Machine Learning for Aero-Optical Wavefront Characterization and Forecasting

Shervin Sahba

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:
J. Nathan Kutz, Chair
Arka Majumdar
Shih-Chieh Hsu

Program Authorized to Offer Degree:
Physics

University of Washington

Abstract

Machine Learning for Aero-Optical Wavefront Characterization and Forecasting

Shervin Sahba

Chair of the Supervisory Committee:
J. Nathan Kutz
Department of Applied Mathematics

The laser is a masterwork of the previous century of physics, but to harness its power coherently in the turbulent wilds of the sky remains a challenge. Free-space lasing hosts myriad applications, from direct energy transmissions for defense to secure communication channels that robustly support quantum entanglement. Aero-optics is the multi-disciplinary field underpinning such optical transmissions through atmosphere and fluid flows. Turbulent variations in the index of refraction, like those forming around boundary layers of airborne optical platforms, manifest aberrated wavefronts. Forecasting these rapid phase distortions allows us to rectify laser transmissions via adaptive optic engineering. Doing so hinges on the development of low-latency predictive techniques, for which we turn to advancements in data-driven algorithms and deep learning.

This thesis thus introduces key concepts of aero-optics, wavefront sensing, and machine learning for data-driven physics. We then demonstrate three machine learning methodologies - optimized Dynamic Mode Decomposition, sensor fusion through shallow decoder networks, and forecasting via recurrent neural networks with shallow decoder outputs — for robust aero-optical wavefront sensing.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

AAOL-T: Airborne Aero-Optics Laboratory Transonic

AEL: Aero Effects Laboratory

AO: Adaptive Optics

DH: Digital Holography

DHWFS: Digital Holography Wavefront Sensor

DMD: Dynamic Mode Decomposition

FFT: Fast Fourier Transform

LSTM: Long Short-Term Memory

NN: Neural Network

OPD: Optical Path Difference

OPL: Optical Path Length

PCA: Principal Component Analysis

POD: Proper Orthogonal Decomposition

RELU: Rectified Linear Unit

RNN: Recurrent Neural Network

SDN: Shallow Decoder (Neural) Network

SHWFS: Shack Hartmann Wavefront Sensor

SHRED: Shallow Recurrent Decoder

SH: Shack Hartmann

WFS: Wavefront Sensor

ACKNOWLEDGMENTS

I must admit that this has been a journey, and like most voyages there were highs and lows. Ultimately, the difference between riding a crest versus a trough relied less on the ship and more with the people onboard. Thus for the journey through my doctorate, I first and foremost would like to thank my advisor, Nathan Kutz, for providing a non-stop dynamo of enthusiasm and a remarkable well of patience. Nathan's encouragement and optimism went a long way, probably further than he has realized. I hope to pay it forward. Second, I would like to thank my coauthors and colleagues at the Air Force Research Laboratory, without whose insight and shared discussions the bulk of this work could not have been created. Chris Wilcox, especially, I thank for his candor, support, and mentorship.

To my friends and family: you know who you are, and I thank you all. My parents were the ones who gave me this opportunity, among so many others. I thank my dad for inspiring me to seek knowledge, tinker, and build. I thank my mom for those things and everything else too. My successes are the product of hers. Lastly, I thank my fiancée, Megan, for every ounce of kindness and every quiet wisdom. I am so very grateful.

DEDICATION

to Bahman

Chapter 1

AERO-OPTICS

To fly, we have to have resistance. It's all about turbulence. Reacting to images of wave patterns in fluid motion.

—Maya Lin

Aero-optics, at its core, is the study of how light propagates through turbulent air flows. It is a field that sits at the intersection of fluid dynamics and optics, and has garnered significant attention due to its implications for optical systems deployed in high-speed environments, such as aircraft to aircraft or aircraft to ground transmissions, directed energy systems, and free-space optical communication systems.

The term aero-optics refers to effects on the optical field from a high-speed turbulent flow, where air is forced around the optical platform, possibly resulting in flow separation and shock formation. [48] The aero-optic effects that determine wavefront aberrations are dependent on such factors as the laser platform shape and aircraft geometry, the speed of the craft, the Reynolds number, and the direction of beam propagation, which will be further discussed in context of the Airborne Aero-Optics Laboratory Transonic experiment in Chapters 4 and 6.

In the most general sense, electromagnetic wave propagation is determined by the Maxwell equations [55], and, when the medium is a fluid, the compressible Navier-Stokes equations [76, 71]. Depending on the relevant length and time scales of the problem, differing simplifications can be made. In aero-optical problems, the optical propagation duration is negligibly small relative to flow advection timescales, which leads to solutions in a frozen flow field at each moment in time. The smallest flow length-scale, the Kolmogorov scale [35],

typically far exceeds the optical wavelength of interest so polarization effects can be safely ignored. [101] The electromagnetic field components are thereby decoupled, and scalar components for a given frequency ω may be expressed time-independently as a Helmholtz wave equation

$$\boldsymbol{\nabla}^2 U + \frac{\omega^2 n^2}{c_0^2} U = 0 \tag{1.1}$$

where c_0 is the speed of light in vacuum and n is the index of refraction. The latter is directly linked to air density fluctuations by

$$n\left(\mathbf{r}\right) = 1 + K_{GD}\left(\lambda_0\right)\rho\left(\mathbf{r}\right), \tag{1.2}$$

where K_{GD} is the wavelength-dependent Gladstone-Dale constant, λ is the optical wavelength upon which K_{GD} weakly depends, and $\rho\left(\mathbf{r}\right)$ is the air density as a function of the spatial variable $\mathbf{r}$. [101]

Despite the index of refraction's central role in aero-optical aberrations, its fluctuations are small ($\delta n \sim 10^{-4}$) and are of a larger scale than the optical wavelength, λ. Therefore off-axis propagation is negligible and the paraxial wave approximation is customarily invoked to model laser in aero-optical environments. That is, we assume

$$U\left(\mathbf{r}\right) = A\left(\mathbf{r}\right) e^{-ikz} \tag{1.3}$$

and

$$\left|\frac{\partial^2 A}{\partial z^2}\right| \ll \left|k\frac{\partial A}{\partial z}\right| \tag{1.4}$$

where $k = \omega / c_\infty$ is the free stream wave number. This leads to the paraxial wave equation

$$-2ik\frac{\partial A}{\partial z} + \boldsymbol{\nabla}_\perp^2 A + k^2\left(\frac{n^2}{n_\infty^2} - 1\right)A = 0. \tag{1.5}$$

without any terms for large-angle scattering, dissipation, or absorption which are negligible in aero-optics. [76] Several approximations [84] can be made to pursue solutions of Equation 1.5. For our purposes, we will proceed by considering two regimes of aero-optical effects: the turbulent boundary layer in the near-field and the free space extending to the far-field.

1.1 Aero-Optical Effects and Turbulent Boundary Layers

Two distinct regions exist in terms of aero-optical effects: the turbulent near field from $z = 0$ to z_{TBL} and the free atmosphere extending to the far field from $z = z_{TBL}$ to L with $L \gg z_1$. For the turbulent boundary layer (TBL) region, if we assume diffraction, modeled by the $\boldsymbol{\nabla}_{\perp}^2 A$ term, is negligible, and take the linear approximation $\frac{n^2}{n_\infty^2} \approx 1 + 2\frac{n'}{n_\infty}$, then integration of Equation 1.5 gives

$$U\left(x, y, z_{TBL}\right) = U\left(x, y, 0\right) \exp\left(-ik_0 \int_0^{z_{TBL}} n\left(\mathbf{r}\right) \mathrm{d}z\right) \tag{1.6}$$

with $k_0 = \omega/c_0$ being the wave number in vacuum. The equation suggests that the dominant aero-optical effects after transmission through the TBL are phase aberrations of the wavefront, while the amplitude is roughly unperturbed.

The aerodynamic environment of airborne laser platforms motivates high-fidelity computational fluid dynamics models and experimental techniques in the study of aero-optics effects. The high-speed, high Reynolds number compressible flows around airborne platforms can contain TBLs, shear layers, and wakes, as well as shock waves in the case of transonic and supersonic flows. [31] As a laser beam propagates through this turbulent flow surrounding the aperture, refractive index fluctuations cause phase aberrations, and the resulting distortions of the optical field are referred to as aero-optical effects [104].

Characterizing propagating beam wavefront dynamics in the TBL is critical to correcting the outgoing phase profile of the beam. From an applied standpoint, despite the relatively short, centimeter scale, distance traveled in the TBL, the beam quality is immediately and often heavily degraded within this region. [103] A typical method to quantify the aero-optic wavefront aberrations from a given refractive index field is by calculating optical path difference (OPD). OPD is computed by first calculating the OPL, which is proportional to the travel time for corresponding rays. OPL, as seen in Equation 1.6, is given as

$$\mathrm{OPL}\left(x, y, t\right) = \int_0^{z_1} n\left(x, y, z, t\right) dz. \tag{1.7}$$

Subtracting the mean OPL over the spatial coordinates of the aperture produces the OPD,

$$\mathrm{OPD}\left(x,y,t\right)=\mathrm{OPL}\left(x,y,t\right)-\left\langle \mathrm{OPL}\left(x,y,t\right)\right\rangle . \tag{1.8}$$

1.1.1 Far-field Distortions

The OPD can be decomposed to facilitate further analysis with

$$\mathrm{OPD}\left(x,y,t\right)=\left\langle \mathrm{OPD}\right\rangle \left(x,y\right)+\left(A\left(t\right)x+B\left(t\right)y\right)+\mathrm{OPD}_{\text{high-order}}\left(x,y,t\right) \tag{1.9}$$

where the angle brackets denote a time averaging of the OPD giving a steady defocus term, the functions $A\left(t\right)$ and $B\left(t\right)$ define a tilt that does not change the spatial distributions of the beam but simply redirects it in the given directions, and a high-order term that causes changes in the beam shape and intensity distributions. [102] The steady defocusing term (also known as a coma) and the tilt terms do not change the spatial distribution of the outgoing beam. Rather, the higher-order OPD fluctuations account for aberrations of the wavefront shape and intensity distribution.

1.2 Adaptive Optics

Let us picture the perfect scenario for any optical system. Imagine a flat, calm, undistorted wavefront – the optical equivalent of a day at the beach. But now, add some windy conditions (like turbulent air flow over aircraft windows or turrets in aero-optics) and all of a sudden your day at the beach isn't so calm anymore. These unruly gusts of wind or, technically speaking, density variations in the air flow, cause changes in the refractive index of air. This results in optical aberrations or a less-than-perfect wavefront.

Adaptive optic techniques [99] are used to correct for such optical distortions in real-time, and the study applies to a variety of domains and scales ranging from microscopy applications [10] to astronomy [26]. Adaptive optic techniques often involve deformable mirrors, whose shape mimics the predicted conjugate wavefront in order to rectify the outgoing beam. Wavefront information is processed in real-time predictor to control the deformable mirror shape and counteract the optical distortions caused by the aero-optic phenomena.

The ideal scenario for most optical system is to have a flat, undistorted wavefront. In the field of aero-optics, these distortions often come from the turbulent boundary layer around the airborne optical platform. The density variations in the airflow cause changes in the refractive index of air, resulting in optical aberrations. The goal of an adaptive optics system is to correct these wavefront aberrations, allowing the optical system to approximate the performance it would have in ideal conditions. This is done by using a deformable mirror or spatial light modulator that can change shape in real time, often on the MHz scale, guided by a control algorithm. The control algorithm is driven by data from a wavefront sensor, which measures the distortions in the incoming wavefront.

The adaptive optics system, like that sketched in Figure 1.1 operates in a feedback loop:

1. Wavefront Sensing: First, the distorted wavefront is measured using a wavefront sensor. This could be a Shack-Hartmann sensor, pyramid sensor, or other types of wavefront sensors, depending on the specific requirements of the system. The sensor measures the deviations from the ideal wavefront and converts these deviations into a set of aberration coefficients, typically using Zernike polynomials [100, 77].

2. Deformable Mirror Spatial Light Modulator [53]: The calculated aberration coefficients are used by the control system to determine the necessary shape of the deformable mirror. The mirror's shape is then adjusted in real time to correct the wavefront distortions.

3. Correction and Feedback: The corrected wavefront is then propagated through the system. Some of the light is sampled by the wavefront sensor to verify the accuracy of the corrections. If any residual errors are detected, they are fed back into the system, and the process is repeated until the wavefront error is minimized.

The effectiveness of an adaptive optics system lies in its ability to rapidly and accurately correct wavefront distortions. The temporal dynamics of the turbulence and the deformable

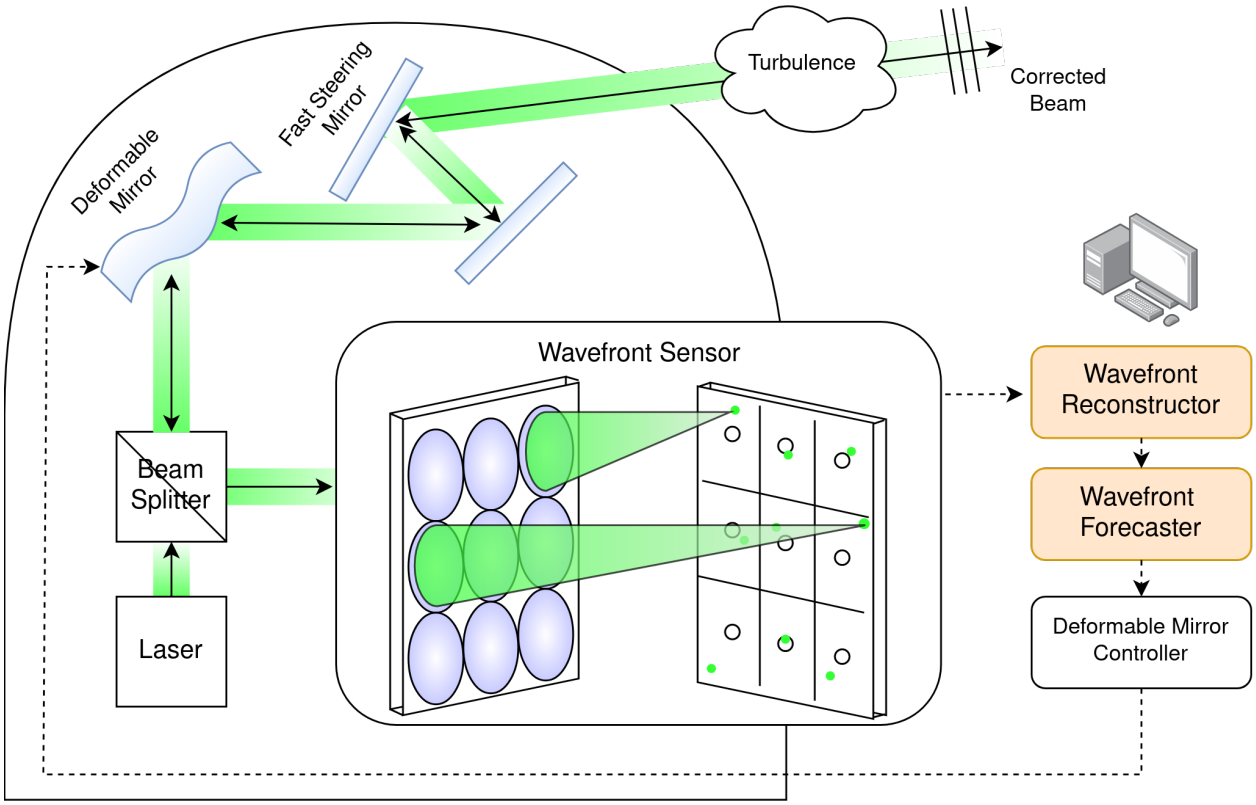

Figure 1.1: Adaptive optic beam control system diagram.

mirror's mechanical response limit the speed of correction. To increase this speed, machine learning techniques [93, 19] can be used to forecast the turbulence evolution, allowing for predictive, rather than reactive, corrections. Our goal is to increase the speed and accuracy of wavefront correction in aero-optic applications, ultimately enhancing the performance of optical systems operating in turbulent aerodynamic environments.

Chapter 2

SENSORS

Somewhere in me is a curiosity sensor. I want to know what's over the next hill. You know, people can live longer without food than without information. Without information, you'd go crazy.

—Arthur C. Clarke

Underpinning modern aero-optical research are the different advancements in wavefront sensor (WFS) technology. Beginning in 1992 with the Malley Probe [74], the first aero-optical sensors operated under the assumption that wavefront aberrations convected downstream simultaneously with the aberrating turbulent flow structures abiding by Reynolds flow theory. Named the Malley Principle, this assumption proved to be simple in hindsight but a powerful tool at the time to create a single-aperture laser-based WFS [59]. Using the same physics as Hartmann sensing[68], the beam's time-varying refraction as it propagates through the turbulent shear layer is recorded to determine wavefront slope and aberrating structure convection speed. The method allowed for 1-D wavefront sensing, as depicted in Figure 2.1. While the Malley Probe is a ubiquitous tool now in aero-optical wavefront sensing [21], desirable due to its minimally-invasive nature and capture rates up to 1 MHz, the desire to create 2-D high resolution WFS data led to Shack-Hartmann wavefront sensing, data from which is used in each of this dissertation's studies.

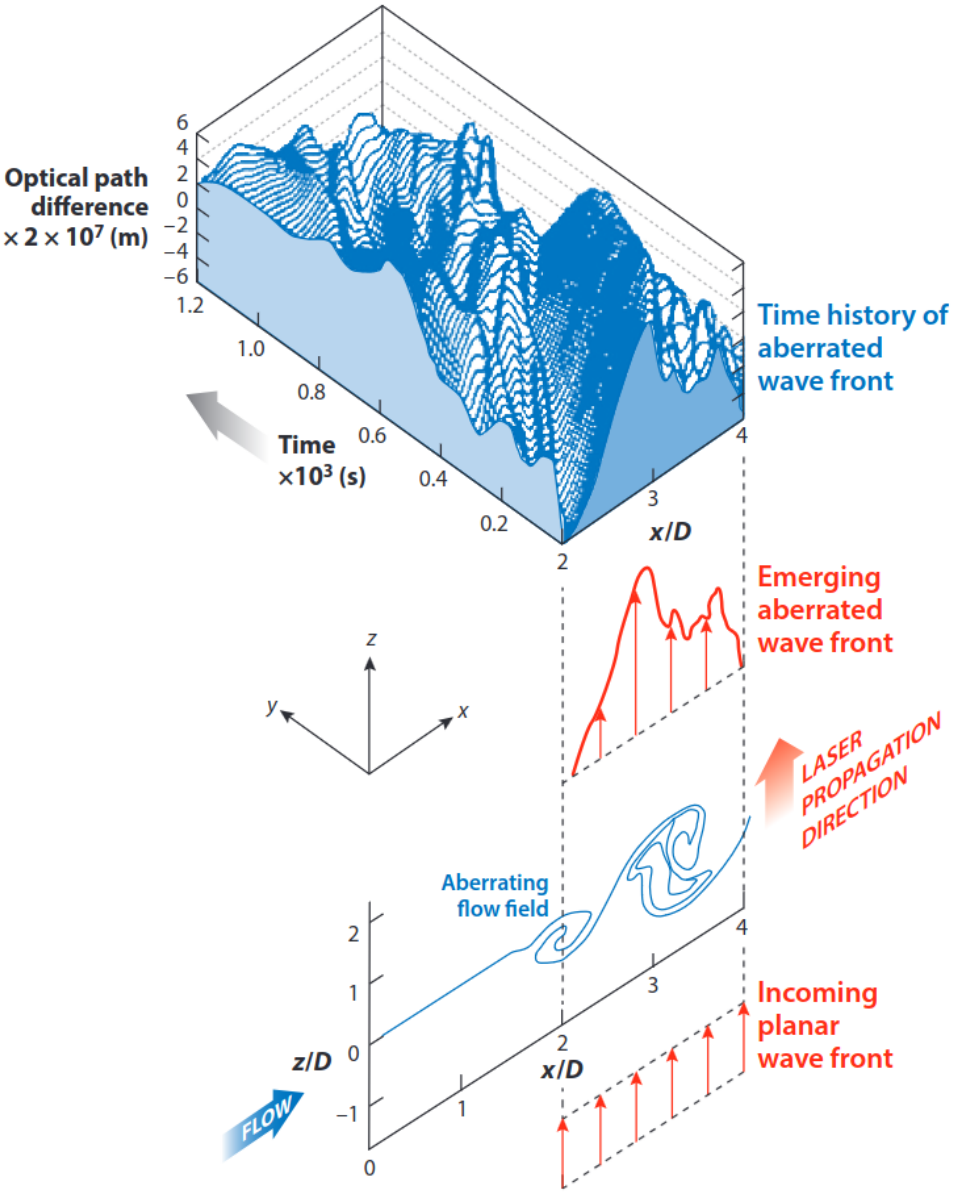

Figure 2.1: Illustration depicting a planar wavefront propagating through a flow field and subject to aero-optical effects and the accompanying time series of 1-D aberrated wavefronts. *Figure adapted from References [24, 57].*

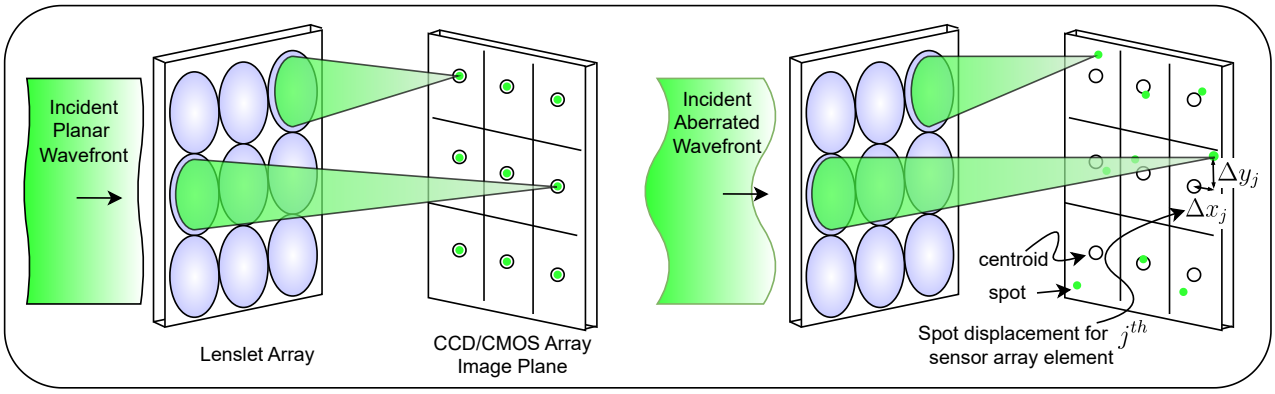

Figure 2.2: Shack Hartmann WFS operational schematic. *Left*: Incident planar wavefronts transit through the lenslet array onto the image plane, individual rays forming spots that coincide with CCD/CMOS array element centroids. *Right*: Aberrated wavefronts result in spot displacements from the centroids, $(\Delta x_j, \Delta y_j)$, on the image plane.

2.1 Shack-Hartmann Wavefront Sensor

A Shack-Hartmann wavefront sensor (SHWFS) [79] is a device used to measure 2-D wavefront time series. It's commonly used in adaptive optics systems, such as those in astronomical telescopes and in ophthalmology to measure aberrations of the human eye. The working principle of a SHWFS is based on dividing the incoming wavefront into a grid of smaller sections, each of which is independently analyzed. This is achieved through a lenslet array — a grid of small lenses — that splits the incoming light into multiple beamlets in order to sample the incident wavefront, as seen in Figure 2.2.

