*, International Journal of Computer Vision, 92 (2011), pp. 211–229.
[8] E. J. Candés, Y. C. Eldar, T. Strohmer, and V. Voroninski, *Phase retrieval via matrix completion*, SIAM J. Imaging Sci., 6 (2013), pp. 199–225.
[9] E. J. Candes, X. Li, and M. Soltanolkotabi, *Phase retrieval from coded diffraction patterns*, Applied and Computational Harmonic Analysis, 39 (2015), pp. 277–299.
[10] E. J. Candes, X. Li, and M. Soltanolkotabi, *Phase retrieval via wirtinger flow: Theory and algorithms*, IEEE Trans. Inf. Theory, 61 (2015), pp. 1985–2007.
[11] E. J. Candes, T. Strohmer, and V. Voroninski, *Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming*, Commu. Pure Applied Math., 66 (2013), pp. 1241–1274.
[12] R. Chan, H. Yang, and T. Zeng, *A two-stage image segmentation method for blurry images with poisson or multiplicative gamma noise*, SIAM Journal on Imaging Sciences, 7 (2014), pp. 98–127.
[13] R. H. Chan and K. Chen, *Multilevel algorithm for a poisson noise removal model with total-variation regularization*, International Journal of Computer Mathematics, 84 (2007), p-p. 1183–1198.
[14] R. H. Chan, M. Tao, and X.M. Yuan, *Constrained total variational deblurring models and fast algorithms based on alternating direction method of multipliers*, SIAM J. Imaging Sci., 6 (2013), pp. 680–697.
[15] H. Chang, Y. Lou, M. Ng, and T. Zeng, *Phase retrieval from incomplete magnitude information via total variation regularization*, CAM Report, 16-39 (2016).
[16] P. Chen and A. Fannjiang, *Fourier phase retrieval with a single mask by douglas-rachford algorithm*, arXiv preprint arXiv:1509.00888, (2015).
[17] Y. Chen and E. Candes, *Solving random quadratic systems of equations is nearly as easy as solving linear systems*, in Advances in Neural Information Processing Systems, 2015, pp. 739–747.
[18] A. Conca, D. Edidin, M. Hering, and C. Vinzant, *An algebraic characterization of injectivity in phase retrieval*, Applied and Computational Harmonic Analysis, 38 (2015), pp. 346–356.
[19] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, *Image denoising by sparse 3-d transform-domain collaborative filtering*, Image Processing, IEEE Transactions on, 16 (2007), pp. 2080–2095.
[20] J.C. Dainty and J.R. Fienup, *Phase retrieval and image reconstruction for astronomy*, Image Recovery: Theory and Application, ed. by H. Stark, Academic Press, (1987), pp. 231–275.
[21] V. Elser, *Phase retrieval by iterated projections*, J. Opt. Soc. Am. A, 20 (2003), pp. 40–55.
[22] J. R. Fienup, *Phase retrieval algorithms: a comparison*, Appl. Opt., 21 (1982), pp. 2758–2769.
[23] M. Figueiredo and J. Bioucas-Dias, *Restoration of poissonian images using alternating direction optimization*, Image Processing, IEEE Transactions on, 19 (2010), pp. 3133–3145.
[24] R. W. Gerchberg and W. O. Saxton, *A practical algorithm for the determination of the phase from image and diffraction plane pictures*, Optik, 35 (1972), pp. 237–246.
[25] Raja Giryes and Michael Elad, *Sparsity-based poisson denoising with dictionary learning*, Image Processing, IEEE Transactions on, 23 (2014), pp. 5057–5069.
[26] R. Glowinski and P. L. Tallec, *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*, SIAM Studies in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1989.
[27] R. W. Harrison, *Phase problem in crystallography*, J. Opt. Soc. Am. A, 10 (1993), pp. 1046–1055.
[28] M. Hayes, *The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform*, IEEE Trans. Acoust. Speech Signal Process., 30 (1982), pp. 140–154.
[29] K. Jaganathan, Y. C. Eldar, and B. Hassibi, *Phase retrieval: An overview of recent developments*, arXiv preprint arXiv:1510.07713, (2015).
[30] K. Jaganathan, S. Oymak, and B. Hassibi, *Sparse phase retrieval: Convex algorithms and limitations*, in Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on, IEEE, 2013, pp. 1022–1026.
[31] Xin Jiang, Patricia Reynaud-Bouret, Vincent Rivoirard, Laure Sansonnet, and Re-

becca Willett, *A data-dependent weighted lasso under poisson noise*, arXiv preprint arXiv:1509.08892, (2015).
