Analysis of positive descriptor systems

vorgelegt von

Dipl.-Math. oec. Elena Virnik

von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Stefan Felsner
Berichter:	Prof. Dr. Volker Mehrmann
Berichter:	Prof. Dr. Reinhard Nabben
Gutachter:	Prof. Dr. Daniel B. Szyld

Tag der wissenschaftlichen Aussprache: 30. Mai 2008

Berlin 2008
D83

Acknowledgements

There are a few people whose help was invaluable in completing this thesis. Lacking a language understood by everyone, the author chooses to differentiate between the cases.

Als erstes möchte ich mich herzlich bei Volker Mehrmann bedanken. Das Thema der Dissertation, für das er einen Rahmen vorgegeben hat, der mir gleichzeitig Richtungsfreiheit ließ, hat mir viel Freude bereitet. Ich bin dankbar für seine stetige Unterstützung, die zahlreichen Diskussionen, Hinweise und konstruktive Kritik. Darüber hinaus habe ich die wissenschaftlich inspirierende Atmospäre, in der ich arbeiten durfte, zu schätzen gelernt und auch die vielen Möglichkeiten, die mir geboten wurden, meine Arbeit international zu präsentieren und mich mit anderen Forschern auszutauschen.

Estoy muy agradecida a la persona que me alentó a emprender mi doctorado, Daniel B. Szyld. Valoro mucho que siempre haya tenido ideas para mis avances de estudio, sea la participación en un concurso de trabajos de investigación estudiantil o una estadía de estudios en Philadelphia, EEUU. Me significa mucho su asistencia constante como mentor y como amigo.

Ich bedanke mich recht herzlich bei Reinhard Nabben für viele wertvolle Diskussionen und für die Begutachtung dieser Arbeit.

I am much obliged to Shmuel Friedland for very fruitful discussions on nonnegative projectors and Schur complements that were invaluable for completing some parts of this thesis.

I would like to express my gratitude to all colleagues in the TU-Berlin Department of Mathematics, where I have very much enjoyed the friendly atmosphere and the productive work environment.

Mein spezieller Dank gilt meinen Büronachbarn, Christian Schröder und Lena Wunderlich für ein harmonisches Zusammenleben und ihre Hilfsbereitschaft.

Weiterhin bedanke ich mich insbesondere bei Timo Reis für viele produktive Diskussionen nicht nur zum Thema Modellreduktion.

Я благодарна моим родителям за безусловную и непрерывную поддержку во всём.

Nicht zuletzt bedanke ich mich bei Wido Weise, der mir in den letzten Jahren viel Liebe, moralische Unterstützung und Verständnis entgegenbrachte.

Contents

Introduction

One is tempted to assert that positive systems are the most often encountered systems in almost all areas of science and technology.
- Lorenzo Farina / Sergio Rinaldi

We consider linear time-invariant positive descriptor systems in continuous-time

$$E\dot{x}(t) = Ax(t) + Bu(t),\ x(0) = x_0 \tag{1a}$$
$$y(t) = Cx(t) + Du(t), \tag{1b}$$

and in discrete-time

$$Ex(t+1) = Ax(t) + Bu(t),\ x(0) = x_0 \tag{2a}$$
$$y(t) = Cx(t) + Du(t), \tag{2b}$$

where $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}, D \in \mathbb{R}^{p\times m}$, are real constant coefficient matrices. In the continuous-time case, the state x, input u and output y are real-valued vector functions. In the discrete-time case x, u and y are real-valued vector sequences. Positive systems, and here we mean internally positive systems, are systems whose state and output variables take only nonnegative values at all times t for any nonnegative initial state and any nonnegative input [2], [38], [64], [83].

Positive systems arise naturally in many applications such as pollutant transport, chemotaxis, pharmacokinetics, Leontief input-output models, population models and compartmental systems [2], [3], [14], [15], [19], [31], [38], [64]. In these models, the variables represent concentrations, population numbers of bacteria or cells or, in general, measures that are per se nonnegative. Positive standard systems, i.e., where E is the identity matrix, are subject to ongoing research by many authors [1], [32], [33], [38], [64], [100], [102], [110], [118], [119]. Recent advances on control theoretical issues have been made especially in the positive discrete-time case. Yet, there are still many open problems, especially for standard positive systems in continuous-time.

Control theory of linear time-invariant descriptor systems without the nonnegativity restriction was studied in [34], [95]. Very little is known about positive descriptor systems up to now, however, some properties mainly in the discrete-time case were studied in [23], [24], [25], [31], [64].

With this work we aim to lay the foundation for understanding positive descriptor systems in the continuous-time as well as in the discrete-time case. We present a cohesive framework that allows to generalise many results from standard positive systems to the descriptor case. In the following paragraphs we briefly explain the main constituent parts of our framework.

It is well known that many properties of standard systems, where $E = I$, are closely related to the spectral properties of the system matrix A. For instance, asymptotic stability of the system is equivalent to the eigenvalues of A being located in the open left complex half-plane in the continuous-time case, or to the eigenvalues of A being located in the open unit ball around the origin in the discrete-time case. If the dynamics of the system, however, is described by an implicit differential or difference equation, then such properties are determined by the eigenvalues and eigenvectors associated with the matrix pencil $\lambda E - A$, or just the matrix pair (E, A).

Most pertinent to the spectral analysis of standard positive systems is the well-known Perron-Frobenius theory. The classical Perron-Frobenius Theorem states that for an elementwise nonnegative matrix the spectral radius, i.e., the largest modulus of an eigenvalue is itself an eigenvalue and has a nonnegative eigenvector, see Chapter 2, Section 2.1. For the analysis of positive descriptor systems, it is therefore essential to have available a meaningful counterpart of the Perron-Frobenius theory for matrix pairs (E, A).

Due to the many applications, several approaches have been presented in the literature to generalise the classical Perron-Frobenius theory to matrix pencils or further to matrix polynomials [9], [41], [87], [107]. However, none of these generalisations is suitable in the case of positive descriptor systems.

In Chapter 2, Section 2.2, we propose a new approach to extend the classical Perron-Frobenius theory to matrix pairs (E, A), where a sufficient condition guarantees that the finite spectral radius of (E, A) is an eigenvalue with a corresponding nonnegative eigenvector. For the special case $E = I$ our new approach reduces to the classical Perron-Frobenius theorem for matrices. We present several examples where the new condition holds, whereas previous conditions in the literature are not satisfied.

Another notion from the theory of nonnegative matrices that we focus on in Chapter 2 are nonnegative projectors, i.e. nonnegative idempotent matrices. In the descriptor case, the choice of the right projector onto the deflating subspace that corresponds to the finite eigenvalues of the matrix pair (E, A) is crucial for the analysis [89]. As it turns out, nonnegative projectors play an important role in the analysis of positive descriptor systems, see Chapter 3 and Chapter 4. Furthermore, Schur complements constitute a fundamental tool in applications [129], in particular such as algebraic multi-grid methods [123] or model reduction [82]. However, it is important to ensure that the main properties of the matrix, the Schur complement is applied to, are preserved. In our case, in a positivity preserving model reduction method proposed in Chapter 6, two variations of the Schur complement will be applied to nonnegative projectors. Therefore, in Section 2.4 we show that for a nonnegative projector, if the corresponding part of the matrix is invertible, the Schur complement is again a nonnegative projector. Otherwise, if the corresponding part has a positive diagonal, the Moore-Penrose inverse Schur complement is again a nonnegative projector. Also the nonnegativity of a shifted Schur complement needed for discrete-time systems is shown.

