Aspects of damping optimization in vibrational systems using model order reduction

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von **M. Sc. Jennifer Przybilla**

geb. am **12.04.1995** in Berlin

genehmigt durch die Fakultät für Mathematik

der Otto-von-Guericke-Universität Magdeburg

Gutachter: **Prof. Dr. Peter Benner**

Prof. Dr. Tatjana Stykel

Dr. Igor Pontes Duff

eingereicht am: **23.02.2024**

Verteidigung am: **05.07.2024**

PUBLICATIONS

[105] J. Przybilla, I. Pontes Duff, and P. Benner, *Semi-active damping optimization of vibrational systems using the reduced basis method*, Adv. Comput. Math., 50 (2024).

[106] J. Przybilla, I. Pontes Duff, P. Goyal, and P. Benner, *Balanced Truncation of Descriptor Systems with a Quadratic Output*, e-print 2402.14716, arXiv, 2024. math.DS.

[107] J. Przybilla, I. Pontes Duff, and P. Benner, *Model reduction for second-order systems with inhomogeneous initial conditions*, Systems Control Lett., 183 (2024).

ABSTRACT

When constructing infrastructures like buildings or bridges, we need to consider the influence of external forces such as wind perturbations, moving pedestrians, or even earthquakes. These forces can cause vibrations or damage within the structure. With advancements in engineering making structures lighter and more refined, they have become more susceptible to large deflections and fatigue when external forces, especially those close to the structure's natural eigenfrequencies, come into play. To prevent these effects, we include external dampers in the system structure. In this thesis, we aim to find the best way to position and adjust these dampers so they can absorb the most critical forces.

We use models to describe the respective constructions and to compute the system responses, such as changes in the system behavior with applied external damping. However, when constructions are described in detail, the respective models are of high dimension. Therefore, evaluating the system's behavior or optimizing dampers within them becomes numerically very demanding. Hence, we derive and apply different reduction methods, depending on the problem settings, generating reduced surrogate models that are evaluated instead of the full-order model.

In this work, we consider two problem classes: The first one considers inhomogeneous systems with a given external damper, which require suitable reduction methods. The second challenge is to optimize the external dampers and the respective parameters within a vibrational system, where we also need to derive reduction techniques tailored to parameter-dependent systems.

When considering vibrational systems with a given external damper, inhomogeneous initial conditions appear that further define the respective displacements and velocities. Furthermore, linear and quadratic output equations are of interest, while the state equation can have a first- or second-order structure. Moreover, the state equation can include physical constraints, which lead to differential–algebraic equations. Most of these system structures are non-standard forms that have not been discussed in the literature, yet they are relevant. Hence, in this work, we derive algorithms and respective error bounds that determine surrogate models for large-scale systems in a non-standard form. To approach the problem of reducing systems with inhomogeneous initial conditions while considering linear and quadratic output equations, we use the superposition principle, which allows us to decompose the system behavior into independent components. The first component corresponds to the transfer between the input and output having zero initial conditions. In contrast, the others correspond to the system behavior resulting from the initial conditions. Based on this superposition of the system, we propose model reduction schemes that preserve the structure in the surrogate models. To this aim, we introduce tailored Gramians for the different system structures that incorporate the controllability and observability properties of each system component. We propose two resulting methodologies. The first one consists of reducing each of the components independently using a suitable balanced truncation procedure, which allows flexibility in the order of the reduced-order models. The sum of these reduced systems provides an approximation of the original system. The second proposed methodology consists in extracting the dominant subspaces from the sum of Gramians to construct one surrogate model. Additionally, we discuss error bounds for the overall output approximation and illustrate the proposed methods using benchmark problems.

In addition, this thesis investigates the problem of optimizing dampers in vibrational systems. The aim is to determine the positions and viscosities of external dampers in such a way that the influence of the input on the output is minimized. We use the energy response as an optimization criterion, whose calculation involves solving Lyapunov equations. Hence, the optimization of external dampers can be computationally demanding. Therefore, we derive reduction techniques suitable for parameter-dependent systems that determine surrogate models of significantly smaller dimensions. We describe reduced basis methods that approximate the solution space of the Lyapunov equations, coinciding with the controllability space of the system, for all possible external dampers. To improve these methods, we also decouple the solution spaces of the problem to obtain a space that corresponds to the system without external dampers and serves as a starting point for the reduction of the optimization problem. Furthermore, we derive spaces that correspond to the different damper positions and that are used to extend the reduced basis if necessary. This decomposition additionally provides an error estimator that evaluates the approximation to the controllability space. Moreover, we derive an adaptive scheme that generates the reduced solution space by adding the subspaces of interest during the optimization process, resulting in the corresponding reduced optimization problem. Our new technique leads to a reduced optimization problem with a significantly smaller dimension, which is fast solvable, especially compared to the original system, which we illustrate with numerical examples.

ZUSAMMENFASSUNG

Beim Bau von Infrastruktur wie Gebäuden oder Brücken müssen wir den Einfluss äußerer Kräfte durch Fußgänger, Windereignisse oder sogar Erdbeben berücksichtigen. Diese Kräfte können Vibrationen oder Schäden innerhalb der Struktur verursachen. Durch die Fortschritte in der Technik sind die Bauwerke leichter geworden, aber sie sind auch anfälliger für starke Auslenkungen und Ermüdung der Strukturen, wenn äußere Kräfte wirken, insbesondere, wenn diese nahe an den natürlichen Eigenfrequenzen des Bauwerks liegen. Um diese Effekte zu verhindern, werden externe Dämpfer in die Systemstruktur eingebaut. In dieser Arbeit wollen wir die Positionen und die Stärke dieser Dämpfer optimieren, damit sie die kritischsten Kräfte abdämpfen können.

Wir verwenden Modelle, um die jeweiligen Konstruktionen zu beschreiben und das Systemverhalten und die Änderungen des Systemverhaltens bei angewandter externer Dämpfung zu berechnen. Wenn die Konstruktionen jedoch detailliert beschrieben werden, haben die entsprechenden Modelle sehr große Dimensionen. Daher wird die Auswertung des Systemverhaltens oder die Optimierung von Dämpfern in diesen Modellen numerisch sehr anspruchsvoll. Aus diesem Grund leiten wir verschiedene Reduktionsmethoden her und wenden sie je nach Problemstellung an, um reduzierte Ersatzmodelle zu erzeugen, die anstelle des ursprünglichen Modells ausgewertet werden.

Wir betrachten in dieser Arbeit zwei Problemklassen: Die erste betrachtet inhomogene Systeme mit gegebenen externen Dämpfern. Da diese Systeme große Dimensionen haben, leiten wir entsprechende Reduktionsverfahren her. Die zweite Problematik besteht darin, die externen Dämpfer und die entsprechenden Parameter innerhalb eines schwingenden Systems zu optimieren. Auch hier müssen wir Reduktionsverfahren anwenden, die auf parameterabhängige Systeme mit einer bestimmten Struktur zugeschnitten sind.

Bei der Betrachtung von schwingenden Systemen mit gegebenen externen Dämpfern spielen auch inhomogene Anfangsbedingungen eine Rolle, da sie die Verschiebungen und Geschwindigkeiten beeinflussen. Außerdem sind lineare und quadratische Ausgangsgleichungen von Interesse, während die Zustandsgleichung eine Differentialgleichung erster oder zweiter Ordnung sein kann. Darüber hinaus kann die Zustandsgleichung physikalische Bedingungen enthalten, die zu differential-algebraischen Gleichungen führen. All diese Systemstrukturen führen zu mehreren Systemtypen, die nicht in Standardform sind und in der Literatur kaum berücksichtigt wurden, aber von großer Bedeutung sind. Daher leiten wir in dieser Arbeit Algorithmen her, welche für große Systeme in Nicht-Standardform reduzierte Modelle bestimmen, die das Systemverhalten approximieren. Um mit inhomogenen Anfangsbedingungen umzugehen und gleichzeitig lineare und quadratische Ausgangsgleichungen zu berücksichtigen, verwenden wir das Superpositionsprinzip. Dies ermöglicht es uns, das Systemverhalten in unabhängige Komponenten zu zerlegen. Das erste System entspricht der Übertragung zwischen dem Eingang und dem Ausgang bei homogenen Ausgangsbedingungen. Die restlichen Komponenten entsprechen dem Systemverhalten unter Berücksichtigung der Anfangsbedingungen. Auf der Grundlage dieser Überlagerung von Systemen ist es unser Ziel, Modellreduktionsverfahren herzuleiten, welche die relevanten Strukturen erhalten. Dafür führen wir maßgeschneiderte Matrizen, sogenannte Gramschen, für jede Systemkomponente ein und berechnen diese numerisch, indem wir die Lyapunov Gleichungen lösen. Daraus resultieren zwei Methoden. Die erste besteht darin, jede der Komponenten unabhängig voneinander durch ein geeignetes balanciertes Trunkierungsverfahren zu reduzieren, was Flexibilität bei den Dimensionen der reduzierten Modelle ermöglicht. Die Summe dieser reduzierten Systeme liefert eine Annäherung an das ursprüngliche System. Die zweite vorgeschlagene Methode besteht darin, die dominanten Unterräume aus der Summe der Gramschen zu extrahieren, um die Projektionsmatrizen zu erstellen, die zu einem Ersatzmodell führen. Darüber hinaus werden Fehlerschranken für die Approximation der Ausgänge diskutiert. Schließlich werden die vorgeschlagenen Methoden anhand von Benchmark-Problemen illustriert.

Des Weiteren wird in dieser Arbeit das Problem der Optimierung von Dämpfern in schwingungsfähigen Systemen untersucht. Ziel ist es, die Positionen und Viskositäten von externen Dämpfern so zu bestimmen, dass der Einfluss des Eingangs auf den Ausgang minimiert wird. Als Optimierungskriterium verwenden wir die Energieantwort. Um die optimalen externen Dämpfer zu finden, müssen viele dieser Gleichungen gelöst werden. Daher kann der Minimierungsprozess sehr rechenaufwendig sein. Aus diesem Grund leiten wir Reduktionsverfahren her, um dieses Problem zu lösen. Um den Prozess der Suche nach den optimalen Dämpfern zu beschleunigen, schlagen wir reduzierte-Basen-Methoden vor. Unsere Algorithmen erzeugen eine Basis, die den Lösungsraum der Lyapunov Gleichungen, der mit dem Steuerbarkeitsraum des Systems übereinstimmt, für alle möglichen Positionen der Dämpfer approximiert. Wir entkoppeln die Lösungsräume des Problems, um einen Raum zu erhalten, der dem System ohne externe Dämpfer entspricht und als Ausgangspunkt für die Reduktion des Optimierungsproblems dient. Darüber hinaus leiten wir Räume her, die den verschiedenen Dämpferpositionen entsprechen und bei Bedarf zur Erweiterung der reduzierten Basis verwendet werden. Diese Zerlegung liefert zusätzlich einen Fehlerschätzer, der die Approximation des Steuerbarkeitsraums bewertet. Darüber hinaus leiten wir ein adaptives Schema her, das den reduzierten Lösungsraum durch Hinzufügen der relevanten Unterräume während des Optimierungsprozesses erzeugt, was zu dem entsprechenden reduzierten Optimierungsproblem führt. Unsere neuen Methoden führen zu reduzierten Optimierungsproblemen mit einer deutlich geringeren Dimension, das schneller zu lösen ist als das ursprüngliche Problem, was wir anhand numerischer Beispiele veranschaulichen.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

LIST OF SYMBOLS

$\mathbb{R}$, $\mathbb{C}$	fields of real and complex numbers
$\mathbb{C}_+$, $\mathbb{C}_-$	open right/open left complex half plane
$\mathbb{R}^n$, $\mathbb{C}^n$	vector space of real/complex n-tuples
$\mathbb{R}^{m\times n}$, $\mathbb{C}^{m\times n}$	real/complex $m \times n$ matrices
$\lvert\xi\rvert$	absolute value of real or complex scalar
i	imaginary unit ($\mathrm{i}^2 = -1$)
$\mathrm{Re}(\mathbf{A})$, $\mathrm{Im}(\mathbf{A})$	real and imaginary part of a complex quantity $\mathbf{A} = \mathrm{Re}(\mathbf{A}) + \mathrm{i}\,\mathrm{Im}(\mathbf{A}) \in \mathbb{C}^{n\times m}$
a_{ij}	the (i,j)-th entry of $\mathbf{A}$
a_i	the i-th column of $\mathbf{A}$
$\mathbf{A}(i:j,:)$, $\mathbf{A}(:,k:\ell)$	rows $i,\dots,j$ of $\mathbf{A}$ and columns $k,\dots,\ell$ of $\mathbf{A}$
$\mathbf{A}(i:j,k:\ell)$	rows $i,\dots,j$ of columns $k,\dots,\ell$ of $\mathbf{A}$
$\mathbf{A}^{\mathrm{T}}$	the transpose of $\mathbf{A}$
$\mathbf{A}^{\mathrm{H}}$	$:= (\overline{\mathbf{A}})^{\mathrm{T}}$, the complex conjugate transpose
$\mathbf{A}^{-1}$	the inverse of nonsingular $\mathbf{A}$
$\mathbf{A}^{-\mathrm{T}}$, $\mathbf{A}^{-\mathrm{H}}$	the inverse of $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{\mathrm{H}}$
$\mathbf{I}_n$	identity matrix of dimension n
$\mathbf{1}_{n\times 1}$	n dimensional vector containing one values
e_j	j-th unit vector, i.e., the j-th column of $\mathbf{I}_n$
$(\mathbf{A},\mathbf{E})$	matrix pencil $\lambda\mathbf{E} - \mathbf{A}$
$\Lambda(\mathbf{A})$, $\Lambda(\mathbf{A},\mathbf{E})$	spectrum of matrix $\mathbf{A}$ or the matrix pencil $\lambda\mathbf{E} - \mathbf{A}$
$\lambda_j(\mathbf{A})$	j-th eigenvalue of $\mathbf{A}$
$\lambda_{\min}(\mathbf{A})$, $\lambda_{\max}(\mathbf{A})$	the smallest, largest eigenvalue of $\mathbf{A}$

$\rho(\mathbf{A})$	$:= \max_j \lvert\lambda_j(\mathbf{A})\rvert$, spectral radius of $\mathbf{A}$
$\sigma_j(\mathbf{A})$	j-th singular value of $\mathbf{A}$
$\sigma_{\min}(\mathbf{A})$, $\sigma_{\max}(\mathbf{A})$	the smallest, largest singular value of $\mathbf{A}$
$\mathbf{A} \otimes \mathbf{B}$	the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$
$\mathrm{vec}(\mathbf{A})$	vectorization operator applied to matrix $\mathbf{A}$
$\mathrm{orth}(\mathbf{A})$	orthonormalization of the columns of a matrix $\mathbf{A}$
$\mathrm{tr}(\mathbf{A})$	$:= \sum_{i=1}^{n} a_{ii}$, trace of $\mathbf{A}$
$\Vert\mathbf{u}\Vert_2$, $\Vert\mathbf{A}\Vert_2$	Euclidean vector or subordinate matrix norm $\Vert\cdot\Vert_2$
$\Vert\mathbf{A}\Vert_{\mathrm{F}}$	$:= \sqrt{\sum_{i,j} \lvert a_{ij}\rvert^2} = \sqrt{\mathrm{tr}(\mathbf{A}^{\mathrm{H}}\mathbf{A})}$, the Frobenius norm of matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$
$\Vert\mathbf{u}\Vert_\infty$	$:= \max_i \lvert u_i\rvert$, the maximum norm of $\mathbf{u} = [u_1, \ldots, u_n]^{\mathrm{T}} \in \mathbb{C}^n$
$\Vert\mathbf{u}\Vert_p$	$:= (\sum_{i=1}^{n} \lvert u_i\rvert^{p})^{1/p}$ for $\mathbf{u} = [u_1, \ldots, u_n]^{\mathrm{T}} \in \mathbb{C}^n$ and $1 \le p < \infty$
$L_2([0,\infty), \mathbb{R}^m)$	$:= \{\mathbf{u} : [0,\infty) \to \mathbb{R}^m \mid \int_0^\infty \Vert\mathbf{u}(t)\Vert_2^2 \mathrm{d}t < \infty\}$, L_2-space
$\Vert\mathbf{u}\Vert_{L_2}$	$:= \sqrt{\int_0^\infty \Vert\mathbf{u}(t)\Vert_2^2 \mathrm{d}t}$, L_2-norm of $\mathbf{u} \in L_2([0,\infty), \mathbb{R}^m)$
$L_\infty([0,\infty), \mathbb{R}^m)$	$:= \{\mathbf{u} : [0,\infty) \to \mathbb{R}^m \mid \sup_{t\ge 0} \Vert\mathbf{u}(t)\Vert_2 < \infty\}$, L_∞-space
$\Vert\mathbf{u}\Vert_{L_\infty}$	$:= \sup_{t\ge 0} \Vert\mathbf{u}(t)\Vert_2$, L_∞-norm of $\mathbf{u} \in L_\infty([0,\infty), \mathbb{R}^m)$
$\mathcal{H}_2([0,\infty), \mathbb{R}^m)$	$:= \{\mathcal{G} : \mathbb{C}^+ \to \mathbb{C}^{m\times p} \mid \int_{-\infty}^{\infty} \Vert\mathcal{G}(\mathrm{i}\omega)\Vert_{\mathrm{F}}^2 \mathrm{d}\omega < \infty\}$, $\mathcal{H}_2$-space
$\Vert\mathcal{G}\Vert_{\mathcal{H}_2}$	$:= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \Vert\mathcal{G}(\mathrm{i}\omega)\Vert_{\mathrm{F}}^2 \mathrm{d}t}$, $\mathcal{H}_2$-norm of $\mathcal{G} \in \mathcal{H}_2([0,\infty), \mathbb{R}^m)$
$\mathcal{H}_\infty([0,\infty), \mathbb{R}^m)$	$:= \{\mathcal{G} : \mathbb{C}^+ \to \mathbb{C}^{m\times p} \mid \sup_{\omega\in\mathbb{R}} \Vert\mathcal{G}(\mathrm{i}\omega)\Vert_2 < \infty\}$, $\mathcal{H}_\infty$-space
$\Vert\mathcal{G}\Vert_{\mathcal{H}_\infty}$	$:= \sup_{\omega\in\mathbb{R}} \Vert\mathcal{G}(\mathrm{i}\omega)\Vert_2$, $\mathcal{H}_\infty$-norm of $\mathcal{G} \in \mathcal{H}_\infty([0,\infty), \mathbb{R}^m)$
$\mathcal{C}^{\nu-1}([0,\infty), \mathbb{R}^m)$	$:= \{\mathbf{u} : [0,\infty) \to \mathbb{R}^m \mid \mathbf{u}$ is $(\nu-1)$-times continuously differentiable $\}$, $\mathcal{C}^{\nu-1}$-space
$\Vert\mathbf{u}\Vert_{\mathcal{C}^{\nu-1}}$	$:= \max_{k=0,\ldots,\nu-1} \sup_{t\ge 0} \Vert\mathbf{u}^{(k)}(t)\Vert_2$, $\mathcal{C}^{\nu-1}$-norm of $\mathbf{u} \in C^{\nu-1}([0,\infty), \mathbb{R}^m) \cup L_2([0,\infty), \mathbb{R}^m)$, $\mathcal{C}^{\nu-1}$-norm

LIST OF ABBREVIATIONS

ODE	ordinary differential equation
DAE	differential–algebraic equation
BT	balanced truncation
IRKA	iterative rational Krylov algorithm
sym2IRKA	symmetric second-order iterative rational Krylov algorithm
SVD	singular value decomposition
ADI	alternating-direction implicit method
WCF	Weierstraß-canonical form
C-stable	completely stable
C-controllable	completely controllable
C-observable	completely observable
SISO	single-input single-output
MIMO	multi-input multi-output
RBM	reduced basis method
EE-RBM	error equation reduced basis method

CHAPTER 1 INTRODUCTION

Contents

1.1 Motivation

When constructing large civil engineering infrastructure such as buildings or bridges, external vibrational forces like wind perturbations or earthquakes need to be taken into account. These disturbances can cause vibrations, deflection, or even damage in the construction, which can be prevented by adding external dampers. Due to the continuous improvement in engineering construction, which provides for lighter and finer structures, corresponding infrastructures have become more susceptible to large deflections and fatigue when external forces, with dominant frequencies close to the eigenfrequencies of the construction, are applied. We eliminate this effect by designing dampers to remove critical forces from the physical system. In this thesis, we investigate the problem of optimizing external dampers in vibrational systems. The objective is to determine the viscosities and positions of external dampers in such a way that the influence of the input on the output is minimized using the energy response as an optimization criterion.

To model these infrastructures, we consider vibrational systems of the form

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}(c,g)\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0,$$

where $\mathbf{M},\ \mathbf{D}(c,g),\ \mathbf{K} \in \mathbb{R}^{n\times n}$ are the mass matrix, the damping matrix, and the stiffness matrix, respectively, for parameters $(c,g) \in \mathfrak{D}$, where $\mathfrak{D}$ is a parameter set. The vectors

$\mathbf{u}(t) \in \mathbb{R}^m$ and $\mathbf{x}(t) \in \mathbb{R}^n$ represent the input and state of the system, respectively, and $\mathbf{x}_0$, $\dot{\mathbf{x}}_0 \in \mathbb{R}^n$ are the position and velocity initial conditions. Naturally, the matrices $\mathbf{M}$, $\mathbf{D}(c, g)$, and $\mathbf{K}$ are symmetric and positive semidefinite. Because of their structure, these systems are asymptotically stable, i.e., all eigenvalues $\lambda(c, g)$ of the polynomial eigenvalue problem $(\lambda(c, g)^2\mathbf{M} + \lambda(c, g)\mathbf{D}(c, g) + \mathbf{K})\mathbf{x}(c, g) = 0$ have a negative real part. Moreover, the mass matrix $\mathbf{M}$ can be singular. In this case, we consider differential–algebraic equations (DAEs) as state equations.

The damping matrix $\mathbf{D}(c, g)$ consists of two parts, a parameter-independent internal damping $\mathbf{D}_{\text{int}}$ and a parameter-dependent external damping $\mathbf{D}_{\text{ext}}(c, g)$, i.e.,

$$\mathbf{D}(c, g) = \mathbf{D}_{\text{int}} + \mathbf{D}_{\text{ext}}(c, g). \tag{1.1}$$

There are several different models for internal damping. In this work, we use a small multiple of the critical damping defined as

$$\mathbf{D}_{\text{int}} := 2\alpha\mathbf{M}^{\frac{1}{2}} \left(\mathbf{M}^{-\frac{1}{2}}\mathbf{K}\mathbf{M}^{-\frac{1}{2}}\right)^{\frac{1}{2}} \mathbf{M}^{\frac{1}{2}}, \tag{1.2}$$

where $\alpha \ll 1$ and $\mathbf{M}$ is assumed to be nonsingular, see [34, 36]. However, the theory presented in this work is more general and can be applied to all modal dampers, which include, e.g., Rayleigh damping defined in [81, 155].

The external damping $\mathbf{D}_{\text{ext}}(c, g)$ depends on two types of parameters. The first ones are the damping positions $c = \left[c_1, \ldots, c_\ell\right]^{\mathrm{T}} \in \mathfrak{D}_c \subset \{1, \ldots, n\}^\ell$, which are stored in the matrix $\mathbf{F}(c) \in \mathbb{R}^{n\times\ell}$. The structure of the matrix $\mathbf{F}(c)$ depends on the damper type so that, e.g., grounded dampers are described by unit vectors $e_{c_1}, \ldots, e_{c_\ell}$, which are concatenated to build the matrix $\mathbf{F}(c)$. The second parameters are the damping gains $g = \left[g_1, \ldots, g_\ell\right]^{\mathrm{T}} \in \mathfrak{D}_g \subset \mathbb{R}^\ell_+$, which represent the viscosities of the dampers. We assume that the viscosities g_j are fixed over time and lie in given intervals $[g_j^-, g_j^+]$, for all $j = 1, \ldots, \ell$. We encode these conditions by setting $g \in \mathfrak{D}_g$, where the parameter set $\mathfrak{D}_g$ contains all given conditions. The different external dampers are described in more detail for different numerical examples later in this work. The resulting external damper is then given as

$$\mathbf{D}_{\text{ext}}(c, g) := \mathbf{F}(c)\mathbf{G}(g)\mathbf{F}(c)^{\mathrm{T}}, \qquad \mathbf{G}(g) := \operatorname{diag}(g_1, \ldots, g_\ell).$$

We assume that the number of external dampers ℓ is significantly smaller than the dimension n, i.e., $\ell \ll n$.

Since it is infeasible to measure and evaluate the behavior of all states individually, if n is large, we need to define an output function. In this work, two different output types are investigated. The first one is a linear output equation, which results in a system

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}(c, g)\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) &= \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0, \\ \mathbf{y}_{\mathrm{L}}(t) &= \mathbf{C}_1\mathbf{x}(t) + \mathbf{C}_2\dot{\mathbf{x}}(t) \end{aligned} \tag{1.3}$$

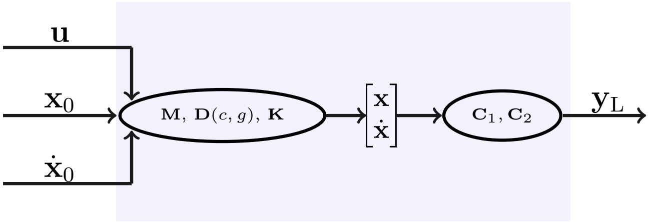

$\mathbf{u}$
$\mathbf{x}_0$
$\dot{\mathbf{x}}_0$
$\mathbf{M},\ \mathbf{D}(c,g),\ \mathbf{K}$
$\begin{bmatrix}\mathbf{x}\\ \dot{\mathbf{x}}\end{bmatrix}$
$\mathbf{C}_1, \mathbf{C}_2$
$\mathbf{y}_\mathrm{L}$

Figure 1.1: Structure of a second-order system with a linear output.

with matrices $\mathbf{C}_1,\ \mathbf{C}_2 \in \mathbb{R}^{p\times n}$ so that we observe the displacements by evaluating $\mathbf{C}_1\mathbf{x}(t)$ and velocities by evaluating $\mathbf{C}_2\dot{\mathbf{x}}(t)$. In practice, we are often only interested in the displacement properties, i.e., we set $\mathbf{C}_2 = 0$. The system (1.3) is depicted in Figure 1.1.

The second output type considered in this work is a quadratic output equation that is described by

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}(c,g)\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0, \\ \mathbf{y}_\mathrm{Q}(t) = \begin{bmatrix}\mathbf{x}(t)^\mathrm{T} & \dot{\mathbf{x}}(t)^\mathrm{T}\end{bmatrix} \mathcal{M} \begin{bmatrix}\mathbf{x}(t)\\ \dot{\mathbf{x}}(t)\end{bmatrix}, \end{aligned} \tag{1.4}$$

where $\mathcal{M} \in \mathbb{R}^{2n\times 2n}$. These systems can be interpreted as a special class of Wiener models. These output equations arise when investigating the variance or deviation of the state and velocity variable from a certain reference point, which can be represented as a quadratic function of the state. Also, when considering the potential and kinetic energy of the system as an output, which is given by

$$\mathbf{E}_\mathrm{pot} := \frac{1}{2}\mathbf{x}(t)^\mathrm{T}\mathbf{K}\mathbf{x}(t), \qquad \mathbf{E}_\mathrm{kin} := \frac{1}{2}\dot{\mathbf{x}}(t)^\mathrm{T}\mathbf{M}\dot{\mathbf{x}}(t),$$

we consider quadratic output equations. Moreover, when considering, e.g., the 2-norm of the output or some weighted norms, we measure quadratic output equations. Examples can be found in [12, 40, 69, 70, 99, 100].

In Figure 1.2, the structure of system (1.4) is depicted, where two inputs and outputs are added to the system to indicate the quadratic output equation.

We want to clarify that images in Figure 1.1 and Figure 1.2 deviate from the typical used diagrams in the control engineering literature. Nevertheless, they are used in this dissertation as they serve as a convenient tool for vividly illustrating the approaches introduced in this work.

To simplify computations, the second-order systems in (1.3) and (1.4) can also be

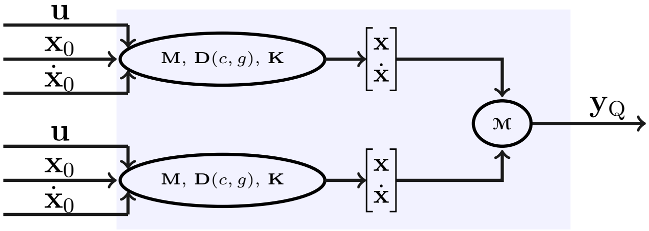

$\mathbf{u}$
$\mathbf{x}_0$
$\dot{\mathbf{x}}_0$
$\mathbf{M}$, $\mathbf{D}(c,g)$, $\mathbf{K}$
$\begin{bmatrix}\mathbf{x}\\\dot{\mathbf{x}}\end{bmatrix}$
$\mathcal{M}$
$\mathbf{y}_\mathrm{Q}$
$\mathbf{u}$
$\mathbf{x}_0$
$\dot{\mathbf{x}}_0$
$\mathbf{M}$, $\mathbf{D}(c,g)$, $\mathbf{K}$
$\begin{bmatrix}\mathbf{x}\\\dot{\mathbf{x}}\end{bmatrix}$

Figure 1.2: Structure of a second-order system with a quadratic output.

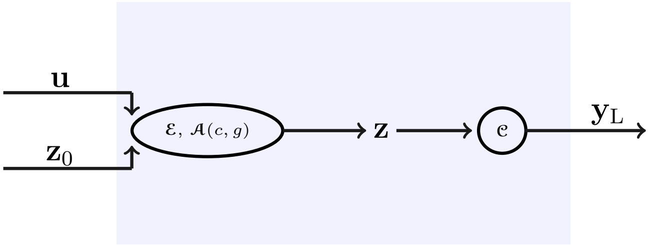


Figure 1.3: Structure of a first-order system with a linear output.

written in first-order form, i.e.,

$$\begin{aligned} \boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) &= \boldsymbol{\mathcal{A}}(c,g)\mathbf{z}(t) + \boldsymbol{\mathcal{B}}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0, \\ \mathbf{y}_\mathrm{L}(t) &= \boldsymbol{\mathcal{C}}\mathbf{z}(t) \end{aligned} \tag{1.5}$$

and

$$\begin{aligned} \boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) &= \boldsymbol{\mathcal{A}}(c,g)\mathbf{z}(t) + \boldsymbol{\mathcal{B}}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0, \\ \mathbf{y}_\mathrm{L}(t) &= \mathbf{z}(t)^\mathrm{T}\boldsymbol{\mathcal{M}}\mathbf{z}(t), \end{aligned} \tag{1.6}$$

respectively, with first-order matrices

$$\begin{aligned} &\boldsymbol{\mathcal{E}} := \begin{bmatrix} \mathbf{I}_n & 0 \\ 0 & \mathbf{M} \end{bmatrix}, \qquad \boldsymbol{\mathcal{A}}(c,g) := \begin{bmatrix} 0 & \mathbf{I}_n \\ -\mathbf{K} & -\mathbf{D}(c,g) \end{bmatrix}, \qquad \boldsymbol{\mathcal{B}} := \begin{bmatrix} 0 \\ \mathbf{B} \end{bmatrix}, \qquad \mathbf{z}_0 = \begin{bmatrix} \mathbf{x}_0 \\ \dot{\mathbf{x}}_0 \end{bmatrix}, \\ &\boldsymbol{\mathcal{C}} := \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix}, \qquad \boldsymbol{\mathcal{M}} := \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{12}^\mathrm{T} & \mathbf{M}_{22} \end{bmatrix}. \end{aligned} \tag{1.7}$$

The inputs $\mathbf{u}(t) \in \mathbb{R}^m$ and the outputs $\mathbf{y}_\mathrm{L}(t) \in \mathbb{R}^p$, $\mathbf{y}_\mathrm{Q}(t) \in \mathbb{R}$ are equal to those in (1.3) and (1.4), and the state in first-order representation is $\mathbf{z}(t) \in \mathbb{R}^N$ with $N = 2n$. The structures of the two first-order systems (1.5) and (1.6) are depicted in Figure 1.3 and Figure 1.4, respectively. In the following, we consider the systems in first-order and second-order representations since both can be advantageous depending on the application.

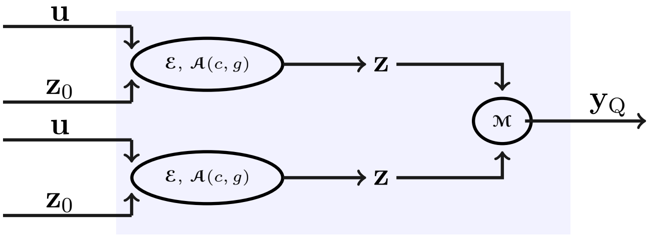


Figure 1.4: Structure of a first-order system with a quadratic output.

Exhibiting complex dynamic behavior may result in high-fidelity models, i.e., the dimension of the state vector n or N is large. Hence, engineering design processes become computationally very demanding. As a remedy, we seek to employ model reduction techniques that allow us to construct a low-dimensional model that closely resembles the dynamic behaviors of the high-fidelity model. Our goal is to construct reduced-order surrogate models while preserving the original structure. We consider, in this work, parameter-independent systems as well as parameter-dependent ones.

First, we consider systems whose external dampers are already defined and are, therefore, parameter-independent. This situation appears, e.g., when we want to investigate the system behavior for a given external damper. Because of the high dimension of the original system, we aim to derive a reduced model that approximates the effect of the input and the initial condition on the output. There are several methods to reduce dynamical systems in the literature. However, we consider inhomogeneous systems with linear and quadratic output equations. Since most of these systems are not considered in the literature so far, in this work, we derive reduction methods tailored for these non-standard system structures. Therefore, we derive system matrices that are called *Gramians* and encode the controllability and observability behavior. These Gramians are used to identify significant controllability and observability subspaces, which define the reduced surrogate models. Moreover, we derive respective error bounds for the presented methods, which are used to evaluate the quality of the system approximations.

Second, we consider the problem of finding optimal external dampers, for which we have to investigate parameter-dependent systems. These parameter-dependent systems need to be evaluated at every step of the optimization process. Our goal is to design the damping values based on the optimization of an objective function $\boldsymbol{\mathcal{J}}(c, g)$. For a given vibrational system, we determine the best damping $\mathbf{D}(c, g)$ that ensures optimal attenuation of the output $\mathbf{y}_{\mathrm{L}}$ or $\mathbf{y}_{\mathrm{Q}}$. The L_∞-norm of $\mathbf{y}_{\mathrm{L}}$ or $\mathbf{y}_{\mathrm{Q}}$ is bounded by the *system response* and defined as

$$\boldsymbol{\mathcal{J}}_{\mathrm{L}}(c, g) := \operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}(c, g)\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big)$$

when we consider a system (1.3) with a linear output equation, and

$$\boldsymbol{\mathcal{J}}_{\mathrm{Q}}(c, g) := \operatorname{tr}(\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}(c, g)\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}(c, g))$$

when considering the system (1.4) with a quadratic output equation. The matrix $\boldsymbol{\mathcal{P}}(c, g)$ is the *controllability Gramian* that spans the controllability space of both systems (1.3) and (1.4). We aim to optimize the damping values in such a way that the system response is minimized. This criterion was also used in [25, 60, 140]. The Gramian $\boldsymbol{\mathcal{P}}(c, g)$ is computed by solving the continuous-time Lyapunov equation

$$\boldsymbol{\mathcal{A}}(c, g)\boldsymbol{\mathcal{P}}(c, g)\boldsymbol{\mathcal{E}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}(c, g)\boldsymbol{\mathcal{A}}(c, g)^{\mathrm{T}} = -\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}. \tag{1.8}$$

To find the damping gains $(c, g) \in \boldsymbol{\mathcal{D}}$ that minimize the energy response $\boldsymbol{\mathcal{J}}(c, g)$, we have to solve a Lyapunov equation (1.8) in every step of the optimization method. Since the Lyapunov equation solves are computationally very demanding if the matrices are of large dimensions, the minimization process would lead to high computational cost and, hence, be inefficient or unfeasible in a large-scale setup. To accelerate the optimization process, we propose new reduction methods. We derive offline-online methods to generate bases spanning an approximation to the solution space of the Lyapunov equations for all possible positions and viscosities of the dampers. Furthermore, we derive an adaptive scheme that generates the reduced solution space by adding the subspaces of interest. Then, we define the corresponding reduced optimization problem that is solvable in a reasonable amount of time. Also, we decouple the solution spaces of the problem to obtain a space that corresponds to the system without external dampers and serves as a starting point for the reduction of the optimization problem. In addition, we derive spaces corresponding to the different damper positions, which are used to expand the reduced basis if needed. To evaluate the quality of the basis, we introduce different error estimators. Our new techniques produce reduced optimization problems of significantly smaller dimensions, which are faster to solve than the original problem.

1.2 Literature overview

Vibrational systems and their contained dampers have been studied in the last decades, for example, in [14, 49, 55, 71, 73, 76, 85, 96, 153], where external dampers are considered in systems that already contain internal damping of small magnitude. In this work, we consider model reduction schemes for parameter-independent and parameter-dependent vibrational systems.

There is a vast amount of literature that considers parameter-independent systems. For ordinary differential equation (ODE) systems with a linear output equation and homogenous initial conditions, there exist several methods to construct reduced-order models, e.g., singular value-based approaches such as balanced truncation [26, 93, 138] and Hankel norm approximations [56]. Also, the authors in [20] extend the BT method for systems with a quadratic output equation. Moreover, moment matching methods [5, 60, 84] and Krylov subspace methods, e.g., the iterative rational Krylov algorithm (IRKA) [26, 51, 60, 61] are used frequently. An overview of these methods is given, e.g., in [5, 22, 23, 26]. Moreover, the authors in [13, 15, 66, 121] provide methods to reduce systems with inhomogeneous initial conditions.

All methods mentioned above treat systems with a nonsingular matrix $\mathcal{E}$. Therefore, they are not directly applicable to systems with a DAE as a state equation. This issue is addressed in, e.g., [33, 61, 91, 130]. Existing methods that deal with the DAE case include interpolatory projection methods [1–3, 61] and balancing-based methods [28, 67, 91, 130, 131]. In this work, we mostly focus on a balancing-based method. DAE systems require the corresponding projection matrices that describe the deflating subspaces. Such projection matrices are difficult to form explicitly. However, the structure of the DAE systems is often known and can be used to define and implicitly apply the projection matrices in practice. For details, we refer to [28, 31, 67, 116, 133]. Also, the classic BT method is not directly applicable to the case of quadratic output equations since the observability space is not of the same form as in the linear output case. Hence, the observability Gramian, defined in [91], can not be used in this setup. In [20], the authors derived Gramians corresponding to ODE systems (meaning $\mathcal{E} = \mathbf{I}$ in (1.5)) with quadratic output equations. However, the methodology proposed in [20] cannot be directly applied to DAEs due to the singularity of the matrix $\mathcal{E}$. Therefore, there is a necessity to modify BT to incorporate the differential-algebraic structure, which is investigated in this dissertation.

For second-order systems, there exist different tailored model order reduction methods. BT and balancing-based approaches for second-order systems were introduced in [44, 92, 112]. Also, Krylov methods tailored for second-order systems were derived in [17, 134] and generalized to rational interpolation in [9, 10, 53, 149]. Overviews of these methods can be found, e.g., in [43, 117]. However, none of these methods considers inhomogeneous systems.

To deal with parameter-dependent systems, in this work, we apply the reduced basis method (RBM) and modifications of it. We reduce the Lyapunov equation in (1.8) to derive a surrogate equation that is solvable in a reasonable time. The RBM was first introduced to reduce parameter-dependent partial differential equations, see [68, 109, 150–152]. Later, it was used for Riccati equations [119], and, finally, the RBM was applied to Lyapunov equations by Son and Stykel in [126]. In [108], the authors use the RBM to reduce parametric differential–algebraic systems.

We aim to apply the RBM to optimize the effect of the input on the system output. Therefore, we want to choose external dampers that stabilize the system and shift eigenfrequencies so that possible external loads do not lead to resonances. The problem of finding optimal external dampers was widely investigated in the literature, see [55, 71, 74, 95, 135, 153]. In this work, we use the RBM to accelerate the optimization process. In the literature, other approaches were applied to the problem of damping optimization. Depending on the application, different criteria are chosen to quantify the stability of systems and the response to external disturbances. When systems (1.3) with

$\mathbf{B} \equiv 0$ are considered, then the spectral abscissa or the total average energy are used as described in [36, 52, 97, 144, 147]. In [35, 142, 148], the authors present different reduction techniques to optimize the related problem of minimizing the total average energy for the system (1.3) with no input.

In the case $\mathbf{B} \neq 0$, as considered in this work, external disturbances are taken into account, which potentially plays an important role in real-life scenarios. In these cases, the average displacement amplitude can be evaluated, which minimizes the square of the norm of the displacement $\mathbf{x}(t)$ averaged over a certain time period, see [82, 145]. Another criterion used in this work is the average energy amplitude corresponding to the minimization of the *system response*, $\mathfrak{J}_{\mathrm{L}}(c, g)$ or $\mathfrak{J}_{\mathrm{Q}}(c, g)$, of the system describing the input-to-output behavior in the frequency domain. This optimization criterion was also used in [25, 140].

Moreover, the authors in [25] utilize the dominant pole algorithm to build a reduced minimization problem that is quickly solvable. In [140], an efficient optimization approach using structure-preserving parametric model reduction based on the iterative rational Krylov algorithm (`sym2IRKA`) is used to derive an efficient optimization algorithm. In [16], a sampling-free approach is presented that reduces the system (1.3) for all admissible parameters. Alternatively, in [37, 141], the authors optimize the $\mathcal{H}_\infty$-norm of the systems constraining the L_2-norm of the output $\mathbf{y}_{\mathrm{L}}$ of the corresponding system, which can be interpreted as the worst-case amplification of the output energy caused by an input signal. Most of the established methods consider the optimization of the damping viscosities.

The optimization of the discrete damper positions is still a challenging problem, especially for large systems, which has been studied in [34, 50, 50, 62, 62, 75, 136, 139, 143]. In particular, in [34, 139], the authors describe the optimization using a discrete-to-continuous approach, which is modified and used in this work.

1.3 Goal of this thesis

In this work, we consider the problem of model reduction and optimization of external dampers for large-scale vibrational systems. Therefore, the two main goals of this work are the following.

System theory and model reduction methods for systems in non-standard form

The model order reduction of parameter-independent systems is needed to evaluate the behavior of (damped) systems, where we consider BT as well as the IRKA method. We investigate first-order ODE systems, first-order DAE systems, and second-order ODE systems with inhomogeneous initial conditions. Also, we consider systems with linear and quadratic output equations. In this work, we derive BT methods for systems in these non-standard forms that appear when considering vibrational systems. In particular, the novelties of this work include the introduction of BT methods for inhomogeneous ODE systems with quadratic output equations, inhomogeneous DAE systems with linear and quadratic output equations, and inhomogeneous second-order ODE systems with linear and quadratic output equations. Therefore, we derive respective suitable system representations, tailored Gramians for the different system types, the corresponding energy interpretations, and error bounds that describe the quality of the approximations based on the respective Gramians. We demonstrate the effectiveness of the derived algorithms by applying them to some numerical examples.

RBM and damping optimization for vibrational systems We also solve the problem of reducing parameter-dependent systems that arise when optimizing positions and viscosities of external dampers in vibrational systems. Therefore, we apply RBM approaches that generate a basis that spans an approximation of the controllability space of the system, which defines a reduced surrogate model. First, we use the offline-online RBM introduced in [126] and extend this method to second-order systems. Moreover, we derive a decoupling of the controllability space of the respective systems. This decoupling can accelerate our RBM for first-order and second-order systems. Afterwards, we tailor the derived RBM schemes to be more suitable for the damping optimization process in vibrational systems. In addition, we derive an adaptive scheme in which we enrich the basis within the optimization process. Therefore, prior knowledge of the assumed parameters is not necessary. Additionally, we derive several error estimates suitable for the different methods.

For both topics, similar system theoretical considerations need to be done beforehand. Hence, first, we investigate the three types of dynamical systems (first-order ODE systems, first-order DAE systems, and second-order ODE systems) and their controllability, observability, and the corresponding system energies to have a theoretical foundation for the rest of the thesis.

1.4 Overview of the author's contributions

The theory and results presented in this thesis have been partially published in [105–107]. The theoretical results from [105, 106] are part of Chapter 3, and the resulting reduction methods are introduced in Chapter 4. The main contributions from [107] are described in Chapter 5 and Chapter 6. All of the chapters presented in this thesis are extended versions of these papers.

In [105], the author derives a BT method that reduces DAE systems with quadratic output equations. New proper and improper Gramians are derived with suitable energy interpretations that result in a BT method. Also, error bounds are determined to quantify the quality of the system approximation. This work is a natural extension of the theory in [20], where ODE systems with quadratic output equations are considered. In this thesis, [105] is extended to DAE systems with quadratic output equations and inhomogeneous initial conditions.

Moreover, in [106], the authors investigate inhomogeneous second-order systems with nonzero initial conditions. They derive tailored Gramians, energy functionals, and error bounds, which result in a BT method that reduces second-order systems. In this thesis, we also derive a BT scheme that reduces inhomogeneous second-order systems with a quadratic output equation, which is an extension of the published work in [106].

In [107], the authors derive a reduction scheme to optimize the viscosities of some external dampers in vibrational systems. Therefore, they reduce the respective parametric homogeneous second-order systems using the RBM. Together with an error estimator, this method exceeded the acceleration rates from [140], where the authors use an IRKA-based reduction scheme. This method was the fastest so far in the literature. The RBM-based method from [107] is extended in this dissertation. Hence, we also consider second-order systems with quadratic output equations.

Finally, in collaboration with Matea Ugrica, Ninoslav Truhar, and Peter Benner, the author derived a decoupling in the controllability space of parametric homogeneous second-order systems that is used to derive approximations of the controllability spaces of the respective systems. These controllability space approximations are used to derive reduced parametric systems in which the external dampers' viscosities and positions are optimized. Also, these theories are extended to systems with quadratic output equations in this work.

1.5 Outline

This work is organized as follows. In Chapter 2, we review existing theories and methods, including system theoretical concepts, resulting model reduction schemes, and solution strategies for Lyapunov equations as part of the reduction methods.

Afterwards, in Chapter 3, we derive different system theoretical concepts for dynamical systems in a non-standard form. They include transfer functions, system equivalences, corresponding Gramians, and the respective energy interpretations. These concepts are used throughout the remaining work.

In Chapter 4, we study model reduction schemes for different parameter-independent system types, in particular, BT and IRKA methods, where the main focus lies on the BT method. We extend existing model reduction schemes to systems in non-standard form and derive respective error estimators. Using some numerical examples, we demonstrate the efficiency of these methods. These methods are applied in the context of damping optimization when we have a trial external damper for which we aim to analyze the respective system behavior.

In Chapter 5, we revisit and extend the RBM to the different system structures.

Moreover, we derive a decoupling of the controllability space, which is used to accelerate the basis-building process. For the two resulting RBM methods, we also derive suitable error estimators.

Afterwards, in Chapter 6, the RBM methods are used to reduce the problem of finding optimal external dampers concerning the system response. Moreover, we derive an adaptive scheme that enriches the respective basis within the optimization process. Hence, we can ensure that no unnecessary information is contained in the reduced basis. Moreover, we extend this approach by a controllability space decomposition that accelerates the methods. Moreover, this decoupling leads to an RBM optimization process that does not require a given parameter set. Again, we derive suitable error estimators and apply the resulting optimization methods to some numerical examples.

Finally, in Chapter 7, we conclude the work and give an outlook on future work perspectives.

CHAPTER 2

MATHEMATICAL BACKGROUND

Contents

In this chapter, we give an overview of various mathematical theories and methods that form the mathematical background of this thesis. First, we describe system properties and theoretical concepts in Section 2.1.1 for different system structures. Afterwards, we present model reduction methods for these classes of systems in Section 2.2. One of the reduction methods, balanced truncation, uses solutions of Lyapunov equations to identify the dominant controllability and observability subspaces. Therefore, in Section 2.3, we describe existing numerical methods to solve Lyapunov equations, especially for those with large dimensions.

In the remaining course of this thesis, the theoretical concepts and methods from this chapter are extended to systems with inhomogeneous initial conditions and to systems with quadratic output equations. These extended concepts are needed to solve the problem of optimizing external dampers in mechanical systems presented in Chapter 1.

2.1 System theoretical concepts

In this section, we consider several classes of systems with linear output equations and provide an overview of the respective basic system theoretical concepts. These will be used in the remainder of this work to identify significant states and the resulting dominant controllability and observability spaces corresponding to these systems. The concepts introduced in this section were originally derived in the field of control theory, see [4, 88, 127, 159], where the aim is to provide a mathematical implementation of real-life dynamical systems through analysis of input-output behavior.

We analyze the systems presented in Chapter 1 that arise from mechanical systems considered in the context of damping optimization. We study parameter-independent systems, which means we consider the second-order system (1.3) for a fixed external damper $\mathbf{D}(c, g) \equiv \mathbf{D}$. Also, the resulting first-order system (1.5) is assumed to be parameter independent such that $\boldsymbol{\mathcal{A}}(c, g) \equiv \boldsymbol{\mathcal{A}}$. Moreover, we allow the matrix $\boldsymbol{\mathcal{E}}$ to be singular in systems with first-order structure (1.5). This situation occurs when the mass matrix $\mathbf{M}$ of the second-order system (1.3) is singular. In this case, we consider systems with differential-algebraic equations (DAEs) as state equations. However, we only consider the DAE case in its first-order representation, as considering second-order descriptor systems is beyond the scope of this work. For system theoretical concepts for second-order DAE systems, we refer to the work [86] that was further used and extended in [1, 30, 32, 77] for particular index classes.

The different system structures are analyzed separately below. In Section 2.1.1, we consider system theoretical aspects of first-order systems with an ODE as a state equation. In Section 2.1.2, we investigate first-order systems with a DAE as a state equation, and, finally, in Section 2.1.3, we study second-order systems.

2.1.1 First-order ODE systems

In this subsection, we repeat selected, well-known system theoretical concepts for first-order systems, that are dynamical systems of the form

$$\begin{aligned} \boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) &= \boldsymbol{\mathcal{A}}\mathbf{z}(t) + \boldsymbol{\mathcal{B}}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0, \\ \mathbf{y}_{\mathrm{L}}(t) &= \boldsymbol{\mathcal{C}}\mathbf{z}(t), \end{aligned} \tag{2.1}$$

where $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{N\times N}$, $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{N\times m}$ and $\boldsymbol{\mathcal{C}} \in \mathbb{R}^{p\times N}$. The matrix $\boldsymbol{\mathcal{E}}$ is assumed to be nonsingular so that the state equation in (2.1) is an ODE. The input, the state and the output are $\mathbf{u}(t) \in \mathbb{R}^m$, $\mathbf{z}(t) \in \mathbb{R}^N$, and $\mathbf{y}_{\mathrm{L}}(t) \in \mathbb{R}^p$, respectively, with $\mathbf{u} \in L_2([0, \infty), \mathbb{R}^m)$. The theory repeated in this section is based on [4, 88, 127, 159]. The solution of the first-order state equation in (2.1) is equal to

$$\mathbf{z}(t) = \int_0^t e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}(t-\tau)}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\mathbf{u}(\tau)\mathrm{d}\tau + e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\mathbf{z}_0. \tag{2.2}$$

From the state trajectory, stability properties can be derived, i.e., convergence to an equilibrium state when no external force is acting. Equation (2.2) also describes the controllability behavior of the system that indicates which states are reachable. Both properties are significant for the analysis of the system. Moreover, the observability of (2.1) is of interest in this work, which describes whether states are uniquely identifiable based on the output observations. Hence, we define these properties formally in the following.

Definition 2.1:
The system (2.1) is called

1. *asymptotically stable* if all the solutions $\mathbf{z}(t) = e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\mathbf{z}_0$ of the linear ODE
$$\boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) = \boldsymbol{\mathcal{A}}\mathbf{z}(t)$$
satisfy $\lim_{t\to\infty} \mathbf{z}(t) = 0$ for all initial states $\mathbf{z}(0) = \mathbf{z}_0$;

2. *controllable* if for all initial conditions $\mathbf{z}(0) = \mathbf{z}_0 \in \mathbb{R}^N$ and all $\mathbf{z}_1 \in \mathbb{R}^N$ there exists a time $t_1 > 0$ and a control function $\mathbf{u} \in L_2([0,\infty),\mathbb{R}^m)$ in the set of all admissible inputs so that the state trajectory in (2.2) yields
$$\mathbf{z}(t_1) = \mathbf{z}_1;$$

3. *observable* if for two solution trajectories $\mathbf{z}(\cdot)$ and $\widetilde{\mathbf{z}}(\cdot)$ from (2.2) resulting from the same input $\mathbf{u} \in L_2([0,\infty),\mathbb{R}^m)$ it holds that
$$\boldsymbol{\mathcal{C}}\mathbf{z}(t) = \boldsymbol{\mathcal{C}}\widetilde{\mathbf{z}}(t) \qquad \text{for all} \qquad t \geq 0$$
implies that $\mathbf{z}(t) = \widetilde{\mathbf{z}}(t)$ for all $t \geq 0$. ◊

We call a system to be in *minimal realization* if it is controllable and observable. Since it is difficult to check these properties by definition, we will repeat some equivalent properties that will help us to characterize the dynamical system (2.1).

Theorem 2.2:
Consider the system (2.1). The following equivalences hold.

1. The system is asymptotically stable if and only if all eigenvalues of the matrix pencil $\lambda\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}$ lie in the negative complex half-plane, that means if
$$\Lambda(\boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{A}}) \subset \mathbb{C}^- := \{\lambda \in \mathbb{C} \mid \operatorname{Re}(\lambda) < 0\}.$$

2. The system is controllable if and only if
$$\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}} & \boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}} & \dots & (\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}})^{N-1}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\end{bmatrix}\right) = N.$$

3. The system is observable if and only if

$$\operatorname{rank}\left(\begin{bmatrix} \mathfrak{C} \\ \mathfrak{C}\mathcal{E}^{-1}\mathcal{A} \\ \vdots \\ \mathfrak{C}(\mathcal{E}^{-1}\mathcal{A})^{N-1} \end{bmatrix}\right) = N.$$

◊

After introducing the basic system properties, we derive some tools used to describe the overall controllability and observability behavior. For this purpose, we study the system dynamics in the frequency domain, which means we apply the Laplace transform to the system (2.1) with zero initial conditions, which leads to the state equation

$$\mathbf{Z}(s) = (s\mathcal{E} - \mathcal{A})^{-1}\mathfrak{B}\mathbf{U}(s),$$

where $\mathbf{Z}$ and $\mathbf{U}$ denote the Laplace transforms of $\mathbf{z}$ and $\mathbf{u}$, respectively. Inserting $\mathbf{Z}(s)$ into the output equation in the frequency domain with $\mathbf{Y}_{\mathrm{L}}$ being the Laplace transform of $\mathbf{y}_{\mathrm{L}}$ results in

$$\mathbf{Y}_{\mathrm{L}}(s) = \mathfrak{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathfrak{B}\mathbf{U}(s). \tag{2.3}$$

Based on this frequency domain representation of the output, we can define the systems transfer function that encodes the input-to-output behavior.

Definition 2.3:
Consider the system (2.1). Then the corresponding *transfer function* is defined as

$$\mathfrak{G}_{\mathrm{L}}(s) := \mathfrak{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathfrak{B}. \tag{2.4}$$

◊

We use the following definition to describe the system behavior of (2.1) concerning its transfer function.

Definition 2.4:
Consider the system (2.1). The corresponding transfer function $\mathfrak{G}_{\mathrm{L}}(s)$ as defined in (2.4) is called

a) *strictly proper* if $\lim_{\omega\to\infty} \|\mathfrak{G}_{\mathrm{L}}(\mathrm{i}\omega)\|_2 = 0$,

b) *proper* if $\lim_{\omega\to\infty} \|\mathfrak{G}_{\mathrm{L}}(\mathrm{i}\omega)\|_2 < \infty$,

c) *improper* if $\lim_{\omega\to\infty} \|\mathfrak{G}_{\mathrm{L}}(\mathrm{i}\omega)\|_2 = \infty$. ◊

In the following, we introduce some matrices, the so-called Gramians, which provide information about the controllability and observability spaces of the system, including all reachable and observable states. To give an intuition of how these Gramians are defined, we first introduce the input-to-state mapping and the state-to-output mapping

$$\boldsymbol{c}(t) := e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{B} \qquad \text{and} \qquad \boldsymbol{o}_{\mathrm{L}}(t) := \mathcal{C}e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}.$$

We add the subscript L to the state-to-output mapping $\boldsymbol{o}_{\mathrm{L}}(s)$ to emphasize that we consider a linear output equation since later in this work, we also investigate systems with quadratic output equations. Since the mappings $\boldsymbol{c}$ and $\boldsymbol{o}_{\mathrm{L}}$ encode the reachable and observable states of the system, the integration over the entire time domain provides the Gramians that span the respective spaces.

Definition 2.5:
Consider the asymptotically stable system (2.1). The respective *controllability* and *observability Gramian* are defined as

$$\begin{aligned} \mathcal{P} &:= \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{B}\mathcal{B}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t}\mathrm{d}t, \\ \mathcal{Q}_{\mathrm{L}} &:= \int_0^\infty e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t}\mathcal{C}^{\mathrm{T}}\mathcal{C}e^{\mathcal{E}^{-1}\mathcal{A}t}\mathrm{d}t. \end{aligned} \tag{2.5}$$

◊

As stated, e.g. in [4], these Gramians are computed by solving the Lyapunov equations

$$\mathcal{A}\mathcal{P}\mathcal{E}^{\mathrm{T}} + \mathcal{E}\mathcal{P}\mathcal{A}^{\mathrm{T}} = -\mathcal{B}\mathcal{B}^{\mathrm{T}}, \qquad \mathcal{A}^{\mathrm{T}}\widetilde{\mathcal{Q}}_{\mathrm{L}}\mathcal{E} + \mathcal{E}^{\mathrm{T}}\widetilde{\mathcal{Q}}_{\mathrm{L}}\mathcal{A} = -\mathcal{C}^{\mathrm{T}}\mathcal{C}, \tag{2.6}$$

where $\mathcal{E}^{-\mathrm{T}}\mathcal{Q}_{\mathrm{L}}\mathcal{E}^{-1} = \widetilde{\mathcal{Q}}_{\mathrm{L}}$. The Gramians introduced in (2.5) are used in the next section to identify dominant subspaces and derive respective reduced surrogate models that approximate the input-to-output behavior of the original system (2.1) described by the transfer functions introduced in (2.4). We recall the definition of *Hardy spaces*, the corresponding scalar products, and norms that we utilize to quantify the output errors between the original system and the reduced approximation by evaluating the respective transfer functions. The first Hardy space, we consider, is the $\mathcal{H}_2^{p\times m}$-space that is defined as

$$\mathcal{H}_2^{p\times m} := \left\{\mathcal{G} : \mathbb{C}^+ \to \mathbb{C}^{p\times m} : \mathcal{G} \text{ is analytic in } \mathbb{C}^+ \text{ and } \int_{-\infty}^{\infty} \|\mathcal{G}(\mathrm{i}\omega)\|_{\mathrm{F}}^2 \mathrm{d}\omega < \infty\right\}. \tag{2.7}$$

Its scalar product is

$$\langle \mathcal{H}, \mathcal{G}\rangle_{\mathcal{H}_2} := \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{tr}\big(\mathcal{H}(\mathrm{i}\omega)^{\mathrm{H}}\mathcal{G}(\mathrm{i}\omega)\big)\,\mathrm{d}\omega$$

and the resulting norm is

$$\|\mathcal{G}\|_{\mathcal{H}_2} := \langle \mathcal{G}, \mathcal{G} \rangle_{\mathcal{H}_2}^{\frac{1}{2}} = \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \|\mathcal{G}(\mathrm{i}\omega)\|_{\mathrm{F}}^2 \mathrm{d}\omega \right)^{\frac{1}{2}}.$$

This norm provides an upper bound on the L_∞-norm of the output, which characterizes the system's response to an input, as shown in the following proposition from [4, Proposition 5.2].

Proposition 2.6:
Consider the system (2.1) with the corresponding transfer function $\mathcal{G}_{\mathrm{L}}(s) \in \mathcal{H}_2^{p \times m}$. Then it holds

$$\|\mathbf{y}\|_{L_\infty} \leq \|\mathcal{G}_{\mathrm{L}}\|_{\mathcal{H}_2} \|\mathbf{u}\|_{L_2}. \qquad \diamondsuit$$

We see that the $\mathcal{H}_2$-norm of the transfer function serves as a criterion to estimate the maximal output. This output bound was used in the context of damping optimization in [25, 107, 140]. We choose this particular bound if we want to limit or minimize the maximum deflections, and therefore consider the L_∞-norm of the output.

2.1.2 First-order DAE systems

In this subsection, we consider differential-algebraic systems that are of the structure

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}(t) &= \mathcal{A}\mathbf{z}(t) + \mathcal{B}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0, \\ \mathbf{y}_{\mathrm{L}}(t) &= \mathcal{C}\mathbf{z}(t), \end{aligned} \tag{2.8}$$

with matrices as in (2.1) and a singular matrix $\mathcal{E}$. Hence, the state equation contains differential equations as well as algebraic ones. These systems arise when modeling industrial processes, e.g., electrical circuits, thermal and diffusion processes, multibody systems, and certain discretized partial differential equations [39, 41]. Throughout this work, the pencil $\lambda\mathcal{E} - \mathcal{A}$ is assumed to be *regular*, i.e., the polynomial $\det(\lambda\mathcal{E} - \mathcal{A})$ is not identically zero.

To deal with differential-algebraic systems, we first repeat the Weierstrass canonical form (WCF). According to [79], there exist matrices $\mathbf{W}$ and $\mathbf{T}$ that transform the differential equation of the system (2.8) into WCF, that is

$$\mathcal{E} = \mathbf{W} \begin{bmatrix} \mathbf{I}_{N_f} & 0 \\ 0 & \mathbf{N} \end{bmatrix} \mathbf{T}, \quad \mathcal{A} = \mathbf{W} \begin{bmatrix} \mathbf{J} & 0 \\ 0 & \mathbf{I}_{N_\infty} \end{bmatrix} \mathbf{T}, \quad \mathcal{B} = \mathbf{W} \begin{bmatrix} \widetilde{\mathbf{B}}_1 \\ \widetilde{\mathbf{B}}_2 \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} \widetilde{\mathbf{C}}_1 & \widetilde{\mathbf{C}}_2 \end{bmatrix} \mathbf{T} \tag{2.9}$$

where N_f and N_∞ are the numbers of the finite and infinite eigenvalues of the matrix pencil $(\mathbf{A}, \mathbf{E})$. The matrix $\mathbf{J} \in \mathbb{R}^{n_f \times n_f}$ represents a Jordan block associated with the finite eigenvalues, 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* (2.8) that is also called the *Kronecker index*. In practice, for large-scale systems, we do not calculate this transformed form of the system explicitly. Based on this form we derive certain theoretical concepts.

Moreover, we define the matrices

$$\mathbf{P}_{\mathrm{r}} = \mathbf{T}^{-1} \begin{bmatrix} \mathbf{I}_{N_f} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{T} \quad \text{and} \quad \mathbf{P}_{\mathrm{l}} = \mathbf{W} \begin{bmatrix} \mathbf{I}_{N_f} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{W}^{-1} \tag{2.10}$$

that are the spectral projectors onto the right and left deflating subspaces of the pencil $\lambda\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}$, corresponding to the finite eigenvalues, that describe these subspaces. However, such projection matrices are challenging to form explicitly. Alternatively, approaches, as introduced in [89], can be used to derive the deflating subspaces, which is numerically unfeasible when large-scale systems are considered as the respective computations include several matrix decompositions, also of dense matrices. Even if one manages, they can destroy the sparsity of the original matrices and, therefore, increase 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 [31, 67, 116, 133].

By multiplying the system (2.8) from the left by $\mathbf{W}^{-1}$ and replacing $\mathbf{z}(t) =: \mathbf{T}^{-1} \left[\begin{smallmatrix} \mathbf{z}_1(t) \\ \mathbf{z}_2(t) \end{smallmatrix}\right]$, we obtain the following system in WCF

$$\begin{aligned} \dot{\mathbf{z}}_1(t) &= \mathbf{J}\mathbf{z}_1(t) + \widetilde{\mathbf{B}}_1\mathbf{u}(t), & \mathbf{z}_1(0) &= \mathbf{z}_{1,0}, \\ \mathbf{N}\dot{\mathbf{z}}_2(t) &= \mathbf{z}_2(t) + \widetilde{\mathbf{B}}_2\mathbf{u}(t), & \mathbf{z}_2(0) &= \mathbf{z}_{2,0}. \end{aligned} \tag{2.11}$$

The system (2.11) provides the decoupled differential and algebraic states $\mathbf{z}_1(t)$ and $\mathbf{z}_2(t)$ that are

$$\mathbf{z}_1(t) = \int_0^t e^{\mathbf{J}(t-\tau)}\widetilde{\mathbf{B}}_1\mathbf{u}(\tau)\mathrm{d}\tau + e^{\mathbf{J}t}\mathbf{z}_{1,0}, \qquad \mathbf{z}_2(t) = \sum_{k=0}^{\nu-1} -\mathbf{N}^k\widetilde{\mathbf{B}}_2\mathbf{u}^{(k)}(t), \tag{2.12}$$

where $\mathbf{u}^{(k)}(t)$ describes the k-th derivative of the function $\mathbf{u} \in \mathcal{C}^{\nu-1}([0,\infty),\mathbb{R}^m)$ evaluated in the time variable t where we assume that the input is sufficiently differentiable. Furthermore, we define

$$\boldsymbol{\mathfrak{F}}_{\mathbf{J}}(t) := \mathbf{T}^{-1} \begin{bmatrix} e^{\mathbf{J}t} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{W}^{-1} \qquad \text{and} \qquad \boldsymbol{\mathfrak{F}}_{\mathbf{N}}(k) := \mathbf{T}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & -\mathbf{N}^k \end{bmatrix} \mathbf{W}^{-1} \tag{2.13}$$

and transform $\mathbf{z}_1(t)$ and $\mathbf{z}_2(t)$ into the original state space of system (2.8) to obtain the differential and algebraic states

$$\begin{aligned}\mathbf{z}_{\mathrm{p}}(t) &= \mathbf{T}^{-1}\begin{bmatrix}\mathbf{z}_1(t)\\ 0\end{bmatrix} = \int_0^t \boldsymbol{\mathfrak{F}}_{\mathbf{J}}(t-\tau)\boldsymbol{\mathfrak{B}}\mathbf{u}(\tau)\mathrm{d}\tau + \boldsymbol{\mathfrak{F}}_{\mathbf{J}}(t)\boldsymbol{\mathfrak{E}}\mathbf{z}_{\mathrm{p},0},\\ \mathbf{z}_{\mathrm{i}}(t) &= \mathbf{T}^{-1}\begin{bmatrix}0\\ \mathbf{z}_2(t)\end{bmatrix} = \sum_{k=0}^{\nu-1}\boldsymbol{\mathfrak{F}}_{\mathbf{N}}(k)\boldsymbol{\mathfrak{B}}\mathbf{u}^{(k)}(t)\end{aligned} \tag{2.14}$$

with $\mathbf{z}(t) = \mathbf{z}_{\mathrm{p}}(t) + \mathbf{z}_{\mathrm{i}}(t)$ and $\mathbf{z}_{\mathrm{p},0} = \mathbf{P}_{\mathrm{r}}\mathbf{z}_0$. We see that for the improper state $\mathbf{z}_2(t)$, the initial conditions need to satisfy

$$\mathbf{z}_2(0) = \sum_{k=0}^{\nu-1} -\mathbf{N}^k\widetilde{\mathbf{B}}_2\mathbf{u}^{(k)}(0)$$

to ensure solvability, that is, an initial state $\mathbf{z}_0 = \mathbf{z}(0)$ needs to satisfy

$$(\mathbf{I}_N - \mathbf{P}_{\mathrm{r}})\mathbf{z}_0 = \sum_{k=0}^{\nu-1}\boldsymbol{\mathfrak{F}}_{\mathbf{N}}(k)\boldsymbol{\mathfrak{B}}\mathbf{u}^{(k)}(0). \tag{2.15}$$

If the system fulfills this condition, it is called *consistent*. 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.

According to controllability and observability for ODE systems, introduced in Section 2.1.1, we introduce here the concepts of C-stability, C-controllability, and C-observability that were defined in [129, 130].

Definition 2.7:
The system (2.8) is called

1. *C-stable* if it has a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathfrak{E}})$ and all the finite eigenvalues of $\lambda\boldsymbol{\mathfrak{E}} - \boldsymbol{\mathcal{A}}$ lie in the open-left half-plane $\mathbb{C}^- := \{\lambda \in \mathbb{C} \mid \operatorname{Re}(\lambda) < 0\}$.

2. *C-controllable* (completely controllable) if
$$\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{\mathfrak{E}} & \boldsymbol{\mathfrak{B}}\end{bmatrix}\right) = N \quad \text{and} \quad \operatorname{rank}\left(\begin{bmatrix}\lambda\boldsymbol{\mathfrak{E}} - \boldsymbol{\mathcal{A}} & \boldsymbol{\mathfrak{B}}\end{bmatrix}\right) = N \quad \text{for all finite } \lambda \in \mathbb{C}.$$

3. *C-observable* (completely observable) if
$$\operatorname{rank}\left(\begin{bmatrix}\boldsymbol{\mathfrak{E}}\\ \boldsymbol{\mathfrak{C}}\end{bmatrix}\right) = N \quad \text{and} \quad \operatorname{rank}\left(\begin{bmatrix}\lambda\boldsymbol{\mathfrak{E}} - \boldsymbol{\mathcal{A}}\\ \boldsymbol{\mathfrak{C}}\end{bmatrix}\right) = N \quad \text{for all finite } \lambda \in \mathbb{C}.$$ ◊

We can also derive the transfer function for the DAE system (2.8), given that the matrix pencil $(\mathbf{A}, \mathbf{E})$ is regular. Since the input-to-output behavior is invariant under transformation, we use the system in WCF from (2.11) to define $\mathcal{G}_{\mathrm{L}}(s) = \mathcal{G}_{\mathrm{L,p}}(s) + \mathcal{G}_{\mathrm{L,i}}(s)$ with

$$\mathcal{G}_{\mathrm{L,p}}(s) := \widetilde{\mathbf{C}}_1(s\mathbf{I}_{N_f} - \mathbf{J})^{-1}\widetilde{\mathbf{B}}_1, \qquad \mathcal{G}_{\mathrm{L,i}}(s) := \widetilde{\mathbf{C}}_2(s\mathbf{N} - \mathbf{I}_{N_\infty})^{-1}\widetilde{\mathbf{B}}_2 \tag{2.16}$$

where $\mathcal{G}_{\mathrm{L,p}}$ is the strictly proper component of the transfer function and $\mathcal{G}_{\mathrm{L,i}}(s)$ is called the polynomial component. Summing over both transfer function components yields the following definition.

Definition 2.8:
Consider the system (2.8) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$. Its *transfer function* is defined as

$$\mathcal{G}_{\mathrm{L}}(s) := \mathcal{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathcal{B}. \qquad \diamondsuit$$

To describe the properties of the system related to this transfer function, we can apply the system theoretical concepts introduced in Definition 2.4 and Proposition 2.6. For more details, we refer to [79, 91, 130].

As for the ODE case, we can derive controllability and observability Gramians corresponding to the proper and improper part of the system as introduced in [91] based on the input-to-state mappings in the time domain

$$\boldsymbol{c}_{\mathrm{p}}(t) = \mathcal{F}_{\mathbf{J}}(t)\mathcal{B} \qquad \text{and} \qquad \boldsymbol{c}_{\mathrm{i}}(k) = \mathcal{F}_{\mathbf{N}}(k)\mathcal{B}.$$

The corresponding proper and improper controllability Gramians result when integrating over the entire time domain and summing over all indices $k = 0, \ldots, \nu - 1$, which leads to the following Gramian definition.

Definition 2.9:
Consider the C-stable system (2.8). The corresponding *proper and improper controllability Gramians* are defined as

$$\mathcal{P}_{\mathrm{p}} := \int_0^\infty \mathcal{F}_{\mathbf{J}}(t)\mathcal{B}\mathcal{B}^{\mathrm{T}}\mathcal{F}_{\mathbf{J}}(t)^{\mathrm{T}}\mathrm{d}t, \qquad \mathcal{P}_{\mathrm{i}} := \sum_{k=0}^{\nu-1}\mathcal{F}_{\mathbf{N}}(k)\mathcal{B}\mathcal{B}^{\mathrm{T}}\mathcal{F}_{\mathbf{N}}(k)^{\mathrm{T}}. \tag{2.17}$$
$$\diamondsuit$$

The ranges of the matrices $\mathcal{P}_{\mathrm{p}}$ and $\mathcal{P}_{\mathrm{i}}$ provide the controllability spaces associated with the states $\mathbf{z}_{\mathrm{p}}(t)$ and $\mathbf{z}_{\mathrm{i}}(t)$, respectively. Furthermore, inserting the definitions of $\mathcal{F}_{\mathbf{J}}(t)$ and $\mathcal{F}_{\mathbf{N}}(k)$ into (2.5) yields

$$\mathcal{P}_{\mathrm{p}} := \mathbf{T}^{-1}\begin{bmatrix}\mathbf{P}_1 & 0\\ 0 & 0\end{bmatrix}\mathbf{T}^{-\mathrm{T}}, \qquad \mathcal{P}_{\mathrm{i}} := \mathbf{T}^{-1}\begin{bmatrix}0 & 0\\ 0 & \mathbf{P}_2\end{bmatrix}\mathbf{T}^{-\mathrm{T}} \tag{2.18}$$

where $\mathbf{P}_1 := \int_0^\infty e^{\mathbf{J}t}\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_1^{\mathrm{T}}e^{\mathbf{J}^{\mathrm{T}}t}\mathrm{d}t$ and $\mathbf{P}_2 := \sum_{k=0}^{\nu-1}\mathbf{N}^k\widetilde{\mathbf{B}}_2\widetilde{\mathbf{B}}_2^{\mathrm{T}}(\mathbf{N}^k)^{\mathrm{T}}$ are the controllability Gramians corresponding to the states in (2.12) with matrices from (2.9). Using the controllability Gramians, we can characterize the hard-to-reach or unreachable states that play an important role in the reduction of the system. To compute the Gramians we use that $\boldsymbol{\mathcal{P}}_\mathrm{p}$ and $\boldsymbol{\mathcal{P}}_\mathrm{i}$ defined in (2.5) are the unique solutions of the following projected continuous-time and discrete-time projected Lyapunov equations

$$\begin{aligned} \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}_\mathrm{p}\boldsymbol{\mathcal{A}}^\mathrm{T} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}_\mathrm{p}\boldsymbol{\mathcal{E}}^\mathrm{T} &= -\mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^\mathrm{T}\mathbf{P}_\mathrm{l}^\mathrm{T}, & \boldsymbol{\mathcal{P}}_\mathrm{p} &= \mathbf{P}_\mathrm{r}\boldsymbol{\mathcal{P}}_\mathrm{p}\mathbf{P}_\mathrm{r}^\mathrm{T}, \\ \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}_\mathrm{i}\boldsymbol{\mathcal{A}}^\mathrm{T} - \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}_\mathrm{i}\boldsymbol{\mathcal{E}}^\mathrm{T} &= (\mathbf{I}-\mathbf{P}_\mathrm{l})\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^\mathrm{T}(\mathbf{I}-\mathbf{P}_\mathrm{l})^\mathrm{T}, & 0 &= \mathbf{P}_\mathrm{r}\boldsymbol{\mathcal{P}}_\mathrm{i}\mathbf{P}_\mathrm{r}^\mathrm{T}. \end{aligned} \tag{2.19}$$

To describe the observability behavior of the DAE system (2.8), we derive the corresponding state-to-output mappings

$$\boldsymbol{o}_\mathrm{p}(t) = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{F}}_\mathbf{J}(t) \qquad \text{and} \qquad \boldsymbol{o}_\mathrm{i}(k) = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{F}}_\mathbf{N}(k)$$

that are used to derive the respective observability Gramians by integration over the entire time domain and summation over all indices.

Definition 2.10:

Consider the C-stable system (2.8). The corresponding *proper and improper observability Gramians* are defined as

$$\boldsymbol{\mathcal{Q}}_\mathrm{L,p} := \int_0^\infty \boldsymbol{\mathcal{F}}_\mathbf{J}(t)^\mathrm{T}\boldsymbol{\mathcal{C}}^\mathrm{T}\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{F}}_\mathbf{J}(t)\mathrm{d}\tau, \qquad \boldsymbol{\mathcal{Q}}_\mathrm{L,i} := \sum_{k=0}^{\nu-1}\boldsymbol{\mathcal{F}}_\mathbf{N}(k)^\mathrm{T}\boldsymbol{\mathcal{C}}^\mathrm{T}\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{F}}_\mathbf{N}(k). \tag{2.20}$$

◊

We insert the definitions of $\boldsymbol{\mathcal{F}}_\mathbf{J}$ and $\boldsymbol{\mathcal{F}}_\mathbf{N}$ from (2.13) to derive

$$\boldsymbol{\mathcal{Q}}_\mathrm{L,p} := \mathbf{W}^{-\mathrm{T}}\begin{bmatrix}\mathbf{Q}_\mathrm{L,1} & 0 \\ 0 & 0\end{bmatrix}\mathbf{W}^{-1}, \qquad \boldsymbol{\mathcal{Q}}_\mathrm{L,i} := \mathbf{W}^{-\mathrm{T}}\begin{bmatrix}0 & 0 \\ 0 & \mathbf{Q}_\mathrm{L,2}\end{bmatrix}\mathbf{W}^{-1}, \tag{2.21}$$

where $\mathbf{Q}_\mathrm{L,2} = \int_0^\infty e^{\mathbf{J}^\mathrm{T}t}\widetilde{\mathbf{C}}_1^\mathrm{T}\widetilde{\mathbf{C}}_1 e^{\mathbf{J}t}\mathrm{d}t$ and $\mathbf{Q}_\mathrm{L,2} = \sum_{k=0}^{\nu-1}(\mathbf{N}^k)^\mathrm{T}\widetilde{\mathbf{C}}_2^\mathrm{T}\widetilde{\mathbf{C}}_2\mathbf{N}^k$, with matrices from (2.9), are the observability Gramians corresponding to the states $\mathbf{z}_1(t)$ and $\mathbf{z}_2(t)$ defined in (2.12). This equation describes the connection between the Gramians of the subsystems in the Weierstraß-canonical form corresponding to the states $\mathbf{z}_1(t)$ and $\mathbf{z}_2(t)$ and the Gramians corresponding to the original state spaces.

To compute the Gramians $\boldsymbol{\mathcal{Q}}_\mathrm{L,p}$ and $\boldsymbol{\mathcal{Q}}_\mathrm{L,i}$ we utilize that they are the unique solutions of the following continuous-time and discrete-time projected Lyapunov equations

$$\begin{aligned} \boldsymbol{\mathcal{E}}^\mathrm{T}\boldsymbol{\mathcal{Q}}_\mathrm{L,p}\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{A}}^\mathrm{T}\boldsymbol{\mathcal{Q}}_\mathrm{L,p}\boldsymbol{\mathcal{E}} &= -\mathbf{P}_\mathrm{r}^\mathrm{T}\boldsymbol{\mathcal{C}}^\mathrm{T}\boldsymbol{\mathcal{C}}\mathbf{P}_\mathrm{r}, & \boldsymbol{\mathcal{Q}}_\mathrm{L,p} &= \mathbf{P}_\mathrm{L}^\mathrm{T}\boldsymbol{\mathcal{Q}}_\mathrm{L,p}\mathbf{P}_\mathrm{L}, \\ \boldsymbol{\mathcal{A}}^\mathrm{T}\boldsymbol{\mathcal{Q}}_\mathrm{i}\boldsymbol{\mathcal{A}} - \boldsymbol{\mathcal{E}}^\mathrm{T}\boldsymbol{\mathcal{Q}}_\mathrm{L,i}\boldsymbol{\mathcal{E}} &= (\mathbf{I}-\mathbf{P}_\mathrm{r})^\mathrm{T}\boldsymbol{\mathcal{C}}^\mathrm{T}\boldsymbol{\mathcal{C}}(\mathbf{I}-\mathbf{P}_\mathrm{r}), & 0 &= \mathbf{P}_\mathrm{L}^\mathrm{T}\boldsymbol{\mathcal{Q}}_\mathrm{L,i}\mathbf{P}_\mathrm{L}. \end{aligned} \tag{2.22}$$

As in the ODE case, the Gramians encode the reachability and observability behavior as stated in the following theorem from [129].

Theorem 2.11:

Consider a C-stable DAE system of the form (2.8). Then the following equivalences hold.

a) The system is C-controllable, if and only if $\boldsymbol{\mathcal{P}}_\mathrm{p} + \boldsymbol{\mathcal{P}}_\mathrm{i}$ is positive definite.

b) The system is C-observable, if and only if $\boldsymbol{\mathcal{Q}}_\mathrm{L,p} + \boldsymbol{\mathcal{Q}}_\mathrm{L,i}$ is positive definite. ◊

2.1.3 Second-order ODE systems

Finally, we consider the second-order system

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0, \\ \mathbf{y}_{\mathrm{L}}(t) = \mathbf{C}_1\mathbf{x}(t) + \mathbf{C}_2\dot{\mathbf{x}}(t) \end{aligned} \tag{2.23}$$

with a mass matrix $\mathbf{M} \in \mathbb{R}^{n\times n}$, a damping matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$, a stiffness matrix $\mathbf{K} \in \mathbb{R}^{n\times n}$, an input matrix $\mathbf{B} \in \mathbb{R}^{n\times m}$, and output matrices $\mathbf{C}_1$, $\mathbf{C}_2 \in \mathbb{R}^{p\times n}$. We assume that the matrices $\mathbf{M}$, $\mathbf{D}$, and $\mathbf{K}$ are symmetric and positive semi definite, so that the state equation in (2.23) is an ODE. The input, the state, and the output are given as $\mathbf{u}(t) \in \mathbb{R}^m$, $\mathbf{x}(t) \in \mathbb{R}^n$, and $\mathbf{y}_{\mathrm{L}}(t) \in \mathbb{R}^p$, respectively.

One possible way to handle second-order systems is to transform them into first-order systems of the form (2.1) with first-order matrices

$$\mathcal{E} := \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{M} \end{bmatrix}, \quad \mathcal{A} := \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{K} & -\mathbf{D} \end{bmatrix}, \quad \mathcal{B} := \begin{bmatrix} 0 \\ \mathbf{B} \end{bmatrix}, \quad \text{and} \quad \mathcal{C} := \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix}. \tag{2.24}$$

Then, the respective first-order system has the same input-to-output behavior as the second-order system and hence can be analyzed instead. The disadvantage of the first-order representation is that the second-order structure, which characterizes the physical properties, is not retained. Therefore, in this subsection, we repeat the system theoretical results for second-order systems introduced in [43, 44, 112]. According to controllability and observability for first-order ODE systems, introduced in Section 2.1.1, we define the concepts of asymptotic stability, controllability, and observability as introduced in [112]. Those are equivalent to the asymptotic stability, controllability, and observability of their first-order representation with matrices (2.24).

Definition 2.12:
The system (2.23) is called

1. *asymptotically stable* if all zeros of the matrix polynomial $\lambda^2\mathbf{M} + \lambda\mathbf{D} + \mathbf{K}$ lie in the open-left half-plane $\mathbb{C}^- := \{\lambda \in \mathbb{C} \mid \mathrm{Re}(\lambda) < 0\}$.

2. *controllable* if
$$\operatorname{rank}\left(\begin{bmatrix} \lambda^2\mathbf{M} + \lambda\mathbf{D} + \mathbf{K} & \mathbf{B} \end{bmatrix}\right) = n \qquad \text{for all} \quad \lambda \in \mathbb{C}.$$

3. *observable* if
$$\operatorname{rank}\left(\begin{bmatrix} \lambda^2\mathbf{M}^{\mathrm{T}} + \lambda\mathbf{D}^{\mathrm{T}} + \mathbf{K}^{\mathrm{T}} & \mathbf{C}_1^{\mathrm{T}} + \lambda\mathbf{C}_2^{\mathrm{T}} \end{bmatrix}\right) = n \qquad \text{for all} \quad \lambda \in \mathbb{C}. \qquad \diamondsuit$$

To investigate the input-to-output behavior of the system and the respective controllability and observability properties, we apply the Laplace transform to the homogenous system (2.23), i.e., we set $\mathbf{x}(0) = 0$, $\dot{\mathbf{x}}(0) = 0$, which yields

$$\mathbf{Y}_{\mathrm{L}}(s) = \mathbf{C}_1\mathbf{X}(s) + \mathbf{C}_2\dot{\mathbf{X}}(s) = (\mathbf{C}_1 + s\mathbf{C}_2)\boldsymbol{\Lambda}(s)\mathbf{B}\mathbf{U}(s)$$

where $\boldsymbol{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$. The corresponding transfer function that encodes the input-to-output mapping is extracted and defined in the following.

Definition 2.13:
Consider the second-order system (2.1.3). Its *transfer function* is defined as

$$\mathcal{G}_{\mathrm{L}}(s) := (\mathbf{C}_1 + s\mathbf{C}_2)\boldsymbol{\Lambda}(s)\mathbf{B} \tag{2.25}$$

where $\boldsymbol{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$. ◊

To describe the system properties that result from that transfer function, we can apply the system theoretical concepts introduced in Definition 2.4 and Proposition 2.6.

We can derive systems Gramians tailored for systems of second-order structures describing the controllability and observability properties. First, to describe the controllability behavior, we introduce the input-to-state mappings in the frequency domain corresponding to the displacement (position) and to the velocity, which are

$$\mathcal{C}_{\mathrm{pos}}(s) = \boldsymbol{\Lambda}(s)\mathbf{B} \qquad \text{and} \qquad \mathcal{C}_{\mathrm{vel}}(s) = s\boldsymbol{\Lambda}(s)\mathbf{B}.$$

As described in [43, 112], we can derive the respective second-order controllability Gramians as introduced in the following.

Definition 2.14:
Consider the asymptotically stable system (2.23) and define $\boldsymbol{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$. Then the respective *position* and *velocity controllability Gramians* are defined as

$$\mathbf{P}_{\mathrm{pos}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \boldsymbol{\Lambda}(\mathrm{i}\omega)\mathbf{B}\mathbf{B}^{\mathrm{H}}\boldsymbol{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega, \qquad \mathbf{P}_{\mathrm{vel}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \omega^2\boldsymbol{\Lambda}(\mathrm{i}\omega)\mathbf{B}\mathbf{B}^{\mathrm{H}}\boldsymbol{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega. \tag{2.26}$$

◊

Since we consider second-order Gramians, the methods from Section 2.1 can not be applied to compute them. However, one can show that $\mathbf{P}_{\mathrm{pos}}$ and $\mathbf{P}_{\mathrm{vel}}$ are the upper-left and the lower-right block, respectively, of the first-order controllability Gramian $\mathcal{P}$ as defined in (2.5) with matrices as introduced in (2.24), see [44].

To derive the second-order observability Gramians, we extract the state-to-output mappings from $\mathcal{G}_{\mathrm{L}}$ as defined in (2.25)

$$\mathcal{O}_{\mathrm{pos}}(s) = \mathbf{C}_1\boldsymbol{\Lambda}(s)(s\mathbf{M} + \mathbf{D}) - \mathbf{C}_2\boldsymbol{\Lambda}(s)\mathbf{K}, \qquad \mathcal{O}_{\mathrm{vel}}(s) = (\mathbf{C}_1 + s\mathbf{C}_2)\boldsymbol{\Lambda}(s).$$

These mappings are now used to define the respective second-order Gramians by integrating over the entire frequency domain, which leads to the following definition.

Definition 2.15:
Consider the asymptotically stable system (2.23) and define $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$. Then the respective *position* and *velocity observability Gramians* are defined as

$$\begin{aligned}\mathbf{Q}_{\mathrm{L,pos}} &= \frac{1}{2\pi}\int_{-\infty}^{\infty}((\mathrm{i}\omega\mathbf{M}+\mathbf{D})^{\mathrm{H}}\mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathbf{C}_1^{\mathrm{T}} - \mathbf{K}\mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathbf{C}_2^{\mathrm{T}}) \\ &\qquad\qquad \cdot (\mathbf{C}_1\mathbf{\Lambda}(\mathrm{i}\omega)(\mathrm{i}\omega\mathbf{M}+\mathbf{D}) - \mathbf{C}_2\mathbf{\Lambda}(\mathrm{i}\omega)\mathbf{K})\mathrm{d}\omega, \qquad (2.27) \\ \mathbf{Q}_{\mathrm{L,vel}} &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}(\mathbf{C}_1+\mathrm{i}\omega\mathbf{C}_2)^{\mathrm{H}}(\mathbf{C}_1+\mathrm{i}\omega\mathbf{C}_2)\mathbf{\Lambda}(\mathrm{i}\omega)\mathrm{d}\omega.\end{aligned}$$

◊

These Gramians are the upper-left and the lower-right block of the first-order observability Gramian $\mathcal{Q}_{\mathrm{L}}$ from (2.5) with matrices as defined in (2.24), see again [44].

2.2 Model order reduction methods

Engineering applications such as modeling electrical circuits, structural dynamics, vibration analysis, thermal and diffusion processes, or multibody systems lead to different types of dynamical systems. Models that exhibit complex dynamic behavior or are derived from the discretization of PDEs are often high-fidelity models, i.e., the dimension of the state vector is large, lead to computationally expensive engineering design processes. As a remedy, we seek to employ model reduction techniques that allow us to construct a low-dimensional model that closely resembles the dynamic behaviors of the high-fidelity model. We present some well-established model order reduction techniques for homogeneous systems as considered in Section 2.1. There are several classes of methods for reducing the order of a model.

For first-order ODE systems (2.1), examples include singular value-based approaches such as balanced truncation [26, 93, 138] and Hankel norm approximations [56]. In addition, there are Krylov subspace-based methods, such as the iterative rational Krylov algorithm (IRKA) [26, 51, 60, 61] and moment matching, as well as data-driven methods such as the Loewner framework [57, 90]. A comprehensive overview of these methods can be found, i.e., in [5, 22, 23, 26].

The methods presented above treat systems in which $\mathcal{E}$ is nonsingular and is therefore not directly applicable to the DAE case introduced in (2.8). Several challenges arise due to the algebraic equations. Since the matrix $\mathbf{E}$ is singular, the transfer function $\mathcal{G}_{\mathrm{L}}(s) := \mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}$, defining the input-to-output mapping in the frequency domain, can have a non-zero polynomial part. A model reduction scheme for DAEs must preserve the polynomial part of its transfer function when constructing a reduced-order model as addressed in, e.g., [61, 91, 130]. There exist several methods that deal with DAE systems, i.e., interpolatory projection methods [2, 3, 61] and balancing-based methods [67, 91, 130, 131]. Also, data-driven approaches have been recently extended to differential-algebraic systems, see, e. g., [6, 58, 94].

We also consider methods tailored for systems (2.23) of second-order structure. In the literature exist several methods enabling model order reduction preserving the second-order structure [43, 53]. These techniques range from balanced truncation as well as balancing related model order reduction [44, 112, 128] to moment matching approximations based on the Krylov subspace method [17, 118]. The work [115] provides a comprehensive comparison between common second-order model reduction methods applied to a large-scale mechanical fishtail model. Additionally, [18] proposed interpolation-based methods for systems possessing very general dynamical structures. More recently, the authors in [19] propose a new philosophy to find the dominant reachability and observability subspaces, enabling very accurate reduced-order models preserving the structure. Moreover, an extension of the Loewner framework was proposed in [21] for the class of Rayleigh damped systems and in [122] for general structured systems. Second-order systems were also considered in a vast amount of literature by now, where some Krylov space-based methods are derived in [9, 10, 43, 53, 117] and balancing reduction methods are introduced in [43, 44, 92, 112].

In this work, we focus on Balanced Truncation (BT) and Iterative Rational Krylov Iteration (IRKA). Both methods construct projection matrices for the reduction so that the multiplication of the system matrices by these projection matrices then yields a ROM. BT generates an $\mathcal{H}_\infty$-optimal surrogate system and has the advantage of guaranteed asymptotic stability of the ROM, the existence of an error bound, and respective numerical techniques for the Lyapunov equations involved [29, 48, 120, 126]. IRKA, on the other hand, generates an $\mathcal{H}_2$-optimal reduced surrogate system. In the following, we introduce the balanced truncation method in Section 2.2.1 and the iterative rational Krylov iteration method in Section 2.2.2 to generate the reduced surrogate systems.

2.2.1 Balanced truncation

In this subsection, we repeat the balanced truncation (BT) method. First, we explain the original method for a first-order system with an ODE as a state equation from [20, 26, 93, 138]. Afterwards, we briefly show BT for first-order systems with DAE state equations as introduced in [91, 130], and for second-order systems as shown in [112].

2.2.1.1 Balanced truncation for first-order ODE systems

We consider systems of the form (2.1) with $\mathbf{z}(0) = 0$ and aim to generate a surrogate model

$$\begin{aligned} \boldsymbol{\mathcal{E}}_\mathrm{r}\dot{\mathbf{z}}_\mathrm{r}(t) &= \boldsymbol{\mathcal{A}}_\mathrm{r}\mathbf{z}_\mathrm{r}(t) + \boldsymbol{\mathcal{B}}_\mathrm{r}\mathbf{u}(t), \qquad \mathbf{z}_\mathrm{r}(0) = 0, \\ \mathbf{y}_{\mathrm{L},\mathrm{r}}(t) &= \boldsymbol{\mathcal{C}}_\mathrm{r}\mathbf{z}_\mathrm{r}(t) \end{aligned} \tag{2.28}$$

where $\boldsymbol{\mathcal{E}}_\mathrm{r}$, $\boldsymbol{\mathcal{A}}_\mathrm{r} \in \mathbb{R}^{R\times R}$, $\boldsymbol{\mathcal{B}}_\mathrm{r} \in \mathbb{R}^{R\times m}$, and $\boldsymbol{\mathcal{C}}_\mathrm{r} \in \mathbb{R}^{p\times R}$. The reduced state and output are denoted by $\mathbf{z}_\mathrm{r}(t) \in \mathbb{R}^R$ and $\mathbf{y}_{\mathrm{L},\mathrm{r}}(t) \in \mathbb{R}^p$, respectively.

To derive such a reduced system, we use projecting matrices $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r}} \in \mathbb{R}^{N\times R}$ so that

$$\boldsymbol{\mathcal{E}}_{\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}, \qquad \boldsymbol{\mathcal{A}}_{\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}, \qquad \boldsymbol{\mathcal{B}}_{\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}, \qquad \boldsymbol{\mathcal{C}}_{\mathrm{r}} = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}. \tag{2.29}$$

We aim to find such projecting matrices defining a surrogate model that satisfy the Petrov-Galerkin orthogonality conditions, where $\boldsymbol{\mathcal{T}}_{\mathrm{r}}$ approximates the space of reachable states, i.e., for every $\mathbf{z}(t)$ generated by system (2.1) there exists a

$$\mathbf{z}_{\mathrm{r}}(t) \in \mathbb{R}^{R} \qquad \text{with} \qquad \mathbf{z}(t) \approx \boldsymbol{\mathcal{T}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t). \tag{2.30}$$

This approximation defines the residual $\boldsymbol{\mathfrak{R}}(\mathbf{z}_{\mathrm{r}}(t)) := \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}\dot{\mathbf{z}}_{\mathrm{r}}(t) - \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t) - \boldsymbol{\mathcal{B}}_{\mathrm{r}}\mathbf{u}(t)$. The Petrov-Galerkin condition then imposes that $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$ is chosen so that

$$\boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathfrak{R}}(\mathbf{z}_{\mathrm{r}}(t)) = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\left(\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}\dot{\mathbf{z}}_{\mathrm{r}}(t) - \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t) - \boldsymbol{\mathcal{B}}_{\mathrm{r}}\mathbf{u}(t)\right) = 0. \tag{2.31}$$

We want to build the projecting matrices $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r}} \in \mathbb{R}^{N\times R}$ in such a way, that their dimension R is significantly smaller than the original dimension N, i.e. $R \ll N$, and so that the input-to-output behavior is well-approximated, that means that $\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|$ is small in a suitable norm. The main idea of BT is to truncate states of the systems that are simultaneously hard to reach and to observe to obtain surrogate models (2.28) of significantly smaller dimensions.

We derive energy functionals that indicate the controllability and observability properties of the states in (2.2). First, we define the input energy corresponding to an input $\mathbf{u} \in L_2((-\infty, 0], \mathbb{R}^m)$ that is

$$E_{\mathbf{u}} := \int_{-\infty}^{0} \|\mathbf{u}(t)\|_2^2 \mathrm{d}t = \|\mathbf{u}\|^2_{L_2((-\infty,0],\mathbb{R}^m)}.$$

We evaluate the minimal amount of energy needed to reach a state $\mathbf{z}(0) = \mathbf{z}_0$ starting from $\mathbf{z}(-\infty) = 0$ which is equal to

$$E_{\mathbf{u}}(\mathbf{z}_0) = \inf_{\substack{\widetilde{\mathbf{u}} \in L_2((-\infty,0],\mathbb{R}^m) \\ \mathbf{z}(-\infty)=0,\ \mathbf{z}(0)=\mathbf{z}_0}} \int_{-\infty}^{0} \|\widetilde{\mathbf{u}}(t)\|_2^2 \mathrm{d}t.$$

The following lemma from [4, Lemma 4.29] describes how the $E_{\mathbf{u}}(\mathbf{z}_0)$ is computed.

Lemma 2.16:
Consider the asymptotically stable system (2.1) with zero initial conditions. The minimal energy needed to reach a state $\mathbf{z}_0 \in \mathbb{R}^N$ is equal to

$$E_{\mathbf{u}}(\mathbf{z}_0) = \mathbf{z}_0^{\mathrm{T}}\boldsymbol{\mathcal{P}}^{-1}\mathbf{z}_0 \tag{2.32}$$

with $\mathbf{u}(t) = \boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{-\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-T}t}\boldsymbol{\mathcal{P}}^{-1}\mathbf{z}_0$ and $\boldsymbol{\mathcal{P}}$ as defined in (2.5). ◊

From (2.32), it follows that states $\mathbf{z}_0$ corresponding to small eigenvalues of $\boldsymbol{\mathcal{P}}$ are harder to reach since more energy $E_{\mathbf{u}}(\mathbf{z}_0)$ is needed to attain them.

Moreover, we evaluate the output energy of the system (2.1) corresponding to an output function $\mathbf{y}_{\mathrm{L}} \in L_2\left([0,\infty), \mathbb{R}^p\right)$ that is defined as

$$E_{\mathbf{y}_{\mathrm{L}}} := \int_0^\infty \|\mathbf{y}_{\mathrm{L}}(t)\|_2^2 \mathrm{d}t = \|\mathbf{y}_{\mathrm{L}}\|^2_{L_2([0,\infty),\mathbb{R}^p)}.$$

We denote the energy generated by system (2.1) with an initial state $\mathbf{z}(0) = \mathbf{z}_0$ and no input, i.e., $\mathbf{u} \equiv 0$, by

$$E_{\mathbf{y}_{\mathrm{L}}}(\mathbf{z}_0) = \|\mathbf{y}_{\mathrm{L}}\|^2_{L_2([0,\infty),\mathbb{R}^p)}.$$

Again, the lemma from [4, Lemma 4.29] is used to compute the energy $E_{\mathbf{y}_{\mathrm{L}}}(\mathbf{z}_0)$.

Lemma 2.17:
Consider the asymptotically stable system (2.1) with zero input $\mathbf{u} \equiv 0$. The energy generated by the system with an initial state $\mathbf{z}(0) = \mathbf{z}_0$ is equal to

$$E_{\mathbf{y}_{\mathrm{L}}}(\mathbf{z}_0) = \mathbf{z}_0^{\mathrm{T}} \boldsymbol{\mathcal{E}}^{\mathrm{T}} \boldsymbol{\mathcal{Q}}_{\mathrm{L}} \boldsymbol{\mathcal{E}} \mathbf{z}_0 \tag{2.33}$$

where the output is $\mathbf{y}_{\mathrm{L}}(\cdot) = \boldsymbol{\mathcal{C}} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}(\cdot)} \mathbf{z}_0$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ as defined in (2.5). ◊

It follows that states corresponding to small singular values of the observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ lead to small amounts of energies that can be observed and are, therefore, neglectable. These states are truncated in the following. However, the controllability Gramian $\boldsymbol{\mathcal{P}}$ and the observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ are, in general, not equal, and hence, the states that are hard to reach are not necessarily hard to observe and vice versa. Therefore, we balanced the system so that the Gramians coincide.

Definition 2.18:
Consider the asymptotically stable dynamical system (2.1) with the controllability Gramian $\boldsymbol{\mathcal{P}}$ and observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ as defined in (2.5). Then the dynamical system is called *balanced* if the corresponding Gramians are equal, i.e., it holds

$$\boldsymbol{\mathcal{P}} = \boldsymbol{\mathcal{Q}}_{\mathrm{L}} = \boldsymbol{\Sigma},$$

where $\boldsymbol{\Sigma} = \operatorname{diag}\left(\sigma_1, \ldots, \sigma_N\right)$ is a diagonal matrix with $\sigma_1 \geq \cdots \geq \sigma_N$. ◊

We can balance the system by applying simple transformations that generate an equivalent system, i.e., it has the same input-to-output behavior as the systems in (2.1). The transformed Gramians then coincide and are even diagonal matrices. For that, assume that $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{S}}$ are Cholesky factors (or if available low-rank factors) of the Gramians of our original system in (2.1), i.e. $\boldsymbol{\mathcal{P}} = \boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}^{\mathrm{T}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{L}} = \boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^{\mathrm{T}}$. We compute the following singular value decomposition (SVD)

$$\boldsymbol{\mathcal{S}}^{\mathrm{T}} \boldsymbol{\mathcal{E}} \boldsymbol{\mathcal{R}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_1 & 0 \\ 0 & \boldsymbol{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^{\mathrm{T}} \\ \mathbf{V}_2^{\mathrm{T}} \end{bmatrix}.$$

The matrix $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \dots, \sigma_N)$ contains the so-called *Hankel singular values* in decreasing order, i.e. $\sigma_1 \geq \dots \geq \sigma_N$. The transformation matrices

$$\boldsymbol{\mathcal{V}}_{\mathrm{b}} = \boldsymbol{\mathcal{S}}\mathbf{U}\boldsymbol{\Sigma}^{-\frac{1}{2}}, \qquad \boldsymbol{\mathcal{T}}_{\mathrm{b}} = \boldsymbol{\mathcal{R}}\mathbf{V}\boldsymbol{\Sigma}^{-\frac{1}{2}} \tag{2.34}$$

satisfy $\boldsymbol{\mathcal{V}}_{\mathrm{b}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{b}} = \mathbf{I}_N$ and generate the transformed system

$$\begin{aligned}\dot{\mathbf{z}}_{\mathrm{b}}(t) &= \boldsymbol{\mathcal{V}}_{\mathrm{b}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{b}}\mathbf{z}_{\mathrm{b}}(t) + \boldsymbol{\mathcal{V}}_{\mathrm{b}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}\mathbf{u}(t),\\ \mathbf{y}_{\mathrm{L}}(t) &= \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{T}}_{\mathrm{b}}\mathbf{z}_{\mathrm{b}}(t)\end{aligned}$$

with new Gramians

$$\boldsymbol{\mathcal{P}}_{\mathrm{b}} = \boldsymbol{\mathcal{Q}}_{\mathrm{L,b}} = \boldsymbol{\Sigma}.$$

The remaining step is to truncate states corresponding to small singular values of $\boldsymbol{\Sigma}$. For that, we build projecting matrices

$$\boldsymbol{\mathcal{V}}_{\mathrm{r}} = \boldsymbol{\mathcal{S}}\mathbf{U}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}, \qquad \boldsymbol{\mathcal{T}}_{\mathrm{r}} = \boldsymbol{\mathcal{R}}\mathbf{V}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}} \tag{2.35}$$

that project the system onto the state spaces spanned by $\mathbf{U}_1$ and $\mathbf{V}_1$ corresponding to the largest singular values stored in $\boldsymbol{\Sigma}_1$. Multiplying the original system in (2.1) by $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r}}$ results in the reduced system in (2.28) with the reduced matrices defined in (2.29) and $\boldsymbol{\mathcal{E}}_{\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{r}} = \mathbf{I}_R$. This method results in Algorithm 1.

There exists an error bound, described, e.g., in [4], that quantifies the error in the output of the reduced system, i.e., the error between $\mathbf{y}_{\mathrm{L}}$ and $\mathbf{y}_{\mathrm{L,r}}$ that is

$$\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_2} \leq \|\boldsymbol{\mathcal{G}}_{\mathrm{L}} - \boldsymbol{\mathcal{G}}_{\mathrm{L,r}}\|_{\mathcal{H}_\infty}\|\mathbf{u}\|_{L_2} \leq \left(2\sum_{k=R+1}^{N}\sigma_k\right)\|\mathbf{u}\|_{L_2} \tag{2.36}$$

where $\boldsymbol{\mathcal{G}}_{\mathrm{L}}$ and $\boldsymbol{\mathcal{G}}_{\mathrm{L,r}}$ are the transfer functions of the original and the reduced system (2.1) and (2.28), respectively.

2.2.1.2 Balanced truncation for first-order DAE systems

We consider systems of the form (2.8) and aim to generate a surrogate model

$$\begin{aligned}\boldsymbol{\mathcal{E}}_{\mathrm{r}}\dot{\mathbf{z}}_{\mathrm{r}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t) + \boldsymbol{\mathcal{B}}_{\mathrm{r}}\mathbf{u}(t), \qquad \mathbf{z}_{\mathrm{r}}(0) = \mathbf{z}_{\mathrm{r},0},\\ \mathbf{y}_{\mathrm{L,r}}(t) &= \boldsymbol{\mathcal{C}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t),\end{aligned} \tag{2.37}$$

where $\boldsymbol{\mathcal{E}}_{\mathrm{r}}$, $\boldsymbol{\mathcal{A}}_{\mathrm{r}} \in \mathbb{R}^{R\times R}$, $\boldsymbol{\mathcal{B}}_{\mathrm{r}} \in \mathbb{R}^{R\times m}$, and $\boldsymbol{\mathcal{C}}_{\mathrm{r}} \in \mathbb{R}^{p\times R}$. Also the reduced state and output are $\mathbf{z}_{\mathrm{r}}(t) \in \mathbb{R}^R$ and $\mathbf{y}_{\mathrm{L}}(t) \in \mathbb{R}^p$, respectively, and the initial state is $\mathbf{z}_{\mathrm{r},0} \in \mathbb{R}^R$ satisfying the consistency conditions (2.15) . We generate the reduced matrices as described in (2.29) using projecting matrices $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r}} \in \mathbb{R}^{N\times R}$. To generate these projecting matrices, we follow the method introduced in [91, 130] and investigate the

Algorithm 1 BT method for the first-order ODE system (2.1).

```
Require: The original system (2.1) and the reduced order $R$.
Ensure: The reduced system (2.28).
1: Compute factors of the Gramians $\mathcal{P} \approx \mathcal{R}\mathcal{R}^\mathrm{T}$ and $\mathcal{Q}_\mathrm{L} \approx \mathcal{S}\mathcal{S}^\mathrm{T}$ from Definition (2.5).
2: Perform the SVD of $\mathcal{S}^\mathrm{T}\mathcal{E}\mathcal{R}$, and decompose as
       $$\mathcal{S}^\mathrm{T}\mathcal{E}\mathcal{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\mathrm{T} = \begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_1 & 0 \\ 0 & \boldsymbol{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_1^\mathrm{T} \\ \mathbf{V}_2^\mathrm{T}\end{bmatrix}.$$
   with $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{R\times R}$.
3: Construct the projection matrices
       $$\mathcal{V}_\mathrm{r} = \mathcal{S}\mathbf{U}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}, \qquad \mathcal{T}_\mathrm{r} = \mathcal{R}\mathbf{V}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}.$$
4: Determine the reduced matrices (2.29) of the reduced system (2.28).
```

proper components of the system and the improper ones separately. The goal is to reduce the differential parts of the system as they correspond to an ODE in WCF (2.9). Since the algebraic components encode algebraic constraints, reducing those could generate results that are physically difficult to interpret. Hence, we aim to find a minimal realization corresponding to the improper components.

To reduce the differential parts of the system, we consider the energy functional of the proper component of the DAE system in WCF (2.11). Since a differential state can be written as $\mathbf{z}_\mathrm{p}^* = \mathbf{T}^{-1}\left[\begin{smallmatrix}\mathbf{z}_1^* \\ 0\end{smallmatrix}\right]$, where $\mathbf{z}_1^*$ is a proper state from (2.11), the energy needed to reach a differential state $\mathbf{z}_\mathrm{p}^*$ is

$$E_\mathbf{u} = (\mathbf{z}_1^*)^\mathrm{T}\mathbf{P}_1^{-1}\mathbf{z}_1^* = \begin{bmatrix}(\mathbf{z}_1^*)^\mathrm{T} & 0\end{bmatrix}\mathbf{T}^\mathrm{T}\mathbf{T}^{-\mathrm{T}}\begin{bmatrix}\mathbf{P}_1^{-1} & 0 \\ 0 & 0\end{bmatrix}\mathbf{T}^{-1}\mathbf{T}\begin{bmatrix}\mathbf{z}_1^* \\ 0\end{bmatrix}$$

where we make use of (2.32) with $\mathbf{P}_1$ as introduced in (2.18). Hence, we obtain

$$E_\mathbf{u} = (\mathbf{z}_\mathrm{p}^*)^\mathrm{T}\mathcal{P}_\mathrm{p}^\mathbf{I}\mathbf{z}_\mathrm{p}^* \qquad \text{with} \qquad \mathcal{P}_\mathrm{p}^\mathbf{I} := \mathbf{T}^{-\mathrm{T}}\begin{bmatrix}\mathbf{P}_1^{-1} & 0 \\ 0 & 0\end{bmatrix}\mathbf{T}^{-1}. \tag{2.38}$$

It follows that proper states $\mathbf{z}_\mathrm{p}^*$ corresponding to eigenvalues of $\mathcal{P}_\mathrm{p}$ with small magnitudes, as indicated in (2.17), require large amounts of energy to be reached and are therefore truncated in the following analysis. Conversely, states corresponding to large eigenvalues are easier to attain and thus define the dominant proper controllability subspace.

To evaluate the observability behavior of system (2.8), we determine the energy generated by the system with a differential initial state $\mathbf{z}_\mathrm{p}^*$ and no input, i.e., $\mathbf{u} \equiv 0$. For that, we consider the WCF of the system, as presented in (2.11) and investigate the energy corresponding to the proper state $\mathbf{z}_1^*$ as shown in (2.33), which yields

$$E_{\mathbf{y}_\mathrm{L}} = (\mathbf{z}_1^*)^\mathrm{T}\mathbf{Q}_{\mathrm{L},1}\mathbf{z}_1^* = (\mathbf{z}_\mathrm{p}^*)^\mathrm{T}\mathbf{T}^\mathrm{T}\begin{bmatrix}\mathbf{I} & 0\\ 0 & \mathbf{N}^\mathrm{T}\end{bmatrix}\mathbf{W}^\mathrm{T}\mathbf{W}^{-\mathrm{T}}\begin{bmatrix}\mathbf{Q}_1 & 0\\ 0 & 0\end{bmatrix}\mathbf{W}^{-1}\mathbf{W}\begin{bmatrix}\mathbf{I} & 0\\ 0 & \mathbf{N}\end{bmatrix}\mathbf{T}\mathbf{z}_\mathrm{p}^*$$

with $\mathbf{Q}_{\mathrm{L},1}$ as defined in (2.21) and the proper state as $\mathbf{z}_\mathrm{p}^* = \mathbf{T}^{-1}\left[\begin{smallmatrix}\mathbf{z}_1^*\\ 0\end{smallmatrix}\right]$. From (2.21), it follows, that

$$E_{\mathbf{y}_\mathrm{L}} = (\mathbf{z}_\mathrm{p}^*)^\mathrm{T}\boldsymbol{\mathcal{E}}^\mathrm{T}\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}\boldsymbol{\mathcal{E}}\mathbf{z}_\mathrm{p}^* \tag{2.39}$$

for $\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$ as defined in (2.20). Hence, proper states corresponding to small eigenvalues of $\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$ generate small amounts of output energy and are hard to observe, while states corresponding to large eigenvalues are easy to observe and span the dominant observability subspaces.

In general, the states corresponding to small eigenvalues of the proper controllability Gramian $\boldsymbol{\mathcal{P}}_\mathrm{p}$ do not coincide with those corresponding to small eigenvalues of the proper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$. Therefore, we need to balance the system as in the previous paragraph, i.e., generate an equivalent system for which the controllability Gramians and the observability Gramians coincide.

Definition 2.19:
Consider the C-stable system in (2.8), the corresponding proper and improper controllability Gramians $\boldsymbol{\mathcal{P}}_\mathrm{p}$ and $\boldsymbol{\mathcal{P}}_\mathrm{i}$ as defined in (2.18), and the proper and improper observability Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}$ from (2.21). We call the system *balanced* if the Gramians fulfill

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

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

We follow the methodology presented in [91] to derive a balanced and truncated system. Since all Gramians are symmetric and positive semi-definite, there exist factorizations

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

We compute the singular value decompositions

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

where $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \dots, \sigma_n)$, $\sigma_1 \geq \dots \geq \sigma_n$ 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 $\mathbf{\Sigma}_2$. We truncate the corresponding states that lie in the spaces spanned by $\mathbf{U}_{\mathrm{p},2}$ and $\mathbf{V}_{\mathrm{p},2}$ by building the projection matrices

$$\mathcal{V}_{\mathrm{r}} = \begin{bmatrix} \mathcal{S}_{\mathrm{p}}^{\mathrm{T}} \mathbf{U}_{\mathrm{p},1} \mathbf{\Sigma}_1^{-\frac{1}{2}} & \mathcal{S}_{\mathrm{i}}^{\mathrm{T}} \mathbf{U}_{\mathrm{i},1} \mathbf{\Theta}_1^{-\frac{1}{2}} \end{bmatrix}, \qquad \mathcal{T}_{\mathrm{r}} = \begin{bmatrix} \mathcal{R}_{\mathrm{p}} \mathbf{V}_{\mathrm{p},1} \mathbf{\Sigma}_1^{-\frac{1}{2}} & \mathcal{R}_{\mathrm{i}} \mathbf{V}_{\mathrm{i},1} \mathbf{\Theta}_1^{-\frac{1}{2}} \end{bmatrix}. \tag{2.40}$$

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 system in (2.8) with singular $\mathcal{E}$ by $\mathcal{V}_{\mathrm{r}}$ and $\mathcal{T}_{\mathrm{r}}$ leads to a reduced system (2.37) with

$$\mathcal{E}_{\mathrm{r}} = \mathcal{V}_{\mathrm{r}}^{\mathrm{T}} \mathcal{E} \mathcal{T}_{\mathrm{r}} = \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \widehat{\mathbf{E}}_2 \end{bmatrix}, \qquad \mathcal{A}_{\mathrm{r}} = \mathcal{V}_{\mathrm{r}}^{\mathrm{T}} \mathcal{A} \mathcal{T}_{\mathrm{r}} = \begin{bmatrix} \widehat{\mathbf{A}}_1 & 0 \\ 0 & \mathbf{I} \end{bmatrix}, \tag{2.41}$$

$$\mathcal{B}_{\mathrm{r}} = \mathcal{V}_{\mathrm{r}}^{\mathrm{T}} \mathcal{B} = \begin{bmatrix} \widehat{\mathbf{B}}_1 \\ \widehat{\mathbf{B}}_2 \end{bmatrix}, \qquad \mathcal{C}_{\mathrm{r}} = \mathcal{C} \mathcal{T}_{\mathrm{r}} = \begin{bmatrix} \widehat{\mathbf{C}}_1 & \widehat{\mathbf{C}}_2 \end{bmatrix} \tag{2.42}$$

and with $\widehat{\mathbf{A}}_1 \in \mathbb{R}^{R_f \times R_f}$ and $\widehat{\mathbf{E}}_2 \in \mathbb{R}^{N_\infty \times R_\infty}$ being nilpotent. Consequently, the reduced system is inherently decoupled into a proper and improper reduced state. The output error is bounded as

$$\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_2} \leq \|\mathcal{G}_{\mathrm{L}} - \mathcal{G}_{\mathrm{L,r}}\|_{\mathcal{H}_\infty} \|\mathbf{u}\|_{L_2} \leq \left(2 \sum_{k=R_f+1}^{N_f} \sigma_k \right) \|\mathbf{u}\|_{L_2}, \tag{2.43}$$

where $\mathcal{G}_{\mathrm{L}}$ is the transfer function corresponding to the original system (2.8) and $\mathcal{G}_{\mathrm{L,r}}$ is the transfer function corresponding to the reduced system (2.37).

Remark 2.20:
The BT method presented above decouples the proper and the improper states where the proper states are reduced while for the improper states, only a minimal realization is found. ♢

2.2.1.3 Balanced truncation for second-order systems

In this subsection, we describe BT for homogeneous systems of second-order structure (2.23), i.e., $\mathbf{x}(0) = 0$, $\dot{\mathbf{x}}_{\mathrm{r}}(0) = 0$, where we aim to find a second-order surrogate model of the form

$$\begin{aligned} \mathbf{M}_{\mathrm{r}} \ddot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{D}_{\mathrm{r}} \dot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{K}_{\mathrm{r}} \mathbf{x}_{\mathrm{r}}(t) &= \mathbf{B}_{\mathrm{r}} \mathbf{u}(t), \qquad \mathbf{x}_{\mathrm{r}}(0) = 0, \quad \dot{\mathbf{x}}_{\mathrm{r}}(0) = 0, \\ \mathbf{y}_{\mathrm{L,r}}(t) &= \mathbf{C}_{1,\mathrm{r}} \mathbf{x}_{\mathrm{r}}(t) + \mathbf{C}_{2,\mathrm{r}} \dot{\mathbf{x}}_{\mathrm{r}}(t) \end{aligned} \tag{2.44}$$

with reduced matrices $\mathbf{M}_{\mathrm{r}}$, $\mathbf{D}_{\mathrm{r}}$, $\mathbf{K}_{\mathrm{r}} \in \mathbb{R}^{r \times r}$, $\mathbf{B}_{\mathrm{r}} \in \mathbb{R}^{r \times m}$, $\mathbf{C}_{1,\mathrm{r}}$, $\mathbf{C}_{2,\mathrm{r}} \in \mathbb{R}^{p \times r}$, $\mathbf{x}_{\mathrm{r}} \in \mathbb{R}^{r}$, and $\mathbf{y}_{\mathrm{L,r}}(t)^{\mathrm{T}} \in \mathbb{R}^{p}$, $r \ll n$. The reduced system (2.44) shall satisfy that $\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|$ is small

Algorithm 2 BT method for the first-order DAE system (2.8).

```
Require: The original system (2.8) and the reduced order $R = R_f + R_\infty$.
Ensure: The reduced system (2.37).
1: Compute factors of the Gramians $\mathcal{P}_\mathrm{p} \approx \mathcal{R}_\mathrm{p}\mathcal{R}_\mathrm{p}^\mathrm{T}$, $\mathcal{P}_\mathrm{i} \approx \mathcal{R}_\mathrm{i}\mathcal{R}_\mathrm{i}^\mathrm{T}$ and $\mathcal{Q}_\mathrm{L,p} \approx \mathcal{S}_\mathrm{p}^\mathrm{T}\mathcal{S}_\mathrm{p}$,
   $\mathcal{Q}_\mathrm{L,i} \approx \mathcal{S}_\mathrm{i}^\mathrm{T}\mathcal{S}_\mathrm{i}$ from Definition (2.17) and (2.20).
2: Perform the two SVDs and decompose them as
   $$\mathcal{S}_\mathrm{p}\mathcal{E}\mathcal{R}_\mathrm{p} = \begin{bmatrix}\mathbf{U}_\mathrm{p,1} & \mathbf{U}_\mathrm{p,2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_1 & \\ & \boldsymbol{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_\mathrm{p,1}^\mathrm{T} \\ \mathbf{V}_\mathrm{p,2}^\mathrm{T}\end{bmatrix}, \qquad \mathcal{S}_\mathrm{i}\mathcal{A}\mathcal{R}_\mathrm{i} = \begin{bmatrix}\mathbf{U}_\mathrm{i,1} & \mathbf{U}_\mathrm{i,2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Theta}_1 & \\ & 0\end{bmatrix}\begin{bmatrix}\mathbf{V}_\mathrm{i,1}^\mathrm{T} \\ \mathbf{V}_\mathrm{i,2}^\mathrm{T}\end{bmatrix}.$$
   with $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{R_f \times R_f}$.
3: Construct the projection matrices
   $$\mathcal{V}_\mathrm{r} = \begin{bmatrix}\mathcal{S}_\mathrm{p}^\mathrm{T}\mathbf{U}_\mathrm{p,1}\boldsymbol{\Sigma}_1^{-\frac{1}{2}} & \mathcal{S}_\mathrm{i}^\mathrm{T}\mathbf{U}_\mathrm{i,1}\boldsymbol{\Theta}_1^{-\frac{1}{2}}\end{bmatrix}, \qquad \mathcal{T}_\mathrm{r} = \begin{bmatrix}\mathcal{R}_\mathrm{p}\mathbf{V}_\mathrm{p,1}\boldsymbol{\Sigma}_1^{-\frac{1}{2}} & \mathcal{R}_\mathrm{i}\mathbf{V}_\mathrm{i,1}\boldsymbol{\Theta}_1^{-\frac{1}{2}}\end{bmatrix}.$$
4: Determine the reduced matrices (2.41) of the reduced system (2.37).
```

in an appropriate norm. To derive such a system, we build the respectively reduced matrices using two projection matrices $\mathbf{T}_\mathrm{r}$, $\mathbf{W}_\mathrm{r} \in \mathbb{R}^{n\times r}$ that fulfill the Petrov-Galerkin conditions from (2.30), (2.31) so that

$$\begin{aligned} \mathbf{M}_\mathrm{r} &= \mathbf{W}_\mathrm{r}^\mathrm{T}\mathbf{M}\mathbf{T}_\mathrm{r}, & \mathbf{D}_\mathrm{r} &= \mathbf{W}_\mathrm{r}^\mathrm{T}\mathbf{D}\mathbf{T}_\mathrm{r}, & \mathbf{K}_\mathrm{r} &= \mathbf{W}_\mathrm{r}^\mathrm{T}\mathbf{K}\mathbf{T}_\mathrm{r}, \\ \mathbf{B}_\mathrm{r} &= \mathbf{W}_\mathrm{r}^\mathrm{T}\mathbf{B}, & \mathbf{C}_\mathrm{1,r} &= \mathbf{C}_1\mathbf{T}_\mathrm{r}, & \mathbf{C}_\mathrm{2,r} &= \mathbf{C}_2\mathbf{T}_\mathrm{r}. \end{aligned} \tag{2.45}$$

We want to emphasize that the reduction using the matrices $\mathbf{T}_\mathrm{r}$ and $\mathbf{W}_\mathrm{r}$ preserves the second-order structure of the system.

