UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

APPLYING DYNAMIC MODE DECOMPOSITION TO INTERCONNECTED SYSTEMS FOR FORECASTING AND SYSTEM IDENTIFICATION

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

NOAH BRIDGES

Norman, Oklahoma

2023

APPLYING DYNAMIC MODE DECOMPOSITION TO INTERCONNECTED SYSTEMS FOR FORECASTING AND SYSTEM IDENTIFICATION

A THESIS APPROVED FOR THE
GALLOGLY COLLEGE OF ENGINEERING

BY THE COMMITTEE CONSISTING OF

Chair: Dr. Andrés D. González

Dr. Charles Nicholson

Dr. Talayeh Razzaghi

Table of Contents

List Of Tables

List Of Figures

Abstract

Dynamic Mode Decomposition (DMD) describes a family of dynamical systems analysis approaches that approximate complex, likely non-linear behaviors with a low-rank linear operator. DMD has traditionally been used in a systems-identification context and was originally developed as a method of modeling fluid flows using the Koopman Operator. In contrast to these original applications, this work explores DMD's ability to produce high-fidelity forecasts using small training sets in an effort to flexibly model two complex, real-world systems. In particular, a novel, iterative implementation of DMD is tested and validated on 18 years of trading price data for constituent companies of the S&P 500 and on 2 years of per-capita COVID-19 case counts throughout the continental US. The novel combination of DMD with blocked time-series cross-validation described in this work was found to consistently produce forecasts with an average MAPE of approximately 0.1 (in the case of the financial model) and RMSE of approximately 0.2 cases per 1000 citizens (for the COVID-19 model). In addition to reliably predicting the complex behaviors characteristic of real-world systems, this approach was leveraged to identify robust, distinct dynamical trends and construct networks which provided insights into central system elements. This work has illustrated the utility of applying DMD in an iterative approach facilitates forecasting accuracy across a variety of systems without compromising its ability to uncover fundamental characteristics of these underlying systems.

Chapter 1

Introduction

Dynamical systems are ubiquitous elements of modern life. A wide range of phenomena, ranging from the natural to the human engineered, are well-described by these systems. Despite their frequent occurrences in life, these systems can be incredibly complex, making high-fidelity modeling challenging (even with prior knowledge of the functional forms that drive system behavior). Before further describing dynamical systems and their existing modeling approaches, I will first define and introduce notation for dynamical systems (following (1), (2)).

Dynamical systems are characterized by two fundamental objects: a state space and defining function. A dynamical system's state space M (assumed to be non-observable) comprises all possible realizations (or states) of the dynamical system. Since one's understanding of the system is given by observed measurements, $\exists f \in \mathcal{F}$, an arbitrary function space, such that $f : M \to \mathbb{C}$. Here, $\mathbb{C}$ is taken to be a complex-valued observable state space of the dynamical system.

Let the system's location in the state space M at time t be given by $\mathbf{z}_t$. Then the defining function of the dynamical system, T, is defined as $T : M \to M$ according to the relation $T(\mathbf{z}_t) = \mathbf{z}_{t+1}$. In the case that the system is deterministic, a single measurement $\mathbf{z}_t \in M$ and knowledge of the function T fully defines the behavior of the dynamical systems at all times $t \in \mathcal{T}$. If the system is stochastic (that is, subject to noise) then the future states of the system cannot be fully characterized by the two objects discussed above given the variable, unpredictable effect of noise in the system.

Swinging pendulums, evolving population sizes, migration patterns, and fluid flows may all be described using a dynamical systems framework. While this framework has demonstrated its relevance in a variety of contexts, these systems can often become intractably complex using a strictly analytical approach. This complexity has given rise to alternative perspectives on dynamical systems analysis; in this work, I consider an operator-theoretic perspective and its potential to aid in understanding the complex behaviors seen in financial markets and the spread of contagious diseases.

I specifically explore the characteristics of a family of operator approximation techniques known as Dynamic Mode Decomposition (DMD). DMD shares many connections with Koopman Operator Theory, complex analysis, and advanced techniques in linear algebra. DMD has been traditionally been primarily leveraged in a systems identification context, with minimal emphasis being placed on its forecasting fidelity. In contrast, I explore DMD's potential to generate high-fidelity forecasts using small training sets without sacrificing its strengths in the realm of system identification. DMD is particularly suited to this task as it produces linear, decoupled models of non-linear, coupled systems using an equation-free, data-driven framework.

The remainder of this work proceeds as follows. Chapter 2 introduces Koopman Operator Theory, the DMD algorithm and some important variations thereof, as well as relevant research in time-series, financial, and epidemiological modeling. Chapter 3 describes the analytical methods and the data used to study the systems of interest – namely the S&P 500 and COVID-19's US spread. Chapters 4 and 5 include the results and analysis for financial markets and disease spread, respectively. Finally, Chapter 6 completes the work with a summary of the key conclusions and the opportunities for future work.

Chapter 2

Literature Review

2.1 Mathematical Theory

DMD is intimately connected with a variety of deep fields of mathematical research. Among the most significant of these are Koopman Operator Theory, Fourier Analysis (which study how to approximate functions with damped or driven sinusoidal functions), and linear algebra. In the following sections, I explore relevant features of these fields and how they contribute to a fuller knowledge and appreciation for the power of DMD's analytical approach.

2.1.1 Koopman Operator Theory

In 1931, Bernard Koopman (3) proved that the dynamics of any Hamiltonian system (such as an oscillating spring or swinging pendulum) on a Hilbert space (a possibly infinite-dimensional space equipped with some measure of distance) could be fully described by an infinite dimensional linear operator. Using the notation of the discrete-time dynamical system described earlier, one may define this operator (now known as the discrete-time Koopman operator) as:

$$Uf\mathbf{z}_t = f \circ T(\mathbf{z}_t) \tag{2.1}$$

That is, the infinite-dimensional, linear Koopman operator precisely describes the unknown and likely non-linear dynamics that arise directly from the unobserved space M by acting on the function $f : M \to \mathbb{C}$. In recent years with increasing data availability and computational resources, Koopman-adjacent decomposition techniques have arisen for modeling dynamical systems. DMD itself is one such technique, being originally proposed in (4) to obtain temporally-consistent spatial structures in the modeling of fluid flows. Early researchers into DMD (5) noted the strong connections between this technique and Koopman theory.

The success of data-driven modal decomposition techniques are driven by the fact that, in practice, it is unnecessary to fully identify the infinite-dimensional operator U. Rather, U may be well-approximated by a low-rank operator with clearly interpretable spectral components. In particular, as U is an operator on functions, its effects may be described in terms of eigenfunctions and corresponding eigenvalues. In particular, let $\phi_i : M \to \mathbb{C}$ denote an eigenfunction of U. Then,

$$U\phi_i = \lambda_i \phi_i \tag{2.2}$$

Making the (reasonable) assumption that the function f lies within the span of the eigenfunctions of U, it follows that:

$$f(\mathbf{z}_t) = \sum_{i=1}^{\infty} \phi_i(\mathbf{z}_t)\mathbf{v}_j \tag{2.3}$$

In this case, where $\mathbf{z}_t$ is a vector of system measurements, $\mathbf{v}_j$ may be considered as the vector coefficients of the linear combination of eigenfunctions needed to describe f. Following this line of reasoning to its logical end, we see that:

$$Uf(\mathbf{z}_t) = \sum_{i=1}^{\infty} U^t \phi_i(\mathbf{z}_0)\mathbf{v}_j = \sum_{i=1}^{\infty} \lambda_i^t \phi_i(\mathbf{z}_0)\mathbf{v}_j \tag{2.4}$$

That is, at any point in time $t \in \mathcal{T}$, one may describe the observed behavior of the system as a linear combination of eigenfunctions of U whose effect is modulated by repeated applications of their corresponding eigenvalues. Observing too that the operator U essentially describes a discrete-time derivative within this context, its eigenfunctions will be complex-valued exponential functions. Then the growth or decay of a mode will be determined by the Euclidean norm of the eigenvalue while the oscillatory frequency of the mode is described by the imaginary component of the eigenvalue. Describing the Koopman operator in terms of its eigendecomposition is especially convenient as it allows complex, coupled, and non-linear dynamics to be described in terms of simple, decoupled functions whose characteristics are fully determined by the linear operator. For a more comprehensive review of Koopman Operator theory, see (2) (6) (7).

2.1.2 DMD Algorithm and Koopman Theory

To further cement the connections between Koopman Theory and DMD, here I discuss a common implementation of the algorithm and its connections to the Koopman operator. The algorithm described below closely follows the approach described in chapter 1 of (8).

To begin, Let $\mathbf{X}$ be a matrix containing the observed measurements of a dynamical system. Presume that each row corresponds to a specific system element (e.g., a specific point in space) and each column represents a point in time. Then,

$$\mathbf{X} = \begin{bmatrix} | & | & | & ... & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & ... & \mathbf{x}_m \\ | & | & | & ... & | \end{bmatrix}$$

The matrix X is then split into two overlapping matrices – in essence, an input set of measurements and a set of measurements to predict, as described below:

$$\mathbf{Y} = \begin{bmatrix} | & | & | & ... & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & ... & \mathbf{x}_{m-1} \\ | & | & | & ... & | \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} | & | & | & ... & | \\ \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 & ... & \mathbf{x}_m \\ | & | & | & ... & | \end{bmatrix}$$

Note that this offset need not be a single time step; rather, an approximation may be generated for an arbitrary number of steps into the future. Regardless of the time-difference separating these matrices, we seek to identify the optimal operator A that transforms $\mathbf{Y}$ into $\mathbf{Z}$, such that:

$$\mathbf{Z} = A\mathbf{Y} \implies A = \mathbf{Z}\mathbf{Y}^{\dagger} \tag{2.5}$$

Where $\mathbf{Y}^{\dagger}$ denotes the Moore-Penrose pseudoinverse of the matrix. As an aside, under this characterization, A may be described as a best-fit linear operator minimizing $\|Z - AY\|_2$. While this characterization of DMD as essentially high-dimensional regression provides greater flexibility in the types of systems one may model, in this work I focus on the operator-theoretic interpretation of A, where A approximates the real

dynamics of the system. Returning to the algorithm, a singular value decomposition (SVD) is performed on $\mathbf{Y}$, yielding:

$$\mathbf{Y} = \Psi \Sigma V^*$$

Suppose that $\mathbf{Y}$ be an n x m-1 matrix. Then Ψ is an n x n unitary matrix whose columns denote "directions" in the observed state-space with maximum variation. Σ is a diagonal n x m matrix with $\text{Min}\{n, m-1\}$ singular values arranged in descending order. Finally, V^* is an m-1 x m-1 unitary matrix, whose rows describe the relative contribution of the left singular vectors to each measurement in $\mathbf{Y}$. Since many dynamical systems display low-rank characteristics, a truncated SVD is performed, retaining only r columns of Ψ, the r largest singular values, and the first r rows of V^* (one robust approach to judiciously selecting r is found in (9)). Then, based on the SVD and the operator's definition in (2.5), the reduced-rank DMD operator is precisely:

$$\tilde{A} = V X_2 \Sigma^{-1} U^* \tag{2.6}$$

Next, one takes the eigendecomposition of $\tilde{A}$, revealing the set of r eigenvectors $\mathbb{W}$ and corresponding eigenvalues Λ such that $\tilde{A}\mathbb{W} = \Lambda\mathbb{W}$. The r eigenvalues approximate the eigenvalues of the Koopman operator for the discrete-time dynamical system from which they are derived. The discrete-time eigenvalues Λ may be converted to corresponding continuous-time eigenvalues, Ω, via $\Omega = \frac{ln(\Lambda)}{\Delta t}$ to aid in system interpretability and reduce computational complexity. Note that this transformation assumes a constant Δt, the difference in time between two subsequent measurements. The eigenvalue λ_i fully defines the eigenfunction $\phi_i = e^{\lambda_i}$.

To obtain the "dynamic modes" (which approximate the Koopman modes) the eigenvectors of the reduced-order $\tilde{A}$ must be broadcast back to the original state space. These, following (10), are precisely:

$$\Phi = \mathbf{Z}V\Sigma^{-1}\mathbb{W} \tag{2.7}$$

The original state of the system (taken at time $t = 0$) is obtained by computing the vector of modal amplitudes, b, via the relation:

$$x_1 = \Phi b \implies b = x_1\Phi^{\dagger} \tag{2.8}$$

where, since Φ may not be directly invertible, $\Phi^{\dagger}$ denotes the pseudoinverse. Having recovered the modes, amplitude vector, and continuous-time eigenvalues of the system, the DMD approximation of the system in the observable space at time t is computed by:

$$\hat{\mathbf{x}}_t = \Phi e^{\Omega t} b \tag{2.9}$$

Then, the projection over all times $t \in \mathcal{T}$ is:

$$X_{DMD} = \begin{bmatrix} | & | & | & ... & | \\ \hat{\mathbf{x}}_1 & \hat{\mathbf{x}}_2 & \hat{\mathbf{x}}_3 & ... & \hat{\mathbf{x}}_m \\ | & | & | & ... & | \end{bmatrix}$$

To cement the connection of DMD to Koopman theory, I follow the work of (5). In particular, consider a matrix of vector-valued observables from a dynamical system,

$\mathbf{X}$, as defined previously. Then following the DMD algorithm defined above, we can approximate $\hat{\mathbf{x}}_t$ as:

$$\hat{\mathbf{x}}_t = \sum_{i=1}^{r} \lambda_i^t \Phi_i b_i + \epsilon \tag{2.10}$$

The above expression is simply a summation expansion of the matrix computation given in equation (2.9), where Φ_i denotes the i^{th} column of Φ and ϵ denotes the error associated with this projection. Observe that the error is necessarily outside the span of Φ, so $\epsilon \perp \text{Span}(\Phi)$. (5) note that in the case that $\epsilon = 0$ the modes and eigenvalues produced via DMD are exactly a finite number of eigenvalue - eigenfunction pairs from the Koopman operator. In the case that $\epsilon \neq 0$, then the eigenvalue - eigenfunction pairs are the best possible approximations, in the least-squares sense, that can be derived from the given input data.

Before exploring approaches to augment the efficacy of the DMD algorithm, I briefly discuss the interpretation (within a systems-oriented context) of the dynamic modes of a system. By assumption, each row of the data matrix corresponds to a specific element of (or location in) the system. If $\mathbf{X}$ contains m rows, each dynamic mode will contain m elements corresponding to each row. The full DMD forecast value at time $t \in \mathcal{T}$ is a linear combination of modes (which are scaled by eigenvalues, raised to a time-varying power). Additionally, these modes are independent (or decoupled) from each other. Thus, one can interpret these modes as describing a set of independent relationships, or meaningful structures, among system elements which form a low-rank "eigenbasis" for the observable state-space. This interpretation key dynamical relationships to be automatically identified while reducing modeling complexity. In this way, the dynamic modes of the system share similarities to the "components" identified by Principal Components Analysis (PCA), in that the modes allow each

system measurement to be described with a reduced-dimension representation. Unlike PCA, classifying modal importance is a more ambiguous than identifying the most critical principal components.

2.1.2.1 Identifying Dominant Modes

In the literature, the importance of DMD modes has been traditionally analyzed at only a single point in time (when $t = 0$) by calculating the DMD analog of the Fast Fourier Transform (FFT) power spectrum. This power spectrum is computed as:

$$DMD_{pow} = |b| * \frac{2}{\sqrt{s}} \tag{2.11}$$

Where s is the number of times the signal is delay-embedded (discussed in greater detail below). In the case that the raw data is used, $s = 1$. However, this power spectrum only describes the relative contribution of modes at time $t = 0$, and neglects the role of oscillatory behavior and the evolving interactions between the columns of Φ through time. In an effort to more faithfully capture the intricate interactions between DMD triplets (modes, amplitudes, and eigenvalues), (11) propose an alternative analysis approach that yields more insightful visualizations and facilitates greater understanding of global system behaviors.