The positivity condition that the state and output variables take only nonnegative values at all times $t \geq 0$, per se, is not easy to check. In the standard case, however, an equivalent characterisation is available in the continuous-time as well as in the discrete-time case that allows to determine positivity by just considering the system matrices, see, e.g., [38], [64].

In the descriptor case discussed in Chapter 3, the situation is more complicated. Not every initial value is consistent and consistency depends on the choice of the input [26], [34], [74]. Furthermore, the properties of the matrices in the standard case have to hold for special matrix products and on the deflating subspace that corresponds to the finite eigenvalues of the matrix pair (E, A). Assuming that the spectral projector onto this subspace is a nonnegative matrix allows to establish equivalent characterisations of positivity that directly correspond to the characterisations in the standard case. We reason why the nonnegativity assumption on the spectral projector is meaningful from the point of view of applications.

The characterisation of positivity for descriptor systems established in Section 3.2 imposes a very special structure on the system matrices. In Section 3.2.2 we analyse and specify this structure. For instance, this becomes important in Section 6.3, where we prove positivity of the reduced order descriptor system for the proposed model reduction technique.

In the case of standard positive systems, classical stability criteria take a simple form [38], [64]. In Chapter 4 we present generalisations of these stability criteria for the case of positive descriptor systems. It turns out, that if the spectral projector onto the finite deflating subspace of the matrix pair (E, A) is nonnegative, then all stability criteria for standard positive systems take a comparably simple form in the positive descriptor case.

Stability properties and also many other control theoretical issues such as model reduction methods or the quadratic optimal control problem are, furthermore, closely related to the solution of Lyapunov equations, see. e.g., [5], [44], [45], [79], [95]. For descriptor systems, projected generalised Lyapunov equations were presented in [115]. In the context of positive systems one is interested not only in positive (semi)definite solutions of such Lyapunov equations but rather in doubly nonnegative solutions, i.e., solutions that are both positive semidefinite and entry-wise nonnegative. Such results for standard Lyapunov equations, e.g., can easily be deduced from a more general discussion in [35]. In Chapter 5, we provide sufficient conditions that guarantee the existence of doubly nonnegative solutions of projected generalised Lyapunov equations.

A very important issue in systems and control theory is the development of model reduction techniques [5], [101], [116]. The need for highly detailed and accurate models leads to very large and complex systems. On the other hand, when dealing with simulation and especially control of such systems, limitations in computational and storage capabilities pose the problem of finding a simpler model that approximates the complex one in some sense. A crucial issue in model reduction is the preservation of special system properties such as stability, passivity or in our case the positivity of the system.

In Chapter 6 we propose a model reduction method for positive systems. In Section 6.2 we generalise the model reduction methods of standard balanced truncation [46], and singular perturbation balanced truncation [82], such that positivity is preserved. The proposed technique uses a linear matrix inequality (LMI) approach [37], [81], and we show that stability is preserved and an error bound in the H_∞-norm is provided.

Furthermore, in Section 6.3 we generalise this technique to the case of positive descriptor systems. Here, the procedure involves the additive decomposition of the transfer function into a strictly proper and a polynomial part as in [116]. The reduced order model is obtained via positivity preserving reduction of the strictly proper part of the transfer function, whereas the polynomial part remains unchanged.

In Chapter 7 we reflect the studied topics and give some concluding remarks. This is followed by Chapter 8 where we discuss open problems and point out possible directions of future research.

Chapter 1

Preliminaries and notation

Consistency is the last refuge of the unimaginative.
- Oscar Wilde

In the present chapter, we introduce concepts from systems theory of descriptor systems as well as from nonnegative matrix theory, that are essential for the analysis of positive descriptor systems. Throughout this thesis, we adapt the following standard notation. $\mathbb{R}$ and $\mathbb{C}$ are the fields of real and complex numbers, respectively. $\mathbb{N}$ is the set of nonnegative integers and $\mathbb{R}_+$ denotes the nonnegative real numbers. $\mathbb{C}_-$ and $\mathbb{C}_+$ are the open left and right complex half-planes. For $\lambda \in \mathbb{C}$ we denote by $\Re(\lambda)$ its real part and by $|\lambda|$ its absolute value. For the image space of a matrix $A \in \mathbb{R}^{n\times n}$ we write $\mathrm{im}(A)$ and for the nullspace of A we write $\ker(A)$. The rank of A is denoted by $\mathrm{rank}(A)$. For a matrix $A \in \mathbb{C}^{n\times m}$ we denote by A^T the transpose of A and by A^* the conjugate transpose of A. Let $\mathbb{I}$ denote a real interval. The space of k-times continuously differentiable functions from the real interval $\mathbb{I}$ into the real vector space $\mathbb{R}^n$ is denoted by $C^k(\mathbb{I}, \mathbb{R}^n)$, or shortly by C^k.

We define submatrices of a matrix as follows. Let $\langle n\rangle := \{1, \ldots, n\}$ and let $\alpha = \{\alpha_1, \ldots, \alpha_l\}$, $\beta = \{\beta_1, \ldots, \beta_m\} \subseteq \langle n\rangle$ be two nonempty sets. For $A \in \mathbb{R}^{n\times n}$ we denote by $A[\alpha, \beta]$ the submatrix of A composed of the rows and columns indexed by the sets α and β respectively, i.e.,

$$A[\alpha, \beta] := [a_{\alpha_i\beta_j}]_{i,j=1}^{l,m} \in \mathbb{R}^{l\times m}.$$

For square matrices $A_1, \ldots, A_n$ we define the block diagonal matrix

$$\mathrm{diag}(A_1, \ldots, A_n) := \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{bmatrix},$$

and for matrix products AB, we denote the (i,j)-th (block) entry by $[AB]_{ij}$.

1.1 Matrix pencils and the generalised eigenvalue problem

Let $E, A \in \mathbb{R}^{n\times m}$. A matrix pair (E,A), or a matrix pencil $\lambda E - A$, is called *regular* if E and A are square $(n=m)$ and $\det(\lambda E - A) \neq 0$ for some $\lambda \in \mathbb{C}$. It is called *singular* otherwise. In the present work we only consider square and regular pencils. The terms matrix pair and matrix pencil will be used interchangeably throughout this work.

Note that all notions defined in this section have the usual definitions for a single matrix A as a special case when setting $E = I$.

