Wavefront Prediction of Convective Boundary Layers Using Multiple-Aperture Digital Holography

Nathaniel Brownell,[1] James Mitchell,[1] and Stanislav Gordeyev[1]

[1]Aerospace and Mechanical Engineering Department, University of Notre Dame, Notre Dame, Indiana 46556

ABSTRACT

A major limiting factor for adaptive-optic systems is the cumulative latency inherent in the control loop. One way to address the latency problem is to use multiple apertures, where upstream wavefront measurements can be used to predict the wavefronts over the main aperture. In this work, the multiple aperture prediction method is extended to a dual-wavelength digital holography sensing system, and machine learning is used to predict downstream turbulent boundary layer wavefronts from upstream measurements. Wavefronts are decomposed using Joint Proper Orthogonal Decomposition (JPOD), and a ridge least squares predictor and a multilayer perceptron (MLP) are trained to predict downstream time coefficients from a window of upstream coefficients. Predictions are compared to a time shift convective baseline across three aperture separations (2.9δ, 3.6δ, and 4.3δ), which target the regime where Taylor's frozen flow hypothesis breaks down, and at three Mach numbers (0.58, 0.70, and 0.80). Both predictors substantially outperform the time shift baseline across all nine conditions, recovering 64% of the downstream wavefront energy at the easiest condition compared to 30% for time shift, and 34% compared to -21% at the most extreme condition. They also converge to nearly identical predictions, indicating that the flow's upstream to downstream dynamics are well described by a linear mapping. Analysis of the ridge weight matrix shows that these predictions are largely enabled by the aperture windowing effect, where convecting structures project onto different modes at different lags as they move through the downstream aperture, allowing the predictor to distinguish coherent convecting structures from unpredictable turbulent fluctuations. For a real time predictive controller, retaining approximately 60 modes captures nearly all the predictable energy. The prediction limits that remain appear to stem from the path-integrated nature of the wavefront measurement itself, in which contributions from the opposite wall boundary layer are integrated into the same signal and cannot be resolved.

Keywords: adaptive optics, wavefront prediction, digital holography, machine learning, multilayer perceptron, ridge regression, turbulent boundary layer, joint proper orthogonal decomposition

1. INTRODUCTION

In the field of aero-optics, we seek to understand the optical aberrations caused when light travels through density fluctuating regions from turbulent boundary layers and shear layers. Critical to this understanding is the relationship between density and index of refraction. The Gladstone-Dale equation quantifies this relationship,

$$n(x, y, z, t) = 1 + K_{GD}\,\rho(x, y, z, t), \tag{1}$$

where n is the media's index of refraction, K_{GD} is the Gladstone-Dale constant, and ρ is the fluid density.[1] Within a turbulent boundary layer, local density fluctuations from hairpin vortices lead to local changes in index of refraction, causing the wavefront to accumulate local phase lags. According to Huygen's principle,[2] light propagates normal to the wavefront, which means that the rays from the distorted wavefront deviate from the original optical axis, which in the far field corresponds to less energy on target. To quantify the resulting

Further author information: (Send correspondence to N.B.)
N.B.: E-mail: nbrownel@nd.edu

Unconventional Imaging, Sensing, and Adaptive Optics 2026, edited by Jean Dolne, Santasri Bose-Pillai, Matthew Kalensky, Proc. of SPIE Vol. 14229, 1422912 · 2026 Published by SPIE
0277-786X · doi: 10.1117/12.3113690

distortion, we first compute the optical path length (OPL), defined as the integral of the index of refraction along the beam propagation direction z,

$$OPL(x, y, t) = \int_0^{z_1} n(x, y, z, t)\, dz. \tag{2}$$

The optical path difference (OPD) is then obtained by subtracting the spatial mean,

$$OPD(x, y, t) = OPL(x, y, t) - \langle OPL(x, y, t) \rangle, \tag{3}$$

where the angular brackets denote the spatial mean over the aperture. Combining Eqs. (1–3), the OPD reduces to

$$OPD(x, y, t) = K_{GD} \int_0^{z_1} \rho'(x, y, z, t)\, dz, \tag{4}$$

where ρ' is the fluctuating density.[1] This shows that aero-optical aberrations are a spatial integral of turbulent density fluctuations in the boundary layer.[1,3] These aberrations degrade far field beam intensity, which becomes costly in aero-optical applications such as directed energy and optical communications. Thus, understanding these effects and ways to mitigate them has become the central focus of aero-optical studies.

One way to mitigate aero-optical effects is through adaptive optics (AO) systems. An AO systems works by using a wavefront sensor to quantify the optical aberrations, then applying the conjugate of the aberrations to a deformable mirror, so that the correct signal beam can be sent through the flow. While such systems are well suited for low speed atmospheric flows, aero-optical environments are characterized by rapidly evolving turbulent structures with frequencies on the order of 1 kHz. Thus, inherent in aero-optical AO systems is the latency problem, where the flow has changed significantly by the time the wavefront correction is applied, severely degrading AO controller performance.[4] One approach to addressing this problem is the use of multiple apertures, where an upstream aperture measures the wavefront before it convects over the main lasing aperture.[5] Under Taylor's frozen flow hypothesis, turbulent structures are assumed to convect downstream without evolving, such that once convective speed is found, a simple time shift can be applied to predict downstream flow. However, prior dual aperture Shack-Hartmann studies show that small-scale boundary-layer structures lose cross aperture correlation rapidly with increasing separation (or equivalently increased prediction time), making convection only models insufficient beyond 3δ,[5,6] where δ is boundary layer thickness. Previous work has also shown that decomposing wavefronts into a set of basis vectors using Proper Orthogonal Decomposition (POD), and then applying a linear convolution to predict time coefficients can drastically improve prediction capabilities for all aperture separations.[6]

In this work, we extend the multiple-aperture prediction method to a dual-wavelength digital holography sensing system developed in previous work by Mitchell et al.,[7] and apply machine learning to the prediction problem. Wavefronts are decomposed using Joint Proper Orthogonal Decomposition (JPOD), and both a least squares predictor and a multilayer perceptron (MLP) are trained to predict downstream time coefficients from a window of upstream coefficients. By comparing the two predictors across a range of aperture separations and Mach numbers, we test whether nonlinear modeling improves the prediction. Our separations of 2.9δ, 3.6δ, and 4.3δ specifically target the regime where convection only models are known to break down. Finally, we discuss limits imposed by our sensing system.

2. EXPERIMENTAL SETUP

All measurements were conducted in the Trisonic Wind Tunnel (TWT) at the Hessert Laboratory for Aerospace Research at the University of Notre Dame. The TWT draws air through a 150:1 contraction ratio inlet into a rectangular test section of 10 cm × 9.9 cm that spans 160 cm from the inlet to the diffuser. A pair of optical access windows span the tunnel walls from 130 cm to 150 cm downstream of the inlet, providing optical access for the holographic measurements. A schematic of the TWT facility is shown in Fig. 1. The freestream Mach number was determined from stagnation and static pressure measurements taken by a Pitot probe mounted just ahead of the optical window. The isentropic relation,

$$\frac{p}{p_t} = \left(1 + \frac{\gamma - 1}{2} M^2\right)^{-\frac{\gamma}{\gamma - 1}}, \tag{5}$$

relates the freestream static pressure p to the Pitot stagnation pressure p_t, where M is the freestream Mach number and $\gamma = 1.4$ is the specific heat ratio of air. The freestream velocity was then obtained from $U = M \cdot a$, where the speed of sound is given by the adiabatic relation,

$$a = \sqrt{\frac{\gamma R T_0}{1 + \frac{\gamma - 1}{2} M^2}}, \tag{6}$$

with $R = 287\ J/(kg \cdot K)$ the specific gas constant for air and $T_0 = 295$ K the measured stagnation temperature, at freestream Mach numbers of $M = 0.58$, 0.70, and 0.80. The flow of interest is a turbulent boundary layer developing along the tunnel wall, with a measured thickness of $\delta = 18.5 \pm 1.5$ mm. The Reynolds number based on momentum thickness, Re_θ, was computed using Sutherland's law for dynamic viscosity,[8] and ranges from approximately 21,600 to 23,800 across the three Mach numbers, which is well into the fully turbulent boundary layer regime.[9]

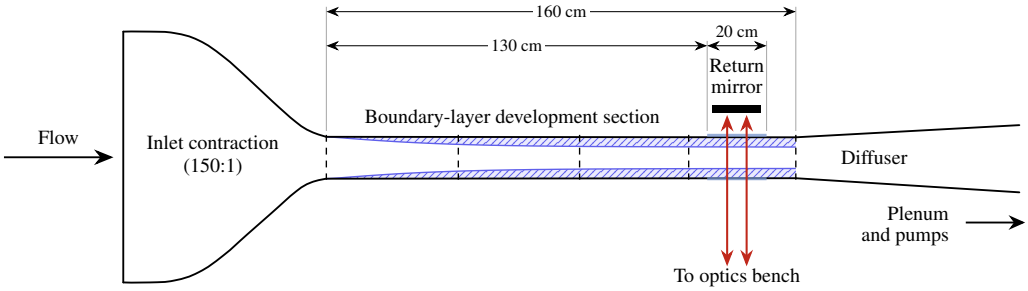

Figure 1. Schematic of the Notre Dame Tri-sonic Wind Tunnel facility.

