

---

# Comparison of Model Reduction Methods with Applications to Circuit Simulation\*

Roxana Ionutiu, Sanda Lefteriu, and Athanasios C. Antoulas

Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA  
rlonutiu@rice.edu, slefteri@rice.edu, aca@rice.edu

**Summary.** We compare different model reduction methods applied to the dynamical system of a coupled transmission line: balanced truncation (BT), truncation by balancing one gramian (or PMTBR - poor man's truncated balanced reduction), positive real balanced truncation (PRBT) and its Hamiltonian implementation (PRBT-Ham), PRIMA, spectral zero method (SZM) and its Hamiltonian implementation (SZM-Ham), and finally, optimal  $\mathcal{H}_2$ . Their performance is analyzed in terms of several criteria such as: preservation of controllability, observability, stability and passivity, relative  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norms, and the computational cost involved.

## 1 Introduction

This paper presents different reduction methods together with results obtained by applying each method on a dynamical system given by a coupled transmission line. In Sect. 2, a modified nodal analysis (MNA)-similar representation of the system is derived. The model reduction methods are grouped in two main categories, *gramian based* and *Krylov based*, discussed in sections 3 and 4 respectively. Sect. 3 outlines the theory behind gramian based reduction methods: BT, PMTBR and PRBT. Krylov based reduction methods PRIMA, SZM and optimal  $\mathcal{H}_2$  are described in Sect. 4. In Sect. 5 we compare all methods in terms of: preservation of some important properties like controllability, observability, stability and passivity, the relative  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norms and in terms of the computational cost. In Sect. 6, error systems resulting from different methods are compared. This allows us to identify frequency ranges where one particular method approximates the original system more accurately. Sect. 7 presents additional results obtained with the optimal  $\mathcal{H}_2$  method. Finally, Sect. 8 summarizes our analysis and motivates further research.

## 2 State-space representation

The model reduction problem of transmission lines has been studied extensively, see for instance [8]. Our system consists of two transmission lines with inductive

---

\* This work was supported in part by the NSF through Grants CCR-0306503, ACI-0325081, and CCF-0634902. Invited Paper at SCEE-2006

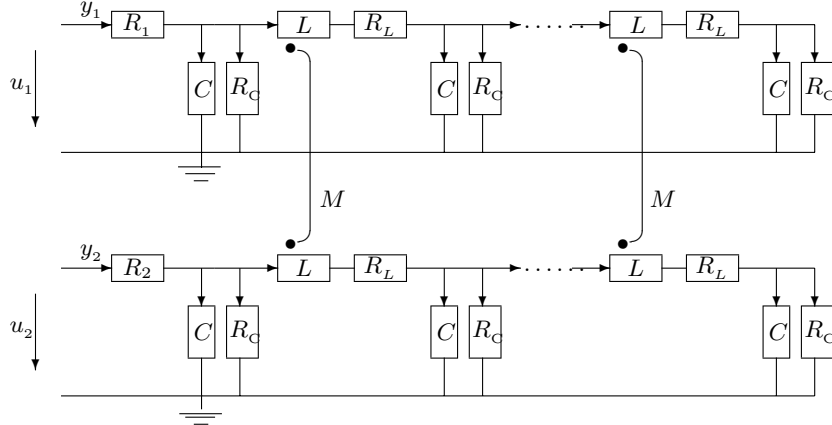


Fig. 1: Two coupled transmission lines

coupling as shown in Fig. 1. Each section consists of an inductor and its associated resistor, in series with a capacitor and its associated resistor. The first section has no inductor. All capacitor values  $C_i$  are equal. The same holds for the inductors  $L_i$ , the coupling inductors  $M_{ij}$ , the resistors associated with the capacitors  $R_{C_i}$ , the resistors associated with the inductors  $R_{L_i}$  and the input resistors,  $R_1$  and  $R_2$ .

To simulate this circuit, the *state-space representation* of the system needs to be derived. Choosing the state variables as the currents through the inductors and the voltages across the capacitors, we obtain a system of order  $n = 4N - 2$ , where  $N$  is the number of sections of the circuit. The state-space representation in *modified nodal analysis (MNA)*-similar form is the following:

$$\left. \begin{aligned} \mathbb{C}\dot{\mathbf{x}}(t) &= \mathbf{G}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{L}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \right\} \quad (1)$$

where  $\mathbb{C} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{G} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times 2}$ ,  $\mathbf{L} \in \mathbb{R}^{2 \times n}$ ,  $\mathbf{D} \in \mathbb{R}^{2 \times 2}$  and  $\mathbf{x}(t) \in \mathbb{R}^n$ ,  $\mathbf{u}(t) \in \mathbb{R}^2$ ,  $\mathbf{y}(t) \in \mathbb{R}^2$ .

The problem will be studied under the following simplifying assumptions:

- (1) the equations are in an MNA-similar form so that the resulting  $\mathbb{C}$  matrix in (1) is nonsingular and positive definite (this means that all variables are state variables and none is redundant). In general,  $\mathbb{C}$  resulting from circuit simulation is singular, due to additionally generated variables at the nodes between  $L_i$  and  $R_{L_i}$ .
- (2) The transmission line has one input and one output, that is  $u_2 = 0$  and only  $y_1$  is observed, so that  $\mathbf{u} = u_1$  and  $\mathbf{y} = y_1$ .

These assumptions are made to ease certain technical issues and allow a comparison of all reduction methods enumerated above; for instance, the optimal  $\mathcal{H}_2$  method is currently available for single-input-single-output (SISO) systems only. None of these assumptions is essential for the validity of the results presented. Similar results for a system with MNA equations (where  $\mathbb{C}$  is singular), using in part results from [5], will be reported in a future analysis.

For simplicity we will show the form of the equations by deriving them for  $N = 3$  sections, namely for a circuit with 6 capacitors and 4 inductors, resulting in 10 states. In particular, the elements of the first line, from left to right will be

$$R_1, C_1, R_{C_1}; L_1, R_{L_1}, C_2, R_{C_2}; L_2, R_{L_2}, C_3, R_{C_3},$$

and those of the second line from left to right

$$R_2, C_4, R_{C_4}; L_3, R_{L_3}, C_5, R_{C_5}; L_4, R_{L_4}, C_6, R_{C_6}.$$

The state variables are:

$$\mathbf{x}_{C_1}, \mathbf{x}_{L_1}, \mathbf{x}_{C_2}, \mathbf{x}_{L_2}, \mathbf{x}_{C_3}, \mathbf{x}_{C_4}, \mathbf{x}_{L_3}, \mathbf{x}_{C_5}, \mathbf{x}_{L_4}, \mathbf{x}_{C_6},$$

and the state is chosen as:

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_C \\ \mathbf{x}_L \end{pmatrix}, \quad \mathbf{x}_C = \begin{pmatrix} \mathbf{x}_{C_1} \\ \mathbf{x}_{C_2} \\ \mathbf{x}_{C_3} \\ \mathbf{x}_{C_4} \\ \mathbf{x}_{C_5} \\ \mathbf{x}_{C_6} \end{pmatrix}, \quad \mathbf{x}_L = \begin{pmatrix} \mathbf{x}_{L_1} \\ \mathbf{x}_{L_3} \\ \mathbf{x}_{L_2} \\ \mathbf{x}_{L_4} \end{pmatrix}.$$

The associated system matrices are<sup>2</sup>:

$$\mathbf{C} = \begin{pmatrix} \tilde{\mathbf{C}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{L}} \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} -\mathbf{R}_C & \tilde{\mathbf{E}} \\ -\tilde{\mathbf{E}}^* & -\mathbf{R}_L \end{pmatrix}, \quad \mathbf{B} = \left( \frac{1}{R_1} \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \right)^*,$$

$\mathbf{L} = -\mathbf{B}^*$  and  $\mathbf{D} = \frac{1}{R_1}$ , where:

$$\tilde{\mathbf{C}} = \text{diag}(C_1, C_2, C_3, C_4, C_5, C_6), \quad \tilde{\mathbf{L}} = \begin{pmatrix} L_1 & M_{13} & & & & \\ M_{13} & L_3 & & & & \\ & & L_2 & M_{24} & & \\ & & M_{24} & L_4 & & \end{pmatrix} \text{ and}$$

$$\tilde{\mathbf{E}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{R}_C = \text{diag}\left(\frac{1}{R_1} + \frac{1}{R_{C_1}}, \frac{1}{R_{C_2}}, \frac{1}{R_{C_3}}, \frac{1}{R_2} + \frac{1}{R_{C_4}}, \frac{1}{R_{C_5}}, \frac{1}{R_{C_6}}\right)$$

$$\mathbf{R}_L = \text{diag}(R_{L_1}, R_{L_3}, R_{L_2}, R_{L_4}).$$

The values of the elements used in the simulation are as follows: the input resistors are  $R_1 = R_2 = 10\Omega$ , the capacitors are  $C_i = 5.4 \cdot 10^{-12}F$  and the associated resistors  $R_{C_i} = 10^3\Omega$ , ( $i = 1, \dots, 6$ ), the inductors are  $L_i = 0.25 \cdot 10^{-9}H$ , ( $i = 1, \dots, 4$ ), the mutual inductors are  $M_{ij} = 0.2L_i$  ( $i = 1, 2, j = 3, 4$ ) of that value. The associated resistors are zero  $R_{L_i} = 0$ , ( $i = 1, \dots, 4$ ).

### 3 Gramian based methods

Gramian based methods involve diagonalization of gramians by congruence. These can either be the positive definite solutions to the Lyapunov equations (called *controllability* and *observability gramians*) or the positive definite solutions to algebraic Riccati equations (called *positive real controllability* and *observability gramians*). The methods that we discuss are balanced truncation (BT) in Sect. 3.1 which

<sup>2</sup> For a matrix  $\mathbf{M}$ ,  $\mathbf{M}^*$  denotes transposition followed by complex conjugation if the matrix is complex.

performs simultaneous diagonalization of the controllability and the observability gramians, an equivalent of poor man's truncated balanced reduction (PMTBR) in Sect. 3.2 in which only one of the gramians is diagonalized and positive real balanced truncation (PRBT) in Sect. 3.3 in which positive definite solutions to the algebraic Riccati equations are simultaneously diagonalized.

### 3.1 Balanced truncation (BT)

The idea behind balanced truncation is to simultaneously diagonalize the two infinite gramians,  $\mathcal{P}$  and  $\mathcal{Q}$  [1]. These are the solutions to the controllability and observability *Lyapunov equations* respectively, which are associated with the state space formulation (1). The mathematical model of the system may come in two representations: standard state-space and MNA-similar representation (or invertible descriptor form), respectively. We describe the application of model reduction methods for both cases of models.

