2022-04-25

Characterizing Equivalence and Correctness Properties of Dynamic Mode Decomposition and Subspace Identification Algorithms

Samuel Gregory Neff
Brigham Young University

Characterizing Equivalence and Correctness Properties of Dynamic Mode Decomposition and Subspace Identification Algorithms

Samuel Gregory Neff

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Sean Warnick, Chair
Mark Transtrum
Chris Archibald

Department of Computer Science

Brigham Young University

ABSTRACT

Characterizing Equivalence and Correctness Properties of Dynamic Mode Decomposition and Subspace Identification Algorithms

Samuel Gregory Neff
Department of Computer Science, BYU
Master of Science

We examine the related nature of two identification algorithms, subspace identification (SID) and Dynamic Mode Decomposition (DMD), and their correctness properties over a broad range of problems. This investigation begins by noting the strong relationship between the two algorithms, both drawing significantly on the pseudoinverse calculation using singular value decomposition, and ultimately revealing that DMD can be viewed as a substep of SID. We then perform extensive computational studies, characterizing the performance of SID on problems of various model orders and noise levels. Specifically, we generate 10,000 random systems for each model order and noise level, calculating the average identification error for each case, and then repeat the entire experiment to ensure the results are, in fact, consistent. The results both quantify the intrinsic algorithmic error at zero-noise, monotonically increasing with model complexity, as well as demonstrate an asymptotically *linear* degradation to noise intensity, at least for the range under study. Finally, we close by demonstrating DMD's ability to recover system matrices, because its access to full state measurements makes them identifiable. SID, on the other hand, can't possibly hope to recover the original system matrices, due to their fundamental unidentifiability from input-output data. This is true even when SID delivers excellent performance identifying a correct set of equivalent system matrices.

Keywords: subspace identification, dynamic mode decomposition, state space model

ACKNOWLEDGMENTS

I'd like to thank Dr. Warnick and his continuous support throughout the entire process of writing this thesis. He is completely dedicated to the growth and success of his students. I have grown in the professional and researcher I am today in thanks in large part to his guidance. I would also like to thank my wonderful wife Natalie. She has never yielded in her support for me and my academic pursuits and I couldn't have done this without her support. Lastly, thanks to my Heavenly Father and the opportunities I have had at this amazing institution. BYU holds a dear place in my heart and I will always be grateful for all of the opportunities I have had while studying here.

Table of Contents

Chapter 1

In Preparation: Characterizing Equivalence and Correctness Properties of Dynamic Mode Decomposition and Subspace Identification Algorithms

This manuscript has not yet been accepted for publication.

Characterizing Equivalence and Correctness Properties of Dynamic Mode Decomposition and Subspace Identification Algorithms

Samuel Neff
Information and Decision Algorithms Laboratories
Brigham Young University
Provo, U.S.A
neff.samuelg@byu.edu

Sean Warnick
Information and Decision Algorithms Laboratories
Brigham Young University
Provo, U.S.A
sean@cs.byu.edu

***Abstract*—We examine the related nature of two identification algorithms, subspace identification (SID) and Dynamic Mode Decomposition (DMD), and their correctness properties over a broad range of problems. This investigation begins by noting the strong relationship between the two algorithms, both drawing significantly on the pseudoinverse calculation using singular value decomposition, and ultimately revealing that DMD can be viewed as a substep of SID. We then perform extensive computational studies, characterizing the performance of SID on problems of various model orders and noise levels. Specifically, we generate 10,000 random systems for each model order and noise level, calculating the average identification error for each case, and then repeat the entire experiment to ensure the results are, in fact, consistent. The results both quantify the intrinsic algorithmic error at zero-noise, monotonically increasing with model complexity, as well as demonstrate an asymptotically *linear* degradation to noise intensity, at least for the range under study. Finally, we close by demonstrating DMD's ability to recover system matrices, because its access to full state measurements makes them identifiable. SID, on the other hand, can't possibly hope to recover the original system matrices, due to their fundamental unidentifiability from input-output data. This is true even when SID delivers excellent performance identifying a correct set of equivalent system matrices.**

***Index Terms*—Subspace Identification, Dynamic Mode Decomposition, State Space Model**

I. Introduction

Various methods for identifying dynamic systems from data have developed somewhat independently from one another, resulting in a wide array of methods and world views around the topic. System Identification methods, developed in the controls community, are popular methods for identifying systems from data. The principle categories of System Identification methods include general autoregressive models, which are often defined as prediction error methods [14, 27], set-membership identification [30], and subspace identification (SID) [12]. This work focuses specifically on SID and it's non-autonomous variants as described in [18]. SID estimates a dynamic system by observing input-output data and generating a state sequence and observability matrix via the singular value decomposition. The state sequence and observability matrix are then used to estimate the system matrices of the observed system.