To identify the states that have the least influence on the system dynamics and that are truncated within this method, we derive the input energies of the second-order system (2.23). For that, we apply the theory derived in [43, 44, 112] where we consider the first-order controllability Gramian $\mathcal{P}_\mathrm{so}$ corresponding to the first-order matrices in (2.24) as introduced in (2.5) which has the upper-left block $\mathbf{P}_{11} = \mathbf{P}_\mathrm{pos}$ and the lower-right block $\mathbf{P}_{22} = \mathbf{P}_\mathrm{vel}$. We apply the Schur complement to obtain

$$\mathcal{P}_\mathrm{so}^{-1} = \begin{bmatrix}\mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^\mathrm{T} & \mathbf{P}_{22}\end{bmatrix}^{-1} = \begin{bmatrix}\mathbf{P}_{11}^{-1} + \mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{S}^{-1}\mathbf{P}_{12}^\mathrm{T}\mathbf{P}_{11}^{-1} & -\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{P}_{12}^\mathrm{T}\mathbf{P}_{11}^{-1} & \mathbf{S}^{-1}\end{bmatrix} =: \begin{bmatrix}\mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{12}^\mathrm{T} & \mathbf{R}_{22}\end{bmatrix},$$

where $\mathbf{S} := \mathbf{P}_{22} - \mathbf{P}_{12}^\mathrm{T}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}$. As shown in [44], this first-order representation is used to derive the energy needed to reach a second-order state $\mathbf{x}(0) = \mathbf{x}_0$ with a varying velocity at time zero $\dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0$, that is

$$\begin{aligned} E_{\mathbf{u}}(\mathbf{x}_0) &= \inf_{\substack{\mathbf{u}\in L_2((-\infty,0],\mathbb{R}^m),\\ \mathbf{x}(0)=\mathbf{x}_0,\mathbf{x}(-\infty)=0,\\ \dot{\mathbf{x}}(-\infty)=0}} \int_{-\infty}^{0} \|\mathbf{u}(t)\|_2^2 \mathrm{d}t = \begin{bmatrix} \mathbf{x}_0^{\mathrm{T}} & \dot{\mathbf{x}}_0^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{12}^{\mathrm{T}} & \mathbf{R}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \dot{\mathbf{x}}_0 \end{bmatrix} \\ &= \mathbf{x}_0^{\mathrm{T}}\mathbf{R}_{11}\mathbf{x}_0 + 2\dot{\mathbf{x}}_0^{\mathrm{T}}\mathbf{R}_{12}^{\mathrm{T}}\mathbf{x}_0 + \dot{\mathbf{x}}_0^{\mathrm{T}}\mathbf{R}_{22}\dot{\mathbf{x}}_0. \end{aligned}$$

Since we choose $\dot{\mathbf{x}}_0$ to be variable, we can minimize with respect to this vector which yields

$$\nabla_{\dot{\mathbf{x}}_0} E_{\mathbf{u}}(\mathbf{x}_0) = 2\mathbf{R}_{12}^{\mathrm{T}}\mathbf{x}_0 + 2\mathbf{R}_{22}\dot{\mathbf{x}}_0,$$

and hence the minimal input energy is attained for $\dot{\mathbf{x}}_0 = -\mathbf{R}_{22}^{-1}\mathbf{R}_{12}^{\mathrm{T}}\mathbf{x}_0$. This yields the energy

$$E_{\mathbf{u}}(\mathbf{x}_0) = \mathbf{x}_0^{\mathrm{T}}\mathbf{R}_{11}\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{T}}\mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{12}^{\mathrm{T}}\mathbf{x}_0 = \mathbf{x}_0^{\mathrm{T}}(\mathbf{R}_{11} - \mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{12}^{\mathrm{T}})\mathbf{x}_0.$$

Inserting now the matrices $\mathbf{R}_{11}$, $\mathbf{R}_{12}$, $\mathbf{R}_{22}$ yields

$$E_{\mathbf{u}}(\mathbf{x}_0) = \mathbf{x}_0^{\mathrm{T}}\mathbf{P}_{11}^{-1}\mathbf{x}_0 \tag{2.46}$$

where $\mathbf{P}_{11} = \mathbf{P}_{\mathrm{pos}}$ is the position controllability Gramian as defined in (2.26). The equation (2.46) describes that states corresponding to small singular values of $\mathbf{P}_{\mathrm{pos}}$ are hard to reach while states corresponding to large singular values need only little amounts of energies to be reached so that the eigenvalues of $\mathbf{P}_{\mathrm{pos}}$ describe which states to truncate in reduction methods.

To investigate the state derivative $\dot{\mathbf{x}}(t)$ in more detail, the authors in [44] determine the energy needed to reach a velocity $\dot{\mathbf{x}}_0$ at time zero for a variable displacement at time zero $\mathbf{x}_0$, which is

$$\begin{aligned} E_{\mathbf{u}}(\mathbf{x}_0) &= \inf_{\substack{\mathbf{u}\in \mathcal{L}((-\infty,0],\mathbb{R}^m),\\ \dot{\mathbf{x}}(0)=\dot{\mathbf{x}}_0,\dot{\mathbf{x}}(-\infty)=0,\\ \mathbf{x}(-\infty)=0}} \int_{-\infty}^{0} \|\mathbf{u}(t)\|_2^2 \mathrm{d}t = \begin{bmatrix} \mathbf{x}_0^{\mathrm{T}} & \dot{\mathbf{x}}_0^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{12}^{\mathrm{T}} & \mathbf{R}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \dot{\mathbf{x}}_0 \end{bmatrix} \\ &= \mathbf{x}_0^{\mathrm{T}}\mathbf{R}_{11}\mathbf{x}_0 + 2\mathbf{x}_0^{\mathrm{T}}\mathbf{R}_{12}\dot{\mathbf{x}}_0 + \dot{\mathbf{x}}_0^{\mathrm{T}}\mathbf{R}_{22}\dot{\mathbf{x}}_0. \end{aligned}$$

We minimize this energy with respect to the vector $\mathbf{x}_0$, which yields

$$\nabla_{\mathbf{x}_0} E_{\mathbf{u}}(\mathbf{x}_0) = 2\mathbf{R}_{11}\mathbf{x}_0 + 2\mathbf{R}_{12}\dot{\mathbf{x}}_0$$

and hence the minimal input energy is attained for $\mathbf{x}_0 = -\mathbf{R}_{22}^{-1}\mathbf{R}_{12}\dot{\mathbf{x}}_0$. This results in the input energy

$$E_{\mathbf{u}}(\dot{\mathbf{x}}_0) = \dot{\mathbf{x}}_0^{\mathrm{T}}\mathbf{R}_{22}\dot{\mathbf{x}}_0 - \dot{\mathbf{x}}_0^{\mathrm{T}}\mathbf{R}_{12}^{\mathrm{T}}\mathbf{R}_{11}^{-1}\mathbf{R}_{12}\dot{\mathbf{x}}_0 = \dot{\mathbf{x}}_0^{\mathrm{T}}(\mathbf{R}_{22} - \mathbf{R}_{12}^{\mathrm{T}}\mathbf{R}_{11}^{-1}\mathbf{R}_{12})\dot{\mathbf{x}}_0 = \dot{\mathbf{x}}_0^{\mathrm{T}}\mathbf{P}_{22}^{-1}\dot{\mathbf{x}}_0, \tag{2.47}$$

where $\mathbf{P}_{22} = \mathbf{P}_{\text{vel}}$ is the velocity controllability Gramian as defined in (2.26), and hence, states corresponding to small eigenvalues of $\mathbf{P}_{\text{vel}}$ are hard to reach.

To investigate the output energy for second-order systems (2.23), the authors in [44] consider the respective first-order observability Gramian $\mathfrak{Q}_{\text{L}}$ as defined in (2.5) corresponding to the first-order matrices (2.24). This Gramian $\mathfrak{Q}_{\text{L}}$ includes the position observability Gramian on its upper-left block $\mathbf{Q}_{11} = \mathbf{Q}_{\text{L,pos}}$ and the velocity observability Gramian on its lower-right block $\mathbf{Q}_{22} = \mathbf{Q}_{\text{L,vel}}$. Again, the first-order system analysis from (2.33) is used to obtain the energy that is generated by a first-order state $\mathbf{z}_0 = \left[\begin{smallmatrix}\mathbf{x}_0\\ \dot{\mathbf{x}}_0\end{smallmatrix}\right]$ that is

$$E_{\mathbf{y}_{\text{L}}}(\mathbf{z}_0) = \begin{bmatrix}\mathbf{x}_0^{\text{T}} & \dot{\mathbf{x}}_0^{\text{T}}\end{bmatrix} \begin{bmatrix}\mathbf{Q}_{11} & \mathbf{Q}_{12}\\ \mathbf{Q}_{12}^{\text{T}} & \mathbf{Q}_{22}\end{bmatrix} \begin{bmatrix}\mathbf{x}_0\\ \dot{\mathbf{x}}_0\end{bmatrix} = \mathbf{x}_0^{\text{T}}\mathbf{Q}_{11}\mathbf{x}_0 + 2\dot{\mathbf{x}}_0^{\text{T}}\mathbf{Q}_{12}^{\text{T}}\mathbf{x}_0 + \dot{\mathbf{x}}_0^{\text{T}}\mathbf{Q}_{22}\dot{\mathbf{x}}_0.$$

The energy functional $E_{\mathbf{y}_{\text{L}}}$ evaluates the output energy generated by the system if we are at a state $\mathbf{z}_0$ without any input $\mathbf{u}$. Since we are interested in the energy generated by an initial state $\mathbf{x}_0$, we first set $\dot{\mathbf{x}}_0 = 0$ to evaluate the energy generated by $\mathbf{x}_0$ that is

$$E_{\mathbf{y}_{\text{L}}}(\mathbf{x}_0) = \mathbf{x}_0^{\text{T}}\mathbf{Q}_{11}\mathbf{x}_0. \tag{2.48}$$

To evaluate the energy generated by an initial velocity $\dot{\mathbf{x}}_0$, the authors in [44] set $\mathbf{x}_0 = 0$ which yields

$$E_{\mathbf{y}_{\text{L}}}(\dot{\mathbf{x}}_0) = \dot{\mathbf{x}}_0^{\text{T}}\mathbf{Q}_{22}\dot{\mathbf{x}}_0. \tag{2.49}$$

The equations (2.48) and (2.49) show that states corresponding to small singular values of $\mathbf{Q}_{11}$ and $\mathbf{Q}_{22}$ generate small amounts of observable energies while states corresponding to large singular values define the dominant observability subspaces.

Since states corresponding to small singular values of $\mathbf{P}_{\text{pos}}$, $\mathbf{P}_{\text{vel}}$ and $\mathbf{Q}_{\text{L,pos}}$, $\mathbf{Q}_{\text{L,vel}}$ are hard to reach and hard to observe, respectively, we aim to truncate theses states to generate a reduced surrogate model. For this purpose, various Gramian combinations can be used, see [112]. As for first-order systems, the controllability and observability Gramians need to coincide.

Definition 2.21:
Consider the asymptotically stable second-order system (2.23) and the corresponding position and velocity controllability Gramians $\mathbf{P}_{\text{pos}}$, $\mathbf{P}_{\text{vel}}$ as defined in (2.26) and the observability Gramians $\mathbf{Q}_{\text{L,pos}}$, $\mathbf{Q}_{\text{L,vel}}$ as introduced in (2.27). The system is called *balanced* if it holds

$$\mathbf{P} = \mathbf{Q}_{\text{L}} = \mathbf{\Sigma}$$

where $\mathbf{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$ is a diagonal matrix, $\mathbf{P}$ denotes either $\mathbf{P}_{\text{pos}}$ or $\mathbf{P}_{\text{vel}}$, and $\mathbf{Q}_{\text{L}}$ denotes $\mathbf{Q}_{\text{L,pos}}$ or $\mathbf{Q}_{\text{L,vel}}$. ◊

Assume that the Gramians can be factorized as $\mathbf{P} = \mathbf{R}\mathbf{R}^{\mathrm{T}}$ and $\mathbf{Q}_{\mathrm{L}} = \mathbf{S}\mathbf{S}^{\mathrm{T}}$ where $\mathbf{R}$ and $\mathbf{S}$ are either Cholesky factors or low-rank factors of $\mathbf{P}$ and $\mathbf{Q}_{\mathrm{L}}$, respectively, and $\mathbf{P}$ represents either the position controllability Gramian $\mathbf{P}_{\mathrm{pos}}$ or the velocity controllability Gramian $\mathbf{P}_{\mathrm{vel}}$ while $\mathbf{Q}_{\mathrm{L}}$ represents either the position observability Gramian $\mathbf{Q}_{\mathrm{L,pos}}$ or the velocity observability Gramian $\mathbf{Q}_{\mathrm{L,vel}}$. Considering the position Gramians, we compute the singular value decomposition of

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

which is used to define the balancing transformation matrices

$$\mathbf{W}_{\mathrm{b}} := \mathbf{S}\mathbf{U}\mathbf{\Sigma}^{-\frac{1}{2}}, \qquad \mathbf{T}_{\mathrm{b}} := \mathbf{R}\mathbf{V}\mathbf{\Sigma}^{-\frac{1}{2}}. \tag{2.50}$$

For the velocity Gramians, we proceed analogously, but we use the singular value decomposition of $\mathbf{S}^{\mathrm{T}}\mathbf{M}\mathbf{R}$.

BT balances the system and truncates the states corresponding to the smallest singular values stored in $\mathbf{\Sigma}$. Therefore, states that are simultaneously hardest to reach and hardest to observe are removed from the system dynamics. To do so, we define the balancing and truncating bases

$$\mathbf{W}_{\mathrm{r}} := \mathbf{S}\mathbf{U}_1\mathbf{\Sigma}_1^{-\frac{1}{2}}, \qquad \mathbf{T}_{\mathrm{r}} := \mathbf{R}\mathbf{V}_1\mathbf{\Sigma}_1^{-\frac{1}{2}}. \tag{2.51}$$

We multiply the system matrices by the two bases $\mathbf{W}_{\mathrm{r}}^{\mathrm{T}}$ and $\mathbf{T}_{\mathrm{r}}$ to generate the reduced matrices as shown in (2.45) and to define the reduced system (2.44). Up to now, there exists no a priori error bound for second-order BT methods.

2.2.2 Iterative rational Krylov algorithm

In this paragraph, we introduce the iterative rational Krylov algorithm (IRKA) as described in [60] to reduce systems of the form (2.1) and extended to DAE systems in [61]. The authors in [140] derive an IRKA method for second-order systems.

2.2.2.1 IRKA for first-order ODE systems

Within the IRKA method, we aim to derive a reduced surrogate model (2.28) approximating the original dynamical system (2.1). To explain the IRKA method, we consider the transfer function $\mathcal{G}_{\mathrm{L}}$ as introduced in (2.4) that encodes the input-to-output behavior of the original system (2.1). We consider the transfer function of the reduced system that is $\mathcal{G}_{\mathrm{L,r}}$. From Proposition 2.6, it follows that

$$\|\mathbf{y} - \mathbf{y}_{\mathrm{r}}\|_{L_\infty} \leq \|\mathcal{G}_{\mathrm{L}} - \mathcal{G}_{\mathrm{L,r}}\|_{\mathcal{H}_2}\|\mathbf{u}\|_{L_2}$$

Algorithm 3 BT method for the second-order ODE system (2.23).

```
Require: The system (2.23) and the reduced order $r$.
Ensure: The reduced system (2.44).
Compute factors of the Gramians $\mathbf{P} \approx \mathbf{R}\mathbf{R}^{\mathrm{T}}$ and $\mathbf{Q} \approx \mathbf{S}\mathbf{S}^{\mathrm{T}}$ from (2.26), (2.27).
Perform the SVD of $\mathbf{S}^{\mathrm{T}}\mathbf{R}$ or $\mathbf{S}^{\mathrm{T}}\mathbf{M}\mathbf{R}$, and decompose as
    $\mathbf{S}^{\mathrm{T}}\mathbf{R} \text{ or } \mathbf{S}^{\mathrm{T}}\mathbf{M}\mathbf{R} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_1 & \\ & \boldsymbol{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^{\mathrm{T}} \\ \mathbf{V}_2^{\mathrm{T}} \end{bmatrix},$
    with $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{r \times r}$.
Construct the projection matrices
    $\mathbf{W}_{\mathrm{r}} = \mathbf{S}\mathbf{U}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}} \text{ and } \mathbf{T}_{\mathrm{r}} = \mathbf{R}\mathbf{V}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}.$
Construct reduced matrices (2.45).
```

if both transfer functions live in $\boldsymbol{\mathcal{H}}_2^{p \times m}$. That bound provides that reduced outputs $\mathbf{y}_{\mathrm{r}}(t)$ are uniformly close to $\mathbf{y}(t)$ over all inputs $\mathbf{u} \in L_2([0,\infty), \mathbb{R}^m)$ if the transfer functions are close in the $\mathcal{H}_2$-norm. Hence, we aim at constructing a reduced order model that minimizes the $\mathcal{H}_2$ approximation error as follows

$$\|\mathcal{G}_{\mathrm{L}} - \mathcal{G}_{\mathrm{L,r}}\|_{\mathcal{H}_2} = \min_{\substack{\dim \widehat{\mathcal{G}}_{\mathrm{L}} = R \\ \widehat{\mathcal{G}}_{\mathrm{L}} \text{ is stable}}} \|\mathcal{G}_{\mathrm{L}} - \widehat{\mathcal{G}}_{\mathrm{L}}\|_{\mathcal{H}_2}, \tag{2.52}$$

where $\dim \widehat{\mathcal{G}}_{\mathrm{L}}$ denotes the McMillan degree that is the number of the poles of $\widehat{\mathcal{G}}_{\mathrm{L}}$. For a given degree N rational function $\mathcal{G}_{\mathrm{L}}$, we seek a degree R rational function $\mathcal{G}_{\mathrm{L,r}}$ that approximates $\mathcal{G}_{\mathrm{L}}$ w.r.t. the $\boldsymbol{\mathcal{H}}_2$-norm. This optimization problem is non-convex. Therefore, the search for a global optimum is infeasible, so we aim to find local minimizers. To solve this problem, we inspect the optimality conditions that are tangential interpolatory conditions. In the multi-input multi-output (MIMO) case, we require that $\mathcal{G}_{\mathrm{L}}(s)$ and $\mathcal{G}_{\mathrm{L,r}}(s)$ coincide for the interpolation points s along determined directions, the *tangential directions.* We call $\mathcal{G}_{\mathrm{L,r}}(s)$ a *right tangential interpolant to* $\mathcal{G}_{\mathrm{L}}(s)$ *at* σ *along the right tangential direction* $\mathbf{b} \in \mathbb{C}^m$ if

$$\mathcal{G}_{\mathrm{L}}(\sigma)\mathbf{b} = \mathcal{G}_{\mathrm{L,r}}(\sigma)\mathbf{b} \tag{2.53}$$

and accordingly *a left tangential interpolant to* $\mathcal{G}_{\mathrm{L}}(s)$ *at* σ *along the left tangential direction* $\mathbf{c} \in \mathbb{C}^p$ if

$$\mathbf{c}^{\mathrm{T}}\mathcal{G}_{\mathrm{L}}(\sigma) = \mathbf{c}^{\mathrm{T}}\mathcal{G}_{\mathrm{L,r}}(\sigma). \tag{2.54}$$

Also $\mathcal{G}_{\mathrm{L,r}}(s)$ is called a bitangential Hermite interpolant to $\mathcal{G}_{\mathrm{L}}(s)$ at σ along the right tangential direction $\mathbf{b} \in \mathbb{C}^m$ and the left tangential direction $\mathbf{c} \in \mathbb{C}^p$ if

$$\mathbf{c}^{\mathrm{T}}\mathcal{G}'(\sigma)\mathbf{b} = \mathbf{c}^{\mathrm{T}}\mathcal{G}'_{\mathrm{L,r}}(\sigma)\mathbf{b} \tag{2.55}$$

where $(\cdot)'$ denotes the derivative with respect to s. For a set of given interpolation points $\sigma_1, \ldots, \sigma_r$, and left and right tangential directions $\mathbf{c}_1, \ldots, \mathbf{c}_r$ and $\mathbf{b}_1, \ldots, \mathbf{b}_r$, we aim to find a reduced model that matches these interpolation conditions.

For that, we aim to find projecting matrices $\boldsymbol{\mathcal{V}}_{\mathrm{r}} \in \mathbb{R}^{n \times r}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r}} \in \mathbb{R}^{n \times r}$ defining a surrogate model (2.28) with reduced matrices (2.29) and a transfer function $\mathcal{G}_{\mathrm{L,r}}$ that satisfy the Petrov-Galerkin orthogonality conditions (2.30) and (2.31). As described in, e.g., [59, 60, 154], the following theorem gives a criterion for generating the reduction bases.

Theorem 2.22:
Consider the system in (2.1) with the transfer function $\mathcal{G}_{\mathrm{L}}$ and a reduced order model (2.28) with the respective transfer function $\mathcal{G}_{\mathrm{L,r}}$ generated by the left and right basis $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$, $\boldsymbol{\mathcal{T}}_{\mathrm{r}}$ as described in (2.29). Assume that $\sigma_1, \ldots, \sigma_R$ and $\mu_1, \ldots \mu_r$ are right and left interpolation points and $\mathbf{b}_1, \ldots, \mathbf{b}_R$ and $\mathbf{c}_1, \ldots, \mathbf{c}_R$ are given right and left tangential directions. Then the following statements hold:

a) if $(\sigma_k\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_k \in \operatorname{range}(\boldsymbol{\mathcal{T}}_{\mathrm{r}})$ then $\mathcal{G}_{\mathrm{L}}(\sigma_k)\mathbf{b}_k = \mathcal{G}_{\mathrm{L,r}}(\sigma_k)\mathbf{b}_k$,

b) if $(\mu_k\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\mathrm{T}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\mathbf{c}_k \in \operatorname{range}(\boldsymbol{\mathcal{V}}_{\mathrm{r}})$ then $\mathbf{c}_k\mathcal{G}_{\mathrm{L}}(\mu_k) = \mathbf{c}_k\mathcal{G}_{\mathrm{L,r}}(\sigma_k)$,

c) if a) and b) are satisfied at $\sigma_k = \mu_k$ then $\mathbf{c}_k\mathcal{G}'_{\mathrm{L}}(\sigma_k)\mathbf{b}_k = \mathbf{c}_k\mathcal{G}'_{\mathrm{L,r}}(\sigma_k)\mathbf{b}_k$

for $k = 1, \ldots, R$. ◊

Using Theorem 2.22, we define the bases $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r}}$ that satisfy the tangential interpolation conditions as

$$\begin{aligned}
\boldsymbol{\mathcal{T}}_{\mathrm{r}} &= \left[(\sigma_1\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_1 \quad \ldots \quad (\sigma_R\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_R\right], \\
\boldsymbol{\mathcal{V}}_{\mathrm{r}} &= \left[(\sigma_1\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\mathrm{H}}\boldsymbol{\mathcal{C}}^{\mathrm{H}}\mathbf{c}_1 \quad \ldots \quad (\sigma_R\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\mathrm{H}}\boldsymbol{\mathcal{C}}^{\mathrm{H}}\mathbf{c}_R\right]
\end{aligned} \tag{2.56}$$

for some interpolation points $\sigma_1, \ldots, \sigma_R$ and right and left tangential directions $\mathbf{b}_1, \ldots, \mathbf{b}_r$ and $\mathbf{c}_1, \ldots, \mathbf{c}_r$, respectively. The reduced system and the respective transfer function generated by these bases satisfy the interpolation conditions without computing the interpolated values, which is a significant advantage.

After presenting the interpolation of the transfer functions, we use these results to approach the problem of finding a $\mathcal{H}_2$ optimal replacement model. To do so, we consider the pole-residue representation of the reduced transfer function $\mathcal{G}_{\mathrm{L,r}}(s)$, which is

$$\mathcal{G}_{\mathrm{L,r}}(s) = \boldsymbol{\mathcal{C}}_{\mathrm{r}}(s\boldsymbol{\mathcal{E}}_{\mathrm{r}} - \boldsymbol{\mathcal{A}}_{\mathrm{r}})^{-1}\boldsymbol{\mathcal{C}}_{\mathrm{r}} = \sum_{k=1}^{R} \frac{\widehat{\mathbf{c}}_k\widehat{\mathbf{b}}_k^{\mathrm{T}}}{s - \lambda_k}, \tag{2.57}$$

where $\lambda_1, \dots, \lambda_R$ are assumed to be distinct poles of $\mathcal{G}_{\mathrm{L,r}}(s)$, the vectors $\widehat{\mathbf{c}}_1, \dots, \widehat{\mathbf{c}}_R$ and $\widehat{\mathbf{b}}_1, \dots, \widehat{\mathbf{b}}_R$ are the respective left and right residue directions, and $\widehat{\mathbf{c}}_1\widehat{\mathbf{b}}_1^{\mathrm{T}}, \dots, \widehat{\mathbf{c}}_R\widehat{\mathbf{b}}_R^{\mathrm{T}}$ the respective matrix residues of $\mathcal{G}_{\mathrm{L,r}}(s)$ at $s = \lambda_1, \dots, \lambda_R$. We determine the poles and residue directions by computing the generalized eigenvalue decomposition of $\lambda\boldsymbol{\mathcal{E}}_{\mathrm{r}} - \boldsymbol{\mathcal{A}}_{\mathrm{r}}$. Since the dimension R is assumed to be small, the computation of the eigenvalue decomposition is feasible. To derive the optimal interpolation points and directions corresponding to the pole-residue representation of $\mathcal{G}_{\mathrm{L,r}}$ described in (2.57), we consider the following theorem.

Theorem 2.23:
Let $\mathcal{G}_{\mathrm{L,r}}$ in pole-residue form (2.57) be a minimizer of the optimization problem (2.52) with respect to a transfer function $\mathcal{G}_{\mathrm{L}}$ and assume that $\mathcal{G}_{\mathrm{L,r}}$ has only simple poles $\lambda_1, \dots, \lambda_R$. Then $\mathcal{G}_{\mathrm{L,r}}$ interpolates $\mathcal{G}_{\mathrm{L}}$ and $\mathcal{G}'_{\mathrm{L,r}}$ interpolates $\mathcal{G}'_{\mathrm{L}}$ at $-\lambda_1, \dots, -\lambda_R$ along the right and left tangential directions $\widehat{\mathbf{b}}_1, \dots, \widehat{\mathbf{b}}_R$ and $\widehat{\mathbf{c}}_1, \dots, \widehat{\mathbf{c}}_R$, i.e.,

$$\begin{aligned} \mathcal{G}_{\mathrm{L,r}}(-\lambda_k)\widehat{\mathbf{b}}_k &= \mathcal{G}_{\mathrm{L}}(-\lambda_k)\widehat{\mathbf{b}}_k, \qquad \widehat{\mathbf{c}}_k^{\mathrm{T}}\mathcal{G}_{\mathrm{L,r}}(-\lambda_k) = \widehat{\mathbf{c}}_k^{\mathrm{T}}\mathcal{G}_{\mathrm{L}}(-\lambda_k), \\ \widehat{\mathbf{c}}_k^{\mathrm{T}}\mathcal{G}'_{\mathrm{L,r}}(-\lambda_k)\widehat{\mathbf{b}}_k &= \widehat{\mathbf{c}}_k^{\mathrm{T}}\mathcal{G}'_{\mathrm{L}}(-\lambda_k)\widehat{\mathbf{b}}_k \end{aligned} \tag{2.58}$$

holds for $k = 1, \dots, R$. ◊

It follows that if the transfer function $\mathcal{G}_{\mathrm{L,r}}$ is a local minimizer of (2.52), the interpolation conditions in (2.58) are satisfied. Hence, to build the bases in (2.56), we use the poles and residue direction of the reduced transfer function $\mathcal{G}_{\mathrm{L,r}}$ as interpolation points and tangential directions. Assume that $\mathbf{s} := \{s_1, \dots, s_R\}$ is the set of the currently considered expansion point and $\boldsymbol{\lambda}(\mathbf{s}) = \{\lambda_1, \dots, \lambda_R\}$ the resulting poles of $\boldsymbol{\mathcal{E}}_{\mathrm{r}}^{-1}\boldsymbol{\mathcal{A}}_{\mathrm{r}}$. Then we can define the function $\mathbf{g}(\mathbf{s}) = \mathbf{s} + \boldsymbol{\lambda}(\mathbf{s})$. Aside from reordering, if $\mathbf{g}(\mathbf{s}) = 0$, then the optimization problem (2.52) is solved by the current basis, which means that $\mathcal{G}_{\mathrm{L,r}}(s)$ corresponds to a $\mathcal{H}_2$-optimal reduced system (2.28). Hence, we apply Newton's method to the function $\mathbf{g}$ to determine iteratively the optimal expansion points $\mathbf{s} = \boldsymbol{\lambda}(\mathbf{s})$. This results in the following iteration

$$\mathbf{s}^{k+1} = \mathbf{s}^k - (\mathbf{I} - \nabla_{\mathbf{s}}\boldsymbol{\lambda}(\mathbf{s}))^{-1}(\mathbf{s}^k - \boldsymbol{\lambda}(\mathbf{s}^k))$$

where $\nabla_{\mathbf{s}}\boldsymbol{\lambda}(\mathbf{s})$ is the Jacobian of $\boldsymbol{\lambda}(\mathbf{s})$. Since often the entries of $\nabla_{\mathbf{s}}\boldsymbol{\lambda}(\mathbf{s})$ are small, we set $\nabla_{\mathbf{s}}\boldsymbol{\lambda}(\mathbf{s}) = 0$ and obtain

$$\mathbf{s}^{k+1} = \boldsymbol{\lambda}(\mathbf{s}^k).$$

This iteration defines the IRKA method, described in Algorithm 4.

2.2.2.2 IRKA for first-order DAE systems

For systems of DAE structure as defined in (2.8), we aim to derive a reduced system (2.37) that approximates the input-to-output behavior described by the $\mathcal{H}_2$-norm of

Algorithm 4 IRKA method for the first-order ODE system (2.1).

```
Require: The original system (2.1), maximum number of iterations $N_{\max}$, tolerance tol, reduced dimension $R$.
Ensure: A reduced system (2.28) that satisfies (2.52)
Choose initial expansion points $s_1, \dots, s_R$, left tangential direction $\mathbf{c}_1, \dots, \mathbf{c}_R$ and right tangential directions $\mathbf{b}_1, \dots, \mathbf{b}_R$.
while Iteration number $\leq N_{\max}$ and $s_1, \dots, s_R$ did not converge do
    Choose $\mathcal{V}_{\mathrm{r}}$ and $\mathcal{T}_{\mathrm{r}}$ so that
        $\mathcal{T}_{\mathrm{r}} = \left[(s_1\mathcal{E} - \mathcal{A})^{-1}\mathcal{B}\mathbf{b}_1, \dots, (s_R\mathcal{E} - \mathcal{A})^{-1}\mathcal{B}\mathbf{b}_R\right]$
        $\mathcal{V}_{\mathrm{r}} = \left[(s_1\mathcal{E} - \mathcal{A})^{-\mathrm{H}}\mathcal{C}^{\mathrm{H}}\mathbf{c}_1, \dots, (s_R\mathcal{E} - \mathcal{A})^{-\mathrm{H}}\mathcal{C}^{\mathrm{H}}\mathbf{c}_R\right].$
    Build reduced matrices as in (2.29) using $\mathcal{T}_{\mathrm{r}}$ and $\mathcal{V}_{\mathrm{r}}$.
    Compute the pole-residue expansion (2.57) of $\mathcal{G}_{\mathrm{L,r}}$ corresponding to the by $\mathcal{V}_{\mathrm{r}}$ and $\mathcal{T}_{\mathrm{r}}$ reduced system (2.28).
    Set $s_j = -\lambda_j$, $\mathbf{b}_j = \widehat{\mathbf{b}}_j$ and $\mathbf{c}_j = \widehat{\mathbf{c}}_j$, $j = 1, \dots, R$.
end while
```

the transfer function error. The interpolation conditions described in Theorem 2.22 for ODE system also hold for DAE systems. However, applying the IRKA method as described in Algorithm 4 to generate a reduced system (2.37) might lead to unbounded error measures. This is because the transfer function $\mathcal{G}_{\mathrm{L}}(s)$ corresponding to the original system (2.8) consist of a strictly proper component $\mathcal{G}_{\mathrm{L,p}}(s)$ and a polynomial one $\mathcal{G}_{\mathrm{L,i}}(s)$ as described in (2.16). Hence, the reduced transfer function also requires to have a strictly proper component $\mathcal{G}_{\mathrm{L,r,p}}(s)$ and a polynomial one $\mathcal{G}_{\mathrm{L,r,i}}(s)$, with $\mathcal{G}_{\mathrm{L,i}}(s) = \mathcal{G}_{\mathrm{L,r,i}}(s)$, so that

$$\mathcal{G}_{\mathrm{L}}(s) - \mathcal{G}_{\mathrm{L,r}}(s) = \mathcal{G}_{\mathrm{L,p}}(s) - \mathcal{G}_{\mathrm{L,r,p}}(s) + \mathcal{G}_{\mathrm{L,i}}(s) - \mathcal{G}_{\mathrm{L,r,i}}(s) = \mathcal{G}_{\mathrm{L,p}}(s) - \mathcal{G}_{\mathrm{L,r,p}}(s).$$

Otherwise, the error in the transfer functions is not bounded. However, utilizing the bases $\mathcal{V}_{\mathrm{r}}$ and $\mathcal{T}_{\mathrm{r}}$ determined with Algorithm 4 for the matrices of the DAE system (2.8) is likely to lead to an ODE system (2.37) if R is smaller that the rank of $\mathcal{E}$ so that the reduced transfer function has no polynomial component or a constant one if we add a feed-through term $\mathcal{D}$. Therefore, we need to maintain the polynomial component $\mathcal{G}_{\mathrm{L,i}}(s)$ of the original transfer function $\mathcal{G}_{\mathrm{L}}(s)$. The authors in [61, 156] discuss a procedure to generate a reduced surrogate model that preserves the polynomial system components as described in the following theorem.

Theorem 2.24:
Consider the transfer function $\mathcal{G}_{\mathrm{L}}(s) = \mathcal{G}_{\mathrm{L,p}}(s) + \mathcal{G}_{\mathrm{L,i}}(s)$ corresponding to the DAE system (2.8) where $\mathcal{G}_{\mathrm{L,p}}(s)$ is the strictly proper component and $\mathcal{G}_{\mathrm{L,i}}(s)$ the polynomial

one, respectively. Let $\mathbf{P}_\mathrm{r}$ and $\mathbf{P}_\mathrm{l}$ be the spectral projectors defined in (2.10), and let the columns of $\boldsymbol{\mathcal{T}}_\infty$ and $\boldsymbol{\mathcal{V}}_\infty$ span the right and left deflating subspaces of $(\mathbf{A}, \mathbf{E})$ corresponding to the eigenvalues λ at infinity, i.e., $\boldsymbol{\mathcal{T}}_\infty$ and $\boldsymbol{\mathcal{V}}_\infty$ span the same spaces as $(\mathbf{I} - \mathbf{P}_\mathrm{r})$ and $(\mathbf{I} - \mathbf{P}_\mathrm{l})$. Assume that for right and left interpolation points $\sigma_1, \ldots, \sigma_R$ and $\mu_1, \ldots, \mu_R$ the matrices $\sigma_k\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}$, $\mu_k\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}$ are invertable for $k = 1, \ldots, R$ and that $\mathbf{b}_1, \ldots, \mathbf{b}_R$ and $\mathbf{c}_1, \ldots, \mathbf{c}_R$ are the right and left nonzero tangent directions. Construct

$$\boldsymbol{\mathcal{T}}_{N_f} = \left[(\sigma_1\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{B}}\mathbf{b}_1, \ldots, (\sigma_R\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{B}}\mathbf{b}_R\right], \tag{2.59}$$

$$\boldsymbol{\mathcal{V}}_{N_f} = \left[(\mu_1\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\mathrm{H}}\mathbf{P}_\mathrm{r}^\mathrm{T}\boldsymbol{\mathcal{C}}^\mathrm{T}\mathbf{c}_1, \ldots, (\mu_R\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\mathrm{H}}\mathbf{P}_\mathrm{r}^\mathrm{T}\boldsymbol{\mathcal{C}}^\mathrm{T}\mathbf{c}_R\right], \tag{2.60}$$

and define

$$\boldsymbol{\mathcal{T}}_\mathrm{r} = \begin{bmatrix}\boldsymbol{\mathcal{T}}_{N_f} & \boldsymbol{\mathcal{T}}_\infty\end{bmatrix}, \qquad \boldsymbol{\mathcal{V}}_\mathrm{r} = \begin{bmatrix}\boldsymbol{\mathcal{V}}_{N_f} & \boldsymbol{\mathcal{V}}_\infty\end{bmatrix}. \tag{2.61}$$

Assume that $\mathcal{G}_{\mathrm{L,r}}(s) = \mathcal{G}_{\mathrm{L,r,p}}(s) + \mathcal{G}_{\mathrm{L,r,i}}(s)$ is the transfer function corresponding to the reduced system (2.37) with matrices (2.24) generated by the bases $\boldsymbol{\mathcal{V}}_\mathrm{r}$ and $\boldsymbol{\mathcal{T}}_\mathrm{r}$, where $\mathcal{G}_{\mathrm{L,r,p}}(s)$ is the respective strictly proper component and $\mathcal{G}_{\mathrm{L,r,i}}(s)$ is the polynomial one. Then it holds $\mathcal{G}_{\mathrm{L,i}}(s) = \mathcal{G}_{\mathrm{L,r,i}}(s)$ and a) and b) from Theorem 2.22 are fulfilled. If also $\sigma_k = \mu_k$ holds for $k = 1, \ldots, R$, then also c) from Theorem 2.22 is fulfilled. ◊

Using this theorem, we define an IRKA method tailored for the first-order DAE system. For that, we apply Algorithm 4 to derive $\boldsymbol{\mathcal{T}}_{N_f}$ and $\boldsymbol{\mathcal{V}}_{N_f}$ by replacing $\boldsymbol{\mathcal{B}}$ by $\mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{B}}$ and $\boldsymbol{\mathcal{C}}$ by $\boldsymbol{\mathcal{C}}\mathbf{P}_\mathrm{r}$. Using these bases we derive the bases $\boldsymbol{\mathcal{T}}_\mathrm{r}$ and $\boldsymbol{\mathcal{V}}_\mathrm{r}$ as defined in (2.61).

2.2.2.3 IRKA for second-order ODE systems

In this paragraph, we briefly present the iterative rational Krylov method (IRKA) suitable for the case of second-order systems (2.23), as introduced in [156], where we aim to find a reduced system (2.44) with a transfer function

$$\mathcal{G}_{\mathrm{L,r}}(s) := (\mathbf{C}_{1,\mathrm{r}} + s\mathbf{C}_{2,\mathrm{r}})(s^2\mathbf{M}_\mathrm{r} + s\mathbf{D}_\mathrm{r} + \mathbf{K}_\mathrm{r})^{-1}\mathbf{B}_\mathrm{r}. \tag{2.62}$$

Within the IRKA approach, we determine a reduced system that maintains the second-order structure while following an approach similar to the one for first-order systems. The author in [156] derives projecting matrices $\mathbf{W}_\mathrm{r}$ and $\mathbf{T}_\mathrm{r}$ to construct the reduced matrices (2.45) and the respective reduced system (2.44). However, the choice of the projecting bases also depends on additional conditions applied to the reduced systems. For mechanical systems (2.23), the aim is to find bases that preserve the symmetry and the positive definiteness of the mass matrix, the damping matrix, and the stiffness matrix to obtain an asymptotically stable reduced system. Hence, we set $\mathbf{V}_\mathrm{r} = \mathbf{T}_\mathrm{r} = \mathbf{W}_\mathrm{r}$. Also, the reduced system is supposed to be of second-order structure. Hence, the methods presented in [140, 156], the authors generate a reduced transfer function of the structure

Algorithm 5 IRKA method for the first-order DAE system (2.8).

```
Require: The original system (2.8), maximum number of iterations $N_{\max}$, tolerance tol, reduced dimension $R$.
Ensure: A reduced system (2.37) that satisfies (2.52)
Choose initial expansion points $s_1, \dots, s_R$, left tangential direction $\mathbf{c}_1, \dots, \mathbf{c}_R$ and right tangential directions $\mathbf{b}_1, \dots, \mathbf{b}_R$.
while s did not converge do
    Choose $\mathcal{T}_{\mathrm{r}} = \begin{bmatrix} \mathcal{T}_{N_f} & \mathcal{T}_{\infty} \end{bmatrix}$ and $\mathcal{V}_{\mathrm{r}} = \begin{bmatrix} \mathcal{V}_{N_f} & \mathcal{V}_{\infty} \end{bmatrix}$, where
        $$\begin{aligned} \mathcal{T}_{N_f} &= \left[ (\sigma_1 \mathcal{E} - \mathcal{A})^{-1} \mathbf{P}_{\mathrm{l}} \mathcal{B} \mathbf{b}_1, \dots, (\sigma_R \mathcal{E} - \mathcal{A})^{-1} \mathbf{P}_{\mathrm{l}} \mathcal{B} \mathbf{b}_R \right] \\ \mathcal{V}_{N_f} &= \left[ (\mu_1 \mathcal{E} - \mathcal{A})^{-\mathrm{H}} \mathbf{P}_{\mathrm{r}} \mathcal{C}^{\mathrm{H}} \mathbf{c}_1, \dots, (\mu_R \mathcal{E} - \mathcal{A})^{-\mathrm{H}} \mathbf{P}_{\mathrm{r}} \mathcal{C}^{\mathrm{H}} \mathbf{c}_R \right] \end{aligned}$$
    and $\mathcal{T}_{\infty}$, $\mathcal{V}_{\infty}$ are chosen so that they span the right and left deflating subspaces of $(\mathbf{A}, \mathbf{E})$ corresponding to $\lambda = \infty$.
    Build reduced matrices as in (2.29) using $\mathcal{V}_{\mathrm{r}}$ and $\mathcal{T}_{\mathrm{r}}$.
    Compute the pole-residue expansion (2.57) of $\mathcal{G}_{\mathrm{L,r,p}}$ corresponding to the by $\mathcal{V}_{N_f}$ and $\mathcal{T}_{N_f}$ reduced system (2.37).
    Set $s_j = -\lambda_j$, $\mathbf{b}_j = \widehat{\mathbf{b}}_j$ and $\mathbf{c}_j = \widehat{\mathbf{c}}_j$, $j = 1, \dots, R$.
end while
Build reduced matrices as in (2.29) using $\mathcal{T}_{\mathrm{r}}$ and $\mathcal{V}_{\mathrm{r}}$.
```

shown in (2.62) that represents a second-order system (2.44). For that, they use a one-sided projection approach that generates a basis $\mathbf{V}_{\mathrm{r}}$ with

$$\mathbf{V}_{\mathrm{r}} = \left[(s_1^2 \mathbf{M} + s_1 \mathbf{D} + \mathbf{K})^{-1} \mathbf{B} \mathbf{b}_1 \quad \dots \quad (s_r^2 \mathbf{M} + s_r \mathbf{D} + \mathbf{K})^{-1} \mathbf{B} \mathbf{b}_r \right], \tag{2.63}$$

for interpolation points $s_1, \dots, s_r$ and tangential directions $\mathbf{b}_1, \dots, \mathbf{b}_r$. The interpolation points and tangential directions are updated in each step of the method. After a basis $\mathbf{V}_{\mathrm{r}}$ is built, the reduced matrices, which correspond to a reduced second-order system with a transfer function of order $2r$, are built. Since this order is twice the dimension we aim for, we apply an internal reduction step. By applying a second IRKA or BT method to the system defined by the matrices in (2.45), we obtain a reduced-order system of dimension r with a transfer function $\mathcal{G}_{\mathrm{L,r}}$. We determine the respective poles and residues to obtain the interpolation points and tangential directions, which are used in the next step to derive the basis $\mathbf{V}_{\mathrm{r}}$ and the respective reduced system. This procedure results in Algorithm 6 from [140].

Algorithm 6 IRKA method for the second-order ODE system (2.23).

```
Require: The original system (2.23), maximum number of iterations $N_{\max}$, tolerance tol, reduced dimension $r$.
Ensure: A reduced system (2.44) that satisfies (2.58).
Choose initial expansion points $s_1, \dots, s_r$ and right tangential directions $\mathbf{b}_1, \dots, \mathbf{b}_r$.
while Iteration number $\leq N_{\max}$ and $s_1, \dots, s_r$ did not converge do
    Set
    $$\mathbf{V}_{\mathrm{r}} = \left[(s_1^2\mathbf{M} + s_1\mathbf{D} + \mathbf{K})^{-1}\mathbf{B}\mathbf{b}_1 \quad \dots \quad (s_r^2\mathbf{M} + s_r\mathbf{D} + \mathbf{K})^{-1}\mathbf{B}\mathbf{b}_r\right].$$
    Determine reduced matrices as in (2.45).
    Compute the pole-residue expansion (2.57) of $\mathcal{G}_{\mathrm{L,r}}$ corresponding to the reduced system (2.44).
    Determine new interpolation points and tangential directions
    $$s_1, \dots, s_r = -\lambda_1, \dots, -\lambda_r, \qquad \mathbf{b}_1, \dots, \mathbf{b}_r = \widehat{\mathbf{b}}_1, \dots, \widehat{\mathbf{b}}_r.$$
end while
Determine reduced matrices as in (2.45).
```

2.3 Lyapunov equations

As shown in Section 2.2.1, we need to solve certain Lyapunov equations to compute the Gramians of the respective system. Hence, this section we aim to solve the Lyapunov equations from (2.6) and the projected Lyapunov equations from (2.19), where we focus on the controllability case while the observability Lyapunov equations are solved similarly. There are multiple methods to solve this kind of equation. If the matrix dimensions are sufficiently small, we use Hammarling's method [65] or the Bartels-Steward algorithm [11]. However, these methods are unfeasible if the matrix dimensions are large. In this case, the alternating-direction implicit (ADI) method [80, 83, 87, 101], the sign function method [27] and Krylov subspace methods [72, 123, 125] are the state of the art. Those methods require that the system representation is sparse. An overview and comparison of those methods is given in [29, 47, 124].

Since we consider in this section Lyapunov equations, including system matrices (2.24) corresponding to a mechanical system of the structure (2.1.3), we can exploit the structure of these matrices in the following. Therefore, we consider a decomposition of the matrix $\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}$ that is

$$\begin{aligned}
&\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}} = \widetilde{\boldsymbol{\mathcal{A}}} - \widetilde{\mathbf{U}}\mathbf{G}\widetilde{\mathbf{V}}^{\mathrm{T}}, \quad \text{with}\\
&\qquad \widetilde{\boldsymbol{\mathcal{A}}} := \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{D}_{\mathrm{int}} \end{bmatrix}, \ \widetilde{\mathbf{U}} := \begin{bmatrix} 0 \\ \mathbf{M}^{-1}\mathbf{F} \end{bmatrix}, \ \widetilde{\mathbf{V}} := \begin{bmatrix} 0 \\ \mathbf{F} \end{bmatrix},
\end{aligned} \tag{2.64}$$

with $\mathbf{D} = \mathbf{D}_{\mathrm{int}} + \mathbf{F}\mathbf{G}\mathbf{F}^{\mathrm{T}}$, into a sparse matrix $\widetilde{\boldsymbol{\mathcal{A}}}$ and into a low-rank matrix $\widetilde{\mathbf{U}}\mathbf{G}\widetilde{\mathbf{V}}^{\mathrm{T}}$.

In this work, the methods of choice are the alternating direction implicit (ADI) method, introduced in Section 2.3.1, and the sign function method shown in Section 2.3.2, which are both iterative methods that derive the solution of Lyapunov equations that can make use of the decomposition (2.64) to decrease their computational costs.

2.3.1 Alternating direction implicit method

In this subsection, the alternating direction implicit method (ADI) from [80, 102] is presented. This method is applied in this manuscript to solve ordinary Lyapunov equations as defined in (2.6) and projected Lyapunov equations (2.19) that arise when considering DAE systems. The author in [80] has also derived an ADI method for second-order systems, which we omit in this work as it converges too slowly in the considered problem setting.

2.3.1.1 ADI method for classic Lyapunov equations

We aim to compute the solution $\boldsymbol{\mathcal{P}}$ of the Lyapunov equation (2.6) of large dimension. To do so, we utilize that this Lyapunov equation is equivalent to the Stein equation

$$\boldsymbol{\mathcal{P}} = \boldsymbol{\mathcal{T}}(p)\boldsymbol{\mathcal{S}}(p)\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{S}}(p)^{\mathrm{H}}\boldsymbol{\mathcal{T}}(p)^{\mathrm{H}} - 2\sqrt{\mathrm{Re}(p)}\boldsymbol{\mathcal{S}}(p)\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{S}}(p)^{\mathrm{H}}. \tag{2.65}$$

for $\boldsymbol{\mathcal{S}}(p) := (\mathbf{A} + p\boldsymbol{\mathcal{E}})^{-1}$ and $\boldsymbol{\mathcal{T}}(p) := (\boldsymbol{\mathcal{A}} - \overline{p}\boldsymbol{\mathcal{E}})$. We choose several shift parameters $p_1, \ldots, p_\ell \in \mathbb{C}^-$ so that the spectral radius

$$\rho\left(\boldsymbol{\mathcal{T}}(p_k)\boldsymbol{\mathcal{S}}(p_k)\right) < 1$$

is as small as possible for all $k = 1, \ldots, \ell$, and obtain the resulting ADI iteration

$$\begin{aligned}
\boldsymbol{\mathcal{P}}_0 &:= 0,\\
\boldsymbol{\mathcal{P}}_k &:= \boldsymbol{\mathcal{T}}(p_k)\boldsymbol{\mathcal{S}}(p_k)\boldsymbol{\mathcal{P}}_{k-1}\boldsymbol{\mathcal{S}}(p_k)^{\mathrm{H}}\boldsymbol{\mathcal{T}}(p_k)^{\mathrm{H}} - 2\mathrm{Re}(p_k)\boldsymbol{\mathcal{S}}(p_k)\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{S}}(p_k)^{\mathrm{H}}.
\end{aligned} \tag{2.66}$$

Since the right-hand side of the Lyapunov equation consists of the low-rank factor $\boldsymbol{\mathcal{B}}$ with a dimension $m \ll N$, the solution $\boldsymbol{\mathcal{P}}$ can be well-approximated by $\boldsymbol{\mathcal{P}} \approx \boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{Z}}^{\mathrm{T}}$ with the tall and skinny matrix $\boldsymbol{\mathcal{Z}} \in \mathbb{C}^{N \times N_{\boldsymbol{\mathcal{Z}}}}$, $N_{\boldsymbol{\mathcal{Z}}} \ll N$. Hence, by introducing $\boldsymbol{\mathcal{P}}_k = \boldsymbol{\mathcal{Z}}_k \boldsymbol{\mathcal{Z}}_k^{\mathrm{T}}$, the iteration (2.66) is equal to

$$\begin{aligned} \boldsymbol{\mathcal{Z}}_k &:= \left[\boldsymbol{\mathcal{Z}}_{k-1} \quad -\sqrt{2\mathrm{Re}(p_k)}\mathbf{V}_k\right], \qquad \boldsymbol{\mathcal{Z}}_0 := [\,], \\ \mathbf{V}_k &:= \mathbf{V}_{k-1} - (p_k + \overline{p}_{k-1})\boldsymbol{\mathcal{S}}(p_k)\boldsymbol{\mathcal{E}}\mathbf{V}_{k-1} = \boldsymbol{\mathcal{S}}(p_k)\mathbf{W}_{k-1}, \\ \mathbf{W}_k &:= \mathbf{W}_{k-1} - 2\mathrm{Re}(p_k)\boldsymbol{\mathcal{E}}\mathbf{V}_k, \qquad \mathbf{W}_0 := \boldsymbol{\mathcal{B}}, \end{aligned}$$

where $\boldsymbol{\mathcal{Z}}_0$ is an empty matrix. Using the decomposition from (2.64) we compute the inverse of $(\boldsymbol{\mathcal{A}} + p_k\boldsymbol{\mathcal{E}})$ efficiently by applying the Sherman-Morrison-Woodbury formula as

$$\begin{aligned} \boldsymbol{\mathcal{S}}(p_k) &= (\boldsymbol{\mathcal{A}} + p_k\boldsymbol{\mathcal{E}})^{-1} \\ &= \left(\widetilde{\boldsymbol{\mathcal{A}}} + p_k\mathbf{I} - \widetilde{\mathbf{U}}\boldsymbol{\mathcal{G}}\widetilde{\mathbf{V}}^{\mathrm{T}}\right)^{-1}\boldsymbol{\mathcal{E}}^{-1} \\ &= \Big((\widetilde{\boldsymbol{\mathcal{A}}} + p_k\mathbf{I})^{-1} \\ &\qquad + (\widetilde{\boldsymbol{\mathcal{A}}} + p_k\mathbf{I})^{-1}\widetilde{\mathbf{U}}\left(\boldsymbol{\mathcal{G}}^{-1} - \widetilde{\mathbf{V}}^{\mathrm{T}}(\widetilde{\boldsymbol{\mathcal{A}}} + p_k\mathbf{I})^{-1}\widetilde{\mathbf{U}}\right)^{-1}\widetilde{\mathbf{V}}^{\mathrm{T}}(\widetilde{\boldsymbol{\mathcal{A}}} + p_k\mathbf{I})^{-1}\Big)\boldsymbol{\mathcal{E}}^{-1} \end{aligned}$$

with matrices defined in (2.64). Since $\widetilde{\boldsymbol{\mathcal{A}}} + p_k\mathbf{I}$ and $\boldsymbol{\mathcal{E}}$ are easy to invert and $\widetilde{\mathbf{U}}$, $\widetilde{\mathbf{V}}$ are of small dimension, this structure accelerates the computation of $\boldsymbol{\mathcal{S}}(p_k)$.

One can show that the norm of the residual after the k-th step $\mathfrak{R}_k = \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}_k\boldsymbol{\mathcal{E}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}_k\boldsymbol{\mathcal{A}}^{\mathrm{T}} + \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}$ is given by $\|\mathfrak{R}_k\| = \|\mathbf{W}_k^{\mathrm{H}}\mathbf{W}_k\|$, and hence, the residual is used as a stopping criterion that does not require additional computational costs.

If we require a real-valued approximation of $\boldsymbol{\mathcal{P}}$, the shifts have to occur in pairs of complex conjugate shifts, i.e., if $p_k \in \mathbb{C}^- \setminus \mathbb{R}$ then $p_{k+1} = \overline{p}_k$. In that case one can show, that the $(k+1)$-st iterates $\mathbf{V}_{k+1}$ and $\mathbf{W}_{k+1}$ are equal to

$$\begin{aligned} \mathbf{V}_{k+1} &:= \overline{\mathbf{V}}_k + 2\frac{\mathrm{Re}(p_k)}{\mathrm{Im}(p_k)}\mathrm{Im}(\mathbf{V}_k), \\ \mathbf{W}_{k+1} &:= \mathbf{W}_{k-1} - 4\mathrm{Re}(p_k)\boldsymbol{\mathcal{E}}\left(\mathrm{Re}(\mathbf{V}_k) + \frac{\mathrm{Re}(p_k)}{\mathrm{Im}(p_k)}\mathrm{Im}(\mathbf{V}_k)\right). \end{aligned}$$

The remaining task is to determine the shift parameters $p_1, \ldots, p_\ell$ where we use the self-generating shifts presented in [24]. The main idea of these shifts is that we generate the Ritz values from certain spaces. For that, we assume that we have an orthonormal basis $\mathbf{U} \in \mathbb{R}^{n \times n_U}$ so that the shifts are

$$\{p_1, \ldots, p_\ell\} = \Lambda(\mathbf{U}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\mathbf{U}) \cap \mathbb{C}^-.$$

The initial shifts are generated using a first basis $\mathbf{U}$ that spans the columns of the low-rank factors of the right-hand side $\boldsymbol{\mathcal{B}}$. For the next iteration steps –assume we are in the k-th step of the iteration– we use a basis $\mathbf{U}$ that fulfills

$$\mathrm{span}(\mathbf{U}) = \mathrm{span}(\mathbf{V}_k) \qquad \text{or} \qquad \mathrm{span}(\mathbf{U}) = \mathrm{span}(\mathrm{Re}(\mathbf{V}_k), \mathrm{Im}(\mathbf{V}_k)).$$

If the dimension of $\mathbf{V}_k$ is not large enough to generate a required number of shift parameters, the columns of the previous iterates $\mathbf{V}_{k-1}$, $\mathbf{V}_{k-2}, \ldots$ can be used. This ADI procedure, including the displacement parameters, is implemented in [114] and is used in this work for the various numerical examples.

2.3.1.2 ADI method for projected Lyapunov equations

This subsection aims to present numerical techniques to solve the projected Lyapunov equations (2.19) and (2.22) to approximate the Gramians of system (2.1) with the singular matrix $\mathcal{E}$. We utilize the ADI method to solve the projected continuous-time Lyapunov equations and the generalized Smith method to solve the discrete-time Lyapunov equations.

Here, we follow the ideas of [133] to derive an equation equivalent to the projected continuous-time Lyapunov equation (2.19). First, we extend the Stein equation from (2.65) to the projected Lyapunov equation as shown in the following lemma.

Lemma 2.25 ([133]):
Let the matrix pencil $s\mathcal{E} - \mathcal{A}$ with $\mathcal{E}$, $\mathcal{A} \in \mathbb{R}^{N\times N}$ be regular. Let further the matrix $\mathcal{A}$ be nonsingular and $\mathcal{B} \in \mathbb{R}^{N\times m}$. Assume that the left and right spectral projectors onto the finite spectrum of $s\mathcal{E} - \mathcal{A}$ from (2.10) are denoted by $\mathbf{P}_\mathrm{l}$, $\mathbf{P}_\mathrm{r} \in \mathbb{R}^{N\times N}$. If $p \in \mathbb{C}$ is not an eigenvalue of the pencil $s\mathcal{A} - \mathcal{E}$, then the projected discrete-time Lyapunov equation

$$\mathcal{P}_\mathrm{p} = \mathcal{S}(p)\mathcal{R}(p)\mathcal{P}_\mathrm{p}\mathcal{R}(p)^\mathrm{H}\mathcal{S}(p)^\mathrm{H} - 2\mathrm{Re}(p)\mathcal{S}(p)\mathbf{P}_\mathrm{l}\mathcal{B}\mathcal{B}^\mathrm{T}\mathbf{P}_\mathrm{l}^\mathrm{T}\mathcal{S}(p)^\mathrm{H}, \qquad \mathcal{P}_\mathrm{p} = \mathbf{P}_\mathrm{r}\mathcal{P}_\mathrm{p}\mathbf{P}_\mathrm{r}^\mathrm{T} \tag{2.67}$$

with $\mathcal{S}(p) := (\mathcal{E}+p\mathcal{A})^{-1}$ and $\mathcal{R}(p) := (\mathcal{E}-\overline{p}\mathcal{A})$ is equivalent to the projected continuous-time Lyapunov equation (2.19), i. e., their solution sets coincide. ◊

The projected Stein equation (2.67) motivates the ADI iteration similar to (2.66) that is

$$\begin{aligned} &\mathcal{P}_0 := 0, \\ &\mathcal{P}_k := \mathcal{S}(p_k)\mathcal{R}(p_k)\mathcal{P}_{k-1}\mathcal{R}(p_k)^\mathrm{H}\mathcal{S}(p_k)^\mathrm{H} - 2\mathrm{Re}(p_k)\mathcal{S}(p_k)\mathbf{P}_\mathrm{l}\mathcal{B}\mathcal{B}^\mathrm{T}\mathbf{P}_\mathrm{l}^\mathrm{T}\mathcal{S}(p_k)^\mathrm{H}. \end{aligned} \tag{2.68}$$

As shown in [133], given a sequence of shift parameters $(p_k)_{k\geq 0}$ in $\mathbb{C}^-$ with $p_{k+\ell} = p_k$ for some $\ell \geq 1$ and all $k = 0, 1, 2, \ldots$, the iteration (2.68) converges to the solution $\mathcal{P}_\mathrm{p}$ of the projected Lyapunov equation (2.19).

To work with the ADI iteration more efficiently, we aim to compute low-rank factors of $\mathcal{P}_\mathrm{p}$, i.e., we aim to determine a tall and skinny matrix $\mathcal{Z} \in \mathbb{C}^{N\times N_\mathcal{Z}}$, $N_\mathcal{Z} \ll N$, such that $\mathcal{P}_\mathrm{p} \approx \mathcal{Z}\mathcal{Z}^\mathrm{H}$. We can represent the iteration (2.68) by the low-rank factors of $\mathcal{P}_k = \mathcal{Z}_k\mathcal{Z}_k^\mathrm{H}$ with

$$\begin{aligned} \mathcal{Z}_k &= \big[\kappa(p_k)\mathbf{S}(p_k)\mathbf{P}_\mathrm{l}\mathcal{B} \quad \mathcal{S}(p_k)\mathcal{R}(p_k)\mathcal{Z}_{k-1}\big] \\ &= \big[\kappa(p_k)\mathcal{S}(p_k)\mathbf{P}_\mathrm{l}\mathcal{B} \quad \kappa(p_{k-1})\mathcal{S}(p_k)\mathcal{R}(p_k)\mathcal{S}(p_{k-1})\mathbf{P}_\mathrm{l}\mathcal{B} \\ &\qquad\qquad \ldots \quad \kappa(p_1)\mathcal{S}(p_k)\mathcal{R}(p_k)\cdot\ldots\cdot\mathcal{S}(p_2)\mathcal{R}(p_2)\mathbf{S}(p_1)\mathbf{P}_\mathrm{l}\mathcal{B}\big], \end{aligned}$$

where $\kappa(p_k) = \sqrt{-\mathrm{Re}(p_k)}$ and $\mathcal{Z}_0$ is chosen to be an empty matrix in $\mathbb{R}^{N\times 0}$. We note that the following properties hold for all j, $k = 0, 1, \ldots$:

$$\begin{aligned} &\mathcal{S}(p_k)\mathcal{A}\mathcal{S}(p_j) = \mathcal{S}(p_j)\mathcal{A}\mathcal{S}(p_k), \qquad \mathcal{R}(p_k)\mathcal{A}^{-1}\mathcal{R}(p_j) = \mathcal{R}(p_j)\mathcal{A}^{-1}\mathcal{R}(p_k),\\ &\mathcal{S}(p_k)\mathcal{R}(p_j) = \mathcal{A}^{-1}\mathcal{R}(p_j)\mathcal{S}(p_k)\mathcal{A}. \end{aligned} \tag{2.69}$$

We further define

$$\mathcal{B}_0 := \kappa(p_k)\mathcal{S}(p_k)\mathbf{P}_\mathrm{l}\mathcal{B} \qquad \text{and} \qquad \mathcal{F}_j := \frac{\kappa(p_j)}{\kappa(p_{j-1})}\mathcal{S}(p_j)\mathcal{R}(p_{j+1}), \quad j = 1, \ldots, k.$$

Using (2.69), we obtain

$$\mathcal{Z}_k = \begin{bmatrix}\mathcal{B}_0 & \mathcal{F}_{k-1}\mathcal{B}_0 & \ldots & \mathcal{F}_1 \cdot \ldots \cdot \mathcal{F}_{k-1}\mathcal{B}_0\end{bmatrix}.$$

It remains to solve the discrete-time Lvapunov equation (2.19). Under the assumption that $\mathcal{A}$ is nonsingular, (2.19) is equivalent to the transformed discrete-time Lyapunov equation

$$\mathcal{P}_\mathrm{i} - \mathcal{A}^{-1}\mathcal{E}\mathcal{P}_\mathrm{i}\mathcal{E}^\mathrm{T}\mathcal{A}^{-\mathrm{T}} = \mathcal{A}^{-1}(\mathbf{I}_N - \mathbf{P}_\mathrm{l})\mathcal{B}\mathcal{B}^\mathrm{T}(\mathbf{I}_N - \mathbf{P}_\mathrm{l})^\mathrm{T}\mathcal{A}^{-\mathrm{T}}, \qquad 0 = \mathbf{P}_\mathrm{r}\mathcal{P}_\mathrm{i}\mathbf{P}_\mathrm{r}^\mathrm{T}.$$

This equation is solved using the Smith method [133]. Since $\mathcal{A}^{-1}(\mathbf{I} - \mathbf{P}_\mathrm{l}) = (\mathbf{I}_N - \mathbf{P}_\mathrm{r})\mathcal{A}^{-1}$ and the matrix $(\mathbf{I}_N - \mathbf{P}_\mathrm{r})\mathcal{A}^{-1}\mathcal{E} = \mathcal{A}^{-1}\mathcal{E}(\mathbf{I}_N - \mathbf{P}_\mathrm{r})$ is nilpotent with the nilpotency index ν, the iteration stops after ν steps. The Smith method then leads to the unique solution

$$\mathcal{P}_\mathrm{i} = \sum_{k=0}^{\nu-1}(\mathcal{A}^{-1}\mathcal{E})^k(\mathbf{I}_N - \mathbf{P}_\mathrm{r})\mathcal{A}^{-1}\mathcal{B}\mathcal{B}^\mathrm{T}\mathcal{A}^{-\mathrm{T}}(\mathbf{I}_N - \mathbf{P}_\mathrm{r})^\mathrm{T}((\mathcal{A}^{-1}\mathcal{E})^\mathrm{T})^k.$$

Instead of computing the full matrix $\mathcal{P}_\mathrm{i}$ we can also generate the low-rank factors $\mathcal{P}_\mathrm{i} = \mathcal{Y}\mathcal{Y}^\mathrm{T}$ as

$$\mathcal{Y} = \begin{bmatrix}(\mathbf{I} - \mathbf{P}_\mathrm{r})\mathcal{A}^{-1}\mathcal{B} & \mathcal{A}^{-1}\mathcal{E}(\mathbf{I} - \mathbf{P}_\mathrm{r})\mathcal{A}^{-1}\mathcal{B} & \ldots & (\mathcal{A}^{-1}\mathcal{E})^{\nu-1}(I - \mathbf{P}_\mathrm{r})\mathcal{A}^{-1}\mathcal{B}\end{bmatrix}.$$

2.3.2 Sign function method

As a second Lyapunov equation-solving method, we consider the sign function method introduced in [27] and extended for system structures that appear in this work by [46].

2.3.2.1 Sign function method for classic Lyapunov equations

To describe the sign function method, we first define the sign function for an arbitrary, quadratic matrix $\mathbf{A}$.

Definition 2.26:
For a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ with $\Lambda(\mathbf{A}) \cap \mathrm{i}\mathbb{R} = \emptyset$ and a Jordan decomposition $\mathbf{A} = \mathbf{Z}\mathbf{J}\mathbf{Z}^{-1}$ with $\mathbf{J} = \operatorname{diag}(\mathbf{J}^{-},\ \mathbf{J}^{+})$, $\mathbf{J}^{-} \in \mathbb{C}^{n_{-} \times n_{-}}$, $\Lambda(\mathbf{J}^{-}) \subset \mathbb{C}^{-}$ and $\mathbf{J}^{+} \in \mathbb{C}^{n_{+} \times n_{+}}$, $\Lambda(\mathbf{J}^{+}) \subset \mathbb{C}^{+}$, the *sign function* of $\mathbf{A}$ is defined as

$$\operatorname{sign}(\mathbf{A}) := \mathbf{Z} \begin{bmatrix} -\mathbf{I}_{n_{-}} & 0 \\ 0 & \mathbf{I}_{n_{+}} \end{bmatrix} \mathbf{Z}^{-1}.$$

The sign function of a matrix is unique and independent of the eigenvalue order of the Jordan decomposition. ♢

The sign function is computed by applying Newton's method:

$$\mathbf{A}_0 := \mathbf{A}, \quad \mathbf{A}_{k+1} = \frac{1}{2} c_k \mathbf{A}_k + \frac{1}{2c_k} \mathbf{A}_k^{-1} \quad \to \quad \operatorname{sign}(\mathbf{A}) \tag{2.70}$$

where c_k denotes acceleration factor which is chosen to be $c_k = \sqrt{\|\mathbf{A}_k^{-1}\|_{\mathrm{F}} \|\mathbf{A}_k\|_{\mathrm{F}}^{-1}}$ within this work. The convergence of this method is shown in [113].

Within the sign function method, the authors in [46] utilize the structure introduced in (2.64) to solve the Lyapunov equations more efficiently. Additionally, we notice that the solution $\boldsymbol{\mathcal{P}}$ can be approximated by the low-rank factor $\boldsymbol{\mathcal{Z}} \in \mathbb{R}^{N \times N_{\mathcal{Z}}}$ with $\boldsymbol{\mathcal{P}} \approx \boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{Z}}^{\mathrm{T}}$ due to the structure of the right-hand side of the corresponding Lyapunov equation (2.19). Hence, this subsection aims to determine the low-rank factor $\boldsymbol{\mathcal{Z}}$ that approximately solves the Lyapunov equation (2.19). We exploit the fact that this Lyapunov equation is equivalent to

$$\begin{bmatrix} \mathbf{I} & 0 \\ -\boldsymbol{\mathcal{P}} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{A}}^{\mathrm{T}} & 0 \\ 0 & -\boldsymbol{\mathcal{A}} \end{bmatrix} \begin{bmatrix} \mathbf{I} & 0 \\ \boldsymbol{\mathcal{P}} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{A}}^{\mathrm{T}} & 0 \\ \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}} & -\boldsymbol{\mathcal{A}} \end{bmatrix} =: \boldsymbol{\mathcal{W}} \quad \text{and that} \quad \operatorname{sign}(\boldsymbol{\mathcal{W}}) = \begin{bmatrix} -\mathbf{I} & 0 \\ 2\boldsymbol{\mathcal{P}} & \mathbf{I} \end{bmatrix}.$$

We observe that the sign function of $\boldsymbol{\mathcal{W}}$ provides the solution $\boldsymbol{\mathcal{P}} \approx \boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{Z}}^{\mathrm{T}}$ of the Lyapunov equation (2.19) with $\boldsymbol{\mathcal{E}} = \mathbf{I}$ in its lower left block. We apply Newton's method described in equation (2.70) to compute $\operatorname{sign}(\boldsymbol{\mathcal{W}})$. We set $\boldsymbol{\mathcal{A}}_0 := \boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}_0 := \boldsymbol{\mathcal{B}}$ to obtain the following iterations:

$$\boldsymbol{\mathcal{A}}_{k+1} = \frac{1}{2}\left(c_k \boldsymbol{\mathcal{A}}_k + \frac{1}{c_k} \boldsymbol{\mathcal{A}}_k^{-1} \right), \qquad \boldsymbol{\mathcal{B}}_{k+1} = \frac{1}{\sqrt{2}} \left[\sqrt{c_k} \boldsymbol{\mathcal{B}}_k, \ \frac{1}{\sqrt{c_k}} \boldsymbol{\mathcal{A}}_k^{-1} \boldsymbol{\mathcal{B}}_k \right], \tag{2.71}$$

where $\boldsymbol{\mathcal{B}}_k$ converges to $\frac{1}{\sqrt{2}}\boldsymbol{\mathcal{Z}}$.

In order to improve the efficiency while computing the inverse $\boldsymbol{\mathcal{A}}^{-1}$ we make use of the decomposition presented in equation (2.64) and apply the Sherman-Morrison-Woodbury formula as described in [46] to obtain

$$\boldsymbol{\mathcal{A}}_{k+1} = \frac{1}{2} c_k \boldsymbol{\mathcal{A}}_k + \frac{1}{2c_k} \boldsymbol{\mathcal{A}}_k^{-1} = \widetilde{\boldsymbol{\mathcal{A}}}_{k+1} - \widetilde{\mathbf{U}}_{k+1} \mathbf{G}_{k+1} \widetilde{\mathbf{V}}_{k+1}^{\mathrm{T}},$$

with

$$\widetilde{\boldsymbol{\mathcal{A}}}_{k+1} = \frac{1}{2}\left(c_k\widetilde{\boldsymbol{\mathcal{A}}}_k + \frac{1}{c_k}\widetilde{\boldsymbol{\mathcal{A}}}_k^{-1}\right), \; \widetilde{\mathbf{U}}_{k+1} = \left[\widetilde{\mathbf{U}}_k,\; \widetilde{\boldsymbol{\mathcal{A}}}_k^{-1}\widetilde{\mathbf{U}}_k\right], \; \widetilde{\mathbf{V}}_{k+1} = \left[\widetilde{\mathbf{V}}_k,\; \widetilde{\boldsymbol{\mathcal{A}}}_k^{-\mathrm{T}}\widetilde{\mathbf{V}}_k\right],$$

$$\mathbf{G}_{k+1} = \frac{1}{2}\mathrm{diag}\left(c_k\mathbf{G}_k,\; -\frac{1}{c_k}(\mathbf{G}_k^{-1} - \widetilde{\mathbf{V}}_k^{\mathrm{T}}\widetilde{\boldsymbol{\mathcal{A}}}_k^{-1}\widetilde{\mathbf{U}}_k)^{-1}\right).$$

We stop this method if $\|\boldsymbol{\mathcal{A}}_k + \mathbf{I}\|^2 \leq \mathrm{tol}$ since $\boldsymbol{\mathcal{A}}_k$ converges to $-\mathbf{I}$, or if a maximum number of iterations $\mathrm{iter}_{\max}$ is exceeded. The disadvantage of this method is the high growth rate of the dimension of the low-rank factor $\boldsymbol{\mathcal{Z}}_k$. Therefore, even with internal truncation techniques, the method must converge after a few steps, stop, or become slow when calculating the next steps.

2.3.2.2 Sign function method for projected Lyapunov equations

As described in [132], the sign function method can be extended for projected Lyapunov equations of the form (2.19). If the matrix pencil $(\mathbf{A}, \mathbf{E})$ is C-stable, i.e., all respective finite eigenvalues have a negative real part, then the WCF presented in (2.11) contains a nilpotent matrix $\mathbf{N}$ and a matrix $\mathbf{J}$ that also only has eigenvalues with a negative real part, according to [54, 132] . We make use of this decomposition to derive a sign function extension for projected Lyapunov equations with an arbitrary nilpotency index ν corresponding to $\mathbf{N}$. For that, we need to remove the $\mathbf{N}$ matrix from the iteration by multiplying the matrix $\boldsymbol{\mathcal{E}}$ from the left or right by the projecting matrices $\mathbf{P}_\mathrm{l}$ or $\mathbf{P}_\mathrm{r}$ from (2.10), respectively, and apply Newton's method from (2.71). To avoid that $\boldsymbol{\mathcal{A}}_k$ converges to a singular matrix we add an additional summand $(\mathbf{I} - \mathbf{P}_\mathrm{l})\boldsymbol{\mathcal{A}}(\mathbf{I} - \mathbf{P}_\mathrm{r})$ that does not affect the method. This leads to the generalized iteration as introduced by [132]:

$$\boldsymbol{\mathcal{A}}_0 := \boldsymbol{\mathcal{A}}, \qquad \boldsymbol{\mathcal{P}}_0 = \mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^\mathrm{T}\mathbf{P}_\mathrm{l}^\mathrm{T},$$

$$\boldsymbol{\mathcal{A}}_k := \frac{1}{2c_k}\left(\boldsymbol{\mathcal{A}}_{k-1} + c_k^2\mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{A}}_{k-1}^{-1}\boldsymbol{\mathcal{E}}^\mathrm{T}\mathbf{P}_\mathrm{r} + (2c_k - 1)(\mathbf{I} - \mathbf{P}_\mathrm{l})\boldsymbol{\mathcal{A}}(\mathbf{I} - \mathbf{P}_\mathrm{r})\right) = \mathbf{W}\begin{bmatrix}\mathbf{J}_k^{-} & 0\\ 0 & \mathbf{I}\end{bmatrix}\mathbf{T},$$

$$\boldsymbol{\mathcal{P}}_k = \frac{1}{2c_k}\left(\boldsymbol{\mathcal{P}}_{k-1} + c_k^2\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{A}}_k^{-1}\boldsymbol{\mathcal{P}}_{k-1}\boldsymbol{\mathcal{A}}_k^{-\mathrm{T}}\boldsymbol{\mathcal{E}}^\mathrm{T}\right).$$

It holds that $\boldsymbol{\mathcal{P}}_k = \mathbf{P}_\mathrm{r}\boldsymbol{\mathcal{P}}_k\mathbf{P}_\mathrm{r}^\mathrm{T}$ and that $\lim_{k\to\infty}\boldsymbol{\mathcal{A}}_k^{-\mathrm{T}}\boldsymbol{\mathcal{P}}_k\boldsymbol{\mathcal{A}}_k^{-1} = 2\boldsymbol{\mathcal{P}}_\mathrm{p}$ with $\boldsymbol{\mathcal{P}}_\mathrm{p} = \mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{P}}_\mathrm{p}\mathbf{P}_\mathrm{l}$. Also it holds that $\lim_{k\to\infty}\boldsymbol{\mathcal{A}}_k = -\boldsymbol{\mathcal{E}}\mathbf{P}_\mathrm{r} + \boldsymbol{\mathcal{A}}(\mathbf{I} - \mathbf{P}_\mathrm{r})$, which results in the stopping criterion

$$\|\boldsymbol{\mathcal{A}}_k + \boldsymbol{\mathcal{E}}\mathbf{P}_\mathrm{r} - \boldsymbol{\mathcal{A}}(\mathbf{I} - \mathbf{P}_\mathrm{r})\| \leq \mathrm{tol}.$$

When we aim to find low-rank factors $\boldsymbol{\mathcal{Z}}_\mathrm{p} \in \mathbb{R}^{N\times N_{\boldsymbol{\mathcal{Z}}}}$ with $\boldsymbol{\mathcal{P}}_\mathrm{p} \approx \boldsymbol{\mathcal{Z}}_\mathrm{p}\boldsymbol{\mathcal{Z}}_\mathrm{p}^\mathrm{T}$, the iteration of $\boldsymbol{\mathcal{P}}_k$ can be replaced by

$$\boldsymbol{\mathcal{B}}_k := \frac{1}{\sqrt{2c_k}}\begin{bmatrix}\boldsymbol{\mathcal{B}}_{k-1} & c_k\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{A}}_k^{-1}\boldsymbol{\mathcal{B}}_{k-1}\end{bmatrix}, \qquad \boldsymbol{\mathcal{B}}_0 := \mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{B}}$$

with $\lim_{k\to\infty}\boldsymbol{\mathcal{A}}_k^{-1}\boldsymbol{\mathcal{B}}_k = \sqrt{2}\boldsymbol{\mathcal{Z}}_\mathrm{p}$.

CHAPTER 3

INHOMOGENEOUS SYSTEMS AND THEIR SYSTEM THEORETICAL ASPECTS

Contents

The analysis of the dynamical systems presented in Chapter 1 can provide an understanding of their behavior and is therefore used to identify the most significant system components. In Section 2.1, the properties of homogeneous systems with linear output equations investigated in established literature were repeated. In this work, however, we deal with inhomogeneous systems that are evaluated using linear and quadratic output equations. Therefore, in this chapter, we extend the concepts of Section 2.1.

For first-order systems with an ODE as a state equation and inhomogeneous initial conditions, some methods have already been developed in [13, 15, 66, 121]. In [13], the authors propose to shift the state by the initial condition $\mathbf{z}_0$, i.e., the new state is given as $\tilde{\mathbf{z}}(t) := \mathbf{z}(t) - \mathbf{z}_0$. In this way, the initial condition is included in the input and output equations and thus is taken into account in the reduction process. In [66], the input $\mathbf{B}\mathbf{u}(t)$ is extended by the initial condition space $\mathbf{Z}_0$, i.e., $\mathbf{z}_0 = \mathbf{Z}_0\zeta_0$. More precisely, a new input matrix $\widetilde{\mathbf{B}} := [\mathbf{B}\ \mathbf{Z}_0]$ and a new input $[\mathbf{u}(t)^{\mathrm{T}}\ \zeta_0^{\mathrm{T}}]^{\mathrm{T}}$ are defined such that the initial condition is taken into account using reduction techniques. In [15], the author's strategy is to decompose the system into one with no initial conditions and one with initial conditions but no input. The sum of the two corresponding outputs is equal to the original output. This superposition is used to reduce these two systems separately. Extensions of that methodology for the class of bilinear systems are proposed in [42] and [110], based on different splittings. A recent approach [121] introduces a new balanced truncation method based on the shift transformation of the state. This transformation depends on designing parameters that allow some flexibility and the generalization of the methods proposed in [66] and [15]. In addition, these parameters can be optimized, leading to accurate reduced-order models.

In this chapter, the superposition ideas from [15] and the extended-input approach from [66] are extended to the system classes that are relevant to this work. The introduction of the controllability and observability spaces and the respective Gramians are novelties of this manuscript, which are used in the remaining chapters to reduce the systems of the respective structures. Many of these systems have never been analyzed in the literature, so the theory in this chapter is a valuable addition to existing theories.

Especially for systems with a quadratic output equation and those with inhomogeneous initial conditions, new definitions are derived, which are relevant for the rest of this thesis. Hence, the main contributions of this chapter are the definition of transfer functions describing the input-to-output behavior and the respective tailored controllability and observability Gramians for systems in non-standard form, which also take into account the effects of initial conditions on the system dynamics. In particular, first-order ODE systems with quadratic output equations, first-order DAE systems with linear and quadratic initial equations, and second-order systems with linear and quadratic initial equations have not been studied until now in the literature. Therefore, these Gramian derivatives are a significant contribution to the analysis of dynamical systems. The energy expressions derived in this chapter also provide the basis for the reduction methods applied later in this manuscript. The concepts presented in this chapter are partially published in [106].

This chapter is structured as follows. In Section 3.1, we repeat and extend the theory to first-order systems with an ODE as state equation. Afterwards, in Section 3.2, we introduce different methods for first-order systems with DAEs as state equations, and finally, in Section 3.3, we analyze systems with a second-order structure.

3.1 Inhomogeneous first-order ODE systems

In this section, we consider first-order systems with a state equation

$$\boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) = \boldsymbol{\mathcal{A}}\mathbf{z}(t) + \boldsymbol{\mathcal{B}}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0 \tag{3.1}$$

with $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{N\times N}$, and $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{N\times m}$. The matrix $\boldsymbol{\mathcal{E}}$ is assumed to be nonsingular, so we consider an ODE as state equation. The vectors $\mathbf{z}(t) \in \mathbb{R}^N$ and $\mathbf{u}(t) \in \mathbb{R}^m$ denote the state and the input, respectively. These systems arise when modeling mechanical systems, as described in the introduction, but also in circuit simulation, heat transfer simulations, fluid simulations, and several biological and chemical fields of application. Examples are shown, e.g., in [38, 45, 158]. The solution trajectory of the dynamical system (3.1) is given as

$$\mathbf{z}(t) = \int_0^t e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}(t-\tau)}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\mathbf{u}(\tau)\mathrm{d}\tau + e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\mathbf{z}_0. \tag{3.2}$$

We observe that the state $\mathbf{z}(t) = \mathbf{z}_{\boldsymbol{\mathcal{B}}}(t) + \mathbf{z}_{\mathbf{z}_0}(t)$ consists of two components, one corresponding to the input and one that results from the initial condition, that are

$$\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t) := \int_0^t e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}(t-\tau)}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\mathbf{u}(\tau)\mathrm{d}\tau \qquad \text{and} \qquad \mathbf{z}_{\mathbf{z}_0}(t) := e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\mathbf{z}_0, \tag{3.3}$$

respectively.

We assume that there exists a matrix $\mathbf{Z}_0 \in \mathbb{R}^{N\times N_{\mathbf{z}_0}}$, $N_{\mathbf{z}_0} \in \mathbb{N}_+$, so that all admissible initial conditions satisfy

$$\mathbf{z}(0) = \mathbf{z}_0 = \mathbf{Z}_0\zeta_0 \tag{3.4}$$

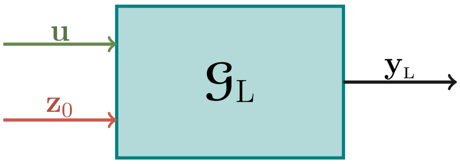


Figure 3.1: Structure of a first-order ODE system with a linear output.

for a vector $\zeta_0 \in \mathbb{R}^{N_{\mathbf{z}_0}}$, i.e., all possible initial states $\mathbf{z}_0$ lie in the space spanned by the matrix $\mathbf{Z}_0$. This assumption allows us to analyze all the initial conditions collectively.

In this section, we consider systems with a linear output equation and those with a quadratic one separately. First, in Section 3.1.1, systems with linear output equations are considered, where we mainly repeat the concepts presented in [15] and [66]. Afterwards, in Section 3.1.2, this theory is extended to systems with a quadratic output equation.

3.1.1 Inhomogeneous first-order ODE systems with a linear output

We consider the first-order system with a linear output equation of the form

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}(t) &= \mathcal{A}\mathbf{z}(t) + \mathcal{B}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0, \\ \mathbf{y}_{\mathrm{L}}(t) &= \mathcal{C}\mathbf{z}(t), \end{aligned} \tag{3.5}$$

including a state equation (3.1), an output matrix $\mathcal{C} \in \mathbb{R}^{q\times N}$, and an output $\mathbf{y}_{\mathrm{L}}(t) \in \mathbb{R}^q$. The corresponding system structure is depicted in Figure 3.1, where we see that the system, denoted by $\mathcal{G}_{\mathrm{L}}$, receives an input $\mathbf{u}$ and an initial state $\mathbf{z}_0$ to generate an output $\mathbf{y}_{\mathrm{L}}$.

We review two concepts that treat the inhomogeneous initial conditions in this subsection. The first one is explained in Section 3.1.1.1 and was introduced in [15], where the system (3.5) is decomposed into two subsystems, one including the input-to-output behavior and one including the initial condition-to-output behavior. These two systems are then analyzed separately. The second approach, discussed in Section 3.1.1.2, derives a surrogate model that incorporates the initial conditions into the input that is analyzed instead of the original system, see [66].

3.1.1.1 Multi-system approach for inhomogeneous first-order ODE systems with a linear output

We describe the multi-system approach, introduced in [15], where the authors utilize the superposition principle to derive two subsystems that describe the input- and initial condition-to-output behavior. This approach has the advantage that the subsystems can be analyzed and reduced separately. When applying reduction techniques, the reduced dimensions can, therefore, be chosen more flexibly so that the user can ensure that all the required information is preserved during the reduction, i.e., the reduced systems are accurate enough but also choose dimensions that are as small as possible. In the following, we repeat the steps of this approach, which will be extended to different system structures throughout this work.

Transfer function To derive an input-to-output mapping that describes the system dynamics, we investigate the system in its frequency-domain representation. Therefore, we consider the state components $\mathbf{z}_{\mathcal{B}}(t)$ and $\mathbf{z}_{\mathbf{Z}_0}(t)$ from (3.3) and apply the Laplace transform, which yields

$$\mathbf{Z}_{\mathcal{B}}(s) := (s\mathcal{E} - \mathcal{A})^{-1}\mathcal{B} \qquad \text{and} \qquad \mathbf{Z}_{\mathbf{Z}_0}(s) := (s\mathcal{E} - \mathcal{A})^{-1}\mathcal{E}\mathbf{Z}_0. \tag{3.6}$$

Applying the Laplace transform to the output $\mathbf{y}_{\mathrm{L}}(t)$ from (3.5) and inserting $\mathbf{Z}_{\mathcal{B}}(s)$ and $\mathbf{Z}_{\mathbf{Z}_0}(s)$ from (3.6) results in the output $\mathbf{Y}_{\mathrm{L}}(s) = \mathbf{Y}_{\mathrm{L},\mathcal{B}}(s) + \mathbf{Y}_{\mathrm{L},\mathbf{Z}_0}(s)$, where $\mathbf{Y}_{\mathrm{L}}$ is the Laplace transform of the linear output $\mathbf{y}_{\mathrm{L}}$ and

$$\mathbf{Y}_{\mathrm{L},\mathcal{B}}(s) := \mathcal{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathcal{B}\mathbf{U}(s) \qquad \text{and} \qquad \mathbf{Y}_{\mathrm{L},\mathbf{Z}_0}(s) := \mathcal{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathbf{Z}_0\zeta_0$$

are the two output components in the frequency domain that correspond to the input and the initial state, respectively. From these two outputs, we can extract the respective input- and initial condition-to-output mappings, leading to the following definition.

Definition 3.1:
Consider the asymptotically stable system (3.5) with an initial condition as defined in (3.4). Then the *transfer functions* corresponding to this system are defined as

$$\mathcal{G}_{\mathrm{L},\mathcal{B}}(s) := \mathcal{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathcal{B} \qquad \text{and} \qquad \mathcal{G}_{\mathrm{L},\mathbf{Z}_0}(s) := \mathcal{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathcal{E}\mathbf{Z}_0. \tag{3.7}$$

◊

The first transfer function $\mathcal{G}_{\mathrm{L},\mathcal{B}}(s)$ has the homogeneous system representation

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathcal{B}}(t) &= \mathcal{A}\mathbf{z}_{\mathcal{B}}(t) + \mathcal{B}\mathbf{u}(t), \qquad \mathbf{z}_{\mathcal{B}}(0) = 0, \\ \mathbf{y}_{\mathrm{L},\mathcal{B}}(t) &= \mathcal{C}\mathbf{z}_{\mathcal{B}}(t), \end{aligned} \tag{3.8}$$

which coincides with the system considered in (2.1). The second transfer function $\mathcal{G}_{\mathrm{L},\mathbf{Z}_0}(s)$ corresponds the system

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathbf{Z}_0}(t) &= \mathcal{A}\mathbf{z}_{\mathbf{Z}_0}(t), \qquad \mathbf{z}_{\mathbf{Z}_0}(0) = \mathbf{Z}_0\zeta_0, \\ \mathbf{y}_{\mathrm{L},\mathbf{Z}_0}(t) &= \mathcal{C}\mathbf{z}_{\mathbf{Z}_0}(t). \end{aligned} \tag{3.9}$$

As depicted in Figure 3.2, the sum of the two outputs coincides with the output of the original system. In the following, we make use of this superposition and study the two systems (3.8) and (3.9) separately. Therefore, we derive the respective controllability and observability Gramians that encode the controllability and observability behavior.

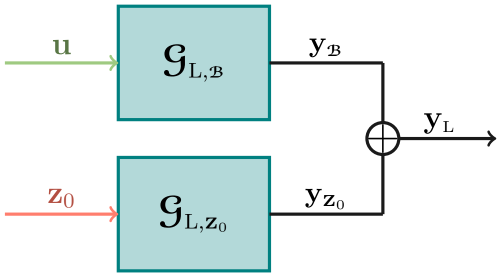

$\mathbf{u}$
$\mathcal{G}_{\mathrm{L},\mathcal{B}}$
$\mathbf{y}_{\mathcal{B}}$
$\mathbf{y}_{\mathrm{L}}$
$\mathbf{z}_0$
$\mathcal{G}_{\mathrm{L},\mathbf{z}_0}$
$\mathbf{y}_{\mathbf{z}_0}$

Figure 3.2: Structure of two separated first-order ODE systems with a linear output.

Controllability Gramian To describe the controllability behavior of the system (3.8), we consider its state equation with the respective input-to-state mapping that is given as

$$\boldsymbol{c}_{\mathcal{B}}(t) := e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{B}. \tag{3.10}$$

This mapping encodes all reachable states and can, therefore, be used to define a matrix $\mathcal{P}_{\mathcal{B}} := \int_0^\infty \boldsymbol{c}_{\mathcal{B}}(t)\boldsymbol{c}_{\mathcal{B}}(t)^{\mathrm{T}}\mathrm{d}t$ that spans the overall controllability space by integrating over the whole time domain.

Definition 3.2:
Consider the asymptotically stable system (3.8). The *controllability Gramian* is defined as

$$\mathcal{P}_{\mathcal{B}} := \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{B}\mathcal{B}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\mathrm{d}t. \tag{3.11}$$

◊

The controllability Gramian $\mathcal{P}_{\mathcal{B}}$ spans the controllability space of system (3.8), i.e., every reachable state of this system lies in the space, spanned by $\mathcal{P}_{\mathcal{B}}$. Consequently, if $\mathcal{P}_{\mathcal{B}}$ has full rank, then the system is controllable. As shown in (2.6), the Gramian $\mathcal{P}_{\mathcal{B}}$ is computed by solving the Lyapunov equation

$$\mathcal{E}\mathcal{P}_{\mathcal{B}}\mathcal{A}^{\mathrm{T}} + \mathcal{A}\mathcal{P}_{\mathcal{B}}\mathcal{E}^{\mathrm{T}} = -\mathcal{B}\mathcal{B}^{\mathrm{T}}. \tag{3.12}$$

Analogously, we extract the initial condition-to-state mapping from the system in (3.9) that is

$$\boldsymbol{c}_{\mathbf{z}_0}(t) := e^{\mathcal{E}^{-1}\mathcal{A}t}\mathbf{Z}_0. \tag{3.13}$$

This mapping encodes all reachable states resulting from initial conditions, lying in a space spanned by $\mathbf{Z}_0$. Hence, evaluating this mapping over the entire time domain leads to a matrix $\mathcal{P}_{\mathbf{z}_0} := \int_0^\infty \boldsymbol{c}_{\mathbf{z}_0}(t)\boldsymbol{c}_{\mathbf{z}_0}(t)^{\mathrm{T}}\mathrm{d}t$ that spans the respective reachability space.

Definition 3.3:
Consider the asymptotically stable system in (3.9). The corresponding *controllability Gramian* is defined as

$$\mathcal{P}_{\mathbf{Z}_0} := \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A}t}\mathbf{Z}_0\mathbf{Z}_0^{\mathrm{T}}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\mathrm{d}t. \tag{3.14}$$

◊

Again, we compute the Gramians $\mathcal{P}_{\mathbf{Z}_0}$ by solving the Lyapunov equation

$$\mathcal{E}\mathcal{P}_{\mathbf{Z}_0}\mathcal{A}^{\mathrm{T}} + \mathcal{A}\mathcal{P}_{\mathbf{Z}_0}\mathcal{E}^{\mathrm{T}} = -\mathcal{E}\mathbf{Z}_0\mathbf{Z}_0^{\mathrm{T}}\mathcal{E}^{\mathrm{T}}. \tag{3.15}$$

Observability Gramian To describe the observability behavior of the two subsystems (3.8) and (3.9), we first note that their observability behavior coincides. The corresponding state-to-output mapping is defined as

$$\boldsymbol{o}_{\mathrm{L}}(t) := \mathcal{C}e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}. \tag{3.16}$$

This mapping describes the observability behavior of the respective systems and is therefore used to define a matrix $\mathcal{Q}_{\mathrm{L}} := \int_0^\infty \boldsymbol{o}_{\mathrm{L}}(t)^{\mathrm{T}}\boldsymbol{o}_{\mathrm{L}}(t)\mathrm{d}t$ that spans the observability space.

Definition 3.4:
Consider the asymptotically stable systems (3.8) and (3.9). Then, the corresponding *observability Gramian* is defined as

$$\mathcal{Q}_{\mathrm{L}} := \int_0^\infty \mathcal{E}^{-\mathrm{T}}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\mathcal{C}^{\mathrm{T}}\mathcal{C}e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathrm{d}t. \tag{3.17}$$

◊

As shown in (2.6), one can compute the observability Gramian $\mathcal{Q}_{\mathrm{L}}$ by solving a Lyapunov equation

$$\mathcal{E}^{\mathrm{T}}\mathcal{Q}_{\mathrm{L}}\mathcal{A} + \mathcal{A}^{\mathrm{T}}\mathcal{Q}_{\mathrm{L}}\mathcal{E} = -\mathcal{C}^{\mathrm{T}}\mathcal{C}. \tag{3.18}$$

The Gramians introduced above describe the controllability and observability of the two subsystems (3.8) and (3.9). They are, therefore, used to identify states with significant influence on the system dynamics, which comprise the dominant controllability and observability subspaces. Therefore, we evaluate the system energies corresponding to the input- and initial condition-to-state mappings and the state-to-output mapping. We consider the controllability and observability energies separately.

Controllability Energies First, we analyze the controllability energies of the two subsystems (3.8) and (3.9). For that, we investigate the input-to-state mappings $\boldsymbol{c}_{\mathcal{B}}$ and $\boldsymbol{c}_{\mathbf{Z}_0}$ that are defined in (3.10) and (3.13), respectively. They describe the overall system behavior so that evaluating their energy norms leads to an energy measure that can be used to identify the most dominant controllability subspaces.

We define the energy norm of a function $\boldsymbol{c} \in L_2([0,\infty),\mathbb{R}^{N\times m})$ as

$$E(\boldsymbol{c}) := \|\boldsymbol{c}\|^2_{L_2([0,\infty),\mathbb{R}^{N\times m})} = \int_0^\infty \operatorname{tr}\big(\boldsymbol{c}(t)\boldsymbol{c}(t)^{\mathrm{T}}\big)\,\mathrm{d}t. \tag{3.19}$$

Accordingly, the energy norm encoding the controllability energy of subsystem (3.8) is given as

$$E(\boldsymbol{c}_{\mathcal{B}}) = \|\boldsymbol{c}_{\mathcal{B}}\|^2_{L_2([0,\infty),\mathbb{R}^{N\times m})} = \int_0^\infty \operatorname{tr}\big(\boldsymbol{c}_{\mathcal{B}}(t)\boldsymbol{c}_{\mathcal{B}}(t)^{\mathrm{T}}\big)\,\mathrm{d}t = \operatorname{tr}(\boldsymbol{\mathcal{P}}_{\mathcal{B}})\,. \tag{3.20}$$

On the other hand, the energy norm corresponding to system (3.9) is described by

$$E(\boldsymbol{c}_{\mathbf{Z}_0}) = \|\boldsymbol{c}_{\mathbf{Z}_0}\|^2_{L_2\left([0,\infty),\mathbb{R}^{N\times N_{\mathbf{Z}_0}}\right)} = \int_0^\infty \operatorname{tr}\big(\boldsymbol{c}_{\mathbf{Z}_0}(t)\boldsymbol{c}_{\mathbf{Z}_0}(t)^{\mathrm{T}}\big)\,\mathrm{d}t = \operatorname{tr}(\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0})\,. \tag{3.21}$$

For a symmetric matrix $\boldsymbol{\mathcal{P}}$ it holds that $\operatorname{tr}(\boldsymbol{\mathcal{P}}) = \sigma_1 + \dots + \sigma_N$ for the eigenvalues $\sigma_1 \geq \dots \geq \sigma_N \geq 0$ of $\boldsymbol{\mathcal{P}}$. Since the two Gramians $\boldsymbol{\mathcal{P}}_{\mathcal{B}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ are by definition symmetric, it follows from (3.20) and (3.21) that the largest eigenvalues of $\boldsymbol{\mathcal{P}}_{\mathcal{B}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ have the most effect on the system dynamics. Therefore, the states corresponding to the largest eigenvalues of $\boldsymbol{\mathcal{P}}_{\mathcal{B}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ span the most dominant controllability subspaces.

Observability energies In this paragraph, we aim to analyze the observability energies of the two subsystems (3.8) and (3.9) to identify the most observable states that span the dominant observability spaces. To provide an energy measure describing the observability properties of the two subsystems, we evaluate the energy norm of the state-to-output mapping $\boldsymbol{o}_{\mathrm{L}}$ defined in (3.16) according to (3.19), which is

$$E(\boldsymbol{o}_{\mathrm{L}}) = \|\boldsymbol{o}_{\mathrm{L}}\|^2_{L_2([0,\infty),\mathbb{R}^{p\times N})} = \int_0^\infty \operatorname{tr}\big(\boldsymbol{o}_{\mathrm{L}}(t)^{\mathrm{T}}\boldsymbol{o}_{\mathrm{L}}(t)\big)\,\mathrm{d}t = \operatorname{tr}(\boldsymbol{\mathcal{Q}}_{\mathrm{L}})\,. \tag{3.22}$$

Since the mapping $\boldsymbol{o}_{\mathrm{L}}$ encodes the observability of all states $\mathbf{z}(t)$, the evaluation of its energy norm describes the observability properties of the respective systems. Since the trace of the Gramian is equal to the sum of its eigenvalues, the states corresponding to the largest eigenvalues of $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ span the most dominant observability subspace. Moreover, the eigenvalues of the Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ that are small have the least influence on the system. Hence, the corresponding states are neglectable for the system dynamics and are truncated in the model reduction procedures.

In this paragraph, we derived two subsystems that encode the input- and initial condition-to-output behavior represented by their transfer functions. Moreover, we derived respective controllability and observability Gramians and the resulting energy norms, which are summarized Table 3.1.

	System (3.8)	System (3.9)
Transfer function	$\boldsymbol{\mathcal{G}}_{\mathrm{L},\boldsymbol{\mathcal{B}}}(s)$	$\boldsymbol{\mathcal{G}}_{\mathrm{L},\mathbf{Z}_0}(s)$
Controllability Gramian	$\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$	$\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$
Observability Gramian	$\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$	$\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$
Controllability energies	$E(\boldsymbol{c}_{\boldsymbol{\mathcal{B}}}) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}})$	$E(\boldsymbol{c}_{\mathbf{Z}_0}) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0})$
Observability energies	$E(\boldsymbol{o}_{\mathrm{L}}) = \mathrm{tr}(\boldsymbol{\mathcal{Q}}_{\mathrm{L}})$	$E(\boldsymbol{o}_{\mathrm{L}}) = \mathrm{tr}(\boldsymbol{\mathcal{Q}}_{\mathrm{L}})$

Table 3.1: Properties of system (3.5) corresponding to its multi-system representation.

3.1.1.2 Extended-input approach for inhomogeneous first-order ODE systems with a linear output

In this section, we consider a different approach to include the initial conditions in the analysis of the system dynamics. We describe the method presented in [66], where the authors add the initial conditions space to the input matrix. As a result, the respective initial conditions space is taken into account when describing the controllability space.

Transfer functions First, we consider the state of the system in the frequency domain. Therefore, we apply the Laplace transform to the state $\mathbf{z}(t)$ from (3.2), which yields

$$\mathbf{Z}(s) = (s\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\left(\boldsymbol{\mathcal{B}}\mathbf{U}(s) + \boldsymbol{\mathcal{E}}\mathbf{Z}_0\zeta_0\right) = (s\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{W}}\widetilde{\mathbf{U}}(s), \tag{3.23}$$

for an extended input matrix and an output defined as

$$\boldsymbol{\mathcal{W}} := \begin{bmatrix}\boldsymbol{\mathcal{B}} & \boldsymbol{\mathcal{E}}\mathbf{Z}_0\end{bmatrix} \qquad \text{and} \qquad \widetilde{\mathbf{U}}(s) := \begin{bmatrix}\mathbf{U}(s)\\ \zeta_0\end{bmatrix}, \tag{3.24}$$

respectively. Using the state expression from (3.23), we derive the input- and initial condition-to-output mapping of the original system (3.5). To do so, we apply the Laplace transform to the linear output equation in (3.5) and insert the state $\mathbf{Z}(s)$ from (3.23), which results in the output

$$\mathbf{Y}_{\mathrm{L}}(s) = \boldsymbol{\mathcal{C}}(s\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{W}}\widetilde{\mathbf{U}}(s). \tag{3.25}$$

We observe that the output $\mathbf{Y}_{\mathrm{L}}(s)$ is of the same structure as the output in (2.3) that describes a homogeneous system. Hence, the same theoretical considerations apply, while the matrix $\boldsymbol{\mathcal{W}}$ spans both the input space and the initial condition space. From the output $\mathbf{Y}_{\mathrm{L}}(s)$ in (3.25), we derive the input-to-output mapping as defined in the following.

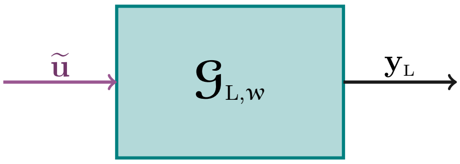


Figure 3.3: Structure of a first-order ODE system with an extended input and a linear output.

Definition 3.5:
Consider an asymptotically stable system (3.5) with initial conditions as defined in (3.4). Also consider the input matrix $\boldsymbol{\mathcal{W}}$ from (3.24). Then, the *transfer function* corresponding to that system is defined as

$$\mathcal{G}_{\mathrm{L},w}(s) := \boldsymbol{\mathcal{C}}(s\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{W}}. \tag{3.26}$$

◊

The transfer function defined in (3.26) satisfies $\mathbf{Y}_{\mathrm{L}}(s) = \mathcal{G}_{\mathrm{L}}(s)\widetilde{\mathbf{U}}(s)$ and, hence, encodes the system dynamics of the original system (3.5). Since the transfer function $\mathcal{G}_{\mathrm{L}}$ has multiple system representations, the authors in [66] derive the homogeneous system representation

$$\begin{aligned}\boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) &= \boldsymbol{\mathcal{A}}\mathbf{z}(t) + \boldsymbol{\mathcal{W}}\widetilde{\mathbf{u}}(t), \qquad \mathbf{z}(0) = 0,\\ \mathbf{y}_{\mathrm{L}}(t) &= \boldsymbol{\mathcal{C}}\mathbf{z}(t),\end{aligned} \tag{3.27}$$

where $\widetilde{\mathbf{u}} \in L_2([0,\infty),\mathbb{R}^m)$ is assumed to be a suitable output. The structure of the surrogate system (3.27) is depicted in Figure 3.3, where we see that only one input enters the system as it includes the initial conditions.

Instead of investigating the inhomogeneous system (3.5), in the following, we analyze the homogeneous surrogate model (3.27) and apply the system theoretical concepts from Section 2.1.1 for homogeneous systems. Note that the surrogate system is only used to derive controllability spaces and corresponding Gramians that incorporate the influence of initial conditions on the system. However, to derive a surrogate model of a smaller dimension, later in this work, we apply reduction techniques to the original system (3.5) utilizing the controllability spaces derived in this section.

Controllability Gramian To investigate the controllability properties of the surrogate system (3.27), in this paragraph, we aim to derive the controllability Gramian which spans the respective controllability space, i.e., the space in which all reachable states lie. For that, we extract an input-to-state mapping of system (3.27) that is

$$\boldsymbol{c}_{\mathrm{W}}(t) := e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{W}}. \tag{3.28}$$

Since this mapping encodes all reachable states, it is used to derive a matrix $\boldsymbol{\mathcal{P}}_{\mathrm{W}} := \int_0^\infty \boldsymbol{c}_{\mathrm{W}}(t)\boldsymbol{c}_{\mathrm{W}}(t)^{\mathrm{T}}\mathrm{d}t$ that spans the respective controllability space.

Definition 3.6:
Consider the asymptotically stable system (3.27) with $\mathcal{W}$ as defined in (3.24). Then the corresponding *controllability Gramian* is defined as

$$\mathcal{P}_{\mathcal{W}} := \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{W}\mathcal{W}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\mathrm{d}t. \tag{3.29}$$

◊

The Gramian $\mathcal{P}_{\mathcal{W}}$ spans the controllability space of system (3.27), i.e., all controllable states lie in the space spanned by $\mathcal{P}_{\mathcal{W}}$. Hence, if the controllability Gramian has full rank, the respective system (3.27) is controllable. As explained in Section 2.1.1, the controllability Gramian can be computed by solving a Lyapunov equation of the form (2.6) that is

$$\mathcal{A}\mathcal{P}_{\mathcal{W}}\mathcal{E}^{\mathrm{T}} + \mathcal{E}\mathcal{P}_{\mathcal{W}}\mathcal{A}^{\mathrm{T}} = -\mathcal{W}\mathcal{W}^{\mathrm{T}}.$$

To solve the Lyapunov equation, we use the methods presented in Section 2.3.

Finally, we describe the connection between the Gramian $\mathcal{P}_{\mathcal{W}}$, and the Gramians $\mathcal{P}_{\mathcal{B}}$ and $\mathcal{P}_{\mathbf{Z}_0}$ that result from the extended-input approach and the multi-system approach, respectively.

Theorem 3.7:
Consider the asymptotically stable system (3.27) with the corresponding controllability Gramian $\mathcal{P}_{\mathcal{W}}$ from (3.11). Also, consider the asymptotically stable systems (3.8) and (3.9) with the controllability Gramians $\mathcal{P}_{\mathcal{B}}$ and $\mathcal{P}_{\mathbf{Z}_0}$ defined in (3.11) and (3.14), respectively. Then the following relation holds

$$\mathcal{P}_{\mathcal{W}} = \mathcal{P}_{\mathcal{B}} + \mathcal{P}_{\mathbf{Z}_0}.$$

◊

Proof. We insert the definition of $\mathcal{W}$ into the definition of $\mathcal{P}_{\mathcal{W}}$ to obtain

$$\begin{aligned}
\mathcal{P}_{\mathcal{W}} &= \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\begin{bmatrix}\mathcal{B} & \mathcal{E}\mathbf{Z}_0\end{bmatrix}\begin{bmatrix}\mathcal{B}^{\mathrm{T}} \\ \mathbf{Z}_0^{\mathrm{T}}\mathcal{E}^{\mathrm{T}}\end{bmatrix}\mathcal{E}^{-\mathrm{T}}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\mathrm{d}t \\
&= \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A}t}\left(\mathcal{E}^{-1}\mathcal{B}\mathcal{B}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}} + \mathbf{Z}_0\mathbf{Z}_0^{\mathrm{T}}\right)e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\mathrm{d}t \\
&= \mathcal{P}_{\mathcal{B}} + \mathcal{P}_{\mathbf{Z}_0}.
\end{aligned}$$

□

Observability Gramian We aim to describe the observability behavior of the surrogate system (3.27). However, we observe that the state-to-output mapping of system (3.27) coincides with the mapping of the two subsystems (3.8) and (3.9). Hence, the same observability Gramian defined in (3.17) encodes the observability behavior of the surrogate system (3.27) as stated in the following definition.

Definition 3.8:
Consider the asymptotically stable system (3.27). The corresponding *observability Gramian* is defined as

$$\mathcal{Q}_{\mathrm{L}} = \int_0^\infty \boldsymbol{\mathcal{E}}^{-\mathrm{T}} e^{(\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}})^{\mathrm{T}}t} \boldsymbol{\mathcal{C}}^{\mathrm{T}} \boldsymbol{\mathcal{C}} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t} \boldsymbol{\mathcal{E}}^{-1} \mathrm{d}t. \qquad \Diamond$$

As described in (3.18), the observability Gramian is computed by solving a Lyapunov equation.

Controllability energies In this paragraph, we derive the controllability energies corresponding to the surrogate system (3.27) including the input-to-state and the initial condition-to-state behavior of the original system (3.5). To do so, we derive an energy measure that describes the controllability behavior of the system (3.27). As energy measure, we use the energy norm introduced in (3.19). We evaluate the energy norm of the input-to-state mapping $\boldsymbol{c}_{\mathrm{W}}$ from (3.28), which yields

$$E(\boldsymbol{c}_{\mathrm{W}}) = \|\boldsymbol{c}_{\mathrm{W}}\|^2_{L_2\left([0,\infty),\mathbb{R}^{N\times(m+N_{\mathbf{Z}_0})}\right)} = \int_0^\infty \mathrm{tr}\left(\boldsymbol{c}_{\mathrm{W}}(t)\boldsymbol{c}_{\mathrm{W}}(t)^{\mathrm{T}}\right) \mathrm{d}t = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{W}}). \tag{3.30}$$

The trace of the Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{W}}$ coincides with the sum of its eigenvalues. Hence, large eigenvalues have a high influence on the system's energy, while small eigenvalues are neglectable. Therefore, the states corresponding to the highest eigenvalues span the most dominant subspaces.

Observability energies As described above, the observability behavior of system (3.27) is equal to the ones of the subsystems (3.8) and (3.9). Hence, we describe the observability energies using the observability Gramian $\mathcal{Q}_{\mathrm{L}}$. The energy norm of the state-to-output mapping $\boldsymbol{o}_{\mathrm{L}}$ from (3.16) is equal to

$$E(\boldsymbol{o}_{\mathrm{L}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{L}}),$$

as shown in (3.22). From this energy measure, we follow that states corresponding to large eigenvalues of the observability Gramian $\mathcal{Q}_{\mathrm{L}}$ span the dominant observability subspaces. On the other hand, states corresponding to small eigenvalues are neglectable when describing the system dynamics.

In the extended-input approach, we derive a model incorporating the initial condition space into the input expression. That way, we can derive suitable Gramians and energies that encode the system properties that are summarized in Table 3.2.

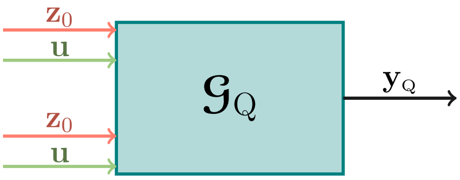


Figure 3.4: Structure of a first-order ODE system with a quadratic output.

	System (3.27)
Transfer function	$\mathcal{G}_{\mathrm{L},w}$
Controllability Gramian	$\mathcal{P}_w$
Observability Gramian	$\mathcal{Q}_{\mathrm{L}}$
Controllability energies	$E(\boldsymbol{c}_w) = \operatorname{tr}(\mathcal{P}_w)$
Observability energies	$E(\boldsymbol{o}_{\mathrm{L}}) = \operatorname{tr}(\mathcal{Q}_{\mathrm{L}})$

Table 3.2: Properties of system (3.5) corresponding to its extended-input representation.

3.1.2 Inhomogeneous first-order ODE systems with a quadratic output

In this subsection, we study first-order systems of the form

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}(t) &= \mathcal{A}\mathbf{z}(t) + \mathcal{B}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0, \\ \mathbf{y}_{\mathrm{Q}}(t) &= \mathbf{z}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}(t), \end{aligned} \tag{3.31}$$

with a state equation as defined in (3.1), and a quadratic output equation with a symmetric output matrix $\mathcal{M} \in \mathbb{R}^{N\times N}$ and an output $\mathbf{y}_{\mathrm{Q}}(t) \in \mathbb{R}$. The resulting system (3.31) is depicted in Figure 3.4, where we insert two inputs $\mathbf{u}$ and an initial condition $\mathbf{z}_0$ into the system, denoted $\mathcal{G}_{\mathrm{Q}}$, to indicate the quadratic output equation. The system (3.31) occurs, particularly in the study of the variance, or deviation, of the state variable from a given reference point, which can be represented as a quadratic function of the state. Quadratic output equations also arise when, e.g., evaluating system energies as output variables.

The authors in [20] derived concepts to evaluate quadratic output equations. However, they only consider systems with homogeneous initial conditions. Hence, in this section, we extend the approaches of [66] and [15] to systems with quadratic output equations (3.31) using the ideas from [20]. Therefore, in Section 3.1.2.1, we apply the superposition principles to derive four subsystems that describe the overall system behavior and are analyzed separately. Afterwards, in Section 3.1.2.2, we apply an extended-input approach that results in a homogeneous system with an input matrix that includes the input and the initial condition space. This system is then analyzed instead of the original system (3.31).

3.1.2.1 Multi-system approach for inhomogeneous first-order ODE systems with a quadratic output

In this section, we apply the multi-system approach to analyze the system (3.31) while taking into account the inhomogeneous initial conditions. We apply the superposition principles to derive subsystems with outputs that sum up to the original output expression $\mathbf{y}_{\mathrm{Q}}(t)$. By considering the subsystems individually, we gain more flexibility in analyzing the influences of the initial conditions $\mathbf{z}_0$ and inputs $\mathbf{u}(t)$ to reduce the respective subsystems later in this work.

We consider the state $\mathbf{z}(t) = \mathbf{z}_{\mathcal{B}}(t) + \mathbf{z}_{\mathbf{z}_0}(t)$ from (3.2) that consist of two components defined in (3.3). Using these components, the output equation of system (3.5) can be written as

$$\begin{aligned}\mathbf{y}_{\mathrm{Q}}(t) &= \mathbf{z}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}(t)\\ &= \mathbf{z}_{\mathcal{B}}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathcal{B}}(t) + \mathbf{z}_{\mathbf{z}_0}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathcal{B}}(t) + \mathbf{z}_{\mathcal{B}}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathbf{z}_0}(t) + \mathbf{z}_{\mathbf{z}_0}(s)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathbf{z}_0}(t) \qquad (3.32)\\ &=: \mathbf{y}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t) + \mathbf{y}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}(t) + \mathbf{y}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}(t) + \mathbf{y}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0}(t),\end{aligned}$$

where we identify four output components, as the sketch in Figure 3.5 shows. We note that the terms $\mathbf{y}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}(t)$ and $\mathbf{y}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}(t)$ coincide because of the symmetry of $\mathcal{M}$. However, to describe the output behavior, we need to analyze both components separately, as they describe different observability spaces.

Transfer function To study the behavior of the system (3.31) concerning the initial conditions and the input, we consider the four output components defined in (3.32) separately.

Inserting $\mathbf{z}_{\mathcal{B}}(t)$ from (3.3) into the first output component $\mathbf{y}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t)$ from (3.32) yields

$$\mathbf{y}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t) = \int_0^t\int_0^t \mathbf{u}(\tau_1)^{\mathrm{T}}\mathcal{B}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}(t-\tau_1)}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}(t-\tau_2)}\mathcal{E}^{-1}\mathcal{B}\mathbf{u}(\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2.$$

From this output expression, we can extract the kernel

$$\mathbf{g}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t_1,t_2) := \mathcal{B}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t_1}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}t_2}\mathcal{E}^{-1}\mathcal{B}.$$

Since the kernel $\mathbf{g}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t_1,t_2)$ encodes the input-to-output mapping, it is used to describe the respective system behavior. Therefore, we analyze the respective system dynamics in the frequency domain using the 2-dimensional Laplace transform.

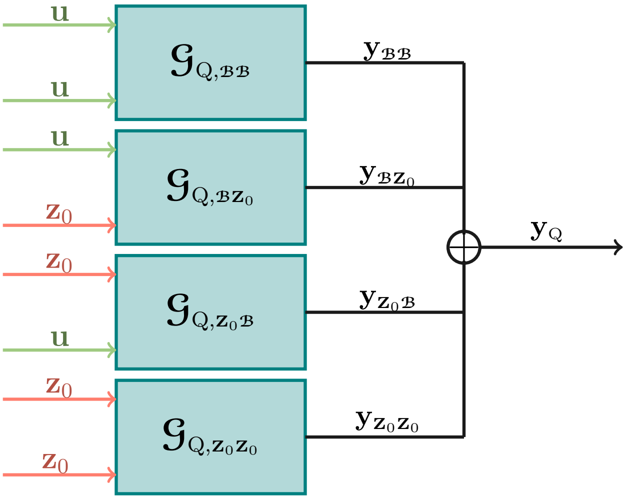

$\mathbf{u}$
$\mathbf{u}$
$\mathbf{u}$
$\mathbf{z}_0$
$\mathbf{z}_0$
$\mathbf{u}$
$\mathbf{z}_0$
$\mathbf{z}_0$
$\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathcal{B}}$
$\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}$
$\mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}$
$\mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0}$
$\mathbf{y}_{\mathcal{B}\mathcal{B}}$
$\mathbf{y}_{\mathcal{B}\mathbf{z}_0}$
$\mathbf{y}_{\mathbf{z}_0\mathcal{B}}$
$\mathbf{y}_{\mathbf{z}_0\mathbf{z}_0}$
$\mathbf{y}_{\mathrm{Q}}$

Figure 3.5: Structure of four separated first-order ODE systems with a quadratic output.

Definition 3.9:
Let $f(t_1,t_2) : [0,\infty)^2 \to \mathbb{R}^n$ be a function that is exponentially bounded, i.e., there exist numbers M and α so that

$$\|f(t_1,t_2)\|_2 \leq Me^{\alpha t_1} \quad \text{and} \quad \|f(t_1,t_2)\|_2 \leq Me^{\alpha t_2}, \quad \text{for all } t_1,t_2 \geq 0.$$

Then the 2-*dimensional Laplace transform* is defined as

$$F(s_1,s_2) := \mathcal{L}\{f\}(s_1,s_2) := \int_0^\infty \int_0^\infty e^{-s_2t_2-s_1t_1} f(t_1,t_2)\mathrm{d}t_1\mathrm{d}t_2. \qquad \Diamond$$

Applying the 2-dimensional Laplace transform to $\mathbf{g}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t_1,t_2)$ leads to the input-to-output mapping

$$\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(s_1,s_2) := \mathcal{B}^{\mathrm{T}}(s_1\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathcal{M}(s_2\mathcal{E}-\mathcal{A})^{-1}\mathcal{B}$$

in the frequency domain, which is the transfer function corresponding to $\mathbf{y}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t)$.

Using the same procedure, we derive the transfer functions for the remaining output components in (3.32), which yields the following definition.

