UNIVERSITY OF CALIFORNIA,
IRVINE

Dynamical Feature Extraction of Atomization Phenomena Using Deep Koopman Analysis

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

In Mechanical and Aerospace Engineering

by

Scott Benjamin Leask

Dissertation Committee:
Professor Scott Samuelsen, Chair
Professor Vincent McDonell
Professor Pierre Baldi

2021

DEDICATION

F&F

TABLE OF CONTENTS

Page

LIST OF FIGURES

NOMENCLATURE

Roman Symbols

$\mathbf{b}$	Bias vector
d	Diameter (m)
D_{ij}	Mean diameter (μm)
f	Frequency (Hz)
$\mathbf{g}$	Observable function
$\mathbf{K}$	Koopman operator approximation
$\mathcal{K}$	Koopman operator
$\mathcal{L}$	Loss
Oh	Ohnesorge number
Re	Reynolds number
St	Stokes number
T	Time steps for prediction
u	Velocity (m/s)
$\mathbf{w}$	Vector of weights
We	Weber number
$\mathbf{x}$	State vector

Greek Symbols

α	Learning rate or constant
Δt	Time step or interval
λ	Complex eigenvalue or Ritz value
μ	Eigenvalue real part, or dynamic viscosity (Pa·s)
ρ	Density (kg/m^3)
σ	Activation function, or surface tension (N/m)
ϕ	Koopman eigenfunctions
ω	Eigenvalue imaginary part

Subscripts

∞	gas phase
l	liquid phase
s	Sampling frequency
ψ	Modal frequency

Acronyms

BPOD	Balanced POD
CFD	Computational fluid dynamics
CKN	Convolutional Koopman network
CNN	Convolutional neural network
DFT	Discrete Fourier transform
DMD	Dynamic mode decomposition
DMD-NN	DMD neural network
DNS	Direct numerical simulation
EDMD	Extended DMD
GAN	Generative adversarial network
LIF	Laser-induced fluorescence
MSE	Mean squared error
NN	Neural network
PCA	Principal component analysis
PDI	Phase Doppler interferometry
PIV	Particle image velocimetry
POD	Proper orthogonal decomposition
POD-NN	POD neural network
PSD	Power spectral density
ReLU	Rectified linear unit
RNN	Recurrent neural network
SMD	Sauter mean diameter
SPOD	Spectral POD
SVD	Singular value decomposition

Glossary

Big data	Large datasets which are amenable to data-driven techniques
Dimensionality reduction	Identifying a low-dimensional subspace which captures the dominant structures or trends of the full-dimensional data
Generalization	The ability of a model to perform successfully on unseen test data
Koopman modes	Projection of the observable function onto the Koopman eigenfunctions; captures the system coherent structures
Observables	The variables whose dynamics evolve linearly via the observable function, $\mathbf{g}$
Representational capacity	The size of the set of all learnable functions of a given model

ACKNOWLEDGEMENTS

I am tremendously grateful to both Professor Samuelsen and Professor McDonell, for all of your support, guidance, and discussion. In particular, I am especially thankful for the freedom to investigate challenging and fruitful ideas and the journey that entailed.

Thank you to Professor Baldi for agreeing to serve on my dissertation committee, for the insightful academic and career discussions, and for the matches.

Thank you to my friends, family, and colleagues.

CURRICULUM VITAE

EDUCATION

University of California, Irvine, USA *June 2017 – May 2021*
Ph.D – Mechanical and Aerospace Engineering GPA: 3.99/4.00

University of Glasgow, UK *Sept 2012 – June 2017*
Master of Engineering, Aeronautical Engineering ***Honors of the First-Class***

RESEARCH EXPERIENCE

Graduate Student Researcher | UCI Combustion Laboratory *June 2017 – May 2021*

Ph.D focus: **Nonlinear modal extraction of optical spatiotemporal data for fuel injection analysis.**

Computer Vision | Koopman theory-inspired deep learning for nonlinear modal extraction; novel deep 2D-/3D-CNN autoencoder with auxiliary network to extract continuous spectra and improve nonlinear future state prediction [**J4**, **J6**].

Injector Analysis | Published dynamical analyses of jet in crossflow and pressure-swirl atomizers [**J1**, **J2**, **J3**]. Master's thesis on doublet injectors.

Optical Diagnostics | Experimental setup and use of laser-induced fluorescence, phase Doppler interferometry, laser diffraction, and high-speed video [**J3**].

Undergraduate/Graduate Researcher | UCI Combustion Laboratory *Jan 2015 – Dec 2016*
Laser diagnostics, high-speed video, POD, and DMD for efficiency and stability of doublet injectors.

WORK EXPERIENCE

Software Engineering Intern | MathWorks *June 2020 – Sept 2020*
Deep Learning Toolbox team - researched and developed deep learning implementations:

- Independently developed first internal transformer network using MATLAB functionality
- Demonstrated cross-lingual embedding for unsupervised programming language translation

Teaching Assistant *Sept 2018 – Dec 2020*
Combustion and Fuel Cell Systems and *Senior Design Projects*. Give lectures, discussion sessions.

AWARDS : MAE Graduate Student of the Year | UC Irvine *March 2019*

PEER-REVIEWED JOURNAL PUBLICATIONS

[J6] **Leask, S.B.**, McDonell, V.G., Samuelsen, S., "Modal Extraction of Spatiotemporal Atomization Data Using a Deep Convolutional Koopman Network", Physics of Fluids, 33(3), 2021.

[J5] **Leask, S.B.**, McDonell, V.G., Samuelsen, S., "Neural Network Prediction of Boundary Layer Flashback", J. Gas Turb. and Power, 143(5), 2021.

[J4] **Leask, S.B.**, McDonell, V.G., Samuelsen, S., "On the Use of Dynamic Mode Decomposition for Liquid Injection", Atomization and Sprays, 29(11), 2019.

[J3] **Leask, S.B.**, Li, A.K., McDonell, V.G., Samuelsen, S., "Preliminary Development of a Measurement Reference Using a Research Simplex Atomizer", J. of Fluids Engineering, 141(12), 2019.

[J2] **Leask, S.B.**, McDonell, V.G., Samuelsen, S., "Emulsion Jet in Crossflow Atomization Characteristics and Dynamics", J. Gas Turb. and Power, 141(4), 2019.

[J1] **Leask, S.B.**, McDonell, V.G., Samuelsen, S., "Critical Evaluation of Momentum Flux Ratio Relative to a Liquid Jet in Crossflow", Atomization and Sprays, 28(7), 2018.

PROGRAMMING

Languages Python and MATLAB (>5 years) Java and C++ (secondary)
Machine Learning TensorFlow, PyTorch, Keras, Scikit-Learn, Pandas, Numpy, Scipy, OpenCV

ABSTRACT OF THE DISSERTATION

Dynamical Feature Extraction of Atomization Phenomena Using Deep Koopman Analysis

By

Scott Benjamin Leask

Doctor of Philosophy in Mechanical and Aerospace Engineering

University of California, Irvine, 2021

Professor Scott Samuelsen, Chair

Liquid injection systems and subsequent atomization behaviors are vital in many power generation and propulsion systems. These systems are inherently complex, owing to coupling of nonlinear processes in turbulent, multiphase flows. As a result, understanding and predicting the dynamical behaviors is inhibited through traditional system analysis which greatly impacts the ability to predict future states or control the nonlinear flow.

Koopman analysis has emerged as a data-driven approach in extracting dynamical features and physical understanding of nonlinear flows through a modal decomposition. This dissertation investigates such Koopman analysis techniques to improve the understanding of atomization systems, with an emphasis on generating compact and interpretable representations. Commonplace techniques of proper orthogonal decomposition (POD) and dynamic mode decomposition (DMD) are applied and evaluated on canonical atomization systems, highlighting their respective strengths and disadvantages as a means to perform a modal decomposition. Although POD and DMD have seen widespread use, they are inherently limited in capturing underlying dynamical processes of nonlinear data, which is demonstrated and verified on a hierarchy of system complexity.

To overcome the limitations of POD and DMD, a deep learning-based extension of Koopman analysis, in the form of a deep convolutional Koopman network (CKN), is proposed for extracting dynamical features of spatio-temporal atomization data. The CKN is end-to-end trainable with an architecture which can

successfully be applied to a wide range of fluid flows. The CKN admits a more compact and, importantly, a more interpretable modal decomposition for improving the physical understanding of these systems. Highly accurate long-term future state predictions are achieved across multiple systems using an identical architecture. Indeed, these findings extend to spatio-temporal data in general which exhibit periodic, dynamical behaviors.

Chapter 1

Introduction

1.1 Motivation

1.1.1 Developing Models

For hundreds of years, scientists and engineers have strived to understand and represent phenomena within the natural sciences. These phenomena are general, defined as any system or process whose behaviors can be observed experimentally, or through simulation during the current era of computational progress. This encompasses physical, biological, and chemical processes, combinations thereof, with no bounds of complexity (or abstraction, as the frame of reference is arbitrary). Understanding these processes and representing them have necessarily gone hand-in-hand for centuries; one can only represent the behaviors of a system or process by first improving their understanding, to some fundamental level, while an understanding can only be developed by first establishing the assumptions, boundary conditions, or initial conditions for the representation.

Progress in scientific innovation is enabled by the ability to both understand systems and by, ideally, representing their behaviors with as simple and few irreducible basic elements that capture all collected data [1]. The scientist or engineer collects data about the system in a methodical approach, revealing the fundamental behaviors which compactly dictate the system behaviors, to then distill their observations into scientific laws and theories which hold generally, save for conditions outside of the system assumptions. This has led to remarkably elegant formulae, capturing the equations of motion with a handful of governing variables, to equally generalizable equations like the Navier-Stokes equations, though the latter has evaded a general analytical solution due to the closure problem.

Unfortunately, not all systems admit a clean and concise mathematical expression which elegantly captures all underlying truths about the state and time-evolution properties, let alone a closed-form analytical solution. In fact, this is the case *in general*. Natural phenomena are tremendously complex: they exhibit spatial and temporal scales over many orders of magnitude; they exhibit apparent randomness or chaos; they are highly nonlinear; expressive state variables are either unknown or intractable to measure; or any combination of the aforementioned.

Nonlinearity, in particular, greatly inhibits human intuition and understanding of these systems. Only linear systems, oftentimes with the added constraint of constant coefficient linear systems [2], or (often impractical) simple nonlinear systems, can be solved analytically. This has led to a growing interest in the field of nonlinear science, despite the ubiquity and abundance of nonlinear problems. "Using a term like nonlinear science is like referring to the bulk of zoology as the study of non-elephant animals" [3] – scientists and engineers have still to make significant progress to advance beyond only understanding elephants.

1.1.2 Liquid Injection and Atomization

A particular area of interest in the nonlinear sciences is liquid injection and atomization systems, which have been a vital component for power generation and propulsion for over 100 years. They must be safe and stable while maximizing efficiency and performance and, of paramount importance, they must burn cleanly. Almost surprisingly, there is a lack of deep understanding for these systems and even simplified representations of these systems. Multiphase flows, interfacial forces, complex dynamical coupling on many spatio-temporal scales with chaotic behaviors have occluded understanding. Yet, society relies on these systems. Regardless of any phasing out of liquid injection systems, for example in favor of electric propulsion, the lack of understanding is found elsewhere.

The lack of understanding is particularly problematic on the dynamics of physical structures in a flow. In these systems, a range of physical structures interact in multiphase flows, and predicting their future

paths and interactions is necessary for controlling or predicting overall system performance. Currently, extracting the spatio-temporal dynamics of atomization processes, which often exhibit continuous spectra, is not possible.

It may seem that complete and true understanding of these systems is unnecessary. After all, the human race has landed their own kind on the moon using these systems despite an apparent lack of understanding. Unfortunately, for all great successes, there is a list of drawbacks, failures, and tragedies. If one truly understands a system, one can avoid detrimental consequences, notwithstanding computational intractability. Improving the methods for improving system understanding, especially for these liquid atomization systems, therefore, has huge potential.

1.1.3 Data-Driven Modeling

Traditional model development relies on experimentation and data collection. All models, whether empirical or derived from first principles, whether general or greatly constrained, have been developed by collecting data or validating against data. To leverage these data, scientists and engineers attempt to distill the data down into general models which can be applied *ad infinitum* for their associated application. Inaccuracies and uncertainties are expected, often with surrogate models used to account for primary model errors.

The past five decades have seen an explosion in computational capabilities, which has fueled a paradigm shift which has permeated, and continues to permeate, every field of study. One of the most significant contributions of this new computational power is in rapid data handling, processing, and analysis at ever growing scales. While humans can only handle working with and maintaining intuition for a small number of variables and manually process a small amount of data, these limitations are removed, subject of course to reasonable computational constraints, by using computers. In fact, this has enabled an entirely new, and exceedingly relevant, approach to developing useful models: data-driven modeling.

Data-driven modeling eliminates, or reduces, the need for human intuition or knowledge in generating a general model directly from data. Instead, an algorithm is used which *learns* structure from data in order to minimize a cost function. Of course, these algorithms are developed by humans and are therefore preconditioned on human knowledge, however, they are often designed to be as general and flexible as possible to remove human bias from the model development.

The progress, power, development, interest, and indeed hype, around data-driven approaches has grown dramatically in the 21st century. Their success has been demonstrated in everything from healthcare to finance to high-energy physics to social media to fluid mechanics. This success has been facilitated by growing computational power but is largely attributed to the great representational capacity of these approaches. These frameworks have allowed for highly nonlinear processes to be identified; a near-impossible challenge using traditional model development techniques. This does come at the cost of model opacity where "opening the black box" of these models is a nascent and active area of research. Ensuring a model, or the results of such a model, is interpretable is conducive to improving understanding of the target system.

Data-driven modeling is well-suited for high-dimensional, nonlinear data which makes it the perfect candidate for investigating the complex, multiphase flow systems comprising much of atomization phenomena. As this is a growing area where applications to fluid mechanics in general have only just begun, substantial progress has yet to be made. Koopman analysis is one of these data-driven approaches receiving growing attention which attempts to globally linearize nonlinear systems, thereby enabling linear analysis. Common Koopman analysis techniques are currently limited, however, extending these techniques into a deep learning framework can improve representational capacity, which has the potential to enable transformative progress in fluid mechanics.

1.2 Goals and Objectives

The goals of this dissertation are to (1) assess the extracted dynamical features of atomization phenomena using existing Koopman analysis techniques, and (2) improve on existing techniques' performance and generalizability while promoting result interpretability by developing and applying a novel deep learning-based Koopman analysis framework: the deep convolutional Koopman network.

The following objectives provide a systematic approach to achieve these goals:

Objective 1: Identify and select Koopman analysis techniques typically used for liquid injection systems

Objective 2: Critically evaluate algorithm design, apply selected techniques, and evaluate extracted results and interpretability

Objective 3: Replicate selected techniques using a deep learning framework

Objective 4: Develop a general deep learning-based Koopman network

Objective 5: Refine the deep Koopman network and training requirements to perform an interpretable modal decomposition

Objective 6: Evaluate the deep Koopman network and compare to traditional techniques

1.3 Dissertation Outline

Chapter 2 provides the reader with the background necessary to understand the contributions of this work. In particular, the relevance and complexity of liquid injection systems and atomization phenomena is given, however, more focus is placed on Koopman theory and associated algorithms, and the pertinent deep learning theory for extending Koopman analysis into a deep learning framework.

The approach to meet the aforementioned objectives is then presented in Chapter 3, which, in turn, enables the dissertation goals to be accomplished.

Chapter 4 covers the methodology of the work capable of achieving the dissertation goals. Relevant Koopman analysis theory, deep Koopman network design and training details, and the methods for comparing Koopman techniques on image and video data are discussed.

In Chapter 5, results are presented and discussed. First, the standard Koopman analysis techniques are implemented, with a focus on their utility, compactness, and interpretability. Next, a proposed deep convolutional Koopman network (CKN) is provided with insight into model development and design choices. The effectiveness of standard Koopman techniques and the CKN on a variety of spatio-temporal data of dynamical systems are evaluated, with a growing emphasis on practical atomization systems.

Finally, Chapter 6 provides a summary of the dissertation with the main contributions and conclusions developed. This work provides early insights into an exciting and growing area, so recommendations for future work to be carried out are also given.

Chapter 2

Background

This dissertation focuses on three distinct areas with appropriate accompanying background. The first area discusses the practical significance and relevant processes and behaviors associated with liquid atomization. Traditional approaches to analyze these fluid systems have yet to advance understanding and technologies sufficiently, owing largely to the highly complex flow phenomena present. Pertinent flow phenomena are discussed based on example applications.

The background of Koopman analysis is then provided. Koopman analysis is the leading approach to dynamical feature extraction for fluid systems, which has demonstrated many advantages in improving fluid flow understanding compared to traditional approaches within the fluid mechanics community.

Finally, an overview of the relevant deep learning background and theory is presented as there is a need to afford existing Koopman analysis techniques with the representational flexibility, and other attractive properties, of deep learning.

2.1 Liquid Atomization Processes

2.1.1 Broader Impact

Environmental Impacts

Liquid transport and subsequent liquid breakup and atomization are essential processes which drive modern life, with example liquid injection systems given in Figure 2-1. These processes are important in many applications including, but not limited to, propulsion, power generation, agriculture, transportation, and pharmaceuticals. All of these applications heavily rely on liquid atomization characteristics [4].

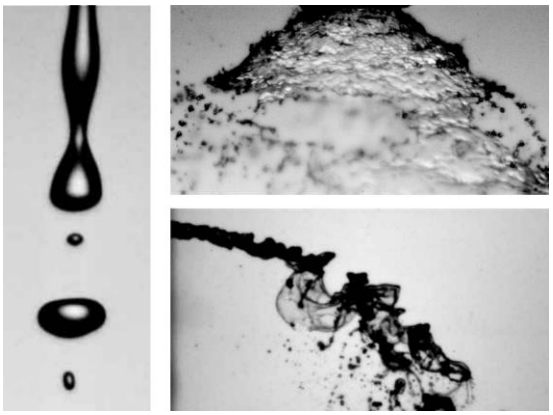

Figure 2-1: Example fuel injection and atomization systems. (Left) A canonical jet breakup case, (top right) flow from a pressure swirl atomizer, and (bottom right) a liquid jet in crossflow system.

Of great significance is the emissions which are produced when combusting liquids through fuel injection systems. Around 80% of the world's energy comes from the burning of fossil fuels [5], of which a third is in liquid form where atomization is necessary to achieve high combustion performance. The properties and characteristics of the fuel atomization have a major impact on emissions and, therefore, climate change, with 80% of emitted greenhouse gases in the U.S. originating from the combustion of fossil fuels [6]. Greenhouse gas and criteria air pollutant emissions are of great societal significance as they have severe environmental and human health impacts, respectively. With an ever-growing concern of climate change, it is a necessity to reduce greenhouse gas emissions, in particular CO_2 [7]. Similarly, nitrogen oxides (NO_x) and particulate matter have a significant impact on air quality, contributing to smog and respiratory issues [8]. Almost half of NO_x emissions arise from liquid fueled engines, with many of these harmful pollutants forming through flame instability, flashback, and equivalence ratio heterogeneity [9]; transient behaviors deriving from atomization mechanisms which are not yet well understood and, currently, cannot be controlled accurately.

Even with a societal shift towards the electrification of, for example, light- and heavy-duty transportation and renewable and sustainable power generation, the presence of these liquid systems, and need to understand them, is necessary. There are numerous reasons for this. First, these systems will continue to be necessary and relevant to non-combusting applications, such as in agriculture and pharmaceuticals. For

combustion applications, liquid fueled injection systems are likely a mainstay for rocket and aircraft applications, owing to their higher energy densities and thrust capabilities required for the aerospace industry, compared to their electric counterparts [10]. Second, for power generation and light- and heavy-duty transportation, there is still a transition period required before becoming fully electric [11], allowing for significant improvements in current combustion technologies in the next few decades. Finally, any fundamental understanding of these systems is likely to transfer to many other areas of fluid mechanics.

In all cases, it is imperative to optimize and understand system properties and behaviors. By improving the understanding of liquid injection and atomization mechanisms and dynamics, societal and technological benefits can be enabled. The production of greenhouse gases or criteria air pollutants can be reduced, the efficiency of liquid dispersion and subsequent combustion can be improved, while ensuring safe and stable system operation.

Operational Impacts

Issues associated with emissions can be mitigated through two avenues: cleaner combustion of existing fuels or the practical implementation of advanced, green fuels. The former relies on understanding the pathways of harmful emission formation, for example through turbulent liquid transport and breakup characteristics which govern the combustion process, in order to design cleaner systems. This dependence has been noted for decades (e.g., [12]) and gives rise to the need for experimental and computational work on liquid breakup phenomena. Despite decades of work, traditional analytic approaches for understanding the dynamical behaviors of these systems have not been met with great practical improvements.

The practical implementation of advanced fuels relies on the tailored understanding of specific candidate renewable fuels which inherently have better molecular properties, e.g., carbon-free molecules. Currently, these types of fuels have unfavorable performance relative to typical fuels, due to, for example, considerably higher viscosity and surface tension values [13]. This inhibits liquid breakup which consequently affects operation stability making them unsuitable as a drop-in replacement for fossil fuels [14]. More

fundamentally, the inability to transition to new liquid fuels stems from the poor generalization of current models or understanding of these systems and how they behave. Generalizability is particularly important for models, both in terms of a given model's ability to perform well on new test conditions and its practical implementation to many systems or areas.

Addressing stability issues not only ensures expected performance but also promotes system lifetime and operational safety. A fuel injection system which operates close to steady-state limits material fatigue and failure associated with oscillatory behavior. Combustion instabilities often arise through liquid injection, atomization, and mixing, which greatly impact combustion stability and operational safety. Moreover, these instabilities and processes promote nonlinear behavior and inhibit principled understanding [15]; the governing liquid atomization theory and processes is a vast and largely unsolved area.

2.1.2 Liquid Atomization Theory

In its own right, liquid atomization is a subject that requires numerous books to even begin to address its complexity. Not only is turbulence of a single phase flow yet to be solved, and one of the most important unsolved problems in classical physics [16], liquid atomization deals with turbulence of *multiphase* flows. This section addresses only the most relevant details for future chapters.

Injector Design

The goal of a liquid injector is to atomize and disperse a given liquid throughout a continuous gas environment according to some pre-defined objective. The geometric and operational design of an injector must take into consideration the internal and external flows, liquid breakup mechanisms, and environmental conditions to achieve this goal.

An example liquid injection system is shown in Figure 2-2; a pressure-swirl atomizer commonly used for aerospace applications. While the internal flow and geometrical artifacts have considerable effects, an

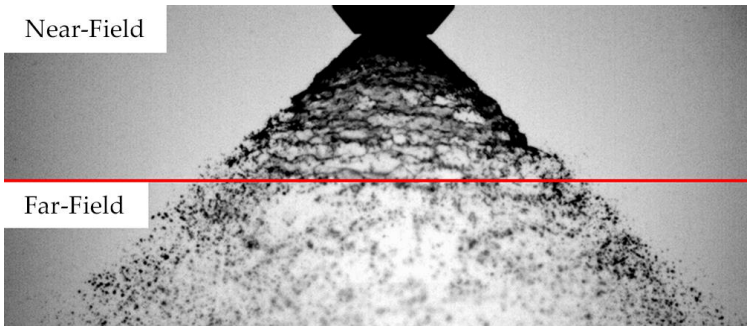

Figure 2-2: An example atomization system from a pressure-swirl atomizer. Atomization systems can often be divided into the near-field, close to the nozzle, and the far-field.

example being cavitation [17, 18, 19], often both the experimentalist and computational fluid dynamicist will assume idealized exit conditions from the injector exit. Boundary and initial conditions are then imposed which, along with environmental conditions, promote liquid breakup into the desired atomization characteristics.

The atomization system is often characterized in two different sections: the near-nozzle near-field and the far-field, also shown in Figure 2-2. While the extent and exact properties of these regions vary dramatically across different injector designs, there are many similarities relevant to understanding these systems. In the near-field, *primary* atomization occurs; the intact liquid structures emanating from the injector undergo initial atomization processes. This region is predominantly made up of large segments of liquid ligaments and liquid sheets, and large amorphous droplets, creating a highly dynamic and optically dense region.

In the far-field, *secondary* atomization occurs. As the initial intact liquid breaks down into high volume structures, aerodynamic and hydrodynamic forces continue to disintegrate these structures further, resulting in full atomization where only droplets, stable under their own surface tension forces, are present. In contrast to the near-field, the far-field has predictable liquid structures, approaching spherical droplets, and has a low optical density. This enables the use of optical diagnostics for probing the entirety of the far-field volume.

Once secondary atomization is complete, a droplet size distribution has been produced. Generally, a specific liquid droplet size distribution or dispersion homogeneity is desired. These particle sizes are often

distributed according to the Rosin-Rammler, or Weibull, distribution. Consider the experimentally obtained distribution of droplet sizes *at a point* in the far-field of a pressure-swirl atomizer shown in Figure 2-3. A wide range of physical scales exist throughout the entire atomization system. Coupled with a size distribution target, injectors are also subject to velocity and momenta criteria. This is a complex challenge which requires an understanding of atomization phenomena and dynamics spanning all spatial and temporal scales.

Multiphase Flows

All atomization systems are multiphase; a dispersed liquid phase is injected into a continuous gas phase. The ultimate goal is to atomize the liquid and disperse the liquid mass throughout the gas subject to problem requirements. Typically for combustion processes, monodispersed droplets with a small diameter, homogeneously distributed throughout the gas phase, is desired. In addition, these systems are often *multicomponent*; the liquid and gas phases have dissimilar molecular constituents. Of course, any flow utilizing air, which is the most common gas phase for transportation and power generation, is a multicomponent flow.

What is often of interest is the dynamics of the final liquid droplet size distribution as it is suspended in the gas phase. As the distribution represents a continuum of values, it is often represented by a *mean diameter* of the form:

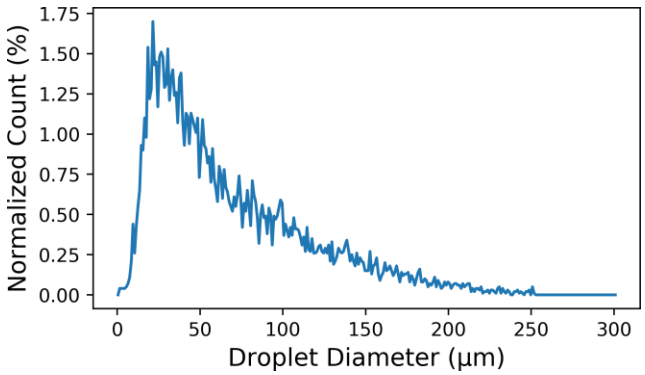

Figure 2-3: An experimentally measured droplet size distribution at a point in an atomization system.

$$D_{jk} = \left[\frac{\int_0^\infty D^j N'(D) dD}{\int_0^\infty D^k N'(D) dD}\right]^{\frac{1}{j-k}}. \quad (1.1)$$

A commonly used mean diameter is the *Sauter mean diameter* (SMD), which is representative of the system volume to surface area ratio. These mean diameters provide a simplifying step for modelers in the hope that they are representative of the entire system.

The evolution of both phases is encoded in the *Navier-Stokes equations*. The equations elegantly capture the relationship between the velocity, pressure, density, and temperature of moving fluids. Unfortunately, the form of these equations is as coupled differential equations which, particularly in the presence of turbulence, are intractable to solve analytically. Even more concerning, numerically solving the Navier-Stokes equations for practical conditions is outside of computational capabilities, and so simplifications, such as using a mean diameter, are essential to model the multiphase flow behaviors [20].

Other parameters exist for practicality. The Stokes number, St, is an important parameter which captures the relationship in an appropriate characteristic time between a liquid particle and the gas flow it is suspended in:

$$\mathrm{St} = \frac{\tau_l}{\tau_\infty} \quad (1.2)$$

where τ_l is a characteristic time for the liquid, and τ_∞ is a characteristic time of the gas. For sufficiently small values, i.e., $\mathrm{St} << 1$, then the liquid particle velocity field will approach the velocity field of the gas flow. While for large values, $\mathrm{St} >> 1$, the gas flow has negligible effect on the liquid. For all other values, which is where typical droplet size distributions lie [21], there is a competing effect between the two phases. The Stokes number provides a useful heuristic for the bulk behaviors of the flow, but it does not provide dynamical insight.

In addition, there is often phase coupling [22]. Coupling of phase mass, momentum, and energy transfer greatly complicates individual phase behaviors which are only compounded by interfacial forces between the two phases [23]. These complexities have meant prediction and computation of multiphase behaviors are confined to a wide array of idealized or simplified conditions or assumptions to admit practical

implementation. An overarching problem is the inherent complexity of multiphase flows and, as such, a detailed description of the many theoretical approaches is beyond the scope of this dissertation, though readers are directed to [24] for further reading. Each aforementioned aspect is still actively being researched and only exemplifies the complexity of the subject.

Waves and Oscillatory Behaviors

Despite the complexity of multiphase flows, it is often the case that a number of simple processes contribute to the overall system complexity. The majority of liquid atomization systems are comprised of liquid droplets, jets or ligaments, or sheets which may be any 2D surface, and intermediate transient states thereof. Example systems in Figure 2-1 contain these fundamental elements.

One of the prevailing governing forces for these simple elements is surface tension forces, while additional hydrodynamic and aerodynamic forces compound complexity but represent a more realistic system [4]. The presence of numerous oscillating modes within droplet breakup processes has been noted since the 1980's [25], with acknowledgement of decaying energy content of particular modes. The effect of Weber number, We, and Ohnesorge number, Oh, on the transition between droplet oscillatory deformation and bag breakup was presented in [26]. Indeed, these dimensionless numbers are commonly used for liquid studies, and are represented by

$$\mathrm{We} = \frac{\rho u^2 d}{\sigma} \tag{1.3}$$

$$\mathrm{Oh} = \frac{\mu}{\sqrt{\rho \sigma d}} = \frac{\sqrt{\mathrm{We}}}{\mathrm{Re}} \tag{1.4}$$

$$\mathrm{Re} = \frac{\rho u d}{\mu} \tag{1.5}$$

where ρ is density, d is a characteristic length, which is often the diameter of the injector outlet, σ is surface tension, μ is dynamic viscosity, u is velocity, and Re is Reynolds number. Units are arbitrary so long as they are consistent and cancel out. The above dimensionless numbers have significant utility in all fluid

flows, despite them representing *bulk* values. As a result, dynamical information and understanding is not considered in favor of general flow behaviors.

A specific multiphase variant of the Weber number is the *aerodynamic* Weber number, which is the ratio of the aerodynamic drag force on a liquid segment to the liquid's surface tension force. It is given by

$$\mathrm{We}_{\infty} = \frac{\rho_{\infty} u_{\infty}^2 d}{\sigma_l} \tag{1.6}$$

where the subscript ∞ and l denote properties of the gas and liquid phases, respectively. Surprisingly, these dimensionless numbers are not free from their own complexities, as they have shown to be highly sensitive in their utility for jet in crossflow systems [27].

From an atomizer, the liquid is often injected as either a liquid circular jet or a planar or conical sheet. Liquid jets are perhaps the most common and well-studied flow from atomization systems, going back to Lord Rayleigh [28] and Haenlein [29] in the late 19th century and early 20th century, respectively. Surface tension forces drive diameter oscillations with known wavelengths in the absence of air friction, while the presence of aerodynamic forces can lead to dilational or sinuous oscillations prior to liquid breakup. This has led to a number of simple analytical expressions for stable and unstable wavelengths of idealized conditions.

Of course, a spectrum of breakup types and complexities can be achieved, often governed empirically by Oh and Re values [30]. The atomization mechanisms of liquid jets from plain orifices were studied extensively by Reitz and Bracco [31], focusing on various hydrodynamic and aerodynamic instabilities. These instabilities are primarily investigated experimentally with attempts to tie findings to previous developments of atomization theory, specifically from the incompressible, linearized Navier-Stokes equations. Agreeing with work by Taylor [32], they present the wave growth rate, ω, of a wave with wavelength k to be

$$\frac{\omega}{ku} = 2\left(\frac{\rho_l}{\rho_{\infty}}\right)^{\frac{1}{2}} f(\rho_l, \rho_{\infty}, u, \sigma, \mu) \tag{1.7}$$

where u is the jet injection velocity. They correlate experimental atomization findings with growth rates of surface waves and the spreading angle of a liquid jet, though strong dependency was found on nozzle geometry. While the prediction of atomization behaviors and mechanisms of liquid jets are limited to simplified cases, such as for approximately inviscid flow, theoretical analysis continues to provide insights [33].

Likewise, a liquid sheet can induce sinuous or dilational waves with known wavelengths, as excited by aerodynamic forces [34], with theory governing growth rates of surface waves being identical to the liquid jet case [31]. Liquid sheets undergo successive stages of atomization, often first to large ligament structures, which share many properties with liquid jets, prior to ligament disintegration into droplets. These behaviors are shared with conical sheets.

Although the above represents simplified cases, they often admit analytical understanding and they are both ubiquitous in practical applications and serve as fundamental processes which contribute to all subsequent, more complex processes. Further, they serve as a *ground truth*; they are repeatable and periodic with known behaviors. Any attempt to analyze such an atomization system, whether complex or simplified, must extract, or be validated with, these processes.

Turbulence

Even for the simplest injectors, for example plain orifice injectors, analytical system understanding is confined to laminar or transitional flow of a cylindrical jet which gives rise to the above oscillatory behaviors. In fact, these findings date back over 100 years ago without much analytic improvement [35]. Indeed, characterization of more complex systems has improved, but the presence of turbulence, coupled with multiphase interactions, has marred the development of analytical theory or a truly generalizable understanding of flow behaviors.

In reality, practical atomization systems are highly nonlinear. For multiphase interactions, there is a velocity differential at the phase interface which creates Kelvin-Helmholtz instabilities, these instabilities

may be damped due to surface tension forces which, in turn, also induces Rayleigh-Taylor instabilities [4]. There is a cascade of processes producing or affecting new processes resulting in atomization systems which are highly nonlinear. Of course, this is exacerbated by high velocity instabilities and the energy cascade found in turbulence flows, occurring on spatio-temporal scales over many orders of magnitude [36]. For numerical analysis, many works have addressed the coupling of Kelvin-Helmholtz and Rayleigh-Taylor instabilities through linear stability analysis, which has only recently been expanded to include nonlinear effects [37].

A particular practical system of interest in this work is that of a jet in crossflow, where multiple breakup mechanisms are shown in Figure 2-4. The liquid jet in crossflow case is commonly found in high velocity air-breathing engines [4] and shares many behaviors with the well-studied liquid jet. In these systems, a liquid fuel is injected perpendicular into a gaseous oxidant flow, which is commonly air, where the velocity differential gives rise to many nonlinear, turbulent interactions. As these systems have been intractable to study analytically, works have focused on generating correlations to predict characteristics and behaviors on a macroscale. Works for gaseous jets in crossflow [38, 39, 40, 41] and liquid jets in crossflow [42, 43, 44] have identified the importance of We_∞ and the momentum flux ratio, given by

$$q = \frac{\rho_l v_l^2}{\rho_\infty v_\infty^2}, \tag{1.8}$$

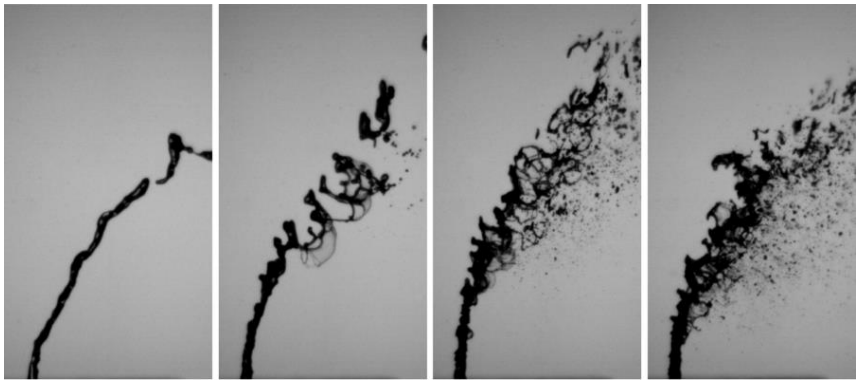

Figure 2-4: Breakup mechanisms of a jet in crossflow system. From left to right: column, bag, multimode, and shear breakup.

for governing jet penetration depth, droplet size characteristics, and breakup regime. Work by Wu et al. [44], especially, has been influential as they provide a regime map for the breakup mechanisms given in Figure 2-4 based only on We_∞ and q, though turbulent effects were not considered. These works do not account for the dynamical behaviors of the system in favor of capturing general trends. Reviews and summaries of jet in crossflow investigations have been provided (e.g., [45] and [46]) which have discussed discrepancies within the literature. A particular concern is the lack of a single, generally applicable model as existing correlations simply do not generalize well or capture dynamical behaviors.

2.1.3 Diagnostics and Analytical Tools

In the absence of analytical solutions for solving the flow evolution of liquid injection systems, these systems can be probed experimentally to directly capture state variable behaviors and processes, which can subsequently be used to enhance and validate computational fluid dynamics (CFD) models or used directly for data-driven approaches.

Unfortunately, no diagnostics currently exist which are able to provide high spatio-temporal resolution global data of an atomization system. Intrusive methods are often avoided as they directly impact the flow characteristics and behaviors. As a result, the majority of diagnostic tools are optical so as to be non-intrusive. A wide variety of diagnostics are available to the experimentalist, however, there is no panacea. In particular, the multiphase flows are often optically dense or interfere with the diagnostics, inhibiting measurement fidelity.

Experiments for characterizing and understanding liquid atomization systems typically investigate either the near-nozzle near-field or the downstream far-field, recalling Figure 2-2. The former is of critical importance as it is where primary atomization begins and large liquid segments breakup. An understanding of the phenomena occurring in the near-field should, in theory, allow for prediction of far-field behavior.

Near-Field Diagnostics

Primary atomization occurs in the near-field; the emanating liquid from the injector begins to disintegrate into large segments. Amorphous liquid sheets, ligaments, and large, unstable droplets are common structures present within this region, combining to create an optically dense and stochastic region.

Currently, no diagnostic is able to successfully collect instantaneous atomization characteristics, e.g., droplet sizes, velocities, or liquid breakup monitoring, in the near-field. Improvements have been made using high fidelity simulations [47] for this purpose, but these simulations are prohibitively expensive for practical investigations. Methods to probe the near-field are generally image- or simulation-based, because other diagnostic methods, such as those commonly used in the far-field region, rely upon unreliable assumptions of liquid fragment morphology or require low optical density.

High-speed shadowgraphy can be applied but only to capture peripheral information. To probe optically thick systems, advanced time-gating methods can be applied, such as ballistic imaging (e.g., [48]), which rejects multiply-scattered photons and has shown promise in imaging very dense spray regions. An example comparison of high-speed shadowgraphy and ballistic imaging for the same system is shown in Figure 2-5. Ballistic imaging is able to capture intact liquid structures despite high optical density but suffers from low spatial resolution, a lack of portability, and may require image-averaging to increase signal-to-noise ratio. Regardless, it is unable to capture fine time-resolution data [49].

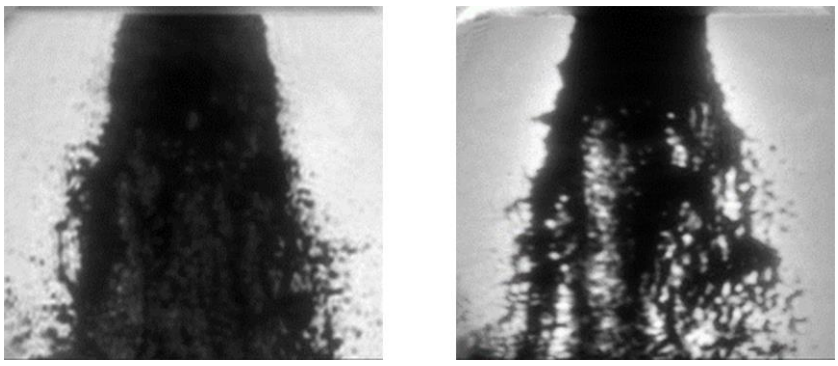

Figure 2-5: A comparison of (left) high-speed imaging and (right) ballistic imaging for the same system. Image from [48].

Other near-field diagnostics include digital particle field holography and x-ray radiography, but these suffer from the same issues or are prohibitively costly. Not only is analytical progress inhibited by system complexity, but as is the collection of meaningful upstream data.

Far-Field Diagnostics

Despite the theoretical benefits of probing the near-field, the far-field is much easier to characterize as optically dense liquid ligaments and sheets have undergone secondary atomization, dispersed into a larger spatial volume, and form a dilute system. Laser-based diagnostics have had great success in characterizing the droplet size distribution and velocity field of dilute atomization systems, such as through laser-induced fluorescence (LIF), particle image velocimetry (PIV), phase Doppler interferometry (PDI), and laser diffraction.

All of these techniques rely on the scattering behaviors of light as it is transported, non-intrusively, throughout the system of interest. LIF and PIV are relatively insensitive to the liquid morphology, though the former requires long time scales to record data due to the fluorescence time scales and weak intensity signals. PDI and laser diffraction are two common approaches for measuring droplet sizes, where PDI can also measure velocity, but they are highly sensitive to liquid morphology as they have strong assumptions on droplet sphericity [50]. In addition, they have small measurements volumes. PDI can provide highly accurate particle size and velocity measurements with high temporal resolution, but only at a point. Laser diffraction has a greater measurement volume, expanding to a line of sight measurement, but with low temporal resolution. To collect quasi-global information about the system with these techniques is very time expensive and will not have temporal coherence. This, in itself, encourages the use of bulk or averaged values to reconcile the asynchronous measurements.

Unfortunately, while PDI and laser diffraction are extensively used for capturing atomization characteristics, they also suffer from sensitive operation. A large number of works (e.g., [51, 52, 53, 54, 55, 56, 57, 58, 59] have evaluated the robustness of these techniques on nominally identical diagnostics and

experimental systems. Numerous caveats and inconsistencies were presented in their findings with many instruments requiring calibration but with few robust calibration procedures available. There is then a circular problem which affects liquid atomization research: a system with known characteristics is required to fully calibrate a diagnostic, but knowing the characteristics relies on already having fully calibrated diagnostics. Indeed, alongside the development of more robust diagnostics, *reference* liquid injection systems have been established in an attempt to provide the fluids community with systems with known characteristics and dynamics at different test conditions [60, 61, 62, 63].

Extracting the underlying true characteristics or behaviors is currently intractable with current diagnostics, however, using the law of large numbers, measured characteristics can be assumed correct if a large and varied enough dataset is produced for a given atomizer.

Computational Modeling

Numerical modeling is heavily constrained for turbulent flows as, unfortunately, the Navier-Stokes equations lack a closed-form solution when modeling turbulence. Additionally, the lack of instantaneous, global, quantitative characterization of the near-field and primary atomization hinders the already challenging problem of numerical modeling. Only simple cases have been sufficiently characterized experimentally, such as individual droplet breakup and dynamics [64] and laminar flow fragmentation [65], so near-field primary atomization numerical modeling is rare and can lack experimental validation [47]. Recent progress has been made in applying direct numerical simulation (DNS) to understand primary atomization for turbulent inflow conditions of a circular liquid jet [66]. For that study, DNS had utility in identifying turbulent structures, as shown in Figure 2-6. Asymmetric surface waves of the emanating jet promote the generation of liquid rims creating high vorticity turbulent features. This provides additional insight into turbulent mechanisms beyond existing experimental work. While DNS does not require the same experimental validation as other CFD models, it requires unique access to expensive computational resources to simulate practical systems (see, e.g., [67, 68, 69]). Any practical system is likely unable to be

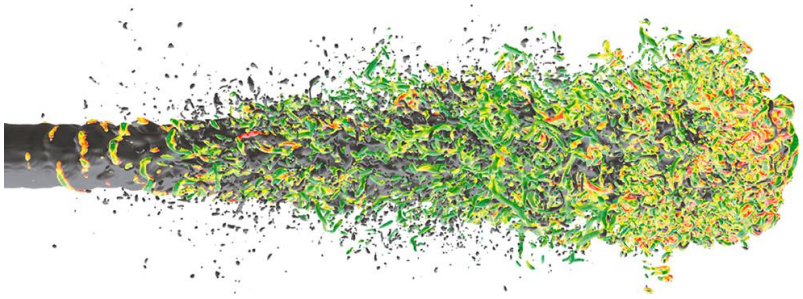

Figure 2-6: DNS of the breakup of a turbulent jet. Colored regions represent turbulent structures. Image modified from [66].

understood or predicted with sufficient accuracy for the desired application so, instead, approximations are required. Therefore, a need exists for methods that are able to quantitatively characterize the near-field, which includes extracting system characteristics and dynamical features, which can aid in the development of more computationally accessible numerical models.

