

Fast reconstruction and prediction of frozen flow turbulence based on structured Kalman filtering

Rufus Fraanje,^{1,2,*} Justin Rice,¹ Michel Verhaegen,¹ and Niek Doelman²

¹*Delft University of Technology, Delft Center for Systems and Control, Mekelweg 2, 2628 CD Delft, The Netherlands*

²*TNO Science & Industry, Precision Motion Systems, Stieltjesweg 1, 2628 CK Delft, The Netherlands*

**Corresponding author: p.r.fraanje@tudelft.nl*

Received April 1, 2010; revised July 10, 2010; accepted August 3, 2010;
posted August 26, 2010 (Doc. ID 126366); published September 27, 2010

Efficient and optimal prediction of frozen flow turbulence using the complete observation history of the wavefront sensor is an important issue in adaptive optics for large ground-based telescopes. At least for the sake of error budgeting and algorithm performance, the evaluation of an accurate estimate of the optimal performance of a particular adaptive optics configuration is important. However, due to the large number of grid points, high sampling rates, and the non-rationality of the turbulence power spectral density, the computational complexity of the optimal predictor is huge. This paper shows how a structure in the frozen flow propagation can be exploited to obtain a state-space innovation model with a particular sparsity structure. This sparsity structure enables one to efficiently compute a structured Kalman filter. By simulation it is shown that the performance can be improved and the computational complexity can be reduced in comparison with auto-regressive predictors of low order. © 2010 Optical Society of America

OCIS codes: 000.5490, 010.1080, 010.7060, 010.7350, 110.1080.

1. INTRODUCTION

The accurate and efficient reconstruction and prediction of wavefront errors is an essential ingredient of adaptive optics (AO) systems, especially for application in large ground-based telescopes. In AO the objective is to maximize the Strehl ratio, which can be approximated by minimizing the mean square residual wavefront error (Marechal approximation). The wavefront of the light observed by ground-based telescopes is distorted by time varying inhomogeneities in the diffraction index of the air caused by turbulence in the earth's atmosphere. This distortion can be compensated by means of a deformable mirror (DM). Controls for the DM are derived from measurements of the wavefront by a wavefront sensor (WFS). This WFS, however, does not measure the wavefront phase directly, but only a derived quantity, as its slope averaged over some region, such as in the Shack–Hartmann WFS. Light detectors in WFSs suffer from photon shot noise and read noise and need some exposure time to collect a sufficient amount of photons, depending on the intensity of the observed light.

The estimation of the wavefront phase based on the WFS observations is usually called wavefront reconstruction. Because the turbulence is changing in time, the compensation for the observed wavefront may not apply for the wavefront error a moment later after applying the DM controls. Hence, given the WFS observations a *prediction* needs to be made of the wavefront phase, such that any delays in the control loop are accounted for. In AO systems, wavefront phase predictors are used as a part of an (optimal) controller and can be used in both feedforward and feedback control configurations. In feedback control configurations the predictor is used as part of

an internal model controller (see, e.g., [1]), where the feedback control problem is reformulated in a feedforward control problem by subtraction of the influence of the control signal from the measurements. In the AO control literature this method is better known as pseudo open loop control (c.f. [2]).

The problem of efficiently and accurately reconstructing or predicting the wavefront phase has been raised especially in the context of AO for extremely large telescopes, where the number of actuators and sensors is expected to be on the order of 10^4 – 10^5 and the sampling frequencies on the order of kilohertz to achieve sufficient performance [3]. When no structure is exploited the computational complexity of the reconstruction or prediction of the wavefront phase will be extremely high; even an estimator consisting of just a single matrix-vector multiplication already takes several Tflops (1 flop is one floating point operation per second).

The algorithms proposed in the literature for wavefront reconstruction and/or prediction can be classified in two categories: (1) without memory and (2) with memory. The algorithms without memory compute the wavefront only on the basis of the last observation of the WFS. The methods with memory also take into account information from previous observations.

The first category is dominated by fast algorithms that exploit structures in the WFS matrix and turbulence correlation, e.g., the sparse iterative solvers proposed in [4,5], the FFT solution proposed in [6], and the fast solver exploiting the self-similarity of the turbulence proposed in [7]. We consider the warm start iterative solvers, which use the estimate of the previous sampling instant to initialize the iteration of the current sampling instant (see

[8]) to be in this category. All these methods provide a reconstruction of the wavefront that generated the last observation.

The second category is dominated by methods that make an optimal estimation, i.e., they predict the wavefront that needs to be compensated a moment later. The prediction is made using past observations together with some disturbance model quantifying its spatiotemporal correlation. For AO in astronomy applications the first predictive algorithm using past observations was proposed in [9]. In [10] a detailed analysis of minimum variance (MV) auto-regressive (AR) predictors has been given, showing that performance improves or alternatively that spatial- and temporal-bandwidth requirements can be relaxed. In [11] the analysis is further generalized taking into account the WFS sensing. In [12], in the context of multi-conjugate AO, a diagonal AR-1 (AR of order 1) model has been used to describe the turbulence, and predictions have been made using Kalman filtering. The turbulence model has been further generalized into a state-space innovation model as suggested in [13], and given in detail in [14,15], to adequately model the complete spatiotemporal correlation. In [14] the prediction is obtained by the Kalman filter, and to reduce the computational complexity the wavefront phase is transformed into a lower dimensional space by principal component analysis (PCA). In [16] a prediction has been made of the wavefront reconstructed by taking the pseudo-inverse of the WFS matrix, which was fully based on measured data. Recursive and adaptive alternatives of [16] have been proposed in [17].

Generally speaking, the methods in the first category sacrifice performance at the benefit of computational efficiency, whereas the methods of the second category sacrifice computational efficiency at the benefit of performance. The ultimate question is which method performs best given limited resources of WFSs, data-transfer, computational power, and memory? Most likely, the ideal approach will be somewhere in between the two categories. More recently, a structure from the frozen flow propagation has been exploited by some of the first category methods. In [2] it is shown that the optimal predictor can be implemented by means of an optimal reconstructor in series with a spatial shift (assuming an infinite dimensional spatial grid). In [18] the fast Fourier transform (FFT) method of [6] has been extended to include decoupled adaptive predictors. The solution of [18] purports to be optimal for Taylor frozen flow propagation of a phase screen with Kolmogorov turbulence statistics, provided that circular apertures are periodically extended to a square grid.

In this paper, our goal is to develop a method that combines the advantages of both categories, i.e., an optimal prediction based method that is computationally efficient. This goal is achieved by exploiting the structure from the frozen flow propagation to efficiently determine a structured Kalman filter. In this way, it is shown that (almost) optimal predictions can be made even with less computational complexity than the matrix-vector multiplication used in the optimal AR-1 predictor. The price to be paid is an increase in memory, which is about twice as much as

for the AR-1 predictor. However, delayed observations do not need to be stored in memory, as is the case for general AR- n predictors.

To be more precise, the frozen flow propagation of the wavefront can be described by a line by line propagation in time along one direction in space (the line not necessarily needs to be straight). This time-space propagation results in a string of interconnected subsystems, which together form a large structured state-space model. The structure in the state-space matrices can be considered as a special case of the more general *sequentially semi-separable* (SSS) matrices [19]. A SSS matrix is a matrix with a special structure that results from a string interconnection of a series of subsystems. Efficient algorithms for addition, (matrix-vector and matrix-matrix) multiplication, and inversion have been derived that exploit the fact that the calculations can be performed on the level of the local subsystems [19]. In [20] efficient methods, based on sign iterations, have been proposed for solving Riccati equations as well. The sign iterations preserve the SSS structure, such that the solution of the Riccati equation has a SSS structure as well. Because the Kalman gain matrix can be determined from the solution of a Riccati equation, this approach enables one to efficiently calculate the Kalman filter for frozen flow turbulence. Moreover, this Kalman filter will have a SSS structure as well, such that it can be efficiently implemented by a string of interconnected small subfilters.

The paper is organized as follows. Section 2 presents the turbulence model, the wavefront prediction problem, and its MV solution. Section 3 reformulates the frozen flow propagation in terms of a state-space system whose matrices have SSS structures. The solution of the SSS structured Kalman filter is presented, and its computational complexity is analyzed. Section 4 presents simulation results that compare the SSS structured Kalman filter with the optimal Kalman filter as well as a number of first order AR predictors.

