Delft University of Technology
sc42025 Filtering & Identification
Practical Assignment 2017-2018

Turbulence modeling for adaptive optics

Abstract

This assignment is part of the course *Filtering and Identification* (SC42025). It is intended to be solved by a group of *two* students. The goal is to use the methods learned throughout the course to solve control problems that arise in adaptive optics (AO) system. It covers weighted least squares, Kalman filtering, Vector- Auto-Regressive modeling and stochastic subspace identification. The document is organized as follows. A first part describes the requirements for the practical assignment, a second part introduces the adaptive optics system, and the third part consists of the questions.

1 Requirements

1.1 Course material

In solving this assignment you are encouraged to use the following material:

- Course book: *Filtering and System Identification: A Least Squares Approach*; [1]
- 1 Matlab file entitled: *AOloop-nocontrol.m.*
- 2 datasets:
  - the dataset *systemMatrices.mat* contains the matrices G, H and the signal-to-noise ratio (SNR).
  - the dataset *turbulenceData.mat* contains atmospheric turbulence data collected in open-loop for 20 realizations.

It is **neither** allowed to use the *System Identification* toolbox from Matlab nor the *LTI Toolbox* from DCSC.

1.2 Desired output

One report -clearly structured and written- must be written by each group. A few recommendations follow:

- The cover page of the report must contain the names of the students, the student numbers as well as the date and a title of your document plus a short abstract stating the main conclusions of the work.
- x- and y-axis should be named, and units declared. Plots should be commented and there should be a caption with the observations displayed by the figure.
- The Matlab scripts *written by yourself* shall be included in an Appendix, and indexed by question number. There is no need to copy and paste the code from the AO simulator that was provided.

1.3 Handing in your solutions

Please leave a *hard copy* (printed in double-page printed format) of your solutions in the mail box of DCSC next to the secretariat. The deadline of the assignment is on January, 17^{th} 2018 at 5pm.

1.4 Grading

You will receive a grade between 1 and 10. To pass the course *Filtering and Identification*, your grade for this assignment must be larger or equal to 6.

2 Introduction

High resolution imaging in the visible spectrum from ground-based telescopes is seriously hampered by atmospheric turbulence. *Adaptive optics* (AO) is the system that corrects in real-time atmospheric aberrations as displayed in Figure 1. Incoming light is split in two beams and directed toward a wavefront sensor (WFS) and the instrument camera. The information provided by the sensors is processed to estimate a future wavefront and converted into voltage for the actuators located under a deformable mirror (DM). The DM flattens the wavefront in order to retrieve the original image quality as if there was no atmosphere.

2.1 Data-driven turbulence modeling

Modeling the atmospheric turbulence is of key importance for predicting the near-future of the wavefront, and hence achieve optimal control performances. We study in this assignment the crucial role of modeling *from data* in order to improve the control performances. Each part focuses on a different assumption on the model. In the first section we assume that the spatio-temporal dynamics are decoupled: the wavefront is first reconstructed and is then mapped onto the mirror by assuming that the wavefront has not evolved during the computation time. Second spatio-temporal dynamics of the wavefront are considered, first with a Vector- Auto-Regressive (VAR) model of order 1, and second with a stochastic state-space model which we identify from open-loop sensor data using stochastic subspace identification.

2.2 An adaptive optics model

While modeling real-life systems, a first direction is to use the knowledge available on the system and write corresponding equations. This approach is said to be based on *first principles.* In this part we introduce the AO equations based on first principles. It introduces the key elements in an AO system and helps to understand the AO Matlab-based simulator. First we start with describing the wavefront sensor, and then the deformable mirror.

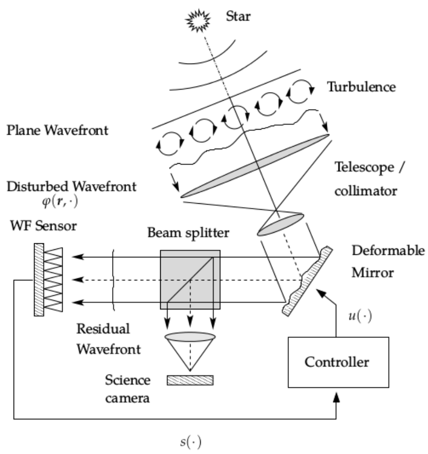

Figure 1: Schematic of an adaptive optics system. Courtesy: [3].