Bohr et al. [70] discusses turbulence extensively with a dynamical systems approach with consideration of chaos. While *some* turbulent systems can be simulated directly, this does not necessarily provide any understanding of the underlying phenomena. A growing approach, avoiding analytic derivation or tackling the Navier-Stokes equations directly, is to identify the underlying spatio-temporal dynamical features which govern the behaviors of a flow. This may not provide the desired generality of the Navier-Stokes equations, but is more amenable to extracting physical understanding and enabling nonlinear control. Indeed, extracting dynamical features is a general approach to solving nonlinear dynamical systems as a whole [71], which operates on automatically identifying system approximations directly from data opposed to traditional approaches from first principles [2]. For the aforementioned flows, it is intractable to capture and model their full, high-dimensional state. Instead, a possible alternative is to identify a low-dimensional subspace which captures the most energetic dynamical processes. This is discussed to greater detail later.

High-Speed Imaging

An ability to determine the characteristics and dynamics of atomization systems is needed to improve physical understanding of these systems, to improve CFD models, and to design and control improved liquid injection systems. Current best practices for extracting this information rely on laser-based diagnostics. As covered previously, these systems are often very expensive, have small measurement volumes, and make assumptions about particle sphericity. In contrast, imaging methods, such as high-speed video, provide an instantaneous, quasi-global, 2D characterization of the system of interest. Indeed, where optical density is too high, particularly at the near field, only peripheral information can be recorded with high fidelity.

Use of high-speed imaging is often only used qualitatively, see Figure 2-1 for example high-speed images, yet it inherently captures the underlying system dynamics and, depending on data resolution, system characteristics. High-speed recordings can exceed 1 megapixel, and their temporal resolution can be on the order of 10^{-5} seconds, generating a vast amount of data and spanning a large probe area. Importantly, the use of high-speed imaging is very robust and common. Recordings do not require extensive calibrations and are not sensitive to user operation.

A primary concern with high-speed imaging is that it does not directly measure state variables of interest. However, if it were possible to leverage this data quantitatively, a more flexible and robust diagnostic would be available, yielding high resolution, quasi-global spatio-temporal dynamics, further improving system understanding and modeling efforts. Data-driven approaches are an attractive and growing area to directly learn from these datasets, with Koopman analysis techniques being at the forefront in the fluids community.

2.2 Koopman Analysis

The previous section has discussed how fluid flow problems are typically very difficult to model accurately and generally. Difficulties arise due to certain characteristics which are inherently found in these systems; the systems exhibit high-dimensionality, complex coupling of system processes, and often suffer from noise or incomplete data [4]. In a practical experiment, system processes abound. A variety of procedures can be taken to isolate underlying system processes or to limit the effect of undesirable processes, however the recorded data will still likely be an entangled combination of numerous process modes or behaviors. Even the simplest of these systems lack analytical theory based off first principles as multiphase interactions and turbulence greatly inhibit detailed physical interactions [72].

Of course, collecting high fidelity measurements throughout the entire atomization system is itself a great challenge which only hinders physical understanding and model development for these phenomena. However, recent decades have seen tremendous progress in computational power, data acquisition, and data analysis. This has enabled data-driven analysis: the ability to learn and discover system behaviors and patterns directly from collected data. Data-driven approaches provide an alternative route for understanding and modeling systems where analytic theory is hindered. In particular, their performance improves as the amount of data available increases. Focus in this dissertation is placed on data collected from high-speed video as it is especially amenable to data-driven techniques due to the large quantities of data acquired.

Koopman analysis is an example data-driven approach based on Koopman theory which takes advantage of high-dimensional data, attempts to decouple underlying system processes, and enable linear analysis for nonlinear systems. Generally, Koopman theory states that all nonlinear systems can be *globally* linearized through a possibly infinite-dimensional coordinate transform where the dynamics evolve according to a linear operator: the Koopman operator. This has the potential to facilitate future state prediction and control of highly nonlinear systems.

2.2.1 Data-Driven Approaches

Large quantities of data are becoming ubiquitous in all science and engineering fields. As a result, learning from these big data collections, i.e., datasets which are amenable to data-driven analysis, has become a serious undertaking to facilitate system analysis and to identify low-dimensional representations of high-dimensional structures. With the explosion of big data, machine learning has flourished and is quickly becoming a popular data-driven toolset in solving heretofore difficult or intractable problems throughout science and engineering.

Consider Figure 2-7 for a general overview of a data-driven pipeline. The core of these techniques is the learning algorithm; computational procedures for processing large datasets which learn parameters of a given model. The parameters are updated according to a relevant cost function, which may be solved analytically or iteratively, depending on the complexity of the model. Given enough computational time and data, a trained model is produced which can be used for prediction, e.g., regression or classification, of new input data.

Although impressive progress has been made in addressing problems through machine learning techniques, much of the current progress is merely a means to an end; solving a problem is prioritized over understanding the problem itself. Neural networks, as an example machine learning framework, have been

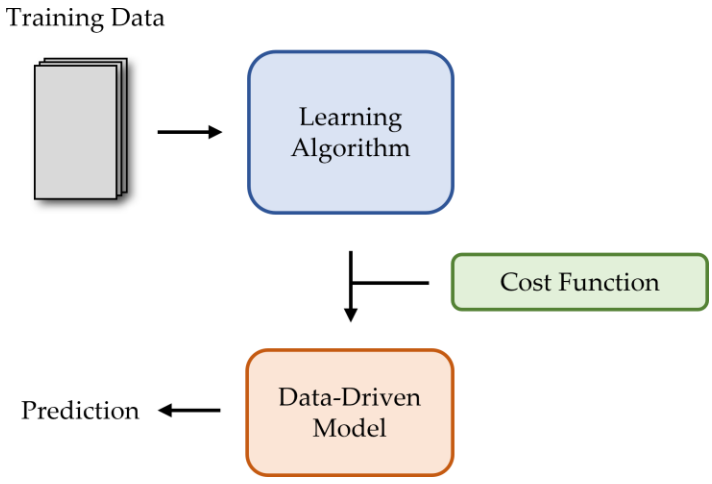

Figure 2-7: A general framework for a data-driven technique.

particularly criticized for this as it is often a *black box* which achieves the task at hand without relying on concrete understanding of how the network is learning. Researchers have noted this issue for over two decades [73, 74], yet this has continued even to current efforts without transformative consequences [75]. A better understanding and interpretation of how these data-driven techniques learn has the potential to aid understanding of the systems analyzed, to improve how the models solve problems, and to aid understanding in how the models generalize. The latter is of great interest [76].

As much of the work in liquid transport and breakup relies on experiments and simulations which generate large amounts of data, e.g., image-based measurements or DNS, these problems lend themselves very well to machine learning techniques. In fluid mechanics, many flows are of very high dimension where, for example, a large number of spatially distributed measurements are required to capture the global behaviors of the flow. This is commonly found in high-speed video and simulations from CFD. Therefore, many researchers have taken advantage of machine learning techniques for dimensionality reduction so to capture low-dimensional structures of the original high-dimensional datasets (e.g., [77, 78, 79, 80]). Dimensionality reduction takes high-dimensional data and eliminates low information-content dimensions to reduce dataset size and to facilitate data analysis. Variants of these techniques aim to extract the dynamical features of independent processes, called *modes*, which govern flow behavior and their corresponding modal dynamics. Hence, many of these dimensionality reduction approaches perform a *modal decomposition*. An example modal decomposition for a pressure-swirl atomizer is given in Figure 2-8. Instead of representing a video as a collection of frames, dominant spatio-temporal structures are identified through modes where, ideally, the modes capture distinct system phenomena, e.g., surface waves and accompanying frequencies.

Many modal decomposition techniques are founded in abstract theory whose application purports to extract physical understanding. In an atomization system, for example, although a large number of modes may be affecting the system dynamics, it is often the case that the flow is dominated by only a small subset of fundamental and periodic modes and that an accurate reconstruction of the entire system may be computed with knowledge of only a few modes. Dimensionality reduction can then be performed, whereby

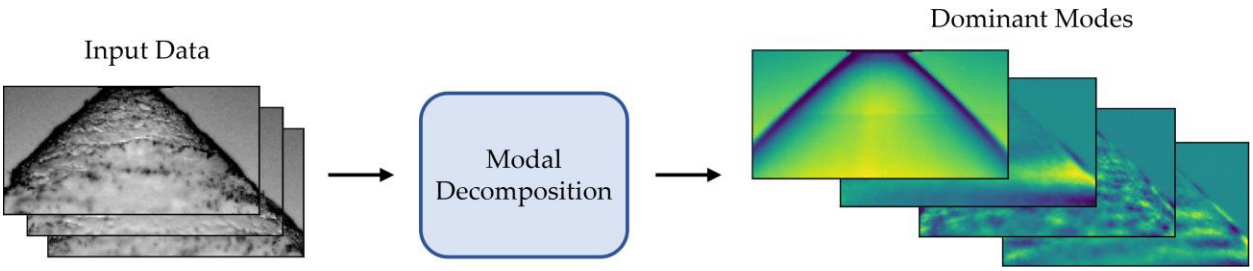

Figure 2-8: Modal decompositions transform input data into spatio-temporally coherent modes.

insignificant modes are removed based on, for example, an energy content metric, which can reduce redundancy in datasets and illuminate dominant flow processes [81, 82].

Oftentimes, image-based recordings of a dynamical system are used for modal decomposition as they are a convenient and non-intrusive approach to capture a large spatial extent of the system at a, potentially, very high time resolution. For liquid injection systems, workers are greatly limited in their diagnostic tools due to complex multiphase interactions and instabilities which plague both the near- and far-field areas. Instead, learning physical structures through high-speed video offers a convenient way of extracting very high spatial- and temporal-resolution data from the probed system. Further, the field of computer vision has already seen large improvements in capabilities from learning from pixel-value data alone [83]. Of course, generally, high-speed video captures a two-dimensional representation of a three-dimensional dynamical system, but this does not prevent meaningful results from being extracted. Three-dimensional imaging techniques may provide several advantages, for example through digital holography [84].

2.2.2 Koopman Mode Decomposition

Koopman analysis, or Koopman operator theory, is a data-driven approach able to perform a modal decomposition: the *Koopman mode decomposition*. This approach has gained increasing attention recently (e.g., [85, 86, 87, 88, 89]), owing to its ability to allow for linear analysis of nonlinear systems. Koopman analysis seeks to identify the Koopman operator [90]; a possibly infinite-dimensional linear map whose eigenfunctions describe the coordinate transformation to globally linearize the system dynamics. These

eigenfunctions are linked to the Koopman modes. This transformation into a linear system is not restrictive; nonlinear dynamics are still captured through the linear system, only the variables governing the system have changed. Once the system is globally linearized, analysis of the system is greatly enhanced. Intuitively, Koopman theory says that there are a set of system observables, which are a function of measurable variables, which behave linearly.

The problem then becomes finding the observables, or observable function, where they dynamics evolve linearly. Consider a discrete-time system, chosen due to its practicality over continuous-time systems. A dynamical system of the spatial distribution of variables $\mathbf{x} \in \mathbb{C}^n$, such as velocity or pixel-values, can be written

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) \tag{1.9}$$

where $\mathbf{f}$ is some possibly nonlinear function which maps from states at time step k to $k+1$. The Koopman operator, $\mathcal{K}$, acts on, generally, nonlinear observable functions, $\mathbf{g}(\mathbf{x})$. This can be written

$$\mathbf{g}(\mathbf{f}(\mathbf{x}_k)) = \mathcal{K}\mathbf{g}(\mathbf{x}_k). \tag{1.10}$$

Instead of modeling Equation 1.10, one can directly model the eigenfunctions of $\mathcal{K}$, the Koopman eigenfunctions, through

$$\phi(\mathbf{x}_{k+1}) = \mathcal{K}\phi(\mathbf{x}_k) \tag{1.11}$$

for a given eigenfunction, ϕ. This bypasses modeling the observable functions, and instead projects onto the eigenfunctions which yields the Koopman modes, $\boldsymbol{\psi}_j$, associated with eigenfunction ϕ_j, according to

$$\mathbf{g}(\mathbf{x}) = \sum_{j=0}^{r} \phi_j(\mathbf{x})\boldsymbol{\psi}_j\,, \tag{1.12}$$

where r represents the number of modes used for the prediction or reconstruction. Additionally, the eigenfunctions are orthogonal which yields a block diagonal Koopman operator, where each block can be associated with its own Koopman mode.

A low-dimensional approximation of the Koopman operator, denoted here as $\mathbf{K}$, is desirable as it is generally infinite-dimensional. Ideally, the low-dimensional approximation represents a Koopman-

invariant subspace, whereby $\mathbf{K}^m\phi(S) \subseteq S$ for an invariant set of states, S. Therefore, any trajectories within S are confined to S, which also includes useful dynamical features such as fixed points, limit cycles, and attractors.

In the study of dynamical systems, it is these dynamical features which are sought, with each feature associated with a given Koopman eigenfunction or mode. Indeed, each mode contributes linearly to the overall Koopman mode decomposition according to

$$\mathbf{g}(\mathbf{x}_m) = \mathbf{K}^m\mathbf{g}(\mathbf{x}_0) = \sum_{j=0}^{r} \lambda_j^m \phi_j(\mathbf{x}_0)\boldsymbol{\psi}_j \tag{1.13}$$

where λ_j is the associated eigenvalue of the Koopman eigenfunction. The eigenvalues capture the stability and frequency of the associated modes, with general trends depicted in Figure 2-9. Eigenvalues which lie close to the unit circle capture the long-term dynamics of the observables [86] and are often of greatest interest. Modal energy can also be extracted which identifies the most dominant oscillatory behaviors and transients present in the system.

Through knowledge of the observable function, **g**, the Koopman modes can be projected in either the state space or the observable space. The choice of which projection is dependent on the interpretability of the mode.

With access to an approximation of the Koopman operator, nonlinear prediction, control, and system understanding can be improved.

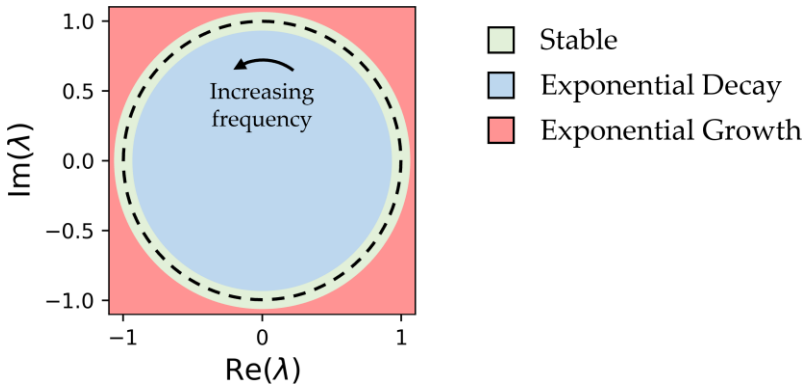

Figure 2-9: The distribution of modal eigenvalues, generally for modal decompositions, describe the stability and oscillatory behaviors of the corresponding mode.

2.2.3 Dynamical Feature Extraction

Koopman theory lends itself well to the extraction of dynamical features for a system of interest; spatio-temporal coherent structures can be extracted through the Koopman modes. The need for dynamical feature extraction for liquid atomization systems is three-fold. First, in the absence of analytical theory or improved understanding from empirical correlations, the extraction of spatio-temporal dynamical features can provide a simplified but accurate approximation of dominant system behaviors.

Second, while extracting characteristics has its own merits, they do not give insight into the dynamical behaviors and, therefore, future state prediction of the system. By extracting dynamical features, the spatio-temporal coherence of system structures can be obtained which naturally allows for the temporal evolution of features to be determined.

Third, the dynamical features are more readily able to improve *interpretability*. Extracting system size characteristics, for example, lacks context for practical experiments, as either spatial or temporal resolution, or both, is prohibitively small to understand the full system behaviors. By construct, dynamical features admit clear context and interpretation. They are often extracted from high spatial resolution images or videos, thereby preserving the spatial topology of dynamically-similar physical structures, with the temporal evolution succinctly described through a dynamical system, whose time steps can be set as desired.

Of course, the above relies on physically meaningful dynamical features to be extracted, with an accurate model of their temporal evolution. However, there are many groups both actively researching and demonstrating the capabilities of such an approach.

2.2.4 Algorithms

While Koopman theory affords linear analysis of nonlinear systems, both $\mathcal{K}$ and ϕ are unknown in Equation 1.11. The goal of any Koopman algorithm is to approximate $\mathcal{K}$ and to extract its eigenfunctions and eigenvalues to improve the understanding of the system dynamics and future state prediction or control. Approaches to this are growing in popularity, with many of the most common approaches stemming from

two related techniques. As the accompanying theory are particularly relevant to the work of this dissertation, it is left for more thorough review, analysis, and discussion in Chapter 4.

Proper Orthogonal Decomposition

Although not an approach to approximate the Koopman operator, proper orthogonal decomposition (POD) is a related modal decomposition technique and is also a fundamental component of many Koopman techniques. POD has a long history, going back to 1901 with Pearson [91] who established principal component analysis (PCA). Since then, PCA developed into the Karhunen-Loève transform [92, 93], the Hotelling transform [94], empirical orthogonal functions [95], and then Lumley [96] provided the first introduction to POD. All of these techniques are approximately identical, up to nomenclature and mean-subtraction variations. In addition, while there are a number of methods for calculating POD, it shares the same formulation as the singular value decomposition (SVD) and, hence, it is valid for any data matrix. While the SVD can be computed very efficiently, the data matrix is usually prohibitively high-dimensional with great redundancy and is, thus, limited by computational expense to even modest datasets [88].

POD was established to address the need of understanding turbulence [96]. POD finds the orthogonal low-rank structures which optimally capture the entire system energy from a time-series of spatially-distributed measurements, e.g., time-resolved PIV or high-speed video data. These low-rank structures are captured through the POD modes, with each successive mode capturing the highest variance dimensions which are orthogonal to the previous modes. The orthogonality constraint ensures energy optimality at the cost of wide power spectral densities (PSDs).

POD's first application in fluid mechanics was in representing large eddies in turbulent velocity fields [96], which was subsequently applied to extracting physical information of eddies within the viscous sublayer of turbulent pipe flow [97]. These initial works demonstrated POD's ability to identify streamwise velocity fluctuations, attributed to a low effective Reynolds number close to the wall, for the analyzed flow. Lower energy modes, however, were not attributed to any physical process, owing to a lack of accuracy

and computational concerns. The original formulation was later refined by Sirovich [98] in the popular POD snapshot method. Many formulations exist but, ultimately, the application and results remain constant.

Decades after its first introduction, POD's popularity began to increase. Berkooz et al. [99] revisited the applications of POD to turbulent flows more generally. They discuss conditions for modal energy content to be meaningful in relation to POD's PSD but, significantly, they conclude that an *a priori* physical understanding of the system is essential for a successful application of POD. An early application of POD to simulation was given by Kirby et al. [100] in the study of supersonic shear layers. They demonstrate POD's utility at data compression, as it is able to capture large-scale features of the simulated flow with orders of magnitude decrease in the dataset size. While physical meaning is attributed to extracting POD modes, they acknowledge the inability to directly extract physics, as the POD formulation does not generally adhere to the governing flow equations. Additionally, time-dependence is removed from POD which is a considerable detriment to an analytical tool for highly dynamical systems. As will be a common theme, they recommend an adapted POD formulation to maintain the spatio-temporal relationships of the data.

Since the above early applications, a wide variety of fluid and atomization systems have been analyzed using POD. Annular jets were analyzed by Patte-Rouland et al. [101], where the most dominant modes were considered as probable realizations of the flow. The most dominant mode, as is commonly found with POD's application, captures the mean flow, while the remaining modes are interpreted as capturing pairs of vortical structures in the flow. Jets in crossflow have also been studied with POD [102, 103]. POD was used to identify breakup mechanisms, wake vortical structures, and jet penetration fluctuations with different degrees of bending. While the most energetically dominant modes are often discussed the most, considering their relevance from a system energy point of view, higher order modes (i.e., modes with lower energy content) are also used to facilitate discussion. The POD modes were also used to extract system wavelengths and frequencies attributed to upstream Kelvin-Helmholtz instabilities [103].

Indeed, the application of POD has found extensive use for more specific atomization-related phenomena. Recently, the morphology and dynamics of immiscible liquids emanating from a co-axial

injector was investigated [104], POD was used to aid flame dynamic understanding from an industrial fuel nozzle [105], cavitation processes were identified in a sharp-edged orifice using POD [106], and vortical structures governing the flow of cryogenic liquids injected at transcritical conditions have also been captured [107]. Further, owing more to its ability to compress data rather than extract physics, POD has been used as a pre-processing tool for image analysis of atomization systems [108].

While many works have been conducted for many different, but related, applications, most of the results are interpretive and qualitative. Specifically, physical interpretation is often given but is not sufficiently validated. These concerns were noted in the early applications of POD, yet its use for fluid flows continues extensively.

Dynamic Mode Decomposition

Dynamic mode decomposition (DMD) is a recent modal decomposition technique, introduced by Schmid [88], which approximates the Koopman operator by assuming the system evolves linearly [88, 89, 109, 110]. It was formed to overcome the wide PSDs and temporal independence of POD by relating extracted spatial modes with a single oscillatory frequency with exponential growth or decay. DMD and its variants build upon POD and can be considered an extension of POD with Fourier analysis (see, e.g., [89]). Schmid [88] discusses how POD attempts a decomposition based on orthogonality in space whereas DMD does this in time. It provides an approximation of the Koopman eigenfunctions, which is exact when the observable function **g** is just the identity function. DMD can be considered an alternative to the Arnoldi method for calculating the eigendecomposition of a linear operator. It is noted to be the first decomposition technique to provide temporal dynamical characteristics with spatially coherent structures for experimental data.

DMD and POD have become ubiquitous tools in the analysis of nonlinear dynamical systems, even though they are linear analysis tools. The seminal work on DMD [88] highlights differences between POD and DMD and provides example applications, and whose follow up work gives insight to extracted

representative dynamic modes for fluid flow cases [111]. Detailed differences of the POD and DMD formulation are explored in Chapter 4.

DMD also has a close relation to the discrete Fourier transform (DFT), beyond its ability to extract oscillatory behavior. Chen et al. [112] showed that when data are first mean-subtracted, DMD reduces exactly to the DFT. This is clearly restrictive as, in its fullest generality, DMD is able to extract oscillatory dynamics which can also be coupled with exponential growth or decay. Subtracting the mean of the data, while a common approach in POD [113], has the consequence of also removing important dynamical information. Mean-subtraction is therefore not encouraged for most DMD studies.

Despite the linearity assumption, DMD has been applied extensively to numerous problems in an attempt to extract physical understanding of a system, such as predictive modeling [114, 115, 116], disease modeling [117], nonlinear control [109], and a wide range of fluid flow problems [80, 111, 77, 118]. Specifically, for liquid injection and atomization, which demonstrate wave-like oscillatory transport and breakup mechanisms even in complex flows [4], DMD appears well-suited to aid in the understanding of the underlying processes due to its relationship with Fourier analysis.

There are numerous example applications within atomization. An early DMD study is that by Rowley et al. [80], who applied DMD to a simulated jet in crossflow, and discuss the differences between POD and DMD results. While POD extracts similar modal structures to DMD, it assigns numerous peak frequencies for each mode, given by its PSD, which hinders insight into the temporal evolution of the corresponding modal structures. DMD, on the other hand, decouples different frequency components, whose modes capture shear layer vortices and the shedding of wall vortices.

A recent application of DMD to a simulated axial jet suggests that DMD is not always able to fully separate independent dynamical behaviors [119], even when modal structures appear to be intuitive. The wavelengths, frequencies, and growth rates of prominent instabilities are identified, though parts of the discussion focus on higher order modes which should not contribute significantly to the system's energy.

Hua et al. [77] used DMD in the analysis of shear coaxial jets and discussed POD's inability to differentiate low-energy robust flow constituents from noise. Further studies have identified cyclic

variations of pulsing fuel injectors [120], unsteady internal flow structures and surface wave instabilities from pressure-swirl atomizers [121, 122] and annular liquid sheets [123], and the morphology and dynamics of emulsions from a plain orifice [124]. Example extracted DMD modes from [122] are given in Figure 2-10 which purport to extract surface instabilities and their associated frequencies. As with POD, many results found with DMD have yet to be validated and are interpretive.

Due to the overlapping relationship of POD and DMD, many studies have directly explored their respective efficacies comparatively in a number of areas. Huang et al. [125] compared POD and DMD on the analysis of simulated self-excited longitudinal combustion instabilities. DMD was found to be more useful at providing system modes that correspond to frequencies relating to pressure and heat-release. Tirunagari et al. [126] applied POD and DMD to data from a large eddy simulation of injected gaseous jets at subsonic conditions and were able to compare the techniques' efficacy at identifying Kelvin-Helmholtz instabilities. POD is preferred for preserving behaviors corresponding to high kinetic energies and more global structures, whereas DMD focuses on finer scale details.

Additionally, a comparative analysis of POD and DMD in capturing the dynamic evolution of flow past a marine propeller has been performed [78], differences between the techniques in extracting acoustic resonance behaviors from a compression-ignited combustion chamber have been investigated [79], and extracted coherent flow structures have been compared for the wake of a high-speed train [127]. Hua et al. [77] investigated the application of DMD on shear coaxial jets and looked into the robustness of extracted modes, following on previous work by Roy et al. [128]. These works note the folly of POD and DMD modes but with their respective advantages and acknowledge the lack of verification in successfully differentiating between true underlying system modes and inherent system noise.

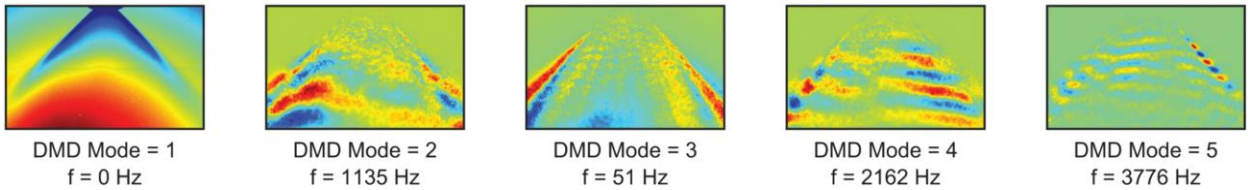

Figure 2-10: Extracted modes from [122], capturing surface wave instabilities and frequencies of a pressure swirl atomizer system.

It is important to remember that these techniques are inherently linear techniques. As the systems of interest are multiphase, turbulent, and high-dimensional, the underlying physics is highly nonlinear. Even with these techniques being commonly applied to experimental data of interest, no new validated physical insight has been found, potentially due to applying linear analytical tools to nonlinear processes.

2.2.5 Modal Interpretation

Certainly, there are a plethora of works which have applied POD, DMD, or both in parallel, especially to nonlinear systems. Interestingly, in certain cases, the agreement of DMD with POD is used as validation that the modal decompositions are extracting meaningful results, irrespective of their close relationship in theory. Although POD and DMD are becoming more prevalent in the literature, there are limited works which focus on the physical interpretation of extracted modes, seek to validate this interpretation, or identify decomposition dependence to user parameters, such as the number of data points used for analysis. As a result, there is a risk that interpretations of results, or the results themselves, do not represent true underlying system processes.

Physical interpretation of POD modes has been discussed by Kerschen and Golinval [129] for structural dynamics problems where a focus on linear systems revealed that POD will converge to the system eigenmodes. For the nonlinear case tested, however, no intuition was given on what the best linear representation means physically. Similarly, Feeny and Liang [130] interpreted POD modes of a linear vibration system by correlating them with the true vibrational modes under random excitation. Indeed, good agreement in the POD modes and the normal modes of the system have been found for linear systems, but greater difficulty arises through the linearization of nonlinear systems.

Chen et al. [131, 132] noted the difficulty in interpreting POD modes for nonlinear systems, particularly when comparing the likeness between two datasets as POD energy distributions or spectra are not a good correlation metric. Further, they demonstrate the highly coupled nature of processes in the system as all flow structures in every data matrix snapshot contributes, to some extent, to every POD mode. Arienti and

Soteriou [103] conducted an investigation of POD for liquid injection cases. While a large number of modes and their interpretations are given, this is not the case for all modes where some are left out of the discussion. As is commonly found in practical applications, certain modes are particularly easy to discuss, while others are not. Distinguishing between a valid mode or noise can be a significant challenge [133], particularly when the data are resolution-limited.

Of the large number of papers which have implemented both POD and DMD, there is still not a sound understanding of what extracted modes represent and how to interpret them in a validated way for nonlinear systems. Interpretative claims, in most cases, have not been validated which hinders the progression of understanding flow processes. It is vital for these models, and extracted modes, to be interpretable to truly improve physical understanding of the system. To this end, simpler systems with multiple processes with known spatio-temporal coherence can serve as an initial validation step prior to application of more complex systems.

2.2.6 POD and DMD Variants

While there have been numerous successes with implementing both POD and DMD, many works rely on adapting or supplementing the standard algorithms. Of course, this has led to significant algorithmic progress being made in Koopman analysis, with many variants being successfully implemented [134, 112, 135, 136, 137, 86, 138]. However, a more general approach which performs well on many problems is desirable. Instead, researchers have often adapted algorithms specifically for their problem of interest as the onus lies on them to identify the technique which is most appropriate for their dataset. While prominent and relevant alternatives are discussed here, readers are directed to the recent reviews of POD and DMD, and similar techniques, provided by Rowley and Dawson [139] and Taira et al. [140] for an in-depth overview of modal decomposition techniques.

A variant of POD specifically for dynamical fluid systems is the balanced POD (BPOD), first introduced by Rowley [141], which is designed to be an approximation of balanced truncation [142] which becomes

intractable on high-dimensional data. Stemming from control theory, BPOD attempts to extract modes which are both highly controllable and observable. The weakly controllable or observable modes can then be truncated. Error bounds for BPOD can be set *a priori*, and this approach can overcome POD's sensitivity to the empirical data used and inner product definition. However, this technique has not seen much attention in the fluid dynamics community.

Attempts have been made to incorporate frequency information into POD creating a *spectral* POD algorithm (SPOD). There are, however, at least two procedures both called SPOD which attempt this in different ways. Sieber et al. [143] take the data correlation matrix, which is discussed in Chapter 4, and apply a filter which smooths the values along the diagonals of the matrix. By varying the strength of this filter, SPOD can vary between a pure POD and a pure DFT to find the data augmentation which extracts modes whose spatial structures and frequencies are easily interpretable. This approach forces structure into the raw data matrix so it becomes difficult to ensure the extracted results are truly a consequence of the measured data, and not of the applied filtering process. The other procedure, where a recent discussion on POD and SPOD is given by Towne et al. [134], tries to capture space-time coherence in the data, rather than the space-only coherence captured by POD. The main change is in the definition of the inner product for the correlation matrix which incorporates a sum over the temporal extent of the acquired data, not just the spatial extent.

Many variants exist to overcome concerns with applying DMD to nonlinear flows, as it can be ill-suited in its standard form. For example, Alessandri et al. [144] demonstrated the need to modify the standard DMD algorithm to improve the detection of stable to unstable, or unstable to stable, flow regime changes. A higher-order variant of DMD has been shown to extract salient flow patterns and improve on the standard algorithm's extrapolation performance [145], and a composite-based DMD, utilizing snapshots of different variable distributions, was shown to capture Reynolds stresses in turbulent channel flows [146].

While not necessarily an issue with high-speed video data, Brunton et al. [147] provided a DMD approach to compressive sensing where requirements of complete and uniformly sampled data are relaxed. This is important for systems where spatial measurements are affected by occlusion or the probing

diagnostic is subject to non-uniform time-sampling. This technique is particularly useful in two scenarios. First, if the full system of interest is only sub-sampled, it is possible to approximate the full-state DMD through what is called *compressive-sampling* DMD. Second, if information on the full-state is available, *compressed* DMD may be performed where the full-state is first compressed prior to applying DMD. Also borrowing ideas from compressive sensing, Jovanović et al. [137] suggest using a *sparsity-promoting* DMD to get the best trade-off between quality of approximation and the number of modes required for a suitable approximation. To achieve this, a regularizer term is added to the approximation cost function, which is typically the L_2 deviation between the original data matrix and the DMD reconstruction. By including an L_1 regularizer, the cost function is then weighted in favor of reconstructions using fewer modes.

Optimized DMD [148] was developed to overcome DMD's sensitivity to noise while also having improved ability at extracting hidden dynamics. This added robustness is balanced with an increased computation cost, which may increase by 50%. The approach rephrases DMD into an exponential data fitting problem through the variable projection method which considers all snapshots of the data at the same time. This also removes any bias from the original DMD formulation and is recommended for use by the authors over the standard DMD formulation.

While DMD is an approximation of a Koopman decomposition with an observable function, **g**, equal to the identity, Williams et al. [86] generalize this notion to a dictionary of potential scalar functions with *extended* DMD (EDMD). A dictionary of potential functions is defined *a priori*, and the EDMD algorithm identifies which combination of functions yield the best finite-dimensional linear representation of the nonlinear system. This gives DMD greater flexibility to generalize to a variety of problems, but this is necessarily accompanied by a much greater computational burden. As the dictionary must be defined *a priori*, there is also the problem of how to select the appropriate set of candidate functions. Dietrich et al. [138] demonstrate how to learn a suitable dictionary using a neural network to make EDMD a fully data-driven method.

2.3 Deep Learning

Prior discussions into the background of liquid atomization system complexity, traditional dynamical analysis for these systems, and the conventional diagnostics used to probe these systems have highlighted the continuing difficulty in successfully understanding and predicting their behaviors. Growing interest has been towards Koopman analysis: a data-driven approach for linearizing nonlinear systems which can also learn dynamical features directly from data. These approaches fall into the category of machine learning; an algorithmic technique which learns to optimize a desired cost function automatically and improves with more data. However, while successes have been demonstrated for POD, DMD, and variants thereof, each algorithm either struggles with computational cost, in generalizing to many disparate datatypes and systems of interest, or in validating its efficacy on nonlinear systems.

Koopman analysis is just one such machine learning technique. Recently, a vast array of machine learning approaches has emerged which have been shown to effectively analyze large datasets from various fields including, but not limited to, finance, business, computational biology, healthcare, and signal processing, wherever it can learn from large datasets in order to elucidate patterns and make predictions. Recently, *deep learning*, a subset of machine learning, is a specific approach leveraging neural networks and has made tremendous progress in almost every field it has been applied in. The great utility of deep learning is in its rich representational capacity; it is a universal function approximator. As it learns directly from data, no prior information about the system of interest is required.

The flexibility and generality afforded by neural networks in system analysis can be used to overcome concerns with existing Koopman approaches. This section introduces the most relevant deep learning background required to present the deep learning contributions in this work.

2.3.1 Learning

All machine learning algorithms *learn* directly from data. Data are collected about a system with many different examples or snapshots in time and are split into *training* and *test* datasets. A further *validation*

dataset is also created, serving as a proxy test dataset while training a given model. The selected machine learning model optimizes its model parameters, i.e., it *learns* the optimal model parameters, to minimize a loss function. This follows the schematic presented in Figure 2-7.

Often, the model has both model parameters and hyperparameters. The former is what is learned automatically by minimizing the given loss function. Hyperparameters, in comparison, are manually selected and govern the specific model configuration or design, thus, are not learned directly. Instead, experimentation is done using the training and validation data, either manually or programmatically, to perform a hyperparameter search to identify the optimal model design [149, 150]. Hyperparameter optimization is not necessary for all machine learning approaches, but it is vital for neural networks.

Learning of the model parameters is often done analytically for simple algorithms, otherwise an iterative approach is taken. A basic example is the use of the normal equations to analytically optimize the parameters for the generalized least squares method. Similarly, both POD and DMD have analytical solutions due to their linear formulation; only linear algebra computations are required. In contrast, iterative methods update the model parameters incrementally, often through a gradient descent algorithm to traverse the loss surface.

Two of the most prolific types of learning methodologies are supervised and unsupervised learning. In supervised learning, algorithms are trained on inputs with known outputs such that the algorithm is then able to classify or make predictions on input data with unknown outputs. Unsupervised learning, also called self-supervised learning, however, does not require matching input-output data. Instead, it takes large datasets and identifies characteristics and salient trends present without known outputs. Achieving high performance with unsupervised learning algorithms is generally more sought after than with supervised learning algorithms, as relying on hand-labeled training data can be prohibitively expensive [151].

Learning without labels has great implications in engineering as much of the data collected are impossible to label due to the desired network output not being known *a priori*. Consider the goal of extracting physics from an observed fluids system. Many processes and behaviors dominate the system

dynamics of which only a few are known analytically, others are understood qualitatively, while many others are obfuscated by system complexity. As a result, no known output data exist.

2.3.2 Neural Networks

Recently, neural networks (NNs) have seen a tremendous increase in interest and application due to a coupled increase in data acquisition with computational power. While POD and DMD learn the best linear representation of a dataset, NNs relax this constraint by allowing the network to find the best nonlinear representation. Using a NN framework for Koopman analysis, therefore, offers much greater flexibility in identifying dynamical features compared to POD and DMD.

There is no shortage of areas where NN-based models have performed at or exceeded state-of-the-art levels. While specialized architectures do exist, the prototypical NN is a *feed-forward* design which, in theory, is a universal function approximator [152]. Of course, in practice, alternative designs are found to perform better or be more robust for specific problems. Image-data are best probed by convolutional neural networks (CNNs) which have excelled in object recognition and classification, regardless of noise or occlusion. Recurrent neural networks (RNNs) are predominantly used on sequential data that can learn to remember salient features over a certain sequence interval. Models can be either deterministic or generative, where generative models learn the underlying data distribution of input data which can then be used to predict or generate further data samples.

Deep learning has seen great success since the beginning of the 2010's, despite active research into neural networks for decades prior (see, e.g., [153, 154, 155] for early developments). Much of the recent progress has been attributed to leveraging large datasets and in creating *deep* neural networks; networks which have many hidden layers. Early insight into the performance of neural networks for image recognition was given by Aizenberg et al. [156], prior to the watershed image classification performance on the ImageNet dataset by Krizhevsky et al. [157]. Since then, deep learning has proven successful in medical diagnosis [158], high-energy physics [159, 160, 161], genomics [162], finance [163], as well as

countless other areas. In many of these areas, NNs have quickly outperformed the previous state-of-the-art often leading to new, fast, and inexpensive predictive or diagnostic tools.

Feed-Forward Neural Networks

The basic neural network design is that of a feed-forward neural network, with an example 5-layer neural network given in Figure 2-11. The network is composed of neurons, arranged in layers, with each neuron taking in inputs, calculating outputs, and passing the outputs in one direction. The input layer has the same number of neurons as the dimension of the input data, likewise for the output layer and the output data. All layers in between are *hidden* layers, as they only have access to latent variables. For clarity, the network design in Figure 2-11 is arbitrary; the number of inputs and outputs are defined by the specific problem.

Each neuron performs the same basic calculation, given by

$$o_i = \sigma\left(\mathbf{w}^{\mathrm{T}}\mathbf{x} + \mathbf{b}\right) \tag{1.14}$$

where the scalar output of neuron i, o_i, is a function of the dot product of the inputs into that neuron, $\mathbf{x}$, with the associated weights, $\mathbf{w}$, plus a bias term, $\mathbf{b}$. The σ is the activation function of the neuron, which is used to introduce nonlinearity into the model to provide the rich representational capacity. Common choices

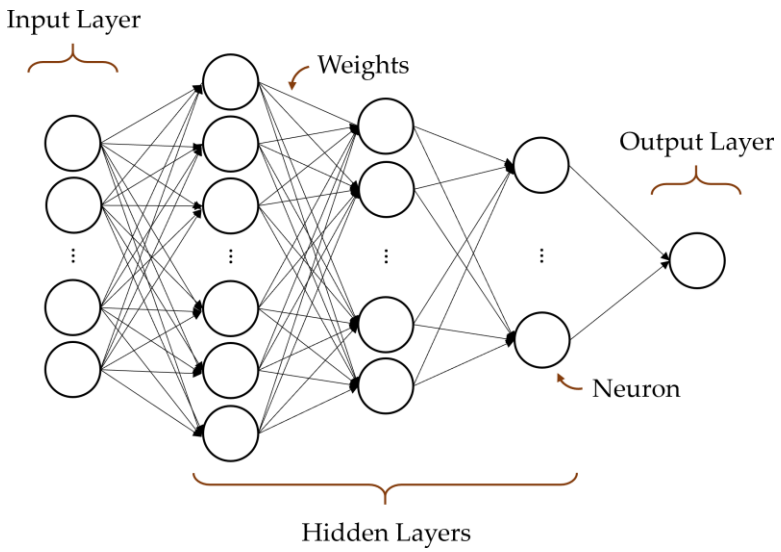

Figure 2-11: An example feed-forward neural network predicting a single output value.

of the activation function include the rectified linear unit (ReLU) [164], the sigmoid function, and the hyperbolic tangent.

The weights and biases, $\mathbf{w}$ and $\mathbf{b}$, respectively, are the model parameters which are optimized to minimize the chosen loss function. The choice of loss function is dependent on the formulation of the problem, producing a loss $\mathcal{L}$. For regression problems, the mean squared error (MSE) is commonly used, given by

$$\mathcal{L} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2 \tag{1.15}$$

where $\tilde{y}_i$ is the prediction of the true i^{th} output, y_i, which is averaged over n outputs.

For each training iteration, a batch of input data are used to update the weights. The inputs propagate through the network according to Equation 1.14 until a prediction is made which produces an error, $\mathcal{L}$, on that batch. The error is used to update all of the weights in the network through backpropagation [165, 166]. In a simplistic view, this is done by calculating the contribution of each weight to the overall error by calculating its gradient, $\frac{\partial \mathcal{L}}{\partial w_i}$, and updating each weight using gradient descent,

$$w_i = w_i - \alpha\frac{\partial \mathcal{L}}{\partial w_i}, \tag{1.16}$$

where α is the learning rate. In practice, a variation for updating the weights is used, of which stochastic gradient descent is core. Momentum is often incorporated, whereby previously calculated gradients are used to inform the present calculation. Commonly used variants include AdaGrad [167], RMSProp, and Adam [168]. The learning rate, α, is a significant hyperparameter in neural network training. As training progresses, the general weight update form in Equation 1.16 often uses a decaying learning rate to help model convergence.

Convolutional Neural Networks

While feed-forward NNs are universal function approximators, they do not always perform well on image and video data in their standard configuration. Referring back to the feed-forward design in Figure 2-11, feed-forward NNs will first flatten a 2D input, such as an image, into a 1D vector. This design, therefore, does not take advantage of the spatio-temporal topology of image and video data as the neurons in a given layer are independent of each other. This is often useful when no *a priori* knowledge of the input is available to the network designer, but this is not the case for image or video data.

Convolutional NNs (CNNs) were designed specifically for image and video data which leverage the local neighborhood dependency of pixels in the data. They have famously been applied to image data [157] and video data [169] and related tasks [157, 170, 171, 172], for the purposes of object recognition, tracking, and prediction, or in tandem with deep autoencoders for compression and low-dimensional representations of image and video data (e.g., [173]).

Instead of a regular fully connected network in the feed-forward NN where the weights between neurons are optimized, CNNs instead optimize filters; small grids of weights, imitating a small image of pixels, which convolves over the entire input image, see Figure 2-12. Each filter in a convolutional layer is applied to the input image and produces a feature map in the next layer, where each feature map can be considered the new input image for the next convolutional layer. In Figure 2-12, a 3×3 filter is applied to an 8×8 input image, producing a 6×6 feature map, when no input padding is used. The filter strides across both

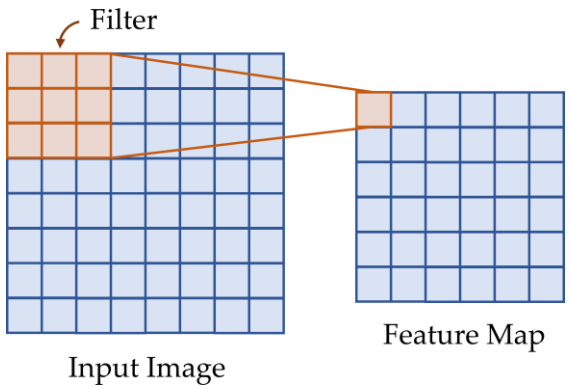

Figure 2-12: A 3×3 filter convolves over an input to produce a feature map in a convolutional layer.

input dimensions, where a stride of 1 is used in this example. Padding is commonly added around the input image for the convolutions, so the feature map dimensions match the input image dimensions. The basic convolution computation is identical to Equation 1.14; the weight vector is now the filter weights and the input vector is the corresponding neighborhood where the convolution is applied.

