Principle Component Analysis and Model Reduction for Dynamical Systems

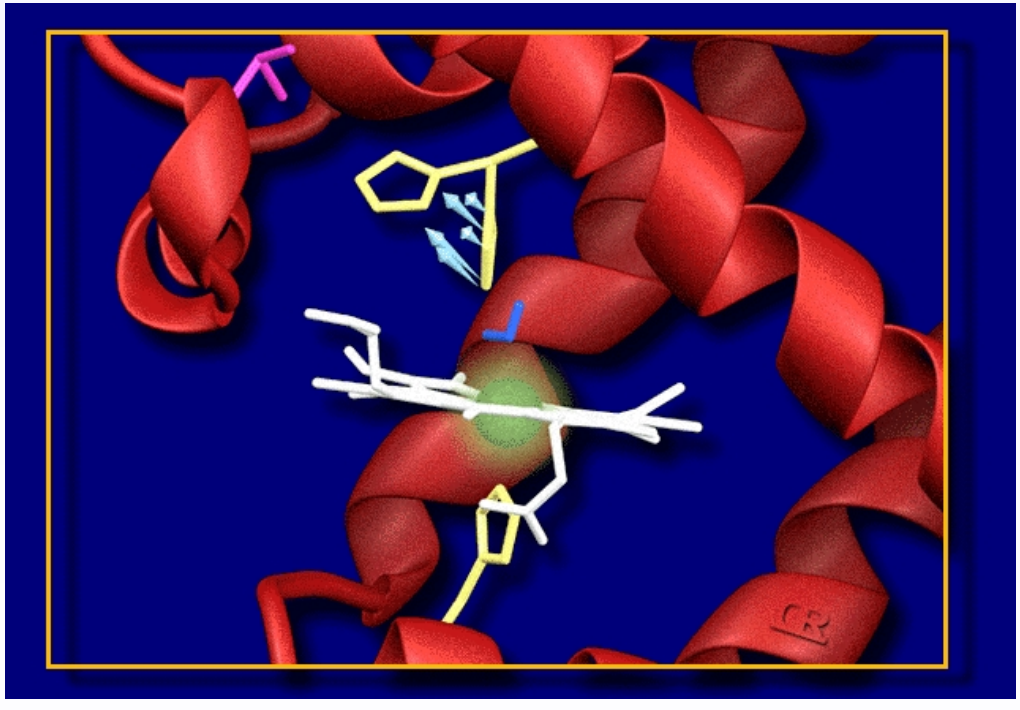
D.C. Sorensen

Virginia Tech 12 Nov 2004

Protein Substate Modeling and Identification

PCA Dimension Reduction Using the SVD

- Tod Romo
- George Phillips



T.D. Romo, J.B. Clarage, D.C. Sorensen and G.N. Phillips, Jr., Automatic Identification of Discrete Substates in Proteins: Singular Value Decomposition Analysis of Time Averaged Crystallographic Refinements, *Proteins: Structure, Function, and Genetics* **22**,311-321,(1995).

Outline

- Brief Intro to Model Reduction for Dynamical Systems
- Reduced Basis Trajectory Time Integration for MD
- The Symmetric SVD: Reduced Dimension MD Simulation

LTI Systems and Model Reduction

Time Domain

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} \end{aligned}$$

$$\mathbf{A} \in \mathbb{R}^{n\times n}, \quad \mathbf{B} \in \mathbb{R}^{n\times m}, \quad \mathbf{C} \in \mathbb{R}^{p\times n}, \quad n >> m, p$$

Frequency Domain

$$\begin{aligned} s\mathbf{x} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} \end{aligned}$$

Transfer Function

$$\mathbf{H}(s) \equiv \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}, \qquad \mathbf{y}(s) = \mathbf{H}(s)\mathbf{u}(s)$$

Model Reduction

Construct a new system $\{\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}\}$ with LOW dimension $k << n$

$$\begin{aligned} \dot{\hat{\mathbf{x}}} &= \hat{\mathbf{A}}\hat{\mathbf{x}} + \hat{\mathbf{B}}\mathbf{u} \\ \hat{\mathbf{y}} &= \hat{\mathbf{C}}\hat{\mathbf{x}} \end{aligned}$$

Goal: Preserve system response

$$\underline{\hat{\mathbf{y}} \text{ should approximate } \mathbf{y}}$$

Projection: $\mathbf{x}(t) = \mathbf{V}\mathbf{x}\big(\hat{t}\big)$ and $\mathbf{V}\dot{\hat{\mathbf{x}}} = \mathbf{A}\mathbf{V}\hat{x} + \mathbf{B}\mathbf{u}$

Model Reduction by (Krylov) Projection

Approximate $\mathbf{x} \in \mathcal{S}_V = Range(\mathbf{V})$ k-diml. subspace
i.e. Put $\mathbf{x} = \mathbf{V}\hat{\mathbf{x}}$, and then force

$$\begin{aligned} \mathbf{W}^T[\mathbf{V}\dot{\hat{\mathbf{x}}} &- (\mathbf{A}\mathbf{V}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u})] = 0 \\ \hat{\mathbf{y}} &= \mathbf{C}\mathbf{V}\hat{\mathbf{x}} \end{aligned}$$

If $\mathbf{W}^T\mathbf{V} = \mathbf{I}_k$, then the k dimensional reduced model is

$$\begin{aligned} \dot{\hat{\mathbf{x}}} &= \hat{\mathbf{A}}\hat{\mathbf{x}} + \hat{\mathbf{B}}\mathbf{u} \\ \hat{\mathbf{y}} &= \hat{\mathbf{C}}\hat{\mathbf{x}} \end{aligned}$$

where $\hat{\mathbf{A}} = \mathbf{W}^T\mathbf{A}\mathbf{V}$, $\hat{\mathbf{B}} = \mathbf{W}^T\mathbf{B}$, $\hat{\mathbf{C}} = \mathbf{C}\mathbf{V}$.

Moment Matching $\leftrightarrow$ Krylov Subspace Projection

Padé via Lanczos (PVL)

Freund, Feldmann

Bai

Multipoint Rational Interpolation

Grimme

Gallivan, Grimme, Van Dooren

Gugercin, Antoulas, Beattie

Gramian Based Model Reduction

Proper Orthogonal Decomposition (POD)
Principle Component Analysis (PCA)

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{y} = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t))$$

The gramian

$$\mathcal{P} = \int_o^\infty \mathbf{x}(\tau)\mathbf{x}(\tau)^T d\tau$$

Eigenvectors of $\mathcal{P}$

$$\mathcal{P} = \mathbf{V}\mathbf{S}^2\mathbf{V}^T$$

Orthogonal Basis

$$\mathbf{x}(t) = \mathbf{V}\mathbf{S}\mathbf{w}(t)$$

PCA or POD Reduced Basis

Low Rank Approximation

$$\mathbf{x} \approx \mathbf{V}_k \hat{\mathbf{x}}_k(t)$$

Galerkin condition – Global Basis

$$\dot{\hat{\mathbf{x}}}_k = \mathbf{V}_k^T \mathbf{f}(\mathbf{V}_k \hat{\mathbf{x}}_k(t), \mathbf{u}(t))$$

Global Approximation Error ($\mathcal{H}_2$ bound for LTI)

$$\|\mathbf{x} - \mathbf{V}_k \hat{\mathbf{x}}_k\|_2 \approx \sigma_{k+1}$$

Snapshot Approximation to $\mathcal{P}$

$$\mathcal{P} \approx \frac{1}{m} \sum_{j=1}^{m} \mathbf{x}(t_j)\mathbf{x}(t_j)^T$$

SVD of Snapshot Trajectory (Conformations)

$$\mathbf{X} = [\mathbf{x}(t_1), \mathbf{x}(t_2), \ldots, \mathbf{x}(t_m)]$$

SVD of $\mathbf{X}$:

