Data-driven methods for identifying nonlinear models of fluid flows

Clancy Rowley

12 January 2017

PRINCETON UNIVERSITY

Goals

- Determine models of dynamical systems directly from data.
- Use structure of known governing equations when helpful.
- Apply this to turbulence? Well, at least fluids.

Acknowledgments:

- Scott Dawson (Caltech)
- Matt Williams (UTRC)
- Yannis Kevrekidis (Princeton)
- Funding from AFOSR, DARPA

Outline

Motivation and overview
- Jet in crossflow
- Dynamic Mode Decomposition and the Koopman operator

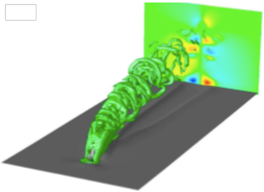
Data-driven approximations of the Koopman operator
- Data-driven inner product
- Determining the projected Koopman operator
- Example: two-dimensional map
- Example: basins of attraction in the Duffing equation

Determining nonlinear models from data
- Choice of observables
- Example: flow past a cylider
- Energy-conserving constraints

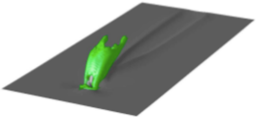
Example: jet in crossflow

Linearize a jet in crossflow about an unstable equilibrium.[1]

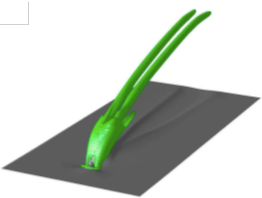
($Re_{\delta_0^*} = 165$, $V_{\mathbf{jet}}/U_\infty = 3$, $\delta_0^*/D = 1/3$)



Instantaneous snapshot



Mean



Unstable equilibrium

Compute eigenvalues and compare with observed frequencies:

	Observed	Linear theory
Shear layer	St = 0.141	St = 0.169
Near wall	St = 0.0174	St = 0.043

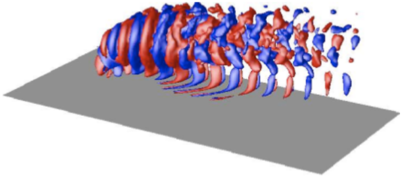
Frequency mismatch for near-wall structures: failure of linear theory.

[1] Bagheri, Schlatter, Schmid, Henningson, JFM 2009

Dynamic Mode Decomposition for jet in crossflow

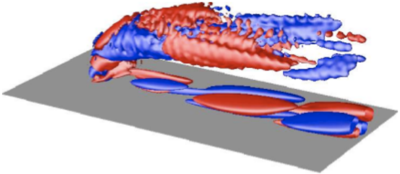
- Dynamic Mode Decomposition (DMD) modes capture relevant structures and frequencies

High-frequency mode captures structures in the shear layer.



St = 0.141

Low-frequency mode captures near-wall structures associated with horseshoe vortex.



St = 0.017

Main point: can use this method to separate the structure from the randomness in a turbulent flow.

Many applications of DMD in fluids[2]

