Technische Universität Berlin

Institut für Mathematik

Balanced Truncation Model Reduction for Semidiscretized Stokes Equation

Tatjana Stykel

Technical Report 04-2003

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin

Report 04-2003 **February 2003**

Abstract

We discuss model reduction of linear continuous-time descriptor systems that arise in the control of semidiscretized Stokes equations. Balanced truncation model reduction methods for descriptor systems are presented. These methods are closely related to the proper and improper controllability and observability Gramians and Hankel singular values of descriptor systems. The Gramians can be computed by solving projected generalized Lyapunov equations. Important properties of the balanced truncation approach are that the asymptotic stability is preserved in the reduced order system and there is a priori bound on the approximation error. We demonstrate the application of balanced truncation model reduction to the semidiscretized Stokes equation.

Key words. Stokes equation, descriptor system, model reduction, balanced truncation, controllability and observability Gramians, projected generalized Lyapunov equation, Hankel singular values.

AMS subject classification. 15A24, 37N10, 93A15

Author address:

Tatjana Stykel
Institut für Mathematik, MA 4-5
Technische Universität Berlin
Straße des 17. Juni 136
D-10623 Berlin
Germany

stykel@math.tu-berlin.de

1 Introduction

Consider the instationary Stokes equation describing the flow of an incompressible fluid

$$\begin{aligned} \frac{\partial v}{\partial t} &= \vartheta \Delta v - \nabla p + f, && (\xi,t) \in \Omega \times (0,t_f), \\ 0 &= \operatorname{div} v, && (\xi,t) \in \Omega \times (0,t_f) \end{aligned} \tag{1.1}$$

with the initial and boundary conditions

$$\begin{aligned} v(\xi,t) &= g(\xi,t), && (\xi,t) \in \partial\Omega \times (0,t_f), \\ v(\xi,0) &= v_0(\xi), && \xi \in \Omega. \end{aligned}$$

Here $v \in \mathbb{R}^d$ is the velocity vector ($d = 2$ or 3 is the dimension of the model), $p \in \mathbb{R}$ is the pressure, ϑ is the kinematic viscosity, $f \in \mathbb{R}^d$ is the vector of external forces, $\Omega \subset \mathbb{R}^d$ is a bounded open domain with boundary $\partial\Omega$ and $t_f > 0$ is the endpoint of the time interval. The spatial discretization of the Stokes equation by the finite difference or the finite element method produces a system in the generalized state space (or descriptor) form

$$\begin{aligned} M\dot{\mathbf{v}}_h(t) &= L\mathbf{v}_h(t) - D^T\mathbf{p}_h(t) + B_0 u(t), \\ 0 &= -D\mathbf{v}_h(t) + B_2 u(t), \end{aligned} \tag{1.2}$$

where $\mathbf{v}_h(t) \in \mathbb{R}^{n_\mathbf{v}}$ and $\mathbf{p}_h(t) \in \mathbb{R}^{n_\mathbf{p}}$ are the semidiscretized vectors of velocities and pressures, respectively, $M \in \mathbb{R}^{n_\mathbf{v},n_\mathbf{v}}$ is the symmetric, positive definite mass matrix and $L \in \mathbb{R}^{n_\mathbf{v},n_\mathbf{v}}$ is the discrete Laplace operator times ϑ. The matrices $D^T \in \mathbb{R}^{n_\mathbf{v},n_\mathbf{p}}$ and $D \in \mathbb{R}^{n_\mathbf{p},n_\mathbf{v}}$ are the discrete gradient and divergence operators. Due to the non-uniqueness of the pressure, the matrix D has a rank defect which in most spatial discretization methods is equal to one. In this case instead of D we can take a full row rank matrix obtained from D by discarding the last row. Therefore, in the following we will assume without loss of generality that D has full row rank. The matrices $B_0 \in \mathbb{R}^{n_\mathbf{v},m}$, $B_2 \in \mathbb{R}^{n_\mathbf{p},m}$ and the control input $u(t) \in \mathbb{R}^m$ are resulting from the boundary conditions and external forces.

System (1.2) together with the output equation $y(t) = C_0\mathbf{v}_h(t) + C_2\mathbf{p}_h(t)$ can be rewritten as a descriptor system

$$\begin{aligned} E\,\dot{x}(t) &= A\,x(t) + B\,u(t), \\ y(t) &= C\,x(t) \end{aligned} \tag{1.3}$$

where E, $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$, $C \in \mathbb{R}^{q,n}$, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the control input, $y(t) \in \mathbb{R}^q$ is the output. The order $n = n_\mathbf{v} + n_\mathbf{p}$ of system (1.3) depends on the fineness of the discretization and is usually very large, whereas the number m of inputs and the number q of outputs are small. Note that the matrices E and A in (1.3) are sparse and have a special block structure.

Simulation, control and optimization of large-scale sparse dynamical systems arising from semidiscretization of partial differential equations become prohibitive because of computational complexity and storage requirements. This motivates model order reduction that consists in an approximation of the descriptor system (1.3) by a reduced order system

$$\begin{aligned} \widetilde{E}\,\dot{\widetilde{x}}(t) &= \widetilde{A}\,\widetilde{x}(t) + \widetilde{B}\,u(t), \quad \widetilde{x}(0) = \widetilde{x}^0, \\ \widetilde{y}(t) &= \widetilde{C}\,\widetilde{x}(t), \end{aligned} \tag{1.4}$$

where $\widetilde{E}$, $\widetilde{A} \in \mathbb{R}^{\ell,\ell}$, $\widetilde{B} \in \mathbb{R}^{\ell,m}$, $\widetilde{C} \in \mathbb{R}^{q,\ell}$ and the order ℓ of this system is much smaller than the order n of (1.3). Note that systems (1.3) and (1.4) have the same input $u(t)$. One requires that the approximate system (1.4) preserves properties of the original system (1.3) like regularity

and stability. It is also desirable to estimate how well system (1.3) is approximated by (1.4). Moreover, the computation of the reduced order system should be numerically stable and efficient. Among other things the model reduction method has to be suitable for large-scale and sparse systems.

There exist various model reduction approaches for standard ($E = I$) state space systems such as balanced truncation [18, 24, 32, 37, 38], singular perturbation approximation [22], optimal Hankel norm approximation [14], proper orthogonal decomposition [30, 39] and moment matching approximation [1, 12]. Unfortunately, there is no general approach that can be considered as optimal. Surveys on system approximation and model reduction can be found in [1, 10].

Model reduction of descriptor systems based on Padé approximation via the Lanczos process has been considered in [9, 11, 13, 16]. This approach consists in computing the Krylov subspaces and projecting the dynamical system onto these subspaces. Krylov subspace methods are attractive for large-scale sparse systems, since only matrix-vector multiplications are required. Drawbacks of this technique are that there is no approximation error bound for the reduced order system and stability is not necessarily preserved.

The balanced truncation approach [18, 24, 29, 32, 37, 38] related to the controllability and observability Gramians is free from these disadvantages. Balanced truncation methods are based on transforming the dynamical system to a balanced form whose controllability and observability Gramians become diagonal and equal, together with truncation of states that are both difficult to reach and to observe. The diagonal elements of the transformed Gramians are known as the Hankel singular values of the dynamical system, and the truncated states correspond to the small Hankel singular values, see [24] for details. Important advantages of this approach are that asymptotic stability is preserved in the reduced order system and a priori bounds on the approximation error can be derived [8, 14].

The computation of the controllability and observability Gramians as well the Hankel singular values involves the solution of two Lyapunov equations. Recently, effective iterative methods based on the ADI method and the Smith method have been proposed [19, 20, 25, 26] to compute a low rank approximation for the solution of standard Lyapunov equations with large-scale sparse matrices.

The balanced truncation approach has been extended to descriptor systems in [35, 36]. The method proposed there is based on decoupling the descriptor system (1.3) into slow and fast subsystems [5] that correspond to the deflating subspaces of the pencil $\lambda E - A$ associated with the finite and infinite eigenvalues, respectively, and then reducing the order only of the slow subsystem. Thereby the fast subsystem remains unchanged. However, in many applications, including the semidiscretized Stokes equation, the order of the fast subsystem may be much larger than the order of the slow subsystem, see [35, Example 7.6]. In this paper we discuss how the order of the fast subsystem can also be reduced by using the balanced truncation technique.

Section 2 contains some basic concepts of model reduction via balanced truncation for descriptor systems. In particular, we consider generalizations of controllability and observability Gramians as well as Hankel singular values for descriptor systems. The latter play an important role in estimating the approximation error. In Section 3 we apply these results to the semidiscretized Stokes equation (1.2). We also discuss how the block structure of this equation may be used to reduce the computational effort. A numerical example is presented in Section 4 to illustrate the applicability and effectiveness of the proposed model reduction algorithms.

2 Model reduction for descriptor systems

In this subsection we briefly review some of the results from [35, 36].

Consider the continuous-time descriptor system (1.3). Assume that the matrix pencil $\lambda E - A$ is *regular*, that is, $\det(\lambda E - A) \neq 0$ for some $\lambda \in \mathbb{C}$. In this case $\lambda E - A$ can be reduced to the Weierstrass canonical form [34]. There exist nonsingular matrices W and T such that

$$E = W \begin{bmatrix} I_{n_f} & 0 \\ 0 & N \end{bmatrix} T \qquad \text{and} \qquad A = W \begin{bmatrix} J & 0 \\ 0 & I_{n_\infty} \end{bmatrix} T, \tag{2.1}$$

where I_k denotes the identity matrix of order k, J and N are matrices in Jordan canonical form and N is nilpotent with index of nilpotency ν. The numbers n_f and n_∞ are the dimensions of the deflating subspaces of $\lambda E - A$ corresponding to the finite and infinite eigenvalues, respectively, and ν is the *index* of the pencil $\lambda E - A$ and of the descriptor system (1.3). The matrices

$$P_r = T^{-1} \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} T, \qquad P_l = W \begin{bmatrix} I_{n_f} & 0 \\ 0 & 0 \end{bmatrix} W^{-1} \tag{2.2}$$

are the *spectral projections* onto the right and left deflating subspaces of the pencil $\lambda E - A$ corresponding to the finite eigenvalues.

Applying the Laplace transform [7] to the descriptor system (1.3), we find that

$$\mathbf{y}(s) = C(sE - A)^{-1}B\mathbf{u}(s) + C(sE - A)^{-1}Ex(0),$$

where $\mathbf{x}(s)$, $\mathbf{u}(s)$ and $\mathbf{y}(s)$ are the Laplace transforms of $x(t)$, $u(t)$ and $y(t)$, respectively. The rational matrix-valued function $\mathbf{G}(s) = C(sE - A)^{-1}B$, $s \in \mathbb{C}$, is called the *transfer function* of the continuous-time descriptor system (1.3). For $Ex(0) = 0$, $\mathbf{G}(s)$ gives the transfer relation between the Laplace transforms of the input $u(t)$ and the output $y(t)$. Using the Weierstrass canonical form (2.1), we have the following Laurent expansion at infinity for the transfer function

$$\mathbf{G}(s) = \sum_{k=-\infty}^{\infty} CF_kBs^{-k-1}, \tag{2.3}$$

where the matrices F_k have the form

$$F_k = \begin{cases} T^{-1} \begin{bmatrix} J^k & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, & k = 0, 1, 2 \ldots, \\ T^{-1} \begin{bmatrix} 0 & 0 \\ 0 & -N^{-k-1} \end{bmatrix} W^{-1}, & k = -1, -2, \ldots . \end{cases} \tag{2.4}$$

Note that $F_k = 0$ for $k < -\nu$, where ν is the index of the pencil $\lambda E - A$. The transfer function $\mathbf{G}(s)$ is called *proper* if it has no poles at infinity. Clearly, $\mathbf{G}(s)$ is proper if and only if $CF_kB = 0$ for $k < -1$.

For any rational matrix-valued function $\mathbf{G}(s)$, there exist matrices E, A, B and C such that $\mathbf{G}(s) = C(sE - A)^{-1}B$, see [5]. A continuous-time descriptor system (1.3) with these matrices is called a *realization* of $\mathbf{G}(s)$. We will also denote a realization of $\mathbf{G}(s)$ by $\mathbf{G} = [\, E,\, A,\, B,\, C\,]$ or by

$$\mathbf{G} = \left[\begin{array}{c|c} sE - A & B \\ \hline C & \end{array}\right].$$

Note that the realization of $\mathbf{G}(z)$ is, in general, not unique [5].