A scalar $\lambda \in \mathbb{C}$ is said to be a *(finite) eigenvalue* of the matrix pair (E,A) if $\det(\lambda E - A) = 0$. A vector $x \in \mathbb{C}^n \setminus \{0\}$ such that $(\lambda E - A)x = 0$ is called *(right) eigenvector* of (E,A) corresponding to λ. If E is singular and $v \in \mathbb{C}^n \setminus \{0\}$, such that $Ev = 0$ holds, then v is called *(right) eigenvector of* (E,A) *corresponding to the eigenvalue* ∞.

The equation

$$\lambda E v = A v, \tag{1.1}$$

is called *generalised eigenvalue problem*.

The set of all eigenvalues is called *spectrum* of (E,A) and is defined by

$$\sigma(E,A) := \begin{cases} \sigma_f(E,A), & \text{if } E \text{ is invertible,} \\ \sigma_f(E,A) \cup \{\infty\}, & \text{if } E \text{ is singular,} \end{cases}$$

where $\sigma_f(E,A)$ is the set of all finite eigenvalues. We denote by

$$\rho_f(E,A) = \max_{\lambda \in \sigma_f(E,A)} |\lambda|,$$

the *finite spectral radius* of (E,A).

A k-dimensional subspace $\mathcal{V} \subset \mathbb{C}^n$ is called *right deflating subspace of* (E,A), if there exists a k-dimensional subspace $\mathcal{W} \subset \mathbb{C}^n$ such that $E\mathcal{V} \subset \mathcal{W}$ and $A\mathcal{V} \subset \mathcal{W}$. A k-dimensional subspace $\mathcal{V} \subset \mathbb{C}^n$ is called *left deflating subspace of* (E,A), if it is right deflating subspace of (E^T, A^T) [65], [125]. Note that deflating subspaces are sometimes also termed eigenspaces [113].

Vectors $v_1, \ldots, v_k$ form a *right Jordan chain* of the matrix pair (E,A) corresponding to the finite eigenvalue λ if

$$(\lambda E - A)v_i = -Ev_{i-1},$$

for all $1 \leq i \leq k$ and $v_0 = 0$. Note that vectors $w_1, \ldots, w_k$ form a *right Jordan chain* of the matrix pair (E, A) corresponding to the eigenvalue ∞ if they form a right Jordan chain of the matrix pair (A, E) corresponding to the eigenvalue 0. A subspace $S_\lambda^{def} \subset \mathbb{C}^n$ is called *right deflating subspace of* (E, A) *corresponding to the eigenvalue* λ, if it is spanned by all right Jordan chains corresponding to λ.

Let $\lambda_1, \ldots, \lambda_p$, be the pairwise distinct finite eigenvalues of (E, A) and let $S_{\lambda_i}^{def}$, $i = 1, \ldots, p$, be the right deflating subspaces corresponding to these eigenvalues. We call the subspace defined by

$$S_f^{def} := S_{\lambda_1}^{def} \oplus \ldots \oplus S_{\lambda_p}^{def} \tag{1.2}$$

the *right finite deflating subspace* of (E, A). We call the subspace S_∞^{def} *right infinite deflating subspace*.

Vectors $z_1, \ldots, z_k$ form a *left Jordan chain* of the matrix pair (E, A) corresponding to the finite eigenvalue λ if

$$w_i^*(\lambda E - A) = -w_{i-1}^* E,$$

for all $1 \leq i \leq k$ and $w_0 = 0$. A subspace $\mathcal{V}_\lambda^{def} \subset \mathbb{C}^n$ is called *left deflating subspace of* (E, A) *corresponding to the eigenvalue* λ if it is spanned by all left Jordan chains corresponding to λ.

Let $\lambda_1, \ldots, \lambda_p$, be the pairwise distinct finite eigenvalues of (E, A) and let $\mathcal{V}_{\lambda_i}^{def}$, $i = 1, \ldots, p$, be the left deflating subspaces corresponding to these eigenvalues. We call the subspace defined by

$$\mathcal{V}_f^{def} := \mathcal{V}_{\lambda_1}^{def} \oplus \ldots \oplus \mathcal{V}_{\lambda_p}^{def} \tag{1.3}$$

the *left finite deflating subspace* of (E, A). We call the subspace $\mathcal{V}_\infty^{def}$ *left infinite deflating subspace*.

Often it is useful to consider the regular matrix pair (E, A) in the Weierstraß canonical form [26], [34] that is a special case of the Kronecker canonical form [45].

Theorem 1.1 (Weierstraß canonical form (WCF)) *Let* (E, A) *with* $E, A \in \mathbb{R}^{n \times n}$ *be a regular matrix pair. Then, there exist regular matrices* $W, T \in \mathbb{R}^{n \times n}$ *such that*

$$(E, A) = \left(W \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T, W \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} T \right), \tag{1.4}$$

where J *is a matrix in Jordan canonical form and* N *is a nilpotent matrix also in Jordan canonical form.*

The smallest nonnegative integer ν such that

$$\mathrm{rank}(E^\nu) = \mathrm{rank}(E^{\nu+1})$$

is called the *index* of the matrix E and is denoted by $\operatorname{ind}(E)$. For a nilpotent matrix N, i.e., there exists $k \in \mathbb{N}$ with $N^k = 0$, the smallest k with this property is called the *index of nilpotency*. Note that for a nilpotent matrix the index and the index of nilpotency coincide.

For a matrix pair (E, A) the *index* is defined by the index of nilpotency of the nilpotent block N in the Weierstraß canonical form and is denoted by $\operatorname{ind}(E, A)$. For a descriptor system with constant coefficients as in (1a) or (2a), we define the *index of the descriptor system* by $\operatorname{ind}(E, A)$. Note that for a regular matrix pair (E, A) with $EA = AE$, we have $\operatorname{ind}(E, A) = \operatorname{ind}(E)$.

1.2 Projectors and index of (E, A)

In this subsection, we present an alternative derivation of the index of (E, A) using projectors, that is due to [52].

A matrix $Q \in \mathbb{R}^{n\times n}$ is called *projector* if $Q^2 = Q$. A projector Q is called projector *onto* a subspace $S \subseteq \mathbb{R}^n$ if $\operatorname{im} Q = S$. It is called projector *along* a subspace $S \subseteq \mathbb{R}^n$ if $\ker Q = S$. The following Lemma 1.2, in particular, states that every projector is diagonalisable, see, e.g., [58].

Lemma 1.2 *Let P be a projector. Then, there exists a regular matrix T such that*

$$P = T^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T.$$

Furthermore, we will make use of the following well-known property of projectors, see, e.g., [58].

Lemma 1.3 *Let P_1, P_2 be projectors. Then,*

1. *P_1, P_2 project onto the same subspace S if and only if $P_1 = P_2P_1$ and $P_2 = P_1P_2$.*
2. *P_1, P_2 project along the same subspace S if and only if $P_1 = P_1P_2$ and $P_2 = P_2P_1$.*

A simple consequence of Lemma 1.3 is that a projector is uniquely defined by two subspaces, the one it projects onto and the one along which it projects. Consider the Weierstraß canonical form of the pair (E, A) given in (1.4). The matrices

$$P_r := T^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T \quad \text{and} \quad P_l := W \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} W^{-1} \tag{1.5}$$

are the unique spectral projectors onto the right and left finite deflating subspaces along the right and left infinite deflating subspaces, respectively.