Study	Applications, findings, and variants
Rowley et al. (2009)	Jet in crossflow (DNS)
Schmid (2010)	Plane Poiseuille flow; linearized two-dimensional flow over a square cavity; wake of a flexible membrane (PIV); jet between two cylinders (PIV)
Chen et al. (2011)	Transitional cylinder flow (DNS)
Nastase et al. (2011)	Lobed jet from three-dimensional diffusers (experiment)
Pan et al. (2011)	Wake of a NACA 0015 airfoil with Gurney flap (PIV)
Schmid et al. (2011)	Schlieren snapshots of a helium jet; PIV snapshots of an acoustically forced jet
Schmid (2011)	Passive tracer in flame simulation and axisymmetric water jet experiment
Seena & Sung (2011)	Turbulent cavity flow (DNS)
Duke et al. (2012a)	Annular liquid sheet instabilities (experiment)
Grilli et al. (2012)	Shockwave turbulent boundary layer interaction (DNS)
Jardin & Bury (2012)	Flow past a cylinder, with forcing near the mean separation point (DNS)
Lee et al. (2012)	Developing turbulent boundary layers over roughened walls (DNS)
Muld et al. (2012a)	Wake of high-speed train model (detached eddy simulation)
Muld et al. (2012b)	Flow over a surface-mounted cube (detached eddy simulation)
Schmid et al. (2012)	Transitional water jet with tomographic PIV
Semeraro et al. (2012)	Confined turbulent jet with coflow (PIV)
Bagheri (2013)	Cylinder wake approaching limit cycle (DNS)
Ghommem et al. (2013)	Flows in high-contrast porous media (DNS)
He et al. (2013)	Boundary layer and cylinder configuration (experiment)
Meslem et al. (2013)	Impinging circular jet (PIV)
Motheau et al. (2013)	Gas turbine combustion instability (LES)
Sarkar et al. (2013)	Nanofluid flow past a square cylinder (DNS)
Tu et al. (2013, 2014b)	Wake of a cylinder (DNS) and finite-thickness flat plate (PIV)
Wynn et al. (2013)	Flow over a backward-facing step (PIV) using optimal mode decomposition
Carlsson et al. (2014)	Flow-flame interactions (LES)
Gómez et al. (2014)	Turbulent pipe flow (DNS)
Jovanović et al. (2014)	Sparsity-promoting DMD applied to two-dimensional plane Poiseuille flow; screeching supersonic jet (LES); jet between two cylinders (PIV)
Ma & Liu (2014)	Flow over high angle of attack, slender bodies (DNS)
Markovich et al. (2014)	Swirling, confined flames and jets (PIV)
Tu et al. (2014a)	Flow past a cylinder (PIV), temporally sparse data
Sarmast et al. (2014)	Wind turbine wakes (LES)
Sayadi et al. (2014)	Flat plate boundary layer transition to turbulence (DNS and LES)
Subbareddy et al. (2014)	Transition of Mach 6 boundary layer with roughness element (DNS)
Thompson et al. (2014)	Flow past elliptic cylinders (DNS)
Tissot et al. (2014)	Flow past a cylinder (experiment), mode extraction for reduced-order modeling
Dunne & McKeon (2015)	Dynamic stall on a pitching and surging airfoil (PIV)
Kramer et al. (2015)	Flow in a two-dimensional differential heated cavity (DNS), for identification of flow regimes
Roy et al. (2015)	Reacting flows behind bluff bodies (experiment)
Sayadi et al. (2015)	Thermo-acoustic instabilities in ducted and bifurcating flames (numerical and experimental), using parameterized DMD

[2] Rowley and Dawson, *Annual Rev. Fluid Mech.*, 2017.

DMD modes and the Koopman operator

Dynamic Mode Decomposition (DMD) is a method for approximating eigenvalues and eigenfunctions of linear dynamics, given snapshots sampled from the system.[3]
It turns out that the DMD modes shown on the previous slide are related to a linear operator called the Koopman operator.[4]

- Consider a state space M, with discrete-time dynamics given by a map $T : M \to M$.
- Let V be a vector space of functions from M to $\mathbb{C}$. We call elements of V *observables*. A measurement of a state $x \in V$ consists of the value $f(x)$ of a particular observable $f \in V$.
- The Koopman operator is an operator $U : V \to V$, defined by

  $$(Uf)(x) = f(Tx).$$

  That is, U maps a function f to another function Uf.
- U is a linear operator: for any $f, g \in V$,

  $$U(f+g)(x) = (f+g)(Tx) = f(Tx) + g(Tx) = Uf(x) + Ug(x).$$

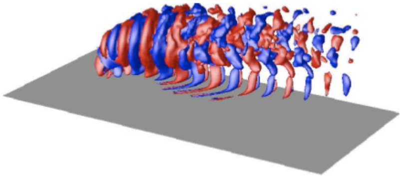
[3] Schmid. *J. Fluid Mech.* 2010.
[4] Rowley, Mezic, Bacheri, Schlatter, and Henningson. *J. Fluid Mech.* 2009.

Why is this useful?

