SIAM J. IMAGING SCIENCES
Vol. 13, No. 2, pp. 775–806

Reconstruction of the High Resolution Phase in a Closed Loop Adaptive Optics System*

Rihuan Ke†, Roland Wagner‡, Ronny Ramlau§, and Raymond Chan¶

Abstract. Adaptive optics is a commonly used technique to correct the phase distortions caused by the Earth's atmosphere to improve the image quality of the ground-based imaging systems. However, the observed images still suffer from the blur caused by the adaptive optics residual wavefront. In this paper, we propose a model for reconstructing the residual phase in high resolution from a sequence of deformable mirror data. Our model is based on the turbulence statistics and the Taylor frozen flow hypothesis with knowledge of the wind velocities in atmospheric turbulence layers. A tomography problem for the phase distortions from different altitudes is solved in order to get a high quality phase reconstruction. We also consider inexact tomography operators resulting from the uncertainty in the wind velocities. The wind velocities are estimated from the deformable mirror data and, additionally, by including them as unknowns in the objective function. We provide a theoretical analysis on the existence of a minimizer of the objective function. To solve the associated joint optimization problem, we use an alternating minimization method which results in a high resolution reconstruction algorithm with adaptive wind velocities. Numerical simulations are carried out to show the effectiveness of our approach.

Key words. image improvement, adaptive optics, astronomical imaging

AMS subject classification. 85-08

DOI. 10.1137/19M1258426

1. Introduction.

Modern telescopes allow the acquisition of high resolution images of astronomical objects seen in the night sky. To increase the resolution, ground-based extremely large telescopes (ELTs) with a primary mirror bigger than 30 m in diameter are currently under construction. However, the image quality is still degraded due to the turbulent atmosphere

*Received by the editors April 26, 2019; accepted for publication (in revised form) January 27, 2020; published electronically May 5, 2020.
https://doi.org/10.1137/19M1258426

Funding: The work of the first author was supported by the Hong Kong Research Grants Council grant CUHK14306316 and by the Austrian Science Fund (FWF) in the Doctoral Program "Computational Mathematics" grant W1214 - Project 8. The work of the second and third authors was supported by the Austrian Science Fund (FWF) project F 6805-N36: "SFB Tomography Across the Scales" and by the Austrian Ministry of Research (Hochschulraumstrukturmittel) in the project "Beobachtungsorientierte Astrophysik in der E-ELT Ara." The work of the fourth author was supported by the Hong Kong Research Grants Council grants CUHK14306316, CUHK14301718, CityU Grant 9380101, CRF grant C1007-15G, and AoE/M-05/12.

†Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK (krhuan@gmail.com).

‡Industrial Mathematics Institute, Johannes Kepler University, Altenberger Strasse 69, 4040 Linz, Austria (wagner@indmath.uni-linz.ac.at).

§Industrial Mathematics Institute, Johannes Kepler University, Altenberger Strasse 69, 4040 Linz, Austria and Johann Radon Institute for Computational and Applied Mathematics, Altenberger Strasse 69, 4040 Linz, Austria (ronny.ramlau@jku.at).

¶College of Science, City University of Hong Kong, Kowloon, Hong Kong (rchan.sci@cityu.edu.hk).

above these telescopes. As a remedy for this degradation, adaptive optics (AO) systems were introduced; see, e.g., [15, 12].

An astronomical AO system is a hardware-based device that uses deformable mirrors (DMs) to compensate for distortions of the incoming phase due to atmospheric turbulence and therefore making the observed images sharper. The DMs are adjusted according to the data from wavefront sensors (WFS) which measure the incoming wavefront from so-called guide stars, which is the phase measured in a specific wavelength. However, due to several physical limitations, the distortions by atmospheric turbulence can never be compensated for completely. First, the WFS have a rather coarse resolution as enough light of the guide star has to reach each pixel of the WFS to acquire accurate data. As a consequence, the higher spatial frequencies of the wavefront are not sensed and thus cannot be corrected by the deformable mirror. Second, the correction of the wavefront is also limited by the DM, the shape of which is controlled by a finite number of actuators. Finally, a time gap exists between the acquisition of wavefront data and the update of the DM shape. The atmospheric turbulence is changing rapidly and the increment of the distorted wavefront within the time gap is not taken into account. The uncorrected part of the wavefront, called the residual wavefront, results in a residual blur of the observed image.

Further improvement on the quality of the observed image can be made by image post-processing with data from the AO systems. According to the Fourier optics model (see, e.g., [20, 23]), in an incoherent imaging system, the blurred image is formed by the convolution of the point spread function (PSF) and the true image. The PSF can be determined by the residual wavefront, or the residual phase. Several techniques known as deconvolution methods (see, e.g., [24, 9, 10, 13]) have been developed. Such methods rely on accurate PSF knowledge in order to remove the blur from the observed image. An estimate of the PSF can be computed from the WFS data (see, e.g., [36]) or the reconstructed residual wavefronts [39, 38], which always requires the use of a simulation step for the spatial frequencies of the residual wavefront that are not sensed by the WFS.

The problem of reconstructing a high resolution phase from low resolution data has been investigated by several authors in the literature. In particular, Chu, Jefferies, and Nagy [10] consider an open loop system and high resolution gradients of the phase are computed from WFS data at multiple time steps by doing a least squares data fitting with Tikhonov regularization. The high resolution phases are then reconstructed using the computed gradients. Chan, Yuan, and Zhang [9] suggest that the phase gradients are not smooth and propose an l_1 regularization term for recovering the high resolution gradients. The method developed in [8], in contrast, reconstructs the high resolution phase directly from low resolution gradients. All these works consider the reconstruction problems in an open loop setting, and they require precisely known tomography operators.

The core issue of this paper is to find a high resolution residual phase in a closed loop AO system which can then be used to obtain a high quality PSF estimate in order to further mitigate the blur of the images. The high resolution residual phase has the capacity of capturing the higher spatial frequencies needed for a more accurate PSF. We make use of the commonly adopted Taylor frozen flow (TFF) assumption [34] to relate a sequence of low resolution AO data to the high resolution phase with the wind velocities. A variational regularization method based on the turbulence statistics and the TFF assumption is developed.

The wind velocities, which define the tomography operator, are not available in high accuracy in reality. This introduces a perturbation of the tomography operator. As consistency between the TFF and the data helps to recover the operator, we take the wind velocities as variables of the objective function instead of fixed values. The minimization of the objective function produces estimations of the high resolution phase and the wind velocities simultaneously. To our knowledge, only an instantaneous low resolution reconstruction of the wavefront on the same grid as the WFS is performed in projects for AO assisted telescopes. This is also due to the fact that usually the AO data are not saved and thus only one set of data is available at every time. However, for upcoming telescopes such as ESO's ELT this will change and at least parts of the data will be available for reconstructing the PSF in a postprocessing step. In our approach we solve a joint model, estimating high resolution wavefronts as well as wind velocities of atmospheric layers.

To the best of our knowledge, this is the first time that a high resolution wavefront reconstruction for AO systems, based on a tomographic approach and measurements from a shifted atmosphere, is considered. In particular, we propose to jointly reconstruct the wavefront and the imprecisely known wind velocities from noisy measurement using a unified variational approach. This also helps to refine the inaccurate tomography operator that we start with. We prove that there exists a minimizer of the objective function of the joint model. We also present an algorithm for solving the underlying nonconvex optimization problem. Note that our approach differs from [27, 9], where observations without AO systems were considered with the goal to recover an instantaneous PSF.

In image processing, the high resolution problem which aims for restoring a high resolution image from a sequence of the low resolution ones has been attracting much interest (see, e.g., [7, 16]). It has been used in a variety of practical applications such as video enhancement, facial image analysis, or medical image processing [28].

The remainder of this paper is organized as follows: In section 2, we describe the problem setting and clarify the notation. In section 3, a model is proposed based on the turbulence statistics and the TFF hypothesis, by assuming the wind velocities are known in advance. In order to handle the underlying tomography problem with inexact tomography operators due to the uncertainty in the wind velocities, we also consider adjusting the wind velocities from the DM data using our model. In section 4, the minimization problem induced by our model is considered. We implement an alternating direction algorithm for computing estimations of the residual phase and the wind velocities. Finally, section 5 displays the simulated AO system and the numerical results.

2. Problem modeling.

Throughout this paper, we focus on so-called single-conjugate adaptive optics (SCAO) systems. An SCAO system consists of one WFS, one DM, and a control unit, which in particular computes the DM shape from data obtained by the WFS. AO systems can be constructed and run in two different settings: open and closed loop. In open loop, the WFS sees the full, uncorrected incoming wavefront, i.e., the WFS is located *in front* of the DM in the optical path. In contrast, in closed loop the WFS only sees the residual wavefront after correction by the DM, i.e., the WFS is located *behind* the DM in the optical path. In current AO systems, closed loop is the preferred setting, as this realizes measurements closer to zero where the behavior of the sensor is linear and much better understood. We

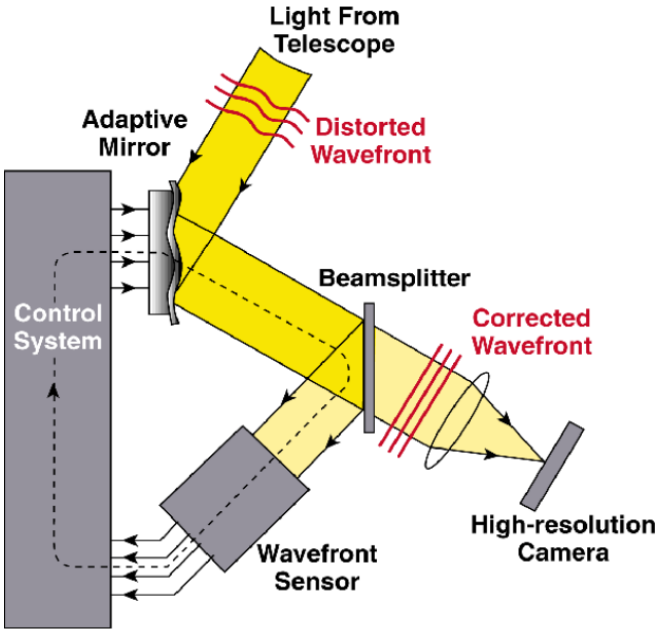


Figure 1. *Illustration of a closed loop AO system from* [11]. Adaptive mirror *is equivalent to DM,* distorted wavefront *is the incoming wavefront, the* control system *computes the reconstruction of* $\phi_t^{res,\ coarse}$ *from WFS data and uses it to adjust the DM. The* corrected wavefront*, which is called the* residual wavefront *throughout the paper, refers to the residual errors after the AO correction and it is the main quantity of interest.*

assume that the SCAO system under consideration is run in a *closed loop*, meaning that WFS is located in the optical path after the DM and thus measures only a residual of the incoming phase after correction through the DM, where the incoming phase is the phase arriving at the telescope after being distorted by the atmospheric turbulence. These components and their interaction are illustrated in Figure 1. We use a time discrete setting, i.e., t always indicates the tth time step and is therefore used as a subscript. Each time step takes only a few milliseconds of real time in our application. Let us denote the incoming phase at t as $\boldsymbol{\phi}_t(\boldsymbol{x})$, where $\boldsymbol{x} \in \Omega \subset \mathbb{R}^2$ in the pupil plane and Ω is the telescope aperture. Let $\boldsymbol{\phi}_t^{\mathrm{DM}}(\boldsymbol{x})$ be the DM shape. Then the (unknown) residual phase $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$ can be written as

$$\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x}) = \boldsymbol{\phi}_t(\boldsymbol{x}) - \boldsymbol{\phi}_t^{\mathrm{DM}}(\boldsymbol{x}) \tag{1}$$

for $\boldsymbol{x} \in \Omega$ inside the telescope pupil. Recall that in a closed loop system the WFS measures the residual phase, i.e., $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$. In contrast, in an open loop the full incoming wavefront, i.e., $\boldsymbol{\phi}_t(\boldsymbol{x})$, is measured as the light beam is split before being reflected (and thus corrected) on the DM. The goal of every AO system is to perfectly match the DM to the incoming phase, i.e., $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x}) = 0$, but due to a system inherent time delay stemming from measuring, calculating, and adjusting the DM and the coarse resolution of both WFS and DM; this is not possible in reality. Further errors are introduced as $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$ can only be measured indirectly by a WFS and thus needs to be reconstructed from these data.

Fortunately, $\boldsymbol{\phi}_t^{\mathrm{DM}}(\boldsymbol{x})$ is updated over time and the information of $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$ is encoded in the updates. A coarse resolution version $\boldsymbol{\phi}_t^{\mathrm{DM}}$ of $\boldsymbol{\phi}_t^{\mathrm{DM}}(\boldsymbol{x})$ is stored and can be used to recover

Table 1
Data for high resolution reconstruction.

Data	Description
$\boldsymbol{\phi}_t^{\mathrm{DM}}$	The coarse grid DM shape at time step t.
$\boldsymbol{\phi}_t^{\text{res, coarse}}$	The computed phase residual on the coarse grid at time step t, an estimate for $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$, obtained as a solution to (2). This quantity is not saved by the AO system, but it can be exactly recovered from $\boldsymbol{\phi}_t^{\mathrm{DM}}$.
$\tilde{\boldsymbol{s}}_t$	The pseudo-WFS measurement given by (5). It is used in our refined WFS model (4) for finding the high resolution residual phase.
$\boldsymbol{v}$	The wind velocity, which is constant over the aperture and time.

$\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$ in postprocessing. The main problem in this paper is to reconstruct $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$ in high resolution from a sequence of $\boldsymbol{\phi}_t^{\mathrm{DM}}$ at different time steps.

In the remaining part of this section, we will introduce a mathematical model to link $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$ to the updates of DM shapes $\boldsymbol{\phi}_t^{\mathrm{DM}}$, and then solve an inverse problem for a high resolution $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$. The main structural idea is as follows. First, we formulate the relationship between $\boldsymbol{\phi}_t^{\mathrm{DM}}$ and $\boldsymbol{\phi}_t^{\text{res, coarse}}$ which essentially means a coarse resolution approximation of $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$. Second, from $\boldsymbol{\phi}_t^{\text{res, coarse}}$ we derive pseudomeasurements $\tilde{\boldsymbol{s}}_t$, which are then used as input to our algorithm for finding $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$ from a refined WFS model. Additionally, we assume a constant wind velocity $\boldsymbol{v}$ given in units [m/(time step)] which describe the motion of $\boldsymbol{\phi}_t(\boldsymbol{x})$. All these quantities are summarized in Table 1.

2.1. WFS model and low resolution DM data. Let us start with the forward model connecting the WFS measurements and the residual of the incoming wavefront. We assume a closed loop AO system using a Shack–Hartmann WFS (SH-WFS). In such an AO system, the WFS is located behind the DM (as Figure 1 shows). The WFS is modeled as an operator $\Gamma : H^{11/6}(\Omega) \to \mathbb{R}^{n^2\times 2}$ (cf., e.g., [22]) mapping a phase $\boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$ onto measurements $\boldsymbol{s}_t$ obtained on $n \times n$ subapertures, i.e.,

$$\boldsymbol{s}_t = \Gamma \boldsymbol{\phi}_t^{\mathrm{res}}(\boldsymbol{x})$$

with $\Gamma := (\Gamma_x, \Gamma_y)^T$, defined via its action on subapertures $\Omega_{i,j}$, $i,j = 1,\dots,n$ giving as measurements the discrete averaged gradients $\boldsymbol{s}_{t,x}$ and $\boldsymbol{s}_{t,y}$, given as

$$(\Gamma_z \boldsymbol{\phi}_t^{\mathrm{res}})_{i,j} = \frac{1}{|\Omega_{i,j}|} \int_{\Omega_{i,j}} \frac{\partial}{\partial z} \boldsymbol{\phi}^{\mathrm{res}}(\boldsymbol{x})\, d\boldsymbol{x},\ z \in \{x,y\},\ \boldsymbol{x} = (x,y).$$

Clearly, Γ is bounded in $H^1(\Omega)$. Note that the use of the space $H^{11/6}(\Omega)$ for the definition of the operator Γ stems from the fact that the atmospheric turbulence, and thus the wavefronts, follows a von Karman power law, which basically states that the wavefronts belong to $H^{11/6}$ with high probability (cf., e.g., [37, 25], and section 3.1.2). These SH-WFS measurements $\boldsymbol{s}_t$ are gradients averaged over the subapertures $\Omega_{i,j}$, $i,j = 1,\dots,n$ with $\cup_{i,j=1}^{n} \Omega_{i,j} = \Omega$.

