DYNAMIC MODE DECOMPOSITION WITH APPLICATION TO OPTIMAL CONTROL

BY

MICHAEL J. BANKS

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Aerospace Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Advisor:

Associate Professor Daniel J. Bodony

ABSTRACT

High-dimensional fluid dynamics systems are central to a variety of modern engineering challenges. Direct numerical simulation of these processes is possible, but the high computational cost of these simulations is always an important consideration. Optimal control of these high-dimensional fluid dynamics problems is especially cost prohibitive. In order to keep computational cost within a reasonable level, we explore the potential of merging data-driven model reduction with optimal control. In particular, we demonstrate the application of Dynamic Mode Decomposition (DMD) to the adjoint-based optimal control of the Ginzburg-Landau equation.

The adjoint method uses the governing equations of the system to derive a gradient in the space of admissible inputs. This gradient points toward the choice of input which minimizes a cost functional. A gradient descent algorithm can then be employed to arrive at the optimum. Since the gradient comes from the governing equations, it can be difficult to derive and costly to compute. Dynamic Mode Decomposition computes a linear, low-dimensional approximation of the dynamics which allows the adjoint gradient to be computed more easily. We demonstrate that a ten-fold dimensional reduction of the Ginzburg-Landau system can be used to approximate the full-state gradient to within a margin of error of 0.02%.

The error in the reduced order gradient is dependent on the quality of the reduced order model (ROM). In order to compute an accurate gradient, we require that the ROM be robust to the range of possible dynamics in the search path of the gradient descent algorithm. This is difficult because one data set does not always capture the rich variety of possible behaviors of the system. Moreover, a reduced order model constructed from data acquired at a single operating point may not be robust to changes in control inputs or boundary conditions, thereby limiting the model's utility in control.

In order to evaluate the quality of a given data set, we use a condition called

Persistence of Excitation (PE). When a data sample satisfies the PE property, it guarantees that the data represents the dynamics well and that the hidden model parameters of the system can be approximated using methods from adaptive control. We prove that the persistence of excitation condition ensures that DMD-based reduced order models derived from PE data optimally approximate the true low-rank dynamics of the system. This method is system agnostic and is based on the idea of Persistence of Excitation. Since PE is often not possible to achieve for many systems, we propose an optimization problem which, when solved, specifies an input designed to drive the dynamical system toward a more excited state. We call this Optimally Persistent Excitation (OPE). The act of applying our OPE-enriched data to DMD is called PE-informed DMD. To demonstrate our method, we apply PE-informed DMD to the simulation and closed loop control of the Ginzburg-Landau equation. Our results show that when we start with a poorly representative baseline data set, we can improve the resulting DMD approximation of the low-rank state transition matrix by 20%.

To my family

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Daniel J. Bodony for his constant support and encouragement and for the wonderful opportunity of working in his research group.

This work was sponsored by the Office of Naval Research (ONR) as part of the Multidisciplinary University Research Initiatives (MURI) Program, under grant number N00014-16-1-2617.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

DMD	Dynamic Mode Decomposition
DMDc	Dynamic Mode Decomposition with Control
PE	Persistence of Excitation
OPE	Optimally Persistent Excitation
ROM	Reduced Order Model
POD	Proper Orthogonal Decomposition
bPOD	Balanced Proper Orthogonal Decomposition
ERA	Eigensystem Realization Algorithm
LQR	Linear Quadratic Regulator

CHAPTER 1

INTRODUCTION

Given a governing dynamical system and a cost functional, the simplest approach to solving an optimal control problem is to do an exhaustive search over all possible values of the control parameters. These types of searches involve performing many experiments or simulations and quickly become impractically expensive. Gradient-based search methods are an alternative approach. Given some guess for the optimal value of the input, a gradient is computed that points in the direction of a better choice of input. Then, a gradient descent algorithm, such as the conjugate gradient method, can be used to determine the next choice of input. One way to compute the gradient is by using a finite difference. This method is simple, but at every iteration of the search, it requires the evaluation of the governing equation at least $p + 1$ times where p is the number of input parameters.

In order to mitigate this computational cost, we employ the adjoint method [Teo et al., 1991]. In this method, the governing equations and the cost functional are manipulated to construct another system of differential equations called the adjoint equations. This system is then evaluated backward in time and the resulting adjoint data is collected. The adjoint data and the forward state data from the governing equations are then used to compute the gradient. At each iteration of the search method, the gradient computation only requires each of these evaluations to be completed once, regardless of the number of input parameters. Adjoint-based optimal control was successfully applied to the optimal control of the noise in a jet [Kim et al., 2014]. An introduction to the use of the adjoint method in aircraft design can be found in Giles and Pierce (2000). This method, however, requires detailed knowledge of the governing equations and can be difficult to execute on complicated problems. Instead of working with the full complexity of a high-dimensional fluid dynamics problem, we propose using data-driven system identification and model reduction techniques to simplify the computation.

Data-driven model reduction techniques are able to infer approximate low-rank models of a dynamical system without any a priori knowledge of the governing equations. Proper Orthogonal Decomposition (POD) is a data-driven system identification technique which seeks to find an approximate model which captures as much of a suitably-defined energy of the system as possible [Holmes et al., 1996]. However, this technique misses the low-energy components of the dynamics which can be critical to the overall evolution of the flow. There is a version of POD which is optimized for use in control called Balanced Proper Orthogonal Decomposition (bPOD). This method is inspired by balanced truncation [Moore, 1981] and attempts to find a projection onto which the system will be both controllable and observable. This method was applied to the control of transitional channel flow in [Ilak and Rowley, 2008]. A third method is called the Eigensystem Realization Algorithm (ERA) and is used to find models of impulse response data. A thorough discussion of these methods can be found in Rowley (2005).

Our main tool for constructing data-driven reduced order models is Dynamic Mode Decomposition (DMD)[Schmid, 2010]. This method computes modes from data and associates each of those modes with a frequency and a growth rate. We choose DMD because of its commonality, low computational cost, and because it can be modified for use with input-output systems using DMDc [Proctor et al., 2016]. DMD can easily be extended to nonlinear systems using Koopman Mode Decomposition (KMD) [Arbabi and Mezić, 2017]. Future work will attempt to use KMD to strengthen our methods and extend them to nonlinear control problems.

Both POD and DMD were applied to the adjoint-based optimal control of the FitzHugh-Nagumo equation in Karaszen et al. (2017). In this thesis, we apply DMD to a dynamical system based on the linearized Ginzburg-Landau equation [Chen and Rowley, 2011] and apply the adjoint method to optimally stabilize the system. It is found that the DMD-based reduced order model (ROM) is capable of approximating the adjoint gradients with sufficient precision to perform optimal control.

DMD and other related algorithms create reduced order models based only on the data they are provided. When this data poorly represents the full dynamics of the system, the reduced order model is not predictive. We want to prescribe an input to drive the system toward a better representation of the underlying dynamics. In order to determine which types of inputs will lead to

better data-driven reduced order models, we turn to the field of system identification. Connections between DMD and system identification have been pointed out, namely its connection to the eigensystem realization algorithm [Kutz et al., 2014]. We prove that when the data satisfies the persistence of excitation condition (PE), the resulting low-rank state transition matrix is the best possible low-rank approximation of the governing state transition matrix.

Since the PE condition is usually too strong to be enforced in full, an optimization problem is posed involving a PE-inspired cost functional. When the optimization problem is solved, we see a significant increase in the accuracy of our reduced order model when this method is applied to the linearized Ginzburg-Landau equation.

CHAPTER 2

MATHEMATICAL PRELIMINARIES

In this chapter, we describe the mathematical tools of model reduction and optimal control required to understand the proceeding analysis. We also introduce the Ginzburg-Landau system which we will use to test our methods.

2.1 Dynamic Mode Decomposition

Dynamic mode decomposition is a data-driven method that extracts dynamical information from data in the form of modes. Each mode is associated with a frequency and a growth rate [Schmid, 2010]. In order to compute the DMD of a system, we first collect M time snapshots of the state data. The resulting time series $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_M\}$ is then split into the unshifted $Q_1 = [\mathbf{q}_1\ \mathbf{q}_2\ \cdots\ \mathbf{q}_{M-1}]$ and shifted $Q_2 = [\mathbf{q}_2\ \mathbf{q}_3\ \cdots\ \mathbf{q}_M]$ snapshot matrices.

