arXiv:2402.14716v1 [math.DS] 22 Feb 2024

Balanced Truncation of Descriptor Systems with a Quadratic Output

Jennifer Przybilla[*, ▲] Igor Pontes Duff[*, †]
Pawan Goyal[*, △] Peter Benner[*, **, ⊛]

[*] *Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany.*
[**] *Otto von Guericke University Magdeburg, Fakultät für Mathematik, Universitätsplatz 2, 39106 Magdeburg, Germany.*
[▲]Email: przybilla@mpi-magdeburg.mpg.de, ORCID: 0000-0002-8703-8735
[†]Email: pontes@mpi-magdeburg.mpg.de, ORCID: 0000-0001-6433-6142
[△]Email: goyalp@mpi-magdeburg.mpg.de, ORCID: 0000-0003-3072-7780
[⊛]Email: benner@mpi-magdeburg.mpg.de, ORCID: 0000-0003-3362-4103

Abstract: This work discusses model reduction for differential-algebraic systems with quadratic output equations. Under mild conditions, these systems can be transformed into a Weierstraß canonical form and, thus, be decoupled into differential equations and algebraic equations. The corresponding decoupled states are referred to as proper and improper states. Due to the quadratic function of the state as an output, the proper and improper states are coupled in the output equation, which imposes a challenge from a model reduction viewpoint. Keeping the coupling in mind, our goal in this work is to find important subspaces of the proper and improper states and to reduce the system accordingly. To that end, we first propose the system's matrices, the so-called Gramians, to characterize the system's dominant subspaces. We pay particular attention to the computation of the observability Gramians that take into account the nonlinear coupling between the proper and the improper states. We furthermore show that the proposed Gramians are related to certain kernel functions, which are used to identify important subspaces. This allows us to propose a reduction algorithm to obtain reduced-order systems by removing the subspaces that are difficult to reach, as well as, difficult to observe. Moreover, we quantify the error between the full-order and reduced-order models and demonstrate the proposed methodology using three numerical experiments.

Keywords: model order reduction, balanced truncation, differential-algebraic systems, quadratic output systems, system Gramians, reduced-order models.

Novelty statement:

- We discuss a balanced truncation approach for linear differential-algebraic equations with a quadratic output.
- For this, we propose new Gramians, characterizing the importance of the state from the input-output view-point and show their connections to corresponding kernel functions that are used to identify important subspaces.
- Moreover, we discuss an algorithm to construct reduced-order models using these Gramians, and characterize the error between the full-order and reduced-order models due to the truncation.
- The performance of the proposed methodology is demonstrated using three numerical examples.

1 Introduction

In this paper, we discuss balanced truncation for a class of descriptor systems with a quadratic output function of the form

$$\begin{aligned} \mathbf{E}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}(t), \end{aligned} \tag{1}$$

where $\mathbf{E},\ \mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, and $\mathbf{M} \in \mathbb{R}^{n\times n}$, where the matrix $\mathbf{E}$ is singular, and $\mathbf{M}$ is assumed to be symmetric, i.e., $\mathbf{M} = \mathbf{M}^{\mathrm{T}}$. Additionally, we assume that the matrix pencil $s\mathbf{E} - \mathbf{A}$ is regular—that is, $\det(s\mathbf{E} - \mathbf{A})$ is not the zero polynomial. The input vector, the state vector, and the output are denoted by $\mathbf{u}(t) \in \mathbb{R}^m$, $\mathbf{x}(t) \in \mathbb{R}^n$ and $\mathbf{y}(t) \in \mathbb{R}$, respectively. In the following, we also assume that all the finite eigenvalues of the matrix pencil $s\mathbf{E} - \mathbf{A}$ lie in the negative half-plane, i.e., the system is asymptotically stable. Note, that these systems can be interpreted as a special class of Wiener models.

Differential algebraic systems (DAEs) arise when for example, electrical circuits, thermal and diffusion processes, or multibody systems are modeled by methods such as finite elements or finite volumes. These systems involve dynamic constraints that lead to algebraic equations, and therefore analysis tools must be developed for them. The system (1) appears particularly while investigating the variance or deviation of the state variable from a certain reference point, which can be represented as a quadratic function of the state.

Models exhibiting complex dynamic behavior, or coming from PDEs discretization, are often high-fidelity models, i.e., the dimension of the state vector n is large, which makes the engineering design process computationally infeasible. As a remedy, we seek to employ model reduction techniques that allow us to construct a low-dimensional model which closely resembles the dynamic behaviors of the high-fidelity model. Our goal, in this paper, is to construct reduced-order models for the original models (1) while preserving the original structure. Precisely, we aim to determine the reduced-order models of the form

$$\widehat{\mathbf{E}}\dot{\widehat{\mathbf{x}}}(t) = \widehat{\mathbf{A}}\widehat{\mathbf{x}}(t) + \widehat{\mathbf{B}}\mathbf{u}(t), \tag{2a}$$

$$\widehat{\mathbf{y}}(t) = \widehat{\mathbf{x}}(t)^{\mathrm{T}}\widehat{\mathbf{M}}\widehat{\mathbf{x}}(t), \tag{2b}$$

where $\widehat{\mathbf{E}},\ \widehat{\mathbf{A}} \in \mathbb{R}^{r\times r}$, $\widehat{\mathbf{B}} \in \mathbb{R}^{r\times m}$ and $\widehat{\mathbf{M}} \in \mathbb{R}^{r\times r}$, with $\widehat{\mathbf{M}} = \widehat{\mathbf{M}}^{\mathrm{T}}$ and $r \ll n$. We obtain the reduced matrices in (2) by multiplying the matrices of system (1) using two projection matrices, namely, $\mathbf{W}_{\mathrm{r}},\ \mathbf{T}_{\mathrm{r}} \in \mathbb{R}^{n\times r}$, i.e.,

$$\widehat{\mathbf{E}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{E}\mathbf{T}_{\mathrm{r}}, \quad \widehat{\mathbf{A}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{A}\mathbf{T}_{\mathrm{r}}, \quad \widehat{\mathbf{B}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{B}, \quad \widehat{\mathbf{M}} = \mathbf{T}_{\mathrm{r}}^{\mathrm{T}}\mathbf{M}\mathbf{T}_{\mathrm{r}}.$$

Furthermore, the reduced state and the approximated output are denoted by $\widehat{\mathbf{x}}(t) \in \mathbb{R}^r$ and $\widehat{\mathbf{y}}(t) \in \mathbb{R}$, respectively. The reduced-order model (2) shall approximate the input–to–output behavior of the full-order model (1), i.e., $\|\mathbf{y} - \widehat{\mathbf{y}}\| \leq \mathtt{tol}$, where $\mathtt{tol}$ is a user-defined tolerance, for all admissible inputs $\mathbf{u}$.

For ordinary differential equation (ODE) systems with a linear output equation, there exist several methods to construct reduced-order models, e.g., singular value-based approaches such as balanced truncation [8, 20, 28] and Hankel norm approximations [13]. Moreover, moment matching methods [4, 14, 18] and Krylov subspace methods, e.g., the iterative rational Krylov algorithm (IRKA) [8, 12, 14, 15] are frequently used. A vast overview of these methods is given, e.g., in [4, 6–8].

All of the above-mentioned methods treat the case in which $\mathbf{E}$ is nonsingular, and, therefore, are not directly applicable to the DAE case. There are several challenges that arise due to the algebraic equations. Since the matrix $\mathbf{E}$ is singular, the transfer function $\boldsymbol{\mathcal{G}}(s) := \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$, defining the input-output mapping in the frequency domain, can have a non-zero polynomial part that needs to be preserved. Therefore, a model reduction scheme for DAEs must preserve the polynomial part of its transfer function in the construction of reduced-order models. This issue is addressed in, e.g., [11, 15, 19, 24]. Several existing methods that deal with the DAE case include interpolatory projection methods [1–3, 15] and balancing-based methods [9, 16, 19, 24, 25]. In this work, we focus on a balancing-based method. In balanced truncation (BT), one has to solve large

generalized Lyapunov equations, also referred to as *projected Lyapunov equations*, that act in the left and right deflating subspaces corresponding to the finite or infinite eigenvalues of the matrix pencil $s\mathbf{E} - \mathbf{A}$. It requires to define the corresponding projection matrices that describe these deflating subspaces. However, such projection matrices are difficult to form explicitly. Even if one manages, they can destroy the sparsity of the original matrices, thus, increasing the computational burden. However, the structure of the DAE systems is often known and can be used to define, and implicitly apply the projection matrices in theory without the need of explicitly forming or multiplying by these projection matrices. For details, we refer to [9, 10, 16, 22, 26].

However, BT is not directly applicable to the case of quadratic output functionals since the observability space is not of the same form as in the linear output case. Hence. the observability Gramian, defined in [19], is not usable. In this work, we develop BT for DAEs with quadratic output equations. To that end, we derive new Gramians and corresponding kernel functions that allow us to characterize the controllability and observability behavior. It is worth mentioning that in [5], the authors derived Gramians corresponding to ODE systems (meaning $\mathbf{E} = \mathbf{I}$ in (1)) with quadratic output equations. However, the methodology proposed in [5] cannot be directly applied to DAEs due to the singularity of the matrix $\mathbf{E}$. Therefore, there is a necessity to modify BT to incorporate the differential-algebraic structure. Precisely, we tailor the Gramians, corresponding to the proper and improper states of the system in (1) that describe the controllability and observability spaces. Based on this, we proposed a balancing scheme to determine projection matrices $\mathbf{W}_\mathrm{r}$ and $\mathbf{T}_\mathrm{r}$, leading to the construction of reduced-order models.

The remainder of the paper is organized as follows. In Section 2, we briefly recap BT for differential-algebraic systems with linear output equations from [24]. Moreover, we provide an overview of the Gramians for ODE systems with quadratic output equations, proposed in [5]. In Section 3, we then derive the Gramians for the system (1). An algorithm to construct reduced-order models by truncating unimportant subspaces is discussed in Section 4. In Section 5, we derive an error estimator that bounds the output error between the original and reduced-order models. Furthermore, in Section 6, we extend the theory presented in Sections 3 to 5 for the systems with multiple inputs. We demonstrate the efficiency of the method in Section 7 using three numerical examples. We conclude the paper with a summary and future directions.

2 Preliminaries

In this section, we summarize previous works that form the basis for this paper. We begin by discussing the Weierstraß-canonical form for DAEs in Subsection 2.1 followed by the introduction to the BT method for DAE systems with linear output equation in Subsection 2.2 and ODE systems with quadratic output equation in Subsection 2.3.

2.1 Weierstraß-canonical form

We consider descriptor systems with a quadratic output function described in (1). According to [17], there exist matrices $\mathbf{W}$ and $\mathbf{T}$ that transform the differential-algebraic equation of the system (1) into a Weierstraß-canonical form, i.e.,

$$\mathbf{E} = \mathbf{W}\begin{bmatrix}\mathbf{I}_{n_f} & 0\\ 0 & \mathbf{N}\end{bmatrix}\mathbf{T}, \quad \mathbf{A} = \mathbf{W}\begin{bmatrix}\mathbf{J} & 0\\ 0 & \mathbf{I}_{n_\infty}\end{bmatrix}\mathbf{T}, \quad \mathbf{B} = \mathbf{W}\begin{bmatrix}\mathbf{B}_1\\ \mathbf{B}_2\end{bmatrix}, \quad \mathbf{M} = \mathbf{T}^\mathrm{T}\begin{bmatrix}\mathbf{M}_{11} & \mathbf{M}_{12}\\ \mathbf{M}_{12}^\mathrm{T} & \mathbf{M}_{22}\end{bmatrix}\mathbf{T},$$

with n_f and n_∞ being the respective numbers of the finite and infinite eigenvalues of the matrix pencil $s\mathbf{E} - \mathbf{A}$. The matrix $\mathbf{J} \in \mathbb{R}^{n_f \times n_f}$ is in Jordan normal form, and $\mathbf{N} \in \mathbb{R}^{n_\infty \times n_\infty}$ is nilpotent of nilpotency index ν. Typically, the index ν is referred to as the index of the system (1) as well. Moreover, we define the spectral projection matrices

$$\mathbf{P}_\mathrm{r} = \mathbf{T}^{-1}\begin{bmatrix}\mathbf{I}_{n_f} & 0\\ 0 & 0\end{bmatrix}\mathbf{T} \quad \text{and} \quad \mathbf{P}_\mathrm{l} = \mathbf{W}\begin{bmatrix}\mathbf{I}_{n_f} & 0\\ 0 & 0\end{bmatrix}\mathbf{W}^{-1} \tag{3}$$

onto the right and left deflating subspaces of the pencil $\lambda\mathbf{E} - \mathbf{A}$, corresponding to the finite eigenvalues. By multiplying the system (1) from the left by $\mathbf{W}^{-1}$ and replacing $\mathbf{x}(t) =: \mathbf{T}^{-1}\begin{bmatrix}\mathbf{x}_1(t)\\ \mathbf{x}_2(t)\end{bmatrix}$,

we obtain the following system:

$$\dot{\mathbf{x}}_1(t) = \mathbf{J}\mathbf{x}_1(t) + \mathbf{B}_1\mathbf{u}(t), \tag{4a}$$
$$\mathbf{N}\dot{\mathbf{x}}_2(t) = \mathbf{x}_2(t) + \mathbf{B}_2\mathbf{u}(t), \tag{4b}$$
$$\mathbf{y}(t) = \begin{bmatrix}\mathbf{x}_1(t)\\ \mathbf{x}_2(t)\end{bmatrix}^{\mathrm{T}} \begin{bmatrix}\mathbf{M}_{11} & \mathbf{M}_{12}\\ \mathbf{M}_{12}^{\mathrm{T}} & \mathbf{M}_{22}\end{bmatrix} \begin{bmatrix}\mathbf{x}_1(t)\\ \mathbf{x}_2(t)\end{bmatrix}. \tag{4c}$$

The system (4) provides the decoupled proper and improper states $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$. Additionally, the solution trajectories of (4a) and (4b) are given as follows:

$$\mathbf{x}_1(t) = \int_0^t e^{\mathbf{J}(t-\tau)}\mathbf{B}_1\mathbf{u}(\tau)\mathrm{d}\tau, \qquad \mathbf{x}_2(t) = \sum_{k=0}^{\nu-1} -\mathbf{N}^k\mathbf{B}_2\mathbf{u}^{(k)}(t), \tag{5}$$

where $\mathbf{u}^{(k)}(t)$ describes the k-th derivative of the function $\mathbf{u}(\cdot)$ evaluated in the time variable t. It is easy to see that initial conditions must satisfy $\mathbf{x}_2(0) = \sum_{k=0}^{\nu-1} -\mathbf{N}^k\mathbf{B}_2\mathbf{u}^{(k)}(0)$. In this case the initial state is called *consistent*.

Furthermore, we define

$$\mathbf{F}_J(t) := \mathbf{T}^{-1}\begin{bmatrix}e^{\mathbf{J}t} & 0\\ 0 & 0\end{bmatrix}\mathbf{W}^{-1} \qquad \text{and} \qquad \mathbf{F}_N(k) := \mathbf{T}^{-1}\begin{bmatrix}0 & 0\\ 0 & -\mathbf{N}^k\end{bmatrix}\mathbf{W}^{-1} \tag{6}$$

and transform $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ into the original state space of system (1) to obtain the proper and improper states

$$\mathbf{x}_{\mathrm{p}}(t) = \int_0^t \mathbf{F}_J(t-\tau)\mathbf{B}\mathbf{u}(\tau)\mathrm{d}\tau \quad \text{and} \quad \mathbf{x}_{\mathrm{i}}(t) = \sum_{k=0}^{\nu-1} \mathbf{F}_N(k)\mathbf{B}\mathbf{u}^{(k)}(t) \tag{7}$$

with $\mathbf{x}(t) = \mathbf{x}_{\mathrm{p}}(t) + \mathbf{x}_{\mathrm{i}}(t)$.

In this paper, we aim to define Gramians that describe the controllability and the observability of the proper and improper states $\mathbf{x}_{\mathrm{p}}(t)$ and $\mathbf{x}_{\mathrm{i}}(t)$. Using these Gramians, we then derive tailored energy functional estimations that are used to identify irrelevant states. Note that the Weierstraß-canoncial form will only serve as a tool for analysis, but will not be computed in practice as its numerical determination is known to be difficult.

2.2 BT for DAEs with linear output

Here, we describe model reduction by BT for descriptor systems as introduced in [19]. We consider the continuous-time descriptor system with a linear output equation

$$\begin{aligned}\mathbf{E}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = 0,\\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t),\end{aligned} \tag{8}$$

where the matrices $\mathbf{E}$, $\mathbf{A}$, $\mathbf{B}$, and the vectors $\mathbf{x}(t)$, $\mathbf{u}(t)$ are as in the system (1). The output equation provides the output, $\mathbf{y}(t) \in \mathbb{R}^q$, that results from the output matrix $\mathbf{C} \in \mathbb{R}^{q\times n}$ and the state $\mathbf{x}(t)$. We assume that system (8) is asymptotically stable, i.e., all finite eigenvalues of the matrix pencil $s\mathbf{E} - \mathbf{A}$ lie in the negative half-plane.

As described in the above subsection, the state $\mathbf{x}(t)$ of system (8) can be decomposed as $\mathbf{x}(t) = \mathbf{x}_{\mathrm{p}}(t) + \mathbf{x}_{\mathrm{i}}(t)$ with $\mathbf{x}_{\mathrm{p}}(t)$ and $\mathbf{x}_{\mathrm{i}}(t)$ as defined in (7). We define the corresponding input–to–state mappings

$$\mathcal{C}_{\mathrm{p}}(t) := \mathbf{F}_J(t)\mathbf{B} \qquad \text{and} \qquad \mathcal{C}_{\mathrm{i}}(k) := \mathbf{F}_N(k)\mathbf{B}$$

of the system (8) that describe the controllability of the corresponding states. With the help of these mappings, we can define the corresponding controllability Gramians of the system (8). They are defined as $\mathcal{P}_{\mathrm{p}} := \int_0^\infty \mathcal{C}_{\mathrm{p}}(t)\mathcal{C}_{\mathrm{p}}(t)^{\mathrm{T}}\mathrm{d}t$ and $\mathcal{P}_{\mathrm{i}} := \sum_{k=0}^{\nu-1} \mathcal{C}_{\mathrm{i}}(k)\mathcal{C}_{\mathrm{i}}(k)^{\mathrm{T}}$ and result in the following Gramian expressions:

$$\mathcal{P}_{\mathrm{p}} := \int_0^\infty \mathbf{F}_J(t)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t)^{\mathrm{T}}\mathrm{d}t, \quad \mathcal{P}_{\mathrm{i}} := \sum_{k=0}^{\nu-1} \mathbf{F}_N(k)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{F}_N(k)^{\mathrm{T}}. \tag{9}$$

The matrices $\mathcal{P}_\mathrm{p}$ and $\mathcal{P}_\mathrm{i}$ span the controllability space of the states $\mathbf{x}_\mathrm{p}(t)$ and $\mathbf{x}_\mathrm{i}(t)$. Furthermore, inserting the definitions of $\mathbf{F}_J(t)$ and $\mathbf{F}_N(k)$ in (9) indicates the Gramians to be of the following form

$$\mathcal{P}_\mathrm{p} = \mathbf{T}^{-1}\begin{bmatrix}\mathcal{P}_1 & 0\\ 0 & 0\end{bmatrix}\mathbf{T}^{-\mathrm{T}}, \qquad \mathcal{P}_\mathrm{i} = \mathbf{T}^{-1}\begin{bmatrix}0 & 0\\ 0 & \mathcal{P}_2\end{bmatrix}\mathbf{T}^{-\mathrm{T}} \tag{10}$$

where $\mathcal{P}_1 = \int_0^\infty e^{\mathbf{J}t}\mathbf{B}_1\mathbf{B}_1^\mathrm{T}e^{\mathbf{J}^\mathrm{T}t}\mathrm{d}t$ and $\mathcal{P}_2 = \sum_{k=0}^{\nu-1}\mathbf{N}^k\mathbf{B}_2\mathbf{B}_2^\mathrm{T}(\mathbf{N}^k)^\mathrm{T}$ are the controllability Gramians corresponding to $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$. Using the controllability Gramians, we can characterize the states that are difficult to reach or even unreachable, which play an important role in the reduction of the system.