- Eigenfunctions of the Koopman operator determine coordinates in which a system evolves linearly
  - For dynamics given by a nonlinear system $x(k+1) = T(x(k))$, we have
  $$Uf(x) = f(Tx).$$
  - Suppose we have an eigenfunction of the linear operator U:
  $$U\varphi = \lambda\varphi.$$
  - Define a new coordinate $z(k) = \varphi(x(k))$. Then z evolves as
  $$z(k+1) = \varphi(x(k+1)) = \varphi(Tx(k)) = U\varphi(x(k)) = \lambda\varphi(x(k)) = \lambda z(k).$$
  - The evolution is linear! An eigenfunction represents a "structured" part of the nonlinear dynamics.
  - If U has enough eigenfunctions so that we can reconstruct the state x from the values of the eigenfunctions, then there is a coordinate change in which the system is linear. (However, for chaotic systems, there is not a "full set" of eigenfunctions.)

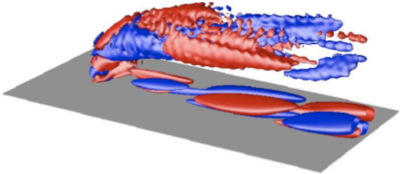
Jet in crossflow revisited

High-frequency mode evolves according to $e^{i\omega_1 t}$



$\mathrm{St}_1 = 0.141$

Low-frequency mode evolves according to $e^{i\omega_2 t}$.



$\mathrm{St}_2 = 0.017$

Outline

Data-driven approximations of the Koopman operator

We will consider a new approach to data-driven approximations of the Koopman operator:

- A data-driven inner product
- A subspace S spanned by a set of observables
- A projection theorem $V = S \oplus S^{\perp}$
- Approximate U by projection onto the subspace S

Spoiler: In the end, the numerical method we obtain is the same as DMD (actually, Extended DMD[5]). But the path we take to get there is different.

Why a new path?

- Derivation is more natural, less ad hoc
- We will be able to say more about the correspondence between DMD and Koopman.

[5] Williams, Kevrekidis, and Rowley, *J. Nonlinear Sci.*, 2015.

Goal

As before:

- State space M, with dynamics given by $T : M \to M$.
- V is the vector space of functions from M to $\mathbb{C}$, called *observables*.
- The operator $U : V \to V$ is given by $Uf(x) = f(Tx)$.

Now, suppose we are given observables $f_1, \dots, f_n \in V$. At a particular state $x \in M$, our measurements consist of the n values $f_j(x)$.

Goal

Determine an approximation of U directly from data sampled from the system, without explicit knowledge of T or U. In particular, we will obtain the projection of U onto the subspace spanned by $\{f_1, \dots, f_n\}$.

A data-driven inner product

- To determine a projection, we would like some additional structure on the function space V (e.g., an inner product).
- Here, we will not assume any structure on V *a priori*. Instead, we will define structure based on some available data.
- Suppose we have sample points $x_1, x_2, x_3, \ldots, x_m \in M$. For any functions $f, g \in V$, define

$$\langle f, g \rangle = \frac{1}{m} \sum_{k=1}^{m} f(x_k) \overline{g(x_k)}.$$

- If M is a probability space, with probability measure μ, and the points x_k are drawn at random with probability μ, then by the law of large numbers, as $m \to \infty$,

$$\langle f, g \rangle \to \int_M f \bar{g} \, d\mu,$$

  which is the usual inner product on $L^2(M, \mu)$.
- Similarly, the above holds if the points x_k are sampled from a measure-preserving dynamical system $x_{k+1} = Tx_k$ and T is ergodic.

The data

The data we use consist of values of our observables $f_1, \ldots, f_n$ at the given sample points $x_1, \ldots, x_m$. Collect the data into an $m \times n$ matrix

$$X = \begin{bmatrix} f_1(x_1) & \cdots & f_n(x_1) \\ \vdots & & \vdots \\ f_1(x_m) & \cdots & f_n(x_m) \end{bmatrix}.$$

Let

$$S = \operatorname{span}\{f_1, \ldots, f_n\}, \qquad S^\perp = \{g \in V : \langle f, g \rangle = 0\},$$

using our inner product just defined.

Note that $\operatorname{rank} X$ cannot be more than $\dim S$. If $\operatorname{rank} X = \dim S$, have some useful properties:

- $\langle \cdot, \cdot \rangle$ is a strictly positive-definite inner product on S.
- $V = S \oplus S^\perp$. That is, any $f \in V$ may be written uniquely as a sum of a function in S and a function in $S^\perp$.

This latter property lets us define a projection $P : V \to S$.

Henceforth, we shall always assume $\operatorname{rank} X = \dim S$.

(If not true, gather more data.)