It is assumed that there exists a state transition matrix A such that the following relationship is satisfied:

$$Q_2 = AQ_1. \tag{2.1}$$

The matrix A transforms each snapshot of data in Q_1 to its time-shifted counterpart. Our goal is to find a rank r approximation of A called $\bar{A}$ which solves the following minimization problem:

$$\bar{A} = \overset{argmin}{A} \|Q_2 - AQ_1\|^2 \tag{2.2}$$

The Young-Eckart Theorem [Eckart and Young, 1936] states that this can be solved exactly using the Moore-Penrose pseudo-inverse. To accomplish this, we first take the economy-sized singular value decomposition of Q_1,

$$Q_1 = U\Sigma V^H. \tag{2.3}$$

The columns of the unitary matrices U and V are the left-singular and right-singular vectors of Q_1, respectively. The superscript H signifies the Hermitian transpose of V. Each vector corresponds to one of the non-negative, real singular values along the diagonal of Σ. By truncating all but the first r columns of U and V and all but the largest r singular values from Σ, we get the following low-rank approximation of the data:

$$Q_1 \approx U_r \Sigma_r V_r^H. \tag{2.4}$$

From here we compute the Moore-Penrose pseudo-inverse:

$$Q_1^\dagger = V_r \Sigma_r^{-1} U_r^H. \tag{2.5}$$

We can now approximate the state transition matrix A by left multiplying Q_2.

$$A \approx \bar{A} = Q_2 Q_1^\dagger = Q_2 V_r \Sigma_r^{-1} U_r^H \tag{2.6}$$

The matrix $\bar{A}$ has rank r but the same dimension as the original matrix A. In order to reduce the computational cost involved with working with our reduced order model, we can project our low-rank matrix onto the basis defined by the columns of U_r. The resulting low-rank, low-dimension state transition matrix is $\tilde{A}$,

$$\tilde{A} = U_r^H Q_2 V_r \Sigma_r^{-1}. \tag{2.7}$$

This operator governs the following reduced dimension dynamical system:

$$\tilde{\mathbf{q}}_{k+1} = \tilde{A} \tilde{\mathbf{q}}_k. \tag{2.8}$$

The corresponding reduced dimension state variable $\tilde{\mathbf{q}}$ can be projected back onto the full dimension state space by computing $\mathbf{q}_k \approx \tilde{U} \tilde{\mathbf{q}}_k$.

2.2 Dynamic Mode Decomposition with Control

DMD can be modified for use in linear input-output dynamical systems. This is called dynamic mode decomposition with control (DMDc). This method attempts to isolate the natural dynamics of the system from the contribution

of input [**?**]. This method only works when the input is purely additive.

In order to construct a reduced order model using DMDc, we start with both a time series of state data $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_M\}$ and input data $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_M\}$. Assume that the dynamical system has the form $\mathbf{q}_{k+1} = A\mathbf{q}_k + B\mathbf{u}_k$ where A is the unknown state transition matrix, and B is the known input matrix. Collect these into the matrices $Q_1 = [\mathbf{q}_1\ \mathbf{q}_2\ \cdots\ \mathbf{q}_{M-1}]$, $Q_2 = [\mathbf{q}_2\ \mathbf{q}_3\ \cdots\ \mathbf{q}_M]$, and $U = [\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_{M-1}]$. The data matrices satisfy the following relationship:

$$Q_2 = AQ_1 + BU. \tag{2.9}$$

We wish to find a low rank approximation of A. The minimization problem is now:

$$\bar{A} = \overset{argmin}{A}\ \|Q_2 - BU - AQ_1\|^2. \tag{2.10}$$

To solve this, we rearrange the system into $AQ_1 = Q_2 - BU$. Now, we take the truncated singular value decomposition of Q_1,

$$Q_1 = U_r \Sigma_r V_r^H. \tag{2.11}$$

As in standard DMD, we approximate A by using the Moore-Penrose pseudo-inverse:

$$A \approx \bar{A} = (Q_2 - BU)V_r \Sigma_r^{-1} U_r^H. \tag{2.12}$$

If a dimensional reduction is desired, $\bar{A}$ can be projected onto the modes in U_r,

$$\tilde{A} = U_r^H (Q_2 - BU)V_r \Sigma_r^{-1}. \tag{2.13}$$

The reduced dimension dynamical system is now:

$$\tilde{\mathbf{q}}_{k+1} = \tilde{A}\tilde{\mathbf{q}}_k + \tilde{B}\mathbf{u}_k. \tag{2.14}$$

2.3 Discrete Time Adjoint Method

In order to perform optimal control, we use the adjoint method. In particular, we use a discrete-time variant which is better suited to be implemented with DMD. The adjoint method involves manipulating the governing equations and cost functional to construct a system of differential equations called the adjoint equations. These adjoint equations produce adjoint data which when combined with the state data can be used to compute a gradient. This gradient points in the space of admissible inputs to the input that best minimizes the cost functional J [Giles and Pierce, 2000]. This saves computational cost because the governing equations and the adjoint system only need to be evaluated once per iteration of the search method. Therefore, the computational cost of the search is independent of the number of input parameters. We follow the analysis in Teo et al. (1991).

Given a discrete-time dynamical system of the form $\mathbf{q}_{k+1} = f(\mathbf{q}_k, \mathbf{u}_k)$ with $k \in \{1, \cdots, M\}$, we wish to find the gradient in the space of all possible control parameters U_{ad} which will lead us to the vector of optimal control parameters $\mathbf{u}^{opt} = [\mathbf{u}_1^{opt}\ \mathbf{u}_2^{opt}\ \cdots\ \mathbf{u}_{M-1}^{opt}]$ which minimizes the cost functional

$$J = \Phi(\mathbf{q}_M) + \sum_{k=1}^{M-1} j(\mathbf{q}_k, \mathbf{u}_k). \tag{2.15}$$

We first perturb the vector of control parameters $\mathbf{u}$ in the direction $\boldsymbol{\rho}$ with magnitude $\epsilon > 0$,

$$\mathbf{u}(\epsilon) = \mathbf{u} + \epsilon\boldsymbol{\rho}. \tag{2.16}$$

The perturbed state is now given by

$$\mathbf{q}_k(\epsilon) = \mathbf{q}_k(\mathbf{u}(\epsilon)). \tag{2.17}$$

The perturbed state is governed by the equation,

$$\mathbf{q}_{k+1}(\epsilon) = f(\mathbf{q}_k(\epsilon), \mathbf{u}_k(\epsilon)). \tag{2.18}$$

The variation in the state is given by

$$\delta\mathbf{q}_{k+1} = \left.\frac{d\mathbf{q}_{k+1}(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \frac{\partial f(\mathbf{q}_k, \mathbf{u}_k)}{\partial \mathbf{q}_k}\delta\mathbf{q}_k + \frac{\partial f(\mathbf{q}_k, \mathbf{u}_k)}{\partial \mathbf{u}_k}\boldsymbol{\rho}, \tag{2.19}$$

with $\delta\mathbf{q}_1 = 0$.

We can now compute

$$\frac{\partial J}{\partial \mathbf{u}}\boldsymbol{\rho} = \left.\frac{dJ(\mathbf{u}(\epsilon))}{d\epsilon}\right|_{\epsilon=0} = \frac{\partial \Phi(\mathbf{q}_M)}{\partial \mathbf{q}_M}\delta\mathbf{q}_M + \sum_{k=1}^{M-1}\left[\frac{\partial j(\mathbf{q}_k,\mathbf{u}_k)}{\partial \mathbf{q}_k}\delta\mathbf{q}_k + \frac{\partial j(\mathbf{q}_k,\mathbf{u}_k)}{\partial \mathbf{u}_k}\boldsymbol{\rho}\right]. \tag{2.20}$$

We define the Hamiltonian sequence

$$H(\mathbf{q}_k,\mathbf{u}_k,\mathbf{z}_{k+1}) = j(\mathbf{q}_k,\mathbf{u}_k) + \mathbf{z}_{k+1}^H f(\mathbf{q}_k,\mathbf{u}_k). \tag{2.21}$$

The set of Lagrange multipliers $\mathbf{z}$ will be known as the adjoint state or costate.

Substitution of the Hamiltonian into the expression for the sensitivity of the cost functional gives

$$\begin{aligned}\frac{\partial J}{\partial \mathbf{u}}\boldsymbol{\rho} = \frac{\partial \Phi(\mathbf{q}_M)}{\partial \mathbf{q}_M}\delta\mathbf{q}_M + \sum_{k=1}^{M-1}\Bigg[&\frac{\partial H(\mathbf{q}_k,\mathbf{u}_k,\mathbf{z}_{k+1})}{\partial \mathbf{q}_k}\delta\mathbf{q}_k - \mathbf{z}_{k+1}^H\frac{\partial f(\mathbf{q}_k,\mathbf{u}_k)}{\partial \mathbf{q}_k}\delta\mathbf{q}_k \\ &+ \frac{\partial H(\mathbf{q}_k,\mathbf{u}_k,\mathbf{z}_{k+1})}{\partial \mathbf{u}_k}\boldsymbol{\rho} - \mathbf{z}_{k+1}^H\frac{\partial f(\mathbf{q}_k,\mathbf{u}_k)}{\partial \mathbf{u}_k}\boldsymbol{\rho}\Bigg].\end{aligned} \tag{2.22}$$

The adjoint data is computed via the adjoint difference equations

$$\mathbf{z}_k^H = \frac{\partial H(\mathbf{q}_k,\mathbf{u}_k,\mathbf{z}_{k+1})}{\partial \mathbf{q}_k} \tag{2.23}$$

which are marched backward in time starting from the adjoint terminal conditions $\mathbf{z}_M^H = \frac{\partial \Phi}{\partial \mathbf{q}_M}$.

