Adv Comput Math (2020) 46:40
https://doi.org/10.1007/s10444-020-09750-w

Hankel-norm approximation of large-scale descriptor systems

Peter Benner[1,2] · **Steffen W. R. Werner**[1]

Received: 1 August 2018 / Accepted: 21 November 2019 /

Abstract
Hankel-norm approximation is a model reduction method for linear time-invariant systems, which provides the best approximation in the Hankel semi-norm. In this paper, the computation of the optimal Hankel-norm approximation is generalized to the case of linear time-invariant continuous-time descriptor systems. A new algebraic characterization of all-pass descriptor systems is developed and used to construct an efficient algorithm by refining the generalized balanced truncation square root method. For a wide practical usage, adaptations of the introduced algorithm towards stable computations and sparse systems are suggested, as well as an approach for a projection-free algorithm. To show the approximation behavior of the introduced method, numerical examples are presented.

Keywords Model order reduction · Hankel singular values · Linear systems · Differential-algebraic equations

Mathematics Subject Classification (2010) 93B40 · 93A15 · 93B11

Communicated by: Anthony Nouy

This article belongs to the Topical Collection: *Model reduction of parametrized Systems*
Guest Editors: Anthony Nouy, Peter Benner, Mario Ohlberger, Gianluigi Rozza, Karsten Urban and Karen Willcox

✉ Steffen W. R. Werner
werner@mpi-magdeburg.mpg.de

Peter Benner
benner@mpi-magdeburg.mpg.de

[1] Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106, Magdeburg, Germany

[2] Faculty of Mathematics, Otto von Guericke University, Universitätsplatz 2, 39106, Magdeburg, Germany

1 Introduction

Many different real-world applications, like chemical processes, electrical circuits and networks, or computational fluid dynamics, naturally lead to models, described by systems of differential-algebraic equations. Since experiments can be very costly, time-consuming, and expensive, these models are used for simulations and the design of controllers. The modeling process often results in linear time-invariant continuous-time descriptor systems of the form

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{1}$$

with $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$. Here, $u(t) \in \mathbb{R}^m$ are the inputs of the system, which influence the generalized states $x(t) \in \mathbb{R}^n$ to get the desired outputs $y(t) \in \mathbb{R}^p$. Throughout this paper, it is assumed that the matrix pencil $\lambda E - A$ is regular, i.e., there exists at least one $\lambda \in \mathbb{C}$ such that $\det(\lambda E - A) \neq 0$. In this case, and with the initial condition $Ex(0) = 0$, the input-output behavior of the system (1) in the frequency domain can be described via the system's transfer function

$$G(s) = C(sE - A)^{-1}B + D. \tag{2}$$

The quintuple (E, A, B, C, D), consisting of the matrices from (1), defines a realization of the system (1) and its transfer function (2). Usually, the numbers of inputs and outputs are very small in contrast to the number of differential-algebraic equations and generalized states n, which quickly enlarges due to different reasons, e.g., the model shall provide a required accuracy. Because of that, the usage of complete models often reaches the limits of computational resources like memory and computation time. Since the acquired data for the model usually contain a huge amount of redundancies, it is possible to approximate the original model by a new system with a much smaller order. The task of model reduction is to construct a reduced-order descriptor system

$$\begin{aligned} \hat{E}\dot{\hat{x}}(t) &= \hat{A}\hat{x}(t) + \hat{B}u(t), \\ \hat{y}(t) &= \hat{C}\hat{x}(t) + \hat{D}u(t), \end{aligned} \tag{3}$$

of order $r \ll n$, such that the input-output behavior of the original system (1) is approximated.

Many model reduction techniques were originally developed for the standard system case, where the descriptor term E is the identity matrix I_n (or at least nonsingular). But in recent years, quite a few of those methods have been extended to the case of descriptor systems with singular E matrices. There are different approaches for the construction of (3), e.g., matrix equations can be used to determine a measure for truncatable states [6], or the transfer function can be approximated by rational interpolation [14]. A special technique of model reduction is the computation of the optimal Hankel-norm approximation (HNA). This technique actually provides a best approximation in the Hankel semi-norm. Based on the work of Adamjan, Arov, and Krein about the approximation of Hankel matrices [1], an algorithm for the computation of the HNA for standard systems was introduced by Glover in [12]. Beside minimizing the Hankel semi-norm, by construction, the HNA is very close to the

solution of the best approximation problem in the $\mathcal{H}_\infty$-norm, which usually also gives a smaller $\mathcal{H}_\infty$ error of the reduced-order models than balanced truncation.

A generalization of the HNA to the descriptor system case was already described by Cao, Saltik, and Weiland in [10]. The authors are using the Weierstrass canonical form for an explicit construction of reduced decoupled subsystems. The main problem of this method is the computation of the Weierstrass canonical form, which is numerically costly and unstable. Also, additional conditions, like C-controllability and C-observability of the system, have to be assumed.

In this paper, a new efficient algorithm for the computation of the generalized Hankel-norm approximation (GHNA) will be proposed. Our main contributions are twofold:

1. We generalize the concept of all-pass transfer functions to descriptor systems (Theorem 1).
2. We derive new and reliable numerical implementations of the GHNA that also allow the application of the Hankel-norm approximation method to large-scale problems with sparse coefficient matrices as they arise, e.g., from systems with dynamics described by semi-discretized unsteady partial differential equations.

In Section 2, the mathematical background of linear descriptor systems is recalled. Then, the HNA method for the standard system case is introduced in the first part of Section 3. Afterwards, the generalized balanced truncation is reviewed and used for the construction of the new GHNA method. The numerical difficulties and adjustments are discussed in Section 4 for usable implementations of the method. Two different implementations of the method are then tested on numerical examples in Section 5. In Section 6, the conclusions of this paper can be found.

2 Mathematical basics

For regular matrix pencils $\lambda E - A$, the Weierstrass canonical form always exists: there are invertible matrices $W, T \in \mathbb{C}^{n\times n}$ such that

$$W(\lambda E - A)T = \lambda \begin{bmatrix} I_{n_f} & 0 \\ 0 & N \end{bmatrix} - \begin{bmatrix} J & 0 \\ 0 & I_{n_\infty} \end{bmatrix}, \tag{4}$$

where J and N are both in Jordan canonical form, J is regular, and N is nilpotent with index ν; see [20]. The numbers n_f and n_∞ are the dimensions of the deflating subspaces corresponding to the finite and infinite eigenvalues of $\lambda E - A$, respectively. Then, the spectral projectors onto the left and right deflating subspaces corresponding to the finite eigenvalues of the matrix pencil $\lambda E - A$ are given by

$$P_\ell = W^{-1} \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} W \quad \text{and} \quad P_r = T \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} T^{-1}, \tag{5}$$

with W and T from the Weierstrass canonical form (4); see, e.g., [21].

Throughout the paper, we assume the c-stability of the matrix pencil $\lambda E - A$, i.e., the matrix pencil $\lambda E - A$ is regular and all finite eigenvalues of $\lambda E - A$ lie in the open left half-plane. In this case, the proper controllability and observability Gramians are

defined as the unique, positive semidefinite solutions of the projected generalized continuous-time Lyapunov equations

$$E\mathcal{G}_{pc}A^{\mathsf{T}} + A\mathcal{G}_{pc}E^{\mathsf{T}} + P_\ell B B^{\mathsf{T}} P_\ell^{\mathsf{T}} = 0, \quad \mathcal{G}_{pc} = P_r \mathcal{G}_{pc} P_r^{\mathsf{T}}, \tag{6}$$

$$E^{\mathsf{T}}\mathcal{G}_{po}A + A^{\mathsf{T}}\mathcal{G}_{po}E + P_r^{\mathsf{T}} C^{\mathsf{T}} C P_r = 0, \quad \mathcal{G}_{po} = P_\ell^{\mathsf{T}} \mathcal{G}_{po} P_\ell, \tag{7}$$

with P_ℓ and P_r the spectral projectors corresponding to the finite eigenvalues (5); see [21]. Furthermore, the improper controllability and observability Gramians are given as the unique, positive semidefinite solutions of the projected generalized discrete-time Lyapunov equations (Stein equations)

$$A\mathcal{G}_{ic}A^{\mathsf{T}} - E\mathcal{G}_{ic}E^{\mathsf{T}} - Q_\ell B B^{\mathsf{T}} Q_\ell^{\mathsf{T}} = 0, \quad \mathcal{G}_{ic} = Q_r \mathcal{G}_{ic} Q_r^{\mathsf{T}}, \tag{8}$$

$$A^{\mathsf{T}}\mathcal{G}_{io}A - E^{\mathsf{T}}\mathcal{G}_{io}E - Q_r^{\mathsf{T}} C^{\mathsf{T}} C Q_r = 0, \quad \mathcal{G}_{io} = Q_\ell^{\mathsf{T}} \mathcal{G}_{io} Q_\ell, \tag{9}$$

with $Q_\ell = I_n - P_\ell$ and $Q_r = I_n - P_r$, the spectral projectors onto the left and right deflating subspaces corresponding to the infinite eigenvalues of the matrix pencil $\lambda E - A$; see [21].

Using the system Gramians, the set of Hankel singular values is defined in the following; see [16].

Definition 1 The square roots of the n_f largest eigenvalues of $\mathcal{G}_{pc}E^{\mathsf{T}}\mathcal{G}_{po}E$ denoted by $\varsigma_1 \geq \varsigma_2 \geq \cdots \geq \varsigma_{n_f}$ are the proper Hankel singular values of (1). The square roots of the n_∞ largest eigenvalues of $\mathcal{G}_{ic}A^{\mathsf{T}}\mathcal{G}_{io}A$ denoted by $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_{n_\infty}$ are the improper Hankel singular values of (1).

In case of a non-singular descriptor term E, the proper Hankel singular values are the classical Hankel singular values of the system. Therefore, an equivalent energy interpretation of the proper Hankel singular values exists, which proposes the truncation of states corresponding to small proper Hankel singular values, which are difficult to control and observe. Unfortunately, this does not hold for the improper Hankel singular values. Those correspond to the algebraic constraints of the system and quantify which constraints are necessary to characterize the system's behavior and which are not, i.e., states corresponding to the non-zero improper Hankel singular values describe necessary constraints and should not be truncated.

There exist diverse concepts of controllability and observability for descriptor systems. For this paper, we restrict ourselves to the following ones; see, e.g., [21].

Definition 2 System (1) is called:

1. R-controllable if $\operatorname{rank}\begin{bmatrix} \lambda E - A, & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C}$.
2. C-controllable if the system is R-controllable and $\operatorname{rank}\begin{bmatrix} E, & B \end{bmatrix} = n$.
3. R-observable if $\operatorname{rank}\begin{bmatrix} \lambda E^{\mathsf{T}} - A^{\mathsf{T}}, & C^{\mathsf{T}} \end{bmatrix} = n$ for all $\lambda \in \mathbb{C}$.
4. C-observable if the system is R-observable and $\operatorname{rank}\begin{bmatrix} E^{\mathsf{T}}, & C^{\mathsf{T}} \end{bmatrix} = n$.

The relation between these controllability, observability notions and the system Gramians is given in [22, Theorem 2.3]. Especially, all proper Hankel singular values are non-zero if and only if the system is R-controllable and R-observable.

The mapping from past inputs $u_- : (-\infty, 0] \to \mathbb{R}^m$ to future outputs $y_+ : (0, +\infty] \to \mathbb{R}^p$ is described by the Hankel operator $y_+ = \mathcal{H}u_-$. A generalization of this operator to the case of descriptor systems can be found in [10]. The measure of the influence of past inputs on future outputs in the $\mathcal{L}_2$-norm leads to the definition of the Hankel semi-norm for descriptor systems.

Definition 3 The Hankel semi-norm of a system G is given by

$$\|G\|_H = \sup_{u_- \in \mathcal{W}_2^{\nu-1}(-\infty,0]} \frac{\|y_+\|_{\mathcal{L}_2}}{\|u_-\|_{\mathcal{L}_2}}, \tag{10}$$

where $\mathcal{W}_2^{\nu-1}(-\infty, 0]$ denotes the Sobolev space of $\nu - 1$ times weakly differentiable functions w.r.t. the $\mathcal{L}_2$ inner product on the interval $(-\infty, 0]$ and $\|.\|_{\mathcal{L}_2}$ is the $\mathcal{L}_2$-norm.

It should be noted that the Hankel semi-norm is independent of the feed-through term D, i.e., even for $D \neq 0$ the semi-norm can be 0. In case of an invertible descriptor term E, the Hankel semi-norm (10) simplifies to

$$\|G\|_H = \varsigma_{\max}(G),$$

where $\varsigma_{\max}(G)$ is the largest Hankel singular value of the system G.

