

Point spread function reconstruction in ground-based astronomy by l^1 - l^p model

Raymond H. Chan¹, Xiaoming Yuan^{2,*}, and Wenxing Zhang³

¹*Department of Mathematics, The Chinese University of Hong Kong, Hong Kong, China*

²*Department of Mathematics, Hong Kong Baptist University, Hong Kong, China*

³*School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China*

*Corresponding author: xmyuan@hkbu.edu.hk

In ground-based astronomy, images of objects in outer space are acquired via ground-based telescopes. However, the imaging system is generally interfered by atmospheric turbulence and hence images so acquired are blurred with unknown point spread function. To restore the observed images, the wavefront of light at the telescope's aperture is utilized to derive the point spread function. A model with the Tikhonov regularization has been proposed to find the high-resolution phase gradients by solving a least-squares system. Here we propose the l^1 - l^p ($p = 1, 2$) model for reconstructing the phase gradients. This model can provide sharper edges in the gradients while removing noise. The minimization models can easily be solved by the Douglas-Rachford alternating direction method of multiplier, and the convergence rate is readily established. Numerical results are given to illustrate that the model can give better phase gradients and hence a more accurate point spread function. As a result, the restored images are much more accurate when compared to the traditional Tikhonov regularization model. © 2012 Optical Society of America

OCIS codes: 100.1830, 100.3020, 100.5070, 110.4155

1. Introduction

In ground-based astronomy, images of objects in outer space are acquired via ground-based telescopes. However, the imaging system is generally interfered by atmospheric turbulence

and the resulting images are usually blurred. The mathematical model of the imaging system in ground-based astronomy is

$$g(x, y) = k(x, y) * f(x, y) + \varepsilon_g(x, y), \quad (1)$$

where $f(x, y)$ is the true object in outer space, $g(x, y)$ is its observation by the ground-based telescope, $k(x, y)$ is the *point spread function* (PSF) related to the atmospheric turbulence, ‘*’ is the convolutional operator, and $\varepsilon_g(x, y)$ is white noise. When $k(x, y)$ is known, it is well known that the problem of finding $f(x, y)$ from $g(x, y)$ is an ill-posed inverse problem (see e.g., [1]). One can invert (1) with a suitable regularization technique to improve the quality of the observed image $g(x, y)$, see e.g., [1–3].

However, in ground-based astronomy, $k(x, y)$ is usually unknown as it is generated by the atmospheric turbulence. The problem thus becomes even more ill-posed. One may consider using blind-deconvolution [4, 5] to simultaneously obtain $k(x, y)$ and $f(x, y)$. But that usually requires some prior knowledge on $k(x, y)$, e.g., smoothness, symmetry or sparsity; and they usually do not hold for atmospheric blurring. The Fourier optics model [6, 7] is an approach to directly approximate $k(x, y)$. It expresses $k(x, y)$ as a function of the incoming wavefronts of light:

$$k(x, y) = |\mathcal{F}^{-1}\{\mathcal{P}(x, y) \exp[\iota(1 - \omega(x, y))]\}|^2 = |\mathcal{F}^{-1}\{\mathcal{P}(x, y) \exp[\iota\phi(x, y)]\}|^2, \quad (2)$$

where \mathcal{F} denotes the Fourier transform and \mathcal{F}^{-1} is its inverse; $\mathcal{P}(x, y)$ is the mirror aperture function of the telescope ($= 1$ inside the mirror aperture and 0 otherwise); $\iota = \sqrt{-1}$; $\omega(x, y)$ is the wavefront of incoming light at the telescope; $\phi(x, y) = 1 - \omega(x, y)$ is the *phase* term which measures the deviation from the planarity of the wavefront. Specially, if $\phi(x, y) = 0$ at any (x, y) , we have the so-called *diffraction limited imaging*, i.e.,

$$k(x, y) = |\mathcal{F}^{-1}\{\mathcal{P}(x, y)\}|^2, \quad (3)$$

which represents the case where the PSF is merely related to the mirror aperture \mathcal{P} but is independent of the phase ϕ . This is the ideal case where the imaging system has no distortion from atmospheric turbulence, and is the goal of astronomers (see e.g., [6]).

However, for many applications, the incoming wavefronts are disturbed by atmospheric turbulence which limits the performance of ground-based astronomy. Commonly, astronomers quantify the deterioration in seeing condition by the ratio d/r_0 , where d is the diameter of the telescope and r_0 is the Fried parameter (see [7]). Generally, good seeing conditions correspond to d/r_0 with small magnitudes (e.g., $d/r_0 \lesssim 10$) and poor seeing conditions correspond to d/r_0 with large magnitudes (e.g., $d/r_0 \gtrsim 20$). When the wavefronts are disturbed, ϕ is no longer zero, and one has to determine ϕ in order to get the PSF $k(x, y)$ in (2). Wavefront

sensor (WFS) is a device that measures wavefront aberrations caused by atmospheric turbulence, i.e., the gradients of phase ϕ at the telescope, denoted by \mathbf{s}_x (horizontal direction) and \mathbf{s}_y (vertical direction), can be derived by WFS. Once the gradients are obtained, then the following linear inverse problem can be solved to obtain the phase ϕ (see [6, 8, 9]):

$$\begin{bmatrix} \mathbf{s}_x \\ \mathbf{s}_y \end{bmatrix} = \begin{bmatrix} D_x \\ D_y \end{bmatrix} \phi + \varepsilon_s. \quad (4)$$

Here D_x, D_y are the derivative operators in the horizontal and vertical directions respectively, and ε_s is the perturbation in the gradient measurements where it is usually modeled as white noise. Two categories of derivative operators can be used for D_x and D_y . One is the Fried geometry [10] and the other is the Hudgin geometry [11]. To simplify the discussions, we will concentrate on the Fried geometry though numerical examples in Section 4 will show that our method works equally well for the Hudgin geometry.

The main obstacle in solving problem (4) is that the phase gradients \mathbf{s}_x and \mathbf{s}_y are generally derived by WFS under the coarse grid. These low-resolution phase gradients so obtained contain relatively less information, therefore it is difficult to reconstruct an accurate phase ϕ . The inadequacies on the phase ϕ affect the accuracy in estimating the PSF k and hence in reconstructing the image f .

In [12], under the frozen flow hypothesis (FFH) of the atmosphere, a method using multiple frames for reconstructing high-resolution phase gradients was proposed. Briefly speaking, some low-resolution frames of the phase gradients within a short time period are first collected. In this way, we obtain the low-resolution phase gradients at different overlapping locations. Then by aligning these low-resolution frames, the high-resolution phase gradients are reconstructed at these locations. Using those reconstructed high-resolution phase gradients, an accurate phase ϕ can be obtained by solving model (4), and hence, the more precise PSF k can be obtained by formula (2).

In [12], the movement of the atmospheric turbulence above the telescope are assumed to be of linear velocity. Nagy et al. [13] have extended that to nonlinear velocity. They acquired the high-resolution phase gradients by using the Tikhonov regularization model and solved it as a least-squares problem. The relationship between the multiple frames of low-resolutions phase gradients and the high-resolution phase gradients is given by the following equations:

$$s_x^i = RWA_i \mathbf{s}_x + \mathbf{n}_x^i \quad \text{and} \quad s_y^i = RWA_i \mathbf{s}_y + \mathbf{n}_y^i, \quad i = 1, 2, \dots, m, \quad (5)$$

where $A_i \in \mathbb{R}^{n \times n}$, $i = 1, 2, \dots, m$, are motion matrices which describe the shift (linear constant velocity or nonlinear velocity between the low-resolution frames) of the phase gradients in the i -th frame; $W \in \mathbb{R}^{n \times n}$ is an indicator matrix which represents the telescope aperture; $R \in \mathbb{R}^{l \times n}$ is a down-sampling matrix which transforms high-resolution phase gradients into

low-resolutions ones; s_x^i and s_y^i , $i = 1, 2, \dots, m$, are the two sequences of low-resolution phase gradients; and n_x^i and n_y^i are additive white noise in each low-resolution frame.

