Flow analysis using Data Decomposition Methods

A Report submitted

on completion of summer internship

Bachelor of Technology

in

Aerospace Engineering

by

Sambhav Jain

(SC17B038)

pursued in

Department of Aerospace Engineering

Indian Institute of Space Science and Technology

to

INDIAN INSTITUTE OF SPACE SCIENCE AND TECHNOLOGY

THIRUVANANTHAPURAM

June-July 2019

CERTIFICATE

This is to certify that the project report entitled **"Flow Analysis using Data Decomposition Methods"** submitted by **Sambhav Jain**, to the Indian Institute of Space Science and Technology, Thiruvananthapuram, in completion of the summer internship June-July 2019, is a bonafide record of the internship work carried out by him under my supervision. The contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. B.R. Vinoth
Assistant Professor
Department of Aerospace Engineering

Dr. Manoj T. Nair
Head of the department
Department of Aerospace Engineering

Place: IIST, Thiruvanathapuram
July 2019

Declaration

I declare that this report has been written by me. No part of it has been plagiarised. All content from any other sources have been duly cited or acknowledged. I aver that if any part of the report is found to be plagiarised, I shall take full responsibility of it. I also declare that no content of this report has been submitted to any other institute or university for the award of any diploma or degree.

Place: IIST, Thiruvanathapuram
July 2019

Sambhav Jain
SC17B038

Acknowledgements

I gratefully acknowledge the support of my mentors, professors, colleagues, friends, and family during this project work. I would like to thank my project supervisor, Dr. B. R. Vinoth, Assistant Professor, Department of Aerospace Engineering, for his extraordinary support, guidance, and planning throughout the project. I appreciate his knowledge, experience, resourcefulness, and dedication to his students. It was a priceless opportunity learning from and with him. My thanks are due also to my good friends- Rajdip podder and Shashank Tomar, to name a few- for excellent discussions and interactions I had with them regarding my project, which helped me resolve several issues faced by me during my work. Most importantly, I express my love and gratitude to my family, who have been the greatest support and guidance I have. My twin brother- Sahaj Jain is a constant support and a source of inspiration for me, who stands for help whenever I need help. I am sure to have missed mentioning several people who have directly, or indirectly, acted in the best interests of my project work, and I am indebted to all of them. I thank you all.

Abstract

As the technological advancement is increasing in all the aspects of science, covering from neurological science to fluid dynamics, we have sophisticated our equipments so much that even the minute details can now be obtained. But, job doesn't end at obtaining data, we have to process it and extract useful information out of it. Going by the conventional way of numerical computing won't be that easy and will cost much higher time. So, to sort out such problems, some methods and techniques have been developed. This report gets into the insight of two of such techniques which are able to compute Low Rank Approximations (LRA)- Proper Orthogonal Decomposition(POD) and Dynamic Mode Decomposition (DMD). Such process are able to approximate huge data with the lower rank approximations, reducing the computation time drastically. One of the most striking features of such decomposition methods is that these methods are data-based methods, not the model-based methods. This means that these algorithms does not need the governing equations to be known and is completely data-driven.

The objective of this report is to introduce the reader of the various decomposition techniques and its mathematical basis. Two chapters are dedicated to make the reader know the fundamentals of this concept and to gain the mathematical background. Also, application of these techniques are used and all aspects of the results are analysed. In, short the report serves the reader the crux of these decomposition methods and how to apply them to a given data. The stability of the modes can also calculated using the eigen values of the DMD modes. The analysis of spectrum is also performed to get into the clear insight into the results.

Keywords: Data decomposition techniques, Low Rank Approximations, data-driven, modes, stability, eigen-values, spectrum

Table of contents

List of figures

Chapter 1

Introduction

1.1 Motivation

High-dimensionality is a common challenge in processing data from large complex systems. These systems may involve large measured data sets including audio, image, or video data. The data may also be generated from a complex physical system, such as neural recordings from the mammalian cortex, or fluid velocity measurements from a simulation or experiment. In any case, it is observed that most data from naturally occurring systems exhibits dominant patterns of activity, which may be characterized by a low-dimensional attractor or manifold.

So, there are many low rank approximation techniques developed and currently developing, which represent the data to a good approximation using much smaller basis. Singular Value Decomposition (SVD), Dynamic Mode Decomposition (DMD), etc. are the various algorithms developed and are able to approximate through lower rank matrices. Most important fact about these algorithms is that all these are purely data driven methods, implying that the governing equations through which the data is being generated is completely unknown and various methods are implemented to obtain the lower rank approximations, which contains most of the predominated dynamics of the system.

1.2 Scope of the project

The data analysed in this report is of the flow of Helium jet, obtained experimentally. The report acquaints the reader to the algorithms used in the lower rank approximation methods and how to analyse the data. The mathematical theory is included in this report. This report deals with two data-driven lower rank approximation techniques, namely :

1. Singular Value Decomposition (SVD)

2. Dynamic Mode Decomposition (DMD)

1.3 Organisation of Study

The following section presents the reader with the conventional data processing techniques used for fluid mechanics measurements and their limitations. Chapter 2 is devoted to present the POD in the context of approximation theory, followed by Singual Value Decomposition (SVD) as an application of the technique to the real data. Also the geometrical interpretation of the technique is shown in subsection 2.4.4. Chapter 3 presents DMD , a recent technique that attempts to extract dynamic information based on the sequence of snapshots. A brief description of DMD algorithms (SVD based) to apply this to both experimental and numerical data follows. Also the advantages of this technique over POD is listed with limitations. Chapter 4 is dedicated to the application of these techniques. The results of the analysis of Helium jet were obtained and then analysed using different aspects. The chronos, singular mode energy distribution, phase shifts, FFT and various such methods are applied on the results to analyse the results and get a better insight to the decomposition methods.

1.4 Some Fundamental Data Processing Techniques

There are some of the basic fundamental data processing techniques, which are listed as follows:

1.4.1 Fourier Transform

The Fourier transform (FT) is an integral transform with orthogonal sinusoidal basis functions of different frequencies. The result represents the frequency spectrum of the signal. Depending on the characteristics of the original (time) signal, different variants of the transform are defined.

A continuous periodic signal x(t) = x(t +T) with the period T can be decomposed into an infinite series of sinusoidal functions (Fourier series), whose linear combination reproduces the original function.

$$x(t) = \sum_{k=-\infty}^{\infty} a_k cos(\frac{2\pi kt}{T}) + \sum_{k=-\infty}^{\infty} b_k sin(\frac{2\pi kt}{T})$$

The Fourier coefficients a_k and b_k are given by:

$$a_k = \frac{1}{T}\int_0^T x(t)cos(\frac{2\pi kt}{T})dt$$

$$b_k = \frac{1}{T}\int_0^T x(t)sin(\frac{2\pi kt}{T})dt$$

In terms of a superposition of sinusoids, the **fundamental frequency** f_o is the lowest frequency sinusoidal in the sum.

Super-harmonics is the signal having frequency which is integral multiple of fundamental frequency.

If the frequency of a signal is less than the fundamental frequency then it is called **sub-harmonics**. This means, the frequency of sub-harmonics will be (1/n)th times of the fundamental frequency.

Continuous Fourier Transform

By using T $\rightarrow \infty$, for a continuous complex signal x(t) with finite energy content the superposition of the Fourier series becomes:

$$x(t) = \int_{-\infty}^{\infty} X(f)exp(2\pi ift)df$$

with

$$X(f) = \int_{-\infty}^{\infty} x(t)exp(2\pi ift)dt$$

It is a continuous, infinite, non-periodic, complex frequency spectrum.

Discrete Fourier Transform

A finite series of complex values x_n = x(t = nδt_s) with n = 0, 1, . . . , N -1, sampled at equal time intervals and over the time duration $0 \leq t < T = N\delta t_s$ can be decomposed into a finite sum of complex Fourier coefficients X_k, yielding the discrete Fourier transform **(DFT)**. The DFT is defined as

$$X_k = X(f = k\delta f)$$
$$= \sum_{n=0}^{N-1} x_n exp(-i\frac{2\pi nk}{N})$$

$$k = 0, 1, ..., (N-1)$$

The **Nyquist Sampling Theorem** states that: A bandlimited continuous-time signal can be sampled and perfectly reconstructed from its samples if the waveform is sampled over twice as fast as it's highest frequency component. A major disadvantage of Fourier transform is that Fourier series give no information on the spatial/temporal localization of features.

1.4.2 Limitations of fundamental techniques

The fundamental data processing techniques are very useful and their application is great. They maintains information on amplitude, harmonics, and phase and uses all parts of the waveform to translate the signal into the frequency domain. But, there are some limitations to them as well, which are enough to promote to look for other techniques to process the data. They are :

1. All the techniques can be only applied on a time series,i.e., a signal captured at a point throughout the duration. They are called **local decomposition techniques**. It's hence not possible to analyse the flow field globally using these preliminary techniques.

2. All the techniques discussed previously require a predefined basis functions (of user choice). So the processing requires proper guidance. With this limitation manipulating fluid flow field data could become cumbersome. This has led to the development of data-driven methods which are completely based on the available data but not on the model.

3. Since the experimental data comes with a huge amount of noise the extraction of physical mechanisms behind is very limited with these preliminary techniques.

4. It was not possible to represent the data through decomposition, we have to process a huge amount of data.

1.5 Literature Review

The POD was introduced in the context of turbulence by Lumley (1967). In other disciplines the same procedure goes by the names Karhunen$-Lo$ve decomposition or principal components analysis and it seems to have been independently rediscovered several times, Sirovich (1987). According to Lumley, quoting A. Yaglom (personal communication), the POD was suggcsted independently by several scientists, e.g. Kosambi (1943), Loève (1945), Karhunen (1946), Pugachev (1953), and Monin (n.d.). For use of the POD in other disciplines

see: Papoulis (1965)–random variables; Rosenfeld and Kak (1982) −−image processing; Algazi and Sakrison (1969) −−signal analysis ; Andrews et al. (1967)−−data compression; Preisendorfer (1988) ocean-ography; and Gay and Ray (1986, 1988) process identification and control in chemical engineering. Introductory discussions of the method in the context of fluid mechanical problems can also be found in Kirby and Sirovich (1990), (Sirovich, 1989), Sirovich (1987) and (Holmes et al., 1998).

The DMD method originated in the fluid dynamics community as a method to decompose complex flows into a simple representation based on spatio temporal coherent structures. Schmid (2010) first defined the DMD algorithm and demonstrated its ability to provide insights from hig$h-d$imensional fluids data. The growing success of DMD stems from the fact that it is an equatio$n-free, data-d$riven method capable of providing an accurate decomposition of a complex system into spatio temporal coherent structures that may be used for short-time futur$e-sta$te prediction and control. More broadly, DMD has quickly gained popularity since Mezić (2013), (Tu et al., 2014), (Chen et al., 2012) showed that it is connected to the underlying non linear dynamics through Koopman operator theory and is readily interpretable using standard dynamical systems techniques.

Chapter 2

Proper Orthogonal Decomposition (POD)

2.1 Introduction

The proper orthogonal decomposition is a procedure for extracting a basisfor a modal decomposition from an ensemble of signals. Its power lies inthe mathematical properties that suggest that it is the preferred basis touse in many circumstances. The proper orthogonal decomposition (POD) is a powerful and elegant method for data analysis aimed at obtaining low-dimensional approximate descriptions of a high-dimensional process. The POD provides a basis for the modal decomposition of an ensemble of functions, such as data obtained in the course of experiments or numerical simulations.

In recent years, there have been many reported applications of the POD methods in engineering fields. The POD has been used in various disciplines including random variables, image processing, signal analysis, data compression, process identification and control in chemical engineering, oceanography, etc. . In the bulk of these applications,the POD is used to analyze experimental data with the objective of extracting dominant features. The POD has been used to obtain approximate, low-dimensional descriptions of turbulent fluid flows , structural vibrations and chaotic dynamical systems, and more recently, micro electromechanical systems (MEMS).

2.2 Special features and properties of POD

The most striking feature of the POD is its optimality: it provides the most efficient way of capturing the dominant components of an infinite-dimensional process with only a finite

number of modes, and often surprisingly few modes. In the field of experimental and numerical fluid mechanics, POD has a special place reserved for it because of its striking properties, like (Berkooz et al., 1993)

1. It is statistically based–extracting data from experi-ments and simulations.

2. Its analytical foundations supply a clear under-standing of its capabilities and limitations.

3. It permits the extraction,from a turbulent field, of spatial and temporal structures judged essential according to predetermined criteria and it provides a rigorous mathematical framework for their description.

As such, it offers not only a tool for the analysis and synthesis of data from experiment or simulation,but also for the construction, from the Navier-Stokes equations, of low-dimensional dynamical models for the interaction of these essential structures. Thus, coming full circle, it is a statistical technique that contributes to deterministic dynamical analysis.

2.3 POD Concept

Suppose we wish to approximate a function z(x, t) over some domain of interest as a finite sum in the variablesseparated form Chatterjee (2000) :

$$z(x,t) \approx \sum_{k=1}^{M} a_k(t)\phi_k(x),$$

with the reasonable expectation that the approximation becomes exact in the limit as M approaches infinity. We can think x as a spatial coordinate (possibly vector valued) and t as a temporal coordinate.

The representation of the function z(x,t) is not unique. If the domain of x is a bounded interval X on the real line, then the functions ϕ_k(x) can be chosen as a Fourier series, or Legendre polynomials, or Chebyshev polynomials, and so on. For each such choice of a sequence $\phi_k(x)$ that forms a basis for some suitable class of functions z(x, t), the sequence of time-functions a_k(t) is different. That is, for sines and cosines we get one sequence of functions $a_k(t)$ is different. That is, for sines and cosines we get one sequence of functions a_k(t), for Legendre polynomials we get another, and so on. The POD is concerned with one possible choice of the functions ϕ_k.