#### Standard state-space representation

The standard state-space representation  $(\mathbf{A}_{ss}, \mathbf{B}_{ss}, \mathbf{C}_{ss}, \mathbf{D}_{ss})$  is obtained from (1) by inverting the  $\mathbb{C}$  matrix.

$$\mathbf{A}_{ss} = \mathbb{C}^{-1}\mathbf{G}, \mathbf{B}_{ss} = \mathbb{C}^{-1}\mathbf{B}, \mathbf{C}_{ss} = -\mathbf{B}^*, \mathbf{D}_{ss} = \mathbf{D}$$

The controllability and observability gramians are given by the symmetric positive definite solutions to the controllability and observability Lyapunov equations:

$$\mathbf{A}_{ss}\mathcal{P} + \mathcal{P}\mathbf{A}_{ss}^* + \mathbf{B}_{ss}\mathbf{B}_{ss}^* = 0 \quad (2)$$

$$\mathbf{A}_{ss}^*\mathcal{Q} + \mathcal{Q}\mathbf{A}_{ss} + \mathbf{C}_{ss}^*\mathbf{C}_{ss} = 0 \quad (3)$$

BT is performed in two steps. First, the balancing projection is computed (both gramians become equal and diagonal, with the Hankel singular values (HSVs) on the diagonal). Second, the states which are equally difficult to reach and to observe are truncated. This amounts to eliminating the states corresponding to the HSVs which are below a certain tolerance. Setting a tolerance for the reduced system a priori defines the number of states to be kept. The procedure is the following.

1. Compute the Cholesky factors of  $\mathcal{P} = \mathbf{U}\mathbf{U}^*$  and  $\mathcal{Q} = \mathbf{L}\mathbf{L}^*$
2. Compute the singular value decomposition of the product  $\mathbf{U}^*\mathbf{L}$

$$\mathbf{U}^*\mathbf{L} = \mathbf{Z}\mathbf{S}\mathbf{Y}^* \quad (4)$$

The diagonal elements:  $\mathbf{S} = \text{diag}(\sigma_1, \dots, \sigma_n)$ ,  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ , where  $\sigma_i = \sqrt{\lambda_i(\mathcal{P}\mathcal{Q})}$  are the *Hankel singular values* of the system. Choosing only the first  $k$  singular values and the first  $k$  columns of  $\mathbf{Z}$  and  $\mathbf{Y}$  gives the reduced system of order  $k$  after applying the projection  $\mathbf{\Pi}$

3.  $\mathbf{\Pi} = \mathbf{V}\mathbf{W}^*$  where  $\mathbf{V} = \mathbf{U}\mathbf{Z}_k\mathbf{S}_k^{-\frac{1}{2}}$ ,  $\mathbf{V} \in \mathbb{R}^{n \times k}$ ,  $\mathbf{W} = \mathbf{L}\mathbf{Y}_k\mathbf{S}_k^{-\frac{1}{2}}$ ,  $\mathbf{W} \in \mathbb{R}^{n \times k}$

4. Compute the representation of the reduced system:

$$\tilde{\mathbf{A}}_{ss} = \mathbf{W}^*\mathbf{A}_{ss}\mathbf{V}, \tilde{\mathbf{B}}_{ss} = \mathbf{W}^*\mathbf{B}_{ss}, \tilde{\mathbf{C}}_{ss} = \mathbf{C}_{ss}\mathbf{V}, \tilde{\mathbf{D}}_{ss} = \mathbf{D}_{ss}$$

The corresponding diagonalized controllability and observability gramians are given by  $\tilde{\mathcal{P}} = \mathbf{W}^*\mathcal{P}\mathbf{W} = \mathbf{S}_k$ ,  $\tilde{\mathcal{Q}} = \mathbf{V}^*\mathcal{Q}\mathbf{V} = \mathbf{S}_k$  where  $\mathbf{S}_k$  is the matrix containing the largest  $k$  HSV's on the diagonal.

### Descriptor form representation

The MNA-similar representation is precisely (1). For simplicity, we rename the matrices in (1) to match the standard descriptor system representation:<sup>3</sup>

$$\mathbf{E}_{ds} = \mathbb{C}, \mathbf{A}_{ds} = \mathbf{G}, \mathbf{B}_{ds} = \mathbf{B}, \mathbf{C}_{ds} = \mathbf{L}, \mathbf{D}_{ds} = \mathbf{D}$$

The gramians are now the solutions to the following Lyapunov equations:

$$\mathbf{A}_{ds} \mathcal{P} \mathbf{E}_{ds}^* + \mathbf{E}_{ds} \mathcal{P} \mathbf{A}_{ds}^* + \mathbf{B}_{ds} \mathbf{B}_{ds}^* = 0 \quad (5)$$

$$\mathbf{A}_{ds}^* \hat{\mathcal{Q}} \mathbf{E}_{ds} + \mathbf{E}_{ds}^* \hat{\mathcal{Q}} \mathbf{A}_{ds} + \mathbf{C}_{ds}^* \mathbf{C}_{ds} = 0, \quad (6)$$

where  $\mathcal{P}$  in (5) is precisely the solution of (2), while the original observability gramian corresponding to the solution of (3) is obtained by means of the

congruence transformation  $\mathcal{Q} = \mathbf{E}_{ds}^* \hat{\mathcal{Q}} \mathbf{E}_{ds}$

The balancing and truncation procedures follow as described in Sect. 3.1, where (4) is replaced by:

$$\mathbf{U}^* \mathbf{E}_{ds} \mathbf{L} = \mathbf{Z} \mathbf{S} \mathbf{Y}^*$$

The system representation in the new basis now becomes:

$$\begin{aligned} \tilde{\mathbf{E}}_{ds} &= \mathbf{W}^* \mathbf{E}_{ds} \mathbf{V} = \mathbf{I}_k, \tilde{\mathbf{A}}_{ds} = \mathbf{W}^* \mathbf{A}_{ds} \mathbf{V}, \\ \tilde{\mathbf{B}}_{ds} &= \mathbf{W}^* \mathbf{B}_{ds}, \tilde{\mathbf{C}}_{ds} = \mathbf{C}_{ds} \mathbf{V}, \tilde{\mathbf{D}}_{ds} = \mathbf{D}_{ds}. \end{aligned}$$

Gramians  $\mathcal{P}$  and  $\mathcal{Q}$  are simultaneously diagonalized as mentioned in Sect. 3.1.

### Solving the Lyapunov equation

There are many methods for solving the Lyapunov equation  $\mathbf{A} \mathcal{P} + \mathcal{P} \mathbf{A}^* = \mathbf{Q}$  [1]. We will use the so-called *square-root method*, which directly computes  $\mathbf{U}$  such that  $\mathcal{P} = \mathbf{U} \mathbf{U}^*$ . In MATLAB, this is implemented by `lyapchol`. Another important tool is the *sign function method*, which is discussed next.

The Lyapunov equation is a particular form of the Sylvester equation  $\mathbf{A} \mathbf{X} + \mathbf{X} \mathbf{B} = \mathbf{C}$ . To treat this generalized case, consider a matrix of the type

$$\mathbf{Z} = \begin{pmatrix} \mathbf{A} & -\mathbf{C} \\ \mathbf{0} & -\mathbf{B} \end{pmatrix},$$

where  $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\Re(\lambda_i(\mathbf{A})) < 0$ ,  $\mathbf{B} \in \mathbb{R}^{k \times k}$ ,  $\Re(\lambda_i(\mathbf{B})) < 0$ , and  $\mathbf{C} \in \mathbb{R}^{n \times k}$ . The sign function iteration  $\mathbf{Z}_{n+1} = (\mathbf{Z}_n + \mathbf{Z}_n^{-1})/2$ ,  $\mathbf{Z}_0 = \mathbf{Z}$  converges to

$$\lim_{j \rightarrow \infty} \mathbf{Z}_j = \begin{pmatrix} -\mathbf{I}_n & 2\mathbf{X} \\ \mathbf{0} & \mathbf{I}_k \end{pmatrix}$$

where  $\mathbf{X}$  is the solution to the equation  $\mathbf{A} \mathbf{X} + \mathbf{X} \mathbf{B} = \mathbf{C}$ .

For the Lyapunov equation  $\mathbf{A} \mathcal{P} + \mathcal{P} \mathbf{A}^* = \mathbf{Q}$ , the starting matrix is

<sup>3</sup> As mentioned earlier, our analysis of the system in descriptor form is restricted to the case in which matrix  $\mathbf{E}_{ds} = \mathbb{C}$  is invertible.

$$\mathbf{Z} = \begin{pmatrix} \mathbf{A} & -\mathbf{Q} \\ \mathbf{0} & -\mathbf{A}^* \end{pmatrix}, \mathbf{A} \in \mathbb{R}^{n \times n}, \Re(\lambda_i(\mathbf{A})) < 0 \Rightarrow \mathbf{Z}_j = \begin{pmatrix} \mathbf{A}_j & -\mathbf{Q}_j \\ \mathbf{0} & -\mathbf{A}_j^* \end{pmatrix}$$

where the iterations can be written as follows

$$\mathbf{A}_{j+1} = \frac{1}{2} (\mathbf{A}_j + \mathbf{A}_j^{-1}), \mathbf{A}_0 = \mathbf{A}; \quad \mathbf{Q}_{j+1} = \frac{1}{2} (\mathbf{Q}_j + \mathbf{A}_j^{-1} \mathbf{Q}_j \mathbf{A}_j^{-*}), \mathbf{Q}_0 = \mathbf{Q}.$$

The limits of these iterations are  $\mathbf{A}_\infty = -\mathbf{I}_n$  and  $\mathbf{Q}_\infty = 2\mathcal{P}$  where  $\mathcal{P}$  is the solution of the Lyapunov equation  $\mathbf{A}\mathcal{P} + \mathcal{P}\mathbf{A}^* = \mathbf{Q}$ .