3 Generalized Hankel-norm approximation

3.1 Algorithm for standard systems

First, the algorithm for the standard system case, introduced by Glover in [12], is reviewed. Therefore, a balanced minimal realization of the given standard system $(I_{n_{\min}}, A, B, C, D)$ is assumed, where $n_{\min}$ is the McMillan degree of the system, i.e., the order of its minimal realization. The computation is usually done by the balanced truncation square root method. Since the resulting system is balanced and minimal, the system Gramians are equal and diagonal

$$\mathcal{G}_{pc} = \mathcal{G}_{po} = \operatorname{diag}(\varsigma_1, \varsigma_2, \ldots, \varsigma_{n_{\min}}),$$

with $\varsigma_1, \ldots, \varsigma_{n_{\min}}$ all non-zero Hankel singular values of the system. Next, the system is partitioned by the order r such that

$$\varsigma_1 \geq \ldots \geq \varsigma_r > \varsigma_{r+1} = \ldots = \varsigma_{r+k} > \varsigma_{r+k+1} \geq \ldots \geq \varsigma_{n_{\min}},$$

with $k \geq 1$ being the multiplicity of the $(r+1)$-st Hankel singular value. The Gramians are reordered to separate the block with the $(r+1)$-st Hankel singular value as

$$\check{\mathcal{G}}_{pc} = \check{\mathcal{G}}_{po} = \begin{bmatrix} \check{\Sigma} & \\ & \varsigma_{r+1} I_k \end{bmatrix}, \tag{11}$$

with $\check{\Sigma} = \operatorname{diag}(\varsigma_1, \ldots, \varsigma_r, \varsigma_{r+k+1}, \ldots, \varsigma_{n_{\min}})$. Accordingly to (11), the remaining system matrices have to be permuted and partitioned

$$\check{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \check{B} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad \check{C} = \begin{bmatrix} C_1, & C_2 \end{bmatrix}.$$

Then, the partitioned system is transformed by the following formulas

$$\begin{aligned} \tilde{A} &= \Gamma^{-1}(\varsigma_{r+1}^2 A_{11}^{\mathsf{T}} + \check{\Sigma} A_{11} \check{\Sigma} + \varsigma_{r+1} C_1^{\mathsf{T}} U B_1^{\mathsf{T}}), \\ \tilde{B} &= \Gamma^{-1}(\check{\Sigma} B_1 - \varsigma_{r+1} C_1^{\mathsf{T}} U), \\ \tilde{C} &= C_1 \check{\Sigma} - \varsigma_{r+1} U B_1^{\mathsf{T}}, \\ \tilde{D} &= D + \varsigma_{r+1} U, \end{aligned} \tag{12}$$

with $\Gamma = \check{\Sigma}^2 - \varsigma_{r+1}^2 I_{n_{\min}-k}$ and $U = (C_2^{\mathsf{T}})^{\dagger} B_2$. Here, $M^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of a matrix M. This system is constructed such that the error transfer function $\mathcal{E} = G - \tilde{G}$ is scaled all-pass with $\tilde{G}$ the transfer function of (12), i.e., it holds

$$\mathcal{E}(s)\mathcal{E}^{\mathsf{T}}(-s) = \varsigma_{r+1}^2 I_p, \tag{13}$$

for all $s \in \mathbb{C}$ that are no poles of $\mathcal{E}(s)$ or $\mathcal{E}^{\mathsf{T}}(-s)$. In this case, the approximation error satisfies

$$\|\mathcal{E}\|_H = \|\mathcal{E}\|_{\mathcal{L}_\infty} = \varsigma_{r+1}. \tag{14}$$

The transfer function $\tilde{G}$ of (12) has exactly $n_{\min} - k - r$ unstable poles. As last step, an additive decomposition of $\tilde{G}$ is computed such that $\tilde{G} = G_h + G_+$, where G_+ is the anti-stable part of order $n_{\min} - k - r$ and G_h is the stable part of order r. Since the Hankel semi-norm only depends on the stable part of the system, the error (14) in the Hankel semi-norm does not change if the unstable part is removed, such that

$$\|G - G_h\|_H = \varsigma_{r+1}. \tag{15}$$

3.2 Computing a balanced realization for descriptor systems

As for the standard system case, for descriptor systems, a balanced conditionally minimal realization is needed. The term "conditionally" minimal means that the order of the system is minimal except of the reduction of the index-1 parts in E; see [19]. The computation is done using the generalized balanced truncation square root method (GBT(SR)). The basic idea of this method is the computation of a balanced realization and the truncation of unnecessary states.

Definition 4 A realization of a descriptor system (1) is called balanced if

$$\mathcal{G}_{pc} = \mathcal{G}_{po} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{G}_{ic} = \mathcal{G}_{io} = \begin{bmatrix} 0 & 0 \\ 0 & \Theta \end{bmatrix}$$

hold, with the proper Hankel singular values $\Sigma = \operatorname{diag}(\varsigma_1, \ldots, \varsigma_{n_f})$ and the improper Hankel singular values $\Theta = \operatorname{diag}(\theta_1, \ldots, \theta_{n_\infty})$.

The system states are truncated with respect to the computed Hankel singular values. The proper Hankel singular values have the same meaning as the classical

Hankel singular values in the standard case, i.e., states corresponding to small proper Hankel singular values are difficult to control and observe at the same time and can be omitted. In case of the improper Hankel singular values, only zeros can be truncated as mentioned in the previous section; see also [16]. The number of non-zero improper Hankel singular values is equal to the rank of the matrix $\mathcal{G}_{ic}A^{\mathsf{T}}\mathcal{G}_{io}A$, which can in fact be bounded by

$$\operatorname{rank}(\mathcal{G}_{ic}A^{\mathsf{T}}\mathcal{G}_{io}A) \leq \min(\nu m, \nu p, n_\infty), \tag{16}$$

with ν, the index of the system, m, the number of inputs, p, the number of outputs, and n_∞, the dimension of the deflating subspace corresponding to the infinite eigenvalues of $\lambda E - A$. So for large n_∞ and usually small ν, the descriptor system (1) can be reduced significantly by the truncation of zero improper Hankel singular values.

One method to compute the balanced truncation of a descriptor system is the square root method. Therefore, consider the skinny singular value decompositions

$$L_p^{\mathsf{T}} E R_p = \begin{bmatrix} U_1, & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^{\mathsf{T}} \\ V_2^{\mathsf{T}} \end{bmatrix}, \tag{17}$$

$$L_i^{\mathsf{T}} A R_i = U_3 \Theta_3 V_3^{\mathsf{T}}, \tag{18}$$

with $\mathcal{G}_{pc} = R_p R_p^{\mathsf{T}}$, $\mathcal{G}_{po} = L_p L_p^{\mathsf{T}}$, $\mathcal{G}_{ic} = R_i R_i^{\mathsf{T}}$, and $\mathcal{G}_{io} = L_i L_i^{\mathsf{T}}$. The matrices $\begin{bmatrix} U_1, & U_2 \end{bmatrix}$, $\begin{bmatrix} V_1, & V_2 \end{bmatrix}$, U_3 and V_3 have orthonormal columns and the diagonal matrices Σ_1, Σ_2, and Θ_3 contain the non-zero proper and improper Hankel singular values, respectively. The partition of the proper Hankel singular values is chosen such that Σ_1 contains all the desired Hankel singular values and Σ_2 the undesired ones. By using the singular value decompositions (17) and (18), the following projection matrices can be defined

$$\begin{aligned} W_\ell &= \begin{bmatrix} L_p U_1 \Sigma_1^{-\frac{1}{2}}, & L_i U_3 \Theta_3^{-\frac{1}{2}} \end{bmatrix} \in \mathbb{R}^{n \times \ell}, \\ T_\ell &= \begin{bmatrix} R_p V_1 \Sigma_1^{-\frac{1}{2}}, & R_i V_3 \Theta_3^{-\frac{1}{2}} \end{bmatrix} \in \mathbb{R}^{n \times \ell}, \end{aligned} \tag{19}$$

where $\ell = \ell_f + \ell_\infty$ is the sum of the number of desired proper Hankel singular values ℓ_f and the non-zero improper Hankel singular values ℓ_∞. The projected realization

$$(\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{D}) = (W_\ell^{\mathsf{T}} E T_\ell, W_\ell^{\mathsf{T}} A T_\ell, W_\ell^{\mathsf{T}} B, C T_\ell, D) \tag{20}$$

is of order ℓ and balanced with the set of Hankel singular values contained in Σ_1 and Θ_3. The resulting matrix pencil $\lambda \hat{E} - \hat{A}$ resembles the Weierstrass canonical form (4), in that

$$\hat{E} = \begin{bmatrix} I_{\ell_f} & 0 \\ 0 & E_\infty \end{bmatrix} \quad \text{and} \quad \hat{A} = \begin{bmatrix} A_f & 0 \\ 0 & I_{\ell_\infty} \end{bmatrix} \tag{21}$$

hold, where $A_f \in \mathbb{R}^{\ell_f \times \ell_f}$ is non-singular and $E_\infty \in \mathbb{R}^{\ell_\infty \times \ell_\infty}$ is nilpotent with index ν.

Due to the reason that only the zero improper Hankel singular values are truncated, the polynomial part of the system G has not changed. So it can be shown that the same error bound as for the classical balanced truncation method holds. Let $\hat{G}$ be the

reduced descriptor system (20), then it holds

$$\|G - \hat{G}\|_{\mathcal{H}_\infty} \leq 2 \sum_{k=\ell_f+1}^{n_f} \varsigma_k(G),$$

with $\varsigma_k(G)$ the k-th proper Hankel singular value of G.

3.3 Hankel-norm approximation of descriptor systems

As for the standard case, the GHNA method for descriptor systems is based on the construction of an error system with all-pass transfer function (13). The following theorem provides an algebraic characterization of descriptor systems with all-pass transfer functions.

Theorem 1 *Let (E, A, B, C, D) be a realization of a descriptor system (1) with a regular matrix pencil $\lambda E - A$, the same number of inputs and outputs, $m = p$, the system's transfer function $G(s)$ and $\varsigma > 0$ a real constant. Also, it is assumed that the descriptor system is R-controllable and R-observable. Then $G(s)$ is all-pass, i.e., $G(s)G^{\mathsf{T}}(-s) = \varsigma^2 I_m$ holds, if and only if the following conditions are satisfied:*

1. *There are symmetric matrices $\mathcal{G}_{pc}$ and $\mathcal{G}_{po}$ with*

$$\mathcal{G}_{pc} = P_r \mathcal{G}_{pc} P_r^{\mathsf{T}}, \tag{22}$$
$$\mathcal{G}_{po} = P_\ell^{\mathsf{T}} \mathcal{G}_{po} P_\ell. \tag{23}$$

2. *The matrices $\mathcal{G}_{pc}$ and $\mathcal{G}_{po}$ are the solutions of the projected generalized continuous-time Lyapunov equations*

$$E\mathcal{G}_{pc}A^{\mathsf{T}} + A\mathcal{G}_{pc}E^{\mathsf{T}} + P_\ell B B^{\mathsf{T}} P_\ell^{\mathsf{T}} = 0, \tag{24}$$
$$E^{\mathsf{T}}\mathcal{G}_{po}A + A^{\mathsf{T}}\mathcal{G}_{po}E + P_r^{\mathsf{T}} C^{\mathsf{T}} C P_r = 0. \tag{25}$$

3. *The proper Hankel singular values satisfy*

$$\mathcal{G}_{pc}E^{\mathsf{T}}\mathcal{G}_{po}E = \varsigma^2 P_r, \tag{26}$$
$$\mathcal{G}_{po}E\mathcal{G}_{pc}E^{\mathsf{T}} = \varsigma^2 P_\ell^{\mathsf{T}}. \tag{27}$$

4. *Let $G = G_{sp} + P$ be the decomposition of G into the strictly proper part G_{sp} and the polynomial part P. Then it holds $P(s) = \sum_{k=0}^{\infty} M_k s^k$ with*

$$M_0 M_0^{\mathsf{T}} = \varsigma^2 I_m, \tag{28}$$
$$M_k = 0 \quad \text{for} \quad k \geq 1. \tag{29}$$

5. *Also, the following constraints hold*

$$M_0^{\mathsf{T}} C P_r + B^{\mathsf{T}}\mathcal{G}_{po}E = 0, \tag{30}$$
$$M_0 B^{\mathsf{T}} P_\ell^{\mathsf{T}} + C\mathcal{G}_{pc}E^{\mathsf{T}} = 0. \tag{31}$$

Proof The proof can be found in the Appendix. □

Theorem 1 gives us various implications for the computation of the Hankel-norm approximation for descriptor systems. In [8], this algebraic characterization has been used to derive a set of more general transformation formulas for all-pass systems than (12) to handle invertible E matrices. Also it describes the main idea for an algorithm to compute the Hankel-norm approximation. Like in [12], we aim for the construction of an error system that is all-pass by using Theorem 1. While Theorem 1 only considers square transfer functions, the extension to the case $m \neq p$ is based on the all-pass embedding of rational matrix functions and allowing for bounded-real transfer functions instead of all-pass, i.e., $G(s)G^{\mathsf{T}}(-s) \leq \varsigma^2 I_m$. Nevertheless, we refer the reader to [12, Corollary 7.3], which states the extension to non-square transfer functions using the algebraic characterization of all-pass transfer functions and shows that the construction of the Hankel-norm approximation for non-square transfer functions is also based on the formulas developed in Theorem 1. Therefore, we will still denote the transformation based on Theorem 1 as all-pass even for non-square transfer functions.