The wavefront sensing approach used is essentially the same off-axis, dual-wavelength digital holography system developed by Mitchell et al.[7] Each of the two laser is split into a tilted reference beam and a signal beam. The signal beam double-passes through the wind tunnel to accumulate the boundary-layer wavefront distortion, then interferes with the reference beam on the camera sensor to produce a hologram, which is later processed to recover the wavefront as described in Section 3. The sensor implemented for this experiment uses two slight modifications. First, the signal beams are expanded to 2 inches in diameter before entering the test section to improve spatial resolution, and second, the re-imaging telescopes are redesigned to compress the beam to 0.5 inches at the camera sensor, increasing the maximum frame rate and thereby the temporal resolution. A schematic and photograph of the sensor are shown in Fig. 2 and Fig. 3, respectively. A full explanation of the setup is given in Mitchell et al.[7]

The lasers chosen for this optical setup were the Cobolt 05-01 Series at 532 nm and 640 nm, both with coherence lengths exceeding 100 m and output powers of 600 mW. The two wavelengths were selected for their ability to be cleanly separated by dichroic filters and, once combined on the camera sensor, to produce non-interfering fringe patterns that can be isolated in the Fourier plane.

In order to vary the separation between the two apertures, the green (532 nm) beam expander is mounted on a translation rail, allowing its position to be adjusted without disturbing the rest of the optical path. Three center to center separations were tested in this study, corresponding to 2.9δ, 3.6δ, and 4.3δ.

The resulting holograms were recorded on a Phantom v2512 high-speed camera with a pixel width of 28 μm, frame rate of $f_s = 83$ kHz, resolution of 512 $\times$ 464 pixels, and exposure time of 1.1 μs. For each case, 20,000 frames were collected.

3. WAVEFRONT RECONSTRUCTION

Raw holograms recorded by the camera are processed to reconstruct the OPD wavefront for each frame and at each aperture location. When the signal and reference beams interfere on the camera sensor, the total intensity recorded is

$$I = |U_S + U_R|^2 = |U_S|^2 + |U_R|^2 + U_S U_R^* + U_S^* U_R, \tag{7}$$

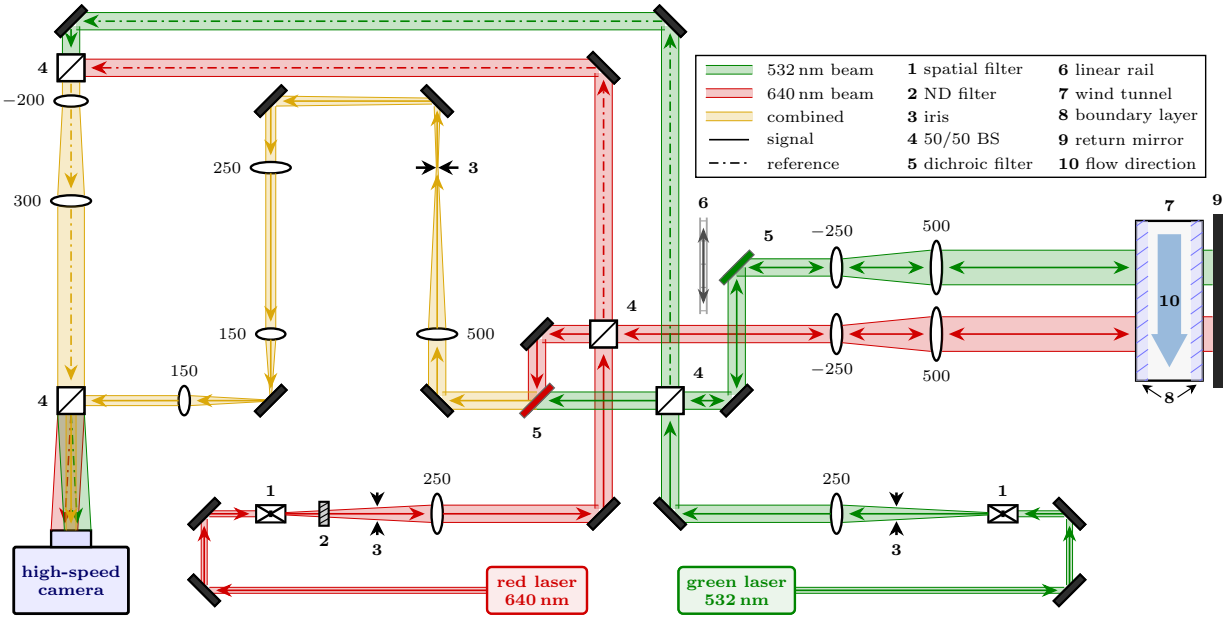

Figure 2. Schematic of the dual-wavelength digital holography wavefront sensor.

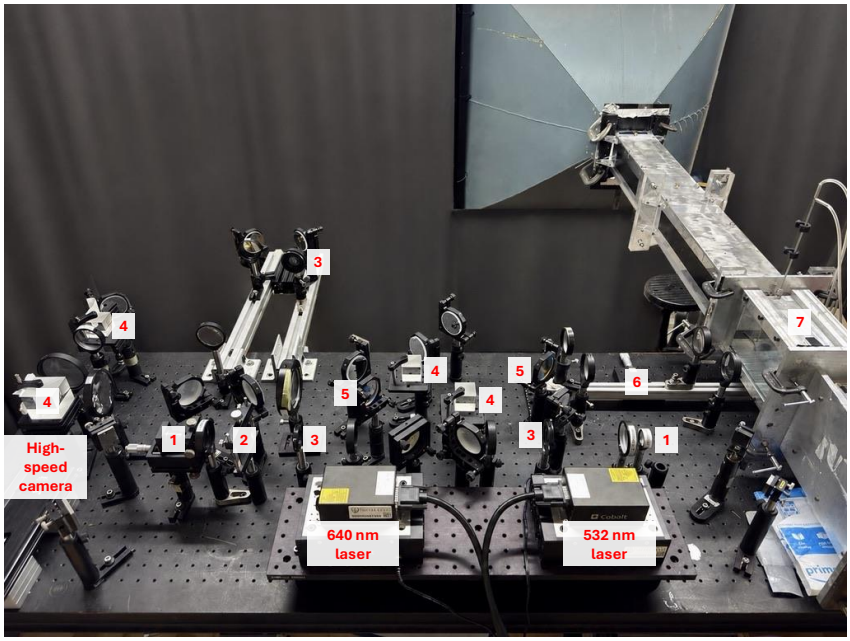

Figure 3. Photograph of the dual-wavelength digital holography sensor on the optical table, with numbered components matching the legend in Fig. 2.

where U_S and U_R are the complex optical fields of the signal and reference beams, and the asterisk denotes the complex conjugate terms.[7,10] The terms $|U_S|^2$ and $|U_R|^2$ are real, zero-order terms that do not carry phase information and are therefore ignored. The cross-correlation terms $U_S U_R^*$ and $U_S^* U_R$ each contain phase information, but are complex conjugates of one another, so only one is needed to recover the wavefront. The

signal and reference complex fields can be written as

$$U_S(x,y) = A_S \exp[i\psi_S(x,y)]\,, \tag{8}$$

$$U_R^*(x,y) = A_R \exp[-i\psi_R(x,y)] \exp\left[-i\frac{2\pi}{\lambda z}(x_R x + y_R y)\right], \tag{9}$$

where A_S and A_R are the beam amplitudes, ψ_S and ψ_R are the signal and reference phases, λ is the laser wavelength, z is the propagation distance, and x_R, y_R describe the tip and tilt of the reference beam.[10]

To isolate $U_S U_R^*$, a two-dimensional spatial fast Fourier transform (FFT2) is applied to the hologram. Because the reference beam is tilted relative to the signal beam (off-axis setup), the interference fringes on the sensor carry an additional spatial frequency proportional to the tilt angle, which drives the cross-correlation term $U_S U_R^*$ away from the central zero-order term in Fourier space. Thus, a circular aperture mask is applied to isolate this peak and all other content is zeroed out. The peak is then shifted to the center of the Fourier plane which removes the reference beam frequency $\exp[-i\frac{2\pi}{\lambda z}(x_R x + y_R y)]$. An inverse FFT2 is performed, yielding the complex optical field

$$\mathcal{F}^{-1}\{U_S U_R^*\} = A_S A_R \exp[i(\psi_S - \psi_R)]\,, \tag{10}$$

whose argument is the wrapped phase $\psi_S - \psi_R$, which is confined between $-\pi$ and π. An example Fourier plane showing both wavelength peaks is presented in Fig. 4.