Now, we define the matrix chain as introduced in [52]. Let (E, A) be a regular matrix pair and set

$$\begin{aligned} &E_0 := E, \quad A_0 := A \quad \text{and} \\ &E_{i+1} := E_i - A_i\tilde{Q}_i, \quad A_{i+1} := A_i\tilde{P}_i, \quad \text{for} \quad i \geq 0, \end{aligned} \tag{1.6}$$

where $\tilde{Q}_i$ are projectors onto $\ker E_i$ and $\tilde{P}_i = I - \tilde{Q}_i$. Since we have assumed (E, A) to be regular, there exists an index ν such that E_ν is nonsingular and all E_i are singular for $i < \nu$ [52], [88]. Note, that ν is independent of a specific choice of the projectors Q_i. We say that the matrix pair (E, A) has *tractability index* ν. It is well known that for regular pairs (E, A) with constant coefficients the tractability index is equal to the (differentiation) index as defined in Section 1.1, see, e.g., [27], and it can be determined as the size of the largest Jordan block associated with the eigenvalue infinity in the Weierstraß canonical form of the pair (E, A), see [74] , [88].

For the analysis of descriptor system via projector chains as defined in (1.6), it is essential to use specific, so called *canonical* projectors Q_i, for $i = 0, \ldots, \nu - 1$, that have additional properties [89]. Note that for $\operatorname{ind}(E, A) > 1$, even these specific projectors are not unique. However, the canonical projector $Q_{\nu-1}$ is unique and can be calculated as in the following Lemma 1.4 [88].

Lemma 1.4 *Let (E, A) be a matrix pair and define a matrix chain as in (1.6). Furthermore, define sets S_i by*

$$S_i := \{y \in \mathbb{R}^n : A_i y \in \operatorname{im} E_i\}. \tag{1.7}$$

Then, if E_{i+1} is nonsingular, we have that

$$Q_i = -\tilde{Q}_i E_{i+1}^{-1} A_i$$

is a projector onto $\ker E$ along S_i.

For the construction of canonical projectors in the higher index cases in Section 2.2.4, we need the following properties.

Lemma 1.5 *Let (E, A) be a regular matrix pair of $\operatorname{ind}(E, A) = \nu$ and define a matrix chain as in (1.6), where the projectors $\tilde{Q}_i$ are chosen such that $\tilde{Q}_j\tilde{Q}_i = 0$ holds for all $0 \leq i < j \leq \nu - 1$. For $0 \leq i \leq \nu - 1$ define new projectors Q_i onto $\ker E_i$ by setting $Q_i = -\tilde{Q}_i E_\nu^{-1} A_i$ and $P_i = I - Q_i$. Then, $Q_j Q_i = 0$ holds for all $0 \leq i < j \leq \nu - 1$.*

Proof. The matrix $-\tilde{Q}_i E_\nu^{-1} A_i$ is a projector for all $0 \le i \le \nu - 1$, since

$$(-\tilde{Q}_i E_\nu^{-1} A_i)^2 = \tilde{Q}_i E_\nu^{-1}(E_i - E_{i+1})\tilde{Q}_i E_\nu^{-1} A_i = -\tilde{Q}_i E_\nu^{-1} E_{i+1}\tilde{Q}_i E_\nu^{-1} A_i = -Q_i E_\nu^{-1} A_i,$$

where we have used that $E_\nu \tilde{Q}_i = (E_{i+1} - A_{i+1}\tilde{Q}_{i+1} - \ldots - A_{\nu-1}Q_{\nu-1})\tilde{Q}_i = E_{i+1}\tilde{Q}_i$.

To show that $Q_j Q_i = 0$ holds for all $0 \le i < j \le \nu - 1$, let i, j be arbitrarily chosen fixed indeces $0 \le i < j \le \nu - 1$. Then, we have that

$$\begin{aligned} Q_j Q_i &= \tilde{Q}_j E_\nu^{-1} A_j \tilde{Q}_i E_\nu^{-1} A_i = \tilde{Q}_j E_\nu^{-1} A_i \tilde{P}_i \cdots \tilde{P}_{j-1} \tilde{Q}_i E_\nu^{-1} A_i \\ &= \tilde{Q}_j E_\nu^{-1} A_i (I - \tilde{Q}_i) \cdots (I - \tilde{Q}_{j-1}) \tilde{Q}_i E_\nu^{-1} A_i = \tilde{Q}_j E_\nu^{-1} A_i (\tilde{Q}_i - \tilde{Q}_i) \\ &= 0. \end{aligned}$$

□

1.3 Nonnegative matrices

A vector $x \in \mathbb{R}^n, x = [x_i]_{i=1}^n$ is called *nonnegative (positive)* and we write $x \ge 0$ $(x > 0)$ if all entries x_i are nonnegative (positive). A matrix $A \in \mathbb{R}^{n \times n}$, $A = [a_{ij}]_{i,j=1}^n$ is called *nonnegative (positive)* and we write $A \ge 0$ $(A > 0)$ if all entries a_{ij} are nonnegative (positive). A matrix A is called *nonnegative on a subset* $S \subset \mathbb{R}^n$ if for all $x \in S \cap \mathbb{R}^n_+$, we have $Ax \in \mathbb{R}^n_+$ [17].

A matrix $P \in \mathbb{R}^{n \times n}_+$ that has precisely one entry in each column and each row whose value is 1 and all other entries are zero is called *permutation matrix*. Denote by $\Pi_n \subset \mathbb{R}^{n \times n}_+$ the set of permutation matrices of order n.

In the theory of nonnegative matrices, irreducibility of a matrix plays an important role. We call a matrix A *reducible* if there exists a permutation matrix $P \in \Pi_n$, such that

$$PAP^T = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

where $A_{11},\ A_{22}$ are square. Otherwise A is called *irreducible*.

The direct sum of n matrices $A_i \in \mathbb{R}^{n_i \times n_i}$ is defined by

$$\oplus_{i=1}^n A_i := \operatorname{diag}(A_1, \ldots, A_n) = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{bmatrix}.$$

The following Theorem 1.6 states a canonical form for nonnegative projectors [17], [39].

Theorem 1.6 *Let $A \in \mathbb{R}_+^{n\times n}$ be a nonnegative projector, i.e., $A^2 = A$ and $A \geq 0$, and let A be of rank k. Then, there exists a permutation matrix Π such that*

$$\Pi A \Pi^T = \begin{bmatrix} J & JU & 0 & 0 \\ 0 & 0 & 0 & 0 \\ VJ & VJU & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{1.8}$$

where $U, V \geq 0$ are arbitrary matrices and $J = \oplus_{i=1}^k J_i$, where the matrices $J_i \in \mathbb{R}_+^{l_i\times l_i}$ are positive projectors of rank 1, i.e., $J_i = u_i v_i^T$, where $0 < u_i, v_i \in \mathbb{R}_+^{l_i}$ with $v_i^T u_i = 1$ for $i = 1, \ldots, k$. Conversely, every matrix of the form in (1.8), where $J = \oplus_{i=1}^k J_i$ and $U, V \geq 0$, is a nonnegative projector of rank k.

