



MAX-PLANCK-GESELLSCHAFT

**Max Planck Institute Magdeburg
Preprints**

Peter Benner Jens Saak M. Monir Uddin

**Balancing based model reduction for
structured index-2 unstable descriptor
systems with application to flow control**



Abstract

Stabilizing a flow around an unstable equilibrium is a typical problem in flow control. Model-based designed of modern controllers like LQR/LQG or H_∞ compensators is often limited by the large-scale of the discretized flow models. Therefore, model reduction is usually needed before designing such a controller. Here we suggest an approach based on applying balanced truncation for unstable systems to the linearized flow equations usually used for compensator design. For this purpose, we modify the ADI iteration for Lyapunov equations to deal with the index-2 structure of the underlying descriptor system efficiently in an implicit way. The resulting algorithm is tested for model reduction and control design of a linearized Navier-Stokes system describing *von Kármán vortex shedding*.

Keywords: Flow control, model reduction, balanced truncation, descriptor system, index-2, ADI iteration, von Kármán vortex shedding.

Imprint:

Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg

Publisher:

Max Planck Institute for
Dynamics of Complex Technical Systems

Address:

Max Planck Institute for
Dynamics of Complex Technical Systems
Sandtorstr. 1
39106 Magdeburg

<http://www.mpi-magdeburg.mpg.de/preprints/>

1 Introduction

Stokes, Navier-Stokes, or Oseen equations play an important role in describing various problems in fluid dynamics and engineering applications. The spatial discretization of such equations using finite difference or finite element methods produces large-scale structured index-2 descriptor systems [21, 4]. In case of controller design, simulation and design optimization, working with these large-scale systems is often problematic due to computational complexity and storage requirements. The idea of model order reduction (MOR) [2, 11, 25] is to approximate a large scale dynamical system by a substantially lower dimensional system which has nearly the same input-output behavior. This lower dimensional model then, e.g. serves as the basis for feedback control design. The system theoretic method balanced truncation (BT) [2, 11] is a particular technique which preserves important properties of the system (such as asymptotic stability) and features an a priori bound on the approximation error.

The theory of BT for large-scale *asymptotically stable* descriptor systems has been developed first by Stykel in [27, 28]. There the author discusses the general framework of the BT method for a descriptor system. In principle, the proposed method is based on splitting the descriptor system into proper and improper subsystems corresponding to the deflating subspaces of the associated matrix pencil with respect to finite and infinite eigenvalues, and on then reducing only the order of the proper subsystem. This approach requires the availability of spectral projectors onto the respective subspaces. Recently, a method for BT of structured large-scale descriptor systems of index 2 has been developed [19] that avoids the computation of spectral projectors. Instead it implicitly performs an index reduction by projection to the inherent or hidden manifold on which the solution evolves.

In this paper, we study the BT based model reduction of an *unstable* structured index-2 descriptor system. In principle, one can apply the above BT technique to this model by stabilizing the system first using a proper stabilizing feedback matrix (SFM) and then following the approach in [19]. Picking up the ideas in [31] we apply the truncating transformations computed with respect to the stabilized system to the original unstable system. To compute the controllability and observability Gramian factors (which are the main ingredients in determining the balancing transformations), we need to solve two projected algebraic Lyapunov equations of the stabilized system. Again following [31] we employ Bernoulli stabilization to derive the SFM. The main advantage of the Bernoulli stabilization is that it only changes the anti-stable eigenvalues of the system. Thus the required Bernoulli equation can be restricted to these and is of the same dimension as the corresponding eigenspaces.

The main computational bottleneck in BT methods is the efficient solution of large-scale Lyapunov equations, where here, the additional difficulty of the involved (implicit) spectral projection has to be treated. Compared to the presentation in [19], this paper also presents an updated version of the low-rank Cholesky factor-alternating direction implicit (LRCF-ADI) algorithm to solve these projected Lyapunov equations. Here, we call this approach the generalized sparse (GS-)LRCF-ADI method. In order to ensure fast convergence of GS-LRCF-ADI, proper selection of the ADI shift parameters is crucial. In Section 6.2 we also discuss and resolve the difficulties in computing

shift parameters for the models of flow control considered here.

Stabilization of the non-stationary incompressible Navier-Stokes equations around a stationary solution using a Riccati-based feedback has received considerable attention regarding theory as well as numerical methods during the recent years. In the Riccati-based boundary feedback stabilization technique [5], the most challenging task is to solve the corresponding algebraic Riccati equation (ARE) for the high dimensional model. In this paper we numerically investigate the potential of a feedback control for the original model computed using a reduced-order model. In direct comparison this can be computationally much cheaper if the number of Newton steps for solving the Riccati equation is larger than 2.

The proposed method is applied to data for a spatially FEM semi-discretized linearized Navier-Stokes model. Numerical results are discussed for both the BT model reduction as well as the reduced-order model based stabilization.

This article is organized as follows. In Section 2, we describe the model which is considered for our numerical test. The fundamental idea of BT for both stable and unstable systems and the LRCF-ADI methods are repeated briefly in Section 3. Section 4 collocates the ideas for converting the index-2 system into an ODE system applying the projector onto the hidden manifold together with the properties of the projector. In Section 5 we discuss BT based MOR techniques for structured index-2 unstable systems. Section 6 consists of the GS-LRCF-ADI algorithm for the implicit solution of the projected generalized algebraic Lyapunov equations and also the techniques for selecting ADI shift parameters for models of the given structure. In the subsequent section we illustrate numerical results, obtained by applying our approaches to the model introduced in Section 2 with different spatial resolutions and consequently different matrix dimensions.

2 Model example

In his seminal book [26] Sontag, among others, presents the linearization principle. That principle basically states that a general nonlinear model can be stabilized by a linear quadratic regulator (LQR) for a linearization of itself in the vicinity of the linearization point. The basic idea is that if the regulator is working properly, the vicinity where the linearization is a proper approximation of the nonlinear system is never left. This principle has been exploited by the authors in [5] for a Navier-Stokes model for the *von Kármán vortex street*. The linearized Navier-Stokes equations that arose there and that we consider in this paper are

$$\begin{aligned} \frac{\partial}{\partial t} \vec{v} - \frac{1}{Re} \Delta \vec{v} + (\vec{w} \cdot \nabla) \vec{v} + (\vec{v} \cdot \nabla) \vec{w} + \nabla p &= 0, \\ \nabla \cdot \vec{v} &= 0, \end{aligned} \tag{1}$$

where \vec{v}, \vec{w} denote velocity vectors, p the pressure and Re is the Reynolds number. The vector \vec{w} represents the stationary solution of the incompressible nonlinear Navier-Stokes equations and \vec{v} is the deviation of the original state from the stationary solu-

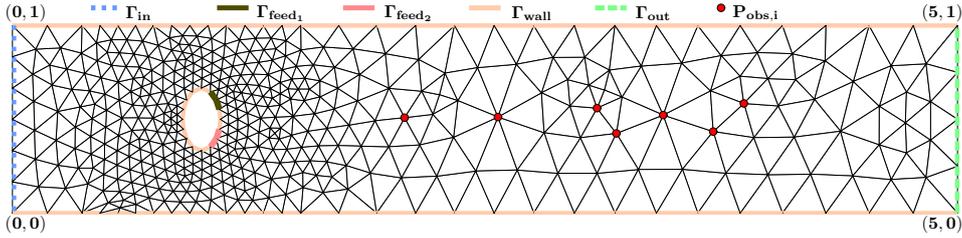


Figure 1: Initial discretization of the *von Kármán vortex street* with coordinates, boundary parts and observation points (source [5]).

tion. The boundary and initial conditions, as well as the derivation of this model, are given in [5]. There the authors apply a mixed finite element method based on the well known Taylor-Hood finite elements [20] to discretize equation (1). The coarsest discretization of the domain for the *von Kármán vortex street* example from [5] is shown in Figure 1. This yields the differential-algebraic equations

$$E_1 \dot{v}(t) = A_1 v(t) + A_2 p(t) + B_1 u(t), \quad (2a)$$

$$A_2^T v(t) = 0, \quad (2b)$$

where $v(t) \in \mathbb{R}^{n_1}$ denotes the nodal vector of discretized velocity deviations, $p(t) \in \mathbb{R}^{n_2}$ the discretized pressure, $u(t) \in \mathbb{R}^m$ are the inputs, and $E_1, A_1 \in \mathbb{R}^{n_1 \times n_1}$, $A_2 \in \mathbb{R}^{n_1 \times n_2}$, $B_1 \in \mathbb{R}^{n_1 \times m}$ are all sparse matrices.