Each of these beamlets strikes a detector, typically a CCD (Charge-Coupled Device) or CMOS (Complementary Metal-Oxide-Semiconductor) sensor. If the incoming wavefront is a plane wave and thus free of distortion, the focal spots of these beamlets form a regular grid pattern on the SH detector array. However, if the incoming wavefront is not a plane wave but has some shape due to aberrations in the optical path, each lenslet still focuses the light to a point, but the point is now shifted relative to the expected grid position. By measuring the

amount and direction of these spot displacements for each lenslet and corresponding sensor array element, we can build up a map of the local slopes of the wavefront across the entire lenslet array. This map quantifies the distortions present in the wavefront, which can then be used to correct the optical system. The most common way to reconstruct the shape of the entire wavefront is probably least-squares fitting of Zernike polynomials. The reconstruction can then be passed to an AO system.

Thus, say the wavefront phase is given by $\phi(x, y)$ and the focal length of the lenslets, or distance from the lenslets to the image plane, is L, then sub-aperture local slope measurements are given by

$$\boldsymbol{\nabla}\phi_j(x_i, y_j) = \Delta_j(x_i, y_j)/L \tag{2.1}$$

Because this spatial resolution is rather low, equal to the number of array elements, a Zernike polynomial basis, matrix $\mathbf{Z}$, is often used to reconstruct the wavefronts with

$$\phi = \mathbf{Z}a \tag{2.2}$$

where coefficients a could be solved via a L_2 minimization problem

$$a = \arg\min_a \|\boldsymbol{\nabla}\mathbf{Z}a - \Delta/L\|^2 = \left((\boldsymbol{\nabla}Z)^T \boldsymbol{\nabla}Z\right)^{-1} (\boldsymbol{\nabla}Z)^T \Delta/L \tag{2.3}$$

Now this Zernike-fitted wavefront's conjugate can be mapped onto a deformable mirror for adaptive optic control. For reference, the Zernike polynomials are an orthogonal sequence on the unit disk given by

$$Z_n^m(\rho, \varphi) = R_n^m(\rho)\cos(m\varphi) \tag{2.4}$$

$$Z_n^{-m}(\rho, \varphi) = R_n^m(\rho)\sin(m\varphi) \tag{2.5}$$

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k!\left(\frac{n+m}{2}-k\right)!\left(\frac{n-m}{2}-k\right)!}\rho^{n-2k}. \tag{2.6}$$

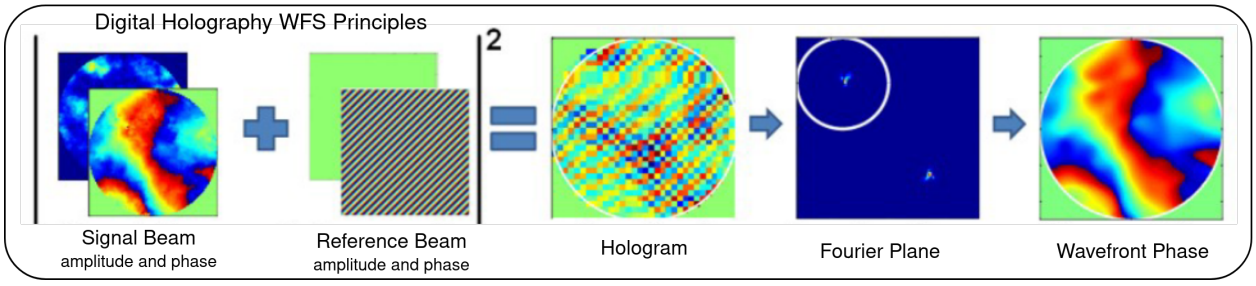

Figure 2.3: Digital Holography WFS operational schematic.

2.1.1 Shock waves and the SHWFS

One potential shortcoming of the SHWFS is in the presence of shock waves. As shown in 2023 work by Kalensky, Kemnetz, and Spencer, [60], shock-induced phase discontinuities may alter the irradiance pattern on the SHWFS image plane. The SHWFS operates on the assumption that each lenslet focuses a single ray, yet shocks may cause a beam bifurcation that does not map as a proportional shift to the image. For phase discontinuities in excess of $\pi/2$, the SHWFS begins to severely underestimate the resultant OPD. Issues concerning this detail may have arisen in the study in Chapter 5.

2.2 Digital Holography Wavefront Sensor

Digital holography wavefront sensing is a technique that enables the measurement of the phase and amplitude of light waves, which is essential for several applications in optics, including imaging, microscopy, and beam shaping. The technique is based on the principles of holography, a method for recording and reconstructing the wavefront of light.

A signal beam of light from the object of interest interferes with a known reference wave. That interference pattern contains both the amplitude and phase information of the object wave, essentially creating a hologram of the wavefront. A digital sensor, typically a CCD (Charge Coupled Device) or CMOS (Complementary Metal-Oxide-Semiconductor) camera, are used to record the interference pattern. The recorded hologram is a 2-D intensity image

that encodes the 3D information of the object wavefront. Using Fourier analysis, the phase information can be reconstructed, providing the full complex field of the signal wave. This process is depicted in Figure 2.3.

In the context of aero-optics, for example, a DHWFS can be used to sense and correct for the aberrations induced by the TBL on an optical beam propagated through it. The signal beam's electric field, U_S, interacts with the volume of interest and carries phase and amplitude information, while the reference beam's field, U_R, is usually undisturbed. When re-collimated and captured on a CCD or CMOS sensor, the resulting interference pattern creates a hologram whose intensity, I_H, is derived from the superposition of the beam's signal and reference electric fields such that

$$I_H = |U_S + U_R|^2 = I_S + I_R + U_S U_R^* + U_S^* U_R, \tag{2.7}$$

from which the phase of the signal beam can be extracted via Fourier analysis.

2.3 Sensor Fusion

Sensor fusion refers to the process of combining data from different sensors to improve the performance of a system. The aim of sensor fusion is to consolidate the information gathered from multiple sensors, each with its own operational modality, in order to provide a comprehensive understanding of the environment or system under scrutiny.

Sensor fusion is grounded in the core belief that the use of multiple sensors can provide more accurate, reliable, and comprehensive information than a single sensor operating in isolation. This notion is particularly apt in the context of aero-optic wavefront control, where the complexities of fluid dynamics, the variability of environmental conditions, and the high-speed nature of the problems at hand all conspire to make accurate prediction and control a challenging task.

The integration of data from various sensors isn't a straightforward task, and it requires effective computational algorithms. This is where machine learning comes into the picture. Machine learning algorithms, particularly those based on neural networks, have proven to be

incredibly adept at the task of sensor fusion. These algorithms can learn to extract features from raw sensor data, find correlations and patterns among the data from different sensors, and integrate this information into a cohesive whole.

In aero-optic wavefront control, these algorithms might be used to learn the relationships between temperature gradients, pressure variations, and wavefront distortions. Once these relationships are learned, they can be used to predict future wavefront distortions under different conditions and guide the control systems to preemptively mitigate these distortions.

Chapter 3

MACHINE LEARNING

In three years, Cyberdyne will become the largest supplier of military computer systems. All stealth bombers are upgraded with Cyberdyne computers, becoming fully unmanned. Afterwards, they fly with a perfect operational record. The Skynet funding bill is passed. The system goes online on August 4th, 1997. Human decisions are removed from strategic defense. Skynet begins to learn at a geometric rate. It becomes self-aware 2:14 AM, Eastern time, August 29th.

—The Terminator

I visualize a time when we will be to robots what dogs are to humans. And I am rooting for the machines.

—Claude Shannon

In this chapter we highlight several methods, both general and specific, to leverage data-driven algorithms for physics discovery. Notably we discuss

- **Dynamic Mode Decomposition (DMD)**: a nonlinear dynamic model reduction framework [14], where time series data may be decomposed into spatial and temporal modes. A variety of DMD algorithms exist, including specialized tools outfitting compressed sensing [18] and physical constraints [6]. We focus on the classic, or exact, DMD method and compare it to the Optimized DMD method [3].

- **Supervised Learning via Shallow Decoder Networks (SDNs)**: minimally layered feed-forward neural networks for supervised learning in regression based frameworks.

- **Supervised Deep Learning via Recurrent Neural Networks (RNNs)**: bi-directional, cell based neural networks like the LSTM that are capable of tracking temporal changes in time series data for supervised learning tasks.

3.1 Dynamic Mode Decomposition

DMD is a feature extraction and dimensionality reduction framework for time series data, first developed by Peter Schmid [88] for use in the fluid dynamics communities. The DMD algorithm efficiently approximates the best-fit linear matrix operator, $\mathbf{A}$, that advances high-dimensional measurements of a system forward in time [65].

Suppose we are studying a linear, time-invariant dynamical system governed by

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}\mathbf{x}, \tag{3.1}$$

then DMD, when provided with vectorized snapshots of time series data, $\mathbf{x}_j = \mathbf{x}(t_j)$ for $j \in \mathbb{N}_{>0}$ and $\mathbf{x} \in \mathbb{R}^n$, computes a regression over snapshots to seek the best fit linear matrix operator $\mathbf{A}$ that evolves states forward in time. A reduced-order model is often then determined by rank truncating the operator, $\mathbf{A}_r$ with $r = \mathrm{rank}(\mathbf{A}_r) < \mathrm{rank}(\mathbf{A})$.

Stemming from the composition operator, $\mathbf{A}$, are modal decompositions and accompanying oscillation frequencies or decay rates. Like the Fast Fourier Transform (FFT), DMD provides a form of frequency analysis, and like Principal Component Analysis (PCA) [12], DMD generates a set of spatial modes. It is, in a sense, providing the best of both worlds. DMD's modal decomposition is non-orthogonal, yet the modes often have value in their spatial interpretability. For purely linear systems, the modal decomposition are analogous to normal modes of the system, but in general, the eigen-decomposition approximates the otherwise infinite-dimensional Koopman operator, $\mathcal{K}$. Indeed, underpinning DMD is Koopman theory, only within the DMD frame we are restricted to the measurement subspace spanned by the directly measurements snapshots, $\{\mathbf{x}_j\}$ [15, 63].

Computing DMD is efficient. DMD is based on the singular value decomposition (SVD) [42], which also runs by the name of proper orthogonal decomposition (POD) in fluid dynamics

communities. A typical SVD computation takes $\mathcal{O}(mn^2)$ floating-point operations for a $m \times n$ array, assuming $m \geq n$, first reducing the matrix to bi-diagonal form, often via Householder reflections, and then performing an eigenvalue computation via a QR algorithm. [96]. The first step in performing DMD, though, is collecting sequential measurements, vectorizing them, and assembling a pair of time-shifted snapshot matrices given by

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_{m-1} \\ | & | & & | \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} | & | & & | \\ \mathbf{x}_2 & \mathbf{x}_3 & \dots & \mathbf{x}_m \\ | & | & & | \end{bmatrix} \tag{3.2}$$

for m snapshots, $\mathbf{x}_j$, with n features each. The exact DMD algorithm, as referred to by Tu et. al [98] as the classical DMD formulation, typically assumes that snapshots have a discrete temporal spacing such that $\mathbf{x}_j = \mathbf{x}(j\delta t)$, though they are not required to be of this form. Because data is likely sampled from a nonlinear dynamic system, we solve the locally linear approximation

$$\mathbf{Y} \approx \mathbf{A}\mathbf{X} \tag{3.3}$$

to find

$$\mathbf{A} \equiv \mathbf{Y}\mathbf{X}^{\dagger} = \underset{\mathbf{A}}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F \tag{3.4}$$

where the † symbolizes the Moore-Penrose pseudoinverse which minimizes the error via the Frobenius norm $\|\cdot\|_F$. In practice, when the n features of each state is large, the $\mathbf{A}$ may be intractable. Therefore DMD circumvents its decomposition by calculating a rank-reduced representation using a projection of the matrix onto POD modes, $\tilde{\mathbf{A}}$. The eigenmode matrix, $\mathbf{W}$, and diagonal eigenvalue matrix, $\mathbf{\Lambda}$, of $\tilde{\mathbf{A}}$ deliver the DMD modes and eigenvalues of $\mathbf{A}$. The eigenvalues remain the same, and the DMD modes may be verified to be

$$\mathbf{\Phi} = \mathbf{Y}\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{W} \tag{3.5}$$

Algorithm 1 shows the procedure as pseudocode.

Algorithm 1 Dynamic Mode Decomposition

$\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_{m-1}]$ ▷ snapshot matrix

$\mathbf{Y} = [\mathbf{x}_2, \ldots, \mathbf{x}_m]$ ▷ time-shifted snapshot matrix

function DMD$(\mathbf{X}, \mathbf{Y}, r)$

  $\mathbf{U}, \mathbf{\Sigma}, \mathbf{V} \leftarrow \text{SVD}(\mathbf{X}, r)$ ▷ r-rank truncated SVD

  $\tilde{\mathbf{A}} \leftarrow \mathbf{U}^*\mathbf{Y}\mathbf{V}\mathbf{\Sigma}_r^{-1}$ ▷ POD mode projected matrix, where $\tilde{\mathbf{A}} = \mathbf{U}^*\mathbf{A}\mathbf{U}$

  $\mathbf{\Lambda}, \mathbf{W} \leftarrow \text{EIG}\left(\tilde{\mathbf{A}}\right)$ ▷ Eigendecomposition of $\mathbf{A}$ into frequencies $\mathbf{\Lambda}$ and modes $\mathbf{W}$

  $\mathbf{\Phi} \leftarrow \mathbf{Y}\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{W}$ ▷ Eigenmodes of $\mathbf{A}$

end function

3.2 Optimized Dynamic Mode Decomposition

The *optimized* DMD algorithm of Askham and Kutz [3], which uses a variable projection method [4] for nonlinear least squares, provides the best performance of any DMD algorithm currently available. The opt-DMD algorithm minimally de-biases the eigenvalues of the system by allowing the data matrix, $\mathbf{X}$, to be reconstructed as

$$\mathbf{X} \approx \underbrace{\begin{bmatrix} | & & | \\ \boldsymbol{\phi}_1 & \cdots & \boldsymbol{\phi}_r \\ | & & | \end{bmatrix}}_{\mathbf{\Phi}} \underbrace{\begin{bmatrix} b_1 & & \\ & \ddots & \\ & & b_r \end{bmatrix}}_{\text{diag}(\mathbf{b})} \underbrace{\begin{bmatrix} e^{\lambda_1 t_1} & \cdots & e^{\lambda_1 t_m} \\ \vdots & \ddots & \vdots \\ e^{\lambda_r t_1} & \cdots & e^{\lambda_r t_m} \end{bmatrix}}_{\mathbf{T}(\boldsymbol{\lambda})}, \tag{3.6}$$

where the i^{th} eigenmode, $\boldsymbol{\phi}_i$, has a corresponding mode amplitude b_i and eigenvalue λ_i. The opt-DMD algorithm directly solves the exponential time dynamics fitting problem,

$$\min_{\boldsymbol{\lambda}, \mathbf{\Phi_b}} \|\mathbf{X} - \mathbf{\Phi_b}\mathbf{T}(\boldsymbol{\lambda})\|_F, \tag{3.7}$$

where modes and amplitudes are combined into $\mathbf{\Phi_b} = \mathbf{\Phi}\text{diag}(\mathbf{b})$. The optimized dynamic mode decomposition (opt-DMD) method has been demonstrated to provide a superior decomposition by its capability to optimally suppress bias, a feature not inherent in exact DMD and its variants. Unlike its counterparts, opt-DMD simultaneously computes an optimization across data snapshots, thereby obviating the need for evenly spaced sampling

intervals. However, a drawback of opt-DMD is that it requires solving a nonlinear and non-convex optimization problem, which is computationally more demanding than the exact DMD. Moreover, the solution inherently provides only a local minimum to the optimization problem, yet to date no DMD algorithm provides a global minimizer with this scheme. Despite these challenges, opt-DMD is considered a relatively lightweight algorithm and has proven to be numerically performant in practice [3]. As a result, the usage of exact DMD and its variants, such as fbDMD [27], has been largely relegated to scenarios where low precision is acceptable and low computational complexity is an exceptional priority.

3.3 Feed-forward Neural Networks

Artificial neural networks, or neural nets as they are often referred to now, have emerged as powerful tools for various machine learning tasks, ranging from image recognition to natural language processing and, notably, turbulent fluid dynamics [36, 38, 17]. In particular, deep neural networks with multiple hidden layers have shown remarkable success in learning intricate feature representations. However, these deeper architectures may require substantial computational resources and data to train effectively, making them less suitable for scenarios with limited resources or small datasets.

Shallow Decoder Neural Networks (SDNs) [33] present an alternative approach that focuses on simplicity and efficiency with a reduced number of layers. It should be noted that "shallow" as opposed to "deep" in the general machine learning context are not well-defined numbers, but rather, heuristical assessments of a neural net's size. A visualization is provided in Figure 3.1 In context of this manuscript, we explicitly focus on the minimal layer requirements: three layer SDNs. Despite their relative simplicity, SDNs can still achieve competitive performance, especially in low-complexity tasks where feature extraction from multiple layers is not imperative.

SDNs prove themselves to be quicker to train (fewer weights), faster to tune (fewer hyper-parameters), and simpler to interpret (fewer layers). Regarding the final point of interpretability, an SDN's dense neural layers are not easily translated to image output, yet

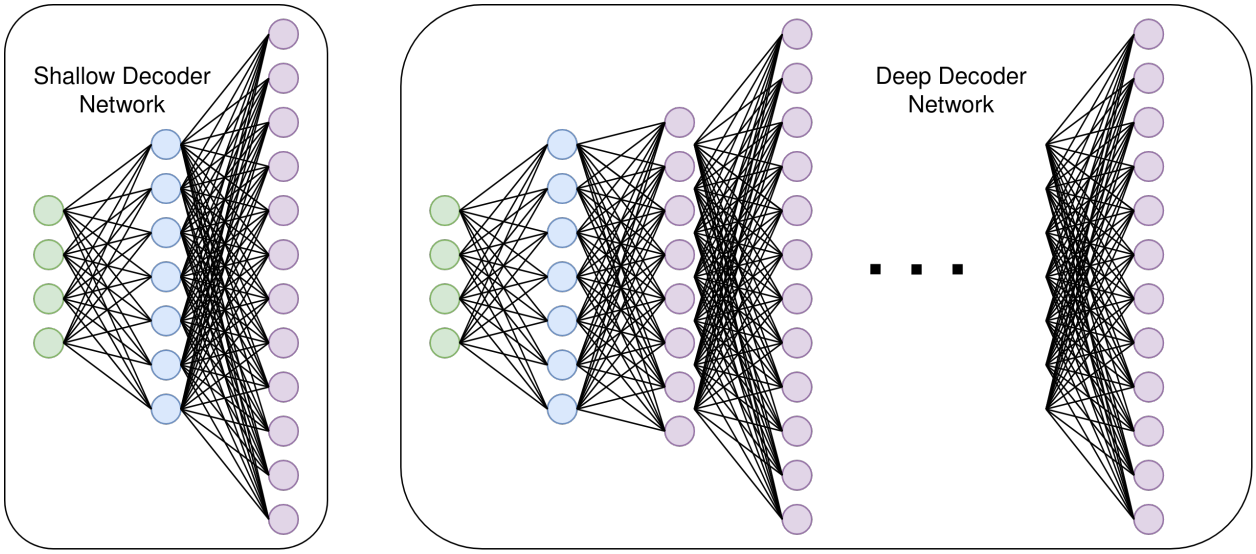

Figure 3.1: A Shallow Decoder Network (SDN). Shallow versus deep is not a well-defined layer size but a heuristical assessment. For the purpose of this thesis, we consider SDNs to be 3-layer networks.

using a single hidden layer at the same size as the output, for example, can allow for a sort of window into the SDN's m̈odal decomposition".

Fundamental in training a neural network is a choice of activation function, loss function, and optimizer for the backpropagation algorithm [43, 81]. The Rectified Linear Unit (ReLU) [1] activation function,

$$\sigma_{\mathrm{ReLU}}(x) = \max(0, x) = \frac{x + |x|}{2} = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.8}$$

is widely considered one of the strongest candidate activation functions, which is why it will be exclusively used in this work. ReLU displays several desirable properties:

1. ReLU is idempotent for any finite composition: $\sigma_{\mathrm{ReLU}} \circ \sigma_{\mathrm{ReLU}} \circ \cdots \circ \sigma_{\mathrm{ReLU}} = \sigma_{\mathrm{ReLU}}$. Thus, it circumvents some problems with vanishing gradients when compared to, say, sigmoidal activation functions.

2. ReLU is computationally inexpensive, requiring only an element comparison, addition, and multiplication. Network-wide activation rates are sparse as well, given that half of all ReLU activations are zero for symmetric, demeaned datasets.

3. ReLU is scale invariant: $\sigma_{\text{ReLU}}(ax) = a\sigma_{\text{ReLU}}(x)$ for $a \geq 0$.

Furthering the elegance of the ReLU in mitigating vanishing gradient issues, an SDN, versus a deeper network, also minimizes gradient issues by construction. While there are downsides to the ReLU activation function, like non-differentiability at zero and unboundedness, many of these can be addressed, for example with a 0 or 1 sub-gradient and with proper data normalization, respectively.

3.4 Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are a type of artificial neural network well-suited for sequential data, and they play a significant role in our research on aero-optic wavefront prediction and control because of their abiolity to recognize temporal trends. Unlike traditional feedforward neural networks, which process inputs independently, RNNs maintain hidden states that encapsulate information about previous inputs in the sequence. This makes RNNs inherently dynamic and capable of handling data where temporal ordering matters, such as time-series data or sequential data.

A Long Short-Term Memory (LSTM) [52, 2] RNN, a cell of which is depicted in Figure 3.2, consists of four gated layers of neurons with feedback connections that allow information to cycle through the network over time steps. The output at a given time step is a function of the current input as well as the output from the previous time step's cell. In this way, the network retains a form of memory about past inputs that influences future computations. LSTMs were designed to mitigate issues of vanishing and exploding gradients.