Definition 3.10:
Consider the asymptotically stable dynamical system (3.31) with an initial condition as defined in (3.4). Then the four *transfer functions* corresponding to that system are defined as

$$\begin{aligned}
\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(s_1,s_2) &:= \mathcal{B}^{\mathrm{T}}(s_1\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathcal{M}(s_2\mathcal{E}-\mathcal{A})^{-1}\mathcal{B},\\
\mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}(s_1,s_2) &:= \mathbf{Z}_0^{\mathrm{T}}\mathcal{E}^{\mathrm{T}}(s_1\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathcal{M}(s_2\mathcal{E}-\mathcal{A})^{-1}\mathcal{B},\\
\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}(s_1,s_2) &:= \mathcal{B}^{\mathrm{T}}(s_1\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathcal{M}(s_2\mathcal{E}-\mathcal{A})^{-1}\mathcal{E}\mathbf{Z}_0,\\
\mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0}(s_1,s_2) &:= \mathbf{Z}_0^{\mathrm{T}}\mathcal{E}^{\mathrm{T}}(s_1\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathcal{M}(s_2\mathcal{E}-\mathcal{A})^{-1}\mathcal{E}\mathbf{Z}_0.
\end{aligned} \tag{3.33}$$

$\Diamond$

The first transfer function $\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(s_1, s_2)$ corresponding to the first output component $\mathbf{y}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t)$ has the following homogeneous system realization

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathcal{B}}(t) &= \mathcal{A}\mathbf{z}_{\mathcal{B}}(t) + \mathcal{B}\mathbf{u}(t), \qquad \mathbf{z}_{\mathcal{B}}(0) = 0, \\ \mathbf{y}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(t) &= \mathbf{z}_{\mathcal{B}}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathcal{B}}(t). \end{aligned} \tag{3.34}$$

This system has the same structure analyzed in [20]. Therefore, we can treat it the same way. The two transfer functions $\mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}(s_1, s_2)$ and $\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}(s_1, s_2)$ each include an input-to-state mapping and an initial condition-to-state mapping. Consequently, two state equations are needed to define the respective system realizations, which are

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathcal{B}}(t) &= \mathcal{A}\mathbf{z}_{\mathcal{B}}(t) + \mathcal{B}\mathbf{u}(t), \qquad && \mathbf{z}_{\mathcal{B}}(0) = 0, \\ \mathcal{E}\dot{\mathbf{z}}_{\mathbf{z}_0}(t) &= \mathcal{A}\mathbf{z}_{\mathbf{z}_0}(t), && \mathbf{z}_{\mathbf{z}_0}(0) = \mathbf{Z}_0\zeta_0, \\ \mathbf{y}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}(t) &= \mathbf{z}_{\mathbf{z}_0}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathcal{B}}(t), \end{aligned} \tag{3.35}$$

and

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathcal{B}}(t) &= \mathcal{A}\mathbf{z}_{\mathcal{B}}(t) + \mathcal{B}\mathbf{u}(t), \qquad && \mathbf{z}_{\mathcal{B}}(0) = 0, \\ \mathcal{E}\dot{\mathbf{z}}_{\mathbf{z}_0}(t) &= \mathcal{A}\mathbf{z}_{\mathbf{z}_0}(t), && \mathbf{z}_{\mathbf{z}_0}(0) = \mathbf{Z}_0\zeta_0, \\ \mathbf{y}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}(t) &= \mathbf{z}_{\mathcal{B}}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathbf{z}_0}(t). \end{aligned} \tag{3.36}$$

We observe that both systems lead to the same output $\mathbf{y}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}(t) = \mathbf{y}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}(t)$ as the matrix $\mathcal{M}$ is assumed to be symmetric. However, for consideration later in this work, we distinguish between them. The remaining transfer function $\mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0}(s_1, s_2)$ that generates the output component $\mathbf{y}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0}(t)$ corresponds to the system realization

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathbf{z}_0}(t) &= \mathcal{A}\mathbf{z}_{\mathbf{z}_0}(t), \qquad \mathbf{z}_{\mathbf{z}_0}(0) = \mathbf{Z}_0\zeta_0, \\ \mathbf{y}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0} &= \mathbf{z}_{\mathbf{z}_0}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathbf{z}_0}(t). \end{aligned} \tag{3.37}$$

We observe that no input is acting on the system as the system behavior only depends on the initial condition $\mathbf{z}_0$.

In the following, we investigate the four subsystems separately instead of analyzing the inhomogeneous system (3.31) to describe the overall system behavior. These considerations are used later in Section 4.1.1 to reduce all the subsystems separately within a model reduction scheme.

Controllability Gramians In this paragraph, we aim to derive controllability Gramians encoding the controllability properties to the four subsystems (3.34), (3.35), (3.36), and (3.37). We observe that only two different state equations appear within the four subsystems that also coincide with the state equations of the systems (3.10) and (3.13). Also their input- and initial condition-to-output mappings corresponding to the input $\mathcal{B}\mathbf{u}(t)$ and the initial condition $\mathbf{Z}_0\zeta_0$ that are

$$\boldsymbol{c}_{\mathcal{B}}(t) := e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{B} \qquad \text{and} \qquad \boldsymbol{c}_{\mathbf{z}_0}(t) := e^{\mathcal{E}^{-1}\mathcal{A}t}\mathbf{Z}_0,$$

respectively, coincide with those from (3.10) and (3.13). Hence, evaluating these mappings over the time domain leads to the respective controllability Gramians as defined in (3.11) and (3.14).

Definition 3.11:
Consider the asymptotically stable systems (3.34), (3.35), (3.36), and (3.37). The respective *controllability Gramians* are defined as

$$\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}} := \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{(\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}})^{\mathrm{T}}t}\mathrm{d}t, \qquad \boldsymbol{\mathcal{P}}_{\mathbf{Z}_0} := \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\mathbf{Z}_0\mathbf{Z}_0^{\mathrm{T}}e^{(\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}})^{\mathrm{T}}t}\mathrm{d}t. \ \diamond$$

These Gramians are determined by solving Lyapunov equations (3.15) and (3.18).

Observability Gramian We aim to determine tailored Gramians encoding observability subspaces for ODE systems with quadratic output equations (3.31) that describe their observability properties. However, extensions of the Gramian definitions for systems with linear output equations to systems with quadratic output equations are not straightforward. Hence, in this paragraph, we propose new Gramians that describe the observability based on the output decomposition (3.32). Therefore, we investigate the four output components separately. Since the output is a superposition of the four components, the Gramians that describe the output components sum up to a Gramian that describes the overall observability of the system.

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

$$\mathbf{y}_{\mathrm{Q}}(t) = \boldsymbol{\mathcal{C}}(\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t))\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t) + \boldsymbol{\mathcal{C}}(\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t))\mathbf{z}_{\mathbf{Z}_0}(t) + \boldsymbol{\mathcal{C}}(\mathbf{z}_{\mathbf{Z}_0}(t))\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t) + \boldsymbol{\mathcal{C}}(\mathbf{z}_{\mathbf{Z}_0}(t))\mathbf{z}_{\mathbf{Z}_0}(t).$$

We observe that the observability of the state $\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t)$ in the output $\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t) = \boldsymbol{\mathcal{C}}(\mathbf{z}_{\mathbf{Z}_0}(t))\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t)$ also depends on the reachability of $\mathbf{z}_{\mathbf{Z}_0}(t)$. On the other hand, the observability of the state $\mathbf{z}_{\mathbf{Z}_0}(t)$ corresponding to $\mathbf{y}_{\mathrm{Q},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}(t) = \boldsymbol{\mathcal{C}}(\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t))\mathbf{z}_{\mathbf{Z}_0}(t)$ depends on the reachability of $\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t)$. Hence, the outputs $\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t) = \mathbf{y}_{\mathrm{Q},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}(t)$ encode two different observability properties. Analogously, the outputs $\mathbf{y}_{\mathrm{Q},\boldsymbol{\mathcal{B}}}(t) = \boldsymbol{\mathcal{C}}(\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t))\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t)$ and $\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0}(t) = \boldsymbol{\mathcal{C}}(\mathbf{z}_{\mathbf{Z}_0}(t))\mathbf{z}_{\mathbf{Z}_0}(t)$ encode the observability of the state $\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t)$ depending on the reachability of the same, and the observability of the state $\mathbf{z}_{\mathbf{Z}_0}(t)$ depending on the reachability of the same state $\mathbf{z}_{\mathbf{Z}_0}(t)$, respectively.

In this paragraph, we define observability Gramians encoding the observability behavior of state the $\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t)$ corresponding to $\boldsymbol{\mathcal{C}}(\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t))$ and $\boldsymbol{\mathcal{C}}(\mathbf{z}_{\mathbf{Z}_0}(t))$ and observability Gramians describing the observability of the state $\mathbf{z}_{\mathbf{Z}_0}(t)$ corresponding to $\boldsymbol{\mathcal{C}}(\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t))$ and $\boldsymbol{\mathcal{C}}(\mathbf{z}_{\mathbf{Z}_0}(t))$. Because of the dependencies on the reachability of $\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t)$ and $\mathbf{z}_{\mathbf{Z}_0}(t)$ encoded by $\boldsymbol{\mathcal{C}}(\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t))$ and $\boldsymbol{\mathcal{C}}(\mathbf{z}_{\mathbf{Z}_0}(t))$, we expect that the observability Gramians will depend on the controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$.

The first subsystem (3.34) encodes the input-to-output behavior from which we extract the input-to-state mapping $\boldsymbol{c}_{\mathcal{B}}(t) = e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{B}$ defined in (3.10). The remaining mapping encodes the state-to-output mapping that is

$$\boldsymbol{o}_{\mathrm{Q},\mathcal{B}}(t_1, t_2) = \mathcal{B}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t_1}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}t_2}\mathcal{E}^{-1}. \tag{3.38}$$

The output of the third subsystem (3.36) includes the input-to-state mapping $\boldsymbol{c}_{\mathbf{Z}_0}(t) = e^{\mathcal{E}^{-1}\mathcal{A}t}\mathbf{Z}_0$ from (3.13) so that the remaining state-to-output mapping is equal to $\boldsymbol{o}_{\mathrm{Q},\mathcal{B}}$ as defined in (3.38). Using the mapping $\boldsymbol{o}_{\mathrm{Q},\mathcal{B}}$, we can construct a matrix $\mathcal{Q}_{\mathrm{Q},\mathcal{B}}$ that spans the respective observability space by integrating over the time domain, which yields

$$\begin{aligned}
\mathcal{Q}_{\mathrm{Q},\mathcal{B}} &:= \int_0^\infty \int_0^\infty \boldsymbol{o}_{\mathrm{Q},\mathcal{B}}(t_1, t_2)^{\mathrm{T}}\boldsymbol{o}_{\mathrm{Q},\mathcal{B}}(t_1, t_2)\mathrm{d}t_1\mathrm{d}t_2 \\
&= \int_0^\infty \int_0^\infty \mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t_2}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}t_1}\mathcal{E}^{-1}\mathcal{B}\mathcal{B}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t_1}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}t_2}\mathcal{E}^{-1}\mathrm{d}t_1\mathrm{d}t_2 \\
&= \int_0^\infty \mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t_2}\mathcal{M}\mathcal{P}_{\mathcal{B}}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}t_2}\mathcal{E}^{-1}\mathrm{d}t_2
\end{aligned}$$

according to the definition of $\mathcal{P}_{\mathcal{B}}$ in (3.11). This consideration is summarized in the following definition.

Definition 3.12:
We consider the asymptotically stable systems (3.34) and (3.36) and the controllability Gramian $\mathcal{P}_{\mathcal{B}}$ as defined in (3.11). Then the *observability Gramian* corresponding to those systems is defined as

$$\mathcal{Q}_{\mathrm{Q},\mathcal{B}} := \int_0^\infty \mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t}\mathcal{M}\mathcal{P}_{\mathcal{B}}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathrm{d}t. \tag{3.39}$$

◊

We compute the observability Gramian $\mathcal{Q}_{\mathrm{Q},\mathcal{B}}$ by solving the Lyapunov equation

$$\mathcal{E}^{\mathrm{T}}\mathcal{Q}_{\mathrm{Q},\mathcal{B}}\mathcal{A} + \mathcal{A}^{\mathrm{T}}\mathcal{Q}_{\mathrm{Q},\mathcal{B}}\mathcal{E} = -\mathcal{M}\mathcal{P}_{\mathcal{B}}\mathcal{M}.$$

Since we investigate the controllability and observability behavior of the right state in the different output components in (3.32), the behavior of the system (3.34) is described by the two Gramians $\mathcal{P}_{\mathcal{B}}$ and $\mathcal{Q}_{\mathrm{Q},\mathcal{B}}$. Moreover, the behavior of the system (3.36) is described by the two Gramians $\mathcal{P}_{\mathbf{Z}_0}$ and $\mathcal{Q}_{\mathrm{Q},\mathcal{B}}$. Note that we can also investigate the controllability and observability behavior of the left states of the quadratic output expressions, which would lead to similar results, where the Gramians from the two outputs $\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0\mathcal{B}}$ and $\mathbf{y}_{\mathrm{Q},\mathcal{B}\mathbf{Z}_0}$ are swapped.

Analogously, we investigate the observability of the subsystem (3.35). We extract the input-to-state mapping $\boldsymbol{c}_{\mathcal{B}}(t)$ from (3.10) and the remaining state-to-output mapping that is defined as

$$\boldsymbol{o}_{\mathrm{Q},\mathbf{Z}_0}(t_1, t_2) = \mathbf{Z}_0^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}t_1}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}t_2}\mathcal{E}^{-1}. \tag{3.40}$$

The subsystem (3.37) also results in the state-to-output mapping $\boldsymbol{o}_{\mathrm{Q},\mathbf{Z}_0}$ as defined in (3.40) after the respective input-to-state mapping $\boldsymbol{c}_{\mathbf{Z}_0}(t)$ from (3.13) is identified. The mapping $\boldsymbol{o}_{\mathrm{Q},\mathbf{Z}_0}$ is used to derive a matrix that spans the observability space as

$$\begin{aligned}
\boldsymbol{o}_{\mathrm{Q},\mathbf{Z}_0}(t_1, t_2) &= \int_0^\infty \int_0^\infty \boldsymbol{o}_{\mathrm{Q},\mathbf{Z}_0}(t_1, t_2) \boldsymbol{o}_{\mathrm{Q},\mathbf{Z}_0}(t_1, t_2) \mathrm{d}t_1 \mathrm{d}t_2 \\
&= \int_0^\infty \int_0^\infty \boldsymbol{\mathcal{E}}^{-\mathrm{T}} e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}} \boldsymbol{\mathcal{E}}^{-\mathrm{T}} t} \boldsymbol{\mathcal{M}} e^{\boldsymbol{\mathcal{E}}^{-1} \boldsymbol{\mathcal{A}} t_1} \mathbf{Z}_0 \mathbf{Z}_0^{\mathrm{T}} e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}} \boldsymbol{\mathcal{E}}^{-\mathrm{T}} t_1} \boldsymbol{\mathcal{M}} e^{\boldsymbol{\mathcal{E}}^{-1} \boldsymbol{\mathcal{A}} t} \boldsymbol{\mathcal{E}}^{-1} \mathrm{d}t_1 \mathrm{d}t_2 \\
&= \int_0^\infty \boldsymbol{\mathcal{E}}^{-\mathrm{T}} e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}} \boldsymbol{\mathcal{E}}^{-\mathrm{T}} t_2} \boldsymbol{\mathcal{M}} \boldsymbol{\mathcal{P}}_{\mathbf{Z}_0} \boldsymbol{\mathcal{M}} e^{\boldsymbol{\mathcal{E}}^{-1} \boldsymbol{\mathcal{A}} t_2} \boldsymbol{\mathcal{E}}^{-1} \mathrm{d}t_2
\end{aligned}$$

by inserting the definition of $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ from (3.14).

Definition 3.13:
We consider the asymptotically stable systems (3.35) and (3.37) and the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ as defined in (3.14). Then the corresponding *observability Gramian* is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{Q},\mathbf{Z}_0} := \int_0^\infty \boldsymbol{\mathcal{E}}^{-\mathrm{T}} e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}} \boldsymbol{\mathcal{E}}^{-\mathrm{T}} t} \boldsymbol{\mathcal{M}} \boldsymbol{\mathcal{P}}_{\mathbf{Z}_0} \boldsymbol{\mathcal{M}} e^{\boldsymbol{\mathcal{E}}^{-1} \boldsymbol{\mathcal{A}} t} \boldsymbol{\mathcal{E}}^{-1} \mathrm{d}t. \tag{3.41}$$

◊

The observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{Q},\mathbf{Z}_0}$ is the unique solution of the Lyapunov equation

$$\boldsymbol{\mathcal{E}}^{\mathrm{T}} \boldsymbol{\mathcal{Q}}_{\mathrm{Q},\mathbf{Z}_0} \boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{A}}^{\mathrm{T}} \boldsymbol{\mathcal{Q}}_{\mathrm{Q},\mathbf{Z}_0} \boldsymbol{\mathcal{E}} = -\boldsymbol{\mathcal{M}} \boldsymbol{\mathcal{P}}_{\mathbf{Z}_0} \boldsymbol{\mathcal{M}}.$$

Controllability Energy In this paragraph, we analyze the controllability energies of the four subsystems (3.34), (3.35), (3.36), and (3.37) to describe their controllability properties and identify the dominant controllability subspaces accordingly. Since the controllability behavior of the different state equations within the four subsystems coincides with those of the two subsystems (3.8) and (3.9), we obtain equal energy expressions. Therefore, we analyze the energy norm of the different input- and initial condition-to-state mappings $\boldsymbol{c}_{\boldsymbol{\mathcal{B}}}(t)$ and $\boldsymbol{c}_{\mathbf{Z}_0}(t)$ from (3.10) and (3.13), respectively, which leads to the energy norms

$$E(\boldsymbol{c}_{\boldsymbol{\mathcal{B}}}) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}) \qquad \text{and} \qquad E(\boldsymbol{c}_{\mathbf{Z}_0}) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0})$$

as described in (3.20) and (3.21). Since the traces of Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ are equal to the sum of their eigenvalues, the states corresponding to the large eigenvalues of the controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ span the most dominant controllability subspaces since the Gramians are by definition symmetric.

Observability energy Analogous to the controllability energies, we want to analyze the observability energies to identify the states that encode the dominant observability subspaces. To this end, we consider the state-to-output mappings from (3.38) and (3.40) describing the observability behavior of the different subsystems with a quadratic output equation. We apply the respective energy norm which yields the energy expressions

$$\begin{aligned} E(\boldsymbol{o}_{\mathrm{Q},\boldsymbol{\mathfrak{B}}}) &= \|\boldsymbol{o}_{\mathrm{Q},\boldsymbol{\mathfrak{B}}}\|^2_{L_2([0,\infty)^2,\mathbb{R}^{m\times N})} = \int_0^\infty \int_0^\infty \operatorname{tr}\big(\boldsymbol{o}_{\mathrm{Q},\boldsymbol{\mathfrak{B}}}(t_1,t_2)^{\mathrm{H}}\boldsymbol{o}_{\mathrm{Q},\boldsymbol{\mathfrak{B}}}(t_1,t_2)\big)\,\mathrm{d}t_1\mathrm{d}t_2 \\ &= \operatorname{tr}(\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\boldsymbol{\mathfrak{B}}}) \end{aligned} \tag{3.42}$$

and

$$\begin{aligned} E(\boldsymbol{o}_{\mathrm{Q},\mathbf{z}_0}) &= \|\boldsymbol{o}_{\mathrm{Q},\mathbf{z}_0}\|^2_{L_2\left([0,\infty)^2,\mathbb{R}^{N_{\mathbf{z}_0}\times N}\right)} = \int_0^\infty \int_0^\infty \operatorname{tr}\big(\boldsymbol{o}_{\mathrm{Q},\mathbf{z}_0}(t_1,t_2)^{\mathrm{H}}\boldsymbol{o}_{\mathrm{Q},\mathbf{z}_0}(t_1,t_2)\big)\,\mathrm{d}t_1\mathrm{d}t_2 \\ &= \operatorname{tr}(\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\mathbf{z}_0})\,. \end{aligned} \tag{3.43}$$

Again, we note that the traces of the observability Gramians $\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\boldsymbol{\mathfrak{B}}}$ and $\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\mathbf{z}_0}$, which are the summands of their eigenvalues, indicate which states are significant for the system dynamics. Since the largest eigenvalues of the observability Gramians have the most influence on the output energies, the corresponding states span the dominant observability subspaces.

In this section, we have derived four transfer functions with corresponding system representations, that describe the overall system behavior. Corresponding to these subsystems, we have derived suitable Gramians and energy expressions that incorporate the controllability and observability properties of the original system. Table 3.3 depicts the derived subsystems with the corresponding transfer functions, the respective Gramians, and the resulting energies.

	System (3.34)	System (3.35)	System (3.36)	System (3.37)
Transfer function	$\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(s_1, s_2)$	$\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}(s_1, s_2)$	$\mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}(s_1, s_2)$	$\mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0}(s_1, s_2)$
Controllability Gramian	$\mathcal{P}_{\mathcal{B}}$	$\mathcal{P}_{\mathbf{z}_0}$	$\mathcal{P}_{\mathcal{B}}$	$\mathcal{P}_{\mathbf{z}_0}$
Observability Gramian	$\mathcal{Q}_{\mathrm{Q},\mathcal{B}}$	$\mathcal{Q}_{\mathrm{Q},\mathcal{B}}$	$\mathcal{Q}_{\mathrm{Q},\mathbf{z}_0}$	$\mathcal{Q}_{\mathrm{Q},\mathbf{z}_0}$
Controllability energies	$E(\boldsymbol{c}_{\mathcal{B}}) = \mathrm{tr}(\mathcal{P}_{\mathcal{B}})$	$E(\boldsymbol{c}_{\mathbf{z}_0}) = \mathrm{tr}(\mathcal{P}_{\mathbf{z}_0})$	$E(\boldsymbol{c}_{\mathcal{B}}) = \mathrm{tr}(\mathcal{P}_{\mathcal{B}})$	$E(\boldsymbol{c}_{\mathbf{z}_0}) = \mathrm{tr}(\mathcal{P}_{\mathbf{z}_0})$
Observability energies	$E(\boldsymbol{o}_{\mathrm{Q},\mathcal{B}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{Q},\mathcal{B}})$	$E(\boldsymbol{o}_{\mathrm{Q},\mathcal{B}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{Q},\mathcal{B}})$	$E(\boldsymbol{o}_{\mathrm{Q},\mathbf{z}_0}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{Q},\mathbf{z}_0})$	$E(\boldsymbol{o}_{\mathrm{Q},\mathbf{z}_0}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{Q},\mathbf{z}_0})$

Table 3.3: Properties of system (3.31) corresponding to its multi-system representation.

3.1.2.2 Extended-input approach for inhomogeneous first-order ODE systems with a quadratic output

As an alternative to the multi-system approach presented before, we apply the extended-input method from [66] and modify it so that it is suitable for systems with a quadratic output. Therefore, we derive the transfer function of the original system (3.31) and a respective homogeneous system representation that is analyzed instead of the original system.

Transfer function Our objective is to describe the relationship between inputs, initial conditions, and the output behavior of the system (3.31). To achieve this, we combine the transfer functions from (3.33), as each encodes a part of the overall input– and initial condition–to–output behavior, which yields

$$\begin{aligned}\mathcal{G}_{\mathrm{Q},\mathrm{ww}}(s_1, s_2) &:= \mathcal{G}_{\mathrm{Q},\mathcal{B}\mathcal{B}}(s_1, s_2) + \mathcal{G}_{\mathrm{Q},\mathcal{B}\mathbf{z}_0}(s_1, s_2) + \mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathcal{B}}(s_1, s_2) + \mathcal{G}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0}(s_1, s_2)\\ &= \mathcal{W}^{\mathrm{T}}(s_1\mathcal{E} - \mathcal{A})^{-\mathrm{H}}\mathcal{M}(s_2\mathcal{E} - \mathcal{A})^{-1}\mathcal{W}\end{aligned}$$

for the input matrix $\mathcal{W}$ defined in (3.24).

Definition 3.14:
Consider the system (3.31) with initial conditions as defined in (3.4). The *transfer function* corresponding to this system is defined as

$$\mathcal{G}_{\mathrm{Q},\mathrm{ww}}(s_1, s_2) := \mathcal{W}^{\mathrm{T}}(s_1\mathcal{E} - \mathcal{A})^{-\mathrm{H}}\mathcal{M}(s_2\mathcal{E} - \mathcal{A})^{-1}\mathcal{W}. \tag{3.44}$$

◊

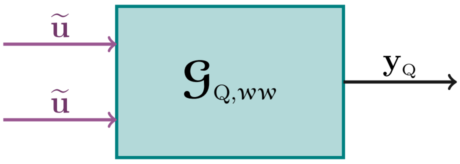


Figure 3.6: Structure of a first-order ODE system with an extended input and a quadratic output.

Since the transfer function $\mathcal{G}_{\mathrm{Q},ww}$ has multiple system realizations, we use a homogeneous one, that is

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}(t) &= \mathcal{A}\mathbf{z}(t) + \mathcal{W}\widetilde{\mathbf{u}}(t), \qquad \mathbf{z}(0) = 0, \\ \mathbf{y}_{\mathrm{Q}}(t) &= \mathbf{z}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}(t), \end{aligned} \tag{3.45}$$

where $\widetilde{\mathbf{u}} \in L_2([0,\infty),\mathbb{R}^{n_{\mathcal{W}}})$ is a suitable input function. In the following, we consider the homogeneous system (3.45) instead of the inhomogeneous original one (3.31). The homogeneous system is depicted in Figure 3.6, where no initial conditions are added to the homogeneous system as they are embedded in the input $\widetilde{\mathbf{u}}$. Note, however, that we only use the surrogate system to describe the controllability and observability behavior. Later in this work, when we apply model reduction techniques, the resulting spaces are used to reduce the original system (3.31) and find an inhomogeneous reduced model.

Controllability Gramian To investigate the input-to-output behavior of the homogeneous system (3.45), we aim to derive the corresponding controllability Gramian spanning the controllability space. We observe that the input-to-output mapping of this system is equal to the input-to-output mapping of the ODE system with a linear output equation from (3.27). Hence, the controllability Gramian of the system (3.45) is equal to the one defined in (3.29), which leads to the following definition.

Definition 3.15:
Consider the asymptotically stable system (3.45) with $\mathcal{W}$ as defined in (3.24). Then the corresponding *controllability Gramian* is defined as

$$\mathcal{P}_{\mathrm{w}} = \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{W}\mathcal{W}^{\mathrm{T}}\mathcal{E}^{-1}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\mathrm{d}t. \qquad \Diamond$$

Observability Gramians In this paragraph, we aim to derive an observability Gramian that encodes the observability properties of the homogenous system (3.27), which can be used to identify dominant observability spaces. Therefore, we consider the output equation

$$\mathbf{y}_{\mathrm{Q}}(t) = \int_0^t\int_0^t \widetilde{\mathbf{u}}(\tau_1)^{\mathrm{T}}\mathcal{W}^{\mathrm{T}}\mathcal{E}^{-1}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}\tau_1}\mathcal{M}e^{\mathcal{E}^{-1}\mathcal{A}\tau_2}\mathcal{E}^{-1}\mathcal{W}\widetilde{\mathbf{u}}(\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2$$

of system (3.45) and identify the input-to-state mapping $\boldsymbol{c}_{\mathrm{w}}(s)$ from (3.28). The remaining mapping within the output $\mathbf{y}_{\mathrm{Q}}(t)$ is the state-to-output mapping, which is defined as

$$\boldsymbol{o}_{\mathrm{Q,w}}(t_1,t_2) := \boldsymbol{\mathcal{W}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_1}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_2}\boldsymbol{\mathcal{E}}^{-1}. \tag{3.46}$$

From this mapping, we derive the matrix

$$\begin{aligned}\boldsymbol{\mathcal{Q}}_{\mathrm{Q,w}} &:= \int_0^\infty\int_0^\infty \boldsymbol{o}_{\mathrm{Q,w}}(t_1,t_2)^{\mathrm{T}}\boldsymbol{o}_{\mathrm{Q,w}}(t_1,t_2)\mathrm{d}t_1\mathrm{d}t_2\\ &= \int_0^\infty\int_0^\infty \boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_2}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_1}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{W}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-1}e^{(\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}})^{\mathrm{T}}t_1}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_2}\boldsymbol{\mathcal{E}}^{-1}\mathrm{d}t_1\mathrm{d}t_2\\ &= \int_0^\infty \boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_2}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{w}}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_2}\boldsymbol{\mathcal{E}}^{-1}\mathrm{d}t_2,\end{aligned}$$

that spans the corresponding observability space using the definition of $\boldsymbol{\mathcal{P}}_{\mathrm{w}}$ from (3.29).

Definition 3.16:
Consider the asymptotically stable system (3.45) and the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{w}}$ defined in (3.29). Then, the corresponding *observability Gramian* is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{Q,w}} := \int_0^\infty \boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{w}}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\mathrm{d}t. \tag{3.47}$$

◊

We observe that the observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,w}}$ also contains the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{w}}$, as the observability behavior of the right state in the output expression in (3.45) depends also on the controllability state of the left state encoded by $\boldsymbol{\mathcal{P}}_{\mathrm{w}}$. To compute the observability Gramian, we solve the Lyapunov equation

$$\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_{\mathrm{Q,w}}\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_{\mathrm{Q,w}}\boldsymbol{\mathcal{E}} = -\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{w}}\boldsymbol{\mathcal{M}}$$

as described in [20, Lemma 2.1].

Controllability energies To identify the dominant controllability subspaces of the homogeneous system (3.45), we aim to derive the controllability energies of this system. We note that the input-to-state mapping, and hence the corresponding controllability Gramian coincide with those corresponding to the system (3.27) with a linear output equation. Therefore, as derived in (3.30), we apply the energy norm to state-output mapping $\boldsymbol{c}_{\mathrm{w}}$ defined in (3.28), which yields

$$E(\boldsymbol{c}_{\mathrm{w}}) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{w}})\,.$$

This energy expression indicates that the states corresponding to the highest eigenvalues of the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{w}}$ span the most dominant controllability subspaces.

Observability Energies In this paragraph, we derive some observability energies to identify the dominant observability subspaces of the homogeneous system (3.45). We derive an energy expression based on the state-to-output mapping $\boldsymbol{o}_{\mathrm{Q},w}$ defined in (3.46). We evaluate the energy norm from (3.19) of this mapping, that is

$$\begin{aligned} E(\boldsymbol{o}_{\mathrm{Q},w}) = \|\boldsymbol{o}_{\mathrm{Q},w}\|^2_{L_2\left([0,\infty)^2,\mathbb{R}^{m+N\mathbf{z}_0\times N}\right)} &= \int_0^\infty\int_0^\infty \mathrm{tr}\big(\boldsymbol{o}_{\mathrm{Q},w}(t_1,t_2)^{\mathrm{H}}\boldsymbol{o}_{\mathrm{Q},w}(t_1,t_2)\big)\,\mathrm{d}t_1\mathrm{d}t_2 \\ &= \mathrm{tr}(\mathcal{Q}_{\mathrm{Q}})\,. \end{aligned} \tag{3.48}$$

This energy norm shows, which states are the most significant ones describing the system dynamics. Since the trace of the observability Gramian $\mathcal{Q}_{\mathrm{Q},w}$ is equal to the sum of its eigenvalues, it follows that the states corresponding to the largest eigenvalues of $\mathcal{Q}_{\mathrm{Q},w}$ encode the dominant observability subspaces.

To summarize the extended-input approach that derives a homogeneous system (3.45) to describe the system dynamics, Table 3.4 depicts the considered transfer function, the resulting Gramians, and the respective energies.

	System (3.45)
Transfer function	$\mathcal{G}_{\mathrm{Q},ww}(s_1,s_2)$
Controllability Gramian	$\mathcal{P}_w$
Observability Gramian	$\mathcal{Q}_{\mathrm{Q},w}$
Controllability energies	$E(\boldsymbol{c}_w)=\mathrm{tr}(\mathcal{P}_w)$
Observability energies	$E(\boldsymbol{o}_{\mathrm{Q},w})=\mathrm{tr}(\mathcal{Q}_{\mathrm{Q},w})$

Table 3.4: Properties of system (3.31) corresponding to its extended-input representation.

3.2 Inhomogeneous first-order DAE systems

In this section, we generalize the theory presented above to dynamical systems with a differential-algebraic equation as a state equation that has the form

$$\mathcal{E}\dot{\mathbf{z}}(t) = \mathcal{A}\mathbf{z}(t) + \mathcal{B}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0 \tag{3.49}$$

with $\mathcal{E}$, $\mathcal{A} \in \mathbb{R}^{N\times N}$, and $\mathcal{B} \in \mathbb{R}^{N\times m}$, where we assume that $\mathcal{E}$ is a singular matrix. Moreover, we assume in this work that the matrix pencil $(\mathcal{A},\mathcal{E})$ is regular, i.e., $\lambda\mathcal{E}-\mathcal{A}$ is not a zero polynomial, and that the consistency conditions in (2.15) are satisfied.

DAE systems arise when, for example, electrical circuits, heat and diffusion processes, or multibody systems are modeled using methods such as finite elements or finite volumes. These systems involve physical constraints, which lead to algebraic equations. Therefore, tailored analysis tools need to be developed, see [79, 91, 130].

We assume that the matrix $\mathbf{Z}_0 \in \mathbb{R}^{N \times N_{\mathbf{Z}}}$ spans a space containing all admissible initial states, i.e., for all initial states $\mathbf{z}_0$ there exists a vector $\zeta_0 \in \mathbb{R}^{N_{\mathbf{Z}}}$ such that

$$\mathbf{z}_0 = \mathbf{Z}_0 \zeta_0. \tag{3.50}$$

The matrix $\mathbf{Z}_0$ is composed of a proper part $\mathbf{Z}_{\mathrm{p},0}$ and an improper part $\mathbf{Z}_{\mathrm{i},0}$, so that $\mathbf{Z}_0 = \mathbf{Z}_{\mathrm{p},0} + \mathbf{Z}_{\mathrm{i},0}$ and

$$\mathbf{Z}_{\mathrm{p},0} = \mathbf{P}_{\mathrm{r}} \mathbf{Z}_0 \qquad \text{and} \qquad \mathbf{Z}_{\mathrm{i},0} = (\mathbf{I}_N - \mathbf{P}_{\mathrm{r}}) \mathbf{Z}_0 \tag{3.51}$$

holds for a projection matrix $\mathbf{P}_{\mathrm{r}}$ as defined in (2.10). The matrices $\mathbf{Z}_{\mathrm{p},0}$ and $\mathbf{Z}_{\mathrm{i},0}$ are used in the following to study the system properties while considering all admissible initial conditions. According to that initial condition matrix decomposition in (3.51), the initial state is composed of

$$\mathbf{z}_0 = \mathbf{z}_{\mathrm{p},0} + \mathbf{z}_{\mathrm{i},0} \qquad \text{with} \qquad \mathbf{z}_{\mathrm{p},0} = \mathbf{P}_{\mathrm{r}} \mathbf{z}_0, \quad \mathbf{z}_{\mathrm{i},0} = (\mathbf{I}_N - \mathbf{P}_{\mathrm{r}}) \mathbf{z}_0 \tag{3.52}$$

Note that the initial condition for the algebraic state component $\mathbf{z}_{\mathrm{i}}(t)$ is already included in the state trajectory due to the consistency condition (2.15). Hence, $\mathbf{Z}_{\mathrm{i},0}$ can be chosen as

$$\mathbf{Z}_{\mathrm{i},0} = \begin{bmatrix} \mathcal{F}_{\mathbf{N}}(0)\mathcal{B} & \cdots & \mathcal{F}_{\mathbf{N}}(\nu - 1)\mathcal{B} \end{bmatrix} \qquad \text{with} \qquad \mathbf{z}_{\mathrm{i},0} = \mathbf{Z}_{\mathrm{i},0} \begin{bmatrix} \mathbf{u}^{(0)}(0) \\ \vdots \\ \mathbf{u}^{(\nu-1)}(0) \end{bmatrix}. \tag{3.53}$$

In the following, we analyze the behavior of DAE systems with a state equation (3.49). Therefore, we generalize existing methods for systems with linear output equations to include the initial conditions when analyzing the system behavior. Moreover, we investigate DAE systems with quadratic output equations, which have not been studied in the literature. For both classes of systems, we consider a multi-system approach and an extended-input approach, modifying the methods presented in the previous section.

First, in Section 3.2.1, we derive the system properties of DAE systems with linear output equations, and then study systems with quadratic output equations in Section 3.2.2.

3.2.1 Inhomogeneous first-order DAE systems with a linear output

We consider DAE systems with a linear output equation of the form

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}(t) &= \mathcal{A}\mathbf{z}(t) + \mathcal{B}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0, \\ \mathbf{y}_{\mathrm{L}}(t) &= \mathcal{C}\mathbf{z}(t), \end{aligned} \tag{3.54}$$

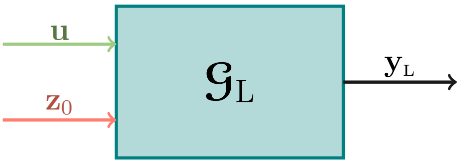


Figure 3.7: Structure of a first-order DAE system with a linear output.

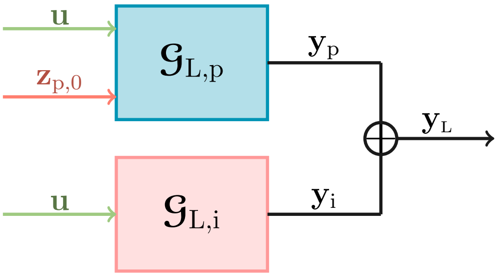


Figure 3.8: Structure of a first-order DAE system with a linear output - differential and algebraic components decoupled.

where the state equation is as defined in (3.49), and the output equation includes the output matrix $\mathcal{C} \in \mathbb{R}^{p\times N}$ and an output $\mathbf{y}_\mathrm{L}(t) \in \mathbb{R}^p$. The input- and initial condition-to-output structure of this system is depicted in Figure 3.7, where we have not yet separated the differential and algebraic components so that it is of the same form as for the ODE system with a linear output equation depicted in Figure 3.1. Decomposing the system into its differential and algebraic components as defined in (2.14) leads to the two outputs

$$\mathbf{y}_{\mathrm{L,p}}(t) := \mathcal{C}\mathbf{z}_\mathrm{p}(t), \qquad \mathbf{y}_{\mathrm{L,i}}(t) := \mathcal{C}\mathbf{z}_\mathrm{i}(t) \qquad \text{with} \qquad \mathbf{y}_\mathrm{L}(t) = \mathbf{y}_{\mathrm{L,p}}(t) + \mathbf{y}_{\mathrm{L,i}}(t),$$

as depict in Figure 3.8. We add an input $\mathbf{u}$ and an initial condition $\mathbf{z}_0$ to the differential system, that are needed to derive the respective proper output $\mathbf{y}_\mathrm{p}$. To generate the improper output $\mathbf{y}_\mathrm{i}$ only the input $\mathbf{u}$ is needed since the initial conditions satisfy the consistency conditions (2.15).

In the following, we analyze the input- and initial condition-to-output behavior of the differential and the algebraic system separately. Therefore, we generate the corresponding Gramians that describe the respective controllability and observability spaces. These Gramians are then used to derive the system energies, which define the dominant controllability and observability subspaces. We again utilize two different approaches to treat the inhomogeneous initial conditions, namely the multi-system approach shown in Section 3.2.1.1 and the extended-input approach presented in Section 3.2.1.2.

3.2.1.1 Multi-system approach for inhomogeneous first-order DAE systems with a linear output

In this paragraph, we derive a multi-system representation of the system (3.54) to treat the inhomogeneous initial conditions. Therefore, we modify the method introduced in Section 3.1.1.1 to incorporate the differential and algebraic components of the DAE system. We again derive some subsystems that are analyzed separately and derive the respective Gramians and system energies.

Transfer function To evaluate the behavior of the system (3.54), we consider the state components defined in (2.14) where we decompose the differential state additionally into

$$\mathbf{z}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t) = \int_0^t \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t-\tau)\boldsymbol{\mathcal{B}}\mathbf{u}(\tau)\mathrm{d}\tau \qquad \text{and} \qquad \mathbf{z}_{\mathrm{p},\mathbf{Z}_0}(t) = \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\boldsymbol{\mathcal{E}}\mathbf{Z}_{\mathrm{p},0}\zeta_0. \tag{3.55}$$

We apply the Laplace transform to each of the components $\mathbf{z}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)$, $\mathbf{z}_{\mathrm{p},\mathbf{Z}_0}(t)$, and $\mathbf{z}_{\mathrm{i}}(t)$, which yields

$$\begin{aligned} &\mathbf{Z}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(s) := \mathbf{P}_{\mathrm{r}}(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}\mathbf{U}(s), \qquad \mathbf{Z}_{\mathrm{p},\mathbf{Z}_0}(s) := \mathbf{P}_{\mathrm{r}}(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{E}}\mathbf{Z}_{\mathrm{p},0}\zeta_0, \\ &\mathbf{Z}_{\mathrm{i}}(s) := (\mathbf{I}-\mathbf{P}_{\mathrm{r}})(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}\mathbf{U}(s). \end{aligned} \tag{3.56}$$

To describe not only the input- and initial condition-to-state behavior but also the respective output, we apply the Laplace transform to the output equation in (3.54) and insert the three state components from (3.56), to obtain the three outputs

$$\begin{aligned} &\mathbf{Y}_{\mathrm{L},\mathrm{p},\boldsymbol{\mathcal{B}}}(s) := \boldsymbol{\mathcal{C}}\mathbf{P}_{\mathrm{r}}(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}\mathbf{U}(s), \qquad \mathbf{Y}_{\mathrm{L},\mathrm{p},\mathbf{Z}_0}(s) : \boldsymbol{\mathcal{C}}\mathbf{P}_{\mathrm{r}}(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{E}}\mathbf{Z}_{\mathrm{p},0}\zeta_0, \\ &\mathbf{Y}_{\mathrm{L},\mathrm{i}}(s) := \boldsymbol{\mathcal{C}}(\mathbf{I}-\mathbf{P}_{\mathrm{r}})(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}\mathbf{U}(s). \end{aligned}$$

From these output representations, we can extract the respective transfer functions of the system (3.54), that encode the proper and improper input- and initial condition-to-output mappings.

Definition 3.17:
Consider the system (3.54) with a regular matrix pencil $s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}}$ and the projection matrix $\mathbf{P}_{\mathrm{r}}$ defined in (2.10). Assume that the consistency conditions in (2.15) are satisfied. Then the respective *transfer functions* are defined as

$$\begin{aligned} &\boldsymbol{\mathcal{G}}_{\mathrm{L},\mathrm{p},\boldsymbol{\mathcal{B}}}(s) := \boldsymbol{\mathcal{C}}\mathbf{P}_{\mathrm{r}}(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}}, \qquad \boldsymbol{\mathcal{G}}_{\mathrm{L},\mathrm{p},\mathbf{Z}_0}(s) := \boldsymbol{\mathcal{C}}\mathbf{P}_{\mathrm{r}}(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{E}}\mathbf{Z}_{\mathrm{p},0}, \\ &\boldsymbol{\mathcal{G}}_{\mathrm{L},\mathrm{i}}(s) := \boldsymbol{\mathcal{C}}(\mathbf{I}-\mathbf{P}_{\mathrm{r}})(s\boldsymbol{\mathcal{E}}-\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{B}} \end{aligned} \tag{3.57}$$

◊

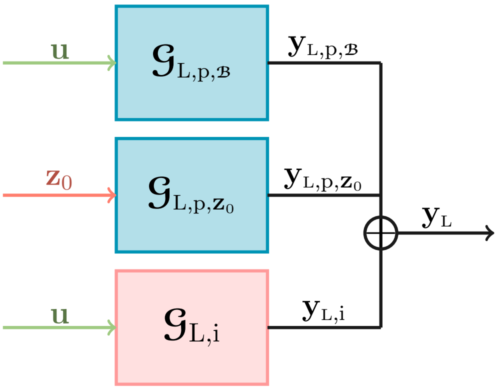

$\mathbf{u}$
$\mathcal{G}_{\mathrm{L,p},\mathcal{B}}$
$\mathbf{y}_{\mathrm{L,p},\mathcal{B}}$
$\mathbf{z}_0$
$\mathcal{G}_{\mathrm{L,p},\mathbf{z}_0}$
$\mathbf{y}_{\mathrm{L,p},\mathbf{z}_0}$
$\mathbf{y}_{\mathrm{L}}$
$\mathbf{u}$
$\mathcal{G}_{\mathrm{L,i}}$
$\mathbf{y}_{\mathrm{L,i}}$

Figure 3.9: Structure of three separated first-order DAE systems with a linear output.

The first transfer function, $\mathcal{G}_{\mathrm{L,p},\mathcal{B}}$, results from the differential system with the homogeneous differential initial condition, i.e., $\mathbf{P}_{\mathrm{r}}\mathbf{z}_{\mathrm{p},\mathcal{B}}(0) = \mathbf{z}_{\mathrm{p},0} = 0$ as described in (3.52), which yields the system realization

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathrm{p},\mathcal{B}}(t) &= \mathcal{A}\mathbf{z}_{\mathrm{p},\mathcal{B}}(t) + \mathbf{P}_{\mathrm{l}}\mathcal{B}\mathbf{u}(t), \qquad \mathbf{P}_{\mathrm{r}}\mathbf{z}_{\mathrm{p},\mathcal{B}}(0) = 0,\\ \mathbf{y}_{\mathrm{L,p},\mathcal{B}}(t) &= \mathcal{C}\mathbf{z}_{\mathrm{p},\mathcal{B}}(t). \end{aligned} \tag{3.58}$$

The second transfer function, $\mathcal{G}_{\mathrm{L,p},\mathbf{z}_0}$, corresponds to the system representation

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathrm{p},\mathbf{z}_0}(t) &= \mathcal{A}\mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t), \qquad \mathbf{P}_{\mathrm{r}}\mathbf{z}_{\mathrm{p},\mathbf{z}_0}(0) = \mathbf{Z}_{\mathrm{p},0}\zeta_0,\\ \mathbf{y}_{\mathrm{L,p},\mathbf{z}_0}(t) &= \mathcal{C}\mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t), \end{aligned} \tag{3.59}$$

where no input is added to the system and a differential initial condition $\mathbf{P}_{\mathrm{r}}\mathbf{z}(0) = \mathbf{z}_{\mathrm{p},0} = \mathbf{Z}_{\mathrm{p},0}\zeta_0$ is applied. This system describes the differential initial condition-to-output mapping. The third transfer function, $\mathcal{G}_{\mathrm{L,i}}$, has the system representation

$$\begin{aligned} \mathcal{E}\dot{\mathbf{z}}_{\mathrm{i}}(t) &= \mathcal{A}\mathbf{z}_{\mathrm{i}}(t) + (\mathbf{I} - \mathbf{P}_{\mathrm{l}})\mathcal{B}\mathbf{u}(t), \qquad (\mathbf{I} - \mathbf{P}_{\mathrm{r}})\mathbf{z}_{\mathrm{i}}(0) = \mathbf{z}_{\mathrm{i},0},\\ \mathbf{y}_{\mathrm{L,i}}(t) &= \mathcal{C}\mathbf{z}_{\mathrm{i}}(t) \end{aligned} \tag{3.60}$$

where we assume that the initial condition $(\mathbf{I} - \mathbf{P}_{\mathrm{r}})\mathbf{z}(0) = \mathbf{z}_{\mathrm{i},0} = \mathbf{Z}_{\mathrm{i},0}\zeta_0$ from (3.52) satisfies the consistency conditions.

The decomposition of the system (3.54) into the three subsystems (3.58), (3.59), and (3.60) describing the dynamics of the overall system is depicted in Figure 3.9.

In the following, we investigate the controllability and observability properties of the three subsystems separately. For that, we derive the respective Gramians spanning their controllability and observability spaces.

Controllability Gramian To describe the controllability behavior of the three subsystems, we derive respective input- and initial condition-to-state mappings. The first subsystem (3.58) has the input-to-state mapping

$$\boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t) := \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\boldsymbol{\mathcal{B}}, \tag{3.61}$$

where $\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)$ is as defined in (2.13). Since this mapping encodes all reachable states of the subsystem (3.58), it is used to define a matrix $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} := \int_0^\infty \boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)\boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)^{\mathrm{T}}\mathrm{d}t$ that spans the corresponding controllability space.

Definition 3.18:
Consider the C-stable system (3.58) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$ and $\boldsymbol{\mathcal{F}}_{\mathbf{J}}$ as defined in (2.13). Then the corresponding *proper controllability Gramian* is defined as

$$\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} := \int_0^\infty \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)^{\mathrm{T}}\mathrm{d}t. \tag{3.62}$$

◊

Furthermore, inserting the definition of $\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)$ into (3.62) yields the following lemma.

Lemma 3.19:
Consider the C-stable system (3.58) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$. Assume that $\mathbf{T}$, $\mathbf{W}$ are matrices that transform system (3.58) into Weierstraß canonical form as described in (2.9). Then the Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ from (3.62) is of the form

$$\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} := \mathbf{T}^{-1}\begin{bmatrix}\mathbf{P}_{1,\boldsymbol{\mathcal{B}}} & 0\\ 0 & 0\end{bmatrix}\mathbf{T}^{-\mathrm{T}}, \qquad \mathbf{P}_{1,\boldsymbol{\mathcal{B}}} = \int_0^\infty e^{\mathbf{J}t}\mathbf{B}_1\mathbf{B}_1^{\mathrm{T}}e^{\mathbf{J}^{\mathrm{T}}t}\mathrm{d}t \tag{3.63}$$

with $\boldsymbol{\mathcal{B}} = \mathbf{W}\left[\begin{smallmatrix}\mathbf{B}_1\\ \mathbf{B}_2\end{smallmatrix}\right]$. ◊

This lemma vividly shows that the proper controllability Gramian is connected to the Gramian of the differential states in the WCF from (2.9). Since the differential state results from an ODE state equation, the theory from Section 3.1 applies to this state component. Using the controllability Gramian, we can characterize the states that are difficult to reach or even unreachable, which play a significant role when applying reduction methods to the system. It remains to compute the Gramian. For that, we use that the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ defined in (3.63) is the unique solution of the continuous-time projected Lyapunov equation

$$\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{A}}^{\mathrm{T}} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{E}}^{\mathrm{T}} = -\mathbf{P}_{\mathrm{l}}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\mathbf{P}_{\mathrm{l}}^{\mathrm{T}}, \qquad \boldsymbol{\mathcal{P}}_{\mathrm{p}} = \mathbf{P}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}} \tag{3.64}$$

as described in (2.19).

To describe the controllability behavior of the subsystem (3.59), we derive the respective input-to-state mapping

$$\boldsymbol{c}_{\mathrm{p},\mathbf{Z}_0}(t) := \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\boldsymbol{\mathcal{E}}\mathbf{Z}_0, \tag{3.65}$$

where $\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)$ is as defined in (2.13). This mapping is used to describe the controllability space of the subsystem (3.59). For that, we define a matrix $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{Z}_0} := \int_0^\infty \boldsymbol{c}_{\mathrm{p},\mathbf{Z}_0}(t)\boldsymbol{c}_{\mathrm{p},\mathbf{Z}_0}(t)^{\mathrm{T}}\mathrm{d}t$ that encodes the respective controllability space by integrating over the entire times domain.

Definition 3.20:
Consider the C-stable system (3.59) with a regular matrix pencil $(\mathcal{A}, \mathcal{E})$ and $\mathcal{F}_{\mathbf{J}}(t)$ as defined in (2.13). Then the respective *proper controllability Gramian* is defined as

$$\mathcal{P}_{\mathrm{p},\mathbf{Z}_0} := \int_0^\infty \mathcal{F}_{\mathbf{J}}(t)\mathcal{E}\mathbf{Z}_0\mathbf{Z}_0^{\mathrm{T}}\mathcal{E}^{\mathrm{T}}\mathcal{F}_{\mathbf{J}}(t)^{\mathrm{T}}\mathrm{d}t. \tag{3.66}$$

◊

To describe the connection between the Gramian $\mathcal{P}_{\mathrm{p},\mathbf{Z}_0}$ and the respective WCF, we insert the definition of $\mathcal{F}_{\mathbf{J}}(t)$ into (3.66) to derive the following lemma.

Lemma 3.21:
Consider the C-stable system (3.59) with a regular matrix pencil $(\mathcal{A}, \mathcal{E})$. Assume that $\mathbf{T}$, $\mathbf{W}$ are matrices that transform system (3.59) into Weierstraß canonical form. Then the Gramian $\mathcal{P}_{\mathrm{p},\mathbf{Z}_0}$ from (3.66) is of the form

$$\mathcal{P}_{\mathrm{p},\mathbf{Z}_0} = \mathbf{T}^{-1}\begin{bmatrix}\mathbf{P}_{1,\mathbf{Z}_0} & 0\\ 0 & 0\end{bmatrix}\mathbf{T}^{-\mathrm{T}}, \qquad \mathbf{P}_{1,\mathbf{Z}_0} = \int_0^\infty e^{\mathbf{J}t}\mathbf{Z}_1\mathbf{Z}_1^{\mathrm{T}}e^{\mathbf{J}^{\mathrm{T}}t}\mathrm{d}t \tag{3.67}$$

with $\mathbf{Z}_0 = \mathbf{W}\left[\begin{smallmatrix}\mathbf{Z}_1\\ \mathbf{Z}_2\end{smallmatrix}\right]$. ◊

To compute the Gramian we use that the controllability Gramian $\mathcal{P}_{\mathrm{p},\mathbf{Z}_0}$ as defined in (3.67) is the unique solution of the continuous-time projected Lyapunov equation

$$\mathcal{E}\mathcal{P}_{\mathrm{p},\mathbf{Z}_0}\mathcal{A}^{\mathrm{T}} + \mathcal{A}\mathcal{P}_{\mathrm{p},\mathbf{Z}_0}\mathcal{E}^{\mathrm{T}} = -\mathbf{P}_{\mathrm{l}}\mathcal{E}\mathbf{Z}_0\mathbf{Z}_0^{\mathrm{T}}\mathcal{E}^{\mathrm{T}}\mathbf{P}_{\mathrm{l}}^{\mathrm{T}}, \qquad \mathcal{P}_{\mathrm{p},\mathbf{Z}_0} = \mathbf{P}_{\mathrm{r}}\mathcal{P}_{\mathrm{p},\mathbf{Z}_0}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}. \tag{3.68}$$

Finally, we describe the controllability of the remaining subsystem (3.60), that corresponds to the improper system dynamics of the original system (3.54). Therefore, we consider the corresponding input-to-state mapping

$$\boldsymbol{c}_{\mathrm{i}}(k) := \mathcal{F}_{\mathbf{N}}(k)\mathcal{B}, \tag{3.69}$$

with $\mathcal{F}_{\mathbf{N}}(k)$ as defined in (2.13). This mapping is used to derive a matrix $\mathcal{P}_{\mathrm{i}} := \sum_{k=0}^{\nu-1}\boldsymbol{c}_{\mathrm{i}}(k)\boldsymbol{c}_{\mathrm{i}}(k)^{\mathrm{T}}$ that spans the controllability space including all reachable algebraic states by summing over all discrete matrices defined by $\boldsymbol{c}_{\mathrm{i}}(k)$.

Definition 3.22:
Consider the system (3.60) with a regular matrix pencil $(\mathcal{A}, \mathcal{E})$ and $\mathcal{F}_{\mathbf{N}}(k)$ as defined in (2.13). The corresponding *improper controllability Gramian* is defined as

$$\mathcal{P}_{\mathrm{i}} := \sum_{k=0}^{\nu-1}\mathcal{F}_{\mathbf{N}}(k)\mathcal{B}\mathcal{B}^{\mathrm{T}}\mathcal{F}_{\mathbf{N}}(k)^{\mathrm{T}}. \tag{3.70}$$

◊

The matrix $\mathcal{P}_{\mathrm{i}}$ spans the controllability space of the subsystem (3.60). Since the Gramian $\mathcal{P}_{\mathrm{i}}$ spans the improper controllability space, it is connected to the algebraic components of the WCF from (2.9). The relation is described in the following lemma.

Lemma 3.23:
Consider the system (3.60) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$. Assume that $\mathbf{T}$, $\mathbf{W}$ are matrices that transform system (3.54) into Weierstraß canonical form. Then the Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ from (3.70) is of the form

$$\boldsymbol{\mathcal{P}}_{\mathrm{i}} := \mathbf{T}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{P}_{2,\boldsymbol{\mathcal{B}}} \end{bmatrix} \mathbf{T}^{-\mathrm{T}}, \qquad \mathbf{P}_{2,\boldsymbol{\mathcal{B}}} = \sum_{k=0}^{\nu-1} \mathbf{N}^k \mathbf{B}_2 \mathbf{B}_2^{\mathrm{T}} (\mathbf{N}^k)^{\mathrm{T}} \tag{3.71}$$

with $\boldsymbol{\mathcal{B}} = \mathbf{W}\left[\begin{smallmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{smallmatrix}\right]$. ◊

To compute the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ defined in (3.67), we use that $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ is the unique solution of the discrete-time projected Lyapunov equation

$$\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\mathbf{A}^{\mathrm{T}} - \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\boldsymbol{\mathcal{E}}^{\mathrm{T}} = (\mathbf{I} - \mathbf{P}_{\mathrm{l}})\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}(\mathbf{I} - \mathbf{P}_{\mathrm{l}})^{\mathrm{T}}, \qquad 0 = \mathbf{P}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}. \tag{3.72}$$

The three controllability Gramians derived in this paragraph are used in the following to identify states that are dominant in the dynamics of the system and states that are neglectable.

Observability Gramians To describe the observability behavior of the three subsystems (3.58), (3.59), and (3.60), we derive their state-to-output mappings and utilize them to define the corresponding observability Gramians which encode the observability properties of the respective system. We observe that the state-to-output mappings coincide for the two systems (3.58) and (3.59) that describe the differential components of the system dynamics. Hence, we derive a proper state-to-output mapping corresponding to these two systems, which is

$$\boldsymbol{o}_{\mathrm{L,p}}(t) := \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t) \tag{3.73}$$

where $\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)$ is as defined in (2.13). Using $\boldsymbol{o}_{\mathrm{L,p}}(t)$, we derive a matrix $\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}} := \int_0^\infty \boldsymbol{o}_{\mathrm{L,p}}(t)^{\mathrm{T}}\boldsymbol{o}_{\mathrm{L,p}}(t)\mathrm{d}t$ that spans the proper observability space.

Definition 3.24:
Consider the two proper C-stable systems (3.58) and (3.59) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$ and $\boldsymbol{\mathcal{F}}_{\mathbf{J}}$ as defined in (2.13). Then the corresponding *proper observability Gramian* is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}} := \int_0^\infty \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\mathrm{d}t. \tag{3.74}$$

◊

We insert the Weierstraß-canonical form from (2.9) to derive the following lemma.

Lemma 3.25:
Consider the proper C-stable systems (3.58) and (3.59) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathfrak{E}})$. The observability Gramian $\boldsymbol{\mathfrak{Q}}_{\mathrm{L,p}}$ as defined in (3.74) satisfies

$$\boldsymbol{\mathfrak{Q}}_{\mathrm{L,p}} := \mathbf{W}^{-\mathrm{T}} \begin{bmatrix} \mathbf{Q}_{\mathrm{L,1}} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{W}^{-1}, \qquad \mathbf{Q}_{\mathrm{L,1}} = \int_0^\infty e^{\mathbf{J}^{\mathrm{T}}t} \widetilde{\mathbf{C}}_1^{\mathrm{T}} \widetilde{\mathbf{C}}_1 e^{\mathbf{J}t} \mathrm{d}t \tag{3.75}$$

with $\widetilde{\mathbf{C}}_1$ and $\mathbf{J}$ corresponding to the WCF of the respective systems defined in (2.9). ♢

This lemma vividly describes the connection between the observability Gramians of the differential component of the state $\mathbf{z}_{\mathrm{p}}$ and the respective differential state $\mathbf{z}_1$ of the transformed system in WCF. To compute the observability Gramian $\boldsymbol{\mathfrak{Q}}_{\mathrm{L,p}}$ defined in (3.74), we solve the continuous-time projected Lyapunov equation

$$\boldsymbol{\mathfrak{E}}^{\mathrm{T}} \boldsymbol{\mathfrak{Q}}_{\mathrm{L,p}} \boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{A}}^{\mathrm{T}} \boldsymbol{\mathfrak{Q}}_{\mathrm{L,p}} \boldsymbol{\mathfrak{E}} = -\mathbf{P}_{\mathrm{r}}^{\mathrm{T}} \boldsymbol{\mathfrak{C}}^{\mathrm{T}} \boldsymbol{\mathfrak{C}} \mathbf{P}_{\mathrm{r}}, \qquad \boldsymbol{\mathfrak{Q}}_{\mathrm{L,p}} = \mathbf{P}_{\mathrm{l}}^{\mathrm{T}} \boldsymbol{\mathfrak{Q}}_{\mathrm{L,p}} \mathbf{P}_{\mathrm{l}}. \tag{3.76}$$

Now we evaluate the observability behavior of the subsystem (3.60), which encodes the improper components of the original system (3.54). For that, we extract the state-to-output mapping, that is

$$\boldsymbol{o}_{\mathrm{L,i}}(k) := \boldsymbol{\mathfrak{C}} \boldsymbol{\mathfrak{F}}_{\mathbf{N}}(k), \tag{3.77}$$

where $\boldsymbol{\mathfrak{F}}_{\mathbf{N}}(k)$ is as defined in (2.13). We derive a matrix $\boldsymbol{\mathfrak{Q}}_{\mathrm{L,i}} := \sum_{k=0}^{\nu-1} \boldsymbol{o}_{\mathrm{L,i}}(k)^{\mathrm{T}} \boldsymbol{o}_{\mathrm{L,i}}(k)$ that spans the improper observability space by summing over all discrete matrices defined by $\boldsymbol{o}_{\mathrm{L,i}}(k)$.

Definition 3.26:
Consider the improper system (3.60) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathfrak{E}})$. Then the corresponding *improper observability Gramian* is defined as

$$\boldsymbol{\mathfrak{Q}}_{\mathrm{L,i}} := \sum_{k=0}^{\nu-1} \boldsymbol{\mathfrak{F}}_{\mathbf{N}}(k)^{\mathrm{T}} \boldsymbol{\mathfrak{C}}^{\mathrm{T}} \boldsymbol{\mathfrak{C}} \boldsymbol{\mathfrak{F}}_{\mathbf{N}}(k). \tag{3.78}$$

♢

Again, we insert the Weierstraß-canonical form from (2.9) to derive the following lemma.

Lemma 3.27:
Consider the improper system (3.60) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathfrak{E}})$. The observability Gramian $\boldsymbol{\mathfrak{Q}}_{\mathrm{L,i}}$ as defined in (3.78) satisfies

$$\boldsymbol{\mathfrak{Q}}_{\mathrm{L,i}} := \mathbf{W}^{-\mathrm{T}} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{Q}_{\mathrm{L,2}} \end{bmatrix} \mathbf{T}^{-1}, \qquad \mathbf{Q}_{\mathrm{L,2}} = \sum_{k=0}^{\nu-1} (\mathbf{N}^k)^{\mathrm{T}} \mathbf{C}_2^{\mathrm{T}} \mathbf{C}_2 \mathbf{N}^k, \tag{3.79}$$

with $\widetilde{\mathbf{C}}_2$ and $\mathbf{N}$ corresponding to the WCF of the respective systems defined in (2.9). ♢

The improper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}$, defined in (3.78), is computed by solving a projected Lyapunov equation as it is the unique solution of the discrete-time projected Lyapunov equation

$$\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_{\mathrm{i}}\boldsymbol{\mathcal{A}} - \boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}\boldsymbol{\mathcal{E}} = (\mathbf{I}-\mathbf{P}_{\mathrm{r}})^{\mathrm{T}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\boldsymbol{\mathcal{C}}(\mathbf{I}-\mathbf{P}_{\mathrm{r}}), \qquad 0 = \mathbf{P}_{\mathrm{l}}^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}\mathbf{P}_{\mathrm{l}}. \tag{3.80}$$

As for the controllability Gramians, we can characterize states that are hard to observe or unobservable using the observability Gramians. Such states correspond to the small or zero eigenvalues of the corresponding Gramians, as described in the following paragraphs.

Controllability energies We now use the Gramians, defined above, to describe the controllability behavior and the respective energies in more detail. To provide an energy measure based on the proper input- and initial condition-to-state mappings $\boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{c}_{\mathrm{p},\mathbf{z}_0}$ as defined in (3.61) and (3.65), respectively, we evaluate their energy norms as defined in (3.19) to obtain the following energy expressions

$$E(\boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}) = \|\boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\|^2_{L_2([0,\infty),\mathbb{R}^{N\times m})} = \int_0^\infty \mathrm{tr}\big(\boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)\boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)^{\mathrm{T}}\big)\,\mathrm{d}t = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}})\,, \tag{3.81}$$

and

$$E(\boldsymbol{c}_{\mathrm{p},\mathbf{z}_0}) = \|\boldsymbol{c}_{\mathrm{p},\mathbf{z}_0}\|^2_{L_2\left([0,\infty),\mathbb{R}^{N\times N_{\mathbf{z}_0}}\right)} = \int_0^\infty \mathrm{tr}\big(\boldsymbol{c}_{\mathrm{p},\mathbf{z}_0}(t)\boldsymbol{c}_{\mathrm{p},\mathbf{z}_0}(t)^{\mathrm{T}}\big)\,\mathrm{d}t = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}) \tag{3.82}$$

for the Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}$ defined in (3.62) and (3.66).

To apply such an energy measure to the improper component of the system encoded by the controllability mapping $\boldsymbol{c}_{\mathrm{i}}$, we define a discrete energy norm. For that, we consider a sequence $(\boldsymbol{c}(k))_k$, $\boldsymbol{c} : \mathbb{N} \to \mathbb{R}^{N\times m}$ and assume that $\boldsymbol{c} \in \ell_2(N, \mathbb{R}^{N\times m})$, i.e., that

$$\sum_{k=0}^{\infty} \|\boldsymbol{c}(k)\|_{\mathrm{F}} < \infty.$$

Then the ℓ_2-norm of $\boldsymbol{c}$ is defined as

$$E(\boldsymbol{c}) := \|\boldsymbol{c}\|^2_{\ell_2(N,\mathbb{R}^{N\times m})} := \sum_{k=0}^{\infty} \mathrm{tr}\big(\boldsymbol{c}(k)\boldsymbol{c}(k)^{\mathrm{T}}\big)\,. \tag{3.83}$$

Applying the ℓ_2-norm from (3.83) to the input-to-state mapping $\boldsymbol{c}_{\mathrm{i}}$ from (3.69) yields

$$E(\boldsymbol{c}_{\mathrm{i}}) = \|\boldsymbol{c}_{\mathrm{i}}\|^2_{\ell_2(N,\mathbb{R}^{N\times m})} = \sum_{k=0}^{\infty} \mathrm{tr}\big(\boldsymbol{c}_{\mathrm{i}}(k)\boldsymbol{c}_{\mathrm{i}}(k)^{\mathrm{T}}\big) = \sum_{k=0}^{\nu-1} \mathrm{tr}\big(\boldsymbol{c}_{\mathrm{i}}(k)\boldsymbol{c}_{\mathrm{i}}(k)^{\mathrm{T}}\big) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{i}}) \tag{3.84}$$

where $\boldsymbol{\mathcal{P}}_{\mathrm{i}}$ is as defined in (3.70).

Since the trace of a Gramian is equal to the sum of its eigenvalues, it follows from (3.81), (3.82), and (3.84) that the most dominant proper controllability subspaces are those corresponding to the largest eigenvalues of the two proper Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}$. However, the dynamics of the improper system must be captured precisely, and only the states corresponding to zero eigenvalues are negligible.

Observability energies In this paragraph, we aim to analyze the observability behavior of the three subsystems (3.61), (3.65), and (3.69) to identify their dominant observability subspaces. To derive an energy measure, we evaluate the state-to-output mappings $\boldsymbol{o}_{\mathrm{L,p}}$ and $\boldsymbol{o}_{\mathrm{L,i}}$ as defined in (3.73) and (3.77), respectively. These mappings describe the observability behavior of the proper and improper components of the subsystems (3.58), (3.59), and (3.60). Hence, they are used to identify significant subspaces of the state space. For that, we evaluate the energy norms defined in (3.19) and (3.83) of these mappings, which yields

$$E(\boldsymbol{o}_{\mathrm{L,p}}) = \|\boldsymbol{o}_{\mathrm{L,p}}\|^2_{L_2([0,\infty),\mathbb{R}^{p\times N})} = \int_0^\infty \mathrm{tr}\big(\boldsymbol{o}_{\mathrm{L,p}}(t)^{\mathrm{T}}\boldsymbol{o}_{\mathrm{L,p}}(t)\big)\,\mathrm{d}t = \mathrm{tr}(\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}) \tag{3.85}$$

and

$$E(\boldsymbol{o}_{\mathrm{L,i}}) = \|\boldsymbol{o}_{\mathrm{L,i}}\|^2_{\ell_2(\mathbb{N},\mathbb{R}^{p\times N})} = \sum_{k=0}^{\nu-1} \mathrm{tr}\big(\boldsymbol{o}_{\mathrm{L,i}}(k)^{\mathrm{T}}\boldsymbol{o}_{\mathrm{L,i}}(k)\big) = \mathrm{tr}(\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}})\,. \tag{3.86}$$

We observe again that the output energies described by the traces of the proper and improper observability Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}$ are equal to the sum of their eigenvalues, which allows the conclusion that states corresponding to the largest eigenvalues define the most dominant observability subspaces. As the improper system (3.59) needs to maintain the complete system dynamics, only states corresponding to zero eigenvalues of the respective Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}$ can be removed, as they do not change the system behavior.

To summarize the multi-system approach presented in this section, Table 3.5 describes the three derived transfer functions, the respective Gramians, and the resulting energies. They are used in the following chapters to reduce systems of this structure.

	System (3.58)	System (3.59)	System (3.60)
Transfer function	$\boldsymbol{\mathcal{G}}_{\mathrm{L,p},\boldsymbol{\mathcal{B}}}(s)$	$\boldsymbol{\mathcal{G}}_{\mathrm{L,p},\mathbf{z}_0}(s)$	$\boldsymbol{\mathcal{G}}_{\mathrm{L,i}}(s)$
Controllability Gramian	$\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$	$\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}$	$\boldsymbol{\mathcal{P}}_{\mathrm{i}}$
Observability Gramian	$\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$	$\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$	$\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}$
Controllability energies	$E(\boldsymbol{c}_{\mathrm{p},\boldsymbol{\mathcal{B}}}) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}})$	$E(\boldsymbol{c}_{\mathrm{p},\mathbf{z}_0}) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0})$	$E(\boldsymbol{c}_{\mathrm{i}}) = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{i}})$
Observability energies	$E(\boldsymbol{o}_{\mathrm{L,p}}) = \mathrm{tr}(\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}})$	$E(\boldsymbol{o}_{\mathrm{L,p}}) = \mathrm{tr}(\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}})$	$E(\boldsymbol{o}_{\mathrm{L,i}}) = \mathrm{tr}(\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}})$

Table 3.5: Properties of system (3.54) corresponding to its multi-system representation.

3.2.1.2 Extended-input approach for inhomogeneous first-order DAE systems with a linear output

In this paragraph, we derive an extended-input approach that introduces a model with a homogeneous differential initial condition that is evaluated instead of the inhomogeneous original system (3.54). This process will be advantageous if we need one system that captures the overall system behavior rather than three subsystems.

Transfer function To derive proper and improper transfer functions, which also include the initial condition space, we first apply the Laplace transform to the differential state $\mathbf{z}_\mathrm{p}(t)$ from (2.14) to obtain

$$\mathbf{Z}_\mathrm{p}(s) = \mathbf{P}_\mathrm{r}(s\mathcal{E} - \mathcal{A})^{-1}(\mathcal{B}\mathbf{U}(s) + \mathcal{E}\mathbf{Z}_{\mathrm{p},0}\zeta_0) = \mathbf{P}_\mathrm{r}(s\mathcal{E} - \mathcal{A})^{-1}\mathcal{W}_\mathrm{p}\widetilde{\mathbf{U}}(s) \tag{3.87}$$

for the input matrix and the input in the frequency domain

$$\mathcal{W}_\mathrm{p} = \begin{bmatrix}\mathcal{B} & \mathcal{E}\mathbf{Z}_{\mathrm{p},0}\end{bmatrix} \qquad \text{and} \qquad \widetilde{\mathbf{U}}(s) = \begin{bmatrix}\mathbf{U}(s)\\ \zeta_0\end{bmatrix}, \tag{3.88}$$

respectively. Applying the Laplace transform to the output equation in (3.54) and inserting the state $\mathbf{Z}_\mathrm{p}(s)$ from (3.87) leads to the proper output $\mathbf{Y}_{\mathrm{L,p},\mathcal{W}_\mathrm{p}}(s) := \mathcal{C}\mathbf{P}_\mathrm{r}\left(s\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{W}_\mathrm{p}\widetilde{\mathbf{U}}$ from which we extract the respective transfer function

$$\mathcal{G}_{\mathrm{L,p},\mathcal{W}_\mathrm{p}}(s) := \mathcal{C}\mathbf{P}_\mathrm{r}\left(s\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{W}_\mathrm{p}$$

that encodes the input- and initial condition-to-output behavior of the proper components of the system (3.54).

To derive a transfer function that encodes the improper input-to-output mapping, we apply the Laplace transform to the algebraic state $\mathbf{z}_\mathrm{i}(t)$ from (2.14), which yields

$$\begin{aligned}
\mathbf{Z}_{\mathrm{i},\mathcal{W}_\mathrm{p}}(s) &:= (\mathbf{I} - \mathbf{P}_\mathrm{r})\left(s\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{B}\mathbf{U}(s)\\
&= (\mathbf{I} - \mathbf{P}_\mathrm{r})\left(s\mathcal{E} - \mathcal{A}\right)^{-1}(\mathbf{I} - \mathbf{P}_\mathrm{l})\begin{bmatrix}\mathcal{B} & 0\end{bmatrix}\begin{bmatrix}\mathbf{U}(s)\\ \zeta_0\end{bmatrix}\\
&= (\mathbf{I} - \mathbf{P}_\mathrm{r})\left(s\mathcal{E} - \mathcal{A}\right)^{-1}(\mathbf{I} - \mathbf{P}_\mathrm{l})\begin{bmatrix}\mathcal{B} & \mathcal{E}(\mathbf{I} - \mathbf{P}_\mathrm{r})\mathbf{Z}_{\mathrm{p},0}\end{bmatrix}\widetilde{\mathbf{U}}(s)\\
&= (\mathbf{I} - \mathbf{P}_\mathrm{r})\left(s\mathcal{E} - \mathcal{A}\right)^{-1}\begin{bmatrix}\mathcal{B} & (\mathbf{I} - \mathbf{P}_\mathrm{l})\mathcal{E}\mathbf{Z}_{\mathrm{p},0}\end{bmatrix}\widetilde{\mathbf{U}}(s)\\
&= (\mathbf{I} - \mathbf{P}_\mathrm{r})\left(s\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{W}_\mathrm{p}\widetilde{\mathbf{U}}(s)
\end{aligned} \tag{3.89}$$

where we make use of the properties $(\mathbf{I} - \mathbf{P}_\mathrm{r})\mathbf{Z}_{\mathrm{p},0} = 0$ and $(\mathbf{I} - \mathbf{P}_\mathrm{r})(s\mathcal{E} - \mathcal{A})^{-1} = (s\mathcal{E} - \mathcal{A})^{-1}(\mathbf{I} - \mathbf{P}_\mathrm{l})$.

Now, we apply the Laplace transform to the output equation in (3.54) and insert $\mathbf{Z}_{\mathrm{i},\mathcal{W}_\mathrm{p}}(s)$, which leads to the output $\mathbf{Y}_{\mathrm{L,i},\mathcal{W}_\mathrm{p}}(s) := \mathcal{C}(\mathbf{I} - \mathbf{P}_\mathrm{r})\left(s\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{W}_\mathrm{p}\widetilde{\mathbf{U}}$. This improper output encodes the input-to-output mapping described by the following improper

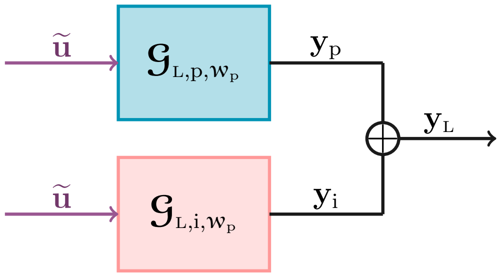

$\widetilde{\mathbf{u}}$
$\mathcal{G}_{\mathrm{L,p},\mathcal{W}_{\mathrm{p}}}$
$\mathbf{y}_{\mathrm{P}}$
$\mathbf{y}_{\mathrm{L}}$
$\widetilde{\mathbf{u}}$
$\mathcal{G}_{\mathrm{L,i},\mathcal{W}_{\mathrm{p}}}$
$\mathbf{y}_{\mathrm{i}}$

Figure 3.10: Structure of a first-order DAE system with an extended input and a linear output - differential and algebraic components decoupled.

transfer function

$$\mathcal{G}_{\mathrm{L,i},\mathcal{W}_{\mathrm{p}}}(s) := \mathcal{C}(\mathbf{I}-\mathbf{P}_{\mathrm{r}})\left(s\mathcal{E}-\mathcal{A}\right)^{-1}\mathcal{W}_{\mathrm{p}}.$$

The proper and improper transfer functions sum up to the following transfer function describing the overall system behavior.

Definition 3.28:
Consider the system (3.54) with a regular matrix pencil $(\mathcal{A},\mathcal{E})$. Also, consider the matrix $\mathcal{W}_{\mathrm{p}}$ from (3.88) and assume that the consistency conditions from (2.15) are satisfied. Then the *transfer function* corresponding to this system is defined as

$$\mathcal{G}_{\mathrm{L},\mathcal{W}_{\mathrm{p}}}(s) :=\mathcal{C}\left(s\mathcal{E}-\mathcal{A}\right)^{-1}\mathcal{W}_{\mathrm{p}}. \tag{3.90}$$

◊

Since the transfer function $\mathcal{G}_{\mathrm{L},\mathcal{W}_{\mathrm{p}}}$ results from multiple system realizations, we consider the following one with a homogeneous differential initial condition

$$\begin{aligned}\mathcal{E}\dot{\mathbf{z}}(t) &= \mathcal{A}\mathbf{z}(t)+\mathcal{W}_{\mathrm{p}}\widetilde{\mathbf{u}}(t), \qquad \mathbf{P}_{\mathrm{r}}\mathbf{z}(0)=0,\\ \mathbf{y}_{\mathrm{L}}(t) &= \mathcal{C}\mathbf{z}(t),\end{aligned} \tag{3.91}$$

with $\widetilde{\mathbf{u}} \in L_2([0,\infty),\mathbb{R}^{m\times N_{\mathbf{z}_0}})$ suitably chosen. The sketch in Figure 3.10 depicts this surrogate system. We note that no initial conditions are needed to evaluate the system dynamics since the differential part of the system has homogeneous initial conditions and the algebraic components satisfy the consistency conditions (2.15).

Controllability Gramian We aim to describe the controllability properties of system (3.91), which also encodes the controllability behavior of the original system (3.54). Therefore, we evaluate the differential and the algebraic states separately.

As described in (2.14), the state $\mathbf{z}(t)$ of system (3.91) decomposes into

$$\mathbf{z}(t)=\mathbf{z}_{\mathrm{p}}(t)+\mathbf{z}_{\mathrm{i}}(t)=\int_0^t \mathcal{F}_{\mathbf{J}}(t-\tau)\mathcal{W}_{\mathrm{p}}\widetilde{\mathbf{u}}(\tau)\mathrm{d}\tau+\sum_{k=0}^{\nu-1}\mathcal{F}_{\mathbf{N}}(k)\mathcal{W}_{\mathrm{p}}\widetilde{\mathbf{u}}^{(k)}(t).$$

We extract the proper and improper input-to-state mappings that are

$$\boldsymbol{c}_{\mathrm{p},\mathcal{W}_\mathrm{p}}(t) := \boldsymbol{\mathcal{F}}_\mathbf{J}(t)\boldsymbol{\mathcal{W}}_\mathrm{p} \qquad \text{and} \qquad \boldsymbol{c}_{\mathrm{i},\mathcal{W}_\mathrm{p}}(k) := \boldsymbol{\mathcal{F}}_\mathbf{N}(k)\boldsymbol{\mathcal{W}}_\mathrm{p}, \tag{3.92}$$

where $\boldsymbol{\mathcal{F}}_\mathbf{J}(t)$ and $\boldsymbol{\mathcal{F}}_\mathbf{N}(k)$ are as defined in (2.13). These mappings are used to define matrices $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{W}_\mathrm{p}} := \int_0^\infty \boldsymbol{c}_{\mathrm{p},\mathcal{W}_\mathrm{p}}(t)\boldsymbol{c}_{\mathrm{p},\mathcal{W}_\mathrm{p}}(t)^\mathrm{T}\mathrm{d}t$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{W}_\mathrm{p}} := \sum_{k=0}^{\nu-1} \boldsymbol{c}_{\mathrm{i},\mathcal{W}_\mathrm{p}}(k)\boldsymbol{c}_{\mathrm{i},\mathcal{W}_\mathrm{p}}(k)^\mathrm{T}$ that span the proper and improper controllability space, respectively.

Definition 3.29:
Consider the C-stable system (3.91) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$ and $\boldsymbol{\mathcal{F}}_\mathbf{J}(t)$ and $\boldsymbol{\mathcal{F}}_\mathbf{N}(k)$ as defined in (2.13). Then the corresponding *proper and improper controllability Gramians* are defined as

$$\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{W}_\mathrm{p}} := \int_0^\infty \boldsymbol{\mathcal{F}}_\mathbf{J}(t)\boldsymbol{\mathcal{W}}_\mathrm{p}\boldsymbol{\mathcal{W}}_\mathrm{p}^\mathrm{T}\boldsymbol{\mathcal{F}}_\mathbf{J}(t)^\mathrm{T}\mathrm{d}t, \qquad \boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{W}_\mathrm{p}} := \sum_{k=0}^{\nu-1}\boldsymbol{\mathcal{F}}_\mathbf{N}(k)\boldsymbol{\mathcal{W}}_\mathrm{p}\boldsymbol{\mathcal{W}}_\mathrm{p}^\mathrm{T}\boldsymbol{\mathcal{F}}_\mathbf{N}(k)^\mathrm{T}. \tag{3.93}$$

◊

Since the Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{W}_\mathrm{p}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{W}_\mathrm{p}}$ span the controllability spaces, all reachable states $\mathbf{z}_\mathrm{p}(t)$ and $\mathbf{z}_\mathrm{i}(t)$ lie in the spaces spanned by these matrices. Furthermore, inserting the definitions of $\boldsymbol{\mathcal{F}}_\mathbf{J}(t)$ and $\boldsymbol{\mathcal{F}}_\mathbf{N}(k)$ into (3.93) yields the following Lemma.

Lemma 3.30:
Consider the C-stable system (3.91) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$. Assume that $\mathbf{T}$, $\mathbf{W}$ are matrices that transform system (3.91) into WCF as introduced in (2.9). Then the controllability Gramians $\boldsymbol{\mathcal{P}}_\mathrm{p}$ and $\boldsymbol{\mathcal{P}}_\mathrm{i}$ defined in (3.93) are of the following form

$$\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{W}_\mathrm{p}} = \mathbf{T}^{-1}\begin{bmatrix}\mathbf{P}_1 & 0\\ 0 & 0\end{bmatrix}\mathbf{T}^{-\mathrm{T}}, \qquad \boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{W}_\mathrm{p}} = \mathbf{T}^{-1}\begin{bmatrix}0 & 0\\ 0 & \mathbf{P}_2\end{bmatrix}\mathbf{T}^{-\mathrm{T}} \tag{3.94}$$

where

$$\mathbf{P}_1 = \int_0^\infty e^{\mathbf{J}t}\widehat{\mathbf{W}}_1\widehat{\mathbf{W}}_1^\mathrm{T}e^{\mathbf{J}^\mathrm{T}t}\mathrm{d}t \qquad \text{and} \qquad \mathbf{P}_2 = \sum_{k=0}^{\nu-1}\mathbf{N}^k\widehat{\mathbf{W}}_2\widehat{\mathbf{W}}_2^\mathrm{T}(\mathbf{N}^k)^\mathrm{T} \tag{3.95}$$

with $\boldsymbol{\mathcal{W}}_\mathrm{p} = \mathbf{W}\begin{bmatrix}\widehat{\mathbf{W}}_1\\ \widehat{\mathbf{W}}_2\end{bmatrix}$ and $\mathbf{J}$, $\mathbf{N}$ as defined in (2.9). ◊

Note that the Gramians $\mathbf{P}_1$ and $\mathbf{P}_2$ are the proper and improper controllability Gramians corresponding to the states $\mathbf{z}_1(t)$ and $\mathbf{z}_2(t)$ as defined in (2.12), respectively. Using the controllability Gramians, we can characterize the states that are difficult to reach or even unreachable, which play a significant role when reducing the system.

To compute the Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{W}_\mathrm{p}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{W}_\mathrm{p}}$ defined in (3.94), we utilize that they are the unique solutions of the following continuous-time and discrete-time projected Lyapunov equations

$$\begin{aligned}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{W}_\mathrm{p}}\boldsymbol{\mathcal{A}}^\mathrm{T} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{W}_\mathrm{p}}\boldsymbol{\mathcal{E}}^\mathrm{T} &= -\mathbf{P}_\mathrm{l}\boldsymbol{\mathcal{W}}_\mathrm{p}\boldsymbol{\mathcal{W}}_\mathrm{p}^\mathrm{T}\mathbf{P}_\mathrm{l}^\mathrm{T}, & \boldsymbol{\mathcal{P}}_\mathrm{p} &= \mathbf{P}_\mathrm{r}\boldsymbol{\mathcal{P}}_\mathrm{p}\mathbf{P}_\mathrm{r}^\mathrm{T},\\ \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{W}_\mathrm{p}}\boldsymbol{\mathcal{A}}^\mathrm{T} - \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{W}_\mathrm{p}}\boldsymbol{\mathcal{E}}^\mathrm{T} &= (\mathbf{I}-\mathbf{P}_\mathrm{l})\boldsymbol{\mathcal{W}}_\mathrm{p}\boldsymbol{\mathcal{W}}_\mathrm{p}^\mathrm{T}(\mathbf{I}-\mathbf{P}_\mathrm{l})^\mathrm{T}, & 0 &= \mathbf{P}_\mathrm{r}\boldsymbol{\mathcal{P}}_\mathrm{i}\mathbf{P}_\mathrm{r}^\mathrm{T}.\end{aligned} \tag{3.96}$$

Observability Gramians When considering systems with a linear output equation as in system (3.91), the initial conditions do not affect the observability behavior. Hence, the surrogate system (3.91) has equal observability properties as the homogeneous system in (2.8). Therefore, the same observability Gramians introduced in (2.20) encode the observability behavior of system (3.91), leading to the following definition.

Definition 3.31:
Consider the C-stable system (3.91) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$. Then the corresponding *proper and improper observability Gramians* are defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}} := \int_0^\infty \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)^{\mathrm{T}} \boldsymbol{\mathcal{C}}^{\mathrm{T}} \boldsymbol{\mathcal{C}} \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t) \mathrm{d}\tau, \qquad \boldsymbol{\mathcal{Q}}_{\mathrm{L,i}} := \sum_{k=0}^{\nu-1} \boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)^{\mathrm{T}} \boldsymbol{\mathcal{C}}^{\mathrm{T}} \boldsymbol{\mathcal{C}} \boldsymbol{\mathcal{F}}_{\mathbf{N}}(k), \tag{3.97}$$

where $\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)$ and $\boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)$ are as defined in (2.13). ♢

We compute these Gramians by solving the projected Lyapunov equations (2.22).

Controllability energies In this paragraph, we use the controllability Gramians from (3.93) to describe the controllability behavior of the system (3.91) and corresponding energies in more detail. To derive an energy measure based on the proper and improper input-to-state mappings defined in (3.92), we evaluate their energy norms defined in (3.19) and (3.83) and obtain the expressions

$$E(\boldsymbol{c}_{\mathrm{p,w_p}}) = \|\boldsymbol{c}_{\mathrm{p,w_p}}\|^2_{L_2\left([0,\infty),\mathbb{R}^{N\times(m+N_{\mathbf{Z}_0})}\right)} = \int_0^\infty \mathrm{tr}\left(\boldsymbol{c}_{\mathrm{p,w_p}}(t)\boldsymbol{c}_{\mathrm{p,w_p}}(t)^{\mathrm{T}}\right) \mathrm{d}t = \mathrm{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{p,w_p}}\right), \tag{3.98}$$

and

$$E(\boldsymbol{c}_{\mathrm{i,w_p}}) = \|\boldsymbol{c}_{\mathrm{i,w_p}}\|^2_{\ell_2\left(\mathbb{N},\mathbb{R}^{N\times(m+N_{\mathbf{Z}_0})}\right)} = \sum_{k=0}^{\nu-1} \mathrm{tr}\left(\boldsymbol{c}_{\mathrm{i,w_p}}(k)\boldsymbol{c}_{\mathrm{i,w_p}}(k)^{\mathrm{T}}\right) = \mathrm{tr}\left(\boldsymbol{\mathcal{P}}_{\mathrm{i,w_p}}\right). \tag{3.99}$$

Since the trace of the Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p,w_p}}$ is equal to the sum of its eigenvalues, it follows from the energy norm in (3.98) that small eigenvalues contribute only small amounts of energy to the system dynamics and correspond to the least significant states. The same properties can be observed for the improper Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i,w_p}}$. However, the improper system components encode algebraic system constraints. Neglecting states corresponding to nonzero eigenvalues could lead to physically meaningless dynamics so that only the states corresponding to zero eigenvalues are negligible.

Observability energies To describe the observability behavior of system (3.91), we evaluate the respective observability energies in this paragraph. Therefore, we evaluate the energy norms of the state-to-output mappings $\boldsymbol{o}_{\mathrm{L,p}}$ and $\boldsymbol{o}_{\mathrm{L,i}}$, as defined in (3.73). Since they coincide with the mappings introduced in (3.73) and (3.77), they provide the same energy norms as in (3.85) and (3.86), that are

$$E(\boldsymbol{o}_{\mathrm{L,p}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{L,p}}), \qquad E(\boldsymbol{o}_{\mathrm{L,i}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{L,i}}).$$

The states corresponding to large eigenvalues of $\mathcal{Q}_{\mathrm{L,p}}$ encode the dominant observability subspaces. Among the improper states $\mathbf{z}_{\mathrm{i}}(t)$, only those corresponding to zero eigenvalues of the improper Gramian $\mathcal{Q}_{\mathrm{L,i}}$ can be neglected since they do not affect the system dynamics.

The extended-input approach from this paragraph derives a system realization that is used to define suitable Gramians encoding the controllability and observability spaces. These Gramians and the resulting energy norms are summarized in Table 3.6.

	System (3.91) – differential component	System (3.91) – algebraic component
Transfer function	$\mathcal{G}_{\mathrm{L,p},\mathrm{w_p}}(s)$	$\mathcal{G}_{\mathrm{L,i},\mathrm{w_p}}(s)$
Controllability Gramian	$\mathcal{P}_{\mathrm{p},\mathrm{w_p}}$	$\mathcal{P}_{\mathrm{i},\mathrm{w_p}}$
Observability Gramian	$\mathcal{Q}_{\mathrm{L,p}}$	$\mathcal{Q}_{\mathrm{L,i}}$
Controllability energies	$E(\boldsymbol{c}_{\mathrm{p},\mathrm{w_p}}) = \mathrm{tr}\big(\mathcal{P}_{\mathrm{p},\mathrm{w_p}}\big)$	$E(\boldsymbol{c}_{\mathrm{i},\mathrm{w_p}}) = \mathrm{tr}\big(\mathcal{P}_{\mathrm{i},\mathrm{w_p}}\big)$
Observability energies	$E(\boldsymbol{o}_{\mathrm{L,p}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{L,p}})$	$E(\boldsymbol{o}_{\mathrm{L,i}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{L,i}})$

Table 3.6: Properties of system (3.54) corresponding to its extended-input representation.

3.2.2 Inhomogeneous first-order DAE systems with a quadratic output

As a second class of DAE systems, we investigate systems with a quadratic output that are of the form

$$\begin{aligned} \boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) &= \boldsymbol{\mathcal{A}}\mathbf{z}(t) + \boldsymbol{\mathcal{B}}\mathbf{u}(t), \qquad \mathbf{z}(0) = \mathbf{z}_0, \\ \mathbf{y}_{\mathrm{Q}}(t) &= \mathbf{z}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}\mathbf{z}(t), \end{aligned} \tag{3.100}$$

with a state equation as defined in (3.49), and an output equation including the symmetric output matrix $\boldsymbol{\mathcal{M}} \in \mathbb{R}^{N\times N}$ and the output $\mathbf{y}_{\mathrm{Q}}(t) \in \mathbb{R}$. We assume that the

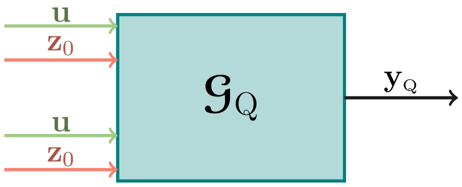


Figure 3.11: Structure of a first-order DAE system with a quadratic output.

matrix pencil $(\mathcal{A}, \mathcal{E})$ is regular and that the consistency conditions in (2.15) are satisfied. Figure 3.11 provides a sketch of the system structure where the inputs $\mathbf{u}$ and initial conditions $\mathbf{z}_0$ appear twice to indicate the quadratic output equation.

We decompose the output matrix $\mathcal{M}$ according to the WCF introduced in (2.9) into

$$\mathcal{M} = \mathbf{T}^{\mathrm{T}} \begin{bmatrix} \widetilde{\mathbf{M}}_{11} & \widetilde{\mathbf{M}}_{12} \\ \widetilde{\mathbf{M}}_{12}^{\mathrm{T}} & \widetilde{\mathbf{M}}_{22} \end{bmatrix} \mathbf{T}. \tag{3.101}$$

This decomposition is used in the following for theoretical considerations. As described for DAE systems with linear output equations in Section 3.2.1, we consider the differential and algebraic components of the system separately. Therefore, we decompose the output equation in (3.100) as

$$\begin{aligned} \mathbf{y}_{\mathrm{Q}}(t) &= \mathbf{z}_{\mathrm{p}}(t)^{\mathrm{T}} \mathcal{M} \mathbf{z}_{\mathrm{p}}(t) + \mathbf{z}_{\mathrm{p}}(t)^{\mathrm{T}} \mathcal{M} \mathbf{z}_{\mathrm{i}}(t) + \mathbf{z}_{\mathrm{i}}(t)^{\mathrm{T}} \mathcal{M} \mathbf{z}_{\mathrm{p}}(t) + \mathbf{z}_{\mathrm{i}}(t)^{\mathrm{T}} \mathcal{M} \mathbf{z}_{\mathrm{i}}(t) \\ &= \mathbf{z}_1(t)^{\mathrm{T}} \widetilde{\mathbf{M}}_{11} \mathbf{z}_1(t) + \mathbf{z}_1(t)^{\mathrm{T}} \widetilde{\mathbf{M}}_{12} \mathbf{z}_2(t) + \mathbf{z}_2(t)^{\mathrm{T}} \widetilde{\mathbf{M}}_{12}^{\mathrm{T}} \mathbf{z}_1(t) + \mathbf{z}_2(t)^{\mathrm{T}} \widetilde{\mathbf{M}}_{22} \mathbf{z}_2(t) \\ &=: \mathbf{y}_{\mathrm{pp}}(t) + \mathbf{y}_{\mathrm{pi}}(t) + \mathbf{y}_{\mathrm{ip}}(t) + \mathbf{y}_{\mathrm{ii}}(t) \end{aligned} \tag{3.102}$$

using the state components from (2.14) and (2.12). We observe that the output consists of four components. We note that the two output components $\mathbf{y}_{\mathrm{pi}}(t)$ and $\mathbf{y}_{\mathrm{ip}}(t)$ coincide. However, they are analyzed separately in this work as they span different observability spaces. Moreover, both components depend on a differential state and an algebraic one. Hence, there is no obvious categorization of the outputs into proper and improper ones.

In the following, we consider the subsystems corresponding to the different output components defined in (3.102) and investigate them individually. Figure 3.12 depicts the subsystem structure, where we again only add an input $\mathbf{u}$ to derive the algebraic components of the quadratic output since we assume the system satisfies the consistency conditions. Hence, the initial conditions are included implicitly.

To consider the differential initial conditions while analyzing the system with a quadratic output, we aim to apply the multi-system and extended-input approach introduced above. Therefore, in Section 3.2.2.1, we describe the multi-system approach for inhomogeneous DAE systems with a quadratic output. Afterwards, in Section 3.2.2.2, we apply the extended-input approach for this class of systems. For this purpose, we derive suitable transfer functions, the corresponding Gramians, and the resulting energy interpretations.

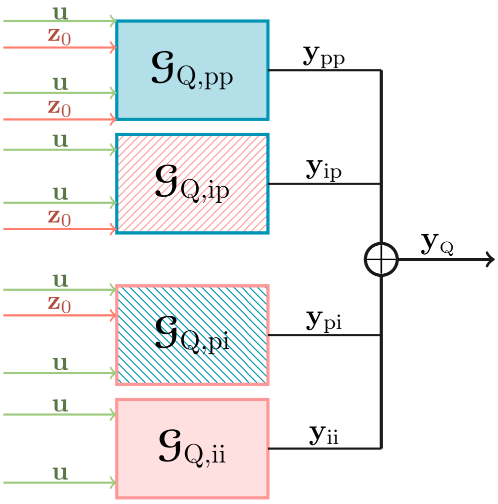


Figure 3.12: Structure of a first-order DAE system with a quadratic output - proper and improper components decoupled.

3.2.2.1 Multi-system approach for inhomogeneous first-order DAE systems with a quadratic output

We aim to modify the multi-system approach presented in [15] so that it applies to DAE systems (3.100) with a quadratic output equation. Within this approach, we derive subsystems for the different input- and initial condition-to-output mappings. Inserting the three state components $\mathbf{z}_{\mathrm{p},\mathfrak{B}}(t)$, $\mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t)$, and $\mathbf{z}_{\mathrm{i}}(t)$ from (3.55) and (2.14), respectively, into the quadratic output equation from (3.102) yields

$$\begin{aligned}\mathbf{y}_{\mathrm{Q}}(t) = {}& \mathbf{z}_{\mathrm{p},\mathfrak{B}}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{p},\mathfrak{B}}(t) + \mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{p},\mathfrak{B}}(t) + \mathbf{z}_{\mathrm{p},\mathfrak{B}}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t)\\ &+ \mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t) + \mathbf{z}_{\mathrm{p},\mathfrak{B}}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{i}}(t) + \mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{i}}(t)\\ &+ \mathbf{z}_{\mathrm{i}}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{p},\mathfrak{B}}(t) + \mathbf{z}_{\mathrm{i}}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t) + \mathbf{z}_{\mathrm{i}}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{i}}(t).\end{aligned} \tag{3.103}$$

We note that the decomposition of the output consists of nine components. Examining each respective subsystem individually would lead to extensive computations. Therefore, for the sake of simplicity, we consider only the extended-input approach presented below.

3.2.2.2 Extended-input approach for inhomogeneous first-order DAE systems with a quadratic output

We apply, in this paragraph, the extended-input approach, to consider the differential initial conditions while analyzing the respective system. For that, we derive a system representation with homogeneous differential initial conditions.

Transfer function Our objective is to describe the input- and initial condition-to-output behavior of the system (3.100). Therefore, we consider the different output components in (3.103), where, e.g., the first output component is

$$\mathbf{y}_{\mathrm{pp},\mathcal{BB}}(t) = \int_0^t \int_0^t \mathbf{u}(\tau_1)^{\mathrm{T}} \mathcal{B}^{\mathrm{T}} \mathcal{F}_{\mathbf{J}}(t-\tau_1)^{\mathrm{T}} \mathcal{M} \mathcal{F}_{\mathbf{J}}(t-\tau_2) \mathcal{B} \mathbf{u}(\tau_2) \mathrm{d}\tau_1 \mathrm{d}\tau_2.$$

We extract the kernel

$$\mathbf{g}_{\mathrm{Q,pp},\mathcal{BB}}(t_1, t_2) := \mathcal{B}^{\mathrm{T}} \mathcal{F}_{\mathbf{J}}(t_1)^{\mathrm{T}} \mathcal{M} \mathcal{F}_{\mathbf{J}}(t_2) \mathcal{B},$$

which encodes the input-to-output mapping corresponding to the first output component $\mathbf{z}_{\mathrm{p},\mathcal{B}}(t)^{\mathrm{T}} \mathcal{M} \mathbf{z}_{\mathrm{p},\mathcal{B}}(t)$. To derive the respective transfer function, we apply the 2-dimensional Laplace transform, which yields

$$\mathcal{G}_{\mathrm{Q,pp},\mathcal{BB}}(s_1, s_2) := \mathcal{B}^{\mathrm{T}} \mathbf{P}_{\mathrm{l}}^{\mathrm{T}} (s_1 \mathcal{E} - \mathcal{A})^{-\mathrm{T}} \mathcal{M} (s_2 \mathcal{E} - \mathcal{A})^{-1} \mathbf{P}_{\mathrm{l}} \mathcal{B}.$$

Applying this procedure to all output components in (3.103) and summing over the resulting transfer functions leads to the transfer function

$$\begin{aligned}
\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{p}}\mathcal{W}_{\mathrm{p}}}(s_1, s_2) &:= (\mathcal{B}^{\mathrm{T}} \mathbf{P}_{\mathrm{l}}^{\mathrm{T}} + \mathbf{Z}_0^{\mathrm{T}} \mathbf{P}_{\mathrm{l}}^{\mathrm{T}} + \mathcal{B}^{\mathrm{T}} (\mathbf{I} - \mathbf{P}_{\mathrm{l}})^{\mathrm{T}}) (s_1 \mathcal{E} - \mathcal{A})^{-\mathrm{T}} \\
&\qquad \cdot \mathcal{M} (s_2 \mathcal{E} - \mathcal{A})^{-1} (\mathbf{P}_{\mathrm{l}} \mathcal{B} + \mathbf{P}_{\mathrm{l}} \mathbf{Z}_0 + (\mathbf{I} - \mathbf{P}_{\mathrm{l}}) \mathcal{B}) \\
&= (\mathcal{B}^{\mathrm{T}} + \mathbf{Z}_{0,\mathrm{p}}^{\mathrm{T}}) (s_1 \mathcal{E} - \mathcal{A})^{-\mathrm{T}} \mathcal{M} (s_2 \mathcal{E} - \mathcal{A})^{-1} (\mathbf{Z}_{0,\mathrm{p}} + \mathcal{B}) \\
&= \begin{bmatrix} \mathbf{Z}_{0,\mathrm{p}}^{\mathrm{T}} \\ \mathcal{B}^{\mathrm{T}} \end{bmatrix} (s_1 \mathcal{E} - \mathcal{A})^{-\mathrm{T}} \mathcal{M} (s_2 \mathcal{E} - \mathcal{A})^{-1} \begin{bmatrix} \mathbf{Z}_{0,\mathrm{p}} & \mathcal{B} \end{bmatrix},
\end{aligned}$$

which encodes the overall input-to-output behavior. Using the definition of $\mathcal{W}_{\mathrm{p}}$ from (3.88) leads to the following definition.

Definition 3.32:
Consider the system (3.100) with a regular matrix pencil $(\mathcal{A}, \mathcal{E})$. Also, consider the matrix $\mathcal{W}_{\mathrm{p}}$ from (3.88) and assume that the consistency conditions in (2.15) are satisfied. Then the *transfer function* corresponding to the system is defined as

$$\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{p}}\mathcal{W}_{\mathrm{p}}}(s_1, s_2) = \mathcal{W}_{\mathrm{p}}^{\mathrm{T}} (s_1 \mathcal{E} - \mathcal{A})^{-\mathrm{T}} \mathcal{M} (s_2 \mathcal{E} - \mathcal{A})^{-1} \mathcal{W}_{\mathrm{p}}. \tag{3.104}$$

◊

The transfer function $\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{p}}\mathcal{W}_{\mathrm{p}}}$ from (3.104) has several system realizations. One of them is the following DAE system with a homogeneous differential initial condition,

$$\begin{aligned}
\mathcal{E} \dot{\mathbf{z}}(t) &= \mathcal{A} \mathbf{z}(t) + \mathcal{W}_{\mathrm{p}} \widetilde{\mathbf{u}}(t), \qquad \mathbf{P}_{\mathrm{r}} \mathbf{z}(0) = 0, \\
\mathbf{y}_{\mathrm{Q}}(t) &= \mathbf{z}(t)^{\mathrm{T}} \mathcal{M} \mathbf{z}(t),
\end{aligned} \tag{3.105}$$

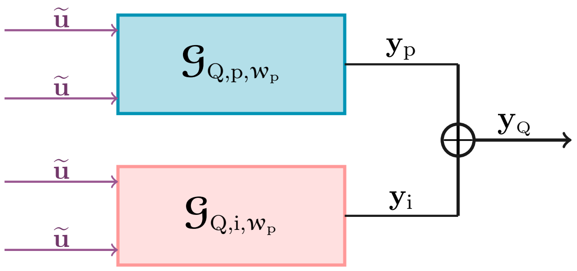

$\widetilde{\mathbf{u}}$
$\widetilde{\mathbf{u}}$
$\mathcal{G}_{\mathrm{Q,p},\mathcal{W}_{\mathrm{p}}}$
$\mathbf{y}_{\mathrm{p}}$
$\mathbf{y}_{\mathrm{Q}}$
$\widetilde{\mathbf{u}}$
$\widetilde{\mathbf{u}}$
$\mathcal{G}_{\mathrm{Q,i},\mathcal{W}_{\mathrm{p}}}$
$\mathbf{y}_{\mathrm{i}}$

Figure 3.13: Structure of a first-order DAE system with an extended input and a quadratic output - differential and algebraic components decoupled.

with $\mathcal{W}_{\mathrm{p},\mathcal{W}_{\mathrm{p}}\mathcal{W}_{\mathrm{p}}}$ as defined in (3.88) and a suitable input function $\widetilde{\mathbf{u}} \in L([0,\infty),\mathbb{R}^{m+n_{\mathbf{z}_0}})$. In the following, the surrogate system (3.105) is analyzed instead of the original system (3.100) so that the system analysis involves the initial conditions.

In the following, we investigate the controllability behavior of the right state added to the quadratic output equation in (3.105). Moreover, we investigate the observability of the right state under consideration of the left one. Therefore, we distinguish between the differential and the algebraic right state components in the following. Hence, we divide the transfer function from (3.104) into a transfer function corresponding to the differential right state and one that corresponds to the algebraic one while considering all left states generated by the system (3.105), which results in the proper and the improper transfer function

$$\begin{aligned}
\mathcal{G}_{\mathrm{Q,p},\mathcal{W}_{\mathrm{p}}}(s_1,s_2) &:= \mathcal{W}_{\mathrm{p}}^{\mathrm{T}}(s_1\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathcal{M}(s_2\mathcal{E}-\mathcal{A})^{-1}\mathbf{P}_{\mathrm{l}}\mathcal{W}_{\mathrm{p}},\\
\mathcal{G}_{\mathrm{Q,i},\mathcal{W}_{\mathrm{p}}}(s_1,s_2) &:= \mathcal{W}_{\mathrm{p}}^{\mathrm{T}}(s_1\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathcal{M}(s_2\mathcal{E}-\mathcal{A})^{-1}(\mathbf{I}-\mathbf{P}_{\mathrm{l}})\mathcal{W}_{\mathrm{p}},
\end{aligned}$$

respectively. The respective structure is depicted in Figure 3.13.

Controllability Gramians To investigate the controllability behavior of the surrogate system (3.105), we aim to derive its controllability Gramians that encode the respective controllability space. We note that the state equation of system (3.100) coincides with the one corresponding to the DAE system (3.54) with a linear output equation. Therefore, the same mappings $\boldsymbol{c}_{\mathrm{p}}$ and $\boldsymbol{c}_{\mathrm{i}}$ from (3.92) encode the input-to-state behavior, and hence, the same controllability Gramians defined in (3.29) encode the controllability spaces.

Definition 3.33:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathcal{A},\mathcal{E})$. Then the corresponding *proper and improper controllability Gramians* are defined as

$$\mathcal{P}_{\mathrm{p},\mathcal{W}_{\mathrm{p}}} := \int_0^{\infty} \mathcal{F}_{\mathbf{J}}(t)\mathcal{W}_{\mathrm{p}}\mathcal{W}_{\mathrm{p}}^{\mathrm{T}}\mathcal{F}_{\mathbf{J}}(t)^{\mathrm{T}}\mathrm{d}t, \quad \mathcal{P}_{\mathrm{i},\mathcal{W}_{\mathrm{p}}} := \sum_{k=0}^{\nu-1}\mathcal{F}_{\mathbf{N}}(k)\mathcal{W}_{\mathrm{p}}\mathcal{W}_{\mathrm{p}}^{\mathrm{T}}\mathcal{F}_{\mathbf{N}}(k)^{\mathrm{T}},$$

where $\mathcal{F}_{\mathbf{J}}(t)$ and $\mathcal{F}_{\mathbf{N}}(k)$ are as defined in (2.13). ◊

We compute these Gramians solving the projected Lyapunov equations in (3.96).

Observability Gramians To describe the controllability behavior of system (3.105), we aim to derive the respective observability Gramians. Therefore, we decompose the output as described in (3.102) and investigate the proper and improper components separately. For a better understanding, we can rewrite $\mathbf{y}_{\mathrm{Q}}(t)$ by defining the state-dependent function $\mathcal{C}(\mathbf{z}(t)) := \mathbf{z}(t)^{\mathrm{T}}\mathcal{M}$. Applying this representation to the decomposed output in (3.102) leads to

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

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

In this paragraph, we define proper and improper observability Gramians encoding the observability behavior of state the $\mathbf{z}_{\mathrm{p}}(t)$ and $\mathbf{z}_{\mathrm{i}}(t)$, respectively, corresponding to $\mathcal{C}(\mathbf{z}_{\mathrm{p}}(t))$ and $\mathcal{C}(\mathbf{z}_{\mathrm{i}}(t))$. That way, we obtain a proper observability Gramian corresponding to the outputs $\mathbf{y}_{\mathrm{pp}}(t)$ and $\mathbf{y}_{\mathrm{ip}}(t)$ and an improper Gramian corresponding to $\mathbf{y}_{\mathrm{pi}}(t)$ and $\mathbf{y}_{\mathrm{ii}}(t)$. We want to emphasize that the observability of the right state $\mathbf{z}_{\mathrm{p}}(t)$ (or $\mathbf{z}_{\mathrm{i}}(t)$) does not only depend on the matrix $\mathcal{M}$ but also on the space in which the left state $\mathbf{z}_{\mathrm{p}}(t)$ or $\mathbf{z}_{\mathrm{i}}(t)$ live. So it is expected that the observability Gramian for $\mathbf{z}_{\mathrm{p}}(t)$ (or $\mathbf{z}_{\mathrm{i}}(t)$) depend on $\mathcal{P}_{\mathrm{p}}$ and $\mathcal{P}_{\mathrm{i}}$ as well.

Proper observability Gramian In this paragraph, 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.

We start investigating the first component of the output $\mathbf{y}_{\mathrm{pp}}(t) = \mathbf{z}_{\mathrm{p}}(t)^{\mathrm{T}}\mathcal{M}\mathbf{z}_{\mathrm{p}}(t)$ that includes two proper states. Inserting the solution trajectories of the states leads to

$$\begin{aligned}
\mathbf{y}_{\mathrm{pp}}(t) &= \int_0^t \int_0^t \widetilde{\mathbf{u}}(\tau_1)^{\mathrm{T}} \mathcal{W}_{\mathrm{p}}^{\mathrm{T}} \mathcal{F}_{\mathbf{J}}(t-\tau_1)^{\mathrm{T}} \mathcal{M} \mathcal{F}_{\mathbf{J}}(t-\tau_2) \mathcal{W}_{\mathrm{p}} \widetilde{\mathbf{u}}(\tau_2) \mathrm{d}\tau_2 \mathrm{d}\tau_1 \\
&= \int_0^t \int_0^t \operatorname{vec}\big(\mathcal{W}_{\mathrm{p}}^{\mathrm{T}} \mathcal{F}_{\mathbf{J}}(t-\tau_1)^{\mathrm{T}} \mathcal{M} \mathcal{F}_{\mathbf{J}}(t-\tau_2) \mathcal{W}_{\mathrm{p}}\big)^{\mathrm{T}} \left(\widetilde{\mathbf{u}}(\tau_2) \otimes \widetilde{\mathbf{u}}(\tau_1)\right) \mathrm{d}\tau_2 \mathrm{d}\tau_1.
\end{aligned}$$

We identify the input-to-state mapping $\boldsymbol{c}_{\mathrm{p},\mathrm{w}_\mathrm{p}}(t) = \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\boldsymbol{\mathcal{W}}_\mathrm{p}$ defined in (3.92) within the output $\mathbf{y}_\mathrm{pp}(t)$ and extract the remaining state-to-output mapping

$$\boldsymbol{o}_{\mathrm{pp},\mathrm{w}_\mathrm{p}}(t_1, t_2) := \boldsymbol{\mathcal{W}}_\mathrm{p}^\mathrm{T}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t_1)\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t_2),$$

which encodes the observability of the differential right state while considering the left differential state. Based on this mapping, we define a matrix

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

that spans the respective observability space using the definition of the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathrm{w}_\mathrm{p}}$ from (3.93).

Definition 3.34:
Consider the C-stable system (3.100) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$ and the corresponding proper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathrm{w}_\mathrm{p}}$ as defined in (3.93). The *proper-proper observability Gramian* $\boldsymbol{\mathcal{Q}}_{\mathrm{pp},\mathrm{w}_\mathrm{p}}$ corresponding to the output $\mathbf{y}_\mathrm{pp}$ is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{pp},\mathrm{w}_\mathrm{p}} := \int_0^\infty \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t_2)^\mathrm{T}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathrm{w}_\mathrm{p}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t_2)\mathrm{d}t_2 \tag{3.106}$$

where $\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)$ is defined as in (2.13). ♢

To describe the connection between the Weierstraß-canonical representation (2.11) and the system (3.100) in the observability Gramian $\boldsymbol{\mathcal{Q}}_\mathrm{pp}$, we insert the function $\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)$, which leads to the following Lemma.

Lemma 3.35:
Consider the C-stable system (3.100) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$ and the corresponding proper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathrm{w}_\mathrm{p}}$ as defined in (3.93). The proper-proper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{pp},\mathrm{w}_\mathrm{p}}$ corresponding to the output $\mathbf{y}_\mathrm{pp}$ is of the form

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

where

$$\mathbf{Q}_{11} := \int_0^\infty e^{\mathbf{J}^\mathrm{T}t}\widetilde{\mathbf{M}}_{11}\mathbf{P}_1\widetilde{\mathbf{M}}_{11}e^{\mathbf{J}t}\mathrm{d}t \tag{3.107}$$

with the proper controllability Gramian $\mathbf{P}_1$ as defined in (3.95) and $\widetilde{\mathbf{M}}_{11}$ as defined in (3.101). ♢

Note that $\mathbf{Q}_{11}$ is the proper-proper observability Gramian corresponding to the state $\mathbf{z}_1(t)$ defined in (2.12). The following theorem describes how the Gramian $\mathcal{Q}_{\mathrm{pp},w_\mathrm{p}}$ is computed.

Theorem 3.36:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathcal{A}, \mathcal{E})$ and the corresponding proper controllability Gramian $\mathcal{P}_{\mathrm{p},w_\mathrm{p}}$ as defined in (3.93). The proper observability Gramian $\mathcal{Q}_{\mathrm{pp},w_\mathrm{p}}$ as defined in (3.106) solves the projected Lyapunov equation

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

where the projection matrices $\mathbf{P}_\mathrm{l}$ and $\mathbf{P}_\mathrm{r}$ are defined as in (2.10). ◊

Proof. We first observe that the projection condition is naturally satisfied since $\mathcal{Q}_{\mathrm{pp}}$ is by definition equal to $\mathbf{W}^{-\mathrm{T}}\left[\begin{smallmatrix}\mathbf{Q}_{11} & 0\\ 0 & 0\end{smallmatrix}\right]\mathbf{W}^{-1}$ with $\mathbf{Q}_{11}$ as defined in (3.107). To prove that $\mathcal{Q}_{\mathrm{pp},w_\mathrm{p}}$ satisfies the remaining Lyapunov equation, we show that $\mathbf{Q}_{11}$ solves the Lyapunov equation

$$\mathbf{J}^\mathrm{T}\mathbf{Q}_{11} + \mathbf{Q}_{11}\mathbf{J} = -\widetilde{\mathbf{M}}_{11}\mathbf{P}_1\widetilde{\mathbf{M}}_{11}. \tag{3.108}$$

For that we insert $\mathbf{Q}_{11}$ into (3.108) and obtain

$$\begin{aligned}\int_0^\infty \left(\mathbf{J}^\mathrm{T}e^{\mathbf{J}^\mathrm{H}t}\widetilde{\mathbf{M}}_{11}\mathbf{P}_1\widetilde{\mathbf{M}}_{11}e^{\mathbf{J}t} + e^{\mathbf{J}^\mathrm{H}t}\widetilde{\mathbf{M}}_{11}\mathbf{P}_1\widetilde{\mathbf{M}}_{11}e^{\mathbf{J}t}\mathbf{J}\right)\mathrm{d}t &= \left[e^{\mathbf{J}^\mathrm{H}t}\widetilde{\mathbf{M}}_{11}\mathbf{P}_1\widetilde{\mathbf{M}}_{11}e^{\mathbf{J}t}\right]_0^\infty\\ &= -\widetilde{\mathbf{M}}_{11}\mathbf{P}_1\widetilde{\mathbf{M}}_{11}.\end{aligned}$$

Moreover, we insert the WCF of $\mathcal{E}$ and $\mathcal{A}$ and the definition of $\mathbf{P}_\mathrm{r}$ into the Lyapunov equation to obtain

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

such that (3.108) implies the statement, since $\mathbf{T}$ is a regular matrix. □

Now we consider the third output component $\mathbf{y}_{\mathrm{ip}}(t) = \mathbf{z}_\mathrm{i}(t)^\mathrm{T}\mathcal{M}\mathbf{z}_\mathrm{p}(t)$. We insert the states $\mathbf{z}_\mathrm{p}(t)$ and $\mathbf{z}_\mathrm{i}(t)$ and obtain

$$\begin{aligned}\mathbf{y}_{\mathrm{ip}}(t) &= -\sum_{k=0}^{\nu-1}\int_0^t (\widetilde{\mathbf{u}}^{(k)}(t))^\mathrm{T}\mathcal{W}_\mathrm{p}^\mathrm{T}\mathcal{F}_\mathbf{N}(k)^\mathrm{T}\mathcal{M}\mathcal{F}_\mathbf{J}(t-\tau)\mathcal{W}_\mathrm{p}\widetilde{\mathbf{u}}(\tau)\mathrm{d}\tau\\ &= -\sum_{k=0}^{\nu-1}\int_0^t \mathrm{vec}\big(\mathcal{W}_\mathrm{p}^\mathrm{T}\mathcal{F}_\mathbf{N}(k)^\mathrm{T}\mathcal{M}\mathcal{F}_\mathbf{J}(t-\tau)\mathcal{W}_\mathrm{p}\big)^\mathrm{T}\left(\widetilde{\mathbf{u}}(\tau)\otimes\widetilde{\mathbf{u}}^{(k)}(t)\right)\mathrm{d}\tau.\end{aligned}$$

We identify the controllability mapping $\boldsymbol{c}_{\mathrm{p},\mathcal{W}_\mathrm{p}}(t) = \boldsymbol{\mathcal{F}}_\mathbf{J}(t)\boldsymbol{\mathcal{W}}_\mathrm{p}$ within the output $\mathbf{y}_\mathrm{ip}(t)$ and define the remaining observability mapping

$$\boldsymbol{o}_{\mathrm{ip},\mathcal{W}_\mathrm{p}}(t,k) := \boldsymbol{\mathcal{W}}_\mathrm{p}^\mathrm{T}\boldsymbol{\mathcal{F}}_\mathbf{N}(k)^\mathrm{T}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_\mathbf{J}(t),$$

that is used to describe the observability behavior of the output $\mathbf{y}_\mathrm{ip}(t)$. Therefore, we construct a matrix

$$\begin{aligned}\boldsymbol{\mathcal{Q}}_{\mathrm{ip},\mathcal{W}_\mathrm{p}} &:= \sum_{k=0}^{\nu-1}\int_0^\infty \boldsymbol{o}_{\mathrm{ip},\mathcal{W}_\mathrm{p}}(t,k)^\mathrm{T}\boldsymbol{o}_{\mathrm{ip},\mathcal{W}_\mathrm{p}}(t,k)\mathrm{d}t\\ &= \sum_{k=0}^{\nu-1}\int_0^\infty \boldsymbol{\mathcal{F}}_\mathbf{J}(t)^\mathrm{T}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_\mathbf{N}(k)\boldsymbol{\mathcal{W}}_\mathrm{p}\boldsymbol{\mathcal{W}}_\mathrm{p}^\mathrm{T}\boldsymbol{\mathcal{F}}_\mathbf{N}(k)^\mathrm{T}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_\mathbf{J}(t)\mathrm{d}t\\ &= \int_0^\infty \boldsymbol{\mathcal{F}}_\mathbf{J}(t)^\mathrm{T}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_\mathrm{i}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_\mathbf{J}(t)\mathrm{d}t,\end{aligned}$$

that spans the observability space of the state $\mathbf{z}_\mathrm{p}(t)$ while considering the controllability space of $\mathbf{z}_\mathrm{i}(t)$.