Generating the low resolution DM data. The AO system needs to reconstruct the residual phase in real time in order to control $\boldsymbol{\phi}_t^{\mathrm{DM}}(\boldsymbol{x})$ since the $\boldsymbol{\phi}_t(\boldsymbol{x})$ in (1) changes quickly over the time steps t. This raises an inverse problem in which the 2-dimensional function

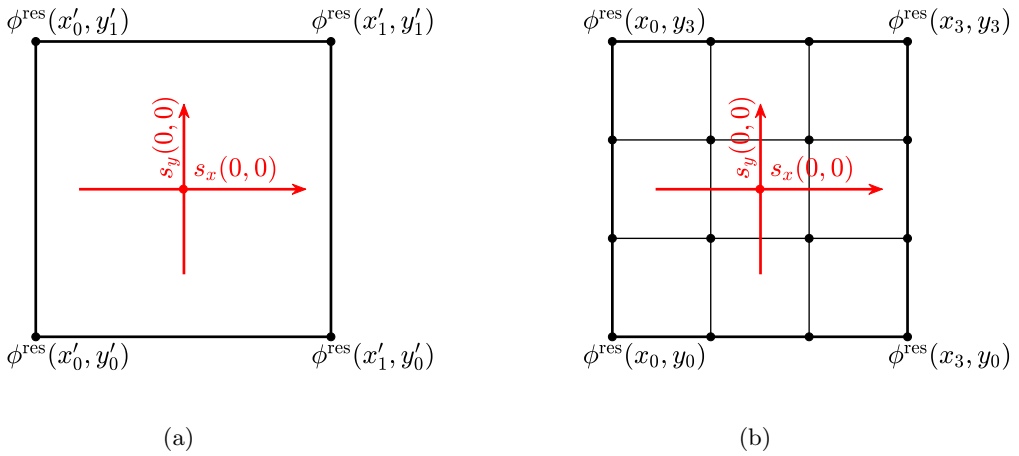


Figure 2. (a) *The SH-WFS measurement and the low resolution residual phase in a subaperture. In this example,* (x_i', y_j') *are the grid points which satisfy* $x_i' = x_{3i}, y_j' = y_{3j}$. (b) *The SH-WFS measurement and the high resolution residual phase in a subaperture of the WFS. The subaperture is bounded by the outermost square whose side length is* 3 *times the grid spacing of* $\boldsymbol{\phi}^{res}$, (x_i, y_j) *are the grid points. The measurement* $s_x(0,0)$ *and* $s_y(0,0)$ *is modeled as the average slope in the* x *direction and* y *direction in the subaperture, respectively.*

$\boldsymbol{\phi}_t^{\text{res}}(\boldsymbol{x})$ has to be reconstructed for each time step t from given measurements $\boldsymbol{s}_t$. However, the DM has a finite number of actuators only and thus it is sufficient to perform a discrete phase reconstruction instead of inverting the continuous operator. We view Γ as a combination of a projector $P_n : H^{11/6}(\Omega) \to \mathbb{R}^{(n+1)\times(n+1)}$ and a discrete SH-WFS operator $\Gamma_n : \mathbb{R}^{(n+1)\times(n+1)} \to \mathbb{R}^{n^2\times 2}$, i.e., $\Gamma = \Gamma_n P_n$. Several ways to choose the $(n+1)\times(n+1)$ grid exist.

The discrete representation of the residual phase can then be obtained as a solution $\boldsymbol{\phi}_t^{\text{res, coarse}}$ to the discrete inverse problem

$$\boldsymbol{s}_t = \Gamma_n \boldsymbol{\phi}_t^{\text{res, coarse}} + \boldsymbol{\eta}_t, \tag{2}$$

where $\boldsymbol{\phi}_t^{\text{res, coarse}} := P_n \boldsymbol{\phi}_t^{\text{res}}(\boldsymbol{x})$ is in low resolution, and $\boldsymbol{\eta}_t$ models the noise. Figure 2(a) illustrates the discretization for the SH-WFS case. During an AO run, the measurements $\boldsymbol{s}_t$ are obtained and the residual phase $\boldsymbol{\phi}_t^{\text{res, coarse}}$ needs to be reconstructed, i.e., (2) has to be inverted. To solve this problem several direct methods were developed and studied in simulation and on-sky [43, 32, 33, 29, 40, 3, 4]. These methods neglect the motion of the atmosphere during the measurement and computation time due to the small shift. The distance of the shift is determined by the wind velocity $\boldsymbol{v} \in \mathbb{R}^2$. However, to reconstruct a high resolution phase from multiple frames, the shift of the atmosphere has to be taken into account. We delay the discussion on the shift to section 3.

Once a solution to (2) is obtained, it is used to compute the new DM shape $\boldsymbol{\phi}_{t+1}^{\text{DM}}$, minimizing the residual in (1), as

$$\boldsymbol{\phi}_{t+1}^{\text{DM}} = \boldsymbol{\phi}_t^{\text{DM}} + \alpha \cdot \boldsymbol{\phi}_t^{\text{res, coarse}}, \tag{3}$$

where $\alpha > 0$ is the so-called loop gain and the initial DM shape is given as $\boldsymbol{\phi}_{-1}^{\text{DM}} = 0$. These data are generated by the AO system and stored for image postprocessing.

2.2. Refined WFS model. We want to reconstruct the unkown $\boldsymbol{\phi}_t^{\text{res}}$ on a finer grid than the WFS or DM, meaning the grid of $\boldsymbol{\phi}_t^{\text{DM}}$. Therefore, we need to introduce additional operators coarsening functions on the respective grids. Let Q_{nk}, for $k \in \mathbb{N}$, be the projector from $H^{11/6}(\Omega)$ onto $\mathbb{R}^{(nk+1)\times(nk+1)}$ and C_k the coarsening operator mapping a fine resolution phase from $\mathbb{R}^{(nk+1)\times(nk+1)}$ to $\mathbb{R}^{(n+1)\times(n+1)}$. From this we can model the measurements of a higher resolution phase $\boldsymbol{\phi}_t^{\text{res}} := Q_{nk}\boldsymbol{\phi}_t^{\text{res}}(\boldsymbol{x})$ with $P_n = C_k Q_{nk}$, so using (2) we have

$$\boldsymbol{s}_t = \Gamma_n C_k \boldsymbol{\phi}_t^{\text{res}} + \boldsymbol{\eta}_t, \tag{4}$$

where $\boldsymbol{\eta}_t$ models the noise. Note that the coarsening operator is not uniquely defined.

An illustration of the grids of the SH-WFS measurement and the high resolution residual phase is given in Figure 2(b) in which the SH-WFS measurement is considered to be located at the center of the subaperture following the Fried geometry [17].

2.3. From low resolution DM data to high resolution phase. As AO systems are running at a frequency of up to 500 Hz, meaning 500 sets of WFS measurements and DM shapes are measured and applied, respectively, each second, a huge amount of data needs to be saved. If one wants to use these data in a postprocessing step, then it might be beneficial to reduce the amount of saved data as much as possible. One option is to save the applied DM shapes only, which are represented by point values of the corresponding function, instead of the WFS data, being averaged gradients, as this already reduces the amount of data by a factor 2.

This means that $\boldsymbol{\phi}_t^{\text{res, coarse}}$, the solution to (2), computed in real time during the AO run, is treated as the input to our problem instead of $\boldsymbol{s}_t$. Therefore, we rewrite (3) for recovering $\boldsymbol{\phi}_t^{\text{res, coarse}}$ from the DM shapes as

$$\boldsymbol{\phi}_t^{\text{res, coarse}} = \frac{1}{\alpha}(\boldsymbol{\phi}_{t+1}^{\text{DM}} - \boldsymbol{\phi}_t^{\text{DM}}).$$

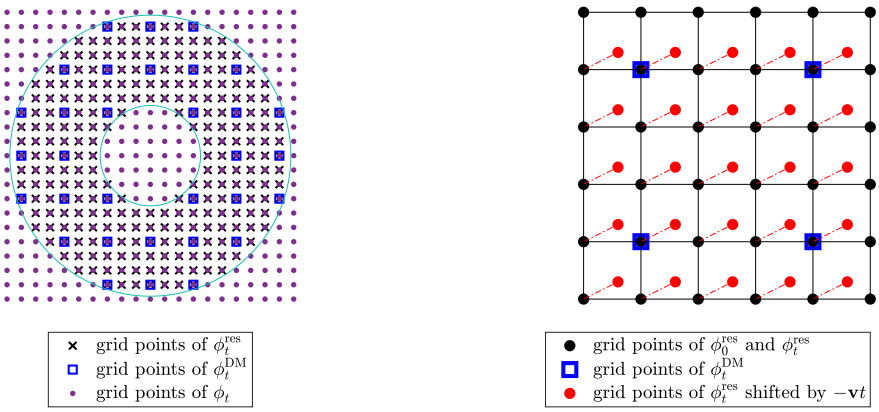
Since $\boldsymbol{\phi}_t^{\text{res, coarse}}$ represent low resolution copies of $\boldsymbol{\phi}_t^{\text{res}}$, one may want to compute $\boldsymbol{\phi}_t^{\text{res}}$ by upsampling them. However, $\boldsymbol{\phi}_t^{\text{res, coarse}}$ is known only up to an additive constant, given that the constant function is in the null space of Γ_n, due to the underlying physical structure of the WFS. To manage the issue of unknown constant, we define a pseudo-WFS measurement that respects the low resolution WFS model (2) via

$$\tilde{\boldsymbol{s}}_t := \Gamma_n \boldsymbol{\phi}_t^{\text{res, coarse}}, \tag{5}$$

and use it instead of $\boldsymbol{\phi}_t^{\text{res, coarse}}$ in the high resolution reconstruction.

In summary, our approach is based on the refined observation model (4), using pseudo-WFS measurements created from the saved DM shapes via (5). Recall that both $\boldsymbol{\phi}_t^{\text{DM}}$ and $\boldsymbol{\phi}_t^{\text{res, coarse}}$ are in low resolution, and the quantity of interest $\boldsymbol{\phi}_t^{\text{res}}$ is in high resolution. The grids for different quantities are compared in Figure 3(a). So the problem is highly ill-posed. The idea for dealing with the ill-posedness of this problem is to combine multiple measurements from neighboring time steps, and to use prior knowledge of the phase.

3. Estimating the residual phase in high resolution. In order to estimate the residual phase in high resolution, we have to make assumptions on the flow and statistics of the atmosphere. For this purpose, we present and adopt the TFF hypothesis from [34] and the



(a) (b)

Figure 3. (a) *Three different discretization grids. The purple dots, the blue square, and the black cross represent the points on the grids of* $\boldsymbol{\phi}_t$, $\boldsymbol{\phi}_t^{res,\ coarse}$, *and* $\boldsymbol{\phi}_t^{res}$, *respectively. In this example the grid of* $\boldsymbol{\phi}^{res}$ *is* 3 *times finer than that of* $\boldsymbol{\phi}^{res,\ coarse}$. *The residual phase* $\boldsymbol{\phi}_t^{res}$ *and computed phase* $\boldsymbol{\phi}_t^{res,\ coarse}$ *lie on the annular aperture only.* (b) *The grid of* $\boldsymbol{\phi}_t^{res}$ *shifted by* $-\boldsymbol{v}t$ *(i.e., the small red dots) does not fall on the grid of* $\boldsymbol{\phi}_0^{res}$. *To define the discrete motion operator, the shifted* $\boldsymbol{\phi}_t^{res}$ *should be interpolated onto the grid of* $\boldsymbol{\phi}_0^{res}$ *(i.e., the black dots).*

von Karman power spectral density from [31] in the following sections. Furthermore, we first investigate the simplified problem of estimating the residual phase in the presence of only one atmospheric layer and known wind velocities. However, measurements of the wind velocities are not available all the time, thus we will introduce a more evolved setting, allowing one to adapt the wind velocities by using the current WFS measurements. In a last step, we move to the realistic situation of several atmospheric layers moving with different wind speeds and directions. Thus a tomographic problem arises which needs to be solved.

3.1. Assumptions on the atmosphere.

To model the properties of the atmospheric turbulence, we make assumptions on the flow within the atmospheric layers and on their statistics.

3.1.1. TFF hypothesis.

In this section, we introduce the TFF hypothesis in order to be able to connect subsequently measured information of the incoming phase. In particular, we want to use this assumption to find a shift operator mapping the phase at time step t back to its position at time step 0. This hypothesis was introduced in [34]: The atmospheric turbulence is composed of a number of layers located at different altitudes above the ground. Each of the layers does not change its shape on small time scales and moves across the telescope aperture with a constant velocity due to wind. As shown in [2] and the references therein, the TFF holds up to approx 100 ms, which corresponds to 50 time frames in an AO system with loop frequency of 500 Hz (i.e., 2 ms per time frame). In a first step, we assume a one-layer model. As a consequence of the TFF hypothesis, the incoming phase reaching the telescope has a translational motion as

$$\boldsymbol{\phi}_t(\boldsymbol{x}) = \boldsymbol{\phi}_0(\boldsymbol{x} - t\boldsymbol{v}), \tag{6}$$

where $\boldsymbol{v} \in \mathbb{R}^2$ is the wind velocity and t denotes the tth time step, with $0 < t < T$, e.g., for a loop frequency of 500 Hz, we get $t \leq 50$. Equation (6) together with (1) implies that $\boldsymbol{\phi}_0(\boldsymbol{x})$ is measured repeatedly by the WFS, provided that the DM shapes $\boldsymbol{\phi}_t^{\mathrm{DM}}(\boldsymbol{x})$ are known and T is not too big, i.e., $|T\boldsymbol{v}| < D$, where D is the telescope diameter. Based on this fact a high resolution reconstruction of $\boldsymbol{\phi}_0(\boldsymbol{x})$ is possible.

Discrete motion operator. Let us now define the discrete motion operator $M_{\boldsymbol{v},t}$ acting on $\boldsymbol{\phi}_t$ associated with the translation of the form $f(\boldsymbol{x}) \to f(\boldsymbol{x} + t\boldsymbol{v})$. If $t\boldsymbol{v}$ is not a multiple of the grid spacing, we use bilinear interpolation, as shown in Figure 3(b). Thus, we can approximate (6) in a discrete sense as

$$M_{\boldsymbol{v},t}\boldsymbol{\phi}_t \approx \boldsymbol{\phi}_0. \tag{7}$$

3.1.2. Turbulence statistics. In order to use the properly weighted terms in our functional, we need to gather knowledge on the turbulence statistics of the atmosphere. Fortunately, the statistical properties of the atmospheric turbulence have been well studied. In Kolmogorov's theory, atmospheric turbulence is assumed to be a homogeneous and isotropic random process [25, 41]. The turbulence statistics are usually described by its power spectral density (PSD) and there are various versions of PSDs in the literature. In this paper, we assume the von Karman PSD of the phase $\boldsymbol{\phi}(\boldsymbol{x})$ (see, e.g., [31]) given by

$$\mathscr{P}_\phi(\boldsymbol{\kappa}) = \frac{0.023 r_0^{-5/3}}{\left(\kappa_0^2 + |\boldsymbol{\kappa}|^2\right)^{11/6}},$$

where r_0 is the Fried parameter, $\kappa_0 = 1/L_0$, and L_0 is the atmospheric turbulence outerscale. The covariance function of the phase is the inverse Fourier transform of the PSD, i.e.,

$$C_\phi = \mathcal{F}^{-1}\left(\mathscr{P}_\phi\right), \tag{8}$$

where $\mathcal{F}$ denotes the Fourier transform.