[32] T. Le, R. Chartrand, and T.J. Asaki, *A variational approach to reconstructing images corrupted by poisson noise*, Journal of mathematical imaging and vision, 27 (2007), pp. 257–263.
[33] X. Li and V. Voroninski, *Sparse signal recovery from quadratic measurements via convex programming*, SIAM Journal on Mathematical Analysis, 45 (2013), pp. 3019–3033.
[34] D. R. Luke, *Relaxed averaged alternating reflections for diffraction imaging*, Inverse Probl., 21 (2005), pp. 37–50.
[35] D. R. Luke, J. V. Burke, and R. G. Lyon, *Optical wavefront reconstruction: Theory and numerical methods*, SIAM review, 44 (2002), pp. 169–224.
[36] M. Lustig, *SparseMRI v0.2*. http://www.eecs.berkeley.edu/~mlustig/Software.html.
[37] M. Lustig, D.L. Donoho, J. M. Santos, and J. M. Pauly, *Compressed sensing mri*, IEEE Signal Processing Magazine, 25 (2008), pp. 72–82.
[38] Markku Mäkitalo and Alessandro Foi, *Optimal inversion of the anscombe transformation in low-count poisson image denoising*, Image Processing, IEEE Transactions on, 20 (2011), pp. 99–109.
[39] S. Marchesini, *Invited article: A unified evaluation of iterative projection algorithms for phase retrieval*, Review of scientific instruments, 78 (2007), p. 011301.
[40] S. Marchesini, Y.-C. Tu, and H.-T. Wu, *Alternating projection, ptychographic imaging and phase synchronization*, Applied and Computational Harmonic Analysis, (2015).
[41] J. Miao, T. Ishikawa, Q. Shen, and T. Earnest, *Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes*, Annu Rev Phys Chem., 59 (2008), pp. 387–410.
[42] P. Netrapalli, P. Jain, and S. Sanghavi, *Phase retrieval using alternating minimization*, Signal Processing, IEEE Transactions on, 63 (2015), pp. 4814–4826.
[43] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, *Cprl–an extension of compressive sensing to the phase retrieval problem*, in Advances in Neural Information Processing Systems, 2012, pp. 1367–1375.
[44] J. Qian, C. Yang, A. Schirotzek, F. Maia, and S Marchesini, *Efficient algorithms for ptychographic phase retrieval*, Inverse Problems and Applications, Contemp. Math, 615 (2014), pp. 261–280.
[45] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation noise removal algorithm*, Phys. D, 60 (1992), pp. 259–268.
[46] Joseph Salmon, Zachary Harmany, Charles-Alban Deledalle, and Rebecca Willett, *Poisson noise reduction with non-local pca*, Journal of mathematical imaging and vision, 48 (2014), pp. 279–294.
[47] Y. Shechtman, A. Beck, and Y. C Eldar, *Gespar: Efficient phase retrieval of sparse signals*, Signal Processing, IEEE Transactions on, 62 (2014), pp. 928–938.
[48] 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*, Signal Processing Magazine, IEEE, 32 (2015), pp. 87–109.
[49] I. Waldspurger, A. Aspremont, and S. Mallat, *Phase recovery, maxcut and complex semidefinite programming*, Math. Program., Ser. A, (DOI 10.1007/s10107-013-0738-9, 2012), pp. 1–35.
[50] A. Walther, *The question of phase retrieval in optics*, Journal of Modern Optics, 10 (1963), pp. 41–49.
[51] Y. Wang and Z. Xu, *Phase retrieval for sparse signals*, Applied and Computational Harmonic Analysis, 37 (2014), pp. 531–544.
[52] YouWei Wen, Raymond Honfu Chan, and TieYong Zeng, *Primal-dual algorithms for total variation based image restoration under poisson noise*, Science China Mathematics, 59 (2016), pp. 141–160.
[53] Z. Wen, C. Yang, X. Liu, and S. Marchesini, *Alternating direction methods for classical and ptychographic phase retrieval*, Inverse Probl., 28 (2012), p. 115010.
[54] C. Wu and X.-C. Tai, *Augmented Lagrangian method, dual methods and split-Bregman iterations for ROF, vectorial TV and higher order models*, SIAM J. Imaging Sci., 3 (2010), pp. 300–339.
[55] P. Yin and J. Xin, *Phaseliftoff: an accurate and stable phase retrieval method based on difference of trace and frobenius norms*, Comm. Math. Sci., 13 (2015), pp. 1033–1049.
[56] Xiaoqun Zhang, Yujie Lu, and Tony Chan, *A novel sparsity reconstruction method from poisson data for 3d bioluminescence tomography*, Journal of scientific computing, 50 (2012), pp. 519–535.