Order Reduction of (Truly) Large-Scale Linear Dynamical Systems

Roland W. Freund

Department of Mathematics
University of California, Davis, USA
http://www.math.ucdavis.edu/~freund/

Supported in part by NSF

Motivation

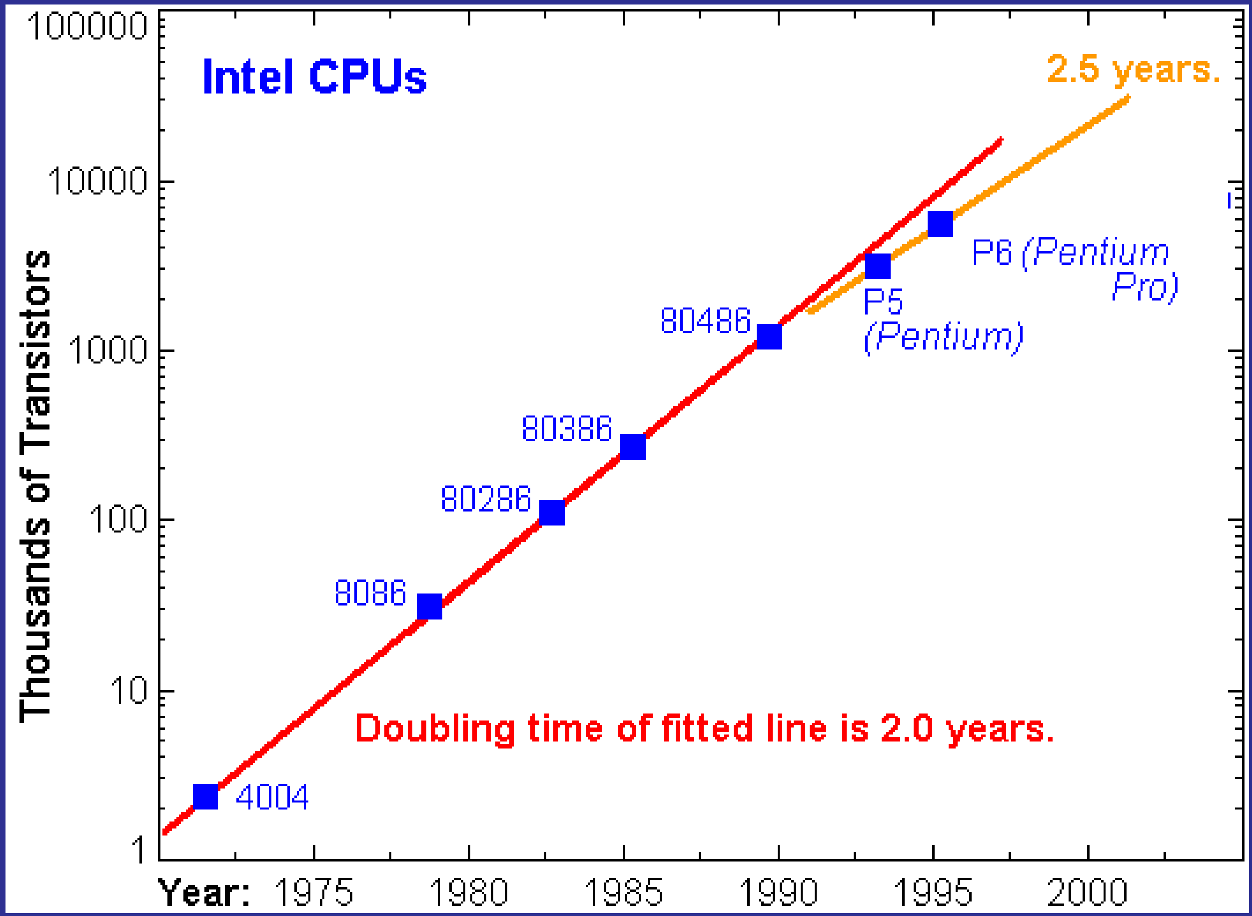
- Need for order reduction in VLSI circuit simulation

- Corollary to Moore's Law

- **RCL** networks:
  Electric networks consisting of only resistors (**R**'s), capacitors (**C**'s), and inductors (**L**'s)

