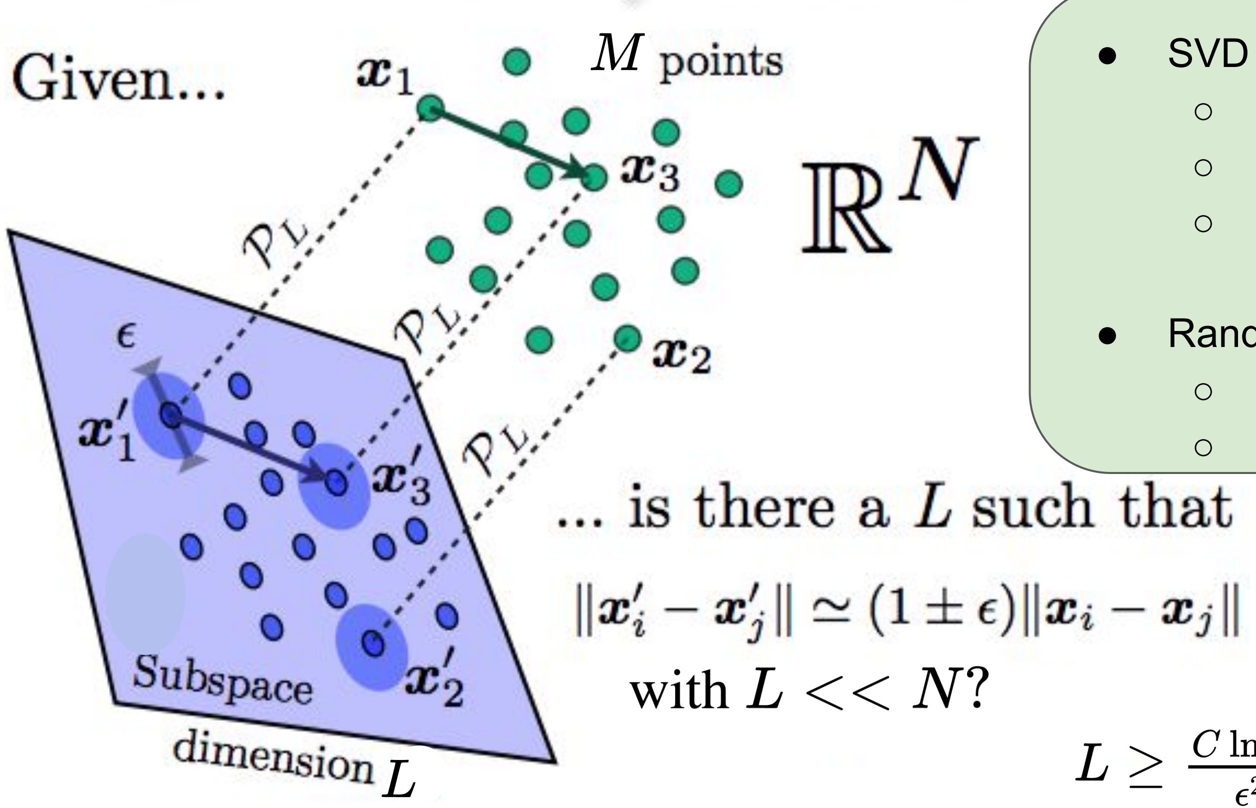
Learning a Reduced Order Dynamic Mode Decomposition by Random Observable Features

Sudam Surasinghe[1], Erik Bollt[2].

[1] Department of Mathematics, [2] Department of Electrical and Computer Engineering, Clarkson Center for Complex Systems Science (C^{3}S^2),Clarkson University, Potsdam, NY

Presented in, CP3: Data and Computation, SIAM DS21.

Rework DMD through Johnson-Lindenstrauss theorem and Random Projection



- SVD based algorithms
  - need more memory
  - Computationally expensive
  - Optimal Projection

- Random Projection
  - Just matrix multiplication.
  - High quality

$$L \geq \frac{C \ln M}{\epsilon^2}$$

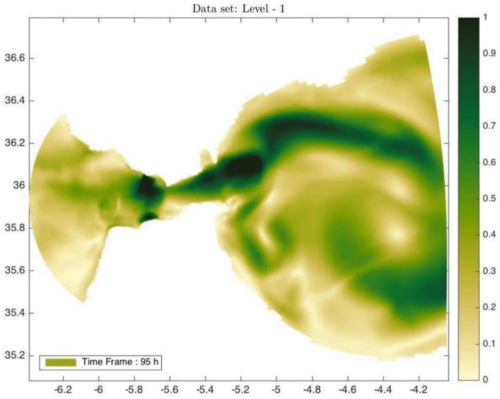
Source: https://perso.uclouvain.be/laurent.jacques/uploads/Main/buffon-present-120314-MLG.pdf

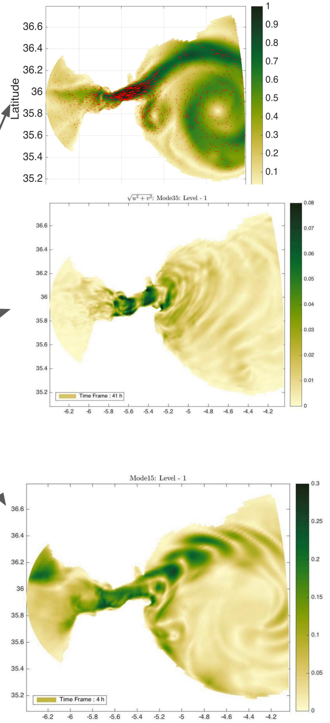
Objective & Overview

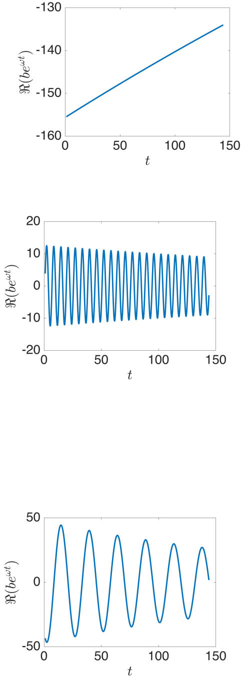
- DMD separate the variables(space & time) and isolate dynamic structures by data.
- We will use Random projection for efficient calculations.

$$X(t) \approx \sum_{i=1}^{r} b_i \Phi_i(z) e^{\omega_i t}$$

Example:







Analysis of data from Strait of Gibraltar by standard DMD.
Collaborate with: Erik Bollt, Marko Budišić, Kanaththa Priyankara,Sathsara Dias, Larry Pratt, Jose Sanchez-Garrido.

