A GRADIENT-BASED METHOD FOR ATMOSPHERIC TOMOGRAPHY

DANIELA SAXENHUBER* AND RONNY RAMLAU
Altenbergerstrasse 69
Linz, 4040, Austria
(Communicated by Xuecheng Tai)

Abstract. In Adaptive Optics (AO) systems for ground-based telescopes, one aims at mechanically correcting for atmospheric aberrations by means of quickly moving deformable mirrors. In complex AO systems, which are using several light sources and aim at a good reconstruction in a large field of view, the derivation of optimal mirror commands from measured light typically includes the problem of atmospheric tomography. As the computational effort for such a limited-angle tomography problem is strongly increasing for growing telescope sizes, fast algorithms are needed. We present a novel algorithm for atmospheric tomography that takes real-life effects such as tip/tilt indetermination, cone effect and spot elongation into account. Furthermore, we discuss two models for the tip and tilt components of an incoming wavefront and incorporate them into the reconstruction. We find a fast step size choice for our Gradient-based iteration and compare it with different existing step size choices. Numerical results are demonstrated for two different AO systems on a 42 m telescope, using the European Southern Observatory’s end-to-end simulation tool, OCTOPUS.

1. Introduction. The problem of atmospheric tomography arises in astronomical Adaptive Optics (AO), a technology which is primarily used in large ground-based telescopes in order to correct for atmospheric aberrations. Due to layers of warm and cold air, originally plane light wavefronts from astronomical objects get distorted while travelling through the Earth’ atmosphere, resulting in blurred images of observed astronomical objects. AO compensates for these atmospheric perturbations by means of deformable mirrors that are adapting in real time to the shape of the distortions. Thus, the reflected light wavefronts are flattened and image quality improves. For a detailed description of Adaptive Optics, we refer to, e.g., [28, 29].

In complex AO systems, several light sources are used: bright astronomical objects as Natural Guide Stars (NGS) and artificially created laser beacons as Laser Guide Stars (LGS). By means of wavefront sensors (WFS), the incoming light of each guide star is measured. The calculation of the shape of a deformable mirror from these WFS measurements typically involves a limited-angle tomography problem. Currently, the standard method for solving this ill-posed inverse problem is by a precomputation of the (generalized) inverse of the matrix that maps wavefront measurements to mirror shapes, the so called matrix-vector-multiplication (MVM). A thorough description of the mathematical challenges in this field can be found, e.g. in [8].

2010 Mathematics Subject Classification. Primary: 65F22, 65F20; Secondary: 65Y20, 85A04.
Key words and phrases. Regularization, iterative method, tomography.
* Corresponding author: Daniela Saxenhuber.
In particular, for the new generation of Extremely Large Telescopes (ELT), the computational complexity of the reconstruction algorithm becomes a major issue. The deformable mirrors have to be controlled in real-time, e.g. at around 500 Hertz. This reduces the time for one reconstruction step to 1 ms. For the presented simulation results, one second of real time, i.e. 500 time steps, are performed. The dimension of the planned European ELT (E-ELT), along with a realistic setting of six LGS and three NGS arranged in a field of view of one arc minute, would yield a system matrix of the order of $10^4 \times 10^4$ to be inverted in each time step.

To solve the atmospheric tomography problem in AO, many methods have been developed: In [10], the authors present a minimum-mean-square-error estimator, while [11] gives a back-projection tomography algorithm. Iterative solvers, such as conjugate gradient (CG) based methods are presented, e.g. in [6]. Efficient multigrid preconditioners were investigated in [6, 12, 13], whereas in [45, 43] Fourier domain preconditioners were found. Atmospheric reconstruction for future ELTs has to be performed within an extremely tight time frame, which is barely reachable even on future hardware. Recent reconstruction methods attempt to fulfil the stringent speed requirements. Methods for solving the Bayesian MAP estimator are the CG-based fractal iterative method (FrIM) [5, 35, 38, 37], and the wavelet-based Finite Element Wavelet Hybrid Algorithm (FEWHA) [47, 48, 46]. Additionally, the atmospheric tomography has been solved in a matrix-free approach via the Kaczmarz algorithm, presented in [24, 33].

This paper is based on the idea that the problem of determining the DM shapes from WFS data is split in three independent steps [24]: the reconstruction of the incoming wavefronts, the problem of atmospheric tomography with wavefronts as input data and the computation of the optimal mirror correction. Due to the limited time frame for determining the DM shapes, there is a growing need for methods that yield good qualitative performance with very low computational complexity. We propose a Gradient-based method to solve the subproblem of atmospheric tomography iteratively, in order to achieve a fast reconstruction. Furthermore, we include LGS deficiencies, such as spot elongation, cone effect and tip/tilt indetermination in our models. We propose two different models for tip/tilt indetermination, the Gradient-tilt and the Zernike-tilt, and investigate their differences and similarities. For the forward operators, we give the corresponding adjoint operators for two differently weighted inner products. We summarize the noise model for spot elongated measurements and review the prior statistics of the atmosphere. Two approaches for the Gradient-based iteration are presented, a least-squares model without statistical information and a Tikhonov-type model including prior information.

In order to speed up the method, we develop a simplified and accelerated step size choice for the Gradient iteration and compare it to several existing step size methods. For the numerical demonstration of the algorithm, we consider two different AO systems, a Multi-Conjugate Adaptive Optics (MCAO) system [26, 8] and a Multi-Object Adaptive Optics (MOAO) system [1, 42]. We chose this combination to demonstrate the versatility of our algorithm, because the two systems differ not only in their size of field of view and guide star setting, but also intrinsically in the design and arrangement of deformable mirrors. We study the quality performance of the Gradient-based method in detail in the context of an E-ELT setting on OCTOPUS, the ESO end-to-end simulation tool.

The paper is organized as follows: In Section 2, we introduce the problem setting, develop a forward model including real-life effects such as tip/tilt indetermination,
Gradient-Based Method

2. Atmospheric tomography. Contrary to many standard methods in AO, we follow a 3-step approach, first introduced in [24], that splits the problem of determining the optimal mirror shape(s) from WFS measurement data into three independent steps that can be solved with different algorithms. In the first step, the incoming wavefronts $\varphi_{\alpha_g}$ of guide stars in directions $\alpha_g$ are determined from WFS sensor data $s_{\alpha_g} = (s_{\alpha_g}^x, s_{\alpha_g}^y)$ via a wavefront reconstruction algorithm. Then, in the second step, a limited angle tomography problem is solved, i.e. discrete atmospheric layers are reconstructed from the incoming wavefronts of all guide stars. The final step is the determination of the optimal mirror shape(s) from the discrete atmosphere, i.e. by simple projection in the case of MOAO or LTAO (Laser Tomography Adaptive Optics) systems, or by more elaborate algorithms for MCAO systems, see e.g. [24, 47]. For a graphical illustration of the 3-step approach, see Figure 1.

For an accurate reconstruction of the atmosphere, light sources are needed, i.e. an astronomical object or natural guide star (NGS). However, only a limited sky-coverage can be reached with NGS, therefore, laser guide stars (LGS) are created by strong laser beacons that stimulate the sodium layer of the atmosphere.

The emitted light is affected by several problems [8]: first, the light source is only finitely high, thus, the light travels down in a cone. This cone effect can be treated...
rather easily in the tomography step, see Section 2.2. Second, due to the fact that
the laser beam travels up and down the same way, the tip/tilt component of the
atmospheric aberrations cannot be measured correctly. This problem of tip/tilt
indetermination is treated in Section 2.3. Third, the Shack-Hartman (SH) WFS
measurements from LGS are affected by spot elongation, an effect explained in
Section 2.4 together with a turbulence model for the atmosphere.

In this paper, we will mainly focus on the second step of the 3-step approach, the
atmospheric tomography. This is a limited angle tomography problem and, thus,
an ill-posed inverse problem. Due to the discontinuous dependence of the data on
the solution, regularization techniques are needed in order to obtain a stable and
efficient reconstruction.

2.1. Wavefront reconstruction. There are many algorithms available for the re-
construction of an incoming wavefront from WFS data, e.g. for the Shack-Hartman
WFS [44, 21, 13, 7, 19, 30, 31, 49, 18] or for the Pyramid WFS [22, 34, 20].

In the following, we assume a SH WFS, but want to point out that this choice is by
no means necessary and the Gradient-based algorithm for atmospheric tomography
with incoming wavefronts as data presented below can be easily used with a Pyramid
WFS as well.

For a SH WFS, one has to solve

\[ \Gamma \varphi = s, \]

and \( \Gamma = (\Gamma_1, \ldots, \Gamma_G) \) the combined SH operator for all guide star directions,
\( \varphi = (\varphi_{\alpha_1}, \ldots, \varphi_{\alpha_G}) \) the incoming wavefronts and \( s = (s_{\alpha_1}, \ldots, s_{\alpha_G}) \) the WFS
measurement data. In the following we omit the subscript \( g \) for the SH operator,
as it is independent of the specific guide star direction:

\[
\Gamma : H^s(\Omega_D) \to \mathbb{R}^{2|I|},
\]

\[
s^x_{\alpha_g} = \frac{1}{|\Omega_{ij}|} \int_{\Omega_{ij}} \frac{\partial \varphi_{\alpha_g}}{\partial x} d(x, y), \ (i, j) \in I
\]

\[
s^y_{\alpha_g} = \frac{1}{|\Omega_{ij}|} \int_{\Omega_{ij}} \frac{\partial \varphi_{\alpha_g}}{\partial y} d(x, y), \ (i, j) \in I
\]

with \( I \) the index set of all active subapertures of the SH WFS, \(|I|\) the number of
active subapertures, \( \Omega_{ij} \) the area of the \( ij \)th subaperture.