The Gramians $\mathcal{P}_\mathrm{p}$ and $\mathcal{P}_\mathrm{i}$ satisfy the following time-continuous and time-discrete projected Lyapunov equation

$$\mathbf{E}\mathcal{P}_\mathrm{p}\mathbf{A}^\mathrm{T} + \mathbf{A}\mathcal{P}_\mathrm{p}\mathbf{E}^\mathrm{T} = -\mathbf{P}_\mathrm{l}\mathbf{B}\mathbf{B}^\mathrm{T}\mathbf{P}_\mathrm{l}^\mathrm{T}, \qquad \mathcal{P}_\mathrm{p} = \mathbf{P}_\mathrm{r}\mathcal{P}_\mathrm{p}\mathbf{P}_\mathrm{r}^\mathrm{T}, \tag{11a}$$

$$\mathbf{A}\mathcal{P}_\mathrm{i}\mathbf{A}^\mathrm{T} - \mathbf{E}\mathcal{P}_\mathrm{i}\mathbf{E}^\mathrm{T} = (I-\mathbf{P}_\mathrm{l})\mathbf{B}\mathbf{B}^\mathrm{T}(I-\mathbf{P}_\mathrm{l})^\mathrm{T}, \qquad 0 = \mathbf{P}_\mathrm{r}\mathcal{P}_\mathrm{i}\mathbf{P}_\mathrm{r}^\mathrm{T} \tag{11b}$$

where the projection matrices $\mathbf{P}_\mathrm{l}$ and $\mathbf{P}_\mathrm{r}$ are as defined in (3).

To describe the observability of the system (8), we consider the state–to–output mappings

$$\mathcal{O}_\mathrm{p}(t) := \mathbf{C}\mathbf{F}_J(t) \qquad \text{and} \qquad \mathcal{O}_\mathrm{i}(k) := \mathbf{C}\mathbf{F}_N(k).$$

These mappings are used to describe the observability of certain states and are therefore used to define the proper and improper observability Gramians as $\mathcal{Q}_\mathrm{p} := \int_0^\infty \mathcal{O}_\mathrm{p}(t)^\mathrm{T}\mathcal{O}_\mathrm{p}(t)\mathrm{d}t$ and $\mathcal{Q}_\mathrm{i} := \sum_{k=0}^{\nu-1}\mathcal{O}_\mathrm{i}(k)^\mathrm{T}\mathcal{O}_\mathrm{i}(k)$ such that we obtain

$$\mathcal{Q}_\mathrm{p} := \int_0^\infty \mathbf{F}_J(t)^\mathrm{T}\mathbf{C}^\mathrm{T}\mathbf{C}\mathbf{F}_J(t)\mathrm{d}\tau, \qquad \mathcal{Q}_\mathrm{i} := \sum_{k=0}^{\nu-1}\mathbf{F}_N(k)^\mathrm{T}\mathbf{C}^\mathrm{T}\mathbf{C}\mathbf{F}_N(k).$$

The goal of BT is to determine simultaneously the states which are both hard to reach and hard to observe. In general, the states corresponding to small singular values of the controllability Gramians do not coincide with the states corresponding to small singular values of the observability Gramians. Therefore, we need to balance the system first, i.e., we transform the system to obtain a balanced one.

Definition 2.1. We call the system (8) *balanced* if the Gramians satisfy

$$\mathcal{P}_\mathrm{p} = \mathcal{Q}_\mathrm{p} = \begin{bmatrix}\mathbf{\Sigma} & 0\\ 0 & 0\end{bmatrix}, \qquad \mathcal{P}_\mathrm{i} = \mathcal{Q}_\mathrm{i} = \begin{bmatrix}0 & 0\\ 0 & \mathbf{\Theta}\end{bmatrix},$$

where $\mathbf{\Sigma} = \operatorname{diag}\left(\sigma_1, \ldots, \sigma_{n_f}\right)$, and $\mathbf{\Theta} = \operatorname{diag}\left(\theta_1, \ldots, \theta_{n_\infty}\right)$.

Since all Gramians are symmetric and positive semi-definite, there exist factorizations

$$\mathcal{P}_\mathrm{p} = \mathbf{R}_\mathrm{p}\mathbf{R}_\mathrm{p}^\mathrm{T}, \quad \mathcal{Q}_\mathrm{p} = \mathbf{L}_\mathrm{p}^\mathrm{T}\mathbf{L}_\mathrm{p}, \quad \mathcal{P}_\mathrm{i} = \mathbf{R}_\mathrm{i}\mathbf{R}_\mathrm{i}^\mathrm{T}, \quad \mathcal{Q}_\mathrm{i} = \mathbf{L}_\mathrm{i}^\mathrm{T}\mathbf{L}_\mathrm{i}.$$

Next, We compute the singular value decompositions

$$\mathbf{L}_\mathrm{p}\mathbf{E}\mathbf{R}_\mathrm{p} = \mathbf{U}_\mathrm{p}\mathbf{\Sigma}\mathbf{V}_\mathrm{p}^\mathrm{T} = \begin{bmatrix}\mathbf{U}_{\mathrm{p},1} & \mathbf{U}_{\mathrm{p},2}\end{bmatrix}\begin{bmatrix}\mathbf{\Sigma}_1 & \\ & \mathbf{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{p},1}^\mathrm{T}\\ \mathbf{V}_{\mathrm{p},2}^\mathrm{T}\end{bmatrix},$$

$$\mathbf{L}_\mathrm{i}\mathbf{A}\mathbf{R}_\mathrm{i} = \mathbf{U}_\mathrm{i}\mathbf{\Theta}\mathbf{V}_\mathrm{i}^\mathrm{T} = \begin{bmatrix}\mathbf{U}_{\mathrm{i},1} & \mathbf{U}_{\mathrm{i},2}\end{bmatrix}\begin{bmatrix}\mathbf{\Theta}_1 & \\ & 0\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{i},1}^\mathrm{T}\\ \mathbf{V}_{\mathrm{i},2}^\mathrm{T}\end{bmatrix}$$

where $\mathbf{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$, $\sigma_1 \geq \cdots \geq \sigma_n$, includes the proper Hankel singular values of the system. The proper states that are simultaneously difficult to reach and to observe correspond to the smallest Hankel singular values that are $\mathbf{\Sigma}_2$. We truncate the corresponding states that lie in the spaces spanned by $\mathbf{U}_{p,2}$ and $\mathbf{V}_{p,2}$ by building the projection matrices

$$\mathbf{W}_\mathrm{r} = [\mathbf{L}_\mathrm{p}^\mathrm{T}\mathbf{U}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}},\ \mathbf{L}_\mathrm{i}^\mathrm{T}\mathbf{U}_{\mathrm{i},1}\mathbf{\Theta}^{-\frac{1}{2}}], \qquad \mathbf{T}_\mathrm{r} = [\mathbf{R}_\mathrm{p}^\mathrm{T}\mathbf{V}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}},\ \mathbf{R}_\mathrm{i}^\mathrm{T}\mathbf{V}_{\mathrm{i},1}\mathbf{\Theta}^{-\frac{1}{2}}].$$

Note that additionally improper states that correspond to zero singular values in $\mathbf{\Theta}$, i.e., the states that lie in the spaces spanned by $\mathbf{U}_{\mathrm{i},2}$ and $\mathbf{V}_{\mathrm{i},2}$, are truncated. Multiplying the matrices of the full-order model in (8) by $\mathbf{W}_{\mathrm{r}}$ and $\mathbf{T}_{\mathrm{r}}$ leads to a reduced-order model

$$\begin{aligned}\widehat{\mathbf{E}}\dot{\widehat{\mathbf{x}}}(t) &= \widehat{\mathbf{A}}\widehat{\mathbf{x}}(t) + \widehat{\mathbf{B}}\mathbf{u}(t),\\ \widehat{\mathbf{y}}(t) &= \widehat{\mathbf{C}}\widehat{\mathbf{x}}(t),\end{aligned}$$

where it can be shown that $\widehat{\mathbf{E}} = \begin{bmatrix}\mathbf{I} & 0\\ 0 & \widehat{\mathbf{E}}_2\end{bmatrix}$ and $\widehat{\mathbf{A}} = \begin{bmatrix}\widehat{\mathbf{A}}_1 & 0\\ 0 & \mathbf{I}\end{bmatrix}$ with $\widehat{\mathbf{A}}_1$ being nonsingular and $\widehat{\mathbf{E}}_2$ being nilpotent. Consequently, the reduced-order model is inherently decoupled into a proper and an improper reduced state. The quality of the approximation can be estimated as

$$\|\boldsymbol{\mathcal{G}} - \widehat{\boldsymbol{\mathcal{G}}}\|_{\mathcal{H}_\infty} \leq 2(\sigma_{r+1} + \cdots + \sigma_{n_f}),$$

where $\boldsymbol{\mathcal{G}}(s) := \mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}$ and $\widehat{\boldsymbol{\mathcal{G}}}(s) := \widehat{\mathbf{C}}(s\widehat{\mathbf{E}}-\widehat{\mathbf{A}})^{-1}\widehat{\mathbf{B}}$ are the transfer functions of the original and reduced-order models, respectively, and σ_{i} is the i-th largest singular value, that is the i-th diagonal element of $\mathbf{\Sigma}$.

We emphasize that the controllability behavior of the model in (8) is the same as for the model in (1), as the input–to–state mapping is the same. However, the observability behavior of (8) is not the same as for (1), due to the quadratic form of the output equations instead of the linear one in (8).

2.3 BT for linear dynamical systems with quadratic outputs

In this subsection, we briefly summarize BT for linear systems with quadratic outputs—introduced in [5]—for $\mathbf{E} = \mathbf{I}$. These systems are of the form

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = 0,\\ \mathbf{y}(t) &= \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}(t),\end{aligned} \tag{12}$$

where the matrices are defined as in (1), except that the matrix $\mathbf{E}$ is replaced by the identity matrix. The goal is to find projection matrices $\mathbf{W}_{\mathrm{r}},\ \mathbf{T}_{\mathrm{r}} \in \mathbb{R}^{n\times r}$ such that the reduced-order system

$$\begin{aligned}\dot{\widehat{\mathbf{x}}}(t) &= \widehat{\mathbf{A}}\widehat{\mathbf{x}}(t) + \widehat{\mathbf{B}}\mathbf{u}(t), \qquad \widehat{\mathbf{x}}(0) = 0,\\ \widehat{\mathbf{y}}(t) &= \widehat{\mathbf{x}}(t)^{\mathrm{T}}\widehat{\mathbf{M}}\widehat{\mathbf{x}}(t)\end{aligned} \tag{13}$$

with $\widehat{\mathbf{A}} := \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{A}\mathbf{T}_{\mathrm{r}},\ \widehat{\mathbf{B}} := \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{B},\ \widehat{\mathbf{M}} := \mathbf{T}_{\mathrm{r}}^{\mathrm{T}}\mathbf{M}\mathbf{T}_{\mathrm{r}}$, approximates the input–to–output behavior of the model in (12) well. To achieve this goal, using BT, a new pair of Gramians is discussed in [5], tailored for the model in (12).

The state $\mathbf{x}(t)$ of the model in (12) is given by $\mathbf{x}(t) = \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\mathrm{d}\tau$. We define the input–to–state mapping as $\boldsymbol{\mathcal{C}}(t) := e^{\mathbf{A}t}\mathbf{B}$. The controllability space of the model in (12) is described by the controllability Gramian $\boldsymbol{\mathcal{P}}$, that is

$$\boldsymbol{\mathcal{P}} := \int_0^\infty \boldsymbol{\mathcal{C}}(t)\boldsymbol{\mathcal{C}}(t)^{\mathrm{T}}\mathrm{d}t = \int_0^\infty e^{\mathbf{A}t}\mathbf{B}\mathbf{B}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}t}\mathrm{d}t.$$

To describe the observability of the system, we consider the output equation

$$\begin{aligned}\mathbf{y}(t) = \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}(t) &= \int_0^t\int_0^t \mathbf{u}(t_1)^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}t_1}\mathbf{M}e^{\mathbf{A}t_2}\mathbf{B}\mathbf{u}(t_2)\mathrm{d}t_1\mathrm{d}t_2\\ &= \int_0^t\int_0^t \mathrm{vec}\left(\mathbf{B}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}(t-\tau_1)}\mathbf{M}e^{\mathbf{A}(t-\tau_2)}\mathbf{B}\right)(\mathbf{u}(\tau_2)\otimes\mathbf{u}(\tau_1))\mathrm{d}\tau_1\mathrm{d}\tau_2.\end{aligned}$$

We identify the input–to–state mapping $\boldsymbol{\mathcal{C}}(t) = e^{\mathbf{A}t}\mathbf{B}$ that was defined above and define the state–to–output mapping

$$\boldsymbol{\mathcal{O}}(t_1,t_2) := \mathbf{B}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}t_1}\mathbf{M}e^{\mathbf{A}t_2}.$$

The corresponding observability Gramian $\boldsymbol{\mathcal{Q}}$ is defined as follows:

$$\begin{aligned}\boldsymbol{\mathcal{Q}} &= \int_0^\infty \int_0^\infty \boldsymbol{\mathcal{O}}(t_1,t_2)^{\mathrm{T}}\boldsymbol{\mathcal{O}}(t_1,t_2)\mathrm{d}t_1\mathrm{d}t_2 = \int_0^\infty e^{\mathbf{A}^{\mathrm{T}}t_2}\mathbf{M}\int_0^\infty e^{\mathbf{A}t_1}\mathbf{B}\mathbf{B}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}t_1}\mathrm{d}t_1\mathbf{M}e^{\mathbf{A}t_2}\mathrm{d}t_2\\ &= \int_0^\infty e^{\mathbf{A}^{\mathrm{T}}t_2}\mathbf{M}\boldsymbol{\mathcal{P}}\mathbf{M}e^{\mathbf{A}t_2}\mathrm{d}t_2.\end{aligned}$$

Because of the positive definiteness of $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{Q}}$, we can compute Cholesky factorizations $\boldsymbol{\mathcal{P}} = \mathbf{R}\mathbf{R}^{\mathrm{T}}$ and $\boldsymbol{\mathcal{Q}} = \mathbf{S}\mathbf{S}^{\mathrm{T}}$. We make use of the energy functionals $E_{\mathrm{c}}(\mathbf{x}_0)$ that is the minimal energy that is needed to reach a state $\mathbf{x}_0$ and $E_{\mathrm{o}}(\mathbf{x}_0)$ that is the output energy that is produced by a nonzero initial condition $\mathbf{x}_0$. They satisfy

$$\begin{aligned}E_{\mathrm{c}}(\mathbf{x}_0) &:= \min_{\substack{\mathbf{x}(-\infty)\,=\,0\\ \mathbf{x}(0)\,=\,0}} \|\mathbf{u}\|_{L_2}^2 = \mathbf{x}_0^{\mathrm{T}}\boldsymbol{\mathcal{P}}^{-1}\mathbf{x}_0,\\ E_{\mathrm{o}}(\mathbf{x}_0) &:= \int_0^\infty \|\mathbf{y}(t)\|^2\mathrm{d}t \le \mathbf{x}_0^{\mathrm{T}}\boldsymbol{\mathcal{Q}}\mathbf{x}_0 \quad \text{for} \quad \mathbf{x}_0 = \mathbf{R}\mathbf{w},\ \|\mathbf{w}\| \le 1\end{aligned}$$

and can be used to characterize the hard to reach and hard to observe states, which correspond to small singular values of $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{Q}}$, respectively. To truncate such states simultaneously, we compute the singular value decomposition

$$\mathbf{S}^{\mathrm{T}}\mathbf{E}\mathbf{R} = \mathbf{U}_{\mathrm{p}}\boldsymbol{\Sigma}\mathbf{V}_{\mathrm{p}}^{\mathrm{T}} = \begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_1 & \\ & \boldsymbol{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_1^{\mathrm{T}}\\ \mathbf{V}_2^{\mathrm{T}}\end{bmatrix}.$$

The singular values in $\boldsymbol{\Sigma}_1$ and the corresponding most observable and reachable states lying in the spaces spanned by $\mathbf{V}_1$ and $\mathbf{U}_1$ are used to derive the projection matrices

$$\mathbf{W}_{\mathrm{r}} = \mathbf{S}^{\mathrm{T}}\mathbf{U}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}, \qquad \mathbf{T}_{\mathrm{r}} = \mathbf{R}^{\mathrm{T}}\mathbf{V}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}.$$

We multiply the model in (12) by $\mathbf{W}_{\mathrm{r}}$ and $\mathbf{T}_{\mathrm{r}}$ and obtain the reduced-order model in (13) that satisfies the error bound

$$\|\mathbf{y} - \widehat{\mathbf{y}}\|_{\mathcal{L}_\infty} \le \sqrt{\operatorname{tr}\left(\mathbf{B}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}\mathbf{B} - 2\mathbf{B}^{\mathrm{T}}\boldsymbol{\mathcal{Z}}\widehat{\mathbf{B}} + \widehat{\mathbf{B}}^{\mathrm{T}}\widehat{\boldsymbol{\mathcal{Q}}}\widehat{\mathbf{B}}\right)}\|\mathbf{u}\otimes\mathbf{u}\|_{\mathcal{L}_2}$$

with

$$\mathbf{A}^{\mathrm{T}}\boldsymbol{\mathcal{Z}} + \boldsymbol{\mathcal{Z}}\widehat{\mathbf{A}} = -\mathbf{M}\boldsymbol{\mathcal{X}}\widehat{\mathbf{M}}, \qquad \mathbf{A}\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{X}}\widehat{\mathbf{A}}^{\mathrm{T}} = -\mathbf{B}\widehat{\mathbf{B}}^{\mathrm{T}}.$$

The matrix $\widehat{\boldsymbol{\mathcal{Q}}}$ is the observability Gramian of the reduced-order model in (13).

In the following sections, we extend the methodology for the DAE systems as described in (1) and derive the corresponding proper and improper observability Gramians.