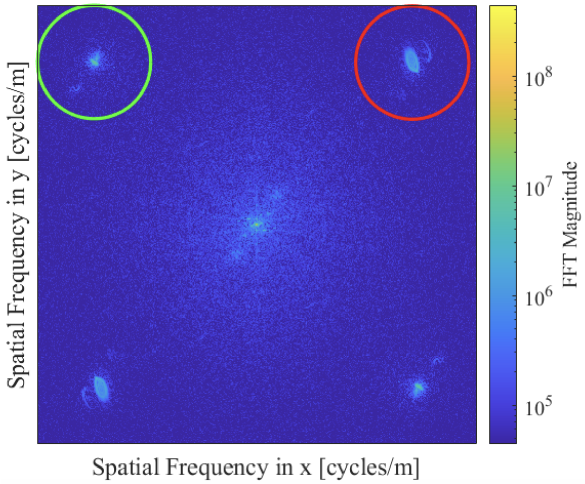

Figure 4. Representative Fourier plane of a raw dual-wavelength hologram, showing the central DC peak and the off-axis cross-correlation peaks for the 532 nm (green) and 640 nm (red) signal beams.

Phase unwrapping is performed using the least-squares unwrapping method of Ghiglia and Romero.[11] The method finds the unwrapped phase field ϕ that minimizes the total squared error between its gradient and the measured wrapped phase gradient

$$\mathcal{L} = \|\nabla\phi - \mathbf{g}\|^2, \tag{11}$$

where $\mathbf{g}$ is the wrapped gradient of ψ in both spatial directions. Setting $\partial\mathcal{L}/\partial\phi = 0$ yields the 2D Poisson equation,

$$\nabla^2\phi = \nabla \cdot \mathbf{g}, \tag{12}$$

which is solved using the preconditioned conjugate gradient method to find ϕ.[11] The unwrapped phase is then converted to OPD via

$$OPD(x,y,t) = \frac{\lambda}{2\pi}\,\phi(x,y,t). \tag{13}$$

A least-squares plane fit is applied to each reconstructed wavefront to remove tip, tilt, and piston, consistent with the OPD definition in Eq. (3). Static aberrations are removed by subtracting the time-averaged no-flow wavefront, computed from frames acquired when the tunnel was off. Finally, the result is divided by $2\sqrt{2}$, where

the factor of 2 corrects for the the beam passing twice through the tunnel, and the factor of $\sqrt{2}$ accounts for the statistically independent boundary layer on the opposite tunnel wall.[1,7] The current configuration means that the measured signal contains contributions from both tunnel walls that cannot be individually resolved, imposing a noise floor on the wavefront measurements, which is discussed in Section 6.

4. PREDICTION METHODS

4.1 POD Analysis

Predicting the full downstream wavefront directly in pixel space is impractical, since each OPD field contains N_{pix} degrees of freedom, which becomes computationally expensive for prediction methods. Additionally, turbulent structures that produce aero-optical distortions are spatially coherent, meaning that pixels within the same structure fluctuate together.[5] Therefore, we apply Proper Orthogonal Decomposition (POD) to decompose the wavefronts into a set of orthogonal spatial modes, reducing the dimensionality of the prediction problem and ordering structures by variance.[12]

For a given flow condition, wavefronts are stacked into a single snapshot matrix $\mathbf{X} \in \mathbb{R}^{N_{\text{pix}} \times N_s}$, where $N_{\text{pix}} = N_x \times N_y$ is the number of pixels and N_s is the number of training frames, which in our case was 16000. The spatial correlation matrix $\mathbf{R}$

$$\mathbf{R} = \frac{1}{N_s}\mathbf{X}\mathbf{X}^T, \tag{14}$$

encodes how much any two pixels fluctuate together over time. POD finds the eigenvectors of $\mathbf{R}$, or equivalently the spatial modes $\boldsymbol{\phi}_n$, and the corresponding eigenvalues λ_n, which represent the variance of each mode. The modes are computed efficiently with Singular Value Decomposition (SVD) of $\mathbf{X}$,

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T, \tag{15}$$

where the columns of $\mathbf{U}$ are the spatial modes, $\boldsymbol{\Sigma}$ is diagonal with values of s_n, and $\lambda_n = s_n^2/N_s$. $\mathbf{V}^T$ is a matrix whose rows encode the temporal behavior of each mode, and thus the temporal coefficient matrix is

$$\mathbf{A} = \boldsymbol{\Sigma}\mathbf{V}^T = \mathbf{U}^T\mathbf{X}, \tag{16}$$

where row n of $\mathbf{A}$ gives the time series $a_n(t)$, or the strength of each mode n at a given timestep t. Each wavefront is then written as[13]

$$W(x, y, t) = \sum_n a_n(t)\, \phi_n(x, y). \tag{17}$$

Under Taylor's frozen-flow hypothesis, the upstream and downstream OPD fields share similar spatial structures and can be defined using the same spatial modes.[5] To project onto a joint basis, a set of shared modes is constructed from both apertures using the Joint POD (JPOD) technique,[14] where the joint correlation matrix is the average of the upstream and downstream correlation matrices computed using Eq. (14),

$$\mathbf{R}_{\text{joint}} = \frac{1}{2}\left(\mathbf{R}_{\text{up}} + \mathbf{R}_{\text{dn}}\right). \tag{18}$$

The shared JPOD spatial modes $\boldsymbol{\phi}_n$ are the eigenvectors of $\mathbf{R}_{\text{joint}}$, computed via SVD following the same procedure as Eq. (15). Then, the upstream and downstream temporal coefficients are obtained by projecting each aperture's snapshot matrix onto the shared modes, as in Eq. (16), so that each aperture is decomposed as

$$W^{\text{UP}}(x, y, t) = \sum_n a_n^{\text{UP}}(t)\, \phi_n(x, y), \qquad W^{\text{DN}}(x, y, t) = \sum_n a_n^{\text{DN}}(t)\, \phi_n(x, y), \tag{19}$$

and the prediction task reduces to learning the mapping between time coefficients.

The first 30 JPOD spatial modes for the 2.9δ, $M = 0.58$ case are shown in Fig. 5. The low-order modes are dominated by large-scale structures that capture the bulk of the wavefront energy, while the higher-order modes resolve smaller-scale structures that contain less energy. This is reflected in the decay of the normalized eigenvalues shown in Fig. 6, where about 90% of the cumulative energy is contained within the first 100 modes.

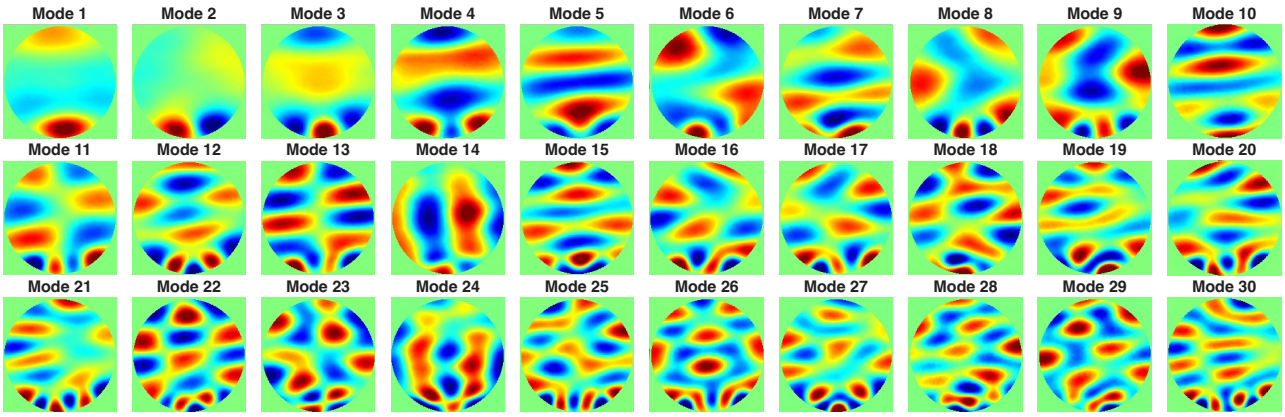

Figure 5. First 30 JPOD spatial modes (2.9δ, $M = 0.58$) ordered by decreasing eigenvalue λ_n, where flow is from left to right.