Often, the constant term in the Lyapunov equation above is provided in factored form  $\mathbf{Q} = \mathbf{R}\mathbf{R}^*$ . As a consequence, it is possible to obtain the solution in factored form. In particular, the  $(j+1)^{st}$  iterate in factored form is

$$\mathbf{Q}_{j+1} = \mathbf{R}_{j+1} \mathbf{R}_{j+1}^* \text{ where } \mathbf{R}_{j+1} = \frac{1}{\sqrt{2}} [\mathbf{R}_j, \mathbf{A}_j^{-1} \mathbf{R}_j] \Rightarrow \mathbf{Q}_\infty = \mathbf{R}_\infty \mathbf{R}_\infty^* = 2\mathcal{P}$$

$\mathbf{R}_\infty$  has infinitely many columns, although its rank cannot exceed  $n$ . This can be avoided by performing at each step a rank revealing RQ factorization  $\mathbf{R}_j \mathbf{P}_j = \mathbf{T}_j \mathbf{U}_j$  with  $\mathbf{P}_j$  the permutation matrix and  $\mathbf{T}_j \mathbf{P}_j = [\Delta_j^*, \mathbf{0}]^*$ .  $\Delta_j$  is upper triangular and  $\mathbf{U}_j \mathbf{U}_j^* = \mathbf{I}_j$ . Thus, at the  $j^{th}$  step,  $\mathbf{R}_j$  is replaced by  $\Delta_j$  which has as many columns as the rank of  $\mathbf{R}_j$ . For accelerating convergence, the eigenvalues of  $\mathbf{A}$  can be scaled [3]: at each step,  $\mathbf{A}_j$  is replaced by  $\frac{1}{\gamma_j} \mathbf{A}_j$  where the factors  $\gamma_j$  can be chosen as  $\gamma_j = |\det(\mathbf{A}_j)|^{\frac{1}{n}}$  in order to minimize the distance of the geometric mean of the eigenvalues of  $\mathbf{A}_j$  from 1.

Convergence of the iteration which uses scaling is quadratic. The time required to compute the Cholesky factor by MATLAB's `lyapchol` function versus the iterative implementation of the sign function method in [3] is as follows: on a Pentium M at 1.3Ghz with 768MB RAM, `lyapchol` runs in 0.751s for a matrix  $\mathbf{A}$  of dimension 242, while the implementation in [3] requires 5.423s and converges in  $16 \approx \sqrt{242}$  steps. Even if, in theory, no scaling should also give quadratic convergence, in practice, due to numerical issues, convergence occurs after 20 steps.

### 3.2 Truncation by diagonalization of one gramian or poor man's truncated balanced reduction (PMTBR)

For the standard state-space representation, the procedure is the following [1].

1. Compute the gramian to be diagonalized (controllability gramian  $\mathcal{P}$  in our case)
2. Compute the eigenvalue decomposition of  $\mathcal{P} = \mathbf{V}\Sigma\mathbf{V}^*$
3. Choose the eigenvectors corresponding to the largest  $k$  eigenvalues to obtain the transformation  $\mathbf{T} = \mathbf{V}_k^*$
4. The reduced system is

$$\tilde{\mathbf{A}}_{ss} = \mathbf{T}\mathbf{A}_{ss}\mathbf{T}^*, \tilde{\mathbf{B}}_{ss} = \mathbf{T}\mathbf{B}_{ss}, \tilde{\mathbf{C}}_{ss} = \mathbf{C}_{ss}\mathbf{T}^*, \tilde{\mathbf{D}}_{ss} = \mathbf{D}_{ss}$$

PMTBR is presented in [10] and uses numerical quadrature to approximate the gramian  $\mathcal{P}$ , without solving the Lyapunov equation. The algorithm used in our analysis, however, diagonalizes the exact solution  $\mathcal{P}$  of the Lyapunov equation. As mentioned in Sect. 3.1, the solution to the Lyapunov equation can be computed either by using the *sign function method* or by using MATLAB's `lyapchol` function.

### 3.3 Positive real balanced truncation (PRBT)

Coupled transmission lines such as the one in Fig. 1 are passive systems, with *positive real* transfer functions (further information on passivity and positive realness is provided in [1]). We are therefore interested in reduced order models that are passive. In general, BT is not a passivity preserving method, since the resulting reduced system may have a non-positive real transfer function. PRBT, however, is a passivity preserving method. It yields reduced order models with positive real transfer functions by simultaneously diagonalizing the positive definite solutions  $\mathcal{P}$  and  $\mathcal{Q}$  of the controllability and observability algebraic *Riccati equations* respectively. This desirable result cannot be guaranteed with BT, where the solutions to the Lyapunov equations are diagonalized, rather than the solutions the Riccati equations. Riccati equations have a different form depending on whether the system is in standard state-space form or in descriptor form.

*Historical note:* this method was first introduced by Ober [6] and rediscovered by Phillips, Daniel and Silveira [9]. For an overview see also [1].

#### Standard state-space representation

The controllability and observability positive real Riccati equations are:

$$\mathbf{A}_{ss}\mathcal{P} + \mathcal{P}\mathbf{A}_{ss}^* + (\mathcal{P}\mathbf{C}_{ss}^* - \mathbf{B}_{ss})\Delta(\mathcal{P}\mathbf{C}_{ss}^* - \mathbf{B}_{ss})^* = 0 \quad (7)$$

$$\mathbf{A}_{ss}^*\mathcal{Q} + \mathcal{Q}\mathbf{A}_{ss} + (\mathcal{Q}\mathbf{B}_{ss} - \mathbf{C}_{ss}^*)\Delta(\mathcal{Q}\mathbf{B}_{ss} - \mathbf{C}_{ss}^*)^* = 0 \quad (8)$$

where  $\Delta = (\mathbf{D}_{ss} + \mathbf{D}_{ss}^*)^{-1}$ .

The procedure is the same as for BT (see Sect. 3.1), except that now balancing is performed on the minimal solutions of the Riccati equations. The diagonal elements of  $\mathbf{S}$  in (4) are the *positive real singular values* of the system, which we denote by  $\pi_i$ :  $\mathbf{S} = \text{diag}(\pi_1, \dots, \pi_n)$ , where  $\pi_1 \geq \pi_2 \geq \dots \geq \pi_n$ .

#### Descriptor form representation

The corresponding algebraic Riccati equations in descriptor form are

$$\mathbf{A}_{ds}\mathcal{P}\mathbf{E}_{ds}^* + \mathbf{E}_{ds}\mathcal{P}\mathbf{A}_{ds}^* + (\mathbf{E}_{ds}\mathcal{P}\mathbf{C}_{ds}^* - \mathbf{B}_{ds})\Delta(\mathbf{E}_{ds}\mathcal{P}\mathbf{C}_{ds}^* - \mathbf{B}_{ds})^* = 0 \quad (9)$$

$$\mathbf{A}_{ds}^*\hat{\mathcal{Q}}\mathbf{E}_{ds} + \mathbf{E}_{ds}^*\hat{\mathcal{Q}}\mathbf{A}_{ds} + (\mathbf{E}_{ds}^*\hat{\mathcal{Q}}\mathbf{B}_{ds} - \mathbf{C}_{ds}^*)\Delta(\mathbf{E}_{ds}^*\hat{\mathcal{Q}}\mathbf{B}_{ds} - \mathbf{C}_{ds}^*)^* = 0 \quad (10)$$

where  $\Delta = (\mathbf{D}_{ds} + \mathbf{D}_{ds}^*)^{-1}$ . The observability gramian given by the solution of (8) is obtained via the congruence transformation  $\mathcal{Q} = \mathbf{E}_{ds}^*\hat{\mathcal{Q}}\mathbf{E}_{ds}$ .

Balancing and truncation are now performed on the solutions to (9) and (10) and the procedure follows as in 3.1.

#### Hamiltonian Riccati Balanced Truncation (PRBT-Ham)

Solutions to Riccati equations ((7),(8)) (or ((9),(10)) for MNA-similar form) can be obtained using the MATLAB function `care`. This can be applied to a system in usual state space form or in descriptor form. An alternative is to solve for  $\mathcal{P}$  and  $\hat{\mathcal{Q}}$  by means of the Hamiltonian eigenvalue problem [11]:

$$\begin{bmatrix} \mathbf{A}_{ds} - \mathbf{B}_{ds}\Delta\mathbf{C}_{ds} & -\mathbf{B}_{ds}\Delta\mathbf{B}_{ds}^* \\ \mathbf{C}_{ds}^*\Delta\mathbf{C}_{ds} & -\mathbf{A}_{ds}^* + \mathbf{C}_{ds}^*\Delta\mathbf{B}_{ds}^* \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{E}_{ds} \\ \mathbf{E}_{ds}^* \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \begin{bmatrix} \Lambda_- \\ \Lambda_+ \end{bmatrix} \quad (11)$$

where  $\Delta = (\mathbf{D}_{ds} + \mathbf{D}_{ds}^*)^{-1}$ , and  $\Lambda_-$ ,  $\Lambda_+$  are the Hamiltonian eigenvalues, with negative and positive real parts respectively (i.e. the *stable* and *antistable spectral zeros* of the system). We can partition  $\mathbf{X}$  and  $\mathbf{Y}$  according to the stable and antistable eigenvalues of the Hamiltonian into

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_- & \mathbf{X}_+ \\ \mathbf{Y}_- & \mathbf{Y}_+ \end{bmatrix}$$

The minimal solutions to (9) and (10) are given by:

$$\mathcal{P} = -\mathbf{X}_+(\mathbf{Y}_+)^{-1}\mathbf{E}_{ds}^{-*} \quad (12)$$

$$\hat{\mathcal{Q}} = -\mathbf{Y}_-(\mathbf{X}_-)^{-1}\mathbf{E}_{ds}^{-1} \quad (13)$$

and are the same as the ones resulting from the MATLAB `care` routine. The stabilizing solution (corresponding to the stable spectral zeros) is  $\hat{\mathcal{Q}}$  while  $\mathcal{P}$  is the antistabilizing solution (corresponding to the antistable spectral zeros). Both  $\hat{\mathcal{Q}}$  and  $\mathcal{P}$  are obtained from the same Hamiltonian eigenvalue computation (11). The original positive real observability gramian as solution to (8) is  $\mathcal{Q} = \mathbf{E}_{ds}^* \hat{\mathcal{Q}} \mathbf{E}_{ds}$ , so the positive real Hankel singular values are  $\pi_i = \sqrt{\lambda_i(\mathcal{P}\mathcal{Q})}$ , i.e. the diagonal elements of  $\mathbf{X}_+(\mathbf{Y}_+)^{-1}\mathbf{Y}_-(\mathbf{X}_-)^{-1}$ . We see that the positive real Hankel singular values can be computed without any inversion of  $\mathbf{E}_{ds}$ . The reduction procedure follows as in Sect. 3.1 using the computed (12) and (13).