At the core of their approach is an alternative algorithm for analyzing the dominance structure of the DMD modes. Before discussing their specific approach, recall that the contribution of each mode at time $t \in \mathcal{T}$ is given by

$$\lambda_i^t \Phi_i b_i$$

That is, the mode's contribution depends not only on the initial amplitude, b_i, but also on the eigenvalue and the particular time step t. While there may be cases where

the impact of the eigenvalue and time-step are minimal, in cases where pronounced periodicity or growth or decay exist, the dynamic behavior described by the eigenvalue will substantively effect the mode's contribution to the overall forecast. To take a holistic view of each mode's relative contribution through time, the authors propose taking a subset of times, $T \subset \mathcal{T}$, and then for each mode i and each $t \in T$, evaluating:

$$\|\lambda_i^t \Phi_i b_i\|$$

These norms may then be plotted versus the modal frequency to provide insight into which frequencies are most dominant through time (as well as identify any persistently dominant modes). The authors also recommend two clustering approaches to identify "similar" modes, one based on eigenvalue proximity (denoting similar dynamic behavior), and the other on harmonic multiplicities (where the frequency of modal oscillation is related via integer multiplication).

While (11) provide insights into the analysis of the systemic effects of modal behavior, they largely overlook the spatial relationships depicted within the modes themselves. As is noted in (12), modes automatically identify latent relationships between system elements. These come in two forms, one related to the magnitude of the element's measurement and the other related to the "phase", or how "in-sync" the elements are. These two components of similarity can be readily visualized in a variety of methods (e.g., choropleth maps for geospatial data) to more quickly assess the fundamental structures of the system.

2.1.3 Koopman-Adjacent Techniques

While the operator-theoretic approach that forms the foundation of Koopman analysis provides a flexible framework through which to analyze dynamical systems, its efficacy

and interpretability may be augmented by leveraging insights from complex analysis and linear algebra. In particular, I describe some of the connections between methods of approximating functions via complex-valued exponential functions (like the FFT and variations of Prony's method) and properties of Hankel, Toeplitz, and Vandermonde systems to DMD to facilitate modeling techniques for more complex, noisy systems.

2.1.3.1 Function Approximation

Methods to approximate functions with linear combinations of complex-valued exponential functions first arose in mathematical analysis in the late 18th century with the work of Gasparde Priche de Prony. While Joseph Fourier's alternative methodology for decomposing functions into their constituent signals is far more influential (in part because Prony's method requires a digital computer to be feasible), insights drawn from the approaches of both researchers may enhance the capability and interpretability of DMD-adjacent models.

For example, a primary assumption of the SVD-based DMD algorithm is that the data is highly sampled along the spatial dimension (that is, the data matrix is "tall and skinny"). DMD's capability to identify robust, coherent spatial structures from matrices with sparse spatial sampling may be compromised, and many real systems of interest may be composed of few elements, limiting the spatial dimension of the data. While I will discuss alternative data processing techniques to mitigate this challenge, (13) used a vector-valued version of Prony's method to implement Koopman Mode Decomposition for data with low spatial dimension. The researchers utilized a variation of Prony's method rather than the Arnoldi-like DMD algorithm (the SVD-based approach, while similar, is more numerically stable than the Arnoldi approach) to find the best linear combination of damped and driven sinusoids to model the observed behavior of European power grids. They note that this approach is well-fitted for data

with heavy temporal sampling (where $n > 2m$, resulting in a "short and fat" matrix), as it provides a unique decomposition where some DMD approaches may not.

Prony's original technique for function approximation is prone to significant numerical errors when used to model stochastic systems. Consequently, many algorithmic variations with greater robustness to noise have been developed. While these techniques are most often related to scalar-valued rather than vector-valued functions, the principles required for developing high-fidelity, minimal rank approximations remain consistent. For example, (14) describe an approach for identifying the minimum number of terms required to approximate an arbitrary smooth function with complex exponential functions using the SVD of a Hankel matrix. Additional insights into Prony- and Fourier-based approximation methods may be found in (15) (16) (17).

2.1.3.2 Hankel and Vandermonde Systems

While one may adopt alternative Prony-derived modeling approaches to address low levels of spatial sampling, an alternative approach - known as *delay embedding* within time-series analysis - can address these issues. Delay embedding draws on characteristics of Hankel matrices. Given vector-valued system states comprising $\mathbf{X}$, where

$$\mathbf{X} = \begin{bmatrix} | & | & | & ... & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & ... & \mathbf{x}_n \\ | & | & | & ... & | \end{bmatrix}$$

The data may be delay-embedded to form an augmented data matrix, $\mathbf{X}_{aug}$ defined below:

$$\mathbf{X}_{aug} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & ... & \mathbf{x}_{n-j} \\ \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 & ... & \mathbf{x}_{n-j+1} \\ \vdots & \vdots & \vdots & ... & \vdots \\ \mathbf{x}_j & \mathbf{x}_{j+1} & \mathbf{x}_{j+2} & ... & \mathbf{x}_n \end{bmatrix} \tag{2.12}$$

This matrix shares the form (albeit in a vectorized fashion) of a Hankel matrix, where each off-diagonal is a constant value (in this case a constant vector). This data transformation technique dramatically increases the spatial dimensionality of the data, enabling an SVD-based DMD algorithm to more faithfully describe the behaviors observed within a sparsely-observed system. While this technique has been empirically shown to improve the modeling capability of DMD, the form of $\mathbf{X}_{aug}$, as a vectorized Hankel matrix, implies that this technique carries deeper connections than simply increasing the spatial dimensional of a data set.

Interestingly, Hankel matrices frequently arise in computing sparse approximations of functions with complex-valued exponentials. Indeed, (14) create a Hankel matrix from oversampled data points to estimate the number of terms needed to approximate a scalar-valued function within a desired accuracy. Their use of Hankel matrices aligns with the fundamental research of Hankel matrices described in AAK Theory (named for researchers Adamyan, Arov, and Krein's result for approximating infinite-dimensional, but finite-rank, Hankel operators). The core insight arising from AAK theory (which is leveraged in the approximation of functions) is the following:

Theorem 1 *Let Γ be an infinite-dimensional Hankel matrix, with singular values arranged in descending order $\{\sigma_1, \sigma_2, ..., \sigma_n, ...\}$. Then there exists a rank-n Hankel matrix K, such that $\sigma_n = \|\Gamma - K\|$. (18)*

This result is leveraged to find approximations of fixed size whose worst-case deviance from observed behaviors is tightly controlled. Indeed, (1) have proposed an extension of DMD termed *Hankel-DMD* with close connections to a Prony-approximation of Koopman Mode Decomposition. They demonstrated that applying the DMD algorithm to an embedding of (potentially vector-valued) time-series data could produce higher fidelity results than traditional methods and provided a strong mathematical framework for creating such models. A core assumption of their mathematical foundation is that the observables lie on a finite-dimensional subspace that is invariant under a Koopman operator U. They suggest implementing this assumption by assigning a hard threshold for the SVD truncation in the DMD algorithm to ensure a desired numerical accuracy. That is, one discards all singular values less than some cutoff value.

While Hankel matrices are leveraged to improve function approximation techniques, they have also been used in other data-driven modeling techniques like Singular Spectrum Analysis (SSA) (19) (20). The authors of (19) in particular found that they were able to more clearly distinguish distinct signals within their (scalar-valued) data when the Hankel matrices were near-square. Hankel matrices are also a critical feature in stochastic modeling approaches like Hidden Markov Models (HMMs) which share clear similarities with the operator-theoretic background of DMD. For more a more thorough review of the properties of Hankel matrices can be found in (21) (22).

Before considering common modeling approaches in the fields of financial and epidemiological modeling, I briefly discuss Vandermonde systems, which share connections both with Koopman theory and with function approximation techniques more generally. In general, one defines a Vandermonde matrix, V, as:

$$V = \begin{bmatrix} 1 & x_0 & x_0^2 & ... & x_0^n \\ 1 & x_1 & x_1^2 & ... & x_1^n \\ \vdots & \vdots & \vdots & ... & \vdots \\ 1 & x_m & x_m^2 & ... & x_m^n \end{bmatrix} \tag{2.13}$$

The sharp similarities with the general dynamical system defined earlier are clear if one considers that a data matrix, $\mathbf{X}$ may be defined as:

$$\mathbf{X} = \begin{bmatrix} f_1(\mathbf{z}_0) & f_1 \circ T(\mathbf{z}_0) & f_1 \circ T^2(\mathbf{z}_0) & ... & f_1 \circ T^n(\mathbf{z}_0) \\ f_2(\mathbf{z}_0) & f_2 \circ T(\mathbf{z}_0) & f_2 \circ T^2(\mathbf{z}_0) & ... & f_2 \circ T^n(\mathbf{z}_0) \\ \vdots & \vdots & \vdots & ... & \vdots \\ f_m(\mathbf{z}_0) & f_m \circ T(\mathbf{z}_0) & f_m \circ T^2(\mathbf{z}_0) & ... & f_m \circ T^n(\mathbf{z}_0) \end{bmatrix}$$

Vandermonde matrices are intimately connected with Discrete- and Fast Fourier Transforms, allowing for their efficient computation. A (possibly overdetermined) Vandermonde matrix is also used by (14) as one step of their algorithmic approach to creating minimal approximations of functions using complex-valued exponential functions. While I do not leverage specific properties of the Vandermonde matrix in this work, close connections to the Discrete Fourier Transform and function approximation more generally point to the rigorous connections between DMD and advanced algebraic tools.

2.2 Example Application of DMD Algorithm

Finally, before exploring the domain-specific literature in both finance and epidemiology, I first follow the basic example (found in Chapter 1 of (8)) of DMD's capability of identifying distinct functions and provide concrete examples into the interpretation of the dynamic modes and eigenvalues.

For this example, consider $f(x,t) = g(x,t) + h(x,t)$ as a linear combination of two functions with two input variables - x representing a spatial dimension and t representing the temporal dimension. In particular then, $g(x,t) = \frac{1}{cosh(x+3)} * e^{2.3i*t}$ and $f(x,t) = \frac{1}{cosh(x)} * e^{2.8i*t}$. Ignoring the complex component of this function, the composite resulting linear combination (that is, $f(x,t)$ is shown in the figure below.

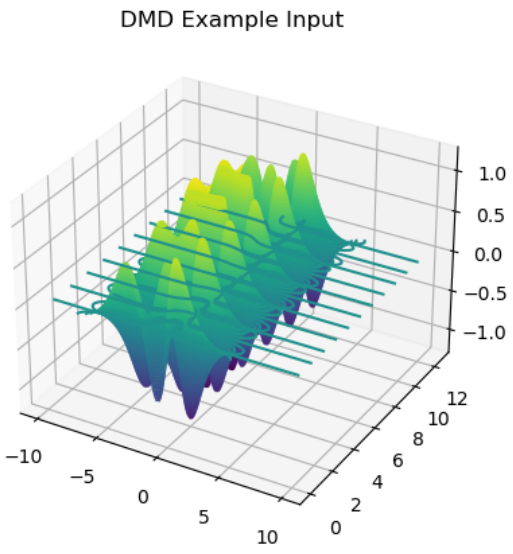
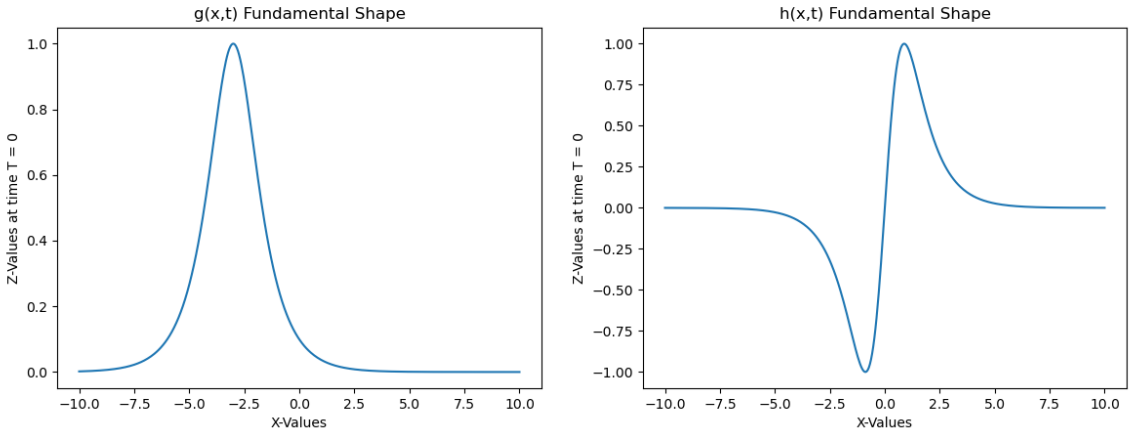


Figure 2.1: Input Function Visualization

Before advancing to the DMD outputs for this particular input, I consider the fundamental structures of the two input functions, $g(x,t)$ and $f(x,t)$. These are both visualized below.



(a) $g(x,t)$ Fundamental Structure (b) $h(x,t)$ Fundamental Structure

Figure 2.2: Decoupled Data Structures

Combining all of this information, it is clear that the input function displays oscillatory behavior in the temporal dimension while retaining distinct structures along the spatial dimension. In particular, it is important to observe that $g(x,t)$ will oscillate 2.3 times for every 2π units of time. Similarly, $h(x,t)$ will oscillate 2.8 times every 2π units of time. Having identified these core characteristics, consider the results from a rank 2 DMD analysis of this input data.

First, the ability of the DMD algorithm to create a high-fidelity rendering of the input data is remarkable. The figure below displaying the DMD rendering of the input data is indistinguishable from the input data.

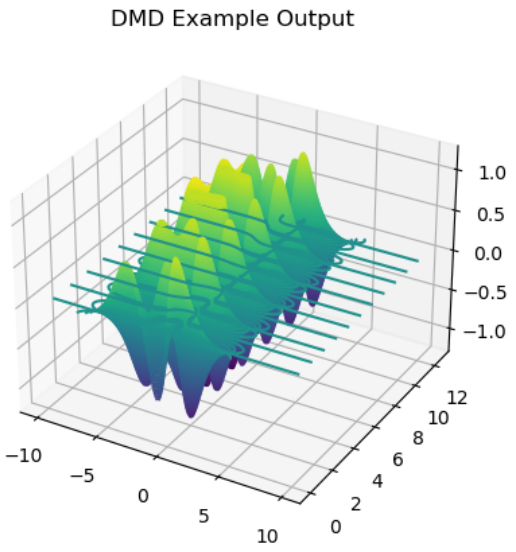
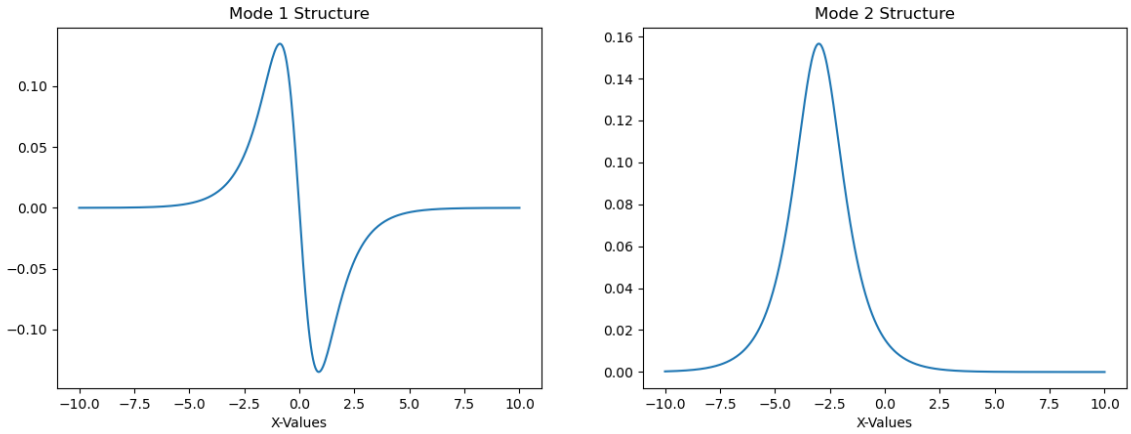


Figure 2.3: DMD Example Forecast

Equally remarkable is DMD's ability to describe the functions and structures responsible for the observed behavior. Consider that the continuous-time eigenvalues returned from this analysis are precisely $\omega_1 = 0 + 2.8i$ and $\omega_2 = 0 + 2.3i$ – exactly corresponding to the two frequencies of the input functions. Further, consider the structures revealed by the modes.