2.1 Controllability and observability Gramians

The descriptor system (1.3) is *asymptotically stable* if the pencil $\lambda E - A$ is *c-stable*, i.e, it is regular and all the finite eigenvalues of $\lambda E - A$ lie in the open left half-plane, see [5]. In this case the integrals

$$\mathcal{G}_{pc} = \int_0^\infty \mathcal{F}(t) B B^T \mathcal{F}^T(t)\, dt$$

and

$$\mathcal{G}_{po} = \int_0^\infty \mathcal{F}^T(t) C^T C \mathcal{F}(t)\, dt$$

exist, where $\mathcal{F}(t)$ is the *fundamental solution matrix* of system (1.3) given by

$$\mathcal{F}(t) = T^{-1} \begin{bmatrix} e^{tJ} & 0 \\ 0 & 0 \end{bmatrix} W^{-1}$$

and T, W and J are as in (2.1). The matrix $\mathcal{G}_{pc}$ is called the *proper controllability Gramian* and the matrix $\mathcal{G}_{po}$ is called the *proper observability Gramian* of the continuous-time descriptor system (1.3). The *improper controllability Gramian* of system (1.3) is defined by

$$\mathcal{G}_{ic} = \sum_{k=-\nu}^{-1} F_k B B^T F_k^T,$$

and the *improper observability Gramian* of system (1.3) is given by

$$\mathcal{G}_{io} = \sum_{k=-\nu}^{-1} F_k^T C^T C F_k,$$

where the matrices F_k are as in (2.4). Note that the improper controllability and observability Gramians $\mathcal{G}_{ic}$ and $\mathcal{G}_{io}$ are, up to the sign, the same as those defined in [3, 36]. If $E = I$, then $\mathcal{G}_{pc}$ and $\mathcal{G}_{po}$ are the usual controllability and observability Gramians for the standard state space system [14].

It has been shown in [35] that the proper controllability and observability Gramians are the unique symmetric, positive semidefinite solutions of the *projected generalized continuous-time Lyapunov equations*

$$\begin{aligned} &E\mathcal{G}_{pc}A^T + A\mathcal{G}_{pc}E^T = -P_l B B^T P_l^T, \\ &\mathcal{G}_{pc} = P_r \mathcal{G}_{pc} \end{aligned} \tag{2.5}$$

and

$$\begin{aligned} &E^T\mathcal{G}_{po}A + A^T\mathcal{G}_{po}E = -P_r^T C^T C P_r, \\ &\mathcal{G}_{po} = \mathcal{G}_{po} P_l, \end{aligned} \tag{2.6}$$

respectively, where P_l and P_r are given in (2.2). Furthermore, the improper controllability and observability Gramians are the unique symmetric, positive semidefinite solutions of the *projected generalized discrete-time Lyapunov equations*

$$\begin{aligned} &A\mathcal{G}_{ic}A^T - E\mathcal{G}_{ic}E^T = (I - P_l) B B^T (I - P_l)^T, \\ &P_r \mathcal{G}_{ic} = 0 \end{aligned} \tag{2.7}$$

and

$$\begin{aligned} &A^T\mathcal{G}_{io}A - E^T\mathcal{G}_{io}E = (I - P_r)^T C^T C (I - P_r), \\ &\mathcal{G}_{io} P_l = 0, \end{aligned} \tag{2.8}$$

respectively.

Recall that the descriptor system (1.3) is called *completely controllable* if for all $\lambda \in \mathbb{C}$,

$$\operatorname{rank}[\,\lambda E - A,\, B] = n \qquad \text{and} \qquad \operatorname{rank}[\,E,\, B] = n.$$

System (1.3) is called *completely observable* if

$$\operatorname{rank}\begin{bmatrix} \lambda E - A \\ C \end{bmatrix} = n \quad \text{for all } \lambda \in \mathbb{C} \qquad \text{and} \qquad \operatorname{rank}\begin{bmatrix} E \\ C \end{bmatrix} = n.$$

The controllability and observability Gramians can be used to characterize complete controllability and complete observability properties of system (1.3).

Theorem 2.1. [3, 35] *Consider the descriptor system* (1.3)*. Assume that $\lambda E - A$ is c-stable.*

1. *System* (1.3) *is completely controllable if and only if the proper controllability Gramian $\mathcal{G}_{pc}$ is positive definite on* $\operatorname{Im} P_r^T$ *and the improper controllability Gramian $\mathcal{G}_{ic}$ is positive definite on* $\operatorname{Ker} P_r^T$*.*

2. *System* (1.3) *is completely observable if and only if the proper observability Gramian $\mathcal{G}_{po}$ is positive definite on* $\operatorname{Im} P_l$ *and the improper observability Gramian $\mathcal{G}_{io}$ is positive definite on* $\operatorname{Ker} P_l$*.*

2.2 Hankel singular values

Similar to standard state space systems [14], the controllability and observability Gramians can be used to define Hankel singular values for the descriptor system (1.3) that are of great importance in model reduction via balanced truncation.

Consider the matrices $\mathcal{G}_{pc}E^T\mathcal{G}_{po}E$ and $\mathcal{G}_{ic}A^T\mathcal{G}_{io}A$. These matrices play the same role for descriptor systems as the product of the controllability and observability Gramians for standard state space systems [14]. It has been shown in [35, 36] that all the eigenvalues of $\mathcal{G}_{pc}E^T\mathcal{G}_{po}E$ and $\mathcal{G}_{ic}A^T\mathcal{G}_{io}A$ are real and non-negative.

Definition 2.2. Let $\lambda E - A$ be a c-stable pencil and let n_f and n_∞ be the dimensions of the deflating subspaces of $\lambda E - A$ corresponding to the finite and infinite eigenvalues, respectively. The square roots of the n_f largest eigenvalues of the matrix $\mathcal{G}_{pc}E^T\mathcal{G}_{po}E$, denoted by ς_j, are called the *proper Hankel singular values* of the descriptor system (1.3). The square roots of the n_∞ largest eigenvalues of the matrix $\mathcal{G}_{ic}A^T\mathcal{G}_{io}A$, denoted by θ_j, are called the *improper Hankel singular values* of system (1.3).

The proper and improper Hankel singular values together form the set of Hankel singular values of the descriptor system (1.3). For $E = I$, the proper Hankel singular values are the classical Hankel singular values of the standard state space system [14].

Since the proper and improper controllability and observability Gramians are symmetric and positive semidefinite, there exist full rank factorizations

$$\begin{aligned} \mathcal{G}_{pc} &= R_p R_p^T, & \mathcal{G}_{po} &= L_p^T L_p, \\ \mathcal{G}_{ic} &= R_i R_i^T, & \mathcal{G}_{io} &= L_i^T L_i, \end{aligned} \tag{2.9}$$

where the matrices R_p, L_p, R_i and L_i are full rank Cholesky factors [17]. The following lemma gives a connection between the proper and improper Hankel singular values and the standard singular values of the matrices L_pER_p and L_iAR_i.

Lemma 2.3. [36] *Let $\lambda E - A$ be a c-stable pencil. Consider the full rank factorizations* (2.9) *of the Gramians of the descriptor system* (1.3)*. The non-zero proper Hankel singular values of* (1.3) *are the non-zero singular values of the matrix $L_p E R_p$, while the non-zero improper Hankel singular values of* (1.3) *are the non-zero singular values of the matrix $L_i A R_i$.*

2.3 Balanced truncation

As mentioned above, for a given transfer function $\mathbf{G}(s)$, there are many different realizations. Here we are interested only in particular realizations that are most useful in model reduction.

Definition 2.4. A realization $[\, E, A, B, C \,]$ of the transfer function $\mathbf{G}(s)$ is called *minimal* if the dimension of the matrices E and A is small as possible.

The following theorem gives necessary and sufficient conditions for a realization of $\mathbf{G}(s)$ to be minimal.

Theorem 2.5. [5] *A realization $\mathbf{G} = [\, E, A, B, C \,]$ is minimal if and only if the descriptor system (1.3) is completely controllable and completely observable.*

From Theorems 2.1 and 2.5 we obtain the following result.

Corollary 2.6. *Consider the descriptor system* (1.3)*, where the pencil $\lambda E - A$ is c-stable. The following statements are equivalent:*

1. *the realization $[\, E, A, B, C \,]$ is minimal;*
2. $\operatorname{rank}(\mathcal{G}_{pc}) = \operatorname{rank}(\mathcal{G}_{po}) = \operatorname{rank}(\mathcal{G}_{pc} E^T \mathcal{G}_{po} E) = n_f$ *and*
   $\operatorname{rank}(\mathcal{G}_{ic}) = \operatorname{rank}(\mathcal{G}_{io}) = \operatorname{rank}(\mathcal{G}_{ic} A^T \mathcal{G}_{io} A) = n_\infty$*;*
3. *the proper and improper Hankel singular values of* (1.3) *are non-zero.*

Definition 2.7. A realization $\mathbf{G} = [\, E, A, B, C \,]$ with the c-stable pencil $\lambda E - A$ is called *balanced* if

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

with $\Sigma = \operatorname{diag}(\varsigma_1, \ldots, \varsigma_{n_f})$ and $\Theta = \operatorname{diag}(\theta_1, \ldots, \theta_{n_\infty})$.

We will show that for a minimal realization $[\, E, A, B, C \,]$ with the c-stable pencil $\lambda E - A$, there exist transformation matrices W_b and T_b such that the realization

$$[\, W_b^T E T_b,\ W_b^T A T_b,\ W_b^T B,\ C T_b \,] \tag{2.10}$$

is balanced.

Consider the full rank factorizations (2.9) of the controllability and observability Gramians. Let

$$L_p E R_p = U_p \Sigma_p V_p^T, \qquad L_i A R_i = U_i \Sigma_i V_i^T,$$

be the 'thin' singular value decompositions [15] of the matrices $L_p E R_p$ and $L_i A R_i$, where U_p, V_p, U_i, V_i have orthonormal columns, Σ_p and Σ_i are diagonal and nonsingular. By Lemma 2.3 and Corollary 2.6 we have that $\Sigma_p = \operatorname{diag}(\varsigma_1, \ldots, \varsigma_{n_f}) = \Sigma$ and $\Sigma_i = \operatorname{diag}(\theta_1, \ldots, \theta_{n_\infty}) = \Theta$. It follows from $\mathcal{G}_{pc} = P_r \mathcal{G}_{pc}$, $\mathcal{G}_{po} = \mathcal{G}_{po} P_l$, $P_r \mathcal{G}_{ic} = \mathcal{G}_{io} P_l = 0$ and $P_l E = E P_r$, $P_l A = A P_r$ that

$\mathcal{G}_{io}E\mathcal{G}_{pc} = \mathcal{G}_{po}E\mathcal{G}_{ic} = \mathcal{G}_{po}A\mathcal{G}_{pc} = 0$. Hence, $L_iER_p = L_pER_i = 0$ and $L_pAR_i = L_iAR_p = 0$. Consider now the matrices

$$W_b = \left[L_p^T U_p \Sigma^{-1/2},\ L_i^T U_i \Theta^{-1/2} \right], \qquad \check{W}_b = \left[ER_p V_p \Sigma^{-1/2},\ AR_i V_i \Theta^{-1/2} \right]. \tag{2.11}$$

We get

$$W_b^T \check{W}_b = \begin{bmatrix} \Sigma^{-1/2} U_p^T L_p E R_p V_p \Sigma^{-1/2} & \Sigma^{-1/2} U_p^T L_p A R_i V_i \Theta^{-1/2} \\ \Theta^{-1/2} U_i^T L_i E R_p V_p \Sigma^{-1/2} & \Theta^{-1/2} U_i^T L_i A R_i V_i \Theta^{-1/2} \end{bmatrix} = I_n,$$

i.e., W_b and $\check{W}_b$ are nonsingular and $W_b^{-1} = \check{W}_b^T$. Similarly, we can show that the matrices

$$T_b = \left[R_p V_p \Sigma^{-1/2},\ R_i V_i \Theta^{-1/2} \right], \qquad \check{T}_b = \left[E^T L_p^T U_p \Sigma^{-1/2},\ A^T L_i^T U_i \Theta^{-1/2} \right] \tag{2.12}$$

are also nonsingular and $T_b^{-1} = \check{T}_b^T$. The Gramians of the transformed system (2.10) with W_b and T_b as in (2.11) and (2.12), respectively, have the form

$$T_b^{-1} \mathcal{G}_{pc} T_b^{-T} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} = W_b^{-1} \mathcal{G}_{po} W_b^{-T},$$
$$T_b^{-1} \mathcal{G}_{ic} T_b^{-T} = \begin{bmatrix} 0 & 0 \\ 0 & \Theta \end{bmatrix} = W_b^{-1} \mathcal{G}_{po} W_b^{-T}.$$

Hence, W_b and T_b are the balancing transformation matrices and realization (2.10) is balanced.