Problem (5) is similar to that of multi-frame super-resolution in the realm of image reconstruction, see e.g., [14–17]. The Tikhonov regularization model was considered in [13] to derive the high-resolution phase gradients by solving the following linear least-squares problems:

$$\min_{\mathbf{s}} \alpha \|\mathbf{s}\|_2^2 + \frac{1}{2} \sum_{i=1}^m \|RWA_i \mathbf{s} - d_i\|_2^2. \quad (6)$$

where $\mathbf{s} \in \mathbb{R}^n$ is the high-resolution phase gradient (either \mathbf{s}_x or \mathbf{s}_y); and $d_i \in \mathbb{R}^l$, $i = 1, 2, \dots, m$, are the sequence of low-resolution phase gradients (s_x^i or s_y^i accordingly). Model (6) uses the Tikhonov regularization to generate phase gradients in both horizontal and vertical directions. It is well known that the Tikhonov regularization smoothes the edges in the resulting images, see e.g., [1]. However, the phase gradients are not smooth in general (see for example the phase gradients of a typical ϕ in Fig. 1). Thus in this paper, instead of using model (6), we consider the l^1-l^p models ($p = 1, 2$) for reconstructing them. More precisely, our mathematical model for finding the high-resolution phase gradients in (5) is

$$\min_{\mathbf{s}} \alpha \|\mathbf{s}\|_1 + \frac{1}{p} \sum_{i=1}^m \|RWA_i \mathbf{s} - d_i\|_p^p, \quad p = 1, 2. \quad (7)$$

The l^1 -regularization term (the first term) is known to be able to sharpen the edges in the solution, see e.g., [1, 2]. The p -norm in the fidelity term can cater for different possible noise in the measurements. When collecting the low-resolution phase gradients d_i by the WFS, the measurements inevitably possess readout noise which contaminates the phase gradients measurements. It is well known in image denoising (see e.g., [2, 18, 19]) that the p -norm in the fidelity term can handle impulsive or Laplacian noise when $p = 1$ or Gaussian noise when $p = 2$. The numerical results in Section 4 show that the phase ϕ and the PSF k obtained by model (7) are more accurate than those by model (6). Hence the deblurred images are 1 to 1.6 dB higher in signal-to-noise ratio (SNR) than those by model (6). The SNR is defined as $\text{SNR} := 20 \log_{10}(\|x\|_2 / \|\bar{x} - x\|_2)$, where \bar{x} is the restored image and x is the true image.

From the optimization point of view, model (7) is harder to solve than model (6) due to the nonsmoothness of its l^1 -regularization term. We will show that the nonsmooth model (7) can be tackled efficiently by the Douglas-Rachford alternating direction method of multipliers (ADM for short) proposed in [20, 21]. In fact, by introducing additional variables, model (7) can be reformulated as an optimization problem with separable structures (i.e., both the objective function and linear constraint are separable, see problem (8) in Section 2), and hence can be solved by ADM. When ADM is applied to solve (7), the computation at each iteration is dominated by a shrinkage operation and solving a system of linear

equations. Moreover, by invoking standard theory in [20–22], the convergence of ADM can be easily proven; and by [23], a worst-case $O(1/n)$ convergence rate of ADM can be established straightforwardly.

The rest of the paper is organized as follows. In Section 2, we review briefly the background of ADM. In Section 3, we elaborate on the details of solving model (7) by ADM, for both $p = 1$ and $p = 2$. Then, in Section 4 we compare the Tikhonov regularization model (6) with our l^1-l^p model (7). Numerical results show that the phase gradients obtained by our models are better than those by model (6), and hence the images recovered are better too. Some conclusions are given in Section 5.

2. Preliminaries

Our method of solving the model (7) is by ADM. In this section, we recall ADM in solving structured optimization problem, specifically, when the objective function is a sum of two convex functions. Consider the following optimization problem with special structure

$$\begin{aligned} \min \quad & \theta_1(x) + \theta_2(y) \\ \text{s.t.} \quad & Bx + Cy = b, \quad x \in \mathcal{X}, \quad y \in \mathcal{Y}, \end{aligned} \tag{8}$$

where $\theta_1 : \mathbb{R}^{n_1} \rightarrow (-\infty, +\infty]$ and $\theta_2 : \mathbb{R}^{n_2} \rightarrow (-\infty, +\infty]$ are proper convex functions; $B \in \mathbb{R}^{l \times n_1}$ and $C \in \mathbb{R}^{l \times n_2}$ are given matrices; $b \in \mathbb{R}^l$ is a known vector; and $\mathcal{X} \subseteq \mathbb{R}^{n_1}$ and $\mathcal{Y} \subseteq \mathbb{R}^{n_2}$ are nonempty closed convex sets. Then, the augmented Lagrangian function of problem (8) on $\mathcal{X} \times \mathcal{Y} \times \mathbb{R}^l$ is

$$\mathcal{L}(x, y, \lambda) = \theta_1(x) + \theta_2(y) - \langle \lambda, Bx + Cy - b \rangle + \frac{\beta}{2} \|Bx + Cy - b\|_2^2,$$

where $\lambda \in \mathbb{R}^l$ is the Lagrange multiplier and $\beta > 0$ is a penalty parameter on the linear constraint, see e.g., [24]. The iterative scheme of ADM proposed in [20, 21] is summarized in Algorithm 1 below.

Algorithm 1 : Douglas-Rachford alternating direction method of multipliers

Choose arbitrary $\beta > 0$, $y^0 \in \mathbb{R}^{n_2}$ and $\lambda^0 \in \mathbb{R}^l$

for $k = 1, 2, \dots$ **do**

- $x^{k+1} = \arg \min_x \mathcal{L}(x, y^k, \lambda^k)$
 - $y^{k+1} = \arg \min_y \mathcal{L}(x^{k+1}, y, \lambda^k)$
 - $\lambda^{k+1} = \lambda^k - \beta (Bx^{k+1} + Cy^{k+1} - b)$
-

In [20–22], it was guaranteed that the ADM converges under the condition that $\beta > 0$. Recently, a worst-case $O(1/n)$ convergence rate of ADM was established in [23]. We refer to [25] and references therein for numerous practical applications of ADM in diversified areas. In particular, we refer to [26–28] for some applications of ADM in image processing.

