PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Kaczmarz and Cimmino: iterative and layer-oriented approaches to atmospheric tomography

Garbellotto, Chiara, Donini, Michele, Ragazzoni, Roberto, Arcidiacono, Carmelo, Baruffolo, Andrea, et al.

Chiara Garbellotto, Michele Donini, Roberto Ragazzoni, Carmelo Arcidiacono, Andrea Baruffolo, Jacopo Farinato, "Kaczmarz and Cimmino: iterative and layer-oriented approaches to atmospheric tomography," Proc. SPIE 9909, Adaptive Optics Systems V, 99094J (27 July 2016); doi: 10.1117/12.2232494

SPIE.

Event: SPIE Astronomical Telescopes + Instrumentation, 2016, Edinburgh, United Kingdom

Kaczmarz and Cimmino: iterative and layer-oriented approaches to atmospheric tomography

Chiara Garbellotto[a,b], Michele Donini[c,d], Roberto Ragazzoni[e], Carmelo Arcidiacono[f], Andrea Baruffolo[e], and Jacopo Farinato[e]

[a]Department of Physics and Astronomy "Galileo Galilei", University of Padova, Vicolo dell' Osservatorio, 2, 35122 Padova PD, Italy
[b]School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK
[c]Department of Computer Science, University College London, London WC1E 6BT, UK
[d]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London WC1B 5EH, UK
[e]INAF - Astronomical Observatory of Padova, Vicolo dell'Osservatorio, 5, 35141 Padova PD, Italy
[f]INAF - Astronomical Observatory of Bologna, Via Ranzani, 1, 40127 Bologna BO, Italy

ABSTRACT

Multi Conjugated Adaptive Optics is based upon tomographic reconstruction of the atmospheric turbulence over the line of sight of a telescope, achieved by combining measurements from different directions in the sky. Using deformable mirrors optically conjugated to different altitudes, a correction can be performed directly on the reconstructed turbulence layers. Different approaches have been developed so far, notably the so called layer-oriented one, experienced with success at the VLT (Very Large Telescope) through MAD (Multi Conjugate Adaptive Optics Demonstrator). It was later shown that the tomography problem, once posed in terms of solving a set of linear equations describing the turbulence layers with respect to the observables, can be solved in an iterative manner through a technique first proposed by Kaczmarz in 1937. It was then speculated that a layer-oriented iteration would asymptotically converge to the same solution. In this paper, we placed the two approaches in the same theoretical framework, identifying them as two different iterative methods to solve the same system of linear equations. We found that the layer-oriented approach can be seen as a weighted form of the iterative method proposed by Cimmino in 1938. By using the known mathematical results relative to Kaczmarz's and the weighted Cimmino methods, we were able to demonstrate the validity of the initial speculation.

1. INTRODUCTION

Multi Conjugate Adaptive Optics (MCAO) was introduced for the first time in 1988[1] in order to obtain a more uniform correction over a larger field of view with respect to classical AO systems.[2] The principle of MCAO is to obtain information about the atmospheric turbulence volume in more than one direction by using multiple guide stars, and to use these measurements to reconstruct the main turbulence layers.[3] Wavefront corrections are then applied on more than one plane, using multiple deformable mirrors (DMs) optically conjugated to different altitudes. Adopting the classical approach to MCAO, called star-oriented, each guide star is assigned to a wavefront sensor (WFS) and the computation of the optimal mirror shape involves the solution of the three sub-problems of wavefront reconstruction, where WFS measurements are used to reconstruct the wavefronts collected in each sensed direction, atmospheric turbulence tomography, where the reconstructed wavefronts are used to derive the shape of the main turbulence layers, and mirror fitting, where the optimal shape of the DMs is

Further author information:
C.G.: E-mail: c.garbellotto.1@research.gla.ac.uk, Telephone: +44 7761 830690

Adaptive Optics Systems V, edited by Enrico Marchetti, Laird M. Close, Jean-Pierre Véran, Proc. of SPIE Vol. 9909, 99094J
 · doi: 10.1117/12.2232494

determined from the reconstructed layers. These three sub-problems are summarized into one operator, usually called R, which maps the WFS measurements to the mirror shape (or related mirror commands). For the new generation of extremely large telescopes, as the ESO E-ELT, using a MCAO system means dealing with the inversion of a matrix-vector system of the size of $\sim 10000 \times 10000$. Moreover, since the atmospheric turbulence rapidly changes, an update for the mirror commands must be computed approximately every millisecond. The size of the system and the short time in which it needs to be solved make its numerical solution particularly challenging, and two alternative approaches to the classical star-oriented MCAO have been proposed. In the so called layer-oriented approach,[4,5] each WFS is conjugated to a specific atmospheric layer and coupled to the DM conjugated to that layer. Each DM applies to its layer a correction proportional to what the corresponding WFS measures, optically combining the light from all the guide stars. This approach does not require the computation of a global reconstruction matrix and it is therefore computationally more efficient. Another advantage is an increase of the sky coverage: optically combining the light from all the guide stars allows for the use of fainter stars, whose light is added to the light coming from the other guide stars. The second alternative approach is the one presented in Ref. 6 by Ramlau and Rosensteiner, who proposed a way to compute the mirror shape without the use of matrix-vector multiplications, aiming at decreasing the overall numerical effort. The proposed method solves the three sub-problems (wavefront reconstruction, atmospheric tomography, mirror deformation) sequentially, using a Kaczmarz iterative algorithm to solve the atmospheric tomography problem.

Despite their different nature, "age" and origin, these two methods show some similarities, and a conjecture was made about their achievable solution, believing that they asymptotically lead to the same one. As it was originally thought and realized, the layer-oriented method does not perform several iterations before finding the appropriate mirror shape, but the conjecture is that iterating the layer-oriented process of WFS "reading" and DM "shaping", considering a fixed turbulence profile (as in Ramlau and Rosensteiner's approach), would asymptomatically lead to the same correction achievable by using the iterative atmospheric reconstruction proposed by Ramlau and Rosensteiner.

This work is organized as follows. We start by introducing the notation and mathematical framework we use to describe and compare the two iterative methods. Section 3 explains the layer-oriented approach and its reformulation as an iterative method. The the second method, the Kaczmarz approach, is described in Sec. 4. In Sec. 5 we compare the two methods and discuss their solutions. Conclusions and ideas for future works are presented in the final section.

2. ADOPTED FORMALISM

The mathematical framework we decided to work with is the one introduced in Ref. 6 to describe the Kaczmarz approach to MCAO, which perfectly fits our need to write the two approaches in terms of iterative methods to solve a system of linear equations. For the sake of completeness we include rigorous mathematical definitions, but we also provide hints for an intuitive understanding of the formalism.

Throughout this work we use bold letters to indicate vectors, matrices, angles and domains (e.g. $\boldsymbol{\Phi}$, $\boldsymbol{S}$, $\boldsymbol{\alpha}_g$, $\boldsymbol{\Omega}_m$). Light letters are used for functions and operators (e.g. w, $T^{\boldsymbol{\alpha}_g h_m}$). We collect some operators into vectors, which we define as new operators; for these new operators we use bold letters, in order to emphasize their vector nature (e.g. $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$). Light subscripts define different elements of a vector ($\boldsymbol{\Phi} = [\Phi_1, \dots, \Phi_M]$) and, when dealing with iterative methods, bold subscripts are used to indicate different iterations (e.g. $\boldsymbol{\Phi_k}$, $\boldsymbol{\Phi_{k+1}}$, $\dots$). When we deal with the norm of a vector or an operator, and use symbol $\|\cdot\|$, we always refer to the norm induced by the inner product defined for that particular space (e.g. the Euclidean norm for the space $\mathbb{R}^2$). Different norms are explained and identified with a subscript (as for $\|\cdot\|_{S^{-1}}$).

Consider and MCAO system which involves G guide stars, all assumed to be at infinite altitude. Each star is referred to using its direction, defined by the angle $\boldsymbol{\alpha}_g$ measured from the optical axis of the telescope. Let M be the number of DMs, each of which is conjugated to a different height h_m. Let (x, y) denote the plane of the telescope aperture, with the origin $(0, 0)$ on the telescope optical axis. Let $\boldsymbol{\Omega}_T$ be a disc with radius $\boldsymbol{r}_T$, centered at the origin and representing the telescope aperture, i.e.

$$\boldsymbol{\Omega}_T = \{\boldsymbol{r} = (x, y) \in \mathbb{R}^2 : \|\boldsymbol{r}\| \leq \boldsymbol{r}_T\}. \tag{1}$$

The shape of the m-th DM can be described as a function Φ_m of the space $L_2(\boldsymbol{\Omega}_m)$, the space of square integrable functions defined on $\boldsymbol{\Omega}_m$. Its domain $\boldsymbol{\Omega}_m$ corresponds to the intersection of a plane parallel to the pupil plane but positioned at height h_m and the light beams of the guide stars, as shown in Fig. 1. The m-th domain $\boldsymbol{\Omega}_m$ can therefore be defined as the union of G shifted version of $\boldsymbol{\Omega}_T$:

$$\boldsymbol{\Omega}_m = \bigcup_{g=1}^{G} \boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g), \tag{2}$$

where

$$\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g) = \{\boldsymbol{r} \in \mathbb{R}^2 : \boldsymbol{r} - h_m \boldsymbol{\alpha}_g \in \boldsymbol{\Omega}_T\} \tag{3}$$

represents the footprint of the g-th guide star on the plane conjugated to the m-th DM, as in Fig. 1, and can also be called a pupil image, because it corresponds to the projection of the pupil in direction $\boldsymbol{\alpha}_g$ on plane h_m. $\boldsymbol{\Omega}_T$ is $\boldsymbol{\Omega}_m$ for $m = 1$. Having Φ_m equal to zero everywhere on $\boldsymbol{\Omega}_m$ means for example that no deformation is assigned to the m-th DM, i.e. it is completely flat.

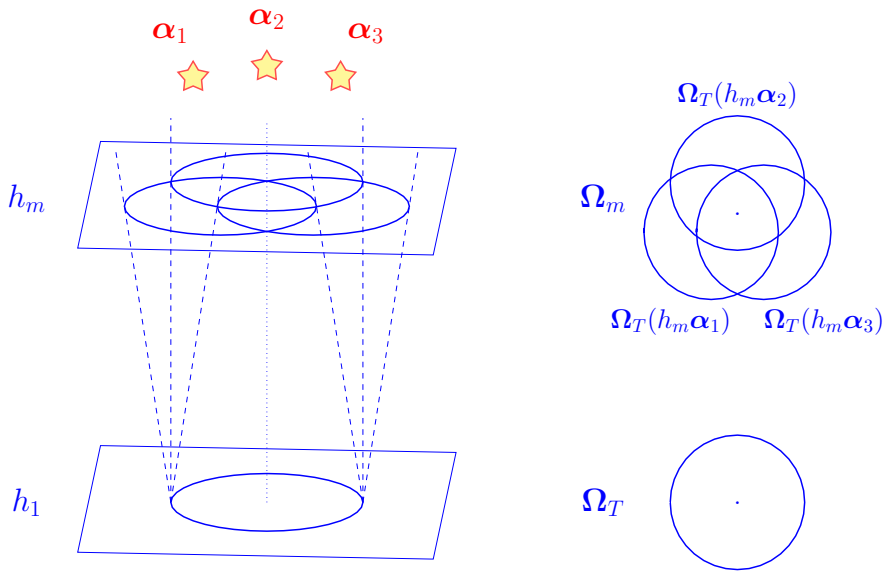

Figure 1. On the left: three guide stars, in directions $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$ and $\boldsymbol{\alpha}_3$, and two planes in the atmosphere, at heights h_m and $h_1 = 0$. On the right: the overlap, on each plane, of the light beams coming from the three observed directions. At $h_1 = 0$, the beams perfectly overlap on the telescope aperture $\boldsymbol{\Omega}_T$. On a higher plane, each direction defines a different pupil image $\boldsymbol{\Omega}_T(h_m\boldsymbol{\alpha}_g)$, and the shape of the illuminated area $\boldsymbol{\Omega}_m$ therefore depends on its height h_m and on the directions $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$ and $\boldsymbol{\alpha}_3$.