It should be noted that just as for standard state space systems [14, 24], the balancing transformation for descriptor systems is not unique. Indeed, if W_b and T_b transform the descriptor system (1.3) to a balanced form, then for any diagonal matrix S with diagonal entries ± 1, the matrices W_bS and T_bS are also a balancing transformation.

Remark 2.8. For the matrices W_b and T_b as in (2.11) and (2.12), we have

$$E_b = W_b^T E T_b = \begin{bmatrix} \Sigma^{-1/2} U_p^T L_p E R_p V_p \Sigma^{-1/2} & \Sigma^{-1/2} U_p^T L_p E R_i V_i \Theta^{-1/2} \\ \Theta^{-1/2} U_i^T L_i E R_p V_p \Sigma^{-1/2} & \Theta^{-1/2} U_i^T L_i E R_i V_i \Theta^{-1/2} \end{bmatrix} = \begin{bmatrix} I_{n_f} & 0 \\ 0 & E_2 \end{bmatrix},$$
$$A_b = W_b^T A T_b = \begin{bmatrix} \Sigma^{-1/2} U_p^T L_p A R_p V_p \Sigma^{-1/2} & \Sigma^{-1/2} U_p^T L_p A R_i V_i \Theta^{-1/2} \\ \Theta^{-1/2} U_i^T L_i A R_p V_p \Sigma^{-1/2} & \Theta^{-1/2} U_i^T L_i A R_i V_i \Theta^{-1/2} \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & I_{n_\infty} \end{bmatrix},$$

where $E_2 = \Theta^{-1/2} U_i^T L_i E R_i V_i \Theta^{-1/2}$ is nilpotent and $A_1 = \Sigma^{-1/2} U_p^T L_p A R_p V_p \Sigma^{-1/2}$ is nonsingular. Thus, the pencil $\lambda E_b - A_b$ is in Weierstrass-like canonical form. It is regular, c-stable and has the same index as $\lambda E - A$.

If the descriptor system (1.3) is not minimal, then it has states that are uncontrollable or/and unobservable. These states correspond to the zero proper and improper Hankel singular values and can be truncated without changing the input-output relation in the system. Note that the number of non-zero improper Hankel singular values of (1.3) is equal to $\operatorname{rank}(\mathcal{G}_{ic} A^T \mathcal{G}_{io} A)$ which can in turn be estimated as

$$\operatorname{rank}(\mathcal{G}_{ic} A^T \mathcal{G}_{io} A) \leq \min(\nu m, \nu q, n_\infty).$$

This estimate shows that if the index ν of the pencil $\lambda E - A$ times the number m of inputs or the number q of outputs is much smaller than the dimension n_∞ of the deflating subspace

of $\lambda E - A$ corresponding to the infinite eigenvalues, then the order of system (1.3) can be reduced significantly.

Furthermore, taking into account the input-output energy characterization via the proper controllability and observability Gramians, see [35, 36], we conclude that the truncation of the states related to the small proper Hankel singular values does not change the system properties essentially.

Remark 2.9. Unfortunately, this does not hold for the improper Hankel singular values. If we truncate the states that correspond to the small improper Hankel singular values, then the pencil of the reduced order system may have no infinite eigenvalues or may get finite eigenvalues in the closed right half-plane, see such an example in [21]. As a result the approximation will be inaccurate.

In summary, we have the following algorithm which is a generalization of the *square root balanced truncation method* [18, 37] for the descriptor system (1.3).

Algorithm 2.1. *Generalized Square Root Balanced Truncation (GSRBT) method.*
Input: *A realization* $[\,E,\,A,\,B,\,C\,]$ *such that* $\lambda E - A$ *is c-stable.*
Output: *A reduced order system* $[\,\widetilde{E},\,\widetilde{A},\,\widetilde{B},\,\widetilde{C}\,]$.
1. *Compute the full rank Cholesky factors* R_p *and* L_p *of the proper controllability and observability Gramians* $\mathcal{G}_{pc} = R_p R_p^T$ *and* $\mathcal{G}_{po} = L_p^T L_p$ *that satisfy the projected generalized continuous-time Lyapunov equations* (2.5) *and* (2.6)*, respectively.*
2. *Compute the full rank Cholesky factors* R_i *and* L_i *of the improper controllability and observability Gramians* $\mathcal{G}_{ic} = R_i R_i^T$ *and* $\mathcal{G}_{io} = L_i^T L_i$ *that satisfy the projected generalized discrete-time Lyapunov equations* (2.7) *and* (2.8)*, respectively.*
3a. *Compute the 'thin' singular value decomposition*

$$L_p E R_p = [\,U_1,\,U_2\,] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} [\,V_1,\,V_2\,]^T, \tag{2.13}$$

where the matrices $[\,U_1,\,U_2\,]$ *and* $[\,V_1,\,V_2\,]$ *have orthonormal columns,* $\Sigma_1 = \mathrm{diag}(\varsigma_1, \ldots, \varsigma_{\ell_f})$ *and* $\Sigma_2 = \mathrm{diag}(\varsigma_{\ell_f+1}, \ldots, \varsigma_{r_p})$ *with* $r_p = \mathrm{rank}(L_p E R_p)$ *and* $\varsigma_1 \geq \ldots \geq \varsigma_{\ell_f} > \varsigma_{\ell_f+1} \geq \ldots \geq \varsigma_{r_p}$.
3b. *Compute the 'thin' singular value decomposition*

$$L_i A R_i = U_3 \Theta_3 V_3^T, \tag{2.14}$$

where U_3 *and* V_3 *have orthonormal columns,* $\Theta_3 = \mathrm{diag}(\theta_1, \ldots, \theta_{\ell_\infty})$ *with* $\ell_\infty = \mathrm{rank}(L_i A R_i)$.
4. *Compute the matrices*

$$W_\ell = [\,L_p^T U_1 \Sigma_1^{-1/2},\;\; L_i^T U_3 \Theta_3^{-1/2}\,], \qquad T_\ell = [\,R_p V_1 \Sigma_1^{-1/2},\;\; R_i V_3 \Theta_3^{-1/2}\,]. \tag{2.15}$$

5. *Compute the reduced order system* $[\,\widetilde{E},\,\widetilde{A},\,\widetilde{B},\,\widetilde{C}\,] = [\,W_\ell^T E T_\ell,\, W_\ell^T A T_\ell,\, W_\ell^T B,\, C T_\ell\,]$.

If the original system (1.3) is highly unbalanced or if the deflating subspaces of the pencil $\lambda E - A$ corresponding to the finite and infinite eigenvalues are close, then the projection matrices W_ℓ and T_ℓ are ill-conditioned. To avoid accuracy loss in the reduced system, a *square root balancing free method* has been proposed in [38] for standard state space systems. This method can be generalized for descriptor systems as follows.

Algorithm 2.2. *Generalized Square Root Balancing Free (GSRBF) method.*
Input: *A realization* $[\,E,\,A,\,B,\,C\,]$ *such that* $\lambda E - A$ *is c-stable.*
Output: *A reduced order system* $[\,\breve{E},\,\breve{A},\,\breve{B},\,\breve{C}\,]$.
1. *Compute the full rank Cholesky factors* R_p *and* L_p *of the proper controllability and observability Gramians* $\mathcal{G}_{pc} = R_p R_p^T$ *and* $\mathcal{G}_{po} = L_p^T L_p$ *that satisfy* (2.5) *and* (2.6)*, respectively.*
2. *Compute the full rank Cholesky factors* R_i *and* L_i *of the improper controllability and observability Gramians* $\mathcal{G}_{ic} = R_i R_i^T$ *and* $\mathcal{G}_{io} = L_i^T L_i$ *that satisfy* (2.7) *and* (2.8)*, respectively.*
3. *Compute the 'thin' singular value decompositions* (2.13) *and* (2.14).
4. *Compute the 'economy size' QR decompositions*

$$[\,R_p V_1,\, R_i V_3\,] = Q_R R, \qquad [\,L_p^T U_1,\, L_i^T U_3\,] = Q_L L, \tag{2.16}$$

where $Q_R,\, Q_L \in \mathbb{R}^{n,\ell}$ *have orthonormal columns and* $R,\, L \in \mathbb{R}^{\ell,\ell}$ *are upper triangular and nonsingular.*
5. *Compute the reduced order system* $[\,\breve{E},\,\breve{A},\,\breve{B},\,\breve{C}\,] = [\,Q_L^T E Q_R,\, Q_L^T A Q_R,\, Q_L^T B,\, C Q_R\,]$.

The GSRBT and GSRBF methods are mathematically equivalent in the sense that they return reduced systems with the same transfer function. Indeed, taking into account (2.15) and (2.16), we obtain that

$$Q_L = [\,L_p^T U_1,\, L_i^T U_3\,] L^{-1} = W_\ell \begin{bmatrix} \Sigma_1^{1/2} & 0 \\ 0 & \Theta_3^{1/2} \end{bmatrix} L^{-1} = W_\ell S_1,$$
$$Q_R = [\,R_p V_1,\; R_i^T V_3\,] R^{-1} \;=\; T_\ell \begin{bmatrix} \Sigma_1^{1/2} & 0 \\ 0 & \Theta_3^{1/2} \end{bmatrix} R^{-1} = T_\ell S_2,$$

where $S_1 = \mathrm{diag}(\Sigma_1^{1/2}, \Theta_3^{1/2}) L^{-1}$ and $S_2 = \mathrm{diag}(\Sigma_1^{1/2}, \Theta_3^{1/2}) R^{-1}$ are nonsingular. Then

$$\breve{\mathbf{G}}(s) = \breve{C}(s\breve{E} - \breve{A})^{-1}\breve{B} = \widetilde{C} S_2 (s S_1^T \widetilde{E} S_2 - S_1^T \widetilde{A} S_2)^{-1} S_1^T \widetilde{B} = \widetilde{\mathbf{G}}(s).$$

Since the projection matrices Q_L and Q_R computed by the GSRBF method have orthonormal columns, they may be significantly better conditioned than the projection matrices W_ℓ and T_ℓ computed by the GSRBT method. It should be noted that the realization $[\,\breve{E},\,\breve{A},\,\breve{B},\,\breve{C}\,]$ is, in general, not balanced and the pencil $\lambda\breve{E} - \breve{A}$ is not in Weierstrass-like canonical form any more.

2.4 Approximation error

Computing the reduced order descriptor system via balanced truncation can be interpreted as transforming at first the system (1.3) to the block diagonal form

$$\left[\begin{array}{c|c} \breve{W}(sE - A)\breve{T} & \breve{W}B \\ \hline C\breve{T} & \end{array}\right] = \left[\begin{array}{cc|c} sE_f - A_f & & B_f \\ & sE_\infty - A_\infty & B_\infty \\ \hline C_f & C_\infty & \end{array}\right],$$

where $\breve{W}$ and $\breve{T}$ are nonsingular, the pencil $\lambda E_f - A_f$ has the finite eigenvalues only and all eigenvalues of $\lambda E_\infty - A_\infty$ are infinite, and then reducing the order of the subsystems $[\,E_f,\, A_f,\, B_f,\, C_f\,]$ and $[\,E_\infty,\, A_\infty,\, B_\infty,\, C_\infty\,]$ separately. Clearly, the reduced order system (1.4) is asymptotically stable and minimal.