3. Solving the l^1 - l^p Model by ADM

We now elaborate on how to use ADM to reconstruct the high-resolution phase gradients using the l^1 - l^p model (7). Let $I_m \in \mathbb{R}^{m \times m}$ be the identity matrix and “ \otimes ” denote the Kronecker product of matrices. Let $\mathbf{R} = I_m \otimes R$ and $\mathbf{W} = I_m \otimes W$. We stack the motion matrices A_i , $i = 1, 2, \dots, m$, in model (7) as $\mathbf{A} = [A_1; A_2; \dots; A_m] \in \mathbb{R}^{mn \times n}$. The m frames of the low-resolution phase gradients d_i , $i = 1, 2, \dots, m$, are rearranged as $\mathbf{d} = [d_1; d_2; \dots; d_m] \in \mathbb{R}^{lm}$. The model (7) therefore can be reformulated as

$$\min_{\mathbf{s}} \alpha \|\mathbf{s}\|_1 + \frac{1}{p} \|\mathbf{RWAs} - \mathbf{d}\|_p^p, \quad p = 1, 2. \quad (9)$$

3.A. The case of $p = 2$

When $p = 2$, the model (9) can be reformulated as a constrained optimization problem with separable structure by introducing an additional variable $\mathbf{v} \in \mathbb{R}^n$:

$$\begin{aligned} \min_{\mathbf{s}, \mathbf{v}} & \alpha \|\mathbf{s}\|_1 + \frac{1}{2} \|\mathbf{RWAv} - \mathbf{d}\|_2^2 \\ \text{s.t. } & \mathbf{s} = \mathbf{v}. \end{aligned} \quad (10)$$

By defining $(x, y) := (\mathbf{s}, \mathbf{v})$, $\theta_1(x) := \alpha \|\mathbf{s}\|_1$, $\theta_2(y) := \frac{1}{2} \|\mathbf{RWAv} - \mathbf{d}\|_2^2$, $B := I$, $C := -I$ and $b := 0$, we see that (10) is the same as the optimization problem (8). Consequently, ADM can be used to solve it. More concretely, the augmented Lagrangian function of (10) is

$$\mathcal{L}(\mathbf{s}, \mathbf{v}, \lambda) = \alpha \|\mathbf{s}\|_1 + \frac{1}{2} \|\mathbf{RWAv} - \mathbf{d}\|_2^2 - \langle \lambda, \mathbf{s} - \mathbf{v} \rangle + \frac{\beta}{2} \|\mathbf{s} - \mathbf{v}\|_2^2,$$

where $\lambda \in \mathbb{R}^n$ is the Lagrange multiplier and $\beta > 0$ is the penalty parameter. According to Algorithm 1, the \mathbf{s} - and \mathbf{v} -related subproblems can be re-written as follows:

- The \mathbf{s} -related subproblem is

$$\mathbf{s}^{k+1} = \arg \min_{\mathbf{s}} \alpha \|\mathbf{s}\|_1 + \frac{\beta}{2} \left\| \mathbf{s} - \mathbf{v}^k - \frac{\lambda^k}{\beta} \right\|_2^2,$$

and its closed-form solution can be obtained by the well-known shrinkage (soft-threshold) operator [29], i.e., $\mathbf{s}^{k+1} = \text{shrink}_{\alpha/\beta}(\mathbf{v}^k + \lambda^k/\beta)$.

- The \mathbf{v} -related subproblem is

$$\mathbf{v}^{k+1} = \arg \min_{\mathbf{v}} \frac{1}{2} \|\mathbf{RWAv} - \mathbf{d}\|_2^2 - \langle \lambda^k, \mathbf{s}^{k+1} - \mathbf{v} \rangle + \frac{\beta}{2} \|\mathbf{s}^{k+1} - \mathbf{v}\|_2^2,$$

which is a least-squares problem. It is equivalent to solving the following linear system

$$[\beta I + (\mathbf{RWA})^T \mathbf{RWA}] \mathbf{v}^{k+1} = (\beta \mathbf{s}^{k+1} - \lambda^k) + (\mathbf{RWA})^T \mathbf{d}. \quad (11)$$

Since the coefficient matrix of linear system (11) is positive definite, the linear system can be solved by the conjugate gradient method, see, e.g., [30].

3.B. The case of $p = 1$

When $p = 1$, model (9) can be reformulated as the following constrained optimization problem by introducing two additional variables \mathbf{u} and \mathbf{v} :

$$\begin{aligned} \min \quad & \alpha \|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 \\ \text{s.t.} \quad & \mathbf{s} - \mathbf{u} = 0 \text{ and } \mathbf{RWAs} - \mathbf{v} = \mathbf{d}. \end{aligned} \quad (12)$$

By defining $(x, y) := (\mathbf{s}, (\mathbf{u}, \mathbf{v}))$, $\theta_1(x) := 0$ and $\theta_2(y) := \alpha \|\mathbf{u}\|_1 + \|\mathbf{v}\|_1$,

$$B := \begin{pmatrix} I \\ \mathbf{RWAs} \end{pmatrix}, C := \begin{pmatrix} -I & 0 \\ 0 & -I \end{pmatrix} \text{ and } b := \begin{pmatrix} 0 \\ \mathbf{d} \end{pmatrix},$$

we see that (12) is of the form of (8). Consequently, we can still use ADM to solve it. The augmented Lagrangian function of (12) is

$$\mathcal{L}(\mathbf{s}, \mathbf{u}, \mathbf{v}, \lambda_1, \lambda_2) = \alpha \|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 + \frac{\beta_1}{2} \left\| \mathbf{u} - \mathbf{s} - \frac{\lambda_1}{\beta_1} \right\|_2^2 + \frac{\beta_2}{2} \left\| \mathbf{RWAs} - \mathbf{v} - \mathbf{d} - \frac{\lambda_2}{\beta_2} \right\|_2^2,$$

where $(\lambda_1, \lambda_2) \in \mathbb{R}^n \times \mathbb{R}^{lm}$ is the Lagrange multiplier and $\beta_1 > 0$, $\beta_2 > 0$ are the penalty parameters. According to Algorithm 1, the \mathbf{s} - and (\mathbf{u}, \mathbf{v}) -related subproblems are as follows:

- The \mathbf{s} -related subproblem is

$$\mathbf{s}^{k+1} = \arg \min_{\mathbf{s}} \beta_1 \left\| \mathbf{s} - \mathbf{u}^k + \frac{\lambda_1^k}{\beta_1} \right\|_2^2 + \beta_2 \left\| \mathbf{RWAs} - \mathbf{v}^k - \mathbf{d} - \frac{\lambda_2^k}{\beta_2} \right\|_2^2,$$

which is a least-squares problem. It is equivalent to the following linear system

$$[\beta_1 I + \beta_2 (\mathbf{RWAs})^T \mathbf{RWAs}] \mathbf{s}^{k+1} = (\beta_1 \mathbf{u}^k - \lambda_1^k) + (\mathbf{RWAs})^T [\beta_2 (\mathbf{v}^k + \mathbf{d}) + \lambda_2^k]. \quad (13)$$

Since the coefficient matrix of the linear system (13) is positive definite, the linear system can be solved easily as in (11) by using the conjugate gradient method.

- The (\mathbf{u}, \mathbf{v}) -related subproblem is

$$\begin{aligned} (\mathbf{u}^{k+1}, \mathbf{v}^{k+1}) = & \arg \min_{\mathbf{u}, \mathbf{v}} \alpha \|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 + \frac{\beta_1}{2} \left\| \mathbf{u} - \mathbf{s}^{k+1} - \frac{\lambda_1^k}{\beta_1} \right\|_2^2 \\ & + \frac{\beta_2}{2} \left\| \mathbf{v} - \mathbf{RWAs}^{k+1} + \mathbf{d} + \frac{\lambda_2^k}{\beta_2} \right\|_2^2, \end{aligned}$$

and the \mathbf{u} -, \mathbf{v} -related subproblems can be solved parallelly as follows:

- The \mathbf{u} -related subproblem is:

$$\mathbf{u}^{k+1} = \arg \min_{\mathbf{u}} \alpha \|\mathbf{u}\|_1 + \frac{\beta_1}{2} \left\| \mathbf{u} - \mathbf{s}^{k+1} - \frac{\lambda_1^k}{\beta_1} \right\|_2^2.$$

Hence, its closed-form solution can be obtained by the shrinkage operator, i.e., $\mathbf{u}^{k+1} = \text{shrink}_{\alpha/\beta_1} (\mathbf{s}^{k+1} + \lambda_1^k / \beta_1)$.