2.2.1 The Shack-Hartmann sensor

A Shack-Hartmann sensor is a sensor that measures the first derivative of the wavefront. It consists of a 2D-array of lenses, each of which project the wavefront on the camera located parallel to the lenses, at their focal distance. See Figure 2 for an illustration. A flat wavefront results in an image on the camera with dots well aligned one with another. However local tilts of the wavefront deviate the point where lights rays focus as displayed in Figure 3. These local tilts are measured and are called the slopes. These slopes are lifted in a vector which is denoted with $s(k)$.

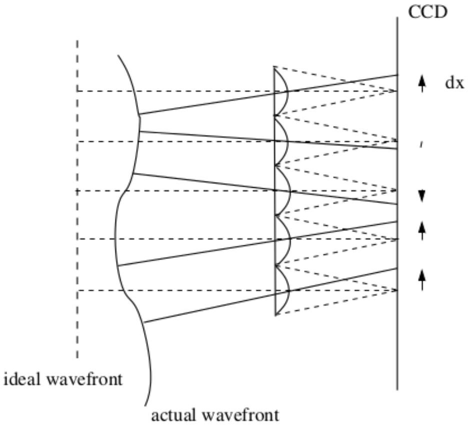

Figure 2: Schematic of an Shack- Hartmann array of lenses to measure the slopes of the wavefront aberration distributed in space. Courtesy: [3].

To clarify the ideas, we consider a square array of $p \times p$ lenslets. The wavefront is evaluated (sampled) as the corners of the subapertures of the Shack-Hartmann sensor. In other words, it is spatially sampled with $(p+1) \times (p+1)$ points:

$$\Phi(k) = \begin{bmatrix} \phi_{1,1}(k) & \dots & \phi_{1,(p+1)}(k) \\ \vdots & & \vdots \\ \phi_{(p+1),1}(k) & & \phi_{(p+1),(p+1)}(k) \end{bmatrix}, \qquad \phi(k) = \mathrm{vec}\Big(\Phi(k)\Big) \tag{1}$$

The slopes are related to the lifted wavefront at time instant k, $\phi(k)$, with the linear measurement equation:

$$s(k) = G\phi(k) + e(k) \tag{2}$$

where $G \in \mathbb{R}^{2p^2 \times (p+1)^2}$ and $e(k)$ is zero-mean white Gaussian noise and with covariance matrix $\sigma_e^2 I$. The matrix G contains +1,-1 entries and corresponds to a finite-difference operation.

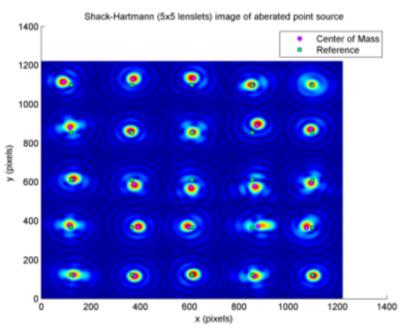

Figure 3: Simulated reading of the Shack-Hartmann sensor for an aberrated wavefront. Courtesy: [3].

2.2.2 The deformable mirror

We consider a deformable mirror with continuous face-sheets along with a square array of $m \times m$ actuators. The temporal dynamics of the mirror are neglected because the control frequency is much smaller than the first resonance frequency. Each actuator is coupled to its closest neighbors: the interaction with the neighborhood is modeled with a 2D Gaussian influence function, which is identical for each actuator. By setting the actuators to some values, a wavefront is induced by the mirror and is denoted as $\phi_{DM} \in \mathbb{R}^{m^2}$. A one-step delay is assumed between the time at which we apply the control inputs and the time at which the wavefront is induced, i.e the relationship between the control inputs and the induced wavefront is:

$$\phi_{DM}(k) = Hu(k-1) \tag{3}$$

where $H \in \mathbb{R}^{(p+1)^2 \times m^2}$. In the following we consider the special case where H is a square matrix, i.e $p+1 = m$. The matrix H is assumed invertible. The actuator signal $u(k-1)$ is the lifted vector of the 2D input data, with the same reordering as in (1). It means the actuators are located at the same position as the phase points. This particular actuator-sensor geometry is called the *Fried geometry.*

NB: For more details on the AO system, you can read e.g the reference [2] or attend the course sc4045: Control for High Resolution Imaging.

3 Static wavefront reconstruction

1. [**Chapter 2-4.**] We first collect wavefront sensor data $\{s(k)\}_{1..N_t}$ in open-loop. We would like to reconstruct the wavefront. Determine a closed-form expression for the estimate of the wavefront $\phi(k)$ using the measurement $s(k)$ in the noise-free case. Is the solution unique?

2. [**Chapter 2-4.**] We now take into account the prior knowledge of the statistical properties. Determine a closed-form expression for the estimate of the wavefront $\phi(k)$ using the measurement $s(k)$, the noise covariance σ_e and the prior wavefront information: $E[\phi(k)] = 0$ and $E\left[\phi(k)\phi(k)^T\right] = C_\phi(0) > 0$.