(a) Mode 1 Fundamental Structure (b) Mode 2 Fundamental Structure

Figure 2.4: DMD Example Resulting Mode Structures

While these structures do not align perfectly with the initial spatial conditions of the input data (the first mode is the mirror image of $h(x,t)$), they clearly depict the underlying structure of the data. Additionally, the corresponding eigenvalues precisely align with the underlying frequencies which accompany these structures. The images above also provide a strong intuition for the interpretation of modes – they depict how, when aligned along a meaningful spatial axis, robust structures and relationships between elements may be identified.

2.3 Domain-Specific Research

While my research is primarily focused on applying dynamical systems theory and analysis to financial modeling and epidemiology, this work would be incomplete without an overview of existing techniques within these fields. In the subsequent sections I outline some of the dominant modeling approaches in both these fields, introducing recent applications of DMD or related methods, and highlight gaps in the research that my work will begin to address.

2.3.1 Financial Markets Modeling

While the modeling approaches utilized by financial analysts are diverse, most adhere to traditional statistical methods (e.g., ARIMA or autoregressive conditionally heteroscedastic (ARCH) models) (23),(24) or computational machine-learning algorithms (like artificial neural networks (ANNs) (25). While ARCH and ARIMA models can account for market volatility, they make assumptions regarding the functional behavior of the market. For example, ARIMA models may struggle to account for seasonality or a lack of stationarity in the data (24). Meanwhile, ANNs make no assumptions regarding the functional form of the system, but offer little insight into the functional behavior that guides stock price evolution through time. In recent years, few researchers are pursuing models that strictly adhere to these classic delineations between models. Many approaches, including wavelet-based decomposition methods (26), stochastic and agent-based models (27), and physics-informed Brownian motion models (24), have been tested individually or combined with more traditional techniques to increase the fidelity of their predictions.

While both statistical models and machine-learning methods have had success in modeling observed market behavior, no technique has achieved supremacy over another. In their review of popular financial modeling approaches, (24) observe that researchers have claimed that both ARIMA and ANNs are, under certain circumstances, superior to the other. (24) ultimately found that Brownian motion and ARIMA models both outperformed an ANN in forecasting the price of the S&P 500 Index. While financial modeling has traditionally adhered to either a physics-informed, computational, and statistical approach, many modern approaches integrate tools from each of these areas to produce more accurate forecasts. For example, (28) created a hybrid model using ARIMA and an ANN that outperformed either model individually. Other research groups have implemented ensemble models using wavelet decomposition, ARIMA, and

ANNs to produce models for exchange rates (29) and the DOW-Jones Industrial Average (30). Additionally, a variety of time-series analysis techniques (besides ARIMA and finance-specific ARCH models) have been used to model financial markets. The flexible, time-series approach SSA has been leveraged to forecast, smooth, and decompose financial time-series data (31). Given its non-parametric, data-driven approach, this approach shares several modeling similarities to DMD which augment its utility.

In contrast to these other approaches, (32) used Dynamic Mode Decomposition (DMD) to study the long-term cyclic and periodic behaviors within financial markets. While wavelet-based analyses of financial markets are not a new phenomenon (26), these are often incorporated as a data-processing step to account for non-stationary data (29) (30). In contrast, (32) used DMD to identify several robust modes whose frequencies range from one to three cycles per year. Other researchers have also utilized DMD in their analysis of stock trends. (33) used DMD to identify stock trading strategies that could, at certain times and in specific sectors, generate returns greater than the "buy and hold" strategy. While the existing applications of DMD to financial data point to its potential to provide new insights, neither approach explores whether there exist fundamental relationships between companies or sectors. Further, neither research group explored DMD's capability of generating accurate forecasts without sacrificing system identification.

2.3.2 Epidemiological Modeling

I now briefly review two of the dominant paradigms within epidemiological modeling: compartmental models - like the Susceptible-Infected-Recovered (SIR) model proposed in 1927 (34) - and time-series techniques. The compartmental model framework (epitomized by SIR) take a population-level view of disease spread. In the SIR model

for example, each individual in the population is either susceptible to infection, currently infected, or recovered from infection. Compartmental models may be based on stochastic or deterministic differential equations, with deterministic models frequently being sufficient for modeling large-scale disease dynamics (35). Compartmental models have been extended to include more disease states and account for disease-mitigation measures. (36) found that model structure (i.e., the number of compartments) can dramatically alter the projected effect of interventions; they urge discretion in determining an appropriate degree of model complexity given the dynamics one is attempting to describe. (37) provide a control theoretic perspective on compartmental models, ascertaining whether model parameters can be identified from inputs and whether states may be observed from inputs or outputs (in short, whether distinct inputs yield unique outputs). The authors observe that allowing parameters to vary in time can increase model observability and emphasize the importance of restricting model use cases to their identifiable parameters and observable states.

In contrast to compartmental models which are constructed on epidemiological explanations for observed behavior, time-series adjacent are generally concerned with predicting behavior and (occasionally) decomposing signals into components. (38) compared several time-series analysis techniques for predicting the spread of a variety of viral diseases with varying success. (39) related compartmental models to time-series analyses by discretizing the continuous SIR models to produce a high-fidelity representation of measles cases from 1944-1964 in Wales. (40) describe a wavelet-based decomposition method for describing time-varying relationships in non-stationary epidemiological data sets. DMD shares the dynamical systems perspective of many classical epidemiological approaches, but the specific methodology shares more in common with discrete time-series modeling approaches.

While extensive applications of DMD to epidemiology remain rare, in (12) the researchers explored the potential of DMD to explain infectious disease spread using three case studies: Google Flu trends, Measles in England, and Polio in Nigeria. They primarily leveraged the tool to identify geographic relationships in the evolution of the disease rather than focusing on the capability of their model to closely model recorded disease burdens. That is, while DMD was demonstrated to draw out latent relationships among distinct regions (a system identification application), it's ability to accurately predict future disease states - the standard for compartmental modeling success - was overlooked. Connecting these distinct modeling capabilities of DMD is critical in demonstrating its full potential to enhance existing modeling approaches.

Chapter 3

Data & Methods

Having established the mathematical foundation of the DMD algorithm and overviewed existing techniques within financial and epidemiological modeling, I now address the specific data sets and methods used for my analysis in this work.

3.1 Data Description

I used one data set for each of the application areas in this work. The first data set, for the study of financial markets, is comprised of daily trading prices of S&P 500 constituent companies. The second data set, for studying the spread of COVID-19, contains daily state-level COVID-19 case-count data.

3.1.1 Financial Data

The data for S&P 500 constituents was obtained from Yahoo Finance's historical data. Every company in the S&P 500 as of September, 2022 was originally included in the analysis. As this data was too sparse, compromising the efficacy of the algorithm, companies with initial public offerings after August 12, 2004 were excluded. At the same time, data from prior to August 12, 2004 was excluded. Thus, the data is composed of the trading prices (measured at close of day) for 405 companies, beginning August 12, 2004 and ending on September 7, 2022.

3.1.1.1 Data Overview

As the data were collected from a reliable source and filtered to eliminate missingness, two common data problems (veracity and missingness) were eliminated. However, the trading prices of some companies were separated by orders of magnitude, a data artifact that could compromise the numerical stability of the Singular Value Decomposition, thereby impacting the DMD forecasts. Consider the price discrepancies between the average lowest priced stock (Ford Motor Company), with an average price of approximately $11, with the highest priced stock (on average) NVR, with an average price of approximately $1,700.

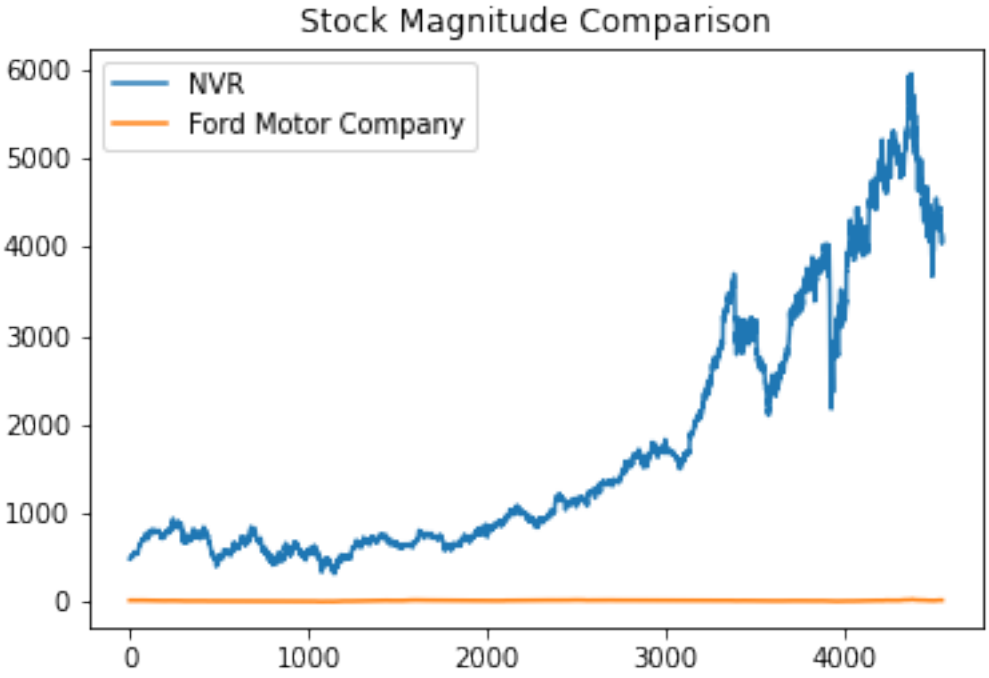


Figure 3.1: Magnitude of Stock Price Differentials

To ameliorate the numerical instability resulting from these significant differentials, the natural log was applied to the data, dramatically reducing the spread between the highest and lowest priced stocks. This transformation is commonly applied to modeling

financial data and has the additional benefit of being easily interpretable and invertible. Thus, when analyzing this data set, the DMD algorithm is applied to log-transformed data, but I assess the forecast accuracy based on the true stock price.

Finally, I considered the long-term trends in the data. While there is clearly some noise and stochasticity within the data, the overall behavior is relatively smooth with a fairly consistent growth trend, with the most pronounced dips for the 2008 financial crisis and COVID crash in 2020.

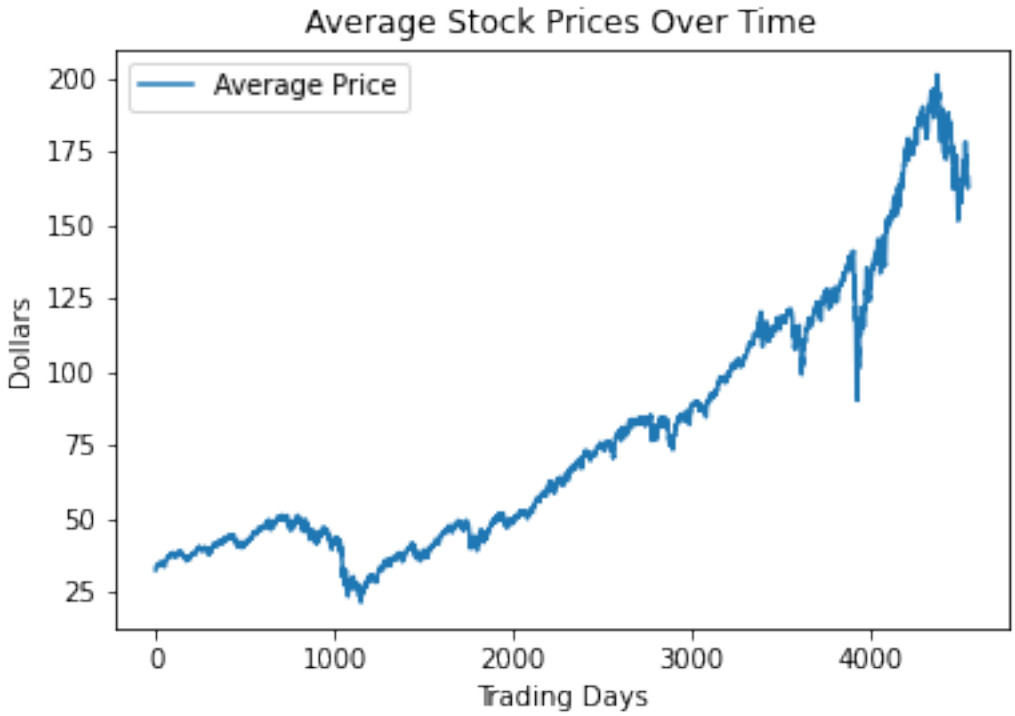


Figure 3.2: Average Daily Stock Price

3.1.2 COVID-19 Data

Unlike the data for the S&P 500 constituent companies, the data describing the spread of COVID-19 through the US is, even when collected from the best-available sources, somewhat unreliable and incredibly noisy. The data were obtained from a publicly

available repository, maintained by Johns Hopkins University, with daily state-level COVID-19 case and death counts. The data were transformed to record the incidence rate on a per-1000 citizen basis to facilitate one-to-one comparisons between states. Data from the continental 48 US states and the District Columbia were included. Evolving reporting procedures, limited availability of testing (especially early in the pandemic), and changing definitions of positive cases all contribute to the significant noise within the data, calling into question its veracity. The raw per-capita case counts (for the entire continental US) are visualized below.

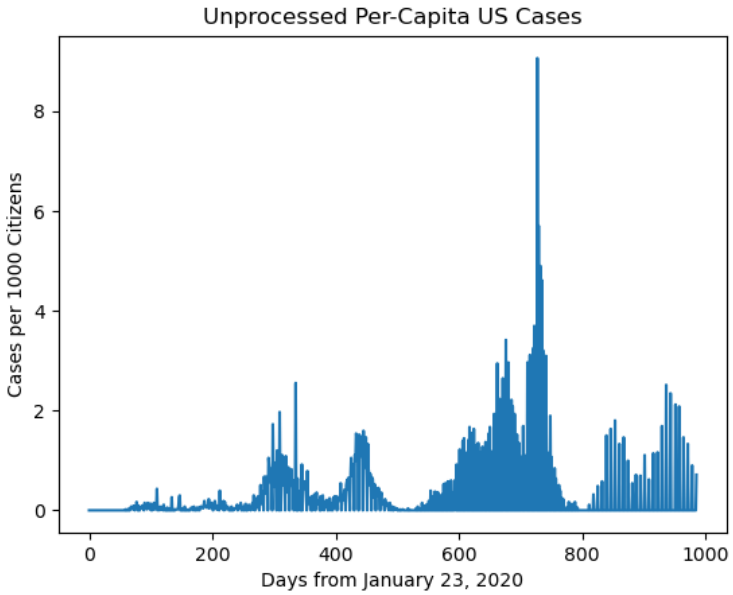


Figure 3.3: Unprocessed US Daily Cases per 1000 People

3.1.2.1 Data Cleaning and Exploration

As an artifact of local reporting procedures, case counts were occasionally reported as negative; in these cases, the number was corrected to zero, an incorrect (but better) estimation of real system behavior. To further ameliorate the effect of noise and smooth the data, a 7-day moving average was computed and used for analysis. Additionally,

data collected prior to April 1, 2020 and after May 31, 2022 were excluded as data quality deteriorated significantly outside of these dates. The effect of these processing is shown below:

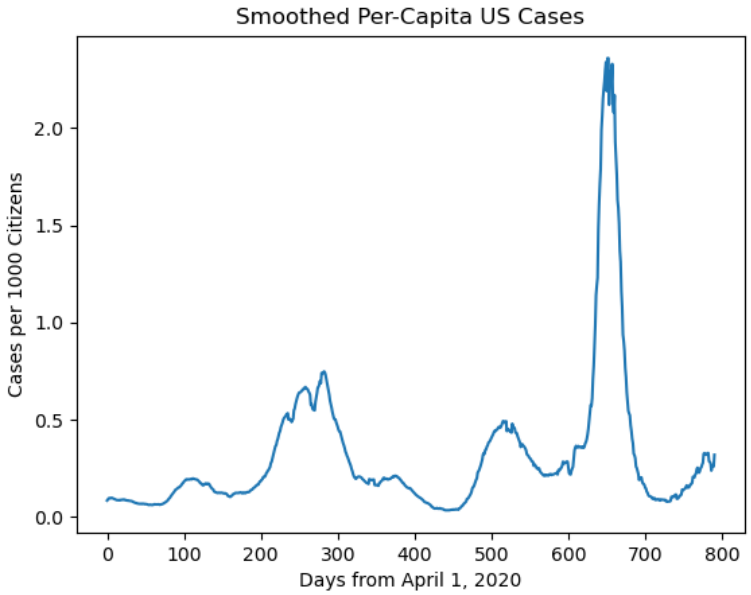


Figure 3.4: 7-Day Average of US Cases per 1000 People

As can be seen in the figure above, there are three main peaks in the data. The first corresponds to the winter of 2020, the second small peak is the rise of the 'Delta' variant, and the largest peak corresponds to the rapid spread of the highly contagious 'Omicron' variant. Preliminary tests of the DMD's capability to model the observed behavior were largely unsuccessful (given the significant swings and volatility of the disease burden), necessitating augmentations to the SVD-based DMD. Specifically, the data set was delay embedded (with the number of embeddings being empirically determined). While this processing step increases both the computational complexity (a minor challenge in this case) and the amount of data required to identify the underlying dynamics, the improvements in forecast accuracy far outweigh these costs.