A matrix A is called Z*-matrix* if its off-diagonal entries are nonpositive. In the literature, a matrix for which $-A$ is a Z-matrix sometimes is called L-matrix, Metzler matrix or essentially positive matrix, see, e.g., [17], [38], [64], [122]. Throughout this work we will use the term $-Z$-matrix.

Lemma 1.7 *Let $A \in \mathbb{R}^{n\times n}$. Then, $e^{At} \geq 0$ for all $t \geq 0$ if and only if A is a $-Z$-matrix.*

Proof. See, e.g., [122]. □

Let $B \geq 0$ with spectral radius $\rho(B)$. A matrix A of the form $A = sI - B$, with $s > 0$, and $s \geq \rho(B)$ is called M*-matrix*. If $s > \rho(B)$ then A is a *nonsingular* M-matrix, if $s = \rho(B)$ then A is a *singular* M-matrix. The class of M-matrices is a subclass of the Z-matrices. Accordingly, a matrix for which $-A$ is an M-matrix we call a $-M$-matrix.

The following Lemma 1.8 and Theorem 1.9 are well-known properties of M-matrices and can be found, e.g., in [17].

Lemma 1.8 *Let A be a $-Z$-matrix with $\sigma(A) \in \mathbb{C}_-$. Then, A is a $-M$-matrix.*

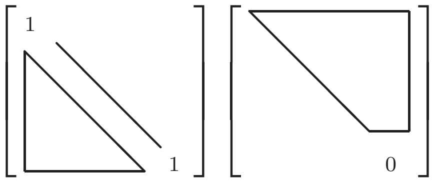
Theorem 1.9 *Let $A \in \mathbb{R}^{n\times n}$ be an irreducible singular M-Matrix. Then,*

- $\operatorname{rank}(A) = n - 1$.
- *Every principal submatrix of A other than A is **nonsingular** M-matrix.*

By Theorem 1.9, for an irreducible singular M-matrix, one can deduce the existence of an LU-decomposition that takes a special form.

Corollary 1.10 *Let $A \in \mathbb{R}^{n\times n}$ be an irreducible singular M-matrix. Then, there exists a unit lower triangular nonsingular M-matrix L and an upper triangular M-matrix U such that*

$$A = LU =$$



. (1.9)

Proof. See [76]. □

A symmetric matrix A is called *positive (semi)definite* and we write $(A \succeq 0)$ $A \succ 0$ if for all $x \neq 0$ we have $(x^T Ax \geq 0)$ $x^T Ax > 0$. If this holds for $-A$ then A is called *negative (semi)definite* and we write $(A \preceq 0)$ $A \prec 0$. For matrices A, B we write $(A \preceq B)$ $A \prec B$ if $(B - A \succeq 0)$ $B - A \succ 0$.

1.4 Generalised inverses

Since we consider descriptor systems, where the matrix E in systems (1a) or (2a) is singular, we will need the concept of generalised inverses. The definitions and notation we adopt here are taken from [17]. Generalised inverses are matrices for which some of the properties of the standard inverses do not hold whereas some others do. If a matrix A is nonsingular, then the matrix $X = A^{-1}$ satisfies the properties

$$AXA = A, \tag{1.10a}$$
$$XAX = X, \tag{1.10b}$$
$$(AX)^T = AX, \tag{1.10c}$$
$$(XA)^T = XA, \tag{1.10d}$$
$$AX = XA, \tag{1.10e}$$
$$XA^{k+1} = A^k \quad \text{for } k \in \mathbb{N}. \tag{1.10f}$$

We only introduce three special cases of generalised inverses that we need in the following.

Definition 1.11 (Semi-inverse) *A matrix that satisfies conditions (1.10a) and (1.10b) is called* semi-inverse *of A and we denote it by A^{ginv}.*

Note that a semi-inverse is not unique. The Moore-Penrose inverse of a matrix A is a special case of a semi-inverse and it is uniquely defined by two additional properties, see, e.g., [29].

Definition 1.12 (Moore-Penrose inverse) *A* Moore-Penrose inverse $A^\dagger$ *of* A *is defined by the properties (1.10a)-(1.10d).*

The following Lemma 1.13 is well known and gives explicit formulas for any semi-inverse and the Moore-Penrose inverse of a special matrix A, see, e.g., [29].

Lemma 1.13 *Let* $A = xy^T,\ x, y \in \mathbb{R}^n \setminus \{0\}$*. Then any semi-inverse is of the form*

$$A^{\text{ginv}} = zw^T,\ z, w \in \mathbb{R}^n,$$

such that $(y^T z)(w^T x) = 1$*. In particular, the Moore-Penrose inverse is given by*

$$A^\dagger = \frac{1}{(x^T x)(y^T y)} yx^T.$$

In general, the Moore-Penrose inverse may be calculated via the reduced *singular value decomposition (SVD)*, see, e.g., [49].

Theorem 1.14 *Let* $A \in \mathbb{R}^{n\times n}$ *with* $\operatorname{rank}(A) = r$*. Then, there exist matrices* $U = [u_1, \ldots, u_r] \in \mathbb{R}^{n\times r}$ *and* $V = [v_1, \ldots, v_r] \in \mathbb{R}^{n\times r}$*, such that* $U^T U = V^T V = I_r$ *and*

$$A = U\Sigma V^T,$$

where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$ *and* $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$*.*

Lemma 1.15 *Let* $A \in \mathbb{R}^{n\times n}$ *and let* $A = U\Sigma V^T$*, where* U, Σ, V *are the matrices of the reduced SVD as in Theorem 1.14. Then, the Moore-Penrose inverse of* A *is given by*

$$A^\dagger = V\Sigma^{-1}U^T.$$

For the explicit solution representation of the systems (1a) or (2a) that we introduce in Section 1.6 we need the Drazin generalised inverse, first introduced in [36]. For a matrix theoretical approach and applications, see, e.g., [29].

Definition 1.16 (Drazin inverse) *Let* $A \in \mathbb{R}^{n\times n}$ *with* $\operatorname{ind}(A) = \nu$*. The Drazin inverse* $A^D \in \mathbb{R}^{n\times n}$ *of* A *is defined by the properties (1.10b), (1.10e) and (1.10f).*

The following result states the Jordan canonical form representation of the Drazin inverse [29], which gives a more intuitive idea of this generalised inverse.