Additionally, the vertical velocities in the observation nodes depicted in Figure 1 in the domain are modeled by the output equation

$$y(t) = C_1 v(t), \quad (3)$$

with the output $y(t) \in \mathbb{R}^p$ and the output matrix $C_1 \in \mathbb{R}^{p \times n_1}$.

The above system remains stable, i.e., the finite spectrum of the *matrix pencil*

$$\mathcal{P}(\lambda) = \lambda \begin{bmatrix} E_1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A_1 & A_2 \\ A_2^T & 0 \end{bmatrix} \quad (4)$$

is located in the negative half plane \mathbb{C}^- , as long as the Reynolds number Re is small. However, already for moderate Reynolds numbers (e.g., in the configuration of Figure 1 $Re \geq 300$) a few finite eigenvalues move to the positive half plane, \mathbb{C}^+ [1]. The key problem in the LQR approach for the model under investigation is to compute the boundary feedback stabilization matrix K_1 (see e.g., [5]), such that the stabilized system has the following form:

$$\begin{aligned} E_1 \dot{v}(t) &= (A_1 - B_1 K_1) v(t) + A_2 p(t) + B_1 u(t), \\ A_2^T v(t) &= 0. \end{aligned} \quad (5)$$

The authors in [5] presented an algorithm (see [5, Algorithm 2]) to compute K_1 which is based on the standard linear-quadratic regulator approach [26, 14] for a projected

semidiscretized state-space system. The most challenging part in this algorithm is to solve the usually very large, generalized, projected algebraic Riccati equation (GARE) based on the full order semidiscretized model.

In this paper our aim is to replace this full order system (2) together with its associated output equation (3) by a substantially lower dimensional model

$$\begin{aligned}\hat{E}\dot{\hat{v}}(t) &= \hat{A}\hat{v}(t) + \hat{B}u(t), \\ \hat{y}(t) &= \hat{C}\hat{v}(t),\end{aligned}\tag{6}$$

where the responses (both in frequency and time domain) of the reduced-order model are good approximations to those of the full model. Note that for balanced truncation, we can always guarantee that $\hat{E} = I$, which we thus assume in the following. We employ the reduced-order model to compute an approximation to the optimal LQR feedback matrix of the full system. The main advantage of this approach is that we only need to solve two projected algebraic Lyapunov equations in order to derive the reduced-order model instead of one Lyapunov equation per Newton step in the solver for the GARE, which are usually many more [12].

Based on the reduced model (6) the GARE

$$\hat{A}^T \hat{X} + \hat{X} \hat{A} - \hat{X} \hat{B} \hat{B}^T \hat{X} = -\hat{C}^T \hat{C}\tag{7}$$

is now much smaller in dimension. It can thus easily be solved for \hat{X} using classical solvers as, e.g., the MATLAB `care` command. The stabilizing feedback matrix for the reduced model (6) then is

$$\hat{K} = \hat{B}^T \hat{X}.\tag{8}$$

The ROM-based approximation to the stabilizing feedback matrix for the full order model can now be retrieved as

$$K_1 = \hat{B}^T \hat{X} T_L^T E_1 = \hat{K} T_L^T E_1$$

where T_L, T_R are the balancing transformations (see Section 5) used to compute the reduced-order model.

3 Background

In this section we briefly review the basic idea of standard balanced truncation for both stable and unstable systems. We also repeat the basics for the LRCF-ADI method for computing the required Gramian factors in large scale settings, to put our new contributions for the index-2 case in Algorithm 2 (see Section 6) into perspective.

3.1 Standard BT and Low-Rank-ADI

Consider a generalized linear time-invariant (LTI) continuous time system

$$\mathcal{E}\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t), \quad y(t) = \mathcal{C}x(t) + \mathcal{D}u(t),\tag{9}$$

in which $\mathcal{E} \in \mathbb{R}^{n \times n}$ is non-singular, and $\mathcal{A} \in \mathbb{R}^{n \times n}$, $\mathcal{B} \in \mathbb{R}^{n \times p}$, $\mathcal{C} \in \mathbb{R}^{m \times n}$ and $\mathcal{D} \in \mathbb{R}^{m \times p}$. Assume that the original system is asymptotically stable, i.e., all its eigenvalues lie in \mathbb{C}^- . Then the controllability Gramian $\mathcal{P} \in \mathbb{R}^{n \times n}$ and the observability Gramian $\mathcal{Q} \in \mathbb{R}^{n \times n}$, which are the solutions of the two Lyapunov equations (see, e.g., [2])

$$\mathcal{A}\mathcal{P}\mathcal{E}^T + \mathcal{E}\mathcal{P}\mathcal{A}^T = -\mathcal{B}\mathcal{B}^T \quad \text{and} \quad \mathcal{A}^T\mathcal{Q}\mathcal{E} + \mathcal{E}^T\mathcal{Q}\mathcal{A} = -\mathcal{C}^T\mathcal{C},$$

respectively, are symmetric and positive semi-definite, and therefore have symmetric decompositions

$$\mathcal{P} = \mathcal{R}\mathcal{R}^T \quad \text{and} \quad \mathcal{Q} = \mathcal{L}\mathcal{L}^T. \quad (10)$$

The balancing transformation can be formed using the SVD

$$\mathcal{R}^T\mathcal{E}\mathcal{L} = \mathcal{U}\mathcal{S}\mathcal{V}^T = \begin{bmatrix} \mathcal{U}_1 & \mathcal{U}_2 \end{bmatrix} \begin{bmatrix} \mathcal{S}_1 & \\ & \mathcal{S}_2 \end{bmatrix} \begin{bmatrix} \mathcal{V}_1^T \\ \mathcal{V}_2^T \end{bmatrix},$$

and defining

$$\mathcal{T}_L := \mathcal{R}\mathcal{U}_1\mathcal{S}_1^{-\frac{1}{2}}, \quad \mathcal{T}_R := \mathcal{L}\mathcal{V}_1\mathcal{S}_1^{-\frac{1}{2}}. \quad (11)$$

Here, \mathcal{U}_1 and \mathcal{V}_1 are composed of the leading k columns of \mathcal{U} and \mathcal{V} , respectively, \mathcal{S}_1 is the first $k \times k$ block of the matrix $\mathcal{S} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k, \dots, \sigma_n)$, where σ_i , $i = 1, 2, \dots, n$ are the Hankel singular values of the system (9) [2, 16]. Now the reduced-order model (ROM) is derived as

$$\hat{\mathcal{E}}\hat{\dot{x}}(t) = \hat{\mathcal{A}}\hat{x}(t) + \hat{\mathcal{B}}u(t), \quad \hat{y}(t) = \hat{\mathcal{C}}\hat{x}(t) + \mathcal{D}u(t), \quad (12)$$

where

$$\hat{\mathcal{E}} = \mathcal{T}_L^T\mathcal{E}\mathcal{T}_R = I_k, \quad \hat{\mathcal{A}} = \mathcal{T}_L^T\mathcal{A}\mathcal{T}_R, \quad \hat{\mathcal{B}} = \mathcal{T}_L^T\mathcal{B}, \quad \text{and} \quad \hat{\mathcal{C}} = \mathcal{C}\mathcal{T}_R. \quad (13)$$

The ROM obtained in this way satisfies the global error bound [2, 16]

$$\|\mathbb{G}(\cdot) - \hat{\mathbb{G}}(\cdot)\|_{\mathcal{H}^\infty} \leq 2 \sum_{i=k+1}^n \sigma_i, \quad (14)$$

where $\mathbb{G}(s) := \mathcal{C}(s\mathcal{E} - \mathcal{A})^{-1}\mathcal{B}$ and $\hat{\mathbb{G}}(s) := \hat{\mathcal{C}}(s\hat{\mathcal{E}} - \hat{\mathcal{A}})^{-1}\hat{\mathcal{B}}$ are called the transfer functions [2] of the full and reduced model, respectively, and $\|\cdot\|_{\mathcal{H}^\infty}$ denotes the \mathcal{H}^∞ -norm. The relation (14) is indeed an *a priori* error bound. Hence, one can easily determine the required dimension of the ROM for a given error tolerance. In [29] this procedure for the BT method is summarized as the *square-root* (SR) algorithm.

If the numbers of inputs and outputs are very small compared to the number of degrees of freedom (DoFs), then usually the Gramians \mathcal{P} and \mathcal{Q} can be expressed in low-rank factored form [23, 3, 18], i.e., \mathcal{R} and \mathcal{L} are thin rectangular matrices. They can for example, be computed by the LRCF-ADI iteration [23, 22, 10]. Recently, the formulation of the LRCF-ADI method has been updated in [7], such that real Gramian factors can be computed by clever handling of complex shift parameters. Moreover

Algorithm 1 LRCF-ADI to solve $FX\tilde{E}^T + \tilde{E}XF^T = -NN^T$

Input: \tilde{E}, F, N , ADI shift parameters $\{\mu_1, \mu_2, \dots, \mu_J\}$.