Definition 3.37:
Consider the C-stable system (3.100) with a regular matrix pencil $(\boldsymbol{\mathcal{A}},\boldsymbol{\mathcal{E}})$ and the corresponding improper controllability Gramian $\boldsymbol{\mathcal{P}}_\mathrm{i}$ as defined in (3.93). The *improper-proper observability Gramian* $\boldsymbol{\mathcal{Q}}_{\mathrm{ip},\mathcal{W}_\mathrm{p}}$ corresponding to the output $\mathbf{y}_\mathrm{ip}$ is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{ip},\mathcal{W}_\mathrm{p}} = \int_0^\infty \boldsymbol{\mathcal{F}}_\mathbf{J}(t)^\mathrm{T}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_\mathrm{i}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_\mathbf{J}(t)\mathrm{d}t, \tag{3.109}$$

where $\boldsymbol{\mathcal{F}}_\mathbf{J}(t)$ is defined in(2.13). ◇

Lemma 3.38:
Consider the C-stable system (3.100) with a regular matrix pencil $(\boldsymbol{\mathcal{A}},\boldsymbol{\mathcal{E}})$ and the corresponding improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{W}_\mathrm{p}}$ as defined in (3.93). The improper-proper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{ip},\mathcal{W}_\mathrm{p}}$ corresponding to the output $\mathbf{y}_\mathrm{ip}$ is of the form

$$\boldsymbol{\mathcal{Q}}_{\mathrm{ip},\mathcal{W}_\mathrm{p}} := \mathbf{W}^{-\mathrm{T}}\begin{bmatrix}\mathbf{Q}_{21} & 0\\ 0 & 0\end{bmatrix}\mathbf{W}^{-1}$$

where

$$\mathbf{Q}_{21} := \int_0^\infty e^{\mathbf{J}^\mathrm{T}t}\widetilde{\mathbf{M}}_{12}\mathbf{P}_2\widetilde{\mathbf{M}}_{12}^\mathrm{T}e^{\mathbf{J}t}\mathrm{d}t \tag{3.110}$$

with the improper controllability Gramian $\mathbf{P}_2$ as defined in (3.95) and $\widetilde{\mathbf{M}}_{12}$ as defined in (3.101). ◇

Theorem 3.39:
Consider the C-stable system (3.100) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$ and the corresponding improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{w}_{\mathrm{p}}}$ as defined in (3.93). The improper-proper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{ip},\boldsymbol{w}_{\mathrm{p}}}$ solves the projected Lyapunov equation

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

where the projection matrices $\mathbf{P}_{\mathrm{l}}$ and $\mathbf{P}_{\mathrm{r}}$ are defined as in (2.10). ♢

Proof. The proof follows the same argumentation as for Theorem 3.36. □

We can combine the two proper observability Gramians to obtain one Gramian that encodes the observability behavior of the differential states $\mathbf{z}_{\mathrm{p}}(t)$ independent of the second state, that is, the observability of the output $\mathbf{y}_{\mathrm{P}}(t) = \mathbf{z}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}\mathbf{z}_{\mathrm{p}}(t)$ for an arbitrary state $\mathbf{z}(t)$ generated by system (3.100). Since the sum $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{w}_{\mathrm{p}}} + \boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{w}_{\mathrm{p}}}$ spans the full controllability space of the state $\mathbf{z}(t)$, the proper observability Gramian corresponding to proper and improper left states is given by

$$\boldsymbol{\mathcal{Q}}_{\mathrm{Q,p},\boldsymbol{w}_{\mathrm{p}}} = \int_0^{\infty} \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{w}_{\mathrm{p}}} + \boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{w}_{\mathrm{p}}})\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\mathrm{d}t = \boldsymbol{\mathcal{Q}}_{\mathrm{pp},\boldsymbol{w}_{\mathrm{p}}} + \boldsymbol{\mathcal{Q}}_{\mathrm{ip},\boldsymbol{w}_{\mathrm{p}}}.$$

We summarize this paragraph with the following definition.

Definition 3.40:
Consider the C-stable system (3.100) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$, and the corresponding proper and improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{w}_{\mathrm{p}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{w}_{\mathrm{p}}}$ as defined in (3.93). The proper observability Gramian corresponding to the output $\mathbf{y}_{\mathrm{P}}(t) = \mathbf{y}_{\mathrm{pp}}(t) + \mathbf{y}_{\mathrm{ip}}(t)$ is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{Q,p},\boldsymbol{w}_{\mathrm{p}}} := \boldsymbol{\mathcal{Q}}_{\mathrm{pp},\boldsymbol{w}_{\mathrm{p}}} + \boldsymbol{\mathcal{Q}}_{\mathrm{ip},\boldsymbol{w}_{\mathrm{p}}}, \tag{3.111}$$

with $\boldsymbol{\mathcal{Q}}_{\mathrm{pp},\boldsymbol{w}_{\mathrm{p}}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{ip},\boldsymbol{w}_{\mathrm{p}}}$ as defined in (3.106) and (3.109), respectively. ♢

Improper observability Gramians In this paragraph, we investigate the observability behavior of the outputs $\mathbf{y}_{\mathrm{pi}}(t) := \mathbf{z}_{\mathrm{p}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}\mathbf{z}_{\mathrm{i}}(t)$ and $\mathbf{y}_{\mathrm{ii}}(t) := \mathbf{z}_{\mathrm{i}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}\mathbf{z}_{\mathrm{i}}(t)$. Both outputs describe the observability of an algebraic (right) state $\mathbf{z}_{\mathrm{i}}(t)$ while considering either a differential state or an algebraic one multiplied from the left.

The state $\mathbf{y}_{\mathrm{pi}}(t)$ is equal to

$$\begin{aligned}
\mathbf{y}_{\mathrm{pi}}(t) &= \int_0^t \sum_{k=0}^{\nu-1} \tilde{\mathbf{u}}(\tau)^{\mathrm{T}}\boldsymbol{\mathcal{W}}_{\mathrm{p}}^{\mathrm{T}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t-\tau)^{\mathrm{T}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)\boldsymbol{\mathcal{W}}_{\mathrm{p}}\tilde{\mathbf{u}}^{(k)}(t)\mathrm{d}\tau \\
&= \int_0^t \sum_{k=0}^{\nu-1} \mathrm{vec}\big(\boldsymbol{\mathcal{W}}_{\mathrm{p}}^{\mathrm{T}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t-\tau)^{\mathrm{T}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)\boldsymbol{\mathcal{W}}_{\mathrm{p}}\big)^{\mathrm{T}} \left(\tilde{\mathbf{u}}^{(k)}(t) \otimes \tilde{\mathbf{u}}(\tau)\right) \mathrm{d}\tau.
\end{aligned}$$

We identify the improper controllability mapping $\boldsymbol{c}_{\mathrm{i},\mathcal{W}_\mathrm{p}}(k) = \mathcal{F}_\mathbf{N}(k)\mathcal{W}_\mathrm{p}$ within the output $\mathbf{y}_\mathrm{pi}(t)$ and the remaining observability mapping

$$\boldsymbol{o}_{\mathrm{pi},\mathcal{W}_\mathrm{p}}(t,k) = \mathcal{W}_\mathrm{p}^\mathrm{T}\mathcal{F}_\mathbf{J}(t)^\mathrm{T}\mathcal{M}\mathcal{F}_\mathbf{N}(k)$$

that is used to define a matrix

$$\begin{aligned}\mathcal{Q}_{\mathrm{pi},\mathcal{W}_\mathrm{p}} &= \int_0^\infty \sum_{k=0}^{\nu-1} \boldsymbol{o}_{\mathrm{pi},\mathcal{W}_\mathrm{p}}(t,k)^\mathrm{T}\boldsymbol{o}_{\mathrm{pi},\mathcal{W}_\mathrm{p}}(t,k)\mathrm{d}t \\ &= \int_0^\infty \sum_{k=0}^{\nu-1} \mathcal{F}_\mathbf{N}(k)^\mathrm{T}\mathcal{M}\mathcal{F}_\mathbf{J}(t)\mathcal{W}_\mathrm{p}\mathcal{W}_\mathrm{p}^\mathrm{T}\mathcal{F}_\mathbf{J}(t)^\mathrm{T}\mathcal{M}\mathcal{F}_\mathbf{N}(k)\mathrm{d}t \\ &= \sum_{k=0}^{\nu-1} \mathcal{F}_\mathbf{N}(k)^\mathrm{T}\mathcal{M}\mathcal{P}_{\mathrm{p},\mathcal{W}_\mathrm{p}}\mathcal{M}\mathcal{F}_\mathbf{N}(k),\end{aligned}$$

that spans the observability space of the state $\mathbf{z}_\mathrm{i}(t)$ while considering the controllability space of the state $\mathbf{z}_\mathrm{p}(t)$.

Definition 3.41:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathcal{A}, \mathcal{E})$ and the corresponding proper controllability Gramian $\mathcal{P}_{\mathrm{p},\mathcal{W}_\mathrm{p}}$ as defined in (3.93). The *proper-improper observability Gramian* $\mathcal{Q}_{\mathrm{pi},\mathcal{W}_\mathrm{p}}$ corresponding to the output $\mathbf{y}_\mathrm{pi}$ is defined as

$$\mathcal{Q}_{\mathrm{pi},\mathcal{W}_\mathrm{p}} = \sum_{k=0}^{\nu-1} \mathcal{F}_\mathbf{N}(k)^\mathrm{T}\mathcal{M}\mathcal{P}_{\mathrm{p},\mathcal{W}_\mathrm{p}}\mathcal{M}\mathcal{F}_\mathbf{N}(k), \tag{3.112}$$

where $\mathcal{F}_\mathbf{N}(k)$ is defines as in (2.13). ◊

We insert the mapping $\mathcal{F}_\mathbf{N}(k)$ to derive the following Lemma.

Lemma 3.42:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathcal{A}, \mathcal{E})$ and the corresponding proper controllability Gramian $\mathcal{P}_{\mathrm{p},\mathcal{W}_\mathrm{p}}$ as defined in (3.93). The proper-improper observability Gramian $\mathcal{Q}_{\mathrm{pi},\mathcal{W}_\mathrm{p}}$ is equal to the following representation

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

where

$$\mathbf{Q}_{12} := \sum_{k=0}^{\nu-1}(-\mathbf{N}^k)^\mathrm{T}\widetilde{\mathbf{M}}_{12}^\mathrm{T}\mathbf{P}_1\widetilde{\mathbf{M}}_{12}(-\mathbf{N}^k), \tag{3.113}$$

with the proper controllability Gramian $\mathbf{P}_1$ defined as in (3.95) and $\widetilde{\mathbf{M}}_{12}$ as in (3.101).◊

The Gramian $\mathbf{Q}_{12}$ is the improper observability Gramian corresponding to the algebraic (right) state $\mathbf{z}_2(t)$ and the differential (left) state $\mathbf{z}_1(t)$ defined in (2.12). Hence, Lemma 3.42 describes the relation between the observability of the system in WCF (2.11) and the original system (3.100).

The following theorem is used to compute the Gramian $\mathcal{Q}_{\mathrm{pi},w_\mathrm{p}}$.

Theorem 3.43:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathcal{A}, \mathcal{E})$ and the corresponding proper controllability Gramian $\mathcal{P}_{\mathrm{p},w_\mathrm{p}}$ as defined in (3.93). The proper-improper observability Gramian $\mathcal{Q}_{\mathrm{pi},w_\mathrm{p}}$ solves the projected Lyapunov equation

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

where the projection matrices $\mathbf{P}_\mathrm{l}$ and $\mathbf{P}_\mathrm{r}$ are as defined in (2.10). ◊

Proof. To prove the projection condition, we derive

$$\mathbf{P}_\mathrm{L}^\mathrm{T}\mathcal{Q}_{\mathrm{pi},w_\mathrm{p}}\mathbf{P}_\mathrm{L} = \mathbf{W}^{-\mathrm{T}}\begin{bmatrix}\mathbf{I} & 0\\ 0 & 0\end{bmatrix}\mathbf{W}^\mathrm{T}\mathbf{W}^{-\mathrm{T}}\begin{bmatrix}0 & 0\\ 0 & \mathbf{Q}_{12}\end{bmatrix}\mathbf{W}^{-1}\mathbf{W}\begin{bmatrix}\mathbf{I} & 0\\ 0 & 0\end{bmatrix}\mathbf{W}^{-1} = 0.$$

Moreover, we show that the Gramian $\mathbf{Q}_{12}$ defined in (3.113) solves the discrete-time Lyapunov equation

$$\mathbf{Q}_{12} - \mathbf{N}^\mathrm{T}\mathbf{Q}_{12}\mathbf{N} = \widetilde{\mathbf{M}}_{12}^\mathrm{T}\mathbf{P}_1\widetilde{\mathbf{M}}_{12}. \tag{3.114}$$

For that, we insert the definition of $\mathbf{Q}_{12}$ into (3.114). This results in

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

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

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

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

which proves the statement. □

Now, we consider the fourth output component $\mathbf{y}_{\mathrm{ii}}(t)$ that describes the observability space of the algebraic (right) state $\mathbf{z}_{\mathrm{i}}(t)$ for an algebraic (left) state. The respective output component is equal to

$$\begin{aligned}\mathbf{y}_{\mathrm{ii}}(t) &= \sum_{k=0}^{\nu-1}\sum_{\ell=0}^{\nu-1}(\widetilde{\mathbf{u}}^{(k)}(t))^{\mathrm{T}}\mathcal{W}_{\mathrm{p}}^{\mathrm{T}}\mathfrak{F}_{\mathbf{N}}(k)^{\mathrm{T}}\mathcal{M}\mathfrak{F}_{\mathbf{N}}(\ell)\mathcal{W}_{\mathrm{p}}\widetilde{\mathbf{u}}^{(\ell)}(t)\\ &= \sum_{k=0}^{\nu-1}\sum_{\ell=0}^{\nu-1}\mathrm{vec}\big(\mathcal{W}_{\mathrm{p}}^{\mathrm{T}}\mathfrak{F}_{\mathbf{N}}(k)^{\mathrm{T}}\mathcal{M}\mathfrak{F}_{\mathbf{N}}(\ell)\mathcal{W}_{\mathrm{p}}\big)^{\mathrm{T}}\left(\widetilde{\mathbf{u}}^{(\ell)}(t)\otimes\widetilde{\mathbf{u}}^{(k)}(t)\right).\end{aligned}$$

We identify the improper controllability mapping $\boldsymbol{c}_{\mathrm{i},\mathcal{W}_{\mathrm{p}}}(\ell) = \mathfrak{F}_{\mathbf{N}}(\ell)\mathcal{W}_{\mathrm{p}}$ and the remaining observability mapping

$$\boldsymbol{o}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}}(k,\ell) = \mathcal{W}_{\mathrm{p}}^{\mathrm{T}}\mathfrak{F}_{\mathbf{N}}(k)^{\mathrm{T}}\mathbf{M}\mathfrak{F}_{\mathbf{N}}(\ell).$$

Based on this mapping $\boldsymbol{o}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}}(k,\ell)$, we define a matrix

$$\begin{aligned}\mathcal{Q}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}} &:= \sum_{k=0}^{\nu-1}\sum_{\ell=0}^{\nu-1}\boldsymbol{o}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}}(k,\ell)^{\mathrm{H}}\boldsymbol{o}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}}(k,\ell)\\ &= \sum_{k=0}^{\nu-1}\sum_{\ell=0}^{\nu-1}\mathfrak{F}_{\mathbf{N}}(\ell)^{\mathrm{T}}\mathcal{M}\mathfrak{F}_{\mathbf{N}}(k)\mathcal{W}_{\mathrm{p}}\mathcal{W}_{\mathrm{p}}^{\mathrm{T}}\mathfrak{F}_{\mathbf{N}}(k)^{\mathrm{T}}\mathcal{M}\mathfrak{F}_{\mathbf{N}}(\ell)\\ &= \sum_{\ell=0}^{\nu-1}\mathfrak{F}_{\mathbf{N}}(\ell)^{\mathrm{T}}\mathcal{M}\mathcal{P}_{\mathrm{i},\mathcal{W}_{\mathrm{p}}}\mathcal{M}\mathfrak{F}_{\mathbf{N}}(\ell),\end{aligned}$$

which spans the observability space of the (right) algebraic state $\mathbf{z}_{\mathrm{i}}(t)$ and a (left) algebraic state $\mathbf{z}_{\mathrm{i}}(t)$.

Definition 3.44:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathcal{A},\mathcal{E})$ and the corresponding improper controllability Gramian $\mathcal{P}_{\mathrm{i},\mathcal{W}_{\mathrm{p}}}$ as defined in (3.93). The *improper-improper observability Gramian* $\mathcal{Q}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}}$ corresponding to the output $\mathbf{y}_{\mathrm{ii}}$ is defined as

$$\mathcal{Q}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}} := \sum_{\ell=0}^{\nu-1}\mathfrak{F}_{\mathbf{N}}(\ell)^{\mathrm{T}}\mathcal{M}\mathcal{P}_{\mathrm{i},\mathcal{W}_{\mathrm{p}}}\mathcal{M}\mathfrak{F}_{\mathbf{N}}(\ell),$$

where $\mathfrak{F}_{\mathbf{N}}(\ell)$ is defined as in (2.13). ◊

Lemma 3.45:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathcal{A},\mathcal{E})$ and the corresponding improper controllability Gramian $\mathcal{P}_{\mathrm{i},\mathcal{W}_{\mathrm{p}}}$ as defined in (3.93). The improper-improper observability Gramian $\mathcal{Q}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}}$ is equal to the following representation

$$\mathcal{Q}_{\mathrm{ii},\mathcal{W}_{\mathrm{p}}} = \mathbf{W}^{-\mathrm{T}}\begin{bmatrix}0 & 0\\ 0 & \mathbf{Q}_{22}\end{bmatrix}\mathbf{W}^{-1},$$

where

$$\mathbf{Q}_{22} := \sum_{k=0}^{\nu-1} (-\mathbf{N}^k)^{\mathrm{T}} \widetilde{\mathbf{M}}_{22} \mathbf{P}_2 \widetilde{\mathbf{M}}_{22} (-\mathbf{N}^k), \tag{3.115}$$

with the improper controllability Gramian $\mathbf{P}_2$ defined as in (3.95) and $\widetilde{\mathbf{M}}_{22}$ as in (3.101).◊

Theorem 3.46:
Consider the C-stable system (3.100) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$ and the corresponding improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{w}_\mathrm{p}}$ as defined in (3.93). The improper-improper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{ii},\mathcal{w}_\mathrm{p}}$ solves the projected Lyapunov equation

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

where $\mathbf{P}_\mathrm{l}$ and $\mathbf{P}_\mathrm{r}$ are defined as in (2.10). ◊

Proof. The proof is similar to the one of Theorem 3.43. □

We can combine the two improper output Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{pi},\mathcal{w}_\mathrm{p}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{ii},\mathcal{w}_\mathrm{p}}$ to obtain an improper Gramian that encodes the observability of the output $\mathbf{y}_\mathrm{i}(t) = \mathbf{z}(t)^{\mathrm{T}} \boldsymbol{\mathcal{M}} \mathbf{z}_\mathrm{i}(t)$ for an arbitrary state $\mathbf{z}(t)$ generated by system (3.100). Since the sum $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{w}_\mathrm{p}} + \boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{w}_\mathrm{p}}$ spans the full controllability space of the state $\mathbf{z}(t)$, the proper observability Gramian corresponding to both, differential and algebraic left states, is given by

$$\boldsymbol{\mathcal{Q}}_{\mathrm{Q,i},\mathcal{w}_\mathrm{p}} = \sum_{k=0}^{\nu-1} \boldsymbol{\mathcal{F}}_{\mathbf{N}}(t)^{\mathrm{T}} \boldsymbol{\mathcal{M}} (\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{w}_\mathrm{p}} + \boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{w}_\mathrm{p}}) \boldsymbol{\mathcal{M}} \boldsymbol{\mathcal{F}}_{\mathbf{N}}(t) = \boldsymbol{\mathcal{Q}}_{\mathrm{pi},\mathcal{w}_\mathrm{p}} + \boldsymbol{\mathcal{Q}}_{\mathrm{ii},\mathcal{w}_\mathrm{p}}.$$

We summarize this paragraph with the following definition.

Definition 3.47:
Consider the C-stable system (3.100) with a regular matrix pencil $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{E}})$, and the corresponding proper and improper controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{w}_\mathrm{p}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{w}_\mathrm{p}}$ as defined in (3.93). The improper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{i},\mathcal{w}_\mathrm{p}}$ corresponding to the output $\mathbf{y}_\mathrm{i}$ is defined as

$$\boldsymbol{\mathcal{Q}}_{\mathrm{Q,i},\mathcal{w}_\mathrm{p}} := \boldsymbol{\mathcal{Q}}_{\mathrm{pi},\mathcal{w}_\mathrm{p}} + \boldsymbol{\mathcal{Q}}_{\mathrm{ii},\mathcal{w}_\mathrm{p}}, \tag{3.116}$$

where the Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{pi},\mathcal{w}_\mathrm{p}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{ii},\mathcal{w}_\mathrm{p}}$ are as defined in (3.112) and (3.44). ◊

Controllability energies As described above, the controllability behavior and hence the controllability energies of the system (3.105) with a quadratic output equation are equal to those in (3.91) with a linear output equation. Hence, we consider the energy measure derived in (3.98) and (3.99) for systems with a linear output equation based on the proper and improper input-to-state mappings defined in (3.92). The energy norms of these mappings are

$$E(\boldsymbol{c}_{\mathrm{p},\mathrm{w}_\mathrm{p}}) = \operatorname{tr}\big(\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathrm{w}_\mathrm{p}}\big)\,, \qquad E(\boldsymbol{c}_{\mathrm{i},\mathrm{w}_\mathrm{p}}) = \operatorname{tr}\big(\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathrm{w}_\mathrm{p}}\big)\,.$$

We observe that states corresponding to the large eigenvalues of the Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathrm{w}_\mathrm{p}}$ span the most dominant proper controllability subspaces. On the other hand, the smallest eigenvalues, including the zero eigenvalues, are negligible to describe the system dynamics. Moreover, it follows from the evaluation of $E(\boldsymbol{c}_{\mathrm{i},\mathrm{w}_\mathrm{p}})$ that zero eigenvalues of the improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathrm{w}_\mathrm{p}}$ are negligible since they do not change the energy of the system.

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

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

We follow the same methodology as above and evaluate the energy norm of the proper observability mapping. However, this mapping depends on continuous and discrete variables. Therefore, we need to define an energy norm that considers both. For a function $\boldsymbol{o} : \mathbb{N} \times [0,\infty) \to \mathbb{R}^{m\times N}$ with $\boldsymbol{o}(k,\cdot) \in L_2\left([0,\infty), \mathbb{R}^{m\times N}\right)$ for all $k \in \mathbb{N}$, we can evaluate the L_2-norm as

$$E(k,\cdot) = \|\boldsymbol{o}(k,\cdot)\|^2_{L_2([0,\infty),\mathbb{R}^{m\times N})}.$$

Also these norm values define a sequence $(E(k,\cdot))_k$, $E(k,\cdot) : \mathbb{N} \to \mathbb{R}$. If it holds that $\sum_0^\infty \|E(k,\cdot)\|_{\mathbf{F}} < \infty$, we can define mixed energy norm as the ℓ_2-norm of $E(k,\cdot)$ that is defined as

$$E(\boldsymbol{c}) := \|E(k,\cdot)\|_{\ell_2(\mathbb{N},\mathbb{R})} = \sum_{k=0}^{\infty} |E(k,\cdot)| = \sum_{k=0}^{\infty} \int_0^\infty \operatorname{tr}\big(\boldsymbol{o}(k,t)\boldsymbol{o}(k,t)^{\mathrm{T}}\big)\,\mathrm{d}t. \tag{3.117}$$

Applying the energy norm from (3.117) to the state-to-output mapping $\boldsymbol{o}_{\mathrm{p},\mathsf{w}_{\mathrm{p}}}$ yields

$$\begin{aligned}
E(\boldsymbol{o}_{\mathrm{p},\mathsf{w}_{\mathrm{p}}}) :=& \sum_{k=0}^{\nu-1} \|\boldsymbol{o}_{\mathrm{p},\mathsf{w}_{\mathrm{p}}}(k,\cdot,\cdot)\|^2_{L_2\left([0,\infty)^2,\mathbb{R}^{2(m+\mathrm{N}_{\mathbf{Z}_0})\times N}\right)} \\
=& \sum_{k=0}^{\nu-1} \int_0^\infty \int_0^\infty \operatorname{tr}\left(\boldsymbol{o}_{\mathrm{p},\mathsf{w}_{\mathrm{p}}}(k,t_1,t_2)^{\mathrm{T}} \boldsymbol{o}_{\mathrm{p},\mathsf{w}_{\mathrm{p}}}(k,t_1,t_2)\right) \mathrm{d}t_1 \mathrm{d}t_2 \\
=& \int_0^\infty \int_0^\infty \operatorname{tr}\left(\boldsymbol{o}_{\mathrm{pp},\mathsf{w}_{\mathrm{p}}}(t_1,t_2)^{\mathrm{T}} \boldsymbol{o}_{\mathrm{pp},\mathsf{w}_{\mathrm{p}}}(t_1,t_2)\right) \mathrm{d}t_1 \mathrm{d}t_2 \\
&\qquad + \sum_{k=0}^{\nu-1} \int_0^\infty \operatorname{tr}\left(\boldsymbol{o}_{\mathrm{ip},\mathsf{w}_{\mathrm{p}}}(k,t_2)^{\mathrm{T}} \boldsymbol{o}_{\mathrm{ip},\mathsf{w}_{\mathrm{p}}}(k,t_2) \mathrm{d}t_2\right) \\
=& \operatorname{tr}\left(\boldsymbol{\mathfrak{Q}}_{\mathrm{pp},\mathsf{w}_{\mathrm{p}}}\right) + \operatorname{tr}\left(\boldsymbol{\mathfrak{Q}}_{\mathrm{ip},\mathsf{w}_{\mathrm{p}}}\right) = \operatorname{tr}\left(\boldsymbol{\mathfrak{Q}}_{\mathrm{p},\mathsf{w}_{\mathrm{p}}}\right)
\end{aligned}$$

with $\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\mathrm{p},\mathsf{w}_{\mathrm{p}}}$ is the proper observability Gramian defined in (3.111). We observe, that the largest eigenvalues of the Gramian $\boldsymbol{\mathfrak{Q}}_{\mathrm{p},\mathsf{w}_{\mathrm{p}}}$ have the highest influence on the observability energy while the influence of the smallest eigenvalues is negligible.

We also apply the energy norm from (3.117) to the improper state-to-output mapping $\boldsymbol{o}_{\mathrm{i},\mathsf{w}_{\mathrm{p}}}$ to obtain

$$\begin{aligned}
E(\boldsymbol{o}_{\mathrm{i},\mathsf{w}_{\mathrm{p}}}) :=& \sum_{\ell=0}^{\nu-1} \sum_{k=0}^{\nu-1} \int_0^\infty \|\boldsymbol{\mathcal{O}}_{\mathrm{i},\mathsf{w}_{\mathrm{p}}}(\ell,k,\cdot)\|^2_{L_2\left([0,\infty),\mathbb{R}^{2(m+\mathrm{N}_{\mathbf{Z}_0})\times N}\right)} \\
=& \sum_{\ell=0}^{\nu-1} \sum_{k=0}^{\nu-1} \int_0^\infty \operatorname{tr}\left(\boldsymbol{o}_{\mathrm{i},\mathsf{w}_{\mathrm{p}}}(\ell,k,t)^{\mathrm{T}} \boldsymbol{o}_{\mathrm{i},\mathsf{w}_{\mathrm{p}}}(\ell,k,t)\right) \mathrm{d}t \\
=& \sum_{k=0}^{\nu-1} \int_0^\infty \operatorname{tr}\left(\boldsymbol{o}_{\mathrm{pi},\mathsf{w}_{\mathrm{p}}}(k,t)^{\mathrm{T}} \boldsymbol{o}_{\mathrm{pi},\mathsf{w}_{\mathrm{p}}}(k,t)\right) \mathrm{d}t \\
&\qquad + \sum_{\ell=0}^{\nu-1} \sum_{k=0}^{\nu-1} \operatorname{tr}\left(\boldsymbol{o}_{\mathrm{ii},\mathsf{w}_{\mathrm{p}}}(\ell,k)^{\mathrm{T}} \boldsymbol{o}_{\mathrm{ii},\mathsf{w}_{\mathrm{p}}}(\ell,k)\right) \\
=& \operatorname{tr}\left(\boldsymbol{\mathfrak{Q}}_{\mathrm{pi},\mathsf{w}_{\mathrm{p}}}\right) + \operatorname{tr}\left(\boldsymbol{\mathfrak{Q}}_{\mathrm{ii},\mathsf{w}_{\mathrm{p}}}\right) = \operatorname{tr}\left(\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\mathrm{i},\mathsf{w}_{\mathrm{p}}}\right),
\end{aligned}$$

where $\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\mathrm{i},\mathsf{w}_{\mathrm{p}}}$ is the improper observability Gramian as defined in (3.116). We observe, that the largest eigenvalues of the Gramian $\boldsymbol{\mathfrak{Q}}_{\mathrm{p},\mathsf{w}_{\mathrm{p}}}$ have the highest influence on the observability energy. However, since the algebraic states encode physical restrictions on the system dynamics, only zero eigenvalues are negligible.

From both energy expressions, we follow that the states corresponding to the largest eigenvalues of the respective Gramians span the most dominant observability subspaces. Consequently, when reducing the respective system (3.100), we truncate the states corresponding to small eigenvalues of $\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\mathrm{p},\mathsf{w}_{\mathrm{p}}}$ and zero eigenvalues of $\boldsymbol{\mathfrak{Q}}_{\mathrm{Q},\mathrm{i},\mathsf{w}_{\mathrm{p}}}$.

In Table 3.7, we summarize the transfer functions, the derived Gramians, and the respective energies that were introduced in this Section.

	System (3.105) – differential component	System (3.105) – algebraic component
Transfer function	$\mathcal{G}_{\mathrm{Q,p},w_\mathrm{p}}(s_1, s_2)$	$\mathcal{G}_{\mathrm{Q,i},w_\mathrm{p}}(s_1, s_2)$
Controllability Gramian	$\mathcal{P}_{\mathrm{p},w_\mathrm{p}}$	$\mathcal{P}_{\mathrm{i},w_\mathrm{p}}$
Observability Gramian	$\mathcal{Q}_{\mathrm{Q,p},w_\mathrm{p}}$	$\mathcal{Q}_{\mathrm{Q,i},w_\mathrm{p}}$
Controllability energies	$E(\boldsymbol{c}_{\mathrm{p},w_\mathrm{p}}) = \mathrm{tr}\big(\mathcal{P}_{\mathrm{p},w_\mathrm{p}}\big)$	$E(\boldsymbol{c}_{\mathrm{i},w_\mathrm{p}}) = \mathrm{tr}\big(\mathcal{P}_{\mathrm{i},w_\mathrm{p}}\big)$
Observability energies	$E(\boldsymbol{o}_{\mathrm{p},w_\mathrm{p}}) = \mathrm{tr}\big(\mathcal{Q}_{\mathrm{Q,p},w_\mathrm{p}}\big)$	$E(\boldsymbol{o}_{\mathrm{i},w_\mathrm{p}}) = \mathrm{tr}\big(\mathcal{Q}_{\mathrm{Q,i},w_\mathrm{p}}\big)$

Table 3.7: Properties of system (3.100) corresponding to its extended-input representation.

3.3 Inhomogeneous second-order ODE systems

In this section, we extend the theory from Section 3.1 to second-order systems with a state equation

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0 \tag{3.118}$$

where the mass, damping, and stiffness matrix are $\mathbf{M}$, $\mathbf{D}$, $\mathbf{K} \in \mathbb{R}^{n\times n}$, respectively, and the input matrix is $\mathbf{B} \in \mathbb{R}^{n\times m}$. The matrices $\mathbf{M}$, $\mathbf{D}$, and $\mathbf{K}$ are naturally symmetric and positive semi-definite. Throughout this work, however, we assume positive definiteness. Also, the state is given as $\mathbf{x}(t) \in \mathbb{R}^n$, the input as $\mathbf{u}(t) \in \mathbb{R}^m$, and initial values as $\mathbf{x}_0$, $\dot{\mathbf{x}}_0 \in \mathbb{R}^n$. We assume that there are matrices $\mathbf{X}_0 \in \mathbb{R}^{n\times n_{\mathbf{x}_0}}$ and $\mathbf{V}_0 \in \mathbb{R}^{n\times n_{\mathbf{v}_0}}$ so that all admissible initial states and velocities can be written as

$$\mathbf{x}(0) = \mathbf{x}_0 = \mathbf{X}_0\chi_0, \qquad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0 = \mathbf{V}_0\nu_0, \tag{3.119}$$

for suitable vectors $\chi_0 \in \mathbb{R}^{n_{\mathbf{x}_0}}$ and $\nu_0 \in \mathbb{R}^{n_{\mathbf{v}_0}}$.

Using the matrices introduced in (2.24), the second-order state equation in (3.118) can be written as first-order equation (3.1) and, hence, the respective system properties derived in Section 3.1 can be used to describe the system dynamics. However, we aim to maintain the second-order structure to derive physically meaningful results. Therefore, in this section, we derive the transfer functions, Gramians, and respective energy expressions for second-order systems. To do so, we apply the Laplace transform to the

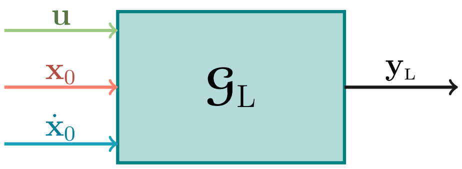


Figure 3.14: Structure of a second-order ODE system with a linear output
.

state equation in (3.118) and obtain

$$\mathbf{X}(s) = \boldsymbol{\Lambda}(s)\left(\mathbf{B}\mathbf{U}(s) + (\mathbf{D} + s\mathbf{M})\mathbf{X}_0\chi_0 + \mathbf{M}\mathbf{V}_0\nu_0\right) \tag{3.120}$$

for $\boldsymbol{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$, where $\mathbf{X}(s)$ and $\mathbf{U}(s)$ are the Laplace transforms of $\mathbf{x}(t)$ and $\mathbf{u}(t)$, respectively. We note that the state $\mathbf{X}(s)$ consists of three components: one arising from the input, one arising from the displacement initial condition, and one arising from the velocity initial condition. This state composition will be used in the following to describe the behavior of the system (3.118).

We study second-order systems with different output structures. First, in Section 3.3.1, we investigate second-order systems with linear output equations, and afterward, in Section 3.3.2, systems with quadratic output equations. For this purpose, we derive tailored second-order systems Gramians and the resulting system energies that describe the respective system behavior. Therefore, we modify the concepts introduced in Section 3.1 for first-order systems.

3.3.1 Inhomogeneous second-order ODE systems with a linear output

We first consider second-order systems with a linear output equation of the form

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) &= \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0, \\ \mathbf{y}_{\mathrm{L}}(t) &= \mathbf{C}_1\mathbf{x}(t) + \mathbf{C}_2\dot{\mathbf{x}}(t), \end{aligned} \tag{3.121}$$

with a state equation as introduced in (3.118), and an output equation containing the output matrices $\mathbf{C}_1, \mathbf{C}_2 \in \mathbb{R}^{p\times n}$ and the output $\mathbf{y}_{\mathrm{L}}(t) \in \mathbb{R}^p$. In this section, we assume that $\mathbf{C}_2 = 0$. Otherwise, if $\mathbf{C}_2 \neq 0$, the first-order representation (3.5) with matrices (2.24) is used so that we apply the theory from Section 3.1.1.

Figure 3.14 describes the structure of these systems, where we note that the system is affected by the input, the displacement initial condition, and the velocity initial condition. In the following, we analyze the system dynamics and consider, in particular, the initial conditions. For this purpose, we extend the superposition ideas from [15] to the class of second-order systems. Given the superposition principle, we show that the original system is decomposable into three subsystems. The first subsystem considers the mapping between the input $\mathbf{u}(t)$ and the output, where the initial conditions are zero. The second and the third subsystems correspond to the outputs resulting from the position initial condition $\mathbf{x}_0$ and the velocity initial condition $\dot{\mathbf{x}}_0$. Based on the representation of these subsystems in the frequency domain presented in this thesis, we design tailored controllability and observability Gramians for the input and initial conditions. They are valuable tools to describe the controllability spaces corresponding to the initial conditions since they allow the preservation of physically meaningful second-order structures. Moreover, these Gramians can be concatenated so that one controllability Gramian and one observability Gramian encode the overall system behavior.

We propose two methods to study the inhomogeneous systems. The first one analyzes each subsystem independently using the respective Gramians, as introduced in Section 3.3.1.1. The second proposed method, introduced in Section 3.3.1.2, analyzes the system as a whole using the extended-input method and the associated Gramians.

3.3.1.1 Multi-system approach for inhomogeneous second-order ODE systems with a linear output

In this section, we apply the superposition principles to the state equation in (3.120) to deal with inhomogeneous initial conditions in second-order systems. Therefore, we derive three subsystems, one corresponding to the input, one to the displacement initial condition, and one to the velocity initial condition. These subsystems are then analyzed independently to describe the respective controllability and observability behavior.

Transfer function We apply the Laplace transform to the output equation in (3.121) and insert the state $\mathbf{X}(s)$ as defined in (3.120) to obtain the output in the frequency domain, that is

$$\mathbf{Y}_{\mathrm{L}}(s) = \mathbf{C}_1\mathbf{\Lambda}(s)\mathbf{B}\mathbf{U}(s) + \mathbf{C}_1\mathbf{\Lambda}(s)(\mathbf{D} + s\mathbf{M})\mathbf{X}_0\chi_0 + \mathbf{C}_1\mathbf{\Lambda}(s)\mathbf{M}\mathbf{V}_0\nu_0. \tag{3.122}$$

We observe that the output is a superposition of the input-to-output mapping, the position initial condition-to-output mapping, and the velocity initial condition-to-output mapping. Corresponding to the three output components, we define the transfer functions describing the input- and initial condition-to-output mappings, which follow directly from the output decomposition above.

Definition 3.48:
Consider the asymptotically stable second-order system in (3.121) with initial conditions (3.119) and define $\mathbf{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$. The three *transfer functions* describing the system behavior are defined as

$$\mathcal{G}_{\mathrm{L},\mathbf{B}}(s) := \mathbf{C}_1\mathbf{\Lambda}(s)\mathbf{B}, \qquad \mathcal{G}_{\mathrm{L},\mathbf{X}_0}(s) := \mathbf{C}_1\mathbf{\Lambda}(s)(\mathbf{D} + s\mathbf{M})\mathbf{X}_0, \qquad \mathcal{G}_{\mathrm{L},\mathbf{V}_0}(s) := \mathbf{C}_1\mathbf{\Lambda}(s)\mathbf{M}\mathbf{V}_0. \tag{3.123}$$

◊

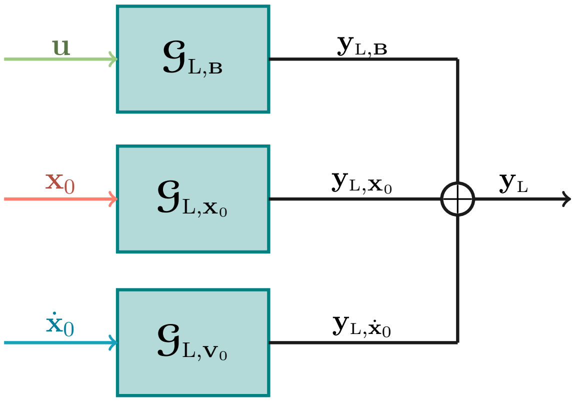

$\mathbf{u}$
$\mathcal{G}_{\mathrm{L},\mathbf{B}}$
$\mathbf{y}_{\mathrm{L},\mathbf{B}}$
$\mathbf{x}_0$
$\mathcal{G}_{\mathrm{L},\mathbf{x}_0}$
$\mathbf{y}_{\mathrm{L},\mathbf{x}_0}$
$\mathbf{y}_{\mathrm{L}}$
$\dot{\mathbf{x}}_0$
$\mathcal{G}_{\mathrm{L},\mathbf{v}_0}$
$\mathbf{y}_{\mathrm{L},\dot{\mathbf{x}}_0}$

Figure 3.15: Structure of three separated second-order systems with a linear output.

We generate three subsystems that are treated individually in the following as depicted in Figure 3.15. The first transfer function $\mathcal{G}_{\mathrm{L},\mathbf{B}}(s)$ with the output component $\mathbf{y}_{\mathrm{L},\mathbf{B}}(t)$ corresponds to the homogeneous system representation

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) &= \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = 0, \quad \dot{\mathbf{x}}(0) = 0, \\ \mathbf{y}_{\mathrm{L},\mathbf{B}}(t) &= \mathbf{C}_1\mathbf{x}(t). \end{aligned} \tag{3.124}$$

The second transfer function $\mathcal{G}_{\mathrm{L},\mathbf{x}_0}(t)$ has a system representation with an inhomogeneous displacement initial condition, that is

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) &= 0, \qquad \mathbf{x}(0) = \mathbf{X}_0\chi_0, \quad \dot{\mathbf{x}}(0) = 0, \\ \mathbf{y}_{\mathrm{L},\mathbf{x}_0}(t) &= \mathbf{C}_1\mathbf{x}(t). \end{aligned} \tag{3.125}$$

Finally, the transfer function $\mathcal{G}_{\mathrm{L},\mathbf{v}_0}(t)$ corresponds to the mapping between the velocity initial condition and the output and has the system realization

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) &= 0, \qquad \mathbf{x}(0) = 0, \quad \dot{\mathbf{x}}(0) = \mathbf{V}_0\nu_0, \\ \mathbf{y}_{\mathrm{L},\mathbf{v}_0}(t) &= \mathbf{C}_1\mathbf{x}(t). \end{aligned} \tag{3.126}$$

The three systems are treated individually to describe the overall system dynamics. For that, we derive the corresponding controllability and observability Gramians encoding the behavior of the subsystems.

Controllability Gramians To derive the controllability Gramians that encode the controllability behavior for the subsystems in (3.124), (3.125), and (3.126), we consider the respective input- and initial condition-to-state mappings separately.

First, we investigate system (3.124). From the transfer function $\mathcal{G}_{\mathrm{L},\mathbf{B}}(s)$ in (3.123), we extract the input-to-state mapping

$$\mathcal{C}_{\mathbf{B}}(s) := \mathbf{\Lambda}(s)\mathbf{B}, \tag{3.127}$$

where $\mathbf{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$. Using this mapping, we define a matrix $\mathbf{P}_{\mathbf{B}} := \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{C}_{\mathbf{B}}(\mathrm{i}\omega)\mathcal{C}_{\mathbf{B}}(-\mathrm{i}\omega)^{\mathrm{T}}\mathrm{d}\omega$ spanning the respective controllability space.

Definition 3.49:
Consider the asymptotically stable second-order system (3.124). Define $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$, then the corresponding *second-order controllability Gramian* is defined as

$$\mathbf{P}_{\mathbf{B}} := \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{\Lambda}(\mathrm{i}\omega)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{\Lambda}(-\mathrm{i}\omega)^{\mathrm{T}}\mathrm{d}\omega. \tag{3.128}$$

◊

As we have seen in the previous sections, first-order Gramians are computed by solving Lyapunov equations. However, for second-order systems, the computation of the respective Gramians is not straightforward. We define the first-order matrices as in (2.24) to derive a connection between second-order and first-order Gramians. The following theorem from [44, 112] describes how to compute the second-order controllability Gramian $\mathbf{P}_{\mathbf{B}}$ as a component of a first-order matrix.

Theorem 3.50:
Consider the asymptotically stable system (3.124) with the second-order controllability Gramian $\mathbf{P}_{\mathbf{B}}$ defined in (3.128). Then the Gramian $\mathbf{P}_{\mathbf{B}}$ is the upper-left block $\mathbf{P}_{\mathbf{B},1}$ of the first-order controllability Gramian

$$\mathcal{P}_{\mathbf{B}} = \begin{bmatrix} \mathbf{P}_{\mathbf{B},1} & \mathbf{P}_{\mathbf{B},2} \\ \mathbf{P}_{\mathbf{B},2}^{\mathrm{T}} & \mathbf{P}_{\mathbf{B},3} \end{bmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty} (\mathrm{i}\omega\mathcal{E} - \mathcal{A})^{-1}\begin{bmatrix} 0 \\ \mathbf{B} \end{bmatrix}\begin{bmatrix} 0 & \mathbf{B}^{\mathrm{T}} \end{bmatrix}(-\mathrm{i}\omega\mathcal{E} - \mathcal{A})^{-\mathrm{T}}\mathrm{d}\omega \tag{3.129}$$

with first-order matrices $\mathcal{E}$ and $\mathcal{A}$ as defined in (2.24). ◊

Proof. We first apply the Schur complement to $(s\mathcal{E} - \mathcal{A})^{-1}$ and obtain

$$(s\mathcal{E} - \mathcal{A})^{-1} = \begin{bmatrix} s\mathbf{I} & -\mathbf{I} \\ \mathbf{K} & \mathbf{D} + s\mathbf{M} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{\Lambda}(s)(s\mathbf{M}+\mathbf{D}) & \mathbf{\Lambda}(s) \\ -\mathbf{\Lambda}(s)\mathbf{K} & s\mathbf{\Lambda}(s) \end{bmatrix} \tag{3.130}$$

for $\mathbf{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$. Applying this formula provides that its upper-right block is equal to $\mathbf{\Lambda}(\mathrm{i}\omega)$, and hence it holds that

$$\mathbf{P}_{\mathbf{B},1} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{\Lambda}(\mathrm{i}\omega)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega = \mathbf{P}_{\mathbf{B}}.$$

□

From this theorem, it follows that the controllability Gramian of the corresponding first-order realization is determined to compute the second-order controllability Gramian, which is done by solving a Lyapunov equation as described in Section 2.3.

Remark 3.51:
Note, that the mapping $\mathcal{C}_{\mathrm{B}}(s)$ from (3.127) is the Laplace transform of the mapping $\begin{bmatrix}\mathbf{I} & 0\end{bmatrix} \boldsymbol{c}_{\mathcal{B}}(t)$ with $\boldsymbol{c}_{\mathcal{B}}(t)$ as defined in (3.10) for first-order matrices from (2.24). Moreover, the first-order Gramian $\mathcal{P}_{\mathrm{B}}$ from (3.129) is equal to the Gramian defined in (3.11). Hence, we can also define the second-order Gramian in the time domain as

$$\mathbf{P}_{\mathrm{B}} = \int_0^\infty \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{B}\mathcal{B}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\begin{bmatrix}\mathbf{I}\\ 0\end{bmatrix}\mathrm{d}t.$$ ◊

Similarly, we investigate the system (3.125), where we extract from the transfer function $\mathcal{G}_{\mathrm{L},\mathbf{x}_0}$ in (3.123) the position initial condition-to-state mapping

$$\mathcal{C}_{\mathbf{x}_0}(s) := \mathbf{\Lambda}(s)(s\mathbf{M}+\mathbf{D})\mathbf{X}_0, \tag{3.131}$$

that we use to define a matrix $\mathbf{P}_{\mathbf{x}_0} := \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathcal{C}_{\mathbf{x}_0}(\mathrm{i}\omega)\mathcal{C}_{\mathbf{x}_0}(-\mathrm{i}\omega)^{\mathrm{T}}\mathrm{d}\omega$ that spans the controllability space corresponding to the position initial condition.

Definition 3.52:
Consider the asymptotically stable second-order system (3.125). Define $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$, then the corresponding *second-order controllability Gramian* is defined as

$$\mathbf{P}_{\mathbf{x}_0} := \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{\Lambda}(\mathrm{i}\omega)(\mathrm{i}\omega\mathbf{M}+\mathbf{D})\mathbf{X}_0\mathbf{X}_0^{\mathrm{T}}(-\mathrm{i}\omega\mathbf{M}+\mathbf{D})^{\mathrm{T}}\mathbf{\Lambda}(-\mathrm{i}\omega)^{\mathrm{T}}\mathrm{d}\omega. \tag{3.132}$$ ◊

The following theorem shows that the Gramian $\mathbf{P}_{\mathbf{x}_0}$ is computed by determining the respective first-order controllability Gramian.

Theorem 3.53:
Consider the asymptotically stable second-order system (3.125) with the second-order controllability Gramian $\mathbf{P}_{\mathbf{x}_0}$ defined in (3.132). Then the Gramian $\mathbf{P}_{\mathbf{x}_0}$ is the upper-left block $\mathbf{P}_{\mathbf{x}_0,1}$ of the first-order Gramian

$$\mathcal{P}_{\mathbf{x}_0} = \begin{bmatrix}\mathbf{P}_{\mathbf{x}_0,1} & \mathbf{P}_{\mathbf{x}_0,2}\\ \mathbf{P}_{\mathbf{x}_0,2}^{\mathrm{T}} & \mathbf{P}_{\mathbf{x}_0,3}\end{bmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-1}\begin{bmatrix}\mathbf{X}_0\\ 0\end{bmatrix}\begin{bmatrix}\mathbf{X}_0^{\mathrm{T}} & 0\end{bmatrix}(-\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathrm{d}\omega, \tag{3.133}$$

with first-order matrices $\mathcal{E}$ and $\mathcal{A}$ as defined in (2.24). ◊

Proof. Applying the Schur complement to $(\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-1}$ as shown in (3.130) provides that its upper-left block of $(\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-1}$ is $\mathbf{\Lambda}(\mathrm{i}\omega)(\mathrm{i}\omega\mathbf{M}+\mathbf{D})$ for $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$, and hence, it holds that

$$\mathbf{P}_{\mathbf{x}_0,1} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{\Lambda}(\mathrm{i}\omega)(\mathrm{i}\omega\mathbf{M}+\mathbf{D})\mathbf{X}_0\mathbf{X}_0^{\mathrm{T}}(-\mathrm{i}\omega\mathbf{M}+\mathbf{D})^{\mathrm{T}}\mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega = \mathbf{P}_{\mathbf{x}_0}.$$ □

Theorem 3.53 shows that the second-order controllability Gramian $\mathbf{P}_{\mathbf{X}_0}$ of a system (3.125) is the upper-left block $\mathbf{P}_1$ of the controllability Gramian $\mathcal{P}_{\mathbf{X}_0}$ of the first-order system (2.1) with $\mathcal{B} := \left[\begin{smallmatrix}\mathbf{X}_0\\ 0\end{smallmatrix}\right]$.

Remark 3.54:
Note, that the mapping $\mathcal{C}_{\mathbf{X}_0}(s)$ from (3.131) is the Laplace transform of the mapping $\begin{bmatrix}\mathbf{I} & 0\end{bmatrix} \boldsymbol{c}_{\mathbf{Z}_0}(t)$ with $\boldsymbol{c}_{\mathbf{Z}_0}(t)$ as defined in (3.13) for first-order matrices from (2.24) and $\mathbf{Z}_0 = \left[\begin{smallmatrix}\mathbf{X}_0\\ 0\end{smallmatrix}\right]$. Moreover, the first-order Gramian $\mathcal{P}_{\mathbf{X}_0}$ from (3.133) is equal to the Gramian defined in (3.14) with that matrix $\mathbf{Z}_0$. Hence, we can also define the second-order Gramian in the time domain as

$$\mathbf{P}_{\mathbf{X}_0} = \int_0^\infty \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\begin{bmatrix}\mathbf{X}_0\mathbf{X}_0^{\mathrm{T}} & 0\\ 0 & 0\end{bmatrix}\mathcal{E}^{-\mathrm{T}}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\begin{bmatrix}\mathbf{I}\\ 0\end{bmatrix}\mathrm{d}t.$$

◊

Now we consider the remaining system (3.126) with the transfer function $\mathcal{G}_{\mathrm{L},\mathbf{V}_0}$ as defined in (3.123). From this output, we extract the velocity initial condition-to-state mapping

$$\mathcal{C}_{\mathbf{V}_0}(s) := \mathbf{\Lambda}(s)\mathbf{M}\mathbf{V}_0, \tag{3.134}$$

which is used to define a matrix $\mathbf{P}_{\mathbf{V}_0} := \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathcal{C}_{\mathbf{V}_0}(\mathrm{i}\omega)\mathcal{C}_{\mathbf{V}_0}(-\mathrm{i}\omega)^{\mathrm{T}}\mathrm{d}\omega$ that spans the controllability space corresponding to the velocity initial condition.

Definition 3.55:
Consider the asymptotically stable second-order system (3.126). Define $\mathbf{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$, then the corresponding *second-order controllability Gramian* is defined as

$$\mathbf{P}_{\mathbf{V}_0} := \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{\Lambda}(\mathrm{i}\omega)\mathbf{M}\mathbf{V}_0\mathbf{V}_0^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{\Lambda}(-\mathrm{i}\omega)^{\mathrm{T}}\mathrm{d}\omega. \tag{3.135}$$

◊

The Gramian $\mathbf{P}_{\mathbf{V}_0}$ corresponding to the systems (3.126) is of the same structure as the Gramian $\mathbf{P}_{\mathbf{B}}$, and hence, can be computed similarly.

Theorem 3.56:
Consider the asymptotically stable second-order system (3.126) with the second-order position controllability Gramian $\mathbf{P}_{\mathbf{V}_0}$ defined in (3.135). Then the Gramian $\mathbf{P}_{\mathbf{V}_0}$ is the upper-left block $\mathbf{P}_{\mathbf{V}_0,1}$ of the first-order Gramian

$$\mathcal{P}_{\mathbf{V}_0} = \begin{bmatrix}\mathbf{P}_{\mathbf{V}_0,1} & \mathbf{P}_{\mathbf{V}_0,2}\\ \mathbf{P}_{\mathbf{V}_0,2}^{\mathrm{T}} & \mathbf{P}_{\mathbf{V}_0,3}\end{bmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\mathrm{i}\omega\mathcal{E} - \mathcal{A})^{-1}\begin{bmatrix}0\\ \mathbf{M}\mathbf{V}_0\end{bmatrix}\begin{bmatrix}0 & \mathbf{V}_0^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\end{bmatrix}(-\mathrm{i}\omega\mathcal{E} - \mathcal{A})^{-\mathrm{T}}\mathrm{d}\omega \tag{3.136}$$

with first-order matrices $\mathcal{E}$ and $\mathcal{A}$ as defined in (2.24). ◊

Remark 3.57:
Note, that the mapping $\boldsymbol{\mathfrak{C}}_{\mathrm{V}_0}(s)$ from (3.134) is the Laplace transform of the mapping $\begin{bmatrix}\mathbf{I} & 0\end{bmatrix} \boldsymbol{c}_{\mathbf{Z}_0}(t)$ with $\boldsymbol{c}_{\mathbf{Z}_0}(t)$ as defined in (3.13) for first-order matrices from (2.24) and $\mathbf{Z}_0 = \begin{bmatrix} 0 \\ \mathbf{M}\mathbf{V}_0 \end{bmatrix}$. Moreover, the first-order Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{V}_0}$ from (3.136) is equal to the Gramian defined in (3.14) with that matrix $\mathbf{Z}_0$. Hence, we can also define the second-order Gramian in the time domain as

$$\mathbf{P}_{\mathrm{V}_0} = \frac{1}{2\pi} \int_0^\infty \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t} \boldsymbol{\mathcal{E}}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{M}\mathbf{V}_0\mathbf{V}_0^{\mathrm{T}}\mathbf{M}^{\mathrm{T}} \end{bmatrix} \boldsymbol{\mathcal{E}}^{-\mathrm{T}} e^{(\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}})^{\mathrm{T}}t} \begin{bmatrix} \mathbf{I} \\ 0 \end{bmatrix} \mathrm{d}t.$$ ◊

Observability Gramians In this paragraph, we aim to derive the observability Gramians that encode the observability behavior of the three subsystems. Therefore, we define the respective state-to-output mappings and the resulting Gramians from the transfer functions in (3.123). The observability behavior of the three subsystems (3.124), (3.125), and (3.126) is encoded by the state-to-output mapping

$$\boldsymbol{\mathfrak{O}}_{\mathrm{L}}(s) := \mathbf{C}_1 \boldsymbol{\Lambda}(s). \tag{3.137}$$

We integrate over the frequency domain to define a matrix $\mathbf{Q}_{\mathrm{L}} := \frac{1}{2\pi}\int_{-\infty}^{\infty} \boldsymbol{\mathfrak{O}}_{\mathrm{L}}(-\mathrm{i}\omega)^{\mathrm{T}} \boldsymbol{\mathfrak{O}}_{\mathrm{L}}(\mathrm{i}\omega)\mathrm{d}\omega$ that includes all observable states, which leads to the following definition.

Definition 3.58:
Consider the asymptotically stable second-order systems (3.124), (3.125), and (3.126). Also, define $\boldsymbol{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$, then the corresponding *second-order observability Gramian* is defined as

$$\mathbf{Q}_{\mathrm{L}} := \frac{1}{2\pi} \int_{-\infty}^{\infty} \boldsymbol{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}} \mathbf{C}_1^{\mathrm{T}} \mathbf{C}_1 \boldsymbol{\Lambda}(\mathrm{i}\omega) \mathrm{d}\omega. \tag{3.138}$$

◊

The Gramian $\mathbf{Q}_{\mathrm{L}}$ can be computed as a component of a first-order Gramian, as shown in the following theorem.

Theorem 3.59:
Consider the asymptotically stable second-order systems (3.124), (3.125), and (3.126) with the second-order observability Gramian $\mathbf{Q}_{\mathrm{L}}$ as defined in (3.138). Then this Gramian is equal to the lower-right block $\mathbf{Q}_3$ of the first-order Gramian $\boldsymbol{\mathfrak{Q}}_{\mathrm{L}}$ that is

$$\boldsymbol{\mathfrak{Q}}_{\mathrm{L}} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{Q}_2^{\mathrm{T}} & \mathbf{Q}_3 \end{bmatrix} = \frac{1}{2\pi} \int_{-\infty}^{\infty} (\boldsymbol{\mathcal{A}} + \mathrm{i}\omega\boldsymbol{\mathcal{E}})^{-\mathrm{T}} \boldsymbol{\mathfrak{C}}^{\mathrm{T}} \boldsymbol{\mathfrak{C}} (\boldsymbol{\mathcal{A}} - \mathrm{i}\omega\boldsymbol{\mathcal{E}})^{-1} \mathrm{d}\omega \tag{3.139}$$

with first-order matrices $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{A}}$, and $\boldsymbol{\mathfrak{C}}$ as defined in (2.24) with $\mathbf{C}_2 = 0$. ◊

Proof. We apply the Schur complement for $(\mathrm{i}\omega\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}$ as shown in (3.130) to obtain

$$\boldsymbol{\mathfrak{C}}(\mathrm{i}\omega\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1} = \begin{bmatrix}\mathbf{C}_1\boldsymbol{\Lambda}(\mathrm{i}\omega)(\mathrm{i}\omega\mathbf{M} + \mathbf{D}) & \mathbf{C}_1\boldsymbol{\Lambda}(\mathrm{i}\omega)\end{bmatrix}.$$

Considering the right block of this matrix provides that the lower-right block $\boldsymbol{\mathfrak{Q}}_3$ of the first-order Gramian $\boldsymbol{\mathfrak{Q}}_{\mathrm{L}}$ is equal to the second-order Gramian $\mathbf{Q}_{\mathrm{L}}$, which proofs the statement. □

Remark 3.60:
Note, that the mapping $\boldsymbol{\mathfrak{O}}_{\mathrm{L}}(s)$ from (3.137) is the Laplace transform of the mapping $\boldsymbol{o}_{\mathrm{L}}(t)\left[\begin{smallmatrix}0\\ \mathbf{I}\end{smallmatrix}\right]$ with $\boldsymbol{o}_{\mathrm{L}}(t)$ as defined in (3.16) with first-order matrices from (2.24) and $\mathbf{C}_2 = 0$. Moreover, the first-order Gramian $\boldsymbol{\mathfrak{Q}}_{\mathrm{L}}$ from (3.139) is equal to the Gramian defined in (3.17). Hence, we can also define the second-order Gramian in the time domain as

$$\mathbf{Q}_{\mathrm{L}} = \int_0^\infty \begin{bmatrix}0 & \mathbf{I}\end{bmatrix} \boldsymbol{\mathcal{E}}^{-\mathrm{T}} e^{(\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}})^{\mathrm{T}}t} \begin{bmatrix}\mathbf{C}_1^{\mathrm{T}}\mathbf{C}_1 & 0\\ 0 & 0\end{bmatrix} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t} \boldsymbol{\mathcal{E}}^{-1} \begin{bmatrix}0\\ \mathbf{I}\end{bmatrix} \mathrm{d}t. \qquad \diamond$$

Controllability energy We aim to derive the controllability energies of the three subsystems to identify the respective important controllability subspaces. To do so, we consider the three subsystems (3.124), (3.125), and (3.126) separately.

First, we derive an energy measure corresponding to subsystem (3.124) by evaluating the energy norm of the input-to-state mapping $\boldsymbol{\mathfrak{C}}_{\mathbf{B}}$ from (3.127). Therefore, we consider the displacement component of the respective first-order input-to-state mapping $\boldsymbol{c}_{\boldsymbol{\mathcal{B}}}(t)$ (3.10) in the time domain

$$\boldsymbol{c}_{\mathbf{B}}(t) = \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} \boldsymbol{c}_{\boldsymbol{\mathcal{B}}}(t) = \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}$$

as the Laplace transform of $\boldsymbol{c}_{\mathbf{B}}(t)$ is equal to the mapping $\boldsymbol{\mathfrak{C}}_{\mathbf{B}}$ from (3.127), for the first-order matrices $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{A}}$, and $\boldsymbol{\mathcal{B}}$ as defined in (2.24). We determine the energy norm from (3.19) of the mapping $\boldsymbol{c}_{\mathbf{B}}$ to describe the controllability energy that is

$$\begin{aligned}E(\boldsymbol{c}_{\mathbf{B}}) = \|\boldsymbol{c}_{\mathbf{B}}\|^2_{L_2([0,\infty),\mathbb{R}^{n\times m})} &= \int_0^\infty \mathrm{tr}\left(\begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t}\begin{bmatrix}\mathbf{I}\\ 0\end{bmatrix}\right)\mathrm{d}t\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{tr}\big(\boldsymbol{\Lambda}(\mathrm{i}\omega)\mathbf{B}\mathbf{B}^{\mathrm{T}}\boldsymbol{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\big)\,\mathrm{d}\omega\\ &= \mathrm{tr}(\mathbf{P}_{\mathbf{B}}),\end{aligned} \tag{3.140}$$

see Remark 3.51. We observe that the energy norm of the mapping $\boldsymbol{c}_{\mathbf{B}}$ is equal to the trace of the respective second-order controllability Gramian $\mathbf{P}_{\mathbf{B}}$. Since this Gramian is symmetric, its trace is equal to the sum of its eigenvalues, i.e., $\mathrm{tr}(\mathbf{P}_{\mathbf{B}}) = \sigma_1 + \cdots +$

σ_n. Hence, large eigenvalues have a significant effect on the energy expression while smaller eigenvalues are negligible. It follows that the states corresponding to the large eigenvalues span the most dominant controllability subspaces, which is used later in this work when we reduce these systems.

Analogously, we derive the initial condition-to-output mappings in the time-domain corresponding to the remaining two subsystems (3.125) and (3.126) that are

$$\boldsymbol{c}_{\mathbf{X}_0}(t) = \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t} \begin{bmatrix}\mathbf{X}_0\\ 0\end{bmatrix} \qquad \text{and} \qquad \boldsymbol{c}_{\mathbf{V}_0}(t) = \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t} \begin{bmatrix}0\\ \mathbf{V}_0\end{bmatrix},$$

respectively, see Remark 3.54 and Remark 3.57. It holds that the initial condition-to-state mappings $\boldsymbol{\mathcal{C}}_{\mathbf{V}_0}(s)$ and $\boldsymbol{\mathcal{C}}_{\mathbf{V}_0}(s)$ are the Laplace transforms of $\boldsymbol{c}_{\mathbf{X}_0}(t)$ and $\boldsymbol{c}_{\mathbf{V}_0}(t)$. Hence, to derive the energies that encode the controllability behavior of the two subsystems (3.125) and (3.126), we define the respective energy norms corresponding to these initial condition-to-state mappings according to (3.19) that are

$$E(\boldsymbol{c}_{\mathbf{X}_0}) = \operatorname{tr}(\mathbf{P}_{\mathbf{X}_0}), \qquad E(\boldsymbol{c}_{\mathbf{V}_0}) = \operatorname{tr}(\mathbf{P}_{\mathbf{V}_0}). \tag{3.141}$$

Since the traces of $\mathbf{P}_{\mathbf{X}_0}$ and $\mathbf{P}_{\mathbf{V}_0}$ contain the eigenvalues of the corresponding second-order Gramians, it follows that states corresponding to large eigenvalues span the dominant controllability subspaces of the respective systems and states corresponding to small eigenvalues are negligible as they only have little effect on the system dynamics.

Observability energies In this paragraph, we evaluate the observability energies of the second-order subsystems (3.124), (3.125), and (3.126) encoded by the second-order observability Gramian $\mathbf{Q}_{\mathrm{L}}$. Therefore, we consider the energy norm of the respective state-to-output mapping in the time domain, which is defined using the first-order matrices $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{A}}$, and $\boldsymbol{\mathcal{C}}$ from (1.7) with $\mathbf{C}_2 = 0$, as

$$\boldsymbol{o}_{\mathrm{L}}(t) = \boldsymbol{\mathcal{C}} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\begin{bmatrix}0\\ \mathbf{I}\end{bmatrix}, \tag{3.142}$$

see Remark 3.60. Note that applying the Laplace transform to this mapping yields the mapping $\boldsymbol{\mathcal{O}}_{\mathrm{L}}$ in the frequency domain as defined in (3.137). We apply the energy norm from (3.19) that results in the energy expressions

$$\begin{aligned} E(\boldsymbol{o}_{\mathrm{L}}) = \|E(\boldsymbol{o}_{\mathrm{L}})\|^2_{L_2([0,\infty),\mathbb{R}^{p\times n})} &= \int_0^\infty \operatorname{tr}\left(\begin{bmatrix}0 & \mathbf{I}\end{bmatrix}\boldsymbol{\mathcal{E}}^{-\mathrm{T}} e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\boldsymbol{\mathcal{C}} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\begin{bmatrix}0\\ \mathbf{I}\end{bmatrix}\right)\mathrm{d}t\\ &= \operatorname{tr}(\mathbf{Q}_{\mathrm{L}}). \end{aligned} \tag{3.143}$$

The trace of the Gramian $\mathbf{Q}_{\mathrm{L}}$ coincides with the sum of its eigenvalues. Hence, the states corresponding to large eigenvalues of the Gramian $\mathbf{Q}_{\mathrm{L}}$ encode the most dominant observability subspaces. On the other hand, states corresponding to small eigenvalues are negligible as they have negligible effects on the system dynamics.

In this section, we have derived three subsystems encoding the behavior of the original system (3.121). For these subsystems, we introduced transfer functions, tailored Gramians, and energies, which are summarized in Table 3.8.

	System (3.124)	System (3.125)	System (3.126)
Transfer function	$\mathcal{G}_{\mathrm{L},\mathbf{B}}$	$\mathcal{G}_{\mathrm{L},\mathbf{x}_0}$	$\mathcal{G}_{\mathrm{L},\mathbf{v}_0}$
Controllability Gramian	$\mathbf{P}_{\mathbf{B}}$	$\mathbf{P}_{\mathbf{x}_0}$	$\mathbf{P}_{\mathbf{v}_0}$
Observability Gramian	$\mathbf{Q}_{\mathrm{L}}$	$\mathbf{Q}_{\mathrm{L}}$	$\mathbf{Q}_{\mathrm{L}}$
Controllability energies	$E(\boldsymbol{c}_{\mathbf{B}}) = \mathrm{tr}(\mathbf{P}_{\mathbf{B}})$	$E(\boldsymbol{c}_{\mathbf{x}_0}) = \mathrm{tr}(\mathbf{P}_{\mathbf{x}_0})$	$E(\boldsymbol{c}_{\mathbf{v}_0}) = \mathrm{tr}(\mathbf{P}_{\mathbf{v}_0})$
Observability energies	$E(\boldsymbol{o}_{\mathrm{L}}) = \mathrm{tr}(\mathbf{Q}_{\mathrm{L}})$	$E(\boldsymbol{o}_{\mathrm{L}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{L}})$	$E(\boldsymbol{o}_{\mathrm{L}}) = \mathrm{tr}(\mathcal{Q}_{\mathrm{L}})$

Table 3.8: Properties of system (3.121) corresponding to its multi-system representation.

3.3.1.2 Extended-input approach for inhomogeneous second-order ODE systems with a linear output

In this paragraph, we apply the extended-input approach to treat the inhomogeneous initial conditions. Therefore, we reformulate the state $\mathbf{X}(s)$ from (3.120), using a modified input matrix that includes the input and initial condition spaces as described in the following theorem.

Theorem 3.61:
Consider the asymptotically stable second-order system (3.121) with initial conditions as defined in (3.119). Define the input matrix and the modified input in the frequency domain as

$$\mathcal{W}_{\mathrm{so}} := \begin{bmatrix} 0 & \mathbf{X}_0 & 0 \\ \mathbf{B} & 0 & \mathbf{M}\mathbf{V}_0 \end{bmatrix} \qquad \text{and} \qquad \widetilde{\mathbf{U}}_{\mathrm{so}}(s) := \begin{bmatrix} \mathbf{U}(s) \\ \chi_0 \\ \nu_0 \end{bmatrix}, \tag{3.144}$$

respectively. Then the state $\mathbf{X}(s)$ from (3.120) is equal to

$$\mathbf{X}(s) = \mathbf{\Lambda}(s) \begin{bmatrix} (\mathbf{D} + s\mathbf{M}) & \mathbf{I} \end{bmatrix} \mathcal{W}_{\mathrm{so}} \widetilde{\mathbf{U}}(s), \tag{3.145}$$

with $\mathbf{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$. ◊

Proof. By inserting the definition of $\mathcal{W}_{\text{so}}$ and $\widetilde{\mathbf{U}}(s)$ from (3.144), we obtain

$$\begin{aligned}
\mathbf{X}(s) &= \mathbf{\Lambda}(s)(\mathbf{B}\mathbf{U}(s) + (\mathbf{D} + s\mathbf{M})\mathbf{X}_0\chi_0 + \mathbf{M}\mathbf{V}_0\nu_0) \\
&= \mathbf{\Lambda}(s)\begin{bmatrix}(s\mathbf{M} + \mathbf{D}) & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{X}_0\chi_0 \\ \mathbf{B}\mathbf{U}(s) + \mathbf{M}\mathbf{V}_0\nu_0\end{bmatrix} \\
&= \mathbf{\Lambda}(s)\begin{bmatrix}(s\mathbf{M} + \mathbf{D}) & \mathbf{I}\end{bmatrix}\begin{bmatrix}0 & \mathbf{X}_0 & 0 \\ \mathbf{B} & 0 & \mathbf{M}\mathbf{V}_0\end{bmatrix}\begin{bmatrix}\mathbf{U}(s) \\ \chi_0 \\ \nu_0\end{bmatrix} \\
&= \mathbf{\Lambda}(s)\begin{bmatrix}(s\mathbf{M} + \mathbf{D}) & \mathbf{I}\end{bmatrix}\mathcal{W}_{\text{so}}\widetilde{\mathbf{U}}(s)
\end{aligned}$$ □

Transfer function Since we aim to derive a surrogate model with the same input- and initial condition-to-output mapping as the original system (3.121), we first derive the respective transfer function. For that, we apply the Laplace transform to the output equation in (3.121) and insert the state $\mathbf{X}(s)$ from (3.145) to obtain the output

$$\mathbf{Y}_{\text{L}}(s) = \mathbf{C}_1\mathbf{X}(s) = \mathbf{C}_1\mathbf{\Lambda}(s)\begin{bmatrix}(\mathbf{D} + s\mathbf{M}) & \mathbf{I}\end{bmatrix}\mathcal{W}_{\text{so}}\widetilde{\mathbf{U}}(s),$$

which is the Laplace transform of $\mathbf{y}_{\text{L}}(t)$. We extract the input-to-output mapping from $\mathbf{Y}_{\text{L}}(s)$ to define the respective transfer function in the following.

Definition 3.62:
Consider the asymptotically stable second-order system (3.121) with initial conditions as defined in (3.119) and the matrix $\mathcal{W}_{\text{so}}$ as defined in (3.144). Then the *transfer function* corresponding to this system is defined as

$$\mathcal{G}_{\text{L},\mathcal{W}_{\text{so}}}(s) := \mathbf{C}_1\mathbf{\Lambda}(s)\begin{bmatrix}(\mathbf{D} + s\mathbf{M}) & \mathbf{I}\end{bmatrix}\mathcal{W}_{\text{so}}, \tag{3.146}$$

with $\mathbf{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$. ◊

Since we aim to maintain the second-order structure while considering a first-order input matrix $\mathcal{W}_{\text{so}}$, we are not able to write down a suitable system realization. However, for theoretical considerations, we assume that there is a homogeneous system representation of the transfer function $\mathcal{G}_{\text{L},\mathcal{W}_{\text{so}}}(s)$ to derive the controllability spaces. This relation is depicted in Figure 3.16, where the input and the initial conditions are applied by the matrix $\mathcal{W}_{\text{so}}$ and a suitable input $\widetilde{\mathbf{u}} \in L_2([0,\infty), \mathbf{R}^{m+n_{\mathbf{X}_0}+n_{\mathbf{V}_0}})$. In the following, we derive the controllability and observability Gramians of the second-order system from (3.121) using the transfer function $\mathcal{G}_{\text{L},\mathcal{W}_{\text{so}}}(s)$ from (3.146).

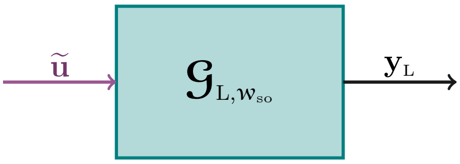

$\widetilde{\mathbf{u}}$
$\mathcal{G}_{\mathrm{L},\mathcal{W}_{\mathrm{so}}}$
$\mathbf{y}_{\mathrm{L}}$

Figure 3.16: Structure of a second-order ODE system with an extended input and a linear output.

Controllability Gramian To derive a controllability Gramian that spans the controllability space of the original system (3.121), we extract the input-to-state mapping from the transfer function $\mathcal{G}_{\mathrm{L},\mathcal{W}_{\mathrm{so}}}(s)$ in (3.146), which yields

$$\mathcal{C}_{\mathcal{W}_{\mathrm{so}}}(s) = \mathbf{\Lambda}(s)\begin{bmatrix}(\mathbf{D}+s\mathbf{M}) & \mathbf{I}\end{bmatrix}\mathcal{W}_{\mathrm{so}}. \tag{3.147}$$

Since this mapping encodes the controllability behavior of system (3.121), it is used to define a matrix $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}} := \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathcal{C}_{\mathcal{W}_{\mathrm{so}}}(\mathrm{i}\omega)\mathcal{C}_{\mathcal{W}_{\mathrm{so}}}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega$ that spans the respective controllability space.

Definition 3.63:
Consider the asymptotically stable second-order system (3.121) with initial conditions as defined in (3.119) and the input matrix $\mathcal{W}_{\mathrm{so}}$ as defined in (3.144). Then the corresponding *second-order controllability Gramian* is defined as

$$\mathbf{P}_{\mathcal{W}_{\mathrm{so}}} := \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{\Lambda}(\mathrm{i}\omega)\begin{bmatrix}(\mathbf{D}+\mathrm{i}\omega\mathbf{M}) & \mathbf{I}\end{bmatrix}\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}^{\mathrm{H}}\begin{bmatrix}(\mathbf{D}+\mathrm{i}\omega\mathbf{M})^{\mathrm{H}} \\ \mathbf{I}\end{bmatrix}\mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega, \tag{3.148}$$

with $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$. ◊

The Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ spans the controllability space of the state $\mathbf{x}(t)$ in system (3.121) without considering its derivative $\dot{\mathbf{x}}(t)$, as it does not affect the output. Hence, this Gramian is called *position controllability Gramian*.

The following theorem describes that the second-order Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ is determined by computing a first-order Gramian.

Theorem 3.64:
Consider the asymptotically stable second-order system (3.121) with initial conditions as defined in (3.119) and the input matrix $\mathcal{W}_{\mathrm{so}}$ as defined in (3.144). The second-order controllability Gramians $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ as defined in (3.148) is equal to the upper-left block $\mathbf{P}_1$ of the first-order controllability Gramian

$$\mathcal{P}_{\mathcal{W}_{\mathrm{so}}} = \begin{bmatrix}\mathbf{P}_{1,\mathcal{W}_{\mathrm{so}}} & \mathbf{P}_{2,\mathcal{W}_{\mathrm{so}}} \\ \mathbf{P}_{2,\mathcal{W}_{\mathrm{so}}}^{\mathrm{T}} & \mathbf{P}_{3,\mathcal{W}_{\mathrm{so}}}\end{bmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-1}\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}^{\mathrm{T}}(-\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-\mathrm{T}}\mathrm{d}\omega \tag{3.149}$$

with first-order matrices $\mathcal{E}$ and $\mathcal{A}$ as defined in (2.24). ◊

Proof. Applying the Schur complement to $(\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-1}$ as shown in (3.130) leads to

$$\begin{aligned}\mathcal{P}_{\mathcal{W}_{\mathrm{so}}} &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{bmatrix}\mathbf{\Lambda}(\mathrm{i}\omega)(\mathrm{i}\omega\mathbf{M}+\mathbf{D}) & \mathbf{\Lambda}(\mathrm{i}\omega)\\ -\mathbf{\Lambda}(\mathrm{i}\omega)\mathbf{K} & \mathrm{i}\omega\mathbf{\Lambda}(\mathrm{i}\omega)\end{bmatrix}\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}^{\mathrm{T}}(\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-\mathrm{H}}\mathrm{d}\omega\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{bmatrix}\mathbf{\Lambda}(\mathrm{i}\omega)\begin{bmatrix}(\mathrm{i}\omega\mathbf{M}+\mathbf{D}) & \mathbf{I}\end{bmatrix}\\ \mathbf{\Lambda}(\mathrm{i}\omega)\begin{bmatrix}-\mathbf{K} & \mathrm{i}\omega\mathbf{I}\end{bmatrix}\end{bmatrix}\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}^{\mathrm{T}}(\mathrm{i}\omega\mathcal{E}-\mathcal{A})^{-\mathrm{H}}\mathrm{d}\omega \qquad (3.150)\\ &= \begin{bmatrix}\mathbf{P}_{\mathcal{W}_{\mathrm{so}}} & *\\ * & *\end{bmatrix}\end{aligned}$$

for $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$ and with $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ as defined in (3.148). □

To compute the position controllability Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$, we can solve a Lyapunov equation of the form (3.12) with $\mathcal{B} = \mathcal{W}_{\mathrm{so}}$ to compute a first-order Gramian $\mathcal{P}_{\mathcal{W}_{\mathrm{so}}}$ and to extract $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ from its upper-left block. The Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ derived in this paragraph is used later in this work to apply balanced truncation for systems with a second-order structure.

Remark 3.65:
Note, that the mapping $\mathcal{C}_{\mathcal{W}_{\mathrm{so}}}(s)$ from (3.147) is the Laplace transform of the mapping $\begin{bmatrix}\mathbf{I} & 0\end{bmatrix}\boldsymbol{c}_{\mathcal{W}}(t)$ with $\boldsymbol{c}_{\mathcal{W}}(t)$ as defined in (3.28) for first-order matrices from (2.24) and $\mathcal{W} = \mathcal{W}_{\mathrm{so}}$. Moreover, the first-order Gramian $\mathcal{P}_{\mathcal{W}_{\mathrm{so}}}$ from (3.149) is equal to the Gramian $\mathcal{P}_{\mathcal{W}}$ defined in (3.29) for $\mathcal{W} = \mathcal{W}_{\mathrm{so}}$. Hence, we can also define the second-order Gramian in the time domain as

$$\mathbf{P}_{\mathcal{W}_{\mathrm{so}}} = \int_0^{\infty}\begin{bmatrix}\mathbf{I} & 0\end{bmatrix}e^{\mathcal{E}^{-1}\mathcal{A}t}\mathcal{E}^{-1}\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}^{\mathrm{T}}\mathcal{E}^{-\mathrm{T}}e^{(\mathcal{E}^{-1}\mathcal{A})^{\mathrm{T}}t}\begin{bmatrix}\mathbf{I}\\ 0\end{bmatrix}\mathrm{d}t.$$

◊

Observability Gramian To derive the second-order observability Gramian of the system (3.121) that describes its observability properties, we extract the state-to-output mapping from the respective transfer function $\mathcal{G}_{\mathrm{L},\mathcal{W}_{\mathrm{so}}}(s)$ in (3.146), that coincides with $\mathcal{O}_{\mathrm{L}}(s)$ from (3.137). Hence, from that mapping, we derive the same observability Gramian as defined in (3.138).

Definition 3.66:
Consider the asymptotically stable second-order system (3.121) and define $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$. Then the corresponding *second-order observability Gramian* is defined as

$$\mathbf{Q}_{\mathrm{L}} := \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathbf{C}_1^{\mathrm{H}}\mathbf{C}_1\mathbf{\Lambda}(\mathrm{i}\omega)\mathrm{d}\omega. \qquad (3.151)$$

◊

To compute that Gramian, we apply Theorem 3.59, which describes that $\mathbf{Q}_{\mathrm{L}}$ is equal to the lower-right block of the first-order controllability Gramian $\mathcal{Q}_{\mathrm{L}}$ from (3.139). Hence, we compute that Gramian by solving a Lyapunov equation.

Controllability energies To describe the controllability behavior of the system (3.121), we derive the respective system energies. Therefore, we define the input-to-state mapping in the time domain, which is

$$\boldsymbol{c}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}(t) = \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{W}}_{\mathrm{so}}, \tag{3.152}$$

with first-order matrices $\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\mathcal{A}}$ as defined in (2.24) and $\boldsymbol{\mathcal{W}}_{\mathrm{so}}$ as defined in (3.144), so that the mapping $\mathfrak{C}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$ defined in (3.147) is the Laplace transform of $\boldsymbol{c}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$, see Remark 3.65. To evaluate the controllability behavior of the system, we apply the energy norm from (3.19) to the input-to-state mapping $\boldsymbol{c}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$, which yields

$$\begin{aligned} E(\boldsymbol{c}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}) &= \|\boldsymbol{c}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}\|^2_{L_2\left([0,\infty),\mathbb{R}^{2n\times(n+n_{\mathbf{x}_0}+n_{\mathbf{v}_0})}\right)} = \int_0^\infty \operatorname{tr}\left(\boldsymbol{c}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}(t)\boldsymbol{c}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}(t)^{\mathrm{T}}\right)\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{tr}\left(\mathfrak{C}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}(\mathrm{i}\omega)\mathfrak{C}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}(\mathrm{i}\omega)^{\mathrm{H}}\right)\mathrm{d}\omega \\ &= \operatorname{tr}(\mathbf{P}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}})\,. \end{aligned} \tag{3.153}$$

Since the trace of a Gramian is equal to the sum of its eigenvalues, it follows that the states corresponding to large eigenvalues of $\mathbf{P}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$ have the highest impact on the respective system and, hence, encode the dominant controllability subspaces of the system (3.121). On the other hand, states corresponding to small eigenvalues of the Gramians $\mathbf{P}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$ have little effect on the system dynamics and are, therefore, negligible.

Observability energies To investigate the output energies of the second-order system (3.121), we evaluate the energy norm of the state-to-output mapping

$$\boldsymbol{o}_{\mathrm{L}}(t) = \boldsymbol{\mathfrak{C}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\begin{bmatrix}0\\ \mathbf{I}\end{bmatrix}$$

in the time domain. The Laplace transform to the mapping $\boldsymbol{o}_{\mathrm{L}}$ is equal to the mapping $\mathfrak{O}_{\mathrm{L}}$ from (3.137) in the frequency domain. This mapping coincides with the mapping introduced in (3.142). Hence, applying the energy norm leads to the same energy expression as in (3.143) that is

$$E(\boldsymbol{o}_{\mathrm{L}}) = \operatorname{tr}(\mathbf{Q}_{\mathrm{L}})\,.$$

The trace of the Gramian $\mathbf{Q}_{\mathrm{L}}$ is equal to the sum of its eigenvalues, which indicates that the largest eigenvalues have the greatest impact on the energy norm values and the system's dynamics. Consequently, the states associated with these large eigenvalues significantly influence the system's behavior. Therefore, the states corresponding to the large eigenvalues of the Gramian $\mathbf{Q}_{\mathrm{L}}$ form the dominant observability subspaces.

We summarize the extended-input approach and the resulting properties in Table 3.10, which depicts the transfer function, the derived Gramians, and the respective energies used in the following chapters to reduce systems of this structure.

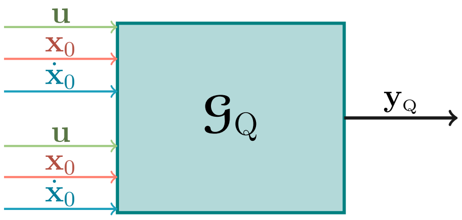


Figure 3.17: Structure of a second-order ODE system with a quadratic output.

	System (3.121)
Transfer function	$\mathcal{G}_{\mathrm{L},\mathbf{w}_{\mathrm{so}}}(s)$
Controllability Gramian	$\mathbf{P}_{\mathbf{w}_{\mathrm{so}}}$
Observability Gramian	$\mathbf{Q}_{\mathrm{L}}$
Controllability energies	$E(\boldsymbol{c}_{\mathbf{w}_{\mathrm{so}}}) = \mathrm{tr}(\mathbf{P}_{\mathbf{w}_{\mathrm{so}}})$
Observability energies	$E(\boldsymbol{o}_{\mathrm{L}}) = \mathrm{tr}(\mathfrak{Q}_{\mathrm{L}})$

Table 3.9: Properties of system (3.121) corresponding to its extended-input representation.

3.3.2 Inhomogeneous second-order ODE systems with a quadratic output

In this subsection, we consider the class of second-order systems with a quadratic output equation of the form

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) &= \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0, \\ \mathbf{y}_{\mathrm{Q}}(t) &= \begin{bmatrix} \mathbf{x}(t)^{\mathrm{T}} & \dot{\mathbf{x}}(t)^{\mathrm{T}} \end{bmatrix} \mathcal{M} \begin{bmatrix} \mathbf{x}(t) \\ \dot{\mathbf{x}}(t) \end{bmatrix} \end{aligned} \tag{3.154}$$

with a state equation as defined in (3.118) and a quadratic output equation that includes a symmetric output matrix $\mathcal{M} \in \mathbb{R}^{2n \times 2n}$ and the output $\mathbf{y}_{\mathrm{Q}}(t) \in \mathbb{R}$. Figure 3.17 depicts the system structure where we again indicate the quadratic output equation by adding the input, the displacement initial condition, and the velocity initial condition twice to the system dynamics.

The output matrix $\mathcal{M}$ is decomposed as described in (1.7), where we additionally assume that $\mathbf{M}_{12} = 0$ and $\mathbf{M}_{22} = 0$. Otherwise, if one of these submatrices is not equal to the zero matrices, we consider the respective system in first-order representation

(2.24) and apply the theory from Section 3.1.2. Hence, in the following, we investigate the output equation

$$\mathbf{y}_{\mathrm{Q}}(t) = \begin{bmatrix}\mathbf{x}(t)^{\mathrm{T}} & \dot{\mathbf{x}}(t)^{\mathrm{T}}\end{bmatrix} \begin{bmatrix}\mathbf{M}_{11} & 0 \\ 0 & 0\end{bmatrix} \begin{bmatrix}\mathbf{x}(t) \\ \dot{\mathbf{x}}(t)\end{bmatrix} = \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}_{11}\mathbf{x}(t). \tag{3.155}$$

We aim to preserve the second-order structure and include initial conditions in the analysis so that the effects of initial conditions on the output are considered.

We have already used two approaches that consider the initial conditions while evaluating the controllability and observability behavior. The first approach is the multi-system approach, in which the superposition principles are used to derive subsystems for each input and initial condition component. This approach is discussed for second-order systems with a quadratic output equation in Section 3.3.2.1. The second approach incorporates the initial conditions into the input matrix and is called the extended-input approach, which we present for this class of systems in (3.154).

3.3.2.1 Multi-system approach for inhomogeneous second-order ODE systems with a quadratic output

We consider the output in (3.155). As described in (3.120), the state consists of three components: one corresponding to the input $\mathbf{u}(t)$, one to the position initial condition $\mathbf{x}_0$, and one to the velocity initial condition $\dot{\mathbf{x}}_0$. Inserting the three components of $\mathbf{x}(t)$ leads to 9 different output components that we aim to analyze separately in the multi-system approach. Analyzing those systems and applying reduction methods to each of them separately will be numerically prohibitive. Hence, applying the extended-input approach that includes the initial conditions in the input matrix, presented in the following subsection, is the preferred strategy for systems of the structure introduced in (3.154).

3.3.2.2 Extended-input approach for inhomogeneous second-order ODE systems with a quadratic output

In the extended-input approach, we derive an extended input matrix $\boldsymbol{\mathcal{W}}_{\mathrm{so}}$ that includes the input space and the initial condition spaces as defined in (3.144). Using this matrix, we derive a transfer function, tailored Gramians, and the respective energies that encode the system behavior.

Transfer function To derive the transfer function of the system (3.154), we consider its first-order representation (3.31) with matrices from (2.24), with $\boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{W}}_{\mathrm{so}}$. We insert the respective matrices into the transfer function as defined in (3.44) and using the Schur complement from (3.130) to obtain

$$\begin{aligned}\mathcal{G}_{\mathrm{Q},\mathcal{W}\mathcal{W}}(s_1,s_2) &:= \mathcal{W}^{\mathrm{T}}(s_1\mathcal{E}-\mathcal{A})^{-\mathrm{H}}\mathcal{M}(s_2\mathcal{E}-\mathcal{A})^{-1}\mathcal{W} \\ &= \begin{bmatrix} 0 & \mathbf{X}_0 & 0 \\ \mathbf{B} & 0 & \mathbf{M}\mathbf{V}_0 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \mathbf{\Lambda}(s)(s\mathbf{M}+\mathbf{D}) & \mathbf{\Lambda}(s) \\ -\mathbf{\Lambda}(s)\mathbf{K} & s\mathbf{\Lambda}(s) \end{bmatrix}^{\mathrm{H}} \\ &\qquad \cdot \begin{bmatrix} \mathbf{M}_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}(s)(s\mathbf{M}+\mathbf{D}) & \mathbf{\Lambda}(s) \\ -\mathbf{\Lambda}(s)\mathbf{K} & s\mathbf{\Lambda}(s) \end{bmatrix} \begin{bmatrix} 0 & \mathbf{X}_0 & 0 \\ \mathbf{B} & 0 & \mathbf{M}\mathbf{V}_0 \end{bmatrix} \\ &= \mathcal{W}_{\mathrm{so}}^{\mathrm{T}} \begin{bmatrix} (\mathbf{D}+s_1\mathbf{M})^{\mathrm{H}} \\ \mathbf{I} \end{bmatrix} \mathbf{\Lambda}(s_1)^{\mathrm{H}}\mathbf{M}_{11}\mathbf{\Lambda}(s_2) \begin{bmatrix} (\mathbf{D}+s_2\mathbf{M}) & \mathbf{I} \end{bmatrix} \mathcal{W}_{\mathrm{so}}.\end{aligned}$$

Since we consider matrices corresponding to the second-order system representation, we denote the respective transfer function $\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}}(s_1,s_2)$ in the following, which yields the following definition.

Definition 3.67:
Consider the asymptotically stable second-order system in (3.154) with initial conditions as defined in (3.119). Also consider the input matrix $\mathcal{W}_{\mathrm{so}}$ as defined in (3.144) and define $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$. Then the *transfer function* of this system is defined as

$$\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}}(s_1,s_2) := \mathcal{W}_{\mathrm{so}}^{\mathrm{T}} \begin{bmatrix} (\mathbf{D}+s_1\mathbf{M})^{\mathrm{H}} \\ \mathbf{I} \end{bmatrix} \mathbf{\Lambda}(s_1)^{\mathrm{H}}\mathbf{M}_{11}\mathbf{\Lambda}(s_2) \begin{bmatrix} (\mathbf{D}+s_2\mathbf{M}) & \mathbf{I} \end{bmatrix} \mathcal{W}_{\mathrm{so}}. \quad (3.156)$$

◊

We observe that the inhomogeneous second-order system (3.154) has a transfer function of the same structure as a homogeneous system with the first-order input matrix $\mathcal{W}_{\mathrm{so}}$. However, since we aim to maintain the second-order structure, we are not able to write down a suitable system realization. For theoretical considerations, we assume that there is a homogeneous system representation of the transfer function $\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}}(s_1,s_2)$ to derive the controllability and observability spaces in the following. Figure 3.18 depicts that we analyze the system while considering an input matrix $\mathcal{W}_{\mathrm{so}}$ that includes the input and initial condition spaces, indicated by a suitable input $\widetilde{\mathbf{u}}$. In the following, we describe the behavior of this system in terms of controllability and observability. For this purpose, we derive the corresponding controllability and observability Gramians that encode these behaviors.

Controllability Gramians To describe the controllability behavior of system (3.154), we first derive the input-to-state mapping $\mathcal{C}_{\mathcal{W}_{\mathrm{so}}}(t)$ from the transfer function $\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}}(s_1,s_2)$ in (3.144), that coincides with the one defined in (3.147) as the state equation coincides for the systems (3.121) and (3.154). Hence, the same controllability Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ as defined in (3.148) encodes the controllability space of system (3.154).

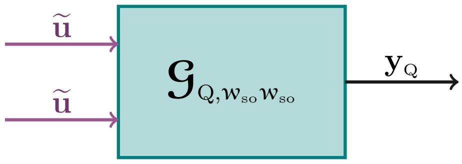

$\widetilde{\mathbf{u}}$
$\widetilde{\mathbf{u}}$
$\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}}$
$\mathbf{y}_{\mathrm{Q}}$

Figure 3.18: Structure of a second-order ODE system with an extended input and a quadratic output.

Definition 3.68:
Consider the asymptotically stable second-order system (3.154) with initial conditions as defined in (3.119) and the input matrix $\mathcal{W}_{\mathrm{so}}$ as defined in (3.144). Then the corresponding *second-order controllability Gramian* is defined as

$$\mathbf{P}_{\mathcal{W}_{\mathrm{so}}} := \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{\Lambda}(\mathrm{i}\omega)\left[(\mathbf{D}+\mathrm{i}\omega\mathbf{M}) \quad \mathbf{I}\right]\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}^{\mathrm{H}}\begin{bmatrix}(\mathbf{D}+\mathrm{i}\omega\mathbf{M})^{\mathrm{H}}\\ \mathbf{I}\end{bmatrix}\mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega,$$

with $\mathbf{\Lambda}(s) := (s^2\mathbf{M}+s\mathbf{D}+\mathbf{K})^{-1}$. ◊

The Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ spans the controllability space of the state $\mathbf{x}(t)$ from system (3.154). This Gramian is computed as the upper-left block of a first-order Gramian $\mathcal{P}_{\mathcal{W}_{\mathrm{so}}}$ as described in Theorem 3.64.

Observability Gramians Now, we describe the observability properties of the system (3.154). Therefore, we aim to derive an observability Gramian that encodes the respective observability space. Since we consider a quadratic output equation, we describe the controllability properties of the state $\mathbf{x}(t)$ multiplied from the right to the quadratic output expression in (3.154), taking into account the controllability space of the left state. For that, we can rewrite $\mathbf{y}_{\mathrm{Q}}(t)$ by defining the state-dependent function

$$\mathbf{C}_{11}(\mathbf{x}(t)) := \mathbf{x}(t)^{\mathrm{T}}\mathbf{M}_{11}.$$

Applying that representation to the output yields $\mathbf{y}_{\mathrm{Q}}(t) = \mathbf{C}_{11}\left(\mathbf{x}(t)\right)\mathbf{x}(t)$. We observe, that the observability of the (right) state $\mathbf{x}(t)$ in the output $\mathbf{y}_{\mathrm{Q}}(t) = \mathbf{C}_{11}(\mathbf{x}(t))\mathbf{x}(t)$ also depends on the reachability of the (left) state $\mathbf{x}(t)$. Hence, we expect that the observability Gramian will depend on the controllability Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ defined in (3.148).

From the transfer function $\mathcal{G}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so}}}(s_1,s_2)$ in (3.144), we identify the input-to-state mapping $\mathcal{C}_{\mathcal{W}_{\mathrm{so}}}(s)$ from (3.147) corresponding to the right state $\mathbf{X}(s)$ in the frequency domain, so that the remaining state-to-output mapping is

$$\mathcal{O}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}(s_1,s_2) := \mathcal{W}_{\mathrm{so}}^{\mathrm{T}}\begin{bmatrix}(\mathbf{D}+s_1\mathbf{M})^{\mathrm{H}}\\ \mathbf{I}\end{bmatrix}\mathbf{\Lambda}(s_1)^{\mathrm{H}}\mathbf{M}_{11}\mathbf{\Lambda}(s_2). \tag{3.157}$$

Since the mapping $\boldsymbol{\mathcal{O}}_{\mathrm{Q},\boldsymbol{\mathcal{W}}_{\mathrm{so}}}(s_1, s_2)$ spans the observability space of the right state $\mathbf{X}(s)$ considering the space in which the left state $\mathbf{x}(s)$ lives, it is used to define a matrix

$$\begin{aligned}
\mathbf{Q}_{\mathrm{Q},\boldsymbol{\mathcal{W}}_{\mathrm{so}}} &:= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{\mathcal{O}}_{\mathrm{Q},\boldsymbol{\mathcal{W}}_{\mathrm{so}}}(\mathrm{i}\omega_1, \mathrm{i}\omega_2)^{\mathrm{H}}\boldsymbol{\mathcal{O}}_{\mathrm{Q},\boldsymbol{\mathcal{W}}_{\mathrm{so}}}(\mathrm{i}\omega_1, \mathrm{i}\omega_2)\mathrm{d}\omega_1\mathrm{d}\omega_2 \\
&= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{\Lambda}(\mathrm{i}\omega_2)^{\mathrm{H}}\mathbf{M}_{11}\mathbf{\Lambda}(\mathrm{i}\omega_1)\begin{bmatrix}(\mathbf{D} + \mathrm{i}\omega_1\mathbf{M}) & \mathbf{I}\end{bmatrix}\boldsymbol{\mathcal{W}}_{\mathrm{so}} \\
&\qquad\qquad\qquad\qquad \cdot \boldsymbol{\mathcal{W}}_{\mathrm{so}}^{\mathrm{T}}\begin{bmatrix}(\mathbf{D} + \mathrm{i}\omega_1\mathbf{M})^{\mathrm{H}} \\ \mathbf{I}\end{bmatrix}\mathbf{\Lambda}(\mathrm{i}\omega_1)^{\mathrm{H}}\mathbf{M}_{11}\mathbf{\Lambda}(\mathrm{i}\omega_2)\mathrm{d}\omega_1\mathrm{d}\omega_2 \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{\Lambda}(\mathrm{i}\omega_2)^{\mathrm{H}}\mathbf{M}_{11}\mathbf{P}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}\mathbf{M}_{11}\mathbf{\Lambda}(\mathrm{i}\omega_2)\mathrm{d}\omega_2
\end{aligned}$$

that spans the observability space of the right state $\mathbf{X}(s)$ in the frequency domain or $\mathbf{x}(t)$ in the time domain.

Definition 3.69:
Consider the asymptotically stable second-order system (3.154) with initial conditions as defined in (3.119), the corresponding controllability Gramian $\mathbf{P}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$ as introduced in (3.148), and the input matrix $\boldsymbol{\mathcal{W}}_{\mathrm{so}}$ as defined in (3.144). Then the corresponding *second-order observability Gramian* is defined as

$$\mathbf{Q}_{\mathrm{Q},\boldsymbol{\mathcal{W}}_{\mathrm{so}}} := \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{\Lambda}(\mathrm{i}\omega)^{\mathrm{H}}\mathbf{M}_{11}\mathbf{P}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}\mathbf{M}_{11}\mathbf{\Lambda}(\mathrm{i}\omega)\mathrm{d}\omega \tag{3.158}$$

with $\mathbf{\Lambda}(s) := (s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}$. ◊

To compute the second-order observability Gramian $\mathbf{Q}_{\mathrm{Q},\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$, we apply the following theorem.

Theorem 3.70:
Consider the asymptotically stable second-order system (3.154) with initial conditions as defined in (3.119), the corresponding controllability Gramian $\mathbf{P}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$ from (3.148), the first-order matrices $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{A}}$ as defined in (2.24), and the input matrix $\boldsymbol{\mathcal{W}}_{\mathrm{so}}$ as defined in (3.144). Then the second-order observability Gramian $\mathbf{Q}_{\mathrm{Q},\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$ as defined in (3.158) is the lower-right block $\mathbf{Q}_{3,\boldsymbol{\mathcal{W}}_{\mathrm{so}}}$ of the first-order matrix

$$\mathbf{Q}_{\mathrm{Q},\boldsymbol{\mathcal{W}}_{\mathrm{so}}} := \begin{bmatrix}\boldsymbol{\mathcal{Q}}_{1,\boldsymbol{\mathcal{W}}_{\mathrm{so}}} & \boldsymbol{\mathcal{Q}}_{2,\boldsymbol{\mathcal{W}}_{\mathrm{so}}} \\ \boldsymbol{\mathcal{Q}}_{2,\boldsymbol{\mathcal{W}}_{\mathrm{so}}}^{\mathrm{T}} & \boldsymbol{\mathcal{Q}}_{3,\boldsymbol{\mathcal{W}}_{\mathrm{so}}}\end{bmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\mathrm{i}\omega\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\mathrm{H}}\begin{bmatrix}\mathbf{M}_{11}\mathbf{P}_{\boldsymbol{\mathcal{W}}_{\mathrm{so}}}\mathbf{M}_{11} & 0 \\ 0 & 0\end{bmatrix}(\mathrm{i}\omega\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathrm{d}\omega \diamond$$

Proof. The proof of this theorem is similar to the one for Theorem 3.59. □

Controllability energies To identify state spaces that encode the dominant system dynamics, we derive the respective controllability energies. For that, we derive the input-to-state mapping in the time domain

$$\boldsymbol{c}_{\mathcal{W}_{\mathrm{so}}}(t) = \begin{bmatrix}\mathbf{I} & 0\end{bmatrix} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{W}}_{\mathrm{so}}$$

which is equal to the mapping defined in (3.152). Hence, we obtain the same energy expression as in (3.153) applying the energy norm from (3.19) that is

$$\begin{aligned}E(\boldsymbol{c}_{\mathcal{W}_{\mathrm{so}}}) = \|\boldsymbol{c}_{\mathcal{W}_{\mathrm{so}}}\|^2_{L_2\left([0,\infty),\mathbb{R}^{2n\times(n+n_{\mathbf{X}_0}+n_{\mathbf{V}_0})}\right)} &= \int_{-\infty}^{\infty} \mathrm{tr}\left(\boldsymbol{c}_{\mathcal{W}_{\mathrm{so}}}(t)\boldsymbol{c}_{\mathcal{W}_{\mathrm{so}}}(t)^{\mathrm{T}}\right)\mathrm{d}t\\ &= \mathrm{tr}(\mathbf{P}_{\mathcal{W}_{\mathrm{so}}})\,.\end{aligned}$$

The trace of the Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ coincides with the sum of its eigenvalues. Therefore, the eigenvalues of the Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ indicate which states are significant for the system dynamics. The states corresponding to large eigenvalues span the dominant controllability spaces, while states corresponding to small eigenvalues have a negligible influence on the system dynamics.

Observability energies In this paragraph, we derive the observability energies to identify the dominant observability subspaces of the system (3.154). Therefore, we consider the state-to-output mapping

$$\boldsymbol{o}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}(t_1,t_2) := \boldsymbol{\mathcal{W}}_{\mathrm{so}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_1}\begin{bmatrix}\mathbf{M}_{11} & 0\\ 0 & 0\end{bmatrix}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_2}\boldsymbol{\mathcal{E}}^{-1}\begin{bmatrix}0\\ \mathbf{I}\end{bmatrix}$$

whose 2-dimensional Laplace transform is equal to $\boldsymbol{\mathcal{O}}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}(s_1,s_2)$ from (3.157). Applying the energy norm from (3.19) to the mapping $\boldsymbol{o}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}$ leads to

$$\begin{aligned}E(\boldsymbol{o}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}) &= \|\boldsymbol{o}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}\|^2_{L_2\left([0,\infty)^2,\mathbb{R}^{(n+n_{\mathbf{X}_0}+n_{\mathbf{V}_0})\times n}\right)}\\ &= \int_0^{\infty}\int_0^{\infty}\mathrm{tr}\left(\boldsymbol{o}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}(t_1,t_2)^{\mathrm{T}}\boldsymbol{o}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}(t_1,t_2)\right)\mathrm{d}t_1\mathrm{d}t_2\\ &= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{tr}\left(\boldsymbol{\mathcal{O}}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}(\mathrm{i}\omega_1,\mathrm{i}\omega_2)^{\mathrm{H}}\boldsymbol{\mathcal{O}}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}(\mathrm{i}\omega_1,\mathrm{i}\omega_2)\right)\mathrm{d}\omega_1\mathrm{d}\omega_2\\ &= \mathrm{tr}(\mathbf{Q}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}})\end{aligned}$$

with $\mathbf{Q}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}$ as defined in (3.158). Since the trace of the Gramian $\mathbf{Q}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}$ coincides with the sum of its eigenvalues, the most dominant observability subspaces are determined by the largest eigenvalues.

We summarize the extended-input approach for system (3.154) in Table 3.10 where we list the transfer function, the tailored controllability and observability Gramian, and the resulting energies. They are used in the following chapters to reduce systems of this type.

	System (3.154)
Transfer function	$\mathcal{G}_{\mathrm{Q},\boldsymbol{w}_{\mathrm{so}}\boldsymbol{w}_{\mathrm{so}}}(s_1, s_2)$
Controllability Gramian	$\mathbf{P}_{\boldsymbol{w}_{\mathrm{so}}}$
Observability Gramian	$\mathbf{Q}_{\mathrm{Q},\boldsymbol{w}_{\mathrm{so}}}$
Controllability energies	$E(\boldsymbol{c}_{\boldsymbol{w}_{\mathrm{so}}}) = \mathrm{tr}(\mathbf{P}_{\boldsymbol{w}_{\mathrm{so}}})$
Observability energies	$E(\boldsymbol{o}_{\mathrm{Q},\boldsymbol{w}_{\mathrm{so}}}) = \mathrm{tr}(\mathbf{Q}_{\mathrm{Q},\boldsymbol{w}_{\mathrm{so}}})$

Table 3.10: Properties of system (3.154) corresponding to its extended-input representation.

CHAPTER 4

MODEL ORDER REDUCTION FOR SYSTEMS IN NON-STANDARD FORM

Contents

As described in Chapter 1, we aim to optimize external dampers added to vibrational systems describing civil engineering infrastructure, such as buildings or bridges, to suppress external vibration forces caused by, e.g., wind disturbances or earthquakes. In this chapter, we consider parameter-independent systems, i.e., we assume that one set of external dampers is given for which we aim to evaluate the system behavior. However, detailed modeling of these structures leads to systems with large dimensions that make their evaluation computationally expensive. Therefore, we aim to derive methods that reduce the model dimension while maintaining or approximating the system dynamics. Several classes of model order reduction methods for parameter-independent systems are listed in Section 2.2. Since the methods of choice presented in Section 2.2, BT and IRKA, consider only homogeneous first-order systems with a linear output equation, in this chapter, we introduce BT and IRKA for the different systems in the non-standard form presented in Chapter 3.

The authors in [66] and [15] introduce the BT method and the IRKA method for inhomogeneous first-order ODE systems with a linear output equation. We describe these methods and derive new model reduction schemes for inhomogeneous first-order ODE systems with a quadratic output equation, for inhomogeneous first-order DAE systems with a linear and a quadratic output equation, and inhomogeneous second-order ODE systems with a linear and a quadratic output equation. The main contribution of this section is the introduction of BT schemes for these system types. Moreover, we derive suitable error bounds, which are needed to evaluate the quality of the approximation. The IRKA methods for systems with linear output equations are a byproduct of the modified and decomposed system structures presented in Chapter 3 and are, therefore, also explained in this chapter, but at a low level of detail.

First, in Section 4.1, we consider first-order systems with an ODE as a state equation. Then, in Section 4.2, we study first-order systems with a DAE as a state equation, and finally, in Section 4.3, we consider the case of second-order systems. In all the sections, we consider systems with linear and quadratic output equations. Moreover, we illustrate the proposed methods on benchmark problems.

4.1 Model order reduction for inhomogeneous first-order ODE systems

In this section, we reduce first-order systems with a linear output equation and with a quadratic output equation as described in (3.5) and (3.31), respectively. We assume that the matrix $\mathcal{E}$ is nonsingular, i.e., we consider an ODE as a state equation, and the initial state is equal to $\mathbf{z}(0) = \mathbf{z}_0 = \mathbf{Z}_0\zeta_0$, see (3.4). We aim to reduce the systems (3.5) and (3.31) to obtain surrogate models with significantly smaller dimensions. These reduced surrogate models are then supposed to approximate the input- and initial condition-to-output behavior of the original systems.

We review in this section the BT and IRKA methods widely used in practice. The authors in [15] and [66] already derived reduction schemes for inhomogeneous first-order ODE systems with a linear output equation, i.e., for systems of the form (3.5). Therefore, in Section 4.1.1, we repeat the respective BT method and extend it to systems with a quadratic output equation. Then, in Section 4.1.2, we describe the IRKA method for inhomogeneous first-order systems with linear output equations.

4.1.1 BT for inhomogeneous first-order ODE systems

For systems with homogeneous initial conditions and a linear output equation, BT was introduced in [20, 26, 93, 138] and repeated in Section 2.2.1. Also, in [20], the authors derive a BT method for homogeneous systems with a quadratic output equation. However, in this section, we consider the class of inhomogeneous first-order ODE systems. In the literature, there are some approaches to reduce these inhomogeneous first-order systems with a linear output equation, see [13, 15, 66, 121]. In this work, we focus on the methods from [66], where the input $\mathcal{B}\mathbf{u}(t)$ is extended by the initial condition space $\mathbf{Z}_0$, and from [15], where the author's strategy is to decompose the system into a zero initial condition subsystem and a subsystem with initial conditions but no input. Since both methods are introduced for systems with a linear output equation, we extend these methods to systems with a quadratic output equation.

This subsection is structured as follows: First, in Section 4.1.1.1, we apply the method from [15] to derive some reduced surrogate models that sum up to an output that approximates the original output. We also extend this method to inhomogeneous first-order systems with a quadratic output equation. Afterwards, in Section 4.1.1.2, the method from [66] is applied for systems with a linear output equation and extended to those with quadratic output equations.

4.1.1.1 Multi-system approach for inhomogeneous first-order ODE systems

When applying the BT method for inhomogeneous first-order ODE systems using the multi-system approach, we distinguish between systems with linear and quadratic output equations since the subsystems differ. However, the reduction methodology is similar for both system classes.

BT for systems with a linear output equation First, we repeat the reduction approach from [15] for the class of inhomogeneous first-order systems (3.5) with a linear output equation. As we have seen in Section 3.1, the system (3.5) can be decomposed into two subsystems that are given in (3.8) and (3.9) so that the output is composed of

$$\mathbf{y}_{\mathrm{L}}(t) = \mathbf{y}_{\mathrm{L},\boldsymbol{\mathcal{B}}}(t) + \mathbf{y}_{\mathrm{L},\mathbf{z}_0}(t)$$

where $\mathbf{y}_{\mathrm{L},\boldsymbol{\mathcal{B}}}(t)$ and $\mathbf{y}_{\mathrm{L},\mathbf{z}_0}(t)$ are the outputs of the subsystems (3.8) and (3.9), respectively. The idea of the multi-system approach is to reduce both subsystems independently to obtain two reduced surrogate systems that are

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\dot{\mathbf{z}}_{\mathrm{r}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\mathbf{z}_{\mathrm{r}}(t) + \boldsymbol{\mathcal{B}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\mathbf{u}(t), \qquad \mathbf{z}_{\mathrm{r}}(0) = 0, \\ \mathbf{y}_{\mathrm{L,r},\boldsymbol{\mathcal{B}}}(t) &= \boldsymbol{\mathcal{C}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\mathbf{z}_{\mathrm{r}}(t) \end{aligned} \tag{4.1}$$

and

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{z}_0}\dot{\mathbf{z}}_{\mathrm{r}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{z}_0}\mathbf{z}_{\mathrm{r}}(t), \qquad \mathbf{z}_{\mathrm{r}}(0) = \mathbf{Z}_{0,\mathrm{r}}\zeta_0, \\ \mathbf{y}_{\mathrm{L,r},\mathbf{z}_0}(t) &= \boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{z}_0}\mathbf{z}_{\mathrm{r}}(t) \end{aligned} \tag{4.2}$$

with reduced matrices

$$\boldsymbol{\mathcal{E}}_{\mathrm{r},*} = \boldsymbol{\mathcal{V}}_{\mathrm{r},*}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{r},*}, \quad \boldsymbol{\mathcal{A}}_{\mathrm{r},*} = \boldsymbol{\mathcal{V}}_{\mathrm{r},*}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{r},*}, \quad \boldsymbol{\mathcal{B}}_{\mathrm{r},\boldsymbol{\mathcal{B}}} = \boldsymbol{\mathcal{V}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}, \quad \mathbf{Z}_{0,\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r},\mathbf{z}_0}^{\mathrm{T}}\mathbf{Z}_0, \quad \boldsymbol{\mathcal{C}}_{\mathrm{r},*} = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{T}}_{\mathrm{r},*} \tag{4.3}$$

encoded by the subscript $*$ that represents either '$\boldsymbol{\mathcal{B}}$' or '$\mathbf{Z}_0$'. We generate the reduced matrices (4.3) using projecting matrices $\boldsymbol{\mathcal{V}}_{\mathrm{r},*}, \boldsymbol{\mathcal{T}}_{\mathrm{r},*} \in \mathbb{R}^{N\times R_*}$ satisfying the Petrov-Galerkin conditions (2.30) and (2.31), with $R_* \ll N$. Then the output of the original system (3.5) is approximated by

$$\mathbf{y}_{\mathrm{L}}(t) \approx \mathbf{y}_{\mathrm{L,r}}(t) := \mathbf{y}_{\mathrm{L,r},\boldsymbol{\mathcal{B}}}(t) + \mathbf{y}_{\mathrm{L,r},\mathbf{z}_0}(t). \tag{4.4}$$

To derive such surrogate systems, we utilize the properties that were derived in Section 3.1.1 and summarized in Table 3.1. From the energy expressions in (3.20), (3.21), and (3.22) it follows that the dominant controllability and observability subspaces of the subsystems (3.8) and (3.9) are spanned by the states corresponding to the largest eigenvalues of the controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{z}_0}$, and the observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ introduced in (3.11), (3.14), and (3.17), respectively. Hence, states corresponding to the smallest eigenvalues of the respective Gramians are truncated in the following, resulting in Algorithm 7.

Algorithm 7 BT method for the first-order ODE system (3.5) with a linear output using the multi-system approach.

```
Require: The original system (3.5), the reduced dimensions $R_*$, where $*$ is '$\boldsymbol{\mathcal{B}}$' or '$\mathbf{Z}_0$'
    corresponding to subsystem (3.8) or (3.9).
Ensure: The reduced systems (4.1) and (4.2).
1: Compute factors of the Gramians $\boldsymbol{\mathcal{P}}_* \approx \boldsymbol{\mathcal{R}}_*\boldsymbol{\mathcal{R}}_*^{\mathrm{T}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{L}} \approx \boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^{\mathrm{T}}$ from (3.11), (3.14),
   and (3.17).
2: Perform the two SVDs of $\boldsymbol{\mathcal{S}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}}_*$ and decompose as
       $\boldsymbol{\mathcal{S}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}}_* = \mathbf{U}_*\boldsymbol{\Sigma}_*\mathbf{V}_*^{\mathrm{T}} = \begin{bmatrix}\mathbf{U}_{1,*} & \mathbf{U}_{2,*}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_{1,*} & 0\\ 0 & \boldsymbol{\Sigma}_{2,*}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{1,*}^{\mathrm{T}}\\ \mathbf{V}_{2,*}^{\mathrm{T}}\end{bmatrix}.$
   with $\boldsymbol{\Sigma}_{1,*} \in \mathbb{R}^{R_*\times R_*}$, $* \in \{\boldsymbol{\mathcal{B}}, \mathbf{Z}_0\}$.
3: Construct the projection matrices
       $\boldsymbol{\mathcal{V}}_{\mathrm{r},*} = \boldsymbol{\mathcal{S}}\mathbf{U}_{1,*}\boldsymbol{\Sigma}_{1,*}^{-\frac{1}{2}}, \qquad \boldsymbol{\mathcal{T}}_{\mathrm{r},*} = \boldsymbol{\mathcal{R}}_*\mathbf{V}_{1,*}\boldsymbol{\Sigma}_{1,*}^{-\frac{1}{2}}.$
4: Construct reduced matrices (4.3).
```

To evaluate the the matrices

$$\mathbf{B}_2 = \left(\boldsymbol{\mathcal{S}}\mathbf{U}_{2,\mathbf{z}_0}\boldsymbol{\Sigma}_{2,\mathbf{z}_0}^{-\frac{1}{2}}\right)^{\mathrm{T}}\boldsymbol{\mathcal{B}}, \qquad \mathbf{A}_{12} = \boldsymbol{\mathcal{V}}_{\mathrm{r},\mathbf{z}_0}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{R}}_{\mathbf{z}_0}\mathbf{V}_{2,\mathbf{z}_0}\boldsymbol{\Sigma}_{2,\mathbf{z}_0}^{-\frac{1}{2}},$$

and $\mathbf{Y}_2$ that is the lower block of $\boldsymbol{\mathcal{Y}} = \left[\begin{smallmatrix}\mathbf{Y}_1\\ \mathbf{Y}_2\end{smallmatrix}\right]$, which solves the Sylvester equation

$$\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{Y}}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{z}_0} + \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{Y}}\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{z}_0} = -\boldsymbol{\mathcal{C}}^{\mathrm{T}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{z}_0}.$$

Then the error bound given in [15, Theorem 3.2] is equal to

$$\begin{aligned}\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_\infty} &\leq \|\mathbf{y}_{\mathrm{L},\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{L,r},\boldsymbol{\mathcal{B}}}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{L},\mathbf{z}_0} - \mathbf{y}_{\mathrm{L,r},\mathbf{z}_0}\|_{L_\infty}\\ &\leq \left(2\sum_{k=R_{\boldsymbol{\mathcal{B}}}+1}^{N}\sigma_{k,\boldsymbol{\mathcal{B}}}\right)\|\mathbf{u}\|_{L_2([0,\infty),\mathbb{R}^m)} + \sqrt{\operatorname{tr}((\mathbf{B}_2\mathbf{B}_2^{\mathrm{T}} + 2\mathbf{Y}_2\mathbf{A}_{12})\boldsymbol{\Sigma}_{2,\mathbf{z}_0})}\|\zeta_0\|_2,\end{aligned}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\mathcal{B}}} = \operatorname{diag}(\sigma_{1,\boldsymbol{\mathcal{B}}}, \ldots, \sigma_{N,\boldsymbol{\mathcal{B}}})$ and $\boldsymbol{\Sigma}_{2,\mathbf{z}_0}$ result from the SVD in Step 2 of the algorithm.