⋮ ⋮

Background

Dynamic Mode Decomposition(DMD)
[Schmid, 2008]

Numerical procedure to exact dynamical features from flow based on Krylov sequence.

DMD & Koopman operator
[Rowley & Mezić 2009]

Connect the DMD with Koopman operator

Extended DMD(EDMD)
[Williams, 2015]

Better Approximation for Koopman operator

Kernel DMD
[Williams, 2015]

Use kernel trick to reduce calculation in EDMD

Application of Koopman Operator
[Mezić, 2002]

Dynamics of Physical systems that they model based on the spectral properties of the Koopman operator.

Koopman Operator
[Koopman, 1931]

Infinite dimensional linear operator which describe the evolution of observable functions.

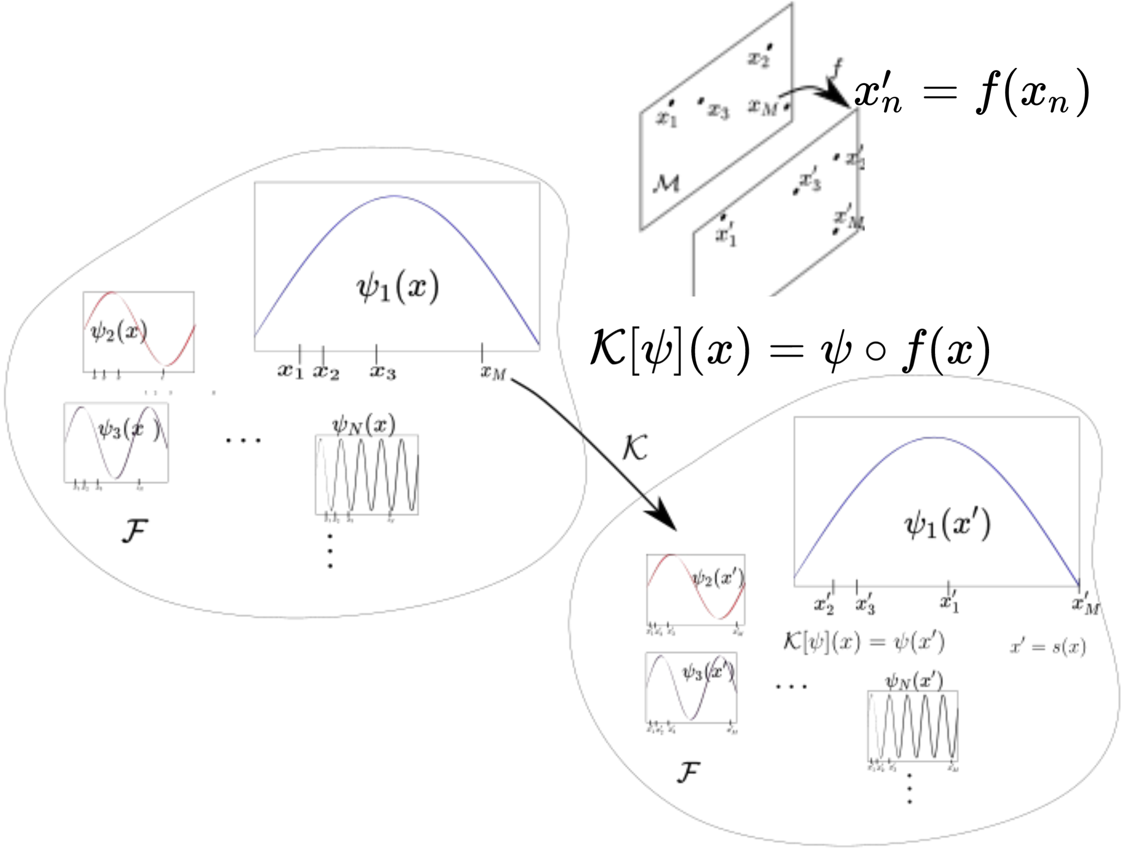
randomized Dynamic Mode Decomposition(rDMD)

- ❑ Random projection will projected data into reasonable space
- ❑ Simple and accurate calculation
- ❑ Low computational cost
- ❑ Reduce the storage cost

We are looking at DMD in lens of Johnson-Lindenstrauss theorem and random projection.

Koopman Operator

Koopman operator is a tool to analyze global dynamics of a dynamical system



Koopman operator is defined by,

$\mathcal{K} : \mathcal{F} \to \mathcal{F}$, which acts on

function space $\mathcal{F} = \{\text{set of functions } \psi : \mathcal{M} \to \mathbb{R}\}$

$$\psi'(x) = \mathcal{K}[\psi](x) = \int \delta(x' - f(x))\psi(x')dx' = \psi \circ f(x)$$

- Meaning measure f but downstream by ψ.
- Adjoint operator of Frobenius-Perron operator.