(ii) The \mathbf{v} -related subproblem is:

$$\mathbf{v}^{k+1} = \arg \min_{\mathbf{v}} \|\mathbf{v}\|_1 + \frac{\beta_2}{2} \left\| \mathbf{v} - \mathbf{RWAs}^{k+1} + \mathbf{d} + \frac{\lambda_2^k}{\beta_2} \right\|_2^2, \quad (14)$$

which can also be solved by the shrinkage operator as

$$\mathbf{v}^{k+1} = \text{shrink}_{\frac{1}{\beta_2}} \left(\mathbf{RWAs}^{k+1} - \mathbf{d} - \frac{\lambda_2^k}{\beta_2} \right).$$

By the theory in [20–23], our ADM schemes for the l^1-l^p models, $p = 1, 2$, are convergent and the convergent rate is at least $O(1/n)$.

3.C. The choice of parameters

For all the three models we consider here, i.e. the l^2-l^2 model in (6) and the l^1-l^p , $p = 1, 2$ models in (7), there is a regularization parameter α balancing the data-fidelity term and the regularization term. For the l^2-l^2 model, which is equivalent to a least-squares problem, there are well-known methods for choosing a reasonable α , e.g. the generalized cross validation method [31] and the discrepancy principle [32]. For the l^1-l^p models, how to choose a good α is still an on-going research, e.g. the recent paper [33] can give a good α for l^1-l^2 model provided that the level of the white noise can be estimated. Note that models (6) and (7) are models for solving the phase gradients, and hence the best α obtained for one set of phase gradients under one environment should be a good estimate of α for another set of phase gradients under a similar environment. Our experience shows that the choice of α is less relevant to the objects in outer space and the seeing conditions of atmospheric turbulence.

For our l^1-l^p models, we have the additional penalty parameters β 's from ADM. Theoretically any positive values ensure the convergence of ADM, and we usually have two ways to determine them. One is to try some values first and pick up a value with satisfactory performance, and then use it throughout the iterations. The other is to apply some self-adaptive adjustment rules (e.g., [34, 35]) with an arbitrary initial guess. The latter strategy requires no tuning but may require expensive computation for imaging applications. We recommend therefore the former strategy. Our experience is that using well-tuned constant values performs almost the same as using the self-adaptive strategy.

4. Numerical Experiments

In this section, we test the accuracy of the three models: the l^2-l^2 model (6) proposed in [13], and our l^1-l^2 , l^1-l^1 model in (7). We also test the robustness of our models with respect to the seeing conditions [7] and the choice of parameters α and β 's. We first generate a 256-by-256 true phase ϕ by the method proposed in [6], and as in [36] we set the seeing condition $d/r_0 = 50$ where d is the diameter of the telescope and r_0 is the Fried parameter. Then its x

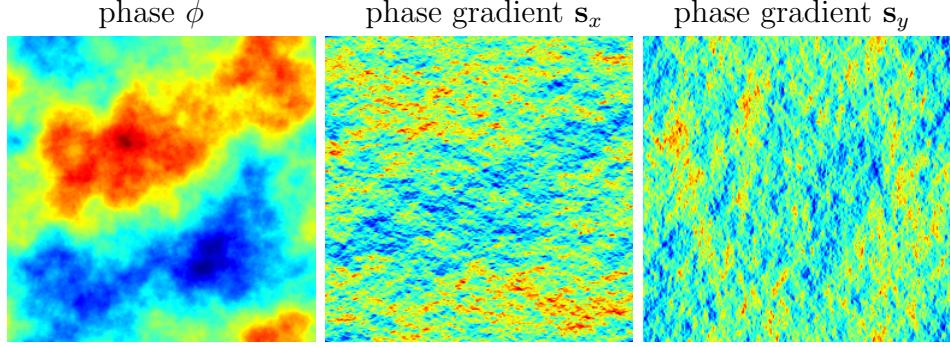


Fig. 1. True phase and its gradients.

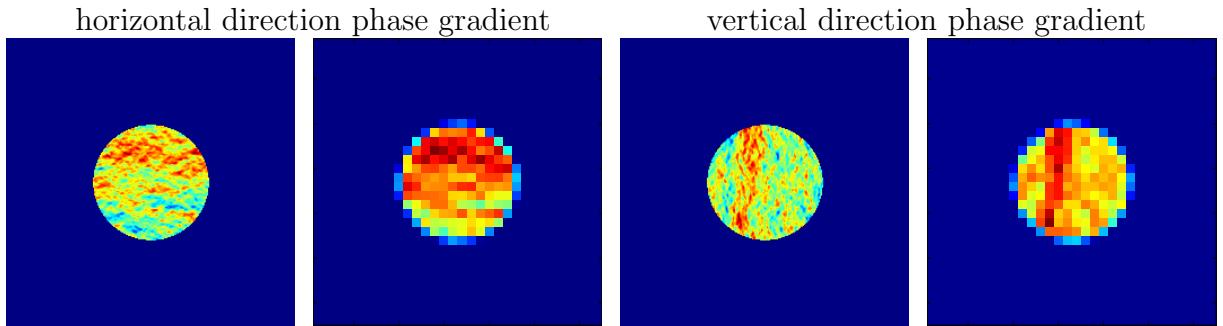


Fig. 2. The first frame high-resolution (256-by-256) and low-resolution (32-by-32) phase gradients within the telescope aperture.

and y -direction gradients \mathbf{s}_x and \mathbf{s}_y are derived by formula (4) using the Fried geometry [10]. White noise ε_s with noise variance 0.05 is then added to both phase gradients. The true phase ϕ and its x - and y -direction gradients are given in Fig. 1.

In the next two subsections, we separate the tests into two parts according to the velocity of the moving frames: linear constant velocity and nonlinear velocity. We downsample the phase gradients by a factor of 8 to generate the low-resolution phase gradients, i.e., s_x^i and s_y^i , $i = 1, 2, \dots, m$, are of size 32-by-32. The number of frames m of low-resolution phase gradients is set to 16. Furthermore, these frames are restricted by mirror aperture of the telescope. Figure 2 depicts one such frame for the high- and also for the low-resolution phase gradients.

4.4. Linear constant velocity

We first consider the scenario that the low-resolution frames move under linear constant velocity. As aforementioned, we generate sixteen 32-by-32 frames of the low-resolution phase gradients in both horizontal and vertical directions. The diameter of the mirror aperture of the telescope is set to 0.4. By solving model (7), we aim at reconstructing the high-resolution phase gradient \mathbf{s}_x and \mathbf{s}_y based on the information of the sixteen low-resolution frames. We choose the regularization parameter α and the penalty parameter β manually so as to maximize the SNR of the respective models. The values are $\alpha = 5 \times 10^{-4}$ in model (6), $\alpha = 10^{-4}$ in model (7); and $\beta = 10^{-3}$ for the l^1-l^2 model and $\beta_1 = 10^{-4}$, $\beta_2 = 0.1$ for the l^1-l^1 model. We employ “pcg” in MATLAB with the tolerance 10^{-4} to solve the model (6) and the subproblems (11) and (13) in ADM. The stopping rules for ADM is $\|\mathbf{s}^{k+1} - \mathbf{s}^k\|/(1 + \|\mathbf{s}^k\|) < 5 \times 10^{-2}$. The reconstructed phase gradients \mathbf{s}_x and \mathbf{s}_y by the three models are given in Fig. 3.