3 Gramians for DAE systems with quadratic output

We aim to extend BT for DAE systems with quadratic output equations described in (1). Therefore, we require tailored Gramians encoding controllability and observability subspaces for this class of systems. Since the nonlinearities appear only on the output equation, the controllability Gramians from (9) can be used to characterize controllability. However, the extension of observability Gramians for DAE systems with quadratic output is not straightforward. Hence, in this section, we propose new Gramians that describe the observability of the proper and improper states based on an output decomposition. Later, we use them in our proposed BT method for DAE systems with quadratic output. To derive observability Gramians, we decompose the output in the following way

$$\begin{aligned}\mathbf{y}(t) &= \mathbf{x}_{\mathrm{p}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{p}}(t) + \mathbf{x}_{\mathrm{p}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{i}}(t) + \mathbf{x}_{\mathrm{i}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{p}}(t) + \mathbf{x}_{\mathrm{i}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{i}}(t)\\ &=: \mathbf{y}_{\mathrm{pp}}(t) + \mathbf{y}_{\mathrm{pi}}(t) + \mathbf{y}_{\mathrm{ip}}(t) + \mathbf{y}_{\mathrm{ii}}(t),\end{aligned}$$

where $\mathbf{x}_{\mathrm{p}}$ and $\mathbf{x}_{\mathrm{i}}$ are, respectively, the proper and improper states. Note that the components $\mathbf{x}_{\mathrm{p}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{i}}(t)$ and $\mathbf{x}_{\mathrm{i}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{p}}(t)$ coincide. However, we will treat them independently for the

purposes of the following derivations. The idea in the following is to investigate the four components of the output separately. Since the output is a superposition of the four components, the Gramians that describe the output components sum up to Gramians that describe the overall observability of the system.

For a better understanding, we can rewrite $\mathbf{y}(t)$ by defining the state depending function $\mathbf{C}(\mathbf{x}(t)) := \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}$. Applying this representation to the decomposed output yields

$$\mathbf{y}(t) = \mathbf{C}(\mathbf{x}_{\mathrm{p}}(t))\mathbf{x}_{\mathrm{p}}(t) + \mathbf{C}(\mathbf{x}_{\mathrm{p}}(t))\mathbf{x}_{\mathrm{i}}(t) + \mathbf{C}(\mathbf{x}_{\mathrm{i}}(t))\mathbf{x}_{\mathrm{p}}(t) + \mathbf{C}(\mathbf{x}_{\mathrm{i}}(t))\mathbf{x}_{\mathrm{i}}(t).$$

We observe, that the observability of the state $\mathbf{x}_{\mathrm{p}}(t)$ in the output $\mathbf{y}_{\mathrm{ip}}(t) = \mathbf{C}(\mathbf{x}_{\mathrm{i}}(t))\mathbf{x}_{\mathrm{p}}(t)$ also depends on the reachability of $\mathbf{x}_{\mathrm{i}}(t)$. On the other hand, the observability of the improper state $\mathbf{x}_{\mathrm{i}}(t)$ corresponding to $\mathbf{y}_{\mathrm{pi}}(t) = \mathbf{C}(\mathbf{x}_{\mathrm{p}}(t))\mathbf{x}_{\mathrm{i}}(t)$ depends on the reachability of $\mathbf{x}_{\mathrm{p}}(t)$. Hence, the outputs $\mathbf{y}_{\mathrm{ip}}(t) = \mathbf{y}_{\mathrm{pi}}(t)$ encode two different observability properties. Analogously, the outputs $\mathbf{y}_{\mathrm{pp}}(t)$ and $\mathbf{y}_{\mathrm{ii}}(t)$ encode the observability of a proper state depending on the reachability of the same, and the observability of an improper state depending on the reachability of an improper one.

In this section, we define proper observability Gramians encoding the observability behavior of the proper state $\mathbf{x}_{\mathrm{p}}(t)$ corresponding to $\mathbf{C}(\mathbf{x}_{\mathrm{p}}(t))$ and $\mathbf{C}(\mathbf{x}_{\mathrm{i}}(t))$ and improper observability Gramians describing the observability of the improper states $\mathbf{x}_{\mathrm{i}}(t)$ corresponding to $\mathbf{C}(\mathbf{x}_{\mathrm{p}}(t))$ and $\mathbf{C}(\mathbf{x}_{\mathrm{i}}(t))$. Because of the dependencies on the reachability of $\mathbf{x}_{\mathrm{p}}(t)$ and $\mathbf{x}_{\mathrm{i}}(t)$ encoded by $\mathbf{C}(\mathbf{x}_{\mathrm{p}}(t))$ and $\mathbf{C}(\mathbf{x}_{\mathrm{i}}(t))$, we expect that the observability Gramians will depend on the controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$.

3.1 Proper observability Gramian

In this subsection, we investigate the two outputs $\mathbf{y}_{\mathrm{pp}}(t)$ and $\mathbf{y}_{\mathrm{ip}}(t)$ and their observability properties. We aim to describe the observability of the right proper state depending on the second (left) state in the quadratic output equation, which is in the first case proper and in the second case improper.

Proper-proper output We start investigating the first component of the output $\mathbf{y}_{\mathrm{pp}}(t) = \mathbf{x}_{\mathrm{p}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{p}}(t)$ that includes two proper states. We define the state–to–output mapping

$$\boldsymbol{\mathcal{O}}_{\mathrm{pp}}(t_1, t_2) := \mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t_1)\mathbf{M}\mathbf{F}_J(t_2)$$

that is used to describe the observability corresponding to $\mathbf{y}_{\mathrm{pp}}(t)$. Based on this mapping, we define in the following the proper-proper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{pp}}$ as

$$\begin{aligned}\boldsymbol{\mathcal{Q}}_{\mathrm{pp}} &:= \int_0^\infty \int_0^\infty \boldsymbol{\mathcal{O}}_{\mathrm{pp}}(t_1, t_2)^{\mathrm{T}}\boldsymbol{\mathcal{O}}_{\mathrm{pp}}(t_1, t_2)\mathrm{d}t_1\mathrm{d}t_2 \\ &= \int_0^\infty \int_0^\infty \mathbf{F}_J(t_2)^{\mathrm{T}}\mathbf{M}\mathbf{F}_J(t_1)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t_1)^{\mathrm{T}}\mathbf{M}\mathbf{F}_J(t_2)\mathrm{d}t_1\mathrm{d}t_2 \\ &= \int_0^\infty \mathbf{F}_J(t_2)^{\mathrm{T}}\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{p}}\mathbf{M}\mathbf{F}_J(t_2)\mathrm{d}t_2\end{aligned}$$

which results in the following definition.

Definition 3.1. Consider the asymptotically stable DAE system with a quadratic output equation from (1) and the corresponding proper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ as defined in (9). The proper-proper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{pp}}$ corresponding to the output $\mathbf{y}_{\mathrm{pp}}$ is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{pp}} = \int_0^\infty \mathbf{F}_J(t)^{\mathrm{T}}\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{p}}\mathbf{M}\mathbf{F}_J(t)\mathrm{d}t,$$

where $\mathbf{F}_J(t)$ is defined as in (6).

The above defined Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{pp}}$ can be transformed into the Weierstraß-canonical representation (4). For that we insert the function $\mathbf{F}_J(t)$ leading to

$$\boldsymbol{\mathcal{Q}}_{\mathrm{pp}} := \mathbf{W}^{-\mathrm{T}}\begin{bmatrix}\mathbf{Q}_{11} & 0\\ 0 & 0\end{bmatrix}\mathbf{W}^{-1},$$

where

$$\mathbf{Q}_{11} := \int_0^\infty e^{\mathbf{J}^\mathrm{T} t} \mathbf{M}_{11} \boldsymbol{\mathcal{P}}_1 \mathbf{M}_{11} e^{\mathbf{J} t} \mathrm{d}t \tag{14}$$

is the proper-proper observability Gramian of the state $\mathbf{x}_1(t)$ from system (4). The Gramian $\mathbf{Q}_{11}$ corresponding to the ODE system was analyzed in [5].

Theorem 3.1. Consider the asymptotically stable DAE system with a quadratic output equation from (1) and the corresponding proper controllability Gramian $\boldsymbol{\mathcal{P}}_\mathrm{p}$ as defined in (9). The proper observability Gramian $\boldsymbol{\mathcal{Q}}_\mathrm{pp}$ solves the projected Lyapunov equation

$$\mathbf{E}^\mathrm{T} \boldsymbol{\mathcal{Q}}_\mathrm{pp} \mathbf{A} + \mathbf{A}^\mathrm{T} \boldsymbol{\mathcal{Q}}_\mathrm{pp} \mathbf{E} = -\mathbf{P}_\mathrm{r}^\mathrm{T} \mathbf{M} \boldsymbol{\mathcal{P}}_\mathrm{p} \mathbf{M} \mathbf{P}_\mathrm{r}, \qquad \boldsymbol{\mathcal{Q}}_\mathrm{pp} = \mathbf{P}_\mathrm{l}^\mathrm{T} \boldsymbol{\mathcal{Q}}_\mathrm{pp} \mathbf{P}_\mathrm{l}.$$

where the projection matrices $\mathbf{P}_\mathrm{l}$ and $\mathbf{P}_\mathrm{r}$ are defined as in (3).

Proof. We first show that the Gramian $\mathbf{Q}_{11}$ defined in (14) solves the Lyapunov equation

$$\mathbf{J}^\mathrm{T} \mathbf{Q}_{11} + \mathbf{Q}_{11} \mathbf{J} = -\mathbf{M}_{11} \boldsymbol{\mathcal{P}}_1 \mathbf{M}_{11}. \tag{15}$$

Therefore, we insert $\mathbf{Q}_{11}$ into (15) and obtain

$$\int_0^\infty \left(\mathbf{J}^\mathrm{T} e^{\mathbf{J}^\mathrm{T} t} \mathbf{M}_{11} \boldsymbol{\mathcal{P}}_1 \mathbf{M}_{11} e^{\mathbf{J} t} + e^{\mathbf{J}^\mathrm{T} t} \mathbf{M}_{11} \boldsymbol{\mathcal{P}}_1 \mathbf{M}_{11} e^{\mathbf{J} t} \mathbf{J} \right) \mathrm{d}t = \left[e^{\mathbf{J}^\mathrm{T} t} \mathbf{M}_{11} \boldsymbol{\mathcal{P}}_1 \mathbf{M}_{11} e^{\mathbf{J} t} \right]_0^\infty = -\mathbf{M}_{11} \boldsymbol{\mathcal{P}}_1 \mathbf{M}_{11}.$$

To prove the statement of the theorem, we first observe that the projection condition is naturally satisfied since $\boldsymbol{\mathcal{Q}}_\mathrm{pp}$ is by definition equal to $\mathbf{W}^{-\mathrm{T}} \begin{bmatrix} \mathbf{Q}_{11} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{W}^{-1}$. To prove that $\boldsymbol{\mathcal{Q}}_\mathrm{pp}$ satisfies the remaining Lyapunov equation, we insert the Weierstraß-canonical form of $\mathbf{E}$ and $\mathbf{A}$ and the definition of $\mathbf{P}_\mathrm{r}$ into the equation to obtain

$$\begin{aligned}
\mathbf{T}^\mathrm{T} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{N}^\mathrm{T} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{J} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \mathbf{T} + \mathbf{T}^\mathrm{T} \begin{bmatrix} \mathbf{J}^\mathrm{T} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{N} \end{bmatrix} \mathbf{T} \\
&= \mathbf{T}^\mathrm{T} \begin{bmatrix} \mathbf{Q}_{11} \mathbf{J} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{T} + \mathbf{T}^\mathrm{T} \begin{bmatrix} \mathbf{J}^\mathrm{T} \mathbf{Q}_{11} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{T} \\
&= -\mathbf{T}^\mathrm{T} \begin{bmatrix} \mathbf{M}_{11} \boldsymbol{\mathcal{P}}_1 \mathbf{M}_{11} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{T} \\
&= -\mathbf{P}_\mathrm{r}^\mathrm{T} \mathbf{M} \boldsymbol{\mathcal{P}}_\mathrm{p} \mathbf{M} \mathbf{P}_\mathrm{r}.
\end{aligned}$$

such that (15) implies the statement since $\mathbf{T}$ is a nonsingular matrix. □

Theorem 3.1 states that we can calculate Gramians by solving certain projected Lyapunov equations. Methods to solve projected Lyapunov equations can be found, e.g., in [23, 26, 27].

Improper-proper output Now we consider the third output component $\mathbf{y}_\mathrm{ip}(t) = \mathbf{x}_\mathrm{i}(t)^\mathrm{T} \mathbf{M} \mathbf{x}_\mathrm{p}(t)$. We define the state–to–output mapping $\boldsymbol{\mathcal{O}}_\mathrm{ip}(t,k) := \mathbf{B}^\mathrm{T} \mathbf{F}_N(k)^\mathrm{T} \mathbf{M} \mathbf{F}_J(t)$ that is used to describe the observability corresponding to $\mathbf{y}_\mathrm{ip}(t)$. We note that we can also consider the transposed kernel to derive an improper Gramian, which is done later in Subsection 3.2. However, we derive a Gramian that encodes the observability properties of the proper component of the output $\mathbf{y}_\mathrm{ip}(t)$. We define the improper-proper observability Gramian as

$$\begin{aligned}
\boldsymbol{\mathcal{Q}}_\mathrm{ip} &:= \sum_{k=0}^{\nu-1} \int_0^\infty \boldsymbol{\mathcal{O}}_\mathrm{ip}(t,k)^\mathrm{T} \boldsymbol{\mathcal{O}}_\mathrm{ip}(t,k) \mathrm{d}t = \sum_{k=0}^{\nu-1} \int_0^\infty \mathbf{F}_J(t)^\mathrm{T} \mathbf{M} \mathbf{F}_N(k) \mathbf{B} \mathbf{B}^\mathrm{T} \mathbf{F}_N(k)^\mathrm{T} \mathbf{M} \mathbf{F}_J(t) \mathrm{d}t \\
&= \int_0^\infty \mathbf{F}_J(t)^\mathrm{T} \mathbf{M} \boldsymbol{\mathcal{P}}_\mathrm{i} \mathbf{M} \mathbf{F}_J(t) \mathrm{d}t.
\end{aligned}$$

Definition 3.2. Consider the asymptotically stable DAE system with a quadratic output equation from (1) and the corresponding improper controllability Gramian $\boldsymbol{\mathcal{P}}_\mathrm{i}$ as defined in (9). The improper-proper observability Gramian $\boldsymbol{\mathcal{Q}}_\mathrm{ip}$ corresponding to the output $\mathbf{y}_\mathrm{ip}$ is defined as

$$\boldsymbol{\mathcal{Q}}_\mathrm{ip} := \int_0^\infty \mathbf{F}_J(t)^\mathrm{T} \mathbf{M} \boldsymbol{\mathcal{P}}_\mathrm{i} \mathbf{M} \mathbf{F}_J(t) \mathrm{d}t$$

where $\mathbf{F}_J(t)$ is defined as in (6).

The following theorem describes how the Gramian $\mathcal{Q}_{\text{ip}}$ can be computed in practice.

Theorem 3.2. Consider the asymptotically stable DAE system with a quadratic output equation from (1) and the corresponding improper controllability Gramian $\mathcal{P}_{\text{i}}$ as defined in (9). The improper-proper observability Gramian $\mathcal{Q}_{\text{ip}}$ solves the projected Lyapunov equation

$$\mathbf{E}^{\mathrm{T}}\mathcal{Q}_{\text{ip}}\mathbf{A} + \mathbf{A}^{\mathrm{T}}\mathcal{Q}_{\text{ip}}\mathbf{E} = -\mathbf{P}_{\text{r}}^{\mathrm{T}}\mathbf{M}\mathcal{P}_{\text{i}}\mathbf{M}\mathbf{P}_{\text{r}}, \quad \mathcal{Q}_{\text{ip}} = \mathbf{P}_{\text{l}}^{\mathrm{T}}\mathcal{Q}_{\text{ip}}\mathbf{P}_{\text{l}},$$

where the projection matrices $\mathbf{P}_{\text{l}}$ and $\mathbf{P}_{\text{r}}$ are defined as in (3).

Proof. The proof follows the same argumentation as for Theorem 3.1. □

Joined proper observability Gramian. We can combine the two proper output Gramians to obtain a Gramian that covers the observability of the proper states independent of the second state, that is the observability of the output $\mathbf{y}_{\text{p}}(t) = \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\text{p}}(t)$ for an arbitrary state $\mathbf{x}(t)$, generated by the model in (1). Since the sum $\mathcal{P}_{\text{p}} + \mathcal{P}_{\text{i}}$ spans the full controllability space of the state $\mathbf{x}(t)$, the proper observability Gramian corresponding to both proper and improper left states is given by

$$\begin{aligned}
\mathcal{Q}_{\text{p}} &= \int_0^\infty \mathbf{F}_J(t)^{\mathrm{T}}\mathbf{M}(\mathcal{P}_{\text{p}} + \mathcal{P}_{\text{i}})\mathbf{M}\mathbf{F}_J(t)\mathrm{d}t \\
&= \int_0^\infty \mathbf{W}^{-\mathrm{T}} \begin{bmatrix} e^{\mathbf{J}^{\mathrm{T}}t} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{12}^{\mathrm{T}} & \mathbf{M}_{22} \end{bmatrix} \begin{bmatrix} \mathcal{P}_1 & 0 \\ 0 & \mathcal{P}_2 \end{bmatrix} \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{12}^{\mathrm{T}} & \mathbf{M}_{22} \end{bmatrix} \begin{bmatrix} e^{\mathbf{J}t} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{W}^{-1}\mathrm{d}t \\
&= \mathbf{W}^{-\mathrm{T}} \begin{bmatrix} \mathbf{Q}_{11} + \mathbf{Q}_{21} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{W}^{-1} = \mathcal{Q}_{\text{pp}} + \mathcal{Q}_{\text{ip}}.
\end{aligned}$$

We summarize this subsection with the following definition.

Definition 3.3. Consider the asymptotically stable DAE system with a quadratic output equation from (1) and the corresponding proper and improper controllability Gramians $\mathcal{P}_{\text{p}}$ and $\mathcal{P}_{\text{i}}$ as defined in (9). The proper observability Gramian $\mathcal{Q}_{\text{p}}$ corresponding to the output $\mathbf{y}_{\text{p}} = \mathbf{y}_{\text{pp}}(t) + \mathbf{y}_{\text{ip}}(t)$ is defined as

$$\mathcal{Q}_{\text{p}} := \mathcal{Q}_{\text{pp}} + \mathcal{Q}_{\text{ip}}$$

with $\mathcal{Q}_{\text{pp}}$ and $\mathcal{Q}_{\text{ip}}$ as in the Definition 3.1 and 3.2.

The summed proper Gramian $\mathcal{Q}_{\text{p}}$ provides a criterion describing the states which are easiest and hardest to observe. We will show later in Subsection 4.1 that the hard to observe states are connected to the smallest non-zero singular values of the Gramian and thus serve as a truncation criterion to reduce the full-order model.

3.2 Improper observability Gramians

In this subsection, we investigate the observability behavior of the outputs $\mathbf{y}_{\text{pi}} := \mathbf{x}_{\text{p}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\text{i}}(t)$ and $\mathbf{y}_{\text{ii}} := \mathbf{x}_{\text{i}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\text{i}}(t)$. Both outputs describe the observability of the improper state $\mathbf{x}_{\text{i}}(t)$ corresponding to a proper and an improper state multiplied from the left.