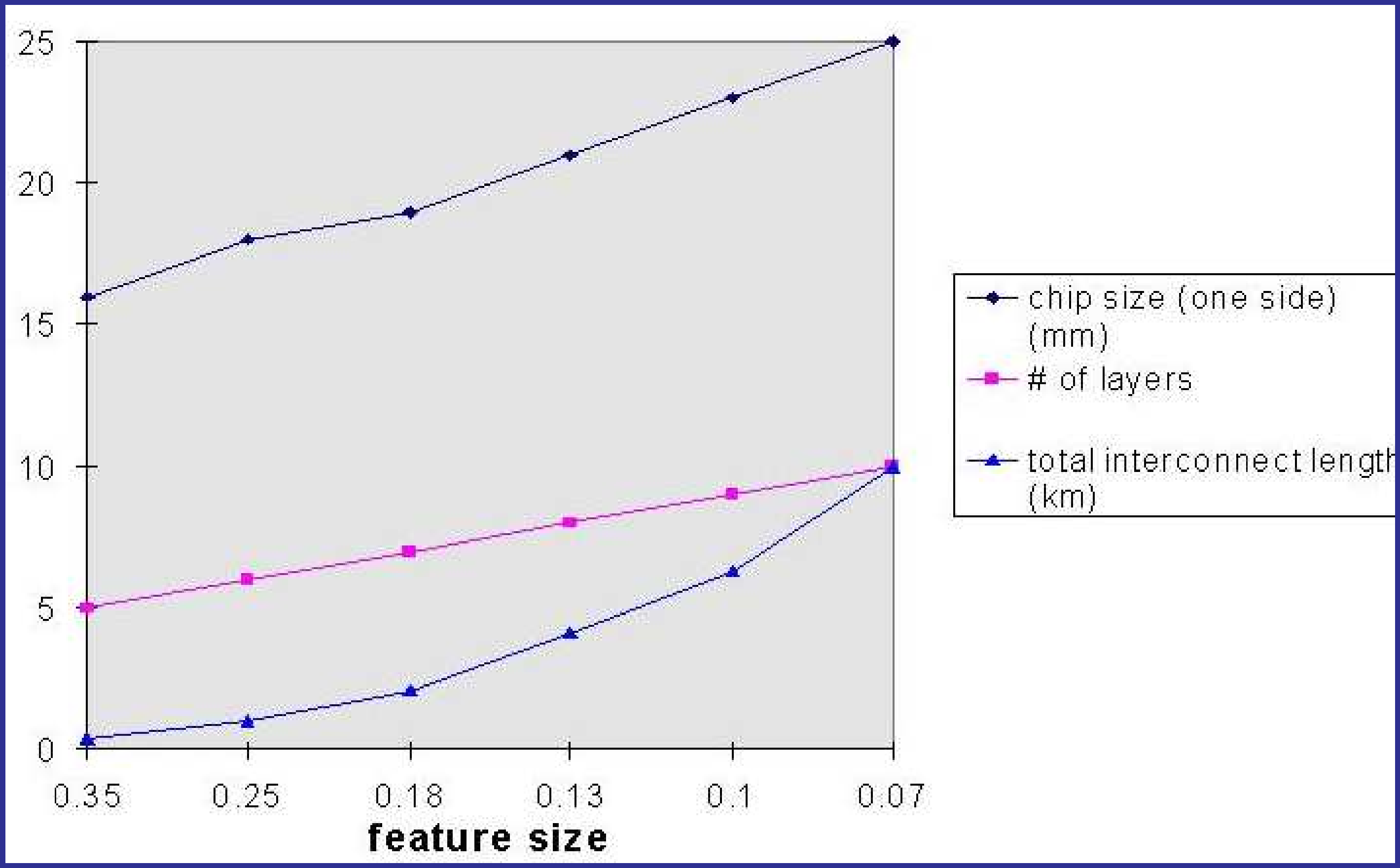
- These networks are (truly) large

Moore's law



Intel CPUs
2.5 years.
Thousands of Transistors
100000
10000
1000
100
10
1
4004
8086
80286
80386
80486
P5
(Pentium)
P6 (Pentium
Pro)
Doubling time of fitted line is 2.0 years.
Year: 1975 1980 1985 1990 1995 2000

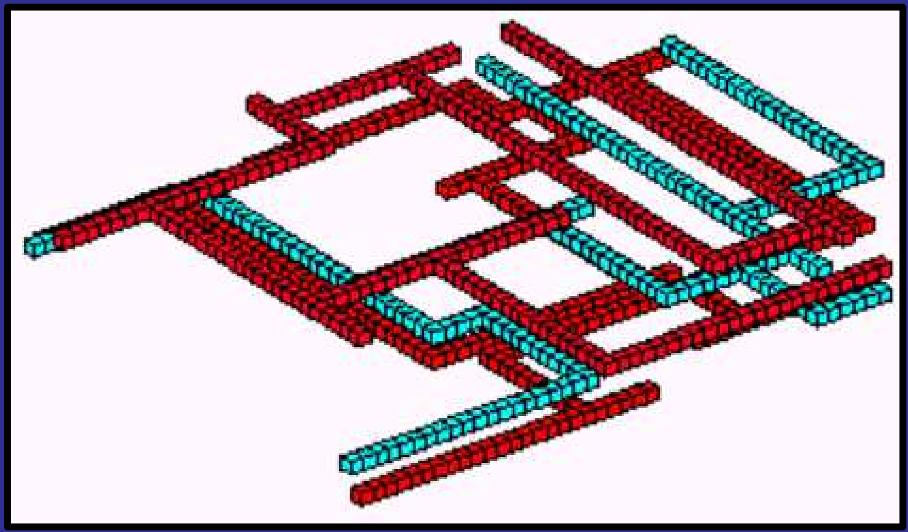
VLSI chip scaling



VLSI interconnect

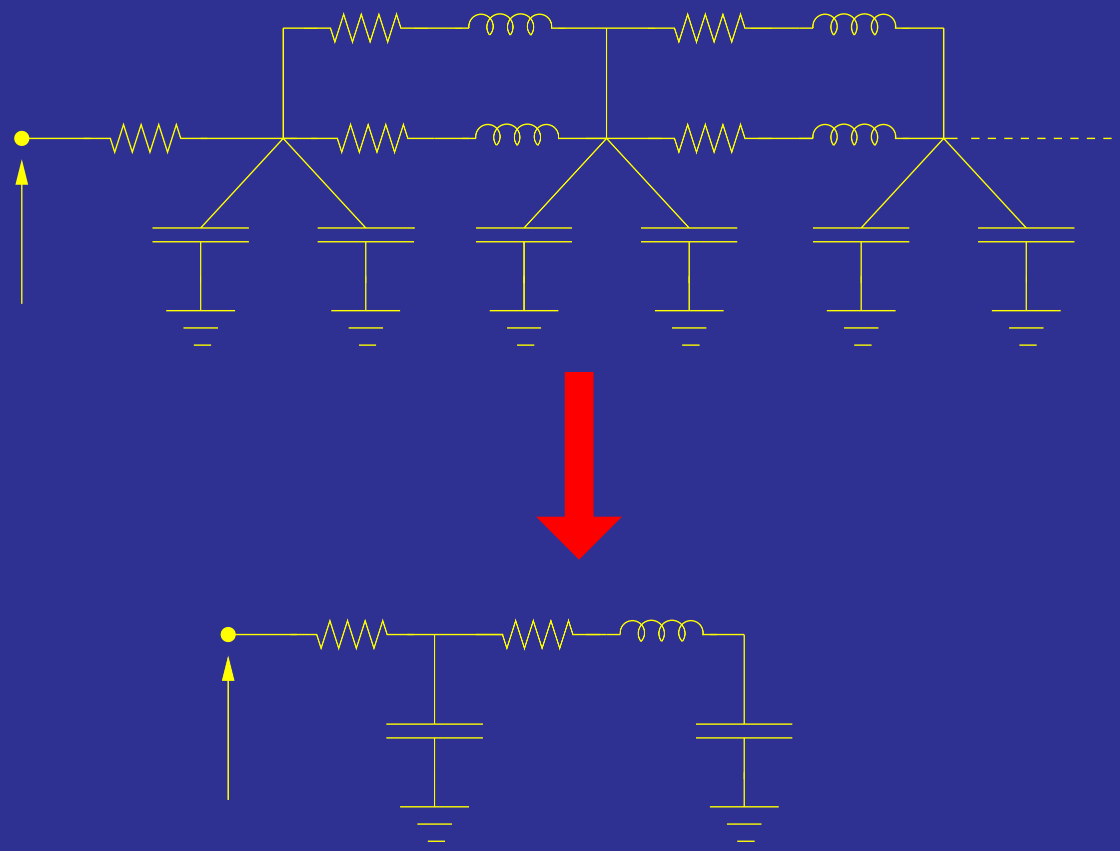
- Wires are not ideal:
  - Resistance
  - Capacitance
  - Inductance
- Consequences:
  - Timing behavior
  - Noise
  - Energy consumption
  - Power distribution

Lumped-circuit paradigm



- Replace ‘pieces’ of the interconnect by RCL networks
- Up to $\mathcal{O}(10^6)$ circuit elements per network
- Up to $\mathcal{O}(10^6)$ networks

Need for order reduction



Outline

- The order reduction problem
- Projection + Krylov = Padé-type reduction
- SPRIM for general RCL networks
- SPRIM–SVD
- Padé-type approximation properties of SPRIM
- Concluding remarks

Outline

- *The order reduction problem*
- Projection + Krylov = Padé-type reduction
- SPRIM for general RCL networks
- SPRIM–SVD
- Padé-type approximation properties of SPRIM
- Concluding remarks

RCL networks as descriptor systems

- System of linear time-invariant DAEs of the form

$$\begin{aligned} \mathbf{C}\,\frac{d}{dt}\mathbf{x}(t) + \mathbf{G}\,\mathbf{x}(t) &= \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{B}^{\top}\mathbf{x}(t) \end{aligned}$$

  where $\mathbf{C},\ \mathbf{G} \in \mathbb{R}^{N \times N}$ and $\mathbf{B} \in \mathbb{R}^{N \times m}$