In our context, we use RNNs to model and predict the sequential changes in aero-optic wavefronts. By capturing temporal dependencies, RNNs can identify trends and patterns over time, providing predictions for wavefront distortions based on historical data. This

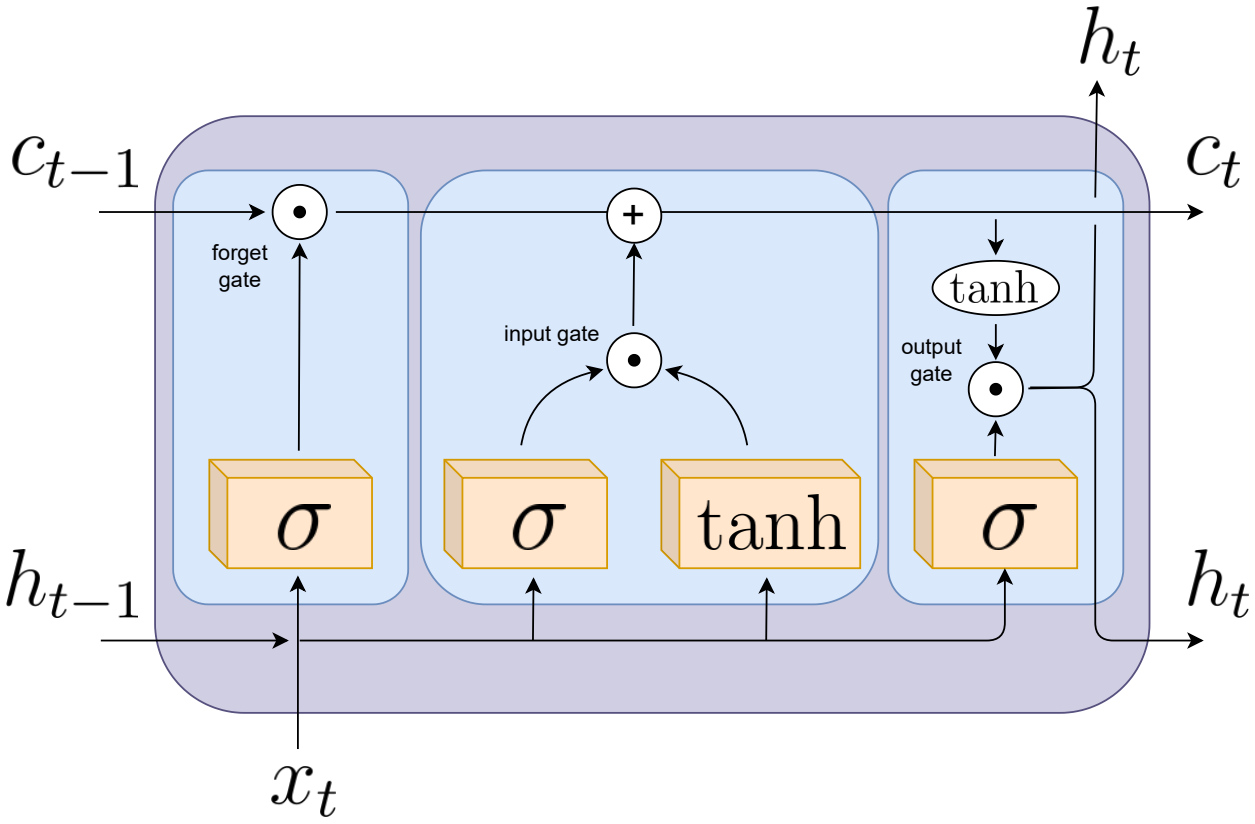

Figure 3.2: Single cell of an LSTM recurrent neural network. Outputs feed from and into the next cell(s), the number of which is generally determined by the length (e.g. number of temporal snapshots) of the input vector. Outputs may also be fed upwards to another LSTM layer.

ability to anticipate changes makes RNNs an ideal choice for our predictive control system, enabling it to react proactively to incoming distortions.

3.4.1 Long Short-Term Memory (LSTM)

LSTM networks are a form of RNN that aim to deal with the problem of vanishing gradients [52, 50]. The vanishing gradient problem arises during the back-propagation process, where errors tend to decay as they are propagated back in time. This problem severely

impairs credit assignment in classic RNNs, making it difficult to correctly identify relevant events that may have occurred far back in the input sequence, thus making it challenging for RNNs to handle long-term dependencies.

The LSTM solution to the vanishing gradient problem is through memory cells, which enable the flow of constant error during training. The common LSTM cell includes an input gate, output gate, and a forget gate [51, 41]. While the cell retains values over arbitrary time periods, the three gates regulate the flow of information into and out of the cell. The input gate, for instance, decides what pieces of new information to store in the current state, after comparing the current input to that of a previous state and assigned a value in $[0, 1]$. Similarly, the forget gate uses such a score to either keep or discard information, while the output gate uses such a score to decide whether to pass relevant information to the next unit.

The PyTorch [78] implementation of the LSTM, which is what is used in this work, uses the following logic. For each element in an input sequence, x_t at time t, every layer in an LSTM performs

$$\begin{aligned} i_t &= \sigma\left(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}\right) \\ f_t &= \sigma\left(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}\right) \\ g_t &= \tanh\left(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}\right) \\ o_t &= \sigma\left(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}\right) \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\ h_t &= o_t \odot \tanh\left(c_t\right) \end{aligned} \tag{3.9}$$

where h_t is the hidden state at time t, c_t is the cell state at time t, and i_t, f_t, g_t, o_t are the input, forget, cell, and output gates, respectively. The $\odot$ symbol denotes the element-wise, or Hadamard, product. Weights, W_qp, and biases, b_qp are sub-scripted to indicate which recurrent connection is being tracked, with $q, p \in i, f, g, o$ to match the gates. The activation functions within the LSTM cell are the sigmoid function, σ.

Chapter 4

DYNAMIC MODE DECOMPOSITION FOR AERO-OPTIC WAVEFRONT CHARACTERIZATION

This work presents Reference [82], reproduced with permission, and is largely self-contained. A proceedings paper over-viewing methods and results is also available in Reference [66].

4.1 Summary

Aero-optical beam control relies on the development of low-latency forecasting techniques to quickly predict wavefronts aberrated by the Turbulent Boundary Layer (TBL) around an airborne optical system, and its study applies to a multi-domain need from astronomy to microscopy for high-fidelity laser propagation. We leverage the forecasting capabilities of the Dynamic Mode Decomposition (DMD) – an equation-free, data-driven method for identifying coherent flow structures and their associated spatiotemporal dynamics – in order to estimate future state wavefront phase aberrations to feed into an adaptive optic (AO) control loop. We specifically leverage the optimized DMD (opt-DMD) algorithm on a subset of the Airborne Aero-Optics Laboratory Transonic (AAOL-T) experimental dataset, characterizing aberrated wavefront dynamics for 23 beam propagation directions via the spatiotemporal decomposition underlying DMD. Critically, we show that opt-DMD produces an optimally de-biased eigenvalue spectrum with imaginary eigenvalues, allowing for arbitrarily long forecasting to produce a robust future-state prediction, while exact DMD loses structural information due to modal decay rates.

4.2 Introduction

Free-space communication, high-resolution imaging, and directed energy are sought-after lasing applications, all of which are desirable on aircraft. Achieving high-fidelity laser beam propagation in the air requires mitigating the phase distortion of the outgoing wavefront. In atmospheric and free-space laser propagation, such as in observational astronomy [26] and satellite quantum key distribution [22], and even in biological specimen microscopy [10], spatiotemporal variations of the index of refraction must be compensated for via adaptive optic (AO) control systems. These AO systems typically use deformable mirrors to correct the outgoing beam by preemptively deforming the wavefront to cancel out any subsequent perturbation. Developing robust and responsive predictive controllers for AO systems is a highly desired enhancement with application well beyond airborne optics.

For an airborne optical platform, the AO system must correct wavefront distortions resulting from three primary sources: mechanical jitter of the platform, near-field effects where the Turbulent Boundary Layer (TBL) around the airborne platform rapidly alters the refractive index, and atmospheric effects where inhomogeneities and turbulence alter the propagation medium. [31, 29] This paper focuses on the near-field wavefront distortions that are referred to as aero-optical effects.

The term aero-optics refers to the intersection of optical and aerodynamic phenomena, such as the effects on the optical field from a high-speed turbulent flow, where air is forced around the optical system, possibly resulting in flow separation and shock formation. Characterizing these rapid wavefront aberrations is the goal of this study, using experimental data from the Airborne Aero-Optics Laboratory Transonic (AAOL-T) [58]. As depicted in Figure 4.1, the AAOL-T consists of a pair of Falcon 10 aircraft measuring in-flight laser transmission. The aero-optic effects that determine wavefront deformation are dependent on such factors as the laser platform shape and aircraft geometry, the speed of the craft, the Reynolds number, and the direction of beam propagation, which will be further discussed in Section 4.3.

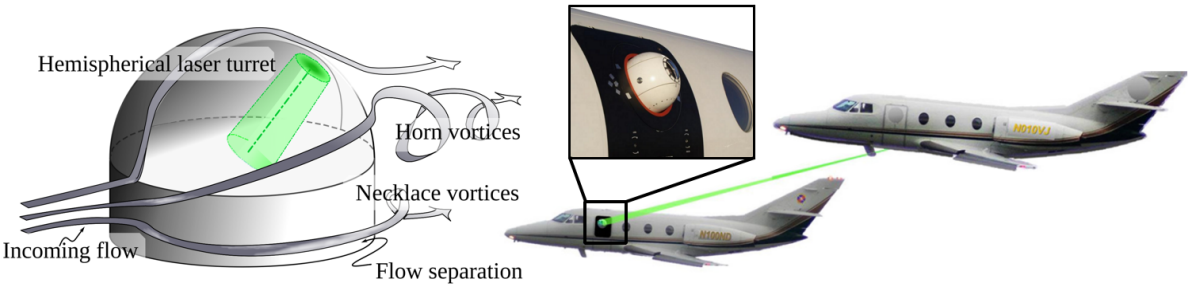

Figure 4.1: AAOL-T aircraft with hemispherical laser turret. The turret geometry produces a turbulent flow field in the few centimeters surrounding the sensors. Because the flow dynamics may not be accounted for with on-board sensors across applications, their effects must be predictively controlled in order to properly produce high-fidelity, coherent transmission of the electric field. *Figure of the AAOL-T aircraft adapted from Reference [58].*

The aerodynamic environment of airborne laser platforms motivates high-fidelity computational fluid dynamics (CFD) models and experimental techniques in the study of aero-optics effects. The high-speed, high Reynolds number compressible flows around airborne platforms can contain TBLs, shear layers, and wakes, as well as shock waves in the case of transonic and supersonic flows. [101, 31] As a laser beam propagates through this turbulent flow surrounding the aperture, refractive index fluctuations cause phase aberrations, and the resulting distortions of the optical field are referred to as aero-optical effects [104].

The index of refraction, n, is directly linked to air density fluctuations by

$$n\left(\mathbf{r}\right) = 1 + K_{GD}\left(\lambda_0\right)\rho\left(\mathbf{r}\right), \tag{4.1}$$

where K_{GD} is the wavelength-dependent Gladstone-Dale factor, λ_0 is the laser wavelength, and $\rho\left(\mathbf{r}\right)$ is the air density as a function of the spatial variable $\mathbf{r}$ [101].

Characterizing propagating beam wavefront dynamics in the TBL is critical to correcting the outgoing phase profile of the beam. From an applied standpoint, despite the relatively short, centimeter scale, distance traveled in the TBL, the beam quality is immediately and

often heavily degraded within this region. [103] A typical method to quantify the aero-optic wavefront aberrations from a given refractive index field is by calculating optical path difference (OPD). OPD is computed by first calculating the optical path length (OPL), which is proportional to the travel time for corresponding rays. OPL is often computed as the integral of the index of refraction along the propagation direction,

$$\mathrm{OPL}\,(x, y, t) = \int_0^{z_1} n\,(x, y, z, t)\,dz. \tag{4.2}$$

Subtracting the mean OPL over the spatial coordinates of the aperture produces the OPD,

$$\mathrm{OPD}\,(x, y, t) = \mathrm{OPL}\,(x, y, t) - \langle \mathrm{OPL}\,(x, y, t) \rangle\,. \tag{4.3}$$

We have let z in Equation 4.2 be the optical axis of the beam with x and y coordinates covering the aperture as seen in Figure 4.2. Assuming the dominant contribution to the OPD occurs within the TBL over short transmission distances, we may let the upper bound of integration, z_1, match the extent of the TBL. The root mean square of OPD across each dataset provides a metric to assess the severity of wavefront distortions for the given experiment. To compare OPD across experiments, we then normalize it as a dimensionless quantity, assigning

$$\mathrm{OPD} \leftarrow \frac{\mathrm{OPD}}{M^2 D \rho/\rho_0}, \tag{4.4}$$

where M is the Mach number as a ratio of the speed of sound, D is the turret diameter as seen in Figure 4.2, and ρ/ρ_0 is a ratio of in-flight air density to sea-level air density. For the AAOL-T data, we analyze a subset such that each trial is taken at $M = 0.6$ and $\rho = 0.812\,\mathrm{kg/m^3}$. The sea level air density is set to the standard $\rho_0 = 1.225\,\mathrm{kg/m^3}$, and the AAOL-T turret diameter is $D = 0.3048\,\mathrm{m}$. To characterize each trial in our dataset, we compute the root-mean-square, $\mathrm{OPD_{rms}}$. This scaling allows extrapolation across various turret diameters and altitudes, since the same spatial wavefront characteristics are retained for subsonic flow ($M \lessapprox 0.6$) in general and for transonic ($M \gtrapprox 0.6$) and supersonic ($M > 1$) flow as long as Mach number is matched [58]. When referring to OPD and $\mathrm{OPD_{rms}}$ in figures and elsewhere in this paper, we imply the normalized formulation in Equation 4.4.

The analysis of aero-optical wavefront reconstruction leverages time-series measurements collected through the TBL. Here we highlight the measurement and sensor technologies used for characterizing aero-optic interactions. We describe the underlying mathematical architecture that leverages these measurements in order to develop dynamic models for wavefront reconstruction.

4.3 AAOL-T Experimental Data

The AAOL-T was run by researchers at the University of Notre Dame to obtain live aero-optical data in flight. A 532 nm source beam propagates from a hemispherical laser turret of diameter 0.3048 m mounted on a Falcon 10 aircraft, as depicted in Figure 4.1. The beam overfills the pupil aperture on the receiver laboratory aircraft. A Shack-Hartmann wavefront sensor (SHWFS), depicted in Figure 4.4a and described in Section 4.4, is used to capture wavefront phase aberrations between the source and destination beams.

The AAOL-T dataset involves measurements up to Mach 0.8 taken at a sampling rate of 25 kHz. Shock-formation on the hemispherical turret is observed at transonic Mach numbers [47]. The distance between source and receiver aircraft is approximately 50 m. The beam direction is recorded in terms of its azimuth and elevation angles, as visualized in Figure 4.2, with respect to the cylindrical base of the hemispherical turret. It is useful to reparametrize the beam direction in terms of a "look-back" angle and inclination angle, α and β respectively, where

$$\alpha = \cos^{-1}\left[\cos(\mathrm{Az})\cos(\mathrm{El})\right] \tag{4.5}$$

$$\beta = \tan^{-1}\left[\frac{\tan(\mathrm{El})}{\sin(\mathrm{Az})}\right]. \tag{4.6}$$

A look-back angle, α, of zero is a beam propagating in the forward direction of the aircraft, while an angle of 180° would designate propagation towards the rear, through the outgoing turbulent wake. An inclination angle, β, of zero describes a beam facing the earth, while 180° would be skyward. Figure 4.3a measures the effects of α and β on $\mathrm{OPD}_{\mathrm{rms}}$ for

all 23 data sets. For backwards looking angles $\alpha > 90°$, an increasing OPD indicates a heightening level of aberrations in the wavefront. With Figure 4.3b, we can visualize the effects of this look-back angle when also considering β. The dark red region indicates angles that lie where horn vortices exist; OPD_{rms} tends to be greatest for these data points.

In-flight measurements from the AAOL-T experiment, the aircraft, and hemispherical laser turret with conformal window, which are depicted in Figures 4.1 and 4.2, have contributed to a database of aero-optic disturbance measurements. Atmospheric aberrations are often characterized by Zernike modes [95], an orthogonal sequence of polynomials that span the unit disk and possess odd or even radial symmetries. Zernike modes offer interpretability to optical dynamics and can yield insights where radially symmetric aberrations are concerned. Yet this is often not the case for aero-optical disturbances in the TBL which are prone to quickly-varying nonlinearities in the index of refraction at transonic flow speeds [44]. An analysis of the temporal phase structure function and other statistics of AAOL-T wavefront data was performed by Brennan and Wittich in 2013 [11]. Proper Orthogonal Decomposition (POD) and Dynamic Mode Decomposition (DMD) modes have been used to provide a spatio-temporal characterization of the flow dynamics [45]. Predictive control methods for aero-optics have been analyzed on these data as well [46, 20].

4.4 Sensors and Data Acquisition

Depicted in Figure 4.4a, a SHWFS is used to capture wavefront phase aberrations in the AAOL-T experiment. [89] A lenslet array in the pupil plane and at a focal distance away from an optical sensor focuses an incoming wavefront into sub-regions on the detector plane. Any deviations from a planar wavefront manifest as displacements, Δx_i and Δy_i, from the optical axis in the i^{th} sub-region. The wavefront phase can be reconstructed by a least-squares fit of the average intensity-weighted gradients across the subapertures [89]. These gradients are proportional to the centroid tilts, the displacement of the centroid of the focused rays. The difference between the centroid tilts and the true gradients can alter SHWFS readings. This source of measurement error cannot be analytically evaluated outside

low Mach numbers where weak scintillation and Rytov theory apply [80]. A performant SHWFS typically requires the length of each subaperture to be less than a fourth of the Fried coherence length for atmospheric turbulence [9].

A 3-D representation of the SHWFS on the AAOL-T laser turret is depicted in Figure 4.4a. The incoming beam's aberrated wavefront is focused from the gridded lenslet array onto sub-regions on the detector plane as depicted by the larger green dots. The displacements, Δx_i and Δy_i, of each focused sub-beam from the centroid of each sub-region, shown by the smaller black dot, is used to compute the local tilt of the incoming wavefront, from which the wavefront may be reconstructed. With a planar, unaberrated incoming wavefront, the focused spots would be in a perfect grid matching the lenslet array geometry. The central circular region in Figure 4.4b represents the secondary mirror obscuration of the optical system inside the turret, which includes a telescope used to align the turret on one AAOL-T aircraft to the incoming beam from the other and is not an explicit feature of a general SHWFS. Because of the 1-inch diameter obscuration [29], the unprocessed SHWFS data is taken as a set of points lying in an annulus.

Figure 4.4b shows a single frame of unprocessed SHWFS data from the AAOL-T platform used in this study. The data were acquired using a v1610 Vision Research Phantom camera at 30 kHz for a total of 21,504 frames captured per dataset. Figure 4.4c is an example of the processed data and the wavefront from the local tilts of the SHWFS that we use in our analysis. As will be described in the upcoming sections, this study investigated 23 sets of data with varying α and β angles, as defined in Figure 4.2. Each (α, β) pair defines a beam direction.

4.5 Optimized Dynamic Mode Decomposition

DMD was an algorithm developed by Schmid [87, 86] in the fluid dynamics community to identify spatio-temporal coherent structures from high-dimensional data. DMD is based on Proper Orthogonal Decomposition (POD), which utilizes the computationally efficient singular value decomposition (SVD) so that it scales well to provide effective dimensionality

reduction in high-dimensional systems. DMD provides a modal decomposition where each mode consists of spatially correlated structures that have the same linear behavior in time (e.g., oscillations at a given frequency with growth or decay). Thus, DMD not only provides dimensionality reduction in terms of a reduced set of modes, but also provides a model for how these modes evolve in time.

Several algorithms have been proposed for DMD, with the *exact* DMD framework developed by Tu et al. [97] being the simplest, least-squares regression to produce the decomposition. DMD is inherently data-driven, and the first step is to collect a number of pairs of snapshots of the state of a system as it evolves in time. These snapshot pairs may be denoted by $\{\mathbf{x}(t_k), \mathbf{x}(t'_k)\}_{k=1}^{m}$, where $t'_k = t_k + \Delta t$, and the timestep, Δt, must be sufficiently small to resolve the highest frequencies in the dynamics. As before, a snapshot may be the state of a system, such as a three-dimensional fluid velocity field sampled at a number of discretized locations that is reshaped into a high-dimensional column vector. These snapshots are then arranged into two data matrices, $\mathbf{X}$ and $\mathbf{X}'$,

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}(t_1) & \mathbf{x}(t_2) & \cdots & \mathbf{x}(t_m) \\ | & | & & | \end{bmatrix} \tag{4.7a}$$

$$\mathbf{X}' = \begin{bmatrix} | & | & & | \\ \mathbf{x}(t'_1) & \mathbf{x}(t'_2) & \cdots & \mathbf{x}(t'_m) \\ | & | & & | \end{bmatrix}. \tag{4.7b}$$

If we assume uniform sampling in time, we will adopt the notation $\mathbf{x}_k = \mathbf{x}(k\Delta t)$.

The DMD algorithm seeks the leading spectral decomposition (i.e. eigenvalues and eigenvectors) of the best-fit linear operator, $\mathbf{A}$, that relates the two snapshot matrices in time by

$$\mathbf{X}' \approx \mathbf{A}\mathbf{X}. \tag{4.8}$$

The best-fit operator, $\mathbf{A}$, then establishes a linear dynamical system that best advances snap-

shot measurements forward in time. If we assume uniform sampling in time, this becomes

$$\mathbf{x}_{k+1} \approx \mathbf{A}\mathbf{x}_k. \tag{4.9}$$

Mathematically, the best-fit operator $\mathbf{A}$ is defined as

$$\mathbf{A} = \underset{\mathbf{A}}{\operatorname{argmin}} \|\mathbf{X}' - \mathbf{A}\mathbf{X}\|_F = \mathbf{X}'\mathbf{X}^{\dagger} \tag{4.10}$$

where $\|\cdot\|_F$ is the Frobenius norm and † denotes the Moore-Penrose pseudo-inverse. The matrix $\mathbf{A}$ is an operator that advances the measurements in $\mathbf{x}$ forward in time. It is often helpful to convert the eigenvalues of this discrete-time operator into continuous time, resulting in eigenvalues $\lambda = \mu + i\omega$.

Alternative and better approaches are available [23, 56, 3] to the exact DMD algorithm. Bagheri [8] first highlighted that DMD is particularly sensitive to the effects of noisy data, with systematic biases introduced to the eigenvalue distribution [30, 7, 28, 49]. For example, when additive white noise is present in the measurements of an n-dimensional system with m snapshots, the bias in exact DMD will be the dominant component of DMD error whenever the signal-to-noise ratio exceeds $\sqrt{n/m}$, implying that the effects of noise cannot always be mitigated by increasing the number of snapshots [28]. As a result, a number of methods have been introduced to stabilize performance, including total least-squares DMD [49], forward-backward DMD [28], variational DMD [5], subspace DMD [94], time-delay embedded DMD [13] and robust DMD methods [3, 85].

However, the *optimized* DMD algorithm of Askham and Kutz [3], which uses a variable projection method [4] for nonlinear least squares, provides the best performance of any algorithm currently available. This is not surprising given that it actually is constructed to both generalize and optimally satisfy the DMD problem formulation. In opt-DMD, the data matrix, $\mathbf{X}$, may be reconstructed as

$$\mathbf{X} \approx \underbrace{\begin{bmatrix} | & & | \\ \boldsymbol{\phi}_1 & \cdots & \boldsymbol{\phi}_r \\ | & & | \end{bmatrix}}_{\boldsymbol{\Phi}} \underbrace{\begin{bmatrix} b_1 & & \\ & \ddots & \\ & & b_r \end{bmatrix}}_{\operatorname{diag}(\mathbf{b})} \underbrace{\begin{bmatrix} e^{\lambda_1 t_1} & \cdots & e^{\lambda_1 t_m} \\ \vdots & \ddots & \vdots \\ e^{\lambda_r t_1} & \cdots & e^{\lambda_r t_m} \end{bmatrix}}_{\mathbf{T}(\boldsymbol{\lambda})}, \tag{4.11}$$

where the i^{th} eigenmode, $\boldsymbol{\phi}_i$, has a corresponding mode amplitude b_i and eigenvalue λ_i. The opt-DMD algorithm directly solves the exponential time dynamics fitting problem,

$$\min_{\boldsymbol{\lambda},\boldsymbol{\Phi}_\mathbf{b}} \|\mathbf{X} - \boldsymbol{\Phi}_\mathbf{b}\mathbf{T}(\boldsymbol{\lambda})\|_F, \tag{4.12}$$

where modes and amplitudes are combined into $\boldsymbol{\Phi}_\mathbf{b} = \boldsymbol{\Phi}\text{diag}(\mathbf{b})$. This has been shown to provide a superior decomposition due to its ability to optimally suppress bias. Unlike exact DMD and its variants, opt-DMD also computes an optimization contemporaneously across snapshots, eliminating the need for evenly timed samples. The disadvantage of optimized DMD is that one must solve a nonlinear, nonconvex optimization problem. This is both computationally more expensive than exact DMD and the solution, by construction, only guarantees a local minimal to the optimization. In practice, opt-DMD is a relatively lightweight algorithm that has demonstrated itself to be numerically performant [3], relegating exact DMD and its variants like fbDMD to scenarios where low-precision is acceptable and low computational complexity is paramount.

