A new temporal control approach for SCAO systems

Markus Pöttninger, Ronny Ramlau and Günter Auzinger

1 Doctoral Program Computational Mathematics, Johannes Kepler University, Linz, Austria
2 Industrial Mathematics Institute, Johannes Kepler University, Linz, and Johann Radon Institute for Computational and Applied Mathematics, Linz, Austria
3 Industrial Mathematics Institute, Johannes Kepler University, Linz, Austria

E-mail: markus.poettinger@dk-compmath.jku.at, ronny.ramlau@jku.at and guenter.auzinger@mathconsult.co.at

Received 15 February 2019, revised 8 August 2019
Accepted for publication 16 September 2019
Published 3 December 2019

Abstract
Adaptive optics (AO) systems for ground-based telescopes use deformable mirrors to physically correct wavefront distortions induced by atmospheric turbulence. Due to time delays caused by different parts of the AO system, the process of turbulence correction becomes even more difficult since the earth’s atmosphere changes continuously. In this paper we propose a new temporal control approach for the computation of optimal mirror configurations based on the solution of a sequence of inverse problems for the wavefront sensor operator. Our mathematical formulation of the underlying problem allows the incorporation of computationally efficient wavefront reconstruction methods and a wavefront prediction step. Based on the frozen flow assumption, the prediction of a future wavefront relies on a suitable shift of the reconstructed wavefront. The performance of our temporal control algorithm is demonstrated in the context of a single conjugate adaptive optics system on a 37 meter telescope using a Shack–Hartmann wavefront sensor. Numerical results of the proposed control method are provided using OCTOPUS, the official end-to-end simulation tool of the European Southern Observatory.

Keywords: adaptive optics, inverse problems, temporal control, wavefront prediction, wind shift

(Some figures may appear in colour only in the online journal)
1. Introduction

In observing the sky with ground-based telescopes the atmosphere of the earth poses a major challenge. Through the interaction with the solar radiation and the wind, certain regions in the atmosphere are heated differently and get mixed up, leading to turbulent eddies [31], which introduce spatial and temporal variations in the air’s refractive index. As a result, the image taken by a ground-based telescope appears blurred in the scientific camera.

A remedy to counteract the turbulence-related image degradation is provided by a technology which is known as adaptive optics (AO). AO physically compensates the distorted incoming light by reflecting it on one or several deformable mirror(s) (DMs), which are steered according to measurements provided by wavefront sensors (WFSs). An essential part of the AO system is the control unit, which represents the link between the WFSs and the DMs. For a classical system (using only one WFS and one DM), the task of the AO control unit is basically concerned with the calculation of the new DM actuator commands from WFS measurements. However, due to unavoidable temporal delays in the AO correction loop, the control of the DM becomes even more complicated. In particular, the control algorithm should have a prediction ability in order to follow the rapid changes in the atmospheric turbulence. Different control strategies have been proposed, ranging from simple, low-computational integrator-based control methods to complex linear quadratic Gaussian (LQG) optimal control algorithms which utilize statistical priors on the atmosphere and the measurement noise.

In this paper, we present a new temporal control method. It relies on a more general approach, where the calculation of the DM actuator commands is done by first solving an inverse problem of reconstructing the incoming WF from noisy WFS data and then projecting the reconstruction onto the DM. The proposed method is based on the consideration of the whole AO correction process in terms of operator equations and allows to incorporate computationally efficient reconstruction methods. For the wavefront reconstruction, we particularly focus on the cumulative reconstructor with domain decomposition (CuReD) [32]. The CuReD method is a highly efficient reconstructor which does not require any statistical priors. Furthermore, the proposed approach allows to include an additional prediction step, which is carried out by a shift of the reconstructed wavefront. The required shift vector is calculated based on a sequence of previously reconstructed wavefronts. In contrast to the LQG-based control method, the proposed temporal control method is fully deterministic and does not need any additional statistical knowledge on the atmosphere nor the measurement noise. All informations used in this approach are purely obtained from the reconstructed WFs.

The reminder of this paper is organized as follows. In section 2 we give an introduction to atmospheric turbulence and discuss its effect on telescope imaging. Section 3 describes the principle of AO in more detail. In section 4 we give a mathematical model for the SCAO system and describe our temporal control approach. Section 5 presents the results obtained in simulation and in section 6 we give our conclusions.

2. Atmospheric turbulence and perturbation of light

From the physical point of view, light is described as an electromagnetic wave

\[ \mathcal{U}(r, t) = \mathcal{A}(r, t)e^{i\phi(r, t)} \]
with amplitude $A$ and phase $\phi$, depending on spatial coordinates $r \in \mathbb{R}^2$ and time $t \in [0, \infty)$.

In case of a wave with globally constant amplitude, a surface of equal phase is called wavefront (WF), travelling through the three-dimensional space. Before the light (stemming from a distant point source) enters the earth’s exosphere, it can be described as a planar WF. While the light is passing through several atmospheric regions with different temperature inhomogeneities, it gets perturbed by random variations in the air’s refractive index and fluctuations in the intensity, leading to a perturbed incoming WF $\varphi_{\text{tur}}$ on the telescope aperture $\Omega_4$.

$$\Omega_4 := \{ r = (r_1, r_2) \in \mathbb{R}^2 : a_{\text{int}} \leq |r| \leq a_{\text{out}} \}$$

with exterior radius $a_{\text{out}}$ and radius $a_{\text{int}}$ of the central obstruction (through which the collected light is led to the optical devices). While variations in the refractive index mainly affect the phase (in form of local phase shifts), intensity fluctuations (called scintillation) causes variations of the amplitude. However, due to the aperture averaging effect [14], the effect of scintillation is usually very low and thus is neglected, i.e. $A = 1$.

In astronomical observations, the image quality is degraded by the phenomenon of diffraction (depending on the size of the telescope aperture) and the effect of atmospheric turbulence. While the image quality is always limited by diffraction, the effect of the atmosphere can be reduced by compensating the shifts in the phase $\phi$.

An important characteristic of the turbulence is represented by the so-called refractive index structure coefficient $C_n^2$, which varies in dependence of the height and describes the turbulence distribution in the atmosphere [10, 31]. Based on the $C_n^2$ profile integrated over height, the Fried parameter $r_0$ [30] represents another important astronomical parameter which provides a measure of the optical strength of the turbulence. While the atmosphere extends over a height of about 300 000 m, the refractive index structure coefficient decreases for increasing altitude. Actually, the $C_n^2$ profile is non-zero only up to a certain height which is located at about $h_{\text{max}} = 30 000$ m [10]. Furthermore, the turbulence strength of the atmosphere is not distributed uniformly in height. In fact, the $C_n^2$ profile shows that most of the turbulence is located within separated layers at different altitudes [30]. Therefore, the atmosphere can be modeled by a finite number $n_L \in \mathbb{N}$ of thin and statistically independent turbulent layers located at heights $0 \leq h_1 < \cdots < h_{n_L} \leq h_{\text{max}}$, where each layer is described by an appropriate statistical atmospheric model (von Karman or Kolmogorov model, see e.g. [10]). While the light propagates through turbulent regions with different refractive indices, parts of the wavefront will be delayed compared to others, leading to local shifts in the WF. According to the layered structure of the atmosphere, the incoming WF gets aberrated by the turbulence contributions $\varphi_{\text{tur},l}$ of the single turbulent layers along the optical path in direction $\alpha \in \mathbb{R}^2$ (measured in radians from the zenith axis), i.e. the shape of the distorted incoming WF can be written as the sum [29]

$$\varphi_{\alpha}(r, t) = \sum_{l=1}^{n_L} \gamma_l \varphi_{\text{tur},l}(r + h_l \alpha, t),$$

where $\gamma_l$ denotes the normalized turbulence strength of the $l$th layer (according to the $C_n^2$ profile). Note that $\sum \gamma_l = 1$ [29]. The phase is related to the WF aberrations via $\phi = (2\pi/\lambda) \varphi_{\alpha}$, depending on the wavelength $\lambda$.

Similar to the motion pattern of clouds, the temporal fluctuations of atmospheric turbulence are mainly driven by two processes: fast deterministic changes due to wind shifts and

---

4 For the future generation of extremely large telescopes with aperture diameters up to 40 meters, the diameter of the central obstruction is about 30% of the outer diameter (www.eso.org). As telescope aperture one can also consider more general geometries than a circular disk.
slow random motions due to heat dissipation in the air, known as boiling [3]. A common approach to model the temporal evolution of the atmosphere is the Taylor frozen flow assumption [10, 30, 31], which basically states that, for short time periods (in the order of 10 ms), the temperature inhomogeneities are considered as frozen within the turbulence layers, and the only change of such a layer is due to the wind shift. In particular, boiling effects (which evolve on time scales greater than 50 ms [4]) are neglected. According to the frozen flow assumption, each turbulent layer propagates with its own wind speed and shift direction. Considering that the $l$th layer is moving with a wind velocity vector $\mathbf{v}_l$, then $\varphi^{\text{turb}}_{l}$ at time $t + \tau$ can be related to the same layer function at time $t$ with an additional shift $\tau \mathbf{v}_l$ in the spatial coordinate [31], i.e.

$$\varphi^{\text{turb}}_{l}(\mathbf{r}, t + \tau) = \varphi^{\text{turb}}_{l}(\mathbf{r} - \tau \mathbf{v}_l, t).$$

Typical values of the wind velocities range from 5 m s$^{-1}$ up to a peak velocity of 40 m s$^{-1}$. An important temporal parameter represents the Greenwood time delay [30], defined as

$$\tau_0 = 0.314 \frac{r_0}{|\mathbf{v}|}$$

with

$$|\mathbf{v}| = \left( \sum_{l=1}^{n_l} \gamma_l |\mathbf{v}_l|^{5/3} \right)^{3/5}$$

beeing the turbulence weighted average of the layers velocities. $\tau_0$, also known as turbulence coherence time, represents the characteristic time scale for changes in the WF aberrations. For time scales shorter than $\tau_0$, the perturbations in the WF and thus the turbulence in the atmosphere are considered to be approximately constant.