2. TURBULENCE DESCRIPTION AND MINIMUM VARIANCE PREDICTION

A. Turbulence Description

We assume that, for simplicity of presentation, the atmospheric turbulence consists of a single layer and only distorts the phase of the wavefront (amplitude variations are neglected). We further assume that the statistics of the turbulence are well described by the von Kármán turbulence model with the Taylor assumption of frozen flow propagation.

Let the phase of the wavefront at time t and position (x,y) in the pupil plane be denoted by $\phi(t,x,y)$, then the von Kármán turbulence assumption specifies the spatial correlation as follows [21]

$$\begin{aligned} C_\phi(\delta_x, \delta_y) &:= E[\phi(t,x,y)\phi(t,x-\delta_x,y-\delta_y)] \\ &= c(2\pi r/L_0)^{5/6} K_{5/6}(2\pi r/L_0), \end{aligned} \quad (1)$$

where

$$c = \frac{\Gamma(11/6)}{2^{5/6}\pi^{8/3}} \left(\frac{24}{5} \Gamma(6/5) \right)^{5/6} \left(\frac{L_0}{r_0} \right)^{5/3},$$

$$r = \sqrt{(\delta_x)^2 + (\delta_y)^2},$$

L_0 is the outer scale of the turbulence, r_0 is the Fried parameter, $\Gamma(\cdot)$ is the gamma function, and $K_{5/6}(\cdot)$ is the modified Bessel function of the second kind of order $5/6$. The Taylor frozen flow assumption assumes that the wavefront phase propagates in time as a wave with constant velocity (v_x, v_y) , such that

$$\phi(t + \tau, x, y) = \phi(t, x - v_x \tau, y - v_y \tau). \quad (2)$$

The wavefront phase is sampled in time and space with a sampling time of ΔT and resolutions of ΔX and ΔY along the x - and y -axes, respectively. Then, with some abuse of notation we denote the discretized samples by

$$\phi(k, i, j) = \phi(k\Delta T, i\Delta X, j\Delta Y), \quad (3)$$

where k, i, j are integers. Note that the spectrum of $\phi(t, x, y)$ is nonzero for all temporal and spatial frequencies, such that sampling in time and space always yields aliasing errors that can be arbitrarily reduced by decreasing ΔT , ΔX , and ΔY .

The frozen flow propagation can have an arbitrary direction on the spatial grid. This flow can be modeled by simple local subsystems located on the spatial grid points that only interact with their neighbors and locally propagate the wavefront. Efficient control and filtering methods for this class of interconnected systems over a two-dimensional (2D) grid has been presented in [22]. Here, for reasons of simplicity, we consider the situation where the frozen flow propagation can be modeled as an interconnection of a series of subsystems along a string, i.e., along one dimension. This requires us to choose the spatial grid to be aligned with the propagation of the wavefront phase, such that $v_y = 0$ and $v_x \geq 0$. As a consequence, there may be a (rotational) misalignment with the spatial sampling by the grid of the WFS. But by means of an interpolation the WFS measurements can be fitted to the grid aligned with the wavefront propagation, e.g., by means of a linear, polynomial, or B -spline interpolation. The interpolation can also be performed by means of a MV estimation of the wavefront phase (or slopes) on one grid from the known equivalents on the other grid using the spatial covariance and the cross-correlation of the wavefront phase on both grids. When the interpolation is based only on the nearest neighbors the additional computational complexity scales linearly with the number of spatial grid points. The interpolation will cause an error, which however can be arbitrarily reduced by reducing ΔX and ΔY , such that there will be grid points on the grid aligned with the wavefront propagation that are arbitrarily close to the WFS grid points. This may result in a greater number of grid points for the wavefront propagation than the number of grid points of the WFS. In this case, the resulting problem of estimating or predicting the wavefront phase on a dense spatial grid using WFS measurements on a course spatial grid is in fact a super-resolution problem. Simulations with WFS measurements on a course grid have been included in Section 4.

With the grid aligned to the propagation direction, the time that is needed for the wavefront to propagate over one grid point along the x -axis direction is equal to $\Delta X/v_x$. When $\Delta X/v_x$ is exactly a multiple of the sampling time ΔT , the propagation of the wavefront can be modeled by just shifting the upstream samples after a delay of an integer number of sampling times. In general, however, $\Delta X/(v_x \Delta T)$ will not be integer, and approximation of this propagation delay by an integer number of sampling instants will give rise to errors and suboptimal reconstruction and prediction. This error can be arbitrarily reduced by adjusting ΔX and ΔT such that $\Delta X/(v_x \Delta T)$ will be integer, but this may lead to high dimensional spatial/temporal grids; moreover the sampling in time and space will become dependent of the propagation velocity, which is not desirable. A better alternative is to model the fractional part of $\Delta X/(v_x \Delta T)$ by means of an approximation of a fractional delay. In [23] several methods have been discussed for modeling of fractional delay. For the sake of simplicity, however, but without lack of generality, we assume here that $\Delta X/(v_x \Delta T) = 1$, i.e., at every sampling instant the wavefront is propagated over one grid point along the x -axis direction, which implies that the frozen flow propagation satisfies the relation

$$\phi(k + 1, i, j) = \phi(k, i - 1, j). \quad (4)$$

Then, using both Eqs. (1) and (4) it can be verified that the temporal and spatial correlation coefficients are given by (c.f. [15])

$$E[\phi(k, i, j)\phi(k - \ell, i - m, j - n)] = C_\phi((m - \ell)\Delta X, n\Delta Y). \quad (5)$$

Suppose that we have $N_x \times N_y$ grid points, and for each time instant k we stack the phase at these grid points in one big vector:

$$\varphi(k) := [\phi(k, 1, 1), \dots, \phi(k, 1, N_y), \dots, \phi(k, N_x, N_y)]^T \in \mathbb{R}^{N_x N_y}. \quad (6)$$

The correlation coefficients of the random vector process $\varphi(k)$ are denoted by

$$C_\varphi(\ell) := E[\varphi(k)\varphi(k - \ell)^T]. \quad (7)$$

Unfortunately, $\varphi(k)$ cannot be measured directly, and only a derived quantity,

$$y(k) := G\varphi(k) + \nu(k) \in \mathbb{R}^{N_y}, \quad (8)$$

is measured by a WFS, where G is the WFS sensing matrix and $\nu(k)$ is the measurement noise, assumed to be zero-mean white noise, independent of $\varphi(k - \ell)$ for all ℓ, k and with covariance $E[\nu(k)\nu(k - \ell)^T] = \sigma_\nu^2 I_{N_y} \delta(k)$ [with $\delta(\cdot)$ being the Kronecker delta function with $\delta(0) = 1$ and $\delta(k) = 0$ for $k \neq 0$]. For example, G may contain the coefficients of the finite difference approximation of the spatial derivatives in the case of Shack–Hartmann sensors.

B. Minimum Variance Prediction

Because the delay in the control loop is usually about two samples (about one sample delay in the WFS and one sample delay in the discrete-time control system), we want to estimate $\varphi(k + 2)$, i.e., to make a 2-step-ahead pre-

diction, given the measurements $y(k-\ell)$ for $\ell=0,1,\dots$. The AR-1 2-step-ahead predictor, which provides the MV prediction of $\varphi(k+2)$ given only the last measurement $y(k)$, is given by

$$\hat{\varphi}(k+2|k) = A_1 y(k), \quad (9)$$

where

$$A_1 = C_\varphi(2)G^T(GC_\varphi(0)G^T + \sigma_v^2 I_{N_y})^{-1}. \quad (10)$$

Assuming an infinite dimensional spatial grid, this predictor is the same as the AR-1 0-step-ahead predictor with a shift over two grid points in the x -direction as was shown in [2].

However, these AR-1 predictors do not take into account the information about $\varphi(k+2)$ that is present in $y(k-\ell)$ for $\ell > 1$. Because of the frozen flow propagation, it may actually be that the “downstream” elements in $\varphi(k+2)$ are correlated much stronger with “upstream” elements in $y(k-\ell)$ for some $k \geq 1$ than with the elements in the measurement $y(k)$. At least when there is correlation between $\varphi(k+2)$ and $y(k-\ell)$ for $k \geq 1$ the estimate of $\varphi(k+2)$ might be improved.