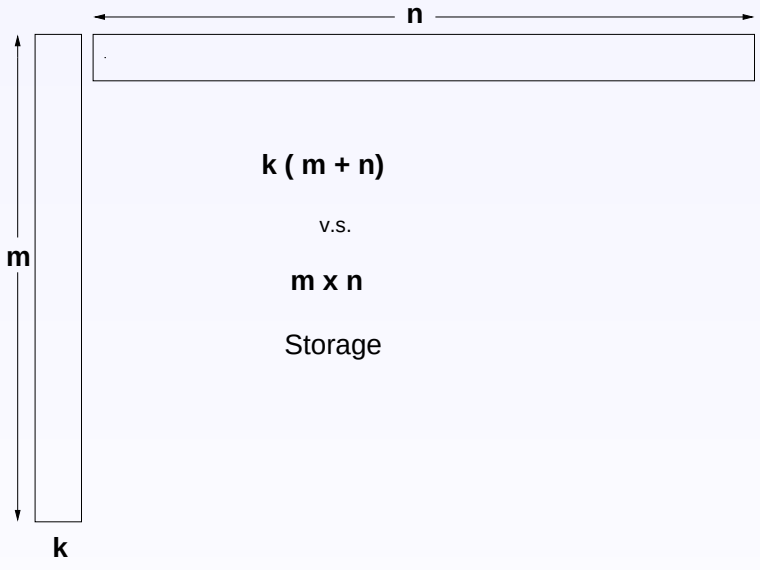
$$\boxed{\mathbf{X} = \mathbf{V}\mathbf{S}\mathbf{W}^T \approx \mathbf{V}_k\mathbf{S}_k\mathbf{W}_k^T}$$

where

$$\mathbf{V}^T\mathbf{V} = \mathbf{W}^T\mathbf{W} = \mathbf{I}_n \quad \mathbf{S} = diag(\sigma_1, \sigma_2, \cdots, \sigma_n)$$

with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

SVD Compression

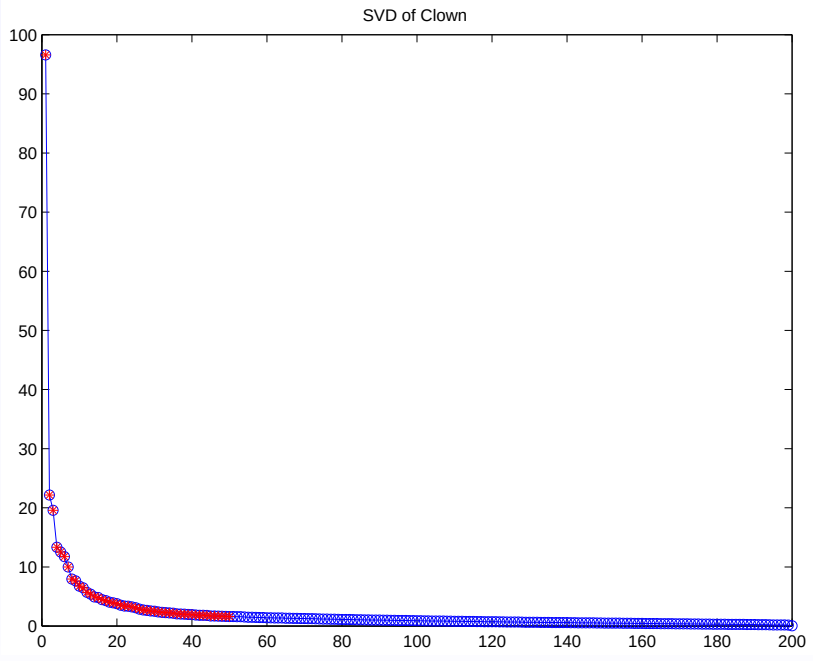


k (m + n)

v.s.

m x n

Storage

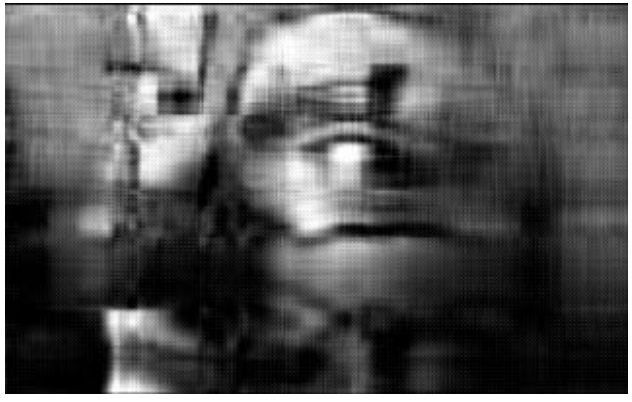


Advantage of SVD Compression

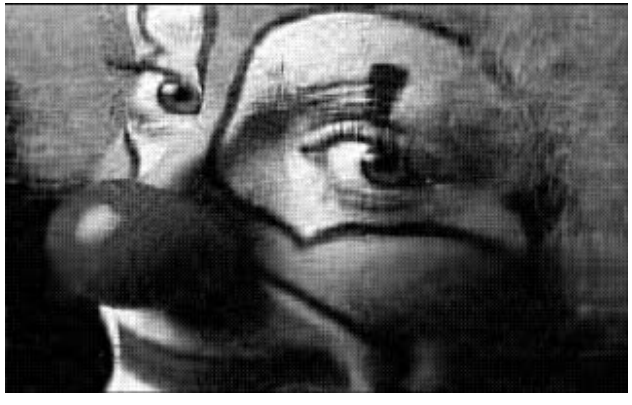
Image Compression - Feature Detection

original

rank = 10



rank = 30



rank = 50

POD in CFD

Extensive Literature

- Karhunen-Loéve, L. Sirovich
- Burns, King
- Kunisch and Volkwein
- Many, many others

Incorporating Observations – Balancing

- Lall, Marsden and Glavaski
- K. Willcox and J. Peraire

POD for LTI systems

Impulse Response: $\mathcal{H}(t) = \mathbf{C}(t\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}, \quad t \geq 0$

Input to State Map: $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{B}$

Controllability Gramian:

$$\mathcal{P} = \int_o^\infty \mathbf{x}(\tau)\mathbf{x}(\tau)^T d\tau = \int_o^\infty e^{\mathbf{A}\tau}\mathbf{B}\mathbf{B}^T e^{\mathbf{A}^T\tau} d\tau$$

State to Output Map: $\mathbf{y}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{x}(0)$

Observability Gramian:

$$\mathcal{Q} = \int_o^\infty e^{\mathbf{A}^T\tau}\mathbf{C}^T\mathbf{C}e^{\mathbf{A}\tau} d\tau$$

Balanced Reduction (Moore 81)

Lyapunov Equations for system Gramians

$$\mathbf{A}\mathcal{P} + \mathcal{P}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = 0 \quad \mathbf{A}^T\mathcal{Q} + \mathcal{Q}\mathbf{A} + \mathbf{C}^T\mathbf{C} = 0$$

With $\mathcal{P} = \mathcal{Q} = \mathbf{S}$: Want Gramians Diagonal and Equal

States Difficult to Reach are also Difficult to Observe

Reduced Model $\mathbf{A}_k = \mathbf{W}_k^T\mathbf{A}\mathbf{V}_k$, $\mathbf{B}_k = \mathbf{W}_k^T\mathbf{B}$, $\mathbf{C}_k = \mathbf{C}_k\mathbf{V}_k$

- $\mathcal{P}\mathbf{V}_k = \mathbf{W}_k\mathbf{S}_k$ $\qquad\qquad$ $\mathcal{Q}\mathbf{W}_k = \mathbf{V}_k\mathbf{S}_k$
- Reduced Model Gramians $\mathcal{P}_k = \mathbf{S}_k$ and $\mathcal{Q}_k = \mathbf{S}_k$.

Hankel Norm Error estimate (Glover 84)

Why Balanced Realization?