3. Adaptive optics

Essentially, an AO system consists of three main components. WFSs measure the incoming light of reference light sources, i.e. bright and distant astronomical objects. Usually, they correspond to stars (called natural guide stars, short NGSs) located in the near vicinity of the observed object. If NGSs are not available, or if they are too faint for measuring, artificially produced laser guide stars can be used. The second AO component are DMs, each conjugated to a certain altitude. The shapes of the DMs are calculated based on WFS measurement data and are adjusted via actuators. The link between the sensors and mirrors is given by the control unit, which represents the third AO system component. In dependence of the arrangement of the sensors and mirrors, an AO system can be driven in different modes. If the WFSs are located in front of the DMs, then one speaks of an open loop configuration. In contrast, the AO system operates in closed loop if the mirrors and sensors are placed the other way around. In the latter case, the residual wavefront (=

$\varphi^{\text{turb}}_{l}(\mathbf{r}, t + \tau) = \varphi^{\text{turb}}_{l}(\mathbf{r} - \tau \mathbf{v}_l, t).$
For a classical AO system like SCAO, the task of the AO control unit is basically concerned with the calculation of the new DM actuator commands from WFS measurements. Note that the WFS, which is also optically conjugated to the ground layer, measures the cumulation of the contributions of the turbulence layers in the direction of the optical axis (see (4)). This is also the only information available to control the DM. Based on the WFS sampling period, the whole AO correction process is considered in discrete time steps (see section 4.3). Let \( \mathbf{a}_i = (a_{i,1}, \ldots, a_{i,n_{\text{act}}}) \) be the reconstructed DM actuator commands with respect to the \( i \)th time instant, where \( n_{\text{act}} \in \mathbb{N} \) denotes the number of actuators. Then, the corresponding DM shape is given by

\[
\varphi_{i}^{\text{dm}}(\mathbf{r}) = \sum_{j=1}^{n_{\text{act}}} a_{i,j} h_{j}(\mathbf{r}),
\]

SCAO highly depends on the direction of the observation and only provides a good correction in an extremely limited area surrounding the guide star. Furthermore, by using SCAO, only small areas of the sky can be observed because there are not enough bright guide stars.

Figure 1. Graphical illustration of the wavefront perturbation. The atmospheric cylinder, through which the light of the observed object is passing, gets spanned by the telescope aperture \( \Omega_A \) and the shifted apertures \( \Omega_A(h_{\alpha, \eta}) := \{ \mathbf{r} \in \mathbb{R}^2 : \mathbf{r} - h_{\alpha, \eta} \in \Omega_A \} \).
where $h_j : \Omega_A \rightarrow \mathbb{R}$ denotes the DM influence function of the $j$th actuator. In general, the mirror shape extracted from currently available measurements does not correspond to an appropriate approximation of the incoming WF at a future time, since the atmosphere changes rapidly. Hence, any time delay in the control loop results in a temporal error, leading to a degraded image correction. Time delays are unavoidable and are mainly caused by the WFS read-out time, the time used for the calculation of the DM actuator commands and the adjustment of the DM. Another error source, that contributes to the overall temporal error, arises due to the finite bandwidth of the AO system. The WFS sampling rate represents the clocking period of the AO system and thus specifies the correction bandwidth. In order to provide accurate measurement data, the WFS has to collect a sufficient large amount of photons within the sampling period. This requires a sampling period of a certain length, which, on the negative side, causes a larger temporal error. Basically, the bandwidth-related temporal error can be reduced by shortening the sampling period. However, speeding up the frame rate will in general lead to inaccurate measurement data as the WFS can only collect a smaller number of photons. A sufficiently bright reference light source would provide a remedy. But, as the sky is barely covered with very bright stars, an appropriate NGS is not available in general. Due to the considerations above, control of the DM poses some challenges. In order to keep up with the rapidly changing turbulence, the DM has to be updated approximately every millisecond. For this reason, the control action has to be sufficiently simple to allow a fast computation of the new DM shape such that the delay-related error is small. Nevertheless, a control strategy should also be able to reduce the overall temporal error that limits the performance of the AO system. This requires a certain complexity of the control algorithm. In order to keep the computational costs low, many AO systems utilize simple control strategies in which the spatial and temporal dynamics of the AO control system are considered separately. Hence, the action of the control system can be split into a spatial control step and a temporal control step. At first, the DM actuator commands are calculated from WFS data. To this end, the so-called interaction matrix, which maps the actuator commands to the WFS measurements, gets measured. The DM commands are then computed by multiplying the (stabilized) pseudo-inverse of the poke matrix (called control matrix) onto the measurements [22]. It is noted that, in general, the poke matrix and thus also the control matrix varies in time. Therefore, the control matrix has to be updated every time step. In the second step, a temporal control method is applied to ensure a good disturbance rejection and temporal stability in the AO correction process. A standard temporal control method is the so-called integrator control. In this approach, the previous DM actuator commands are combined with the currently reconstructed ones to obtain the new commands (see e.g. [15, 22]). For stability reasons, the actuator commands are multiplied with damping factors before they get combined. For an AO system running in closed loop configuration, another approach known as pseudo open loop control (POLC) [28] has been introduced. Here, open loop like data are generated from originally given closed loop data. From these data, the new actuator commands are computed, which are then used to update the previous commands. In addition, the whole process is stabilized over time using e.g. integrator control (see e.g. [2]).

In order to reduce the overall temporal error, optimal control strategies have been introduced. In this approach, the AO system is represented by a state-space model which describes the dynamical behaviour of the system and its outputs. Such a state-space model consists of a state evolution equation and a measurement equation and includes (statistical) prior informations on the system’s dynamics and uncertainties. A very successful control technique is the linear quadratic Gaussian (LQG) controller [5, 23], which is based on a linear state-space model including priors that follows Gaussian statistics and aims at the minimization of a quadratic cost function corresponding to the residual phase variance. Under the assumption of
a DM with instantaneous response, i.e. neglecting dynamics of the DM, the action of the LQG controller consists of a two-step procedure [15, 33]. At first, the WF is reconstructed from noisy measurements by solving a stochastic minimum variance estimation problem. This is done in the framework of the Kalman filter (e.g. [15, 20, 25, 33]). The corresponding mirror actuator commands are then computed by projecting the reconstructed WF onto the DM space, which is done by solving an optimal deterministic control problem [15, 33].

The Kalman filter (KF) corresponds to a system of linear, finite-dimensional equations (including statistical priors) that evaluates the best state estimate in the least-squares sense including all measurements available up to the current time step. The discrete representation of the phase can either be done in a zonal approach (the phase is represented by a set of discrete values at given points of the telescope aperture, e.g. at the mirror actuator positions) or a modal approach (the phase is represented as an expansion with respect to some set of basis functions (modes), e.g. Zernike polynomials) [8, 9, 30, 33, 15]. The dynamics of the turbulence can be modeled by a simple first-order autoregressive model [15, 33] or a more sophisticated model that reflects the turbulence evolution according to the Taylor frozen flow assumption [26, 27, 36]. For the latter case, a near-Markov approach has been proposed [36], which allows to not keep all the prior history of measurement data in memory. In its original form, the KF involves the computation of the so-called Kalman gain, which has to be calculated for each time step and is obtained by the inversion of a generally dense matrix. However, in AO applications the asymptotic formulation of the KF is usually implemented (see e.g. [33, 35]), without any remarkable loss of optimality. In this case, the state transition matrix (representing the evolution of the state vector) and the measurement process matrix are considered as stationary during a part of the observation, which allows the use of the constant, asymptotic Kalman gain. As a consequence, the Kalman gain can be calculated off-line by solving a discrete algebraic Riccati equation (DARE) and only has to be updated if the atmosphere is changing. However, since the off-line computational costs are still proportional to the 6th power of the telescope diameter [8, 20], the calculation of the Kalman gain can take a long time and will be challenging for the future generation of extremely large telescopes.

Several approaches have been proposed to reduce the computational costs of solving the DARE. For example, in [26], a Fourier decomposition of the atmospheric turbulence is used. Under the frozen flow assumption, the Fourier modes can be assumed to be statistically independent, which reduces the computational burden. Another Fourier-based method has been proposed in [20]. Here, the Fourier transform (FT) is used to decompose an infinite-order system into an infinite set of low, finite-order ones. The FT is used as an off-line tool which accelerates the computation of the Kalman gain. The on-line computation is improved by sparsity properties in the zonal representation of the phase. In [5], the computational costs are reduced by replacing vector-matrix multiplications by spectral iterative methods that use a sparse approximation to the Riccati equation solution. Further considerations have been done, e.g. in [21]. Also improvements in the on-line computation of the Kalman gain have been made, see e.g. [8, 9].

In comparison to simple control strategies, LQG control has many advantages. As already stated, the KF provides the best estimate of the turbulence phase in the mean-square sense, since prior knowledge on the turbulence statistics and the measurement noise are taken into account in its formalism. Actually, it has been demonstrated that LQG control provides superior performance in terms of noise rejection and phase reconstruction in classical AO [5]. It has also been demonstrated that LQG can be used to compensate for non-turbulent perturbations like wind shake and vibrations [24, 35] or DM dynamics [6], and also allows to consider arbitrary, real-valued time delays in the AO loop [27]. A significant benefit of Kalman filter-based controllers rely on the ability of phase prediction. In [4, 12, 19], wind predictive control
methods have been proposed. According to the frozen flow assumption, the prediction relies on a suitable shift of the turbulence according to a wind shift profile. On the plus side, prediction-based optimal control efficiently reduces the temporal related WF error, allows longer integration times and weakens brightness requirements on the reference light sources [11, 12]. In particular, (wind) predictive control supports the optimal use of limited photons collected by the WFS from faint guide stars. However, optimal control also has its price. Beside the computational burdens, these methods require prior informations on the atmospheric statistics and the measurement noise and rely on an accurate dynamical model that reflects the temporal behaviour of the turbulence.

As described above, the reconstruction of the new mirror commands can be basically classified into two categories, where the computation is either done on only the previously obtained measurement data or on the whole sequence of obtained data up to the current time step [7]. In general, the performance of 1st category reconstructors is diminished at the benefit of the computational efficiency. For reconstructors of the 2nd category, the behaviour is exactly the opposite. In order to optimize the control action, any control method should be able to balance the performance and the computational complexity, i.e. the method should be somewhere in between this two categories. One way in this direction is to combine a fast and efficient reconstruction method with an explicit prediction step.

We present a fully deterministic temporal control method, which relies on the more general approach (similar to the two-step approach for LQG control, mentioned above), where the calculation of the DM actuator commands is split into two separate steps: at first, the incoming wavefront is reconstructed from noisy WFS measurement data $\mathbf{s}_i$. In this regard, we basically have to solve an inverse problem of the form $\Gamma \varphi_i = \mathbf{s}_i$ for $\varphi_i$, where $\Gamma$ corresponds to the mathematical model of the WFS. The reconstructed WF is then used to calculate the actuator commands for the DM. This is done via a deterministic projection step. The decoupling of the WF reconstruction process from the DM has two main advantages: on the one hand, by exploiting the mathematical model of the WFS, this approach allows to include matrix-free, model-based reconstruction methods. On the other hand, this approach also provides more flexibility regarding the geometry or the used influence functions of the DM. This approach also allows to include an additional prediction step.

4. Temporal control for SCAO systems

As mentioned above, an AO system does not have the ability to respond instantaneously to changes in the wavefront. As a result, any delay between the beginning of the measurement process right up to the DM adjustment causes a deteriorated image correction. Furthermore, also the finite bandwidth of the AO system contributes to the total temporal error budget. In order to counteract temporal related errors, we need to modify the reconstructed incoming wavefront such that the resulting mirror shape represents a good approximation of the unknown turbulent wavefront at a future time.