The MV (linear) prediction, taking into account all past measured data up to k , is given by

$$\hat{\varphi}(k+2|k) = A_k y_k(k), \quad (11)$$

where

$$y_k(k) = [y(k)^T, y(k-1)^T, \dots, y(1)^T]^T, \quad (12)$$

and A_k is obtained from

$$A_1 = C_\varphi(2)G^T P_1, \quad (13)$$

where

$$P_1 = (GC_\varphi(0)G^T + \sigma_v^2 I_{N_x N_y})^{-1}, \quad (14)$$

$$S_1 = GC_\varphi(1)G^T, \quad (15)$$

and for $k \geq 2$

$$A_k = \begin{bmatrix} A_{k-1} & 0 \end{bmatrix} + E_{k-1} \begin{bmatrix} K_{k-1}^T & -R_{k-1}^{-1} \end{bmatrix}, \quad (16)$$

with

$$K_{k-1} = P_{k-1} S_{k-1} R_{k-1}^{-1}, \quad (17)$$

$$R_{k-1} = GC_\varphi(0)G^T + \sigma_v^2 I_{N_x N_y} - S_{k-1}^T P_{k-1} S_{k-1}, \quad (18)$$

$$E_{k-1} = A_{k-1} S_{k-1} - C_\varphi(k+1)^T G^T, \quad (19)$$

$$P_k = \begin{bmatrix} P_{k-1} + K_{k-1} R_{k-1} K_{k-1}^T & -K_{k-1} \\ -K_{k-1}^T & R_{k-1}^{-1} \end{bmatrix}, \quad (20)$$

$$S_k = \begin{bmatrix} GC(k)G^T \\ S_{k-1} \end{bmatrix}. \quad (21)$$

This predictor is readily obtained using linear (recursive) prediction theory (cf. [24]).

Unfortunately, the computational complexity and memory of the predictor in Eq. (11) grow unboundedly. Only when E_k given by Eq. (19) satisfies $E_k=0$ for all

$k > p$ the predictor remains constant for $k > p$, and the predictions are based on the finite data window $\{y(k), \dots, y(k-p)\}$. It can be shown that for AR processes, indeed, $E_k=0$ for $k > p$, where p is the order of the AR process. However, the correlation coefficients determined by the von Kármán turbulence model cannot be obtained by a finite order AR process because of its non-rational spectrum. Of course, one may approximate the turbulence by a finite order AR process, but the order may be very high. A better choice would be to model the turbulence statistics by a finite order state-space innovation model, which also comprises the more general class of AR moving average models [[25], Section 8.3]. The state-space innovation model is given by

$$x(k+1) = Ax(k) + Ke(k), \quad (22)$$

$$\varphi(k) = Cx(k) + e(k), \quad (23)$$

where $e(k) \in \mathbb{R}^{N_x N_y}$ is a zero-mean white noise stochastic process with variance $E[e(k)e(k)^T] = R_e$; $x(k) \in \mathbb{R}^n$ is the state of the innovation model of dimension n ; and the matrices $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{N_x N_y \times n}$, and $K \in \mathbb{R}^{n \times N_x N_y}$ are the state-transition matrix, the output matrix, and the Kalman gain, respectively. This approach for modeling turbulent phase screens was proposed in [15], where the state-space matrices were determined from the correlation coefficients $C_\varphi(k)$ using stochastic realization theory (c.f. [13,26]). In the simulations in Section 4 we have used the stochastic realization method as outlined in [26] to determine the matrices A , K , C , and R_e from the correlation coefficients $C_\varphi(i)$ for $i=0,1,2,\dots,N$.

It can be guaranteed that the matrix A is a stability matrix, such that the pair (A, GC) is always detectable. Then, given the state-space description (22) and (23) the MV predictor of $\varphi(k+1)$ using the measurements (8) can be directly obtained from discrete-time Kalman filtering theory [27]

$$\hat{x}(k+1|k) = A\hat{x}(k|k-1) + K_y(y(k) - GC\hat{x}(k|k-1)), \quad (24)$$

$$\hat{\varphi}(k+1|k) = C\hat{x}(k+1|k), \quad (25)$$

where

$$K_y = (APC^T G^T + S)(GCPC^T G^T + R)^{-1}, \quad (26)$$

and P is the stabilizing positive definite solution of the following discrete algebraic Riccati equation (DARE):

$$P = APA^T - (APC^T G^T + S)(GCPC^T G^T + R)^{-1}(APC^T G^T + S)^T + Q, \quad (27)$$

where

$$Q = KR_e K^T, \quad (28)$$

$$R = GR_e G^T + \sigma_v^2 I_{N_y}, \quad (29)$$

$$S = KR_e G^T. \quad (30)$$

The 2-step-ahead prediction is easily obtained by evaluating one more time-update of the state:

$$\hat{x}(k+2|k) = A\hat{x}(k+1|k), \quad (31)$$

$$\hat{\phi}(k+2|k) = C\hat{x}(k+2|k). \quad (32)$$

For the stochastic realization a singular value decomposition (SVD) needs to be computed for a matrix whose rows and columns scale with the number of grid points $N_x N_y$. Because the computational complexity of the SVD scales with the cube of its number of rows and columns, the complexity to determine model (22) and (23) will scale with $(N_x N_y)^3$, and thus will be infeasible for large numbers of grid points. From simulations it follows that the order n scales approximately with $N_x N_y$. Because the computational complexity of the DARE (27) scales with n^3 , solving Eq. (27) will also be infeasible for large numbers of grid points. Fortunately, the frozen flow propagation yields a particular structure, which can be exploited to efficiently model and predict the wavefront phase, as will be discussed in the next section.

3. EFFICIENT AND DISTRIBUTED PREDICTION BY SEQUENTIALLY SEMI-SEPARABLE SYSTEMS

A. Structure of Frozen Flow Propagation

Recall the frozen flow assumption in Eq. (4), which is given again here:

$$\varphi(k+1, i, j) = \varphi(k, i-1, j). \quad (33)$$

Suppose that $v_x > 0$, and $\varphi(k, i, j)$ is known for all k and j and for $i=1$, then we can also derive $\varphi(k, i, j)$ for all k and j and $i > 1$ using Eq. (33), i.e., by properly shifting in time and space. These shifts can be modeled by a string of interconnected subsystems as depicted in Fig. 1. As a consequence, if a state-space innovation model is given for $\varphi(k, i, j)$ for $j=1, \dots, N_y$ and $i=1$ using Eq. (33) the state-space innovation model describing $\varphi(k, i, j)$ for $j=1, \dots, N_y$ and $i=1, \dots, N_x$ can be derived as well for $N_x > 1$. Moreover, thanks to the interconnection structure of the frozen flow this new state-space description has a particular nicely distributed structure along the x -coordinate.

A similar argument holds true for $v_x < 0$, for which a distributed model can be easily derived by either reordering the spatial grid or adjusting the indexing and propagation direction in the following derivation for $v_x > 0$. The case $v_x = 0$ will not be considered here, because for constant wavefronts the reconstruction problem reduces to a static estimation problem that can be easily solved by averaging the estimates obtained with (fast) static reconstruction methods.

To derive the distributed state-space model over the whole grid, let us first stack all N_y phases at the grid points along the y -dimension in the following vector:

$$\varphi_i(k) := [\varphi(k, i, 1), \varphi(k, i, 2), \dots, \varphi(k, i, N_y)]^T. \quad (34)$$

Then, by the frozen flow propagation (33) we have $\varphi_i(k+1) = \varphi_{i-1}(k)$.

Second, for $i=1$ an n th order state-space innovation model can be determined using the correlation in time and space along the y -dimension, as outlined in Subsection 2.B. This state-space model will be denoted by

$$\xi_1(k+1) = A\xi_1(k) + Ke_1(k), \quad (35)$$

$$\varphi_1(k) = C\xi_1(k) + e_1(k), \quad (36)$$

where $\xi_1(k) \in \mathbb{R}^n$ is the state, and $e_1(k) \in \mathbb{R}^{N_y}$ is a zero-mean white noise signal with covariance $R_{e1} \in \mathbb{R}^{N_y \times N_y}$. Since $\varphi_1(k) \in \mathbb{R}^{N_y}$ is just an N_y -dimensional vector, the computational complexity of the realization of model (35) and (36) just scales with N_y^3 rather than $(N_x N_y)^3$.