Main result: projected Koopman operator

- The data: In addition to the data $f_j(x_k)$, assume we also have measurements $f_j(Tx_k)$ (i.e., at the following "timestep"). Define matrices
$$X_{kj} = f_j(x_k), \qquad X^{\#}_{kj} = f_j(Tx_k).$$
- The subspace: $S = \operatorname{span}\{f_1, \ldots, f_n\}$.
  - Assume $\operatorname{rank} X = \dim S$
  - Let $P : V \to S$ denote the projection onto S.
  - Define a map $F : \mathbb{C}^n \to S$ by
$$F(v) = \sum_{j=1}^{n} v_j f_j, \qquad v = (v_1, \ldots, v_n).$$

Theorem (Data-driven projection)

Let $A = X^+ X^{\#}$. Then, for any $v \in \mathbb{C}^n$,
$$PU\,F(v) = F(Av).$$

That is, A is the matrix representation of the projection $PU : S \to S$.

Projected Koopman operator

According to the theorem, as long as $\operatorname{rank} X = \dim S$, we have

$$PU\,F(v) = F(Av).$$

In other words, the following diagram commutes:

$$\begin{array}{ccc} S & \xrightarrow{PU} & S \\ {\scriptstyle F}\uparrow & & {\scriptstyle F}\uparrow \\ \mathbb{C}^n & \xrightarrow{A} & \mathbb{C}^n \end{array}$$

- The matrix $A = X^+ X^\#$ is determined solely from the data. We did not need to know the map T or the operator U.
- The projected Koopman operator PU is computed exactly, without approximation.

Connection with Dynamic Mode Decomposition (DMD)

- The matrix $A = X^+ X^\#$ turns out to be identical to the matrix computed in Extended Dynamic Mode Decomposition[6].
  - In that work, it was shown that A corresponds to an approximation of U by a weighted residual method, with a particular choice of test functions.
  - Here, we see that A arises naturally as a Galerkin method (orthogonal projection onto the subspace S), with a natural choice of inner product.
- "Standard" Dynamic Mode Decomposition (DMD) is a special case of this:
  - If the state is $x = (x_1, \ldots, x_n)$, one considers the observables $f_j(x) = x_j$ (the "full state observable").
  - The DMD eigenvalues are then the eigenvalues of our matrix A, and the DMD modes are the left eigenvectors of A.

[6] Williams, Kevrekidis, and Rowley, *J. Nonlinear Sci.*, 2015.

What if an eigenfunction is in S?

$$\begin{array}{ccc} S & \xrightarrow{PU} & S \\ F\uparrow & & \uparrow F \\ \mathbb{C}^n & \xrightarrow{A} & \mathbb{C}^n \end{array}$$

Corollary 1

Suppose $U\varphi = \lambda\varphi$ for some $\varphi \in S$. Then there is a $v \in \mathbb{C}^n$ such that $Av = \lambda v$ and $\varphi = F(v)$.

So if a Koopman eigenfunction lies in the subspace $S = \text{span}\{f_1, \ldots, f_n\}$, there will be a corresponding eigenvalue and eigenvector of A.
(This is a restatement of a previously known result[7].)

[7] Tu, Rowley, Luchtenburg, Brunton, and Kutz. *J. Comput. Dyn.*, 2014.

What if S is invariant?

Corollary 2

Suppose S is invariant under U (i.e., $Uf \in S$ whenever $f \in S$). If $Av = \lambda v$, then $U\varphi = \lambda\varphi$, with $\varphi = F(v)$.

- If S is invariant, then any eigenvalue of A will correspond to a Koopman eigenvalue, and $\varphi = F(v)$ will be a Koopman eigenfunction (provided φ is nonzero).
- It is helpful to compare this with a recent result[2], which considered the special case that T is ergodic, the observables are $\{f, Uf, U^2f, \dots, U^{n-1}f\}$, and the sample points are $x_1, x_2, \dots, x_m$ with $x_{k+1} = Tx_k$. The authors showed that if S is invariant, then in the limit $m \to \infty$, the eigenvalues/eigenfunctions determined by DMD converge to Koopman eigenvalues/eigenfunctions.
- Corollary 2 strengthens this in several ways: the eigenvalues and eigenfunctions are computed exactly with only a finite amount of data (need only $\operatorname{rank} X = \dim S$), and T need not be ergodic.

