Model reduction of large-scale systems

An overview and some recent results

Thanos Antoulas

Rice University

email: aca@rice.edu
URL: www.ece.rice.edu/˜aca

POD Workshop, Bordeaux, 31 March - 2 April 2008

Collaborators

- Dan Sorensen, CAAM Rice
- Andrew Mayo, ECE Rice
- Eduardo Gildin, UT Austin
- Mehboob Alam, ECE Rice
- Kai Sun, CAAM Rice
- Roxana Ionutiu, ECE Rice
- Sanda Lefteriu, ECE Rice
- Serkan Gugercin, Virginia Tech
- Chris Beattie, Virginia Tech
- Matthias Heinkenschloss, ECE Rice
- Cosmin Ionita, ECE Rice

Contents

Contents

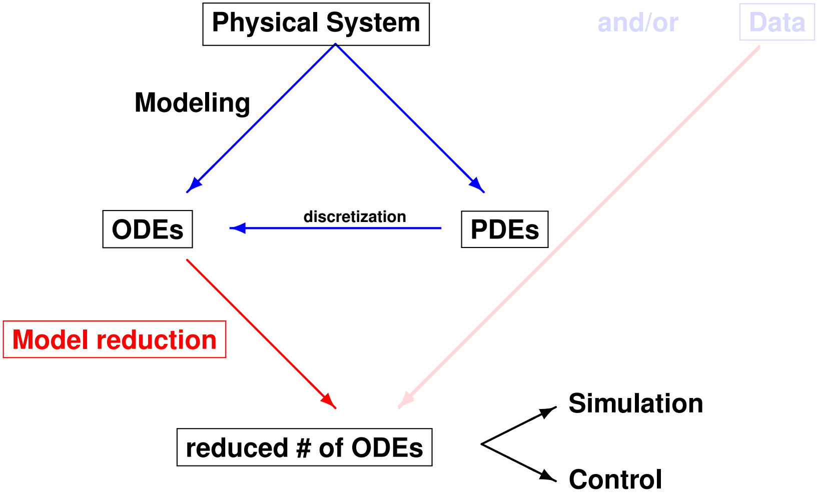
Contents

Contents

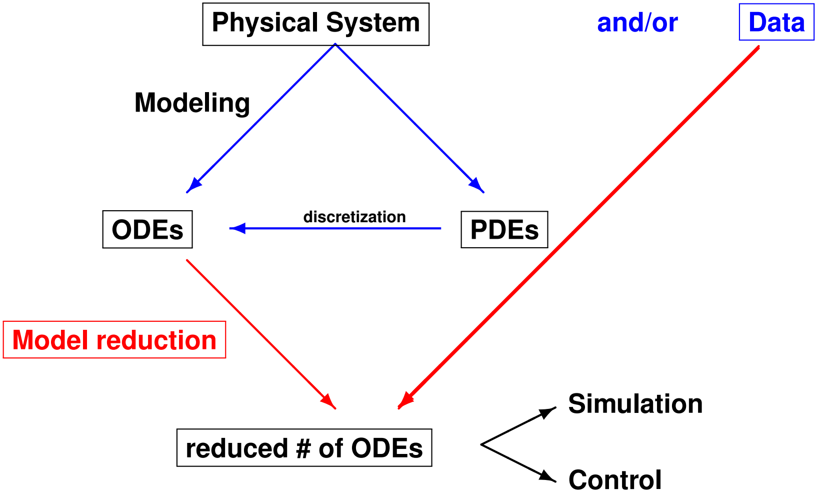
Part I

Introduction and model reduction problem

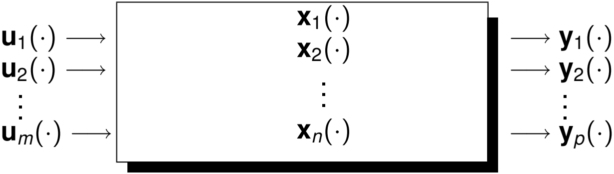
The big picture



The big picture



Dynamical systems



We consider explicit state equations

$$\Sigma : \quad \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \;\; \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t))$$

with **state** $\mathbf{x}(\cdot)$ of dimension $n \gg m, p$.

Problem statement

Given: dynamical system

$$\Sigma = (\mathbf{f}, \mathbf{h}) \text{ with: } \mathbf{u}(t) \in \mathbb{R}^m,\ \mathbf{x}(t) \in \mathbb{R}^n,\ \mathbf{y}(t) \in \mathbb{R}^p.$$

Problem: Approximate Σ with:

$$\hat{\Sigma} = (\hat{\mathbf{f}}, \hat{\mathbf{h}}) \text{ with : } \mathbf{u}(t) \in \mathbb{R}^m,\ \hat{\mathbf{x}}(t) \in \mathbb{R}^k,\ \hat{\mathbf{y}}(t) \in \mathbb{R}^p,\ k \ll n :$$

(1) Approximation error small - global error bound

(2) Preservation of stability/passivity

(3) Procedure must be computationally efficient

Approximation by projection

Unifying feature of approximation methods: **projections**.

Let $\mathbf{V},\ \mathbf{W} \in \mathbb{R}^{n\times k}$, such that $\mathbf{W}^*\mathbf{V} = \mathbf{I}_k \Rightarrow \Pi = \mathbf{V}\mathbf{W}^*$ is a projection.
Define $\hat{\mathbf{x}} = \mathbf{W}^*\mathbf{x}$. Then

$$\hat{\Sigma} : \begin{cases} \frac{d}{dt}\hat{\mathbf{x}}(t) &= \mathbf{W}^*\mathbf{f}(\mathbf{V}\hat{\mathbf{x}}(t),\ \mathbf{u}(t)) \\ \mathbf{y}(t) &= \mathbf{h}(\mathbf{V}\hat{\mathbf{x}}(t),\ \mathbf{u}(t)) \end{cases}$$

Thus $\hat{\Sigma}$ is "good" approximation of Σ, if $\mathbf{x} - \Pi\mathbf{x}$ is "small".

Special case: linear dynamical systems

Σ: $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$$

Problem: Approximate Σ by **projection**: $\Pi = \mathbf{V}\mathbf{W}^*$

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right),\ k \ll n$$

Norms:

- $\mathcal{H}_\infty$-norm: worst output error $\|\mathbf{y}(t) - \hat{\mathbf{y}}(t)\|$ for $\|\mathbf{u}(t)\| = 1$.
- $\mathcal{H}_2$-norm: $\|\mathbf{h}(t) - \hat{\mathbf{h}}(t)\|$

Σ: $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \;\; \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$$

Problem: Approximate Σ by **projection**: $\Pi = \mathbf{V}\mathbf{W}^*$

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right), \;\; k \ll n$$

Norms:

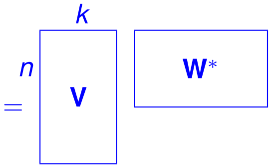
- $\mathcal{H}_\infty$-norm: worst output error $\|\mathbf{y}(t) - \hat{\mathbf{y}}(t)\|$ for $\|\mathbf{u}(t)\| = 1$.
- $\mathcal{H}_2$-norm: $\|\mathbf{h}(t) - \hat{\mathbf{h}}(t)\|$

Special case: linear dynamical systems

Σ: $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),\ \ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$$

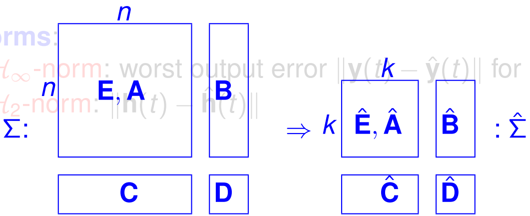
Problem: Approximate Σ by **projection**: $\Pi = \mathbf{V}\mathbf{W}^*$



$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right),\ k \ll n$$

Norms:

- $\mathcal{H}_\infty$-norm: worst output error $\|\mathbf{y}(t) - \hat{\mathbf{y}}(t)\|$ for $\|\mathbf{u}(t)\| = 1$.
- $\mathcal{H}_2$-norm: $\|\mathbf{h}(t) - \hat{\mathbf{h}}(t)\|$



Special case: linear dynamical systems

Σ: $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \;\; \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$$

Problem: Approximate Σ by **projection**: $\Pi = \mathbf{V}\mathbf{W}^*$

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right), \; k \ll n$$

Norms:

- $\mathcal{H}_\infty$-norm: worst output error $\|\mathbf{y}(t) - \hat{\mathbf{y}}(t)\|$ for $\|\mathbf{u}(t)\| = 1$.
- $\mathcal{H}_2$-norm: $\|\mathbf{h}(t) - \hat{\mathbf{h}}(t)\|$

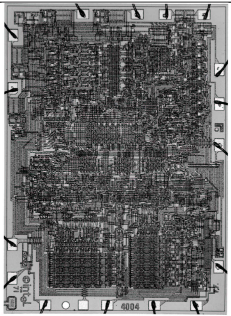
Part II

Motivating examples

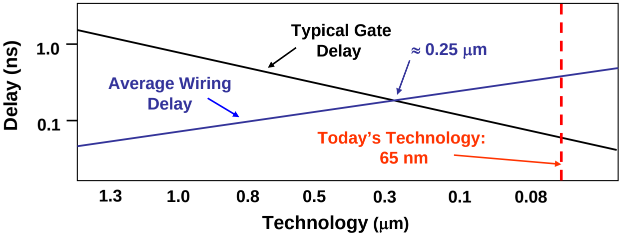
Motivating Examples: Simulation/Control

1. **Passive devices**	• VLSI circuits • Thermal issues • Power delivery networks
2. **Data assimilation**	• North sea forecast • Air quality forecast
3. **Molecular systems**	• MD simulations • Heat capacity
4. **CVD reactor**	• Bifurcations
5. **Mechanical systems**:	•Windscreen vibrations • Buildings
6. **Optimal cooling**	• Steel profile
7. **MEMS**: **Micro Electro- -Mechanical Systems**	• Elf sensor
8. **Nano-Electronics**	• Plasmonics

Passive devices: VLSI circuits

		
1960's: IC	1971: Intel 4004	2001: Intel Pentium IV
	10μ details 2300 components 64KHz speed	0.18μ details 42M components 2GHz speed 2km interconnect 7 layers

Passive devices: VLSI circuits



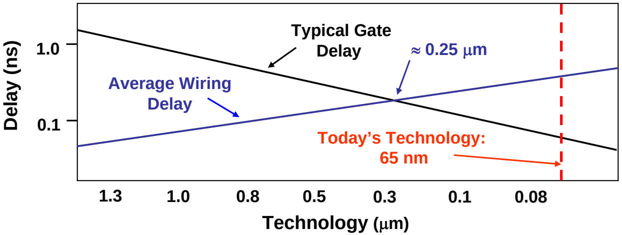
65nm technology: gate delay $<$ interconnect delay!

Conclusion: Simulations are required to verify that internal electromagnetic fields do not significantly delay or distort circuit signals. Therefore interconnections must be modeled.