- Hankel singular values $= \sqrt{\lambda(\mathcal{P}\mathcal{Q})}$
- Model reduction $\mathcal{H}_\infty$ error (Glover)

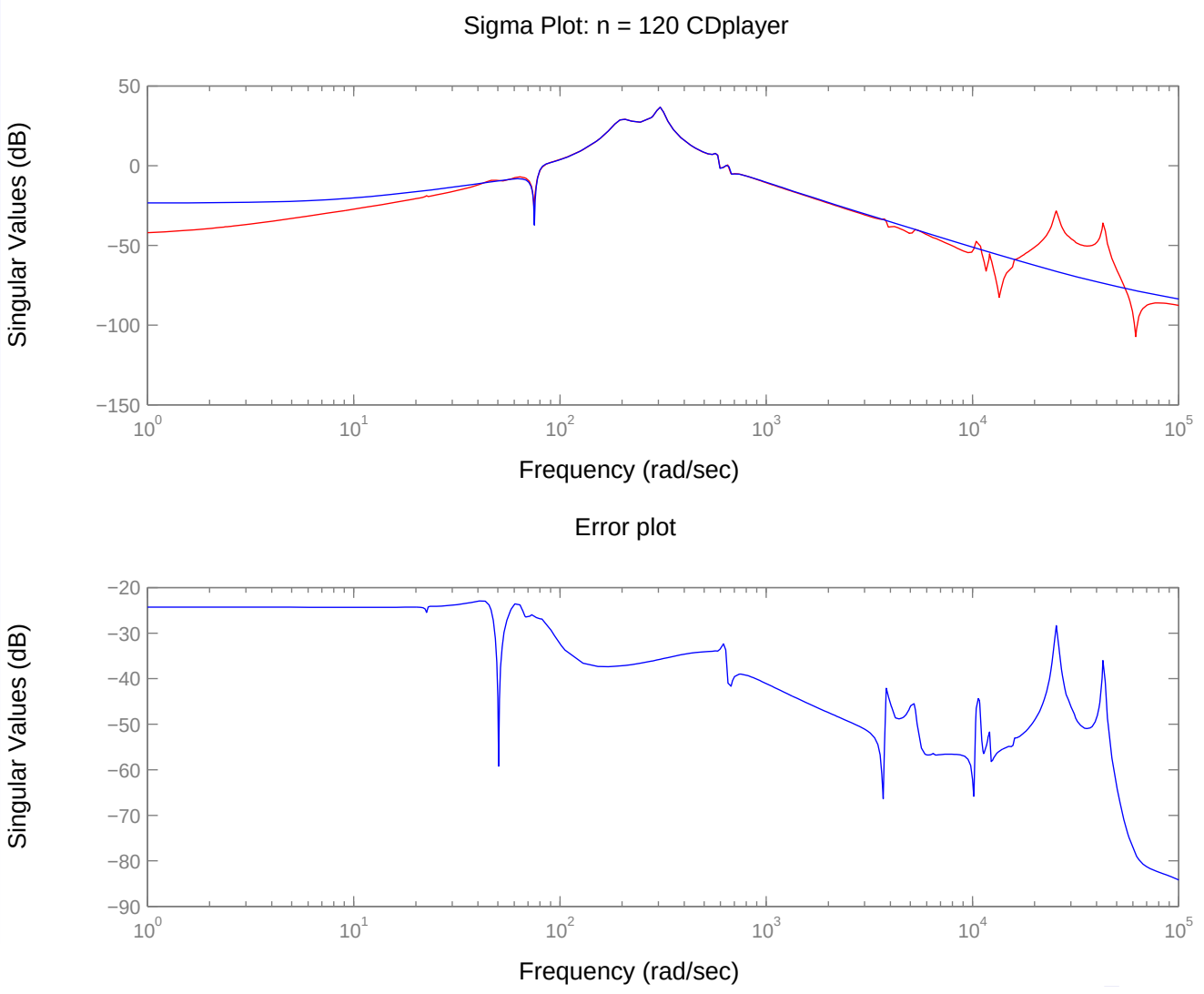
  $\|\mathbf{y} - \hat{\mathbf{y}}\|_2 \leq 2 \times (\text{sum neglected singular values})\|u\|_2$
- Extends to MIMO
- Preserves Stability