Output: Z such that $X \approx ZZ^T$

```

1:  $Z = []$ 
2:  $W_0 = N$ 
3:  $i = 1$ 
4: while ( $\|W_{i-1}^T W_{i-1}\| > \text{tol}$  or  $i \leq i_{\max}$ ) do
5:   Solve  $(F + \mu_i \tilde{E})V_i = W_{i-1}$  for  $V_i$ 
6:   if  $\text{Im}(\mu_i) = 0$  then
7:      $Z = [Z \quad \sqrt{-2 \text{Re}(\mu_i)} V_i]$ 
8:      $W_i = W_{i-1} - 2 \text{Re}(\mu_i) \tilde{E} V_i$ 
9:   else
10:     $\gamma = 2\sqrt{-\text{Re}(\mu_i)}, \beta = \frac{\text{Re}(\mu_i)}{\text{Im}(\mu_i)}$ 
11:    Update low-rank solution factor
12:     $Z = [Z \quad \gamma(\text{Re}(V_i) + \beta \text{Im}(V_i)) \quad \gamma\sqrt{(\beta^2 + 1)} \cdot \text{Im}(V_i)]$ 
13:     $W_i = W_{i-1} - 4 \text{Re}(\mu_i) \tilde{E}(\bar{V}_i + 2\beta \text{Im}(V_i))$ 
14:     $i = i + 1$ 
15:   end if
16: end while

```

the same authors [8] introduced an efficient technique to compute a low-rank factor of the ADI residual as an integral part of each iteration step and thus cheaply evaluate stopping criteria for the algorithm.

Combining the ideas from [7] and [8] the LRCF-ADI is summarized in Algorithm 1. In this Algorithm either $(\tilde{E}, F, N) = (\mathcal{E}, \mathcal{A}, \mathcal{B})$ is applied to generate $Z = \mathcal{R}$, or $(\tilde{E}, F, N) = (\mathcal{E}^T, \mathcal{A}^T, \mathcal{C}^T)$ is used for $Z = \mathcal{L}$. Note that if $i_{\max} > J$, then the shift parameters are repeated cyclically.

3.2 BT for unstable systems

This section is concerned with BT for unstable systems via Bernoulli stabilization. Consider the case where the system in (9) is unstable, i.e., some eigenvalues of the system lie in \mathbb{C}^+ . The helpful feature of our investigated example is that still the number of anti-stable eigenvalues is very small. This is exactly the property we exploit for fast computation of the ROMs and ROM based feedback matrices. In the previous subsection we have recalled classical (Lyapunov based) balancing for stable systems. The main ingredients there are the two Gramians (e.g., [2])

$$\mathcal{P} = \int_0^\infty e^{\mathcal{E}^{-1}\mathcal{A}t} \mathcal{E}^{-1} \mathcal{B} \mathcal{B}^T \mathcal{E}^{-T} e^{\mathcal{A}^T \mathcal{E}^{-T} t} dt,$$

$$\mathcal{Q} = \int_0^\infty e^{\mathcal{E}^{-T} \mathcal{A}^T t} \mathcal{E}^{-T} \mathcal{C}^T \mathcal{C} \mathcal{E}^{-1} e^{\mathcal{A} \mathcal{E}^{-1} t} dt,$$

which obviously do not exist if the system's unstable poles are controllable. This, however, is exactly the desired case in our motivating example (see Section 2). In [31], the authors use the frequency domain representations of these integrals

$$\begin{aligned}\mathcal{P} &= \frac{1}{2} \int_{-\infty}^{\infty} (i\omega\mathcal{E} - \mathcal{A})^{-1} \mathcal{B}\mathcal{B}^T (i\omega\mathcal{E}^T - \mathcal{A}^T) d\omega \\ \mathcal{Q} &= \frac{1}{2} \int_{-\infty}^{\infty} (i\omega\mathcal{E}^T - \mathcal{A}^T)^{-1} \mathcal{C}^T \mathcal{C} (i\omega\mathcal{E} - \mathcal{A}) d\omega\end{aligned}$$

to extend the definition of the Gramians to systems with no poles on the imaginary axis.

Following the theory in [31] the generalized controllability and observability Gramians $\mathcal{P}_s, \mathcal{Q}_s$ for such systems can be computed by solving the algebraic Lyapunov equations

$$\begin{aligned}(\mathcal{A} - \mathcal{B}\mathcal{K}_c)\mathcal{P}_s\mathcal{E}^T + \mathcal{E}\mathcal{P}_s(\mathcal{A} - \mathcal{B}\mathcal{K}_c)^T &= -\mathcal{B}\mathcal{B}^T, \\ (\mathcal{A} - \mathcal{K}_o\mathcal{C})^T\mathcal{Q}_s\mathcal{E} + \mathcal{E}^T\mathcal{Q}_s(\mathcal{A} - \mathcal{K}_o\mathcal{C}) &= -\mathcal{C}^T\mathcal{C},\end{aligned}\tag{15}$$

where $\mathcal{K}_c = \mathcal{B}^T \mathcal{X}_c \mathcal{E}$ and $\mathcal{K}_o = \mathcal{E} \mathcal{X}_o \mathcal{C}^T$ are called Bernoulli stabilizing feedback matrices, due to the fact that the matrices \mathcal{X}_c and \mathcal{X}_o are the respective stabilizing solutions of the generalized algebraic Bernoulli equations

$$\begin{aligned}\mathcal{E}^T \mathcal{X}_c \mathcal{A} + \mathcal{A}^T \mathcal{X}_c \mathcal{E} &= \mathcal{E}^T \mathcal{X}_c \mathcal{B} \mathcal{B}^T \mathcal{X}_c \mathcal{E}, \\ \mathcal{A} \mathcal{X}_o \mathcal{E}^T + \mathcal{E} \mathcal{X}_o \mathcal{A}^T &= \mathcal{E} \mathcal{X}_o \mathcal{C}^T \mathcal{C} \mathcal{X}_o \mathcal{E}^T.\end{aligned}\tag{16}$$

Now, since the Bernoulli stabilization only mirrors the anti-stable eigenvalues across the imaginary axis [6], it is sufficient to solve these Bernoulli equations only on the invariant subspaces corresponding to those eigenvalues. That is, for orthogonal matrices V and W spanning the left and right eigenspaces corresponding to the anti-stable eigenvalues, respectively, we define the Petrov-Galerkin projected system $(\check{\mathcal{E}}, \check{\mathcal{A}}, \check{\mathcal{B}}, \check{\mathcal{C}})$ by

$$\check{\mathcal{E}} := V^T \mathcal{E} W, \quad \check{\mathcal{A}} := V^T \mathcal{A} W, \quad \check{\mathcal{B}} := V^T \mathcal{B}, \quad \check{\mathcal{C}} := \mathcal{C} W.$$

Then we solve

$$\begin{aligned}\check{\mathcal{E}}^T \check{\mathcal{X}}_c \check{\mathcal{A}} + \check{\mathcal{A}}^T \check{\mathcal{X}}_c \check{\mathcal{E}} &= \check{\mathcal{E}}^T \check{\mathcal{X}}_c \check{\mathcal{B}} \check{\mathcal{B}}^T \check{\mathcal{X}}_c \check{\mathcal{E}}, \\ \check{\mathcal{A}} \check{\mathcal{X}}_o \check{\mathcal{E}}^T + \check{\mathcal{E}} \check{\mathcal{X}}_o \check{\mathcal{A}}^T &= \check{\mathcal{E}} \check{\mathcal{X}}_o \check{\mathcal{C}}^T \check{\mathcal{C}} \check{\mathcal{X}}_o \check{\mathcal{E}}^T,\end{aligned}\tag{17}$$

and construct $\mathcal{K}_c = \mathcal{B}^T V \check{\mathcal{X}}_c V^T \mathcal{E}$ and $\mathcal{K}_o = \mathcal{E} W^T \check{\mathcal{X}}_o W \mathcal{C}^T$. The projected Bernoulli equations in (17) are solved by the *Matrix Sign Function* method presented in [6].

The low-rank factors of \mathcal{P}_s and \mathcal{Q}_s can also be computed by solving (15) using Algorithm 1, but avoiding to form the closed loop matrices stays crucial. We discuss this issue in more detail in Section 6. Then using the projection matrices as in (11) and applying them to the original unstable system we can compute an unstable ROM that satisfies the error bound as in (14), but with the \mathcal{H}^∞ -norm replaced by the \mathcal{L}^∞ -norm. Therefore, this error bound can not be translated to a global time domain error bound, as in the classic setting, due to the lack of a Parseval-identity-like result. In fact in the numerical experiments we observed that the error may very well grow over time.

