Adaptive Distributed Kalman Filtering with Wind Estimation for Astronomical Adaptive Optics

Paolo Massioni, Luc Gilles, Brent Ellerbroek

Adaptive Distributed Kalman Filtering with Wind Estimation for Astronomical Adaptive Optics

PAOLO MASSIONI[1*], LUC GILLES[2], AND BRENT ELLERBROEK[2]

[1]Laboratoire Ampère (UMR CNRS 5005), INSA de Lyon, Université de Lyon, 69621 Villeurbanne CEDEX, France
[2]Thirty Meter Telescope Observatory Corp., 1200 E. California Blvd., MC 102-8, Pasadena, California 91125, USA
[*]Corresponding author: paolo.massioni@insa-lyon.fr

Compiled April 30, 2019

In the framework of Adaptive Optics (AO) for astronomy, it is a common assumption to consider the atmospheric turbulent layers as "frozen flows" sliding according to the wind velocity profile. For this reason, having knowledge of such a velocity profile is beneficial in terms of AO control system performance. In this paper we show that it is possible to exploit the phase estimate from a Kalman filter running on an AO system in order to estimate wind velocity. This allows the update of the Kalman filter itself with such a knowledge, making it adaptive. We have implemented such an adaptive controller based on the distributed version of the Kalman filter, for a realistic simulation of a multi-conjugate AO system with laser guide stars on a 30 m telescope. Simulation results show that this approach is effective, promising, and the additional computational cost with respect to the distributed filter is negligible. Comparisons with a previously published SLOpe Detection And Ranging (SLODAR) wind profiler are made and the impact of turbulence profile quantization is assessed. One of the main findings of the paper is that all flavors of the adaptive distributed Kalman filter are impacted more significantly by turbulence profile quantization than the static minimum mean square estimator which does not incorporate wind profile information.

OCIS codes: *(010.1080) Active or adaptive optics; (010.1330) Atmospheric turbulence; (010.7350) Wave-front sensing.*

1. INTRODUCTION

Adaptive optics (AO) systems compensate for wavefront distortions due to atmospheric turbulence by means of one or more deformable mirrors (DMs) inserted in the optical path [1]. The DMs can partially cancel these distortions if the shape of the wavefront is known. Wavefront information can be obtained from wavefront sensor (WFS) measurements and a wavefront reconstruction algorithm. As the DM response time is usually much shorter than the AO loop sampling rate, the problem of optimal AO control boils down to an optimal wavefront estimation problem, which, in order to beat down the servo lag error, requires a one- or two-step ahead prediction. The classical solution to this estimation problem is the so-called minimum mean square error (MMSE) estimator [2], also known as static minimum variance reconstructor (MVR), which uses the last available WFS measurement and knowledge of turbulence spatial statistics at a frozen instant in time to compute the optimal estimation. The MMSE estimator can be improved by incorporating a dynamical turbulence model, which then leads to a Kalman filter [3]. Such a filter computes a turbulence estimate based upon all prior WFS measurements and a dynamical turbulence model which takes into account atmospheric wind velocity (speed and direction) at different altitudes. Provided the filter is constructed from an accurate dynamical model, it is expected to yield the best performance due to its optimality property. Experimental verification of the frozen flow hypothesis has recently been reported [4], and a Kalman filter using a simple autoregressive turbulence temporal evolution model without wind velocity knowledge has recently been demonstrated on the sky [5].

The standard version of the Kalman filter proves to be impractical in terms of implementation for large-scale systems, like extremely large telescopes (ELTs) currently under design and construction [6–8], due to the excessive cost to solve the Riccati equation, which is a central step in the computation of the filter. Several approaches have been proposed to overcome this computational bottleneck [9–12], among which, a distributed version of the Kalman filter [13]. Such a distributed Kalman filter (DKF) can also easily accommodate wind velocity information, which, if available and accurate, has the potential to provide a boost in performance [14]. The problem of robust and accurate wind velocity profile estimation is still open, in particular assessing the impact of the vertical turbulence profile quantization, i.e. the discretization by the reconstruction algorithm of the turbulence profile into a number of layers much smaller than the number of real (or simulated) layers, while maintain-

ing computational cost reasonable; a few solutions have already been developed [15–18]. In this paper, we propose using the distributed Kalman filter itself to estimate wind velocity, making the filter adaptive. We will show that this capability can be implemented with a simple algorithm at a negligible additional cost and a sensible boost in performance [19]. Comparison with a previously published SLOpe Detection And Ranging (SLODAR) wind profiler [15] is made and the impact of turbulence profile quantization is assessed. We will also show that for multi-object adaptive optics (MOAO), this approach has the potential to be cheaper in terms of both offline and online computational cost compared to the MVR solution.

The article is organized as follows. Section 2 reviews the distributed Kalman filter (DKF), Section 3 shows how wind profile information can be integrated into it, and Section 4 describes the wind profile estimation from the Kalman filter phase estimate, which leads to the adaptive distributed Kalman filter (ADKF) described in Section 5. Section 6 discusses the stability of the proposed filter, Section 7 the computational cost and Section 8 covers sample simulation results including a comparison with a previously published SLODAR wind profiler [15] and the assessment of the impact of turbulence profile quantization. Finally, conclusions are drawn in Section 9.

2. REVIEW OF THE DISTRIBUTED KALMAN FILTER

We start by reviewing the distributed Kalman filter originally proposed in [13]. We consider a tomographic multi-conjugate (MCAO) or multi-object (MOAO) adaptive optics system. We assume that the wavefront is perturbed by n_L turbulence phase screens at altitudes $h^{(l)}$, and sensed by n_{gs} laser guide star (LGS) WFS in directions $(\alpha_1^{(s)}, \alpha_2^{(s)})$ measured in radians from the center of the field of view (FoV). We also assume that the first layer coincides with the telescope aperture plane at ground level, i.e. $h^{(1)} = 0$.

All phase screens are sampled on $n \times n$ square grids in the cone coordinate system [20, 21] so as to yield shift invariance for guide stars at finite range [14]. This means that, if LGSs are at a mean altitude h_{gs}, and if the ground layer has sampling $\Delta x = \Delta x^{(1)}$, where Δx denotes the LGS WFS spatial sampling, then layer l at altitude $h^{(l)}$ has sampling $\Delta x^{(l)} = \xi^{(l)} \Delta x^{(1)}$, where the cone compression factor is given by $\xi^{(l)} = (1 - h^{(l)}/h_{gs}) < 1$. We denote by $n_1, n_2 = -n/2, -n/2+1, \cdots, n/2-1$ the 2-dimensional spatial domain (SD) or Fourier domain (FD) index of a phase point on a given screen and by $\phi_{n_1,n_2} \in \mathbb{R}^{n_L}$ the concatenated vector of the n_L phase values at grid point (n_1, n_2) on the n_L phase screens, i.e. a vertical slice of phase through all atmospheric layers. SD coordinates are given by $(x_1^{(l)}, x_2^{(l)}) = (n_1, n_2)\Delta x^{(l)}$, whereas FD coordinates are given by $(f_1^{(l)}, f_2^{(l)}) = (n_1, n_2)\Delta f^{(l)}$, with $\Delta f^{(l)} = \Delta f^{(1)}/\xi^{(l)}$ and $\Delta f^{(1)} = 1/(n\Delta x^{(1)})$. Under the hypothesis of a turbulence autoregressive model of order 1 (AR1) [3], the state space model takes the following form in the SD:

$$\begin{cases} \phi_{n_1,n_2}(k+1) = a_{n_1,n_2}(\phi(k)) + v_{n_1,n_2}(k) \\ y_{n_1,n_2}(k) = f_{n_1,n_2}(\phi(k)) + w_{n_1,n_2}(k) \end{cases} \quad \textbf{(1)}$$

where k is temporal index, $\phi_{n_1,n_2} \in \mathbb{R}^{n_L}$ the concatenated turbulence phase vector, ϕ the concatenated turbulence phase vector for all grid points and phase screens, $v_{n_1,n_2} \in \mathbb{R}^{n_L}$ the process noise, $y_{n_1,n_2} \in \mathbb{R}^{2n_{gs}}$ the WFS measurement vector, i.e. a slice of measurements through all WFSs, $w_{n_1,n_2} \in \mathbb{R}^{2n_{gs}}$ the WFS measurement noise, a_{n_1,n_2} the state transition operator and f_{n_1,n_2} the phase-to-measurement operator further described below. WFS measurements are sampled at spatial resolution Δx on the same grid as the ground level phase screen. For a Shack-Hartmann (SH) WFS, f_{n_1,n_2} consists of a stack of n_{gs} functions $f_{n_1,n_2}^{(s)}$ (with $s = 1, \ldots, n_{gs}$), defined as follows:

$$f_{n_1,n_2}^{(s)}(\phi) = \frac{1}{2} \sum_{l=1}^{n_L} \begin{bmatrix} f_1^{(s,l)} \\ f_2^{(s,l)} \end{bmatrix} \quad \textbf{(2)}$$

$$f_1^{(s,l)} = \phi^{(l)}_{n_1+1+\delta_1^{(s,l)}, n_2+\delta_2^{(s,l)}} + \phi^{(l)}_{n_1+1+\delta_1^{(s,l)}, n_2+1+\delta_2^{(s,l)}} - \phi^{(l)}_{n_1+\delta_1^{(s,l)}, n_2+\delta_2^{(s,l)}} - \phi^{(l)}_{n_1+\delta_1^{(s,l)}, n_2+1+\delta_2^{(s,l)}} \quad \textbf{(3)}$$

$$f_2^{(s,l)} = \phi^{(l)}_{n_1+\delta_1^{(s,l)}, n_2+1+\delta_2^{(s,l)}} + \phi^{(l)}_{n_1+1+\delta_1^{(s,l)}, n_2+1+\delta_2^{(s,l)}} - \phi^{(l)}_{n_1+\delta_1^{(s,l)}, n_2+\delta_2^{(s,l)}} - \phi^{(l)}_{n_1+1+\delta_1^{(s,l)}, n_2+\delta_2^{(s,l)}} \quad \textbf{(4)}$$

where $\phi^{(l)}_{n_1,n_2}$ denotes component l of ϕ_{n_1,n_2}, $\delta_1^{(s,l)} = \alpha_1^{(s)} h^{(l)}/\Delta x^{(l)}$ and $\delta_2^{(s,l)} = \alpha_2^{(s)} h^{(l)}/\Delta x^{(l)}$ are the normalized spatial shifts of ray intercepts at layer l from LGS s. In general $\delta_1^{(s,l)}$ and $\delta_2^{(s,l)}$ will not be integers; $\phi^{(l)}_{n_1+\delta_1^{(s,l)}, n_2+\delta_2^{(s,l)}}$ is then obtained from bilinear interpolation with nearest neighbors at integer spatial coordinates. Eq. (3) and Eq. (4) illustrate that f_{n_1,n_2} models the accumulated phase difference along the propagation path from the guide star to the LGS WFS pupil; this ray-trace is the only process introducing cross-coupling between atmospheric layers. Note that Eq. (1) is for an "open-loop" AO configuration, i.e. the DM correction is not seen in the WFS measurements; nevertheless, if the DM response is perfectly known (i.e. hysteresis and nonlinearities are perfectly known and can be calibrated out), then Eq. (1) describes also a "closed-loop" configuration, as the DM correction can be removed from the closed loop WFS measurements in post-processing to yield "pseudo open-loop" measurements [22, 23]. Note also that for LGS WFSs, tip/tilt (TT) needs to be removed from the measurement vector of each LGS WFS. Although this operation is low-rank, it cross-couples all measurement grid points for each LGS WFS, both in the SD and FD. For this reason, it is omitted from the state space model in Eq. (1), which for LGS WFSs is an approximation.