4.6 Results and Analysis

Figures 4.5-4.7 show the result of an opt-DMD analysis for a total of nine different turret angles (α, β):

1. $(95°, 83°)$, $(127°, 81°)$, and $(153°, 83°)$ in Figure 4.5
2. $(88°, 65°)$, $(119°, 70°)$, and $(141°, 65°)$ in Figure 4.6
3. $(82°, 50°)$, $(94°, 52°)$, and $(108°, 53°)$ in Figure 4.7

These beam directions parameters were chosen to cover a variety of points along the AAOL-T's hemispherical turret. Grouping these trials by inclination angle, β, allows us to better compare look-back angles α, as seen in the figures.

Each row of Figures 4.5-4.7 represents an individual (α, β) data set's opt-DMD analysis, showing the dominant eight eigenvalues and the corresponding modes 1, 3, 5, and 7

from which the even numbered modes may be inferred as complex conjugates. Note in all cases, the eigenvalue spectrum is completely de-biased, lying along the imaginary axis. This is the critical takeaway of opt-DMD: with nearly perfect imaginary eigenvalues, the presented modes and their exponential time dynamics experience little time decay, allowing for arbitrarily long-lasting forecasts.

To compare with the precision of opt-DMD, we consider in Figure 4.8 an exact DMD analysis of the $(\alpha = 153°, \beta = 83°)$ dataset. As shown by the turret geometry Figure 4.2, this angle is roughly along the mid-line of the turret with a high look-back angle, pointing into the turbulent region prone to aero-optical effects but just outside regions with prominent horn vortices. The singular value spectrum and corresponding cumulative energy plots in Figure 4.8a suggest an optimal rank truncation $r = k = 296$, [40] which is an overwhelming amount of modal detail to retain.

Figure 4.8b shows the continuous-time eigenvalue spectrum of the system at the given rank truncation. The parabolic envelope $\mu(\omega) = -0.11\omega^2 - 0.09$ of the continuous-time eigenvalues ought to be compared with the spectrum of opt-DMD in Figures 4.5-4.7, whose eigenvalues lie on the imaginary axis. The deformed envelope seen in Figure 4.8b is consistent with weak noise on self-sustaining oscillating flow fields. [8]. While truncating the exact DMD analysis at a lower rank may produce modes closer to the imaginary axis, a parabolic envelope remains and the performance of opt-DMD remains superior by construction.

Computing the modal half-life gives us a window into the shortcomings of exact DMD. The mean half-life is found to be

$$\left\langle t_{1/2} \right\rangle = \frac{1}{r} \sum_{j=1}^{r} \frac{-\log(2)\Delta t}{\mu_j} = 104\,\mu\text{s} \tag{4.13}$$

and the amplitude-weighted mean half-life is

$$\left\langle t^b_{1/2} \right\rangle = \frac{1}{\sum_{i=1}^{r} |b_i|} \sum_{j=1}^{r} \frac{-|b_j| \log(2)\Delta t}{\mu_j} = 138\,\mu\text{s}. \tag{4.14}$$

The modal half-life indicates a window of opportunity for a predictor to interact with an AO control loop. For the particular example in Figure 4.8, the modal half-lives individually

spanned a range from $50 - 500\,\mu s$. When examining all 23 trials for various (α, β), we discovered the mean half-lives were consistently on the order of a hundred microseconds. With these decay timescales, pertinent coherent turbulent structures become treated as transient effects, diminishing the ability of DMD to forecast the dominant spatiotemporal structures on a long horizon. Figure 4.8c characterizes the power spectrum of the exact DMD modes. Note that many powerful modes have lower than average half-lives, further compromising the ability of exact DMD to forecast turbulent flow dynamics.

4.7 Conclusion

Data-driven methods are becoming increasingly important to model complex spatio-temporal systems whose evolution dynamics are not well known or only characterized by time-series measurements. In the case of aero-optic interactions, modeling the induced turbulent wake from a turret is exceptionally challenging. Unless dynamics are characterized in an appropriate manner, the wavefront aberrations cannot be corrected in the AO system. We proposed a data-driven algorithmic architecture which aims to model the aero-optic interactions in an adaptive and real-time manner. Specifically, we introduced the opt-DMD algorithm to produce an unbiased modal analysis of the AAOL-T dataset which captures the wavefront aberrations induced by a turbulent flow around a turret. The imaginary-valued eigenspectrum of the opt-DMD operator permits longer forecasting in an AO loop when compared to an exact DMD algorithm. Exact DMD suffers from modal decay rates due to the real components introduced in the spectrum, whereas opt-DMD modes display no decay rate and permit forecasting as long as the model remains physically meaningful with respect to the environment. Indeed, traditional DMD algorithms have forecasting horizons which decay on the order of hundreds of microseconds whereas opt-DMD by construction allows for forecasting on scales required for control algorithms for AO corrections.

Further studies ought to assess the performance of opt-DMD on turret geometries beyond hemispherical, as well as compare opt-DMD's forecasting ability to existing aero-optical predictors that rely on time-invariant POD modes for dimensionality reduction or neural

network architectures [20]. Importantly, opt-DMD's minimal bias as well as its freedom in sampling variable time steps make it a promising predictor for aero-optical phenomena.

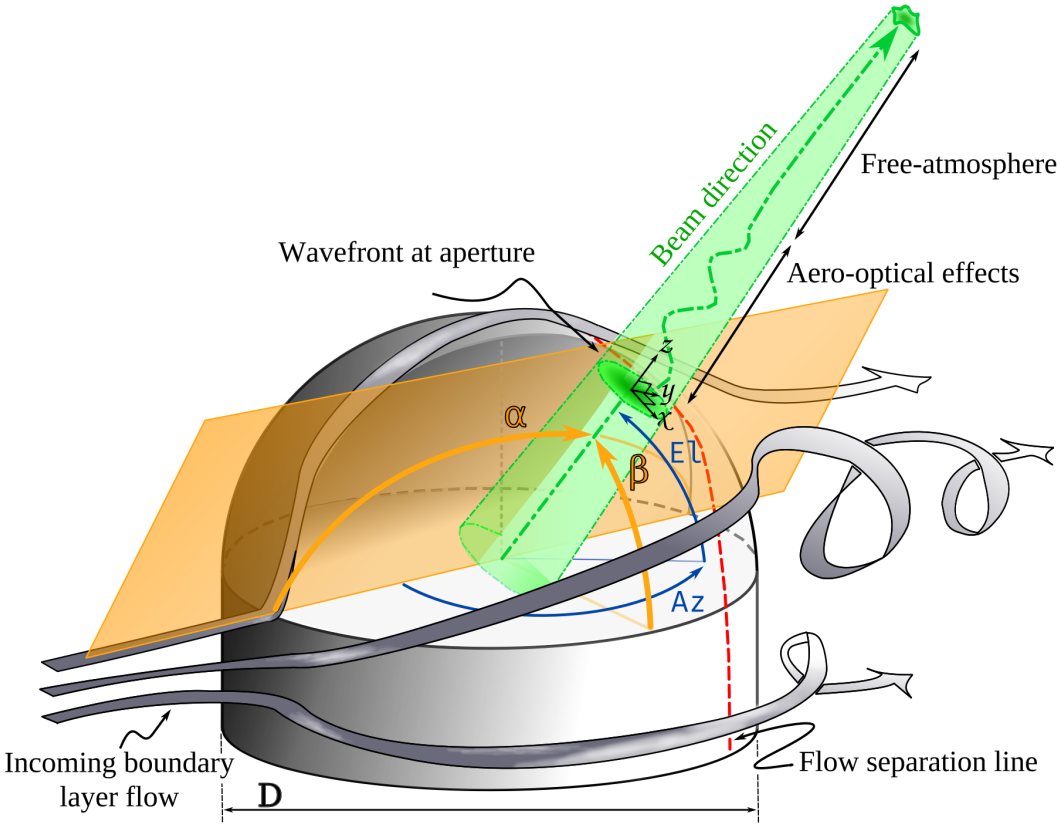

Figure 4.2: Detail of the turret geometry. The aperture (green disc with local coordinate system x-y-z) images the aberrated wavefronts. Wavefront distortion is highly dependent on beam direction, parametrized by α and β, the look-back angle and inclination respectively. These two parameters are mapped from the azimuth and elevation angles, Az and El, of the cylindrical base using the transformation in Equation 4.5. The red dotted line indicates the beginnings of the flow separation region for $\alpha > \pi/2$ and turbulent coherent structure formation.

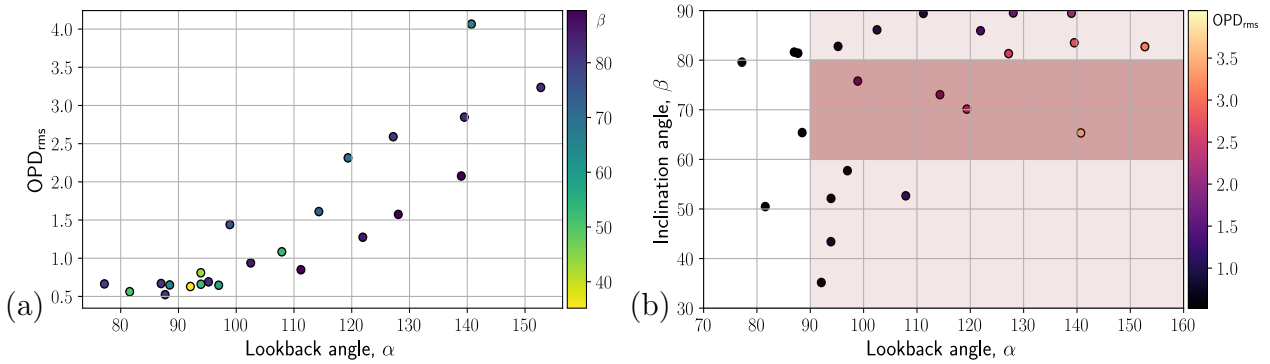

Figure 4.3: Beam propagation direction affects OPD, as seen in the AAOL-T data. **(a)** The effect of look-back angle, α, on OPD_{rms} with inclination angle, β given by the color bar. OPD increases as the turret direction α looks to the trailing edge of the flow field. **(b)** Look-back angle versus inclination angle with OPD_{rms} given by the color bar. The light red shaded region denotes look-back angles $\alpha > 90°$. Beam directions in this region may be in the flow separation region of the index of refraction. The dark red shaded region denotes inclination angles $60° < \beta < 80°$. This coloring serves as a rough guide to where horn vortices interrupt the boundary layer flow, resulting in greater OPD.

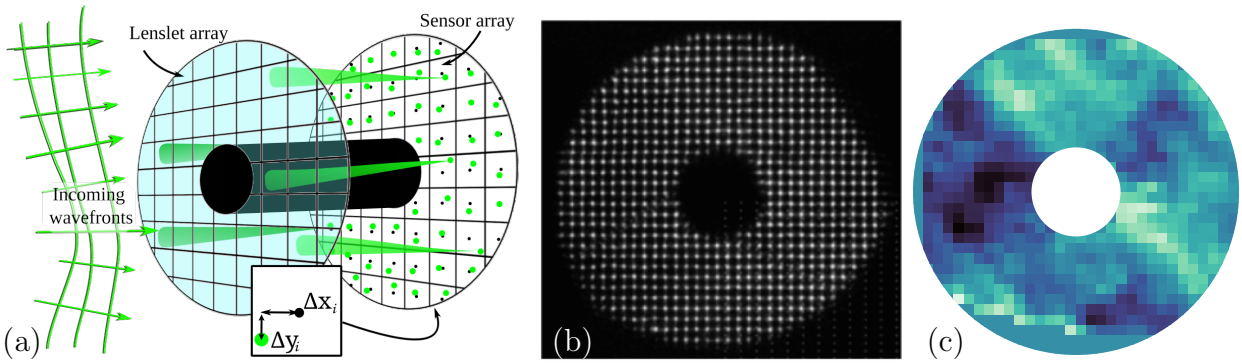

Figure 4.4: **(a)** Geometry of the SHWFS [89] on the AAOL-T laser turret with an incident aberrated wavefront. The lenslet arrays project to the sensor array where the displacements from the sensor centroids, measured by Δx_i and Δy_i, is used to compute the local tilts of the wavefront for reconstruction. **(b)** Unprocessed SHWFS data showing intensities projected on the 32x32 subapertures sensors from the AAOL-T. **(c)** Processed SHWFS data used in the Dynamic Mode Decomposition (DMD) analysis.

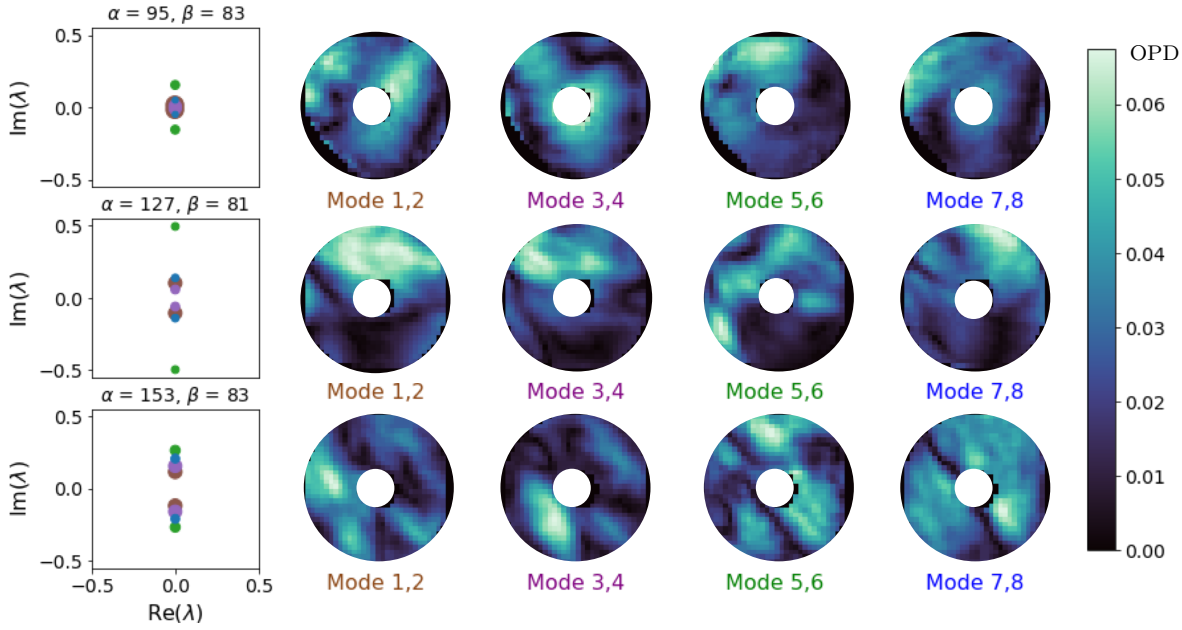

Figure 4.5: Experiments with $\beta \approx 80°$ for various α, all in degrees. Each row depicts the truncated eigenvalue spectrum and first eight modes of the OPD for beam direction (α, β) as indicated above the eigenvalue plot. As seen across all trials, opt-DMD determines completely imaginary eigenvalues, as seen on the leftmost column of plots. These long-lasted modes enable long-time prediction. Mode labels are colored to match the associated eigenvalue in the eigenvalue plot. Note that even modes are not displayed but are the complex conjugates of the odd modes.

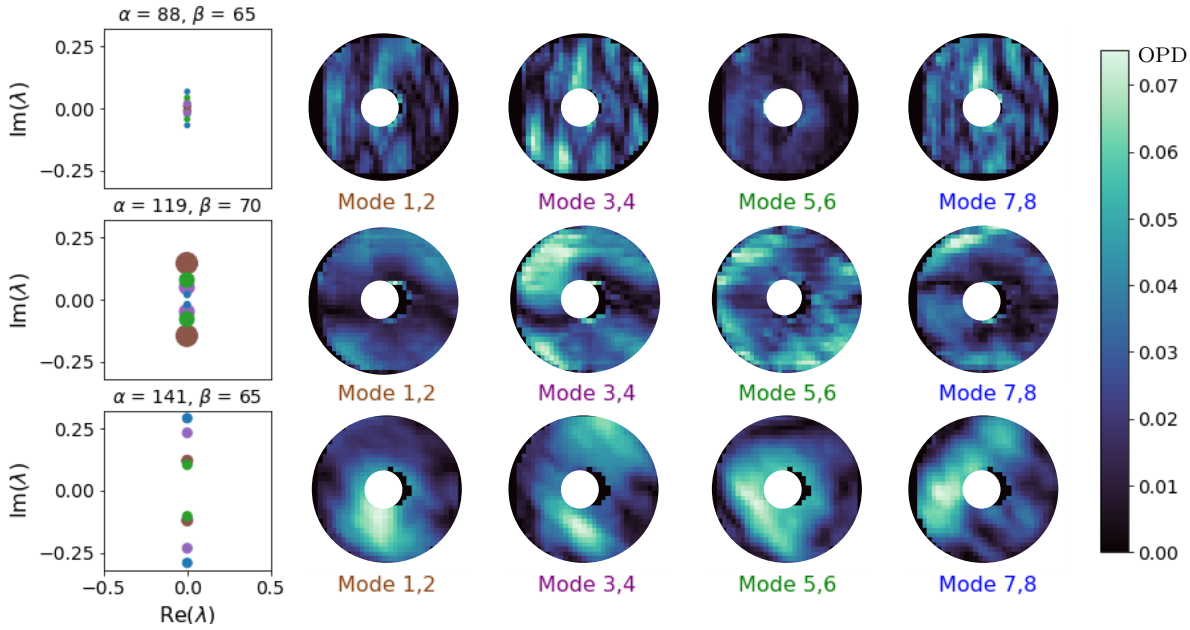

Figure 4.6: Experiments with $\beta \approx 65°$ for various α. Compare with Figures 4.5 and 4.7. Each row depicts the truncated eigenvalue spectrum and first eight modes of the OPD for beam direction (α, β) as indicated above the eigenvalue plot.

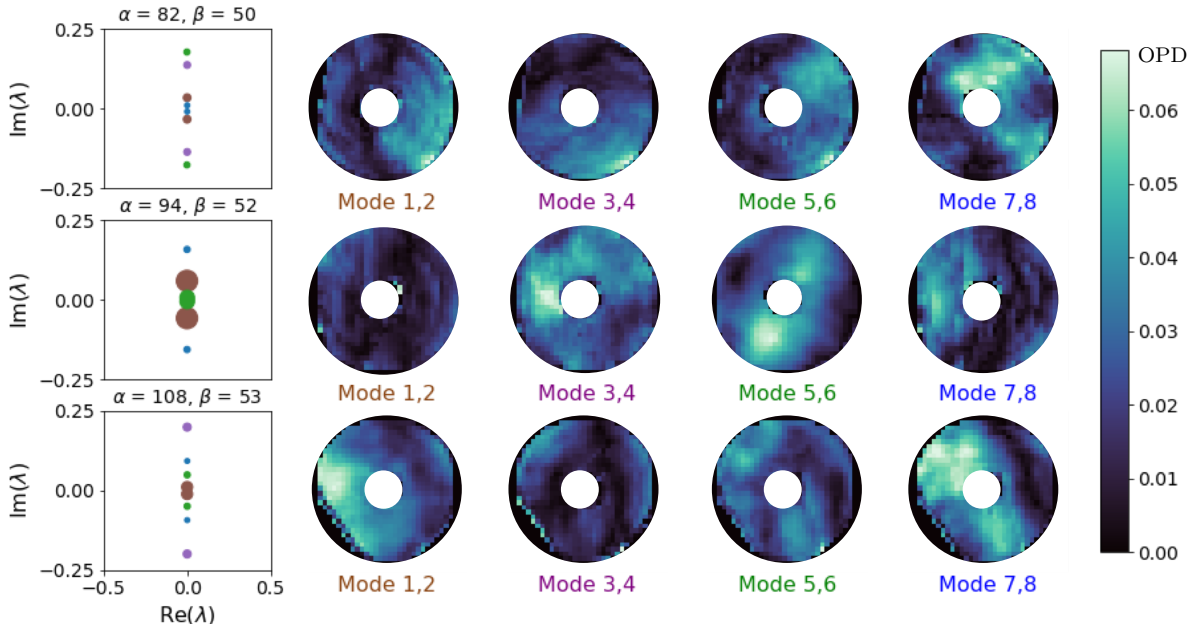

Figure 4.7: Experiments with $\beta \approx 50°$ for various α. Compare with Figures 4.5 and 4.6. Each row depicts the truncated eigenvalue spectrum and first eight modes of the OPD for beam direction (α, β) as indicated above the eigenvalue plot.

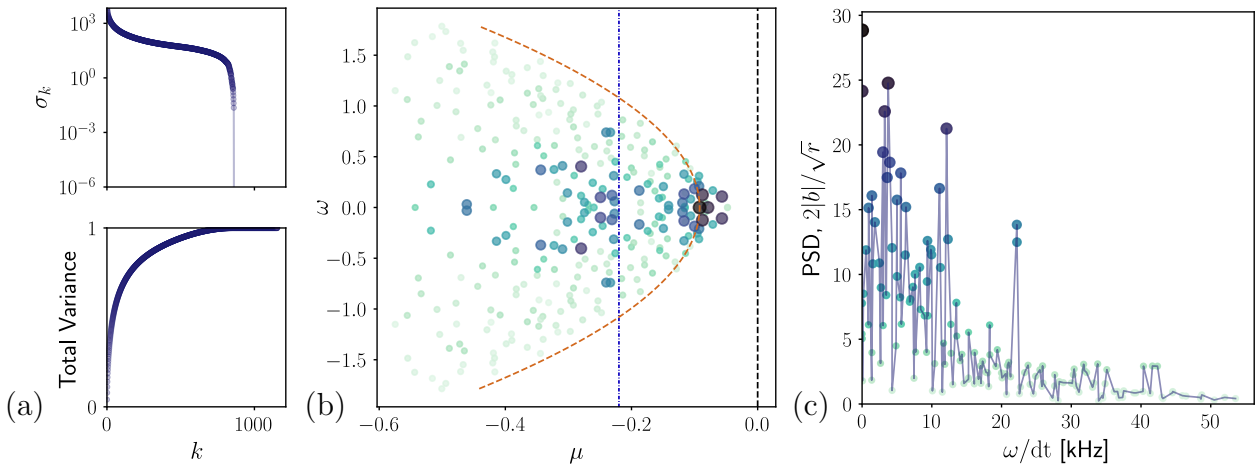

Figure 4.8: Demonstration of bias in the exact DMD algorithm for a beam direction with $\alpha = 153°, \beta = 83°$. The bias produces a half-life decay for the forecast on the order of a hundred microseconds. **(a)** SVD and cumulative energy of singular values. **(b)** Continuous-time eigenvalue spectrum. The orange dashed parabola which forms an envelope around the DMD eigenvalues is characteristic of noisy bias. In the opt-DMD spectra this curve becomes a vertical line. The blue dash-dotted line represents a cutoff to the left of which exist modes whose half-life exceeds the mean half-life. Color and size matches eigenvalues to corresponding points in the next figure, (c), and the power spectral density's vertical axis. **(c)** One-sided power spectrum of the DMD modes.