Approximate covariance operator of the phase. In a discrete setting, the covariance matrix $C_{\boldsymbol{\phi}}$ derived from (8) and its inverse are dense and, therefore, not efficient in real computation especially for large scale problems. Many approximations of the covariance matrix have been developed in the past decades in order to achieve fast computations; see for instance [35, 42]. In particular, Ellerbroek showed in [14] that the biharmonic operator Δ^2 provides a good approximation to the inverse covariance operator. The discrete biharmonic operator has sparsity which contributes to efficient matrix-vector multiplications and shows great advantages when iterative solvers are applied. Thus, we will use it in our method to approximate the covariance matrix.

3.2. Mathematical models for high resolution estimation. We establish a mathematical model for estimating a high resolution wavefront in the following sections. In particular, we first assume known wind velocities, being the ideal case in ground-based astronomy. An additional enhancement of the model leads us to a joint estimation of wavefront and wind velocities.

3.2.1. High resolution reconstruction with known wind velocities. In the following, we will set up a functional which is minimized to obtain a high resolution reconstruction of the residual phase. We use the previously introduced operators C_k, which relates the two different discretization grids in (4), and $M_{\boldsymbol{v},t}$, which accounts for atmospheric motion due to wind in (7). As data, we use the set of pseudo-WFS measurements $\{\tilde{\boldsymbol{s}}_t\}$ to replace the set of WFS data $\{\boldsymbol{s}_t\}$. Therefore, the high resolution reconstruction does not require any AO data other than the DM shapes. Due to the fact that the number of high resolution pixels is bigger than the dimension of the measurements, the high resolution reconstruction problem is ill-posed. As in practice, $\tilde{\boldsymbol{s}}_t$ always contains noise from the sensor and the real time reconstruction algorithm, regularization techniques are necessary to reach an accurate and stable reconstruction. Additionally, we would like to use the statistics of the turbulence (cf., section 3.1.2). To this end, let $H_1 = H + \epsilon I$, where H is the discrete Laplacian operator, $\epsilon > 0$, and I the identity matrix. The matrix ϵI is an analogy to κ_0 in $\mathscr{P}_\phi$ and avoids zero eigenvalues appearing in H. Assuming the von Karman phase PSD, H_1^2 approximates the inverse covariance matrix of $\boldsymbol{\phi}_t$ up to a scaling constant.

As the TFF hypothesis and von Karman PSD are valid for the incoming phase rather than the residual phase we first reconstruct $\boldsymbol{\phi}_t$ and then an estimate of $\boldsymbol{\phi}_t^{\text{res}}$ using (1). Recall that the DM shape $\boldsymbol{\phi}_t^{\text{DM}}$ lies on a coarse grid. However, through the so-called influence functions of the DM actuators, we are able to interpolate $\boldsymbol{\phi}_t^{\text{DM}}$ to the same resolution as $\boldsymbol{\phi}_t^{\text{res}}$. Denoting the high resolution DM shape as $\tilde{\boldsymbol{\phi}}_t^{\text{DM}}$, we get a discrete version of (1) as

$$\boldsymbol{\phi}_t^{\text{res}} = \boldsymbol{\phi}_t - \tilde{\boldsymbol{\phi}}_t^{\text{DM}}.$$

This together with (4) gives a formulation of the observation model

$$\tilde{\boldsymbol{s}}_t = W\Gamma_n C_k \left(\boldsymbol{\phi}_t - \tilde{\boldsymbol{\phi}}_t^{\text{DM}}\right) + \boldsymbol{\eta}_t,$$

where the matrix W defines the pupil mask for a nonrectangular telescope aperture (represented as a ring in Figure 3(a)).

In summary, we adopt the following model for the reconstruction of the high resolution wavefront:

$$\begin{aligned} &\min_{\Phi} \mathcal{J}_0^\beta(\Phi) := \frac{1}{2}\sum_{t=0}^{T}\left(\left\|W\Gamma_n C_k\left(\boldsymbol{\phi}_t - \tilde{\boldsymbol{\phi}}_t^{\text{DM}}\right) - \tilde{\boldsymbol{s}}_t\right\|^2 + \beta\left\|H_1\boldsymbol{\phi}_t\right\|^2\right), \\ &\text{subject to} \quad \boldsymbol{\phi}_0 = M_{\boldsymbol{v},t}\boldsymbol{\phi}_t, \qquad t = 0,\dots,T, \end{aligned} \tag{9}$$

where Φ is the column stacking of the high resolution phases $\boldsymbol{\phi}_0,\dots,\boldsymbol{\phi}_T$, β is the regularization parameter that has to be chosen according to the noise level and the strength of the atmospheric turbulence, and H_1, Γ_n, C_k, $\tilde{\boldsymbol{s}}_t$, $\tilde{\boldsymbol{\phi}}_t^{\text{DM}}$ being the previously defined weighting matrix, coarse SH-WFS operator, coarsening operator, derived pseudo-WFS measurements, and DM shape interpolated to a fine grid, respectively. Note that the motion operator $M_{\boldsymbol{v},t}$ is known as we assume $\boldsymbol{v}$ is given, but in general $\boldsymbol{v}$ is not known precisely.

To connect our method with Bayesian models, let us consider replacing H_1 by the square root of inverse covariance matrix $C_{\boldsymbol{\phi}}^{-1/2}$ and impose the following conditions:

- the noise $\tilde{\boldsymbol{\eta}}_t := \tilde{\boldsymbol{s}}_t - W\Gamma\boldsymbol{\phi}_t^{\text{res}}$ is white Gaussian;
- $\boldsymbol{\phi}_t$ is a realization of Gaussian random variables with covariance matrix $C_{\boldsymbol{\phi}}$ and zero mean;
- the regularization parameter β is properly chosen;
- the wind velocity $\boldsymbol{v}$ satisfies that the components of $t\boldsymbol{v}$ are multiples of the grid spacing,

then the solution to our model can be interpreted as the conditional mean or the minimum mean squared error estimate under the Bayesian framework. The corresponding minimizer in this case is also known as the maximum a posteriori estimate.

3.2.2. Adaptive wind velocities. We now move on to the more realistic situation, where only an estimate of the wind velocity is available, and we need to account for these imprecise measurements. Let us denote the estimate of the wind velocity by $\boldsymbol{v}^{(0)}$. We need to modify the model such that it allows variation of this estimate. For imprecise information on wind speed and direction, the associated motion operator, denoted by $M_{\boldsymbol{v}^{(0)},t}$, becomes imprecise. The model in (9) relies on the assumption that

$$\boldsymbol{\phi}_0 - M_{\boldsymbol{v}^{(0)},t}\boldsymbol{\phi}_t \approx 0,$$

which does not hold when $\boldsymbol{v}^{(0)}$ is not close enough to $\boldsymbol{v}$ and $|t| \gg 1$, due to the resulting cumulative error in the motion operator $M_{\boldsymbol{v}^{(0)},t}$. In fact, $M_{\boldsymbol{v}^{(0)},t}$ defines a displacement of $\boldsymbol{v}^{(0)} \cdot t$ and the error in the displacement is $(\boldsymbol{v} - \boldsymbol{v}^{(0)}) \cdot t$. In this case, an approximation based on the successive difference, relating $\boldsymbol{\phi}_{t+1}$ to its predecessor rather than to the first instance ϕ_0,

$$M_{\boldsymbol{v}^{(0)},t}\boldsymbol{\phi}_t - M_{\boldsymbol{v}^{(0)},t+1}\boldsymbol{\phi}_{t+1} \approx 0,$$

is more reliable. This helps to get a better estimate of $\boldsymbol{v}$ from inexact wind velocity data.

Furthermore, in order to preserve the consistency between the (pseudo)measurement and the TFF hypothesis with given wind velocity which is not accurate enough, the wind velocity is considered as a variable in the functional. Therefore, it can be adjusted appropriately during the minimization process.

In summary, we solve the joint minimization problem for both an estimated high resolution phase and an estimated wind velocity

$$\min_{\Phi,\boldsymbol{v}} \mathcal{J}_1^{\beta}(\Phi,\boldsymbol{v}) := \frac{1}{2}\sum_{t=0}^{T}\left(\left\|W\Gamma_n C_k\left(\boldsymbol{\phi}_t - \tilde{\boldsymbol{\phi}}_t^{\text{DM}}\right) - \tilde{\boldsymbol{s}}_t\right\|^2 + \beta\left\|H_1\boldsymbol{\phi}_t\right\|^2\right), \tag{10}$$
$$\text{subject to} \quad M_{\boldsymbol{v},t}\boldsymbol{\phi}_t = M_{\boldsymbol{v},t+1}\boldsymbol{\phi}_{t+1}, \qquad t = 0,\dots,T-1.$$

3.2.3. Existence of a minimizer for adaptive wind velocities. The objective function $\mathcal{J}_1^{\beta}$ is strongly convex with respect to Φ. If $\boldsymbol{v}$ is fixed, then one can easily conclude that it has a unique minimizer. However, the involvement of $\boldsymbol{v}$ makes (10) a nonconvex problem and the properties of minimizers less straightforward. For the theoretical analysis of the minimizers, here we consider the optimization problem in a continuous setting. Specifically, the problem

(10) is reformulated as

$$\begin{aligned} &\min_{\phi,\boldsymbol{v}} \mathcal{J}_1^{\beta}(\phi,\boldsymbol{v}) := \frac{1}{2}\sum_{t=0}^{T}\left(\left\|W\Gamma\left(\phi_t(\cdot)-\phi_t^{\mathrm{DM}}(\cdot)\right)-\tilde{\boldsymbol{s}}_t\right\|^2+\beta\left\|(\Delta+\epsilon\mathrm{Id})\phi_t(\cdot)\right\|^2_{L^2(\mathbb{R}^2)}\right), \\ &\text{subject to} \quad \phi_t(\cdot+\boldsymbol{v}t)=\phi_{t+1}(\cdot+\boldsymbol{v}(t+1)), \qquad t=0,\dots,T-1, \end{aligned} \tag{11}$$

in which Δ is the Laplacian operator, Id denotes an identity operator, W restricts the data to the telescope aperture, Γ is a bounded linear operator from $H^1(\mathbb{R}^2)$ to $\mathbb{R}^{n^2\times 2}$, and $\phi_t(\cdot)$ is the continuous version of $\boldsymbol{\phi}_t$. For a given time step t, $\phi_t(\cdot)$ is a function mapping from $\mathbb{R}^2$ to $\mathbb{R}$.

As described in section 3.1.2, Δ is an approximation to the (scaled) inverse covariance operator of the phase. Correspondingly, we seek for the solution of (11) in the subspace H^2 of $H^{11/6}$. In the subsequent discussion, we consider $\phi_t(\cdot)$ to be in the Sobolev space $H^2(\mathbb{R}^2)$ for $t=0,1,\dots,T$.

Theorem 1. *Let $\beta,\epsilon>0$ and $\phi_t^{\mathrm{DM}}(\cdot)\in H^1(\mathbb{R}^2)$. Assuming that the wind velocity is bounded $\|\boldsymbol{v}\|\le m$ for some constant m, then the minimization problem in (11) has a minimizer in $(H^2(\mathbb{R}^2))^{T+1}\times B_m$, where $B_m:=\{\boldsymbol{v}\in\mathbb{R}^2 \mid \|\boldsymbol{v}\|\le m\}$ is a ball in $\mathbb{R}^2$.*

Proof. Let us rewrite the constraint in (11) in an equivalent form as $\phi_t(\cdot+\boldsymbol{v}t)=\phi_0(\cdot)$ for $t=1,\dots,T$. It is sufficient to prove that the objective function

$$\hat{\mathcal{J}}_1^{\beta}(\phi_0,\boldsymbol{v}) := \frac{1}{2}\sum_{t=0}^{T}\left(\left\|W\Gamma\left(\phi_0(\cdot-\boldsymbol{v}t)-\phi_t^{\mathrm{DM}}(\cdot)\right)-\tilde{\boldsymbol{s}}_t\right\|^2+\beta\left\|(\Delta+\epsilon\mathrm{Id})\phi_0(\cdot-\boldsymbol{v}t)\right\|^2_{L^2(\mathbb{R}^2)}\right) \tag{12}$$

has a minimizer in $H^2(\mathbb{R}^2)\times B_m$. Clearly $\hat{\mathcal{J}}_1^{\beta}$ is lower bounded. Let $(\phi_0^{(0)},\boldsymbol{v}^{(0)})$, $(\phi_0^{(1)},\boldsymbol{v}^{(1)}),\dots$ be a minimizing sequence of $\hat{\mathcal{J}}_1^{\beta}$ in $H^2(\mathbb{R}^2)\times B_m$. Then $\{\hat{\mathcal{J}}_1^{\beta}(\phi_0^{(i)},\boldsymbol{v}^{(i)}) \mid i=0,1,\dots\}$ is bounded. As for any i we have

$$\begin{aligned} \frac{2}{\beta}\hat{\mathcal{J}}_1^{\beta}\left(\phi_0^{(i)},\boldsymbol{v}^{(i)}\right) &\ge \left((\Delta+\epsilon\mathrm{Id})\phi_0^{(i)},(\Delta+\epsilon\mathrm{Id})\phi_0^{(i)}\right)_{L^2(\mathbb{R}^2)} \\ &= \left(\Delta\phi_0^{(i)},\Delta\phi_0^{(i)}\right)_{L^2(\mathbb{R}^2)}+2\epsilon\left(\nabla\phi_0^{(i)},\nabla\phi_0^{(i)}\right)_{L^2(\mathbb{R}^2)}+\epsilon^2\left(\phi_0^{(i)},\phi_0^{(i)}\right)_{L^2(\mathbb{R}^2)}, \end{aligned} \tag{13}$$

the sequence $\{\phi_0^{(i)}\}$ is bounded in $H^1(\mathbb{R}^2)$. Furthermore, given that $\phi_0^{(i)}\in H^2(\mathbb{R}^2)$, $|\phi_0^{(i)}|^2_{H^2(\mathbb{R}^2)}\le \xi(\Delta\phi_0^{(i)},\Delta\phi_0^{(i)})_{L^2(\mathbb{R}^2)}$ for some constant ξ. Therefore, $\{\phi_0^{(i)}\}$ is a bounded sequence in $H^2(\mathbb{R}^2)$. This observation, together with $\{\boldsymbol{v}^{(i)}\}$ being bounded in $\mathbb{R}^2$, implies that there exists a subsequence $\{(\phi_0^{(i_j)},\boldsymbol{v}^{(i_j)})\}$ and $(\phi_0^{(*)},\boldsymbol{v}^{(*)})$ satisfying

$$\phi_0^{(i_j)} \rightharpoonup \phi_0^{(*)} \quad \text{in} \quad H^2(\mathbb{R}^2) \tag{14}$$

and $\boldsymbol{v}^{(i_j)}\to\boldsymbol{v}^{(*)}\in B_m$.