Now we will discuss the results of Theorem 1. First we observe that (29) enforces the non-constant polynomial part of the error system to be zero, i.e., the system we want to construct as well as the resulting reduced-order model must have the same polynomial part as the original system. Also (24) and (25) necessarily need to be solved for the construction. Therefore, we make use of the generalized balanced truncation method from the previous section, which is additionally used to satisfy the necessary condition of Theorem 1 to get an R-controllable and R-observable realization of the system by the truncation of the zero proper Hankel singular values. Beside that, we can exploit the structure of the balanced reduced-order model (21). So, let the matrices $\hat{B}$ and $\hat{C}$ be partitioned accordingly to (21) as

$$\hat{B} = \begin{bmatrix} B_f \\ B_\infty \end{bmatrix} \quad \text{and} \quad \hat{C} = \begin{bmatrix} C_f, & C_\infty \end{bmatrix}.$$

Using this block partition, the system $(\hat{E}, \hat{A}, \hat{B}, \hat{C}, D)$ automatically decouples into its slow subsystem

$$\begin{aligned} \dot{x}_f(t) &= A_f x_f(t) + B_f u(t), \\ y_f(t) &= C_f x_f(t) \end{aligned} \tag{32}$$

and its fast subsystem

$$\begin{aligned} E_\infty \dot{x}_\infty(t) &= x_\infty(t) + B_\infty u(t), \\ y_\infty(t) &= C_\infty x_\infty(t) + Du(t). \end{aligned} \tag{33}$$

First, the fast subsystem (33) is considered. Since the GBT(SR) was used to compute the system (33), there are no zero improper Hankel singular values anymore. As mentioned in the previous section and concerning (29) from Theorem 1, there is no meaningful further reduction concerning the improper Hankel singular values, so the fast subsystem stays unchanged.

Now, let us consider the slow subsystem (32). It is easy to see that (32) is in standard form. Also beneficial properties, resulting from the applied balanced truncation method, still hold for this subsystem, which means it is stable and balanced.

Let the original system be decomposed as $G = G_{sp} + P$ into its strictly proper part G_{sp} and its polynomial subsystem P. By the truncation of only zero proper

Hankel singular values, the system (32) is a minimal realization of the original slow subsystem G_{sp}. Now, the standard HNA method, mentioned in the previous section, can be applied to (32). As result, an r-th order HNA is computed

$$\begin{aligned} E_h\dot{x}_h(t) &= A_h x_h(t) + B_h u(t), \\ y_h(t) &= C_h x_h(t) + D_h u(t), \end{aligned} \tag{34}$$

where E_h results from the additive decomposition of the all-pass transformed system, if we rewrite the formulas (12) as

$$\begin{aligned} \tilde{E} &= \Gamma, \\ \tilde{A} &= \varsigma_{r+1}^2 A_{11}^{\mathsf{T}} + \breve{\Sigma} A_{11} \breve{\Sigma} + \varsigma_{r+1} C_1^{\mathsf{T}} U B_1^{\mathsf{T}}, \\ \tilde{B} &= \breve{\Sigma} B_1 - \varsigma_{r+1} C_1^{\mathsf{T}} U, \end{aligned} \tag{35}$$

to avoid the disadvantageous scaling by the matrix Γ^{-1} as in the standard system case. This is further discussed in Section 4.1. More general transformation formulas for invertible E matrices have been developed in [8]. To get an optimal HNA of the descriptor system (1), the computed HNA (34) and the reduced-order fast subsystem (33) are coupled

$$\begin{aligned} \begin{bmatrix} E_h & 0 \\ 0 & E_\infty \end{bmatrix} \dot{\hat{x}}(t) &= \begin{bmatrix} A_h & 0 \\ 0 & I_{\ell_\infty} \end{bmatrix} \hat{x}(t) + \begin{bmatrix} B_h \\ B_\infty \end{bmatrix} u(t), \\ \hat{y}(t) &= \begin{bmatrix} C_h, & C_\infty \end{bmatrix} \hat{x}(t) + (D_h + D) u(t). \end{aligned} \tag{36}$$

In the following theorem, the properties of the resulting GHNA are summarized.

Theorem 2 *Let G be a c-stable descriptor system (1) with a regular matrix pencil. The ℓ-th order generalized Hankel-norm approximation (36), with its transfer function $\hat{G}$ and $\ell = r + \ell_\infty$, has the following properties:*

1. *The realization of $\hat{G}$ is conditionally minimal and c-stable.*
2. *The absolute error in the Hankel semi-norm is given by*

$$\|G - \hat{G}\|_H = \varsigma_{r+1}(G),$$

   where $\varsigma_{r+1}(G)$ is the $(r+1)$-st proper Hankel singular value of G.
3. *The absolute error in the $\mathcal{H}_\infty$-norm can be bounded by*

$$\|G - \hat{G}\|_{\mathcal{H}_\infty} \leq 2 \sum_{k=r+1}^{n_f} \varsigma_k(G),$$

   where $\varsigma_k(G)$ is the k-th proper Hankel singular values of G.

Proof Let $G = G_{sp} + P$ be the original system and $\tilde{G} = G_b + P_b$ the balanced, conditionally minimal realization obtained by the GBT(SR) method. Here, G_{sp}, G_b denote the strictly proper parts and P, P_b the polynomial ones. The GHNA is constructed by

$$\hat{G} = G_h + P_b, \tag{37}$$

where G_h is the r-th order HNA (34) of the standard system G_b.

First, we consider part 1. The balanced realization $\tilde{G}$ is conditionally minimal and c-stable. So by construction (37), both of these properties are transferred to the GHNA.

Now we consider the error formulas in 2 and 3 and let $\mathcal{E} = G - \hat{G}$ be the error system of the GHNA. Then it holds

$$\mathcal{E} = G - \hat{G} = G_{sp} + P - G_h - P_b = G_b - G_h,$$

since the balanced realization $\tilde{G}$ is conditionally minimal and therefore, $G_b = G_{sp}$ and $P_b = P$. Using the error bound of the standard method (15), one obtains

$$\|G - \hat{G}\|_H = \|G_b - G_h\|_H = \varsigma_{r+1}(G_b) = \varsigma_{r+1}(G).$$

Using the same approach, the error in the $\mathcal{H}_\infty$-norm is given by

$$\|G - \hat{G}\|_{\mathcal{H}_\infty} = \|G_b - G_h\|_{\mathcal{H}_\infty} \leq 2 \sum_{k=r+1}^{n_f} \varsigma_k(G),$$

if the $\mathcal{H}_\infty$-norm error bound for the standard r-th order HNA from [2] is used. □

In Algorithm 1, the complete GHNA method is summarized. Here, we also point out the differences to the method described in [10]. The two main assumptions, that are necessary for the construction in [10], are the C-controllability and C-observability of the system as well as the system (1) to be given in the Weierstrass canonical form (4). Both assumptions are numerically unfeasible. The computation of the Weierstrass canonical form is numerically highly unstable and should never be done explicitly; see, e.g., [13] for comments on the computation of Jordan blocks necessary for the Weierstrass canonical form. Also, the assumption of C-controllability and C-observability is in general not fulfilled, e.g., all examples presented in this paper satisfy none of those two properties. In contrast to this, the method in Algorithm 1 can avoid those two assumptions. By using the results of Theorem 1 and the generalized balanced truncation, we need neither any assumption on the system structure nor on properties of the system like the C-controllability and C-observability. It should be noted that even with the assumption of the Weierstrass canonical form, for an efficient reduction via the Hankel-norm approximation of the strictly proper part in [10], also a balancing algorithm is needed. Therefore, in terms of computational effort and numerical accuracy our approach is in general less demanding, more stable, and faster than (4). Additionally, the next section shows some numerical extensions and remarks for an even more efficient implementation of Algorithm 1.

4 Numerical methods for GHNA

4.1 Approximate GHNA

The GHNA method can quickly become numerically unstable. This problem arises from the transformation formulas (12) for the construction of a scaled all-pass error

Algorithm 1 Generalized Hankel-Norm Approximation (GHNA) Method.

Input: Realization (E, A, B, C, D) of (1) such that $\lambda E - A$ is c-stable.
Output: Reduced-order system $(\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{D})$.

1: Solve the continuous-time Lyapunov equations (6) and (7) for the Cholesky factorizations $\mathcal{G}_{pc} = R_p R_p^{\mathsf{T}}$ and $\mathcal{G}_{po} = L_p L_p^{\mathsf{T}}$.
2: Solve the discrete-time Lyapunov equations (8) and (9) for the Cholesky factorizations $\mathcal{G}_{ic} = R_i R_i^{\mathsf{T}}$ and $\mathcal{G}_{io} = L_i L_i^{\mathsf{T}}$.
3: Compute the two skinny singular value decompositions

$$L_p^{\mathsf{T}} E R_p = U_1 \Sigma V_1^{\mathsf{T}} \quad \text{and} \quad L_i^{\mathsf{T}} A R_i = U_2 \Theta V_2^{\mathsf{T}}.$$

4: Compute the transformation matrices

$$W_p = L_p U_1 \Sigma^{-\frac{1}{2}}, \quad T_p = R_p V_1 \Sigma^{-\frac{1}{2}},$$
$$W_i = L_i U_2 \Theta^{-\frac{1}{2}}, \quad T_i = R_i V_2 \Theta^{-\frac{1}{2}}.$$

5: Compute the minimal balanced realization of the slow subsystem

$$(I_{\ell_f}, A_f, B_f, C_f, 0) = (W_p^{\mathsf{T}} E T_p, W_p^{\mathsf{T}} A T_p, W_p^{\mathsf{T}} B, C T_p, 0).$$

6: Choose the proper Hankel singular value ς_{r+1}.
7: Permute and partition the Gramians of the slow subsystem

$$\check{\mathcal{G}}_{pc} = \check{\mathcal{G}}_{po} = \operatorname{diag}(\check{\Sigma}, \varsigma_{r+1} I_k),$$

and the corresponding system matrices

$$\check{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \check{B} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad \check{C} = \begin{bmatrix} C_1, & C_2 \end{bmatrix}.$$

8: Compute the all-pass transformation

$$\begin{aligned} \tilde{E} &= \check{\Sigma}^2 - \varsigma_{r+1}^2 I_{\ell_f - k}, \\ \tilde{A} &= \varsigma_{r+1}^2 A_{11}^{\mathsf{T}} + \check{\Sigma} A_{11} \check{\Sigma} + \varsigma_{r+1} C_1^{\mathsf{T}} U B_1^{\mathsf{T}}, \\ \tilde{B} &= \check{\Sigma} B_1 - \varsigma_{r+1} C_1^{\mathsf{T}} U, \\ \tilde{C} &= C_1 \check{\Sigma} - \varsigma_{r+1} U B_1^{\mathsf{T}}, \\ \tilde{D} &= \varsigma_{r+1} U, \end{aligned}$$

with $U = (C_2^{\mathsf{T}})^{\dagger} B_2$.
9: Compute the additive decomposition

$$\tilde{G}(s) = \tilde{C}(s\tilde{E} - \tilde{A})\tilde{B} + \tilde{D} = G_h(s) + F(s),$$

where F is anti-stable and G_h stable with the realization $(E_h, A_h, B_h, C_h, D_h)$.
10: Compute the balanced realization of the fast subsystem

$$(E_\infty, I_{\ell_\infty}, B_\infty, C_\infty, D) = (W_i^{\mathsf{T}} E T_i, W_i^{\mathsf{T}} A T_i, W_i^{\mathsf{T}} B, C T_i, D).$$

11: Couple the resulting subsystems

$$(\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{D}) = \left(\begin{bmatrix} E_h & 0 \\ 0 & E_\infty \end{bmatrix}, \begin{bmatrix} A_h & 0 \\ 0 & I_{\ell_\infty} \end{bmatrix}, \begin{bmatrix} B_h \\ B_\infty \end{bmatrix}, \begin{bmatrix} C_h, & C_\infty \end{bmatrix}, D_h + D \right).$$

transfer function. It is easy to see that the inversion of the diagonal matrix $\Gamma = \check{\Sigma}^2 - \varsigma_{r+1}^2 I_{n_{\min}-k}$ can lead to large numerical errors for small proper Hankel singular values in further computations. This happens if either the chosen value ς_{r+1}, the remaining proper Hankel singular values in $\check{\Sigma}$ or the gap between the chosen value ς_{r+1} and the surrounding ones ς_r and ς_{r+k+1} is small. One preventive measure was the usage of the descriptor system structure (34) to avoid unnecessary scaling by Γ leading to the new transformation formulas (35). While the first problematic case, choosing too small ς_{r+1}, is mainly up to the user of the method, we try to avoid the case of too small gaps between the chosen proper Hankel singular value and its surroundings by considering in the algorithm the region $|\varsigma_{r+1} - \varsigma_j| < \varepsilon$, for ε small and $j = 1, 2, ..., n_{\min}$, to be equal to ς_{r+1}. This increases the number of cut-off Hankel singular values during the all-pass transformation and avoids also too small scaling terms. In the following considerations, we address the case of too small remaining Hankel singular values.