Key Challenge

- Approximately solve large scale Lyapunov Equations
  in Low Rank Factored Form

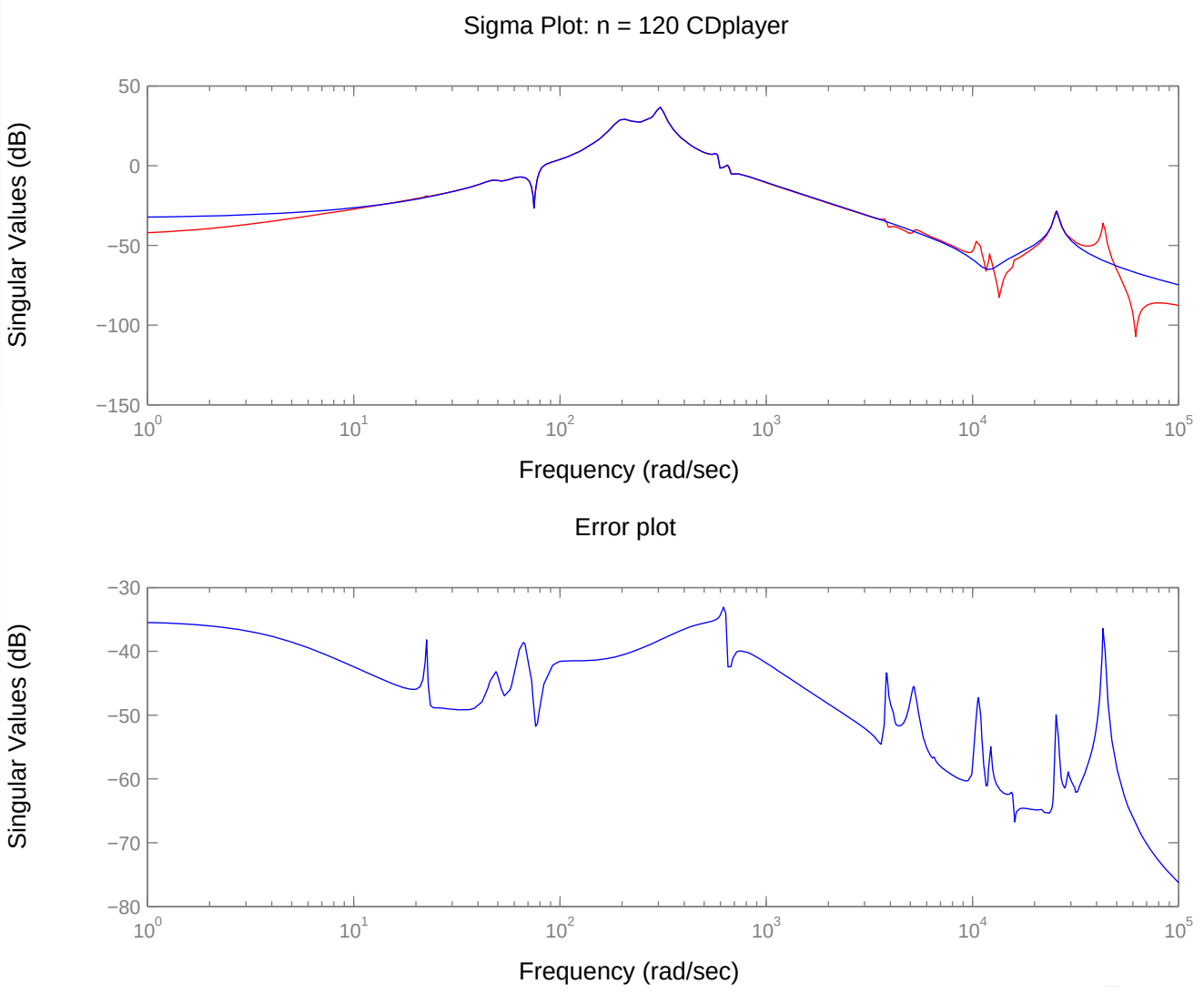
CD Player Impluse Response

$n = 120$ k $= 11$, tol $=$ 5e-3



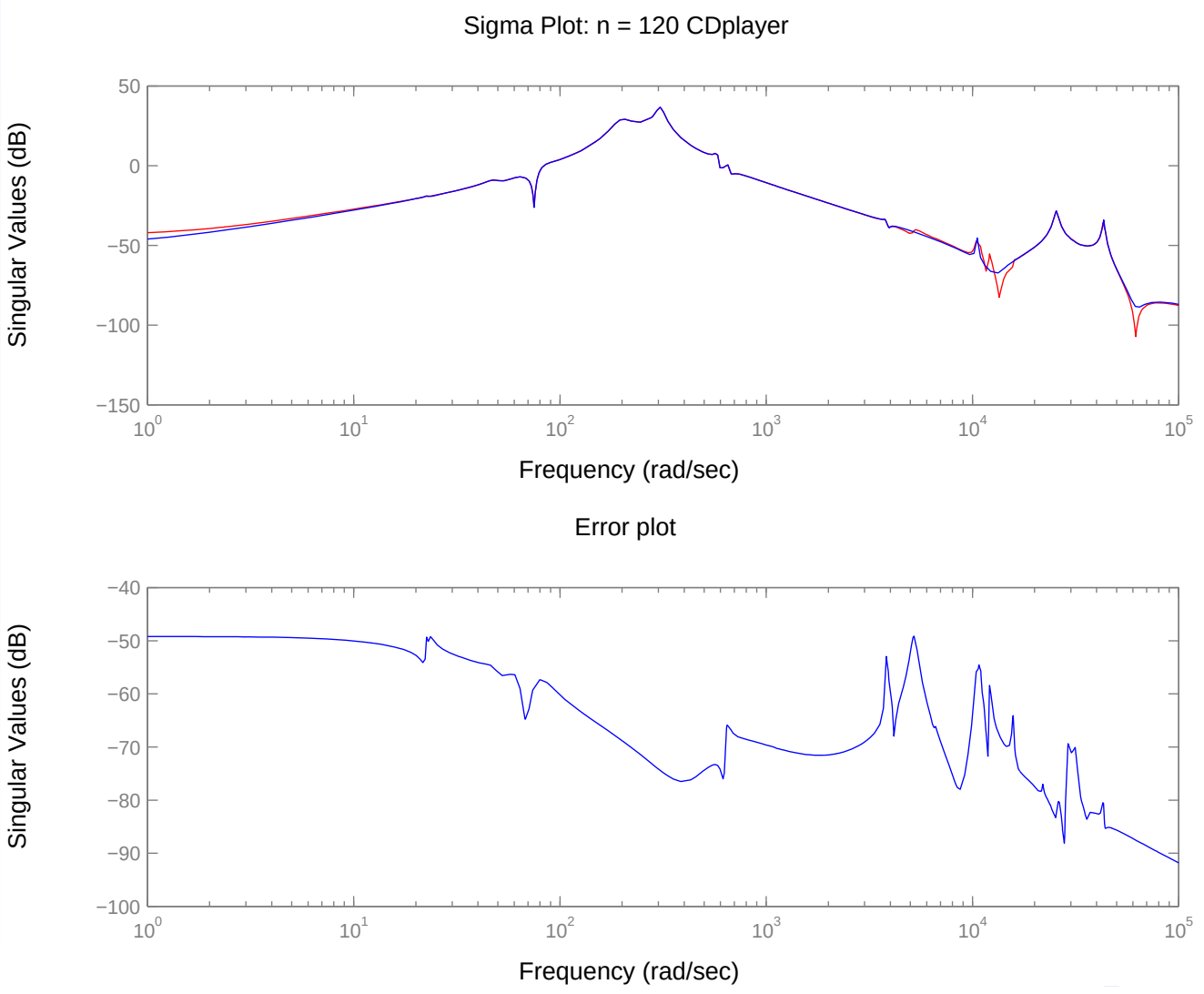
CD Player Impluse Response

$$k = 17 \text{ , tol} = 5\text{e-}4$$

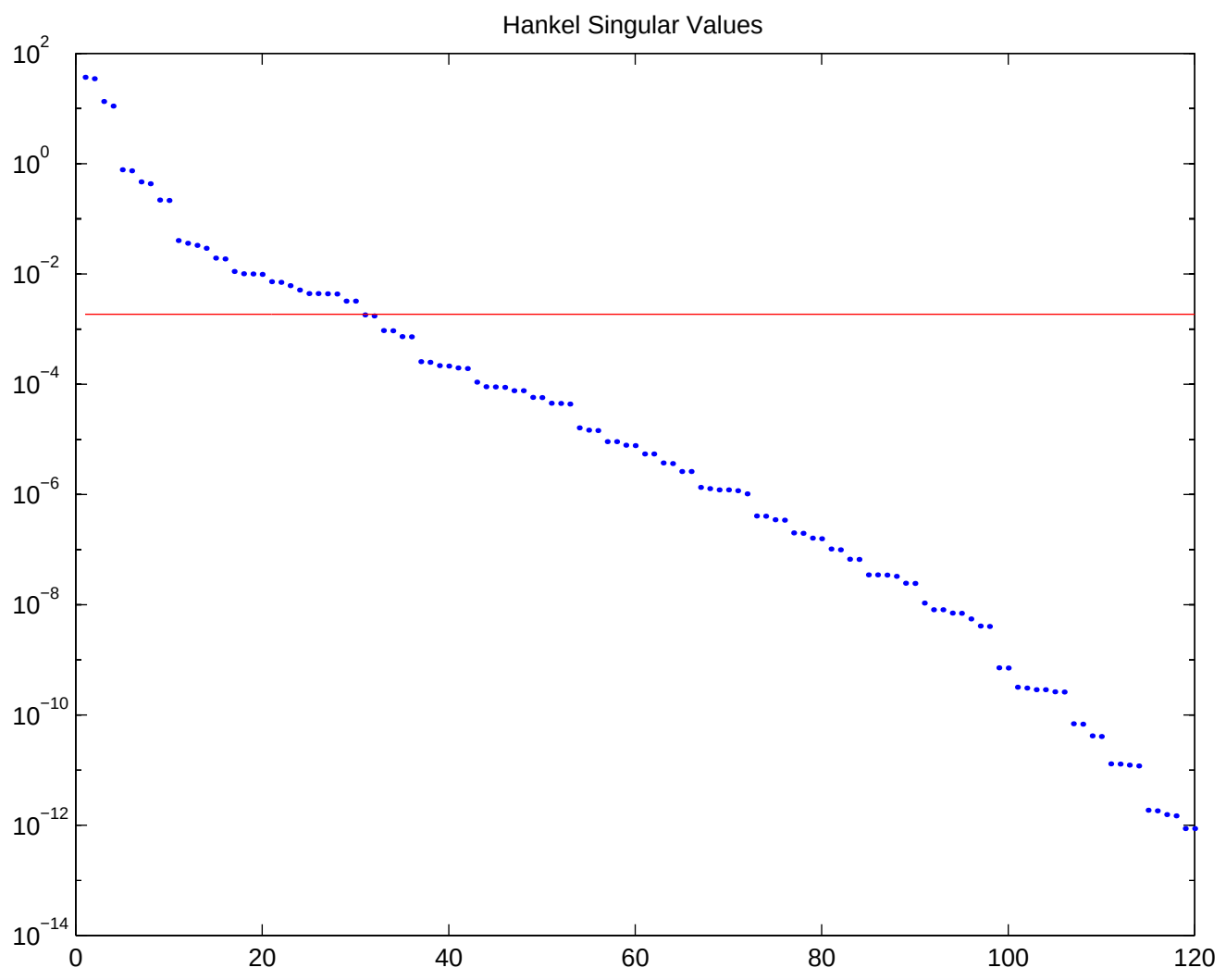


CD Player Impluse Response

$$k = 31 \text{ , tol} = 5\text{e-}5$$



CD Player - Hankel Singular Values



Reduction of Second Order Systems

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}} + \mathbf{G}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} &= \mathbf{B}\mathbf{u} \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \quad \rightarrow \quad \begin{aligned} \hat{\mathbf{M}}\ddot{\hat{\mathbf{x}}} + \hat{\mathbf{G}}\dot{\hat{\mathbf{x}}} + \hat{\mathbf{K}}\hat{\mathbf{x}} &= \hat{\mathbf{B}}\mathbf{u} \\ \hat{\mathbf{y}}(t) &= \hat{\mathbf{C}}\hat{\mathbf{x}}(t) \end{aligned}$$

where $\hat{\mathbf{M}} = \mathbf{V}^T\mathbf{M}\mathbf{V}$, etc. with $\mathbf{V}^T\mathbf{V} = \mathbf{I}$.