Chapter 5

SENSOR FUSION FOR AERO-OPTICS

This work presents an in-review submission that followed the conference proceedings of Reference [83], and is largely self-contained. A completed publication titled after this chapter is thus expected for later in 2023.

5.1 Summary

Sensor fusion technologies and algorithms allow for leveraging sensors with different modalities and/or performance trade-offs to estimate the current system state, produce robust forecasts, and enact control. Modern deep neural network architectures allow one to map between and fuse different measurements types, for instance, between high-resolution slow sensors and low-resolution fast sensors. This mapping can be performed across sensor modalities or within a single sensing modality. In this manuscript, we show that shallow decoder networks are an ideal mathematical architecture for producing high-quality mappings between measurement types, allowing the use of low-resolution fast sensors to produce accurate estimates of high-resolution fast sensing, which is beyond the capability of any single sensing modality. We explicitly demonstrate that our shallow decoder sensor fusion network can measure the complex flow physics of aero-optics data using digital holography and Shack-Hartmann wavefront sensing.

5.2 Introduction

Sensor technologies are critical for myriad diverse applications in the engineering and physical sciences. Sensors are the primary way to estimate the state of a system, produce robust forecasts, and enact control [75]. The diversity of sensor types has led to significant efforts in

sensor fusion technologies, whereby the benefits of multimodality for state estimation can be exploited [67]. The structural monitoring community has been especially active in integrating multimodal sensors in practice [110, 111, 72, 109, 92], with the placement of sensors with differing noise levels or costs having been recently considered [61, 25]. Sensor fusion offers the prospect of leveraging different features acquired by various sensor types in combination in order to produce an improved characterization of the system being measured. It is especially advantageous to use low-fidelity measurements from lower cost sensors to achieve the high-fidelity results that would have previously required expensive sensors. Emerging deep learning algorithms allow us to leverage the full capabilities of data fusion by mapping sensor modalities and sensor trade-offs in advantageous ways. In the present work, we demonstrate this approach explicitly on aero-optics data where different sensing modalities, specifically Digital Holography (DH) and Shack-Hartmann (SH) wavefront sensing, are mapped to each other using shallow decoder networks (SDNs).

There are two types of data fusion paradigms considered in this work. First, we consider mapping one sensor modality to another, in this case DH and SH wavefront sensors (WFS). Second, we consider the more classic fusion relating to the trade-offs between temporal sampling rate and spatial resolution quality. The ultimate goal is simple: use the simplest or least costly combinations of sensors possible to produce high-resolution state estimates at high-sampling rates. In the case of aero-optics, for instance, the DHWFS is able to produce high spatial resolution images at the expense of lower sampling rates. In contrast, the SHWFS can sample at exceptionally high rates, but only when adjusted to a relatively coarse spatial resolution. Mapping from a DH to SHWFS, and vice-versa, is one of the emerging modalities of sensor fusion addressed here via deep learning. For a single sensing modality, there is also the classic trade off between temporal and spatial resolution. High-speed cameras are a canonical example of this, whereby increased frame-rates correspond to coarser video resolution. For example, a modern smartphone can record 1080p HD (1920x1080) video at 240fps or 4K (3840x2160) at 60fps. Thus one can increase the temporal sampling by a factor of four at the expense of three fourths the spatial resolution. Of course, what is desired is the

capability of sampling at full spatial and temporal resolution simultaneously, which is the goal of the current work. There are a number of mathematical methods for addressing the single modality sensor fusion problem, most notably super-resolution techniques [91] which can also be applied to resolve complex fluid flows [37, 39].

Machine learning algorithms provide opportunities to improve on this sensor fusion problem. More broadly, these emerging algorithms are being widely used across the engineering and physical sciences [14]. The recent success of deep learning has been enabled by the continued growth of computational power and exceptionally large labeled datasets which are capable of leveraging multi-layer (deep) architectures. So, although neural network models have a four decade long history, it was the ImageNet dataset in 2012 [64] that provided a watershed for deep learning [69]. Importantly, multi-layer feed-forward networks are a class of universal approximators [54] that are capable of providing exceptional characterizations of labeled training data. Deep learning thus exploits low-dimensional feature spaces in high-dimensional data. These features can be exploited for various tasks, including classification, reconstruction, and control.

Our specific aim in this work is to use machine learning methods for aero-optic sensor fusion with the goal of producing high-resolution imaging at high-sampling rates. Thus, we ultimately aim to integrate SH and DHWFS modalities in order to produce the desired outcome. Figure 5.1 highlights our overall sensor fusion strategy. Specifically, a SDN is used to map low-resolution measurements to their high-resolution counterparts. The SDN is only a few layers, three in our design, thus allowing for less training data in order to achieve the desired performance [34]. Unlike previous work mapping sensors to full state estimates, the SDN structure can map across datasets, that is, from low-resolution SHWFS to high-resolution DHWFS measurements. Alternatively, it can be used to map a single modality from its low-resolution counterpart to its high-resolution counterpart. As we show on both synthetic data and aero-optics data, our fusion paradigm ultimately allows low-resolution, high-sampling rate data streams to produce an estimate of high-resolution, high-sampling rate images.

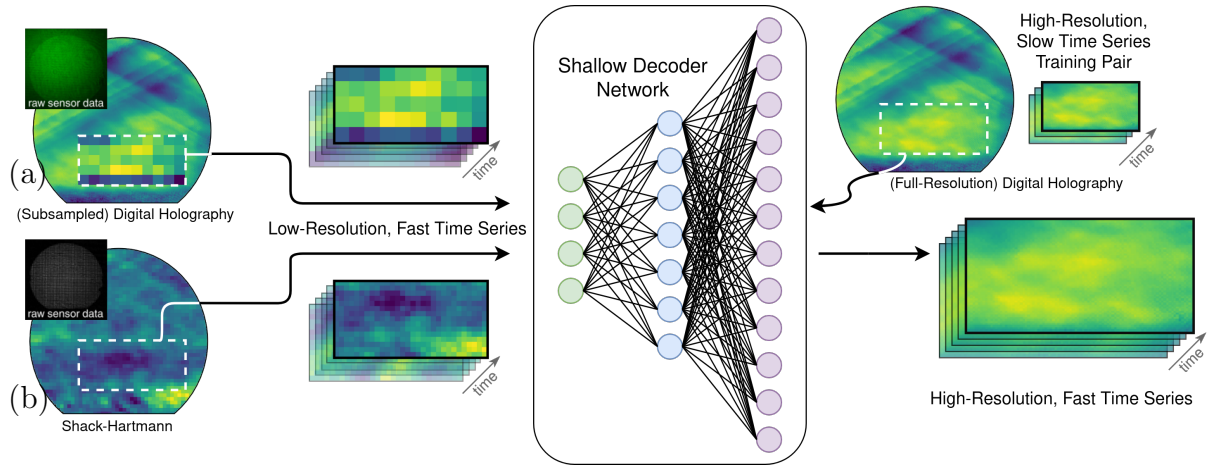

Figure 5.1: Overview of sensor fusion via a SDN. There are two possible fusion mappings to consider: (a) mapping between low-resolution and high-resolution captures of the same sensor type, such as mapping a coarse/fast DHWFS to a fine/slow DHWFS in order to achieve a fine/fast amalgamation, or (b) mapping between varying sensor types, such as between a relatively coarse/fast SHWFS to a fine/slow DHWFS.

5.3 Aero-Optics

For airborne laser systems, the high speed and high Reynolds number compressible flows around the airborne optical platform results both aero-mechanical and aero-optical perturbative effects. Aero-mechanical interactions impact beam stability and present as beam jitter, diminishing far-field intensity. Aero-optical disturbances create distortions both in the far-field as well as in the immediate vicinity of the aircraft or optical platform geometry within turbulent boundary layers (TBLs), shear layers, and wakes, as well as shock waves in the case of transonic and supersonic flows. [101, 31]. Our paper focuses on such aero-optical effects.

The aerodynamic environment of airborne laser platforms motivates high-fidelity computational fluid dynamics models which can, for example, describe boundary layers along surfaces and over any object within the flow volume. As a laser beam propagates through

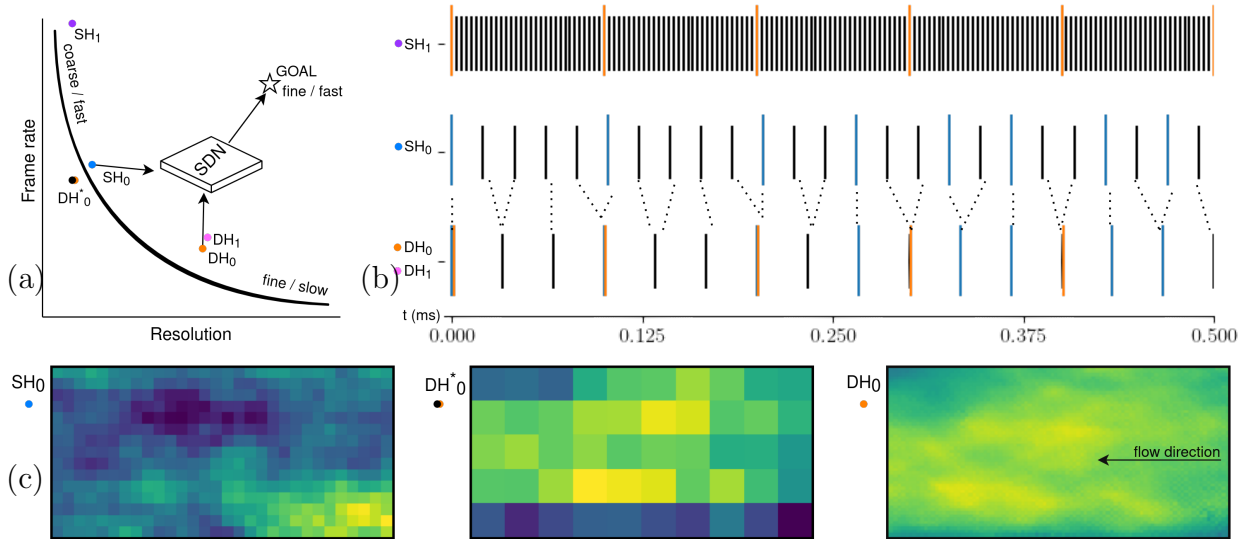

Figure 5.2: (a) A heuristic demonstrating the frame rate and resolution trade-off for image sensors. AEL datasets for SH and DH sensors are noted and correspond to the information in Table 5.1. The DH_0^* dataset represents subsampled DH data used for resolution fusion in Figure 5.6. (b) An event plot of when snapshots from four datasets occur over their first 0.5 m s of run time. (c) Sliced OPD data from frames of the indicated datasets.

the turbulent flow and TBLs, refractive index fluctuations create phase aberrations, and the resulting distortions of the optical field are referred to as aero-optical effects. Air density gradients lead to deviations in the index of refraction, n, such that

$$n(\mathbf{r}) = 1 + K_{GD}(\lambda_0)\,\rho(\mathbf{r}), \tag{5.1}$$

where K_{GD} is the wavelength-dependent Gladstone-Dale factor, λ_0 is the laser wavelength, and $\rho(\mathbf{r})$ is the air density as a function of the spatial variable $\mathbf{r}$ [101]. The deviations of a wavefront from an otherwise planar wavefront may be quantified by its optical path difference (OPD). OPD is the mean-subtracted optical path length (OPL) of the beam, the latter of which is proportional to the travel time for corresponding rays. OPL may be calculated as

the integral of the index of refraction along the propagation axis,

$$\mathrm{OPL}\,(x,y,t) = \int_0^{z_1} n\,(x,y,z,t)\,dz. \tag{5.2}$$

OPD follows as the mean-subtracted OPL,

$$\mathrm{OPD}\,(x,y,t) = \mathrm{OPL}\,(x,y,t) - \langle \mathrm{OPL}\,(x,y,t) \rangle\,, \tag{5.3}$$

where the angle brackets denote a spatial average. With respect to our study, OPD serves as a 2-D spatial map that forms a time series of wavefront measurements.

Once the wavefront is measured, the sensor data can be processed and fed into an adaptive optic (AO) control mechanism to counteract the aero-optical effects [101]. For example, in a mounted laser platform, the conjugate of the detected OPD can be used to shape a deformable mirror and pre-correct an outgoing beam. Unfortunately, with many AO applications, latency in the control loop may bottleneck the ability of the system to quickly respond and adequately compensate for the distortions. The struggle to correct for supersonic OPD fluctuations that contain frequencies in the range of several kHz [70] requires especially low-latency predictors. Furthermore, the compressible fluid flow around optical platforms with flat or conformal windows results in complicated structural phenomena - horn vortices, flow separation, shear layers, Kelvin-Helmholtz vortices, re-circulation, and von Karman vortex shedding - resulting in angle-dependent aero-optical effects [46, 20] that require either extremely generalizable or fast, online AO predictors.

5.3.1 Aero-Optics Data & Sensors

Our study examines supersonic wind tunnel measurements taken by the Aero-Effects Laboratory (AEL) at the US Air Force Research Laboratory Directed Energy Directorate. [105] The AEL wind tunnel provides a 8"x8" optical glass window for highly-accurate wavefront sensor measurements, and its modular design allows for a variety of tests and measurements. The supersonic optical transmission experiments we investigate occur at Mach 2.07 and record

the optical propagation through three sensors: a SHWFS, a DHWFS, and a schlieren imaging setup, depicted in Figure 5.3. Notable is the inclusion of the DHWFS, which provides higher spatial resolution aero-optical data than predecessor experiments, such as data from the Airborne Aero-Optics Laboratory Transonic. [58] For the purpose of our sensor fusion attempts, we focus on the first two sensors, SH and DH, as they both perform a quantitative measurement of the wavefront phase using the same laser source illuminating the test section of the wind tunnel.

The physics behind the two sensors is notably disparate. A typical SHWFS consists of a lenslet array and an image sensor, such as a charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) sensor. As an incoming wavefront passes through the lenslet array, each lenslet creates a small focal spot on the image sensor. The position of these spots relates directly to the local gradients of the wavefront. For a planar wavefront, the spots form a regular grid pattern. Any deviations, $(\Delta x, \Delta y)$, away from the centroid of each lenslet indicate aberrations away from a planar wavefront. These "local tilts" provide a differential measurement of the wavefront phase, ϕ, as

$$\boldsymbol{\nabla}\phi_k \approx \frac{1}{f_L}\left(\Delta x_k \hat{x} + \Delta y_k \hat{y}\right) \tag{5.4}$$

for each k^{th} lenslet where f_L is the lenslet focal length. A least-squares fitting may be used to reconstruct a continuous wavefront as a whole.

DH uses the principles of interferometry to reconstruct both the phase and amplitude of an optical field. In this process, a coherent beam is divided into a signal and a reference optical field. The signal beam's electric field, U_S, interacts with the volume of interest and carries phase and amplitude information, while the reference beam's field, U_R, is usually undisturbed. When re-collimated and captured on a CCD or CMOS sensor, the resulting interference pattern creates a hologram whose intensity, I_H, is derived from the superposition of the beam's signal and reference electric fields such that

$$I_H = |U_S + U_R|^2 = I_S + I_R + U_S U_R^* + U_S^* U_R, \tag{5.5}$$

from which the phase of the signal beam can be extracted via Fourier analysis.

	SH_j WFS			DH_j WFS		
j^{th}	Rate (kHz)	Res. (px)	Run time (s)	Rate (kHz)	Res. (px)	Run time (s)
0	49	(15,30)	0.6594	30	(50,100)	0.7279
1	310	(14,14)	0.1613	30	(48,48)	0.7279
2	320	(11,28)	0.1563	30	(38,97)	0.7279

Table 5.1: Overview of AEL datasets for both the SHWFS and DHWFS. Each of the j experiments simultaneously captures the AEL wind tunnel optical flow in datasets SH_j and DH_j. For each sensor dataset, we display the corresponding frame rate, effective spatial resolution used in this paper, and overall run time.

The three datasets we considered from the AEL experiments are noted in Table 5.1. In each experiment, the DHWFS is running at a comparatively slower frame rate of 30 kHz but at higher spatial resolution of the OPD, and the SHWFS is comparatively running at a faster frame rate with coarser resolution. The DHWFS captures 3.84 pixels in each linear direction per pixel in the SHWFS, making it approximately 15 times better resolved on a per-pixel basis, but the overall resolution difference also depends on the window size. On the left of Figure 5.1 we show frames from the SH_0 and DH_0 datasets, both of which image the entire wind tunnel window. The presence of shock wave artifacts within the tunnel prompt us to crop smaller, shock-free regions for this study to create the time series we need for training. While those datasets both image the entirety of the wind tunnel, the sampling window for the SH sensor in datasets SH_1 and SH_2 was reduced to accommodate their higher frame rate. These two datasets use smaller, centered SH windows which were placed directly over such shocks, the effect of which we note in our results. For the purpose of training our SDN, we sliced DH_1 and DH_2 to match these centered SH windows.

The raw data were preprocessed prior to model training. The preprocessing steps included slicing the AEL OPD data from both the SHWFS and DHWFS to omit, when possible,

volumes of the flow dominated by shock waves. Additionally, for training a map from a SH to DHWFS for sensor modality fusion, we cropped the corresponding datasets to contain only temporally matching snapshots. As can be seen in Table 5.1, the total run time of SH and DHWFS captures differs. Perhaps more pertinently, the framerates are incommensurate ratios as well, meaning we were faced with the challenge of preparing our data to have meaningful and matching input-output training pairs. This is best visualized in Figure 5.2, where snapshots over the first 0.5 ms of run time are displayed. We elected to use two methods to create adequate training datasets for machine learning, both of which match input training snapshots to their nearest output counterpart. The first method enforced the uniqueness of training input and output data pairs across SH_k and DH_k datasets, which comes at the cost of less usable data. The second method matches every frame from a SH_k dataset for a given $k = 0, 1, 2$ to its nearest counterpart in DH_k, creating much more usable data but introducing redundancies that drive overfitting. An ideal sensor fusion dataset would have commensurate framerates to ensure perfect training input-output pair matches, and an ideal sensor fusion framework would be able to robustly accommodate for timing mismatches.

5.4 Sensor Fusion: Resolution fusion

Here we demonstrate the procedure of resolution fusion, first with a surrogate model and then with AEL experiment data. That is, we examine how a single type of sensor can be used at two frame rates and two accompanying spatial resolutions, to create a neural network map that will give us the best of both worlds: a high-speed and high-resolution sensor.

5.4.1 Example: Kuramoto-Sivashinsky

Before examining AEL data, we look at a surrogate model simulated from the chaotic flow of the Kuramoto-Sivashinsky (KS) equation. The KS equation is a nonlinear scalar wave partial differential equation originally derived to model the thermal instability in laminar

flame fronts. Its 1-D form commonly used in fluid dynamics is

$$\begin{aligned} u_t + uu_x + u_{xx} + \gamma u_{xxxx} &= 0, \\ (x,t) &\in \mathbb{R} \times \mathbb{R}^+, \\ u(x,0) &= u_0(x), \\ u(x+2\pi,t) &= u(x,t) \end{aligned} \tag{5.6}$$

with a positive γ serving as a viscosity parameter. When considered on the periodic domain $0 \leq x < 2\pi$, chaotic dynamics present themselves after a cascade of period-doubling bifurcations. To simulate KS flow, we chose a viscosity value of $\gamma = 0.005$ to ensure a periodic-doubling chaotic transition and sampled 2048 points from the standard normal distribution, $\mathcal{N}(0,1)$, to seed the spatial initial condition, such that $u_0 \leftarrow \mathcal{N}(0,1)$. We ran the simulation from $0 \leq t \leq 10$ taking $100,000$ snapshots and discarding the first 2000 of those snapshots to dispose of any transient effects from the initial condition. We then subsample the simulation to create analogues of two sensor readings, a 32×98000 set with high temporal resolution and a 2048×12250 set with high spatial resolution.

Training our SDN involved a PyTorch framework where both grid searches and Bayesian hyper-parameter optimization were used to determine a suitable number of neurons per layer. The hyper-parameter training tool `Guild AI` was used for this process. Training was performed on a single NVIDIA 3080Ti GPU with CUDA 11.7, and individual runs elapsed an average of 25s. Early stopping was elected to truncate validation loss when increasing by greater than 1% after 5 epochs. The SDN for KS resolution fusion performed optimally with a first layer of $l_1 = 16384$ neurons and a second bottleneck layer of $l_2 = 64$ neurons with training ceasing typically within 100 epochs. A consistent dropout of 0.1 was used after the training the first layer, and unstructured L_1 pruning of 0.2 of the neurons was used after the 20th and 50th epochs.

Figure 5.4 highlights the results the SDN's reconstructive ability. Depicted are 500 snapshots from the validation set, X, as well as the desired high-resolution result, Y. The resulting reconstruction, $\hat{Y}$, is shown in the final panel, and a singular snapshot from all

three panels respectively (x_i, y_i, and $\hat{y}_i$) are shown superimposed on the image at an arbitrary time. We see that the resulting fit has super-resolved our input data for each available snapshot, creating a high-speed, highly-resolved fusion. Despite the chaotic instabilities of the KS equation, the temporal trends of our reconstruction appear well-resolved.

5.4.2 AEL aero-optical wind tunnel data

Figures 5.5 and 5.6 respectively demonstrate a resolution fusion mapping from SH_0 to itself and DH_0 to itself. Here we have achieved a temporally continuous mapping at a high frame rate of heavily subsampled sensor readings to their corresponding high fidelity training pairs.

Again, a Bayesian and grid-search of search of hyper-parameters tuned the model. SDNs with a first layer of approximately 2048 neurons and a bottle-necked second layer of 64 or 256 neurons appeared to best generalize to the data. We maintained a dropout of 0.1 after the first layer, and added unstructured pruning stages of strength 0.2 after the 20th, 50th, and 100th epochs.

In Figure 5.5, the SH_0 dataset has been subsampled to procure coarse training data in the first frame, $\hat{x}_i$, compressing the data by a factor of 3 in each spatial dimension for a 9 fold reduction in the spatial resolution of the SHWFS. The reconstruction, $\hat{y}_i$, successfully resolved the higher resolution features seen in the truth data in the center panel, $\hat{y}$.

Figure 5.6, similarly displays the resulting resolution fusion using the DH_0 dataset. The subsampled data shown in the first panel, $\hat{x}_i$, is analogous to a frame coming from the aforementioned DH_0^* dataset of Figure 5.2. The compression here is approximately a factor of 50 when compared to the truth data, y_i. The reconstruction, $\hat{y}_i$, resolves the higher resolution features and returns a finely resolved image.

The data in Figure 5.7 showcases the results of various dataset compression ratios versus a normalized loss. The loss tends to plateau after the compression has significantly degraded the training input dataset, beyond which the SDN seems to only capture background statistics.