Applying the adjoint difference equations to our equation for the sensitivity of the cost functional gives

$$\frac{\partial J}{\partial \mathbf{u}}\boldsymbol{\rho} = \sum_{k=1}^{M-1}\frac{\partial H(\mathbf{q}_k,\mathbf{u}_k,\mathbf{z}_{k+1})}{\partial \mathbf{u}_k}\boldsymbol{\rho}. \tag{2.24}$$

Since $\boldsymbol{\rho}$ is arbitrary, we can now compute the gradient of J with respect to $\mathbf{u}$ directly with the formula

$$\nabla_{\mathbf{u}} J = \frac{\partial J}{\partial \mathbf{u}} = \sum_{k=1}^{M-1}\frac{\partial H(\mathbf{q}_k,\mathbf{u}_k,\mathbf{z}_{k+1})}{\partial \mathbf{u}_k}. \tag{2.25}$$

This gradient can now be used in a gradient-descent search method until convergence on the optimal set of control parameters $\mathbf{u}^{opt}$. These control parameters can be used to approximately minimize the cost functional J.

The following is the algorithm used in the adjoint method. A simple gradient descent method is used to update the value of the input at each iteration, but more sophisticated methods can be used as well.

Algorithm 1: Discrete-Time Adjoint Method

Result: The optimal input $\mathbf{u}^{opt}$
Set initial guess $\mathbf{u}^1$;
Choose tolerance $\epsilon > 0$ and search method step size $\delta > 0$;
while $\|\nabla_{\mathbf{u}^i} J\| > \epsilon$ **do**
  Evaluate the system $\mathbf{q}_{k+1} = f(\mathbf{q}_k, \mathbf{u}_k^i)$ forward in time starting at initial condition $\mathbf{q}_1$;
  Evaluate the system $\mathbf{z}_k^H = \frac{\partial H(\mathbf{q}_k, \mathbf{u}_k^i, \mathbf{z}_{k+1})}{\partial \mathbf{q}_k}$ backward in time starting at terminal condition $\mathbf{z}_M^H = \frac{\partial \Phi}{\partial \mathbf{q}_M}$;
  Compute $\nabla_{\mathbf{u}^i} J = \sum_{k=1}^{M-1} \frac{\partial H(\mathbf{q}_k, \mathbf{u}_k^i, \mathbf{z}_{k+1})}{\partial \mathbf{u}_k}$;
  Compute $\mathbf{u}^{i+1} = \mathbf{u}^i + \delta \nabla_{\mathbf{u}^i} J$;
end

2.4 The Ginzburg-Landau Equation

We demonstrate our methods on the linearized Ginzburg-Landau system with an additive input term. The Ginzburg-Landau equation is a simplified version of the Navier-Stokes equation which can be tuned to exhibit a wide range of stability characteristics [Chen and Rowley, 2011]. The continuous-time governing equations are given as follows:

$$\frac{\partial \mathbf{q}}{\partial t} = -\nu \frac{\partial \mathbf{q}}{\partial x} + \boldsymbol{\mu} \circ \mathbf{q} + \gamma \frac{\partial^2 \mathbf{q}}{\partial x^2} + \mathcal{B} u(t) \tag{2.26}$$

$$\mathbf{q}(0) = \mathbf{q}_0. \tag{2.27}$$

The state $\mathbf{q}$ is complex. The tuneable parameters $(\nu, \boldsymbol{\mu}, \gamma)$ are the complex advection speed, amplification factor, and complex diffusion parameter, respectively. The amplification factor modifies the state by the Hadamard product $\circ$. The term $u(t)$ is a time-dependent, scalar input term applied to the system by the input matrix $\mathcal{B}$.

In the following analysis, we will use the "subcritical" formulation of the Ginzburg-Landau equation. Here, the amplification factor $\boldsymbol{\mu}$ (Figure 2.1) is

negative in the extremes of the domain and therefore has a stabilizing effect on perturbations in this region. The center of the domain experiences a small, positive amplification which defines the part of the domain that experiences oscillatory dynamics. The subcritical parameters are $\nu = 2 + 2i$, $\mu(x) = 0.34 - 0.005x^2$, and $\gamma = 1 - i$. More detail on the implementation of the Ginzburg-Landau system including the Hermite-polynomial based discretization of the derivative operators can be found in [Chen and Rowley, 2011].

The operator $\mathcal{A} = -\nu\frac{\partial}{\partial x} + \boldsymbol{\mu} + \gamma\frac{\partial^2}{\partial x^2}$ represents the state component of the continuous-time dynamics. The discrete-time operator A is derived using the relation $A = e^{\mathcal{A}\Delta t}$ where Δt is the time step. The discrete-time counterpart to the input matrix is given by $B = \mathcal{A}^{-1}(e^{\mathcal{A}\Delta t} - I)\mathcal{B}$. The resulting discrete-time dynamical system is of the form:

$$\mathbf{q}_{k+1} = A\mathbf{q}_k + Bu_k. \tag{2.28}$$

The discrete-time system starts at time step $k = 1$ with initial condition $\mathbf{q}_1$. For illustrative purposes, we define two initial conditions $\mathbf{q}_{\{1,1\}}$ and $\mathbf{q}_{\{1,2\}}$ which are Gaussian humps with centers at $x_1 = -20$ and $x_2 = -10$, respectively. These are shown in Figure 2.1.

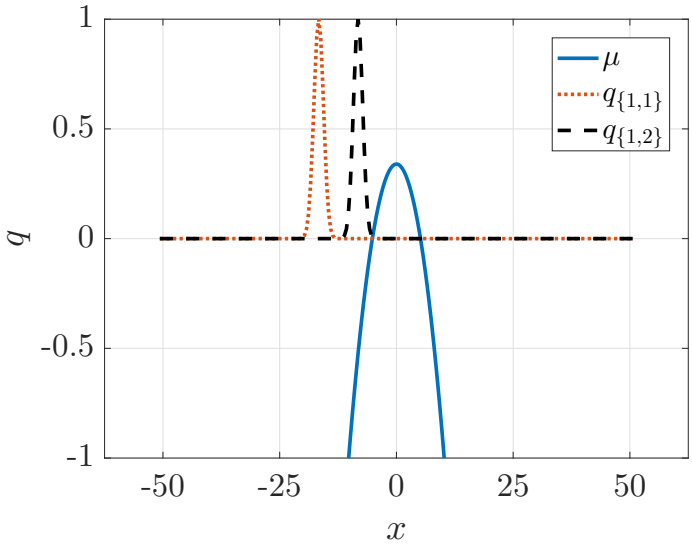

Figure 2.1: The amplification factor $\boldsymbol{\mu}$ is given by the blue solid line. The initial conditions $\mathbf{q}_{\{1,1\}}$ and $\mathbf{q}_{\{1,2\}}$ are given by the red dotted line and the black dashed line, respectively.

The resulting uncontrolled dynamics of the system with initial conditions $\mathbf{q}_{\{1,1\}}$ and $\mathbf{q}_{\{1,2\}}$ are shown in Figs. 2.2 and 2.3, respectively. Note that the initial condition $\mathbf{q}_{\{1,1\}}$ is far into the negative amplification region and therefore decays rapidly in time. However, the initial condition $\mathbf{q}_{\{1,2\}}$ is close enough to the center of the domain that it is convected rightward into the positive amplification region and displays ongoing, oscillatory dynamics. We will establish the effect that the choice of data has on the quality of the resulting DMD-based ROM and DMD-informed adjoint gradients.

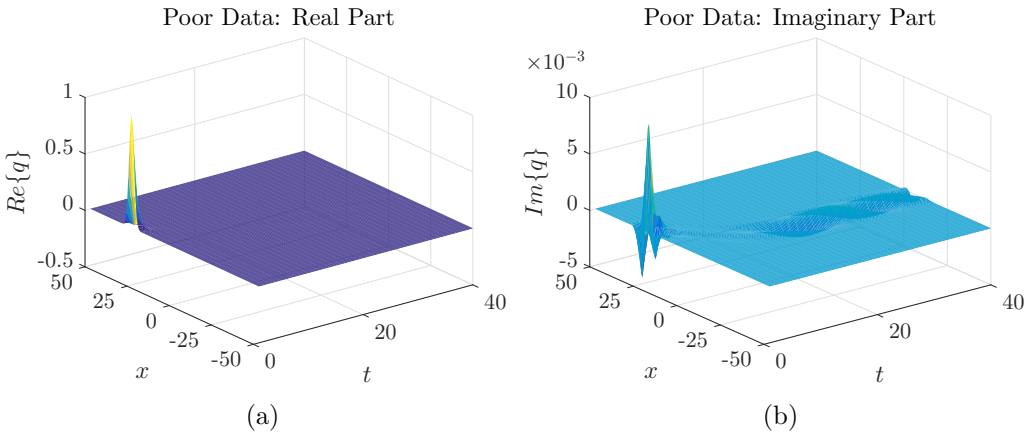

Figure 2.2: The real and imaginary parts of the uncontrolled, poorly representative data with initial condition $\mathbf{q}_{\{1,1\}}$.