Key Point: *Preserve Second Order Form*
– DO NOT convert to First Order Sys.

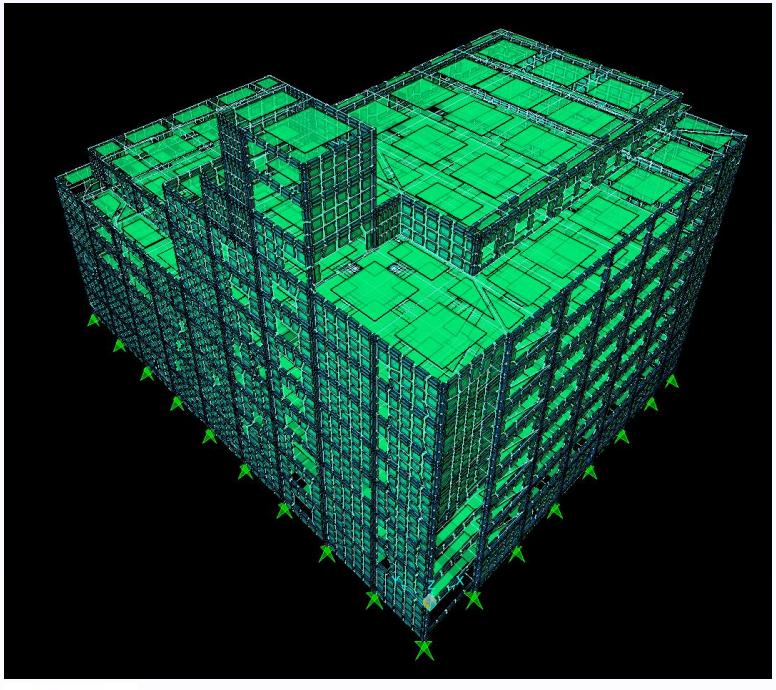
Keeps Physical Meaning - can be built

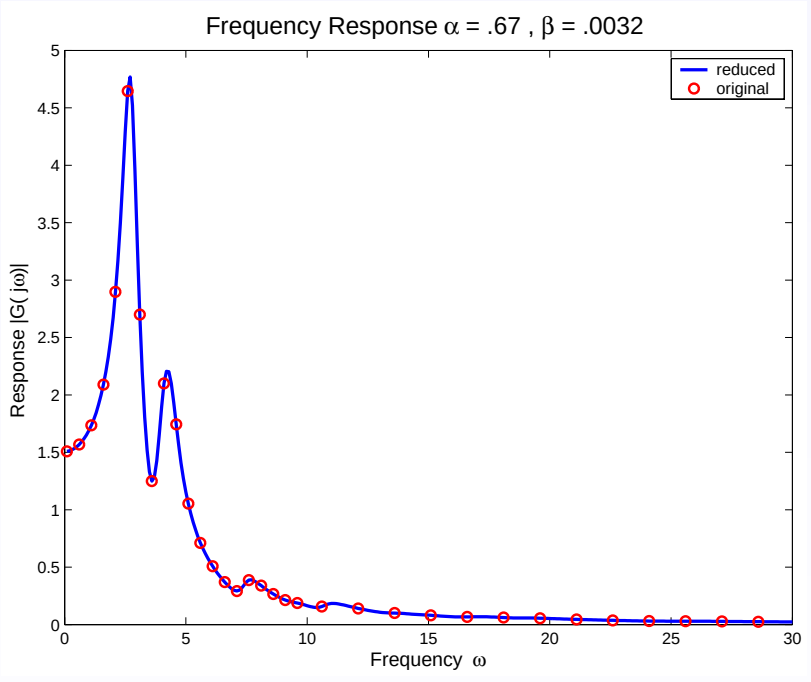
Applications

Mechanical Systems Electrical Systems MEMS devices

e.g. Building Model

N = 26394, k = 200 (ROM)





Error Bound for Second Order Systems

A. Antoulas, C. Teng

Controllability Gramian - Impulse Response

$$\mathbf{P} := \int_0^\infty \mathbf{x}(t)\mathbf{x}(t)^* \mathrm{d}t.$$

Reduce with Dominant Eigenspace $\mathbf{P}$: $\mathbf{P}\mathbf{V}_1 = \mathbf{V}_1\mathbf{S}_1$

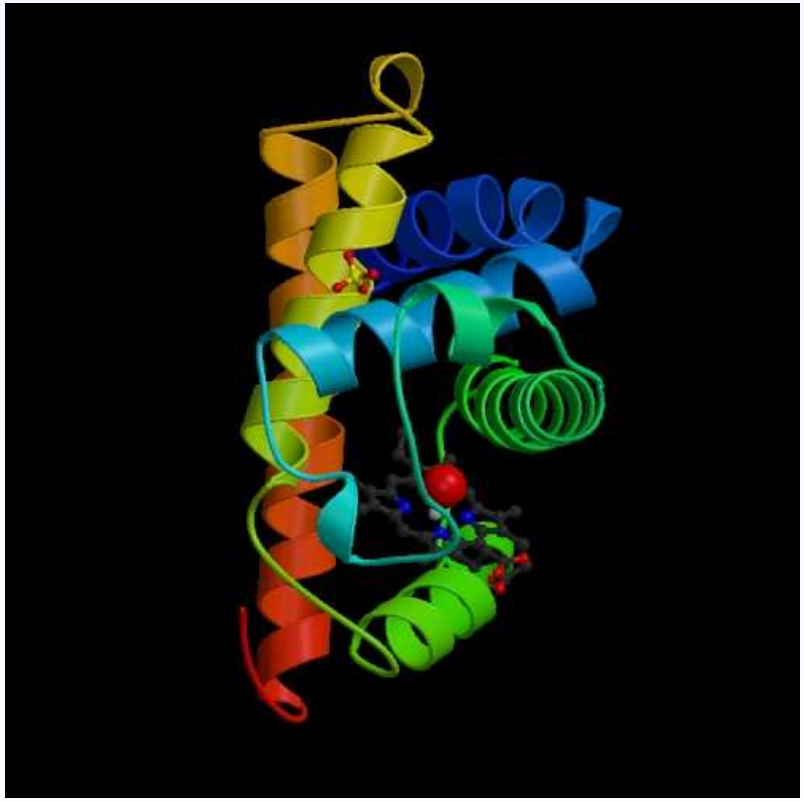
Bounded $\mathcal{H}_2$ norm of error system $\mathcal{E} = \Sigma - \hat{\Sigma}$

$$\boxed{\|\mathcal{E}\|_{\mathcal{H}_2}^2 \leq C_o tr\{\mathbf{S}_2\}}$$

Key: Expression for $\mathbf{P}$ in frequency domain.

PCA Model Reduction for Molecular Dynamics

- Rachel Vincent
- Monte Pettitt



Classical Equations of Motion

Molecular dynamics (MD) simulation is a computational tool used to study a molecular system as it evolves through time. Newton's second law of motion governs atomic motion in MD:

$$\mathrm{M}\ddot{\mathbf{r}}(t) = -\nabla \mathcal{V}(\mathbf{r}(t)).$$

- $\mathbf{r}(t)$ = vector of atomic coordinates at time t
  $= [x_{1t}\ y_{1t}\ z_{1t} \cdots x_{N_a t}\ y_{N_a t}\ z_{N_a t}]^T$
- M = diagonal matrix of atomic masses
- $\mathcal{V}(\mathbf{r}(t))$ = potential energy function

Time Step Barrier $\propto$ *fs*

Example: DHFR (dihydrofolate reductase), 23,558 atoms

To realize a microsecond simulation with a time step of 2 *fs* would require about **13 months** of simulation time when utilizing **126 processors**.

Time with respect to simulation using the NAMD program (**N**ot **A**nother **M**olecular **D**ynamics program) on an Origin 2000 R10000/250.