5.5 Sensor Fusion: Multi-modal sensors

Now we examine multi-modal fusion, the sensor fusion paradigm where differing sensors can be combined to achieve superior results. The end-goal for our paper remains the same: create a neural network map to produce an ultra-fast and ultra-fine sensor. Unlike resolution fusion, however, we are now incorporating differing physics through the varying measurement latent spaces of the SH and DH sensors.

Figures 5.8, 5.10, and 5.9 demonstrate the results of mapping SH_k to DH_k for AEL experiments $k = 0, 1, 2$ respectively. Here we note a mild success, an intriguing disparity of results, and a lesson for future sensor fusion attempts respectively. Our SH_0/DH_0 fusion attempt was the only dataset pairing where we managed to avoid prominent shock waves in the background of the OPD data. This training set, however, ran at the lowest framerate, and with about 9200 usable snapshots the resulting mapping mostly picks up background oscillations with high noise.

The SH_1/DH_1 and SH_2/DH_2 fusion attempts began with SHWFS windows sitting in the center of the AEL wind tunnel test section. That is, the 14×14 frame resolution of the SH_1 WFS and the 11×28 resolution of the SH_2 WFS are unavoidably capturing OPD data in a region exposed to shocks from the wind tunnel nozzle's reported defects. Unsurprisingly, the resulting sensor fusion captures these shocks exquisitely, while neglecting to map higher frequency flow information of interest.

A frequency analysis of these mappings is elucidating, however. We perform a radial power spectral density (PSD) analysis of the SH_1/DH_1 fusion mapping as an example in Figure 5.11. This dataset maps the 14×14 spatial pixels of each frame of a SHWFS running at 310kHz to a corresponding region of 48×48 pixels of a DHWFS running at 30kHz. A PSD analysis of each frame is computed, as shown for several example frames at the right of the figure. We sum the PSD at each radius, r, from the center of the frame as measured in pixels of the higher resolution dataset. The radial PSD for all frames span the banded plot on the left panel, and the mean radial PSD is highlighted by a thick curve. The takeaway is that

our desired DH output, y_i, and our initial SH input, x_i, are not capturing the same frequency information. What we successfully are able to achieve is a recovery of the higher-resolution frequency content, as seen by the curve for our reconstructed DH sensor, $\hat{y}_i$.

5.6 Conclusions

Machine learning algorithms can be leveraged to exploit multiple sensor modalities for state estimation and forecasting, even within challenging environments such as aero-optical wavefront forecasting. Emerging algorithms from the machine learning community can be integrated with many traditional scientific computing approaches to enhance sensor fusion capabilities. To address this challenge, we have proposed an SDN in this manuscript that maps spatially coarse but temporally high-speed sensors to their spatially fine but temporally lower-speed counterparts. Thus a low-resolution sensor, such as a windowed SHWFS operating at a high framerate, can be fused either with results from a more highly-resolved SHWFS or another sensor altogether like a DHWFS to produce high-resolution, high-speed estimates from a quickly-trained SDN model.

With resolution fusion, such as the mappings presented in Figures 5.5-5.6, there can be diminished need to limit a sensor's spatial span as well. The AEL datasets SH_1 and SH_2, for example, operated at 310 kHz and 320 kHz respectively by limiting the sensor window to record a fraction of the wind tunnel test section. While we presented results for a resolution fusion mapping between similarly sized test section windows, our methods can be adapted in future work to map across window sizes at varying framerates, effectively increasing both the spatial resolution and also the spatial span. With multi-modal fusion, the SDN provides a flexible, promising platform to map disparate sensor types so that we not only combine the spatial and temporal resolutions but also potentially benefit from the differences in sensor physics underlying each modality.

A key challenge in our framework is the limited ability to incorporate incommensurate framerate ratios when constructing a training set. An ideal training set would consist of input/output pairs taken at precisely the same time, and the slight timing mismatch we

encountered when constructing multi-modal training pairs from the SHWFS to the DHWFS proved challenging to work around. An architecture that would relax timing requirements for its training data would be a great boon to many sensor deployments.

Our analysis shows an underlying sensor fusion paradigm that is easily leveraged in other application areas, however. Indeed, the SDN architecture is agnostic to the underlying application. A method is simply provided whereby sensor modalities, or different sensor performance characteristics, can be mapped to each other using a representative training data set. The SDN ensures that the data demands for the training remain reasonable. The mathematical architecture also helps achieve the ultimate goal of using the cheapest sensor possible to produce high-resolution state estimates at the high-sampling rates.

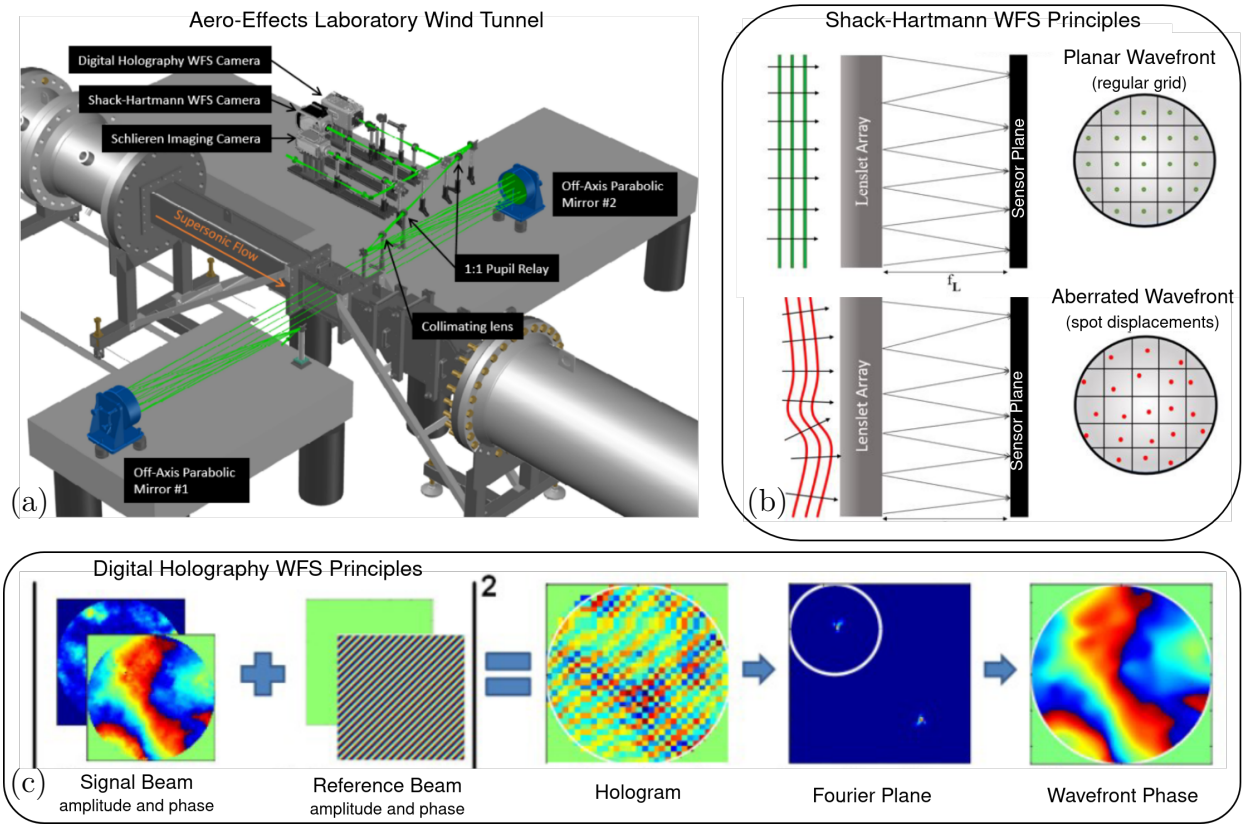

Figure 5.3: (a) Aero Effects Laboratory wind tunnel. (b) Shack-Hartmann WFS operational diagram. A planar wavefront results in a regular grid of spots when transmitted through the lenslet array to the sensor plane, while aberrated wavefronts manifest as spot displacements. (c) Outline of Digital Holography for wavefront phase estimation. *Figures adapted with permission from Reference [107]*

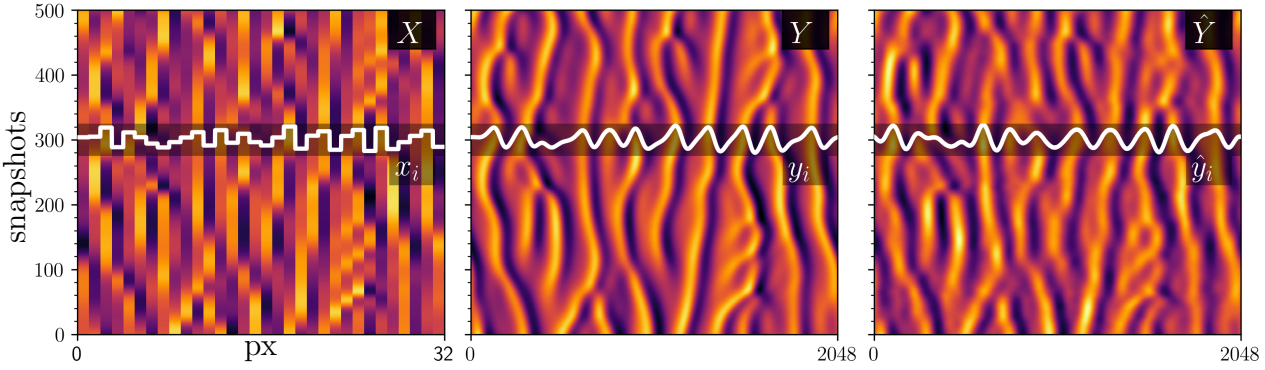

Figure 5.4: KS sensor fusion process demonstrating a 64 fold increase in spatial resolution, mapping 32 pixel input vectors to 2048 pixels outputs. All panels have the same colormap and plot limit scaling. Pictured to the left are 500 snapshots from the validation input set, X, with each row, x_i, depicting a 32 pixel, 1-D snapshot. The second panel depicts the desired results or ground truth, Y, and the third panels shows the SDN's predicted reconstruction, $\hat{Y}$. Additionally, a single snapshot arbitrarily chosen at $t = 300$ is overlaid as an inset in each figure.

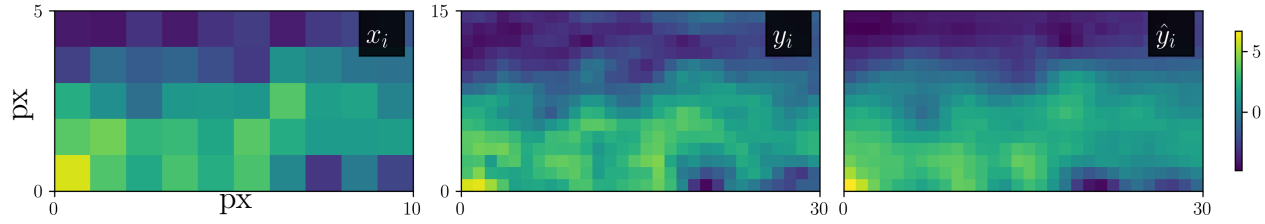

Figure 5.5: Mapping a subsampled SHWFS input with low-spatial/high-temporal resolution to a higher spatial resolution result, after training against high-spatial/low-temporal resolution SHWFS frames. Pictured to the left is a single snapshot frame from the validation input set, x_i. In the center panel, we have the ground truth y_i, and the SDN's mapped result, $\hat{y}_i$, appears in the third panel. All figures have the same colormap scaling.

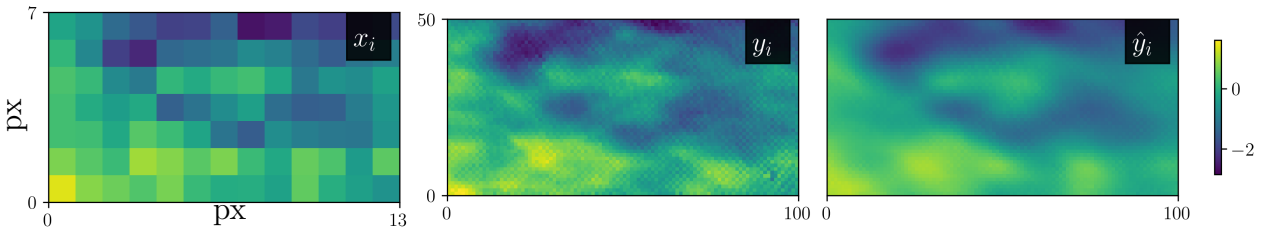

Figure 5.6: Mapping a subsampled DHWFS input with low-spatial/high-temporal resolution to a higher spatial resolution result, after training against high-spatial/low-temporal resolution DHWFS frames. Pictured to the left is a single snapshot frame from the validation input set, x_i. In the center panel, we have the ground truth y_i, and the SDN's mapped result, $\hat{y}_i$, appears in the third panel. All figures have the same colormap scaling.

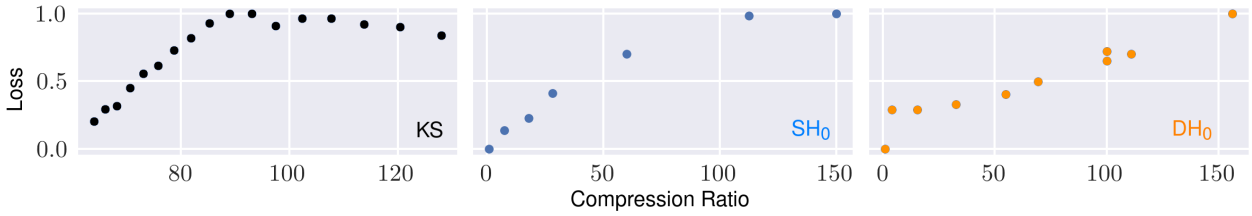

Figure 5.7: Resolution fusion loss versus data compression ratio. Each point indicates the normalized validation loss of an experiment with similar SDN settings for a subsampled, and therefore compressed, input training dataset.

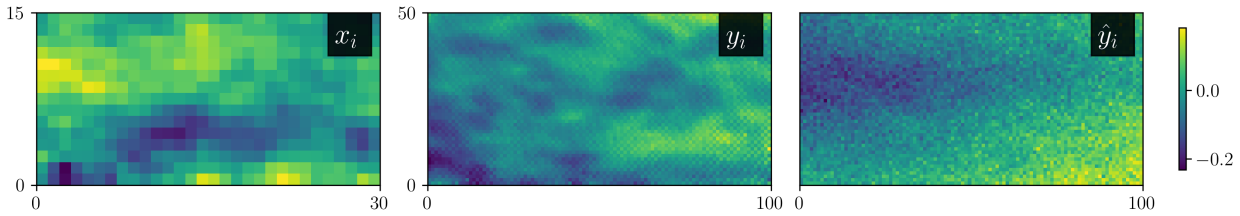

Figure 5.8: Sensor fusion mapping of SH_0 to DH_0.

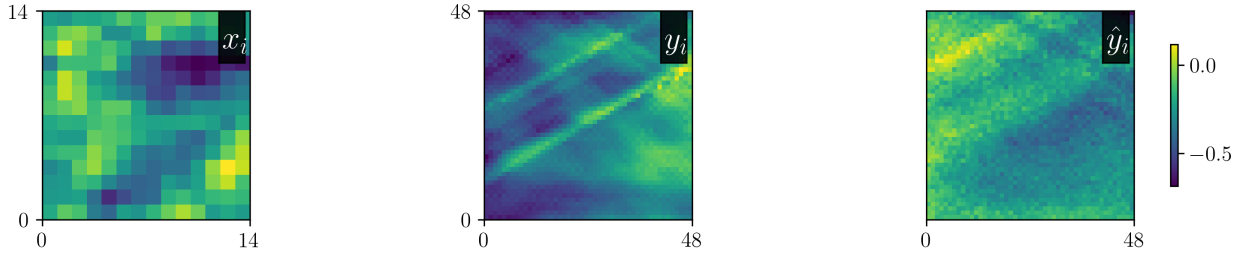

Figure 5.9: Sensor fusion mapping of SH_1 to DH_1.

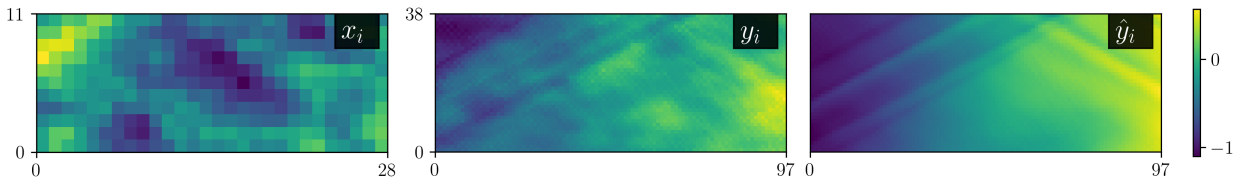

Figure 5.10: Sensor fusion mapping of SH_2 to DH_2.

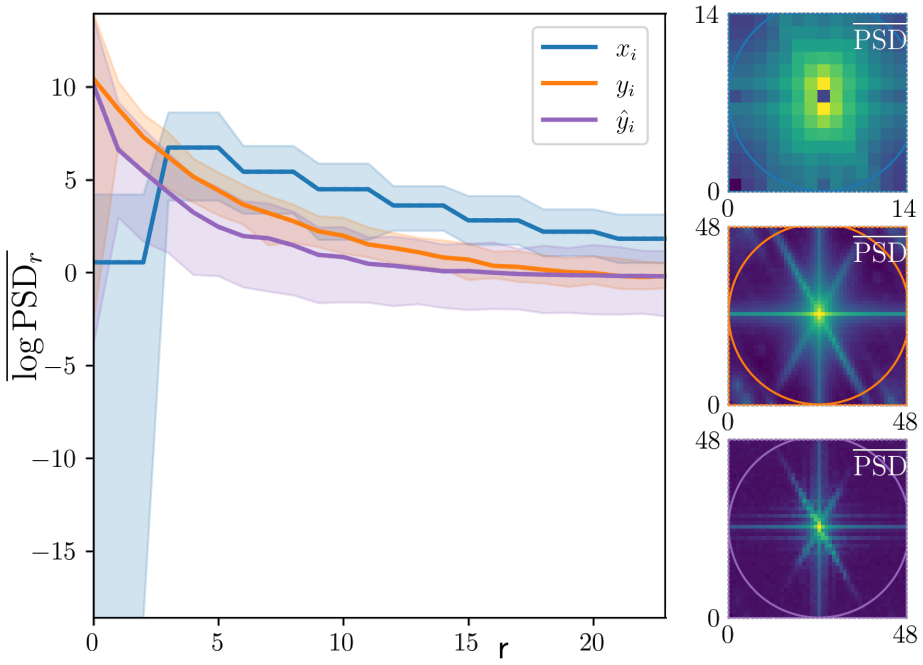

Figure 5.11: PSD analysis of the SH_1 to DH_1 sensor fusion mapping. On the left, we see the mean radial PSD for the input training set from the SHWFS (x_i), desired output from the DHWFS (y_i), and reconstructed mapping from sensor fusion ($\hat{y}_i$). The radius, r, is measured in pixels of the high-resolution dataset. On the right we see sample PSDs for each of the frames in Figure 5.9. The circular outline depicts where the mean radial PSD is computed for the maximal radius. As the curves in the left panel indicate, the PSDs for the low-resolution SH_1 dataset deviates from that of the higher resolution DH_1 dataset. Our multi-modal mapping generates a spectrum closer to the high-resolution result.

Chapter 6

SHALLOW RECURRENT NETWORK WAVEFRONT FORECASTING

This work presents an in-review submission and is largely self-contained. A completed publication titled after this chapter is thus expected for later in 2023.

6.1 Summary

Sensing is a fundamental task in the monitoring, forecasting and control of engineering systems. For a given temporal and spatial resolution, sensors are tasked with extracting as much information as possible from a limited number of sensing locations. Recently, it has been shown that the measurement time history encodes a significant amount of information that can be extracted for critical tasks. Most model-free sensing paradigms aim to map current sparse sensor measurements to the high-dimensional state space, ignoring the time-history all together. Using modern deep learning architectures, we show that a sequence-to-vector model that encodes the time history, such as an LSTM (long, short-term memory) network, with a decoder network that maps limited sensors to the full state space estimates, we can (i) accurately reconstruct the full state space using a limited number of arbitrarily placed sensors, and (ii) use the LSTM to predict future states for adaptive control architectures such as model predictive controllers. In this manuscript, our *shallow recurrent decoder*, or SHRED, is used to augment aero-optical wavefront sensors which monitor the turbulent optical field produced inside transonic flow fields. The neural network achieves high-fidelity state reconstruction and forecasting using sparse, robust sensor placements, useful for data compression and adaptive optic predictive control.

6.2 Introduction

Aero-optical physics is characterized by the propagation of a laser beam in a high-dimensional turbulent flow which exhibit a continuous cascade of multiscale temporal and spatial features. Advances in optical sensor technology, such as digital holography for aero-optical wavefront characterization[107, 106], have driven the quality and quantity of collected data in such systems, allowing for new opportunities for accurate flow characterization with the ultimate goal of real-time beam control. In airborne optical environments, aero-optical effects such as turbulent boundary layers cause rapid index of refraction fluctuations that induce aberrations in transmitted wavefronts [101]. The fast scale physics in the turbulent flows lead to an extremely low latency for control applications. Thus computationally efficient prediction and forecasting is critical in airborne and free space optical systems in order to enact control. Using modern deep learning algorithms, specifically *shallow recurrent decoder* (SHRED) networks, a computationally efficient algorithm can be constructed for meeting the latency tolerance constraints of modern aero-optic sensing in the Mach one regime.

The deployment of modern sensors and the processing of high-resolution data streams create exceptional computational demands. Often such data is approximated by low-rank linear estimation techniques such as the singular value decomposition (SVD) and proper orthogonal decomposition (POD) so as to improve computational efficiency [75]. Such techniques have are well-known for reconstructing fluid and optical flow fields [16, 82, 46]. Such linear methods are sensitive to sensor placement [75], however, with greedy algorithms having been developed in order to maximally extract information from sensors for reconstruction and flow characterization. In turbulent flows such as aero-optics, POD decompositions are not low rank due to the continuous cascade of scales, and thus traditional dimensionality reduction techniques are difficult to implement.

An alternative to linear dimensionality reduction is the use of modern deep learning algorithms. Specifically, it has been recently shown that SHRED with both stationary [108] and mobile [32] sensors are capable of leveraging sensor trajectory information for exceptional

performance. Such lightweight deep learning architectures, like the three layer shallow decoder networks (SDNs) in SHRED, work well for super-resolving and fusing sensor data via latent space mappings [34, 83, 20, 90]. While robust to sensor placement, SDNs are however unreliable in extrapolating data due to their inability to encode temporal trends.

In this work, we present a recurrent neural network using long short-term memory (LSTM) units to integrate the sensor time history to a SDN. SHRED is critically enabling for the turbulent flows encountered in aero-optics, allowing one to not only reconstruct aero-optical data from a sparse sampling of sensors but to forecast wavefronts for a limited duration. Indeed, the forecast accuracy is capable of meeting the stringent latency demands required of aero-optic beam control. As such, SHRED as a lightweight deep learning algorithm can provide a critically enabling technology capable of meeting the performance demands of supersonic aero-optical flows.