A full representation of the state transition matrix a_{n_1,n_2} in Eq. (1) has been proposed in [24, 25], but it appears to be prohibitively expensive to compute and implement at run time on ELT scale tomography problems. As explained in [26] and [13], in the limit of an infinite telescope aperture, i.e. $D \to \infty$, where D denotes the telescope aperture diameter, the system in Eq. (1) with periodic (wrap-around) boundary conditions becomes spatially invariant and therefore amenable to a (block) diagonal Fourier decomposition. This means that if state, input and output variables are replaced by their Discrete Fourier Transform (DFT), then for each spatial frequency, the system consists of n_L cross-coupled equations (cross-coupling arising from ray-tracing from the guide stars to the pupil). If we indicate Fourier transformed variables by capital letters, the Fourier shift theorem yields:

$$\phi^{(l)}_{n_1+\alpha_1, n_2+\alpha_2} \overset{\text{DFT}}{\longleftrightarrow} X_1^{\alpha_1} X_2^{\alpha_2} \Phi^{(l)}_{n_1,n_2} \quad \textbf{(5)}$$

where $\alpha_{1,2}$ are arbitrary real numbers, and $X_{1,2} = \exp(-2j\pi n_{1,2}/n) \in \mathbb{C}$ are complex phasors ($j = \sqrt{-1}$).

Thus, in FD, the state space model Eq. (1) is expressed as follows:

$$\begin{cases} \Phi_{n_1,n_2}(k+1) = A_{n_1,n_2}\Phi_{n_1,n_2}(k) + V_{n_1,n_2}(k) \\ Y_{n_1,n_2}(k) = F_{n_1,n_2}(\Phi_{n_1,n_2}(k)) + W_{n_1,n_2}(k) \end{cases} \tag{6}$$

where $A_{n_1,n_2} \in \mathbb{C}^{n_L \times n_L}$ is the Fourier representation of the state transition matrix a_{n_1,n_2} in Eq. (1), Y, V and W are the Fourier transforms of y, v and w, $\Phi_{n_1,n_2} \in \mathbb{C}^{n_L}$ and F_{n_1,n_2} is the concatenated function with component $s = 1, \ldots, n_{gs}$ defined as follows:

$$F^{(s)}_{n_1,n_2}(\Phi_{n_1,n_2}) = \frac{1}{2}\sum_{l=1}^{n_L} \begin{bmatrix} X_1 + X_1X_2 - 1 - X_2 \\ X_2 + X_1X_2 - 1 - X_1 \end{bmatrix} X_1^{\delta_1^{(s,l)}} X_2^{\delta_2^{(s,l)}} \Phi^{(l)}_{n_1,n_2} \tag{7}$$

Since F_{n_1,n_2} is a linear function of the phase spatial spectrum, by rearranging the above expression we can write:

$$F_{n_1,n_2}(\Phi_{n_1,n_2}(k)) = C_{n_1,n_2}\Phi_{n_1,n_2}(k) \tag{8}$$

with $C_{n_1,n_2} \in \mathbb{C}^{2n_{gs} \times n_L}$. Note that each aperture-plane spatial frequency coordinate n_1, n_2 in Eq. (6) is cross-coupled to the same spatial frequency coordinate n_1, n_2 at layer l. The $n_L \times n_L$ covariance of V_{n_1,n_2} is diagonal and given by:

$$(\Sigma_V)_{n_1,n_2} = (I_{n_L \times n_L} - A_{n_1,n_2}A^H_{n_1,n_2})(\Sigma_\Phi)_{n_1,n_2} \tag{9}$$

where superscript H indicates the Hermitian operator (transpose and complex conjugate of a matrix or vector), I is the identity matrix, $(\Sigma_\Phi)_{n_1,n_2} \in \mathbb{R}^{n_L \times n_L}$ is the diagonal Von Kármán power spectral density (PSD) matrix whose entries are:

$$[(\Sigma_\Phi)_{n_1,n_2}]_{l,l'} = 0.023\delta_{l,l'}c_l r_{0,l}^{-\frac{5}{3}} \left(\frac{n_1^2 + n_2^2}{(n\xi^{(l)}\Delta x^{(1)})^2} + \frac{1}{L_{0,l}^2} \right)^{-\frac{11}{6}} \tag{10}$$

where $\delta_{l,l'}$ is the Kronecker symbol, c_l is the weight of layer l, $r_{0,l}$ is the Fried parameter for layer l, and $L_{0,l}$ is the outer scale for layer l. In Fourier space, the covariance matrix of the measurement noise spatial spectrum W_{n_1,n_2} is modeled as a constant, σ_w^2, multiple of the identity matrix, i.e. $(\Sigma_W)_{n_1,n_2} = \sigma_w^2 I_{2n_{gs} \times 2n_{gs}}$, which for LGS WFSs is an approximation on account of sub-aperture spot elongation.

Thus, the DFT transforms the global state-space model into a block-diagonal system, with blocks of size $n_L \times n_L$ decoupled from one another. The steady-state Kalman filter gain is obtained by solving for each spatial frequency coordinate (n_1, n_2) the Discrete Algebraic Riccati Equation (DARE), which corresponds for the Fourier model Eq. (6) to frequency-depending matrix equations:

$$P_{n_1,n_2} = A_{n_1,n_2}P_{n_1,n_2}A^H_{n_1,n_2} + \Sigma_{n_1,n_2} - K_{n_1,n_2}C_{n_1,n_2}P_{n_1,n_2}A^H_{n_1,n_2} \tag{11}$$

where P_{n_1,n_2} is the $n_L \times n_L$ phase estimation error covariance matrix Fourier representation to be solved for numerically, and

$$K_{n_1,n_2} = A_{n_1,n_2}P_{n_1,n_2}C^H_{n_1,n_2}(C_{n_1,n_2}P_{n_1,n_2}C^H_{n_1,n_2} + \sigma_w^2 I)^{-1} \tag{12}$$

denotes the frequency-dependent Kalman gain, $K_{n_1,n_2} \in \mathbb{C}^{n_L \times n_L}$. Solving for the set of frequency-dependent DAREs is computationally inexpensive compared to the SD DARE, which cross-couples all phase grid points. This is due to the fact that the cost for solving Riccati equations grows with the third power of system size; therefore, solving many low-order Riccati equations in the FD is much more efficient than solving a single high-order DARE in the SD. This is the main advantage of using the distributed approach for computing the Kalman filter. Moreover, for large-scale ill-conditioned systems (like in MCAO or MOAO), solving DAREs in the SD, besides being prohibitively expensive, can stagnate and never fully converge.

At run time, the Kalman filter can be implemented in the FD as follows:

$$\begin{cases} \hat{\Phi}_{n_1,n_2}(k+1|k) = A_{n_1,n_2}\hat{\Phi}_{n_1,n_2}(k|k-1) + K_{n_1,n_2}E_{n_1,n_2}(k|k-1) \\ \hat{Y}_{n_1,n_2}(k|k-1) = C_{n_1,n_2}\hat{\Phi}_{n_1,n_2}(k|k-1) \\ E(k|k-1) = \Pi_{TT}\left(Y(k) - \hat{Y}(k|k-1)\right) \end{cases} \tag{13}$$

where the hat symbols indicate estimates of a variable as a function of past measurements, e.g. $\hat{\Phi}_{n_1,n_2}(k|k-1)$ is the estimate of $\Phi_{n_1,n_2}(k)$ based on all measurements up to time step $k-1$; E is the FD tip/tilt removed (TTR) innovation vector, and Π_{TT} is the FD tip/tilt removal operator. Again, as for ϕ, symbols with subscript (n_1, n_2) indicate vectors of a vertical slice of the physical quantity at coordinate n_1, n_2, whereas symbols without subscript indicate stacks of the aforementioned slice vectors for all valid coordinates. As seen before, any spatially-invariant operator in the SD corresponds to a block diagonal operator in the FD. The converse is true as well, so the FD block diagonal Kalman gain corresponds to a spatially-invariant operator in the SD, i.e. an operator which depends only on the relative position between grid points.

The Kalman filter at run time can also be implemented in the SD as follows:

$$\begin{cases} \hat{\phi}_{n_1,n_2}(k+1|k) = a_{n_1,n_2}(\hat{\phi}(k|k-1)) + \\ \qquad\qquad \sum_{n_1',n_2'=-n/2}^{n/2-1} k_{n_1-n_1',n_2-n_2'}\, e_{n_1',n_2'}(k|k-1) \\ \hat{y}_{n_1,n_2}(k|k-1) = f_{n_1,n_2}(\hat{\phi}(k|k-1)) \\ e(k|k-1) = \pi_{TT}\left(y(k) - \hat{y}(k|k-1)\right) \end{cases} \tag{14}$$

where e is the TTR SD innovation vector, and π_{TT} is the SD tip/tilt removal operator. At run time, the filter can thus be implemented either in the SD with Eq. (14) or in the FD with Eq. (13). According to the real time controller hardware architecture, one or the other implementation might be more advantageous; the FD implementation requires fewer computations [14], whereas the SD implementation may be more straightforward to parallelize. Note also that for LGS AO, in addition to removing tip/tilt from pseudo open loop measurements and measurements estimates, low-order modes (tip, tilt, focus, plate scale) are removed at the output of the DM fitting step that follows tomographic wavefront estimation, in an architecture known as split tomography [27]. These low-order modes are separately estimated by a low-order loop driven by a few low-order (tip/tip, focus) natural guide star (NGS) WFSs.

Note that the TT removal operators Π_{TT} and π_{TT} are required for LGS WFSs. As a result, the Kalman filter model used at run time is slightly different from the model in Eq. (1) that has been used to generate it. Note also that the TT removal operator acts on the LGS WFS measurement space, not on the phase space which can be of much higher dimensionality.