PCA Reduced Basis Simulation

$$\mathbf{M}\mathbf{V}\ddot{\mathbf{y}}(t) = -\nabla\mathcal{V}(\mathbf{V}\mathbf{y}(t)) \quad \rightarrow \quad \hat{\mathbf{M}}\ddot{\mathbf{y}}(t) = -\mathbf{V}^T\nabla\mathcal{V}(\mathbf{V}\mathbf{y}(t))$$

1. Initial Basis $\mathbf{V}$: truncated SVD of short traditional MD trajectory using ARPACK.
2. Approximate the reduced basis potential energy with Radial Basis Fit.
3. Update reduced basis positions $\mathbf{y}$ and velocities $\dot{\mathbf{y}}$ in k dimensions using the approximate potential.
4. Reconstruct $3ND$ trajectory $\mathbf{r} = \mathbf{V}\mathbf{y}$.
5. Update and truncate reduced basis and perform full space correction as needed.

Remarks

- Butane (n = 42):
  80% to 90% of the total motion with 10 - 15 LSVs (24% - 36% DOF)
  Reduced Simulation times order of seconds
  Traditional MD simulation took several minutes.

- BPTI(n = 2700):
  80% to 90% of the total motion with 300 - 500 LSVs (11% - 19% DOF)
  Reduced Simulation times order of minutes
  Traditional MD simulation took several hours.

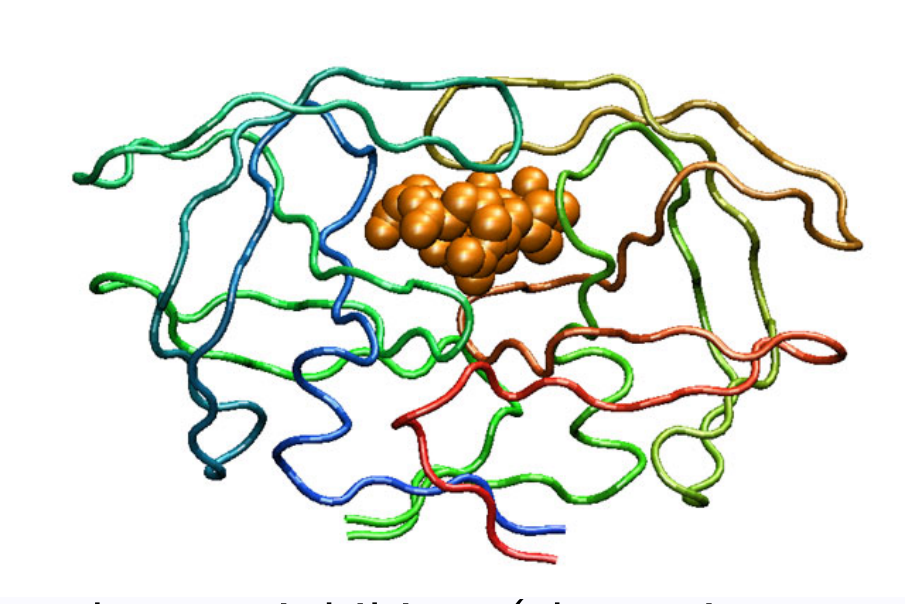
Symmetry Preserving SVD (Mili Shah)

Collaboration with the *Physical and Biological Computing Group*

- Lydia Kavraki
- Mark Moll
- David Schwarz
- Amarda Shehua
- Allison Heath

Symmetry in HIV-1 protease

Backbone representation of HIV-1 protease (from M. Moll)

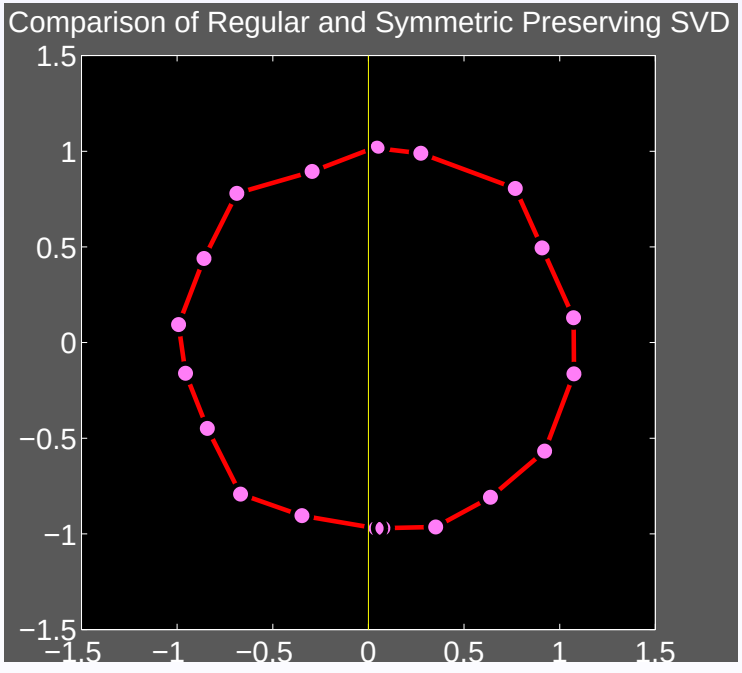


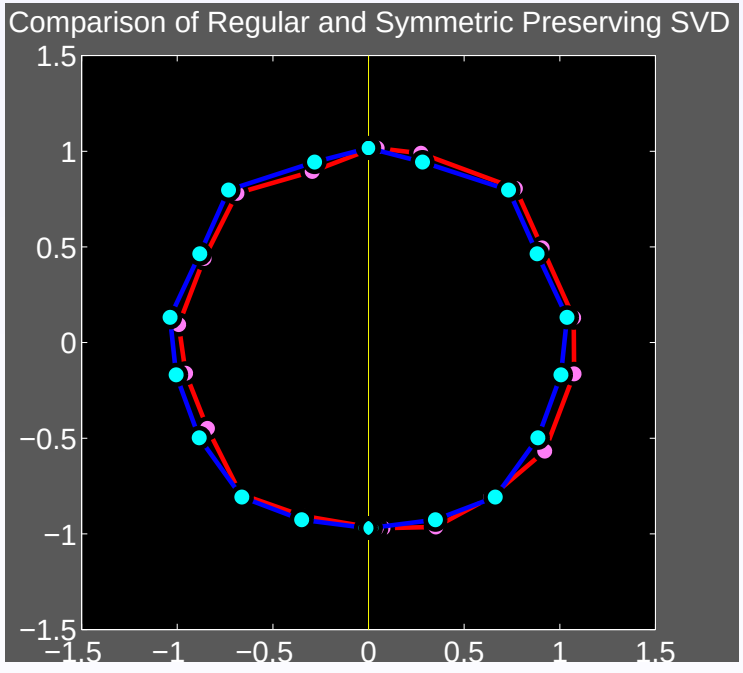
bound to an inhibitor (shown in orange)

Uses PCA dimension reduction of Molecular Dynamics Simulations

Symmetry across a plane should be present

Animation: Symmetric SVD Approximation





click below figures for movies

Finding the Plane of Symmetry

Suppose

$$\mathbf{X} = [x_1, x_2, \ldots, x_n] \quad \text{and} \quad \mathbf{Y} = [y_1, y_2, \ldots, y_n]$$

are two sets of points symmetric across a plane

Exact symmetry condition:

$$\mathbf{Y} = (\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)\mathbf{X},$$

where $\mathbf{w}$ is the normal to the (hyper-) Plane of Symmetry

$$\mathcal{H} = \{\mathbf{x} : \mathbf{w}^T\mathbf{x} = 0\}$$

Remark: In Numerical Linear Algebra $(\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)$ is a Householder Transformation or Elementary Reflector

Best Approximate Plane of Symmetry

Symmetry condition with Noise:

$$\mathbf{Y} = (\mathbf{I} - 2\mathbf{w}_o\mathbf{w}_o^T)\mathbf{X} + \mathbf{E},$$

Problem: Compute a unit vector $\mathbf{w}$ that gives the best *Approximate Plane of Symmetry*

$$\min_{\mathbf{w}} \|\mathbf{Y} - (\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)\mathbf{X}\|_F,$$