Dynamic Mode Decomposition (DMD) is an identification algorithm developed by the fluid dynamics community [13]. DMD provides a scalable approach to deal with high-dimensional systems including fluid flows, smart-grids, and financial markets [8]. In contrast to SID, DMD assumes full-state observability and estimates system matrices directly from the observed state sequence as opposed to generating the state sequence from input-output data.

Though SID and DMD developed independently of each other, in recent years research around both algorithms has highlighted important similarities and even equivalences between the two. Brunton et al. [29] described the relationship between DMD and the eigensystem realization algorithm, a form of SID. It was shown that the low-order linear operators central to each method are related by a similarity transform. More recently DMD was expanded to use time delay Hankel matrices in [4] and [15]. Time delay embedding is a powerful method for geometric reconstruction of attractors of nonlinear systems based on measurements of observables. This expansion of DMD allowed it to be effective at extracting information from the state sequence that cannot be computed by geometric reconstruction [4]. These variants of DMD are often referred to as Hankel DMD or Higher Order DMD. Hankel matrices are fundamental data matrices that are also used in various SID algorithms. Also of note is the result by Brunton et al. [20] which extends DMD to consider control input data as well and derive system matrices. Finally, Shin et al. [24] presented a unifying theorem for autonomous SID and DMD algorithms.

This work extends previous research in a few ways. First, this paper provides a reduction of the SID computation to the DMD computation, followed by oblique projection. This result follows from the observation that solutions to DMD algorithms are computed from the orthogonal projection via the psuedo inverse, while solutions to SID algorithms are computed from oblique projections. In practice, one can compute an orthogonal projection preceeding an oblique projection. In other words, DMD and SID methods share fundamental similarities in their calculations. The reduction of SID and the connection to DMD opens up paths of research into various directions including the analysis of Koopman operators and oblique projections, connections to Krylov subspace theory, and extensions to DMD that don't assume full-state feedback of systems.

Second, this work presents a concrete performance analysis of SID and DMD via the following experiments:

1) SID is applied to identification of fifth, tenth, and fifteenth-order systems with increasing levels of added noise. This highlights this algorithm class's ability to handle additive noise of increasing intensity.
2) SID and DMD are fitted to a series of random fifth-order systems. It is shown that DMD, with full-state feedback, is able to exactly identify the A and B system matrices. By contrast, SID is unable to return the exact system matrices A, B, C, D, but it is able to return state space models reliably given no state space data. This analysis sheds light on the differences in estimation methods, especially in regards to what information is available in the output data as opposed to the state sequence data. The eigenvalues and eigenvectors of SID and DMD are then analyzed for an example system. It is shown that SID cannot accurately identify the dynamics of the state transition matrix A, however, it can correctly extract the exact eigenvalues. This motivates a discussion around the utility of SID in systems with unobserved dynamics and other possible future research directions.

As a final note, all analyses done in this paper assume that the model order of the observed system is known, giving the algorithms the best possible chance to identify the unknown system.

II. Background

A. State Space Models

A stochastic linear system can be written as:

$$x_{k+1} = Ax_k + Bu_k + v_k \quad (1)$$
$$y_k = Cx_k + Du_k + w_k, \quad (2)$$

known as the process form, where $u_k \in \mathbb{R}^m$ are inputs, $x_k \in \mathbb{R}^n$ are the states, $y_k \in \mathbb{R}^l$ are the outputs, and $v_k \in \mathbb{R}^n$ and $w_k \in \mathbb{R}^l$ are i.i.d. random sequences modeling process noise and output measurement noise, respectively. The matrices A, B, C, and D are the system matrices characterizing the resulting stochastic processes [21].

If one knew the state matrices of such a system, one could then design a Kalman filter to estimate its state sequence:

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + K(y_k - C\hat{x}_k - Du_k), \quad (3)$$

where K is chosen to be the Kalman gain of the system. We define e_k as the system innovations:

$$e_k = y_k - C\hat{x}_k - Du_k, \quad (4)$$

and note that it is i.i.d. and independent of the input and output processes. Rearranging the Kalman filter equations, one can then define the innovation form of the original linear system

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + Ke_k \quad (5)$$
$$y_k = C\hat{x}_k + Du_k + e_k. \quad (6)$$

This form will be crucial to establishing the extended state space model, which is the starting place of the SID algorithm.