$\Rightarrow$ Electromagnetic modeling of packages and interconnects $\Rightarrow$ resulting models very complex: using PEEC methods (discretization of Maxwell's equations): $n \approx 10^5 \cdots 10^6 \Rightarrow$ SPICE: inadequate

- Source: van der Meijs (Delft)

Passive devices: VLSI circuits



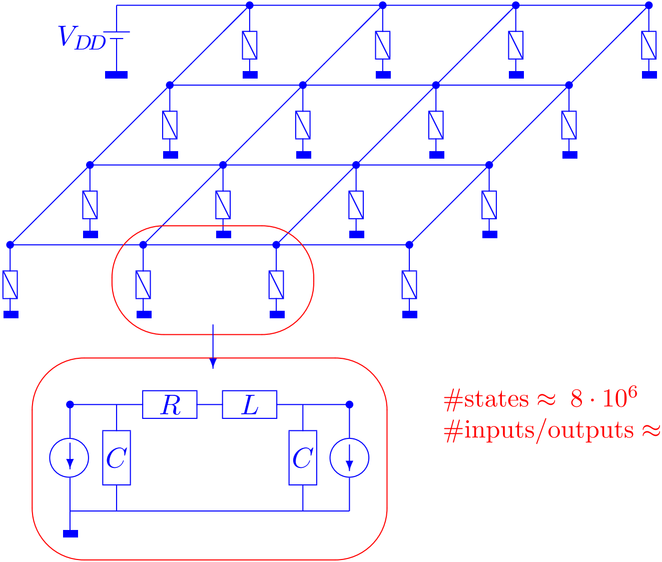
65nm technology: gate delay $<$ interconnect delay!

Conclusion: Simulations are required to verify that internal electromagnetic fields do not significantly delay or distort circuit signals. Therefore interconnections must be modeled.

$\Rightarrow$ Electromagnetic modeling of packages and interconnects $\Rightarrow$ resulting models very complex: using PEEC methods (discretization of Maxwell's equations): $n \approx 10^5 \cdots 10^6 \Rightarrow$ SPICE: inadequate

- Source: van der Meijs (Delft)

Power delivery network for VLSI chips



#states $\approx$ $8 \cdot 10^6$
#inputs/outputs $\approx 1 \cdot 10^6$

Mechanical systems: cars

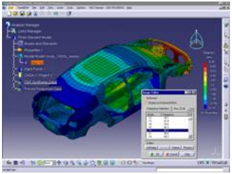
Car windscreen simulation subject to acceleration load.

Problem: compute noise at points away from the window.
PDE: describes deformation of a structure of a specific material; FE discretization: 7564 nodes (3 layers of 60 by 30 elements). Material: glass with Young modulus $7 \cdot 10^{10}\,\mathrm{N/m^2}$; density $2490\,\mathrm{kg/m^3}$; Poisson ratio 0.23 $\Rightarrow$ coefficients of FE model determined experimentally.
The discretized problem has dimension: 22,692.

Notice: this problem yields 2nd order equations:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{f}(t).$$

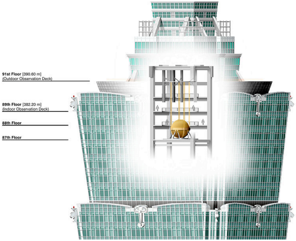
- Source: Meerbergen (Free Field Technologies)



Earthquake prevention

Taipei 101: 508m

Damper between 87-91 floors

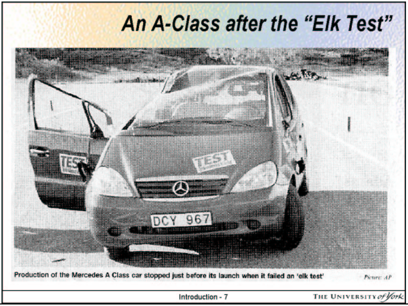


730 ton damper

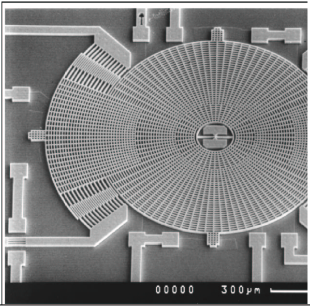
Building	Height	Control mechanism	Damping frequency Damping mass
CN Tower, Toronto	533 m	Passive tuned mass damper	
Hancock building, Boston	244 m	Two passive tuned dampers	0.14Hz, 2x300t
Sydney tower	305 m	Passive tuned pendulum	0.1,0.5z, 220t
Rokko Island P&G, Kobe	117 m	Passive tuned pendulum	0.33-0.62Hz, 270t
Yokohama Landmark tower	296 m	Active tuned mass dampers (2)	0.185Hz, 340t
Shinjuku Park Tower	296 m	Active tuned mass dampers (3)	330t
TYG Building, Atsugi	159 m	Tuned liquid dampers (720)	0.53Hz, 18.2t

- Source: S. Williams

MEMS: Elk sensor



Mercedes A Class



semiconductor rotation sensor

- Source: Laur (Bremen)

Part III

Overview of approximation methods

Approximation methods: Overview

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Krylov/SVD Methods

Approximation methods: Overview

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods • Empirical Gramians	• Balanced truncation • Hankel approximation

Krylov/SVD Methods

Approximation methods: Overview

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Krylov/SVD Methods

Approximation methods: Overview

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods • Empirical Gramians	• Balanced truncation • Hankel approximation

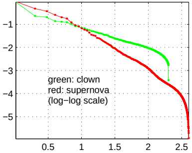
Krylov/SVD Methods

SVD Approximation methods

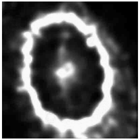
A prototype approximation problem – the **SVD**
(Singular Value Decomposition): $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^*$.

Supernova

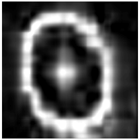
Singular values of Clown and Supernova



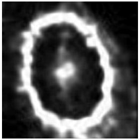
Supernova: original picture



Supernova: rank 6 approximation



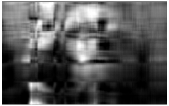
Supernova: rank 20 approximation



Clown

Clown: original picture

Clown: rank 6 approximation



Clown: rank 12 approximation

Clown: rank 20 approximation

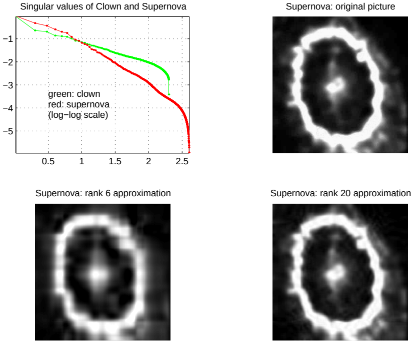
Singular values provide trade-off between accuracy and complexity

SVD Approximation methods

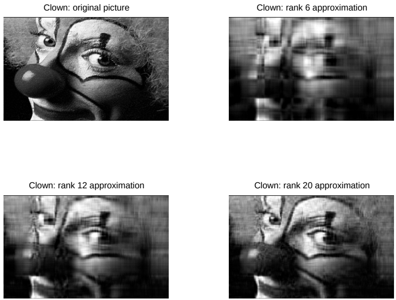
A prototype approximation problem – the SVD

(Singular Value Decomposition): $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^*$.

Supernova



Clown



Singular values provide trade-off between accuracy and complexity

POD: Proper Orthogonal Decomposition

Consider: $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \ \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t))$.

Snapshots of the state:

$$\mathcal{X} = [\mathbf{x}(t_1) \ \ \mathbf{x}(t_2) \ \ \cdots \ \ \mathbf{x}(t_N)] \in \mathbb{R}^{n \times N}$$

SVD: $\mathcal{X} = \mathbf{U}\Sigma\mathbf{V}^* \approx \mathbf{U}_k\Sigma_k\mathbf{V}_k^*, \quad k \ll n$. Approximate the state:

$$\hat{\mathbf{x}}(t) = \mathbf{U}_k^*\mathbf{x}(t) \quad \Rightarrow \quad \mathbf{x}(t) \approx \mathbf{U}_k\hat{\mathbf{x}}(t), \ \ \hat{\mathbf{x}}(t) \in \mathbb{R}^k$$

Project state and output equations. Reduced order system:

$$\boxed{\dot{\hat{\mathbf{x}}}(t) = \mathbf{U}_k^*\mathbf{f}(\mathbf{U}_k\hat{\mathbf{x}}(t), \mathbf{u}(t)), \ \ \mathbf{y}(t) = \mathbf{h}(\mathbf{U}_k\hat{\mathbf{x}}(t), \mathbf{u}(t))}$$

$\Rightarrow \hat{\mathbf{x}}(t)$ evolves in a **low-dimensional** space.

Issues with POD:
(a) Choice of snapshots, (b) singular values not I/O invariants.

POD: Proper Orthogonal Decomposition

Consider: $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)),\ \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t))$.

Snapshots of the state:

$$\mathcal{X} = [\mathbf{x}(t_1)\ \ \mathbf{x}(t_2)\ \ \cdots\ \ \mathbf{x}(t_N)] \in \mathbb{R}^{n \times N}$$

SVD: $\mathcal{X} = \mathbf{U}\Sigma\mathbf{V}^* \approx \mathbf{U}_k\Sigma_k\mathbf{V}_k^*,\ \ k \ll n$. Approximate the state:

$$\hat{\mathbf{x}}(t) = \mathbf{U}_k^*\mathbf{x}(t) \ \ \Rightarrow \ \ \mathbf{x}(t) \approx \mathbf{U}_k\hat{\mathbf{x}}(t),\ \ \hat{\mathbf{x}}(t) \in \mathbb{R}^k$$

Project state and output equations. Reduced order system:

$$\boxed{\dot{\hat{\mathbf{x}}}(t) = \mathbf{U}_k^*\mathbf{f}(\mathbf{U}_k\hat{\mathbf{x}}(t), \mathbf{u}(t)),\ \ \mathbf{y}(t) = \mathbf{h}(\mathbf{U}_k\hat{\mathbf{x}}(t), \mathbf{u}(t))}$$

$\Rightarrow$ $\hat{\mathbf{x}}(t)$ evolves in a **low-dimensional** space.

Issues with POD:

(a) Choice of snapshots, (b) singular values not I/O invariants.

POD: Proper Orthogonal Decomposition

Consider: $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t))$, $\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t))$.

Snapshots of the state:

$$\mathcal{X} = [\mathbf{x}(t_1)\ \ \mathbf{x}(t_2)\ \ \cdots\ \ \mathbf{x}(t_N)] \in \mathbb{R}^{n \times N}$$

SVD: $\mathcal{X} = \mathbf{U}\Sigma\mathbf{V}^* \approx \mathbf{U}_k\Sigma_k\mathbf{V}_k^*$, $k \ll n$. Approximate the state:

$$\hat{\mathbf{x}}(t) = \mathbf{U}_k^*\mathbf{x}(t) \ \Rightarrow\ \mathbf{x}(t) \approx \mathbf{U}_k\hat{\mathbf{x}}(t),\ \ \hat{\mathbf{x}}(t) \in \mathbb{R}^k$$

Project state and output equations. Reduced order system:

$$\boxed{\dot{\hat{\mathbf{x}}}(t) = \mathbf{U}_k^*\mathbf{f}(\mathbf{U}_k\hat{\mathbf{x}}(t), \mathbf{u}(t)),\ \ \mathbf{y}(t) = \mathbf{h}(\mathbf{U}_k\hat{\mathbf{x}}(t), \mathbf{u}(t))}$$

$\Rightarrow$ $\hat{\mathbf{x}}(t)$ evolves in a **low-dimensional** space.

Issues with POD:

(a) Choice of snapshots, (b) singular values not I/O invariants.

SVD methods: balanced truncation

Trade-off between accuracy and complexity for linear dynamical systems is provided by the **Hankel Singular Values**. Define the **gramians** as solutions of the Lyapunov equations

$$\left.\begin{array}{l} \mathbf{AP} + \mathbf{PA}^* + \mathbf{BB}^* = \mathbf{0}, \quad \mathbf{P} > \mathbf{0} \\ \mathbf{A}^*\mathbf{Q} + \mathbf{QA} + \mathbf{C}^*\mathbf{C} = \mathbf{0}, \quad \mathbf{Q} > \mathbf{0} \end{array}\right\} \Rightarrow \boxed{\sigma_i = \sqrt{\lambda_i(\mathbf{PQ})}}$$

σ_i: **Hankel singular values** of the system. There exists balanced basis where $\mathbf{P} = \mathbf{Q} = \mathbf{S} = \mathrm{diag}\,(\sigma_1, \cdots, \sigma_n)$. In this basis partition:

$$\mathbf{A} = \left(\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right), \ \mathbf{B} = \left(\begin{array}{c} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \end{array}\right), \ \mathbf{C} = (\mathbf{C}_1 \mid \mathbf{C}_2), \ \mathbf{S} = \left(\begin{array}{c|c} \Sigma_1 & 0 \\ \hline 0 & \Sigma_2 \end{array}\right).$$

The reduced system is obtained by balanced truncation

$\left(\begin{array}{c|c} \mathbf{A}_{11} & B_1 \\ \hline C_1 & \end{array}\right)$, where Σ_2 contains the small Hankel singular values.