In order to illustrate the accuracy of the reconstructed phase gradients, we also compute the corresponding PSF’s from them and compare them with the true PSF. The process is as follows: for each frame i , $1 \leq i \leq m$, we first compute the high-resolution phase gradients corresponding to this frame by applying the mirror aperture W and motion operator A_i (see formula (5)) onto \mathbf{s}_x and \mathbf{s}_y . For example, the horizontal phase gradient on frame i is given by $\mathbf{s}_x^i = WA_i\mathbf{s}_x$. Then we solve the inverse problem (4) by the algorithm in [6] to obtain the reconstructed phase ϕ corresponding to this frame. Finally, by substituting the ϕ into (2), we get the PSF corresponding to the i -th frame. The reconstructed PSF’s from the 8-th frame for the three models and the true PSF are given in the bottom row of Fig. 3. In Fig. 4, we plot the relative error of the reconstructed PSF’s (measured in 2-norm) for all sixteen frames by the three models. We see that our models give more accurate PSF’s than the l^2-l^2 model. Notice that the errors from the l^2-l^2 model vary from 16% to 43% whereas the errors in our models hover around 10% to 14%. It shows that our models can give more accurate PSF and is more robust to the motion of atmospheric turbulence.

To see whether these PSF’s lead to more accurate restored images, we use them in deblurring an observed “Satellite” image. The observed “Satellite” image is obtained as follows: first we blur the 256-by-256 clean “Satellite” image by the true PSF in Fig. 3 (bottom right column). Then we add to the blurred image a zero-mean Gaussian noise with variance 0.003. The degraded image with SNR 3.84dB is depicted in Fig. 5 (leftmost one). We restored the degraded image by Algorithm 4 in [37] using the reconstructed PSF’s in Fig. 3. The restorations are given in Fig. 5. It shows that the images obtained by our l^1-l^p model has a gain of 0.9dB in SNR over the image from the l^2-l^2 model. It indicates that our models are better in reconstructing the high-resolution phase gradients.

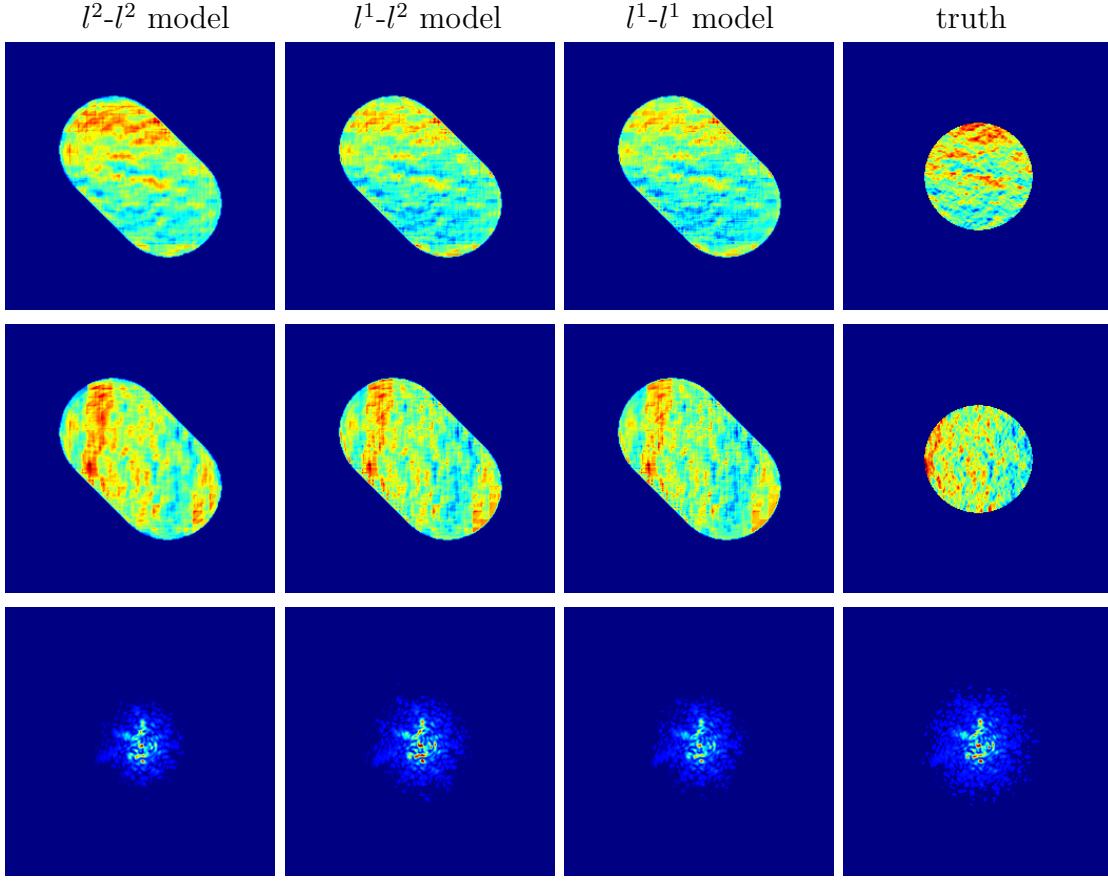


Fig. 3. Reconstructed horizontal (top row), vertical (center row) direction phase gradients and the PSF from the 8th frame (bottom row).

4.B. Nonlinear velocity

We repeat the test for the scenario that the low-resolution phase gradients is moving under nonlinear velocity. Similarly, we generate sixteen 32-by-32 frames of the low-resolution phase gradients. The diameter of the mirror aperture of the telescope is set to 0.28. Again, we choose the regularization parameter and the penalty parameters manually to maximize the SNR. In particular, we choose $\alpha = 5 \times 10^{-4}$ in model (6) and $\alpha = 10^{-4}$ in model (7); $\beta = 10^{-4}$ for the l^1-l^2 model and $\beta_1 = 10^{-3}$, $\beta_2 = 10^{-1}$ for the l^1-l^1 model. We solve the models as in Section 4.A. The reconstructed phase gradients by the three models are given in Fig. 6. We also reconstruct the PSF's by using the 8-th frame of the reconstructed phase and they are given in the bottom row of Fig. 6. Figure 7 plots the relative error of the reconstructed PSF's by the three models. Again, we see that our models give smaller errors. As in Section 4.A, the reconstructed PSF's are used for deblurring the observed ‘‘Satellite’’ image. The restored

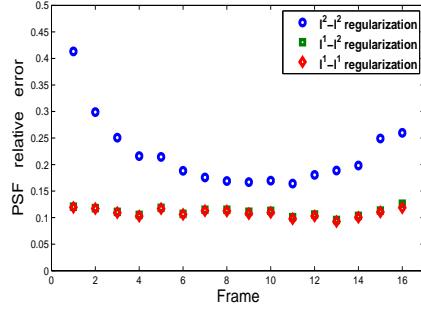


Fig. 4. The relative errors for all reconstructed PSF's.

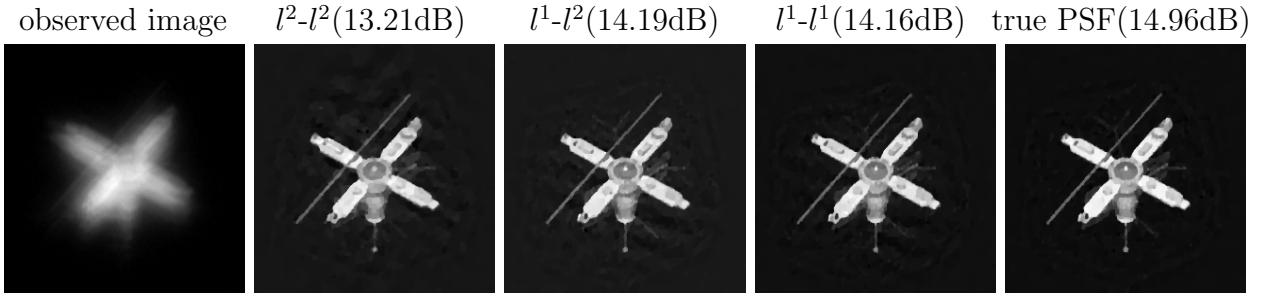


Fig. 5. Image restoration using PSF reconstructed by different models.

images are given in Fig. 8. Here we see that the PSF's from our models are also better than the l^2-l^2 model and the restored images are 1.6 dB higher in SNR than that from the l^2-l^2 model.