The described decoupling of system matrices is equivalent to the additive decomposition of the transfer function as $\mathbf{G}(s) = \mathbf{G}_{sp}(s) + \mathbf{P}(s)$, where $\mathbf{G}_{sp}(s) = C_f(sE_f - A_f)^{-1}B_f$ is the strictly proper part and $\mathbf{P}(s) = C_\infty(sE_\infty - A_\infty)^{-1}B_\infty$ is the polynomial part of $\mathbf{G}(s)$. The reduced order system (1.4) has the transfer function $\widetilde{\mathbf{G}}(s) = \widetilde{\mathbf{G}}_{sp}(s) + \widetilde{\mathbf{P}}(s)$, where $\widetilde{\mathbf{G}}_{sp}(s) = \widetilde{C}_f(s\widetilde{E}_f - \widetilde{A}_f)^{-1}\widetilde{B}_f$ and $\widetilde{\mathbf{P}}(s) = \widetilde{C}_\infty(s\widetilde{E}_\infty - \widetilde{A}_\infty)^{-1}\widetilde{B}_\infty$ are the transfer functions of the reduced subsystems. For the subsystem $[\,E_f,\, A_f,\, B_f,\, C_f\,]$ with nonsingular E_f, we have the following upper bound on the $\mathbb{H}_\infty$-norm of the absolute error

$$\|\mathbf{G}_{sp}(s) - \widetilde{\mathbf{G}}_{sp}(s)\|_{\mathbb{H}_\infty} := \sup_{\omega\in\mathbb{R}} \|\mathbf{G}_{sp}(i\omega) - \widetilde{\mathbf{G}}_{sp}(i\omega)\|_2 \leq 2(\varsigma_{\ell_f+1} + \ldots + \varsigma_{n_f})$$

that can be derived as in [8, 14]. Here $\|\cdot\|_2$ denotes the spectral matrix norm.

Reducing the order of the subsystem $[\,E_\infty,\, A_\infty,\, B_\infty,\, C_\infty\,]$ is equivalent to the balanced model reduction of the discrete-time system

$$\begin{aligned} A_\infty z_{k+1} &= E_\infty z_k + B_\infty \eta_k, \\ w_k &= C_\infty z_k \end{aligned}$$

with the nonsingular matrix A_∞. The Hankel singular values of this system are just the improper Hankel singular values of (1.3). Since we truncate only the states corresponding to the zero improper Hankel singular values, the equality $\mathbf{P}(s) = \widetilde{\mathbf{P}}(s)$ holds and the index of the reduced order system is equal to $\deg(\mathbf{P}) + 1$, where $\deg(\mathbf{P})$ denotes the degree of the polynomial $\mathbf{P}(s)$, or, equivalently, the multiplicity of the pole at infinity of the transfer function $\mathbf{G}(s)$. In this case the error system $\mathbf{G}(s) - \widetilde{\mathbf{G}}(s) = \mathbf{G}_{sp}(s) - \widetilde{\mathbf{G}}_{sp}(s)$ is strictly proper, and we have the following $\mathbb{H}_\infty$-norm error bound

$$\|\mathbf{G}(s) - \widetilde{\mathbf{G}}(s)\|_{\mathbb{H}_\infty} \leq 2(\varsigma_{\ell_f+1} + \ldots + \varsigma_{n_f}).$$

Existence of this error bound is an important property of the balanced truncation model reduction approach for descriptor systems. It makes this approach preferable compared, for instance, to moment matching techniques [9, 11, 13, 16] or the proper orthogonal decomposition method [30, 39].

3 Balanced truncation for the Stokes equation

As follows from the previous considerations, computing the spectral projections P_l and P_r as well as solving the projected generalized Lyapunov equations take the highest computational expenses. To compute the full rank factors of the Gramians we can use the generalized Schur-Hammarling method as proposed in [35]. Since this method is based on computing the generalized upper triangular form [6] of the pencil $\lambda E - A$, it costs $O(n^3)$ flops and has the memory complexity $O(n^2)$. Thus, the generalized Schur-Hammarling method can be applied to problems of small or medium size only. Moreover, this method does not take into account the sparsity or any structure of the system. In this section we will discuss how the block structure and sparsity of the semidiscretized Stokes equation (1.2) can be used to reduce the computational cost and memory requirements.

Consider the semidiscretized Stokes equation (1.2), where M is symmetric, positive definite and D has full row rank. Assume that (1.2) is asymptotically stable. Computing a Cholesky factorization $M = U_M^T U_M$, we define new matrices $A_{11} = U_M^{-T} L U_M^{-1}$, $A_{12} = -U_M^{-T} D^T$, $B_1 = U_M^{-T} B_0$ and $C_1 = C_0 U_M^{-1}$. Then system (1.2) together with the output equation

$y(t) = C_0\mathbf{v}_h(t) + C_2\mathbf{p}_h(t)$ can be rewritten as the descriptor system (1.3) with matrix coefficients

$$E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \qquad C = [\, C_1,\, C_2\,], \tag{3.1}$$

where $A_{12} \in \mathbb{R}^{n_{\mathbf{v}},n_{\mathbf{p}}}$ has full column rank and the pencil $\lambda E - A$ is c-stable. Note that these matrices will never be computed explicitly, since the results would be dense matrices unless M is diagonal.

3.1 Computing the projectors P_r and P_l

To compute the spectral projections P_r and P_l onto the right and left deflating subspaces of $\lambda E - A$ corresponding to the finite eigenvalues we use the canonical projection technique proposed in [23]. Let

$$\begin{aligned} &E_0^r = E, \quad A_0^r = -A, \qquad E_{k+1}^r = E_k^r + A_k^r Q_k^r, \quad A_{k+1}^r = A_k^r(I - Q_k^r), \\ &E_0^l = E, \quad A_0^l = -A, \qquad E_{k+1}^l = E_k^l + Q_k^l A_k^l, \quad A_{k+1}^l = (I - Q_k^l)A_k^l, \end{aligned} \tag{3.2}$$

where Q_k^r and $(Q_k^l)^T$ are projections onto $\operatorname{Ker} E_k^r$ and $\operatorname{Ker} (E_k^l)^T$, respectively. Since the matrix A_{12} has full column rank, the matrix $A_{12}^T A_{12}$ is nonsingular and the pencil $\lambda E - A$ is of index two. In this case the matrices E_2^r and E_2^l are nonsingular and the projections Q_1^r and Q_1^l can be chosen such that $Q_1^r = Q_1^r(E_2^r)^{-1}A_1^r$ and $Q_1^l = A_1^l(E_2^l)^{-1}Q_1^l$. Then the spectral projections P_r and P_l onto the right and left deflating subspaces of $\lambda E - A$ corresponding to the finite eigenvalues can be computed as

$$P_r = \left(I - Q_0^r(I - Q_1^r)(E_2^r)^{-1}A_0^r\right)(I - Q_1^r), \tag{3.3}$$

$$P_l = (I - Q_1^l)\left(I - A_0^l(E_2^l)^{-1}(I - Q_1^l)Q_0^l\right), \tag{3.4}$$

see [23] for details.

For the pencil $\lambda E - A$ with matrices E and A as in (3.1), we have

$$Q_0^r = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}, \qquad E_1^r = \begin{bmatrix} I & -A_{12} \\ 0 & 0 \end{bmatrix}, \qquad A_1^r = \begin{bmatrix} -A_{11} & 0 \\ -A_{12}^T & 0 \end{bmatrix}.$$

The projection Q_1^r onto $\operatorname{Ker} E_1^r$ has the form

$$Q_1^r = \begin{bmatrix} A_{12}(A_{12}^T A_{12})^{-1}A_{12}^T & 0 \\ (A_{12}^T A_{12})^{-1}A_{12}^T & 0 \end{bmatrix}$$

It is easy to verify that the matrix

$$E_2^r = \begin{bmatrix} I - A_{11}A_{12}(A_{12}^T A_{12})^{-1}A_{12}^T & -A_{12} \\ -A_{12}^T & 0 \end{bmatrix}$$

is nonsingular and $Q_1^r = Q_1^r(E_2^r)^{-1}A_1^r$. Therefore, from (3.3) we obtain that

$$P_r = \begin{bmatrix} \Pi & 0 \\ -(A_{12}^T A_{12})^{-1}A_{12}^T A_{11}\Pi & 0 \end{bmatrix}, \tag{3.5}$$

where $\Pi = I - A_{12}(A_{12}^T A_{12})^{-1} A_{12}^T$ is the orthogonal projection onto $\operatorname{Ker} A_{12}^T$ along $\operatorname{Im} A_{12}$. Analogously, we find from (3.2) and (3.4) that

$$P_l = \begin{bmatrix} \Pi & -\Pi A_{11} A_{12}(A_{12}^T A_{12})^{-1} \\ 0 & 0 \end{bmatrix}.$$

Note that if A is symmetric, then $P_r = P_l^T$.

3.2 Computing the proper controllability and observability Gramians

Consider now the projected generalized continuous-time Lyapunov equation (2.5), where the pencil $\lambda E - A$ is c-stable. Let the proper controllability Gramian

$$\mathcal{G}_{pc} = \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{bmatrix}, \tag{3.6}$$

with $X_{11} = X_{11}^T$ and $X_{22} = X_{22}^T$ be partitioned in blocks conformally to E and A in (3.1). Using (3.1), (3.5) and (3.6) we obtain from the first equation in (2.5) that

$$\begin{aligned} &A_{11}X_{11} + X_{11}A_{11}^T + A_{12}X_{12}^T + X_{12}A_{12}^T = -\Pi B_{12} B_{12}^T \Pi, \\ &X_{11}A_{12} = 0, \end{aligned} \tag{3.7}$$

where $B_{12} = B_1 - A_{11}A_{12}(A_{12}^T A_{12})^{-1} B_2$. From the equations $\mathcal{G}_{pc} = P_r \mathcal{G}_{pc} = P_r \mathcal{G}_{pc} P_r^T$ it follows that

$$\begin{aligned} &X_{11} = \Pi X_{11}, \qquad X_{12} = -X_{11}A_{11}^T A_{12}(A_{12}^T A_{12})^{-1}, \\ &X_{22} = (A_{12}^T A_{12})^{-1} A_{12}^T A_{11} X_{11} A_{11}^T A_{12}(A_{12}^T A_{12})^{-1}. \end{aligned} \tag{3.8}$$

Combining (3.7) and (3.8), we find that X_{11} satisfies the projected Lyapunov equation

$$\Pi A_{11} X_{11} + X_{11} A_{11}^T \Pi = -\Pi B_{12} B_{12}^T \Pi, \qquad X_{11} = \Pi X_{11}. \tag{3.9}$$

This equation is equivalent to the projected Lyapunov equation

$$\Pi A_{11} \Pi X_{11} + X_{11} \Pi A_{11}^T \Pi = -\Pi B_{12} B_{12}^T \Pi \tag{3.10}$$

in the sense that (3.9) and (3.10) have the same unique symmetric, positive semidefinite solution.

Let R_1 be a full column rank Cholesky factor of $X_{11} = R_1 R_1^T$. Then the proper controllability Gramian of the semidiscretized Stokes equation (1.2) can be computed in factored form $\mathcal{G}_{pc} = R_p R_p^T$, where

$$R_p = \begin{bmatrix} R_1 \\ -(A_{12}^T A_{12})^{-1} A_{12}^T A_{11} R_1 \end{bmatrix}. \tag{3.11}$$

Analogously, we obtain that the proper observability Gramian of system (1.2) has the form $\mathcal{G}_{po} = L_p^T L_p$ with

$$L_p = [\, L_1, \;\; -L_1 A_{11} A_{12}(A_{12}^T A_{12})^{-1} \,], \tag{3.12}$$

where L_1 is a full row rank Cholesky factor of the solution $Y_{11} = L_1^T L_1$ of the projected continuous-time Lyapunov equation

$$\Pi A_{11}^T \Pi Y_{11} + Y_{11} \Pi A_{11} \Pi = -\Pi C_{12}^T C_{12} \Pi. \tag{3.13}$$

Here $C_{12} = C_1 - C_2(A_{12}^T A_{12})^{-1} A_{12}^T A_{11}$. Using (3.1), (3.11) and (3.12), we find $L_p E R_p = L_1 R_1$. Thus, the proper Hankel singular values of the semidiscretized Stokes equation (1.2) can be computed from the singular value decomposition of the matrix $L_1 R_1$.