Small proper Hankel singular values can arise from numerical errors during the computation of the minimal realization. Therefore, one approach to solve this problem is to compute a smaller balanced truncation approximation of the slow subsystem than the minimal realization such that too small proper Hankel singular values are cut off. In this case, an additional error is made since the balanced realization is only an approximation of the original system. To get a measure for the additional error, let G_b be the computed balanced truncation of order n_b of the slow subsystem G_{sp}. Then it has been shown in [12] that in the Hankel semi-norm it holds

$$\|G_{sp} - G_b\|_H \leq 2 \sum_{k=n_b+1}^{n_f} \varsigma_k(G_{sp}), \tag{38}$$

with n_f the order of the slow subsystem G_{sp}. For the overall error, let $G = G_{sp} + P$ be the original descriptor system and $\tilde{G} = G_b + P_b$ the balanced realization with G_b of order n_b. The generalized Hankel-norm approximation is denoted by $\hat{G} = G_h + P_b$, where the r-th order standard Hankel-norm approximation G_h was computed from the balanced realization G_b. Using (38) one obtains

$$\begin{aligned} \|G - \hat{G}\|_H &= \|G_{sp} - G_h\|_H \\ &= \|G_{sp} - G_b + G_b - G_h\|_H \\ &\leq \|G_b - G_h\|_H + \|G_{sp} - G_b\|_H \\ &\leq \varsigma_{r+1}(G_b) + 2 \sum_{k=n_b+1}^{n_f} \varsigma_k(G_{sp}). \end{aligned} \tag{39}$$

Since balancing the system does not change the Hankel singular values, the Hankel singular values of G_b and G_{sp} are also the proper Hankel singular values of G. The resulting error can be bounded by

$$\|G - \hat{G}\|_H \leq \varsigma_{r+1}(G) + 2 \sum_{k=n_b+1}^{n_f} \varsigma_k(G).$$

Concerning the $\mathcal{H}_\infty$-norm, the approach (39) can be used to get

$$\|G - \hat{G}\|_{\mathcal{H}_\infty} \leq 2 \sum_{k=r+1}^{n_f} \varsigma_k(G),$$

which is the same error bound as for the exact method.

This approximate version of the GHNA takes advantage of the use of the GBT(SR) method in form of the adaptive choice of the order n_b. It is possible to choose the order n_b with respect to the chosen singular value ς_{r+1} such that

$$2 \sum_{k=n_b+1}^{n_f} \varsigma_k(G) \ll \varsigma_{r+1}(G).$$

In this case, the resulting additional error becomes negligibly small concerning the original Hankel semi-norm error and the corresponding matrix Γ leads to a better conditioned problem. The algorithmic adjustments in the implementation of the GHNA method are small, since only the truncation of non-zero proper Hankel singular values has to be allowed in the generalized balanced truncation method. In this case, the Σ_2 term in (17) with the undesired proper Hankel singular values is not empty and only the matrices U_1, Σ_1, and V_1 are used for further computations.

Another advantage of the approximate algorithm can be found in the computation of the balanced truncation. The GBT(SR) method needs to scale the transformation matrices (19) using the inverse remaining Hankel singular values, which is more accurate if the small proper Hankel singular values are truncated. Also with regard to computational costs, this approximate method has advantages. The further steps of the algorithm, i.e., the all-pass transformation and additive decomposition, are extremely costly for large-scale matrices in terms of computational time and memory usage. Therefore, it is advantageous to already have a small balanced realization for the further computations.

4.2 Application to sparse systems

A frequently appearing case in practice is the model reduction of large-scale sparse descriptor systems. In this case, the system matrices E and A from the descriptor system (1) are in a large-scale sparse form, i.e., the dimension n is large, the matrices can be stored using $\mathcal{O}(n)$ memory and the matrix-vector multiplication can be computed in $\mathcal{O}(n)$ effort. Often such matrices result from the discretization of partial differential equations.

The transformation into a balanced realization does not preserve the sparsity of the system matrices. Therefore, the GHNA method can only be adapted to sparse systems in the first two steps. This concerns the computation of the solutions of the generalized projected Lyapunov equations (6)–(9). It has been observed that the eigenvalues of the symmetric positive semidefinite solutions of Lyapunov equations with low-rank right-hand sides generally decay rapidly. The same result holds for the generalized projected Lyapunov equations [24]. Therefore, the system Gramians can be approximated by low-rank Cholesky factorizations, e.g., $\mathcal{G}_{pc} \approx Z_{pc} Z_{pc}^{\mathsf{T}}$ with $Z_{pc} \in \mathbb{R}^{n \times k}$ and $k \ll n$.

For the proper system Gramians, the computation is done by adapting existing low-rank methods, e.g., Krylov subspace methods or low-rank ADI methods. In this case, the right-hand side has to be replaced by the projected form from the Lyapunov equations (6) and (7). Additionally, it is recommended to project the solution back into the corresponding subspace after some steps of the methods due to a drift-off effect.

In contrast to the proper case, full-rank factorizations of the improper Gramians can be constructed explicitly such that $G_{ic} = Z_{ic} Z_{ic}^{\mathsf{T}}$ and $G_{io} = Z_{io} Z_{io}^{\mathsf{T}}$, with

$$\begin{aligned} Z_{ic} &= \left[Q_r A^{-1} B,\ A^{-1} E Q_r A^{-1} B,\ \ldots,\ (A^{-1}E)^{\nu-1} Q_r A^{-1} B \right], \\ Z_{io} &= \left[Q_\ell^{\mathsf{T}} A^{-\mathsf{T}} C^{\mathsf{T}},\ A^{-\mathsf{T}} E^{\mathsf{T}} Q_\ell^{\mathsf{T}} A^{-\mathsf{T}} C^{\mathsf{T}},\ \ldots,\ (A^{-\mathsf{T}} E^{\mathsf{T}})^{\nu-1} Q_\ell^{\mathsf{T}} A^{-\mathsf{T}} C^{\mathsf{T}} \right]; \end{aligned}$$

see [24] for more details. Thereby, the size of the full-rank factors is bounded by the number of inputs m or outputs p times the system's index ν. This corresponds to the overall bound of the non-zero improper Hankel singular values (16).

Still for using these methods, the spectral projections P_ℓ, P_r, Q_ℓ and Q_r have to be computed. But for many problems, these spectral projections can be applied by exploiting the special structure of the problem; see [24] for some examples.

4.3 The projection-free approach

In case of unstructured problems, there are no explicit construction formulas for the spectral projectors P_ℓ, P_r, Q_ℓ and Q_r, so they have to be explicitly computed for the use in the generalized projected Lyapunov equations (6)–(9). But as for the GBT(SR) method, an alternative approach to the use of spectral projectors can be given; see [22].

As already used in the GHNA algorithm, the GBT(SR) method can be interpreted as a decoupling of the original system into the slow and fast subsystems and the individual reduction of both. Therefore, consider the following generalized block triangular form. There are orthogonal matrices $U, V \in \mathbb{R}^{n\times n}$ such that

$$E = V \begin{bmatrix} E_f & E_u \\ 0 & E_\infty \end{bmatrix} U^{\mathsf{T}} \quad \text{and} \quad A = V \begin{bmatrix} A_f & A_u \\ 0 & A_\infty \end{bmatrix} U^{\mathsf{T}},$$

where the matrix pencil $\lambda E_f - A_f$ contains all the finite eigenvalues of $\lambda E - A$ and the matrix pencil $\lambda E_\infty - A_\infty$ has only infinite eigenvalues. For the computation of a block diagonalization of the system, the coupled Sylvester equations

$$\begin{aligned} E_f Y - Z E_\infty &= -E_u, \\ A_f Y - Z A_\infty &= -A_u, \end{aligned}$$

have to be solved for Y and Z; see [5]. Using all of these matrices for the restricted system equivalence transformation

$$W_{dec} = V \begin{bmatrix} I_{n_f} & 0 \\ -Z^{\mathsf{T}} & I_{n_\infty} \end{bmatrix}, \quad T_{dec} = U \begin{bmatrix} I_{n_f} & Y \\ 0 & I_{n_\infty} \end{bmatrix}$$

of the original descriptor system (1), one obtains

$$\begin{aligned} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} \dot{\tilde{x}}(t) &= \begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} B_f \\ B_\infty \end{bmatrix} u(t), \\ y(t) &= \begin{bmatrix} C_f & C_\infty \end{bmatrix} \tilde{x}(t) + Du(t), \end{aligned} \tag{40}$$

where the remaining matrices are constructed as

$$\begin{aligned} V^{\mathsf{T}} B &= \begin{bmatrix} B_u \\ B_\infty \end{bmatrix}, \qquad B_f = B_u - Z B_\infty, \\ CU &= \begin{bmatrix} C_f \\ C_u \end{bmatrix}, \qquad C_\infty = C_f Y + C_u. \end{aligned} \tag{41}$$

Obviously, the realization in (40) decouples into the fast and slow subsystems of (1). Since the spectral projectors of the subsystems are identity matrices, the corresponding Lyapunov equations (6)–(9) simplify to

$$\begin{aligned} E_f X_{pc} A_f^{\mathsf{T}} + A_f X_{pc} E_f^{\mathsf{T}} + B_f B_f^{\mathsf{T}} &= 0, \\ E_f^{\mathsf{T}} X_{po} A_f + A_f^{\mathsf{T}} X_{po} E_f + C_f^{\mathsf{T}} C_f &= 0, \end{aligned}$$

for the slow subsystem and

$$\begin{aligned} A_\infty X_{ic} A_\infty^{\mathsf{T}} - E_\infty X_{ic} E_\infty^{\mathsf{T}} - B_\infty B_\infty^{\mathsf{T}} &= 0, \\ A_\infty^{\mathsf{T}} X_{io} A_\infty - E_\infty^{\mathsf{T}} X_{io} E_\infty - C_\infty^{\mathsf{T}} C_\infty &= 0, \end{aligned}$$

for the fast subsystem. These Lyapunov equations can be computed without the spectral projections. The matrices X_{pc} and X_{po} correspond to the parts of the proper controllability and observability Gramians, which contain the potentially non-zero proper Hankel singular values. The same holds for X_{ic}, X_{io} and the improper system Gramians. For the rest of the algorithm, only the transformations have to be restricted to the subsystems.

The projection-free approach is implemented in the version 3.0 of the MORLAB toolbox [7]. In this special implementation, the block diagonalization of the system is done by using a block transformation approach based on the following generalization of Theorem 4.1 from [15].

Theorem 3 *Let $\Delta \subset \mathbb{C}$ be a region in the complex plane, which contains n_1 eigenvalues of the matrix pencil $\lambda E - A$. Let $Q, Z \in \mathbb{R}^{n\times n}$ be orthogonal matrices that transform the matrix pencil $\lambda E - A$ into the upper block triangular form*

$$Q^{\mathsf{T}}(\lambda E - A)Z = \begin{bmatrix} Q_1^{\mathsf{T}} \\ Q_2^{\mathsf{T}} \end{bmatrix} (\lambda E - A) \begin{bmatrix} Z_1, Z_2 \end{bmatrix} = \begin{bmatrix} \lambda E_{11}^{(1)} - A_{11}^{(1)} & \lambda E_{12}^{(1)} - A_{12}^{(1)} \\ 0 & \lambda E_{22}^{(1)} - A_{22}^{(1)} \end{bmatrix},$$

with $\Lambda(A_{11}^{(1)}, E_{11}^{(1)}) \subseteq \Delta$ and $\Lambda(A_{11}^{(1)}, E_{11}^{(1)}) \cap \Lambda(A_{22}^{(1)}, E_{22}^{(1)}) = \emptyset$. Similarly, let $U, V \in \mathbb{R}^{n\times n}$ be orthogonal matrices that transform the matrix pencil $\lambda E - A$ into the upper block triangular form

$$U^{\mathsf{T}}(\lambda E - A)V = \begin{bmatrix} U_1^{\mathsf{T}} \\ U_2^{\mathsf{T}} \end{bmatrix} (\lambda E - A) \begin{bmatrix} V_1, V_2 \end{bmatrix} = \begin{bmatrix} \lambda E_{11}^{(2)} - A_{11}^{(2)} & \lambda E_{12}^{(2)} - A_{12}^{(2)} \\ 0 & \lambda E_{22}^{(2)} - A_{22}^{(2)} \end{bmatrix},$$

with $\Lambda(A_{22}^{(2)}, E_{22}^{(2)}) \subseteq \Delta$ *and* $\Lambda(A_{11}^{(2)}, E_{11}^{(2)}) \cap \Lambda(A_{22}^{(2)}, E_{22}^{(2)}) = \emptyset$. *Then*

$$X = \begin{bmatrix} U_2, & Q_2 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} Z_1, & V_1 \end{bmatrix}$$

are transformation matrices, such that $X^{\mathsf{T}}(\lambda E - A)Y$ *has a block diagonal structure where the upper block contains the* n_1 *eigenvalues lying inside* Δ *and the lower block has the remaining* $n - n_1$ *eigenvalues of* $\lambda E - A$ *outside of* Δ.