4 Reformulation of the structured dynamical systems

The purpose of this section is to convert the index-2 system (2) into an equivalent ODE system in order to make it fit into the framework for BT based model order reduction as discussed in the previous section. The idea is exactly the same as in [19, Section 3]. For convenience we briefly review it here again. Let us consider a projector of the form

$$\Pi = I - A_2(A_2^T E_1^{-1} A_2)^{-1} A_2^T E_1^{-1}. \quad (18)$$

Clearly one then has

$$\Pi E_1 = E_1 \Pi^T, \quad (19a)$$

$$\text{Null}(\Pi) = \text{Range}(A_2), \text{ and} \quad (19b)$$

$$\text{Range}(\Pi) = \text{Null}(A_2^T E_1^{-1}). \quad (19c)$$

These properties imply

$$A_2^T Y = 0 \quad \text{if and only if} \quad \Pi^T Y = Y, \quad (20)$$

i.e., the image of Π^T is exactly the subspace where the algebraic condition (2b) of the DAE investigated here is satisfied. Now applying the projector in (2) and the output equation (3), and thus reformulating the system on this exact subspace, we obtain the following projected system

$$\Pi E_1 \Pi^T \dot{v}(t) = \Pi A_1 \Pi^T v(t) + \Pi B_1 u(t), \quad (21a)$$

$$y(t) = C_1 \Pi^T v(t). \quad (21b)$$

The system dynamics of (21) are projected onto the $n_m := n_1 - n_2$ dimensional subspace $\text{Range}(\Pi^T)$ [19]. This subspace is, however, still represented in the coordinates of the n_1 dimensional space. The n_m dimensional representation can be made explicit by employing the *thin* singular value decomposition (SVD)

$$\Pi = [U_1 \quad U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 \Sigma_1 V_1^T = \Theta_l \Theta_r^T, \quad (22)$$

where $\Theta_l = U_1$ and $\Theta_r = V_1$ and in which $U_1, V_1 \in \mathbb{R}^{n_1 \times n_m}$ consist of the corresponding leading n_m columns of $U, V \in \mathbb{R}^{n_1 \times n_1}$. Moreover, Θ_l, Θ_r satisfy

$$\Theta_l^T \Theta_r = I. \quad (23)$$

This representation is always possible since Π is a projector and therefore, $\Sigma_1 = I_{n_m}$ (n_m dimensional identity matrix). Inserting the decomposition of Π from (22) into (21) and considering $\tilde{v} = \Theta_l^T v$, we get

$$\begin{aligned} \Theta_r^T E_1 \Theta_r \dot{\tilde{v}}(t) &= \Theta_r^T A_1 \Theta_r \tilde{v}(t) + \Theta_r^T B_1 u(t), \\ y(t) &= C_1 \Theta_r \tilde{v}(t). \end{aligned} \quad (24)$$

System (24) practically is system (9) with the redundant equations removed by the Θ_r projection. We observe that the dynamical systems (2), (21) and (24) are equivalent in the sense that their finite spectra are the same [15, Theorem 2.7.3] and the input-output relations are the same, i.e., they realize the same transfer function.

5 BT for index-2 unstable descriptor systems

In this section we show how to avoid forming (24) explicitly. Suppose that we want to apply balanced truncation to the system (24). To this end, we need to solve the Lyapunov equations

$$\begin{aligned}\Theta_r^T A_c \Theta_r \tilde{P} \Theta_r^T E_1^T \Theta_r + \Theta_r^T E_1 \Theta_r \tilde{P} \Theta_r^T A_c^T \Theta_r &= -\Theta_r^T B_1 B_1^T \Theta_r, \\ \Theta_r^T A_o^T \Theta_r \tilde{Q} \Theta_r^T E_1 \Theta_r + \Theta_r^T E_1^T \Theta_r \tilde{Q} \Theta_r^T A_o \Theta_r &= -\Theta_r^T C_1^T C_1 \Theta_r,\end{aligned}\quad (25)$$

where $A_c = A_1 - B_1 K_c^1$, $A_o = A_1 - K_o^1 C_1$ and $\tilde{P} \in \mathbb{R}^{n_m \times n_m}$, $\tilde{Q} \in \mathbb{R}^{n_m \times n_m}$ are the corresponding projected controllability and observability Gramians. Again, K_c^1 and K_o^1 are the Bernoulli stabilizing feedback matrices and can be computed as described in Section 3.2. The solutions \tilde{P} , \tilde{Q} of (25) are unique since we assured that the respective dynamical system is asymptotically stable and symmetric positive (semi-)definite since the right hand side is semi-definite.

Now multiplying (25) by Θ_l from the left and Θ_l^T from the right and exploiting that $\Theta_r = \Pi^T \Theta_r$ (e.g., due to (22), (23)) we obtain

$$\begin{aligned}\Pi A_c \Pi^T P \Pi E_1^T \Pi^T + \Pi E_1 \Pi^T P \Pi A_c^T \Pi^T &= -\Pi B_1 B_1^T \Pi^T, \\ \Pi A_o^T \Pi^T Q \Pi E_1 \Pi^T + \Pi E_1^T \Pi^T Q \Pi A_o \Pi^T &= -\Pi C_1^T C_1 \Pi^T,\end{aligned}\quad (26)$$

where $P = \Theta_r \tilde{P} \Theta_r^T$ and $Q = \Theta_r \tilde{Q} \Theta_r^T$. These are the respective controllability and observability Lyapunov equations for the realization (21) and the solutions P , $Q \in \mathbb{R}^{n_1 \times n_1}$ are the corresponding controllability and observability Gramians. The system (26) is singular due to the fact that Π is a projection. Uniqueness of solutions is reestablished by the condition that the solutions satisfy $P = \Pi^T P \Pi$ and $Q = \Pi^T Q \Pi$.

It is also an easy exercise to go back to (25) from (26). Let us consider $P \approx R R^T$, $Q \approx L L^T$ and $\tilde{P} \approx \tilde{R} \tilde{R}^T$, $\tilde{Q} \approx \tilde{L} \tilde{L}^T$. Then R , L , \tilde{R} and \tilde{L} are called approximate low-rank Cholesky factors. They fulfill the relation

$$R = \Theta_r \tilde{R} \quad \text{and} \quad L = \Theta_r \tilde{L}. \quad (27)$$

For large-scale problems, however, computing Θ_r is usually impossible due to memory limitations. Therefore, R and L are computed by solving (26). The balancing transformations for (24) are

$$\tilde{T}_L = \tilde{R} U_k S_k^{-\frac{1}{2}}, \quad \tilde{T}_R = \tilde{L} V_k S_k^{-\frac{1}{2}},$$

where $U_k, V_k \in \mathbb{R}^{n_m \times k}$ consist of the corresponding leading k columns of $U, V \in \mathbb{R}^{n_m \times n_m}$, and $S_k \in \mathbb{R}^{k \times k}$ is the upper left $k \times k$ block of S in the SVD

$$\tilde{R}^T \Theta_r^T E_1 \Theta_r \tilde{L} = U S V^T.$$

Observing further that $R^T \Pi E_1 \Pi^T L = \tilde{R}^T \Theta_r^T E_1 \Theta_r \tilde{L} = USV^T$, the projection matrices for the system (21) can be formed as

$$T_L = RU_k S_k^{-\frac{1}{2}} \quad \text{and} \quad T_R = LV_k S_k^{-\frac{1}{2}}. \quad (28)$$

As in [19] we find that

$$\begin{aligned} T_L &= RU_k S_k^{-\frac{1}{2}} = \Theta_r \tilde{R} U_k S_k^{-\frac{1}{2}} = \Theta_r \tilde{T}_L = \Theta_r \Theta_l^T \Theta_r \tilde{T}_L = \Pi^T T_L, \\ T_R &= LV_k S_k^{-\frac{1}{2}} = \Theta_r \tilde{L} V_k S_k^{-\frac{1}{2}} = \Theta_r \tilde{T}_R = \Theta_r \Theta_l^T \Theta_r \tilde{T}_R = \Pi^T T_R. \end{aligned} \quad (29)$$

Now we apply the transformations T_R and T_L in (21) to find the reduced-order model as in (6) where

$$\hat{E} = T_L^T \Pi E_1 \Pi^T T_R, \quad \hat{A} = T_L^T \Pi A_1 \Pi^T T_R, \quad \hat{B} = T_L^T \Pi B_1 \quad \text{and} \quad \hat{C} = C_1 \Pi^T T_R. \quad (30)$$

Due to (29) we can avoid the explicit usage of Π and find

$$I = \hat{E} = T_L^T E_1 T_R, \quad \hat{A} = T_L^T A_1 T_R, \quad \hat{B} = T_L^T B_1 \quad \text{and} \quad \hat{C} = C_1 T_R. \quad (31)$$

Eventually, we see that the reduced-order model (6) is obtained without forming the projected system (21). In the next section we will show how to compute R and L using a tailored version of the LRCF-ADI iteration without using Π explicitly.