(I) As a consequence of (14),

$$\underline{\lim}_{j\to\infty}\left(\Delta\phi_0^{(i_j)},\Delta\phi_0^{(i_j)}\right)_{L^2(\mathbb{R}^2)}\geq\left(\Delta\phi_0^{(*)},\Delta\phi_0^{(*)}\right)_{L^2(\mathbb{R}^2)},$$
$$\underline{\lim}_{j\to\infty}\left(\nabla\phi_0^{(i_j)},\nabla\phi_0^{(i_j)}\right)_{L^2(\mathbb{R}^2)}\geq\left(\nabla\phi_0^{(*)},\nabla\phi_0^{(*)}\right)_{L^2(\mathbb{R}^2)},$$

and also

$$\underline{\lim}_{j\to\infty}\left(\phi_0^{(i_j)},\phi_0^{(i_j)}\right)_{L^2(\mathbb{R}^2)}\geq\left(\phi_0^{(*)},\phi_0^{(*)}\right)_{L^2(\mathbb{R}^2)}.$$

Then it holds that $\underline{\lim}_{j\to\infty}\|(\Delta+\epsilon\mathrm{Id})\phi_0^{(i_j)}\|^2_{L^2(\mathbb{R}^2)}\geq\|(\Delta+\epsilon\mathrm{Id})\phi_0^{(*)}\|^2_{L^2(\mathbb{R}^2)}$ according to the equality in (13). Moreover, based on the invariant property

$$\left\|(\Delta+\epsilon\mathrm{Id})\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(i_j)}t\right)\right\|^2_{L^2(\mathbb{R}^2)}=\left\|(\Delta+\epsilon\mathrm{Id})\phi_0^{(i_j)}\right\|^2_{L^2(\mathbb{R}^2)},$$

we get that for $t=0,2,\ldots,T$,

$$\underline{\lim}_{j\to\infty}\left\|(\Delta+\epsilon\mathrm{Id})\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(i_j)}t\right)\right\|^2_{L^2(\mathbb{R}^2)}\geq\left\|(\Delta+\epsilon\mathrm{Id})\phi_0^{(*)}\left(\cdot-\boldsymbol{v}^{(*)}t\right)\right\|^2_{L^2(\mathbb{R}^2)}. \tag{15}$$

(II) To obtain a similar result to (15) for the data fidelity term, we need to prove that for a fixed t,

$$\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(i_j)}t\right)\rightharpoonup\phi_0^{(*)}\left(\cdot-\boldsymbol{v}^{(*)}t\right)\quad\text{in}\quad H^1\left(\mathbb{R}^2\right). \tag{16}$$

We split the proof into a few steps.

(i) Recalling that $\phi_0^{(i_j)}\in H^2(\mathbb{R}^2)\subset H^1(\mathbb{R}^2)$, we have

$$\left\|\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(i_j)}t\right)-\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(*)}t\right)\right\|_{L^2(\mathbb{R}^2)}\leq t\left\|\boldsymbol{v}^{(*)}-\boldsymbol{v}^{(i_j)}\right\|\cdot\left|\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(*)}t\right)\right|_{H^1(\mathbb{R}^2)}. \tag{17}$$

(ii) For any $g\in L^2(\mathbb{R}^2)$, as $\phi_0^{(i_j)}$ are uniformly bounded in $H^1(\mathbb{R}^2)$, it follows from (17) that $(g,\phi_0^{(i_j)}(\cdot-\boldsymbol{v}^{(i_j)}t)-\phi_0^{(i_j)}(\cdot-\boldsymbol{v}^{(*)}t))_{L^2(\mathbb{R}^2)}\to 0$ as $j\to\infty$. The weak convergence of $\phi_0^{(i_j)}$ in $H^1(\mathbb{R}^2)$ implies that

$$(g,\phi_0^{(i_j)}(\cdot-\boldsymbol{v}^{(*)}t)-\phi_0^{(*)}(\cdot-\boldsymbol{v}^{(*)}t))_{L^2(\mathbb{R}^2)}=(g(\cdot+\boldsymbol{v}^{(*)}t),\phi_0^{(i_j)}-\phi_0^{(*)})_{L^2(\mathbb{R}^2)}\to 0.$$

(iii) Following (ii), we have

$$\begin{aligned}&\left(g,\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(i_j)}t\right)-\phi_0^{(*)}\left(\cdot-\boldsymbol{v}^{(*)}t\right)\right)_{L^2(\mathbb{R}^2)}\\&\quad=\left(g,\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(i_j)}t\right)-\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(*)}t\right)\right)_{L^2(\mathbb{R}^2)}\\&\qquad+\left(g,\phi_0^{(i_j)}\left(\cdot-\boldsymbol{v}^{(*)}t\right)-\phi_0^{(*)}\left(\cdot-\boldsymbol{v}^{(*)}t\right)\right)_{L^2(\mathbb{R}^2)}\to 0\end{aligned} \tag{18}$$

for any $g\in L^2(\mathbb{R}^2)$.

(iv) Finally, the same arguments in (i), (ii), and (iii) apply to all the first order weak derivatives of $\phi_0^{(i_j)}(\cdot - \boldsymbol{v}^{(i_j)}t)$ and $\phi_0^{(*)}(\cdot - \boldsymbol{v}^{(*)}t)$, as the weak derivatives are also in $H^1(\mathbb{R}^2)$ and, thus, clearly in $L^2(\mathbb{R}^2)$. This together with (18) gives the weak convergence (16).

(v) Since the linear operator $W\Gamma$ is bounded and $\phi_t^{\mathrm{DM}}(\cdot) \in H^1(\mathbb{R}^2)$, the weak convergence in (16) admits that the sequence $W\Gamma(\phi_0^{(i_j)}(\cdot - \boldsymbol{v}^{(i_j)}t) - \phi_t^{\mathrm{DM}}(\cdot))$ converges to $W\Gamma(\phi_0^{(*)}(\cdot - \boldsymbol{v}^{(i*)}t) - \phi_t^{\mathrm{DM}}(\cdot))$ and, hence

$$\lim_{j\to\infty} \left\| W\Gamma\left(\phi_0^{(i_j)}\left(\cdot - \boldsymbol{v}^{(i_j)}t\right) - \phi_t^{\mathrm{DM}}(\cdot)\right) - \tilde{\boldsymbol{s}}_t \right\| = \left\| W\Gamma\left(\phi_0^{(*)}\left(\cdot - \boldsymbol{v}^{(*)}t\right) - \phi_t^{\mathrm{DM}}(\cdot)\right) - \tilde{\boldsymbol{s}}_t \right\|.$$

(III) Following (I) and (II), we have $(\phi_0^{(*)}, \boldsymbol{v}^{(*)}) \in H^2(\mathbb{R}^2) \times B_m$ such that

$$\hat{\mathcal{J}}_1^{\beta}(\phi_0^{(*)}, \boldsymbol{v}^{(*)}) \leq \hat{\mathcal{J}}_1^{\beta}(\phi_0^{(i_j)}, \boldsymbol{v}^{(i_j)})$$

for all j. So $(\phi_0^{(*)}, \boldsymbol{v}^{(*)})$ is a minimizer of (12) which completes the proof. ∎

3.3. Atmospheric tomography with an inexact operator. Up to now we focused on the incoming phase $\boldsymbol{\phi}_t(\boldsymbol{x})$ yielded by one single thin atmospheric turbulence layer. In a more realistic situation, the distorted phase $\boldsymbol{\phi}_t(\boldsymbol{x})$ is a result of the 3-dimensional atmospheric turbulence. It is commonly considered that the distortions only happen at some discrete heights in the Earth's atmosphere (see, e.g., [31]). In such a setting, the incoming phase is a projection of the phase distortions at these layers, i.e.,

$$\boldsymbol{\phi}_t(\boldsymbol{x}) = \sum_{l=1}^{L} \boldsymbol{\phi}_{l,t}(\boldsymbol{x}), \tag{19}$$

where $\boldsymbol{\phi}_{l,t}(\boldsymbol{x})$ is the distortion of the phase corresponding to the lth layer and L is the number of layers (see, e.g., [21]). However, we cannot simply use the existing methods as they do not foresee using a time series of WFS data, but several WFS measurements at the same time and, additionally, these methods give a coarse resolution of the wavefront only. Therefore, we need to adapt the model to perform a time dynamic tomographic reconstruction.

Assuming the distortions of the phases $\boldsymbol{\phi}_{l,t}$ are independent of each other, for different l, the TFF is valid for the translational motion on each layer, i.e.,

$$\boldsymbol{\phi}_{l,t}(\boldsymbol{x}) = \boldsymbol{\phi}_{l,0}(\boldsymbol{x} - t\boldsymbol{v}_l) \tag{20}$$

in which $\boldsymbol{v}_l \in \mathbb{R}^2$ is the wind velocity of the lth layer. If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_L$ are different, then $\boldsymbol{\phi}_t(\cdot)$ does not preserve its form as t changes, i.e., (6) does not hold. Our idea to obtain a high resolution incoming phase is first reconstructing $\boldsymbol{\phi}_{l,t}(\boldsymbol{x})$ in high resolution with the TFF hypothesis (20), and then doing projections to get $\boldsymbol{\phi}_t(\boldsymbol{x})$ according to (19). Note that this results in a tomography problem as in an AO system where only information about $\boldsymbol{\phi}_t(\boldsymbol{x})$ is available. Figure 4(a) displays an example of the tomography problem in a three-layer setting. Unfortunately, with some uncertainty in the wind velocities, the exact unknown-to-data operator is unavailable. Small perturbations on the operator due to the errors in the wind velocities are illustrated in Figure 4(b).

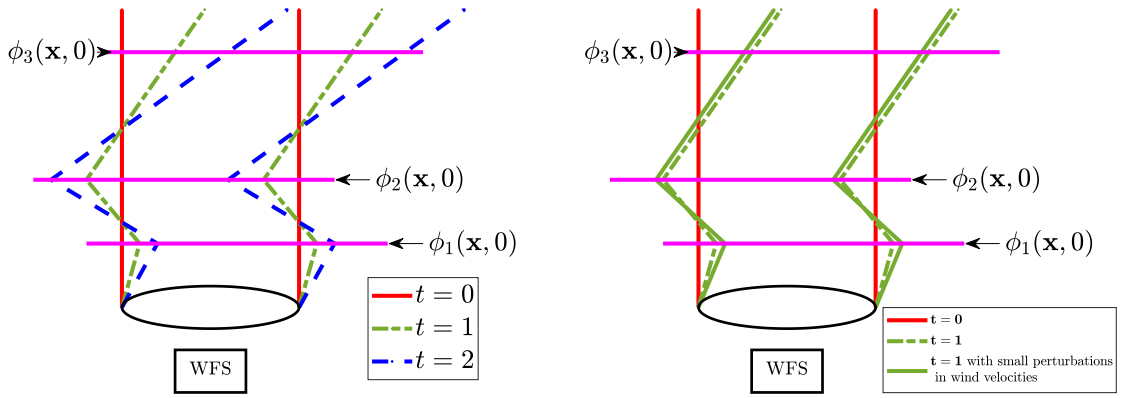


Figure 4. (a) *The tomography problem in a three-layer case. The winds of the layers shift the observed area of the distorted phase with the time, as illustrated by the bended paths corresponding to* $t = 1$ *and* $t = 2$. (b) *The path is perturbed by the inexact wind velocity, resulting in an inaccurate tomography operator.*

Let Φ be a concatenation of $\boldsymbol{\phi}_{1,0}, \boldsymbol{\phi}_{1,1}, \ldots, \boldsymbol{\phi}_{L,T}$. As a direct extension of the previous model, we consider the following optimization problem:

$$\min_{\Phi,\boldsymbol{v}} \mathcal{J}_2^{\beta}(\Phi,\boldsymbol{v}) := \frac{1}{2}\sum_{t=0}^{T}\left(\left\|W\Gamma_n C_k\left(\sum_{l=1}^{L}\boldsymbol{\phi}_{l,t}-\tilde{\boldsymbol{\phi}}_t^{\mathrm{DM}}\right)-\tilde{\boldsymbol{s}}_t\right\|^2+\beta\sum_{l=1}^{L}\|H_l\boldsymbol{\phi}_{l,t}\|^2\right), \tag{21}$$
$$\text{subject to}\quad M_{\boldsymbol{v}_l,t}\boldsymbol{\phi}_{l,t}=M_{\boldsymbol{v}_l,t+1}\boldsymbol{\phi}_{l,t+1},\qquad t=0,\ldots,T-1,\quad l=1,2,\cdots,L,$$

where $\boldsymbol{v}$ is a column stacking of $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_L$, $H_l=\frac{1}{c_l}(H+\epsilon I)$, and $\sum_{l=1}^{L}c_l^2=1$, where c_l^2 are the so-called c_n^2-values of the atmospheric layers (see section 5.1) . Once the minimizer of (21), denoted by $(\{\hat{\boldsymbol{\phi}}_{l,t}\},\hat{\boldsymbol{v}})$, is obtained, the residual phase $\boldsymbol{\phi}_t^{\mathrm{res}}$ is estimated as $\sum_{l=1}^{L}\hat{\boldsymbol{\phi}}_{l,t}-\tilde{\boldsymbol{\phi}}_t^{\mathrm{DM}}$ according to (19) and (1).

4. Numerical minimization. A numerical minimization procedure is needed to compute the estimate $\boldsymbol{\phi}_t^{\mathrm{res}}$ from (10) or, respectively, (21). Let us consider the one-layer case (10) first. It is clear that the optimization problem is not convex as the constraints are not convex functions in $\boldsymbol{v}$. To deal with such a problem, we implement an alternating direction method to find a minimizer of the objective function with respect to Φ and $\boldsymbol{v}$.

We define $M_{\boldsymbol{v}}$ and $\tilde{M}_{\boldsymbol{v}}$ as block matrices of the form

$$M_{\boldsymbol{v}}=\begin{bmatrix}M_{\boldsymbol{v},0} & & \\ & \ddots & \\ & & M_{\boldsymbol{v},T-1}\quad 0\end{bmatrix},\quad \tilde{M}_{\boldsymbol{v}}=\begin{bmatrix}0 & M_{\boldsymbol{v},1} & \\ & \ddots & \\ & & M_{\boldsymbol{v},T}\end{bmatrix},$$

and let $A_{\boldsymbol{v}}=M_{\boldsymbol{v}}^T(\tilde{M}_{\boldsymbol{v}}-M_{\boldsymbol{v}})$, where the superscript T denotes the transpose of a matrix. Then the constraint in (10) can be equivalently written as $A_{\boldsymbol{v}}\Phi=0$. The corresponding augmented

Lagrangian functional is given by

$$\mathcal{L}_1(\Phi, \boldsymbol{v}, \boldsymbol{\lambda}) = \mathcal{J}_1^{\beta}(\Phi, \boldsymbol{v}) + \boldsymbol{\lambda}^* A_{\boldsymbol{v}} \Phi + \frac{\tau}{2} \|A_{\boldsymbol{v}} \Phi\|^2,$$

in which $\tau \geq 0$ is the augmented Lagrangian parameter (see, e.g., [5]). Starting from $\boldsymbol{v} = \boldsymbol{v}^{(0)}$, an alternating update on Φ, $\boldsymbol{v}$, and $\boldsymbol{\lambda}$ leads to estimates of the phase and wind velocities. To be specific, the method requires solving the following two subproblems alternatively:

$$\Phi^{(k+1)} = \arg\min_{\Phi} \left(\mathcal{J}_1^{\beta}(\Phi, \boldsymbol{v}^{(k)}) + \frac{\tau}{2} \left\| A_{\boldsymbol{v}^{(k)}} \Phi + \frac{\boldsymbol{\lambda}^{(k)}}{\tau} \right\|^2 \right), \tag{22}$$

$$\boldsymbol{v}^{(k+1)} = \arg\min_{\boldsymbol{v}} \frac{\tau}{2} \left\| A_{\boldsymbol{v}} \Phi^{(k+1)} + \frac{\boldsymbol{\lambda}^{(k)}}{\tau} \right\|^2 \tag{23}$$

with $\boldsymbol{\lambda}^{(k)}$ being updated as

$$\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} + \tau A_{\boldsymbol{v}^{(k+1)}} \Phi^{(k+1)}. \tag{24}$$

The subproblem (22) is a least squares problem, the solution of which is obtained by solving a linear system. The second subproblem (23) is more difficult as it is neither linear nor convex. We will show that it is related to an image registration problem and the optical flow algorithm proposed by Gilliam and Blu [19, 18] can be adapted for finding a good approximation to the solution.