Proof The proof can be found in [25, Section 5.2]. □

In contrast to the approach above, it is not necessary to compute the solution of the coupled Sylvester equations and, due to the block orthogonal structure of the transformation matrices, the right-hand sides are usually better conditioned than (41). In MORLAB, the right matrix pencil disk function method is used to generate the block transformation matrices; see [25] for more details on the implementation. Additionally, Theorem 3 can be used to compute the additive decomposition in Step 9 of Algorithm 1 by separating the eigenvalues with negative and positive real-parts.

5 Numerical examples

Three typical benchmarks from the descriptor systems model reduction literature have been chosen to demonstrate the introduced GHNA method. All the computations were done on a machine with one Intel(R) Core(TM) i7-6700 CPU processor running at 3.40GHz and equipped with 8 GB total main memory. The computer is running on Ubuntu 16.04.4 LTS and uses MATLAB 9.1.0.441655 (R2016b).

5.1 Semi-discretized Stokes equation

First, the method is tested on a large-scale sparse example. The Stokes equation describes the flow of fluids at very low velocities without convection and coincides with the linearization of the Navier-Stokes equation around the zero-state. The spatial discretization of the Stokes equation by the finite volume method leads to a descriptor system of the form

$$\begin{aligned} \dot{v}_h(t) &= A_{11}v_h(t) + A_{12}p_h(t) + B_1u(t), \\ 0 &= A_{12}^{\mathsf{T}}v_h(t) + B_2u(t), \\ y(t) &= C_1v_h(t) + C_2p_h(t), \end{aligned} \tag{42}$$

where v_h and p_h are the semi-discretized vectors of velocity and pressure, respectively, and the matrices B_1, B_2, C_1, C_2 are all vectors, i.e., the system has $m = p = 1$ inputs and outputs. For matrix pencils like in (42), the spectral projectors P_ℓ and P_r are given by explicit construction formulas

$$\begin{aligned} P_\ell &= \begin{bmatrix} \Pi & -\Pi A_{11}A_{12}(A_{12}^{\mathsf{T}}A_{12})^{-1} \\ 0 & 0 \end{bmatrix}, \\ P_r &= \begin{bmatrix} \Pi & 0 \\ -(A_{12}^{\mathsf{T}}A_{12})^{-1}A_{12}^{\mathsf{T}}A_{11}\Pi & 0 \end{bmatrix}, \end{aligned}$$

where $\Pi = I_{n_v} - A_{12}(A_{12}^{\mathsf{T}}A_{12})^{-1}A_{12}^{\mathsf{T}}$ is the orthogonal projector onto the kernel of A_{12}^{T} along the image of A_{12}; see [23]. The generation of data is based on the test example 3.3 in [18]. The Stokes equation was discretized on a uniform staggered grid of 80×80 points, which leads to a descriptor system of the size $n = 19\,039$, where the matrix pencil $\lambda E - A$ has $n_f = 6\,241$ finite and $n_\infty = 12\,798$ infinite eigenvalues. The data was generated to get a full-rank A_{12} such that the system (42) is of index 2.

For the computation, the implementation of the GHNA method was adjusted to the sparse system case, as described in Section 4.2, and for the solution of the projected continuous-time Lyapunov equations (6) and (7), the solvers from version 1.0.1 of the M-M.E.S.S. toolbox have been used [17]. See the demo file `bt_mor_DAE2.m` in [17] for the applied parameter settings. An approximation of the non-zero proper Hankel singular values has been computed and plotted in Fig. 1 using the low-rank factorizations of the proper system Gramians.

As mentioned before, it is numerically more stable to use a balanced truncation of the slow subsystem than the minimal realization. For this reason, a tolerance for the allowed proper Hankel singular values was computed as $2 \cdot \log(n) \cdot \epsilon$ and multiplied with the largest proper Hankel singular value, with n the order of the system and ϵ the machine epsilon. The resulting bound is also shown in Fig. 1 and the computed balanced realization is of order 20.

To compute a fourth-order Hankel-norm approximation of the slow subsystem, the fifth proper Hankel singular value $\varsigma_5 = 1.8370 \cdot 10^{-6}$ was chosen. The additive decomposition of the transformed realization (12) was achieved by using the `ml_adtf_dss` routine from version 3.0 of the MORLAB toolbox [7]. The solutions of the projected generalized discrete-time Lyapunov equations (8) and (9) were constructed as shown in Section 4.2. In contrast to the continuous-time case, every iteration step was reprojected since the iteration converges after 2 steps at maximum. As result only one non-zero improper Hankel singular value $\theta_1 = 5.3046 \cdot 10^{-18}$ was computed. This implies that the reduced-order system would be of index 1. In this case, the fast subsystem (33) is equivalent to a feed-through term of the form $-C_\infty B_\infty = -1.875 \cdot 10^{-17}$. Since this value is negligible small compared with the

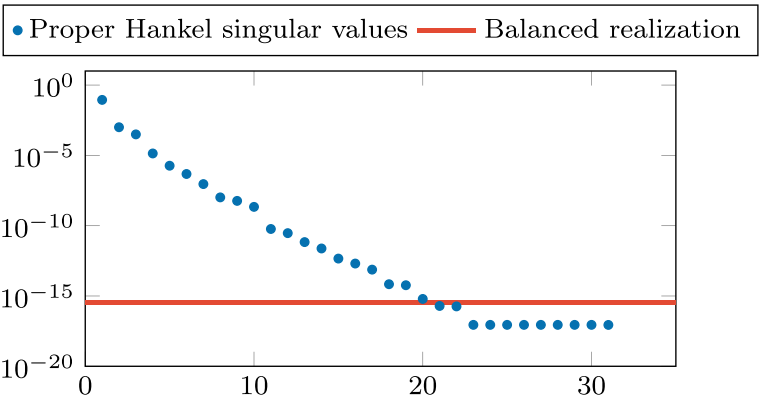

Fig. 1 Computed proper Hankel singular values and the tolerance for the balanced realization for the Stokes example

resulting feed-through term $\hat{D} = \varsigma_5$ from the GHNA method, the state corresponding to this improper Hankel singular value was truncated, too.

To sum up, the original semi-discretized Stokes equation is approximated by a GHNA of order 4 ($r = 4$, $\ell_\infty = 0$). In Fig. 2a, the transfer functions of the full-order model, the GHNA reduced-order model and, for comparison, a reduced-order model of the same order generated by GBT(SR) are plotted. The corresponding errors in the spectral norm are shown in Fig. 2b with the $\mathcal{H}_\infty$ error bound. The shown error behavior of the GHNA is very typical. Since the reduced-order model is based on an all-pass error transfer function, the error becomes nearly all-pass if the influence of the anti-stable part is negligible small. Also, the error of the GHNA approaches the chosen proper Hankel singular value ς_5, which is exactly the error of the approximation in the Hankel semi-norm.

Additional examples and tests of the sparse implementation of the GHNA method can also be found in [25].

5.2 Brazilian interconnected power system model

As second example, we consider the model of a power system network. In general, those power systems have a specific index-1 structure

$$\begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix} \dot{x}(t) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} x(t) + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t), \qquad y(t) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} x(t) + Du(t), \tag{43}$$

where E_{11} and A_{22} are both invertible. By solving the second block line in (43) for the states x_2, the constraints can be eliminated, leading to a state-space realization of the system. This formulation and the corresponding spectral projectors can be found in [6].

For our setup, we use the example data `bips07_3078` from [11], which describes the Brazilian interconnected power system for the year 2007. This system has overall $n = 21\,128$ ($n_f = 3\,078$, $n_\infty = 18\,050$) states and $m = p = 4$ inputs and

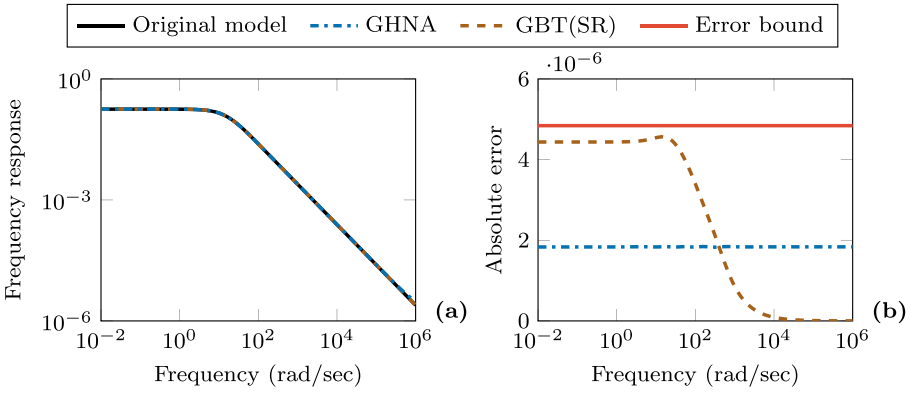

Fig. 2 Transfer functions of the full and reduced-order models of order 4 generated by GHNA and GBT(SR) (**a**), and the absolute approximation error in the spectral norm (**b**) for the Stokes example

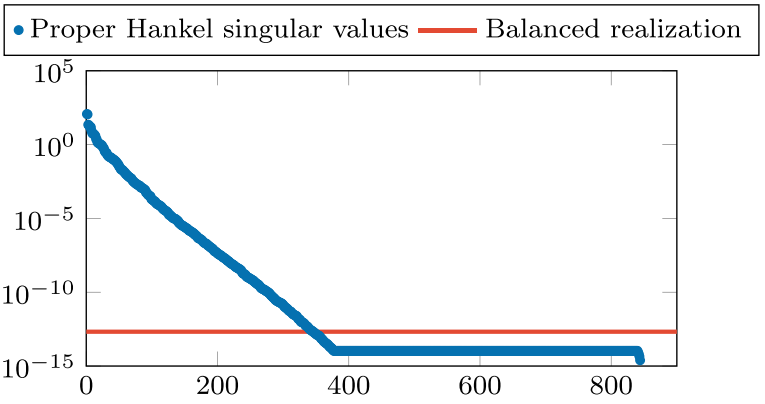

Fig. 3 Computed proper Hankel singular values and the tolerance for the balanced realization for the Brazilian interconnected power system example

outputs. As before we use the M-M.E.S.S. toolbox, with a small code modification to continue the solver iterations for increasing residues, as well as MORLAB for the computation of the generalized Hankel-norm approximation. As suggested by the authors of [11], we will use a numerical trick to stabilize the system for the matrix equation solvers and model reduction methods. The α-shift approach uses a small shift of the A matrix, i.e., $\check{A} = A - \alpha E$, to push the eigenvalues further away from the imaginary axis. Instead of the suggested shift of 0.08 in [11], we use $\alpha = 10^{-4}$, which is small enough to have no visible difference between the shifted and original transfer function. Therefore, we will not shift the system back after the reduction but use the shifted system as our example.

Figure 3 shows the decay of the proper Hankel singular values with the truncation tolerance for the minimal realization. There were only zero improper Hankel singular values computed, which was expected for a strictly proper system. For the

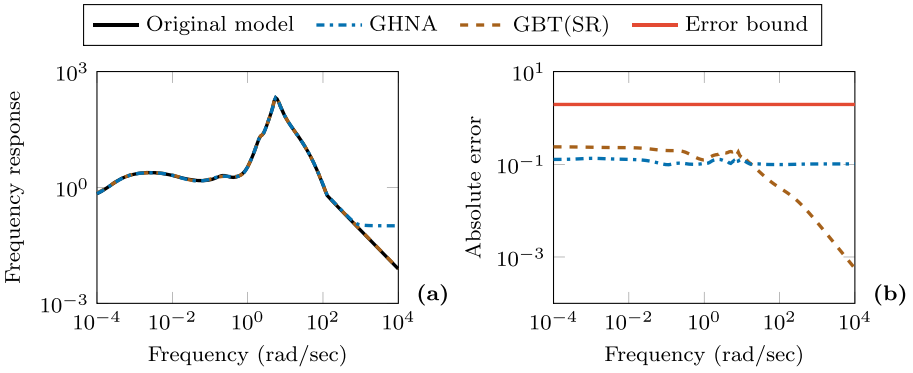

Fig. 4 Transfer functions of the full and reduced-order models of order 40 generated by GHNA and GBT(SR) (**a**), and the absolute approximation error in the spectral norm (**b**) for the Brazilian interconnected power system example

reduced-order model, $r = 40$ ($\ell_\infty = 0$) with $\varsigma_{41} = 1.0175 \cdot 10^{-1}$ has been chosen. Figure 4a shows the transfer functions, while the absolute approximation error with the $\mathcal{H}_\infty$ error bound is plotted in Fig. 4b. As in the previous example, the spectral norm error of the GHNA lies around the chosen proper Hankel singular value and is smaller for low frequencies than the GBT(SR) error. Since by construction the GHNA has a proper transfer function, it will not match the behavior of a strictly proper transfer function for high frequencies. It can be said, that the behavior will diverge, if the frequency response is smaller than the chosen proper Hankel singular value. Even so, up to 10^3 rad/sec the GHNA is matching the original transfer function's behavior just like the GBT(SR).