Proper-improper output We describe $\mathbf{y}_{\text{ip}}(t)$ so that we derive an improper Gramian in this subsection. This just means that we derive a Gramian that encodes the observability of the improper component of $\mathbf{y}_{\text{pi}}(t)$. We identify the improper controllability mapping $\mathcal{C}_{\text{i}}(k) = \mathbf{F}_N(k)\mathbf{B}$ and the remaining observability mapping $\mathcal{O}_{\text{pi}}(t,k) = \mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t)^{\mathrm{T}}\mathbf{M}\mathbf{F}_N(k)$ which is used to define the proper-improper observability Gramian corresponding to the state $\mathbf{x}_{\text{i}}(t)$ and the output $\mathbf{y}_{\text{pi}}(t)$ as

$$\begin{aligned}
\mathcal{Q}_{\text{pi}} &= \int_0^\infty \sum_{k=0}^{\nu-1} \mathcal{O}_{\text{pi}}(t,k)^{\mathrm{T}}\mathcal{O}_{\text{pi}}(t,k)\mathrm{d}t = \int_0^\infty \sum_{k=0}^{\nu-1} \mathbf{F}_N(k)^{\mathrm{T}}\mathbf{M}\mathbf{F}_J(t)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t)^{\mathrm{T}}\mathbf{M}\mathbf{F}_N(k)\mathrm{d}t \\
&= \sum_{k=0}^{\nu-1} \mathbf{F}_N(k)^{\mathrm{T}}\mathbf{M}\mathcal{P}_{\text{p}}\mathbf{M}\mathbf{F}_N(k).
\end{aligned}$$

This results in the following definition.

Definition 3.4. Consider the asymptotically stable DAE system with a quadratic output equation from (1) and the corresponding proper controllability Gramian $\mathcal{P}_{\mathrm{p}}$ as defined in (9). The proper-improper observability Gramian $\mathcal{Q}_{\mathrm{pi}}$ corresponding to the output $\mathbf{y}_{\mathrm{pi}}$ is defined as

$$\mathcal{Q}_{\mathrm{pi}} = \sum_{k=0}^{\nu-1} \mathbf{F}_N(k)^{\mathrm{T}} \mathbf{M} \mathcal{P}_{\mathrm{p}} \mathbf{M} \mathbf{F}_N(k),$$

where $\mathbf{F}_N(k)$ is defined as in (6).

We insert the mapping $\mathbf{F}_N(k)$ and obtain that the improper observability Gramian $\mathcal{Q}_{\mathrm{pi}}$ can be written as

$$\mathcal{Q}_{\mathrm{pi}} = \mathbf{W}^{-\mathrm{T}} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{Q}_{12} \end{bmatrix} \mathbf{W}^{-1}$$

where

$$\mathbf{Q}_{12} := \sum_{k=0}^{\nu-1} (-\mathbf{N}^k)^{\mathrm{T}} \mathbf{M}_{12}^{\mathrm{T}} \mathcal{P}_1 \mathbf{M}_{12} (-\mathbf{N}^k) \tag{16}$$

is the proper-improper observability Gramian corresponding to the improper state $\mathbf{x}_2(t)$ from the transformed model in (4). This describes a relation between the Weierstraß-canonical form (4) and the full-order model in (1) in this observability Gramain $\mathcal{Q}_{\mathrm{pi}}$.

Theorem 3.3. Consider the asymptotically stable DAE system with a quadratic output equation from (1) and the corresponding proper controllability Gramian $\mathcal{P}_{\mathrm{p}}$ as defined in (9). The proper-improper observability Gramian $\mathcal{Q}_{\mathrm{pi}}$ solves the projected generalized discrete-time Lyapunov equation

$$\mathbf{A}^{\mathrm{T}} \mathcal{Q}_{\mathrm{pi}} \mathbf{A} - \mathbf{E}^{\mathrm{T}} \mathcal{Q}_{\mathrm{pi}} \mathbf{E} = (\mathbf{I} - \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}) \mathbf{M} \mathcal{P}_{\mathrm{p}} \mathbf{M} (\mathbf{I} - \mathbf{P}_{\mathrm{r}}), \quad \mathbf{P}_{\mathrm{l}}^{\mathrm{T}} \mathcal{Q}_{\mathrm{pi}} \mathbf{P}_{\mathrm{l}} = 0$$

where the projection matrices $\mathbf{P}_{\mathrm{l}}$ and $\mathbf{P}_{\mathrm{r}}$ are defined as in (3).

Proof. We first show that the Gramian $\mathbf{Q}_{12}$ defined in (16) solves the discrete-time Lyapunov equation

$$\mathbf{Q}_{12} - \mathbf{N}^{\mathrm{T}} \mathbf{Q}_{12} \mathbf{N} = \mathbf{M}_{12}^{\mathrm{T}} \mathcal{P}_1 \mathbf{M}_{12}. \tag{17}$$

That follows if we insert the definition of $\mathbf{Q}_{12}$ into (17). This results in

$$\begin{aligned} \sum_{k=0}^{\nu-1} (-\mathbf{N}^k)^{\mathrm{T}} \mathbf{M}_{12}^{\mathrm{T}} \mathcal{P}_1 \mathbf{M}_{12} (-\mathbf{N}^k) - \sum_{k=0}^{\nu-1} (-\mathbf{N}^{k+1})^{\mathrm{T}} \mathbf{M}_{12}^{\mathrm{T}} \mathcal{P}_1 \mathbf{M}_{12} (-\mathbf{N}^{k+1}) &= (-\mathbf{N}^0)^{\mathrm{T}} \mathbf{M}_{12}^{\mathrm{T}} \mathcal{P}_1 \mathbf{M}_{12} (-\mathbf{N}^0) \\ &= \mathbf{M}_{12}^{\mathrm{T}} \mathcal{P}_1 \mathbf{M}_{12} \end{aligned}$$

since $\mathbf{N}$ has the nilpotency index $\nu - 1$, i.e., $\mathbf{N}^{\nu} = 0$.

To prove the projection condition, we derive

$$\mathbf{P}_{\mathrm{l}}^{\mathrm{T}} \mathcal{Q}_{\mathrm{pp}} \mathbf{P}_{\mathrm{l}} = \mathbf{W}^{-\mathrm{T}} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{W}^{\mathrm{T}} \mathbf{W}^{-\mathrm{T}} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{Q}_{12} \end{bmatrix} \mathbf{W}^{-1} \mathbf{W} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{W}^{-1} = 0.$$

To finalize the proof, we insert the Weierstraß-canonical form of $\mathbf{E}$ and $\mathbf{A}$ and the definition of $\mathbf{P}_{\mathrm{r}}$ into the remaining Lyapunov equation to obtain

$$\begin{aligned} \mathbf{T}^{\mathrm{T}} \begin{bmatrix} \mathbf{J}^{\mathrm{T}} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{Q}_{12} \end{bmatrix} \begin{bmatrix} \mathbf{J} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \mathbf{T} - \mathbf{T}^{\mathrm{T}} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{N}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{Q}_{12} \end{bmatrix} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{N} \end{bmatrix} \mathbf{T} & \\ = \mathbf{T}^{\mathrm{T}} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{Q}_{12} - \mathbf{N}^{\mathrm{T}} \mathbf{Q}_{12} \mathbf{N} \end{bmatrix} \mathbf{T} & \\ = \mathbf{T}^{\mathrm{T}} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{M}_{12}^{\mathrm{T}} \mathcal{P}_1 \mathbf{M}_{12} \end{bmatrix} \mathbf{T} & \\ = (\mathbf{I} - \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}) \mathbf{M} \mathcal{P}_{\mathrm{p}} \mathbf{M} (\mathbf{I} - \mathbf{P}_{\mathrm{r}}) & \end{aligned}$$

which proves the statement. □

Improper-improper output Now, we consider the fourth and last component of the output that describes the observability space of the improper state $\mathbf{x}_{\mathrm{i}}(t)$ if the second state is also improper. We identify the state–to–output mapping $\boldsymbol{\mathcal{O}}_{\mathrm{ii}}(k,\ell) = \mathbf{B}^{\mathrm{T}}\mathbf{F}_N(k)^{\mathrm{T}}\mathbf{M}\mathbf{F}_N(\ell)$. Based on this mapping $\boldsymbol{\mathcal{O}}_{\mathrm{ii}}(k,\ell)$ we define the observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{ii}}$ as

$$\begin{aligned}\boldsymbol{\mathcal{Q}}_{\mathrm{ii}} &:= \sum_{k=0}^{\nu-1}\sum_{\ell=0}^{\nu-1}\boldsymbol{\mathcal{O}}(k,\ell)^{\mathrm{T}}\boldsymbol{\mathcal{O}}(k,\ell) = \sum_{k=0}^{\nu-1}\sum_{\ell=0}^{\nu-1}\mathbf{F}_N(\ell)^{\mathrm{T}}\mathbf{M}\mathbf{F}_N(k)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{F}_N(k)^{\mathrm{T}}\mathbf{M}\mathbf{F}_N(\ell)\\ &= \sum_{\ell=0}^{\nu-1}\mathbf{F}_N(\ell)^{\mathrm{T}}\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\mathbf{M}\mathbf{F}_N(\ell)\end{aligned}$$

which results in the following definition.

Definition 3.5. Consider the DAE system with a quadratic output equation from (1) and the corresponding improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ as defined in (9). The improper-improper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{ii}}$ corresponding to the output $\mathbf{y}_{\mathrm{ii}}$ is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{ii}} = \sum_{\ell=0}^{\nu-1}\mathbf{F}_N(\ell)^{\mathrm{T}}\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\mathbf{M}\mathbf{F}_N(\ell),$$

where $\mathbf{F}_N(\ell)$ is defined as in (6).

Theorem 3.4. Consider the DAE system with a quadratic output equation from (1) and the corresponding improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ as defined in (9). The improper-improper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{ii}}$ solves the projected generalized discrete-time Lyapunov equation

$$\mathbf{A}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_{\mathrm{ii}}\mathbf{A} - \mathbf{E}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_{\mathrm{ii}}\mathbf{E} = (\mathbf{I}-\mathbf{P}_{\mathrm{r}}^{\mathrm{T}})\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\mathbf{M}(\mathbf{I}-\mathbf{P}_{\mathrm{r}}), \quad \mathbf{P}_{\mathrm{l}}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_{\mathrm{ii}}\mathbf{P}_{\mathrm{l}} = 0$$

where $\mathbf{P}_{\mathrm{l}}$ and $\mathbf{P}_{\mathrm{r}}$ are defined as in (3).

Proof. The proof is similar to the proof of Theorem 3.3. □

Jointed improper observability Gramain We can combine the two improper output Gramians to obtain an improper Gramian that covers the observability of an improper state independent of the second state, that is, the observability of the output $\mathbf{y}_{\mathrm{i}}(t) = \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{i}}(t)$ for an arbitrary state $\mathbf{x}(t)$ generated by system (1). Since the sum $\boldsymbol{\mathcal{P}}_{\mathrm{p}} + \boldsymbol{\mathcal{P}}_{\mathrm{i}}$ spans the full controllability space of the state $\mathbf{x}(t)$, the improper observability Gramian corresponding to both proper and improper left states is given by

$$\begin{aligned}\boldsymbol{\mathcal{Q}}_{\mathrm{i}} &= \sum_{k=0}^{\nu-1}\mathbf{F}_N(t)^{\mathrm{T}}\mathbf{M}(\boldsymbol{\mathcal{P}}_{\mathrm{p}}+\boldsymbol{\mathcal{P}}_{\mathrm{i}})\mathbf{M}\mathbf{F}_N(t)\\ &= \sum_{k=0}^{\nu-1}\mathbf{W}^{-\mathrm{T}}\begin{bmatrix}0 & 0\\ 0 & -(\mathbf{N}^k)^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathbf{M}_{11} & \mathbf{M}_{12}\\ \mathbf{M}_{12}^{\mathrm{T}} & \mathbf{M}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{\mathcal{P}}_1 & 0\\ 0 & \boldsymbol{\mathcal{P}}_2\end{bmatrix}\begin{bmatrix}\mathbf{M}_{11} & \mathbf{M}_{12}\\ \mathbf{M}_{12}^{\mathrm{T}} & \mathbf{M}_{22}\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & -\mathbf{N}^k\end{bmatrix}\mathbf{W}^{-1}\\ &= \mathbf{W}^{-\mathrm{T}}\begin{bmatrix}0 & 0\\ 0 & \mathbf{Q}_{12}+\mathbf{Q}_{22}\end{bmatrix}\mathbf{W}^{-1} = \boldsymbol{\mathcal{Q}}_{\mathrm{pi}} + \boldsymbol{\mathcal{Q}}_{\mathrm{ii}}.\end{aligned}$$

We summarize this subsection with the following definition.

Definition 3.6. Consider the DAE system with a quadratic output equation from (1) and the corresponding proper and improper controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ as defined in (9). The improper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{i}}$ corresponding to the output $\mathbf{y}_{\mathrm{i}}$ is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{i}} := \boldsymbol{\mathcal{Q}}_{\mathrm{pi}} + \boldsymbol{\mathcal{Q}}_{\mathrm{ii}}$$

where the Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{pi}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{ii}}$ are defined as in the Definition 3.4 and 3.5.

For the improper case, there are no energy functionals describing the connection of hard to observe states and the singular values of $\boldsymbol{\mathcal{Q}}_{\mathrm{i}}$. However, improper states that lie in the kernel of $\boldsymbol{\mathcal{Q}}_{\mathrm{i}}$ on the subspace $\ker(\mathbf{P}_{\mathrm{l}})$ are unobservable and, hence, can be removed from the dynamics without changing the input to output behavior. Therefore, the states corresponding to zero singular values of $\boldsymbol{\mathcal{Q}}_{\mathrm{i}}$ are removed in the reduction method presented in Subsection 4.2.

4 Kernel functions and balanced truncation

In the previous section, we have proposed Gramians that describe the proper and improper controllability and observability spaces. In this section, the goal is to propose a BT method based on those Gramians. The proper and the improper states are considered separately. We extend the methodology presented in [19] to address systems with quadratic output equations. For that, we first derive controllability and observability energies in Subsection 4.1 that are used in Subsection 4.2 to generate a reduced surrogate model that approximates the input–to–output behavior of the full-order model.

4.1 Energy norms of kernel functions

In standard BT theory, energy functionals are investigated to provide a criterion that states to truncate in the reduction step. To evaluate the output energy, the L_2-norm of $\mathbf{y}(t)$ for an initial condition $\mathbf{x}^*$ and a zero input $\mathbf{u}(t) \equiv 0$ is considered. However, since we consider consistent initial conditions, a vanishing input implies $\mathbf{x}_{\mathrm{i}}(t) \equiv 0$, and hence, $\mathbf{y}(t) = \mathbf{y}_{\mathrm{pp}}(t)$. The investigation of the output does not take into account the improper parts of the system and does not represent the complete system dynamics.

Hence, in this section, we investigate the dominant subspaces of the controllability mappings $\boldsymbol{\mathcal{C}}_{\mathrm{p}}$ and $\boldsymbol{\mathcal{C}}_{\mathrm{i}}$ and of the observability mappings $\boldsymbol{\mathcal{O}}_{\mathrm{pp}}$, $\boldsymbol{\mathcal{O}}_{\mathrm{pi}}$, $\boldsymbol{\mathcal{O}}_{\mathrm{ip}}$ and $\boldsymbol{\mathcal{O}}_{\mathrm{ii}}$. Afterward, in Subsection 4.2, we truncate the states living in the least important subspaces.

Controllability energy

Consider the controllability mapping $\boldsymbol{\mathcal{C}}_{\mathrm{p}}(t)$. We evaluate the energy norm of $\boldsymbol{\mathcal{C}}_{\mathrm{p}}$ that is

$$E(\boldsymbol{\mathcal{C}}_{\mathrm{p}}) = \|\boldsymbol{\mathcal{C}}_{\mathrm{p}}\|^2 = \operatorname{tr}\left(\int_0^\infty \mathbf{F}_J(t)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t)^{\mathrm{T}}\mathrm{d}t\right) = \operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{p}}\right) = \sigma_1 + \cdots + \sigma_{n_f},$$

where $\sigma_1 \geq \cdots \geq \sigma_{n_f} \geq 0$ are eigenvalues of $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$. Since $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ is symmetric and positive semi-definite, there exists $\mathbf{V} \in \mathbb{R}^{n\times n}$ with $\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}_n$ so that $\boldsymbol{\mathcal{P}}_{\mathrm{p}} = \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}$ and $\boldsymbol{\Sigma} = \operatorname{diag}\left(\sigma_1, \ldots, \sigma_{n_f}, 0, \ldots\right)$. We see that the first r columns of $\mathbf{V}$ span the most dominant proper controllability subspace since they correspond to the largest eigenvalues of $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ producing the largest energy values. This observation justifies that in Subsection 4.2, the states corresponding to the smallest singular values of $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ are truncated to reduce the model.

Similarly, we investigate the energy norm of the improper controllability mapping, that is

$$E(\boldsymbol{\mathcal{C}}_{\mathrm{i}}) = \|\boldsymbol{\mathcal{C}}_{\mathrm{i}}\|^2 = \operatorname{tr}\left(\sum_{k=0}^{\nu-1} \mathbf{F}_N(k)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{F}_N(k)^{\mathrm{T}}\right) = \operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{i}}\right) = \theta_1 + \cdots + \theta_{n_\infty}.$$

Again, because of the positive semi-definiteness of $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$, we can decompose the Gramian into $\boldsymbol{\mathcal{P}}_{\mathrm{i}} = \mathbf{W}\boldsymbol{\Theta}\mathbf{W}^{\mathrm{T}}$ where $\mathrm{J}\boldsymbol{\Theta} = \operatorname{diag}\left(\theta_1, \ldots, \theta_{n_\infty}, 0\ldots\right)$ includes the eigenvalues of $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ and $\mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}$. As stated in [19], truncating states corresponding to small singular values of the improper Gramians already leads to inaccurate approximations since the non-zero singular values describe constraints to the model, and hence truncation of those could lead to physically meaningless results. However, states corresponding to zero singular values of the improper Gramians correspond to unreachable improper states and can be truncated without changing the input–to–output behavior.

Observability energy

To investigate the observability energies, we first gather the proper and improper observability mappings as

$$\boldsymbol{\mathcal{O}}_{\mathrm{p}}(k, t_1, t_2) := \begin{bmatrix} \boldsymbol{\mathcal{O}}_{\mathrm{pp}}(t_1, t_2) \\ \boldsymbol{\mathcal{O}}_{\mathrm{ip}}(k, t_2) \end{bmatrix}, \qquad \boldsymbol{\mathcal{O}}_{\mathrm{i}}(\ell, k, t) := \begin{bmatrix} \boldsymbol{\mathcal{O}}_{\mathrm{pi}}(\ell, t) \\ \boldsymbol{\mathcal{O}}_{\mathrm{ii}}(\ell, k) \end{bmatrix}.$$

We follow the same methodology as above and evaluate the energy norm of the proper observability mapping, which yields

$$\begin{aligned} E(\mathcal{O}_{\mathrm{p}}) :=& \|\mathcal{O}_{\mathrm{p}}\|^2 = \operatorname{tr}\left(\sum_{k=0}^{\nu-1}\int_0^\infty\int_0^\infty \mathcal{O}_{\mathrm{p}}(k,t_1,t_2)^{\mathrm{T}}\mathcal{O}_{\mathrm{p}}(k,t_1,t_2)\mathrm{d}t_1\mathrm{d}t_2\right) \\ &= \operatorname{tr}\left(\int_0^\infty\int_0^\infty \mathcal{O}_{\mathrm{pp}}(t_1,t_2)^{\mathrm{T}}\mathcal{O}_{\mathrm{pp}}(t_1,t_2)\mathrm{d}t_1\mathrm{d}t_2 + \sum_{k=0}^{\nu-1}\int_0^\infty \mathcal{O}_{\mathrm{ip}}(k,t_2)^{\mathrm{T}}\mathcal{O}_{\mathrm{ip}}(k,t_2)\mathrm{d}t_2\right) \\ &= \operatorname{tr}(\mathcal{Q}_{\mathrm{pp}}) + \operatorname{tr}(\mathcal{Q}_{\mathrm{ip}}) = \operatorname{tr}(\mathcal{Q}_{\mathrm{p}}) \end{aligned}$$

with $\mathcal{Q}_{\mathrm{p}}$ as defined in Definition 3.3. Accordingly, the improper energy norm is defined as

$$\begin{aligned} E(\mathcal{O}_{\mathrm{i}}) :=& \|\mathcal{O}_{\mathrm{i}}\|^2 = \operatorname{tr}\left(\sum_{\ell=0}^{\nu-1}\sum_{k=0}^{\nu-1}\int_0^\infty \mathcal{O}_{\mathrm{i}}(\ell,k,t)^{\mathrm{T}}\mathcal{O}_{\mathrm{i}}(\ell,k,t)\mathrm{d}t\right) \\ &= \operatorname{tr}\left(\sum_{k=0}^{\nu-1}\int_0^\infty \mathcal{O}_{\mathrm{pi}}(k,t)^{\mathrm{T}}\mathcal{O}_{\mathrm{pi}}(k,t)\mathrm{d}t + \sum_{\ell=0}^{\nu-1}\sum_{k=0}^{\nu-1}\mathcal{O}_{\mathrm{ii}}(\ell,k)^{\mathrm{T}}\mathcal{O}_{\mathrm{ii}}(\ell,k)\right) \\ &= \operatorname{tr}(\mathcal{Q}_{\mathrm{pi}}) + \operatorname{tr}(\mathcal{Q}_{\mathrm{ii}}) = \operatorname{tr}(\mathcal{Q}_{\mathrm{i}}), \end{aligned}$$

where $\mathcal{Q}_{\mathrm{i}}$ is the improper observability Gramian as defined in Definition 3.6. Consequently, in the BT method in the following section, we truncate subspaces corresponding to small eigenvalues of $\mathcal{Q}_{\mathrm{p}}$ and zero singular values of $\mathcal{Q}_{\mathrm{i}}$.