We will now discuss the numerical solution of the projected Lyapunov equations (3.10) and (3.13). Since these equations have the same structure, we restrict ourself to consideration of equation (3.10) only.

Let V be a matrix whose columns form an orthonormal basis of $\operatorname{Im}\Pi = \operatorname{Ker} A_{12}^T$. Then the projector Π is represented as $\Pi = VV^T$. Multiplying equation (3.10) from the left by V^T and from the right by V, we obtain that the solution of (3.10) is given by $X_{11} = VX_0V^T$, where X_0 satisfies the Lyapunov equation

$$V^T A_{11} V X_0 + X_0 V^T A_{11}^T V = -V^T B_{12} B_{12}^T V. \tag{3.14}$$

If the pencil $\lambda E - A$ is c-stable, then all eigenvalues of the matrix $V^T A_{11} V$ have negative real part and, hence, equation (3.14) has a unique symmetric positive definite solution X_0. It was observed that in many cases the eigenvalues of the solution X_0 of (3.14) with a low rank right-hand side decay very fast, see [2, 27, 33]. Then the matrix X_0 and also $X_{11} = VX_0V^T$ can be well approximated by low rank matrices. In other words, it is possible to find a matrix X with a few columns such that $X = \Pi X$ and $X_{11} \approx XX^T$. The matrix X is referred to as the *low rank Cholesky factor* of X_{11}.

To compute the low rank Cholesky factor X of the solution of the Lyapunov equation (3.10) we do not need to compute the matrix V and solve the Lyapunov equation (3.14). Instead, we can apply a *low rank Cholesky factor alternating direction implicit* (LRCF-ADI) method [19, 20, 25, 26] to equation (3.10) directly. This method can be written as

$$\begin{aligned} X^{(1)} &= \sqrt{-2\tau_1}\,(\Pi A_{11}\Pi + \tau_1 I)^{-1}\Pi B_{12}, \qquad X_1 = X^{(1)}, \\ X^{(k)} &= \sqrt{\frac{\tau_k}{\tau_{k-1}}}\left(I - (\tau_{k-1}+\tau_k)(\Pi A_{11}\Pi + \tau_k I)^{-1}\right) X^{(k-1)}, \quad X_k = [\,X_{k-1},\ X^{(k)}\,], \end{aligned} \tag{3.15}$$

where $\tau_1, \ldots, \tau_k$ are real and negative ADI shift parameters that satisfy the ADI minimax problem

$$\{\tau_1, \ldots, \tau_k\} = \operatorname*{argmin}_{\tau_1,\ldots,\tau_k \in \mathbb{R}^-} \ \max_{t \in \operatorname{Sp}(\Pi A_{11}\Pi)\setminus\{0\}} \frac{|r_k(t)|}{|r_k(-t)|}.$$

Here $\mathbb{R}^- = (-\infty, 0)$, $r_k(t) = (t-\tau_1)\cdot\ldots\cdot(t-\tau_k)$ and $\operatorname{Sp}(\Pi A_{11}\Pi)$ denotes the spectrum of $\Pi A_{11}\Pi$. If the matrix $\Pi A_{11}\Pi$ is symmetric and lower and upper bounds on its non-zero spectrum are available, then the optimal parameters can be computed by a parameter selection procedure described in [40, Section 6.2]. Otherwise, we can calculate the suboptimal ADI shift parameters by using a heuristic algorithm [26, Algorithm 5.1] that is based on an Arnoldi iteration applied to $\Pi A_{11}\Pi$. For the stopping criteria and complexity of the LRCF-ADI method, see [19, 26, 28].

It should be noted that neither the matrix $\Pi A_{11}\Pi$ nor the matrices $(\Pi A_{11}\Pi + \tau_k I)^{-1}$ in (3.15) are computed explicitly. Instead, we exploit the product structure and solve linear systems of type $(\Pi A_{11}\Pi + \tau_k I)x = \Pi f$ or, equivalently, $\Pi(A_{11} + \tau_k I)\Pi x = \Pi f$. Since all non-zero eigenvalues of $\Pi A_{11}\Pi$ have negative real part and all τ_k are negative, these systems have unique solutions that can efficiently be computed by iterative Krylov subspace methods, see [31]. Note further that in exact arithmetic we have $X^{(k)} = \Pi X^{(k)}$. However, due to roundoff errors and approximate solution of linear systems it may happen that the columns

of $X^{(k)}$ drift off from $\operatorname{Im}\Pi$. To avoid this we need to correct the computed matrix $X^{(k)}$ by multiplication from the right by Π.

3.3 Computing the improper controllability and observability Gramians

We will now compute the improper controllability and observability Gramians $\mathcal{G}_{ic}$ and $\mathcal{G}_{io}$ of the semidiscretized Stokes equation (1.2). Let the improper controllability Gramian

$$\mathcal{G}_{ic} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^T & Z_{22} \end{bmatrix} \tag{3.16}$$

with $Z_{11} = Z_{11}^T$ and $Z_{22} = Z_{22}^T$ be partitioned in blocks conformally to E and A in (3.1). Substituting (3.1), (3.5) and (3.16) in the projected generalized discrete-time Lyapunov equation (2.7), we obtain that

$$\begin{aligned} A_{11}Z_{11}A_{11}^T + A_{12}Z_{12}^TA_{11}^T + A_{11}Z_{12}A_{12}^T + A_{12}Z_{22}A_{12}^T - Z_{11} &= B_{21}B_{21}^T, & (3.17)\\ A_{12}^TZ_{11}A_{11}^T + A_{12}^TZ_{12}A_{12}^T &= B_2B_{21}^T, & (3.18)\\ A_{12}^TZ_{11}A_{12} &= B_2B_2^T, & (3.19) \end{aligned}$$

where $B_{21} = B_1 - \Pi B_{12}$. Moreover, from $P_r\mathcal{G}_{ic} = \mathcal{G}_{ic}P_r^T = 0$ we have $\Pi Z_{11} = Z_{11}\Pi = 0$ and $\Pi Z_{12} = 0$. Then it follows from (3.18) and (3.19) that

$$\begin{aligned} Z_{11} &= A_{12}(A_{12}^TA_{12})^{-1}B_2B_2^T(A_{12}^TA_{12})^{-1}A_{12}^T,\\ Z_{12} &= A_{12}(A_{12}^TA_{12})^{-1}B_2B_{12}^TA_{12}(A_{12}^TA_{12})^{-1}. \end{aligned}$$

Finally, if we substitute Z_{11} and Z_{12} in (3.17) and multiply this equation by $(A_{12}^TA_{12})^{-1}A_{12}^T$ from the left and by $A_{12}(A_{12}^TA_{12})^{-1}$ from the right, we obtain that

$$Z_{22} = (A_{12}^TA_{12})^{-1}A_{12}^TB_{12}B_{12}^TA_{12}(A_{12}^TA_{12})^{-1} + (A_{12}^TA_{12})^{-1}B_2B_2^T(A_{12}^TA_{12})^{-1}.$$

Thus, the improper controllability Gramian of the semidiscretized Stokes equation (1.2) can be computed in factored form $\mathcal{G}_{ic} = R_iR_i^T$, where

$$R_i = \begin{bmatrix} A_{12}(A_{12}^TA_{12})^{-1}B_2 & 0 \\ (A_{12}^TA_{12})^{-1}A_{12}^TB_{12} & (A_{12}^TA_{12})^{-1}B_2 \end{bmatrix}. \tag{3.20}$$

Analogously, we obtain from the projected generalized discrete-time Lyapunov equation (2.8) that the improper observability Gramian of (1.2) has the form $\mathcal{G}_{io} = L_i^TL_i$, where

$$L_i = \begin{bmatrix} C_2(A_{12}^TA_{12})^{-1}A_{12}^T & C_{12}A_{12}(A_{12}^TA_{12})^{-1} \\ 0 & C_2(A_{12}^TA_{12})^{-1} \end{bmatrix}. \tag{3.21}$$

Note that the factors $R_i \in \mathbb{R}^{n,2m}$ and $L_i^T \in \mathbb{R}^{n,2q}$ in (3.20) and (3.21) are, in general, not of full rank, but they have only a few columns if m and q are small.

It follows from (3.1), (3.20) and (3.21) that

$$L_iAR_i = \begin{bmatrix} B_{11} & C_2(A_{12}^TA_{12})^{-1}B_2 \\ C_2(A_{12}^TA_{12})^{-1}B_2 & 0 \end{bmatrix}. \tag{3.22}$$

where $B_{11} = C_1A_{12}(A_{12}^TA_{12})^{-1}B_2 + C_2(A_{12}^TA_{12})^{-1}A_{12}^TB_{12}$. Hence, to determine the improper Hankel singular values of (1.2) we have to compute the singular value decomposition of the matrix $L_iAR_i \in \mathbb{R}^{2q,2m}$ as in (3.22).

3.4 Three model reduction methods for the Stokes equation

To reduce the order of the semidiscretized Stokes equation (1.2) we use Algorithm 2.1, where the full rank factors R_p and L_p as in (3.11) and (3.12) are replaced by the low rank factors.

Let $X = \Pi X$ and $Y = \Pi Y$ be low rank Cholesky factors of the solutions $X_{11} \approx XX^T$ and $Y_{11} \approx YY^T$ of the projected Lyapunov equations (3.10) and (3.13), respectively. Then the dominant proper Hankel singular values of (1.2) can be approximated by the dominant singular values of the matrix Y^TX. Consider the 'thin' singular value decompositions

$$Y^TX = [\,U_1,\, U_2\,] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} [\,V_1,\, V_2\,]^T, \tag{3.23}$$

$$\begin{bmatrix} B_{11} & C_2(A_{12}^TA_{12})^{-1}B_2 \\ C_2(A_{12}^TA_{12})^{-1}B_2 & 0 \end{bmatrix} = U_3\Theta_3V_3^T, \tag{3.24}$$

where the matrices $[\,U_1,\, U_2\,]$, $[\,V_1,\, V_2\,]$, U_3 and V_3 have orthonormal columns, Θ_3 is nonsingular, $\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_{\ell_f})$ and $\Sigma_2 = \mathrm{diag}(\sigma_{\ell_f+1}, \ldots, \sigma_r)$ with $\sigma_1 \geq \ldots \geq \sigma_{\ell_f} > \sigma_{\ell_f+1} \geq \ldots \geq \sigma_r$ and $r = \mathrm{rank}(Y^TX)$. Then the projection matrices W_ℓ and T_ℓ in (2.15) are rewritten as

$$W_\ell = [\,SYU_1\Sigma_1^{-1/2},\ L_i^TU_3\Theta_3^{-1/2}\,], \qquad T_\ell = [\,SXV_1\Sigma_1^{-1/2},\ R_iV_3\Theta_3^{-1/2}\,], \tag{3.25}$$

where $S = \begin{bmatrix} I \\ -(A_{12}^TA_{12})^{-1}A_{12}^TA_{11} \end{bmatrix}$, R_i and L_i are as in (3.20) and (3.21), respectively. Taking into account that $X = \Pi X$ and $Y = \Pi Y$, we compute the reduced order system $[\,\widetilde{E},\ \widetilde{A},\ \widetilde{B},\ \widetilde{C}\,] = [\,W_\ell^TET_\ell,\ W_\ell^TAT_\ell,\ W_\ell^TB,\ CT_\ell\,]$ with

$$\begin{gathered} \widetilde{E} = \begin{bmatrix} I_{\ell_f} & 0 \\ 0 & W_\infty^TC_2(A_{12}^TA_{12})^{-1}B_2T_\infty \end{bmatrix}, \qquad \widetilde{A} = \begin{bmatrix} W_f^TA_{11}T_f & 0 \\ 0 & I_{\ell_\infty} \end{bmatrix}, \\ \widetilde{B} = \begin{bmatrix} W_f^TB_{12} \\ \Theta_3T_\infty^T \end{bmatrix}, \qquad \widetilde{C} = \begin{bmatrix} C_{12}T_f,\ W_\infty\Theta_3^{1/2} \end{bmatrix}, \end{gathered} \tag{3.26}$$

and $W_f = YU_1\Sigma_1^{-1/2}$, $T_f = XV_1\Sigma_1^{-1/2}$, $W_\infty = [\,I_q,\ 0\,]U_3\Theta_3^{-1/2}$, $T_\infty = [\,I_m,\ 0\,]V_3\Theta_3^{-1/2}$.