5.3 A damped mass-spring system

The last example is a damped mass-spring system with a holonomic constraint. The detailed construction of the system can be found in [16]. The vibrations of the resulting system are described by a system of second-order equations

$$\begin{aligned} M\ddot{p}(t) &= Kp(t) + D\dot{p}(t) - G^{\mathsf{T}}\lambda(t) + B_u u(t), \\ 0 &= Gp(t), \\ y(t) &= C_p p(t), \end{aligned} \tag{44}$$

where $p(t)$ is the vector of positions, $\lambda(t) \in \mathbb{R}$ is the Lagrange multiplier, $K, D \in \mathbb{R}^{g\times g}$ are the tridiagonal stiffness and damping matrices, $M = \operatorname{diag}(m_1, \ldots, m_g)$ is the mass matrix, and $G = \begin{bmatrix} 1, 0, \ldots, 0, -1 \end{bmatrix}$ is the constraint matrix. The input is given by $B_u = e_1$ and three positions of masses are measured by $C_p = \begin{bmatrix} e_1, \ e_2, \ e_{g-1} \end{bmatrix}^{\mathsf{T}}$, where e_i is the i-th column of I_g.

For the application of the GHNA method, the system (44) has to be rewritten in first-order form. Therefore, the velocity vector $v(t) = \dot{p}(t)$ is introduced and all

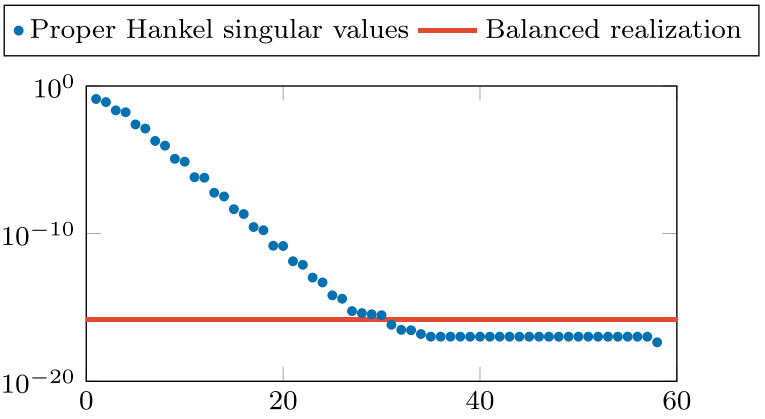

Fig. 5 Computed proper Hankel singular values and the tolerance for the balanced realization for the damped mass-spring example

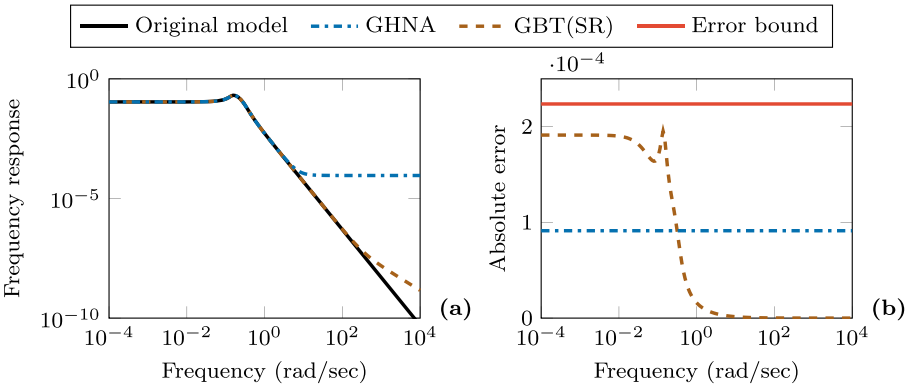

Fig. 6 Transfer functions of the full and reduced-order models of order 7 generated by GHNA and GBT(SR) (**a**), and the absolute approximation error in the spectral norm (**b**) for the damped mass-spring example.

states are collected in $x(t) = \left[\, p(t)^{\mathsf{T}}, v(t)^{\mathsf{T}}, \lambda(t) \,\right]^{\mathsf{T}}$, such that the system (44) can be rewritten in the form

$$\begin{bmatrix} I_g & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & 0 \end{bmatrix} \dot{x}(t) = \begin{bmatrix} 0 & I_g & 0 \\ K & D & -G^{\mathsf{T}} \\ G & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ B_u \\ 0 \end{bmatrix} u(t), \qquad (45)$$
$$y(t) = \begin{bmatrix} C_p & 0 & 0 \end{bmatrix} x(t).$$

This linearization is an index-3 descriptor system. The number of masses was chosen as $g = 1\,500$, which leads to $n = 3\,001$ ($n_f = 2\,998$, $n_\infty = 3$) states in the linearized system (45). For the computation of the GHNA, the `ml_hna_dss` routine from version 3.0 of the MORLAB toolbox has been used [7]. In this function, the projection-free approach from Section 4.3 is implemented as mentioned there. For the computation of the additive decompositions, the right matrix pencil disk function is used and the generalized Lyapunov equations are solved via the matrix sign function method; see, for example, [3] and [4]. More details on handling descriptor systems with the MORLAB toolbox can be found in [9]. The computed proper Hankel singular values and the used bound for the minimal realization of the system can be seen in Fig. 5.

The computed reduced-order model is of order 7 ($r = 7$, $\ell_\infty = 0$) with a Hankel-norm error of $\varsigma_8 = 9.1301 \cdot 10^{-5}$, which gives us again a reduced-order model with only ordinary differential equations, i.e., a regular matrix $\hat{E}$. Figure 6 a shows the transfer functions and the corresponding errors with their $\mathcal{H}_\infty$ error bound can be found in Fig. 6b.

6 Conclusion

An algebraic characterization of descriptor systems with all-pass transfer function was proven and based on this, an efficient algorithm for the computation of the generalized Hankel-norm approximation was developed by exploiting the generalized

balanced truncation square root method. To get a numerically more stable algorithm, an approximate version of the Hankel-norm approximation was introduced. For an efficient practical usage, the introduced method was considered for sparse large-scale systems as well as for unstructured dense systems. The approximation behavior of the method was illustrated for large- and medium-scale examples.

Compared with the approach suggested in [10], the method introduced in this paper has several numerical advantages. It has a more stable and efficient computational behavior, due to the fact that the Weierstrass canonical form does not have to be computed. Also, the introduced method can be applied to more general descriptor systems since C-controllability and C-observability are not required.

Funding information Open access funding provided by Projekt DEAL. This work was supported by the German Research Foundation (DFG) Priority Program 1897: "Calm, Smooth and Smart – Novel Approaches for Influencing Vibrations by Means of Deliberately Introduced Dissipation" and the German Research Foundation (DFG) Research Training Group 2297 "MathCoRe", Magdeburg.

Appendix

For the proof of Theorem 1, the following lemma is used.

Lemma 1 *Every all-pass transfer function $G(s) \in \mathbb{C}^{m\times m}$ is proper.*

Proof From the definition of all-pass transfer functions follows that the product $G(s)G^{\mathsf{T}}(-s) = I_m$ has to be proper.

Improper Case: First, we assume that G is an improper transfer function. The entries of $G(s)$ are rational functions with real coefficients. Since G is improper, at least one entry of G must have a higher polynomial degree in the numerator than in the denominator. Also, one can observe that for $G^{\mathsf{T}}(-s)$ the entries of the matrix are only transposed and coefficients of odd polynomial order change their signs. In the single-input single-output (SISO) case $G(s)$ is given by $G(s) = \frac{n(s)}{d(s)}$, with $\deg(n) > \deg(d)$. Let the numerator and denominator of the para-Hermitian function be denoted by $\tilde{n}(s)$ and $\tilde{d}(s)$. In this case, it is obvious that for the product it holds $2 \cdot \deg(n) = \deg(n\tilde{n}) > \deg(d\tilde{d}) = 2 \cdot \deg(d)$. So the product is always improper.

In the multi-input multi-output (MIMO) case, for simplicity, it is assumed that $m = 2$ and that the denominator is equal for all entries and is factored out such that

$$G(s) = \frac{1}{d(s)} \begin{bmatrix} n_{11}(s) & n_{12}(s) \\ n_{21}(s) & n_{22}(s) \end{bmatrix} \quad \text{and} \quad G^{\mathsf{T}}(-s) = \frac{1}{\tilde{d}(s)} \begin{bmatrix} \tilde{n}_{11}(s) & \tilde{n}_{21}(s) \\ \tilde{n}_{12}(s) & \tilde{n}_{22}(s) \end{bmatrix}.$$

The resulting product is then

$$G(s)G^{\mathsf{T}}(-s) = \frac{1}{d(s)\tilde{d}(s)} \begin{bmatrix} n_{11}(s)\tilde{n}_{11}(s) + n_{12}\tilde{n}_{12} & n_{11}(s)\tilde{n}_{21}(s) + n_{12}\tilde{n}_{22} \\ n_{21}(s)\tilde{n}_{11}(s) + n_{22}\tilde{n}_{12} & n_{21}(s)\tilde{n}_{21}(s) + n_{22}\tilde{n}_{22} \end{bmatrix}.$$

If only one of the entries of $G(s)$ would have a higher polynomial degree than the denominator, the argumentation from the SISO case would follow. Therefore, we can assume w.l.o.g. that

$$\deg(n_{11}) = \deg(n_{12}) = \deg(d) + 1 = g + 1.$$

For simplicity, we concentrate on the $(1, 1)$ entry of the matrix product. For the resulting polynomial degrees it holds $\deg(n_{11}\tilde{n}_{11}) = 2g + 2$, $\deg(n_{12}\tilde{n}_{12}) = 2g + 2$, $\deg(n_{11}\tilde{n}_{11} + n_{12}\tilde{n}_{12}) \leq 2g + 2$, $\deg(d\tilde{d}) = 2g$. To get a proper transfer function in the product, we need that the coefficients in $n_{11}\tilde{n}_{11} + n_{12}\tilde{n}_{12,}$ corresponding to the two highest exponents, cancel out. If we now develop the polynomials as

$$\begin{aligned} n_{11}(s) &= \sum_{k=0}^{g+1} n_{11,k}s^k,\ n_{12}(s) = \sum_{k=0}^{g+1} n_{12,k}s^k, \\ \tilde{n}_{11}(s) &= \sum_{k=0}^{g+1} \tilde{n}_{11,k}s^k,\ \tilde{n}_{12}(s) = \sum_{k=0}^{g+1} \tilde{n}_{12,k}s^k, \end{aligned}$$

we get for the first coefficients $n_{11,g+1}\tilde{n}_{11,g+1} = -n_{12,g+1}\tilde{n}_{12,g+1}$, with $|n_{11,g+1}| = |\tilde{n}_{11,g+1}|$ and $|n_{12,g+1}| = |\tilde{n}_{12,g+1}|$. Now, if $g + 1$ is even then

$$n_{11,g+1} = \tilde{n}_{11,g+1},\ n_{12,g+1} = \tilde{n}_{12,g+1} \ \Rightarrow n_{11,g+1}^2 = -\tilde{n}_{12,g+1}^2,$$

and if $g + 1$ is odd

$$n_{11,g+1} = -\tilde{n}_{11,g+1},\ n_{12,g+1} = -\tilde{n}_{12,g+1}, \ \Rightarrow -n_{11,g+1}^2 = \tilde{n}_{12,g+1}^2.$$

Both cases are a contradiction to the condition that the coefficients are real and non-zero. Therefore, an all-pass transfer function cannot be improper.

Strictly Proper Case: Now, let us assume that G is a strictly proper transfer function. Using the same argumentation as in the improper case, we get that the product of a strictly proper transfer function with its para-Hermitian is also strictly proper. □

Now, we proof the results of Theorem 1.

Proof At first, we can assume w.l.o.g. that $\varsigma = 1$, since the system can be scaled to that case by $\tilde{B} = \varsigma^{-\frac{1}{2}}B$, $\tilde{C} = \varsigma^{-\frac{1}{2}}C$ and $\tilde{D} = \varsigma^{-1}D$.