In both layer-oriented and star-oriented approaches the first step of the MCAO system is to reconstruct the wavefronts from the WFS measurements. Let $w_{\boldsymbol{\alpha}_1}, \dots w_{\boldsymbol{\alpha}_g}, \dots w_{\boldsymbol{\alpha}_G}$ be the G wavefronts observed in the star-oriented approach. Each $w_{\boldsymbol{\alpha}_g}$ is a function of $L_2(\boldsymbol{\Omega}_T)$, i.e. it is a square integrable function defined on the pupil plane Ω_T. In the layer-oriented approach we have M reconstructed wavefronts, since the WFSs are conjugated to the M planes instead of the G directions. In this case we use the functions $W_1, \dots W_m, \dots W_M$ to describe the wavefronts reconstructed by the different WFSs, whith each W_m being a function of $L_2(\boldsymbol{\Omega}_m)$, the same domain as for the function Φ_m which describes the shape of the m-th DM.

In the following sections we use the introduced functions to create vectors of functions. For example, the M

functions Φ_m representing the shape of the DMs can be collected to create the following vector:

$$\boldsymbol{\Phi} = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_M \end{bmatrix}.$$

We define the inner product

$$\langle \boldsymbol{\Phi}, \boldsymbol{\Psi} \rangle_M = \sum_{m=1}^{M} \frac{1}{\gamma_m} \langle \Phi_m, \Psi_m \rangle_{L_2(\boldsymbol{\Omega}_m)},$$

where γ_m is the strength assigned to the m-th turbulence layer, with

$$\sum_{m=1}^{M} \gamma_m = 1. \tag{4}$$

By introducing this inner product we have defined the Hilbert space

$$L_2^M = \bigoplus_{m=1}^{M} L_2(\boldsymbol{\Omega}_m) \tag{5}$$

vector $\boldsymbol{\Phi}$ belongs to. In the same way, the wavefront functions $W_1, \dots W_m, \dots W_M$ can be collected in a vector $\boldsymbol{W}$, also belonging to the space L_2^M. Another useful tool we use in the following discussion are the shift operators $T^{\boldsymbol{\alpha}_g h_m}$, defined as

$$(T^{\boldsymbol{\alpha}_g h_m} \Psi_m)(\boldsymbol{r}) := \Psi_m(\boldsymbol{r} + h_m \boldsymbol{\alpha}_g), \quad \boldsymbol{r} \in \boldsymbol{\Omega}_T, h_m \in \mathbb{R}, \boldsymbol{\alpha}_g \in \mathbb{R}^2. \tag{6}$$

Operator $T^{\boldsymbol{\alpha}_g h_m}$ acts on functions defined on $\boldsymbol{\Omega}_m$, such as the wavefront function W_m, and gives as an output a function defined on the telescope aperture $\boldsymbol{\Omega}_T$. It selects from the input function W_m only what is defined on the footprint of the g-th star (the domain $\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g)$), and gives that "cut-out" of the function as an output, but shifted to the domain $\boldsymbol{\Omega}_T$, as can be seen in Fig. 2. We can combine the action of shift operators that act on different planes and define the operator $\boldsymbol{A}_{\boldsymbol{\alpha}_g} : L_2^M \to L_2(\boldsymbol{\Omega}_T)$ defined as

$$\boldsymbol{A}_{\boldsymbol{\alpha}_g} \boldsymbol{\Psi} = (T^{\boldsymbol{\alpha}_g h_1}, \dots, T^{\boldsymbol{\alpha}_g h_M}) \begin{bmatrix} \Psi_1 \\ \vdots \\ \Psi_M \end{bmatrix} := \sum_{m=1}^{M} T^{\boldsymbol{\alpha}_g h_m} \Psi_m. \tag{7}$$

Operator $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$ acts as a vector of shift operators on a vector of functions: each $T^{\boldsymbol{\alpha}_g h_m}$ selects from the m-th function of the input vector the part defined on the footprint of the g-th star on the m-th layer. The M selected "cut-outs" are all shifted to $\boldsymbol{\Omega}_T$ and added together. The easiest way to understand the use of $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$ is considering a star-oriented system. Imagine an atmosphere ideally composed of M turbulent layers, collected in a vector $\boldsymbol{\Psi}$. What would the g-th star-oriented WFS see of the turbulence represented by $\boldsymbol{\Psi}$? The answer is $\boldsymbol{A}_{\boldsymbol{\alpha}_g} \boldsymbol{\Psi}$: by observing in direction $\boldsymbol{\alpha}_g$, the g-th WFS sees, projected on $\boldsymbol{\Omega}_T$, a sum of the M wavefront deformations introduced by the turbulence layers in that direction.

Finally we introduce the adjoint operator of $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$. The adjoint of an operator A: X $\to$ Y, with X and Y Hilbert spaces, is a mapping $A^* : \mathrm{Y} \to \mathrm{X}$ fulfilling $\langle Ax, y \rangle_Y = \langle x, A^* y \rangle_X$; therefore, operator A^* significantly depends on the chosen inner products of the spaces X and Y. The adjoints of our operators

$$\boldsymbol{A}_{\boldsymbol{\alpha}_g} : \bigoplus_{m=1}^{M} L_2(\boldsymbol{\Omega}_m) \to L_2(\boldsymbol{\Omega}_T), \qquad g = 1, \dots, G,$$

are given as the mappings

$$\boldsymbol{A}^*_{\boldsymbol{\alpha}_g} : L_2(\boldsymbol{\Omega}_T) \to \bigoplus_{m=1}^{M} L_2(\boldsymbol{\Omega}_m), \qquad g = 1, \dots, G,$$

and defined as

$$\boldsymbol{A}^*_{\boldsymbol{\alpha}_g}\varphi = \begin{bmatrix} \gamma_1 (T^{\boldsymbol{\alpha}_g h_1})^* \varphi \\ \gamma_2 (T^{\boldsymbol{\alpha}_g h_2})^* \varphi \\ \vdots \\ \gamma_L (T^{\boldsymbol{\alpha}_g h_M})^* \varphi \end{bmatrix},$$

where φ is a function defined on the telescope aperture $\boldsymbol{\Omega}_T$ and $(T^{\boldsymbol{\alpha}_g h_m})^*$ is the adjoint of operator $T^{\boldsymbol{\alpha}_g h_m}$:

Proposition 1. The adjoint of $T^{\boldsymbol{\alpha}_g h_m}$ is given by

$$(T^{\boldsymbol{\alpha}_g h_m})^* \varphi = w(\boldsymbol{r} - h_m \boldsymbol{\alpha}_g) \chi_{\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g)}(\boldsymbol{r}).$$

In the above proposition (see Ref. 6 for the proof) $\boldsymbol{r} \in \boldsymbol{\Omega}_m$, $h_m \in \mathbb{R}$, $\boldsymbol{\alpha}_g \in \mathbb{R}^2$, and $\chi_{\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g)}$ denotes the indicator function of the subset $\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g)$ of $\boldsymbol{\Omega}_m$, and is defined as follows:

$$\chi : \boldsymbol{\Omega}_m \mapsto \{0, 1\}$$
$$\chi(x) := \begin{cases} 1 & \text{if} \quad x \in \boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g), \\ 0 & \text{if} \quad x \notin \boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g). \end{cases}$$

Operator $(T^{\boldsymbol{\alpha}_g h_m})^*$ maps φ into a new function, defined on $\boldsymbol{\Omega}_m$, that equals zero everywhere on the domain $\boldsymbol{\Omega}_m$, except on $\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g)$, where it equals a shifted copy of φ (see Fig. 3). A graphic explanation of how operator $\boldsymbol{A}^*_{\boldsymbol{\alpha}_g}$ acts on a function φ defined on $\boldsymbol{\Omega}_T$ can be seen in Fig. 4.

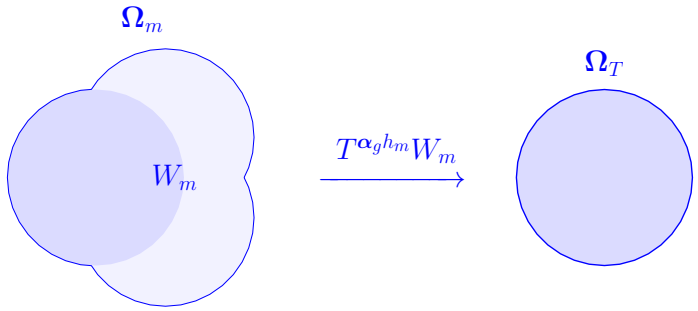

Figure 2. Let W_m be a function defined on $\boldsymbol{\Omega}_m$ representing a wavefront reconstructed from measurements taken on plane h_m. Operator $T^{\boldsymbol{\alpha}_g h_m}$ acts on W_m giving a new function, defined on $\boldsymbol{\Omega}_T$, that equals a shifted copy of the part of W_m defined on the pupil image $\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g)$ (in darker blue).

3. LAYER-ORIENTED APPROACH

In the layer-oriented approach, the correction $\Phi_m(\boldsymbol{r}, t)$ (with $\boldsymbol{r} \in \boldsymbol{\Omega}_m$) applied by each DM is proportional to the measurement $W_m(\boldsymbol{r}, t)$ taken by its coupled WFS. The dynamic behavior of the closed loop system can be described by the following system of differential equations:

$$\frac{d\Phi_m(\boldsymbol{r}, t)}{dt} = \gamma_m W_m(\boldsymbol{r}, t) \quad m = 1, \ldots, M, \tag{8}$$

where γ_m is the gain of the m-th independent loop, with $\sum_{m=1}^{M} \gamma_m = 1$. In Ref. 7 this system was shown to be stable, i.e. capable of returning to the initial conditions when perturbed by an instantaneous input. Furthermore, the achievable correction was studied, under the simplifying hypothesis that the response time of the system is much shorter than the evolution time of the turbulence. Considering this hypothesis and looking for the solution

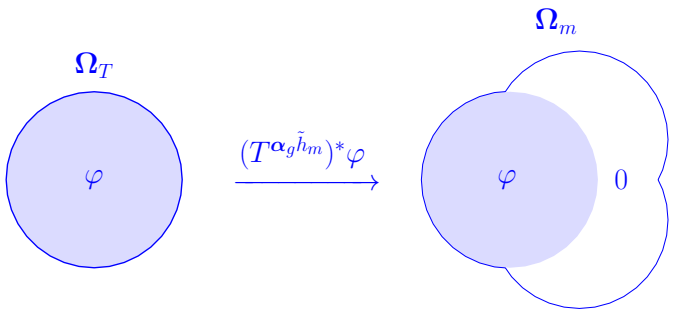

Figure 3. Let φ be a function defined on $\boldsymbol{\Omega}_T$, representing a wavefront reconstruted from star-oriented wavefront measurements. Operator $(T^{\boldsymbol{\alpha}_g h_m})^*$ acts on φ giving a new function, defined on $\boldsymbol{\Omega}_m$, that equals zero on the entire domain except for $\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g)$ (in blue), where it is equal to a shifted copy of function φ.