4.2 Balanced truncation

In this subsection, we propose an extension of the BT method to the class of systems (1). The Gramians and energies from the previous subsections provide a criterion that states to truncate.

We aim to truncate states corresponding to small eigenvalues of $\mathcal{P}_{\mathrm{p}}$ and $\mathcal{Q}_{\mathrm{p}}$. However, in general, the states corresponding to small eigenvalues of $\mathcal{P}_{\mathrm{p}}$ and those corresponding to small eigenvalues of $\mathcal{Q}_{\mathrm{p}}$ do not coincide. Hence, we need to balance the system before we truncate states.

Definition 4.1. The model in (1) is called *balanced* if the Gramians that are defined as in (9) and in Definition 3.3, and Definition 3.6 satisfy

$$\mathcal{P}_{\mathrm{p}} = \mathcal{Q}_{\mathrm{p}} = \begin{bmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{bmatrix}, \qquad \mathcal{P}_{\mathrm{i}} = \mathcal{Q}_{\mathrm{i}} = \begin{bmatrix}0 & 0\\ 0 & \boldsymbol{\Theta}\end{bmatrix},$$

where $\boldsymbol{\Sigma} = \operatorname{diag}\left(\sigma_1,\ldots,\sigma_{n_f}\right)$ and $\boldsymbol{\Theta} = \operatorname{diag}\left(\theta_1,\ldots,\theta_{n_\infty}\right)$.

To balance the system (1), we need to define the following (low-rank) factors

$$\mathcal{P}_{\mathrm{p}} = \mathbf{R}_{\mathrm{p}}\mathbf{R}_{\mathrm{p}}^{\mathrm{T}}, \quad \mathcal{P}_{\mathrm{i}} = \mathbf{R}_{\mathrm{i}}\mathbf{R}_{\mathrm{i}}^{\mathrm{T}}, \quad \mathcal{Q}_{\mathrm{p}} = \mathbf{S}_{\mathrm{p}}\mathbf{S}_{\mathrm{p}}^{\mathrm{T}}, \quad \mathcal{Q}_{\mathrm{i}} = \mathbf{S}_{\mathrm{i}}\mathbf{S}_{\mathrm{i}}^{\mathrm{T}}.$$

Using these factors, we can compute the following singular value decompositions

$$\begin{aligned} \mathbf{L}_{\mathrm{p}}\mathbf{E}\mathbf{R}_{\mathrm{p}} = \mathbf{U}_{\mathrm{p}}\boldsymbol{\Sigma}\mathbf{V}_{\mathrm{p}}^{\mathrm{T}} &= \begin{bmatrix}\mathbf{U}_{\mathrm{p},1} & \mathbf{U}_{\mathrm{p},2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_1 & \\ & \boldsymbol{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{p},1}^{\mathrm{T}}\\ \mathbf{V}_{\mathrm{p},2}^{\mathrm{T}}\end{bmatrix}, \\ \mathbf{L}_{\mathrm{i}}\mathbf{A}\mathbf{R}_{\mathrm{i}} = \mathbf{U}_{\mathrm{i}}\boldsymbol{\Theta}\mathbf{V}_{\mathrm{i}}^{\mathrm{T}} &= \begin{bmatrix}\mathbf{U}_{\mathrm{i},1} & \mathbf{U}_{\mathrm{i},2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Theta}_1 & \\ & 0\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{i},1}^{\mathrm{T}}\\ \mathbf{V}_{\mathrm{i},2}^{\mathrm{T}}\end{bmatrix}. \end{aligned}$$

We can transform the system by the left and right projection matrix

$$\mathbf{W}_b = [\mathbf{L}_{\mathrm{p}}^{\mathrm{T}}\mathbf{U}_{\mathrm{p}}\boldsymbol{\Sigma}^{-\frac{1}{2}},\ \mathbf{L}_{\mathrm{i}}^{\mathrm{T}}\mathbf{U}_{\mathrm{i}}\boldsymbol{\Theta}^{-\frac{1}{2}}], \quad \mathbf{T}_b = [\mathbf{R}_{\mathrm{p}}^{\mathrm{T}}\mathbf{V}_{\mathrm{p}}\boldsymbol{\Sigma}^{-\frac{1}{2}},\ \mathbf{R}_{\mathrm{i}}^{\mathrm{T}}\mathbf{V}_{\mathrm{i}}\boldsymbol{\Theta}^{-\frac{1}{2}}]$$

to obtain a balanced system that has the form

$$\begin{aligned} \begin{bmatrix}\mathbf{I} & 0\\ 0 & \widetilde{\mathbf{E}}_2\end{bmatrix}\begin{bmatrix}\dot{\widetilde{\mathbf{x}}}_1(t)\\ \dot{\widetilde{\mathbf{x}}}_2(t)\end{bmatrix} &= \begin{bmatrix}\widetilde{\mathbf{A}}_1 & 0\\ 0 & \mathbf{I}\end{bmatrix}\begin{bmatrix}\widetilde{\mathbf{x}}_1(t)\\ \widetilde{\mathbf{x}}_2(t)\end{bmatrix} + \begin{bmatrix}\widetilde{\mathbf{B}}_1\\ \widetilde{\mathbf{B}}_2\end{bmatrix}\mathbf{u}(t), \\ \mathbf{y}(t) &= \begin{bmatrix}\widetilde{\mathbf{x}}_1(t)^{\mathrm{T}} & \widetilde{\mathbf{x}}_2(t)^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\widetilde{\mathbf{M}}_{11} & \widetilde{\mathbf{M}}_{12}\\ \widetilde{\mathbf{M}}_{12}^{\mathrm{T}} & \widetilde{\mathbf{M}}_{22}\end{bmatrix}\begin{bmatrix}\widetilde{\mathbf{x}}_1(t)\\ \widetilde{\mathbf{x}}_2(t)\end{bmatrix}. \end{aligned}$$

Algorithm 1 BT method for for DAE systems with quadratic output.

Input: The full-order model (1) and the order r.

Output: The reduced-order model (18).

1: Compute the proper and improper controllability Gramians $\mathcal{P}_\mathrm{p}$ and $\mathcal{P}_\mathrm{i}$ by solving the Lyapunov equations in (11a), (11b).

2: Compute the proper and improper observability Gramians $\mathcal{Q}_\mathrm{p}$ and $\mathcal{Q}_\mathrm{i}$ by solving the Lyapunov equations from Theorems 3.1 to 3.4.

3: Perform the singular values decompositions

$$\mathbf{L}_\mathrm{p}\mathbf{E}\mathbf{R}_\mathrm{p} = \begin{bmatrix}\mathbf{U}_{\mathrm{p},1} & \mathbf{U}_{\mathrm{p},2}\end{bmatrix}\begin{bmatrix}\mathbf{\Sigma}_1 & \\ & \mathbf{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{p},1}^\mathrm{T} \\ \mathbf{V}_{\mathrm{p},2}^\mathrm{T}\end{bmatrix}, \quad \mathbf{L}_\mathrm{i}\mathbf{A}\mathbf{R}_\mathrm{i} = \begin{bmatrix}\mathbf{U}_{\mathrm{i},1} & \mathbf{U}_{\mathrm{i},2}\end{bmatrix}\begin{bmatrix}\mathbf{\Theta}_1 & \\ & 0\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{i},1}^\mathrm{T} \\ \mathbf{V}_{\mathrm{i},2}^\mathrm{T}\end{bmatrix}.$$

4: Construct the projection matrices

$$\mathbf{W}_\mathrm{r} = [\mathbf{L}_\mathrm{p}^\mathrm{T}\mathbf{U}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}},\ \mathbf{L}_\mathrm{i}^\mathrm{T}\mathbf{U}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}], \quad \mathbf{T}_\mathrm{r} = [\mathbf{R}_\mathrm{p}^\mathrm{T}\mathbf{V}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}},\ \mathbf{R}_\mathrm{i}^\mathrm{T}\mathbf{V}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}].$$

5: Construct reduced matrices

$$\widehat{\mathbf{E}} := \mathbf{W}_\mathrm{r}^\mathrm{T}\mathbf{E}\mathbf{T}_\mathrm{r}, \quad \widehat{\mathbf{A}} := \mathbf{W}_\mathrm{r}^\mathrm{T}\mathbf{A}\mathbf{T}_\mathrm{r}, \quad \widehat{\mathbf{B}} := \mathbf{W}_\mathrm{r}^\mathrm{T}\mathbf{B}, \quad \widehat{\mathbf{M}} := \mathbf{T}_\mathrm{r}^\mathrm{T}\mathbf{M}\mathbf{T}_\mathrm{r}.$$

We see again a decomposition into a proper state $\widetilde{\mathbf{x}}_1(t)$ and an improper state $\widetilde{\mathbf{x}}_2(t)$, where the matrix $\widetilde{\mathbf{E}}_2$ is nilpotent.

We reduced the system by truncating the proper states $\widetilde{\mathbf{x}}_1(t)$ that are most difficult to reach and to observe. As we have seen in the previous subsection, these states correspond to the smallest singular values in $\mathbf{\Sigma}$, i.e. $\mathbf{\Sigma}_2$. The projection matrices that balance the system and truncate these states simultaneously are

$$\mathbf{W}_\mathrm{r} = [\mathbf{L}_\mathrm{p}^\mathrm{T}\mathbf{U}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}},\ \mathbf{L}_\mathrm{i}^\mathrm{T}\mathbf{U}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}], \quad \mathbf{T}_\mathrm{r} = [\mathbf{R}_\mathrm{p}^\mathrm{T}\mathbf{V}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}},\ \mathbf{R}_\mathrm{i}^\mathrm{T}\mathbf{V}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}].$$

We multiply the full-order model in (1) by the projection matrices $\mathbf{W}_\mathrm{r}$ and $\mathbf{T}_\mathrm{r}$ to obtain a reduced-order model (2) that is of the form

$$\begin{aligned}\begin{bmatrix}\mathbf{I} & 0 \\ 0 & \widetilde{\mathbf{E}}_2\end{bmatrix}\begin{bmatrix}\dot{\widehat{\mathbf{x}}}_1(t) \\ \dot{\widetilde{\mathbf{x}}}_2(t)\end{bmatrix} &= \begin{bmatrix}\widehat{\mathbf{A}}_1 & 0 \\ 0 & \mathbf{I}\end{bmatrix}\begin{bmatrix}\widehat{\mathbf{x}}_1(t) \\ \widetilde{\mathbf{x}}_2(t)\end{bmatrix} + \begin{bmatrix}\widehat{\mathbf{B}}_1 \\ \widetilde{\mathbf{B}}_2\end{bmatrix}\mathbf{u}(t), \\ \widehat{\mathbf{y}}(t) &= \begin{bmatrix}\widehat{\mathbf{x}}_1(t)^\mathrm{T} & \widetilde{\mathbf{x}}_2(t)^\mathrm{T}\end{bmatrix}\begin{bmatrix}\widehat{\mathbf{M}}_{11} & \widehat{\mathbf{M}}_{12} \\ \widehat{\mathbf{M}}_{12}^\mathrm{T} & \widehat{\mathbf{M}}_{22}\end{bmatrix}\begin{bmatrix}\widehat{\mathbf{x}}_1(t) \\ \widetilde{\mathbf{x}}_2(t)\end{bmatrix}\end{aligned} \tag{18}$$

where r is the dimension of the reduced proper space. The reduction by BT is summarized in Algorithm 1. This algorithm follows the same line as the BT method for differential-algebraic systems with linear output presented in [19], besides the fact that the observability Gramians are different in our algorithm.

Remark 1. The BT method presented above decouples the proper and improper states as described in (18) where the proper states are reduced while for the improper states, only a minimal realization is found. That means that improper states corresponding to zero singular values of the improper Gramians are truncated since they are not reachable or not observable and do not change the input–to–output behavior of the system.

5 Error Estimation

We aim to estimate the error between the output $\mathbf{y}$ and the reduced output $\widehat{\mathbf{y}}$ we obtain when we evaluate the reduced-order model (18). We estimate

$$\|\mathbf{y} - \widehat{\mathbf{y}}\|_{L_\infty} \leq \|\mathbf{y}_\mathrm{pp} - \widehat{\mathbf{y}}_\mathrm{pp}\|_{L_\infty} + \|\mathbf{y}_\mathrm{pi} - \widehat{\mathbf{y}}_\mathrm{pi}\|_{L_\infty} + \|\mathbf{y}_\mathrm{ip} - \widehat{\mathbf{y}}_\mathrm{ip}\|_{L_\infty} + \|\mathbf{y}_\mathrm{ii} - \widehat{\mathbf{y}}_\mathrm{ii}\|_{L_\infty}$$

and consider the four summands separately. Since we do not truncate the improper states the summand $\|\mathbf{y}_\mathrm{ii} - \widehat{\mathbf{y}}_\mathrm{ii}\|_{L_\infty}$ is equal to zero. The remaining summands are investigated in the following.

5.1 The proper-proper output error

In this section, we aim to analyze the error between the proper-proper output $\mathbf{y}_{\mathrm{pp}}(t)$ and its approximation $\widehat{\mathbf{y}}_{\mathrm{pp}}(t)$. To do so, we define

$$\mathbf{h}_{\mathrm{pp}}(t_1,t_2) := \operatorname{vec}\left(\mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t_1)^{\mathrm{T}}\mathbf{M}\mathbf{F}_J(t_2)\mathbf{B}\right) \quad \text{and} \quad \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2) := \operatorname{vec}\left(\widehat{\mathbf{B}}^{\mathrm{T}}\widehat{\mathbf{F}}_J(t_1)^{\mathrm{T}}\widehat{\mathbf{M}}\widehat{\mathbf{F}}_J(t_2)\widehat{\mathbf{B}}\right), \tag{19}$$

where $\widehat{\mathbf{F}}_J(t) := \begin{bmatrix} e^{\widehat{\mathbf{A}}_1 t} & 0 \\ 0 & 0 \end{bmatrix}$, so that the outputs can be represented as

$$\begin{aligned}
\mathbf{y}_{\mathrm{pp}}(t) &= \mathbf{x}_{\mathrm{p}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{p}}(t) = \int_0^t \int_0^t \mathbf{h}_{\mathrm{pp}}(t_1,t_2)(\mathbf{u}(t_2)\otimes\mathbf{u}(t_1))\mathrm{d}t_1\mathrm{d}t_2,\\
\widehat{\mathbf{y}}_{\mathrm{pp}}(t) &= \widehat{\mathbf{x}}_{\mathrm{p}}(t)^{\mathrm{T}}\widehat{\mathbf{M}}\widehat{\mathbf{x}}_{\mathrm{p}}(t) = \int_0^t \int_0^t \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2)(\mathbf{u}(t_2)\otimes\mathbf{u}(t_1))\mathrm{d}t_1\mathrm{d}t_2.
\end{aligned}$$

Using these representations of $\mathbf{y}_{\mathrm{pp}}$ and $\widehat{\mathbf{y}}_{\mathrm{pp}}$ the following lemma provides an upper bound of the L_∞-error in the proper proper output.