We now define Hankel matrices, which are important data structures for SID and DMD algorithms. Hankel matrices can be constructed in a few different ways, the main differences being the size of time delay and the staggering of past and future time periods. In this work we assume input-output data samples u_k, y_k for $k \in \{i, i+1, ..., j\}$, where $(j-i) > 0$ characterizes the sequence length. We can then arrange the data into Hankel matrices with s block rows, indicating the desired delay, and where $0 \leq s < (j-i)$, and $(j-s-i)$ columns of the form:

$$X := \begin{bmatrix} x_i & x_{i+1} & \dots & x_{j-s+1} \\ x_{i+1} & x_{i+2} & \dots & x_{j-s+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i+s-1} & x_{i+s} & \dots & x_j \end{bmatrix} = \quad (7)$$
$$\begin{bmatrix} \mathbf{x}_i & \mathbf{x}_{i+1} & \dots & \mathbf{x}_{j-s+1} \end{bmatrix}. \quad (8)$$

From this (original) Hankel matrix, past and future matrices can also be defined:

$$X_p := \begin{bmatrix} \mathbf{x}_i & \mathbf{x}_{i+1} & \dots & \mathbf{x}_{j-s} \end{bmatrix}, \quad (9)$$
$$X_f := \begin{bmatrix} \mathbf{x}_{i+1} & \mathbf{x}_{i+2} & \dots & \mathbf{x}_{j-s+1} \end{bmatrix}, \quad (10)$$

where p, and f denote past and future horizons respectively. X_p retains all the rows of X and all columns up to the penultimate, while X_f retains all the rows of X and all columns except for the first. This employs the concept of delay embedding of order s. Time delay embedding is a well established method for reconstruction of attractors for nonlinear systems; interested readers can also refer to Taken's classic theory [26].

Using the innovation form defined in (5) and (6), we also define the matrix input-output form, or extended state space model, as:

$$Y_f = \Gamma_i X_f + H_i U_f + G_i E_f, \quad (11)$$

with X_f defined as in (10) (and Y_f, U_f, and E_f defined analogously), and the toepliz matrices H_f and G_f defined as:

$$H_i = \begin{bmatrix} D & 0 & \dots & 0 \\ CB & D & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{i-2} & CA^{i-3} & \dots & D \end{bmatrix} \quad (12)$$

$$G_i = \begin{bmatrix} I & 0 & \dots & 0 \\ CK & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{i-2}K & CA^{i-3}K & \dots & I \end{bmatrix} \quad (13)$$

with the extended observability matrix Γ_f given by:

$$\Gamma_i = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{bmatrix} \quad (14)$$

as in [12].

B. Orthogonal and Oblique Projections

For this section, we generally define matrices $A \in \mathbb{R}^{p\times j}$, $B \in \mathbb{R}^{q\times j}$, and $C \in \mathbb{R}^{r\times j}$. The definition of these matrices and their dimensions are confined to this subsection. DMD and SID algorithms rely on notions of orthogonal and oblique projections to compute solutions to regression problems of the form

$$\min_{\Theta} ||A - \Theta B||_F^2, \quad (15)$$

where $||\cdot||_F$ is the Frobenius norm. The well-known solution is:

$$\hat{\Theta} = AB^\intercal(BB^\intercal)^{-1}. \quad (16)$$

The pseudo inverse of B is given by

$$B^\dagger = B^\intercal(BB^\intercal)^{-1} \quad (17)$$

where the † operator denotes the pseudo inverse. The psuedo inverse is a well known orthogonal projection operator.

We also define oblique projection as the projection of the row space of A along the row space of B on the row space of C

$$A/_B C = \quad (18)$$

$$A\begin{pmatrix} C^\intercal & B^\intercal \end{pmatrix}\left[\left(\begin{pmatrix} C \\ B \end{pmatrix}\begin{pmatrix} C^\intercal & B^\intercal \end{pmatrix}\right)^{-1}\right]C. \quad (19)$$

III. Dynamic Mode Decomposition

In this section we present Dynamic Mode Decomposition with control (DMDc), originally developed by Brunton et. al.[20].