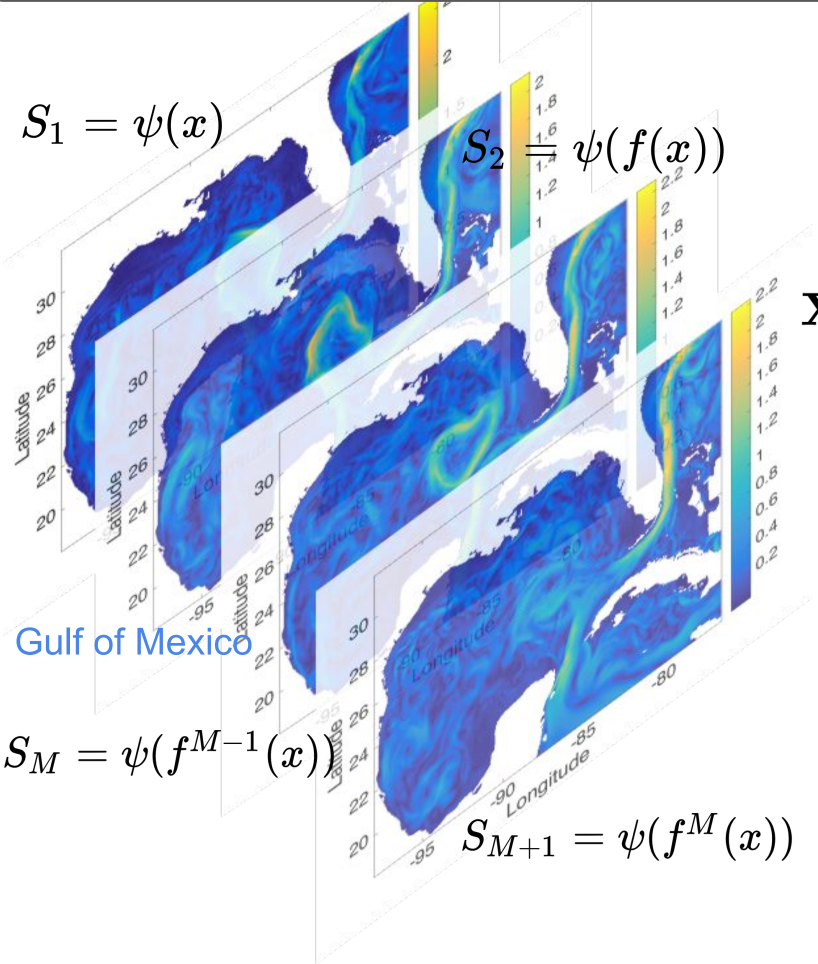
Koopman operator is a

- Linear,
- infinite dimensional

operator.

[Koopman, 1931]

Snapshot Matrix and Estimating Koopman operator



$S_1 = \psi(x)$

$S_2 = \psi(f(x))$

$S_M = \psi(f^{M-1}(x))$

$S_{M+1} = \psi(f^M(x))$

- S_i contains measures of all states at step i.

$$\mathbf{X} = \begin{bmatrix} \vdots & & \vdots \\ S_1 & \cdots & S_M \\ \vdots & & \vdots \end{bmatrix} \qquad \mathbf{Y} = \begin{bmatrix} \vdots & & \vdots \\ S_2 & \cdots & S_{M+1} \\ \vdots & & \vdots \end{bmatrix}$$

- Koopman operator acts as a time shift on columns

$$X \xrightarrow{\mathcal{K}} Y$$

$$\mathbf{K_D} = \arg\min ||\mathbf{K}X - Y||_F$$

$$\mathbf{K_D} = YX^\dagger$$

[Rowley & Mezić 2009]

Commonly Used DMD Algorithm

$$\mathbf{X} = \begin{bmatrix} \vdots & & \vdots \\ S_1 & \cdots & S_M \\ \vdots & & \vdots \end{bmatrix}_{N\times M} \quad \mathbf{Y} = \begin{bmatrix} \vdots & & \vdots \\ S_2 & \cdots & S_{M+1} \\ \vdots & & \vdots \end{bmatrix}_{N\times M}$$

$$\mathbf{K}_{\mathbf{D}\,(N\times N)} = YX^{\dagger}$$

Computing spectrum of $\mathbf{K}_{\mathbf{D}\,(N\times N)}$ is expensive

Therefore, it estimated by

$\hat{\mathbf{K}}_{\mathbf{D}\,(L\times L)} = U^{*}Y(U^{*}X)^{\dagger}$

where $X = U_L \Sigma_L V_L^{*}$.

$L << N$

Exact DMD

Input: $\mathbf{X}, \mathbf{Y}$

1 $[U_r, \Sigma_r, V_r] = \mathrm{SVD}(\mathbf{X})$; `// Truncated SVD`

2 $\hat{\mathbf{K}}_D = U_r^{*}\mathbf{Y}V_r\Sigma_r^{-1}$; `// Compression` $\mathcal{K} \downarrow \mathbf{K}_D$

3 $[\Psi, \Lambda] = \mathrm{eig}(\hat{\mathbf{K}}_D)$;`// diag(`$\Lambda$`) equals to eigenvalues of` $\mathcal{K}$

4 $\Phi = \mathbf{Y}V_r\Sigma_r^{-1}\Psi$; `//` Φ `estimates eigenvectors of` $\mathcal{K}$

5 $\mathbf{b} = \Phi^{\dagger}\mathbf{X}[:,1]$; `// Estimates the coefficients`

Output: $\Phi, \quad \Lambda, \quad \mathbf{b}$

$$X(t) = \approx \sum_{i=1}^{r} b_i \Phi_i(z) e^{\omega_i t} \qquad \omega_i = \ln(\lambda_i)/\Delta t$$

- There are a few variations of the basic algorithm.
- DMD spectrum is a numerical estimator for Koopman spectrum.

Why DMD?