Properties of balanced reduction

1. **Stability is preserved**
2. **Global error bound**:

$$\sigma_{k+1} \leq \| \Sigma - \hat{\Sigma} \|_\infty \leq 2(\sigma_{k+1} + \cdots + \sigma_n)$$

Drawbacks

1. **Dense** computations, matrix factorizations and inversions $\Rightarrow$ may be ill-conditioned
2. Need **whole** transformed system in order to truncate $\Rightarrow$ number of operations $\mathcal{O}(n^3)$
3. **Bottleneck**: **solution of two Lyapunov equations**

Properties of balanced reduction

1. **Stability is preserved**
2. **Global error bound**:

   $$\sigma_{k+1} \leq \| \Sigma - \hat{\Sigma} \|_\infty \leq 2(\sigma_{k+1} + \cdots + \sigma_n)$$

Drawbacks

1. **Dense** computations, matrix factorizations and inversions $\Rightarrow$ may be ill-conditioned
2. Need **whole** transformed system in order to truncate $\Rightarrow$ number of operations $\mathcal{O}(n^3)$
3. **Bottleneck**: **solution of two Lyapunov equations**

Approximation methods: Krylov methods

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods • Empirical Gramians	• Balanced truncation • Hankel approximation

Krylov/SVD Methods

Approximation methods: Krylov methods

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	**Linear systems**
• POD methods • Empirical Gramians	• Balanced truncation • Hankel approximation

Krylov/SVD Methods

The basic Krylov iteration

Given $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{b} \in \mathbb{R}^n$, let $\mathbf{v}_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|}$. At the k^{th} step:

$$\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_k\mathbf{H}_k + \mathbf{f}_k\mathbf{e}_k^*} \qquad \text{where}$$

$\mathbf{e}_k \in \mathbb{R}^k$: canonical unit vector
$\mathbf{V}_k = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k] \in \mathbb{R}^{k\times k}$, $\mathbf{V}_v^*\mathbf{V}_k = \mathbf{I}_k$
$\mathbf{H}_k = \mathbf{V}_k^*\mathbf{A}\mathbf{V}_k \in \mathbb{R}^{k\times k}$

$\Rightarrow \quad \mathbf{v}_{k+1} = \frac{\mathbf{f}_k}{\|\mathbf{f}_k\|} \in \mathbb{R}^n$

Computational complexity for k steps: $\mathcal{O}(n^2k)$; storage $\mathcal{O}(nk)$.

The **Lanczos** and the **Arnoldi** algorithms result.

The **Krylov iteration** involves the subspace $\mathcal{R}_k = \left[\mathbf{b}, \mathbf{A}\mathbf{b}, \cdots, \mathbf{A}^{k-1}\mathbf{b}\right]$.

- Arnoldi iteration $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ upper Hessenberg.
- Symmetric (one-sided) Lanczos iteration $\Rightarrow$ symmetric $\mathbf{A} = \mathbf{A}^*$
  $\Rightarrow \mathbf{H}_k$ tridiagonal and symmetric.
- Two-sided Lanczos iteration with two starting vectors $\mathbf{b}$, $\mathbf{c}$
  $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ tridiagonal.

The basic Krylov iteration

Given $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{b} \in \mathbb{R}^n$, let $\mathbf{v}_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|}$. At the k^{th} step:

$$\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_k\mathbf{H}_k + \mathbf{f}_k\mathbf{e}_k^*} \qquad \text{where}$$

$\mathbf{e}_k \in \mathbb{R}^k$: canonical unit vector
$\mathbf{V}_k = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k] \in \mathbb{R}^{k\times k}$, $\mathbf{V}_v^*\mathbf{V}_k = \mathbf{I}_k$
$\mathbf{H}_k = \mathbf{V}_k^*\mathbf{A}\mathbf{V}_k \in \mathbb{R}^{k\times k}$

$$\Rightarrow \quad \mathbf{v}_{k+1} = \frac{\mathbf{f}_k}{\|\mathbf{f}_k\|} \in \mathbb{R}^n$$

Computational complexity for k steps: $\mathcal{O}(n^2k)$; storage $\mathcal{O}(nk)$.

The **Lanczos** and the **Arnoldi** algorithms result.

The **Krylov iteration** involves the subspace $\mathcal{R}_k = \left[\mathbf{b}, \mathbf{A}\mathbf{b}, \cdots, \mathbf{A}^{k-1}\mathbf{b}\right]$.

- Arnoldi iteration $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ upper Hessenberg.
- Symmetric (one-sided) Lanczos iteration $\Rightarrow$ symmetric $\mathbf{A} = \mathbf{A}^*$
  $\Rightarrow \mathbf{H}_k$ tridiagonal and symmetric.
- Two-sided Lanczos iteration with two starting vectors $\mathbf{b}$, $\mathbf{c}$
  $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ tridiagonal.

The basic Krylov iteration

Given $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{b} \in \mathbb{R}^n$, let $\mathbf{v}_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|}$. At the k^{th} step:

$$\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_k\mathbf{H}_k + \mathbf{f}_k\mathbf{e}_k^*}$$ where

$\mathbf{e}_k \in \mathbb{R}^k$: canonical unit vector
$\mathbf{V}_k = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k] \in \mathbb{R}^{k\times k}$, $\mathbf{V}_v^*\mathbf{V}_k = \mathbf{I}_k$
$\mathbf{H}_k = \mathbf{V}_k^*\mathbf{A}\mathbf{V}_k \in \mathbb{R}^{k\times k}$

$\Rightarrow \quad \mathbf{v}_{k+1} = \frac{\mathbf{f}_k}{\|\mathbf{f}_k\|} \in \mathbb{R}^n$

Computational complexity for k steps: $\mathcal{O}(n^2k)$; storage $\mathcal{O}(nk)$.

The **Lanczos** and the **Arnoldi** algorithms result.

The **Krylov iteration** involves the subspace $\mathcal{R}_k = \left[\mathbf{b}, \mathbf{A}\mathbf{b}, \cdots, \mathbf{A}^{k-1}\mathbf{b}\right]$.

- Arnoldi iteration $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ upper Hessenberg.
- Symmetric (one-sided) Lanczos iteration $\Rightarrow$ symmetric $\mathbf{A} = \mathbf{A}^*$
  $\Rightarrow \mathbf{H}_k$ tridiagonal and symmetric.
- Two-sided Lanczos iteration with two starting vectors $\mathbf{b}$, $\mathbf{c}$
  $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ tridiagonal.

The basic Krylov iteration

Given $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{b} \in \mathbb{R}^n$, let $\mathbf{v}_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|}$. At the k^{th} step:

$$\boxed{\mathbf{AV}_k = \mathbf{V}_k \mathbf{H}_k + \mathbf{f}_k \mathbf{e}_k^*} \qquad \text{where}$$

$\mathbf{e}_k \in \mathbb{R}^k$: canonical unit vector
$\mathbf{V}_k = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k] \in \mathbb{R}^{k \times k}$, $\mathbf{V}_v^* \mathbf{V}_k = \mathbf{I}_k$
$\mathbf{H}_k = \mathbf{V}_k^* \mathbf{A} \mathbf{V}_k \in \mathbb{R}^{k \times k}$

$$\Rightarrow \quad \mathbf{v}_{k+1} = \frac{\mathbf{f}_k}{\|\mathbf{f}_k\|} \in \mathbb{R}^n$$

Computational complexity for k steps: $\mathcal{O}(n^2 k)$; storage $\mathcal{O}(nk)$.

The **Lanczos** and the **Arnoldi** algorithms result.

The **Krylov iteration** involves the subspace $\mathcal{R}_k = \left[\mathbf{b}, \mathbf{Ab}, \cdots, \mathbf{A}^{k-1}\mathbf{b}\right]$.

- Arnoldi iteration $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ upper Hessenberg.
- Symmetric (one-sided) Lanczos iteration $\Rightarrow$ symmetric $\mathbf{A} = \mathbf{A}^*$
  $\Rightarrow \mathbf{H}_k$ tridiagonal and symmetric.
- Two-sided Lanczos iteration with two starting vectors $\mathbf{b}$, $\mathbf{c}$
  $\Rightarrow$ arbitrary $\mathbf{A} \Rightarrow \mathbf{H}_k$ tridiagonal.

Three uses of the Krylov iteration

(1) Iterative solution of $\mathbf{Ax} = \mathbf{b}$: approximate the solution $\mathbf{x}$ iteratively.

(2) Iterative approximation of the eigenvalues of $\mathbf{A}$. In this case $\mathbf{b}$ is not fixed apriori. The eigenvalues of the projected $\mathbf{H}_k$ approximate the dominant eigenvalues of $\mathbf{A}$.

(3) Approximation of linear systems by moment matriching.

$\Rightarrow$ **Item (3) is of interest in the present context.**

Three uses of the Krylov iteration

(1) Iterative solution of $\mathbf{Ax} = \mathbf{b}$: approximate the solution $\mathbf{x}$ iteratively.

(2) Iterative approximation of the eigenvalues of $\mathbf{A}$. In this case $\mathbf{b}$ is not fixed apriori. The eigenvalues of the projected $\mathbf{H}_k$ approximate the dominant eigenvalues of $\mathbf{A}$.

(3) Approximation of linear systems by moment matriching.

$\Rightarrow$ **Item (3) is of interest in the present context.**

Three uses of the Krylov iteration

(1) Iterative solution of $\mathbf{Ax} = \mathbf{b}$: approximate the solution $\mathbf{x}$ iteratively.

(2) Iterative approximation of the eigenvalues of $\mathbf{A}$. In this case $\mathbf{b}$ is not fixed apriori. The eigenvalues of the projected $\mathbf{H}_k$ approximate the dominant eigenvalues of $\mathbf{A}$.

(3) Approximation of linear systems by moment matriching.

$\Rightarrow$ **Item (3) is of interest in the present context.**

Three uses of the Krylov iteration

(1) Iterative solution of $\mathbf{Ax} = \mathbf{b}$: approximate the solution $\mathbf{x}$ iteratively.

(2) Iterative approximation of the eigenvalues of $\mathbf{A}$. In this case $\mathbf{b}$ is not fixed apriori. The eigenvalues of the projected $\mathbf{H}_k$ approximate the dominant eigenvalues of $\mathbf{A}$.

(3) Approximation of linear systems by moment matriching.

$\Rightarrow$ **Item (3) is of interest in the present context.**

Approximation by moment matching

Given $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),\ \ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$, expand transfer function around s_0:

$$\mathbf{G}(s) = \eta_0 + \eta_1(s - s_0) + \eta_2(s - s_0)^2 + \eta_3(s - s_0)^3 + \cdots$$

Moments at s_0: η_j.