4.C. Hudgin geometry

Though the discussions so far have been on Fried geometry [10], we emphasize that the proposed models and algorithms are also valid to wavefront reconstruction under the Hudgin geometry [11]. We illustrate this by using numerical results for the scenario under linear constant velocity. Figure 9 gives the reconstructed phase gradients and the 8-th frame PSF's under the Hudgin geometry. The best parameters chosen manually are $\alpha = 3 \times 10^{-3}$ in model (6) and $\alpha = 10^{-4}$ in model (7); $\beta = 10^{-4}$ for the l^1-l^2 model and $\beta_1 = 10^{-4}$, $\beta_2 = 10^{-1}$ for the l^1-l^1 model. The reconstructed PSF's in Fig. 9 are again used in an image deblurring problem. Specifically, the observed "Moon" image is obtained as follows: the clean image is first blurred by the true PSF and is then contaminated by a zero-mean Gaussian noise with noise variance 0.003. The observed image and the restored images by the respective

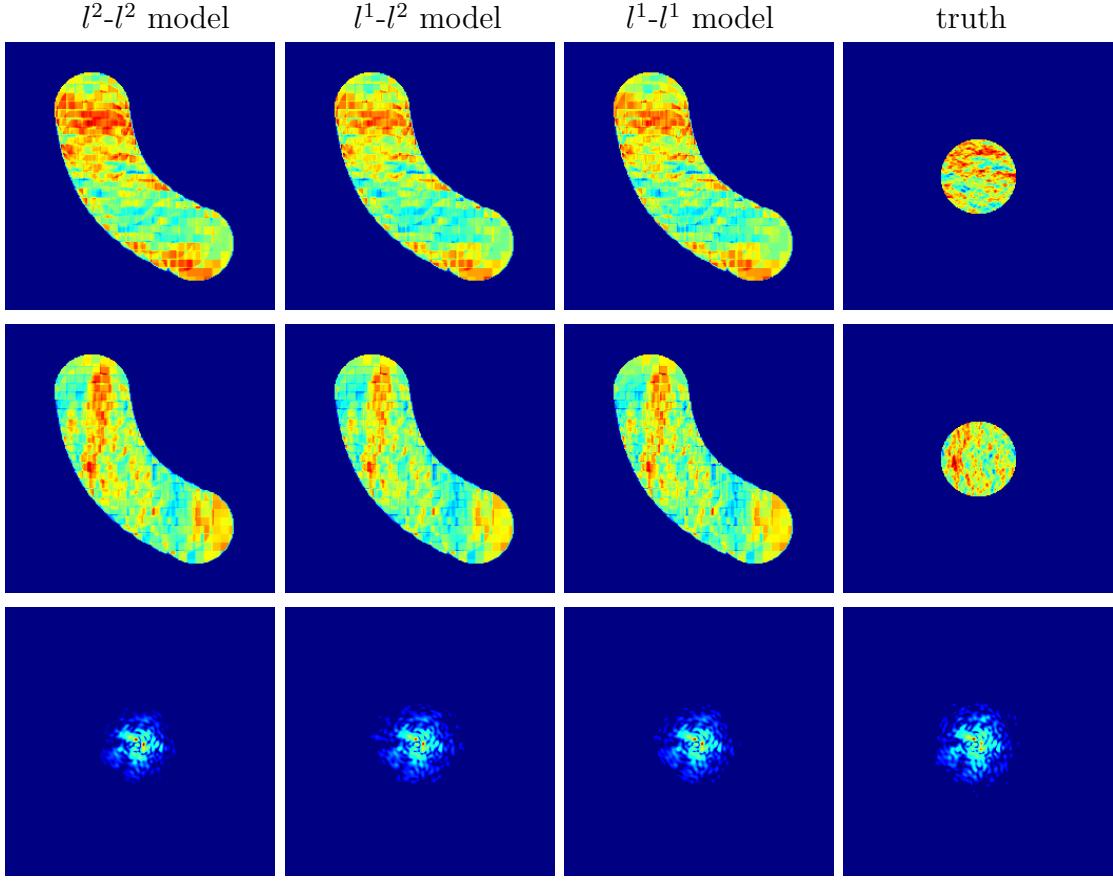


Fig. 6. Reconstructed horizontal (top row), vertical (center row) direction phase gradients and the PSF (bottom row).

reconstructed PSF's are given in Fig. 10. We see that we have about 1.6dB improvement in this case.

4.D. Wavefront reconstruction under different seeing conditions

As mentioned in the introduction, distortions of wavefront are related to seeing conditions, i.e., the ratio of d/r_0 . In this subsection, we test wavefront reconstructions with different seeing conditions for the models under consideration. The low-resolution phase gradients are generated as in Subsection 4.A but with $d/r_0 \in \{5, 40, 80, 120\}$. The models are then solved with values of the parameters taken to be the same as those in Subsection 4.A. We compare the performance of these three models by using the reconstructed PSF in the “Satellite” image deblurring problem in Section 4.A.

In Table 1 we list the numbers of iterations (“Iter”), computing times in seconds (“CPU”) and SNR values (“SNR”) of the restored “Satellite” images for these three models. For the

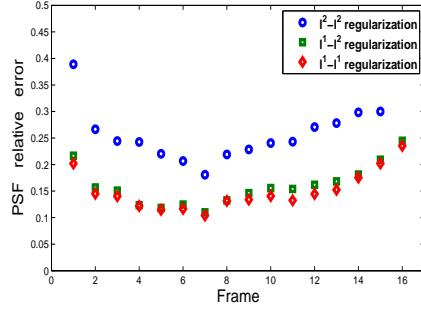


Fig. 7. The relative errors for all reconstructed PSF's.

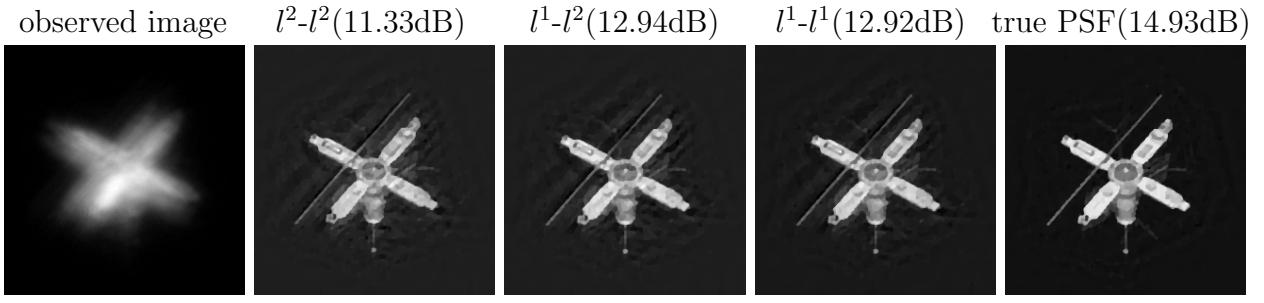


Fig. 8. Image restoration using PSF reconstructed by different models.