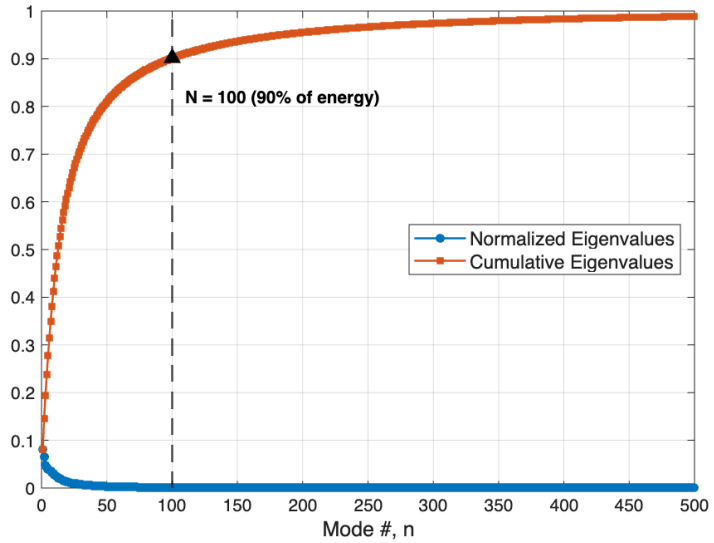

Figure 6. Normalized eigenvalues $\lambda_n / \sum_{n=1}^{\infty} \lambda_n$ and cumulative energy as a function of mode number (2.9δ, $M = 0.58$).

4.2 Error Definitions

To quantify prediction accuracy for each mode, we adopt the framework of Kemnetz et al.[6] Let $\tilde{a}_n^{\mathrm{DN}}(t)$ denote the predicted downstream coefficient for mode n, where the definition of $\tilde{a}_n^{\mathrm{DN}}$ depends on the predictor. The per-mode relative error is defined as[6]

$$\epsilon(n) = \frac{\left\langle \left[a_n^{\mathrm{DN}}(t) - \tilde{a}_n^{\mathrm{DN}}(t) \right]^2 \right\rangle}{\lambda_n}, \tag{20}$$

where angle brackets denote a time average over the test set. Normalizing by λ_n makes the per-mode errors directly comparable, with $\epsilon = 0$ indicating a perfect prediction and $\epsilon = 1$ corresponding to a prediction no better than predicting zero.

The global error across the first N modes is then defined as[6]

$$\Sigma(N) = \frac{\sum_{n=1}^{N} \epsilon_n \lambda_n + \sum_{n=N+1}^{\infty} \lambda_n}{\sum_{n=1}^{\infty} \lambda_n}, \tag{21}$$

where the first term in the numerator is the unpredicted energy from the retained modes and the second term is the residual energy from the truncated modes. When $N \to \infty$, the truncation term vanishes and Eq. (21) reduces to the traditional normalized mean-squared wavefront error,

$$\Sigma = \frac{\left\langle \left[W^{\mathrm{DN}}(x,y,t) - \tilde{W}^{\mathrm{DN}}(x,y,t) \right]^2 \right\rangle}{\left(W_{\mathrm{rms}}^{\mathrm{DN}}\right)^2}, \tag{22}$$

where $\tilde{W}^{\mathrm{DN}}$ is the predicted downstream wavefront reconstructed from the first N modes, $W_{\mathrm{rms}}^{\mathrm{DN}}$ is the root mean square of the downstream wavefront, and the angle brackets denote a spatial and temporal average over the test set. In practice, Σ is computed directly in pixel space via Eq. (22), which implicitly accounts for the truncation error.

4.3 TIME-SHIFT CONVECTIVE BASELINE

According to Taylor's frozen-flow hypothesis, turbulent structures convect downstream without evolving, giving the time-shift prediction for mode n,

$$\tilde{a}_n^{\mathrm{DN}}(t) = a_n^{\mathrm{UP}}(t - \tau), \tag{23}$$

where τ is the convective time delay. Sweeping τ and converting to normalized convective speed $U_c/U_\infty = \Delta f_s / \tau U_\infty$, where $f_s = 83$ kHz and Δ is the aperture spacing, yields the global error curve $\Sigma(U_c)$ defined in Eq. (22), whose interpolated minimum gives the optimal convective speed.

As shown in Fig. 7(a), including 100 modes vs 20 raises the global error, indicating that higher modes have evolved significantly between apertures, and proving that a convective only prediction model is insufficient especially for higher order modes. Fig. 8(b) confirms this, where the per-mode error ϵ_n found using Eq. (20) crosses 1 near $n = 20$, above which the time-shift is no better than predicting zero. Fig. 8(a) shows that convective speed is generally constant across separations and modes below 20, then drifts as noise dominates at higher modes, consistent with findings from previous experiments.[6]

As shown in Fig. 7(b), the time shift baseline predicts only 30%, 15%, and 5% of the downstream wavefront energy at $M = 0.58$ for 2.9δ, 3.6δ, and 4.3δ respectively, with worse performance at higher Mach numbers. A clear minimum near $U_c/U_\infty \approx 0.91$ is observed in each case, consistent with the known convective speed of large-scale boundary layer structures farther from the wall.[1,6] This confirms that time shifting is sufficient to extract the convective speed for these conditions. A summary of convective speeds found for each flow condition is shown in Table 1.

Table 1. Optimal normalized convective speed U_c/U_∞ for all nine flow conditions.

	$M = 0.58$	$M = 0.70$	$M = 0.80$
2.9δ	0.93	0.93	0.93
3.6δ	0.91	0.91	0.91
4.3δ	0.90	0.90	0.90

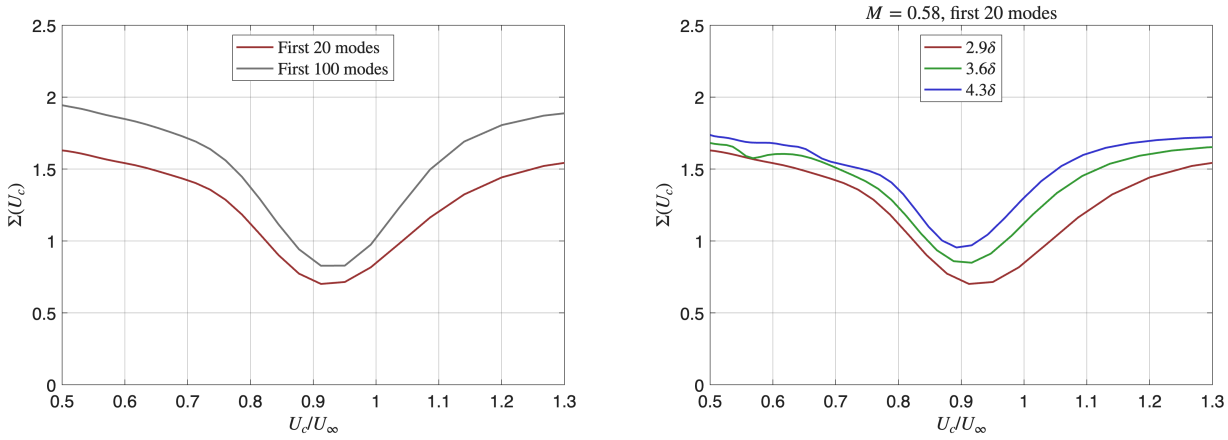

Figure 7. Global time-shift error $\Sigma(U_c)$ at $M = 0.58$. (a) 20 vs. 100 JPOD modes at 2.9δ. (b) First 20 modes for all three aperture separations.