Third, the complete model is obtained by extending Eqs. (35) and (36) with the frozen flow propagation and can be written in terms of the following 2D state-space model:

$$\begin{aligned} \text{for } i=1: \quad \Sigma_i: \begin{bmatrix} \xi_i(k+1) \\ v_{i+1}^m(k) \\ \varphi_i(k) \end{bmatrix} &= \begin{bmatrix} A & 0 & K \\ C & 0 & I_{N_y} \\ C & 0 & I_{N_y} \end{bmatrix} \begin{bmatrix} \xi_i(k) \\ v_i^m(k) \\ e_i(k) \end{bmatrix}, \\ \\ \text{for } i=2, \dots, N_x: \quad \Sigma_i: \begin{bmatrix} \xi_i(k+1) \\ v_{i+1}^m(k) \\ \varphi_i(k) \end{bmatrix} &= \begin{bmatrix} 0 & I_{N_y} & 0 \\ I_{N_y} & 0 & 0 \\ I_{N_y} & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_i(k) \\ v_i^m(k) \\ e_i(k) \end{bmatrix}. \end{aligned} \quad (37)$$

Note that in fact $v_i^m(k)$ and $e_i(k)$ for $i > 1$ do not influence $\varphi_i(k)$ for all k and i and were just included for uniformity in the 2D state-space descriptions, but could be removed from the equations as well. Also note that $v_i^m(k) = \varphi_{i-1}(k)$ and $\xi_i(k) = v_i^m(k-1)$ for $i > 1$.

Equation (37) describes a system with N_x spatially interconnected subsystems. If the N_x subsystems are connected and the interconnection variables are resolved, we obtain the interconnected system

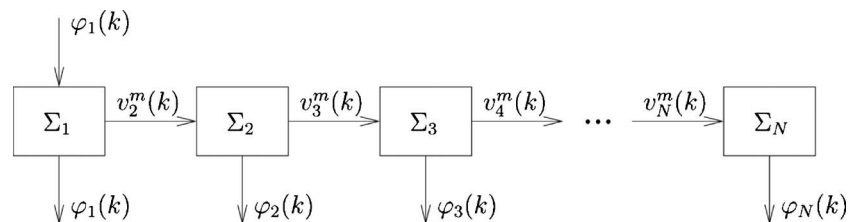


Fig. 1. Frozen flow propagation by subsystem string interconnection.

$$\begin{bmatrix} \bar{\xi}(k+1) \\ \bar{\varphi}(k) \end{bmatrix} = \begin{bmatrix} \bar{A} & \bar{K} \\ \bar{C} & \bar{D} \end{bmatrix} \begin{bmatrix} \bar{\xi}(k) \\ \bar{e}(k) \end{bmatrix}, \quad (38)$$

where $\bar{\xi}(k) = [\xi_1(k)^T, \dots, \xi_N(k)^T]^T$ is the lifted state vector, and $\bar{e}(k)$ and $\bar{\varphi}(k)$ are defined similarly. The matrices \bar{A} , \bar{K} , \bar{C} , and \bar{D} are defined as

$$\bar{A} = \begin{bmatrix} A & & & & \\ C & 0 & & & \\ & I_{N_y} & 0 & & \\ & & \ddots & \ddots & \\ & & & I_{N_y} & 0 \end{bmatrix}, \quad \bar{K} = \begin{bmatrix} K & & & & \\ I_{N_y} & 0 & & & \\ & 0 & 0 & & \\ & & \ddots & \ddots & \\ & & & 0 & 0 \end{bmatrix},$$

$$\bar{C} = \begin{bmatrix} C & & & & \\ & I_{N_y} & & & \\ & & \ddots & & \\ & & & I_{N_y} & \end{bmatrix}, \quad \bar{D} = \begin{bmatrix} I_{N_y} & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \end{bmatrix}. \quad (39)$$

The order of the state-space model (38) is $n + (N_x - 1)N_y$. Note that Eq. (38) describes the same turbulence as Eqs. (22) and (23), but the models were obtained in a different manner. Ideally, when all approximations are exact it holds that $\varphi(k) = \bar{\varphi}(k)$ if $e(k) = \bar{e}(k)$ for all k .

Note that this approach is not limited to square apertures, but can be applied to circular apertures as well. To illustrate this, consider Fig. 2 which shows a square grid on a circular aperture, and the propagation direction is along the x -axis as indicated by the arrows. Then, all the phases at all grid points at the extreme left points along the x -axis indicated by the squares will be stacked in one vector similar to Eq. (34). A state-space innovation model can be determined for this vector-signal as in Eqs. (35) and (36). Then, this vector can be propagated along the x -axis as in Eq. (37). Note that after a number of steps, for

the grid in Fig. 2 after five steps, the dimension of the vector that is propagated can be reduced because of the circular aperture. The essential point is that the grid along the propagation direction needs to be equidistant.

B. Frozen Flow as a Sequentially Semi-Separable System

The matrices \bar{A} , \bar{K} , \bar{C} , and \bar{D} are sparse and have structures that can be exploited by efficiently computing the predictor. Actually, the structures in \bar{A} , \bar{K} , \bar{C} , and \bar{D} are a special case of SSS matrices. State-space models whose matrices have SSS structures naturally result from subsystem string interconnections as in Fig. 1. SSS matrices have the nice property that the multiplication of a SSS matrix with an arbitrary vector can be calculated as a (state-space) recursion in terms of the low dimensional subsystems that generate the SSS matrix. The “state” in this state-space recursion is the interconnection variables between the subsystems, whose dimension determines the complexity of the SSS matrix-vector product.

As an example, consider a SSS matrix \bar{A} which can be block-partitioned as [19]

$$\bar{A} = (A_{ij}), \quad \text{where } A_{ij} \in \mathbb{R}^{\ell_i \times m_j} \text{ satisfies } A_{ij} = \begin{cases} D_i, & \text{if } i = j \\ U_i W_{i+1} \cdots W_{j-1} V_j, & \text{if } j > i \\ P_i R_{i-1} \cdots R_{j+1} Q_j, & \text{if } j < i, \end{cases} \quad (40)$$

for $i, j = 1, \dots, N$. Then, the matrix-vector multiplication $\bar{y} = \bar{A}\bar{u}$, with $\bar{y} \in \mathbb{R}^\ell$ ($\ell = \sum_{i=1}^N \ell_i$) and $\bar{u} \in \mathbb{R}^m$ ($m = \sum_{i=1}^N m_i$), block-partitioned corresponding with the partitioning of \bar{A} , can be performed by the following recursion:

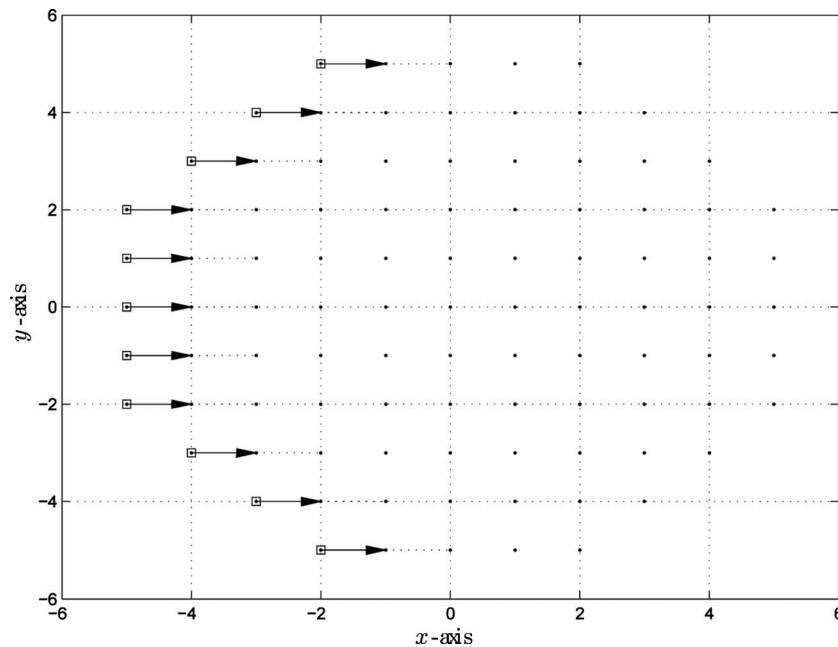


Fig. 2. Frozen flow propagation along a square grid in a circular aperture.

$$\Sigma_i: \begin{bmatrix} v_{i+1}^m \\ v_{i-1}^p \\ y_i \end{bmatrix} = \begin{bmatrix} R_i & 0 & Q_i \\ 0 & W_i & V_i \\ P_i & U_i & D_i \end{bmatrix} \begin{bmatrix} v_i^m \\ v_i^p \\ u_i \end{bmatrix}, \quad (41)$$

for $i=1, \dots, N$ and $v_1^m=0$, and $v_N^p=0$ (cf. Fig. 3).