3. [**Misc.**] Is the parameter σ_e known in practice? What role does it play for reconstructing the wavefront in the above least-squares?

4. [**Chapter 2-4.**] We first assume a random-walk propagation of the atmosphere, i.e the estimate of the atmospheric wavefront at time $k+1$ is such that:

$$\phi(k+1) = \phi(k) + \eta(k) \tag{4}$$

   where $\eta(k)$ is a zero-mean white Gaussian noise with identity covariance matrix. Based on this prediction model, we can now study how to close the loop. The residual wavefront is defined with:

$$\epsilon(k) = \phi(k) - \phi_{DM}(k) \tag{5}$$

   The two following assumptions are made (in this section only):

$$E[\epsilon(k)] = 0, \qquad E\left[\epsilon(k)\epsilon(k)^T\right] = C_\phi(0)$$

   In adaptive optics, we are interested in minimizing the variance of the residual wavefront to obtain the highest image resolution as possible. To do so, determine the optimal input commands $u(k)$ as a function of the estimated residual $\epsilon(k)$, the past command $u(k-1)$ and the matrices H and C_ϕ from a least-squares problem.

5. [**Chapter 2-4.**] Denote $\delta u(k) := u(k) - u(k-1)$. The control matrix M relates the increment in the (lifted) deformable mirror commands $\delta u(k)$ to the slopes measurements, such that

$$\delta u(k) = Ms(k)$$

   See the schematic in Figure 1. Give an expression for M as a function of the system matrices.

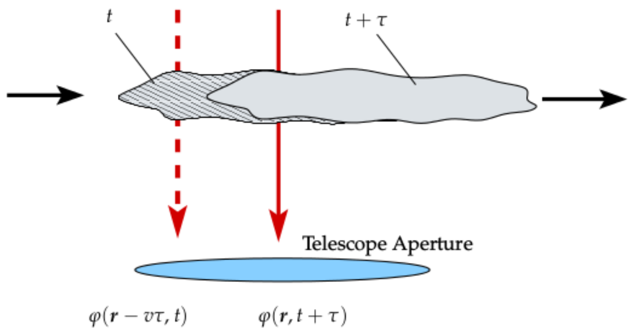

Figure 4: A layer of frozen turbulence is passing over the telescope aperture. Courtesy: [3].

6. [**Matlab.**] Write the Matlab routines to close the loop as depicted in the schematic in Figure 1. It is asked to insert the code in a function entitled:

$$[var_{eps}] = \text{AOloopMVM}(G, H, C_\phi(0), \sigma_e, \phi)$$

where $\phi \in \mathbb{R}^{(p+1)^2 \times N_t}$ and var_{eps} is the variance of the residual wavefront when taking N_t time points within the closed-loop operation.
Hint: It is important to remove the mean of the residual wavefront before computing its variance. Explain why. You can use the Matlab function awgn to add white Gaussian noise on the slopes.

7. [**Matlab.**] Compare the performances of this method with the case when no control is applied.

8. [**Misc.**] For large-scale systems, it is important that the number of elementary operations $(+, \times)$ to be performed online doesn't sharply increase when the number of actuators m^2 or the number of sensors measurements $2p^2$ increases. How many of these elementary operations are necessary for the online computation of the PI controller of question 6? Conclude on the scalability of the method.

4 Vector Auto-Regressive model of Order 1

The method studied in Section 3 ignores all kind of spatio-temporal correlation. It is nonetheless reasonable to assume that the turbulence flows over the telescope aperture with a given pattern, see Figure 4 for an illustration. The time it takes for the turbulence to cross the line of sight of the telescope is way longer than a sampling period. Therefore in this section we study a Vector Auto-Regressive (VAR) model as first approximation:

$$\phi(k+1) = A\phi(k) + w(k) \tag{6}$$

where $A \in \mathbb{R}^{(p+1)^2 \times (p+1)^2}$, and $w(k) \sim \mathcal{N}(0, C_w)$. Moreover, the signal $w(k)$ is uncorrelated with both the measurement noise $e(k)$ and the turbulent wavefront $\phi(k)$:

$$E\left[w(k)e(k)^T\right] = 0, \qquad E\left[w(k)\phi(k)^T\right] = 0 \tag{7}$$

1. [**Misc.**] Denote $C_\phi(0) = E\left[\phi(k)\phi(k)^T\right]$ and $C_\phi(1) = E\left[\phi(k+1)\phi(k)^T\right]$. Relate $C_\phi(0), C_\phi(1)$ and A from the V-AR1 model (6). Determine a closed-form expression for A assuming the covariance information is known.