4.1. All-pass filters for wind estimation. First, we consider the explicit objective function of subproblem (23), i.e.,

$$\left\| A_{\boldsymbol{v}} \Phi^{(k+1)} + \frac{\boldsymbol{\lambda}^{(k)}}{\tau} \right\|^2 = \sum_{t=0}^{T-1} \left\| M_{\boldsymbol{v},t}^T M_{\boldsymbol{v},t+1} \boldsymbol{\phi}_{t+1}^{(k+1)} - M_{\boldsymbol{v},t}^T M_{\boldsymbol{v},t} \boldsymbol{\phi}_t^{(k+1)} + \boldsymbol{\lambda}_t^{(k)} / \tau \right\|^2.$$

To reduce the complexity of notations, let $\boldsymbol{p}_{\boldsymbol{v},t} := M_{\boldsymbol{v},t}^T M_{\boldsymbol{v},t+1} \boldsymbol{\phi}_{t+1}^{(k+1)}$ and $\boldsymbol{q}_{\boldsymbol{v},t} := M_{\boldsymbol{v},t}^T M_{\boldsymbol{v},t} \boldsymbol{\phi}_t^{(k+1)}$ and $\boldsymbol{v} := \boldsymbol{v}^{(k)} + \boldsymbol{u}$. In the frequency domain, for any displacement vector $\boldsymbol{u} = [u_1, u_2]^T$,

$$\widehat{\boldsymbol{p}}_{\boldsymbol{v}^{(k)}+\boldsymbol{u},t} \approx \widehat{\boldsymbol{f}}_{\boldsymbol{u}} \circ \widehat{\boldsymbol{p}}_{\boldsymbol{v}^{(k)},t} \quad \text{and} \quad \widehat{\boldsymbol{q}}_{\boldsymbol{v}^{(k)}+\boldsymbol{u},t} \approx \widehat{\boldsymbol{q}}_{\boldsymbol{v}^{(k)},t}, \tag{25}$$

where $\widehat{\boldsymbol{f}}_{\boldsymbol{u}} := [e^{i(\omega_1 u_1 + \omega_2 u_2)}]_{\boldsymbol{\omega}}$ for $\boldsymbol{\omega} = [\omega_1, \omega_2]^T \in \mathbb{R}^2$, $\widehat{\boldsymbol{p}}_{\boldsymbol{v},t}$ is the Fourier transform of $\boldsymbol{p}_{\boldsymbol{v},t}$, and $\circ$ denotes the Hadamard product. It is clear that the filter $\widehat{\boldsymbol{f}}_{\boldsymbol{u}}$ is an all pass filter. From (25), the objective function

$$\begin{aligned} \left\| A_{\boldsymbol{v}^{(k)}+\boldsymbol{u}} \Phi^{(k+1)} + \frac{\boldsymbol{\lambda}}{\tau} \right\|^2 &\approx \sum_{t=0}^{T-1} \left\| \widehat{\boldsymbol{f}}_{\boldsymbol{u}} \circ \widehat{\boldsymbol{p}}_{\boldsymbol{v}^{(k)},t} - \widehat{\boldsymbol{q}}_{\boldsymbol{v}^{(k)},t} + \widehat{\boldsymbol{\lambda}}_t^{(k)} / \tau \right\|^2 \\ &= \sum_{t=0}^{T-1} \left\| \widehat{\boldsymbol{f}}_{\boldsymbol{u}/2} \circ \widehat{\boldsymbol{p}}_{\boldsymbol{v}^{(k)},t} - \widehat{\boldsymbol{f}}_{-\boldsymbol{u}/2} \circ \left(\widehat{\boldsymbol{q}}_{\boldsymbol{v}^{(k)},t} - \widehat{\boldsymbol{\lambda}}_t^{(k)} / \tau \right) \right\|^2. \end{aligned} \tag{26}$$

The problem is converted into registering $\boldsymbol{p}_{\boldsymbol{v}^{(k)},t}$ and $\boldsymbol{q}_{\boldsymbol{v}^{(k)},t} - \boldsymbol{\lambda}_t^{(k)}/\tau$ for all t simultaneously. The last equality of (26) expresses the shifting motion of $\boldsymbol{u}$ in a forward-backward form, and later we will show that such an expression is crucial in constructing an all pass filter like $\widehat{\boldsymbol{f}_{\boldsymbol{u}}}$.

We use the idea of the local all pass (LAP) algorithm [18] which is based on an approximation of the filter $\boldsymbol{f}$ in a space with basis $\{\boldsymbol{f}^{(j)}\}$. A typical choice of the filter basis is $\boldsymbol{f}^{(j)} := [f^{(j)}]_{x_1,x_2}$, where

$$f^{(0)} = \exp\left(-\frac{x_1^2+x_2^2}{2\xi^2}\right), \quad f^{(1)} = x_1 f^{(0)}, \quad \text{and} \quad f^{(2)} = x_2 \circ f^{(0)}.$$

Here ξ is a parameter that can be determined according to the size of displacement. A larger basis is possible, but here we consider only the 3-dimensional case. The problem of estimating $\boldsymbol{f}_{\boldsymbol{u}}$ in (26) is linearized into finding an optimal filter of the form $\boldsymbol{f}^\dagger = (\sum_{j=0}^2 a_j^\dagger \boldsymbol{f}^{(j)})$ with $\boldsymbol{a}^\dagger = [a_1^\dagger, a_2^\dagger, a_3^\dagger]^T$ being a solution to

$$\min_{\boldsymbol{a}\in\mathbb{R}^3, a_1=1} \sum_{t=0}^{T-1} \left\| \left(\sum_{j=0}^2 a_j \boldsymbol{f}^{(j)}\right) * \boldsymbol{p}_{\boldsymbol{v}^{(k)},t} - \left(\sum_{j=0}^2 a_j \boldsymbol{f}^{(-j)}\right) * \left(\boldsymbol{q}_{\boldsymbol{v}^{(k)},t} - \boldsymbol{\lambda}_t^{(k)}/\tau\right) \right\|^2, \tag{27}$$

in which $*$ denotes the convolution operation and $\boldsymbol{f}^{(-j)} = [f^{(j)}(-x_1,-x_2)]_{x_1,x_2}$. The minimization can be done by solving a small linear system in $\boldsymbol{a}$. Note that up to a scaling constant $\widehat{\boldsymbol{f}^\dagger}$ approximates $\widehat{\boldsymbol{f}}_{\boldsymbol{u}/2}$, and $\widehat{\boldsymbol{f}^{(-\dagger)}}$, defined as the Fourier transform of $\boldsymbol{f}^{(-\dagger)} := [\sum a^\dagger f^{(j)}(-x_1,-x_2)]_{x_1,x_2}$, approximates $\widehat{\boldsymbol{f}}_{-\boldsymbol{u}/2}$. Therefore $\widehat{\boldsymbol{f}_{\boldsymbol{u}}}$, as an all pass filter, is approximated by

$$\left[\sum_{j=0}^2 a_j^\dagger \widehat{f^{(j)}}(\omega_1,\omega_2) \Big/ \sum_{j=0}^2 a_j^\dagger \widehat{f^{(j)}}(-\omega_1,-\omega_2)\right]_{\boldsymbol{\omega}}$$

which is also an all pass filter.

Finally, the new wind velocity is $\boldsymbol{v}^{(k+1)} := \boldsymbol{v}^{(k)} + \boldsymbol{u}^{(k+1)}$ and the update $\boldsymbol{u}^{(k+1)}$ is computed as

$$u_1^{(k+1)} = \frac{2\sum_{x_1,x_2}\sum_{j=0}^2 x_1 a_j^\dagger f^{(j)}(x_1,x_2)}{\sum_{x_1,x_2}\sum_{j=0}^2 a_j^\dagger f^{(j)}(x_1,x_2)} \quad \text{and} \quad u_2^{(k+1)} = \frac{2\sum_{x_1,x_2}\sum_{j=0}^2 x_2 a_j^\dagger f^{(j)}(x_1,x_2)}{\sum_{x_1,x_2}\sum_{j=0}^2 a_j^\dagger f^{(j)}(x_1,x_2)}. \tag{28}$$

The method described here is a simplified version of the LAP algorithm, as the LAP algorithm [18] was designed to estimate a motion field, i.e., $\boldsymbol{u}$ is a vector field rather than a constant vector. To do this, the algorithm computes the velocity for each pixel by restricting the problem in a small window centered at the pixel under the assumption that the motion field within each window is nearly constant. However, this treatment is unnecessary in our problem, since $\boldsymbol{u}$ is a constant velocity according to the TFF assumption. Hence we solve (27) for a single motion vector over the entire domain.

4.2. The alternating direction method. In summary, the minimization process (23) results in a self-adaptive algorithm which updates the inexact wind velocity with the information

Algorithm 1

Input: ϕ_t^{DM} for $t = 0, \dots, T$, $\boldsymbol{v}^{(0)}$

Output: Estimates of residual phases $\hat{\phi}_0^{\mathrm{res}}, \dots, \hat{\phi}_T^{\mathrm{res}}$ and the estimated wind velocity $\hat{\boldsymbol{v}}$.

Step 1. Initialization.
1. recover $\phi_t^{\mathrm{res,\ coarse}}$ from the DM data ϕ_t^{DM},
2. compute the high resolution DM shape $\tilde{\phi}_t^{\mathrm{DM}}$ from ϕ_t^{DM},
3. compute the pseudo-WFS measurement $\tilde{\boldsymbol{s}}_t$ with $\phi_t^{\mathrm{res,\ coarse}}$ by (5),
4. set $\boldsymbol{\lambda}^{(0)} = 0$.

Step 2. For $k = 1, 2, \dots, m$, repeat:
1. Compute $\Phi^{(k)}$ as minimizer of the Lagrangian functional

$$\mathcal{J}_1^{\beta}(\Phi, \boldsymbol{v}^{(k-1)}) + \left(\boldsymbol{\lambda}^{(k-1)}\right)^* A_{\boldsymbol{v}^{(k-1)}}\Phi + \frac{\tau}{2}\left\|A_{\boldsymbol{v}^{(k-1)}}\Phi\right\|^2$$

   over all Φ. Here $\Phi^{(k)}$ is a column stacking of $\phi_0^{(k)}, \dots, \phi_T^{(k)}$.
2. Estimate the wind velocity $\boldsymbol{v}^{(k)} = \boldsymbol{v}^{(k-1)} + \boldsymbol{u}^{(k)}$, where $\boldsymbol{u}^{(k)}$ is given by (28).
3. $\boldsymbol{\lambda}^{(k)} = \boldsymbol{\lambda}^{(k-1)} + \tau A_{\boldsymbol{v}^{(k)}}\Phi^{(k)}$.

Step 3. The output is given by $\hat{\phi}_0^{\mathrm{res}} = \phi_0^{(m)} - \tilde{\phi}_0^{\mathrm{DM}}, \dots, \hat{\phi}_T^{\mathrm{res}} = \phi_T^{(m)} - \tilde{\phi}_T^{\mathrm{DM}}$, $\hat{\boldsymbol{v}} = \boldsymbol{v}^{(m)}$.

from the observed data based on an implicit use of the TFF hypothesis. The algorithm is described in Algorithm 1.

Note that Algorithm 1 is not an exact alternating direction method of multipliers (see, e.g., [5]) as the optimization problem has a nonlinear constraint and the subproblem in $\boldsymbol{v}$ is linearized as in (27). We investigate the convergence of the method from initial wind velocities with errors in section 5.4. In real applications, only values of $\boldsymbol{\phi}_t$ on the telescope aperture are useful. Although in our model $\boldsymbol{\phi}_t$ can be defined on a domain that is much bigger than the telescope aperture, it is not necessary to do so as the part outside the telescope aperture does not contribute to the blur of the observed image. For computational reasons, $\boldsymbol{\phi}_t$ is defined on a grid in a small rectangular domain covering the telescope aperture.

4.3. Solving the tomography problem. Let us now consider problem (21) which is also nonconvex. We introduce an algorithm similar to Algorithm 1 for the minimization.

We define $A_{\boldsymbol{v}}$ as a block diagonal matrix with diagonal blocks $A_{\boldsymbol{v}_1}, A_{\boldsymbol{v}_2}, \dots, A_{\boldsymbol{v}_L}$. Then the equality constraints in (21) have a compact form $A_{\boldsymbol{v}}\Phi = 0$. The corresponding augmented Lagrangian functional is given by

$$\mathcal{L}_2(\Phi, \boldsymbol{v}, \boldsymbol{\lambda}) = \mathcal{J}_2^{\beta}(\Phi, \boldsymbol{v}) + \boldsymbol{\lambda}^* A_{\boldsymbol{v}}\Phi + \frac{\tau}{2}\left\|A_{\boldsymbol{v}}\Phi\right\|^2.$$

Again, the idea of updating $\Phi, \boldsymbol{v}, \boldsymbol{\lambda}$ alternatively can be applied, and the main issue remains in the subproblem of minimizing $\mathcal{L}_2$ with respect to $\boldsymbol{v}$. For fixed $\Phi^{(k+1)}$ and $\boldsymbol{\lambda}^{(k)}$, the subproblem can be decomposed into

$$\min_{\boldsymbol{v}_l} \frac{\tau}{2}\left\|A_{\boldsymbol{v}_l}\Phi_l^{(k+1)} + \frac{\boldsymbol{\lambda}_l^{(k)}}{\tau}\right\|^2, \quad l = 1, \dots, L, \tag{29}$$

where $\Phi_l^{(k+1)}$ and $\boldsymbol{\lambda}_l^{(k)}$ are subvectors of $\Phi^{(k+1)}$ and $\boldsymbol{\lambda}^{(k)}$ corresponding to the lth layer, $l = 1, 2, \ldots, L$. The problem (29) has the same form as (23), which means that the update of $\boldsymbol{v}$ can be performed in a layer-by-layer manner using the technique described in subsection 4.1. We use the update rule for $\boldsymbol{\lambda}$ from (24). The resulting method is summarized in Algorithm 2.

Algorithm 2

Input: $\boldsymbol{\phi}_t^{\mathrm{DM}}$ for $t = 0, \ldots, T$, $\boldsymbol{v}^{(0)}$

Output: Estimates of residual phases $\hat{\boldsymbol{\phi}}_0^{\mathrm{res}}, \ldots, \hat{\boldsymbol{\phi}}_T^{\mathrm{res}}$ and the estimated wind velocity $\hat{\boldsymbol{v}}$.

Step 1. Initialization.
1. recover $\boldsymbol{\phi}_t^{\mathrm{res,\ coarse}}$ from the DM data $\boldsymbol{\phi}_t^{\mathrm{DM}}$,
2. compute the high resolution DM shape $\tilde{\boldsymbol{\phi}}_t^{\mathrm{DM}}$ from $\boldsymbol{\phi}_t^{\mathrm{DM}}$,
3. compute the pseudo-WFS measurement $\tilde{\boldsymbol{s}}_t$ of $\boldsymbol{\phi}_t^{\mathrm{res,\ coarse}}$ by (5),
4. set $\boldsymbol{\lambda}^{(0)} = 0$.

Step 2. For $k = 1, 2, \ldots, m$, repeat:
1. Compute $\boldsymbol{\phi}_{1,0}^{(k)}, \ldots, \boldsymbol{\phi}_{1,T}^{(k)}, \ldots, \boldsymbol{\phi}_{L,0}^{(k)}, \ldots, \boldsymbol{\phi}_{L,T}^{(k)}$ as a minimizer of the Lagrangian functional
$$\mathcal{J}_2^{\beta}(\Phi, \boldsymbol{v}^{(k-1)}) + \left(\boldsymbol{\lambda}^{(k-1)}\right)^* A_{\boldsymbol{v}^{(k-1)}}\Phi + \frac{\tau}{2}\left\|A_{\boldsymbol{v}^{(k-1)}}\Phi\right\|^2$$
over all Φ. Concatenate the vectors $\boldsymbol{\phi}_{1,0}^{(k)}, \ldots, \boldsymbol{\phi}_{1,T}^{(k)}, \ldots, \boldsymbol{\phi}_{L,0}^{(k)}, \ldots, \boldsymbol{\phi}_{L,T}^{(k)}$ as $\Phi^{(k)}$.
2. For $l = 1, 2, \ldots, L$, compute the new estimated wind velocity at the lth layer as
$$\boldsymbol{v}_l^{(k)} = \boldsymbol{v}_l^{(k-1)} + \boldsymbol{u}_l^{(k)},$$
where $\boldsymbol{u}_l^{(k)}$ is given by (28).
3. $\boldsymbol{\lambda}^{(k)} = \boldsymbol{\lambda}^{(k-1)} + \tau A_{\boldsymbol{v}^{(k)}}\Phi^{(k)}$.