Lemma 5.1. We consider the asymptotically stable DAE system with a quadratic output equation from (1), the reduced-order model in (18), and $\mathbf{h}_{\mathrm{pp}}(t_1,t_2)$, $\widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2)$ as defined in (19). Then, the following inequality holds

$$\|\mathbf{y}_{\mathrm{pp}} - \widehat{\mathbf{y}}_{\mathrm{pp}}\|_{L_\infty} \le \left(\int_0^\infty \int_0^\infty \|\mathbf{h}_{\mathrm{pp}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2)\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \|\mathbf{u}\otimes\mathbf{u}\|_{L_2}.$$

Proof. We consider the output error at time $t \ge 0$ that is

$$\begin{aligned}
\left|\mathbf{y}_{\mathrm{pp}}(t) - \widehat{\mathbf{y}}_{\mathrm{pp}}(t)\right| &= \left|\int_0^t \int_0^t \mathbf{x}_{\mathrm{p}}(t_1)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{p}}(t_2) - \widehat{\mathbf{x}}_{\mathrm{p}}(t_1)^{\mathrm{T}}\widehat{\mathbf{M}}\widehat{\mathbf{x}}_{\mathrm{p}}(t_2)\mathrm{d}t_1\mathrm{d}t_2\right|\\
&= \left|\int_0^t \int_0^t \left(\mathbf{h}_{\mathrm{pp}}(t-t_1,t-t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t-t_1,t-t_2)\right)(\mathbf{u}(t_2)\otimes\mathbf{u}(t_1))\mathrm{d}t_1\mathrm{d}t_2\right|.
\end{aligned}$$

Applying the Cauchy-Schwarz inequality multiple times yields

$$\begin{aligned}
|\mathbf{y}_{\mathrm{pp}}(t) - \widehat{\mathbf{y}}_{\mathrm{pp}}(t)| &\le \int_0^t \int_0^t \| \left(\mathbf{h}_{\mathrm{pp}}(t-t_1,t-t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t-t_1,t-t_2)\right)(\mathbf{u}(t_2)\otimes\mathbf{u}(t_1))\|\mathrm{d}t_1\mathrm{d}t_2\\
&\le \int_0^t \int_0^t \|\mathbf{h}_{\mathrm{pp}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2)\|_2 \|(\mathbf{u}(t_2)\otimes\mathbf{u}(t_1))\|_2 \mathrm{d}t_1\mathrm{d}t_2\\
&\le \left(\int_0^t \int_0^t \|\mathbf{h}_{\mathrm{pp}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2)\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \left(\int_0^t \int_0^t \|(\mathbf{u}(t_2)\otimes\mathbf{u}(t_1))\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}}.
\end{aligned}$$

Hence, we can bound the L_∞-norm of the output error as

$$\begin{aligned}
&\|\mathbf{y}_{\mathrm{pp}} - \widehat{\mathbf{y}}_{\mathrm{pp}}\|_{L_\infty}\\
&\quad\le \left(\int_0^\infty \int_0^\infty \|\mathbf{h}_{\mathrm{pp}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2)\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \left(\int_0^\infty \int_0^\infty \|(\mathbf{u}(t_2)\otimes\mathbf{u}(t_1))\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}}\\
&\quad= \left(\int_0^\infty \int_0^\infty \|\mathbf{h}_{\mathrm{pp}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2)\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \|\mathbf{u}\otimes\mathbf{u}\|_{L_2}.
\end{aligned}$$

□

Lemma 5.2. We consider the asymptotically stable DAE system with a quadratic output equation from (1), the reduced-order model in (18), the corresponding proper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ as defined in (9), and the reduced proper controllability Gramian

$$\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} := \int_0^\infty \begin{bmatrix} e^{\widehat{\mathbf{A}}_1 t}\widehat{\mathbf{B}}_1\widehat{\mathbf{B}}_1^{\mathrm{T}}e^{\widehat{\mathbf{A}}_1^{\mathrm{T}} t} & 0 \\ 0 & 0 \end{bmatrix} \mathrm{d}t.$$

The functionals $\mathbf{h}_{\mathrm{pp}}(t_1, t_2)$ and $\widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)$ are as defined in (19). Then, the following equalities hold

$$\int_0^\infty \int_0^\infty \|\mathbf{h}_{\mathrm{pp}}(t_1, t_2)\|_2^2 \mathrm{d}t_1 \mathrm{d}t_2 = \operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{p}} \mathbf{M} \boldsymbol{\mathcal{P}}_{\mathrm{p}} \mathbf{M}\right), \tag{20}$$

$$\int_0^\infty \int_0^\infty \|\widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\|_2^2 \mathrm{d}t_1 \mathrm{d}t_2 = \operatorname{tr}\left(\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{M}} \widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{M}}\right), \tag{21}$$

$$\int_0^\infty \int_0^\infty \langle \mathbf{h}_{\mathrm{pp}}(t_1, t_2), \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2) \rangle \mathrm{d}t_1 \mathrm{d}t_2 = \operatorname{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}^{\mathrm{T}} \mathbf{M} \widetilde{\mathbf{P}}_{\mathrm{p}} \widehat{\mathbf{M}}\right) \tag{22}$$

where $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} := \int_0^\infty \mathbf{F}_J(t) \mathbf{B} \widehat{\mathbf{B}}^{\mathrm{T}} \widehat{\mathbf{F}}_J(t)^{\mathrm{T}} \mathrm{d}t$ satisfies the projected Sylvester equation

$$\mathbf{A} \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{E}}^{\mathrm{T}} + \mathbf{E} \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{A}}^{\mathrm{T}} = -\mathbf{P}_{\mathrm{l}} \mathbf{B} \widehat{\mathbf{B}}^{\mathrm{T}} \widehat{\mathbf{P}}_{\mathrm{l}}^{\mathrm{T}}, \qquad \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} = \mathbf{P}_{\mathrm{r}} \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{P}}_{\mathrm{r}}^{\mathrm{T}} \tag{23}$$

with $\widehat{\mathbf{P}}_{\mathrm{l}} = \widehat{\mathbf{P}}_{\mathrm{r}} = \begin{bmatrix} \mathbf{I}_{\mathrm{r}} & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{P}_{\mathrm{l}}$ and $\mathbf{P}_{\mathrm{r}}$ as defined in (3), and $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product with $\langle v_1, v_2 \rangle = v_1^{\mathrm{T}} v_2$ for $v_1, v_2 \in \mathbb{R}^n$.

Proof. We make use of the property $\|\mathrm{vec}(X)\|_2^2 = \|X\|_F^2$ and the Kronecker product properties to obtain

$$\begin{aligned}
\int_0^\infty \int_0^\infty \|\mathbf{h}_{\mathrm{pp}}(t_1, t_2)\|_2^2 \mathrm{d}t_1 \mathrm{d}t_2 &= \int_0^\infty \int_0^\infty \operatorname{tr}\left(\mathbf{B}^{\mathrm{T}} \mathbf{F}_J(t_2)^{\mathrm{T}} \mathbf{M} \mathbf{F}_J(t_1) \mathbf{B} \mathbf{B}^{\mathrm{T}} \mathbf{F}_J(t_1)^{\mathrm{T}} \mathbf{M} \mathbf{F}_J(t_2) \mathbf{B}\right) \mathrm{d}t_1 \mathrm{d}t_2 \\
&= \int_0^\infty \operatorname{tr}\left(\mathbf{B}^{\mathrm{T}} \mathbf{F}_J(t_2)^{\mathrm{T}} \mathbf{M} \boldsymbol{\mathcal{P}}_{\mathrm{p}} \mathbf{M} \mathbf{F}_J(t_2) \mathbf{B}\right) \mathrm{d}t_2 \\
&= \int_0^\infty \operatorname{tr}\left(\mathbf{F}_J(t_2) \mathbf{B} \mathbf{B}^{\mathrm{T}} \mathbf{F}_J(t_2)^{\mathrm{T}} \mathbf{M} \boldsymbol{\mathcal{P}}_{\mathrm{p}} \mathbf{M}\right) \mathrm{d}t_2 \\
&= \operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{p}} \mathbf{M} \boldsymbol{\mathcal{P}}_{\mathrm{p}} \mathbf{M}\right),
\end{aligned}$$

which proves (20) while (21) is proven analogously. To show the last equation given in (22) we make use of the property $\langle \mathrm{vec}(X), \mathrm{vec}(Y) \rangle = \operatorname{tr}\left(X^{\mathrm{T}} Y\right)$ and obtain

$$\begin{aligned}
&\int_0^\infty \int_0^\infty \langle \mathbf{h}_{\mathrm{pp}}(t_1, t_2), \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2) \rangle \mathrm{d}t_1 \mathrm{d}t_2 \\
&\qquad = \int_0^\infty \int_0^\infty \operatorname{tr}\left(\mathbf{B}^{\mathrm{T}} \mathbf{F}_J(t_2)^{\mathrm{T}} \mathbf{M} \mathbf{F}_J(t_1) \mathbf{B} \widehat{\mathbf{B}}^{\mathrm{T}} \widehat{\mathbf{F}}_J(t_1)^{\mathrm{T}} \widehat{\mathbf{M}} \widehat{\mathbf{F}}_J(t_2) \widehat{\mathbf{B}}\right) \mathrm{d}t_1 \mathrm{d}t_2 \\
&\qquad = \int_0^\infty \int_0^\infty \operatorname{tr}\left(\widehat{\mathbf{F}}_J(t_2) \widehat{\mathbf{B}} \mathbf{B}^{\mathrm{T}} \mathbf{F}_J(t_2)^{\mathrm{T}} \mathbf{M} \mathbf{F}_J(t_1) \mathbf{B} \widehat{\mathbf{B}}^{\mathrm{T}} \widehat{\mathbf{F}}_J(t_1)^{\mathrm{T}} \widehat{\mathbf{M}}\right) \mathrm{d}t_1 \mathrm{d}t_2 \\
&\qquad = \operatorname{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}^{\mathrm{T}} \mathbf{M} \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{M}}\right).
\end{aligned}$$

To show that $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}$ solves the Sylvester equation in (23), we follow the same argumentation as for Theorem 3.1. □

Lemma 5.1 and 5.2 result in the following theorem.

Theorem 5.1. We consider the asymptotically stable DAE system with a quadratic output equation from (1), the reduced-order model in (18), the corresponding proper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ as defined in (9), the reduced proper controllability Gramian $\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}$, and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}$ as defined in Lemma 5.2. The error between the proper-proper output $\mathbf{y}_{\mathrm{pp}}(t)$ of the full-order model (1) and the reduced output $\widehat{\mathbf{y}}_{\mathrm{pp}}(t)$ satisfies the following bound:

$$\|\mathbf{y}_{\mathrm{pp}} - \widehat{\mathbf{y}}_{\mathrm{pp}}\|_{L_\infty}^2 \leq \left(\operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{p}} \mathbf{M} \boldsymbol{\mathcal{P}}_{\mathrm{p}} \mathbf{M}\right) - 2 \operatorname{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}^{\mathrm{T}} \mathbf{M} \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{M}}\right) + \operatorname{tr}\left(\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{M}} \widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} \widehat{\mathbf{M}}\right)\right) \|\mathbf{u} \otimes \mathbf{u}\|_{L_2}.$$

5.2 The improper-proper output error

We want to estimate the improper-proper output error, i.e., the error between the improper-proper output $\mathbf{y}_{\mathrm{ip}}(t)$ and the reduced improper-proper output $\widehat{\mathbf{y}}_{\mathrm{ip}}(t)$. We define

$$\mathbf{h}_{\mathrm{ip}}(t,k) := \mathbf{B}^{\mathrm{T}}\mathbf{F}_N(k)^{\mathrm{T}}\mathbf{M}\mathbf{F}_J(t)\mathbf{B} \quad \text{and} \quad \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k) := \widehat{\mathbf{B}}^{\mathrm{T}}\widehat{\mathbf{F}}_N(k)^{\mathrm{T}}\widehat{\mathbf{M}}\widehat{\mathbf{F}}_J(t)\widehat{\mathbf{B}} \tag{24}$$

to obtain the output representations

$$\begin{aligned}
\mathbf{y}_{\mathrm{ip}}(t) &= \mathbf{x}_{\mathrm{i}}(t)^{\mathrm{T}}\mathbf{M}\mathbf{x}_{\mathrm{p}}(t) = \int_0^t \sum_{k=0}^{\nu-1} \mathbf{h}_{\mathrm{ip}}(t-\tau,k)(\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t))\mathrm{d}\tau,\\
\widehat{\mathbf{y}}_{\mathrm{ip}}(t) &= \widehat{\mathbf{x}}_{\mathrm{i}}(t)^{\mathrm{T}}\widehat{\mathbf{M}}\widehat{\mathbf{x}}_{\mathrm{p}}(t) = \int_0^t \sum_{k=0}^{\nu-1} \widehat{\mathbf{h}}_{\mathrm{ip}}(t-\tau,k)(\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t))\mathrm{d}\tau.
\end{aligned}$$

This representation of the improper-proper output can be used to obtain a bound of the L_∞-error.

Lemma 5.3. We consider the DAE system with a quadratic output equation from (1), the reduced-order model in (18), and $\mathbf{h}_{\mathrm{ip}}(t,k)$, $\widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)$ as defined in (24). Then, the following bound holds

$$\|\mathbf{y}_{\mathrm{ip}} - \widehat{\mathbf{y}}_{\mathrm{ip}}\|_{L_\infty} \leq \left(\int_0^\infty \sum_{k=0}^{\nu-1} \|\mathbf{h}_{\mathrm{ip}}(t,k) - \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}} \left(\int_0^\infty \sum_{k=0}^{\nu-1} \|\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}}.$$

Proof. Using the definition (24), we obtain

$$\left|\mathbf{y}_{\mathrm{ip}}(t) - \widehat{\mathbf{y}}_{\mathrm{ip}}(t)\right| = \left|\int_0^t \sum_{k=0}^{\nu-1} \left(\mathbf{h}_{\mathrm{ip}}(t-\tau,k) - \widehat{\mathbf{h}}_{\mathrm{ip}}(t-\tau,k)\right)\left(\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\right)\mathrm{d}\tau\right|.$$

By applying the Cauchy-Schwarz inequality multiple times, we obtain the following estimations

$$\begin{aligned}
\left|\mathbf{y}_{\mathrm{ip}}(t) - \widehat{\mathbf{y}}_{\mathrm{ip}}(t)\right| &\leq \int_0^t \left|\sum_{k=0}^{\nu-1} \left(\mathbf{h}_{\mathrm{ip}}(t-\tau,k) - \widehat{\mathbf{h}}_{\mathrm{ip}}(t-\tau,k)\right)\left(\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\right)\right|\mathrm{d}\tau\\
&\leq \int_0^t \left(\sum_{k=0}^{\nu-1} \|\mathbf{h}_{\mathrm{ip}}(t-\tau,k) - \widehat{\mathbf{h}}_{\mathrm{ip}}(t-\tau,k)\|_2^2\right)^{\frac{1}{2}} \left(\sum_{k=0}^{\nu-1} \|\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\|_2^2\right)^{\frac{1}{2}} \mathrm{d}\tau\\
&\leq \left(\int_0^t \sum_{k=0}^{\nu-1} \|\mathbf{h}_{\mathrm{ip}}(t,k) - \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}} \left(\int_0^t \sum_{k=0}^{\nu-1} \|\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}}.
\end{aligned}$$

such that the L_∞-norm of the output error is bounded by

$$\|\mathbf{y}_{\mathrm{ip}} - \widehat{\mathbf{y}}_{\mathrm{ip}}\|_{L_\infty} \leq \left(\int_0^\infty \sum_{k=0}^{\nu-1} \|\mathbf{h}_{\mathrm{ip}}(t,k) - \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}} \left(\int_0^\infty \sum_{k=0}^{\nu-1} \|\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}}.$$

□

Lemma 5.4. We consider the asymptotically stable DAE system with a quadratic output equation from (1), the reduced-order model in (18), the corresponding proper and improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ as defined in (9), and the reduced proper and improper controllability Gramians

$$\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}} := \int_0^\infty \begin{bmatrix} e^{\widehat{\mathbf{A}}_1 t}\widehat{\mathbf{B}}_1\widehat{\mathbf{B}}_1^{\mathrm{T}}e^{\widehat{\mathbf{A}}_1^{\mathrm{T}} t} & 0\\ 0 & 0\end{bmatrix}\mathrm{d}t, \qquad \widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{i}} := \sum_{k=0}^{\nu-1} \begin{bmatrix} 0 & 0\\ 0 & \widetilde{\mathbf{E}}_2^k\widetilde{\mathbf{B}}_2\widetilde{\mathbf{B}}_2^{\mathrm{T}}\left(\widetilde{\mathbf{E}}_2^k\right)^{\mathrm{T}}\end{bmatrix}.$$

The functionals $\mathbf{h}_{\mathrm{ip}}(t,k)$ and $\widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)$ are as defined in (24). Then the following equalities hold

$$\int_0^\infty \sum_{k=0}^{\nu-1} \|\mathbf{h}_{\mathrm{ip}}(t,k)\|_2^2 \mathrm{d}t = \operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{i}}\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{p}}\mathbf{M}\right), \tag{25}$$

$$\int_0^\infty \sum_{k=0}^{\nu-1} \|\widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\|_2^2 \mathrm{d}t_1 \mathrm{d}t_2 = \operatorname{tr}\left(\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}\widehat{\mathbf{M}}\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}\widehat{\mathbf{M}}\right), \tag{26}$$

$$\int_0^\infty \sum_{k=0}^{\nu-1} \langle \mathbf{h}_{\mathrm{ip}}(t,k), \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\rangle \mathrm{d}t = \operatorname{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}^{\mathrm{T}}\mathbf{M}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}\widehat{\mathbf{M}}\right) \tag{27}$$

where $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}$ is as in Lemma 5.2 and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i}} := \sum_{k=0}^{\nu-1} \mathbf{F}_N(k)\mathbf{B}\widehat{\mathbf{B}}^{\mathrm{T}}\widehat{\mathbf{F}}_N(k)^{\mathrm{T}}$ satisfies the projected Sylvester equation

$$\mathbf{A}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}\widehat{\mathbf{A}}^{\mathrm{T}} - \mathbf{E}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}\widehat{\mathbf{E}}^{\mathrm{T}} = (\mathbf{I}-\mathbf{P}_{\mathrm{l}})\mathbf{B}\widehat{\mathbf{B}}^{\mathrm{T}}(\mathbf{I}-\widehat{\mathbf{P}}_{\mathrm{l}}^{\mathrm{T}}), \qquad 0 = \mathbf{P}_{\mathrm{r}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}\widehat{\mathbf{P}}_{\mathrm{r}}^{\mathrm{T}} \tag{28}$$

with $\widehat{\mathbf{P}}_{\mathrm{l}} = \widehat{\mathbf{P}}_{\mathrm{r}} = \begin{bmatrix} \mathbf{I}_{\mathrm{r}} & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{P}_{\mathrm{l}}$ and $\mathbf{P}_{\mathrm{r}}$ as defined in (3), and $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product with $\langle v_1, v_2\rangle = v_1^{\mathrm{T}} v_2$ for $v_1, v_2 \in \mathbb{R}^n$.

Proof. The proof is analogous to the one from Lemma 5.2. □

Theorem 5.2. For $\mathbf{u} \in \mathbb{C}^{\nu-1}([0,\infty),\mathbb{R}^m)$ it holds

$$\int_0^t \sum_{k=0}^{\nu-1} \|\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\|_2^2 \mathrm{d}\tau \le \nu \|\mathbf{u}\|_{\boldsymbol{\mathcal{C}}^{\nu-1}}^2 \|\mathbf{u}\|_{L_2}^2.$$

Proof. Applying Kronecker product properties and Cauchy-Schwarz inequality yields

$$\begin{aligned}
\int_0^t \sum_{k=0}^{\nu-1} \|\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\|_2^2 \mathrm{d}\tau &= \int_0^t \sum_{k=0}^{\nu-1} (\mathbf{u}^{(k)}(t)\otimes\mathbf{u}(\tau))^{\mathrm{T}}(\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t))\mathrm{d}\tau \\
&= \int_0^t \sum_{k=0}^{\nu-1} (\mathbf{u}^{(k)}(t)^{\mathrm{T}}\otimes\mathbf{u}(\tau)^{\mathrm{T}})(\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t))\mathrm{d}\tau \\
&= \int_0^t \sum_{k=0}^{\nu-1} (\mathbf{u}^{(k)}(t)^{\mathrm{T}}\mathbf{u}(\tau))\otimes(\mathbf{u}(\tau)^{\mathrm{T}}\mathbf{u}^{(k)}(t))\mathrm{d}\tau \\
&= \int_0^t \sum_{k=0}^{\nu-1} \mathbf{u}^{(k)}(t)^{\mathrm{T}}\mathbf{u}(\tau)\mathbf{u}(\tau)^{\mathrm{T}}\mathbf{u}^{(k)}(t)\mathrm{d}\tau \\
&\le \sum_{k=0}^{\nu-1} \int_0^\infty \|\mathbf{u}(\tau)\|^2 \mathrm{d}\tau \|\mathbf{u}^{(k)}(t)\|^2 \\
&= \sum_{k=0}^{\nu-1} \|\mathbf{u}(\tau)\|_{L_2}^2 \|\mathbf{u}^{(k)}(t)\|^2 \le \nu\|\mathbf{u}\|_{\boldsymbol{\mathcal{C}}^{\nu-1}}^2\|\mathbf{u}\|_{L_2}^2
\end{aligned}$$

for $\|\mathbf{u}\|_{\boldsymbol{\mathcal{C}}^{\nu-1}} := \max_{k=0,\dots,\nu-1} \sup_{t\ge 0} \|\mathbf{u}\|_2$. □

Together with Theorem 5.2, Lemma 5.3 and Lemma 5.4 we obtain the following theorem.