- $\mathbf{x}(t) \in \mathbb{R}^N$ is the unknown vector of state variables

- m inputs, m outputs

- $s\,\mathbf{C} + \mathbf{G}$ is nonsingular except for finitely many values of $s \in \mathbb{C}$

Reduced-order models

- System of DAEs of the same form:

$$\mathbf{C}_n \frac{d}{dt}\mathbf{z}(t) + \mathbf{G}_n\,\mathbf{z}(t) = \mathbf{B}_n\,\mathbf{u}(t)$$
$$\tilde{\mathbf{y}}(t) = \mathbf{B}_n^{\top}\,\mathbf{z}(t)$$

- But now:

$$\mathbf{C}_n,\ \mathbf{G}_n \in \mathbb{R}^{n\times n} \quad \text{and} \quad \mathbf{B}_n \in \mathbb{R}^{n\times m}$$

where $n \ll N$

Transfer functions

- Original descriptor system:

$$\mathbf{H}(s) = \mathbf{B}^{\top} (s\,\mathbf{C} + \mathbf{G})^{-1}\,\mathbf{B}$$

- Reduced-order model:

$$\mathbf{H}_n(s) = \mathbf{B}_n^{\top} (s\,\mathbf{C}_n + \mathbf{G}_n)^{-1}\,\mathbf{B}_n$$

- ‘Good’ reduced-order model

$\Longleftrightarrow$ ‘Good’ approximation $\mathbf{H}_n \approx \mathbf{H}$

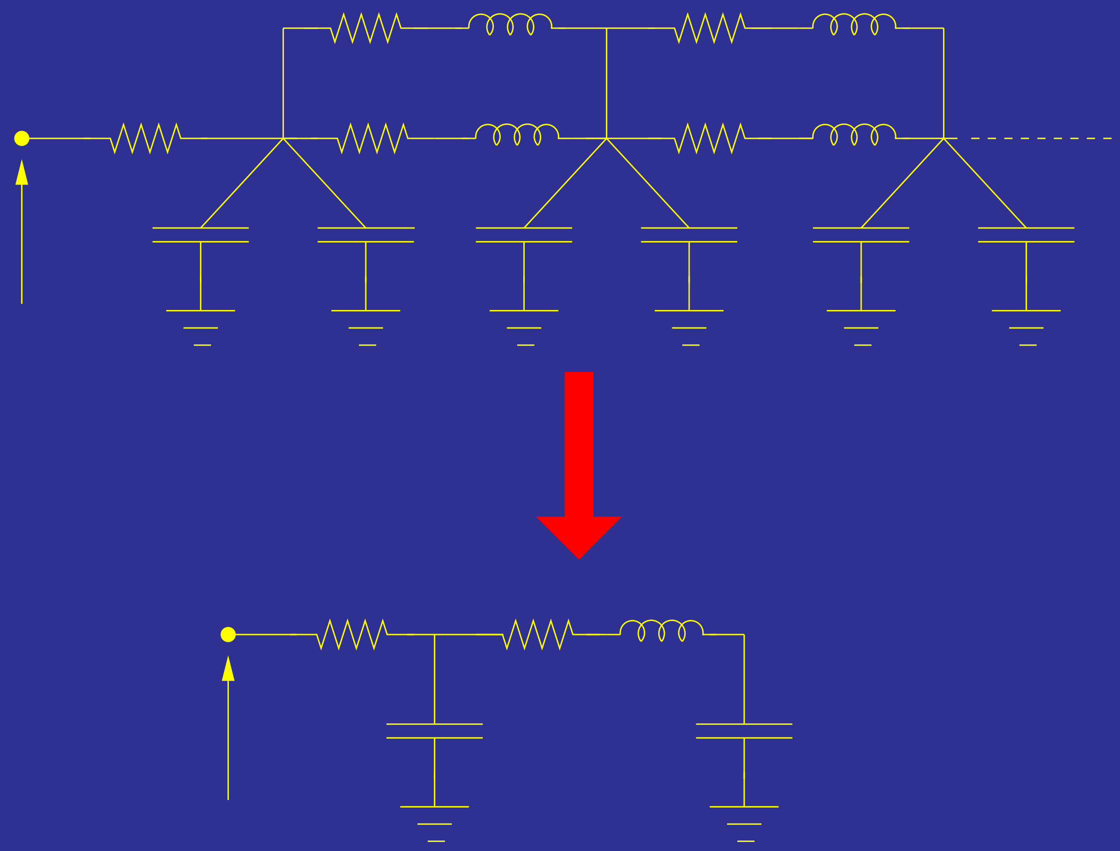
Problem of structure preservation

- Any RCL network is stable, passive, . . .

- Reduced-order model should be stable, passive, . . .

- More difficult problem:
  Reduced-order model of an RCL network should be synthesizable as an RCL network

Preservation of RCL structure



General RCL network equations

- System of linear time-invariant DAEs of the form

$$\mathbf{C}\frac{d}{dt}\mathbf{x}(t) + \mathbf{G}\,\mathbf{x}(t) = \mathbf{B}\,\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{B}^{\mathsf{T}}\mathbf{x}(t)$$

where

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \mathbf{G}_3 \\ -\mathbf{G}_2^{\mathsf{T}} & \mathbf{0} & \mathbf{0} \\ -\mathbf{G}_3^{\mathsf{T}} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 \end{bmatrix}$$

- Moreover:

$$\mathbf{C} \succeq \mathbf{0} \quad \text{and} \quad \mathbf{G} + \mathbf{G}^{\mathsf{T}} \succeq \mathbf{0}$$

(This implies passivity!)

Outline

- The order reduction problem
- *Projection + Krylov = Padé-type reduction*
- SPRIM for general RCL networks
- SPRIM–SVD
- Padé-type approximation properties of SPRIM
- Concluding remarks

Projection-based reduction

- Let $\mathbf{V}_n \in \mathbb{R}^{N \times n}$ be any matrix with full column rank n

- Use $\mathbf{V}_n$ to explicitly project the data matrices of

$$\begin{aligned} \mathbf{C}\,\frac{d}{dt}\mathbf{x}(t) + \mathbf{G}\,\mathbf{x}(t) &= \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{B}^{\top}\mathbf{x}(t) \end{aligned}$$

  onto the subspace spanned by the columns of $\mathbf{V}_n$

Projection-based reduction, continued

- Resulting reduced-order model

$$\mathbf{C}_n \frac{d}{dt}\mathbf{z}(t) + \mathbf{G}_n\,\mathbf{z}(t) = \mathbf{B}_n\,\mathbf{u}(t)$$
$$\tilde{\mathbf{y}}(t) = \mathbf{B}_n^\top \mathbf{z}(t)$$

  where

$$\mathbf{C}_n = \mathbf{V}_n^\top\,\mathbf{C}\,\mathbf{V}_n, \quad \mathbf{G}_n = \mathbf{V}_n^\top\,\mathbf{G}\,\mathbf{V}_n, \quad \mathbf{B}_n = \mathbf{V}_n^\top\,\mathbf{B}$$