Step 3. The output is computed as $\hat{\boldsymbol{\phi}}_0^{\mathrm{res}} = \sum_{l=1}^{L} \boldsymbol{\phi}_{l,0}^{(m)} - \tilde{\boldsymbol{\phi}}_0^{\mathrm{DM}}, \ldots, \hat{\boldsymbol{\phi}}_T^{\mathrm{res}} = \sum_{l=1}^{L} \boldsymbol{\phi}_{l,T}^{(m)} - \tilde{\boldsymbol{\phi}}_T^{\mathrm{DM}}$, $\hat{\boldsymbol{v}} = \boldsymbol{v}^{(m)}$.

5. Simulation results.

In order to validate our algorithm, we used a MATLAB-based AO simulation tool to obtain the required data. As a benefit of simulation, we can compare the phase recovered on a fine grid to the true incoming phase. As no alternative algorithms to solve this problem exist in the literature, we demonstrate the reconstruction quality by comparing the results to the ground truth. Additionally, we compare our method in section 5.4 to the use of optical flow constraints.

5.1. Simulation setting.

The simulated system is an SCAO system on an 8 m telescope, equipped with one 40×40 SH-WFS, described in Table 2. We perform simulations for different atmospheric settings to demonstrate the power of our method. As a starting point, we consider an atmosphere consisting just of one layer with a wind speed of 15 m/s in direction 0°. In the

Table 2
Description of the simulated SCAO system.

Telescope diameter	8 m
Central obstruction	2.2 m
1 NGS SH-WFS	40×40 subapertures
WFS wavelength λ	0.589 μ m
WFS integration time	2 ms
1 DM at height 0 m	closed loop
DM actuator spacing	0.2 m

Table 3
Nine-layer median atmosphere.

Layer	1	2	3	4	5	6	7	8	9
Height(m)	47	140	281	562	1125	2250	4500	9000	18000
c_n^2-profile	0.522	0.026	0.044	0.116	0.098	0.029	0.059	0.043	0.06
Wind speed	15	13	13	9	9	15	25	40	21
Direction	0	$\pi/2$	π	$3\pi/2$	0	$\pi/2$	π	$3\pi/2$	0

next steps, first we change the direction to 30° and, second, we move to a two layer profile, with relative strengths 0.65 and 0.35. The altitudes of the layers are 0 and 5000 m, the wind speeds 15 m/s, and the directions 0° and 90°, respectively.

As a last step, we take an atmospheric profile with nine layers, often referred to as an *ESO standard profile* from [26]. This model is based on measurements at ESO's site Paranal in the Atacama desert with a Fried parameter $r_0 = 12.9$ cm. In Table 3, the values for the nine-layer medium seeing atmosphere are given.

The AO loop is controlled using the cumulative reconstructor with domain decomposition algorithm [32, 33] with optimized loop gain. We take the first 20 frames of an AO simulation using these parameters to show the performance of our algorithm.

Note that using more than 20 frames might improve the results further, however, it also clearly increases the computational costs. The maximum number of frames that can be used is related to the wind speed. If the atmospheric turbulence seen in the first frame is blown over the whole telescope and no part of it can be seen anymore in the Tth frame, it is reasonable to stop taking into account any frame after T steps. Note that, therefore, it is required that the frozen flow hypothesis holds for at least T frames. For the setting of Table 2 and a wind speed of 15 m/s, we have $T = 267$. However, one may also think of reconstructing the atmosphere on an ever bigger domain, and thus take into account more than T frames. In real observations this idea will most likely fail as the TFF hypothesis will not hold for 267 frames being equivalent to more than 0.5 s of real time. Therefore, we did not follow this idea further as it also increases the computational costs.

5.2. Numerical considerations. The residual of the incoming phase $\boldsymbol{\phi}_t^{\text{res}}(\boldsymbol{x})$ is computed by the simulation software on a fine level with 400×400 pixels across the telescope aperture, translating into 2 cm/pixel. The SH-WFS has 40×40 subapertures and the DM 41×41 actuators, which limits the resolution of the reconstructed incoming phase for the usual AO control algorithm to this level as the DM cannot use higher resolved incoming phases.

Our method is able to provide a high resolution reconstruction, however, we cannot choose it arbitrarily large as the resolution of the reconstruction should still be smaller than the resolution of the simulated incoming phase, in order to avoid an inverse crime. In the following simulations, the underlying grid for the high resolution reconstruction is chosen to be four times finer than the grid of the DM actuators, resulting in a 161×161 reconstruction.

To judge the quality of our fine resolution reconstruction, we compare it to the ground truth. As a quality criterion, we take the l_2-norms of the reconstruction error, i.e.,

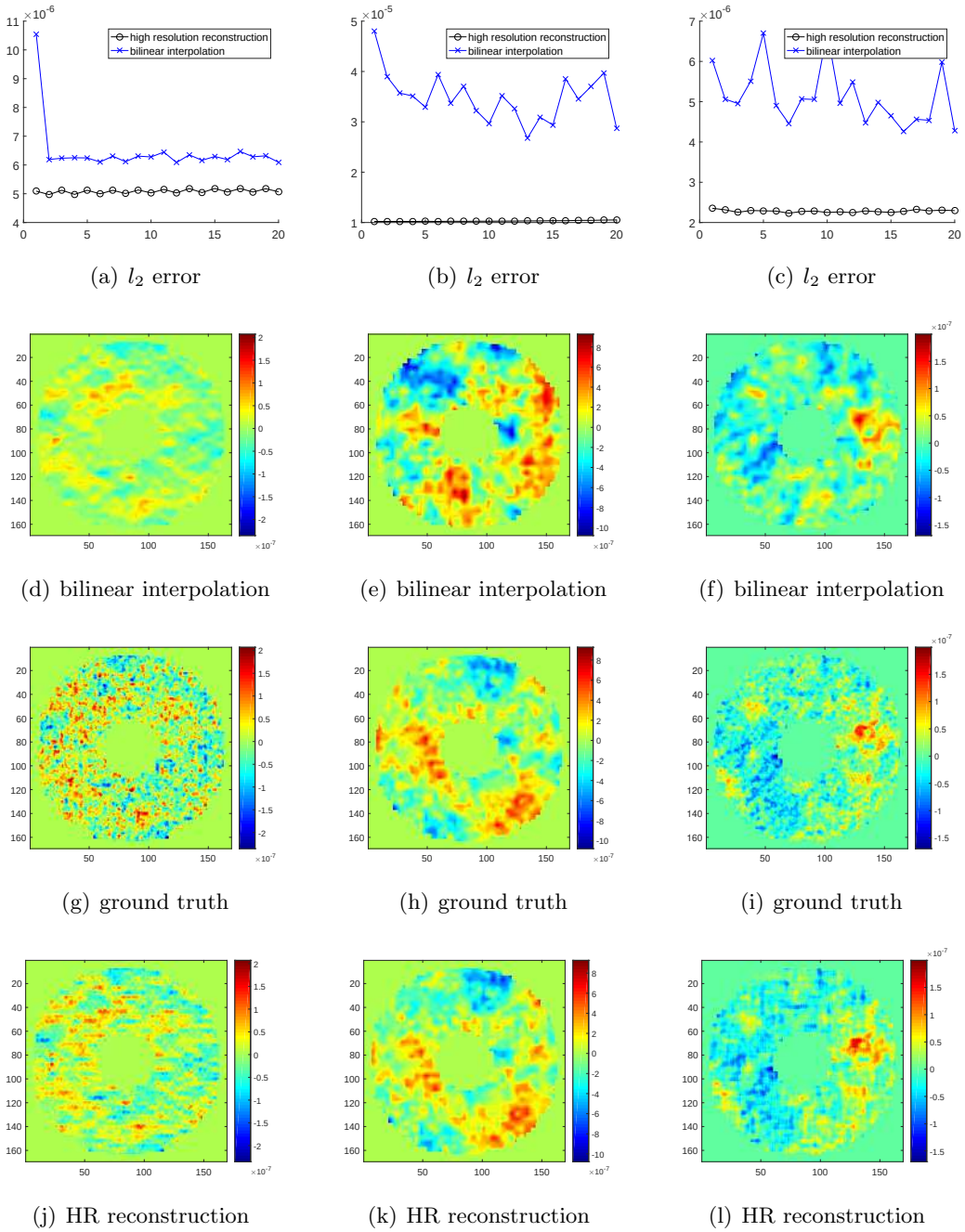
$$\|\hat{\boldsymbol{\phi}}_t^{\text{res}} - \boldsymbol{\phi}_t^{\text{res}}\|_2, \tag{30}$$

where t indicates the time step and $\hat{\boldsymbol{\phi}}_t^{\text{res}}$ is the result of our algorithm. Due to the different resolutions, we have to interpolate the simulated $\boldsymbol{\phi}_t^{\text{res}}$ to the same grid as $\hat{\boldsymbol{\phi}}_t^{\text{res}}$. For this purpose, we use the MATLAB function *interp*2. To make a comparison, we also interpolate $\boldsymbol{\phi}_t^{\text{res, coarse}}$ to the same grid as $\hat{\boldsymbol{\phi}}_t^{\text{res}}$ and compute the corresponding l_2-error. Since only the part of the phase inside the telescope aperture contributes to the image quality, in (30), $\hat{\boldsymbol{\phi}}_t^{\text{res}}$ and $\boldsymbol{\phi}_t^{\text{res}}$ are restricted to the telescope domain with the piston mode (i.e., the additive constant complement) removed.

5.3. Numerical results. In this section we present the numerical results from several test runs. We always use the first 20 frames for our algorithm to compute the high resolution residual phase. We will demonstrate the performance of our algorithm by first assuming that exact wind velocities are given and solving problem (9). After that, we consider the situation where the exact values of wind velocities are unavailable which usually happens in real applications. Algorithms 1 and 2 with adaptive wind velocities are employed to find the high resolution reconstructions in this case.

5.3.1. Exact wind velocity case. To highlight the strength of our method, we start with the simplest possible case, i.e., simulation 1 (S1): All atmospheric turbulence is located in one layer close to the ground, with known wind speed and direction, and a bright star can be used as a guide star. In this case the used AO control algorithm provides a correction in the K-band close to the diffraction limit. Using our model (9), we can still reduce the l_2-error between the true residual phase $\boldsymbol{\phi}_t^{\text{res}}$ and the reconstructed residual phase $\boldsymbol{\phi}_t^{\text{res, coarse}}$ by 15 to 20%; see Figure 5(a). The l_2-error of the bilinear interpolation at the first time step is much bigger than the others because the control loop of the system has to be closed and no other information of the phase is available at this time within the AO control algorithm. Our high resolution (HR) reconstruction still results in a small l_2-error at the first time step since we have used the data from the following time steps as well. The l_2-error curve of our reconstruction has significantly less oscillation over the time than the l_2-error curve of the bilinear interpolation, which is reasonable as our reconstructed phase does not rely heavily on data of a specific time step and hence prevents enormous errors.

As a next step, in simulation 2 (S2), we change the guide star flux to simulate a faint star. This leads to a decreased correction quality through the AO control algorithm but leaves more room for improvement of the reconstructed residual phase. With low flux the real time reconstruction for the adaptive system has correction errors that are bigger than the reconstruction with high flux. So for the low flux case, the AO system is unable to use the



(a) l_2 error (b) l_2 error (c) l_2 error

(d) bilinear interpolation (e) bilinear interpolation (f) bilinear interpolation

(g) ground truth (h) ground truth (i) ground truth

(j) HR reconstruction (k) HR reconstruction (l) HR reconstruction

Figure 5. *First column: simulation* 1*; second column: simulation* 2*; third column: simulation* 3*. The plots in the first row are l_2-errors of the reconstruction, with the x-axis being the frame number. All the images from the second row to the last row are corresponding to $t = 16$ (i.e., at the 16th time step) and show wavefronts in m.*

data efficiently enough, and we expect that a more significant improvement can be made by our algorithm which takes full consideration of the relation of the data from all time steps. Indeed, our method reduces the residual l_2-error by 70 to 80%, as shown in Figure 5(b).

In the first two simulations the wind direction was parallel to an axis, which gives more information for HR reconstruction in one direction than in the other. The best possible situation would be a wind direction of 45° to the axes, leading to an equally spread gain of information in x- and y-direction. However, such a direction might not occur in practice, thus we choose an angle of 30° with respect to the first axis for simulation 3 (S3). Changing again to high flux, even in this case our algorithm gives an improvement of 45% to 65%, shown in Figure 5(c). Compared to the accuracy improvement for S1 (which is around 20%), the improvement for this one is much more significant. However, this is reasonable as the sampling points from different time steps have less overlap. Recall that the angle between the wind velocity and the first axis is 30°, which ensures that the trajectories of coarse grid points do not have an overlap with each other. In S1, in contrast, the coarse grid points always fall in $n+1$ lines parallel to the first axis where n is the number of grid point on the second axis.

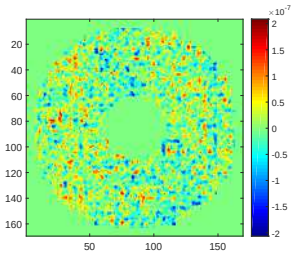
In Figures 5(d)–5(l), the images of the reconstruction by bilinear interpolation on the DM data, the true phase in HR and the HR reconstruction by the proposed method are given. As expected, the bilinear interpolations do not include the fine details of the residual phase well. Our approach is able to recover some missing high frequency components in the data, therefore, shows an improvement from the bilinear interpolation.

However, the HR reconstruction is limited by several factors. First, the time interval in which the TFF hypothesis holds is typically small. This limits the amount of low resolution data and therefore sets an upper bound for the resolution of the reconstruction. Second, the low resolution data are not acquired in a random manner. Instead, the motion vector is constant and the coarse grid points propagate in a fixed direction with a constant rate, which means that the two-dimensional phase is not evenly sampled in a short time.

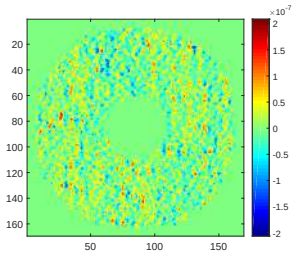
To take a closer look, Figure 6 displays the images of reconstruction errors for simulations 1 and 3. The error images of the HR reconstruction have some special patterns. In Figure 6(b), the pixels with a bigger error are concentrated on some lines parallel to the vertical line, while in Figure 6(d) the relatively big errors are mainly distributed on the lines having an angle of 30 degrees with the vertical line. The directions of the lines coincide with the directions of the winds in both cases, which indicates that the phase at some grid points is not detected as well as other points over the process.

We take two additional steps to get closer to a realistic on-sky scenario. For a two-layer atmosphere and high photon flux, i.e., simulation 4 (S4), the performance of our algorithm is a bit worse than in the one-layer case, as now the problem becomes a tomography problem. As the wind directions for the two layers are different, recovering the information on each layer is much more challenging. Still, the improvement by of our algorithm ranges from 20% to 30%. The results for this case are plotted in Figure 7(a).

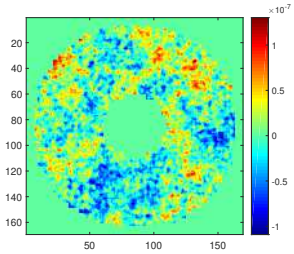
The final step is to use the ESO standard profile from Table 3 in simulation 5 (S5). In this case, the AO control algorithm still performs on a high level. Applying our algorithm leads to an improvement between 10% and 25%, showing that our model can cope with the multilayered nature of the atmosphere. The results are shown in Figure 7(b).