3. DISTRIBUTED KALMAN FILTER INTEGRATING A KNOWLEDGE OF THE WIND VELOCITY

Denoting the wind velocity components at layer l by $\omega_1^{(l)}$ and $\omega_2^{(l)}$ in units of mesh size per time step ($\Delta x^{(l)}/\Delta t$), the FD model in Eq. (6) becomes:

$$\begin{cases} \Phi_{n_1,n_2}(k+1) = a\,\Omega_{n_1,n_2}\Phi_{n_1,n_2}(k) + V_{n_1,n_2}(k) \\ Y_{n_1,n_2}(k) = F_{n_1,n_2}(\Phi_{n_1,n_2}(k))W_{n_1,n_2}(k) \end{cases} \tag{15}$$

where

$$a = \begin{bmatrix} a^{(1)} & 0 & \dots & 0 \\ 0 & a^{(2)} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a^{(n_L)} \end{bmatrix} \tag{16}$$

with $0 < a^{(l)} < 1$; Ω_{n_1,n_2} is a diagonal shift matrix given by

$$\Omega_{n_1,n_2} = \begin{bmatrix} X_1^{\omega_1^{(1)}} X_2^{\omega_2^{(1)}} & 0 & \dots & 0 \\ 0 & X_1^{\omega_1^{(2)}} X_2^{\omega_2^{(2)}} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & X_1^{\omega_1^{(n_L)}} X_2^{\omega_2^{(n_L)}} \end{bmatrix} \tag{17}$$

As it is commonly done in AR state-space models for turbulence, the coefficients $a^{(l)}$ are a damping or forgetting factor which account for the speed at which a phase screen can change while translating. A value of 1 would indicate a frozen flow, values close to 1 mean very slow changes, values closer to zero mean faster changes. Eq. (13) then becomes:

$$\begin{cases} \hat{\Phi}_{n_1,n_2}(k+1|k) = a\Omega_{n_1,n_2}\hat{\Phi}_{n_1,n_2}(k|k-1) + K_{n_1,n_2}E_{n_1,n_2}(k|k-1) \\ \hat{Y}_{n_1,n_2}(k|k-1) = C_{n_1,n_2}\hat{\Phi}_{n_1,n_2}(k|k-1) \\ E(k|k-1) = \Pi_{TT}\left(Y(k) - \hat{Y}(k|k-1)\right) \end{cases} \tag{18}$$

Defining $\hat{\Phi}_{n_1,n_2}(k+1|k-1) = a\Omega_{n_1,n_2}\hat{\Phi}_{n_1,n_2}(k|k-1)$, we have in the SD:

$$\hat{\phi}^{(l)}_{n_1,n_2}(k+1|k-1) = a^{(l)}\hat{\phi}^{(l)}_{n_1-\omega_1^{(l)},n_2-\omega_2^{(l)}}(k|k-1) \tag{19}$$

As illustrated in Fig. 1, $\hat{\phi}^{(l)}_{n_1-\omega_1^{(l)},n_2-\omega_2^{(l)}}$ can be approximated by bilinear interpolation as follows:

$$\begin{aligned} \hat{\phi}^{(l)}_{n_1-\omega_1^{(l)},n_2-\omega_2^{(l)}} = {} & (1-|\omega_1^{(l)}|)(1-|\omega_2^{(l)}|)\hat{\phi}^{(l)}_{n_1,n_2} + \\ & |\omega_1^{(l)}|(1-|\omega_2^{(l)}|)\hat{\phi}^{(l)}_{n_1-\mathrm{sgn}(\omega_1^{(l)}),n_2} + \\ & |\omega_2^{(l)}|(1-|\omega_1^{(l)}|)\hat{\phi}^{(l)}_{n_1,n_2-\mathrm{sgn}(\omega_2^{(l)})} + \\ & |\omega_1^{(l)}||\omega_2^{(l)}|\hat{\phi}^{(l)}_{n_1-\mathrm{sgn}(\omega_1^{(l)}),n_2-\mathrm{sgn}(\omega_2^{(l)})} \end{aligned} \tag{20}$$

Thus, in the SD, the Kalman filter incorporating a wind knowledge can be expressed as in Eq. (21), where $\hat{\phi}_{n_1,n_2} \in \mathbb{R}^{n_L}$ is the concatenation of $\hat{\phi}^{(l)}_{n_1,n_2}$ for $l = 1, \cdots, n_L$, and $k_{m_1,m_2} \in \mathbb{R}^{n_L \times n_L}$.

4. WIND VELOCITY ESTIMATION

Integrating wind velocity knowledge into the Kalman filter has been shown to greatly improve performance [14, 28]. Estimating wind velocity consists in estimating the interpolation coefficients which propagate a phase screen in time. Namely, if we define the nine displacements $\mathcal{D} = \{(0,0), (1,0), (-1,0), (0,1), (0,-1), (1,1), (-1,-1), (-1,1), (1,-1)\}$ corresponding to 1 pixel shifts in the horizontal and/or vertical directions (or no shift), the interpolation is given by

$$\hat{\phi}^{(l)}_{n_1,n_2}(k+1-j|k-1-j) \approx a^{(l)} \sum_{d\in\mathcal{D}} \alpha_d^{(l)}\, \hat{\phi}^{(l)}_{(n_1,n_2)+d}(k-j|k-1-j) \tag{22}$$

for $j \geqslant 0$. This means that we can use past phase estimates in order to estimate the α_d coefficients, using a numerical algorithm operating on $N \geqslant 1$ samples of recently reconstructed phase screens. Notice that the nine coefficients α_d are not independent from one another: under the hypothesis of (damped) frozen flow, five of them will be equal to 0 (those corresponding to pixels which are along the direction of the wind, see again Fig. 1), and the other four will depend on the two components of the wind velocity according to Eq. (20).

The effect of wind velocity can be small from one time step to the consecutive one, meaning that the coefficients for $\alpha_{(0,0)}$ will be much bigger than the others (i.e. the phase screen does not translate much in only a single time step). For this reason, Eq. (20) can be replaced by

$$\begin{aligned} & \hat{\phi}^{(l)}_{n_1,n_2}(k+1-j|k-1-j) \\ & \approx a^{(l)} \textstyle\sum_{d\in\mathcal{D}} \alpha_d^{(l)}\, \hat{\phi}^{(l)}_{(n_1,n_2)+d}(k-j+1-n_d|k-n_d-j) \end{aligned} \tag{23}$$

in which the interpolation is based on $n_d \geqslant 1$ time steps, which will make the translation effect of the wind more visible with respect to the noise of the measurement and estimation errors. Note that the non-linear dependence between the coefficients α_d and the components of the wind velocity makes finding an estimate of such components non trivial. For this reason, as first step, we propose an approach based on an approximation which will lead to a solution through least squares.

Note that in Eq. (20) the only non-zero coefficient involving two displacements at the same time, i.e. one among $\alpha_{(1,1)}$, $\alpha_{(1,-1)}$, $\alpha_{(-1,1)}$ and $\alpha_{(-1,-1)}$, will have a quadratic dependence on wind velocity, namely $|\omega_1^{(l)}||\omega_1^{(l)}|$ (or $|\omega_1^{(l)}||\omega_1^{(l)}|n_d^2$ if $n_d > 1$). As typical winds velocities do not exceed ~ 35 m/s, for a typical 0.5 m subaperture size and a sampling rate of 800 Hz, we have $|\omega_1^{(l)}|, |\omega_2^{(l)}| \leqslant 0.08 \ll 1$, which makes quadratic terms smaller with respect to the first-order ones, if n_d is restricted to values smaller than 5 for example.

The problem of estimating wind velocity can then be cast as a least squares problem as in Eq. (24), in which we limit the search to five coefficients $\alpha_d^{(l)}$ with $d \in \mathcal{D}' = \{(0,0), (1,0), (-1,0), (0,1), (0,-1)\}$. The minimization problem features the constraint $\alpha_d^{(l)} \geqslant 0$ due to the fact that, as seen earlier, such coefficients are either zero or equal to the absolute value of one component of the wind velocity times n_d. To reduce computational complexity, N should be chosen small ($N \sim 2-10$) and the estimation can be done locally, e.g. over a small set $\Pi^{(l)}$ of $n_0 \times n_0$ grid points sampling a specific layer region.

This least squares problem is similar to the one proposed in [28], the difference being that we do not cast it on priors, but

$$\begin{cases} \hat{\phi}^{(l)}_{n_1,n_2}(k+1|k-1) = a^{(l)}(1-|\omega_1^{(l)}|-|\omega_2^{(l)}|+|\omega_1^{(l)}||\omega_2^{(l)}|)\hat{\phi}^{(l)}_{n_1,n_2}(k|k-1) + a^{(l)}(|\omega_1^{(l)}|-|\omega_1^{(l)}||\omega_2^{(l)}|)\hat{\phi}^{(l)}_{n_1-\mathrm{sgn}(\omega_1^{(l)}),n_2}(k|k-1) \\ \qquad\qquad + a^{(l)}(|\omega_2^{(l)}|-|\omega_1^{(l)}||\omega_2^{(l)}|)\hat{\phi}^{(l)}_{n_1,n_2-\mathrm{sgn}(\omega_2^{(l)})}(k|k-1) + a^{(l)}|\omega_1^{(l)}||\omega_2^{(l)}|\hat{\phi}^{(l)}_{n_1-\mathrm{sgn}(\omega_1^{(l)}),n_2-\mathrm{sgn}(\omega_2^{(l)})}(k|k-1) \\ \hat{\phi}_{n_1,n_2}(k+1|k) = \hat{\phi}^{(l)}_{n_1,n_2}(k+1|k-1) + \sum_{n_1',n_2'=-n/2}^{n/2-1} k_{n_1-n_1',n_2-n_2'} e_{n_1',n_2'}(k|k-1) \\ \hat{y}_{n_1,n_2}(k|k-1) = f_{n_1,n_2}(\hat{\phi}_{n_1,n_2}(k|k-1)) \\ e(k|k-1) = \pi_{TT}\left(y(k)-\hat{y}(k|k-1)\right) \end{cases} \tag{21}$$

$$\min_{\alpha_d \geqslant 0} \sum_{(n_1,n_2)\in\Pi^{(l)}} \sum_{1\geqslant j\geqslant N} \left|\left| \hat{\phi}^{(l)}_{n_1,n_2}(k+1-jn_d|k-1-jn_d) - a^{(l)} \sum_{d\in\mathcal{D}'} \alpha_d^{(l)}\, \hat{\phi}^{(l)}_{(n_1,n_2)+d}(k-jn_d+1-n_d|k-jn_d-n_d) \right|\right|^2 \tag{24}$$

on the data generated by the Kalman filter itself. Such a constrained least squares problem can be solved with 2 iterations of unconstrained least squares, the first iteration computing the signs of the coefficients, the second computing the magnitudes. In more details, rewriting Eq. (24) in matrix form, we have:

$$\Psi^{(l)} = \underbrace{a^{(l)}\begin{bmatrix} \Psi^{(l)}_{(0,0)} & \Psi^{(l)}_{(1,0)} & \Psi^{(l)}_{(-1,0)} & \Psi^{(l)}_{(0,1)} & \Psi^{(l)}_{(0,-1)} \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} \alpha^{(l)}_{(0,0)} \\ \alpha^{(l)}_{(1,0)} \\ \alpha^{(l)}_{(-1,0)} \\ \alpha^{(l)}_{(0,1)} \\ \alpha^{(l)}_{(0,-1)} \end{bmatrix}}_{p} \tag{25}$$

where:

$$\Psi^{(l)} = \begin{bmatrix} \mathrm{vec}\{\hat{\phi}^{(l)}_{(n_1,n_2)}(k|k-1)\} \\ \mathrm{vec}\{\hat{\phi}^{(l)}_{(n_1,n_2)}(k-n_d|k-n_d-1)\} \\ \vdots \\ \mathrm{vec}\{\hat{\phi}^{(l)}_{(n_1,n_2)}(k-Nn_d+n_d|k-Nn_d+n_d-1)\} \end{bmatrix}, \tag{26}$$

$$\Psi^{(l)}_{(d_1,d_2)} = \begin{bmatrix} \mathrm{vec}\{\hat{\phi}^{(l)}_{(n_1+d_1,n_2+d_2)}(k-n_d|k-n_d-1)\} \\ \mathrm{vec}\{\hat{\phi}^{(l)}_{(n_1+d_1,n_2+d_2)}(k-2n_d|k-2n_d-1)\} \\ \vdots \\ \mathrm{vec}\{\hat{\phi}^{(l)}_{(n_1+d_1,n_2+d_2)}(k-Nn_d|k-Nn_d-1)\} \end{bmatrix}, \tag{27}$$

where $(d_1,d_2)\in\mathcal{D}$, and the operator vec{ } indicates the vector obtained by concatenating all elements satisfying $(n_1,n_2)\in\Pi^{(l)}$.

The least-squares solution of Eq. (25) is given by $p = M^\dagger\,\Psi^{(l)}$ where $M^\dagger = (M^TM)^{-1}M^T$ denotes the pseudoinverse of M. The pairs $\alpha^{(l)}_{(1,0)}$ and $\alpha^{(l)}_{(-1,0)}$ or $\alpha^{(l)}_{(0,1)}$ and $\alpha^{(l)}_{(0,-1)}$ will in general each have one non-zero, significant, positive value and the other a near zero, positive or negative value: this second one has to be considered as a noise artifact as there can be only one non-zero coefficient per axis depending on the sign of the wind velocity. A positive value indicates the presence of a fit of the data for the corresponding wind direction. A reduced rank matrix M is then constructed by keeping only the columns of M corresponding those positive values. The overall algorithm thus consists of two sequential least-squares fits on each tomography layer, the first to estimate the sign of each component of the velocity vector, the second to estimate the magnitude of each component. The detailed procedure is given in Algorithm 1.

Algorithm 1. Wind estimation

For turbulence layer l at time k, execute the following.

1: Assemble M and Ψ at time k for layer l according to Eq. (25)
2: $p = M^\dagger\Psi$
3: **if** $\alpha^{(l)}_{(1,0)} > \alpha^{(l)}_{(-1,0)}$ **then**
4: remove column $\Psi^{(l)}_{(-1,0)}$ from M
5: $s_{\omega_1} = +1$
6: **else**
7: remove column $\Psi^{(l)}_{(1,0)}$ from M
8: $s_{\omega_1} = -1$
9: **if** $\alpha^{(l)}_{(0,1)} > \alpha^{(l)}_{(0,-1)}$ **then**
10: remove column $\Psi^{(l)}_{(0,-1)}$ from M
11: $s_{\omega_2} = +1$
12: **else**
13: remove column $\Psi^{(l)}_{(0,1)}$ from M
14: $s_{\omega_2} = -1$
15: $p' = [\beta\ \gamma_1\ \gamma_2]^T = M^\dagger\Psi$ (At this point, M has only 3 columns)
16: The velocity estimate is $\hat{\omega}_1^{(l)} = s_{\omega_1}\gamma_1/n_d$, $\hat{\omega}_2^{(l)} = s_{\omega_2}\gamma_2/n_d$

The values of $\hat{\omega}_1^{(l)}$ and $\hat{\omega}_2^{(l)}$ obtained from the least squares problem can be further improved by going back to the original optimization problem and taking into account all terms including the quadratic ones, as in Eq. (28).

If we define a non-linear cost function J in Eq. (28), Newton iterations [29] can be used to minimize J with respect to the components of the wind velocity, using $\hat{\omega}_1^{(l)}$ and $\hat{\omega}_2^{(l)}$ obtained above as initial guess:

$$\begin{bmatrix} \hat{\omega}_1^{(l)} \\ \hat{\omega}_2^{(l)} \end{bmatrix} \leftarrow \begin{bmatrix} \hat{\omega}_1^{(l)} \\ \hat{\omega}_2^{(l)} \end{bmatrix} - (\mathbf{H}J)^{-1}\,\nabla J \tag{29}$$

where ∇J and $\mathbf{H}J$ are the Jacobian and Hessian of J evaluated at the current value of $\hat{\omega}_1^{(l)}$ and $\hat{\omega}_2^{(l)}$; these matrices can be computed analytically from Eq. (28) (details not shown for brevity).

$$
\begin{aligned}
J(\omega_1^{(l)},\omega_2^{(l)}) = \textstyle\sum_{(n_1,n_2)\in\Pi^{(l)}} \sum_{1\geqslant j\geqslant N} \Big|\Big| & \hat{\phi}^{(l)}_{n_1,n_2}(k+1-jn_d|k-1-jn_d) \\
-a^{(l)}(1-n_d|\omega_1^{(l)}|-n_d|\omega_2^{(l)}|+n_d^2|\omega_1^{(l)}||\omega_2^{(l)}|)&\hat{\phi}^{(l)}_{n_1,n_2}(k-jn_d+1-n_d|k-jn_d-n_d) \\
+a^{(l)}(1-n_d|\omega_1^{(l)}|-||+n_d^2|\omega_1^{(l)}||\omega_2^{(l)}|)&\hat{\phi}^{(l)}_{n_1-\mathrm{sgn}(\omega_1^{(l)}),n_2}(k-jn_d+1-n_d|k-jn_d-n_d) \\
+a^{(l)}(1-n_d|\omega_2^{(l)}|-||+n_d^2|\omega_1^{(l)}||\omega_2^{(l)}|)&\hat{\phi}^{(l)}_{n_1,n_2-\mathrm{sgn}(\omega_2^{(l)})}(k-jn_d+1-n_d|k-jn_d-n_d) \\
+\ a^{(l)}n_d^2|\omega_1^{(l)}||\omega_2^{(l)}|&\hat{\phi}^{(l)}_{n_1-\mathrm{sgn}(\omega_1^{(l)}),n_2-\mathrm{sgn}(\omega_2^{(l)})}(k-jn_d+1-n_d|k-jn_d-n_d)\Big|\Big|^2
\end{aligned}
\tag{28}
$$

The arrow operator means that terms on the left hand side have to be replaced with those on the right hand side.

In practice, we have verified that one or two Newton iterations are sufficient to ensure convergence and enhanced robustness of the algorithm (thanks to the good starting point provided by Algorithm 1).

The final estimate $(\hat{\omega}_1^{(l)},\hat{\omega}_2^{(l)})$ should be low-pass filtered to provide a temporally smooth wind velocity estimate. The algorithm shows little sensitivity to tuning of n_d and N. In this paper, we used $n_d = 4$, $N = 4$, and $n_i = 2$ Newton iterations. Note that the above wind profiler operates in the SD and not in the FD, which is a difference compared to other algorithms operating in the FD discussed in the literature, e.g. [18, 30]. We can also remark that the above algorithm is not specific to the Kalman filter, but can be coupled to any tomographic wavefront reconstruction algorithm providing an estimate of the phase at all layers. Finally, atmospheric layer strengths and velocities can also be adaptively estimated using SLOpe Detection And Ranging (SLODAR) [15, 31]. Comparison with a previously published SLODAR wind profiler [15] will be discussed in Section 8.

5. ADAPTIVE DISTRIBUTED KALMAN FILTER

Integrating wind estimation, obtained at run-time, into the distributed Kalman filter turns the algorithm into an adaptive filter [32]. This would imply updating the Kalman gain in Eq. (18) or Eq. (21) and the interpolation coefficients in Eq. (21). The update can be performed at regular intervals to track changes in wind velocity. In fact, we have verified that updating the prediction step (which is computationally inexpensive) without updating the Kalman gain (which is the most computationally intensive step), provides no loss in performance improvement. Algorithm 2 summarizes the adaptive Kalman filter with a regular update of the estimation step every k_u time steps and a first-order low-pass filter with parameter $0 < q < 1$ (steps 5 and 6).

6. STABILITY ANALYSIS

Stability issues may arise when using a different state transition matrix for the prediction step than the one used to compute the Kalman gain. In fact, the Kalman filter is guaranteed to be stable only if both matrices are identical. We can prove that if both matrices do not deviate "too much" (more on this later), the system will still be stable, as a result of the properties of the solution of the Riccati equation. We can assume that the Kalman filter has 60 degrees of phase margin [33] (this is strictly correct only for the continuous-time Kalman filter; in our case this limit is approached at high temporal sampling rate). Assuming

Algorithm 2. Adaptive distributed Kalman filter - ADKF

Initialization:

- set $k = 0$
- set $\hat{\phi}_{n_1,n_2}(k+1|k-1) = 0$, $\omega_1^{(l)} = \omega_2^{(l)} = 0$, $\tilde{\omega}_1^{(l)}(0) = \tilde{\omega}_2^{(l)}(0) = 0$
- compute K_{n_1,n_2} for a zero wind speed.

At run time, execute the following.