While the convolution process captures short-range dependencies in the data, the network also requires long-range dependencies to be captured so general features and behaviors can be extracted. To encourage this, convolutional layers are often paired with pooling layers. These pooling layers downsample the input image or feature map by only preserving a characteristic value of a given region of the partitioned image. The pooling also uses a filter of a given size and stride, but this is typically set to a 2×2 filter with a stride equal to the size of the filter to give non-overlapping regions which downsamples by a factor of 2. For example, in max pooling with a 2×2 filter with a stride of 2, the image is first partitioned into 2×2 non-overlapping blocks where only the maximum value of each block is kept for further computation.

Very often, the goal of a CNN is in image classification or recognition. Convolutional and pooling layers form the core of the network architecture, but they are, generally, not sufficient by themselves. After the convolutional and pooling layers, the final feature maps are flattened, and fully connected layers, such as in the feed-forward design, are used to map to the correct output dimension. This design naturally constrains the CNN to have a fixed input size to ensure all subsequent layers operate on the same dimension size. This constraint can be relaxed through, for example, the use of spatial pyramid pooling which operates on proportions of regions rather than fixed dimensions [174]. This has been successfully applied to a range of applications, in particular in region proposal networks for proposing bounding boxes to localize and detect objects [175, 176, 177].

Autoencoders

Techniques related to principal component analysis (PCA), e.g., POD, have been the most developed and used unsupervised learning algorithms for combustion research. An alternative, deep learning-based

approach, is to use deep autoencoders to convert high-dimensional data to low-dimensional representations which can have improved performance over PCA [178]. An autoencoder is an encoder-decoder network which learns a low-dimensional representation of inputted data; an encoder network learns to transform high-dimensional data to a low-dimensional latent space, or embedding space, where salient patterns are extracted, before a decoder network attempts to recover the original high-dimensional input.

The autoencoder can be thought of as a generalization of PCA, as it can perform nonlinear mappings unlike PCA approaches [179], thus it is an architecture of great interest in extending existing POD and DMD algorithms for engineering datasets. Note, the previous descriptions of feed-forward and convolutional NNs are very general. In their typical use cases, supervised learning is used and the output dimension is predetermined by the problem being solved. For autoencoders, as with PCA, no known outputs accompany the inputs, hence the autoencoder tries to reconstruct the given input in its output layer. Autoencoders are central to the development of a Koopman-based deep learning framework, and more detailed descriptions are given in Chapter 4.

Autoencoders have been used for image compression, an example of dimensionality reduction, where they enable the efficient storing and transmission of large image sizes [180]. Further, Le [181] used deep sparse autoencoders to produce a face detector on a large image dataset of random images. Le found, contrary to widely-held intuition, an unsupervised learning algorithm was able to achieve this, regardless of face scaling, rotation, or position. Unsupervised learning circumvents the need for labels and classifies features autonomously. It is possible to develop high-level features from unlabeled data, giving credence to the possibility of identifying low- and high-level features from fluids-related image datasets using these algorithms.

2.3.3 Neural Network Design

Beyond the fundamental components presented above for feed-forward NNs, CNNs, and autoencoders, there are a tremendous number of adaptations to network design, all striving to improve generalization or

model performance for specific tasks. Every week a new design or feature is presented, oftentimes superseding what was previously state-of-the-art. Acknowledging this, relevant adaptations pertinent to the tasks of this dissertation are presented here which are attractive design choices *as of writing*. While not all of the following discussions are explored or implemented directly in this work, their utility is particularly relevant for future work.

Regularization and Normalization

The above provides the fundamental concepts which are needed to successfully train a neural network. In practice, there are many variants and additional tools used to improve training convergence and overall performance. This is a primary focus with deep learning as, to take advantage of large swathes of data, increasingly large networks are designed to ensure the model has sufficiently high representational capacity.

Consider the relationship between test error and model complexity in Figure 2-13. The Pareto front provides a lower bound on the minimum test error achievable given a model's complexity, where the complexity can be considered the number of parameters in the model. Of course, there are diminishing returns as model complexity exceeds the true, underlying complexity of the problem being solved. What is desired is a parsimonious model; the least complex model whose performance does not improve significantly as complexity is increased. This model is said to be Pareto optimal.

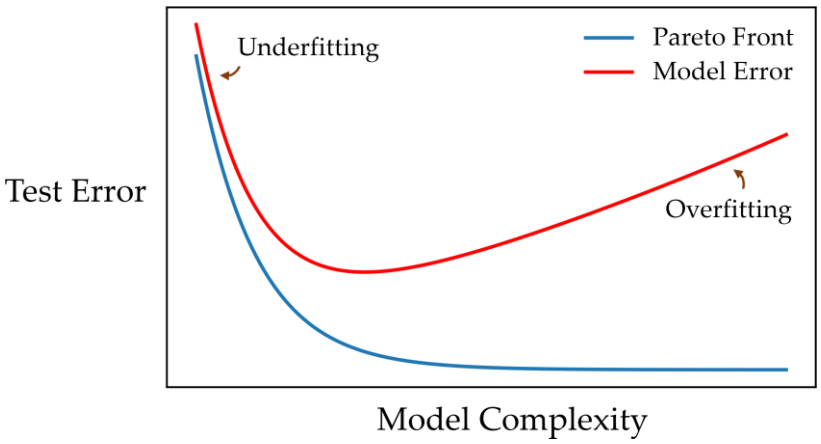

Figure 2-13: Underfitting and overfitting are significant concerns in training NNs, which prevent models from approaching Pareto optimality.

One of the greatest concerns with deep learning is in overfitting; due to the high number of parameters available to the neural network, the network is often able to memorize the training data without learning the underlying truths of how the data are generated. This is the cause for the increase in model error as complexity increases in Figure 2-13. Similarly, underfitting occurs when the network has not been trained to sufficiently high performance or whose representational capacity is too low for the problem it is solving. In theory, underfitting is trivial to solve as additional neurons or layers can be added *ad infinitum* until its representational capacity is sufficiently large. However, as the dimension of the parameter space increases, it becomes much easier to get trapped in saddle points [182].

A number of techniques and heuristics have therefore been introduced to improve training properties and to regularize the model; approaches to the latter aim to reduce overfitting and, thus, generalization error. One of the most popular regularization procedures is the use of dropout [183, 170]. During training, neurons are stochastically dropped from the network to prevent the co-learning of neuron groups. A more intuitive regularization technique is early stopping. By monitoring training and validation error as training progresses, insignificant improvements in model performance can be observed which indicate sufficient learning has been achieved. Any further learning is susceptible to overfitting.

Pre-training, or transfer learning, may also be used for regularization purposes. The core idea is to reuse network weights which have been applied successfully to a similar problem. For example, in many image-based problems, low level features are shared across datasets, such as edges, corners, or gradients. By leveraging the learned feature detectors, i.e., filters, from a related problem, the apparent training dataset size has increased, and more general features are used.

However, interestingly, He et al. [150] demonstrate that pre-training versus random initialization, on the ImageNet dataset, does not affect overall performance once fully trained. Additionally, while pre-training gives faster convergence at the beginning of training, additional fine-tuning to get to peak model performance was found to take the same length of time as training from random initialization. Their findings hold for multiple dataset sizes and show how ImageNet pre-training does not automatically give improved regularization.

Finally, normalization procedures are often taken to make input or latent data more amenable to learning. Batch normalization is a common technique to maintain inputs with statistics that are beneficial for training [184]. Unfortunately, the accuracy of batch normalization severely degrades when applied to small batch sizes [150, 185]. This is particularly troublesome for applications which typically act on high-resolution images or video data. An alternative is to use group normalization which is invariant to changes in batch size [186]. Reducing the batch size can increase the model error dramatically when using batch normalization, but this is not an issue for group normalization. Other alternatives do exist, such as layer normalization or instance normalization, however, they have not approached the accuracy of group normalization in many visual recognition tasks.

Skip Connections

The discussion on the basic feed-forward and convolutional NN designs have had layers feed unidirectionally into the following layer. Neural network design has few topological limitations, although here we consider only acyclic networks, and so many works have demonstrated the use of so-called skip connections, where the connections from one layer may feed into any downstream layer.

A highly successful use of skip connections is that of residual networks, or ResNets [187]. It is motivated by when networks are tasked to learn the identity, which is a trivial task for a network with any number of parameters, it becomes more difficult as the network gets deeper. Performance is found to deteriorate as the number of layers increase beyond a certain threshold. Residual networks overcome this limitation by preconditioning on the idea that, in the worst case, deeper redundant layers should learn an identity mapping. To achieve this, instead of a layer attempting to model $\mathrm{H}(\mathbf{x})$, where $\mathbf{x}$ is the layer input, it tries to model the residual $\mathrm{H}(\mathbf{x}) - \mathbf{x}$ in a so-called residual block. The input is both fed into the residual block and directly into the output to give each residual block a reference point. By including the input, the block is tasked with learning the residual which improves the overall mapping performance and learning the identity becomes trivial.

Taking the idea of passing inputs into the output of blocks further, DenseNets are similar but instead pass all of the previous feature maps, with the restriction of feature maps from convolutional layers of the same dimension [188]. This greatly increases the number of connections in the network thereby creating a densely connected network.

Dilated Convolutions

In many image and video problems, the presence and dynamics of features exist on a range of spatial scales. To successfully identify all features from short- to long-range information (i.e., pixels close and far away from each other, respectively), a network must account for these different scales. The most prevalent technique for this is to use an image pyramid: the original image is downsampled n times, such as through pooling layers for CNNs. Each successive downsampling loses finer details but allows coarser details to be identified. For regular convolutions without a subsequent pooling layer, downsampling is avoided to maintain high feature map resolution, but the receptive field of each filter will always be small throughout the network and will be unable to capture long-range dependencies [189].

Instead of downsampling the feature maps to identify coarse features, dilated convolutions, also called atrous convolutions, maintain the same input resolution but instead dilate the applied filter, i.e., the filter is padded with zeros, to capture coarse features. This is demonstrated in Figure 2-14, where a dilation rate of 2 is used, effectively transforming a 3×3 filter into a 5×5 filter at no extra computational cost. Note, a dilation rate of 1 is identical to a regular convolution. In this case, it is necessary to pad the input with zeros to prevent any downsampling. Successively higher dilation rates can be used to build up coarser features as the depth of the network increases.

Dilated convolutions provide *dense* feature extraction whereas regular convolutions provide *sparse* feature extraction [190]. Extracting dense features is desirable as it allows each pixel in the input to be mapped directly to a discrete or continuous label, rather than trying to infer or interpolate values from sparse features. Consider the task of image segmentation, for example, where the goal is to segment the original

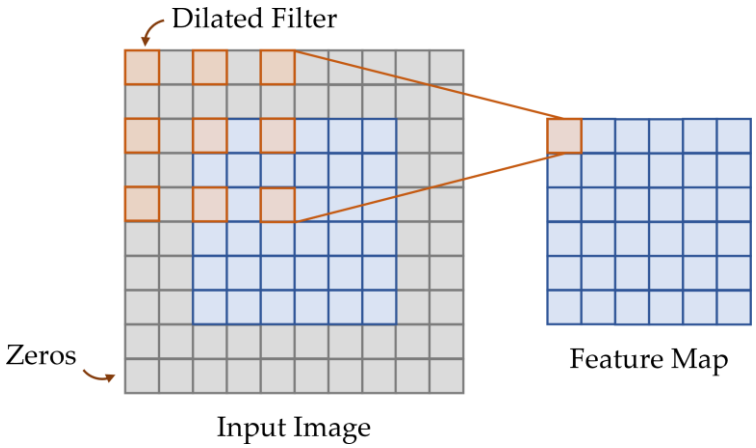

Figure 2-14: Dilated convolutions pad a regular convolutional filter to increase its receptive field over high-resolution inputs.

image at full resolution. For a regular CNN, upsampling must be performed, often through deconvolution, but finer details have been lost through downsampling. Hence, this approach will produce sparse feature extraction. This downsampling is necessary to reduce the number of high-level parameters, which grows exponentially, and allows for multi-scale feature extraction. Instead, dilated convolutions use sparse filters which provide dense feature extraction on the high resolution image directly. In addition, although full resolution feature maps are maintained unlike the regular case, this does not actually introduce more parameters or operations. Not all of the zero-multiplications are computed, instead, each filter weight is strided at the dilation rate in the original image.

Yu and Koltun [191] use dilated convolutions to overcome the sparse feature extraction of regular CNNs, which is more suitable for image classification problems. The authors use dilated convolutions for image segmentation, as it is able to aggregate contextual information at multiple scales without a loss of resolution.

A drawback of using dilations, however, is that they produce grid-like artifacts, since the convolutions are done sparsely. There are a number of approaches for reducing the artifacts produced by dilations (see, e.g., [192]), where the dilated convolutions can be directly smoothed, or by first increasing the dilation factor in the middle of the network before decreasing the dilation factor at the end of the network.

Physics-Informed Design

Neural networks have seen extensive application and there have been a large number of variations presented in literature. Many of these works are able to achieve state-of-the-art performance, yet are unable to soundly explain the rationale behind architecture design or gain an understanding of the effectiveness of their networks. While neural networks were developed to reduce dependence on human understanding of a system, physics-informed designs, which incorporate inductive biases into the network, are growing in popularity to promote interpretability and understanding of model behavior.

Chen et al. [193] investigated a method of encouraging features to be interpretable through their so-called InfoGAN: an information-theoretic extension of generative adversarial networks (GANs). By incorporating a mutual-information cost, latent codes (generator features) can be forced to be similar to the desired output. This causes features of little semantic significance to be rejected as they have poor correspondence to input or output images. As latent codes change discretely, there is clear semantic meaning behind what each code represents. This is achieved at no extra computation cost and is something which can provide greater understanding to how the model is operating.

Deep learning for identifying and learning partial differential equations (PDE) have seen widespread interest (e.g., [194, 195, 196, 197]). In these works, PDEs are used in the cost function of networks to incorporate prior scientific knowledge, or NNs are used directly to learn the structure of initial value problems of a form such as

$$u' = \mathrm{NN}_\theta(u, t), \tag{1.17}$$

governed by the NN with parameters θ. The generality afforded by NNs, either as a subcomponent or as the entire model, has shown promise in turbulence modeling [198] and in overcoming computational constraints of DNS for practical, 3D turbulent flows [199].

2.3.4 Related Deep Learning Applications

Liquid Atomization and Combustion

The application of neural networks is receiving widespread attention, however, it has yet to have a substantial impact in the field of liquid atomization and subsequent combustion. As liquid transport and breakup problems seek global characterization and dynamical understanding, NNs which leverage the spatial- and temporal-dependence of data are of particular interest. Based on the previous discussions, there is a suspected symbiosis between the needs of liquid atomization works and the solutions that NNs can provide.

Very few works have investigated deep learning for liquid atomization work. Focusing on particle size distributions of solids, Ko and Shang [200] developed a NN to help estimate size distributions from images. The images require preprocessing and no attempt at resolution or magnification invariance was made, however, they provided initial insight into the problem. Work by Yu et al. [201] also used a neural network approach for approximating particle size distributions. Instead of using images, dimensionless parameter groups were defined and used as inputs to reduce the size of the input space. Many algorithms for automatic particle sizing leveraging high-speed video do not use neural networks (e.g., [202]), or only use neural networks for segmentation and identification of particles prior to sizing [203].

For datasets which rely on their temporal characteristics, recurrent NNs (RNNs) allow previously used information to persist in future calculations. Arsie et al. [204] used an RNN as a virtual NO_x sensor in an internal combustion engine, which was shown to give good predictions of NO_x dynamics.

An early application of deep learning for combustion work is by Sarkar et al. [205] who used NNs for the early detection of combustion instabilities. The network could extract the coherent structures of the flames and then model the temporal variation at a reduced dimension. This was followed up by Akintayo et al. [206] who developed a convolutional autoencoder for the detection of thermo-acoustic instabilities in gas turbine engines. They leveraged high-speed video recordings to detect low energy instability features which performed favorably to existing image processing techniques. Jha et al. [207]

used a CNN and Gaussian processes to analyze video data of swirl-stabilized combustor flames. Initially, a CNN is used to extract features from the videos with the feature dynamics being modeled using a Gaussian process. Spatio-temporal results of combustion flame instability are produced. They discuss the difficulty in interpreting results with the main conclusion that a framework has been presented for future work to be based off. Additionally, NNs have been used to predict boundary layer flashback [208].

Video Data

One of deep learning's most attractive traits is in its general formulation and flexibility; even seemingly disparate fields or problems can be solved with minimal adjustment to a given network design. This is especially true for problems working with image and video datasets as many of the same spatio-temporal features are required to understand and learn mappings. Here, additional insight into relevant advances in learning from image and video data, beyond the fields of liquid atomization and combustion, are provided.

CNNs have shown great promise in the area of image segmentation and recognition, so its application to video data shortly followed. In this case, data become spatio-temporally correlated. Initial insights to the problem of video classification were given by Karpathy et al. [209]. Millions of video clips were used to train a number of CNN architectures which attempted to incorporate the temporal structure of the data. An obvious concern of applying deep learning to video data is in the increased computational expense, which can be many orders of magnitude greater than image classification. Their work provides a solution to this problem using a multiresolution design. Instead of inputting each video frame at full resolution, two inputs are used: a context stream which is at half resolution of the full image, and a fovea stream which is at full resolution of the center 50% of the frame. This approach is shown to greatly reduce computation demand while having negligible impact on the classification performance. Although the features learned through these streams are not semantically significant, the context stream primarily learns colors while the fovea stream learns high-frequency, grayscale patterns.

To infuse time information into the CNN, three approaches are considered, as shown in Figure 2-15. Late fusion uses two "single frame" paths and connects their outputs in the last layer whereas early fusion immediately combines adjacent frame information. The slow fusion method is a combination of the late and early designs. Interestingly, the single frame method is able to achieve high classification accuracy without the need of explicitly incorporating temporal information. However, this accuracy is increased using the fusion techniques, with the slow fusion case performing the best.

An alternative approach is to naturally extend 2D convolutions into the temporal dimension, thereby creating 3D convolutions. A 3D CNN design was presented by Ji et al. [210] and compared to a 2D CNN for motion recognition of low quality video data. When comparing 2D and 3D designs, both are capable of extracting motion information but the latter has superior performance in recognizing desired actions. Of course, the use of 3D convolutions comes at a cost of increased computational time and more model parameters and hyperparameters, but it allows spatio-temporal information to be fused together directly. Perhaps surprisingly, 2D CNNs have been shown to outperform 3D CNNs for learning video representations, without any explicit encoding of temporal information [211].

A ResNet formulation for spatio-temporal feature learning was used by Tran et al. [212], which was found to have improvements over a regular 3D CNN implementation [213]. To perform the 3D convolutions, the network first performs a 2D spatial convolution followed by a 1D temporal convolution, with a nonlinearity in between which was found necessary to improve performance. By separating the spatial and temporal convolutions, performance improvements were found over regular 3D convolutional ResNets.

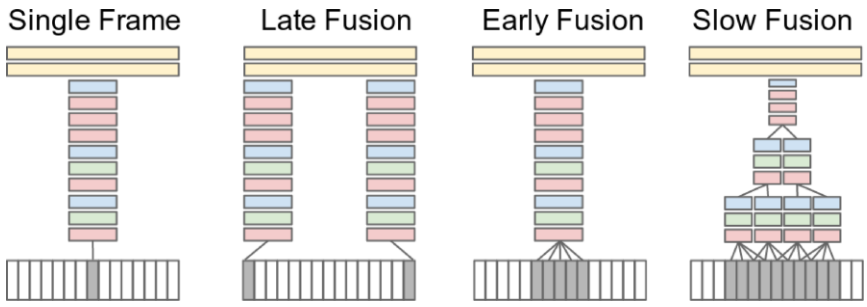

Figure 2-15: Approaches to fusing spatio-temporal information in a CNN. Image from [209].

While ResNets provide benefits over networks without the residual connections, dilated residual networks [189], which utilize dilated convolutions in place of regular convolutions, have demonstrated improved accuracy over regular ResNets at the same depth, or shallower, with no increase in parameters. Both ResNets and dilated convolutions have been, and continue to be, successful for a wide range of image- and video-based problems. An example application is for object segmentation using the FusionSeg architecture, which can combine motion and appearance within videos using an adapted ResNet-101 model with dilations [214].

The work by Carreira and Zisserman [215] focuses on video classification and investigates a variety of ways of fusing temporal information with spatial information. In particular, they present a two-stream 3D convolutional network, where one stream has the regular input images as input while the other stream has optical flow images, before fusing them together prior to classification. They show that while 3D CNNs can directly learn temporal patterns from RGB inputs, also incorporating an optical-flow stream greatly improves model performance. A central idea of the paper is in inflating pre-trained 2D CNNs into a 3D CNN; they use pre-trained filters from ImageNet 2D networks, and stack them to become 3D. Using this pre-training scheme was found to outperform all other networks tested.

2.4 Summary Literature Review

Liquid atomization systems are complex, yet display, and are fundamentally built on, a large number of simpler, oscillatory behaviors with known spatio-temporal coherence. Despite advances in diagnostics, experimentation, simulation, and computational tools and power, understanding of and an ability to predict and control these systems has not greatly improved or led to great advancements in atomization technologies. Indeed, liquid atomization systems are just an example, albeit rather complex, in fluid mechanics as a whole. Contributions and new physical insights into atomization systems have limitless transferability to other fluid systems.

High-speed imaging and videos are used extensively for qualitative purposes for atomization problems; high spatio-temporal resolution, capturing a large spatial volume of the fluid system, can be captured. Not only are these data often only used qualitatively, but only a small fraction are actually used for discussion or to generate results. Yet, these data are very revealing and capture many fundamental processes, with the main drawback being the data are a 2D projection of a 3D system. Methods of extracting quantitative information, or illuminating system physics, directly from these data can overcome the large swathe of existing diagnostics and their respective limitations which are typically used.

A particular problem with improving physical understanding and extracting dynamical features from atomization datasets is that they are extremely high-dimensional. In excess of millions of variables can be recorded whose spatio-temporal dependence is unknown *a priori*. This high-dimensionality has greatly hindered advancements through traditional analytical or dynamical analysis. As a result, techniques from machine learning have seen growing use as they can be used in a purely data-driven approach, offering flexibility in model design.

Proper orthogonal decomposition (POD) and dynamic mode decomposition (DMD) have been two approaches leveraging high quantities of system data to provide dynamical feature extraction from fluid flow data. These techniques are linear, yet have been applied extensively to many different problems in an attempt to improve physical understanding of the applied system. In reality, the systems analyzed are highly nonlinear and, while workers *interpret* extracted modes and results, they are seldom validated. Further, while a reduced order linear model can be formulated using both POD and DMD, future state prediction is rarely done and, for nonlinear systems, are unlikely to be successful.

Many of the aforementioned limitations can be addressed through the use of deep learning. Deep learning has demonstrated state-of-the-art performance for problems in many disparate fields; it is a universal function approximator with high flexibility and generality, particularly in identifying nonlinear relationships; and it excels on tasks which generate large quantities of data. For the purpose of Koopman analysis, only a few recent advances have focused on extending current Koopman ideas into a deep learning framework.

There is great potential in further developing Koopman techniques for extracting dynamical features and improving physical understanding of complex, multiphase, liquid atomization systems. Current Koopman techniques require further analysis to evaluate and validate their efficacy for these challenging problems while deep learning offers many attractive attributes for addressing accompanying concerns. This dissertation provides a critical evaluation of the efficacy of existing Koopman techniques in extracting system physics and associated concerns tied to their underlying theory for spatio-temporal data. Further, a new deep learning-based Koopman framework is presented specifically for spatio-temporal data, enabling the extraction of spatially and temporally dynamic liquid atomization fundamental processes.

Chapter 3

Approach

The previous chapter has highlighted the need to understand atomization systems and the associated challenges in addressing this need. The development of analytical theory is currently restricted to idealized and impractical cases due to the inherent complexity of the multiphase, turbulent flows. Instead, one can leverage high-speed imaging to capture quasi-global system data at a high spatio-temporal resolution. Existing Koopman analysis techniques have sought to leverage these datasets, with numerous works purportedly extracting and identifying fundamental system processes. While these works have shown promise, there is a lack of result validation, which is of particular interest as many common techniques are linear tools.

Therefore, there is a clear and justified need in evaluating and further developing existing Koopman analysis techniques for the dynamical feature extraction of atomization phenomena. Current concerns have the potential to be addressed by leveraging the flexibility and high representational capacity of deep learning.

To meet the dissertation goals and objectives laid out in Chapter 1, it is imperative that tasks are established and reasoned. These tasks form the overall approach of the dissertation work, where each task logically follows the preceding task, enabling a successful and cogent path to achieve the goals.

Task 1: Identify and select Koopman analysis techniques typically used for liquid injection systems

A large number of variants for performing Koopman analysis have previously been, and currently are, implemented for a large number of disparate engineering problems. With a focus on liquid injection studies, the most typical and successful techniques are first identified, with the most relevant techniques being selected for implementation in this work.

Due to the growing interest in modal decomposition and the rapid development of variants, it is essential to conduct analysis on only the most relevant approaches. These approaches will form a baseline with which further analysis and deep learning developments will be compared against.

Task 2: Critically evaluate algorithm design, apply selected techniques, and evaluate extracted results and interpretability

The selected techniques will be critically evaluated for their application in extracting dynamical features. Their algorithmic formulation and design are analyzed to gain a deeper understanding of their intrinsic behaviors prior to applying the techniques to both practical and simplified systems. The extracted results are then validated in a hierarchical fashion, with a focus on the physical interpretation and understanding that is gained.

Despite the extensive application of Koopman analysis techniques, few works validate their findings or contribute effectively to improving the physical understanding of a specific application. Due to this, it is essential to identify any algorithmic reason for the results extracted and the physical interpretation gained or lack thereof. Results should be explained trivially for basic systems with known spatio-temporal coherence and thus serve as a strong foundation to evaluate the techniques.

Task 3: Replicate selected techniques using a deep learning framework

A deep learning analog of the selected techniques is developed and demonstrated. Algorithmic variations and performance details are compared and discussed. This introduces a deep learning baseline and network design fundamentals for incorporating Koopman theory directly into the architecture. Further development is enabled through this task.

Task 4: Develop a general deep Koopman network

A more general and flexible deep learning approach to Koopman analysis is developed, removing the linearity assumptions of traditional Koopman analysis techniques in favor of the arbitrary nonlinear complexity afforded by neural networks. The appropriate design and training choices and potential network variants and modifications for a variety of purposes are described.

Generality ensures that future variants are limited in terms of necessity and modification scope. A general Koopman network provides the necessary flexibility to analyze a large number of flow problems. Indeed, there are many variants and alternative designs that require investigation and evaluation, either for network or problem optimization.

Task 5: Refine the deep Koopman network and training requirements to perform an interpretable modal decomposition

Further deep Koopman network developments are made specifically to improve the extraction of modes and their interpretation. How the network is able to extract the contribution of individual modes, with a focus on spatio-temporal data, is presented and discussed. Approaches for improving the interpretability and robustness of the Koopman network are provided.

Neural networks in general benefit from their representational capacity and flexibility, which often comes at the cost of model opacity. Therefore, while the network can easily minimize the specific cost function, this only provides benefits to physical understanding if the modal decomposition is intuitive and interpretable.

Task 6: Evaluate deep Koopman network and compare to traditional techniques

Finally, the developed deep Koopman network is compared to the selected, traditional techniques in their ability to predict future flow states and in modal interpretability. The previous results of the Koopman

analysis techniques are revisited, again in a hierarchical fashion, to demonstrate the utility and limitations of all approaches.

Where traditional techniques are evaluated and have their limitations identified, it is important to perform the same validation to the deep Koopman network and demonstrate its ability to overcome the limitations. The deep Koopman network's performance on a variety of problems is presented and discussed to serve as a guide for future work.

Chapter 4

Methodology

The background provided in Chapter 2 has highlighted existing concerns for advancing atomization understanding which led to the formulation of tasks for this work, presented in Chapter 3. Specifically, techniques related to Koopman analysis are of interest for atomization systems as they provide a data-driven approach to extract physical understanding from large quantities of data; a task traditional dynamical analysis and analytical development have struggled with. Data from high-speed video are the focus of this work, owing to their ability to capture quasi-global, high spatio-temporal resolution data. While existing Koopman analysis techniques have had widespread use within the fluids community, they have yet to be evaluated critically and validated on even simple systems. Further, many immediate concerns with common techniques can be alleviated through a deep learning framework.

This chapter covers the methodology taken to critically evaluate proper orthogonal decomposition (POD) and dynamic mode decomposition (DMD), highlight their limitations, which subsequently guides the development of a more general and flexible deep learning-based model for video data: the deep convolutional Koopman network (CKN).

Due to contributions from this work being in algorithm and model development, design, and consequent analysis, general theoretical and methodology discussions are also provided here. More specific experimental details are found with their accompanying results in the next chapter.

4.1 Image and Video Data

An ability to extract quantitative information directly from image or video data greatly facilitates the global analysis of atomization systems. While there are drawbacks such as optical density limits and

capturing a 2D projection of a 3D system, important dynamical and physical insights have the potential to be gathered even for the practical systems presented. Of course, this is contingent on a technique capable of providing these insights. With advances in computer vision and data-driven dynamical analysis, however, this is now a feasible approach.

For the analyzed datasets, the spatial resolution is to be kept reasonably low, while still capturing all necessary spatial structures, to limit computational expense. Temporal resolution is largely dependent on spatial resolution for high-speed recordings, but temporal resolution is to be sufficiently low to prevent subpixel motion which may adversely affect the decomposition techniques. This is particularly the case for DMD as a linear mapping is unable to map a single input to multiple outputs. Evaluation of techniques accounting for subpixel motion is beyond the scope of this work, but is highly desirable for future works.

Training data for the decompositions can either be simulated, and extracted from the simulated results, or captured experimentally using a high-speed camera. A number of dynamical systems are analyzed, ordered hierarchically in terms of system complexity. Simulated systems are used to generate data for the simplest systems, as the exact spatio-temporal behaviors can be known *a priori*. Experimental systems are left as examples of more complex systems due to the number of uncontrollable variables in a practical setting.

4.1.1 Data Collection

All of the simulated systems in this work have simple formulations to simulate their spatio-temporal evolution. All spatial and temporal scales are known precisely and selected to be conducive to providing meaningful analysis. The datasets consist of 10,000 snapshots of the system evolution, as if through experimentation, with random noise added with a standard deviation of up to 10% of the peak value. The time interval between frames was sufficiently high to limit sub-pixel motion which can lead to identical adjacent frames.

A Vision Research Phantom V7.2 monochrome high-speed camera was used to capture all experimental systems. Backlighting was used to produce a shadowgraph, which was made uniform by using a diffuser.

As there is a trade-off between the spatial and temporal resolutions for recording data, they were optimized for each system. An Infinity long-distance microscope with accompanying objective lenses were used to provide the appropriate magnification, particularly for the simplest injection systems.

4.2 Koopman Theory

Although many variants of Koopman analysis techniques are available in the literature, the standard POD and DMD algorithms are by far the most prevalent and continue to be used in current works. A deeper understanding of these standard algorithms can have an impact on all current and future variants. The standard DMD formulation provided by, for example, Tu et al. [89] is considered for use in this work. The link between POD and DMD theory and their ability to extract physics has not been verified, despite this link being a major motivation for workers to use the techniques. As such, a thorough discussion and understanding of the algorithms is essential prior to developing further methodology.

4.2.1 Understanding POD

For POD, while mean-subtraction does not have an effect on the basic calculations, it can affect the interpretation of the results [113], hence it is not recommended here. POD can be calculated from the singular value decomposition (SVD) which can be applied to any arbitrary data matrix, $\mathbf{X} \in \mathbb{C}^{m \times n}$, to give

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{*} \tag{4.1}$$

where $\mathbf{U} \in \mathbb{C}^{m \times n}$ has orthonormal columns, $\mathbf{\Sigma} \in \mathbb{C}^{n \times n}$ is a diagonal matrix, $\mathbf{V} \in \mathbb{C}^{n \times n}$ has orthonormal columns, and the asterisk denotes the complex conjugate transpose. The columns of the matrix $\mathbf{U}$ are typically denoted the POD modes, however, this is not necessarily the case. This formulation is brilliantly simple, but some intuition is lost if the POD modes are found blindly from the SVD.

Understanding the data matrix and how the SVD is calculated aids the intuition of how and why POD works. First, the matrix $\mathbf{U}$ is calculated as the eigenvectors of the matrix $\mathbf{X}\mathbf{X}^{*} \in \mathbb{C}^{m \times m}$. The matrix

multiplication $\mathbf{XX}^*$ yields a correlation matrix: each row of $\mathbf{X}$ is dotted with every other row, including itself, which provides a similarity measure. Therefore, column j and row j of the correlation matrix both describe the j^{th} row of the data matrix's similarity with every other row of the data matrix.

The correlation matrix can now be thought of as m vectors living in an m dimensional space, where each dimension corresponds to a row of the data matrix. The vectors' magnitude and direction indicate which rows correlate highly with which other rows in the data matrix. For example, consider the distribution of points in Figure 4-1, where the coordinates of data point d are governed by the d^{th} row's similarity with every other row. The red points correspond to highly correlated data points arising from every k^{th} row. The black points show how these rows share low information content with all other rows. For clarity, only three-dimensions are visualized, but this idea is readily extended to include all k rows. Then, there exists a correlation matrix vector which has large component magnitudes in the direction of all k^{th} dimensions, and small component magnitudes in all other directions.

Further, the resultant matrix is symmetric and therefore has orthogonal eigenvectors, which form a basis in which data are only scaled by corresponding real eigenvalues. Because of this property, and when eigenvalues are sorted in descending order where $\{\lambda_1 > \lambda_2 > \cdots > \lambda_m\}$, the corresponding eigenvectors identify directions of maximal variance, with every subsequent eigenvector capturing the maximal variance in a direction orthogonal to all previous eigenvectors. In this sense, the eigenvectors capture the principal components of the data.

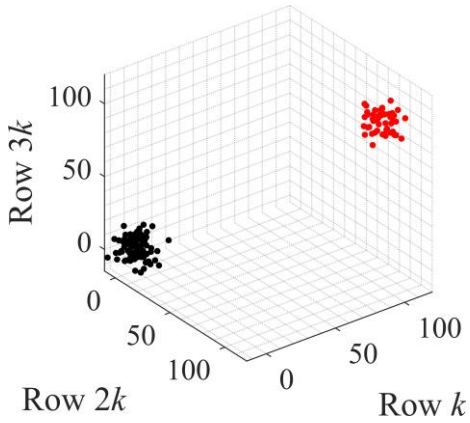

Figure 4-1: Plot of row similarities. Red points show how every k^{th} row is highly correlated, but generally uncorrelated with all other rows, given by the black points.

Similarly, the matrix $\mathbf{V}$ is calculated as the eigenvectors of the matrix $\mathbf{X}^*\mathbf{X} \in \mathbb{C}^{n\times n}$. All of the previous intuition holds but this time the correlation matrix generated captures the similarity between columns of the data matrix.

Finally, the matrix $\mathbf{\Sigma}$ holds the singular values of the data matrix, which are the square roots of the non-zero eigenvalues of both $\mathbf{X}\mathbf{X}^*$ and $\mathbf{X}^*\mathbf{X}$, where the square root arises from the repeated multiplication of the data matrix $\mathbf{X}$ to form the correlation matrix. As the singular values are related to the eigenvalues of the correlation matrix, which is a measure of the variance captured by the corresponding eigenvectors, the singular values capture the relative modal energy content of the system. POD therefore provides a convenient and interpretable energy metric for ordering the importance of the modes.

Since $\mathbf{\Sigma}$ can be calculated with knowledge of only one of $\mathbf{U}$ or $\mathbf{V}$, only one eigendecomposition is necessary since the remaining matrix can be calculated through a rearrangement of Equation 4.1. While convenient, this forgoes the intuition that $\mathbf{U}$ and $\mathbf{V}$ are in fact both correlation-dependent matrices and whose utility depends on the structure of the data matrix $\mathbf{X}$.

More concretely, the data matrix is defined as

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x_2} & \dots & \mathbf{x}_n \\ | & | & & | \end{bmatrix}, \tag{4.2}$$

which is made up of n vectors, each of which belong to $\mathbb{C}^m$. Video data is of interest, which is made up of a collection of video frames of equal time-spacing. As is typical in modal decomposition studies, the spatial measurements at a given time are all stacked into a column vector of identical size with arbitrary order but whose order is consistent across all samples. In this case, video frame j is flattened into a column and forms column j, $\mathbf{x}_j$, of the data matrix, where each frame has m pixels. Recalling Section 2.3.2, this flattening removes the spatial topological order of the dataset, but this is a necessary step of the standard POD algorithm.

In this format, each row represents the measured time history of a given spatial location and each column represents the spatial distribution of measurements at a given time step. Performing the SVD on $\mathbf{X}$ therefore

produces matrices **U** and **V** which capture the spatial and temporal correlative structure of the data, respectively. Here, the spatial measurements are pixel intensity values but could be any variable of interest. The matrix **U** therefore captures the spatial structures which are most dominant in the data, hence why it is often termed the POD modes, while **V** captures the time dynamics, or temporal coefficients, of the corresponding modes. To be clear, the matrix **U** is *only* termed the POD modes because of the format of the data matrix given in Equation 4.2. Further, the spatial- and temporal-correlation matrices are completely independent and, as such, only spatial-coherence in captured in this formulation. For this reason, this algorithm is also called space-only POD.

A common use case for POD is in reduced order modeling or low-dimensional reconstructions of an original dataset. It is often the case that a high percentage of the explainable variance in the data, on the order of 95%+, can be captured within the first 10 or 20 modes. An r-mode truncation can be performed, allowing a reconstruction of time step k through

$$\mathbf{x}_{\mathrm{POD},k} = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j^T \mathbf{v}_j. \tag{4.3}$$

The POD modes, **U**, capture the dynamical features with energy, **Σ**, whose dynamics evolve according to **V**.

A number of important points need to be made here. First, discussion on the SVD highlighted how the columns of **U**, the POD modes, capture the orthogonal directions of maximal variance, which can be physically interpreted as the energy of the system. Thus, the first POD mode is the spatial structure which optimally captures the entire system energy, in an L_2 sense, with one vector. The second mode, together with the first mode, then optimally captures the entire system energy with two orthogonal vectors, and so on. There is then nothing in this formulation which attempts to decouple independent or overlapping processes; rather, any coupling present in a system is likely to be mixed together in the POD modes.

Secondly, POD is not dependent on data ordering. The calculated eigenvectors of the correlation matrices will always identify the spatial measurements which are highly correlated. Clearly, this is not an ideal feature of a dynamical analysis tool, but it does mean that time-sampling does not need to be uniform.

The time dynamic information, captured by $\mathbf{V}$, is, however, highly dependent on the coupling of system processes and on the time-sampling procedure.

4.2.2 Understanding DMD

Dynamic mode decomposition arose from the desire to learn the physical structures and behaviors of turbulent systems directly from measured data [88]. DMD aims to capture data, infer linear models from the data, and then give physical insight into processes which currently have no analytical theory available. DMD approximates the system of interest as being linear in a high-dimensional space, attempting to capture nonlinear dynamics. As this is generally not the case, concern arises over the extent to which DMD is considered valid when applied to nonlinear systems.

Generally, DMD is the procedure for performing the eigendecomposition of the linear operator which maps the matrix

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}_i & \mathbf{x}_j & \cdots & \mathbf{x}_k \\ | & | & & | \end{bmatrix} \tag{4.4}$$

to the matrix

$$\mathbf{X}' = \begin{bmatrix} | & | & & | \\ \mathbf{x}_{i+1} & \mathbf{x}_{j+1} & \cdots & \mathbf{x}_{k+1} \\ | & | & & | \end{bmatrix} \tag{4.5}$$

where the indices i, j, and k are arbitrary time steps but highlight that the only ordering that matters is that corresponding columns in $\mathbf{X}$ and $\mathbf{X}'$ are shifted by one time step. Note how, likewise with POD, the spatial topology of video data is not preserved in the DMD algorithm. In this work, there is a uniform time step, Δt, between adjacent columns, so $\mathbf{X}$ can be written $\{\mathbf{x}_t\}_{t=1}^{n-1}$ and the time-shifted data matrix, $\mathbf{X}'$, can be written $\{\mathbf{x}_t\}_{t=2}^{n}$, so $\mathbf{X}$, $\mathbf{X}' \in \mathbb{C}^{m \times n-1}$, where m is the number of state variables and n is the number of snapshots in time. In general, the data matrices may be complex, however only real data are used for video data. The linear operator, $\mathbf{A} \in \mathbb{C}^{m \times m}$, then maps the columns of $\mathbf{X}$ to the corresponding columns of $\mathbf{X}'$,

$$\mathbf{A}\mathbf{X} = \mathbf{X}', \tag{4.6}$$

which can be solved through

$$\mathbf{A} = \mathbf{X}'\mathbf{X}^{\dagger} \tag{4.7}$$

where the superscript † denotes the Moore-Penrose pseudoinverse. Using the definition of the SVD from Equation 4.1, this can be written as

$$\mathbf{A} = \mathbf{X}'\mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^{*}. \tag{4.8}$$

Typically, this is intractable to compute, so a low-dimensional approximation of **A** is desired from which the behavioral properties of the linear system can be extracted.

DMD builds on the foundations of POD, leveraging the utility of optimally representing the system's energy with a given number of POD modes. It is stressed that POD is purely mathematical in nature; it is a linear technique which provides orthogonal vectors which maximize explained variance of the dataset, independent of data ordering. To reduce the dimensionality of the approximation of **A**, DMD truncates the number of modes used for all of the analysis. An r-truncated SVD of **X** is used for DMD,

$$\mathbf{X} \approx \mathbf{U}_r\boldsymbol{\Sigma}_r\mathbf{V}_r^{*} \tag{4.9}$$

with the r-most dominant POD modes given by $\mathbf{U}_r \in \mathbb{C}^{m\times r}$, associated relative energy contributions given by $\boldsymbol{\Sigma}_r \in \mathbb{C}^{r\times r}$, and modal temporal dynamics given by $\mathbf{V}_r \in \mathbb{C}^{n\times r}$, where r is the number of modes kept to produce an $r \times r$ approximation of **A**.

The approximation of **A** is calculated by projecting onto the column space of $\mathbf{U}_r^{*}$ to give

$$\widetilde{\mathbf{A}} = \mathbf{U}_r^{*}\mathbf{A}\mathbf{U}_r \tag{4.10}$$

where $\widetilde{\mathbf{A}}$ is the $r \times r$ approximation of **A** which can tractably be analyzed. The projection onto the POD modes defines spatial groupings of state variables whose behaviors have similar coherence in time. $\widetilde{\mathbf{A}}$ therefore captures the dynamics of the system with the POD modes forming the basis of the transformation.

An eigendecomposition of $\widetilde{\mathbf{A}}$ can now be calculated through

$$\widetilde{\mathbf{A}}\mathbf{W} = \mathbf{W}\boldsymbol{\Lambda} \tag{4.11}$$

where $\mathbf{W} \in \mathbb{C}^{r\times r}$ is given by

$$\mathbf{W} = \begin{bmatrix} | & | & & | \\ \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_r \\ | & | & & | \end{bmatrix} \tag{4.12}$$

and $\mathbf{\Lambda} \in \mathbb{C}^{r \times r}$ is a diagonal matrix given by

$$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_r \end{bmatrix}. \tag{4.13}$$

$\mathbf{w}_j$ and λ_j are approximations of the j^{th} eigenvector and eigenvalue of $\mathbf{A}$, respectively. The approximate eigenvalues, λ_j, of the full system are known as the Ritz values [88]. Knowing that this eigendecomposition was performed in a low-dimensional representation of $\mathbf{A}$ through the POD modes, the eigenvectors now identify contributions of the POD modes which result in oscillatory behavior with exponential growth or decay. Therefore, when the column space of $\mathbf{X}$ and $\mathbf{X}'$ are approximately equal [71], the DMD modes can be defined through

$$\mathbf{\Psi} = \mathbf{U}\mathbf{W} \tag{4.14}$$

where $\mathbf{\Psi} \in \mathbb{C}^{m \times r}$ is given by

$$\mathbf{\Psi} = \begin{bmatrix} | & | & & | \\ \boldsymbol{\psi}_1 & \boldsymbol{\psi}_2 & \cdots & \boldsymbol{\psi}_r \\ | & | & & | \end{bmatrix}. \tag{4.15}$$

The columns of $\mathbf{\Psi}$, $\boldsymbol{\psi}_j$, represent the DMD modes which can be considered a linear combination of the POD modes, $\mathbf{U}$, whose coefficients are given by the eigenvectors, $\mathbf{W}$. Similar to POD, the DMD modes capture dynamical features, but whose dynamics evolve sinusoidally owing to its linear transformation formulation. Intuitively, the DMD modes use the energetically-optimal POD modes to form a basis for the approximation of $\mathbf{A}$ whose eigenvectors identify contributions to the linear dynamics in the lower-dimensional basis.