2.2. Forward model. The second step of the 3-step approach is the solution of the
atmospheric tomography problem. Given the reconstructed wavefronts \( \varphi_{\alpha_g} \) in
each guide star direction, we have to solve a limited angle tomography problem, see
Figure 2, where we want to reconstruct a discrete atmosphere \( \Phi = (\Phi^{(1)}, \ldots, \Phi^{(L)}) \)
with layers at heights \( h_l \). This is an ill-posed inverse problem and requires regu-
larization. In the following, we present a Gradient–based approach that allows to
regularize the problem on the one hand by the number of iterations, on the other
by the optional inclusion of a statistics-driven penalty term.

Following [32, 33] and the principles of geometrical optics, we can model the light
propagation through the turbulence. For NGS and LGS we assume geometric light
Gradient-Based Method

Figure 2. Illustration of the atmospheric tomography problem with three guide star directions, three WFS and three atmospheric layers.

propagation with the forward operators:

$A_{\alpha_g} : \bigotimes_{l=1}^L L^2(\Omega_l) \rightarrow L^2(\Omega_D)$

(2)

$A_{\alpha_g} \Phi := \sum_{l=1}^L \Phi^{(l)}(c_l r + h_l \alpha_g) = \varphi_{\alpha_g}(r), \ r \in \Omega_D$

(3)

$c_l := \begin{cases} 
1, & \text{for NGS} \\
1 - \frac{h_l}{h_{LGS}}, & \text{for LGS and } h_{LGS} \cdots \text{LGS height}
\end{cases}$

$\Omega_D := \{ r = (x, y) \in \mathbb{R}^2 : d \leq \|r\| \leq D \}$

$\Omega_l := \bigcup_{g=1}^G \Omega_{D}(h_l \alpha_g)$,

where $\Omega_D$ is the (ring-shaped) aperture of the telescope and $\Omega_l$ the area on layer $l$ “seen” by the wavefront sensors, with the shifted version of the aperture

$\Omega_{D}(h_l \alpha_g) := \left\{ r \in \mathbb{R}^2 : \frac{r - h_l \alpha_g}{c_l} \in \Omega_D \right\}$.

The direction of GS $g$ is given by the two-dimensional vector $\alpha_g = (\alpha_g^{(1)}, \alpha_g^{(2)})$ which corresponds to the 3D directional vector $(\alpha_g^{(1)}, \alpha_g^{(2)}, 1)$. The forward model presented above can be used for incoming wavefronts from LGS and NGS and just differs in the values of the parameter $c_l$ denoting whether the cone effect is taken into account.
Note that it is also possible to consider the tomography operator (2) between the spaces \( \bigotimes_{l=1}^{L} H^1(\Omega_l) \) to \( H^1(\Omega_D) \), see [9].

2.3. Tip/tilt indetermination. When introducing LGS with SH WFS in an AO system, one faces the problem of tip/tilt indetermination, i.e. the planar component of the LGS data is wrongly measured. Thus, additional NGS measurements are used to obtain correct tip/tilt information. In some AO systems, instead of full NGS sensor data, so-called tip/tilt sensors (TTS) are used, which have only a small number of subapertures with two values on each, i.e. the x- and y-coordinate of the normal vector of the plane describing the tip and tilt.

There are several possible models for the tip/tilt data (see e.g. [41]), we will follow two of them: the Gradient-tilt and the Zernike-tilt.

2.3.1. Gradient-tilt. The Gradient definition of tip/tilt has been formulated, e.g. in [23]. The values \( t^x_\alpha g \) and \( t^y_\alpha g \) can be modelled as the averages of the derivatives of the incoming wavefronts. The forward model for \( N \) tip/tilt sensors with 1 subaperture reads as

\[
G_{\alpha g} \Phi = (t^x_{\alpha g}, t^y_{\alpha g}) \in \mathbb{R}^2, g = G + 1, \ldots, G + N.
\]

\[
t^x_{\alpha g} = \frac{1}{|\Omega_D|} \sum_{l=1}^{L} \int_{\Omega_D} \frac{\partial}{\partial x} \Phi(l)(r + h_l(\alpha g)) \, dr
\]

\[
t^y_{\alpha g} = \frac{1}{|\Omega_D|} \sum_{l=1}^{L} \int_{\Omega_D} \frac{\partial}{\partial y} \Phi(l)(r + h_l(\alpha g)) \, dr,
\]

with \( r = (x, y) \in \Omega_D \) and \( c_l \) as in (4). Note that \( c_l \) can be omitted as the light source for TTS are NGS, i.e. \( c_l = 1 \). The operator \( G_{\alpha g} \) is well defined for

\[
G_{\alpha g} : \bigotimes_{l=1}^{L} H^s(\Omega_l) \to \mathbb{R}^2,
\]

with \( s > \frac{1}{2} \). This smoothness condition, i.e. \( H^1(\Omega_l) \), makes the formal derivation and computation of the adjoint \( G_{\alpha g}^* \) costly, see [23].

2.3.2. Zernike-tilt. Contrary to the Gradient-tilt, the Zernike-tilt, or best-fit tilt, is defined on the space of square integrable functions. The Zernike-tilt is described by the normal vector to the best-fit plane of the wavefront distortion. Furthermore, the two components are equal exactly the coefficients of the two low-order Zernike polynomials [41]. The corresponding forward operator for \( N \) tip/tilt sensors with 1 subaperture each reads as

\[
Z_{\alpha g} : \bigotimes_{l=1}^{L} L^2(\Omega_l) \to \mathbb{R}^2
\]

(4) \( Z_{\alpha g} \Phi = (t^x_{\alpha g}, t^y_{\alpha g}) \in \mathbb{R}^2, g = G + 1, \ldots, G + N. \)

\[
t^x_{\alpha g} = \frac{1}{c_l|\Omega_D|} \sum_{l=1}^{L} \int_{\Omega_D} \Phi(l)(r + h_l(\alpha g)) \cdot x \, dr = \frac{1}{c_l|\Omega_D|} \int_{\Omega_D} \varphi_{\alpha g}(x, y) \cdot x \, dr.
\]
with a constant $c > 0$, $\mathbf{r} = (x, y) \in \Omega_D$.

To derive the constant $c$, we use the fact that the Zernike-tilt is a least-squares linear approximation to the phase over the aperture, i.e., for a plane the Zernike-tilt is exactly the corresponding normal vector. Thus, the choice of $c$ depends on the geometry of the aperture. First, assume a square aperture $\Omega_S = \left[-\frac{d}{2}, \frac{d}{2}\right]^2$, then for an arbitrary plane with coefficients $A, B, C \in \mathbb{R}$ it holds that

$$\int_{\Omega_S} (A - Bx - Cy) \cdot x \, d(x, y) = \frac{B d^4}{12} = |\Omega_S| \frac{d^2}{12} B,$$

and thus, $c = \frac{d^2}{12}$. For circular apertures $\Omega_D = \{\mathbf{r} \in \mathbb{R}^2 : ||\mathbf{r}|| \leq D\}$, we have

$$\int_{\Omega_D} (A - Bx - Cy) \cdot x \, d(x, y) = \frac{B \pi D^4}{4} = |\Omega_D| \frac{D^2}{4} B,$$

i.e., $c = \frac{D^2}{4}$.

For circular apertures with central obstruction, i.e., a ring $\{\mathbf{r} \in \mathbb{R}^2 : d \leq ||\mathbf{r}|| \leq D\}$ and $0 < d < D$,

$$\int_{\Omega_D} (A - Bx - Cy) \cdot x \, d(x, y) = \frac{B \pi (D^4 - d^4)}{4} = |\Omega_D| \frac{D^4 + d^4}{4} B,$$

i.e., $c = \frac{D^2 + d^2}{4}$.

For a wavefront $\varphi$ sufficiently smooth, we can expand it by means of Taylor series and compare the two tip/tilt definitions:

**Proposition 1.** Let the wavefront $\varphi \in C^4(B_D)$, with $B_D = \{\mathbf{r} \in \mathbb{R}^2 : ||\mathbf{r}|| \leq D\}$, and $\Omega_D = \{\mathbf{r} \in \mathbb{R}^2 : d \leq ||\mathbf{r}|| \leq D\}$ be an annular aperture with $0 \leq d < D$. Then, for the Zernike-tilt $t_G$ and the Gradient-tilt $t_G$, it holds that

$$\|t_G - t_Z\|^2 = C \left( \left( \frac{\partial^3 \varphi(0,0)}{\partial x^3} + \frac{\partial^3 \varphi(0,0)}{\partial x \partial y^2} \right)^2 + \left( \frac{\partial^3 \varphi(0,0)}{\partial y^3} + \frac{\partial^3 \varphi(0,0)}{\partial x^2 \partial y} \right)^2 \right),$$

with the $C = \frac{D^4 + 4D^2d^2 + d^4}{24(D^2 + d^2)} > 0$ only depending on the geometry of the aperture.