If we have chosen orthonormal basis function,i.e.,

$$\int_x \phi_{k1}(x)\phi_{k2}(x)dx = \begin{cases} 1\ if\ k_1 = k_2 \\ 0\ otherwise \end{cases}$$

then ,

$$a_k(t) = \int_x z(x,t)\phi_k(x)dx,$$

i.e., for orthonormal basis function $a_k(t)$ depends only on $\phi_k(x)$ and not on the other ϕ's. For all, a sequence of orthonormal functions ϕ_k(x) such that the first two of these functions give the best possible two-term approximation, the first seven give the best possible seven term approximation, and so on. These special ordered, orthonormal functions are called the proper orthogonal nodes for the function z(x, t). With these functions, the expression of z(x,t) is called the POD of z(x, t).In literature, the spatial modes are referred by the name topos, while the temporal mode by the name chronos.

2.4 Singular Value Decomposition (SVD)

2.4.1 Introduction

The SVD provides a systematic way to determine the dominant patterns underlying a high-dimensional system, providing a low-rank approximation to high-dimensional data. This technique is *data-driven* in that patterns are discovered purely from the data, without the addition of any expert knowledge or intuition. The SVD may be thought of as a numerically stable computation that provides a hierarchical representation of the data in terms of a new coordinate system defined by dominant correlations within the data.
The SVD has many powerful applications beyond dimensionality reduction of high dimensional data. It will be useful in computing the pseudo-inverse of non-square matrices, providing solutions to under determined or overdetermined matrix equations that are either minimum norm solutions, or solutions that minimize the sum-squared error. We will also use the SVD in a principled approach to de-noising data sets. The SVD is also important to characterize the input and output geometry of a linear map between vector spaces. These applications will all be explored in this chapter, providing an intuition for matrices and high-dimensional data.

2.4.2 Definition

Generally, we are interested in analysing a large data set X:

$$\begin{bmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_n \\ | & | & & | \end{bmatrix}$$

The columns $x_k \in \mathbb{C}^n$ may be measurements from simulations or experiments. For example, columns may represent images that have been reshaped into column vectors with as many elements as pixels in the image. The column vectors may also represent the state of a physical system that is evolving in time, such as the fluid velocity at each point in a discretized simulation or at each measurement location in a wind-tunnel experiment. The index k is a label indicating the kth distinct set of measurements; for many of the examples in this book X will consist of a time-series of data, and $x_k = x(k \triangle t)$. Often the *state-dimension* n is very large, on the order of millions or billions in the case of fluid systems. The columns are often called *snapshots*, and m is the number of snapshots in X. For many systems $n \gg m$, resulting in a *tall-skinny matrix*, as opposed to a *short-fat matrix* when $n \ll m$.

The SVD is a unique matrix decomposition that exists for every complex valued matrix X$\in \mathbb{C}^{n\times m}$;

$$X = U\Sigma V^*$$

where U$\in \mathbb{C}^{n\times n}$ and V $\in \mathbb{C}^{m\times m}$ are unitary matrices.[1] and $\sum \in \mathbb{C}^{n\times m}$ is a matrix with non-negative entries on the diagonal and zeroes off the diagonal.
The matrix $\sum$ has at most m non-zero elements on the diagonal, and may therefore be written as:

$$\begin{bmatrix} \widehat{\Sigma} \\ 0 \end{bmatrix}$$

. Therefore, it is possible to exactly represent X using the reduced SVD:

$$X = U\Sigma V^* = [\widehat{U} \;\; \widehat{U}^{\perp}] \left[\frac{\widehat{\Sigma}}{0}\right] V^* = \widehat{U}\widehat{\Sigma}V^*$$

Here, $\widehat{U}^{\perp}$ is complementary to $\widehat{U}$.
The columns of U are called left singular vectors of X and the columns of V are right singular vectors. The diagonal elements of $\widehat{\Sigma} \in \mathbb{C}^{m\times m}$ are called singular values and the are ordered from largest to smallest.

[1] A complex square matrix U is unitary if its conjugate transpose U^* is also its inverse. For real matrices, unitary is same as orthogonal. (Strang, 2003)

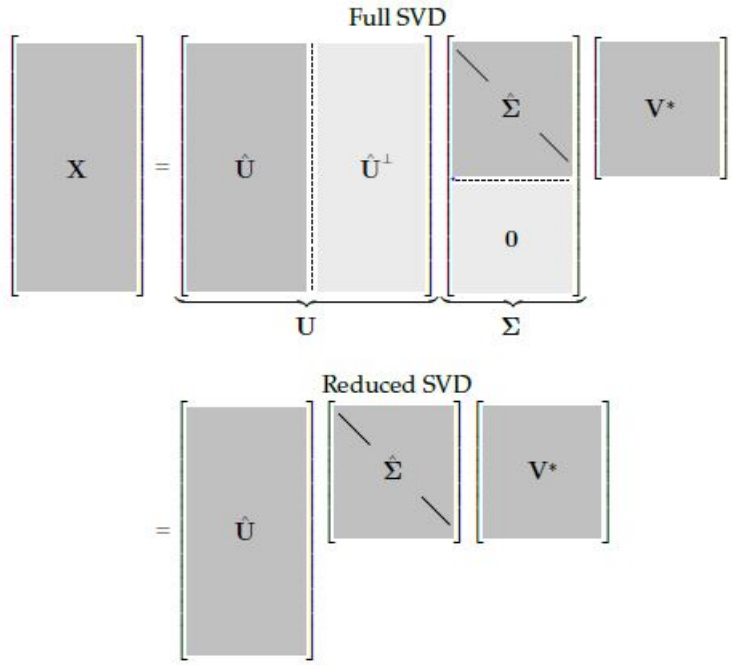

Fig. 2.1 Schematic of the matrices involved in SVD and reduced SVD

The SVD is the basis for many related techniques in dimensionality reduction used to obtain reduced order models (ROMs).

2.4.3 Matrix approximation

Perhaps the most useful and defining property of the SVD is that it provides an optimal low-rank approximation to a matrix X. In fact, the SVD provides a hierarchy of low rank approximations, since a rank-r approximation is obtained by keeping the leading r singular values and vectors, and discarding the rest.

Schmidt (of Gram-Schmidt) generalized the SVD to function spaces and developed an approximation theorem, establishing truncated SVD as the optimal low rank approximation of the underlying matrix X . Schmidt's approximation theorem was rediscovered by Eckart and Young (Eckart and Young, 1936) , and is sometimes referred to as the Eckart-Young theorem.

Eckart-Young theorm states that the optimal rank-r approximation to X, in an L_2 sense, is given by the rank-r SVD truncation $\tilde{X}$:

$$\underset{\tilde{X}}{\text{argmin}} ||X - \tilde{X}||_2 = \tilde{U}\tilde{\Sigma}\tilde{V}^*$$

Here, $\tilde{U}$ and $\tilde{V}$ denotes the leading r columns of U and V, and $\tilde{\Sigma}$ contains the leading $r \times r$ sub-block of Σ.
Here, we establish the notation that a truncated SVD basis (and the resulting approximated matrix $\tilde{X}$) will be denoted by $\tilde{X} = \tilde{U}\tilde{\Sigma}\tilde{V}^*$. Because Σ is diagonal, the rank-r SVD approximation is given by the sum of r distinct rank-1 matrices:

$$\tilde{X} = \sum_{k=1}^{r} \sigma_k u_k v_k^* = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + ... + \sigma_r u_r v_r^*$$

For a given rank r, there is no better approximation for X, in the L_2 sense, than the truncated SVD approximation $\tilde{X}$.

2.4.4 Geometric Interpretation

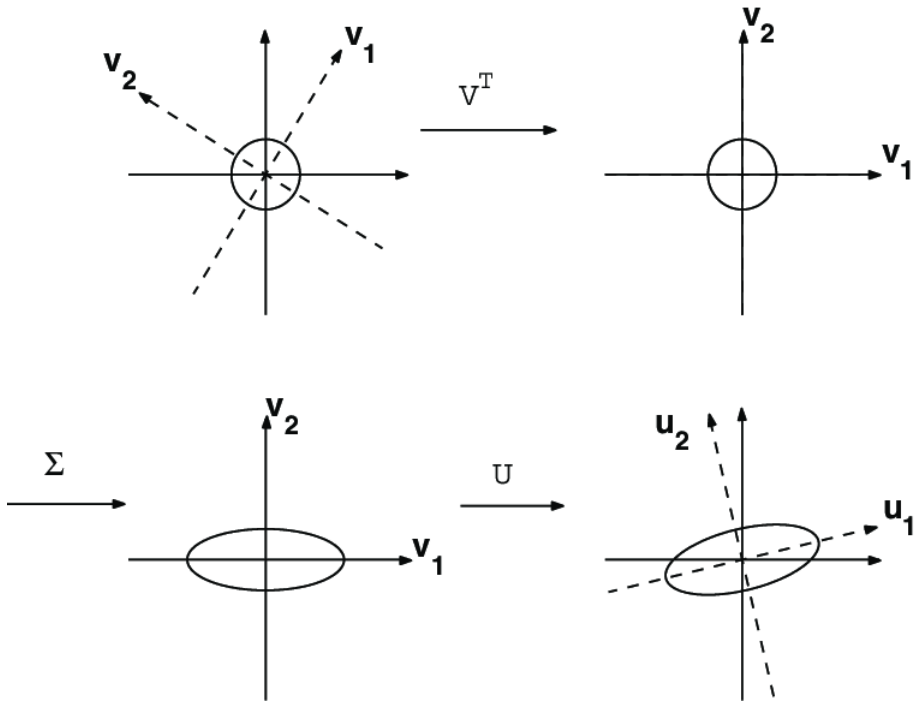

Fig. 2.2 Geometric interpretation of SVD for matrices m=n=2 (Muller et al., 2004)

An important geometrical interpretation of SVD is given in Figure 2.2 for m = n = 2. The image of the unit sphere under any m $\times$ n matrix multiplication is an ellipse. Considering the three factors of the SVD separately, note that V^T is a pure rotation of the circle. Figure 2.2 shows how the axes v1 and v2 are first rotated by V^T to coincide with the coordinate axes. Second, the circle is stretched by Σ in the directions of the coordinate axes to form an ellipse.

The third step rotates the ellipse by U into its final position. Note how v1 and v2 are rotated to end up as u1 and u2, the principal axes of the final ellipse .

2.4.5 Algorithm for computing SVD

Assume a dynamical system with m state variables. We take the measurements at m points in the space (these could be color intensity of the image, velocity of fluid particle or strain value)at n instants of time. We can arrange this data into matrix with i^{th} column containing the spatial information at i^{th} instant of time and j^{th} row contains the temporal information of a particular j^{th} state in space.

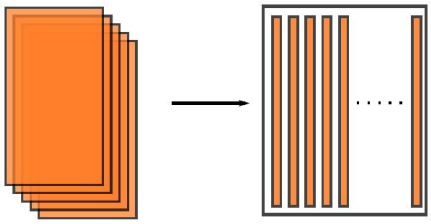

Fig. 2.3 Formation of A matrix from the probe measurement data

It is a common practice to subtract the column mean from each element of that particular column although this does not affect the basic calculation. This mean shift ensures that the data points are centered around origin. Because by definition a subspace should pass through the origin. This will affect the way in which the results are interpreted (Tropea and Yarin, 2007).

Computation of SVD for the matrix A can be done by the eigen value decomposition of AA^T or A^TA. Since, A is a skinny rectangular matrix, it is transformed to a symmetric, semi-positive definite matrix by multiplying with A^T. Now, its possible to diagonalise AA^T or A^TA using eigen-value decomposition.

We know that eigen vectors of a square matrix diagonalise it with eigen values being the diagonal elements, i.e.,

$$A = S\lambda S_{-1}$$

where, S is the matrix with columns as eigen vectors and corresponding eigen value as the diagonal element of λ matrix. Since, our target matrix are symmetric, the eigen vectors produced by eigen value decomposition is orthogonal.

In fluid dynamics, mostly the number spacial points of analysis are much larger than the number of instants at which we take readings, therefore, mostly matrices of skinny form ($m \gg n$) are dealt.

In skinny matrices, its possible to compute singular values with lower computation cost by finding eigen values of A^TA. Substituting A as $U\Sigma V^T$ in A^TA:

$$A^TA = (U\Sigma V^T)^T U\Sigma V^T$$

$$A^TA = V\Sigma^2 V^T$$

$$A^TAV = V\Sigma^2$$

From the above equation, it can be seen that V matrix is the orthogonal eigen vector matrix of A^TA and Σ matrix has singular values equal to the square root of the eigen values.

Once, V and Σ is known, we can compute the U matrix according to the r rank approximation required.

Similarly, we can do the computation of SVD for fat matrices ($n \ll m$) by computing the eigen value decomposition of AA^T as that would be computationally cost efficient. We can do the same as for skinny matrix as:

$$AA^T = U\Sigma V^T (U\Sigma V^T)^T$$

$$AA^T = U\Sigma^2 U^T$$

$$AA^TU = U\Sigma^2$$

Hence, the U matrix is the eigen vector of the AA^T and Σ matrix-singular values is equal to the square root of the eigen values.

Algorithm to find SVD: For A matrix containing the spatial data varying with time , finding a lower rank k^th approximation for the data.