If the system is in the usual state space form rather than in descriptor form,  $\mathbf{E}_{ds}$  in (11) is simply replaced by  $\mathbf{I}$ . The resulting solutions  $\mathcal{P}$  and  $\hat{\mathcal{Q}}$  computed as (12) and (13) respectively, are precisely the positive real gramians solving (7) and (8). They are also the same as the solutions obtained with the MATLAB `care` routine in the usual state space form. The reduction procedure follows as in Sect. 3.3.

**NOTE:** The gramians used in balanced truncation, i.e. the solutions to the Lyapunov equations ((2), (3)) (and correspondingly ((5), (6)) for descriptor form) can be obtained using (11) with  $\Delta = \mathbf{I}$ ,  $\mathbf{C} = \mathbf{0}$  (for controllability) and  $\mathbf{B} = \mathbf{0}$  (for observability).

## 4 Krylov based methods

Krylov based reduction methods exploit the use of Krylov subspace iterations to achieve system approximation by *moment matching* [1]. Three such methods are: PRIMA, the spectral zero method (SZM) and optimal  $\mathcal{H}_2$ . As outlined next, PRIMA matches  $k$  moments at zero by means of an *orthogonal projection*. SZM matches  $2k$  moments of the original system, at  $k$  stable spectral zeros and their mirror images (the corresponding  $k$  antistable spectral zeros), by means of an *oblique projection*. Finally, using an oblique projection, the optimal  $\mathcal{H}_2$  method matches  $2k$  moments of the original system at the mirror images of the  $k$  poles of the reduced system ( $2$  moments are matched at each pole). Hence an iteration is required.

### 4.1 PRIMA

For PRIMA, the moments of the transfer function  $\mathbf{H}(s) = \mathbf{L}(s\mathbf{C} - \mathbf{G})^{-1}\mathbf{B} + \mathbf{D}$  are defined as the coefficients of the Taylor expansion of  $\mathbf{H}(s)$  around  $s_0 = 0$ :  $\mathbf{H}(s) = \mathbf{M}_0 + \mathbf{M}_1s + \mathbf{M}_2s^2 + \dots$ , where



$$\mathbf{M}_0 = \mathbf{D} - \mathbf{L}\mathbf{G}^{-1}\mathbf{B} \text{ and } \mathbf{M}_k = (-1)^{(k+1)}\mathbf{L}(\mathbb{C}^{-1}\mathbf{G})^{-(k+1)}\mathbb{C}^{-1}\mathbf{B}, \text{ for } k > 0.$$

PRIMA computes a  $k^{th}$  order reduced system by matching  $k$  moments of the original system. This is achieved by computing the orthogonal projection  $\mathbf{\Pi} = \mathbf{X}_k\mathbf{X}_k^*$  such that  $\mathbf{X}_k^*\mathbb{C}^{-1}\mathbf{G}\mathbf{X}_k = \mathbf{H}_k$  with  $\mathbf{H}_k$  upper Hessenberg; the *column span* of  $\mathbf{X}_k$  is the same as the *column span* of:

$$[\mathbb{C}^{-1}\mathbf{B}, (\mathbb{C}^{-1}\mathbf{G})^{-1}\mathbb{C}^{-1}\mathbf{B}, (\mathbb{C}^{-1}\mathbf{G})^{-2}\mathbb{C}^{-1}\mathbf{B}, \dots, (\mathbb{C}^{-1}\mathbf{G})^{-(k-1)}\mathbb{C}^{-1}\mathbf{B}].$$

The procedure is as follows [7].

1. Solve  $\mathbf{G}\mathbf{R} = \mathbf{B}$  for  $\mathbf{R}$ .
2.  $(\mathbf{X}_0, \mathbf{T}) = \text{QR}(\mathbf{R})$ ; QR Factorization of  $\mathbf{R}$
3. For  $i = 1, 2, \dots, k$ 
  - Set  $\mathbf{V} = \mathbb{C}\mathbf{X}_{i-1}$
  - Solve  $\mathbf{G}\mathbf{X}_i^{(0)} = \mathbf{V}$  for  $\mathbf{X}_i^{(0)}$
  - For  $j = 1, 2, \dots, i$ 
    - $\mathbf{H} = \mathbf{X}_{i-j}^*\mathbf{X}_i^{(j-1)}$
    - $\mathbf{X}_i^{(j)} = \mathbf{X}_i^{(j-1)} - \mathbf{X}_{i-j}\mathbf{H}$
  - $(\mathbf{X}_i, \mathbf{T}) = \text{QR}(\mathbf{X}_i^{(i)})$ ; QR Factorization of  $\mathbf{X}_i^{(i)}$
4. Set  $\mathbf{X} = [\mathbf{X}_0 \ \mathbf{X}_1, \dots, \mathbf{X}_{i-1}]$  and truncate  $\mathbf{X}$  so that it has  $k$  columns only
5. Compute  $\hat{\mathbf{C}} = \mathbf{X}^*\mathbb{C}\mathbf{X}$ ,  $\hat{\mathbf{G}} = \mathbf{X}^*\mathbf{G}\mathbf{X}$ ,  $\hat{\mathbf{B}} = \mathbf{X}^*\mathbf{B}$  and  $\hat{\mathbf{L}} = \mathbf{L}\mathbf{X}$

#### 4.2 Spectral zero method (SZM)

With PRIMA, system approximation was achieved by matching  $k$  moments of the transfer function at zero. In the general case, using the *rational Krylov* approach [1], reduced systems are obtained which match moments at preassigned *interpolation points* in the complex plane. SZM is a rational Krylov reduction method, in which the interpolation points are chosen as a subset of the spectral zeros of the original system [2], [11]. This selection guarantees the stability and passivity of the reduced system [2], [11]. The spectral zeros are given by  $\Lambda$  in (11). The real spectral zeros  $s_i$  come in pairs  $(s_i, -s_i)$  while the complex spectral zeros come in quadruples of the form:

$$\begin{aligned} s_i &= \Re(s_i) + j \cdot \Im(s_i), \\ s_{i+1} &= \Re(s_i) - j \cdot \Im(s_i) = s_i^*, \\ s_{i+2} &= -\Re(s_i) + j \cdot \Im(s_i) = -s_i^*, \\ s_{i+3} &= -\Re(s_i) - j \cdot \Im(s_i) = -s_i, \end{aligned}$$

where without loss of generality, we assume  $\Re(s_i) < 0$ .

#### The usual procedure

The usual procedure for obtaining a  $k^{th}$  order reduced system with SZM is as follows.

1. Construct matrices  $\mathbf{V}$  and  $\mathbf{W}$  using  $2k$  interpolation points:

$$\mathbf{V} = [(s_1\mathbf{E}_{ds} - \mathbf{A}_{ds})^{-1}\mathbf{B}_{ds}, (s_2\mathbf{E}_{ds} - \mathbf{A}_{ds})^{-1}\mathbf{B}_{ds}, \dots, (s_k\mathbf{E}_{ds} - \mathbf{A}_{ds})^{-1}\mathbf{B}_{ds}]$$

$$\mathbf{W} = \begin{bmatrix} \mathbf{C}_{ds}(-s_1\mathbf{E}_{ds} - \mathbf{A}_{ds})^{-1} \\ \mathbf{C}_{ds}(-s_2\mathbf{E}_{ds} - \mathbf{A}_{ds})^{-1} \\ \vdots \\ \mathbf{C}_{ds}(-s_k\mathbf{E}_{ds} - \mathbf{A}_{ds})^{-1} \end{bmatrix}$$

where  $s_i, i = 1, 2, \dots, k$  are  $k$  spectral zeros that we select a priori. In our case, we selected the spectral zeros which are closest to the real axis.

2. The reduced system:  $\tilde{\mathbf{E}}_{ds} = \mathbf{W}\mathbf{E}_{ds}\mathbf{V}$ ,  $\tilde{\mathbf{A}}_{ds} = \mathbf{W}\mathbf{A}_{ds}\mathbf{V}$ ,  $\tilde{\mathbf{B}}_{ds} = \mathbf{W}\mathbf{B}_{ds}$ ,  $\tilde{\mathbf{C}}_{ds} = \mathbf{C}_{ds}\mathbf{V}$ ,  $\tilde{\mathbf{D}}_{ds} = \mathbf{D}_{ds}$ , matches the chosen  $2k$  spectral zeros of the original system,  $s_i$  and  $-s_i, i = 1, \dots, k$ .

**Note:** Since no inversion of  $\mathbf{E}_{ds}$  is involved, the SZM method is also applicable to systems with singular  $\mathbf{E}_{ds}$ . Also, the oblique projection that reduces the system is  $\mathbf{\Pi} = \bar{\mathbf{V}}\bar{\mathbf{W}}$ , where:  $\mathbf{W}\mathbf{E}_{ds}\mathbf{V} = \mathbf{L}\mathbf{U}$ ,  $\bar{\mathbf{V}} = \mathbf{V}\mathbf{U}^{-1}$  and  $\bar{\mathbf{W}} = \mathbf{L}^{-1}\mathbf{W}$ .

### Hamiltonian spectral zero method (SZM-Ham)

An alternative way to build  $\mathbf{V}$  and  $\mathbf{W}$  is without inverting the matrix  $s_i\mathbf{E}_{ds} - \mathbf{A}_{ds}$  for each spectral zero  $s_i$  we choose. This is achieved as presented in [11], by solving the Hamiltonian eigenvalue problem (11).

Once the eigenvectors and eigenvalues of the Hamiltonian are obtained, the spectral zeros which are closest to the real axis are chosen. Matrices  $\mathbf{V}$  and  $\mathbf{W}$  are now computed from the eigenvectors corresponding to the  $k$  chosen spectral zeros:

$$\mathbf{W}^* = \mathbf{Y}_k^*, \mathbf{V} = \mathbf{X}_k$$

The reduced system is  $\tilde{\mathbf{E}}_{ds} = \mathbf{W}^*\mathbf{E}_{ds}\mathbf{V}$ ,  $\tilde{\mathbf{A}}_{ds} = \mathbf{W}^*\mathbf{A}_{ds}\mathbf{V}$ ,  $\tilde{\mathbf{B}}_{ds} = \mathbf{W}^*\mathbf{B}_{ds}$ ,  $\tilde{\mathbf{C}}_{ds} = \mathbf{C}_{ds}\mathbf{V}$ ,  $\tilde{\mathbf{D}}_{ds} = \mathbf{D}_{ds}$ .

