Principal Component Analysis of Near-Wall Turbulent Flow

Vivek Mathew

November 14, 2025

1 Introduction

This report details the application of Principal Component Analysis (PCA), referred to in fluid dynamics as Proper Orthogonal Decomposition (POD), to a three-dimensional turbulent velocity field. The objective is to decompose the complex flow data into a linear combination of orthogonal basis functions (modes) effectively ordered by their energy content. This allows for the separation of dominant coherent structures from stochastic noise and provides insight into the dimensionality of the system.

2 Mean Flow Characterization

Prior to performing the decomposition, the statistical mean of the flow was computed to separate the steady state from the turbulent fluctuations. Figure 1 illustrates the mean streamwise velocity profile averaged over all temporal snapshots.

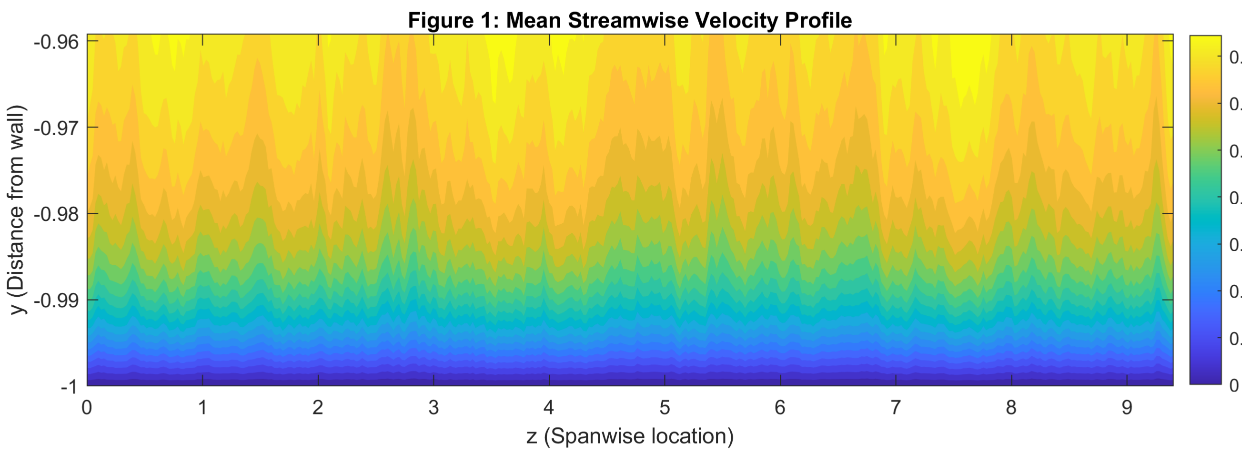

Figure 1: Contour plot of the mean streamwise velocity. The vertical axis represents the wall-normal direction (y), and the horizontal axis represents the spanwise direction (z). The gradient shows the boundary layer profile, where velocity transitions from zero at the wall to the free-stream velocity.

The profile exhibits homogeneity in the spanwise (z) direction, meaning the statistics do not vary across the horizontal axis. The primary gradient exists in the wall-normal (y) direction. PCA was subsequently performed on the fluctuation field, defined as the instantaneous velocity minus this mean profile.

3 Covariance Matrix and Dimensionality

Standard PCA requires the computation of a covariance matrix, which describes the correlation between every pair of features in the dataset. In this context, the "features" are the spatial grid points.

The simulation grid consists of $N_y = 48$ points in the wall-normal direction and $N_z = 384$ points in the spanwise direction. To format this for PCA, the 2D spatial field is flattened into a single feature vector of length M:

$$M = N_y \times N_z = 18,432$$

Computing the standard spatial covariance matrix would result in a matrix of dimensions $M \times M$, or $18,432 \times 18,432$. This matrix contains approximately 339.7 million elements. Storing and performing an eigendecomposition on a matrix of this size is computationally prohibitive for standard workstations due to memory constraints.

To resolve this, the Method of Snapshots was employed. Instead of computing the large $M \times M$ spatial covariance matrix, we computed the inner product of the temporal snapshots. Since the number of snapshots ($N_x = 512$) is significantly smaller than the number of spatial points ($M = 18,432$), the resulting matrix has dimensions of only 512×512. The eigenvectors of this smaller matrix were then mapped back to the full spatial domain to recover the exact POD modes.

4 Energy Spectrum Analysis

The eigenvalues obtained from the decomposition represent the energy (variance) captured by each corresponding mode. Figure 2 presents the distribution of this energy.