1. $A \leftarrow v_1, v_2, ...v_n$
2. $[m,n] \leftarrow size(A)$
3. if m$>$ n
4. then $[V,\lambda] \leftarrow eig(A^TA)$
5. Sort V based on the singular values$(\sqrt{\lambda})$
6. $U_k \leftarrow AV_{(:,1:k)}$
7. $V_k \leftarrow V_{(:,1:k)}$
8. $A_k \leftarrow U_kV_k^H$

9. else
10. $[U, \lambda] \leftarrow eig(AA^T)$
11. Sort U based on the singular values($\sqrt{\lambda}$)
12. $V_k \leftarrow U_{(:,1:k)}A$
13. $U_k \leftarrow U_{(:,1:k)}$
14. $A_k \leftarrow U_k V_k^H$

2.4.6 Limitations of POD

Despite its widespread use, the POD may sometimes be too simple for dealing with real-world data.Two weaknesses of the method must be highlighted. On the one hand, the linear nature of the method may represent a restriction for some data sets. If the data lie on a non-linear manifold, the method then overestimates the intrinsic dimensionality. For instance, the covariance matrix of data sampled from a helix in$\mathbb{R}^3$ has full-rank, and the POD requires the use of three variables for the description of the data. The helix however is a one-dimensional manifold and can be smoothly parametrized with a single variable. In this case, the method represents the data in a space of higher dimension than the number of intrinsic degrees-of-freedom. On the other hand, the POD tries to describe all the data using the same global features. Generally, a rich data set often has varying characteristics in different regions of the space. This suggests the use of local features for efficient representation of qualitatively different domains of the data.The POD can thus determine an appropriate embedding space for a low-dimensional structure, but it cannot provide the most efficient description of a data set where non-linear dependencies are present. Since the POD removes linear correlations among variables(i.e. diagonalizes the covariance matrix), it is only sensitive to second-order statistics. Accordingly, it does not necessarily yield statistical independence; decorrelation implies statistical independence only under the assumption that the variables are Gaussian.(Kerschen et al., 2005)

Chapter 3

Dynamic Mode Decomposition (DMD)

3.1 Introduction

DMD, at its core, can be thought of as an ideal combination of spatial dimensionality-reduction techniques, such as the proper orthogonal decomposition (POD), with Fourier transforms in time. Thus, correlated spatial modes are also now associated with a given temporal frequency, possibly with a growth or decay rate. The method relies simply on collecting snapshots of data x_k from a dynamical system at a number of times t_k, where k=1,2,3,.....m. DMD is algorithmically a regression of data onto locally linear dynamics $x_{k+1} = Ax_k$, where A is chosen to minimize $||x_{k+1} - Ax_k||_2$ over the k=1,2,3,....,m-1 snapshots. The advantages of this method are that it is very simple to execute and it makes almost no assumption about the underlying system. The cost of the algorithm is a singular value decomposition (SVD) of the snapshot matrix constructed from the data x_k.

The development of DMD is timely due to the concurrent rise of data science, encompassing a broad range of techniques, from machine learning and statistical regression to computer vision and compressed sensing. Improved algorithms, abundant data, vastly expanded computational resources, and interconnectedness of data streams make this a fertile ground for rapid development.

3.2 Uses and Interpretations

DMD has a number of uses and interpretations. Specifically, the DMD algorithm can generally be thought to enable three primary tasks.(Kutz et al., 2016)

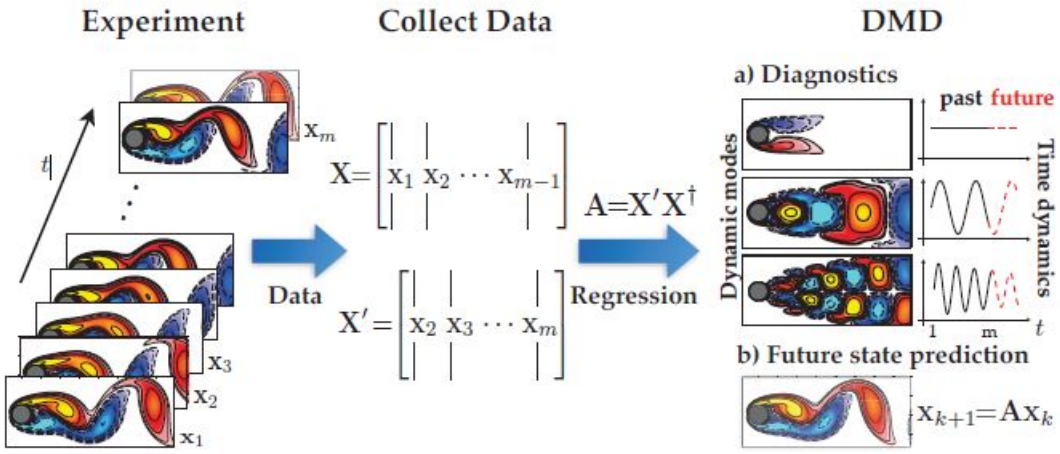

Fig. 3.1 DMD overall picture

- **Diagnostics** At its inception, the DMD algorithm was used as a diagnostic tool to characterize complex fluid flows. In particular, the algorithm extracts key low-rank spatio temporal features of many high-dimensional systems, allowing for physically interpretable results in terms of spatial structures and their associated temporal responses. Interestingly, this is still perhaps the primary function of DMD in many application areas. The diagnostic nature of DMD allows for the data-driven discovery of fundamental, low-rank structures in complex systems analogous to POD analysis in fluid flows, plasma physics, atmospheric modeling, etc.

- **State Estimation and future prediction**: A more sophisticated and challenging use of the DMD algorithm is associated with using the spatio temporal structures that are dominant in the data to construct dynamical models of the underlying processes observed. This is a much more difficult task, especially as DMD is limited to constructing the best-fit (least-square) linear dynamical system to the non-linear dynamical system generating the data. Thus, unlike the diagnostic objective, the goal is to anticipate the state of the system in a regime where no measurements were made. Confounding the regressive nature of DMD is the fact that the underlying dynamics can exhibit multiscale dynamics in both time and space. Regardless, there are a number of key strategies, including intelligent sampling of the data and updating the regression, that allow DMD to be effective for generating a useful linear dynamical model. This generative model approach can then be used for future-state predictions of the dynamical systems and has been used with success in many application areas.

- **Control**: Enabling viable and robust control strategies directly from data sampling is the ultimate, and most challenging, goal of the DMD algorithm. Given that we are using a linear dynamical model to predict the future of a non-linear dynamical system, it is reasonable to expect that there is only a limited, perhaps short-time, window in the future where the two models will actually agree. The hope is that this accurate prediction window is long enough in duration to enable a control decision capable of influencing the future state of the system. The DMD algorithm in this case allows for a completely data-driven approach to control theory, thus providing a compelling mathematical framework for controlling complex dynamical systems whose governing equations are not known or are difficult to model computationally

3.3 Defining DMD

The DMD method provides a spatiotemporal decomposition of data into a set of dynamic modes that are derived from snapshots or measurements of a given system in time. The data collection process involves two parameters:

$$n = number of spatial points saved per time snapshot,$$

$$m = number of snapshots taken$$

Tu and Rowley (2014) provides the most moder definition of the DMD method and algorithm.

Definition: Dynamic Mode Decomposition Suppose we have dynamical system and two sets of data,

$$X = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_{m-1} \\ | & | & & | \end{bmatrix}$$

,

$$X' = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_n \\ | & | & & | \end{bmatrix}$$

, so that $x_k^{'}$ is generated data. DMD computes the leading eigendecomposition of the best-fit linear operator A relating the data X'$\approx$ AX:

$$A = X'X^T$$

The DMD modes, also called dynamic modes are the eigenvectors of A, and each DMD mode corresponds to a particular eigenvalue of A. The SVD approach, slightly modified by Tu and Rowley (2014) is now known as exact DMD algorithm.

3.4 Exact DMD Algorithm

This section adapts the approach mentioned in(Brunton et al., 2016) and (Tu and Rowley, 2014) to explain the DMD algorithm. Let x_k represents the general flow field data (a snapshot) at k^{th} time step. Then it is possible to form a sequential set of data vectors

$$\{x_0, x_1, ..., x_m\}$$

where each x_k was sampled at discrete time interval ($\triangle t$) from a continuously evolving system x(t),i.e., x_k=x(k$\triangle t$).

We assume there exists an unknown linear dynamics that approximate the data generated, be it from linear or a non linear system. This assumption of linear mapping A from one snapshot to other is critical, because usually the data observed from experiments are produced by a non-linear process. Still, DMD is able to capture the dominant dynamics from the experimental data. We can write:

$$x_{k+1} \approx Ax_k$$

Extending this to all snapshots, we get

$$\{x_1, x_2, ..., x_m\} \approx A\{x_0, x_1, ..., x_{m-1}\}$$

For simplicity we form data matrices X=$\{x_0, x_1, ..., x_{m-1}\}$ and X'=$\{x_1, x_2, ..., x_m\}$. It becomes

$$X' = AX$$

The next step involves the computation of reduced SVD of X to get

$$X = U\Sigma V^*$$

Now, defining the matrix $\tilde{A}$ by projecting the least square fit A onto POD modes U

$$\tilde{A} = U^*AU = U^*X'V\Sigma^{-1}$$

Computing the eigenvalues and eigenvectors of $\tilde{A}$, we get

$$\tilde{A}W = W\lambda$$

where λ are the DMD eigen values. The harmonically averaged spatial mode, called as DMD mode(Φ) can be obtained as

$$\Phi = X'V\Sigma^{-1}W$$

Each column of Φ is a DMD mode ϕ_i corresponding to the eigenvalue λ_i. Each pair (ϕ, λ) obtained through this algorithm is an eigenvector-eigenvalue of A. This algorithm produces modes with unit norm. To obtain the DMD spectrum, it is necessary to scale the modes and select the ones with highest power/energy. The following section modifies the algorithm to obtain this appropriate scaling.

3.5 DMD spectrum

From (Brunton et al., 2016), we note that the singular values from SVD can itself be used to get the required scaling. After computing $\tilde{A}$, we compute

$$\widehat{A} = \Sigma^{-1/2}\tilde{A}\Sigma^{1/2}$$

The eigen-decomposition of $\widehat{A}$ yields

$$\widehat{A}\widehat{W} = \widehat{W}\lambda$$

where $\widehat{W}$ contains the eigenvector and the λ the eigen values in its diagonal. Note that the eigen values of $\widehat{A}$ and $\tilde{A}$ are identical[1] and

$$W = \Sigma^{1/2}\widehat{W}$$

Now, we can obtain the modes via equation of Φ in the previous section using the W obtained above and they will not be of unit norm. This norm is a measure of the magnitude of the mode. The DMD spectrum is now obtained by computing the frequency of the mode

[1] This is because $\tilde{A}$ and $\widehat{A}$ are similar matrices and similar matrices have identical eigen values and the eigen vector follow the rule: If B = $M^{-1}AM$ and x is an eigenvector of A, then $M^{-1}x$ is an eigenvector of B = $M^{-1}AM$

from the phase of the eigenvalues .

$$f_i = \frac{imag(log(\lambda_i)/\triangle t)}{2\pi}$$

and the power of the mode from the norm of the mode given by

$$P_i = ||\phi_i||^2$$

Algorithm to find DMD : For the sequence of m+1 snapshots $\{x_0, x_1, ...x_m\}$ sampled at equal time interval $\triangle t$, finding the dynamic modes Φ_j and corresponding frequency f_j for j=1,....,m

1. $X \leftarrow \{x_0, x_1, ...x_{m-1}\}$
2. $X' \leftarrow \{x_1, x_2, ...x_m\}$
3. $[U, \Sigma, V] \leftarrow svd(X, 0)$
4. $\tilde{A} \leftarrow U^* X' V \Sigma^{-1}$
5. $\widehat{A} \leftarrow \Sigma^{-1/2} \tilde{A} \Sigma^{1/2}$
6. $[\widehat{W}, \lambda] \leftarrow eig(\widehat{W})$
7. $W \leftarrow \Sigma^{1/2} \widehat{W}$
8. $\Phi \leftarrow X' V \Sigma^{-1} W$
9. $\omega \leftarrow log(\lambda_{jj})/\triangle t$
10. $f_i = \frac{imag(\omega_j)}{2\pi}$

3.6 Limitations of DMD

The DMD method, much like PCA, is based on an underlying SVD that extracts correlated patterns in the data. It is well known that a fundamental weakness of such SVD-based approaches is the inability to efficiently handle invariances in the data. Specifically, translational and/or rotational invariances of low-rank objects embedded in the data are not well captured. Moreover, transient time phenomena are also not well characterized by such methods.

3.7 Difference between POD and DMD

There are several major differences between the two decomposition methods. The spatial basis functions, for DMD and POD offer an insight of the coherent structures in the flow field. The difference between the two basis functions are due to the principles of the decomposition methods. The POD modes are orthogonal in space with the energy inner product. In DMD,each mode oscillates at a single frequency, hence the expression that the DMD modes are orthogonal in time.It is also possible to extract the phase and growth rate information using DMD, which is not possible in case of POD(Bistrian and Navon, 2014).

Chapter 4

Applications and Results

4.1 Helium jet experiment

The experiment performed on helium jet was recorded using flow visualization device to obtain schieleren images at an fps of 60kHz. Thousand images were used for the application of the decomposition methods.

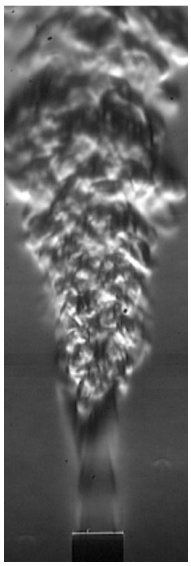

Fig. 4.1 Helium jet image sample

4.2 Application of POD

POD method was applied on the images by taking different areas of the image. Two regions were selected: One with the flow was almost laminar and second image which contains both

laminar as well as turbulent region. It was done to realize the effect of decomposition on different regions of flow. Image at each instant is is stored in the column of matrix A and mean was subtracted from each flow.