**Note:** As for SZM,  $\mathbf{E}_{ds}$  may be singular, so the SZM-Ham method also applies to the general case of descriptor systems since it involves no inversion of  $\mathbf{E}_{ds}$ . We emphasize that SZM-Ham gives the same reduced model as the usual procedure in Sect. 4.2, the difference is only in how  $\mathbf{W}$  and  $\mathbf{V}$  are computed.

### 4.3 Optimal $\mathcal{H}_2$

The optimal  $\mathcal{H}_2$  method, as the name suggests, produces reduced order models which minimize the  $\mathcal{H}_2$  norm of the error system. The problem formulation follows [4].

Given an  $n$ -dimensional single-input, single-output dynamical system in the MNA-similar form (1) (where  $\mathbb{C}$  may be singular), with transfer function  $\mathbf{H}(s) = \mathbf{L}(s\mathbb{C} - \mathbf{G})^{-1}\mathbf{B}$ , find a stable reduced system of order  $k < n$  such that its transfer function  $\tilde{\mathbf{H}}_k(s) = \mathbf{L}_k(s\mathbb{C}_k - \mathbf{G}_k)^{-1}\mathbf{B}_k$  minimizes the  $\mathcal{H}_2$  error, i.e.:

$$\tilde{\mathbf{H}}_k(s) = \arg \min_{\deg(\hat{\mathbf{H}})=k} \|\mathbf{H}(s) - \hat{\mathbf{H}}(s)\|_{\mathcal{H}_2}, \quad \|\mathbf{H}(s)\|_{\mathcal{H}_2}^2 := \frac{1}{2\pi} \int_{-\infty}^{\infty} |\mathbf{H}(j\omega)|^2 d\omega$$

The reduced order model that achieves this is constructed using the *iterative rational Krylov algorithm (IRKA)* [4]:

1. Make an initial shift selection  $\sigma_i \in \mathbb{C}, i = 1, \dots, k$

2. Construct  $\mathbf{W} = [(\sigma_1 \mathbf{C}^* - \mathbf{G}^*)^{-1} \mathbf{L}^*, \dots, (\sigma_k \mathbf{C}^* - \mathbf{G}^*)^{-1} \mathbf{L}^*]$  and  $\mathbf{V} = [(\sigma_1 \mathbf{C} - \mathbf{G})^{-1} \mathbf{B}, \dots, (\sigma_k \mathbf{C} - \mathbf{G})^{-1} \mathbf{B}]$
3. Repeat:
  - a)  $\mathbf{C}_k = \mathbf{W}^* \mathbf{C} \mathbf{V}$ ,  $\mathbf{G}_k = \mathbf{W}^* \mathbf{G} \mathbf{V}$
  - b)  $\sigma_i \leftarrow -\lambda_i(\mathbf{G}_k, \mathbf{C}_k)$  for  $i = 1, \dots, k$
  - c)  $\mathbf{W} = [(\sigma_1 \mathbf{C}^* - \mathbf{G}^*)^{-1} \mathbf{L}^*, \dots, (\sigma_k \mathbf{C}^* - \mathbf{G}^*)^{-1} \mathbf{L}^*]$
  - d)  $\mathbf{V} = [(\sigma_1 \mathbf{C} - \mathbf{G})^{-1} \mathbf{B}, \dots, (\sigma_k \mathbf{C} - \mathbf{G})^{-1} \mathbf{B}]$
 until  $\sum_{i=1}^k |\sigma_i - \bar{\sigma}_i| < \epsilon$ , where  $\sigma_i$  and  $\bar{\sigma}_i$ ,  $i = 1, \dots, k$ , are the shifts at iterations  $j$  and  $j + 1$  respectively and  $\epsilon$  is the desired convergence tolerance.
4. Project the system matrices

$$\mathbf{C}_k = \mathbf{W}^* \mathbf{C} \mathbf{V}, \mathbf{G}_k = \mathbf{W}^* \mathbf{G} \mathbf{V}, \mathbf{B}_k = \mathbf{W}^* \mathbf{B}, \mathbf{L}_k = \mathbf{L} \mathbf{V}$$

Upon convergence, the reduced order model satisfies the necessary  $\mathcal{H}_2$  optimality conditions:

$$\mathbf{H}(-\hat{\lambda}_i) = \mathbf{H}_k(-\hat{\lambda}_i), \quad \frac{d}{ds} \mathbf{H}(s) \Big|_{s=-\hat{\lambda}_i} = \frac{d}{ds} \mathbf{H}_k(s) \Big|_{s=-\hat{\lambda}_i} \quad i = 1, \dots, k,$$

where  $\hat{\lambda}_i$  are the eigenvalues of  $(\mathbf{G}_k, \mathbf{C}_k)$  (*Ritz values*). The reduced system therefore matches  $2k$  moments of the original at the mirror images of the reduced order poles. Initial shifts  $\sigma_i$  can be arbitrarily chosen and influence the convergence rate. Since this algorithm produces a *locally optimal* reduced model, some initial shifts may not lead to convergence. Future work will investigate the optimal choice of initial shifts, how they influence the convergence rate of the Ritz values and the approximation error of the resulting reduced model.

**Note:** If the initial shift selection in step 1. is a subset of the spectral zeros of the original system, reducing the system directly after step 2. makes optimal  $\mathcal{H}_2$  equivalent to SZM. The resulting reduced system will, however, not be optimal in the  $\mathcal{H}_2$  norm.  $\mathcal{H}_2$  norm optimality is guaranteed only through the iterative procedure; this however cannot guarantee passivity for the reduced model like SZM does.

Table 1: Classification of all methods

Reduction Method	Type	Iterative	Moments Matched	Projection
Balanced truncation (BT)	Gramian based	No	-	Oblique
One Gramian Method (PMTBR)	Gramian based	No	-	Orthogonal
Riccati Balanced Truncation (PRBT)	Gramian based	No	-	Oblique
Hamiltonian Riccati Balanced Truncation (PRBT-Ham)	Gramian based	No	-	Oblique
PRIMA	Krylov based	No	$k$ moments at 0	Orthogonal
Spectral Zero Method (SZM)	Krylov based	No	$2k$ spectral zeros	Oblique
Hamiltonian Spectral Zero Method (SZM-Ham)	Krylov based	No	$2k$ spectral zeros	Oblique
Optimal $\mathcal{H}_2$	Krylov based	Yes	$2k$ moments at mirror images of reduced order poles	Oblique

A classification of all methods used in our analysis is presented in Tab. 1.

## 5 Comparison of all methods: performance

A first indication of how easily these systems can be approximated is given by the Hankel singular values and the positive real Hankel singular values. Figure 2 shows a logarithmic plot of the normalized Hankel singular values and the eigenvalues of the gramians for the system associated with the circuit in Fig. 1 with  $N = 61$  sections, resulting in  $n = 242$  states. The eigenvalues of  $\mathcal{P}$  and  $\mathcal{Q}$  decay at about the same rate. The Hankel singular values and the positive real Hankel singular values (see Sect. 3.1 and 3.3) also decay at about the same rate, but not as fast as the controllability or observability gramian eigenvalues.

Since the gramian eigenvalues decay much faster than the Hankel singular values, a reduction method which balances the system by diagonalizing only one gramian seems justifiable. However, as will be shown in Sect. 6, the decay rate of the gramian eigenvalues do not provide sufficient information for the efficiency of the reduction algorithm. For example, a method that balances the system by simultaneously diagonalizing both gramians is more efficient, even though it exploits the slower decay rate of the Hankel singular values. The relative  $\mathcal{H}_\infty$  and  $\mathcal{H}_2$  norms of the associated error systems in Sect. 6 support the above statement.

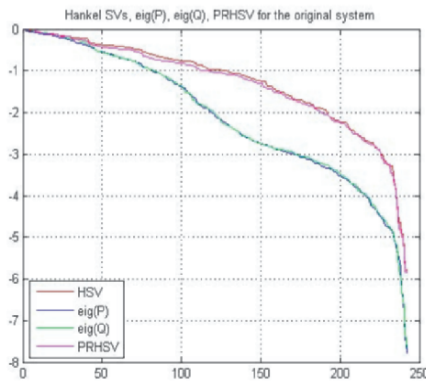


Fig. 2: Hankel singular values, positive real Hankel singular values, eigenvalues of  $\mathcal{P}$ , eigenvalues of  $\mathcal{Q}$ .

Fig. 2 also shows the trade-off between accuracy and complexity [1]. Choosing a larger order  $k$  of the reduced system by truncating the last  $n - k$  states gives a smaller approximation error. In particular, for an error of 10%, one needs to keep about half the states, namely 120. We conclude that our circuit is difficult to approximate as the decay of both Lyapunov and Riccati Hankel singular values is slow.

We further investigate whether the properties of the original system are preserved. We would like to check which of the methods used produce a reduced system which is controllable, observable, stable and passive.

*Controllability* and *observability* are equivalent to the controllability and observability gramians  $\mathcal{P}$  and  $\mathcal{Q}$  having full rank. *Stability* is equivalent to all poles lying in the left half plane. *Passivity* is ensured by the nonexistence of spectral zeros on the  $j\omega$  axis.

Starting from an initial system with  $N = 61$  sections ( $n = 242$  states), both reduced systems ( $k = 11$  and  $k = 21$ ) preserve the same characteristics.

Table 2: Preservation of Controllability, Observability, Stability and Passivity for each Reduction Method

Reduction Method	Controllable	Observable	Stable	Passive
Balanced truncation (BT)	Yes	Yes	Yes	Yes
One Gramian Method (PMTBR)	Yes	Yes	Yes	Yes
Riccati Balanced Truncation (PRBT)	Yes	Yes	Yes	Yes
Hamiltonian Riccati Balanced Truncation (PRBT-Ham)	Yes	Yes	Yes	Yes
PRIMA	No	No	Yes	Yes
Spectral Zero Method (SZM)	Yes	Yes	Yes	Yes
Hamiltonian Spectral Zero Method (SZM-Ham)	Yes	Yes	Yes	Yes
Optimal $\mathcal{H}_2$	Yes	Yes	Yes	Yes

Tab. 2 shows that all methods produce controllable and observable reduced systems, except for PRIMA. However, all reduced systems are stable. Even though only PMTBR, PRBT, PRIMA and SZM are *passivity preserving* methods for MNA representations, all the methods preserve passivity for our system.