In summary, we have the following algorithm that is a generalization of a *low rank square root method* [19, 29] for the semidiscretized Stokes equation (1.2).

Algorithm 3.1. *Generalized Low Rank Square Root (GLRSR) method for the semidiscretized Stokes equation.*
Input: *Matrices* M, L, D, B_0, B_2, C_0, C_2.
Output: *A reduced order system* $[\,\widetilde{E},\ \widetilde{A},\ \widetilde{B},\ \widetilde{C}\,]$.
0. *If* $M \neq I$, *then compute the Cholesky factorization* $M = U_M^TU_M$, *where* U_M *is upper triangular. Otherwise,* $U_M = I$.
1. *Use the LRCF-ADI method* (3.15) *to compute the low rank Cholesky factors* $X = \Pi X$ *and* $Y = \Pi Y$ *such that* XX^T *and* YY^T *are approximate solutions of the projected Lyapunov equations* (3.10) *and* (3.13), *respectively, where* $A_{11} = U_M^{-T}LU_M^{-1}$, $A_{12} = -U_M^{-T}D^T$ *and* $B_{12} = U_M^{-T}B_0 - A_{11}A_{12}(A_{12}^TA_{12})^{-1}B_2$, $C_{12} = C_0U_M^{-1} - C_2(A_{12}^TA_{12})^{-1}A_{12}^TA_{11}$.
2. *Compute the 'thin' singular value decompositions* (3.23) *and* (3.24).
3. *Compute the reduced order system* $[\,\widetilde{E},\ \widetilde{A},\ \widetilde{B},\ \widetilde{C}\,]$ *as in* (3.26).

Note that if A is symmetric and $B \neq C^T$ in (3.1), then the matrices $\widetilde{E}$ and $\widetilde{A}$ as in (3.26) are, in general, not symmetric. The reduced order system with the symmetric matrices $\widetilde{E}$ and $\widetilde{A}$ can be computed by applying the GLRSR method to a symmetrized system

$$\widehat{\mathbf{G}} = \left[\begin{array}{c|c} sE - A & \widehat{B} \\ \hline \widehat{B}^T & \end{array}\right], \qquad \widehat{B} = \left[\, B,\ C^T \,\right] \tag{3.27}$$

with more inputs and outputs but the same number of state variables. Such a system is square in the sense that it has the equal number $m+q$ of inputs and outputs, and its transfer function is symmetric, i.e., $\widehat{\mathbf{G}}(s) = \widehat{\mathbf{G}}^T(s)$. In this case the proper controllability Gramian $\widehat{\mathcal{G}}_{pc}$ of (3.27) is equal to the proper observability Gramian $\widehat{\mathcal{G}}_{po}$ and the improper controllability and observability Gramians $\widehat{\mathcal{G}}_{ic}$ and $\widehat{\mathcal{G}}_{io}$ of (3.27) are also equal.

Using the same computational technique as in Sections 3.2 and 3.3, we obtain that the full rank Cholesky factors $\widehat{R}_p$ and $\widehat{R}_i$ of the Gramians $\widehat{\mathcal{G}}_{pc} = \widehat{\mathcal{G}}_{po}$ and $\widehat{\mathcal{G}}_{ic} = \widehat{\mathcal{G}}_{io}$, respectively, have the form

$$\widehat{R}_p = \begin{bmatrix} \widehat{R} \\ -(A_{12}^T A_{12})^{-1} A_{12}^T A_{11} \widehat{R} \end{bmatrix}, \tag{3.28}$$

$$\widehat{R}_i = \begin{bmatrix} A_{12}(A_{12}^T A_{12})^{-1} \hat{B}_2 & 0 \\ (A_{12}^T A_{12})^{-1} A_{12}^T \widehat{B}_{12} & (A_{12}^T A_{12})^{-1} \widehat{B}_2 \end{bmatrix},$$

where $\widehat{B}_2 = [\, B_2,\ C_2^T \,]$, $\widehat{B}_{12} = [\, B_{12},\ C_{12}^T \,]$ and $\widehat{R}$ is a full rank Cholesky factor of the solution $\widehat{X}_{11} = \widehat{R}\widehat{R}^T$ of the projected Lyapunov equation

$$\Pi A_{11} \Pi \widehat{X}_{11} + \widehat{X}_{11} \Pi A_{11} \Pi = -\Pi \widehat{B}_{12} \widehat{B}_{12}^T \Pi. \tag{3.29}$$

Since $\widehat{R}_p^T E \widehat{R}_p = \widehat{R}^T \widehat{R}$ and $\widehat{R}_i^T A \widehat{R}_i$ are symmetric, the projection matrices W_ℓ and T_ℓ can be chosen as $W_\ell = T_\ell = [\, \widehat{R}_p \widehat{V}_1,\ \widehat{R}_i \widehat{V}_3 \,]$, where the columns of $\widehat{V}_1$ are the right singular vectors corresponding to the dominant singular values of $\widehat{R}$ and the columns of $\widehat{V}_3$ are the right singular vectors corresponding to the non-zero singular values of the symmetric matrix $\widehat{R}_i^T A \widehat{R}_i \in \mathbb{R}^{2(m+q),2(m+q)}$ or $\widehat{R}_i \in \mathbb{R}^{n_\mathbf{v},2(m+q)}$. Since the matrix $\widehat{R}_i$ has much more rows than $\widehat{R}_i^T A \widehat{R}_i$, we will compute $\widehat{V}_3$ from the singular value decomposition of the matrix

$$\widehat{R}_i^T A \widehat{R}_i = \begin{bmatrix} \widehat{B}_{11} & \widehat{B}_2^T (A_{12}^T A_{12})^{-1} \widehat{B}_2 \\ \widehat{B}_2^T (A_{12}^T A_{12})^{-1} \widehat{B}_2 & 0 \end{bmatrix},$$

where $\widehat{B}_{11} = [\, B_1,\ C_1^T \,]^T A_{12} (A_{12}^T A_{12})^{-1} \widehat{B}_2 + \widehat{B}_2^T (A_{12}^T A_{12})^{-1} A_{12}^T \widehat{B}_{12}$.

If we replace the full rank factor $\widehat{R}$ by the low rank Cholesky factor $\widehat{X}$ of the solution $\widehat{X}_{11} \approx \widehat{X}\widehat{X}^T$ of (3.29), then we obtain the following algorithm that is, in fact, a generalization of a *dominant subspace projection method* proposed in [19, 29] for standard state space systems.

Algorithm 3.2. *Generalized Dominant Subspace Projection (GDSP) method for the semidiscretized Stokes equation.*
Input: *Matrices M, L, D, B_0, B_2, C_0, C_2.*
Output: *A reduced order system $[\, \widetilde{E},\ \widetilde{A},\ \widetilde{B},\ \widetilde{C} \,]$.*

0. *If $M \neq I$, then compute the Cholesky factorization $M = U_M^T U_M$, where U_M is upper triangular. Otherwise, $U_M = I$.*
1. *Use the LRCF-ADI method* (3.15) *to compute the low rank Cholesky factor $\widehat{X} = \Pi\widehat{X}$ such that $\widehat{X}\widehat{X}^T$ is an approximate solution of the projected Lyapunov equation* (3.29)*, where $A_{11} = U_M^{-T} L U_M^{-1}$, $A_{12} = -U_M^{-T} D^T$ and $\widehat{B}_{12} = U_M^{-T}[\,B_0,\ C_0^T\,] - A_{11}A_{12}(A_{12}^T A_{12})^{-1}[\,B_2,\ C_2^T\,]$.*
2a. *Compute the 'thin' singular value decomposition*

$$\widehat{X} = \left[\,\widehat{U}_1,\ \widehat{U}_2\,\right] \begin{bmatrix} \widehat{\Sigma}_1 & 0 \\ 0 & \widehat{\Sigma}_2 \end{bmatrix} \left[\,\widehat{V}_1,\ \widehat{V}_2\,\right]^T, \tag{3.30}$$

where the matrices $[\,\widehat{U}_1,\ \widehat{U}_2\,]$ and $[\,\widehat{V}_1,\ \widehat{V}_2\,]$ have orthonormal columns, $\widehat{\Sigma}_1 = \mathrm{diag}(\widehat{\sigma}_1, \ldots, \widehat{\sigma}_{\hat{\ell}_f})$ and $\widehat{\Sigma}_2 = \mathrm{diag}(\widehat{\sigma}_{\hat{\ell}_f+1}, \ldots, \widehat{\sigma}_{\hat{r}})$ with $\widehat{\sigma}_1 \geq \ldots \geq \widehat{\sigma}_{\hat{\ell}_f} > \widehat{\sigma}_{\hat{\ell}_f+1} \geq \ldots \geq \widehat{\sigma}_{\hat{r}} > 0$, $\hat{r} = \mathrm{rank}(\widehat{X})$.
2b. *Compute the 'thin' singular value decomposition*

$$\begin{bmatrix} \widehat{B}_{11} & \widehat{B}_2^T (A_{12}^T A_{12})^{-1} \widehat{B}_2 \\ \widehat{B}_2^T (A_{12}^T A_{12})^{-1} \widehat{B}_2 & 0 \end{bmatrix} = \widehat{U}_3 \widehat{\Theta}_3 \widehat{V}_3^T,$$

where $\widehat{B}_2 = [\,B_2,\ C_2^T\,]$ and $\widehat{B}_{11} = [\,B_0,\ C_0^T\,]^T U_M^{-1} A_{12}(A_{12}^T A_{12})^{-1}\widehat{B}_2 + \widehat{B}_2^T (A_{12}^T A_{12})^{-1} A_{12}^T \widehat{B}_{12}$, $\widehat{U}_3^T$ and $\widehat{V}_3^T$ have orthonormal columns and $\widehat{\Theta}_3$ is nonsingular.
3. *Compute the reduced order system $[\,\widetilde{E},\ \widetilde{A},\ \widetilde{B},\ \widetilde{C}\,] = [\,\widehat{T}_{\hat{\ell}}^T E \widehat{T}_{\hat{\ell}},\ \widehat{T}_{\hat{\ell}}^T A \widehat{T}_{\hat{\ell}},\ \widehat{T}_{\hat{\ell}}^T B,\ C\widehat{T}_{\hat{\ell}}\,]$, where E, A, B, C are as in* (3.1) *and*

$$\widehat{T}_{\hat{\ell}} = \left[\, S\widehat{U}_1, \quad \widehat{R}_i \widehat{V}_3 \,\right] \tag{3.31}$$

with $S = \begin{bmatrix} I \\ -(A_{12}^T A_{12})^{-1} A_{12}^T A_{11} \end{bmatrix}$ and $\widehat{R}_i$ as in (3.28).

The following theorem gives a connection between the GLRSR and GDSP methods.