**Proof.** We expand $\varphi$ by means of Taylor series, i.e., there exists $\theta \in (0, 1)$ such that for all $\mathbf{r} \in B_D$,

$$\varphi(\mathbf{r}) = \varphi(0,0) + x \frac{\partial}{\partial x} \varphi(0,0) + y \frac{\partial}{\partial y} \varphi(0,0)$$
$$+ \frac{1}{2} \left( x^2 \frac{\partial^2}{\partial x^2} \varphi(0,0) + 2xy \frac{\partial^2}{\partial x \partial y} \varphi(0,0) + y^2 \frac{\partial^2}{\partial y^2} \varphi(0,0) \right)$$
$$+ \frac{1}{6} \left( x^3 \frac{\partial^3}{\partial x^3} \varphi(0,0) + 3x^2y \frac{\partial^3}{\partial x^2 \partial y} \varphi(0,0) + 3xy^2 \frac{\partial^3}{\partial x \partial y^2} \varphi(0,0) + y^3 \frac{\partial^3}{\partial y^3} \varphi(0,0) \right)$$
$$+ \frac{1}{24} \left( x^4 \frac{\partial^4}{\partial x^4} \varphi(\theta \mathbf{r}) + 4x^3y \frac{\partial^4}{\partial x^3 \partial y} \varphi(\theta \mathbf{r}) + 6x^2y^2 \frac{\partial^4}{\partial x^2 \partial y^2} \varphi(\theta \mathbf{r}) \right.$$
$$+ 4xy^3 \frac{\partial^4}{\partial x \partial y^3} \varphi(\theta \mathbf{r}) + y^4 \frac{\partial^4}{\partial y^4} \varphi(\theta \mathbf{r}) \right).$$
Then, the components of the Zernike-tilt are
\[
t_x = \frac{1}{c|\Omega_D|} \int_{\Omega_D} \varphi(x, y) \cdot x \, d\mathbf{r} = \frac{\partial}{\partial x} \varphi(0, 0) + c \left( \frac{\partial^3}{\partial x^3} \varphi(0, 0) + \frac{\partial^3}{\partial x \partial y^2} \varphi(0, 0) \right),
\]
\[
t_y = \frac{1}{c|\Omega_D|} \int_{\Omega_D} \varphi(x, y) \cdot y \, d\mathbf{r} = \frac{\partial}{\partial y} \varphi(0, 0) + c \left( \frac{\partial^3}{\partial y^3} \varphi(0, 0) + \frac{\partial^3}{\partial x^2 \partial y} \varphi(0, 0) \right),
\]
with \( c = \frac{D^4 + D^2 d^2 + d^4}{12(D^2 + d^2)} \), while the components of the Gradient-tilt are
\[
t_x^G = \frac{1}{|\Omega_D|} \int_{\Omega_D} \frac{\partial}{\partial x} \varphi(x, y) \, d\mathbf{r} = \frac{\partial}{\partial x} \varphi(0, 0) + \frac{c}{2} \left( \frac{\partial^3}{\partial x^3} \varphi(0, 0) + \frac{\partial^3}{\partial x \partial y^2} \varphi(0, 0) \right),
\]
\[
t_y^G = \frac{1}{|\Omega_D|} \int_{\Omega_D} \frac{\partial}{\partial y} \varphi(x, y) \, d\mathbf{r} = \frac{\partial}{\partial y} \varphi(0, 0) + \frac{c}{2} \left( \frac{\partial^3}{\partial y^3} \varphi(0, 0) + \frac{\partial^3}{\partial x^2 \partial y} \varphi(0, 0) \right),
\]
with \( c = \frac{D^2 + d^2}{4} \) from Definition (4) above. Then, the Euclidean norm of the difference in the two tilt/tilt descriptions fulfils
\[
\|t_G - t_Z\|^2 = C \left( \left( \frac{\partial^3 \varphi(0, 0)}{\partial x^3} + \frac{\partial^3 \varphi(0, 0)}{\partial x \partial y^2} \right)^2 + \left( \frac{\partial^3 \varphi(0, 0)}{\partial y^3} + \frac{\partial^3 \varphi(0, 0)}{\partial x^2 \partial y} \right)^2 \right),
\]
with \( C = \frac{c}{2} - \frac{D^4 + 4D^2 d^2 + d^2}{24(D^2 + d^2)} \).

Thus, for plane, bilinear and bi-quadratic wavefronts, Zernike- and Gradient-tilt yield the same values. Only when introducing cubic or higher order terms, the tilt/tilt values differ. A realistic value for \( C \), as used in our numerical examples in Section 4, is close to \( C \sim 17 \).

In AO systems with a combination of LGS and NGS, both with SH WFS, one needs to remove the incorrect tilt/tilt information from LGS data. Therefore, the operator \( \Pi \) is introduced for Gradient-tilt (as in [33]) and for Zernike-tilt with \( \mathbf{r} = (x, y) \in \Omega_D \):
\[
\Pi^G \varphi_{\alpha_s}(r) := \varphi_{\alpha_s}(r) - x \frac{1}{|\Omega_D|} \int_{\Omega_D} \frac{\partial}{\partial x} \varphi_{\alpha_s}(r) \, d\mathbf{r} - y \frac{1}{|\Omega_D|} \int_{\Omega_D} \frac{\partial}{\partial y} \varphi_{\alpha_s}(r) \, d\mathbf{r},
\]
\[
\Pi^Z \varphi_{\alpha_s}(r) := \varphi_{\alpha_s}(r) - x \frac{1}{c|\Omega_D|} \int_{\Omega_D} \varphi_{\alpha_s}(r) \cdot x \, d\mathbf{r} - y \frac{1}{c|\Omega_D|} \int_{\Omega_D} \varphi_{\alpha_s}(r) \cdot y \, d\mathbf{r}.
\]

2.4. Turbulence and spot elongation modelling. In an atmospheric tomography problem, one aims to accurately reconstruct the turbulent atmosphere [40, 39, 8]. To model these turbulences, we make the standard assumption that an atmosphere \( \Phi \) can be described by a finite number of thin layers \( (\Phi^{(1)}, \ldots, \Phi^{(L)}) \) at heights \( h_l \) for \( l = 1, \ldots, L \) [27]. Let us assume that the incoming wavefronts are measured by a Shack-Hartmann (SH) wavefront sensor [28]. Following the random setting in [14, 47, 48], SH WFS data can be modelled as random variable \( S \), the turbulent layers as random variable \( \phi \) and the additive noise due to spot elongation as random variable \( \eta \sim \mathcal{N}(0, C_{\eta}) \) with zero mean, i.e.
\[
S = \Gamma A \phi + \eta,
\]
Gradient-Based Method

\[ C_\Phi = \begin{pmatrix} C_{\Phi(1)} & 0 \\ \vdots & \ddots \\ 0 & C_{\Phi(L)} \end{pmatrix} \]

and \( C_{\Phi(l)} \) is the covariance matrix of layer \( l \). In the following, we will employ a wavelet approximation of the Kolmogorov model \([27, 28]\) with covariance matrices \( C_{\Phi(l)} \) described in \([14, 47, 48]\).

The additive noise due to spot elongation in (7) arises due to the thickness of the sodium layer and results in a light source that is an elongated spot rather than a point. Depending on the laser launch position, the height of the LGS and the sodium layer density, one can model the corresponding covariance matrix \( C_\eta \) for the noise \([4, 36]\). The error in the measurements due to spot elongation follows a normal distribution with zero mean and covariance matrix \( C_\eta \), thus, noisy SH WFS data can be modelled as

\[ s_\delta = s + C_\eta^{1/2} \eta, \]

with \( \eta \) white noise and \( s \) being the concatenation of SH WFS data \( s_{\alpha_g} = [s_x^{\alpha_g} s_y^{\alpha_g}] \) in all LGS directions \( \alpha_g \).

The noise covariance matrix \( C_\eta \) is a block diagonal matrix with blocks \( C_{\alpha_g} \) for each LGS \( \alpha_g \) (and blocks \( \sigma^2 I \) of weighted identity for NGS) and has been described, e.g. in \([36]\). We use the approach presented in \([3]\) where \( C_{\alpha_g} \) is given as block diagonal matrix with \( 2 \times 2 \) blocks for each subaperture \( i \),

\[ C_{\alpha_g}^i = \sigma^2 \begin{pmatrix} 1 + \alpha_\eta^2 \beta_1^2 / \theta^2 & \alpha_\eta^2 \beta_1 \beta_2 / \theta^2 \\ \alpha_\eta^2 \beta_1 \beta_2 / \theta^2 & 1 + \alpha_\eta^2 \beta_2^2 / \theta^2 \end{pmatrix}, \]

with \( \theta \) the full width at half maximum of a non-elongated spot, the elongation vector

\[ \beta = \frac{\text{FWHM}_{Na}}{h_{LGS}} (x - x_{LLT}), \]

and \( x_{LLT} \) the laser launch position, FWHM_{Na} the full width at half maximum of the sodium density profile, \( h_{LGS} \) the height of the LGS at 90km and \( \sigma \sim \sqrt{1/\text{photons per subap}} \) the noise variance of a non-elongated measurement. The noise variance \( \sigma \) in (8) can also vary for each guide star direction, i.e., \( \sigma_g \) might be different for different directions \( \alpha_g \). Note that this does not affect the presented Gradient-based algorithm. The easily computable inverse of (8) reads as

\[ (C_{\alpha_g}^i)^{-1} = \frac{\theta^2}{\sigma^2 \theta^2 + \alpha_\eta^2 \beta_1^2 + \alpha_\eta^2 \beta_2^2} \begin{pmatrix} 1 + \alpha_\eta^2 \beta_1^2 / \theta^2 & -\alpha_\eta^2 \beta_1 \beta_2 / \theta^2 \\ -\alpha_\eta^2 \beta_1 \beta_2 / \theta^2 & 1 + \alpha_\eta^2 \beta_2^2 / \theta^2 \end{pmatrix}. \]

The parameter \( \alpha_\eta \) allows to tune the influence of the spot elongation. For \( \alpha_\eta = 0 \) the covariance in equation (8) equals the covariance for a uniform and decorrelated case.