BT for systems with a quadratic output equation Now, we derive a BT method for a system (3.31) with a quadratic output equation. As presented in Section 3.1.2, this system can be decomposed into four subsystems (3.34), (3.35), (3.36), and (3.37) so that the respective outputs satisfy

$$\mathbf{y}_{\mathrm{Q}}(t) = \mathbf{y}_{\mathrm{Q},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) + \mathbf{y}_{\mathrm{Q},\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t) + \mathbf{y}_{\mathrm{Q},\boldsymbol{\mathcal{B}}\mathbf{z}_0}(t) + \mathbf{y}_{\mathrm{Q},\mathbf{z}_0\mathbf{z}_0}(t).$$

Hence, in the multi-system approach, we derive four surrogate models approximating the input- and initial condition-to-output behavior of these subsystems. The first surrogate model is given as

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\dot{\mathbf{z}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\mathbf{z}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) + \boldsymbol{\mathcal{B}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\mathbf{u}(t), &\qquad \mathbf{z}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(0) &= 0,\\ \mathbf{y}_{\mathrm{Q},\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) &= \mathbf{z}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\mathbf{z}_{\mathrm{r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t), \end{aligned} \tag{4.5}$$

and approximates the input-to-output behavior of the subsystem (3.34). The second reduced model that approximates the input- and initial condition-to-output behavior of subsystem (3.35) is

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\dot{\mathbf{z}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r}\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\mathbf{z}_{\mathrm{r},\boldsymbol{\mathcal{B}}}(t) + \boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\mathbf{u}(t), &\qquad \mathbf{z}_{\mathrm{r},\boldsymbol{\mathcal{B}}}(0) &= 0,\\ \boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\dot{\mathbf{z}}_{\mathrm{r},\mathbf{Z}_0}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\mathbf{z}_{\mathrm{r},\mathbf{Z}_0}(t), & \mathbf{z}_{\mathrm{r},\mathbf{Z}_0}(0) &= \mathbf{Z}_{0,\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\zeta_0,\\ \mathbf{y}_{\mathrm{Q},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t) &= \mathbf{z}_{\mathbf{Z}_0}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\mathbf{z}_{\boldsymbol{\mathcal{B}}}(t) \end{aligned} \tag{4.6}$$

and the third one approximating the subsystem (3.36) is given as

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\dot{\mathbf{z}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\mathbf{z}_{\mathrm{r},\boldsymbol{\mathcal{B}}}(t) + \boldsymbol{\mathcal{B}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\mathbf{u}(t), &\qquad \mathbf{z}_{\mathrm{r},\boldsymbol{\mathcal{B}}}(0) &= 0,\\ \boldsymbol{\mathcal{E}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\dot{\mathbf{z}}_{\mathrm{r},\mathbf{Z}_0}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\mathbf{z}_{\mathrm{r},\mathbf{Z}_0}(t), & \mathbf{z}_{\mathrm{r},\mathbf{Z}_0}(0) &= \mathbf{Z}_{0,\mathrm{r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\zeta_0,\\ \mathbf{y}_{\mathrm{Q},\mathrm{r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}(t) &= \mathbf{z}_{\boldsymbol{\mathcal{B}}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\mathbf{z}_{\mathbf{Z}_0}(t). \end{aligned} \tag{4.7}$$

Finally, the surrogate system

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{Z}_0\mathbf{Z}_0}\dot{\mathbf{z}}_{\mathrm{r},\mathbf{Z}_0}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{Z}_0\mathbf{Z}_0}\mathbf{z}_{\mathrm{r},\mathbf{Z}_0}(t), &\qquad \mathbf{z}_{\mathrm{r},\mathbf{Z}_0}(0) &= \mathbf{Z}_{0,\mathrm{r},\mathbf{Z}_0\mathbf{Z}_0}\zeta_0,\\ \mathbf{y}_{\mathrm{Q},\mathrm{r},\mathbf{Z}_0\mathbf{Z}_0}(t) &= \mathbf{z}_{\mathbf{Z}_0}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\mathbf{Z}_0}\mathbf{z}_{\mathbf{Z}_0}(t), \end{aligned} \tag{4.8}$$

approximates the behavior of the subsystem (3.37). The respective reduced matrices are build using projecting matrices $\boldsymbol{\mathcal{V}}_{\mathrm{r},*\circ}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r},*\circ} \in \mathbb{R}^{N\times R_{*\circ}}$ satisfying the Petrov-Galerkin conditions (2.30), (2.31) with $R_{*\circ} \ll N$ where the subscripts $*$ and $\circ$ represent either '$\boldsymbol{\mathcal{B}}$'or '$\mathbf{Z}_0$', which yields the matrices

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ} &= \boldsymbol{\mathcal{V}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{r},*\circ}, &\qquad \boldsymbol{\mathcal{A}}_{\mathrm{r},*\circ} &= \boldsymbol{\mathcal{V}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{r},*\circ}, &\qquad \boldsymbol{\mathcal{B}}_{\mathrm{r},*\circ} &= \boldsymbol{\mathcal{V}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{B}},\\ \mathbf{Z}_{0,\mathrm{r},*\circ} &= \boldsymbol{\mathcal{V}}_{\mathrm{r},*\circ}^{\mathrm{T}}\mathbf{Z}_0, & \boldsymbol{\mathcal{M}}_{\mathrm{r},*\circ} &= \boldsymbol{\mathcal{T}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{T}}_{\mathrm{r},*\circ}. \end{aligned} \tag{4.9}$$

To derive the four reduced subsystems (4.5), (4.6), (4.7), and (4.8), we identify which states are most significant to describe the controllability and observability of the system. Therefore, we utilize the system energies summarized in Table 3.3. From the energy expression in (3.20), it follows that states corresponding to large eigenvalues of the controllability Gramian $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$ from (3.11) span the dominant controllability subspace of the two subsystems (3.34) and (3.35). Also, it follows from (3.21) that states corresponding to large eigenvalues of the Gramian $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ from (3.14) span the dominant controllability subspace of the subsystems (3.36) and (3.37). Hence, within the four BT methods

Algorithm 8 BT method for the first-order ODE system (3.31) with a quadratic output using the multi-system approach.

```
Require: The original system (3.31), the reduced orders $R_{*\circ}$, where $*$ and $\circ$ are '$\boldsymbol{\mathcal{B}}$' or
    '$\mathbf{Z}_0$' corresponding to the subsystem (3.34), (3.35), (3.36), and (3.37).
Ensure: The reduced systems (4.5), (4.6), (4.7), and (4.8).
1: Compute factors of the Gramians $\boldsymbol{\mathcal{P}}_* \approx \boldsymbol{\mathcal{R}}_*\boldsymbol{\mathcal{R}}_*^{\mathrm{T}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{Q},\circ} \approx \boldsymbol{\mathcal{S}}_\circ\boldsymbol{\mathcal{S}}_\circ^{\mathrm{T}}$ from Definition
   (3.11), (3.14) and (3.17).
2: Perform the four SVDs of $\boldsymbol{\mathcal{S}}_\circ^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}}_*$, and decompose as
   $\boldsymbol{\mathcal{S}}_\circ^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}}_* = \mathbf{U}_{*\circ}\mathbf{\Sigma}_{*\circ}\mathbf{V}_{*\circ}^{\mathrm{T}} = \begin{bmatrix}\mathbf{U}_{1,*\circ} & \mathbf{U}_{2,*\circ}\end{bmatrix}\begin{bmatrix}\mathbf{\Sigma}_{1,*\circ} & 0 \\ 0 & \mathbf{\Sigma}_{2,*\circ}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{1,*\circ}^{\mathrm{T}} \\ \mathbf{V}_{2,*\circ}^{\mathrm{T}}\end{bmatrix}.$
   with $\mathbf{\Sigma}_{1,*\circ} \in \mathbb{R}^{R_{*\circ}\times R_{*\circ}}$, $*,\circ \in \{\boldsymbol{\mathcal{B}}, \mathbf{Z}_0\}$.
3: Construct the projection matrices
   $\boldsymbol{\mathcal{V}}_{\mathrm{r},*\circ} = \boldsymbol{\mathcal{S}}_\circ\mathbf{U}_{1,*\circ}\mathbf{\Sigma}_{1,*}^{-\frac{1}{2}}, \qquad \boldsymbol{\mathcal{T}}_{\mathrm{r},*\circ} = \boldsymbol{\mathcal{R}}_*\mathbf{V}_{1,*\circ}\mathbf{\Sigma}_{1,*\circ}^{-\frac{1}{2}}.$
4: Construct reduced matrices (4.9).
```

applied to the four subsystems states corresponding to small eigenvalues of $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ are truncated as they are negligible when describing the system dynamics. To investigate the output energies of the four subsystems, we evaluate the energy norms of the respective state-to-output mappings from (3.42), (3.43), which show that states corresponding to small eigenvalues of the observability Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{Q},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{Q},\mathbf{Z}_0}$ from (3.39) and (3.41), respectively, are difficult to observe. Hence, in the following, we apply BT from Algorithm 1 extended by [20] to systems with quadratic output equations to truncate the states which are simultaneously hard to reach and to observe which leads to Algorithm 8. The reduced systems (4.5), (4.6), (4.7), and (4.8) generated by Algorithm 8 approximate the original outputs in the following way

$$\mathbf{y}_{\mathrm{Q}}(t) \approx \mathbf{y}_{\mathrm{Q,r}}(t) := \mathbf{y}_{\mathrm{Q,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) + \mathbf{y}_{\mathrm{Q,r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t) + \mathbf{y}_{\mathrm{Q,r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}(t) + \mathbf{y}_{\mathrm{Q,r},\mathbf{Z}_0\mathbf{Z}_0}(t).$$

We now aim to derive an error bound for the presented BT method. Therefore, we make use of the following decomposition

$$\begin{aligned}\|\mathbf{y}_{\mathrm{Q}} - \mathbf{y}_{\mathrm{Q,r}}\|_{L_\infty} \leq{}& \|\mathbf{y}_{\mathrm{Q},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{Q,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{Q,r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\|_{L_\infty} \\ &+ \|\mathbf{y}_{\mathrm{Q},\boldsymbol{\mathcal{B}}\mathbf{Z}_0} - \mathbf{y}_{\mathrm{Q,r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0\mathbf{Z}_0} - \mathbf{y}_{\mathrm{Q,r},\mathbf{Z}_0\mathbf{Z}_0}\|_{L_\infty}. \end{aligned} \tag{4.10}$$

The authors in [20, Equations (22), (23)] derive an error bound for homogeneous first-order ODE systems with quadratic output equations that is applicable for the error

$\|\mathbf{y}_{\mathrm{Q},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{Q,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\|_{L_\infty}$. To evaluate the remaining output errors, we apply the same methodology, which we demonstrate for the error component $\|\mathbf{y}_{\mathrm{Q},\mathbf{z}_0\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{Q,r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}\|_{L_\infty}$. Therefore, we define the mappings

$$\begin{aligned}
\mathbf{h}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2) &:= \operatorname{vec}\Big(\mathbf{Z}_0^{\mathrm{T}} e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_1}\boldsymbol{\mathcal{M}} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_2}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\Big),\\
\widehat{\mathbf{h}}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2) &:= \operatorname{vec}\Big(\mathbf{Z}_{0,\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}^{\mathrm{T}} e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}^{-\mathrm{T}}t_1}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}} e^{\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}^{-1}\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}t_2}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}^{-1}\boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}\Big).
\end{aligned} \tag{4.11}$$

Using these mappings, the outputs of system (3.35) and (4.6) can be rewritten as

$$\begin{aligned}
\mathbf{y}_{\mathrm{Q},\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t) &= \int_0^t \mathbf{h}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau)^{\mathrm{T}}(\mathbf{u}(\tau)\otimes\zeta_0)\mathrm{d}\tau,\\
\mathbf{y}_{\mathrm{Q,r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t) &= \int_0^t \widehat{\mathbf{h}}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau)^{\mathrm{T}}(\mathbf{u}(\tau)\otimes\zeta_0)\mathrm{d}\tau.
\end{aligned}$$

Using these representations of $\mathbf{y}_{\mathrm{Q},\mathbf{z}_0\boldsymbol{\mathcal{B}}}$ and $\mathbf{y}_{\mathrm{Q,r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}$, the following lemma provides an upper bound of the respective L_∞-error.

Lemma 4.1:
Consider the asymptotically stable system (3.35) with initial conditions as defined in (3.4), the reduced system (4.6) with matrices (4.9), and the mappings $\mathbf{h}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}$, $\widehat{\mathbf{h}}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}$ as defined in (4.11). Then, the following inequality holds

$$\|\mathbf{y}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}\|_{L_\infty} \leq \left(\int_0^\infty\int_0^\infty \|\mathbf{h}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\|_2^2\mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \|\mathbf{u}\otimes\zeta_0\|_{L_2}. \quad \Diamond$$

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

$$\big|\mathbf{y}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t) - \mathbf{y}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t)\big| = \left|\int_0^t \Big(\mathbf{h}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau) - \widehat{\mathbf{h}}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau)\Big)^{\mathrm{T}}(\mathbf{u}(\tau)\otimes\zeta_0)\mathrm{d}\tau\right|.$$

Applying the Cauchy-Schwarz inequality multiple times yields

$$\begin{aligned}
\big|\mathbf{y}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t) - \mathbf{y}_{\mathrm{r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t)\big| &\leq \int_0^t \left\|\Big(\mathbf{h}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau) - \widehat{\mathbf{h}}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau)\Big)(\mathbf{u}(\tau)\otimes\zeta_0)\right\|\mathrm{d}\tau\\
&\leq \int_0^t \left\|\mathbf{h}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau) - \widehat{\mathbf{h}}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau)\right\|_2 \|(\mathbf{u}(\tau)\otimes\zeta_0)\|_2\mathrm{d}\tau\\
&\leq \left(\int_0^t \left|\mathbf{h}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau) - \widehat{\mathbf{h}}_{\mathbf{z}_0\boldsymbol{\mathcal{B}}}(t,t-\tau)\right|_2^2\mathrm{d}\tau\right)^{\frac{1}{2}}\\
&\qquad\qquad\cdot\left(\int_0^t \|(\mathbf{u}(\tau)\otimes\zeta_0)\|_2^2\mathrm{d}\tau\right)^{\frac{1}{2}}.
\end{aligned}$$

Since we only consider nonnegative values in the integral, the values on the right-hand side of the bound increase for larger values of t and can be bounded by choosing $t = \infty$, which leads to the following L_∞-norm bound of $\mathbf{y}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$:

$$\begin{aligned}\|\mathbf{y}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\|_{L_\infty} &\leq \left(\int_0^\infty \int_0^\infty \left\|\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\right\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \\ &\qquad \cdot \left(\int_0^\infty \|(\mathbf{u}(\tau)\otimes\zeta_0)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}} \\ &= \left(\int_0^\infty \int_0^\infty \left\|\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\right\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \|\mathbf{u}\otimes\zeta_0\|_{L_2}. \qquad \square\end{aligned}$$

For further consideration, we define the following matrices

$$\begin{aligned}\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{B}},*\circ} &:= \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}_{\mathrm{r},*\circ}\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-\mathrm{T}} e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-\mathrm{T}}t}\mathrm{d}t, \\ \widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0,*\circ} &:= \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\mathbf{Z}_0\mathbf{Z}_{0,\mathrm{r},*\circ} e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-\mathrm{T}}t}\mathrm{d}t,\end{aligned} \tag{4.12}$$

and the reduced Gramians

$$\begin{aligned}\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}},\mathrm{r},*\circ} &:= \int_0^\infty e^{\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-1}\boldsymbol{\mathcal{A}}_{\mathrm{r},*\circ}t}\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-1}\boldsymbol{\mathcal{B}}_{\mathrm{r},*\circ}\boldsymbol{\mathcal{B}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-\mathrm{T}} e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-\mathrm{T}}t}\mathrm{d}t, \\ \boldsymbol{\mathcal{P}}_{\mathbf{Z}_0,\mathrm{r},*\circ} &:= \int_0^\infty e^{\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-1}\boldsymbol{\mathcal{A}}_{\mathrm{r},*\circ}t}\mathbf{Z}_{0,\mathrm{r},*\circ}\mathbf{Z}_{0,\mathrm{r},*\circ}^{\mathrm{T}} e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*\circ}^{-\mathrm{T}}t}\mathrm{d}t\end{aligned} \tag{4.13}$$

for $*, \circ$ equal to '$\boldsymbol{\mathcal{B}}$' or '$\mathbf{Z}_0$'. Since the bound presented in Lemma 4.1 includes the expression

$$\begin{aligned}\int_0^\infty \int_0^\infty \|\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2) - \widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\|_2^2 \mathrm{d}t_1\mathrm{d}t_2 &= \int_0^\infty \int_0^\infty \Big(\|\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\|_2^2 \\ &\quad - 2\langle \mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2), \widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\rangle + \left\|\widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\right\|_2^2\Big)\mathrm{d}t_1\mathrm{d}t_2,\end{aligned}$$

the following lemma is used to determine the different components of the right-hand side of this bound.

Lemma 4.2:
Consider the asymptotically stable system (3.35) with initial conditions as defined in (3.4), the reduced system (4.6) with matrices (4.9), the corresponding controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ as defined in (3.11) and (3.14), respectively, the matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{B}},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0,\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ from (4.12), and the reduced controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0,\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$

from (4.13). The mappings $\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$, $\widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ are as defined in (4.11). Then, the following equations are fulfilled

$$\int_0^\infty\int_0^\infty \|\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\|_2^2\mathrm{d}t_1\mathrm{d}t_2 = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}\boldsymbol{\mathcal{M}})\,, \tag{4.14a}$$

$$\int_0^\infty\int_0^\infty \|\widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\|_2^2\mathrm{d}t_1\mathrm{d}t_2 = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0,\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}})\,, \tag{4.14b}$$

$$\int_0^\infty\int_0^\infty \langle\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2),\widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\rangle\mathrm{d}t_1\mathrm{d}t_2 = \mathrm{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}^{\mathrm{T}}_{\boldsymbol{\mathcal{B}},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}\widetilde{\mathbf{P}}_{\mathbf{Z}_0,\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\Big)\,. \tag{4.14c}$$

$\lozenge$

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

$$\begin{aligned}
&\int_0^\infty\int_0^\infty \|\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\|_2^2\mathrm{d}t_1\mathrm{d}t_2\\
&\qquad= \int_0^\infty\int_0^\infty \mathrm{tr}\Big(\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_2}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_1}\mathbf{Z}_0\mathbf{Z}_0^{\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_1}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_2}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\Big)\,\mathrm{d}t_1\mathrm{d}t_2\\
&\qquad= \int_0^\infty \mathrm{tr}\Big(\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_2}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_2}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\Big)\,\mathrm{d}t_2\\
&\qquad= \int_0^\infty \mathrm{tr}\Big(e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_2}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_2}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}\boldsymbol{\mathcal{M}}\Big)\,\mathrm{d}t_2\\
&\qquad= \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}\boldsymbol{\mathcal{M}})\,,
\end{aligned}$$

what proves (4.14a) while (4.14b) is proven analogously. To show that the remaining equation in (4.14c) is satisfied, we make use of the property $\langle\mathrm{vec}(\mathbf{X}),\mathrm{vec}(\mathbf{Y})\rangle = \mathrm{tr}\big(\mathbf{X}^{\mathrm{T}}\mathbf{Y}\big)$ and obtain

$$\begin{aligned}
&\int_0^\infty\int_0^\infty \langle\mathbf{h}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2),\widehat{\mathbf{h}}_{\mathbf{Z}_0\boldsymbol{\mathcal{B}}}(t_1,t_2)\rangle\mathrm{d}t_1\mathrm{d}t_2\\
&\qquad= \int_0^\infty\int_0^\infty \mathrm{tr}\Big(\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_2}\boldsymbol{\mathcal{M}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t_1}\mathbf{Z}_0\\
&\qquad\qquad\qquad\cdot\mathbf{Z}^{\mathrm{T}}_{0,\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}t_1}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}e^{\boldsymbol{\mathcal{E}}^{-1}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}t_2}\boldsymbol{\mathcal{E}}^{-1}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\Big)\mathrm{d}t_1\mathrm{d}t_2\\
&\qquad= \int_0^\infty\int_0^\infty \mathrm{tr}\Big(\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t_2}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0,\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}e^{\boldsymbol{\mathcal{E}}^{-1}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}t_2}\boldsymbol{\mathcal{E}}^{-1}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\Big)\,\mathrm{d}t_1\mathrm{d}t_2\\
&\qquad= \mathrm{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}^{\mathrm{T}}_{\boldsymbol{\mathcal{B}},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0,\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\mathbf{M}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\Big)\,.
\end{aligned}$$

$\square$

From Lemma 4.1 and Lemma 4.2, we derive the following theorem, which provides a bound of the L_∞-error $\|\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{Q},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\|_{L_\infty}$.

Theorem 4.3:
Consider the asymptotically stable system (3.35) with initial conditions as defined in (3.4), the reduced system (4.6) with matrices (4.9), the corresponding controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$, $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ as defined in (3.11) and (3.14), respectively, the matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{B}},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0,\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ from (4.12), and the reduced controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0,\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ from (4.13). The error between the output $\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ and the reduced output $\mathbf{y}_{\mathrm{Q},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ satisfies the following bound

$$\begin{aligned}\|\mathbf{y}_{\mathrm{Q},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}-\mathbf{y}_{\mathrm{Q},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\|_{L_\infty}^2 \leq \Big(&\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}\boldsymbol{\mathcal{M}})-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{B}},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0,\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\Big)\\ &+\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0,\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}_{\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}})\Big)\|\mathbf{u}\otimes\zeta_0\|_{L_2}^2. \qquad (4.15)\end{aligned}$$

◊

We apply this error bound to all four error components in (4.10) to obtain an overall error bound.

Corollary 4.4:
Consider the asymptotically stable system (3.5) with initial conditions as defined in (3.4), the reduced subsystems (4.5), (4.6), (4.7), and (4.8) with matrices (4.9), the corresponding controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$, $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ as defined in (3.11) and (3.14), respectively, the matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{B}},*\circ}$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0,*\circ}$ from (4.12), and the reduced controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}},\mathrm{r},\mathbf{Z}_0\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0,\mathrm{r},*\circ}$ from (4.13), for $*,\circ$ equal to '$\boldsymbol{\mathcal{B}}$' or '$\mathbf{Z}_0$'. Then, the error between the output $\mathbf{y}_{\mathrm{Q}}$ and the reduced output $\mathbf{y}_{\mathrm{Q},\mathrm{r}}$ satisfies the following bound

$$\begin{aligned}\|\mathbf{y}_{\mathrm{Q}}-\mathbf{y}_{\mathrm{Q},\mathrm{r}}\|_{L_\infty} &\leq \sum_{*,\circ\in\{\text{'}\boldsymbol{\mathcal{B}}\text{','}\mathbf{Z}_0\text{'}\}}\|\mathbf{y}_{\mathrm{Q},*\circ}-\mathbf{y}_{\mathrm{Q},\mathrm{r},*\circ}\|_{L_\infty}\\ &\leq \sum_{*,\circ\in\{\text{'}\boldsymbol{\mathcal{B}}\text{','}\mathbf{Z}_0\text{'}\}}\Big(\operatorname{tr}(\boldsymbol{\mathcal{P}}_*\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_\circ\boldsymbol{\mathcal{M}})-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\circ,*\circ}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{*,*\circ}\boldsymbol{\mathcal{M}}_{\mathrm{r},*\circ}\Big) \qquad (4.16)\\ &\qquad+\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\circ,\mathrm{r},*\circ}\boldsymbol{\mathcal{M}}_{\mathrm{r},*\circ}\boldsymbol{\mathcal{P}}_{*,\mathrm{r},*\circ}\boldsymbol{\mathcal{M}}_{\mathrm{r},*\circ})\Big)\|\mathbf{u}_*\otimes\mathbf{u}_\circ\|_{L_2}^2,\end{aligned}$$

where $\mathbf{u}_{\boldsymbol{\mathcal{B}}} := \mathbf{u}$ and $\mathbf{u}_{\mathbf{Z}_0} := \zeta_0$. ◊

4.1.1.2 Extended-input approach for inhomogeneous first-order ODE systems

In this paragraph, we aim to derive reduced surrogate models of the inhomogeneous first-order systems with a linear output equation (3.5) and with a quadratic one (3.31), both of significantly smaller dimension $R \ll N$. In contrast to the previous paragraph, we intend to find one reduced system

$$\begin{aligned}\boldsymbol{\mathcal{E}}_{\mathrm{r}}\dot{\mathbf{z}}_{\mathrm{r}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t)+\boldsymbol{\mathcal{B}}_{\mathrm{r}}\widetilde{\mathbf{u}}(t), \qquad \mathbf{z}_{\mathrm{r}}(0)=\mathbf{Z}_{0,\mathrm{r}}\zeta_0,\\ \mathbf{y}_{\mathrm{L},\mathrm{r}}(t) &= \boldsymbol{\mathcal{C}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t),\end{aligned} \qquad (4.17)$$

which approximates the original system (3.5) and one reduced system

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r}}\dot{\mathbf{z}}_{\mathrm{r}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t) + \boldsymbol{\mathcal{B}}_{\mathrm{r}}\widetilde{\mathbf{u}}(t), \qquad \mathbf{z}_{\mathrm{r}}(0) = \mathbf{Z}_{0,\mathrm{r}}\zeta_0, \\ \mathbf{y}_{\mathrm{Q,r}}(t) &= \mathbf{z}_{\mathrm{r}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t), \end{aligned} \tag{4.18}$$

which approximates the system (3.31). The surrogate models include the reduced matrices

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r}} &:= \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}, & \boldsymbol{\mathcal{A}}_{\mathrm{r}} &:= \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}, & \boldsymbol{\mathcal{B}}_{\mathrm{r}} &:= \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}, \\ \mathbf{Z}_{0,\mathrm{r}} &:= \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\mathbf{Z}_0, & \boldsymbol{\mathcal{C}}_{\mathrm{r}} &:= \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}, & \boldsymbol{\mathcal{M}}_{\mathrm{r}} &:= \boldsymbol{\mathcal{T}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{T}}_{\mathrm{r}}, \end{aligned} \tag{4.19}$$

generated using the projecting bases $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$, $\boldsymbol{\mathcal{T}}_{\mathrm{r}} \in \mathbb{R}^{N\times R}$ that satisfy the Petrov-Galerkin conditions (2.30) and (2.31).

To derive such surrogate systems, we consider the respective homogeneous systems from (3.27) and (3.45). The inhomogeneous and the homogeneous systems (3.5) and (3.27), such as (3.31) and (3.45) have the same input- and initial condition-to-output behavior in the frequency-domain. Hence, homogeneous systems are used to identify the states that are hard to reach and to observe. To determine these states, we evaluate the system energies summarized in Table 3.6 and Table 3.7. The controllability energies from (3.30) show that states corresponding to small eigenvalues of the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{W}}$ from (3.29) have only negligible influence on the system dynamics, and hence, are truncated in the following. Also, we investigate the observability energies in (3.22) for systems (3.27) with a linear output equation, as well as (3.48) for systems (3.31) with a quadratic output equation. It follows that states corresponding to small eigenvalues of the observability Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,W}}$ from (3.17) and (3.47), respectively, are hard to observe and, hence, truncated within the BT method. The controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{W}}$ and the observability Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{L,W}}$, $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,W}}$ are in general not equal. Therefore, we transform the system so that the controllability and observability Gramian coincide before truncating the negligible states.

We apply BT as introduced in Algorithm 1 to balance the system and truncate the states spanning the least dominant subspaces. For that, we assume that $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{S}}$ are Cholesky factors (or, if available, low-rank factors) of the Gramians of the homogenous systems system (3.27) or (3.45), i.e., $\boldsymbol{\mathcal{P}}_{\mathrm{W}} \approx \boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}^{\mathrm{T}}$ and $\boldsymbol{\mathcal{Q}} \approx \boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^{\mathrm{T}}$, where $\boldsymbol{\mathcal{Q}}$ represents either $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ or $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,W}}$. We compute the singular value decomposition

$$\boldsymbol{\mathcal{S}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_1 & 0 \\ 0 & \boldsymbol{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^{\mathrm{T}} \\ \mathbf{V}_2^{\mathrm{T}} \end{bmatrix}. \tag{4.20}$$

where the matrix $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_N)$ contains the so called Hankel eigenvalues. The remaining step is to truncate states corresponding to small eigenvalues from $\boldsymbol{\Sigma}$. For that, we define the projecting matrices

$$\boldsymbol{\mathcal{V}}_{\mathrm{r}} = \boldsymbol{\mathcal{S}}\mathbf{U}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}, \qquad \boldsymbol{\mathcal{T}}_{\mathrm{r}} = \boldsymbol{\mathcal{R}}\mathbf{V}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}} \tag{4.21}$$

Algorithm 9 BT method for the first-order ODE systems (3.5) and (3.31) with a linear or quadratic output using the extended-input approach.

```
Require: The original system (3.5) or (3.31), reduced dimension $R$.
Ensure: The reduced system (4.17) or (4.18).
1: Compute factors of the Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{w}} \approx \boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}^{\mathrm{T}}$ and $\boldsymbol{\mathcal{Q}} \approx \boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^{\mathrm{T}}$, with $\boldsymbol{\mathcal{Q}}$ equal to $\boldsymbol{\mathcal{Q}}_{\mathrm{L}}$ or
   $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,w}}$ from (3.29), (3.17), and (3.47).
2: Perform the SVD of $\boldsymbol{\mathcal{S}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}}$, and decompose as
   $$\boldsymbol{\mathcal{S}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}} = \begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_1 & \\ & \boldsymbol{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_1 & \mathbf{V}_2\end{bmatrix}^{\mathrm{T}},$$
   with $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{R \times R}$.
3: Construct the projection matrices
   $$\boldsymbol{\mathcal{V}}_{\mathrm{r}} = \boldsymbol{\mathcal{S}}\mathbf{U}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}, \qquad \boldsymbol{\mathcal{T}}_{\mathrm{r}} = \boldsymbol{\mathcal{R}}\mathbf{V}_1\boldsymbol{\Sigma}_1^{-\frac{1}{2}}.$$
4: Construct reduced matrices (4.19).
```

that balance and truncate the system by projecting the state space onto a space spanned by $\mathbf{U}_1$ and $\mathbf{V}_1$ corresponding to the largest eigenvalues stored in $\boldsymbol{\Sigma}_1$. Multiplying the original system (3.5) or (3.31) by $\boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r}}$ results in the reduced system (4.17) or (4.18), respectively, with the reduced matrices (4.19). This methodology leads to Algorithm 9.

Error bound for systems with a linear output equation For the linear output case, there exists an error bound for the error between the output $\mathbf{y}_{\mathrm{L}}$ of the original system (3.5) and the output $\mathbf{y}_{\mathrm{L,r}}$ of the reduced system (4.17), as shown in [66, Theorem 2], that is

$$\begin{aligned}\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_2} \leq{}& \left(2\sum_{k=R+1}^{N}\sigma_k\right)\|\mathbf{u}\|_{L_2} \\ &+ 3\cdot 2^{-\frac{1}{3}}\left(\|\boldsymbol{\mathcal{S}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\mathbf{Z}_0\|_2 + \|\boldsymbol{\Sigma}_1\boldsymbol{\mathcal{A}}_{\mathrm{r}}\mathbf{Z}_{0,\mathrm{r}}\|_2\right)^{\frac{1}{3}}\left(2\sum_{k=R+1}^{N}\sigma_k\right)^{\frac{2}{3}}\|\zeta_0\|_2.\end{aligned}$$

Error bound for systems with a quadratic output equation For the case of quadratic output equations, we again apply the bound from [20, Equations (22), (23)], which was already used to derive the bound in (4.16). However, in the extended-input case, we set $\boldsymbol{\mathcal{V}}_{\mathrm{r},*\circ} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}$, $\boldsymbol{\mathcal{T}}_{\mathrm{r},*\circ} = \boldsymbol{\mathcal{T}}_{\mathrm{r}}$, with $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{r}}$ as defined in (4.21) for all $*, \circ \in \{$'$\boldsymbol{\mathcal{B}}$', '$\mathbf{Z}_0$'$\}$. We define the reduced controllability Gramian and the matrix

$$\begin{aligned} \boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}},\mathrm{r}} &= \int_0^\infty e^{\boldsymbol{\mathcal{E}}_\mathrm{r}^{-1}\boldsymbol{\mathcal{A}}_\mathrm{r}t}\boldsymbol{\mathcal{E}}_\mathrm{r}^{-1}\boldsymbol{\mathcal{W}}_\mathrm{r}\boldsymbol{\mathcal{W}}_\mathrm{r}^\mathrm{T}\boldsymbol{\mathcal{E}}_\mathrm{r}^{-\mathrm{T}}e^{(\boldsymbol{\mathcal{E}}_\mathrm{r}^{-1}\boldsymbol{\mathcal{A}}_\mathrm{r})^\mathrm{T}t}\mathrm{d}t, \\ \widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{W}}} &= \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{W}}_\mathrm{r}^\mathrm{T}\boldsymbol{\mathcal{E}}_\mathrm{r}^{-\mathrm{T}}e^{(\boldsymbol{\mathcal{E}}_\mathrm{r}^{-1}\boldsymbol{\mathcal{A}}_\mathrm{r})^\mathrm{T}t}\mathrm{d}t. \end{aligned} \tag{4.22}$$

This leads to the following error bound.

Theorem 4.5:
Consider the asymptotically stable system (3.31) with initial conditions as defined in (3.4) and the reduced system (4.18). Also consider the respective controllability Gramian $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}}}$ as defined in (3.29), the reduced controllability Gramian $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}},\mathrm{r}}$ as defined in (4.22), and the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{W}}}$ from (4.22). Then, the respective output error is bounded by

$$\begin{aligned} \|\mathbf{y}_\mathrm{Q} - \mathbf{y}_{\mathrm{Q},\mathrm{r}}\|_{L_\infty} \leq \Big(&\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}}}\boldsymbol{\mathcal{M}}) - 2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{W}}}^\mathrm{T}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{W}}}\boldsymbol{\mathcal{M}}_\mathrm{r}\Big) \\ &+ \operatorname{tr}(\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}},\mathrm{r}}\boldsymbol{\mathcal{M}}_\mathrm{r}\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}},\mathrm{r}}\boldsymbol{\mathcal{M}}_\mathrm{r})\Big)\left(\|\mathbf{u}\|_{L_2} + \|\zeta_0\|_2\right)^2. \end{aligned} \tag{4.23}$$

◊

Proof. We apply the bound from (4.16) to obtain

$$\begin{aligned} \|\mathbf{y}_\mathrm{Q} - \mathbf{y}_{\mathrm{Q},\mathrm{r}}\|_{L_\infty} &\leq \sum_{*,\circ\in\{\text{'}\boldsymbol{\mathcal{B}}\text{','}\mathbf{Z}_0\text{'}\}} \|\mathbf{y}_{\mathrm{Q},*\circ} - \mathbf{y}_{\mathrm{Q},\mathrm{r},*\circ}\|_{L_\infty} \\ &\leq \sum_{*,\circ\in\{\text{'}\boldsymbol{\mathcal{B}}\text{','}\mathbf{Z}_0\text{'}\}} \Big(\operatorname{tr}(\boldsymbol{\mathcal{P}}_*\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_\circ\boldsymbol{\mathcal{M}}) - 2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_\circ^\mathrm{T}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_*\boldsymbol{\mathcal{M}}_\mathrm{r}\Big) \\ &\qquad + \operatorname{tr}(\boldsymbol{\mathcal{P}}_{\circ,\mathrm{r}}\boldsymbol{\mathcal{M}}_\mathrm{r}\boldsymbol{\mathcal{P}}_{*,\mathrm{r}}\boldsymbol{\mathcal{M}}_\mathrm{r})\Big)\|\mathbf{u}_* \otimes \mathbf{u}_\circ\|_{L_2}^2 \end{aligned} \tag{4.24}$$

for the reduced matrices from (4.9), which coincide with those in (4.19), where $\mathbf{u}_{\boldsymbol{\mathcal{B}}} := \mathbf{u}$, $\mathbf{u}_{\mathbf{Z}_0} := \zeta_0$. The respective controllability Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$ are as defined in (3.11) and (3.14), the matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{B}}}$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0}$ are defined in (4.12), and the reduced Gramians $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}},\mathrm{r}}$ and $\boldsymbol{\mathcal{P}}_{\mathbf{Z}_0,\mathrm{r}}$ are from (4.13). Since $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}}} = \boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}} + \boldsymbol{\mathcal{P}}_{\mathbf{Z}_0}$, $\widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{W}}} = \widetilde{\boldsymbol{\mathcal{P}}}_{\boldsymbol{\mathcal{B}}} + \widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{Z}_0}$, and $\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{W}},\mathrm{r}} = \boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}},\mathrm{r}} + \boldsymbol{\mathcal{P}}_{\mathbf{Z}_0,\mathrm{r}}$ holds, the right-hand side of (4.24) is bounded by the one in (4.23), which proves the statement. □

4.1.2 IRKA for inhomogeneous first-order ODE systems

In this section, we briefly describe the application of the IRKA method presented in [60] and in Section 2.2.2 to first-order ODE systems with inhomogeneous initial conditions and linear output equations. However, we will not describe this method in detail since this is not the main contribution of this work. Also, since there is no IRKA approach for systems with quadratic output equations, we only consider the system (3.5) with a linear output equation. For this purpose, we use the equivalent systems represented in Section 3.1, again distinguishing between different approaches to embed the initial conditions, which are described in the following subsections.

4.1.2.1 Multi-system approach for inhomogeneous first-order ODE systems

To derive reduced surrogate models approximating the input-to-output behavior of the original system (3.5) using the IRKA method, we decompose this system into two subsystems that are (3.8) and (3.9), as described in [15]. We aim to derive two respective surrogate models (4.1) and (4.2), approximating the behavior of the two subsystems. Since the transfer functions (3.8) and (3.9) of the two subsystems are of the same structure as the transfer function of system (2.1) with homogeneous initial conditions, the IRKA method introduced in Algorithm 4 can be applied to both subsystems individually, to derive the surrogate systems. To apply the IRKA method to subsystem (3.9), we would replace the input matrix $\boldsymbol{\mathcal{B}}$ by $\boldsymbol{\mathcal{E}}\mathbf{Z}_0$.

4.1.2.2 Extended-input approach for inhomogeneous first-order ODE systems

As described in [66], the inhomogeneous system (3.5) has the same transfer function as the homogeneous system (3.27) with $\boldsymbol{\mathcal{W}}$ as defined in (3.24). Hence, we can apply the IRKA method presented in Algorithm 4 to the system (3.27) and to derive bases $\boldsymbol{\mathcal{V}}_\mathrm{r}$ and $\boldsymbol{\mathcal{T}}_\mathrm{r}$ that define the matrices (4.19) of a reduced surrogate model (4.18) that approximates the input-to-output behavior of the original one.

4.2 Model order reduction for inhomogeneous first-order DAE systems

In this section, we consider dynamical systems with differential-algebraic equations as state equations as defined in (3.54) and (3.100) where the matrix $\boldsymbol{\mathcal{E}}$ is singular. We assume that the matrix pencil $(\mathbf{A}, \mathbf{E})$ is regular, i.e., $\det(\lambda\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})$ is not a zero polynomial. We seek to employ model reduction techniques that allow us to construct a low-dimensional model that closely resembles the dynamic behaviors of the original high-fidelity model.

The BT and IRKA methods presented in the previous Section 4.1 derived for ODE systems treat the case where $\boldsymbol{\mathcal{E}}$ is nonsingular. Therefore, they are not directly applicable to the DAE case. In this section, we extend the methods presented above to the case of inhomogeneous DAE systems (3.54) and (3.100) with a linear and a quadratic output equation, respectively. To that end, we use the new Gramians presented in Section 3.2.1

and Section 3.2.2 that allow us to characterize the controllability and observability behavior.

We derive in Section 4.2.1 the BT method for inhomogeneous DAE systems with linear and quadratic output equations and derive respective error bounds to evaluate the quality of the resulting reduced surrogate systems. Afterwards, in Section 4.2.2, we briefly describe the application of the IRKA approach for inhomogeneous DAE systems with linear output equations.

4.2.1 BT for inhomogeneous first-order DAE systems

To derive BT methods that reduce the DAE systems from (3.54) and (3.100), we identify the least dominant controllability and observability subspaces truncated in this method, as their effect on the system dynamics is negligible. To determine these subspaces, we use respective system energies determined by Gramians tailored for these systems. Based on this, we propose a balancing scheme to determine projection matrices, leading to the construction of reduced-order models.

We again consider the multi-system and extended-input approaches derived in Section 3.2. Both approaches consider the proper and the improper components separately, while the multi-system representation also derives subsystems corresponding to the input and subsystems corresponding to the initial condition.

4.2.1.1 Multi-system approach for inhomogeneous first-order DAE systems

In this paragraph, we derive a reduced system representation of the original system (3.54) with a linear output equation using the multi-system approach introduced in Section 3.2.1.1. For systems (3.100) with quadratic output equations, the multi-system approach leads to too many subsystems as explained in Section 3.2.2.1 so that we apply for systems of this structure the extended-input approach presented later in Section 4.2.1.2.

We decompose the inhomogeneous system (3.54) into three subsystems. The first two subsystems (3.58) and (3.59) encode the system dynamics corresponding to the differential state component and the third subsystem (3.60) includes the algebraic state component, as described in Section 3.2.1. Instead of reducing the original system, we apply the BT method introduced in Algorithm 2 to each subsystem to derive three reduced surrogate models approximating the input- and initial condition-to-output behaviors. The first subsystem (3.58) corresponding to the input $\mathfrak{B}\mathbf{u}(t)$ is approximated by a surrogate model

$$\begin{aligned} \dot{\widehat{\mathbf{z}}}_1(t) &= \widehat{\mathbf{A}}_{1,\mathfrak{B}}\widehat{\mathbf{z}}_1(t) + \widehat{\mathbf{B}}_{1,\mathfrak{B}}\mathbf{u}(t), \qquad \widehat{\mathbf{z}}_1(0) = 0, \\ \mathbf{y}_{\mathrm{L,p,r},\mathfrak{B}}(t) &= \widehat{\mathbf{C}}_{1,\mathfrak{B}}\widehat{\mathbf{z}}_1(t), \end{aligned} \tag{4.25}$$

while the initial condition-to-output behavior of the second subsystem (3.59) is approximated by the surrogate model

$$\begin{aligned} \dot{\widehat{\mathbf{z}}}_1(t) &= \widehat{\mathbf{A}}_{1,\mathbf{Z}_0}\widehat{\mathbf{z}}_1(t), \qquad \widehat{\mathbf{z}}_1(0) = \widehat{\mathbf{Z}}_{0,\mathrm{r}}\zeta_0, \\ \mathbf{y}_{\mathrm{L,p,r},\mathbf{Z}_0}(t) &= \widehat{\mathbf{C}}_{1,\mathbf{Z}_0}\widehat{\mathbf{z}}_1(t), \end{aligned} \tag{4.26}$$

with reduced matrices

$$\widehat{\mathbf{A}}_{1,*} = \boldsymbol{\mathcal{V}}_{\mathrm{p},*}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{p},*}, \qquad \widehat{\mathbf{B}}_{1,\boldsymbol{\mathcal{B}}} = \boldsymbol{\mathcal{V}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}, \qquad \widehat{\mathbf{Z}}_{0,\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{p},\mathbf{Z}_0}^{\mathrm{T}}\mathbf{Z}_0, \qquad \widehat{\mathbf{C}}_{1,*} = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{T}}_{\mathrm{p},*} \tag{4.27}$$

where the subscript $*$ represents either '$\boldsymbol{\mathcal{B}}$' or '$\mathbf{Z}_0$'. The projecting bases $\boldsymbol{\mathcal{V}}_{\mathrm{p},*}, \boldsymbol{\mathcal{T}}_{\mathrm{p},*} \in \mathbb{R}^{N\times R_*}$ are assumed to project on the deflating subspaces of the matrix pencil $(\mathbf{A}, \mathbf{E})$ corresponding to the finite eigenvalues of the matrix pencil. Also, we assume that the dimensions of the reduced systems are significantly smaller than the original system dimension, i.e., $R_* \ll N$.

Moreover, we derive a surrogate model corresponding to the third system (3.60), that encodes the algebraic system dynamics

$$\begin{aligned} \widehat{\mathbf{E}}_2\dot{\widehat{\mathbf{z}}}_2(t) &= \widehat{\mathbf{z}}_2(t) + \widehat{\mathbf{B}}_2\mathbf{u}(t), \qquad \widehat{\mathbf{z}}_2(0) = \widehat{\mathbf{z}}_{2,0}, \\ \mathbf{y}_{\mathrm{L,i,r}}(t) &= \widehat{\mathbf{C}}_2\widehat{\mathbf{z}}_2(t) \end{aligned} \tag{4.28}$$

with a nilpotent matrix $\widehat{\mathbf{E}}_2 \in \mathbb{R}^{R_\mathrm{i}\times R_\mathrm{i}}$, an input matrix $\widehat{\mathbf{B}}_2 \in \mathbb{R}^{R_\mathrm{i}\times m}$ and an output matrix $\widehat{\mathbf{C}}_2 \in \mathbb{R}^{p\times R_\mathrm{i}}$. Note that this reduced model is supposed to have the same input-to-output behavior as the subsystem (3.60) encoding the algebraic component of the system dynamic. The reduced system (4.28) is generated using projecting matrices $\boldsymbol{\mathcal{V}}_{\mathrm{i,r}}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{i,r}} \in \mathbb{R}^{N\times N_{\mathrm{inf}}}$ as

$$\widehat{\mathbf{E}}_2 = \boldsymbol{\mathcal{V}}_{\mathrm{i,r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{i,r}}, \qquad \widehat{\mathbf{B}}_2 = \boldsymbol{\mathcal{V}}_{\mathrm{i,r}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}, \qquad \widehat{\mathbf{C}}_2 = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{T}}_{\mathrm{i,r}}, \qquad \widehat{\mathbf{z}}_{2,0} = \boldsymbol{\mathcal{V}}_{\mathrm{i,r}}^{\mathrm{T}}\mathbf{z}_0. \tag{4.29}$$

The three reduced subsystems approximate the overall input- and initial condition-to-output behavior as

$$\mathbf{y}_{\mathrm{L}}(t) \approx \mathbf{y}_{\mathrm{L,p,r},\boldsymbol{\mathcal{B}}}(t) + \mathbf{y}_{\mathrm{L,p,r},\mathbf{Z}_0}(t) + \mathbf{y}_{\mathrm{L,i,r}}(t).$$

To derive such surrogate systems, we apply Algorithm 2 to the three subsystems (3.58), (3.59), and (3.60). First, we compute Cholesky factors or low-rank factors of the proper controllability and observability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} = \boldsymbol{\mathcal{R}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{R}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}^{\mathrm{T}}$, $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{Z}_0} = \boldsymbol{\mathcal{R}}_{\mathrm{p},\mathbf{Z}_0}\boldsymbol{\mathcal{R}}_{\mathrm{p},\mathbf{Z}_0}^{\mathrm{T}}$, and $\mathbf{Q}_{\mathrm{p,L}} = \boldsymbol{\mathcal{S}}_{\mathrm{p}}^{\mathrm{T}}\boldsymbol{\mathcal{S}}_{\mathrm{p}}$ corresponding to the first two subsystems (3.58) and (3.59) defined in (3.62), (3.66), and (3.74), respectively. Computing the improper controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathbf{Z}_0}$ corresponding to these subsystems results in

$$\boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{\mathcal{B}}} = \boldsymbol{\mathcal{R}}_{\mathrm{i},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{R}}_{\mathrm{i},\boldsymbol{\mathcal{B}}}^{\mathrm{T}} = \sum_{k=0}^{\nu-1}\boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)\mathbf{P}_{\mathrm{l}}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\mathbf{P}_{\mathrm{l}}^{\mathrm{T}}\boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)^{\mathrm{T}} = 0,$$

$$\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathbf{Z}_0} = \boldsymbol{\mathcal{R}}_{\mathrm{i},\mathbf{Z}_0}\boldsymbol{\mathcal{R}}_{\mathrm{i},\mathbf{Z}_0}^{\mathrm{T}} = \sum_{k=0}^{\nu-1}\boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)\mathbf{P}_{\mathrm{l}}\boldsymbol{\mathcal{E}}\mathbf{Z}_0\mathbf{Z}_0^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{P}_{\mathrm{l}}^{\mathrm{T}}\boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)^{\mathrm{T}} = 0$$

since $\mathcal{F}_{\mathbf{N}}\mathbf{P}_{\mathrm{l}} = 0$ for $\mathcal{F}_{\mathbf{N}}$ and $\mathbf{P}_{\mathrm{l}}$ as defined in (2.13) and (2.10), respectively. Therefore, only the proper Gramians are considered when reducing the two subsystems (3.58) and (3.59).

As described in (3.81), (3.82), and (3.85), and later summarized in Table 3.5, states corresponding large eigenvalues of the proper Gramians $\mathcal{P}_{\mathrm{p},\mathcal{B}}$, $\mathcal{P}_{\mathrm{p},\mathbf{z}_0}$, and $\mathcal{Q}_{\mathrm{p,L}}$ span the most dominant controllability and observability subspaces. The states corresponding to small eigenvalues, on the other hand, are hard to reach and observe and are therefore negligible. We truncate these states in the following. To do so, we derive the respective singular value decompositions

$$\begin{aligned}
\mathcal{S}_{\mathrm{p},\mathcal{B}}\mathcal{E}\mathcal{R}_{\mathrm{p},\mathcal{B}} &= \begin{bmatrix}\mathbf{U}_{\mathrm{p},1,\mathcal{B}} & \mathbf{U}_{\mathrm{p},2,\mathcal{B}}\end{bmatrix}\begin{bmatrix}\mathbf{\Sigma}_{1,\mathcal{B}} & \\ & \mathbf{\Sigma}_{2,\mathcal{B}}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{p},1,\mathcal{B}}^{\mathrm{T}} \\ \mathbf{V}_{\mathrm{p},2,\mathcal{B}}^{\mathrm{T}}\end{bmatrix},\\
\mathcal{S}_{\mathrm{p},\mathbf{z}_0}\mathcal{E}\mathcal{R}_{\mathrm{p},\mathbf{z}_0} &= \begin{bmatrix}\mathbf{U}_{\mathrm{p},1,\mathbf{z}_0} & \mathbf{U}_{\mathrm{p},2,\mathbf{z}_0}\end{bmatrix}\begin{bmatrix}\mathbf{\Sigma}_{1,\mathbf{z}_0} & \\ & \mathbf{\Sigma}_{2,\mathbf{z}_0}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{p},1,\mathbf{z}_0}^{\mathrm{T}} \\ \mathbf{V}_{\mathrm{p},2,\mathbf{z}_0}^{\mathrm{T}}\end{bmatrix}.
\end{aligned}$$

Since the improper Gramians corresponding to these two systems are equal to zero, the respective BT projection matrices corresponding to system (3.58) and (3.59) as derived in (2.40) have the structure

$$\begin{aligned}
\mathcal{V}_{\mathrm{r},\mathcal{B}} &= \begin{bmatrix}\mathcal{S}_{\mathrm{p},\mathcal{B}}^{\mathrm{T}}\mathbf{U}_{\mathrm{p},1,\mathcal{B}}\mathbf{\Sigma}_{1,\mathcal{B}}^{-\frac{1}{2}}\end{bmatrix}, & \mathcal{T}_{\mathrm{r},\mathcal{B}} &= \begin{bmatrix}\mathcal{R}_{\mathrm{p},\mathcal{B}}\mathbf{V}_{\mathrm{p},1,\mathcal{B}}\mathbf{\Sigma}_{1,\mathcal{B}}^{-\frac{1}{2}}\end{bmatrix},\\
\mathcal{V}_{\mathrm{r},\mathbf{z}_0} &= \begin{bmatrix}\mathcal{S}_{\mathrm{p},\mathbf{z}_0}^{\mathrm{T}}\mathbf{U}_{\mathrm{p},1,\mathbf{z}_0}\mathbf{\Sigma}_{1,\mathbf{z}_0}^{-\frac{1}{2}}\end{bmatrix}, & \mathcal{T}_{\mathrm{r},\mathbf{z}_0} &= \begin{bmatrix}\mathcal{R}_{\mathrm{p},\mathbf{z}_0}\mathbf{V}_{\mathrm{p},1,\mathbf{z}_0}\mathbf{\Sigma}_{1,\mathbf{z}_0}^{-\frac{1}{2}}\end{bmatrix}.
\end{aligned}$$

For the third subsystem (3.60), we derive the improper controllability and observability Gramians $\mathcal{P}_{\mathrm{i}} = \mathcal{R}_{\mathrm{i}}\mathcal{R}_{\mathrm{i}}^{\mathrm{T}}$ and $\mathcal{Q}_{\mathrm{L,i}} = \mathcal{S}_{\mathrm{i}}^{\mathrm{T}}\mathcal{S}_{\mathrm{i}}$ as defined in (3.70) and (3.78), respectively. Note that the proper controllability Gramian is

$$\mathcal{P}_{\mathrm{p}} = \mathcal{R}_{\mathrm{p}}\mathcal{R}_{\mathrm{p}}^{\mathrm{T}} = \int_0^\infty \mathcal{F}_{\mathbf{J}}(t)(\mathbf{I}-\mathbf{P}_{\mathrm{l}})\mathcal{B}\mathcal{B}^{\mathrm{T}}(\mathbf{I}-\mathbf{P}_{\mathrm{l}})^{\mathrm{T}}\mathcal{F}_{\mathbf{J}}(t)^{\mathrm{T}}\mathrm{d}t = 0$$

as $\mathcal{F}_{\mathbf{J}}(t)(\mathbf{I}-\mathbf{P}_{\mathrm{l}}) = 0$ with $\mathcal{F}_{\mathbf{J}}$ and $\mathbf{P}_{\mathrm{l}}$ from (2.13) and (2.10). We derive the singular value decomposition

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

and the resulting projecting matrices

$$\mathcal{V}_{\mathrm{i,r}} = \begin{bmatrix}\mathcal{S}_{\mathrm{i}}^{\mathrm{T}}\mathbf{U}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}\end{bmatrix}, \qquad \mathcal{T}_{\mathrm{i,r}} = \begin{bmatrix}\mathcal{R}_{\mathrm{i}}\mathbf{V}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}\end{bmatrix}.$$

Computing and applying the projecting matrices to the three different subsystems leads to Algorithm 10.

Algorithm 10 BT method for the first-order DAE system (2.8) with a linear output using the multi-system approach.

```
Require: The original system (3.54) and the reduced orders $R_{\mathrm{p},\mathcal{B}}$, $R_{\mathrm{p},\mathbf{z}_0}$, and $R_{\mathrm{i}}$.
Ensure: The reduced systems (4.25), (4.26), and (4.28).
1: Compute factors of the Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B}} = \boldsymbol{\mathcal{R}}_{\mathrm{p},\mathcal{B}}\boldsymbol{\mathcal{R}}_{\mathrm{p},\mathcal{B}}^{\mathrm{T}}$, $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0} = \boldsymbol{\mathcal{R}}_{\mathrm{p},\mathbf{z}_0}\boldsymbol{\mathcal{R}}_{\mathrm{p},\mathbf{z}_0}^{\mathrm{T}}$, and $\mathbf{Q}_{\mathrm{p,L}} = \boldsymbol{\mathcal{S}}_{\mathrm{p},\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{S}}_{\mathrm{p},\mathcal{B}}$ corresponding to the first two subsystems (3.58) and (3.59) and $\boldsymbol{\mathcal{P}}_{\mathrm{i}} = \boldsymbol{\mathcal{R}}_{\mathrm{i}}\boldsymbol{\mathcal{R}}_{\mathrm{i}}^{\mathrm{T}}$, and $\boldsymbol{\mathcal{Q}}_{\mathrm{i,L}} = \boldsymbol{\mathcal{S}}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{\mathcal{S}}_{\mathrm{i}}$ corresponding to subsystem (3.60).
2: Perform the two SVDs
   $$\begin{aligned}
   \boldsymbol{\mathcal{S}}_{\mathrm{p},\mathcal{B}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}}_{\mathrm{p},\mathcal{B}} &= \begin{bmatrix}\mathbf{U}_{\mathrm{p},1,\mathcal{B}} & \mathbf{U}_{\mathrm{p},2,\mathcal{B}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_{1,\mathcal{B}} & \\ & \boldsymbol{\Sigma}_{2,\mathcal{B}}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{p},1,\mathcal{B}}^{\mathrm{T}} \\ \mathbf{V}_{\mathrm{p},2,\mathcal{B}}^{\mathrm{T}}\end{bmatrix},\\
   \boldsymbol{\mathcal{S}}_{\mathrm{p},\mathbf{z}_0}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{R}}_{\mathrm{p},\mathbf{z}_0} &= \begin{bmatrix}\mathbf{U}_{\mathrm{p},1,\mathbf{z}_0} & \mathbf{U}_{\mathrm{p},2,\mathbf{z}_0}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_{1,\mathbf{z}_0} & \\ & \boldsymbol{\Sigma}_{2,\mathbf{z}_0}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{p},1,\mathbf{z}_0}^{\mathrm{T}} \\ \mathbf{V}_{\mathrm{p},2,\mathbf{z}_0}^{\mathrm{T}}\end{bmatrix},\\
   \boldsymbol{\mathcal{S}}_{\mathrm{i}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{R}}_{\mathrm{i}} &= \begin{bmatrix}\mathbf{U}_{\mathrm{i},1} & \mathbf{U}_{\mathrm{i},2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Theta}_1 & \\ & 0\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{i},1}^{\mathrm{T}} \\ \mathbf{V}_{\mathrm{i},2}^{\mathrm{T}}\end{bmatrix}
   \end{aligned}$$
   with $\boldsymbol{\Sigma}_{1,\mathcal{B}} \in \mathbb{R}^{R_{\mathrm{p},\mathcal{B}}\times R_{\mathrm{p},\mathcal{B}}}$, $\boldsymbol{\Sigma}_{1,\mathbf{z}_0} \in \mathbb{R}^{R_{\mathrm{p},\mathbf{z}_0}\times R_{\mathrm{p},\mathbf{z}_0}}$.
3: Construct the projection matrices
   $$\begin{aligned}
   \boldsymbol{\mathcal{V}}_{\mathrm{p,r},\mathcal{B}} &= \left[\boldsymbol{\mathcal{S}}_{\mathrm{p},\mathcal{B}}^{\mathrm{T}}\mathbf{U}_{\mathrm{p},1,\mathcal{B}}\boldsymbol{\Sigma}_{1,\mathcal{B}}^{-\frac{1}{2}}\right], & \boldsymbol{\mathcal{T}}_{\mathrm{p,r},\mathcal{B}} &= \left[\boldsymbol{\mathcal{R}}_{\mathrm{p},\mathcal{B}}\mathbf{V}_{\mathrm{p},1,\mathcal{B}}\boldsymbol{\Sigma}_{1,\mathcal{B}}^{-\frac{1}{2}}\right],\\
   \boldsymbol{\mathcal{V}}_{\mathrm{p,r},\mathbf{z}_0} &= \left[\boldsymbol{\mathcal{S}}_{\mathrm{p},\mathbf{z}_0}^{\mathrm{T}}\mathbf{U}_{\mathrm{p},1,\mathbf{z}_0}\boldsymbol{\Sigma}_{1,\mathbf{z}_0}^{-\frac{1}{2}}\right], & \boldsymbol{\mathcal{T}}_{\mathrm{p,r},\mathbf{z}_0} &= \left[\boldsymbol{\mathcal{R}}_{\mathrm{p},\mathbf{z}_0}\mathbf{V}_{\mathrm{p},1,\mathbf{z}_0}\boldsymbol{\Sigma}_{1,\mathbf{z}_0}^{-\frac{1}{2}}\right],\\
   \boldsymbol{\mathcal{V}}_{\mathrm{i,r}} &= \left[\boldsymbol{\mathcal{S}}_{\mathrm{i}}^{\mathrm{T}}\mathbf{U}_{\mathrm{i},1}\boldsymbol{\Theta}_1^{-\frac{1}{2}}\right], & \boldsymbol{\mathcal{T}}_{\mathrm{i,r}} &= \left[\boldsymbol{\mathcal{R}}_{\mathrm{i}}\mathbf{V}_{\mathrm{i},1}\boldsymbol{\Theta}_1^{-\frac{1}{2}}\right].
   \end{aligned}$$
4: Determine the reduced system matrices (4.27) and (4.29) of the subsystems (3.58), (3.59), and (3.60).
```

To evaluate the quality of the output approximation using the three surrogate systems, we consider the two proper components separately, while the improper one has an error equal to zero, so that we estimate

$$\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_\infty} \leq \|\mathbf{y}_{\mathrm{L,p},\mathcal{B}} - \mathbf{y}_{\mathrm{L,p,r},\mathcal{B}}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{L,p},\mathbf{z}_0} - \mathbf{y}_{\mathrm{L,p,r},\mathbf{z}_0}\|_{L_\infty}.$$

To derive bounds for the first error component $\|\mathbf{y}_{\mathrm{L,p},\mathcal{B}} - \mathbf{y}_{\mathrm{L,p,r},\mathcal{B}}\|_{L_\infty}$, we define the mappings

$$\mathbf{h}_{\mathrm{p},\mathcal{B}}(t) := \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\boldsymbol{\mathcal{B}}, \qquad \widehat{\mathbf{h}}_{\mathrm{p},\mathcal{B}}(t) := \widehat{\mathbf{C}}_{1,\mathcal{B}}e^{\widehat{\mathbf{A}}_{1,\mathcal{B}}t}\widehat{\mathbf{B}}_{1,\mathcal{B}}, \tag{4.30}$$

so that the outputs are equal to

$$\mathbf{y}_{\mathrm{L,p},\mathcal{B}}(t) = \int_0^t \mathbf{h}_{\mathrm{p},\mathcal{B}}(t-\tau)\mathbf{u}(\tau)\mathrm{d}\tau, \qquad \mathbf{y}_{\mathrm{L,p,r},\mathcal{B}}(t) = \int_0^t \widehat{\mathbf{h}}_{\mathrm{p},\mathcal{B}}(t, t-\tau)\mathbf{u}(\tau)\mathrm{d}\tau.$$

Using these representations of $\mathbf{y}_{\mathrm{L,p},\mathfrak{B}}$ and $\mathbf{y}_{\mathrm{L,p,r},\mathfrak{B}}$ the following Lemma provides an upper bound of the respective L_∞-error.

Lemma 4.6:
Consider the C-stable system (3.58) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$, the reduced system (4.25) with matrices (4.27), and $\mathbf{h}_{\mathrm{p},\mathfrak{B}}$ and $\widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}$ as defined in (4.30). Then, the following inequality holds

$$\|\mathbf{y}_{\mathrm{L,p},\mathfrak{B}} - \mathbf{y}_{\mathrm{L,p,r},\mathfrak{B}}\|_{L_\infty} \leq \left(\int_0^\infty \left\|\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t) - \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t)\right\|_2^2 \mathrm{d}t\right)^{\frac{1}{2}} \|\mathbf{u}\|_{L_2}. \tag{4.31}$$

◊

Proof. We consider the norm of the output error at time $t \geq 0$ that is

$$\left\|\mathbf{y}_{\mathrm{L,p},\mathfrak{B}}(t) - \mathbf{y}_{\mathrm{L,p,r},\mathfrak{B}}(t)\right\|_2 = \left\| \int_0^t \left(\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t-\tau) - \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t-\tau)\right) \mathbf{u}(\tau)\mathrm{d}\tau\right\|_2.$$

Applying the Cauchy-Schwarz inequality multiple times yields

$$\begin{aligned}
\left\|\mathbf{y}_{\mathrm{L,p},\mathfrak{B}}(t) - \mathbf{y}_{\mathrm{L,p,r},\mathfrak{B}}(t)\right\|_2 &\leq \int_0^t \left\| \left(\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t-\tau) - \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t-\tau)\right) \mathbf{u}(\tau)\right\|_2 \mathrm{d}\tau \\
&\leq \int_0^t \left\|\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t-\tau) - \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t-\tau)\right\|_2 \|\mathbf{u}(\tau)\|_2 \mathrm{d}t \\
&\leq \left(\int_0^t \left\|\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t-\tau) - \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t-\tau)\right\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}} \left(\int_0^t \|\mathbf{u}(\tau)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}}.
\end{aligned}$$

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

$$\begin{aligned}
\|\mathbf{y}_{\mathrm{L,p},\mathfrak{B}} - \mathbf{y}_{\mathrm{L,p,r},\mathfrak{B}}\|_{L_\infty} &\leq \left(\int_0^\infty \|\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t) - \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t)\|_2^2 \mathrm{d}t\right)^{\frac{1}{2}} \left(\int_0^\infty \|\mathbf{u}(\tau)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}} \\
&= \left(\int_0^\infty \|\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t) - \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t)\|_2^2 \mathrm{d}t\right)^{\frac{1}{2}} \|\mathbf{u}\|_{L_2}.
\end{aligned}$$

□

The bound in (4.31) includes the expression

$$\int_0^\infty \left\|\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t) - \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t)\right\|_2^2 \mathrm{d}t \leq \int_0^\infty \|\mathbf{h}_{\mathrm{p},\mathfrak{B}}(t)\|_{\mathrm{F}}^2 - 2\langle \mathbf{h}_{\mathrm{p},\mathfrak{B}}(t), \widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t)\rangle + \left\|\widehat{\mathbf{h}}_{\mathrm{p},\mathfrak{B}}(t)\right\|_{\mathrm{F}}^2 \mathrm{d}t.$$

It is used in the following lemma to determine the different components of the bound (4.31) using the respective system Gramians. For that, we define

$$\widetilde{\mathfrak{P}}_{\mathrm{p},\mathfrak{B}} := \int_0^\infty e^{\mathfrak{E}^{-1}\mathfrak{A}t} \mathfrak{E}^{-1}\mathfrak{B}\widehat{\mathbf{B}}_{1,\mathfrak{B}}^{\mathrm{T}} e^{\widehat{\mathbf{A}}_{1,\mathfrak{B}}^{\mathrm{T}} t} \mathrm{d}t, \qquad \widehat{\mathbf{P}}_{1,\mathfrak{B}} := \int_0^\infty e^{\widehat{\mathbf{A}}_{1,\mathfrak{B}} t} \widehat{\mathbf{B}}_{1,\mathfrak{B}} \widehat{\mathbf{B}}_{1,\mathfrak{B}}^{\mathrm{T}} e^{\widehat{\mathbf{A}}_{1,\mathfrak{B}}^{\mathrm{T}} t} \mathrm{d}t. \tag{4.32}$$

Lemma 4.7:
Consider the C-stable system (3.58) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$, the reduced system (4.25) with matrices (4.27), the corresponding controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ as defined in (3.62), the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ from (4.32), and the reduced controllability Gramian $\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}$ from (4.32). The functions $\mathbf{h}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\widehat{\mathbf{h}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ are as defined in (4.30). Then, the following equations hold

$$\int_0^\infty \|\mathbf{h}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)\|_{\mathrm{F}}^2 \mathrm{d}t = \operatorname{tr}\left(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\right), \qquad \int_0^\infty \left\|\widehat{\mathbf{h}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)\right\|_{\mathrm{F}}^2 \mathrm{d}t = \operatorname{tr}\left(\widehat{\mathbf{C}}_{1,\boldsymbol{\mathcal{B}}}\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}\widehat{\mathbf{C}}_{1,\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\right), \tag{4.33a}$$

$$\int_0^\infty \langle \mathbf{h}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t), \widehat{\mathbf{h}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)\rangle \mathrm{d}t = \operatorname{tr}\left(\boldsymbol{\mathcal{C}}\widetilde{\mathbf{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{C}}_{1,\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\right). \tag{4.33b}$$

◊

Proof. We derive

$$\int_0^\infty \|\mathbf{h}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)\|_{\mathrm{F}}^2 \mathrm{d}t = \int_0^\infty \operatorname{tr}\left(\boldsymbol{\mathcal{C}} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}} e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\right) \mathrm{d}t = \operatorname{tr}\left(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\right),$$

what proves the first equation in (4.33a) while the second one is proven analogously. To show the last equation (4.33b), we derive

$$\int_0^\infty \langle \mathbf{h}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t), \widehat{\mathbf{h}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t)\rangle \mathrm{d}t = \int_0^\infty \operatorname{tr}\left(\boldsymbol{\mathcal{C}} e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t}\boldsymbol{\mathcal{B}}\widehat{\mathbf{B}}_{1,\boldsymbol{\mathcal{B}}}^{\mathrm{T}} e^{\widehat{\mathbf{A}}_{1,\boldsymbol{\mathcal{B}}}^{\mathrm{T}}t}\widehat{\mathbf{C}}_{1,\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\right) \mathrm{d}t = \operatorname{tr}\left(\boldsymbol{\mathcal{C}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{C}}_{1,\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\right)$$

what proves the lemma. □

From Lemma 4.6 and Lemma 4.7, we derive the following theorem, which provides a bound of the L_∞-error $\|\mathbf{y}_{\mathrm{L,p},\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{L,p,r},\boldsymbol{\mathcal{B}}}\|_{L_\infty}$.

Theorem 4.8:
Consider the C-stable system (3.58) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$, the reduced system (4.25) with matrices (4.27). Also, consider the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ as defined in (3.62), the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ from (4.32), and the reduced controllability Gramian $\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}$ from (4.32). Then, the error between the output $\mathbf{y}_{\mathrm{L,p},\boldsymbol{\mathcal{B}}}$ and the reduced output $\mathbf{y}_{\mathrm{L,p,r},\boldsymbol{\mathcal{B}}}$ satisfies the following bound

$$\|\mathbf{y}_{\mathrm{L,p},\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{L,p,r},\boldsymbol{\mathcal{B}}}\|_{L_\infty}^2 \leq \left(\operatorname{tr}\left(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\right) - 2\operatorname{tr}\left(\boldsymbol{\mathcal{C}}\widetilde{\mathbf{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{C}}_{1,\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\right) + \operatorname{tr}\left(\widehat{\mathbf{C}}_{1,\boldsymbol{\mathcal{B}}}\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}\widehat{\mathbf{C}}_{1,\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\right) \right) \|\mathbf{u}\|_{L_2}^2. \quad ◊$$

We apply the same bounds to the second error component $\|\mathbf{y}_{\mathrm{L},\mathbf{z}_0} - \mathbf{y}_{\mathrm{L,r},\mathbf{z}_0}\|_{L_\infty}$. Therefore, we define

$$\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0} := \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\mathbf{Z}_0\mathbf{Z}_{0,\mathrm{r}}^{\mathrm{T}} e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{z}_0}^{\mathrm{T}}t} \mathrm{d}t, \qquad \widehat{\mathbf{P}}_{1,\mathbf{z}_0} := \int_0^\infty e^{\widehat{\mathbf{A}}_1 t}\widehat{\mathbf{Z}}_{0,1}\widehat{\mathbf{Z}}_{0,1}^{\mathrm{T}} e^{\widehat{\mathbf{A}}_1^{\mathrm{T}}t} \mathrm{d}t. \tag{4.34}$$

Applying Theorem 4.8 to the second error component yields the following corollary.

Corollary 4.9:
Consider the C-stable system (3.58) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$, the reduced systems (4.25) and (4.26) with matrices (4.27). Also, consider the controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}$ are as defined in (3.62) and (3.66), respectively, the matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}$ from (4.32), and the matrices $\widehat{\mathbf{P}}_{1,\mathbf{z}_0}$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}$ from (4.34). Then, the error between the output $\mathbf{y}_{\mathrm{L}}$ and the reduced output $\mathbf{y}_{\mathrm{L,r}}$ satisfies the following bound

$$\begin{aligned}\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|^2_{L_\infty} \leq{}& \Big(\operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big) - 2\operatorname{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\mathbf{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\Big) + \operatorname{tr}\Big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\Big) \Big)\|\mathbf{u}\|^2_{L_2} \\ &+ \Big(\operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big) - 2\operatorname{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\mathbf{P}}_{\mathrm{p},\mathbf{z}_0}\boldsymbol{\mathcal{C}}^{\mathrm{T}}_{\mathrm{r},\mathbf{z}_0}\Big) + \operatorname{tr}\Big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{P}}_{1,\mathbf{z}_0}\boldsymbol{\mathcal{C}}^{\mathrm{T}}_{\mathrm{r},\mathbf{z}_0}\Big) \Big)\|\zeta_0\|^2_2 \qquad (4.35)\end{aligned}$$

◊

4.2.1.2 Extended-input approach for inhomogeneous first-order DAE systems

In this paragraph, we apply the extended-input approach to derive surrogate models of the DAE systems (3.54) and (3.100) with a linear and a quadratic output equation, respectively, to incorporate the initial conditions into the reduction process. More precisely, we are concerned with deriving reduced-order models of the form

$$\begin{aligned}\boldsymbol{\mathcal{E}}_{\mathrm{r}}\dot{\mathbf{z}}_{\mathrm{r}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t) + \boldsymbol{\mathcal{B}}_{\mathrm{p,r}}\mathbf{u}(t), \qquad \mathbf{z}_{\mathrm{r}}(0) = \mathbf{Z}_{0,\mathrm{r}}\zeta_0, \\ \mathbf{y}_{\mathrm{L,r}}(t) &= \boldsymbol{\mathcal{C}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t),\end{aligned} \tag{4.36}$$

and

$$\begin{aligned}\boldsymbol{\mathcal{E}}_{\mathrm{r}}\dot{\mathbf{z}}_{\mathrm{r}}(t) &= \boldsymbol{\mathcal{A}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t) + \boldsymbol{\mathcal{B}}_{\mathrm{p,r}}\mathbf{u}(t), \qquad \mathbf{z}_{\mathrm{r}}(0) = \mathbf{Z}_{0,\mathrm{r}}\zeta_0, \\ \mathbf{y}_{\mathrm{Q,r}}(t) &= \mathbf{z}_{\mathrm{r}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}_{\mathrm{r}}\mathbf{z}_{\mathrm{r}}(t),\end{aligned} \tag{4.37}$$

with matrices

$$\begin{aligned}&\boldsymbol{\mathcal{E}}_{\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{T}}_{\mathrm{r}} = \begin{bmatrix}\mathbf{I} & 0\\ 0 & \widehat{\mathbf{E}}_2\end{bmatrix}, \quad \boldsymbol{\mathcal{A}}_{\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{T}}_{\mathrm{r}} = \begin{bmatrix}\widehat{\mathbf{A}}_1 & 0\\ 0 & \mathbf{I}\end{bmatrix}, \qquad \mathbf{Z}_{0,\mathrm{r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\mathbf{Z}_0 = \begin{bmatrix}\widehat{\mathbf{Z}}_{0,1}\\ \widehat{\mathbf{Z}}_{0,2}\end{bmatrix}, \\ &\boldsymbol{\mathcal{B}}_{\mathrm{p,r}} = \boldsymbol{\mathcal{V}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{B}} = \begin{bmatrix}\widehat{\mathbf{B}}_1\\ \widehat{\mathbf{B}}_2\end{bmatrix}, \quad \boldsymbol{\mathcal{C}}_{\mathrm{r}} = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{T}}_{\mathrm{r}} = \begin{bmatrix}\widehat{\mathbf{C}}_1 & \widehat{\mathbf{C}}_2\end{bmatrix}, \quad \boldsymbol{\mathcal{M}}_{\mathrm{r}} = \boldsymbol{\mathcal{T}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{T}}_{\mathrm{r}} = \begin{bmatrix}\widehat{\mathbf{M}}_{11} & \widehat{\mathbf{M}}_{12}\\ \widehat{\mathbf{M}}_{12}^{\mathrm{T}} & \widehat{\mathbf{M}}_{22}\end{bmatrix}\end{aligned} \tag{4.38}$$

generated using projecting matrices $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$, $\boldsymbol{\mathcal{T}}_{\mathrm{r}} \in \mathbb{R}^{N\times R}$, where the reduced dimension R is significantly smaller than the original dimension N, i.e., $R \ll N$. Consequently, we aim for a reduced system, which is inherently decoupled into a differential and an algebraic reduced state, i.e., the reduced state $\mathbf{z}_{\mathrm{r}}(t)$ consists of a differential component $\mathbf{z}_{\mathrm{p,r}}(t) := \left[\begin{smallmatrix}\mathbf{z}_{1,\mathrm{r}}(t)\\ 0\end{smallmatrix}\right]$ and an algebraic one $\mathbf{z}_{\mathrm{i,r}}(t) := \left[\begin{smallmatrix}0\\ \mathbf{z}_{2,\mathrm{r}}(t)\end{smallmatrix}\right]$ with $\mathbf{z}_{\mathrm{r}}(t) = \mathbf{z}_{\mathrm{p,r}}(t) + \mathbf{z}_{\mathrm{i,r}}(t)$.

We aim to find reduced-order models (4.36) and (4.37) that approximate the input-to-output behavior of the full-order models (3.54) and (3.100), i.e., the expressions $\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|$ and $\|\mathbf{y}_{\mathrm{Q}} - \mathbf{y}_{\mathrm{Q,r}}\|$ are small in an appropriate norm. As described in Section 3.2.1.2, the original system (3.54) with a linear output equation corresponds to the same transfer function as the surrogate system (3.91), and as shown in Section 3.2.2.2, the original system (3.100) with a quadratic output equation corresponds to the same transfer function as the surrogate system introduced in (3.105). Both surrogate systems incorporate the initial condition spaces into the input so that the respective controllability Gramians describe the input- and initial condition-to-output behavior. Hence, in the following, the respective Gramians are utilized to derive the corresponding reduced surrogate models (4.36) and (4.37).

As summarized in Table 3.6 and Table 3.7, the states corresponding to large eigenvalues of the Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}$, $\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$, and $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,p},w_{\mathrm{p}}}$ from (3.93), (3.97), and (3.111), respectively, span the most dominant controllability and observability subspaces of the respective systems. On the other hand, states corresponding to small eigenvalues are negligible and, hence, truncated in the BT method. To evaluate the algebraic components of the systems, the improper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{i},w_{\mathrm{p}}}$ from (3.94) and the improper observability Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,i},w_{\mathrm{p}}}$ introduced in (3.97) and (3.116) are used to identify states that are not reachable or not observable, i.e., states that do not affect the dynamics of the system. These states are then removed to find a minimal realization of the algebraic system components.

We want to mention that the systems (3.54) and (3.100) have the same proper and improper controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},w_{\mathrm{p}}}$. The observability Gramians, on the other hand, differ. However, since the BT method corresponding to both systems types is the same, we denote the proper observability Gramians $\boldsymbol{\mathcal{Q}}_{\mathrm{L,p}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,p},w_{\mathrm{p}}}$ in the following as $\boldsymbol{\mathcal{Q}}_{\mathrm{p}}$ and the improper ones $\boldsymbol{\mathcal{Q}}_{\mathrm{L,i}}$ and $\boldsymbol{\mathcal{Q}}_{\mathrm{Q,i},w_{\mathrm{p}}}$ in the following as $\boldsymbol{\mathcal{Q}}_{\mathrm{i}}$ so that the user can choose the correct observability Gramian according to the considered system.

We aim to truncate states corresponding to the small eigenvalues of the proper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}$ and the proper observability Gramian $\boldsymbol{\mathcal{Q}}_{\mathrm{p}}$. Therefore, we follow the methodology presented in Algorithm 2 to derive a balanced and truncated system. Since all Gramians are symmetric and positive semi-definite, there exist factorizations

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

We compute the singular value decompositions

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

where $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_{N_f}, 0, \ldots)$, $\sigma_1 \geq \cdots \geq \sigma_{N_f}$, includes the proper Hankel singular values of the system. The differential states that are simultaneously difficult to reach

Algorithm 11 BT method for the first-order DAE systems (3.54) and (3.100) with a linear or quadratic output equations using the extended-input approach.

```
Require: The original system (3.54) or (3.100) and the order $R$.
Ensure: The reduced system (4.36) or (4.37).
Compute the proper and improper controllability Gramians $\mathcal{P}_{\mathrm{p},\mathcal{W}_\mathrm{p}}$ and $\mathcal{P}_{\mathrm{i},\mathcal{W}_\mathrm{p}}$.
Compute the proper observability Gramians $\mathcal{Q}_\mathrm{p}$ equal to $\mathcal{Q}_{\mathrm{L},\mathrm{p}}$ or $\mathcal{Q}_{\mathrm{Q},\mathrm{p},\mathcal{W}_\mathrm{p}}$ and the
    improper one $\mathcal{Q}_\mathrm{i}$ equal to $\mathcal{Q}_{\mathrm{L},\mathrm{i}}$ or $\mathcal{Q}_{\mathrm{Q},\mathrm{i},\mathcal{W}_\mathrm{p}}$.
Perform the singular values decomposition
    $\mathcal{S}_\mathrm{p}\mathcal{E}\mathcal{R}_\mathrm{p} = \begin{bmatrix}\mathbf{U}_{\mathrm{p},1} & \mathbf{U}_{\mathrm{p},2}\end{bmatrix}\begin{bmatrix}\mathbf{\Sigma}_1 & \\ & \mathbf{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{p},1}^\mathrm{T} \\ \mathbf{V}_{\mathrm{p},2}^\mathrm{T}\end{bmatrix}, \quad \mathcal{S}_\mathrm{i}\mathcal{A}\mathcal{R}_\mathrm{i} = \begin{bmatrix}\mathbf{U}_{\mathrm{i},1} & \mathbf{U}_{\mathrm{i},2}\end{bmatrix}\begin{bmatrix}\mathbf{\Theta}_1 & \\ & 0\end{bmatrix}\begin{bmatrix}\mathbf{V}_{\mathrm{i},1}^\mathrm{T} \\ \mathbf{V}_{\mathrm{i},2}^\mathrm{T}\end{bmatrix}.$
Construct the projection matrices
    $\mathcal{V}_\mathrm{r} = \begin{bmatrix}\mathcal{S}_\mathrm{p}^\mathrm{T}\mathbf{U}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}} & \mathcal{S}_\mathrm{i}^\mathrm{T}\mathbf{U}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}\end{bmatrix}, \qquad \mathcal{T}_\mathrm{r} = \begin{bmatrix}\mathcal{R}_\mathrm{p}\mathbf{V}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}} & \mathcal{R}_\mathrm{i}\mathbf{V}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}\end{bmatrix}.$
Construct reduced matrices as defined in (4.38).
```

and to observe correspond to the smallest Hankel singular values, which are the diagonal elements of $\mathbf{\Sigma}_2$. We truncate the corresponding states that lie in the spaces spanned by $\mathbf{U}_{\mathrm{p},2}$ and $\mathbf{V}_{\mathrm{p},2}$ by building the projection matrices

$$\mathcal{V}_\mathrm{r} = \begin{bmatrix}\mathcal{S}_\mathrm{p}^\mathrm{T}\mathbf{U}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}} & \mathcal{S}_\mathrm{i}^\mathrm{T}\mathbf{U}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}\end{bmatrix}, \qquad \mathcal{T}_\mathrm{r} = \begin{bmatrix}\mathcal{R}_\mathrm{p}\mathbf{V}_{\mathrm{p},1}\mathbf{\Sigma}_1^{-\frac{1}{2}} & \mathcal{R}_\mathrm{i}\mathbf{V}_{\mathrm{i},1}\mathbf{\Theta}_1^{-\frac{1}{2}}\end{bmatrix}.$$

Note that additionally improper states that correspond to zero eigenvalues 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 removed. Multiplying the system matrices of the system in (3.91) and (3.105) by $\mathcal{V}_\mathrm{r}^\mathrm{T}$ and $\mathcal{T}_\mathrm{r}$ leads to a reduced system in (4.36) and (4.37) with matrices (4.38) where $\widehat{\mathbf{A}}_1$ is nonsingular and $\widehat{\mathbf{E}}_2$ is a nilpotent matrix. This method results in the Algorithm 11.

Error bound for systems with a linear output equation To evaluate the quality of the approximation by the reduced system, we again have to distinguish between systems with linear output equations and those with quadratic ones. First, we evaluate the error for systems with a linear output equation. For that, we define the matrix and the reduced controllability Gramian

$$\begin{aligned}\widetilde{\mathcal{P}}_{\mathrm{p},\mathcal{W}_\mathrm{p}} &:= \int_0^\infty \mathcal{F}_\mathbf{J}(t)\mathcal{W}_\mathrm{p}\mathcal{W}_\mathrm{p}^\mathrm{T}\mathcal{F}_\mathbf{J}(t)^\mathrm{T}\mathrm{d}t,\\ \mathcal{P}_{\mathrm{p},\mathcal{W}_\mathrm{p},\mathrm{r}} &:= \int_0^\infty e^{\mathcal{E}_\mathrm{r}^{-1}\mathcal{A}_\mathrm{r}t}\mathcal{E}^{-1}\mathcal{W}_{\mathrm{p},\mathrm{r}}\mathcal{W}_{\mathrm{p},\mathrm{r}}^\mathrm{T}\mathcal{E}^{-\mathrm{T}}e^{\mathcal{A}_\mathrm{r}^\mathrm{T}\mathcal{E}_\mathrm{r}^{-\mathrm{T}}t}\mathrm{d}t.\end{aligned} \tag{4.39}$$

We apply the bound from (4.35), where all the subsystems are generated using the same bases $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$, $\boldsymbol{\mathcal{T}}_{\mathrm{r}}$, which results in the following theorem.

Theorem 4.10:
Consider the C-stable system (3.54) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$ and the surrogate system (4.36). Also consider the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}$ as defined in (3.93), the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},w_{\mathrm{p}}}$ from (4.39), and the reduced Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}},\mathrm{r}}$ from (4.39). Then the L_∞-error of the outputs is bounded by

$$\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_\infty}^2 \leq \Big(\operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big) \\ - 2\operatorname{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\mathbf{P}}_{w_{\mathrm{p}}}\boldsymbol{\mathcal{C}}_{\mathrm{r}}^{\mathrm{T}}\Big) + \operatorname{tr}\big(\boldsymbol{\mathcal{C}}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}},\mathrm{r}}\boldsymbol{\mathcal{C}}_{\mathrm{r}}^{\mathrm{T}}\big) \Big) \big(\|\mathbf{u}\|_{L_2}^2 + \|\zeta_0\|_2^2\big)\,. \tag{4.40}$$

◊

Proof. We apply the bound from (4.35) to obtain

$$\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_\infty}^2 \leq \Big(\operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big) - 2\operatorname{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\mathbf{P}}_{\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\Big) + \operatorname{tr}\big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{P}}_{\mathrm{p,r},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\big) \Big)\|\mathbf{u}\|_{L_2}^2 \\ + \Big(\operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathbf{z}_0}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big) - 2\operatorname{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\mathbf{P}}_{\mathbf{z}_0}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{z}_0}^{\mathrm{T}}\Big) + \operatorname{tr}\big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{z}_0}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{z}_0}^{\mathrm{T}}\big) \Big)\|\zeta_0\|_2^2$$

for the controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}$ defined in (3.62) and (3.66), respectively, the reduced controllability Gramians $\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}$ and $\widehat{\mathbf{P}}_{1,\mathbf{z}_0}$, and the matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}$ as defined in (4.32) and (4.34). Since it holds that $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}} = \boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} + \boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}$, $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},w_{\mathrm{p}}} = \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} + \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}$, and $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}},\mathrm{r}} = \boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B},\mathrm{r}} + \boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{Z}_0,\mathrm{r}}$, the right-hand side can be bounded by the one in (4.40), which proves the statement. □

Error bound for systems with a quadratic output equation To describe the output error for the systems (3.100) and (4.37) with a quadratic output equation, we bound the error between $\mathbf{y}_{\mathrm{Q}}$ and $\mathbf{y}_{\mathrm{Q,r}}$ as

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

according to the four summands defined in (3.102). In the following, we consider the respective error norms separately. Since we do not truncate the improper states the error $\|\mathbf{y}_{\mathrm{ii}} - \mathbf{y}_{\mathrm{ii,r}}\|_{L_\infty}$ is equal to zero. Also the components $\|\mathbf{y}_{\mathrm{pi}} - \mathbf{y}_{\mathrm{pi,r}}\|_{L_\infty}$ and $\|\mathbf{y}_{\mathrm{ip}} - \mathbf{y}_{\mathrm{ip,r}}\|_{L_\infty}$ coincide so that only one of them needs to be evaluated. We investigate the remaining summands in the following. Since we consider inhomogeneous systems, the input-related behavior and the initial value-related behavior of the system need to be taken into account. Therefore, we evaluate the input-related components and the initial value-related component separately. We define the different states

$$\mathbf{z}_{\mathrm{p},\boldsymbol{\mathcal{B}}}(t) := \int_0^t \boldsymbol{\mathfrak{F}}_{\mathbf{J}}(t-\tau)\boldsymbol{\mathcal{B}}\mathbf{u}(\tau)\mathrm{d}\tau, \quad \mathbf{z}_{\mathrm{p},\mathbf{z}_0}(t) := \boldsymbol{\mathfrak{F}}_{\mathbf{J}}(t)\mathbf{Z}_0\zeta_0, \quad \mathbf{z}_{\mathrm{i}}(t) := \sum_{k=0}^{\nu-1}\boldsymbol{\mathfrak{F}}_{\mathbf{N}}(t)\boldsymbol{\mathcal{B}}\mathbf{u}^{(k)}(t)$$

with $\mathfrak{F}_{\mathbf{J}}$ and $\mathfrak{F}_{\mathbf{N}}$ as defined in (2.13), and the reduced state approximations

$$\mathbf{z}_{\mathrm{p,r},\mathfrak{B}}(t) := \int_0^t e^{\widehat{\mathbf{A}}_1(t-\tau)}\widehat{\mathbf{B}}_1\mathbf{u}(\tau)\mathrm{d}\tau, \quad \mathbf{z}_{\mathrm{p,r},\mathbf{z}_0}(t) := e^{\widehat{\mathbf{A}}_1 t}\widehat{\mathbf{Z}}_{0,1}\zeta_0, \quad \mathbf{z}_{\mathrm{i,r}}(t) := \sum_{k=0}^{\nu-1} -\widehat{\mathbf{E}}_2^k\widehat{\mathbf{B}}_2\mathbf{u}^{(k)}(t)$$

including the reduced matrices from (4.38) to define the output components.

The proper-proper output error First, we describe the proper-proper output error $\|\mathbf{y}_{\mathrm{pp}} - \mathbf{y}_{\mathrm{pp,r}}\|_{L_\infty}$ that includes the output components

$$\mathbf{y}_{\mathrm{pp},*\circ}(t) := \mathbf{z}_{\mathrm{p},*}(t)^{\mathrm{T}}\mathfrak{M}\mathbf{z}_{\mathrm{p},\circ}(t), \qquad \mathbf{y}_{\mathrm{pp,r},*\circ}(t) := \mathbf{z}_{\mathrm{p,r},*}(t)^{\mathrm{T}}\widehat{\mathbf{M}}_{11}\mathbf{z}_{\mathrm{p,r},\circ}(t)$$

where the subscripts $*$ and $\circ$ are equal to '$\mathfrak{B}$' and '$\mathbf{Z}_0$'. We evaluate the input-related components and the initial value-related components separately so that

$$\begin{aligned}\|\mathbf{y}_{\mathrm{pp}} - \mathbf{y}_{\mathrm{pp,r}}\|_{L_\infty} \leq \|\mathbf{y}_{\mathrm{pp},\mathfrak{B}\mathfrak{B}} - \mathbf{y}_{\mathrm{pp,r},\mathfrak{B}\mathfrak{B}}\|_{L_\infty}& \\ + 2\|\mathbf{y}_{\mathrm{pp},\mathbf{z}_0\mathfrak{B}} - \mathbf{y}_{\mathrm{pp,r},\mathbf{z}_0\mathfrak{B}}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{pp},\mathbf{z}_0\mathbf{z}_0} - \mathbf{y}_{\mathrm{pp,r},\mathbf{z}_0\mathbf{z}_0}\|_{L_\infty}&.\end{aligned}$$

Since the three components are analyzed analogously, we only show the derivation of a bound for $\|\mathbf{y}_{\mathrm{pp},\mathfrak{B}\mathfrak{B}} - \mathbf{y}_{\mathrm{pp,r},\mathfrak{B}\mathfrak{B}}\|_{L_\infty}$. Afterwards, we apply the same methodology for the remaining two components.

To analyze the error $\|\mathbf{y}_{\mathrm{pp},\mathfrak{B}\mathfrak{B}} - \mathbf{y}_{\mathrm{pp,r},\mathfrak{B}\mathfrak{B}}\|_{L_\infty}$ between the proper-proper output $\mathbf{y}_{\mathrm{pp},\mathfrak{B}\mathfrak{B}}$ and its approximation $\mathbf{y}_{\mathrm{pp,r},\mathfrak{B}\mathfrak{B}}$, we define the mappings

$$\mathbf{h}_{\mathrm{pp}}(t_1,t_2) := \mathrm{vec}\big(\mathfrak{B}^{\mathrm{T}}\mathfrak{F}_{\mathbf{J}}(t_1)^{\mathrm{T}}\mathfrak{M}\mathfrak{F}_{\mathbf{J}}(t_2)\mathfrak{B}\big), \quad \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1,t_2) := \mathrm{vec}\Big(\widehat{\mathbf{B}}_1^{\mathrm{T}}e^{\widehat{\mathbf{A}}_1^{\mathrm{T}}t_1}\widehat{\mathbf{M}}_{11}e^{\widehat{\mathbf{A}}_1 t_2}\widehat{\mathbf{B}}_1\Big), \tag{4.41}$$

so that the outputs can be written as

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

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

Lemma 4.11:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathbf{A},\mathbf{E})$, the reduced system (4.37), and the mappings $\mathbf{h}_{\mathrm{pp}}$ and $\widehat{\mathbf{h}}_{\mathrm{pp}}$ as defined in (4.41). Then, the following inequality holds

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

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

$$\begin{aligned}&\left|\mathbf{y}_{\mathrm{pp},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) - \mathbf{y}_{\mathrm{pp,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t)\right| \\ &\qquad = \left|\int_0^t \int_0^t \Big(\mathbf{h}_{\mathrm{pp}}(t-t_1, t-t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t-t_1, t-t_2)\Big)^{\mathrm{T}} (\mathbf{u}(t_2) \otimes \mathbf{u}(t_1))\mathrm{d}t_1\mathrm{d}t_2\right|.\end{aligned}$$

Applying the Cauchy-Schwarz inequality multiple times yields

$$\begin{aligned}|\mathbf{y}_{\mathrm{pp},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) - \mathbf{y}_{\mathrm{pp,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t)| &\leq \int_0^t \int_0^t \Big\| \Big(\mathbf{h}_{\mathrm{pp}}(t-t_1, t-t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t-t_1, t-t_2)\Big)^{\mathrm{T}} \\ &\qquad\qquad \cdot (\mathbf{u}(t_2) \otimes \mathbf{u}(t_1))\Big\|_2 \mathrm{d}t_1\mathrm{d}t_2 \\ &\leq \int_0^t \int_0^t \Big\|\mathbf{h}_{\mathrm{pp}}(t_1, t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\Big\|_2 \|(\mathbf{u}(t_2) \otimes \mathbf{u}(t_1))\|_2 \mathrm{d}t_1\mathrm{d}t_2 \\ &\leq \left(\int_0^t \int_0^t \Big\|\mathbf{h}_{\mathrm{pp}}(t_1, t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\Big\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \\ &\qquad\qquad \cdot \left(\int_0^t \int_0^t \|(\mathbf{u}(t_2) \otimes \mathbf{u}(t_1))\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}}.\end{aligned}$$

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

$$\begin{aligned}\|\mathbf{y}_{\mathrm{pp},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{pp,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\|_{L_\infty} &\leq \left(\int_0^\infty \int_0^\infty \Big\|\mathbf{h}_{\mathrm{pp}}(t_1, t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\Big\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \\ &\qquad\qquad \cdot \left(\int_0^\infty \int_0^\infty \|(\mathbf{u}(t_2) \otimes \mathbf{u}(t_1))\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \\ &= \left(\int_0^\infty \int_0^\infty \Big\|\mathbf{h}_{\mathrm{pp}}(t_1, t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\Big\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\right)^{\frac{1}{2}} \|\mathbf{u} \otimes \mathbf{u}\|_{L_2}. \qquad \square\end{aligned}$$

Note, that the factor $\|\mathbf{u} \otimes \mathbf{u}\|_{L_2}$ is replaced by $\|\mathbf{u} \otimes \zeta_0\|_{L_2}$ when considering the output $\|\mathbf{y}_{\mathrm{pp},\mathbf{z}_0\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{pp,r},\mathbf{z}_0\boldsymbol{\mathcal{B}}}\|_{L_\infty}$ and by $\|\zeta_0\|_2^2$ when considering $\|\mathbf{y}_{\mathrm{pp},\mathbf{z}_0\mathbf{z}_0} - \mathbf{y}_{\mathrm{pp,r},\mathbf{z}_0\mathbf{z}_0}\|_{L_\infty}$. Also the mappings $\mathbf{h}_{\mathrm{pp}}$ and $\widehat{\mathbf{h}}_{\mathrm{pp}}$ need to be replaced accordingly.

Since the bound presented in Lemma 4.11 includes the expression

$$\begin{aligned}&\int_0^\infty \int_0^\infty \Big\|\mathbf{h}_{\mathrm{pp}}(t_1, t_2) - \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\Big\|_2^2 \mathrm{d}t_1\mathrm{d}t_2 \\ &\qquad = \int_0^\infty \int_0^\infty \|\mathbf{h}_{\mathrm{pp}}(t_1, t_2)\|_2^2 - 2\langle\mathbf{h}_{\mathrm{pp}}(t_1, t_2), \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\rangle + \Big\|\widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\Big\|_2^2 \mathrm{d}t_1\mathrm{d}t_2\end{aligned}$$

the following lemma is used to determine the different components of this bound using the respective system Gramians. For that, we also define the matrices

$$\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} := \int_0^\infty \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\boldsymbol{\mathcal{B}}\widehat{\mathbf{B}}_1^{\mathrm{T}} e^{\mathbf{A}_1^{\mathrm{T}} t}\mathrm{d}t, \qquad \widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}} := \int_0^\infty e^{\mathbf{A}_1 t}\widehat{\mathbf{B}}_1\widehat{\mathbf{B}}_1^{\mathrm{T}} e^{\mathbf{A}_1^{\mathrm{T}} t}\mathrm{d}t. \tag{4.42}$$

Lemma 4.12:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$, the reduced system (4.37), the corresponding proper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ as defined in (3.62), the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$, and the reduced proper controllability Gramian $\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}$ from (4.42). The mappings $\mathbf{h}_{\mathrm{pp}}$ and $\widehat{\mathbf{h}}_{\mathrm{pp}}$ are as defined in (4.41). Then, the following equations are fulfilled

$$\int_0^\infty \int_0^\infty \|\mathbf{h}_{\mathrm{pp}}(t_1, t_2)\|_2^2 \mathrm{d}t_1 \mathrm{d}t_2 = \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} \boldsymbol{\mathcal{M}} \boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} \boldsymbol{\mathcal{M}}) \,, \tag{4.43a}$$

$$\int_0^\infty \int_0^\infty \left\|\widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\right\|_2^2 \mathrm{d}t_1 \mathrm{d}t_2 = \mathrm{tr}\left(\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}} \widehat{\mathbf{M}}_{11} \widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}} \widehat{\mathbf{M}}_{11}\right) , \tag{4.43b}$$

$$\int_0^\infty \int_0^\infty \langle \mathbf{h}_{\mathrm{pp}}(t_1, t_2), \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\rangle \mathrm{d}t_1 \mathrm{d}t_2 = \mathrm{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}^{\mathrm{T}} \boldsymbol{\mathcal{M}} \widetilde{\mathbf{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} \widehat{\mathbf{M}}_{11}\right) . \tag{4.43c}$$

◊

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

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

what proves the equation in (4.43a), while the one in (4.43b) is proven analogously. To show that the equation in (4.43c) holds, we make use of the property $\langle \mathrm{vec}(\mathbf{X}), \mathrm{vec}(\mathbf{Y})\rangle = \mathrm{tr}\big(\mathbf{X}^{\mathrm{T}} \mathbf{Y}\big)$ and obtain

$$\begin{aligned}
&\int_0^\infty \int_0^\infty \langle \mathbf{h}_{\mathrm{pp}}(t_1, t_2), \widehat{\mathbf{h}}_{\mathrm{pp}}(t_1, t_2)\rangle \mathrm{d}t_1 \mathrm{d}t_2 \\
&\qquad = \int_0^\infty \int_0^\infty \mathrm{tr}\Big(\boldsymbol{\mathcal{B}}^{\mathrm{T}} \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t_2)^{\mathrm{T}} \boldsymbol{\mathcal{M}} \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t_1) \boldsymbol{\mathcal{B}} \widehat{\mathbf{B}}_1^{\mathrm{T}} e^{\widehat{\mathbf{A}}^{\mathrm{T}} t_1} \widehat{\mathbf{M}}_{11} e^{\widehat{\mathbf{A}} t_2} \widehat{\mathbf{B}}_1\Big) \, \mathrm{d}t_1 \mathrm{d}t_2 \\
&\qquad = \int_0^\infty \int_0^\infty \mathrm{tr}\Big(e^{\widehat{\mathbf{A}}_1 t_2} \widehat{\mathbf{B}}_1 \boldsymbol{\mathcal{B}}^{\mathrm{T}} \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t_2)^{\mathrm{T}} \boldsymbol{\mathcal{M}} \boldsymbol{\mathcal{F}}_{\mathbf{J}}(t_1) \boldsymbol{\mathcal{B}} \widehat{\mathbf{B}}_1^{\mathrm{T}} e^{\widehat{\mathbf{A}}_1^{\mathrm{T}} t_1} \widehat{\mathbf{M}}_{11}\Big) \, \mathrm{d}t_1 \mathrm{d}t_2 \\
&\qquad = \mathrm{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}^{\mathrm{T}} \boldsymbol{\mathcal{M}} \widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}} \widehat{\mathbf{M}}_{11}\Big) \,.
\end{aligned}$$

□

From Lemma 4.11 and 4.12, we derive the following theorem, which provides a bound of the L_∞-error of the proper-proper output component.

Theorem 4.13:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$, the reduced system (4.37), the corresponding proper controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ as defined in (3.62), the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$, and the reduced proper controllability Gramian $\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}$ from (4.42). The error between the proper-proper output $\mathbf{y}_{\mathrm{pp},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}$ of the original system (3.100) and the reduced output $\mathbf{y}_{\mathrm{pp,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}$ satisfies the following bound

$$\begin{aligned}
&\|\mathbf{y}_{\mathrm{pp},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{pp,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\|_{L_\infty}^2 \\
&\qquad \le \Big(\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}) - 2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{M}}_{11}\Big) \\
&\qquad\qquad + \operatorname{tr}\Big(\widehat{\mathbf{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{M}}_{11}\widehat{\mathbf{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{M}}_{11}\Big)\Big)\|\mathbf{u}\otimes\mathbf{u}\|_{L_2}^2. \quad \diamondsuit
\end{aligned}$$

The improper-proper output error Now, we bound the improper-proper output error $\|\mathbf{y}_{\mathrm{ip}} - \mathbf{y}_{\mathrm{ip,r}}\|_{L_\infty}$ that includes the output components

$$\mathbf{y}_{\mathrm{ip},\boldsymbol{\mathcal{B}}\circ}(t) := \mathbf{z}_{\mathrm{i}}(t)^{\mathrm{T}}\boldsymbol{\mathcal{M}}\mathbf{z}_{\mathrm{p},\circ}(t), \qquad \mathbf{y}_{\mathrm{ip,r},\boldsymbol{\mathcal{B}}\circ}(t) := \mathbf{z}_{\mathrm{i,r}}(t)^{\mathrm{T}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\mathbf{z}_{\mathrm{p,r},\circ}(t)$$

where the subscript $\circ$ is equal to '$\boldsymbol{\mathcal{B}}$' and '$\mathbf{Z}_0$'. We evaluate the input-related components and the initial value-related components separately so that

$$\|\mathbf{y}_{\mathrm{ip}} - \mathbf{y}_{\mathrm{ip,r}}\|_{L_\infty} \le \|\mathbf{y}_{\mathrm{ip},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{ip,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\|_{L_\infty} + 2\|\mathbf{y}_{\mathrm{ip},\boldsymbol{\mathcal{B}}\mathbf{Z}_0} - \mathbf{y}_{\mathrm{ip,r},\boldsymbol{\mathcal{B}}\mathbf{Z}_0}\|_{L_\infty}.$$

We derive an error bound for the two components following the same theory. Hence, we only investigate the error $\|\mathbf{y}_{\mathrm{ip},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}} - \mathbf{y}_{\mathrm{ip,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}\|_{L_\infty}$, while the remaining one is computed analogously. To bound the improper-proper output error, i.e., the error between the improper-proper output $\mathbf{y}_{\mathrm{ip},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t)$ and the reduced improper-proper output $\mathbf{y}_{\mathrm{ip,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t)$, we define the mappings

$$\mathbf{h}_{\mathrm{ip}}(t,k) := \operatorname{vec}\Big(\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)^{\mathrm{T}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{F}}_{\mathbf{J}}(t)\boldsymbol{\mathcal{B}}\Big) \quad \text{and} \quad \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k) := \operatorname{vec}\Big(\widehat{\mathbf{B}}_2^{\mathrm{T}}(\widehat{\mathbf{E}}_2^k)^{\mathrm{T}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}e^{\widehat{\mathbf{A}}_1 t}\widehat{\mathbf{B}}_1\Big). \tag{4.44}$$

Using the mappings $\mathbf{h}_{\mathrm{ip}}$ and $\widehat{\mathbf{h}}_{\mathrm{ip}}$ from (4.44), we rewrite the outputs as

$$\begin{aligned}
\mathbf{y}_{\mathrm{ip},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) &= \int_0^t \sum_{k=0}^{\nu-1}\mathbf{h}_{\mathrm{ip}}(t-\tau,k)^{\mathrm{T}}\left(\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\right)\mathrm{d}\tau,\\
\mathbf{y}_{\mathrm{ip,r},\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}}(t) &= \int_0^t \sum_{k=0}^{\nu-1}\widehat{\mathbf{h}}_{\mathrm{ip}}(t-\tau,k)^{\mathrm{T}}\left(\mathbf{u}(\tau)\otimes\mathbf{u}^{(k)}(t)\right)\mathrm{d}\tau.
\end{aligned}$$

We use this representation of the improper-proper outputs to derive the following lemma, which provides a bound of the respective L_∞-error.

Lemma 4.14:
We consider the C-stable system (3.100) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$, the reduced system (4.37), and $\mathbf{h}_{\text{ip}}$ and $\widehat{\mathbf{h}}_{\text{ip}}$ as defined in (4.44). Then, the following bound holds

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

Proof. Using the mappings $\mathbf{h}_{\text{ip}}$ and $\widehat{\mathbf{h}}_{\text{ip}}$ from (4.44), we obtain

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

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

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

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

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

When considering the second component $\|\mathbf{y}_{\text{ip},\mathfrak{B}\mathbf{z}_0} - \mathbf{y}_{\text{ip,r},\mathfrak{B}\mathbf{z}_0}\|_{L_\infty}$, we replace the output norm by $\left(\sum_{k=0}^{\nu-1}\|\zeta_0\otimes\mathbf{u}^{(k)}(t)\|_2^2\right)^{\frac{1}{2}}$ and choose the respective mappings $\mathbf{h}_{\text{ip}}$ and $\widehat{\mathbf{h}}_{\text{ip}}$ accordingly.

The output error bound from Lemma 4.14 contains the following expression

$$\begin{aligned}\int_0^\infty \sum_{k=0}^{\nu-1} \left\|\mathbf{h}_{\mathrm{ip}}(t,k) - \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\right\|_2^2 \mathrm{d}t \\ = \int_0^\infty \sum_{k=0}^{\nu-1} \|\mathbf{h}_{\mathrm{ip}}(t,k)\|_2^2 - 2\langle \mathbf{h}_{\mathrm{ip}}(t,k), \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\rangle + \left\|\widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\right\|_2^2 \mathrm{d}t.\end{aligned}$$

In the following lemma, we derive formulas of the different components of this expression, which contain the Gramians of the respective systems. To do so, we define the matrix and the reduced Gramian

$$\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\boldsymbol{\mathcal{B}}} := \sum_{k=0}^{\nu-1} \boldsymbol{\mathcal{F}}_{\mathbf{N}}(k)\boldsymbol{\mathcal{B}}\widehat{\mathbf{B}}_2^{\mathrm{T}}(\mathbf{N}^k)^{\mathrm{T}}, \qquad \widehat{\mathbf{P}}_{2,\boldsymbol{\mathcal{B}}} := \sum_{k=0}^{\nu-1} \widehat{\mathbf{E}}_2^k\widehat{\mathbf{B}}_2\widehat{\mathbf{B}}_2^{\mathrm{T}}\left(\widehat{\mathbf{E}}_2^k\right)^{\mathrm{T}}. \tag{4.45}$$

Lemma 4.15:
We consider the C-stable system (3.100) with a regular matrix pencil $(\mathbf{A}, \mathbf{E})$, the reduced system (4.37). Also, consider the proper and improper controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{\mathcal{B}}}$ as defined in (3.62) and (3.70), respectively, the matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\boldsymbol{\mathcal{B}}}$, and the reduced proper and improper controllability Gramians $\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}$ and $\widehat{\mathbf{P}}_{2,\boldsymbol{\mathcal{B}}}$ as defined in (4.42) and (4.45). The functionals $\mathbf{h}_{\mathrm{ip}}$ and $\widehat{\mathbf{h}}_{\mathrm{ip}}$ are as defined in (4.44). Then, the following equations hold

$$\begin{aligned}\int_0^\infty \sum_{k=0}^{\nu-1} \|\mathbf{h}_{\mathrm{ip}}(t,k)\|_2^2 \mathrm{d}t &= \mathrm{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{M}})\,, \\ \int_0^\infty \sum_{k=0}^{\nu-1} \|\widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\|_2^2 \mathrm{d}t_1\mathrm{d}t_2 &= \mathrm{tr}\left(\widehat{\mathbf{P}}_{2,\boldsymbol{\mathcal{B}}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\widehat{\mathbf{P}}_{1,\boldsymbol{\mathcal{B}}}\widehat{\mathbf{M}}_{12}\right), \\ \int_0^\infty \sum_{k=0}^{\nu-1} \langle \mathbf{h}_{\mathrm{ip}}(t,k), \widehat{\mathbf{h}}_{\mathrm{ip}}(t,k)\rangle \mathrm{d}t &= \mathrm{tr}\left(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\boldsymbol{\mathcal{B}}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}\widehat{\mathbf{M}}_{12}\right). \qquad \lozenge\end{aligned}$$

Proof. The proof is analogous to the one from Lemma 4.12. □

We use Lemma 4.14 and Lemma 4.15 to derive the following bound of the L_∞ error corresponding to the improper-proper output.

Theorem 4.16:
We consider the C-stable system (3.100) with the nilpotency index ν and a regular matrix pencil, and the reduced system (4.37). Also, consider the proper and improper controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{P}}_{\mathrm{i},\boldsymbol{\mathcal{B}}}$ as defined in (3.62) and (3.70), respectively, the

matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathcal{B}}$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\mathcal{B}}$, and the reduced proper and improper controllability Gramians $\widehat{\mathbf{P}}_{1,\mathcal{B}}$ and $\widehat{\mathbf{P}}_{2,\mathcal{B}}$ as defined in (4.42) and (4.45). The error between the improper-proper output $\mathbf{y}_{\mathrm{ip},\mathcal{B}\mathcal{B}}(t)$ of the original system (3.100) and the reduced output $\mathbf{y}_{\mathrm{ip,r},\mathcal{B}\mathcal{B}}(t)$ satisfies the following bound

$$
\begin{aligned}
&\|\mathbf{y}_{\mathrm{ip},\mathcal{B}\mathcal{B}}(t) - \mathbf{y}_{\mathrm{ip,r},\mathcal{B}\mathcal{B}}(t)\|_{L_\infty}^2 \\
&\qquad \le \Big(\mathrm{tr}\Big(\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{B}}\boldsymbol{\mathcal{M}}\Big) - 2\mathrm{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\mathcal{B}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\Big) \\
&\qquad\qquad + \mathrm{tr}\Big(\widehat{\mathbf{P}}_{1,\mathcal{B}}\widehat{\mathbf{M}}_{12}\widehat{\mathbf{P}}_{2,\mathcal{B}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\Big)\Big)\nu\|\mathbf{u}\|_{\mathfrak{C}^{\nu-1}}^2\|\mathbf{u}\|_{L_2}^2
\end{aligned}
$$

for $\|\mathbf{u}\|_{\mathfrak{C}^{\nu-1}} := \max_{k=0,\ldots,\nu-1}\sup_{t\ge 0}\|\mathbf{u}^{(k)}(t)\|_2$ and output functions $\mathbf{u} \in C^{\nu-1}([0,\infty),\mathbb{R}^m) \cup L_2([0,\infty),\mathbb{R}^m)$. ◊

Proof. We apply Lemma 4.14 and 4.15 to derive the first multiplier of the right-hand side. Moreover, applying Kronecker product properties and Cauchy-Schwarz inequality to the second factor from Lemma 4.14 yields

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

which proves the statement. □

The total output error Finally, we use the bounds for the different error components introduced in Theorem 4.13 and Theorem 4.16 to derive an expression that bounds the total error between the output $\mathbf{y}_{\mathrm{Q}}$ and $\mathbf{y}_{\mathrm{Q,r}}$.

Theorem 4.17:
Consider the C-stable system (3.100) with a regular matrix pencil $(\mathbf{A},\mathbf{E})$ and the reduced approximation (4.37). Also, consider the Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathrm{w}_{\mathrm{p}}}$ from (3.93), the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathrm{w}_{\mathrm{p}}}$ from (4.39), and the reduced Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathrm{w}_{\mathrm{p}},\mathrm{r}}$ from (4.39). Then, the L_∞-error between

the two outputs is bounded by

$$\begin{aligned}
&\|\mathbf{y}_{\mathrm{Q}}-\mathbf{y}_{\mathrm{Q,r}}\|_{L_\infty}^2\\
&\leq \Bigg(\operatorname{tr}\big(\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}\boldsymbol{\mathcal{M}}\big)-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},w_{\mathrm{p}}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},w_{\mathrm{p}}}\widehat{\mathbf{M}}_{11}\Big)+\operatorname{tr}\Big(\widehat{\mathbf{P}}_{1,w_{\mathrm{p}}}\widehat{\mathbf{M}}_{11}\widehat{\mathbf{P}}_{1,w_{\mathrm{p}}}\widehat{\mathbf{M}}_{11}\Big)\Bigg)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\cdot\left(\|\mathbf{u}\|_{L_2}^2+\|\zeta_0\|_2^2\right)^2\\
&+2\Bigg(\operatorname{tr}\big(\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{B}}\boldsymbol{\mathcal{M}}\big)-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},w_{\mathrm{p}}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\mathcal{B}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\Big)+\operatorname{tr}\Big(\widehat{\mathbf{P}}_{1,w_{\mathrm{p}}}\widehat{\mathbf{M}}_{12}\widehat{\mathbf{P}}_{2,\mathcal{B}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\Big)\Bigg)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\cdot\nu\|\mathbf{u}\|_{\mathfrak{C}^{\nu-1}}^2\left(\|\mathbf{u}\|_{L_2}^2+\|\zeta_0\|_2^2\right).
\end{aligned}\tag{4.46}$$

◊

Proof. We apply Theorem 4.13 and Theorem 4.16 to all the components of the output to obtain

$$\begin{aligned}
&\|\mathbf{y}_{\mathrm{Q}}-\mathbf{y}_{\mathrm{Q,r}}\|_{L_\infty}^2\\
&\leq \|\mathbf{y}_{\mathrm{pp},\mathcal{B}\mathcal{B}}-\mathbf{y}_{\mathrm{pp,r},\mathcal{B}\mathcal{B}}\|_{L_\infty}^2+2\|\mathbf{y}_{\mathrm{pp},\mathbf{z}_0\mathcal{B}}-\mathbf{y}_{\mathrm{pp,r},\mathbf{z}_0\mathcal{B}}\|_{L_\infty}^2\\
&\qquad+\|\mathbf{y}_{\mathrm{pp},\mathbf{z}_0\mathbf{z}_0}-\mathbf{y}_{\mathrm{pp,r},\mathbf{z}_0\mathbf{z}_0}\|_{L_\infty}^2+2\|\mathbf{y}_{\mathrm{ip},\mathcal{B}\mathcal{B}}-\mathbf{y}_{\mathrm{ip,r},\mathcal{B}\mathcal{B}}\|_{L_\infty}^2+4\|\mathbf{y}_{\mathrm{ip},\mathcal{B}\mathbf{z}_0}-\mathbf{y}_{\mathrm{ip,r},\mathcal{B}\mathbf{z}_0}\|_{L_\infty}^2\\
&\leq\Bigg(\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B}}\boldsymbol{\mathcal{M}})-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathcal{B}}\widehat{\mathbf{M}}_{11}\Big)+\operatorname{tr}\Big(\widehat{\mathbf{P}}_{1,\mathcal{B}}\widehat{\mathbf{M}}_{11}\widehat{\mathbf{P}}_{1,\mathcal{B}}\widehat{\mathbf{M}}_{11}\Big)\Bigg)\|\mathbf{u}\|_{L_2}^4\\
&+2\Bigg(\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}\boldsymbol{\mathcal{M}})-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}\widehat{\mathbf{M}}_{11}\Big)+\operatorname{tr}\Big(\widehat{\mathbf{P}}_{1,\mathcal{B}}\widehat{\mathbf{M}}_{11}\widehat{\mathbf{P}}_{1,\mathbf{z}_0}\widehat{\mathbf{M}}_{11}\Big)\Bigg)\|\zeta_0\|_2^2\|\mathbf{u}\|_{L_2}^2\\
&+\Bigg(\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}\boldsymbol{\mathcal{M}})-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}\widehat{\mathbf{M}}_{11}\Big)+\operatorname{tr}\Big(\widehat{\mathbf{P}}_{1,\mathbf{z}_0}\widehat{\mathbf{M}}_{11}\widehat{\mathbf{P}}_{1,\mathbf{z}_0}\widehat{\mathbf{M}}_{11}\Big)\Bigg)\|\zeta_0\|_2^4\\
&+2\Bigg(\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{B}}\boldsymbol{\mathcal{M}})-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\mathcal{B}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\Big)+\operatorname{tr}\Big(\widehat{\mathbf{P}}_{1,\mathcal{B}}\widehat{\mathbf{M}}_{12}\widehat{\mathbf{P}}_{2,\mathcal{B}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\Big)\Bigg)\nu\|\mathbf{u}\|_{\mathfrak{C}^{\nu-1}}^2\|\mathbf{u}\|_{L_2}^2\\
&+4\Bigg(\operatorname{tr}(\boldsymbol{\mathcal{P}}_{\mathrm{p}}\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{P}}_{\mathrm{i}}\boldsymbol{\mathcal{M}})-2\operatorname{tr}\Big(\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}^{\mathrm{T}}\boldsymbol{\mathcal{M}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\mathcal{B}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\Big)+\operatorname{tr}\Big(\widehat{\mathbf{P}}_{1,\mathbf{z}_0}\widehat{\mathbf{M}}_{12}\widehat{\mathbf{P}}_{2,\mathcal{B}}\widehat{\mathbf{M}}_{12}^{\mathrm{T}}\Big)\Bigg)\nu\|\mathbf{u}\|_{\mathfrak{C}^{\nu-1}}^2\|\zeta_0\|_2^2
\end{aligned}$$

for $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B}}$ from (3.62), $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathcal{B}}$ from (4.42), $\widehat{\mathbf{P}}_{1,\mathcal{B}}$ from (4.42), $\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}$ from (3.66), $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}$ from (4.34), $\widehat{\mathbf{P}}_{1,\mathbf{z}_0}$ from (4.34), $\boldsymbol{\mathcal{P}}_{\mathrm{i},\mathcal{B}}$ from (3.70), $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{i},\mathcal{B}}$ from (4.45), and $\widehat{\mathbf{P}}_{2,\mathcal{B}}$ from (4.45). Since $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}}}=\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathcal{B}}+\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{z}_0}$, $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},w_{\mathrm{p}}}=\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathcal{B}}+\widetilde{\boldsymbol{\mathcal{P}}}_{\mathrm{p},\mathbf{z}_0}$, and $\boldsymbol{\mathcal{P}}_{\mathrm{p},w_{\mathrm{p}},\mathrm{r}}=\boldsymbol{\mathcal{P}}_{\mathrm{p},\boldsymbol{\mathcal{B}},\mathrm{r}}+\boldsymbol{\mathcal{P}}_{\mathrm{p},\mathbf{Z}_0,\mathrm{r}}$ hold, this expression can be reduced to (4.46), which proves the statement. □

4.2.2 IRKA for inhomogeneous first-order DAE systems

In this subsection, we extend the IRKA method presented in Algorithm 5 to inhomogeneous DAE systems (3.54) with linear output equations. For that, we make use of the multi-system approach introduced in Section 3.2.1.1 and the extended-input approach from Section 3.2.1.2. We restrict this subsection to the case of linear output systems (3.54) since there exists no IRKA approach for systems with quadratic output equations. IRKA for homogeneous DAE systems was derived in [61]. Since the IRKA method is not the main topic of this work, we only consider the broad idea of these approaches.

4.2.2.1 Multi-system approach for inhomogeneous first-order DAE systems

In this paragraph, we extend the IRKA method to DAE systems presented in Algorithm 5 to systems (3.54) with inhomogeneous differential initial conditions. For that, we use the multi-system representation from Section 3.2.1.1 and consider the two proper subsystems (3.58) and (3.59) individually. We apply Algorithm 5 to these subsystems and derive two reduced surrogate models of the form (4.25) and (4.26) using the projecting bases

$$\boldsymbol{\mathcal{V}}_{\mathrm{r},\boldsymbol{\mathcal{B}}} = \begin{bmatrix}\boldsymbol{\mathcal{V}}_{N_f,\boldsymbol{\mathcal{B}}} & 0\end{bmatrix}, \quad \boldsymbol{\mathcal{T}}_{\mathrm{r},\boldsymbol{\mathcal{B}}} = \begin{bmatrix}\boldsymbol{\mathcal{T}}_{N_f,\boldsymbol{\mathcal{B}}} & 0\end{bmatrix}, \quad \boldsymbol{\mathcal{V}}_{\mathrm{r},\mathbf{Z}_0} = \begin{bmatrix}\boldsymbol{\mathcal{V}}_{N_f,\mathbf{Z}_0} & 0\end{bmatrix}, \quad \boldsymbol{\mathcal{T}}_{\mathrm{r},\mathbf{Z}_0} = \begin{bmatrix}\boldsymbol{\mathcal{T}}_{N_f,\mathbf{Z}_0} & 0\end{bmatrix}$$

with

$$\begin{aligned}
\boldsymbol{\mathcal{V}}_{N_f,\boldsymbol{\mathcal{B}}} &= \left[(\sigma_{1,\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathbf{P}_{\mathrm{l}}\boldsymbol{\mathcal{B}}\mathbf{b}_{1,\boldsymbol{\mathcal{B}}}, \dots, (\sigma_{R_{\boldsymbol{\mathcal{B}}},\boldsymbol{\mathcal{B}}}\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathbf{P}_{\mathrm{l}}\boldsymbol{\mathcal{B}}\mathbf{b}_{R_{\boldsymbol{\mathcal{B}}},\boldsymbol{\mathcal{B}}}\right], \\
\boldsymbol{\mathcal{V}}_{N_f,\mathbf{Z}_0} &= \left[(\sigma_{1,\mathbf{Z}_0}\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathbf{P}_{\mathrm{l}}\boldsymbol{\mathcal{E}}\mathbf{Z}_0\mathbf{b}_{1,\mathbf{Z}_0}, \dots, (\sigma_{R_{\mathbf{Z}_0},\mathbf{Z}_0}\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathbf{P}_{\mathrm{l}}\boldsymbol{\mathcal{E}}\mathbf{Z}_0\mathbf{b}_{R_{\mathbf{Z}_0},\mathbf{Z}_0}\right], \\
\boldsymbol{\mathcal{T}}_{N_f,*} &= \left[(\sigma_{1,*}\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\mathrm{H}}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{C}}^{\mathrm{H}}\mathbf{c}_{1,*}, \dots, (\sigma_{R,*}\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\mathrm{H}}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{C}}^{\mathrm{H}}\mathbf{c}_{R,*}\right]
\end{aligned}$$

for interpolation points $\sigma_{1,*}, \dots, \sigma_{R,*}$ and tangential directions $\mathbf{b}_{1,*}, \dots, \mathbf{b}_{R_*,*}$ and $\mathbf{c}_{1,*}, \dots, \mathbf{c}_{R,*}$ that are chosen individually for the two subsystems represented by the subscript $*$ that is either '$\boldsymbol{\mathcal{B}}$' or '$\mathbf{Z}_0$'. The reduced system (4.25) corresponding to $* =$'$\boldsymbol{\mathcal{B}}$' and the reduced system (4.26) corresponding to $* =$'$\mathbf{Z}_0$' are generated as described in (4.27).