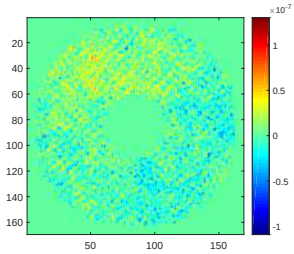
(a) bilinear interpolation for simulation 1



(b) our method for simulation 1

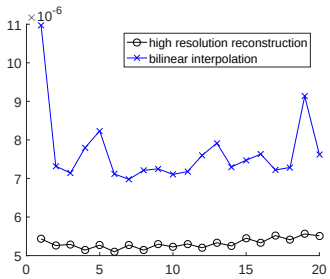


(c) bilinear interpolation for simulation 3

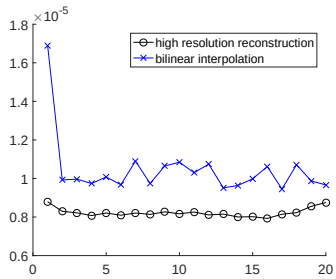


(d) our method for simulation 3

Figure 6. *Reconstruction error images at $t = 16$ of the bilinear interpolation and our method.*



(a)



(b)

Figure 7. (a) *and* (b) *are the l_2-errors of two methods for simulations* 4 *and* 5*, respectively. The x-axis is the frame number.*

5.3.2. Reconstructions with inexact wind velocities. We further study the performance of the proposed algorithms when the initial wind velocities (IWVs) are imprecise. The adaptive wind velocity (AWV) algorithm is used. To make the results comparable to the previous reconstruction, we let the simulation settings be the same as before except that the precise wind velocities (PWVs) are replaced by the imprecise ones. The proposed method is employed on the five simulations and the IWVs are assumed to have 10% to 20% error for all simulated cases.

For the first three simulations, Algorithm 1 is used and one wind velocity is computed for each simulation. The l_2-errors of the HR reconstructions are shown in Figures 8(a)–8(c). For

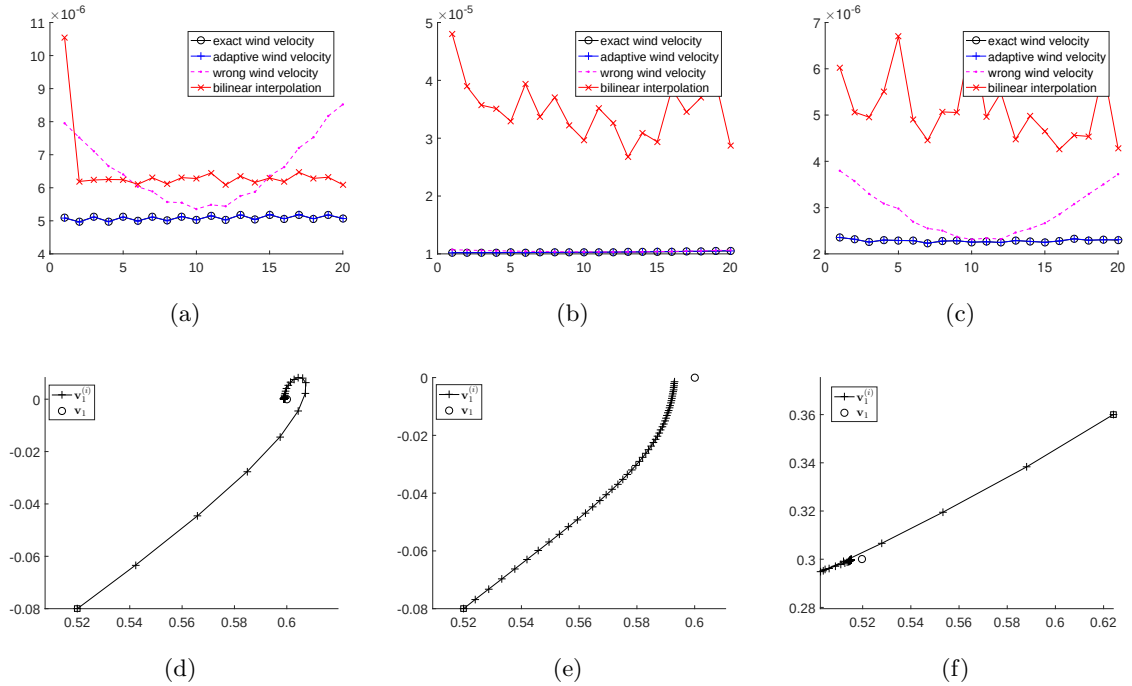


Figure 8. (a)–(c): *The l_2-errors of the reconstructions for simulations* 1 *to* 3*, respectively. In these plots the x-axis is the frame number and y-axis is the l_2-error. The exact wind speed means reconstruction with fixed PWV, while the wrong wind speed means using IWV without adaption.* (d)–(e): *the AWV for simulations* 1 *to* 3*. The unit of the x- and y-axis is pixels per time step, given as* 25 *m/s (i.e., the number of HR pixels (with length* 0.05 *m) per time step (being* 1/500 *s).*

comparison the previous test results on exact wind velocity cases are also displayed. According to the figures, with IWV the minimizer of (9) is computed and the corresponding l_2-error is significantly bigger than the ones obtained with the exact wind velocity at the first and last few frames, except for the low flux case (i.e., the second simulation). An IWV does not play a key role in the reconstruction accuracy in the low flux case where the observed data are very unreliable.

The estimated wind velocities at all iterations are plotted in Figures 8(d)–8(f). The unit of the x- and y-axis is relating the number of HR pixels (with length 0.05 m) to the length of a time step, being 1/500 s, thus giving 25 m/s. In the high flux cases, Algorithm 1 starts with an IWV that has around 20% error, but finally outputs improved ones with less than 1% error. This implies that the algorithm is capable of extracting the wind velocity information from the AO data itself, and explains why it is still able to have a comparable reconstruction accuracy to the one obtained with the exact wind velocity, as shown Figures 8(a) and 8(c). In the low flux case, the error of the estimated wind velocity is reduced from 13.4% to around 1.4% as shown in Figure 8(e).

For simulations 1 and 3, the error images of the reconstruction from the last subsection, the reconstruction by the proposed algorithm, and the reconstruction with fixed IWV are given in Figure 9. Significantly larger errors can be observed from the algorithm without AWVs

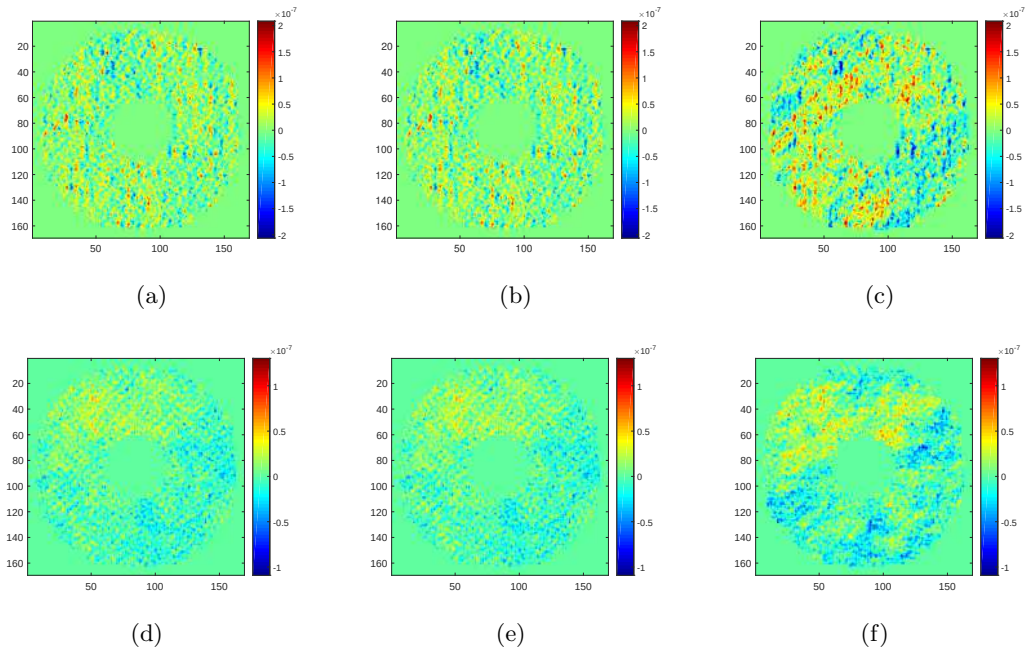
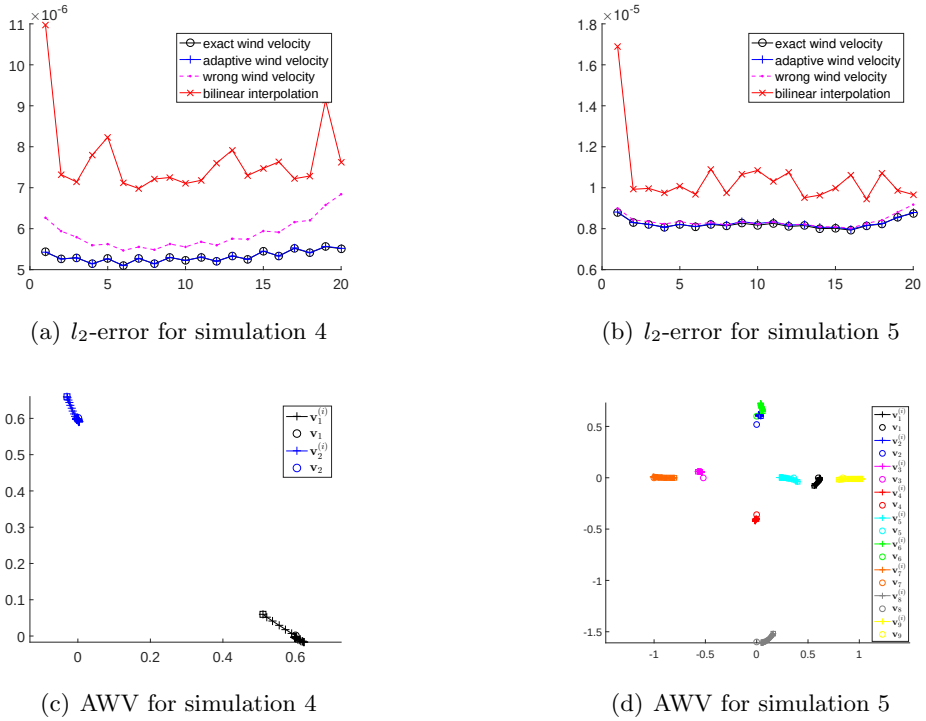


Figure 9. *Error images of different reconstructions at* $t = 16$. (a)–(c): *With exact wind velocity, with AWV, and without AWV for simulation* 1. (d)–(f): *With exact wind velocity, AWV, and without AWV for simulation* 3.

while the proposed algorithm with adaptive velocities results in errors that look similar to the reconstructions with PWV.

For the two-layer case, i.e., simulation 4, we used Algorithm 2 which solves for the HR residual phase as well as the wind velocity for each of the two layers. IWVs $\{\boldsymbol{v}_l^{(0)}\}$, having around 10% to 20% error, are given. A reconstruction from (9) using the wrong wind velocities $\{\boldsymbol{v}_l^{(0)}\}$ is also computed and a large error is observed at the first and last few frames as shown in Figure 10(a). The reconstruction at the middle frames, however, has a close accuracy to reconstruction for the exact wind velocity case. On the other hand, Algorithm 2 with only inexact wind velocities $\{\boldsymbol{v}_l^{(0)}\}$ gives almost the same accuracy as the solution of (9) with exact wind velocities. In Figure 10(c), the wind velocity estimates for each of the two layers are displayed. It suggests that our algorithm can recover information of the wind velocities from the DM data itself using imprecise IWV.

Finally, Algorithm 2 is used to estimate estimates of the nine wind velocities for nine layers in the fifth simulation. Though in this case the l_2-error of the minimizer of (9) with the fixed IWV is still reasonably small compared to the minimizer with the PWV, Algorithm 2 can still make an improvement with the implementation of AWV. The resulting accuracy is shown in Figure 10(b) and the estimated wind velocities for all layers are given in Figure 10(d). The algorithm fails to distinguish the layers with close wind velocities, such as the second layer and the sixth layer. The isolated wind velocities, however, are captured well, such as that of the fourth layer and the eighth layer. The reconstruction does not depend heavily on separating



(a) l_2-error for simulation 4

(b) l_2-error for simulation 5

(c) AWV for simulation 4

(d) AWV for simulation 5

Figure 10. *The l_2-errors and the estimated wind velocities for simulation 4 and 5.*

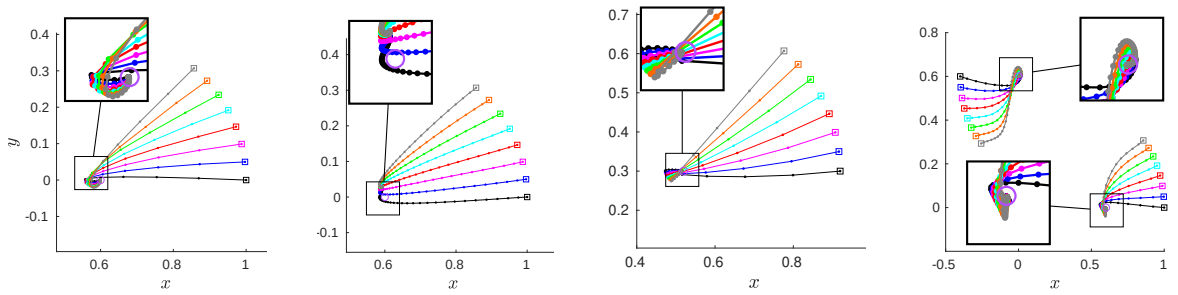
the wind velocities of those layers with similar motions, as the reconstruction accuracy shown in Figure 10(b) is still comparable to the one obtained with the PWV.

5.4. Numerical convergence and comparison of our method. As discussed in section 4.2, the constraints of our method are nonconvex. Therefore no convergence to a global minimum is guaranteed for our method theoretically. We demonstrate that our method converges numerically to the correct solution for different levels of errors in the IWV and compare our method to the optical flow constraints.

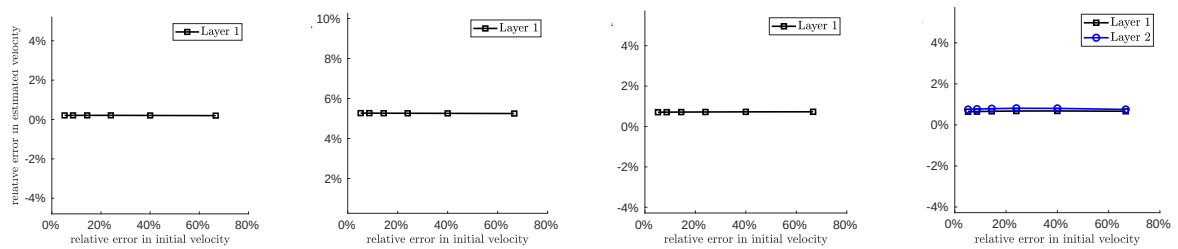
5.4.1. Different levels of error in the initial wind speed. In this subsection we test different initialization and different levels of noise in wind velocities and investigate their impacts on the algorithm.

For each simulation, 8 different initializations with 66.7% error (i.e., the squares in Figure 11(a)) are applied. The trajectories of the estimated wind velocity by the algorithm are plotted in Figure 11(a). The iterations of the wind velocities converge to a point close to the ground truth value (i.e., the circle in Figure 11(a)) consistently for all the simulations.