Although DMD utilizes the orthogonal modal decomposition of POD, the DMD modes themselves are generally not orthogonal [88]. For practical considerations, the eigenvalues of $\widetilde{\mathbf{A}}$ are distinct, which indicates that DMD is able to decouple spatial behaviors which have different oscillatory and decay

properties. The frequencies and exponential growth or decay rates of the DMD modes can easily be given by the corresponding phase and magnitude of the eigenvalues, respectively. Unfortunately, the convenient energy hierarchy afforded by POD is no longer available for DMD and, therefore, ranking the DMD modes by energy content can be difficult.

A modal energy ranking suggested by Tu et al. [216] utilizes the modal norm, given by the L_2 norm, for mode j, $\|\boldsymbol{\psi}_j\|$, scaled by its corresponding eigenvalue to the $n-1^{\text{th}}$ power, such that

$$b_j = \lambda_j^{n-1}\|\boldsymbol{\psi}_j\| \tag{4.16}$$

where b_j is the energy content of the j^{th} DMD mode. The scaling by the eigenvalue assumes that quickly decaying eigenvectors in the system have reduced coherence which are not necessarily representative of the system as a whole. Higham et al. [217] demonstrate the utility of using the Fourier transform of the POD temporal coefficients, creating a power spectral density (PSD), to guide the selection of dominant DMD modes.

A popular DMD energy metric is to solve the least squares problem

$$\mathbf{b} = \min_{\mathbf{b}}\|\boldsymbol{\Psi}\mathbf{b} - \mathbf{x}_1\| \tag{4.17}$$

relative to the first column in the data matrix $\mathbf{X}$. Although in practice this formulation has been successful, it is dependent on data ordering; different values of b_j or modal ordering, or both, can arise. Instead, for this work, an average energy metric over v randomly selected vectors is introduced and used, to reduce sensitivity on data ordering. Thus, the formulation used here is

$$\mathbf{b} = \frac{1}{v}\sum_{\mathbf{x}\in V}\min_{\mathbf{b}}\|\boldsymbol{\Psi}\mathbf{b} - \mathbf{x}\| \tag{4.18}$$

where V is the set of v randomly selected vectors from $\mathbf{X}$. A value of $v = 100$ was used as it provided a robust modal ordering without excessive additional computational cost.

Finally, generating reconstructions of the original dataset is a common approach to quantify the accuracy of the truncated linear model, which is given by

$$\mathbf{x}_{\mathrm{DMD},k} = \sum_{j=1}^{r} b_j \, \boldsymbol{\psi}_j \exp\left(\omega_j t_k\right) \tag{4.19}$$

for an r-mode truncation, where $\mathbf{x}_{\mathrm{DMD},k}$ is the DMD reconstruction of the k^{th} time step and $\omega_i = \ln\left(\lambda_j\right)/\Delta t$.

As DMD relies on the formulation of POD, and both techniques are inherently linear, the representational capacity is clearly limited. Further, nothing about either formulation is physically-informed [218], so their ability to extract physics is not tied to theory.

4.2.3 Evaluating POD and DMD

Despite POD and DMD being linear tools, they are almost exclusively applied to nonlinear systems. These techniques purport to extract physical understanding and allow the user to identify dominant dynamical features. While the idea of the modal decomposition is a useful one, the lack of validation on the results is concerning. Further, while interpretability of modes is likely to be an ongoing issue for highly complex processes and behaviors, any postulation should be built upon the validation of known, ground truth processes. Although a seemingly trivial step, no published works have critically evaluated and validated the application of POD and DMD for even simple systems with known spatio-temporal coherence.

This serves as a fundamental aspect of this dissertation: all algorithms and models in this work are to be applied, analyzed, evaluated, and validated on a hierarchy of problem complexity. As such, initial contributions of this work focus on the application of POD and DMD on problems with simple dynamics before investigating practical systems as is traditionally done, with a focus on evaluating the efficacy of the decompositions.

For clarity, for all the systems analyzed, they are recorded in spatio-temporal form optically, or optically through simulation.

Simulated Systems

While there are many simple liquid injection systems which behave predictably with established formulae for instability wavelengths and breakup lengths, many experimental variables must be controlled which are likely to make the system less idealized. Instead, a simulated system can be controlled *exactly*, serving as a starting point with which to analyze POD and DMD.

As the data are observed from an optical viewpoint, the intrinsic state variables are not recorded directly. Of course, the true underlying dynamics are known *a priori* as they are specified to simulate the systems. For many such cases, this change in perspective causes nonlinear variables to be recorded, regardless of the true underlying linearity. Hence, the performance of POD and DMD for these systems, without access to the intrinsic state variables, is of great interest.

Due to the presence of oscillatory waves or wave-like behaviors in liquid atomization systems, the application of POD and DMD to a fundamental wave problem serves as a strong baseline. Probably the simplest wave system is that of a vibrating string. The analyzed system is shown in Figure 4-2 with its constituent sinusoidal waveforms. Only two sinusoids contribute to the overall waveform: one sinusoid is one wavelength long with a frequency of 2 Hz and the other is three wavelengths long with a frequency of 3 Hz. As these are the only structures contributing to the overall system dynamics, a successful modal decomposition should produce two modes, one for each waveform, capturing the correct spatio-temporal behaviors. The resolution of the data is 40×60 pixels.

Indeed, vibrating strings are ubiquitous problems with well-known analytical solutions. However, when viewed optically, the response of each pixel is nonlinear, often capturing a square-like wave, and so POD and DMD are tasked with identifying the characteristics of a linear system from a nonlinear representation.

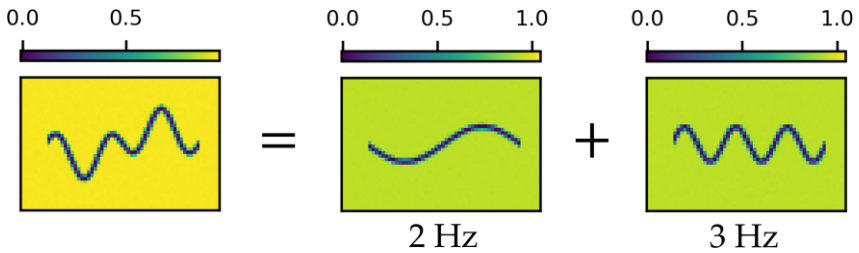

Figure 4-2: The vibrating string system, made up of two superimposed sinusoids.

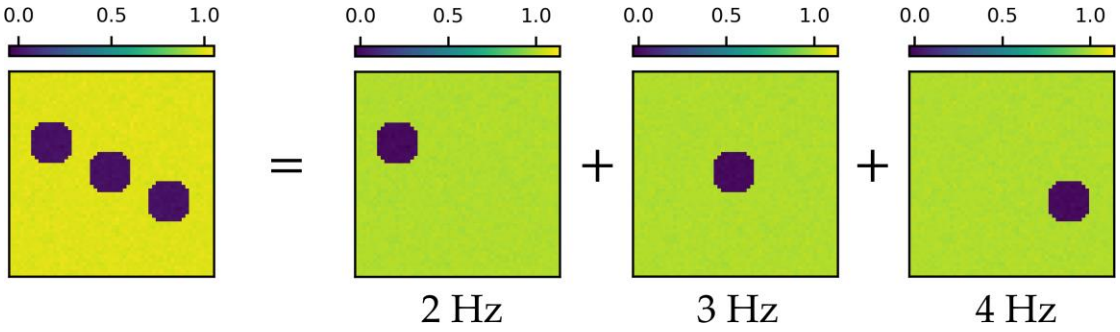

Figure 4-3: The particle stream system, made up of three distinct particle streams.

Evaluating these results clarifies to what extent POD and DMD can analyze nonlinear representations of linear systems or if they are confined to linear representations.

The next simulated system analyzed is that of a group of observed particles moving on linear trajectories at constant velocity; a basic representation of behaviors typically found in atomization systems. The particle stream system is shown in Figure 4-3 with a snapshot of its independent constituent processes. Three particle streams move at different velocities and, thus, have different periods as they wrap in and out of the frame (i.e., motion is top to bottom). The cycling frequency of the particle streams from left to right is 2 Hz, 3 Hz, and 4 Hz. Therefore, the right particle stream cycles at the second harmonic frequency of the left particle stream. The resolution of the data is 56×56 pixels.

The particle stream system tests fundamental spatial and temporal behaviors of the decompositions. Each stream is spatially independent; no overlapping or cross-stream interaction occurs between the streams. As POD and DMD do not preserve the spatial ordering of the analyzed video, their ability to separate spatially distinct dynamical features is of great interest. In addition, two of the streams are related by an integer multiple frequency. While this ensures strong temporal correlation between the two streams, one is interested in the extraction of spatio-temporal coherence.

Liquid Atomization Systems with Known Spatio-Temporal Coherence

Next, a simple experimental atomization system is considered: the dilational jet, shown in Figure 4-4. The behavior of the dilational jet can be carefully controlled, it is a fundamental process in many

atomization systems, and has accompanying analytical theory. To ensure predictable behavior, the emanating jet is laminar, cavitation is avoided in the nozzle, and recorded data are close to nozzle exit [219]. The issuing liquid is water, injected from a plain sharp-edged orifice with an orifice diameter of 0.6604 mm and an orifice length-to-diameter ratio of 4, and at a low velocity so low Ohnesorge and Reynolds numbers encourage Rayleigh breakup. In this case, only a single mode governs the system: a 2700 Hz dilational wave induced by a 900 Hz pulse. The perturbed wave only supports the third harmonic as its wavelength, $\lambda_{\text{opt}} \approx 4.51d$, where d is the injector diameter, is the most dominant unstable wavelength [4, 220]. Of course, although the 900 Hz perturbing frequency does not provide its own sustaining instability, it may indirectly affect the flow in a practical system. The resolution of the data is 200×64 pixels.

Subsequently, the dilational jet will begin to breakup as surface tension forces amplify the instabilities, leading to jet breakup. Thus, a jet breakup case is analyzed, but with a 4000 Hz perturbation frequency to produce primary droplets, along with the satellite droplets, which are large enough to be captured by the optical setup [220]. A snapshot of the jet breakup process is also given in Figure 4-4; the dilational jet breaks up into a large primary droplet with an accompanying smaller satellite droplet. The jet breakup case is the first system analyzed which has a continuous spectrum [221] due to fluctuations in the jet breakup length. As a result of the inconsistent breakup length, the location of the primary droplet formation, and accompanying satellite droplet, also fluctuates. Certainly, a mode which captures the breakup and formation of both the primary and satellite droplets is expected due to the periodic formation. It is of interest as to

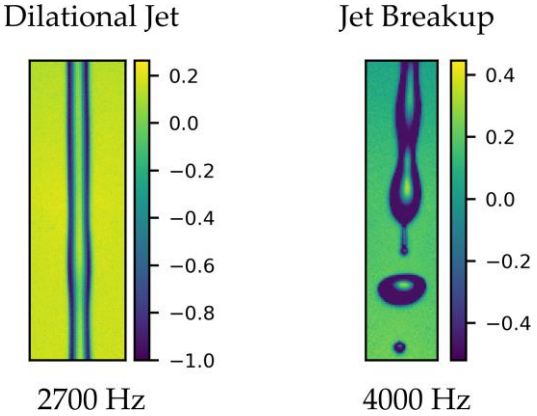

Figure 4-4: Snapshots from the (left) dilational jet and (right) jet breakup, and corresponding instability frequencies.

whether POD and DMD are able to capture the continuous spectrum of the breakup length as they only directly capture discrete spectra phenomena. Both of these systems are recorded at 20,000 frames per second.

Jet in Crossflow

From the above cases, from a linear system analyzed from a nonlinear representation to a basic experimental system exhibiting a continuous spectrum, a verifiable hierarchy is provided with which to evaluate POD and DMD on a fundamental level that heretofore has yet to be presented. The final system provides a more practical use case while still admitting tangible physical insights and understanding: the jet in crossflow. Data are recorded at 9708 frames per second with a (downsampled) resolution of 128×200 pixels.

Specifically, the system analyzed is a liquid jet undergoing bag breakup from a transverse gaseous crossflow, shown in Figure 4-5. Again, a plain sharp-edged orifice with an orifice diameter of 0.6604 mm and an orifice length-to-diameter ratio of 4 was used to inject water into an air crossflow. The air enters a 25.4 mm × 25.4 mm square channel with the liquid injected perpendicular to the air flow. Waves in the liquid column are produced primarily through aerodynamic instabilities. The conditions were run at an aerodynamic Weber number, We_∞, of 30, and a momentum flux ratio, q, of 10. No underlying frequencies of this system are known. Liquid segments flatten out and then expand into a large bag before fully atomizing. This regime has alternating segments which are undisturbed or have undergone bag breakup.

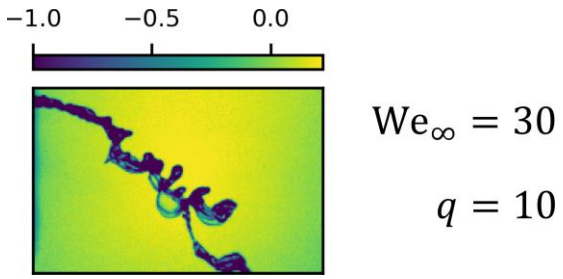

Figure 4-5: A snapshot of a jet in crossflow undergoing bag breakup.

Visualizing Modes

Both POD and DMD modes are very natural to visualize as they represent spatial structures in the flow. In both cases, each mode captures the relative spatio-temporal contributions of structures whose dynamics evolve according to their temporal coefficients or modal frequency for POD and DMD, respectively. Modal values are centered around zero and the positive (resp. negative) regions oscillate between positive and negative (resp. negative and positive) values according to their dynamics. In addition, the modes come in complex conjugate pairs whose values are phase lagged by 90°, analogous to pairs of sines and cosines. Only the real part of the modes is plotted which is conventionally done.

An example pair of complex conjugate modal structures, corresponding to a single POD mode, for the particle stream system is given in Figure 4-6. Note, that this is still representative of the visualization of DMD modes. It is stressed that modal values are relative and should be treated qualitatively. Quantitative information is extracted through the temporal coefficients or modal frequency, the modal power, or a combination of these values with the modal values.

By combining the mode with its temporal behaviors, a modal video can be created which provides intuitive insight into the contribution of each mode. However, representing the modes as a single image is still clear. For example, for DMD, structures present in positive (resp. negative) regions will then move to adjacent negative (resp. positive) regions at twice the modal frequency. This is easy to show by considering the relevant temporal scale of the mode, found by taking the reciprocal of twice the modal frequency. As

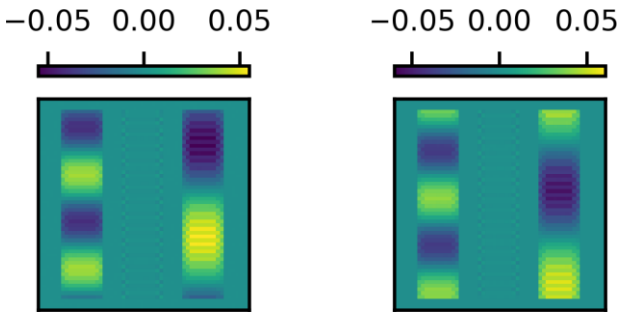

Figure 4-6: An example pair of complex conjugate modal structures. Note the 90° phase offset in the positive-negative regions between the modal pair.

DMD models the dynamics of an arbitrary system linearly, the matrix **A** progresses the system one time step forward and, hence,

$$\mathbf{A}^{f_s/(2f_\psi)}\boldsymbol{\psi} = -|\lambda|^{f_s/(2f_\psi)}\boldsymbol{\psi} \tag{4.20}$$

where f_s is the data sampling frequency and f_ψ is the modal frequency. This occurs, by definition, since $\boldsymbol{\psi}$ is an eigenvector of the linear operator **A**, however this provides clear physical insight into how the modal structures and frequencies interact. The direction of movement of modal structures can be found by decreasing the exponents in Equation 4.20 and comparing the output.

Both POD and DMD capture the mean flow of the system as their most dominant mode. As this does not usually provide any notable insights, this mode is omitted and considered the zeroth mode. Instead, all modes are presented in order of energy starting from the most dominant mode not capturing the mean flow.

The modal structures are compared qualitatively to known spatio-temporal structures and dynamics, particularly for the simplest systems, as the underlying behaviors are known. A focus is placed on the physical meaning represented by each mode *individually*, as this is commonly done in the literature.

To evaluate the modal decomposition holistically, an r-mode reconstruction and future state prediction are performed and compared to the original dataset. The quantitative metric for the comparison is the mean squared error (MSE) of pixel values for both the reconstruction and the future state prediction, given by

$$\text{MSE} = \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(x_{ij} - \tilde{x}_{ij}\right)^2 \tag{4.21}$$

where the MSE is calculated over all n columns and m elements of the data matrix **X**, x_{ij} is the j^{th} element of the i^{th} vector, and $\tilde{x}_{ij}$ is the modal reconstruction or prediction of the corresponding element.

While POD provides temporal coefficients for how each mode evolves, it does not, generally, allow for future state prediction. A power spectral density (PSD) can be produced by taking the Fourier transform of the modal temporal coefficients. If the PSD has an appropriately narrow bandwidth, a single frequency can be attributed to the mode and future state prediction is possible. For even simple systems, however, the

PSD for each mode can cover a broad range of frequencies. As a result, only state reconstructions are performed for POD.

In contrast, DMD is spectrally clean in that it outputs a single modal frequency per DMD mode. This enables both reconstruction and future state prediction as well as assigning each modal structure clear dynamical behaviors. This is particularly important for physical interpretation discussions of the extracted modes.

In addition, DMD modes are accompanied by their Ritz values, which are the approximations of the true eigenvalues of **A**. The Ritz values are plotted relative to the unit circle, whose magnitudes capture the decay or growth rates and whose angle captures the modal frequency. The marker size for the Ritz plots is directly proportional to the DMD modal energy. No scale is accompanied, so the marker size scaling is used qualitatively.

4.3 Deep Koopman Architecture

Revisiting the general linearized dynamical system through the Koopman operator:

$$\phi(\mathbf{x}_{k+1}) = \mathcal{K}\phi(\mathbf{x}_k), \tag{4.22}$$

$\mathcal{K}$ is the full-dimensional Koopman operator and ϕ is the generally nonlinear coordinate transformation which linearizes the dynamics. Both $\mathcal{K}$ and ϕ are unknown; only the recorded variables, **x**, are known through time and so approximations are sought, particularly for a low-dimensional representation of the Koopman operator, given by **K**.

DMD approximates $\mathcal{K}$ by assuming the dynamics evolve linearly, such that ϕ is just the identity. Instead, with deep learning, there is the possibility to learn both ϕ and an approximate **K** simultaneously.

4.3.1 POD and DMD Neural Networks

As will be discussed fully in the next chapter, the logical intermediate step between the analytic POD and DMD approaches and the full deep Koopman framework is to represent both POD and DMD in neural

network form. The understanding gained by analyzing the underlying theory enables their implementation via a deep learning framework, with example comparative results acting as validation for the network design.

Example networks for learning POD and DMD modes through a POD neural network (POD-NN) and a DMD neural network (DMD-NN) are given in Figure 4-7. In this specific figure with four neurons in the hidden layers, two complex conjugate pairs of modes can be learned from the system by using an autoencoder framework. POD attempts to learn an efficient data compression while maximizing the reconstruction accuracy, which implicitly preserves the energy-optimal formulation. The compressed section of the network will be referred to as the embedding space in this work, which is synonymous to the latent space. Of course, the ability to perform PCA using neural networks has been known for decades [182]. Specific application for identifying spatio-temporal dynamics, however, has not been demonstrated and provides a nice link between the analytical version of POD to further deep learning extensions.

For the given example, the only change in the DMD network is to expand the embedding space with an extra hidden layer with the same dimensionality. The DMD network attempts to learn an approximation of the linear map $\mathbf{A}$ which maps a given input, $\mathbf{x}_k$, into a prediction of the next time step, $\tilde{\mathbf{x}}_{k+1}$. The activation for all neurons in both POD-NN and DMD-NN designs is the linear activation function.

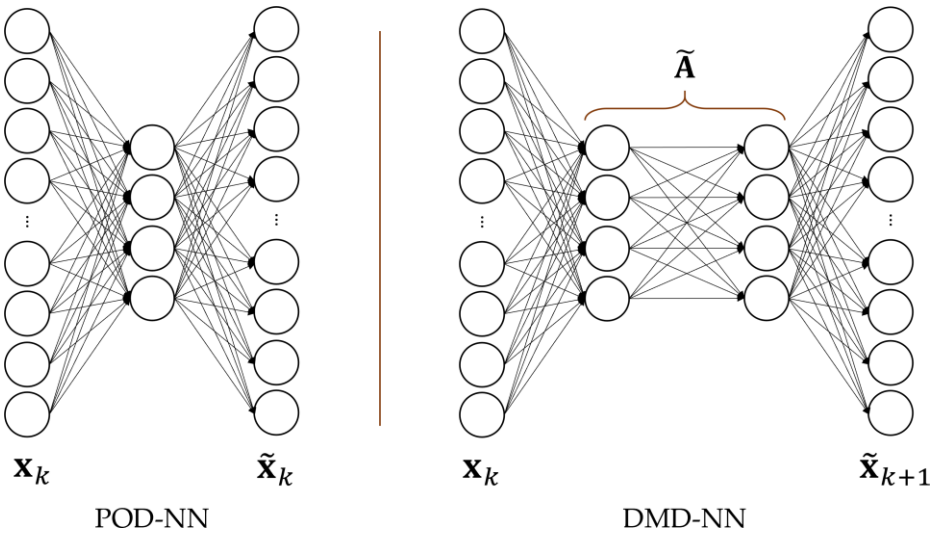

Figure 4-7: Example neural network representations of (left) POD and (right) DMD. The number of layers is fixed, but the number of neurons per layer can change.

4.3.2 Koopman Autoencoder

One only has access to the variables, $\mathbf{x}$, with no knowledge of the target outputs. This lends itself well to an autoencoder design, following on from the basic POD-NN and DMD-NN implementations, with a general diagram given in Figure 4-8, acting on image data. In this case, the network inputs and outputs are data points from the data matrix $\mathbf{X}$. The encoder section learns the transformation ϕ from Equation 4.22 to linearize the system dynamics. Likewise, the decoder section learns the inverse, ϕ^{-1}, to recover the original coordinate system. As the encoder and decoder are inverses of the other, only the parameters of one need to be optimized if an analytical inverse exists. For a feed-forward NN, for example, the transpose of the encoder weights with the inverse of the activation function can be used as the decoder, forgoing any extra parameters. For many cases, such as the choice of activation function or in CNNs, an analytical inverse does not exist. However, a generalized inverse can be approximated by learning parameters in both the encoder and decoder.

In a common autoencoder application, the encoder will take the input space, transform it into an embedding space, then transform back to the input space and try to reconstruct the given input. Indeed, this step can be taken to help learn ϕ. In theory, the only necessary addition is to learn the linear map $\mathbf{K}$ which advances the input forward in time. The linear map can be learned as a, for example, fully connected layer or as a parameterized block diagonal matrix. An example forward propagation pass would be mapping the input to the embedding space, applying $\mathbf{K}$ a number of times, $\mathbf{K}^m$, before the embedding variables are transformed back into the input space where instead of reconstructing $\mathbf{x}_k$, the network is to predict $\mathbf{x}_{k+m}$.

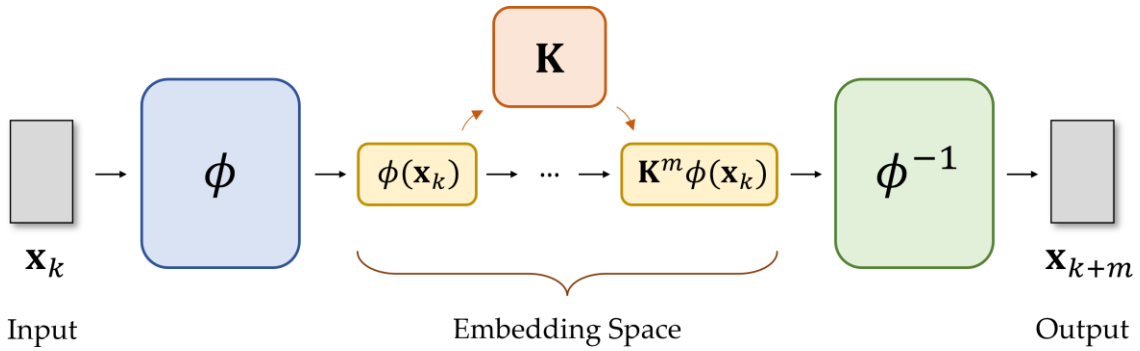

Figure 4-8: A schematic of the fundamental components of the Koopman autoencoder.

A loss function is desired which optimizes input reconstruction to help learn ϕ, with an additional term optimizing future state reconstruction which learns **K** and further helps to learn ϕ. The general loss function for the Koopman autoencoder can thus be written

$$\mathcal{L} = \alpha_1 \left\| \mathbf{x}_k - \phi^{-1}\left(\phi(\mathbf{x}_k)\right) \right\| + \alpha_2 \frac{1}{T} \sum_{m=1}^{T} \left\| \mathbf{x}_{k+m} - \phi^{-1}\left(\mathbf{K}^m \phi(\mathbf{x}_k)\right) \right\| \tag{4.23}$$

where all norms are the mean squared error, T is the number of time steps the network predicts into the future, k is a random time step from the training input, and α_1 and α_2, are constants. The first term is the reconstruction loss and the second term is the future state prediction loss over T time steps.

Fundamentally, this is the structure of the Koopman autoencoder. The transformation to the embedding space captures the Koopman modes which linearizes the system dynamics. The ϕ and **K** sections can have an arbitrary design whose parameters are learned; specific design choices are provided later. While many practical alterations can be applied, they will, currently, be variants of this design.

4.3.3 Deep Koopman Network

Recently, Koopman theory-inspired deep learning networks have demonstrated the greatest model flexibility in generating low-dimensional representations of the Koopman operator [138, 222, 223, 224]. The first published work introducing an end-to-end trainable deep Koopman network is that by Lusch et al. [223], providing the ability to compactly represent dynamical systems with either discrete or continuous spectra. By leveraging the high representational capacity of neural networks, this approach has the potential to perform a modal decomposition successfully on a wide range of problems with minimal algorithmic alteration.

Initial deep learning work in this dissertation, including the above discussions, was developed independently from the publication by Lusch et al. [223], however, important insights from their study were gained to serve as a practical, demonstrable Koopman framework. Their work aligns with the Koopman autoencoder structure and uses a parameterized block diagonal matrix to advance the embedding state

forward in time. A straightforward implementation of learning **K** is to directly learn the real and imaginary parts of the eigenvalues for each mode, $\lambda = \mu + i\omega$. Indeed, this works very well but only for systems with discrete spectra; the eigenvalues do not vary dependent on phase space location. Importantly, they introduce an auxiliary network which learns to parameterize the matrix $\mathbf{K}(\lambda)$**,** dependent on the embedding state location in phase space. Specifically, the matrix $\mathbf{K}(\lambda)$ is of Jordan form

$$\mathbf{K}(\lambda) = \begin{bmatrix} \mathbf{B}_1 & & 0 \\ & \mathbf{B}_2 & \\ 0 & & \mathbf{B}_r \end{bmatrix} \tag{4.24}$$

for an r-mode approximation, where each mode is advanced independently according to

$$\mathbf{B}_i(\mu_i, \omega_i) = \exp(\mu_i \Delta t) \begin{bmatrix} \cos(\omega_i \Delta t) & -\sin(\omega_i \Delta t) \\ \sin(\omega_i \Delta t) & \cos(\omega_i \Delta t) \end{bmatrix} \tag{4.25}$$

for mode i. Not only does this learn a low-dimensional approximation, but it also allows for continuous spectra to be learned directly, which subsumes discrete spectra processes.

The extensions by Lusch et al. to the Koopman autoencoder are shown in Figure 4-9. The main practical contributions are in the auxiliary network and the network's loss function. The auxiliary network takes the first variables in the embedding space directly from the transformation, ϕ, where the eigenvalues of a given mode are a function of their respective embedded modal coordinates. In addition to the loss terms introduced in Equation 4.23, they also include a linearity loss, an infinity norm loss, and an explicit L_2

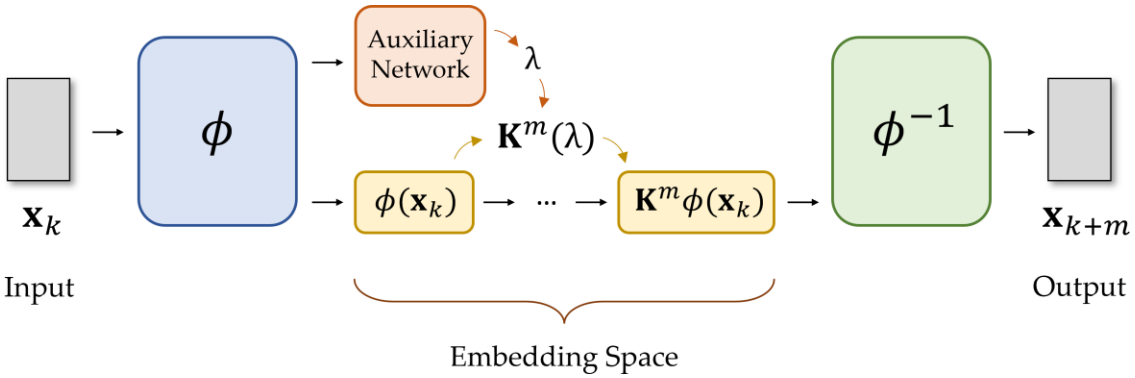

Figure 4-9: The deep Koopman network from [223] suggests learning an auxiliary network to parameterize $\mathbf{K}(\lambda)$.

regularization term on the network weights. The linearity loss acts very similar to the future prediction loss, but helps to ensure the learned mapping is linear in the embedding space. The infinity norm loss penalizes examples with unusually large errors.

This design of a deep Koopman network was validated independently in this work for the same systems presented by Lusch et al. [223]. Due to this deep learning framework only recently being developed, this both helps to verify the results of the publication and ensure independently developed models for this work are learning as expected. A comparison of results is shown in Figure 4-10 for two eigenvectors and their corresponding eigenvalues. Readers are directed to their work for a detailed discussion of the results and problems investigated.

Notable variations between the previous work and this include the use of TensorFlow 2.0 for the current work compared to TensorFlow 1.14, computational hardware and time, and normalization and regularization schemes. However, strong agreement is found and verify both the results of the initial deep Koopman network and the development of the current network.

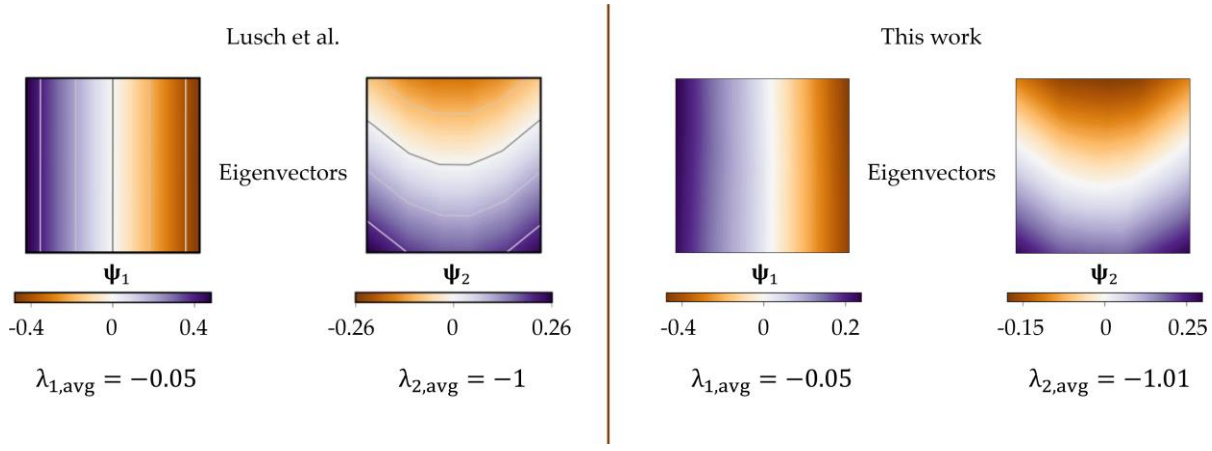

Figure 4-10: A comparison of results replicating the deep Koopman network from [223]. Given enough training, the results will converge to the same solution.

4.3.4 Convolutional Koopman Network

This dissertation presents a deep convolutional Koopman network (CKN) architecture, which naturally adapts the Koopman framework provided above for video data [225]. The primary CKN used in this work is shown in Figure 4-11. Specific details about this network are given in Chapter 5.

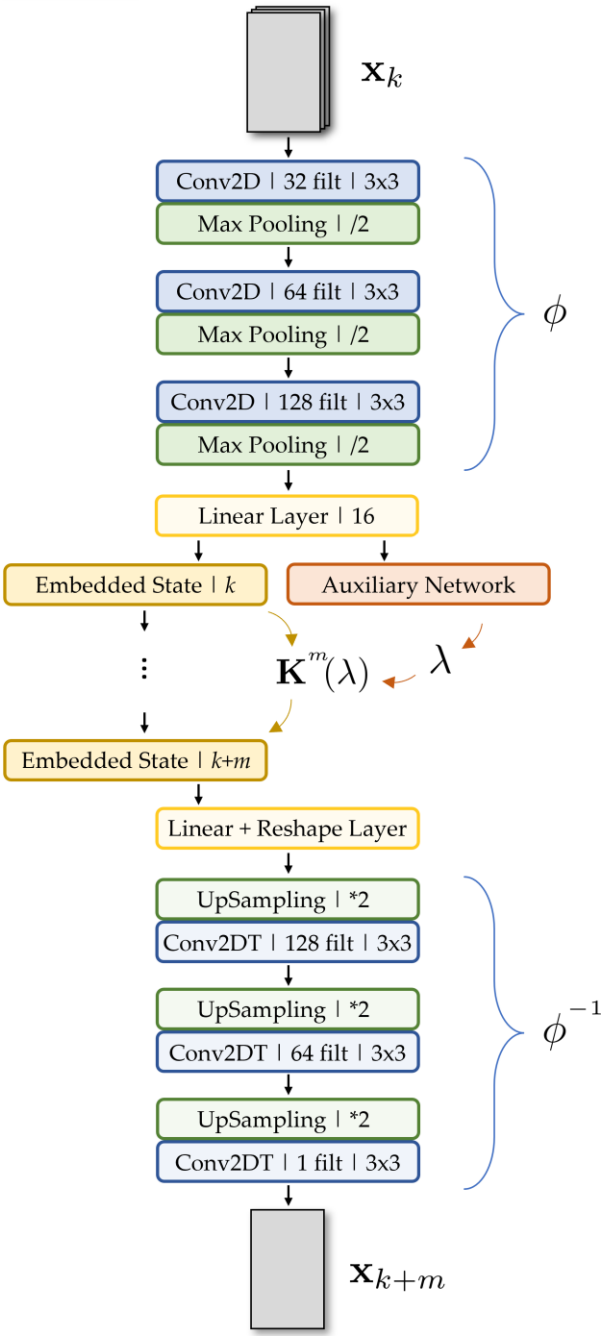

Figure 4-11: One of the CKN architectures used in this work.

To handle image data, the encoder and decoder sections are composed of convolutional and pooling layers, and their transposes, respectively, which learn the coordinate transformation, ϕ, in Equation 4.22. The auxiliary network is a feed-forward NN which learns the eigenvalues which parameterize an approximation of the Koopman operator, $\mathbf{K}(\lambda)$, which, again, allows both discrete and continuous spectra to be identified.

For modal learning, all data are used when training, similar to POD and DMD. Using the CKN architecture primarily for future state prediction, for example, may warrant the use of validation and test data. However, for the purposes of extracting dynamical features and gaining physical insights, this is not considered necessary with the embedding size used in this work.

Prior to training, all data are scaled to have their values in the range $(-1, 1)$ with the mean of the full dataset equal to zero. While videos of experimental systems include noise (e.g., sensor noise), artificial noise was added to the simulated systems.

Loss Function

The general loss function used for the CKN is

$$\begin{aligned}\mathcal{L} = \alpha_1 \left\| \mathbf{x}_k - \phi^{-1}\left(\phi(\mathbf{x}_k)\right) \right\| \\ + \alpha_2 \frac{1}{T} \sum_{m=1}^{T} \left\| \mathbf{x}_{k+m} - \phi^{-1}\left(\mathbf{K}^m \phi(\mathbf{x}_k)\right) \right\| \\ + \alpha_3 \frac{1}{T} \sum_{m=1}^{T} \left\| \phi(\mathbf{x}_{k+m}) - \mathbf{K}^m \phi(\mathbf{x}_k) \right\|\end{aligned} \tag{4.26}$$

where all norms are the mean squared error, T is the number of time steps the network predicts into the future, k is a random time step for the training input, and α_1, α_2, and α_3 are constants. Constant values of unity were found to work well for most networks analyzed, though additional tuning can be made to optimize model performance and convergence further. The first term is the reconstruction loss, the second term is the future state prediction loss, and the last term is the linearity loss in the embedding space as proposed by Lusch et al. [223].

Visualizing CKN Modes

POD and DMD modes are often presented as a single image where, with each associated modal temporal coefficients or frequency and energy, the time evolution of the mode can be inferred from the modal values. The CKN modes, though, provide greater spatial localization than POD and DMD modes and evolve nonlinearly on the input space, as will be discussed to a greater extent in the next chapter. Therefore, the most revealing visualization is to show snapshots of the modal videos and reconstructions at different time steps to demonstrate modal dynamics and to promote interpretability.

For the CKN, producing the modes is straightforward because of the network design. As the modes in the embedding space are forced to be independent through the Jordan matrix $\mathbf{K}(\lambda)$, extracting mode i only requires keeping block i in $\mathbf{K}(\lambda)$, zeroing all other blocks, and passing an example input through the network. The output will produce a modal video in the original input coordinate system which can subsequently be fed back into the input for further state prediction. Each CKN mode, therefore, performs a partial state reconstruction which identifies the direct contribution of each mode.

While it is possible to rank POD and DMD modes by their energy contribution to the reconstruction, the modes for the CKN are not ranked. However, for the systems presented in this work, this is not necessarily a drawback. For the case where the dimension of the CKN embedding space is larger than the number of independent system processes, redundant modes are easy to identify due to their aperiodicity. The remaining modes capture the underlying independent processes which govern the system and, therefore, all are required to achieve an accurate reconstruction or prediction. It is possible to rank the modes by their respective contribution to the mean squared error for future state prediction, which may or may not be suitable for a given system. As a result, no modal ranking is performed for the systems analyzed in this work, but it is acknowledged that this is of future interest, particularly when the embedding space is restrictively small.

4.3.5 Evaluating the Convolutional Koopman Network

As there are many design choices to the CKN, an essential step in its development is a systematic evaluation of candidate design choices. While neural network architectures in literature are optimized specifically for the investigated problem to maximize model performance, this approach is not a high priority for the CKN design choices, in favor of a single design which performs well on numerous problems. Of course, any given metric can be optimized for a specific problem, such as future state prediction for the CKN, but an emphasis here is placed on developing a model which is successful without significant alteration and remaining interpretable. Interpretability is subjective and so optimizing a given metric is unlikely to also make the results more interpretable.

Network Design and Training

General network design choices are evaluated on representative datasets to identify sufficiently performant configurations. Specific design requirements include the number of neurons or filters per layer, the number of layers, filter size, choice of activation function, and choice of normalization and regularization procedures. The goal is not to identify an exact, optimal configuration, but rather to investigate the sensitivity and effects of varying these design options. A deep investigation of hyperparameters, optimizers, and training routines is beyond the scope of this work but should be conducted in the future.

For video data, a number of options for fusing spatial and temporal information exist, where different approaches often lead to various impacts on model performance [212, 213]. As this is central to the problem of applying convolutions to video data, a number of convolutional approaches are compared to identify the most appropriate option in terms of model performance and computational cost.

Specifically, regular 2D and 3D convolutions are tested, as the former is the predominant choice for image-based problems while the latter is also selected as it is a natural extension of 2D convolutions to

integrate temporal information. Additionally, dilated 2D convolutions are also tested, motivated by their ability to identify dense features without downsampling [190, 226].

Koopman Comparison

Once a CKN design and training method have been developed, a direct comparison against POD and DMD can be made. The compactness and interpretability, in particular, will be compared for the same systems analyzed by POD and DMD. As DMD builds on POD and is more suitable for dynamically evolving systems, a focus is placed on comparing the CKN to DMD. As DMD has numerous limitations and drawbacks, the ability of the CKN to overcome these issues is of great interest.

Quantitatively, the reconstruction and future state prediction errors will be discussed. While this is not a measure of gain in physical understanding, it is a desirable trait which can help verify that a mode's spatio-temporal structures are physically meaningful.

Chapter 5

Results and Discussion

5.1 POD and DMD

5.1.1 Decomposition Sensitivity

Before introducing results for specific systems, a brief discussion is given on the sensitivity of both POD and DMD to decomposition parameters. Of particular interest are the number of data points used in the decomposition and the choice of r-mode truncation for DMD. Despite a wide understanding of the lack of robustness in DMD modal energy ranking, sensitivity analyses are seldom accompanied with published works. The common energy ranking in Equation 4.17 is sensitive to the initial condition of the dataset which has been analyzed and improved on by Kou and Zhang [227]. This same analysis has yet to be applied to other aspects of the decompositions.

Here, only data which fully capture oscillatory behavior is considered, i.e., the number of data points used for the decomposition is much larger than the period of the slowest, persisting system process. The first system analyzed, the vibrating string, is used here to facilitate exposition, however an understanding of the system is not necessary.

Due to the energy optimal formulation of POD, the number of frames used for the decomposition has little effect on the ordering of modes, modal structures, or modal values. This is particularly the case for the most energetic modes which capture in excess of 95% of the system's energy content. The only requirement for this to hold is that no anomalous or unrepresentative behaviors of the system are included with a change in the number of frames analyzed. For the vibrating string example, testing 500, 1000, and 2000 frames analyzed had no effect on the first 50 POD modes, which account for more than 99.7% of the system's energy. Transient behaviors generally do not affect the most dominant POD results as they have

relatively low system energy content, but they will affect the ordering of higher order modes, i.e., modes which contribute to the remaining <5% of the system's energy.

While DMD builds on POD, it does not maintain the same energy optimality. DMD is additionally subject to the choice of the r-mode truncation; a necessary dimensionality reduction step for calculating $\tilde{\mathbf{A}}$. As POD is generally robust within the choice of r, this does not cause additional concern, however subsequent DMD calculations are dependent on the chosen value of r.

Figure 5-1 demonstrates the sensitivity of DMD modal ordering as the number of frames analyzed varies. These results are shown for $r = 10$, but similar results are found for other r values. Even for the third most energetic mode, the modal ordering is inconsistent across the three cases. As the number of frames analyzed changes, the modal energy can fluctuate by in excess of 25%, which greatly impacts its ordering with modes at similar energy levels. The change in modal ordering, therefore, is localized to distinct energy groupings. This often leads to, for example, the most 50 dominant DMD modes being consistent across decompositions, but the specific ordering of those 50 changes.

Similarly, Figure 5-2 demonstrates the sensitivity of DMD modal ordering as the value of r varies with a constant 2000 frames analyzed. As with the prior results, the third mode is not consistent across the different cases. DMD is dependent on the POD decomposition, as its modes must be a linear combination of POD modes. Therefore, DMD is constrained to the expressiveness of the set of the most dominant r POD modes, which increases as the value of r increases. For example, the third DMD mode for $r = 20$

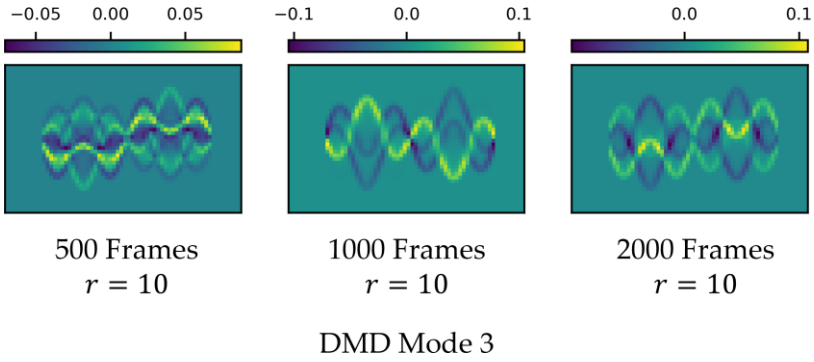

Figure 5-1: DMD modal ordering sensitivity on number of frames analyzed for the third DMD mode.