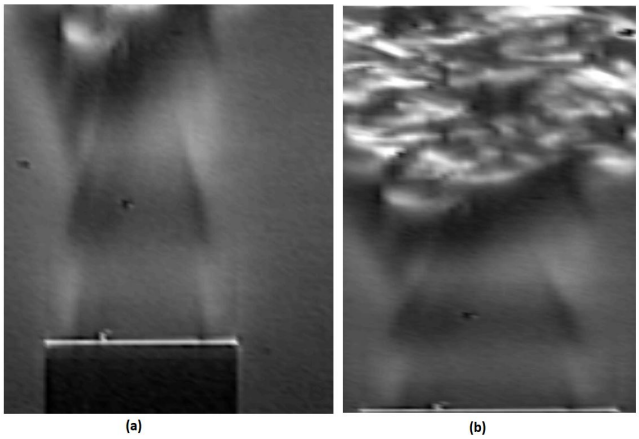

Fig. 4.2 Different regions selected for the analysis (a) laminar region: consisting of mostly laminar flow, and (b) turbulent region: consisting turbulent and laminar flow regime of flow

4.2.1 For laminar region

The POD method was applied for the space consisting laminar flow and the energy distribution was obtained and plotted as shown in Figure 4.3.

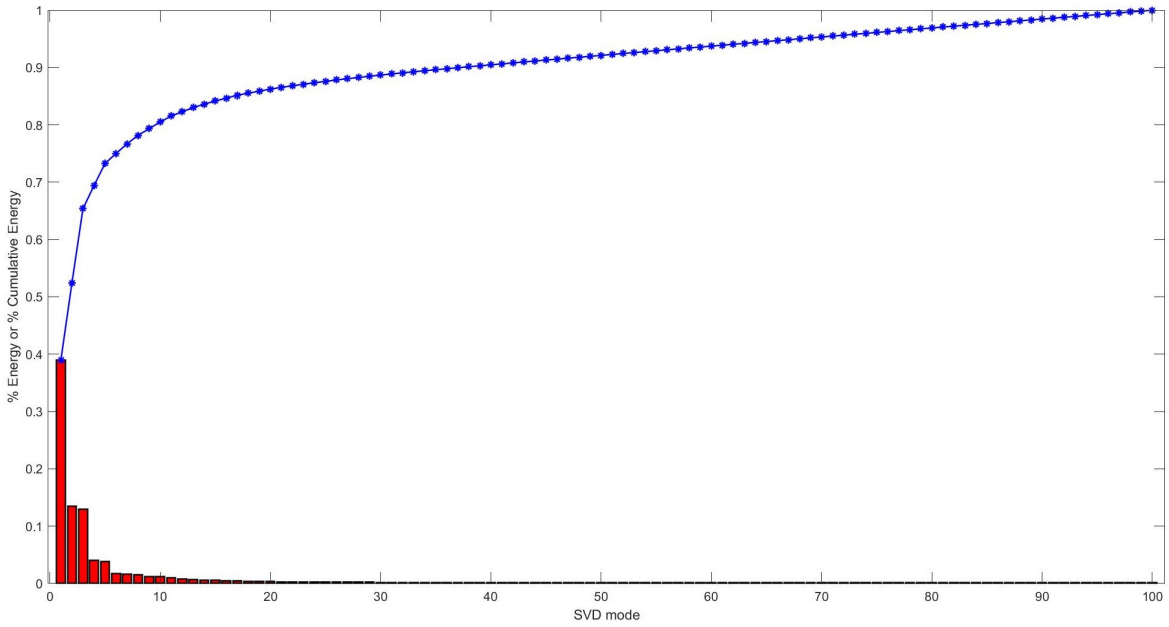

Fig. 4.3 % Modal energy distribution obtained by SVD (where, blue line shows cumulative energy and red bars show individual modal energy

The energy distribution of the singular values shows that the first mode itself contains nearly 40% of the total energy and first 5 modes contain nearly 75% of the total energy. It shows that the data can be well approximated using first few modes only. The concentration of the energy within few modes can be clearly seen and POD method applies very efficiently in this data set.

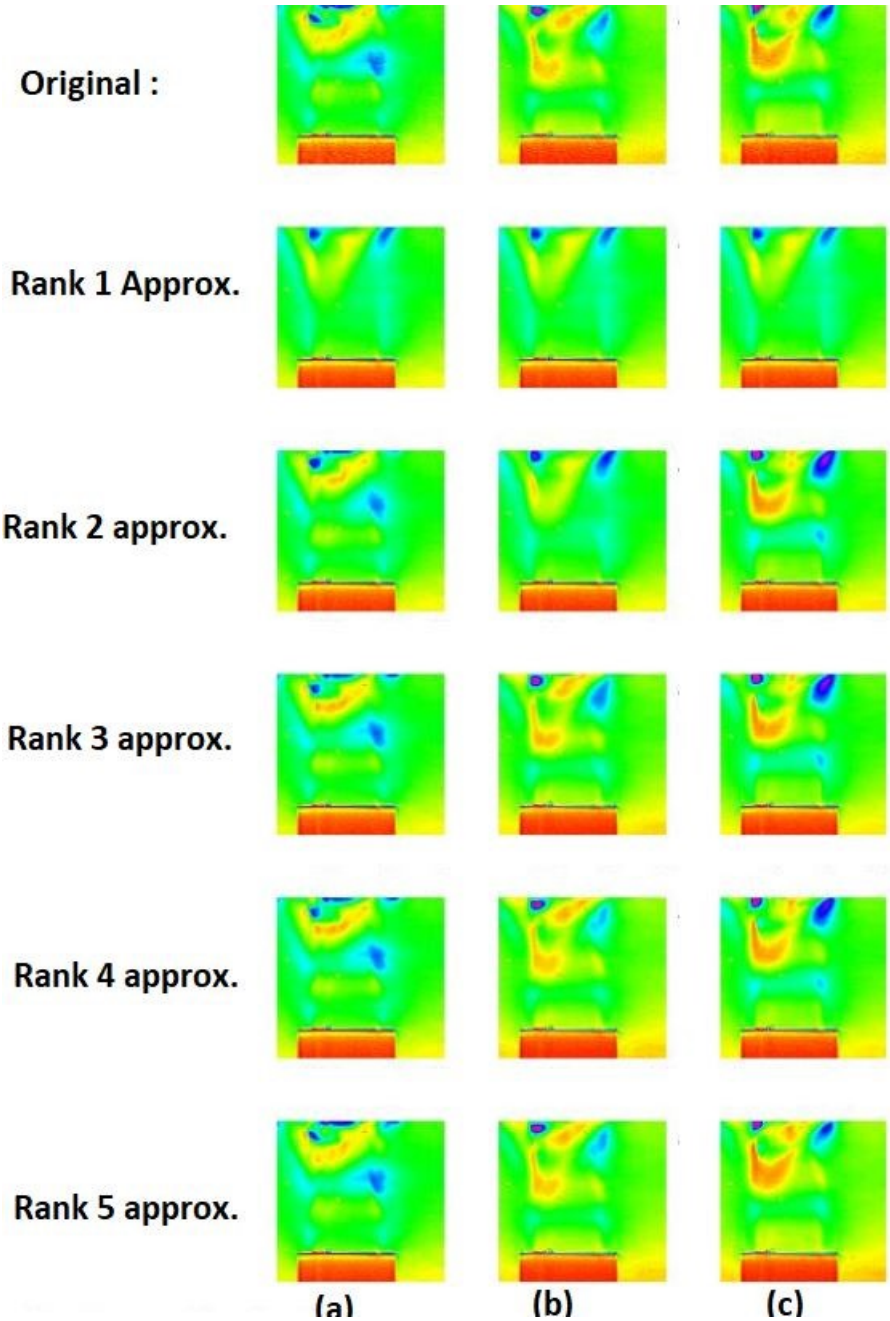

Fig. 4.4 POD results obtained at (a) at t= 1ms, (b) at t=1.2ms, and (c) t=1.4 ms

It can be seen from the approximated images in Figure 4.4 that neither rank 1 and rank 2 can approximate the original flow field completely. At least 3 modes are require to get a good approximation of the data.

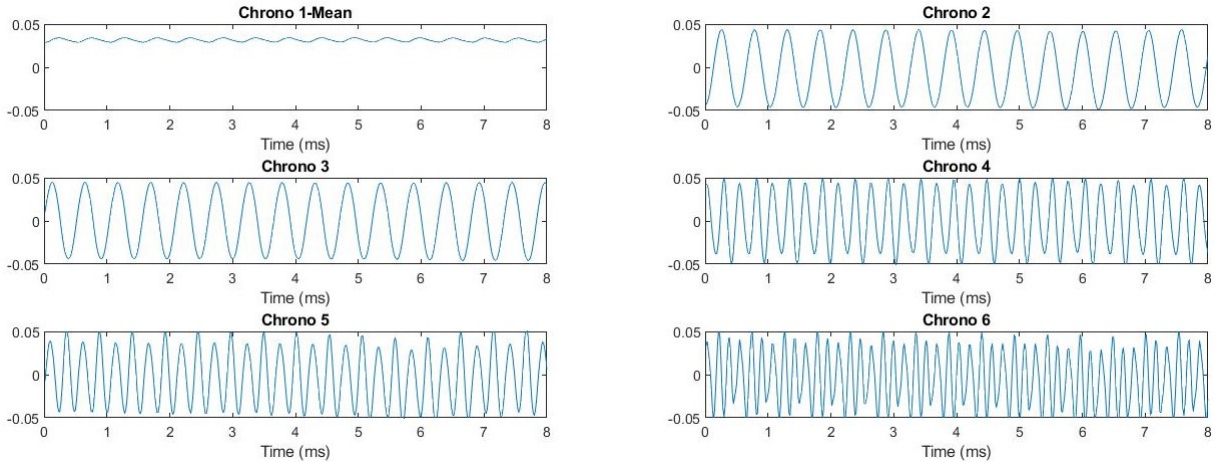

Fig. 4.5 Chronos of the helium jet oscillation (amplitude vs time); chrono 1 represents the mean of the data

Different chronos 1 to 6 were plotted as shown in Figure 4.5. It can be observed from the Figure 4.5 that the chronos 2 and 3 contains the most clear sinusoidal signal, whereas further chronos show superimposition of sinusoidal waves of different harmonics. To obtain better understanding of the chronos, FFT of each chronos were obtained (see Figure 4.6).

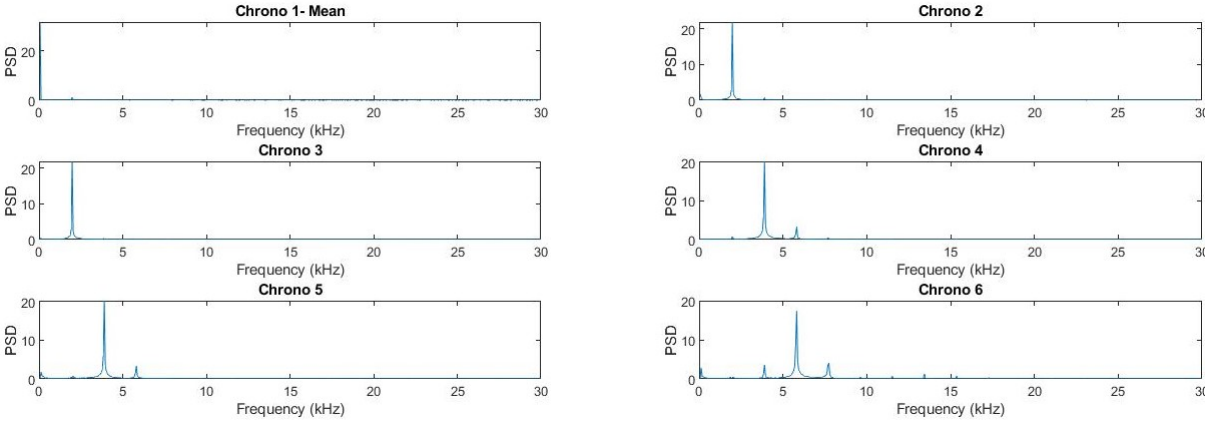

Fig. 4.6 FFT of the Chronos (PSD represents Power Spectral Density)

FFT gives the frequency of the waves present in the chronos-Mode 2 and 3 have peak at 1980 Hz and subsequent modes 4,5,6 have the peak at 3900 Hz and 5820 Hz which shows the super-harmonics of the fundamental frequency (of mode 2 and 3) as they are the nearly integral multiple of it.

To get the further clear insight of the chronos, phase properties of the waves were obtained by finding the Lissajous figures of the chronos as shown in Figure 4.7 .

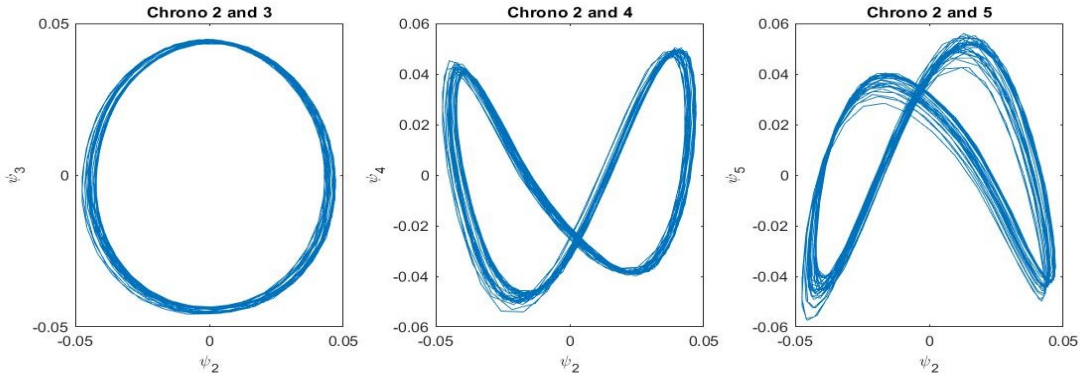

Fig. 4.7 Lissajous figures of different mode pairs

On studying the Lissajous figure, it can be noted that the phase difference between the chrono 2 and 3 is 90^o. Chronos 2 and 4 shows the chronos of the fundamental frequency and its first harmonic.