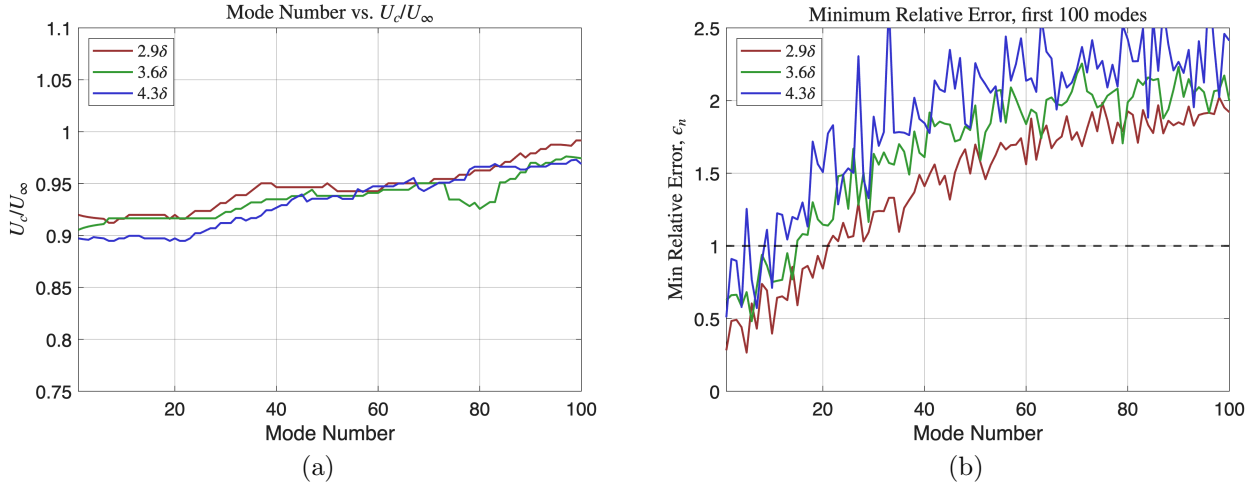

Figure 8. Per-mode analysis at $M = 0.58$. (a) Convective speed U_c/U_∞ for each mode. (b) Minimum per-mode error ϵ_n, where the dashed line marks $\epsilon_n = 1$.

4.4 RIDGE LEAST SQUARES LINEAR PREDICTOR

The time-shift baseline predicts each downstream coefficient from a single upstream coefficient at one time delay. A natural extension of this is to predict each downstream coefficient from a window of upstream frames and all other modes, allowing the predictor to account for the temporal evolution of each mode and the diffusion of energy between modes that may be present. At each prediction timestep t, a window of $M = 30$ consecutive upstream frames centered on the convective delay τ is extracted and flattened into the row vector $\mathbf{x}(t) \in \mathbb{R}^{MN}$. The entries of $\mathbf{x}(t)$ are the upstream coefficients $a_n^{\mathrm{UP}}(t-\tau+k)$ for every mode n up to $N = 100$ and every frame offset $k \in [-M/2, M/2-1]$. Stacking these row vectors across all N_s training timesteps gives the input matrix $\mathbf{X} \in \mathbb{R}^{N_s \times MN}$, with corresponding downstream coefficient matrix $\mathbf{Y} \in \mathbb{R}^{N_s \times N}$. A weight matrix $\mathbf{W} \in \mathbb{R}^{MN \times N}$ is sought such that

$$\mathbf{Y} \approx \mathbf{X}\mathbf{W}, \tag{24}$$

so that the predicted downstream coefficients for all N modes at a given timestep t are

$$\tilde{\mathbf{a}}^{\mathrm{DN}}(t) = \mathbf{x}(t)\,\mathbf{W}, \tag{25}$$

where the nth entry of $\tilde{\mathbf{a}}^{\mathrm{DN}}(t)$ is the predicted coefficient $\tilde{a}_n^{\mathrm{DN}}(t)$ for mode n.

Because consecutive frames are strongly correlated in time, the columns of $\mathbf{X}$ carry nearly redundant information, and ordinary least squares overfits noise in the training data and yields non-optimal predictions. To address this, ridge regression[15] is used, which implements an ℓ_2 penalty on the weights to the squared training error,

$$\mathcal{L}_{\text{ridge}} = \|\mathbf{Y} - \mathbf{X}\mathbf{W}\|^2 + \lambda\|\mathbf{W}\|^2, \tag{26}$$

where λ controls how strongly large weights are penalized. Setting the gradient with respect to $\mathbf{W}$ to zero gives the solution

$$\mathbf{W} = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}\right)^{-1}\mathbf{X}^T\mathbf{Y}, \tag{27}$$

where $\mathbf{I}$ is the identity matrix. λ is chosen by sweeping over a range and selecting the value that minimizes the mean squared error (MSE) of the validation set. Of the 4,000 frames not used for training, 2,000 form this validation set and the remaining 2,000 form the test set on which the final prediction error is reported. This split is also used for the MLP.

All input and output coefficients are normalized by per-mode training standard deviation before solving Eq. (27), and predictions are de-normalized before wavefront reconstruction. The cross-aperture correlation tensor is defined as

$$C(m,n,k) = \frac{\mathbf{X}^T\mathbf{Y}}{N_s} = \frac{1}{N_s}\sum_t \hat{a}_m^{\text{UP}}(t-\tau+k)\,\hat{a}_n^{\text{DN}}(t), \tag{28}$$

where $\hat{a}_n(t) = a_n(t)/\sigma_n$ denotes the normalized coefficient, and the $k \neq 0$ entries encode the temporal history of each upstream mode across the window. Because the JPOD spatial modes are orthogonal, upstream mode coefficients are uncorrelated with one another at any given time instant, so for $M = 1$, $\mathbf{X}^T\mathbf{X} \approx N_s\mathbf{I}$ and the ridge solution reduces to

$$W(m,n) \approx \frac{C(m,n)}{1+\lambda/N_s}, \tag{29}$$

where $C(m,n) = C(m,n,k=0)$ is the single-frame cross-aperture correlation between upstream mode m and downstream mode n at the convective delay.

4.5 MULTILAYER PERCEPTRON NONLINEAR PREDICTOR

A multilayer perceptron (MLP) is a feedforward, fully connected neural network in which the input passes through one or more hidden layers of neurons before reaching an output layer. Each neuron computes a weighted sum of its inputs, adds a learned bias, and passes the result through a nonlinear activation function which allows the network to approximate nonlinear functions. Fully connected means each neuron receives an input from every neuron in the previous layer.

An MLP is trained on the same sliding window inputs as the ridge predictor to determine whether nonlinear mode interactions are present and lead to better predictions. The MLP takes the input vector $\mathbf{x}(t) \in \mathbb{R}^{MN}$ and produces the predicted downstream coefficient vector

$$\tilde{\mathbf{a}}^{\text{DN}}(t) = f(\mathbf{x}(t)), \tag{30}$$

where $f(\cdot)$ is obtained by the following forward pass through two hidden layers with ELU activations,

$$\begin{aligned}
\mathbf{z}_1 &= \text{ELU}(\mathbf{P}_1\mathbf{x}(t) + \mathbf{b}_1),\\
\mathbf{z}_2 &= \text{ELU}(\mathbf{P}_2\mathbf{z}_1 + \mathbf{b}_2),\\
\tilde{\mathbf{a}}^{\text{DN}}(t) &= \mathbf{P}_3\mathbf{z}_2 + \mathbf{b}_3.
\end{aligned} \tag{31}$$

where $\mathbf{z}_1 \in \mathbb{R}^{1024}$ and $\mathbf{z}_2 \in \mathbb{R}^{512}$ are the output vectors for each layer, and $\mathbf{P}_\ell$ and $\mathbf{b}_\ell$ are the learned weight matrix and bias vector at layer ℓ. Given the large number of network parameters ($N_{\text{p}} = 3{,}652{,}196$) relative to the available training data ($N_s = 16{,}000$), controlling overfitting is critical.[16] To prevent overfitting, three regularization strategies are applied: an ℓ_2 weight decay to move the network toward small weights (similar to ridge regression in Eq. (26)), dropout of random neurons to help generalize, and batch normalization.[16] All

input and output coefficients are standardized by per-mode training standard deviation, identical to the ridge normalization. The network is trained to minimize the MSE loss

$$\mathcal{L}_{\mathrm{MLP}} = \frac{1}{N}\sum_{n=1}^{N}\left\langle \left[\hat{a}_n^{\mathrm{DN}}(t) - \tilde{a}_n^{\mathrm{DN}}(t)\right]^2 \right\rangle, \tag{32}$$

where the angle brackets denote a time average over the training set. Parameters are updated using the Adam optimizer,[17] which continuously adjusts the learning rate for each parameter, taking larger steps for consistent gradients and smaller steps for noisy gradients. The hyperparameters used are summarized in Table 2. In order to verify that any improved prediction from the MLP is attributable to nonlinearity rather than regularization, a linear control MLP is trained across all nine cases with identical hyperparameters, only replacing the ELU activations with identity functions.