1: Compute $\hat{\phi}_{n_1,n_2}(k+1|k-1)$ with Eq. (21), with current values for $\omega_1^{(l)}$ and $\omega_2^{(l)}$.
2: **if** $k > N$ **then**
3: Use Algorithm 1 to compute $\hat{\omega}_1^{(l)}$, $\hat{\omega}_2^{(l)}$
4: Iterate Eq. (29) up to 3 times to optimize $\hat{\omega}_1^{(l)}$, $\hat{\omega}_2^{(l)}$
5: $\tilde{\omega}_1^{(l)}(k+1) = q\tilde{\omega}_1^{(l)}(k) + (1-q)\hat{\omega}_1^{(l)}$
6: $\tilde{\omega}_2^{(l)}(k+1) = q\tilde{\omega}_2^{(l)}(k) + (1-q)\hat{\omega}_2^{(l)}$
7: **if** $(k-k_u) \bmod k_u = 0$ **then**
8: Set $\omega_1^{(l)} = \tilde{\omega}_1^{(l)}$, $\omega_2^{(l)} = \tilde{\omega}_2^{(l)}$
9: Increment k (i.e. $k \leftarrow k+1$).
10: Go to 1.

no TT components, from Eq. (18) we have:

$$
\hat{\Phi}_{n_1,n_2}(k+1|k) = Q\hat{\Phi}_{n_1,n_2}(k|k-1) + K_{n_1,n_2}Y_{n_1,n_2}(k) \tag{30}
$$

$$
Q = a\Omega_{n_1,n_2} - K_{n_1,n_2}C_{n_1,n_2} \tag{31}
$$

Under the simplifying hypothesis of identical wind velocities (ω_1,ω_2) at all layers, $\Omega_{n_1,n_2} = \mathrm{e}^{-2\pi j(\omega_1 n_1+\omega_2 n_2)/n} I_{n_L\times n_L}$. We rewrite $Q = \mathrm{e}^{-2\pi j(\omega_1 n_1+\omega_2 n_2)/n}Q'$ with $Q' = a - \mathrm{e}^{2\pi j(\omega_1 n_1+\omega_2 n_2)/n}K_{n_1,n_2}C_{n_1,n_2}$. Since the modulus of the eigenvalues of Q and Q' is identical, if Q' is stable (i.e. all its eigenvalues are inside the unit circle), then Q is stable. We know that $a - K_{n_1,n_2}C_{n_1,n_2}$ is guaranteed to be stable with approximately 60 degrees of phase margin, which directly requires the perturbation term multiplying $K_{\nu_1,\nu_2}C_{\nu_1,\nu_2}$ to have a phase of at most 60 degrees, i.e.

$$
2\pi|\omega_1 n_1 + \omega_2 n_2|/n < \frac{\pi}{3} \tag{32}
$$

The largest magnitude is reached at Nyquist frequency, i.e. at $|n_1| = |n_2| = n/2$, leading to:

$$
|\omega_1| + |\omega_2| < \frac{1}{3} \tag{33}
$$

In polar coordinates, the left hand side of Eq. (33) is largest for a wind direction of $\pi/4$ rad, i.e. the phase margin rule becomes:

$$
\|\omega\| < \frac{1}{3\sqrt{2}} \tag{34}
$$

where $\|\omega\| = \sqrt{\omega_1^2 + \omega_2^2}$. For a typical 800 Hz frame rate and $\Delta x^{(1)} = 1/2$ m, $\Delta x^{(1)}/\Delta t = 400$ m/s, i.e. Eq. (34) is met for wind speeds $\|\omega\| < 94$ m/s, which is well above typical atmospheric wind speeds. At finer sampling, e.g. $\Delta x^{(1)} = 1/4$ m, the threshold is reduced to $\|\omega\| < 47$ m/s, which is still acceptable. We would like to stress that Eq. (34) does not provide an exact stability requirement due to the underlying simplifying assumptions. As described in [34], exact stability margins can be derived from the singular value decomposition of Q, but this would not result in a practical rule as simple as Eq. (34). Note also that the phase margin approach is conservative and provides a sufficient but not necessary condition for stability, i.e. the filter might be stable even for wind velocities not satisfying Eq. (34).

7. COMPUTATIONAL COST

With the advent of powerful, off-the-shelf, cost-effective, central processing units (CPUs) and graphics processing units (GPUs), there has been a recent trend towards hardware solutions favoring the massively parallel matrix vector multiplication (MVM) solution of the static minimum variance reconstructor (MVR) encapsulating both tomography and DM actuator fitting step into a single reconstruction matrix [35]. The ADKF, although non-iterative, does not belong to the MVM algorithm category, and therefore goes against this trend, as it requires the explicit computation of the tomography step at each time step while the MVR approach directly yields the DM command vector from WFS measurements.

Taking the Thirty Meter Telescope (TMT) LGS MCAO system [8] as a reference, for $n_L = 6$ tomography layers reconstructed at $\Delta x^{(l)} = (\xi^{(l)}/4)$ m from $n_{gs} = 6$ LGS WFSs providing measurements sampled at $\Delta x = 2\Delta x^{(1)} = (1/2)$ m and a $n\Delta x^{(1)} = 60$ m wide aperture-plane grid, the online cost of DKF's FD implementation is approximately 44 millions of multiplication/accumulations (MMACs) per frame, and that of a hybrid sparse/MVM DM fitting step is approximately $5 + N_a^2/10^6$ per frame [14], where N_a denotes the total number of DM actuators, the factor 5 accounts for the sparse ray-trace to the aperture-plane along all fitting directions of interest (typically 9 directions) followed by backward ray-trace from the aperture-plane to all DM actuator grids, and the factor N_a^2 accounts for the fitting MVM. Upgrading DKF to ADKF is computationally inexpensive. For step 2 of Algorithm 1, the cost of assembling $M^T\Psi^{(l)}$ is $5n_0^2N$ MACs (M has a rank equal to 5, and n_0^2 is the number of phase points used for the estimation), the one of assembling M^TM is $15n_0^2N$ MACs, and the one of inverting M^TM is $5^3/3$. For step 15 of Algorithm 1, the cost of assembling $M^T\Psi^{(l)}$ is $3n_0^2N$ MACs (M has a rank equal to 3 at this stage), that of assembling M^TM is $6n_0^2N$ MACs, and that of inverting M^TM is $3^3/3$. The total cost per frame per atmospheric layer is thus $29\,n_0^2\,N + (5^3 + 3^3)/3$ MACs, which for $n_0 = 20$ and $N = 4$ amounts to 0.0465 MMACs. For $n_L = 6$ turbulence layers, the grand total is 0.28 MMACs, which represents only a 0.6 % cost increase from the 44 MMACs reported in [14] without wind profiler. For a Newton iteration, (step 3 of Algorithm 2) the cost is on the order of $28Nn_0^2$ MACs per layer, which for $n_i = 2$ iterations gives a total of 0.5 MMACs, i.e. an additional 1.2 % increment. Recall that these Newton iterations are not stricly necessary, their main role being to provide extra precision to the solution. For LGS MOAO, the grand total online cost becomes $44 + N_f(0.5 + N_a^2/10^6)$ MMACs, where N_a now denotes now the number of actuators per DM, N_f the number of science channels over which DM fitting is performed, and the factor 0.5 accounts for the sparse ray-trace to the aperture-plane along a single fitting direction followed by backward ray-trace from the aperture-plane to a single DM actuator grid. For the MVR solution, the online costs are $N_sN_a/10^6$ and $N_sN_aN_f/10^6$ MMACs for the MCAO and MOAO configurations respectively, where N_s denotes the total number of LGS WFS measurements. For LGS MOAO, in the limit of a large number of fitting directions, we have:

$$\frac{\text{computational cost of MVR}}{\text{computational cost of ADKF+fitting}} \rightarrow \frac{N_s}{N_a} \approx 2n_{gs}, \qquad \textbf{(35)}$$

i.e. the cost ratio asymptotes to approximately twice the number of LGS WFSs used to perform tomography. This asymptotic cost ratio is achieved if $N_fN_a^2/10^6 \gg 44 + 0.5N_f$, which for order 60×60 DMs is met if $N_f \gg 4$. This shows that ADKF has the potential to be significantly cheaper in terms of online computational cost than the MVR solution, which requires separate control matrices for each MOAO fitting direction.

8. SIMULATION RESULTS

We have implemented Algorithm 2 in LAOS, a MATLAB, high-fidelity, end-to-end AO simulation tool, particularly well suited for algorithm development and prototyping [36]. An overview of the code's LGS wave-optics capabilities can be found in [37]. We simulated LGS MCAO observations with TMT at zenith under median turbulence conditions (0.58 arcsec seeing at 500 nm at Mauna Kea) and split tomography with pseudo open-loop control at 800 Hz frame rate [27]. Note that off-zenith simulations would require to rescale the Fried parameter, the distance to the atmospheric layers and to the sodium layer, and the LGS WFS signal level, but the basic principle of the ADKF algorithm would not be affected. The simulated system features 6 LGSs, projected from a 0.4 m diameter laser launch telescope (LLT) located behind the secondary mirror of the telescope, forming a regular 70 arcsec diameter pentagon plus one LGS on-axis. The total number of LGS WFS measurements is 31,000. A LGS WFS signal level of 750 photo-detected electrons (PDEs) per subaperture (0.25 m^2) per frame (1.25 ms) was simulated (expected return for a 25 W sodium guide star laser under median conditions), and a constrained matched filter was used to compute the average wavefront gradient over each LGS WFS subaperture; the LGS WFS camera is a polar coordinate custom-made charge coupled device (CCD) with 16×6 pixels per subaperture on the outermost subapertures and 3 electrons read-out noise per pixel (for the details, see [38]). Wavefront sensing is simulated using wave optics to properly account of spot elongation, whereas wavefront propagation (laser and science beams) through the atmosphere and AO system is simulated using geometric ray-tracing. The simulated AO system has a ground conjugate deformable mirror (DM) with an equivalent 0.5 m actuator spacing at the aperture plane, and a 11.8 km conjugate DM, also with an equivalent 0.5 m inter-actuator pitch at the aperture plane. The upper DM has its actuator grid interlaced by half a subaperture with respect to the ground DM actuator grid in order to reduce the on-axis fitting error, and both DMs have 30% inter-actuator coupling and bicubic influence functions. The total number of DM actuators is 8,000. Performance is assessed in terms of residual LGS mode wavefront error (WFE), i.e. WFE over all modes orthogonal to tip/tilt and

3 tilt anisoplanatism modes consisting of quadratic modes on both DMs with opposite signs so as to yield pure tip/tilt for a point source at the centroid of the sodium layer (90 km for these simulations), averaged over a 34 arcsec ×34 arcsec science field of view (FoV), which corresponds to the field of view of IRIS, TMT's first light diffraction-limited imager/spectrograph [39]. The DM fitting step is common to all algorithms and consists of a least-squares projection of the tomography solution, optimized for the 34 arcsec ×34 arcsec science FoV.

We tested two different situations: an ideal (unrealistic) scenario excluding the turbulence profile quantization error, for which a 6-layer atmospheric turbulence/wind profile is simulated and the same atmospheric profile is estimated ("matched layers"), and a more realistic scenario including the quantization error, for which a high-resolution 65-layer atmospheric profile is simulated with a 300 m altitude sampling and the former 6-layer profile is reconstructed ("mismatched layers"). Each simulated turbulence layer spans 128 m ×128 m and has 1/64 m spatial sampling (8192 × 8192 grid points), fine enough to capture high spatial frequencies, whereas reconstructed layers have 0.25 m, 0.242 m, 0.235 m, 0.228 m, 0.221 m and 0.215 m sampling on account of the cone coordinate system used. Note that all reconstructed layers are oversampled by a factor of 2 with respect to the LGS WFS resolution of 0.5 m, which significantly reduces the spatial aliasing error at the expense of a higher computational cost. The altitude of the reconstructed 6 layers was chosen to match the SLODAR altitudes calculated for the on-axis and top LGS (35 arcsec angular separation on the sky), and are equal to 0, 2.85, 5.5, 8, 10.4 and 12.6 km above ground-level.