6 Solution of the projected Lyapunov equations

In order to apply the aforementioned balancing based MOR we need to solve the projected Lyapunov equations (26). We have seen above that the Π^T invariant solution factors enable us to compute the corresponding truncating transformations. The approach here is different from the spectral projection based work by Stykel in that we are applying the E_1 -orthogonal projection to the hidden manifold, where Stykel uses the orthogonal projection (in the Euclidean sense) to the eigenspaces corresponding to the finite poles of the system. In fact both methods project to the same subspace only considering orthogonality in different inner products. Here we are concerned with two main issues. First we discuss the reformulation of the basic low-rank ADI Algorithm for the projected Lyapunov equation that ensures the invariance of the solution factor and the computation of the correct corresponding residual factors. We are lifting the ideas of [19] to the reformulation of the LR-ADI iteration in Algorithm 1. For the spectral projection methods the analogue procedure has been discussed in [12]. In the second part we address the important issue of ADI shift parameter computation. There the main issue in the DAE setting is to avoid the subspaces corresponding to infinite eigenvalues in order to correctly compute the large magnitude Ritz values involved in many parameter choices. The crucial point in both parts is to provide methods that use the projection Π^T at most implicitly and never form the projected system (21).

Algorithm 2 GS-LRCF-ADI

Input: E_1, A_1, A_2, B_1, K_1 and ADI shift parameters $\{\mu_1, \mu_2, \dots, \mu_{i_{max}}\}$

Output: \tilde{R} , such that $\tilde{P} \approx \tilde{R}\tilde{R}^T$

```
1:  $\tilde{R} = []$ 
2: Solve the linear system (39) for  $\Xi$  and compute  $W_0 = E_1\Xi$ 
3:  $i = 1$ 
4: while  $\|W_{i-1}^T W_{i-1}\| > \text{tol}$  do
5:   solve the linear system (36) for  $V_i$ 
6:   if  $\text{Im}(\mu_i) = 0$  then
7:      $\tilde{R} = [\tilde{R} \quad \sqrt{-2 \text{Re}(\mu_i)} V_i]$ 
8:      $W_i = W_{i-1} - 2 \text{Re}(\mu_i) E_{11} V_i$ 
9:   else
10:     $\gamma = 2\sqrt{-\text{Re}(\mu_i)}, \beta = \frac{\text{Re}(\mu_i)}{\text{Im}(\mu_i)}$ 
11:    Update low-rank solution factor
12:     $\tilde{R} = [\tilde{R} \quad \gamma(\text{Re}(V_i) + \beta \text{Im}(V_i)) \quad \gamma\sqrt{(\beta^2 + 1)} \cdot \text{Im}(V_i)]$ 
13:     $W_i = W_{i-1} - 4 \text{Re}(\mu_i) E_{11}(\overline{V_i} + 2\beta \text{Im}(V_i))$ 
14:     $i = i + 1$ 
15:   end if
16: end while
```

6.1 GS-LRCF-ADI for index-2 unstable systems

This section is concerned with the efficient solution of the Lyapunov equations in (26) to compute the low-rank Gramian factors using LRCF-ADI as discussed in Section 3. Here, we consider the projected controllability equation elaborately. The observability equation can be handled analogously. For convenience we rewrite the Lyapunov equations (26) as

$$\begin{aligned} \tilde{A}\tilde{P}\tilde{E}^T + \tilde{E}\tilde{P}\tilde{A}^T &= -\tilde{B}\tilde{B}^T, \\ \tilde{A}^T\tilde{Q}\tilde{E} + \tilde{E}^T\tilde{Q}\tilde{A} &= -\tilde{C}^T\tilde{C}, \end{aligned} \quad (32)$$

with $\tilde{E} = \Pi E_1 \Pi^T$, $\tilde{A} = \Pi A_c \Pi^T$, $\tilde{B} = \Pi B_1$ and $\tilde{C} = C_1 \Pi^T$.

In the i -th iteration step of the ADI method, the residual of the controllability Lyapunov equation (32) can be written as

$$\tilde{\mathcal{R}}(\tilde{P}_i) = \tilde{A}\tilde{P}_i\tilde{E}^T + \tilde{E}\tilde{P}_i\tilde{A}^T + \tilde{B}\tilde{B}^T = \tilde{W}_i\tilde{W}_i^T,$$

where

$$\tilde{W}_i = \prod_{k=1}^i (\tilde{A} - \bar{\mu}_k \tilde{E})(\tilde{A} + \mu_k \tilde{E})^{-1} \tilde{B}.$$

To compute the low-rank controllability Gramian factor \tilde{R} we follow Algorithm 1. In the i -th iteration step, V_i is computed from

$$(\tilde{A} + \mu_i \tilde{E})V_i = \tilde{W}_{i-1}, \quad (33)$$

which enables us to update the residual factor according to

$$\tilde{W}_i = (\tilde{A} - \mu^* \tilde{E})V_i = \tilde{W}_{i-1} - 2 \operatorname{Re}(\mu_i) \tilde{E}V_i. \quad (34)$$

In complete analogy to [19, Lemma 5.2], we observe that instead of solving (33), one can compute V_i by solving

$$\begin{bmatrix} A_c + \mu_i E_1 & A_2 \\ A_2^T & 0 \end{bmatrix} \begin{bmatrix} V_i \\ \star \end{bmatrix} = \begin{bmatrix} \tilde{W}_{i-1} \\ 0 \end{bmatrix}, \quad (35)$$

where the special case $i = 1$, here especially the computation of the initial residual factor \tilde{W}_0 , is discussed in detail below.

Inserting $A_c = A_1 - B_1 K_c^1$ in (35),

$$\begin{bmatrix} A_1 + \mu_i E_1 - B_1 K_c^1 & A_2 \\ A_2^T & 0 \end{bmatrix} \begin{bmatrix} V_i \\ \star \end{bmatrix} = \begin{bmatrix} \tilde{W}_{i-1} \\ 0 \end{bmatrix},$$

implies

$$\left(\underbrace{\begin{bmatrix} A_1 + \mu_i E_1 & A_2 \\ A_2^T & 0 \end{bmatrix}}_{\underline{A}} - \underbrace{\begin{bmatrix} B_1 \\ 0 \end{bmatrix}}_{\underline{B}} \underbrace{\begin{bmatrix} K_c^1 & 0 \end{bmatrix}}_{\underline{K}} \right) \begin{bmatrix} V_i \\ \star \end{bmatrix} = \begin{bmatrix} \tilde{W}_{i-1} \\ 0 \end{bmatrix}. \quad (36)$$

In this equation the inversion of $(\underline{A} - \underline{B} \underline{K})$ should in practice be computed using the *Sherman-Morrison-Woodbury formula* (see, e.g. [17]):

$$(\underline{A} - \underline{B} \underline{K})^{-1} = \underline{A}^{-1} + \underline{A}^{-1} \underline{B} (I - \underline{K} \underline{A}^{-1} \underline{B})^{-1} \underline{K} \underline{A}^{-1}, \quad (37)$$

to avoid explicit forming of the (usually dense) matrix $\underline{A} - \underline{B} \underline{K}$. In accordance with [19, Lemma 5.2], again the computed V_i in (35) satisfies $V_i = \Pi^T V_i$. Therefore, the correct projected residual factor in (34) can be obtained by

$$\tilde{W}_i = \tilde{W}_{i-1} - 2 \operatorname{Re}(\mu_i) E_1 V_i, \quad (38)$$

since we have $\Pi E_1 = E_1 \Pi^T$.

In order to really compute the correct residual, the initial residual must be computed as $\tilde{W}_0 = \Pi B_1$ to ensure $\tilde{W}_0 = \Pi \tilde{W}_0$. This can be performed cheaply using the following observation.