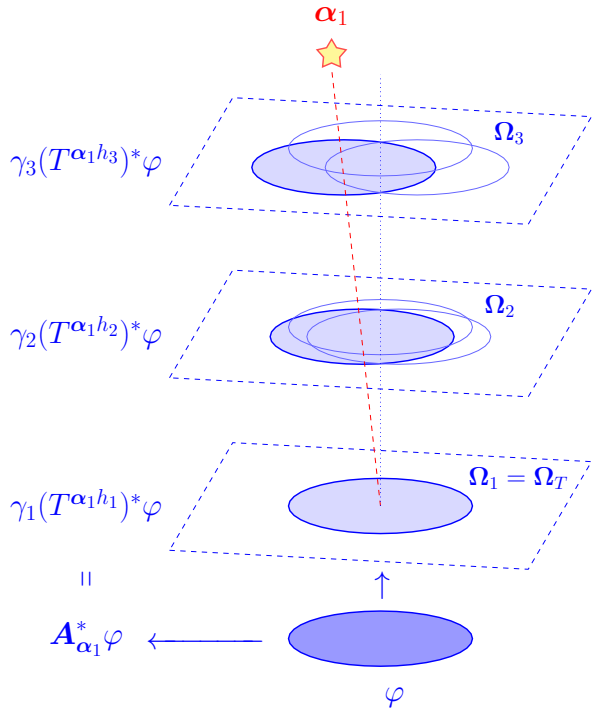

Figure 4. Example of how operator $\boldsymbol{A}^*_{\boldsymbol{\alpha}_1}$ acts on a function φ defined on $\boldsymbol{\Omega}_T$, considering $M = 3$. It maps the function into a vector of the space $\bigoplus_{m=1}^{3} L_2(\boldsymbol{\Omega}_m)$, i.e. in this case a vector composed of 3 layers. Each component of vector $\boldsymbol{A}^*_{\boldsymbol{\alpha}_1} \varphi$ is a function of $\boldsymbol{\Omega}_m$ that equals zero on the entire domain except for the pupil image $\boldsymbol{\Omega}_T(h_m \boldsymbol{\alpha}_g)$, where it equals a shifted copy of $\gamma_m \varphi$.

achievable as $t \to \infty$, starting with flat DMs, is the same as looking for the mirror deformations $\Phi_1, \Phi_2, \ldots, \Phi_M$ that solve system (8) when considering time constant turbulence and

$$\frac{d\Phi_m(\boldsymbol{r}, t)}{dt} = 0 \quad m = 1, \ldots, M. \tag{9}$$

This means that we want to find the mirror deformations that ideally give:

$$\gamma_m W_m(\boldsymbol{r}, t) = 0 \quad m = 1, \ldots, M, \tag{10}$$

i.e. perfectly corrected wavefronts in all sensed planes.

3.1 Iterative layer-oriented

As said, our purpose is to analyze the properties of the layer-oriented approach and comparing it, in terms of achievable correction, to the Kaczmarz approach. In order to do this we here consider a time constant turbulence and we build an iterative method to find the MCAO correction that proceeds at each step in a layer-oriented way. Let ${}_{\boldsymbol{\gamma}}\boldsymbol{W}$ be the vector containing the M closed loop wavefront measurements, weighted with weights γ_m

$$ {}_{\boldsymbol{\gamma}}\boldsymbol{W} = \begin{bmatrix} \gamma_1 W_1 \\ \vdots \\ \gamma_M W_M \end{bmatrix}, $$

and let $\boldsymbol{\Phi_0}$ be the vector whose M components represent the initial corrections applied by the M DMs. This vector is the initial point of our layer-oriented iterative method. To obtain the first iteration, $\boldsymbol{\Phi_1}$, we add to the initial $\boldsymbol{\Phi_0}$ a correction that is proportional to what the WFSs see when observing the stars corrected by $\boldsymbol{\Phi_0}$ (closed loop). Since we are considering time independent incoming wavefronts, the WFS measurements depend on those wavefronts and on the current DM deformations: $W_m = W_m(\boldsymbol{\Phi_0})$. Note that what each WFS measures depends on all the DMs, i.e. on the entire vector $\boldsymbol{\Phi_0}$. With this notation, we are able to write the first iteration of the layer-oriented iterative method as

$$ \begin{bmatrix} \Phi_{1,\mathbf{1}} \\ \vdots \\ \Phi_{M,\mathbf{1}} \end{bmatrix} = \begin{bmatrix} \Phi_{1,\mathbf{0}} \\ \vdots \\ \Phi_{M,\mathbf{0}} \end{bmatrix} + \begin{bmatrix} \gamma_1 W_1(\boldsymbol{\Phi_0}) \\ \vdots \\ \gamma_M W_M(\boldsymbol{\Phi_0}) \end{bmatrix}, \tag{11} $$

or, shortly:

$$ \boldsymbol{\Phi_1} = \boldsymbol{\Phi_0} + {}_{\boldsymbol{\gamma}}\boldsymbol{W_0}. \tag{12} $$

The current DM deformations are now contained in vector $\boldsymbol{\Phi_1}$. To find the second iteration we proceed in the same way: we add to the current DM deformations a term proportional to the WFS measurements. The M WFSs now see the incoming wavefronts (constant in time) as corrected by the DMs described by $\boldsymbol{\Phi_1}$: $W_m = W_m(\boldsymbol{\Phi_1})$. The second iteration is therefore given by

$$ \boldsymbol{\Phi_2} = \boldsymbol{\Phi_1} + {}_{\boldsymbol{\gamma}}\boldsymbol{W_1}. $$

More generally, iteration $k+1$ can be written as

$$ \boldsymbol{\Phi_{k+1}} = \boldsymbol{\Phi_k} + {}_{\boldsymbol{\gamma}}\boldsymbol{W_k}, \tag{13} $$

and a perfect solution is reached when $\boldsymbol{W_k} = \mathbf{0}$, which means that the WSFs observe perfectly corrected wavefronts and the mirror deformations do not need any further update (the so called static response). By using the operators introduced in the previous section, we can write the layer-oriented measurements $\boldsymbol{W_k}$ in terms of the time constant wavefronts coming in from directions $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_G$. Let $\varphi_{\boldsymbol{\alpha}_g}$ be a function defined on the telescope aperture $\boldsymbol{\Omega}_T$ and representing the wavefront deformation observed in direction $\boldsymbol{\alpha}_g$. What a WFS conjugated to height h_m sees of $\varphi_{\boldsymbol{\alpha}_g}$ can be written as

$$ W_m = s_m \sum_{g=1}^{G} (T^{\boldsymbol{\alpha}_g h_m})^* \varphi_{\boldsymbol{\alpha}_g}, \tag{14} $$

where s_m is a function, defined on $\boldsymbol{\Omega}_m$, that equals a different constant value $1/s$ on different parts of the domain $\boldsymbol{\Omega}_m$. The weighting given by s_m comes from the fact that, on each point of the m-th WFS, the incident light is given by the sum of the light coming from different directions. Considering stars of equal intensities, the total light that reaches a certain point of a WFS can be calculated by summing the different contributions and dividing by the number of incident beams. As a next step, we can write each wavefront $\varphi_{\boldsymbol{\alpha}_g}$, observed from the telescope aperture in direction $\boldsymbol{\alpha}_g$, in terms of the incoming wavefront $w_{\boldsymbol{\alpha}_g}$ and the currently applied DM correction $\boldsymbol{\Phi_k}$. Without any mirror deformation we would have $\varphi_{\boldsymbol{\alpha}_g} = w_{\boldsymbol{\alpha}_g}$ (open loop), while taking the mirror deformations into account we have

$$ \varphi_{\boldsymbol{\alpha}_g} = w_{\boldsymbol{\alpha}_g} - \boldsymbol{A}_{\boldsymbol{\alpha}_g} \boldsymbol{\Phi_k}. \tag{15} $$

In the above equation, we use operator $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$ to select and add together the parts of the DMs which influence the light coming from direction $\boldsymbol{\alpha}_g$, and we then subtract this correction to the incoming wavefront. We substitute Eq. (15) into Eq. (14) and multiply each W_m by the gain γ_m, obtaining the following expression for the entire vector of measurements ${}_{\boldsymbol{\gamma}}\boldsymbol{W}_{\boldsymbol{k}}$:

$$\begin{aligned} {}_{\boldsymbol{\gamma}}\boldsymbol{W}_{\boldsymbol{k}} &= \boldsymbol{S}\sum_{g=1}^{G}\boldsymbol{A}^*_{\boldsymbol{\alpha}_g}(w_{\boldsymbol{\alpha}_g} - \boldsymbol{A}_{\boldsymbol{\alpha}_g}\boldsymbol{\Phi}_{\boldsymbol{k}}) \\ &= \boldsymbol{S}\boldsymbol{A}^*(\boldsymbol{w} - \boldsymbol{A}\boldsymbol{\Phi}_{\boldsymbol{k}}), \end{aligned} \tag{16}$$

where $\boldsymbol{A}^*$ is the adjoint of operator $\boldsymbol{A}$ and is represented by a matrix whose G columns are the operators $\boldsymbol{A}^*_{\boldsymbol{\alpha}_g}$, adjoints of operators $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$ (which are the rows of $\boldsymbol{A}$), $\boldsymbol{w}$ is a column vector containing the G star-oriented wavefronts $w_{\boldsymbol{\alpha}_1}, \ldots, w_{\boldsymbol{\alpha}_G}$, and $\boldsymbol{S}$ =diag(s_m) is the matrix of weighting functions (see Fig. 5). Iteration $k+1$ can

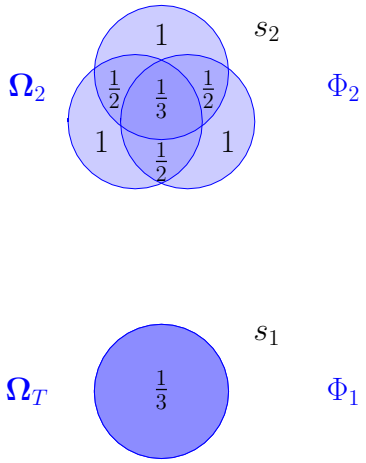

Figure 5. Let $\boldsymbol{\Phi}$ be a vector of two components, Φ_1 and Φ_2, defined on the two domains $\boldsymbol{\Omega}_T$ and $\boldsymbol{\Omega}_2$. Matrix $\boldsymbol{S}$ is in this case a 2×2 diagonal matrix; its two diagonal values s_1 and s_2 are two functions, defined on $\boldsymbol{\Omega}_T$ and $\boldsymbol{\Omega}_2$ respectively, that equal a different constant value on each differently colored area of the two domains. Function s_2 equals 1/3 on the central area of $\boldsymbol{\Omega}_2$, where three footprints overlap, 1/2 on the three lighter blue areas where two footprints overlap, and 1 in the three light blue areas where the detector is illuminated only by one guide star. Function s_1 is simply a constant function that equals 1/3 on the whole domain $\boldsymbol{\Omega}_T$, since it is completely illuminated by the three guide stars.