“Iter” column of l^1-l^2 (respectively l^1-l^1), the numbers in parentheses denote the iteration numbers for solving the subproblem (11) (respectively (13)). Note that the l^2-l^2 model is equivalent to a linear system while the l^1-l^2 and l^1-l^1 models are tricky to minimize because of the non-smooth l^1 term. The l^1-l^2 and l^1-l^1 models in principle require more computing time to solve than the l^2-l^2 model, but we see from the table that using ADM, our two models can also be solved very fast. In addition, the SNR values of the restored images by either l^1-l^2 or l^1-l^1 model are much higher than those by the l^2-l^2 model.

4.E. Sensitivity of parameters

In order to compare which model gives the best reconstructed PSF's and restored images, in our tests in the previous subsections, we have used the best choices of parameters α and β 's which are obtained manually. In this section, we report the robustness of our algorithms with respect to these parameters. We take the l^1-l^2 model as an example. The seeing condition d/r_0 is set as 50, and the phase ϕ is generated as in Subsection 4.A. In Figure 11, we compare

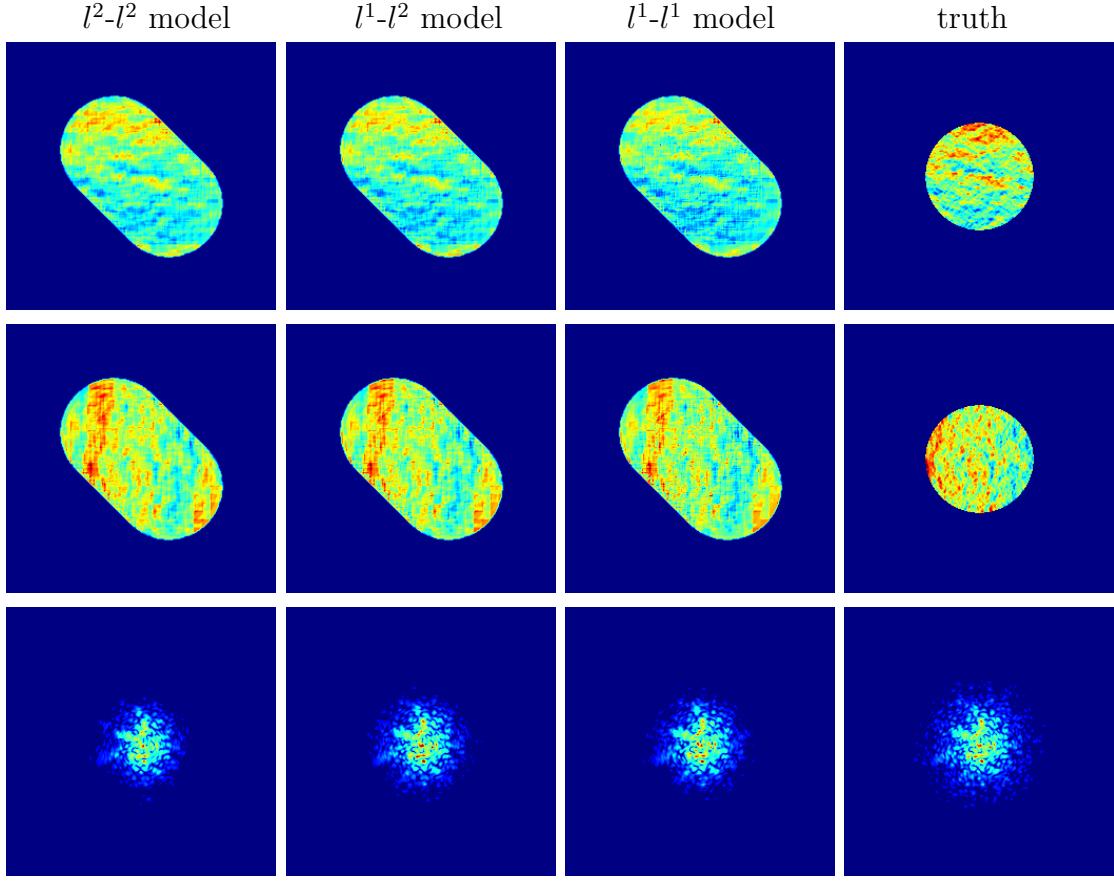


Fig. 9. Reconstructed horizontal (top row), vertical (center row) direction phase gradients and the PSF (bottom row).

the relative errors of PSF's reconstructed by the model with different values of α and β . For the left figure, β is fixed as 10^{-3} ; and for the right figure, α is fixed as 10^{-4} . We see that values of α in the order of 10^{-3} to 10^{-6} and β in the order of 10^{-2} to 10^{-3} are good for the l^1-l^2 model being tested. Hence the ranges for good choice of parameters are not small. We also notice that once a parameter is chosen correctly, the relative errors are very stable with respect to the other parameter.

5. Conclusion

In this paper, we discuss the problem of reconstructing the point spread function, or rather, the phase gradients under the frozen flow hypothesis in ground-based astronomy. We propose the l^1-l^p ($p = 1, 2$) model for solving the problem instead of the traditional Tikhonov regularization model. Our model can be solved effectively by the Douglas-Rachford alter-

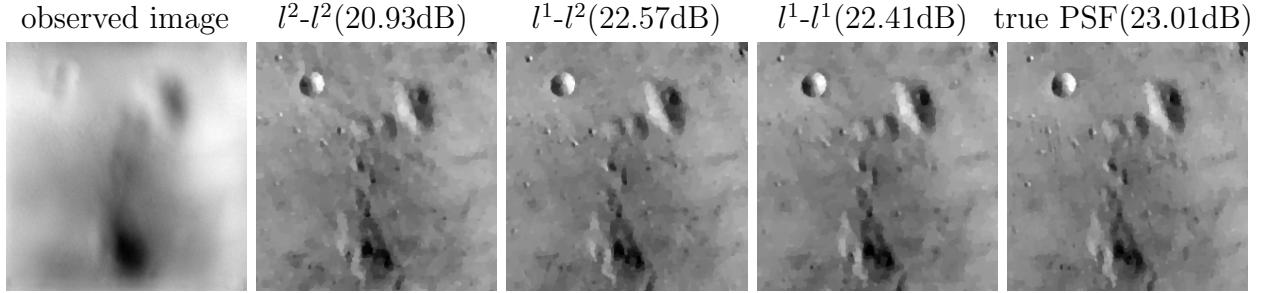


Fig. 10. Image restoration using PSF reconstructed by different models.

Table 1. Wavefront reconstruction with different seeing conditions.

d/r_0	l^2-l^2 model			l^1-l^2 model			l^1-l^1 model		
	Iter	CPU	SNR	Iter	CPU	SNR	Iter	CPU	SNR
5	50	0.94	15.46	11(147)	2.02	16.14	10(144)	2.50	16.03
40	51	0.89	13.47	11(163)	2.31	14.77	12(153)	2.64	14.71
80	51	0.87	11.04	13(180)	2.24	12.82	11(156)	2.72	12.83
120	51	0.90	9.69	12(177)	2.29	10.74	11(167)	2.87	10.72

nating direction method of multiplier. Experimental results showed that when compared to the Tikhonov regularization model, the new model can render more accurate phase gradients, and hence more accurate point spread functions. Therefore the images so deblurred are more accurate, with 1 to 1.6 dB higher in SNR. We also notice that the performance of the l^1-l^2 and l^1-l^1 models are about the same. Since the l^1-l^2 model is easier to solve than the l^1-l^1 model and has one less penalty parameter, it is the model that we recommend for this application.