Lemma 6.1. *The matrix Ξ satisfies $\Xi = \Pi^T \Xi$ and $E_1 \Xi = \Pi B_1 \Leftrightarrow$*

$$\begin{bmatrix} E_1 & A_2 \\ A_2^T & 0 \end{bmatrix} \begin{bmatrix} \Xi \\ \Lambda \end{bmatrix} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}. \quad (39)$$

Proof. If $\Xi = \Pi^T \Xi$, then $E_1 \Xi = \Pi B_1$ implies $\Pi(E_1 \Xi - B_1) = 0$. Since $\operatorname{Null}(\Pi) = \operatorname{Range}(A_2)$, there exists Λ such that $E_1 \Xi - B_1 = -A_2 \Lambda$, or $E_1 \Xi + A_2 \Lambda = B_1$. Again

applying the properties in (20), we have $A_2^T \Xi = 0$. These two relations give (39). Conversely, we assume (39) holds. From the first block row of (39) we get

$$\Xi = E_1^{-1} B_1 - E_1^{-1} A_2 \Lambda,$$

and thus from the second row it follows

$$0 = A_2^T \Xi = A_2^T E_1^{-1} B_1 - A_2^T E_1^{-1} A_2 \Lambda,$$

such that

$$\Lambda = (A_2^T E_1^{-1} A_2)^{-1} A_2^T E_1^{-1} B_1.$$

Inserting this in the first block row we get as desired

$$E_1 \Xi = B_1 - (A_2^T E_1^{-1} A_2)^{-1} A_2^T E_1^{-1} B_1 = \Pi B_1.$$

□

This especially ensures $\Xi = \Pi^T \Xi$, since

$$E_1 \Xi = \Pi B_1 = \Pi^2 B_1 = \Pi E_1 \Xi = E_1 \Pi^T \Xi,$$

and thus using $\tilde{W}_0 = E_1 \Xi$, we get the desired invariance $\tilde{W}_0 = \Pi \tilde{W}_0$.

The above findings on the residual factor are summarized in the following lemma.

Lemma 6.2. *The residual factor in every step of Algorithm 2 fulfills the relation*

$$\tilde{W}_i = \Pi \tilde{W}_i.$$

The whole procedure to compute the low-rank factor of the controllability Gramian \tilde{R} is summarized in Algorithm 2. Analogous to the derivation in [19], our algorithm computes the correct solution factor. In contrast to the version presented there, we guarantee to compute a real solution factor even if the shifts occur in complex conjugate pairs and we have the low rank residual factors in hand to evaluate stopping criteria cheaply. Still one issue remains open that has not been tackled in the original paper [19]. The shifts that guarantee fast convergence of the algorithm are closely related to the spectrum of the original pencil. The question how these can be computed is answered in the next section.

6.2 ADI-parameter selection

The appropriate shift parameter selection is one of the crucial tasks for fast convergence of the GS-LRCF-ADI iteration. Recently, most of the papers followed the *heuristic* procedure introduced by Penzl [23] to compute sub-optimal ADI shift parameters μ_i , $i = 1, 2, \dots, J$, for a large-scale problem. Very recently new shift computation ideas considering adaptive and automatic computation of shifts during the iteration [9, 30] have come up. We present the basic ideas to adapt both the classic and the new methods to our framework in the following two paragraphs.

Heuristic shift selection. The main ingredient of the heuristic method is the computation of a number of large and small magnitude Ritz values. In the case of DAE systems the computation of Ritz values of large magnitude is causing difficulties due to the existence of infinite eigenvalues. We need to make sure that the infinite eigenvalues are avoided. This can be achieved by the following fact that is a direct consequence of [13, Theorem 3.1].

Corollary 6.3. *The matrix pencil*

$$\mathcal{P}_\delta(\lambda) = \lambda \begin{bmatrix} E_1 & \delta A_2 \\ \delta A_2^T & 0 \end{bmatrix} - \begin{bmatrix} A_1 & A_2 \\ A_2^T & 0 \end{bmatrix} \quad (40)$$

transforms all infinite eigenvalues of the pencil $\mathcal{P}(\lambda)$ (see (4)) to $\frac{1}{\delta}$ while at the same time preserving its finite eigenvalues.

Thus from the pencil \mathcal{P}_δ we can compute the desired Ritz values of large magnitude via an Arnoldi iteration [24]. The parameter δ can easily be determined after the small Ritz values β_i have been computed with respect to the original pencil. In order to ensure that $\frac{1}{\delta}$ is avoided by the Arnoldi process for the large magnitude Ritz values, and thus only finite eigenvalues of the original pencil are approximated, one could, e.g., set $\delta = \frac{1}{\min_i \operatorname{Re}(\beta_i)}$. For the unstable case the corollary obviously has to be applied with A_1 replaced by A_c .

Adaptive shift selection. A second shift computation strategy we use in the numerical experiments follows the lines of the adaptive shift strategy proposed in [9]. There, the shifts are initialized by the eigenvalues of the pencil projected to the span of W_0 . Then, whenever all shifts in the current set have been used, the pencil is projected, e.g., to the span of the current V_i and the eigenvalues are used as the next set of shifts. Here, we use the same initialization. For the update step, however, we extend the subspace to all the V_i generated with the current set of shifts and then choose the next shifts following Penzl's heuristic with the Ritz values replaced by the projected eigenvalues computed with respect to this larger subspace. Note that in lack of the conditions for Bendixon's theorem, we cannot guarantee that the projected eigenvalues will be contained in \mathbb{C}_- . Should any of them end up in the wrong half-plane, we neglect them. Should the resulting set of shifts become empty due to this, we reuse the previous set as in [9]. Note further that the problem with the infinite eigenvalues does not exist in this case. Since we have $\Pi^T Z = Z$, for any orthogonal bases U of a set of columns of Z , we also have $\Pi^T U = U$. As an immediate result of this fact, the projected pencil $(U^T A_1 U - \lambda U^T E_1 U)$ automatically resides on the hidden manifold and can thus only have finite eigenvalues.

7 Numerical results

To assess the performance of the techniques, this section discusses some numerical tests. The method is applied to a set of linearized semi-discretized Navier-Stokes

model	n_1	n_2
Mod-1	3 142	453
Mod-2	8 268	1 123
Mod-3	19 770	2 615
Mod-4	44 744	5 783
Mod-5	98 054	12 566

Table 1: The number of differential and algebraic variables of different discretization levels of the model.

model	heuristic shift			adaptive shift		
	iteration		time (sec)	iteration		time (sec)
	\tilde{R}	\tilde{L}	$\tilde{R} + \tilde{L} + \mu$	\tilde{R}	\tilde{L}	$\tilde{R} + \tilde{L}$
Mod-1	240	210	67	116	88	25
Mod-2	170	133	165	106	77	87
Mod-3	257	182	625	114	99	305
Mod-4	307	196	1 922	146	111	1 063
Mod-5	368	238	5 839	147	120	2 551

Table 2: The performances of the heuristic and adaptive shifts in the GS-LRCF-ADI iteration.

equations as described in Section 2. All the computations were carried out using MATLAB[®] 7.11.0 (R2010b) on a board with 2 Intel[®] Xeon[®] X5650 CPUs with a 2.67-GHz clock speed.

The authors of [5] generate different sized models using Reynolds number $R_e = 500$. Table 1 shows the different sizes of the model and distinguishes the dimensions n_1 of the velocity vector (differential variable) and n_2 of the pressure vector (algebraic variable).

In all the sets $B_1 \in \mathbb{R}^{n_1 \times 2}$ and $C_1 \in \mathbb{R}^{7 \times n_1}$. For Reynolds numbers of 400 and more the described linearized model is unstable. Thus, especially the Reynolds number 500 case discussed here is unstable. The Bernoulli stabilizing feedback matrices K_c^1 and K_o^1 for all models are computed applying the procedure from [1] and [5, Section 2]. It uses 2 calls of the MATLAB `eigs` implementation of the Arnoldi method (for P_δ and P_δ^T) to compute the rightmost eigenvalues together with their left and right eigenvectors.

We apply the GS-LRCF-ADI iteration (Algorithm 2) to all aforementioned models to compute the low-rank factors \tilde{R} and \tilde{L} considering the tolerance 10^{-6} . We investigate the performances of both the *heuristic* and *adaptive* shifts to implement this algorithm. The results are shown in Table 2. For all models we chose 30 optimal heuristic shifts out of 10 large and 80 small magnitude Ritz-values. In the case of the adaptive shifts, in each cycle, 10 proper shift parameters are selected following the procedure discussed above. For computing the initial shifts, first we project the pencil $(A_1 - \lambda E_1)$, onto the column space of an $n_1 \times 100$ random matrix. For all the models, looking at the iteration

model	tolerance	system dimension	
		original	reduced
Mod-1		3 595	145
Mod-2		9 391	147
Mod-3	10^{-5}	22 385	163
Mod-4		50 527	178
Mod-5		110 620	184

Table 3: Dimensions of original and reduced systems of the differently sized original models for a fixed balanced truncation tolerance.

model	tolerance	dimension of ROM
	10^{-4}	161
	10^{-3}	138
Mod-5	10^{-2}	115
	10^{-1}	93

Table 4: Balanced truncation tolerances and dimensions of reduced models.

numbers and and the CPU time computations, the performance of the adaptive shifts is much better than that of the heuristic shifts. The performances of the heuristic and adaptive shifts are also depicted in Figure 2 for the largest dimensional system Mod-5. This figure shows the convergence rate of the low-rank controllability and observability Gramian factors with respect to iterations (Figures 2a, 2b) and time (Figures 2c, 2d). In both cases the convergence rate for the adaptive shifts is much faster than the heuristic shifts. Note that we use the Frobenius norm to compute the residual norm. We compute the reduced-order models for all the data sets. The dimensions of the original and reduced models are shown in Table 3. If nothing else is stated, the truncation tolerance is set to 10^{-5} . The dimension of the ROM can, however, be decreased by increasing the tolerance if desired or required. This is shown in Table 4. Since the numerical results are all comparable, we exemplarily present only selected plots (e.g., for the largest model Mod-5).