In the following we outline a temporal control law which is capable of reducing the temporal error of a SCAO system in closed loop configuration, sketched in figure 2.

4.1. AO performance criterion

In the sequel let $\varphi_{\text{dm}}$ denote the shape of the DM. In order to assess the performance of an AO system, the Strehl ratio, which is related to the residual wavefront $\varphi_{\text{tur}} - \varphi_{\text{dm}}$, provides the most important quality measure. The Strehl ratio will also be used for the AO performance
optimization. For a classical AO system like SCAO, the performance criterion is given by the minimization of the space-averaged square of the residual wavefront, averaged over a sufficiently large exposure time $T \in \mathbb{R}^+$ [15], i.e.

$$
\frac{1}{T} \int_{[0,T]} \frac{1}{|\Omega_A|} \int_{\Omega_A} (\phi_{\text{tur}}(r,t) - \phi_{\text{dm}}(r,t))^2 \, dr \, dt \to \min!
$$

According to this criterion, the turbulence correction aims at the determination of an appropriate mirror shape $\phi_{\text{dm}}$ such that $J(\phi_{\text{dm}}, \phi_{\text{tur}})$ gets minimized, where

$$
J(\phi, \phi_{\text{tur}}) := \int_{[0,T]} \int_{\Omega_A} (\phi_{\text{tur}}(r,t) - \phi(r,t))^2 \, dr \, dt .
$$

Since $1/T$ and $1/|\Omega_A|$ only represent scaling factors we may neglect them. As indicated above, the functional $J$ also depends on the unknown incoming wavefront. Additionally required information on $\phi_{\text{tur}}$ are obtained from measurements provided by the WFS.

4.2. Function spaces for wavefronts and mirror shapes

Note that $\Omega_A$ actually defines a compact set in $\mathbb{R}^2$. However, for the theoretical considerations we will focus on an open and bounded set $\Omega_A = \{ r \in \mathbb{R}^2 : a_{\text{in}} - \epsilon < |r| < a_{\text{out}} + \epsilon \}$ with $\epsilon > 0$ sufficiently small. In the following let $H^1(\Omega_A)$ be the Hilbert space

$$
H^1(\Omega_A) = \left\{ \phi \in L^2(\Omega_A) : \nabla r \phi \in [L^2(\Omega_A)]^2 \right\}
$$

equipped with the inner product

$$
\langle \phi, \psi \rangle_{H^1(\Omega_A)} = \int_{\Omega_A} \phi(r) \psi(r) + \nabla r \phi(r) \cdot \nabla r \psi(r) \, dr , \quad \phi, \psi \in H^1(\Omega_A) ,
$$

and the induced norm $\| \phi \|_{H^1(\Omega_A)} = \sqrt{\langle \phi, \phi \rangle_{H^1(\Omega_A)}}$. Here, $\nabla r \phi$ corresponds to the weak formulation of the gradient, satisfying the condition

$$
\int_{\Omega_A} \nabla r \phi(r) \cdot \phi(r) \, dr = - \int_{\Omega_A} \phi(r) \text{div}_r \phi(r) \, dr
$$

Figure 2. Diagram of an AO system in closed loop configuration. The residual WF is split up into two paths using a beam splitter (BS). While one path is led to the WFS the other one leads to the scientific camera (SC). Based on noisy measurement data, the control unit computes the mirror commands, which are used to steer the DM. The introduced correction WF (partially) compensates for the perturbations in the incoming WF.
for all $\varrho \in [C^\infty_0(\Omega_A)]^2$, where $C^\infty_0(\Omega_A)$ defines the space of all infinitely differentiable functions with compact support in $\Omega_A$.

Since any function $\varphi(\mathbf{r}, t)$ can also be interpreted as a family of functions $\varphi(\mathbf{r})(t)$ parametrized in time, the definition of Bochner-measurable functions will be of great importance for our mathematical modeling. A function $\sigma : [0, T) \to H^1(\Omega_A)$ is said to be simple if there exists a finite number of pairwise disjoint Lebesgue-measurable sets $T_j \subset [0, T)$ and functions $\sigma_j$ in $H^1(\Omega_A)$ such that

$$\sigma(t) = \sum_j \chi_{T_j}(t) \sigma_j.$$  \hspace{1cm} (9)

Here, $\chi_{T_j}$ correspond to the index function, which is 1 if $t \in T_j$ and zero otherwise. Furthermore, a function $\varphi : [0, T) \to H^1(\Omega_A)$ is said to be Bochner-measurable if there exists a sequence of simple functions $(\sigma_k)_{k \geq 1}$ such that $\lim_{k \to \infty} \|\sigma_k(t) - \varphi(t)\|_{H^1(\Omega_A)} = 0$ for almost every $t \in [0, T)$. For our purposes, we consider the linear Bochner–Lebesgue space

$$L^2([0, T), H^1(\Omega_A)) = \left\{ \varphi \mid \varphi \text{ Bochner–measureable, } \int_{[0, T)} \|\varphi(t)\|_{H^1(\Omega_A)}^2 \, dt < \infty \right\}$$  \hspace{1cm} (10)

containing all Bochner-measurable functions $\varphi : [0, T) \to H^1(\Omega_A)$ such that $\|\varphi(\cdot)\|_{H^1(\Omega_A)} : [0, T) \to \mathbb{R}$ is square integrable. It is noted that $L^2([0, T), H^1(\Omega_A))$ is also a Hilbert space with inner product given by (7) integrated over time.

For the rest of this paper we consider $\varphi^{\text{tur}}, \varphi^{\text{dm}} \in L^2([0, T), H^1(\Omega_A))$.

### 4.3. Temporal modeling

While the turbulent wavefront $\varphi^{\text{tur}}$ evolves over time, it gets measured by the WFS within fixed sampling periods. For the sake of simplicity we consider equidistant sampling intervals of length $\Delta_T \in \mathbb{R}^+$. More precisely, $\Delta_T$ corresponds to the length of the period between the beginning of a measurement phase and the start of the next measurement process. The measurement read-out of the first sampling period takes place in the subsequent sampling phase (see figure 3).

Strictly speaking, a WFS is not able to collect data continuously: due to periods of sensor resetting, so-called dead times arise. Within these phases no measurement data are obtained. However, for all commonly used sensors, the actual dead times are close to zero and they are usually not considered in the modeling process.

In the following we use these sampling periods to discretize the control functional (6) in time. Introducing the temporal nodes

$$0 = t_0 < t_1 < \cdots < t_{n_F-1} < t_{n_F} = T, \quad n_F \in \mathbb{N},$$

with $t_i := i \Delta_T$, and letting $F_i := [t_{i-1}, t_i)$ define the $i$th frame spanned by two consecutive nodes, we obtain a decomposition

![Figure 3. Sampling periods and read-out phases.](image)
of the exposure time $T$ into a number of $n_F$ time frames $F_i$, each of length $|F_i| = \Delta_T$. Since $\Delta_T$ is usually very small (at about 0.0003 to 0.002 s), the incoming turbulent wavefront $\varphi_{\text{tur}}$ restricted to $F_i$ can be approximately described by its mean value, i.e.

$$\varphi_{\text{tur}}^i(r) := \frac{1}{\Delta_T} \int_{F_i} \varphi_{\text{tur}}(r, t) \, dt, \quad 1 \leq i \leq n_F. \tag{11}$$

Since $F_i \subset [0, T)$, it trivially holds that $\varphi_{\text{tur}}$ restricted to $F_i$ is contained in $L^2(F_i, H^1(\Omega_A))$. Together with the theorem of Bochner [34], it immediately follows that $\varphi_{\text{tur}}^i \in H^1(\Omega_A)$. Due to representation (4), we further obtain

$$\varphi_{\text{tur}}^i(r) = \sum_{l=1}^{n_L} \gamma_l \varphi_{\text{tur}, l}^i(r) \tag{12}$$

with $\varphi_{\text{tur}, l}^i \in H^1(\Omega_A)$ (according to the linearity of the Bochner–Lebesgue space (10)).

As already mentioned above, the DM gets adjusted according to WFS measurements. Due to dynamics caused by the readjustment of the DM, the shape of the mirror will, in general, not be constant over a whole time frame. However, in practice, the DM shaping process requires much less time compared to the sampling period $\Delta_T$, which itself is very short. Thus, for the sake of simplicity, the temporal dynamics of the DM are usually neglected. Defining $\varphi_{\text{dm}}^i$ as the mirror shape corresponding to the $i$th frame $F_i$, we may write

$$\varphi_{\text{dm}}^i(r, t) = \sum_{j=1}^{n_F} \chi_{F_j}(t) \varphi_{\text{dm}, j}^i(r). \tag{13}$$

Since (13) corresponds to a simple function (see (9)), it immediately follows from the definition of the Bochner-Lebesgue space (10) that $\varphi_{\text{dm}}^i$ is in $L^2([0, T), H^1(\Omega_A))$ if and only if $\varphi_{\text{dm}}^i \in H^1(\Omega_A)$ for all $1 \leq i \leq n_F$.

For the choice $\tau = j\Delta_T$ with $j \in \mathbb{N}_0$, the frozen flow equation (1) averaged over $F_i$ is given by

$$\frac{1}{\Delta_T} \int_{F_i} \varphi_{\text{tur}, l}^i(r, t + j\Delta_T) \, dt = \frac{1}{\Delta_T} \int_{F_i} \varphi_{\text{tur}, l}^i(r - j\Delta_T \mathbf{v}_l, t) \, dt.$$

Since

$$\frac{1}{\Delta_T} \int_{F_i} \varphi_{\text{tur}, l}^i(r, t + j\Delta_T) \, dt = \frac{1}{\Delta_T} \int_{F_{i+j}} \varphi_{\text{tur}, l}^j(r, t) \, dt,$$

we obtain the discrete-time frozen flow equation:

$$\varphi_{\text{tur}, l}^{i+j}(r) = \varphi_{\text{tur}, l}^i(r - j\Delta_T \mathbf{v}_l). \tag{14}$$

### 4.4. Measurement data process

In the following let $\varphi_{\text{meas}}^i \in L^2([0, T), H^1(\Omega_A))$ denote the wavefront which gets sampled by the WFS. Independently of the type of the used sensor, only discrete information on $\varphi_{\text{meas}}^i$ is provided. Let $\{\Omega_s\}_{s=1}^{n_S}$ define a decomposition of the telescope aperture into $n_S \in \mathbb{N}$ disjoint and equidistant subapertures. In our case we consider open and square-shaped
subsets in $\mathbb{R}^2$, satisfying the condition $\Omega_S := \bigcup_{s=1}^{n_S} \Omega_S \subseteq \Omega_A$, where $\Omega_t$ and $\Omega_A$ denote the closure of the corresponding sets (see figure 4). For each subaperture we define $r^{(s,c)} := (r_1^{(s,c)}, r_2^{(s,c)})$, $c \in \{1, 2, 3, 4\}$, as the coordinates of its corner nodes and $g \in \mathbb{R}^+$ as the edge length. Furthermore, let $n_N \in \mathbb{N}$ denote the total number of nodes given by the sub-aperture grid $\Omega_S$. Then, for each $n \in \{1, \ldots, n_N\}$, we define $r^{(n)} := (r_1^{(n)}, r_2^{(n)})$ as a grid node, which is connected to a corner node via $s : c \longleftrightarrow n = n(s,c)$.