- Isolate specific dynamic structures
- Equation free modeling
- Reduce the dimension of the data
- Can identify physically meaningful decomposition

[Tu, et. al, 2014]

DMD in Projected Space

$$\mathbf{K_D} = \arg\min ||\mathbf{K}X - Y||_F \quad \Longrightarrow \quad \text{DMD: } \mathbf{K_D}_{(N\times N)} = \mathbf{Y}\mathbf{X}^{\dagger}$$

DMD in projected space with rank L projector $P : \mathbb{R}^N \to \mathbb{R}^L$

$$\hat{\mathbf{K}}_{\mathbf{D}} = \arg\min ||\mathbf{K}PX - PY||_F$$

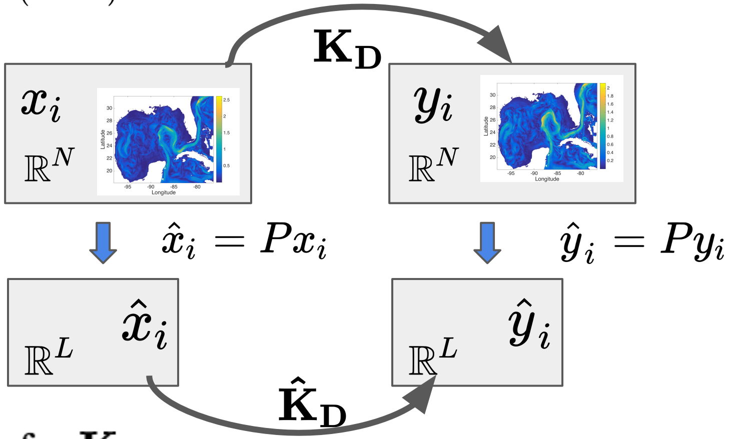
$$\hat{\mathbf{K}}_{\mathbf{D}\,(L\times L)} = P_{L\times N}\mathbf{Y}\left(P_{L\times N}\mathbf{X}\right)^{\dagger}$$

$$\hat{\mathbf{K}}_D = P\mathbf{K_D}P^{\dagger}$$

Therefore $\hat{\mathbf{K}}_{\mathbf{D}}$ and $\mathbf{K_D}$ Have common eigenvalues

If $\hat{\phi}_L$ is eigenvector of $\hat{\mathbf{K}}_{\mathbf{D}}$, then $\hat{\phi} = P^{\dagger}\hat{\phi}_L$ is a eigenvector for $\mathbf{K_D}$

$\hat{\phi} = Y(PX)^{\dagger}\phi_L$ is also a eigenvector of $\mathbf{K}_D$.



DMD with SVD projection

Projection matrix $P = U_L^*$ where $X = U_L\Sigma_L V_L^*$

DMD random projection (rDMD)

Projection matrix $P = R_{L\times N} = (\frac{1}{\sqrt{L}}R_{i,j})$

where elements $R_{i,j}$ distributed $R_{i,j} \overset{iid}{\sim} \mathcal{N}(0,1)$

Johnson-Lindenstrauss Lemma and Random Projection

- The random projection method is based on the Johnson-Lindenstrauss lemma.

Theorem[Johnson-Lindenstrauss Lemma]:

For any $0 < \epsilon < 1$ and any integer $M > 1$ let L be a positive integer such that $L \geq L_0$ with $L_0 = \frac{C \ln M}{\epsilon^2}$,

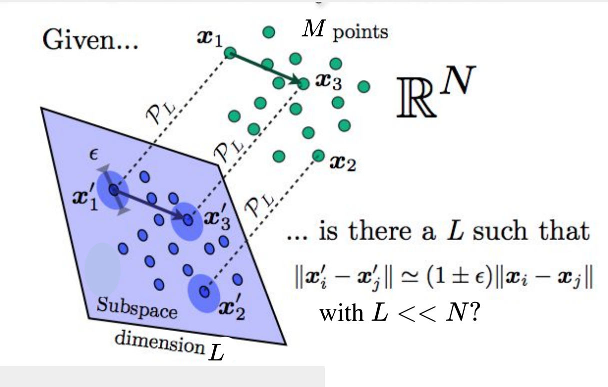
where C is a suitable constant ($C \approx 8$ in practice, $C = 2$ is good enough). Then for any set X of M data points in $\mathbb{R}^N$,

there exists a map $f : \mathbb{R}^N \to \mathbb{R}^L$ such that

for all $x_1,\ x_2 \in X$,

$(1-\epsilon)||x_1 - x_2||^2 \leq ||f(x_1) - f(x_2)||^2 \leq (1+\epsilon)||x_1 - x_2||^2$.

[Johnson et. al, 1984]



Theorem[Random Projection]

For any $0 < \epsilon, \delta < \frac{1}{2}$ and positive integer N,

there exits a random matrix of B of size $L \times N$ such that for $L \geq L_0$ with $L_0 = \frac{C \ln(1/\delta)}{\epsilon^2}$

and for any unit-length vector $x \in R^N$, $Pr\{|||Bx||^2 - 1| > \epsilon\} \leq \delta$

or $Pr\{|||Bx||^2 - 1| > \epsilon\} \leq e^{-CL\epsilon^2}$

[Papadimitriou et. al., 1998]

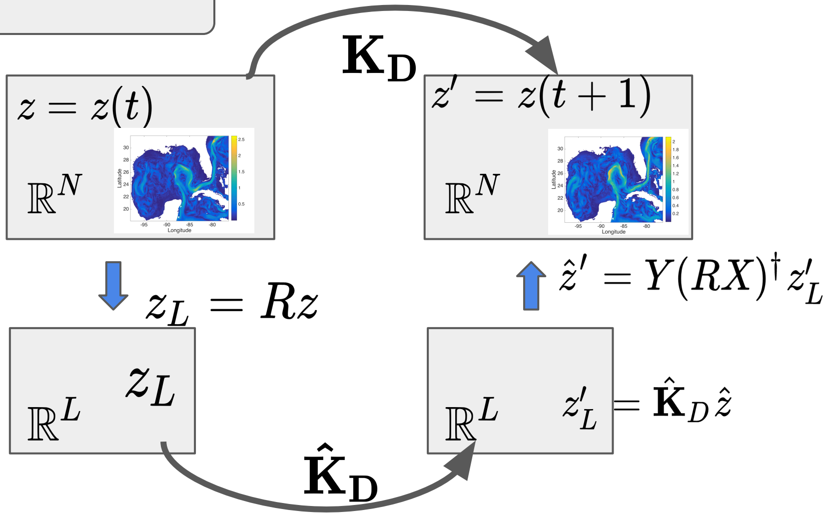
Randomized DMD

DMD in projected space with rank L random projection R

Error bound of estimating z' by using the rDMD

$$||z' - \hat{z}'|| \leq \frac{||Rz' - \hat{\mathbf{K}}_D Rz||}{1-\epsilon}$$

with at least probability of $\mathcal{O}(1/M^2)$ for any $0 < \epsilon < 1$ with $L > \frac{C\log(M)}{\epsilon^2}$.