2. [**Misc.**] We assume the wavefront $\phi(k)$ is a wide-sense stationary signal. Derive a relationship between $C_\phi(0), C_w$ and A in order to estimate the covariance matrix C_w.

3. [**Misc.**] From the equations (2), (3), (5) and (6), formulate a state-space model in which the state is equal to the closed-loop residual wavefront $\epsilon(k)$.

4. [**Chapter 5.**] Write an expression for the Kalman filter associated with the above state-space model.

5. [**Misc.**] We wish to minimize in 2-norm the residual wavefront ϵ at time k+1. What is the control law?

6. [**Matlab.**] Write the Matlab routines to compute A, C_w and the Kalman gain K and embed it into:
$$[A, C_w, K] = \text{computeKalmanAR}(C_\phi(0), C_\phi(1), G, \sigma_e)$$
Write the Matlab code to close the loop and embed the code in the function:
$$[var_{eps}] = \text{AOloopAR}(G, H, C_\phi, \sigma_e, A, C_w, \phi)$$

7. [**Misc.**] Analyze the pattern of the matrix A estimated. Give a condition on A such that the AR-1 assumption is equivalent to the turbulence model used for MVM control. Compare then the control actions theoretically.
Hint: use can be made of the Matlab function imagesc.

5 Subspace identification

We now consider that the turbulence is modeled by a stochastic state-space model:

$$x(k+1) = Ax(k) + Kv(k) \tag{8}$$
$$\phi(k) = Cx(k) + v(k) \tag{9}$$

where $v(k)$ is a zero-mean white noise process with identify covariance matrix.

1. [**Chapter 9.**] Write the data equation and explain how to estimate the system matrices with the N4SID method.

2. [**Matlab.**] Write the Matlab routines to identify the matrices A, C, K from open-loop wavefront sensor data with the N4SID method. Embed the code in the file:
$$[A, C, K, \text{vaf}] = \text{n4sid}(sk, N_{id}, N_{val}, s, n)$$
where sk are the pre-processed sensor measurements according to equation (2) in [4], N_{id} and N_{val} are the number of points used respectively for identification and validation. s is the upper bound on the order n. vaf is the Variance Accounted For between the validation set and the reconstructed output based on the matrices estimation.
Hint: You need to first generated some noisy sensor measurements, then reconstruct the wavefront using the method studied in Section 3, and then use this data for identification. An ad-hoc function for static-wavefront reconstruction shall be created.

3. [**Chapter 2.**] Propose a control algorithm using the matrices A, C, K estimated in the previous question to retrieve the control inputs that minimize the criteria:
$$\min_{u(k)} \|\widehat{\epsilon}(k+1)\|_2^2 + \lambda \|u(k)\|_2^2$$
where λ is a tuning parameter.

4. [**Matlab**] Write the Matlab routines for the latter algorithm and embed the code in the file:
$$[var_{eps}] = \text{phiSid}(G, H, A, K, C, SNR, \lambda, \phi)$$

6 Comparison and critical thinking

1. [**Matlab.**] Compare the closed-loop performances of the three methods studied in this assignment, that corresponds to 3 different turbulence models: random walk (4), V-AR1 (6), state-space (8).

2. [**Misc.**] Discuss the different models used for turbulence modeling and how they affect the closed-loop control performance. Can you infer from the structure of the matrix A or the Markov parameters $C(A-KC)^iK$ the number of turbulence layers used?

3. [**Misc.**] Throughout the exercise, we have derived methods for minimizing the 2-norm of the predicted wavefront that is coming from the atmosphere. The Point-Spread function is the sharpest when the wavefront has no aberrations. We now wonder whether it is necessary to reconstruct the wavefront. Derive a method inspired from MVM for minimizing the residual slopes at time $k+1$ rather than the residual wavefront. Compare it numerically with the results obtained in 6.1 and comment.

References

[1] M. Verhaegen, V. Verdult. "Filtering and System Identification: a least-squares approach", Cambridge University Press, 2007.

[2] F. Roddier. "Adaptive Optics in Astronomy", Cambridge U. Press (1999).

[3] M. Verhaegen, G. Vdovin, O. Soloviev. "Control for High Resolution Imaging", Lecture notes sc4045, TU Delft, 2016.

[4] K. Hinnen, M. Verhaegen, N. Doelman. "A Data-Driven $\mathcal{H}2$ Optimal Control Approach for Adaptive Optics", IEEE Trans. on Control Systems Technology, Vol. 16, No. 3, 2008.