A. Dynamic Mode Decomposition with Control

DMDc expands upon the original formulation of DMD by adding control inputs, giving the form

$$x_{k+1} = Ax_k + Bu_k. \quad (20)$$

Rewritten to include our definition of data matrices we get

$$X_f = AX_p + BU_p \quad (21)$$

or in terms of matrix multiplication

$$X_f = \begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} X_p \\ U_p \end{bmatrix} = G\Omega, \quad (22)$$

where Ω contains the past input and state data. DMDc seeks to find a best-fit solution over parameters G, solving a problem of the form

$$\min_G ||X_f - G\Omega||_F^2. \quad (23)$$

This is accomplished via orthogonal decomposition, calculated using the pseudo inverse of Ω, which is denoted as $\Omega^\dagger$. In practice, the pseudo inverse is calculated by way of the SVD, $\Omega = U\Sigma V^*$.

$$G = X_f\Omega^\dagger, \quad (24)$$

where $\Omega^\dagger := V\Sigma^{-1}U^*$ is the calculated pseudo inverse of Ω. The matrices A and B are then derived from the calculated G [20].

As we make continual references to DMD in this work, we are refering specifically to DMDc as defined above.

IV. Subspace Identification

A. Problem Formulation

SID algorithms seek to estimate system matrices A, B, C, and D that are equivalent (in a well defined sense) to those of the original process, given input-output data alone and where the input-output data is arranged into Hankel matrices. This involves solving a regression or projection problem in order to estimate a matrix O, where

$$O = Y_f/_{U_f}W_p = \Gamma_i X_f, \quad (25)$$

and

$$W_p = \begin{bmatrix} Y_p \\ U_p \end{bmatrix}. \quad (26)$$

To approximate O, one sets up a regression problem utilizing the matrix input-output form defined previously in Equation (11). For deterministic systems, this can be shortened to

$$\hat{Y}_f = \Gamma_i X_f + H_i U_f, \quad (27)$$

where $\hat{Y}_f$ is the best fit linear estimate given the limited data set. We restrict ourselves to the deterministic case in order to simplify the comparison with DMD.

It is well known that this equation can be rewritten as

$$\hat{Y}_f = L_p W_p + H_i U_f. \quad (28)$$

where X_f is replaced by a linear combination of past input-output data W_p, and $L_p = \begin{bmatrix} B & AB & ... & AB^{p-1} \end{bmatrix}$ [21].

Solving for O requires us to find matrices H_f and L_p. We therefore formulate the problem as

$$\min_{L_p H_i} ||Y_f - \begin{bmatrix} L_p H_i \end{bmatrix} \begin{bmatrix} W_p \\ U_f \end{bmatrix} ||_F^2 \tag{29}$$

minimizing L_p and H_i, in the least squares sense.

B. SID Algorithm

We now review the SID algorithm, which will give a solution to the minimization problem introduced in Equation (29).

1) **Arrange input-output data into Hankel matrices** Given inputs u_k and outputs y_k, one arranges this data into past and future hankel matrices as in Equation (3). We also define the combined input data as $Q = \begin{bmatrix} W_p \\ U_f \end{bmatrix}$.

2) **Calculate the orthogonal projection via the pseudo inverse of Q**
The solution to Equation (29) can be found by using the pseudo inverse of Q.
$$\begin{bmatrix} L_p H_i \end{bmatrix} = Y_f Q^\dagger \tag{30}$$
In practice, we are only concerned with the portion of Equation (29) that yeilds $\begin{bmatrix} L_p W_p \end{bmatrix}$ because it is equal to $\Gamma_i X_f$ which is equivalent to the matrix O. We can therefore utilize only the first r columns(where r is at least the number of rows in W_p) [28] of our pseudo inverse calculation
$$L_p = Y_f Q^\intercal \left[(QQ^\intercal)^{-1}\right]_{\text{first } r \text{ columns}} \tag{31}$$
or more succinctly
$$L_p = Y_f \tilde{Q}^\dagger, \tag{32}$$
where $\tilde{Q}$ denotes the trimmed version of the pseudo inverse.