In this paper, we consider only single-layer atmospheric turbulence. The models can be extended to multi-layer atmospheric turbulence [38] too, and it will be our future research direction.

Acknowledgments

The first author is supported in part by HKRGC Grant No. CUHK400510 and CUHK Direct Allocation Grant 2060408, and the second author is supported by HKRGC Grant No. HKBU203311.

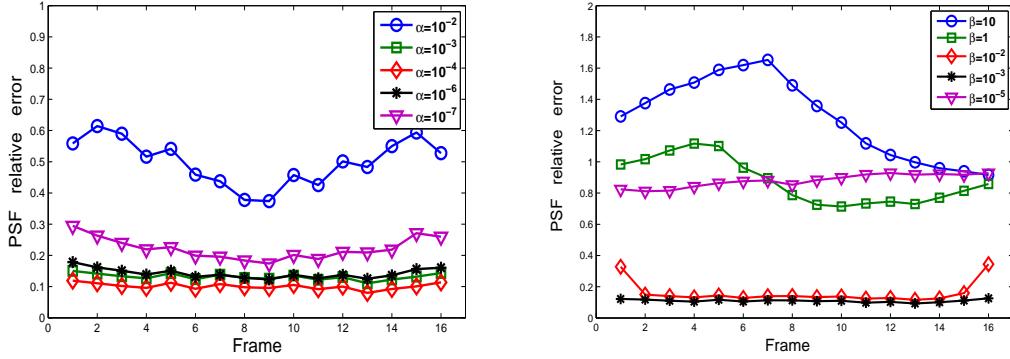


Fig. 11. Performances of l^1-l^2 model with different parameter values. Left: for different α with $\beta = 10^{-3}$. Right: for different β with $\alpha = 10^{-4}$.

References

1. T. F. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet and Stochastic Methods* (SIAM, Philadelphia, 2005).
2. L. Rudin, S. Osher, and F. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*. **60**, 259–268 (1992).
3. A. Tikhonov and V. Arsenin, *Solutions of Ill-posed Problems* (New York, 1997).
4. T. F. Chan and C. K. Wong, “Total variation blind deconvolution,” *IEEE Trans. Image Process.* **7**, 370–375 (1998).
5. E. Y. Lam and J. W. Goodman, “Iterative statistical approach to blind image deconvolution,” *J. Opt. Soc. Am. A*. **17** (2000), pp. 1177–1184.
6. J. M. Bardsley, “Wavefront reconstruction methods for adaptive optics systems on ground-based telescopes,” *SIAM J. Matrix Anal. Appl.* **30**, 67–83 (2008).
7. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 1996).
8. S. M. Jefferies, M. L. Hart, E. K. Hege, and J. Georges, “Sensing wave-front amplitude and phase with phase diversity,” *Appl. Opt.* **41**, 2095–2102 (2002).
9. L. M. Mugnier, C. Robert, J. M. Conan, V. Michau, and S. Salem, “Myopic deconvolution from wave-front sensing,” *J. Opt. Soc. Am.* **18**, 862–872 (2001).
10. D. L. Fried, “Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements,” *J. Opt. Soc. Am.* **67**, 370–375 (1977).
11. R. H. Hudgin, “Wavefront reconstruction for compensated imaging,” *J. Opt. Soc. Am.* **67**, 375–378 (1977).
12. S. M. Jefferies and M. Hart, “Deconvolution from wave front sensing using the frozen flow hypothesis,” *Opt. Express*, **19**, 1975–1984 (2010).

13. J. Nagy, S. Jefferies, and Q. Chu, “Fast PSF reconstruction using the frozen flow hypothesis,” *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*, (2010).
14. N. K. Bose and K. Boo, “High-resolution image reconstruction with multisensors,” *Int. J. Imaging Syst. Technol.* **9**, 294–304 (1998).
15. R. H. Chan, T. F. Chan, L. X. Shen, and Z. W. Shen, “Wavelet algorithms for high-resolution image reconstruction,” *SIAM J. Sci. Comput.* **24**, 1408–1432 (2003).
16. R. H. Chan, Z. W. Shen, and T. Xia, “A framelet algorithm for enhancing video stills,” *Appl. Comput. Harmon. Anal.* **23**, 153–170 (2007).
17. R. Y. Tsai and T. S. Huang, “Multiframe image restoration and registration,” *Adv. Computer Vis. Image Process.* 317–339 (1984).
18. S. Aliney, “A property of the minimum vectors of a regularizing functional defined by means of the absolute norm,” *IEEE Trans. Signal Process.* **45**, 913–917 (1997).
19. M. Nikolova, “A variational approach to remove outliers and impulse noise,” *J. Math. Imaging. Vis.*, **20**, 99–120 (2004).
20. D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite element approximations,” *Comput. Math. Appl.* **2**, 17–40 (1976).
21. R. Glowinski and A. Marrocco, “Approximation par éléments finis d’ordre un et résolution par pénalisation dualité d’une classe de problèmes non linéaires,” *R.A.I.R.O., R2*, 41–76 (1975).
22. R. Glowinski, *Numerical Methods for Nonlinear Variational Problems* (Springer-Verlag, 1984).
23. B. S. He and X. M. Yuan, “On the $O(1/n)$ convergence rate of Douglas-Rachford alternating direction method,” *SIAM J. Numer. Anal.* **50**, 700–709 (2012).
24. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer Verlag, 2006).
25. S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.* **3**, 1–122 (2010).
26. M. Afonso, J. Bioucas-Dias, and M. Figueiredo, “Fast image recovery using variable splitting and constrained optimization,” *IEEE Trans. Image Process.* **19**, 2345–2356 (2010).
27. R. H. Chan, J. F. Yang, and X. M. Yuan, “Alternating direction method for image inpainting in wavelet domain,” *SIAM J. on Imaging Sci.* **4**, 807–826 (2011).
28. M. Ng, P. A. Weiss, and X. M. Yuan, “Solving constrained total-variation problems via alternating direction methods,” *SIAM J. Sci. Comput.* **32**, 2710–2736 (2010).
29. D. Donoho, “De-noising by soft-thresholding,” *IEEE Trans. Inform. Theory* **41**, 613–627 (1995).

30. G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins University Press, 1996).
31. G. H. Golub, M. Heath, and G. Wahba, “Generalized cross-validation as a method for choosing a good ridge parameter,” *Technometrics*. **21**, 215–223 (1979).
32. V. A. Morozov, *Methods for Solving Incorrectly Posed Problems*. (Springer-Verlag, New York, 1984).
33. Y. W. Wen and R. H. Chan, “Parameter selection for total variation based image restoration using discrepancy principle,” *IEEE Trans. Image Process.* **21**, 1770–1781 (2012).
34. B. S. He, H. Yang and S. L. Wang, “Alternating direction method with self-Adaptive penalty parameters for monotone variational inequalities,” *J. Opt. Theory Appl.* **106**, 337–356 (2000).
35. B. S. He, L. Z. Liao, D. R. Han and H. Yang, “A new inexact alternating directions method for monotone variational inequalities,” *Math. Program.* **92**, 103–118 (2002).
36. C. M. Harding, R. A. Johnston and R.G. Lane, “Fast simulation of a Kolmogorov phase screen,” *App. Opt.* **38**, 2161–2170 (1999).
37. B. S. He and X. M. Yuan, “Convergence analysis of primal-dual algorithms for total variation image restoration,” *SIAM J. Imaging Sci.* **5**, 119–149 (2012).
38. H. Jakobsson, “Simulations of time series of atmospherically distorted wave fronts,” *Appl. Opt.*, 35, 1561–1565 (1996).