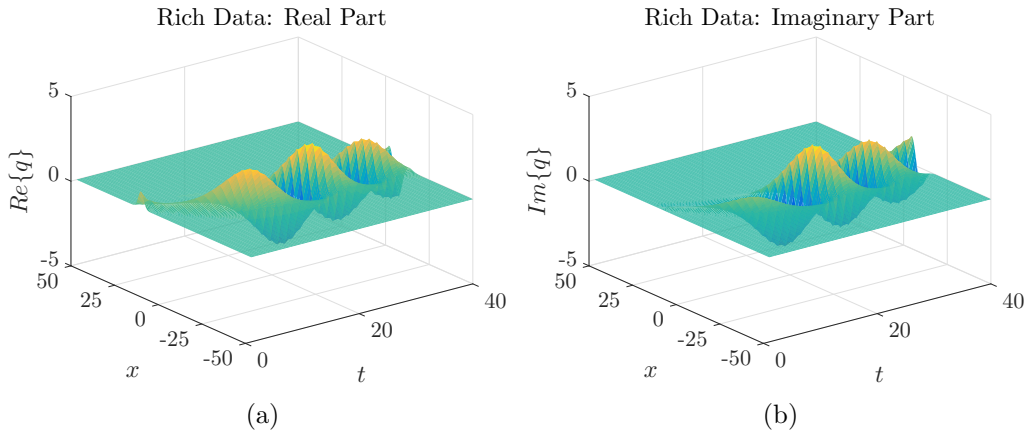

Figure 2.3: The real and imaginary parts of the uncontrolled, rich data with initial condition $\mathbf{q}_{\{1,2\}}$.

CHAPTER 3

DMD-BASED OPTIMAL CONTROL

We now apply DMD to the computation of the adjoint gradient in order to stabilize the Ginzburg-Landau input-output system.

3.1 Formulation

We are given the following discretization of the Ginzburg-Landau equation

$$\mathbf{q}_{k+1} = f(\mathbf{q}_k, u_k) = A\mathbf{q}_k + Bu_k. \tag{3.1}$$

We want find an input u which stabilizes the system. To this end, we define the cost functional

$$J = \sum_{k=1}^{M-1} \left[\mathbf{q}_k^H Q \mathbf{q}_k + \bar{u}_k R u_k \right]. \tag{3.2}$$

The diagonal matrix Q contains the state penalty weights, and the scalar R is the input penalty weight. We take $Q = I$ and $R = 1$. The pair (f, J) defines an optimization problem with optimal solution u^{opt}. In order to compute u^{opt}, we follow the analysis in Section 2.3 and construct the adjoint system:

$$\mathbf{z}_k^H = \mathbf{z}_{k+1}^H A + 2Q\mathbf{q}_k^H. \tag{3.3}$$

The resulting gradient is:

$$\nabla_{u_k} J = \mathbf{z}_{k+1}^H B + 2Ru_k. \tag{3.4}$$

We want to investigate whether the low-dimensional, DMD-approximated operators $\tilde{A}$ and $\tilde{B}$ can be used to estimate $\nabla_{u_k} J$. First, we define the low-dimension dynamical system:

$$\tilde{\mathbf{q}}_{i+1} = \tilde{A}\tilde{\mathbf{q}}_k + \tilde{B}u_k. \tag{3.5}$$

The corresponding reduced dimension cost is

$$\tilde{J} = \sum_{k=1}^{M-1} \left[\tilde{\mathbf{q}}_k^H \tilde{Q} \tilde{\mathbf{q}}_k + \bar{u}_k R u_k \right], \tag{3.6}$$

where $\tilde{Q}$ is dimensionally reduced using the expression $\tilde{Q} = \tilde{U}^H Q \tilde{U}$. Following the discrete time adjoint method, the following DMD-informed reduced order adjoint equations can be constructed:

$$\tilde{\mathbf{z}}_k^H = \tilde{\mathbf{z}}_{k+1}^H \tilde{A} + 2\tilde{Q}\tilde{\mathbf{q}}_k^H, \tag{3.7}$$

with terminal conditions $\tilde{\mathbf{z}}_M^H = 2\tilde{Q}\tilde{\mathbf{q}}_M^H$. After computing the reduced-dimension adjoint data backward in time, the DMD-based adjoint gradient is given by:

$$\nabla_{u_k} \tilde{J} = \tilde{\mathbf{z}}_{k+1}^H \tilde{B} + 2Ru_k. \tag{3.8}$$

The following tests address the question of whether $\nabla_{u_k} J \approx \nabla_{u_k} \tilde{J}$.

3.2 Results

We want to verify whether the DMD-based adjoint gradient $\nabla_{u_k} \tilde{J}$ approximates the full-state adjoint gradient $\nabla_{u_k} J$ well enough to solve optimal control problems. In order to do this, we compare both to the finite difference-based gradient

$$\nabla_{u_k} J_{FD} = \frac{J(u_k + \epsilon) - J(u_k)}{\epsilon}, \tag{3.9}$$

which is computed separately for each time sample k. The finite difference step size ϵ varies between 10^{-10} and 10^0. The resulting gradient error is given by

$$e_{grad} = \frac{\|\nabla_u J - \nabla_u J_{FD}\|}{\|\nabla_u J_{FD}\|}, \tag{3.10}$$

with the norms taken over all time.

Figure 3.1 shows the dependence of e_{grad} on the finite difference step size

used to compute $\nabla_u J_{FD}$. For comparison, the gradient error resulting from ROMs informed by the poor data from Fig. 2.2 is compared to that of the rich data in Fig. 2.3. The rich data approximates the adjoint gradient with an error multiple orders of magnitude better than the poor data. The converged value of the rich data gradient error is shown for a wide selection of modes in Figure 3.2. For this particular configuration, 21 seems to be the optimal choice for the rank of the DMD approximation. The resulting error at this point is $e_{grad} = 1.062 \cdot 10^{-4}$. If any more modes are taken, the gradient error increases. This is because the data used to compute DMD becomes rank deficient whereby the singular values in Σ become small, their corresponding modes represent noise as opposed to true dynamics. The trend in singular values is illustrated in Figure 3.3.

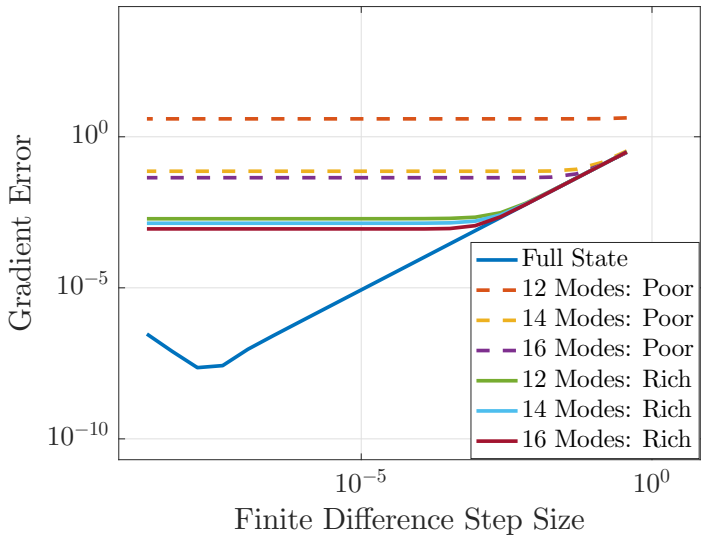

Figure 3.1: The error in the DMD-reduced and full-state gradients compared to a finite difference gradient. The dashed lines represent gradients derived from the poor data in 2.2, and the solid lines use the rich data from 2.3.

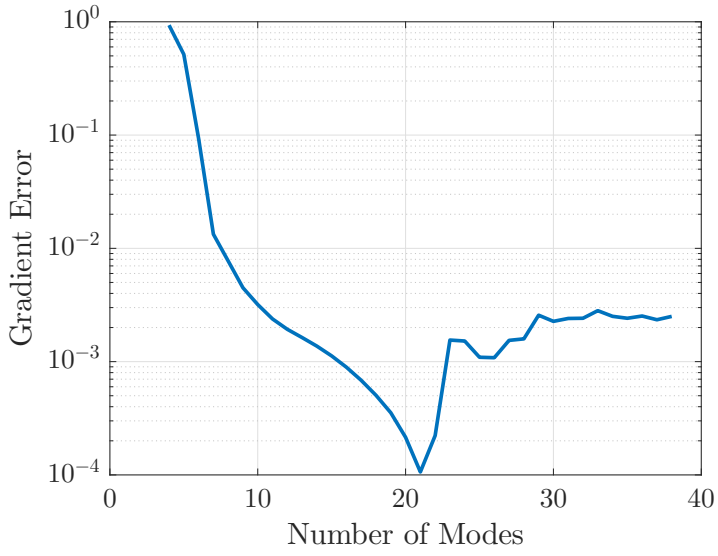

Figure 3.2: The converged DMD-reduced gradient error as a function of the number of modes. The error grows after the number of modes reaches 21 because of the rank deficiency of the data matrix Q_1.