In case of a Shack–Hartmann (SH) wavefront sensor, only discrete information about the derivative of $\varphi_{\text{meas}}$ gets measured (see [29]), i.e.

$$s(t) = \left(\frac{1}{|\Omega_s|} \int_{\Omega_s} \nabla r_{r_{\varphi_{\text{meas}}}}(r,t) \, dr\right)_{s=1}^{n_S}.$$ (15)

In practice, the measurement data provided by the WFS include information on $\varphi_{\text{meas}}$ within a period of time (sampling period). Therefore, analogously to (11), we consider the mean value of $s$ with respect to one time frame, i.e.

$$s_i := \frac{1}{\Delta T} \int_{F_i} s(t) \, dt = \left(\frac{1}{\Delta T |\Omega_s|} \int_{F_i} \int_{\Omega_s} \nabla r_{r_{\varphi_{\text{meas}}}}(r,t) \, dr \, dt\right)_{s=1}^{n_S}.$$ (16)

Based on the theorem of Fubini–Tonelli [1] (in combination with Bochner’s theorem [34] and the definition of the weak gradient (8)), the spatial derivative and the integral with respect to time can be interchanged, i.e.

$$\frac{1}{\Delta T |\Omega_s|} \int_{F_i} \int_{\Omega_s} \nabla r_{r_{\varphi_{\text{meas}}}}(r,t) \, dr \, dt = \frac{1}{|\Omega_s|} \int_{\Omega_s} \left(\frac{1}{\Delta T} \int_{F_i} \nabla r_{r_{\varphi_{\text{meas}}}}(r,t) \, dt\right) \, dr.$$ (17)
Introducing the operator
\[ \Gamma : H^1(\Omega_A) \rightarrow \mathbb{R}^{n_S \times 2}, \]
\[ \phi \mapsto \Gamma \phi := \left( \frac{1}{|\Omega_s|} \int_{\Omega_s} \nabla r r r \phi(r) \, dr \right)_{s=1}^{n_S}, \]
which corresponds to the Shack–Hartmann wavefront sensor, the discrete-time measurement equation (derived from (15)) is then given by the formula
\[ s_i = \Gamma \varphi_{\text{meas}}^i, \quad 1 \leq i \leq n_F, \]
with \( \varphi_{\text{meas}}^i \) analogously defined as (11). In case of an AO system driven in closed loop configuration we have \( \varphi_{\text{meas}}^i = \varphi_{\text{tur}}^i - \varphi_{\text{dm}}^i \). Due to the data read-out time, measurements are usually not obtained at the end of the sampling period, i.e. at time \( t_i \). Actually, \( s_i \) is available at a later time point
\[ t_i + d_1 \Delta T \quad (16) \]
with \( d_1 > 0 \).

4.5. Calculation of the mirror shapes considering time delays

Concerning the calculation of the new DM shape, we first consider that, for \( \varphi(r, t) = \sum_{i=1}^{n_F} \chi_{F_i}(t) \varphi_i(r) \), the functional \( J(\varphi, \varphi_{\text{tur}}) \) can be rewritten as
\[ J(\varphi, \varphi_{\text{tur}}) = J(\varphi_{\text{tur}}, \varphi_{\text{tur}}) + \Delta T \sum_{i=1}^{n_F} \| \varphi_{\text{tur}}^i - \varphi_{\text{tur}} \|_{L^2(\Omega_A)}^2, \quad (17) \]
where
\[ \varphi_{\text{tur}}(r, t) := \sum_{i=1}^{n_F} \chi_{F_i}(t) \varphi_{\text{tur}}^i(r) \quad (18) \]
and \( \| \cdot \|_{L^2(\Omega_A)} \) denotes the \( L^2 \)-norm on the telescope aperture. The proof of (17) is based on the considerations in [15]. According to the temporal discretization (11) and the definition (18), \( \varphi_{\text{tur}} \) is piecewise constant and thus defines only an approximation of the turbulent WF \( \varphi_{\text{tur}} \) (which evolves over time). Hence, the first term in (17) does not vanish in general, leading to a loss of optimality in the DM shape calculation. However, since \( J(\varphi_{\text{tur}}, \varphi_{\text{tur}}) \) does not depend on \( \varphi \), the original AO performance criterion can be considered in discrete time without any additional loss of optimality.

4.5.1. Wavefront reconstruction. In order to keep \( J(\varphi, \varphi_{\text{tur}}) \) small, it is sufficient to determine for all \( 1 \leq i \leq n_F \) an appropriate approximation of the turbulent wavefront \( \varphi_{\text{tur}}^i \). Hence, the turbulence correction aims at the calculation of \( \varphi_{\text{dm}}^i, i \in \{1, \ldots, n_F\} \), that optimally fits the unknown turbulent wavefront \( \varphi_{\text{tur}}^i \). In order to extract information on \( \varphi_{\text{tur}}^i \) from WFS measurements, we basically have to solve an inverse problem of the form
\[ \Gamma \varphi_{\text{meas}}^i = s_i, \]
for all \( 1 \leq i \leq n_F \), where \( \Gamma \) corresponds to the mathematical model of the underlying WFS. However, due to imperfections in the sensing process, a WFS only provides noisy data. For our considerations we assume additive noise described by the random process \( \eta_i \in \mathbb{R}^{n_S \times 2} \). Hence, instead of the exact measurement data \( s_i \), only an approximation \( s_i^\eta = s_i + \eta_i \) is known. Note
that, in closed loop configuration, any dynamics of the DM causes distorted measurements. However, since mirror dynamics are neglected in the modeling process, these disturbances are implicitly considered in the measurement process noise $\eta_i$. Since the problem of WF reconstruction is in general ill-posed (see [13]), any reconstruction method, described as operator $R$, has to meet the following requirements:

- $R$ has to be stable with respect to noise in the measurements, i.e. small perturbations in the data should not cause arbitrarily large changes in the solution $R s^\eta$.
- $R$ has to be an appropriate approximation of the (generalized) inverse $\Gamma^\dagger$ of the WFS operator, i.e. $R \Gamma \varphi \approx \varphi$ for all $\varphi$.

A reliable reconstruction algorithm for Shack–Hartmann sensors is the **cumulative reconstructor with domain decomposition** (CuReD, [32]), which will be the reconstruction method of choice in this paper. Basically, the CuReD method provides discrete information on the measured wavefront with respect to the corner nodes $r^{(n)}$ of the subaperture grid. Introducing the reconstruction operator

$$R : \mathbb{R}^{n_s \times 2} \rightarrow \mathbb{R}^{n_y},$$

$$s^\eta_j \mapsto Rs^\eta_j : \varphi^\mathrm{rec},$$

the discrete reconstruction $\tilde{\varphi}^\mathrm{tur}$ of the turbulent wavefront $\varphi^\mathrm{tur}$ is then given by

$$\tilde{\varphi}^\mathrm{tur}_i = \gamma^\mathrm{gain} \varphi^\mathrm{rec}_i + P \varphi^\mathrm{dm}_i, \quad 1 \leq i \leq n_F,$$

where $P$ defines the collocation operator which evaluates the continuous mirror shape on the nodes $r^{(n)}$ of the subaperture grid and $\gamma^\mathrm{gain} \in (0, 1]$ defines the so-called loop gain, which is used for the stability control of the AO correction process.

The parameter $\gamma^\mathrm{gain}$ is found empirically and mainly depends on the reconstruction method and the accuracy of the measurement data. If we assume that the measurements are not contaminated by noise, $R$ perfectly reconstructs the measured WF and the DM response has no effect on the data, then

$$\tilde{\varphi}^\mathrm{rec}_i = Rs^\eta_i = R s_i = R (\tilde{\varphi}^\mathrm{tur}_i - \varphi^\mathrm{dm}_i) = \varphi^\mathrm{tur}_i - \varphi^\mathrm{dm}_i.$$

In this case, (20) combined with $\gamma^\mathrm{gain} = 1$ would provide an exact reconstruction of the incoming WF, i.e. $\tilde{\varphi}^\mathrm{tur} = \varphi^\mathrm{tur}$. However, in real world applications, none of these assumptions will be fulfilled. Hence, the reconstruction process needs to be stabilized, which is controlled by a parameter $\gamma^\mathrm{gain} < 1$. In particular, stabilization is necessary in order to damp error propagations caused by uncertainties in the reconstruction. If not stabilized, errors in the reconstruction gets transferred to the DM shape which further affects the measurement data.

Note that, due to the time used for the reconstruction, $\tilde{\varphi}^\mathrm{tur}_i$ is available at time $t_i + (d_1 + d_2) \Delta_T$ with $d_1$ given in (16) and $d_2 \in \mathbb{R}^+$.  

**4.5.2. Determination of the new mirror shape.** Based on the sampling period, the WFS read-out time and the reconstruction process, we have to consider a time delay of length $(1 + d_1 + d_2) \Delta_T$. However, due to the time $d_2 \Delta_T$, $d_1 \in \mathbb{R}^+$, required for the calculation and adjustment of the new mirror shape, we have to expect a total delay of length $d \Delta_T$, where $d := 1 + d_1 + d_2 + d_3$ (see figure 5). For the sake of simplicity we assume that $d \in \mathbb{N}$, i.e. the delay extends over a fixed number of frames, and is considered as constant.

In this paper we use a simplified DM model, where the actuators are placed on the corner nodes of the subapertures (according to the Fried geometry), and the shape is determined via (bilinear) interpolation between the actuators. According to (5), the actuator commands $a_j$ are simply given by the values of the WF at the corresponding subaperture corner nodes and the
DM space gets spanned by continuous shape functions $h_j$. To a certain approximation, such a DM model is sufficient, especially for an AO system running in closed loop configuration [38]. Furthermore, this choice enables to consider the continuous function $\hat{\phi}_{i}^{\text{tur}}$, which is generated via interpolation between the values $\hat{\phi}_{i}^{\text{tur}}$ at the subaperture corner nodes, directly as DM shape, i.e. $\phi_{i}^{\text{dm}} = \hat{\phi}_{i}^{\text{tur}}$.