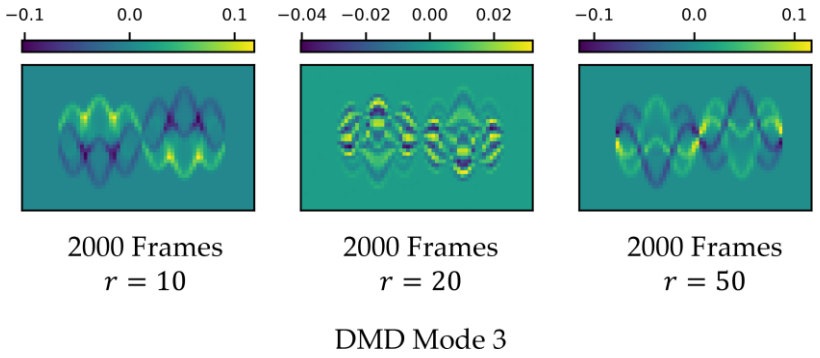

Figure 5-2: DMD modal ordering sensitivity on r-mode truncation for the third DMD mode.

requires a POD mode which is unavailable to the $r = 10$ case. A sufficiently high value of r should be used to limit this effect. Unfortunately, as will be seen later in this chapter, this value grows large for even simple systems.

The above results were also observed for different systems. However, no consistencies were found with the cause of fluctuation and relationship with number of frames analyzed.

5.1.2 Vibrating String

An overarching qualitative discussion for the modal decomposition techniques in general is that of physical interpretation of extracted modes. This can be addressed through consideration of a reverse problem: *given the modes of a system, what are the characteristics of the physical system?* The characteristics depend on the type of problem, but can generally be considered relevant to physical spatial scales, their motion and trajectories, and periodic behaviors.

Generally, system processes combine with or occlude other processes and it is visually unclear what the governing processes are. POD and DMD are first tasked with a simple wave decomposition problem: given optical access to a periodic vibrating string, i.e., without direct access to amplitude measurements, perform a Fourier decomposition of the string vibration.

In this system, a string vibrates linearly according to the superposition of two sinusoids. One sinusoid is one wavelength long with a frequency of 2 Hz and the other is three wavelengths long with a frequency of

3 Hz. Thus, it is expected that two modes are extracted: one for each of the sinusoids. The data are recorded observing the changes in amplitude on the y-axis of the image data. This change in perspective for a Fourier problem makes the problem nonlinear in the observed variables. The interval between frames is 8.8×10^{-3} seconds, 2000 frames were analyzed for both decompositions, and a 50-mode truncation was used for DMD.

The four most dominant POD and DMD modes are shown in Figure 5-3, with the POD PSD for each mode. It is noted that preferential discussion is given to the most dominant modes, according to the used energy metric. While higher order (i.e., lower energy) modes might be meaningful to the analyzed system, this either requires domain-specific knowledge to verify or speculation which is not a useful feature of a decomposition technique.

While some clear patterns of the original string vibration are identified, no mode successfully identifies one of the sinusoids which contribute to the system behavior; the modes are more representative of

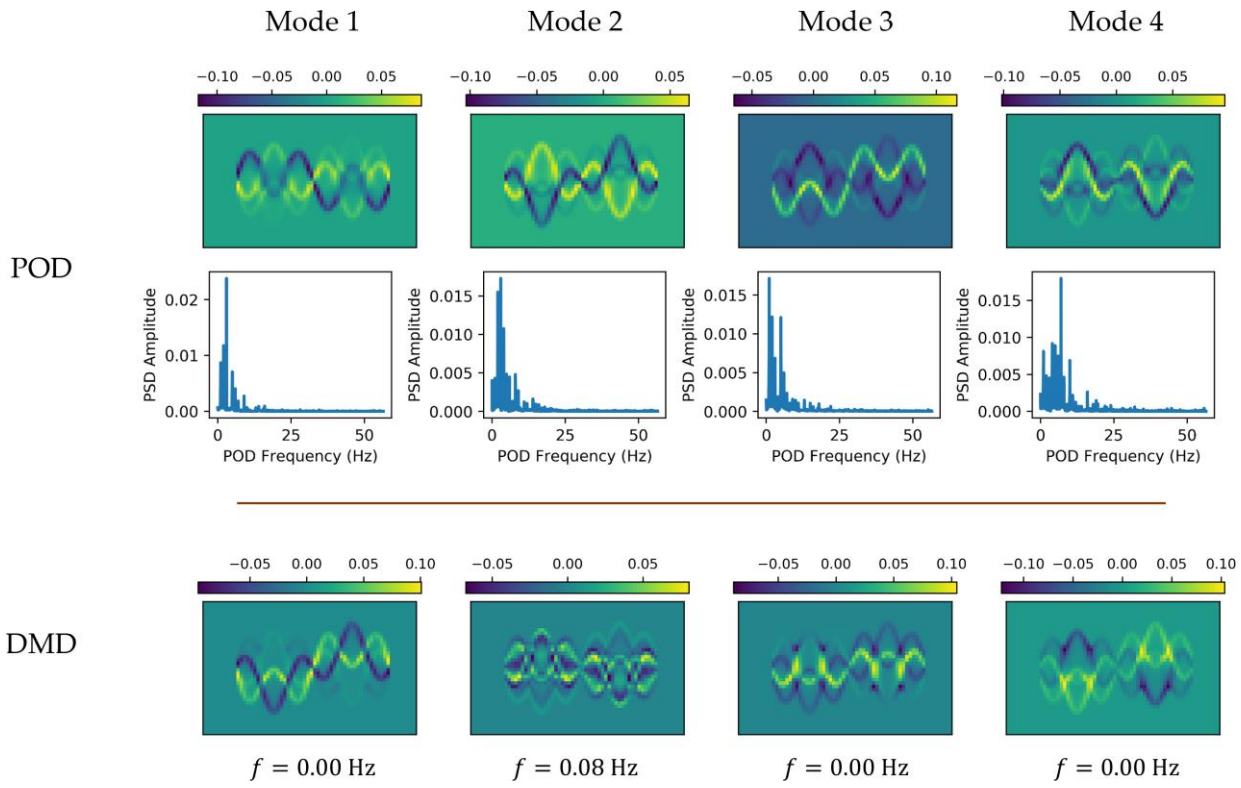

Figure 5-3: The first four POD and DMD modes (columns) for the vibrating string.

snapshots of the data. This becomes apparent by looking at the DMD modal frequencies and the POD PSD for the corresponding modes. The majority of the most dominant DMD modes have frequencies close to zero, which corresponds to quasi-steady state behaviors. The POD PSD for mode 1 and 2 capture a peak signal at 3 Hz, however the PSD is not localized at this peak, the frequencies do not correspond to the underlying spatial structures of the system, and many other modes have their peak signal between 0 Hz and 10 Hz.

The above findings lead to a decomposition which, verifiably, does not capture the correct, independent dynamical features. This is further verified by looking at the POD energy distribution and DMD Ritz values in Figure 5-4. While POD provides a useful energy hierarchy, the modal energies decrease continuously, save for a minor drop at mode 30, precluding a clear truncation which captures the necessary system behaviors. For context of results found in the literature, a drop in energy content of half an order of magnitude is not uncommon for POD modes in adjacent energy groupings. It is therefore difficult, in this case, to justify that fewer than, for example, 30 modes are necessary in both reconstructing and extracting the necessary physics of the system. The latter point is hindered regardless due to the inability to capture the correct spatial structures.

The DMD Ritz values are grouped closely to the real axis, corresponding to frequencies at 0 Hz or the Nyquist limit. In addition, the majority of modes have very high decay rates (i.e., low eigenvalue

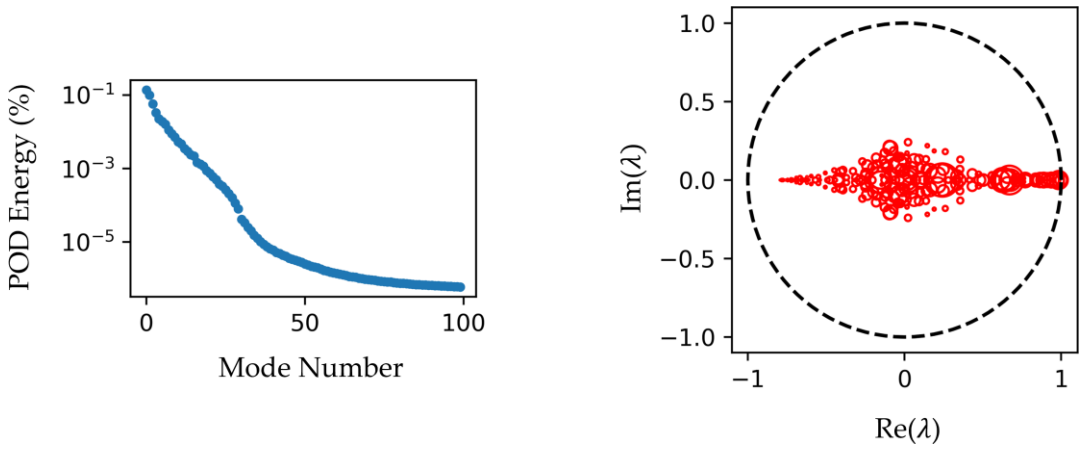

Figure 5-4: (Left) POD energy distribution and (right) DMD Ritz values for the vibrating string.

magnitudes), with the exception of steady or quasi-steady state modes. This suggests DMD will have poor future state prediction, which is indeed the case as evidenced by Figure 5-5. $0\Delta t$ corresponds to input reconstruction, while $50\Delta t$ is predicting 50 time steps into the future. DMD can reconstruct the input image, but since the majority of the modes are static, DMD is unable to reconstruct the time evolution of the system as modes which are not static decay rapidly.

Both POD and DMD are able to reconstruct a given snapshot, though the POD reconstruction is not shown for brevity, but neither are able to identify the correct spatial structures and corresponding modal frequencies. DMD is also incapable of providing any semblance of an accurate future state prediction. The modes and modal characteristics from both POD and DMD suggest that, physically, numerous static processes govern the analyzed system which have complex spatial structures which, of course, is not the case.

These findings are surprising as the underlying system is truly linear, and these techniques have seen widespread use for highly nonlinear problems. As demonstrated by the next system, part of the complication with the string vibration system is that the governing processes are superimposed and are neither spatially nor temporally distinct. POD and DMD provide more physical insight when independent processes are spatially uncorrelated.

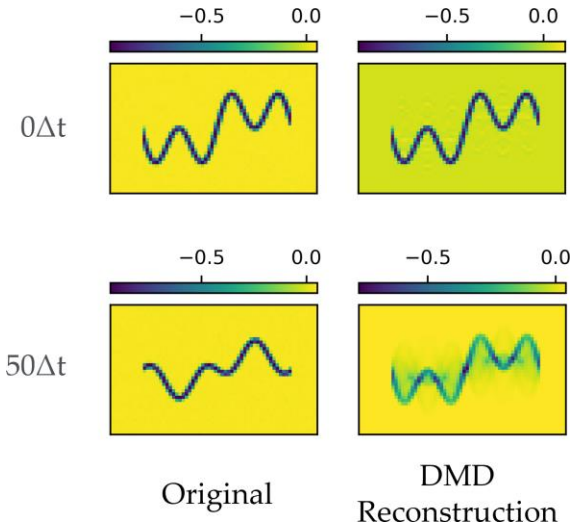

Figure 5-5: A 50-mode DMD reconstruction and future state prediction for the vibrating string.

5.1.3 Particle Streams

The vibrating string problem is trivial when one has direct access to amplitude measurements, otherwise the superposition of the constituent sinusoids greatly affects the utility of POD and DMD beyond snapshot reconstruction. The next system is a group of constant velocity particles whose spatio-temporal scales, trajectories, and frequencies are precisely known. The frequencies of interest, for this case, are related to the periodic looping of the particles in the vertical direction of the data. The interval between frames is 0.02 seconds, 2000 frames were analyzed for both decompositions, and a 50-mode truncation was used for DMD.

The first four POD modes, with accompanying PSDs, and DMD modes are shown in Figure 5-6. Both decompositions extract the same modes, although with a slightly different ordering. Additionally, the POD PSDs identify a single dominant frequency for each of these modes, which agrees exactly with the DMD results. Because of this agreement, POD and DMD are functionally identical for this system, at least for the most dominant modes. The modes are also a lot more intuitive than the vibrating string case: the left and

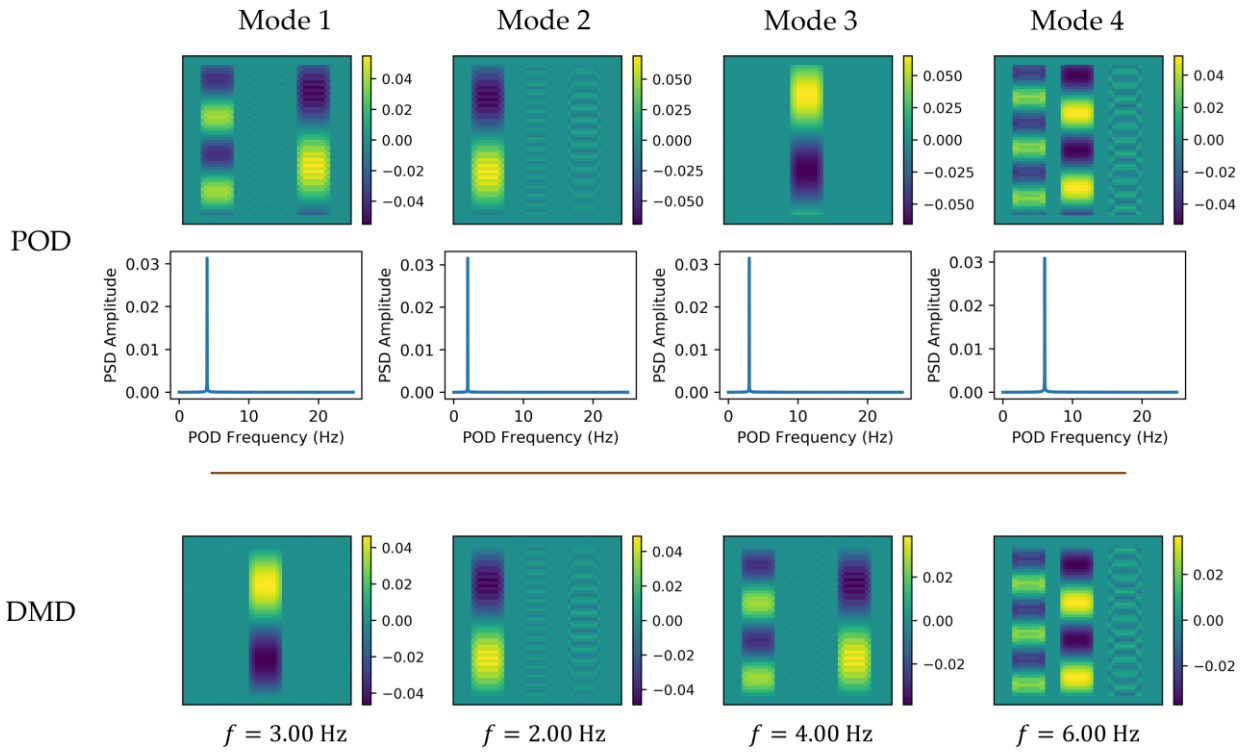

Figure 5-6: The first four POD and DMD modes (columns) for the particle stream system.

center streams are isolated with their own mode with its accompanying modal frequency, and the modal structures span the full length of the particle trajectory. This is also the case for the right stream, although this is coupled with contributions from the left stream.

These results appear a lot more promising than for the vibrating string problem which demonstrates the effect of analyzing spatially independent processes. Not only are the modal structures more interpretable, but the modes are also spectrally clean and capture the correct corresponding modal frequency.

Of course, while more promising, the right stream is always coupled with the left stream. Further, both POD and DMD identify mode 4 as being prominent, oscillating at 6 Hz, despite this frequency's absence in the governing dynamics. In fact, mode 4 is highly representative of other higher order modes: many low energy modes have frequencies which are harmonics of the first three modes and whose modal spatial scales decrease as the frequency increases. This relationship is commonly found in the literature. For ease of explanation, any modes whose frequency is a harmonic of a more dominant mode, and which shares modal spatial structures, is termed a *harmonic mode*. Conversely, other modes are termed *fundamental modes*. For this system, the first three modes are fundamental modes, while all other modes are harmonic modes.

The fourth POD and DMD mode identify contributions to two trajectories as the modal frequency, 6 Hz, is a harmonic of two fundamental modes. Returning to the reverse problem of inferring what system behaviors occur by only investigating the extracted modes, this would suggest cross-trajectory coupling at this frequency. This is not the case for this system, as no cross-trajectory effects were added. The reason for this modal structure covering two trajectories is only because the modal frequency is a higher harmonic of multiple fundamental modes. Suppose it were not obvious that this frequency was directly tied to fundamental modes, it would be easy to interpret this as a distinct fundamental mode which has a single, complex structure.

A reasonable question here is to ask whether the spatial scales of the fundamental or harmonic modes are representative of the physical spatial scales of the system. Certainly, for extracting physics, an understanding of the physical scales and the associated dynamics is highly desirable. Interestingly, this question has not been discussed at length in the literature; modal structures often appear intuitive without a

critical investigation of if their spatial scales exist in the physical system. This leads to many works using modal spatial scales as representing physical scales without any validation.

Clearly, while the fundamental modes capture the correct frequency, they lack the spatial localization to identify the particle size. The modal structure, instead, describes the particle's *trajectory*, as the same mode will be extracted even for particles with reduced vertical height. This trajectory is described by a single positive-negative region over the span of the trajectory. These fundamental modes, therefore, should not be used to infer physical lengths along the trajectory. The horizontal spatial scale can be extracted, but that is trivial as only dynamics in the vertical direction exist for this system.

Analogously to sinusoids with integer wavelengths, the harmonic modal spatial structures provide an integer number of positive-negative region pairs *along the trajectory*, with the fundamental frequency being represented by a single positive-negative region pair which spans the entire periodic trajectory. The j^{th} harmonic is therefore represented by j positive-negative region pairs. In fact, this continues up until the Nyquist limit. This is evidenced by the DMD Ritz values in Figure 5-7 where the most dominant points (i.e., largest markers) are the fundamental modes while all other modes are distributed evenly around the unit circle. All other modes are harmonics of either 2 Hz, 3 Hz, or 4 Hz, hence the uniform distribution, with prominent harmonic modes occurring at multiples of 6 Hz.

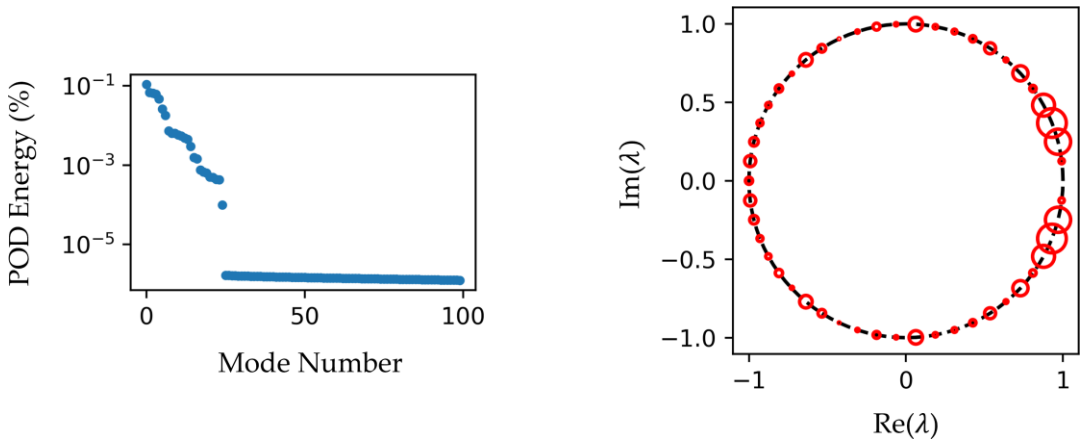

Figure 5-7: (Left) POD energy distribution and (right) DMD Ritz values for the particle stream system.

The harmonic spatial scales decrease by a factor of two with each new harmonic mode up until the Nyquist limit. As there are many harmonic modes present for this system, interpreting modal scales as representing physical scales suggests that a spectrum of physical scales exist in the data. Of course, this is verifiably not true. Further, as the harmonic modes produce integer wavelength structures based on the trajectory defined by the corresponding fundamental mode, it is possible that the correct spatial scale of the corresponding dynamics is never identified for an *individual* mode. For this example, no single mode is able to extract the vertical size of the particles.

Instead, the harmonics are used to help with the spatial localization of the particles. Referring to Figure 5-7, all Ritz values of harmonic modes lie on the unit circle, indicating they are all necessary for a reconstruction. This is the case for both POD and DMD, as more than 20 modes are needed to accurately reconstruct the particles, as demonstrated in Figure 5-8. While 20 modes provide a reasonable reconstruction, more harmonic modes are still needed to sharpen the final reconstruction. The necessary r-mode truncation is clearly depicted by the POD energy spectrum, as 25 modes have a meaningful contribution to the overall system energy. Indeed, this also corresponds to the 25 complex conjugate Ritz value pairs that DMD provides, though this relationship is generally not the case.

The DMD reconstruction error, for different r-mode reconstructions, is shown in Figure 5-9. Both of the reconstructions in Figure 5-8 produce a low error, which reaches a minimum at around 30 modes, which agrees with the expected r-mode truncation from the Ritz values and POD energy spectrum. The reconstruction error also holds for DMD future state prediction error, since the most dominant modes do

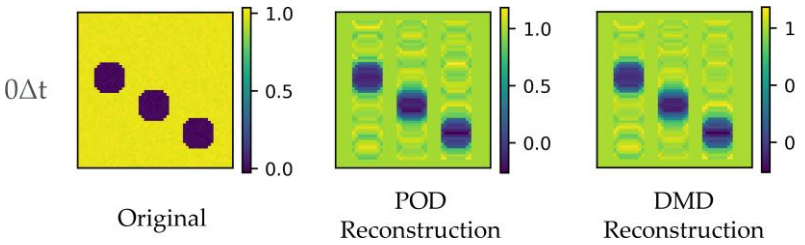

Figure 5-8: Comparison of a 20-mode POD and DMD reconstruction for the particle stream system.

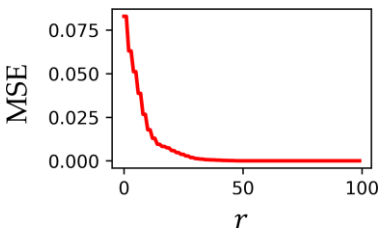

Figure 5-9: Effect of r on DMD reconstruction error

not decay over time, i.e., they lie on the unit circle. The final error is nonzero, which is attributed to unresolved noise. In linear superposition with all other harmonic modes, the harmonics greatly aid in the reconstruction of the original dataset, which seem to indicate they hold physical meaning. For physical length information, this is not the case.

Further, modes can be considered as acting over *temporal* scales, rather than oscillating at a fixed frequency. As discussed, Equation 4.20 can be applied to modes to map from positive (resp. negative) to adjacent negative (resp. positive) regions. A cycle is only completed for the fundamental mode by applying **A** over the temporal scale, whereas the harmonics require an integer multiple more mappings. For this case, the harmonic frequencies do not represent periodic behavior, but rather define the temporal scale with which positive (resp. negative) regions move to negative (resp. positive) regions. This is more descriptive of physical *motion* rather than frequency.

These harmonics can be interpreted as identifying regions where if liquid is present, such as at maximal values, then the liquid will move to the adjacent downstream minimal values in $f_s/(2f)$ time steps. Care must then be taken to identify modes which capture pure motion content or a combination of motion and structure content. Harmonic modes, therefore, do not necessarily capture physical spatial scales or oscillatory behavior; instead, they are all effective in capturing motion behavior over temporal scales associated with their modal frequency and in spatial localization. In addition, the fundamental mode does not necessarily capture the correct physical spatial scales.

Emergence of Harmonics

Harmonics appear to arise very naturally for the particle stream system, and similar structures are found throughout the literature. The emergence of harmonics can be attributed to DMD's close ties with the temporal discrete Fourier transform (DFT) [228], as mean-subtracted DMD has been shown to be mathematically equivalent to the temporal DFT [112]. DMD can be considered a combination of POD with the temporal DFT, where the POD modes define dominant spatial groupings whose temporal coherence is

governed by their temporal DFT, but with exponential growth or decay from their Ritz values. Here, conditions under which harmonics may arise for a DFT are investigated to help explain, quantitatively, the preceding empirical results.

The forward DFT is defined by

$$\hat{x}_k = \frac{1}{n-1}\sum_{r=0}^{n-2} x_n \exp\left(-\frac{2\pi i}{n-1}kr\right) \tag{5.1}$$

for the data matrices in Equation 4.4 with $n-1$ columns. Here, k is representative of frequency with the contribution of frequency k to the signal x being $\hat{x}_k$. There exists $(n-1)/2+1$ distinct frequency values which are linearly spaced.

The temporal DFT decomposes an arbitrary signal into discretized sinusoids of a finite frequency set, $f_k \in \{f_k\}_{k=0}^{k=\frac{n-1}{2}}$. A perfect decomposition therefore relies on a signal $\{f_s\} \subseteq \{f_k\}_{k=0}^{k=\frac{n-1}{2}}$ where $\{f_s\}$ is the set of constituent frequencies in the signal. Suppose the signal of interest, x_s, is a sinusoid with frequency s, the inner product with the DFT sinusoids is given by

$$\begin{aligned}\langle x_s, x_{\mathrm{DFT}}\rangle &= \sum_{\mathrm{r}=0}^{\mathrm{n}-2} \exp\left(-\frac{2\pi i}{n-1}sr\right)\exp\left(\frac{2\pi i}{n-1}kr\right)\\ &= \sum_{\mathrm{r}=0}^{\mathrm{n}-2} \exp\left(-\frac{2\pi i}{n-1}(k-s)r\right)\\ &= \frac{1-\exp\big(2\pi i(k-s)\big)}{1-\exp\left(\frac{2\pi i}{n-1}(k-s)\right)}\end{aligned} \tag{5.2}$$

where the last line is a consequence of the inner product producing a geometric series. Orthogonality is maintained only when $(k-s)$ is a non-zero integer. Should $f_s \notin \{f_k\}_{k=0}^{k=\frac{n-1}{2}}$, the result in Equation 5.2 is never zero. Harmonics therefore necessarily arise, even for the approximation of pure sinusoids.

For spatially-distributed data, a further issue arises. By leveraging POD modes, DMD attempts to identify spatial groupings with shared temporal coherence, whose respective signal is analyzed similar to a

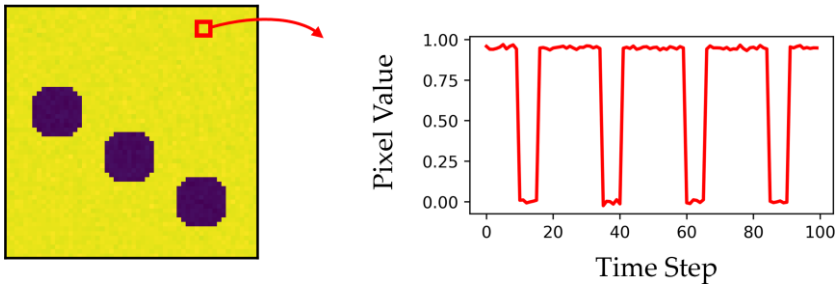

Figure 5-10: Normalized pixel value at a single point in the particle stream system. Note how a square wave is formed.

temporal DFT. However, consider data for the particle streams, or image-based data in general, where state points are uniformly and statically distributed throughout the probed domain. The time history of most state points contains discontinuities; for example, zero or non-zero values are measured when a physical structure is present or absent, respectively. An example pixel-value time history of a single pixel for the particle stream case is shown in Figure 5-10. This signal is close to a square wave and these are the temporal signatures which are analyzed by POD and DMD. As is well known, the Fourier series approximation of a square wave requires an infinite number of harmonics to resolve the discontinuities. As discontinuities are present in the analyzed datasets, harmonics necessarily arise.

Most DMD algorithms do not retain the topological ordering of the underlying data which therefore relies heavily on the independent state point time histories. As discussed here, and shown for the particle stream system, this yields harmonics which do not necessarily capture physical frequencies or structures in the data.

5.1.4 Dilational Jet

Next, a weakly nonlinear experimental system is investigated: a dilational jet perturbed at a known frequency. As there is a velocity differential at the interface of the two phases, Kelvin-Helmholtz instabilities arise, damped due to surface tension which also induces Rayleigh instabilities [4], providing weak nonlinearity.

Numerous works have shown the dependence of sustainable wavelengths with liquid jet diameter (see, e.g., [229]). Therefore, to further test the physical interpretation of harmonic modes, a 900 Hz perturbation was applied to a liquid jet as this frequency induces wave instabilities at a harmonic frequency: 2700 Hz. The interval between frames is 5×10^{-5} seconds, 5000 frames were analyzed for both decompositions, and a 50-mode truncation was used for DMD.

The four most energetic POD modes, with accompanying PSD, and DMD modes are shown in Figure 5-11. This time, only two of the modes agree between the decompositions: the first two DMD modes with POD mode 1 and 3, respectively. For POD, these modes isolate the same modal frequency as DMD with high confidence. POD mode 2 and mode 4, however, are not present in the first 20 DMD modes. The discussion of POD theory in the previous chapter suggested that POD is unlikely to decouple processes

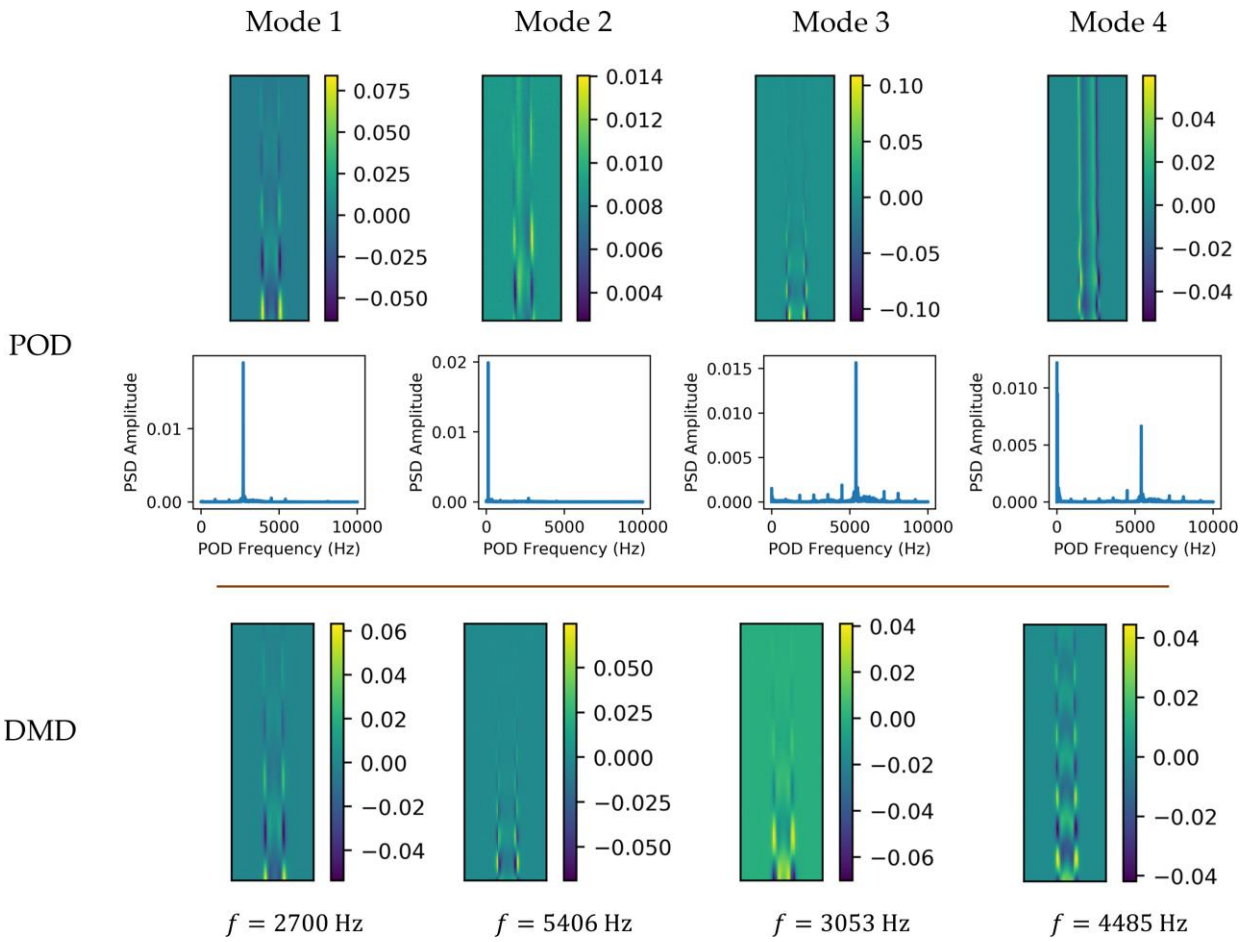

Figure 5-11: The first four POD and DMD modes (columns) for the dilational jet.

which share spatio-temporal coherence. It is hypothesized here that this is the reason for the fourth POD mode's structure: it has peak signals close to 0 Hz, which captures mean flow behaviors, and 5401 Hz, which is captured by the third POD mode. The resulting mode is therefore a combination of these two modal structures. This finding is found for other, higher order modes. POD modes which have multiple localized peaks are therefore likely an artifact of cross-modal coupling.

Both POD and DMD identify the induced 2700 Hz instability and describes wave structures in the first mode. As this is the most dominant mode, this appears to be a fundamental mode, however this conflicts with prior discussion as this mode does not have a single positive-negative spatial wavelength. The modal spatial scales, again, do not agree exactly with the spatial scales of the experimental data, though they are similar. Further, the most dominant modes are harmonics of the 900 Hz perturbation frequency.

Instead, DMD mode 11, shown in Figure 5-12, is considered the fundamental mode: its modal structures span the full trajectory and it oscillates at a frequency which gives rise to higher harmonics. Attributing meaning to this mode, however, is not easy to validate. While the jet is perturbed by a 900 Hz signal, the modal structures are not representative of distinct physical structures of the data. Each subsequent harmonic mode decreases the modal structures by a factor of two which is necessary for spatial localization for reconstructions.

Also, this apparent fundamental mode has a considerably low modal energy, marked in blue on the Ritz plot in Figure 5-12. Harmonics of the 900 Hz mode are distributed around the unit circle, but most of the

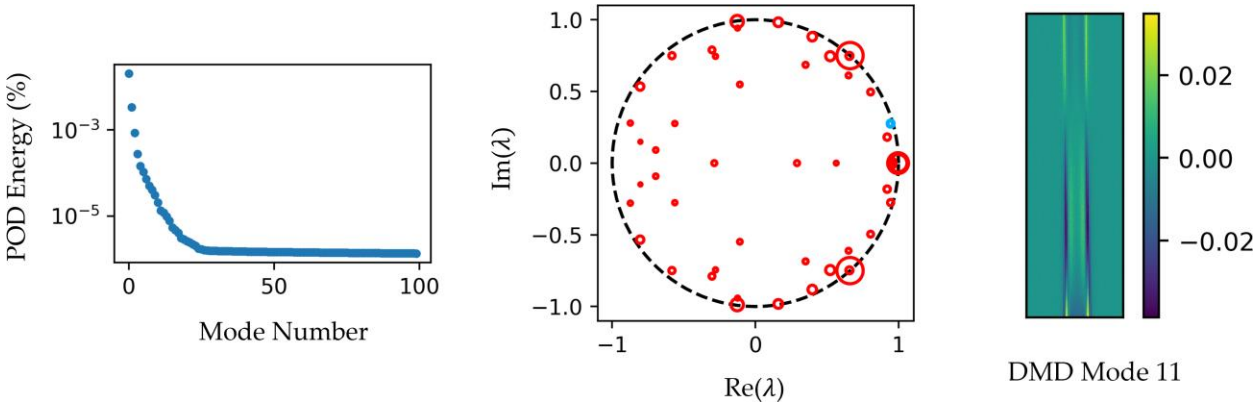

Figure 5-12: (Left) POD energy distribution and (middle) DMD Ritz values for the particle stream system. (Right) The 900 Hz DMD mode, marked blue on the Ritz plot.

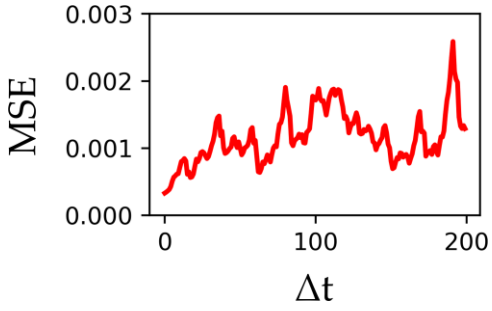

Figure 5-13: Future state prediction error using 50 DMD modes for the dilational jet.

energy is in the first two modes. The POD energy spectrum confirms this, as there are four distinct, dominant modes.

Of course, this is a weakly nonlinear system, so there is a potential asymptotic need for higher harmonics for reconstruction, and harmonics are often present in these systems [230, 231, 232], but this is difficult to validate experimentally. Further, care was taken to ensure only a single dominant instability was present in the system. Considering the results of the previous systems, harmonics will extend up until the Nyquist limit, regardless of if they represent physical properties other than motion. This makes it difficult to verify these modes' physical contribution, but they can still be used to characterize the jet motion over a given temporal scale. Many of these modes lie within the unit circle, so they will decay with increasing future state prediction using DMD. As the prediction extends to longer time scales, the nonlinearities compound and begin to increase the overall reconstruction error, shown in Figure 5-13.

Despite the most dominant mode not capturing the precise spatial scales of the system, they are sufficiently close to provide low reconstruction errors using only a few modes. This is also the case for future state prediction, however all harmonics up until the Nyquist limit are needed for highly accurate and improved long-term predictions.

5.1.5 Jet Breakup

The analysis of the breakup of a liquid jet is of particular interest as it has many similarities with both primary and secondary atomization in practical systems. A perturbation frequency of 4000 Hz is applied to

a liquid jet producing a dilational jet which subsequently breaks up producing large droplets with accompanying satellite droplets, both of which are expected to be captured through modal analysis. The interval between frames is 5×10^{-5} seconds, 5000 frames were analyzed for both decompositions, and a 50-mode truncation was used for DMD.

All of the previous systems have had discrete spectra, so the modal decompositions provide a single frequency for each extracted mode. The jet breakup case, however, as demonstrated in this work, has a continuous spectrum [221] due to fluctuations in the jet breakup length. As a result of the inconsistent breakup length, the location of the primary droplet formation, and accompanying satellite droplet, also fluctuates.

The most dominant POD and DMD modes are shown in Figure 5-14. The first POD and DMD modes correctly capture the perturbation frequency and agree on the spatial structure. This mode can be segmented into three parts. The top part, just prior to droplet formation, correlates well with the pure dilational structures identified in the dilational jet discussed previously. This indicates that this segment primarily undergoes a dilational process rather than jet breakup. The bottom segment identifies clear regions and shapes of formed droplets indicating that jet breakup has generally occurred further upstream. In between these two sections, however, there is a bifurcation of these droplet regions, as if representing a transition from the pure dilational behavior to the pure droplet formation. From the video of this system, the location of jet breakup and consequent droplet formation fluctuates greatly within the captured field of view. This fluctuation is, in part, captured by this mode as the centerline bifurcation is caused by an uncertainty of whether the liquid column is still intact or has broken up within this region. Even the formed droplets in the bottom section are not true droplet shapes but instead may indicate an uncertainty that liquid is present along the jet centerline.

Though not validated, the alternating high and low valued regions in the DMD modes have been attributed to traveling waves or surface instabilities [233]. This interpretation makes sense for mode 1 in capturing the jet dilational behavior and subsequent droplet formation, however, waves of the spatial scales in DMD mode 2 are not observed in the original data, similar to the discussion of harmonic modes of

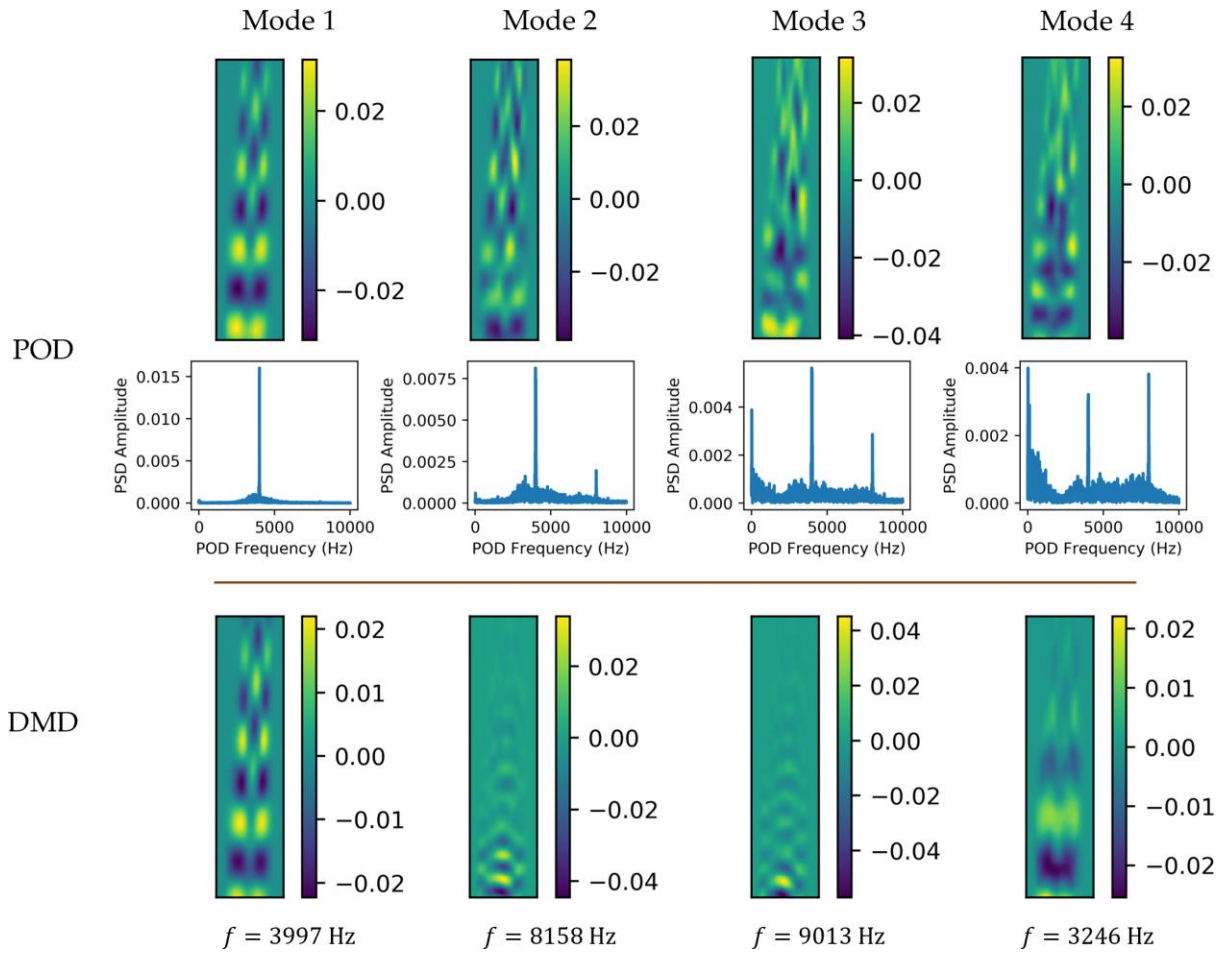

Figure 5-14: The first four POD and DMD modes (columns) for the jet breakup case.

previous systems. Based on previous findings, DMD mode 2 is necessary to improve spatial localization and to even attempt to reconstruct the satellite droplets, though, this is poorly done. Since these smaller spatial scales are not observed through observation of the original data, this suggests that DMD mode 2 does not capture independent physical behaviors by itself.