Now, let us compare the computational complexity of directly evaluating $\bar{y}=\bar{X}\bar{u}$ with the recursions of Eq. (41). Let $\ell_i=\ell'$ and $m_i=m'$ for all $i=1, \dots, N$, then $\bar{X} \in \mathbb{R}^{N\ell' \times Nm'}$ such that the number of floating point operations (additions and multiplications) is $2N^2\ell'm'$. Further, let n_m and n_p be the dimensions of v_i^m and v_i^p , respectively, which are typically much smaller than N . Then evaluating N iterations of Eq. (41) takes $2N(n_m^2+n_p^2+(n_m+n_p) \times (\ell'+m')+\ell'm')$ floating point operations. Hence, the complexity of evaluating Eq. (41) scales with N rather than N^2 . For the case that $\ell'=m'=1$ and $n_m=n_p=2$, then evaluating $\bar{X}\bar{u}$ by one big matrix-vector multiplication takes $2N^2$, whereas evaluating Eq. (41) takes $34N$ floating point operations; so for $N>4$ evaluating Eq. (41) is more efficient than the matrix-vector multiplication.

It can also be shown that the class of SSS matrices is closed under addition, multiplication, and inversion [19]. However, when adding or multiplying two SSS matrices, the dimension of the subsystem interconnections v_i^m and v_i^p will be the sum of the dimensions of the subsystem interconnections of both terms. In [20] efficient order reduction techniques have been proposed to limit the dimensions of the interconnections, which can be applied after each addition or multiplication.

Moreover, it has been shown in [20] that the solution of Riccati equations can be approached by means of the so-called sign iterations that also preserve the SSS structure. In this way Riccati equations can be solved efficiently, with a computational complexity that scales with the number of subsystems N rather than N^3 . In addition the solutions are in terms of SSS matrices as well, such that SSS structured predictors and controllers can be designed.

C. Augmenting the Model with Wavefront Sensing

Before the Kalman filter can be derived to predict the wavefront phase in Eq. (38), the model needs to be augmented with the WFS output. The WFS will be spatially distributed as well; it will provide measurements of spatial derivatives or higher order derivatives averaged only over a local region in space. Hence, it is still possible to derive an interconnected subsystem description as in Eq. (37), which is essential to obtain the SSS structure in the lifted system. The main difference with Eq. (37) will be that the WFS will provide linear combinations of the wavefront phases at a number of grid points, such that

each subsystem needs to be augmented to include all grid points necessary for determining the (local) WFS output.

To illustrate this, consider a WFS that measures the slopes in the x - and y -directions, which are approximated by the following averaged finite differences:

$$y^x(k, i, j) = \frac{1}{\Delta} \sum_{n=1}^{\Delta} \varphi(k, i\Delta, (j-1)\Delta + n) - \varphi(k, (i-1)\Delta + 1, (j-1)\Delta + n) + v^x(k, i, j), \quad (42)$$

$$y^y(k, i, j) = \frac{1}{\Delta} \sum_{n=1}^{\Delta} \varphi(k, (i-1)\Delta + n, j\Delta) - \varphi(k, (i-1)\Delta + n, (j-1)\Delta + 1) + v^y(k, i, j), \quad (43)$$

where v_x and v_y are the measurement noise processes, $\Delta \in \mathbb{N}$ is the resolution of the WFS, and i and j range from 1 to N_x/Δ and N_y/Δ , respectively. Equations (42) and (43) can also be easily adjusted according to the Fried geometry. Here Eqs. (42) and (43) have been used for the sake of simplicity as well as to allow for different resolutions Δ of the WFS.

Then, by stacking the vector $\varphi_\ell(k)$ defined in Eq. (34) locally over $\ell=(i-1)\Delta+1, \dots, i\Delta$,

$$\bar{\varphi}_i(k) = [\varphi_{(i-1)\Delta+1}(k)^T, \dots, \varphi_{i\Delta}(k)^T]^T, \quad (44)$$

and stacking the observations $y^x(k, i, j)$ and $y^y(k, i, j)$ for $j=1, \dots, N_y/\Delta$ into

$$y_i^x(k) = [y^x(k, i, 1)^T, \dots, y^x(k, i, N_y)^T]^T, \quad (45)$$

$$y_i^y(k) = [y^y(k, i, 1)^T, \dots, y^y(k, i, N_y)^T]^T, \quad (46)$$

we can rewrite the averaged finite differences as follows:

$$\bar{y}_i(k) = \begin{bmatrix} y_i^x(k) \\ y_i^y(k) \end{bmatrix} = \begin{bmatrix} G^x \\ G^y \end{bmatrix} \bar{\varphi}_i(k) + \bar{v}_i(k), \quad (47)$$

where the matrices G^x and G^y are sparse matrices determining the averaged finite differences in the x - and y -directions, and $\bar{v}_i(k)$ contains the measurement noise terms.

To generate the lifted vectors $\bar{\varphi}_i(k)$ rather than $\varphi_i(k)$ the subsystems in Eq. (37) need to be lifted as well, resulting in N_x/Δ subsystems whose sparsity structure is rather similar to Eq. (37). The process of lifting a series of subsystems is very similar to the lifting over all N_x grid points along the x -axis in Eq. (38) whose state-space matrices are defined in Eq. (39). The difference now is that the lifting is performed a number of times over Δ subsystems resulting in N_x/Δ larger subsystems.

Then the resulting string of subsystems with $\bar{\varphi}_i(k)$ and $\bar{y}_i(k)$ as outputs can be written as

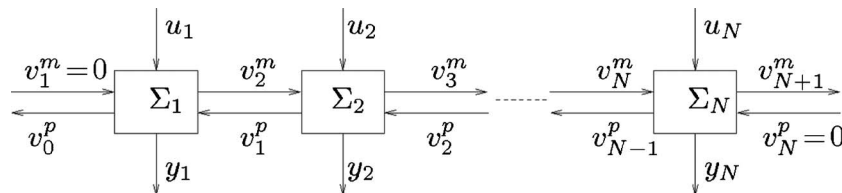


Fig. 3. SSS matrix-vector multiplication as a subsystem string interconnection.

$$\Sigma_i: \begin{bmatrix} \bar{\xi}_i(k+1) \\ \bar{v}_{i+1}^m(k) \\ \bar{\varphi}_i(k) \\ \bar{y}_i(k) \end{bmatrix} = \begin{bmatrix} A_i & B_i^m & K_i^e & 0 \\ C_i^m & 0 & D_i^m & 0 \\ C_i^e & 0 & D_i^e & 0 \\ C_i^y & 0 & D_i^y & I_{2N_y/\Delta} \end{bmatrix} \begin{bmatrix} \bar{\xi}_i(k) \\ \bar{v}_i^m(k) \\ \bar{e}_i(k) \\ \bar{v}_i(k) \end{bmatrix}, \quad (48)$$

for $i=1, \dots, N_x/\Delta$, where $\bar{\xi}_i(k)$ and $\bar{e}_i(k)$ are the vectors stacking the states $\xi_\ell(k)$ and the innovations $e_\ell(k)$ over $\ell = (i-1)\Delta + 1, \dots, i\Delta$, and \bar{v}_i^m is the N_y -dimensional communication channel between subsystems, with $\bar{v}_1^m(t)=0$. The matrices $(A_i, B_i^m, K_i^e, C_i^m, D_i^m, C_i^e, D_i^e, C_i^y, D_i^y)$ are derived from Eqs. (37) and (47) and have some additional sparsity structures which will not be denoted explicitly but can be exploited in the implementation of the predictor. Lifting Eq. (48) further along all points $i=1, \dots, N_x/\Delta$ will result in system (38) augmented with the WFS outputs:

$$\begin{bmatrix} \bar{\xi}(k) \\ \bar{\varphi}(k) \\ \bar{y}(k) \end{bmatrix} = \begin{bmatrix} \bar{A} & \bar{K} & 0 \\ \bar{C} & \bar{D} & 0 \\ \bar{C}^y & \bar{D}^y & I \end{bmatrix} \begin{bmatrix} \bar{\xi}(k) \\ \bar{e}(k) \\ \bar{v}(k) \end{bmatrix}, \quad (49)$$

where $\bar{C}^y = G\bar{C}$ and $\bar{D}^y = G\bar{D}$, where G is the WFS matrix.