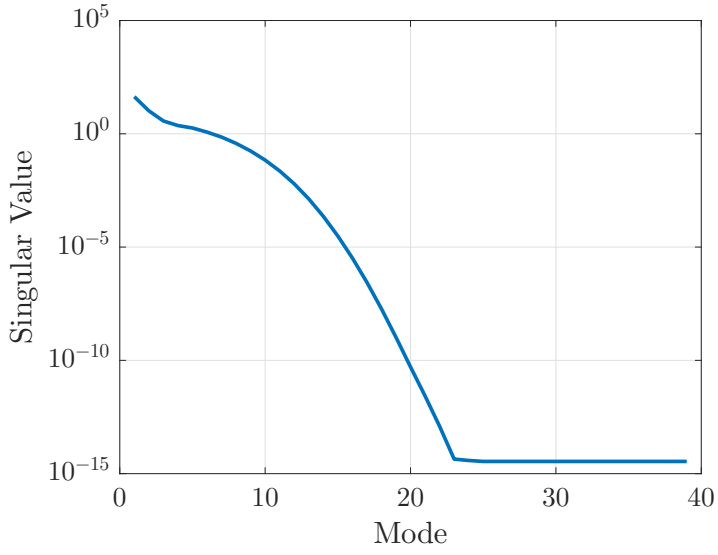

Figure 3.3: The singular values of the un-shifted data matrix Q_1.

Next, we use these gradients along with a conjugate gradient minimization method to find the converged control. The Ginzburg-Landau system starts with the input set to $u_k = 0 \ \forall k$, and three different types of adjoint-based

gradient are used to converge to approximations of the optimal input u^{opt} which minimizes J. The trend in J as a function of the number of iterations of the search method is shown in Figure 3.4. For comparison, the adjoint gradients of the 6 and 7 modes DMD-reduced system are used alongside the full state adjoint gradient.

The uncontrolled system resulted in a cost of $J = 2.019 \cdot 10^3$. The 6 and 7 mode DMD-based adjoint methods converged to values of $J = 32.67$ and $J = 10.30$, respectively. Using fewer than 6 modes causes a divergence in the cost while more than 7 modes results in almost perfect convergence of the adjoint search. The full state adjoint method converged to a value of $J = 9.405$. For comparison, the exact optimal solution was computed by solving the continuous algebraic Ricatti equations and constructing a linear quadratic regulator (LQR) [Corless, 2003]. The resulting cost was $J_{LQR} = 6.047$.

This test demonstrates the effectiveness of DMD in solving the optimal control of the Ginzburg-Landau system. Only 7 modes were required to approximate the full state adjoint gradient well enough to converge to nearly the same optimal input and cost J.

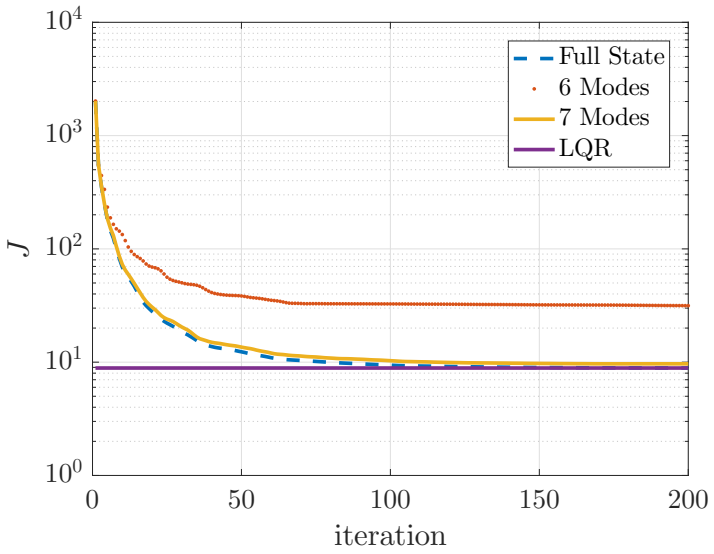

Figure 3.4: Convergence of the adjoint method to an approximate optimal control. The dotted red line and the solid yellow line are the 6 and 7 mode DMD-based adjoint method. The dashed blue line is the full state adjoint method. The horizontal purple line is placed at the value of the best possible J given by the LQR solution.

CHAPTER 4

CONTROL-INFORMED DMD

The quality of our DMD approximations and by extension, our optimal control depends on the quality of the data. The data must represent the dynamics relevant to the system near the initial guess of input $\mathbf{u}^1$, the optimal input $\mathbf{u}^{opt}$, and all points along the search path. In order to ensure this robustness to inputs, we need to ensure that the reduced order models we construct with DMD represent the true dynamics of the system. We prove that Persistence of Excitation (PE) is a sufficient condition on the data to ensure that the resulting DMD-based dynamical system represents the optimal low-rank approximation of the dynamics. In order to enrich our data, we propose the use of an input. The choice of input is inspired by PE and is computed by solving an optimal control problem. We demonstrate the effectiveness of this method on the Ginzburg-Landau equation.

4.1 Persistence of Excitation

We want to use an input to enrich our data sets. What exactly does it mean for our data to be "rich"? We will use the idea of persistent excitation from the field of system identification [Bai and Sastry, 1985]. In adaptive control, persistence of excitation is a condition on data that guarantees that properly chosen adaptive laws will cause estimates of system parameters to converge uniformly and exponentially to the actual underlying system parameters [Ioannou and Sun, 2012].

Define the regressor vector $\boldsymbol{\phi} = [\mathbf{q}^H(t) \ \cdots \ \mathbf{q}^H(t-P)]^H$ where P is a positive integer. A signal $\boldsymbol{\phi}(t)$ is persistently exciting in P time steps if $\exists P \in \mathbb{Z}_+$ and $\exists \alpha > 0$ such that

$$\sum_{t=t_0+1}^{t_0+P} \boldsymbol{\phi}(t)\boldsymbol{\phi}^H(t) \geq \alpha I \ \forall t_0. \tag{4.1}$$

In order to relate this to DMD, we will prove that if the state data $\mathbf{q}$ satisfies the persistence of excitation condition, then the DMD-estimated state transition matrix $\bar{A}$ is the best possible low-rank estimate A_r of the exact discrete-time state transition matrix A.

Suppose we are given the discrete-time, linear dynamical system $\mathbf{q}_{k+1} = A\mathbf{q}_k$ with initial condition $\mathbf{q}_1$. This system generates the time series of data $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_M\}$. For convenience, we parameterize the data matrices Q_1 and Q_2 by the initial condition $\mathbf{q}_1$ as follows:

$$Q_1(\mathbf{q}_1) = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_{M-1}] \tag{4.2}$$

$$Q_2(\mathbf{q}_1) = [\mathbf{q}_2 \ \mathbf{q}_3 \ \cdots \ \mathbf{q}_M]. \tag{4.3}$$

We now define the exact low-rank version of this dynamical system. First, we compute the low-rank version of the original state transition matrix A by computing the svd: $A = U\Sigma V^H$. If we truncate all but the r largest singular values in Σ as well as all the corresponding modes from U and V, we get the exact low rank state transition matrix $A \approx A_r = U_r\Sigma_r V_r^H$. The corresponding low rank dynamical system is $\mathbf{q}_{k+1}^r = A_r\mathbf{q}_k^r$. Given initial condition $\mathbf{q}_1$, the data matrices generated by the low-rank system are $Q_1^r(\mathbf{q}_1)$ and $Q_2^r(\mathbf{q}_1)$.

The third operator we need to define is the DMD-approximated, low-rank state transition matrix $\bar{A}$. In order to perform DMD, we specify a particular initial condition $\hat{\mathbf{q}}_1$. DMD minimizes the following error norm $\|Q_2(\hat{\mathbf{q}}_1) - \bar{A}Q_1(\hat{\mathbf{q}}_1)\|$. Using the svd and the pseudo-inverse, we find the matrix $\bar{A}$ of specified rank r such that $\|Q_2(\hat{\mathbf{q}}_1) - \bar{A}Q_1(\hat{\mathbf{q}}_1)\|$ is minimized.

By definition,

$$\|Q_2(\mathbf{q}_1) - AQ_1(\mathbf{q}_1)\| = 0 \ \forall \mathbf{q}_1 \tag{4.4}$$

and

$$\|Q_2^r(\mathbf{q}_1) - A_rQ_1^r(\mathbf{q}_1)\| = 0 \ \forall \mathbf{q}_1 \tag{4.5}$$

The Young-Eckart Theorem guarantees that the norm $\|Q_2(\hat{\mathbf{q}}_1) - \bar{A}Q_1(\hat{\mathbf{q}}_1)\|$ is minimized when $\bar{A}$ is computed using the pseudo-inverse. However, this does not guarantee that the norm is identically 0; the quality of $\bar{A}$ is dependent on the quality of the data.