The sensitivity of POD to cross-modal coupling is exacerbated for the jet breakup case. The corresponding modal PSDs show how only the first mode has a single localized peak. The remaining modes have wide bandwidths with numerous spikes at dominant frequencies, namely 0 Hz, 4000 Hz, and 8000 Hz. This cross-modal dependence is clear when observing the modal structures qualitatively for these modes, as modes 2, 3, and 4 have high correlation to the first mode. Indeed, POD also captures the second harmonic

of the first mode, similar to DMD mode 2, but the modal energy of this mode is considerably lower and contributes to the modal structures of, for example, POD mode 4.

Again, as system complexity is increased, the appropriate modal truncation becomes less clear through the POD energy distribution, beyond the first two dominant modes, and DMD Ritz values in Figure 5-15. All remaining modes have significant energy content and are necessary in an accurate POD reconstruction. Although the Ritz values do cluster close to the unit circle, which relates to the 4000 Hz modal frequency of mode 1 and its second harmonic, mode 2, the remaining Ritz values appear to be distributed randomly within the unit circle. In addition, many of the other Ritz values lie close to the unit circle or have very significant modal energies while providing no clear physical interpretation to what they represent. The seemingly random distribution is attributed to DMD trying to account for the continuous spectrum of frequencies needed to represent the jet breakup length fluctuations.

The poor interpretability of the modes and the distributions of the modal energies and Ritz values lead to very poor DMD future state prediction, and even fairly poor reconstruction performance, as shown in Figure 5-16. With 50 modes, DMD and POD are able to reconstruct an input image, though many artifacts are present. This performance deteriorates rapidly as DMD predicts future states, as even after 10 time steps, the 50-mode truncation converges to a steady-state prediction. An exponential number of modes is required to improve the prediction which dramatically limits the utility of DMD for this system. Too many rapidly decaying modes dominate the decomposition.

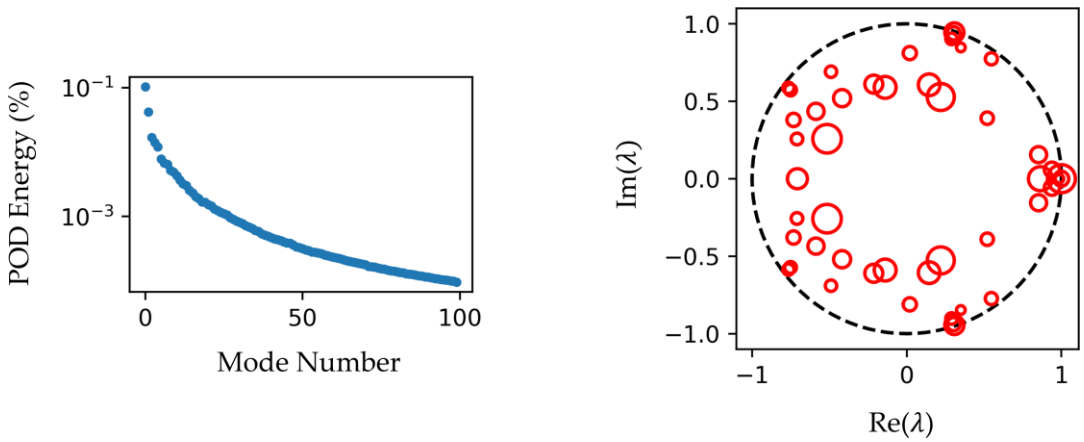

Figure 5-15: (Left) POD energy distribution and (right) DMD Ritz values for the jet breakup case.

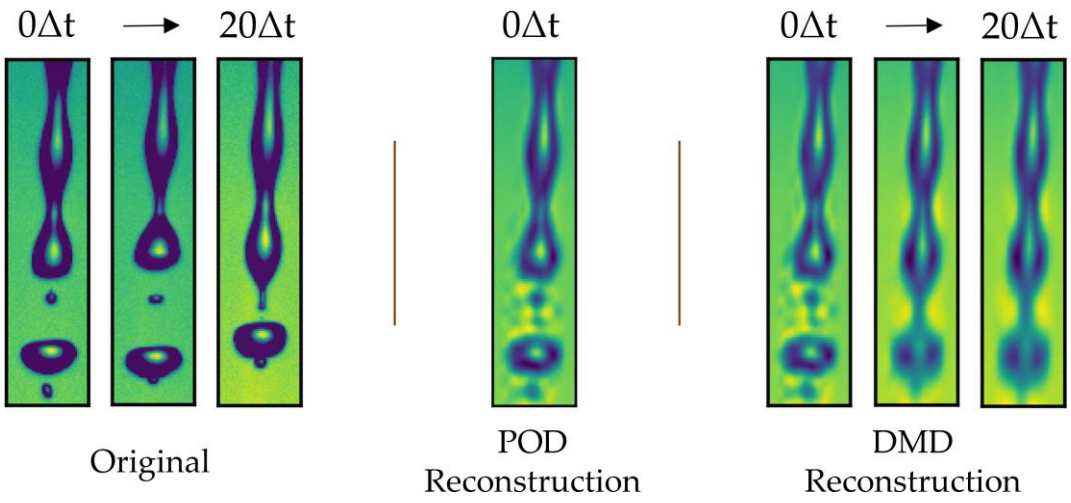

Figure 5-16: A 50-mode POD and DMD reconstruction for the jet breakup case. POD and DMD can reconstruct an input, but future state prediction performance deteriorates rapidly for DMD, even within 10 time steps.

5.1.6 Jet in Crossflow

The previous systems have served to illuminate the behaviors and limitations of both POD and DMD in a verifiable setting. Concerns have been shown for systems which exhibit processes which combine spatio-temporally, where neither system spatial structures or frequencies can be extracted, or for spatially-independent processes, which requires many harmonics to resolve the system spatial characteristics. Yet, both POD and DMD are applied to systems with much greater complexity and with strong nonlinearities.

The final system analyzed is a more practical system which still has some behaviors which can be validated: the jet in crossflow. Specifically, the jet undergoes bag breakup; a regime where liquid segments of the jet are periodically flattened and subsequently ballooned into a bag by the crossflowing gas before atomizing. The periodic behaviors lead to a range of physical scales in the system: large intact ligaments are present in the upstream and between the periodic bag breakup; large primary droplets and smaller satellite droplets are formed as the ligaments disintegrate, as in the jet breakup case; and very small droplets are formed through bag breakup. It is desirable for extracted modes to capture all of these spatial scales and their accompanying dynamics. The interval between frames is 1×10^{-4} seconds, 5000 frames were analyzed for both decompositions, and, initially, a 50-mode truncation was used for DMD.

The most dominant POD and DMD modes are presented in Figure 5-17. Both techniques provide very similar modal structures in their decomposition, which extends even to lower energy modes. The first two POD and DMD modes are commonly found in the literature (see, e.g., [102, 103]), which is often attributed to bending degrees of freedom for the jet as the jet penetration depth fluctuates. Inspecting the original data confirms this interpretation of the first mode, which does oscillate at a low frequency. The second mode, contributing an additional bending degree of freedom, however, is not easy to verify through observation. Additionally, note that the POD modes in general do not provide a single localized frequency, which makes validating any interpretation difficult.

POD mode 4 and DMD mode 3 are also commonly found in the literature (see, e.g., [103, 234]). These structures are attributed to the periodic shedding of larger liquid segments or, conversely, the periodic bag breakup formations in the jet. Indeed, these approximate structures are found in the original data, but this is not sufficient validation for these modes. One issue is that the POD PSD has a very wide bandwidth. It

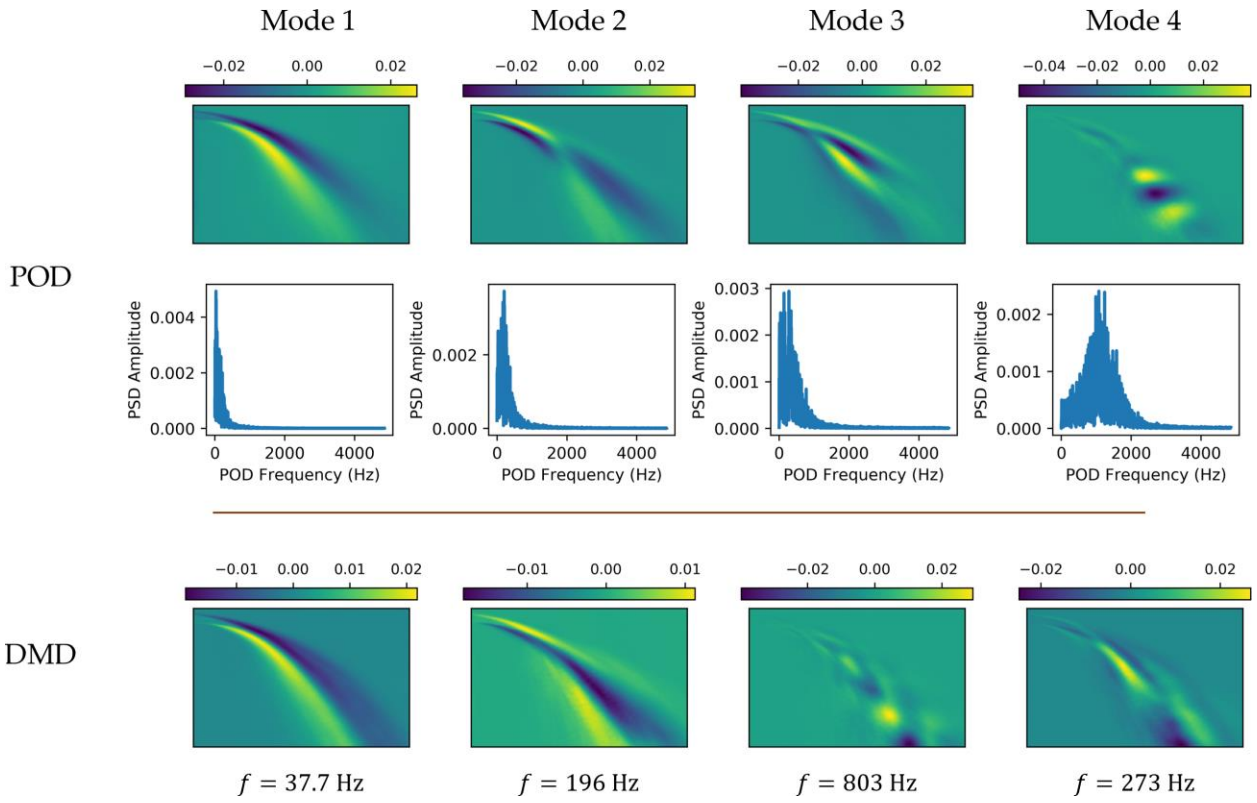

Figure 5-17: The first four POD and DMD modes (columns) for the jet in crossflow system.

centers at 1000 Hz but there is high variance around the mean. Verifying these high frequencies is also difficult, including the spectrally clean DMD frequency of mode 3.

Another issue is in the number of modes which capture the same formations. Four DMD modes, with relatively high modal energy, with the same formations as DMD mode 3 are shown in Figure 5-18, with a wide range of frequencies. As was seen with the previous systems, as the modal frequency increases, the modal spatial scales continuously decrease, but this does not necessarily capture separate processes or physical spatial scales. Further, there is no mode which serves as a clear fundamental mode; many of these modes do not follow an integer multiple of a given frequency. Certainly, this is a complex system whose processes may have continuous spectra, but it is difficult to gain additional insight here.

The POD energy spectrum and DMD Ritz values in Figure 5-19 do not provide extra clarity. Beyond the first two modes, the next 50 are all within one order of magnitude and contribute significantly to the overall system energy. Similarly, the Ritz values are distributed continuously within the unit circle, with only the mean flow mode lying on the unit circle. No dominant frequencies are identified or long-term modes, which may suggest a continuous spectrum is needed to capture the system dynamics.

Unsurprisingly, the POD and DMD reconstructions and prediction are poor for the jet in crossflow system, even when using 100 modes, shown in Figure 5-20. Of particular note is how POD and DMD resolve the different spatial scales, from the large intact ligaments to the small droplets produced via bag breakup. These smallest scales are not resolved as preference is given to larger structures which contain most of the system's energy. This arises from the energy-optimality afforded by POD, but this prevents

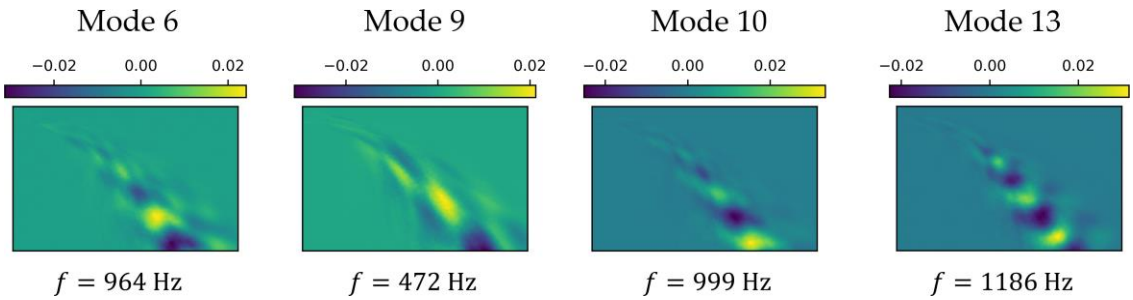

Figure 5-18: Four DMD modes with significant energy content which purport to capture periodic shedding of the jet.

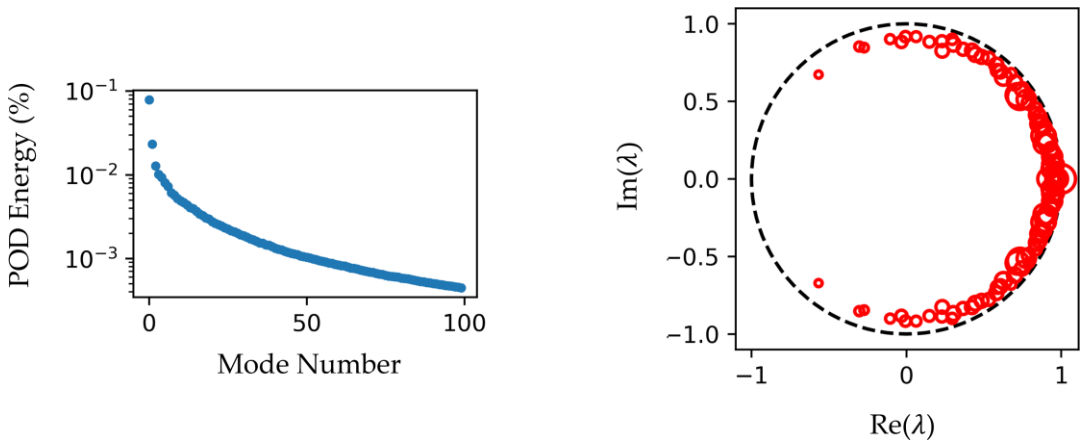

Figure 5-19: (Left) POD energy distribution and (right) DMD Ritz values for the jet in crossflow system.

small scales from being analyzed sufficiently, which is of great interest in atomization studies. Regardless, the time evolution of even large structures is poor, again due to rapidly decaying modes until the prediction converges to the mean flow. The mean squared error of the DMD prediction grows rapidly until the mean flow is predicted.

The presented results for this system do not provide an interpretation or discussion which is easily validated. Due to the concerns with comparatively simple systems, this is expected, but conflicts with the apparent utility of the approaches found in the literature. Certainly, some basic processes can be extracted using these techniques, such as the flapping and bending of the jet or the periodic shedding and bag breakup, but only this high level understanding is extracted, which is already evident from observation of the data. POD and DMD do not provide *additional* insight that can be validated, some modes have been shown to provide neither spatial or temporal information about the system, and their inability to reconstruct or predict the future states of practical systems brings concern to their general utility for these applications.

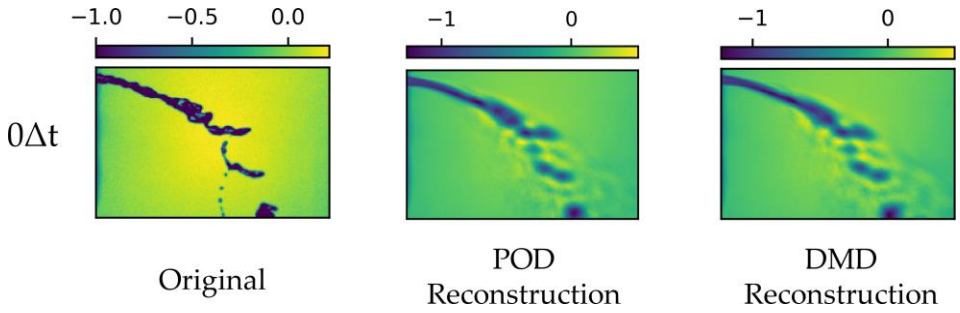

Figure 5-20: Comparison of a 100-mode POD and DMD reconstruction for the jet in crossflow system.

5.2 POD and DMD Neural Networks

Before applying the deep convolutional Koopman network (CKN) to compare against prior POD and DMD results, a logical intermediate step is in representing both POD and DMD in neural network form. POD and DMD have precise mathematical definitions which enables their translation into a neural network framework.

5.2.1 POD Neural Network

Here, POD and principal component analysis (PCA) are considered synonymous, which is often the case for mean-subtracted data. Within computer science, PCA has long been a robust approach for dimensionality reduction and in making predictive models. PCA identifies the principal components (resp. modes) of a dataset, where each subsequent principal component (resp. mode) captures the maximum explained variance of the dataset, in a direction orthogonal to all previous principal components (resp. modes). Only a linear transformation is needed to perform the change of basis into the principal components, then a truncation step can be performed maintaining an appropriate number of components, often keeping >95% of the explained variance. This truncation step performs the dimensionality reduction through knowledge of the principal components and their associated singular values. Finally, the basis is changed back to the original coordinate system.

The concept of a POD neural network (POD-NN) mimics the Koopman autoencoder described in Chapter 4: the input space is transformed into an embedding space before being inversely transformed back into the input space. The main difference is that the transformation is linear, i.e., linear activations are used for the encoder and decoder, and, thus, only a single layer is needed for each. No linear mapping is required in the embedding space. Instead, an r-mode truncation is enforced by compressing the input space into an embedding space with $2r$ neurons, where each neuron represents a complex conjugate mode by reshaping the weights tied to that neuron. By training the autoencoder to reconstruct the input, an energy optimal solution with r modes can be learned.

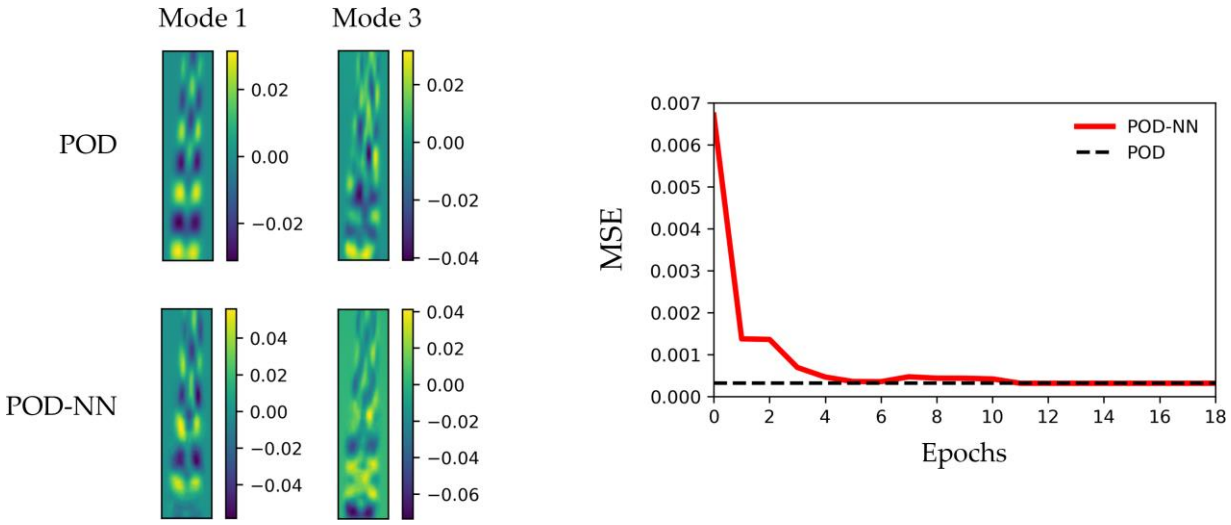

Figure 5-21: The POD-NN modes converge to the POD modes given enough training. Note, mode 3 represents the same structures, but with inverted values.

A comparison of results between extracted POD and POD-NN modes with the 20-mode NN reconstruction error with training time for the jet breakup system are shown in Figure 5-21. As POD treats each pixel independently, only fully connected layers are needed in the POD-NN. The POD-NN results converge to the POD reconstruction error after 10 epochs, with 2000 iterations per epoch. The most dominant modes are extracted accurately once the error converges, though some artifacts are present in the NN results. From experimentation, lower energy modes take considerably longer to learn and may produce different results from the SVD approach. The orthogonality of modes is not explicitly enforced through the POD-NN which does not provide a unique solution.

Additionally, the POD-NN does not itself extract the temporal coefficients, $\mathbf{V}$, nor the modal energy content, but these can be approximated in post-processing once the modes have been extracted.

5.2.2 DMD Neural Network

A major drawback of the standard DMD algorithm is in its assumption that the dynamics evolve linearly. Extended DMD [86] and DMD with dictionary learning [138] address this concern by providing a dictionary of candidate functions which attempt to approximate the dynamics of the system. The latter work was the first to incorporate deep learning into a Koopman framework, acting as a supplement to the standard

DMD approach. Works following up on this idea focused on an end-to-end trainable network for learning Koopman representations, specifically the work by Yeung et al. [222] and Lusch et al. [223]. The intermediate step of performing DMD through deep learning has yet to be demonstrated which, again, acts as a logical connection between the standard algorithm, the POD-NN, and the more general deep Koopman framework.

As DMD uses POD as a foundation, the DMD neural network (DMD-NN) shares the same architectural components as the POD-NN, namely the linear encoder and decoder are also present in the DMD-NN. Recalling the DMD formulation, the dynamics of the system are assumed to evolve according to

$$\mathbf{AX} = \mathbf{X}' \tag{5.3}$$

where $\mathbf{A}$ is the DMD approximation of the Koopman operator, as per convention. DMD utilizes the POD modes to learn a low dimensional approximation of $\mathbf{A}$ through

$$\widetilde{\mathbf{A}} = \mathbf{U}^{*}\mathbf{A}\mathbf{U}. \tag{5.4}$$

Naively, the encoder and decoder capture the POD modes as in the POD-NN, and form the basis of the above transformation to calculate $\widetilde{\mathbf{A}}$. The only addition to the DMD-NN is to learn $\widetilde{\mathbf{A}}$; the linear map which advances the embedded POD space forward a single step in time. Similar to discussions of the Koopman autoencoder, this can be achieved either through a single fully connected layer with linear activations or by learning the parameters of a block diagonal matrix.

Results comparing DMD and the DMD-NN using a parameterized diagonal block matrix are shown in Figure 5-22 for the jet breakup case. The block matrix is favored as it is more intuitive; the eigenvalues are directly learned and one has access to the approximation $\widetilde{\mathbf{A}}$ without further processing. The DMD-NN results converge to a 20-mode DMD reconstruction after 7 epochs, again with 2000 iterations per epoch.

The standard DMD formulation leverages the POD modes. In practice, the DMD-NN design attempts to bypass this step and, instead, learn the DMD modes and modal frequencies directly. Using POD modes as a basis can be forced by first converting the training images into their constituent POD modes, however, this is both an expensive and undesirable step and is not considered. Empirically, while the DMD-NN can

be successful, the extracted DMD modal structures are not unique across training instances, likely due to the reduced POD constraint. The final modal structures are often a random combination of dominant structures identified from the data.

This further impacts the extracted modal frequencies; as the extracted modes have no POD basis and are therefore less constrained, the corresponding modal frequencies are more volatile. The extracted frequencies are often on the same order as the corresponding DMD modes, particularly for the most dominant patterns such as the first DMD mode in Figure 5-22, but fluctuations in excess of 50% are possible.

The current DMD-NN formulation is very simple and computationally and memory inexpensive. The latter point is significant for large r-mode decompositions as the number of parameters for the DMD-NN embedding space scales with $\mathcal{O}(n)$ in space, where n is the number of modes, as only the modal Ritz values are learned. In contrast, standard DMD scales with $\mathcal{O}(n^2)$ in space due to matrix computations.

Of course, the DMD-NN can be adapted to retain the additional constraints of the standard algorithm which would provide more consistent modal structures and frequencies to be extracted. This route is not taken for the current work as the goal is to overcome demonstrated limitations of POD and DMD. Instead, the previous results serve as a starting point on the development of a more general and flexible deep Koopman network.

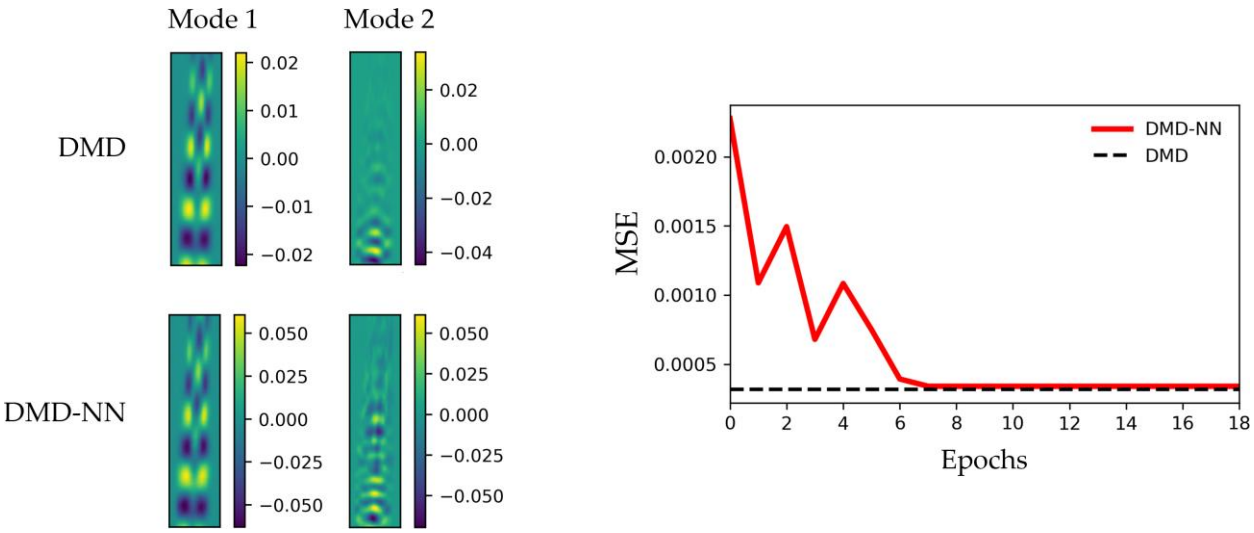

Figure 5-22: The DMD-NN will approach the same loss as DMD given enough training.

5.3 Deep Convolutional Koopman Network

The original deep Koopman network is not well-suited for image and video data but can be naturally adapted with convolutional and pooling layers to form a deep convolutional Koopman network (CKN). The goal of the modal decompositions is to successfully separate independent physical processes in the form of modes, maintain high interpretability, and to reconstruct or predict the system evolution accurately. Results from Section 5.1 have highlighted many similarities in POD and DMD's strengths and drawbacks. Because of this, the developed CKN is compared only to DMD to limit duplicated discussion and to allow for a comparison of future state prediction performance. Many of the benefits of the CKN over DMD is also experienced compared to POD.

For the presented reconstruction results, results at $0\Delta t$ show the networks ability to reconstruct the last input image, while results at, for example, $80\Delta t$ are predictions 80 time steps into the future. The values were chosen to demonstrate long-term future state prediction beyond the number of training time steps and for clarity of exposition.

An arbitrary number of images can be stacked to form the input of the network. To avoid ambiguity in the direction of velocity, three consecutive frames, concatenated together in the channel dimension, were used as inputs to the CKN for all systems investigated.

Generally, for each training iteration, $\mathbf{K}(\lambda)$ is applied up to m-times before being decoded back into the original state space to predict an output. The network is tasked to learn to predict $[0, 1, 2, \ldots, m]$ time steps into the future.

The exact same architecture is used for each system, Figure 4-11 in Chapter 4, except for the jet in crossflow system. Three convolutional layers, each followed by a max pooling layer with pool size 2, make up the encoder, with the transpose of each layer, in reverse order, making up the decoder. Max pooling is non-invertible, so the decoder approximates the inverse of ϕ through up-sampling and transpose convolutions [235]. As will be discussed, the architecture for the jet in crossflow system uses dilated

convolutions instead. The auxiliary network in all designs is a simple feed-forward neural network with three hidden layers with 8, 64, and 64 neurons per layer, respectively.

Eight pairs of complex conjugate eigenvalues were used to parameterize the Koopman operator approximation for each system, which allows the network to identify 8 different modes in the embedding space. The number of modes was chosen to be greater than the number needed to suitably decompose the simplest systems of interest; any redundant modes do not contribute to the system reconstruction, are aperiodic, and either approach the mean flow or are spatially and temporally uniform. For the jet in crossflow system, while a larger embedding space may be beneficial, only 8 modes were used to demonstrate the utility of aggressive dimensionality reduction using the CKN.

Batch normalization [184] and dropout [170] are applied after each layer, with a dropout rate of 0.25. The last layer of the auxiliary network and the decoder and the two linear layers use linear activation functions, all other layers use the ReLU activation function [164]. Models were trained using TensorFlow 2.0 [236].

Batch sizes were kept low as this was found to be less susceptible to converging to the mean flow compared to larger batch sizes. A batch size of 8 was used for all systems except the jet in crossflow system which used a batch size of 4. Many deep learning works have recommended large batch sizes, but an investigation into the effect of batch size on learning from image datasets found that batch sizes between 2 and 32 yielded the best results [237]. Certainly, the optimal batch size will depend on the type of data and problem of interest. It is speculated here that a large batch size identifies general dynamical patterns in video data, i.e., it converges towards mean flow dynamics, so transients or low energy behaviors are more likely to be averaged out.

5.3.1 Input Masking

Without alteration, the CKN will learn efficient Koopman embeddings which allow for accurate nonlinear prediction. However, there is no guarantee that extracted modes are interpretable or capture independent processes. In initial experiments, the network was able to learn embeddings which were *too* compact; a single mode could capture multiple independent processes. Interestingly, this has utility in other applications, such as through neural video compression.

Of course, compression which does not disaggregate independent processes inhibits system understanding and is not desirable here. Consider the string vibration system: a linear system with two constituent linear processes but which is recorded in a nonlinear representation. Not only can the CKN identify the necessary nonlinear transformation to linearize the dynamics, but it is able to capture both processes with a single mode, as shown in Figure 5-23. This is great for compression, however it clearly does not improve one's physical understanding of the system.

Similarly, for the particle stream system, fewer than three modes are needed to accurate predict the future states of the system. Similar to the POD and DMD results, the two streams with an integer multiple frequency were highly coupled. In extreme cases, all three particle streams were captured by a single mode, where a 1-mode reconstruction is compared to the original snapshot in Figure 5-23. The left stream, with the fundamental frequency of the right stream, identifies two particles, again similar to the POD and DMD harmonic modes.

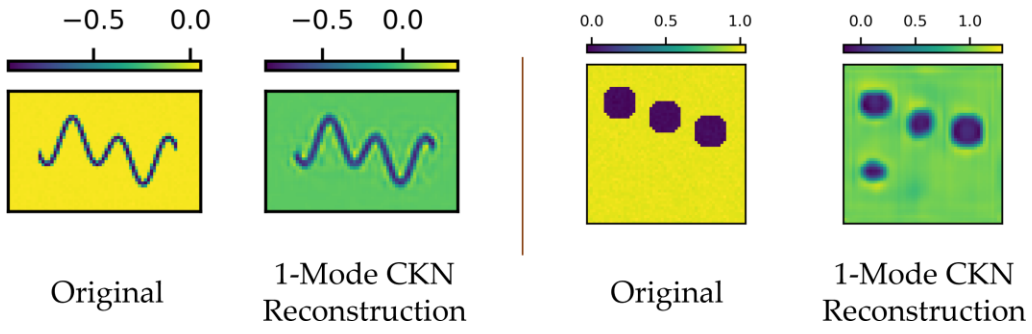

Figure 5-23: Initially, the CKN modes learned are *too* compact. Only a single mode is needed to get high future state prediction for the vibrating string and particle stream systems.

To overcome this, a simple input masking scheme was implemented which robustly identifies system modes and improves interpretability. Input images are masked with a random number of, possibly overlapping, squares of random side length. While the random values depend on the systems of interest, 2 or 3 squares with a side length between 10% to 30% of the image width or height worked well for this study. The masking value was either that of the background or zero. An example of the input masking is shown in Figure 5-24. For all systems analyzed in this work, three sequential images were used as the input and the random masking was applied to each image identically.

For all systems except the simulated particle streams, the masking was applied only to the input images; the network was still to predict the unmasked future time steps. While this approach was still effective for the particle stream system, a more aggressive masking approach was required to fully disaggregate independent processes (i.e., remove the second particle in the left stream). For this case, the same masking was also applied to the target output images.

The intuition behind the effectiveness of this approach is that the masking blocks the contribution of certain independent processes, thus allowing the network to focus on and identify other independent processes. As will be shown, this is successful regardless of the spatial independence, or lack thereof, of the processes.

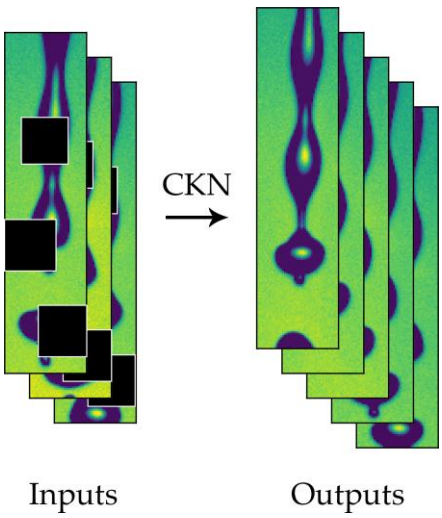

Figure 5-24: Input images are masked at the same spatial locations, represented by the blacked-out squares. The CKN is tasked to predict the unmasked future states, i.e., output images.

5.3.2 Convolutions

The three main choices for convolutions investigated in this work are regular 2D convolutions, 2D dilated convolutions, and 3D convolutions, where only the latter explicitly incorporates the spatial and temporal information together. Although the benefits of directly learning spatio-temporal behaviors together are apparent, this often does not improve model performance or is compromised by a large increase in computational cost (see, e.g., [211]).

Regular 2D and 3D convolutional layers were followed by max pooling with a pool size of 2. Both 2D and 3D pooling were compared when using 3D convolutions, but negligible performance difference was found so 2D pooling operations were selected for all comparisons here.

Agreeing with results by Qiu et al. [211], no improvement was found with using 3D convolutions over 2D. In fact, convergence deteriorated as the model was more prone to falling into suboptimal minima or saddle points. Results, when converged, did match the performance of the 2D convolutions but with no additional advantages. Figure 5-25 compares the training loss progression between using 2D and 3D convolutions, with a batch size of 8 for the particle stream case. The time to converge was considerably longer for 3D convolutions than for 2D convolutions, so 3D convolutions were not selected for the final design.

Dilated convolutions offer an opportunity to capture dense dynamical features by avoiding downsampling and instead convolving filters with successively large dilation factors acting on the full

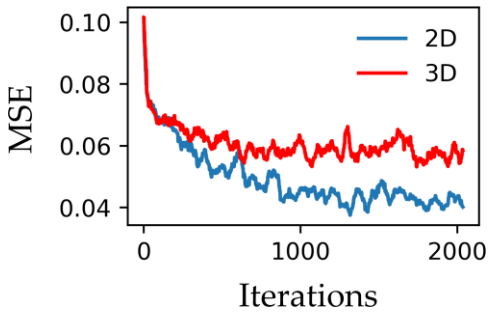

Figure 5-25: Despite directly fusing spatio-temporal information together, 3D convolutions did not outperform regular 2D convolutions for the systems analyzed.

dimensional input image. No increase in model parameters is required, but using dilations allows small scale features and dynamics to be preserved. This is of great importance for liquid atomization systems with small particles, and highly desirable considering POD and DMD's inability to resolve these features.

As is common with using dilations, gridding artifacts may be present in extracted modes or in state reconstruction, but good agreement is still found with regular convolutions. The benefit of dilations, however, is in applications to systems with a wider range of relevant spatial scales. Taking the particle stream system and adapting the size of the particles enables a comparison of how regular and dilated convolutions perform over different spatial scales. The architectures are identical with three convolutional layers in both the encoder and decoder. The only difference is max pooling is used after the regular convolution layers to allow coarser features to be identified, whereas no pooling was used after the dilated convolution layers.

Both networks were trained for 3000 iterations with a batch size of 8. Results comparing the reconstruction of the adapted particle stream system are shown in Figure 5-26. The results highlight the reconstruction efficacy at three scales. The left stream is a 1×1 pixel particle and the middle stream is a 2×2 pixel particle. These scales are prevalent for practical systems, as will be seen for the jet in crossflow system, so an ability to represent these scales and their accompanying dynamics is of great significance. The regular 2D convolutions generally create smoother images as a result of their downsampling and

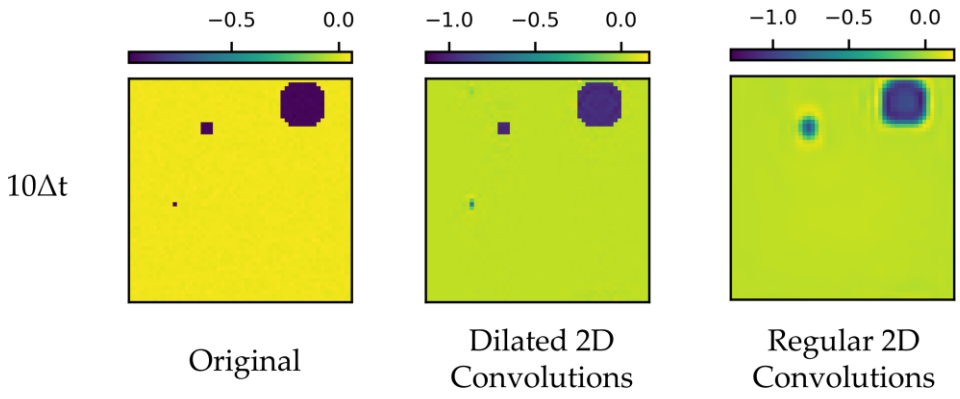

Figure 5-26: Dilated convolutions accurate resolve small scales, even down to the pixel level. Regular convolutions lose these details.

upsampling operations and local neighborhood operations. For many applications, this is acceptable, but it comes at the cost of completely removing very small scale features, as in the left stream, or losing the specific morphology or resolution of small scales, as in the middle stream. Using dilated convolutions, in contrast, maintains the relevance of the smaller scales and produces a highly accurate reconstruction for all scales, particularly the middle stream. In addition, the structures of the right stream are much cleaner compared to the blurred reconstruction from the regular convolutions. The main compromise of dilated convolutions is the presence of extra modal artifacts in extracted modes, which will be explored for the jet in crossflow system.

For the systems analyzed here, only the jet in crossflow system benefits from the use of dilated convolutions due to the wider range of spatial scales. Therefore, all systems use regular 2D convolutions, and identical architectures, except for the jet in crossflow which uses dilated convolutions with no pooling or upsampling layers.

5.3.3 Loss Function

The loss function used for all systems is

$$\begin{aligned}\mathcal{L} = &\left\|\mathbf{x}_k - \phi^{-1}\left(\phi(\mathbf{x}_k)\right)\right\| \\ &+\frac{1}{T}\sum_{m=1}^{T}\left\|\mathbf{x}_{k+m} - \phi^{-1}\left(\mathbf{K}^m\phi(\mathbf{x}_k)\right)\right\| \\ &+\frac{1}{T}\sum_{m=1}^{T}\left\|\phi(\mathbf{x}_{k+m}) - \mathbf{K}^m\phi(\mathbf{x}_k)\right\|\end{aligned} \tag{5.5}$$

which is the same formulation as Equation 4.26 where all constants α_1, α_2, and α_3 are unity. It is emphasized again that a single loss function which performed well on many systems was of interest and, thus, these constant terms are not to be taken as optimized for a given system. Through experimentation, an equal weighting across the three terms was found to give the most consistent results which maintain high interpretability, uncoupled modes, and high future state prediction accuracy for all systems.

$T = 5$ was used for all systems except for the jet breakup and jet in crossflow systems, which used $T = 15$. Low values of T were found to be the most stable in training and robust in convergence. As the complexity of the system increases, a greater value of T was needed to improve long-term reconstruction error but suffered more from local saddle points or minima.

5.3.4 Vibrating String

The results of the modal decomposition by the CKN are given in Figure 5-27, comparing a two-mode reconstruction with the original data. Without access to the intrinsic amplitude measurements, the network is still able to successfully decompose the waveform into its constituent sinusoids. The network has learned to associate pixel row with wave amplitude; the reconstructed wave amplitudes (pixel rows) are a linear combination of the extracted modal wave amplitudes, as highlighted by the dotted red line for the $30\Delta t$ reconstruction. The correct spatial wavelengths have been captured, with modal frequencies 2.00 Hz and 3.00 Hz, and the model maintains high reconstruction accuracy even hundreds of time steps into the future despite being trained on 5 time steps, as shown in Figure 5-28. Interestingly, while the minimum error is

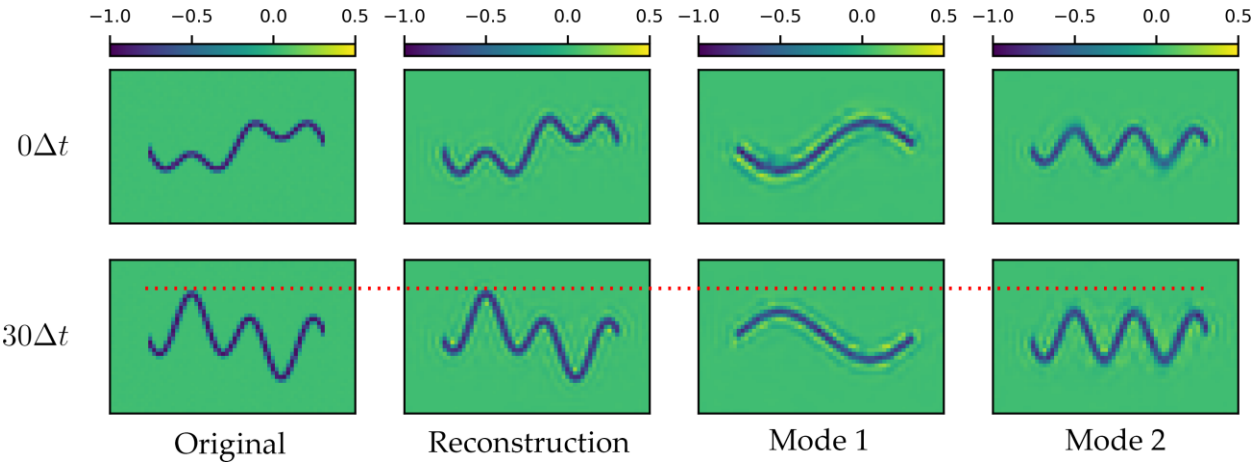

Figure 5-27: Given a nonlinear representation of a superimposed waveform, the CKN learns the nonlinear transformation to linearize the dynamics. The extracted modes are the correct spatio-temporal structures of the constituent sinusoids.

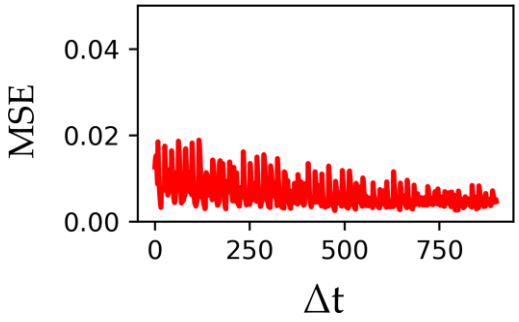

Figure 5-28: Since the CKN extracts the physics of the vibrating string, it can accurately predict hundreds of time steps into the future.

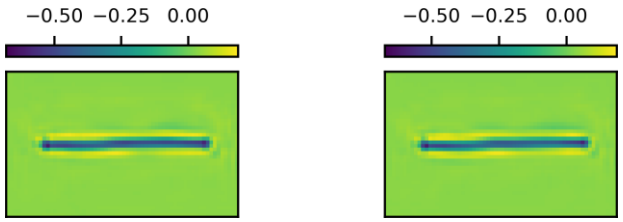

Figure 5-29: Two example redundant modes for the vibrating string. Both modes are aperiodic.

achieved throughout the future state predictions, the envelope of the error decreases as more time steps are predicted. This occurs for different input sequences. The reason for this is currently uncertain, but it may be due to a decaying component of the modes which prevents localized spikes in the error.