- Passivity is preserved:

$$\mathbf{C} \succeq \mathbf{0},\;\; \mathbf{G} + \mathbf{G}^\top \succeq \mathbf{0} \quad \Rightarrow \quad \mathbf{C}_n \succeq \mathbf{0},\;\; \mathbf{G}_n + \mathbf{G}_n^\top \succeq \mathbf{0}$$

Projection-based order reduction

- **PRIMA**
  Passive Reduced Interconnect Macromodeling Algorithm
  (Odabasioglu, '96; Odabasioglu, Celik, and Pileggi, '97)

- Split-congruence transformations
  (Kerns, Yang, '97)

- **SPRIM**
  Structure-Preserving Reduced Interconnect Macromodeling
  (F., '04 and '07)

PRIMA reduced-order models

- Let $\mathbf{V}_n$ be any matrix whose columns span the n-th **Krylov subspace** $\mathcal{K}_n(\mathbf{A}, \mathbf{R})$ where

$$\mathbf{A} := \left(s_0\,\mathbf{C} + \mathbf{G}\right)^{-1}\mathbf{C} \quad \text{and} \quad \mathbf{R} := \left(s_0\,\mathbf{C} + \mathbf{G}\right)^{-1}\mathbf{B}$$

  and $s_0 \in \mathbb{R}$ is a suitably chosen expansion point

- Projection + Krylov subspace = **Padé-type approximant**:

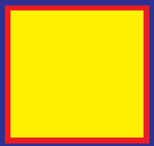
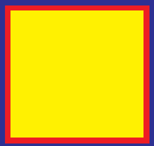
$$\mathbf{H}_n(s) = \mathbf{H}(s) + \mathcal{O}\left((s - s_0)^q\right), \quad \text{where} \quad q \geq \lfloor n/m \rfloor$$

Structure is not preserved

- Structure of the data matrices:

$$C = \begin{bmatrix} C_1 & 0 & 0 \\ 0 & C_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} G_1 & G_2 & G_3 \\ -G_2^T & 0 & 0 \\ -G_3^T & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \\ 0 & B_2 \end{bmatrix}$$

- Structure of PRIMA reduced-order matrices:

$C_n =$  , $G_n =$  , $B_n =$ 

Outline

- The order reduction problem
- Projection + Krylov = Padé-type reduction
- *SPRIM for general RCL networks*
- SPRIM–SVD
- Padé-type approximation properties of SPRIM
- Concluding remarks

SPRIM

- As in PRIMA, let $\mathbf{V}_n$ be any matrix such that

$$\mathcal{K}_n(\mathbf{A}, \mathbf{R}) = \text{colspan}\, \mathbf{V}_n$$

- Key insight that is exploited in SPRIM:
  In order to have a Padé-type property as in PRIMA, we can project with any matrix $\tilde{\mathbf{V}}_n$ such that

$$\mathcal{K}_n(\mathbf{A}, \mathbf{R}) \subseteq \text{colspan}\, \tilde{\mathbf{V}}_n$$

- ... ; Odabasioglu, '96; Grimme, '97; Odabasioglu, Celik, and Pileggi, '97; ...

SPRIM, continued

- Recall:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \mathbf{G}_3 \\ -\mathbf{G}_2^\top & \mathbf{0} & \mathbf{0} \\ -\mathbf{G}_3^\top & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 \end{bmatrix}$$

- Partition $\mathbf{V}_n$ accordingly:

$$\mathbf{V}_n = \begin{bmatrix} \mathbf{V}_n^{(1)} \\ \mathbf{V}_n^{(2)} \\ \mathbf{V}_n^{(3)} \end{bmatrix}$$

SPRIM, continued

- Set

$$\tilde{\mathbf{V}}_n = \begin{bmatrix} \mathbf{V}_n^{(1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_n^{(2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_n^{(3)} \end{bmatrix}$$

- Then: $\mathcal{K}_n(\mathbf{A}, \mathbf{R}) = \text{colspan}\, \mathbf{V}_n \subseteq \text{colspan}\, \tilde{\mathbf{V}}_n$

- This guarantees a Padé-type property!

SPRIM models

- Recall:

$$C = \begin{bmatrix} C_1 & 0 & 0 \\ 0 & C_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} G_1 & G_2 & G_3 \\ -G_2^T & 0 & 0 \\ -G_3^T & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \\ 0 & B_2 \end{bmatrix}$$

and

$$\tilde{V}_n = \begin{bmatrix} V_n^{(1)} & 0 & 0 \\ 0 & V_n^{(2)} & 0 \\ 0 & 0 & V_n^{(3)} \end{bmatrix}$$

SPRIM models, continued

- The projection now preserves this structure:

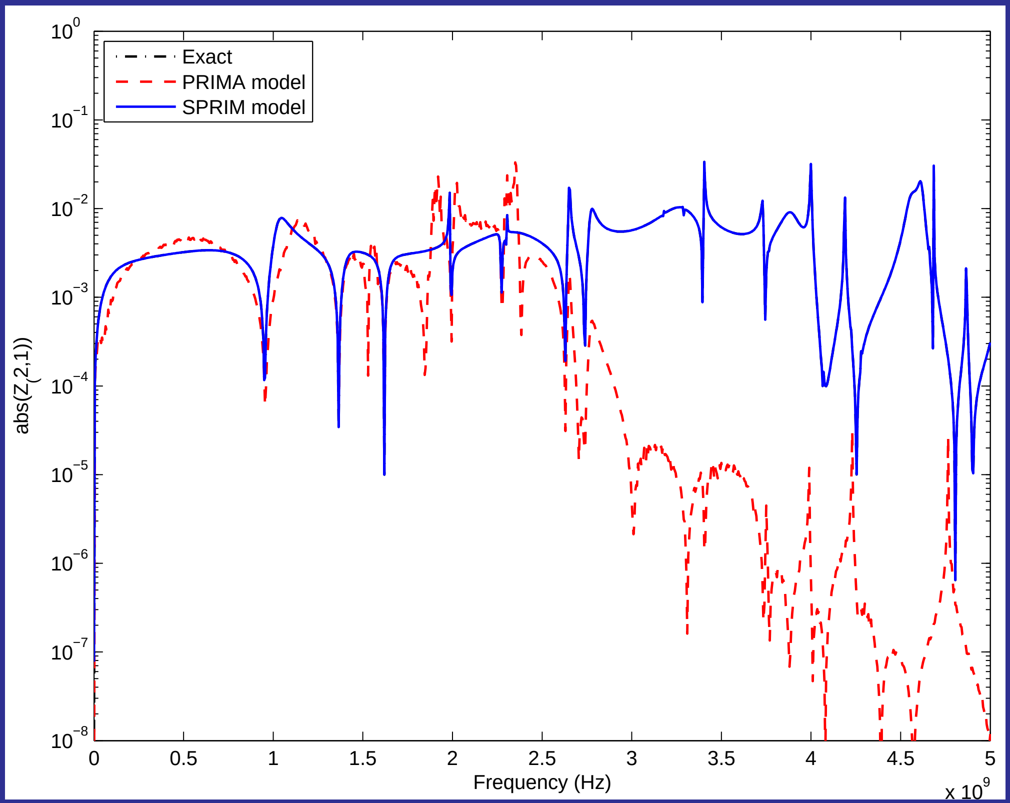
$$\mathbf{C}_n = \begin{bmatrix} \tilde{\mathbf{C}}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \; \mathbf{G}_n = \begin{bmatrix} \tilde{\mathbf{G}}_1 & \tilde{\mathbf{G}}_2 & \tilde{\mathbf{G}}_3 \\ -\tilde{\mathbf{G}}_2^{\mathsf{T}} & \mathbf{0} & \mathbf{0} \\ -\tilde{\mathbf{G}}_3^{\mathsf{T}} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \; \mathbf{B}_n = \begin{bmatrix} \tilde{\mathbf{B}}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{B}}_2 \end{bmatrix}$$