3) **Calculate O via oblique projection** After calculating the initial solution via orthogonal projection, oblique projection can then be calculated by multiplying each side by W_p. This gives an oblique projection as described in Equation (19).
$$L_p W_p = Y_f Q^\intercal \left[(QQ^\intercal)^{-1}\right]_{\text{first } r \text{ columns}} W_p, \tag{33}$$
which is equivalent to
$$L_p W_p = Y_f /_{U_f} W_p = \Gamma_i X_f = O. \tag{34}$$

4) **Calculate the reduced extended observability matrix Γ_f and reduced future state sequence X_f via SVD of O**
After calculating O, Γ_i and X_f can be calculated via the SVD of O
$$O = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{35}$$
where
$$\Gamma_f = U_1 \Sigma_1^{1/2} \tag{36}$$
$$X_f = \Gamma_f^\dagger O. \tag{37}$$

5) **Calculate the State Space Parameters A, B, C, and, D using the reduced extended observability matrix and reduced future state sequence.**
Finally, with the knowledge of the estimated future state sequence X_f and the input-output data, the system matrices can be a approximated in a parameter estimation step.

C. Reducing the SID Algorithm to DMD and Oblique Projection

The principal goal of DMD is to find a best approximation of the linear operator G which describes the system dynamics, while the primary goal of SID is to identify system matrices with input-output dynamics that best fit the available data. However, examination of both algorithms shows an important equivalence. In Equation (32), L_p is calculated in an analogous manner to the calculation of G in Equation (24). The SID calculation is

$$L_p = Y_f \tilde{Q}^\dagger \tag{38}$$

and the DMD calculation is

$$G = X_f \Omega^\dagger. \tag{39}$$

The differences, of course, being the assumptions made about what data is available and how it is pre-processed. Though both algorithms define the data matrices in a distinct manner, they share orthogonal projection as a common step. Figure 1 depicts the reduction of the SID algorithm and its connection to DMD.

V. Analysis

The following section describes two different experiments to characterize the performance of each algorithm.

A. Example 1 - SID applied to a fifth-order system with increasing additive noise

In this example, SID was applied to SISO systems of order five, ten, and fifteen. Noise was added with increasing intensity, with the standard deviation of the distribution being the independent variable. This trial was run for 10000 different random systems at each order and noise level. The error was calculated as the difference between the magnitudes (dB) calculated from the bode plots of the true and derived systems. Figure 2 shows this analysis. The results exhibit two regimes. At the first addition of noise, error increases sharply. The error then levels off and begins to increase slowly at a linear level. When run many times, this analysis returns errors for each order and noise level that share consistent distributions across trials. The calculated difference in standard deviations for distributions from different trials were not statistically significant,

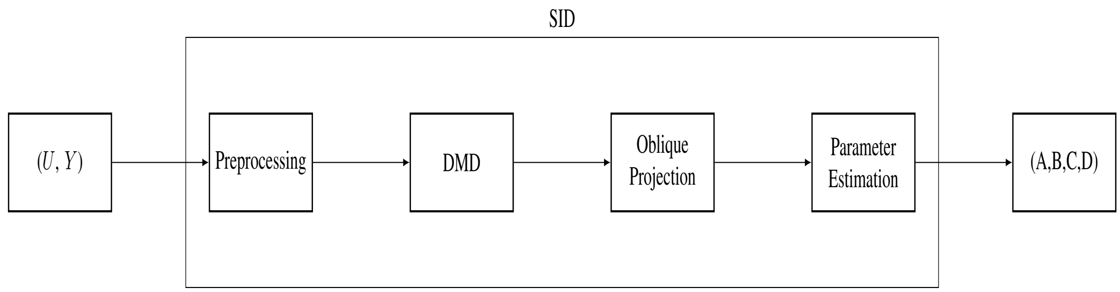


Figure 1: Subspace Identification

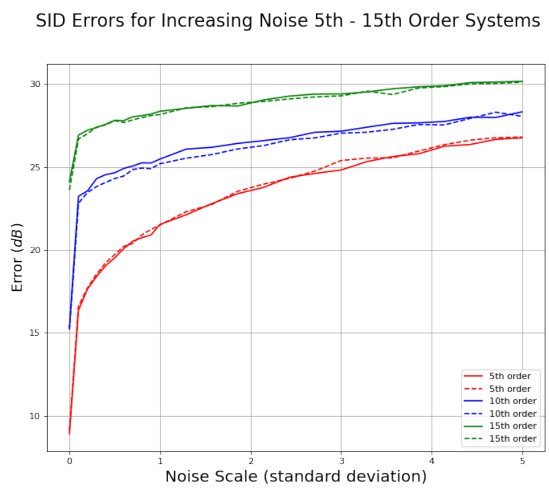


Figure 2: SID evaluated at 25 different noise levels for model orders 5, 10, and 15.

implying that these results are reproducible. Two example experiments were plotted to visually show the consistency between multiple trials.