therefore be written as

$$\boldsymbol{\Phi}_{\boldsymbol{k+1}} = \boldsymbol{\Phi}_{\boldsymbol{k}} + \boldsymbol{S}\boldsymbol{A}^*(\boldsymbol{w} - \boldsymbol{A}\boldsymbol{\Phi}_{\boldsymbol{k}}). \tag{17}$$

If we then add a relaxation parameter $\lambda_{\boldsymbol{k}}$, we obtain

$$\boldsymbol{\Phi}_{\boldsymbol{k+1}} = \boldsymbol{\Phi}_{\boldsymbol{k}} + \lambda_{\boldsymbol{k}}\boldsymbol{S}\boldsymbol{A}^*(\boldsymbol{w} - \boldsymbol{A}\boldsymbol{\Phi}_{\boldsymbol{k}}). \tag{18}$$

This iteration is a weighted version of

$$\boldsymbol{\Phi}_{\boldsymbol{k+1}} = \boldsymbol{\Phi}_{\boldsymbol{k}} + \lambda_{\boldsymbol{k}}\frac{1}{G}\boldsymbol{A}^*(\boldsymbol{w} - \boldsymbol{A}\boldsymbol{\Phi}_{\boldsymbol{k}}), \tag{19}$$

the iterative method introduced by Cimmino[8] in 1938 to solve system

$$\boldsymbol{A}\boldsymbol{\Phi} = \boldsymbol{w}, \tag{20}$$

or, in a more explitic form, the system of linear equations

$$\boldsymbol{A}_{\boldsymbol{\alpha}_g}\boldsymbol{\Phi} = \begin{bmatrix} T^{\boldsymbol{\alpha}_g h_1} & \cdots & T^{\boldsymbol{\alpha}_g h_M} \end{bmatrix} \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_M \end{bmatrix} = w_{\boldsymbol{\alpha}_g} \quad g = 1, \ldots, G. \tag{21}$$

The unknown of this system is the vector of mirror deformations $\boldsymbol{\Phi}$, and the data are the open loop star-oriented wavefont measurements $w_{\boldsymbol{\alpha}_1}, \ldots, w_{\boldsymbol{\alpha}_G}$. A vector $\boldsymbol{\Phi}^*$ solves the above system if, for each direction, the sum of the mirror corrections applied on the light coming from that direction ($\boldsymbol{A}_{\boldsymbol{\alpha}_g}\boldsymbol{\Phi}^*$) equals the wavefront deformation that light experiences going through the atmosphere ($w_{\boldsymbol{\alpha}_g}$). Vector $\boldsymbol{w}$ belongs to the Hilbert space $\boldsymbol{\Omega}_T^G$ defined by the inner product

$$\langle \boldsymbol{w}, \boldsymbol{z} \rangle_G = \sum_{g=1}^{G} \langle w_{\boldsymbol{\alpha}_g}, z_{\boldsymbol{\alpha}_g} \rangle_{L_2(\boldsymbol{\Omega}_T)}.$$

We have shown that iterating a layer-oriented method considering time independent turbulence corresponds to using a weighted form of Cimmino's iterative method to find the mirror deformations vector $\boldsymbol{\Phi}$ that solves system $\boldsymbol{A}\boldsymbol{\Phi} = \boldsymbol{w}$, whose known term is the vector of star-oriented measurements. The created iterative method is in fact based on a *numerical* layer-oriented approach, where star-oriented measurements are combined, using operators $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$ and the weighting functions s_m, in order to recreate the layer-oriented WFS measurements. We must though keep in mind that in a real layer-oriented implementation light coming from the different guide stars is optically combined on the WFSs: the data are not the G wavefronts $w_{\boldsymbol{\alpha}_g}$, but the M measurements W_m. This means that, even though we managed to write these measurements in terms of the star-oriented wavefronts $w_{\boldsymbol{\alpha}_g}$, operators $\boldsymbol{A}$ and $\boldsymbol{A}^*$ and the weighting matrix $\boldsymbol{S}$ (see Eq. (16)), and used this substitution to identify the iterative layer-oriented method with a weighted Cimmino, in a real implementation term ${}_{\boldsymbol{\gamma}}\boldsymbol{W}_{\boldsymbol{k}}$ of Eq. (13) is directly obtained by the WFS measurements, with no need for matrix-vector multiplications nor calculation of inverse operators.

4. KACZMARZ APPROACH

In the Kaczmarz approach, the three steps needed to compute the optimal mirror shape (wavefront reconstruction, atmospheric turbulence tomography and mirror fitting) are performed separately, one after another. The incoming wavefronts are reconstructed from Shack-Hartmann WFS measurements; the atmospheric layers are then reconstructed by using a Kaczmarz iterative method to solve the atmospheric tomography problem and finally the optimal DM shapes are computed. For the wavefront reconstruction Ramlau and Rosensteiner used, in Ref. 6, the CuReD method (Cumulative Reconstructor with Domain decomposition) described in Ref. 9. In the following, we concentrate on the atmospheric tomography and mirror fitting steps, considering that the G wavefronts $w_{\boldsymbol{\alpha}_1}, \ldots, w_{\boldsymbol{\alpha}_G}$ have already been reconstructed with the CuReD method. In Ref. 6, the tomography problem is presented as the problem of finding a solution to the following system of linear equations:

$$\tilde{\boldsymbol{A}}_{\boldsymbol{\alpha}_g}\tilde{\boldsymbol{\Phi}} = w_{\boldsymbol{\alpha}_g} \quad g = 1, \ldots, G, \tag{22}$$

where the wavefronts $w_{\boldsymbol{\alpha}_1}, \ldots, w_{\boldsymbol{\alpha}_G}$ are the star-oriented reconstructed wavefronts, $\tilde{\boldsymbol{\Phi}}$ is the vector containing the L turbulence layers we want to find (approximation of the entire turbulence volume above the telescope)

$$\tilde{\boldsymbol{\Phi}} = \begin{bmatrix} \tilde{\Phi}_1 \\ \vdots \\ \tilde{\Phi}_L \end{bmatrix}, \tag{23}$$

and operators $\tilde{\boldsymbol{A}}_{\boldsymbol{\alpha}_g}$ are similar to the operators $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$ we introduced in Sec. 2, with the only difference that the shift operators they are composed of act on the L turbulence layers at heights $\tilde{h}_1, \ldots, \tilde{h}_L$:

$$\tilde{\boldsymbol{A}}_{\boldsymbol{\alpha}_g} := (T^{\boldsymbol{\alpha}_g \tilde{h}_1}, \ldots, T^{\boldsymbol{\alpha}_g \tilde{h}_L}). \tag{24}$$

These operators and vectors, as well as the Hilbert spaces they belong to, can therefore be defined in the same way we did for those we have used so far, simply changing $m = 1, \dots, M$ for $l = 1, \dots, L$.

Once the tomography problem described by system (22) has been solved, i.e. the L layers contained in $\tilde{\boldsymbol{\Phi}}$ have been reconstructed, we need to find the optimal shape for the DMs. If the number of DMs equals the number of reconstructed layers we can simply conjugate each DM to one turbulence layer. If, instead, we have less mirrors than reconstructed layers, the mirror shape must be found following a chosen optimality criterion. Let $M(< L)$ be the number of DMs and $\boldsymbol{\Phi}$ the vector of DM shapes. In Ref. 6 it is shown that, instead of reconstructing L layers and then finding the M mirror shapes that best correct for them, one can directly reconstruct M artificial layers located at the mirror heights h_m, solving the following system of G linear equations:

$$\boldsymbol{A}_{\boldsymbol{\alpha}_g}\boldsymbol{\Phi} = w_{\boldsymbol{\alpha}_g} \qquad g = 1, \dots, G, \tag{25}$$

where we go back to our operators $\boldsymbol{A}_{\boldsymbol{\alpha}_g}$, which act on the mirror heights $h_1, \dots, h_M$, and vector $\boldsymbol{\Phi}$, which has M components. Ramlau and Rosensteiner proposed to solve the above system using a Kaczmarz iterative algorithm, which reads as follows:

$$\boldsymbol{\Phi_{k+1}} = \boldsymbol{\Phi_k} + \boldsymbol{A}^*_{\boldsymbol{\alpha}_c}(w_{\boldsymbol{\alpha}_c} - \boldsymbol{A}_{\boldsymbol{\alpha}_c}\boldsymbol{\Phi_k}), \tag{26}$$

where index $\boldsymbol{k}$ increases by 1 at each iteration ($\boldsymbol{k} = 0, 1, 2, \dots$), while index c cyclically sweeps through the G sensed directions ($c = 1, 2, \dots, G, 1, 2, \dots, G, \dots$). This algorithm approaches a solution to system (25) by cyclically solving one of its G equations at a time: its g-th iteration gives a mirror deformation vector $\boldsymbol{\Phi_g}$ which perfectly corrects the wavefronts coming from the g-th direction. The following iteration adds to $\boldsymbol{\Phi_g}$ the correction needed in order to obtain a $\boldsymbol{\Phi_{g+1}}$ which perfectly corrects the wavefronts coming from direction $\boldsymbol{\alpha}_{g+1}$, and so on.

5. CONVERGENCE

Now that we have written the layer-oriented and the Kaczmarz approach as two different iterative methods to solve the same system of linear equations, we can analyze and compare their behaviors and convergences. We recall that, given the system

$$\boldsymbol{A}_{\boldsymbol{\alpha}_g}\boldsymbol{\Phi} = w_{\boldsymbol{\alpha}_g} \quad g = 1, 2, \dots, G, \tag{27}$$

the two considered iterative algorithms can be written as follows:

$$\text{Kaczmarz}: \quad \boldsymbol{\Phi_{k+1}} = \boldsymbol{\Phi_k} + \lambda\, \boldsymbol{A}^*_{\boldsymbol{\alpha}_g}(w_{\boldsymbol{\alpha}_g} - \boldsymbol{A}_{\boldsymbol{\alpha}_g}\boldsymbol{\Phi_k}), \tag{28}$$

$$\text{Layer} - \text{oriented}: \quad \boldsymbol{\Phi_{k+1}} = \boldsymbol{\Phi_k} + \lambda\, \boldsymbol{S}\boldsymbol{A}^*(\boldsymbol{w} - \boldsymbol{A}\boldsymbol{\Phi_k}), \tag{29}$$

where $\lambda > 0$ is a relaxation coefficient (which could also assume a different value $\lambda_{\boldsymbol{k}}$ at each iteration), $\boldsymbol{S} = \text{diag}(s_m)$ is the weighting matrix and index g cyclically goes from 1 to G. In Sec. 5.1 and 5.2 we give the mathematical description of the solutions achievable with the two methods. The main idea is that, starting with completely flat mirrors ($\boldsymbol{\Phi_0} = \boldsymbol{0}$), both iterative methods asymptotically ($k \to \infty$) find the mirror shape that minimizes the overall residual wavefront error ($\boldsymbol{\Phi^*} = \text{argmin}\{\|\boldsymbol{A\Phi} - \boldsymbol{w}\|^2\}$), where by "overall" we mean considering all sensed directions. In our convergence analysis we studied the asymptotic solution to which the Kaczmarz and the layer-oriented method converge to, and future works should be dedicated to an analysis of their convergence rates. Even though we here concentrate on *what* the two methods converge to, and not on *how fast* they do it, some interesting considerations on *how* they converge can still be done. Indeed, the way they proceed from one iteration to another is very different. In Fig. 6 we qualitatively illustrate their different behaviors. While the unrelaxed Kaczmarz method forces the solution, at each iteration, to perfectly correct the wavefronts coming from one single direction, the iterative layer-oriented approach applies at each iteration the deformation that best corrects all directions, without being able to perfectly correct none of them, but instead "going for a compromise".