7.1 Model reduction of the unstable system

This section reviews the numerical experiments for the unstable case. Here we present both frequency and time domain error analyses. The frequency domain error analysis is shown in Figure 3. In Figure 3a we see the frequency responses of the full and 184 dimensional reduced-order models for Mod-5 with a nice match in the eyeball norm. The absolute and relative deviations between full and reduced-order models are shown, respectively, in Figures 3b and 3c. Here, we can see that the absolute error is bounded by the prescribed truncation tolerance of 10^{-5} . For higher frequencies the relative error is slightly increasing since the frequency response is decreasing more rapidly than the

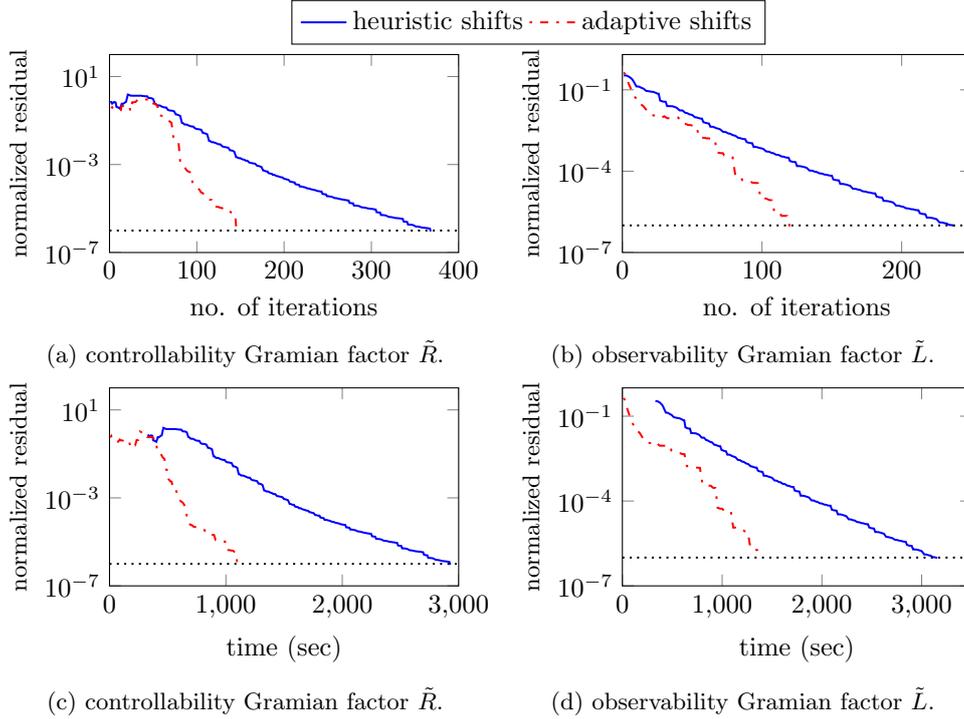


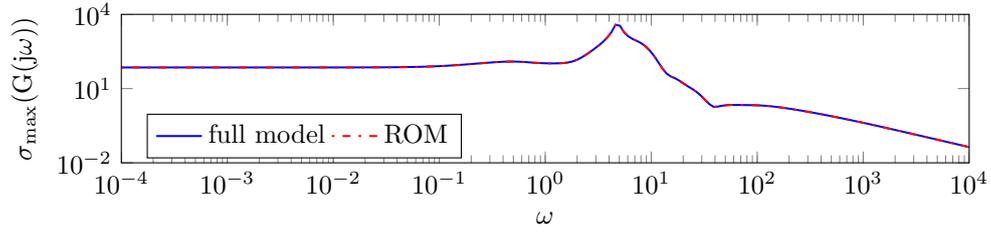
Figure 2: Comparisons of the heuristic and adaptive shifts in computing the low-rank Gramian factors using the GS-LRCF-ADI iteration.

absolute error can.

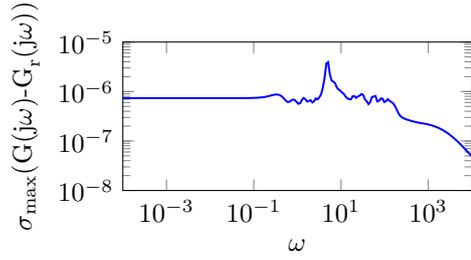
Figure 4 depicts time domain simulation of full and reduced-order models for Mod-5. This figure shows the step responses from Input 1 to Output 1 together with their absolute deviations. To compute the step response we use an implicit Euler method with fixed time step size 10^{-2} . Initially, the imposed control is kept inactive, therefore the responses for both (full and reduced) models are constant within the range 0 to 15s. Switching the control to constant unit actuation on Input 1, the responses are oscillating with increasing amplitude in the higher time domain caused by the instability of the model. Here we also see the issue with the balanced truncation error bound for unstable systems since the absolute error is increasing gradually with increasing time.

7.2 Numerical Experiments for the stabilized system

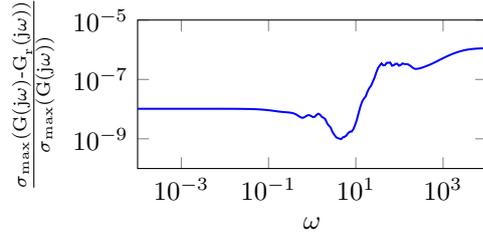
In Section 2 we have mentioned that the stabilizing feedback matrix for the full model can be computed from the reduced-order model. To this end, we solve the corresponding algebraic Riccati equation (7) arising from the linear-quadratic regulator approach using the MATLAB `care` command and compute the optimal stabilizing feedback ma-



(a) Sigma plot.

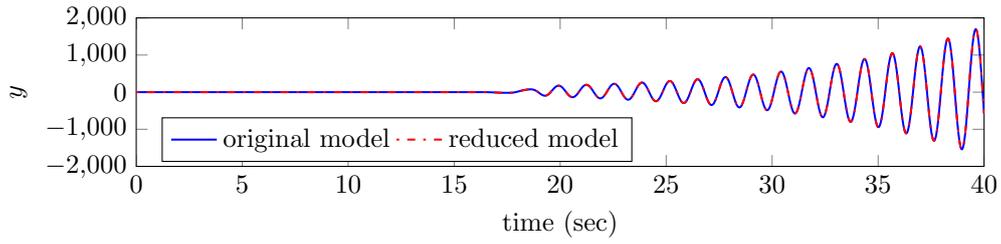


(b) Absolute error.

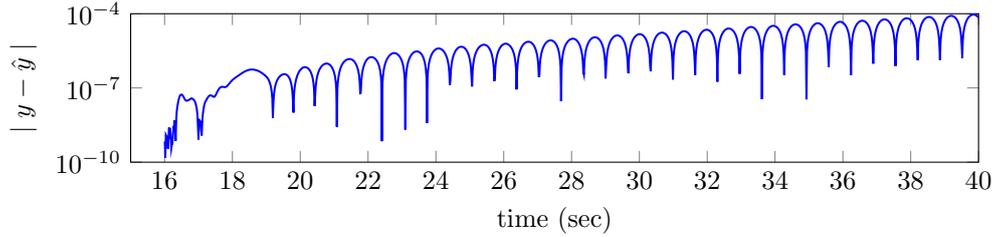


(c) Relative error.

Figure 3: Comparison of the full and reduced models in frequency domain.

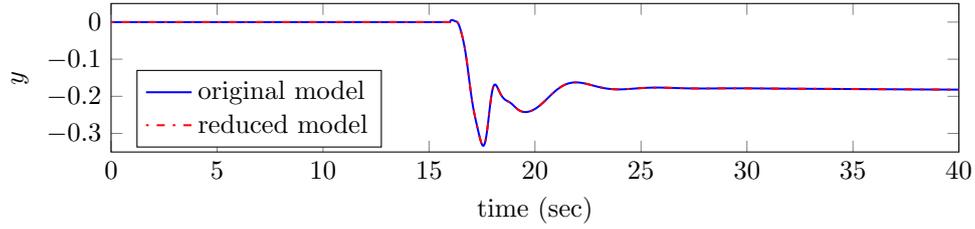


(a) Step response.