The results both quantify the intrinsic algorithmic error at zero-noise, monotonically increasing with model complexity, as well as demonstrate an asymptotically *linear* degradation to noise intensity, at least for the range under study. Presumably, the error increases with model complexity likely due to machine precision or round-off error. At a high level, this analysis shows that SID and by extension DMD are sensitive to the first addition of noise, but are relatively robust to further increases in noise.

B. Example 2 - DMD and SID applied to 500 random fifth-order systems.

The next example explores how well each algorithm is able to return the exact system matrices of the original system from which the data was collected. An important fact to note here is that DMD assumes full-state feedback in order to correctly return the A and B matrices. With this in mind, we apply each algorithm to 500 random fifth-order LTI systems. DMD is given access to the full state sequence and the input while SID is given access only to the inputs and outputs. The system matrix error was determined by concatenating all state space matrices into a joint system matrix (A, B, C, D for SID and A, B for DMD), taking the Frobenius norm of the difference, and then normalizing by the Frobenius norm of the true aggregate system matrix

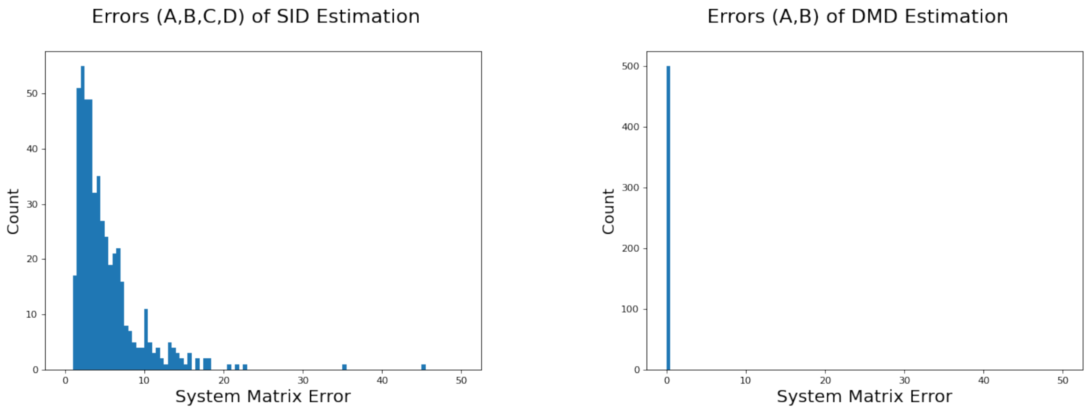
$$\frac{||P - \hat{P}||_F}{||P||_F}, \tag{40}$$

where P is the true aggregate system matrix and $\hat{P}$ is the estimated aggregate system matrix. The results, shown in Figure 3, verify that given full state feedback as in the DMD algorithm, the system matrices A and B are successfully identified. If the full state sequence is not observable as in SID, the state space model cannot be reliably approximated.

To further characterize each algorithms ability to recover the system matrices, a single example system out of the 500 was taken, and the eigendecompostions of the approximated state transition matrices calculated. This example illustrates an important result of this work, namely that without state sequence information, identification algorithms are able to correctly identify eigenvalues of the true system, but unable to identify eigenvectors. This is shown in Figure 4. The implications of this are that without a-prior knowledge of the mechanics of the underlying system, it is impossible to derive the true dynamics of the system. However, a strong advantage of SID is that even without the state-sequence of the true system, SID is able to recover the correct eigenvalues. This suggests that SID could be a strong alternative to DMD when applied to systems with partial or no observability. Additionally, the correct eigenvalues but incorrect eigenvectors suggest that the approximated A matrix may be equivalent to the true A up to a similarity transform. Given more knowledge about the true B or C matrices, the A matrix could be better approximated.