Solution:

$$(\mathbf{X}\mathbf{Y}^T + \mathbf{Y}\mathbf{X}^T)\mathbf{v} = \mathbf{v}\lambda_{min}, \quad \mathbf{w} = \mathbf{v}$$

gives the normal $\mathbf{w}$ to the best approximate plane of symmetry

Supressing Outlier Effects

Iteratively determine diagonal *weighting matrix* $\mathbf{D}_w$

The i-th diagonal of $\mathbf{D}_w$ is $1/\text{discrepancy}$,

$$\text{discrepancy} = \|\mathbf{y}_i - (\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)\mathbf{x}_i\|$$

Problem: Compute a unit vector $\mathbf{w}$ that gives the best *Weighted Approximate Plane of Symmetry*

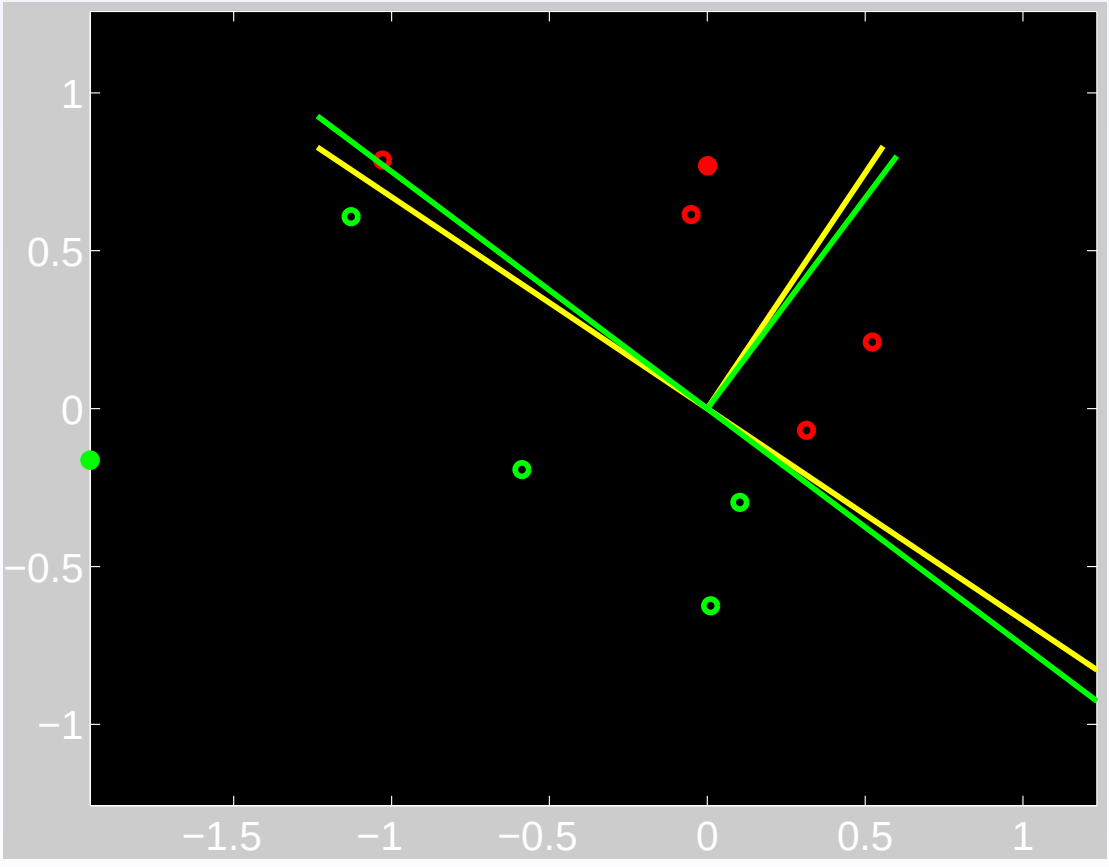
$$\min_{\mathbf{v}} \|[\mathbf{Y} - (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T)\mathbf{X}]\mathbf{D}_w\|_F,$$

Solution:

$$(\mathbf{X}\mathbf{D}_w^2\mathbf{Y}^T + \mathbf{Y}\mathbf{D}_w^2\mathbf{X}^T)\mathbf{v} = \mathbf{v}\lambda_{min}, \quad \mathbf{w} \leftarrow \mathbf{v}$$

gives the normal $\mathbf{w}$ to the best weighted approximate plane of symmetry

Finding Normal to "Best" Plane of Symmetry



click below for movie

The Symmetric SVD Approximation

If $\mathbf{W}\mathbf{X}_2 = \mathbf{X}_1 + \mathbf{E}$ where $\mathbf{W} = \mathrm{blockdiag}(\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)$

$$\min_{\mathbf{w}\hat{\mathbf{x}}_2 = \hat{\mathbf{x}}_1} \left\| \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} - \begin{pmatrix} \hat{\mathbf{X}}_1 \\ \hat{\mathbf{X}}_2 \end{pmatrix} \right\|_F^2 \quad \text{and} \quad \begin{pmatrix} \hat{\mathbf{X}}_1 \\ \hat{\mathbf{X}}_2 \end{pmatrix} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

Solved by:

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{pmatrix}, \quad \mathbf{S} = \sqrt{2}\mathbf{S}_1, \quad \mathbf{V} = \mathbf{V}_1. \quad \text{and} \quad \mathbf{U}_2 = \mathbf{W}\mathbf{U}_1,$$

with

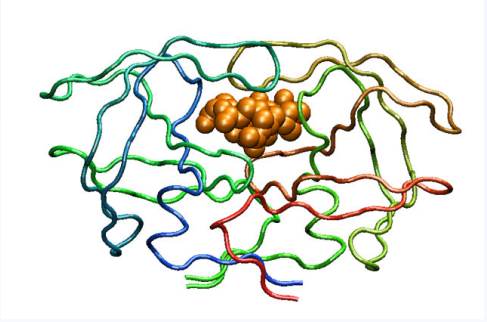
$$\boxed{\mathbf{U}_1 S_1 \mathbf{V}_1^T = \tfrac{1}{2}(\mathbf{X}_1 + \mathbf{W}\mathbf{X}_2)}$$

Symmetric Major Modes: HIV-1 protease

- Major mode regular SVD is red
- Major mode SYMMETRIC SVD is blue
- 3120 atoms (3*3120=9360 degrees of freedom)
- MD trajectory consisted of 10000 conformations (NAMD)
- SVD and SymSVD used P_ARPACK on a Linux cluster
- dual-processor nodes; 1600MHz AMD Athlon processors, 1GB RAM per node. 1GB/s Ethernet connection . 12 Processors = 6 nodes.
- First 10 standard singular vectors: 88 secs.
- First 10 symmetric singular vectors: 131 secs.

Animation: Symmetric SVD on HIV1 Protease

click for movie

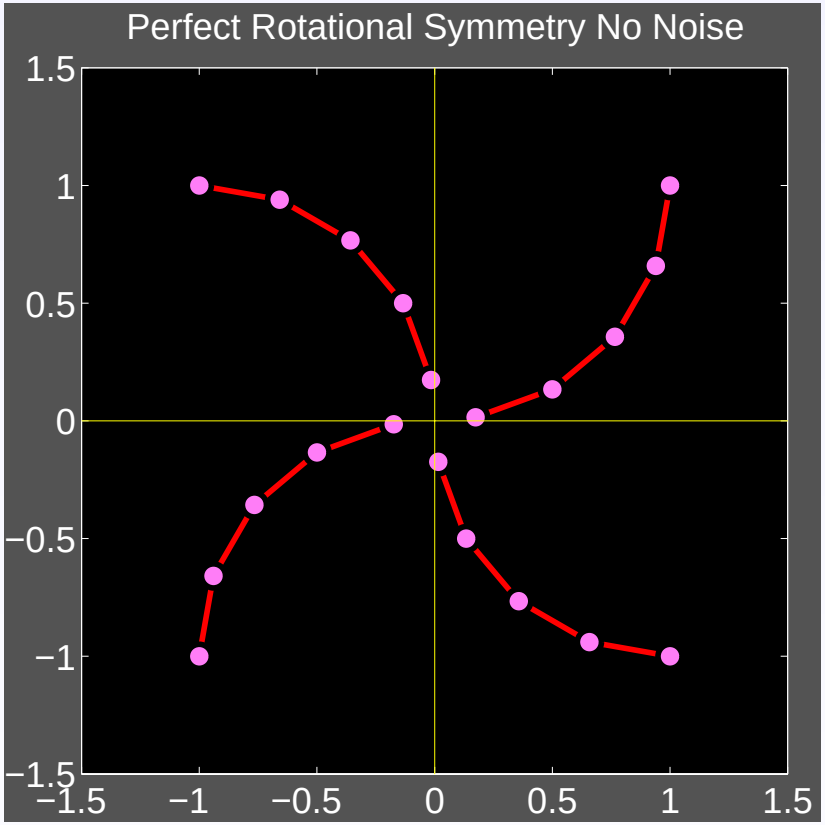


Red = Unsymmetric

Blue = Symmetric

First SVD mode – Symmetric vs. Unsymmetric

Rotational Symmetry



$$\mathbf{X}_j = \mathbf{W}\mathbf{X}_{j-1}, \quad j = 1 : k-1, \quad \text{where} \quad \mathbf{W} = \mathbf{I} - \mathbf{Q}\mathbf{G}\mathbf{Q}^T$$

$\mathbf{I}_p - \mathbf{G}$ is a rotation $\qquad\qquad$ $\mathbf{X}_k = \mathbf{X}_0$.