Why rDMD?

- Simple calculation with high accuracy
- Reduce the computational cost
- Reduce the storage cost
- SVD based existing algorithms need to store high resolution data matrix and may lead to memory issues
- Our proposed algorithm can reduce the dimension of data just using matrix multiplication.

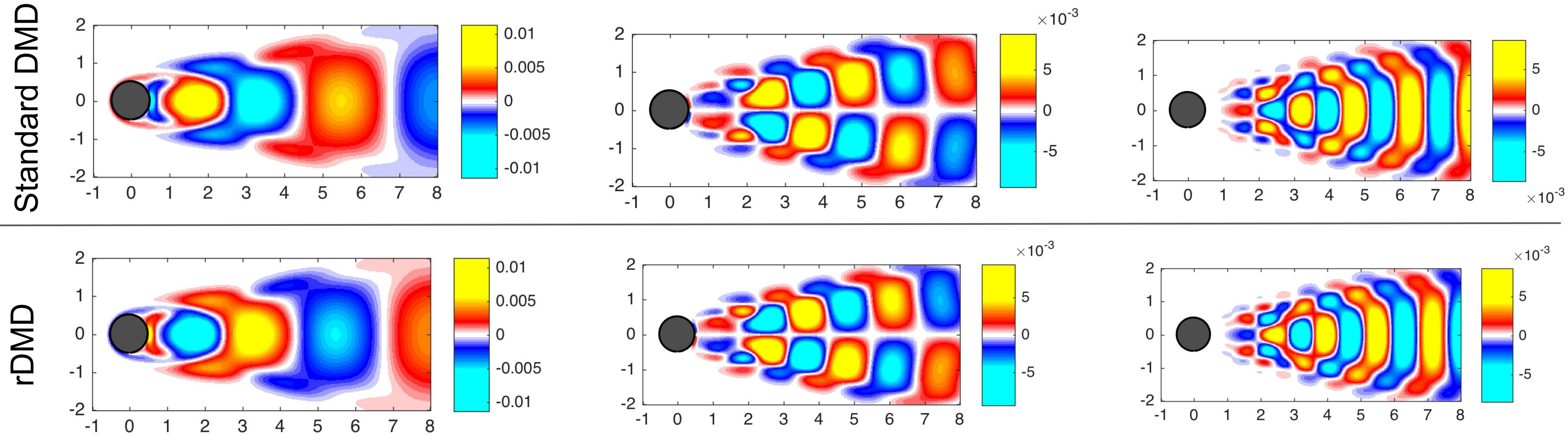
Algorithm 1: Randomized DMD(rDMD)

Data: $X,\ Y \in \mathbb{R}^{N\times M}$
Input: ϵ
$L_0 = \frac{C\ln M}{\epsilon^2}$;
Choose L such that $L \geq L_0$;
Construct a random matrix $R = \frac{1}{\sqrt{L}}(r_{ij}) \in \mathbb{R}^{L\times N}$ such that r_{ij} $N(0,1)$;
Calculate $X_L := RX,\ Y_L := RY$;
Calculate $\hat{\mathbb{K}} = Y_L X_L^\dagger$;
$[\lambda\ \ \Phi_L]$=eigs($\hat{\mathbb{K}}$);
Result: $diag(\Lambda), YX_L^\dagger\Phi_L$

Example: Fluid Flow

Example: FLUID FLOW PAST A CYLINDER AT RE=100

Data source: http://dmdbook.com/

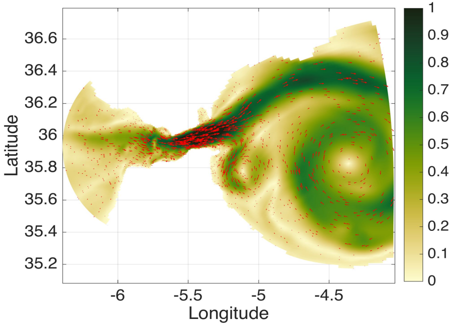


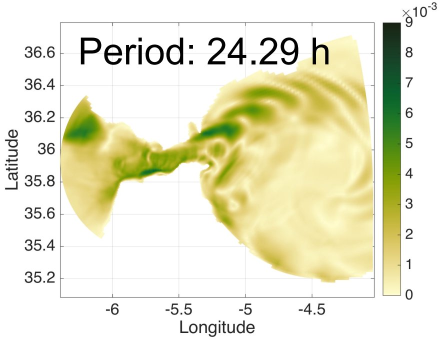
N=89351, M=150, L=25

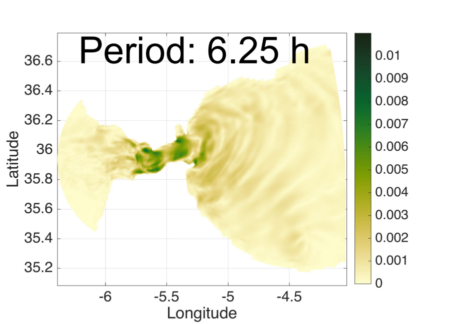
Our proposed algorithm can reduce the dimension of data just using matrix multiplication.