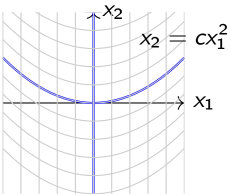
[2] Arbabi & Mezic, arXiv:1611.06664, 2016.

Example: two-dimensional map

Consider the map

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} \lambda x_1 \\ \mu x_2 + (\lambda^2 - \mu) c x_1^2 \end{bmatrix}.$$

This system has an equilibrium at the origin, and invariant manifolds given by $x_1 = 0$ and $x_2 = c x_1^2$:



Koopman eigenvalues are λ, μ with eigenfunctions

$$\varphi_\lambda(\mathbf{x}) = x_1$$
$$\varphi_\mu(\mathbf{x}) = x_2 - c x_1^2.$$

In addition, φ_λ^k is an eigenfunction with eigenvalue λ^k, the product $\varphi_\lambda \varphi_\mu$ is an eigenfunction with eigenvalue $\lambda\mu$, etc.

DMD for two-dimensional map

Apply DMD to this example, with initial states $\mathbf{x}$ given by $(1,1),(5,5),(-1,1),(-5,5)$, with $\lambda = 0.9$, $\mu = 0.5$.

- Case 1: observables $\mathbf{f}(\mathbf{x}) = (x_1, x_2)$. If $c = 0$, so that the problem is linear, then DMD eigenvalues are 0.9 and 0.5: good!
  If $c = 1$, however, then the DMD eigenvalues are 0.9 and 2.002. These do not correspond to Koopman eigenvalues, and one might even presume the equilibrium is *unstable*!
- Case 2: observables $\mathbf{f}(\mathbf{x}) = (x_1, x_2, x_1^2)$. (Note: the subspace $S = \text{span}\{x_1, x_2, x_1^2\}$ is now invariant.) The DMD eigenvalues are 0.9, 0.5, and $0.81 = 0.9^2$, which agree with Koopman eigenvalues.
- Case 3: observables $\mathbf{f}(\mathbf{x}) = (x_1, x_2, x_2^2)$. Now, the DMD eigenvalues are 0.9, 0.822, and 4.767. The eigenvalues do not correspond to Koopman eigenvalues because the Koopman eigenfunction φ_μ is not in the span of the observables, and the subspace S is not invariant.

Choice of observables

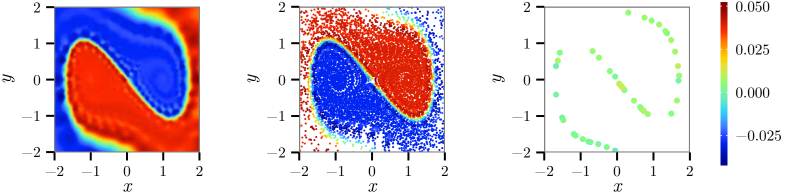
- The previous example illustrates that it is critical to choose an appropriate set of observables $f_1, \ldots, f_n$.
- Some possible choices:
  - Orthogonal basis functions (e.g., Fourier modes, Chebyshev polynomials, Legendre polynomials,...)
  - Indicator functions on small subsets (Ulam's method)
  - Spectral elements
  - Time delay coordinates $(f(x), f(Tx), f(T^2x), \ldots, f(T^{n-1}x))$ ("Takens embedding")
  - Radial basis functions
  - [your idea here!]

Example: basins of attraction in the Duffing equation

- Consider the Duffing equation

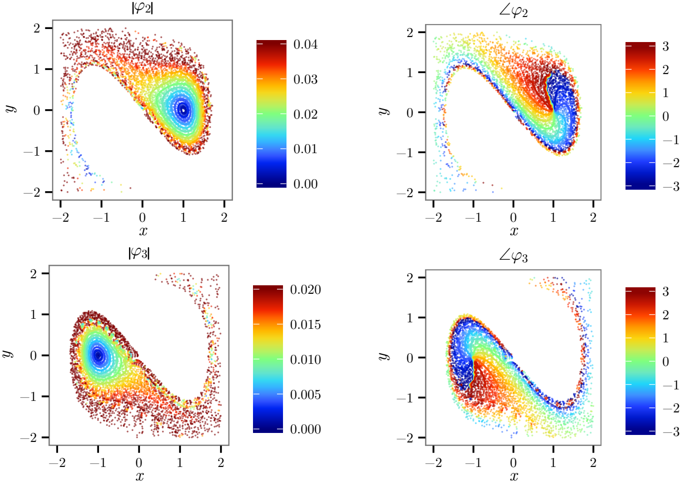
$$\ddot{x} + \delta\dot{x} + x(x^2 - 1) = 0$$