D. Kalman Filtering of Frozen Flow Turbulence

Actually system (49) is equivalent to the state-space model (22) and (23) together with the WFS output (8) such that the 2-step-ahead prediction is given by the following expressions:

$$\hat{e}(k) = \bar{y}(k) - \bar{C}^y \hat{\xi}(k|k-1), \quad (50)$$

$$\hat{\xi}(k+1|k) = \bar{A} \hat{\xi}(k|k-1) + \bar{K}_y \hat{e}(k), \quad (51)$$

$$\hat{\xi}(k+2|k) = \bar{A} \hat{\xi}(k+1|k), \quad (52)$$

$$\hat{\varphi}(k+2|k) = \bar{C} \hat{\xi}(k+2|k), \quad (53)$$

where \bar{K}^y is the Kalman gain given by

$$\bar{K}^y = (\bar{A}\bar{P}(\bar{C}^y)^T + \bar{S})(\bar{C}^y\bar{P}(\bar{C}^y)^T + \bar{R})^{-1}, \quad (54)$$

and \bar{P} is the stabilizing positive definite solution of the DARE:

$$\begin{aligned} \bar{P} = & \bar{A}\bar{P}\bar{A}^T - (\bar{A}\bar{P}(\bar{C}^y)^T + \bar{S})(\bar{C}^y\bar{P}(\bar{C}^y)^T + \bar{R})^{-1}(\bar{A}\bar{P}(\bar{C}^y)^T + \bar{S})^T \\ & + \bar{Q}, \end{aligned} \quad (55)$$

where

$$\bar{Q} = \bar{K}\bar{R}_e\bar{K}^T, \quad (56)$$

$$\bar{R} = \bar{D}^y\bar{R}_e(\bar{D}^y)^T + \sigma_v^2 I, \quad (57)$$

$$\bar{S} = \bar{K}\bar{R}_e(\bar{D}^y)^T. \quad (58)$$

This predictor is the same as in Eqs. (24), (31), and (32), except that the state-space matrices now have SSS structures. Because the SSS structure is preserved under addition, multiplication, and inversion, the matrices \bar{Q} , \bar{R} , and \bar{S} also have SSS structures, i.e., they can be decomposed in a series of subsystems as matrix (40) where the subsystems are given by Eq. (41). Hence, \bar{P} will have a SSS structure as well as \bar{K}_y . In practice, it will be necessary to perform order reductions to limit the number of interconnection variables in \bar{Q} , \bar{R} , \bar{S} and after each addition and multiplication in the iterations to solve for \bar{P} . These order reductions will result in small errors in the solution of \bar{P} after convergence, and thus may result in an approximation of the Kalman gain \bar{K}_y . In the next section, we show by several simulations that the approximation error is relatively small, especially for high signal-to-noise ratios (SNRs).

The subsystem recursions for the SSS matrix-vector multiplication $\bar{u}(k) = \bar{K}_y \hat{e}(k)$ are given by

$$\begin{bmatrix} v_{i+1}^m(k) \\ v_{i-1}^p(k) \\ u_i(k) \end{bmatrix} = \begin{bmatrix} K_i^{mm} & 0 & K_i^{me} \\ 0 & K_i^{pp} & K_i^{pe} \\ K_i^{um} & K_i^{up} & K_i^{ue} \end{bmatrix} \begin{bmatrix} v_i^m(k) \\ v_i^p(k) \\ \hat{e}_i(k) \end{bmatrix}, \quad (59)$$

where the K_i^{xy} matrices are obtained from the solution of the Riccati equation and the expression of the Kalman gain in Eq. (54) together with the order reductions that reduce the dimensions n_m and n_p of v_i^m and v_i^p , respectively. In the simulations in the next section the maximal values of n_m and n_p were set to $N_y/2$ in all simulations, which appeared to be a good trade-off between complexity and accuracy of the wavefront prediction. We also observed in our experiments that the number of iterations of the sign algorithm to solve the Riccati equations should not be too large (typically three to five iterations) to prevent large truncation errors in the order reduction. The accuracy was also improved by slightly increasing the noise covariance, which makes the solution more robust to errors in the order reduction.

The computational complexity of the 2-step-ahead predictor is outlined in Table 1. Because the matrices \bar{A} and

Table 1. Computational Complexity of the SSS Kalman 2-Step-Ahead Predictor

Expression	Complexity
$\bar{e}(k) = \bar{y}(k) - \bar{C}^y \hat{\xi}(k k-1)$	$4nN_y/\Delta$
$\hat{\xi}(k+1 k) = \bar{A} \hat{\xi}(k k-1) + \bar{K}_y \bar{e}(k)$	$2n^2 + \frac{2N_x}{\Delta}(n_m^2 + n_p^2 + (n_m + n_p)(n + \frac{2N_y}{\Delta}) + \frac{2nN_y}{\Delta})$
$\hat{\xi}(k+2 k) = \bar{A} \hat{\xi}(k+1 k)$	$2n^2$
$\hat{\varphi}(k+2 k) = \bar{C} \hat{\xi}(k+2 k)$	$2N_y n$
Total (approximately)	$\frac{2N_x}{\Delta}(n_m^2 + n_p^2 + (n_m + n_p)(n + \frac{2N_y}{\Delta}) + \frac{2nN_y}{\Delta})$

\bar{C} in Eq. (49) are very sparse, the multiplication $\bar{K}_y \hat{e}(k)$ is the most complex part of the algorithm. However, without exploiting the SSS structure, the computational complexity is $4nN_x^2N_y/\Delta^2$ and scales with N_x^2 rather than N_x , which is usually much higher than that obtained with the SSS recursions. In the simulation experiments, the SSS recursions are even more efficient than the matrix-vector multiplication of the AR-1 predictor in Eq. (9), which has a computational complexity of $4N_x^2N_y^2/\Delta^2$.

4. SIMULATIONS

The method is validated on a turbulence model on a grid of $(N_x \times N_y) = (32 \times 32)$ points with a spatial resolution of $\Delta X = \Delta Y = 3.1 \times 10^{-2}$ [m]; the turbulence outer-diameter in the von Kármán model is set to $L_0 = 10$ [m], and the Fried parameter is $r_0 = 0.16$ [m]. The wind velocity is $v_x = 20$ [m/s], and the sampling time is $\Delta T = 1.56 \times 10^{-3}$ [s] (i.e., sampling rate is 640 [Hz]) such that for every sampling time the phase is shifted exactly by one grid point in the x -direction.

The disturbance model at $i=1$ has been determined using stochastic realization (c.f. [15]) on the basis of the correlation coefficients over the $N_y=32$ grid points in the y -direction and 251 time lags. The order of the disturbance model was set to $n=40$.

Experiments have been performed with WFS signals as defined by Eq. (42) with different WFS resolutions $\Delta=2, 4$, and 8, resulting in $2N_xN_y/\Delta^2=1024, 512$, and 256 sensor values, respectively. The variance of the measurement noise v_x and v_y has chosen to be relative to the variance of the finite differences with SNRs of 0 dB up to 40 dB with steps of 10 dB.