Other than analysing the result qualitatively, the quantitative analysis of the result was also performed by selecting some points in the flow domain as shown in Figure 4.8. The gray scale intensity variation with time was plotted for the original data and for the approximated data obtained via SVD as shown in the Figures 4.9 and 4.10 .

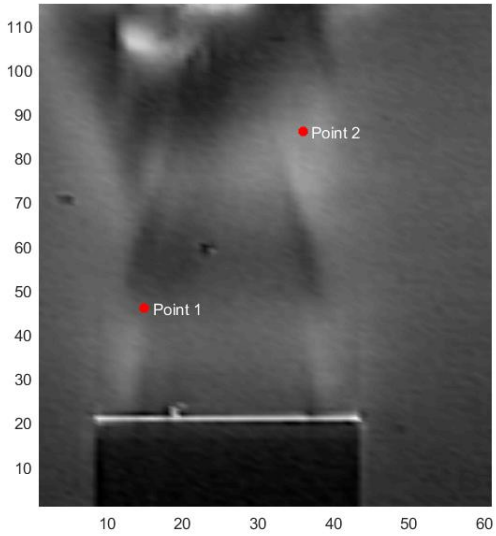

Fig. 4.8 Points selected for quantitative analysis quantitative analysis

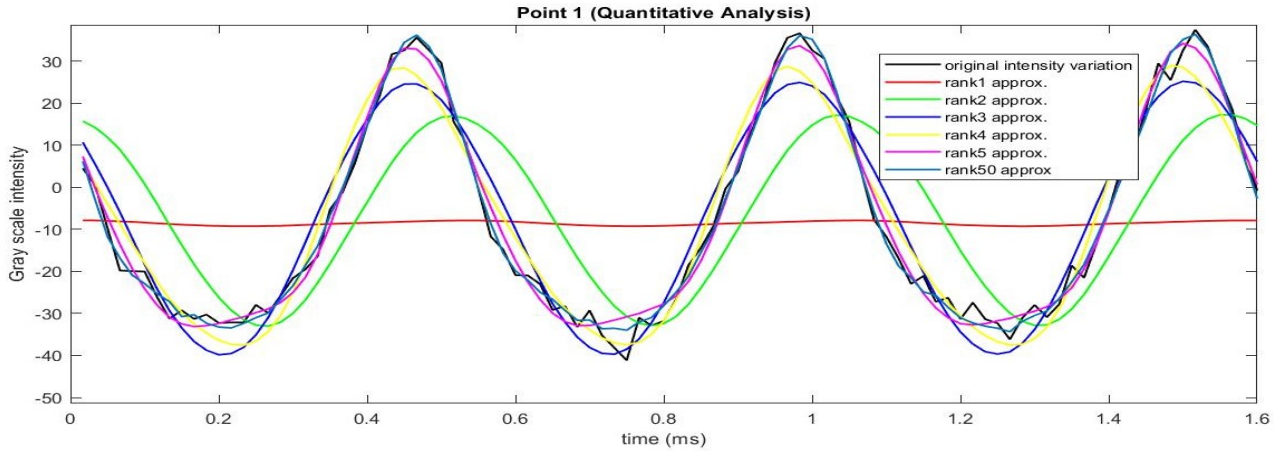

Fig. 4.9 Point 1 Quantitative Analysis

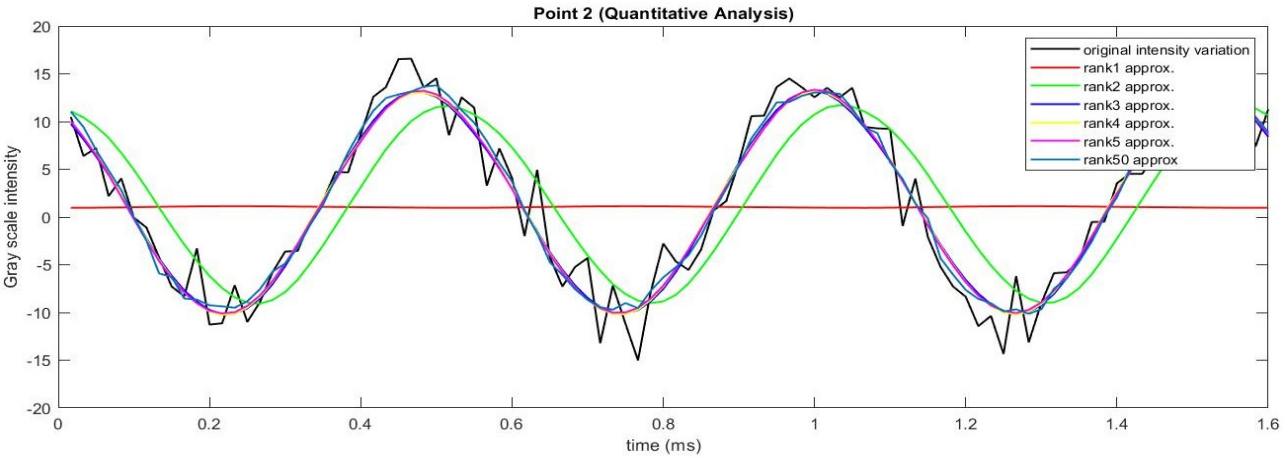

Fig. 4.10 Point 2 Quantitative Analysis

The analysis shows that the approximation becomes better as we increase the rank of the approximation. The quantitative analysis of the point also depends on the fluctuation of the gray scale intensity of the point in the original data. If the fluctuation in the intensity is more random, it will be difficult to approximate using few modes as compared to those points in which the intensity variation is smooth or periodic (the same can be realized by referring to Figures 4.9 and 4.10).

4.2.2 For turbulent region

POD method was also applied to space (Figure 4.2(b)), which contains turbulent flow regime of the flow. The modal energy plot was obtained by performing SVD on the given data set and is shown in Figure 4.11.

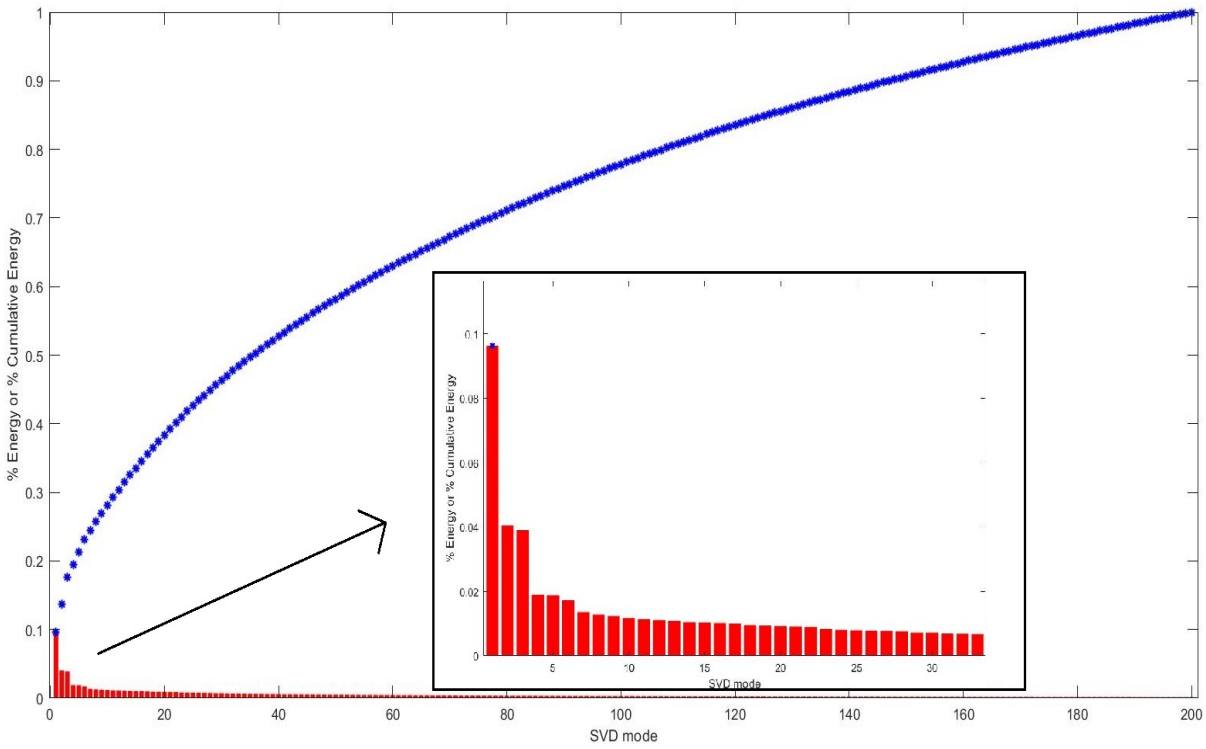

Fig. 4.11 Energy in different modes obtained by SVD

From Figure 4.11, it can be seen that the distribution of energy in modes is not concentrated in few modes as in case of laminar regime. The first mode contains just 10% of the total energy, whereas in case of laminar regime flow, the first mode contains 40% of the energy. Moreover, in this case, 90 modes are required to build up 75% of the total energy, whereas in case of laminar region just 5 modes were able to constitute 75% of the total energy. This shows that as in case of turbulent regime, the fluctuations are more random, it is difficult to approximate the data set through less number of modes as in case of laminar regime. Substantially large number of modes are required to model turbulent region than laminar one (18 times more in this case). The efficiency of POD method decreased drastically as the data loses its structure or randomness(also characteristic of turbulent flows) is induced in the data set. Still, if the analysis of the turbulent flow is seen separately, POD stands out to be an excellent way to generate a low rank approximation or extracting a structure out of a much random flow.

The qualitative aspect of analysing the POD approximated data is shown in Figure 4.12 . The approximated data obtained shows that not even rank 50 approximation is able to approximate the data in a nice manner. The approximation appears to be too much vague for 1,2,3 and 4 rank approximation. This shows that the energy is distributed among modes(which

can also be seen through modal energy plot) and cannot be efficiently represented through few modes as in case of laminar regime data.

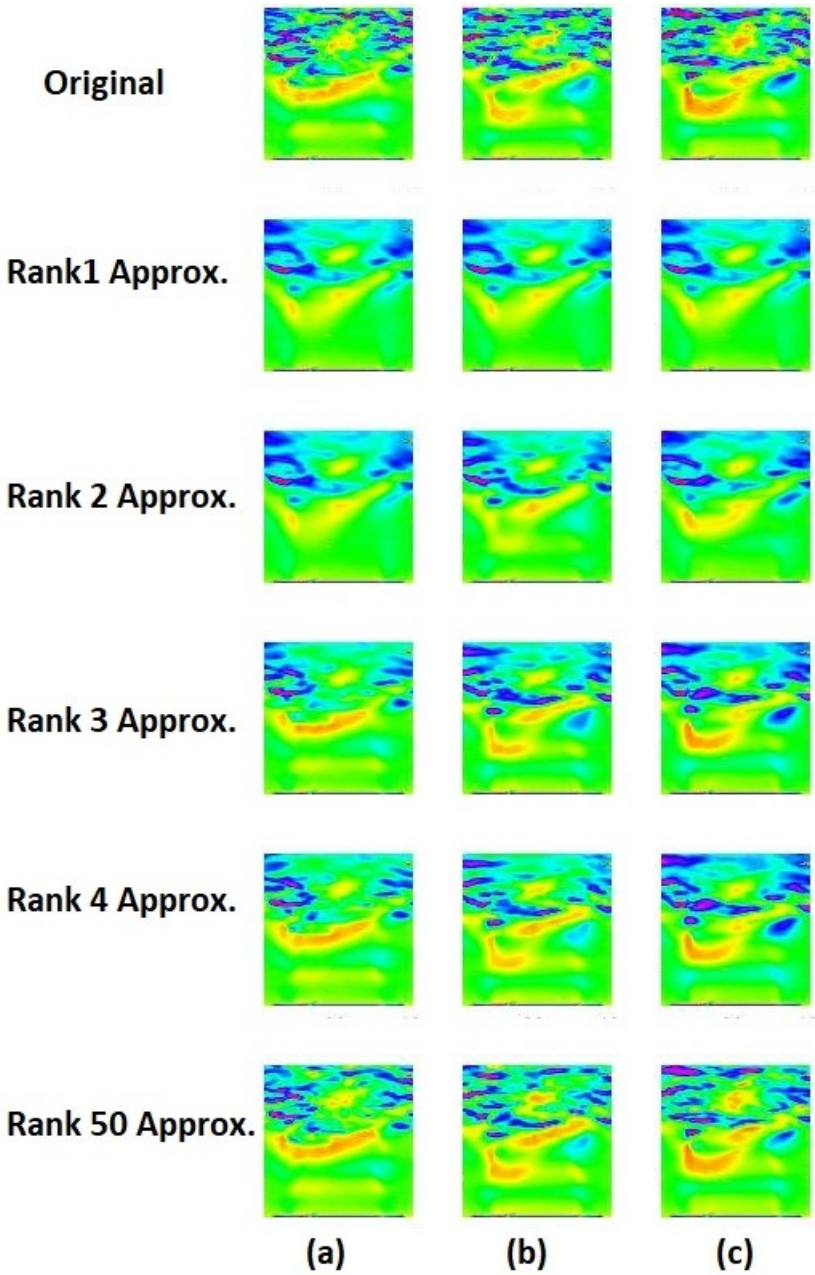

Fig. 4.12 POD results obtained at (a) at t= 1ms, (b) at t=1.2ms, and (c) t=1.4 ms

The chronos from 1 to 6 were also plotted as shown in the Figure 4.13. To get the insight of the chronos, FFT was also performed and results were obtained as shown in Figure 4.14. Mode 2 and 3 shows the peak at 2178 Hz and Mode 4 and 5 shows peak at 4290 Hz. Peak in

mode 4 and 5 are the super-harmonic of the fundamental frequency (mode 2 and 3) as they are nearly equal to its integral multiple.