In total, 8 modes were used for the CKN decomposition, however only two modes contributed to the reconstruction. Representative examples of remaining modes are shown in Figure 5-29, both of which are aperiodic and capture a flat wave.

As shown earlier, DMD is unable to compactly capture the two modes and is unable to reconstruct the system beyond the first time-step. Importantly, no DMD mode successfully identifies one of the sinusoids which contribute to the system behavior nor are the correct string frequencies extracted. Where DMD may fail due to the variable representation analyzed, the CKN is able to learn the suitable nonlinear transformation to linearize the dynamics and extract a compact, interpretable, and physically meaningful selection of modes. Indeed, this is the goal of variants such as extended DMD [86, 138]. Unfortunately, extended DMD and variants suffer from combinatorial explosion, prohibiting high-dimensional analysis, unlike the end-to-end trainable deep Koopman network approach.

5.3.5 Particle Streams

Recalling the performance of POD and DMD for the particle stream system, while both can extract the correct underlying frequencies of the processes and, with a relatively large r-mode reconstruction, reconstruct to high accuracy, there are numerous downsides. First, the modes which capture the correct process frequency are not sufficient in also capturing the correct spatial localization. Instead, POD and DMD rely on higher harmonic modes to provide this localization which reduces the compactness of the modal representation. Second, if multiple processes share an integer multiple frequency, these processes will be coupled by at least one fundamental mode. The CKN overcomes these issues by providing both spatial and temporal localization for each mode, greatly improving the compactness and interpretability of the representation.

For the CKN, the time evolution of the input data, a three-mode reconstruction, and the contribution of each mode are given in Figure 5-30. The network is able to successfully disaggregate each particle stream and its frequency, although some artifacts are present. As well as capturing the correct modal frequencies, each mode provides a clear physical understanding of the system. The partial state reconstruction, particularly when viewed as a video, does not include contributions from other processes so each mode

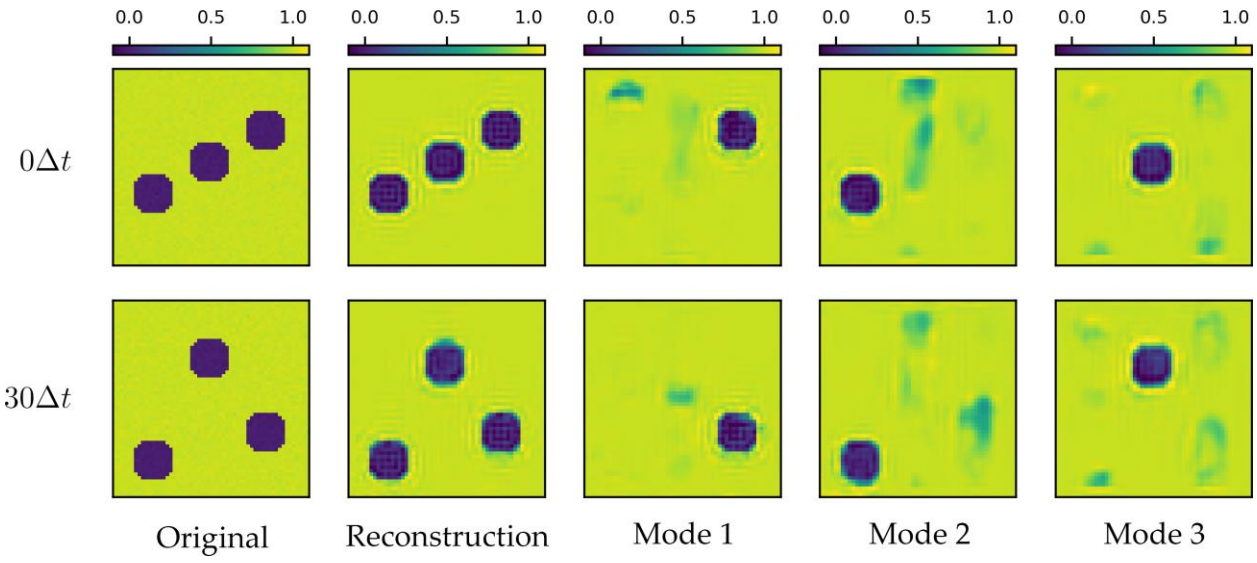

Figure 5-30: The CKN is able to extract the three underlying modal frequencies of the particle streams and the localized spatio-temporal contribution of each associated process.

provides accurate spatio-temporal localization. This provides clear interpretation of each mode and makes validating each mode against the original data trivial.

In initial experiments, contributions of mode 2 were also captured by mode 1 due to the shared harmonic frequency. This was prevented by extending the input masking to the outputs. Because of the masking, however, additional training did not dramatically improve the presence of artifacts in each mode. Fortunately, it is easy to confirm that these attributes do not represent anything physical due to their neutral values, but this does warrant caution for the possibility of artifacts which are difficult to distinguish from true modal structures. For the systems analyzed in this work, these artifacts are only a concern for this system.

Ringing artifacts are also present around the individual particles, as well as slight blurring of the particle edges. These are both well-known consequences of using regular 2D convolutions, with the latter being discussed previously. Ringing is closely related to the Gibbs phenomenon, which often arises at signal discontinuities. There are numerous approaches to suppressing these artifacts (see, e.g., [238]), including the use of dilated convolutions. Of course, dilated convolutions give rise to other artifacts which may be more or less desirable depending on the problem.

Despite the presence of these artifacts, only three modes contribute to the reconstruction which is highly accurate, again for many time steps into the future. Unlike the string vibration case, the reconstruction error does deteriorate over time as shown in Figure 5-31. This is attributed to the non-negligible effect of the artifacts, where small disturbances over time compound until a steady-state condition is predicted, which

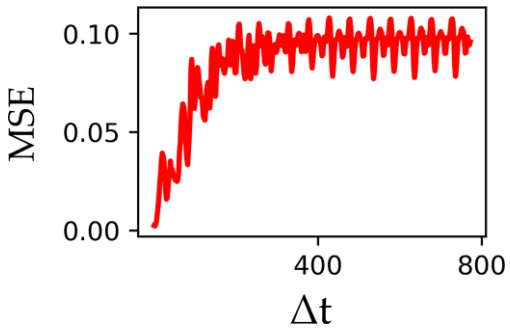

Figure 5-31: The CKN predicts the particle stream future state to high accuracy, but modal artifacts cause deterioration over time.

gives the maximum error. This effect is also found on the extracted modal frequencies: 1.96 Hz, 2.99 Hz, and 3.98 Hz for the left, middle, and right streams, respectively. The values are very close to the true underlying values, but will, over time, deteriorate the reconstruction.

The ability to decouple processes which share harmonic frequencies makes the CKN very attractive, as this is a concern of using both POD and DMD for these spatio-temporal systems. While artifacts are present, the spatio-temporal localization afforded by the CKN maintains interpretability, and easy of validation, for these modes.

5.3.6 Dilational Jet

Both POD and DMD were able to reconstruct the dilational jet to high accuracy, even with only a few modes. Higher harmonics were needed to provide the spatial localization required, but this was not needed to the same extent as for the particle stream system. The DMD results raised questions of the physical presence of a 900 Hz instability, and instabilities at higher harmonics, which were not visibly present in the data; only the 2700 Hz was observed, which should be the dominant unstable frequency.

Only a single mode, oscillating at 2700 Hz, is required to reconstruct the original flow, as shown in Figure 5-32. The CKN captures the correct spatial structures and frequency to predict, again, hundreds of time steps into the future without an increase in prediction error. These results almost appear trivial: a single

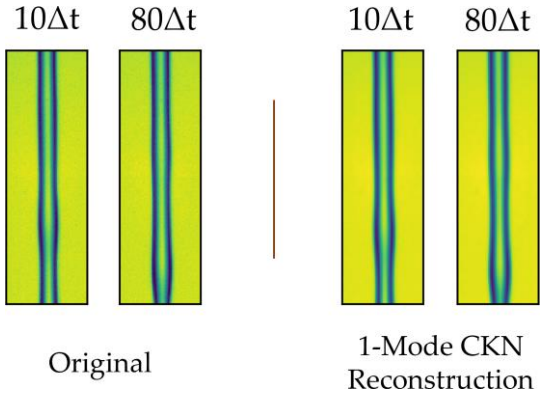

Figure 5-32: Only a single CKN mode, oscillating at 2700 Hz, is needed to reconstruct the dilational jet.

wave disturbance dominates the flow, and this single wave disturbance is identified by the CKN in one mode. Indeed, it *should* be trivial, however, it leads to two valuable insights.

First, the CKN results highlight the lack of compactness and effectiveness of the physical representation of POD and DMD. POD and DMD will provide as many modes as frames analyzed, which will be unique, covering the entire spatial and temporal scales of the system. This is particularly troublesome when discerning the physical meaning behind harmonic modes. In contrast, redundant CKN modes are rarely unique and are aperiodic.

Secondly, the generation of the modes is more intuitive for the CKN. As will be seen in the jet breakup and jet in crossflow systems, each CKN mode provides context for behaviors it has captured. Considering the single mode of Figure 5-32, the mode needs no alteration to reconstruct the original data. DMD, instead, does not provide this context which can hinder the interpretability of a single mode in isolation.

The only concern with the CKN for the presented case is that the jet boundary has blurred due to the small horizontal spatial scales for regular 2D convolutions. At the risk of repetition, the results were repeated for a dilational jet oscillating at 4000 Hz, but at a higher magnification to increase the spatial scales of the jet boundary. The results in Figure 5-33 comparing the CKN and DMD modes agree with the original dilational jet, but with a sharper jet boundary. The CKN and the first DMD mode identify the 4000 Hz frequency while DMD mode 2 is the second harmonic at 8000 Hz.

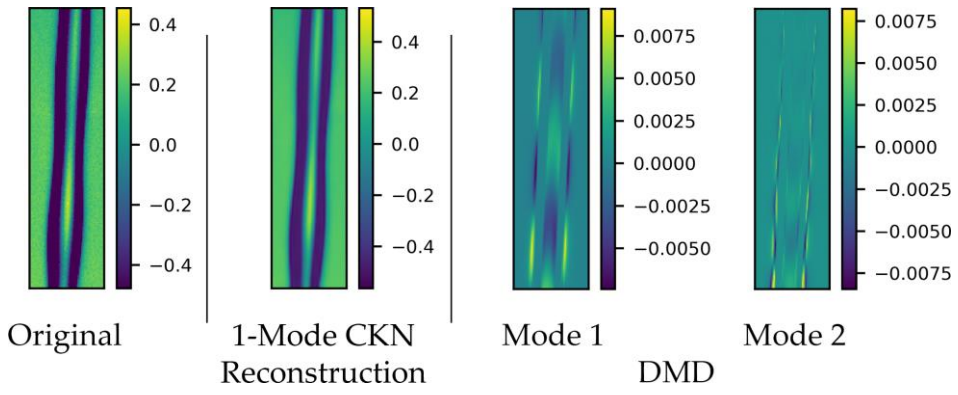

Figure 5-33: Repeatable results are found for a different perturbation frequency and magnification for the dilational jet.

5.3.7 Jet Breakup

All of the previous systems have had discrete spectra, so the modal decompositions provide a single frequency for each extracted mode. While the CKN, in general, learns a continuous spectrum for each mode, it asymptotically approaches a discrete spectrum approximation for these special cases. The jet breakup case, however, as demonstrated here, has a continuous spectrum [221] due to fluctuations in the jet breakup length. As a result of the inconsistent breakup length, the location of the primary droplet formation, and accompanying satellite droplet, also fluctuates. The jet breakup and consequent droplet formation are induced by a 4000 Hz pulse.

The CKN maintains high reconstruction accuracy, even a large number of time steps into the future, using only a 2-mode reconstruction, as demonstrated in Figure 5-34. Importantly, these modes are highly interpretable. Although the breakup length fluctuates, mode 2 captures the 4000 Hz jet breakup and droplet formation, with accompanying satellite droplets, with *stationary* breakup length. Another mode is needed, in this case mode 1, to account for the fluctuation in the breakup length which has a continuous spectrum.

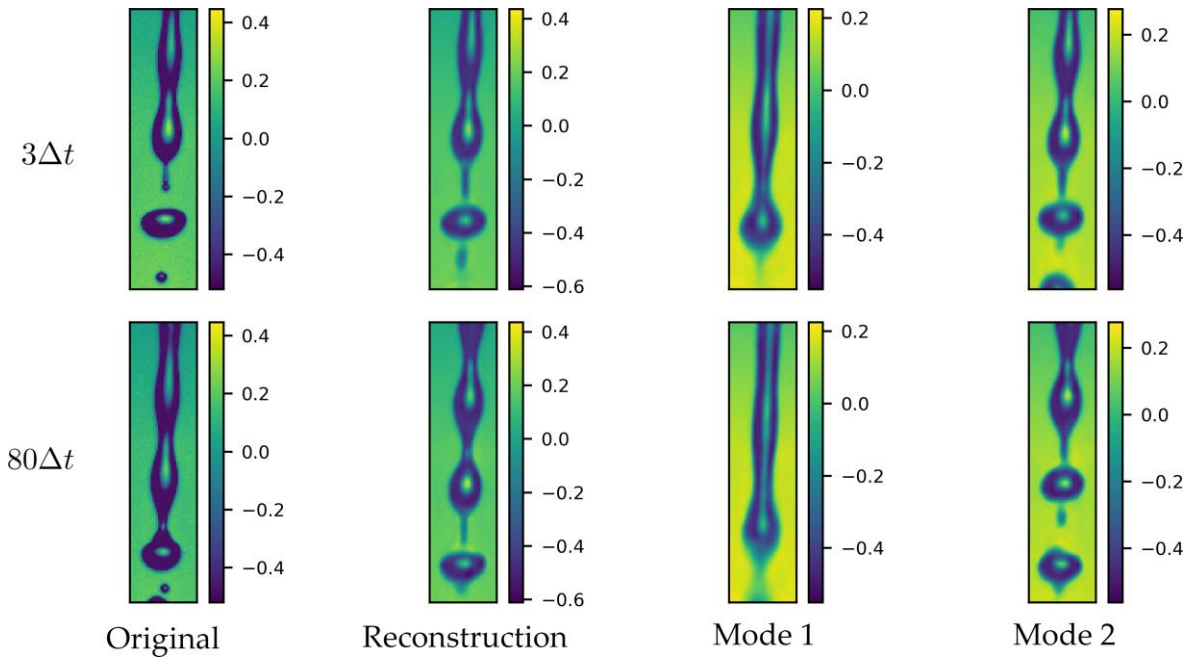

Figure 5-34: Only two CKN modes are needed to reconstruct the jet breakup system. Mode 1 has a continuous spectrum which predicts the breakup length while mode 2 is a pure 4000 Hz breakup with constant breakup length.

Figure 5-35 shows the normalized modal frequency as a function of time as the modal phase space is traversed. For discrete spectra phenomena, such as the pure jet breakup without breakup length fluctuation as identified by mode 2, the normalized modal frequency is constant, regardless of the location in phase space. This is true for all of the previous systems with discrete spectra phenomena. Mode 1, however, exhibits a continuous spectrum to account for the variable breakup length. Indeed, although the spectrum is continuous, a recurring behavior in the modal frequency fluctuations is observed.

The motion of the jet in mode 1 still describes the coarse features of the dilational behavior and subsequent breakup of the jet. Qualitatively, this mode combines with mode 2 to alter the reconstructed breakup length, where thinner jet segments in mode 1 encourages breakup in the reconstruction while thicker jet segments do not. As discussed for the dilational jet, these modes provide the necessary context by identifying their spatio-temporal contributions relative to the overall system. Rather than identifying the static and periodic jet breakup subtractively, i.e., as for DMD modes where skeletal modal structures are provided, the CKN modes are shown as a direct contribution to the underlying system.

The reconstructions are accurate, even 200 time steps into the future, as demonstrated in Figure 5-36, while still capturing variability in the jet breakup length. Beyond 200 time steps, the error begins to increase quickly. The ability for the CKN to predict accurately 200 time steps into the future, when it was only

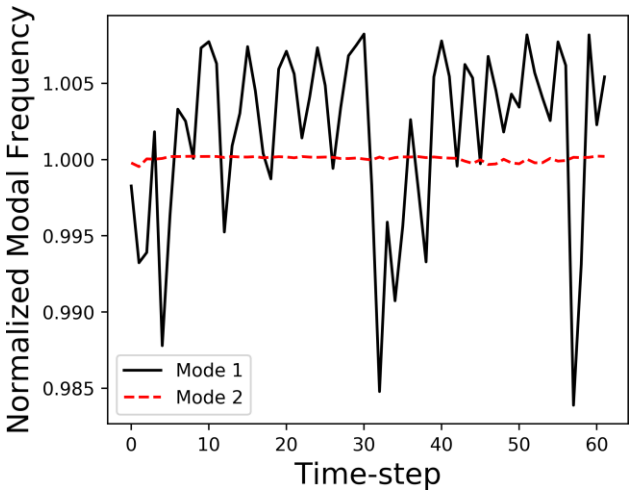

Figure 5-35: A comparison of the change in modal frequency with time (resp. phase space location). For discrete spectra phenomena, like mode 2, the frequency does not change in time. Mode 1, however, requires a continuous spectrum to predict breakup length.

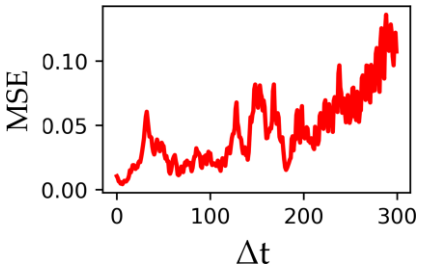

Figure 5-36: Even for systems with continuous spectra, the CKN predicts accurately hundreds of time steps into the future.

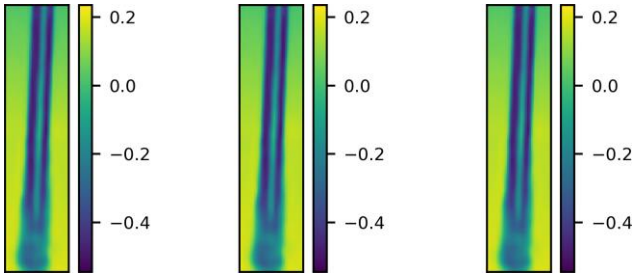

Figure 5-37: Three identical, aperiodic CKN modes which do not contribute to the jet breakup reconstruction.

trained to predict the next 15 time steps, strongly highlights the CKN's ability to identify both discrete and continuous spectra and true underlying system processes. Recalling the DMD results, poor performance was seen after only 10 time steps using 50 modes.

Finally, to reaffirm the properties of the redundant modes, three of the CKN modes which do not contribute to the reconstruction and prediction results discussed are shown in Figure 5-37. These modes are identical and are aperiodic, making identifying their redundancy simple for this case. The remaining redundant modes are also highly similar to these structures and are aperiodic. In fact, in none of the systems analyzed have aperiodic modes been necessary in the reconstruction, unlike for POD and DMD where the mean flow is always separated, captured by the most dominant mode. The CKN naturally incorporates this information into each mode as, by definition, no dynamics are present in static behaviors. In fact, the aperiodicity of CKN modes may be a general metric for identifying redundant modes, however, further work is needed to evaluate this.

5.3.8 Jet in Crossflow

The most practical system is now analyzed with the CKN: a jet in crossflow undergoing bag breakup. Findings from investigating the effects of convolution type led to an adapted architecture which leverages dilated convolutions in order to resolve the finest spatial scales of the system. While large scale dynamics

are very important, the spatio-temporal behaviors of small droplets are equally important for understanding and predicting the atomization performance of a system.

The network architecture for analyzing the jet in crossflow system is shown in Figure 5-38. Though the ultimate goal for the CKN is to provide a single, robust model capable of performing successfully, which

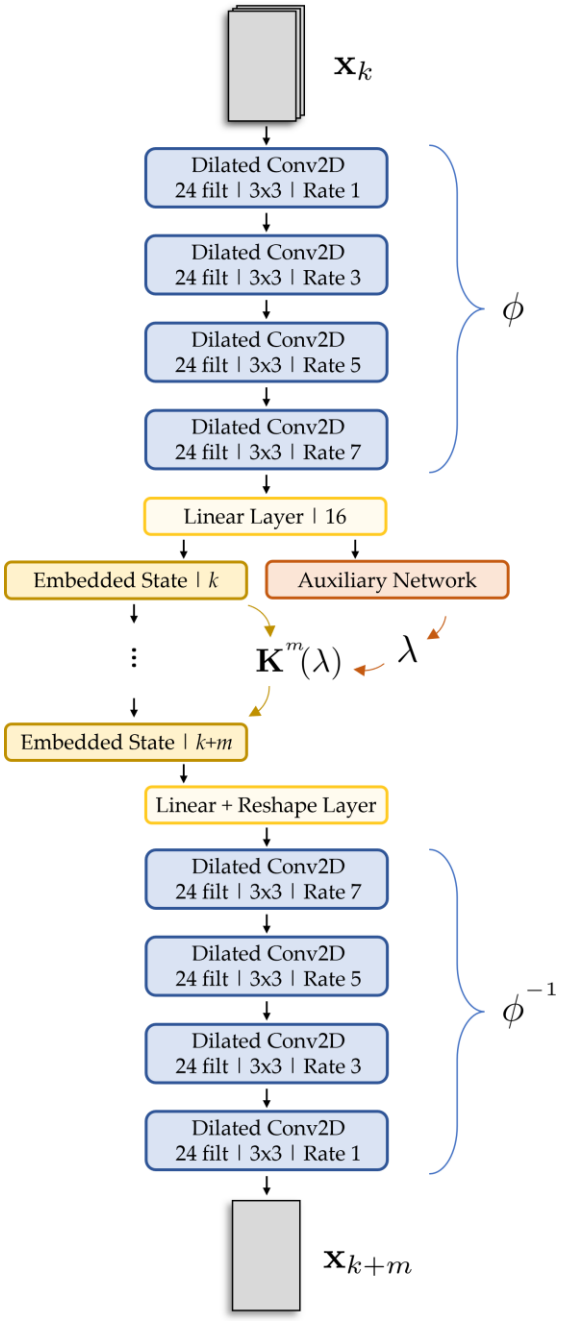

Figure 5-38: The dilational CKN used for the jet in crossflow system.

has already been demonstrated for all previous systems, this additional network is presented as it provides a potential preferred variation which, in initial experiments, can replicate the results seen in the previous sections. The only changes are in the encoder and decoder sections. Pooling layers have been removed, and regular 2D convolutions have been replaced by dilated 2D convolutions. Each dilated convolutional layer uses 24 filters, of size 3×3, which was largely constrained by computational memory requirements as all layers act on the full resolution input. Using 24 filters reduces the number of parameters compared to the previous architecture, so a fourth convolutional layer is used to provide richer representational capacity. The dilation rate grows in increments of 2, starting with a dilation rate of 1, which corresponds to a regular convolution. This increase in dilation rate provides an exponential increase in the receptive field of each neuron [191]. The network was trained to predict 15 time steps into the future and was trained on 8000 frames.

A comparison of the 8-mode CKN future state predictions against the original flow is shown in Figure 5-39. The network is able to capture many of the salient features and behaviors of the original jet given only three consecutive input images. Jet penetration depth, periodic shedding of ligaments, location of bag breakup, and fine droplets are all captured by the CKN reconstruction. Impressively, the jet structure at $120\Delta t$ matches the original jet very closely, with a single larger segment separated from a relatively undisturbed jet.

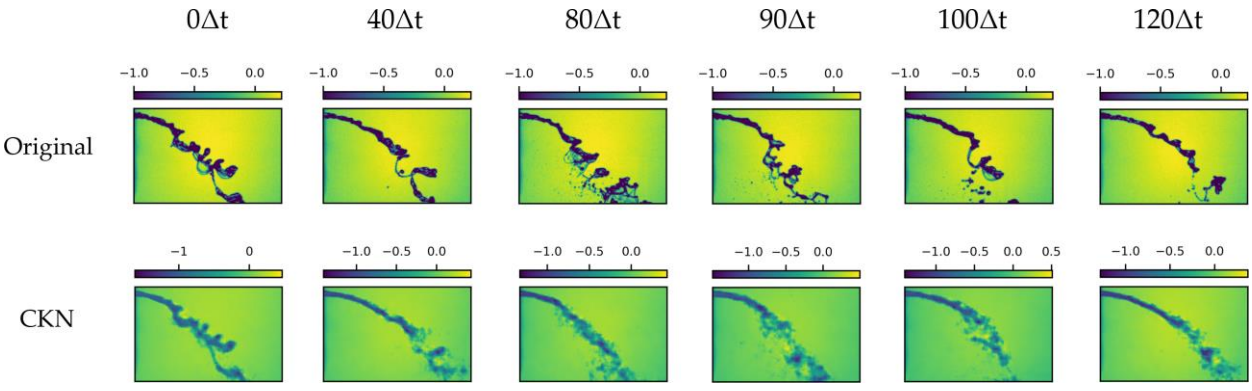

Figure 5-39: CKN future state prediction for the jet in crossflow. Large-scale features and behaviors are predicted correctly even at $120\Delta t$.

This accurate prediction continues hundreds of time steps into the future, with the prediction error growth shown in Figure 5-40. The error grows quickly at the beginning, though on inspection of the reconstruction, many of the salient features are captured. The error then only grows gradually, exhibiting periodic improvements in prediction error as the predicted behaviors align with the actual system. An example of gridding artifacts, magnified for clarity, is also shown in Figure 5-40, which becomes more prominent as future states are predicted. As outputs of the network are re-inputted for subsequent predictions, gridding effects slowly compound, which deteriorates the prediction accuracy over time. Smoothed dilated convolutions offer a potential alternative to suppressing gridding artifacts [192].

The jet in crossflow system is the first application where all 8 modes were required for the reconstruction. A larger number of modes would probably increase the overall performance of the network, and testing with a larger embedding space is a high priority for future work, but this allows evaluation of a restrictively small embedding space. The reader is reminded that modes are labeled numerically out of convenience of discussion. The presented modes are not ordered in terms of modal energy. Further, representing any mode as a still image forgoes its modal dynamics. To address this, modal reconstructions at different snapshots in time are presented which are representative of the general qualitative behaviors observed from their modal videos.

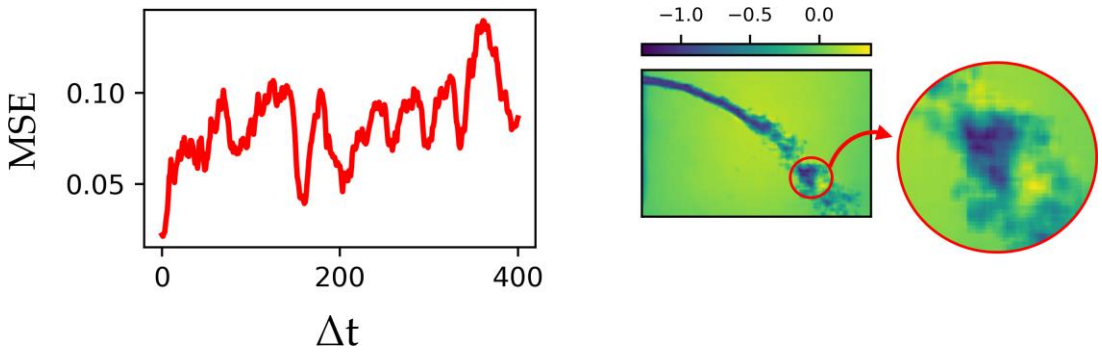

Figure 5-40: Prediction error (left) quickly increases but remains qualitatively accurate for hundreds of time steps. The initial increase in error is attributed to gridding artifacts (right) which compound over time.

Perhaps surprisingly, none of the individual modes require a continuous spectrum; they all converged to a single frequency for their governing process. The reason for this is currently an open question. It is possible that the most dominant processes are truly governed by a single frequency, however, it may also be a consequence of the small embedding space. It is noted that the single frequency behavior occurs in the *embedding space*, not the input space, which is dependent on the input and so the reconstructions are not necessarily periodic. Further, while it is difficult to truly validate the extracted CKN results for this system, the CKN has been validated on systems with known spatio-temporal behaviors with no concerns of spurious modal spatial structures or frequencies. This provides confidence in the extracted results for the jet in crossflow system.

The extracted modes can be split into capturing five distinct features. The first mode, shown in Figure 5-41, captures the general large-scale structures of the flow. In particular, at $30\Delta t$, this mode is able to resolve a large, separated ligament away from the intact jet, which is not identified by any other modes. Subsequent temporal evolutions approximate the shape and overall thickness of the atomization system. The CKN identifies a single frequency for this mode, oscillating at 96.4 Hz.

The second mode, shown in Figure 5-42, attempts to resolve bag breakup behaviors and associated finer particles, though many of the other modes also resolve small scale features as well. Only this mode has a

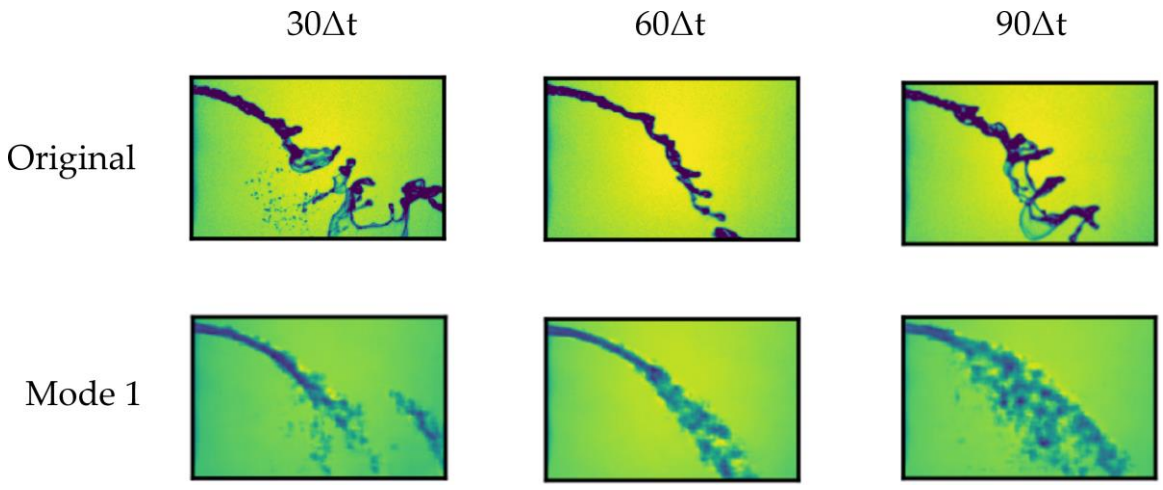

Figure 5-41: The first CKN mode captures the large-scale features of the jet.

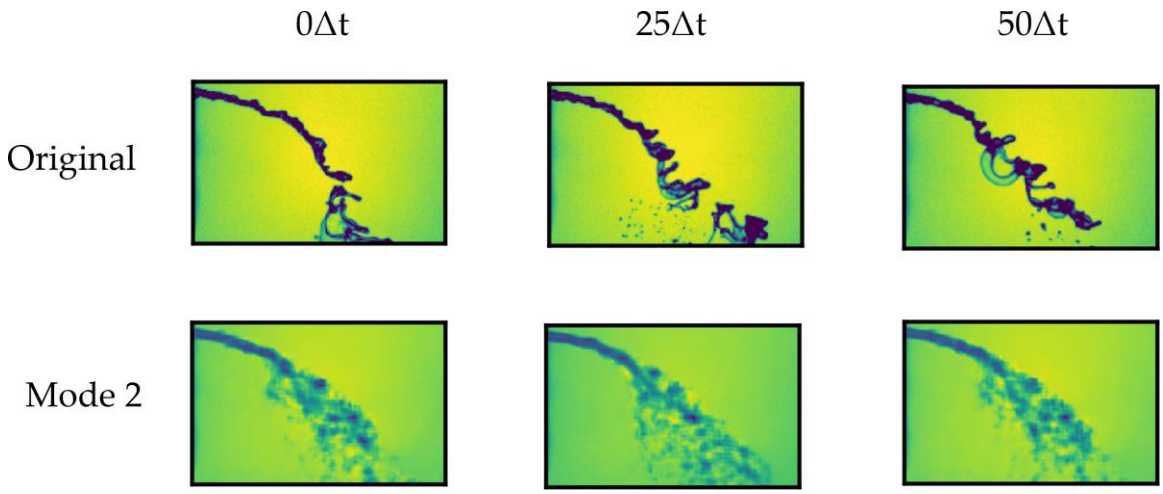

Figure 5-42: The second CKN mode correlates with bag breakup behaviors.

clear correlation with the formation and atomization of bags in the flow, whose modal behaviors are more pronounced during bag breakup. The 25Δt prediction captures the full spatial distribution of an instance of bag breakup. It is difficult to validate this behavior compared to all other modes, but, qualitatively, the above relationship can be found.

The remaining modes come in pairs, which capture the same phenomena but at different frequencies and spatial scales. Modes 3 and 4 both identify the periodic flapping of the upstream jet, shown in Figure 5-43. Mode 3 oscillates at 307 Hz, capturing the full extent of the upstream jet, while mode 4 is a higher frequency flapping at the end of the jet, at 509 Hz. Both POD and DMD extract modes which appear to identify this flapping, which has also been discussed extensively in the literature. One of the recurring concerns with both POD and DMD is in their wide range of represented frequencies. For POD, this comes in the form of wide PSDs, while DMD provides many harmonic-type modes with growing frequency. The CKN identifies two distinct frequencies, whose direct contribution is evident to the overall jet behavior, but at a relatively high frequency compared to the most dominant DMD modes. As DMD requires many modes to represent the jet penetration fluctuations, and with concerns with the physical meaning of extracted frequencies, it is difficult to identify a true underlying frequency.

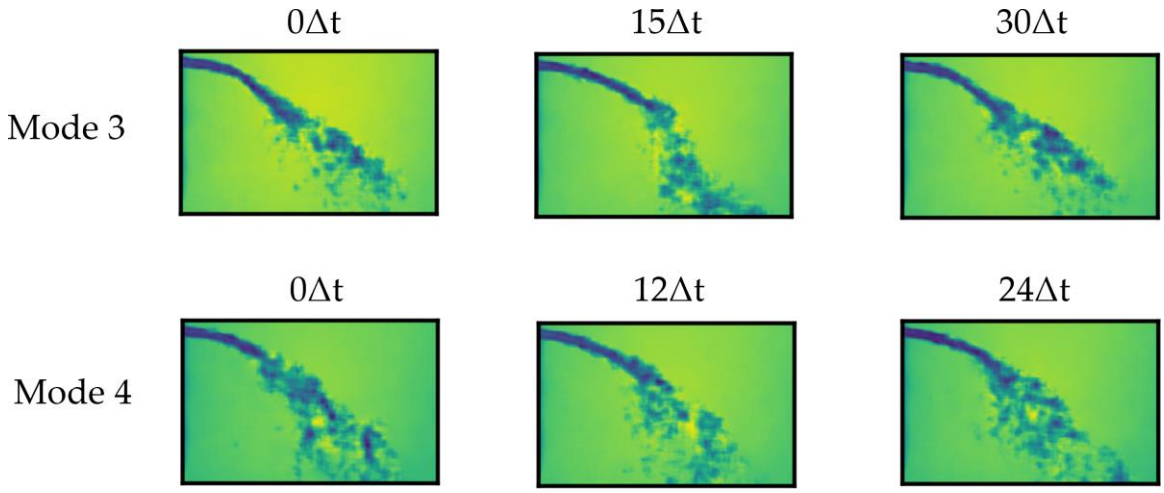

Figure 5-43: The third and fourth modes both capture upstream bending of the jet. Mode 4 is at a higher frequency.

Next, modes 5 and 6 capture downstream penetration fluctuations, as demonstrated in Figure 5-44. Similar to the upstream flapping, mode 5 identifies low frequency behavior at 129 Hz and mode 6 identifies higher frequency behavior at 204 Hz. These modes have reduced frequencies compared to the upstream case. While high frequency flapping may be observed upstream, this is confined to a dense, intact region where small variations in pixel locations provide significant changes in the energy distribution. Further downstream, the liquid atomizes, and the energy is distributed over a greater spatial extent, which is dominated by the bulk flow. This reduces the sensitivity of the downstream distribution, hence reducing the fluctuation frequency.

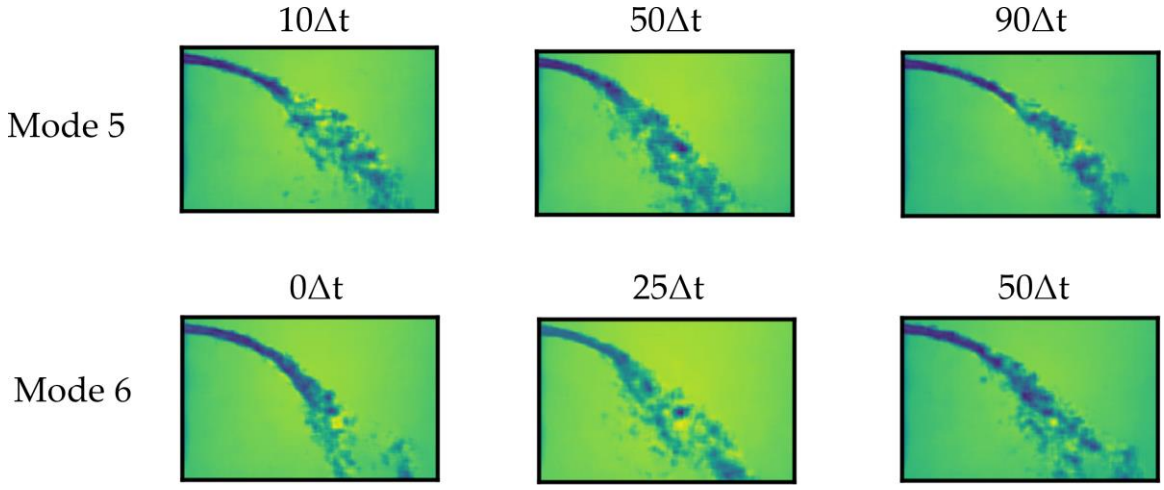

Figure 5-44: CKN modes 5 and 6 capture the fluctuation in jet penetration depth at different frequencies.

Finally, the last two modes both capture the periodic shedding of larger structures, in between which bag breakup has likely occurred. These CKN modes, with similarities to the identified shedding structures from DMD, are shown in Figure 5-45. Recalling the discussion of the DMD modes; they cover a wide range of frequencies and spatial scales, thereby limiting their compactness and utility. The CKN modes provide two distinct frequencies, 855 Hz and 1084 Hz for mode 7 and mode 8, respectively, which fall in the range of the extracted DMD frequencies, and the modal structures and frequency of CKN mode 8 correspond very closely with the DMD results. This appears to indicate that the extracted DMD results are interpretable and accurate, however this agreement is not surprising. DMD identifies a continuum of salient frequencies to capture periodic shedding; the concern is verifying whether all, or some, frequencies are physically meaningful, which is demonstrably challenging. Additionally, DMD modes rarely capture the correct physical spatial structures in the flow, and their relationship with modal frequency is not tied to physics.

The utility of the CKN modes is much more evident than the DMD modes. Again, the CKN modes provide the necessary context via the partial state reconstruction and shows, very clearly, the captured

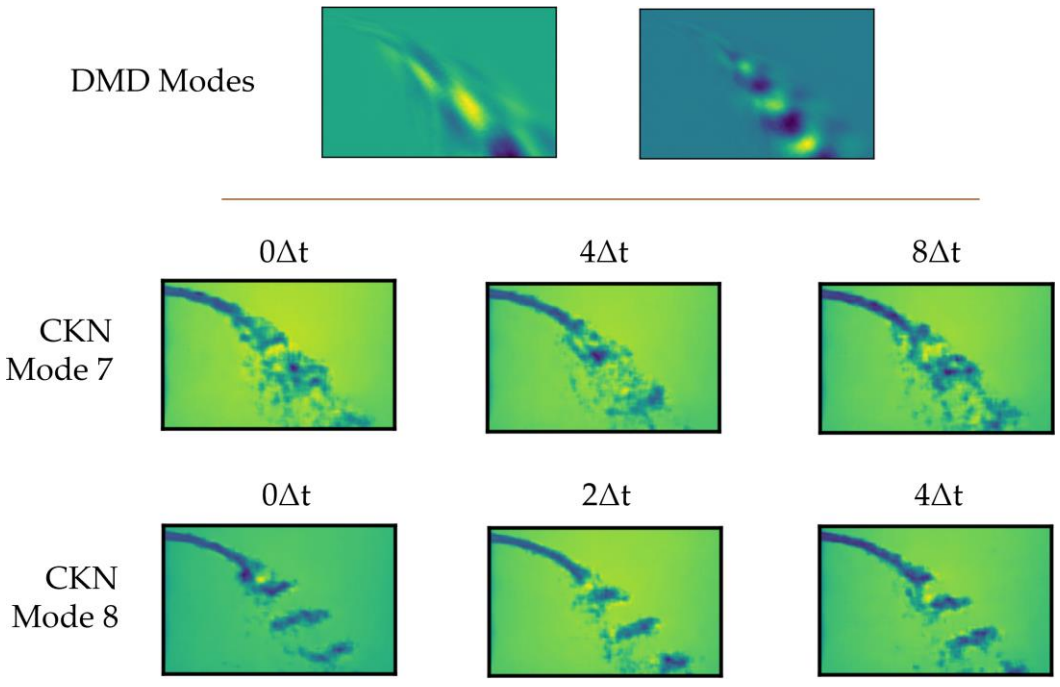

Figure 5-45: CKN modes 7 and 8 identify high frequency periodic shedding from the jet, with example DMD modes also identifying periodic shedding.

spatial behaviors of the jet. This is without the need for further harmonics to improve spatial localization; each mode is spatially and temporally localized. In agreement with the DMD results, the lower frequency does correspond to larger scale shedding of the jet.

Chapter 6

Summary, Conclusions and Recommendations

6.1 Summary

This dissertation has investigated the use of Koopman analysis as a data-driven approach for analyzing and understanding atomization phenomena. Liquid injection and atomization systems are very complex: they have evaded sound analytical understanding and model development due to highly nonlinear multiphase and turbulent interactions. Further, measuring informative state variables from these systems is limited to spatially or temporally small resolutions, or both. Koopman analysis provides a framework to overcome these challenges by leveraging large datasets to automatically learn accurate models and extract physical understanding, and by unlocking high-speed video as a quantitative diagnostic for providing quasi-global, high spatio-temporal resolution datasets.

Proper orthogonal decomposition (POD) and dynamic mode decomposition (DMD) are two prominent decomposition techniques which purport to extract physics from large datasets. DMD assumes a system evolves linearly, and provides an approximation of the Koopman operator; a possibly infinite-dimensional operator which evolves the dynamics of a globally linearized system. This has the potential to enable linear analysis of nonlinear systems. However, POD and DMD are inherently linear tools, which continue to be applied extensively in the fluids community for the analysis of nonlinear flows.

The first part of this work provided a systematic analysis and evaluation of POD and DMD for a number of spatio-temporal systems. Despite their widespread use, extracted results are rarely validated, so their ability to extract physics and reconstruct or predict future states was investigated on a hierarchy of system complexities, so as to validate and verify their results. Even for the simplest systems analyzed, both POD

and DMD have limited utility and do not provide clear interpretations, as the presence of harmonics mars the effectiveness of physical representation of individual modes.

To overcome these concerns, a newly developed deep convolutional Koopman network (CKN) was presented for the first time. The CKN provides a highly compact and interpretable modal decomposition which enables long-term future state prediction for all systems analyzed.

6.2 Conclusions

The main contributions of this work include:

- **POD and DMD are unable to decouple nonlinear processes which are spatially and temporally dependent**

  In practical systems, processes overlap, occlude, and combine resulting in spatially and temporally correlated processes. The simplest example is that of a linear vibrating string made up of sinusoids: the resulting waveform is spatially and temporally dependent on all constituent sinusoids. Despite the underlying linearity, given a nonlinear representation of the system, POD and DMD are unable to extract either the correct spatial or temporal structures. Both techniques appear to require spatial independence, or weak spatial dependence, as a minimum for successful application. In fact, temporal dependence also hinders modal decoupling. At least one spatially distinct and independent process, related to another by a harmonic frequency, will never be uncoupled without algorithmic modification. DMD's orthogonality in time prevents the necessary spatial influence to decouple processes which are spatially *and* temporally distinct. For POD, the PSD can give a strong indication of modal coupling, particularly if multiple localized peaks are identified.