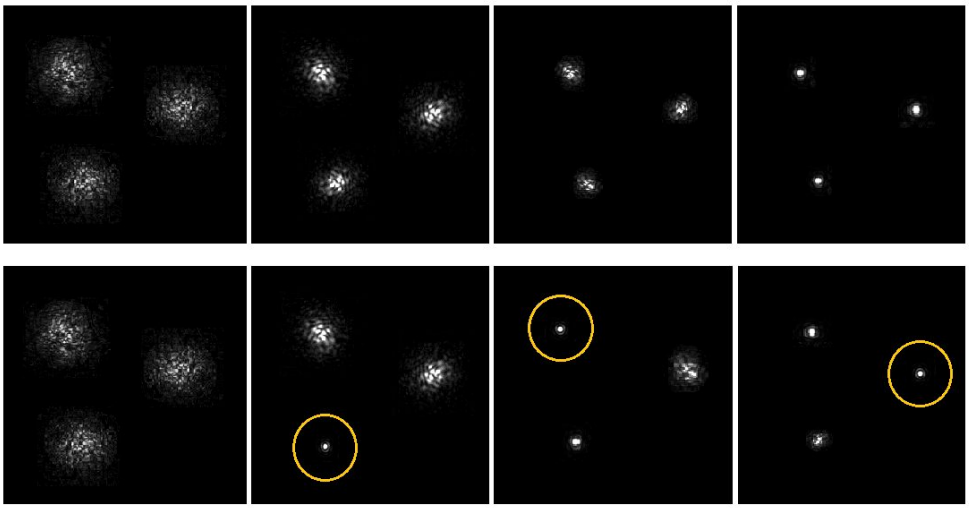

Figure 6. This figure (qualitatively) illustrates the different behaviors of the iterative layer-oriented method (first line of four panels) and the Kaczmarz iterative method (second line). The two panels on the left represent the initial situation (the same for the two methods), with three guide stars, whose PSFs are distorted by the *seeing* effects; they represent how the telescope sees the three guide stars without the MCAO correction. The other three panels represent, for each method, the results of their first three iterations, performed considering a time-fixed turbulence (we remind that we here do not consider their convergence rates, but only their general way of proceeding from one iteration to another). The iterative layer-oriented approach tries, at each iteration, to correct the three directions in a similar way. The Kaczmarz method instead selects at each iteration one direction (pointed with an orange circle), and concentrates on perfectly correcting it, without caring about what happens to the other two. This qualitative analysis of the behaviors of the two methods should of course be related to an analysis of their convergence rates, in order to be able to illustrate how much the correction in the three directions is improved at each iteration.

5.1 Kaczmarz approach

We here provide the convergence results that mathematically express what above explained about the solution reachable with the Kaczmarz method. For each of the proofs we either refer to App. A or to Ref. 10. Let $\boldsymbol{\Phi_k}$ be the k-th iteration of the Kaczmarz method (28), and let us start by considering system (27) to be a consistent system. The following theorem is taken from Ref. 11 (Theorem 3.6 of Sec. V.3) and its proof is provided in the App. A:

Theorem 1. Assume that system (27) has a solution. Then, if $0 < \lambda < 2$ and $\boldsymbol{\Phi_0} \in \text{range}(\boldsymbol{A}^*)$ (e.g. $\boldsymbol{\Phi_0} = \boldsymbol{0}$), $\boldsymbol{\Phi_k}$ converges, as $\boldsymbol{k} \to \infty$, to the solution of system (27) with minimum norm.

In the inconsistent case, this Kaczmarz method does not, in general, converge to the minimum norm solution. The following theorem (Theorem 3.9 of Ref. 11) relates the solution found in the inconsistent case to the relaxation parameter λ (see App. A for the proof):

Theorem 2. Let $0 < \lambda < 2$. Then, for $\boldsymbol{\Phi_0} \in R(\boldsymbol{A}^*)$, e.g. $\boldsymbol{\Phi_0} = \boldsymbol{0}$, the Kackmarz method (28) for an inconsistent system converges to $\boldsymbol{\Phi_\lambda} = \boldsymbol{\Phi_{MN}} + O(\lambda)$, where $\boldsymbol{\Phi_{MN}}$ minimizes $\langle(\boldsymbol{A\Phi} - \boldsymbol{w}), (\boldsymbol{A\Phi} - \boldsymbol{w})\rangle$.

A third theorem tells us something more about the reachable solution. In Ref. 10 it is shown that, assigning to the relaxation parameter λ a different value $\lambda_{\boldsymbol{k}}$ at each iteration, if $\lambda_{\boldsymbol{k}} \to 0$ as $k \to \infty$ the solution reached approaches the least squares solution of $\boldsymbol{A\Phi} = \boldsymbol{w}$ with minimum norm:

Theorem 3. Assume that system (27) has no solutions. Then, if $\lambda_{\boldsymbol{k}} \to 0$ as $k \to \infty$ and $\boldsymbol{\Phi_0} \in \text{range}(\boldsymbol{A}^*)$ (e.g. $\boldsymbol{\Phi_0} = \boldsymbol{0}$), $\boldsymbol{\Phi_k}$ converges, as $\boldsymbol{k} \to \infty$, to the minimum norm least squares solution of system (27), i.e. the minimizer of $\langle(\boldsymbol{A\Phi} - \boldsymbol{w}), (\boldsymbol{A\Phi} - \boldsymbol{w})\rangle$ that has minimum norm.

For the proof of this theorem we refer to Ref. 10, where the argument is extensively treated and the simpler case of $\boldsymbol{\Phi} \in R^M$ largely discussed.

5.2 Layer-oriented approach

In order to study the convergence of the layer-oriented method, we first notice its equivalence to the DROP algorithm (Diagonally-Relaxed Orthogonal Projection) introduced in Ref. 12 in 2008, a diagonally component-wise relaxed version of Cimmino's method. In the DROP method, the diagonal values of $\boldsymbol{S}$ are simple constants $1/s_m$, where s_m is the number of non-zero elements in the m-th column of matrix $\boldsymbol{A}$. In our layer-oriented method these diagonal values are functions which equal a different value on different areas of their domain $\boldsymbol{\Omega}_m$. On each point of $\boldsymbol{\Omega}_m$, the value of the function s_m equals $1/s$, s being the number of light beams that overlap on that point. Even if they may appear different, the two $\boldsymbol{S}$ matrices actually do the same thing: they take into account the sparsity of the system and replace Cimmino's constant division by $1/G$ with more appropriate component-dependent weights. Actually, by slightly modifying the description of the system we have used so far, it is easy to show that the layer-oriented method can be written exactly as the DROP method (see App. B). This simplifies our study of the convergence of method (29), and allows us to use the known convergence results for the DROP method. We can therefore, from now on, think of matrix $\boldsymbol{S}$ as a diagonal matrix of constants, and we start with the observation that, if $\boldsymbol{S} = \mathbb{I}$, iteration (29) becomes

$$\boldsymbol{\Phi_{k+1}} = \boldsymbol{\Phi_k} + \lambda \boldsymbol{A}^*(\boldsymbol{w} - \boldsymbol{A}\boldsymbol{\Phi_k}). \tag{30}$$

This iteration represents the so called Landweber algorithm, which is known to converge to the solution, or the least squares solution, closest to $\boldsymbol{\Phi_0}$ if $0 < \lambda < 2/\rho(\boldsymbol{A}^*\boldsymbol{A})$, being $\rho(\boldsymbol{A}^*\boldsymbol{A})$ the spectral radius of matrix $\boldsymbol{A}^*\boldsymbol{A}$, i.e. its largest, in absolute value, eigenvalue. Given this, and calling for simplicity H_1 the Hilbert space L_2^M of vector $\boldsymbol{\Phi_0}$, we can prove the following theorem (Theorem 2.3 of Ref. 12):

Theorem 4. Let λ in algorithm (29) assume a different value $\lambda_{\boldsymbol{k}}$ at each iteration. If each $\lambda_{\boldsymbol{k}}$ satisfies

$$0 < \epsilon \leq \lambda_k \leq (2-\epsilon)/\rho(\boldsymbol{S}\boldsymbol{A}^*\boldsymbol{A}),$$

where ϵ is an arbitrary small but fixed constant, then any sequence $\{\boldsymbol{\Phi_k}\}_{k=0}^{\infty}$ generated by algorithm (29) converges to a least squares solution

$$\boldsymbol{\Phi^*} = \text{argmin}\{\|\boldsymbol{A}\boldsymbol{\Phi} - \boldsymbol{w}\|^2 : \boldsymbol{\Phi} \in H_1\}.$$

If, in addition, $\boldsymbol{\Phi_0} \in R(\boldsymbol{S}\boldsymbol{A}^*)$, the range of $\boldsymbol{S}\boldsymbol{A}^*$, then $\boldsymbol{\Phi^*}$ has minimum $\boldsymbol{S}^{-1}$-norm.