Basically, the reconstructor allows to extract information on $\phi_{i}^{\text{tur}}$ from given data $s_i^j$. However, due to the temporal delays we actually require an estimate of the unknown turbulent WF $\phi_{i}^{\text{tur}}$ based on the same measurement set. Thus, in general, $\phi_{i}^{\text{tur}}$ does not represent $\phi_{i}^{\text{dm}}$ may not appropriately compensate for WF distortions at time frame $i + d$. Consequently, the reconstructed wavefront has to be further modified. As a remedy, AO correction may be improved by including additional knowledge of the temporal evolution of the incoming wavefronts to obtain a prediction $\bar{\phi}_{i+d}^{\text{tur}}$ of the turbulent wavefront $\phi_{i+d}^{\text{tur}}$ based on the reconstruction $\phi_{i}^{\text{tur}}$. Once the predicted WF is available, the new mirror shape (corresponding to the $(i + d)$th frame) is then given by $\phi_{i+d}^{\text{dm}} = \bar{\phi}_{i+d}^{\text{tur}}$. Due to the considerations above, it immediately follows that $\phi_{i}^{\text{dm}} = \cdots = \phi_{d}^{\text{dm}}$, where $\phi_{1}^{\text{dm}}$ defines the initial mirror shape. Since, in general, no information on $\phi_{1}^{\text{dm}}$ is available, we set $\phi_{1}^{\text{dm}} = 0$.

4.6. Wavefront prediction

In general, the incoming wavefront is affected by several turbulent layers, which, according to the frozen flow assumption, are blown across the telescope aperture by the wind in different horizontal directions with different speeds. However, since the turbulence strength varies in dependence of the altitude according to the $C_n^2$ profile, some of these layers have a stronger contribution to the incoming WF than the others. From (12) together with the discrete-time frozen flow equation (given in the alternative form $\phi_{i}^{\text{tur}}(r) = \phi_{i}^{\text{tur}}(r + j\Delta r \eta_{\ell})$), we have

$$\phi_{i}^{\text{tur}}(r) = \sum_{l=1}^{n_L} \gamma_l \phi_{i+l}^{\text{tur}}(r + j\Delta r \eta_{l}). \quad (21)$$

Assuming one strong turbulent layer $\phi_{i}^{\text{tur},\ell}$ with $\ell \in \{1, \ldots, n_L\}$ dominating all the others, i.e. $\gamma_{\ell} > \gamma_l$ for all $l \in \{1, \ldots, n_L\} \setminus \{\ell\}$, the incoming WF is mainly affected by this layer. In particular, we have that $\phi_{i}^{\text{tur}}$ contributes the most to $\phi_{i}^{\text{tur}}$ and $\phi_{i}^{\text{tur}}$ is dominated by $\phi_{i}^{\text{tur},\ell}$. Hence, according to the frozen flow assumption, the temporal evolution of the incoming wavefront is approximately determined by the temporal evolution of the dominant turbulence layer. Consequently, we may assume that

$$\phi_{i+d}^{\text{tur}}(r + j\Delta r \eta_{\ell}) \approx \sum_{i=1}^{n_L} \gamma_i \phi_{i+l}^{\text{tur}}(r + j\Delta r \eta_{l}), \quad (22)$$

with turbulence weighted average propagation vector $\eta_{\ell}$, which is mainly dictated by the shift vector $\eta_{\ell}$ of the strongest layer. From (21) and (22), we finally obtain

$$\phi_{i+d}^{\text{tur}}(r) \approx \phi_{i}^{\text{tur}}(r - j\Delta r \eta_{\ell}), \quad (23)$$

i.e. the temporal evolution of the incoming wavefront can be approximately described by a shift with a turbulence weighted propagation vector. Since, typically, the ground layer contributes the most to the incoming wavefront, a shift-based predictive control approach for a SCAO system is also possible in presence of multiple turbulent layers.

In order to counteract temporal delays, the prediction of a future wavefront is essential. As described above, a shifted version of $\phi_{i}^{\text{tur}}$ can be taken as a prediction of $\phi_{i+d}^{\text{tur}}$. However, this approach involves two main difficulties: on the one hand, the shift requires the identification
of the actual windshift vector, and on the other hand we need information on the wavefront outside of the telescope aperture \( \Omega_d \). In the following we outline a method which allows to compute an approximation of \( \varphi^\text{int}_{r_d} \) based on the discrete reconstruction \( \varphi^\text{int}_{r_d} \). By using interpolation, \( \varphi^\text{int}_{r_d} \) can be extended to a continuous function on \( \Omega_d \). However, in order to obtain a function over the whole telescope aperture or a shifted version of \( \Omega_d \), we additionally need information outside of \( \Omega_d \). Therefore, we use extrapolation of the reconstructed data.

In the following, let \( \mathbf{v}_i \in \mathbb{R}^2 \) denote the windshift vector corresponding to the \( i \)th frame, which, at the moment, is assumed to be known. In the next subsection we will deal with the extraction of \( \mathbf{v}_i \). The domain of the predicted wavefront is then given by the shifted telescope aperture

\[
\Omega^+_d (d \Delta \mathbf{v}) := \{ \mathbf{r} \in \mathbb{R}^2 : \mathbf{r} + d \Delta \mathbf{v} \in \Omega_d \}.
\]

Now, let \( \Omega^+_d := \bigcup_{i=1}^{n_s} \Omega_s \), \( \Omega_s \) define the smallest extension of the subaperture grid \( \Omega_s \) with \( n_s \in \mathbb{N} \) additional subapertures such that \( \Omega^+_d (d \Delta \mathbf{v}_i) \subseteq \Omega^+_d \) (see figure 6). Furthermore, let \( n_{\lambda j} \in \mathbb{N} \) define the number of grid nodes additionally gained by the extension. In order to extend the reconstructed data we use extrapolation. For any \( n \in \{ n_N + 1, \ldots, n_N + n_{\lambda j} \} \), let \( B_{\rho_n} (r^{(n)}) := \{ \mathbf{r} \in \mathbb{R}^2 : |\mathbf{r} - r^{(n)}| \leq \rho_n \} \) define the circle with radius \( \rho_n > 0 \), surrounding the corner node \( r^{(n)} \) of a newly added subaperture. Based on that, we further define

\[
N_n := \{ \tilde{n} : r^{(\tilde{n})} \in B_{\rho_n} (r^{(n)}) \} \cap \Omega_d.
\]

Then, for given weights \( \omega_{\tilde{n}} \in \mathbb{R} \) satisfying \( \sum_{\tilde{n} \in N_n} \omega_{\tilde{n}} \leq 1 \), we introduce the extrapolation operator

\[
E_i : \mathbb{R}^{n_N} \to \mathbb{R}^{n_{\lambda j} + n_{\lambda j}},
\]

\[
\phi \mapsto E_i(\phi)(n) := \sum_{\tilde{n} \in N_n} \omega_{\tilde{n}} \phi(\tilde{n}), \quad \text{for } n \in \{ n_N + 1, \ldots, n_N + n_{\lambda j} \},
\]

which enables to extend the reconstructed data onto \( \Omega^+_d \). Now, let \( Q_s \) define the linear space of all polynomials of degree \( \leq \kappa \) in each of the coordinateaxes \( r^{(1)}, r^{(2)} \). Furthermore, let

\[
V_i(\Omega^+_d) := \{ \psi \in C^{(s)}(\Omega^+_d) : \psi \big|_{\Omega_s} \in Q_s \text{ for all } s \in \{ 1, \ldots, n_s + n_{s j} \} \}
\]

define the space of all \( \nu \)-times differentiable, piecewise polynomial functions on the extension \( \Omega^+_d \). In order to extend \( \psi = E_i(\phi) \) to a continuous function, we define the operator

\[
B_i : \mathbb{R}^{n_{\lambda j} + n_{\lambda j}} \to V_i(\Omega^+_d),
\]

\[
\psi \mapsto B_i(\psi)(\mathbf{r}) := \frac{\sum_{i=1}^{n_{\lambda j} + n_{\lambda j}} \chi_{\Omega_i}(\mathbf{r}) \beta_i(\psi; \mathbf{r})}{\sum_{i=1}^{n_{\lambda j} + n_{\lambda j}} \chi_{\Omega_i}(\mathbf{r})}, \quad \mathbf{r} \in \Omega^+_d,
\]
where $\beta_s(r) := \sum_{c=1}^{4}\psi(n(s, c))\beta^{(s,c)}(r)$, which generates a function in $V_i(\bar{\Omega}^+)$ by means of shape functions $\beta^{(s,c)} \in Q$, which interpolates the function values between the discrete values $\psi(n(s, c))$ obtained at the corner nodes of a subaperture. Thereby, the shape functions satisfy the conditions $\beta^{(s,c)}(r^{c'}) = \delta_{c,c'}$, where $\delta$ denotes the Kronecker symbol. Based on (27), we finally introduce the operator $S_{d \Delta T}^{v_i}: R_{n N}^+ \rightarrow VA_i$, where $\psi \mapsto \hat{\psi}(r) =$ $S_{d \Delta T}^{v_i}(\psi)(r) := B_i(\psi)(r - d \Delta T v_i), \ r \in \bar{\Omega}_A,$ (28)

where $VA_i := \{ \phi : \phi = \psi(\bar{\Omega}^+_i(d \Delta T v_i))$, where $\psi \in V_i(\bar{\Omega}^+_i) \}$. Note that any function $\phi \in VA_i$ has support $\bar{\Omega}_A$. According to (28), the operator extends $\psi \in R_{n N}$ to a continuous function $\hat{\psi} = S_{d \Delta T}^{v_i}(\psi)$ which corresponds to a shifted version of the continuous extension of $\phi \in R_{n N}$. As one can see from the definition, any function $\phi$ in $VA_i$ inherits the properties of the functions in $V_i(\bar{\Omega}^+_i)$, i.e. $\phi$ is $\nu$-times continuously differentiable and piecewise polynomial. As an immediate result, it follows that $VA_i \subset H^1(\bar{\Omega}_A)$.