Theorem 1.17 (JCF representation of the Drazin inverse) *Let $E \in \mathbb{R}^{n \times n}$ be such that $\operatorname{ind}(E) = \nu > 0$ and let $P \in \mathbb{R}^{n \times n}$ be a regular matrix such that*

$$E = P \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix} P^{-1},$$

where C is regular and N is nilpotent of index ν. Then,

$$E^D = P \begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1}.$$

Corollary 1.18 (Existence and uniqueness of the Drazin inverse) *Every $E \in \mathbb{R}^{n \times n}$ has one and only one Drazin inverse E^D.*

1.5 Transfer function and H_∞-norms

Let the matrix quintuple $[\,E\,, A\,,\, B\,,\, C\,,\, D]$ denote the system (1) or (2) with a regular matrix pair (E, A), respectively. The function

$$G(\lambda) = C(\lambda E - A)^{-1} B + D$$

is called *transfer function* and λ is called *frequency variable*. Conversely, the quintuple $[\,E\,, A\,,\, B\,,\, C\,,\, D]$ is called *realisation* of G. Note that realisations are in general not unique. Typically, the frequency variable is denoted by s in continuous-time and by z in discrete-time. A transfer function $G(\lambda)$ is called *proper* if

$$\lim_{\lambda \to \infty} G(\lambda) < \infty,$$

and *improper* otherwise. If $\lim_{\lambda \to \infty} G(\lambda) = 0$, then $G(\lambda)$ is called *strictly proper*. Let $H_{\infty,c}$ be the space of all transfer functions that are analytic and bounded in the open right complex half-plane $\mathbb{C}_+$ and let $H_{\infty,d}$ be the space of all transfer functions that are analytic and bounded on $\mathbb{C} \backslash \mathbb{D}$, where $\mathbb{D}$ is the closed unit ball around the origin. The continuous-time and discrete-time H_∞*-norms* are defined by

$$\|G\|_{\infty,c} = \sup_{s \in \mathbb{C}_+} \|G(s)\|_2, \qquad \|G\|_{\infty,d} = \sup_{z \in \mathbb{C} \backslash \mathbb{D}} \|G(z)\|_2, \tag{1.11}$$

respectively, where $\|\cdot\|_2$ denotes the spectral matrix norm.

1.6 Explicit solution representation

Consider the systems (1a) and (2a). In this work, we adopt the classical solvability concept. A vector function $x \in C^1$ is called solution of (1a), if for the assigned input u and the given initial condition x_0 it satisfies (1a) pointwise. A vector sequence x is called solution of (2a), if for the assigned input sequence u and the given initial condition x_0 it satisfies (2a) pointwise.

In order to formulate explicit solution representations of (1a) and (2a), respectively, we need that the matrices E and A commute. If they do not commute and the matrix pair (E, A) is regular, we can obtain commuting matrices by multiplication with a scaling factor as stated in the following Lemma, [26].

Lemma 1.19 *Let (E, A) be a regular matrix pair. Let $\hat{\lambda}$ be chosen such that $\hat{\lambda}E - A$ is nonsingular. Then, the matrices*

$$\hat{E} = (\hat{\lambda}E - A)^{-1}E \text{ and } \hat{A} = (\hat{\lambda}E - A)^{-1}A$$

commute.

In the following, we refer to $\hat{E}, \hat{A}$ as defined in Lemma 1.19 independently of the specific choice of $\hat{\lambda}$. Furthermore, for a matrix B from system (1) or (2) we define

$$\hat{B} := (\hat{\lambda}E - A)^{-1}B. \tag{1.12}$$

Note, that for systems (1a) and (2a), respectively, the scaling by a nonsingular factor such as $(\hat{\lambda}E - A)^{-1}$ does not change the solution.

For the matrices $\hat{E}, \hat{A}$ as defined in Lemma 1.19 and their corresponding Drazin inverses, the following properties hold, see, e.g., [74]:

$$\hat{E}\hat{A}^D = \hat{A}^D\hat{E}, \tag{1.13a}$$
$$\hat{E}^D\hat{A} = \hat{A}\hat{E}^D, \tag{1.13b}$$
$$\hat{E}^D\hat{A}^D = \hat{A}^D\hat{E}^D. \tag{1.13c}$$

Note that if we form matrix products such as $\hat{E}^D\hat{E}$, $\hat{E}^D\hat{A}$, $\hat{E}\hat{A}^D$, $\hat{E}^D\hat{B}$, $\hat{A}^D\hat{B}$, the terms in $\hat{\lambda}$ cancel out, so that these products do not depend on the specific choice of $\hat{\lambda}$, see [74, Chapter 2, Exercise 11]. This can be verified by transforming (E, A) into Weierstraß canonical form in Theorem 1.1. Then, we have

$$\begin{aligned}\hat{E} = (\hat{\lambda}E - A)^{-1}E &= \left(\hat{\lambda}W \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T - W \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} T\right)^{-1} W \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T = \\ &= \left(W \begin{bmatrix} \hat{\lambda}I - J & 0 \\ 0 & \hat{\lambda}N - I \end{bmatrix} T\right)^{-1} W \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T = \\ &= T^{-1} \begin{bmatrix} (\hat{\lambda}I - J)^{-1} & 0 \\ 0 & (\hat{\lambda}N - I)^{-1}N \end{bmatrix} T,\end{aligned}$$

and similarly,

$$\hat{A} = (\hat{\lambda}E - A)^{-1}A = T^{-1} \begin{bmatrix} (\hat{\lambda}I - J)^{-1}J & 0 \\ 0 & (\hat{\lambda}N - I)^{-1} \end{bmatrix} T.$$

For the Drazin inverses of $\hat{E}$ and $\hat{A}$ we obtain

$$\hat{E}^D = T^{-1} \begin{bmatrix} \hat{\lambda}I - J & 0 \\ 0 & 0 \end{bmatrix} T \quad \text{and} \quad \hat{A}^D = T^{-1} \begin{bmatrix} J^D(\hat{\lambda}I - J) & 0 \\ 0 & \hat{\lambda}N - I \end{bmatrix} T.$$

Here, we have used that the matrices J and $(\hat{\lambda}I - J)^{-1}$ commute, and for commuting matrices Z_1, Z_2 with Z_1 regular, we have $(Z_1 Z_2)^D = Z_2^D Z_1^{-1}$, see, e.g., [74]. Therefore, the products

$$\begin{aligned}\hat{E}^D\hat{E} &= T^{-1} \begin{bmatrix} \hat{\lambda}I - J & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (\hat{\lambda}I - J)^{-1} & 0 \\ 0 & (\hat{\lambda}N - I)^{-1}N \end{bmatrix} T = T^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T, \\ \hat{E}^D\hat{A} &= T^{-1} \begin{bmatrix} \hat{\lambda}I - J & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (\hat{\lambda}I - J)^{-1}J & 0 \\ 0 & (\hat{\lambda}N - I)^{-1} \end{bmatrix} T = T^{-1} \begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} T, \\ \hat{E}\hat{A}^D &= T^{-1} \begin{bmatrix} (\hat{\lambda}I - J)^{-1} & 0 \\ 0 & (\hat{\lambda}N - I)^{-1}N \end{bmatrix} \begin{bmatrix} (\hat{\lambda}I - J)J^D & 0 \\ 0 & \hat{\lambda}N - I \end{bmatrix} T = T^{-1} \begin{bmatrix} J^D & 0 \\ 0 & N \end{bmatrix} T,\end{aligned}$$