- Padé-type property:

$$\mathbf{H}_n(s) = \mathbf{H}(s) + \mathcal{O}\left((s - s_0)^q\right)$$

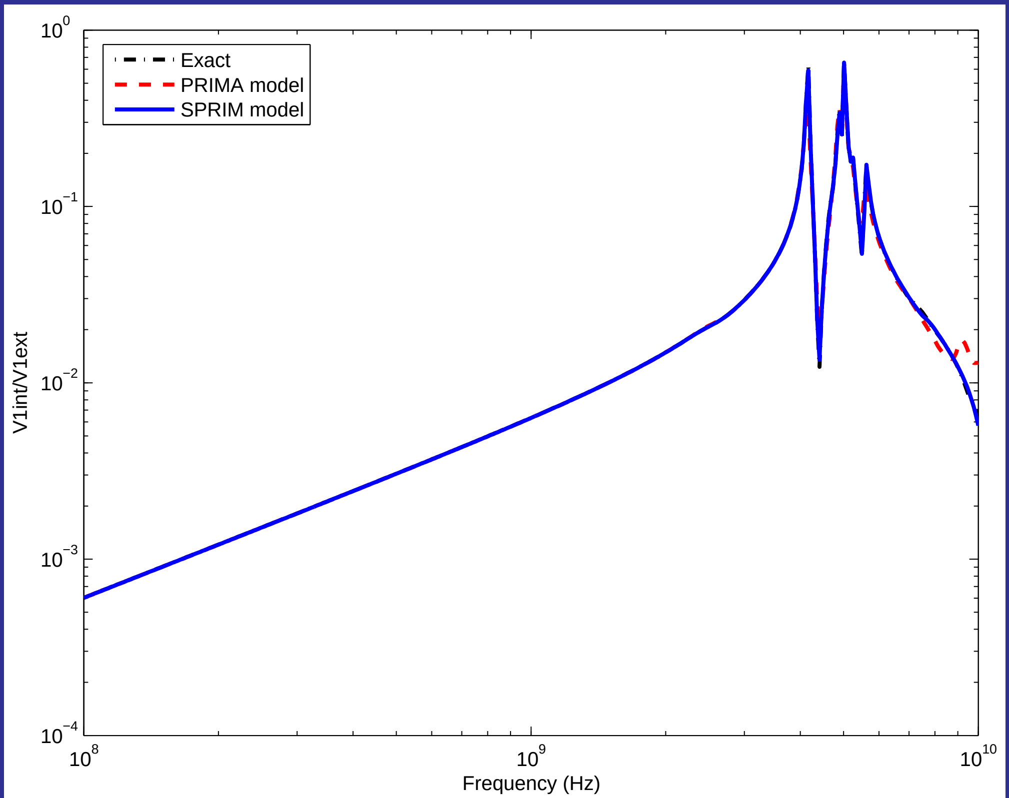
with $q \geq \lfloor n/m \rfloor$

An RCL circuit with mostly C's and L's



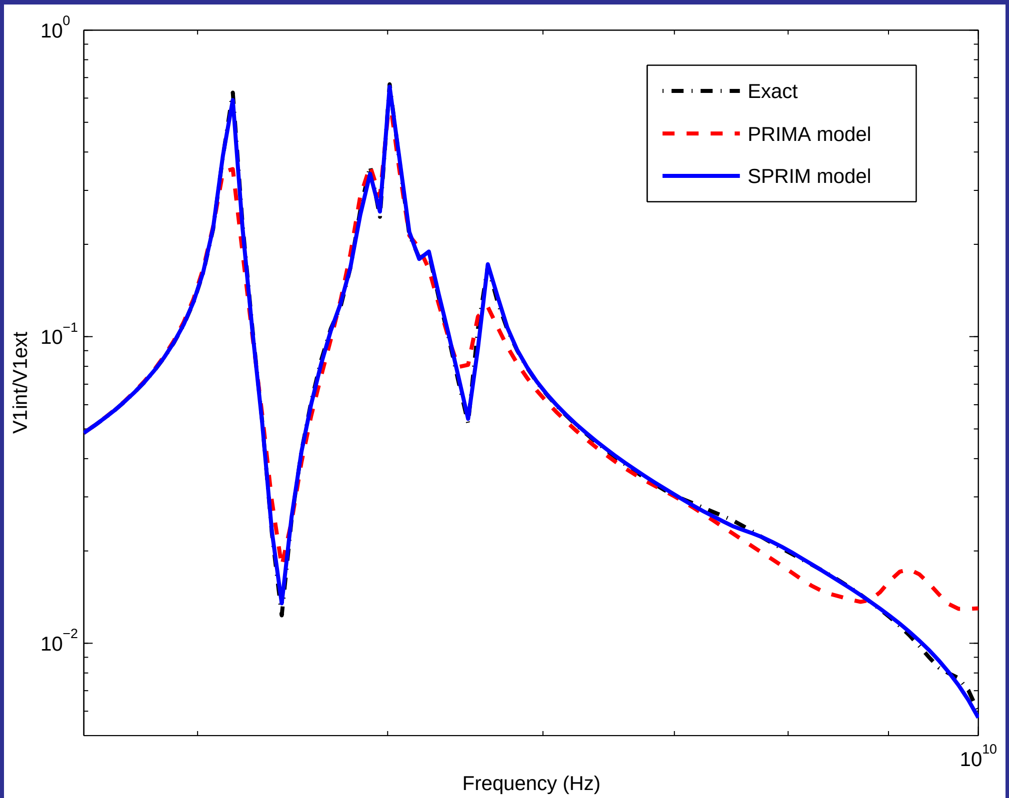
Exact and models corresponding to $n = 120$