Once the windshift vector $v_i$ is available, an approximation of the turbulent WF $\varphi_{i+d}^{\text{tur}}$ (and thus also $\varphi_{i+d}^{\text{dm}}$) can be determined (according to (25) and (28), see figure 7) by

$$\hat{\varphi}_{i+d}^{\text{tur}}(r) = (S_{d \Delta T}^{v_i} \circ E_i)[\hat{\varphi}_{i+d}^{\text{tur}}](r), \ r \in \bar{\Omega}_A, \ 1 \leq i \leq n_F - d.$$ (29)

4.7 Calculation of the propagation vector

A method to obtain information on the windshift profile of the atmospheric layers using multiple Shack-Hartmann wavefront sensors is the so-called slope detection and ranging method.
short SLODAR [37]. In this method, phase aberrations of the wavefront are used to estimate the turbulence profile. More precisely, SLODAR uses wavefront slopes of two or more guide stars and their cross correlation to extract informations on the height, strength and velocity of the turbulence layers. However, in case of SCAO, only one guide star and one WFS is used, i.e. SLODAR can not be used. Furthermore, SCAO only provides information on the sum of the turbulence layer contributions. Therefore, a shift vector, which describes the propagation of the cumulated turbulence contributions, is required. Following the method proposed in [4, 12], the problem of propagation vector identification is related to a cost function

\[
J_1(v) = \sum_{r \in \Omega_{\text{S}} \cap \Omega_{\text{SA}}} \left( \tilde{\varphi}_i^\text{tur}(r) - \tilde{\varphi}_{i-1}^\text{tur}(r - \Delta_T v) \right)^2,
\] (30)

which has to be minimized. Here, the shift vector is calculated based on a reduced telescope aperture \( \Omega_{\text{SA}} \). Beside the reduction of the computational costs, there is another important reason for computing the propagation vector on a reduced set. As we have seen in the wavefront prediction above, we have to extend the data outside of the telescope aperture to apply a shift on it. However, if we focus on a reduced telescope aperture we are able to circumvent this problem (see figure 8). As the turbulent flow is assumed to be constant over a certain number of time frames, it is also possible to calculate the propagation vector based on a sequence of previously reconstructed wavefronts, leading to more sophisticated cost functions, e.g. [19]

\[
J_2(v) = \sum_{r \in \Omega_{\text{S}} \cap \Omega_{\text{SA}}} \sum_{j=0}^{q-1} \left( \tilde{\varphi}_{i-j}^\text{tur}(r) - \tilde{\varphi}_{i-m-j}^\text{tur}(r - m \Delta_T v) \right)^2,
\] (31)

with \( q, m \in \mathbb{N} \). In order to make the translation effect of small shifts more visible, it is recommended to use \( m > 1 \) [19]. In comparison to (30), this approach will in general increase the approximation quality of the reconstructed vector as more informations are used in the calculation. As a drawback, it also increases the computational costs to minimize (31). In [4, 12, 19], the methods above are proposed to calculate the wind velocity vectors of single turbulent layers. However, as discussed above (see (23)), these methods can also be used to determine the propagation vector in case of a SCAO system with multi-layered atmosphere.

Basically, the cost functions (30) and (31) can be considered independently of the way how the reconstructed wavefronts are obtained, i.e. it does not matter if the approximations \( \tilde{\varphi}_i^\text{tur} \) are calculated based on a deterministic or a Bayesian reconstruction approach. However, in case of a deterministic approach, one has to keep in mind that the reconstruction method...
ony minimizes effects of the measurement noise. In particular, the noise is not filtered out. To achieve a smoothing effect, it is recommended to average the reconstructions. Due to that, we consider an alternative approach in which we first compute the mean values of two bundles of reconstructed wavefronts and then use them to determine the propagation vector. Furthermore, the calculation is directly done on the discrete data provided by the reconstructor. For $i \geq u + v$ with $u, v \in \mathbb{N}$, let

$$
\tilde{\Phi}_{i,u,v} = \frac{1}{v} \sum_{j=0}^{v-1} \tilde{\phi}_{i-u-j}^{\text{tur}}
$$

(32)
define the mean change of the reconstructed wavefronts $\tilde{\phi}_{i-u}^{\text{tur}}, \tilde{\phi}_{i-u-1}^{\text{tur}}, \ldots, \tilde{\phi}_{i-(u+v)+1}^{\text{tur}}$ over $v$ time frames. Furthermore, let $\tilde{\phi}_{i-p}^{\text{tur}}, p \in \mathbb{N}$, be the most recent reconstruction among those used to compute the windshift vector. In order to calculate $v$, we focus on a set of reconstructed wavefront data within a fixed, reduced telescope aperture $\Omega_A(\text{area between dashed circles}).$ The magenta and blue dots correspond to the values which are actually used for the calculation of the shift vector. Shifted values (blue dots) only get interpolated based on given data—no extrapolation is needed.

$$
\Omega_A = \{ r \in \mathbb{R}^2 : a_{\text{int}} + a \leq |r| \leq a_{\text{out}} - a \} \subset \Omega_A
$$

(33)

with some a priori chosen constant $a \in \mathbb{R}^+$. For a set $I \subset \{1, \ldots, n_S\}$ of indices let $\Omega_I := \bigcup_{r \in I} \Omega_r$ define a reduced subaperture grid such that $\Omega_I \subseteq \Omega_A$. Furthermore, let $N_I := \{ \tilde{n} \in \{1, \ldots, n_N\} : r^{\tilde{n}} \in \Omega_I \}$ be the set of indices of those nodes which correspond to the corners of the subapertures in $\Omega_I$. Consider $w \in \mathbb{R}^2$ such that $r^{(n)} - w \in \Omega_S$ for all $n \in N_I$. Introducing the operator

$$
T^w : \mathbb{R}^{n_N} \rightarrow \mathbb{R}^{|N_I|}, \quad \phi \mapsto (T^w \phi)(n) := \frac{\sum_{s=1}^{n_S} \chi_{\Omega_r}(r^{(s)} - w) \beta_s(\phi; r^{(s)} - w)}{\sum_{s=1}^{n_S} \chi_{\Omega_r}(r^{(s)} - w)}, \quad n \in N_I,
$$

(34)

the shift vector $v$ is calculated by solving the minimization problem

$$
\arg\min_{v \in \mathbb{R}^2} \left\| T^{0} \Phi_{i,p,q}^{\text{tur}} - T^{w} \nabla \Phi_{i,p,q}^{\text{tur}} \right\|^2_{\mathbb{R}^{|N_I|}}
$$

(35)
with $q_i \coloneqq \min\{i - (m + p), q_{\text{max}}\}$, where $q_{\text{max}} \in \mathbb{N}$ defines the maximal number of reconstructions used for averaging (32), and $m \in \mathbb{N}$ is used to make the translation effect of small windshifts more visible. Note that, if $\mathbf{w}$ equals the zero vector, then $\mathbf{T}^\mathbf{w}$ only corresponds to a restriction operator. By the definition of $q_i$, and (32), $\mathbf{v}_i$ can only be calculated for $i > m + p$. For $i \leq m + q$ we define $\mathbf{v}_i$ as the zero vector. In this case, no shift will be done, i.e. $\tilde{\mathbf{v}}_i = \tilde{\mathbf{v}}_{1,\text{tur}}$.

For any given vector $\mathbf{v} \in \mathbb{R}^2$, it holds that $\{\mathbf{r} \in \mathbb{R}^2 : \mathbf{r} + m\Delta T\mathbf{v} \in \Omega_N^\gamma\} \subset \Omega_N^\gamma$, if and only if $a \geq m\Delta T\gamma\mathbf{v}$. Thus, from the theoretical point of view, $a > m\Delta T\gamma\mathbf{v}_{\text{max}}$ would be an appropriate criterion for the constant, since the windshift hardly changes over a certain amount of time. However, for practical use, we may replace the factor $\max\{|\mathbf{v}_i|\}_{j=m+p+1}^{|\mathbf{v}_i|_{\text{tur}}}^{\text{rel}}$ by an upper bound $\mathbf{v}_{\text{max}}$ for the wind speed, and choose the constant according to $a > m\Delta T\gamma\mathbf{v}_{\text{max}}$. A possible choice for $\mathbf{v}_{\text{max}}$ would be the wind velocity at which the observation is no longer possible. Note that the choice of $a$ represents a compromise between the reduction of the computational cost and the loss of information used for the calculation of the shift vectors.

Note that the parameter $p$ in (35) has no effect on the computational cost of the minimization problem. Actually, $p$ relates to the number of time frames needed to solve a problem of the form (35). If $d_1 \Delta T$, $d_1 \in \mathbb{R}^+$, defines an upper bound on the required time needed to calculate a shift vector, then we may choose $p = \min\{\tilde{p} \in \mathbb{N} : \tilde{p} \geq d_1\}$.

To guarantee a higher stability in the calculation of the shift vectors, one can also think of using exponential smoothing [19], i.e. if $\tilde{\mathbf{v}}_i$ denotes the solution of (35), then, for a fixed $\gamma_{\text{smooth}} \in (0, 1)$, we compute

$$\mathbf{v}_i = \gamma_{\text{smooth}}\tilde{\mathbf{v}}_i + (1 - \gamma_{\text{smooth}})\mathbf{v}_{i-1}. \quad (36)$$

### 4.8. Predictive control algorithm

Summarized, we obtain the following temporal control algorithm:

**Algorithm 1.** Predictive control for SCAO.

1: for $i = 1$ to $n - d$ do
2: $\mathbf{\Phi}_i$ \footnotesize{$\text{rec} =$} $\mathbf{R}_i\mathbf{F}_i$ \hfill $\triangleright$ reconstruction of residual WF based on given data $\mathbf{R}_i$
3: $\tilde{\mathbf{\Phi}}_i$ \footnotesize{$\text{tur} =$} $\gamma_{\text{in}}\mathbf{\Phi}_{i-1}$ + $\mathbf{P}_{\text{dm}}\mathbf{\Phi}_{i-1}$ \hfill $\triangleright$ calculate pseudo open loop data
4: if $i \leq m + p$ then
5: $\mathbf{\phi}_{1,m+p} = \mathbf{B}\mathbf{\Phi}_{i-1}$ \hfill $\triangleright$ projection onto DM (via bilinear interpolation)
6: else
7: $q_i = \min\{i - (m + p), q_{\text{max}}\}$ \hfill $\triangleright$ number of reconstructions used to calculate propagation vector
8: $\tilde{\mathbf{\phi}}_{1,m+p,q_i} = \frac{1}{q_i} \sum_{j=1}^{j=q_i-1} \mathbf{\Phi}^\text{turr}_{i-p,j}$ \hfill $\triangleright$ average value of reconstructions
9: $\tilde{\mathbf{\phi}}_{1,m+p,q_i} = \frac{1}{q_i} \sum_{j=1}^{j=q_i-1} \mathbf{\Phi}^\text{mix}_{i-(m+p)-j}$ \hfill $\triangleright$ calculation of propagation vector
10: $\tilde{\mathbf{v}}_i = \arg\min_{\mathbf{v} \in \mathbb{R}^2} \left\{ \mathbf{r}^T\tilde{\mathbf{\phi}}^\text{turr}_{1,m+p,q_i} - \mathbf{r}^T\Delta \mathbf{\phi}^\text{mix}_{1,m+p,q_i} \right\}_{\mathbb{R}^2}$ \hfill $\triangleright$ smoothing of calculated propagation vector
11: $\mathbf{v}_i = \gamma_{\text{smooth}}\tilde{\mathbf{v}}_i + (1 - \gamma_{\text{smooth}})\mathbf{v}_{i-1}$ \hfill $\triangleright$ projection onto DM
12: $\mathbf{\phi}_{1,m+p,q_i} = (S\tilde{\mathbf{\phi}}_{1,m+p,q_i})[\tilde{\mathbf{\Phi}}^\text{tur}_{i-1}]$ \hfill $\triangleright$ projection onto DM
13: end if
14: end for
5. Numerical validation

For our simulations we consider a SCAO system operated in closed loop configuration on an assumed 37 m telescope with a central obstruction of 28% and a diameter of 18.5 m and 0.028. The WFS measurements from the observed guide star are received by a Shack–Hartmann wavefront sensor with 74 × 74 subapertures, which measure the wavefront slopes at a frequency of 500 Hz (i.e. Δτ = 0.002 s) and a sensing wavelength of 700 nm. The edge length of such a subaperture is 0.5 m. Since not all subapertures are exposed equally, only nS = 3832 of them are active. Overall, we have nN = 4204 grid nodes.