Theorem 3.1. *Suppose that the reduced order systems computed by the GLRSR and GDSP methods have orders $\ell = \ell_f + \ell_\infty$ and $\hat{\ell} = \hat{\ell}_f + \hat{\ell}_\infty$, respectively, where $\ell_f \leq \mathrm{rank}(Y^T X)$, $\ell_\infty = \mathrm{rank}(L_i A R_i)$, $\hat{\ell}_f = \mathrm{rank}(\widehat{X})$ and $\hat{\ell}_\infty = rank(\widehat{R}_i^T A \widehat{R}_i)$. Let W_ℓ, $T_\ell \in \mathbb{R}^{n,\ell}$ be the projection matrices as in* (3.25) *and let $\widehat{T}_{\hat{\ell}} \in \mathbb{R}^{n,\hat{\ell}}$ be the projection matrix as in* (3.31)*. Then*

$$\mathrm{Im}\, W_\ell \subseteq \mathrm{Im}\,[\,Y_p,\ L_i^T\,] \subseteq \mathrm{Im}\, \widehat{T}_{\hat{\ell}}, \qquad \mathrm{Im}\, T_\ell \subseteq \mathrm{Im}\,[\,X_p,\ R_i\,] \subseteq \mathrm{Im}\, \widehat{T}_{\hat{\ell}}, \tag{3.32}$$

where $Y_p = SY$ and $X_p = SX$ are the low rank Cholesky factors of the proper observability and controllability Gramians $\mathcal{G}_{po}$ and $\mathcal{G}_{pc}$ of (1.2)*, respectively.*

Proof. The projection matrices W_ℓ, T_ℓ and $\widehat{T}_{\hat{\ell}}$ can be rewritten as

$$\begin{aligned}
W_\ell &= [\,W_{\ell_f},\ W_{\ell_\infty}\,] \quad \text{with} \quad W_{\ell_f} = SYU_1\Sigma_1^{-1/2}, \quad W_{\ell_\infty} = L_i^T U_3 \Theta_3^{-1/2}, \\
T_\ell &= [\ T_{\ell_f},\ T_{\ell_\infty}\,] \quad \text{with} \quad T_{\ell_f} = SXV_1\Sigma_1^{-1/2}, \quad T_{\ell_\infty} = R_i V_3 \Theta_3^{-1/2}, \\
\widehat{T}_{\hat{\ell}} &= [\ \widehat{T}_{\hat{\ell}_f},\ \widehat{T}_{\hat{\ell}_\infty}\,] \quad \text{with} \quad \widehat{T}_{\hat{\ell}_f} = S[\,\widehat{U}_1,\ \widehat{U}_2\,], \quad \widehat{T}_{\hat{\ell}_\infty} = \widehat{R}_i \widehat{V}_3.
\end{aligned} \tag{3.33}$$

From (3.15), where B_{12} is replaced by $\widehat{B}_{12} = [\,B_{12},\ C_{12}^T\,]$, we obtain that the low rank Cholesky factor $\widehat{X}$ computed in Step 1 of Algorithm 3.2 has the form $\widehat{X} = [X,\ Y]P_{\widehat{X}}$, where

X and Y are the low rank Cholesky factors computed in Step 1 of Algorithm 3.1 and $P_{\widehat{X}}$ is a permutation matrix. In this case

$$\operatorname{Im}(YU_1\Sigma_1^{-1/2}) \subseteq \operatorname{Im} Y \subseteq \operatorname{Im}[X,\ Y] = \operatorname{Im}\widehat{X} = \operatorname{Im}[\widehat{U}_1,\ \widehat{U}_2]$$

and, hence, $\operatorname{Im} W_{\ell_f} \subseteq \operatorname{Im} Y_p \subseteq \operatorname{Im}\widehat{T}_{\hat{\ell}_f}$. Moreover, taking into account that $\widehat{R}_i = [R_i,\ L_i^T]$, we get

$$\operatorname{Im} W_{\ell_\infty} = \operatorname{Im}(L_i^T U_3 \Theta_3^{-1/2}) \subseteq \operatorname{Im} L_i^T \subseteq \operatorname{Im}\widehat{R}_i = \operatorname{Im}\widehat{T}_{\hat{\ell}_\infty}.$$

Analogously, one can show that $\operatorname{Im} T_{\ell_f} \subseteq \operatorname{Im} X_p \subseteq \operatorname{Im}\widehat{T}_{\hat{\ell}_f}$ and $\operatorname{Im} T_{\ell_\infty} \subseteq \operatorname{Im} R_i \subseteq \operatorname{Im}\widehat{T}_{\hat{\ell}_\infty}$. Thus, inclusions (3.32) hold. □

It follows from the proof of Theorem 3.1 that the range of $\widehat{T}_{\hat{\ell}_f}$ as in (3.33) can be considered as an approximation of the union of the dominant subspaces of the proper controllability and observability Gramians $\mathcal{G}_{pc}$ and $\mathcal{G}_{po}$ of system (1.2), whereas the column of $\widehat{T}_{\hat{\ell}_\infty}$ given in (3.33) span the union of the ranges of the improper controllability and observability Gramians $\mathcal{G}_{ic}$ and $\mathcal{G}_{io}$ of (1.2). This justifies why Algorithm 3.2 is called the dominant subspace projection method.

Note that for $\hat{\ell}_\infty > \ell_\infty$, the reduced order system computed by the GDSP method is not minimal and has still redundant state variables. To compute the minimal reduced order system for (1.3), (3.1) with symmetric A and $B_2 = C_1 = 0$, we can combine the GLRSR and GDSP methods in the following way.

Algorithm 3.3. *Generalized Symmetric Low Rank Square Root (GSLRSR) method for the semidiscretized Stokes equation.*
Input: *Matrices M, L, D, B_0, C_2.*
Output: *A reduced order system $[\widetilde{E},\ \widetilde{A},\ \widetilde{B},\ \widetilde{C}]$.*
0. *If $M \neq I$, then compute the Cholesky factorization $M = U_M^T U_M$, where U_M is upper triangular. Otherwise, $U_M = I$.*
1. *Use the LRCF-ADI method* (3.15) *to compute the low rank Cholesky factor $\widehat{X} = \Pi\widehat{X}$ such that $\widehat{X}\widehat{X}^T$ is an approximate solution of the projected Lyapunov equation* (3.29)*, where $A_{11} = U_M^{-T} L U_M^{-1}$, $A_{12} = -U_M^{-T} D^T$ and $\widehat{B}_{12} = [U_M^{-T} B_0,\ -A_{11}A_{12}(A_{12}^T A_{12})^{-1}C_2^T]$.*
2. *Compute the 'thin' singular value decompositions* (3.30) *and* (3.24)*.*
3. *Compute the reduced order system*

$$\widetilde{E} = \begin{bmatrix} I_{\hat{\ell}_f} & 0 \\ 0 & 0 \end{bmatrix}, \qquad \widetilde{A} = \begin{bmatrix} \widehat{U}_1^T A_{11} \widehat{U}_1 & 0 \\ 0 & I_{\ell_\infty} \end{bmatrix}, \qquad \widetilde{B} = \begin{bmatrix} \widehat{U}_1^T B_1 \\ \Theta_3^{1/2} V_3^T \end{bmatrix},$$
$$\widetilde{C} = \left[-C_2(A_{12}^T A_{12})^{-1} A_{12}^T A_{11} \widehat{U}_1,\ U_3 \Theta_3^{1/2} \right]. \tag{3.34}$$

We see that the matrices $\widetilde{E}$ and $\widetilde{A}$ in (3.34) are symmetric, the pencil $\lambda\widetilde{E} - \widetilde{A}$ is of index one and the reduced order system has order $\tilde{\ell} = \hat{\ell}_f + \ell_\infty$.

4 Numerical examples

In this section we demonstrate the reliability and performance of the proposed balanced truncation model reduction methods for the semidiscretized Stokes equation. All of the following results were obtained on a SunOS 5.8 workstation at the Department of Mathematics and

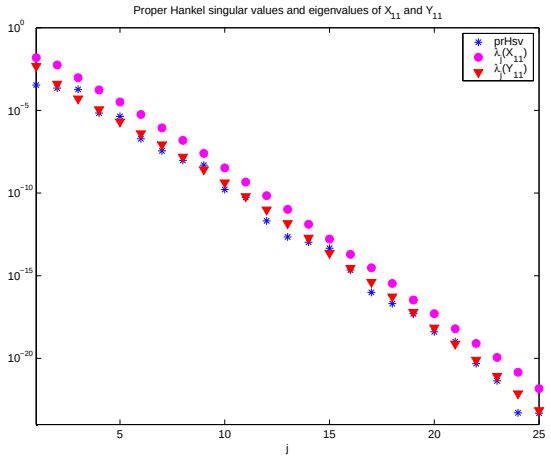

Figure 1: Proper Hankel singular values of (1.2) and eigenvalues of X_{11} and Y_{11}.

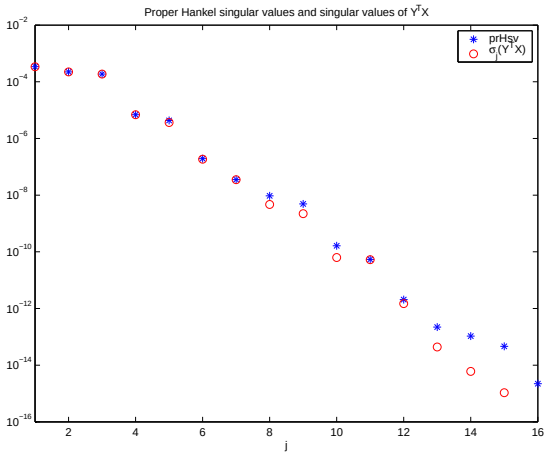

Figure 2: Proper Hankel singular values of (1.2) and singular values of Y^TX.

Statistics of the University of Calgary. The computations were performed with MATLAB 6.5 using IEEE double precision arithmetic with relative machine precision $\epsilon = 2.22 \times 10^{-16}$.

Consider the two dimensional instationary Stokes equation (1.1), where $\Omega = (0,1) \times (0,1)$, $\xi = (\xi_1, \xi_2)^T$ is the vector of space variables and the boundary conditions are non-slip, i.e., $v(\xi, t) = 0$ for $(\xi, t) \in \partial\Omega \times (0, t_f)$. Using a finite volume semidiscretization method on a uniform staggered grid [4, 41] with $n_1 + 1$ points in the ξ_1-direction and $n_2 + 1$ points in ξ_2-direction, we obtain system (1.2) with $M = I$. For simplicity, $B_0 \in \mathbb{R}^{n_\mathbf{v},1}$ is chosen at random, $B_2 = C_0 = 0$ and $C_2 = [1,\ 0] \in \mathbb{R}^{1,n_\mathbf{p}}$ with $n_\mathbf{v} = (n_1 + 1)n_2 + n_1(n_2 + 1)$, $n_\mathbf{p} = (n_1 + 1)(n_2 + 1) - 1$. For $n_1 = n_2 = 22$, we have $n_\mathbf{v} = 1012$, $n_\mathbf{p} = 528$ and the dimensions of the deflating subspaces of the pencil corresponding to the finite and infinite eigenvalues are $n_f = 484$ and $n_\infty = 1056$, respectively.

Note that in this example the proper and improper Hankel singular values of system (1.2) do not depend on the viscosity ϑ. Indeed, let $R_1(\vartheta)$ and $L_1(\vartheta)$ be the full rank Cholesky factors of the solutions of the projected Lyapunov equations (3.10) and (3.13) with $A_{11} = \vartheta L$, where L is the discrete Laplace operator. Then $R_1(\vartheta) = 1/\sqrt{\vartheta} R_1(1)$ and $L_1(\vartheta) = \sqrt{\vartheta} L_1(1)$. Hence, the proper Hankel singular values are given by $\varsigma_j = \sigma_j(L_1(\vartheta)R_1(\vartheta)) = \sigma_j(L_1(1)R_1(1))$. The non-zero improper Hankel singular values are computed as $\theta_j = \sigma_j(C_2(A_{12}^T A_{12})^{-1} A_{12}^T B_0)$, where C_2, A_{12} and B_0 are independent of ϑ. In this case the order of the semidiscretized Stokes equation (1.2) can be reduced equally for small and large values of the viscosity ϑ. We will assume that $\vartheta = 1$.

Figure 1 shows the 25 largest proper Hankel singular values of system (1.2) and eigenvalues of the solutions X_{11} and Y_{11} of the projected Lyapunov equations (3.10) and (3.13), respectively. One can see that the eigenvalues decay very fast and, hence, the matrices X_{11} and Y_{11} can be well approximated by matrices of low rank. Using the LRCF-ADI method we have computed the low rank Cholesky factors X, $Y \in \mathbb{R}^{1012,16}$ of the solutions of equations (3.10), (3.13) and the low rank Cholesky factor $\widehat{X} \in \mathbb{R}^{1012,32}$ of the solution of (3.29).

Figure 2 shows the proper Hankel singular values of (1.2) and the singular values of the matrix Y^TX. We see that the proper Hankel singular values are quite well approximated by the singular values of Y^TX.