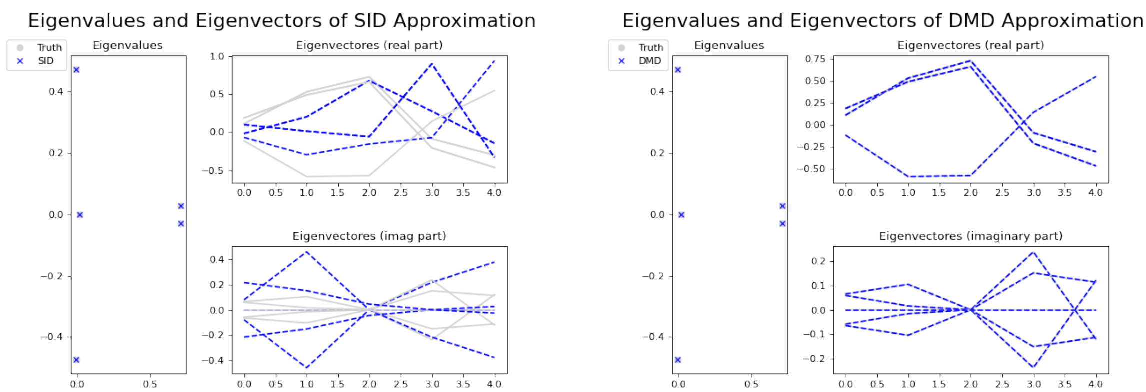
VI. Conclusion and Future Directions

In this paper we showed the relation between non-autonomous Dynamic Mode Decomposition and Subspace Identification. Namely, that DMD and SID share a the step of orthogonal decomposition via the pseudo inverse. We then conduct a computational study quantifying how the performance of Subspace Identification degrades with



(a) (b)

Figure 3: (a) System matrix error for the derived state space model from SID. The Frobenius norm of the error was taken and normalized by the Frobenius norm of the true system. (b) Error from A and B matrices of estimated model. The error was normalized and the the Frobenius norm was taken as in the SID example.



(a) (b)

Figure 4: (a) SID approximation of system identifies the correct eigenvalues, but incorrect eigenvectors. (b) DMD identifies the correct eigenvalues and eigenvectors of the true system.

model order and increasing levels of noise, illustrating two interesting points. First, even with zero noise, the algorithm's performance degrades with model order, presumably because of round-off error in the calculations. Second, the algorithm's performance then degrades linearly with increasing noise levels for a broad range of noise intensities.

Finally, we demonstrate that because DMD is applied to full state measurements, it can, in fact, recover the actual system matrices, (A, B), characterizing the data generation process. In contrast, although Subspace Identification correctly identifies an equivalent realization of the data generation process, we illustrate that it does not recover the system matrices, (A, B, C, D). This is expected since the system matrices are unidentifiable from input-output data alone, but sometimes this fact gets lost in the discussion of Subspace Identification since it ultimately delivers state matrices characterizing the unknown system.

This analysis of DMD and SID is nascent and many avenues are open for future research. Some of the following are ideas that have developed out of this work:

- **Extend SID to utilize proper orthogonal decomposition or balanced truncation when calculating the orthogonal projection for further model reduction.** While DMD leverages orthogonal projections to calculate a best-fit linear system A in the least-squares sense, proper orthogonal decomposition (POD) was recognized as an effective way to reduce the order of the system during estimation [13]. This significantly reduces the order of the system while still calculating the retained eigenvalues accurately as the reduced order $\tilde{A}$ (derived from the approximated G) still has the same non-zero eigenvalues as the unreduced A. In practice, SID algorithms utilize the QR decomposition to efficiently calculate oblique projections. However, more recent work such as [17, 11] has explored algorithms that use POD for efficiently calculating oblique projections which may or may not employ the QR decomposition. Future iterations of SID algorithms could benefit from utilizing POD in their calculations.

- **Analyze SID as a means of identifying dominant dynamics in systems with only partial observability.** DMD assumes full-state feedback of observables. SID algorithms do not have such limitations as they estimate the state sequence directly from data. SID has the potential to succeed where DMD has not in the past, especially with systems without full-state feedback. An interesting analysis will be to apply this method to high-dimensional systems of partial observables, or even completely unobservable states.
- **Explore the implications of utilizing oblique projections vs orthogonal projections in regards to Koopman operator theory.**
  It is well known that DMD approximates the Koopman operator. The theory behind Koopman operators as estimated by DMD was discovered as a result of the equivalence between DMD and the Arnoldi decomposition [29]. This connection to Arnoldi decompositions and Kyrlov subspaces is well established for DMD, but is not yet explored well for SID methods. Future research could focus on exploring the theory behind oblique projections and if the oblique projections calculated by SID has any strong connection to Koopman operator theory.

REFERENCES

[1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.

[2] Carl Andersson. *Deep Learning Applied to System Identification: A Probabilistic Approach*. PhD thesis, Uppsala University, 12 2019.