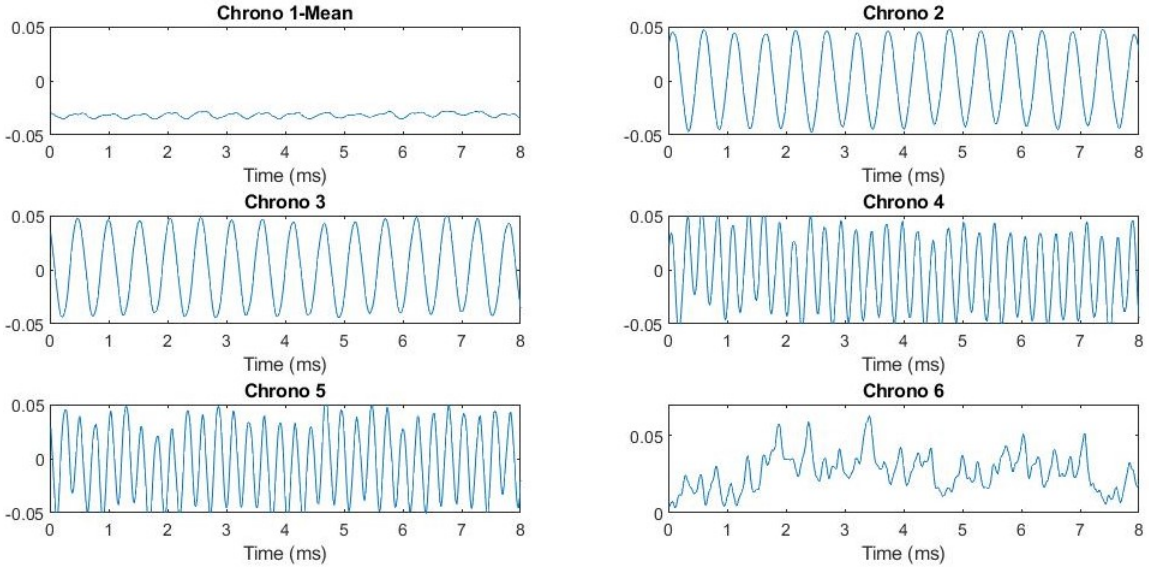

Fig. 4.13 Chronos of the Helium jet (Amplitude Vs Time)

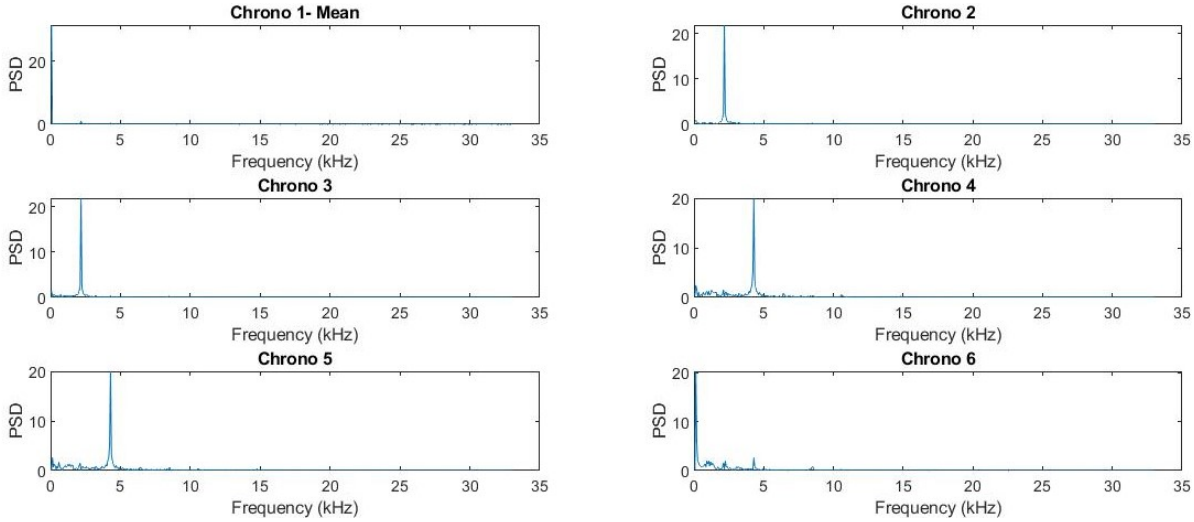

Fig. 4.14 FFT of the Chronos of the Helium jet (Power Spectral Density Vs Frequency)

To know about the phase properties of the chronos, Lissajous figures between different modes were also plotted as shown in Figure 4.15. The phase shift between the mode 2 and 3 is 90^o as depicted by the results.

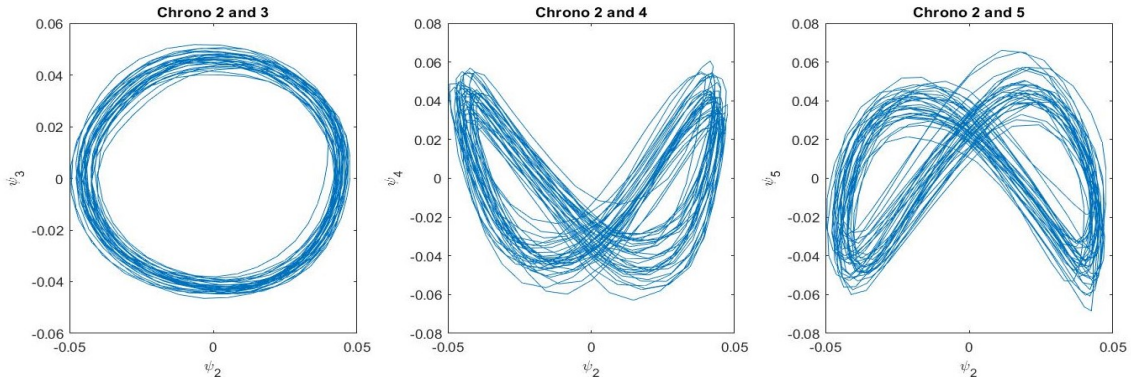

Fig. 4.15 Lissajous figure for the Chronos

Now, in order to validate the results, quantitative analysis was also performed on selected points in the region of analysis as shown in Figure 4.16.

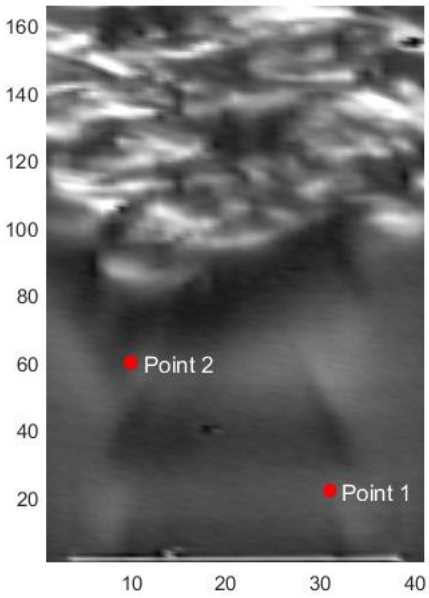

Fig. 4.16 Points selected for the quantitative analysis

The quantitative result shows that the variation in the intensity is originally much random as compared to the laminar case (which is also expected to be as Turbulence implies randomness). To model such randomness, the approximation needs to consist of larger number of modes. It is also evident from the Figure 4.17 and 4.18 that even rank 50 approximation is unable to effectively approximate the data, implying that the energy is much distributed in this case. This shows that as the randomness in the data increases (or no

structure in the flow), the POD method efficiency reduces(as compared to laminar regime data) and is unable to approximate through lower rank matrices. But, if turbulent POD result is standalone analysed, it shows that the rank of the data can be drastically reduced as compared to the original data set rank.

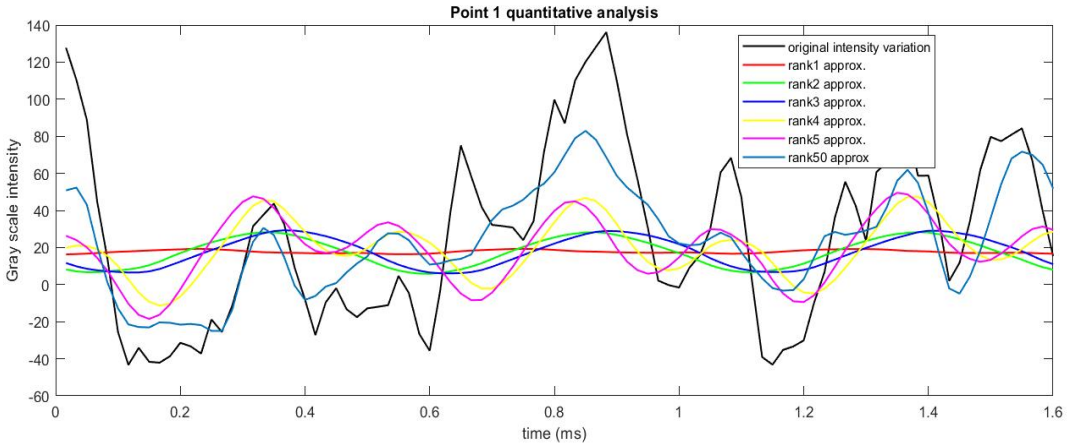

Fig. 4.17 Point 1 Quantitative Analysis

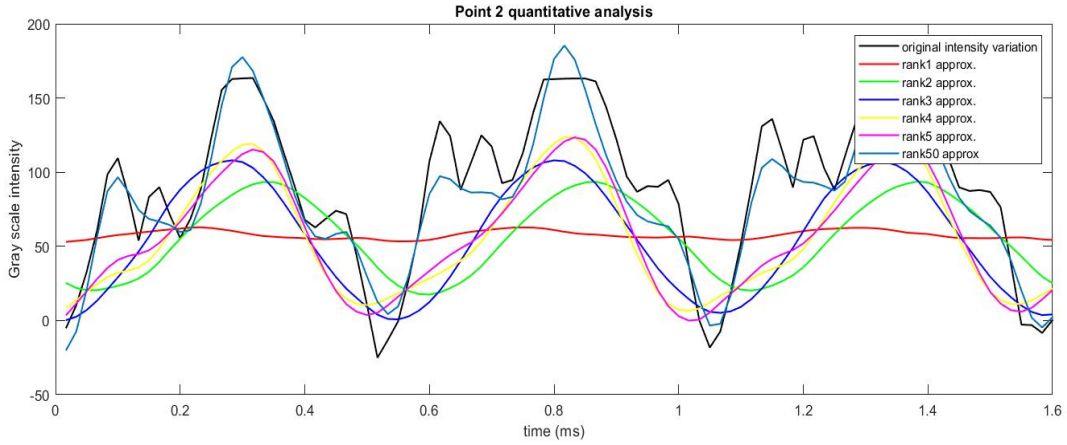

Fig. 4.18 Point 2 Quantitative Analysis

4.3 Local Analysis

For this analysis as well, seperate study was conducted for laminar as well as turbulent flows. To find the intensity of a point, average of the surrounding four plus one itself was taken at each time instant. The points were selected near the shear layer to obtain nearly periodic oscillations.

4.3.1 Laminar Region

The points were selected in the laminar space to realize the oscillations in the flow field(as shown in Figure 4.19) and their gray scale intensity were plotted as shown in figure 4.20.