To demonstrate the qualitative performance of the proposed temporal control method, we use European Southern observatory’s (ESO’s) official end-to-end simulation tool OCTOPUS [17, 18]. The software enables to simulate a certain number of atmospheric turbulence layers located at different altitudes. Each layer is described by a von Karman model with outer scale r0.4 and is included in (36). Furthermore, the algorithm involves the reconstruction operator R (19), the collocation operator P (see (20)), the operator Ei (25) used for extrapolation, and the shift operators T* (34) and Sρ (28), respectively. We also use the operator B, which is similar defined as the interpolation operator Bi but only on the original aperture ΩA.

For wavefront reconstruction we use the CuReD algorithm [32]. The loop gain γgain given in (20) depends on the used reconstruction method among other things and is found empirically. In our case, we consider γgain ∈ {0.1, 0.2, ..., 0.9}. In order to extend the reconstructed data we extrapolate the given values on the boundary of the subaperture grid ΩS. Depending on the position of the node r(n) outside of ΩS, we choose ρn as a multiple of the edge length g such that Bρn(r(n)) ∩ ΩS (see (24)) only contains boundary nodes of the subaperture grid. The weights in (25) are chosen as ωn = 1/2πω with some damping factor ω. In our case, we consider ω = 0.7. The DM, which is conjugated to the ground layer, is modeled by continuous (ν = 0 in (26)), piecewise bilinear functions, i.e. β(ke) ∈ Q1 = span{1, r1, r2, r1r2}, with a total number of 4204 degrees of freedom. In order to calculate the shift vector we focus on a reduced data set with |N| = 1538 given values. This corresponds to an approximately 60% reduced aperture ΩA in (33). Furthermore, we use the parameters p = 4, qmax = 8, m = 2 in (35) and γsmooth = 0.4 in (36) for all test cases. To calculate the shift vectors (i.e. to solve problem (35)) we use the MATLAB® function fminsearch, which is based on the downhill simplex method (see [16]).

We simulate two seconds of real time for different configurations of atmospheric layers, where we assume d = 2, 3, 4 frames of delay. The achieved AO correction quality is evaluated in terms of long exposure Strehl ratios in [0, 1] for a wavelength of 2200 nm after nsc = 1000 iteration steps, where 1 is best and 0 worst. We run simulations for a NGS photon flux of 10000 photons per subaperture per frame (ppspf) (high flux) and also for 10 ppspf (low flux).
Furthermore, we assume that the WFS measurements suffer from photon noise and read-out noise.

As test cases, we focus on single layer atmospheres and atmospheres with $n_L = 2, 3, 9$ turbulence layers. In the former case we use one atmospheric layer located at an altitude of 30 m, which is simulated to move with a wind speed of 25 m s$^{-1}$, 15 m s$^{-1}$ and 5 m s$^{-1}$, respectively. For the two- and three-layer atmospheres, we simulate two different layer configurations for demonstration purposes. Finally, we also consider the nine-layer ESO median seeing atmosphere.

All main parameters used for the simulations are summarized in table 1. A detailed list of the heights, wind speeds and shift directions of the single layers, and also the normalized layer strengths $\gamma_l$ can be found in table 2.

For the numerical validation, we compare the results of the AO correction using the mirror shape $\hat{\phi}_{\text{dm}}^{d_i} = \hat{\phi}_{\text{tur}}^{d_i}$ determined by the prediction step (29) with those obtained by only using $\hat{\phi}_{\text{dm}}^{d_i} = \hat{\phi}_{\text{tur}}^{d_i}$ (integrator-type control), where $\hat{\phi}_{\text{tur}}^{d_i}$ corresponds to the continuous extension of the reconstruction (20).

As it can be seen from figures 9 and 10, we obtain a strong boost in performance in the low flux case for all simulated layer configurations. In the high flux case, we only get an improved AO correction in the cases where the wind speeds and also the time delays are large. It is noted that, in the high flux case, the main temporal limitation is due to time delays. From (14) and (23) together with the Greenwood time delay (2), we obtain that the perturbations in the incoming WF are approximately constant for $j\Delta_T < \tau_0$, i.e. for $j \in \mathbb{N}_0$ with

### Table 1. Parameters of the OCTOPUS simulations for testing the proposed temporal control method.

| OCTOPUS simulation parameters | Value |
|-----------------------------|-------|
| Telescope diameter          | 37 m  |
| Central obstruction         | 28%   |
| Science target              | SCAO on-axis |
| Wavefront sensor            | Shack–Hartmann sensor |
| Sensing wavelength          | 700 nm |
| Number of subapertures      | $74 \times 74 = 5476$ |
| Number of active subapertures | 3832 |
| Subaperture edge length     | 0.5 m |
| Loop frequency              | 500 Hz |
| Frame length (sampling period) | 0.002 s |
| DM influence functions      | Piecewise bilinear functions |
| Number of actuators         | 4204  |
| DM delay                    | 1, 2, 3 time frame(s) |
| Total delay                 | 2, 3, 4 time frames |
| Turbulence model            | von Karman with $L_0 = 25$ m |
| Fried parameter             | 0.157 m (median atmosphere) |
| Number of turbulence layers | 1, 2, 3, 9 (see table 2 for details) |
| Turbulence temporal evolution | Taylor frozen flow assumption |
| Evaluation wavelength       | 2200 nm |
| Photon flux                 | 10 000 ppspf (high flux), 10 ppspf (low flux) |
| Number of time steps        | 1000  |
Table 2. Characteristics of the turbulence layers of the simulated single-layer, two-layer, three-layer and nine-layer atmospheres for different test cases.

| Single-layer atmospheres: | Height | Speed | Direction |
|--------------------------|--------|-------|-----------|
| 1st case                 | 30 m   | 5 m s\(^{-1}\) | (1, 0)    |
| 2nd case                 | 30 m   | 15 m s\(^{-1}\) | (1, 0)    |
| 3rd case                 | 30 m   | 25 m s\(^{-1}\) | (1, 0)    |

| Two-layer atmospheres:               | Height | Speed | Direction | \(\gamma_l\) |
|-------------------------------------|--------|-------|-----------|-------------|
| 1st case                            | 1. layer | 30 m | 15 m s\(^{-1}\) | (1, 0) | 0.85 |
|                                      | 2. layer | 200 m | 5 m s\(^{-1}\) | (−1, 0) | 0.15 |
| 2nd case                            | 1. layer | 30 m | 5 m s\(^{-1}\) | (1, 0) | 0.85 |
|                                      | 2. layer | 200 m | 15 m s\(^{-1}\) | (−1, 0) | 0.15 |

| Three-layer atmospheres:            | Height | Speed | Direction | \(\gamma_l\) |
|-------------------------------------|--------|-------|-----------|-------------|
| 1st case                            | 1. layer | 140 m | 25 m s\(^{-1}\) | (1, 0) | 0.80 |
|                                      | 2. layer | 2500 m | 10 m s\(^{-1}\) | (−1, 0) | 0.15 |
|                                      | 3. layer | 12000 m | 15 m s\(^{-1}\) | (1, 0) | 0.05 |
| 2nd case                            | 1. layer | 140 m | 10 m s\(^{-1}\) | (1, 0) | 0.80 |
|                                      | 2. layer | 2500 m | 20 m s\(^{-1}\) | (−1, 0) | 0.15 |
|                                      | 3. layer | 12000 m | 25 m s\(^{-1}\) | (1, 0) | 0.05 |

| Nine-layer atmosphere:              | Height | Speed | Direction | \(\gamma_l\) |
|-------------------------------------|--------|-------|-----------|-------------|
| 1. layer                            | 30 m   | 15 m s\(^{-1}\) | (1, 0) | 0.5224 |
| 2. layer                            | 140 m  | 13 m s\(^{-1}\) | (−1, 0) | 0.026 |
| 3. layer                            | 281 m  | 13 m s\(^{-1}\) | (1, 0) | 0.0444 |
| 4. layer                            | 562 m  | 9 m s\(^{-1}\) | (−1, 0) | 0.116 |
| 5. layer                            | 1125 m | 9 m s\(^{-1}\) | (1, 0) | 0.0989 |
| 6. layer                            | 2250 m | 15 m s\(^{-1}\) | (−1, 0) | 0.0295 |
| 7. layer                            | 4500 m | 25 m s\(^{-1}\) | (1, 0) | 0.0598 |
| 8. layer                            | 9000 m | 40 m s\(^{-1}\) | (−1, 0) | 0.043 |
| 9. layer                            | 18000 m | 21 m s\(^{-1}\) | (1, 0) | 0.06 |

\[
j < \mathcal{J} := 0.314 \frac{p_0}{\Delta T |\mathbf{p}|} \tag{37}\] and \(|\mathbf{p}|\) given as in (3). Hence, in the high flux case, predictive control may not always lead to an improvement in performance. For all simulated test cases, the values of \(|\mathbf{p}|\) and \(\mathcal{J}\) are given in table 3. However, since prediction-based control uses the limited photons obtained from faint guide stars optimally, an improvement is expected in the low flux case.
As already mentioned, the parameter \( \gamma \) in (20) (which is used for stability purposes) is found empirically and mainly depends on the reconstruction method and the accuracy of the measurement data. In particular, \( \gamma \) will be different for high flux and low flux considerations and has to be determined for each simulation case separately. In Table 4, we have listed the values of \( \gamma \) for all simulated test cases. At a first glance, it may be confusing that \( \gamma \) is smaller in case of low wind velocities and less time delays as for the other cases. However, since SCAO is operated in closed loop configuration, formula (20) can also be seen as update of the discrete DM shape \( P_{\phi_{dm}} \) with \( \gamma_{rec} \). Since perturbations in the incoming WF are approximately constant for low wind velocities and less time delays (see discussion above), the DM shape remain approximately constant with respect to a certain number of time frames and thus an essential update of the DM shape is not required, i.e. \( \gamma \) is usually small.