Our goal is to prove that $\bar{A} = A_r$ when the data used to compute $\bar{A}$ satisfies the PE condition. In order to do this, we need to prove the following:

$$\|Q_2^r(\mathbf{q}_1) - \bar{A}Q_1^r(\mathbf{q}_1)\| = 0 \ \forall \mathbf{q}_1. \tag{4.6}$$

Take the signal $Q^r = [\mathbf{q}_1 \ \mathbf{q}_2^r \ \cdots \ \mathbf{q}_M^r]$ generated by the dynamical system $\mathbf{q}_{k+1}^r = A_r\mathbf{q}_k^r$ with initial condition $\mathbf{q}_1$. If this signal is PE, then a suitable choice of gradient based search (called an adaptive law) will guarantee that $\|\bar{A} - A_r\| \to 0$ exponentially fast. Since $Q_1^r(\mathbf{q}_1)$ is a constant,

$$\|\bar{A} - A_r\| \to 0 \implies \|\bar{A}Q_1^r(\mathbf{q}_1) - A_rQ_1^r(\mathbf{q}_1)\| \to 0. \tag{4.7}$$

Adding $(Q_2^r(\mathbf{q}_1) - Q_2^r(\mathbf{q}_1)) = 0$ inside the norm has no effect, so:

$$\|(Q_2^r(\mathbf{q}_1) - Q_2^r(\mathbf{q}_1)) + \bar{A}Q_1^r(\mathbf{q}_1) - A_rQ_1^r(\mathbf{q}_1)\| \to 0. \tag{4.8}$$

Rearranging, we get:

$$\|(Q_2^r(\mathbf{q}_1) - \bar{A}Q_1^r(\mathbf{q}_1)) - (Q_2^r(\mathbf{q}_1) - A_rQ_1^r(\mathbf{q}_1))\| \to 0. \tag{4.9}$$

Since $\|Q_2^r(\mathbf{q}_1) - A_rQ_1^r(\mathbf{q}_1)\| = 0 \ \forall \mathbf{q}_1$,

$$\|(Q_2^r(\mathbf{q}_1) - \bar{A}Q_1^r(\mathbf{q}_1))\| \to 0 \ \forall \mathbf{q}_1. \tag{4.10}$$

Therefore, the PE condition implies convergence of DMD to the best possible low-rank approximation of the state transition operator of the full-state system. When the PE condition is met, we can expect the resulting DMD-based reduced order model to be more robust to inputs because it more closely represents the actual dynamics of the system.

4.2 Measures of Performance

In order to measure the robustness of ROMs, we propose two choices of normalized error: the operator error e_{op} and the reconstruction error e_{recon}.

Given a DMD-approximated state transition matrix $\bar{A}$ with rank r, the corresponding exact low rank state transition matrix of rank r is A_r. We define the operator error

$$e_{op} = \frac{\|\bar{A} - A_r\|}{\|A_r\|}. \tag{4.11}$$

This provides a way of measuring the degree to which the DMD-approximated ROM represents the true underlying low-rank dynamics of the system.

The next error measure measures the degree to which the dynamics of the DMD-approximated ROM represents the dynamics relevant to the stabilization optimal control problem presented in Section 3.1. This error measure compares the data given by the evolution of the DMD-ROM to the exact data of the full-state system.

For example, the data in Figure 2.2 generated by the uncontrolled Ginzburg-Landau equation with initial condition $\mathbf{q}_{\{1,1\}}$ is fed into DMD. This data is used to compute an approximate state transition matrix $\bar{A}_1$ where $rank(\bar{A}_1) = 12$. The state transition operator $\bar{A}_1$ is applied to the initial condition $\mathbf{q}_{\{1,2\}}$ in order to reconstruct the original data. The results of this reconstruction are shown in 4.1. For comparison, the same process is also done using a model derived from the $\mathbf{q}_{\{1,2\}}$ data and is shown in Figure 4.2.

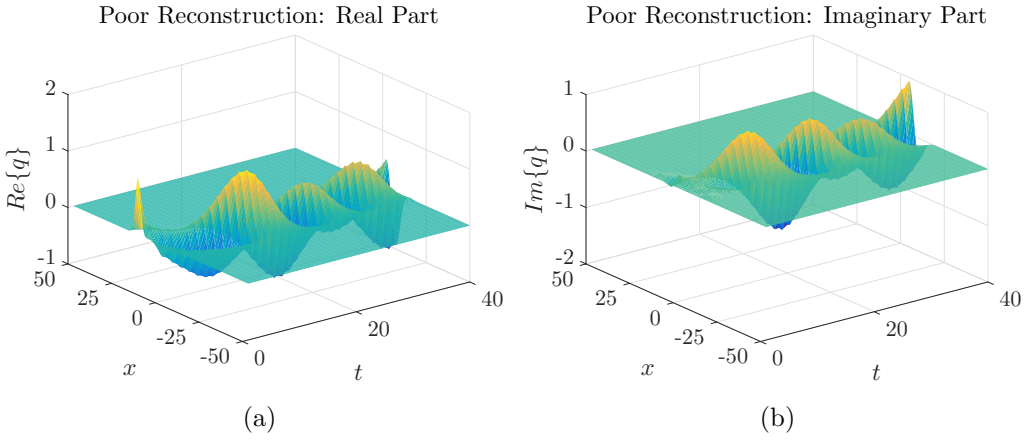

Figure 4.1: Here, a DMD-derived reduced order model is computed using the uncontrolled $\mathbf{q}_{\{1,1\}}$ data. The resulting model is used to reconstruct the data from the uncontrolled system with initial condition $\mathbf{q}_{\{1,2\}}$. The relative error is $e_{recon} = 137\%$.

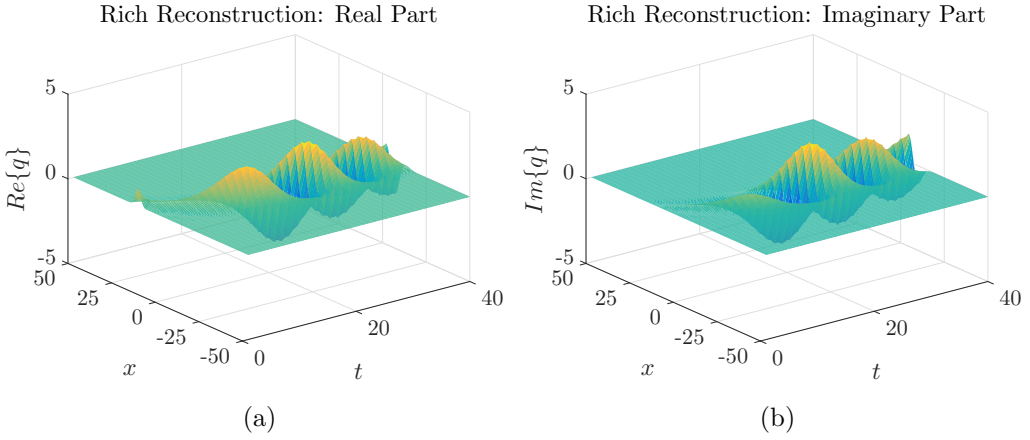

Figure 4.2: Here, a DMD-derived reduced order model is computed using the uncontrolled $\mathbf{q}_{\{1,2\}}$ data. The resulting model is used to reconstruct the same data set. The relative error is $e_{recon} = 2\%$.

The exact data with initial condition $\mathbf{q}_{\{1,2\}}$ is called $\mathbf{q}_{exact}$ and the reconstructed data is $\mathbf{q}_{recon}$. The relative error between the two data sets is given by the following equation:

$$e_{recon} = \frac{\|\mathbf{q}_{recon} - \mathbf{q}_{exact}\|}{\|\mathbf{q}_{exact}\|} \tag{4.12}$$

The reconstruction error given by the reduced order model $\bar{A}_1$ is $e_{recon} = 137\%$, and the error corresponding to $\bar{A}_2$ is $e_{recon} = 2\%$. As can be seen from Figure 4.1, the reduced order model from $\bar{A}_1$ captures some of the qualitative behavior of the system, but is quantitatively wrong.

The question now is this: Can we use an input to enrich the $\mathbf{q}_{\{1,1\}}$ data and construct a better version of $\bar{A}_1$?

4.3 Optimally Persistent Excitation (OPE)

Now, it is time to determine which input would best "enrich" our data set. We will first cast the Ginzburg-Landau equation in the form of a continuous-time input-output system:

$$\dot{\mathbf{q}} = \mathcal{A}\mathbf{q} + \mathcal{B}u. \tag{4.13}$$

The input matrix $\mathcal{B}$ is a Gaussian hump located at the variable location $x_{\mathcal{B}}$ with magnitude 1. The input is taken to be $u(t) = 1\ \forall t$.

The corresponding discrete time state transition and input operators are $A = e^{\mathcal{A}\Delta t}$ and $B = \mathcal{A}^{-1}(e^{\mathcal{A}\Delta t} - I)\mathcal{B}$, respectively. We arrive at the following discrete-time dynamical system:

$$\mathbf{q}_{k+1} = A\mathbf{q}_k + Bu_k. \tag{4.14}$$

The initial condition is taken to be $\mathbf{q}_{\{1,1\}}$. Normally, the uncontrolled system with this initial condition produces a low quality reduced order model. It will be our goal to determine how to choose the location $x_{\mathcal{B}}$ of the Gaussian input with the goal of enriching this data.

Persistency of excitation is a strong assumption on the data, and it is usually not possible to know beforehand how to enforce PE when the dynamics are unknown. In fact, for the Ginzburg-Landau equation, it is unrealistic to ever expect the data to be PE. This is because perturbations of the state are quickly stabilized on the ends of the domain. Instead of strictly enforcing this condition, we will define a cost functional J_{PE} which measures the degree to which the system fails to be persistently excited. We will use a finite difference gradient to find the direction of greatest increase of J_{PE} and use gradient descent to improve J_{PE}.