A package example



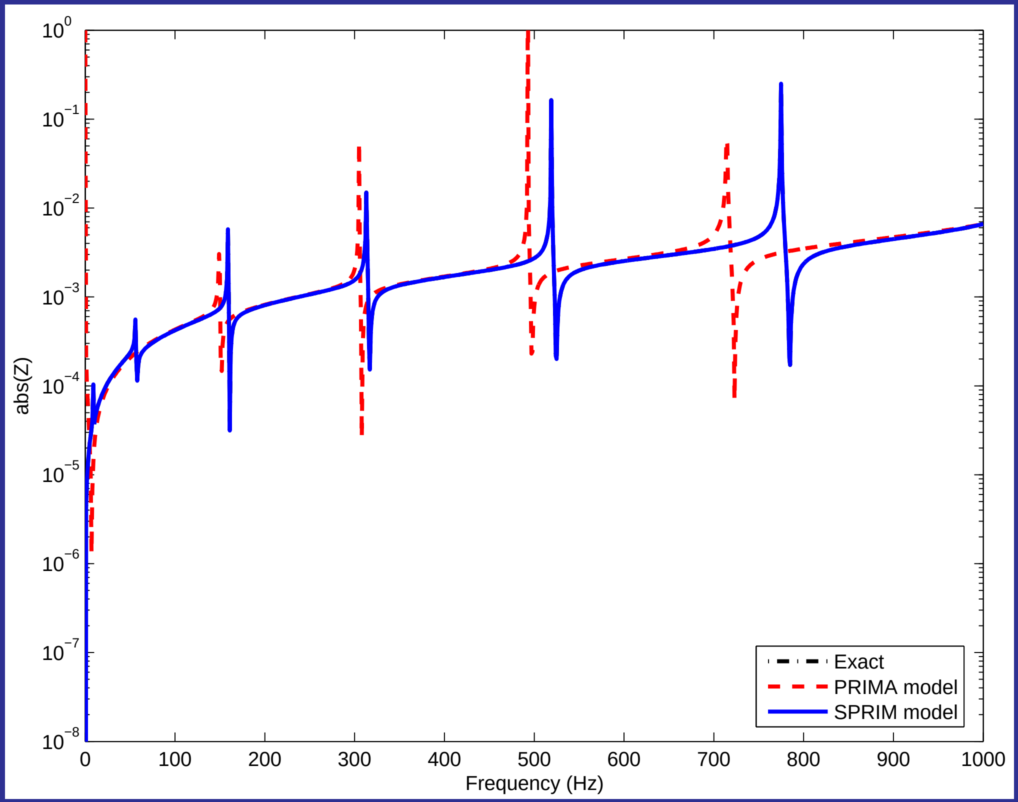
Exact and models corresponding to $n = 80$

Package example, high frequencies



Exact and models corresponding to size $n = 80$

A finite-element model of a shaft



Exact and models corresponding to $n = 15$

SPRIM vs. PRIMA

- Pros:
  - Same computational work
  - SPRIM preserves block structure and reciprocity
  - Higher accuracy
- Cons:
  - SPRIM models are two or three times as large as corresponding PRIMA models

Outline

- The order reduction problem
- Projection + Krylov = Padé-type reduction
- SPRIM for general RCL networks
- *SPRIM–SVD*
- Padé-type approximation properties of SPRIM
- Concluding remarks

SPRIM–SVD

- Columns of $\mathbf{V}_n$ span $\mathcal{K}_n(\mathbf{A}, \mathbf{R})$

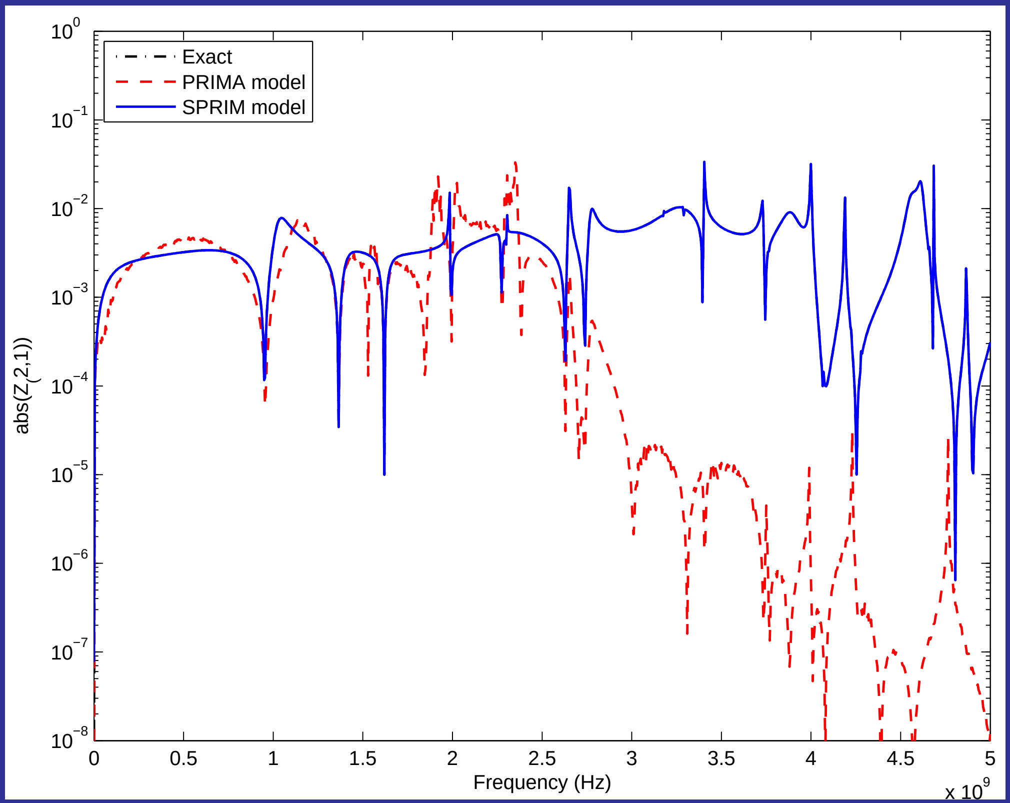
- SPRIM projection:

$$\mathbf{V}_n = \begin{bmatrix} \mathbf{V}_n^{(1)} \\ \mathbf{V}_n^{(2)} \\ \mathbf{V}_n^{(3)} \end{bmatrix} \quad \Longrightarrow \quad \tilde{\mathbf{V}}_n = \begin{bmatrix} \mathbf{V}_n^{(1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_n^{(2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_n^{(3)} \end{bmatrix}$$

- But:

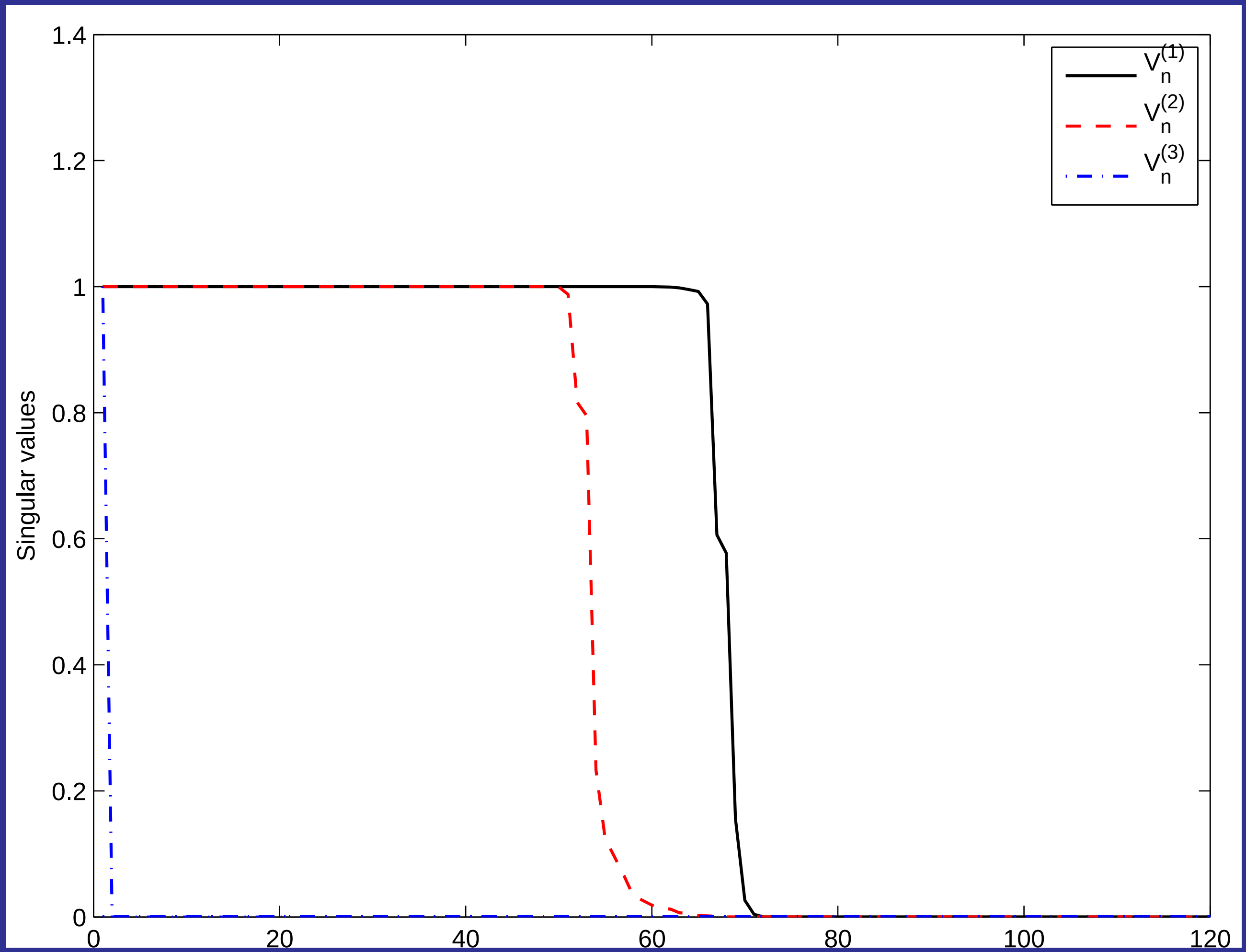
$$\# \text{ of rows of } \mathbf{V}_n^{(3)} \ll \# \text{ of rows of } \mathbf{V}_n^{(1)} \text{ and } \mathbf{V}_n^{(2)}$$

The RCL circuit with mostly C's and L's



Exact and models corresponding to $n = 120$

Singular values of projection subblocks



SPRIM–SVD, continued

- For $l = 1, 2, 3$, replace $\mathbf{V}_n^{(l)}$ by the matrix $\mathbf{U}_n^{(l)}$ containing the left singular vectors corresponding to the 'non-zero' singular values

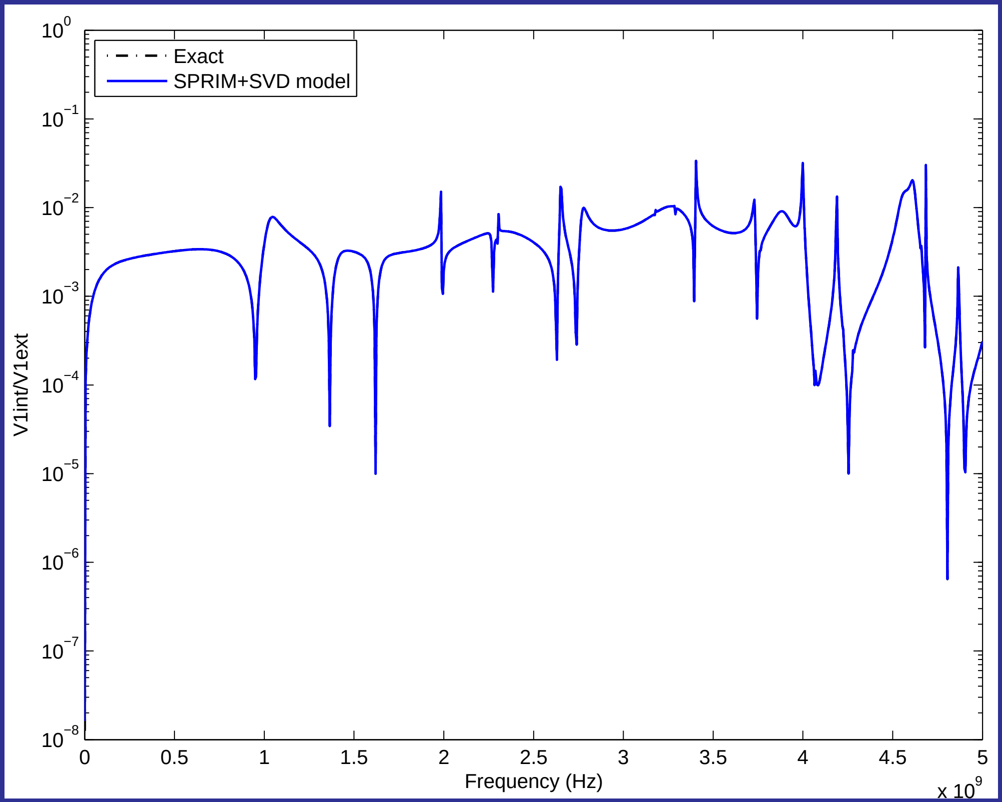
- SPRIM–SVD projection:

$$\tilde{\mathbf{V}}_n = \begin{bmatrix} \mathbf{V}_n^{(1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_n^{(2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_n^{(3)} \end{bmatrix} \quad \Longrightarrow \quad \hat{\mathbf{V}}_n = \begin{bmatrix} \mathbf{U}_n^{(1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_n^{(2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}_n^{(3)} \end{bmatrix}$$

- For the example:

$$3n = 360 \quad \Longrightarrow \quad 74 + 72 + 1 = 147$$

The RCL circuit with mostly C's and L's



Exact and models corresponding to $n = 120$

Theory of SPRIM–SVD?

- Number of small singular values of the subblocks $\mathbf{V}_n^{(1)}$ and $\mathbf{V}_n^{(2)}$?