Application: Oceanographic Data(Strait Of Gibraltar)

Standard DMD

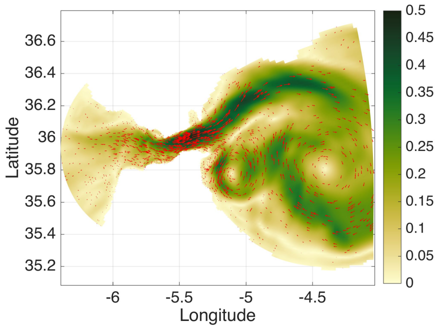


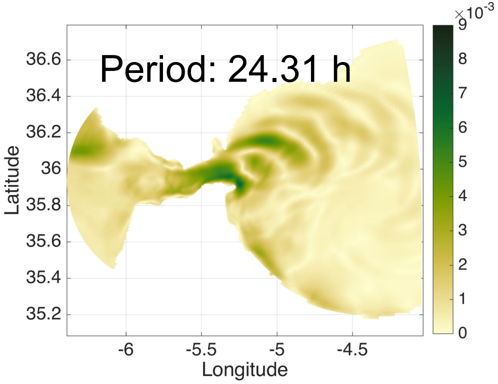


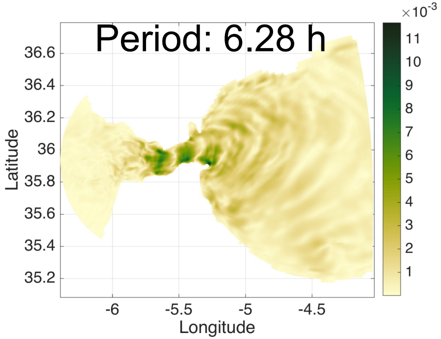


N=1,693,440, M=143, L=119

rDMD







Background Mode

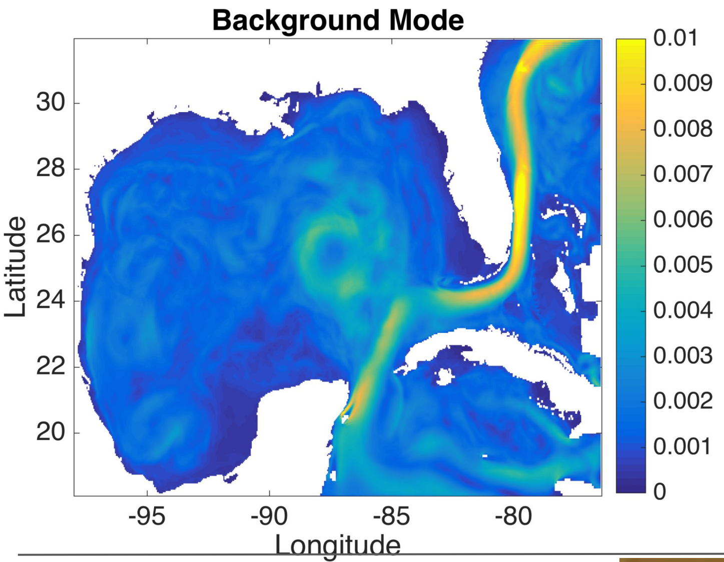
Meandering Mode

Tidal Mode

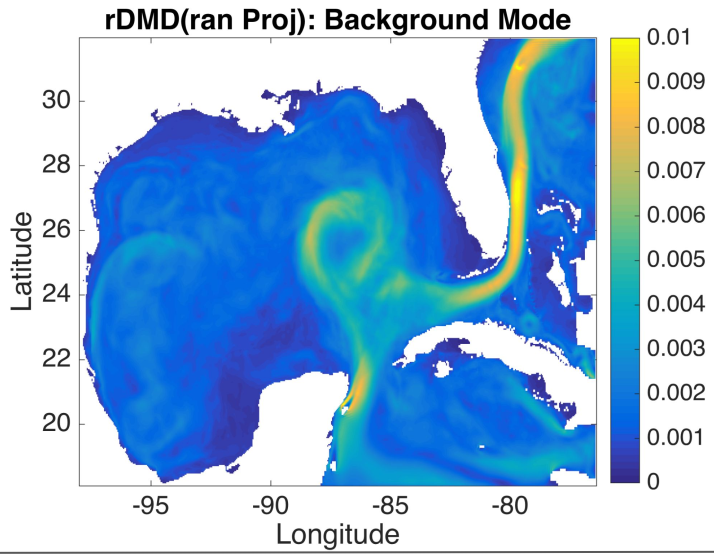
Our proposed algorithm can reduce the dimension of data just using matrix multiplication.

Application: Oceanographic Data(Gulf of Mexico)