Applying Algorithm 5 to the third subsystem (3.60) results in the bases $\boldsymbol{\mathcal{V}}_{\mathrm{i,r}} = \begin{bmatrix}0 & \boldsymbol{\mathcal{V}}_\infty\end{bmatrix}$ and $\boldsymbol{\mathcal{T}}_{\mathrm{i,r}} = \begin{bmatrix}0 & \boldsymbol{\mathcal{T}}_\infty\end{bmatrix}$, where $\boldsymbol{\mathcal{V}}_\infty$, $\boldsymbol{\mathcal{T}}_\infty$ are chosen so that they span the left and right deflating subspaces of $(\mathbf{A}, \mathbf{E})$ corresponding to $\lambda = \infty$. Generating the reduced system matrices according to (4.29) leads to the reduced system (4.28).

4.2.2.2 Extended-input approach for inhomogeneous first-order DAE systems

In this paragraph, the extended-input approach, introduced in Section 3.2.1.2, is used to apply the IRKA method from Algorithm 5 to DAE systems with inhomogeneous differential initial conditions. For that, we consider the system (3.91) instead of the original one (3.54) as it has the same input-to-output behavior in the frequency domain but also is of the system structure introduced in (2.8) so that Algorithm 5 is applicable. Applying that method to the system (3.91) leads to the bases

$$\boldsymbol{\mathcal{V}}_{\rm r} = \begin{bmatrix}\boldsymbol{\mathcal{V}}_{N_f} & \boldsymbol{\mathcal{V}}_{\infty}\end{bmatrix}, \qquad \boldsymbol{\mathcal{T}}_{\rm r} = \begin{bmatrix}\boldsymbol{\mathcal{T}}_{N_f} & \boldsymbol{\mathcal{T}}_{\infty}\end{bmatrix}$$

with

$$\begin{aligned}\boldsymbol{\mathcal{V}}_{N_f} &= \left[(\sigma_1\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathbf{P}_{\rm l}\boldsymbol{\mathcal{W}}_{\rm p}\mathbf{b}_1, \ldots, (\sigma_R\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-1}\mathbf{P}_{\rm l}\boldsymbol{\mathcal{W}}_{\rm p}\mathbf{b}_R\right],\\ \boldsymbol{\mathcal{T}}_{N_f} &= \left[(\sigma_1\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\rm H}\mathbf{P}_{\rm r}^{\rm T}\boldsymbol{\mathcal{C}}^{\rm H}\mathbf{c}_1, \ldots, (\sigma_R\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}})^{-\rm H}\mathbf{P}_{\rm r}^{\rm T}\boldsymbol{\mathcal{C}}^{\rm H}\mathbf{c}_R\right]\end{aligned}$$

for interpolation points $\sigma_1, \ldots, \sigma_R$ and tangential directions $\mathbf{b}_1, \ldots, \mathbf{b}_R$ and $\mathbf{c}_1, \ldots, \mathbf{c}_R$. Again, $\boldsymbol{\mathcal{V}}_{\infty}$ and $\boldsymbol{\mathcal{T}}_{\infty}$ are chosen so that they span the left and right deflating subspaces of $(\mathbf{A}, \mathbf{E})$ corresponding to $\lambda = \infty$. We multiply the system matrices of the original system (3.54) from the left and the right by the projecting bases $\boldsymbol{\mathcal{V}}_{\rm r}$ and $\boldsymbol{\mathcal{T}}_{\rm r}$ according to (4.38) to derive the reduced system (4.36) that approximates the input to output behavior of the original one.

4.2.3 Numerical results

In this section, we discuss the efficiency of the proposed methodology using several examples. For that, we focus on the BT methods for DAE systems (3.100) with quadratic output equations as they are the main focus of this section and the most challenging system structure considered. We apply our BT methods to systems with homogeneous and inhomogeneous initial conditions. First, we introduce a homogeneous example of dimension four and show that the mixed Gramians containing the proper and improper controllability and observability space information are required to approximate the system behavior. Afterwards, we consider an inhomogeneous example of index 2, which takes into account the input and the initial condition space to reduce the respective system. Finally, we consider a homogeneous example of index 3, which describes a mechanical system with additional constraints.

We also verify our theoretical findings in our numerical experiments,e.g., the error bounds. 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 R2019a and examples and methods from M-M.E.S.S.-2.1., see [114]. All results are available at [104].

4.2.3.1 Example 1: an illustrative example

First, we introduce a small toy example with homogeneous initial conditions to highlight that we need to consider mixed Gramians $\boldsymbol{\mathfrak{Q}}_{\rm pi,w_p}$ and $\boldsymbol{\mathfrak{Q}}_{\rm ip,w_p}$, as introduced in (3.109) and (3.112) when considering systems with a quadratic output equation. For this, we consider the following system in Weierstraß canonical form

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

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

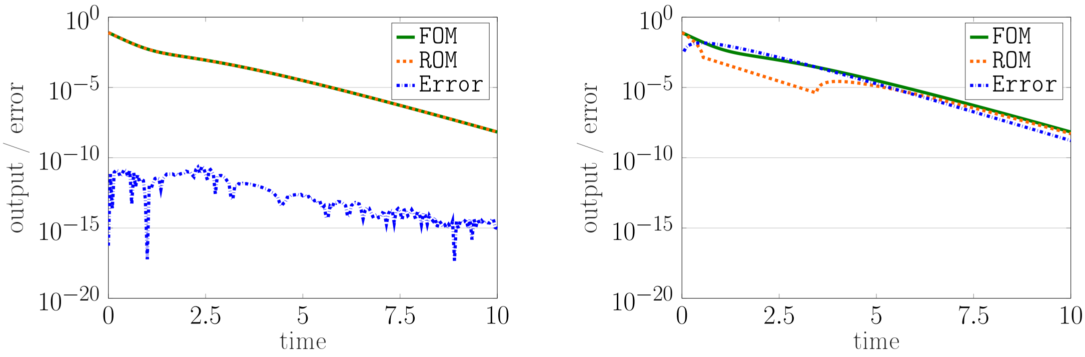
The proper state is then given by $\mathbf{z}_1(t) = \left[\begin{smallmatrix} z_1(t) \\ z_2(t) \end{smallmatrix}\right]$ and the improper one as $\mathbf{z}_2(t) = \left[\begin{smallmatrix} z_3(t) \\ z_4(t) \end{smallmatrix}\right]$. The corresponding system Gramians are

$$\mathbf{P}_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \qquad \mathbf{P}_2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

$$\mathbf{Q}_{11} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathfrak{Q}_{21} = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \qquad \mathfrak{Q}_{12} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}, \qquad \mathfrak{Q}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$$

as defined in (3.95), (3.107), (3.110), (3.113), and (3.115). We note that the proper controllability Gramian has rank one. Therefore, the minimal realization of the differential part of the system is also of rank one, and so is the differential part of the reduced order model for this example. The improper state is described by a rank two controllability Gramian and a rank two observability Gramian, namely $\mathfrak{Q}_{\mathrm{Q,i},\mathfrak{B}} = \mathfrak{Q}_{\mathrm{pi},\mathfrak{B}} + \mathfrak{Q}_{\mathrm{ii},\mathfrak{B}}$, so the minimal realization of the improper part of the system is of rank two. However, we note that the Gramian of improper-improper observability Gramian is of rank one. This fact vividly shows that the mixed Gramians must be taken into account.

To investigate the quality of the reduced surrogate system, we consider the system output obtained by applying the input function $\mathbf{u}(t) = 0.2 \cdot e^{-t}$. The results are shown in Figure 4.1, 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., $\mathfrak{Q}_{\mathrm{Q,p},\mathfrak{B}} := \mathfrak{Q}_{\mathrm{pp},\mathfrak{B}}$ and $\mathfrak{Q}_{\mathrm{Q,i},\mathfrak{B}} := \mathfrak{Q}_{\mathrm{ii},\mathfrak{B}}$. We observe that the mixed observability Gramians $\mathfrak{Q}_{\mathrm{pi},\mathfrak{B}}$ and $\mathfrak{Q}_{\mathrm{ip},\mathfrak{B}}$ must be considered within the reduction process.



(a) Mixed Gramians are used.

(b) Mixed Gramians are not used.

Figure 4.1: Example 1 - Output responses and the corresponding errors.

4.2.3.2 Example 2: an index-2 Stokes example

We consider the creeping flow in capillaries or porous media described by the following equations

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} v(\zeta,t) &= \mu \Delta v(\zeta,t) - \nabla p(\zeta,t) + f(\zeta,t), \\ 0 &= \operatorname{div}(v(\zeta,t)), \end{aligned} \tag{4.47}$$

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 system (4.47) by a finite difference scheme as discussed in [91, 131] and add an output equation to measure our quantity of interest. We choose the matrix $\mathcal{M}$ to be $0.01 \cdot \mathbf{I}_N$, yielding the ℓ_2-norm of the state vector with a scaling factor 0.01, so that we obtain a discretized system of the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \mathbf{I} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \mathbf{z}(t) \\ \boldsymbol{\lambda}(t) \end{bmatrix} &= \begin{bmatrix} \mathbf{A} & \mathbf{G} \\ \mathbf{G}^{\mathrm{T}} & 0 \end{bmatrix}\begin{bmatrix} \mathbf{z}(t) \\ \boldsymbol{\lambda}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix}\mathbf{u}(t), \qquad \begin{bmatrix} \mathbf{z}(0) \\ \boldsymbol{\lambda}(0) \end{bmatrix} = \begin{bmatrix} \mathbf{z}_0 \\ 0 \end{bmatrix}, \\ \mathbf{y}_{\mathrm{Q}}(t) &= \begin{bmatrix} \mathbf{z}(t)^{\mathrm{T}} & \boldsymbol{\lambda}(t)^{\mathrm{T}} \end{bmatrix} \mathcal{M} \begin{bmatrix} \mathbf{z}(t) \\ \boldsymbol{\lambda}(t) \end{bmatrix} \end{aligned} \tag{4.48}$$

with system matrices $\mathbf{A} \in \mathbb{R}^{N_v \times N_v}$, $\mathbf{G} \in \mathbb{R}^{N_v \times N_p}$, and the initial state value $\mathbf{z}_0 \in \mathbb{R}^{N_v \times 1}$. The input matrices are given as $\mathbf{B}_1 \in \mathbb{R}^{N_v \times m}$, $\mathbf{B}_2 \in \mathbb{R}^{N_p \times m}$ and the output matrix is $\mathcal{M} \in \mathbb{R}^{N \times N}$ with $N = N_v + N_p$. The state consists of $\mathbf{z}(t) \in \mathbb{R}^{N_v}$ and $\boldsymbol{\lambda}(t) \in \mathbb{R}^{N_p}$, while the input is $\mathbf{u}(t) \in \mathbb{R}^m$ and the output is $\mathbf{y}_{\mathrm{Q}}(t) \in \mathbb{R}$. 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.

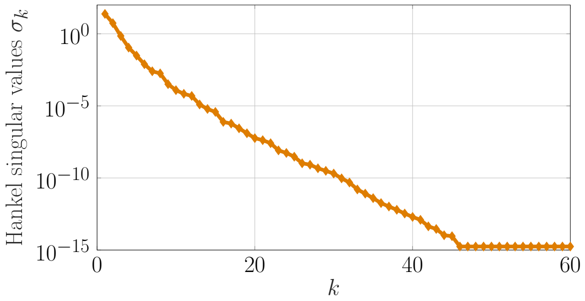


Figure 4.2: Example 2 - Decay of proper Hankel singular values.

As shown in [131], the projection matrices from (2.10) are given as

$$\mathbf{P}_\mathrm{l} = \mathbf{P}_\mathrm{r}^\mathrm{T} = \begin{bmatrix} \Pi & -\Pi\mathbf{A}\mathbf{G}(\mathbf{G}^\mathrm{T}\mathbf{G})^{-1} \\ 0 & 0 \end{bmatrix}$$

where

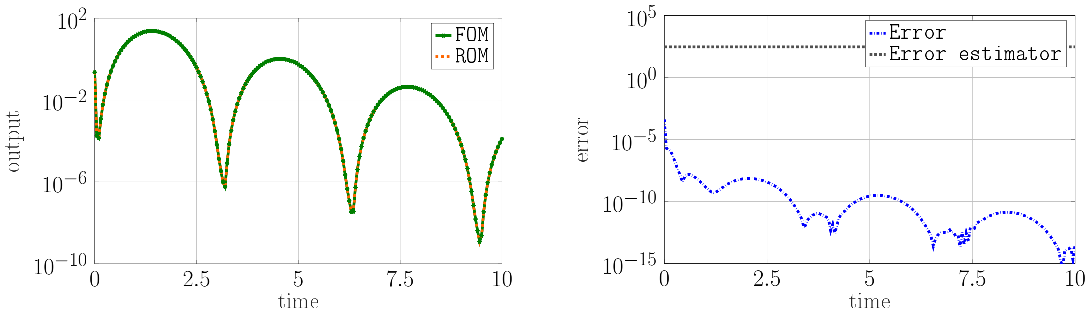
$$\Pi = \mathbf{I}_{N_v} - \mathbf{G}\left(\mathbf{G}^\mathrm{T}\mathbf{G}\right)^{-1}\mathbf{G}^\mathrm{T}.$$

The initial value is chosen to be $\mathbf{z}_0 = \mathbf{Z}_0 = (\Pi \cdot \mathbf{1}_{N_v \times 1})/\|\Pi \cdot \mathbf{1}_{N_v \times 1}\|_2$, where $\mathbf{1}_{N_v \times 1}$ is the vector containing one-values on every entry. That choice for the initial condition $\mathbf{z}_0$ leads to a purely proper initial condition, i.e., $\mathbf{z}_0 = \Pi \cdot \mathbf{z}_0$, while the improper component $(\mathbf{I}_{N_v} - \Pi) \cdot \mathbf{z}_0 = 0$ is equal to zero.

We need to determine the Gramians corresponding to the proper and improper states of the system (4.48). For this purpose, we apply the methods described in [131, 133], noting that the improper Gramians can be computed explicitly. In Figure 4.2, we depict the decay of the Hankel singular values $\sigma_1, \sigma_2, \dots$ corresponding to the proper Gramians $\mathcal{P}_{\mathrm{p},\mathrm{w}_\mathrm{p}}$ and $\mathcal{Q}_{\mathrm{Q},\mathrm{p},\mathrm{w}_\mathrm{p}}$ as described in (3.93) and (3.111), respectively. 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 dimension $R = R_v + R_p$ with $R_v = 18$ and $R_p = 2$. Figure 4.3 shows the output behavior of the full-order model (3.100) and of the reduced-order model (4.37) for an input function $\mathbf{u}(t) = \sin(t)^3 e^{-t/2}$. Additionally, the figure includes the output error and the corresponding error bound. The actual error is below the estimated error for all time, and we observe that the error bound is rather conservative. The error is sufficiently small, and the approximation quality of the reduced-order systems is much better than the estimated one.

4.2.3.3 Example 3: an index-3 mechanical system

Now, we investigate an index-3 system that results from mechanical systems Figure 4.4, which is of specific interest in this work. It is of the form



(a) Output (b) Error

Figure 4.3: Example 2 - Output responses and the corresponding errors.

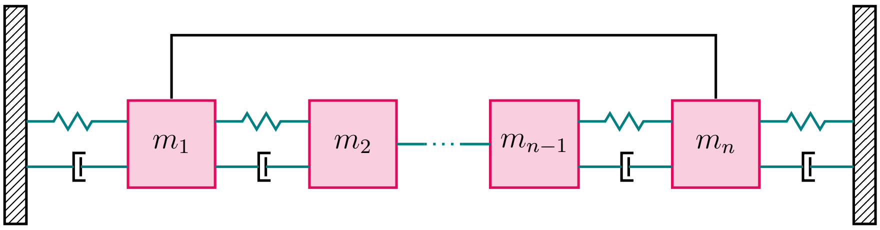


Figure 4.4: Example 3 - Sketch of a mechanical example with one row of masses connected with consecutive springs and one stiff connection between the first and last mass.

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\mathbf{I}_{n_x} & 0 & 0\\ 0 & \mathbf{M} & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\mathbf{x}_1(t)\\ \mathbf{x}_2(t)\\ \boldsymbol{\lambda}(t)\end{bmatrix} = \begin{bmatrix}0 & \mathbf{I}_{n_x} & 0\\ -\mathbf{K} & -\mathbf{D} & \mathbf{G}\\ \mathbf{G}^{\mathrm{T}} & 0 & 0\end{bmatrix}\begin{bmatrix}\mathbf{x}_1(t)\\ \mathbf{x}_2(t)\\ \boldsymbol{\lambda}(t)\end{bmatrix} + \begin{bmatrix}0\\ \mathbf{B}_x\\ 0\end{bmatrix}\mathbf{u}(t), \qquad \begin{bmatrix}\mathbf{x}_1(0)\\ \mathbf{x}_2(0)\\ \boldsymbol{\lambda}(0)\end{bmatrix} = \begin{bmatrix}\mathbf{x}_0\\ \mathbf{x}_0\\ 0\end{bmatrix}$$

$$\mathbf{y}_{\mathrm{Q}}(t) = \begin{bmatrix}\mathbf{x}_1(t)^{\mathrm{T}} & \mathbf{x}_2(t)^{\mathrm{T}} & \boldsymbol{\lambda}(t)^{\mathrm{T}}\end{bmatrix}\boldsymbol{\mathcal{M}}\begin{bmatrix}\mathbf{x}_1(t)\\ \mathbf{x}_2(t)\\ \boldsymbol{\lambda}(t)\end{bmatrix}, \tag{4.49}$$

where $\mathbf{M}$, $\mathbf{D}$, $\mathbf{K} \in \mathbb{R}^{g\times g}$, $\mathbf{B}_x \in \mathbb{R}^{g\times m}$, $\mathbf{G} \in \mathbb{R}^{g\times q}$, and $\boldsymbol{\mathcal{M}} \in \mathbb{R}^{(2g+q)\times(2g+q)}$. The state is given by $\mathbf{x}_1(t)$, $\mathbf{x}_2(t) \in \mathbb{R}^{g}$, $\boldsymbol{\lambda}(t) \in \mathbb{R}^{q}$, the input by $\mathbf{u}(t) \in \mathbb{R}^{m}$ and the output by $\mathbf{y}_{\mathrm{Q}}(t) \in \mathbb{R}$. We consider the index-3 system (4.49), which arises in the modeling of constraint mechanical systems with matrices

$$\begin{aligned}
\mathbf{M} &= \operatorname{diag}(m_1, \ldots, m_g),\\
\mathbf{D} &= \begin{bmatrix} d_1+\delta_1 & -d_1 & & & \\ -d_1 & d_1+d_2+\delta_2 & -d_2 & & \\ & \ddots & \ddots & \ddots & \\ & & -d_{g-2} & d_{g-2}+d_{g-1}+\delta_{g-1} & -d_{g-1} \\ & & & -d_{g-1} & d_{g-1}+\delta_g \end{bmatrix},\\
\mathbf{K} &= \begin{bmatrix} k_1+\kappa_1 & -k_1 & & & \\ -k_1 & k_1+k_2+\kappa_2 & -k_2 & & \\ & \ddots & \ddots & \ddots & \\ & & -k_{g-2} & k_{g-2}+k_{g-1}+\kappa_{g-1} & -k_{g-1} \\ & & & -k_{g-1} & k_{g-1}+\kappa_g \end{bmatrix},
\end{aligned}$$

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

The matrices are generated using the M-M.E.S.S. function `msd_ind3`, see [114], with dimension $g = 600$. We choose

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

The projection matrices (2.10) for this example were introduced in [91]. To compute the Gramians, we follow the same procedure, as presented in [131, 133] modified to the index-3 case. We assume zero-initial conditions.

Figure 4.5 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 that are zero. The resulting reduced dimension is $R = R_v + R_p$ with $R_v = 20$ and $R_p = 1$. The outputs of the full-order model (3.100) and the reduced-order model (4.37) are described in Figure 4.6 for an input function $\mathbf{u}(t) = \sin(2t)^2 e^{-t/2}$. The figure also shows the error between the outputs and the error bound using (4.46). We observe that the output error, which is smaller than 10^{-13} for all $t \in [0,\, 10]$, is sufficiently small and that the error bound is rather conservative.

4.3 Model order reduction for inhomogeneous second-order ODE systems

In this section, we aim to reduce second-order systems of the structure presented in (3.121) and (3.154). One possible approach is to transform these systems into systems

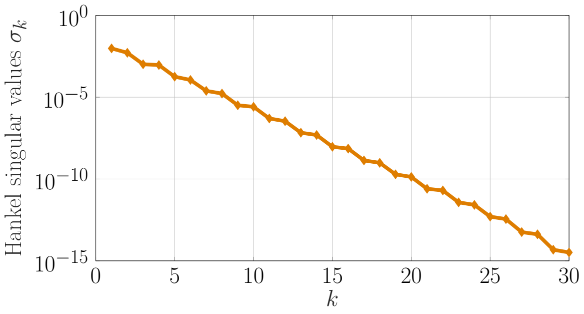
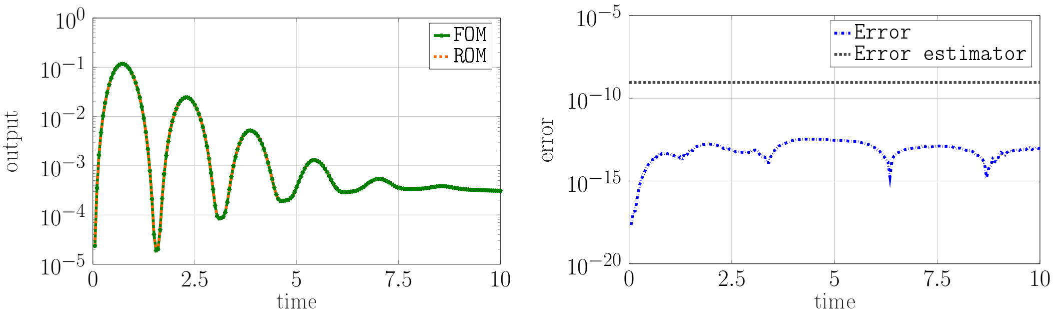


Figure 4.5: Example 3 - Decay of proper Hankel singular values.



(a) Output

(b) Error

Figure 4.6: Example 3 - Output responses and the corresponding errors.

of first-order structure (3.5) and (3.31), and evaluate the behavior of these representations as shown in Section 4.1. However, reducing the first-order systems does not maintain the second-order structure, so the reduced first-order systems might be physically meaningless. Also, a first-order system is generally not transferable into a second-order representation. On the other hand, having a reduced system of second-order structure allows us a meaningful physical interpretation and is therefore desired as described in [111]. Hence, we introduce BT methods for inhomogeneous second-order systems with linear and quadratic output equations.

First, in Section 4.3.1, the BT method for second-order systems is introduced, and afterward, in Section 4.3.2, these methods are evaluated by applying them to some numerical examples.

4.3.1 BT for inhomogeneous second-order ODE systems

BT for second-order systems was derived in [44, 112] for systems (3.121) with linear output equation and homogeneous initial conditions. In this subsection, we extend this method to systems with inhomogeneous initial conditions and to systems with quadratic output equations, i.e., we consider second-order systems (3.121) and (3.154). Therefore, we use the different system representations and the respective tailored Gramians presented in Section 3.3 to construct reduced second-order models via BT.

We first use the multi-system approach in Section 4.3.1.1 to derive surrogate models corresponding to the system (3.121) with a linear output equation. As described in Section 3.3.2.1, applying the multi-system approach for the system (3.154) with a quadratic output equation would lead to 9 subsystems, which makes this approach numerically prohibitive. In Section 4.3.1.2, we utilize the extended-input approach to derive reduced surrogate systems for both systems structures (3.121) and (3.154).

4.3.1.1 Multi-system approach for inhomogeneous second-order ODE systems

To reduce the second-order system (3.121) with a linear output equation while considering the initial conditions, we utilize the superposition properties of this system. Since the input- and initial condition-to-output behavior is represented by the subsystems (3.124), (3.125), and (3.126) as shown in Section 3.3.1.1, we reduce these subsystems separately.

We aim to derive the reduced surrogate system

$$\begin{aligned}\mathbf{M}_{\mathrm{r},\mathbf{B}}\ddot{\mathbf{x}}_{\mathrm{r}}(t)+\mathbf{D}_{\mathrm{r},\mathbf{B}}\dot{\mathbf{x}}_{\mathrm{r}}(t)+\mathbf{K}_{\mathrm{r},\mathbf{B}}\mathbf{x}_{\mathrm{r}}(t)=\mathbf{B}_{\mathrm{r},\mathbf{B}}\mathbf{u}(t),\qquad \mathbf{x}_{\mathrm{r}}(0)=0,\quad \dot{\mathbf{x}}_{\mathrm{r}}(0)=0,\\ \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}(t)=\mathbf{C}_{1,\mathrm{r},\mathbf{B}}\mathbf{x}_{\mathrm{r}}(t)\dot{\mathbf{x}}_{\mathrm{r}}(t),\end{aligned}\tag{4.50}$$

that approximated the input-to-output behavior of the homogeneous subsystem (3.124). The subsystem (3.125) that corresponds to the position initial condition is approximated by the surrogate model

$$\begin{aligned} \mathbf{M}_{\mathrm{r},\mathbf{x}_0}\ddot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{D}_{\mathrm{r},\mathbf{x}_0}\dot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{K}_{\mathrm{r},\mathbf{x}_0}\mathbf{x}(t) = 0, \qquad \mathbf{x}_{\mathrm{r}}(0) = \mathbf{X}_{0,\mathrm{r},\mathbf{x}_0}\chi_0, \quad \dot{\mathbf{x}}_{\mathrm{r}}(0) = 0,& \\ \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{x}_0}(t) = \mathbf{C}_{1,\mathrm{r},\mathbf{x}_0}\mathbf{x}_{\mathrm{r}}(t).& \end{aligned} \tag{4.51}$$

Finally, the subsystem (3.126) corresponding to the velocity initial condition shall be approximated by a reduced system of the structure

$$\begin{aligned} \mathbf{M}_{\mathrm{r},\mathbf{v}_0}\ddot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{D}_{\mathrm{r},\mathbf{v}_0}\dot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{K}_{\mathrm{r},\mathbf{v}_0}\mathbf{x}_{\mathrm{r}}(t) = 0, \qquad \mathbf{x}_{\mathrm{r}}(0) = 0, \quad \dot{\mathbf{x}}_{\mathrm{r}}(0) = \mathbf{V}_{0,\mathrm{r},\mathbf{v}_0}\nu_0,& \\ \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{v}_0}(t) = \mathbf{C}_{1,\mathrm{r},\mathbf{v}_0}\mathbf{x}(t).& \end{aligned} \tag{4.52}$$

We aim to find such subsystems that approximate the output $\mathbf{y}_{\mathrm{L}}(t)$ as

$$\mathbf{y}_{\mathrm{L}}(t) \approx \mathbf{y}_{\mathrm{L},\mathrm{r}}(t) = \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}(t) + \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{x}_0}(t) + \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{v}_0}(t).$$

Therefore, we generate the respectively reduced system matrices of the three subsystems using projecting matrices $\mathbf{V}_{\mathrm{r},*}, \mathbf{T}_{\mathrm{r},*} \in \mathbb{R}^{n \times r_*}$ with $r_* \ll n$, where the subscript $*$ is equal to '$\mathbf{B}$', '$\mathbf{x}_0$', or '$\mathbf{v}_0$', so that

$$\begin{aligned} &\mathbf{M}_{\mathrm{r},*} = \mathbf{V}_{\mathrm{r},*}^{\mathrm{T}}\mathbf{M}\mathbf{T}_{\mathrm{r},*}, \quad &&\mathbf{D}_{\mathrm{r},*} = \mathbf{V}_{\mathrm{r},*}^{\mathrm{T}}\mathbf{D}\mathbf{T}_{\mathrm{r},*}, \quad &&\mathbf{K}_{\mathrm{r},*} = \mathbf{V}_{\mathrm{r},*}^{\mathrm{T}}\mathbf{K}\mathbf{T}_{\mathrm{r},*}, && \\ &\mathbf{B}_{\mathrm{r},\mathbf{B}} = \mathbf{V}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\mathbf{B}, \quad &&\mathbf{C}_{1,\mathrm{r},*} = \mathbf{C}_1\mathbf{T}_{\mathrm{r},*}, \quad &&\mathbf{X}_{0,\mathrm{r}} = \mathbf{V}_{\mathrm{r},\mathbf{x}_0}^{\mathrm{T}}\mathbf{X}_0, \quad &&\mathbf{V}_{0,\mathrm{r},\mathbf{v}_0} = \mathbf{V}_{\mathrm{r},\mathbf{v}_0}^{\mathrm{T}}\mathbf{V}_0. \end{aligned}$$

We aim to apply the BT method for homogeneous systems from Algorithm 3 to generate these projecting matrices. However, two of the three subsystems have inhomogeneous initial conditions. On the other hand, the Gramians and system energies summarized in Table 3.8 have the same structure as for the homogeneous subsystem (3.124). Hence, the BT method for homogeneous second-order systems from Algorithm 3 can be applied to the subsystems (3.124), (3.125), and (3.126) using the suitable Gramians. To do so, we consider the controllability Gramian $\mathbf{P}_*$ that is equal to $\mathbf{P}_{\mathbf{B}}$, $\mathbf{P}_{\mathbf{x}_0}$, or $\mathbf{P}_{\mathbf{v}_0}$, and the observability Gramian $\mathbf{Q}_{\mathrm{L}}$ depending on the considered subsystem according to Table 3.8. As shown in (3.140), (3.141), and (3.143), the states corresponding to the large eigenvalues of the respective controllability Gramians $\mathbf{P}_*$ and the observability Gramian $\mathbf{Q}_{\mathrm{L}}$ span the most dominant controllability and observability subspaces while those corresponding to small eigenvalues are neglectable. Hence, we balance the system to derive $\mathbf{P}_* = \mathbf{Q}_{\mathrm{L}}$ and truncate the least important subspaces within the BT method. We compute the respective SVD

$$\mathbf{S}^{\mathrm{T}}\mathbf{M}\mathbf{R}_* = \begin{bmatrix} \mathbf{U}_{1,*} & \mathbf{U}_{2,*} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{1,*} & \\ & \boldsymbol{\Sigma}_{2,*} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{1,*}^{\mathrm{T}} \\ \mathbf{V}_{2,*}^{\mathrm{T}} \end{bmatrix}$$

for $\mathbf{P}_* = \mathbf{R}_*\mathbf{R}_*^{\mathrm{T}}$ and $\mathbf{Q}_{\mathrm{L}} = \mathbf{S}\mathbf{S}^{\mathrm{T}}$. Then, the projecting matrices are defined as

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

```
Algorithm 12 BT method for the second-order ODE system (3.121) with a linear output using the multi-system approach.
Require: The original system (3.121), the reduced dimensions $r_*$, where $* =$'$\mathbf{B}$', '$\mathbf{X}_0$', '$\mathbf{V}_0$'.
Ensure: The reduced systems (4.50), (4.51), and (4.52).
1: Compute factors of the Gramians $\mathbf{P}_* \approx \mathbf{R}_*\mathbf{R}_*^{\mathrm{T}}$ and $\mathbf{Q}_\circ \approx \mathbf{S}\mathbf{S}^{\mathrm{T}}$ with $* =$'$\mathbf{B}$', '$\mathbf{X}_0$', or '$\mathbf{V}_0$' according to Table 3.8.
2: Perform the SVD of $\mathbf{S}^{\mathrm{T}}\mathbf{M}\mathbf{R}_*$ and decompose as
       $$\mathbf{S}^{\mathrm{T}}\mathbf{M}\mathbf{R}_* = \begin{bmatrix}\mathbf{U}_{1,*} & \mathbf{U}_{2,*}\end{bmatrix} \begin{bmatrix}\mathbf{\Sigma}_{1,*} & \\ & \mathbf{\Sigma}_{2,*}\end{bmatrix} \begin{bmatrix}\mathbf{V}_{1,*}^{\mathrm{T}} \\ \mathbf{V}_{2,*}^{\mathrm{T}}\end{bmatrix}$$
   with $\mathbf{\Sigma}_{1,*} \in \mathbb{R}^{r_* \times r_*}$.
3: Construct the projecting matrices
       $$\mathbf{V}_{*,\mathrm{r}} = \mathbf{S}\mathbf{U}_{1,*}\mathbf{\Sigma}_{1,*}^{-\frac{1}{2}} \text{ and } \mathbf{T}_{*,\mathrm{r}} = \mathbf{R}_*\mathbf{V}_{1,*}\mathbf{\Sigma}_{1,*}^{-\frac{1}{2}}.$$
4: Determine the reduced matrices (4.3.1.1).
```

Using these bases, we derive the reduced surrogate systems (4.50), (4.51), and (4.52) with the respective reduced matrices defined in (4.3.1.1). The detailed reduction procedure for each subsystem is given in Algorithm 12.

To develop an a posteriori error bound for the respective output error, we use the output error decomposition

$$\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_\infty} \leq \|\mathbf{y}_{\mathrm{L},\mathbf{B}} - \mathbf{y}_{\mathrm{L,r},\mathbf{B}}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{L},\mathbf{x}_0} - \mathbf{y}_{\mathrm{L,r},\mathbf{x}_0}\|_{L_\infty} + \|\mathbf{y}_{\mathrm{L},\mathbf{v}_0} - \mathbf{y}_{\mathrm{L,r},\mathbf{v}_0}\|_{L_\infty} \tag{4.53}$$

and analyze the three error norms separately. For that, we make use of the first-order matrices from (2.24) and the reduced first-order matrices

$$\begin{aligned} \boldsymbol{\mathcal{E}}_{\mathrm{r},*} &:= \begin{bmatrix}\mathbf{I}_{r_*} & 0 \\ 0 & \mathbf{M}_{\mathrm{r},*}\end{bmatrix}, & \boldsymbol{\mathcal{A}}_{\mathrm{r},*} &:= \begin{bmatrix}0 & \mathbf{I}_{r_*} \\ -\mathbf{K}_{\mathrm{r},*} & -\mathbf{D}_{\mathrm{r},*}\end{bmatrix}, & \boldsymbol{\mathcal{C}} &:= \begin{bmatrix}\mathbf{C}_{1,\mathrm{r},*} & 0\end{bmatrix}, \\ \boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{B}} &:= \begin{bmatrix}0 \\ \mathbf{B}_{\mathrm{r},\mathbf{B}}\end{bmatrix}, & \mathbf{Z}_{0,\mathrm{r},\mathbf{x}_0} &:= \begin{bmatrix}\mathbf{X}_{0,\mathrm{r},\mathbf{x}_0} \\ 0\end{bmatrix}, & \mathbf{Z}_{0,\mathrm{r},\mathbf{v}_0} &:= \begin{bmatrix}0 \\ \mathbf{V}_{0,\mathrm{r}}\end{bmatrix}, \end{aligned} \tag{4.54}$$

with $*$ equal to '$\mathbf{B}$', '$\mathbf{X}_0$', or '$\mathbf{V}_0$'.

To derive a bound for the first error component $\|\mathbf{y}_{\mathrm{L},\mathbf{B}} - \mathbf{y}_{\mathrm{L,r},\mathbf{B}}\|_{L_\infty}$, we define the mappings

$$\mathbf{h}_{\mathbf{B}}(t) := \boldsymbol{\mathcal{C}} e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}} \qquad \text{and} \qquad \widehat{\mathbf{h}}_{\mathbf{B}}(t) := \boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}} e^{\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-1}\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{B}}t}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-1}\boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{B}}, \tag{4.55}$$

so that we can rewrite the respective outputs as

$$\mathbf{y}_{\mathrm{L},\mathbf{B}}(t) = \int_0^t \mathbf{h}_{\mathbf{B}}(t-\tau)\mathbf{u}(\tau)\mathrm{d}\tau \text{ and } \qquad \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}(t) = \int_0^t \widehat{\mathbf{h}}_{\mathbf{B}}(t-\tau)\mathbf{u}(\tau)\mathrm{d}\tau.$$

These representations of $\mathbf{y}_{\mathrm{L},\mathbf{B}}$ and $\mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}$ are used in the following lemma to derive an upper bound of the respective L_∞-error.

Lemma 4.18:
Consider the asymptotically stable second-order system (3.124) with corresponding first-order matrices as defined in (2.24), the reduced system (4.50) with corresponding reduced first-order matrices as defined in (4.54), and the mappings $\mathbf{h}_{\mathbf{B}}$, $\widehat{\mathbf{h}}_{\mathbf{B}}$ as defined in (4.55). Then, the following bound holds

$$\|\mathbf{y}_{\mathrm{L},\mathbf{B}} - \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}\|_{L_\infty} \leq \left(\int_0^\infty \left\|\mathbf{h}_{\mathbf{B}}(t) - \widehat{\mathbf{h}}_{\mathbf{B}}(t)\right\|_{\mathrm{F}}^2 \mathrm{d}t\right)^{\frac{1}{2}} \|\mathbf{u}\|_{L_2}. \qquad \Diamond$$

Proof. We consider the 2-norm of the output error at time $t \geq 0$ that is

$$\left\|\mathbf{y}_{\mathrm{L},\mathbf{B}}(t) - \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}(t)\right\|_2 = \left\|\int_0^t \left(\mathbf{h}_{\mathbf{B}}(t-\tau) - \widehat{\mathbf{h}}_{\mathbf{B}}(t-\tau)\right)\mathbf{u}(\tau)\mathrm{d}\tau\right\|_2.$$

Applying the Cauchy-Schwarz inequality multiple times yields

$$\begin{aligned}
\left\|\mathbf{y}_{\mathrm{L},\mathbf{B}}(t) - \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}(t)\right\|_2 &\leq \int_0^t \left\|\left(\mathbf{h}_{\mathbf{B}}(t-\tau) - \widehat{\mathbf{h}}_{\mathbf{B}}(t-\tau)\right)\mathbf{u}(\tau)\right\|_2 \mathrm{d}\tau \\
&\leq \int_0^t \left\|\mathbf{h}_{\mathbf{B}}(t-\tau) - \widehat{\mathbf{h}}_{\mathbf{B}}(t-\tau)\right\|_2 \|\mathbf{u}(\tau)\|_2 \mathrm{d}t \\
&\leq \left(\int_0^t \left\|\mathbf{h}_{\mathbf{B}}(t-\tau) - \widehat{\mathbf{h}}_{\mathbf{B}}(t-\tau)\right\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}} \left(\int_0^t \|\mathbf{u}(\tau)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}}.
\end{aligned}$$

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

$$\begin{aligned}
\|\mathbf{y}_{\mathrm{L},\mathbf{B}} - \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}\|_{L_\infty} &\leq \left(\int_0^\infty \left\|\mathbf{h}_{\mathbf{B}}(t) - \widehat{\mathbf{h}}_{\mathbf{B}}(t)\right\|_2^2 \mathrm{d}t\right)^{\frac{1}{2}} \left(\int_0^\infty \|\mathbf{u}(\tau)\|_2^2 \mathrm{d}\tau\right)^{\frac{1}{2}} \\
&\leq \left(\int_0^\infty \left\|\mathbf{h}_{\mathbf{B}}(t) - \widehat{\mathbf{h}}_{\mathbf{B}}(t)\right\|_{\mathrm{F}}^2 \mathrm{d}t\right)^{\frac{1}{2}} \|\mathbf{u}\|_{L_2}. \qquad \square
\end{aligned}$$

The bound presented in Lemma 4.18 includes the expression

$$\int_0^\infty \left\|\mathbf{h}_{\mathbf{B}}(t) - \widehat{\mathbf{h}}_{\mathbf{B}}(t)\right\|_{\mathrm{F}}^2 \mathrm{d}t = \int_0^\infty \|\mathbf{h}_{\mathbf{B}}(t)\|_{\mathrm{F}}^2 - 2\langle \operatorname{vec}(\mathbf{h}_{\mathbf{B}}(t)), \operatorname{vec}(\widehat{\mathbf{h}}_{\mathbf{B}}(t))\rangle + \left\|\widehat{\mathbf{h}}_{\mathbf{B}}(t)\right\|_{\mathrm{F}}^2 \mathrm{d}t.$$

The following lemma bounds the different components of this bound. Therefore, we define the following reduced Gramian and the matrix

$$
\begin{aligned}
\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{B}} &:= \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-\mathrm{T}}t}\mathrm{d}t,\\
\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{B}} &:= \int_0^\infty e^{\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-1}\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{B}}t}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-1}\boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{B}}\boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-\mathrm{T}}t}\mathrm{d}t,
\end{aligned}
\tag{4.56}
$$

respectively.

Lemma 4.19:
Consider the asymptotically stable second-order system (3.124), the reduced system (4.50) with matrices (4.54), the corresponding controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathbf{B}}$ as defined in (3.129), the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{B}}$, and the reduced controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{B}}$ from (4.56). The mappings $\mathbf{h}_{\mathbf{B}}$ and $\widehat{\mathbf{h}}_{\mathbf{B}}$ are as defined in (4.55). Then, the following equations hold

$$
\int_0^\infty \|\mathbf{h}_{\mathbf{B}}(t)\|_{\mathrm{F}}^2\mathrm{d}t = \mathrm{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathbf{B}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big)\,, \qquad \int_0^\infty \|\widehat{\mathbf{h}}_{\mathbf{B}}(t)\|_{\mathrm{F}}^2\mathrm{d}t = \mathrm{tr}\big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{B}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\big)\,, \tag{4.57a}
$$

$$
\int_0^\infty \langle \mathrm{vec}(\mathbf{h}_{\mathbf{B}}(t)), \mathrm{vec}(\widehat{\mathbf{h}}_{\mathbf{B}}(t))\rangle \mathrm{d}t = \mathrm{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{B}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\Big)\,. \tag{4.57b}
$$

◊

Proof. We derive

$$
\int_0^\infty \|\mathbf{h}_{\mathbf{B}}(t)\|_{\mathrm{F}}^2\mathrm{d}t = \int_0^\infty \mathrm{tr}\Big(\boldsymbol{\mathcal{C}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\Big)\,\mathrm{d}t = \mathrm{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathbf{B}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big)\,,
$$

what proves the first equation in (4.57a) while the second one is proven analogously. To show equation (4.57b), we derive

$$
\begin{aligned}
\int_0^\infty \langle \mathrm{vec}(\mathbf{h}_{\mathbf{B}}(t)), \mathrm{vec}(\widehat{\mathbf{h}}_{\mathbf{B}}(t))\rangle \mathrm{d}t &= \int_0^\infty \mathrm{tr}\Big(\boldsymbol{\mathcal{C}}e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{B}}^{-\mathrm{T}}t}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\Big)\,\mathrm{d}t\\
&= \mathrm{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{B}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\Big)\,.
\end{aligned}
$$

□

From Lemma 4.18 and Lemma 4.19, we derive the following theorem, which provides a bound of the L_∞-error $\|\mathbf{y}_{\mathrm{L},\mathbf{B}} - \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}\|_{L_\infty}$.

Theorem 4.20:
Consider the asymptotically stable second-order system (3.124), the reduced system (4.50) with matrices (4.54), the corresponding controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathbf{B}}$ as defined in (3.129), the matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{B}}$, and the reduced controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{B}}$ from (4.56). The L_∞-error between the output $\mathbf{y}_{\mathrm{L},\mathbf{B}}$ and the reduced output $\mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}$ satisfies the following bound

$$
\|\mathbf{y}_{\mathrm{L},\mathbf{B}} - \mathbf{y}_{\mathrm{L},\mathrm{r},\mathbf{B}}\|_{L_\infty}^2 \leq \Big(\mathrm{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathbf{B}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big) - 2\,\mathrm{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{B}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\Big) + \mathrm{tr}\big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{B}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\big)\Big)\|\mathbf{u}\|_{L_2}^2. \tag{4.58}
$$

◊

To derive similar bounds for the remaining error components in (4.53), we consider the first-order controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathbf{X}_0}$ from (3.133) and $\boldsymbol{\mathcal{P}}_{\mathbf{V}_0}$ from (3.136). Also, we define the following reduced Gramians and matrices

$$\begin{aligned}\boldsymbol{\mathcal{P}}_{\mathrm{r},*} &:= \int_0^\infty e^{\boldsymbol{\mathcal{E}}_{\mathrm{r},*}^{-1}\boldsymbol{\mathcal{A}}_{\mathrm{r},*}t}\boldsymbol{\mathcal{E}}_{\mathrm{r},*}^{-1}\boldsymbol{\Gamma}_{\mathrm{r},*}\boldsymbol{\Gamma}_{\mathrm{r},*}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},*}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*}^{-\mathrm{T}}t}\mathrm{d}t,\\ \widetilde{\boldsymbol{\mathcal{P}}}_* &:= \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\Gamma}_*\boldsymbol{\Gamma}_{\mathrm{r},*}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}_{\mathrm{r},*}^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{r},*}^{-\mathrm{T}}t}\mathrm{d}t.\end{aligned} \tag{4.59}$$

with $*$ equal to '$\mathbf{X}_0$', or '$\mathbf{V}_0$', corresponding to the remaining reduced second-order systems (4.51) and (4.52) where

$$\boldsymbol{\Gamma}_{\mathbf{X}_0} := \begin{bmatrix}\mathbf{X}_0\\ 0\end{bmatrix}, \qquad \boldsymbol{\Gamma}_{\mathrm{r},\mathbf{X}_0} := \mathbf{Z}_{0,\mathrm{r},\mathbf{X}_0}, \qquad \boldsymbol{\Gamma}_{\mathbf{V}_0} := \begin{bmatrix}0\\ \mathbf{M}\mathbf{V}_0\end{bmatrix}, \qquad \boldsymbol{\Gamma}_{\mathrm{r},\mathbf{V}_0} := \boldsymbol{\mathcal{E}}_{\mathrm{r},\mathbf{V}_0}\mathbf{Z}_{0,\mathrm{r},\mathbf{V}_0}.$$

Corollary 4.21:
Consider the asymptotically stable second-order system (3.121), the reduced systems (4.50), (4.51), and (4.52) with matrices (4.54), the corresponding controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathbf{B}}$, $\boldsymbol{\mathcal{P}}_{\mathbf{X}_0}$, and , $\boldsymbol{\mathcal{P}}_{\mathbf{V}_0}$ as defined in (3.129), (3.133), and (3.136), respectively, the matrices $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{B}}$, $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{X}_0}$, and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{V}_0}$, and the reduced controllability Gramians $\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{B}}$, $\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{X}_0}$, and $\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{V}_0}$ from (4.56). The L_∞-error between the output $\mathbf{y}_{\mathrm{L}}$ and the reduced output $\mathbf{y}_{\mathrm{L,r}}$ satisfies the following bound

$$\begin{aligned}\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_\infty}^2 \leq &\Big(\operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathbf{B}}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big) - 2\operatorname{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{B}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\Big) + \operatorname{tr}\big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{B}}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{B}}^{\mathrm{T}}\big)\Big)\|\mathbf{u}\|_{L_2}^2\\ &+ \Big(\operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathbf{X}_0}\boldsymbol{\mathcal{C}}_1^{\mathrm{T}}\big) - 2\operatorname{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{X}_0}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{X}_0}^{\mathrm{T}}\Big) + \operatorname{tr}\big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{X}_0}\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{X}_0}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{X}_0}^{\mathrm{T}}\big)\Big)\|\chi_0\|_2^2\\ &+ \Big(\operatorname{tr}\big(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{P}}_{\mathbf{V}_0}\boldsymbol{\mathcal{C}}^{\mathrm{T}}\big) - 2\operatorname{tr}\Big(\boldsymbol{\mathcal{C}}\widetilde{\boldsymbol{\mathcal{P}}}_{\mathbf{V}_0}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{V}_0}^{\mathrm{T}}\Big) + \operatorname{tr}\big(\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{V}_0}\boldsymbol{\mathcal{P}}_{\mathrm{r},\mathbf{V}_0}\boldsymbol{\mathcal{C}}_{\mathrm{r},\mathbf{V}_0}^{\mathrm{T}}\big)\Big)\|\nu_0\|_2^2.\end{aligned} \tag{4.60}$$

◊

4.3.1.2 Extended-input approach for inhomogeneous second-order ODE systems

In this paragraph, we reduce the second-order systems (3.121) and (3.154) using the extended-input representation described in Section 3.2.1.2 and Section 3.2.2.2. To consider the initial conditions within the reduction process, we utilize the extended input matrix $\boldsymbol{\mathcal{W}}_{\mathrm{so}}$ from (3.144).

For the system (3.121) with a linear output equation, we aim to derive a surrogate system

$$\begin{aligned}\mathbf{M}_{\mathrm{r}}\ddot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{D}_{\mathrm{r}}\dot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{K}_{\mathrm{r}}\mathbf{x}_{\mathrm{r}}(t) &= \mathbf{B}_{\mathrm{r}}\mathbf{u}(t), \qquad \mathbf{x}_{\mathrm{r}}(0) = \mathbf{X}_{0,\mathrm{r}}\chi_0, \quad \dot{\mathbf{x}}(0) = \mathbf{V}_{0,\mathrm{r}}\nu_0,\\ \mathbf{y}_{\mathrm{L,r}}(t) &= \boldsymbol{\mathcal{C}}_{1,\mathrm{r}}\mathbf{x}_{\mathrm{r}}(t),\end{aligned} \tag{4.61}$$

which leads to the output approximation $\mathbf{y}_{\mathrm{L}}(t) \approx \mathbf{y}_{\mathrm{L,r}}(t)$. Analogously, for the system (3.154) with a quadratic output equation, we aim to derive a surrogate model

$$\begin{aligned} \mathbf{M}_{\mathrm{r}}\dot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{D}_{\mathrm{r}}\dot{\mathbf{x}}_{\mathrm{r}}(t) + \mathbf{K}_{\mathrm{r}}\mathbf{x}_{\mathrm{r}}(t) = \mathbf{B}_{\mathrm{r}}\mathbf{u}(t), \qquad \mathbf{x}_{\mathrm{r}}(0) = \mathbf{X}_{0,\mathrm{r}}\chi_0, \quad \dot{\mathbf{x}}_{\mathrm{r}}(0) = \mathbf{V}_{0,\mathrm{r}}\nu_0, \\ \mathbf{y}_{\mathrm{r}}(t) = \mathbf{x}_{\mathrm{r}}(t)^{\mathrm{T}}\mathcal{M}_{11,\mathrm{r}}\mathbf{x}_{\mathrm{r}}(t), \end{aligned} \tag{4.62}$$

with $\mathbf{y}_{\mathrm{Q}}(t) \approx \mathbf{y}_{\mathrm{Q,r}}(t)$. To derive the reduced system matrices, we determine projecting matrices $\mathbf{V}_{\mathrm{r}}, \mathbf{T}_{\mathrm{r}} \in \mathbb{R}^{n\times r}$ with $r \ll n$. Then, the reduced matrices are

$$\begin{aligned} &\mathbf{M}_{\mathrm{r}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{M}\mathbf{T}_{\mathrm{r}}, \quad \mathbf{D}_{\mathrm{r}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{D}\mathbf{T}_{\mathrm{r}}, \quad \mathbf{K}_{\mathrm{r}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{K}\mathbf{T}_{\mathrm{r}}, \quad \mathbf{B}_{\mathrm{r}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{B}, \quad \mathbf{C}_{1,\mathrm{r}} = \mathbf{C}_1\mathbf{T}_{\mathrm{r}}, \\ &\mathbf{X}_{0,\mathrm{r}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{X}_0, \quad \mathbf{V}_{0,\mathrm{r}} = \mathbf{W}_{\mathrm{r}}^{\mathrm{T}}\mathbf{V}_0, \quad \mathbf{M}_{11,\mathrm{r}} = \mathbf{T}_{\mathrm{r}}^{\mathrm{T}}\mathbf{M}_{11}\mathbf{T}_{\mathrm{r}}. \end{aligned} \tag{4.63}$$

For systems (3.121) with a linear output equation, Table 3.9 summarizes the system Gramians and respective energies. From those energies, it follows that states corresponding to small eigenvalues of the respective controllability and observability Gramians $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ and $\mathbf{Q}_{\mathrm{L},\mathcal{W}_{\mathrm{so}}}$ from (3.148) and (3.138), respectively, are negligible, while states corresponding to large eigenvalues span the most dominant controllability and observability subspaces.

As described in Table 3.10, the controllability and observability behavior of the system (3.154) with a quadratic output equation is encoded by the corresponding second-order controllability and observability Gramians $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ and $\mathbf{Q}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}$ defined in (3.148) and (3.158), respectively. The corresponding energy norms summarized in Table 3.10 show that states corresponding to large eigenvalues of the respective Gramians encode the dominant controllability and observability spaces, while states corresponding to the small eigenvalues are negligible.

It follows, that for both system classes, we truncate states corresponding to small eigenvalues of the controllability Gramian $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}}$ and of the observability Gramians $\mathbf{Q}_{\mathrm{L},\mathcal{W}_{\mathrm{so}}}$ and $\mathbf{Q}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}$. Since these properties are similar to those for homogeneous systems with a linear output equation, we can apply the BT method for second-order systems as introduced in Algorithm 3. Again, we first balance the system to derive controllability Gramians and observability Gramians that coincide and truncate the states corresponding to the smallest eigenvalues of those Gramians, which results in Algorithm 13.

Error bound for systems with a linear output equation We develop an a posteriori error bound for the reduced systems derived by Algorithm 13. Therefore, we utilize the reduced first-order matrices

$$\mathcal{E}_{\mathrm{r}} := \begin{bmatrix} \mathbf{I}_r & 0 \\ 0 & \mathbf{M}_{\mathrm{r}} \end{bmatrix}, \quad \mathcal{A}_{\mathrm{r}} := \begin{bmatrix} 0 & \mathbf{I}_r \\ -\mathbf{K}_{\mathrm{r}} & -\mathbf{D}_{\mathrm{r}} \end{bmatrix}, \quad \mathcal{W}_{\mathrm{so,r}} := \begin{bmatrix} 0 & \mathbf{X}_{0,\mathrm{r}} & 0 \\ \mathbf{B}_{\mathrm{r}} & 0 & \mathbf{M}_{\mathrm{r}}\mathbf{V}_{0,\mathrm{r}} \end{bmatrix}, \quad \mathcal{C} := \begin{bmatrix} \mathbf{C}_{1,\mathrm{r}} & 0 \end{bmatrix} \tag{4.64}$$

Algorithm 13 BT method for the second-order ODE systems (3.121) and (3.154) with a linear or quadratic output using the extended-input approach.

```
Require: The original system (3.121) or (3.154) and the order $r$.
Ensure: The reduced system (4.61) or (4.62).
1: Build the input matrix
       $\mathcal{W}_{\mathrm{so}} = \begin{bmatrix} 0 & \mathbf{X}_0 & 0 \\ \mathbf{B} & 0 & \mathbf{M}\mathbf{V}_0 \end{bmatrix}.$
2: Compute factors of Gramians $\mathbf{P}_{\mathcal{W}_{\mathrm{so}}} \approx \mathbf{R}\mathbf{R}^{\mathrm{T}}$ from (3.148) and $\mathbf{Q} \approx \mathbf{S}\mathbf{S}^{\mathrm{T}}$, where $\mathbf{Q}$ is
   equal to $\mathbf{Q}_{\mathrm{L}}$ from (3.138) or $\mathbf{Q}_{\mathrm{Q},\mathcal{W}_{\mathrm{so}}}$ from (3.158).
3: Perform the SVD of $\mathbf{S}^{\mathrm{T}}\mathbf{M}\mathbf{R}$, and decompose as
       $\mathbf{S}^{\mathrm{T}}\mathbf{M}\mathbf{R} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_1 & \\ & \mathbf{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix}^{\mathrm{T}},$
   with $\mathbf{\Sigma}_1 \in \mathbb{R}^{r \times r}$.
4: Construct the projection matrices
       $\mathbf{W}_{\mathrm{r}} = \mathbf{S}\mathbf{U}_1\mathbf{\Sigma}_1^{-\frac{1}{2}}$ and $\mathbf{T}_{\mathrm{r}} = \mathbf{R}\mathbf{V}_1\mathbf{\Sigma}_1^{-\frac{1}{2}}.$
5: Construct reduced matrices (4.63).
```

and the reduced Gramian and matrix

$$
\begin{aligned}
\mathcal{P}_{\mathcal{W}_{\mathrm{so}},\mathrm{r}} &:= \int_0^\infty e^{\mathcal{E}_\mathrm{r}^{-1}\mathcal{A}_\mathrm{r} t}\mathcal{E}_\mathrm{r}^{-1}\mathcal{W}_{\mathrm{so,r}}\mathcal{W}_{\mathrm{so,r}}^{\mathrm{T}}\mathcal{E}_\mathrm{r}^{-\mathrm{T}} e^{\mathcal{A}_\mathrm{r}^{\mathrm{T}}\mathcal{E}_\mathrm{r}^{-\mathrm{T}} t}\mathrm{d}t, \\
\widetilde{\mathcal{P}}_{\mathcal{W}_{\mathrm{so}}} &:= \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A} t}\mathcal{E}^{-1}\mathcal{W}_{\mathrm{so}}\mathcal{W}_{\mathrm{so,r}}^{\mathrm{T}}\mathcal{E}_\mathrm{r}^{-\mathrm{T}} e^{\mathcal{A}_\mathrm{r}^{\mathrm{T}}\mathcal{E}_\mathrm{r}^{-\mathrm{T}} t}\mathrm{d}t.
\end{aligned}
\tag{4.65}
$$

Applying the bounds from (4.60) while using the same bases for all subsystems yields the following theorem.

Theorem 4.22:

Consider the asymptotically stable system (3.121) and the reduced surrogate system (4.61). Also, consider the controllability Gramian $\mathcal{P}_{\mathcal{W}_{\mathrm{so}}}$ as defined in (3.149), the reduced Gramian $\mathcal{P}_{\mathcal{W}_{\mathrm{so}},\mathrm{r}}$, and the matrix $\widetilde{\mathcal{P}}_{\mathcal{W}_{\mathrm{so}}}$ from (4.65). Then, the error between the two outputs is bounded by

$$
\begin{aligned}
&\|\mathbf{y}_{\mathrm{L}} - \mathbf{y}_{\mathrm{L,r}}\|_{L_\infty}^2 \\
&\quad \leq \left(\operatorname{tr}\!\left(\mathcal{C}\mathcal{P}_{\mathcal{W}_{\mathrm{so}}}\mathcal{C}^{\mathrm{T}}\right) - 2\operatorname{tr}\!\left(\mathcal{C}\widetilde{\mathcal{P}}_{\mathcal{W}_{\mathrm{so}}}\mathcal{C}_{\mathrm{r}}^{\mathrm{T}}\right) + \operatorname{tr}\!\left(\mathcal{C}_{\mathrm{r}}\mathcal{P}_{\mathcal{W}_{\mathrm{so}},\mathrm{r}}\mathcal{C}_{\mathrm{r}}^{\mathrm{T}}\right)\right)\left(\|\mathbf{u}\|_{L_2} + \|\chi_0\|_2 + \|\nu_0\|_2\right)^2.
\end{aligned}
\tag{4.66}
$$

◊

Proof. Since $\mathcal{P}_{\mathcal{W}_{\rm so}} = \mathcal{P}_{\mathcal{B}} + \mathcal{P}_{\mathbf{x}_0} + \mathcal{P}_{\mathbf{v}_0}$, $\widetilde{\mathcal{P}}_{\mathcal{W}_{\rm so}} = \widetilde{\mathcal{P}}_{\mathcal{B}} + \widetilde{\mathcal{P}}_{\mathbf{x}_0} + \widetilde{\mathcal{P}}_{\mathbf{v}_0}$, and $\mathcal{P}_{\mathcal{W}_{\rm so},\rm r} = \mathcal{P}_{\mathcal{B},\rm r} + \mathcal{P}_{\mathbf{x}_0,\rm r} + \mathcal{P}_{\mathbf{v}_0,\rm r}$, the bounds from (4.60) can be bounded as stated in the theorem. □

Error bound for systems with a quadratic output equation To bound the error for systems with a quadratic output equation that results from the approximation generated by Algorithm 13, we consider the respective first-order representation with matrices from (4.64),

$$\mathcal{M} := \begin{bmatrix} \mathbf{M} & 0 \\ 0 & 0 \end{bmatrix}, \qquad \text{and} \qquad \mathcal{M}_{\rm r} := \begin{bmatrix} \mathbf{M}_{rr} & 0 \\ 0 & 0 \end{bmatrix}.$$

We apply the error bound from (4.24), which leads to the following theorem.

Theorem 4.23:
Consider the asymptotically stable system (3.154) and the reduced surrogate system (4.63). Also, consider the controllability Gramian $\mathcal{P}_{\mathcal{W}_{\rm so}}$ as defined in (3.149), the reduced Gramian $\mathcal{P}_{\mathcal{W}_{\rm so},\rm r}$, and the matrix $\widetilde{\mathcal{P}}_{\mathcal{W}_{\rm so}}$ from (4.65). Then, the error between the two outputs is bounded by

$$\begin{aligned}\|\mathbf{y}_{\rm Q} - \mathbf{y}_{\rm Q,r}\|_{L_\infty}^2 \leq \Big(&\operatorname{tr}(\mathcal{P}_{\mathcal{W}_{\rm so}}\mathcal{M}\mathcal{P}_{\mathcal{W}_{\rm so}}\mathcal{M}) - 2\operatorname{tr}\Big(\widetilde{\mathcal{P}}_{\mathcal{W}_{\rm so}}^{\rm T}\mathcal{M}\widetilde{\mathcal{P}}_{\mathcal{W}_{\rm so}}\mathcal{M}_{\rm r}\Big) \\ &+ \operatorname{tr}(\mathcal{P}_{\mathcal{W}_{\rm so},\rm r}\mathcal{M}_{\rm r}\mathcal{P}_{\mathcal{W}_{\rm so},\rm r}\mathcal{M}_{\rm r})\Big)\left(\|\chi_0\|_2 + \|\nu_0\|_2 + \|\mathbf{u}\|_{L_2}\right)^2. \end{aligned} \tag{4.67}$$

◊

Proof. Applying the error bound from (4.24) yields

$$\begin{aligned}\|\mathbf{y}_{\rm Q} - \mathbf{y}_{\rm Q,r}\|_{L_\infty} \leq \sum_{*,\circ\in\{\text{'}\mathbf{B}\text{'},\text{'}\mathbf{X}_0\text{'},\text{'}\mathbf{V}_0\text{'}\}} \Big(&\operatorname{tr}(\mathcal{P}_*\mathcal{M}\mathbf{P}_\circ\mathcal{M}) - 2\operatorname{tr}\Big(\widetilde{\mathbf{P}}_\circ^{\rm T}\mathcal{M}\widetilde{\mathcal{P}}_*\mathcal{M}_{\rm r}\Big) \\ &+ \operatorname{tr}(\mathcal{P}_{\rm r,\circ}\mathcal{M}_{\rm r}\mathcal{P}_{\rm r,*}\mathcal{M}_{\rm r})\Big)\|\mathbf{u}_* \otimes \mathbf{u}_\circ\|_{L_2}^2\end{aligned}$$

for the controllability Gramians $\mathcal{P}_{\mathbf{B}}$, $\mathcal{P}_{\mathbf{x}_0}$, and , $\mathcal{P}_{\mathbf{v}_0}$ as defined in (3.129), (3.133), and (3.136), respectively, the matrices $\widetilde{\mathcal{P}}_{\mathbf{B}}$, $\widetilde{\mathcal{P}}_{\mathbf{x}_0}$, and $\widetilde{\mathcal{P}}_{\mathbf{v}_0}$, and the reduced controllability Gramians $\mathcal{P}_{\rm r,\mathbf{B}}$, $\mathcal{P}_{\rm r,\mathbf{x}_0}$, and $\mathcal{P}_{\rm r,\mathbf{v}_0}$ from (4.56). Also, the different inputs are $\mathbf{u}_{\mathbf{B}} = \mathbf{u}$, $\mathbf{u}_{\mathbf{x}_0} = \chi_0$, and $\mathbf{u}_{\mathbf{v}_0} = \nu_0$. The statement follows since $\mathcal{P}_{\mathcal{W}_{\rm so}} = \mathcal{P}_{\mathbf{B}} + \mathcal{P}_{\mathbf{x}_0} + \mathcal{P}_{\mathbf{v}_0}$, $\widetilde{\mathcal{P}}_{\mathcal{W}_{\rm so}} = \widetilde{\mathcal{P}}_{\mathbf{B}} + \widetilde{\mathcal{P}}_{\mathbf{x}_0} + \widetilde{\mathcal{P}}_{\mathbf{v}_0}$, and $\mathcal{P}_{\mathcal{W}_{\rm so},\rm r} = \mathcal{P}_{\mathbf{B},\rm r} + \mathcal{P}_{\mathbf{x}_0,\rm r} + \mathcal{P}_{\mathbf{v}_0,\rm r}$ holds. □

Remark 4.24:
Since the transfer function describing the second-order input- and initial condition-to-output behavior is equal to

$$\mathcal{G}(s) = \mathbf{C}_1\boldsymbol{\Lambda}(s)\left[(\mathbf{D} + s\mathbf{M}) \quad \mathbf{I}\right]\mathcal{W}_{\rm so}$$

the IRKA method as presented in [140] is not applicable. Also, the IRKA method from [140] only considers systems with homogeneous initial conditions. Consequently, this method only applies to specific cases of our setting. Hence, to apply the IRKA method we refer to the first-order IRKA method in Section 4.1.2, which can be applied to the system after transforming it into first-order form. ◊

4.3.2 Numerical results

In this section, we illustrate the BT method for second-order systems using two different examples. The first example is a vibrational model of a building, and the second one is a mass-spring-damper system. Both examples are considered with a linear and with a quadratic output equation. We will refer to the original systems (3.121) and (3.154) as `FOM`, in the following, and to the reduced systems generated by standard BT that considers homogeneous systems by `ROM_HOM`. The reduced system approximation that is obtained by applying the multi-system approach from Algorithm 12 is referred to as `ROM_MULT` and the reduced system that is generated by applying Algorithm 13 as `ROM_EXT`.

The computations were done on a computer with 4 Intel® Core™i5-4690 CPUs running at 3.5 GHz. The experiments use Matlab R2021a.

4.3.2.1 Example 4: Building example

We consider the building example from [7, page 17] with dimensions $n = 24$, $m = 1$. The data are available in [98].

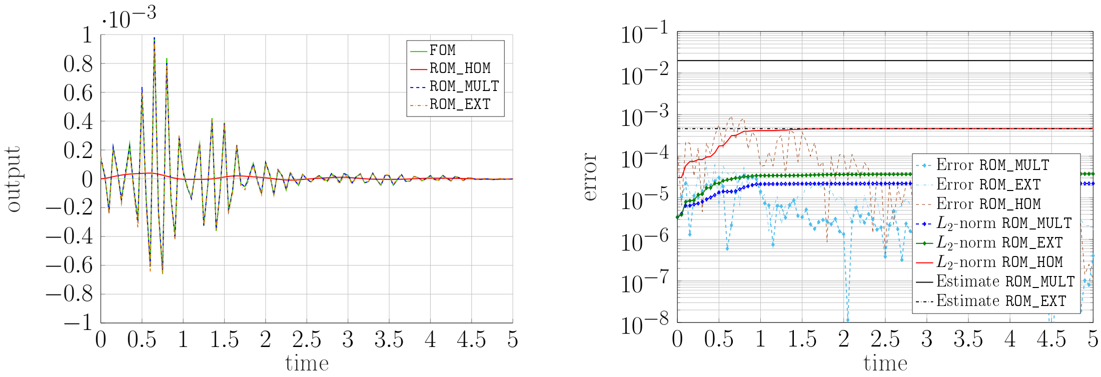
Example 4a: Linear output equation As output matrices, we use

$$\mathbf{C}_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \in \mathbb{R}^{1\times 24}.$$

For the projecting matrix $\mathbf{V}_\mathrm{r}$ that results from the BT procedure for the homogeneous second-order system (3.121), we consider the singular value decomposition $\mathbf{U\Sigma V}^\mathrm{T} = \mathbf{V}_\mathrm{r}$. Assume that $\operatorname{rank}(\mathbf{V}_\mathrm{r}) = \ell$. The position and velocity initial condition are the $(\ell+1)$-st column of $\mathbf{U}$, i.e.,

$$\mathbf{X}_0 = \mathbf{x}_0 = \mathbf{V}_0 = \dot{\mathbf{x}}_0 = \mathbf{U}[\,:\,,\,\ell+1\,].$$

In this example, the reduced dimension is set to $r = 10$ within the multi-system and the extended-input approach. Figure 4.7a shows the output behavior of the original system and the reduced ones for an input $\mathbf{u}(t) = 0.2 \cdot e^{-t}$. We observe that the original output behavior, depicted in green, is well approximated by the separately reduced subsystems (`ROM_MULT`), which is depicted by the blue dashed line. The reduced system `ROM_EXT` using the extended-input approach (depicted by the orange-colored dashed line) provides a proper approximation of the original output as well. Additionally, we see that



(a) Output

(b) Error

Figure 4.7: Example 4a - Output responses and the corresponding errors.

the output of the reduced system ROM_HOM, depicted in red, fails in approximating the original system's transient behavior.

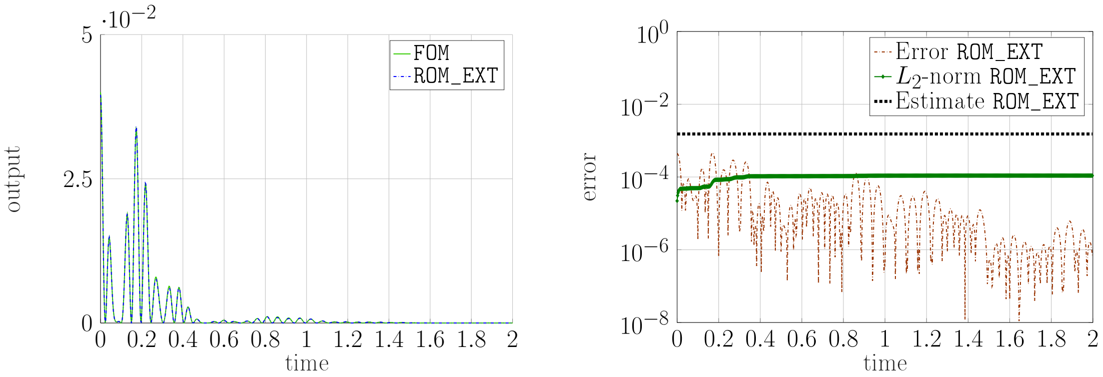
Figure 4.7b depicts the output errors. Additionally, we evaluate the actual L_2-norm error. Therefore, we plot the integral

$$\sqrt{\int_0^t \|\mathbf{y}_\mathrm{L}(\tau) - \mathbf{y}_{\mathrm{L,r}}(\tau)\|_2^2 \mathrm{d}\tau} \tag{4.68}$$

that converges to the L_2-norm of the error. The light blue line with markers depicts the error of the separately reduced system ROM_MULT and the dashed, brown colored line the error of the reduced system ROM_EXT using the combined Gramian. The reduced system ROM_HOM leads to the error depicted by the dashed, orange-colored line. We observe that the multi-system approach and the extended input approach lead to significantly smaller errors than the error corresponding to the reduced system ROM_HOM. The dark blue, dashed line with markers is the integral (4.68) converging to the actual L_2-norm error of the separately reduced system ROM_MULT. The error bound from (4.60) provides a value of $1.99 \cdot 10^{-2}$ (depicted by the black line). This error bound provides a proper upper bound of the actual L_2-norm error. The green line with markers provides the integral (4.68) corresponding to the extended-input approach ROM_EXT and its error estimation $4.5 \cdot 10^{-4}$ from (4.66) is depicted by the dashed, black line. The red line shows the integral (4.68) of the reduced system ROM_HOM. It confirms again that not considering the initial conditions within the reduction method leads to unsatisfactory approximations for this example.

Example 4b: Quadratic output equation Now, we consider the building example with a quadratic output equation. For that, we choose the output matrix to be

$$\mathbf{M}_{11} = \mathbf{C}_1^\mathrm{T}\mathbf{C}_1, \quad \mathbf{C}_1 = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix},$$



(a) Output

(b) Error

Figure 4.8: Example 4b - Output responses and the corresponding errors.

and the position and velocity initial conditions are

$$\mathbf{X}_0 = \mathbf{V}_0 = e_n$$

so that $\mathbf{x}_0 = \dot{\mathbf{x}}_0 = 0.0137 \cdot e_n$ while $\|\mathbf{B}\|_2 = 0.0137$.

We reduce the system (3.154) to obtain a surrogate system of the form (4.62) with matrices of the reduced dimension $r = 10$. Figure 4.8a shows the output behavior of the original system and the reduced ones for an input $\mathbf{u}(t) = 0.2 \cdot e^{-t}$. We observe that the output behavior of the original system depicted in green is well-approximated by the reduced outputs that are derived using the extended-input approach and that is depicted in blue (ROM_EXT).

Figure 4.8b depicts the errors and their L_2-norms. The dashed, brown colored line shows the error of the reduced system ROM_EXT using the extended-input approach. Additionally, we evaluate the actual L_2-norm error. Therefore, we plot the integral

$$\sqrt{\int_0^t \|\mathbf{y}_{\mathrm{Q}}(\tau) - \mathbf{y}_{\mathrm{Q,r}}(\tau)\|_2 \mathrm{d}\tau} \tag{4.69}$$

that converges to the L_2-norm of the error. The error bound from (4.67) provides a value of $1.5 \cdot 10^{-3}$ (depicted by the black line). This error bound provides a conservative upper bound of the actual L_2-norm error. The green line with markers provides the integral (4.69) corresponding to the reduced system ROM_EXT.

4.3.2.2 Example 5: Mass-spring-damper example

We consider the mass-spring-damper model presented in [137] describing the structure depicted in Figure 4.9. A more detailed background can be found in [78].

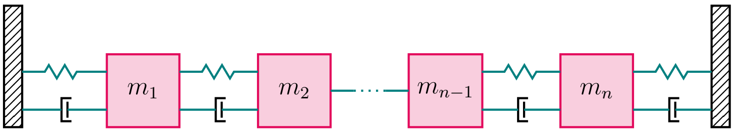


Figure 4.9: Example 5 - Sketch of a mechanical example with one row of masses connected with consecutive springs.

We choose the model of dimensions $n = 2000$, $m = p = 1$. The input is the external forcing on the n-th mass, and the initial conditions are set to be the last and the first unit vectors

$$\mathbf{X}_0 = \mathbf{x}_0 := e_n, \qquad \mathbf{V}_0 = \dot{\mathbf{x}}_0 := e_1.$$

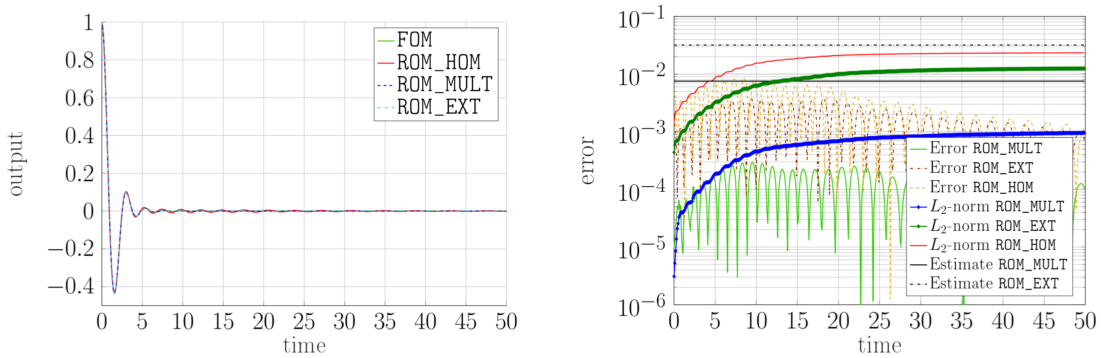
Example 5a: Linear output equation We consider an output that observes the displacement of the n-th mass, i.e.,

$$\mathbf{C}_1 = \begin{bmatrix} 0 & 0 & \dots & 1 \end{bmatrix}.$$

In this example, we truncate the systems with a tolerance of 10^{-4}, i.e., all Hankel singular values smaller than $10^{-4} \cdot \sigma_1$ are neglected. That way, we obtain reduced systems (4.50), (4.51), and (4.52) of dimensions 147, 180, 98, respectively, resulting from the multi-system method. Using the extended-input approach, we obtain a surrogate model (4.61) of dimension 157.

Figure 4.10a shows the output behavior of the systems for the input $\mathbf{u}(t) = 0.2 \cdot e^{-t}$. The output behavior of the original system is depicted in green. The blue, dashed line displays the output composed by the separately reduced systems ROM_MULT and the orange-colored, dashed line the reduced system ROM_EXT using the extended-input approach. The reduced output resulting from the reduced system ROM_HOM is depicted in red. We observe that all outputs approximate the original system behavior. However, ROM_HOM shows oscillations of slightly higher magnitude than the FOM for some timings.

The output errors and their L_2-norms are illustrated in Figure 4.10b. The light blue line with markers, the brown colored dashed line, and the orange colored dashed line show the error of the separately reduced outputs, the output corresponding to the extended-input approach and the output resulting from the reduced system ROM_HOM, respectively. We observe again that the separately reduced system ROM_MULT and the reduced system ROM_EXT using the extended-input approach leads to lower errors. Additionally, we evaluate the actual L_2-norm error and plot the integral (4.68), which converges to the L_2-norm of the error. The dark blue, dashed line with markers shows the integral (4.68) for the separately reduced system ROM_MULT and the green one the integral for the reduced system ROM_EXT using the combined Gramian. The error bounds from (4.60)



(a) Output

(b) Error

Figure 4.10: Example 5a - Output responses and the corresponding errors.

and (4.66) provide L_2-error estimator values of $7.5490 \cdot 10^{-3}$ and $3.1922 \cdot 10^{-2}$ for this example, respectively. The error bounds are depicted by the black line and the black dashed line in Figure 4.10b. We observe that the error bounds are rather conservative. The integral (4.68) of the reduced system `ROM_HOM` is depicted in red. It converges to an L_2-error larger than the errors corresponding to the first two reduction methods.

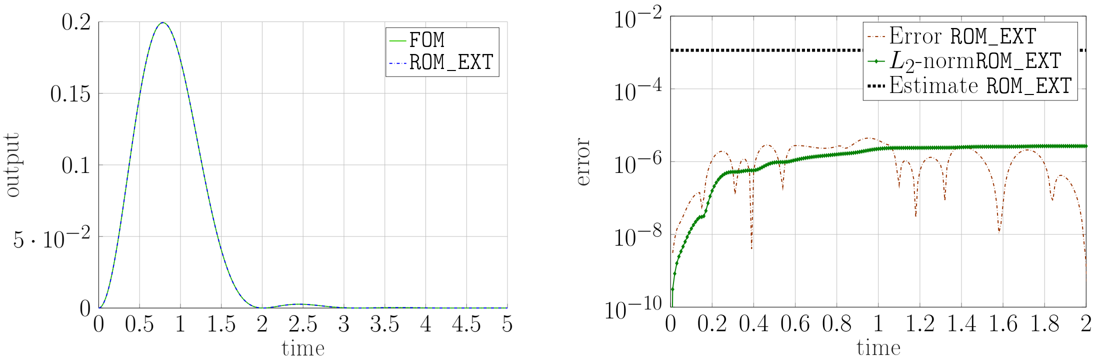
Example 5b: quadratic output equation We consider the mechanical system with a quadratic output equation where the output matrix is

$$\mathcal{M} = \begin{bmatrix} \mathbf{C}_1^T\mathbf{C}_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{C}_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}.$$

We truncate the system with a tolerance of 10^{-4}, i.e., all Hankel singular values smaller than $10^{-4} \cdot \sigma_1$ are neglected. That way, we have a reduced system of dimension 213 using the extended-input approach.

Figure 4.11a shows the output behavior of the systems for the input $\mathbf{u}(t) = 0.2 \cdot e^{-t}$. The output behavior of the original system is depicted in green, and the orange-colored dashed line describes the reduced system `ROM_EXT` using the extended-input approach. We observe that the output approximates the original system behavior well.

The output errors and their L_2-norms are illustrated in Figure 4.10b. The brown-colored dashed line shows the output error corresponding to the extended-input approach. We observe again that the reduced system `ROM_EXT` leads to small errors. Also, we evaluate the actual L_2-norm error and plot the integral (4.69) that converges to the L_2-norm of the error. The green line shows the integral for the reduced system `ROM_EXT` using the combined Gramian. The error bound from (4.67) provides an L_2 error estimation value $1.15 \cdot 10^{-3}$. The black dashed line depicts this bound. We observe that the error bound is conservative.



(a) Output (b) Error

Figure 4.11: Example 5b - Output responses and the corresponding errors.

Summary of reduction methods

In this chapter, we have introduced the BT method and the IRKA method for different system structures. Therefore, we have used the multi-system approach following the ideas introduced in [15] and the extended-input approach initially derived in [66] for first-order ODE systems with linear output equations. Our contributions in this chapter include the introduction of a BT method for inhomogeneous first-order ODE systems with quadratic output equations and the respective error bounds in Section 4.1.1. Moreover, we have derived BT methods for inhomogeneous first-order DAE systems with linear and quadratic output equations, where we again provide a tailored error bound. We applied these methods to some numerical examples to illustrate their effectiveness. Finally, in Section 4.3.1, we have derived a BT approach for inhomogeneous second-order ODE systems with linear and quadratic output equations, including suitable error bounds that maintain the second-order structure of the respective systems. Again, we have applied the respective methods to some numerical examples.

Since not every approach applies to all system structures, in Table 4.1, we summarize the methods available for the different system types.

		BT	IRKA
First-order ODE systems with linear output	Multi-system approach	✓	✓
	Extend-input approach	✓	✓
First-order ODE systems with quadratic output	Multi-system approach	✓	–
	Extend-input approach	✓	–
First-order DAE systems with linear output	Multi-system approach	✓	✓
	Extend-input approach	✓	✓
First-order DAE systems with quadratic output	Multi-system approach	–	–
	Extend-input approach	✓	–
Second-order ODE systems with linear output	Multi-system approach	✓	–
	Extend-input approach	✓	–
Second-order ODE systems with quadratic output	Multi-system approach	–	–
	Extend-input approach	✓	–

Table 4.1: Available for reduction method for different system structures.

CHAPTER 5

REDUCED BASIS METHOD

Contents

As described in Chapter 1, this work aims to optimize external dampers in vibrational systems concerning the system response. However, computing the system response includes the computation of the respective system Gramians and, hence, the solution of certain Lyapunov equations. Solving these Lyapunov equations for several external damper configurations within an optimization process leads to high computational costs, especially when the dimensions are too large. Therefore, we aim to reduce the respective Lyapunov equations for all parameters considered within the optimization process. That way, the Lyapunov equations are approximately solvable in a reasonable time.

To describe the system dynamics of a vibrational system, we can consider the first-order parameter-dependent systems (1.5) and (1.6) with matrices as defined in (1.7). Since it is often advantageous to maintain the second-order system structure to generate physically meaningful results, we also consider the second-order parameter-dependent systems (1.3) and (1.4).

To avoid the high computational costs during the optimization process, in this chapter, the reduced basis method (RBM) is applied to accelerate the computation of the different Gramians. The RBM is a well-established method to reduce parameter-dependent partial differential equations, see [64, 68, 109, 150, 152]. Moreover, the RBM was applied to Riccatti equations, see [119] and to Lyapunov equations in [126]. In [108], the authors derive an RBM method for projected Lyapunov equations corresponding to parametric first-order DAE systems. However, we do not consider this method in this work as this would go beyond the scope of this thesis.

To accelerate the solving of the Lyapunov equations for different external dampers represented by the parameters $(c, g) \in \mathfrak{D}$, we utilize the RBM from [126] where the parametric Lyapunov equations are solved only for a few sampling parameters. Then, based on these solutions, a reduced subspace in which the Lyapunov equation solutions for all $(c, g) \in \mathfrak{D}$ approximately live is constructed. The latter steps form the computationally expensive *offline phase*. Using the reduced basis representation, the Lyapunov equations for all $(c, g) \in \mathfrak{D}$ can be solved much more efficiently in the *online phase*. We also utilize Krylov spaces to determine the corresponding bases, which reduce the respective Lyapunov equations as introduced in [140] for second-order systems.

A crucial question in the offline phase is the choice of the sample parameters. Usually, a grid of test parameters is selected. For this grid, the error is quantified using an a posteriori error bound. Then, new samples are taken at the parameters where the error bound gives the largest error. However, for the type of problems that are considered in this work, which are mechanical systems with small internal damping $\mathbf{D}_{\text{int}}$, the standard error bounds overestimate so significantly that the methods are often not converging. Hence, one of the main contributions of this chapter is to derive several error approximations that we use within the different RBM applications. Also, we derive some decoupling of parameter-independent and parameter-dependent components of the controllability space and, hence, of the solution spaces of the Lyapunov equation.

To simplify the computations and the numerical effort, we describe briefly an advantageous matrix transformation. As shown in [146, 147], there exists a transformation $\mathbf{\Phi}$, called *modal matrix*, such that

$$\mathbf{\Phi}^{\mathrm{T}}\mathbf{M}\mathbf{\Phi} = \mathbf{I}, \qquad \mathbf{\Phi}^{\mathrm{T}}\mathbf{K}\mathbf{\Phi} = \mathbf{\Omega}^2 = \operatorname{diag}\left(\omega_1^2, \dots, \omega_n^2\right).$$

The values $\omega_1, \dots, \omega_n$ are the eigenvalues of the undamped system and are called *eigenfrequencies*. Moreover, it holds that $\mathbf{\Phi}^{\mathrm{T}}\mathbf{D}_{\text{int}}\mathbf{\Phi} = 2\alpha\mathbf{\Omega}$. That means that $\mathbf{\Phi}$ diagonalizes the internal damping $\mathbf{D}_{\text{int}}$. Hence, this damping is called *modal damping*. The transformed mass matrix is the identity matrix, the transformed stiffness and internal damping matrix are diagonal matrices, and the external damping matrix is written using the low-rank factors $\widetilde{\mathbf{F}}(c) := \mathbf{\Phi}^{\mathrm{T}}\mathbf{F}(c)$. Hence, with $\widetilde{\mathbf{x}}(t) := \mathbf{\Phi}^{-1}\mathbf{x}(t)$ and $\widetilde{\mathbf{B}} := \mathbf{\Phi}^{\mathrm{T}}\mathbf{B}$, the state equation of the second-order systems (1.3) and (1.4) is equivalent to

$$\ddot{\widetilde{\mathbf{x}}}(t) + \widetilde{\mathbf{D}}(c, g)\dot{\widetilde{\mathbf{x}}}(t) + \mathbf{\Omega}^2\widetilde{\mathbf{x}}(t) = \widetilde{\mathbf{B}}\mathbf{u}(t) \tag{5.1}$$

where $\widetilde{\mathbf{D}}(c, g) = 2\alpha\mathbf{\Omega} + \widetilde{\mathbf{F}}(c)\mathbf{G}(g)\widetilde{\mathbf{F}}(c)^{\mathrm{T}}$. The respective output equations are given as

$$\mathbf{y}_{\mathrm{L}}(t) = \widetilde{\mathbf{C}}_1\widetilde{\mathbf{x}}(t) + \widetilde{\mathbf{C}}_2\dot{\widetilde{\mathbf{x}}}(t)$$

with $\widetilde{\mathbf{C}}_1 := \mathbf{C}_1\mathbf{\Phi}$ and $\widetilde{\mathbf{C}}_2 := \mathbf{C}_2\mathbf{\Phi}$, or

$$\mathbf{y}_{\mathrm{Q}}(t) = \frac{1}{2}\begin{bmatrix}\widetilde{\mathbf{x}}(t)^{\mathrm{T}} & \dot{\widetilde{\mathbf{x}}}(t)^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\widetilde{\mathbf{M}}_{11} & \widetilde{\mathbf{M}}_{12}\\ \widetilde{\mathbf{M}}_{12}^{\mathrm{T}} & \widetilde{\mathbf{M}}_{22}\end{bmatrix}\begin{bmatrix}\widetilde{\mathbf{x}}(t)\\ \dot{\widetilde{\mathbf{x}}}(t)\end{bmatrix} = \begin{bmatrix}\widetilde{\mathbf{x}}(t)^{\mathrm{T}} & \dot{\widetilde{\mathbf{x}}}(t)^{\mathrm{T}}\end{bmatrix}\widetilde{\boldsymbol{\mathcal{M}}}\begin{bmatrix}\widetilde{\mathbf{x}}(t)\\ \dot{\widetilde{\mathbf{x}}}(t)\end{bmatrix}$$

with $\widetilde{\mathbf{M}}_{11} := \mathbf{\Phi}^{\mathrm{T}}\mathbf{M}_{11}\mathbf{\Phi}$, $\widetilde{\mathbf{M}}_{12} := \mathbf{\Phi}^{\mathrm{T}}\mathbf{M}_{12}\mathbf{\Phi}$, and $\widetilde{\mathbf{M}}_{22} := \mathbf{\Phi}^{\mathrm{T}}\mathbf{M}_{22}\mathbf{\Phi}$.

In the following, we consider first-order and second-order systems separately as their controllability spaces differ and, hence, individual investigations are performed. Therefore, in Section 5.1, Lyapunov equations resulting from first-order ODE systems are considered. Then, we study the Lyapunov equations that result from second-order ODE systems in Section 5.2. For both sections, we follow the same procedure. First, we repeat the RBM with an offline and an online phase and derive an error approximation suitable for the structure of the considered vibrational systems. Afterwards, we derive some decoupling of the controllability spaces into parameter-dependent and parameter-independent components, which is used to derive some offline-online schemes and new error approximations. We want to mention that this method can also be applied and extended for the case of DAEs as presented in [108]. However, this is out of the scope of this thesis. Note that in this chapter, we introduce several algorithms that are applied later in Chapter 6.

5.1 Reduced basis method for first-order systems

This section aims to simplify the computation of the controllability Gramians

$$\boldsymbol{\mathcal{P}}(c,g) := \int_0^\infty e^{\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{A}}(c,g)t}\boldsymbol{\mathcal{E}}^{-1}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}e^{\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{-\mathrm{T}}t}\mathrm{d}t \tag{5.2}$$

for several parameters $(c,g) \in \boldsymbol{\mathcal{D}}$ by solving the Lyapunov equations

$$\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{P}}(c,g)\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}(c,g)\boldsymbol{\mathcal{E}}^{\mathrm{T}} = -\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}. \tag{5.3}$$

Therefore, we apply the RBM to build a basis $\mathbf{V}_{\mathrm{r}}$ that approximately spans the solution space of the Lyapunov equation (5.3) for all parameters $(c,g) \in \boldsymbol{\mathcal{D}}$. This basis is then used to build the approximations of the controllability Gramians $\boldsymbol{\mathcal{P}}(c,g)$ as

$$\boldsymbol{\mathcal{P}}(c,g) \approx \widetilde{\boldsymbol{\mathcal{P}}}(c,g) := \mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}} \tag{5.4}$$

for a suitable matrix $\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g) \in \mathbb{R}^{R_{\mathbf{V}} \times R_{\mathbf{V}}}$, in the online phase.

This section is structured as follows. We first repeat the offline-online approach, presented in [126], and introduce an error bound for this method in Section 5.1.1. In Section 5.1.2, we derive a decoupling strategy for the controllability space. This decoupling is used in Section 5.1.3 to derive an accelerated offline-online scheme. Moreover, we derive an error indicator independent of the parameter set $\boldsymbol{\mathcal{D}}$.

5.1.1 Offline-online RBM for first-order systems

We consider the RBM as presented in [126], which follows the paradigm of decomposing the procedure into an offline and online phase. In the offline phase, we derive subspaces, which approximate the solution spaces $\boldsymbol{\mathcal{V}}(c, g)$ of the Lyapunov equations in (5.3). Therefore, we compute a basis $\mathbf{V}_\mathrm{r} \in \mathbb{R}^{N \times R_\mathbf{V}}$ that spans a subspace $\boldsymbol{\mathcal{V}}_\mathrm{r}$ that approximates the original solution space $\boldsymbol{\mathcal{V}}(c, g)$, which approximately coincides with the controllability space of the corresponding systems (1.5) and (1.6), for all parameters $(c, g) \in \boldsymbol{\mathcal{D}}$. This phase is time-consuming but needs to be performed only once.

In the online phase, the Gramians are then approximated according to (5.4) for any parameter $(c, g) \in \boldsymbol{\mathcal{D}}$. Due to the reduced dimension of the basis $\mathbf{V}_\mathrm{r}$ obtained in the offline phase, this step is fast and can be performed repeatedly for all required parameters.

To describe the two phases in more detail, we need a criterion to evaluate the quality of the reduced space $\boldsymbol{\mathcal{V}}_\mathrm{r}$. Thus, we assume that we have an error approximation $\boldsymbol{\Delta}(c, g)$ that provides a criterion to determine how well the solution space for a parameter $(c, g) \in \boldsymbol{\mathcal{D}}$ is approximated by the current basis $\mathbf{V}_\mathrm{r}$. The error approximations are described later in this section. The following offline-online scheme was presented in [126], while the corresponding error approximation provides a novelty in this work.

Offline phase We aim to find a space $\boldsymbol{\mathcal{V}}_\mathrm{r}$ and the corresponding basis $\mathbf{V}_\mathrm{r}$ that is built as

$$\mathbf{V}_\mathrm{r} = \operatorname{orth}\left(\begin{bmatrix} \boldsymbol{\mathcal{Z}}_\mathbf{V}(c_1, g_1) & \dots & \boldsymbol{\mathcal{Z}}_\mathbf{V}(c_{N_\ell}, g_{N_\ell}) \end{bmatrix}\right) \in \mathbb{R}^{N \times R_\mathbf{V}} \tag{5.5}$$

containing the matrices $\boldsymbol{\mathcal{Z}}_\mathbf{V}(c_k, g_k)$ for $(c_k, g_k) \in \boldsymbol{\mathcal{D}}$, $k = 1, \dots, N_\ell$, where $\boldsymbol{\mathcal{Z}}_\mathbf{V}(c, g)$ spans an approximation of the solution space $\boldsymbol{\mathcal{V}}(c, g) := \operatorname{span}\{\boldsymbol{\mathcal{P}}(c, g)\}$ for the parameter $(c, g) \in \boldsymbol{\mathcal{D}}$. We can approximate the space $\boldsymbol{\mathcal{V}}(c, g)$ by a basis $\boldsymbol{\mathcal{Z}}_\mathbf{V}(c, g)$ that results from the low-rank factor $\boldsymbol{\mathcal{Z}}_\mathrm{BT}(c, g)$ of $\boldsymbol{\mathcal{P}}(c, g)$, i.e.,

$$\boldsymbol{\mathcal{Z}}_\mathbf{V}(c, g) := \boldsymbol{\mathcal{Z}}_\mathrm{BT}(c, g) \qquad \text{with} \qquad \boldsymbol{\mathcal{P}}(c, g) \approx \boldsymbol{\mathcal{Z}}_\mathrm{BT}(c, g)\boldsymbol{\mathcal{Z}}_\mathrm{BT}(c, g)^\mathrm{T}. \tag{5.6}$$

Since the controllability space of the systems (1.5) and (1.6) are spanned by the Gramian $\boldsymbol{\mathcal{P}}(c, g)$, the controllability space and the solution space $\boldsymbol{\mathcal{V}}(c, g)$ of the Lyapunov equation in (5.3) coincide. Hence, alternatively, the space $\boldsymbol{\mathcal{V}}(c, g)$ can be approximated by a basis $\boldsymbol{\mathcal{Z}}_\mathrm{IRKA}(c, g)$ that is given by

$$\boldsymbol{\mathcal{Z}}_\mathbf{V}(c, g) := \boldsymbol{\mathcal{Z}}_\mathrm{IRKA}(c, g) = \begin{bmatrix} (s_1\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}(c, g))^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_1 & \dots & (s_N\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}(c, g))^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_N \end{bmatrix} \tag{5.7}$$

for well chosen interpolation points $s_1, \dots, s_N$ and tangential directions $\mathbf{b}_1, \dots, \mathbf{b}_N$ as described in (2.56). The RBM in [126] only considers the bases $\boldsymbol{\mathcal{Z}}_\mathrm{BT}(c, g)$ from (5.6). However, in this work, we use both options, $\boldsymbol{\mathcal{Z}}_\mathrm{BT}(c, g)$ and $\boldsymbol{\mathcal{Z}}_\mathrm{IRKA}(c, g)$, to approximate the controllability space and denote the corresponding basis as $\boldsymbol{\mathcal{Z}}_\mathbf{V}(c, g)$.

Since we can not evaluate an infinite number of parameters in $\mathfrak{D}$, we define a test-parameter set

$$\mathfrak{D}_{\text{Test},c} \times \mathfrak{D}_{\text{Test},g} = \mathfrak{D}_{\text{Test}} \subset \mathfrak{D} = \mathfrak{D}_c \times \mathfrak{D}_g,$$

which is finite and densely distributed in $\mathfrak{D}$. For this test-parameter set $\mathfrak{D}_{\text{Test}}$, we will derive a space $\mathcal{V}_{\text{r}}$ that approximates the solution space of the Lyapunov equation (5.3). Additionally, this test-parameter set is used to evaluate the quality of the reduced solution space $\mathcal{V}_{\text{r}}$. Since the test-parameter set $\mathfrak{D}_{\text{Test}}$ is assumed to be well chosen in $\mathfrak{D}$, we expect that the space $\mathcal{V}_{\text{r}}$ approximates the solution space for all parameters in $\mathfrak{D}$ if it does for all parameters in $\mathfrak{D}_{\text{Test}}$.

We start constructing a basis $\mathbf{V}_{\text{r}}$ that spans the reduced space $\mathcal{V}_{\text{r}}$ by picking one test-parameter $(c_0,\, g_0) \in \mathfrak{D}_{\text{Test}}$. For this parameter $(c_0,\, g_0)$, we compute a basis $\mathfrak{Z}_{\mathbf{V}}(c_0,\, g_0)$ as described in (5.6) or (5.7), which yields the first orthonormal basis

$$\mathbf{V}_{\text{r}} := \operatorname{orth}(\mathfrak{Z}_{\mathbf{V}}(c_0,\, g_0)).$$

Remark 5.1 describes a detailed implementation of the basis orthonormalization. After forming our first basis, we evaluate the quality of the Gramian approximation for all remaining parameters in $\mathfrak{D}_{\text{Test}}$. To this aim, we compute the error approximation $\boldsymbol{\Delta}(c, g)$ for all these parameters $(c,\, g) \in \mathfrak{D}_{\text{Test}}$ and define the largest one as

$$\boldsymbol{\Delta}^{\max} := \boldsymbol{\Delta}(c_1,\, g_1) := \max_{(c,\, g) \in \mathfrak{D}_{\text{Test}}} \boldsymbol{\Delta}(c, g)$$

where $(c_1,\, g_1)$ is the parameter pair that leads to the largest error approximation value. If $\boldsymbol{\Delta}^{\max}$ is larger than a given tolerance tol, we know that the current basis does not approximate the solution space well enough for at least one pair of parameters $(c_1,\, g_1)$. Hence, we need to enlarge the basis $\mathbf{V}_{\text{r}}$. Therefore, we enrich the basis $\mathbf{V}_{\text{r}}$ by the controllability space approximation $\mathfrak{Z}_{\mathbf{V}}(c_1,\, g_1)$ for the parameters $(c_1,\, g_1)$ that result in this largest error approximation. We compute the basis $\mathfrak{Z}_{\mathbf{V}}(c_1,\, g_1)$ that is equal to $\mathfrak{Z}_{\text{BT}}(c_1,\, g_1)$ or $\mathfrak{Z}_{\text{IRKA}}(c_1,\, g_1)$ and set

$$\mathbf{V}_{\text{r}} = \operatorname{orth}(\begin{bmatrix}\mathbf{V}_{\text{r}} & \mathfrak{Z}_{\mathbf{V}}(c_1,\, g_1)\end{bmatrix}).$$

We continue this procedure until the maximal error approximation $\boldsymbol{\Delta}^{\max}$ is smaller than the tolerance tol. That means that for all parameters $(c, g) \in \mathfrak{D}_{\text{Test}}$, the solution space is well approximated by $\mathcal{V}_{\text{r}}$ which is spanned by the basis $\mathbf{V}_{\text{r}}$. If $\mathfrak{D}_{\text{Test}}$ is chosen well in $\mathfrak{D}$, also, the solution space of the remaining parameters in $\mathfrak{D}$ is approximated by $\mathcal{V}_{\text{r}}$.

Remark 5.1:
The orthonormalization operator $\operatorname{orth}(\mathfrak{Z}_{\mathbf{V}})$ is implemented in such a way that basis vectors of $\mathfrak{Z}_{\mathbf{V}}$ corresponding to singular values close to zero are truncated and not included in the resulting basis. Additionally, we add a maximum for the basis dimension

$N_{\max}$. To implement this, we compute a singular value decomposition $\mathfrak{Z}_{\mathbf{V}} = \mathbf{U}\mathbf{\Sigma}\mathbf{X}^{\mathrm{T}}$ with $\mathbf{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_{n_Z})$. We set

$$\operatorname{orth}(\mathfrak{Z}_{\mathbf{V}}) := \mathbf{U}[\,:\,,\ 1:q\,], \quad \text{with} \quad \mathrm{tol}_{\mathrm{V}} \cdot \sigma_1 > \sigma_{k+1}, \quad q = \min\{k, N_{\max}\},$$

where $\mathrm{tol}_{\mathrm{V}}$ is a given tolerance and k is the smallest index that satisfies $\mathrm{tol}_{\mathrm{V}} \cdot \sigma_1 > \sigma_{k+1}$. That way, we only use the most dominant columns of $\mathfrak{Z}_{\mathbf{V}}$ to form the basis. ◊

Online phase After we have computed a basis $\mathbf{V}_{\mathrm{r}}$ that spans a space $\boldsymbol{\mathcal{V}}_{\mathrm{r}}$ approximating the solution space of the Lyapunov equation (5.3) in the offline phase, we derive a reduced Lyapunov equation that is fast solvable and approximates the solution of (5.3) for all parameters $(c,\ g) \in \mathfrak{D}$. To do so, we define the reduced matrices as

$$\boldsymbol{\mathcal{E}}_{\mathrm{r}} := \mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\mathbf{V}_{\mathrm{r}}, \qquad \boldsymbol{\mathcal{A}}_{\mathrm{r}}(c,g) := \mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{r}}, \qquad \boldsymbol{\mathcal{B}}_{\mathrm{r}} := \mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}. \tag{5.8}$$

Then for all parameters $(c, g) \in \mathfrak{D}$ we can compute an approximation of the solution $\boldsymbol{\mathcal{P}}(c,g)$ as described in (5.4) where $\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)$ is the solution of the reduced Lyapunov equation

$$\boldsymbol{\mathcal{E}}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\boldsymbol{\mathcal{A}}_{\mathrm{r}}(c,g)^{\mathrm{T}} + \boldsymbol{\mathcal{A}}_{\mathrm{r}}(c,g)\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\boldsymbol{\mathcal{E}}_{\mathrm{r}}^{\mathrm{T}} = -\boldsymbol{\mathcal{B}}_{\mathrm{r}}\boldsymbol{\mathcal{B}}_{\mathrm{r}}^{\mathrm{T}}, \tag{5.9}$$

which has dimension r corresponding to the number of vectors in $\mathbf{V}_{\mathrm{r}}$.

Error approximation For the reduced basis method presented above, error approximations are needed to evaluate the quality of the resulting basis $\mathbf{V}_{\mathrm{r}}$. We can estimate different quantities to obtain error approximations. One option is to evaluate the norm of the error in the solution of the Lyapunov equation (5.3), that is $\|\mathfrak{E}(c,g)\|$. The respective error is defined as $\mathfrak{E}(c,g) := \boldsymbol{\mathcal{P}}(c,g) - \widetilde{\boldsymbol{\mathcal{P}}}(c,g)$ where the approximated solution $\widetilde{\boldsymbol{\mathcal{P}}}(c,g)$ is as described in (5.4). There exist various upper bounds of the error norm $\|\mathfrak{E}(c,g)\|$ that are based on the residual

$$\mathfrak{R}(c,g) := \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}} + \boldsymbol{\mathcal{A}}(c,g)\widetilde{\boldsymbol{\mathcal{P}}}(c,g)\boldsymbol{\mathcal{E}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}}\widetilde{\boldsymbol{\mathcal{P}}}(c,g)\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}}. \tag{5.10}$$

Examples are described in [63, 120, 126]. Often, these bounds are rather conservative and might not apply to these examples. Hence, we aim to find another approximation of the error norm $\|\mathfrak{E}(c,g)\|$.

To do so, we consider the error equation

$$\boldsymbol{\mathcal{A}}(c,g)\mathfrak{E}(c,g)\boldsymbol{\mathcal{E}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}}\mathfrak{E}(c,g)\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}} = -\mathfrak{R}(c,g) \tag{5.11}$$

and observe, that the error $\mathfrak{E}(c,g)$ is the solution of this error equation for $\mathfrak{R}(c,g)$ as defined in (5.10). Hence, we can apply a second RBM to approximate the error space spanned by $\boldsymbol{\mathfrak{E}}(c, g)$ and to determine an error approximation $\widetilde{\boldsymbol{\mathfrak{E}}}(c, g)$ with $\boldsymbol{\mathfrak{E}}(c, g) \approx \widetilde{\boldsymbol{\mathfrak{E}}}(c, g)$. Therefore, we derive a basis $\mathbf{V}_{\mathrm{err}}$ that spans an approximation of the solution space of the error equation in (5.11). To avoid confusion, we denote the second RBM that determines the basis for the error approximation in the following as EE-RBM.

We can write the controllability Gramian that is the solution of the Lyapunov equation (5.3) as $\boldsymbol{\mathfrak{P}}(c, g) = \mathbf{V}\boldsymbol{\mathfrak{X}}(c, g)\mathbf{V}^{\mathrm{T}}$ where $\mathbf{V}$ is a basis that spans the (full-order) solution space of the Lyapunov equation (5.3) for all parameters $(c, g) \in \boldsymbol{\mathfrak{D}}$. The respective error is then given as

$$\boldsymbol{\mathfrak{E}}(c, g) = \mathbf{V}\boldsymbol{\mathfrak{X}}(c, g)\mathbf{V}^{\mathrm{T}} - \mathbf{V}_{\mathrm{r}}\boldsymbol{\mathfrak{P}}_{\mathrm{r}}(c, g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}},$$

and, hence, we obtain that the error lies in the space spanned by the basis $\mathbf{V}_{\boldsymbol{\mathfrak{E}}} = \operatorname{orth}(\begin{bmatrix}\mathbf{V}_{\mathrm{r}} & \mathbf{V}\end{bmatrix})$ for all parameters. Since $\mathbf{V}_{\mathrm{r}}$ is computed within the first RBM, the remaining task is to determine the basis $\mathbf{V}$. However, the basis $\mathbf{V}$ is not available. Otherwise, we would have a basis that spans the solution space of the Lyapunov equation (5.3) for all parameters without an error. Hence, we apply the second EE-RBM and derive an approximation of $\mathbf{V}_{\boldsymbol{\mathfrak{E}}}$ called $\mathbf{V}_{\mathrm{err}}$.

Because of the structure of the basis $\mathbf{V}_{\boldsymbol{\mathfrak{E}}}$, adding $\mathbf{V}_{\mathrm{r}}$ and $\boldsymbol{\mathfrak{Z}}_{\mathbf{V}}(c^{\mathrm{e}}, g^{\mathrm{e}})$ to the basis is equivalent to adding a factor $\boldsymbol{\mathfrak{Z}}_{\boldsymbol{\mathfrak{E}}}(c^{\mathrm{e}}, g^{\mathrm{e}})$ with $\boldsymbol{\mathfrak{E}}(c^{\mathrm{e}}, g^{\mathrm{e}}) \approx \boldsymbol{\mathfrak{Z}}_{\boldsymbol{\mathfrak{E}}}(c^{\mathrm{e}}, g^{\mathrm{e}})\boldsymbol{\mathfrak{Z}}_{\boldsymbol{\mathfrak{E}}}(c^{\mathrm{e}}, g^{\mathrm{e}})^{\mathrm{T}}$ to the basis $\mathbf{V}_{\mathrm{err}}$. This computation is of a more advantageous structure because of the low-rank right-hand side $\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}$ in (5.3) compared to $\boldsymbol{\mathfrak{R}}(c, g)$ in (5.11). Hence, in every step of EE-RBM, we compute a basis $\boldsymbol{\mathfrak{Z}}_{\mathbf{V}}(c^{\mathrm{e}}, g^{\mathrm{e}})$ in a parameter pair $(c^{\mathrm{e}}, g^{\mathrm{e}})$ to enrich the basis of the error equation (5.11) as $\mathbf{V}_{\mathrm{err}} = \operatorname{orth}(\begin{bmatrix}\mathbf{V}_{\mathrm{err}} & \mathbf{V}_{\mathrm{r}} & \boldsymbol{\mathfrak{Z}}_{\mathbf{V}}(c^{\mathrm{e}}, g^{\mathrm{e}})\end{bmatrix})$ and therefore build a basis that approximates $\boldsymbol{\mathcal{V}}_{\boldsymbol{\mathfrak{E}}}$.

Using the basis $\mathbf{V}_{\mathrm{err}}$, we determine the approximation

$$\widetilde{\boldsymbol{\mathfrak{E}}}(c, g) = \mathbf{V}_{\mathrm{err}}\widehat{\boldsymbol{\mathfrak{E}}}(c, g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}} \tag{5.12}$$

where $\widehat{\boldsymbol{\mathfrak{E}}}(c, g)$ solves the reduced error equation

$$\begin{aligned}\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c, g)\mathbf{V}_{\mathrm{err}}\widehat{\boldsymbol{\mathfrak{E}}}(c, g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{V}_{\mathrm{err}} + \mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\mathbf{V}_{\mathrm{err}}\widehat{\boldsymbol{\mathfrak{E}}}(c, g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c, g)^{\mathrm{T}}\mathbf{V}_{\mathrm{err}}& \\ = -\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathfrak{R}}(c, g)\mathbf{V}_{\mathrm{err}}.&\end{aligned} \tag{5.13}$$

which results in the error approximation

$$\boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}(c, g) := \|\widetilde{\boldsymbol{\mathfrak{E}}}(c, g)\|_{\mathrm{F}} = \|\mathbf{V}_{\mathrm{err}}\widehat{\boldsymbol{\mathfrak{E}}}(c, g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\|_{\mathrm{F}} = \|\widehat{\boldsymbol{\mathfrak{E}}}(c, g)\|_{\mathrm{F}}. \tag{5.14}$$

Using this procedure, we derive an error approximation $\widetilde{\boldsymbol{\mathfrak{E}}}(c, g)$ that is fast computable if the basis dimension of $\mathbf{V}_{\mathrm{err}}$ is sufficiently small.

Both, RBM and EE-RBM, run in parallel. The first parameters $(c_0,\ g_0)$ and $(c_0^{\mathrm{e}},\ g_0^{\mathrm{e}})$ are chosen arbitrarily in $\boldsymbol{\mathfrak{D}}_{\mathrm{Test}}$ with $(c_0,\ g_0) \neq (c_0^{\mathrm{e}},\ g_0^{\mathrm{e}})$. We compute the basis $\mathbf{V}_{\mathrm{r}}$ as described above and, in addition, determine $\boldsymbol{\mathfrak{Z}}_{\mathrm{IRKA}}(c_0^{\mathrm{e}}, g_0^{\mathrm{e}})$ or $\boldsymbol{\mathfrak{Z}}_{\mathrm{BT}}(c_0^{\mathrm{e}}, g_0^{\mathrm{e}})$ to obtain $\boldsymbol{\mathfrak{Z}}_{\mathbf{V}}(c_0^{\mathrm{e}}, g_0^{\mathrm{e}})$ such that our first error space basis is given as

$$\mathbf{V}_{\mathrm{err}} = \operatorname{orth}\left(\begin{bmatrix}\mathbf{V}_{\mathrm{r}} & \boldsymbol{\mathfrak{Z}}_{\mathbf{V}}(c_0^{\mathrm{e}}, g_0^{\mathrm{e}})\end{bmatrix}\right) = \operatorname{orth}\left(\begin{bmatrix}\boldsymbol{\mathfrak{Z}}_{\mathbf{V}}(c_0, g_0) & \boldsymbol{\mathfrak{Z}}_{\mathbf{V}}(c_0^{\mathrm{e}}, g_0^{\mathrm{e}})\end{bmatrix}\right).$$

As described above, the consecutive parameter $(c_1,\, g_1)$ is the one that leads to the largest error approximation $\boldsymbol{\Delta}(c,g)$, and we use the corresponding controllability space basis $\mathcal{Z}_{\mathbf{V}}(c_1, g_1)$ to enrich the basis $\mathbf{V}_{\mathrm{r}}$. The consecutive parameter $(c_1^{\mathrm{e}},\, g_1^{\mathrm{e}})$ is chosen to be the one that results in the largest residual of the error equation in the Frobenius norm, i.e., the parameters $(c_1^{\mathrm{e}},\, g_1^{\mathrm{e}})$ that lead to the largest value $\|\mathfrak{R}^{\mathrm{e}}(c,g)\|_{\mathrm{F}}$ with

$$\begin{aligned}\mathfrak{R}^{\mathrm{e}}(c,g) &:= \begin{bmatrix}\mathbf{R}_{11}^{\mathrm{e}}(c,g) & \mathbf{R}_{12}^{\mathrm{e}}(c,g)\\ \mathbf{R}_{12}^{\mathrm{e}}(c,g)^{\mathrm{T}} & \mathbf{R}_{22}^{\mathrm{e}}(c,g)\end{bmatrix}\\ &:= \boldsymbol{\mathcal{A}}(c,g)\widetilde{\boldsymbol{\mathcal{P}}}(c,g)\boldsymbol{\mathcal{E}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}}\widetilde{\boldsymbol{\mathcal{P}}}(c,g)\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}}\\ &\qquad\qquad + \boldsymbol{\mathcal{A}}(c,g)\widetilde{\mathfrak{E}}(c,g)\boldsymbol{\mathcal{E}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}}\widetilde{\mathfrak{E}}(c,g)\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}} + \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}.\end{aligned} \tag{5.15}$$

We compute $\mathcal{Z}_{\mathbf{V}}(c_1^{\mathrm{e}}, g_1^{\mathrm{e}})$ equal to $\mathcal{Z}_{\mathrm{IRKA}}(c_1^{\mathrm{e}}, g_1^{\mathrm{e}})$ or $\mathcal{Z}_{\mathrm{BT}}(c_1^{\mathrm{e}}, g_1^{\mathrm{e}})$ and generate the next error equation basis

$$\mathbf{V}_{\mathrm{err}} = \operatorname{orth}\left(\begin{bmatrix}\mathbf{V}_{\mathrm{err}} & \mathbf{V}_{\mathrm{r}} & \mathcal{Z}_{\mathbf{V}}(c_1^{\mathrm{e}},\, g_1^{\mathrm{e}})\end{bmatrix}\right) = \operatorname{orth}\left(\begin{bmatrix}\mathbf{V}_{\mathrm{err}} & \mathcal{Z}_{\mathbf{V}}(c_1,\, g_1) & \mathcal{Z}_{\mathbf{V}}(c_1^{\mathrm{e}},\, g_1^{\mathrm{e}})\end{bmatrix}\right).$$

Again, we continue this procedure until the largest error approximation is smaller than a given tolerance tol. The first RBM combined with the EE-RBM results in Algorithm 14.

We observe that the first steps of the RBM, together with the EE-RBM, lead to rough error approximations since the basis $\mathbf{V}_{\mathrm{err}}$ includes only a few solutions. However, the larger and therefore better the basis $\mathbf{V}_{\mathrm{r}}$ is, the more detailed is $\mathbf{V}_{\mathrm{err}}$, and we know that the stopping criterion $\boldsymbol{\Delta}^{\max} > \mathrm{tol}$ is meaningful.

Remark 5.2:
It turns out that adding $\mathcal{Z}_{\mathbf{V}}(c, 0)$, where 0 is the zero vector of the appropriate dimension, to the basis $\mathbf{V}_{\mathrm{r}}$ leads to a more robust basis. Thus, we add $\mathcal{Z}_{\mathbf{V}}(c, 0)$ corresponding to the undamped system to our basis $\mathbf{V}_{\mathrm{r}}$. Since this basis is independent of the damping values, we calculate it beforehand and do not include this calculation in our procedure.$\diamondsuit$

Remark 5.3:
In practice, we compute $\mathfrak{R}_{\mathrm{r}}(c,g)$ more efficiently. Therefore, we make use of the trace formulation of the Frobenius norm and utilize the low-rank representations $\widetilde{\mathfrak{E}}(c,g) = \mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}$ and $\widetilde{\boldsymbol{\mathcal{P}}}(c,g) = \mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}$, and the trace properties to obtain the fast

Algorithm 14 Offline phase of the first-order RBM.

```
Input: $\boldsymbol{\mathcal{E}} \in \mathbb{R}^{N \times N}$, $\boldsymbol{\mathcal{A}} : \boldsymbol{\mathfrak{D}} \to \mathbb{R}^{N \times N}$ asymptotically stable, $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{N \times m}$, test-parameter set $\boldsymbol{\mathfrak{D}}_{\mathrm{Test}}$, tolerance tol.
Output: Orthonormal bases $\mathbf{V}_{\mathrm{r}}$, $\mathbf{V}_{\mathrm{err}}$.
Choose any $(c_0,\, g_0)$, $(c_0^{\mathrm{e}},\, g_0^{\mathrm{e}}) \in \boldsymbol{\mathfrak{D}}_{\mathrm{Test}}$ with $(c_0,\, g_0) \neq (c_0^{\mathrm{e}},\, g_0^{\mathrm{e}})$.
Determine a basis $\boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_0, g_0)$ either as $\boldsymbol{\mathcal{Z}}_{\mathrm{BT}}(c_0,\, g_0)$ from (5.6) or $\boldsymbol{\mathcal{Z}}_{\mathrm{IRKA}}(c_0,\, g_0)$ from (5.7).
Set $\mathcal{M} := \{(c_0,\, g_0)\}$.
Set $\mathbf{V}_{\mathrm{r}} := \operatorname{orth}(\boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_0,\, g_0))$.
Determine a basis $\boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_0^{\mathrm{e}},\, g_0^{\mathrm{e}})$ either as $\boldsymbol{\mathcal{Z}}_{\mathrm{BT}}(g_0^{\mathrm{e}},\, c_0^{\mathrm{e}})$ from (5.6) or $\boldsymbol{\mathcal{Z}}_{\mathrm{IRKA}}(c_0^{\mathrm{e}},\, g_0^{\mathrm{e}})$ from (5.7).
Set $\mathbf{V}_{\mathrm{err}} := \operatorname{orth}(\begin{bmatrix} \boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_0,\, g_0) & \boldsymbol{\mathcal{Z}}_{\mathbf{V}}(g_0^{\mathrm{e}},\, c_0^{\mathrm{e}}) \end{bmatrix})$.
Set $k := 1$.
Determine $(c_1, g_1) := \operatorname{argmax}_{(c,g) \in \boldsymbol{\mathfrak{D}}_{\mathrm{Test}} \setminus \mathcal{M}} \boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}(c, g)$.
Set $\boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}^{\max} := \boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}(c_1, g_1)$.
Determine $(c_1^{\mathrm{e}},\, g_1^{\mathrm{e}}) := \operatorname{argmax}_{(c,g) \in \boldsymbol{\mathfrak{D}}_{\mathrm{Test}} \setminus \mathcal{M}} \|\boldsymbol{\mathfrak{R}}^{\mathrm{e}}(c, g)\|_{\mathrm{F}}$.
while $\boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}^{\max} > \mathrm{tol}$ do
    Determine a basis $\boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_k, g_k)$ either as $\boldsymbol{\mathcal{Z}}_{\mathrm{BT}}(c_k, g_k)$ from (5.6) or $\boldsymbol{\mathcal{Z}}_{\mathrm{IRKA}}(c_k, g_k)$ from (5.7).
    Set $\mathcal{M} := \mathcal{M} \cup \{(c_k, g_k)\}$.
    Set $\mathbf{V}_{\mathrm{r}} := \operatorname{orth}(\begin{bmatrix} \mathbf{V}_{\mathrm{r}} & \boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_k, g_k) \end{bmatrix})$.
    Determine a basis $\boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_k^{\mathrm{e}},\, g_k^{\mathrm{e}})$ either as $\boldsymbol{\mathcal{Z}}_{\mathrm{BT}}(c_k^{\mathrm{e}},\, g_k^{\mathrm{e}})$ from (5.6) or $\boldsymbol{\mathcal{Z}}_{\mathrm{IRKA}}(c_k^{\mathrm{e}},\, g_k^{\mathrm{e}})$ from (5.7).
    Set $\mathbf{V}_{\mathrm{err}} := \operatorname{orth}(\begin{bmatrix} \mathbf{V}_{\mathrm{err}} & \boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_k, g_k) & \boldsymbol{\mathcal{Z}}_{\mathbf{V}}(c_k^{\mathrm{e}},\, g_k^{\mathrm{e}}) \end{bmatrix})$.
    Determine $(c_{k+1}, g_{k+1}) := \operatorname{argmax}_{(c,g) \in \boldsymbol{\mathfrak{D}}_{\mathrm{Test}} \setminus \mathcal{M}} \boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}(c, g)$.
    Set $\boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}^{\max} := \boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}(c_{k+1},\, g_{k+1})$.
    Determine $(c_{k+1}^{\mathrm{e}},\, g_{k+1}^{\mathrm{e}}) := \operatorname{argmax}_{(c,g) \in \boldsymbol{\mathfrak{D}}_{\mathrm{Test}} \setminus \mathcal{M}} \|\boldsymbol{\mathfrak{R}}^{\mathrm{e}}(c, g)\|_{\mathrm{F}}$.
    Set $k := k + 1$.
end while
```

computable residual representation

$$\begin{aligned}
\|\mathfrak{R}_{\mathrm{r}}(c,g)\|_{\mathrm{F}}^2 = {} & 2\,\mathrm{tr}\Big(\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\Big)\\
& + 2\,\mathrm{tr}\Big(\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\Big)\\
& + 2\,\mathrm{tr}\big(\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\big)\\
& + 2\,\mathrm{tr}\big(\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\big)\\
& + 4\,\mathrm{tr}\Big(\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\Big)\\
& + 4\,\mathrm{tr}\Big(\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}\mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\Big)\\
& + 4\,\mathrm{tr}\Big(\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}\Big)\\
& + 4\,\mathrm{tr}\big(\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{A}}(c,g)\mathbf{V}_{\mathrm{r}}\boldsymbol{\mathcal{P}}_{\mathrm{r}}(c,g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}\big)\\
& + \mathrm{tr}\big(\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\mathcal{B}}\big)\,. \qquad\qquad \diamondsuit
\end{aligned}$$

5.1.2 Decoupling of the controllability space of first-order systems

From Theorem 2.22 and Theorem 2.23 it follows that the controllability space of the first-order systems (1.5) and (1.6) is spanned by

$$\boldsymbol{\mathcal{V}}(c,g) = \mathrm{span}\left\{\left(s_1\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}(c,g)\right)^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_1, \ldots \left(s_M\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}(c,g)\right)^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_M\right\} \tag{5.16}$$

if the interpolation points s_j and the tangential directions $\mathbf{b}_j$ are chosen correctly (e.g., the poles of the system) for $j = 1, \ldots, M$.

