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

The graph shows the number of transistors in Intel CPUs from 1975 to 2000. The doubling time of the fitted line is 2.0 years. The key points are:

- 4004 (1975)
- 8086 (1980)
- 80286 (1985)
- 80386 (1990)
- P6 (Pentium Pro) (2000)

The graph indicates that the number of transistors in Intel CPUs has doubled every 2.0 years according to Moore’s law.
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 $O(10^6)$ circuit elements per network
- Up to $O(10^6)$ 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

\[ C \frac{d}{dt} x(t) + G x(t) = B u(t) \]
\[ y(t) = B^T x(t) \]

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

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

- \( m \) inputs, \( m \) outputs

- \( sC + G \) is nonsingular except for finitely many values of \( s \in \mathbb{C} \)
Reduced-order models

- System of DAEs of the same form:
  \[ C_n \frac{d}{dt} z(t) + G_n z(t) = B_n u(t) \]
  \[ \tilde{y}(t) = B_n^T z(t) \]

- But now:
  \[ C_n, G_n \in \mathbb{R}^{n \times n} \quad \text{and} \quad B_n \in \mathbb{R}^{n \times m} \]

where \( n \ll N \)
Transfer functions

- Original descriptor system:
  \[ H(s) = B^T (s \, C + G)^{-1} \, B \]

- Reduced-order model:
  \[ H_n(s) = B_n^T (s \, C_n + G_n)^{-1} \, B_n \]

- ‘Good’ reduced-order model
  \[ \iff \] ‘Good’ approximation \( H_n \approx H \)
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

\[ C \frac{d}{dt} x(t) + G x(t) = B u(t) \]

\[ y(t) = B^T x(t) \]

where

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

• Moreover:

\[ C \succeq 0 \quad \text{and} \quad G + G^T \succeq 0 \]

(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 $V_n \in \mathbb{R}^{N \times n}$ be any matrix with full column rank $n$

- Use $V_n$ to explicitly project the data matrices of

$$C \frac{d}{dt}x(t) + G x(t) = Bu(t)$$
$$y(t) = B^T x(t)$$

onto the subspace spanned by the columns of $V_n$
Projection-based reduction, continued

• Resulting reduced-order model

\[ C_n \frac{d}{dt} z(t) + G_n z(t) = B_n u(t) \]
\[ \tilde{y}(t) = B_n^T z(t) \]

where

\[ C_n = V_n^T C V_n, \quad G_n = V_n^T G V_n, \quad B_n = V_n^T B \]

• Passivity is preserved:

\[ C \succeq 0, \quad G + G^T \succeq 0 \quad \Rightarrow \quad C_n \succeq 0, \quad G_n + G_n^T \succeq 0 \]
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 $V_n$ be any matrix whose columns span the $n$-th Krylov subspace $K_n(A, R)$ where

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

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

- Projection + Krylov subspace = \textbf{Pade-type approximant}:

  $$H_n(s) = H(s) + O((s - s_0)^q), \quad \text{where} \quad q \geq \lfloor n/m \rfloor$$
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:

  \[
  C_n = \quad
  G_n = \quad
  B_n = \quad
  \]
Outline

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

- As in PRIMA, let $V_n$ be any matrix such that

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

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

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

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

• Recall:

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

• Partition \( V_n \) accordingly:

\[
V_n = \begin{bmatrix}
V_n^{(1)} \\
V_n^{(2)} \\
V_n^{(3)} \\
\end{bmatrix}
\]
SPRIM, continued

- Set

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

- Then: \( K_n(A, R) = \text{colspan } V_n \subseteq \text{colspan } \tilde{V}_n \)

- This guarantees a Padé-type property!
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

\[ \tilde{V}_n = \begin{bmatrix} V_n^{(1)} & 0 & 0 \\ 0 & V_n^{(2)} & 0 \\ 0 & 0 & V_n^{(3)} \end{bmatrix} \]
• The projection now preserves this structure:

\[
C_n = \begin{bmatrix}
\tilde{C}_1 & 0 & 0 \\
0 & \tilde{C}_2 & 0 \\
0 & 0 & 0
\end{bmatrix},
G_n = \begin{bmatrix}
\tilde{G}_1 & \tilde{G}_2 & \tilde{G}_3 \\
-\tilde{G}_2^T & 0 & 0 \\
-\tilde{G}_3^T & 0 & 0
\end{bmatrix},
B_n = \begin{bmatrix}
\tilde{B}_1 & 0 \\
0 & 0 \\
0 & \tilde{B}_2
\end{bmatrix}
\]