do not depend on $\hat{\lambda}$. Note that $P_r = \hat{E}^D\hat{E}$ is the unique spectral projector onto S_f^{def} along the deflating subspace corresponding to the eigenvalue ∞ defined in (1.5). Let $\hat{B}$ be defined as in (1.12) and $B = W\tilde{B}$, where $\tilde{B} = \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix}$ is partitioned according to the Weierstraß canonical form of (E, A). Then, we have

$$\hat{B} = (\hat{\lambda}E - A)^{-1}B = \left(\hat{\lambda}W \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T - W \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} T\right)^{-1} W \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix} = T^{-1} \begin{bmatrix} (\hat{\lambda}I - J)^{-1}\tilde{B}_1 \\ (\hat{\lambda}N - I)^{-1}\tilde{B}_2 \end{bmatrix},$$

and we obtain $\hat{E}^D\hat{B} = T^{-1}\begin{bmatrix}\tilde{B}_1\\ \tilde{0}\end{bmatrix}$ and $\hat{A}^D\hat{B} = T^{-1}\begin{bmatrix}J^D\tilde{B}_1\\ \tilde{B}_2\end{bmatrix}$, which are also independent of $\hat{\lambda}$.

The following Theorem 1.20 gives an explicit solution representation in terms of the Drazin inverse.

Theorem 1.20 *Let (E,A) be a regular matrix pair with $E,A \in \mathbb{R}^{n\times n}$ and $\operatorname{ind}(E,A) = \nu$. Let $\hat{E},\hat{A}$ be defined as in Lemma 1.19 and $\hat{B}$ as in (1.12). Furthermore, for the continuous-time case, let $u \in \mathrm{C}^\nu$ and denote by $u^{(i)}, i = 0\dots,\nu-1,$ the i-th derivative of u. Then, every solution $x \in C^1$ to Equation (1a) has the form:*

$$x(t) = e^{\hat{E}^D\hat{A}t}\hat{E}^D\hat{E}v + \int_0^t e^{\hat{E}^D\hat{A}(t-\tau)}\hat{E}^D\hat{B}u(\tau)d\tau - (I-\hat{E}^D\hat{E})\sum_{i=0}^{\nu-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}u^{(i)}(t), \tag{1.14}$$

for some $v \in \mathbb{R}^n$. In the discrete-time case, every solution sequence $x(t)$ to Equation (2a) has the form:

$$x(t) = (\hat{E}^D\hat{A})^t\hat{E}^D\hat{E}v + \sum_{\tau=0}^{t-1}(\hat{E}^D\hat{A})^{t-1-\tau}\hat{E}^D\hat{B}u(\tau) - (I-\hat{E}^D\hat{E})\sum_{i=0}^{\nu-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}u(t+i). \tag{1.15}$$

for some $v \in \mathbb{R}^n$.

Proof. See, e.g., [26], [74]. □

Corollary 1.21 *Under the same assumptions as in Theorem 1.20, the continuous-time initial value problem (1a) has a (unique) solution corresponding to the initial condition x_0 and to the input $u \in C^\nu$ if and only if there exists a vector $v \in \mathbb{R}^n$ such that*

$$x_0 = \hat{E}^D\hat{E}v - (I-\hat{E}^D\hat{E})\sum_{i=0}^{\nu-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}u^{(i)}(0). \tag{1.16}$$

The discrete-time initial value problem (2a) has a (unique) solution corresponding to the initial condition x_0 and to the input sequence u if and only if there exists a vector $v \in \mathbb{R}^n$ such that

$$x_0 = \hat{E}^D\hat{E}v - (I-\hat{E}^D\hat{E})\sum_{i=0}^{\nu-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}u(i). \tag{1.17}$$

Proof. See, e.g., [26], [74]. □

Unlike in the standard case, Corollary 1.21 shows that not for every initial condition x_0, there exists a solution to systems (1a) or (2a). The set of initial conditions for which a solution exists is restricted and depends on the chosen input. This leads to the following definition of consistent initial values.

Definition 1.22 *We call an initial value x_0 in (1a) or in (2a)* consistent with respect to an assigned input u *if (1.16) or (1.17) holds, respectively.*

Chapter 2

Nonnegative matrix theory for positive descriptor systems

As for everything else, so for a mathematical theory:
beauty can be perceived but not explained.
- Arthur Cayley

2.1 Classical Perron-Frobenius theory

The well-known Perron-Frobenius Theorem states that for an elementwise nonnegative matrix, the spectral radius is itself an eigenvalue and has a nonnegative eigenvector. It is named after Oscar Perron and Georg Ferdinand Frobenius. Perron proved the first part of the theorem for positive matrices in 1907 [106], and Frobenius extended it to irreducible and nonnegative matrices in 1912 [43].

This result has many applications in all areas of science and engineering, in particular in economics and population dynamics, see, e.g., [17]. Also, the Perron-Frobenius theorem is widely used in the analysis of standard positive systems, see, e.g., [2], [14], [31], [38], [120]. For instance, stability properties of positive systems are mainly determined by the Perron-Frobenius theory, see Chapter 4.

The classical Perron-Frobenius Theorem, see, e.g., [17, pp. 26/27], states as follows.

Theorem 2.1 (Perron-Frobenius Theorem) *Let $A \geq 0$ have the spectral radius $\rho(A)$. Then $\rho(A)$ is an eigenvalue of A and A has a nonnegative eigenvector v corresponding to $\rho(A)$. If, in addition, A is irreducible, then $\rho(A)$ is simple and A has a positive eigenvector v corresponding to $\rho(A)$. Furthermore, $v > 0$ is unique up to a scalar multiple,* *i.e. if $w > 0$ is an eigenvector of A, then $w = \alpha v,\ \alpha \in \mathbb{R}_+$.*

Since we consider positive descriptor systems as in (1) or (2), where E, A are real $n \times n$ matrices, the dynamics is described by the eigenvalues and eigenvectors associated with the matrix pair (E, A). Therefore, the next section is devoted to the extension of this important theory to matrix pairs.

2.2 Perron-Frobenius theory for matrix pairs

In this section we present a new approach to extend the classical Perron-Frobenius theory to matrix pairs (E, A), where a sufficient condition guarantees that the finite spectral radius of (E, A) is an eigenvalue with a corresponding nonnegative eigenvector [96]. Our approach is based on the construction of projector chains as they were introduced in the context in [52]. For the special case $E = I$ our new approach reduces to the classical Perron-Frobenius theorem for matrices. We present several examples where the new condition holds, whereas previous conditions are not satisfied.