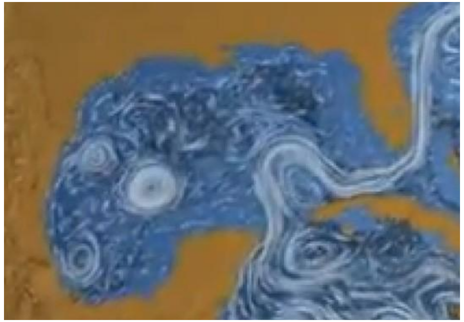
Standard DMD



rDMD



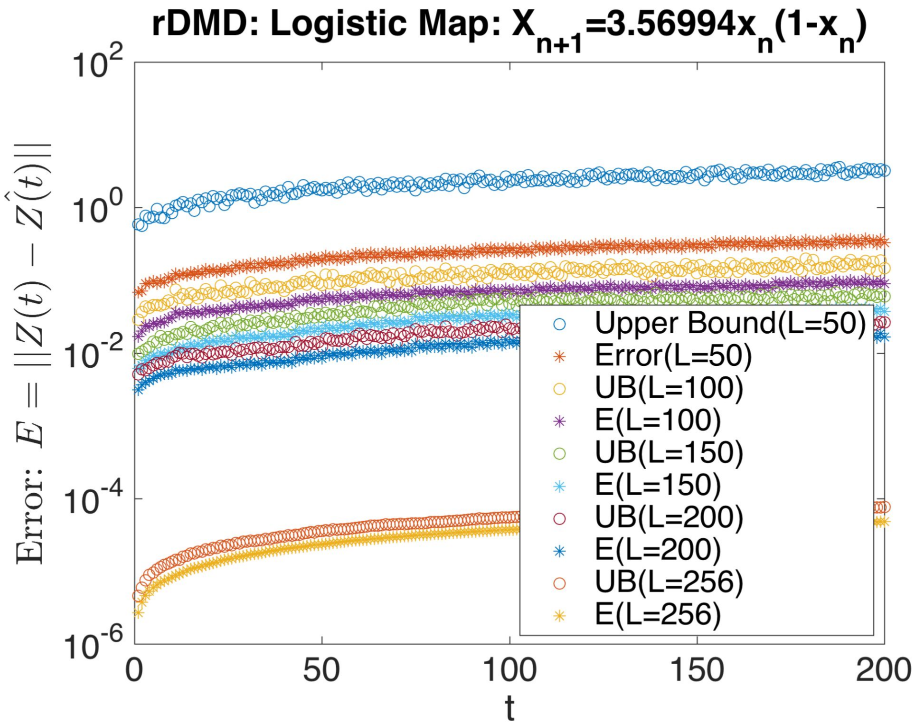
N=208285, M=104, L=90

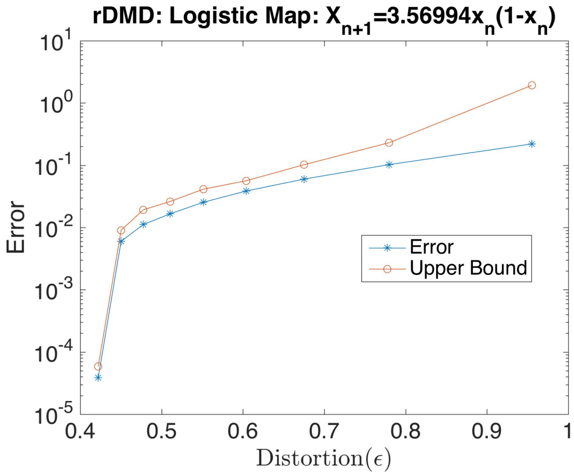


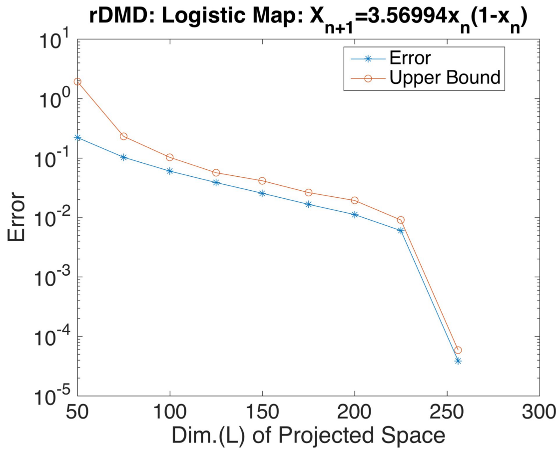
NASA's stunning Perpetual Ocean animation visualizes ocean currents (Image: NASA/Goddard Space Flight Center)

Error Analysis with Logistic map example

$$||z' - \hat{z}'|| \leq \frac{||Rz' - \hat{\mathbf{K}}_D Rz||}{1-\epsilon}$$

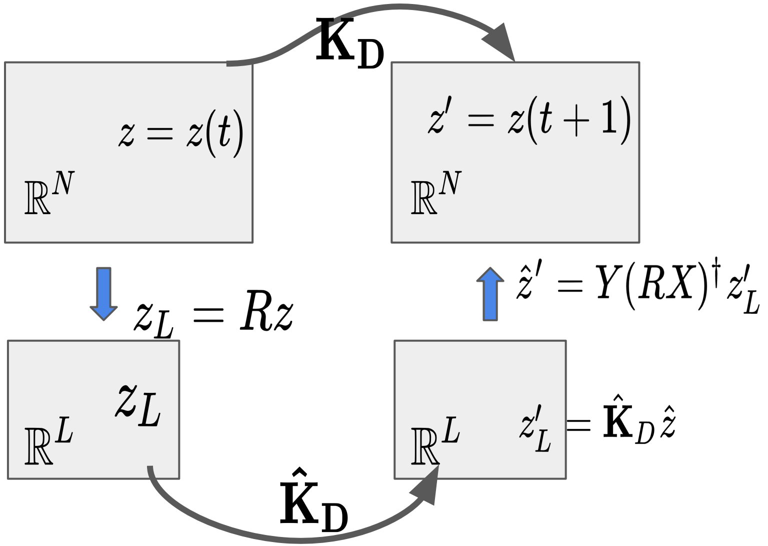






Summary

- DMD on projected space approximate Koopman Operator
- JL-theory grantee a Random projection.
  $L << N. \quad L \geq \frac{C \ln M}{\epsilon^2}$
- rDMD
  - Computationally Efficient simple algorithm
  - Reduce the storage cost.
- We will extend our current **randomized method** to **extended DMD** and **kernel DMD**.