Find $\hat{\mathbf{E}}\dot{\hat{\mathbf{x}}}(t) = \hat{\mathbf{A}}\hat{\mathbf{x}}(t) + \hat{\mathbf{B}}\mathbf{u}(t),\ \ \mathbf{y}(t) = \hat{\mathbf{C}}\hat{\mathbf{x}}(t) + \hat{\mathbf{D}}\mathbf{u}(t)$, with

$$\hat{\mathbf{G}}(s) = \hat{\eta}_0 + \hat{\eta}_1(s - s_0) + \hat{\eta}_2(s - s_0)^2 + \hat{\eta}_3(s - s_0)^3 + \cdots$$

such that for appropriate s_0 and ℓ:

$$\boxed{\eta_j = \hat{\eta}_j,\ \ j = 1, 2, \cdots, \ell}$$

Approximation by moment matching

Given $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$, expand transfer function around s_0:

$$\mathbf{G}(s) = \eta_0 + \eta_1(s - s_0) + \eta_2(s - s_0)^2 + \eta_3(s - s_0)^3 + \cdots$$

Moments at s_0: η_j.

Find $\hat{\mathbf{E}}\dot{\hat{\mathbf{x}}}(t) = \hat{\mathbf{A}}\hat{\mathbf{x}}(t) + \hat{\mathbf{B}}\mathbf{u}(t),\ \mathbf{y}(t) = \hat{\mathbf{C}}\hat{\mathbf{x}}(t) + \hat{\mathbf{D}}\mathbf{u}(t)$, with

$$\hat{\mathbf{G}}(s) = \hat{\eta}_0 + \hat{\eta}_1(s - s_0) + \hat{\eta}_2(s - s_0)^2 + \hat{\eta}_3(s - s_0)^3 + \cdots$$

such that for appropriate s_0 and ℓ:

$$\boxed{\eta_j = \hat{\eta}_j,\quad j = 1, 2, \cdots, \ell}$$

Approximation by moment matching

Given $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),\ \ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$, expand transfer function around s_0:

$$\mathbf{G}(s) = \eta_0 + \eta_1(s - s_0) + \eta_2(s - s_0)^2 + \eta_3(s - s_0)^3 + \cdots$$

Moments at s_0: η_j.

Find $\hat{\mathbf{E}}\dot{\hat{\mathbf{x}}}(t) = \hat{\mathbf{A}}\hat{\mathbf{x}}(t) + \hat{\mathbf{B}}\mathbf{u}(t),\ \ \mathbf{y}(t) = \hat{\mathbf{C}}\hat{\mathbf{x}}(t) + \hat{\mathbf{D}}\mathbf{u}(t)$, with

$$\hat{\mathbf{G}}(s) = \hat{\eta}_0 + \hat{\eta}_1(s - s_0) + \hat{\eta}_2(s - s_0)^2 + \hat{\eta}_3(s - s_0)^3 + \cdots$$

such that for appropriate s_0 and ℓ:

$$\boxed{\eta_j = \hat{\eta}_j,\ \ j = 1, 2, \cdots, \ell}$$

Projectors for Krylov and rational Krylov methods

Given:

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) \text{ by } \textbf{projection}: \Pi = \mathbf{V}\mathbf{W}^*,\ \Pi^2 = \Pi \text{ obtain}$$

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right), \text{ where } k < n.$$

Krylov (Lanczos, Arnoldi): let $\mathbf{E} = \mathbf{I}$ and

$$\mathbf{V} = \left[\mathbf{B},\ \mathbf{AB},\ \cdots,\ \mathbf{A}^{k-1}\mathbf{B}\right] \in \mathbb{R}^{n\times k}$$

$$\bar{\mathbf{W}}^* = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{k-1} \end{bmatrix} \in \mathbb{R}^{k\times n}$$

$$\Rightarrow \mathbf{W}^* = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*$$

then the **Markov parameters match**:

$$\mathbf{C}\mathbf{A}^i\mathbf{B} = \hat{\mathbf{C}}\hat{\mathbf{A}}^i\hat{\mathbf{B}}$$

Rational Krylov: let

$$\mathbf{V} = \left[(\lambda_1\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\ \cdots\ (\lambda_k\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\right] \in \mathbb{R}^{n\times k}$$

$$\bar{\mathbf{W}}^* = \begin{bmatrix} \mathbf{C}(\lambda_{k+1}\mathbf{E}-\mathbf{A})^{-1} \\ \mathbf{C}(\lambda_{k+2}\mathbf{E}-\mathbf{A})^{-1} \\ \vdots \\ \mathbf{C}(\lambda_{2k}\mathbf{E}-\mathbf{A})^{-1} \end{bmatrix} \in \mathbb{R}^{k\times n}$$

$$\Rightarrow \mathbf{W}^* = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*$$

then the **moments of $\hat{\mathbf{G}}$ match those of $\mathbf{G}$ at λ_i**:

$$\mathbf{G}(\lambda_i) = \mathbf{D}+\mathbf{C}(\lambda_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{B} = \hat{\mathbf{D}}+\hat{\mathbf{C}}(\lambda_i\hat{\mathbf{E}}-\hat{\mathbf{A}})^{-1}\hat{\mathbf{B}} = \hat{\mathbf{G}}(\lambda_i)$$

Projectors for Krylov and rational Krylov methods

Given:

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$$ by **projection**: $\Pi = \mathbf{V}\mathbf{W}^*$, $\Pi^2 = \Pi$ obtain

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right), \text{ where } k < n.$$

Krylov (Lanczos, Arnoldi): let $\mathbf{E} = \mathbf{I}$ and

$$\mathbf{V} = \left[\mathbf{B}, \ \mathbf{A}\mathbf{B}, \ \cdots, \ \mathbf{A}^{k-1}\mathbf{B}\right] \in \mathbb{R}^{n \times k}$$

$$\bar{\mathbf{W}}^* = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \vdots \\ \mathbf{C}\mathbf{A}^{k-1} \end{bmatrix} \in \mathbb{R}^{k \times n}$$

$$\Rightarrow \ \mathbf{W}^* = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*$$

then the **Markov parameters match**:

$$\mathbf{C}\mathbf{A}^{\mathbf{i}}\mathbf{B} = \hat{\mathbf{C}}\hat{\mathbf{A}}^{\mathbf{i}}\hat{\mathbf{B}}$$

Rational Krylov: let

$$\mathbf{V} = \left[(\lambda_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} \ \cdots \ (\lambda_k\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\right] \in \mathbb{R}^{n \times k}$$

$$\bar{\mathbf{W}}^* = \begin{bmatrix} \mathbf{C}(\lambda_{k+1}\mathbf{E} - \mathbf{A})^{-1} \\ \mathbf{C}(\lambda_{k+2}\mathbf{E} - \mathbf{A})^{-1} \\ \vdots \\ \mathbf{C}(\lambda_{2k}\mathbf{E} - \mathbf{A})^{-1} \end{bmatrix} \in \mathbb{R}^{k \times n}$$

$$\Rightarrow \ \mathbf{W}^* = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*$$

then the **moments of $\hat{\mathbf{G}}$ match those of $\mathbf{G}$ at λ_i**:

$$\mathbf{G}(\lambda_\mathbf{i}) = \mathbf{D} + \mathbf{C}(\lambda_\mathbf{i}\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} = \hat{\mathbf{D}} + \hat{\mathbf{C}}(\lambda_\mathbf{i}\hat{\mathbf{E}} - \hat{\mathbf{A}})^{-1}\hat{\mathbf{B}} = \hat{\mathbf{G}}(\lambda_\mathbf{i})$$

Projectors for Krylov and rational Krylov methods

Given:

$$\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$$ by **projection**: $\Pi = \mathbf{V}\mathbf{W}^*$, $\Pi^2 = \Pi$ obtain

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{\mathbf{E}}, \hat{\mathbf{A}} & \hat{\mathbf{B}} \\ \hline \hat{\mathbf{C}} & \hat{\mathbf{D}} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{W}^*\mathbf{E}\mathbf{V}, \mathbf{W}^*\mathbf{A}\mathbf{V} & \mathbf{W}^*\mathbf{B} \\ \hline \mathbf{C}\mathbf{V} & \mathbf{D} \end{array}\right),$$ where $k < n$.

Krylov (Lanczos, Arnoldi): let $\mathbf{E} = \mathbf{I}$ and

$$\mathbf{V} = \left[\mathbf{B}, \ \mathbf{A}\mathbf{B}, \ \cdots, \ \mathbf{A}^{k-1}\mathbf{B}\right] \in \mathbb{R}^{n\times k}$$

$$\bar{\mathbf{W}}^* = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \vdots \\ \mathbf{C}\mathbf{A}^{k-1} \end{bmatrix} \in \mathbb{R}^{k\times n}$$

$$\Rightarrow \ \mathbf{W}^* = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*$$

then the **Markov parameters match**:

$$\mathbf{C}\mathbf{A}^{\mathbf{i}}\mathbf{B} = \hat{\mathbf{C}}\hat{\mathbf{A}}^{\mathbf{i}}\hat{\mathbf{B}}$$

Rational Krylov: let

$$\mathbf{V} = \left[(\lambda_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} \ \cdots \ (\lambda_k\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\right] \in \mathbb{R}^{n\times k}$$

$$\bar{\mathbf{W}}^* = \begin{bmatrix} \mathbf{C}(\lambda_{k+1}\mathbf{E} - \mathbf{A})^{-1} \\ \mathbf{C}(\lambda_{k+2}\mathbf{E} - \mathbf{A})^{-1} \\ \vdots \\ \mathbf{C}(\lambda_{2k}\mathbf{E} - \mathbf{A})^{-1} \end{bmatrix} \in \mathbb{R}^{k\times n}$$

$$\Rightarrow \ \mathbf{W}^* = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*$$

then the **moments of $\hat{\mathbf{G}}$ match those of G at λ_i**:

$$\mathbf{G}(\lambda_\mathbf{i}) = \mathbf{D} + \mathbf{C}(\lambda_\mathbf{i}\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} = \hat{\mathbf{D}} + \hat{\mathbf{C}}(\lambda_\mathbf{i}\hat{\mathbf{E}} - \hat{\mathbf{A}})^{-1}\hat{\mathbf{B}} = \hat{\mathbf{G}}(\lambda_\mathbf{i})$$

Properties of Krylov methods

(a) Number of operations: $\mathcal{O}(kn^2)$ or $\mathcal{O}(k^2 n)$ vs. $\mathcal{O}(n^3)$ $\Rightarrow$ efficiency

(b) Only matrix-vector multiplications are required. No matrix factorizations and/or inversions. No need to compute transformed model and then truncate.

(c) **Drawbacks**

- global error bound?
- $\hat{\Sigma}$ may not be stable.

Q: **How to choose the projection points?**

Properties of Krylov methods

(a) Number of operations: $\mathcal{O}(kn^2)$ or $\mathcal{O}(k^2 n)$ vs. $\mathcal{O}(n^3)$ $\Rightarrow$ efficiency

(b) Only matrix-vector multiplications are required. No matrix factorizations and/or inversions. No need to compute transformed model and then truncate.

(c) **Drawbacks**

- global error bound?
- $\hat{\Sigma}$ may not be stable.

Q: **How to choose the projection points?**

Properties of Krylov methods

(a) Number of operations: $\mathcal{O}(kn^2)$ or $\mathcal{O}(k^2n)$ vs. $\mathcal{O}(n^3)$ $\Rightarrow$ efficiency

(b) Only matrix-vector multiplications are required. No matrix factorizations and/or inversions. No need to compute transformed model and then truncate.

(c) **Drawbacks**

- global error bound?
- $\hat{\Sigma}$ may not be stable.

Q: **How to choose the projection points?**

Part IV

Approximation methods: two recent results

Choice of projection points in Krylov methods

1. **Passivity preserving model reduction.**
2. **Optimal $\mathcal{H}_2$ model reduction.**

Passive systems:

$\mathcal{R}e \int_{-\infty}^{t} \mathbf{u}(\tau)^* \mathbf{y}(\tau) d\tau \geq 0, \forall\, t \in \mathbb{R}, \forall\, \mathbf{u} \in \mathcal{L}_2(\mathbb{R})$.