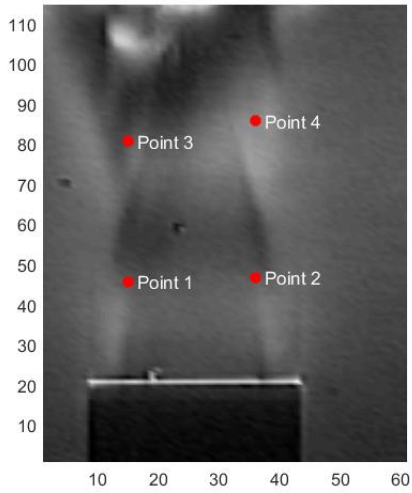

Fig. 4.19 Points selected for local analysis

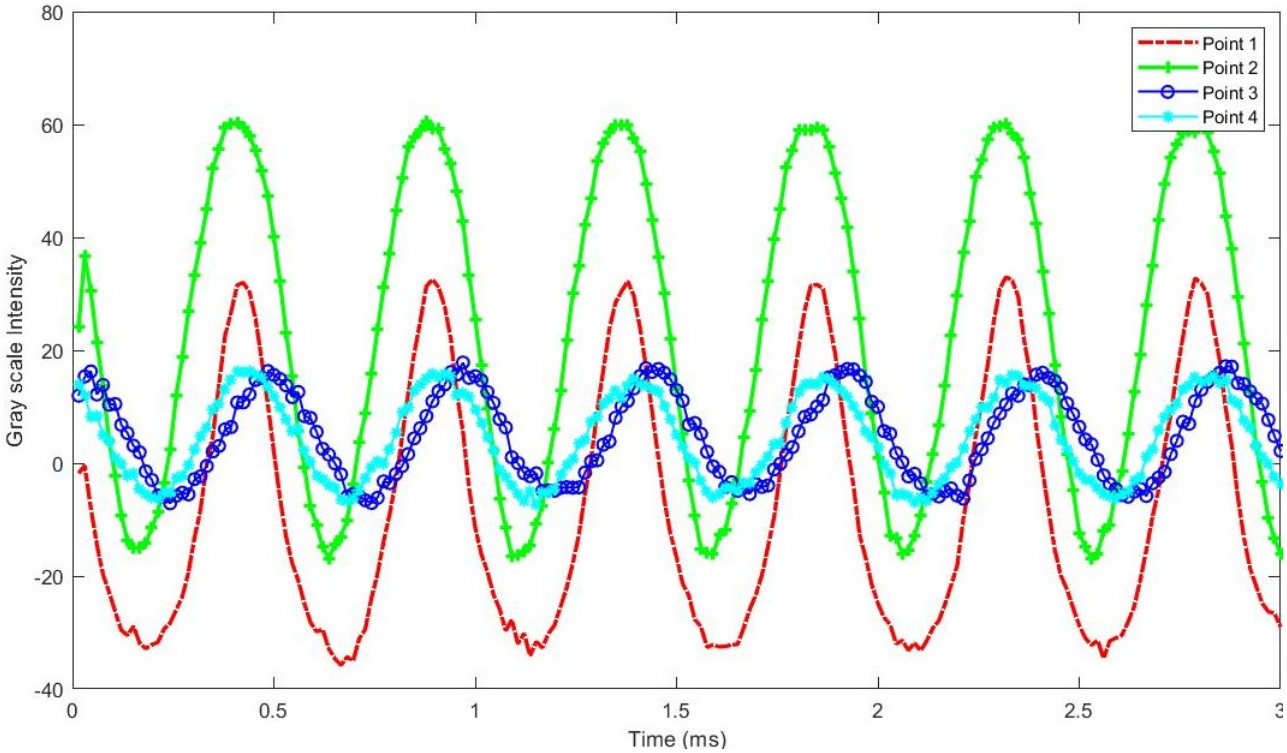

Fig. 4.20 Gray scale intensities of the selected points

FFT of the intensities was done to obtain the dominant frequencies and plotted as shown in Figure 4.21. The fundamental frequency obtained is 1980 Hz, which matches closely with the fundamental frequencies obtained earlier. As the local analysis is not able to present the complete idea of the flow field and the dominant structures in the flow, global analysis was performed in earlier subsections.

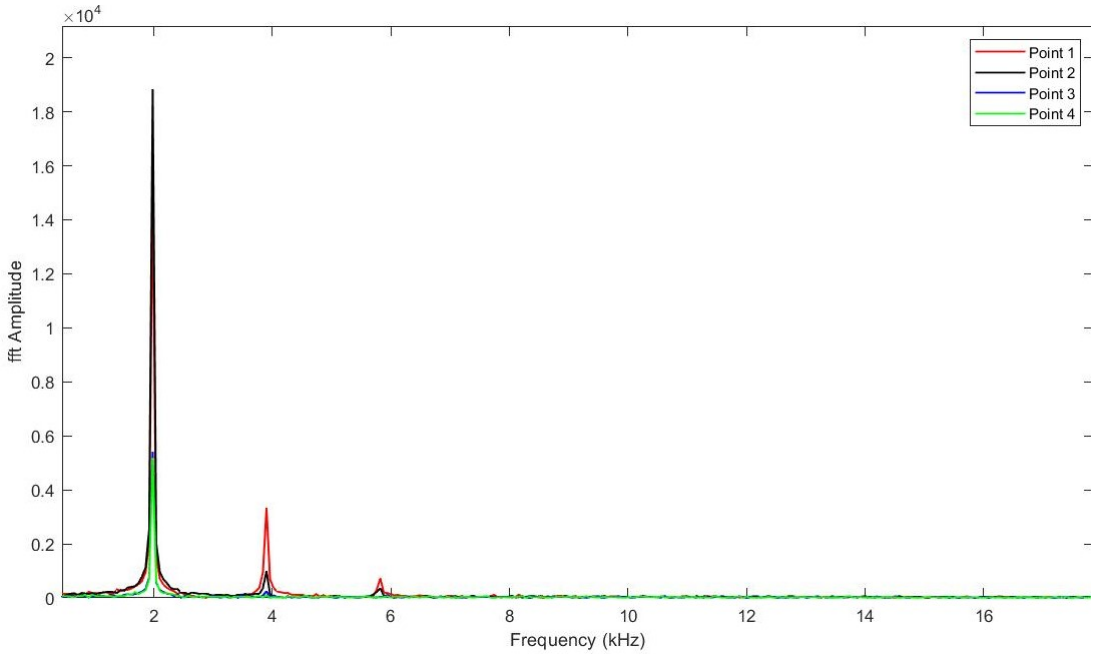

Fig. 4.21 FFT of the intensities of the points selected for analysis

4.3.2 Turbulent Region

The points selected for the local analysis are shown in Figure 4.22, which includes points in both laminar as well as turbulent region. The variation in gray scale intensity of the point with time is shown in Figure 4.23.

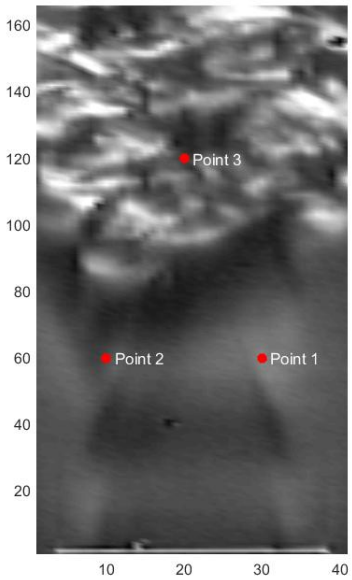

Fig. 4.22 Points selected for analysis

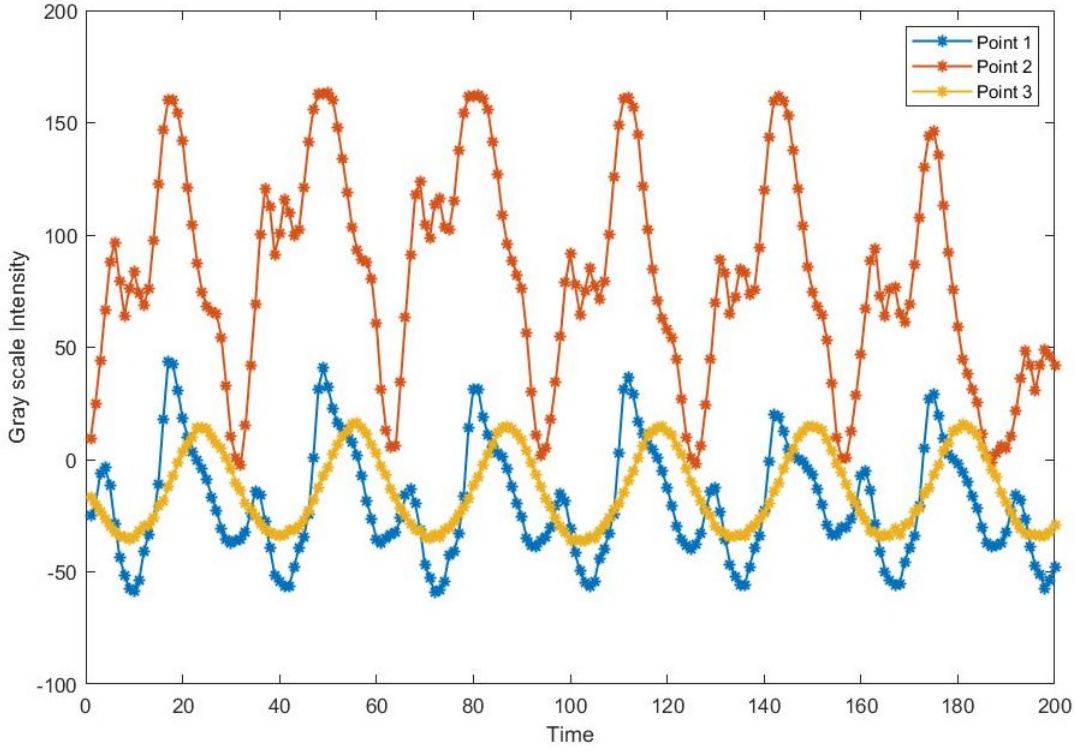

Fig. 4.23 Gray scale intensities of the points selected for analysis

The fft of the gray scale intensities of the point was performed and the fundamental frequency was obtained to be 1980 Hz. The frequency from the local analysis of both the regions was found to be identical. This implies that the local analysis does not gives proper insight into the data, hence, there is a need for the global analysis. Hence, the global analysis was performed on the data.

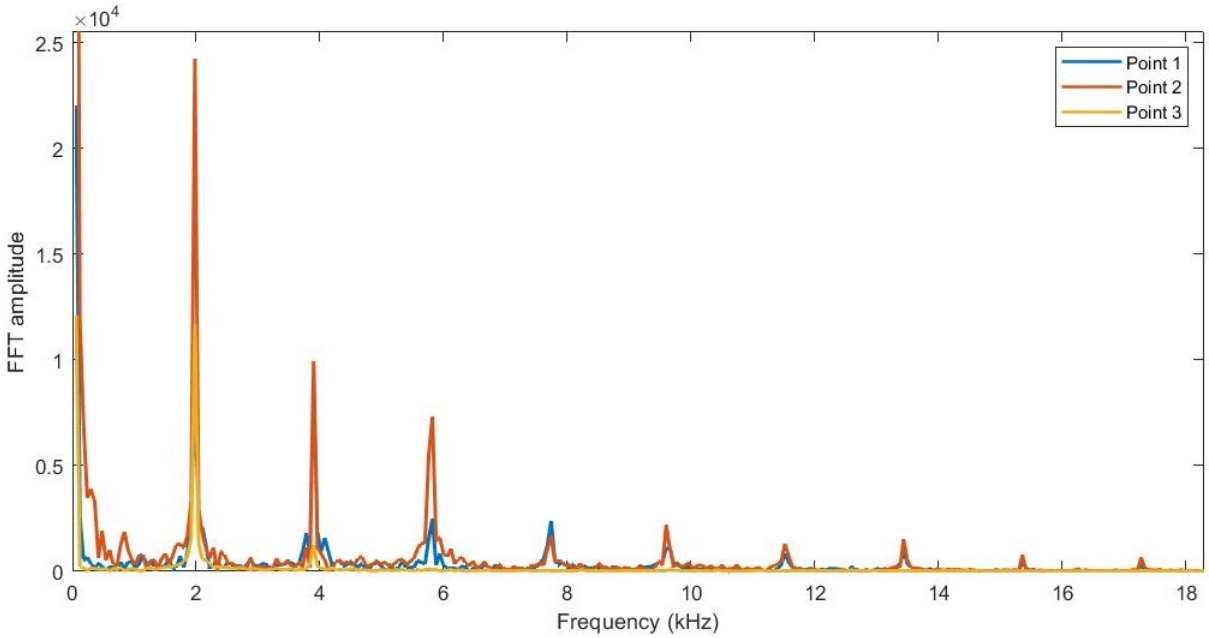

Fig. 4.24 FFT of Gray scale intensities of the points selected for analysis

4.4 Application of DMD

The following section presents the application of DMD on the space as shown in Figure 4.25.

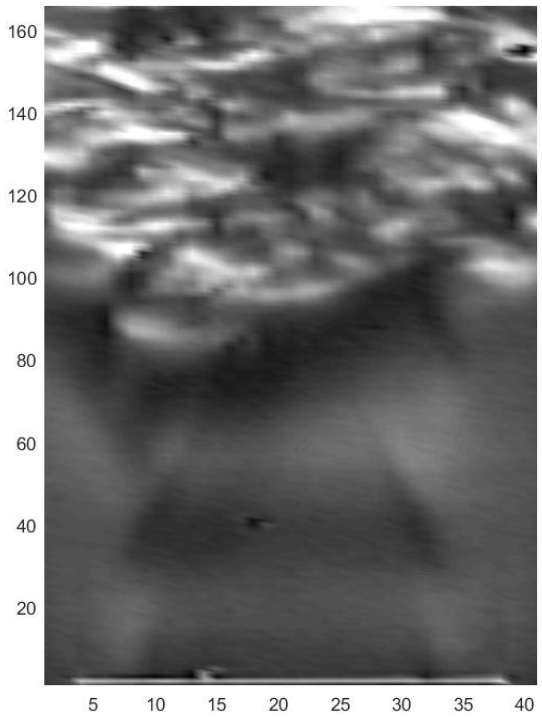

Fig. 4.25 Space used for the application of DMD

The region as shown in Figure 4.25 was extracted from the schlieren images on which DMD was applied. A total of 1000 schlieren images (recorded at an fps of 60000 Hz) were chosen based on the convergence obtained in the eigenvalues which is discussed further. As shown in Figure 4.26(a), the DMD eigenvalues λ can be visualised on the complex plane with respect to the unit circle. This gives the information of the stable, growing and the decaying modes. If the eigen value lies within the unit circle, the mode will decay as the time increases; if the eigen value lies on the circle, the mode will remain neutral and will neither decay or grow as time increases; and if the value lies outside the unit circle, the mode will grow and become unstable. The figure 4.26(b) shows an alternative visualisation of the obtained DMD eigenvalues in the complex ω plane.