Proof. Using the transformations $\boldsymbol{y_k} = \boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\Phi_k}$ and $\bar{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{S}^{\frac{1}{2}}$, the iterative step (29) becomes

$$\boldsymbol{y_{k+1}} = \boldsymbol{y_k} + \lambda_k \bar{\boldsymbol{A}}^*(\boldsymbol{w} - \bar{\boldsymbol{A}}\boldsymbol{y_{k+1}}).$$

Assuming that

$$0 < \epsilon \leq \lambda_k \leq (2-\epsilon)/\rho(\bar{\boldsymbol{A}}^*\bar{\boldsymbol{A}}), \tag{31}$$

we apply the convergence results for the Landweber algorithm (30) to conclude that

$$\lim_{k\to\infty} \boldsymbol{y_k} = \boldsymbol{y}^* \text{ and } \boldsymbol{y}^* = \text{argmin}\{\|\bar{\boldsymbol{A}}\boldsymbol{y} - \boldsymbol{w}\|^2 : \boldsymbol{y} \in \mathbb{H}_1\},$$

which implies that

$$\lim_{k\to\infty} \boldsymbol{\Phi_k} = \boldsymbol{S}^{\frac{1}{2}}\boldsymbol{y}^* = \boldsymbol{\Phi}^* \text{ and } \boldsymbol{\Phi}^* = \text{argmin}\{\|\boldsymbol{A}\boldsymbol{\Phi} - \boldsymbol{w}\|^2 : \boldsymbol{\Phi} \in \mathbb{H}_1\}.$$

Also, if $\boldsymbol{y_0} \in R(\bar{\boldsymbol{A}}^*)$ then $\boldsymbol{y}^*$ has minimum norm. Hence, by using

$$\|\boldsymbol{y}^*\| = \|\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{y}^*\| = \|\boldsymbol{\Phi}^*\|_{\boldsymbol{S}^{-1}},$$

it follows that $\boldsymbol{\Phi}^*$ has minimum $\boldsymbol{S}^{-1}$-norm provided that $\boldsymbol{\Phi_0} = \boldsymbol{S}^{\frac{1}{2}}\boldsymbol{y_0} \in R(\boldsymbol{S}\boldsymbol{A}^*)$.
Finally,

$$\rho(\bar{\boldsymbol{A}}^*\bar{\boldsymbol{A}}) = \rho(\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{S}^{\frac{1}{2}}) = \rho(\boldsymbol{S}^{\frac{1}{2}}(\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{S}^{\frac{1}{2}})\boldsymbol{S}^{-\frac{1}{2}}) = \rho(\boldsymbol{S}\boldsymbol{A}^*\boldsymbol{A}).$$

□

6. CONCLUSIONS

The problem of finding the optimal mirror shape for an MCAO system can be written in terms of a system of linear equations. The known term of this system is given by the wavefront deformations observed in the guide stars' directions, while the unknown is the shape to be given to the deformable mirrors in order to correct for the observed turbulence effects. We wrote the layer-oriented and the star-oriented method proposed in Ref. 6 as two different methods to solve the same MCAO system of linear equations. Considering an atmospheric turbulence not changing with time, we compared the solutions theoretically achievable by an infinite iteration of the two different methods. While the star-oriented approach of Ref. 6 is there identified with Kaczmarz's iterative method, we showed that the layer-oriented approach corresponds to a weighted form of Cimmino's iterative method. Using the theoretical results known for these two methods, we proved that the two MCAO approaches asymptotically tend to minimize the same quantity, which represents the overall (i.e. considering all sensed directions) residual wavefront error.

Having the two methods written in the same theoretical framework, allowed us to further investigate their different behaviors. While the Kaczmarz star-oriented iteratively forces the solution to achieve a perfect matching in one direction at a time, the layer-oriented performs a solution which does not privileges any direction. Although this has just a sort of historical perspective, it is interesting to notice that the second method, the layer-oriented, is better suited for a limited number of iterations. In fact, in its implementation in the sky the number of iterations performed before commands being applied is actually just one. This single-iteration approach would turn out to produce largely non uniform effects in the field of view when using Kaczmarz's method, which, however, has been introduced assuming a large number of iterations to be performed before applying the mirror commands. Future works need to provide these considerations with more quantitative convergence studies. The analysis presented in this work can be used as a theoretical basis for future experimental simulations aimed at studying and comparing the actual behaviors of the two methods. Our preliminary conclusions in terms of speed of convergence further suggest that several other approaches invented almost a century ago could be of some practical interest.

APPENDIX A. MATHEMATICAL PROOFS

We here consider the Kaczmarz iterative algorithm used to solve the system of linear equations

$$\boldsymbol{A}_g\boldsymbol{\Phi} = w_g \quad g = 1, 2, \dots G. \tag{32}$$

We remind that bold letters are used for vectors and matrices. Operators are normally indicated with light letters, except those that are defined as vector of operators, which are instead indicated with bold letters. Let $\mathbb{H}_1$ and $\mathbb{H}_T$ be Hilbert spaces, and let $\boldsymbol{A}_g : \mathbb{H}_1 \to \mathbb{H}_T$ be bounded linear operators from $\mathbb{H}_1$ to $\mathbb{H}_T$. Let $w_g \in \mathbb{H}_T$ for $g = 1, \dots, G$ be given. Let P_g be the orthogonal projection of $\mathbb{H}_1$ onto the affine subspace $\boldsymbol{A}_g\boldsymbol{\Phi} = w_g$, and let

$$P_g^\lambda = (1-\lambda)\mathbb{I} + \lambda P_g, \tag{33}$$

$$P^\lambda = P_G^\lambda \dots P_1^\lambda. \tag{34}$$

where λ is a relaxation parameter. Then, the Kaczmarz method (with relaxation) for the solution of system (32) can be written as

$$\boldsymbol{\Phi_{k+1}} = P^\lambda \boldsymbol{\Phi_k}, \quad \boldsymbol{k} = 0, 1, \dots \tag{35}$$

with $\boldsymbol{\Phi_0} \in \mathbb{H}_1$ arbitrary. Follows the proof for **Theorem 1**, which we here recall:

Theorem 1. Assume that system (32) has a solution. Then, if $0 < \lambda < 2$ and $\boldsymbol{\Phi_0} \in \text{range}(\boldsymbol{A}^*)$ (e.g. $\boldsymbol{\Phi_0} = \boldsymbol{0}$), $\boldsymbol{\Phi_k}$ converges, as $\boldsymbol{k} \to \infty$, to the solution of system (32) with minimum norm.

Proof. Let Q_g be the orthogonal projection onto $\ker(\boldsymbol{A}_g)$, and let $\boldsymbol{\Phi}^*$ be any solution of (32). Then, for $\boldsymbol{\Psi}$ in $\mathbb{H}_1$, we have that

$$P_g\boldsymbol{\Psi} = \boldsymbol{\Phi}^* + Q_g(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*), \tag{36}$$
$$P_g^\lambda\boldsymbol{\Psi} = \boldsymbol{\Phi}^* + Q_g^\lambda(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*), \tag{37}$$
$$P^\lambda\boldsymbol{\Psi} = \boldsymbol{\Phi}^* + Q^\lambda(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*), \tag{38}$$
$$(P^\lambda)^k\boldsymbol{\Psi} = \boldsymbol{\Phi}^* + (Q^\lambda)^k(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*). \tag{39}$$

Equation (36) comes from the fact that, since $\boldsymbol{A}_g(P_g\boldsymbol{\Psi}) = w_g$, we can write $\boldsymbol{A}_g(P_g\boldsymbol{\Psi}) = \boldsymbol{A}_g\boldsymbol{\Phi}^* + \boldsymbol{A}_g(\boldsymbol{\Gamma})$, with $\boldsymbol{\Gamma} \in \ker(\boldsymbol{A}_g)$. It is easy to show that Eq. (37) follows from Eq. (36). Using Eq. (33) we can in fact re-write the right hand side of Eq. (37) as

$$P_g^\lambda\boldsymbol{\Psi} = (1-\lambda)\boldsymbol{\Psi} + \lambda P_g\boldsymbol{\Psi}$$

and its left hand side as

$$\begin{aligned}\boldsymbol{\Phi}^* + Q_g^\lambda(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*) &= \boldsymbol{\Phi}^* + (1-\lambda)(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*) + \lambda Q_g(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*)\\ &= \boldsymbol{\Phi}^* + (1-\lambda)\boldsymbol{\Psi} - (1-\lambda)\boldsymbol{\Phi}^* + \lambda Q_g(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*)\\ &= (1-\lambda)\boldsymbol{\Psi} + \lambda\boldsymbol{\Phi}^* + \lambda Q_g(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*).\end{aligned}$$

By equating these new expressions for the two sides of Eq. (37) we obtain

$$\begin{aligned}(1-\lambda)\boldsymbol{\Psi} + \lambda P_g\boldsymbol{\Psi} &= (1-\lambda)\boldsymbol{\Psi} + \lambda\boldsymbol{\Phi}^* + \lambda Q_g(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*),\\ P_g\boldsymbol{\Psi} &= \boldsymbol{\Phi}^* + Q_g(\boldsymbol{\Psi} - \boldsymbol{\Phi}^*),\end{aligned}$$

which is Eq. (36). As concerns Eq. (38) and (39), they follow from the fact that, being $\boldsymbol{\Phi}^*$ a solution of the system, we have that $P_g^\lambda\boldsymbol{\Phi}^* = \boldsymbol{\Phi}^*$ and also $P^\lambda\boldsymbol{\Phi}^* = \boldsymbol{\Phi}^*$, for every $g = 1, \ldots, G$ and every value of λ. We can now write the Kaczmarz iteration as

$$\boldsymbol{\Phi_k} = (P^\lambda)^k\boldsymbol{\Phi_0} \tag{40}$$
$$= \boldsymbol{\Phi}^* + (Q^\lambda)^k(\boldsymbol{\Phi_0} - \boldsymbol{\Phi}^*). \tag{41}$$

We now need to introduce the following lemma, which tells us what operator $(Q^\lambda)^k$ converges to when $k \to \infty$.

Lemma 1. For $0 < \lambda < 2$, $(Q^\lambda)^k$ converges strongly, as $k \to \infty$, to the orthogonal projection onto

$$\bigcap_{g=1}^{G} \ker(\mathbb{I} - Q_g).$$

Using Lemma 1 we can write

$$\boldsymbol{\Phi_k} = \boldsymbol{\Phi}^* + (Q^\lambda)^k(\boldsymbol{\Phi_0} - \boldsymbol{\Phi}^*) \quad \longrightarrow \quad \boldsymbol{\Phi}^* + T(\boldsymbol{\Phi_0} - \boldsymbol{\Phi}^*) = (\mathbb{I} - T)\boldsymbol{\Phi}^* + T\boldsymbol{\Phi_0}$$

as $k \to \infty$, where T is the orthogonal projection onto

$$\boldsymbol{A} = \bigcap_{g=1}^{G} \ker(\boldsymbol{A}_g).$$

Now, if $\boldsymbol{\Phi_0} \in \text{range}(\boldsymbol{A}^*) = \text{range}(\boldsymbol{A}_1^*) + \text{range}(\boldsymbol{A}_2^*) + \cdots + \text{range}(\boldsymbol{A}_G^*)$, then $T\boldsymbol{\Phi_0} = 0$ and $(\mathbb{I} - T)\boldsymbol{\Phi}^*$ is the solution of system (32) with minimum norm, which ends our proof for Theorem 1. □

We now need to proof the lemma we used. In order to do that, we first need to prove the following lemmas, taken from Ref. 11.