- **Harmonic modes necessarily arise for spatio-temporal data and do not, necessarily, provide additional physical insight**

  An important qualifying term is *additional* physical insight. The only consistency found in this work for the utility of harmonic modes is in capturing motion over discrete temporal scales. Every harmonic mode describes an integer multiple number of positive-negative regions, whose modal frequency can be used to define the temporal scale over which structures will flow from a positive region to a negative region. Beyond this, harmonic modes necessarily arise due to the analyzed spatio-temporal waveforms having discontinuities. While a fundamental mode may capture an underlying system frequency, its modal spatial structures are unlikely to be representative of physical spatial scales. This was demonstrated for a simple particle stream system, where no mode captured the correct spatial scales of the system, as the harmonic scales decrease by a factor of 2 from the fundamental spatial scales.

- **The presented deep convolutional Koopman network provides a highly flexible, compact, and accurate modal decomposition**

  The developed CKN successfully identified all verifiable processes governing the systems analyzed with the correct spatial and temporal behaviors. This has been demonstrated for systems with spatially and temporally dependent processes, and spatially independent processes. Significantly, this has been achieved using an identical model architecture, proving that the intended model flexibility and generality has been achieved for systems with discrete or continuous spectra, or with a redundantly large or restrictively small embedding space. This has led to highly accurate long-term future state prediction, even for practical systems. Although a second architecture was presented, it was not necessary in performing a successful modal

decomposition. However, for the second architecture, only a simple modification was made to layers in the encoder and decoder; no other hyperparameters required modification, which demonstrates the robustness of the overall network design.

- **The presented deep convolutional Koopman network is interpretable**

  Unlike POD and DMD, the CKN modes can provide both spatial and temporal localization. This allows the CKN to identify specific features of a system, such as droplets or ligaments, and evolve them in time according to the captured modal dynamics. This greatly enhances the interpretability of the results, as context is provided with each mode showing the mode's direct contribution to a particular behavior or phenomenon in the system. Single mode reconstructions are therefore sufficient in providing high future state prediction accuracy.

6.3 Recommendations

The development of the deep Koopman network has only recently started, and there is no shortage of directions with which future work can expand into. While key areas that have been identified in this work are addressed here, the reader is invited to take this research into their own path; deep learning provides a tremendous amount of flexibility, with the CKN being no exception.

The following are the most notable recommendations from the investigation:

- **Dimensionality of the embedding space**

  The effects of changing the size of the embedding space are very important to investigate. The size of the embedding space is central to the learning of the network, so the sensitivity of learned

modes or training convergence on this must be evaluated. This is pertinent for modal ordering and robustness, and in understanding the network behaviors for highly complex systems. A jet in crossflow system was analyzed in this work to provide a practical implementation of the CKN. Surprisingly, the CKN predicted only discrete spectra, which raises the question of if the size of the embedding space used has an effect on this.

- **Dilated convolutions**

Of course, there are an endless number of interesting design paths and choices for modifying or optimizing the given CKN, however, the main recommendation is in continued investigation of the effectiveness of dilated convolutions. There were noticeable benefits in transitioning from regular 2D convolutions to dilated convolutions, none of which required fine-tuning. It is generally simpler to identify and represent large-scale features, which can be done with POD and DMD, than for finer details. For atomization systems especially, these finer details are arguably the most important features to analyze, which is precisely why further work into the use of dilated convolutions is of great interest.

- **POD and DMD evaluation**

Finally, despite the limitations of POD and DMD observed in this work, it is highly recommended that further evaluation of both techniques for the purpose of extracting system dynamics or understanding is conducted. At least, it is recommended that future works using either POD or DMD provide sufficient validation for the interpretation attributed to extracted modes.

Bibliography

[1] A. Einstein, *On the Method of Theoretical Physics*, United Kingdom, 1933.

[2] S. Wiggins, Introduction to applied nonlinear dynamical systems and chaos, vol. 2, Springer Science & Business Media, 2003.

[3] D. K. Campbell, Nonlinear Science, Los Alamos Science 15, 1987.

[4] A. H. Lefebvre and V. G. McDonell, Atomization and Sprays, CRC Press, 2017.

[5] G. W. Crabtree and N. S. Lewis, "Solar energy conversion," *Physics today*, vol. 60, p. 37–42, 2007.

[6] T. R. Lakshmanan and X. Han, "Factors underlying transportation CO2 emissions in the USA: a decomposition analysis," *Transportation Research Part D: Transport and Environment*, vol. 2, p. 1–15, 1997.

[7] O. Edenhofer, Climate change 2014: mitigation of climate change, vol. 3, Cambridge University Press, 2015.

[8] T. Boningari and P. G. Smirniotis, "Impact of nitrogen oxides on the environment and human health: Mn-based materials for the NOx abatement," *Current Opinion in Chemical Engineering*, vol. 13, p. 133–141, 2016.

[9] C. T. Bowman, "Control of combustion-generated nitrogen oxide emissions: technology driven by regulation," in *Symposium (International) on Combustion*, 1992.

[10] S. Mazouffre, "Electric propulsion for satellites and spacecraft: established technologies and novel approaches," *Plasma Sources Science and Technology*, vol. 25, p. 033002, 2016.

[11] R. York, "Do alternative energy sources displace fossil fuels?," *Nature Climate Change*, vol. 2, p. 441–443, 2012.

[12] G. M. Faeth, "Evaporation and combustion of sprays," *Progress in Energy and Combustion Science*, vol. 9, p. 1–76, 1983.

[13] M. F. Othman, A. Adam, G. Najafi and R. Mamat, "Green fuel as alternative fuel for diesel engine: A review," *Renewable and Sustainable Energy Reviews*, vol. 80, p. 694–709, 2017.

[14] A. K. Hossain and P. A. Davies, "Pyrolysis liquids and gases as alternative fuels in internal combustion engines–A review," *Renewable and Sustainable Energy Reviews*, vol. 21, p. 165–189, 2013.

[15] S. Candel, "Combustion dynamics and control: Progress and challenges," *Proceedings of the combustion institute*, vol. 29, p. 1–28, 2002.

[16] R. P. Feynman, R. B. Leighton and M. Sands, The Feynman lectures on physics, Vol. I: The new millennium edition: mainly mechanics, radiation, and heat, vol. 1, Basic books, 2011.

[17] T. R. Ohrn, D. W. Senser and A. H. Lefebvre, "Geometrical effects on discharge coefficients for plain-orifice atomizers," *Atomization and Sprays*, vol. 1, 1991.

[18] K. Ramamurthi and K. Nandakumar, "Characteristics of flow through small sharp-edged cylindrical orifices," *Flow measurement and Instrumentation,* vol. 10, p. 133–143, 1999.

[19] F. Payri, V. Bermúdez, R. Payri and F. J. Salvador, "The influence of cavitation on the internal flow and the spray characteristics in diesel injection nozzles," *Fuel,* vol. 83, p. 419–431, 2004.

[20] C. E. Brennen, Fundamentals of Multiphase Flow, Cambridge University Press, 2005.

[21] R. A. Mugele and H. D. Evans, "Droplet size distribution in sprays," *Industrial & Engineering Chemistry,* vol. 43, p. 1317–1324, 1951.

[22] C. T. Crowe, J. D. Schwarzkopf, M. Sommerfeld and Y. Tsuji, Multiphase flows with droplets and particles, CRC press, 2011.

[23] H.-J. Butt, M. Kappl and others, Surface and interfacial forces, Wiley Online Library, 2010.

[24] G. H. Yeoh and J. Tu, "Computational techniques for multiphase flows," in *Computational Techniques for Multiphase Flows,* G. H. Yeoh and J. Tu, Eds., Oxford, Butterworth-Heinemann, 2010, pp. xi-xiii.

[25] P. J. O'Rourke and A. A. Amsden, "The TAB method for numerical calculation of spray droplet breakup," 1987.

[26] L.-P. Hsiang and G. M. Faeth, "Near-limit drop deformation and secondary breakup," *International journal of multiphase flow,* vol. 18, p. 635–652, 1992.

[27] S. B. Leask, V. G. McDonell and S. Samuelsen, "Critical Evaluation of Momentum Flux Ratio Relative to a Liquid Jet in Crossflow," *Atomization and Sprays,* vol. 28, 2018.

[28] L. Rayleigh, "On the instability of jets," *Proceedings of the London mathematical society,* vol. 1, p. 4–13, 1878.

[29] A. Haenlein, "Disintegration of a liquid jet," *NACA Report TN 659,* 1932.

[30] R. D. Reitz, "Atomization and other breakup regimes of a liquid jet," *Ph.D. Thesis,* 1978.

[31] R. D. Reitz and F. V. Bracco, "Mechanism of atomization of a liquid jet," *The physics of Fluids,* vol. 25, p. 1730–1742, 1982.

[32] G. I. Taylor, "Collected Works of GI Taylor, Vol. 3, GK (ed.) Cambridge Univ," 1958.

[33] S. Wang, Y. Huang and Z. L. Liu, "Theoretical analysis of surface waves on a round liquid jet in a gaseous crossflow," *Atomization and Sprays,* vol. 24, 2014.

[34] R. P. Fraser, P. Eisenklam, N. Dombrowski and D. Hasson, "Drop formation from rapidly moving liquid sheets," *AIChE Journal,* vol. 8, p. 672–680, 1962.

[35] D. T. Papageorgiou, "Analytical description of the breakup of liquid jets," *Journal of Fluid Mechanics,* vol. 301, p. 109–132, 1995.

[36] A. J. Chorin, Vorticity and turbulence, vol. 103, Springer Science & Business Media, 2013.

[37] B. J. Olson, J. Larsson, S. K. Lele and A. W. Cook, "Nonlinear effects in the combined Rayleigh-Taylor/Kelvin-Helmholtz instability," *Physics of Fluids,* vol. 23, p. 114107, 2011.

[38] J. F. Keffer and W. D. Baines, "The round turbulent jet in a cross-wind," *Journal of Fluid Mechanics,* vol. 15, p. 481–496, 1963.

[39] Y. Kamotani and I. Greber, "Experiments on a turbulent jet in a cross flow," *AIAA journal,* vol. 10, p. 1425–1429, 1972.

[40] J. Andreopoulos and W. Rodi, "Experimental investigation of jets in a crossflow," *Journal of Fluid Mechanics,* vol. 138, p. 93–127, 1984.

[41] R. M. Kelso, T. T. Lim and A. E. Perry, "An experimental study of round jets in cross-flow," *Journal of fluid mechanics,* vol. 306, p. 111–144, 1996.

[42] M. Adelberg, "Breakup rate and penetration of a liquid jet in a gas stream.," *AIAA Journal,* vol. 5, p. 1408–1415, 1967.

[43] J. A. Schetz and A. Padhye, "Penetration and breakup of liquids in subsonic airstreams," *AIAA Journal,* vol. 15, p. 1385–1390, 1977.

[44] P.-K. Wu, K. A. Kirkendall, R. P. Fuller and A. S. Nejad, "Breakup processes of liquid jets in subsonic crossflows," *Journal of Propulsion and Power,* vol. 13, p. 64–73, 1997.

[45] M. Wang, M. Broumand and M. Birouk, "Liquid jet trajectory in a subsonic gaseous cross-flow: an analysis of published correlations," *Atomization and Sprays,* vol. 26, 2016.

[46] M. Broumand and M. Birouk, "Liquid jet in a subsonic gaseous crossflow: Recent progress and remaining challenges," *Progress in Energy and Combustion Science,* vol. 57, p. 1–29, 2016.

[47] O. Desjardins, H. Pitsch and others, "Detailed numerical investigation of turbulent atomization of liquid jets," *Atomization and Sprays,* vol. 20, p. 311, 2010.

[48] L. Wang, P. P. Ho, C. Liu, G. Zhang and R. R. Alfano, "Ballistic 2-D imaging through scattering walls using an ultrafast optical Kerr gate," *Science,* vol. 253, p. 769–771, 1991.

[49] M. Paciaroni and M. Linne, "Single-shot, two-dimensional ballistic imaging through scattering media," *Applied optics,* vol. 43, p. 5100–5109, 2004.

[50] W. D. Bachalo, "Method for measuring the size and velocity of spheres by dual-beam light-scatter interferometry," *Applied optics,* vol. 19, p. 363–370, 1980.

[51] L. G. Dodge, "Comparison of performance of drop-sizing instruments," *Applied Optics,* vol. 26, p. 1328–1341, 1987.

[52] L. G. Dodge, D. J. Rhodes and R. D. Reitz, "Drop-size measurement techniques for sprays: comparison of Malvern laser-diffraction and Aerometrics phase/Doppler," *Applied optics,* vol. 26, p. 2144–2154, 1987.

[53] V. G. McDonell and S. Samuelsen, "Sensitivity assessment of a phase-Doppler interferometer to user-controlled settings," in *Liquid Particle Size Measurement Techniques: 2nd Volume*, ASTM International, 1990.

[54] R. Payri, L. Araneo, J. Shakal and V. Soare, "Phase doppler measurements: system set-up optimization for characterization of a diesel nozzle," *Journal of Mechanical Science and Technology,* vol. 22, p. 1620–1632, 2008.

[55] W. Olsen, D. Takeuchi and K. Adams, "Experimental comparison of icing cloud instruments," in *21st Aerospace Sciences Meeting*, 1983.

[56] H. C. Simmons and C. F. Harding, "Some effects of using water as a test fluid in fuel nozzle spray analysis," *Journal of Engineering for Power,* vol. 103, p. 118–123, 1981.

[57] D. C. Hammond Jr, "Measurement of the Droplet-Size Distributions Produced by the ASTM 'Round Robin'Nozzle," *General Motors Research Laboratories Report GMR-3668,* 1981.

[58] T. A. Jackson and G. S. Samuelsen, "Performance comparison of two interferometric droplet sizing techniques," in *Particle sizing and spray analysis*, 1985.

[59] T. A. Jackson and G. S. Samuelsen, "Spatially resolved droplet size measurements," *Journal of engineering for gas turbines and power,* vol. 108, p. 196–203, 1986.

[60] V. G. McDonnell, G. S. Samuelsen, M. R. Wang, C. H. Hong and W. H. Lai, "Interlaboratory comparison of phase Doppler measurements in a research simplex atomizer spray," *Journal of Propulsion and Power,* vol. 10, p. 402–409, 1994.

[61] M. Bardi, R. Payri, L. M. C. Malbec, G. Bruneaux, L. M. Pickett, J. Manin, T. Bazyn and C. L. Genzale, "Engine combustion network: comparison of spray development, vaporization, and combustion in different combustion vessels," *Atomization and Sprays,* vol. 22, 2012.

[62] M. Meijer, B. Somers, J. Johnson, J. Naber, S.-Y. Lee, L. M. C. Malbec, G. Bruneaux, L. M. Pickett, M. Bardi, R. Payri and others, "Engine Combustion Network (ECN): Characterization and comparison of boundary conditions for different combustion vessels," *Atomization and Sprays,* vol. 22, 2012.

[63] A. L. Kastengren, F. Z. Tilocco, C. F. Powell, J. Manin, L. M. Pickett, R. Payri and T. Bazyn, "Engine combustion network (ECN): measurements of nozzle geometry and hydraulic behavior," *Atomization and Sprays,* vol. 22, 2012.

[64] A. L. Yarin, "Drop impact dynamics: splashing, spreading, receding, bouncing…," *Annu. Rev. Fluid Mech.,* vol. 38, p. 159–192, 2006.

[65] E. Villermaux, "Fragmentation," *Annu. Rev. Fluid Mech.,* vol. 39, p. 419–446, 2007.

[66] F. J. Salvador, S. Ruiz, M. Crialesi-Esposito and I. Blanquer, "Analysis on the effects of turbulent inflow conditions on spray primary atomization in the near-field by direct numerical simulation," *International Journal of Multiphase Flow,* vol. 102, p. 49–63, 2018.

[67] O. Desjardins, J. McCaslin, M. Owkes and P. Brady, "Direct numerical and large-eddy simulation of primary atomization in complex geometries," *Atomization and Sprays,* vol. 23, 2013.

[68] R. Canu, S. Puggelli, M. Essadki, B. Duret, T. Menard, M. Massot, J. Reveillon and F. X. Demoulin, "Where does the droplet size distribution come from?," *International Journal of Multiphase Flow,* vol. 107, p. 230–245, 2018.

[69] J. Hasslberger, S. Ketterl, M. Klein and N. Chakraborty, "Flow topologies in primary atomization of liquid jets: A direct numerical simulation analysis," *Journal of Fluid Mechanics,* vol. 859, p. 819–838, 2019.

[70] T. Bohr, M. H. Jensen, G. Paladin and A. Vulpiani, Dynamical systems approach to turbulence, Cambridge University Press, 2005.

[71] S. L. Brunton and J. N. Kutz, Data-driven science and engineering: Machine learning, dynamical systems, and control, Cambridge University Press, 2019.

[72] P. Bradshaw, "Turbulence: the chief outstanding difficulty of our subject," *Experiments in fluids,* vol. 16, p. 203–216, 1994.

[73] J. M. Benítez, J. L. Castro and I. Requena, "Are artificial neural networks black boxes?," *IEEE Transactions on neural networks,* vol. 8, p. 1156–1164, 1997.

[74] J. D. Olden and D. A. Jackson, "Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks," *Ecological modelling,* vol. 154, p. 135–150, 2002.

[75] R. Shwartz-Ziv and N. Tishby, "Opening the black box of deep neural networks via information," *arXiv preprint arXiv:1703.00810,* 2017.

[76] C. Zhang, S. Bengio, M. Hardt, B. Recht and O. Vinyals, "Understanding deep learning requires rethinking generalization," *arXiv preprint arXiv:1611.03530,* 2016.

[77] J.-C. Hua, G. H. Gunaratne, D. G. Talley, J. R. Gord and S. Roy, "Dynamic-mode decomposition based analysis of shear coaxial jets with and without transverse acoustic driving," *Journal of Fluid Mechanics,* vol. 790, p. 5–32, 2016.

[78] F. Magionesi, G. Dubbioso, R. Muscari and A. Di Mascio, "Modal analysis of the wake past a marine propeller," *Journal of Fluid Mechanics,* vol. 855, p. 469–502, 2018.

[79] A. J. Torregrosa, A. Broatch, J. García-Tíscar and J. Gomez-Soriano, "Modal decomposition of the unsteady flow field in compression-ignited combustion chambers," *Combustion and Flame,* vol. 188, p. 469–482, 2018.

[80] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of fluid mechanics,* vol. 641, p. 115–127, 2009.

[81] L. Van Der Maaten, E. Postma and J. Van den Herik, "Dimensionality reduction: a comparative," *J Mach Learn Res,* vol. 10, p. 13, 2009.

[82] F. J. Herrmann, M. P. Friedlander and O. Yilmaz, "Fighting the curse of dimensionality: Compressive sensing in exploration seismology," *IEEE Signal Processing Magazine,* vol. 29, p. 88–100, 2012.

[83] J. R. Parker, Algorithms for image processing and computer vision, John Wiley & Sons, 2010.

[84] U. Schnars, C. Falldorf, J. Watson and W. Jüptner, "Digital holography," in *Digital Holography and Wavefront Sensing*, Springer, 2015, p. 39–68.

[85] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dynamics,* vol. 41, p. 309–325, 2005.

[86] M. O. Williams, I. G. Kevrekidis and C. W. Rowley, "A data–driven approximation of the koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science,* vol. 25, p. 1307–1346, 2015.

[87] M. Korda and I. Mezić, "On convergence of extended dynamic mode decomposition to the Koopman operator," *Journal of Nonlinear Science,* vol. 28, p. 687–710, 2018.

[88] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of fluid mechanics,* vol. 656, p. 5–28, 2010.

[89] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," *Journal of Computational Dynamics,* vol. 1, p. 391–421, 2014.

[90] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proceedings of the national academy of sciences of the united states of america,* vol. 17, p. 315, 1931.

[91] K. Pearson, "On lines and planes of closest fit to systems of point in space," *Philosophical Magazine,* vol. 2, p. 559–572, 1901.

[92] K. Karhunen, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, vol. 37, Universitat Helsinki, 1947.

[93] M. Loeve, "Functions aleatoires du second ordre," *Processus stochastique et mouvement Brownien,* p. 366–420, 1948.

[94] H. Hotelling, "Analysis of a complex of statistical variables into principal components.," *Journal of educational psychology,* vol. 24, p. 417, 1933.

[95] E. N. Lorenz, "Empirical orthogonal functions and statistical weather prediction," 1956.

[96] J. L. Lumley, "The structure of inhomogeneous turbulent flows," *Atmospheric turbulence and radio wave propagation,* 1967.

[97] H. P. Bakewell Jr and J. L. Lumley, "Viscous sublayer and adjacent wall region in turbulent pipe flow," *The Physics of Fluids,* vol. 10, p. 1880–1889, 1967.

[98] L. Sirovich, "Turbulence and the dynamics of coherent structures. I. Coherent structures," *Quarterly of applied mathematics,* vol. 45, p. 561–571, 1987.

[99] G. Berkooz, P. Holmes and J. L. Lumley, "The proper orthogonal decomposition in the analysis of turbulent flows," *Annual review of fluid mechanics,* vol. 25, p. 539–575, 1993.

[100] M. Kirby, J. P. Boris and L. Sirovich, "A proper orthogonal decomposition of a simulated supersonic shear layer," *International journal for numerical methods in fluids,* vol. 10, p. 411–428, 1990.

[101] B. Patte-Rouland, G. Lalizel, J. Moreau and E. Rouland, "Flow analysis of an annular jet by particle image velocimetry and proper orthogonal decomposition," *Measurement Science and Technology,* vol. 12, p. 1404, 2001.

[102] K. E. Meyer, J. M. Pedersen and O. Özcan, "A turbulent jet in crossflow analysed with proper orthogonal decomposition," *Journal of Fluid Mechanics,* vol. 583, p. 199–227, 2007.

[103] M. Arienti and M. C. Soteriou, "Time-resolved proper orthogonal decomposition of liquid jet dynamics," *Physics of Fluids,* vol. 21, p. 112104, 2009.

[104] G. Charalampous and Y. Hardalupas, "Application of proper orthogonal decomposition to the morphological analysis of confined co-axial jets of immiscible liquids with comparable densities," *Physics of Fluids,* vol. 26, p. 113301, 2014.

[105] S. Gadiraju, S. Park, D. Gomez-Ramirez, S. V. Ekkad, K. T. Lowe, H.-K. Moon, Y. Kim and R. Srinivasan, "Application of Proper Orthogonal Decomposition to High Speed Imaging for the Study of Combustion Oscillations," in *ASME Turbo Expo 2017: Turbomachinery Technical Conference and Exposition*, 2017.

[106] M. G. De Giorgi, D. Fontanarosa and A. Ficarella, "Characterization of cavitating flow regimes in an internal sharp-edged orifice by means of Proper Orthogonal Decomposition," *Experimental Thermal and Fluid Science,* vol. 93, p. 242–256, 2018.

[107] S. Taghizadeh and D. Jarrahbashi, "Proper orthogonal decomposition analysis of turbulent cryogenic liquid jet injection under transcritical and supercritical conditions," *Atomization and Sprays,* vol. 28, 2018.

[108] V. Narayanan, M. D. Lightfoot, S. A. Schumaker, S. A. Danczyk and B. Eilers, "Use of proper orthogonal decomposition towards time-resolved image analysis of sprays," 2011.

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

[110] J. Kou and W. Zhang, "Dynamic mode decomposition with exogenous input for data-driven modeling of unsteady flows," *Physics of fluids,* vol. 31, p. 057106, 2019.

[111] P. J. Schmid, L. Li, M. P. Juniper and O. Pust, "Applications of the dynamic mode decomposition," *Theoretical and Computational Fluid Dynamics,* vol. 25, p. 249–259, 2011.

[112] K. K. Chen, J. H. Tu and C. W. Rowley, "Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses," *Journal of nonlinear science,* vol. 22, p. 887–915, 2012.

[113] A. Chatterjee, "An introduction to the proper orthogonal decomposition," *Current science,* p. 808–817, 2000.

[114] J.-C. Hua, F. Noorian, D. Moss, P. H. W. Leong and G. H. Gunaratne, "High-dimensional time series prediction using kernel-based Koopman mode regression," *Nonlinear Dynamics,* vol. 90, p. 1785–1806, 2017.

[115] J. Mann and J. N. Kutz, "Dynamic mode decomposition for financial trading strategies," *Quantitative Finance,* vol. 16, p. 1643–1655, 2016.

[116] A. Bao, E. Gildin, A. Narasingam and J. S. Kwon, "Data-Driven Model Reduction for Coupled Flow and Geomechanics Based on DMD Methods," *Fluids,* vol. 4, p. 138, 2019.

[117] J. L. Proctor and P. A. Eckhoff, "Discovering dynamic patterns from infectious disease data using dynamic mode decomposition," *International health,* vol. 7, p. 139–145, 2015.

[118] A. T. Mohan, M. R. Visbal and D. V. Gaitonde, "Model reduction and analysis of deep dynamic stall on a plunging airfoil using dynamic mode decomposition," in *53rd AIAA Aerospace Sciences Meeting*, 2015.

[119] W. Krolick and M. Owkes, "Primary Atomization Instability Extraction Using Dynamic Mode Decomposition," *Atomization and Sprays,* vol. 28, 2018.

[120] H. Chen, D. Hung, M. Xu and J. Zhong, "Analyzing the cycle-to-cycle variations of pulsing spray characteristics by means of the proper orthogonal decomposition," *Atomization and Sprays,* vol. 23, 2013.

[121] F. Vashahi, S. Rezaei and J. Lee, "Large Eddy Simulation and Dynamic Mode Decomposition of Internal Flow Structure of Pressure Swirl Atomizer," in *21st Australasian Fluid Mechanics Conference,* 2018.

[122] S. B. Leask, A. K. Li, V. G. McDonell and S. Samuelsen, "Preliminary Development of a Measurement Reference Using a Research Simplex Atomizer," *Journal of Fluids Engineering,* vol. 141, 2019.

[123] D. Duke, D. Honnery and J. Soria, "Experimental investigation of nonlinear instabilities in annular liquid sheets," *Journal of Fluid Mechanics,* vol. 691, p. 594–604, 2012.

[124] S. B. Leask, V. G. McDonell and S. Samuelsen, "Emulsion Jet in Crossflow Atomization Characteristics and Dynamics," *Journal of Engineering for Gas Turbines and Power,* vol. 141, p. 041025, 2019.

[125] C. Huang, W. E. Anderson, M. E. Harvazinski and V. Sankaran, "Analysis of self-excited combustion instabilities using decomposition techniques," *AIAA Journal*, p. 2791–2807, 2016.

[126] S. Tirunagari, V. Vuorinen, O. Kaario and M. Larmi, "Analysis of proper orthogonal decomposition and dynamic mode decomposition on les of subsonic jets," *CSI Journal of Computing*, vol. 1, p. 20–26, 2012.

[127] T. W. Muld, G. Efraimsson and D. S. Henningson, "Flow structures around a high-speed train extracted using proper orthogonal decomposition and dynamic mode decomposition," *Computers & Fluids*, vol. 57, p. 87–97, 2012.

[128] S. Roy, J.-C. Hua, W. Barnhill, G. H. Gunaratne and J. R. Gord, "Deconvolution of reacting-flow dynamics using proper orthogonal and dynamic mode decompositions," *Physical review E*, vol. 91, p. 013001, 2015.

[129] G. Kerschen and J.-C. Golinval, "Physical interpretation of the proper orthogonal modes using the singular value decomposition," *Journal of Sound and vibration*, vol. 249, p. 849–865, 2002.

[130] B. F. Feeny and Y. Liang, "Interpreting proper orthogonal modes of randomly excited vibration systems," *Journal of Sound and Vibration*, vol. 265, p. 953–966, 2003.

[131] H. Chen, D. L. Reuss and V. Sick, "On the use and interpretation of proper orthogonal decomposition of in-cylinder engine flows," *Measurement Science and Technology*, vol. 23, p. 085302, 2012.

[132] H. Chen, D. L. Reuss, D. L. S. Hung and V. Sick, "A practical guide for using proper orthogonal decomposition in engine research," *International Journal of Engine Research*, vol. 14, p. 307–319, 2013.

[133] S. T. M. Dawson, M. S. Hemati, M. O. Williams and C. W. Rowley, "Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition," *Experiments in Fluids*, vol. 57, p. 42, 2016.

[134] A. Towne, O. T. Schmidt and T. Colonius, "Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis," *Journal of Fluid Mechanics*, vol. 847, p. 821–867, 2018.

[135] B. R. Noack, W. Stankiewicz, M. Morzyński and P. J. Schmid, "Recursive dynamic mode decomposition of transient and post-transient wake flows," *Journal of Fluid Mechanics*, vol. 809, p. 843–872, 2016.

[136] J. N. Kutz, X. Fu and S. L. Brunton, "Multiresolution dynamic mode decomposition," *SIAM Journal on Applied Dynamical Systems*, vol. 15, p. 713–735, 2016.

[137] M. R. Jovanović, P. J. Schmid and J. W. Nichols, "Sparsity-promoting dynamic mode decomposition," *Physics of Fluids*, vol. 26, p. 024103, 2014.

[138] Q. Li, F. Dietrich, E. M. Bollt and I. G. Kevrekidis, "Extended dynamic mode decomposition with dictionary learning: A data-driven adaptive spectral decomposition of the Koopman operator," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 27, p. 103111, 2017.

[139] C. W. Rowley and S. T. M. Dawson, "Model reduction for flow analysis and control," *Annual Review of Fluid Mechanics*, vol. 49, p. 387–417, 2017.

[140] K. Taira, S. L. Brunton, S. T. M. Dawson, C. W. Rowley, T. Colonius, B. J. McKeon, O. T. Schmidt, S. Gordeyev, V. Theofilis and L. S. Ukeiley, "Modal analysis of fluid flows: An overview," *Aiaa Journal*, p. 4013–4041, 2017.

[141] C. W. Rowley, "Model reduction for fluids, using balanced proper orthogonal decomposition," *International Journal of Bifurcation and Chaos*, vol. 15, p. 997–1013, 2005.

[142] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE transactions on automatic control*, vol. 26, p. 17–32, 1981.

[143] M. Sieber, C. O. Paschereit and K. Oberleithner, "Spectral proper orthogonal decomposition," *Journal of Fluid Mechanics*, vol. 792, p. 798–828, 2016.

[144] A. Alessandri, P. Bagnerini, M. Gaggero, D. Lengani and D. Simoni, "Dynamic mode decomposition for the inspection of three-regime separated transitional boundary layers using a least squares method," *Physics of Fluids*, vol. 31, p. 044103, 2019.

[145] S. Le Clainche and J. M. Vega, "Higher order dynamic mode decomposition to identify and extrapolate flow patterns," *Physics of Fluids*, vol. 29, p. 084102, 2017.

[146] J. Garicano-Mena, B. Li, E. Ferrer and E. Valero, "A composite dynamic mode decomposition analysis of turbulent channel flows," *Physics of Fluids*, vol. 31, p. 115102, 2019.

[147] S. L. Brunton, J. L. Proctor and J. N. Kutz, "Compressive sampling and dynamic mode decomposition," *arXiv preprint arXiv:1312.5186*, 2013.

[148] T. Askham and J. N. Kutz, "Variable projection methods for an optimized dynamic mode decomposition," *SIAM Journal on Applied Dynamical Systems*, vol. 17, p. 380–416, 2018.

[149] L. Hertel, J. Collado, P. Sadowski, J. Ott and P. Baldi, "Sherpa: Robust hyperparameter optimization for machine learning," *SoftwareX*, vol. 12, p. 100591, 2020.

[150] X. He, K. Zhao and X. Chu, "AutoML: A Survey of the State-of-the-Art," *Knowledge-Based Systems*, vol. 212, p. 106622, 2019.

[151] D. Yarowsky, "Unsupervised word sense disambiguation rivaling supervised methods," in *33rd annual meeting of the association for computational linguistics*, 1995.

[152] F. Scarselli and A. C. Tsoi, "Universal approximation using feedforward neural networks: A survey of some existing methods, and some new results," *Neural networks*, vol. 11, p. 15–37, 1998.

[153] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, p. 115–133, 1943.

[154] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.," *Psychological review*, vol. 65, p. 386, 1958.

[155] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors," *nature*, vol. 323, p. 533–536, 1986.

[156] I. Aizenberg, N. Aizenberg, C. Butakov and E. Farberov, "Image recognition on the neural network based on multi-valued neurons," in *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, 2000.

[157] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012.

[158] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis,* vol. 42, p. 60–88, 2017.

[159] P. Baldi, K. Bauer, C. Eng, P. Sadowski and D. Whiteson, "Jet substructure classification in high-energy physics with deep neural networks," *Physical Review D,* vol. 93, p. 094034, 2016.

[160] P. Baldi, J. Bian, L. Hertel and L. Li, "Improved energy reconstruction in NOvA with regression convolutional neural networks," *Physical Review D,* vol. 99, p. 012011, 2019.

[161] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski and D. Whiteson, "Parameterized neural networks for high-energy physics," *The European Physical Journal C,* vol. 76, p. 235, 2016.

[162] P. Baldi, "Deep Learning in Biomedical Data Science," *Annual Review of Biomedical Data Science,* vol. 1, p. 181–205, 2018.

[163] J. B. Heaton, N. G. Polson and J. H. Witte, "Deep learning for finance: deep portfolios," *Applied Stochastic Models in Business and Industry,* vol. 33, p. 3–12, 2017.

[164] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *ICML*, 2010.

[165] P. Werbos, "Beyond regression:" new tools for prediction and analysis in the behavioral sciences," *Ph. D. dissertation, Harvard University*, 1974.

[166] Y. LeCun, D. Touresky, G. Hinton and T. Sejnowski, "A theoretical framework for back-propagation," in *Proceedings of the 1988 connectionist models summer school*, 1988.

[167] J. Duchi, E. Hazan and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization.," *Journal of machine learning research,* vol. 12, 2011.

[168] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980,* 2014.

[169] D. Maturana and S. Scherer, "Voxnet: A 3d convolutional neural network for real-time object recognition," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015.

[170] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research,* vol. 15, p. 1929–1958, 2014.

[171] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980,* 2014.

[172] S. Nah, T. Hyun Kim and K. Mu Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[173] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee and A. Y. Ng, "Multimodal deep learning," in *Proceedings of the 28th international conference on machine learning (ICML-11)*, 2011.

[174] K. He, X. Zhang, S. Ren and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence,* vol. 37, p. 1904–1916, 2015.

[175] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014.

[176] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015.

[177] S. Ren, K. He, R. Girshick and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *arXiv preprint arXiv:1506.01497*, 2015.

[178] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science*, vol. 313, p. 504–507, 2006.

[179] S. J. Wetzel, "Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders," *Physical Review E,* vol. 96, p. 022140, 2017.

[180] G. W. Cottrell, "Image compression by back-propagation: An example of extensional programming," *Advances in cognitive science,* vol. 3, p. 208–240, 1988.

[181] Q. V. Le, "Building high-level features using large scale unsupervised learning," in *2013 IEEE international conference on acoustics, speech and signal processing*, 2013.

[182] P. Baldi and K. Hornik, "Neural networks and principal component analysis: Learning from examples without local minima," *Neural networks,* vol. 2, p. 53–58, 1989.

[183] P. Baldi and P. J. Sadowski, "Understanding dropout," *Advances in neural information processing systems,* vol. 26, p. 2814–2822, 2013.

[184] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167,* 2015.

[185] C. Peng, T. Xiao, Z. Li, Y. Jiang, X. Zhang, K. Jia, G. Yu and J. Sun, "MegDet: A Large Mini-Batch Object Detector," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.

[186] Y. Wu and K. He, "Group normalization," in *Proceedings of the European conference on computer vision (ECCV)*, 2018.

[187] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.

[188] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.

[189] F. Yu, V. Koltun and T. Funkhouser, "Dilated residual networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.

[190] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence,* vol. 40, p. 834–848, 2017.

[191] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," *arXiv preprint arXiv:1511.07122*, 2015.

[192] Z. Wang and S. Ji, "Smoothed dilated convolutions for improved dense prediction," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018.

[193] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever and P. Abbeel, "Infogan: Interpretable representation learning by information maximizing generative adversarial nets," *arXiv preprint arXiv:1606.03657*, 2016.

[194] M. Raissi, P. Perdikaris and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, p. 686–707, 2019.

[195] R. T. Q. Chen, Y. Rubanova, J. Bettencourt and D. Duvenaud, "Neural ordinary differential equations," *arXiv preprint arXiv:1806.07366*, 2018.

[196] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan and A. Edelman, "Universal differential equations for scientific machine learning," *arXiv preprint arXiv:2001.04385*, 2020.

[197] N. M. Mangan, J. N. Kutz, S. L. Brunton and J. L. Proctor, "Model selection for dynamical systems via sparse regression and information criteria," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 473, p. 20170009, 2017.

[198] R. Wang, K. Kashinath, M. Mustafa, A. Albert and R. Yu, "Towards physics-informed deep learning for turbulent flow prediction," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020.

[199] A. T. Mohan, D. Tretiak, M. Chertkov and D. Livescu, "Spatio-temporal deep learning models of 3D turbulence with physics informed diagnostics," *Journal of Turbulence*, vol. 21, p. 484–524, 2020.

[200] Y.-D. Ko and H. Shang, "A Neural Network-Based Soft Sensor for Particle Size Distribution using Image Analysis," *Powder Technology*, vol. 212(2), p. 359–366, 2011.

[201] H. Yu, J. Fu, L. Dang, Y. Cheong, H. Tan and H. Wei, "Prediction of the Particle Size Distribution Parameters in a High Shear Granulation Process Using a Key Parameter Definition Combined Artificial Neural Network Model," *Ind. Eng. Chem. Res.*, vol. 54(43), no. 43, p. 10825–10834, 2015.

[202] M. Bombrun, V. Barra and A. Harris, "Algorithm for Particle Detection and Parameterization in High-Frame-Rate Thermal Video," *Journal of Applied Remote Sensing*, vol. 8(1), no. 43, p. 083549, 2014.

[203] T. H. Teng, J. C. Patra and E.-L. Ang, "Identification and Tracking of Dim Moving Targets in FLIR using Artificial Neural Networks," *Seventh International Symposium on Signal Processing and Its Applications*, vol. 2, no. 43, p. 223–226, 2003.

[204] I. Arsie, C. Pianese and M. Sorrentino, "Development of recurrent neural networks for virtual sensing of NOx emissions in internal combustion engines," *SAE International Journal of Fuels and Lubricants*, vol. 2, p. 354–361, 2010.

[205] S. Sarkar, K. G. Lore, S. Sarkar, V. Ramanan, S. R. Chakravarthy, S. Phoha and A. Ray, "Early detection of combustion instability from hi-speed flame images via deep learning and

symbolic time series analysis," in *Annual Conf. of the Prognostics and Health Management*, 2015.

[206] A. Akintayo, K. G. Lore, S. Sarkar and S. Sarkar, "Prognostics of combustion instabilities from hi-speed flame video using a deep convolutional selective autoencoder," *International Journal of Prognostics and Health Management,* vol. 7, p. 1–14, 2016.

[207] D. K. Jha, A. Srivastav and A. Ray, "Temporal learning in video data using deep learning and Gaussian processes," in *Workshop on Machine Learning for Prognostics and Health Managament at 2016 KDD, San Francisco, CA*, 2016.

[208] S. B. Leask, V. G. McDonell and S. Samuelsen, "Neural Network Prediction of Boundary Layer Flashback," *Journal of Engineering for Gas Turbines and Power,* vol. 143, p. 054501, 2021.

[209] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2014.

[210] S. Ji, W. Xu, M. Yang and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE transactions on pattern analysis and machine intelligence,* vol. 35, p. 221–231, 2012.

[211] Z. Qiu, T. Yao and T. Mei, "Learning spatio-temporal representation with pseudo-3d residual networks," in *proceedings of the IEEE International Conference on Computer Vision*, 2017.

[212] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun and M. Paluri, "A closer look at spatiotemporal convolutions for action recognition," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2018.

[213] D. Tran, L. Bourdev, R. Fergus, L. Torresani and M. Paluri, "Learning spatiotemporal features with 3d convolutional networks," in *Proceedings of the IEEE international conference on computer vision*, 2015.

[214] S. D. Jain, B. Xiong and K. Grauman, "Fusionseg: Learning to combine motion and appearance for fully automatic segmentation of generic objects in videos," in *2017 IEEE conference on computer vision and pattern recognition (CVPR)*, 2017.

[215] J. Carreira and A. Zisserman, "Quo vadis, action recognition? a new model and the kinetics dataset," in *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[216] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," *Journal of Computational Dynamics,* vol. 1, p. 391–421, 2014.

[217] J. E. Higham, W. Brevis and C. J. Keylock, "Implications of the selection of a particular modal decomposition technique for the analysis of shallow flows," *Journal of Hydraulic Research,* vol. 56, p. 796–805, 2018.

[218] E. Qian, B. Kramer, B. Peherstorfer and K. Willcox, "Lift & Learn: Physics-informed machine learning for large-scale nonlinear dynamical systems," *Physica D: Nonlinear Phenomena*, vol. 406, p. 132401, 2020.

[219] M. Birouk and N. Lekic, "Liquid jet breakup in quiescent atmosphere: A review," *Atomization and Sprays*, vol. 19, 2009.

[220] N. Ashgriz and A. L. Yarin, "Capillary instability of free liquid jets," in *Handbook of atomization and sprays*, Springer, 2011, p. 3–53.

[221] T. Si, F. Li, X.-Y. Yin and X.-Z. Yin, "Spatial instability of coflowing liquid-gas jets in capillary flow focusing," *Physics of Fluids*, vol. 22, p. 112105, 2010.

[222] E. Yeung, S. Kundu and N. Hodas, "Learning deep neural network representations for Koopman operators of nonlinear dynamical systems," in *2019 American Control Conference (ACC)*, 2019.

[223] B. Lusch, J. N. Kutz and S. L. Brunton, "Deep learning for universal linear embeddings of nonlinear dynamics," *Nature communications*, vol. 9, p. 4950, 2018.

[224] Y. Han, W. Hao and U. Vaidya, "Deep learning of koopman representation for control," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020.

[225] S. B. Leask, V. G. McDonell and S. Samuelsen, "Modal extraction of spatiotemporal atomization data using a deep convolutional Koopman network," *Physics of Fluids*, vol. 33, p. 033323, 2021.

[226] L.-C. Chen, G. Papandreou, F. Schroff and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017.

[227] J. Kou and W. Zhang, "An improved criterion to select dominant modes from dynamic mode decomposition," *European Journal of Mechanics-B/Fluids*, vol. 62, p. 109–129, 2017.

[228] S. B. Leask, V. G. McDonell and S. Samuelsen, "On the Use of Dynamic Mode Decomposition for Liquid Injection," *Atomization and Sprays*, vol. 29, 2019.

[229] R. J. Donnelly and W. Glaberson, "Experiments on the capillary instability of a liquid jet," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 290, p. 547–556, 1966.

[230] M.-C. Yuen, "Non-linear capillary instability of a liquid jet," *Journal of Fluid Mechanics*, vol. 33, p. 151–163, 1968.

[231] K. C. Chaudhary and L. G. Redekopp, "The nonlinear capillary instability of a liquid jet. Part 1. Theory," *Journal of Fluid Mechanics*, vol. 96, p. 257–274, 1980.

[232] K. C. Chaudhary and T. Maxworthy, "The nonlinear capillary instability of a liquid jet. Part 2. Experiments on jet behaviour before droplet formation," *Journal of fluid mechanics*, vol. 96, p. 275–286, 1980.

[233] D. Ayyappan and A. Vaidyanathan, "Study on instability of circular liquid jets at subcritical to supercritical conditions using dynamic mode decomposition," *Physics of Fluids*, vol. 32, p. 014107, 2020.

[234] R. Vernet, L. Thomas and L. David, "Analysis and reconstruction of a pulsed jet in crossflow by multi-plane snapshot POD," *Experiments in Fluids*, vol. 47, p. 707–720, 2009.

[235] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*, 2014.

[236] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin and others, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.

[237] D. Masters and C. Luschi, "Revisiting small batch training for deep neural networks," *arXiv preprint arXiv:1804.07612*, 2018.

[238] L. Cavigelli, P. Hager and L. Benini, "CAS-CNN: A deep convolutional neural network for image compression artifact suppression," in *2017 International Joint Conference on Neural Networks (IJCNN)*, 2017.

[239] C. Zhang, S. Bengio, M. Hardt, B. Recht and O. Vinyals, "Understanding deep learning requires rethinking generalization," *arXiv preprint arXiv:1611.03530*, 2016.