3.2 Methodological Approach

Despite the differences in the systems from which the data are collected, the core features of the data remain unchanged. The data are taken from complex, interconnected systems over an extended period of time. They contain measurements from distinct system elements, and compared with the number of temporal measures, the number of 'spatial' measures is comparatively small. Given the similarities in the core characteristics of the data, the same analytical approach was applied to each data set. The analysis has two primary objectives. First, to validate the veracity of the functional approximations of the model, I propose an iterative implementation of DMD that consistently generates a high-fidelity, cross-validated, forecasts. Second, I identify and analyze the most significant eigenvalues and dynamic modes of the system to more fully characterize latent dynamics and network structures of the system. In the following sections I discuss the process to accomplish each of these objectives.

3.2.1 Model Creation and Validation

Both financial markets and the spread of diseases are highly sensitive to external events – for example, investor risk tolerance or shelter-in-place orders, respectively. This sensitivity gives rise to a reasonable hypothesis that the system's governing dynamics evolve quickly through time. Thus, rather than training the models once on most or all of the data, the models were trained on small subsets of data and evaluated with blocked time-series cross-validation (TSCV).

Blocked TSCV holds the size of the model's training set constant and, based on the training set, forecasts future measurements at a future time (or times). Having trained and evaluated a model on a particular training set, the training set "slides" forward

in time (while remaining the same size) and the process repeats until the full data set has been leveraged.

I used the SVD-based DMD to create the model for each training set, given its straightforward implementation and numerical stability. Blocked TSCV requires two parameters to be tuned: the size of the training set and the model complexity (that is, the rank of the SVD truncation in the DMD algorithm). These parameters are tuned according to the cross-validated (future) forecasts of the model, rather than its fidelity of modeling the training set. I considered three future points for model cross-validation – one in the 'near'-, middle-, and 'far' future. I used mean absolute percentage error (MAPE) to evaluate the accuracy of the financial model and root mean square error (RMSE) for the epidemiological models. The modeling specifics are given in the table below.

Data Set	Training Size	Model Sizes	Forecast Points	Sliding Step
S&P 500	50 - 350 Days (by 10s)	10-40 Modes (by 5s)	5, 20, and 50 Days	20 Days
COVID-19	21-56 Days (by 7s)	5-15 Modes (by 1s)	7, 14, and 28 Days	14 Days

Table 3.1: Modeling Specifics

The process of selecting the best combination of training set size and model size is described below in pseudo code.

```
TrainingSize = [...]  # list of the training sizes
ModelSizes = [...]  # list of model sizes
    for i in TrainingSize:
        numRuns = setRuns(i)  # set the number of runs
        for 1 <= j <= numRuns:
            # for each run, create training / test sets
            tempSet = data(i, j)
            testSet = data(i, j)
            for k in ModelSizes:
                # create a forecast
                forecast = DMD(tempSet, modelSize)
                # assess validity of forecast
                Validate(forecast)
return best(dataSet, TrainingSize, modelSize)
```

While the process described was followed for both data sets, effectively modeling the COVID-19 data required tuning one additional parameter: the number of delay-embeddings. As noted in the overview of the COVID-19 data set, the data was noisy and preliminary testing indicated that the basic SVD-based DMD would be incapable of describing the observed behavior. While delay embedding the data resulted in far superior forecasts, tuning the number of embeddings required empirical testing. Thus, when evaluating the COVID-19 data set, I also considered the number of delay embeddings as an additional parameter. To minimize data loss and additional computational complexity, a maximum of 10 embeddings was permitted.

Once the best combination of model parameters, I assessed how the forecast errors related to the level of turbulence in the system. For the financial markets, particular

attention is given to the model's ability to adjust for periods of volatility (e.g., the 2008 financial crisis). In contrast, I consider the effect of new variants or policy-based mitigation measures on both the spread of COVID and speed with which the model is capable of identifying these trends.

3.2.2 Dynamics and Network Analyses

Producing accurate forecasts only addresses the first goal of this project. To gain a deeper understanding of the fundamental structures and dynamical behavior of the system, a detailed analysis of the modes, eigenvalues, and best-fit operators was also performed. This analysis was complicated by the iterative nature of the cross-validated DMD algorithm. To select only the most critical dynamic modes, I follow a principled, holistic evaluation approach. The process has three key steps: first, identify the 'dominant' dynamic modes (those with the greatest influence on the forecast) following a process inspired by (11). Second, analyze the eigenvalues corresponding to the dominant modes in search of trends or consistent behaviors. Third, construct networks using the best-fit operators and assess the network centrality of specific system elements.

3.2.2.1 Selecting Dominant Modes

I closely follow the process described by (11) to create and visualize the dominance structures of the DMD forecasts at each step the cross-validated forecasts. To begin, the complex conjugates of complex-valued modes are discarded as these do not provide additional information to a high-level analysis. Then, for a subset of times $T \in \mathcal{T}$ associated with the given validation step, the norms of the modal contributions are assessed. That is,

$$\forall i \in |r| \text{ and } t \in T \ , w_{it} = \|\lambda_i^t \Phi_i b_i\|$$

Next, the two modes with the largest average relative contribution to the forecast were selected as 'dominant'. Expressing this in mathematical terminology,

$$\text{ArgMax} \left[\frac{1}{|T|} \sum_{t \in T} \frac{w_{it}}{\sum_{i \in |r|} w_{it}}, \forall i \in |r| \right]$$

where in this case the ArgMax command takes the two largest $i \in |r|$. As each dominant mode is selected, its corresponding (discrete-time) eigenvalue is also recorded.

3.2.2.2 Analyzing Dynamical Behavior

To seek out robust and persistent dynamical characteristics, I plot the dominant eigenvalues on the unit circle in the complex plane. In particular, positive real-valued eigenvalues describe monotonic behaviors, negative real-valued eigenvalues describe high-frequency oscillations, and all complex-valued eigenvalues describe oscillatory behaviors with a variety of frequencies. Further, eigenvalues outside the unit circle describe unstable, divergent behavior while those within the unit circle describe stable or converging dynamics. If a visual analysis of these eigenvalues prompted the conclusion that distinct clusters of dynamical behavior existed, I then create clusters using agglomerative clustering with a Ward linkage.

3.2.2.3 Network Analysis

Once I assessed the dominant dynamical traits of the system, the interdependencies between system elements was modeled using the best-fit operators. In particular, the reduced order operator, described earlier, may be expanded to a full system-size operator describing the relationships across the full system. In particular, this operator A is constructed by taking:

$$A = X_2 V \Sigma^{-1} U^* \tag{3.1}$$

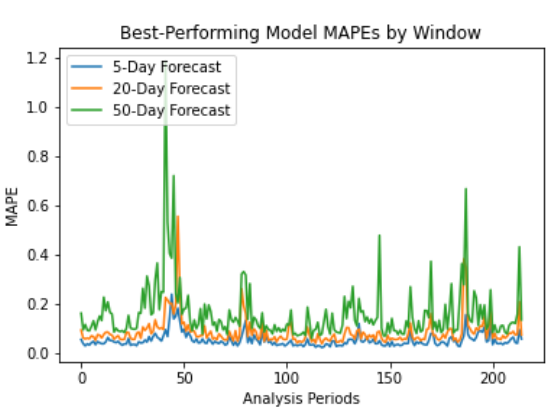
where V, Σ^{-1}, and U^* are of maximum possible rank, given the size of the input matrix. Each cell of A - that is, each a_{ij} - then corresponds to the strength of the connection running from node i to node j. To construct a realistically sparse network, I use the most significant 5% of these cells to construct the directed network describing the strongest connections between system elements. For purposes of this analysis, I prohibited self-loops within the network. This network was then be explored to identify central nodes and the topological characteristics of the underlying system. For one example of using a best-fit operator to identify network-based relationships, see (41).

Chapter 4

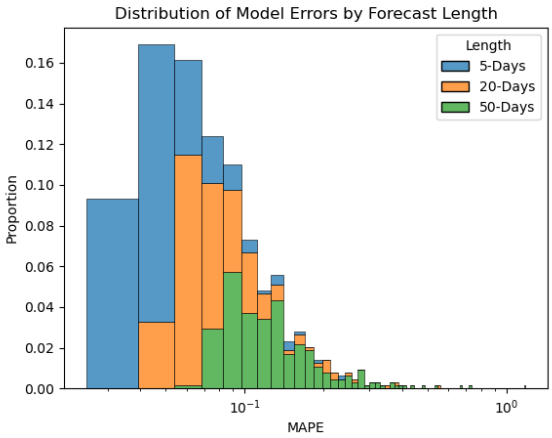
DMD and Financial Markets

4.1 Cross-Validated Forecasting

Constructing an effective model of financial markets using this iterative implementation of DMD required tuning two parameters: the training set size and the model rank. To prepare the data and account for the wide differentials between the prices of different companies, the natural log of the price was taken to reduce the magnitude of stock price differentials between companies. Under the methodology described earlier, it was found that the best training set duration was 190 trading days using a model built with five dynamic modes. This combination of parameters resulted in an average cross-validated MAPE of 0.1005. The forecast errors for each training window and the distribution of overall MAPEs are shown below:



(a) MAPEs by Cross-Validation Window

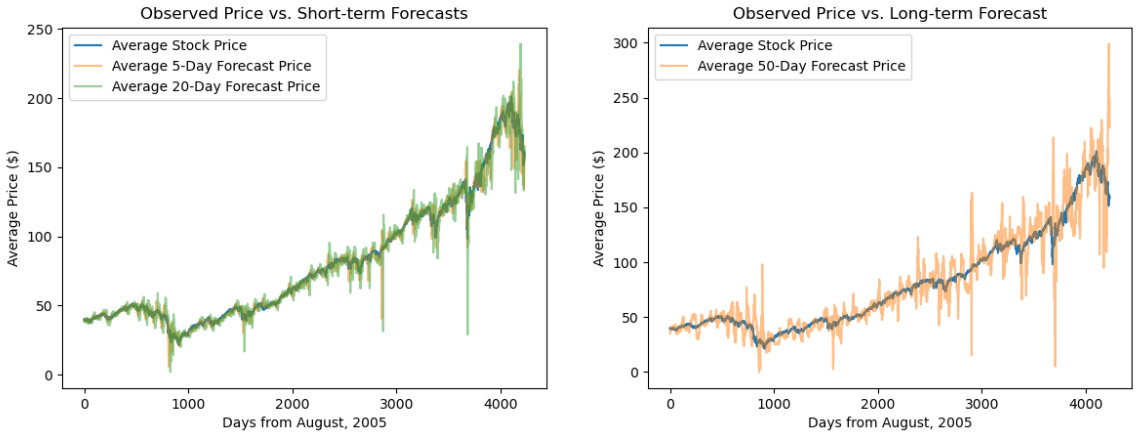


(b) Distribution of Model MAPEs

Figure 4.1: Measures of Financial Model Capability

The best-performing model (with parameter values given above) is capable of generating high-fidelity forecasts for periods of time one to four weeks into future. However, it's ability to project 50 trading days into the future is far more limited, showing an average MAPE of 0.1683 (more than double that of the one-week forecasts). The sharp spike in variability arising from these extended forecasts indicates that the functional drivers identified in the training set begin to evolve between one and three months into the future.

Having established the parameter values to most effectively model the behavior of S&P 500 constituent companies, a closer analysis of the patterns of variability was merited.



(a) 5- and 20-Day Forecasts (b) 50-Day Forecasts

Figure 4.2: Financial Forecasts vs. Observed Values

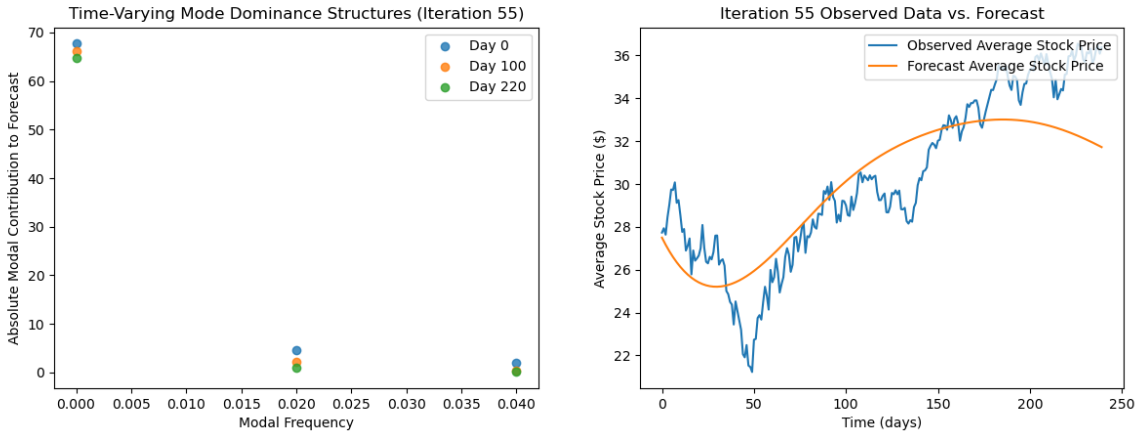
While the model's 5- and 20-day forecasts tend to adhere to observed behaviors, there is a significant degree of stochasticity around periods of significant market volatility. The short-term forecasts maintain a high-level of fidelity until the crash of 2020, at which point volatile behaviors limit one's ability to make accurate predictions regarding future trading prices. Meanwhile, the 50-day forecasts show substantively higher variability from observed behavior (as is expected), with the greatest deviations occurring after 2020. The lack of an obvious trend in the residuals lends credence to the hypothesis that market behavior is stochastic - subject to signal-less noise - and that accurately predicting daily volatility is particularly challenging.