Figure 6.1 overviews the goal and methodology of this manuscript: snapshots of airborne optical wavefront data are sampled at various sensor locations and used as training inputs for the SHRED model. Training output pairs include a time-step advancement of sensor data as well as the full state representation of the wavefront. Therefore, two models are trained for each SHRED iteration: one to forecast the inputted sensors and the other to reconstruct high-resolution wavefront states. Together, we show in this manuscript how SHRED can not only reconstruct full-state wavefront representations from sparse sensor placements but also forecast the full-state for short timescales suitable for use in adaptive optic control systems.

6.3 Methods

The SHRED architecture in Figure 6.1 maps trajectories of sensor measurements to, in our case, the high-dimensional feature state of the full wavefront. The architecture can be written as

$$\text{SHRED}\left(\{\mathbf{s}_i\}_{i=t-m}^{t}\right) = F\left(R\left(\{\mathbf{s}_i\}_{i=t-m}^{t}; \mathbf{W}_{LSTM}\right); \mathbf{W}_{SDN}\right) \tag{6.1}$$

where $\mathbf{s}_i$ is a sensor measurement of the higher-dimensional state $\mathbf{x}_i$ at time t_i, F is a densely connected, feed-forward shallow neural network with weights $\mathbf{W}_{SDN}$, and R is a recurrent,

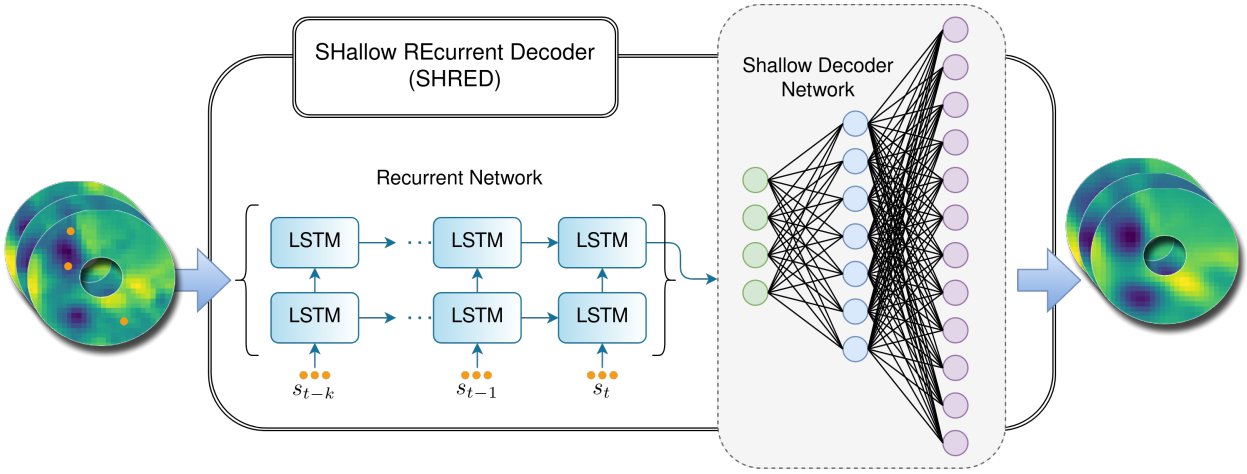

Figure 6.1: SHRED network for aero-optical flow reconstruction from sensor point measurements. SHRED a workflow for sparse sensor measurements using LSTM recurrent networks fed into a SDN for sensor forecasting and full state reconstruction.

LSTM neural network with weights $\mathbf{W}_{LSTM}$. Note that $\{\mathbf{s}_i\}_{i=t-m}^{t}$ is a sparse set of sensor measurements over the past m snapshots. Thus, each sensor has a defined memory and is providing that measurement history to SHRED. When selecting the number of such sensors, SHRED excels at robustly using a sparse selection of sensor placements for reconstruction. For example, the SHRED network is capable of reconstructing simulated isotropic turbulent flow with only three sensors[108]. Unlike the original implementation of SHRED, which joined an LSTM directly to a three layer SDN, we removed the first layer of the SDN because the LSTM itself acts as the input layer. As a result, we have fewer parameters to train and a lighter, faster network. More GPU memory can be allocated to training multiple vertical stacks of LSTM networks instead.

Training SHRED involves minimizing the L_2 loss using the ADAM optimizer [62] given the sparse set of sensor measurements $\{\mathbf{s}_i\}$ and corresponding high-dimensional state measurements $\{\mathbf{x}_i\}$. The PyTorch ADAM weight decay parameter is used in training to provide L_2 regularization of the weight matrix as well. [73] All training was performed on a single

NVIDIA 4070Ti GPU with CUDA 11.7 using grid-searches for hyper-parameter optimization.

6.4 Data

We assess in-flight, transonic data from the Airborne Aero-Optics Laboratory Transonic [58] (AAOL-T) experiment. The experiment uses two Falcon 10 aircraft to record aero-optical phenomena, and, in this study, the craft are flown 50 m apart in tandem at Mach 0.6 while recording 532 nm laser transmissions. Figure 6.2 depicts the AAOL-T experiment, along with the airborne optical platform and its Shack-Hartmann (SH) wavefront sensor (WFS). The optical platform, a hemispherical laser turret, is more closely detailed in Figure 6.3. We were provided 23 datasets of experimental data spanning differing look-back and inclination angles, α and β, that parameterize the beam's direction. The more pertinent parameter for our study is α, with which OPD tends to increase from nearly laminar flow for forward-facing angles, $\alpha \ll \frac{\pi}{2}$, to highly turbulent flow for rear-facing angles, $\alpha \gg \frac{\pi}{2}$. The inclination angle, β, is centered at the mid-line of the airborne optical turret at $\beta = \frac{\pi}{2}$, deviations from which tend to bring the beam propagation direction through necklace and horn vortices away from the hemisphere's mid-line. We aim in this manuscript to demonstrate the ease at which SHRED reconstructs the wavefront OPD in any of these regimes.

We elect to input 12 radially chosen sensors with a memory of $m = 60$ snapshots for this experiment: 9 sensors equiangularly placed along the outer radius of the SHWFS and 3 equiangularly placed around the inner radius. Sensor input vectors therefore have a dimension of 720. Because the AAOL-T SHWFS operates at 30 kHz, the chosen sensor memory histories corresponds to 2 ms of recorded data at each given pixel. Since each pixel in the SHWFS corresponds to a lenslet and CCD sub-aperture, we are effectively using a partial, sparsely distributed SHWFS with our input data composed of approximately 2% of the original sub-apertures.

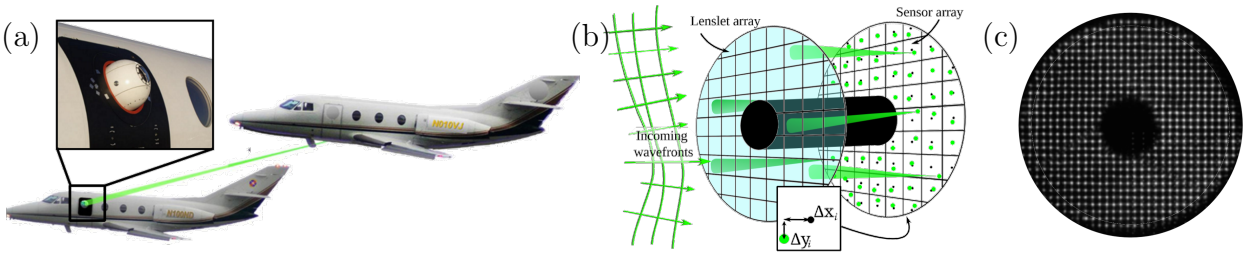

Figure 6.2: The AAOL-T experiment. (a) Dual Falcon 10 aircraft are flown in tandem at Mach 0.6, in the case of this study. Flow separation over the airborne optical platform's geometry results in a turbulent boundary layer. (b) Wavefront aberrations are detected using a SHWFS, whose lenslet array focuses incoming wavefronts onto CCD sensors and measures spot displacements $(\Delta x_i, \Delta y_i)$ for each sensor. (c) A frame of raw SHWFS spot displacement data. The wavefront OPD is reconstructed via a least squares method, stitching together the recorded "local tilts". *Figure of the AAOL-T aircraft adapted from Reference [58]*

6.5 Results

Figure 6.4 shows results from the test set of a highly turbulent AAOL-T dataset, in which the airborne optical system is aiming at the rear of the craft at a look-back angle of $\alpha = 139.0°$ into the turbulent wake with inclination $\beta = 89.5°$. Each row of the panels in Figure 6.4a respectively represents a vectorized snapshot of AAOL-T wavefront data as truth data, a SHRED reconstruction from sensor histories $\{s_i\}_{i=t-m}^{t}$, and a SHRED forecast. The forecasting in the third panel is achieved via one reconstruction that is fed forward through the network incrementally in singular time steps to forecast future sensor states; the sensor states are then reconstructed into a full state forecast. The test dataset truth and reconstruction show a strong match, while the forecast, on a millisecond scale, shows oscillatory trends.

In these panels, the initial sensor placements and memories are depicted by the colored arrows, whose locations may be best visualized as the corresponding colored pixels in Figure 6.4b, which examines a single snapshot reshaped to its original form. The truth value, reconstruction, and forecast of the snapshot are shown at $t = 150$, which is well beyond the

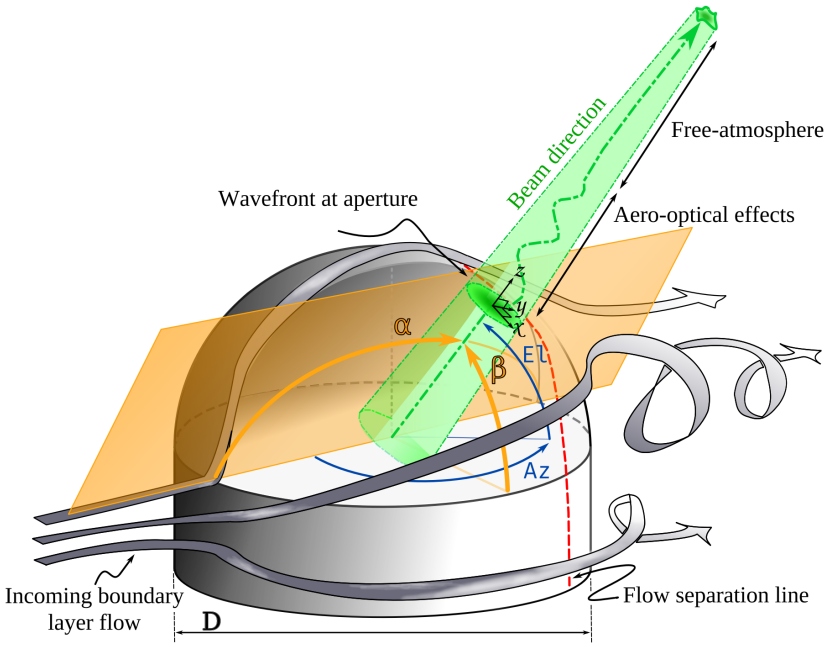

Figure 6.3: Hemispherical laser turret geometry. Aberrated wavefronts are imaged at the aperture (green disc with local coordinate system x-y-z). The look-back angle and inclination angle, α and β respectively, parameterize the beam direction. Wavefront aberrations are highly dependent on beam direction. OPD tends to increase with *alpha* past the red dotted line which indicates the start of flow separation.

sensor memory timescale. Each individual sensor has its values plotted in Figure 6.4c, again with colors corresponding to the previous panels. Therefore, the orange-brown sensors are towards the source of the flow while the blue sensors are near the rear.

This particular experiment continues its forecast seemingly indefinitely. However it should be noted that the typical trend is to settle into an attractor within several milliseconds of the forecast. More importantly, because adaptive optic systems operates on the MHz scale, we emphasize that for aero-optical predictive control we only require several frames of forecast. Figure 6.8 depicts the first 0.5 ms of the forecast for each sensor, overlaid on top of true sensor measurements. The majority of sensors provide sensible predictions within this window.

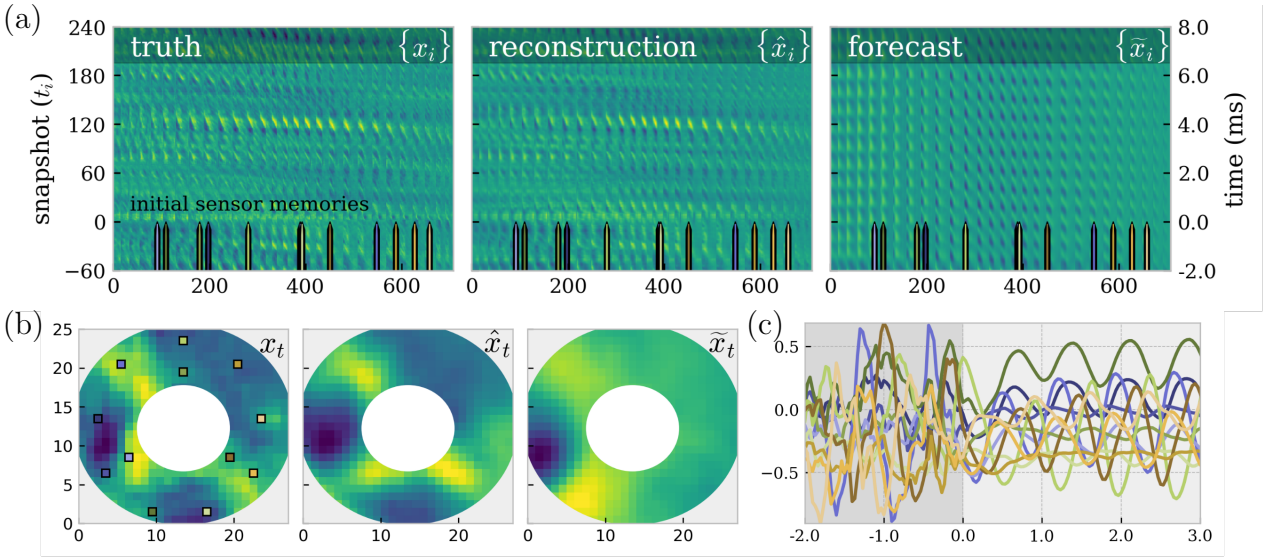

Figure 6.4: (a) Truth, reconstruction, and forecasted snapshots from running SHRED on a AAOL-T test set, each represented by x_i, $\hat{x}_i$, and $\tilde{x}_i$ respectively. The inputs to SHRED are 12 radially dispersed sensor measurement, depicted by the colored arrows, with each sensor history retaining a memory of 60 snapshots, which corresponds to 2 ms of OPD history. (b) A reshaped wavefront snapshot is shown from the top panel at $t = 200$. Sensor locations are displayed, colored to match the remaining sub-figures. (c) The forecasted time series for the 12 sensors is shown. Negative values, shaded in dark gray, represent the sensor memory, after which the proper forecast begins.

Figures 6.7-6.8 depict another dataset, this time with an $\alpha = 87.7°$ and $\beta = 81.4°$. This relatively forward-looking beam direction occurs before the flow separation region of the hemispherical laser turret, and as a result the normalized, root-mean-square OPD of the dataset's wavefront sits at 0.52, one fourth of the previous example's value of 2.08. As expected, the less turbulent flow of this dataset proves simpler to capture using SHRED. The short time sensor forecasts in Figure 6.8 particularly showcase the relative ease of forecasting a full 0.5 ms into the future for all but one of the 12 sensors.

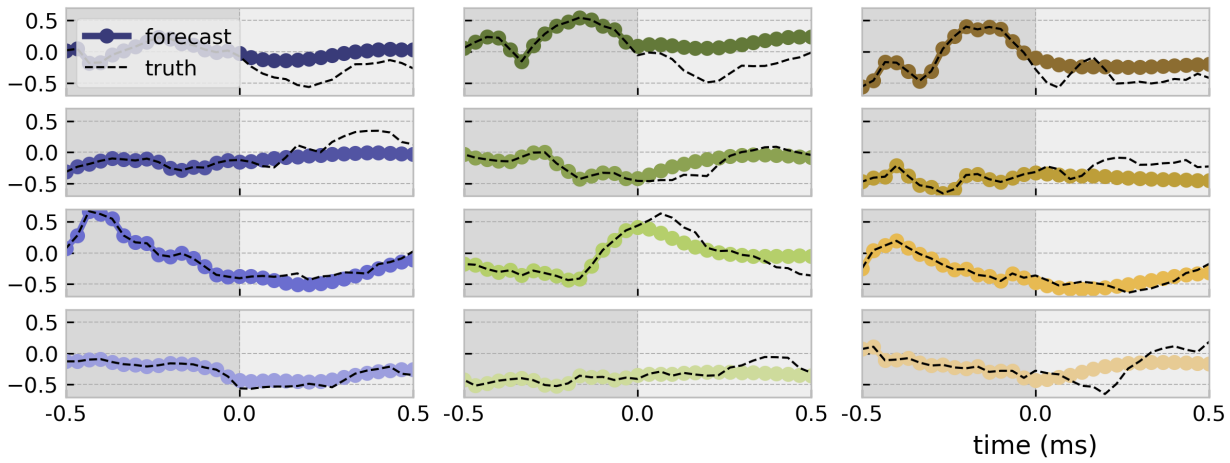

Figure 6.5: Forecasts of each individual sensor depicted in Figure 6.4.

6.6 Conclusion

We demonstrate the deep learning SHRED architecture, a recurrent shallow decoder neural network, on high-dimensional turbulent aero-optical wavefront data. Despite the turbulent, multiscale physics of the flow field, the SHRED method is able to reconstruct the flow physics from limited measurements and produce a future state estimate within the stringent latency tolerances required for controlling laser beams in the turbulent boundary layer. Because of its robust ability to use sparse, flexibly located sensors, SHRED doubles as a high-fidelity data compressor that serves as an alternative to conventional reduced-order models. The ability to leverage temporal memory via LSTM units also allows SHRED to forecast time series, making it a candidate predictive controller in challenging optical environments such as AO for aero-optical correction.

The success of SHRED is based upon its ability to leverage the sensor time history for encoding global information of the spatio-temporal field. Indeed, unlike most sensing modalities where only the current time measurement is used for reconstruction or state estimation, SHRED trains a recurrent neural network (LSTM) with an accompanying decoder network to leverage dimensionality reduction and dynamics jointly. This provides an efficient nonlin-

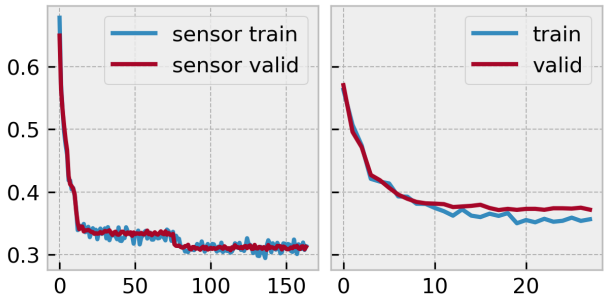

Figure 6.6: The training and validation loss curves for the SHRED model in Figure 6.4. The first set for fitting sensor forecasts and the latter for reconstructing the high-dimensional state from the sensors.

ear dimensionality reduction technique that can accurately encode the flow physics even in aero-optics where turbulence dominates and low-rank linear representation fails. The success in aero-optics and elsewhere [108, 32] are suggestive of the universal applicability of the SHRED architecture.

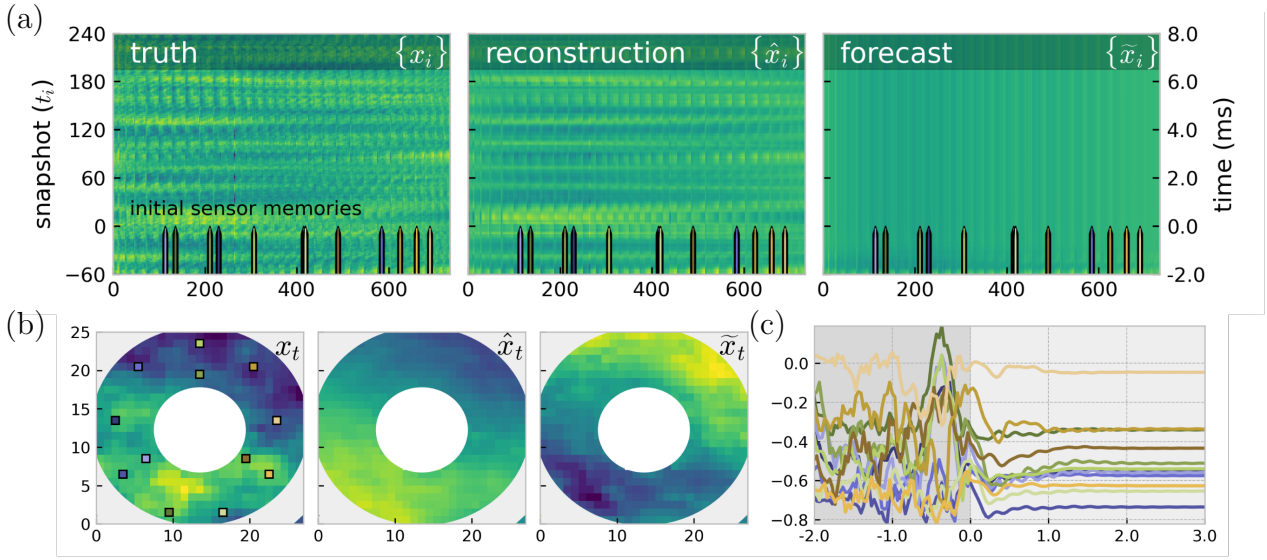

Figure 6.7: (a) Truth, reconstruction, and forecasted snapshots from running SHRED on a AAOL-T test dataset 03, each represented by x_i, $\hat{x}_i$, and $\tilde{x}_i$ respectively. The inputs to SHRED are 12 radially dispersed sensor measurement, depicted by the colored arrows, with each sensor history retaining a memory of 60 snapshots, which corresponds to 2 ms of OPD history. (b) A reshaped wavefront snapshot is shown from the top panel at $t = 200$. Sensor locations are displayed, colored to match the remaining subfigures. (c) The forecasted time series for the 12 sensors is shown. Negative values, shaded in dark gray, represent the sensor memory, after which the proper forecast begins.

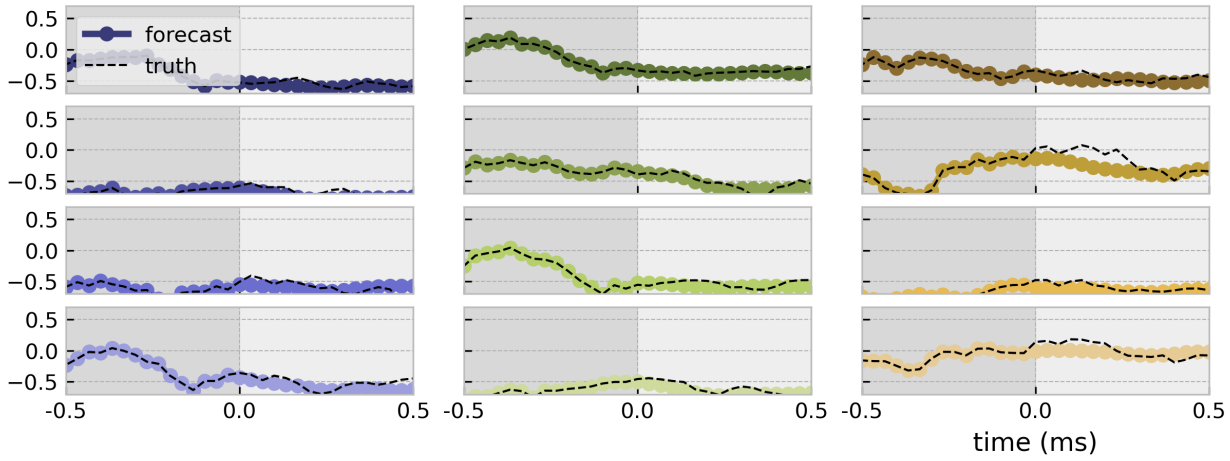

Figure 6.8: Forecasts of each individual sensor depicted in Figure 6.7.