Fig.2 displays both turbulence profiles. The 65-layer profile is a median (50% r_0) Mauna Kea profile obtained from generalized SCIDAR during a site testing campaign for the months of October and December 2002 [40]. Values below the telescope elevation (4200 m) and within 800 m of the telescope elevation have been folded back into the ground layer. The 6-layer turbulence profile has been obtained by binning the high-resolution 65-layer turbulence profile using Multi Aperture Scintillation Sensor (MASS) triangular influence functions (see [31] for details). Both profiles are scaled to yield an integrated Fried parameter, r_0, of 17.9 cm at 500 nm wavelength, and they both have an isoplanatic angle, θ_0, of 2.82 arcsec at 500 nm wavelength and a generalized isoplanatic angle, θ_2, of 10.7 arcsec at 500 nm wavelength for DM conjugates at 0 km and 11.8 km altitude above ground-level. Note that 50%, 80% and 90% of the high-resolution turbulence profile is contained within the first 1 km, 5 km, and 10 km respectively. Fig.3 displays both wind profiles. The 65-layer profile is a sample profile obtained by linear interpolation of balloon wind profile measurements for Hilo, Hawaii (starting at 4.2 km elevation above sea level) available from the University of Wyoming weather database [41]. The raw wind data is sampled at 85 layers non-uniformly distributed between 4.2 km and 19 km elevation above sea level. The 6-layer wind profile was obtained in polar coordinates by (i) binning the turbulence-weighted wind profiles to yield the equivalent wind speed profile, and (ii) binning the wind x-velocity (North direction) and y-velocity (East direction) profiles to yield the equivalent wind direction profile. Mathematically, these binning operations can be expressed as follows:

$$v_k^{eq} = [\,\Sigma_l H_{kl} p_l^{sim} (v_l^{sim})^{5/3}\,]/p_k^{eq} = [\,\Sigma_l H_{kl} w_l^{sim} (v_l^{sim})^{5/3}\,]/w_k^{eq}, \quad \textbf{(36)}$$

where $p_l = r_{0,l}^{-5/3}$ is the turbulence strength for layer l, $r_{0,l}$ the Fried parameter for layer l, $w_l = p_l/\Sigma_l p_l$ the relative weight for layer l, $\Sigma_l p_l = r_0^{-5/3}$, v_l the wind speed for layer l, H_{kl} the binning weight mapping layer l onto layer k (see [31] for details), and

$$v_{x,k}^{eq} = v_k^{eq}\cos(\theta_k^{eq}), \quad v_{y,k}^{eq} = v_k^{eq}\sin(\theta_k^{eq}), \quad \textbf{(37)}$$

where

$$\theta_k^{eq} = \text{atan}(\,[\Sigma_l H_{kl} v_{y,l}^{sim}]/[\Sigma_l H_{kl} v_{x,l}^{sim}]\,). \quad \textbf{(38)}$$

The Greenwood frequency for the 65- and 6-layer wind profiles combined with the turbulence profiles of Fig.2 is $f_G = 20.8$ Hz at 500 nm wavelength.

Five different algorithms are compared: (i) the static minimum variance reconstructor (MVR) which does not incorporate wind profile information, (ii) the distributed Kalman filter assuming knowledge of the equivalent 6-layer profile (DKF*), (iii) the adaptive distributed Kalman filter coupled with the SLODAR wind profiler discussed in [15] assuming knowledge of the equivalent 6-layer profile (ADKF-SLO*), (iv) the adaptive distributed Kalman filter coupled with the SLODAR wind profiler estimating both turbulence and wind profiles (ADKF-SLO), and finally (v) the adaptive distributed Kalman filter coupled with the phase wind profiler discussed in Algorithm 2 operating on $N = 4$ temporal samples with $n_d = 4$ frames between successive samples, $n_i = 2$ Newton iterations (ADKF-4/4/2) and low-pass filter (LPF) parameter $q = e^{-T/\tau_c} = 0.95$ (which corresponds to a time constant $\tau_c \sim 20\,T$ where T is the sampling period, i.e. $\tau_c \sim 25$ ms). All Kalman filters use $a^{(l)} = 0.99$ in Eq. (16), and the SLODAR wind profiler uses 18 baselines spanning a narrow 1-subaperture wide stripe cross-coupling measurements from the on-axis and top LGS pair (35 arcsec angular separation on the sky, see Fig.2 in [15]).

A. Matched reconstruction and simulation layers

We start by analyzing the ideal case of matched reconstruction and simulation layers (unrealistic, excludes turbulence profile quantization error). Simulated and reconstructed profiles are the 6-layer profiles displayed in dashed in Fig.2 and Fig.3.

The top panel of Fig.4 shows the time history of the cumulative LGS mode WFE, i.e. the time history of the square root of the LGS mode wavefront variance averaged from time step 100 to the current time and averaged over a 34 arcsec ×34 arcsec FoV. MVR is for the minimum variance reconstructor (which does not use wind velocity information), DKF* is for the DKF algorithm using perfect knowledge of the wind profile (unrealistic), ADKF-SLO* is for ADKF coupled with the SLODAR wind profiler assuming perfect turbulence profile knowledge (unrealistic), ADKF-SLO is for ADKF coupled with the SLODAR wind profiler estimating both wind and turbulence profiles, and ADKF-4/4/2 is for ADKF coupled with the reconstructed phase wind profiler using $N = 4$ temporal samples, $n_d = 4$ frames between successive samples, and $n_i = 2$ Newton iterations described in Algorithm 2. The bottom panel shows the quadratic incremental WFE root-mean-square (RMS) with respect to the MVR WFE. Negatives values indicate lower WFE (improved performance). The SLODAR wind profiler seems to perform slightly better than the phase wind profiler in this sample simulation, and both algorithms converge after a few seconds from cold start. Note also that there is still a ∼ 23 nm RMS performance gap in quadrature between the adaptive solutions and the case of perfect wind profile knowledge, pointing to more work needed to further improve the accuracy of both wind profilers.

The top panel of Fig.5 displays the estimated Greenwood frequency. Very little deviation is seen between the different estimates and the 20.8 Hz simulated value. The bottom panel plots the weighted average over altitude of the wind speed, x-velocity and y-velocity estimates. Here again, little deviation is seen between the different cases.

B. Mismatched reconstruction and simulation layers

We now turn to the more realistic case of a much larger number of simulated layers, which captures the turbulence profile quantization error arising by discretizing the reconstructed profile into a number of layers much smaller than the number of simulated layers. The 65-layer turbulence/wind profiles of Fig.2 and Fig.3 were simulated for this purpose. Reconstruction layers are as in section A.

The top panel of Fig.6 shows the time history of the cumulative LGS mode WFE averaged over 34 arcsec $\times$ 34 arcsec. The salient feature is the significantly degraded performance of DKF* and of all flavors of ADKF, which illustrates that both wind profilers fail to provide accurate velocity estimates when exposed to finely sampled turbulence/wind profiles. The SLODAR wind profiler is also seen to perform worse than the phase profiler for reasons presently not understood. Compared to the MVR solution (which does not incorporate wind profile information), all algorithms are degraded by $40-50$ nm RMS in quadrature, which is very significant and calls for more work to understand and mitigate this performance loss while maintaining computation cost reasonable.

The top panel of Fig.7 displays the estimated Greenwood frequency. Significant deviation is seen between the different estimates and the 20.8 Hz simulated value. The bottom panel plots the weighted average over altitude of the wind speed, x-velocity and y-velocity estimates. Here again, significant deviation is seen between the different cases.

Finally, the turbulence quantization WFE is plotted in Fig.8, which is obtained by taking the quadrature difference between Fig.6 and Fig.4. For this sample simulation, the turbulence profile quantization error is seen to be negligible for the MVR algorithm, but on the order of $40-50$ nm RMS in quadrature for all flavors of ADKF.

9. CONCLUSIONS

We have developed a computationally efficient recursive, non-iterative, adaptive distributed Kalman filter (ADKF) for LGS atmospheric tomography on ELTs. ADKF incorporates a low-cost adaptive least squares wind profiler operating on a small portion of each estimated phase screen to track temporal dynamics. With this module, ADKF is able to reach a lower WFE than the static minimum variance reconstruction (MVR) when simulation and reconstruction layers coincide (unrealistic case), thereby fully compensating for the loss of performance of the original distributed Kalman filter (DKF) arising from the infinite pupil and uniform measurement noise approximations. For the more realistic case where the simulated turbulence/wind profile is much more finely sampled in altitude than the reconstructed profile, performance is significantly degraded. The reasons behind this loss of performance and the higher turbulence profile quantization error compared to the static MVR solution are presently not understood and call for more work.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the TMT partner institutions. They are the Association of Canadian Universities for Research in Astronomy (ACURA), the California Institute of Technology, the University of California, the National Astronomical Observatory of Japan, the National Astronomical Observatories of China and their consortium partners, and the Department of Science and Technology of India and their supported institutes. This work was supported as well by the Gordon and Betty Moore Foundation, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the National Research Council of Canada, the Natural Sciences and Engineering Research Council of Canada, the British Columbia Knowledge Development Fund, the Association of Universities for Research in Astronomy (AURA), and the U.S. National Science Foundation. We also would like to thank Lewis Roberts (JPL) for providing the wind profile data.