Table 2. MLP hyperparameter summary.

Hyperparameter	**Value**
Input size	3000
Hidden layer widths	[1024, 512]
Activation (nonlinear / linear control)	ELU / Identity
Batch normalization	Yes
Dropout rate	0.2
Learning rate	1×10^{-4}
Weight decay	1×10^{-3}
Batch size	100
Max epochs	500
Early stopping patience	60

5. RESULTS

Having established the time-shift baseline, this section first compares the prediction performance of ridge LS and MLP across the nine flow conditions, then isolates the source of any performance gap between them, and finally examines the structure of the ridge weight matrix to identify the mechanisms each predictor exploits.

5.1 PREDICTION PERFORMANCE OF RIDGE LS AND MLP

The minimum global error $\Sigma_{\min}$ for each flow condition and predictor is shown in Fig. 9, where $\Sigma_{\min}$ is computed using Eq. (22) and found within the first $N = 100$ modes. As can be seen, $\Sigma_{\min}$ increases with aperture separation, since at larger separations the convective time delay τ increases, allowing more time for boundary-layer structures to evolve and decorrelate before reaching the downstream aperture. $\Sigma_{\min}$ also increases with Mach number at each aperture separation, indicating a genuine increase in the decorrelation of boundary-layer structures during convection at higher Mach numbers. The window based predictors are shown to recover substantially more downstream wavefront energy than the time-shift baseline at every condition. At $M = 0.58$, for example, the time-shift baseline predicted only 30%, 15%, and 5% of the downstream wavefront energy for 2.9δ, 3.6δ, and 4.3δ, respectively (Fig. 7b), compared to 64%, 55% and 45% for the same conditions with window based predictors.

Fig. 10 shows a prediction example at 2.9δ and $M = 0.58$, comparing the true downstream wavefront to the ridge LS, MLP, and time-shift predictions, along with each predictor's residual. Ridge LS and MLP both capture the dominant large-scale structures well, while the time-shift prediction recovers only the broadest features.

In comparing the window based predictors, it is found that Ridge LS outperforms MLP at small separations, with the largest advantage of approximately 3% more predicted energy occurring at 2.9δ and $M = 0.58$. This gap narrows as separation and Mach number increase, reversing at 4.3δ and $M = 0.80$ where MLP predicts approximately 2% more energy than LS. This trend is visible at every Mach number, but is most clearly illustrated at $M = 0.80$, shown in Fig. 11(a), which gives the global error $\Sigma(N)$ computed using Eq. (22), as a function of retained mode count for all three aperture separations.

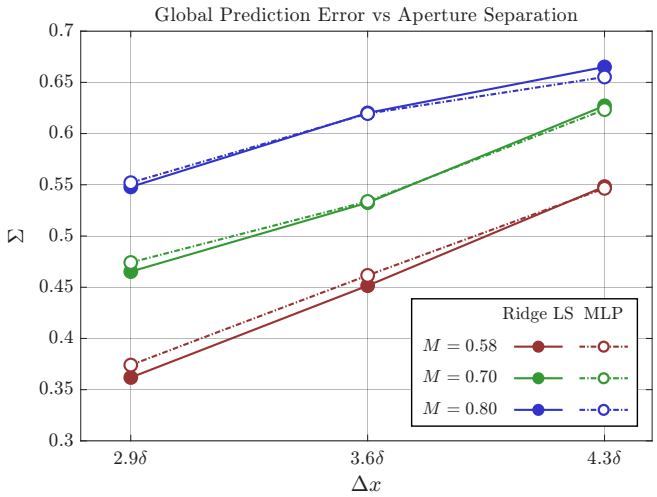

Figure 9. Minimum global prediction error $\Sigma_{\min}$ as a function of aperture separation for all three Mach numbers, comparing ridge LS (solid) and MLP (dash-dot).

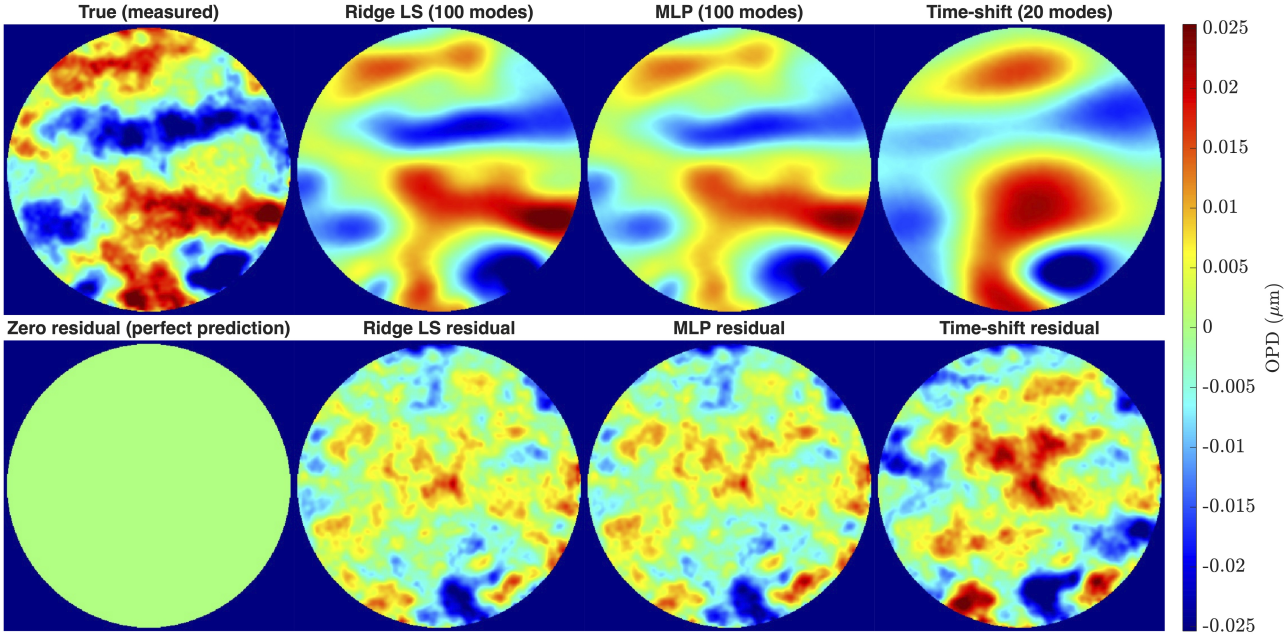

Figure 10. Example measured downstream wavefront and ridge LS, MLP, and time-shift predictions (top row), with corresponding residuals (bottom row) at 2.9δ, $M = 0.58$.

Because the JPOD eigenvalue spectrum decays rapidly (Fig. 6), the higher order modes where MLP shows its strongest advantage over LS, shown in Fig. 11(b), carry little energy λ_n. While these modes show genuine predictive power (ϵ_n largely below 1), the gain in total predicted energy is small. Practically, this indicates that a real controller could retain approximately 60 modes and still recover nearly all of the predictable energy.

The linear control MLP performed virtually identically to the nonlinear MLP at every condition, indicating that the mapping between upstream and downstream mode coefficients is well described by a linear model, and that the MLP's slight edge over LS is not attributable to nonlinear mode interactions.