"⇒": Assume the transfer function $G(s)$ is all-pass. From Lemma 1 it follows that $G(s)$ has to be proper. If we consider now the decomposition of the transfer function into its strictly proper and polynomial part $G(s) = G_{sp}(s) + P(s)$, the polynomial part must satisfy

$$P(s) = \sum_{k=1}^{\infty} M_k s^k,$$

with $M_k = 0$ for all $k \geq 1$. In this case, it holds $\lim_{s\to\infty} G(s) = M_0$, and with the definition of all-pass transfer functions we get $M_0 M_0^{\mathsf{T}} = G(s)G^{\mathsf{T}}(-s) = I_m$. Therefore,

the expressions (28) and (29) hold. Since the matrix pencil $\lambda E - A$ is assumed to be regular, there are non-singular matrices $Q, Z \in \mathbb{R}^{n\times n}$, which resemble the Weierstrass canonical form (4), similar to (21). Using P and Q for a restricted system equivalence transformation leads to the block partitioned realization

$$(QEZ, QAZ, QB, CZ, D) = \left(\begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix}, \begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix}, \begin{bmatrix} B_f \\ B_\infty \end{bmatrix}, \begin{bmatrix} C_f & C_\infty \end{bmatrix}, D\right), \tag{46}$$

where $\lambda E_f - A_f$ contains all the finite eigenvalues of $\lambda E - A$ and $\lambda E_\infty - A_\infty$ contains only infinite eigenvalues. This block diagonal system decouples into its slow and fast subsystems (32), (33), with the realizations $(E_f, A_f, B_f, C_f, 0)$ and $(E_\infty, A_\infty, B_\infty, C_\infty, D)$. While the slow subsystem corresponds to the strictly proper part of the transfer function, the fast subsystem belongs to the polynomial part. The constant part of the transfer function is then given by $M_0 = D - C_\infty A_\infty^{-1} B_\infty$ and the transfer function can also be realized as $G(s) = C_f(sE_f - A_f)^{-1}B_f + M_0$, with invertible matrix E_f.

By the definition of an all-pass transfer function, i.e., $G(s)G^\mathsf{T}(-s) = I_m$, we get the equality $G^{-1}(s) = G^\mathsf{T}(-s)$, which can be written as

$$\begin{aligned} G^{-1}(s) &= M_0^{-1} - M_0^{-1}C_f(sE_f - A_f + B_f M_0^{-1}C_f)^{-1}B_f M_0^{-1} \\ &= M_0^\mathsf{T} + B_f^\mathsf{T}(-sE_f^\mathsf{T} - A_f^\mathsf{T})^{-1}C_f^\mathsf{T} \\ &= G^\mathsf{T}(-s). \end{aligned}$$

The equality $M_0^{-1} = M_0^\mathsf{T}$ was already proven above. From the R-controllability and R-observability assumption together with the regularity of E_f, it follows that there exist invertible matrices $T, W \in \mathbb{R}^{n\times n}$, which transform one realization into the other, i.e.,

$$E_f^\mathsf{T} = W E_f T, \tag{47}$$

$$-A_f^\mathsf{T} = W(A_f - B_f M_0^\mathsf{T} C_f)T, \tag{48}$$

$$C_f^\mathsf{T} = W B_f M_0^\mathsf{T}, \tag{49}$$

$$B_f^\mathsf{T} = M_0^\mathsf{T} C_f T. \tag{50}$$

By reformulating (49) we obtain

$$C_f^\mathsf{T} = W B_f M_0^\mathsf{T} \iff B_f^\mathsf{T} = M_0^\mathsf{T} C_f W^{-\mathsf{T}},$$

and from (50) we get

$$B_f^\mathsf{T} = M_0^\mathsf{T} C_f T \iff C_f^\mathsf{T} = T^{-\mathsf{T}} B_f M_0^\mathsf{T}.$$

The equation (48) can be rewritten as

$$\begin{aligned} -A_f^{\mathsf{T}} &= W(A_f - B_f M_0^{\mathsf{T}} C_f)T \\ \iff -W^{-1}A_f^{\mathsf{T}}T^{-1} &= A_f - B_f M_0^{\mathsf{T}} C_f \\ \iff A_f &= -W^{-1}A_f^{\mathsf{T}}T^{-1} + B_f M_0^{\mathsf{T}} C_f \\ \iff -A_f^{\mathsf{T}} &= T^{-\mathsf{T}}A_f W^{-\mathsf{T}} - C_f^{\mathsf{T}} M_0 B_f^{\mathsf{T}} = T^{-\mathsf{T}}(A_f - B_f M_0^{\mathsf{T}} C_f)W^{-\mathsf{T}}, \end{aligned}$$

and for (47) we analogously get

$$E_f^{\mathsf{T}} = W E_f T \quad \iff E_f^{\mathsf{T}} = T^{-\mathsf{T}} E_f W^{-\mathsf{T}}.$$

Therefore, T and $W^{-\mathsf{T}}$, as well as T^{-1} and W^{T}, satisfy the same set of equations, which means that $W = T^{-\mathsf{T}}$. Using this, the expressions (47)–(50) are equivalent to

$$E_f^{\mathsf{T}} = T^{-\mathsf{T}} E_f T, \tag{51}$$

$$-A_f^{\mathsf{T}} = T^{-\mathsf{T}}(A_f - B_f M_0^{\mathsf{T}} C_f)T, \tag{52}$$

$$C_f^{\mathsf{T}} = T^{-\mathsf{T}} B_f M_0^{\mathsf{T}}, \tag{53}$$

$$B_f^{\mathsf{T}} = M_0^{\mathsf{T}} C_f T. \tag{54}$$

Then, the matrix T is given as the solution of the system of matrix equations following from (51), (52) and (54)

$$A_f T + T^{\mathsf{T}} A_f^{\mathsf{T}} - B_f B_f^{\mathsf{T}} = 0, \; E_f T = T^{\mathsf{T}} E_f^{\mathsf{T}}.$$

Using the fact that E_f is invertible, we can define the symmetric matrix $\tilde{G}_{pc} = -T E_f^{-\mathsf{T}} = -E_f^{-1} T^{\mathsf{T}}$ and replace the system of matrix equations by the following generalized Lyapunov equation

$$A_f \tilde{G}_{pc} E_f^{\mathsf{T}} + E_f \tilde{G}_{pc} A_f^{\mathsf{T}} + B_f B_f^{\mathsf{T}} = 0. \tag{55}$$

For the matrix $\tilde{G}_{pc}$, it then holds that

$$\tilde{G}_{pc} E_f^{\mathsf{T}} \tilde{G}_{po} E_f = (-T E_f^{-\mathsf{T}}) E_f^{\mathsf{T}} (-T^{-1} E_f^{-1}) E_f = T T^{-1} = I_{n_f}. \tag{56}$$

Additionally, from (53) the following constraint is obtained

$$\begin{aligned} T^{\mathsf{T}} C_f^{\mathsf{T}} &= B_f M_0^{\mathsf{T}} \\ \iff M_0 B_f^{\mathsf{T}} - C_f T &= 0 \\ \iff M_0 B_f^{\mathsf{T}} + C_f G_{pc} E_f^{\mathsf{T}} &= 0. \end{aligned} \tag{57}$$

While observing that the dual conditions can be obtained analogously by using (51), (52) and (53), the all-pass characterization is proven for the realization of G with the invertible matrix E_f.

In the next step, the original realization of the system has to be rebuild by using the block diagonal structure (46). Therefore, we need to apply appropriate spectral projectors of the deflating subspaces corresponding to the finite eigenvalues of the

matrix pencil $\lambda E - A$. In case of (46), those left and right spectral projectors are given by

$$\tilde{P}_\ell = \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \tilde{P}_r = \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix}.$$

Since the matrices $\tilde{G}_{pc}$ and $\tilde{G}_{po}$ are only determined by the system parts corresponding to the finite eigenvalues, they have to be expanded accordingly to the spectral projectors by

$$\tilde{G}_{pc} \to \begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \tilde{G}_{po} \to \begin{bmatrix} \tilde{G}_{po} & 0 \\ 0 & 0 \end{bmatrix}.$$

Using this, the equation (55) is equivalent to

$$\begin{aligned} &\begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix} \begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix}^\mathsf{T} + \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} \begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix}^\mathsf{T} \\ &+ \tilde{P}_\ell \begin{bmatrix} B_f \\ B_\infty \end{bmatrix} \begin{bmatrix} B_f \\ B_\infty \end{bmatrix}^\mathsf{T} \tilde{P}_\ell^\mathsf{T} = 0. \end{aligned} \tag{58}$$

Also, the matrix product in (56) becomes

$$\begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix}^\mathsf{T} \begin{bmatrix} \tilde{G}_{po} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} = \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix},$$

and the constraint (57) becomes

$$M_0 \begin{bmatrix} B_f \\ B_\infty \end{bmatrix}^\mathsf{T} \tilde{P}_\ell^\mathsf{T} + \begin{bmatrix} C_f & C_\infty \end{bmatrix} \begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix}^\mathsf{T} = 0.$$

Now, this realization has to be back-transformed into the original one. By multiplying (55) from the left with Q^{-1} and from the right with $Q^{-\mathsf{T}}$ we get

$$\begin{aligned} &Q^{-1} \begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix} \begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix}^\mathsf{T} Q^{-\mathsf{T}} \\ &+ Q^{-1} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} \begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix}^\mathsf{T} Q^{-\mathsf{T}} \\ &+ Q^{-1} \tilde{P}_\ell \begin{bmatrix} B_f \\ B_\infty \end{bmatrix} \begin{bmatrix} B_f \\ B_\infty \end{bmatrix}^\mathsf{T} \tilde{P}_\ell^\mathsf{T} Q^{-\mathsf{T}} \\ = \; &Q^{-1} \begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix} Z^{-1} Z \begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} Z^\mathsf{T} Z^{-\mathsf{T}} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix}^\mathsf{T} Q^{-\mathsf{T}} \\ &+ Q^{-1} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} Z^{-1} Z \begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} Z^\mathsf{T} Z^{-\mathsf{T}} \begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix}^\mathsf{T} Q^{-\mathsf{T}} \\ &+ Q^{-1} \tilde{P}_\ell Q Q^{-1} \begin{bmatrix} B_f \\ B_\infty \end{bmatrix} \begin{bmatrix} B_f \\ B_\infty \end{bmatrix}^\mathsf{T} Q^{-\mathsf{T}} Q^{-\mathsf{T}} \tilde{P}_\ell^\mathsf{T} Q^{-\mathsf{T}} \\ = \; &A \mathcal{G}_{pc} E^\mathsf{T} + E \mathcal{G}_{pc} A^\mathsf{T} + P_\ell B B^\mathsf{T} P_\ell^\mathsf{T} \\ = \; &0, \end{aligned}$$

with the spectral projection

$$P_\ell = Q^{-1}\tilde{P}_\ell Q = Q^{-1}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Q,$$

and the symmetric matrix

$$\mathcal{G}_{pc} = Z\begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} Z^{\mathsf{T}}.$$

Therefore, the conditions (22) and (24) are satisfied. For the condition (22), it holds

$$\begin{aligned} P_r\mathcal{G}_{pc}P_r^{\mathsf{T}} &= Z\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Z^{-1} Z\begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} Z^{\mathsf{T}} Z^{-\mathsf{T}} \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Z^{\mathsf{T}} \\ &= Z\begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} Z^{\mathsf{T}} = \mathcal{G}_{pc}. \end{aligned}$$

The condition (26) for the proper Hankel singular values is then

$$\begin{aligned} \mathcal{G}_{pc}E^{\mathsf{T}}\mathcal{G}_{po}E &= Z\begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} Z^{\mathsf{T}} Z^{-\mathsf{T}} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix}^{\mathsf{T}} Q^{-\mathsf{T}} Q^{\mathsf{T}} \begin{bmatrix} \tilde{G}_{po} & 0 \\ 0 & 0 \end{bmatrix} \\ &\quad \times Q Q^{-1} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} Z^{-1} \\ &= Z\begin{bmatrix} \tilde{G}_{pc} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \tilde{G}_{po} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} Z^{-1} \\ &= Z\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Z^{-1} \\ &= P_r. \end{aligned}$$

For the constraint (30), it holds

$$\begin{aligned} M_0^{\mathsf{T}} C P_r + B^{\mathsf{T}}\mathcal{G}_{po}E &= M_0^{\mathsf{T}}\begin{bmatrix} C_f & C_\infty \end{bmatrix} Z^{-1} Z\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Z^{-1} \\ &\quad + \begin{bmatrix} B_f \\ B_\infty \end{bmatrix}^{\mathsf{T}} Q^{-\mathsf{T}} Q^{\mathsf{T}} \begin{bmatrix} \tilde{G}_{po} & 0 \\ 0 & 0 \end{bmatrix} Q Q^{-1} \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} Z^{-1} \\ &= \left(M_0^{\mathsf{T}}\begin{bmatrix} C_f & C_\infty \end{bmatrix} \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B_f \\ B_\infty \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \tilde{G}_{po} & 0 \\ 0 & 0 \end{bmatrix} \right. \\ &\quad \left. \times \begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} \right) Z^{-1} = 0, \end{aligned}$$

Note, that the conditions (23), (25), (23), (27) and (27) follow analogously by using the dual equations and concepts. Hence, all conditions of the characterization are fulfilled.