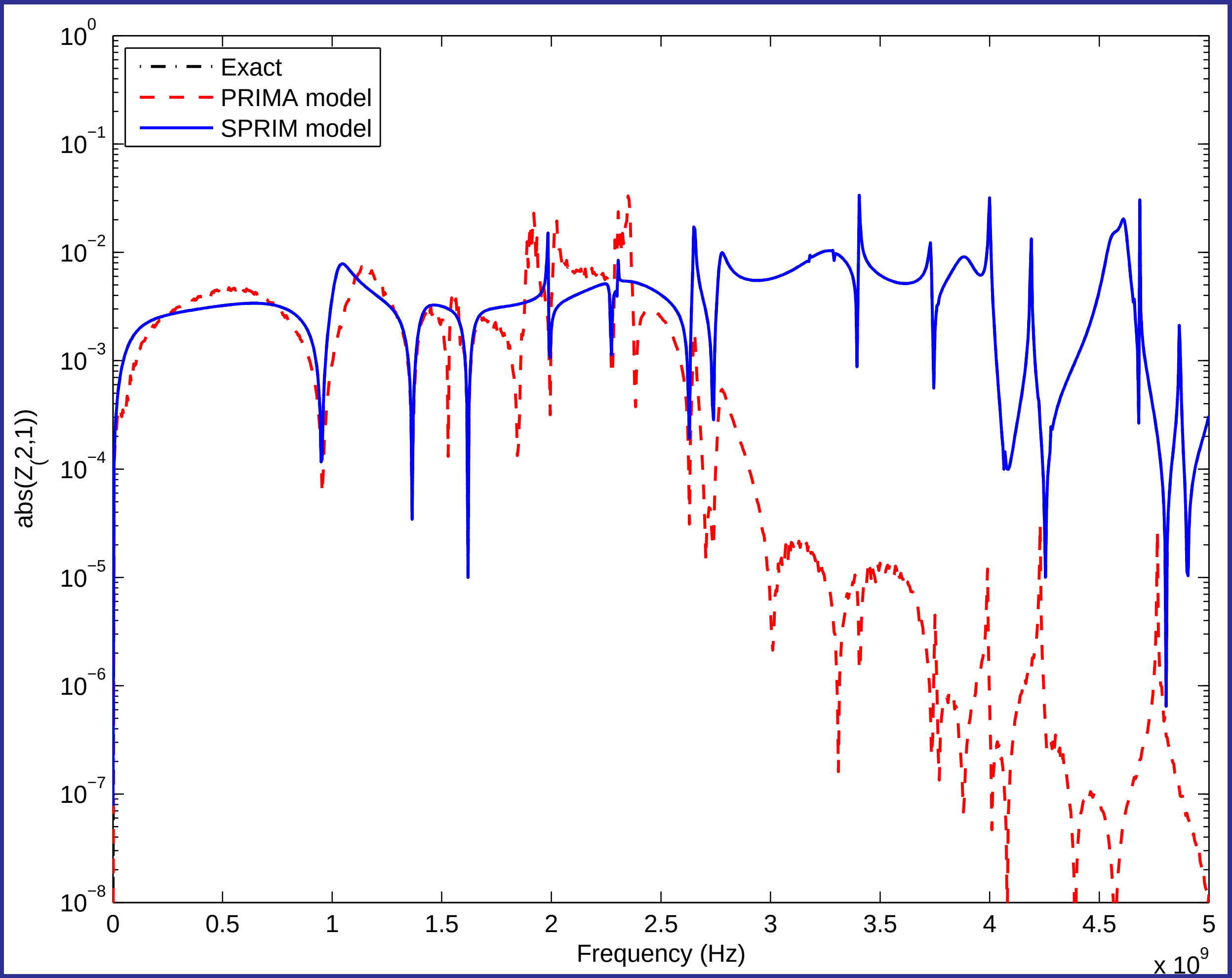
- Structure is understood in the case of no voltage sources, i.e., no thrird subblock $\mathbf{V}_n^{(3)}$
(F. '05)

- Key is the structure of the block Krylov subspaces $\mathcal{K}_n(\mathbf{A}, \mathbf{R})$; but what is it?

Outline

- The order reduction problem
- Projection + Krylov = Padé-type reduction
- SPRIM for general RCL networks
- SPRIM–SVD
- *Padé-type approximation properties of SPRIM*
- Concluding remarks

SPRIM vs. PRIMA



Padé-type property

- So far, we only know that both PRIMA and SPRIM produce Padé-type reduced-order models with

$$\mathbf{H}_n(s) = \mathbf{H}(s) + \mathcal{O}\left((s - s_0)^q\right), \quad \text{where} \quad q \geq \lfloor n/m \rfloor$$

- Can we say more in the case of SPRIM?

- Easy in the case of no third subblock $\mathbf{V}_n^{(3)}$ (F. '05)

- General case: $\mathbf{J}$-symmetric linear dynamical systems (F. '07)

$\mathbf{J}$-symmetry

- Recall:

$$\mathbf{C}\,\frac{d}{dt}\mathbf{x}(t) + \mathbf{G}\,\mathbf{x}(t) = \mathbf{B}\,\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{B}^{\top}\mathbf{x}(t)$$

where

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \mathbf{G}_3 \\ -\mathbf{G}_2^{\top} & \mathbf{0} & \mathbf{0} \\ -\mathbf{G}_3^{\top} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 \end{bmatrix}$$

- $\mathbf{C}$ and $\mathbf{G}$ are $\mathbf{J}$-symmetric:

$$\mathbf{J}\,\mathbf{C} = \mathbf{C}^{\top}\mathbf{J} \quad \text{and} \quad \mathbf{J}\,\mathbf{G} = \mathbf{G}^{\top}\mathbf{J}, \quad \text{where} \quad \mathbf{J} := \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix}$$

$\mathbf{J}$-symmetry, continued

- The input-output matrix $\mathbf{B}$ satisfies

$$\mathsf{Range}(\mathbf{J}\,\mathbf{B}) = \mathsf{Range}(\mathbf{B})$$

$\mathbf{J}_n$-symmetry of SPRIM models

- The SPRIM models

$$\mathbf{C}_n \frac{d}{dt}\mathbf{z}(t) + \mathbf{G}_n \mathbf{z}(t) = \mathbf{B}_n \mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{B}_n^\top \mathbf{z}(t)$$

  preserve the structure of $\mathbf{C}_n$, $\mathbf{G}_n$, $\mathbf{B}_n$

- Therefore, $\mathbf{C}_n$ and $\mathbf{G}_n$ are $\mathbf{J}_n$-symmetric with

$$\mathbf{J}_n := \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix} \quad \text{and} \quad \text{Range}(\mathbf{J}_n \mathbf{B}_n) = \text{Range}(\mathbf{B}_n)$$

- Moreover, the projection matrix $\mathbf{V}_n$ satisfies

$$\mathbf{J}\,\mathbf{V}_n = \mathbf{V}_n \mathbf{J}_n$$

Padé-type property

- **Theorem** (F., '05 and '07)

  For $\mathbf{J}$-symmetric systems and real expansion points s_0, the n-th SPRIM model is $\mathbf{J}_n$-symmetric and satisfies

  $$\mathbf{H}_n(s) = \mathbf{H}(s) + \mathcal{O}\left((s - s_0)^{\tilde{q}}\right), \quad \text{where} \quad \tilde{q} \geq 2\lfloor n/m \rfloor$$

- Twice as accurate as PRIMA!

Outline

- The order reduction problem
- Projection + Krylov = Padé-type reduction
- SPRIM for general RCL networks
- SPRIM–SVD
- Padé-type approximation properties of SPRIM
- *Concluding remarks*

Concluding remarks

- SPRIM and SPRIM–SVD for general RCL networks
- Key property for higher accuracy of SPRIM:
  $\mathbf{J}_n$-symmetry reduced-order models
- Theory of the zero singular values exploited in SPRIM–SVD?
- Projection-based reduction requires the storage of $\mathbf{V}_n \in \mathbb{R}^{N \times n}$ and is thus limited to moderately large N
- Structure-preserving reduction for truly large-scale RCL networks?