Table 3: Relative norms of the error systems,  $k = 21, k = 11$

Reduction Method $N = 61, n = 242$	$\mathcal{H}_\infty$	$\mathcal{H}_2$	$\mathcal{H}_\infty$	$\mathcal{H}_2$
	$k = 21$	$k = 21$	$k = 11$	$k = 11$
Balanced Truncation (BT)	0.4746	0.4230	0.5409	0.4599
One Gramian (PMTBR)	0.7204	0.5284	0.5867	0.5322
Riccati Balanced Truncation (PRBT)	0.5247	0.5318	0.6486	0.7068
Hamiltonian Riccati Balanced Truncation (PRBT-Ham)	0.5247	0.5318	0.6486	0.7068
PRIMA	0.8519	0.6762	0.8147	0.8134
Spectral Zero Method (SZM)	0.6498	0.7259	0.6946	0.8392
Hamiltonian Spectral Zero Method (SZM-Ham)	0.6498	0.7259	0.6946	0.8392
<b>Optimal <math>\mathcal{H}_2</math></b>	<b>0.3554</b>	<b>0.2676</b>	<b>0.3561</b>	<b>0.2909</b>

To assess the performance of these methods, Tab. 3 collects the relative norms of the error systems, that is  $\frac{\|\Sigma_{orig} - \Sigma_k\|}{\|\Sigma_{orig}\|}$ . The surprising result in Tab. 3 is that optimal  $\mathcal{H}_2$  yields the smallest relative error both in the  $\mathcal{H}_\infty$  and  $\mathcal{H}_2$  norms; it is superior for instance to BT, which is usually considered as the overall best approximation method. From Tab. 3, it is also evident that relative errors for PRBT and PRBT-Ham are identical. The same holds for SZM and SZM-Ham. This shows that SZM and PRBT are equivalent to their Hamiltonian counterparts respectively. Reduction by diagonalizing one gramian method yields the best approximant out of all the passivity preserving methods. PRIMA gives a reduced system which approximates the original one very well for low frequencies (much better than any other method). This is a consequence of the fact that the expansion point in the Arnoldi algorithm is the origin.

The  $\mathcal{H}_\infty$  norm of a stable system is the maximum singular value of the transfer function or the value of the highest peak in the frequency response plot. Notice that, in case of one gramian and PRIMA, a larger dimension of the reduced system does

not yield a smaller  $\mathcal{H}_\infty$  relative norm of the error system: the value for the reduced system of size  $k = 21$  is larger than the value obtained for the reduced system of size  $k = 11$ . This could be explained by the fact that a higher dimension of the reduced system yields peaks of higher amplitude. The  $\mathcal{H}_2$  norm of a stable SISO system with the  $\mathbf{D}$  term equal to 0 is a measure of the area frequency response over the entire frequency range. Notice that there is no norm which captures all aspects of the reduced system. The ones used here are the most popular.

Table 4: Elapsed Times

Reduction Method $N = 61, n = 242$	Elapsed Time (s) $k = 21$	Elapsed Time(s) $k = 11$
Balanced Truncation (BT)	1.915538	1.956464
One Gramian (PMTBR)	0.972908	0.974607
Riccati Balanced Truncation (PRBT)	65.47625	66.20935
<i>Hamiltonian Riccati Balanced Truncation (PRBT-Ham)</i>	39.64582	40.97727
<b>PRIMA</b>	<b>0.355047</b>	<b>0.189727</b>
Spectral Zero Method (SZM)	12.208282	8.749104
<i>Hamiltonian Spectral Zero Method (SZM-Ham)</i>	5.017118	4.746759
Optimal $\mathcal{H}_2$	136.5	111.35

Elapsed times are useful for comparing the computational cost of each method versus the quality of the resulting reduced system. The computational times in Tab. 4 were obtained on a Pentium M at 1.3Ghz with 768MB RAM. The most expensive method is optimal  $\mathcal{H}_2$ ; it requires a certain number of iteration steps to converge, depending on the initial shift selection. PRBT is also expensive, when implemented using MATLAB's `care` function for obtaining the positive definite solutions to the algebraic Riccati equations. On the other hand, PRIMA is the most computationally efficient, having the complexity of an iterative Arnoldi algorithm. This is more computationally efficient than performing eigenvalue decompositions, singular value decompositions or solving the Lyapunov or algebraic Riccati equations, which are needed in the other reduction methods.

Another aspect worth noticing in Tab. 4 is that, indeed, for the spectral zero method, computing the projectors from the eigenvectors of the Hamiltonian matrix (SZM-Ham) requires about half the time needed to compute the projectors by inverting the matrix  $s_i E_{ds} - A_{ds}$  for each spectral zero  $s_i$  we choose (SZM). A similar performance improvement is achieved when the Riccati solutions in PRBT were computed with the Hamiltonian eigenvalue problem (PRBT-Ham) rather than with the MATLAB `care` routine (PRBT). Considering that relative error norms for PRBT and PRBT-Ham are identical (the same holds for SZM and SZM-Ham), we conclude that performing PRBT (and SZM) via the Hamiltonian approach is more efficient.

Also, from Tab. 4 we notice the small difference between elapsed times for the two reduced dimensions ( $k = 21$  and  $k = 11$ ), since most of the computational effort is used in computing the projectors, not in obtaining the reduced systems themselves.

## 6 Comparison of all methods: plots

We first provide pairwise comparisons of error systems resulting from applying each reduction procedure on the circuit in Fig. 1. Next, the preservation of stability and passivity is shown in the distributions of poles and spectral zeros for the original and each reduced system (Figs. 13-18). Figures for methods PRBT-Ham and SZM-Ham are omitted, because they are identical to figures for PRBT and SZM respectively. The original system has  $n = 242$  states (resulting from  $N = 61$  sections) and we reduce it to dimension  $k = 21$ .

### 6.1 Error systems

Fig. 3 shows the frequency response of the original system together with all reduced systems. Fig. 4 shows the frequency response of the systems obtained by taking the difference between the original and each of the reduced systems.

Comparing the errors for BT and PRBT in Fig. 5 shows that the first one is a better approximant of the original system since the error plot is almost always below the error plot for PRBT. However, we notice that the shapes of the plots are almost the same, with the second one shifted up by a few decibels.

Comparing the error systems for PRBT and PRIMA in Fig. 6 shows that, even though PRIMA gives small error for low frequencies, PRBT performs better in the middle range, where the response is harder to capture because of the large number of oscillations.

Comparing BT with PMTBR essentially means comparing diagonalization of only one gramian versus simultaneous diagonalization of both controllability and observability gramians. Fig. 7 shows that BT gives a smaller approximation error. Diagonalizing both gramians therefore leads to a better approximation than diagonalizing only one gramian. This is because after simultaneous diagonalization, truncation is performed on states that are equally difficult to reach and to observe.

From Fig. 8, it is clear that the spectral zero method performs comparably to balanced truncation. The advantage of the spectral zero method over balanced truncation is that it guarantees the passivity of the reduced system, irrespective of whether the system is in MNA-similar form. As shown in figures 9 and 10, the spectral zero method also performs similarly to the other two passivity preserving methods, one gramian and PRIMA.

Inspecting Fig. 22, we emphasize that with randomly chosen initial shifts, optimal  $\mathcal{H}_2$  yields an approximation error smaller than BT. Fig. 11 shows that optimal  $\mathcal{H}_2$  gives a smaller approximation error than PRIMA, except for low frequencies, as expected. Optimal  $\mathcal{H}_2$  also provides a better approximant than SZM, as seen from Fig. 12.