We consider the first-order representation of the mechanical systems with matrices as in (1.7) so that for every interpolation point s_j, $j = 1, \ldots, M$ we get

$$(s_j\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}(c,g)) = \boldsymbol{\Gamma}(s_j) + \boldsymbol{\mathcal{F}}(c)\mathbf{G}(g)\boldsymbol{\mathcal{F}}(c)^{\mathrm{T}}$$

with

$$\boldsymbol{\Gamma}(s_j) := \begin{bmatrix} s_j\mathbf{I} & -\mathbf{I} \\ \mathbf{K} & s_j\mathbf{M} + \mathbf{D}_{\mathrm{int}} \end{bmatrix}, \quad \boldsymbol{\mathcal{F}}(c) := \begin{bmatrix} 0 \\ \mathbf{F}(c) \end{bmatrix}. \tag{5.17}$$

This representation is used to decouple parameter-independent from parameter-dependent components as shown in the following lemma.

Lemma 5.4:
Consider the controllability space $\boldsymbol{\mathcal{V}}(c,g)$ as defined in (5.16) with interpolation points $\widetilde{s}_1,\dots,\widetilde{s}_M$ and tangential directions $\widetilde{\mathbf{b}}_1,\dots,\widetilde{\mathbf{b}}_M$ that spans the controllability space of systems (1.5) and (1.6) with matrices (1.7). Then, this space fulfills

$$\boldsymbol{\mathcal{V}}(c,g) \subseteq \boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{B}}} \cup \boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{F}}}(c),$$

with spaces

$$\begin{aligned} \boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{B}}} &:= \operatorname{span}\left\{\boldsymbol{\Gamma}(s_1)^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_1,\dots,\boldsymbol{\Gamma}(s_M)^{-1}\boldsymbol{\mathcal{B}}\mathbf{b}_M\right\}, \qquad (5.18)\\ \boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{F}}}(c) &:= \operatorname{span}\left\{\boldsymbol{\Gamma}(m_1)^{-1}\boldsymbol{\mathcal{F}}(c)\mathbf{f}_1,\dots,\boldsymbol{\Gamma}(m_M)^{-1}\boldsymbol{\mathcal{F}}(c)\mathbf{f}_M\right\} \qquad (5.19)\end{aligned}$$

for interpolation points $s_1,\dots,s_{M_{\mathbf{B}}}$, $m_1,\dots,m_{M_{\mathbf{F}}}$ and tangential directions $\mathbf{b}_1,\dots,\mathbf{b}_{M_{\mathbf{B}}}$, $\mathbf{f}_1,\dots,\mathbf{f}_{M_{\mathbf{F}}}$ that are chosen in such a way, that

$$\boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{B}}} = \operatorname{span}\left\{\boldsymbol{\Gamma}(s)^{-1}\boldsymbol{\mathcal{B}} \,|\, s\in\mathbb{R}\right\} \qquad \text{and} \qquad \boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{F}}}(c) = \operatorname{span}\left\{\boldsymbol{\Gamma}(m)^{-1}\boldsymbol{\mathcal{F}}(c) \,|\, m\in\mathbb{R}\right\}$$

for $\boldsymbol{\Gamma}(s)$ as defined in (5.17). ♢

Proof. We apply the Sherman-Morrison-Woodbury formula for every entry in (5.16) to obtain

$$\begin{aligned}\left(\widetilde{s}_j\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}(c,g)\right)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_j &= \left(\boldsymbol{\Gamma}(s_j) + \boldsymbol{\mathcal{F}}(c)\mathbf{G}(g)\boldsymbol{\mathcal{F}}(c)^{\mathrm{T}}\right)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_j\\ &= \boldsymbol{\Gamma}(\widetilde{s}_j)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_j\\ &\quad - \boldsymbol{\Gamma}(\widetilde{s}_j)^{-1}\boldsymbol{\mathcal{F}}(c)\left(\mathbf{G}(g)^{-1} + \boldsymbol{\mathcal{F}}(c)^{\mathrm{T}}\boldsymbol{\Gamma}(\widetilde{s}_j)^{-1}\boldsymbol{\mathcal{F}}(c)\right)^{-1}\boldsymbol{\mathcal{F}}(c)^{\mathrm{T}}\boldsymbol{\Gamma}(\widetilde{s}_j)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_j.\end{aligned}$$

Hence, the j-th subspace $\operatorname{span}\left\{\left(\widetilde{s}_j\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}(c,g)\right)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_j\right\}$ of $\boldsymbol{\mathcal{V}}(c,g)$ satisfies

$$\operatorname{span}\left\{\left(\widetilde{s}_j\boldsymbol{\mathcal{E}} - \boldsymbol{\mathcal{A}}(c,g)\right)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_j\right\} \subseteq \operatorname{span}\left\{\boldsymbol{\Gamma}(\widetilde{s}_j)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_j,\ \boldsymbol{\Gamma}(\widetilde{s}_j)^{-1}\boldsymbol{\mathcal{F}}(c)\widetilde{\mathbf{f}}_j(c)\right\},$$

where $\widetilde{\mathbf{f}}_j(c) := \left(\mathbf{G}(g)^{-1} + \boldsymbol{\mathcal{F}}(c)^{\mathrm{T}}\boldsymbol{\Gamma}(\widetilde{s}_j)^{-1}\boldsymbol{\mathcal{F}}(c)\right)^{-1}\boldsymbol{\mathcal{F}}(c)^{\mathrm{T}}\boldsymbol{\Gamma}(\widetilde{s}_j)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_j$. From that, it follows that

$$\begin{aligned}\boldsymbol{\mathcal{V}}(c,g) &\subseteq \operatorname{span}\left\{\boldsymbol{\Gamma}^{-1}(\widetilde{s}_1)\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_1,\dots\boldsymbol{\Gamma}(\widetilde{s}_M)^{-1}\boldsymbol{\mathcal{B}}\widetilde{\mathbf{b}}_M\right\}\\ &\qquad\qquad \cup \operatorname{span}\left\{\boldsymbol{\Gamma}^{-1}(\widetilde{s}_1)\boldsymbol{\mathcal{F}}(c)\widetilde{\mathbf{f}}_1(c),\dots\boldsymbol{\Gamma}(\widetilde{s}_M)^{-1}\boldsymbol{\mathcal{F}}(c)\widetilde{\mathbf{f}}_M(c)\right\}\\ &\subseteq \boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{B}}} \cup \boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{F}}}(c).\end{aligned}$$

□

Note that the interpolation points and tangential directions s_k, m_j, $\mathbf{b}_k$, and $\mathbf{f}_j$, where $k = 1,\dots,M_{\mathbf{B}}$ and $j = 1,\dots,M_{\mathbf{F}}$, are chosen such that the spaces $\boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{F}}}(c)$ not only include the controllability space $\boldsymbol{\mathcal{V}}(c,g)$ but also span the controllability spaces defined by $\boldsymbol{\Gamma}_{\mathrm{so}}(s)$, $\boldsymbol{\mathcal{B}}$, and $\boldsymbol{\mathcal{F}}(c)$, respectively (see the upcoming systems (5.21) and (5.23)). This ensures that $\boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{B}}}$ are independent of the chosen interpolation points, and therefore, the derived space $\boldsymbol{\mathcal{V}}_{\boldsymbol{\mathcal{B}}}$ remains consistent for all parameters $(c,g)\in\boldsymbol{\mathfrak{D}}$.

Applying Lemma 5.4 for every parameter in $\boldsymbol{\mathfrak{D}}$ yields the following theorem.

Theorem 5.5:
Assume that $\boldsymbol{\mathcal{V}}(c,g)$, defined in (5.16), spans the controllability space of the systems (1.5) and (1.6) with matrices (1.7) for all parameter pairs $(c,g) \in \boldsymbol{\mathcal{D}}$. Also consider the spaces $\boldsymbol{\mathcal{V}}_{\mathcal{B}}$ and $\boldsymbol{\mathcal{V}}_{\mathcal{F}}(c)$ as defined in (5.18) and (5.19), respectively. Define the space

$$\boldsymbol{\mathcal{V}}_{\mathcal{F}} := \boldsymbol{\mathcal{V}}_{\mathcal{B}} \cup \bigcup_{c \in \boldsymbol{\mathcal{D}}_c} \boldsymbol{\mathcal{V}}_{\mathcal{F}}(c). \tag{5.20}$$

Then the controllability space $\boldsymbol{\mathcal{V}}(c,g)$ from (5.16) satisfies that $\boldsymbol{\mathcal{V}}(c,g) \subseteq \boldsymbol{\mathcal{V}}_{\mathcal{F}}$ for all $(c,g) \in \boldsymbol{\mathcal{D}}$. ♢

This theorem is useful for our considerations as it shows that the space $\boldsymbol{\mathcal{V}} = \bigcup_{(c,g)\in\boldsymbol{\mathcal{D}}} \boldsymbol{\mathcal{V}}(c,g)$ that we aim to approximate lies in the space $\boldsymbol{\mathcal{V}}_{\mathcal{F}}$. Hence, if we approximate $\boldsymbol{\mathcal{V}}_{\mathcal{F}}$ well, also $\boldsymbol{\mathcal{V}}$ is approximated.

We further investigate the spaces $\boldsymbol{\mathcal{V}}_{\mathcal{B}}$ and $\boldsymbol{\mathcal{V}}_{\mathcal{F}}(c)$. The space $\boldsymbol{\mathcal{V}}_{\mathcal{B}}$ is the controllability space of the undamped system

$$\boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) = \boldsymbol{\mathcal{A}}(c,0)\mathbf{z}(t) + \boldsymbol{\mathcal{B}}\mathbf{u}(t), \tag{5.21}$$

where $\boldsymbol{\mathcal{A}}(c,0)$ describes a system where the external damping viscosities g are equal to zero and, hence, no external damping is applied. The corresponding controllability Gramian that spans the controllability space of the undamped system is defined as

$$\boldsymbol{\mathcal{P}}_{\mathcal{B}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \boldsymbol{\Gamma}(\mathrm{i}\omega)\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{\mathrm{T}}\boldsymbol{\Gamma}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega \tag{5.22}$$

with $\boldsymbol{\Gamma}(\mathrm{i}\omega)$ as defined in (5.17). This Gramian is equal to the controllability Gramian $\boldsymbol{\mathcal{P}}$ from (2.5) and $\boldsymbol{\mathcal{P}}(c,0)$ as defined in (5.3) with an external damping value equal to zero. Also, the space $\boldsymbol{\mathcal{V}}_{\mathcal{F}}(c)$ spans the controllability space of the undamped system

$$\boldsymbol{\mathcal{E}}\dot{\mathbf{z}}(t) = \boldsymbol{\mathcal{A}}(c,0)\mathbf{z}(t) + \boldsymbol{\mathcal{F}}(c)\mathbf{u}(t), \qquad \boldsymbol{\mathcal{F}}(c) := \begin{bmatrix} 0 \\ \mathbf{F}(c) \end{bmatrix} \tag{5.23}$$

with a position-dependent input matrix $\boldsymbol{\mathcal{F}}(c)$. The corresponding controllability Gramian that spans the controllability space of system (5.23) is

$$\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \boldsymbol{\Gamma}(\mathrm{i}\omega)\boldsymbol{\mathcal{F}}(c)\boldsymbol{\mathcal{F}}(c)^{\mathrm{T}}\boldsymbol{\Gamma}(\mathrm{i}\omega)^{\mathrm{H}}\mathrm{d}\omega. \tag{5.24}$$

Hence, we can compute the two spaces by setting

$$\boldsymbol{\mathcal{V}}_{\mathcal{B}} = \operatorname{span}\{\boldsymbol{\mathcal{P}}_{\mathcal{B}}\}, \qquad \boldsymbol{\mathcal{V}}_{\mathcal{F}}(c) = \operatorname{span}\{\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c)\}.$$

In what follows, we use this Gramian representation to derive an error indicator that can be used within the RBM to describe the quality of the approximation of the controllability space by a reduced basis $\mathbf{V}_{\mathrm{r}}$.

Error indicator We derive an error indicator from the space decomposition introduced in (5.20). For that, we consider the respective system in modal form as presented in (5.1) that includes diagonal matrices as submatrices. We assume that we have a basis $\mathbf{V}_{\mathrm{r}} \in \mathbb{R}^{N \times 2r}$, $2r \ll N$ with $\boldsymbol{\mathcal{V}}_{\mathcal{B}} \subset \operatorname{span}\{\mathbf{V}_{\mathrm{r}}\}$, so that the controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c)$ is well approximated by a matrix $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c)$ that lies in that space spanned by $\mathbf{V}_{\mathrm{r}}$, i.e, there exist a matrix $\boldsymbol{\mathcal{P}}_{\mathcal{F},\mathrm{r}} \in \mathbb{R}^{2r \times 2r}$ with

$$\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c) \approx \widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c) = \mathbf{V}_{\mathrm{r}} \boldsymbol{\mathcal{P}}_{\mathcal{F},\mathrm{r}}(c) \mathbf{V}_{\mathrm{r}}^{\mathrm{T}}. \tag{5.25}$$

Then, the controllability space lies approximately in

$$\begin{aligned} \boldsymbol{\mathcal{V}}(c,g) &\subset \operatorname{span}\{\boldsymbol{\mathcal{P}}_{\mathcal{B}}\} \cup \operatorname{span}\{\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c)\} \\ &\approx \operatorname{span}\{\boldsymbol{\mathcal{P}}_{\mathcal{B}}\} \cup \operatorname{span}\left\{\widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c)\right\} = \operatorname{span}\{\boldsymbol{\mathcal{P}}_{\mathcal{B}}\} \cup \operatorname{span}\left\{\mathbf{V}_{\mathrm{r}} \boldsymbol{\mathcal{P}}_{\mathcal{F},\mathrm{r}}(c) \mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\right\} \\ &= \operatorname{span}\{\boldsymbol{\mathcal{P}}_{\mathcal{B}}\} \cup \operatorname{span}\{\mathbf{V}_{\mathrm{r}}\} = \operatorname{span}\{\mathbf{V}_{\mathrm{r}}\}. \end{aligned}$$

Hence, to determine the quality of the approximation of the controllability space $\boldsymbol{\mathcal{V}}(c,g)$ by the basis $\mathbf{V}_{\mathrm{r}}$, an appropriate criterion is to determine how good $\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c)$ is approximated by $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c)$, where the reduced Gramian $\boldsymbol{\mathcal{P}}_{\mathcal{F},\mathrm{r}}(c)$ solves the Lyapunov equation

$$\boldsymbol{\mathcal{A}}_{\mathrm{r}}(c,0) \boldsymbol{\mathcal{P}}_{\mathcal{F},\mathrm{r}}(c) \boldsymbol{\mathcal{E}}_{\mathrm{r}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}}_{\mathrm{r}} \boldsymbol{\mathcal{P}}_{\mathcal{F},\mathrm{r}}(c) \boldsymbol{\mathcal{A}}_{\mathrm{r}}(c,0)^{\mathrm{T}} = -\boldsymbol{\mathcal{F}}_{\mathrm{r}}(c) \boldsymbol{\mathcal{F}}_{\mathrm{r}}(c)^{\mathrm{T}}$$

with $\boldsymbol{\mathcal{A}}_{\mathrm{r}}(c,0)$, $\boldsymbol{\mathcal{E}}_{\mathrm{r}}$ as in (5.8), and $\boldsymbol{\mathcal{F}}_{\mathrm{r}}(c) := \mathbf{V}_{\mathrm{r}}^{\mathrm{T}} \boldsymbol{\mathcal{F}}(c)$. For that, we define the submatrices

$$\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c) = \begin{bmatrix} \mathbf{X}_{11}(c) & \mathbf{X}_{12}(c) \\ \mathbf{X}_{12}(c)^{\mathrm{T}} & \mathbf{X}_{22}(c) \end{bmatrix}, \qquad \widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c) = \begin{bmatrix} \mathbf{Y}_{11}(c) & \mathbf{Y}_{12}(c) \\ \mathbf{Y}_{12}(c)^{\mathrm{T}} & \mathbf{Y}_{22}(c) \end{bmatrix}. \tag{5.26}$$

Since the Gramian $\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c)$ satisfies the Lyapunov equation

$$\boldsymbol{\mathcal{A}}(c,0) \boldsymbol{\mathcal{P}}_{\mathcal{F}}(c) \boldsymbol{\mathcal{E}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}} \boldsymbol{\mathcal{P}}_{\mathcal{F}}(c) \boldsymbol{\mathcal{A}}(c,0)^{\mathrm{T}} + \boldsymbol{\mathcal{F}}(c) \boldsymbol{\mathcal{F}}(c)^{\mathrm{T}} = 0, \tag{5.27}$$

the approximation $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c)$ leads to the residual

$$\boldsymbol{\mathcal{A}}(c,0) \widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c) \boldsymbol{\mathcal{E}}^{\mathrm{T}} + \boldsymbol{\mathcal{E}} \widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c) \boldsymbol{\mathcal{A}}(c,0) + \boldsymbol{\mathcal{F}}(c) \boldsymbol{\mathcal{F}}(c)^{\mathrm{T}} := \boldsymbol{\mathfrak{R}}(c) := \begin{bmatrix} \mathbf{R}_{11}(c) & \mathbf{R}_{12}(c) \\ \mathbf{R}_{12}(c)^{\mathrm{T}} & \mathbf{R}_{22}(c) \end{bmatrix}. \tag{5.28}$$

Using the residual $\boldsymbol{\mathfrak{R}}(c)$, we can evaluate the trace of the error between $\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c)$ and $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c)$ as described in the following theorem, which serves as an error indicator within the RBM.

Theorem 5.6:
Consider the first-order system (5.23) with matrices (1.7) corresponding to a second-order system in modal form (5.1), the corresponding controllability Gramian $\boldsymbol{\mathcal{P}}_{\mathcal{F}}(c)$ as defined in (5.24), and the respective approximation $\widetilde{\boldsymbol{\mathcal{P}}}_{\mathcal{F}}(c)$ as defined in (5.25). Also

consider the matrix decompositions as described in (5.26) and the residual $\mathfrak{R}(c)$ as defined in (5.28). Then it holds

$$\begin{aligned}\boldsymbol{\Delta}_{\mathcal{P}_{\mathcal{F}}}(c) :=\; &\mathrm{tr}\Big(\mathcal{P}_{\mathcal{F}}(c) - \widetilde{\mathcal{P}}_{\mathcal{F}}(c)\Big)\\ =\; &\mathrm{tr}\big(\mathbf{R}_{12}(c)\mathbf{\Omega}^{-2}\big) + \frac{1}{2\alpha}\,\mathrm{tr}\big(\mathbf{\Omega}^{-3}\mathbf{R}_{22}(c)\big) + \left(\frac{1}{2\alpha} - \alpha\right)\mathrm{tr}\big(\mathbf{\Omega}^{-1}\mathbf{R}_{11}(c)\big)\\ &+ \frac{1}{4\alpha}\,\mathrm{tr}\big(\mathbf{\Omega}^{-1}\mathbf{R}_{22}(c)\big) + \frac{1}{4\alpha}\,\mathrm{tr}(\mathbf{\Omega}\mathbf{R}_{11}(c))\,.\end{aligned} \tag{5.29}$$

◊

Proof. The Lyapunov equation in (5.27) and the respective residual equation in (5.28) lead to the subequations

$$\mathbf{X}_{12}(c) + \mathbf{X}_{12}(c)^{\mathrm{T}} = 0, \tag{5.30a}$$

$$\mathbf{X}_{22}(c) - \mathbf{X}_{11}(c)\mathbf{\Omega}^2 - 2\alpha\mathbf{X}_{12}(c)\mathbf{\Omega} = 0, \tag{5.30b}$$

$$-\mathbf{\Omega}^2\mathbf{X}_{12}(c) - \mathbf{X}_{12}(c)^{\mathrm{T}}\mathbf{\Omega}^2 - 2\alpha\mathbf{\Omega}\mathbf{X}_{22}(c) - 2\alpha\mathbf{X}_{22}(c)\mathbf{\Omega} + \mathbf{F}(c)\mathbf{F}(c)^{\mathrm{T}} = 0, \tag{5.30c}$$

and

$$\mathbf{Y}_{12}(c) + \mathbf{Y}_{12}(c)^{\mathrm{T}} = \mathbf{R}_{11}(c), \tag{5.31a}$$

$$\mathbf{Y}_{22}(c) - \mathbf{Y}_{11}(c)\mathbf{\Omega}^2 - 2\alpha\mathbf{Y}_{12}(c)\mathbf{\Omega} = \mathbf{R}_{12}(c), \tag{5.31b}$$

$$-\mathbf{\Omega}^2\mathbf{Y}_{12}(c) - \mathbf{Y}_{12}(c)^{\mathrm{T}}\mathbf{\Omega}^2 - 2\alpha\mathbf{\Omega}\mathbf{Y}_{22}(c) - 2\alpha\mathbf{Y}_{22}(c)\mathbf{\Omega} + \mathbf{F}(c)\mathbf{F}(c)^{\mathrm{T}} = \mathbf{R}_{22}(c). \tag{5.31c}$$

From (5.30a) and (5.31a), it follows that $\mathbf{X}_{12}(c) = \mathbf{S}_{\mathbf{X}}(c)$ and $\mathbf{Y}_{12}(c) = \mathbf{S}_{\mathbf{Y}}(c) + \frac{1}{2}\mathbf{R}_{11}(c)$ where $\mathbf{S}_{\mathbf{X}}(c)$ and $\mathbf{S}_{\mathbf{Y}}(c)$ are skew-symmetric matrices so that we define the skew-symmetric matrix $\mathbf{S}_{\mathbf{XY}}(c) := \mathbf{S}_{\mathbf{X}}(c) - \mathbf{S}_{\mathbf{Y}}(c)$. Hence, it holds

$$\mathbf{X}_{12}(c) - \mathbf{Y}_{12}(c) = \mathbf{S}_{\mathbf{XY}}(c) - \frac{1}{2}\mathbf{R}_{11}(c). \tag{5.32}$$

We aim to compute the trace expression

$$\mathrm{tr}\Big(\mathcal{P}_{\mathcal{F}}(c) - \widetilde{\mathcal{P}}_{\mathcal{F}}(c)\Big) = \mathrm{tr}(\mathbf{X}_{11}(c) - \mathbf{Y}_{11}(c)) + \mathrm{tr}(\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c))\,.$$

For that, we consider both trace components separately. First, we consider $\mathrm{tr}(\mathbf{X}_{11}(c) - \mathbf{Y}_{11}(c))$. Therefore, we subtract the equation in (5.31b) from the one in (5.30b), insert (5.32), and multiply from the right by $\mathbf{\Omega}^{-2}$ to obtain

$$(\mathbf{X}_{11}(c) - \mathbf{Y}_{11}(c)) = \mathbf{R}_{12}(c)\mathbf{\Omega}^{-2} + (\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c))\mathbf{\Omega}^{-2} - 2\alpha\mathbf{S}_{\mathbf{XY}}(c)\mathbf{\Omega}^{-1} + \alpha\mathbf{R}_{11}(c)\mathbf{\Omega}^{-1}. \tag{5.33}$$

According to the equation in (5.33), to compute the error between $\mathbf{X}_{11}(c)$ and $\mathbf{Y}_{11}(c)$, we need to describe $(\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c))\boldsymbol{\Omega}^{-2}$ more detailed. We subtract the equation in (5.31c) from the one in (5.30c), insert(5.32), and multiply from the left by $\boldsymbol{\Omega}^{-3}$ to obtain

$$\begin{aligned}
&- \boldsymbol{\Omega}^{-1}\mathbf{S}_{\mathbf{XY}}(c) + \frac{1}{2}\boldsymbol{\Omega}^{-1}\mathbf{R}_{11}(c) - \boldsymbol{\Omega}^{-3}\mathbf{S}_{\mathbf{XY}}(c)^{\mathrm{T}}\boldsymbol{\Omega}^{2} + \frac{1}{2}\boldsymbol{\Omega}^{-3}\mathbf{R}_{11}(c)\boldsymbol{\Omega}^{2} \\
&\qquad - 2\alpha\boldsymbol{\Omega}^{-2}(\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c)) - 2\alpha\boldsymbol{\Omega}^{-3}(\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c))\boldsymbol{\Omega} = -\boldsymbol{\Omega}^{-3}\mathbf{R}_{22}(c).
\end{aligned}$$

Applying the trace operator yields

$$2\operatorname{tr}\big(\boldsymbol{\Omega}^{-2}(\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c))\big) = \frac{1}{2\alpha}\left(\operatorname{tr}\big(\boldsymbol{\Omega}^{-1}\mathbf{R}_{11}(c)\big) + \operatorname{tr}\big(\boldsymbol{\Omega}^{-3}\mathbf{R}_{22}(c)\big)\right) \tag{5.34}$$

since $\operatorname{tr}(\boldsymbol{\Omega}^{-1}(\mathbf{S}_{\mathbf{X}}(c) - \mathbf{S}_{\mathbf{Y}}(c))) = 0$ because of the skew-symmetry of $\mathbf{S}_{\mathbf{XY}}(c)$ and the symmetry of $\boldsymbol{\Omega}^{-1}$. Finally, we apply the trace operator to the equation in (5.33) and insert the equation from (5.34) to obtain

$$\begin{aligned}
&\operatorname{tr}(\mathbf{X}_{11}(c) - \mathbf{Y}_{11}(c)) \\
&\quad = \operatorname{tr}\big(\mathbf{R}_{12}(c)\boldsymbol{\Omega}^{-2}\big) + \operatorname{tr}\big((\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c))\boldsymbol{\Omega}^{-2}\big) + \alpha\operatorname{tr}\big((\mathbf{R}_{11}(c))\boldsymbol{\Omega}^{-1}\big) \\
&\quad = \operatorname{tr}\big(\mathbf{R}_{12}(c)\boldsymbol{\Omega}^{-2}\big) + \frac{1}{2\alpha}\operatorname{tr}\big(\boldsymbol{\Omega}^{-3}\mathbf{R}_{22}(c)\big) + \frac{1}{2\alpha}\operatorname{tr}\big(\boldsymbol{\Omega}^{-1}\mathbf{R}_{11}(c)\big) - \alpha\operatorname{tr}\big((\mathbf{R}_{11}(c))\boldsymbol{\Omega}^{-1}\big) \\
&\quad = \operatorname{tr}\big(\mathbf{R}_{12}(c)\boldsymbol{\Omega}^{-2}\big) + \frac{1}{2\alpha}\operatorname{tr}\big(\boldsymbol{\Omega}^{-3}\mathbf{R}_{22}(c)\big) + \left(\frac{1}{2\alpha} - \alpha\right)\operatorname{tr}\big(\boldsymbol{\Omega}^{-1}\mathbf{R}_{11}(c)\big).
\end{aligned}$$

Now, we derive a formula for the expression $\operatorname{tr}(\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c))$. Therefore, we subtract the equation in (5.31c) from the one in (5.30c), insert (5.32), multiply from the left by $\boldsymbol{\Omega}$, and apply the trace operator which yields

$$\operatorname{tr}((\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c))) = \frac{1}{4\alpha}\operatorname{tr}\big(\boldsymbol{\Omega}^{-1}\mathbf{R}_{22}(c)\big) + \frac{1}{4\alpha}\operatorname{tr}(\boldsymbol{\Omega}\mathbf{R}_{11}(c)).$$

We combine the two trace components to obtain

$$\begin{aligned}
\operatorname{tr}\Big(\mathcal{P}_{\mathcal{F}}(c) - \widetilde{\mathcal{P}}_{\mathcal{F}}(c)\Big) &= \operatorname{tr}(\mathbf{X}_{11}(c) - \mathbf{Y}_{11}(c)) + \operatorname{tr}(\mathbf{X}_{22}(c) - \mathbf{Y}_{22}(c)) \\
&= \operatorname{tr}\big(\mathbf{R}_{12}(c)\boldsymbol{\Omega}^{-2}\big) + \frac{1}{2\alpha}\operatorname{tr}\big(\boldsymbol{\Omega}^{-3}\mathbf{R}_{22}(c)\big) + \left(\frac{1}{2\alpha} - \alpha\right)\operatorname{tr}\big(\boldsymbol{\Omega}^{-1}\mathbf{R}_{11}(c)\big) \\
&\qquad + \frac{1}{4\alpha}\operatorname{tr}\big(\boldsymbol{\Omega}^{-1}\mathbf{R}_{22}(c)\big) + \frac{1}{4\alpha}\operatorname{tr}(\boldsymbol{\Omega}\mathbf{R}_{11}(c)). \qquad \square
\end{aligned}$$

5.1.3 Offline-online RBM with a decoupled controllability space for first-order systems

To accelerate the RBM introduced in Algorithm 14, in this section, we combine the offline-online RBM from Section 5.1.1 and the controllability space decomposition from Section 5.1.2. We again aim to build a basis $\mathbf{V}_\mathrm{r} \in \mathbb{R}^{N \times N_\mathbf{V}}$ that approximates the solution space of the Lyapunov equations in (5.3) for all possible parameters $(c, g) \in \mathfrak{D}$ and that spans approximately the controllability space $\mathcal{V}$ of the first-order system (1.5) or (1.6). Using this basis, we derive an approximation of $\mathfrak{P}(c, g)$ as described in (5.4).

However, in contrast to the first approach presented in (5.5) where we added the basis $\mathfrak{Z}_\mathbf{V}(c, g)$ to build the basis $\mathbf{V}_\mathrm{r}$, in this subsection, we utilize the decomposition of the controllability spaces as presented in Section 5.1.2. For that we repeat that the controllability spaces $\mathcal{V}(c, g)$ satisfy

$$\mathcal{V}(c, g) \subseteq \mathcal{V}_\mathfrak{F} := \mathcal{V}_\mathfrak{B} \cup \bigcup_{c \in \mathfrak{D}_c} \mathcal{V}_\mathfrak{F}(c)$$

as shown in (5.20) with $\mathcal{V}_\mathfrak{B}$ and $\mathcal{V}_\mathfrak{F}(c)$ as defined in (5.18) and(5.19).

The space $\mathcal{V}_\mathfrak{B}$ is equal to the controllability space spanned by the Gramian $\mathfrak{P}_\mathfrak{B}$ defined in (5.22). Similarly, the space $\mathcal{V}_\mathfrak{F}(c)$ coincides with the controllability space spanned by the Gramian $\mathfrak{P}_\mathfrak{F}(c)$ defined in (5.24). Again because of the low-rank structure of $\mathfrak{B}\mathfrak{B}^\mathrm{T}$ and $\mathfrak{F}(c)\mathfrak{F}(c)^\mathrm{T}$ the Gramians $\mathfrak{P}_\mathfrak{B}$ and $\mathfrak{P}_\mathfrak{F}(c)$ are well-approximated by some tall and skinny matrices $\mathfrak{Z}_{\mathfrak{B},\mathrm{BT}}$, $\mathfrak{Z}_{\mathfrak{F},\mathrm{BT}}(c)$, so that

$$\mathfrak{P}_\mathfrak{B} \approx \mathfrak{Z}_{\mathfrak{B},\mathrm{BT}}\mathfrak{Z}_{\mathfrak{B},\mathrm{BT}}^\mathrm{T} \qquad \text{and} \qquad \mathfrak{P}_\mathfrak{F}(c) \approx \mathfrak{Z}_{\mathfrak{F},\mathrm{BT}}(c)\mathfrak{Z}_{\mathfrak{F},\mathrm{BT}}(c)^\mathrm{T}. \tag{5.35}$$

These matrices approximate the controllability spaces $\mathcal{V}_\mathfrak{B}$ and $\mathcal{V}_\mathfrak{F}(c)$.

We also define the approximating bases

$$\mathfrak{Z}_{\mathfrak{B},\mathrm{IRKA}} := \begin{bmatrix}\mathbf{\Gamma}^{-1}(s_1)\mathfrak{B}\mathbf{b}_1 & \dots & \mathbf{\Gamma}(s_{N_\mathfrak{B}})^{-1}\mathfrak{B}\mathbf{b}_{N_\mathfrak{B}}\end{bmatrix}, \tag{5.36a}$$

$$\mathfrak{Z}_{\mathfrak{F},\mathrm{IRKA}}(c) := \begin{bmatrix}\mathbf{\Gamma}^{-1}(s_1)\mathfrak{F}(c)\mathbf{f}_1 & \dots & \mathbf{\Gamma}(s_{N_\mathfrak{F}})^{-1}\mathfrak{F}(c)\mathbf{f}_{N_\mathfrak{F}}\end{bmatrix} \tag{5.36b}$$

generated using the IRKA method from Algorithm 4, so that $\mathcal{V}_\mathfrak{B} \approx \operatorname{span}\{\mathfrak{Z}_{\mathfrak{B},\mathrm{IRKA}}\}$ and $\mathcal{V}_\mathfrak{F}(c) \approx \operatorname{span}\{\mathfrak{Z}_{\mathfrak{F},\mathrm{IRKA}}(c)\}$ with $\mathbf{\Gamma}(s)$ and $\mathfrak{F}(c)$ as defined in (5.17).

In the following, we choose $\mathfrak{Z}_\mathfrak{B}$ to be either $\mathfrak{Z}_{\mathfrak{B},\mathrm{BT}}$ or $\mathfrak{Z}_{\mathfrak{B},\mathrm{IRKA}}$ and $\mathfrak{Z}_\mathfrak{F}(c)$ to be either $\mathfrak{Z}_{\mathfrak{F},\mathrm{BT}}(c)$ or $\mathfrak{Z}_{\mathfrak{F},\mathrm{IRKA}}(c)$. Using these matrices, we build a basis

$$\mathbf{V}_\mathrm{r} = \operatorname{orth}(\begin{bmatrix}\mathfrak{Z}_\mathfrak{B} & \mathfrak{Z}_\mathfrak{F}(c_1) & \dots & \mathfrak{Z}_\mathfrak{F}(c_\ell)\end{bmatrix}),$$

which spans an approximation of the solution space $\mathcal{V}$ using the decomposition of $\mathcal{V}_\mathfrak{F}$ presented in (5.20). That way, the components in $\mathbf{V}_\mathrm{r}$ are independent of the damping gains g, and we only consider the different damping positions c. Building a basis $\mathbf{V}_\mathrm{r}$ by using $\mathfrak{Z}_\mathfrak{B}$ and $\mathfrak{Z}_\mathfrak{F}(c)$, we derive a modified RBM that is described in Algorithm 15. Within this method, we make use of the error approximation $\mathbf{\Delta}_{\mathfrak{P}_\mathfrak{F}}(c)$ (5.29) applied to the position test-set $\mathfrak{D}_{\mathrm{Test},c} \subset \mathfrak{D}_c$.

Algorithm 15 Offline phase of the first-order RBM using a decoupled controllability space.

```
Input: $\boldsymbol{\mathcal{E}} \in \mathbb{R}^{N\times N}$, $\boldsymbol{\mathcal{A}} : \boldsymbol{\mathcal{D}} \to \mathbb{R}^{N\times N}$ asymptotically stable, $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{N\times m}$, test-parameter set $\boldsymbol{\mathcal{D}}_{\mathrm{Test},c}$, tolerance tol.
Output: Orthonormal basis $\mathbf{V}_{\mathrm{r}}$.
Compute the basis $\boldsymbol{\mathcal{Z}}_{\boldsymbol{\mathcal{B}}}$ that is equal to $\boldsymbol{\mathcal{Z}}_{\boldsymbol{\mathcal{B}},\mathrm{BT}}$ as in (5.35) or $\boldsymbol{\mathcal{Z}}_{\boldsymbol{\mathcal{B}},\mathrm{IRKA}}$ as in (5.36a).
Set $\mathbf{V}_{\mathrm{r}} := \mathrm{orth}(\boldsymbol{\mathcal{Z}}_{\boldsymbol{\mathcal{B}}})$.
Set $k := 1$.
Determine $c_1 := \mathrm{argmax}_{c\in\boldsymbol{\mathcal{D}}_{\mathrm{Test},c}} \boldsymbol{\Delta}_{\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{F}}}}(c)$.
Set $\mathcal{M} := \{c_1\}$.
Set $\boldsymbol{\Delta}^{\max}_{\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{F}}}} := \boldsymbol{\Delta}_{\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{F}}}}(c_1)$.
while $\boldsymbol{\Delta}^{\max}_{\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{F}}}} > \mathrm{tol}$ do
    Compute the basis $\boldsymbol{\mathcal{Z}}_{\boldsymbol{\mathcal{F}}}(c_k)$ that is equal to $\boldsymbol{\mathcal{Z}}_{\boldsymbol{\mathcal{F}},\mathrm{BT}}(c_k)$ as in (5.35) or $\boldsymbol{\mathcal{Z}}_{\boldsymbol{\mathcal{F}},\mathrm{IRKA}}(c_k)$ as in (5.36b).
    Set $\mathcal{M} := \mathcal{M} \cup \{c_k\}$.
    Set $\mathbf{V}_{\mathrm{r}} := \mathrm{orth}([\mathbf{V}_{\mathrm{r}},\ \boldsymbol{\mathcal{Z}}_{\boldsymbol{\mathcal{F}}}(c_k)])$.
    Determine $c_{k+1} := \mathrm{argmax}_{c\in\boldsymbol{\mathcal{D}}_{\mathrm{Test},c}\setminus\mathcal{M}} \boldsymbol{\Delta}_{\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{F}}}}(c)$.
    Set $\boldsymbol{\Delta}^{\max}_{\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{F}}}} := \boldsymbol{\Delta}_{\boldsymbol{\mathcal{P}}_{\boldsymbol{\mathcal{F}}}}(c_{k+1})$.
    Set $k := k + 1$.
end while
```

5.2 Reduced basis method for second-order systems

We aim to optimize the system response corresponding to the second-order systems (1.3) and (1.4). The computation of both system response expressions includes the calculation of a second-order controllability Gramian $\mathbf{P}_{\mathrm{pos}}(c, g)$ from (2.26), which is the upper-left block of a first-order Gramian

$$\boldsymbol{\mathcal{P}}(c, g) = \begin{bmatrix} \mathbf{P}_{\mathrm{pos}}(c, g) & \mathbf{P}_{12}(c, g) \\ \mathbf{P}_{12}(c, g)^{\mathrm{T}} & \mathbf{P}_{22}(c, g) \end{bmatrix}, \tag{5.37}$$

from (5.2) as shown in Theorem 3.50. To compute a position controllability Gramian $\mathbf{P}_{\mathrm{pos}}(c, g)$ and the respective low-rank factor $\mathbf{Z}_{\mathrm{so}}(c_k, g_k)$, we have to solve a first-order Lyapunov equation (5.3) with matrices as in (1.7) of dimension $N = 2n$. Hence, in every step of the optimization process, we need to compute the respective Gramian by solving a Lyapunov equation of the form (5.3) for the currently considered parameter $(c, g) \in \boldsymbol{\mathcal{D}}$. Solving a Lyapunov equation for all parameters considered within the optimization process leads to high computational costs or is even unfeasible. Hence, we aim to accelerate solving the Lyapunov equations using an RBM. Therefore, we tailor the RBM presented above for first-order systems to be suitable for second-order systems.

Within this RBM, we then aim to find a basis $\mathbf{V}_{\mathrm{so,r}} \in \mathbb{R}^{n\times r}$ that approximates the controllability space of the second-order systems (1.3) and (1.4) for all admissible parameters $(c, g) \in \mathfrak{D}$ so that there exists a reduced matrix $\mathbf{P}_{\mathrm{pos,r}}(c, g) \in \mathbb{R}^{r \times r}$ with

$$\mathbf{P}_{\mathrm{pos}}(c, g) \approx \widetilde{\mathbf{P}}_{\mathrm{pos}}(c, g) := \mathbf{V}_{\mathrm{so,r}} \mathbf{P}_{\mathrm{pos,r}}(c, g) \mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}. \tag{5.38}$$

The authors in [140] derived an RBM for second-order systems where the bases are generated using an IRKA algorithm. We repeat that method and derive a Gramian-based RBM in this section.

This section is structured as follows. In Section 5.2.1, we derive an RBM including an offline phase in which the basis $\mathbf{V}_{\mathrm{so,r}}$ is computed and an online phase that determines an approximation $\widetilde{\mathbf{P}}_{\mathrm{pos}}(c, g)$ of the position controllability Gramian $\mathbf{P}_{\mathrm{pos}}(c, g)$ for all parameters of interest. Afterwards, we derive a controllability space decomposition presented in Section 5.2.2, that is used in Section 5.2.3 to derive a numerically more advantageous second-order RBM.

5.2.1 Offline-online RBM for second-order systems

To simplify the computation of the position controllability Gramians $\mathbf{P}_{\mathrm{pos}}(c, g)$ for various parameters, we derive a basis $\mathbf{V}_{\mathrm{so,r}}$ that approximately spans the controllability space of the second-order systems (1.3) and (1.4). This basis is constructed in the offline phase. Afterwards, in the online phase, we use this basis to compute an approximation of the position controllability Gramian as described in (5.38) for all requested parameters $(c, g) \in \mathfrak{D}$. To describe the quality of the approximation, we assume that there exists an error approximation $\mathbf{\Delta}(c, g)$ that estimates the error between the Gramian $\mathbf{P}_{\mathrm{pos}}(c, g)$ and the respective approximation $\widetilde{\mathbf{P}}_{\mathrm{pos}}(c, g)$, that is specified later in this subsection.

Offline phase To construct a basis $\mathbf{V}_{\mathrm{so,r}}$, we concatenate the controllability space bases $\mathbf{Z}_{\mathrm{so}}(c, g)$ for several parameters, which leads to

$$\mathbf{V}_{\mathrm{so,r}} = \mathrm{orth}\big(\begin{bmatrix} \mathbf{Z}_{\mathrm{so}}(c_1, g_1) & \dots & \mathbf{Z}_{\mathrm{so}}(c_\ell, g_\ell) \end{bmatrix} \big) \in \mathbb{R}^{n \times r}$$

for $(c_k, g_k) \in \mathfrak{D}$, $k = 1, \dots, \ell$. We add controllability space bases $\mathbf{Z}_{\mathrm{so}}(c, g)$ to the basis $\mathbf{V}_{\mathrm{so,r}}$ until the controllability Gramian $\mathbf{P}_{\mathrm{pos}}(c, g)$ is well-approximated by (5.38) for all admissible parameters $(c, g) \in \mathfrak{D}$.

To compute the controllability space bases $\mathbf{Z}_{\mathrm{so}}(c, g)$, we either use the low-rank factors of the respective Gramians or the IRKA method. Because of the structure of the Gramian $\mathfrak{P}(c, g)$, there exists a low-rank factors

$$\mathfrak{Z}(c, g) = \begin{bmatrix} \mathbf{Z}_{\mathrm{so,BT}}(c, g) \\ \mathbf{Z}_2(c, g) \end{bmatrix}, \qquad \text{with} \qquad \mathfrak{P}(c, g) \approx \mathfrak{Z}(c, g)\mathfrak{Z}(c, g)^{\mathrm{T}}$$

that result when applying one of the Lyapunov equation solvers presented in Section 2.3. Hence, the position controllability Gramian $\mathbf{P}_{\text{pos}}(c, g)$ is approximated by

$$\mathbf{P}_{\text{pos}}(c, g) \approx \mathbf{Z}_{\text{so,BT}}(c, g)\mathbf{Z}_{\text{so,BT}}(c, g)^{\text{T}} \tag{5.39}$$

so that $\mathbf{Z}_{\text{so,BT}}(c, g)$ spans an approximation of the controllability space $\boldsymbol{\mathcal{V}}_{\text{so}}(c, g)$. Alternatively, an approximation of the controllability space is spanned by a basis

$$\begin{aligned}&\mathbf{Z}_{\text{so,IRKA}}(c, g)\\&\quad = \begin{bmatrix}(s_1^2\mathbf{M} + s_1\mathbf{D}(c, g) + \mathbf{K})^{-1}\mathbf{B}\mathbf{b}_1 & \dots & (s_{n_\mathbf{V}}^2\mathbf{M} + s_{n_\mathbf{V}}\mathbf{D}(c, g) + \mathbf{K})^{-1}\mathbf{B}\mathbf{b}_{n_\mathbf{V}}\end{bmatrix}\end{aligned} \tag{5.40}$$

for certain interpolation points $s_1, \dots, s_{n_\mathbf{V}}$ and tangential directions $\mathbf{b}_1, \dots, \mathbf{b}_{n_\mathbf{V}}$ generated by the IRKA procedure from Algorithm 6. Hence, controllability space bases $\mathbf{Z}_{\text{so}}(c, g)$ is computed using $\mathbf{Z}_{\text{so}}(c, g) = \mathbf{Z}_{\text{so,BT}}(c, g)$ or $\mathbf{Z}_{\text{so}}(c, g) = \mathbf{Z}_{\text{so,IRKA}}(c, g)$.

Since we can not evaluate all infinite parameters in $\mathfrak{D}$, as in the first-order case, we define a finite and well-distributed test-parameter set $\mathfrak{D}_{\text{Test}} \subset \mathfrak{D}$. We build the basis $\mathbf{V}_{\text{so,r}}$ by picking an arbitrary parameter pair $(c_0, g_0) \in \mathfrak{D}_{\text{Test}}$ and determine the corresponding basis $\mathbf{Z}_{\text{so}}(c_0, g_0)$ equal to $\mathbf{Z}_{\text{so,BT}}(c_0, g_0)$ or $\mathbf{Z}_{\text{so,IRKA}}(c_0, g_0)$, that is used to define the first basis

$$\mathbf{V}_{\text{so,r}} = \text{orth}(\mathbf{Z}_{\text{so}}(c_0, g_0)).$$

For that basis $\mathbf{V}_{\text{so,r}}$, we evaluate the quality of the resulting approximations (5.38). Therefore, we compute the error approximations for all remaining parameters in $\mathfrak{D}_{\text{Test}}$ and determine the largest one as

$$\boldsymbol{\Delta}^{\max} := \boldsymbol{\Delta}(c_1, g_1) := \max_{(c,g)\in\mathfrak{D}} \boldsymbol{\Delta}(c, g).$$

If $\boldsymbol{\Delta}^{\max}$ is larger than a given tolerance tol, the current basis does not approximate the controllability space $\boldsymbol{\mathcal{V}}_{\text{so}}(c_1, g_1)$ good enough, and, hence, the basis $\mathbf{V}_{\text{so,r}}$ needs to be enlarged. We compute a basis $\mathbf{Z}_{\text{so}}(c_1, g_1)$ in (c_1, g_1) that is either $\mathbf{Z}_{\text{so,BT}}(c_1, g_1)$ or $\mathbf{Z}_{\text{so,IRKA}}(c_1, g_1)$, and enrich the basis

$$\mathbf{V}_{\text{so,r}} = \text{orth}\left(\begin{bmatrix}\mathbf{V}_{\text{so,r}} & \mathbf{Z}_{\text{so}}(c_1, g_1)\end{bmatrix}\right).$$

We continue with this method until we have determined a basis $\mathbf{V}_{\text{so,r}}$ that leads to a maximal error approximation $\boldsymbol{\Delta}^{\max} <$ tol.

Online phase In the online phase, we use the basis $\mathbf{V}_{\text{so,r}}$ to compute an approximation $\widetilde{\mathbf{P}}_{\text{pos}}(c, g)$ of the position controllability Gramian $\mathbf{P}_{\text{pos}}(c, g)$ for all required parameters (c, g). For that, we define the first-order basis

$$\mathbf{V}_{\text{r}} = \begin{bmatrix}\mathbf{V}_{\text{so,r}} & 0\\ 0 & \mathbf{V}_{\text{so,r}}\end{bmatrix}, \tag{5.41}$$

that is used to reduce the matrices (1.7). We derive the reduced matrices (5.8), which define the reduced Lyapunov equation (5.9). The reduced position controllability Gramian $\mathbf{P}_{\mathrm{pos,r}}(c,g)$ is then the upper-left block of the reduced first-order Gramian

$$\mathfrak{P}_{\mathrm{r}}(c,g) = \begin{bmatrix} \mathbf{P}_{\mathrm{pos,r}}(c,g) & \mathbf{P}_{12,\mathrm{r}}(c,g) \\ \mathbf{P}_{12,\mathrm{r}}(c,g)^{\mathrm{T}} & \mathbf{P}_{22,\mathrm{r}}(c,g) \end{bmatrix}. \tag{5.42}$$

After solving the reduced Lyapunov equation (5.9) to determine $\mathbf{P}_{\mathrm{pos,r}}(c,g)$, we compute an approximation of the solution $\mathbf{P}_{\mathrm{pos}}(c,g)$ as defined in (5.38).

Error approximation In the second-order RBM presented above, we require an error approximation to evaluate the quality of the resulting approximations. We follow a similar methodology as for the first-order case but modify it in such a way that only the error in the position controllability Gramian is evaluated, i.e., we aim to approximate the error

$$\mathfrak{E}_{\mathrm{so}}(c,g) := \mathbf{P}_{\mathrm{pos}}(c,g) - \widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g) = \mathbf{P}_{\mathrm{pos}}(c,g) - \mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{pos,r}}(c,g)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}. \tag{5.43}$$

The second-order error $\mathfrak{E}_{\mathrm{so}}(c,g)$ is the upper left block of the first-order error

$$\mathfrak{E}(c,g) = \begin{bmatrix} \mathfrak{E}_{\mathrm{so}}(c,g) & \mathfrak{E}_{12}(c,g) \\ \mathfrak{E}_{12}(c,g)^{\mathrm{T}} & \mathfrak{E}_{22}(c,g) \end{bmatrix}$$

that solves the error equation (5.11) with first-order matrices (1.7) and with the corresponding first-order residual (5.10), which is decomposed as

$$\mathfrak{R}(c,g) = \begin{bmatrix} \mathfrak{R}_{11}(c,g) & \mathfrak{R}_{12}(c,g) \\ \mathfrak{R}_{12}(c,g)^{\mathrm{T}} & \mathfrak{R}_{22}(c,g) \end{bmatrix}. \tag{5.44}$$

We apply a second reduced basis method (EE-RBM) to generate a basis $\mathbf{V}_{\mathrm{so,err}}$ that approximately spans the error space, i.e., the space where the errors $\mathfrak{E}_{\mathrm{so}}(c,g)$ for all $(c,g) \in \mathfrak{D}$ live. Suppose we have a basis $\mathbf{V}_{\mathrm{so}}$ such that for each parameter pair $(c,g) \in \mathfrak{D}$ there exists a matrix $\mathfrak{X}_{\mathrm{so}}(c,g)$ with $\mathbf{P}_{\mathrm{pos}}(c,g) = \mathbf{V}_{\mathrm{so}}\mathfrak{X}_{\mathrm{so}}(c,g)\mathbf{V}_{\mathrm{so}}^{\mathrm{T}}$. Then, we can write the error in the position controllability Gramian as

$$\mathfrak{E}_{\mathrm{so}}(c,g) = \mathbf{V}_{\mathrm{so}}\mathfrak{X}_{\mathrm{so}}(c,g)\mathbf{V}_{\mathrm{so}}^{\mathrm{T}} - \mathbf{V}_{\mathrm{so,r}}\mathfrak{P}_{\mathrm{so,r}}(c,g)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}$$

and therefore the error $\mathfrak{E}_{\mathrm{so}}(c,g)$ lives in the space spanned by the basis $\mathbf{V}_{\mathrm{so},\mathfrak{E}} = \mathrm{orth}([\mathbf{V}_{\mathrm{so}}, \mathbf{V}_{\mathrm{so,r}}])$ for all parameters $(c,g) \in \mathfrak{D}$. Since $\mathbf{V}_{\mathrm{so,r}}$ is known from the first RBM, it remains to determine the basis $\mathbf{V}_{\mathrm{so}}$. To compute an approximation of the space spanned by the basis $\mathbf{V}_{\mathrm{so}}$, we apply a second EE-RBM. In each step of the EE-RBM, we add a basis $\mathbf{Z}_{\mathrm{so}}(c^{\mathrm{e}},g^{\mathrm{e}})$ that approximately spans the controllability space corresponding to a parameter $(c^{\mathrm{e}},g^{\mathrm{e}})$, and set

$$\mathbf{V}_{\mathrm{so,err}} = \mathrm{orth}\left(\begin{bmatrix} \mathbf{V}_{\mathrm{so,err}} & \mathbf{V}_{\mathrm{so,r}} & \mathbf{Z}_{\mathrm{so}}(c^{\mathrm{e}},g^{\mathrm{e}}) \end{bmatrix}\right).$$

To compute such a basis $\mathbf{Z}_{\text{so}}(c^{\text{e}}, g^{\text{e}})$, we solve a second Lyapunov equation (5.3) to obtain a low-rank factor $\mathbf{Z}_{\text{so,BT}}(c^{\text{e}}, g^{\text{e}})$ or derive the corresponding controlability space approximation $\mathbf{Z}_{\text{so,IRKA}}(c^{\text{e}}, g^{\text{e}})$ resulting from the IRKA approach as described in (5.40).

After determining a basis $\mathbf{V}_{\text{so,err}}$, we approximate the error for a parameter pair $(c, g) \in \mathfrak{D}$ as

$$\boldsymbol{\mathfrak{E}}_{\text{so}}(c, g) \approx \widetilde{\boldsymbol{\mathfrak{E}}}_{\text{so}}(c, g) = \mathbf{V}_{\text{so,err}} \widehat{\boldsymbol{\mathfrak{E}}}_{\text{so}}(c, g) \mathbf{V}_{\text{so,err}}^{\text{T}} \tag{5.45}$$

where $\widehat{\boldsymbol{\mathfrak{E}}}_{\text{so}}(c, g)$ is the upper-left block of the first-order error

$$\widehat{\boldsymbol{\mathfrak{E}}}(c, g) = \begin{bmatrix} \widehat{\boldsymbol{\mathfrak{E}}}_{\text{so}}(c, g) & \widehat{\boldsymbol{\mathfrak{E}}}_{12}(c, g) \\ \widehat{\boldsymbol{\mathfrak{E}}}_{12}(c, g)^{\text{T}} & \widehat{\boldsymbol{\mathfrak{E}}}_{22}(c, g) \end{bmatrix}$$

that solves the reduced error equation (5.13) for the first-order basis

$$\mathbf{V}_{\text{err}} := \begin{bmatrix} \mathbf{V}_{\text{so,err}} & 0 \\ 0 & \mathbf{V}_{\text{so,err}} \end{bmatrix}.$$

This error approximation is fast computable if the dimension of the basis $\mathbf{V}_{\text{so,err}}$ is sufficiently small. After we have determined a basis $\mathbf{V}_{\text{so,err}}$, we define the error approximation

$$\boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}_{\text{so}}}(c, g) := \|\widetilde{\boldsymbol{\mathfrak{E}}}_{\text{so}}(c, g)\|_{\text{F}} = \|\mathbf{V}_{\text{so,err}} \widehat{\boldsymbol{\mathfrak{E}}}_{\text{so}}(c, g) \mathbf{V}_{\text{so,err}}^{\text{T}}\|_{\text{F}}. \tag{5.46}$$

Both reduced basis methods run in parallel where the consecutive parameter corresponding to the error basis $\mathbf{V}_{\text{so,err}}$, i.e., $(c^{\text{e}}, g^{\text{e}})$ is chosen to be that one that results in the largest residual of the error equation in the Frobenius norm, i.e.,

$$(c^*, g^*) := \arg \max_{(c,g) \in \mathfrak{D}} \|\mathbf{R}_{11}^{\text{e}}(c, g)\|_{\text{F}}$$

where $\mathbf{R}_{11}^{\text{e}}(c, g)$ is as defined in (5.15). The two parallel second-order RBMs result in Algorithm 16.

5.2.2 Decoupling of the controllability space of second-order systems

The controllability space of the second-order systems (1.3) and (1.4) is spanned by

$$\begin{aligned} \boldsymbol{\mathcal{V}}_{\text{so}}(c, g) = \operatorname{span}\{ & \left((s_1)^2\mathbf{M} + s_1\mathbf{D}(c, g) + \mathbf{K}\right)^{-1} \mathbf{B}\mathbf{b}_1, \\ & \ldots, \left((s_M)^2\mathbf{M} + s_M\mathbf{D}(c, g) + \mathbf{K}\right)^{-1} \mathbf{B}\mathbf{b}_M \} \end{aligned} \tag{5.47}$$

if the interpolation points $s_1, \ldots, s_M$ are chosen very well (e.g., the poles of the system) for $N = 2n$. We decompose the kernel for $j = 1, \ldots, M$ as

$$s_j^2\mathbf{M} + s_j\mathbf{D}(c, g) + \mathbf{K} = \boldsymbol{\Gamma}_{\text{so}}(s_j) + \mathbf{F}(c)\mathbf{G}(g)\mathbf{F}(c)^{\text{T}} \tag{5.48}$$

Algorithm 16 Offline phase of the second-order RBM.

Input: $\mathbf{M}$, $\mathbf{K} \in \mathbb{R}^{n\times n}$, $\mathbf{D} : \mathfrak{D} \to \mathbb{R}^{n\times n}$ asymptotically stable, $\mathbf{B} \in \mathbb{R}^{n\times m}$, test-parameter set $\mathfrak{D}_{\text{Test}}$, tolerance tol.

Output: Orthonormal bases $\mathbf{V}_{\text{so,r}}$, $\mathbf{V}_{\text{so,err}}$.

```
Choose any $(c_0,\, g_0)$, $(g_0^{\text{e}},\, c_0^{\text{e}}) \in \mathfrak{D}_{\text{Test}}$ with $(c_0,\, g_0) \neq (c_0^{\text{e}}, g_0^{\text{e}})$.
Compute $\mathbf{Z}_{\text{so}}(c_0, g_0)$ that is either $\mathbf{Z}_{\text{so,BT}}(c_0, g_0)$ from (5.39) or $\mathbf{Z}_{\text{so,IRKA}}(c_0, g_0)$ from (5.40).
Set $\mathcal{M} := \{(c_0,\, g_0)\}$.
Set $\mathbf{V}_{\text{so,r}} := \operatorname{orth}(\mathbf{Z}_{\text{so}}(c_0,\, g_0))$.
Compute $\mathbf{Z}_{\text{so}}(c_0^{\text{e}}, g_0^{\text{e}})$ that is either $\mathbf{Z}_{\text{so,BT}}(c_0^{\text{e}}, g_0^{\text{e}})$ from (5.39) or $\mathbf{Z}_{\text{so,IRKA}}(c_0^{\text{e}}, g_0^{\text{e}})$ from (5.40).
Set $\mathbf{V}_{\text{so,err}} := \operatorname{orth}([\mathbf{Z}_{\text{so}}(c_0,\, g_0),\ \mathbf{Z}_{\text{so}}(g_0^{\text{e}},\, c_0^{\text{e}})])$.
Set $k := 1$.
Determine $(c_1, g_1) := \operatorname{argmax}_{(c,g)\in\mathfrak{D}_{\text{Test}}\setminus\mathcal{M}} \boldsymbol{\Delta}_{\text{so},\boldsymbol{\mathfrak{E}}}(c, g)$.
Set $\boldsymbol{\Delta}_{\text{so},\boldsymbol{\mathfrak{E}}}^{\max} := \boldsymbol{\Delta}_{\text{so},\boldsymbol{\mathfrak{E}}}(c_1, g_1)$.
Determine $(c_1^{\text{e}},\, g_1^{\text{e}}) := \operatorname{argmax}_{(c,g)\in\mathfrak{D}_{\text{Test}}\setminus\mathcal{M}} \|\mathbf{R}_{11}^{\text{e}}(c, g)\|_{\text{F}}$.
while $\boldsymbol{\Delta}_{\text{so},\boldsymbol{\mathfrak{E}}}^{\max} >$ tol do
    Compute $\mathbf{Z}_{\text{so}}(c_k, g_k)$ that is either $\mathbf{Z}_{\text{so,BT}}(c_k, g_k)$ from (5.39) or $\mathbf{Z}_{\text{so,IRKA}}(c_k, g_k)$ from (5.40).
    Set $\mathcal{M} := \mathcal{M} \cup \{(c_k, g_k)\}$.
    Set $\mathbf{V}_{\text{so,r}} := \operatorname{orth}([\mathbf{V}_{\text{so,r}},\ \mathbf{Z}_{\text{so}}(c_k, g_k)])$.
    Compute $\mathbf{Z}_{\text{so}}(c_k^{\text{e}},\, g_k^{\text{e}})$ that is either $\mathbf{Z}_{\text{so,BT}}(c_k^{\text{e}},\, g_k^{\text{e}})$ from (5.39) or $\mathbf{Z}_{\text{so,IRKA}}(c_k^{\text{e}},\, g_k^{\text{e}})$ from (5.40).
    Set $\mathbf{V}_{\text{soerr}} := \operatorname{orth}([\mathbf{V}_{\text{so,err}},\ \mathbf{Z}_{\text{so}}(c_k, g_k),\ \mathbf{Z}_{\text{so}}(c_k^{\text{e}},\, g_k^{\text{e}})])$.
    Determine $(c_{k+1}, g_{k+1}) := \operatorname{argmax}_{(c,g)\in\mathfrak{D}_{\text{Test}}\setminus\mathcal{M}} \boldsymbol{\Delta}_{\text{so},\boldsymbol{\mathfrak{E}}}(c, g)$.
    Set $\boldsymbol{\Delta}_{\text{so},\boldsymbol{\mathfrak{E}}}^{\max} := \boldsymbol{\Delta}_{\text{so},\boldsymbol{\mathfrak{E}}}(c_{k+1},\, g_{k+1})$.
    Determine $(c_{k+1}^{\text{e}},\, g_{k+1}^{\text{e}}) := \operatorname{argmax}_{(c,g)\in\mathfrak{D}_{\text{Test}}\setminus\mathcal{M}} \|\mathbf{R}_{11}^{\text{e}}(c, g)\|_{\text{F}}$.
    Set $k := k + 1$.
end while
```

with $\mathbf{\Gamma}_{\mathrm{so}}(s_j) := s_j^2\mathbf{M} + s_j\mathbf{D}_{\mathrm{int}} + \mathbf{K}$ where only the low-rank factors $\mathbf{F}(c)\mathbf{G}(g)\mathbf{F}(c)^{\mathrm{T}}$ are parameter-dependent while $\mathbf{\Gamma}_{\mathrm{so}}(s_j)$ is independent of the parameter values. This decomposition is used to derive parameter-independent and parameter-dependent components of the controllability space.

Lemma 5.7:
The controllability space (5.47) with interpolation points $\widetilde{s}_1, \ldots, \widetilde{s}_{2n}$ and tangential directions $\widetilde{\mathbf{b}}_1, \ldots, \widetilde{\mathbf{b}}_{2n}$ of the second-order systems (1.3) and (1.4) satisfies

$$\mathcal{V}_{\mathrm{so}}(c,g) \subseteq \mathcal{V}_{\mathrm{so},\mathbf{F}} := \mathcal{V}_{\mathrm{so},\mathbf{B}} \cup \mathcal{V}_{\mathrm{so},\mathbf{F}}(c), \tag{5.49}$$

with

$$\begin{aligned} \mathcal{V}_{\mathrm{so},\mathbf{B}} &:= \operatorname{span}\left\{\mathbf{\Gamma}_{\mathrm{so}}(s_1)^{-1}\mathbf{B}\mathbf{b}_1, \ldots \mathbf{\Gamma}_{\mathrm{so}}(s_{2n})^{-1}\mathbf{B}\mathbf{b}_{2n}\right\}, \\ \mathcal{V}_{\mathrm{so},\mathbf{F}}(c) &:= \operatorname{span}\left\{\mathbf{\Gamma}_{\mathrm{so}}(m_1)^{-1}\mathbf{F}(c)\mathbf{f}_1, \ldots \mathbf{\Gamma}_{\mathrm{so}}(m_M)^{-1}\mathbf{F}(c)\mathbf{f}_M\right\} \end{aligned} \tag{5.50}$$

for interpolation points $s_1, \ldots, s_{2n}$, $m_1, \ldots, m_M$ and tangential directions $\mathbf{b}_1, \ldots, \mathbf{b}_M$, $\mathbf{f}_1, \ldots, \mathbf{f}_M$ that are chosen in such a way, that

$$\mathcal{V}_{\mathrm{so},\mathbf{B}} = \{\mathbf{\Gamma}_{\mathrm{so}}(s)^{-1}\mathbf{B} |\, s \in \mathbb{R}\} \qquad \text{and} \qquad \mathcal{V}_{\mathrm{so},\mathbf{F}}(c) = \{\mathbf{\Gamma}_{\mathrm{so}}(m)^{-1}\mathbf{F}(c) |\, m \in \mathbb{R}\}.$$

for $\mathbf{\Gamma}_{\mathrm{so}}(s_j) := s_j^2\mathbf{M} + s_j\mathbf{D}_{\mathrm{int}} + \mathbf{K}$. ◊

Proof. For every entry $\left(\widetilde{s}_j^2\mathbf{M} + \widetilde{s}_j\mathbf{D}(c,g) + \mathbf{K}\right)^{-1}\mathbf{B}\widetilde{\mathbf{b}}_j$, $j = 1, \ldots, M$ we can apply the Sherman-Morrison-Woodbury formula to obtain

$$\begin{aligned} &\left(\widetilde{s}_j^2\mathbf{M} + \widetilde{s}_j\mathbf{D}(c,g) + \mathbf{K}\right)^{-1}\mathbf{B}\widetilde{\mathbf{b}}_j = \left(\mathbf{\Gamma}_{\mathrm{so}}(\widetilde{s}_j) + \mathbf{F}(c)\mathbf{G}(g)\mathbf{F}(c)^{\mathrm{T}}\right)^{-1}\mathbf{B}\widetilde{\mathbf{b}}_1 \\ &\quad = \mathbf{\Gamma}_{\mathrm{so}}(\widetilde{s}_j)^{-1}\mathbf{B}\widetilde{\mathbf{b}}_j - \mathbf{\Gamma}_{\mathrm{so}}(\widetilde{s}_j)^{-1}\mathbf{F}(c)\left(\mathbf{G}^{-1} + \mathbf{F}(c)^{\mathrm{T}}\mathbf{\Gamma}_{\mathrm{so}}(\widetilde{s}_j)^{-1}\mathbf{F}(c)\right)^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{\Gamma}_{\mathrm{so}}(\widetilde{s}_j)^{-1}\mathbf{B}\widetilde{\mathbf{b}}_j. \end{aligned} \tag{5.51}$$

Hence, the j-th entry of $\mathcal{V}_{\mathrm{so}}(c,g)$ satisfies

$$\left(\widetilde{s}_j^2\mathbf{M} + \widetilde{s}_j\mathbf{D}(c,g) + \mathbf{K}\right)^{-1}\mathbf{B}\widetilde{\mathbf{b}}_j \in \operatorname{span}\left\{\mathbf{\Gamma}_{\mathrm{so}}(\widetilde{s}_j)^{-1}\mathbf{B},\ \mathbf{\Gamma}_{\mathrm{so}}(\widetilde{s}_j)^{-1}\mathbf{F}(c)\right\}.$$

From that, it follows that

$$\begin{aligned} \mathcal{V}_{\mathrm{so}}(c,g) &\subseteq \operatorname{span}\left\{\mathbf{\Gamma}_{\mathrm{so}}^{-1}(s_1)\mathbf{B}\mathbf{b}_1, \ldots \mathbf{\Gamma}_{\mathrm{so}}(s_M)^{-1}\mathbf{B}\mathbf{b}_M\right\} \\ &\qquad\qquad\qquad \cup \operatorname{span}\left\{\mathbf{\Gamma}_{\mathrm{so}}^{-1}(m_1)\mathbf{F}(c)\mathbf{f}_1, \ldots \mathbf{\Gamma}_{\mathrm{so}}(m_M)^{-1}\mathbf{F}(c)\mathbf{f}_M\right\} \\ &= \mathcal{V}_{\mathrm{so},\mathbf{B}} \cup \mathcal{V}_{\mathrm{so},\mathbf{F}}(c). \end{aligned}$$

□

The following theorem is a direct result of Lemma.

Theorem 5.8:
Consider the second-order systems (1.3) and (1.4), and the space $\mathcal{V}_{\text{so},\mathbf{F}}$ as defined in (5.49). Then the controllability space $\mathcal{V}_{\text{so}}(c,g)$ from (5.47) fulfills

$$\mathcal{V}_{\text{so}}(c,g) \subseteq \mathcal{V}_{\text{so},\mathbf{F}}$$

for all parameters $(c,g) \in \mathfrak{D}$. ◊

Note, that the space $\mathcal{V}_{\text{so},\mathbf{B}}$ is the controllability space of the externally undamped system

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}_{\text{int}}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t). \tag{5.52}$$

The corresponding controllability Gramian, that is called $\mathbf{P}_{\text{so},\mathbf{B}}$, can be well-approximated by a low-rank factor $\mathbf{Z}_{\text{so},\mathbf{B},\text{BT}}$, so that

$$\mathbf{P}_{\text{so},\mathbf{B}}(c,g) \approx \mathbf{Z}_{\text{so},\mathbf{B},\text{BT}}\mathbf{Z}_{\text{so},\mathbf{B},\text{BT}}^{\text{T}}. \tag{5.53}$$

The space $\mathcal{V}_{\text{so},\mathbf{F}}(c)$ is the controllability space of the undamped system

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}_{\text{int}}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{F}(c)\mathbf{u}(t) \tag{5.54}$$

with a position-dependent input matrix $\mathbf{F}(c)$. The corresponding controllability Gramian is called $\mathbf{P}_{\text{so},\mathbf{F}}(c)$ with the low-rank factor $\mathbf{Z}_{\text{so},\mathbf{F},\text{BT}}$, so that

$$\mathbf{P}_{\text{so},\mathbf{F}}(c) \approx \mathbf{Z}_{\text{so},\mathbf{F},\text{BT}}(c)\mathbf{Z}_{\text{so},\mathbf{F},\text{BT}}(c)^{\text{T}}. \tag{5.55}$$

The controllability spaces $\mathcal{V}_{\text{so}}(0)$ and $\mathcal{V}_{\text{so}}(c)$ can also be approximated using the IRKA method, which yields

$$\mathbf{Z}_{\text{so},\mathbf{B},\text{IRKA}} := \begin{bmatrix}\mathbf{\Gamma}_{\text{so}}^{-1}(s_1)\mathbf{B}\mathbf{b}_1 & \ldots & \mathbf{\Gamma}_{\text{so}}(s_{\ell_{\mathbf{B}}})^{-1}\mathbf{B}\mathbf{b}_{\ell_{\mathbf{B}}}\end{bmatrix}, \tag{5.56a}$$

$$\mathbf{Z}_{\text{so},\mathbf{F},\text{IRKA}}(c) := \begin{bmatrix}\mathbf{\Gamma}_{\text{so}}^{-1}(m_1)\mathbf{F}(c)\mathbf{f}_1 & \ldots & \mathbf{\Gamma}_{\text{so}}(m_{\ell_{\mathbf{B}}})^{-1}\mathbf{F}(c)\mathbf{f}_{\ell_{\mathbf{F}}}\end{bmatrix} \tag{5.56b}$$

for interpolation points $s_1,\ldots,s_{\ell_{\mathbf{B}}}$, $m_1,\ldots,m_{\ell_{\mathbf{F}}}$ and tangential directions $\mathbf{b}_1,\ldots,\mathbf{b}_{\ell_{\mathbf{B}}}$, $\mathbf{f}_1,\ldots,\mathbf{f}_{\ell_{\mathbf{F}}}$ as described in (2.63).

Hence, the controllability space $\mathcal{V}_{\text{so},\mathbf{F}}$ is approximately spanned by

$$\mathcal{V}_{\text{so},\mathbf{F}} \approx \operatorname{span}\{\mathbf{Z}_{\text{so},\mathbf{B},\text{BT}}\} \bigcup_{c\in\mathfrak{D}} \operatorname{span}\{\mathbf{Z}_{\text{so},\mathbf{F},\text{BT}}(c)\} \quad \text{and}$$
$$\mathcal{V}_{\text{so},\mathbf{F}} \approx \operatorname{span}\{\mathbf{Z}_{\text{so},\mathbf{B},\text{IRKA}}\} \bigcup_{c\in\mathfrak{D}} \operatorname{span}\{\mathbf{Z}_{\text{so},\mathbf{F},\text{IRKA}}(c)\}.$$

Error indicator We aim to derive an error indicator that results from the space decomposition in (5.49). We assume that we have a basis $\mathbf{V}_{\mathrm{so,r}} \in \mathbb{R}^{n\times r}$, $r \ll n$, with $\boldsymbol{\mathcal{V}}_{\mathrm{so},\mathbf{B}} \subset \operatorname{span}\{\mathbf{V}_{\mathrm{so,r}}\}$, so that the controllability Gramian $\mathbf{P}_{\mathrm{so},\mathbf{F}}(c)$ is well-approximated by a matrix $\widetilde{\mathbf{P}}_{\mathrm{so},\mathbf{F}}(c)$ that lies in that space spanned by $\mathbf{V}_{\mathrm{so,r}}$, i.e,

$$\mathbf{P}_{\mathrm{so},\mathbf{F}}(c) \approx \widetilde{\mathbf{P}}_{\mathrm{so},\mathbf{F}}(c) = \mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{so},\mathbf{F},\mathrm{r}}(c)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}.$$

Then, the controllability space lies approximately in

$$\begin{aligned}\boldsymbol{\mathcal{V}}_{\mathrm{so}}(c,g) &\subset \operatorname{span}\{\mathbf{P}_{\mathrm{so},\mathbf{B}}\} \cup \operatorname{span}\{\mathbf{P}_{\mathrm{so},\mathbf{F}}(c)\} \\ &\approx \operatorname{span}\{\mathbf{P}_{\mathrm{so},\mathbf{B}}\} \cup \operatorname{span}\left\{\widetilde{\mathbf{P}}_{\mathrm{so},\mathbf{F}}(c)\right\} = \operatorname{span}\{\mathbf{P}_{\mathrm{so},\mathbf{B}}\} \cup \operatorname{span}\left\{\mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{so},\mathbf{F},\mathrm{r}}(c)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}\right\} \\ &= \operatorname{span}\{\mathbf{P}_{\mathrm{so},\mathbf{B}}\} \cup \operatorname{span}\{\mathbf{V}_{\mathrm{so,r}}\} = \operatorname{span}\{\mathbf{V}_{\mathrm{so,r}}\}.\end{aligned}$$

Hence, to determine the quality of the approximation of the controllability space $\boldsymbol{\mathcal{V}}_{\mathrm{so}}(c,g)$ by the basis $\mathbf{V}_{\mathrm{so,r}}$, an appropriate criterion is to determine how good $\mathbf{P}_{\mathrm{so},\mathbf{F}}(c)$ is approximated by $\widetilde{\mathbf{P}}_{\mathrm{so},\mathbf{F}}(c) = \mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{so},\mathbf{F},\mathrm{r}}(c)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}$.

We consider the system in modal form, i.e., in the transformed representation from (5.1). We consider the corresponding first-order Gramians (5.26) with $\mathbf{P}_{\mathrm{so},\mathbf{F}}(c) := \mathbf{X}_{11}(c)$ and $\widetilde{\mathbf{P}}_{\mathrm{so},\mathbf{F}}(c) = \mathbf{Y}_{11}(c)$. Since we are only interested in the error $\mathbf{P}_{\mathrm{so},\mathbf{F}}(c) - \widetilde{\mathbf{P}}_{\mathrm{so},\mathbf{F}}(c) = \mathbf{X}_{11} - \mathbf{Y}_{11}$, we modify the error expression from (5.29), which yields the following theorem.

Theorem 5.9:
Consider the second-order system (5.54) corresponding to the modal form introduced in (5.1) and the corresponding second-order controllability Gramian $\mathbf{P}_{\mathrm{so},\mathbf{F}}(c)$ from (5.55). Also consider the respective first-order matrices (5.26) and the residual decomposition as defined in (5.28). Then, it holds

$$\begin{aligned}\boldsymbol{\Delta}_{\mathbf{P}_{\mathbf{F}}}(c) &:= \operatorname{tr}\Big(\mathbf{P}_{\mathrm{so},\mathbf{F}}(c) - \widetilde{\mathbf{P}}_{\mathrm{so},\mathbf{F}}(c)\Big) \\ &= \operatorname{tr}\big(\mathbf{R}_{12}(c)\boldsymbol{\Omega}^{-2}\big) + \frac{1}{2\alpha}\operatorname{tr}\big(\boldsymbol{\Omega}^{-3}\mathbf{R}_{22}(c)\big) + \left(\frac{1}{2\alpha} - \alpha\right)\operatorname{tr}\big(\boldsymbol{\Omega}^{-1}\mathbf{R}_{11}(c)\big).\end{aligned} \tag{5.57}$$
◊

Proof. The statement follows from (5.33) and, hence, a byproduct of the proof of Theorem 5.6. □

5.2.3 Offline-online RBM with a decoupled controllability space for second-order systems

In this subsection, we combine the RBM for second-order systems presented in Section 5.2.1 and the controllability space decomposition from Section 5.2.2. We aim to build a basis $\mathbf{V}_{\mathrm{so,r}}$ that spans an approximation of the controllability space $\boldsymbol{\mathcal{V}}_{\mathrm{so}}(c,g)$ for all admissible parameters $(c,g) \in \mathfrak{D}$ but also aim to exploit the structure of the corresponding second-order system described in (5.48). Hence, we again decompose our method into an offline and an online phase. The online phase is similar to the one in Section 5.2.1, so we only describe the offline phase in this section.

We initialize the basis $\mathbf{V}_{\mathrm{so,r}}$ by setting

$$\mathbf{V}_{\mathrm{so,r}} = \mathrm{orth}(\mathbf{Z}_{\mathrm{so},\mathbf{B}}),$$

where $\mathbf{Z}_{\mathrm{so},\mathbf{B}}$ is equal to $\mathbf{Z}_{\mathrm{so},\mathbf{B},\mathrm{BT}}$ or $\mathbf{Z}_{\mathrm{so},\mathbf{B},\mathrm{IRKA}}$ as defined in (5.53) and (5.56a), respectively. The first basis $\mathbf{V}_{\mathrm{so,r}}$ approximates the controllability space of the undamped system realized by the system in (5.52). Using this basis, we evaluate the error approximations for all sample parameters in $\mathfrak{D}_{\mathrm{Test}}$ to determine the largest one

$$\boldsymbol{\Delta}_{\mathbf{P_F}}^{\max} := \boldsymbol{\Delta}_{\mathbf{P_F}}(c_1) := \max_{c \in \mathfrak{D}_{\mathrm{Test},c}} \boldsymbol{\Delta}_{\mathbf{P_F}}(c),$$

where $\mathfrak{D}_{\mathrm{Test},c} \subset \mathfrak{D}_c$ is the subset of $\mathfrak{D}_{\mathrm{Test}}$ that contains the position parameters c. If $\boldsymbol{\Delta}_{\mathbf{P_F}}^{\max}$ is larger than a given tolerance, the basis $\mathbf{V}_{\mathrm{so,r}}$ does not approximate the controllability space in c_1 sufficiently good, and the basis $\mathbf{V}_{\mathrm{so,r}}$ needs to be enriched. For that, we determine the basis $\mathbf{Z}_{\mathrm{so},\mathbf{F}}(c_1)$ that approximates the controllability space of the system (5.54) with $c = c_1$. The basis $\mathbf{Z}_{\mathrm{so},\mathbf{F}}(c_1)$ is either equal to $\mathbf{Z}_{\mathrm{so},\mathbf{F},\mathrm{BT}}(c_1)$ from (5.55) if we use the low-rank factor of the respective controllability Gramians or $\mathbf{Z}_{\mathrm{so},\mathbf{F},\mathrm{IRKA}}(c_1)$ as defined in (5.56a) if we use the IRKA method to derive such a basis. Then, we enrich the basis by setting

$$\mathbf{V}_{\mathrm{so,r}} = \mathrm{orth}\left(\begin{bmatrix}\mathbf{V}_{\mathrm{so,r}} & \mathbf{Z}_{\mathrm{so},\mathbf{F}}(c_1)\end{bmatrix}\right).$$

As before, we continue with this process until the maximal error approximation $\boldsymbol{\Delta}_{\mathbf{P_F}}^{\max}$ is smaller than a certain tolerance, and therefore, the controllability space for all parameters in $\mathfrak{D}_{\mathrm{Test}}$ is approximated sufficiently good by the basis $\mathbf{V}_{\mathrm{so,r}}$. This method results in Algorithm 17.

The algorithms presented in this chapter are applied to damping optimization problems in the next chapter, where we illustrate their efficiency using various numerical examples.

Algorithm 17 Offline phase of the second-order RBM using a decoupled controllability space.

```
Input: $\mathbf{M}$, $\mathbf{K} \in \mathbb{R}^{n \times n}$, $\mathbf{D} : \mathfrak{D} \rightarrow \mathbb{R}^{n \times n}$ asymptotically stable, $\mathbf{B} \in \mathbb{R}^{n \times m}$, test-parameter set $\mathfrak{D}_{\text{Test},c}$, tolerance tol.
Output: Orthonormal basis $\mathbf{V}_{\text{so,r}}$.
Compute the basis $\mathbf{Z}_{\text{so},\mathbf{B}}$ that is equal to $\mathbf{Z}_{\text{so},\mathbf{B},\text{BT}}$ as in (5.53) or $\mathbf{Z}_{\text{so},\mathbf{B},\text{IRKA}}$ as in (5.56a).
Set $\mathbf{V}_{\text{so,r}} := \text{orth}(\mathbf{Z}_{\text{so},\mathbf{B}})$.
Set $k := 1$.
Determine $c_1 := \text{argmax}_{c \in \mathfrak{D}_{\text{Test},c}} \boldsymbol{\Delta}_{\mathbf{P_F}}(c)$.
Set $\mathcal{M} := \{c_1\}$.
Set $\boldsymbol{\Delta}_{\mathbf{P_F}}^{\max} := \boldsymbol{\Delta}_{\mathbf{P_F}}(c_1)$.
while $\boldsymbol{\Delta}_{\mathbf{P_F}}^{\max} > \text{tol}$ do
    Compute the basis $\mathbf{Z}_{\text{so},\mathbf{F}}(c_k)$ that is equal to $\mathbf{Z}_{\text{so},\mathbf{F},\text{BT}}(c_k)$ as in (5.55) or $\mathcal{Z}_{\text{so},\mathbf{F},\text{IRKA}}(c_k)$ as in (5.56b).
    Set $\mathcal{M} := \mathcal{M} \cup \{c_k\}$.
    Set $\mathbf{V}_{\text{so,r}} := \text{orth}([\mathbf{V}_{\text{so,r}},\ \mathbf{Z}_{\text{so},\mathbf{F}}(c_k)])$.
    Determine $c_{k+1} := \text{argmax}_{c \in \mathfrak{D}_{\text{Test},c} \setminus \mathcal{M}} \boldsymbol{\Delta}_{\mathbf{P_F}}(c)$.
    Set $\boldsymbol{\Delta}_{\mathbf{P_F}}^{\max} := \boldsymbol{\Delta}_{\mathbf{P_F}}(c_{k+1})$.
    Set $k := k + 1$.
end while
```

CHAPTER 6

DAMPING OPTIMIZATION

Contents

In this section, we consider the problem of semiactive damping, in which external dampers are added to a vibrational system to minimize the effect of an external force on the system. In more detail, that means that we consider a system of the form (1.3) or (1.4) with a parameter-dependent damping matrix $\mathbf{D}(c,g)$ as described in (1.1) that consists of a parameter-independent internal damping $\mathbf{D}_{\text{int}}$ and an external damping $\mathbf{D}_{\text{ext}}(c,g) = \mathbf{F}(c)\mathbf{G}(g)\mathbf{F}(c)^{\mathrm{T}}$ for $(c,g) \in \mathfrak{D}$. The goal is to optimize damper viscosities g and damper positions c to minimize the effect of external disturbances on the system and the corresponding output. Various criteria quantify the stability of systems and the response to external disturbances, which are selected according to the application. In this work, the average energy amplitude is used, which is equal to the *system responses*

$$\begin{aligned}\mathcal{J}_{\mathrm{L}}(c,g) &:= \|\mathcal{G}_{\mathrm{L}}(\cdot\,;c,g)\|_{\mathcal{H}_2}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{tr}\big(\mathcal{G}_{\mathrm{L}}(\mathrm{i}\omega;c,g)^{\mathrm{H}}\mathcal{G}_{\mathrm{L}}(\mathrm{i}\omega;c,g)\big)\,\mathrm{d}\omega,\\ \mathcal{J}_{\mathrm{Q}}(c,g) &:= \|\mathcal{G}_{\mathrm{Q}}(\cdot,\cdot\,;c,g)\|_{\mathcal{H}_2}^2\\ &= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\operatorname{tr}\big(\mathcal{G}_{\mathrm{Q}}(\mathrm{i}\omega_1,\mathrm{i}\omega_2;c,g)^{\mathrm{H}}\mathcal{G}_{\mathrm{Q}}(\mathrm{i}\omega_1,\mathrm{i}\omega_2;c,g)\big)\,\mathrm{d}\omega_1\mathrm{d}\omega_2\end{aligned}$$

in the linear and quadratic output case, respectively. The transfer functions $\mathcal{G}_{\mathrm{L}}(s;c,g)$ and $\mathcal{G}_{\mathrm{Q}}(s_1,s_2;c,g)$ are as defined in (2.4) and (3.33) as $\mathcal{G}_{\mathrm{Q},\mathcal{B}\mathcal{B}}$, respectively, and describe the input-to-output behavior in the frequency domain. This optimization criterion was also used in [25, 140] for systems with linear output equations. We choose this particular criterion since we aim to minimize the maximal deflections, or more specifically, the maximal time response magnitude $\max_{t\geq 0}\|\mathbf{y}(t)\|_{\infty}$, and hence, we consider the L_{∞}-norm of the output that satisfies the bound

$$\|\mathbf{y}\|_{L_\infty} \leq \|\mathcal{G}(\cdot;c,g)\|_{\mathcal{H}_2}\|\mathbf{u}\|_{L_2}.$$

To simplify the computation of the system response $\mathcal{J}_{\mathrm{L}}(c,g)$ or $\mathcal{J}_{\mathrm{Q}}(c,g)$, we transform the second-order system (1.3) or (1.4) into a first-order system (1.5) or (1.6) with corresponding matrices defined in (1.7). As described in [159], the $\mathcal{H}_2$-norm of the transfer function $\mathcal{G}_{\mathrm{L}}(s;c,g)$ in the linear output case can be computed as

$$\begin{aligned}\mathcal{J}_{\mathrm{L}}(c,g) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{tr}\big(\mathcal{C}(\mathrm{i}\omega\mathcal{E}-\mathcal{A}(c,g))^{-1}\mathcal{B}\mathcal{B}^{\mathrm{H}}(\mathrm{i}\omega\mathcal{E}-\mathcal{A}(c,g))^{-\mathrm{H}}\mathcal{C}^{\mathrm{T}}\big)\,\mathrm{d}\omega && (6.1)\\ &= \operatorname{tr}\big(\mathcal{C}\mathcal{P}(c,g)\mathcal{C}^{\mathrm{T}}\big)\,. && (6.2)\end{aligned}$$

For the quadratic transfer function $\mathcal{G}_{\mathrm{Q}}(s_1,s_2;c,g)$ the system response is equal to

$$\begin{aligned}\mathcal{J}_{\mathrm{Q}}(c,g) =& \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\operatorname{tr}\Big(\mathcal{B}^{\mathrm{H}}(\mathrm{i}\omega_1\mathcal{E}-\mathcal{A}(c,g))^{-\mathrm{H}}\mathcal{M}(\mathrm{i}\omega_2\mathcal{E}-\mathcal{A}(c,g))^{-1}\mathcal{B}\\ &\qquad\cdot\mathcal{B}^{\mathrm{H}}(\mathrm{i}\omega_2\mathcal{E}-\mathcal{A}(c,g))^{-\mathrm{H}}\mathcal{M}(\mathrm{i}\omega_1\mathcal{E}-\mathcal{A}(c,g))^{-1}\mathcal{B}\Big)\mathrm{d}\omega_1\mathrm{d}\omega_2\\ =&\frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{tr}\big(\mathcal{B}^{\mathrm{H}}(\mathrm{i}\omega_1\mathcal{E}-\mathcal{A}(c,g))^{-\mathrm{H}}\mathcal{M}\mathcal{P}(c,g)\mathcal{M}(\mathrm{i}\omega_1\mathcal{E}-\mathcal{A}(c,g))^{-1}\mathcal{B}\big)\,\mathrm{d}\omega_1\\ =&\frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{tr}\big(\mathcal{M}\mathcal{P}(c,g)\mathcal{M}(\mathrm{i}\omega_1\mathcal{E}-\mathcal{A}(c,g))^{-1}\mathcal{B}\mathcal{B}^{\mathrm{H}}(\mathrm{i}\omega_1\mathcal{E}-\mathcal{A}(c,g))^{-\mathrm{H}}\big)\,\mathrm{d}\omega_1\\ =&\operatorname{tr}(\mathcal{M}\mathcal{P}(c,g)\mathcal{M}\mathcal{P}(c,g))\,. \qquad (6.3)\end{aligned}$$

Both system response expressions include the computation of the parameter-dependent controllability Gramian $\mathcal{P}(c,g)$ defined in (5.2). Hence, within an optimization process, we have to solve multiple Lyapunov equations (5.3) in parameters $(c, g) \in \mathfrak{D}$ to compute $\mathfrak{P}(c, g)$. These computations lead to high computational costs if the respective matrices are of large dimensions.

Hence, we utilize the RBM in the following to approximate the controllability Gramian $\widetilde{\mathfrak{P}}(c, g) \approx \mathfrak{P}(c, g)$, as described in (5.4). This approximation is used to approximate the system responses as

$$\mathfrak{J}_{\mathrm{L,r}}(c, g) := \operatorname{tr}\left(\mathfrak{C}\widetilde{\mathfrak{P}}(c, g)\mathfrak{C}^{\mathrm{T}}\right), \qquad \mathfrak{J}_{\mathrm{Q,r}}(c, g) := \operatorname{tr}\left(\mathfrak{M}\widetilde{\mathfrak{P}}(c, g)\mathfrak{M}\widetilde{\mathfrak{P}}(c, g)\right).$$

If the system response values are well-approximated by $\mathfrak{J}_{\mathrm{L,r}}(c, g)$ and $\mathfrak{J}_{\mathrm{Q,r}}(c, g)$, we can optimize these reduced system response expressions instead of the original ones to accelerate the optimization process.

Optimization and the respective methods are not the primary focus of this thesis. To optimize the system responses and their approximations presented in the following sections, we utilize the Nelder-Mead method, a multi-dimensional simplex method. Therefore, we use the `fminsearch` function in MATLAB, that is, whenever we write that we find an optimizer or minimizer, we mean that we apply fminsearch to the function to be minimized. However, analyzing or improving the optimization method itself is beyond the scope of this thesis. Our focus is on accelerating various computational steps and reducing the dimensions of the respective matrices to accelerate the overall optimization process. In particular, the computation of the systems responses $\mathfrak{J}_{\mathrm{L}}(c, g)$ and $\mathfrak{J}_{\mathrm{Q}}(c, g)$ includes the computation of the controllability Gramian $\mathfrak{P}(c, g)$ for several parameters within the optimization. Consequently, a Lyapunov equation needs to be solved for every parameter evaluated within the optimization procedure.

The main task in this section is to accelerate the optimization process by approximating the Gramian $\mathfrak{P}(c, g)$ for all required parameters (c, g). Since both system response expressions, $\mathfrak{J}_{\mathrm{L}}(c, g)$ and $\mathfrak{J}_{\mathrm{Q}}(c, g)$, depend on the Gramians $\mathfrak{P}(c, g)$, the derived methods coincide for both expressions.

In the following, we distinguish between systems in first-order representation, considered in Section 6.1 and those in second-order representation, analyzed in Section 6.2.

6.1 Damping optimization in the first-order representation

In this section, we apply the RBM to accelerate the computation of the controllability Gramian $\mathfrak{P}(c, g)$ and hence of the system responses $\mathfrak{J}_{\mathrm{L}}(c, g)$ and $\mathfrak{J}_{\mathrm{Q}}(c, g)$ for all parameters (c, g) required during an optimization process. For that, first in Section 6.1.1, we apply the offline-online RBM introduced in Section 5.1. Afterwards, in Section 6.1.2, we derive an adaptive scheme that approximates the system responses but does not require a given parameter domain. Also, this approach is combined with the decomposition presented in Section 5.1.2.

6.1.1 Damping optimization using an offline-online RBM for first-order systems

The Gramian $\mathfrak{P}(c,g)$ that is used to compute the system responses solves the Lyapunov equation (5.3) and hence the RBM from Section 5.1 can be used to approximate the controllability Gramian $\mathfrak{P}(c,g)$ and the system response expressions $\mathfrak{J}_{\mathrm{L}}(c,g)$ and $\mathfrak{J}_{\mathrm{Q}}(c,g)$ from (6.1) and (6.3), respectively. For that, we apply Algorithm 14 to generate a basis $\mathbf{V}_{\mathrm{r}} \in \mathbb{R}^{N \times R_{\mathbf{V}}}$ that spans a space that approximates the controllability space $\mathcal{V}$ of the systems (1.3) and (1.4), or Algorithm 15 to generate an approximation of the space $\mathcal{V}_{\mathcal{F}}$ from (5.20).

Within the optimization process, for every requested parameter pair (c,g), we use this basis $\mathbf{V}_{\mathrm{r}}$, to define the reduced Lyapunov equation from (5.9) with matrices from (5.8). Solving this reduced Lyapunov equation yields the reduced Gramian $\mathfrak{P}_{\mathrm{r}}(c,g)$ with $\mathfrak{P}(c,g) \approx \widetilde{\mathfrak{P}}(c,g) = \mathbf{V}_{\mathrm{r}}\mathfrak{P}_{\mathrm{r}}(c,g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}$. The reduced Gramian $\mathfrak{P}_{\mathrm{r}}(c,g)$ is then used to determine the reduced system responses

$$\begin{aligned}\mathfrak{J}_{\mathrm{L,r}}(c,g) &:= \operatorname{tr}\big(\mathfrak{C}\mathbf{V}_{\mathrm{r}}\mathfrak{P}_{\mathrm{r}}(c,g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\mathfrak{C}^{\mathrm{T}}\big)\,,\\ \mathfrak{J}_{\mathrm{Q,r}}(c,g) &:= \operatorname{tr}\big(\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\mathfrak{M}\mathbf{V}_{\mathrm{r}}\mathfrak{P}_{\mathrm{r}}(c,g)\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\mathfrak{M}\mathbf{V}_{\mathrm{r}}\mathfrak{P}_{\mathrm{r}}(c,g)\big)\end{aligned} \tag{6.4}$$

for the linear and quadratic output case, respectively, that approximate the system response values from (6.1) and (6.3). We make use of the trace properties and reorder the matrices so that we can precompute $\mathfrak{C}_{\mathrm{r}} := \mathfrak{C}\mathbf{V}_{\mathrm{r}}$ and $\mathfrak{M}_{\mathrm{r}} := \mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\mathfrak{M}\mathbf{V}_{\mathrm{r}} \in \mathbb{R}^{N_{\mathbf{V}} \times N_{\mathbf{V}}}$ and only matrices of dimension $N_{\mathbf{V}}$ need to be multiplied in the online phase.

Error approximation To describe the quality of the system response approximation from (6.4) by a basis $\mathbf{V}_{\mathrm{r}}$, we can either use the error approximation $\boldsymbol{\Delta}_{\mathfrak{E}}$ from (5.14) to approximate the error in the controllability Gramian, or the error indicator $\boldsymbol{\Delta}_{\mathfrak{P}_{\mathcal{F}}}$ from (5.29) that indicates the quality of the controllability space approximation.

Also, we can tailor the error approximation from (5.14) to evaluate the error in the system response values. For systems (1.5) with a linear output equation, this error is equal to

$$\mathfrak{E}_{\mathfrak{J}_{\mathrm{L}}}(c,g) := \mathfrak{J}_{\mathrm{L}}(c,g) - \mathfrak{J}_{\mathrm{L,r}}(c,g) = \operatorname{tr}\big(\mathfrak{C}\mathfrak{P}(c,g)\mathfrak{C}^{\mathrm{T}}\big) - \operatorname{tr}\Big(\mathfrak{C}\widetilde{\mathfrak{P}}(c,g)\mathfrak{C}^{\mathrm{T}}\Big) = \operatorname{tr}\big(\mathfrak{C}\mathfrak{E}(c,g)\mathfrak{C}^{\mathrm{T}}\big)$$

for the error $\mathfrak{E}(c,g) := \mathfrak{P}(c,g) - \widetilde{\mathfrak{P}}(c,g)$. For systems (1.6) with a quadratic output equation, the error in the system response is

$$\begin{aligned}
\mathfrak{E}_{\mathfrak{J}_{\mathrm{Q}}}(c,g) &:= \mathfrak{J}_{\mathrm{Q}}(c,g) - \mathfrak{J}_{\mathrm{Q,r}}(c,g) \\
&= \operatorname{tr}(\mathfrak{M}\mathfrak{P}(c,g)\mathfrak{M}\mathfrak{P}(c,g)) - \operatorname{tr}\Big(\mathfrak{M}\widetilde{\mathfrak{P}}(c,g)\mathfrak{M}\widetilde{\mathfrak{P}}(c,g)\Big) \\
&= \operatorname{tr}\Big(\mathfrak{M}\left(\mathfrak{P}(c,g) - \widetilde{\mathfrak{P}}(c,g)\right)\mathfrak{M}\left(\mathfrak{P}(c,g) + \widetilde{\mathfrak{P}}(c,g)\right)\Big) \\
&= \operatorname{tr}\Big(\mathfrak{M}\left(\mathfrak{P}(c,g) - \widetilde{\mathfrak{P}}(c,g)\right)\mathfrak{M}\left(\mathfrak{P}(c,g) - \widetilde{\mathfrak{P}}(c,g) + 2\widetilde{\mathfrak{P}}(c,g)\right)\Big) \\
&= \operatorname{tr}\Big(\mathfrak{M}\mathfrak{E}(c,g)\mathfrak{M}\left(\mathfrak{E}(c,g) + 2\widetilde{\mathfrak{P}}(c,g)\right)\Big),
\end{aligned}$$

where we only need the Gramian approximation $\widetilde{\mathfrak{P}}(c,g)$ and the error $\mathfrak{E}(c,g)$, but not the actual Gramian $\mathfrak{P}(c,g)$.

We notice that both error expressions, $\mathfrak{E}_{\mathfrak{J}_{\mathrm{L}}}(c,g)$ and $\mathfrak{E}_{\mathfrak{J}_{\mathrm{Q}}}(c,g)$, include the error $\mathfrak{E}(c,g)$. Hence, we aim to find an approximation $\widetilde{\mathfrak{E}}(c,g) \approx \mathfrak{E}(c,g)$ to determine the approximations of the system response errors

$$\begin{aligned}
\mathfrak{E}_{\mathfrak{J}_{\mathrm{L}}}(c,g) \approx \widetilde{\mathfrak{E}}_{\mathfrak{J}_{\mathrm{L}}}(c,g) &:= \operatorname{tr}\Big(\mathfrak{C}\widetilde{\mathfrak{E}}(c,g)\mathfrak{C}^{\mathrm{T}}\Big), \\
\mathfrak{E}_{\mathfrak{J}_{\mathrm{Q}}}(c,g) \approx \widetilde{\mathfrak{E}}_{\mathfrak{J}_{\mathrm{Q}}}(c,g) &:= \operatorname{tr}\Big(\mathfrak{M}\widetilde{\mathfrak{E}}(c,g)\mathfrak{M}\left(\widetilde{\mathfrak{E}}(c,g) + 2\widetilde{\mathfrak{P}}(c,g)\right)\Big).
\end{aligned}$$

To do so, we follow the same procedure as described in Section 5.1 and make use of the fact that the error $\mathfrak{E}(c,g)$ solves the error equation given in (5.11). Hence, we apply a second EE-RBM to the error equation (5.11) to determine a basis $\mathbf{V}_{\mathrm{err}}$ that spans an approximation of the solution space of the error equation in (5.11). The basis $\mathbf{V}_{\mathrm{err}}$ is then used to derive the approximation $\widetilde{\mathfrak{E}}(c,g) = \mathbf{V}_{\mathrm{err}}\widehat{\mathfrak{E}}(c,g)\mathbf{V}_{\mathrm{err}}^{\mathrm{T}}$ where $\widehat{\mathfrak{E}}(c,g)$ solves the reduced error equation (5.13). Using $\widetilde{\mathfrak{E}}(c,g)$, we derive the error approximations

$$\begin{aligned}
\mathbf{\Delta}_{\mathfrak{J}_{\mathrm{L}}}(c,g) &:= \left|\operatorname{tr}\Big(\mathfrak{C}\widetilde{\mathfrak{E}}_{\mathrm{L}}(c,g)\mathfrak{C}^{\mathrm{T}}\Big)\right|, \\
\mathbf{\Delta}_{\mathfrak{J}_{\mathrm{Q}}}(c,g) &:= \left|\operatorname{tr}\Big(\mathfrak{M}\widetilde{\mathfrak{E}}(c,g)\mathfrak{M}\left(\widetilde{\mathfrak{E}}(c,g) + 2\widetilde{\mathfrak{P}}(c,g)\right)\Big)\right|.
\end{aligned} \tag{6.5}$$

6.1.2 Damping optimization using an adaptive RBM for first-order systems

If we use the RBM to optimize the system response as described in Section 6.1.1, we need to know the parameter set $\mathfrak{D}$ beforehand, which is, in general, not given. Also, the optimization process might only use parameters from a subset of $\mathfrak{D}$ such that the basis $\mathbf{V}_{\mathrm{r}}$ from Section 6.1.1 may contain unused information and is therefore of too large dimension what motivates an adaptive scheme. The idea of the adaptive RBM is to enrich the basis $\mathbf{V}_\mathrm{r}$ within the optimization process. Consequently, there is no decomposition into an offline and an online phase.

We select a parameter $(c_0,\, g_0)$ as the initial value for the optimization process and compute a basis $\mathfrak{Z}_\mathbf{V}(c_0,\, g_0)$ that approximates the respective controllability space using either $\mathfrak{Z}_\mathrm{BT}(c_0,\, g_0)$ from (5.6) or $\mathfrak{Z}_\mathrm{IRKA}(c_0,\, g_0)$ from (5.7). We set the first basis to be $\mathbf{V}_\mathrm{r} = \mathrm{orth}(\mathfrak{Z}_\mathbf{V}(c_0,\, g_0))$ that is used to define the reduced optimization problem (6.4), where the reduced Gramian $\mathfrak{P}_\mathrm{r}(c, g)$ solves the reduced Lyapunov equation in (5.9).

Using the basis $\mathbf{V}_\mathrm{r}$, we start an optimization process to minimize the system response (6.4). In contrast to the previous method, we add a stopping criterion within the optimization process that interrupts the procedure whenever the solution space corresponding to the current parameter (c, g) is not well-approximated by the basis $\mathbf{V}_\mathrm{r}$. To achieve this stopping, we modify the goal function as described by Algorithm 18. In every iteration of the optimization process, we query the error approximation $\boldsymbol{\Delta}(c, g)$ of the current parameter (c, g) as described in Step 2. The used error approximation $\boldsymbol{\Delta}(c, g)$ is derived later in this subsection. If the $\boldsymbol{\Delta}(c, g)$ is smaller than a given tolerance, we proceed with the function evaluation in Step 5 and 6 to compute the function value $\widetilde{\mathfrak{J}}(c, g)$ and continue with the optimization process. On the other hand, if the error approximation is larger than the tolerance, we know that the current basis $\mathbf{V}_\mathrm{r}$ does not approximate the solution space of the Lyapunov equation (5.3) for the current parameter pair $(c,\, g)$ sufficiently well. Hence, we return that the minimization did not converge and enrich the basis $\mathbf{V}_\mathrm{r}$. Therefore, we compute $\mathfrak{Z}_\mathbf{V}(c, g)$ using either $\mathfrak{Z}_\mathrm{BT}(c,\, g)$ from (5.6) or $\mathfrak{Z}_\mathrm{IRKA}(c,\, g)$ from (5.7) and define the updated basis

$$\mathbf{V}_\mathrm{r} = \mathrm{orth}([\mathbf{V}_\mathrm{r}, \mathfrak{Z}_\mathbf{V}(c, g)]).$$

Consequently, we obtain a new optimization problem (6.4) that is defined with the new basis $\mathbf{V}_\mathrm{r}$ together with the computation in (5.4) and the corresponding Lyapunov equation in (5.9). Since the function that is to be optimized depends on the current basis $\mathbf{V}_\mathrm{r}$, which changes during the optimization procedure, convergence problems may occur. Hence, we start a new optimization procedure whenever we enrich the basis and use the current parameter $(c,\, g)$ as the initial value. We continue with this procedure until the optimum is reached.

Error approximation To derive an error approximation used in the adaptive procedure introduced above, we follow the same idea as for the offline-online scheme and run a second reduced basis method to generate a basis $\mathbf{V}_\mathrm{err}$ that spans an approximation of the error space. The equations in (5.14) and (6.5) define then two possible error approximations $\boldsymbol{\Delta}(c, g)$ corresponding to the bases $\mathbf{V}_\mathrm{r}$ and $\mathbf{V}_\mathrm{err}$. In this adaptive procedure, the basis $\mathbf{V}_\mathrm{err}$ is enlarged whenever $\mathbf{V}_\mathrm{r}$ is expanded.

The detailed procedure is described in Algorithm 19. We first determine a basis $\mathbf{V}_\mathrm{r} = \mathrm{orth}(\mathfrak{Z}_\mathbf{V}(c_0, g_0))$ that spans an approximation of the controllability space, where

Algorithm 18 Reduced first-order system response.

Input: $\boldsymbol{\mathfrak{E}} \in \mathbb{R}^{N\times N}$, $\boldsymbol{\mathfrak{A}} : \boldsymbol{\mathfrak{D}} \to \mathbb{R}^{N\times N}$, $\boldsymbol{\mathfrak{B}} \in \mathbb{R}^{N\times m}$, $\boldsymbol{\mathfrak{C}} \in \mathbb{R}^{p\times N}$ or $\boldsymbol{\mathfrak{M}} \in \mathbb{R}^{N\times N}$, parameters $(c, g) \in \boldsymbol{\mathfrak{D}}$, basis $\mathbf{V}_\mathrm{r}$, tolerance tol.

Output: System response $\boldsymbol{\mathfrak{J}}_\mathrm{r}(c, g)$ equal to $\boldsymbol{\mathfrak{J}}_\mathrm{Lr}(c, g)$ or $\boldsymbol{\mathfrak{J}}_{\mathrm{Q,r}}(c, g)$ from (6.4), variable conv that shows whether the algorithm converged.

```
Set $\mathsf{conv} = \mathsf{true}$.
if $\boldsymbol{\Delta}(c, g) > \mathrm{tol}$ then
    Set $\mathsf{conv} = \mathsf{false}$, $\boldsymbol{\mathfrak{J}}_\mathrm{r}(c, g) = \infty$.
else
    Solve the reduced Lyapunov equation (5.9) to obtain $\boldsymbol{\mathfrak{P}}_\mathrm{r}(c, g)$.
    Compute $\boldsymbol{\mathfrak{J}}_{\mathrm{L,r}}(c, g)$ or $\boldsymbol{\mathfrak{J}}_{\mathrm{Q,r}}(c, g)$ from (6.4).
end if
```

$\boldsymbol{\mathfrak{Z}}_\mathbf{V}(c_0, g_0)$ is either $\boldsymbol{\mathfrak{Z}}_\mathrm{BT}(c_0, g_0)$ from (5.6) or $\boldsymbol{\mathfrak{Z}}_\mathrm{IRKA}(c_0, g_0)$ from (5.7). Then, we choose a parameter $(c_0^\mathrm{r}, g_0^\mathrm{r})$ that is used to determine an error space approximation. To limit the possibilities of choosing $(c_0^\mathrm{r}, g_0^\mathrm{r})$, we again define a finite set $\boldsymbol{\mathfrak{D}}_\mathrm{Test}$ that can be either a subset of $\boldsymbol{\mathfrak{D}}$, if given, or some set that contains arbitrarily chosen parameters with some distance to the currently considered parameter. We pick the parameter $(c_0^\mathrm{r}, g_0^\mathrm{r})$ from this finite set $\boldsymbol{\mathfrak{D}}_\mathrm{Test}$ and determine $\boldsymbol{\mathfrak{Z}}_\mathbf{V}(c_0^\mathrm{r}, g_0^\mathrm{r})$ that approximates the respective controllability space using either $\boldsymbol{\mathfrak{Z}}_\mathrm{BT}(c_0^\mathrm{r}, g_0^\mathrm{r})$ from (5.6) or $\boldsymbol{\mathfrak{Z}}_\mathrm{IRKA}(c_0^\mathrm{r}, g_0^\mathrm{r})$ from (5.7). The bases $\boldsymbol{\mathfrak{Z}}_\mathbf{V}(c_0^\mathrm{r}, g_0^\mathrm{r})$ and $\mathbf{V}_\mathrm{r}$ are then used to define the first error equation basis

$$\mathbf{V}_\mathrm{err} = \mathrm{orth}([\mathbf{V}_\mathrm{r},\ \boldsymbol{\mathfrak{Z}}_\mathbf{V}(c_0^\mathrm{r}, g_0^\mathrm{r})]) = \mathrm{orth}([\boldsymbol{\mathfrak{Z}}_\mathbf{V}(c_0, g_0),\ \boldsymbol{\mathfrak{Z}}_\mathbf{V}(c_0^\mathrm{r}, g_0^\mathrm{r})]).$$

Using $\mathbf{V}_\mathrm{r}$ and $\mathbf{V}_\mathrm{err}$, we define the error approximation $\boldsymbol{\Delta}(c, g)$ that is equal to either $\boldsymbol{\Delta}_{\boldsymbol{\mathfrak{E}}}(c, g)$ from (5.14), or $\boldsymbol{\Delta}_{\boldsymbol{\mathfrak{J}}_\mathrm{L}}(c, g)$ or $\boldsymbol{\Delta}_{\boldsymbol{\mathfrak{J}}_\mathrm{Q}}(c, g)$ from (6.5). The computation of the two bases is described in Step 1 to 5 of Algorithm 19. After computing the first bases $\mathbf{V}_\mathrm{r}$ and $\mathbf{V}_\mathrm{err}$, we define the system response approximation $\boldsymbol{\mathfrak{J}}_{\mathrm{L,r}}(c, g)$ or $\boldsymbol{\mathfrak{J}}_{\mathrm{Q,r}}(c, g)$ from (6.4), which is then optimized instead of the original system response. We apply an optimization method to compute the minimal system response and the respective minimizer (c^*, g^*).

To ensure the optimization process is interrupted if the basis $\mathbf{V}_\mathrm{r}$ is insufficient, we optimize the function defined in Algorithm 18 instead of the actual system response approximation. This approach allows the optimization to yield either the minimizer (c^*, g^*) or, if $\mathsf{conv} = \mathsf{false}$ holds, the information that the optimization process did not converge, indicating the need to enrich the bases. If the bases need to be expanded, in Step 8 to 12, we enlarge the bases $\mathbf{V}_\mathrm{r}$ and $\mathbf{V}_\mathrm{err}$ as

$$\mathbf{V}_\mathrm{r} = \mathrm{orth}([\mathbf{V}_\mathrm{r},\ \boldsymbol{\mathfrak{Z}}_\mathbf{V}(c, g)]), \text{ and } \mathbf{V}_\mathrm{err} = \mathrm{orth}([\mathbf{V}_\mathrm{err},\ \boldsymbol{\mathfrak{Z}}_\mathbf{V}(c, g),\ \boldsymbol{\mathfrak{Z}}_\mathbf{V}(c^\mathrm{r}, g^\mathrm{r})])$$

where we choose in Step 12 the parameter $(c^\mathrm{r}, g^\mathrm{r}) \in \boldsymbol{\mathfrak{D}}_\mathrm{Test}$ that results in the largest residual

$$(c^{\mathrm{r}}, g^{\mathrm{r}}) = \operatorname{argmax}_{(c,g)\in\mathfrak{D}_{\mathrm{Test}}} \|\mathfrak{R}_{\mathrm{r}}(c, g)\|_{\mathrm{F}}$$

with $\mathfrak{R}_{\mathrm{r}}(c, g)$ defined as in (5.15). Afterwards, in Step 13, we compute the approximated energy response value and proceed with the minimization process.

Algorithm 19 Adaptive first-order RBM.

```
Input: $\mathfrak{E} \in \mathbb{R}^{N\times N}$, $\mathfrak{A} : \mathfrak{D} \to \mathbb{R}^{N\times N}$, $\mathfrak{B} \in \mathbb{R}^{N\times m}$, $\mathfrak{C} \in \mathbb{R}^{p\times N}$ or $\mathfrak{M} \in \mathbb{R}^{N\times N}$, tolerance tol.
Output: Minimizer $(c^{\mathrm{opt}}, g^{\mathrm{opt}})$, energy response $\mathfrak{J}_{\mathrm{L,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ or $\mathfrak{J}_{\mathrm{Q,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$.
 1: Choose $(c_0, g_0)$, $(c_0^{\mathrm{r}}, g_0^{\mathrm{r}}) \in \mathfrak{D}_{\mathrm{Test}}$, $(c_0, g_0) \neq (c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$.
 2: Determine a basis $\mathfrak{Z}_{\mathbf{V}}(c_0, g_0)$ either as $\mathfrak{Z}_{\mathrm{BT}}(c_0, g_0)$ from (5.6) or $\mathfrak{Z}_{\mathrm{IRKA}}(c_0, g_0)$ from (5.7).
 3: Set $\mathbf{V}_{\mathrm{r}} := \operatorname{orth}(\mathfrak{Z}_{\mathbf{V}}(c_0, g_0))$.
 4: Determine a basis $\mathfrak{Z}_{\mathbf{V}}(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$ either as $\mathfrak{Z}_{\mathrm{BT}}(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$ from (5.6) or $\mathfrak{Z}_{\mathrm{IRKA}}(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$ from (5.7).
 5: Set $\mathbf{V}_{\mathrm{err}} = \operatorname{orth}([\mathfrak{Z}_{\mathbf{V}}(c_0, g_0),\ \mathfrak{Z}_{\mathbf{V}}(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})])$.
 6: Find the minimizer $(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ of the function Algorithm 18 using fminsearch and obtain $\mathfrak{J}_{\mathrm{L,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ or $\mathfrak{J}_{\mathrm{Q,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$, and conv.
 7: while conv = false do
 8:     Determine a basis $\mathfrak{Z}_{\mathbf{V}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ either as $\mathfrak{Z}_{\mathrm{BT}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ from (5.6) or $\mathfrak{Z}_{\mathrm{IRKA}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ from (5.7).
 9:     Set $\mathbf{V}_{\mathrm{r}} := \operatorname{orth}([\mathbf{V}_{\mathrm{r}}, \mathfrak{Z}_{\mathbf{V}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})])$.
10:     Determine $(c^{\mathrm{r}}, g^{\mathrm{r}}) := \operatorname{argmax}_{(c,g)\in\mathfrak{D}_{\mathrm{Test}}} \|\mathfrak{R}_{\mathrm{r}}(c, g)\|_{\mathrm{F}}$.
11:     Determine a basis $\mathfrak{Z}_{\mathbf{V}}(c^{\mathrm{r}}, g^{\mathrm{r}})$ either as $\mathfrak{Z}_{\mathrm{BT}}(c^{\mathrm{r}}, g^{\mathrm{r}})$ from (5.6) or $\mathfrak{Z}_{\mathrm{IRKA}}(c^{\mathrm{r}}, g^{\mathrm{r}})$ from (5.7).
12:     Set $\mathbf{V}_{\mathrm{err}} = \operatorname{orth}([\mathbf{V}_{\mathrm{err}},\ \mathfrak{Z}_{\mathbf{V}}(c^{\mathrm{opt}}, g^{\mathrm{opt}}),\ \mathfrak{Z}_{\mathbf{V}}(c^{\mathrm{r}}, g^{\mathrm{r}})])$.
13:     Find the minimizer $(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ of the function Algorithm 18 using fminsearch and obtain $\mathfrak{J}_{\mathrm{L,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ or $\mathfrak{J}_{\mathrm{Q,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$, and conv.
14: end while
```

Remark 6.1:
As described in Remark 5.2, we solve the Lyapunov equation (5.3) in $g = 0 \in \mathbb{R}^{\ell}$ (undamped system) to obtain $\mathfrak{Z}_{\mathbf{V}}(c, 0)$. The vectors of $\mathfrak{Z}_{\mathbf{V}}(c, 0)$ are then added to the basis $\mathbf{V}_{\mathrm{r}}$, which turns out to lead to a more robust basis. ♢

6.1.3 Damping optimization using an adaptive RBM with a decoupled controllability space for first-order systems

To derive an adaptive RBM that uses the structure of the considered vibrational system, we combine the adaptive method presented in Algorithm 19 and the decoupling presented in Section 5.1.2. Again, we aim to find a basis $\mathbf{V}_{\mathrm{r}}$ that approximates the solution space for all parameters of interest. Therefore, we again derive a basis $\mathbf{V}_{\mathrm{r}}$ that approximates the controllability space of the systems (1.5) and (1.6) for all parameters (c, g). Hence, we again utilize the function in Algorithm 18 using an error indicator $\boldsymbol{\Delta}(c, g)$ equal to $\boldsymbol{\Delta}_{\mathcal{P}_{\mathcal{F}}}(c)$ from (5.29). We initialize a basis by setting $\mathbf{V}_{\mathrm{r}} = \mathrm{orth}(\mathcal{Z}_{\mathcal{B}})$ with the basis $\mathcal{Z}_{\mathcal{B}}$ equal to $\mathcal{Z}_{\mathcal{B},\mathrm{IRKA}}$ from (5.36a) or $\mathcal{Z}_{\mathcal{B},\mathrm{BT}}$ from (5.35). Using this basis, we start the optimization of the reduced energy response function given in Algorithm 18 until the method either converges or returns `conf` $= \mathit{false}$. If `conf` $= \mathit{false}$ holds, the optimization process stopped since the current basis is not sufficiently good. In that case, we enrich the basis

$$\mathbf{V}_{\mathrm{r}} = \mathrm{orth}(\begin{bmatrix}\mathbf{V}_{\mathrm{r}} & \mathcal{Z}_{\mathcal{F}}(c)\end{bmatrix}),$$

where $\mathcal{Z}_{\mathcal{F}}(c)$ is equal to $\mathcal{Z}_{\mathcal{F},\mathrm{BT}}(c)$ from (5.35) or $\mathcal{Z}_{\mathcal{F},\mathrm{IRKA}}(c)$ from (5.36b), and c is the current position parameter in which the optimization method stopped. Using the new basis, we start a new optimization process and continue with this method until it converges. This results in Algorithm 20.

Algorithm 20 Adaptive first-order RBM using a decoupled controllability space.

```
Input: $\mathcal{E} \in \mathbb{R}^{N \times N}$, $\mathcal{A} : \mathcal{D} \to \mathbb{R}^{N \times N}$, $\mathcal{B} \in \mathbb{R}^{N \times m}$, $\mathcal{C} \in \mathbb{R}^{p \times N}$ or $\mathcal{M} \in \mathbb{R}^{N \times N}$, initial parameters $(c, g)$, tolerance tol.
Output: Minimizer $(c^{\mathrm{opt}}, g^{\mathrm{opt}})$, energy response $\mathcal{J}_{\mathrm{L,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ or $\mathcal{J}_{\mathrm{Q,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$.
Compute the basis $\mathcal{Z}_{\mathcal{B}}$ equal to $\mathcal{Z}_{\mathcal{B},\mathrm{IRKA}}$ from (5.36a) or $\mathcal{Z}_{\mathcal{B},\mathrm{BT}}$ from (5.35).
Set $\mathbf{V}_{\mathrm{r}} := \mathrm{orth}(\mathcal{Z}_{\mathcal{B}})$.
Find the minimizer $(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ of the function Algorithm 18 with error indicator $\boldsymbol{\Delta}_{\mathcal{P}_{\mathcal{F}}}$ from (5.29) using fminsearch and obtain $\mathcal{J}_{\mathrm{L,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ or $\mathcal{J}_{\mathrm{Q,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$, and conv.
while conv = false do
    Compute the basis $\mathcal{Z}_{\mathcal{F}}(c^{\mathrm{opt}})$ equal to $\mathcal{Z}_{\mathcal{F},\mathrm{IRKA}}(c^{\mathrm{opt}})$ from (5.36b) or $\mathcal{Z}_{\mathcal{F},\mathrm{BT}}(c^{\mathrm{opt}})$ from (5.35).
    Set $\mathbf{V}_{\mathrm{r}} := \mathrm{orth}([\mathbf{V}_{\mathrm{r}}, \mathcal{Z}_{\mathcal{F}}(c^{\mathrm{opt}})])$.
    Find the minimizer $(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ of the function Algorithm 18 with error indicator $\boldsymbol{\Delta}_{\mathcal{P}_{\mathcal{F}}}$ from (5.29) using fminsearch and obtain $\mathcal{J}_{\mathrm{L,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$ or $\mathcal{J}_{\mathrm{Q,r}}(c^{\mathrm{opt}}, g^{\mathrm{opt}})$, and conv.
end while
```

6.2 Damping optimization in the second-order representation

We consider second-order systems of the form (1.3) and (1.4). When evaluating the vibrational systems of these structures, often only the displacements are considered to derive an output. Therefore, the first-order output matrices are

$$\mathcal{C} = \begin{bmatrix}\mathbf{C}_1 & 0\end{bmatrix} \qquad \text{and} \qquad \mathcal{M} = \begin{bmatrix}\mathbf{M}_{11} & 0\\ 0 & 0\end{bmatrix}.$$

If also the velocities are evaluated, we apply the methods from Section 6.1.