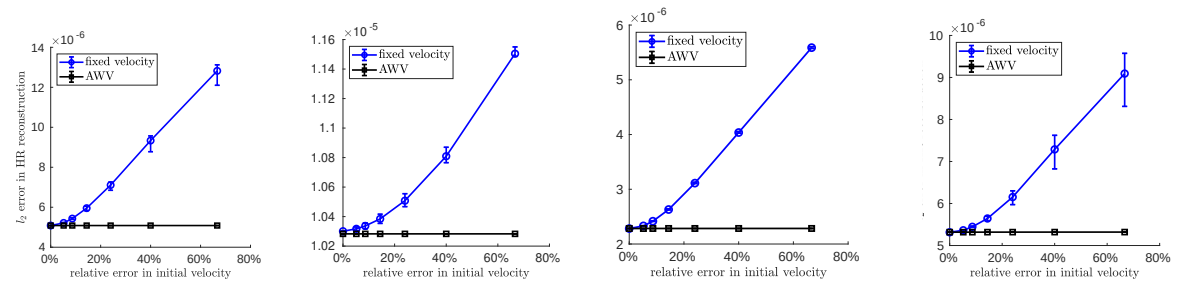
We also compared the accuracy of wind estimation under different levels of errors in IWV (66.7%, 40%, 24%, 14.4%, 8.6%, 5.2%, and 0%, respectively). The estimation errors are plotted against the initialization errors in Figure 11(b) for the four simulations, respectively. This clearly shows that the final errors in wind velocity are stable with respect to the levels of initial errors.



(a). Trajectories of wind estimates starting from different IWV (squares)



(b). l_2-errors of the estimated wind velocities plotted versus the l_2-errors in the IWV



(c). The l_2-error of the HR reconstruction plotted versus the error in IWV

Figure 11. *Reconstruction accuracy with different wind velocity initialization and different levels of IWV errors for simulation* 1 *(1st column), simulation* 2 *(2nd column), simulation* 3 *(3rd column), and simulation* 4 *(4th column), respectively. The results for* (b) *are averaged from* 8 *different initializations at each error level of IWV. The error bars in* (c) *reflect the impacts of* 8 *initializations.*

It is also interesting to note that the error of computed HR residual phase by our (joint phase reconstruction and wind estimation) approach remains small despite different levels of initial wind errors (ranging from 66.7% to 0%), i.e., the black curve in Figure 11(c). The reconstruction using the fixed wind velocity algorithm (i.e., the blue curve in Figure 11(c)), in contrast, significantly degraded as the error in the wind velocity increases.

5.4.2. Comparison with the optical flow constraints. We compare our model with the optical flow constraints (OFC) [6, 1], which is a well-known model for motion estimation in imaging problems. The OFC assumes that the function values (e.g., image intensity) remain

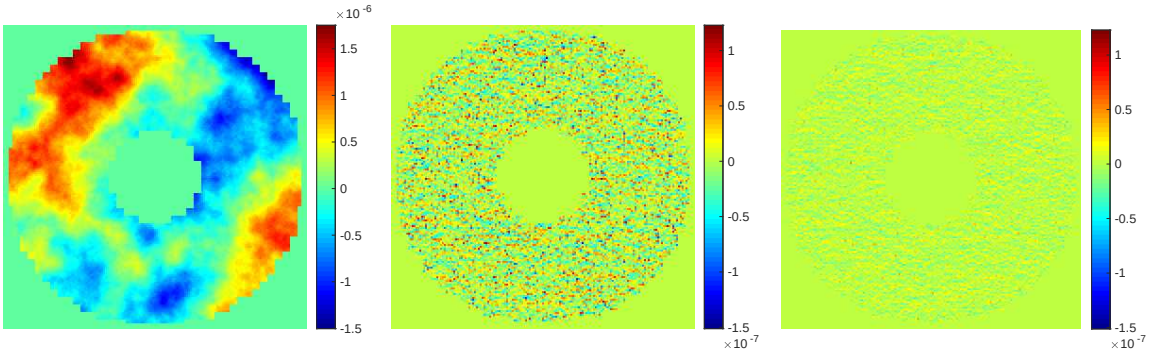


Figure 12. *The S3 open loop wavefront (left), residual of discrete OFC (middle), and ours (right).*

Table 4

l_2 reconstruction error under constraint OFC and SD with different wind velocity initializations.

Error in IWV	Constraint	S1	S2	S3	S4
66.7%	OFC	6.27×10^{-6}	1.04×10^{-5}	2.94×10^{-6}	6.42×10^{-6}
	SD	$\mathbf{5.08 \times 10^{-6}}$	$\mathbf{1.03 \times 10^{-5}}$	$\mathbf{2.28 \times 10^{-6}}$	$\mathbf{5.32 \times 10^{-6}}$
14.4%	OFC	6.34×10^{-6}	1.03×10^{-5}	2.94×10^{-6}	6.69×10^{-6}
	SD	$\mathbf{5.08 \times 10^{-6}}$	$\mathbf{1.03 \times 10^{-5}}$	$\mathbf{2.28 \times 10^{-6}}$	$\mathbf{5.32 \times 10^{-6}}$
0%	OFC	6.37×10^{-6}	1.03×10^{-5}	2.94×10^{-6}	6.82×10^{-6}
	SD	$\mathbf{5.08 \times 10^{-6}}$	$\mathbf{1.03 \times 10^{-5}}$	$\mathbf{2.28 \times 10^{-6}}$	$\mathbf{5.32 \times 10^{-6}}$

unchanged along the trajectory of a moving point and, therefore, satisfy

$$\frac{\partial \phi}{\partial t} + \nabla \phi \cdot \boldsymbol{v} = 0$$

for an image ϕ. This can be discretized as

$$\boldsymbol{\phi}_{t+1} - \boldsymbol{\phi}_t + \nabla_h \frac{\boldsymbol{\phi}_t + \boldsymbol{\phi}_{t+1}}{2} \cdot \boldsymbol{v} = 0 \tag{31}$$

in our setting where the measurements are taken at a fixed time frequency. Here ∇_h is the discrete gradient operator.

We first evaluate the approximation error of OFC (31) and our constraint (9) on a ground truth simulated phase. The residual of our constraint is much smaller than that of OFC (see Figure 12). We denote our constraint (9) by successive difference (SD). The comparison of these two approaches with different initial wind error levels (0%, 14.4%, and 66.7%) is given in Table 4. Our method achieves a better reconstruction error over different simulations and IWVs.

6. Conclusion. We investigated the problem of deriving an HR phase from coarse measurements in a closed loop AO system, arising from the quest of using such an HR phase in postprocessing of the data. After describing the system setting, we developed an approach for solving this problem with consideration of the model error due to imprecise observations (i.e.,

wind velocity). Our approach is based on the TFF hypothesis and uses subsequently applied DM shapes in a minimization procedure to obtain a HR phase. For known wind velocities in the different atmospheric layers the functional to be minimized describes a tomography problem and can be solved with standard tools.

As the wind velocities might not be known in practice, we also investigated the case of imprecisely known wind velocities. This results in an augmented Lagrangian functional, which can be solved using an alternating direction method. Finally, we demonstrated the power of our method in numerical experiments under different atmospheric conditions. The l_2-error is clearly reduced by our method compared to a simple bilinear interpolation from a coarse to a fine grid in all cases. For the more realistic setting with nine atmospheric layers and imprecise knowledge of the wind velocities, the adaptive method outperforms the method using constantly wrong wind velocities and is close to exact knowledge of the wind velocities.

In practice, this method might be useful on the one hand for temporal control, using the calculated HR phase to predict the incoming phase (see [30] and the references therein for details on temporal control), and on the other hand a highly resolved phase could be useful for postprocessing, e.g., in PSF reconstruction methods as in [39]. Bringing our method to this level requires investigating the behavior when scaled to the size of upcoming 40 m class telescopes like ESO's ELT. Furthermore, we will test our algorithm in end-to-end simulation tools for such telescopes to demonstrate the feasibility for the AO community in more detail.

Acknowledgment. We thank the anonymous reviewers for their thorough review helping us to improve the paper.

REFERENCES

[1] F. Becker, S. Petra, and C. Schörr, *Optical flow*, Handbook of Mathematical Methods in Imaging, Springer, New York, 2015, pp. 1945–2004.

[2] N. Bharmal, *Frozen flow or not? Investigating the predictability of the atmosphere*, J. Phys. Conf. Ser., 595 (2015), 012003.

[3] N. Bharmal, U. Bitenc, A. Basden, R. Myers, and N. Dipper, *A hierarchical wavefront reconstruction algorithm for gradient sensors*, in Third AO4ELT Conference – Adaptive Optics for Extremely Large Telescopes, 2013, https://doi.org/10.12839/AO4ELT3.13280.

[4] U. Bitenc, A. Basden, N. A. Bharmal, T. Morris, N. Dipper, E. Gendron, F. Vidal, D. Gratadour, G. Rousset, and R. Myers, *On-sky tests of the CuReD and HWR fast wavefront reconstruction algorithms with CANARY*, Month. Not. R. Astron. Soc., 448 (2015), pp. 1199–1205.

[5] B. 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 (2011), pp. 1–122.

[6] M. Burger, H. Dirks, and L. Frerking, *On optical flow models for variational motion estimation*, Variational Methods, Radon Ser. Comput. Appl. Math. 18, De Gruyter, Berlin 2017, pp. 225–251.

[7] R. Chan, Z. Shen, and T. Xia, *A framelet algorithm for enhancing video stills*, Appl. Comput. Harmon. Anal., 23 (2007), pp. 153–170.

[8] R. Chan, M. Xiao, and X. Wen, *A phase model for point spread function estimation in ground-based astronomy*, Sci. China Math., 56 (2013), pp. 2701–2710.

[9] R. Chan, X. Yuan, and W. Zhang, *Point-spread function reconstruction in ground-based astronomy by $l^1 - l^p$ model*, J Opt. Soc. Amer. A, 29 (2012), pp. 2263–2271.

[10] Q. Chu, S. Jefferies, and J. G. Nagy, *Iterative wavefront reconstruction for astronomical imaging*, SIAM J. Sci. Comput., 35 (2013), pp. S84–S103.

[11] M. Claire, *Introduction to adaptive optics and its history*, in American Astronomical Society 197th Meeting, AAS, Washington, DC, 2001, http://www.ucolick.org/~max/History_AO_Max.htm.
[12] R. Davies and M. Kasper, *Adaptive optics for astronomy*, Annu. Rev. Astron. Astrophys., 50 (2012), pp. 305–351.
[13] L. Dykes, R. Ramlau, L. Reichel, K. Soodhalter, and R. Wagner, *Lanczos-Based Fast Blind Deconvolution Methods*, Technical report, RICAM Report 2017-27, 2017.
[14] B. Ellerbroek, *Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques*, J. Opt. Soc. Amer. A, 19 (2002), pp. 1803–1816.
[15] B. Ellerbroek and C. Vogel, *Inverse problems in astronomical adaptive optics*, Inverse Problems, 25 (2009), 063001.
[16] S. Farsiu, M. Elad, and P. Milanfar, *Multiframe demosaicing and super-resolution of color images*, IEEE Trans. Image Process., 15 (2006), pp. 141–159.
[17] D. Fried, *Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements*, J. Opt. Soc. Amer., 67 (1977), pp. 370–375.
[18] C. Gilliam and T. Blu, *Local all-pass filters for optical flow estimation*, in 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, Piscataway, NJ, 2015, pp. 1533–1537.
[19] C. Gilliam and T. Blu, *Local all-pass geometric deformations*, IEEE Trans. Image Process., 27 (2018), pp. 1010–1025.
[20] J. Goodman, *Introduction to Fourier Optics*, Roberts and Company Publishers, Englewood, CO, 2005.
[21] T. Helin, S. Kindermann, and D. Saxenhuber, *Towards analytical model optimization in atmospheric tomography*, Math. Methods Appl. Sci., 40 (2016), pp. 1153–1169, https://doi.org/10.1002/mma.4042.
[22] T. Helin and M. Yudytskiy, *Wavelet methods in multi-conjugate adaptive optics*, Inverse Problems, 29 (2013), 085003, http://stacks.iop.org/0266-5611/29/i=8/a=085003.
[23] A. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
[24] S. Jefferies and M. Hart, *Deconvolution from wave front sensing using the frozen flow hypothesis*, Optics Express, 19 (2011), pp. 1975–1984.
[25] A. Kolmogorov, *The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers*, Dokl. Akad. Nauk SSSR, 30 (1941), pp. 299–303.
[26] M. L. Louarn, C. Vérinaud, V. Korkiakoski, N. Hubin, and E. Marchetti, *Adaptive optics simulations for the European Extremely Large Telescope*, in Proc. SPIE 6272, Advances in Adaptive Optics II, 2006.
[27] J. Nagy, S. Jefferies, and Q.Chu, *Fast PSF reconstruction using the frozen flow hypothesis*, in Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, 2010.
[28] K. Nasrollahi and T. Moeslund, *Super-resolution: A comprehensive survey*, Mach. Vision Appl., 25 (2014), pp. 1423–1468.
[29] A. Neubauer, *A new cumulative wavefront reconstructor for the Shack-Hartmann sensor*, J. Inverse Ill-Posed Prob., 21 (2013), pp. 451–476.
[30] M. Pöttinger, R. Ramlau, and G. Auzinger, *A new temporal control approach for SCAO systems*, Inverse Problems, 36 (2019), 015002.
[31] M. Roggemann, B. Welsh, and B. Hunt, *Imaging through Turbulence*, CRC, Boca Raton, FL, 1996.
[32] M. Rosensteiner, *Cumulative reconstructor: Fast wavefront reconstruction algorithm for Extremely Large Telescopes*, J. Opt. Soc. Amer. A, 28 (2011), pp. 2132–2138.
[33] M. Rosensteiner, *Wavefront reconstruction for extremely large telescopes via CuRe with domain decomposition*, J. Opt. Soc. Amer. A, 29 (2012), pp. 2328–2336.
[34] G. Taylor, *The spectrum of turbulence*, Roy. Soc. Proc. A Math. Phys. Eng. Sci., 164 (1938), pp. 476–490.
[35] E. Thiébaut and M. Tallon, *Fast minimum variance wavefront reconstruction for extremely large telescopes*, J. Opt. Soc. Amer. A, 27 (2010), pp. 1046–1059.
[36] J.-P. Véran, F. Rigaut, H. Maître, and D. Rouan, *Estimation of the adaptive optics long exposure point spread function using control loop data*, J. Opt. Soc. Amer. A, 14 (1997), pp. 3057–3069, https://doi.org/10.1364/JOSAA.14.003057.
[37] T. von Karman, *Mechanische ähnlichkeit und Turbulenz*, in 3rd International Congress of Applied Mechanics, Wiedmannsche Buchhandlung, Berlin, 1930.

[38] R. Wagner, *From Adaptive Optics systems to Point Spread Function Reconstruction and Blind Deconvolution for Extremely Large Telescopes*, Ph.D. thesis, Johannes Kepler University Linz, Linz, Austria, 2017.

[39] R. Wagner, C. Hofer, and R. Ramlau, *Point spread function reconstruction for single-conjugate adaptive optics*, J. Astron. Telescopes Instrum. Systems, 4 (2018), 049003, https://doi.org/10.1117/1.JATIS.4.4.049003.

[40] R. Wagner, A. Neubauer, and R. Ramlau, *Simulation results for a finite element-based cumulative reconstructor*, J. Astron. Telescopes Instrum. Systems, 3 (2017), 049001, https://doi.org/10.1117/1.JATIS.3.4.049001.

[41] H. Weichel, *Laser Beam Propagation in the Atmosphere*, Tutor. Texts Opt. Eng. TT #3, SPIE, Bellingham, WA, 1990.

[42] M. Yudytskiy, T. Helin, and R. Ramlau, *Finite element-wavelet hybrid algorithm for atmospheric tomography*, J. Opt. Soc. Amer. A, 31 (2014), pp. 550–560.

[43] M. Zhariy, A. Neubauer, M. Rosensteiner, and R. Ramlau, *Cumulative wavefront reconstructor for the Shack-Hartman sensor*, Inverse Prob. Imaging, 5 (2011), pp. 893–913.