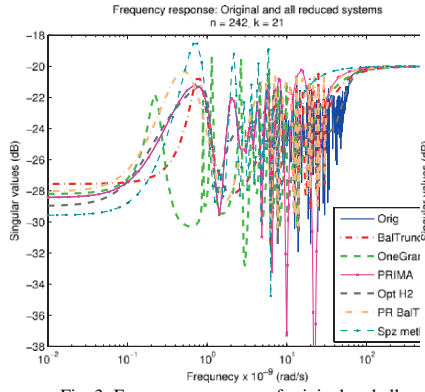


Fig. 3: Frequency response of original and all reduced systems

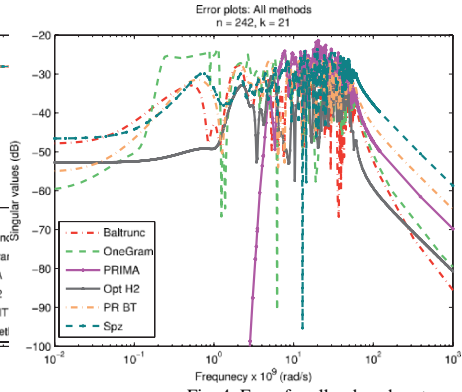


Fig. 4: Error for all reduced systems

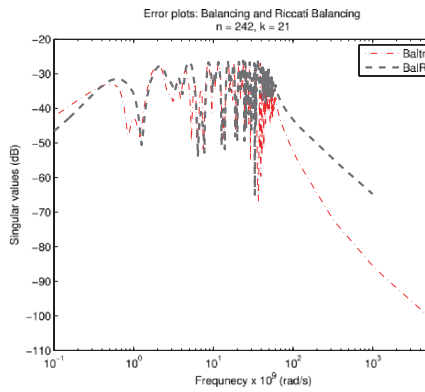


Fig. 5: Error systems: Balanced truncation and Positive Real Balanced Truncation

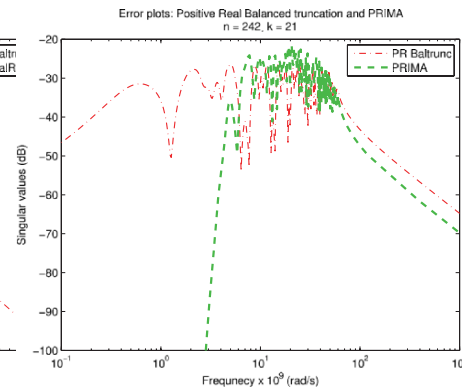


Fig. 6: Error systems: Positive Real Balanced Truncation and PRIMA

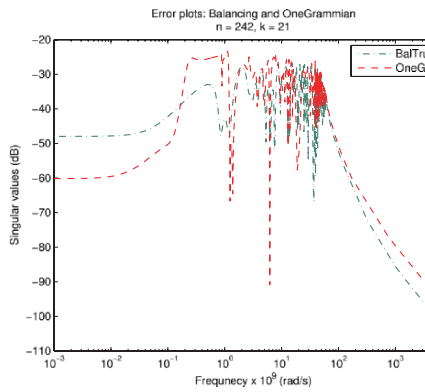


Fig. 7: Error systems: Balanced truncation and One Gramian

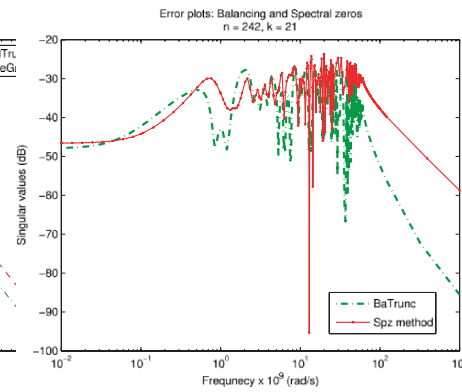


Fig. 8: Error systems: Balanced truncation and Projection using spectral zeros



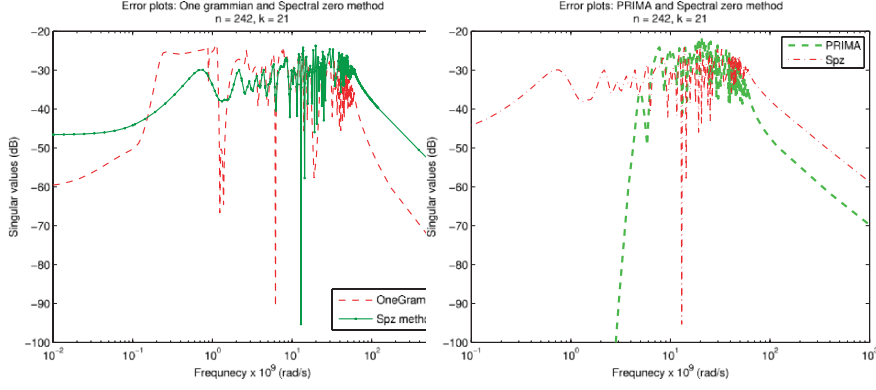


Fig. 9: Error systems: One Gramian and Projection using spectral zeros

Fig. 10: Error systems: PRIMA and Projection using spectral zeros

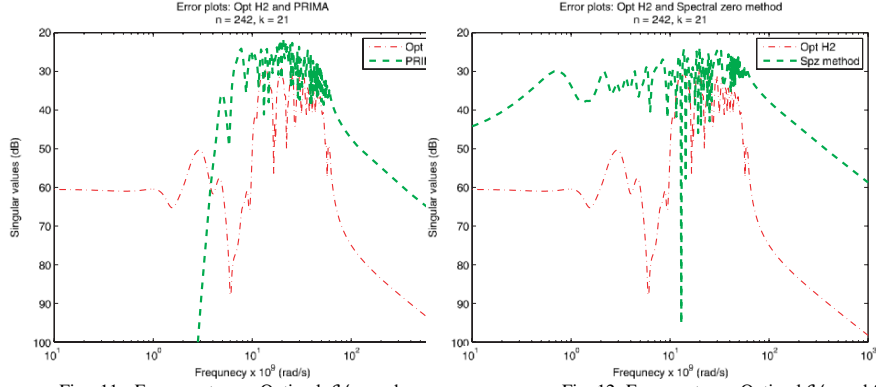

 Fig. 11: Error systems: Optimal  $\mathcal{H}_2$  and PRIMA

 Fig. 12: Error systems: Optimal  $\mathcal{H}_2$  and Spz method

## 6.2 Poles and spectral zeros of reduced systems

The following figures show the location of poles and spectral zeros of the original, stable and passive system together with the poles and spectral zeros of each reduced system. As already indicated in Tab. 2, all reduction methods yielded stable and passive reduced order systems. This is shown in Figs. 13-18, where the poles of all reduced system lie in the left half plane and the spectral zeros of all reduced systems are located away from the  $j\omega$  axis.

No spectral zero matching or similarity in pole distribution occurs for reduced models obtained with BT, one gramian or PRBT, as seen from Figs. 13, 14 and 15. In particular, Fig. 13 shows that poles and spectral zeros resulting from BT are scattered, while in Fig. 14, the poles and spectral zeros from one gramian are clustered close to the  $j\omega$  axis. Spectral zeros resulting from PRBT are aligned along some of the spectral zeros of the original system, as seen in Fig. 15.

However, comparing Figs. 16, 17, and 18, we identify a similarity between the distribution of spectral zeros and poles resulting from PRIMA, SZM, and optimal  $\mathcal{H}_2$  respectively. The poles of these reduced systems follow a pattern, being located close to the real axis. Furthermore, the spectral zeros resulting from optimal  $\mathcal{H}_2$ , match some of the spectral zeros of the original system, similarly to the spectral zero method and PRIMA.

Fig. 16 shows that PRIMA preserves some of the poles as well as some of the spectral zeros of the original system. This is the only method for which some poles of the reduced system are close to the poles of the original system.

In Fig. 18, as expected, the spectral zeros of the reduced system match the chosen spectral zeros of the original system with smaller imaginary parts.

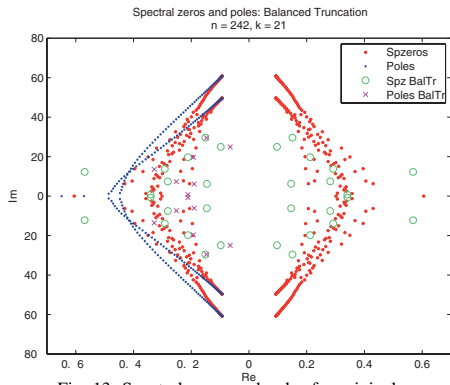


Fig. 13: Spectral zeros and poles for original system and reduced with Balanced truncation

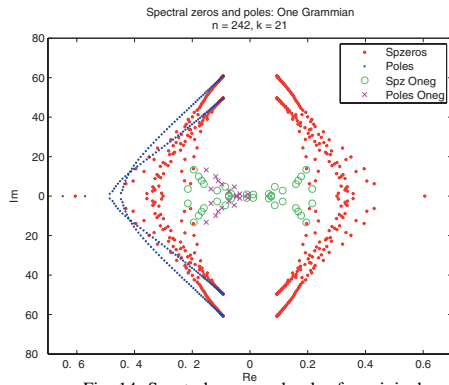


Fig. 14: Spectral zeros and poles for original system and reduced with One Gramian

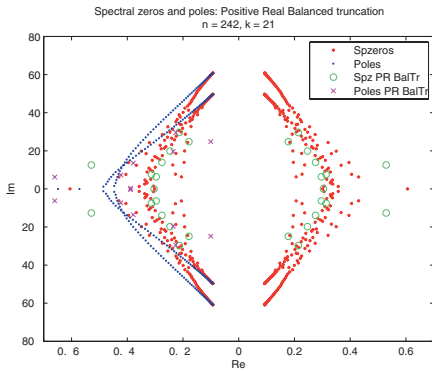


Fig. 15: Spectral zeros and poles for original system and reduced with Riccati Balanced truncation

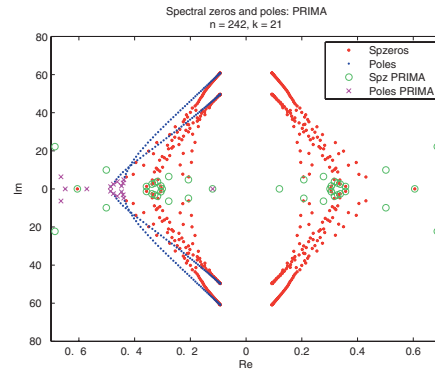


Fig. 16: Spectral zeros and poles for original system and reduced with PRIMA

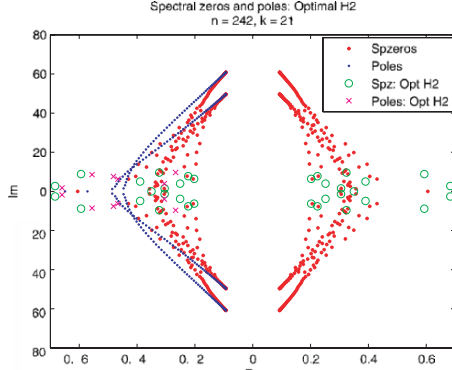


Fig. 17: Random initial shifts: Poles and Spectral zeros of original system and reduced with Optimal  $\mathcal{H}_2$

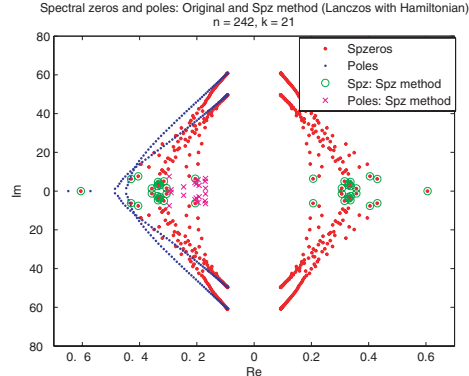


Fig. 18: Spectral zeros and poles of original and reduced with projection using spectral zero selection

## 7 Optimal $\mathcal{H}_2$ results: errors, convergence and initial shifts

We compare reduced order models obtained with balanced truncation and optimal  $\mathcal{H}_2$ , since these methods yield the smallest relative  $\mathcal{H}_2$  error norms. We approximate the initial order  $n = 242$  system with reduced models of orders  $k = 21$  and  $k = 11$ . The corresponding relative error norms are found in Tab. 3, Sect. 5. From these results, it is clear that optimal  $\mathcal{H}_2$  is the overall best method with respect to both relative error norms:  $\mathcal{H}_\infty$  and  $\mathcal{H}_2$ .

The selection of initial shifts can be determined to influence the convergence rate, the approximation error and the distribution of poles and spectral zeros for the reduced system. Results were obtained for two different sets of initial shifts: the poles resulting from BT and randomly generated complex shifts. The table below summarizes the number of iterations needed for IRKA to converge, with a threshold difference of  $10^{-4}$  between successively generated shifts. Choosing random complex numbers as initial shifts yielded convergence which was almost twice as fast.