REFERENCES

1. F. Roddier, ed., *Adaptive optics in astronomy* (Cambridge University Press, 1999).
2. C. Béchet, M. Tallon, and E. Thiébaut, "Closed-loop AO performance with FrIM," in "Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/Information Photonics/Signal Recovery and Synthesis Topical Meetings," OSA (OSA, Vancouver, Canada, 2007).
3. C. Kulcsár, H.-F. Raynaud, C. Petit, J.-M. Conan, and P. Viaris de Lesegno, "Optimal control, observers and integrators in adaptive optics," Optics Express **14**, 7463–8012 (2006).
4. L. Poyneer, M. van Dam, and J.-P. Véran, "Experimental verification of the frozen flow atmospheric turbulence assumption with use of astronomical adaptive optics telemetry," JOSA A **26**, 833–846 (2009).
5. G. Sivo, C. Kulcsár, J.-M. Conan, H.-F. Raynaud, É. Gendron, A. Basden, F. Vidal, T. Morris, S. Meimon, C. Petit, D. Gratadour, O. Martin, Z. Hubert, A. Sevin, D. Perret, F. Chemla, G. Rousset, N. Dipper, G. Talbot, E. Younger, R. Myers, D. Henry, S. Todd, D. Atkinson, C. Dickson, and A. Longmore, "First on-sky SCAO validation of full LQG control with vibration mitigation on the CANARY pathfinder," Optics Express **22**, 23565–23591 (2014).
6. E. Diolaiti, C. Arcidiacono, B. G., R. Butler, M. Lombini, L. Schreiber, A. Baruffolo, A. Basden, M. Bellazzini, E. Cascone, P. Ciliegi, F. Cortecchia, G. Cosentino, V. De Caprio, A. De Rosa, N. Dipper, S. Esposito, I. Foppiani, E. Giro, G. Morgante, R. Myers, F. Patru, R. Ragazzoni, A. Riccardi, M. Riva, F. Zerbi, M. Casali, B. Delabre, N. Hubin, F. Kerber, M. Le Louarn, E. Marchetti, S. Ramsay, S. Stroebele, and E. Vernet, "Preparing for phase B of the E-ELT MCAO module project," in "SPIE Astronomical Adaptive Optics Systems IV," (International Society for Optics and Photonics, 2014), pp. 91480Y1–91480Y8.
7. A. Bouchez, D. Acton, R. Biasi, R. Conan, B. Espeland, S. Esposito, J. Filgueira, D. Gallieni, B. McLeod, E. Pinna, F. Santoro, G. Trancho, and M. Van Dam, "The Giant Magellan Telescope adaptive optics program," in "SPIE Astronomical Adaptive Optics Systems IV," (International Society for Optics and Photonics, 2014), pp. 91480W1–91480W19.
8. G. Herriot, D. Andersen, J. Atwood, C. Boyer, P. Byrnes, K. Caputa, B. Ellerbroek, L. Gilles, A. Hill, Z. Ljusic, J. Pazder, M. Rosensteiner, M. Smith, P. Spano, K. Szeto, J.-P. Véran, I. Wevers, L. Wang, and R. Wooff, "NFIRAOS first light facility AO system for the Thirty Meter Telescope," in "SPIE Astronomical Adaptive Optics Systems IV," (International Society for Optics and Photonics, 2014), pp. 9148101–9148111.
9. A. Beghi, A. Cenedese, and A. Masiero, "On the computation of Kalman gain in large adaptive optics systems," in "21st Mediterranean Conference on Control & Automation (MED)," (IEEE, 2013), pp. 1374–1379.
10. C. Correia, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, and C. Petit,

"Adapting optimal LQG methods to ELT-sized AO systems," in "Proc. of the 1st Conference on Adaptive Optics for Extremely Large Telescopes," (2010).
11. P. Massioni and M. Di Loreto, "Fast finite-horizon Kalman filter in wavefront estimation for adaptive optics," in "European Control Conference (ECC)," (Strasbourg, France, 2014), pp. 2398–2403.
12. P. Massioni, H.-F. Raynaud, C. Kulcsár, and J.-M. Conan, "An approximation of the Riccati equation in large-scale systems with application to adaptive optics," IEEE Transactions on Control Systems Technology **23**, 479–487 (2015).
13. P. Massioni, C. Kulcsár, H.-F. Raynaud, and J.-M. Conan, "Fast computation of an optimal controller for large-scale adaptive optics," J. Opt. Soc. Am. A **28**, 2298–2309 (2011).
14. L. Gilles, P. Massioni, C. Kulcsár, H.-F. Raynaud, and B. Ellerbroek, "Distributed Kalman filtering compared to Fourier domain preconditioned conjugate gradient for laser guide star tomography on extremely large telescopes," J. Opt. Soc. Am. A **30**, 898–909 (2013).
15. L. Gilles and B. Ellerbroek, "Wind profiling via slope detection and ranging: algorithm formulation and performance analysis for laser guide star tomography on extremely large telescopes," in "Proc. of the 3th Conference on Adaptive Optics for Extremely Large Telescopes," (Florence, Italy, 2013).
16. L. Wang, M. Schöck, and G. Chanan, "Atmospheric turbulence profiling with SLODAR using multiple adaptive optics wavefront sensors," Applied Optics **47**, 1880–1892 (2008).
17. A. Cortes, B. Neichel, A. Guesalaga, J. Osborn, F. Rigaut, and D. Guzman, "Atmospheric turbulence profiling using multiple laser star wavefront sensors," Mon. Not. R. Astron. Soc. **427**, 2089–2099 (2012).
18. L. Poyneer, B. Macintosh, and J.-P. Véran, "Fourier transform wavefront control with adaptive prediction of the atmosphere," J. Opt. Soc. Am. A **24**, 2645–2660 (2007).
19. L. Gilles, P. Massioni, and B. Ellerbroek, "Adaptive distributed Kalman filtering for laser guide star tomography on extremely large telescopes," in "Proc. of Adaptive Optics: Analysis, Methods and Systems," (Optical Society of America, Arlington, USA, 2015).
20. Q. Yang, C. Vogel, and B. Ellerbroek, "Fourier domain preconditioned conjugate gradient algorithm for atmospheric tomography," Appl. Opt. **45**, 5281–5293 (2006).
21. C. Vogel and Q. Yang, "Fast optimal wavefront reconstruction for multiconjugate adaptive optics using the fourier domain preconditioned conjugate gradient algorithm," Opt. Exp. **14**, 7487–7498 (2006).
22. B. Ellerbroek and C. Vogel, "Simulations for closed-loop wavefront reconstruction for multiconjugate adaptive optics on giant telescopes," in "SPIE Astronomical Adaptive Optics Systems and Applications," (International Society for Optics and Photonics, 2003), pp. 206–217.
23. L. Gilles, "Closed-loop stability and performance analysis of least-squares and minimum-variance control algorithms for multiconjugate adaptive optics," Appl. Opt. **44**, 993–1002 (2005).
24. P. Piatrou and M. Roggemann, "Performance study of Kalman filter controller for multiconjugate adaptive optics," Appl. Opt. **46**, 1446–1455 (2007).
25. C. Correia, K. Kackson, J. Véran, D. Andersen, O. Lardière, and B. C., "Static and predictive tomographic reconstruction for wide-field multi-object adaptive optics," J. Opt. Soc. Am. A **31**, 101–113 (2014).
26. B. Bamieh, F. Paganini, and M. Dahleh, "Distributed control of spatially invariant systems," IEEE Transactions on Automatic Control **47**, 1091–1107 (2002).
27. L. Gilles and B. Ellerbroek, "Split atmospheric tomography using laser and natural guide stars," J. Opt. Soc. Am. A **25**, 2427–2435 (2008).
28. K. Jackson, C. Correia, O. Lardière, D. Andersen, and C. Bradley, "Linear prediction of atmospheric wave-fronts for tomographic adaptive optics systems: modelling and robustness assessment," Optics Letters **40**, 143–146 (2008).
29. A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics* (Springer, 2000).
30. L. Poyneer and J.-P. Véran, "Predictive wavefront control for adaptive optics with arbitrary control loop delays," J. Opt. Soc. Am. A **25**, 1486–1496 (2008).
31. L. Gilles and B. Ellerbroek, "Real-time turbulence profiling with a pair of laser guide star Shack-Hartmann wavefront sensors for wide-field adaptive optics systems on large to extremely large telescopes," J. Opt. Soc. Am. A **27**, A76–A83 (2010).
32. K. Åström and B. Wittenmark, *Adaptive control* (Courier Dover Publications, 2013).
33. K. Zhou, J. Doyle, and K. Glover, *Robust and optimal control*, vol. 40 (Prentice Hall New Jersey, 1996).
34. U. Shaked, "Guaranteed stability margins for the discrete-time linear quadratic optimal regulator," IEEE Transactions on Automatic Control **31**, 162–165 (1986).
35. L. Wang and B. Ellerbroek, "Computer simulations and real-time control of ELT AO systems using graphical processing units," in "Proc. SPIE," , vol. 8447 (International Society for Optics and Photonics, 2012), vol. 8447, pp. 844723–844723–11.
36. L. Gilles and B. Ellerbroek, "LAOS: linear adaptive optics simulator," Thirty Meter Telescope project **TMT.AOS.TEC.07.003** (2007).
37. L. Gilles, B. Ellerbroek, and V. J.P., "Laser guide star multiconjugate adaptive optics performance of the thirty meter telescope with elongated beacons and matched filtering," in "SPIE Advances in Adaptive Optics II," (International Society for Optics and Photonics, 2006), pp. 6272361–6272368.
38. L. Gilles, L. Wang, and B. Ellerbroek, "Impact of laser launch location on the performance of laser tomography and multiconjugate adaptive optics for extremely large telescopes," Applied Optics **49**, G114–G119 (2010).
39. A. Moore, J. Larkini, S. Wright, B. Bauman, J. Dunn, B. Ellerbroek, A. Phillips, L. Simard, R. Suzuki, K. Zhang, T. Aliado *et al.*, "The infrared imaging spectrograph (iris) for tmt: Instrument overview," in "SPIE Astronomical Telescopes and Instrumentation 2014: Space and Ground-based Telescopes," (International Society for Optics and Photonics, 2014), pp. 914724–1–914724–12.
40. M. Chun, *University of Hawaii* (Institute for Astronomy, Private Communication, 2009).
41. *University of Wyoming, Atmospheric soundings database* (http://weather.uwyo.edu/upperair/sounding.html).

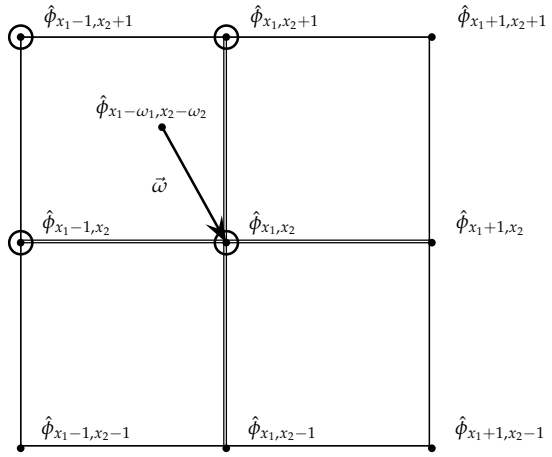

Fig. 1. Interpolation of the wind shift effect on a turbulence layer. Grid points associated to non-zero coefficients are highlighted by a circle, time and layer dependency are omitted.

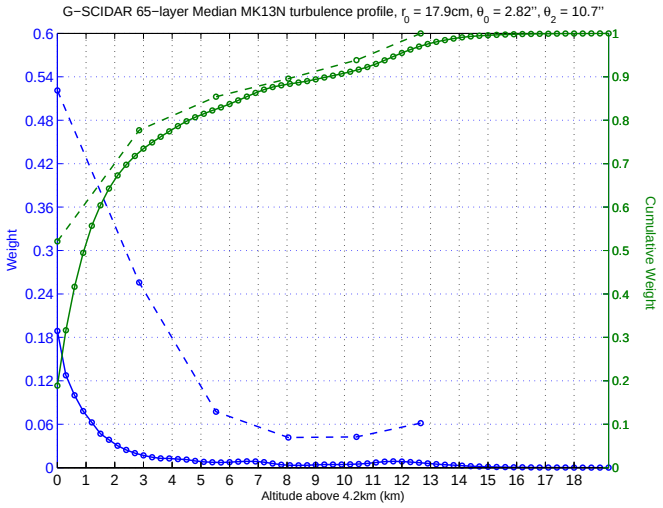

Fig. 2. Left axis: turbulence profile relative weight. Solid blue curve is a generalized SCIDAR 65-layer median ($50\%r_0$) Mauna Kea profile with uniform 300 m altitude sampling. Dashed blue curve is the "equivalent" 6-layer turbulence profile. Right axis: cumulative relative turbulence strength.