- Compute approximation of Koopman operator (with $\delta = 0.5$):
  - Data: 10^3 trajectories with 11 samples each, sampling interval $\Delta t = 0.25$
  - Basis functions: 1000 radial basis functions (thin plate splines)
- $\lambda_0 = -10^{-14}$: corresponding eigenfunction is the constant function
- $\lambda_1 = -10^{-3}$: eigenfunction reveals basins of attraction



Dynamics in each basin

- $\lambda_2 = -0.237 + 1.387i$ (analytically $-0.250 + 1.392i$)



Outline

Determining nonlinear models from data

- The use of data-driven approximations to the Koopman operator shows promise for describing and modeling nonlinear systems.
- Can we use this method to extract nonlinear reduced-order models from data?
- Can we incorporate known properties of the governing equations into these models (e.g., quadratic nonlinearities, energy conservation)?

What are we trying to model?

- While data-driven methods can be powerful, they often do not make full use of what we know about a system.
- Assume that our system is described by the Navier-Stokes equations:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla p + \nu \Delta \mathbf{u}, \qquad \nabla \cdot \mathbf{u} = 0.$$

- How can this help us?
  - Guiding the choice of observables.
  - Enforcing conservation properties.

Choice of observables

- Consider the projection of the velocity field onto a set of basis functions $\mathbf{u}_i$ (e.g., POD modes):

$$\mathbf{u}(x,t) = \mathbf{u}_0(x) + \sum_{i=1}^{N} \mathbf{u}_i(x) a_i(t)$$

- Projecting the governing equations onto these basis functions (under certain assumptions) gives

$$\dot{\mathbf{a}} = \mathbf{L}\mathbf{a} + \mathbf{B}(\mathbf{a},\mathbf{a}),$$

  where $\mathbf{L}$ is linear and $\mathbf{B}$ is bilinear.

- It is hence reasonable to choose observables that include (at least) monomials of POD coefficients, up to second order

Data-driven models

- Suppose we have collected pairs of snapshots of data $(\mathbf{y}_k, \mathbf{y}_k^{\#})$, which are separated by a fixed time interval Δt
- Choose as observables:

$$\mathbf{f}(\mathbf{y}) = \begin{bmatrix} \mathbf{a} \\ \mathrm{vec}(\mathbf{a} \otimes \mathbf{a}) \end{bmatrix}$$

- We seek the discrete propagation matrix (finite-dimensional approximation to the Koopman operator) $\mathbf{A}$ such that

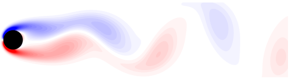
$$\mathbf{f}(\mathbf{y}_k^{\#}) = \mathbf{A}\mathbf{f}(\mathbf{y}_k).$$

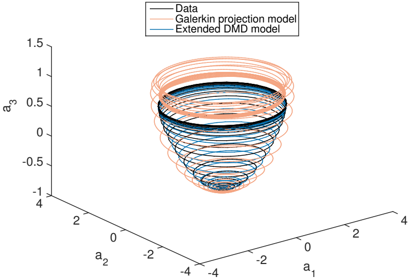
- This may be obtained from the data by

$$\mathbf{A} = \left[\mathbf{f}(\mathbf{y}_1^{\#})\ \mathbf{f}(\mathbf{y}_2^{\#}) \cdots \mathbf{f}(\mathbf{y}_m^{\#})\right] \left[\mathbf{f}(\mathbf{y}_1)\ \mathbf{f}(\mathbf{y}_2) \cdots \mathbf{f}(\mathbf{y}_m)\right]^{+}$$

Example: flow past a cylinder

- Collect data between the unstable equilibrium and limit cycle of the system
- The method works well, and often outperforms Galerkin projection of the governing equations onto POD modes.