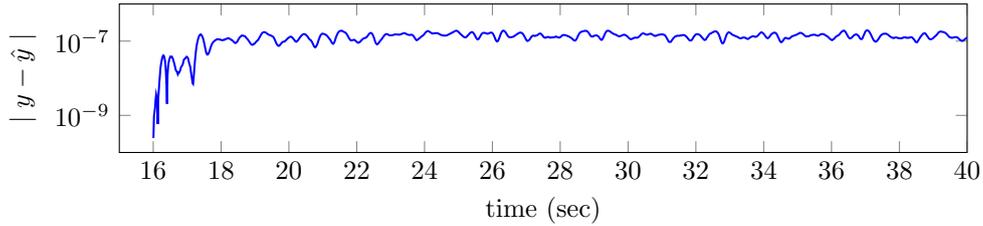


(b) Absolute error.

Figure 4: Step responses of 1st input to 1st output of full and reduced-order models and respective absolute deviations.



(a) Step response.



(b) Absolute error.

Figure 5: Step responses of 1st input to 1st output of stabilized full and reduced-order models and respective absolute deviations.

trix \hat{K} (see, e.g., [14, Chap. 10]) for the reduced-order model (6). The ROM based approximation to the stabilizing feedback matrix for the full order model (5) is then given by $K_1 = K_r T_L^T E_1$. Figure 5 shows the step response (from 1st input to 1st output) of closed loop full and reduced-order models and their absolute error. For the generation of the step response the same procedure has been followed as for the unstable case above. Note that for a stabilizing feedback the step response system has to be viewed as that of an asymptotically stable system with a constant source term. Thus the output settles at constant non zero values.

8 Conclusions

We have discussed the model reduction problem for unstable descriptor systems arising from flow control problems. In particular, we have considered linearized Navier-Stokes equations. Their spatial discretization by finite elements leads to index-2 DAEs. This causes some technical difficulties for the application of model reduction based on balanced truncation. Following [19], we have treated the problem by using projectors onto deflating subspaces corresponding to finite eigenvalues of the associated matrix pencils. The explicit application of the projectors is avoided by a re-formulation of the ADI method used to compute approximate low-rank factors of the system Gramians which is the bottleneck in the numerical implementation of the balanced truncation method. Here we show how this can be done for unstable systems, and we also in-

corporate a number of recent improvements into the resulting ADI method. As an illustrative example, we apply our algorithms to the linearization of the von Kármán vortex shedding at a moderate Reynolds number. It is demonstrated that the resulting reduced-order model can be used to accurately simulate the unstable linearized model and to design a stabilizing controller. Future work will include the realization of the resulting control law for the full nonlinear model.

Acknowledgments

The authors express their thanks to Heiko Weichelt for providing the models used in the numerical experiments.

M. Monir Uddin is also affiliated to the "International Max Planck Research School (IMPRS) for Advanced Methods in Process and System Engineering (Magdeburg)".

References

- [1] L. AMODEI AND J.-M. BUCHOT, *A stabilization algorithm of the Navier-Stokes equations based on algebraic Bernoulli equation*, Numer. Lin. Alg. Appl., 19 (2012), pp. 700–727.
- [2] A. C. ANTOULAS, *Approximation of Large-Scale Dynamical Systems*, SIAM Publications, Philadelphia, PA, 2005.
- [3] A. C. ANTOULAS, D. C. SORENSEN, AND Y. ZHOU, *On the decay rate of Hankel singular values and related issues*, Systems Control Lett., 46 (2002), pp. 323–342.
- [4] U. M. ASCHER AND L. R. PETZOLD, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, SIAM, Philadelphia, PA, 1998.
- [5] E. BÄNSCH, P. BENNER, J. SAAK, AND H. K. WEICHELT, *Riccati-based boundary feedback stabilization of incompressible Navier-Stokes flow*, Preprint SPP1253-154, DFG-SPP1253, Sept. 2013. Available from: <http://www.am.uni-erlangen.de/home/spp1253/wiki/index.php/Preprints>.
- [6] S. BARRACHINA, P. BENNER, AND E. S. QUINTANA-ORTÍ, *Efficient algorithms for generalized algebraic Bernoulli equations based on the matrix sign function*, Numer. Algorithms, 46 (2007), pp. 351–368.
- [7] P. BENNER, P. KÜRSCHNER, AND J. SAAK, *Efficient handling of complex shift parameters in the low-rank Cholesky factor ADI method*, Numer. Algorithms, 62 (2013), pp. 225–251.
- [8] P. BENNER, P. KÜRSCHNER, AND J. SAAK, *An improved numerical method for balanced truncation for symmetric second order systems*, Math. Comput. Model. Dyn. Syst., 19 (2013), pp. 593–615.

- [9] 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., (2014). to appear.
- [10] P. BENNER, J.-R. LI, AND T. PENZL, *Numerical solution of large Lyapunov equations, Riccati equations, and linear-quadratic control problems*, Numer. Lin. Alg. Appl., 15 (2008), pp. 755–777.
- [11] P. BENNER, V. MEHRMANN, AND D. C. SORESENSEN, *Dimension Reduction of Large-Scale Systems*, vol. 45 of Lect. Notes Comput. Sci. Eng., Springer-Verlag, Berlin/Heidelberg, Germany, 2005.
- [12] P. BENNER AND T. STYKEL, *Numerical solution of projected algebraic Riccati equations*, SIAM J. Numer. Anal., 52 (2014), pp. 581–600.
- [13] K. A. CLIFFE, T. J. GARRATT, AND A. SPENCE, *Eigenvalues of block matrices arising from problems in fluid mechanics*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1310–1318.
- [14] B. N. DATTA, *Numerical Methods for Linear Control Systems*, Elsevier Academic Press, 2004.
- [15] E. EICH-SOELLNER AND C. FÜHRER, *Numerical Methods in Multibody Dynamics*, European Consortium for Mathematics in Industry, B. G. Teubner GmbH, Stuttgart, 1998.
- [16] K. GLOVER, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ norms*, Internat. J. Control, 39 (1984), pp. 1115–1193.
- [17] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1983.
- [18] L. GRASEDYCK AND W. HACKBUSCH, *A multigrid method to solve large scale Sylvester equations*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 870–894.
- [19] M. HEINKENSCHLOSS, D. C. SORESENSEN, AND K. SUN, *Balanced truncation model reduction for a class of descriptor systems with applications to the Oseen equations*, SIAM J. Sci. Comput., 30 (2008), pp. 1038–1063.
- [20] P. HOOD AND C. TAYLOR, *Navier-Stokes equations using mixed interpolation*, in Finite Element Methods in Flow Problems, J. T. Oden, R. H. Gallagher, C. Taylor, and O. C. Zienkiewicz, eds., University of Alabama in Huntsville Press, 1974, pp. 121–132.
- [21] P. KUNKEL AND V. MEHRMANN, *Differential-Algebraic Equations: Analysis and Numerical Solution*, Textbooks in Mathematics, EMS Publishing House, 2006.
- [22] J.-R. LI AND J. WHITE, *Low rank solution of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 260–280.

- [23] T. PENZL, *A cyclic low rank Smith method for large sparse Lyapunov equations*, SIAM J. Sci. Comput., 21 (2000), pp. 1401–1418.
- [24] Y. SAAD, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, Manchester, UK, 1992.
- [25] W. H. A. SCHILDERS, H. A. VAN DER VORST, AND J. ROMMES, *Model Order Reduction: Theory, Research Aspects and Applications*, Springer-Verlag, Berlin, Heidelberg, 2008.
- [26] E. D. SONTAG, *Mathematical Control Theory*, Springer-Verlag, New York, NY, 2nd ed., 1998.
- [27] T. STYKEL, *Gramian-based model reduction for descriptor systems*, Math. Control Signals Systems, 16 (2004), pp. 297–319.
- [28] T. STYKEL, *Balanced truncation model reduction for semidiscretized Stokes equation*, Linear Algebra Appl., 415 (2006), pp. 262–289.
- [29] M. S. TOMBS AND I. POSTLETHWAITE, *Truncated balanced realization of a stable nonminimal state-space system*, Internat. J. Control, 46 (1987), pp. 1319–1330.
- [30] T. WOLF, H. K. F. PANZER, AND B. LOHMANN, *ADI iteration for Lyapunov equations: a tangential approach and adaptive shift selection*, e-print 1312.1142, arXiv, Dec. 2013.
- [31] K. ZHOU, G. SALOMON, AND E. WU, *Balanced realization and model reduction for unstable systems*, Internat. J. Robust and Nonlinear Cont., 9 (1999), pp. 183–198.