Positive real rational functions:

(1) $\mathbf{G}(s) = \mathbf{D} + \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$, is analytic for $\mathcal{R}e(s) > 0$,
(2) $\mathcal{R}e\, \mathbf{G}(s) \geq 0$ for $\mathcal{R}e(s) \geq 0$, s not a pole of $G(s)$.

Theorem: $\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ is passive $\Leftrightarrow$ $\mathbf{G}(s)$ is positive real.

Conclusion: Positive realness of $\mathbf{G}(s)$ implies the existence of a **spectral factorization** $\mathbf{G}(s) + \mathbf{G}^*(-s) = \mathbf{W}(s)\mathbf{W}^*(-s)$, where $\mathbf{W}(s)$ is stable rational and $\mathbf{W}(s)^{-1}$ is also stable. The **spectral zeros** λ_i of the system are the zeros of the spectral factor $\mathbf{W}(\lambda_i) = 0$, $i = 1, \cdots, n$.

Choice of Krylov projection points:

Passivity preserving model reduction

Passive systems:
$\mathcal{R}e \int_{-\infty}^{t} \mathbf{u}(\tau)^{*}\mathbf{y}(\tau)\mathrm{d}\tau \geq 0, \forall\, t \in \mathbb{R}, \forall\, \mathbf{u} \in \mathcal{L}_2(\mathbb{R})$.

Positive real rational functions:
(1) $\mathbf{G}(s) = \mathbf{D} + \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$, is analytic for $\mathcal{R}e(s) > 0$,
(2) $\mathcal{R}e\, \mathbf{G}(s) \geq 0$ for $\mathcal{R}e(s) \geq 0$, s not a pole of $G(s)$.

Theorem: $\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ is passive $\Leftrightarrow$ $\mathbf{G}(s)$ is positive real.

Conclusion: Positive realness of $\mathbf{G}(s)$ implies the existence of a **spectral factorization** $\mathbf{G}(s) + \mathbf{G}^{*}(-s) = \mathbf{W}(s)\mathbf{W}^{*}(-s)$, where $\mathbf{W}(s)$ is stable rational and $\mathbf{W}(s)^{-1}$ is also stable. The **spectral zeros** λ_i of the system are the zeros of the spectral factor $\mathbf{W}(\lambda_i) = 0$, $i = 1, \cdots, n$.

Choice of Krylov projection points:

Passivity preserving model reduction

Passive systems:

$\mathcal{R}e \int_{-\infty}^{t} \mathbf{u}(\tau)^* \mathbf{y}(\tau) d\tau \geq 0, \forall\, t \in \mathbb{R}, \forall\, \mathbf{u} \in \mathcal{L}_2(\mathbb{R})$.

Positive real rational functions:

(1) $\mathbf{G}(s) = \mathbf{D} + \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$, is analytic for $\mathcal{R}e(s) > 0$,
(2) $\mathcal{R}e\, \mathbf{G}(s) \geq 0$ for $\mathcal{R}e(s) \geq 0$, s not a pole of $G(s)$.

Theorem: $\Sigma = \left(\begin{array}{c|c} \mathbf{E}, \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ is passive $\Leftrightarrow$ $\mathbf{G}(s)$ is positive real.

Conclusion: Positive realness of $\mathbf{G}(s)$ implies the existence of a **spectral factorization** $\mathbf{G}(s) + \mathbf{G}^*(-s) = \mathbf{W}(s)\mathbf{W}^*(-s)$, where $\mathbf{W}(s)$ is stable rational and $\mathbf{W}(s)^{-1}$ is also stable. The **spectral zeros** λ_i of the system are the zeros of the spectral factor $\mathbf{W}(\lambda_i) = 0$, $i = 1, \cdots, n$.

- **Method**: Rational Krylov
- **Solution**: projection points = spectral zeros

Recall: $$\begin{cases} \mathbf{V} = \left[(\lambda_1 \mathbf{E} - \mathbf{A})^{-1}\mathbf{B} \ \cdots \ (\lambda_k \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\right] \in \mathbb{R}^{n \times k} \\ \mathbf{W}^* = \begin{bmatrix} \mathbf{C}(\lambda_{k+1}\mathbf{E} - \mathbf{A})^{-1} \\ \vdots \\ \mathbf{C}(\lambda_{2k}\ \mathbf{E} - \mathbf{A})^{-1} \end{bmatrix} \in \mathbb{R}^{k \times n} \end{cases}$$

Main result. If **V**, **W** are defined as above, where $\lambda_1, \cdots, \lambda_k$ are **spectral zeros**, and in addition $\boxed{\lambda_{k+i} = -\lambda_i^*}$, the reduced system satisfies:

(i) the interpolation constraints,
(ii) it is stable, and
(iii) it is passive.

Passivity preserving model reduction

New result

- **Method**: Rational Krylov
- **Solution**: projection points = spectral zeros

Recall:
$$\begin{cases} \mathbf{V} = \left[(\lambda_1\mathbf{E}-\mathbf{A})^{-1}\mathbf{B} \ \cdots \ (\lambda_k\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\right] \in \mathbb{R}^{n\times k} \\ \mathbf{W}^* = \begin{bmatrix} \mathbf{C}(\lambda_{k+1}\mathbf{E}-\mathbf{A})^{-1} \\ \vdots \\ \mathbf{C}(\lambda_{2k}\mathbf{E}-\mathbf{A})^{-1} \end{bmatrix} \in \mathbb{R}^{k\times n} \end{cases}$$

Main result. If **V**, **W** are defined as above, where $\lambda_1, \cdots, \lambda_k$ are **spectral zeros**, and in addition $\boxed{\lambda_{k+i} = -\lambda_i^*}$, the reduced system satisfies:

(i) the interpolation constraints,
(ii) it is stable, and
(iii) it is passive.

- **Method**: Rational Krylov
- **Solution**: projection points = spectral zeros

Recall:
$$\begin{cases} \mathbf{V} = \left[(\lambda_1 \mathbf{E} - \mathbf{A})^{-1}\mathbf{B} \ \cdots \ (\lambda_k \mathbf{E} - \mathbf{A})^{-1}\mathbf{B} \right] \in \mathbb{R}^{n\times k} \\ \mathbf{W}^* = \begin{bmatrix} \mathbf{C}(\lambda_{k+1}\mathbf{E} - \mathbf{A})^{-1} \\ \vdots \\ \mathbf{C}(\lambda_{2k}\ \mathbf{E} - \mathbf{A})^{-1} \end{bmatrix} \in \mathbb{R}^{k\times n} \end{cases}$$

Main result. If **V**, **W** are defined as above, where $\lambda_1, \cdots, \lambda_k$ are **spectral zeros**, and in addition $\boxed{\lambda_{k+i} = -\lambda_i^*}$, the reduced system satisfies:

(i) the interpolation constraints,
(ii) it is stable, and
(iii) it is passive.

Spectral zero interpolation preserving passivity

Hamiltonian EVD & projection

- **Hamiltonian eigenvalue problem**

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{0} & -\mathbf{A}^* & -\mathbf{C}^* \\ \mathbf{C} & \mathbf{B}^* & \Delta^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} \Lambda$$

The generalized eigenvalues Λ are the spectral zeros of Σ

- **Partition eigenvectors**

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_- & \mathbf{X}_+ \\ \mathbf{Y}_- & \mathbf{Y}_+ \\ \mathbf{Z}_- & \mathbf{Z}_+ \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \Lambda_- & & \\ & \Lambda_+ & \\ & & \pm\infty \end{bmatrix}$$

Λ_- are the stable spectral zeros

- **Projection**
  - $\mathbf{V} = \mathbf{X}_-, \ \mathbf{W} = \mathbf{Y}_-$
  - $\hat{\mathbf{E}} = \mathbf{W}^*\mathbf{E}\mathbf{V}, \ \hat{\mathbf{A}} = \mathbf{W}^*\mathbf{A}\mathbf{V}, \ \hat{\mathbf{B}} = \mathbf{W}^*\mathbf{B}, \ \hat{\mathbf{C}} = \mathbf{C}\mathbf{V}, \ \hat{\mathbf{D}} = \mathbf{D}$

- **Hamiltonian eigenvalue problem**

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{0} & -\mathbf{A}^* & -\mathbf{C}^* \\ \mathbf{C} & \mathbf{B}^* & \Delta^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} \Lambda$$

The generalized eigenvalues Λ are the spectral zeros of Σ

- **Partition eigenvectors**

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_- & \mathbf{X}_+ \\ \mathbf{Y}_- & \mathbf{Y}_+ \\ \mathbf{Z}_- & \mathbf{Z}_+ \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \Lambda_- & & \\ & \Lambda_+ & \\ & & \pm\infty \end{bmatrix}$$

Λ_- are the stable spectral zeros

- **Projection**
  - $\mathbf{V} = \mathbf{X}_-$, $\mathbf{W} = \mathbf{Y}_-$
  - $\hat{\mathbf{E}} = \mathbf{W}^*\mathbf{E}\mathbf{V}$, $\hat{\mathbf{A}} = \mathbf{W}^*\mathbf{A}\mathbf{V}$, $\hat{\mathbf{B}} = \mathbf{W}^*\mathbf{B}$, $\hat{\mathbf{C}} = \mathbf{C}\mathbf{V}$, $\hat{\mathbf{D}} = \mathbf{D}$

Spectral zero interpolation preserving passivity

Hamiltonian EVD & projection

- **Hamiltonian eigenvalue problem**

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{0} & -\mathbf{A}^* & -\mathbf{C}^* \\ \mathbf{C} & \mathbf{B}^* & \Delta^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} \Lambda$$

The generalized eigenvalues Λ are the spectral zeros of Σ

- **Partition eigenvectors**

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_- & \mathbf{X}_+ \\ \mathbf{Y}_- & \mathbf{Y}_+ \\ \mathbf{Z}_- & \mathbf{Z}_+ \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \Lambda_- & & \\ & \Lambda_+ & \\ & & \pm\infty \end{bmatrix}$$

Λ_- are the stable spectral zeros

- **Projection**
  - $\mathbf{V} = \mathbf{X}_-, \ \mathbf{W} = \mathbf{Y}_-$
  - $\hat{\mathbf{E}} = \mathbf{W}^*\mathbf{E}\mathbf{V}, \ \hat{\mathbf{A}} = \mathbf{W}^*\mathbf{A}\mathbf{V}, \ \hat{\mathbf{B}} = \mathbf{W}^*\mathbf{B}, \ \hat{\mathbf{C}} = \mathbf{C}\mathbf{V}, \ \hat{\mathbf{D}} = \mathbf{D}$

Dominant spectral zeros – SADPA

What is a good choice of k spectral zeros out of n ?

- **Dominance criterion**: Spectral zero s_j is **dominant** if: $\frac{|R_j|}{|\Re(s_j)|}$, is large.
- Efficient computation for large scale systems: we compute the $k \ll n$ **most dominant** eigenmodes of the Hamiltonian pencil.
- **SADPA** (Subspace Accelerated Dominant Pole Algorithm) solves this **iteratively**.

Conclusion:

Passivity preserving model reduction becomes a **structured eigenvalue problem**

Dominant spectral zeros – SADPA

What is a good choice of k spectral zeros out of n ?

- **Dominance criterion**: Spectral zero s_j is **dominant** if: $\frac{|R_j|}{|\Re(s_j)|}$, is large.
- Efficient computation for large scale systems: we compute the $k \ll n$ **most dominant** eigenmodes of the Hamiltonian pencil.
- **SADPA** (Subspace Accelerated Dominant Pole Algorithm) solves this **iteratively**.