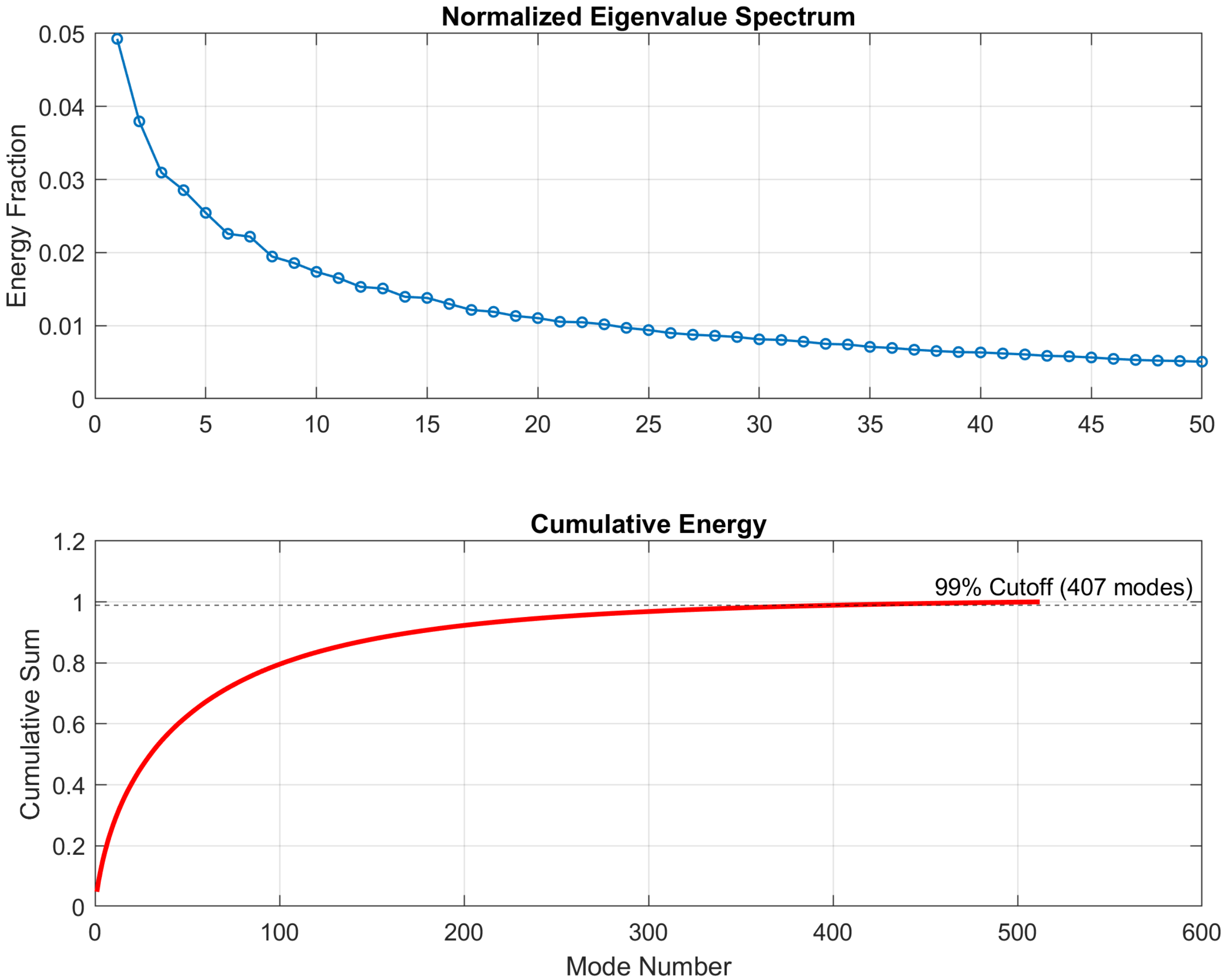

Figure 2: Top: Normalized eigenvalue spectrum showing the fraction of total energy contained in each mode. Bottom: Cumulative energy sum indicating the total variance captured as a function of the number of modes included.

The spectrum shows a slow decay in energy. The first mode captures only a small fraction of the total variance (approximately 5%), and the energy is distributed broadly across the subsequent modes. To capture 99% of the total kinetic energy in the flow, the cumulative sum indicates that 407 modes are required. This result characterizes the flow as a high-dimensional system, where the turbulence cannot be accurately represented by a low-order approximation.

5 Spatial Mode Visualization

The computed POD modes represent the orthogonal basis vectors that maximize the captured energy. Figure 3 visualizes the spatial structure of the first four modes.