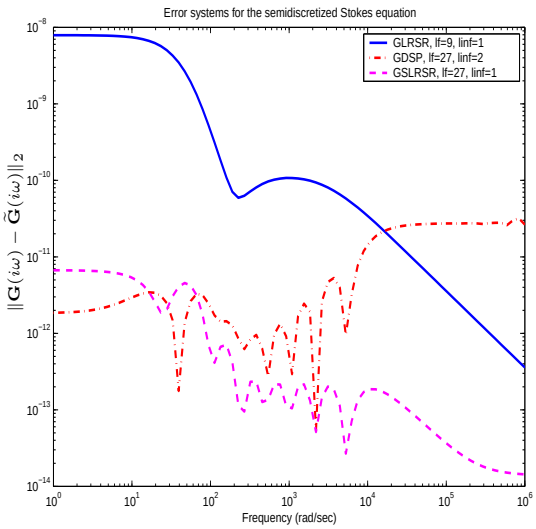

Figure 3: Error systems for the adaptive choice of ℓ_f and $\hat{\ell}_f$ with $\texttt{tol} = 10^{-6}$.

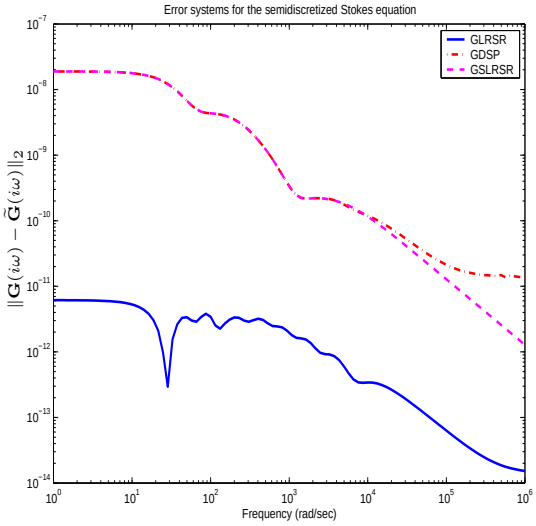

Figure 4: Error systems for the fixed choice $\ell_f = \hat{\ell}_f = 15$.

In Figures 3 and 4 we illustrate how accurate the semidiscretized Stokes equation is approximated by the reduced order models computed by three different model reduction methods described in Section 3.4. We display the spectral norm of the error system $\mathbf{G}(i\omega) - \widetilde{\mathbf{G}}(i\omega)$ for a frequency rang $\omega \in [1, 10^6]$. Two distinct choices of the reduced orders $\ell = \ell_f + \ell_\infty$, $\hat{\ell} = \hat{\ell}_f + \hat{\ell}_\infty$ and $\tilde{\ell} = \hat{\ell}_f + \hat{\ell}_\infty$ were made. In the first experiment, see Figure 3, the values ℓ_f and $\hat{\ell}_f$ are chosen as largest indexes such that $\sigma_{\ell_f}(Y^T X)/\sigma_1(Y^T X) \geq \texttt{tol}$ and $\sigma_{\hat{\ell}_f}(\widehat{X})/\sigma_1(\widehat{X}) \geq \texttt{tol}$ with prescribed tolerance $\texttt{tol}$. In the second experiment, see Figure 4, we choose $\ell_f = \hat{\ell}_f$. The values ℓ_∞ and $\hat{\ell}_\infty$ in both experiments are equal to the numerical rank of the matrices $L_i A R_i$ and $\widehat{R}_i^T A \widehat{R}_i$, respectively.

For the adaptive choice of ℓ_f and $\hat{\ell}_f$ the reduced orders ℓ of the systems computed by the GLRSR method are generally smaller than the orders $\hat{\ell}$ and $\tilde{\ell}$ of the systems delivered by the GDSP and GSLRSR methods, respectively. For $\texttt{tol} = 10^{-6}$, we have $\ell = 10$, $\hat{\ell} = 29$ and $\tilde{\ell} = 28$. We see that the approximation by the GDSP method is better for low frequencies whereas the GSLRSR method delivers the best approximation for the middle and high frequency ranges. However, for fixed $\ell_f = \hat{\ell}_f = 15$, the approximation error for the GLRSR method is considerably smaller than for the GDSP and GSLRSR methods.

5 Conclusion

In this paper we have discussed balanced truncation model reduction for descriptor systems. This approach is related to the proper and improper controllability and observability Gramians and Hankel singular values that can be computed by solving projected generalized Lyapunov equations. The balanced truncation method is based on transforming the descriptor system to a balanced form and reducing the order by truncation of the states that correspond to the small proper and zero improper Hankel singular values. Important properties of this method are that the regularity and stability is preserved in the reduced order system and there is an a priori bound on the approximation error.

We have also discussed the application of the balanced truncation model reduction to the semidiscretized Stokes equation. This equation has a special block structure that can be used to reduce the computational effort. The proper controllability and observability Gramians for the semidiscretized Stokes equation with a small number of inputs and outputs are well approximated by low rank matrices and their low rank Cholesky factors can efficiently be computed by the low rank Cholesky factor alternating direction implicit method. The Cholesky factors of the improper controllability and observability Gramians have been found in explicit form.

Three model reduction methods for the semidiscretized Stokes equation have been presented. The first two methods are generalizations of the low rank square root method and the dominant subspace projection method known for standard state space systems. The third method is a combination of the others. The effectiveness of the proposed model reduction algorithms has been demonstrated by numerical experiments.

Acknowledgement. This work was supported in part by Deutsche Forschungsgemeinschaft, Research Grant ME790/12-1. It was completed at the University of Calgary while the author was a PIMS Postdoctoral Fellow. The author would like to thank P. Lancaster for his hospitality and V. Mehrmann for helpful discussions.

References

[1] A.C. Antoulas, D.C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. In V. Olshevsky, editor, *Structured Matrices in Mathematics, Computer Science and Engineering, Vol. I*, Contemporary Mathematics Series, 280. American Mathematical Society, 2001.

[2] A.C. Antoulas, D.C. Sorensen, and Y. Zhou. On the decay rate of the Hankel singular values and related issues. *Systems Control Lett.*, 46(5):323–342, 2002.

[3] D.J. Bender. Lyapunov-like equations and reachability/observability Gramians for descriptor systems. *IEEE Trans. Automat. Control*, 32(4):343–348, 1987.

[4] T.R. Bewley. Flow control: new challenges for a new Renaissance. *Progress in Aerospace Sciences*, 37:21–58, 2001.

[5] L. Dai. *Singular Control Systems*. Lecture Notes in Control and Information Sciences, 118. Springer-Verlag, Berlin, Heidelberg, 1989.

[6] J.W. Demmel and B. Kågström. The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: Robust software with error bounds and applications. Part I: Theory and algorithms. *ACM Trans. Math. Software*, 19(2):160–174, 1993.

[7] G. Doetsch. *Guide to the Applications of the Laplace and Z-Transforms*. Van Nostrand Reinhold Company, London, 1971.

[8] D. Enns. Model reduction with balanced realization: an error bound and a frequency weighted generalization. In *Proceedings of the 23rd IEEE Conference on Decision and Control (Las Vegas, 1984)*, pages 127–132. IEEE, New York, 1984.

[9] P. Feldmann and R.W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanzcos process. *IEEE Trans. Computer-Aided Design*, 14:639–649, 1995.

[10] L. Fortuna, G. Nunnari, and A. Gallo. *Model Order Reduction Techniques with Applications in Electrical Engineering*. Springer-Verlag, London, 1992.

[11] R.W. Freund. Krylov-subspace methods for reduced-order modeling in circuit simulation. *J. Comput. Appl. Math.*, 123(1-2):395–421, 2000.

[12] K. Gallivan, E. Grimme, and P. Van Dooren. Asymptotic waveform evaluation via a Lanczos method. *Appl. Math. Lett.*, 7:75–80, 1994.

[13] K. Gallivan, E. Grimme, and P. Van Dooren. A rational Lanczos algorithm for model reduction. *Numerical Algorithms*, 12(1-2):33–63, 1996.

[14] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-errors bounds. *Internat. J. Control*, 39(6):1115–1193, 1984.

[15] G.H. Golub and C.F. Van Loan. *Matrix Computations. 3rd ed.* The Johns Hopkins University Press, Baltimore, London, 1996.

[16] E.J. Grimme, D.C. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numerical Algorithms*, 12(1-2):1–31, 1996.

[17] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, Orlando, FL, 2nd edition, 1985.

[18] A.J. Laub, M.T. Heath, C.C. Paige, and R.C. Ward. Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms. *IEEE Trans. Automat. Control*, AC-32(2):115–122, 1987.

[19] J.-R. Li. *Model Reduction of Large Linear Systems via Low Rank System Gramians*. Ph.D. thesis, Department of Mathematics, Massachusetts Institute of Technology, 2000.

[20] J.-R. Li, F. Wang, and J. White. An efficient Lyapunov equation-based approach for generating reduced-order models of interconnect. In *Proceedings of the 36th Design Automation Conference (New Orleans, USA, 1999)*, pages 1–6. IEEE, 1999.

[21] W.Q. Liu and V. Sreeram. Model reduction of singular systems. In *Proceedings of the 39th IEEE Conference on Decision and Control (Sydney, Australia, 2000)*, pages 2373–2378. IEEE, 2000.

[22] Y. Liu and B.D.O. Anderson. Singular perturbation approximation of balanced systems. *Internat. J. Control*, 50:1379–1405, 1989.

[23] R. März. Canonical projectors for linear differential algebraic equations. *Comput. Math. Appl.*, 31(4-5):121–135, 1996.

[24] B.C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Automat. Control*, AC-26(1):17–32, 1981.

[25] T. Penzl. *Numerische Lösung großer Lyapunov-Gleichungen*. Logos Verlag, Berlin, 1998. [German].

[26] T. Penzl. A cyclic low-rank Smith method for large sparse Lyapunov equations. *SIAM J. Sci. Comput.*, 21(4):1401–1418, 1999/00.

[27] T. Penzl. Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case. *Systems Control Lett.*, 40(2):139–144, 2000.

[28] T. Penzl. LYAPACK Users Guide. Preprint SFB393/00-33, Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, August 2000. Available from `http://www.tu-chemnitz.de/sfb393/sfb00pr.html`.

[29] T. Penzl. Algorithms for model reduction of large dynamical systems. Preprint SFB393/99-40, Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, December 1999. Available from `http://www.tu-chemnitz.de/sfb393/sfb99pr.html`.

[30] S.S. Ravindran. A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. *Internat. J. Numer. Methods Fluids*, 34(5):425–448, 2000.

[31] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Company, Boston, MA, 1996.

[32] M.G. Safonov and R.Y. Chiang. A Schur method for balanced-truncation model reduction. *IEEE Trans. Automat. Control*, AC-34(7):729–733, 1989.

[33] D.C. Sorensen and Y. Zhou. Bounds on eigenvalue decay rates and sensitivity of solutions of Lyapunov equations. Technical Report TR02-07, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 2002.

[34] G.W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, New York, 1990.

[35] T. Stykel. *Analysis and Numerical Solution of Generalized Lyapunov Equations*. Ph.D. thesis, Institut für Mathematik, Technische Universität Berlin, Berlin, 2002.

[36] T. Stykel. Model reduction of descriptor systems. Technical Report 720-2001, Institut für Mathematik, Technische Universität Berlin, D-10263 Berlin, Germany, December 2001. Available from `http://www.math.tu-berlin.de/~stykel/Publications/pr_720_01.ps`.

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

[38] A. Varga. Efficient minimal realization procedure based on balancing. In A. EL Moudni, P. Borne, and S.G. Tzafestas, editors, *Proc. of IMACS/IFAC Symposium on Modelling and Control of Technological Systems (Lille, France, May 7-10, 1991)*, volume 2, pages 42–47, 1991.

[39] S. Volkwein. *Optimal and Suboptimal Control of Partial Differential Equations: Augmented Lagrange-SQP Methods and Reduced-Order Modeling with Proper Orthogonal Decomposition*. Grazer Mathematische Berichte, 343. Karl-Franzens-Universität Graz, Graz, 2001.

[40] E.L. Wachspress. *Iterative Solution of Elliptic Systems, and Applications to the Neutron Diffusion Equations of Reactor Physics*. Prentice Hall, Englewood Cliffs, N.J., 1966.

[41] J. Weickert. *Applications of the Theory of Differential-Algebraic Equations to Partial Differential Equations of Fluid Dynamics*. Ph.D. thesis, Technische Universität Chemnitz, Chemnitz, 1997.