Various wavefront predictors have been evaluated: the first order AR reconstructor (i.e., no prediction is made) is denoted by “0-step AR-1.” Further, the 0-step AR-1 reconstructor with a shift according to the frozen flow propagation over two time steps is denoted by “0-step AR-1 with shift.” The first order AR 2-step-ahead predictor is denoted by “2-step AR-1.” The 2-step-ahead predictor based on the Kalman filter with a SSS structure is denoted by “2-step SSS Kalman,” and finally the 2-step-ahead predictor based on the Kalman filter obtained by solving the Riccati equation without restricting to the SSS structure is denoted by “2-step Kalman.”

For $\Delta=2$ the order of the state-space innovation model of the turbulence was 1280 ($=nN_x$). The time to compute the AR predictors was 4 s with MATLAB running on a 2.33GHz dual core personal computer (PC) with 4 Mbytes cache and 3 Gbytes working memory. On the same PC the computation of the SSS Kalman filter was done in 8.8 min, and the solution of the Riccati equation using the MATLAB routine DARE was done in 19.8 min. Here, we have to note that our MATLAB implementation of the SSS Riccati solver has not been optimized, and a significant further reduction in the computation time may be possible.

The performance is evaluated in terms of the variance of the piston term corrected prediction error averaged over all 32×32 grid points. Because in all experiments the full state-space model of the turbulent disturbance is

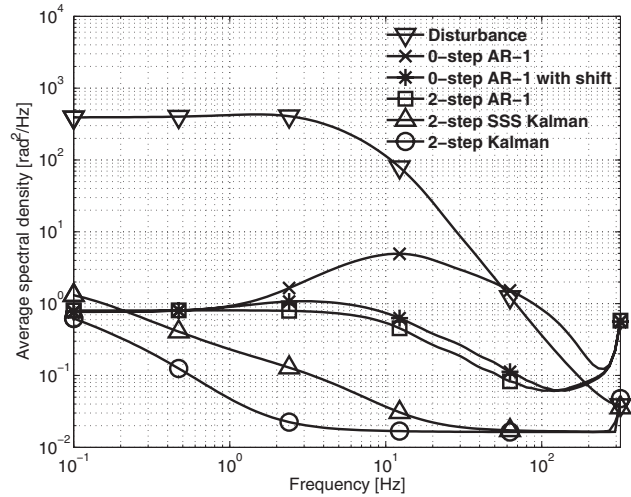


Fig. 4. Spectral density for $\Delta=2$, SNR=40 dB.

known, the variance of the piston term corrected prediction errors can be computed analytically. Also the distribution of the prediction error over the frequency, i.e., the power spectral density, can be computed. Figure 4 shows the power spectral density in [rad²/Hz] of the disturbance and the piston term corrected prediction errors obtained by each predictor averaged over all grid points for a WFS with a resolution of $\Delta=2$ and a 40 dB SNR. Figures 5–7 show the Strehl ratios determined by the Marechal approximation

$$S_r = \exp(-\sigma_e^2), \quad (60)$$

where σ_e^2 is the variance of the piston corrected prediction error averaged over all grid points, versus the SNRs for $\Delta=2, 4$, and 8, respectively.

Table 2 shows the computational complexity of the various predictors. The AR-1 predictors all have the same complexity, which is just listed once and is determined by the complexity of a matrix-vector multiplication. There exist fast implementations whose complexity scales linearly with N_xN_y , such as the robust multigrid implementations as presented in [28]. For the complexity of the SSS Kalman filter we only evaluated the complexity of the SSS recursions for the Kalman gain since the other state-space matrices are very sparse. The row with “unstruc-

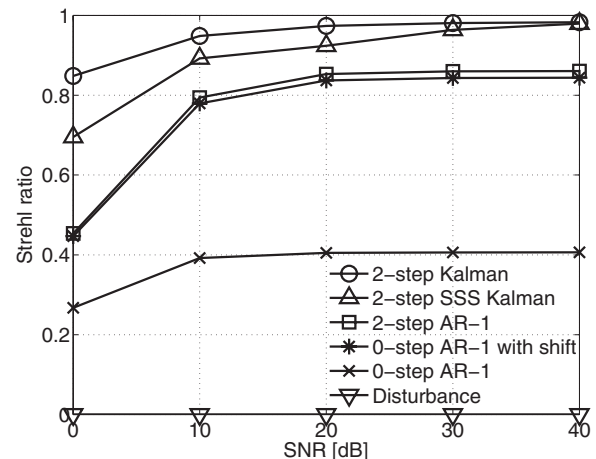
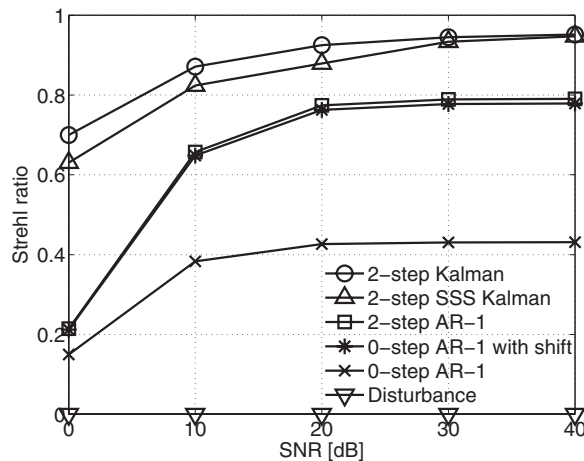


Fig. 5. Strehl ratio versus SNR for WFS resolution $\Delta=2$.

Fig. 6. Strehl ratio versus SNR for WFS resolution $\Delta=4$.

tured Kalman” denotes the complexity of the full matrix-vector multiplication for an unstructured Kalman gain. At the last row we also included the complexity of the Kalman filter provided the state-space matrices are not sparse as would be the case by determining the complete model by means of stochastic realization.

From Figs. 4 and 5 we observe that for the case $\Delta=2$, SNR=40 dB, the Strehl ratio obtained by the “SSS Kalman” is approximately the same as the Strehl ratio of 0.98 obtained by the true Kalman filter, i.e., without imposing the SSS structure. For lower SNRs the SSS Kalman filter loses a bit of performance, but stays significantly above the performance obtained by the 2-step AR-1 and the 0-step AR-1 with shift. This holds for all cases, $\Delta=2, 4, 8$. The 2-step AR-1 and the 0-step AR-1 with shift predictors have about the same performance, with the 0-step AR-1 with shift slightly worse, which is most likely due to errors due to the improper shift at the boundary. The difference between the 0-step AR-1 and the 0-step AR-1 with shift is about a reduction of 0.2–0.4 in the Strehl ratio, indicating that the performance is significantly improved by taking into account the frozen flow propagation. Of course the difference can be reduced by increasing the sampling frequency and thus limiting the spatial shift between subsequent time samples. However, increasing the sampling frequency directly reduces the

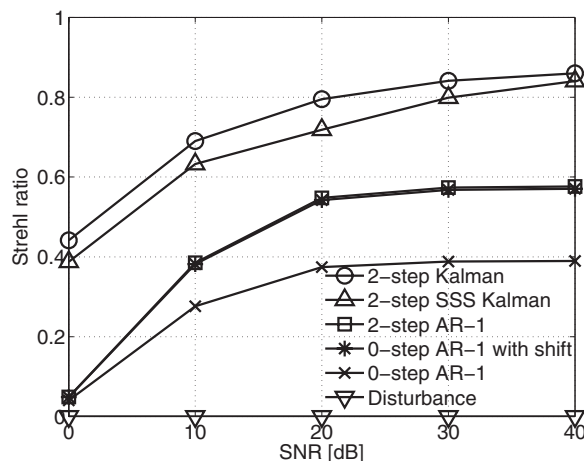
Fig. 7. Strehl ratio versus SNR for WFS resolution $\Delta=8$.