3. A Gradient-based method for atmospheric tomography. We want to reconstruct the discrete atmosphere \( \Phi \), that is the solution to the system

\[ A \Phi = \varphi, \]
with $\varphi$ the incoming wavefronts or tip/tilt data from all guide stars. The operator $A$ maps the turbulent layers onto the incoming data:

$$A : \bigotimes_{l=1}^{L} L^2(\Omega_l) \to \bigotimes_{g=1}^{G} L^2(\Omega_D) \times \bigotimes_{n=1}^{N} Y,$$

with $G$ the number of LGS and $Y \in \{L^2(\Omega_D), \mathbb{R}^2\}$ for the $N$ NGS. For AO systems with full $84 \times 84$ SH WFS for the NGS, the space $Y$ equals $L^2(\Omega_D)$, while for AO systems with TTS with only 1 sub-aperture $Y$ is set to $\mathbb{R}^2$. In the following, we denote the function space for the incoming wavefronts and tip/tilt data as

$$Q := \bigotimes_{g=1}^{G} L^2(\Omega_D) \times \bigotimes_{n=1}^{N} Y.$$

The operator equation (9) can be solved with the Gradient–based method (Algorithm 1) in two possible ways: First, the method of steepest descent can be applied to the least-squares functional (approach without noise models)

$$J(\Phi) = \|\Pi(A\Phi - \varphi)\|_Q^2 \to \min .$$

We call this the approach without noise models. The tip/tilt removal operator $\Pi = (\Pi, \ldots, \Pi, I, \ldots, I)$ equals the operator $\Pi$ from (5) or (6) for LGS and the identity for NGS. The gradient of (10) reads as $J'(\Phi) = -2A^*\Pi(\varphi - A\Phi)$, with $A^*$ the adjoint of the coupled operator $A$. Note that $\Pi$ is a self-adjoint projection.

Second, the method of steepest descent can be applied to the Tikhonov-type functional

$$J(\Phi) = \|\Pi(A\Phi - \varphi)\|_Q^2 + \alpha_\Phi \|\Phi\|_{C^{-1}}^2 \to \min ,$$

with $\alpha_\Phi \in [0, \infty)$. We call this the approach with noise models, where a regularized version of (9) with the a-priori information about the atmosphere is solved. Furthermore, for spot-elongated measurements, the inclusion of the noise covariance matrix into the data misfit term is possible. Note that the approach with noise models, i.e. (11), is similar to the maximum a-posteriori estimation [15]. Contrary to standard Landweber iteration and existing methods for atmospheric tomography, such as the Kaczmarz method [24], the Gradient-based approach allows to include the statistical models for the atmospheric turbulence and the noise due to spot elongation, which is of particular interest in cases of low photon flux.

The incoming wavefronts follow a linear model with additive noise $\varphi^{\delta} = \Gamma^T(s_{\alpha} + C_{\alpha, \eta}^{1/2})$. Thus the inverse of the noise covariance matrix of the incoming wavefronts can be described as $C^{-1}_\eta = \Gamma^T C^{-1}_\eta \Gamma$, see, e.g., [15]. In order to account for the fact that the incoming wavefronts lie in a Sobolev space $H^s$ with $s \approx 2$ [27], we approximate the corresponding transfer matrix $U$ in $H^2$.

However, we do not want to actually calculate $U^{-1}\Gamma^T$ as this would destroy any advantage of separating step 1 and 2 in the three step approach. As $\Gamma^T \Gamma$ is a discrete approximation to the negative Laplace operator with homogeneous Neumann boundary conditions [13], we approximate numerically $U^{-1}\Gamma^T C_{\alpha, \eta}^{-1} \Gamma \sim (\Gamma^T \Gamma) C_{\alpha, \eta}^{-1} \Gamma$, such that for all guide star directions combined $C^{-1}_\eta \sim \Gamma C^{-1}_\eta \Gamma$, which can be easily pre-calculated, e.g., by means of the CuReD algorithm or a
truncated SVD. The gradient of (11) then reads as
\[ J'(\Phi) = -2A^*\Pi\Gamma^{-1}\Pi(\varphi - A\Phi) + 2\alpha_4C_{\varphi}^{-1}\Phi. \]

**Algorithm 1** Gradient-based method for atmospheric tomography

1. Initialize
   \[ j := 0 \]
   \[ \Phi_0 := \text{initial guess} \]
2. Begin steepest descent iterations
   \[ d_j := -J'(\Phi_j) \]
   \[ \tau_j := \min_{t \in [0, \infty)} J(\Phi_j + td_j) \]
   \[ \Phi_{j+1} := \Phi_j + \tau_j d_j \]
3. Increment iteration
   \[ j := j + 1 \]
4. End steepest descent iterations

### 3.1. Spaces and inner products

We will consider the incoming wavefronts as elements of the space \( L^2(\Omega_D) \), and use the canonical inner product

\[ \langle \varphi^{(1)}, \varphi^{(2)} \rangle_Q := \sum_{g=1}^{G+N} \langle \varphi^{(1)}_{\alpha_g}, \varphi^{(2)}_{\alpha_g} \rangle_{L^2(\Omega_D)}, \]

for full NGS measurements,

\[ \langle \varphi^{(1)}, \varphi^{(2)} \rangle_Q := \sum_{g=1}^{G} \langle \varphi^{(1)}_{\alpha_g}, \varphi^{(2)}_{\alpha_g} \rangle_{L^2(\Omega_D)} + \sum_{g=G+1}^{G+N} \langle t^{(1)}_{\alpha_g}, t^{(2)}_{\alpha_g} \rangle_{\mathbb{R}^2}, \]

for TTS.

For the atmospheric layers, we can use the canonical inner product, given by

\[ \langle \Phi, \Psi \rangle_L := \sum_{l=1}^{L} \frac{1}{\gamma_l} \langle \Phi^{(l)}, \Psi^{(l)} \rangle_{L^2(\Omega_l)}, \]

with the \( c^2 \)-profile \( \gamma_l \) as the relative strength of layer \( l \) with \( \sum_{l=1}^{L} \gamma_l = 1 \).

However, in a gradient based method, this inner product leads to optimal updates only within the area that is covered by all guide stars. Thus we would like to define a new, weighted inner product that also gives optimal updates in areas that are covered by less or even just one guide star.

To this end we define on the layer \( l \) the overlap function

\[ \omega_l(r) := \sum_{g=1}^{G} \chi_{\Omega_l}^{(h_\alpha_g)}(r) + \sum_{n=1}^{N} \chi_{\Omega_l}^{(h_n\beta_n)}(r), \]

with \( \omega_l(r) \in \{1, 2, \ldots, G+N\} \), for all \( r \in \Omega_l \). We define \( \Lambda^l_i := \{ x \in \Omega_l : \omega_l(x) = i \} \) for \( i = 1, \ldots, G+N \) and immediately observe:

**Lemma 3.1.** For the sets \( \Lambda^l_i \) holds

(i) \( \bigcup_{i=1}^{G+N} \Lambda^l_i = \Omega_l \), \( l = 1, \ldots, L \),

(ii) \( \Lambda^l_i \cap \Lambda^l_j = \emptyset \) for \( i \neq j \), \( i, j = 1, \ldots, G+N \).

We can now define the new inner product on the layers by

\[ \langle \Phi^{(l)}, \Psi^{(l)} \rangle_\Lambda := \sum_{i=1}^{G+N} \int_{\Lambda^l_i} \omega_l(x) \cdot \Phi^{(l)}(x) \cdot \Psi^{(l)}(x) \, dx, \]
and the weighted inner product on the atmosphere by

\[ \langle \Phi, \Psi \rangle_\Xi := \sum_{l=1}^L \frac{1}{\gamma_l} \langle \Phi^{(l)}, \Psi^{(l)} \rangle_\xi. \]

For the approach without noise models, i.e. the minimization of (10), the stepsize \( \tau \in [0, \infty) \) in the method of steepest descent (Algorithm 1) then reads as

\[ \tau_j = \min_{t \in [0, \infty)} J(\Phi_j + td_j) = \frac{1}{2} \langle \mathrm{d}_j, \mathrm{d}_j \rangle_Z, \]

with \( Z \in \{\Xi, L\} \). Similarly for (11) when noise models are included

\[ \tau_j = \min_{t \in [0, \infty)} J(\Phi_j + td_j) = \frac{1}{2} \langle \mathrm{d}_j, \mathrm{d}_j \rangle_Z \left( \Pi^{-1} \Pi \mathrm{A}_j, \mathrm{A}_j \right)_Q + \alpha_\Phi \langle C_\Phi^{-1} \mathrm{d}_j, \mathrm{d}_j \rangle_Z. \]

In the case of \( Z = \Xi \), i.e. the inner product \( \langle , \rangle_\Xi \), the stepsize in (18) can be simplified to \( \tau_j = \frac{1}{2} \).

### 3.2. The adjoint operators.

The adjoint of the operator depends on the chosen inner products. In this Section we present the adjoint operators for full LGS or NGS data and the adjoint of the Zernike-tilt operator, for the case of Gradient-tilt we refer to [23].