Theorem 5.3. We consider the asymptotically stable DAE system with a quadratic output equation (1), the reduced-order model in (18), the corresponding proper and improper controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p}}$, $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ as defined in (9), the reduced proper and improper controllability Gramians $\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}$, $\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}$, and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}$, $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}$ as defined in Lemma 5.4. The error between the improper-proper output $\mathbf{y}_{\mathrm{ip}}$ of the full-order model (1) and the reduced output $\widehat{\mathbf{y}}_{\mathrm{ip}}$ satisfies the following bound:

$$\|\mathbf{y}_{\mathrm{ip}} - \widehat{\mathbf{y}}_{\mathrm{ip}}\|_{L_\infty} \le \left(\operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{p}}\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\mathbf{M}\right) - 2\operatorname{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}^{\mathrm{T}}\mathbf{M}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}\widehat{\mathbf{M}}\right) + \operatorname{tr}\left(\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}\widehat{\mathbf{M}}\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}\widehat{\mathbf{M}}\right)\right)^{\frac{1}{2}} \nu^{\frac{1}{2}} \|\mathbf{u}\|_{\boldsymbol{\mathcal{C}}^{\nu-1}} \|\mathbf{u}\|_{L_2}$$

for output functions $\mathbf{u} \in \mathbb{C}^{\nu-1}([0,\infty),\mathbb{R}^m)$.

Since $\|\mathbf{y}_{\mathrm{pi}} - \widehat{\mathbf{y}}_{\mathrm{pi}}\|_{L_\infty}^2$ is equal to $\|\mathbf{y}_{\mathrm{ip}} - \widehat{\mathbf{y}}_{\mathrm{ip}}\|_{L_\infty}^2$, and the improper states are not truncated the overall error $\|\mathbf{y} - \widehat{\mathbf{y}}\|_{L_\infty}$ can be estimated as

$$\begin{aligned}\|\mathbf{y} - \widehat{\mathbf{y}}\|_{L_\infty} &\leq \|\mathbf{y}_{\mathrm{pp}} - \widehat{\mathbf{y}}_{\mathrm{pp}}\|_{L_\infty} + 2 \cdot \|\mathbf{y}_{\mathrm{pi}} - \widehat{\mathbf{y}}_{\mathrm{pi}}\|_{L_\infty} \\ &\leq \left(\operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{p}}\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{p}}\mathbf{M}\right) - 2\operatorname{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}^{\mathrm{T}}\mathbf{M}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}\widehat{\mathbf{M}}\right) + \operatorname{tr}\left(\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}\widehat{\mathbf{M}}\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}\widehat{\mathbf{M}}\right)\right)^{\frac{1}{2}} \|\mathbf{u}\otimes\mathbf{u}\|_{L_2}^{\frac{1}{2}} \\ &\quad + 2 \cdot \left(\operatorname{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{p}}\mathbf{M}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\mathbf{M}\right) - 2\operatorname{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}^{\mathrm{T}}\mathbf{M}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}\widehat{\mathbf{M}}\right) + \operatorname{tr}\left(\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{p}}\widehat{\mathbf{M}}\widehat{\boldsymbol{\mathcal{P}}}_{\mathrm{i}}\widehat{\mathbf{M}}\right)\right)^{\frac{1}{2}} \nu^{\frac{1}{2}}\|\mathbf{u}\|_{\boldsymbol{\mathcal{C}}^{\nu-1}}\|\mathbf{u}\|_{L_2}.\end{aligned} \tag{29}$$

6 Extension to the multiple output case

Up to now, we considered systems (1) with a single output. In this section, however, we will extend the previous theory to the multiple output case, where the output is given as

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \begin{bmatrix}\mathbf{x}(t)^{\mathrm{T}}\mathbf{M}_1\mathbf{x}(t)\\ \vdots \\ \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}_p\mathbf{x}(t)\end{bmatrix} \tag{30}$$

where $\mathbf{C} \in \mathbb{R}^{p\times n}$ and $\mathbf{M}_k = \mathbf{M}_k^{\mathrm{T}} \in \mathbb{R}^{n\times n}$ for all $k = 1,\ldots,p$. As we already did for the DAE system with one quadratic output (1) we consider the different parts of the output separately so that we investigate

$$\mathbf{y}_C(t) := \mathbf{C}\mathbf{x}(t), \qquad \mathbf{y}_j(t) := \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}_j\mathbf{x}(t), \quad j = 1,\ldots,p$$

and derive the corresponding Gramians that are summed up in the end to derive Gramians that cover the overall observability behavior.

For the linear term $\mathbf{y}_C(t)$, define the proper and improper observability mapping

$$\boldsymbol{\mathcal{C}}_{\mathrm{p}}^C(t) := \mathbf{C}\mathbf{F}_J(t), \qquad \boldsymbol{\mathcal{C}}_{\mathrm{i}}^C(k) := \mathbf{C}\mathbf{F}_N(k)$$

and apply the theory from [24] to obtain the proper and improper observability Gramian

$$\begin{aligned}\boldsymbol{\mathcal{Q}}_{\mathrm{p}}^C &:= \int_0^\infty \left(\boldsymbol{\mathcal{C}}_{\mathrm{p}}^C(t)\right)^{\mathrm{T}}\boldsymbol{\mathcal{C}}_{\mathrm{p}}^C(t)\mathrm{d}t := \int_0^\infty \mathbf{F}_J(t)^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{C}\mathbf{F}_J(t)\mathrm{d}t,\\ \boldsymbol{\mathcal{Q}}_{\mathrm{i}}^C &:= \sum_{k=0}^{\nu-1}\left(\boldsymbol{\mathcal{C}}_{\mathrm{i}}^C(k)\right)^{\mathrm{T}}\boldsymbol{\mathcal{C}}_{\mathrm{i}}^C(k) := \sum_{k=0}^{\nu-1}\mathbf{F}_N(k)^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{C}\mathbf{F}_N(k).\end{aligned}$$

For each of the quadratic components, we define the observability mappings

$$\begin{aligned}\boldsymbol{\mathcal{O}}_{\mathrm{pp}}^j(t_1,t_2) &:= \mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t_1)^{\mathrm{T}}\mathbf{M}_j\mathbf{F}_J(t_2), & \boldsymbol{\mathcal{O}}_{\mathrm{pi}}^j(t,k) &:= \mathbf{B}^{\mathrm{T}}\mathbf{F}_J(t)^{\mathrm{T}}\mathbf{M}_j\mathbf{F}_N(k),\\ \boldsymbol{\mathcal{O}}_{\mathrm{ip}}^j(t,k) &:= \mathbf{B}^{\mathrm{T}}\mathbf{F}_N(k)^{\mathrm{T}}\mathbf{M}_j\mathbf{F}_J(t), & \boldsymbol{\mathcal{O}}_{\mathrm{ii}}^j(\ell,k) &:= \mathbf{B}^{\mathrm{T}}\mathbf{F}_N(\ell)^{\mathrm{T}}\mathbf{M}_j\mathbf{F}_J(t)\end{aligned}$$

and apply the theory from Section 3 to derive the corresponding observability Gramians

$$\boldsymbol{\mathcal{Q}}_{\mathrm{p}}^j := \int_0^\infty \mathbf{F}_J(t)^{\mathrm{T}}\mathbf{M}_j(\boldsymbol{\mathcal{P}}_{\mathrm{p}} + \boldsymbol{\mathcal{P}}_{\mathrm{i}})\mathbf{M}_j\mathbf{F}_J(t)\mathrm{d}t, \quad \boldsymbol{\mathcal{Q}}_{\mathrm{i}}^j := \sum_{k=0}^{\nu-1}\mathbf{F}_N(k)^{\mathrm{T}}\mathbf{M}_j(\boldsymbol{\mathcal{P}}_{\mathrm{p}} + \boldsymbol{\mathcal{P}}_{\mathrm{i}})\mathbf{M}_j\mathbf{F}_N(k)$$

for $j = 1,\ldots,p$. Finally, we can describe the overall observability behavior using the proper and improper observability Gramians, which we define as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{p}} := \boldsymbol{\mathcal{Q}}_{\mathrm{p}}^C + \sum_{j=1}^p \boldsymbol{\mathcal{Q}}_{\mathrm{p}}^j, \qquad \boldsymbol{\mathcal{Q}}_{\mathrm{i}} := \boldsymbol{\mathcal{Q}}_{\mathrm{i}}^C + \sum_{j=1}^p \boldsymbol{\mathcal{Q}}_{\mathrm{i}}^j. \tag{31}$$

To describe the output energies, we first gather the output mappings defined above to a general proper and improper observability mapping

$$\mathcal{O}_{\mathrm{p}}(k,t_1,t_2) := \begin{bmatrix} \mathcal{Q}_{\mathrm{p}}^{C}(t_1) \\ \mathcal{O}_{\mathrm{pp}}^{1}(t_1,t_2) \\ \vdots \\ \mathcal{O}_{\mathrm{pp}}^{p}(t_1,t_2) \\ \mathcal{O}_{\mathrm{ip}}^{1}(k,t_2) \\ \vdots \\ \mathcal{O}_{\mathrm{ip}}^{p}(k,t_2) \end{bmatrix} \quad \text{and} \quad \mathcal{O}_{\mathrm{i}}(\ell,k,t) := \begin{bmatrix} \mathcal{Q}_{\mathrm{i}}^{C}(\ell) \\ \mathcal{O}_{\mathrm{pi}}^{1}(\ell,t) \\ \vdots \\ \mathcal{O}_{\mathrm{pi}}^{p}(\ell,t) \\ \mathcal{O}_{\mathrm{ii}}^{1}(\ell,k) \\ \vdots \\ \mathcal{O}_{\mathrm{ii}}^{p}(\ell,k) \end{bmatrix}.$$

Together with the derivations from Subsection 4.1 we obtain the proper and improper output energies

$$E(\mathcal{O}_{\mathrm{p}}) = \|\mathcal{O}_{\mathrm{p}}\|^2 = \operatorname{tr}(\mathcal{Q}_{\mathrm{p}}), \quad E(\mathcal{O}_{\mathrm{i}}) = \|\mathcal{O}_{\mathrm{i}}\|^2 = \operatorname{tr}(\mathcal{Q}_{\mathrm{i}}),$$

where $\mathcal{Q}_{\mathrm{p}}$ and $\mathcal{Q}_{\mathrm{i}}$ are as defined in (31). These energy expressions justify the truncation process as described in Subsection 4.2 also for the multiple output case.

To estimate the error $\|\mathbf{y}-\widehat{\mathbf{y}}\|_{L_\infty}$ in the case of multiple outputs, we estimate the output norm by the sum of the norms of the different components of the output, that is

$$\begin{aligned} \|\mathbf{y}-\widehat{\mathbf{y}}\|_{L_\infty} &\leq \|\mathbf{y}_C-\widehat{\mathbf{y}}_C\|_{L_\infty} + \left\| \begin{bmatrix} \mathbf{y}_1-\widehat{\mathbf{y}}_1 \\ \vdots \\ \mathbf{y}_{\mathrm{p}}-\widehat{\mathbf{y}}_{\mathrm{p}} \end{bmatrix} \right\|_{L_\infty} \\ &\leq \|\mathbf{y}_C-\widehat{\mathbf{y}}_C\|_{L_\infty} + \left\| \begin{bmatrix} \mathbf{y}_1-\widehat{\mathbf{y}}_1 \\ 0 \\ \vdots \end{bmatrix} \right\|_{L_\infty} + \cdots + \left\| \begin{bmatrix} \vdots \\ 0 \\ \mathbf{y}_{\mathrm{p}}-\widehat{\mathbf{y}}_{\mathrm{p}} \end{bmatrix} \right\|_{L_\infty} \\ &\leq \|\mathbf{y}_C-\widehat{\mathbf{y}}_C\|_{L_\infty} + \|\mathbf{y}_1-\widehat{\mathbf{y}}_1\|_{L_\infty} + \cdots + \|\mathbf{y}_{\mathrm{p}}-\widehat{\mathbf{y}}_{\mathrm{p}}\|_{L_\infty}. \end{aligned}$$

In the first component $\|\mathbf{y}_C-\widehat{\mathbf{y}}_C\|_{L_\infty}$, the improper part can be neglected since the improper states are not truncated. Hence, it holds

$$\begin{aligned} \|\mathbf{y}_C-\widehat{\mathbf{y}}_C\|_{L_\infty} &= \|\mathbf{C}\mathbf{x}_{\mathrm{p}}-\widehat{\mathbf{C}}\widehat{\mathbf{x}}_{\mathrm{p}}\|_{L_\infty} = \left(\int_0^\infty \|\operatorname{vec}(\mathbf{C}\mathbf{F}_J(t)\mathbf{B}-\widehat{\mathbf{C}}\widehat{\mathbf{F}}_J(t)\widehat{\mathbf{B}})\|_2^2 \mathrm{d}t\right)^{\frac{1}{2}} \left(\int_0^\infty \|\mathbf{u}(\tau)\|_2^2 \mathrm{d}t\right)^{\frac{1}{2}} \\ &= \operatorname{tr}\left(\mathbf{B}^{\mathrm{T}}\mathcal{Q}_{\mathrm{p}}\mathbf{B}\right)\|\mathbf{u}\|_{L_2} \end{aligned}$$

For the other summands, we apply the theory presented in Section 5 to obtain the bound

$$\begin{aligned} &\sum_{j=1}^{p} \|\mathbf{y}_j-\widehat{\mathbf{y}}_j\|_{L_\infty} \\ &\quad\leq \sum_{j=1}^{p} \|\mathbf{y}_{\mathrm{pp},j}-\widehat{\mathbf{y}}_{\mathrm{pp},j}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{ip},j}-\widehat{\mathbf{y}}_{\mathrm{ip},j}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{pi},j}-\widehat{\mathbf{y}}_{\mathrm{pi},j}\|_{L_\infty} \\ &\quad= \sum_{j=1}^{p} \left(\operatorname{tr}(\mathcal{P}_{\mathrm{p}}\mathbf{M}_j\mathcal{P}_{\mathrm{p}}\mathbf{M}_j) - 2\operatorname{tr}\left(\widetilde{\mathcal{P}}_{\mathrm{p}}^{\mathrm{T}}\mathbf{M}_j\widetilde{\mathcal{P}_{\mathrm{p}}}\widehat{\mathbf{M}}_j\right) + \operatorname{tr}\left(\widehat{\mathcal{P}}_{\mathrm{p}}\widehat{\mathbf{M}}_j\widehat{\mathcal{P}}_{\mathrm{p}}\widehat{\mathbf{M}}_j\right)\right)\|\mathbf{u}\otimes\mathbf{u}\|_{L_2} \\ &\qquad+ \left(\operatorname{tr}(\mathcal{P}_{\mathrm{p}}\mathbf{M}_j\mathcal{P}_{\mathrm{i}}\mathbf{M}_j) - 2\operatorname{tr}\left(\widetilde{\mathcal{P}}_{\mathrm{p}}^{\mathrm{T}}\mathbf{M}_j\widetilde{\mathcal{P}}_{\mathrm{i}}\widehat{\mathbf{M}}_j\right) + \operatorname{tr}\left(\widehat{\mathcal{P}}_{\mathrm{p}}\widehat{\mathbf{M}}_j\widehat{\mathcal{P}}_{\mathrm{i}}\widehat{\mathbf{M}}_j\right)\right)^{\frac{1}{2}} \nu^{\frac{1}{2}}\|\mathbf{u}\|_{\mathcal{C}^{\nu-1}}\|\mathbf{u}\|_{L_2} \\ &\qquad+ \left(\operatorname{tr}(\mathcal{P}_{\mathrm{p}}\mathbf{M}_j\mathcal{P}_{\mathrm{i}}\mathbf{M}_j) - 2\operatorname{tr}\left(\widetilde{\mathcal{P}}_{\mathrm{p}}^{\mathrm{T}}\mathbf{M}_j\widetilde{\mathcal{P}}_{\mathrm{i}}\widehat{\mathbf{M}}_j\right) + \operatorname{tr}\left(\widehat{\mathcal{P}}_{\mathrm{p}}\widehat{\mathbf{M}}_j\widehat{\mathcal{P}}_{\mathrm{i}}\widehat{\mathbf{M}}_j\right)\right)^{\frac{1}{2}} \nu^{\frac{1}{2}}\|\mathbf{u}\|_{\mathcal{C}^{\nu-1}}\|\mathbf{u}\|_{L_2}. \end{aligned}$$

7 Numerical Results

In this section, we discuss the efficiency of the proposed methodology using several examples. We also verify our theoretical findings, particularly the error bounds, in our numerical experiments. All the numerical experiments are carried out on a computer with 4 Intel Core i5-4690 CPUs running at 3.5 GHz and equipped with 8 GB total main memory. The experiments use MATLAB®R2017a and examples and methods from M-M.E.S.S.-2.1., see [21].

7.1 An illustrative example

Using this example, we highlight that for quadratic output models, it is necessary to consider mixed Gramians $\mathcal{Q}_{\mathrm{pi}}$ and $\mathcal{Q}_{\mathrm{ip}}$, as discussed in Section 3. For this, we consider the following system in Weierstraß canonical form

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \\ \dot{z}_3(t) \\ \dot{z}_4(t) \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \\ z_4(t) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \mathbf{u}(t),$$

$$\mathbf{y}(t) = \begin{bmatrix} z_1(t) & z_2(t) & z_3(t) & z_4(t) \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \\ z_4(t) \end{bmatrix}.$$

The proper state is then given by $\mathbf{x}_1(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$ and the improper one as $\mathbf{x}_2(t) = \begin{bmatrix} z_3(t) \\ z_4(t) \end{bmatrix}$. The corresponding system Gramians are given as

$$\mathcal{P}_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \mathcal{P}_2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathcal{Q}_{11} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathcal{Q}_{21} = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \mathcal{Q}_{12} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}, \quad \mathcal{Q}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}.$$

We observe that the controllability Gramian of the proper state is of rank one. Hence, the minimal realization of the proper part of the system is of rank one, and so is the proper part of the reduced-order model for this example. The improper state is described by a rank two controllability Gramian and by a rank two observability Gramian that is $\mathcal{Q}_{\mathrm{i}} = \mathcal{Q}_{\mathrm{pi}} + \mathcal{Q}_{\mathrm{ii}}$ so that the minimal realization of the improper system part is of rank two. However, we observe that the improper-improper observability Gramian is of rank one. This fact shows vividly that the mixed Gramians need to be taken into consideration.

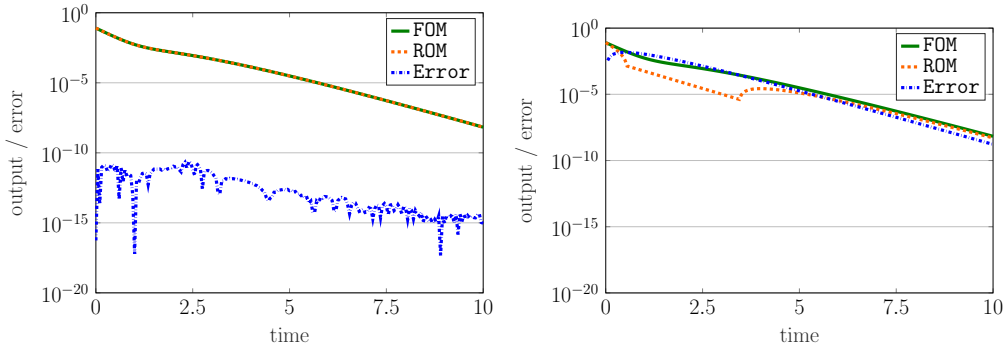
To investigate the quality of the obtained reduced-order system, we consider the system output, which we obtain by applying the input function $\mathbf{u}(t) = 0.2 \cdot e^{-t}$. The results are shown in Figure 1, where the left plot shows the results of the full-order model (`FOM`), the reduced-order model (`ROM`), and the corresponding error (`Error`) when the mixed Gramians are applied in the reduction process. The right plot shows the same values for the case when the mixed Gramians were not part of the reduction step, i.e., $\mathcal{Q}_{\mathrm{p}} := \mathcal{Q}_{\mathrm{pp}}$ and $\mathcal{Q}_{\mathrm{i}} := \mathcal{Q}_{\mathrm{ii}}$. We note that the mixed observability Gramians $\mathcal{Q}_{\mathrm{pi}}$ and $\mathcal{Q}_{\mathrm{ip}}$ must be considered within the reduction process.