First, we collect our state data into the regressor matrix $\Phi = [\mathbf{q}_1 \mathbf{q}_2 \ ... \ \mathbf{q}_M]$. We are interested in tracking where this data shows interesting dynamics in space and time. We define the following collection of Gaussians:

$$G(x,t)_{ij} = e^{\frac{-(x-x_i)^2}{2\sigma_1^2}} - e^{\frac{-(t-t_j)^2}{2\sigma_2^2}}. \tag{4.15}$$

The locations in space and time of the peaks of these Gaussians are given by the values of i and j, respectively. These values are taken to give coarse samplings of the spatial and time domains. The Gaussians are plotted in the $x-t$ plane in Figure 4.3.

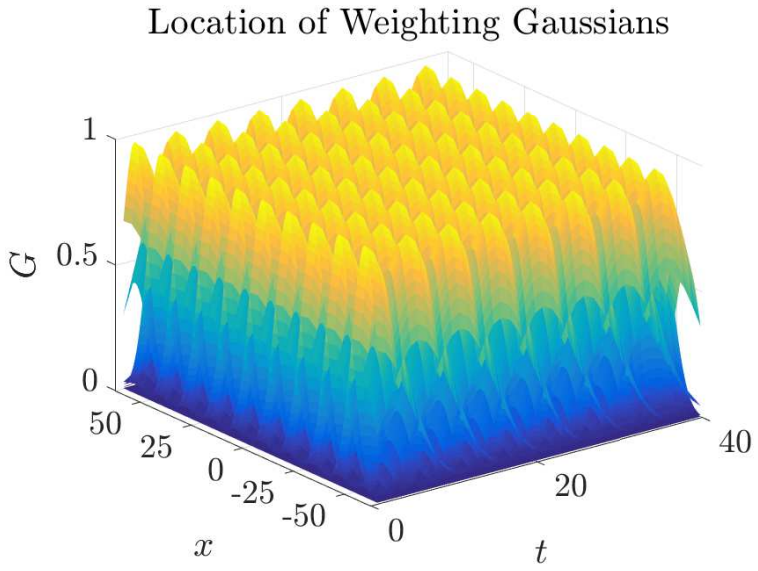

Figure 4.3: Gaussian weighting functions used in the PE-inspired cost functional

We define our cost functional:

$$J_{PE} = \sum_{i,j} \frac{1}{\| \frac{\dot{\Phi}}{\|\dot{\Phi}\|} \circ G_{ij} \|}. \tag{4.16}$$

Since activity in the data is more important that simply the magnitude of the data, we take the derivative of the regressor matrix $\dot{\Phi}$. This is then normalized because we care more about the distribution of activity within the domain than we do the magnitude of this activity. The term $\dot{\Phi}/\|\dot{\Phi}\|$ is then multiplied elementwise with one of the Gaussians G_{ij}. The norm of this term tells us how "active" the state is in the region around the i^{th} point in

space and the j^{th} point in time. If no activity occurs in this region, then the inverse of this term results in a strong penalty, but the Gaussians ensure no blowup occurs. In this way, we see that we want to minimize J_{PE}.

Figure 4.4 shows the trend in J_{PE} plotted against all possible choices of $x_{\mathcal{B}}$. J_{PE} is high at the extremes of the domain because perturbations in these regions are quickly stabilized by the negative amplification factor. However, a broad patch to the left of center of the domain shows the lowest values of J_{PE}. This is because inputs in this region are convected rightward into the positive amplification region. This location for the minimum of J is near the regions where e_{op} and e_{recon} are minimized. Therefore, our choice of J_{PE} is helpful in trying to improve the robustness of our DMD-ROM. Algorithm 2 searches for the input location which minimizes J_{PE}.

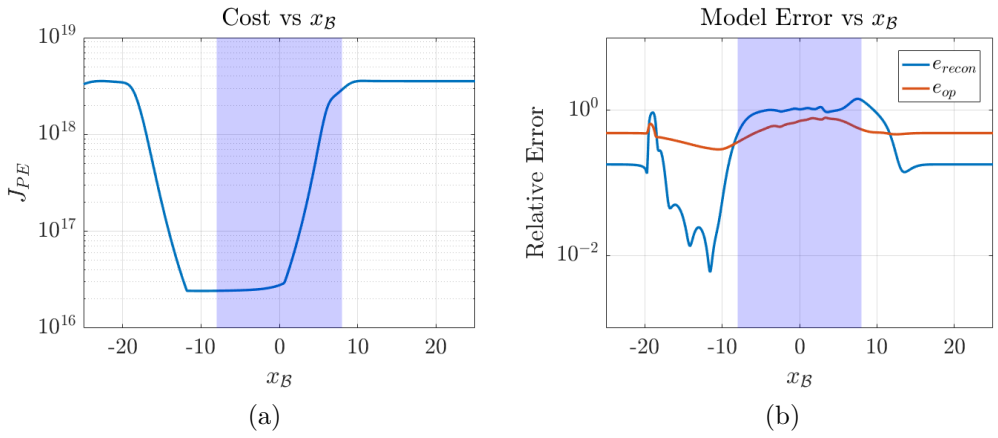

Figure 4.4: (a) The cost defined by J_{PE} as a function of the input Gaussian location. (b) The relative reconstruction and operator errors as a function of the input Gaussian location. Note that the lowest value of J_{PE} occurs just left of $x = -10$ and roughly corresponds to the value of x which shows the lowest errors. The blue rectangle shows the location of the positive amplification region.

Algorithm 2: Optimally Persistent Excitation

Result: The optimally exciting input location $x_{\mathcal{B}}^{PE}$
Set initial guess $x_{\mathcal{B}}^{1} = -20$;
Fix $u_k = 1 \; \forall k$;
Choose tolerance ϵ, perturbation magnitude α, and search method step size δ;
while $\|\nabla_{x_{\mathcal{B}}^{i}} J_{PE}\| > \epsilon$ **do**
  Construct $\mathcal{B}_{pert}$ with peak location at $x_{\mathcal{B}}^{i} + \alpha$;
  $B_{pert} = \mathcal{A}^{-1}(e^{\mathcal{A}\Delta t} - I)\mathcal{B}_{pert}$;
  Advance the system $\mathbf{q}_{k+1} = A\mathbf{q}_k + B_{pert}u_k$ forward in time and collect the resulting data $\mathbf{q}^{pert}$;
  $\nabla_{x_{\mathcal{B}}} J_{PE} = \frac{J_{PE}(\mathbf{q}^{pert}) - J_{PE}(\mathbf{q})}{\alpha}$;
  $x_{\mathcal{B}}^{i+1} = x_{\mathcal{B}}^{i} + \delta \nabla_{x_{\mathcal{B}}} J_{PE}$;
end

4.4 Results

We now apply Algorithm 2 to the enhancement of a DMD-based reduced order model. We use an initial guess for the input location of $x_{\mathcal{B}}^{1} = -20$. This input location results in a data set which stabilizes quickly, resulting in poor data and an inaccurate reduced order model. In Figure 4.5, we see that over the course of 100 iterations of Algorithm 2, the reconstruction error gradually decreases indicating a substantially improved ROM. The converged relative reconstruction error is $e_{recon} = 0.82\%$.

The Ginzburg-Landau system with the converged input and the $\mathbf{q}_{\{1,1\}}$ initial condition was then used to construct the rank 12 state transition matrix $\bar{A}_{PE}$. This reduced order model was then used to compute the data reconstruction of the uncontrolled system with initial condition $\mathbf{q}_{\{1,2\}}$. The result of this reconstruction is given in Figure 4.7. This reconstruction turns out to be a very accurate approximation of the state in Figure 2.3. We conclude that the input derived from the minimization of J_{PE} was enough to greatly enhance the quality and robustness of the resulting DMDc-ROM.

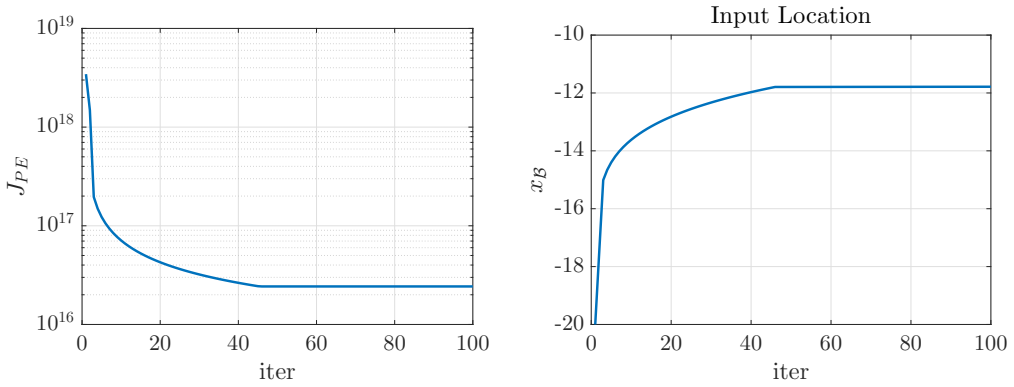

Figure 4.5: (a) The cost J is shown to be decreasing and converging to $J = 2.428 \cdot 10^{16}$. (b) The location x_d of the input converges to $x_d = -11.78$.