"⇐": Now, it is assumed that the conditions (22)–(31) hold. It has to be shown that the resulting transfer function of the linear descriptor system is all-pass. Therefore, a

reformulation of (24) is considered

$$\begin{aligned} P_\ell B B^\mathsf{T} P_\ell^\mathsf{T} &= -A\mathcal{G}_{pc}E^\mathsf{T} - E\mathcal{G}_{pc}A^\mathsf{T} \\ &= -A\mathcal{G}_{pc}E^\mathsf{T} - E\mathcal{G}_{pc}A^\mathsf{T} + sE\mathcal{G}_{pc}E^\mathsf{T} - sE\mathcal{G}_{pc}E^\mathsf{T} \\ &= (sE-A)\mathcal{G}_{pc}E^\mathsf{T} + E\mathcal{G}_{pc}(-sE^\mathsf{T} - A^\mathsf{T}). \end{aligned}$$

The right-hand side of this expression shall be transformed into the form of a transfer function and its para-Hermitian. It holds

$$\begin{aligned} &(sE-A)^{-1}P_\ell B B^\mathsf{T} P_\ell^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1} \\ &= \mathcal{G}_{pc}E^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1} + (sE-A)^{-1}E\mathcal{G}_{pc} \\ \Rightarrow\quad & CP_r(sE-A)^{-1}P_\ell B B^\mathsf{T} P_\ell^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1}P_r^\mathsf{T}C^\mathsf{T} \\ &= CP_r\mathcal{G}_{pc}E^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1}P_r^\mathsf{T}C^\mathsf{T} + CP_r(sE-A)^{-1}E\mathcal{G}_{pc}P_r^\mathsf{T}C^\mathsf{T}. \end{aligned}$$

In the parts with the symmetric matrix $\mathcal{G}_{pc}$, there is also the additional spectral projector P_r. Following (22), we get that $P_r\mathcal{G}_{pc} = \mathcal{G}_{pc}$ and it holds

$$\begin{aligned} &CP_r(sE-A)^{-1}P_\ell B B^\mathsf{T} P_\ell^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1}P_r^\mathsf{T}C^\mathsf{T} \\ &= C\mathcal{G}_{pc}E^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1}P_r^\mathsf{T}C^\mathsf{T} + CP_r(sE-A)^{-1}E\mathcal{G}_{pc}C^\mathsf{T}. \end{aligned}$$

Now, the additional constraint (31) leads to

$$\begin{aligned} &CP_r(sE-A)^{-1}P_\ell B B^\mathsf{T} P_\ell^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1}P_r^\mathsf{T}C^\mathsf{T} \\ &= -M_0 B^\mathsf{T} P_\ell^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1}P_r^\mathsf{T}C^\mathsf{T} - CP_r(sE-A)^{-1}P_\ell B M_0^\mathsf{T}, \end{aligned}$$

and, inserting the definition of the spectral projectors, we get on the left-hand side

$$\begin{aligned} &CP_r(sE-A)^{-1}P_\ell B B^\mathsf{T} P_\ell^\mathsf{T}(-sE^\mathsf{T}-A^\mathsf{T})^{-1}P_r^\mathsf{T}C^\mathsf{T} \\ &= CZ\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Z^{-1}(sE-A)^{-1}Q^{-1}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} QB \\ &\quad \times B^\mathsf{T}Q^\mathsf{T}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Q^{-\mathsf{T}}(-sE^\mathsf{T}-A^\mathsf{T})^{-1}Z^{-\mathsf{T}}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Z^\mathsf{T}C^\mathsf{T} \\ &= CZ\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} (sQEZ - QAZ)^{-1}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} QB \\ &\quad \times B^\mathsf{T}Q^\mathsf{T}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} (-sZ^\mathsf{T}E^\mathsf{T}Q^\mathsf{T} - Z^\mathsf{T}A^\mathsf{T}Q^\mathsf{T})^{-1}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} Z^\mathsf{T}C^\mathsf{T} \\ &= \begin{bmatrix} C_f & C_\infty \end{bmatrix}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix}\left(s\begin{bmatrix} E_f & 0 \\ 0 & E_\infty \end{bmatrix} - \begin{bmatrix} A_f & 0 \\ 0 & A_\infty \end{bmatrix}\right)^{-1}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} B_f \\ B_\infty \end{bmatrix} \\ &\quad \times \begin{bmatrix} B_f^\mathsf{T} & B_\infty^\mathsf{T} \end{bmatrix}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix}\left(-s\begin{bmatrix} E_f^\mathsf{T} & 0 \\ 0 & E_\infty^\mathsf{T} \end{bmatrix} - \begin{bmatrix} A_f^\mathsf{T} & 0 \\ 0 & A_\infty^\mathsf{T} \end{bmatrix}\right)^{-1}\begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} C_f^\mathsf{T} \\ C_\infty^\mathsf{T} \end{bmatrix} \\ &= \left(C_f(sE_f - A_f)^{-1}B_f + 0\cdot(sE_\infty - A_\infty)^{-1}\cdot 0\right) \\ &\quad \times \left(B_f^\mathsf{T}(-sE_f^\mathsf{T} - A_f^\mathsf{T})^{-1}C_f^\mathsf{T} + 0\cdot(-sE_\infty^\mathsf{T} - A_\infty^\mathsf{T})^{-1}\cdot 0\right) \\ &= C_f(sE_f - A_f)^{-1}B_f B_f^\mathsf{T}(-sE_f^\mathsf{T} - A_f^\mathsf{T})^{-1}C_f^\mathsf{T}. \end{aligned}$$

For the right-hand side we get

$$\begin{aligned} &-M_0 B^{\mathsf{T}} P_\ell^{\mathsf{T}} (-sE^{\mathsf{T}} - A^{\mathsf{T}})^{-1} P_r^{\mathsf{T}} C^{\mathsf{T}} - C P_r (sE - A)^{-1} P_\ell B M_0^{\mathsf{T}} \\ = \; &-M_0 B_f^{\mathsf{T}} (-sE_f^{\mathsf{T}} - A_f^{\mathsf{T}})^{-1} C_f^{\mathsf{T}} - C_f (sE_f - A_f)^{-1} B_f M_0^{\mathsf{T}}. \end{aligned}$$

Using the above expressions, the all-pass condition is satisfied

$$\begin{aligned} G(s)G^{\mathsf{T}}(-s) &= (C(sE - A)^{-1}B + D)(B^{\mathsf{T}}(sE^{\mathsf{T}} - A^{\mathsf{T}})^{-1}C^{\mathsf{T}} + D^{\mathsf{T}}) \\ &= (C_f(sE_f - A_f)^{-1}B_f + M_0)(B_f^{\mathsf{T}}(sE_f^{\mathsf{T}} - A_f^{\mathsf{T}})^{-1}C_f^{\mathsf{T}} + M_0^{\mathsf{T}}) \\ &= C_f(sE_f - A_f)^{-1}B_f B_f^{\mathsf{T}}(sE_f^{\mathsf{T}} - A_f^{\mathsf{T}})^{-1}C_f^{\mathsf{T}} \\ &\quad + M_0 B_f^{\mathsf{T}}(sE_f^{\mathsf{T}} - A_f^{\mathsf{T}})^{-1}C_f^{\mathsf{T}} + C_f(sE_f - A_f)^{-1}B_f M_0^{\mathsf{T}} + M_0 M_0^{\mathsf{T}} \\ &= M_0 M_0^{\mathsf{T}} \\ &= I_m. \end{aligned}$$

□

References

1. Adamjan, V.M., Arov, D.Z., Kreĭn, M.G.: Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. Mathematics of the USSR-Sbornik **15**(1), 31–73 (1971). https://doi.org/10.1070/SM1971v015n01ABEH001531
2. Antoulas, A.C.: Approximation of large-scale dynamical systems, Adv. Des Control, vol. 6. SIAM Publications, Philadelphia (2005). https://doi.org/10.1137/1.9780898718713
3. Bai, Z., Demmel, J., Gu, M.: An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems. J Numer Math. **76**(3), 279–308 (1997). https://doi.org/10.1007/s002110050264
4. Benner, P., Quintana-Ortí, E.S.: Model reduction based on spectral projection methods. In: Benner, P., Mehrmann, V., Sorensen, D. (eds.) Dimension Reduction of Large-Scale Systems, Lect. Notes Comput. Sci. Eng., vol. 45, pp. 5–45. Springer, Berlin (2005). https://doi.org/10.1007/3-540-27909-1_1
5. Benner, P., Quintana-Ortí, E.S., Quintana-Ortí, G.: Parallel model reduction of large-scale linear descriptor systems via balanced truncation. In: Daydé, M., Dongarra, J.J., Hernández, V., Palma, J.M.L.M. (eds.) High Performance Computing for Computational Science - VECPAR 2004, Lecture Notes in Comput. Sci., vol. 3402, pp. 340–353. Springer, Berlin (2005). https://doi.org/10.1007/11403937_27
6. Benner, P., Stykel, T.: Model order reduction for differential-algebraic equations: a survey. In: Ilchmann, A., Reis, T. (eds.) Surveys in Differential-Algebraic Equations IV, Differential-Algebraic Equations Forum, pp. 107–160. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-46618-7_3
7. Benner, P., Werner, S.W.R.: MORLAB-3.0 – model order reduction laboratory. https://doi.org/10.5281/zenodo.842659. See also: https://www.mpi-magdeburg.mpg.de/projects/morlab (2017)
8. Benner, P., Werner, S.W.R.: On the transformation formulas of the Hankel-norm approximation. Proc Appl Math Mech. **17**(1), 823–824 (2017). https://doi.org/10.1002/pamm.201710379
9. Benner, P., Werner, S.W.R.: Model reduction of descriptor systems with the MORLAB toolbox. IFAC-PapersOnLine 9th Vienna International Conference on Mathematical Modelling MATHMOD 2018, Vienna Austria, 21–23 February 2018 **51**(2), 547–552 (2018). https://doi.org/10.1016/j.ifacol.2018.03.092
10. Cao, X., Saltik, M.B., Weiland, S.: Hankel model reduction for descriptor systems. In: 2015 54th IEEE Conference on Decision and Control (CDC), pp. 4668–4673 (2015). https://doi.org/10.1109/CDC.2015.7402947
11. Freitas, F., Rommes, J., Martins, N.: Gramian-based reduction method applied to large sparse power system descriptor models. IEEE Trans Power Del. **23**(3), 1258–1270 (2008). https://doi.org/10.1109/TPWRS.2008.926693

12. Glover, K.: All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error norms. Internat J. Control **39**(6), 1115–1193 (1984). https://doi.org/10.1080/00207178408933239
13. Golub, G.H., Van Loan, C.F. Matrix Computations, 4th edn. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore (2013)
14. Gugercin, S., Stykel, T., Wyatt, S.: Model reduction of descriptor systems by interpolatory projection methods. SIAM J. Sci. Comput. **35**(5), B1010–B1033 (2013). https://doi.org/10.1137/130906635
15. Kågström, B., Van Dooren, P.: A generalized state-space approach for the additive decomposition of a transfer matrix. Numer. Lin. Alg. Appl. **1**(2), 165–181 (1992)
16. Mehrmann, V., Stykel, T.: Balanced truncation model reduction for large-scale systems in descriptor form. In: Benner, P., Mehrmann, V., Sorensen, D.C. (eds.) Dimension Reduction of Large-Scale Systems, Lect. Notes Comput. Sci. Eng., vol. 45, pp. 83–115. Springer, Berlin (2005). https://doi.org/10.1007/3-540-27909-1_3
17. Saak, J., Köhler, M., Benner, P.: M-M.E.S.S.-1.0.1 – the matrix equations sparse solvers library. https://doi.org/10.5281/zenodo.50575. See also: https://www.mpi-magdeburg.mpg.de/projects/mess (2016)
18. Schmidt, M.: Systematic discretization of input/output maps and other contributions to the control of distributed parameter systems. Ph.D. Thesis. Technische Universität, Berlin (2007). https://doi.org/10.14279/depositonce-1600
19. Sokolov, V.: Contributions to the minimal realization problem for descriptor systems. Dissertation, Fakultät für Mathematik, TU Chemnitz, Chemnitz (2006). http://nbn-resolving.de/urn:nbn:de:swb:ch1-200600965
20. Stewart, G.W., Sun, J.G.: Matrix Perturbation Theory. Academic Press, New York (1990)
21. Stykel, T.: Analysis and Numerical Solution of generalized Lyapunov equations. Dissertation, TU Berlin (2002). http://webdoc.sub.gwdg.de/ebook/e/2003/tu-berlin/stykel_tatjana.pdf
22. Stykel, T.: Gramian-based model reduction for descriptor systems. Math Control Signals Syst. **16**(4), 297–319 (2004). https://doi.org/10.1007/s00498-004-0141-4
23. Stykel, T.: Balanced truncation model reduction for semidiscretized Stokes equation. Linear Algebra Appl. **415**(2–3), 262–289 (2006). https://doi.org/10.1016/j.laa.2004.01.015
24. Stykel, T.: Low-rank iterative methods for projected generalized Lyapunov equations. Electron Trans Numer Anal. **30**, 187–202 (2008)
25. Werner, S.: Hankel-norm approximation of descriptor systems. Master's thesis, Otto-von-Guericke-Universität, Magdeburg, Germany (2016). http://nbn-resolving.de/urn:nbn:de:gbv:ma9:1-8845