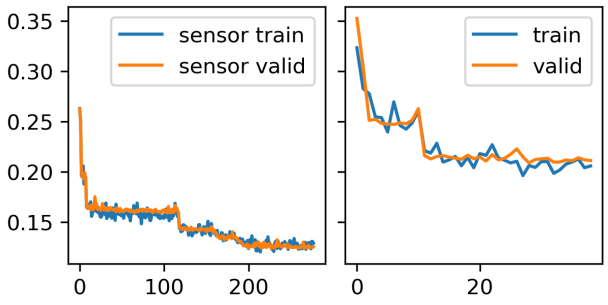

Figure 6.9: The training and validation loss curves for the SHRED model in Figure 6.7. The first set for fitting sensor forecasts and the latter for reconstructing the high-dimensional state from the sensors.

Chapter 7

CONCLUSION

When I meet God, I am going to ask him two questions: Why relativity ? And why turbulence ? I really believe he will have an answer for the first.

—Werner Heisenberg (or, perhaps, Horace Lamb)

In the evolving landscape of aero-optics research, the pertinence of innovative machine learning methodologies has emerged as a crucial step towards enhancing the forecasting and sensing of aberrated wavefronts in airborne optical systems. Through this research, three machine learning techniques are demonstrated, each showcasing an enhancements to the field of aero-optical wavefront sensing and control, potentially with the ability of being used in synergistic workflows.

The first significant accomplishment was the application of opt-DMD to characterize AAOL-T datasets. The opt-DMD algorithm precisely identifies coherent flow dynamics and their temporal and spatial evolutions. By leveraging this method on the AAOL-T datasets, we determined its capability to provide a debiased eigenvalue spectrum, which provides a low-latency predictor for more precise wavefront forecasting, critical for real-time adaptive optic controls.

Secondly, our exploration into sensor fusion revealed the efficacy of SDNs as conduits between diverse sensor modalities. By mapping between sensor resolutions or across sensor modalities with different physics and performance, SDNs showed a path to harness low-resolution sensors for deriving high-resolution insights. A key illustration from our study is the decoding of intricate aero-optics flow physics using techniques like digital holography and Shack-Hartmann wavefront sensing.

Lastly, we used a hybridized LSTM and SDN network, jointly named SHRED, to further

augment wavefront sensor measurements for forecasting. This framework, with its ability to harness temporal trends in the wavefront data, is exemplary at accurately reconstructing the full-state space of the wavefront and predicting future states from sparsely selected sensor data points. The ability to forecast into the near future improves aero-optical wavefront sensing capabilities in an already challenging transonic flow domain.

In summation, this thesis underscores the untapped potential of machine learning techniques in aero-optics. By paving the way for enhanced sensor precision, comprehensive data characterization, and state-of-the-art adaptive optic predictor candidates, we hope to provide a foundation for future research endeavors in this realm, ultimately pushing the boundaries of what's achievable in aero-optical wavefront control and free-space laser transmission research.

BIBLIOGRAPHY

[1] Abien Fred Agarap. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*, 2018.

[2] Leila Arras, José Arjona-Medina, Michael Widrich, Grégoire Montavon, Michael Gillhofer, Klaus-Robert Müller, Sepp Hochreiter, and Wojciech Samek. Explaining and interpreting lstms. *Explainable ai: Interpreting, explaining and visualizing deep learning*, pages 211–238, 2019.

[3] Travis Askham and J Nathan Kutz. Variable projection methods for an optimized dynamic mode decomposition. *SIAM Journal on Applied Dynamical Systems*, 17, 2018.

[4] Travis Askham and J. Nathan Kutz. Variable Projection Methods for an Optimized Dynamic Mode Decomposition. *SIAM Journal on Applied Dynamical Systems*, 17, January 2018.

[5] Omri Azencot, Wotao Yin, and Andrea Bertozzi. Consistent dynamic mode decomposition. *SIAM Journal on Applied Dynamical Systems*, 18, 2019.

[6] Peter J Baddoo, Benjamin Herrmann, Beverley J McKeon, J Nathan Kutz, and Steven L Brunton. Physics-informed dynamic mode decomposition. *Proceedings of the Royal Society A*, 479(2271):20220576, 2023.

[7] Shervin Bagheri. Koopman-mode decomposition of the cylinder wake. *Journal of Fluid Mechanics*, 726, 2013.

[8] Shervin Bagheri. Effects of weak noise on oscillating flows: Linking quality factor, Floquet modes, and Koopman spectrum. *Physics of Fluids*, 26, September 2014.

[9] Jeffrey D. Barchers, David L. Fried, and Donald J. Link. Evaluation of the performance of Hartmann sensors in strong scintillation. *Applied Optics*, 41, 2002.

[10] Martin J Booth. Adaptive optical microscopy: The ongoing quest for a perfect image. *Light: Science & Applications*, 3, 2014.

[11] T. J. Brennan and D. J. Wittich III. Statistical analysis of Airborne Aero-Optical Laboratory optical wavefront measurements. *Optical Engineering*, 52, 2013.

[12] Rasmus Bro and Age K Smilde. Principal component analysis. *Analytical methods*, 6(9):2812–2831, 2014.

[13] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz. Chaos as an intermittently forced linear system. *Nature Communications*, 8, 2017.

[14] S. L. Brunton and J. N. Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2nd edition, 2022.

[15] Steven L. Brunton, Marko Budišić, Eurika Kaiser, and J. Nathan Kutz. Modern Koopman Theory for Dynamical Systems, October 2021.

[16] Steven L Brunton and J Nathan Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.

[17] Steven L Brunton, Bernd R Noack, and Petros Koumoutsakos. Machine learning for fluid mechanics. *Annual review of fluid mechanics*, 52:477–508, 2020.

[18] Steven L Brunton, Joshua L Proctor, Jonathan H Tu, and J Nathan Kutz. Compressed sensing and dynamic mode decomposition. *Journal of computational dynamics*, 2(2):165–191, 2016.

[19] William R Burns, Eric J Jumper, and Stanislav Gordeyev. A robust modification of a predictive adaptive-optic control method for aero-optics. In *47th AIAA plasmadynamics and lasers conference*, page 3529, 2016.

[20] William R. Burns, Eric J. Jumper, and Stanislav Gordeyev. A Robust Modification of a Predictive Adaptive-Optic Control Method for Aero-Optics. In *47th AIAA Plasmadynamics and Lasers Conference*. American Institute of Aeronautics and Astronautics, June 2016.

[21] Luke N Butler and Stanislav Gordeyev. Development of the focused malley probe as a local aero-optical measurement technique. In *AIAA Scitech 2022 Forum*, page 0988, 2022.

[22] Yuan Cao, Yu-Huai Li, Kui-Xing Yang, Yang-Fan Jiang, Shuang-Lin Li, Xiao-Long Hu, Maimaiti Abulizi, Cheng-Long Li, Weijun Zhang, Qi-Chao Sun, Wei-Yue Liu, Xiao Jiang, Sheng-Kai Liao, Ji-Gang Ren, Hao Li, Lixing You, Zhen Wang, Juan Yin, Chao-Yang Lu, Xiang-Bin Wang, Qiang Zhang, Cheng-Zhi Peng, and Jian-Wei Pan. Long-distance free-space measurement-device-independent quantum key distribution. *Physical Review Letters*, 125, 2020.

[23] K. K. Chen, J. H. Tu, and C. W. Rowley. Variants of dynamic mode decomposition: Boundary condition, Koopman, and Fourier analyses. *Journal of Nonlinear Science*, 22, 2012.

[24] James M Cicchiello and Eric J Jumper. Far-field optical degradation due to near-field transmission through a turbulent heated jet. *Applied optics*, 36(25):6441–6452, 1997.

[25] Emily Clark, Steven L Brunton, and J Nathan Kutz. Multi-fidelity sensor selection: Greedy algorithms to place cheap and expensive sensors with cost constraints. *IEEE Sensors Journal*, 21(1):600–611, 2020.

[26] Richard Davies and Markus Kasper. Adaptive optics for astronomy. *Annual Review of Astronomy and Astrophysics*, 50, 2012.

[27] Scott TM Dawson, Maziar S Hemati, Matthew O Williams, and Clarence W Row-

ley. Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition. *Experiments in Fluids*, 57:1–19, 2016.

[28] Scott TM Dawson, Maziar S Hemati, Matthew O Williams, and Clarence W Rowley. Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition. *Experiments in Fluids*, 57, 2016.

[29] Nicholas De Lucca, Stanislav Gordeyev, Eric J. Jumper, and Donald J. Wittich. Effects of engine acoustic waves on optical environment around turrets in-flight on AAOL-T. *Optical Engineering*, 57, 2018.

[30] Daniel Duke, Julio Soria, and Damon Honnery. An error analysis of the dynamic mode decomposition. *Experiments in Fluids*, 52, 2012.

[31] S. Gordeyev E. Jumper. Physics and measurement of aero-optical effects: Past and present. *Annu. Rev. Fluid Mech.*, 49, 2017.

[32] Megan R Ebers, Jan P Williams, Katherine M Steele, and J Nathan Kutz. Leveraging arbitrary mobile sensor trajectories with shallow recurrent decoder networks for full-state reconstruction. *arXiv preprint arXiv:2307.11793*, 2023.

[33] N. B. Erichson, S. L. Brunton, and J. N. Kutz. Compressed dynamic mode decomposition for real-time object detection. *Journal of Real-Time Image Processing*, 2016.

[34] N. Benjamin Erichson, Lionel Mathelin, Zhewei Yao, Steven L. Brunton, Michael W. Mahoney, and J. Nathan Kutz. Shallow Neural Networks for Fluid Flow Reconstruction with Limited Sensors. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476, 2020.

[35] Uriel Frisch. *Turbulence: the legacy of AN Kolmogorov*. Cambridge university press, 1995.

[36] Koji Fukagata. Reduced order modeling of fluid flows using convolutional neural networks. *Journal of Fluid Science and Technology*, 18(1):JFST0002–JFST0002, 2023.

[37] Kai Fukami, Koji Fukagata, and Kunihiko Taira. Super-resolution reconstruction of turbulent flows with machine learning. *Journal of Fluid Mechanics*, 870, 2019.

[38] Kai Fukami, Koji Fukagata, and Kunihiko Taira. Assessment of supervised machine learning methods for fluid flows. *Theoretical and Computational Fluid Dynamics*, 34:497–519, 2020.

[39] Kai Fukami, Koji Fukagata, and Kunihiko Taira. Machine-learning-based spatio-temporal super resolution reconstruction of turbulent flows. *Journal of Fluid Mechanics*, 909, 2021.

[40] M. Gavish and D. L. Donoho. The optimal hard threshold for singular values is $4/\sqrt{3}$. *IEEE Transactions on Information Theory*, 60, 2014.

[41] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. *Neural computation*, 12(10):2451–2471, 2000.

[42] Gene H Golub and Christian Reinsch. Singular value decomposition and least squares solutions. In *Handbook for Automatic Computation: Volume II: Linear Algebra*, pages 134–151. Springer, 1971.

[43] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[44] Joseph W. Goodman. *Introduction to Fourier Optics*. W. H. Freeman, 2017.

[45] D. J. Goorskey, R. Drye, and M. R. Whiteley. Dynamic modal analysis of transonic Airborne Aero-Optics Laboratory conformal window flight-test aero-optics. *Optical Engineering*, 52, 2013.

[46] David J. Goorskey, Jason Schmidt, and Matthew R. Whiteley. Efficacy of predictive wavefront control for compensating aero-optical aberrations. *Optical Engineering*, 52, March 2013.

[47] Stanislav Gordeyev and Eric Jumper. Fluid dynamics and aero-optics of turrets. *Progress in Aerospace Sciences*, 46, 2010.

[48] Stanislav Gordeyev, Eric J Jumper, and Matthew R Whiteley. *Aero-Optical Effects: Physics, Analysis and Mitigation*. John Wiley & Sons, 2023.

[49] Maziar S Hemati, Clarence W Rowley, Eric A. Deem, and Louis N. Cattafesta. Debiasing the dynamic mode decomposition for applied Koopman spectral analysis. *Theoretical and Computational Fluid Dynamics*, 31, 2017.

[50] Sepp Hochreiter. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02):107–116, 1998.

[51] Sepp Hochreiter and Jürgen Schmidhuber. Lstm can solve hard long time lag problems. *Advances in neural information processing systems*, 9, 1996.

[52] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, 1997.

[53] Larry J Hornbeck. Deformable-mirror spatial light modulators. In *Spatial Light Modulators and Applications III*, volume 1150, pages 86–103. Spie, 1990.

[54] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2, 1989.

[55] John David Jackson. *Classical Electrodynamics*. American Association of Physics Teachers, 1999.

[56] Mihailo R Jovanović, Peter J Schmid, and Joseph W Nichols. Sparsity-promoting dynamic mode decomposition. *Physics of Fluids*, 26, 2014.

[57] Eric J. Jumper and Stanislav Gordeyev. Physics and Measurement of Aero-Optical Effects: Past and Present. *Annual Review of Fluid Mechanics*, 49, 2017.

[58] Eric J. Jumper, Stanislav Gordeyev, David Cavalieri, Phil Rollins, Matthew Whiteley, and Matthew Krizo. Airborne Aero-Optics Laboratory - Transonic (AAOL-T). In *53rd AIAA Aerospace Sciences Meeting*. American Institute of Aeronautics and Astronautics, 2015.

[59] Eric J Jumper and Ronald J Hugo. Quantification of aero-optical phase distortion using the small-aperture beam technique. *AIAA journal*, 33(11):2151–2157, 1995.

[60] Matthew Kalensky, Matthew R Kemnetz, and Mark F Spencer. Effects of shock waves on shack–hartmann wavefront sensor data. *AIAA Journal*, 61(6):2356–2368, 2023.

[61] Daniel C Kammer. Effects of noise on sensor placement for on-orbit modal identification of large space structures. *Journal of Dynamic Systems, Measurement and Control*, 114(3):436–443, 1992.

[62] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv:1412.6980 [cs]*, Dec 2014. arXiv: 1412.6980.

[63] B. O. Koopman. Hamiltonian systems and transformation in Hilbert space. *Proceedings of the National Academy of Sciences*, 17, 1931.

[64] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

[65] J. Nathan Kutz, Steven L. Brunton, Bingni W. Brunton, and Joshua L. Proctor.

Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems. Society for Industrial and Applied Mathematics, November 2016.

[66] Nathan Kutz, Diya Sashidhar, Shervin Sahba, Steven L. Brunton, Austin McDaniel, and Christopher Wilcox. Physics-informed machine-learning for modeling aero-optics. In *Applied Optical Metrology IV*. SPIE, August 2021.

[67] Dana Lahat, Tülay Adali, and Christian Jutten. Multimodal data fusion: an overview of methods, challenges, and prospects. *Proceedings of the IEEE*, 103(9):1449–1477, 2015.

[68] RG Lane and M Tallon. Wave-front reconstruction using a shack–hartmann sensor. *Applied optics*, 31(32):6902–6908, 1992.

[69] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521, 2015.

[70] Zijia Li, Yuanxiang Li, Boyang Xing, Bin Zhang, Hongya Tuo, and Hong Liu. OPD analysis and prediction in aero-optics based on dictionary learning. *Aerospace Systems*, 2, 2019.

[71] Hans Wolfgang Liepmann and Anatol Roshko. *Elements of gasdynamics*. Courier Corporation, 2001.

[72] Jian-Fu Lin, You-Lin Xu, and Siu-Seong Law. Structural damage detection-oriented multi-type sensor placement with multi-objective optimization. *Journal of Sound and Vibration*, 422:568–589, 2018.

[73] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv:1711.05101*, Jan 2019. arXiv:1711.05101 [cs, math].

[74] Michael M Malley, George W Sutton, and Nancy Kincheloe. Beam-jitter measurements of turbulent aero-optical path differences. *Applied Optics*, 31(22):4440–4443, 1992.

[75] Krithika Manohar, Bingni W Brunton, J Nathan Kutz, and Steven L Brunton. Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns. *IEEE Control Systems Magazine*, 38(3):63–86, 2018.

[76] Andreĭ Sergeevich Monin and Akiva M Yaglom. *Statistical fluid mechanics, volume II: mechanics of turbulence*, volume 2. Courier Corporation, 2013.

[77] Robert J Noll. Zernike polynomials and atmospheric turbulence. *JOsA*, 66(3):207–211, 1976.

[78] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

[79] Ben C Platt and Roland Shack. History and principles of shack-hartmann wavefront sensing, 2001.

[80] Clélia Robert, Jean-Marc Conan, Vincent Michau, Thierry Fusco, and Nicolas Vedrenne. Scintillation and phase anisoplanatism in Shack–Hartmann wavefront sensing. *Journal of the Optical Society of America A*, 2006.

[81] Raul Rojas and Raúl Rojas. The backpropagation algorithm. *Neural networks: a systematic introduction*, pages 149–182, 1996.

[82] Shervin Sahba, Diya Sashidhar, Christopher C. Wilcox, Austin McDaniel, Steven L. Brunton, and J. Nathan Kutz. Dynamic mode decomposition for aero-optic wavefront characterization. *Optical Engineering*, 61, 2022.

[83] Shervin Sahba, Christopher C Wilcox, Austin McDaniel, Benjamin D Shaffer, Steven L Brunton, and J Nathan Kutz. Wavefront sensor fusion via shallow decoder neural networks for aero-optical predictive control. In *Interferometry XXI*, volume 12223, pages 11–17. SPIE, 2022.

[84] Bahaa EA Saleh and Malvin Carl Teich. *Fundamentals of photonics*. John Wiley & Sons, 2019.

[85] Isabel Scherl, Benjamin Strom, Jessica K Shang, Owen Williams, Brian L Polagye, and Steven L Brunton. Robust principal component analysis for particle image velocimetry. *Physical Review Fluids*, 5, 2020.

[86] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656, August 2010.

[87] P. J. Schmid and J. Sesterhenn. Dynamic mode decomposition of numerical and experimental data. In *61st Annual Meeting of the APS Division of Fluid Dynamics*. American Physical Society, November 2008.

[88] Peter J Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics*, 656, 2010.

[89] Roland V Shack. Production and use of a lecticular Hartmann screen. *J. Opt. Soc. Am.*, 61, 1971.

[90] Benjamin D Shaffer, Jeremy R Vorenberg, Christopher C Wilcox, and Austin J McDaniel. Neural network forecasting of transonic turbulent flow for adaptive optics control. In *Unconventional Imaging and Adaptive Optics 2022*, volume 12239, pages 134–141. SPIE, 2022.

[91] Eli Shechtman, Yaron Caspi, and Michal Irani. Space-time super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):531–545, 2005.

[92] Rohan N Soman, Toula Onoufrioua, Marios A Kyriakidesb, Renos A Votsisc, and Christis Z Chrysostomou. Multi-type, multi-sensor placement optimization for structural health monitoring of long span bridges. *Smart Structures and Systems*, 14(1):55–70, 2014.

[93] Robin Swanson, Masen Lamb, Carlos M Correia, Suresh Sivanandam, and Kiriakos Kutulakos. Closed loop predictive control of adaptive optics systems with convolutional neural networks. *Monthly Notices of the Royal Astronomical Society*, 503(2):2944–2954, 2021.

[94] Naoya Takeishi, Yoshinobu Kawahara, and Takehisa Yairi. Subspace dynamic mode decomposition for stochastic Koopman analysis. *Physical Review E*, 96, September 2017.

[95] William J. Tango. The circle polynomials of Zernike and their application in optics. *Applied Physics*, 13, August 1977.

[96] Lloyd N Trefethan and David Bau. *Numerical Linear Algebra*. Society for Industrical and Applied Mathematics, 1997.

[97] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz. On dynamic mode decomposition: theory and applications. *Journal of Computational Dynamics*, 1, 2014.

[98] Jonathan H. Tu, Clarence W. Rowley, Dirk M. Luchtenburg, Steven L. Brunton, and J. Nathan Kutz. On Dynamic Mode Decomposition: Theory and Applications. *Journal of Computational Dynamics*, 1(2):391–421, 2014.

[99] Robert K Tyson and Benjamin West Frazier. *Principles of adaptive optics*. CRC press, 2022.

[100] JY Wang and D Es Silva. Wave-front interpretation with zernike polynomials. *Applied optics*, 19(9):1510–1518, 1980.

[101] Meng Wang, Ali Mani, and Stanislav Gordeyev. Physics and Computation of Aero-Optics. *Annual Review of Fluid Mechanics*, 44, January 2012.

[102] W. X. Wang, R. Yang, Y. C. Lai, V. Kovanis, and C. Grebogi. Predicting catastrophes in nonlinear dynamical systems by compressive sensing. *Physical Review Letters*, 106, 2011.

[103] Zhu Wang, Imran Akhtar, Jeff Borggaard, and Traian Iliescu. Proper orthogonal decomposition closure models for turbulent flows: a numerical comparison. *Computer Methods in Applied Mechanics and Engineering*, 237, 2012.

[104] C. C. Wilcox, K. P. Healey, A. L. Tuffli, and B. D. Agena. Air Force Research Laboratory Aero-Effects Laboratory system status and capabilities. *Proc. SPIE 11490, Interferometry XX*, 11490, 2020.

[105] Christopher C Wilcox, Keith P Healey, Brian D Agena, and Andrea L Tuffli. Air force research laboratory, aero-effects laboratory optical metrology system and performance. In *Interferometry XX*, volume 11490, pages 49–62. SPIE, 2020.

[106] Christopher C. Wilcox, Keith P. Healey, Brian D. Agena, and Andrea L. Tuffli. Air Force Research Laboratory, Aero-Effects Laboratory optical metrology system and performance. In Michael B. North Morris, Katherine Creath, and Rosario Porras-Aguilar, editors, *Interferometry XX*, volume 11490, page 114900A. International Society for Optics and Photonics, SPIE, 2020.

[107] Christopher C. Wilcox, Cameron Radosevich, Keith Healy, Andrea Tuffli, Brian Agena, Mark Spencer, and Donald J. Wittich. Digital holography wavefront sensing with a supersonic wind tunnel. In Antonio Fimia, Miroslav Hrabovský, and John T. Sheridan, editors, *Holography: Advances and Modern Trends VI*. SPIE, 2019.

[108] Jan P. Williams, Olivia Zahn, and J. Nathan Kutz. Sensing with shallow recurrent decoder networks, 2023.

[109] CD Zhang and YL Xu. Optimal multi-type sensor placement for response and excitation reconstruction. *Journal of Sound and Vibration*, 360:112–128, 2016.

[110] XH Zhang, YL Xu, and S Zhan. Optimal locations of a multi-type sensor system for structural health monitoring. In *Proceedings Of The 8th International Conference On Structural Dynamics*, 2011.

[111] Songye Zhu, Xiao-Hua Zhang, You-Lin Xu, and Sheng Zhan. Multi-type sensor placement for multi-scale response reconstruction. *Advances in Structural Engineering*, 16(10):1779–1797, 2013.