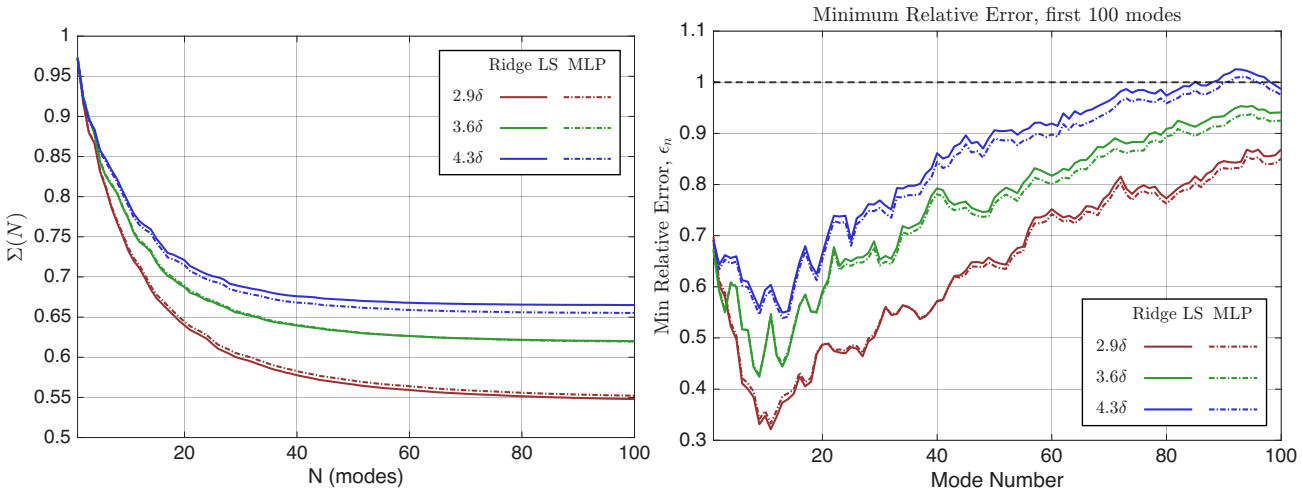

Figure 11. Prediction error of ridge LS (solid) and MLP (dash-dot) at $M = 0.80$ for all three aperture separations. (a) Global error $\Sigma(N)$ vs retained mode count. (b) Per-mode relative error ϵ_n, with dashed line at $\epsilon_n = 1$.

5.2 REGULARIZATION AND THE OPTIMAL RIDGE PARAMETER

Since both predictors reduce to an effectively linear mapping, the remaining distinction between them is how each is regularized. Ridge LS applies a single global penalty λ to every weight (Eq. 26), while the weight decay, dropout, and batch normalization of the MLP allow it to apply different effective regularization to each output mode. The optimal ridge λ is found by sweeping from 0 to 1600 over the validation set as described in Section 4.4, and its value is found to increase with both separation and Mach number, ranging from $\lambda = 200$ at the easiest condition (2.9δ, $M = 0.58$) to $\lambda = 1500$ at the hardest condition (4.3δ, $M = 0.80$). This is because the weaker upstream downstream correlation at these conditions makes the ridge solution more prone to fitting sampling noise in the training data.

Fig. 12 shows the MLP's relative gain over ridge LS as a function of optimal λ for each condition. The MLP transitions from underperforming compared to ridge LS at small λ (up to 3% worse at the easiest condition) to slightly outperforming it at large λ (up to 2% better at the extreme condition), consistent with the trend in Fig. 11(a) where the MLP advantage appears only at the largest separations. This advantage is concentrated in the higher order modes, where the upstream to downstream correlation is weakest. A single global λ must compromise between light shrinkage to preserve the low order modes' strong signal and heavy shrinkage to control noise in the high order modes, whereas the MLP can regularize each mode independently.

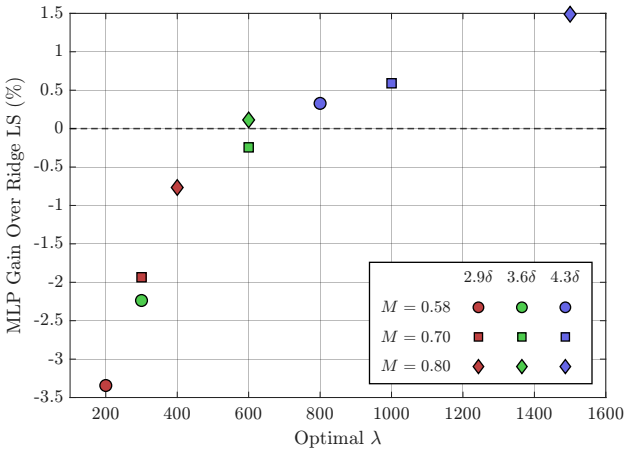

Figure 12. Relative MLP gain over ridge LS as a function of optimal ridge λ across all nine flow conditions.

5.3 STRUCTURE OF THE RIDGE WEIGHT MATRIX

Given that ridge LS and MLP are nearly equivalent predictors, we now examine the underlying cross aperture correlation tensor $C(m, n, k)$ (Eq. 28) from which the ridge predictor's weights are constructed. Two features stand out in Fig. 13, which shows two slices of the tensor for the 2.9δ, $M = 0.58$ case. First, the single frame coupling $|C(m, n)|$ at $k = 0$ in Fig. 13(a) has a dominant diagonal with $|C(n, n)| = 0.31$ on average across modes, while off diagonal elements are individually weak at $|C(m \neq n, n)| = 0.017$. Second, the diagonal temporal profile $|C(n, n, k)|$ in Fig. 13(b) shows that low order modes remain correlated over a broad range of lags around $k = 0$ while high order modes are sharply peaked near the convective delay. Extending to all mode pairs, Fig. 14 plots the full single frame coupling $|C(m, n, k)|$ at representative lags, where it is shown that as $|k|$ grows, the diagonal weakens and the off diagonal content strengthens. At $k = 0$ (the optimal lag τ), a convecting structure sits at the same position inside both apertures and is best captured by the same mode, while at other lags the structure is at a different position within the downstream aperture and projects onto different modes.

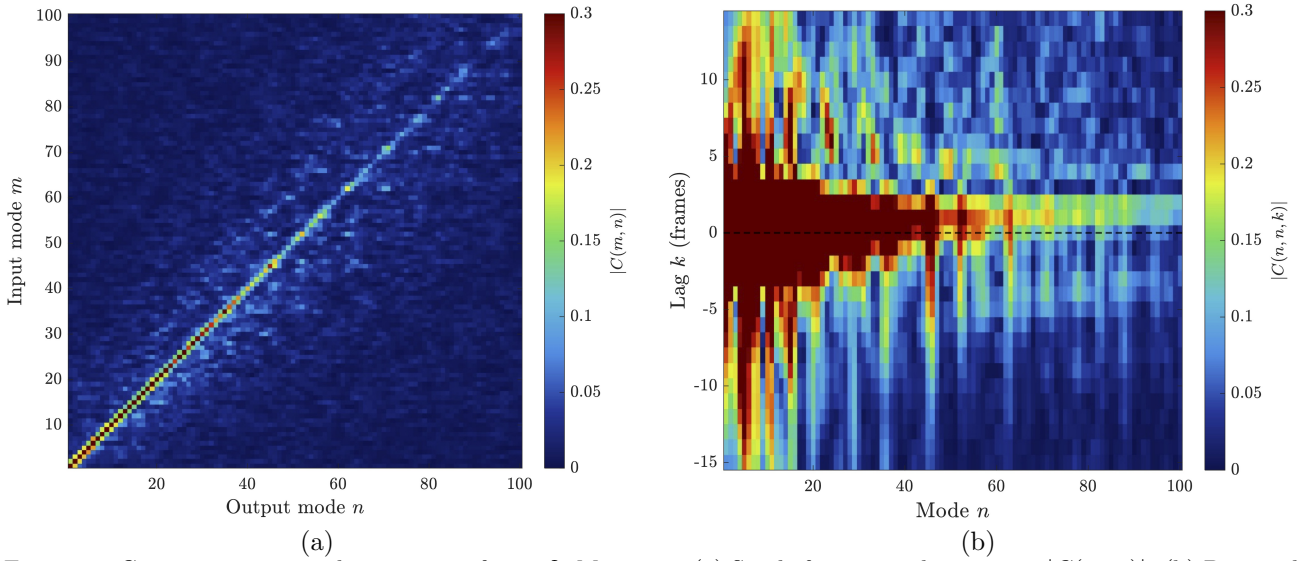

Figure 13. Cross aperture correlation tensor for 2.9δ, $M = 0.58$. (a) Single frame coupling matrix $|C(m, n)|$. (b) Diagonal temporal profile $|C(n, n, k)|$.

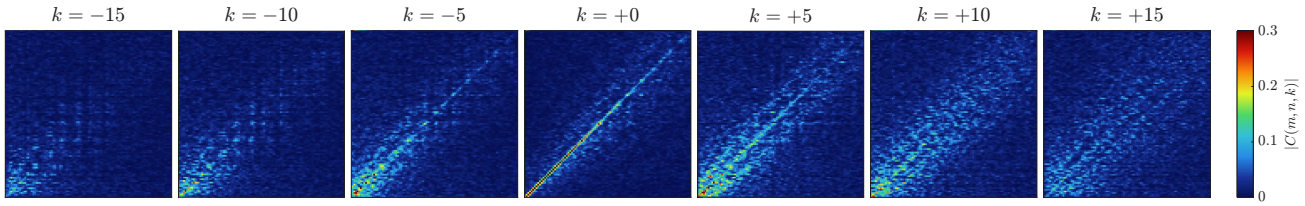

Figure 14. Cross aperture correlation $|C(m, n, k)|$ at representative lags for the 2.9δ, $M = 0.58$ case.