4.2 Modal Analysis

4.2.1 Modal Dominance Structures

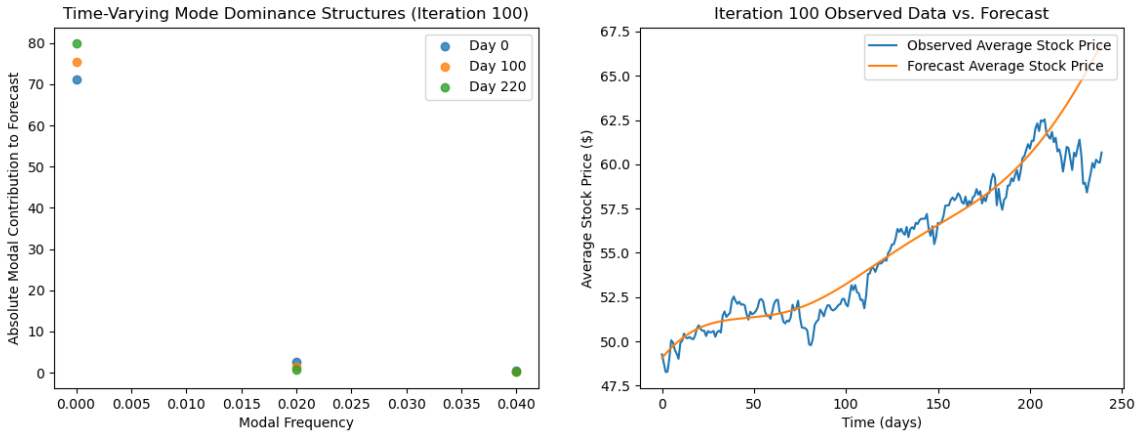
As there are over 200 steps in the cross-validated analysis of the financial market forecasts, providing a comprehensive review of the dominance structures of their consistent

features is untenable. Nonetheless, they do appear to display persistent features, regardless of the periods of the years in which the analysis occurred. First, the most significant modes across time are, in general, the "slowest" moving – that is, modes associated with an eigenvalue with imaginary part at or near zero. Second, while each forecast contains only 10 modes, there are no "fast-moving" modes. Indeed, it appears that the highest frequency modes will oscillate about once every 10 trading years. Given that these are derived from only 170 trading days of data, the overall system behavior may essentially be classified as composed of exponentially growing and decaying signals with minimal seasonality. The figures below highlight two cases of how modal dominance structures interact with different forecasts.



(a) Modal Contributions (Iteration 55) (b) Iteration 55 Forecast Capability

Figure 4.3: Iteration 175 Modal Analysis



(a) Modal Contributions (Iteration 100) (b) Iteration 100 Forecast Capability

Figure 4.4: Iteration 100 Analysis

In both these cases, it can be clearly seen that the real modes have greater magnitudes than the oscillatory modes. Additionally, these images highlight a key trade-off in model sizes. Smaller models, such as those used here, are relatively simple to interpret and allow the interactions between modes to be easily assessed. However, they result in heavily smoothed forecasts that largely ignore real-life variability. These figures also prompt a hypothesis that, at least over the relatively short time horizons used in this work, financial markets are driven by relatively stable trends, rather than highly volatile oscillatory behaviors.

To reinforce this hypothesis, I analyzed the behavior of the eigenvalues corresponding to the dominant modes. In the figure below I plot the discrete-time eigenvalues with non-negative imaginary component (as only the magnitude of the imaginary part is significant in describing dynamic characteristics. These points are colored according to the training set from which they were derived, with darker colors corresponding to earlier times (i.e., closer to August, 2004).

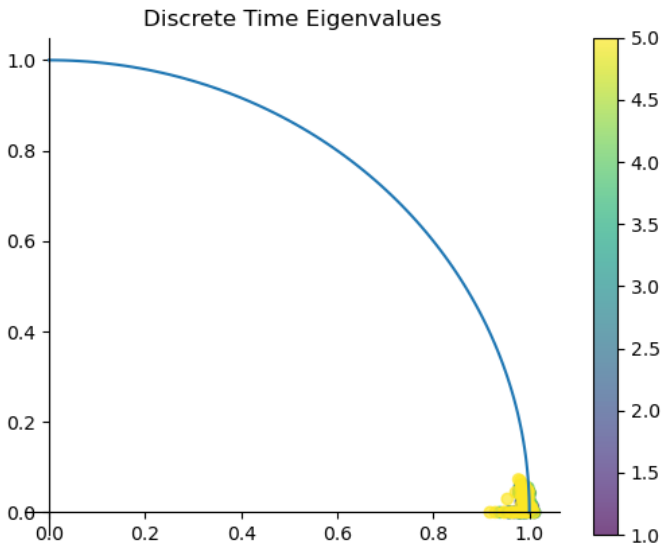


Figure 4.5: Dominant Discrete-Time Eigenvalues

An important initial observation is that all these dominant eigenvalues, regardless of the training set from which they were derived, are tightly clustered, indicating that the dominant dynamics of the market are relatively stable through time. As nearly all the eigenvalues fall on or within the unit circle, they describe stable or convergent dynamical behaviors. Nonetheless, these modes are still capable of describing growth behavior over shorter horizons due to constructive interference between their individual projections. Since the dynamical behavior of the dominant modes was so consistent through time (as evidenced by the minimal spread between eigenvalues), I did not cluster these dominant eigenvalues to search for distinct dynamical trends.

4.2.2 Robust Market Structures

I now consider whether there are robust market structures and interdependencies that emerge from a deeper analysis of the objects returned by the iterative DMD analysis. In particular, as the modes describe relationships which may be difficult to visualize

given the arbitrary ordering of companies, I consider whether there are persistent, network-derived relationships arising from the DMD operators themselves.

To reduce the computational demands while still providing robust results, I analyzed 15 randomly selected networks (from approximately 200 possible) to assess whether persistent trends were evident. Each network contained 8000 links (just under 5% of the possible 164,025 links) with no self-loops being permitted.

The first clear similarity between all these networks was the relatively scale-free degree distribution. For simplicity, I did not consider the differences between in-degree or out-degree of each node (for reasons that will be discussed later). A representative example of the degree distribution of these networks is shown in the figure below.

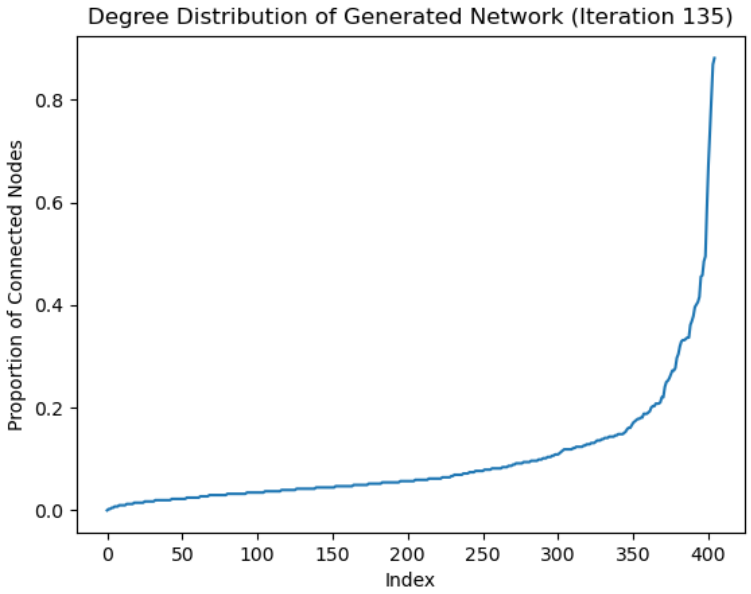


Figure 4.6: Representative Degree Distribution of S&P Constituent Networks

The horizontal axis simply orders the nodes (companies) from least connected to most connected, where the degree of connection is measured by the proportion of nodes with which a given node shares a link. Clearly, only a small proportion of nodes (approximately 20) are highly connected, while the remainder possess a relatively small

number of connections. This same representative network is visualized below, where the size and color of nodes roughly correspond with their degree centrality to the network.

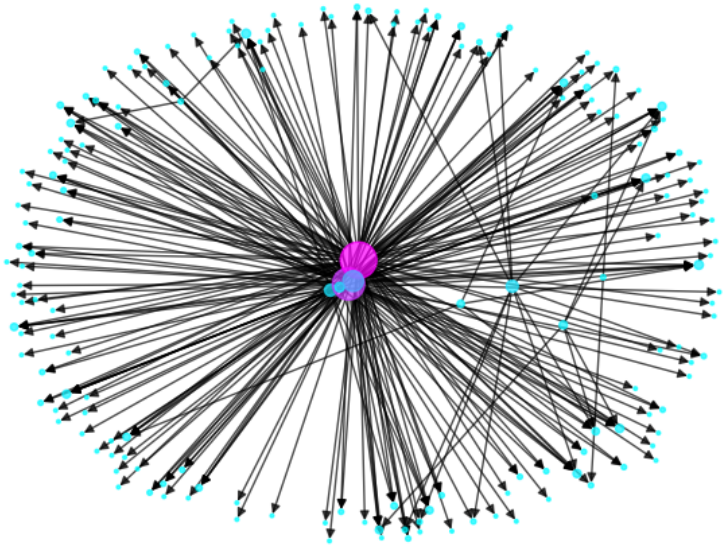


Figure 4.7: Representative Financial Network Visualization

Clearly, a tiny minority of nodes act as central "hubs" for the network – exercising tremendous influence on the peripheral nodes. The potential drivers of network centrality will be further explored in the discussion section, but it is worth noting here that while higher trading prices do tend to correlate with increasing centrality, it appears that factors external to price play a role.

Finally, when the 20 most central nodes (as measured by degree centrality) were analyzed for the 15 randomly selected networks, the top five most connected nodes remained consistent. That is, in addition to sharing similar the degree distributions, the high-ends of those distributions were composed of the same companies.

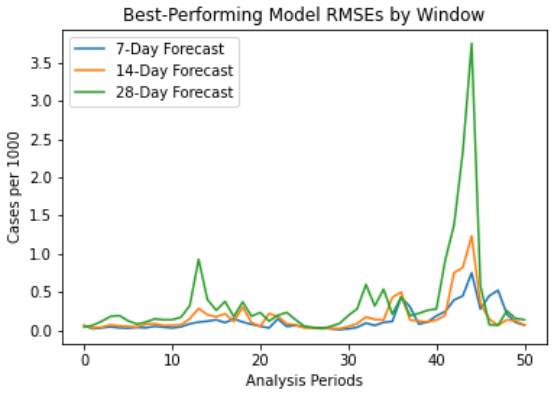
Chapter 5

DMD and COVID-19

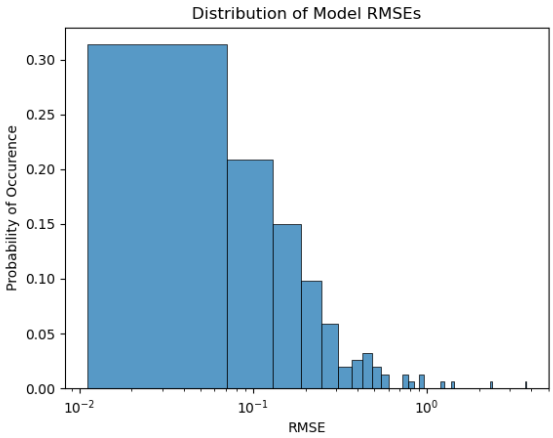
I now address the potential of DMD to identify robust characteristics that drive the spread of COVID-19 through time. I begin by discussing the results and overarching insights resulting from the cross-validated forecasting, then I address the insights regarding latent system characteristics which emerge from a thorough analysis of the operators and dynamic modes.

5.1 Cross-Validated Forecasting

As discussed earlier, effectively modeling the spread of COVID-19 in the US required tuning three parameters: the number of delay-embeddings, the training set size, and the model size. The best combination of parameters was found to be 10 embeddings, a training set size of 28 days, and a model containing 15 dynamic modes. The corresponding RMSE was 0.2281 (while it worth noting that many combinations of parameters generated average RMSEs below 0.3). The model's RMSEs for each forecasting period and overall distribution of RMSEs are shown below.



(a) RMSEs for each Cross-Validation Window



(b) Distribution of Model RMSEs

Figure 5.1: Measures of Epidemiological Model Capability

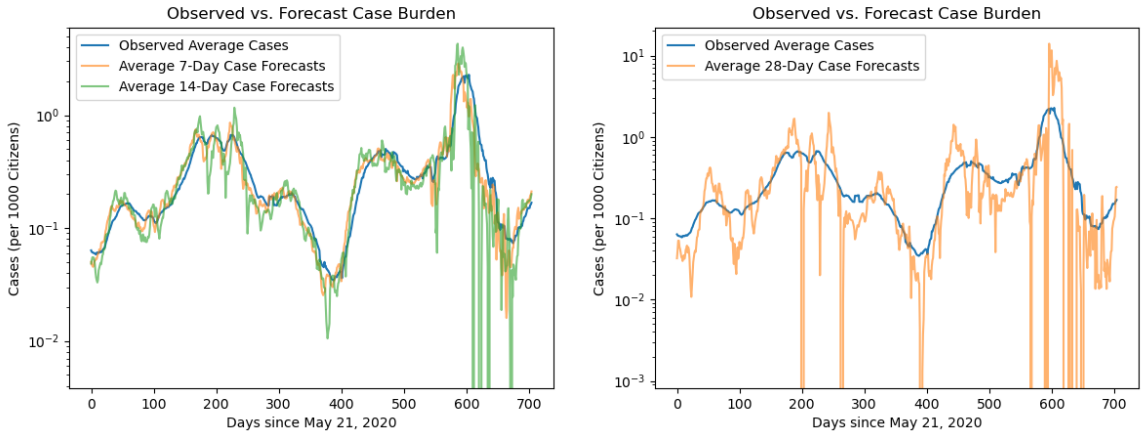
As can be seen in the figures above, the 7- and 14-day forecasts show low levels of variation from observed behaviors, only rarely exceeding an RMSE of 0.5 cases per 1000 citizens. Even the 28-day forecast shows low volatility in accuracy measure until the rise of the Omicron variant. Another factor worth noting is that over 75% of periods have an error rate below 0.25 cases per 1000 citizens, which in real terms is less than 1000 cases per day in a state the size of Oklahoma (arguably an acceptable error, given the disease burden in Oklahoma). While these error rates do not compare with well-designed and tuned compartmental models, this model's capability is satisfactory given its simplicity to develop, train, and tune. Further, the consistent model error measures (excepting the onset of the Omicron variant) lead to the hypothesis that an iterative implementation of DMD is capable of effectively describing small shifts in dynamics (e.g., progressive vaccination deployments or evolving public health interventions).

Before continuing the analysis of the best-fit models, a brief comment on the impact of parameters on the forecast accuracy is warranted. Given the noisiness of the system, increasing the number of embeddings generally improved the forecast accuracies (although the marginal improvement was much smaller after the number of embeddings

surpassed 7). Further, while training sets of approximately one month were most accurate, the size of the training set was less significant to model performance than the number of embeddings used. Additionally, while the most accurate model leveraged the maximum 15 modes, using more complex models did not always result in more accurate models, which speaks to the danger of overfitting DMD models.

5.1.0.1 Time-Varying Errors

While consistently high levels of forecasting accuracy are desired, insights may be derived from analyzing periods of significant volatility. Based on Figure 4.2(a), the three periods of greatest volatility roughly correspond to October 2020 through January 2021, August 2021 through October 2021, and December 2021 through January 2022. The figure below show the 7-, 14-, and 28-day forecasts for each day as compared with the true values to provide more granular insights into the model's time-varying accuracy.



(a) 7- and 14-Day Forecasts (b) 28-Day Forecasts

Figure 5.2: COVID Forecasts compared with Observed Values

These images prompt two thoughts which will inform future investigation into the DMD's forecast. First, it appears that (especially for the 14- and 28-day forecasts)

the model is more likely to produce forecasts orders of magnitude lower than the observed disease burden. While the causes of this consistent underestimation merit further investigation, it certainly implies that the dynamics identified by DMD decay more rapidly than in reality. Second, as the 7-day forecast closely follows the observed disease burden, it seems reasonable to argue that the forces driving the disease spread over the prior four weeks persist for at least an additional week. That is, the dynamical evolution of COVID-19 is a moderately-paced process - one day will not generally show dramatic change.