Initial Shift Choice	Red. order: $k = 21$	Red. order: $k = 11$
Poles from BT	49 steps	85 steps
<b>Random complex</b>	<b>28 steps</b>	<b>47 steps</b>

1.  $n = 242$ ,  $k = 21$

Figs. 19-22 show that the reduced system obtained with optimal  $\mathcal{H}_2$  using randomly generated shifts approximates the original system more accurately than when the initial shifts are the poles of the reduced system obtained from balanced truncation. The error systems in Fig. 22 show that when initial shifts are randomly chosen, optimal  $\mathcal{H}_2$  yields a better approximant than BT. This is not the case when initial shifts are the poles from BT, as seen from the error systems in Fig. 21. Additionally, when the initial shifts are randomly chosen, some of the spectral zeros of the reduced system closely match spectral zeros of the original system, as shown in figure 24. This appealing behavior is not present in Figure 23, where the initial shifts are the poles resulting from BT: no spectral zeros are matched and the corresponding approximation error is larger.

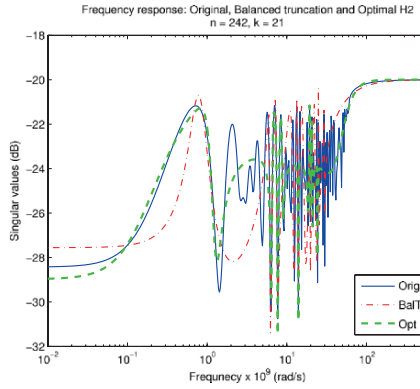


Fig. 19: BT poles as initial shifts: Frequency response of original and reduced systems

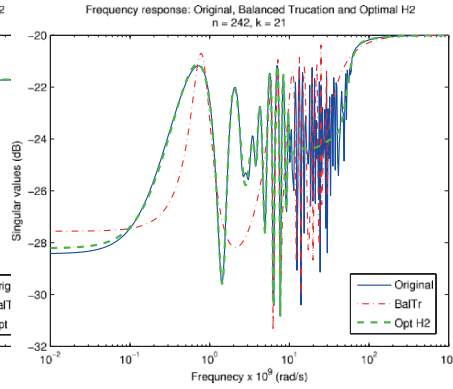


Fig. 20: Random initial shifts: Frequency response of original and reduced systems

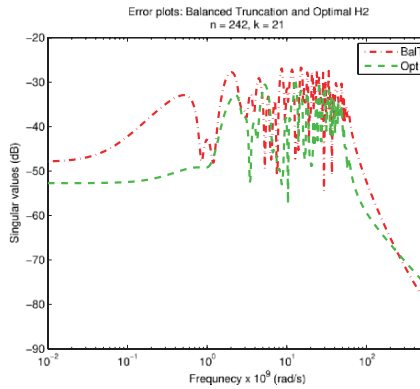


Fig. 21: BT poles as initial shifts: Error systems

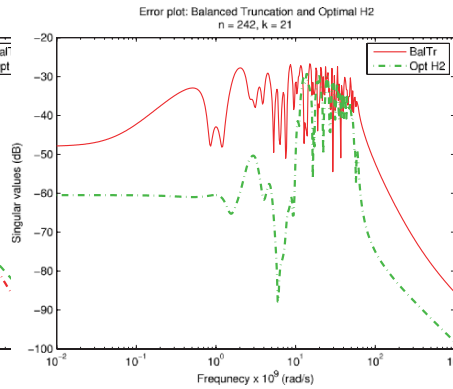


Fig. 22: Random initial shifts: Error systems

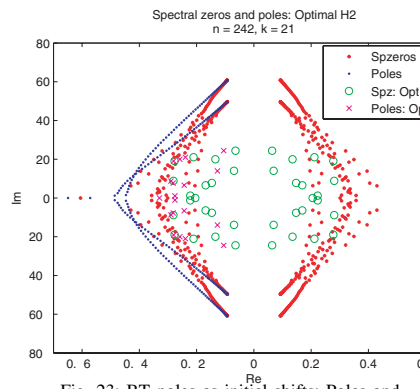


Fig. 23: BT poles as initial shifts: Poles and Spectral zeros of original system and reduced with Optimal  $\mathcal{H}_2$

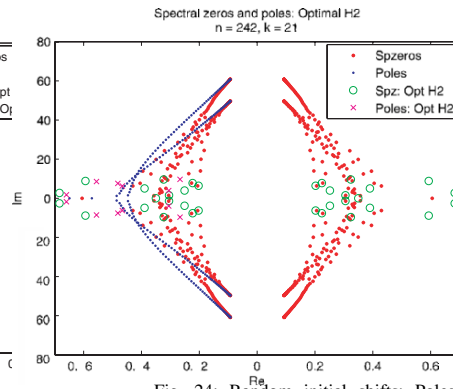


Fig. 24: Random initial shifts: Poles and Spectral zeros of original system and reduced with Optimal  $\mathcal{H}_2$

2.  $n = 1002$ ,  $k = 71$

Promising results for optimal  $\mathcal{H}_2$  were also obtained on a system of dimension  $n = 1002$  ( $N = 251$ ). As shown in Figs. 25 and 26, the reduced model of dimension  $k = 71$  clearly approximates the original more accurately than the reduced model obtained with BT. Again, optimal  $\mathcal{H}_2$  is superior to BT with respect to both relative error norms,  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$ .

Reduction Method: $n = 1002$ , $k = 71$	$\mathcal{H}_\infty$	$\mathcal{H}_2$
Balanced Truncation	0.1488	0.1124
<b>Optimal <math>\mathcal{H}_2</math></b>	<b>0.09467</b>	<b>0.06408</b>

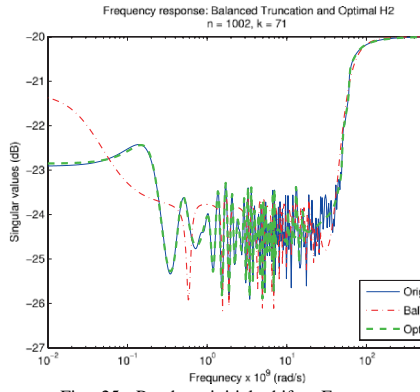


Fig. 25: Random initial shifts: Frequency response: original, reduced with balanced truncation and Optimal  $\mathcal{H}_2$

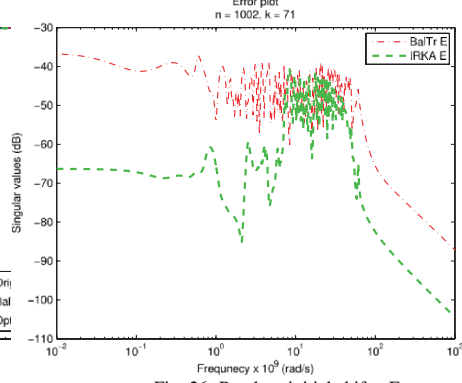


Fig. 26: Random initial shifts: Error systems for balanced truncation and Optimal  $\mathcal{H}_2$

## 8 Conclusion and further research

This paper compares several model reduction methods used in circuit simulation, in particular for systems in invertible descriptor form. The methods are grouped in two categories, gramian and Krylov based respectively. Theoretical considerations for all methods are outlined, and their performance is evaluated by reducing the dynamical system of a coupled transmission line. Approximation error and computational cost analysis for each method shows that while some methods yield better reduced systems, others are computationally cheaper. Furthermore, not all methods that yield small relative approximation errors preserve important properties of the original system, such as controllability, observability or passivity.

Optimal  $\mathcal{H}_2$  is the overall best in terms of both relative  $\mathcal{H}_\infty$  and  $\mathcal{H}_2$  norms, but requires the highest computational complexity and cannot guarantee passivity for the reduced system. It can also be applied to the general class of descriptor systems, where  $\mathbb{C}$  in (1) may be singular. Further research is needed for determining how the choice of initial shifts in optimal  $\mathcal{H}_2$  influences the distribution of poles and spectral zeros of the reduced system and the convergence rate.

Among passivity preserving methods, SZM provides the best trade-off between approximation error, computational cost, and preservation of stability and passivity. Furthermore, SZM can be applied to the general class of descriptor systems. However, determining the optimal selection of spectral zeros in SZM is an open problem.

SZM overcomes the limitations of PRIMA: controllability and observability loss for the reduced system (due to possible pole-zero cancelations) and larger approximation error for high frequencies. PRIMA however provides the best fit for low frequencies from all methods considered. SZM is also computationally cheaper than PRBT. PRBT on the other hand yields an approximation error comparable to BT and has the benefit of preserving passivity.

Since our analysis is conducted on a SISO circuit with invertible descriptor form ( $\mathbb{C}$  in (1) was invertible), a further step would be to reproduce these results for a system in general descriptor form, where  $\mathbb{C}$  is singular, partly using the work in [5]. Applying these reduction methods on a MIMO network is currently under investigation.

## References

- [1] Antoulas, A.C.: *Approximation of large-scale dynamical systems*, SIAM, Philadelphia (2005)
- [2] Antoulas, A.C.: *A new result on passivity preserving model reduction*, Systems and Control Letters, vol. **54**, 361-374 (2005)
- [3] Benner, P., Quintana-Orti, E.S.: *Solving stable generalized Lyapunov equations with the matrix sign function*, Numerical Algorithms, vol. **20**, 75-100, (1999)
- [4] Gugercin, S., Antoulas, A.C.: Beattie, C.A.: *A rational Krylov iteration for optimal  $H_2$  model reduction*, Proceedings of the 17th Int. Symposium on Mathematical Theory of Networks and Systems, MTNS06, Kyoto, 1665-1667 (July, 2006)
- [5] Mehrmann, V., Stykel, T.: *Balanced truncation model reduction for large-scale systems in descriptor form*, Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann and D. Sorensen, Edtrs., LNCSE vol. **45**, Springer Verlag, Heidelberg (2005)
- [6] Ober, R.J.: *Balanced parametrization of classes of linear systems*, SIAM Journal of Control and Optimization, vol. **29**, 1251-1287 (1991)
- [7] Odabasioglu, A., Celik M.: *PRIMA. Passive reduced-order interconnect macromodelling algorithm*, IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems vol. **17**, 8, 645-654 (1998)
- [8] Palenius, T.: *Efficient Time-Domain Simulation of Interconnects Characterized by Large RLC Circuits or Tabulated S Parameters*, Licentiate Thesis (2004)
- [9] Phillips, J.R.; Daniel, L.; Silveira, L.M.: *Guaranteed passive balancing transformations for model order reduction*, IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems vol. **22**, 8, 1027 - 1041 (2003)
- [10] Phillips, J.R., Silveira, L. M.: *Poor man's TBR: a simple model reduction scheme*, IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems vol. **24**, 1, 43-55 (2005)
- [11] Sorensen, D.C.: *Passivity preserving model reduction via interpolation of spectral zeros*, Systems and Control Letters, vol. **54**, 347-360 (2005)