With the inner product in (14), the \( L^2 \)-adjoint \( \mathrm{A}_{\alpha_g}^* : L^2(\Omega_D) \to \bigotimes_{l=1}^L L^2(\Omega_l) \) can be written as

\[ \mathrm{A}_{\alpha_g}^* (\Psi) = \left( (\gamma_1 (T^{\alpha_g h_1})^* \Psi, \ldots, \gamma_L (T^{\alpha_g h_L})^* \Psi \right)^T \]

\[ (T^{\alpha_g h_l})^* \Psi = \frac{1}{c_l} \Psi \left( \frac{1}{c_l} (r - \alpha_g h_l) \right) \chi_{\Omega_D(h_1 \alpha_g)} \left( \frac{1}{c_l} r \right). \]

and \( (T^{\alpha_g h_l} \Phi)(r) := \Phi(c_l r + h_1 \alpha_g) \), with \( c_l \) as above, see also [32, 33]. Note that we use the \( L^2 \)-adjoint here due to its superior computational efficiency. For the choice of \( H^1(\Omega_D) \) instead of \( L^2(\Omega_D) \), we refer to [32, 9]. The adjoint of \( \mathrm{A} \) with respect to the inner product (17) can be formulated as:

**Proposition 2.** The adjoint of the operator \( \mathrm{A} \) is given by

\[ \mathrm{A}^* \varphi = \sum_{g=1}^G \mathrm{A}_{\alpha_g}^* \varphi_{\alpha_g}, \text{ and} \]

\[ \mathrm{A}_{\alpha_g}^* \varphi_{\alpha_g} = \left( (\gamma_1 \cdot \frac{1}{\omega_l(r)} \cdot \varphi_{\alpha_g} \left( \frac{r - h_1 \alpha_g}{c_l} \right) \chi_{\Omega_D(h_1 \alpha_g)} \left( \frac{1}{c_l} r \right) \left( \frac{1}{c_l^2} \right) \right) \right) \]

\[ \int_{\Omega_D} \ldots \]

**Proof.** Setting \( r' = c_l r + h_1 \alpha_g \) we obtain \( dr = \frac{1}{c_l^2} dr \) and \( r \in \Omega_D \Leftrightarrow r' \in \Omega(h_1 \alpha_g) \). Thus we have

\[ \langle \mathrm{A}_{\alpha_g} \Phi, \varphi_{\alpha_g} \rangle_{L^2(\Omega_D)} = \int_{\Omega_D} \left( \sum_{l=1}^L \phi^{(l)}(c_l r + h_1 \alpha_g) \right) \cdot \varphi_{\alpha_g}(r) dr \]

\[ = \sum_{l=1}^L \int_{\Omega_D(h_1 \alpha_g)} \phi^{(l)}(r) \cdot \varphi_{\alpha_g} \left( \frac{r - h_1 \alpha_g}{c_l} \right) \left( \frac{1}{c_l^2} \right) dr. \]
Proof. We use the adjoint of \( \mathbf{Z}(\alpha) \) can be computed as follows:

\[
\mathbf{Z}^*(\alpha) : \mathbb{R}^2 \to \bigotimes_{l=1}^L L^2(\Omega_l)
\]

The inner product on the wavefronts can be rewritten as follows:

\[
\langle A\Phi, \varphi \rangle_W = \sum_{g=1}^G \langle A_{\alpha_g} \Phi, \varphi_{\alpha_g} \rangle_{L^2(\Omega_D)} = \sum_{g=1}^G \langle \Phi, A_{\alpha_g}^* \varphi_{\alpha_g} \rangle_{\Xi}
\]

For the Zernike-tilt operator in (4), the adjoint

\[
\mathbf{Z}^*(\alpha) : \mathbb{R}^2 \to \bigotimes_{l=1}^L L^2(\Omega_l)
\]

can be computed as follows:

**Proposition 3.** The adjoint of \( \mathbf{Z}^*(\alpha) \) is given by

\[
\mathbf{Z}^*(\alpha) \mathbf{t}(\alpha) = \left[ \gamma_l \cdot u \left( \frac{r-h_l \alpha_g}{c_l} \right) \frac{1}{c_l} \chi_{\Omega_D(h_l \alpha_g)} \left( \frac{1}{c_l} \right) \right]_{l=1}^L,
\]

with \( u(x,y) = t^x_\alpha \cdot x + t^y_\alpha \cdot x \).

**Proof.** We use \( \mathbf{t} = (t^x, t^y)^T \) and \( u(x,y) = t^x \cdot x + t^y \cdot x \). Furthermore, \( r' = \frac{r-h_l \alpha_g}{c_l} \)

and, thus, \( dr' = \frac{1}{c_l} dr \). Then, with the definition of the Zernike-tilt in (4), it follows that

\[
\langle \mathbf{Z}(\alpha) \mathbf{t}, \mathbf{e} \rangle_{\mathbb{R}^2} = \sum_{l=1}^L \int_{\Omega_D} \Phi^{(l)}(r) \cdot \varphi_{\alpha_g} \left( \frac{r-h_l \alpha_g}{c_l} \right) \frac{1}{c_l} \chi_{\Omega_D(h_l \alpha_g)} \left( \frac{1}{c_l} \right) dr
\]

Adaptation to different sensor geometries

The theory of the Zernike-tilt can be easily extended to different TTS geometries:

Let us assume a \( p \times p \) sensor with \( p \in \mathbb{N} \). The forward operator then reads as
Figure 3. Example for a possible $2 \times 2$ TTS sensor geometry

(22) $\mathbf{Z}_{\alpha_g}^{p \times p} : \bigotimes_{i=1}^{L} L^2(\Omega_i) \to \mathbb{R}^{2p}$

$\mathbf{Z}_{\alpha_g}^{p \times p} \Phi = \mathbf{t} = (t_1^x, t_1^y, \ldots, t_{2p}^x, t_{2p}^y)^T \in \mathbb{R}^{2p}$, $g = 1, \ldots, G$,

$t_i^x = \frac{1}{c_i |\Omega_i|} \sum_{l=1}^{L} \int_{\Omega_i} \Phi^{(l)}(c_l r + h_l \alpha_g) \cdot x \, d\mathbf{r} = \frac{1}{c_i |\Omega_i|} \int_{\Omega_i} \varphi_{\alpha_g}(x, y) \cdot x \, d\mathbf{r}$,

$t_i^y = \frac{1}{c_i |\Omega_i|} \sum_{l=1}^{L} \int_{\Omega_i} \Phi^{(l)}(c_l r + h_l \alpha_g) \cdot y \, d\mathbf{r} = \frac{1}{c_i |\Omega_i|} \int_{\Omega_i} \varphi_{\alpha_g}(x, y) \cdot y \, d\mathbf{r}$,

for $i = 1, \ldots, 2p$, with $\Omega_i \cap \Omega_j \neq \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^{2p} \Omega_i = \Omega_D$, i.e. a possible sensor geometry would be Figure 3.

Similar as in Proposition 3, the adjoint can be derived as follows:

**Proposition 4.** The adjoint of $\mathbf{Z}_{\alpha_g}^{p \times p}$ is given by

(23) $(\mathbf{Z}_{\alpha_g}^{p \times p})^* \mathbf{t} = \left[ \gamma_l \cdot u \left( \frac{\mathbf{r} - \alpha_g h_l}{c_l} \right) \frac{1}{c_l^2} \chi_{\Omega_D(\alpha_g h_l)} \left( \frac{1}{c_l} \mathbf{r} \right) \right]_{l=1}^{L}$,

with $\mathbf{t} = (t_1^x, t_1^y, \ldots, t_{2p}^x, t_{2p}^y)^T$ and $u : \Omega_D \to \mathbb{R}$, with $u(x, y) := u_i(x, y)$ for $(x, y) \in \Omega_i$ and $i = 1, \ldots, p$. Furthermore,

$u_i : \Omega_i \to \mathbb{R}$, and $u_i(x, y) = \tilde{t}_i^x \cdot x + \tilde{t}_i^y \cdot y$, with $\tilde{t}_i^x := \frac{1}{c_i |\Omega_i|} t_i^x$ and $\tilde{t}_i^y := \frac{1}{c_i |\Omega_i|} t_i^y$

**Proof.** Similarly as in the Proof for Proposition 3 and with the fact that $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^{2p} \Omega_i = \Omega_D$, we have

$$
(\mathbf{Z}_{\alpha_g}^{p \times p} \Phi, \mathbf{t})_{\mathbb{R}^{2p}} = \sum_{i=1}^{p} \sum_{l=1}^{L} \frac{1}{c_i |\Omega_i|} \left( \int_{\Omega_i} (x \cdot \Phi^{(l)}(c_l r + h_l \alpha_g) \cdot t_i^x + y \cdot \Phi^{(l)}(c_l r + h_l \alpha_g) \cdot t_i^y) \, d\mathbf{r} \right)
$$
Gradient-Based Method

\[
\sum_{l=1}^{L} \frac{1}{\gamma_l} \int_{\Omega} \Phi^{(l)}(c_l r + h_l \alpha_y) \cdot \gamma_l \cdot u(x, y) \, dr = \langle \Phi, (Z_{\alpha_y}^{c_p})^n t \rangle_L .
\]

3.3. The Gradient-based method for specific AO systems. We now apply the abstract idea of the Gradient-based method to Multi-Conjugate AO systems as well as Multi-Object AO system and compare different stepsize choices,

3.3.1. Multi-Conjugate Adaptive Optics. For MCAO, the task of atmospheric tomography is to solve the operator equation (9) from LGS measurements and additional tip/tilt measurements. Thus, the operator \( A \) and the vector of measurements \( \varphi \) are given as

\[
A = (A_{\alpha_1}, \ldots, A_{\alpha_G}, N_{\beta_1}, \ldots, N_{\beta_N})^T, \\
\varphi = (\varphi_{\alpha_1}, \ldots, \varphi_{\alpha_G}, t_{\beta_1}, \ldots, t_{\beta_N})^T.
\]

A solution can be obtained via applying the method of steepest descent on the corresponding least squares functional (10). The gradient of the functional then reads as

\[
J'(\Phi) = -2 \left( \sum_{g=1}^{G} A_{\alpha_g}^* \Pi(\varphi_{\alpha_g} - A_{\alpha_g} \Phi) + \sum_{n=1}^{N} N_{\beta_n}^* (t_{\beta_n} - N_{\beta_n} \Phi) \right).
\]

For cases with high noise, i.e. low photon flux with spot elongation, we minimize (11), with the gradient

\[
J'(\Phi) = -2 \left( \sum_{g=1}^{G} A_{\alpha_g}^* \Pi \Pi^{-1} (\varphi_{\alpha_g} - A_{\alpha_g} \Phi) + \sum_{n=1}^{N} N_{\beta_n}^* (t_{\beta_n} - N_{\beta_n} \Phi) - \alpha_{\Phi} C_{\varphi}^{-1} \Phi \right),
\]

and \( \alpha_{\Phi} \in [0, \infty) \). In the case of a combination of LGS and TTS, where the TTS have only 1 subaperture each, we have to reconstruct the tip/tilt information differently than in the case of full NGS data. The separated and integrated tip/tilt reconstruction have been introduced in [23] in the case of Gradient tilt. For Zernike tilt, the algorithms are analogous while the computational effort of the adjoint tip/tilt operator is considerably reduced due to the \( L^2 \)-setting.