5.1.1 Adjusting for Omicron

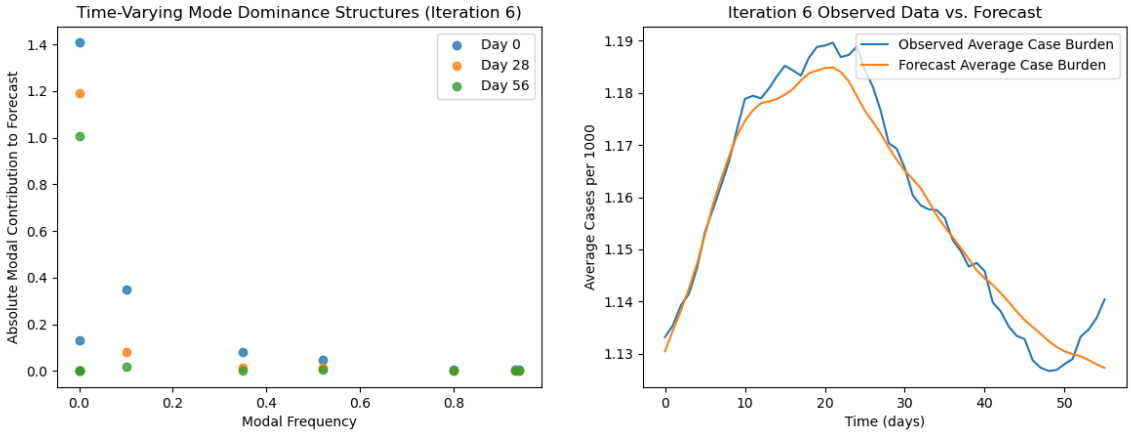
Before advancing to a detailed analysis of the dynamic modes of the system, as the largest deviations in forecast accuracy arose during the initial Omicron surge, I investigated whether the optimal choice of parameters would change with the exclusion of the Omicron data. That is, I assessed if excluding the largest source of variability would indicate that the driving dynamics of the disease's spread were more stable than found earlier. To assess this hypothesis, I followed the same analysis process outlined previously while truncating the analysis at December 1, 2021.

While the resulting models had, on average, lower RMSEs (consistently less than 0.2 cases per 1000 citizens) regardless of parameter values, the best-performing parameters remained unchanged, and the trends in parameter effects remained. Incorporating more delay-embeddings generally increased the cross-validated accuracy while longer training sets generally resulted in lower cross-validated accuracies. Consequently, it does not appear that the Omicron variant caused a shift in the values of the best-fit parameters.

5.2 Modal Analysis

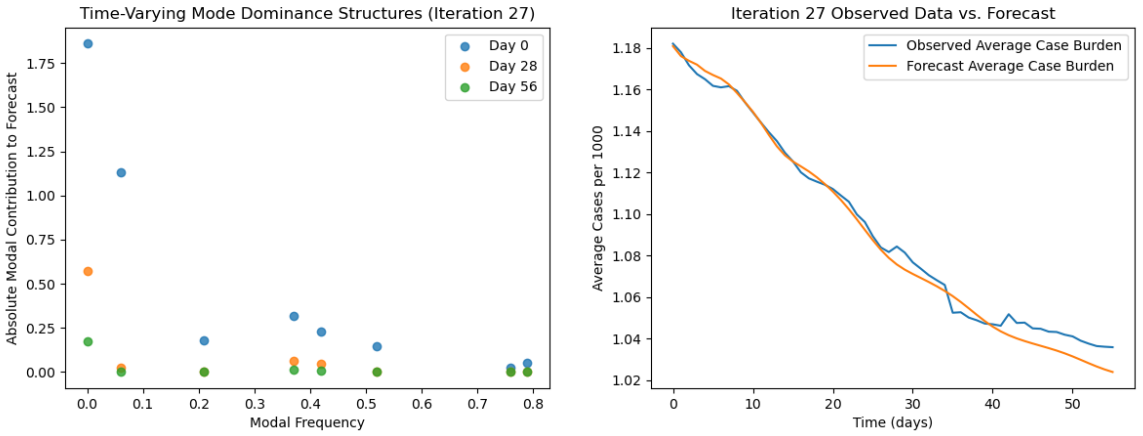
5.2.1 Modal Dominance Structures

To begin analyzing the impacts of the individual modes, I first identified the dominant modes as measured by relative contribution to a forecast measurement. In general, modes with high-frequencies (small imaginary component) or associated with real eigenvalues were dominant; however, during certain windows, slower-moving modes also had substantive impacts. Two examples are shown below. The left figure highlights the relative importance through time of each mode. The points show the norm of each mode at the specified time step. The right figure compares the resulting forecast with the observed values.



(a) Modal Contributions (Iteration 13) (b) Iteration 13 Observed Data

Figure 5.3: Iteration 13 Modal Analysis



(a) Modal Contributions (Iteration 27) (b) Iteration 27 Observed Data

Figure 5.4: Iteration 27 Modal Analysis

These figures highlight a persistent trend across all time periods, that modes associated with real eigenvalues generally have the largest absolute contribution to a forecast (generally at time 0). It is also clear that as the time-step advances, the absolute contribution of each mode diminishes. While this trend follows the observed characteristics of financial markets' models, the slower modes do appear, at least at certain periods, to play a more significant role.

Both figures demonstrate the capability of the model to generate forecasts which closely adhere to the observed dynamics of disease transmission. While Figure 5.4(b) demonstrates the model's capability of capturing almost monotonic behavior, Figure 5.3(b) highlights the model's ability to depict oscillatory behavior as well, even when only trained on a portion of the wave. It is worth observing that while the observed behavior captured in Figure 5.3(b) looks like a single sinusoidal wave, the decomposition models it as a linear combination of several functions, each with some growth and oscillatory components. Combining this figure with 5.3(a) also highlights the insight that the importance of a mode will vary through time, necessitating an evaluation of 'dominance' across a broad time horizon.

Both these figures also highlight a persistent trend of "slow-moving" modes (those with imaginary component near 1) having less impact on the forecasts. This likely occurs because these modes are associated with eigenvalues with small real component, indicating that (a) their growth rate will be small and more influenced by their oscillatory frequency.

Having identified the dominant triplets in the DMD forecasts, I analyzed the dominant eigenvalues alone, independent of their corresponding modal structures. These discrete-time eigenvalues are shown on the complex plane as compared with a unit circle below. The points within the unit circle correspond to stable behavior, while those outside the unit circle will, given adequate time, produce diverging behaviors. The eigenvalues are colored according to the time-stamp of their training data set, with darker colors associated with times earlier in the pandemic.

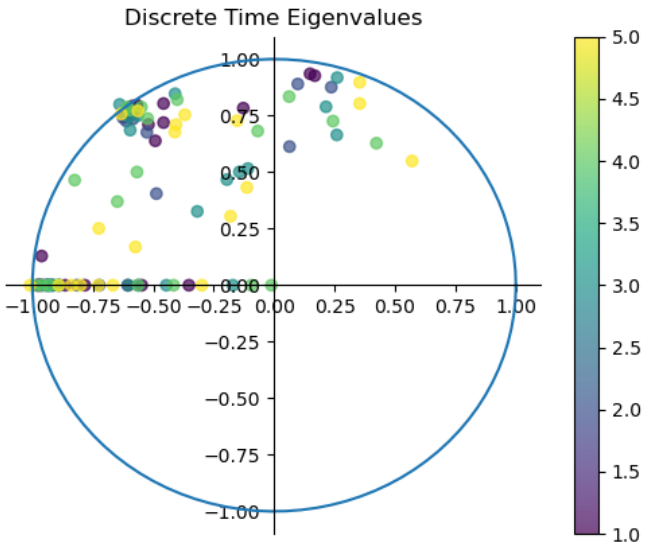


Figure 5.5: Dominant Eigenvalues on the Complex Plane

The eigenvalues corresponding to the dominant modes in the cross-validated time-steps predominantly fall within the unit circle, supporting the hypothesis that the

system persistently displays stable, convergent behavior. Further, as a majority of these eigenvalues have negative real part, it appears that rapid oscillations characterize the behavior of many modes. Despite the rapid growth associated with the Omicron variant, no dominant eigenvalue lies outside the unit circle, implying that even in the most extreme case, the growth rate of case counts was somewhat limited. Further, the model uncovers numerous modes that display a range of oscillatory behaviors; their frequencies range from every 24 days up to slightly more than 2 years. As the frequencies were obtained from only 28 days of training data, additional analysis is needed to determine the robustness of these findings.

To assess whether the dominant eigenvalues should be clustered, I assessed an agglomerative clustering (using a Ward linkage) dendrogram to identify an appropriate number of groups. The dendrogram resulting from this analysis is shown below.

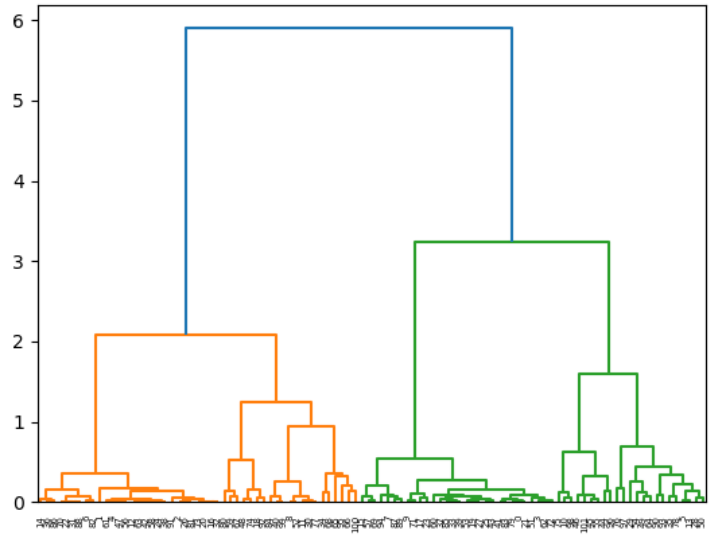


Figure 5.6: Agglomerative Clustering Dendrogram for Dominant Eigenvalues

Based on the figure above, there appears to be good separation among distinct groups of eigenvalues using either two or three clusters. I ultimately decided to use

three clusters as the dominant eigenvalues describe a wide range of behaviors. These three clusters are plotted on the complex plane below.

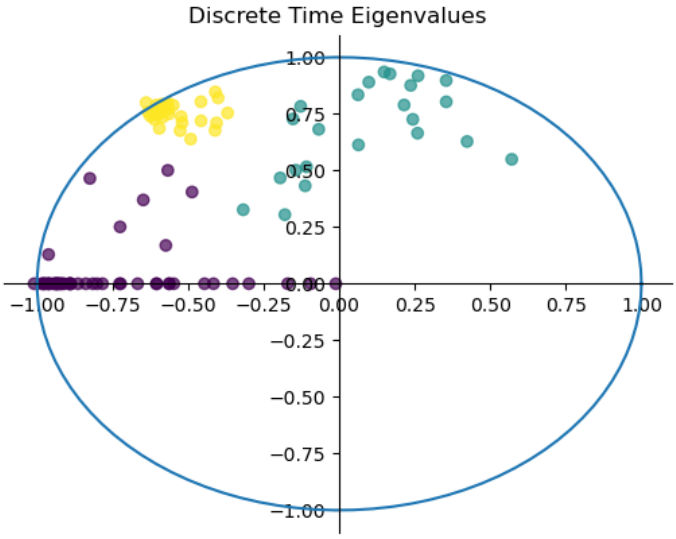


Figure 5.7: Eigenvalue Clusters

One cluster of eigenvalues (colored purple) primarily contains the modes governed by negative, real-valued eigenvalues (which describe rapid oscillations quickly converging toward zero). One cluster (colored yellow) contains a dense cluster of eigenvalues with negative real-component near the boundary of the unit circle, associated with slightly slower oscillations than those in group 1. Finally, the third cluster (colored sage) contains a more disperse collection of eigenvalues, which generally result in smoother oscillatory behavior than those observed in the other clusters. A representative visualization of the dynamics from each cluster is given below.