7.2 Index–2 Stokes example

As the second example, we consider the creeping flow in capillaries or porous media, which can be described by the following equations

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} v(\zeta, t) &= \mu \Delta v(\zeta, t) - \nabla p(\zeta, t) + f(\zeta, t), \\ 0 &= \operatorname{div}(v(\zeta, t)), \end{aligned} \tag{32}$$

with appropriate initial and boundary conditions. The position in the domain $\Omega \subset \mathbb{R}^d$ is described by $\zeta \in \Omega$ and $t \geq 0$ is the time. For simplicity, we use a classical solution concept and assume that the external force $f : \Omega \times [0, \infty) \to \mathbb{R}^d$ is continuous and that the velocities $v : \Omega \times [0, \infty) \to \mathbb{R}^d$ and pressures $p : \Omega \times [0, \infty) \to \mathbb{R}^d$ satisfy the necessary smoothness conditions. We discretize the



(a) Mixed Gramians are used. (b) Mixed Gramians are not used.

Figure 1: An illustrative example: Output responses of the full-order and reduced-order models and the corresponding error.

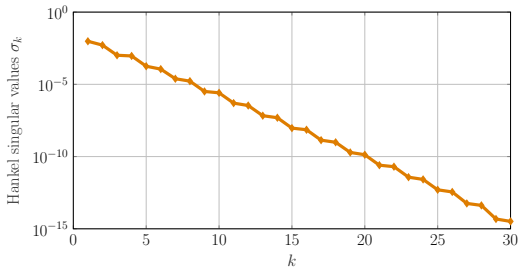


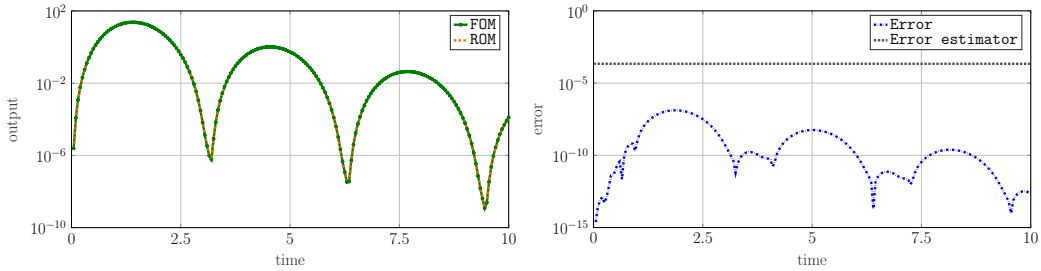
Figure 2: Index–2 Stokes example: the decay of Hankel singular values.

system (32) by a finite difference scheme as discussed in [19, 25] and have an output equation to measure our quantity of interest. We choose the matrix $\mathbf{M}$ to be $0.01 \cdot \mathbf{I}_n$, yielding the l_2-norm of the state vector with a scaling factor 0.01. Consequently, we obtain a discretized system of the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} z(t) \\ \lambda(t) \end{bmatrix} &= \begin{bmatrix} A & G \\ G^{\mathrm{T}} & 0 \end{bmatrix}\begin{bmatrix} z(t) \\ \lambda(t) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}\mathbf{u}(t), \\ \mathbf{y}(t) &= \begin{bmatrix} z(t)^{\mathrm{T}} & \lambda(t)^{\mathrm{T}} \end{bmatrix}\mathbf{M}\begin{bmatrix} z(t) \\ \lambda(t) \end{bmatrix} \end{aligned} \tag{33}$$

with system matrices $A \in \mathbb{R}^{g\times g}$ and $G \in \mathbb{R}^{g,q}$. The input matrices are given as $B_1 \in \mathbb{R}^{g,m}$, $B_2 \in \mathbb{R}^{q,m}$ and the output matrix is $\mathbf{M} \in \mathbb{R}^{n,n}$ with $n = g + q$. The state consists of $z(t) \in \mathbb{R}^g$ and $\lambda(t) \in \mathbb{R}^q$, while the input is $\mathbf{u}(t) \in \mathbb{R}^m$ and the output $\mathbf{y}(t) \in \mathbb{R}$. We consider the system of dimension $n = 645 = n_v + n_p$, where the dimensions of the velocity and pressure vectors are $n_v = 420$ and $n_p = 225$, respectively.

We need to determine the Gramians corresponding to the proper and improper states of the system (33). For this purpose, we apply the methods described in [25, 26], noting that the improper Gramians can be computed explicitly. In Figure 2, we depict the decay of the Hankel singular values $\sigma_1, \sigma_2, \ldots$ of the proper states corresponding to the proper Gramians as described in Section 4.2. We truncate the proper Hankel singular values smaller than $\sigma_1 \cdot 10^{-8}$ and truncate the improper Hankel singular values equal to zero. The reduced-order model has the dimensions $\widehat{n} = \widehat{n}_v + \widehat{n}_p$ with $n_v = 13$ and $\widehat{n}_p = 2$. Figure 3 shows the output behavior of the full-order model (1) and of the reduced-order model (2) for an input function $\mathbf{u}(t) = \sin(t)^3 e^{-t/2}$. Additionally, the figure includes the output error and the corresponding error estimation. The actual error is below the estimated error for all time, and we observe that the error bound is rather conservative. The correct error is sufficiently small, and the approximation quality of the reduced-order systems is much better than the estimated one.



(a) Output (b) Error

Figure 3: Index–2 Stokes example: the outputs of the full-order model and the obtained reduced-order model of order 15 for a test input $\mathbf{u}(t) = \sin(t)^3 e^{-t/2}$. The plot also shows the error between the outputs and our error estimate.

7.3 Index–3 mechanical system

Now, we investigate a system of index–3, that results from mechanical systems and is of the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} I_{n_x} & 0 & 0 \\ 0 & H & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \\ \lambda(t) \end{bmatrix} &= \begin{bmatrix} 0 & I_{n_x} & 0 \\ -K & -D & G \\ G^{\mathrm{T}} & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \\ \lambda(t) \end{bmatrix} + \begin{bmatrix} 0 \\ B_x \\ 0 \end{bmatrix}\mathbf{u}(t), \\ \mathbf{y}(t) &= \begin{bmatrix} x_1(t)^{\mathrm{T}} & x_2(t)^{\mathrm{T}} & \lambda(t)^{\mathrm{T}} \end{bmatrix}\mathbf{M}\begin{bmatrix} x_1(t) \\ x_2(t) \\ \lambda(t) \end{bmatrix}, \end{aligned} \tag{34}$$

where $H,\ D,\ K \in \mathbb{R}^{g\times g}$, $B_x \in \mathbb{R}^{g\times m}$, $G \in \mathbb{R}^{g\times q}$ and $\mathbf{M} \in \mathbb{R}^{2g+\ell\times 2n+q}$. The state is given by $x_1(t),\ x_2(t) \in \mathbb{R}^g,\ \lambda(t) \in \mathbb{R}^g$, the input by $\mathbf{u}(t) \in \mathbb{R}^m$ and the output by $\mathbf{y}(t) \in \mathbb{R}$. We consider the index–3 system (34), which arises in the modeling of constraint mechanical systems with matrices

$$H = \operatorname{diag}(m_1, \ldots, m_g),$$

$$D = \begin{bmatrix} d_1+\delta_1 & -d_1 & & & \\ -d_1 & d_1+d_2+\delta_2 & -d_2 & & \\ & \ddots & \ddots & \ddots & \\ & -d_{g-2} & d_{g-2}+d_{g-1}+\delta_{g-1} & -d_{g-1} & \\ & & -d_{g-1} & d_{g-1}+\delta_g \end{bmatrix},$$

$$K = \begin{bmatrix} k_1+\kappa_1 & -k_1 & & & \\ -k_1 & k_1+k_2+\kappa_2 & -k_2 & & \\ & \ddots & \ddots & \ddots & \\ & -k_{g-2} & k_{g-2}+k_{g-1}+\kappa_{g-1} & -k_{g-1} & \\ & & -k_{g-1} & k_{g-1}+\kappa_g \end{bmatrix},$$

$$G = [1, 0, \ldots, 0, -1]^{\mathrm{T}}, \quad B_x = [1, 0, \ldots, 0]^{\mathrm{T}}, \quad \mathbf{M} = \mathbf{I}_{2g+1}.$$

The matrices are generated using the M-M.E.S.S.-function `msd_ind3`, see [21], with dimension $g = 600$. We choose

$$\begin{aligned} m_1 = \cdots = m_g = 1, \quad k_1 = \cdots = k_{g-1} = 1.5, \quad d_1 = \cdots = d_{g-1} = 0.7, \\ \kappa_1 = \cdots = \kappa_g = 2, \quad \delta_1 = \cdots = \delta_g = 0.9. \end{aligned}$$

In [19], the projection matrices (3) for this example were introduced. To compute the Gramians, we follow the same procedure presented in [25, 26] modified to the index 3 cases.

Figure 4 depicts the proper Hankel singular values. We truncate those smaller than $\sigma_1 \cdot 10^{-8}$. Additionally, we remove the improper states corresponding to improper Hankel singular values

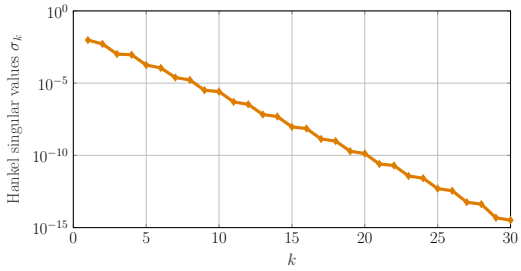
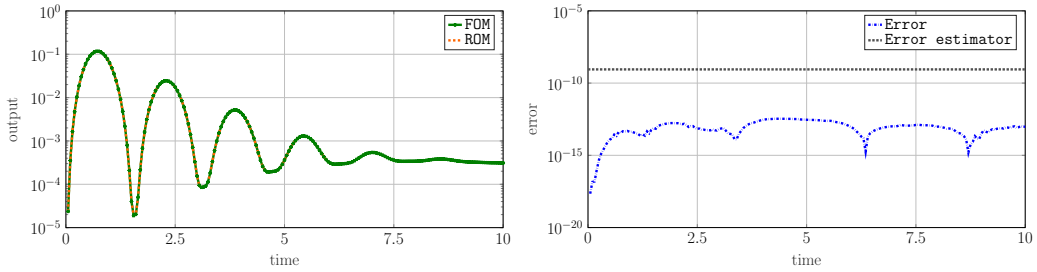


Figure 4: Index–3 Mechanical system example: the decay of Hankel singular values.



(a) Output (b) Error

Figure 5: Index–3 Mechanical system example: the outputs of the full-order model and the obtained reduced-order model of order 21 for the input $\mathbf{u}(t) = \sin(2t)^2 e^{-t/2}$. The plot also shows the error between the outputs and our error estimate.

that are zero. The resulting reduced dimensions are $\widehat{n} = \widehat{n}_v + \widehat{n}_p$ with $n_v = 20$ and $\widehat{n}_p = 1$. The outputs of the full-order and the reduced-order models (1) and (2) are described in Figure 5 for an input function $\mathbf{u}(t) = \sin(2t)^2 e^{-t/2}$ and the figure also shows the error between the outputs and the error estimate using 29.

Here again, we make similar observations as in the previous example–that is, the error bound is conservative, and the system is approximated much better, leading to an output error that is smaller than 10^{-13} for all $t \in [0, 10]$.

Conclusions

This paper has addressed the model reduction problem for differential-algebraic systems with quadratic output equations using BT. To apply the balancing procedure, we have derived new observability Gramians that describe the observability behavior of the proper and improper states. We have shown that these Gramians can be computed by solving certain projected Lyapunov equations. Moreover, we have derived some bounds of the energy functional for the proper states that have been used to derive a truncation criterion. To evaluate the quality of the reduced surrogate model, we have introduced an error estimator.

Our method has been illustrated by numerical examples of indexes one, two, and three. In particular, we were able to derive surrogate models of very small dimensions that approximate the input–to–output behavior of the full-order models very well.

References

[1] M. I. Ahmad and P. Benner. Interpolatory model reduction techniques for linear second-order descriptor systems. In *Proc. European Control Conf. ECC 2014, Strasbourg*, pages 1075–1079. IEEE, 2014. doi:10.1109/ECC.2014.6862210.

[2] M. I. Ahmad, P. Benner, and P. Goyal. Krylov subspace-based model reduction for

a class of bilinear descriptor systems. *J. Comput. Appl. Math.*, 315:303–318, 2017. doi:10.1016/j.cam.2016.11.009.

[3] M. I. Ahmad, P. Benner, P. Goyal, and J. Heiland. Moment-matching based model reduction for Navier-Stokes type quadratic-bilinear descriptor systems. *Z. Angew. Math. Mech.*, 97(10):1252–1267, 2017. doi:10.1002/zamm.201500262.

[4] A. C. Antoulas, C. A. Beattie, and S. Gugercin. *Interpolatory Methods for Model Reduction*. Computational Science & Engineering. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2020. doi:10.1137/1.9781611976083.

[5] P. Benner, P. Goyal, and I. Pontes Duff. Gramians, energy functionals and balanced truncation for linear dynamical systems with quadratic outputs. *IEEE Trans. Autom. Control*, 2021. doi:10.1109/TAC.2021.3086319.

[6] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. H. A. Schilders, and L. M. Silveira, editors. *Model Order Reduction. Volume 1: System- and Data-Driven Methods and Algorithms*. De Gruyter, Berlin, 2021. URL: https://www.degruyter.com/view/title/523453.

[7] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. H. A. Schilders, and L. M. Silveira, editors. *Model Order Reduction. Volume 2: Snapshot-Based Methods and Algorithms*. De Gruyter, Berlin, 2021. doi:10.1515/9783110671490.

[8] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, editors. *Model Reduction and Approximation: Theory and Algorithms*. Computational Science & Engineering. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017. doi:10.1137/1.9781611974829.

[9] P. Benner, E. S. Quintana-Ortí, and G. Quintana-Ortí. Parallel model reduction of large-scale linear descriptor systems via balanced truncation. In M. Daydé, J. J. Dongarra, V. Hernández, and J. M. L. M. Palma, editors, *High Performance Computing for Computational Science - VECPAR 2004*, volume 3402 of *Lecture Notes in Comput. Sci.*, pages 340–353, Berlin/Heidelberg, Germany, 2005. Springer-Verlag. doi:10.1007/11403937_27.

[10] P. Benner, J. Saak, and M. M. Uddin. Balancing based model reduction for structured index-2 unstable descriptor systems with application to flow control. *Numer. Algebra Control Optim.*, 6(1):1–20, March 2016. doi:10.3934/naco.2016.6.1.

[11] P. Benner and T. Stykel. Model order reduction for differential-algebraic equations: A survey. In Achim Ilchmann and Timo Reis, editors, *Surveys in Differential-Algebraic Equations IV*, Differential-Algebraic Equations Forum, pages 107–160. Springer International Publishing, Cham, March 2017. doi:10.1007/978-3-319-46618-7_3.

[12] G. Flagg, C. Beattie, and S. Gugercin. Convergence of the iterative rational Krylov algorithm. *Systems Control Lett.*, 61(6):688–691, 2012. doi:10.1016/j.sysconle.2012.03.005.

[13] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ norms. *Internat. J. Control*, 39:1115–1193, 1984.

[14] S. Gugercin, A. C. Antoulas, and C. Beattie. $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM J. Matrix Anal. Appl.*, 30(2):609–638, 2008. doi:10.1137/060666123.

[15] S. Gugercin, T. Stykel, and S. Wyatt. Model reduction of descriptor systems by interpolatory projection methods. *SIAM J. Sci. Comput.*, 35(5):B1010–B1033, 2013. doi:10.1137/130906635.

[16] M. Heinkenschloss, D. C. Sorensen, and K. Sun. Balanced truncation model reduction for a class of descriptor systems with application to the Oseen equations. *SIAM J. Sci. Comput.*, 30(2):1038–1063, 2008. doi:10.1137/070681910.

[17] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations: Analysis and Numerical Solution*. Textbooks in Mathematics. EMS Publishing House, Zürich, Switzerland, 2006.

[18] Y. Lin, L. Bao, and Y. Wei. A model-order reduction method based on Krylov subspace for MIMO bilinear dynamical systems. *J. Appl. Math. Comput.*, 25(1-2):293–304, 2007.

[19] V. Mehrmann and T. Stykel. Balanced truncation model reduction for large-scale systems in descriptor form. In P. Benner, V. Mehrmann, and D. C. Sorensen, editors, *Dimension Reduction of Large-Scale Systems*, volume 45 of *Lect. Notes Comput. Sci. Eng.*, pages 83–115. Springer-Verlag, Berlin/Heidelberg, Germany, 2005. doi:10.1007/3-540-27909-1_3.

[20] B. C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Autom. Control*, AC–26(1):17–32, 1981. doi:10.1109/TAC.1981.1102568.

[21] J. Saak, M. Köhler, and P. Benner. M-M.E.S.S.-2.1 – the Matrix Equations Sparse Solvers library, April 2021. see also:https://www.mpi-magdeburg.mpg.de/projects/mess. doi:10.5281/zenodo.4719688.

[22] J. Saak and M. Voigt. Model reduction of constrained mechanical systems in M-M.E.S.S. *IFAC-PapersOnLine 9th Vienna International Conference on Mathematical Modelling MATHMOD 2018, Vienna, Austria, 21–23 February 2018*, 51(2):661–666, 2018. doi:10.1016/j.ifacol.2018.03.112.

[23] T. Stykel. *Analysis and Numerical Solution of Generalized Lyapunov Equations*. Dissertation, TU Berlin, 2002. URL: http://webdoc.sub.gwdg.de/ebook/e/2003/tu-berlin/stykel_tatjana.pdf.

[24] T. Stykel. Gramian-based model reduction for descriptor systems. *Math. Control Signals Systems*, 16(4):297–319, 2004. doi:10.1007/s00498-004-0141-4.

[25] T. Stykel. Balanced truncation model reduction for semidiscretized Stokes equation. *Linear Algebra Appl.*, 415(2–3):262–289, 2006. doi:10.1016/j.laa.2004.01.015.

[26] T. Stykel. Low-rank iterative methods for projected generalized Lyapunov equations. *Electron. Trans. Numer. Anal.*, 30:187–202, 2008. URL: http://etna.mcs.kent.edu/vol.30.2008/pp187-202.dir/pp187-202.pdf.

[27] T. Stykel and V. Simoncini. Krylov subspace methods for projected Lyapunov equations. *Appl. Numer. Math.*, 62(1):35–50, January 2012. doi:10.1016/j.apnum.2011.09.007.

[28] M. S. Tombs and I. Postlethwaite. Truncated balanced realization of a stable non-minimal state-space system. *Internat. J. Control*, 46(4):1319–1330, 1987.