• Padé-type property:

\[
H_n(s) = H(s) + \mathcal{O}((s - s_0)^q)
\]

with \( q \geq \left\lfloor \frac{n}{m} \right\rfloor \)
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

• \textit{SPRIM--SVD}

• Padé-type approximation properties of SPRIM

• Concluding remarks
SPRIM–SVD

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

- SPRIM projection:

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

- But:

  $\# \text{ of rows of } V_n^{(3)} \ll \# \text{ of rows of } V_n^{(1)} \text{ and } V_n^{(2)}$
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 $V_n^{(l)}$ by the matrix $U_n^{(l)}$ containing the left singular vectors corresponding to the 'non-zero' singular values.

• SPRIM–SVD projection:

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

• For the example:

$$3n = 360 \quad \Rightarrow \quad 74 + 72 + 1 = 147$$
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 $V_n^{(1)}$ and $V_n^{(2)}$?

• Structure is understood in the case of no voltage sources, i.e., no third subblock $V_n^{(3)}$ (F. ’05)

• Key is the structure of the block Krylov subspaces $\mathcal{K}_n(A, R)$; 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

The graph shows the absolute value of the impedance $|Z(2,1)|$ across different frequencies. The x-axis represents the frequency in Hz, ranging from $10^0$ to $10^5$ Hz, and the y-axis represents the absolute value on a logarithmic scale from $10^{-8}$ to $10^0$. The graph compares the exact solution (dashed line) with the PRIMA model (dotted line) and the SPRIM model (solid line).
Padé-type property

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

\[ H_n(s) = H(s) + O((s - s_0)^q), \quad \text{where} \quad q \geq \lfloor n/m \rfloor \]

• Can we say more in the case of SPRIM?

• Easy in the case of no third subblock \( V_n^{(3)} \) (F. ’05)

• General case: \( J \)-symmetric linear dynamical systems (F. ’07)
J-symmetry

Recall:

\[
C \frac{d}{dt} x(t) + G x(t) = B u(t)
\]

\[
y(t) = B^T x(t)
\]

where

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

C and G are J-symmetric:

\[
J C = C^T J \quad \text{and} \quad J G = G^T J,
\]

where

\[
J := \begin{bmatrix}
I & 0 & 0 \\
0 & -I & 0 \\
0 & 0 & -I
\end{bmatrix}
\]
J-symmetry, continued

- The input-output matrix $B$ satisfies

$$\text{Range}(JB) = \text{Range}(B)$$
\(J_n\)-symmetry of SPRIM models

- The SPRIM models
  \[
  C_n \frac{d}{dt} z(t) + G_n z(t) = B_n u(t)
  \]
  \[
  y(t) = B_n^T z(t)
  \]
  preserve the structure of \(C_n\), \(G_n\), \(B_n\).

- Therefore, \(C_n\) and \(G_n\) are \(J_n\)-symmetric with

\[
J_n := \begin{bmatrix}
  I & 0 & 0 \\
  0 & -I & 0 \\
  0 & 0 & -I
\end{bmatrix}
\]

  and \(\text{Range}(J_n B_n) = \text{Range}(B_n)\)

- Moreover, the projection matrix \(V_n\) satisfies

\[
J V_n = V_n J_n
\]
Padé-type property

- **Theorem** (F., '05 and '07)
  
  For $J$-symmetric systems and real expansion points $s_0$, the $n$-th SPRIM model is $J_n$-symmetric and satisfies
  
  $$H_n(s) = H(s) + O((s - s_0)^{\tilde{q}}), \quad \text{where} \quad \tilde{q} \geq 2 \lfloor n/m \rfloor$$

- 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: $J_n$-symmetry reduced-order models

- Theory of the zero singular values exploited in SPRIM–SVD?

- Projection-based reduction requires the storage of $V_n \in \mathbb{R}^{N \times n}$ and is thus limited to moderately large $N$

- Structure-preserving reduction for truly large-scale RCL networks?