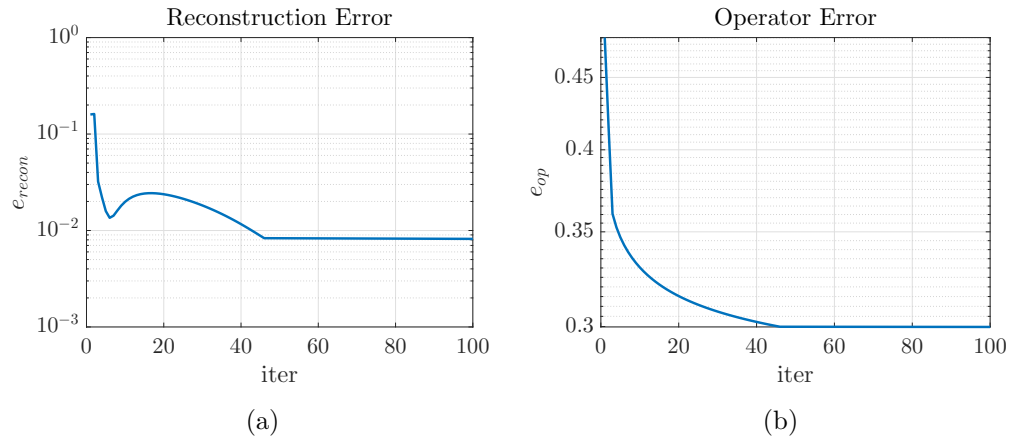

Figure 4.6: (a) The reconstruction error converges to $e_{recon} = 8.189 \cdot 10^{-3}$. (b) The operator error converges to $e_{op} = 0.2998$.

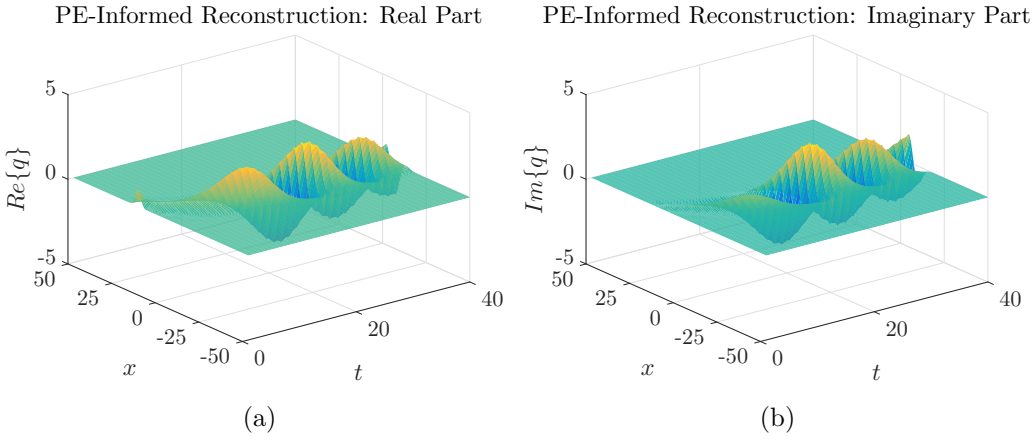

Figure 4.7: The real and imaginary parts of the uncontrolled $\mathbf{q}_{\{1,2\}}$ data reconstructed using PE-informed DMDc.

4.5 OPE Applied to the Stabilization Problem

The motivation for OPE was the construction of robust ROMs which can be used for optimal control. In Section 3.1, we introduced the stabilization problem for the Ginzburg-Landau system. We will now apply OPE to improve the "poor" data from Section 4.2. This realization of the Ginzburg-Landau system has its initial condition at $\mathbf{q}_{\{1,1\}}$ which is to the far left of the domain. Perturbations in this region decay quickly and showcase only a limited range of the dynamics of the system. The resulting DMDc-ROM poorly approximates the gradient of the stabilization problem as shown in Figure 3.1.

An input Gaussian $\mathcal{B}$ with variable location $x_{\mathcal{B}}$ is added to the system as in Section 4.3. As we iterate through the optimally persistent excitation algorithm, the location of the Gaussian changes to better enrich the data. In Figure 4.8, we see the cost J of the stabilization optimal control problem as a function of the number of iterations for a number of choices of $x_{\mathcal{B}}$. As the input Gaussians move closer to the J_{PE} minimizing location of $x_{\mathcal{B}} = -11.78$, the approximation of the gradient is improved. The OPE-informed DMDc-based adjoint method is able to decrease J by 99.4%

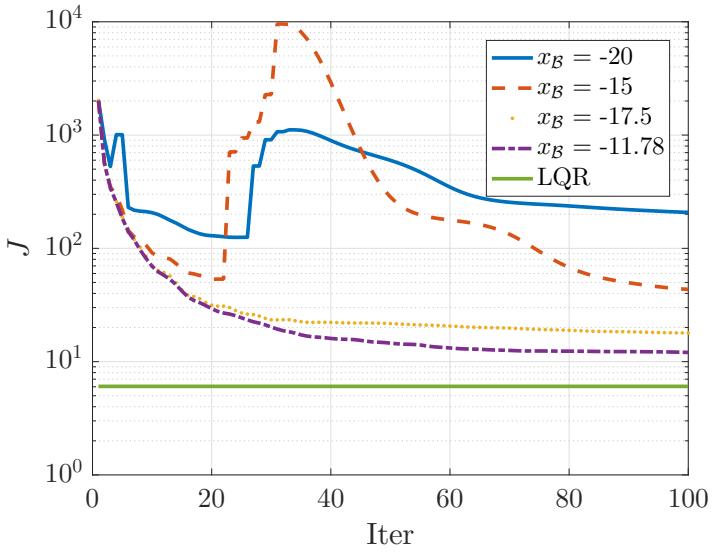
10^4
10^3
10^2
10^1
10^0
J
0
20
40
60
80
100
Iter
$x_{\mathcal{B}}$ = -20
$x_{\mathcal{B}}$ = -15
$x_{\mathcal{B}}$ = -17.5
$x_{\mathcal{B}}$ = -11.78
LQR

Figure 4.8

CHAPTER 5

CONCLUSIONS

This thesis explored the usefulness of DMD in optimal control. In particular, the DMD-based optimal control of the Ginzburg-Landau equation was successfully implemented. Our results show that DMD can be used to compute the adjoint gradient for linear problems and that the resulting optimal control closely resembles that of the full-state optimal control problem.

The problem of finding a data-driven reduced order model which is robust to inputs was also addressed. It was found that control-informed DMDc is a viable way of constructing reduced order models of systems with poorly representative data sets. In order to decide on an appropriate input, the persistency of excitation condition was considered. We chose PE because we were able to prove that PE of the data guarantees the best possible result from DMD. Since it isn't possible to know a priori how to enforce the PE condition on arbitrary systems, a PE-inspired cost functional was proposed which rewards signals which have a high level of excitation. It was found that this cost combined with an appropriate gradient descent algorithm was able to improve the richness of the resulting data. The improved data led to a better data-driven reduced order model.

In the future, we would like to extend this method to nonlinear systems. It is our hope that this work will eventually lead to the construction of robust and controllable reduced order models for two-phase flows. We hope that our method will lead us to the optimal control of spray atomization.

REFERENCES

[Arbabi and Mezić, 2017] Arbabi, H. and Mezić, I. (2017). Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the koopman operator. *SIAM Journal on Applied Dynamical Systems*, 16(4):2096–2126.

[Bai and Sastry, 1985] Bai, E. and Sastry, S. (1985). Persistency of excitation, sufficient richness and parameter convergence in discrete time adaptive control. *Systems & Control Letters*, 6(3):153–163.

[Chen and Rowley, 2011] Chen, K. K. and Rowley, C. W. (2011). H2 optimal actuator and sensor placement in the linearised complex ginzburg–landau system. *Journal of Fluid Mechanics*, 681:241–260.

[Corless, 2003] Corless, M. J. (2003). *Linear Systems and Control*. CRC Press.

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

[Giles and Pierce, 2000] Giles, M. B. and Pierce, N. A. (2000). An introduction to the adjoint approach to design. *Flow, Turbulence and Combustion*, 65(3/4):393–415.

[Holmes et al., 1996] Holmes, P., Lumley, J. L., and Berkooz, G. (1996). *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press.

[Ilak and Rowley, 2008] Ilak, M. and Rowley, C. (2008). Feedback control of transitional channel flow using balanced proper orthogonal decomposition. In *5th AIAA Theoretical Fluid Mechanics Conference*. American Institute of Aeronautics and Astronautics.

[Ioannou and Sun, 2012] Ioannou, P. and Sun, J. (2012). *Robust Adaptive Control*. DOVER PUBN INC.

[Kim et al., 2014] Kim, J., Bodony, D. J., and Freund, J. B. (2014). Adjoint-based control of loud events in a turbulent jet. *Journal of Fluid Mechanics*, 741:28–59.

[Kutz et al., 2014] Kutz, J. N., Brunton, S. L., Luchtenburg, D. M., Rowley, C. W., and Tu, J. H. (2014). On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, 1(2):391–421.

[Moore, 1981] Moore, B. (1981). Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32.

[Proctor et al., 2016] Proctor, J. L., Brunton, S. L., and Kutz, J. N. (2016). Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems*, 15(1):142–161.

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

[Teo et al., 1991] Teo, K. L., Goh, C. J., and Wong, K. H. (1991). *A Unified Computational Approach to Optimal Control Problems*. Longman Sc & Tech.