3.3.2. Multi-Object Adaptive Optics. For MOAO, no separate treatment of the tip/tilt data is necessary as we assume to have full NGS measurements. Thus, equation (9) now reads as

\[
A \Phi = \begin{pmatrix} A_{\alpha_1} \\ \vdots \\ A_{\alpha_{G+N}} \end{pmatrix} \Phi = \begin{pmatrix} \varphi_{\alpha_1} \\ \vdots \\ \varphi_{\alpha_{G+N}} \end{pmatrix} = \varphi.
\]

The gradient of the least-squares functional in (10) reads as

\[
J'(\Phi) = -2 \left( \sum_{g=1}^{G} A_{\alpha_g}^* \Pi(\varphi_{\alpha_g} - A_{\alpha_g} \Phi) + \sum_{g=G+1}^{G+N} A_{\alpha_g}^* (\varphi_{\alpha_g} - A_{\alpha_g} \Phi) \right).
\]
while the corresponding gradient of the Tikhonov functional (11) can be formulated as

\[
J'(\Phi) = -2 \left( \sum_{g=1}^{G} A_{\alpha_g}^* \Pi C_{\alpha_g}^{-1} \Pi (\varphi_{\alpha_g} - A_{\alpha_g} \Phi) \right)
+ \sum_{g=G+1}^{G+N} A_{\alpha_g}^* (\varphi_{\alpha_g} - A_{\alpha_g} \Phi) - \alpha_{\Phi} C_{\Phi}^{-1} \Phi \right),
\]

and \( \alpha_{\Phi} \in [0, \infty) \).

3.4. Fast stepsize. One of the crucial points of a gradient iteration is the choice of the stepsize \( \tau \) as the stepsize is responsible for the convergence and the convergence speed. For the classical Landweber method the stepsize is chosen constantly, which yields very slow convergence speed. The classical steepest descent stepsize as in (18), reads as

\[
\tau_{SD_j} = \frac{1}{\sum_{l=1}^{L} \frac{1}{2\gamma} \|d_l\|^2_{L^2(\Omega_l)}} \left( G + N \right) \sum_{g=1}^{G+N} \left( \varphi_{\alpha_g} - A_{\alpha_g} \Phi_j \right)^2_{L^2(\Omega_D)}.
\]

(25)

which, for the denominator, includes the application of operators \( A_{\alpha_g} \) to each gradient in all directions and the calculation of \( G + N \) scalar products.

A novel and computationally faster stepsize choice is chosen as

\[
\tau_{fast_j} = \frac{1}{\sum_{l=1}^{L} \frac{1}{2\gamma} \|d_l\|^2_{L^2(\Omega_l)}} \left( G + N \right) \sum_{g=1}^{G+N} \left( \varphi_{\alpha_g} - A_{\alpha_g} \Phi_j \right)^2_{L^2(\Omega_D)}.
\]

(26)

Contrary to the classical steepest descent stepsize, the so called fast stepsize choice (26) is a non-monotone stepsize. Due to this fact, the convergence behaviour is improved, see Figure 4. The fast stepsize is applicable to all problems \( Ax = b \), where the operator \( A \) is decomposed in a vector of operators. For any operator equation of this type, the stepsize (26) only involves the evaluation of the two inner products. Here, (26) is tailored to the application of (atmospheric) tomography, as it relies on simplifying the denominator in the presence of several data directions.

The difference can be described as follows: while the classical steepest descent stepsize uses “smoothed” information over all directions, the fast stepsize conserves the information of each direction separately. This is reflected in the choice of the denominator. Let the norms involving the residual phases directions \( \alpha_g \) be defined as

\[
p := \sum_{g=1}^{G+N} \| \varphi_{\alpha_g} - A_{\alpha_g} \Phi \|^2_{L^2(\Omega_D)} \quad \text{and} \quad q := \sum_{g=1}^{G+N} \left( \varphi_{\alpha_g} - A_{\alpha_g} \Phi \right)^2_{L^2(\Omega_D)}.
\]

(27)

The norm \( p \) shows a monotone behaviour over the number of iterations. On the contrary, the norm of the sum over all residuals \( q \) behaves non monotonously and
Gradient-Based Method

797

similarly to $\|d\|_L^2$. This stems from the fact that $q$ equals $\|d^{(1)}\|$ if $h_1 = 0$. Moreover, note that the residual in each direction directly relates to the gradient in each direction defined as

$$d_{\alpha_s} = A_{\alpha_s}^* (\varphi_{\alpha_s} - A_{\alpha_s} \Phi),$$

and

$$A_{\alpha_s} d_{\alpha_s} = A_{\alpha_s} A_{\alpha_s}^* (\varphi_{\alpha_s} - A_{\alpha_s} \Phi) = \varphi_{\alpha_s} - A_{\alpha_s} \Phi.$$

The fast stepsize is chosen as the ratio of $\|d\|_L^2$ and $q$, where the number of guide stars $G + N$ is only needed for correct scaling. The numerical effort for evaluating (26) decreases significantly, as no additional application of $A_{\alpha_s}$ is needed. The already computed residual can be used instead, fewer summations are included and a better possibility for parallelization is given, see also Table 1. This is a crucial improvement, as the computational complexity for an AO reconstruction algorithm is severely restricted due to time constraints.

For comparison, we considered the well-known non-monotone stepsize choice of Barzilai-Borwein [2], i.e.,

$$(28) \quad \tau_j = \left\langle \frac{\Delta \Phi}{\Delta d} \right\rangle,$$

with $\Delta \Phi = \Phi_j - \Phi_{j-1}$ and $\Delta d = J'(\Phi_j) - J'(\Phi_{j-1})$, and compared the performance of the algorithm in Figure 4. According to the numerical results, the stepsize choice in (26) yields so far the best quality results, in particular for few Gradient iterations. For the exact simulation setting and the evaluation criterion, we refer to Section 4.

The pseudo-code for the Gradient–based method with the simplified and accelerated stepsize computation (26) for MOAO can be seen in Algorithm 2.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure4.png}
\caption{Comparison of step size choices for MOAO highflux without spot elongation: classical stepsize choice in (25) w.r.t. $\langle \ldots \rangle_L$ vs. fast stepsize in (26) vs. Barzilai-Borwein stepsize in (28). See Section 4 for simulation details.}
\end{figure}

In Table 1, we provide the computational complexity of the operators used in the Gradient-based method with and without noise models. We focus on a system with full SH WFS for LGS and NGS, i.e. no TTS. The parameter $n$ is the number of active apertures, $L$ is the number of reconstructed layers and $G$ is as the number of guide star directions. The computational effort is given as number of floating point operations in total and per process with the number of parallel processes equal to
Algorithm 2 Gradient–based method with accelerated stepsize computation

Choose $\Phi_0$.

for $i = 1, \ldots$ do

$\Phi_i = \Phi_{i-1}$.

for $g = 1, \ldots, G$ do

residual$_g = [ \overline{C_{\alpha_g}^{-1}} \cdot \Pi ] \cdot (\varphi_{\alpha_g} - A_{\alpha_g} \cdot \Phi_i)$ [if noise models included]

gradient$_g = (A_{\alpha_g})^T \cdot \Pi \cdot $ residual$_g$

$r_i = r_i + L \cdot$ residual$_g$

g$_i = g_i + $ gradient$_g$

end for

for $g = G + 1, \ldots, G + N$ do

residual$_g = \varphi_{\alpha_g} - A_{\alpha_g} \cdot \Phi_i$

gradient$_g = (A_{\alpha_g})^* \cdot $ residual$_g$

$r_i = r_i + L \cdot$ residual$_g$

g$_i = g_i + $ gradient$_g$

end for

$[ g_i = g_i - \alpha_{\Phi} \cdot C_{\Phi}^{-1} \cdot g_i ]$ [if noise models included]

stepsize $= (g_i^T \cdot g_i) / (r_i^T \cdot r_i + \alpha_{\Phi} \cdot g_i^T \cdot C_{\Phi}^{-1} \cdot g_i)$ [if noise models included]

$\Phi_i = \Phi_i + $ stepsize $\cdot \gamma \cdot g_i$

end for

$max(L, G)$. Note that both, the Gradient method with and without noise models, are parallelizable on layers $L$ and in guide star directions $G$.

One can observe that the Gradient method without noise models is considerably faster than the Gradient method on the Tikhonov functional, as no costly noise models (inverse turbulence covariance or inverse spot elongation covariance) are needed. The stepsize choice in (26) yields the lowest computational complexity for the Gradient method without noise models. Under consideration of noise models, the numerical cost is higher due to the inclusion of the inverse turbulence covariance. The Barzilai and Borwein stepsize choice has very low computational cost for both cases, however, needs the storage of the previous update as well as the previous gradient, both on all reconstruction layers $L$. The choice of the weighted inner product $\langle \cdot, \cdot \rangle_{\Xi}$ only slightly reduces the effort of one Gradient iteration.

4. Numerical results. In this Section, we describe the numerical performance of the algorithms presented in terms of quality using OCTOPUS [17], the end-to-end simulation tool of the European Southern Observatory (ESO). We use two test cases, a Multi-Conjugate Adaptive Optics (MCAO) system and a Multi-Object Adaptive Optics (MOAO) system on the European Extremely Large Telescope, simulated with a telescope diameter of 42 m.