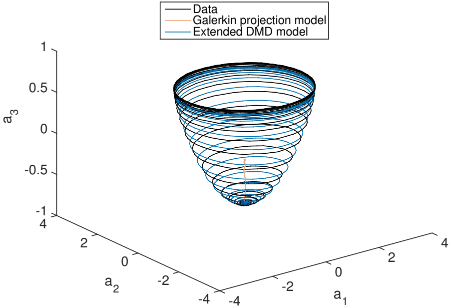




Cylinder example: noisy data

What if the data are noisy?

Data corrupted with Gaussian white noise with standard deviation $\sigma = 0.05U$

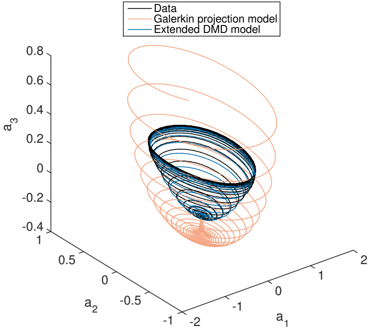


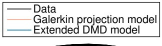
Cylinder example: limited data

What if we only have access to a limited amount of data?

Spatially limited data:

Temporally limited data:





Energy-conserving constraints

- The Navier-Stokes equations may be written

$$\partial_t \mathbf{u} = L\mathbf{u} + B(\mathbf{u}, \mathbf{u}),$$

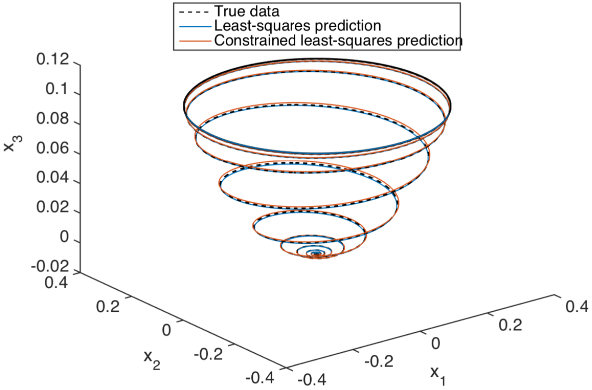
  where L is linear and B is bilinear.
- The quadratic terms satisfy

$$\langle B(\mathbf{u}, \mathbf{u}), \mathbf{u} \rangle = 0.$$

- Can we impose this constraint on the data-driven modeling procedure?
  - Yes. Solve the constrained optimization problem explicitly using Lagrange multipliers.

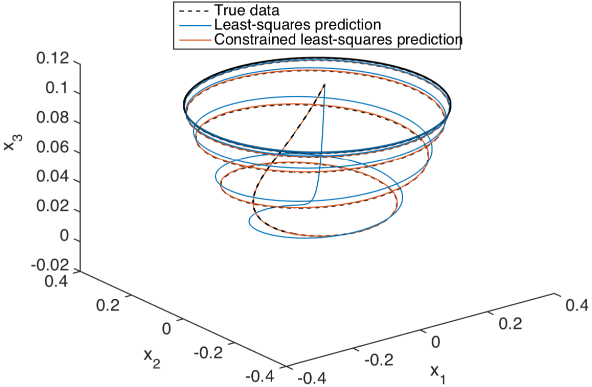
Energy-conserving constraint for cylinder flow

- Predicting the evolution of the system with identified models
- Take initial condition used for the system identification dataset



Energy-conserving constraint for cylinder flow

- Take initial condition away from the system identification dataset



Incorporating the energy-conserving constraint leads to more robust models.

Summary

- Goal: identify structure in dynamics, directly from data
- Can determine a projection of the Koopman operator from data
  - Data-driven inner product determined by sample points $x_1, \dots, x_m$.
  - Subspace S determined by chosen observables $f_1, \dots, f_n$.
  - Data determines an (exact) matrix representation of the projection of the Koopman operator onto this subspace
  - The matrix is the same as that determined by Extended DMD
- Under certain conditions, eigenvalues/eigenvectors of this matrix correspond precisely to Koopman eigenvalues/eigenfunctions
  - Eigenfunction $\varphi \in S$
  - The subspace S is invariant under U.
- Choice of observables is critical.
- Some success using these methods to determine nonlinear models of "simple" (non-turbulent!) fluid flows