Conclusion:

Passivity preserving model reduction becomes a

structured eigenvalue problem

Choice of Krylov projection points:

Optimal $\mathcal{H}_2$ model reduction

The $\mathcal{H}_2$ norm of a (scalar) system is:

$$\|\Sigma\|_{\mathcal{H}_2} = \left(\int_{-\infty}^{+\infty} \mathbf{h}^2(t)dt \right)^{1/2}$$

Goal: construct a **Krylov projection** such that

$$\Sigma_k = \arg \min_{\substack{\deg(\hat{\Sigma}) = r \\ \hat{\Sigma}\,:\,\text{stable}}} \left\| \Sigma - \hat{\Sigma} \right\|_{\mathcal{H}_2}.$$

That is, find a **Krylov projection** $\Pi = \mathbf{V}\mathbf{W}^*$, $\mathbf{V}, \mathbf{W} \in \mathbb{R}^{n \times k}$, $\mathbf{W}^*\mathbf{V} = \mathbf{I}_k$, such that:

$$\boxed{\hat{\mathbf{A}} = \mathbf{W}^*\mathbf{A}\mathbf{V}, \quad \hat{\mathbf{B}} = \mathbf{W}^*\mathbf{B}, \quad \hat{\mathbf{C}} = \mathbf{C}\mathbf{V}}$$

Necessary optimality conditions & resulting algorithm

Let $(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}})$ solve the optimal $\mathcal{H}_2$ problem and let $\hat{\lambda}_i$ denote the eigenvalues of $\hat{\mathbf{A}}$. The necessary optimality conditions are

$$\mathbf{G}(-\hat{\lambda}_i^*) = \hat{\mathbf{G}}(-\hat{\lambda}_i^*) \quad \text{and} \quad \frac{d}{ds}\mathbf{G}(s)\Big|_{s=-\hat{\lambda}_i^*} = \frac{d}{ds}\hat{\mathbf{G}}(s)\Big|_{s=-\hat{\lambda}_i^*}$$

Thus the reduced system has to match the first two moments of the original system at the *mirror images* of the eigenvalues of $\hat{A}$. The proposed algorithm produces such a reduced order system.

1. Make an initial selection of σ_i, for $i = 1, \ldots, k$
2. $\bar{\mathbf{W}} = [(\sigma_1\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*, \cdots, (\sigma_k\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*]$
3. $\mathbf{V} = [(\sigma_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}, \cdots, (\sigma_k\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}]$
4. while (not converged)
   1. $\hat{\mathbf{A}} = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*\mathbf{A}\mathbf{V}$,
   2. $\sigma_i \longleftarrow -\lambda_i(\hat{\mathbf{A}})$ + Newton correction, $i = 1, \ldots, k$,
   3. $\bar{\mathbf{W}} = [(\sigma_1\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*, \cdots, (\sigma_k\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*]$,
   4. $\mathbf{V} = [(\sigma_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}, \cdots, (\sigma_k\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}]$
5. $\hat{\mathbf{A}} = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*\mathbf{A}\mathbf{V}$, $\hat{\mathbf{B}} = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*\mathbf{B}$, $\hat{\mathbf{C}} = \mathbf{C}\mathbf{V}$

Necessary optimality conditions & resulting algorithm

Let $(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}})$ solve the optimal $\mathcal{H}_2$ problem and let $\hat{\lambda}_i$ denote the eigenvalues of $\hat{\mathbf{A}}$. The necessary optimality conditions are

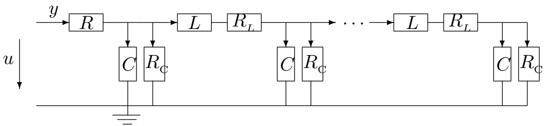
$$\mathbf{G}(-\hat{\lambda}_i^*) = \hat{\mathbf{G}}(-\hat{\lambda}_i^*) \quad \text{and} \quad \frac{d}{ds}\mathbf{G}(s)\Big|_{s=-\hat{\lambda}_i^*} = \frac{d}{ds}\hat{\mathbf{G}}(s)\Big|_{s=-\hat{\lambda}_i^*}$$

Thus the reduced system has to match the first two moments of the original system at the *mirror images* of the eigenvalues of $\hat{A}$. The proposed algorithm produces such a reduced order system.

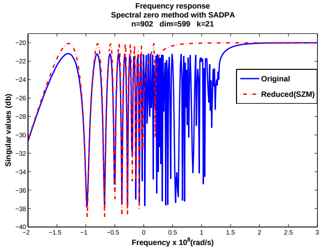
1. **Make an initial selection of σ_i, for $i = 1, \ldots, k$**
2. $\bar{\mathbf{W}} = [(\sigma_1\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*, \cdots, (\sigma_k\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*]$
3. $\mathbf{V} = [(\sigma_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}, \cdots, (\sigma_k\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}]$
4. **while (not converged)**
   1. $\hat{\mathbf{A}} = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*\mathbf{A}\mathbf{V}$,
   2. $\sigma_i \longleftarrow -\lambda_i(\hat{\mathbf{A}}) +$ **Newton correction,** $i = 1, \ldots, k$,
   3. $\bar{\mathbf{W}} = [(\sigma_1\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*, \cdots, (\sigma_k\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*]$,
   4. $\mathbf{V} = [(\sigma_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}, \cdots, (\sigma_k\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}]$
5. $\hat{\mathbf{A}} = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*\mathbf{A}\mathbf{V}$, $\hat{\mathbf{B}} = (\bar{\mathbf{W}}^*\mathbf{V})^{-1}\bar{\mathbf{W}}^*\mathbf{B}$, $\hat{\mathbf{C}} = \mathbf{C}\mathbf{V}$

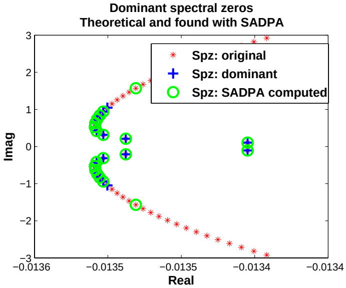
Moderate-dimensional example

SZM with SADPA implementation



- total system variables $\mathrm{n} = 902$, independent variables $\mathrm{dim} = 599$, reduced dimension $\mathrm{k} = 21$
- SADPA computed $2\mathrm{k} = 42$ dominant spectral zeros automatically (95 iterations, CPU time: $\sim$ 16 s)
- reduced model captures dominant modes





$\mathcal{H}_\infty$ and $\mathcal{H}_2$ error norms

Relative norms of the error systems

Reduction Method $n = 902, dim = 599, k = 21$	$\mathcal{H}_\infty$	$\mathcal{H}_2$
PRIMA	1.4775	-
Spectral Zero Method with SADPA	0.9628	0.841
Optimal $\mathcal{H}_2$	**0.5943**	**0.4621**
Balanced truncation (BT)	0.9393	0.6466
Riccati Balanced Truncation (PRBT)	0.9617	0.8164

Approximation methods: Summary

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	**Linear systems**
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Properties

- numerical efficiency
- n $\gg 10^3$
- choice of matching moments

Krylov/SVD Methods

Properties

- Stability
- Error bound
- n $\approx 10^3$

Approximation methods: Summary

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Properties

- numerical efficiency
- n $\gg 10^3$
- choice of matching moments

Krylov/SVD Methods

Properties

- Stability
- Error bound
- n $\approx 10^3$

Approximation methods: Summary

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	Linear systems
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Properties

- **numerical efficiency**
- **n $\gg 10^3$**
- **choice of matching moments**

Krylov/SVD Methods

Properties

- **Stability**
- **Error bound**
- **n $\approx 10^3$**

Approximation methods: Summary

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	**Linear systems**
• POD methods	• Balanced truncation
• Empirical Gramians	• Hankel approximation

Properties

- **numerical efficiency**
- **n $\gg 10^3$**
- **choice of matching moments**

Krylov/SVD Methods

Properties

- **Stability**
- **Error bound**
- **n $\approx 10^3$**

Approximation methods: Summary

Krylov

- Realization
- Interpolation
- Lanczos
- Arnoldi

SVD

Nonlinear systems	**Linear systems**
• POD methods • Empirical Gramians	• Balanced truncation • Hankel approximation

Properties

- **numerical efficiency**
- **n $\gg 10^3$**
- **choice of matching moments**

Krylov/SVD Methods

Properties

- **Stability**
- **Error bound**
- **n $\approx 10^3$**

Model reduction from data:

On-chip analog electronics

Chips for communication systems consist of large analog and RF blocks. To avoid costly re-fabrication, a verification cycle is developed for simulation and design optimization. A common approach to this verification is to replace the circuit block layout by systems of equations and subsequently use their accurate approximants for system simulation. Example: FPGA (Field Programmable Gate Arrays).

Methodology. An input-output approach for modeling of the analog systems can be employed. It treats them as **black boxes**. In the linear passive case, this leads to identification problems using

multi-port **S-parameters**

Chips for communication systems consist of large analog and RF blocks. To avoid costly re-fabrication, a verification cycle is developed for simulation and design optimization. A common approach to this verification is to replace the circuit block layout by systems of equations and subsequently use their accurate approximants for system simulation. Example: FPGA (Field Programmable Gate Arrays).

Methodology. An input-output approach for modeling of the analog systems can be employed. It treats them as **black boxes**. In the linear passive case, this leads to identification problems using

multi-port **S-parameters**

Model reduction from data:

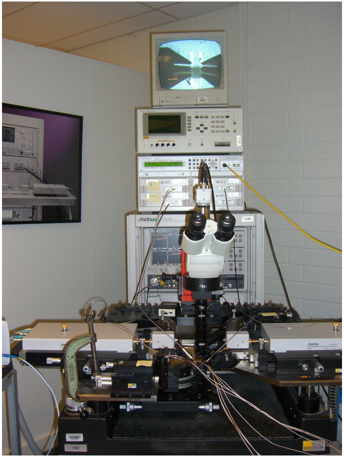
On-chip analog electronics

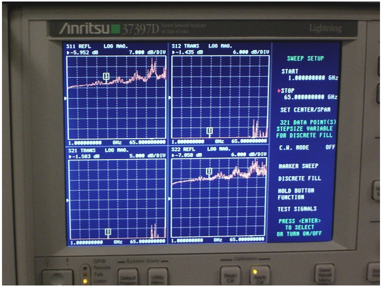
Chips for communication systems consist of large analog and RF blocks. To avoid costly re-fabrication, a verification cycle is developed for simulation and design optimization. A common approach to this verification is to replace the circuit block layout by systems of equations and subsequently use their accurate approximants for system simulation. Example: FPGA (Field Programmable Gate Arrays).

Methodology. An input-output approach for modeling of the analog systems can be employed. It treats them as **black boxes**. In the linear passive case, this leads to identification problems using

multi-port **S-parameters**

Measurement of S-parameters





VNA (Vector Network Analyzer) - Magnitude of S-parameters for 2 ports

Analysis of model reduction from S-parameters

Tangential interpolation

Given:
- right data: $(\lambda_i; \mathbf{r}_i, \mathbf{w}_i)$, $i = 1, \cdots, k$
- left data: $(\mu_j; \ell_j, \mathbf{v}_j)$, $j = 1, \cdots, q$.

We assume for simplicity that all points are distinct.

Problem: Find rational $p \times m$ matrices $\mathbf{H}(s)$, such that

$$\mathbf{H}(\lambda_i)\mathbf{r}_i = \mathbf{w}_i \qquad \ell_j \mathbf{H}(\mu_j) = \mathbf{v}_j$$

Right data:

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \in \mathbb{C}^{k\times k}, \quad \mathbf{R} = [\mathbf{r}_1 \;\; \mathbf{r}_2, \;\; \cdots \;\; \mathbf{r}_k] \in \mathbb{C}^{m\times k}, \quad \mathbf{W} = [\mathbf{w}_1 \;\; \mathbf{w}_2 \;\; \cdots \;\; \mathbf{w}_k] \in \mathbb{C}^{p\times k}$$

Left data:

$$M = \begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_q \end{bmatrix} \in \mathbb{C}^{q\times q}, \; \mathbf{L} = \begin{bmatrix} \ell_1 \\ \vdots \\ \ell_q \end{bmatrix} \in \mathbb{C}^{q\times p}, \; \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_q \end{bmatrix} \in \mathbb{C}^{q\times m}$$

Given:
- right data: $(\lambda_i; \mathbf{r}_i, \mathbf{w}_i)$, $i = 1, \cdots, k$
- left data: $(\mu_j; \ell_j, \mathbf{v}_j)$, $j = 1, \cdots, q$.