Finding the Axis of Rotation

$\mathbf{q}$ is an axis of rotation iff $\mathbf{Q}^T\mathbf{q} = 0$

$$\mathbf{q}^T\mathbf{W} = \mathbf{q}^T(\mathbf{I} - \mathbf{Q}\mathbf{G}\mathbf{Q}^T) = \mathbf{q}^T \quad \Rightarrow \quad \mathbf{q}^T\mathbf{X}_0 = \mathbf{q}^T\mathbf{X}_j$$

Let $\mathbf{M} = (k-1)\mathbf{X}_0 - \sum_{j=1}^{k-1}\mathbf{X}_j$

$$\min_{\|\mathbf{q}\|=1} \|\mathbf{M}^T\mathbf{q}\| \quad (= 0 \ \text{ if exact symmetry holds})$$

Good for *noisy* data

(for another condition see Minovic,Ishikawa and Kato)

Best Rotationally Symmetric Approximation

If $\mathbf{W}^{k-j}\mathbf{X}_j = \mathbf{X}_0 + \mathbf{E}_j, \quad j = 1 : k-1$

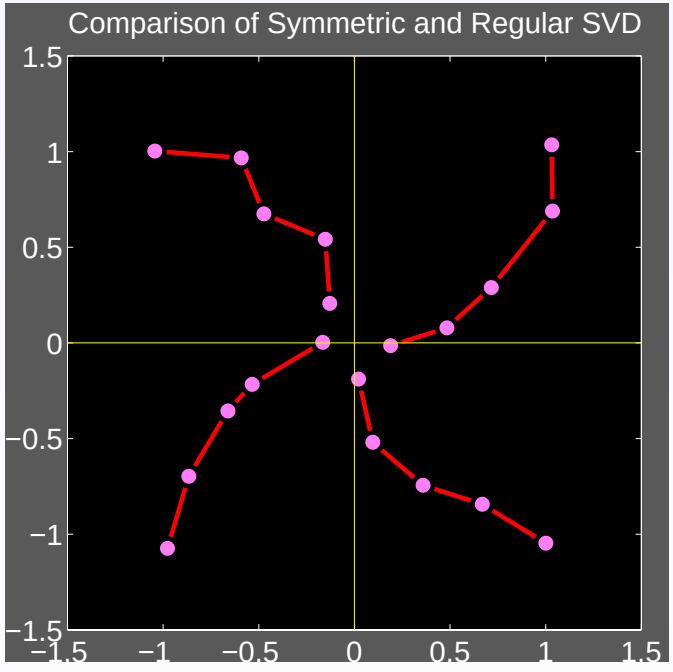
$$\min_{\widehat{\mathbf{X}}_{j+1}=\mathbf{W}\widehat{\mathbf{X}}_j} \left\| \begin{bmatrix} \mathbf{X}_0 \\ : \\ \mathbf{X}_{k-1} \end{bmatrix} - \begin{bmatrix} \widehat{\mathbf{X}}_0 \\ : \\ \widehat{\mathbf{X}}_{k-1} \end{bmatrix} \right\|_F^2 = \frac{1}{k} \sum_{j=0}^{k-1} \|\mathbf{E}_j\|_F^2,$$

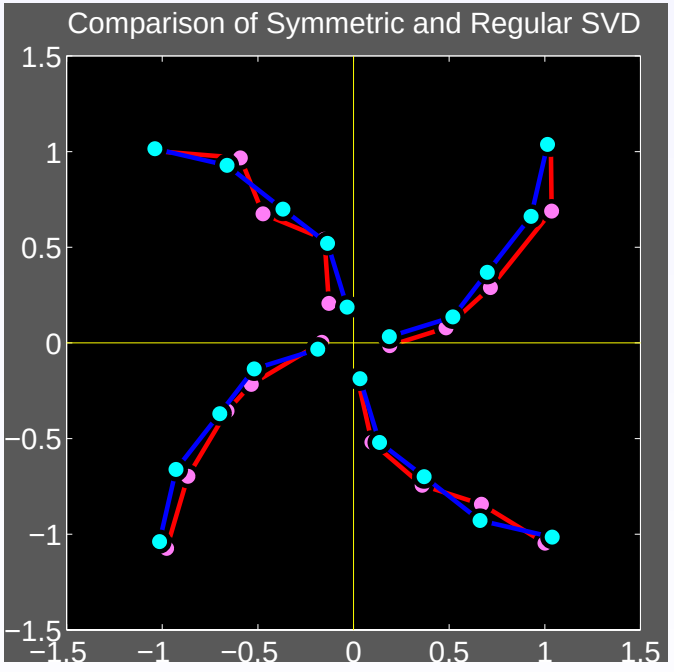
$$\left[\widehat{\mathbf{X}}_0^T \ldots \widehat{\mathbf{X}}_{k-1}^T\right]^T = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad \text{with} \quad \mathbf{U} = \frac{1}{\sqrt{k}} \left[\mathbf{U}_0^T \ldots \mathbf{U}_{k-1}^T\right]^T$$

$$\mathbf{U}_0\mathbf{S}_0\mathbf{V}_0^T = \frac{1}{k}(\mathbf{X}_0 + \mathbf{W}^{k-1}\mathbf{X}_1 + \mathbf{W}^{k-2}\mathbf{X}_2 + \ldots + \mathbf{W}\mathbf{X}_{k-1}).$$

$$\mathbf{S} = \sqrt{k}\mathbf{S}_0 \quad \mathbf{V} = \mathbf{V}_0 \quad \mathbf{U}_j = \mathbf{W}^j\mathbf{U}_0, \quad j = 0, 1, 2, \ldots, k-1$$

Animation: Rotationally Symmetric SVD Approximation

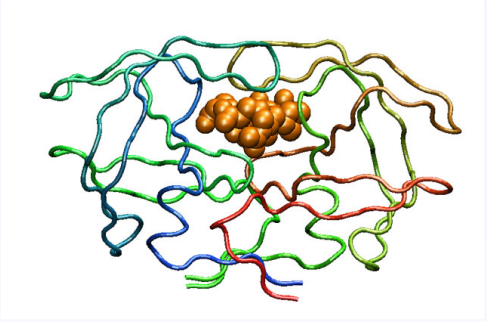




click below figures for movies

Animation: Rotationally Symmetric SVD on HIV1

click for movie



Red = Unsymmetric

Blue = Symmetric

Second SVD mode – Rotationally Symmetric vs. Unsymmetric

Potential for Symmetric SVD

- Obtain a *Symmetric* PCA reduced dimension approximate trajectory
- Test Hypothesis of Symmetry in an Unknown Protein
- Locate Symmeteric Sub-Structures

Things to Do:

- Improve convergence rate for finding **w**
- Give a complete analysis of convergence
- Give a complete analysis of discrepancy weighting
- Extend to more complex symmetries
- Find New Applications

Contact Info.

e-mail: sorensen@rice.edu

web page: www.caam.rice.edu/ ~ sorensen/

ARPACK: www.caam.rice.edu/software/ARPACK/