Table 2. Computational Complexity in Floating Point Operations per Sampling Time for the SSS Structured Kalman Predictor, the Various Types of AR-1 Predictors, the Kalman Filter with Unstructured Kalman Gain but Sparse State-Space Matrices, and the Kalman Filter with Unstructured State-Space Matrices

Method	$\Delta=2$	$\Delta=4$	$\Delta=8$
SSS Kalman	1.3×10^5	4.7×10^4	1.9×10^4
AR-1	1.0×10^6	2.6×10^5	6.6×10^4
Unstructured Kalman	1.3×10^6	3.3×10^5	8.2×10^4
Unstructured state Kalman	7.1×10^6	6.2×10^6	6.0×10^6

time to evaluate the computations and thus increases the computational complexity; moreover increasing the sampling frequency will also limit the exposure time of the WFS and thus increase the SNR. It is interesting to observe that the Strehl ratio of 0.85 obtained by the SSS Kalman predictor for the case of $\Delta=8$ is about the same as the Strehl ratio of 0.86 obtained by the 2-step AR-1 predictor for $\Delta=2$, whereas the computational complexity of the SSS Kalman predictor for this case is more than a factor of 50 smaller than the complexity of the 2-step AR-1 filter.

We should stress here that in the derivation of the structured state-space model some specific assumptions had been made, such as a perfect frozen flow whose propagation matches with the WFS grid and with a known velocity. Further investigation is necessary of how sensitive the performance of the SSS Kalman predictor is with respect to these assumptions. At least, the method may be a useful tool in error budgeting to efficiently compute the (approximate) Kalman filter prediction for comparison with (suboptimal) algorithms.

5. CONCLUSIONS

This paper has addressed the problem of efficiently computing and implementing a Kalman filter to predict frozen flow turbulence with von Kármán spatial correlation. The motivation was to investigate how efficiently a predictor that exploits all information available from the measurement history can be computed and implemented without sacrificing any performance.

Stochastic realization has been used to obtain a state-space innovation model of the turbulence at the upstream edge of the aperture. Then, the frozen flow is efficiently implemented by time delays and spatial shifts along the propagation direction, resulting in a state-space innovation model with relatively sparse matrices. Augmenting this model with a WFS provides the complete description of the turbulence generation and observation. This model has a value in itself, because it enables the analysis of the observability of the turbulence from various wavefront sensors (WFSs). In the simulations where the WFS measures the slopes of the wavefront averaged over a number of grid points, it appears that all states of this model are still observable. A structured Kalman filter predictor has been computed using an efficient Riccati equation solver

that exploits subsystem string interconnection structures and has been recently proposed by two of the authors in [20].

The first simulation results are encouraging, showing both a performance increase and a computational complexity reduction relative to AR-1 predictors. But, there remains a need for further research in this line. The stochastic realization modeling the turbulence at the grid points located on the upstream edge may still be computationally intensive for large numbers of grid points and high sampling ratios. Orthogonal basis transformations over the spatial dimension (such as the PCA in [14]) may be used to divide the stochastic realization in a number of decoupled problems. The structure from the orthogonal basis transformations might also be exploited in the Kalman filter design. Further simulations are necessary for frozen flow propagations that are not aligned with the observation grid; also uncertainties in the turbulence model (e.g., the outer range, the grid size, the jitter in the sampling frequency, and the accuracy of the velocity) need to be studied. It would also be interesting to see whether the structured turbulence model can be identified from the measured data in the line of the data-driven approach followed in [16].

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their constructive comments.

REFERENCES

1. M. Morari and E. Zafriou, *Robust Process Control* (Prentice Hall, 1989).
2. C. Vogel, "Sparse matrix methods for wavefront reconstruction revisited," *Proc. SPIE* **5490**, 1327–1335 (2004).
3. M. Le Louarn, N. Hubin, M. Sarazin, and A. Tokovinin, "New challenges for adaptive optics: Extremely large telescopes," *Mon. Not. R. Astron. Soc.* **317**, 535–544 (2000).
4. B. Ellerbroek, "Efficient computation of minimum variance wavefront reconstructors using sparse matrix techniques," *J. Opt. Soc. Am. A* **19**, 1803–1816 (2002).
5. L. Gilles, C. Vogel, and B. Ellerbroek, "Multigrid preconditioned conjugate-gradient method for large-scale wavefront reconstruction," *J. Opt. Soc. Am. A* **19**, 1817–1822 (2002).
6. L. Poyneer, D. Gavel, and J. Brase, "Fast wave-front reconstruction in large adaptive optics systems with use of the Fourier transform," *J. Opt. Soc. Am. A* **19**, 2100–2111 (2002).
7. M. Tallon, E. Thiébaud, and C. Béchet, "A fractal iterative method for fast wavefront reconstruction for extremely large telescopes," in *Proceedings of Adaptive Optics: Analysis and Methods* (2007), pp. 1–3.
8. L. Lessard, M. West, D. MacMynowski, and S. Lall, "Warm-started wavefront reconstruction for adaptive optics," *J. Opt. Soc. Am. A* **25**, 1147–1155 (2008).
9. M. Jorgenson and G. Aitken, "Prediction of atmospherically induced wave-front degradations," *Opt. Lett.* **17**, 466–468 (1992).
10. C. Schwartz, G. Baum, and E. Ribak, "Turbulence-degraded wave fronts as fractal surfaces," *J. Opt. Soc. Am. A* **11**, 444–451 (1994).
11. D. Gavel and D. Wiberg, "Toward Strehl-optimizing adaptive optics controllers," *Proc. SPIE* **4839**, 890–901 (2003).
12. B. L. Roux, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, L. Mugnier, and T. Fusco, "Optimal control law for classical and multiconjugate adaptive optics," *J. Opt. Soc. Am. A* **21**, 1261–1276 (2004).
13. K. Hinnen, M. Verhaegen, and N. Doelman, "Robust spectral factor approximation of discrete-time frequency domain power spectra," *Automatica* **41**, 1791–1798 (2005).
14. A. Beghi, A. Cenedese, and A. Madiero, "Atmospheric turbulence prediction: a PCA approach," in *Proceedings of the 46th IEEE Conference on Decision and Control* (IEEE, 2007), pp. 566–571.
15. A. Beghi, A. Cenedese, and A. Masiero, "Stochastic realization approach to the efficient simulation of phase screens," *J. Opt. Soc. Am. A* **25**, 515–525 (2008).
16. K. Hinnen, M. Verhaegen, and N. Doelman, "Exploiting the spatiotemporal correlation in adaptive optics using data-driven H_2 -optimal control," *J. Opt. Soc. Am. A* **24**, 1714–1725 (2007).
17. N. Doelman, R. Fraanje, I. Houtzager, and M. Verhaegen, "Real-time optimal control for adaptive optics systems," *Eur. J. Control* **15**, 480–488 (2009).
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. S. Chandrasekaran, P. Dewilde, M. Gu, T. Pals, and A.-J. van der Veen, "Fast stable solvers for sequentially semi-separable linear systems of equations," in *Lecture Notes in Computer Science* (Springer Verlag, 2002), pp. 545–554.
20. J. Rice and M. Verhaegen, "Distributed control: A sequentially semi-separable approach for spatially heterogeneous linear systems," *IEEE Trans. Autom. Control* **54**, 1270–1283 (2009).
21. R. Conan, "Modélisation des effets de l'échelle externe de cohérence spatiale du front d'onde pour l'observation à haute résolution angulaire en astronomie," Ph.D. dissertation (Université de Nice-Sophia Antipolis, 2000).
22. J. Rice and M. Verhaegen, "Distributed control of spatially invariant systems in multiple dimensions: A structure preserving computational technique" (submitted for publication, <http://www.dsc.tudelft.nl/jrice/Publications.htm>).
23. T. Laakso, V. Välimäki, M. Karjalainen, and U. Laine, "Splitting the unit delay—tools for fractional delay filter design," *IEEE Signal Process. Mag.* **13**, 30–60 (1996).
24. P. Brockwell and R. Davies, *Time Series: Theory and Methods* (Springer, 1991).
25. M. Verhaegen and V. Verdult, *Filtering and System Identification—A Least Squares Approach* (Cambridge University Press, 2007).
26. P. Van Overschee and B. De Moor, "Subspace algorithms for the stochastic identification problem," *Automatica* **29**, 649–660 (1993).
27. B. Anderson and J. Moore, *Optimal Filtering* (Prentice-Hall, 1979).
28. C. Vogel and Q. Yang, "Multigrid algorithm for least-squares wave-front reconstruction," *Appl. Opt.* **45**, 705–715 (2006).