We assume for simplicity that all points are distinct.

Problem: Find rational $p \times m$ matrices $\mathbf{H}(s)$, such that

$$\mathbf{H}(\lambda_i)\mathbf{r}_i = \mathbf{w}_i \qquad \ell_j \mathbf{H}(\mu_j) = \mathbf{v}_j$$

Right data:

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \in \mathbb{C}^{k\times k}, \quad \mathbf{R} = [\mathbf{r}_1 \ \ \mathbf{r}_2, \ \cdots \ \ \mathbf{r}_k] \in \mathbb{C}^{m\times k}, \quad \mathbf{W} = [\mathbf{w}_1 \ \ \mathbf{w}_2 \ \ \cdots \ \ \mathbf{w}_k] \in \mathbb{C}^{p\times k}$$

Left data:

$$M = \begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_q \end{bmatrix} \in \mathbb{C}^{q\times q}, \quad \mathbf{L} = \begin{bmatrix} \ell_1 \\ \vdots \\ \ell_q \end{bmatrix} \in \mathbb{C}^{q\times p}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_q \end{bmatrix} \in \mathbb{C}^{q\times m}$$

Given: • right data: $(\lambda_i; \mathbf{r}_i, \mathbf{w}_i)$, $i = 1, \cdots, k$
• left data: $(\mu_j; \ell_j, \mathbf{v}_j)$, $j = 1, \cdots, q$.

We assume for simplicity that all points are distinct.

Problem: Find rational $p \times m$ matrices $\mathbf{H}(s)$, such that

$$\mathbf{H}(\lambda_i)\mathbf{r}_i = \mathbf{w}_i \qquad \ell_j \mathbf{H}(\mu_j) = \mathbf{v}_j$$

Right data:

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \in \mathbb{C}^{k \times k}, \quad \mathbf{R} = [\mathbf{r}_1 \;\; \mathbf{r}_2, \;\; \cdots \;\; \mathbf{r}_k] \in \mathbb{C}^{m \times k}, \quad \mathbf{W} = [\mathbf{w}_1 \;\; \mathbf{w}_2 \;\; \cdots \;\; \mathbf{w}_k] \in \mathbb{C}^{p \times k}$$

Left data:

$$M = \begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_q \end{bmatrix} \in \mathbb{C}^{q \times q}, \quad \mathbf{L} = \begin{bmatrix} \ell_1 \\ \vdots \\ \ell_q \end{bmatrix} \in \mathbb{C}^{q \times p}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_q \end{bmatrix} \in \mathbb{C}^{q \times m}$$

Given: • right data: $(\lambda_i; \mathbf{r}_i, \mathbf{w}_i)$, $i = 1, \cdots, k$

• left data: $(\mu_j; \ell_j, \mathbf{v}_j)$, $j = 1, \cdots, q$.

We assume for simplicity that all points are distinct.

Problem: Find rational $p \times m$ matrices $\mathbf{H}(s)$, such that

$$\mathbf{H}(\lambda_i)\mathbf{r}_i = \mathbf{w}_i \qquad \ell_j \mathbf{H}(\mu_j) = \mathbf{v}_j$$

Right data:

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \in \mathbb{C}^{k\times k}, \quad \begin{array}{l} \mathbf{R} = [\mathbf{r}_1 \;\; \mathbf{r}_2, \;\; \cdots \;\; \mathbf{r}_k] \in \mathbb{C}^{m\times k}, \\ \mathbf{W} = [\mathbf{w}_1 \;\; \mathbf{w}_2 \;\; \cdots \;\; \mathbf{w}_k] \in \mathbb{C}^{p\times k} \end{array}$$

Left data:

$$M = \begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_q \end{bmatrix} \in \mathbb{C}^{q\times q}, \; \mathbf{L} = \begin{bmatrix} \ell_1 \\ \vdots \\ \ell_q \end{bmatrix} \in \mathbb{C}^{q\times p}, \; \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_q \end{bmatrix} \in \mathbb{C}^{q\times m}$$

The Loewner and the shifted Loewner matrices

We define the **Loewner matrix**

$$\mathbb{L} = \begin{bmatrix} \frac{\mathbf{v}_1\mathbf{r}_1 - \ell_1\mathbf{w}_1}{\lambda_1-\mu_1} & \cdots & \frac{\mathbf{v}_1\mathbf{r}_k - \ell_1\mathbf{w}_k}{\lambda_1-\mu_k} \\ \vdots & \ddots & \vdots \\ \frac{\mathbf{v}_q\mathbf{r}_1 - \ell_q\mathbf{w}_1}{\lambda_q-\mu_1} & \cdots & \frac{\mathbf{v}_q\mathbf{r}_k - \ell_q\mathbf{w}_k}{\lambda_q-\mu_k} \end{bmatrix} \in \mathbb{C}^{q\times k}$$

and the **shifted Loewner matrix**

$$\sigma\mathbb{L} = \begin{bmatrix} \frac{\lambda_1\mathbf{v}_1\mathbf{r}_1 - \ell_1\mathbf{w}_1\mu_1}{\lambda_1-\mu_1} & \cdots & \frac{\lambda_1\mathbf{v}_1\mathbf{r}_k - \ell_1\mathbf{w}_k\mu_k}{\lambda_1-\mu_k} \\ \vdots & \ddots & \vdots \\ \frac{\lambda_q\mathbf{v}_q\mathbf{r}_1 - \ell_q\mathbf{w}_1\mu_1}{\lambda_q-\mu_1} & \cdots & \frac{\lambda_q\mathbf{v}_q\mathbf{r}_k - \ell_q\mathbf{w}_k\mu_k}{\lambda_q-\mu_k} \end{bmatrix} \in \mathbb{C}^{q\times k}$$

Remark. For a single interpolation point the Loewner and shifted Loewner matrices reduce to Hankel matrices.

Construction of Interpolants (Models)

Assume that $k = \ell$, and let

$$\det(x\mathbb{L} - \sigma\mathbb{L}) \neq 0, \quad x \in \{\lambda_i\} \cup \{\mu_j\}$$

Then

$$\mathbf{E} = -\mathbb{L}, \quad \mathbf{A} = -\sigma\mathbb{L}, \quad \mathbf{B} = \mathbf{V}, \quad \mathbf{C} = \mathbf{W}$$

is a minimal realization of an interpolant of the data, i.e., the function

$$\mathbf{H}(s) = \mathbf{W}(\sigma\mathbb{L} - s\mathbb{L})^{-1}\mathbf{V}$$

interpolates the data.

Construction of Interpolants (Models)

Assume that $k = \ell$, and let

$$\det(x\mathbb{L} - \sigma\mathbb{L}) \neq 0, \quad x \in \{\lambda_i\} \cup \{\mu_j\}$$

Then

$$\mathbf{E} = -\mathbb{L}, \quad \mathbf{A} = -\sigma\mathbb{L}, \quad \mathbf{B} = \mathbf{V}, \quad \mathbf{C} = \mathbf{W}$$

is a minimal realization of an interpolant of the data, i.e., the function

$$\mathbf{H}(s) = \mathbf{W}(\sigma\mathbb{L} - s\mathbb{L})^{-1}\mathbf{V}$$

interpolates the data.

Construction of Interpolants (Models)

Assume that $k = \ell$, and let

$$\det(x\mathbb{L} - \sigma\mathbb{L}) \neq 0, \quad x \in \{\lambda_i\} \cup \{\mu_j\}$$

Then

$$\mathbf{E} = -\mathbb{L}, \ \ \mathbf{A} = -\sigma\mathbb{L}, \ \ \mathbf{B} = \mathbf{V}, \ \ \mathbf{C} = \mathbf{W}$$

is a minimal realization of an interpolant of the data, i.e., the function

$$\mathbf{H}(s) = \mathbf{W}(\sigma\mathbb{L} - s\mathbb{L})^{-1}\mathbf{V}$$

interpolates the data.

Construction of interpolants: New procedure

Main assumption:

$$\text{rank}\,(x\mathbb{L} - \sigma\mathbb{L}) = \text{rank}\begin{pmatrix} \mathbb{L} & \sigma\mathbb{L} \end{pmatrix} = \text{rank}\begin{pmatrix} \mathbb{L} \\ \sigma\mathbb{L} \end{pmatrix} =: k,\ x \in \{\lambda_i\} \cup \{\mu_j\}$$

Then for some $x \in \{\lambda_i\} \cup \{\mu_j\}$, we compute the *SVD*

$$x\mathbb{L} - \sigma\mathbb{L} = \mathbf{Y\Sigma X}$$

with $\text{rank}\,(x\mathbb{L} - \sigma\mathbb{L}) = \text{rank}\,(\Sigma) = \text{size}\,(\Sigma) =: k$, $\mathbf{Y} \in \mathbb{C}^{\nu \times k}$, $\mathbf{X} \in \mathbb{C}^{k \times \rho}$.

Theorem. A realization $[\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{C}]$, of an interpolant is given as follows:

$\mathbf{E} = -\mathbf{Y}^*\mathbb{L}\mathbf{X}^*$	$\mathbf{B} = \mathbf{Y}^*\mathbf{V}$
$\mathbf{A} = -\mathbf{Y}^*\sigma\mathbb{L}\mathbf{X}^*$	$\mathbf{C} = \mathbf{W}\mathbf{X}^*$

Remark. The singular values of $x\mathbb{L} - \sigma\mathbb{L}$ play a role similar to the that of the Hankel singular values.

Construction of interpolants: **New procedure**

Main assumption:

$$\operatorname{rank}(x\mathbb{L} - \sigma\mathbb{L}) = \operatorname{rank}\begin{pmatrix} \mathbb{L} & \sigma\mathbb{L} \end{pmatrix} = \operatorname{rank}\begin{pmatrix} \mathbb{L} \\ \sigma\mathbb{L} \end{pmatrix} =: k, \;\; x \in \{\lambda_i\} \cup \{\mu_j\}$$

Then for some $x \in \{\lambda_i\} \cup \{\mu_j\}$, we compute the *SVD*

$$x\mathbb{L} - \sigma\mathbb{L} = \mathbf{Y}\Sigma\mathbf{X}$$

with $\operatorname{rank}(x\mathbb{L} - \sigma\mathbb{L}) = \operatorname{rank}(\Sigma) = \operatorname{size}(\Sigma) =: k$, $\mathbf{Y} \in \mathbb{C}^{\nu \times k}$, $\mathbf{X} \in \mathbb{C}^{k \times \rho}$.

Theorem. A realization [**E**, **A**, **B**, **C**], of an interpolant is given as follows:

$\mathbf{E} = -\mathbf{Y}^*\mathbb{L}\mathbf{X}^*$	$\mathbf{B} = \mathbf{Y}^*\mathbf{V}$
$\mathbf{A} = -\mathbf{Y}^*\sigma\mathbb{L}\mathbf{X}^*$	$\mathbf{C} = \mathbf{W}\mathbf{X}^*$

Remark. The singular values of $x\mathbb{L} - \sigma\mathbb{L}$ play a role similar to the that of the Hankel singular values.