References

[1]P.J. Schmid. Dynamic mode decomposition of numerical and experimental data. Journal of Fluid Mechanics 656.1 (2010): 5–28.
[2] C.W. Rowley, I Mezic, S. Bagheri, P. Schlatter, and D.S. Henningson. Spectral analysis of nonlinear flows. Journal of Fluid Mechanics 641 (2009): 85-113
[3] Erik M. Bollt, “Geometric Considerations of a Good Dictionary for Koopman Analysis of Dynamical Systems”, preprint arXiv:1912.09570 (2020)
[4] Mezić, Igor and Banaszuk, Andrzej, "Comparison of systems with complex behavior", Physica D: Nonlinear Phenomena, 197, 1-2: 101--133
[5] M. Budišić, R. Mohr, and I. Mezić. Applied Koopmanism. Chaos. An Interdisciplinary Journal of Nonlinear Science 22, 047510, 33 (2012).
[6] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and, J. N. Kutz.On dynamic mode decomposition: theory and applications., Journal of Computational Dynamics 1, 391–421 (2014).
[7]M.O. Williams , I.G. Kevrekidis, C.W. Rowley. A Data–Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition. Journal of Nonlinear Science 25 (2015): 1307-1346.
[8] Bernard O Koopman. Hamiltonian systems and transformation in Hilbert space. Proceedings of the National Academy of Sciences, 17(5):315–318, 1931.
[9]Schmid, P. & Sesterhenn, J. Dynamic mode decomposition of numerical and experimental data. In Sixty-First Annual Meeting of the APS Division of Fluid Dynamics. San Antonio, Texas, USA. 2008
[10] Matthew O. Williams, Clarence W. Rowley, Ioannis G. Kevrekidis. A kernel-based method for data-driven koopman spectral analysis. Journal of Computational Dynamics, 2015, 2(2): 247-265. doi: 10.3934/jcd.2015005
[11]Johnson, William B.; Lindenstrauss, Joram (1984). "Extensions of Lipschitz mappings into a Hilbert space". In Beals, Richard; Beck, Anatole; Bellow, Alexandra; et al. (eds.). Conference in modern analysis and probability (New Haven, Conn., 1982). Contemporary Mathematics. 26. Providence, RI: American Mathematical Society. pp. 189–206. doi:10.1090/conm/026/737400. ISBN 0-8218-5030-X. MR 0737400.
[12]Papadimitriou, C. H., Raghavan, P., Tamaki, H.,& Vempala, S. (1998). Latent semantic indexing: A probabilistic analysis. Proceedings of the Seventeenth ACM Symposium on the Principles of Database Systems (pp. 159–168). ACM press.

The End

Random Projection

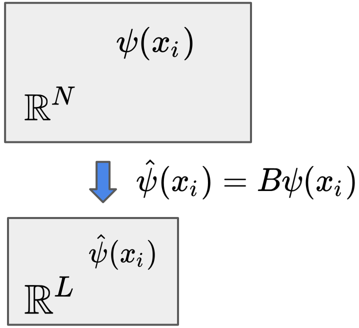
Random projection to EDMD

Extended DMD: $\mathbf{K}_{\mathbf{E}\,(N\times N)} = \mathbf{\Psi}(\mathbf{X}')\mathbf{\Psi}(\mathbf{X})^\dagger$

Random projection matrix B can be use to reduced the dimensionality of input/output space and new operator is given by :

$$\hat{\mathbf{K}}_{\mathbf{E}\,(L\times L)} = B_{L\times N}\mathbf{\Psi}(\mathbf{X}')\big(B_{L\times N}\mathbf{\Psi}(\mathbf{X})\big)^\dagger$$

$$B^\dagger\hat{\mathbf{K}}_{\mathbf{E}} = \mathbf{K}_{\mathbf{E}}B^\dagger$$



Therefore $\hat{\mathbf{K}}_{\mathbf{E}}$ and $\mathbf{K}_{\mathbf{E}}$ Have common eigenvalues

If $\hat{v}$ is eigenvector of $\hat{\mathbf{K}}_{\mathbf{E}}$, then $v = B^\dagger\hat{v}$ is a eigenvector for $\mathbf{K}_{\mathbf{E}}$

Random Projection with kernelized DMD

We can construct $B = \Psi(b)^T = [\psi(b_1)\;\psi(b_2)\ldots\psi(b_L)]^T$ by randomly sampled $b_i \in \mathcal{M}$.

Then entries $\hat{\mathbf{\Psi}}(\mathbf{X})_{ij} = k(b_i, x_j)$ and $\hat{\mathbf{\Psi}}(\mathbf{X})_{ij} = k(b_i, x'_j)$ can be calculated implicitly.

Proposed randomized kernelized DMD

Koopman Operator
on projected space $\mathcal{F}_N$

$\mathcal{A}$

$\mathcal{F}_N = span\{\psi_i : i = 1, \dots N\}$
$\psi_i : \mathcal{M} \to \mathbb{R}$

$\mathcal{F}_N = span\{\psi_i : i = 1, \dots N\}$
$\psi_i : \mathcal{M} \to \mathbb{R}$

$x_1, x_2, \dots, x_M \in \mathcal{M}$

$y_1, y_2, \dots, y_M \in \mathcal{M}$
$y_i = S_t(x_i)$

$$\Psi(X) = \begin{pmatrix} \psi_1(x_1) & \cdots & \psi_1(x_M) \\ \vdots & & \vdots \\ \psi_N(x_1) & \cdots & \psi_N(x_M) \end{pmatrix}_{N\times M} \qquad \Psi(Y) = \begin{pmatrix} \psi_1(y_1) & \cdots & \psi_1(y_M) \\ \vdots & & \vdots \\ \psi_N(y_1) & \cdots & \psi_N(y_M) \end{pmatrix}_{N\times M} \in \mathbb{R}^{N\times M}$$

There are M data points of $\psi = [\psi_1 \dots \psi_N]^T \in \mathbb{R}^N$
Goal is to reduce this dimension to $L \ll N$.
Use random projection matrix $B \in \mathbb{R}^{L\times N}$
$L \geq \frac{C \ln M}{\epsilon^2}$ with $\epsilon -$ isometry.

$\hat{\Psi}(X) = B\Psi(X)$

$\hat{\Psi}(Y) = B\Psi(Y) \in \mathbb{R}^{L\times M}$

We choose $B = \Psi(b)^T = [\psi(b_1), \dots, \psi(b_L)]^T$, where $b_i \in \mathcal{M}$ for $i = 1, \dots, L$ and selected randomly.

$\hat{\Psi}(X) = \Psi(b)^T \Psi(X) = G$

$\hat{\Psi}(Y) = \Psi(b)^T \Psi(Y) = Q$

Now instead of calculate G, Q by inner product of the feature space, we can reduce the calculation by using a kernel k. This kernel should produce L dimensional feature space.

$G_{ij} = k(b_i, x_j)$

$Q_{ij} = k(b_i, y_j)$

$\hat{A}$

$$\mathbf{K_{rk}} = \arg\min ||\mathbf{K}G - Q||_F$$