Lemma 2. Let T be a linear map in the Hilbert space $\mathbb{H}$ with $\|T\| \leq 1$. Then

$$\mathbb{H} = \ker(\mathbb{I} - T) \oplus \overline{\text{range}(\mathbb{I} - T)}.$$

Proof. For an arbitrary linear bounded map A in $\mathbb{H}$ we have

$$\mathbb{H} = \ker(A^*) \oplus \overline{\text{range}(A)}.$$

The lemma follows by putting $A = \mathbb{I} - T$ and showing that

$$\ker(\mathbb{I} - T) = \ker(\mathbb{I} - T^*).$$

Let $\mathbf{\Psi} \in \ker(\mathbb{I} - T)$, i.e. $\mathbf{\Psi} = T\mathbf{\Psi}$. Then

$$(\mathbf{\Psi}, T^*\mathbf{\Psi}) = (T\mathbf{\Psi}, \mathbf{\Psi}) = (\mathbf{\Psi}, \mathbf{\Psi}).$$

It follows that

$$\begin{aligned}
\|\mathbf{\Psi} - T^*\mathbf{\Psi}\|^2 &= (\mathbf{\Psi} - T^*\mathbf{\Psi}, \mathbf{\Psi} - T^*\mathbf{\Psi}) \\
&= (\mathbf{\Psi}, \mathbf{\Psi}) - 2(\mathbf{\Psi}, T^*\mathbf{\Psi}) + (T^*\mathbf{\Psi}, T^*\mathbf{\Psi}) \\
&= -(\mathbf{\Psi}, \mathbf{\Psi}) + \|T^*\mathbf{\Psi}\|^2 \\
&\leq -(\mathbf{\Psi}, \mathbf{\Psi}) + \|T^*\|^2\|\mathbf{\Psi}\|^2 \\
&\leq 0.
\end{aligned}$$

Hence $\mathbf{\Psi} \in \ker(\mathbb{I} - T^*)$. This shows that

$$\ker(\mathbb{I} - T) \subseteq \ker(\mathbb{I} - T^*)$$

and the opposite inclusion follows by symmetry. □

Let now Q_g be a linear orthogonal projection in $\mathbb{H}$, and let Q_g^λ be defined as

$$\begin{aligned}
Q_g^\lambda &= (1 - \lambda)\mathbb{I} + \lambda Q_g, \\
Q^\lambda &= Q_G^\lambda \dots Q_1^\lambda.
\end{aligned}$$

Since $Q_g = Q_g^2 = Q_g^*$ one verifies easily that

$$\|Q_g^\lambda\| \leq 1, \quad 0 < \lambda < 2, \tag{42}$$

$$\|Q_g^\lambda \mathbf{\Phi}\|^2 - \|\mathbf{\Phi}\|^2 = (2 - \lambda)\lambda(\|Q_g\mathbf{\Phi}\|^2 - \|\mathbf{\Phi}\|^2). \tag{43}$$

Lemma 3. Let $(\mathbf{\Phi_k})_{k=0,1,\dots}$ be a sequence in $\mathbb{H}$ such that

$$\|\mathbf{\Phi_k}\| \leq 1, \quad \lim_{k\to\infty} \|Q^\lambda \mathbf{\Phi_k}\| = 1.$$

Then we have, for $0 < \lambda < 2$,

$$\lim_{k\to\infty} (\mathbb{I} - Q^\lambda)\mathbf{\Phi_k} = 0.$$

Proof. The proof is by induction with respect to the number G of factors in Q^λ. For $G = 1$ we have

$$\begin{aligned}
\|(\mathbb{I} - Q^\lambda)\mathbf{\Phi_k}\|^2 &= \|(\mathbb{I} - Q_1^\lambda)\mathbf{\Phi_k}\|^2 = \omega^2\|(\mathbb{I} - Q_1)\mathbf{\Phi_k}\|^2 \\
&= \lambda^2(\|\mathbf{\Phi_k}\|^2 + \|Q_1\mathbf{\Phi_k}\|^2 - 2(\mathbf{\Phi_k}, Q_1\mathbf{\Phi_k})) \\
&= \lambda^2(\|\mathbf{\Phi_k}\|^2 - \|Q_1\mathbf{\Phi_k}\|^2) \\
&= \frac{\lambda}{2 - \lambda}(\|\mathbf{\Phi_k}\|^2 - \|Q_1\mathbf{\Phi_k}\|^2),
\end{aligned}$$

where we used Eq. (43). If $\|\mathbf{\Phi_k}\| \leq 1$ and $\|Q_1^\lambda \mathbf{\Phi_k}\| \to 1$, it follows that, for $0 < \lambda < 2$, $\|(\mathbb{I} - Q^\lambda)\| \to 0$, hence the lemma for $G = 1$. Now assume the lemma to be correct for $G - 1$ factors. We put

$$Q^\lambda = Q_G^\lambda S^\lambda, \quad S^\lambda = Q_{G-1}^\lambda \dots Q_1^\lambda$$

and we then have

$$(\mathbb{I} - Q^\lambda)\mathbf{\Phi_k} = (\mathbb{I} - S^\lambda)\mathbf{\Phi_k} + (S^\lambda - Q^\lambda)\mathbf{\Phi_k} \tag{44}$$

$$= (\mathbb{I} - S^\lambda)\mathbf{\Phi_k} + (\mathbb{I} - Q_G^\lambda)S^\lambda \mathbf{\Phi_k}. \tag{45}$$

Now let $\|\mathbf{\Phi_k}\| \leq 1$ and $\|Q_1^\lambda \mathbf{\Phi_k}\| \to 1$. Because of the inequalities in (42) we have

$$\|Q^\lambda \mathbf{\Phi_k}\| = \|Q_G^\lambda S^\lambda \mathbf{\Phi_k}\| \leq \|S^\lambda \mathbf{\Phi_k}\| \leq \|\mathbf{\Phi_k}\|,$$

hence $\|S^\lambda \mathbf{\Phi_k}\| \to 0$. Applying the lemma for the single factor Q_G^λ to the sequence $\mathbf{\Phi}'_\mathbf{k} = S^\lambda \mathbf{\Phi_k}$ we also obtain $(\mathbb{I} - Q_G^\lambda)S^\lambda \mathbf{\Phi_k} \to 0$. Hence, from Eq. (45), it follows that $(\mathbb{I} - Q^\lambda)\mathbf{\Phi_k} \to 0$. □

Lemma 4. For $0 < \lambda < 2$, $(Q^\lambda)^k$ converges, as $k \to \infty$, strongly to the orthogonal projection onto $\ker(\mathbb{I} - Q^\lambda)$.

Proof. Let T be the orthogonal projection onto $\ker(\mathbb{I} - Q^\lambda)$. From Lemma 2 and inequalities in (42) we know that

$$\mathbb{H} = \ker(\mathbb{I} - Q^\lambda) \oplus \overline{\text{range}(\mathbb{I} - Q^\lambda)}, \tag{46}$$

hence $\mathbb{I} - T$ is the orthogonal projection onto $\overline{\text{range}(\mathbb{I} - Q^\lambda)}$. Thus

$$(\mathbb{I} - T)(\mathbb{I} - Q^\lambda) = \mathbb{I} - Q^\lambda,$$
$$(\mathbb{I} - Q^\lambda)T = 0.$$

From the first equation we get $T = TQ^\lambda$, from the second one $T = Q^\lambda T$. In particular, T and Q^λ commute. Now let $\mathbf{\Phi} \in \mathbb{H}$. The sequence $(\|(Q^\lambda)^k \mathbf{\Phi}\|)_{k=0,1,\dots}$ is decreasing, hence its limit c exists. If $c = 0$ we get, using $T = Q^\lambda T = TQ^\lambda$,

$$T\mathbf{\Phi} = (Q^\lambda)^k T\mathbf{\Phi} = T(Q^\lambda)^k \mathbf{\Phi} \to 0 \ \text{ as } \ k \to \infty.$$

Since $T\mathbf{\Phi}$ doesn't depend on k, having $T\mathbf{\Phi} \to 0$ as $k \to \infty$ implies $T\mathbf{\Phi} = 0$. We therefore have

$$\lim_{k\to\infty} (Q^\lambda)^k \mathbf{\Phi} = 0, \quad T\mathbf{\Phi} = 0,$$

i.e. $(Q^\lambda)^k \mathbf{\Phi} \to T\mathbf{\Phi}$. If $c > 0$ we put

$$\mathbf{\Psi_k} = \|(Q^\lambda)^k \mathbf{\Phi}\|^{-1} (Q^\lambda)^k \mathbf{\Phi}.$$

Then we have

$$\|\mathbf{\Psi_k}\| = 1, \quad \lim_{k\to\infty} \|Q^\lambda \mathbf{\Psi_k}\| = 1,$$

hence we obtain, from Lemma 3,

$$\lim_{k\to\infty} (\mathbb{I} - Q^\lambda)\mathbf{\Psi_k} = 0$$

or

$$\lim_{k\to\infty} (\mathbb{I} - Q^\lambda)(Q^\lambda)^k \mathbf{\Phi} = 0.$$

It follows that $(Q^\lambda)^k$ converges strongly to 0 on $\text{range}(\mathbb{I} - Q^\lambda)$, and this extends to the closure since the $(Q^\lambda)^k$ are uniformly bounded. On $\ker(\mathbb{I} - Q^\lambda)$, $(Q^\lambda)^k$ converges trivially to $\mathbb{I}$. The lemma follows then from Eq. (46). □

Lemma 5. For $0 < \lambda < 2$ we have

$$\ker(\mathbb{I} - Q^\lambda) = \bigcap_{g=1}^{G} \ker(\mathbb{I} - Q_g).$$

Proof. A fixed point of Q_g is also a fixed point of Q_g^λ. This settles the inclusion '$\supset$'. On the other hand, if $Q^\lambda \mathbf{\Phi}$, then, using inequalitiy (42),

$$\|\mathbf{\Phi}\| = \|Q_G^\lambda \dots Q_2^\lambda Q_1^\lambda \mathbf{\Phi}\| = \|Q^\lambda \mathbf{\Phi}\| \leq \|\mathbf{\Phi}\|.$$

Hence, $\|Q^\lambda \mathbf{\Phi}\| = \|\mathbf{\Phi}\|$, and from Eq. (43), $\|Q_1 \mathbf{\Phi}\| = \|\mathbf{\Phi}\|$. Since Q_1 is a projection, $Q_1 \mathbf{\Phi} = \mathbf{\Phi}$. Since now

$$\mathbf{\Phi} = Q_G^\lambda \dots Q_2^\lambda \mathbf{\Phi},$$

we can show in the same way that $Q_2 \mathbf{\Phi} = \mathbf{\Phi}$ and so forth $Q_3 \boldsymbol{x} = \dots = Q_G \mathbf{\Phi} = \boldsymbol{x}$. This settles inclusion '$\subset$'. □

The proof of Lemma 1 immediately follows from the preceding lemmas.

We now prove Theorem 2, which read:

Theorem 2. Let $0 < \lambda < 2$. Then, for $\mathbf{\Phi_0} \in R(\boldsymbol{A}^*)$, e.g. $\mathbf{\Phi_0} = \mathbf{0}$, the Kackmarz method (28) for an inconsistent system converges to $\mathbf{\Phi}_\lambda = \mathbf{\Phi}_{\boldsymbol{MN}} + O(\lambda)$, where $\mathbf{\Phi}_{\boldsymbol{MN}}$ minimizes $\langle (\boldsymbol{A}\mathbf{\Phi} - \boldsymbol{w}), (\boldsymbol{A}\mathbf{\Phi} - \boldsymbol{w}) \rangle$.