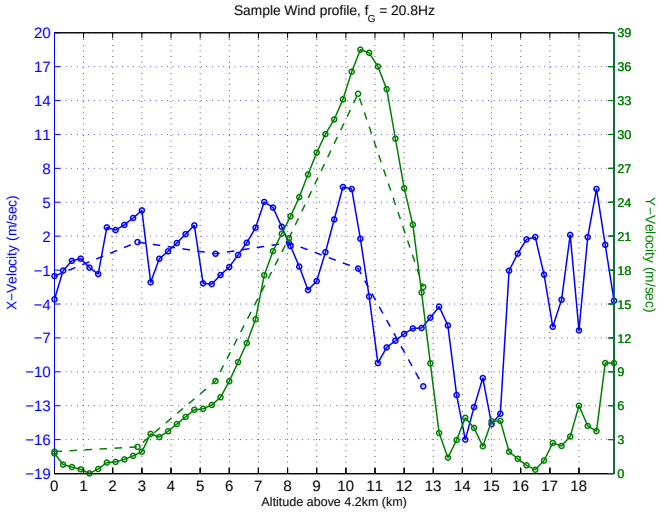

Fig. 3. Left axis: wind x-velocity profile. Solid blue curve is a sample 65-layer profile obtained by linear interpolation of balloon wind profile measurements for Hilo (starting at 4.2 km elevation above sea level) available from the University of Wyoming weather database [41]. Dashed blue curve is an "equivalent" 6-layer wind x-velocity profile. Right axis: same as left axis but for the y-component of the wind velocity vector.

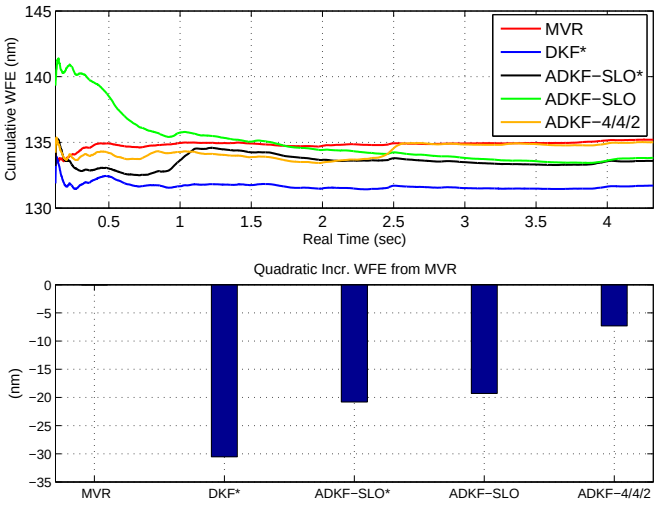

Fig. 4. Top panel: cumulative LGS mode WFE averaged over 34 arcsec $\times$34 arcsec field, matched reconstruction and simulation layers. Turbulence and wind profiles are the 6-layer profiles of Fig.2 and Fig.3 respectively. MVR is for the minimum variance reconstructor (which does not use wind velocity information), DKF* is for the DKF algorithm using perfect knowledge of the wind profile (unrealistic), ADKF-SLO* is for ADKF coupled with the SLODAR wind profiler assuming perfect knowledge of the turbulence profile, ADKF-SLO is for ADKF coupled with the SLODAR wind profiler estimating both wind the turbulence profiles, and ADKF-4/4/2 is for ADKF coupled with the phase profiler using $N = 4$ temporal samples, $n_d = 4$ frames between successive temporal samples, and $n_i = 2$ Newton iterations. Bottom panel: quadratic incremental WFE root-mean-square (RMS) with respect to the MVR WFE. Negatives values indicate lower WFE (improved performance).

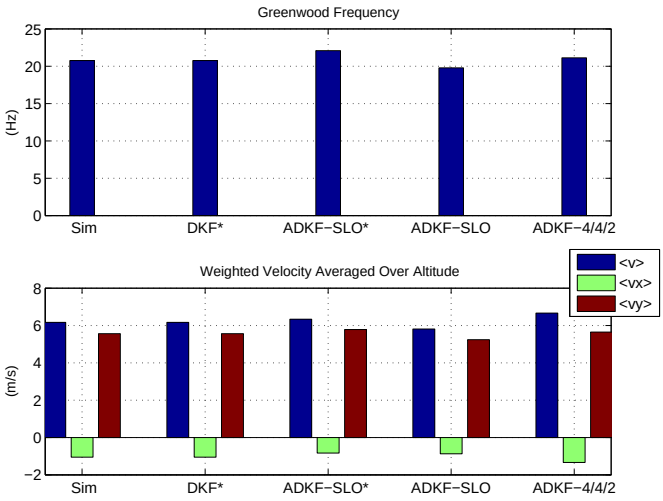

Fig. 5. Top panel: Greenwood frequency, matched reconstruction and simulation layers. Sim is for the simulated turbulence profile, DKF* is for DKF using the equivalent 6-layer wind profile shown in dashed on Fig.3, ADKF-SLO* is for ADKF coupled with the SLODAR wind profiler assuming perfect knowledge of the turbulence profile (unrealistic), ADKF-SLO is for ADKF coupled with the SLODAR wind profiler estimating both wind the turbulence profiles, and ADKF-4/4/2 is for ADKF coupled with the phase profiler using $N = 4$ temporal samples, $n_d = 4$ frames between successive temporal samples, and $n_i = 2$ Newton iterations. Bottom panel: weighted average over altitude of wind speed (blue), x-component of wind velocity (green) and y-component of wind velocity (brown). Weighting given by relative turbulence strength.

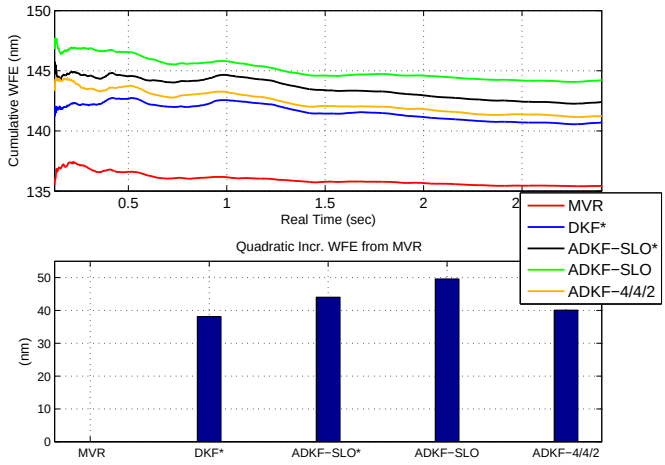

Fig. 6. Top panel: cumulative LGS mode WFE averaged over 34 arcsec $\times$34 arcsec field, mismatched reconstruction and simulation layers (realistic, includes WFE arising from discretizing the vertical turbulence profile into a finite small number of reconstructed layers). Simulated turbulence and wind profiles are the 65-layer profiles of Fig.2 and Fig.3 respectively, reconstructed profiles are the 6-layer profiles. MVR is for the minimum variance reconstructor (which does not use wind velocity information), DKF* is for the DKF algorithm assuming knowledge of the "equivalent" 6-layer wind profile, ADKF-SLO* is for ADKF coupled with the SLODAR wind profiler assuming knowledge of the "equivalent" 6-layer turbulence profile, ADKF-SLO is for ADKF coupled with the SLODAR wind profiler estimating both wind the turbulence profiles, and ADKF-4/4/2 is for ADKF coupled with the phase profiler using $N = 4$ temporal samples, $n_d = 4$ frames between successive temporal samples, and $n_i = 2$ Newton iterations. Bottom panel: quadratic incremental WFE root-mean-square (RMS) with respect to the MVR WFE. Positives values indicate higher WFE (degraded performance).

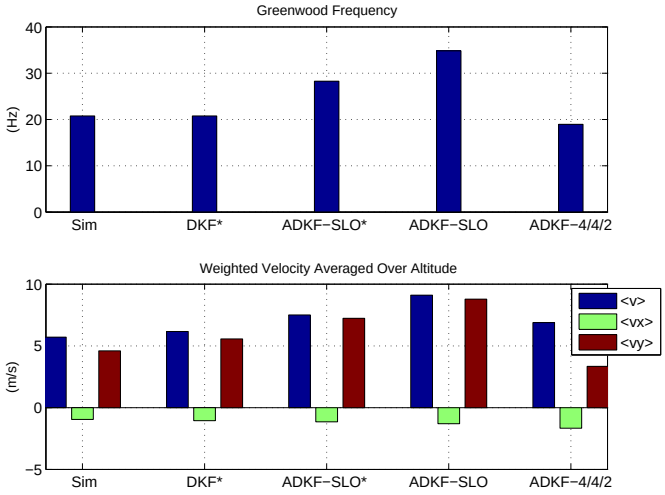

Fig. 7. Top panel: Greenwood frequency, mismatched reconstruction and simulation layers. Sim is for the simulated turbulence profile, DKF* is for DKF using the equivalent 6-layer wind profile (which preserves Greenwood frequency), ADKF-SLO* is for ADKF coupled with the estimate obtained from the SLODAR wind profiler assuming knowledge of the "equivalent" 6-layer turbulence profile, ADKF-SLO is for ADKF coupled with the estimate obtained from the SLODAR wind profiler estimating both wind the turbulence profiles, and ADKF-4/4/2 is for ADKF coupled with the phase wind profiler using $N = 4$ temporal samples, $n_d = 4$ frames between successive temporal samples, and $n_i = 2$ Newton iterations. Bottom panel: weighted average over altitude of wind speed (blue), x-component of wind velocity (green) and y-component of wind velocity (brown). Weighting given by relative turbulence strength.

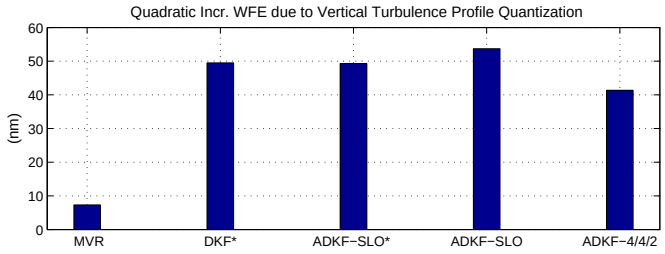

Fig. 8. Quadratic incremental LGS mode WFE averaged over 34 arcsec $\times$34 arcsec field root-mean-square (RMS) due to turbulence profile quantization. The value for each algorithm is obtained as the quadratic difference between the residual WFE achieved in the 65- and 6-layer simulations.