For the sake of completeness, we add the most important parameters used in the simulations: The turbulence is simulated by the ESO median seeing nine-layer atmosphere with a Fried parameter of $r_0 = 0.129$ m. The guide star asterism is shown in Figure 5, with 6 LGS in circles of 2 and 7.5 arcmin of diameter, for MCAO and MOAO respectively, and 3 NGS in circles of 8/3 and 10 arcmin of diameter. For MOAO, both LGS and NGS are observed by $84 \times 84$ Shack-Hartmann WFS. In the MCAO system, the NGS are observed by TTS with 1 subaperture. For the LGS at 90 km height, four laser side-launch positions are used. The LGS are simulated
### Table 1. Upper bounds on the computational cost estimates for the operations in the Gradient-based method with and without noise models.

| Operation | overall complexity | compl. per process |
|-----------|--------------------|---------------------|
| \( A \)   | \( 7(L-1)Gn \)    | \( 7(L-1)n \)      |
| \( C_{n}^{-1} \) | -                   | \( 40Gn \)          |
| \( \Pi_{n} \)    | \( 3Gn \)          | \( 3n \)            |
| \( A^{*}, \langle \ldots \rangle_{L} \) | \( (8L-7)Gn \) | \( (8L-7)n \) |
| \( A^{*}, \langle \ldots \rangle_{\Xi} \) | \( (8L-7)Gn + Ln \) | \( (8L-6)n \) |
| \( C_{\Phi}^{-1}, [47, 48, 25] \) | \( 235nL \) | \( 235n \) |
| SD steps, \( \langle \ldots \rangle_{L} \) | \( 7(L-1)Gn \) + \( (L+G)n \) + \( (236L+G)n \) | \( (7L-5)n \) + \( (7L+271)n \) |
| BB-steps. | \( 5Ln \)          | \( 5n \)            |
| fast steps. | \( (L+1)n \)       | \( (237L+2)n \)     | \( 2n \) + \( 239n \) |
| SD steps. \( \langle \ldots \rangle_{\Xi} \) | -                   | -                   |

with cone effect and tip-tilt indetermination in all test cases. In the MCAO system, three deformable mirrors conjugated to heights 0, 4000 and 12700 m are used, while in MOAO only one \( 85 \times 85 \) ground DM is simulated.

The simulation results include different types and levels of noise: Cases without and with noise due to spot elongation are simulated. In both cases, the level of photon flux is varied, typically from a few photons per subaperture to 10000. The photon flux relates directly to the noise level in the WFS measurements, i.e. the noise variance in the absence of spot elongation is \( \sigma^2 \sim 1/\text{photons per subap} \). In the case of spot elongation the covariance of the WFS measurements is given in (8).

To evaluate the reconstruction quality, we use the long exposure (LE) Strehl ratio [16] and simulate one second of real time, i.e. 500 time steps. Strehl ratio is a measure of image quality \( \in [0,1] \) related to the \( L^2 \)-error, where 1 is best and...
0 worst. As in MCAO, one is interested in a good reconstruction quality over the whole field of view, we evaluate the quality in 25 directions, arranged in a square of 1 arcmin of length. In MOAO, we only observe one direction of interest in the zenith. The reconstruction in MCAO is performed on three artificial layers at the heights of the deformable mirrors, whereas in MOAO nine layers are reconstructed.

We compare the Gradient-based method with the reference results provided by the ESO: for MCAO the reference method is the MVM, for MOAO the Fractal Iterative Method (FrIM) [35], both tuned by ESO. For the Gradient-based method, the initial guess is chosen as zero in the first time step. From then onward, a warm-restarting technique is used: At time step $t > 0$ the initial guess is set to the last iterate from the previous time step, i.e. $\Phi_0 := \Phi_{t-1}$. Choosing $\Phi_0 = 0$ in each time step, would severely decrease quality performance when maintaining the used (low) number iterations, necessary to fulfil the speed requirements. More sophisticated choices of initial guess, e.g. by predictive control have not been tested extensively and are left to future work.

First, we consider LGS without spot elongation. In Figures 6 and 7, one can observe that only 2 to 5 iterations are necessary for a comparable reconstruction quality in MCAO.

In Figure 8, we can observe a superior quality over a variety of flux levels, when using the Zernike-tilt instead of the Gradient-tilt. This holds for MCAO (with
Gradient-Based Method

As expected, one can, in general, observe that the noise level (i.e., the inverse of photon flux) influences all algorithms and settings similarly: the higher the number of photons that reach each subaperture, the better the reconstruction quality.

Figure 8. Comparison of Gradient- and Zernike-tilt. On-axis LE Strehl. Left: MCAO lowflux w/o spot elongation. Right: MOAO lowflux w/o spot elongation.

For the MOAO system we compare the inner products $\langle \cdot, \cdot \rangle_L$ in (14) and $\langle \cdot, \cdot \rangle_\Xi$ in (17) in Figure 9. The results for the inner product $\langle \cdot, \cdot \rangle_\Xi$ lie around 5 percent lower than $\langle \cdot, \cdot \rangle_L$ in the case of Gradient-tilt. For the Zernike-tilt, the difference is less pronounced. The usage of different inner products entails different gradients and, thus, results in different iterations.

Figure 9. MOAO w/o spot elongation: Comparison of inner products for Gradient- and Zernike-tilt. On-axis LE Strehl.

Let us now consider cases with spot elongation, i.e., Figure 10. In the MCAO system, we can observe an improvement when using the Gradient-based method with noise models, especially in the low flux regime. Moreover, the Zernike-tilt reconstruction shows better results than the Gradient-tilt. However, for photon flux levels below 200ph/subaperture/frame, the algorithm still shows deficiencies in the reconstruction quality.

In MOAO, this problem is even more pronounced due to the wide FoV. The Gradient-based method with noise models only yields small improvements for the
5 and 10 photon cases. In view of other methods available to the authors, i.e. results for FrIM (by ESO), FEWHA [48, 25] and the Kaczmarz method [33], the Gradient-based method performs similarly for cases without spot elongation. A detailed comparison in the case of MOAO can be found, e.g., in [25]. We can conclude that the quality performance of the algorithm suffers from the spot elongation noise compared to 2-step methods, such as FEWHA or FrIM. This stems from the separation of the atmospheric tomography step into two individual subproblems - the wavefront reconstruction and the tomography step on incoming phases. In particular, the wavefront reconstruction algorithm used in the numerical tests, the CuReD algorithm, includes the summation of measurements along $x$- and $y$-axes, which removes the information of spot elongation. One possible remedy for the 3-step approach for LGS, is the application of a different wavefront reconstruction algorithm, which is more sensitive to the problem of spot elongation.

The 3-step approach with the Gradient-based method is particularly well suited for AO systems with NGS. The method is very fast and well parallelizable, and without the drawback of spot elongation, it yields good quality results. Furthermore, it is very versatile, as different algorithms for the three individual steps - wavefront reconstruction, tomography and fitting step - can be combined.

5. Conclusion and outlook. In this paper, we have introduced a novel iterative method based on the Gradient method for solving the problem of atmospheric tomography. Based on previous work in iterative reconstruction methods for AO systems, we developed a fast and versatile method and studied its qualitative performance in the context of MCAO and MOAO. Contrary to standard Landweber iteration and existing methods for atmospheric tomography, such as the Kaczmarz method, the Gradient-based approach allows to include the statistical models for the atmospheric turbulence and the noise due to spot elongation. Two types of tip/tilt models have been investigated, the Zernike-tilt and the Gradient-tilt. A new analysis of the Zernike-tilt is presented and simulation results suggest that a quality improvement can be reached at a lower computational cost. In addition, we presented a new stepsize choice for the gradient iteration, which decreases the numerical effort and yields superior quality results in numerical simulations. A thorough analysis of the fast stepsize choice is left to future work. We demonstrated a fast convergence of the algorithm, in particular for MCAO, and illustrated the performance in terms of quality for high- and lowflux regimes.

We believe that the Gradient-based method is a promising algorithm, in particular for NGS systems. Due to its fast and flexible character, it can be applied...
in different AO systems without major changes or precomputations. The improvement of the method with respect to spot elongation in wide field of view AO and the incorporation of predictive schemes are left for future study.

Acknowledgments. This research was done in the framework of the projects Mathematical Algorithms and Software for E-ELT Adaptive Optics in cooperation with the European Southern Observatory and Beobachtungsorientierte Astrophysik, funded by the Austrian Federal Ministry of Science and Research. The authors would like to thank in particular the Austrian Adaptive Optics team for support. In addition, we thank Miska Le Louarn and Andreas Neubauer for fruitful discussions. Furthermore, we would like to thank the unknown referees for valuable comments.