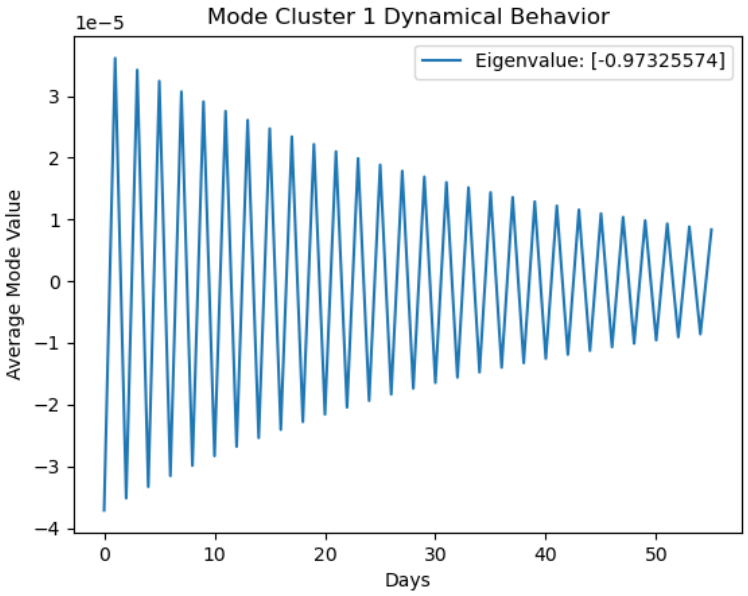


Figure 5.8: Characteristic Dynamics of Cluster 1

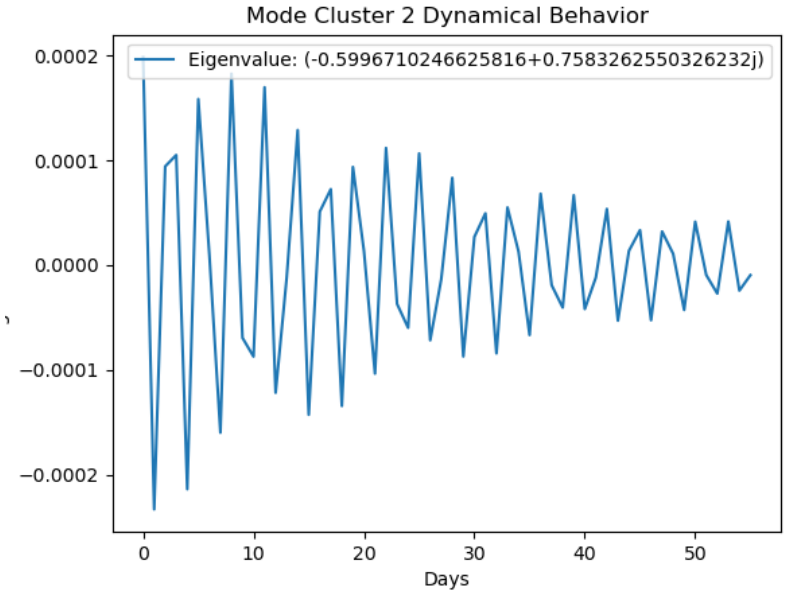


Figure 5.9: Characteristic Dynamics of Cluster 2

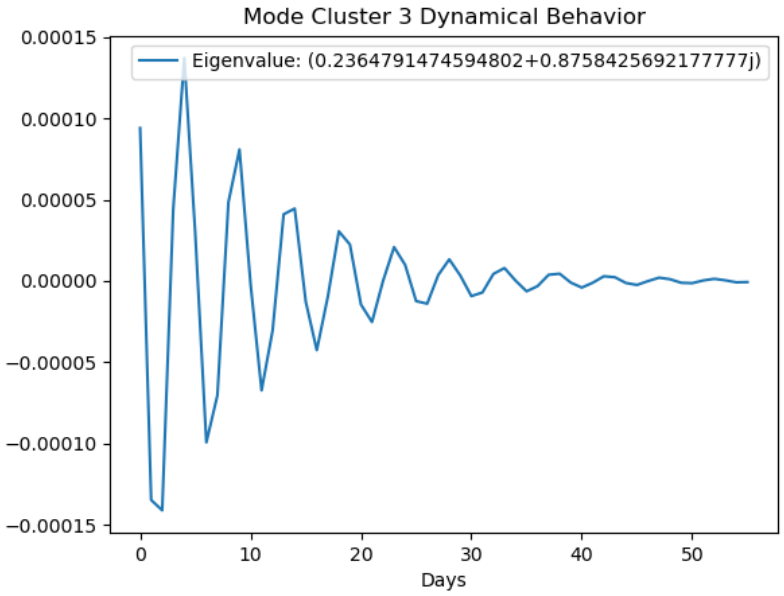


Figure 5.10: Characteristic Dynamics of Cluster 3

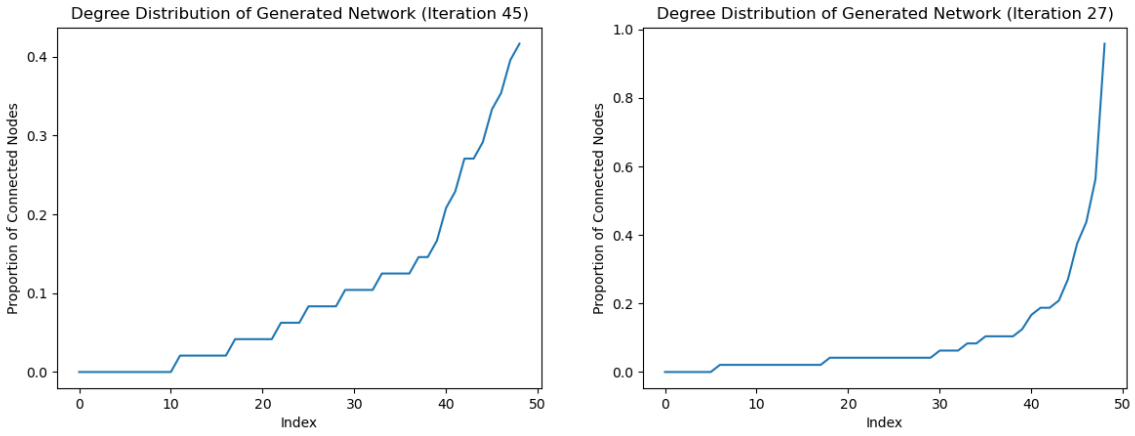
Even a cursory analysis of the above visualizations highlights the remarkable variability in day-to-day forecasts from these modes. In contrast to the comparatively smooth dynamics of the S&P 500 constituents, these modes and eigenvalues highlight the significant stochasticity and noise in the input data set. The high day-to-day variance of these modes also explains the necessity of more complex models than the financial markets; more modes must be combined to accurately describe the evolving burden of COVID-19.

5.2.2 Interstate Network Structures

Next, I considered the network-derived relationships between US States. I constructed the operator-derived networks following the approach described in Chapter 4, here with 120 connections per network. As these networks are smaller and are generated from a smaller number of analysis periods, it was feasible to perform preliminary analyses on

each network. I begin by discussing the overarching trends in the degree distribution and topology of these networks.

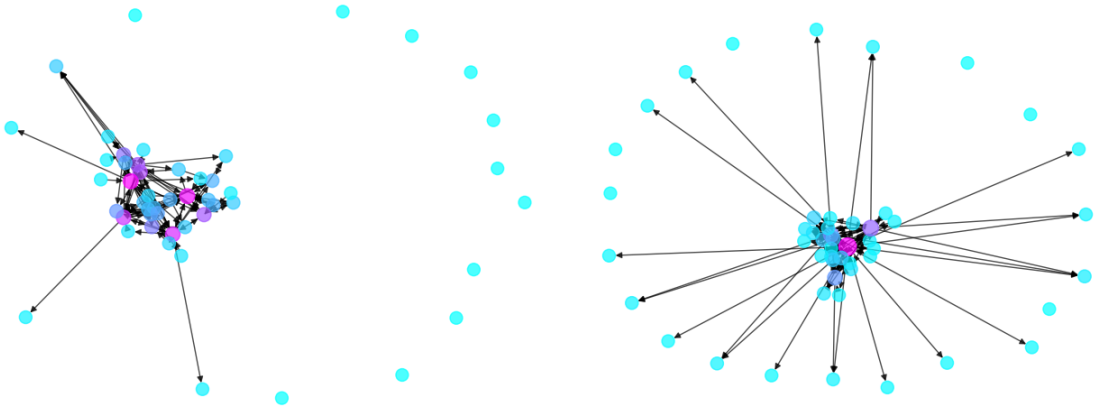
First, the network topology continues to resemble the 'scale-free' property (although the effect is less pronounced than in the financial networks). Where the financial networks persistently modeled this behavior, these epidemiological networks occasionally display a polynomial, rather than exponential, degree distribution as is seen in the figures below.



(a) Iteration 45 Degree Distribution (b) Iteration 27 Degree Distribution

Figure 5.11: Emergent Network Degree Distributions

Clearly, the network generated from the 45th cross-validated analysis step has a slower rate of decline in the connectivity of nodes than the network from the 27th analysis step. The images are further clarified if one considers plots the resultant network, as can be seen in the figure below.



(a) Iteration 45 Network (b) Iteration 27 Network

Figure 5.12: Emergent Network Visualizations

The above figure highlights two key features of the emergent networks mapping the strengths of the connections between states. First, there is a relatively large, dense cluster of states at the center of the network. This cluster is larger in the 45th cross-validation analysis step, when the slope of the degree distribution was more gradual. In both cases, the central cluster is far larger (in relative terms) than that identified in the financial networks, indicating a more disperse distribution of network influence. Second, especially in the 45th iteration, a large proportion of states are completely disconnected from the other states.

Having assessed the topology of these generated networks, I assessed whether specific states were routinely highly central to the network. The first observation that arises from this analysis is that no state is persistently central to the network. The state with the highest network centrality (Rhode Island) has an average degree centrality of only 0.29 - meaning that, on average, it shares connections with only 15 other states. A Choropleth map showing the average network centrality of each state is shown in the figure below.

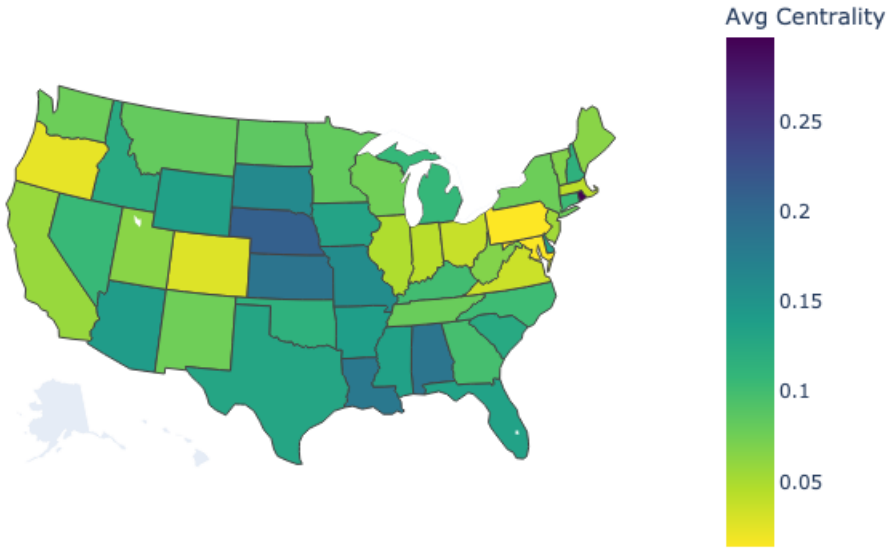


Figure 5.13: Average Network Centrality of US States

This figure is striking in that states that one might anticipate to be particularly significant – e.g., California, Texas, or New York – have comparatively low centrality while smaller states like Nebraska, Alabama, and Louisiana are identified as exercising particular importance. To further emphasize the significant variance in state centrality across time steps, I created a Ridgeline plot of the 10 most central states (on average).

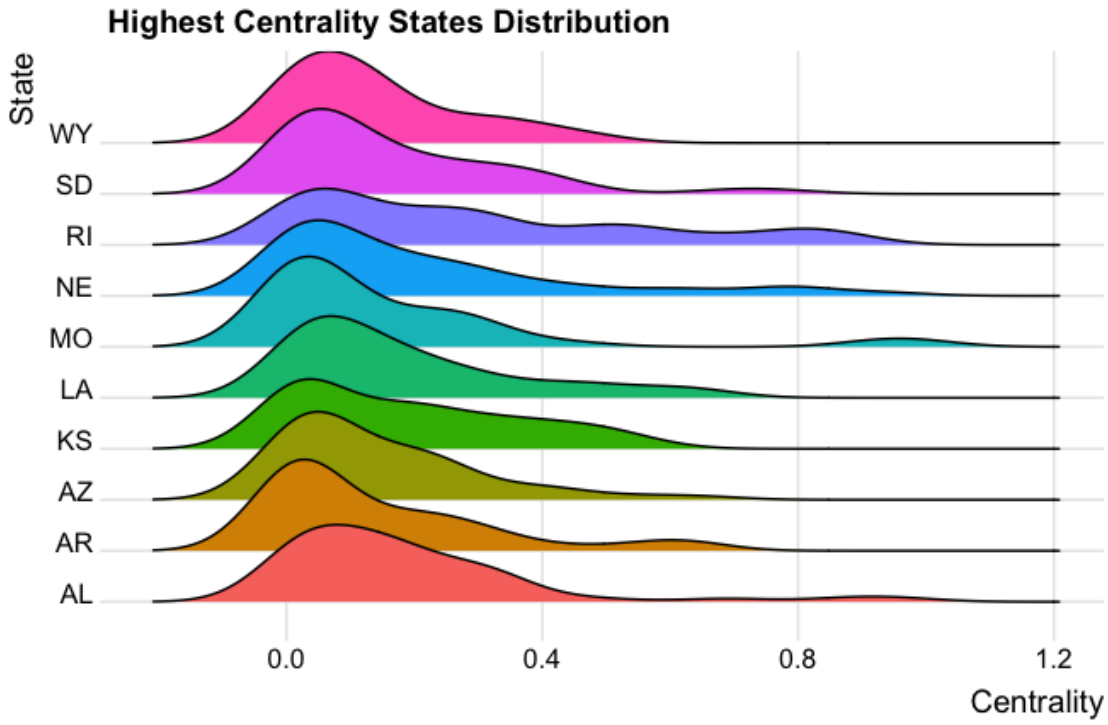


Figure 5.14: Distribution of Network Centrality Measure

The above figure emphasizes that, even among the most central states, the mode of their centrality occurs near zero. That is, their higher average centrality ratings emerge from longer right tails of the distribution, rather than a meaningful shift in the mode.

Chapter 6

Discussion and Conclusions

Having outlined the core results from both case studies, I now discuss the most important implications of the results, some limitations, practical application areas, and several potential avenues for future analysis and study.

6.1 Key Takeaways

DMD produces models with a rich array of outputs which require significant investigation to obtain actionable conclusions. Here I discuss key implications arising from the results discussed in the prior two chapters, with particular emphasis on how these results may be leveraged to improve system understanding and, possibly, system control.

6.1.1 Financial Markets Discussion

While the accuracy of the financial model varied through time, a fascinating feature of the model was its low complexity. While a cursory glance at the overall data reveals significant day-to-day volatility, the best-performing DMD model largely ignored this volatility, instead rewarding a focus on long-term trends. Indeed, more complex models, with greater capability of capturing volatility, consistently under-performed simpler models, indicating that daily price volatility is a function of noise, not a meaningful signal. The success of low-rank models further indicates that financial markets, at

least when observed over a long time-horizon, display low-rank characteristics. Such a conclusion also aligns with the existing literature, in which comparatively simple, statistical models like ARIMA can consistently compete with state-of-the-art neural networks.

Another fascinating feature of DMD's financial models is the consistency of the dominant dynamics. In spite of periodic downturns and intense periods of daily volatility, the dominant eigenvalues are tightly clustered. This remarkable consistency reinforces the importance of long-term trends in describing market behavior rather than daily volatility. Further, the robust consistency of the dominant eigenvalues implies that the real challenge in financial modeling is in describing daily fluctuations (as ARCH-derived models attempt to do) rather than depicting long-term trends. All this leads to the conclusion that, while financial markets are certainly complex systems, their complexity is a product of noise, not an intractable signal.

Considering now the properties of the emergent networks derived from the DMD operators, the most persistent characteristic is its scale-free topology. A scale-free topology is characteristic of many real, complex systems ranging from power grids to social networks, so its emergence in this context is unsurprising. More importantly, regardless of the training period from which the network was generated, the dominant nodes (companies) remained overwhelmingly consistent. In particular, the most significant companies seemed to be involved in the financial sector – whether as traditional financial institutions or more tangentially as mortgage companies. Additionally, while increased trading price does appear to correlate with greater network centrality, price alone is insufficient to explain whether a particular company exercises an out-sized influence on other community members. Consider the figure below, relating average network centrality to average trading price through time.

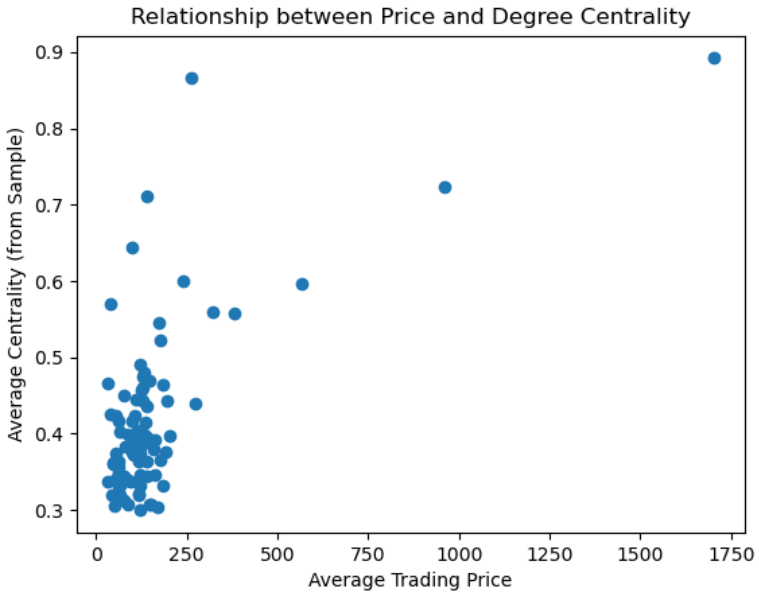


Figure 6.1: Network Centrality as a Function of Trading Price

This figure shows a collection of companies which were, during at least one randomly-selected time step, among the 20 most central companies in the network. In total, approximately 80 companies are included in the figure. While the highest-valued companies in the figure are consistently central to the network, many lower-priced stocks (with average prices over the 18-year period of $250 and less) also have high average levels of centrality. Further exploration into additional explanatory factors driving a particular companies influence is certainly warranted, as it could provide insights into fundamental drivers of market behavior.

Finally, as was noted earlier, the relative centrality of each node was relatively stationary through time. That is, there did not appear to be dramatic swings where a peripheral node moved to a central position or vice versa. While additional analysis is warranted to explore the possible drivers of this persistent behavior, it is worth considering that the companies included in this analysis are all long-standing, successful, and highly influential institutions. Further, given the relative stability of the S&P

500 Index, persistent company-to-company relationships make intuitive sense, as these could facilitate the S&P 500's low-volatility behavior.