Construction of interpolants: New procedure

Main assumption:

$$\operatorname{rank}(x\mathbb{L} - \sigma\mathbb{L}) = \operatorname{rank}\begin{pmatrix} \mathbb{L} & \sigma\mathbb{L} \end{pmatrix} = \operatorname{rank}\begin{pmatrix} \mathbb{L} \\ \sigma\mathbb{L} \end{pmatrix} =: k, \;\; x \in \{\lambda_i\} \cup \{\mu_j\}$$

Then for some $x \in \{\lambda_i\} \cup \{\mu_j\}$, we compute the *SVD*

$$x\mathbb{L} - \sigma\mathbb{L} = \mathbf{Y}\Sigma\mathbf{X}$$

with $\operatorname{rank}(x\mathbb{L} - \sigma\mathbb{L}) = \operatorname{rank}(\Sigma) = \operatorname{size}(\Sigma) =: k$, $\mathbf{Y} \in \mathbb{C}^{\nu \times k}$, $\mathbf{X} \in \mathbb{C}^{k \times \rho}$.

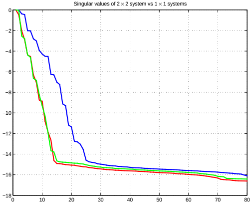
Theorem. A realization $[\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{C}]$, of an interpolant is given as follows:

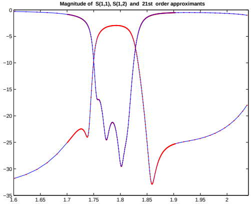
$\mathbf{E} = -\mathbf{Y}^*\mathbb{L}\mathbf{X}^*$	$\mathbf{B} = \mathbf{Y}^*\mathbf{V}$
$\mathbf{A} = -\mathbf{Y}^*\sigma\mathbb{L}\mathbf{X}^*$	$\mathbf{C} = \mathbf{W}\mathbf{X}^*$

Remark. The singular values of $x\mathbb{L} - \sigma\mathbb{L}$ play a role similar to the that of the Hankel singular values.

Example: Four-pole band-pass filter

• **1000 measurements between 40 and 120 GHz; S-parameters** 2×2, **MIMO interpolation** $\Rightarrow$ $\mathbb{L}, \sigma\mathbb{L} \in \mathbb{R}^{2000\times 2000}$.





The singular values of $\mathbb{L}$, $\sigma\mathbb{L}$	The $S(1,1)$ and $S(1,2)$ parameter data 17^{th}-order approximant

Summary: Advantages of this method

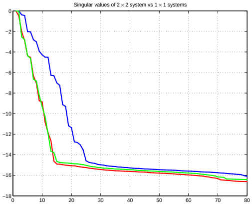
(1): No need to invert **E**.

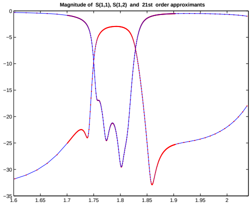
(2): Rank (sing. vals) of $x\mathbb{L} - \sigma\mathbb{L}$ provides the model complexity.

(3): Can handle large-number of inputs/outputs by means of tangential interpolation.

Example: Four-pole band-pass filter

•**1000 measurements between 40 and 120 GHz; S-parameters** 2×2, **MIMO interpolation** $\Rightarrow$ $\mathbb{L}, \sigma\mathbb{L} \in \mathbb{R}^{2000 \times 2000}$.





The singular values of $\mathbb{L}$, $\sigma\mathbb{L}$	The $S(1,1)$ and $S(1,2)$ parameter data 17^{th}-order approximant

Summary: Advantages of this method

(1): No need to invert **E**.

(2): Rank (sing. vals) of $x\mathbb{L} - \sigma\mathbb{L}$ provides the model complexity.

(3): Can handle large-number of inputs/outputs by means of tangential interpolation.

Part V

Challenges in complexity reduction

(Some) Challenges in complexity reduction

- Model reduction of uncertain systems
- Model reduction of differential-algebraic (DAE) systems
- Domain decomposition methods
- Parallel algorithms for sparse computations in model reduction
- Development/validation of control algorithms based on reduced models
- Model reduction and data assimilation (weather prediction)
- Active control of high-rise buildings
- MEMS and multi-physics problems
- VLSI design
- Molecular Dynamics (MD) simulations
- **Nanoelectronics**

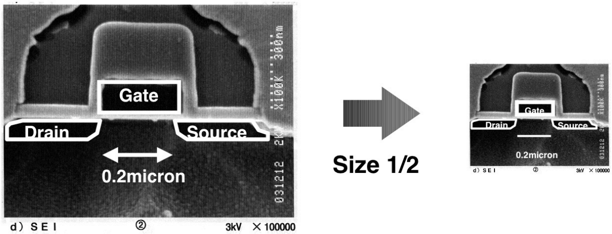
(Some) Challenges in complexity reduction

- Model reduction of uncertain systems
- Model reduction of differential-algebraic (DAE) systems
- Domain decomposition methods
- Parallel algorithms for sparse computations in model reduction
- Development/validation of control algorithms based on reduced models
- Model reduction and data assimilation (weather prediction)
- Active control of high-rise buildings
- MEMS and multi-physics problems
- VLSI design
- Molecular Dynamics (MD) simulations
- **Nanoelectronics**

(Some) Challenges in complexity reduction

- Model reduction of uncertain systems
- Model reduction of differential-algebraic (DAE) systems
- Domain decomposition methods
- Parallel algorithms for sparse computations in model reduction
- Development/validation of control algorithms based on reduced models
- Model reduction and data assimilation (weather prediction)
- Active control of high-rise buildings
- MEMS and multi-physics problems
- VLSI design
- Molecular Dynamics (MD) simulations
- **Nanoelectronics**

Scaling Law



Favorable effects

Size	x1/2
Voltage	x1/2
Electric Field	x1
Speed	x3
Cost	x1/4

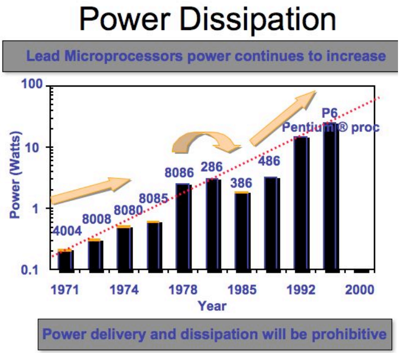
Unfavorable effects

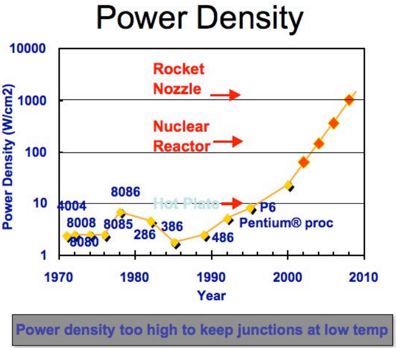
Power density	x1.6
RC delay/Tr. delay	x3.2
Current density	x1.6
Voltage noise	x3.2
Design complexity	x4

Future challenge: Nanoelectronics

Heat generation

Kitchen stove:	18cm diameter,	P ≈ 1.5kW	$\Rightarrow$	$6W/cm^2$
Pentium IV:	Area $\approx 2cm^2$,	P ≈ 88W	$\Rightarrow$	$40W/cm^2$



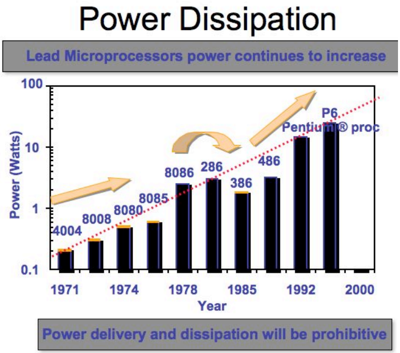


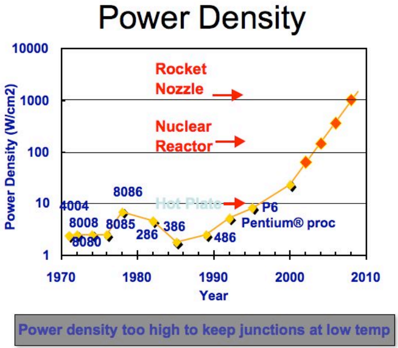
Conclusion: According to the 2006 ITRS, at the present rate of miniaturization, the current technology can be sustained for a few more years (until the feature size reaches 45*nm*).

Future challenge: Nanoelectronics

Heat generation

Kitchen stove:	18cm diameter,	P ≈ 1.5kW	⇒ $6W/cm^2$
Pentium IV:	Area ≈ $2cm^2$,	P ≈ 88W	⇒ $40W/cm^2$





Conclusion: According to the 2006 ITRS, at the present rate of miniaturization, the current technology can be sustained for a few more years (until the feature size reaches $45nm$).

Future challenge: Nanoelectronics

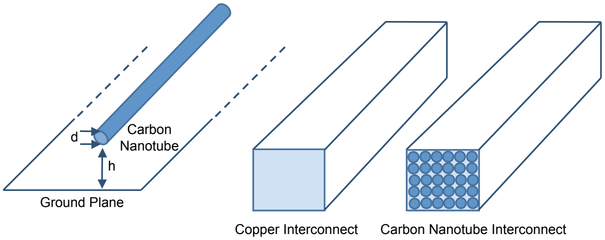
Proposed interconnect solution: carbon nanotubes

- CNTs have been proposed as a replacement for on-chip copper interconnects due to their large conductivity and current carrying capabilities.
- Advantages over copper:

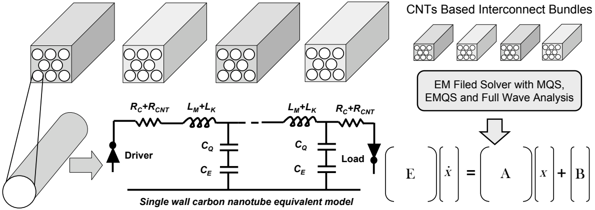
1. **Resistance.** CNTs have lower resistance than standard copper
2. **Current density.** Single-wall Carbon Nanotubes (SWCNTs) with diameters ranging from 0.4nm to 4nm have been reported, with current densities as large as $\mathbf{10^{10} A/cm^2}$, versus traditional metallic interconnect with typical current densities on the order of $\mathbf{10^5 A/cm^2}$.
3. **Electromigration.** CNTs are much less susceptible to electromigration problems with thermal conductivity more than 10 times higher than conventional copper.

Future challenge: Nanoelectronics

Carbon nanotubes (CNTs): modeling



Analytical model of SWCNT: transmission line involving magnetic and kinetic inductance, as well as electrostatic and quantum capacitance.



$$\begin{bmatrix} E \end{bmatrix} \begin{bmatrix} \dot{x} \end{bmatrix} = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x \end{bmatrix} + \begin{bmatrix} B \end{bmatrix} u$$

Future challenge: Nanoelectronics

Some mathematical challenges

- **CNTs**: Develop a scalable state space representation of carbon nanotube circuit models that accurately capture the statistical distribution of single as well as carbon nanotube bundles.

- **CNTs**: Develop model reduction techniques to solve and accurately approximate CNT based interconnects resulting from field solvers. Evaluate the complexity of these methods used for CNT based interconnects and conventional copper interconnects for their suitability in fast simulation.

References

- Passivity preserving model reduction
  - Antoulas SCL (2005)
  - Sorensen SCL (2005)
  - Ionutiu, Rommes, Antoulas IEEE CAD (2008)
- Optimal $\mathcal{H}_2$ model reduction
  - Gugercin, Antoulas, Beattie SIMAX (2008)
- Low-rank solutions of Lyapunov equations
  - Gugercin, Sorensen, Antoulas, Numerical Algorithms (2003)
  - Sorensen (2006)
- Model reduction from data
  - Mayo, Antoulas LAA (2007)
  - Lefteriu, Antoulas, Tech. Report (2007)
- General reference: Antoulas, SIAM (2005)