Proof. To prove this theorem we first write the Kaczmarz method to solve system $\boldsymbol{A}\mathbf{\Phi} = \boldsymbol{w}$ in the following way:

$$\mathbf{\Phi_{k+1}} = \boldsymbol{B}_\lambda \mathbf{\Phi_k} + \boldsymbol{d}_\lambda, \tag{47}$$

where

$$\boldsymbol{A}\boldsymbol{A}^* = \begin{pmatrix} \boldsymbol{A}_1\boldsymbol{A}_1^* & \dots & \boldsymbol{A}_1\boldsymbol{A}_G^* \\ \vdots & & \vdots \\ \boldsymbol{A}_G\boldsymbol{A}_1^* & \dots & \boldsymbol{A}_G\boldsymbol{A}_G^* \end{pmatrix} = \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{L}^*,$$

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{A}_1\boldsymbol{A}_1^* & & 0 \\ & \ddots & \\ 0 & & \boldsymbol{A}_G\boldsymbol{A}_G^* \end{pmatrix} = \begin{pmatrix} \mathbb{I} & & 0 \\ & \ddots & \\ 0 & & \mathbb{I} \end{pmatrix},$$

$$\boldsymbol{L} = \begin{pmatrix} 0 & \dots & \dots & 0 \\ \boldsymbol{A}_2\boldsymbol{A}_1^* & 0 & \dots & 0 \\ \vdots & \ddots & 0 & 0 \\ \boldsymbol{A}_G\boldsymbol{A}_1^* & \dots & \boldsymbol{A}_G\boldsymbol{A}_{G-1}^* & 0 \end{pmatrix},$$

$$\boldsymbol{B}_\lambda = \mathbb{I} - \lambda \boldsymbol{A}^* (\boldsymbol{D} + \lambda \boldsymbol{L})^{-1} \boldsymbol{A},$$

$$\boldsymbol{d}_\lambda = \lambda \boldsymbol{A}^* (\boldsymbol{D} + \lambda \boldsymbol{L})^{-1} \boldsymbol{w}.$$

Since $\mathbf{\Phi_0}$ and $\boldsymbol{d}_\lambda \in R(\boldsymbol{A}^*)$ the iteration takes place in $R(\boldsymbol{A}^*)$ where $\boldsymbol{B}_\lambda$ is a contraction, according to the following lemma:

Lemma 6. (Lemma 3.8 of Ref. 11) Let $0 < \lambda < 2$. Then, the restriction $\boldsymbol{B}'_\lambda$ of $\boldsymbol{B}_\lambda$ to $R(\boldsymbol{A}^*)$ satisfies $\rho(\boldsymbol{B}'_\lambda) < 1$.

Hence the convergence of $\mathbf{\Phi_k}$ to the unique (consistent case) solution $\mathbf{\Phi}_\lambda \in R(\boldsymbol{A}^*)$ of $\mathbf{\Phi}_\lambda = \boldsymbol{B}_\lambda \mathbf{\Phi}_\lambda + \boldsymbol{d}_\lambda$ follows from the theory of iterative methods. $\mathbf{\Phi}_\lambda$ can equivalently be defined as the unique solution of

$$\boldsymbol{A}^* (\boldsymbol{D} + \lambda \boldsymbol{L})^{-1} (\boldsymbol{w} - \boldsymbol{A}\mathbf{\Phi}_\lambda) = 0,$$

which, since in our case $\boldsymbol{D} = \mathbb{I}$,

$$\boldsymbol{A}^* (\mathbb{I} + \lambda \boldsymbol{L})^{-1} (\boldsymbol{w} - \boldsymbol{A}\mathbf{\Phi}_\lambda) = 0. \tag{48}$$

In the case of inconsistent system the minimum norm solution is determined by

$$\boldsymbol{A}^* \boldsymbol{A} \mathbf{\Phi}_{\boldsymbol{MN}} = \boldsymbol{A}^* \boldsymbol{w},$$
$$\boldsymbol{A}^* (\boldsymbol{w} - \boldsymbol{A}\mathbf{\Phi}_{\boldsymbol{MN}}) = 0,$$

which differs from Eq. (48) by $O(\lambda)$.

□

APPENDIX B. LAYER-ORIENTED AS THE DROP METHOD

We remind that the layer-oriented iterative method reads as

$$\boldsymbol{\Phi_{k+1}} = \boldsymbol{\Phi_k} + \lambda\, \boldsymbol{S}\boldsymbol{A^*}(\boldsymbol{w} - \boldsymbol{A}\boldsymbol{\Phi_k}), \tag{49}$$

where the diagonal values of $\boldsymbol{S}$ are functions which equal a different value on different areas of their domain $\boldsymbol{\Omega}_m$. On each point of $\boldsymbol{\Omega}_m$, the value of the function s_m equals $1/s$, s being the number of light beams that overlap on that point. In the DROP method introduced in Ref. 12 the diagonal values of $\boldsymbol{S}$ are simple constants $1/s_m$, where s_m is the number of non-zero elements in the m-th column of matrix $\boldsymbol{A}$. Despite this difference in the matrix $\boldsymbol{S}$, method (49) is equivalent to the DROP method. To explain this we will now formulate a system equivalent to $\boldsymbol{A\Phi} = \boldsymbol{w}$ and a method to solve it that is equivalent to (49), but with the exact $\boldsymbol{S}$ matrix of the DROP method. First, imagine to split each m-th component of vector $\boldsymbol{\Phi}$ into a vector of p_m components, each being a function of the space $L_2(\boldsymbol{\Omega}_m)$ that equals zero on the whole domain $\boldsymbol{\Omega}_m$ except for an area delimited by a change in the value of the function s_m. Referring to Fig. 5, for example, component Φ_1 remains a single component, but Φ_2 splits into 7 components, each of them being a function of $\boldsymbol{\Omega}_m$. The first component, $\Phi_2^{(1)}$,is a function that equals zero on the whole $\boldsymbol{\Omega}_m$ except for the light blue area on the left, defined by $s_m = 1$, where it equals Φ_2; component $\Phi_2^{(2)}$ is a function equal to zero on the whole $\boldsymbol{\Omega}_m$ except for the light blue area on the right, where it equals Φ_2, and so on until component $\Phi_2^{(7)}$, which is different from zero, and equal to Φ_2, only on the dark blue area defined by $s_m = 1/3$. Each vector $\boldsymbol{\Phi}$ can then be replaced by the P-dimensional vector ${}^P\boldsymbol{\Phi}$ defined in a new Hilbert space, where $P = \sum_{m=1}^{M} p_m$. In the example of Fig. 5, $P = 8$, which means that the new vector ${}^8\boldsymbol{\Phi}$ has 8 components. Vector ${}^P\boldsymbol{\Phi}$ contains the same information as $\boldsymbol{\Phi}$: it contains the same M mirror shapes contained in vector $\boldsymbol{\Phi}$, but shuttered and placed on the P components. Indeed, by properly defining the inner product of the new Hilbert space, the new system operators and their adjoints, it is easy to think of the following system:

$${}^P\boldsymbol{A}\,{}^P\boldsymbol{\Phi} = \boldsymbol{w}, \tag{50}$$

which represents the problem of finding a vector ${}^P\boldsymbol{\Phi}$ from the same data $\boldsymbol{w}$ as before. Operator ${}^P\boldsymbol{A}$ is composed of G operators ${}^P\boldsymbol{A}_{\boldsymbol{\alpha}_g}$, each itself composed of P (instead of M) shift operators, in such a way that ${}^P\boldsymbol{A}_{\boldsymbol{\alpha}_g}\,{}^P\boldsymbol{\Phi} = \boldsymbol{A}_{\boldsymbol{\alpha}_g}\boldsymbol{\Phi}$. We can now write, for the new system (50), an iterative method similar to the one given by Eq. (49):

$${}^P\boldsymbol{\Phi_{k+1}} = {}^P\,\boldsymbol{\Phi_k} + \boldsymbol{S}\sum_{g=1}^{G} {}^P\boldsymbol{A}^*_{\boldsymbol{\alpha}_g}(w_{\boldsymbol{\alpha}_g} - {}^P\,\boldsymbol{A}_{\boldsymbol{\alpha}_g}\,{}^P\boldsymbol{\Phi_k}), \tag{51}$$

where now matrix $\boldsymbol{S}$ is the P-dimensional diagonal matrix $\boldsymbol{S} = \mathrm{diag}(1/s_{m,p})$, with values $s_{m,p}$ being constants corresponding to each of the different values that functions s_m had on the different parts of their domains $\boldsymbol{\Omega}_m$. Referring to the example illustrated in Fig. 5, the new $\boldsymbol{S}$ matrix would be $\boldsymbol{S} = \mathrm{diag}(1/3, 1/2, 1/2, 1/2, 1, 1, 1, 1/3)$. Solution ${}^P\boldsymbol{\Phi}^*$, obtained by iterating method (51), contains the same mirror shapes contained in the solution $\boldsymbol{\Phi}^*$ obtained by iterating method (49), but divided in P parts, and we can therefore consider iteration (49) to be a DROP iteration as defined in Ref. 12, i.e. with the "correct" $\boldsymbol{S}$ matrix.

REFERENCES

[1] Beckers, J. M., "Increasing the size of the isoplanatic patch with multiconjugate adaptive optics," in [*European Southern Observatory Conference and Workshop Proceedings*], **30**, 693 (1988).

[2] Rigaut, F. J., Ellerbroek, B. L., and Flicker, R., "Principles, limitations, and performance of multiconjugate adaptive optics," 1022–1031 (2000).

[3] Ragazzoni, R., Marchetti, E., and Valente, G., "Adaptive-optics corrections available for the whole sky," *Nature* **403**(6765), 54–6 (2000).

[4] Ragazzoni, R., "Adaptive optics for giant telescopes: NGS vs. LGS," in [*European Southern Observatory Conference and Workshop Proceedings*], **57**, 175 (2000).

[5] Ragazzoni, R., Farinato, J., and Marchetti, E., "Adaptive optics for 100-m-class telescopes: new challenges require new solutions," in [*Adaptive Optical Systems Technology*], **4007**, 1076–1087 (2000).

[6] Ramlau, R. and Rosensteiner, M., "A Kaczmarz type reconstructor for MCAO data," *Inverse Problems* **28**, 8277–8287 (2012).
[7] Diolaiti, E., Ragazzoni, R., and Tordi, M., "Closed loop performance of a layer-oriented multi-conjugate adaptive optics system," *Astronomy & Astrophysics* **372**(2), 710–718 (2001).
[8] Cimmino, G., "Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari, XVI (9) (1938) 326333," *La Ricerca Scientifica* **16**(9), 326–333 (1938).
[9] Rosensteiner, M., "Wavefront reconstruction for extremely large telescopes via cure with domain decomposition," *JOSA A* **29**(11), 2328–2336 (2012).
[10] Censor, Y., Eggermont, P. P. B., and Gordon, D., "Strong underrelaxation in Kaczmarz's method for inconsistent systems," *Numerische Mathematik* **41**(1), 83–92 (1983).
[11] Natterer, F., [*The mathematics of computerized tomography*], vol. 32, Siam (1986).
[12] Censor, Y., Elfving, T., Herman, G. T., and Nikazad, T., "On diagonally relaxed orthogonal projection methods," *SIAM Journal on Scientific Computing* **30**(1), 473–504 (2008).