6.1.1.1 Future Applications of DMD to Financial Models

While DMD may not generate forecasts with the accuracy of other state-of-the-art models, the wide-ranging insights that arise from DMD add substantively to its value and offer many opportunities for future research. Two areas in particular seem particularly ripe for future exploration. First, I believe there is significant potential for exploring DMD extensions to improve the cross-validated accuracy of this modeling approach. As was noted earlier, my models routinely discarded day-to-day volatility in favor of adhering to long-tern trends. However, daily volatility is a critical feature in modeling financial markets, and one that ought not be discounted. Two existing extensions of DMD – namely multi-resolution DMD (mrDMD) and DMD with Control (DMDc) have the potential to improve DMD's ability to handle increased volatility. mrDMD allows signals from multiple time-horizons to be analyzed, perhaps facilitating the collection of signals that, using my iterative approach, were discarded as noise. Alternatively, DMDc allows external factors (for example, interest rates and other governmental interventions) to be included in the analysis, thereby enhancing the depth and sophistication of system understanding.

Second, additional exploration of the networks generated from the best-fit operators is warranted. In this work, I performed an exploratory analysis of the similarities in network structures through time. While it was evident that the emergent networks showed scale-free tendencies, with certain companies routinely acting as hubs, I have barely scratched the surface of the potential for network-specific analysis to elucidate latent characteristics of financial markets. Comparing the emergent DMD networks with those generated by alternative means (e.g., via the contracts and cash

flows between companies) could be leveraged to validate the fidelity of these generated networks. Further, these networks could be tested to evaluate their robustness to unexpected downturns, the results of which could be readily compared with real economic turbulence.

6.1.2 COVID-19 Modeling Discussion

As described in chapter 5, the 7- and 14-day cross-validated COVID-19 forecasts consistently held to RMSEs of less than 0.15 cases per 1000 citizens. While the DMD models lack the clear and specific interpretability of compartmental models, they were insensitive to poorly tuned parameters and capably described observed system disease burden through a range of disturbances. Given these robust results, one may reasonably argue that DMD can effectively handle the real-world variability arising from a wide range of unconsidered inputs, including vaccines, evolving governmental regulations, and even (comparatively) mild disease variations. The capability of this approach is directly attributable to the short training periods which prevent outdated dynamics from overwhelming more recent behaviors.

Indeed, given the high accuracy of forecasts even two weeks into the future using only one month of training indicates that shifts in the disease burden develop smoothly. As only the arrival of the Omicron variant was capable of decimating the accuracy of the DMD forecast, this iterative implementation readily handled shelter-in-place orders, vaccine roll-outs, and other sociological factors. While this fact could be driven by COVID-19's relatively long incubation period, this feature alone seems unlikely to explain these slow changes in disease evolution.

This line of inquiry points to a potential extension of this work. Given that DMD's approximation of the dynamics of COVID's spread are entirely described by the best-fit operators, one could describe the characteristics of these operators and how they evolve

through time to provide a more granular and quantitative description of COVID-19's evolving dynamics. In particular, this approach could be leveraged to tease out the effect sizes of specific interventions and possibly how these effects vary across states.

Considering now the network topology of the operator-derived networks, two points deserve further elaboration. First, unlike the networks derived from the S&P 500 data, no state was consistently identified as consistently central to the network. Indeed, the mode of each state's frequency was barely higher than zero – indicating that the most likely occurrence was a state sharing one or two edges with additional states. Further, there did not appear to be a temporal trend in the centrality of any state. Consider, for example, the centrality of Rhode Island plotted in time-sequential order.

While there are three distinct peaks in the centrality of Rhode Island in time, additional analysis would be required to assess whether this behavior is associated with meaningful insights into the burden of COVID-19 across the US as a whole. One possible driver of this behavior is that Rhode Island, given its relatively high population density and proximity to major cities like New York had spikes in COVID-19 cases slightly before these same peaks struck the remainder of the US. However, given that Wyoming (a highly disperse, isolated state) also had a relatively high average centrality, such an explanation fails to account for all possible factors driving a state to become particularly central to the network.

The second point worth considering is the relatively large size and density of the central "cluster" of states within the networks. Traditional scale-free networks have a much sparser collection of central hubs, indicating that describing the network of connections between states as scale-free is likely inaccurate. Since traditional metrics of "importance" - like GDP, population, or population-density - do not seem to explain the number of connections shared by a given state, identifying a descriptive topological

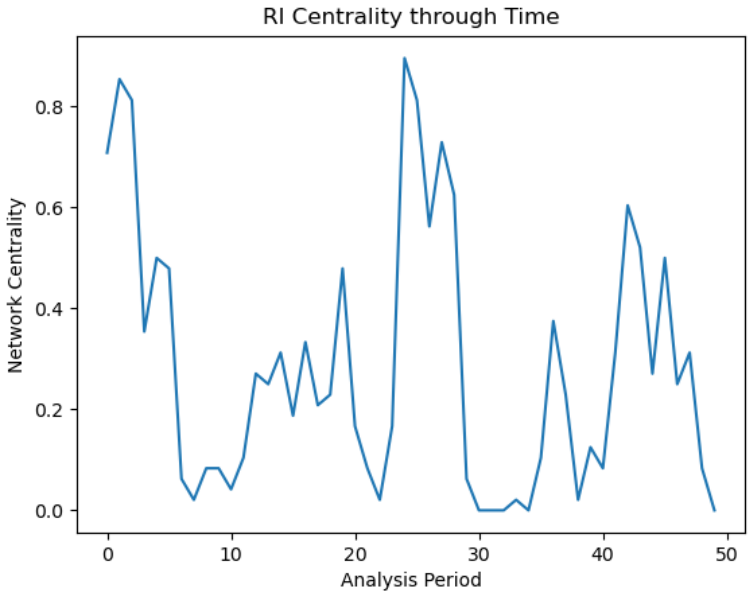


Figure 6.2: Rhode Island's Time-Varying Centrality

characterization of these networks and a corresponding interpretation could elucidate latent factors which promote or impede disease transmission within and across state lines.

6.1.2.1 Opportunities for Future Work

The first clear opportunity for additional exploration of the COVID-19 data set is expanding the network analysis of the operator-derived graphs. Given that including 5% of the edges resulted in networks that were not fully connected, identifying the point at which fully connected graphs emerge could provide insights into the degree to which one state's policies or disease burden influence others. Additionally, permitting self-loops could elucidate whether individual states are the most significant influence on their disease burden. As well, probing the relationships between the training data and

the resultant networks would aid in interpreting the networks and assessing their robustness to noise in the data – a desirable feature in such networks, given the variability in data quality in emerging epidemics.

Other opportunities for future work include adding additional measures to the data set (e.g., deaths from COVID-19). These measures would allow relationships between the different measures to be quantified while also providing insight into the quality of different data sources. An additional area for exploration is in applying this methodology to more granular data (e.g., county-level data) or data from other countries. Exploring whether this analytical approach is capable of producing meaningful results across a range of scenarios is an indispensable step in demonstrating its applicability and value.

6.2 Areas of Practical Application

While this work has been shown to capably model a range of behaviors across a spectrum of test cases, I believe that it has potential to inform real-world decisions and enhance the insight available to decision-makers in a variety of fields. Many practical applications undoubtedly exist; for purposes of this work, I will focus on one recently relevant area: evaluating other models.

At the outset of the COVID-19 pandemic, researchers across the world, in an effort to get ahead of the disease's unpredictable spread, attacked modeling COVID-19 in numerous ways, with varying degrees of success. Given the dearth of information available regarding how the disease might spread and the many conflicting opinions voiced by experts in the field, decision-makers had to guess at which models were best. This uncertainty featured prominently in the public discourse and furthered the anxiety and conflict which characterized the early days of the pandemic. While I do not

claim that the the iterative approach to DMD proposed in this work would eliminate uncertainty regarding which models were most accurate, it has illustrated that a short period of training data, even early in the pandemic, could accurately describe disease burden one-to-two weeks into the future. This fact leads to a practical use of iteratively applied DMD: helping decision-makers identify best-performing models under uncertain conditions. While this tool cannot eliminate uncertainty, it could certainly inform a decision-maker's modeling choices and enhance the utility of their chosen interventions.

6.3 Conclusion

This work implements the traditional, SVD-based DMD algorithm in a novel, iterative framework to facilitate accurate forecasting of complex system behaviors without sacrificing DMD's capability of identifying key explanatory components of system behavior. In addition to providing a thorough review of the DMD algorithm and related mathematical tools, this work describes a novel alteration on the traditional DMD approach to emphasize forecasting capability.

This approach was tested on two distinct data sets - a financial data set comprising 18 years of data on S&P 500 constituents and an epidemiological data set containing 2 years of daily US State COVID-19 case counts. Despite the clear distinctions between these data sets, the DMD algorithm was shown to be capable of consistently generating forecasts with close adherence to observed behaviors. While alternative modeling approaches which specialize in future-state prediction may create more accurate models, this iterative implementation of DMD still permitted a system-identification analysis to be performed.

In particular, using the outputs from the DMD approximations, dominant dynamical trends were identified and robust network structures between system elements were discussed. Distinct differences in the dynamical behavior between the financial and epidemiological data were identified. DMD was able to describe the trends in the financial data using low-rank models, which tended to emphasize smooth, growth-oriented behaviors. In contrast, far more complex models were required to describe the behaviors observed in US COVID-19 cases, perhaps indicating that stochasticity and poor data quality were significant features of the data. One particular strength of iteratively running DMD was evidenced in the network analysis of financial markets, as certainly companies were persistently identified as critical nodes. These companies were generally members of the financial sector, but their stock price alone cannot explain their persistent centrality to the network.

This work has demonstrated the potential of DMD to be utilized in contexts where limited data is available and where rapid evolution of system behavior is expected. Future explorations of this approach to implementing DMD may identify methods of quantifying changes in dynamical behavior and produce standard approaches to analyze and interpret operator-generated networks. The operator-theoretic approach to dynamical systems analysis used as the theoretical foundation for this work shows significant promise for enhancing the tools and approaches used for modeling a wide range of phenomena.

Reference List

[1] H. Arbabi and I. Mezić, "Ergodic theory, dynamic mode decomposition and computation of spectral properties of the koopman operator," 11 2016.

[2] M. Budišić, R. M. Mohr, and I. Mezić, "Applied koopmanism," 6 2012.

[3] B. Koopman, "Mathematics: Hamiltonian systems and transformations in hilbert space," 1931.

[4] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010.

[5] C. W. ROWLEY, I. MEZIĆ, S. BAGHERI, P. SCHLATTER, and D. S. HENNINGSON, "Spectral analysis of nonlinear flows," *Journal of Fluid Mechanics*, vol. 641, pp. 115–127, 12 2009.

[6] I. Mezić, "Koopman operator, geometry, and learning of dynamical systems," *Notices of the American Mathematical Society*, vol. 68, p. 1, 8 2021.

[7] S. L. Brunton, M. Budišić, E. Kaiser, and J. N. Kutz, "Modern koopman theory for dynamical systems," 2 2021.

[8] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*. SIAM, 2016.

[9] M. Gavish and D. L. Donoho, "The optimal hard threshold for singular values is $4/\sqrt{\{}3\}$," *IEEE Transactions on Information Theory*, vol. 60, pp. 5040–5053, 8 2014.

[10] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," 11 2013.

[11] T. Krake, S. Reinhardt, M. Hlawatsch, B. Eberhardt, and D. Weiskopf, "Visualization and selection of dynamic mode decomposition components for unsteady flow," *Visual Informatics*, vol. 5, pp. 15–27, 9 2021.

[12] J. L. Proctor and P. A. Eckhoff, "Discovering dynamic patterns from infectious disease data using dynamic mode decomposition," *International Health*, vol. 7, pp. 139–145, 1 2015.

[13] Y. Susuki and I. Mezic, *2015 54th IEEE Conference on Decision and Control (CDC) date: 15-18 Dec. 2015.*

[14] G. Beylkin and L. Monzón, "On approximation of functions by exponential sums."

[15] R. Zhang and G. Plonka, "Optimal approximation with exponential sums by a maximum likelihood modification of prony's method," 2019.

[16] G. Plonka and M. Tasche, "Prony methods for recovery of structured functions," 2013.

[17] G. Golub and V. Pereyra, "Separable nonlinear least squares: the variable projection method and its applications," *Inverse Problems*, vol. 19, pp. R1–R26, 4 2003.

[18] M. Carlsson, "Aak-theory on weighted spaces," 2009.

[19] R. Mahmoudvand and M. Zokaei, "On the singular values of the hankel matrix with application in singular spectrum analysis," 2012.

[20] H. Hassani, "Munich personal repec archive singular spectrum analysis: Methodology and comparison singular spectrum analysis: Methodology and comparison," 2007.

[21] V. Peller, "Springer monographs in mathematics."

[22] G. Plonka-Hoch, R. Luke, S. Kunis, A. Schöbel, T. Krivobokova, and H. Kriete, "Betreuungsausschuss: Weitere mitglieder der prüfungskommission," 2017.

[23] T. Andersen, T. Bollerslev, P. Christoffersen, and F. Diebold, "Volatility forecasting," 3 2005.

[24] M. R. Islam and N. Nguyen, "Comparison of financial models for stock price prediction," *Journal of Risk and Financial Management*, vol. 13, p. 181, 8 2020.

[25] A. Fadlalla and C.-H. Lin, "An analysis of the applications of neural networks in finance," *Interfaces*, vol. 31, pp. 112–122, 8 2001.

[26] S. M. Raihan, Y. Wen, and B. Zeng, "Wavelet: A new tool for business cycle analysis," 2005.

[27] L. Feng, B. Li, B. Podobnik, T. Preis, and H. E. Stanley, "Linking agent-based models and stochastic models of financial markets," vol. 109, 2012.

[28] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and arima models for time series forecasting," *Applied Soft Computing*, vol. 11, pp. 2664–2675, 3 2011.

[29] Y. Xiao, J. Xiao, J. Liu, and S. Wang, "A multiscale modeling approach incorporating arima and anns for financial market volatility forecasting," *Journal of Systems Science and Complexity*, vol. 27, pp. 225–236, 2014.

[30] H. Boubaker, G. Canarella, R. Gupta, and S. M. Miller, "A hybrid arfima wavelet artificial neural network model for djia index forecasting," *Computational Economics*, 2022.

[31] H. Hassani and D. Thomakos, "A review on singular spectrum analysis for economic and financial time series," 2010.

[32] J.-C. Hua, S. Roy, J. L. McCauley, and G. H. Gunaratne, "Using dynamic mode decomposition to extract cyclic behavior in the stock market," *Physica A: Statistical Mechanics and its Applications*, vol. 448, pp. 172–180, 4 2016.

[33] J. Mann and J. N. Kutz, "Dynamic mode decomposition for financial trading strategies," 8 2015.

[34] W. . Kermack and A. G. Mckendrick, "A contribution to the mathematical theory o f epidemics."

[35] S. Sturniolo, W. Waites, T. Colbourn, D. Manheim, and J. Panovska-Griffiths, "Testing, tracing and isolation in compartmental models," *PLoS Computational Biology*, vol. 17, 3 2021.

[36] S. P. Silal, F. Little, K. I. Barnes, and L. J. White, "Sensitivity to model structure: a comparison of compartmental models in epidemiology," *Health Systems*, vol. 5, pp. 178–191, 10 2016.

[37] G. Massonis, J. R. Banga, and A. F. Villaverde, "Structural identifiability and observability of compartmental models of the covid-19 pandemic," 1 2021.

[38] X. Zhang, T. Zhang, A. A. Young, and X. Li, "Applications and comparisons of four time series models in epidemiological surveillance data," *PLoS ONE*, vol. 9, 2 2014.

[39] B. F. Finkenstädt and B. T. Grenfell, "Time series modelling of childhood diseases: a dynamical systems approach."

[40] B. Cazelles, M. Chavez, G. C. D. Magny, J. F. Guégan, and S. Hales, "Time-dependent spectral analysis of epidemiological time-series with wavelets," 8 2007.

[41] A. D. González, A. Chapman, L. Dueñas-Osorio, M. Mesbahi, and R. M. D'Souza, "Efficient infrastructure restoration strategies using the recovery operator," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, pp. 991–1006, 12 2017. Big Idea:.