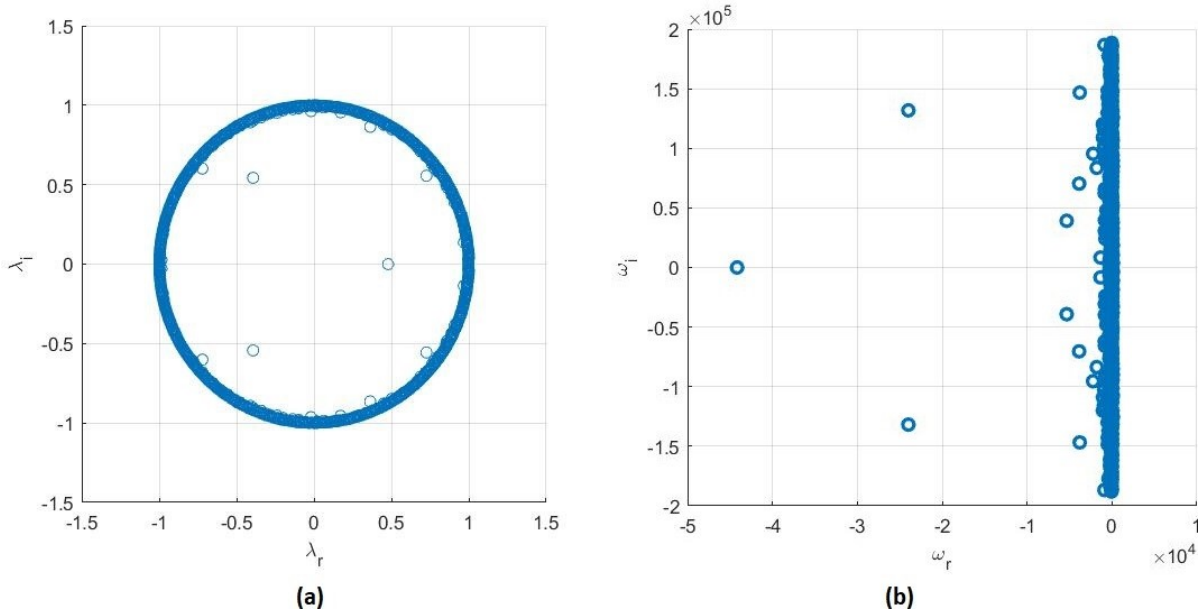

Fig. 4.26 Results obtained from DMD (a) shows the value of eigen values in the complex plane, and (b) shows the ω value in the complex plane

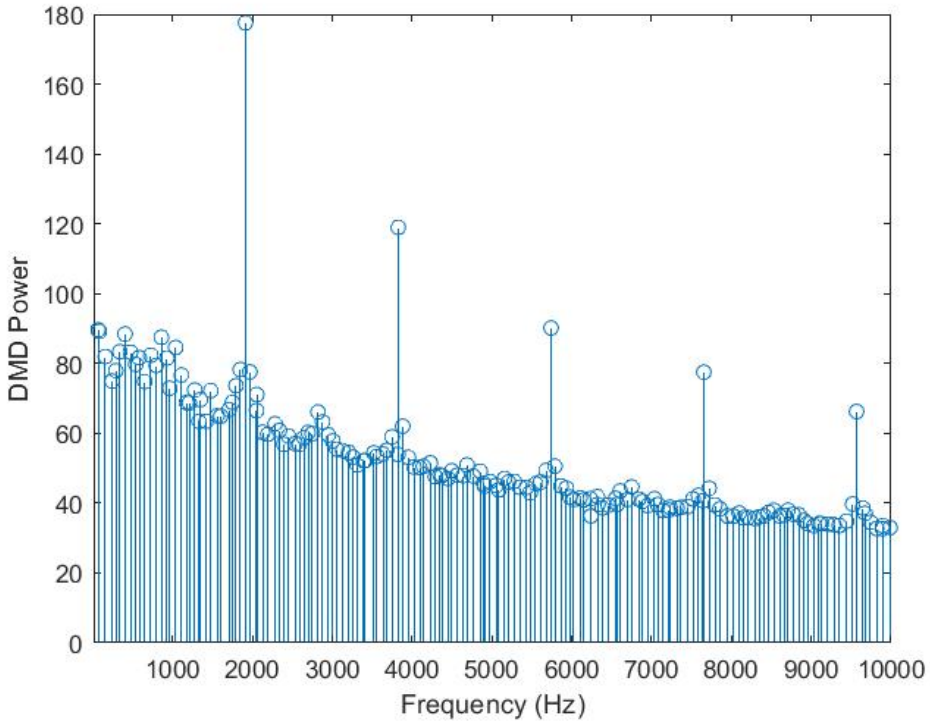

Fig. 4.27 DMD spectrum obtained via algorithm

DMD spectrum is obtained by plotting the “power” of each mode against its frequency obtained from the imaginary part of ω. The highest power DMD mode occurs at 1913 Hz, second highest DMD mode occurs at 3826 Hz and third highest mode occurs at 5740 Hz.

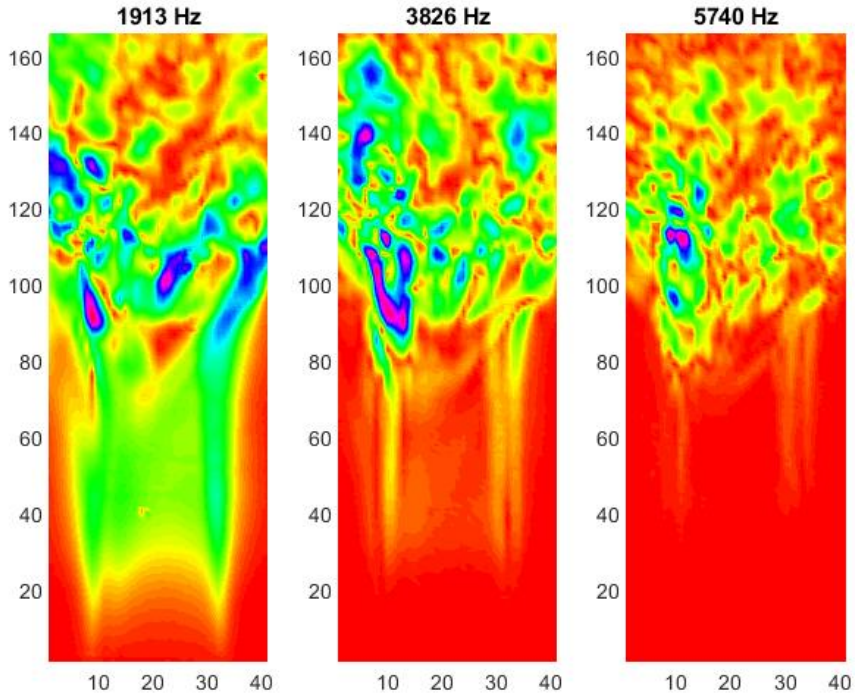

Fig. 4.28 Dominant DMD modes of the flow field

Chapter 5

Results and Conclusion

The main objective of this project was to obtain the clear understanding of the data-based data decomposition methods. The project was initiated by learning the basic concepts of Linear Algebra and Scientific computing (using MATLAB mainly) by (Strang, 2003) and **?** respectively. The data processing techniques mentioned earlier were coded including SVD based on basic eigen-value decomposition approach and DMD based on SVD. These algorithms were used to approximate the data through lower rank matrices. These methods were completely data-based rather than model based. A detailed information regarding the formulation of these algorithms have been presented in the chapters. POD is an energy based decomposition, ranking the flow structures based on second-order statistics. Although POD is able to decompose the dynamics into Spatial and Temporal orthogonal modes, it mixes up the frequency loosing the valuable phase information. It was also observed that when a system dynamics is dominated by a single frequency oscillation, POD produces 2 topos-chronos pair which are quadrature (90^o) shifted to each other. Combination of these 2 modes will then represent the dominant oscillatory motion.

The limitations of the POD were removed by introduction of DMD. DMD technique which decomposes the flow based on the frequency content. DMD supplied the dominant spatial mode called the dynamic mode. It was also noticed that the eigen values of the modes determine the stability of the system.

The application of the decomposition methods on helium jet in the different regimes of the flow gives the better understanding of the flow. It was observed that when the region of analysis was taken to be laminar, the modal energy distribution was concentrated in few modes, whereas the turbulent region shows the distribution of energy to be wider among the modes. It shows that the flow having certain pattern (like laminar flow) have modes with higher energy, whereas as the randomness increases, the flow loses its structure and makes it difficult to be approximated using few modes. Further, the quantitative analysis shows the

same result as obtained qualitatively, verifying the results. The DMD applied on the data was ale to extract dominant frequency from the flow and was able to predict the stability of the flow.

The future scope lies in understanding the decomposition methods to identify coherent structures and extract dynamically relevant processes which quantify the flow behaviour. Newer algorithms like Optimised DMD, Optimal Mode Decomposition and independent modal analysis and many such concepts are yet to be studied.

References

Algazi, V. and Sakrison, D. (1969), 'On the optimality of the karhunen-loève expansion (corresp.)', *IEEE Transactions on Information Theory* **15**(2), 319–321.

Andrews, C., Davies, J. and Schwarz, G. (1967), 'Adaptive data compression', *Proceedings of the IEEE* **55**(3), 267–277.

Berkooz, G., Holmes, P. and Lumley, J. L. (1993), 'The proper orthogonal decomposition in the analysis of turbulent flows', *Annual review of fluid mechanics* **25**(1), 539–575.

Bistrian, D. and Navon, I. (2014), 'Comparison of optimized dynamic mode decomposition vs pod for the shallow water equations model reduction with large-time-step observations', *International Journal for Numerical Methods in Fluids* pp. 1–25.

Brunton, B. W., Johnson, L. A., Ojemann, J. G. and Kutz, J. N. (2016), 'Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition', *Journal of neuroscience methods* **258**, 1–15.

Chatterjee, A. (2000), 'An introduction to the proper orthogonal decomposition', *Current Science* **78**(7), 808–817.

Chen, K. K., Tu, J. H. and Rowley, C. W. (2012), 'Variants of dynamic mode decomposition: boundary condition, koopman, and fourier analyses', *Journal of nonlinear science* **22**(6), 887–915.

Eckart, C. and Young, G. (1936), 'The approximation of one matrix by another of lower rank', *Psychometrika* **1**(3), 211–218.

Gay, D. and Ray, W. (1986, 1988), 'Application of singular value methods for identification and model based control of distributed parameter systems', *IFAC Proceedings Volumes* **21**(4), 95–102.

Holmes, P., Lumley, J. L. and Berkooz, G. (1998), *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press.

Karhunen, K. (1946), 'Zur spektraltheorie stochastischer prozesse', *Ann. Acad. Sci. Fennicae, AI* **34**.

Kerschen, G., Golinval, J.-c., Vakakis, A. F. and Bergman, L. A. (2005), 'The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview', *Nonlinear dynamics* **41**(1-3), 147–169.

Kirby, M. and Sirovich, L. (1990), 'Application of the karhunen-loeve procedure for the characterization of human faces', *IEEE Transactions on Pattern analysis and Machine intelligence* **12**(1), 103–108.

Kosambi, D. D. (1943), 'Statistics in function space', *Journal of the Indian Mathematical Society* **7**, 76–88.

Kutz, J. N., Brunton, S. L., Brunton, B. W. and Proctor, J. L. (2016), *Dynamic mode decomposition: data-driven modeling of complex systems*, SIAM.

Loève, M. (1945), 'Nouvelles classes de lois limites', *Bulletin de la Société Mathématique de France* **73**, 107–126.

Lumley, J. (1967), 'The structure of inhomogeneous turbulence. atmospheric turbulence and wave propagation', *AM Yaglom, VI Tatarski* pp. 166–178.

Mezić, I. (2013), 'Analysis of Fluid Flows via Spectral Properties of the Koopman Operator', *Annual Review of Fluid Mechanics* **45**(1), 357–378.

Monin, A. (n.d.), 'Basic laws of turbulent mixing in the surface layer of the atmosphere'.

Muller, N., Magaia, L. and Herbst, B. (2004), 'Singular value decomposition, eigenfaces, and 3d reconstructions', *Society for Industrial and Applied Mathematics* **46**, 518–545.

Papoulis, A. (1965), 'Probability random variables', *Stochastic Proc* .

Preisendorfer, R. (1988), 'Principal component analysis in meteorology and oceanography', *Elsevier Sci. Publ.* **17**, 425.

Pugachev, V. S. (1953), 'The general theory of correlation of random functions', *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya* **17**(5), 401–420.

Rosenfeld, A. and Kak, A. (1982), 'Digital image processing, vol. 2'.

Schmid, P. J. (2010), 'Dynamic mode decomposition of numerical and experimental data', *Journal of Fluid Mechanics* **656**, 5–28.

Sirovich, L. (1987), 'Turbulence and the dynamics of coherent structures. i. coherent structures', *Quarterly of applied mathematics* **45**(3), 561–571.

Sirovich, L. (1989), 'Chaotic dynamics of coherent structures', *Physica D: Nonlinear Phenomena* **37**(1-3), 126–145.

Strang, G. (2003), *Introduction to Linear Algebra*, Wellesley-Cambridge Press.

Tropea, C. and Yarin, A. L. (2007), *Springer handbook of experimental fluid mechanics*, Springer Science & Business Media.

Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L. and Kutz, J. N. (2014), 'On Dynamic Mode Decomposition: Theory and Applications', *Journal of Computational Dynamics* **1**(2), 391–421. arXiv: 1312.0041.

Tu, J. and Rowley, D. (2014), 'Clarence w. andluchtenburg, sl brunton, jn kutz, on dynamic mode decomposition: Theory and applications', *Journal of Computational Dynamics* **1**, 391–421.