[3] Jennifer Annoni, Pieter Gebraad, and Peter Seiler. Wind farm flow modeling using an input-output reduced-order model. In *American Control Conference*, Boston, MA, USA, 2016.

[4] Hassan Arbabi and Igor Mezić. Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the koopman operator. *SIAM Journal on Applied Dynamical Systems*, 16(4):2096–2126, 2017. doi: 10.1137/17m1125236.

[5] K. J. Astrom and P. Eykhoff. System identification—a survey. *Automatica*, 7(2):123–162, 1971. doi: https://doi.org/10.1016/0005-1098(71)90059-8.

[6] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley and Sons, Hoboken, New Jersey, 2016.

[7] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Compressive sampling dynamic mode decomposition. *submitted for publication, arXiv*, 2013.

[8] Steve L. Brunton and J. Nathan Kutz. *Data-Driven Science and Engineering*. Cambridge University Press, New York, NY, 2019.

[9] Carlton Downey. *Deep Learning Applied to System Identification: A Probabilistic Approach*. PhD thesis, Carnegie Mellon University, 4 2019.

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

[11] Andrea Iannelli, Urban Fasel, and Roy S.Smith. The balanced mode decomposition algorithm for data-driven lpv low-order models of aeroservoelastic systems. *Aerospace Science and Technology*, 115, 2021.

[12] Tohru Katayama. *Subspace Methods for System Identification*. Springer Science, London, 2005.

[13] J. Nathan Kutz, Steven L. Brunton, Bingni W. Brunton, and Joshua L. Proctor. *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*. SIAM, first edition, 2016.

[14] Lennart Ljung. *Sytem Identification Theory for the User*. Pearson, Upper Saddle River, New Jersey, 1998.

[15] Vega José Manuel and Le Soledad Clainche. *Higher Order Dynamic Mode decomposition and its applications*. Academic Press, an imprint of Elsevier, 2021.

[16] L. Mirsky. Symmetric gauge functions and unitarily invariant norms. *Quart. J. Math. Oxford*, 11:50–59, 1960.

[17] Samuel E. Otto, Alberto Padovan, and Clarence W. Rowley. Optimizing oblique projections for nonlinear systems using trajectories. *submitted for publication, arXiv*, 2022.

[18] Peter Van Overschee and Bart De Moor. A unifying theorem for three subspace system identification algorithms. *Automatica*, 31(12):1853–1864, 1995.

[19] Gianluigi Pillonetto, Francesco Dinuzzo, Tianshi Chen, Giuseppe De Nicolao, and Lennart Ljung. Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50(3):657–682, 2014. doi: https://doi.org/10.1016/j.automatica.2014.01.001.

[20] Joshua L. Proctor, Steven L. Brunton, and J. Nathan Kutz. Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems*, 15(1): 142–161, 2016.

[21] S. Joe Qin. An overview of subspace identification. *Computers amp; Chemical Engineering*, 30(10-12):1502–1513, 2006. doi: 10.1016/j.compchemeng.2006.05.045.

[22] C.W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D.S. Henningson. Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics*, 645(115):1–6, 2009.

[23] P. J. Schmid, K. E. Meyer, and O. Pust. Dynamic mode decomposition and proper orthogonal decomposition of flow in a lid-driven cylindrical cavity. In *PIV09-0186, 8th International Symposium on Particle Image Velocimetry*, 2009.

[24] Sungho Shin, Qiugang Lu, and Victor M. Zavala. Unifying theorems for subspace identification and dynamicmode decomposition. *Automatica (preprint)*, 2020.

[25] L. Sirovich. Turbulence and the dynamics of coherent structures, partsi-iii. *Q. Appl. Math*, XLV(3):561–590, 1987.

[26] Floris Takens. Detecting strange attractors in turbulence. *Lecture Notes in Math*, 898:368–381, 1981.

[27] Arun K. Tangirala. *Principles of System Identification*. CRC Press, Boca Raton, Florida, 2015.

[28] Pavel Trnka and Vladimir Havlena. Subspace identification as multi-step predictions optimization. *ACTA Press*, page 223–228, 2005.

[29] Jonathan H. Tu, Clarence W. Rowley, Dirk M. Luchtenburg, Steven L. Brunton, and J. Nathan Kutz. On dynamic mode decomposition: Theory and applications. *Journal of computational dynamics*, 2013.

[30] Saligrama Ramaswamy Venkatesh. *System-Identification for Complex Systems*. PhD thesis, MIT, 6 1997.