REFERENCES

[1] F. Assémat, E. Gendron and F. Hammer, The falcon concept: multi-object adaptive optics and atmospheric tomography for integral field spectroscopy–principles and performance on an 8-m telescope, Monthly Notices of the Royal Astronomical Society, 376 (2007), 287–312.
[2] J. Barzilai and J. Borwein, Two-point step size gradient methods, IMA Journal of Numerical Analysis, 8 (1988), 141–148.
[3] C. Béchet, M. L. Louarn, R. Clare, M. Tallon, I. Tallon-Bosc and E. Thiébaut, Closed-loop ground layer adaptive optics simulations with elongated spots: Impact of modeling noise correlations, in 1st AO4ELT conference - Adaptive Optics for Extremely Large Telescopes, 2010, 03004–03009, URL http://dx.doi.org/10.1051/ao4elt/201003004.
[4] C. Béchet, M. Tallon, I. Tallon-Bosc, Eric Thiébaut, M. L. Louarn and R. M. Clare, Optimal reconstruction for closed-loop ground-layer adaptive optics with elongated spots, J. Opt. Soc. Am. A, 27 (2010), A1–A8, URL http://josa.osa.org/abstract.cfm?URI=josa-27-11-A1.
[5] E. Brunner, C. Béchet and M. Tallon, Optimal projection of reconstructed layers onto deformable mirrors with fractal iterative method for ao tomography, in SPIE Astronomical Telescopes+ Instrumentation, International Society for Optics and Photonics, 847 (2012), 84475I–84475J.
[6] B. Ellerbroek and C. Vogel, Simulations of closed-loop wavefront reconstruction for multi-conjugate adaptive optics on giant telescopes, Proc. SPIE, 5169 (2003), 206–217.
[7] B. Ellerbroek, Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques, J. Opt. Soc. Am., 19 (2002), 1803–1816.
[8] B. Ellerbroek and C. Vogel, Inverse problems in astronomical adaptive optics, Inverse Problems, 25 (2009), 063001, 37pp.
[9] M. Eslitzbichler, C. Pechstein and R. Ramlau, An h1-kaczmarz reconstructor for atmospheric tomography, Journal of Inverse and Ill-Posed Problems, 21 (2013), 431–450.
[10] T. Fusco, J.-M. Conan, G. Rousset, L. Mugnier and V. Michau, Optimal wave-front reconstruction strategies for multi conjugate adaptive optics, J. Opt. Soc. Am. A, 18 (2001), 2527–2538.
[11] D. Gavel, Tomography for multiconjugate adaptive optics systems using laser guide stars, SPIE Astronomical Telescopes and Instrumentation, 5490 (2004), 1356–1373.
[12] L. Gilles and B. Ellerbroek, Split atmospheric tomography using laser and natural guide stars, J. Opt. Soc. Am., 25 (2008), 2427–2435.
[13] L. Gilles, C. Vogel and B. Ellerbroek, Multigrid preconditioned conjugate-gradient method for large-scale wave-front reconstruction, J. Opt. Soc. Am. A, 19 (2002), 1817–1822.
[14] T. Helin and M. Yuditskiy, Wavelet methods in multi-conjugate adaptive optics, Inverse Problems, 29 (2013), 085003, 18pp, URL http://stacks.iop.org/0266-5611/29/i=8/a=085003.
[15] J. Kaipio and E. Somersalo, Statistical and Computational Inverse Problems, vol. 160 of Applied Mathematical Sciences, Springer Science+Business Media, Inc, 2005.
[16] M. Le Louarn, C. Verinaud and V. Korkiakoski, Simulation of mcao on (extremely) large telescopes, Comptes Rendus Physique, 6 (2005), 1070–1080.
[17] M. Le Louarn, C. Vérinaud, V. Korkiakoski, N. Hubin and E. Marchetti, Adaptive optics simulations for the european extremely large telescope, in *SPIE Astronomical Telescopes+Instrumentation*, International Society for Optics and Photonics, 6272 (2006), 627234–627234.

[18] A. Neubauer, On the ill-posedness and convergence of the Shack-Hartmann based wavefront reconstruction, *J. Inv. Ill-Posed Problems*, 18 (2010), 551–576.

[19] A. Neubauer, A new cumulative wavefront reconstructor for the Shack-Hartmann sensor, *J. Inv. Ill-Posed Problems*, 21 (2013), 451–476.

[20] D. W. Phillion and K. Baker, Two-sided pyramid wavefront sensor in the direct phase mode, *Proc. SPIE*, 6272 (2006), 627228.

[21] L. Poyneer, D. Gavel and J. Brase, Fast wave-front reconstruction in large adaptive optics systems with use of the Fourier transform, *J. Opt. Soc. Am. A*, 19 (2002), 2100–2111.

[22] F. Quirós-Pacheco, C. Correia and S. Esposito, Fourier transform-wavefront reconstruction for the pyramid wavefront sensor, in 1st AO4ELT conference - Adaptive Optics for Extremely Large Telescopes, 2010, 07005–07010.

[23] R. Ramlau, A. Obereder, M. Rosensteiner and D. Saxenhuber, Efficient iterative tip/tilt reconstruction for atmospheric tomography, *Inverse Problems in Science and Engineering*, 22 (2014), 1345–1366.

[24] R. Ramlau and M. Rosensteiner, An efficient solution to the atmospheric turbulence tomography problem using Kaczmarz iteration, *Inverse Problems*, 28 (2012), 095004, 23pp.

[25] R. Ramlau, D. Saxenhuber and M. Yudytskiy, Iterative reconstruction methods in atmospheric tomography: FEWHA, Kaczmarz and Gradient-based algorithm, in *Proc. SPIE*, 9148 (2014), 91480Q.

[26] F. Rigaut, B. Ellerbroek and R. Flicker, Principles, limitations and performance of multiconjugate adaptive optics, *Proc. SPIE*, 4007 (2000), 1022–1031.

[27] F. Roddier, *The effects of atmospheric turbulence in optical astronomy*, *Progress in Optics*, 19 (1981), 281–376.

[28] F. Roddier, *Adaptive Optics in Astronomy*, Cambridge, U.K.; New York : Cambridge University Press, Cambridge, 1999.

[29] M. C. Roggemann and B. Welsh, *Imaging Through Turbulence*, [CRC Press laser and optical science and technology series], CRC Press, 1996.

[30] M. Rosensteiner, Cumulative reconstructor: Fast wavefront reconstruction algorithm for Extremely Large Telescopes, *J. Opt. Soc. Am. A*, 28 (2011), 2132–2138.

[31] M. Rosensteiner, Wavefront reconstruction for extremely large telescopes via CuRe with domain decomposition, *J. Opt. Soc. Am. A*, 29 (2012), 2328–2336.

[32] M. Rosensteiner and R. Ramlau, Efficient iterative atmospheric tomography reconstruction from LGS and additional tip/tilt measurements, in *SPIE 8447, Adaptive Optics Systems III*, 2012, 84475S.

[33] M. Rosensteiner and R. Ramlau, The Kaczmarz algorithm for multi-conjugate adaptive optics with laser guide stars, *J. Opt. Soc. Am. A*, 30 (2013), 1680–1686.

[34] I. Shatokhina, A. Obereder, M. Rosensteiner and R. Ramlau, Preprocessed cumulative reconstructor with domain decomposition: A fast wavefront reconstruction method for pyramid wavefront sensor, *Applied Optics*, 52 (2013), 2640–2652.

[35] M. Tallon, I. Tallon-Bosc, C. Béchet, F. Money, M. Fradin and E. Thiébaut, Fractal iterative method for fast atmospheric tomography on extremely large telescopes, in *Proc. SPIE 7736, Adaptive Optics Systems II*, 2010, 77360X.

[36] M. Tallon, I. Tallon-Bosc, C. Béchet and E. Thiébaut, Shack-hartmann wavefront reconstruction with elongated sodium laser guide stars: Improvements with priors and noise correlations, *Proc. SPIE, Adaptive Optics Systems*, 7015 (2008), 70151N, URL: http://link.aip.org/link/?PSI/7015/70151N/1.

[37] M. Tallon, E. Thiébaut and C. Béchet, A fractal iterative method for fast wavefront reconstruction for extremely large telescopes, in *Adaptive Optics: Methods, Analysis and Applications*, Optical Society of America, 2007, paper PMA2

[38] E. Thiébaut and M. Tallon, Fast minimum variance wavefront reconstruction for extremely large telescopes, *J. Opt. Soc. Am. A*, 27 (2010), 1046–1059.

[39] A. Tokovinin, M. L. Louarn and M. Sarazin, Isoplanatism in a multi-conjugate adaptive optics system, *J. Opt. Soc. Am.,* 17 (2000), 1819–1827.

[40] A. Tokovinin and E. Viard, Limiting precision of tomographic phase estimation, *J. Opt. Soc. Am.,* 18 (2001), 873–882.
[41] G. Tyler, Bandwidth considerations for tracking through turbulence, *J. Opt. Soc. Am. A*, **11** (1994), 358–367.
[42] F. Vidal, E. Gendron and G. Rousset, Tomography approach for multi-object adaptive optics, *JOSA A*, **27** (2010), A253–A264.
[43] C. Vogel and Q. Yang, Fast optimal wavefront reconstruction for multi-conjugate adaptive optics using the Fourier domain preconditioned conjugate gradient algorithm, *Optics Express*, **14** (2006), 7487–7498.
[44] C. Vogel and Q. Yang, Multigrid algorithm for least-squares wavefront reconstruction, *Applied Optics*, **45** (2006), 705–715.
[45] Q. Yang, C. Vogel and B. Ellerbroek, Fourier domain preconditioned conjugate gradient algorithm for atmospheric tomography, *Applied Optics*, **45** (2006), 5281–5293.
[46] M. Yudytskiy, *Wavelet Methods in Adaptive Optics*, PhD thesis, Johannes Kepler University Linz, 2014.
[47] M. Yudytskiy, T. Helin and R. Ramlau, A frequency dependent preconditioned wavelet method for atmospheric tomography, in *Third AO4ELT Conference - Adaptive Optics for Extremely Large Telescopes*, 2013.
[48] M. Yudytskiy, T. Helin and R. Ramlau, Finite element-wavelet hybrid algorithm for atmospheric tomography, *J. Opt. Soc. Am. A*, **31** (2014), 550–560, URL http://josaa.osa.org/abstract.cfm?URI=josaa-31-3-550.
[49] M. Zhariy, A. Neubauer, M. Rosensteiner and R. Ramlau, Cumulative wavefront reconstructor for the Shack-Hartman sensor, *Inverse Problems and Imaging*, **5** (2011), 893–913.

Received February 2015; revised January 2016.

E-mail address: saxenhuber@indmath.uni-linz.ac.at
E-mail address: ronny.ramlau@jku.at