To quantify how much predictive information lives in each part of the tensor, we divide $C(m, n, k)$ into four regions: a center element ($m = n$, $k = 0$), the diagonal window ($m = n$, $k \neq 0$), the off diagonal single frame ($m \neq n$, $k = 0$), and the cross mode temporal ($m \neq n$, $k \neq 0$), containing 1, 29, 99, and 2871 elements respectively. A ridge LS predictor is trained on each and evaluated using Σ_{min} from Eq. (22). For the 2.9δ, $M = 0.58$ case, the center element alone recovers 39% of the downstream wavefront energy, adding the diagonal window brings this to 45%, and further adding the off diagonal single frame brings it to 47%. In contrast, using the 2871 cross mode temporal elements recovers 63%, essentially matching the full tensor predictor's 64%, and this near equivalence holds across all nine flow conditions.

The 16% gap between the $C(m \neq n, n, k \neq 0)$ predictor and the combined $C(m = n, k) + C(m, n, k = 0)$

predictor comes from the former's ability to track a convecting structure through mode space as it moves through the aperture. Mode n's own history captures only how well mode n itself represents the structure at each stage of aperture alignment, which allows some noise averaging across redundant time steps but mixes signal with alignment noise indistinguishably. In contrast, $C(m \neq n, n, k \neq 0)$ shows how the structure's projection onto the mode basis evolves across the temporal window, letting ridge LS learn correlations that reflect each convecting structure's trajectory and separate it from unpredictable turbulent fluctuations. This is why $C(n, n, 0)$ can be completely reconstructed from other modes at other lags. The full ridge LS predictor therefore succeeds not by identifying the single strongest correlation (although this is needed to find τ), but by learning the joint mode and lag distribution of each convecting structure across the temporal window.

6. CONCLUSIONS

This work extended the multiple-aperture wavefront prediction method to a dual-wavelength digital holography system and evaluated a linear ridge LS predictor and a nonlinear MLP across three aperture separations (2.9δ, 3.6δ, and 4.3δ) and three Mach numbers (0.58, 0.70, and 0.80). Global prediction error increased with aperture separation, since a longer convective delay allows boundary-layer structures more time to decorrelate before reaching the downstream aperture. Error also increased with Mach number at every separation tested. The time-shift baseline provided the convective speed U_c and delay τ needed to center each upstream window, but in terms of prediction was substantially outperformed by both window predictors, which each recovered 64% and 34% of the downstream wavefront energy at the easiest and most extreme conditions, compared to 30% and -21% from time shift.

Ridge LS and MLP performed nearly identically across all nine conditions, differing by no more than 3%, demonstrating the strong linear behavior of the flow's upstream to downstream dynamics. The learned MLP function therefore reduces to

$$\tilde{\mathbf{a}}^{\mathrm{DN}}(t) = f(\mathbf{x}(t)) \approx \mathbf{x}(t)\mathbf{W}, \tag{33}$$

matching ridge LS (Eq. 25). What remains of the difference between the two predictors is attributable to regularization. The MLP's slight edge at the largest separations reflects its per mode regularization that the single global λ of ridge LS cannot replicate. A per mode regularized linear predictor could be built to take advantage of this benefit, although the advantage would only become significant at the very largest Mach numbers and aperture separations.

With both predictors reducing to the same linear mapping, the predictive information they use is dominated by the aperture windowing effect. As convecting structures move through the downstream aperture, they project onto different modes at different lags, and the cross mode temporal elements of the correlation tensor in Eq. (28) capture this. The apparent temporal evolution of individual modes is largely a consequence of aperture effects, though the weak cross mode coupling that contributes an additional 2% to the prediction suggests that there is weak physical evolution of structures during transit. Overall, ridge LS succeeds by tracking each convecting structure's location within the mode basis rather than by identifying the single strongest correlation, which allows it to distinguish coherent structures from unpredictable turbulent fluctuations. It is also found that retaining approximately 60 modes captures nearly all the predictable energy (Fig. 11a), reducing the computational cost in a mode based predictive controller.

The limits observed in our predictions likely stem from the sensing method itself. Because each wavefront is a path integral of density fluctuations across the tunnel (Eq. 4), contributions from the opposite wall boundary layer are included in the total signal and cannot be resolved, which imposes a noise floor. This path integrated nature may also explain why nonlinear modeling does not offer an advantage, since multiple structures at the same aperture location collapse into a single, seemingly random signal. Isolating each boundary layer's contribution may reduce this noise floor, but the same path integrated nature would still merge structures at different heights within a single boundary layer, so some of this unresolved mixing would likely persist even if the opposite wall's contribution were fully removed.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of the Joint Experimentation Technology Accelerator (JETX) DEWD for the research described in this paper.

REFERENCES

[1] Gordeyev, S., Smith, A. E., Cress, J. A., and Jumper, E. J., "Experimental studies of aero-optical properties of subsonic turbulent boundary layers," *Journal of Fluid Mechanics* **740**, 214–253 (2014).

[2] Pedrotti, F. L., Pedrotti, L. M., and Pedrotti, L. S., [*Introduction to Optics*], Pearson Prentice Hall, Upper Saddle River, NJ, 3rd ed. (2007).

[3] Burns, R., Jumper, E. J., and Gordeyev, S., "A robust modification of a predictive adaptive-optic control method for aero-optics," in [*47th AIAA Plasmadynamics and Lasers Conference*], (2016). AIAA Paper 2016-3529.

[4] Burns, R., Jumper, E. J., and Gordeyev, S., "A latency-tolerant architecture for airborne adaptive optic systems," in [*53rd AIAA Aerospace Sciences Meeting*], (2015). AIAA Paper 2015-0679.

[5] Kemnetz, M. R., Burns, R., and Gordeyev, S., "Multiple aperture approach for the study of large-scale boundary-layer structures," in [*47th AIAA Plasmadynamics and Lasers Conference*], (2016). AIAA Paper 2016-3530.

[6] Kemnetz, M. R. and Gordeyev, S., "Multiple aperture approach to wavefront prediction for adaptive-optic applications," in [*54th AIAA Aerospace Sciences Meeting*], (2016). AIAA Paper 2016-1460.

[7] Mitchell, J. and Gordeyev, S., "Design and validation of a dual-wavelength digital holography wavefront sensor," *Optical Engineering* **65**(5), 055106 (2026).

[8] Sutherland, W., "The viscosity of gases and molecular force," *Philosophical Magazine* **36**(223), 507–531 (1893).

[9] Schlatter, P., Örlü, R., Li, Q., Brethouwer, G., Fransson, J. H. M., Johansson, A. V., Alfredsson, P. H., and Henningson, D. S., "Turbulent boundary layers up to $re_\theta = 2500$ studied through simulation and experiment," *Physics of Fluids* **21**(5), 051702 (2009).

[10] Chu, E. D., Bukowski, T. J., and Gordeyev, S., "Studies of local shock effects on Shack-Hartmann and digital holography wavefront sensors," in [*Unconventional Imaging, Sensing, and Adaptive Optics 2023, Proc. SPIE*], **12693**, 1269317 (2023).

[11] Ghiglia, D. C. and Romero, L. A., "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," *Journal of the Optical Society of America A* **11**(1), 107–117 (1994).

[12] Berkooz, G., Holmes, P., and Lumley, J. L., "The proper orthogonal decomposition in the analysis of turbulent flows," *Annual Review of Fluid Mechanics* **25**, 539–575 (1993).

[13] Holmes, P., Lumley, J. L., and Berkooz, G., [*Turbulence, Coherent Structures, Dynamical Systems and Symmetry*], Cambridge University Press, Cambridge, UK (2012).

[14] Gordeyev, S., De Lucca, N., Jumper, E. J., Hird, K., Juliano, T. J., Gregory, J. W., Thordahl, J., and Wittich, D. J., "Comparison of unsteady pressure fields on turrets with different surface features using pressure-sensitive paint," *Experiments in Fluids* **55**(1661) (2014).

[15] Hoerl, A. E. and Kennard, R. W., "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics* **12**(1), 55–67 (1970).

[16] Goodfellow, I., Bengio, Y., and Courville, A., [*Deep Learning*], MIT Press (2016).

[17] Kingma, D. P. and Ba, J., "Adam: A method for stochastic optimization," in [*3rd International Conference on Learning Representations (ICLR)*], (2015). arXiv:1412.6980.