We reformulate the system response from (6.1) and (6.3) to take advantage of the structure of the underlying second-order system. The system response corresponding to the system (1.3) with a linear output equation is then equal to

$$\begin{aligned}\mathcal{J}_{\mathrm{L}}(c,g) &:= \|\mathcal{G}_{\mathrm{L}}(\cdot;c,g)\|_{\mathcal{H}_2}^2\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{tr}\Big(\mathbf{C}_1\mathbf{\Lambda}(\mathrm{i}\omega;c,g)\mathbf{B}\mathbf{B}^{\mathrm{T}}\mathbf{\Lambda}(\mathrm{i}\omega;c,g)^{\mathrm{H}}\mathbf{C}_1^{\mathrm{T}}\Big)\mathrm{d}\omega\\ &= \operatorname{tr}\big(\mathbf{C}_1\mathbf{P}_{\mathrm{pos}}(c,g)\mathbf{C}_1^{\mathrm{T}}\big)\,.\end{aligned} \tag{6.6}$$

for $\mathbf{\Lambda}(s;c,g) := \left(s^2\mathbf{M} + s\mathbf{D}(c,g) + \mathbf{K}\right)^{-1}$. We see that the system response depends on the second-order controllability Gramian $\mathbf{P}_{\mathrm{pos}}(c,g)$ defined in (2.26). Hence, we can utilize the second-order structure of the underlying system to compute and approximate the system response values.

We also rewrite the system response for a system (1.4) with a quadratic output equation as

$$\begin{aligned}\mathcal{J}_{\mathrm{Q}}(c,g) :=& \|\mathcal{G}_{\mathrm{Q}}(\cdot;c,g)\|_{\mathcal{H}_2}^2\\ =& \frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{tr}\Big(\mathbf{B}^{\mathrm{T}}\mathbf{\Lambda}(\mathrm{i}\omega;c,g)^{\mathrm{H}}\mathbf{M}_{11}\mathbf{P}_{\mathrm{pos}}(c,g)\mathbf{M}_{11}\mathbf{\Lambda}(\mathrm{i}\omega;c,g)\mathbf{B}\Big)\mathrm{d}\omega_1\\ =& \operatorname{tr}(\mathbf{P}_{\mathrm{pos}}(c,g)\mathbf{M}_{11}\mathbf{P}_{\mathrm{pos}}(c,g)\mathbf{M}_{11})\,.\end{aligned} \tag{6.7}$$

Again, the system response representation depends on the second-order controllability Gramian $\mathbf{P}_{\mathrm{pos}}(c,g)$.

In both cases, the system response representation includes the Gramian $\mathbf{P}_{\mathrm{pos}}(c,g)$, which is computed by solving a Lyapunov equation. Hence, in every step of the optimization process such a matrix equation needs to be solved, which leads to high computational costs. Therefore, we attempt to approximate the system response values, including the second-order controllability Gramian, using the RBM presented in Section 5.2.

Within the RBM we build a basis $\mathbf{V}_{\mathrm{so,r}} \in \mathbb{R}^{n\times r}$ that approximates the respective controllability space, i.e.,

$$\mathbf{P}_{\mathrm{pos}}(c,g) \approx \widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g) = \mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{pos,r}}(c,g)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}} \tag{6.8}$$

holds for suitable matrices $\mathbf{P}_{\mathrm{pos,r}}(c,g) \in \mathbb{R}^{r\times r}$ of small dimension $r \ll n$ that are computed as shown in (5.42). This Gramian approximation is then used to approximate the system response as

$$
\begin{aligned}
\mathfrak{J}_{\mathrm{L}}(c,g) &= \operatorname{tr}\big(\mathbf{C}_1\mathbf{P}_{\mathrm{pos}}(c,g)\mathbf{C}_1^{\mathrm{T}}\big) \approx \operatorname{tr}\big(\mathbf{C}_1\mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{pos,r}}(c,g)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}\mathbf{C}_1^{\mathrm{T}}\big)\,, \\
\mathfrak{J}_{\mathrm{Q}}(c,g) &= \operatorname{tr}(\mathbf{M}_{11}\mathbf{P}_{\mathrm{pos}}(c,g)\mathbf{M}_{11}\mathbf{P}_{\mathrm{pos}}(c,g)) \\
&\qquad\qquad \approx \operatorname{tr}\big(\mathbf{M}_{11}\mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{pos,r}}(c,g)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}\mathbf{M}_{11}\mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{pos,r}}(c,g)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}\big)\,.
\end{aligned}
\tag{6.9}
$$

In the following, we derive different RBMs to compute the system response approximations from (6.9). First, in Section 6.2.1, we apply the RBM presented in Section 5.2 and derive suitable error approximations. Afterwards, in Section 6.2.2, we introduce an adaptive scheme to build the reduction basis $\mathbf{V}_{\mathrm{so,r}}$ and finally, in Section 6.2.3, we extent this adaptive scheme using the decomposed controllability space representation from (5.49).

6.2.1 Damping optimization for second-order systems using an offline-online RBM for second-order systems

In this subsection, we apply the RBM, introduced in Section 5.2.1 and Section 5.2.3, to approximate the system responses $\mathfrak{J}_{\mathrm{L}}(c,g)$ and $\mathfrak{J}_{\mathrm{Q}}(c,g)$. Therefore, we use one of the offline-online schemes derived in Section 5.2, where we generate a basis $\mathbf{V}_{\mathrm{so,r}}$ so that (6.8) holds. The basis $\mathbf{V}_{\mathrm{so,r}}$ is generated in the offline phase using Algorithm 16 or Algorithm 17. After building such a basis $\mathbf{V}_{\mathrm{so,r}}$, we define the reduced system responses

$$
\begin{aligned}
\mathfrak{J}_{\mathrm{L,r}}(c,g) &:= \operatorname{tr}(\mathbf{C}_1\mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{pos,r}}(c,g)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}\mathbf{C}_1^{\mathrm{T}}), \\
\mathfrak{J}_{\mathrm{Q,r}}(c,g) &:= \operatorname{tr}(\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}\mathbf{M}_{11}\mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{pos,r}}(c,g)\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}\mathbf{M}_{11}\mathbf{V}_{\mathrm{so,r}}\mathbf{P}_{\mathrm{pos,r}}(c,g)),
\end{aligned}
\tag{6.10}
$$

where $\mathbf{P}_{\mathrm{pos,r}}(c,g)$ is the upper-left block of $\mathfrak{P}_{\mathrm{r}}(c,g)$ from (5.42) which solves the reduced Lyapunov equation (5.9). After determining the basis $\mathbf{V}_{\mathrm{so,r}}$, we compute the reduced matrices $\mathbf{C}_1\mathbf{V}_{\mathrm{so,r}}$ and $\mathbf{V}_{\mathrm{so,r}}^{\mathrm{T}}\mathbf{M}_{11}\mathbf{V}_{\mathrm{so,r}}$ as these matrices do not change in the online phase and are used multiple times.

Afterwards, in the online phase, we apply an optimization method to minimize the reduced system responses from (6.10). Within the optimization process, we solve a $2r$ dimensional Lyapunov equation (5.42) for every considered parameter instead of solving a $2n$ dimensional Lyapunov equation, which accelerates the optimization process.

Error approximation When we apply the offline-online RBM introduced in Section 5.2.1, for the linear output case, we can tailor the error approximation $\boldsymbol{\Delta}_{\mathfrak{E}_{\mathrm{so}}}(c,g)$ to evaluate the error in the system response, i.e.,

$$
\begin{aligned}
\mathfrak{E}_{\mathfrak{J}_{\mathrm{L}}}(c,g) &:= |\mathfrak{J}_{\mathrm{L}}(c,g) - \mathfrak{J}_{\mathrm{L,r}}(c,g)| \\
&\qquad = \left|\operatorname{tr}\big(\mathbf{C}_1\mathbf{P}_{\mathrm{pos}}(c,g)\mathbf{C}_1^{\mathrm{T}}\big) - \operatorname{tr}\Big(\mathbf{C}_1\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\mathbf{C}_1^{\mathrm{T}}\Big)\right| = \left|\operatorname{tr}(\mathbf{C}_1\mathfrak{E}_{\mathrm{so}}(c,g)\mathbf{C}_1)^{\mathrm{T}}\right|
\end{aligned}
$$

for $\mathfrak{E}_{\mathrm{so}}(c,g)$ as defined in (5.43). Also, for the quadratic output case, we can derive the error

$$\begin{aligned}
\mathfrak{E}_{\mathfrak{J}_{\mathrm{Q}}}(c,g) &:= \mathfrak{J}_{\mathrm{Q}}(c,g) - \mathfrak{J}_{\mathrm{Q,r}}(c,g)\\
&= \operatorname{tr}\Big(\mathbf{M}_{11}\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\mathbf{M}_{11}\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\Big) - \operatorname{tr}\Big(\mathbf{M}_{11}\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\mathbf{M}_{11}\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\Big)\\
&= \operatorname{tr}\Big(\mathbf{M}_{11}\left(\mathbf{P}_{\mathrm{pos}}(c,g) - \widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\right)\mathbf{M}_{11}\left(\mathbf{P}_{\mathrm{pos}}(c,g) + \widetilde{\mathbf{P}}_{\mathrm{pos}}\right)\Big)\\
&= \operatorname{tr}\Big(\mathbf{M}_{11}\left(\mathbf{P}_{\mathrm{pos}}(c,g) - \widetilde{\mathbf{P}}_{\mathrm{pos}}\right)\mathbf{M}_{11}\left(\mathcal{P}_{\mathrm{pos}}(c,g) - \widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g) + 2\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\right)\Big)\\
&= \operatorname{tr}\Big(\mathbf{M}_{11}\mathfrak{E}_{\mathrm{so}}(c,g)\mathbf{M}_{11}\left(\mathfrak{E}_{\mathrm{so}}(c,g) + 2\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\right)\Big),
\end{aligned}$$

We notice that, the error $\mathfrak{E}_{\mathrm{so}}(c,g)$ is needed to compute the errors $\mathfrak{E}_{\mathfrak{J}_{\mathrm{L}}}(c,g)$ and $\mathfrak{E}_{\mathfrak{J}_{\mathrm{Q}}}(c,g)$. Hence, we apply a second EE-RBM to find an approximation $\widetilde{\mathfrak{E}}_{\mathrm{so}}(c,g) \approx \mathfrak{E}_{\mathrm{so}}(c,g)$ as introduced in Algorithm 16. Using the error approximation $\widetilde{\mathfrak{E}}_{\mathrm{so}}(c,g)$ from (5.45), we derive the error approximations

$$\begin{aligned}
\boldsymbol{\Delta}_{\mathfrak{J}_{\mathrm{L}}}(c,g) &:= \left|\operatorname{tr}\Big(\mathbf{C}_1^{\mathrm{T}}\widetilde{\mathfrak{E}}_{\mathrm{so}}(c,g)\mathbf{C}_1^{\mathrm{T}}\Big)\right|,\\
\boldsymbol{\Delta}_{\mathfrak{J}_{\mathrm{Q}}}(c,g) &:= \left|\operatorname{tr}\Big(\mathbf{M}_{11}\mathfrak{E}_{\mathrm{so}}(c,g)\mathbf{M}_{11}\left(\mathfrak{E}_{\mathrm{so}}(c,g) + 2\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\right)\Big)\right|,
\end{aligned} \tag{6.11}$$

which are used instead of the approximation $\boldsymbol{\Delta}_{\mathfrak{E}_{\mathrm{so}}}(c,g)$ from (5.46).

6.2.2 Damping optimization for second-order systems using an adaptive RBM for second-order systems

As explained in the previous section for first-order systems, we can use an adaptive scheme to build the basis $\mathbf{V}_{\mathrm{so,r}}$ so that we do not need prior knowledge about the parameter domain $\mathfrak{D}$. The idea of the adaptive RBM is to enrich the basis $\mathbf{V}_{\mathrm{so,r}}$ within the optimization process when the current approximation of the system response, described in (6.10), is not sufficiently good. Consequently, there is no decomposition into an offline and an online phase.

We select a parameter (c_0, g_0) as the initial value for the optimization process and determine a low-rank factor $\mathbf{Z}_{\mathrm{so}}(c_0, g_0)$ that is either $\mathbf{Z}_{\mathrm{so,BT}}(c_0, g_0)$ from (5.39) or $\mathbf{Z}_{\mathrm{so,IRKA}}(c_0, g_0)$ from (5.40). We set the first basis to be $\mathbf{V}_{\mathrm{so,r}} = \operatorname{orth}(\mathbf{Z}_{\mathrm{so}}(c_0, g_0))$ and define the respective reduced system response (6.10). The function that is optimized is defined in Algorithm 21. In every iteration of the minimization, we query the error approximation $\boldsymbol{\Delta}(c,g)$ of the current parameter (c,g) as described in Step 2. If the error approximation is smaller than a given tolerance, we proceed with the function evaluation. In Step 5 and 6, we determine the resulting function value $\mathfrak{J}_{\mathrm{L,r}}(c,g)$ or $\mathfrak{J}_{\mathrm{Q,r}}(c,g)$ as defined in (6.6) or (6.7), respectively, and continue with the minimization. On the other hand, if the error approximation is larger than the tolerance, we know that the current basis $\mathbf{V}_{\mathrm{so,r}}$ does not approximate the controllability space of the systems (1.3) and (1.4) for the current parameter (c, g) sufficiently good. Hence, we return that the minimization did not converge and enrich the basis $\mathbf{V}_{\mathrm{so,r}}$. Therefore, we compute $\mathbf{Z}_{\mathrm{so}}(c, g)$ that is either $\mathbf{Z}_{\mathrm{so,BT}}(c, g)$ from (5.39) or $\mathbf{Z}_{\mathrm{so,IRKA}}(c, g)$ from (5.40) and define the updated basis

$$\mathbf{V}_{\mathrm{so,r}} = \mathrm{orth}(\begin{bmatrix}\mathbf{V}_{\mathrm{so,r}} & \mathbf{Z}_{\mathrm{so}}(c, g)\end{bmatrix}).$$

Consequently, we obtain a new optimization problem (6.10) that is defined using the new basis $\mathbf{V}_{\mathrm{so,r}}$. Since the optimized function depends on the current basis $\mathbf{V}_{\mathrm{so,r}}$, which changes during the optimization procedure, convergence problems may occur. Hence, we start a new optimization procedure whenever we enrich the basis and use the current parameter (c, g) as the initial value. We continue with this procedure until the optimum is reached.

```
Algorithm 21 Reduced second-order system response.
Input: $\mathbf{M}, \mathbf{K} \in \mathbb{R}^{n \times n}$, $\mathbf{D} : \mathfrak{D} \to \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$, $\mathbf{C}_1 \in \mathbb{R}^{p \times n}$ or $\mathbf{M}_{11} \in \mathbb{R}^{n \times n}$,
parameter $(c, g)$, basis $\mathbf{V}_{\mathrm{so,r}}$, tolerance tol.
Output: Energy response $\mathfrak{J}_{\mathrm{L,r}}(c, g)$ or $\mathfrak{J}_{\mathrm{Q,r}}(c, g)$, variable conv.
Set conv = true.
if $\mathbf{\Delta}(c, g) > \mathrm{tol}$ then
    Set conv = false, $\mathfrak{J}_{\mathrm{r}}(c, g) = \infty$.
else
    Solve the reduced Lyapunov equation (5.9) to obtain $\mathcal{P}_{\mathrm{r}}(c, g)$ including $\mathbf{P}_{\mathrm{pos,r}}(c, g)$.
    Compute $\mathfrak{J}_{\mathrm{L,r}}(c, g)$ or $\mathfrak{J}_{\mathrm{Q,r}}(c, g)$ as defined in (6.10).
end if
```

Error approximation Finally, we introduce the error approximation $\mathbf{\Delta}(c, g)$ that is used in the adaptive scheme. We follow the same idea as in Section 5.2.1 to generate a basis $\mathbf{V}_{\mathrm{so,err}}$ that spans an approximation of the error space. Therefore, we run a second EE-RBM to generate such a basis $\mathbf{V}_{\mathrm{so,err}}$ that is enlarged whenever the basis $\mathbf{V}_{\mathrm{so,r}}$ is expanded. In this way, the error approximation, and thus the error approximation from (6.11), becomes more accurate the closer we get to the optimizing parameter. Using the bases $\mathbf{V}_{\mathrm{so,r}}$ and $\mathbf{V}_{\mathrm{so,err}}$, we define the error approximations $\mathbf{\Delta}_{\mathfrak{J}_{\mathrm{L}}}(c, g)$ or $\mathbf{\Delta}_{\mathfrak{J}_{\mathrm{Q}}}(c, g)$ from (6.11).

The detailed procedure is described in Algorithm 22. When we determine the first basis $\mathbf{V}_{\mathrm{so,r}} = \mathrm{orth}(\mathbf{Z}_{\mathrm{so}}(c_0, g_0))$, we solve a second Lyapunov equation in an arbitrary parameter $(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$ with $(c_0^{\mathrm{r}}, g_0^{\mathrm{r}}) \neq (c_0, g_0)$ to obtain the solution $\mathbf{Z}_{\mathrm{so}}(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$. To limit the

Algorithm 22 Adaptive second-order RBM.

Input: $\mathbf{M}, \mathbf{K} \in \mathbb{R}^{n\times n}$, $\mathbf{D} : \mathbb{R}^{\ell} \times \mathbb{N}_{+}^{\ell} \to \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C}_1 \in \mathbb{R}^{p\times n}$ or $\mathbf{M}_{11} \in \mathbb{R}^{n\times n}$, tolerance tol.

Output: Minimizer $(c^{\text{opt}}, g^{\text{opt}})$, energy response $\mathfrak{J}_{\text{L,r}}(c^{\text{opt}}, g^{\text{opt}})$ or $\mathfrak{J}_{\text{Q,r}}(c^{\text{opt}}, g^{\text{opt}})$.

```
Choose $(c_0, g_0)$, $(c_0^{\text{r}}, g_0^{\text{r}}) \in \mathfrak{D}_{\text{Test}}$ with $(c, g) \neq (c^{\text{r}}, g^{\text{r}})$.
Compute $\mathbf{Z}_{\text{so}}(c_0, g_0)$ that is either $\mathbf{Z}_{\text{so,BT}}(c_0, g_0)$ from (5.39) or $\mathbf{Z}_{\text{so,IRKA}}(c_0, g_0)$ from (5.40).
Set $\mathbf{V}_{\text{so,r}} := \operatorname{orth}(\mathbf{Z}_{\text{so}}(c_0, g_0))$.
Compute $\mathbf{Z}_{\text{so}}(c_0^{\text{r}}, g_0^{\text{r}})$ that is either $\mathbf{Z}_{\text{so,BT}}(c_0^{\text{r}}, g_0^{\text{r}})$ from (5.39) or $\mathbf{Z}_{\text{so,IRKA}}(c_0^{\text{r}}, g_0^{\text{r}})$ from (5.40).
Set $\mathbf{V}_{\text{so,err}} = \operatorname{orth}([\mathbf{Z}_{\text{so}}(c_0, g_0),\ \mathbf{Z}_{\text{so}}(c_0^{\text{r}}, g_0^{\text{r}})])$.
Find the minimizer $(c^{\text{opt}}, g^{\text{opt}})$ of the function Algorithm 21 using fminsearch and obtain $\mathfrak{J}_{\text{L,r}}(c^{\text{opt}}, g^{\text{opt}})$ or $\mathfrak{J}_{\text{Q,r}}(c^{\text{opt}}, g^{\text{opt}})$, and conv.
while conv = false do
    Compute $\mathbf{Z}_{\text{so}}(c^{\text{opt}}, g^{\text{opt}})$ that is either $\mathbf{Z}_{\text{so,BT}}(c^{\text{opt}}, g^{\text{opt}})$ from (5.39) or $\mathbf{Z}_{\text{so,IRKA}}(c^{\text{opt}}, g^{\text{opt}})$ from (5.40)
    Set $\mathbf{V}_{\text{so,r}} := \operatorname{orth}([\mathbf{V}_{\text{so,r}}, \mathbf{Z}_{\text{so}}(c^{\text{opt}}, g^{\text{opt}})])$.
    Determine $(c^{\text{r}}, g^{\text{r}}) := \operatorname{argmax}_{(c,g)\in\mathfrak{D}_{\text{Test}}} \|\mathfrak{R}_{\text{so,r}}(c, g)\|_{\text{F}}$.
    Compute $\mathbf{Z}_{\text{so}}(c^{\text{r}}, g^{\text{r}})$ that is either $\mathbf{Z}_{\text{so,BT}}(c^{\text{r}}, g^{\text{r}})$ from (5.39) or $\mathbf{Z}_{\text{so,IRKA}}(c^{\text{r}}, g^{\text{r}})$ from (5.40).
    Set $\mathbf{V}_{\text{so,err}} = \operatorname{orth}([\mathbf{V}_{\text{so,err}},\ \mathbf{Z}_{\text{so}}(c^{\text{opt}}, g^{\text{opt}}),\ \mathbf{Z}_{\text{so}}(c^{\text{r}}, g^{\text{r}})])$.
    Find the minimizer $(c^{\text{opt}}, g^{\text{opt}})$ of the function Algorithm 21 using fminsearch and obtain $\mathfrak{J}_{\text{L,r}}(c^{\text{opt}}, g^{\text{opt}})$ or $\mathfrak{J}_{\text{Q,r}}(c^{\text{opt}}, g^{\text{opt}})$, and conv.
end while
```

possibilities of choosing $(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$, we again define a finite subset $\mathfrak{D}_{\mathrm{Test}} \subset \mathfrak{D}$ and pick the parameter $(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$ from this finite set $\mathfrak{D}_{\mathrm{Test}}$. We use the solution $\mathbf{Z}_{\mathrm{so}}(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})$ and the basis $\mathbf{V}_{\mathrm{so,r}}$ to obtain the first error equation basis

$$\mathbf{V}_{\mathrm{so,err}} = \mathrm{orth}([\mathbf{V}_{\mathrm{so,r}},\ \mathbf{Z}_{\mathrm{so}}(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})]) = \mathrm{orth}([\mathbf{Z}_{\mathrm{so}}(c_0, g_0),\ \mathbf{Z}_{\mathrm{so}}(c_0^{\mathrm{r}}, g_0^{\mathrm{r}})]).$$

Using the bases $\mathbf{V}_{\mathrm{so,r}}$ and $\mathbf{V}_{\mathrm{so,err}}$, we compute the error approximation $\boldsymbol{\Delta}_{\mathfrak{J}_{\mathrm{L}}}(c, g)$ or $\boldsymbol{\Delta}_{\mathfrak{J}_{\mathrm{Q}}}(c, g)$ as in (6.11). Again, we choose in Step 12 the parameter $(c^{\mathrm{r}}, g^{\mathrm{r}}) \in \mathfrak{D}_{\mathrm{Test}}$ that results in the largest residual

$$(c^{\mathrm{r}}, g^{\mathrm{r}}) = \mathrm{argmax}_{(c,g)\in\mathfrak{D}_{\mathrm{Test}}} \|\mathfrak{R}_{\mathrm{so,r}}(c, g)\|_{\mathrm{F}}$$

with $\mathfrak{R}_{\mathrm{so,r}}(c, g) = \mathfrak{R}_{11}(c, g)$ as defined in (5.44).

6.2.3 Damping optimization for second-order systems using an adaptive RBM with a decoupled controllability space for second-order systems

Finally, we can combine the adaptive RBM from Section 6.2.2 with the decoupling of the second-order controllability space presented in Section 5.2.2. For that we again initialize a basis $\mathbf{V}_{\mathrm{so,r}} = \mathrm{orth}(\mathbf{Z}_{\mathrm{so},\mathbf{B}})$ where $\mathbf{Z}_{\mathrm{so},\mathbf{B}}$ is a basis that spans an approximation of the controllability space of the second-order system (5.52) with no external damping. We obtain such a basis either by computing $\mathbf{Z}_{\mathrm{so},\mathbf{B},\mathrm{BT}}$ from (5.53) or by computing $\mathbf{Z}_{\mathrm{so},\mathbf{B},\mathrm{IRKA}}$ from (5.56a). Using the basis $\mathbf{V}_{\mathrm{so,r}}$ we define the reduced optimization problem from (6.10) and run the optimization method to optimize the function defined in Algorithm 21. If the output conv is equal to false, the current basis $\mathbf{V}_{\mathrm{so,r}}$ does not approximate the controllability space for this parameter sufficiently well. Therefore, we enrich the basis $\mathbf{V}_{\mathrm{so,r}}$ by a second basis $\mathbf{Z}_{\mathrm{so},\mathbf{F}}(c)$ that approximates the controllability space of the system (5.54) in c. For that we use $\mathbf{Z}_{\mathrm{so},\mathbf{F}}(c)$ that is either equal to $\mathbf{Z}_{\mathrm{so},\mathbf{F},\mathrm{BT}}(c)$ as defined in (5.55) or $\mathbf{Z}_{\mathrm{so},\mathbf{F},\mathrm{IRKA}}(c)$ as defined in (5.56b). After we built the new basis $\mathbf{V}_{\mathrm{so,r}}$ as

$$\mathbf{V}_{\mathrm{so,r}} = \mathrm{orth}(\begin{bmatrix}\mathbf{V}_{\mathrm{so,r}} & \mathbf{Z}_{\mathrm{so},\mathbf{F}}(c),\end{bmatrix})$$

we define a new reduced optimization problem (6.10) and start an optimization process with (c, g) as initial parameters as described in Algorithm 23. We continue this process until $\mathsf{conv} = \mathsf{true}$. We want to emphasize that we do not need a given parameter set $\mathfrak{D}$ to apply Algorithm 23, which is advantageous compared to the previous methods.

6.3 Numerical results

In this section, we apply the different reduction strategies for optimizing external dampers to selected examples. For that, we first consider vibrational systems, where only the

Algorithm 23 Adaptive second-order RBM using a decoupled controllability space.

```
Input: $\mathbf{M},\mathbf{K}\in\mathbb{R}^{n\times n}$, $\mathbf{D}:\mathbb{R}^{\ell}\times\mathbb{N}_{+}^{\ell}\to\mathbb{R}^{n\times n}$, $\mathbf{B}\in\mathbb{R}^{n\times m}$, $\mathbf{C}_1\in\mathbb{R}^{p\times n}$ or $\mathbf{M}_{11}\in\mathbb{R}^{n\times n}$, tolerance tol.
Output: Minimizer $(c^{\text{opt}},g^{\text{opt}})$, energy response $\mathfrak{J}_{\text{L,r}}(c^{\text{opt}},g^{\text{opt}})$ or $\mathfrak{J}_{\text{Q,r}}(c^{\text{opt}},g^{\text{opt}})$.
Compute the basis $\mathbf{Z}_{\text{so},\mathbf{B}}$ that is equal to $\mathbf{Z}_{\text{so},\mathbf{B},\text{BT}}$ as in (5.53) or $\mathbf{Z}_{\text{so},\mathbf{B},\text{IRKA}}$ as in (5.56a).
Set $\mathbf{V}_{\text{so,r}}:=\operatorname{orth}(\mathbf{Z}_{\text{so},\mathbf{B}})$.
Find the minimizer $(c^{\text{opt}},g^{\text{opt}})$ of the function Algorithm 21 using fminsearch and obtain $\mathfrak{J}_{\text{L,r}}(c^{\text{opt}},g^{\text{opt}})$ or $\mathfrak{J}_{\text{Q,r}}(c^{\text{opt}},g^{\text{opt}})$, and conv.
while conv = false do
    Compute the basis $\mathbf{Z}_{\text{so},\mathbf{F}}(c^{\text{opt}})$ that is equal to $\mathbf{Z}_{\text{so},\mathbf{F},\text{BT}}(c^{\text{opt}})$ as in (5.55) or $\mathbf{Z}_{\text{so},\mathbf{F},\text{IRKA}}(c^{\text{opt}})$ as in (5.56b).
    Set $\mathbf{V}_{\text{so,r}}:=\operatorname{orth}([\mathbf{V}_{\text{so,r}},\mathbf{Z}_{\text{so},\mathbf{F}}(c^{\text{opt}})])$.
    Find the minimizer $(c^{\text{opt}},g^{\text{opt}})$ of the function Algorithm 21 using fminsearch and obtain $\mathfrak{J}_{\text{L,r}}(c^{\text{opt}},g^{\text{opt}})$ or $\mathfrak{J}_{\text{Q,r}}(c^{\text{opt}},g^{\text{opt}})$, and conv.
end while
```

damper's viscosities are optimized as presented, e.g., in [107, 140]. Afterwards, we optimize the damper's positions while fixing the damper's viscosities, and finally, we optimize both parameters simultaneously. Optimizing the first-order representation of vibrational systems (1.3) and (1.4) leads to slower results and computational problems since reducing the first-order matrices $\mathcal{E}$ and $\mathcal{A}(c,g)$ from (1.7) can lead to reduced matrix pencils $\lambda\mathcal{E}_{\text{r}}-\mathcal{A}_{\text{r}}$ where the eigenvalues λ have a nonnegative real-part. Then, the respective Lyapunov equation is not uniquely solvable. Therefore, we only consider the second-order representation and use the first-order reducing basis $\mathbf{V}_{\text{r}}$ as defined in (5.41).

In this section, we illustrate the accelerations that arise when we apply the methods presented in this work to optimize the external dampers. We run the four different algorithms derived in Section 6.2, each with a Gramians-based basis building and with an IRKA-based one. Hence, we evaluate the offline-online RBM from Algorithm 16 using Gramians (`off-on RBM BT`) and using the IRKA method (`off-on RBM IRKA`). Also, we evaluate the offline-online RBM using the decomposition introduced in Section 5.2.2, which leads to Algorithm 17 using again Gramians (`dec off-on RBM BT`) and the IRKA method (`dec off-on RBM IRKA`). Moreover, we apply the adaptive scheme from Algorithm 22 using Gramians (`adpt RBM BT`) and using IRKA (`adpt RBM IRKA`) and the respective decomposed controllability space in Algorithm 23 using Gramians (`dec adpt RBM BT`) and the IRKA method (`dec adpt RBM IRKA`).

The computations have been done on a computer with 2 Intel Xeon Silver 4110 CPUs running at 2.1 GHz and equipped with 192 GB total main memory. The experiments use Matlab 2021a, and the Lyapunov equations were solved using methods from M-

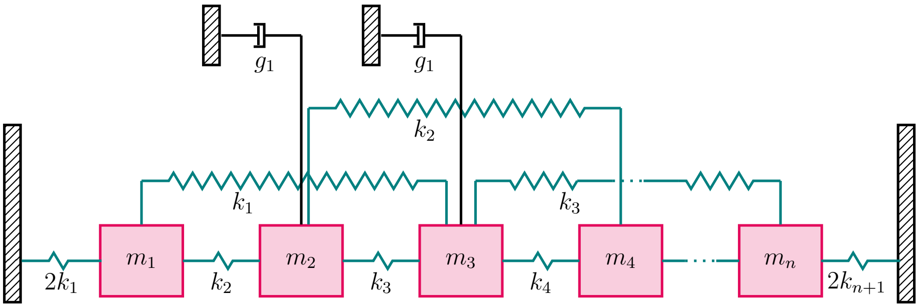


Figure 6.1: Example 6 - Sketch of the system including one row of masses connected by consecutive springs.

M.E.S.S.-2.1., see [114]. All results are available at [103].

6.3.1 Damping value optimization

First, we consider examples where we optimize the dampers viscosities. We evaluate two examples of second-order structure, where the first example has a quadratic output equation as presented in (1.4) while the second example evaluates a linear equation as output so that we consider a system of the structure (1.3).

Example 6 First, we consider a vibrational system (1.4) with a quadratic output equation. It was introduced in [140] and arises in mechanical constructions with n consecutive masses. Each mass m_j is connected to the direct neighbor masses m_{j-1} and m_{j+1} by springs with stiffness values k_j and k_{j+1}. Additionally, each mass is connected by springs with stiffness values k_{j-1} and k_{j+2} to the masses next to the neighbor masses m_{j-2} and m_{j+2}. The outermost masses are connected to fixed objects via springs with constants $2k_1$ and $2k_{n+1}$. This construction is depicted in Figure 6.1 and results in the following mass and stiffness matrix

$$
\begin{aligned}
\mathbf{M} &:= \operatorname{diag}(m_1, \ldots, m_n), \\
\mathbf{K} &:= \begin{bmatrix}
2k_1+2k_2 & -k_2 & -k_3 & & & \\
-k_2 & 2k_2+2k_3 & -k_3 & -k_4 & & \\
-k_3 & -k_3 & 2k_3+2k_4 & -k_4 & -k_5 & \\
 & \ddots & \ddots & \ddots & \ddots & \\
 & & & 2k_{n-2}+2k_{n-1} & -k_{n-2} & -k_{n-1} \\
 & & & -k_{n-1} & 2k_{n-1}+2k_n & -k_n \\
 & & & -k_{n-2} & -k_n & 2k_n+2k_{n+1}
\end{bmatrix}.
\end{aligned}
$$

We consider an example of dimension $n = 1900$ with stiffness constants $k_j = 500,\ j = 1,\ \dots,\ n$. The mass values are chosen as

$$m_j = \begin{cases} 144 - \frac{3}{20}j, & j = 1, \dots, 475, \\ \frac{j}{10} + 25, & j = 476, \dots, 1900. \end{cases}$$

The internal damping $\mathbf{D}_{\mathrm{int}}$ is built as described in (1.2) where the scaling factor is $\alpha = 0.005$. We consider external disturbance forces that attack at the sequential masses from m_{471} to m_{480}. Hence, in the input matrix $\mathbf{B}$ the values at positions 471 to 480 are set to be

$$\mathbf{B}(471 : 480,\ 1 : 10) = \operatorname{diag}(10,\ 20,\ 30,\ 40,\ 50,\ 50,\ 40,\ 30,\ 20,\ 10)\,.$$

The remaining entries of $\mathbf{B}$ are equal to zero. Consequently, we have a $(n \times 10)$-dimensional input matrix $\mathbf{B}$, where the highest magnitude of disturbance is applied to the mass in the center, whereby the disturbance magnitude gets smaller in the outer masses. To observe the system behavior, we consider the displacement of the states $x_{100}(t),\ x_{200}(t),\ \dots,\ x_{1800}(t)$. In contrast to the example in [140], we consider a quadratic output equation. Hence, the output matrix $\mathbf{M}_{11}$ has zero entries everywhere except on the positions

$$(100, 100),\ (200, 200),\ \dots,\ (1800, 1800),$$

where the entries are equal to one. We consider four dampers on the positions $j,\ j+1,\ k,\ k+1$ where j and k can take the following values

$$\{(j, k) \mid j \in \{50,\ 150,\ 250,\ 350\},\ k \in \{850,\ 950,\ \dots,\ 1850\}\}.$$

Hence, we evaluate the system for 44 possible damping configurations. For each damping configuration, we optimize the damping values individually. The damping gains g consist of two values g_1 and g_2 where the dampers on the j-th and the $(j+1)$-th position have the damping value g_1 and the dampers on the k-th and the $(k+1)$-th position the damping value g_2. We assume that the damping values g_1 and g_2 lie in the interval $[500, 4000]$.

To optimize the damping gains for the different damping configurations, we use the Matlab function `fminsearch` where we stop the minimization process if the difference between two successive function values or damping viscosities is smaller than a tolerance $\mathrm{tol} = 10^{-4}$. We start the optimization process at $g_0 = \begin{bmatrix} 1000 & 1000 \end{bmatrix}^{\mathrm{T}}$ for all damping configurations. To solve the Lyapunov equations from (1.8), we use the sign-function method presented in Section 2.3.2 with $\mathrm{tol} = 10^{-6}$ and a maximum iteration number of $\mathrm{iter}_{\max} = 10$ because of the fast dimension growth within the method. As test-parameter set $\mathfrak{D}_{\mathrm{Test}}$, we use 36 uniformly distributed parameters in $[500, 4000] \times [500, 4000]$.

To show that the error approximations $\boldsymbol{\Delta}_{\mathfrak{E}_{\mathrm{so}}}(c, g)$ and $\boldsymbol{\Delta}_{\mathfrak{J}_{\mathrm{Q}}}(c, g)$ from (5.46) and (6.5), respectively, approximate the error well, we evaluate the quality of the error approximation in Figure 6.2 after the first step of the offline-online RBM for the 11-th damper configuration. We observe that the relative error in the position controllability Gramian $\|\mathbf{E}_{11}(c,g)\|_{\mathrm{F}}/\|\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\|_{\mathrm{F}}$ and the corresponding approximation $\|\widetilde{\mathbf{E}}_{11}(c,g)\|_{\mathrm{F}}/\|\widetilde{\mathbf{P}}_{\mathrm{pos}}(c,g)\|_{\mathrm{F}}$ are very close to the actual error. On the other hand, the energy response is underestimated as we evaluate an error approximation and not an error bound. However, for our purposes, the quality was good enough as the error in the energy response and its approximate value have a similar order of magnitude.

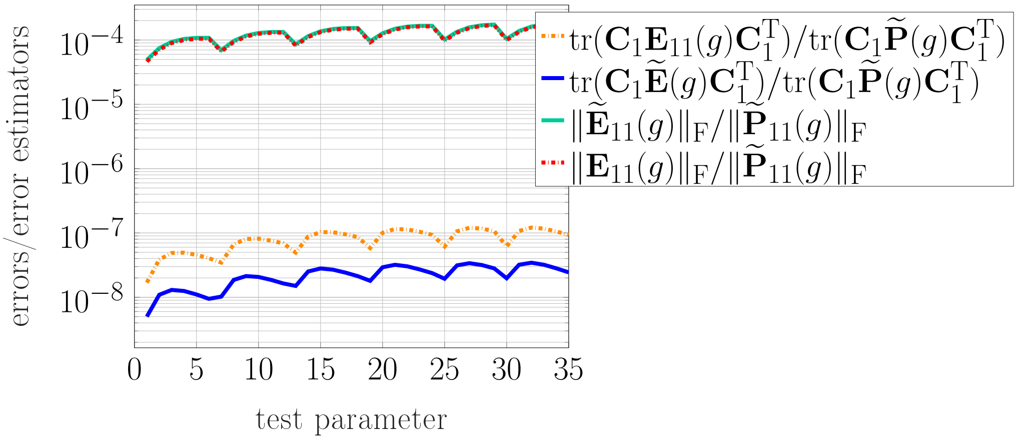


Figure 6.2: Example 6 - Errors approximations for the first damping configuration and the first step of the reduced basis method.

Since the initial value g_0 is known, we choose this parameter as the first one evaluated within the RBM. The first parameter g_0^{r} used to derive a first error equation basis is chosen to be $g_0^{\mathrm{r}} = \begin{bmatrix}100 & 100\end{bmatrix}$ within the offline-online RBM and the adaptive RBM schemes.

The relative errors between the optimal damping gain and the approximations obtained using the methods presented in the previous sections are presented in Figure 6.3. We observe that all methods approximate the optimal viscosity well. However, the methods `dec off-on RBM BT` and `dec adpt RBM BT` each lead for one configuration to an error larger than the tolerance of tol $= 10^{-2}$. We observe that all of the methods approximate, on average, the optimal viscosities sufficiently good, while the methods using IRKA lead to even better results for this example.

We also evaluate the optimization times, which include the offline and the online phases when considering the offline-online schemes and the overall methods when adaptive procedures are applied. We determine a low-rank factor of the solution of the Lyapunov equation (1.8) for the undamped system that is added to all bases considered. Since this low-rank factor is computed beforehand, the computation time of 39 seconds is not taken into account in any of these methods. We compare the optimization times in Figure 6.4 and the respective acceleration rates for the different methods in Figure 6.5. The Matlab-solver `lyapchol` is used to solve the Lyapunov equations from (1.8). The

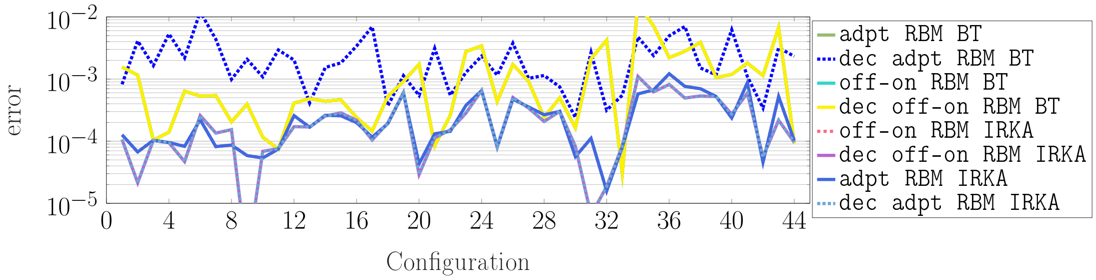


Figure 6.3: Example 6 - Viscosity errors

average errors and acceleration rates are summarized in Table 6.1. We observe that the `off-on RBM BT` and `dec off-on RBM BT` lead to the fastest results.

	Errors	Acceleration rates
`off-on RBM BT`	$1.64 \cdot 10^{-3}$	138
`off-on RBM IRKA`	$2.82 \cdot 10^{-4}$	38
`dec off-on RBM BT`	$1.64 \cdot 10^{-3}$	138
`dec off-on RBM IRKA`	$2.82 \cdot 10^{-4}$	38
`adpt RBM BT`	$1.64 \cdot 10^{-3}$	79
`adpt RBM IRKA`	$3.09 \cdot 10^{-4}$	38
`dec adpt RBM BT`	$2.50 \cdot 10^{-3}$	76
`dec adpt RBM IRKA`	$3.28 \cdot 10^{-4}$	47

Table 6.1: Example 6 - Comparison of the different algorithms.

We want to mention that the adaptive method is still advantageous since we do not need a parameter set $\mathfrak{D}$ in advance for it. Only for the error approximation, the test-parameter set $\mathfrak{D}_{\text{Test}} \subset \mathfrak{D}$ is needed that can be replaced by choosing arbitrary parameters. In particular, when using the decoupled controllability space and the respective error indicator $\boldsymbol{\Delta}_{\mathbf{P_F}}$ from (5.57), no prior knowledge about the parameter set is required. In Figure 6.6, the function values for all 44 damping configurations are evaluated. We observe that the optimal damping configuration is the 34-th one, which has the damping positions $j = 350,\ k = 850$.

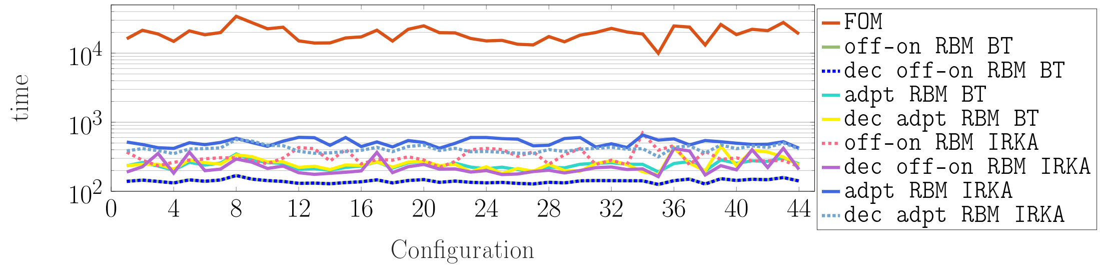


Figure 6.4: Example 6 - Optimization times

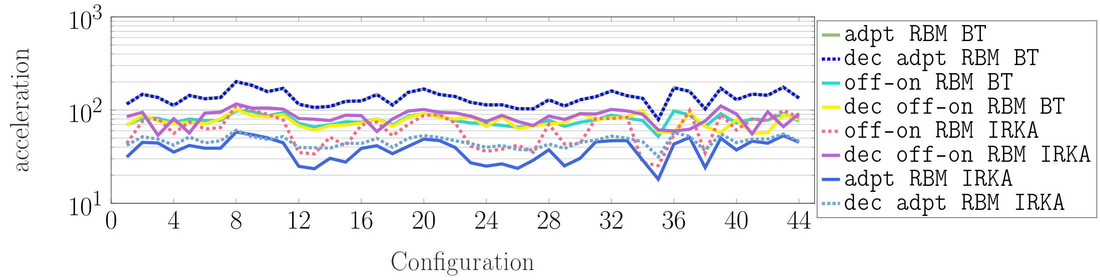


Figure 6.5: Example 6 - Acceleration rates

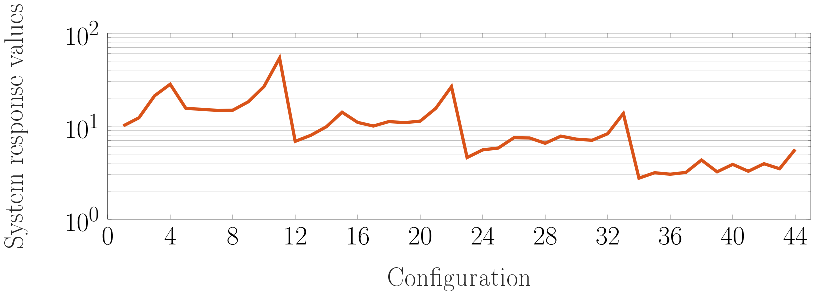


Figure 6.6: Example 6 - Function values

Example 7 The second example that we consider contains a mass oscillator with $2d+1=n$ masses and $n+2$ springs that result in a system (1.3) with a linear output equation as depicted in Figure 6.7. There are two lines of d consecutive masses $m_1, \ldots, m_d$ and

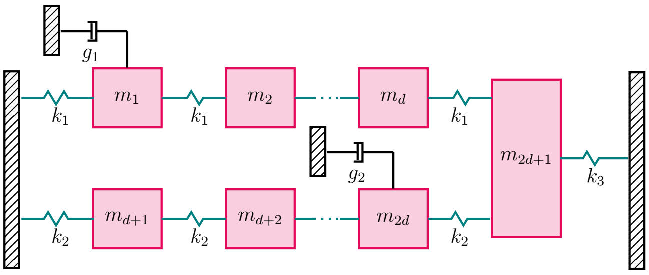


Figure 6.7: Example 7 - Sketch of the system including two rows of masses connected by consecutive springs.

$m_{d+1}, \ldots, m_{2d}$, which are connected by springs. The springs of the first line have all the stiffness value k_1, and the springs in the second line have the stiffness value k_2, where the masses m_1 and m_{d+1} are connected with these springs to a fixed object. The masses m_d and m_{2d} are connected to a mass $m_{2d+1} = m_n$ by springs with stiffness constants k_1 and k_2 while the mass m_n is connected to a fixed object via a spring with a constant $k_1 + k_2 + k_3$. This construction results in a stiffness matrix

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{11} & & \boldsymbol{\kappa}_1 \\ & \mathbf{K}_{22} & \boldsymbol{\kappa}_2 \\ \boldsymbol{\kappa}_1^{\mathrm{T}} & \boldsymbol{\kappa}_2^{\mathrm{T}} & k_1 + k_2 + k_3 \end{bmatrix}, \quad \mathbf{K}_{jj} = k_j \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}, \quad \boldsymbol{\kappa}_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ k_j \end{bmatrix},$$

for $j = 1, 2$. We choose the dimension to be $d = 1000$, $n = 2001$ and set $k_1 = 400$, $k_2 = 100$, $k_3 = 300$. The $n = 2d + 1$ mass values are chosen as follows

$$m_j = \begin{cases} 100 - \frac{j}{10}, & j = 1, \ldots, 500, \\ \frac{j}{30} + 33, & j = 501, \ldots, 1000, \\ 100 - (j - 99)\frac{5}{20} + \frac{(j-999)^2}{5000}, & j = 1001, \ldots, 2000, \end{cases} \qquad m_{2001} = 100.$$

The internal damping $\mathbf{D}_{\text{int}}$ is built as described in (1.2) with the scaling $\alpha = 0.003$. Additionally, some disturbances affect 21 masses. The effect on the masses is described by the matrix $\mathbf{B} \in \mathbb{R}^{n \times 21}$ that consists of zero entries except for the following entries

$$\begin{aligned} \mathbf{B}(1:10,\, 1:10) &= \operatorname{diag}(1000,\, 900,\, \ldots,\, 100), \\ \mathbf{B}(1001:1010,\, 11:20) &= \operatorname{diag}(1000,\, 900,\, \ldots,\, 100), \\ \mathbf{B}(2001, 21) &= 2000. \end{aligned}$$

As output, we observe 42 masses, or more detailed, the displacements of the masses 490 to 510 and those of the masses on positions 1490 to 1510. This is described by the output matrix $\mathbf{C} \in \mathbb{R}^{42 \times n}$:

$$\mathbf{C}(490 : 510,\, 1 : 21) = \mathbf{I}_{21}, \qquad \mathbf{C}(1490 : 1510,\, 1 : 21) = \mathbf{I}_{21}.$$

In this example, we consider four damping values that are optimized. We consider two dampers in the first row that are between the masses m_j and m_{j+5} and between the masses m_{j+20} and m_{j+25}. For the second row, we follow the same pattern and add two dampers between the masses m_k and m_{k+5} and between m_{k+20} and m_{k+25}. Consequently, we optimize four damping values $g_1,\ g_2,\ g_3,\ g_4$ that are saved in $g \in \mathbb{R}^4$. The corresponding damping position matrix is then of the form

$$\mathbf{F} = \begin{bmatrix} e_j - e_{j+5} & e_{j+20} - e_{j+25} & e_k - e_{k+5} & e_{k+20} - e_{k+25} \end{bmatrix},$$

where j and k are from the sets

$$\{(j, k) \mid j \in \{250,\, 450,\, 650,\, 850\},\ k \in \{1150,\, 1250,\, 1350,\, 1450,\, 1550,\, 1650,\, 1750\}\}.$$

This setting leads to 28 damping configurations. We assume that the damping values $g_1,\ g_2,\ g_3,\ g_4$ lie in the interval $[350, 7000]$. For the optimization process we set the tolerance of $5 \cdot 10^{-4}$ and start the optimization process at $g_0 = \begin{bmatrix} 1000 & 1000 & 1000 & 1000 \end{bmatrix}^{\mathrm{T}}$ for all damping configurations. The tolerance for the function value error that indicates whether a basis $\mathbf{V}_{\mathrm{so,r}}$ is sufficiently detailed is tol $= 10^{-2}$. As test-parameter set $\mathfrak{D}_{\mathrm{Test}}$ for the reduced basis method we use 21 uniformly distributed parameters in $[350, 7000]^4$. The first parameter g_0^{r} used to obtain a first error equation basis is chosen to be $g_0^{\mathrm{r}} = \begin{bmatrix} 100 & 100 & 100 & 100 \end{bmatrix}$ within the RBM and the adaptive RBM schemes.

We evaluate the quality of the error approximations $\mathbf{\Delta}_{\mathfrak{C}_{\mathrm{so}}}(c, g)$ and $\mathbf{\Delta}_{\mathfrak{J}_{\mathrm{L}}}(c, g)$ from (5.46) and (6.5), respectively, after the first step of the reduced basis method for the fifth damper configuration in Figure 6.8. We observe that the relative errors in the position controllability Gramian and the corresponding approximation are very close, so the error is well approximated.

In Figure 6.9, we evaluate the relative errors in the damper's viscosities for all considered methods. We observe that most of the methods approximate, on average, the optimal viscosities sufficiently good. However, the IRKA methods using a controllability space decomposition fail in approximating the original system behavior.

Additionally, we evaluate the optimization times. Outside of the applied methods, we determine a low-rank factor of the solution of the Lyapunov equation (1.8) for the undamped system. This low-rank factor is included in all the bases and is not taken into account in the time measures. This solving takes 55 seconds. We compare the optimization times and the acceleration rates for the different methods in Figure 6.10 and Figure 6.11, respectively. The average errors and respective acceleration times for all the methods are summarized in Table 6.2. We observe that for this example the Gramian based methods provide better approximations and acceleration times. The

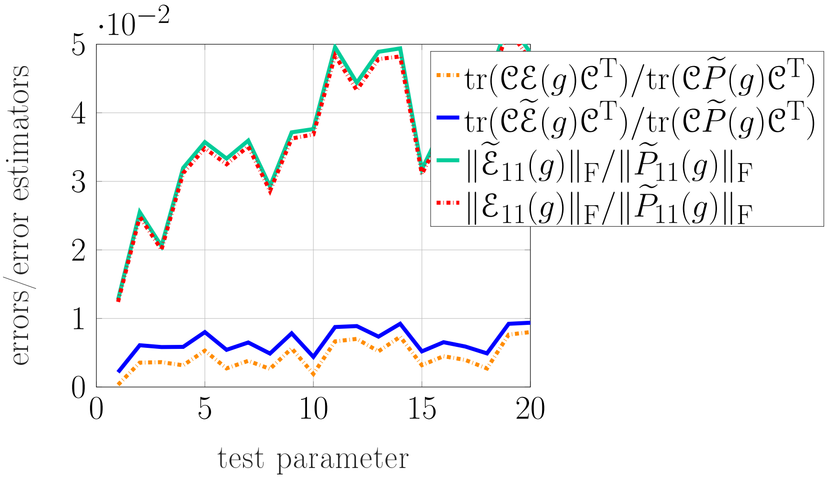


Figure 6.8: Example 7 - Errors approximation for the first damping configuration and the fifth step of the reduced basis method.

IRKA methods that use the decomposed controllability space even fail in approximating the optimal damping values.

	Errors	Acceleration rates
`off-on RBM BT`	$8.08 \cdot 10^{-3}$	209
`off-on RBM IRKA`	$1.17 \cdot 10^{-2}$	43
`dec off-on RBM BT`	$8.08 \cdot 10^{-3}$	209
`dec off-on RBM IRKA`	$2.93 \cdot 10^{-1}$	5
`adpt RBM BT`	$8.08 \cdot 10^{-3}$	95
`adpt RBM IRKA`	$1.17 \cdot 10^{-2}$	43
`dec adpt RBM BT`	$2.13 \cdot 10^{-2}$	148
`dec adpt RBM IRKA`	$2.93 \cdot 10^{-1}$	5

Table 6.2: Example 7 - Comparison of the different algorithms.

In Figure 6.12, the function values for all 28 damping configurations are evaluated. The optimal damping configuration is the 25-th one corresponding to the damping positions $j = 850,\ k = 1450$.

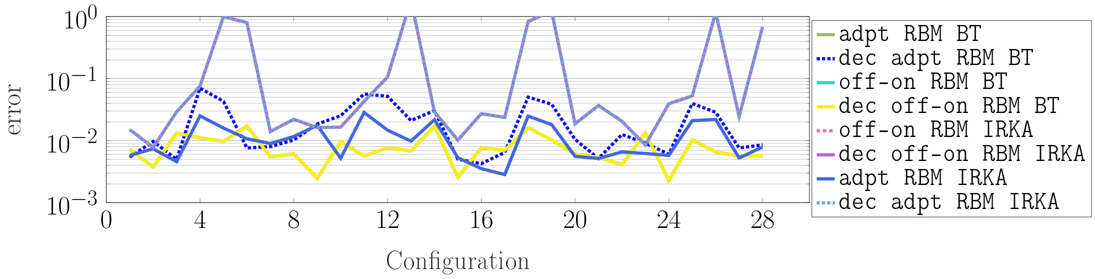


Figure 6.9: Example 7 - Viscosity errors

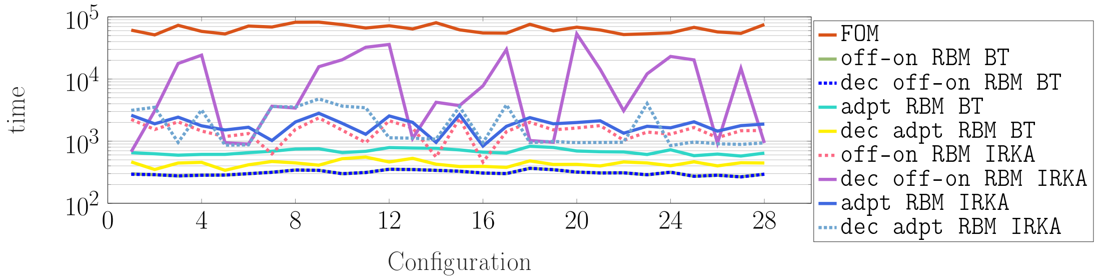


Figure 6.10: Example 7 - Optimization times

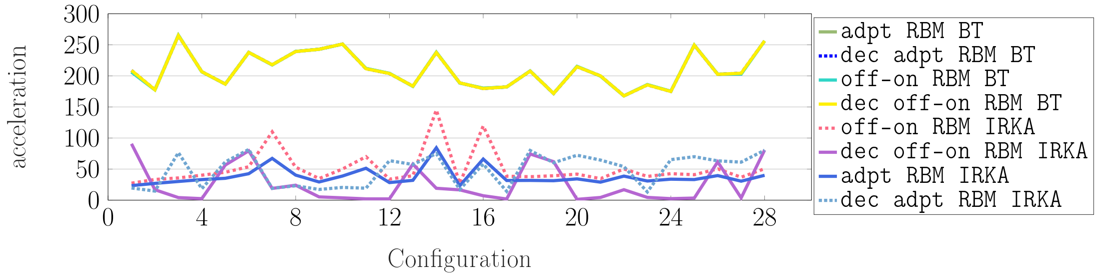


Figure 6.11: Example 7 - Acceleration rates

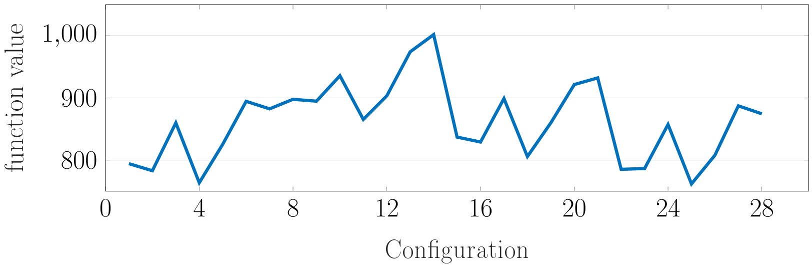


Figure 6.12: Example 7 - Function values

6.3.2 Damper position optimization

In this section, we apply the methods that are presented above to two numerical examples where the dampers' positions are optimized. For both examples, we first optimize only the damper's positions while considering fixed damping gains. Afterwards, we optimize simultaneously the positions and damping gains.

A difficulty in optimizing the position of dampers is that the positions are discrete values. Hence, the authors in [157] reformulate the optimization problem to apply standard optimization methods. Therefore, we define the function

$$\widetilde{\mathbf{J}}(c, g) := \mathbf{J}([c], g) = \mathbf{J}([c_1], \ldots, [c_\ell], g_1, \ldots, g_\ell) \tag{6.12}$$

that is a continuous function. This function is optimized in the following as we can apply, e.g., the Nelder–Mead method encoded in the Matlab function `fminsearch` to minimize (6.12).

We define a second function that assumes that the damping gains $g^* = [g_1^*, \ldots, g_\ell^*] \in \mathbb{R}^\ell$ are given and fixed, that is

$$\widetilde{\mathbf{J}}_{\text{pos}}(c) := \mathbf{J}([c], g^*) = \mathbf{J}([c_1], \ldots, [c_\ell], g_1^*, \ldots, g_\ell^*). \tag{6.13}$$

Optimizing this function using standard optimization methods such as `fminsearch` might cause convergence problems if the step size is too small because of the jumps in the function values in the function definition from (6.13). Hence, we modify this function to make the function values continuous. We first consider the case where we only have one damper with the position $c \in \mathfrak{D}_c \subset \mathbb{R}$. We split the current position value $c = c^{\text{int}} + c^{\text{dec}}$ where $c^{\text{int}} := \lfloor c \rfloor$ and $c^{\text{dec}} = c - c^{\text{int}}$. The corresponding function value is then defined as

$$\widehat{\mathbf{J}}_{\text{pos}}(c) := (1 - c^{\text{dec}})\mathbf{J}(c^{\text{int}}, g^*) + c^{\text{dec}}\mathbf{J}(c^{\text{int}} + 1, g^*),$$

which provides a linear interpolation between the function values corresponding to two discrete damper positions.

This idea is now generalized for ℓ dampers, i.e., $c \in \mathfrak{D}_c \subset \mathbb{R}^\ell$. We define for $c = [c_1, \ldots, c_\ell]$ the values

$$c_j := c_j^{\text{int}} + c_j^{\text{dec}} \quad \text{with} \quad c_j^{\text{int}} := \lfloor c_j \rfloor, \;\; c_j^{\text{dec}} = c_j - c_j^{\text{int}}, \qquad \text{for } j = 1, \ldots, \ell.$$

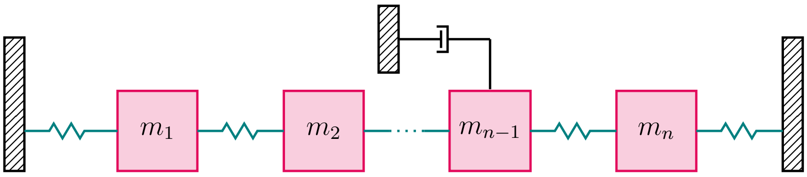


Figure 6.13: Example 8 - Sketch of the system including one row of masses connected by consecutive springs.

Accordingly, we define the time continuous function values as

$$\begin{aligned}\widehat{\mathbf{J}}_{\text{pos}}(c) := (1-c_1^{\text{dec}})\cdot\dots\cdot(1-c_\ell^{\text{dec}})\mathbf{J}([c_1^{\text{int}},\dots,c_\ell^{\text{int}}],g^*)\\
+c_1^{\text{dec}}(1-c_2^{\text{dec}})\cdot\dots\cdot(1-c_\ell^{\text{dec}})\mathbf{J}([c_1^{\text{int}}+1,\,c_2^{\text{int}},\dots,c_\ell^{\text{int}}],g^*)\\
+(1-c_1^{\text{dec}})\cdot\dots\cdot(1-c_{\ell-1}^{\text{dec}})c_\ell^{\text{dec}}\mathbf{J}([c_1^{\text{int}},\dots,c_{\ell-1}^{\text{int}},c_\ell^{\text{int}}+1],g^*)\\
\vdots\\
+(1-c_1^{\text{dec}})c_2^{\text{dec}}\cdot\dots\cdot c_\ell^{\text{dec}}\mathbf{J}([c_1^{\text{int}},c_2^{\text{int}}+1,\dots,c_\ell^{\text{int}}+1],g^*)\\
+c_1^{\text{dec}}\cdot\dots\cdot c_{\ell-1}^{\text{dec}}(1-c_\ell^{\text{dec}})\mathbf{J}([c_1^{\text{int}}+1,\dots,c_{\ell-1}^{\text{int}}+1,c_\ell^{\text{int}}],g^*)\\
+c_1^{\text{dec}}\cdot\dots\cdot c_\ell^{\text{dec}}\mathbf{J}([c_1^{\text{int}}+1,\dots,c_\ell^{\text{int}}+1],g^*).\end{aligned}$$

We observe that the computation of the function values of $\widehat{\mathbf{J}}_{\text{pos}}(c)$ is only accessible for a small number of external dampers or small system dimensions since the number of function evaluations rises exponentially with the numbers of dampers, where we need 2^ℓ Lyapunov equation solves if ℓ is the number of the dampers.

In our examples, the function defined in (6.12) does not have a converging problem, which is why we apply the function reformulation into $\widehat{\mathbf{J}}_{\text{pos}}$ only for the case of fixed viscosities. When viscosities and positions are optimized, such an approach is not needed for the examples considered in this work.

When applying the methods `off-on RBM BT`, `off-on RBM IRKA`, `adpt RBM BT`, and `adpt RBM IRKA`, we use the error approximation $\mathbf{\Delta}_{\mathcal{J}_{\text{L}}}$ or $\mathbf{\Delta}_{\mathcal{J}_{\text{Q}}}$ from (6.11) depending on the considered model. On the other hand, we use the error indicator $\mathbf{\Delta}_{\mathbf{P_F}}$ from (5.57), when the methods `dec off-on RBM BT`, `dec off-on RBM IRKA`, `dec adpt RBM BT`, and `dec adpt RBM IRKA` are applied.

Example 8 The example, we consider in this paragraph, is described in Figure 6.13. We evaluate a system (1.4) with a quadratic output matrix where the mass matrix is defined in Matlab notation as

$$\mathbf{M} = \texttt{sparse(diag([logspace(}-1,1,\texttt{n}/2),\texttt{flip(logspace(}-1,1,\texttt{n}/2))]))).$$

The matrix $\mathbf{M}$ leads to mass values between 0.01 and 0.1. The highest mass values are attained at the middle masses. The outermost masses have the smallest mass values. Moreover, the stiffness matrix is given as

$$\mathbf{K} = \begin{bmatrix} 24 & -20 & & & \\ -20 & 40 & -20 & & \\ & -20 & 40 & -20 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}.$$

We build the internal damping matrix $\mathbf{D}_{\text{int}}$ using a multiple of the critical damping with $\alpha = 0.005$, as described in (1.2). The dimension is $n = 1000$, so the Lyapunov equations of dimension $2n = 2000$ need to be solved multiple times. Additionally, the input matrix is defined as a zero matrix beside the entries

$$\mathbf{B}(1,1) = 1, \qquad \mathbf{B}(500,1) = 1, \qquad \mathbf{B}(1000,1) = 1.$$

Therefore, an external force is applied to the first, the middle, and the last mass. As output, we consider a quadratic function defined by an output matrix $\mathcal{M}$ with submatrices $\mathbf{M}_{11}$, $\mathbf{M}_{12} = 0$, and $\mathbf{M}_{22} = 0$, where $\mathbf{M}_{11}$ has everywhere zero entries besides on the $(10,10)$, $(500,500)$, $(990,990)$ positions where the entries are equal to one so that the output is equal to $\mathbf{y}_{\text{Q}}(t) = \mathbf{x}_{10}(t)^2 + \mathbf{x}_{500}(t)^2 + \mathbf{x}_{990}(t)^2$.

We apply two grounded dampers at positions k and j so that $c = [k,\ j]$ and $\mathbf{F}(c) = [e_k,\ e_l]$, where e_k and e_j are the k-th and the j-th unit vector. We apply the Nelder–Mead method to optimize the system response as defined in (6.13), which is implemented by the Matlab function `fminsearch`. We stop the optimization process when the relative error in the function values or difference between two consecutive values is smaller than the tolerance $\text{tol}_{\text{opt}} = 10^{-3}$.

We run the four different algorithms derived in Section 6.2, each with a Gramian-based basis building and with an IRKA-based one. For both error approximations $\boldsymbol{\Delta}_{\mathcal{J}_{\text{Q}}}$ form (6.11) and $\boldsymbol{\Delta}_{\mathcal{P}_{\text{F}}}$ from (5.57), we use a tolerance of $\text{tol} = 10^{-2}$ to quantify whether the current basis is sufficiently good. Moreover, if $\boldsymbol{\Delta}_{\mathfrak{C}_{\text{so}}}(c,g)$ from (5.46) is smaller than the tolerance $\text{tol} = 10^{-3}$, we stop the method.

Example 8a: Position optimization First, we only optimize the positions of the two dampers and set the damping gains to be the fixed values $g_1 = g_2 = 1000$. The initial positions are $c_0 = [50, 90]$. The respective timings and errors are evaluated in the following. In Figure 6.14, the different optimal positions derived using the full-order model and the reduced models are depicted. We see that the positions are approximated well by our methods except for the RBM methods using the decomposed controllability space. We observe that, in particular, when using the offline-online BT method to generate the basis and the adaptive methods using the respective controllability space decompositions, the approximations are not accurate enough. These relations are depicted in Figure 6.15, where we show the respective errors. Figure 6.16 depicts the

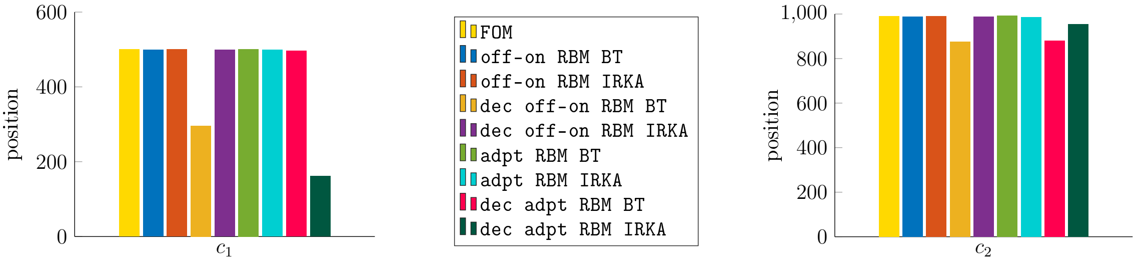


Figure 6.14: Example 8a - Position values.

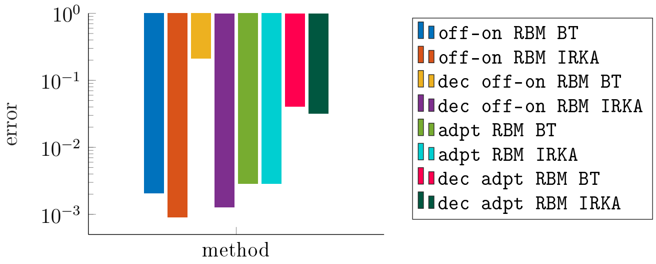


Figure 6.15: Example 8a - Position errors.

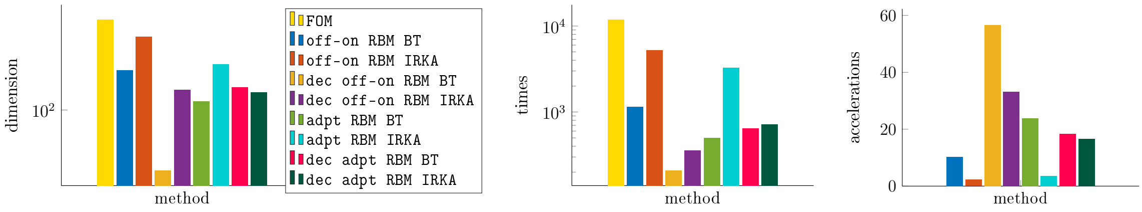


Figure 6.16: Example 8a - Dimensions, times, and acceleration rates.

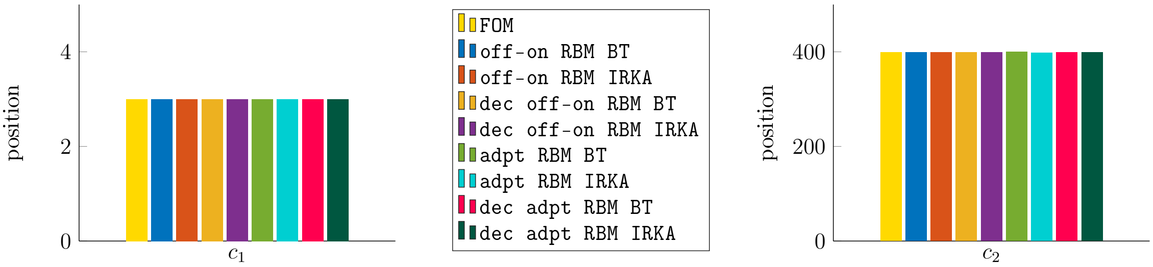


Figure 6.17: Example 8b - Position values.

dimensions, optimization times, and respective acceleration rates that result from the presented methods. The optimization times include the basis building, as well as the optimization of the dampers' positions. Moreover, we present the dimensions of the final reduced systems. We observe that the dimensions of the reduced systems are significantly smaller than the dimensions of the full-order model, where we emphasize that for every reduction approach, the Gramian based methods lead to faster results than the IRKA based ones. Also, we observe that the decomposition of the controllability spaces leads to faster results as the respective Gramians have the smallest dimensions. However, these small dimensions lead to the largest approximation errors.

Example 8b: Position and viscosity optimization Moreover, we optimize the damper's positions and the corresponding gains simultaneously. The initial positions are again chosen to be $c_0 = [50, 90]$, and the initial gains are $g_0 = [1000, 1000]$. The respective positions and gains are depicted in Figure 6.17 and Figure 6.18. Furthermore, the respective errors are shown in Figure 6.19. The bar plots that cover the complete range of the y-axis indicate that the respective error is equal to zero. We observe that the positions are well-approximated or even coincide with the optimal positions of the full-order system. Also, the viscosities are approximated well by the different methods, since all errors are smaller than $1.6 \cdot 10^{-3}$ which is less than 0.16 percent. In Figure 6.20, we depict the dimensions, optimization times, and respective acceleration rates corresponding to the full order and reduced surrogate models. We see that the different reduction approaches accelerate the computations significantly. In particular, the offline-online methods using the BT method to build the basis (with and without a decomposed controllability space) lead to the highest acceleration rates for this example. The offline-online method using the IRKA method and a decomposed controllability space leads to the smallest acceleration rates for this example. However, all the methods are doing sufficiently well.

Example 9 In this subsection, we investigate a system with three rows of masses connected by springs as depicted in Figure 6.21. The masses are given as described by

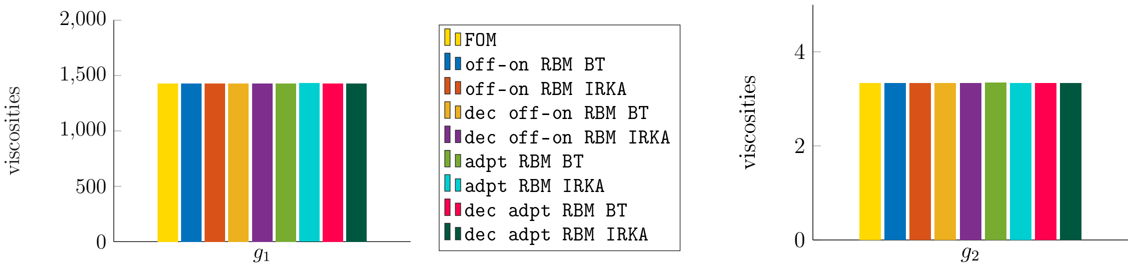


Figure 6.18: Example 8b - Viscosity values.

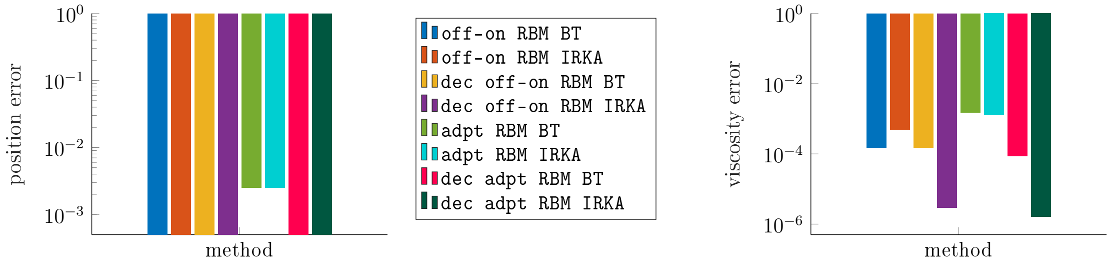


Figure 6.19: Example 8b - Position and viscosity errors.

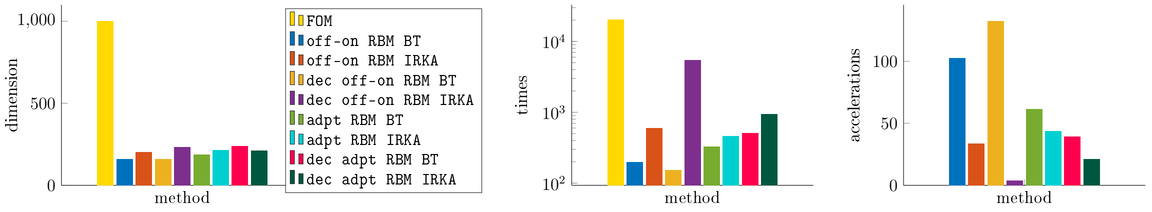


Figure 6.20: Example 8b - Dimensions, times, and acceleration rates.

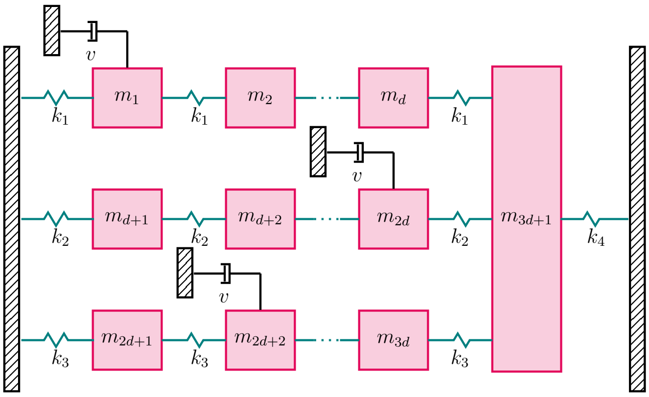


Figure 6.21: Example 9 - Sketch of the system including three rows of masses connected by consecutive springs.

the Matlab expression

$$\mathbf{M} = \texttt{1e4} * \texttt{sparse}(\texttt{diag}([\texttt{logspace}(-1, 1, \texttt{ceil}(\texttt{n}/2)),\\ \texttt{flip}(\texttt{logspace}(-1, 1, \texttt{floor}(\texttt{n}/2)))]));$$

while the stiffness matrix is built as

$$\mathbf{K} = \begin{bmatrix} K_{11} & & & \kappa_1 \\ & K_{22} & & \kappa_2 \\ & & K_{33} & \kappa_3 \\ \kappa_1^{\mathrm{T}} & \kappa_2^{\mathrm{T}} & \kappa_3^{\mathrm{T}} & k_1 + k_2 + k_3 + k_4 \end{bmatrix}, \qquad K_{ii} = k_i \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{d \times d},$$

with $\kappa_i = \begin{bmatrix} 0 & \cdots & 0 & k_i \end{bmatrix}^{\mathrm{T}}$ and $k_1 = 20$, $k_2 = 10$, $k_3 = 5$, $k_4 = 20$. We consider a system of dimension $n = 601 = 3d + 1$, $d = 200$ so that we have to solve Lyapunov equations of dimension $2n = 1202$. The input matrix is chosen to be $\mathbf{B} = -\texttt{ones}(\texttt{n}, \texttt{1})$. We consider a linear output equation defined by the the output matrix $\mathbf{C}$, which is the zero matrix of dimension $3 \times n$ with non-zero entries

$$\mathbf{C}(1, 10) = 1, \qquad \mathbf{C}(1, 450) = 1, \qquad \mathbf{C}(1, 891) = 1.$$

We assume that there are three grounded dampers so that $\mathbf{F} = [e_i,\ e_j,\ e_k]$ for $i,\ j,\ k \in \{1, \ldots, n\}$.

Example 9a: Position optimization Again, we initially consider the case where we only optimize the damper's positions. We stop the optimization process when the relative error in the function values or the difference between two consecutive values is smaller than the tolerance $\mathrm{tol}_{\mathrm{opt}} = 10^{-3}$. The corresponding results of the position optimization are given in Figure 6.22, where we chose the initial positions $c_0 = [100,\ 300,\ 500]$. We observe that the positions obtained by optimizing the full-dimensional problem are still approximated well enough, in the sense that the error is smaller than $1.3 \cdot 10^{-2}$ that is 1.3% for all the methods as shown in Figure 6.23, where we depict the resulting errors of the position optimization. Moreover, Figure 6.24 show the resulting dimensions, optimization times, and the acceleration rates, respectively. We observe that the offline-online scheme using the BT or the IRKA method, and the decoupled controllability space are leading to the highest acceleration rates. Also, the adaptive method using the IRKA method and the offline-online methods using BT or IRKA without the decomposition in the controllability space lead to dimensions that are almost as large as the original ones. For these cases, the acceleration rates are rather small.

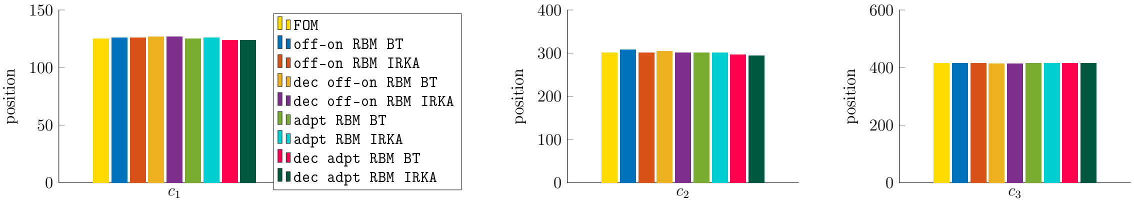


Figure 6.22: Example 9a - Position values.

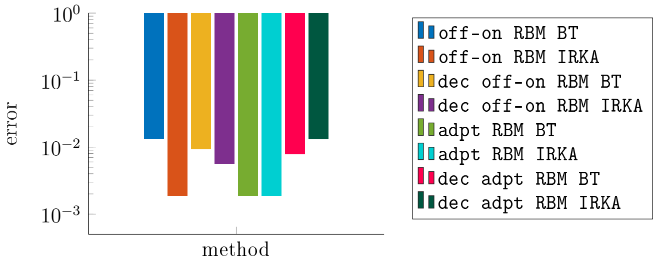


Figure 6.23: Example 9a - Position errors.

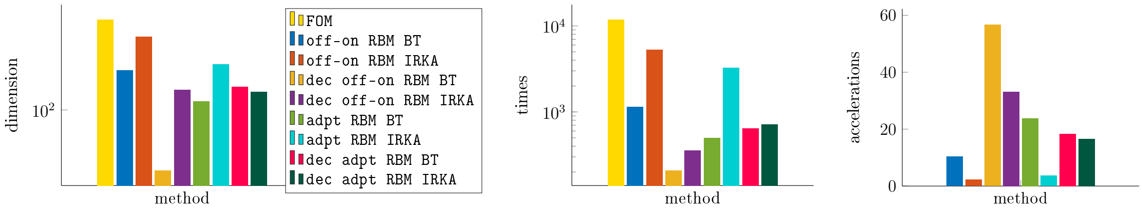


Figure 6.24: Example 9a - Dimensions, times, and acceleration rates.

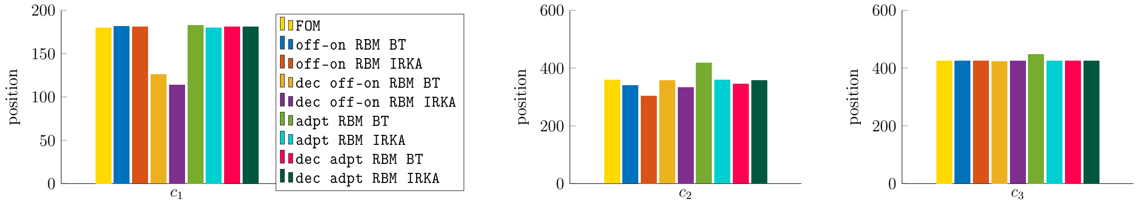


Figure 6.25: Example 9b - Position values.

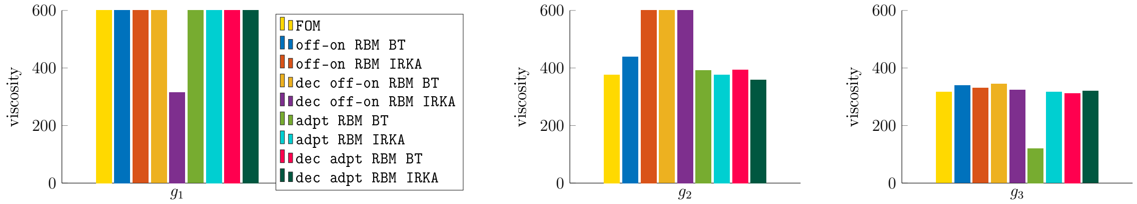


Figure 6.26: Example 9b - Viscosity values.

Example 9b: Position and viscosity optimization Moreover, we optimize the damper's positions and the corresponding gains simultaneously. The initial positions are chosen to be $c_0 = [100,\ 300,\ 500]$, and the initial viscosities are $g_0 = [1000, 1000, 1000]$. The respective positions and gains are depicted in Figure 6.25 and Figure 6.26, respectively. Furthermore, the position and viscosity errors are shown in Figure 6.27. In Figure 6.28, we depict the dimensions, optimization times, and respective acceleration rates corresponding to the full-order and reduced surrogate models. This example shows vividly the limitations of our method. The method (`adpt RBM IRKA`) that approximates the original positions and values so that they coincide with the original ones, requires reduced a dimension of 571 which is almost as large as the original one, and hence, no acceleration is achieved. The method (`dec off-on RBM BT`) leading to the highest acceleration rates of 83 leads to position approximations around 10% and does not approximate the viscosities sufficiently. However, we observe that all methods can give a rough estimation of the optimal positions of the external dampers. Hence, in practice, these could be used to try out different damping position configurations around these positions together with viscosity optimization approaches from this work illustrated in Section 6.3.1 and derived in [106, 140].

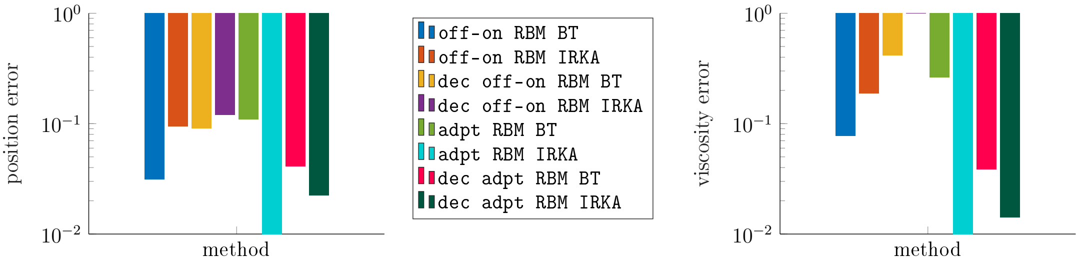


Figure 6.27: Example 9b - Position and viscosity errors.

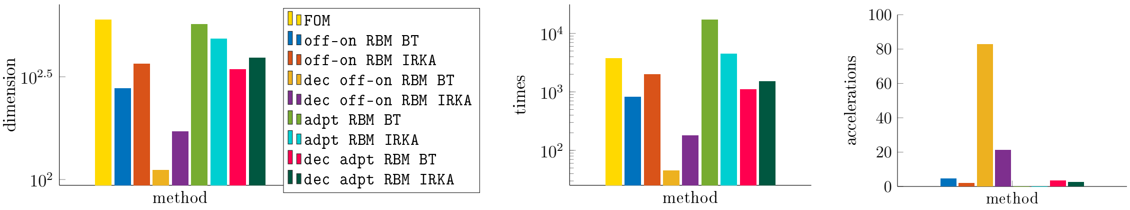


Figure 6.28: Example 9b - Dimensions, times, and acceleration rates.

CHAPTER 7 CONCLUSIONS

Contents

7.1 Summary

In this work, we have considered two main problems. The first has been to reduce various types of inhomogeneous systems, which occur when considering vibrational systems. The aim has been to reduce these systems while considering selected initial values. Therefore, we have extended two approaches from the literature, the multi-system approach and the extended-input approach, to the different system structures. In particular, we have introduced a BT method tailored to an inhomogeneous first-order ODE system with a quadratic output equation and appropriate error bounds. Therefore, we have derived customized observability Gramians and energy expressions that have served as truncation criteria. In addition, we have developed BT methods for inhomogeneous first-order DAE systems with linear and quadratic output equations based on derived Gramians and energy expressions, paying particular attention to the algebraic components of the system. Again, we have derived appropriate error bounds that have served as truncation criteria. We have also introduced a BT method for inhomogeneous second-order systems with linear and quadratic output equations, where the particular focus has been to preserve the system structure. The approach has been based on tailored Gramians and respective energy norms. Also, we have derived appropriate error bounds and have illustrated the efficiency of the methods using various numerical examples.

The second main topic of this work has been the optimization of external dampers based on the reduction of parameter-dependent systems. We have derived RBM schemes tailored to first- and second-order systems arising from vibrational systems with variable external dampers. More detailed, we have used and extended an offline-online scheme, and introduced an adaptive scheme that has been used to build a basis that approximates the respective controllability spaces. This basis has been used to derive reduced system response expressions that have been optimized instead of the original one of large dimensions. Furthermore, we have derived a decomposition of the controllability space, which has led to advantageous computational structures. Moreover, we have derived several error estimators suitable for the different methods that have described the quality of the resulting approximations of the controllability space and the systems response values. The derived RBMs have then been used in the context of damping optimization, where the energy response of the systems has been minimized. In this way, solving the optimization problem has been accelerated significantly, which we have illustrated using different numerical examples.

7.2 Outlook and future research directions

The concepts and methods presented in this manuscript are applicable and extendable to various problems that are out of the scope of this thesis.

For example, in [8], the authors consider the vibration of a plate with tuned vibration dampers added to the system. The methods from Chapter 4 could be applied to reduce systems of similar structures. Also, our methods from Chapter 6 are applicable to optimize the absorbers so that particular frequencies or the maximum response to disturbances are minimized. For these examples, the controllability space decomposition is not applicable as the external attenuators are not of a low-rank structure. Another challenge is that if the Gramians and, hence, the respective controllability spaces are of high numerical rank, good approximations by reduced models could be unfeasible.

A further possible extension of this work concerns the evaluation of inhomogeneous systems in non-standard form investigated in Chapter 3. The authors in [121] introduce a balanced truncation method based on the shift transformation of the respective state for inhomogeneous first-order ODE systems with a linear output equation. This transformation depends on designing parameters that allow some flexibility and the generalization of the multi-system and extended-input approach. Hence, the approach from [121] could be tailored to further inhomogeneous system structures considered in this work to improve the reduction.

Furthermore, the investigation of second-order systems that evaluate not only the displacement but also the velocity as an output component is an interesting research topic for the future. One challenge would be to maintain the second-order structure while taking into account the different initial conditions, which becomes even more challenging when quadratic output equations are used. In the multi-system approach, this would lead to a significant increase in evaluated systems, while in the extended input approach, the derivation of meaningful second-order Gramians that evaluate the displacement and velocity properties is nontrivial. In addition, further work might investigate the IRKA methods for systems with the non-standard forms considered in this thesis. In particular, describing the different observability spaces corresponding to systems with a quadratic output equation is challenging. Considering the state-to-output mappings while building observability space approximations using the IRKA would be an intuitive extension of the IRKA method from [60, 61, 156] and [20].

Moreover, a topic of interest is the extension of model order reduction schemes from this work to second-order systems with a DAE as a state equation. Many approaches have been developed in the literature that deal with second-order DAE systems. However, maintaining a second-order structure while dealing with algebraic equations is still a problem. In particular, we need knowledge about the projecting matrices in the context of DAE systems, which are mostly investigated for the first-order representations of the systems.

BIBLIOGRAPHY

[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, IEEE, 2014, pp. 1075–1079. 7, 14

[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 (2017), pp. 303–318. 25

[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 (2017), pp. 1252–1267. 7, 25

[4] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, vol. 6 of Adv. Des. Control, SIAM Publications, Philadelphia, PA, 2005. 14, 17, 18, 27, 28, 29

[5] 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. 6, 7, 25

[6] A. C. Antoulas, I. V. Gosea, and M. Heinkenschloss, *Data-driven model reduction for a class of semi-explicit DAEs using the Loewner framework*, in Progress in Differential-Algebraic Equations II, T. Reis and A. Ilchmann, eds., Differ.-Algebr. Equ. Forum, Springer, 2020, pp. 185–210. 25

[7] A. C. Antoulas, D. C. Sorensen, and S. Gugercin, *A survey of model reduction methods for large-scale systems*, Contemp. Math., 280 (2001), pp. 193–219. 177

[8] Q. Aumann and S. W. R. Werner, *Structured model order reduction for vibro-acoustic problems using interpolation and balancing methods*, Journal of Sound and Vibration, 543 (2023). 250

[9] Z. Bai, K. Meerbergen, and Y. Su, *Arnoldi methods for structure-preserving dimension reduction of second-order dynamical systems*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. C. Sorensen, eds., vol. 45

of Lect. Notes Comput. Sci. Eng., Springer-Verlag, Berlin/Heidelberg, Germany, 2005, pp. 173–189. 7, 26

[10] Z. Bai and Y.-F. Su, *Dimension reduction of large-scale second-order dynamical systems via a second-order Arnoldi method*, SIAM J. Sci. Comput., 26 (2005), pp. 1692–1709. 7, 26

[11] R. H. Bartels and G. W. Stewart, *Solution of the matrix equation* $AX + XB = C$*: Algorithm 432*, Comm. ACM, 15 (1972), pp. 820–826. 43

[12] K.-J. Bathe, *Finite Element Procesdures*, vol. 2, American Mathematical Society, 2014. 3

[13] U. Baur, P. Benner, and L. Feng, *Model order reduction for linear and nonlinear systems: A system-theoretic perspective*, Arch. Comput. Methods Eng., 21 (2014), pp. 331–358. 7, 52, 129

[14] C. F. Beards, *Structural Vibration: Analysis and Damping*, Arnold, London, 1996. 6

[15] C. Beattie, S. Gugercin, and V. Mehrmann, *Model reduction for systems with inhomogeneous initial conditions*, Systems Control Lett., 99 (2017), pp. 99–106. 7, 52, 54, 63, 91, 106, 128, 129, 130, 131, 141, 182

[16] C. Beattie, S. Gugercin, and Z. Tomljanović, *Sampling-free model reduction of systems with low-rank parameterization*, Adv. Comput. Math., 46 (2020), p. 83. 8

[17] C. A. Beattie and S. Gugercin, *Krylov-based model reduction of second-order systems with proportional damping*, in Proceedings of the 44th IEEE Conference on Decision and Control, Dec. 2005, pp. 2278–2283. 7, 26

[18] ——, *Interpolatory projection methods for structure-preserving model reduction*, Systems Control Lett., 58 (2009), pp. 225–232. 26

[19] P. Benner, P. Goyal, and I. Pontes Duff, *Identification of dominant subspaces for linear structured parametric systems and model reduction*, e-prints 1910.13945, arXiv, 2019. math.NA. 26

[20] ——, *Gramians, energy functionals and balanced truncation for linear dynamical systems with quadratic outputs*, IEEE Trans. Autom. Control, 67 (2021), pp. 886–893. 6, 7, 10, 26, 63, 66, 73, 129, 133, 139, 251

[21] ——, *Data-driven identification of Rayleigh-damped second-order systems*, in Realization and Model Reduction of Dynamical Systems – A Festschrift in Honor of the 70th Birthday of Thanos Antoulas, Springer, 2022. 26

[22] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. H. A. Schilders, and L. M. Silveira, eds., *Model Order Reduction. Volume 1: System- and Data-Driven Methods and Algorithms*, De Gruyter, Berlin, 2021. 7, 25

[23] ——, eds., *Model Order Reduction. Volume 2: Snapshot-Based Methods and Algorithms*, De Gruyter, Berlin, 2021. 7, 25

[24] P. Benner, P. Kürschner, and J. Saak, *Self-generating and efficient shift parameters in ADI methods for large Lyapunov and Sylvester equations*, Electron. Trans. Numer. Anal., 43 (2014), pp. 142–162. 45

[25] P. Benner, P. Kürschner, Z. Tomljanović, and N. Truhar, *Semi-active damping optimization of vibrational systems using the parametric dominant pole algorithm*, Z. Angew. Math. Mech., 96 (2016), pp. 604–619. 6, 8, 18, 214

[26] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, eds., *Model Reduction and Approximation: Theory and Algorithms*, Computational Science & Engineering, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017. 6, 7, 25, 26, 129

[27] P. Benner and E. S. Quintana-Ortí, *Solving stable generalized Lyapunov equations with the matrix sign function*, Numer. Algorithms, 20 (1999), pp. 75–100. 43, 47

[28] P. Benner, E. S. Quintana-Ortí, and G. Quintana-Ortí, *Parallel model reduction of large-scale linear descriptor systems via balanced truncation*, in High Performance Computing for Computational Science - VECPAR 2004, M. Daydé, J. J. Dongarra, V. Hernández, and J. M. L. M. Palma, eds., vol. 3402 of Lecture Notes in Comput. Sci., Berlin/Heidelberg, Germany, 2005, Springer-Verlag, pp. 340–353. 7

[29] P. Benner and J. Saak, *Numerical solution of large and sparse continuous time algebraic matrix Riccati and Lyapunov equations: a state of the art survey*, GAMM Mitteilungen, 36 (2013), pp. 32–52. 26, 43

[30] P. Benner, J. Saak, and M. M. Uddin, *Second order to second order balancing for index-1 vibrational systems*, in 7th International Conference on Electrical & Computer Engineering (ICECE) 2012, IEEE, 2012, pp. 933–936. 14

[31] ——, *Balancing based model reduction for structured index-2 unstable descriptor systems with application to flow control*, Numer. Algebra Control Optim., 6 (2016), pp. 1–20. 7, 19

[32] ——, *Structure preserving model order reduction of large sparse second-order index-1 systems and application to a mechatronics model*, Math. Comput. Model. Dyn. Syst., 22 (2016), pp. 509–523. 14

[33] P. Benner and T. Stykel, *Model order reduction for differential-algebraic equations: A survey*, in Surveys in Differential-Algebraic Equations IV, A. Ilchmann and T. Reis, eds., Differential-Algebraic Equations Forum, Springer International Publishing, Cham, Mar. 2017, pp. 107–160. 7

[34] P. Benner, Z. Tomljanović, and N. Truhar, *Damping optimization for linear vibrating systems using dimension reduction*, in Vibration Problems ICOVP 2011, J. Náprstek, J. Horáček, M. Okrouhlík, B. Marvalová, F. Verhulst, and J. T. Sawicki, eds., vol. 139, Part 5 of Springer Proceedings in Physics, Prag, Czech Republic, 2011, Springer-Verlag, pp. 297–305. 2, 8

[35] ——, *Dimension reduction for damping optimization in linear vibrating systems*, Z. Angew. Math. Mech., 91 (2011), pp. 179–191. 8

[36] ——, *Optimal damping of selected eigenfrequencies using dimension reduction*, Numer. Lin. Alg. Appl., 20 (2013), pp. 1–17. 2, 8

[37] F. Blanchini, D. Casagrande, P. Gardonio, and S. Miani, *Constant and switching gains in semi-active damping of vibrating structures*, Internat. J. Control, 85 (2012), pp. 1886–1897. 8

[38] M. Braun, *Differential Equations and Their Applications*, Springer US, 4 ed., 1993. 53

[39] K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial–Value Problems in Differential–Algebraic Equations*, Elsevier Science Publishing, North-Holland, 1989. 18

[40] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*, Hemisphere Publ. Co., Washington, 1975. 3

[41] S. L. Campbell, *Singular Systems of Differential Equations*, vol. 40 of Research Notes in Mathematics, Pitman Advanced Publishing Program, London, 1980. 18

[42] X. Cao, P. Benner, I. Pontes Duff, and W. Schilders, *Model order reduction for bilinear control systems with inhomogeneous initial conditions*, Internat. J. Control, 94 (2021), pp. 2886–2895. 52

[43] V. Chahlaoui, K. A. Gallivan, A. Vandendorpe, and P. Van Dooren, *Model reduction of second-order systems*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. C. Sorensen, eds., vol. 45 of Lect. Notes Comput. Sci. Eng., Springer-Verlag, Berlin/Heidelberg, Germany, 2005, pp. 149–172. 7, 23, 24, 26, 33

[44] Y. Chahlaoui, D. Lemonnier, A. Vandendorpe, and P. Van Dooren, *Second-order balanced truncation*, Linear Algebra Appl., 415 (2006), pp. 373–384. 7, 23, 24, 25, 26, 33, 34, 35, 109, 168

[45] C. Chicone, *Ordinary Differential Equations with Applications*, Springer, 2nd ed., 2010. 53

[46] J. Denissen, *On Vibrational Analysis and Reduction for Damped Linear Systems*, Dissertation, Otto-von-Guericke-Universität, Magdeburg, Germany, 2019. 47, 48

[47] V. Druskin, L. Knizhnerman, and V. Simoncini, *Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1875–1898. 43

[48] V. Druskin and V. Simoncini, *Adaptive rational Krylov subspaces for large-scale dynamical systems*, Systems Control Lett., 60 (2011), pp. 546–560. 26

[49] C. Du and L. Xie, *Modeling and Control of Vibration in Mechanical Systems*, CRC Press, Boca Raton, 1 ed., 2010. 6

[50] A. Dymarek and T. Dzitkowski, *The use of synthesis methods in position optimisation and selection of tuned mass damper (tmd) parameters for systems with many degrees of freedom*, Archives of Control Sciences, 31(LXVII) (2021), pp. 185–21. 8

[51] G. Flagg, C. Beattie, and S. Gugercin, *Convergence of the iterative rational Krylov algorithm*, Systems Control Lett., 61 (2012), pp. 688–691. 6, 25

[52] P. Freitas and P. Lancaster, *On the optimal spectral abscissa for a system of linear oscillator*, SIAM J. Matrix Anal. Appl., 21 (1999). 8

[53] R. W. Freund, *Padé-type model reduction of second-order and higher-order linear dynamical systems*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. C. Sorensen, eds., vol. 45 of Lect. Notes Comput. Sci. Eng., Springer-Verlag, Berlin/Heidelberg, Germany, 2005, pp. 173–189. 7, 26

[54] F. R. Gantmacher, *Theory of Matrices*, vol. 1, Chelsea Publishing Company, New York, 1959. 49

[55] G. Genta, *Vibration Dynamics and Control*, Mechanical Engineering Series, Springer, 2009. 6, 7

[56] K. Glover, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ norms*, Internat. J. Control, 39 (1984), pp. 1115–1193. 6, 25

[57] I. V. Gosea, C. Poussot-Vassal, and A. C. Antoulas, *On enforcing stability for data-driven reduced-order models*, in 29th Mediterranean Conference on Control and Automation (MED), Virtual, 2021, pp. 487–493. 25

[58] I. V. Gosea, Q. Zhang, and A. C. Antoulas, *Preserving the DAE structure in the Loewner model reduction and identification framework*, Adv. Comput. Math., 46 (2020). 25

[59] E. J. Grimme, *Krylov projection methods for model reduction*, Ph.D. Thesis, Univ. of Illinois at Urbana-Champaign, USA, 1997. 38

[60] 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 (2008), pp. 609–638. 6, 25, 36, 38, 140, 251

[61] S. Gugercin, T. Stykel, and S. Wyatt, *Model reduction of descriptor systems by interpolatory projection methods*, SIAM J. Sci. Comput., 35 (2013), pp. B1010–B1033. 6, 7, 25, 36, 40, 160, 251

[62] M. Gürgöze and P. Müller, *Optimal positioning of dampers in multi-body systems*, J. Sound Vib., 158 (1992), pp. 517–530. 8

[63] B. Haasdonk, M. Dihlmann, and M. Ohlberger, *A training set and multiple basis generation approach for parametrized model reduction based on adaptive grids in parameter space*, Math. Comput. Model. Dyn. Syst., 17 (2011), pp. 423–442. 190

[64] B. Haasdonk and M. Ohlberger, *Adaptive basis enrichment for the reduced basis method applied to finite volume schemes*, in Proc. 5th International Symposium on Finite Volumes for Complex Applications, 2008, pp. 471–478. 186

[65] S. J. Hammarling, *Numerical solution of the stable, non-negative definite Lyapunov equation*, IMA J. Numer. Anal., 2 (1982), pp. 303–323. 43

[66] M. Heinkenschloss, T. Reis, and A. C. Antoulas, *Balanced truncation model reduction for systems with inhomogeneous initial conditions*, Automatica J. IFAC, 47 (2011), pp. 559–564. 7, 52, 54, 59, 60, 63, 71, 128, 129, 139, 141, 182

[67] 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 (2008), pp. 1038–1063. 7, 19, 25

[68] J. S. Hesthaven, G. Rozza, and B. Stamm, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics, Springer, Cham, 2016. 7, 186

[69] D. H. Hodges and G. A. Pierce, *Introduction to Structural Dynamics and Aeroelasticity*. 3

[70] L. R. Hunt, D. A. Linebarger, and R. D. DeGroat, *Realizations of nonlinear systems*, Circuits Syst. Signal Process., 8 (1989), pp. 487–506. 3

[71] D. J. Inman, *Vibration with Control*, John Wiley & Sons Ltd., Virginia Tech, USA, 2006. 6, 7

[72] I. M. Jaimoukha and E. M. Kasenally, *Krylov subspace methods for solving large Lyapunov equations*, SIAM J. Numer. Anal., 31 (1994), pp. 227–251. 43

[73] N. Jakovčević Stor, T. Mitchell, Z. Tomljanović, and M. Ugrica, *Fast optimization of viscosities for frequency-weighted damping of second-order systems*, Z. Angew. Math. Mech., 103 (2023). 6

[74] Y. Kanno, *Damper placement optimization in a shear building model with discrete design variables: a mixed-integer second-order cone programming approach*, Earthquake Engng Struct. Dyn, 42 (2013), pp. 1657–1676. 7

[75] Y. Kanno, M. Puvača, Z. Tomljanović, and N. Truhar, *Optimization of damping positions in a mechanical system*, Rad Hrvat. Akad. Znan. Umjet. Mat. Znan., 23 (2019), pp. 141–157. 8

[76] W. C. Karl, G. C. Verghese, and J. H. Lang, *Control of vibrational systems*, IEEE Trans. Autom. Control, 39 (1994), pp. 222–226. 6

[77] D. T. Kawano, M. Morzfeld, and F. Ma, *The decoupling of second-order linear systems with singular mass matrix*, J. Sound Vib., 332 (2013), pp. 6829–6846. 14

[78] Y. Kawano and J. M. A. Scherpen, *Model reduction by generalized differential balancing*, in Mathematical Control Theory I: Nonlinear and Hybrid Control Systems, vol. 461 of Lect. Notes Control Inf. Sci., 2015, pp. 349–362. 179

[79] P. Kunkel and V. Mehrmann, *Differential-Algebraic Equations: Analysis and Numerical Solution*, Textbooks in Mathematics, EMS Publishing House, Zürich, Switzerland, 2006. 18, 21, 75

[80] P. Kürschner, *Efficient Low-Rank Solution of Large-Scale Matrix Equations*, Dissertation, Otto-von-Guericke-Universität, Magdeburg, Germany, Apr. 2016. 43, 44

[81] I. Kuzmanović, Z. Tomljanović, and N. Truhar, *Optimization of material with modal damping*, Appl. Math. Comput., 218 (2012), pp. 7326–7338. 2

[82] ——, *Damping optimization over the arbitrary time of the excited mechanical system*, J. Comput. Appl. Math., 304 (2016), pp. 120–129. 8

[83] J.-R. Li and J. White, *Low rank solution of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 260–280. 43

[84] 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 (2007), pp. 293–304. 6

[85] Z. Liu, B. Rao, and Q. Zhang, *Polynomial stability of the Rao-Nakra beam with a single internal viscous damping*, Journal of Differential Equations, 269 (2020), pp. 6125–6162. 6

[86] P. Losse and V. Mehrmann, *Controllability and observability of second order descriptor systems*, SIAM J. Control Optim., 47 (2008), pp. 1351–1379. 14

[87] A. Lu and E. L. Wachspress, *Solution of Lyapunov equations by alternating direction implicit iteration.*, Comput. Math. Appl., 21 (1991), pp. 43–58. 43

[88] D. G. Luenberger, *Introduction to dynamic systems. Theory, models, and applications.*, John Wiley & Sons., New York etc., first ed., 1979. 14

[89] R. März, *The index of linear differential algebraic equations with properly stated leading terms*, Results in Mathematics, 42 (2002), pp. 308–338. 19

[90] A. J. Mayo and A. C. Antoulas, *A framework for the solution of the generalized realization problem*, Linear Algebra Appl., 425 (2007), pp. 634–662. Special Issue in honor of P. A. Fuhrmann, Edited by A. C. Antoulas, U. Helmke, J. Rosenthal, V. Vinnikov, and E. Zerz. 25

[91] V. Mehrmann and T. Stykel, *Balanced truncation model reduction for large-scale systems in descriptor form*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. C. Sorensen, eds., vol. 45 of Lect. Notes Comput.

Sci. Eng., Springer-Verlag, Berlin/Heidelberg, Germany, 2005, pp. 83–115. 7, 21, 25, 26, 29, 31, 75, 163, 166

[92] D. G. Meyer and S. Srinivasan, *Balancing and model reduction for second-order form linear systems*, IEEE Trans. Autom. Control, 41 (1996), pp. 1632–1644. 7, 26

[93] B. C. Moore, *Principal component analysis in linear systems: controllability, observability, and model reduction*, IEEE Trans. Autom. Control, AC–26 (1981), pp. 17–32. 6, 25, 26, 129

[94] T. Moser and B. Lohmann, *A Rosenbrock framework for tangential interpolation of port-Hamiltonian descriptor systems*, 2023. 25

[95] P. C. Müller and W. O. Schiehlen, *Linear vibrations: A theoretical treatment of multi-degree-of-freedom vibrating systems*, Martinus Hijhoff publishers, 1985. 7

[96] P. C. Müller and W. E. Weber, *Modal analysis and insights into damping phenomena of a special vibration chain*, Arch. Appl. Mech., 91 (2021), pp. 2179–2187. 6

[97] I. Nakić, *Optimal damping of vibrational systems*, Dissertation, FernUniversität in Hagen, Hagen, Germany, 2003. 8

[98] NICONET Society, *Slicot basic systems and control toolbox*, 2005. See `http://www.slicot.org/`. 177

[99] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, Springer-Verlag, 2003. 3

[100] G. Pandit and S. Gupta, *Structural Analysis: A Matrix Approach*, vol. 2, McGraw Hill Education, 2008. 3

[101] T. Penzl, *A cyclic low rank Smith method for large, sparse Lyapunov equations with applications in model reduction and optimal control*, Tech. Rep. SFB393/98-6, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, FRG, 1998. Available from `http://www.tu-chemnitz.de/sfb393/sfb98pr.html`. 43

[102] ——, *A cyclic low rank Smith method for large sparse Lyapunov equations*, SIAM J. Sci. Comput., 21 (2000), pp. 1401–1418. 44

[103] J. Przybilla, *Semi-active damping optimization of vibrational systems using the reduced basis method*, Jan. 2024. https://doi.org/10.5281/zenodo.10462562. 229

[104] J. Przybilla, I. P. Duff, and P. Benner, *Model reduction for systems in non-standard form*, Jan. 2024. https://doi.org/10.5281/zenodo.10462594. 161

[105] J. Przybilla, I. Pontes Duff, and P. Benner, *Balanced truncation of descriptor systems with a quadratic output*, e-print 2402.14716, arXiv, 2024. math.DS. 9, 10

[106] ——, *Model reduction for second-order systems with inhomogeneous initial conditions*, Systems Control Lett., 183 (2024). 9, 10, 53, 246

[107] ——, *Semi-active damping optimization of vibrational systems using the reduced basis method*, Adv. Comput. Math., 50 (2024). 9, 10, 18, 228

[108] J. Przybilla and M. Voigt, *Model reduction of parametric differential-algebraic systems by balanced truncation*, Math. Comput. Model. Dyn. Syst., 30 (2024). 7, 186, 187

[109] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced Basis Methods for Partial Differential Equations*, vol. 92 of La Matematica per il 3+2, Springer International Publishing, 2016. ISBN: 978-3-319-15430-5. 7, 186

[110] M. Redmann and I. Pontes Duff, *Model order reduction for bilinear systems with non-zero initial states – different approaches with error bounds*, Internat. J. Control, 96 (2022), pp. 1491–1504. 52

[111] T. Reis and T. Stykel, *Stability analysis and model order reduction of coupled systems*, Math. Comput. Model. Dyn. Syst., 13 (2007), pp. 413–436. 168

[112] T. Reis and T. Stykel, *Balanced truncation model reduction of second-order systems*, Math. Comput. Model. Dyn. Syst., 14 (2008), pp. 391–406. 7, 23, 24, 26, 33, 35, 109, 168

[113] J. D. Roberts, *Linear model reduction and solution of the algebraic Riccati equation by use of the sign function*, Internat. J. Control, 32 (1980), pp. 677–687. (Reprint of Technical Report No. TR-13, CUED/B-Control, Cambridge University, Engineering Department, 1971). 48

[114] J. Saak, M. Köhler, and P. Benner, *M-M.E.S.S.-2.1 – The Matrix Equations Sparse Solvers library*, Apr. 2021. see also:`https://www.mpi-magdeburg.mpg.de/projects/mess`. 46, 161, 166, 229

[115] J. Saak, D. Siebelts, and S. W. R. Werner, *A comparison of second-order model order reduction methods for an artificial fishtail*, at-Automatisierungstechnik, 67 (2019), pp. 648–667. 26

[116] 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 (2018), pp. 661–666. 7, 19

[117] B. Salimbahrami, *Structure Preserving Order Reduction of Large Scale Second Order Models*, Dissertation, Technische Universität München, Munich, Germany, 2005. 7, 26

[118] B. Salimbahrami and B. Lohmann, *Order reduction of large scale second-order systems using Krylov subspace methods*, Linear Algebra Appl., 415 (2006), pp. 385–405. 26

[119] A. Schmidt and B. Haasdonk, *Reduced basis approximation of large scale algebraic Riccati equations*, tech. rep., University of Stuttgart, 2015. 7, 186

[120] A. Schmidt and B. Wittwar, D. Haasdonk, *Rigorous and effective a-posteriori error bounds for nonlinear problems—application to RB methods*, Adv. Comput. Math., 46 (2020). 26, 190

[121] C. Schröder and M. Voigt, *Balanced truncation model reduction with a priori error bounds for LTI systems with nonzero initial value*, IEEE Trans. Magn., 420 (2023). 7, 52, 129, 250

[122] P. Schulze, B. Unger, C. A. Beattie, and S. Gugercin, *Data-driven structured realization*, Linear Algebra Appl., 537 (2018), pp. 250–286. 26

[123] V. Simoncini, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comput., 29 (2007), pp. 1268–1288. 43

[124] V. Simoncini, *Computational methods for linear matrix equations*, SIAM Rev., 38 (2016), pp. 377–441. 43

[125] V. Simoncini and V. Druskin, *Convergence analysis of projection methods for the numerical solution of large Lyapunov equations*, SIAM J. Numer. Anal., 47 (2009), pp. 828–843. 43

[126] N. T. Son and T. Stykel, *Solving parameter-dependent Lyapunov equations using the reduced basis method with application to parametric model order reduction*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 478–504. 7, 9, 26, 186, 187, 188, 190

[127] E. D. Sontag, *Mathematical Control Theory*, Texts in Applied Mathematics, Springer-Verlag, New York, NY, 2nd ed., 1998. 14

[128] D. C. Sorensen, *Passivity preserving model reduction via interpolation of spectral zeros*, Systems Control Lett., 54 (2005), pp. 347–360. 26

[129] T. Stykel, *Analysis and Numerical Solution of Generalized Lyapunov Equations*, Dissertation, TU Berlin, 2002. 20, 22

[130] ——, *Gramian-based model reduction for descriptor systems*, Math. Control Signals Systems, 16 (2004), pp. 297–319. 7, 20, 21, 25, 26, 29, 75

[131] ——, *Balanced truncation model reduction for semidiscretized Stokes equation*, Linear Algebra Appl., 415 (2006), pp. 262–289. 7, 25, 163, 164, 166

[132] ——, *A modified matrix sign function method for projected Lyapunov equations*, Systems Control Lett., 56 (2007). 49

[133] ——, *Low-rank iterative methods for projected generalized Lyapunov equations*, Electron. Trans. Numer. Anal., 30 (2008), pp. 187–202. 7, 19, 46, 47, 164, 166

[134] T.-J. Su and R. R. Craig Jr., *Model reduction and control of flexible structures using Krylov vectors*, Journal of Guidance, Control, and Dynamics, 14 (1991), pp. 260–267. 7

[135] I. Takewaki, *Optimal damper placement for minimum transfer functions*, Earthquake Engng Struct. Dyn., 26 (1997), pp. 1113–1997. 7

[136] ——, *Optimal damper positioning in beams for minimum dynamic compliance*, Comp. Meth. Appl. Mech. Eng., 156 (1998), pp. 363–373. 8

[137] The MORwiki Community, *MORwiki - Model Order Reduction Wiki*. http://modelreduction.org. 179

[138] M. S. Tombs and I. Postlethwaite, *Truncated balanced realization of a stable non-minimal state-space system*, Internat. J. Control, 46 (1987), pp. 1319–1330. 6, 25, 26, 129

[139] Z. Tomljanović, *Optimal damping for vibrating systems using dimension reduction*, PhD thesis, Josip Juraj Strossmayer University of Osijek, 2011. 8

[140] Z. Tomljanović, C. Beattie, and S. Gugercin, *Damping optimization of parameter dependent mechanical systems by rational interpolation*, Adv. Comput. Math., 44 (2018), pp. 1797–1820. 6, 8, 10, 18, 36, 41, 42, 177, 186, 202, 214, 228, 229, 230, 246

[141] Z. Tomljanović and M. Voigt, *Semi-active $\mathcal{H}_\infty$ damping optimization by adaptive interpolation*, Numer. Lin. Alg. Appl., 27 (2020), p. e2300. 8

[142] N. Truhar, *An efficient algorithm for damper optimization for linear vibrating systems using Lyapunov equation*, J. Comput. Appl. Math., 127 (2004), pp. 169–182. 8

[143] N. Truhar, Z. Tomljanović, and M. Puvača, *Approximation of damped quadratic eigenvalue problem by dimension reduction*, J. Appl. Math. Comput., 347 (2017), pp. 40–53. 8

[144] ——, *An efficient approximation for optimal damping in mechanical systems*, Internat. J. Numer. Anal. Mod., 14 (2017), pp. 201–217. 8

[145] N. Truhar, Z. Tomljanović, and K. Veselić, *Damping optimization in mechanical systems with external force*, Appl. Math. Comput., 250 (2015), pp. 270–279. 8

[146] N. Truhar and K. Veselić, *Bounds on the trace of a solution to the Lyapunov equation with a general stable matrix*, Systems Control Lett., 56 (2007), pp. 493–503. 186

[147] ——, *An efficient method for estimating the optimal dampers' viscosity for linear vibrating systems using Lyapunov equation*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 18–39. 8, 186

[148] N. Truhar and K. Veselić, *An efficient method for estimating the optimal dampers' viscosity for linear vibrating systems using Lyapunov equation*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 18–39. 8

[149] A. Vandendorpe and P. Van Dooren, *Krylov techniques for model reduction of second-order systems*, tech. rep., Université. 7

[150] K. Veroy and A. T. Patera, *Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: rigorous reduced-basis a posteriori error bounds*, Internat. J. Numer. Methods Fluids, 47 (2005), pp. 773–788. 7, 186

[151] K. Veroy, C. Prud'homme, and A. T. Patera, *Reduced-basis approximation of the viscous Burgers equation: rigorous a posteriori error bounds*, Comptes Rendus Mathematique, 337 (2003), pp. 619–624.

[152] K. Veroy, C. Prud'Homme, D. V. Rovas, and A. T. Patera, *A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations*, in 16th AIAA Computational Fluid Dynamics Conference, Orlando, United States, 2003. 7, 186

[153] K. Veselić, *Damped oscillations of linear systems*, vol. 2023 of Lecture Notes in Math., Springer-Verlag, 2011. 6, 7

[154] D. C. Villemagne and R. E. Skelton, *Model reduction using a projection formulation*, Internat. J. Control, 46 (1987), pp. 2141–2169. 38

[155] E. L. Wilson, *Static and Dynamic Analysis of Structures*, Berkeley, CA: Computers and Structures, 4 ed., 2004. 2

[156] S. Wyatt, *Issues in Interpolatory Model Reduction: Inexact Solves, Second-order Systems and DAEs*, PhD thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA, May 2012. 40, 41, 251

[157] Z. T. Y. Kanno, M. Puvača and N. Truhar, *Optimization of damping positions in a mechanical system*, Rad Hazu. Matemtičke Znanosti, 23 (2019), pp. 141–157. 238

[158] E. C. Zachmanoglou and D. W., *Ordinary Differential Equations with Applications*, Dover Publications, 1987. 53

[159] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice-Hall, Upper Saddle River, NJ, 1996. 14, 214

STATEMENT OF SCIENTIFIC COOPERATIONS

This work is based on articles and reports (published and unpublished) that have been obtained in cooperation with various coauthors. To guarantee a fair assessment of this thesis, this statement clarifies the contributions that each individual coauthor has made. The following people contributed to the content of this work:

- Peter Benner: Provision of study material, materials and IT resources; Acquisition of funding for the projects leading to this thesis; Supervisory and managerial responsibility for the research activity; Introduction of ideas for parts of the thesis; Preparation of the published work, in particular critical review of the included papers and thesis, commentary and revision - including the pre- and post-publication phases.

- Igor Pontes Duff: Supervision- and leadership responsibility for the research activity planning and execution, including mentorship; Introduction of ideas for parts of the thesis; Critical review of the included papers and the thesis, commentary, and revision – including pre-and postpublication stages; Implementation, verification, and mostly reviewing of the computer code and supporting algorithms; testing of existing code components; Application of mathematical and computational, or other formal techniques.

- Ninoslav Truhar: Contribution to the error indicator corresponding to the decoupled controllability space, critically reviewing and revising the corresponding texts; Application of mathematical and computational techniques.

- Matea Ugrica: Preparation and editing of the published paper, in particular critical review, commentary or revision of the position optimization sections. Application of mathematical and computational techniques; Programming, implementation of computer code and supporting algorithms, testing of existing code components.

- Pawan Goyal: Preparation of the published work, specifically critical review and revision.

EHRENERKLÄRUNG

Ich versichere hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; verwendete fremde und eigene Quellen sind als solche kenntlich gemacht.

Ich habe insbesondere nicht wissentlich:

- Ergebnisse erfunden oder widersprüchliche Ergebnisse verschwiegen,
- statistische Verfahren absichtlich missbraucht, um Daten in ungerechtfertigter Weise zu interpretieren,
- fremde Ergebnisse oder Veröffentlichungen plagiiert oder verzerrt wiedergegeben.

Mir ist bekannt, dass Verstöße gegen das Urheberrecht Unterlassungs- und Schadenersatzansprüche des Urhebers sowie eine strafrechtliche Ahndung durch die Strafverfolgungsbehörden begründen kann.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

Berlin, 23.02.2024

Jennifer Przybilla