For a SCAO system, the WFS measures the sum of the turbulence contributions of the single atmospheric layers along the optical axis. This is also the only information available for the computation of the new mirror shape. Therefore, the use of shift-based predictive control methods becomes less accurate as the number of turbulence layers increases, which can move in different directions. Based on the approximations \( \tilde{\phi}_{\text{tur}} \), only one shift vector can be determined, which mainly includes information of the shift direction of the most dominant layer. Nonetheless, we obtain a good AO correction in the low flux case for the nine-layer ESO median seeing atmosphere (see figure 10).

| Table 3. Values of \(|\phi|\) and \(J\) for different layer configurations and different test cases. |
|---------------------------------|----------------|----------------|
| \( |\phi| \)                      | \( J \)         |
| One layer 1st case              | 5 m s\(^{-1}\)  | 4.93           |
| 2nd case                           | 15 m s\(^{-1}\) | 1.64           |
| 3rd case                           | 25 m s\(^{-1}\) | 0.99           |
| Two layers 1st case              | 13.8 m s\(^{-1}\) | 1.78           |
| 2nd case                           | 7.1 m s\(^{-1}\) | 3.48           |
| Three layers 1st case            | 22.7 m s\(^{-1}\) | 1.08           |
| 2nd case                           | 12.8 m s\(^{-1}\) | 1.93           |
| Nine layers                         | 16.4 m s\(^{-1}\) | 1.50           |

| Table 4. Values for \(\gamma_{\text{fin}}\) for all simulated test cases. |
|---------------------------------|----------------|----------------|
| \( d = 2 \) | \( d = 3 \) | \( d = 4 \) | \( d = 2 \) | \( d = 3 \) | \( d = 4 \) |
| One layer 1st case | 0.3 | 0.4 | 0.6 | 0.1 | 0.1 | 0.2 |
| 2nd case                      | 0.6 | 0.6 | 0.6 | 0.2 | 0.2 | 0.2 |
| 3rd case                      | 0.7 | 0.6 | 0.6 | 0.2 | 0.2 | 0.3 |
| Two layers 1st case                   | 0.7 | 0.7 | 0.7 | 0.2 | 0.2 | 0.3 |
| 2nd case                      | 0.6 | 0.6 | 0.7 | 0.2 | 0.2 | 0.3 |
| 1st case                      | 0.8 | 0.7 | 0.7 | 0.3 | 0.3 | 0.4 |
| Three layers 2nd case               | 0.7 | 0.8 | 0.8 | 0.2 | 0.3 | 0.3 |
| Nine layers                       | 0.7 | 0.8 | 0.8 | 0.3 | 0.3 | 0.4 |
Figure 9. Long exposure (LE) Strehl ratios for different configurations of the simulated single-layer (above) and two-layer (below) atmospheres (defined in table 2) using the CuReD method combined with 1) integrator-type control and with 2) shift-based predictive control.
Figure 10. Long exposure (LE) Strehl ratios for different configurations of the simulated three-layer (above) and nine-layer (below) atmospheres (defined in table 2) using the CuReD method combined with 1) integrator-type control and with 2) shift-based predictive control.
6. Conclusion and outlook

In this paper, we consider a new, fully deterministic temporal control approach for SCAO systems operated in closed loop configuration. The derivation of the method is based on a detailed mathematical model of the AO correction process in terms of operator equations. The proposed method is based on the separation of the correction process into subproblems, including a prediction step, which is used to reduce the temporal error in the AO system. In particular, this approach enables to incorporate computationally efficient WF reconstruction algorithms like the CuReD method. The prediction step is carried out by an appropriate shift of the reconstruction. The corresponding shift vector is calculated based on a sequence of previously reconstructed wavefronts. The qualitative performance of the proposed temporal control method is illustrated in high flux and low flux regimes.

In view of the numerical results, the proposed control algorithm is a promising method to counteract temporal delays in SCAO. Due to its simplicity and flexibility, the algorithm can easily be adapted to a pyramid wavefront sensor and other WF reconstruction methods without major changes. One can also incorporate other methods for the computation of the shift vector and it is also possible to adopt the algorithm to other DM models. We believe this is also a very promising temporal control method adaptable to multi conjugate adaptive optics (MCAO) systems, which use multiple numbers of wavefront sensors and DM. Future work includes the development of more efficient methods for calculating the shift vectors and a comparison of the proposed method to other control algorithms—in particular with Kalman filter-based control algorithms.

Acknowledgments

This work was funded by the Austrian Science Fund (FWF): W1214-N15, project DK 8 and F6805-N36, project 5. The authors want to thank Andreas Obrederer, Stefan Raffetseder, Roland Wagner and Simon Hubmer for valuable suggestions and helpful discussions.

ORCID iDs

Markus Pöttinger © https://orcid.org/0000-0001-6471-8917
Ronny Ramlau © https://orcid.org/0000-0002-0277-9615

References

[1] Folland G B 1999 Real Analysis: Modern Techniques and Their Applications 2nd edn (New York: Wiley)
[2] Basden A, Jenkins D, Morris T, Osborn J and Townson M 2019 Efficient implementation of pseudo open loop control for adaptive optics on extremely large telescopes Mon. Not. R. Astron. Soc. 486 1774–80
[3] Berdja A and Borgnino J 2007 Modelling the optical turbulence boiling and its effect on finite-exposure differential image motion Mon. Not. R. Astron. Soc. 378 1177–86
[4] Johnson C L, Gavel T D, Reinig M and Wiberg D M 2008 Wind estimation and prediction for adaptive optics control systems Proc. SPIE 7015 70153F
[5] Correia C, Conan J-M, Kulcsár C, Raynaud H-F and Petit C 2010 Adapting optimal LQG methods to ELT-sized AO systems 1st AO4ELT Conf.—Adaptive Optics for Extremely Large Telescopes 07003
[6] Correia C, Raynaud H-F, Kulcsár C and Conан J-M 2010 On the optimal reconstruction and control of adaptive optical systems with mirror dynamics J. Opt. Soc. Am. A 27 333–49
[7] Fraanje R, Rice J, Verhaegen M and Doelman N 2010 Fast reconstruction and prediction of frozen flow turbulence based on structured Kalman filtering J. Opt. Soc. Am. A 27 A235–45
[8] Gray M, Petit C, Rodionov S, Bertino L, Bocquet M and Fusco T 2013 Local ensemble transform Kalman filter: a non-stationary control law for complex adaptive optics systems on ELTs 3rd AOELT Conf. - Adaptive Optics for Extremely Large Telescopes (Florence, Italy, May 2013) (https://doi.org/10.12839/AO4ELT3.13253)
[9] Gray M, Petit C, Rodionov S, Bocquet M, Bertino L, Ferrari M and Fusco T 2014 Local ensemble transform Kalman filter, a fast non-stationary control law for adaptive optics on ELTs: theoretical aspects and first simulation results Opt. Express 22 20894–913
[10] Hardy J W 1998 Adaptive Optics for Astronomical Telescopes (Oxford: Oxford University Press)
[11] Doelman N J, Hinnen K J G, Stoffelen F J G and Verhaegen M 2004 Optimal control strategy to reduce the temporal wavefront error in AO systems Proc. SPIE 5490 5490
[12] Johnson L C, Gavel D T and Wiberg D M 2011 Bulk wind estimation and prediction for adaptive optics control systems J. Opt. Soc. Am. A 28 1566–77
[13] Kindermann S, Neubauer A and Ramlau R 2012 A singular value decomposition for the Shack–Hartmann based wavefront reconstruction J. Comput. Appl. Math. 236 2186–99
[14] Tyson R K 2000 Introduction to Adaptive Optics (Bellingham, WA: SPIE Optical Engineering Press)
[15] Kulcsár C, Raynaud H-F, Petit C, Conان J-M and de Lesegno P V 2006 Optimal control, observers and integrators in adaptive optics Opt. Express 14 7464–76
[16] Lagarias J, Reeds J, Wright M and Wright P 1998 Convergence properties of the nelder–mead simplex method in low dimensions SIAM J. Optim. 9 112–47
[17] Le Louarn M, Verinaud C, Korkiakoski V, Hubin N and Marchetti E 2006 Adaptive optics simulations for the European extremely large telescope Proc. SPIE 6272 98
[18] Massioni P, Gilles L and Ellerbroek B 2015 Adaptive distributed Kalman filtering with wind estimation for astronomical adaptive optics J. Opt. Soc. Am. A 32 2353–64
[19] Massioni P, Kulcsár C, Raynaud H-F and Conан J-M 2011 Fast computation of an optimal controller for large-scale adaptive optics J. Opt. Soc. Am. 28 2298–309
[20] Massioni P, Raynaud H-F, Kulcsár C and Conан J-M 2014 An approximation of the Riccati equation in large-scale systems with application to adaptive optics IEEE Trans. Control Syst. Technol. 23 1
[21] Parisot A, Petit C and Fusco T 2011 Experimental comparison of wide field AO control schemes using the Homer AO bench 2nd Int. Conf. on Adaptive Optics for Extremely Large Telescopes p 60
[22] Paschall R N and Anderson D J 1993 Linear quadratic Gaussian control of a deformable mirror adaptive optics system with time-delayed measurements Appl. Opt. 32 6347–58
[23] Petit C, Conан J-M, Kulcsár C, Raynaud H-F and Fusco T 2008 First laboratory validation of vibration filtering with LQG control law for adaptive optics Opt. Express 16 87–97
[24] Poyneer L A and Véran J-P 2010 Kalman filtering to suppress spurious signals in adaptive optics control J. Opt. Soc. Am. A 27 A223–34
[25] Poyneer L A, Macintosh B A and Véran J-P 2007 Fourier transform wavefront control with adaptive prediction of the atmosphere J. Opt. Soc. Am. A 24 2645–60
[26] Poyneer L and Véran J-P 2008 Predictive wavefront control for adaptive optics with arbitrary controller loop delays J. Opt. Soc. Am. A 25 1486–96
[27] Vogel C R 2004 Sparse matrix methods for wavefront reconstruction, revisited Proc. SPIE 5490
[28] Ramlau R and Rosensteiner M 2012 An efficient solution to the atmospheric turbulence tomography problem using Kaczmarz iteration Inverse Problems 28 095004
[29] Roddier F 1999 Adaptive Optics in Astronomy (Cambridge: Cambridge University Press)
[30] Roggemann M C and Welsh B 1996 Imaging Through Turbulence (CRC Press Laser and Optical Science and Technology Series) (Boca Raton, FL: CRC Press)
[31] Rosensteiner M 2012 Wavefront reconstruction for extremely large telescopes via CuRe with domain decomposition J. Opt. Soc. Am. A 29 2328–36
[32] Roux B L, Conan J-M, Kulcsár C, Raynaud H-F, Mugnier L M and Fusco T 2004 Optimal control law for classical and multiconjugate adaptive optics J. Opt. Soc. Am. A 21 1261–76
[34] Schweizer B 2013 *Partielle Differentialgleichungen: Eine Anwendungsorientierte Einführung* (Berlin: Springer)

[35] Sivo G et al 2014 First on-sky SCAO validation of full LQG control with vibration mitigation on the CANARY pathfinder *Opt. Express* 22 23565–91

[36] Gavel D T and Wiberg D M 2003 Toward strehl-optimizing adaptive optics controllers *Proc. SPIE* 4839

[37] Wilson R W 2002 SLODAR: measuring optical turbulence altitude with a Shack–Hartmann wavefront sensor *Mon. Not. R. Astron. Soc.* 337 103–8

[38] Yudytskiy M 2014 Wavelet methods in adaptive optics *PhD Thesis* Johannes Kepler University Linz