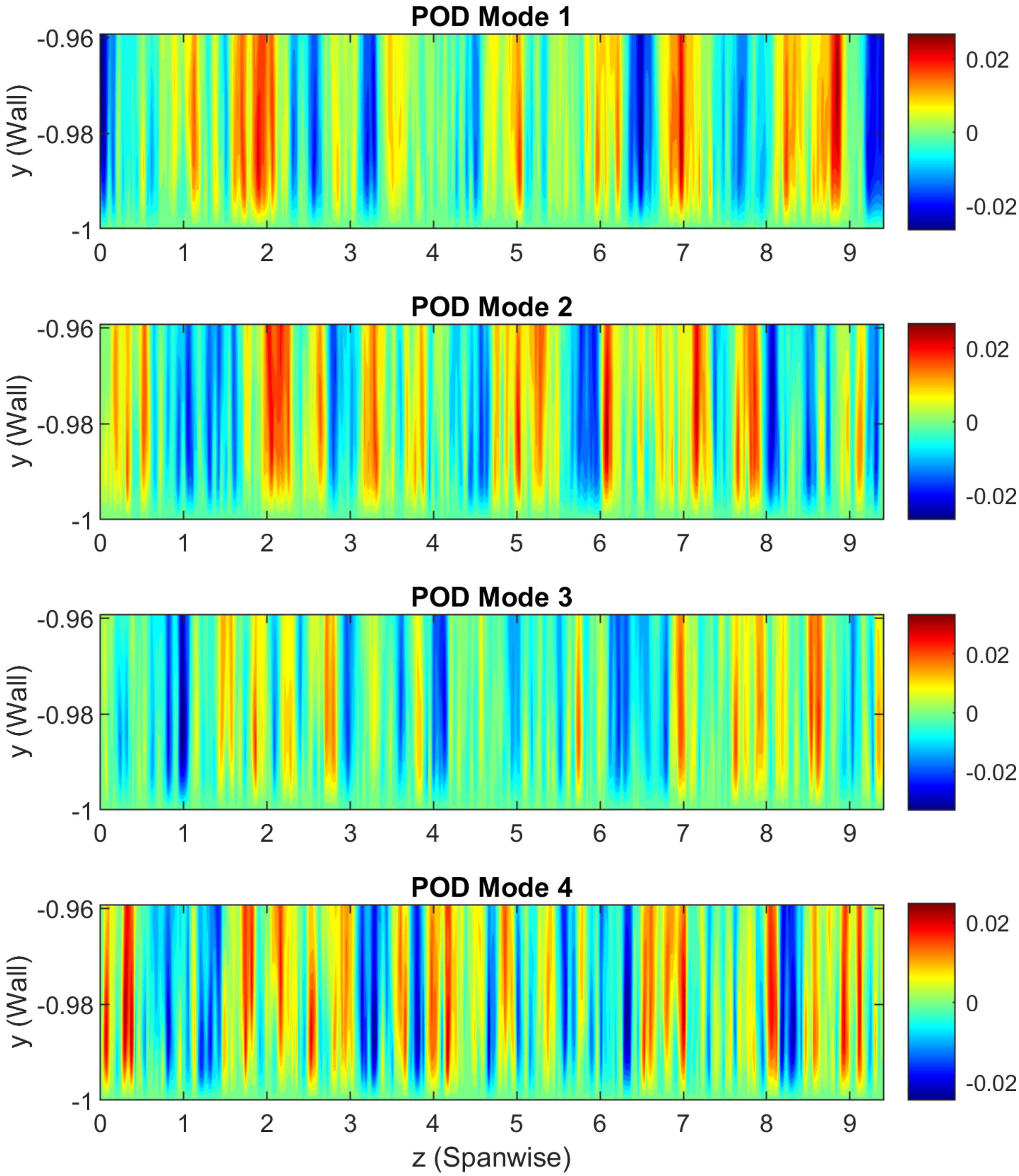

Figure 3: Contour plots of the first four POD spatial modes in the $y - z$ plane. These modes represent the dominant coherent structures in the flow fluctuations.

The modes exhibit distinct alternating positive and negative regions along the spanwise direction. In the context of boundary layer turbulence, these structures correspond to near-wall velocity streaks. The alternating signs indicate regions of high-momentum fluid moving toward the wall and low-momentum fluid moving away from the wall. As the mode number increases, the spatial frequency of these structures increases, representing smaller scales of turbulence.

6 Conclusion

This analysis successfully decomposed the near-wall turbulent flow field using the Method of Snapshots to circumvent the memory limitations associated with the full 340-million-element covariance matrix. The results indicate that the flow is characterized by high dimensionality, requiring 407 orthogonal modes to reconstruct 99% of the turbulent energy. The visualized modes confirm that the energy is organized into spanwise-alternating streak structures consistent with turbulent boundary layer physics.