2.2.1 Previous generalisations and their drawbacks

In the literature, several approaches have been presented to generalise the classical Perron-Frobenius theory to matrix pencils or further to matrix polynomials. In [87] the nonnegativity condition for A, which can be stated as $y \geq 0 \Rightarrow Ay \geq 0$, is directly generalised. For the matrix pair (E, A) the condition $E^T y \geq 0 \Rightarrow A^T y \geq 0$ is given, which is sufficient for the existence of a positive eigenvalue and a corresponding nonnegative eigenvector. In [9] a sufficient condition, $(E - A)^{-1} A \geq 0$, for the existence of a positive eigenvalue in $(0, 1)$ and a corresponding positive eigenvector if $(E - A)^{-1} A$ is irreducible, is proved. The relationship of the two ideas from [9] and [87] is studied in [97]. In [93], the condition from [9] is imposed by requiring $E - A$ to be a nonsingular M-matrix and $A \geq 0$. Here, the structure of nonnegative eigenvectors is studied from the combinatorial point of view. In [107] the Perron-Frobenius theory was extended to matrix polynomials, where the coefficient matrices are entrywise nonnegative. Other extensions concerning matrix polynomials are given in [41].

The main drawbacks of the generalisation in [9] is that on the one hand it is a restrictive condition, since $E - A$ is not necessarily invertible, and on the other hand it does not have the classical Perron-Frobenius theory as a special case, where $E = I$. Furthermore, only the existence of a nonnegative real eigenvalue is guaranteed instead of the spectral radius being an eigenvalue. The condition in [87] has the classical Perron-Frobenius theory as a special case but the condition is not easy to verify. Furthermore, for regular matrix pairs (E, A) with singular E this condition never holds, since one can always find a vector $y \le 0$ in the nullspace of E or E^T such that $Ay \le 0$ or $A^T y \le 0$. However, this is the situation that is studied in this work.

Our extension of the Perron-Frobenius theory to regular matrix pairs (E, A) has a number of advantages over the existing conditions in the literature. In Section 2.2.2, for the case of index 1 pencils, we prove an easily computable sufficient condition in Theorem 2.2 that guarantees that the finite spectral radius of (E, A) is an eigenvalue with a corresponding nonnegative eigenvector. We present several examples where the new condition holds, whereas the conditions in [9] and [87] are not satisfied. In the general case (where the index may be arbitrary) presented in Section 2.2.3, we have to impose an additional condition on the projectors, see Lemma 2.7, that is satisfied naturally in the index 1 case. The general sufficient condition that we then prove in Theorem 2.8 is in the index 1 case the same as in Theorem 2.2 and also guarantees that the finite spectral radius of (E, A) is an eigenvalue with a corresponding nonnegative eigenvector. In Corollary 2.10, we prove two further conditions that are equivalent to the condition in Theorem 2.8. All conditions have the classical Perron-Frobenius theory as a special case when $E = I$.

2.2.2 Regular matrix pairs of index at most one

In this subsection we study regular pairs (E, A) of index at most one. This is a special case of the general result of this section that we present in the next subsection. The aim of this section is to gain a more intuitive idea of the projector-based approach before stating and proving the result in its full generality. The techniques used in the index 1 case go back to [52].

Theorem 2.2 *Let (E, A), with $E, A \in \mathbb{R}^{n\times n}$, be a regular matrix pair with $\operatorname{ind}(E, A) \le 1$. Let Q_0 be a projector onto $\ker E$ along the subspace S_0 defined as in (1.7) for $i = 0$, i.e.,*

$$S_0 := \{y \in \mathbb{R}^n :\ Ay \in \operatorname{im} E\}, \tag{2.1}$$

let $P_0 = I - Q_0$, and $A_1 = AP_0$. Then $E_1 := E - AQ_0$ is nonsingular and if

$$E_1^{-1} A_1 \ge 0, \tag{2.2}$$

and $\sigma_f(E, A) \ne \emptyset$, then the finite spectral radius $\rho_f(E, A)$ is an eigenvalue of the matrix pair (E, A) and if $\rho_f(E, A) > 0$, then there exists a nonnegative eigenvector v corre-

sponding to $\rho_f(E,A)$. If $E_1^{-1}A_1$, in addition, is irreducible, then $\rho_f(E,A)$ is simple and $v>0$ is unique up to a scalar multiple.

Proof. Let $\operatorname{ind}(E,A)=0$. Then, E is regular and we have that $Q_0=0$, $E_1=E$ and the condition in Theorem 2.2 reduces to the one of the classical Perron-Frobenius theorem for $E^{-1}A$.
Let $\operatorname{ind}(E,A)=1$ and consider the generalised eigenvalue problem (1.1). We have that E_1 as defined in (1.6) is nonsingular, see [88], and we can premultiply equation (1.1) by E_1^{-1}. By also using that $P_0+Q_0=I$, we obtain

$$E_1^{-1}(\lambda E-A)(P_0+Q_0)v=0,$$

or equivalently

$$(\lambda E_1^{-1}E-E_1^{-1}AP_0-E_1^{-1}AQ_0)v=0.$$

Furthermore, we have $E_1^{-1}E=P_0$ since $E_1P_0=(E-AQ_0)P_0=E$ and $-E_1^{-1}AQ_0=-Q_0$ since $E_1Q_0=(E-AQ_0)Q_0=-AQ_0$. Hence, we obtain

$$[(\lambda I-E_1^{-1}A)P_0+Q_0]v=0,$$

which after multiplication by P_0 and Q_0 from the left is equivalent to the system of two equations

$$\begin{cases} P_0[(\lambda I-E_1^{-1}A)P_0+Q_0]v=0,\\ Q_0[(\lambda I-E_1^{-1}A)P_0+Q_0]v=0.\end{cases} \tag{2.3}$$

We have that Q_0 is a projector onto $\ker E$ along S_0 and by Lemma 1.4 we conclude that $-Q_0E_1^{-1}A$ is also a projector onto $\ker E$ along S_0. Hence, by Lemma 1.4, we have that $(-Q_0E_1^{-1}A)P_0=Q_0P_0=0$. Therefore, by writing $P_0=I-Q_0$ in the first equation of (2.3), the two equations reduce to

$$\begin{cases} (\lambda I-E_1^{-1}A)P_0v=0,\\ Q_0v=0.\end{cases}$$

Since $P_0=P_0P_0$, this is equivalent to

$$\begin{cases} (\lambda I-E_1^{-1}AP_0)P_0v=0,\\ Q_0v=0.\end{cases} \tag{2.4}$$

Setting $x=P_0v$, $y=Q_0v$ and $v=P_0v+Q_0v=x+y$, we obtain a standard eigenvalue problem in the first equation and a linear system in the second equation. From the first equation we know by the Perron-Frobenius Theorem that if $E_1^{-1}AP_0\geq 0$, then the spectral radius of $E_1^{-1}AP_0$ is an eigenvalue with a corresponding nonnegative eigenvector. If $E_1^{-1}AP_0$ is in addition irreducible, then we have that $\rho(E_1^{-1}AP_0)$ is a simple eigenvalue and there exists a corresponding positive eigenvector that is unique up to a scalar multiple. Set $\hat{\lambda} := \rho(E_1^{-1}AP_0)$ and if $\hat{\lambda} \neq 0$, due to (2.4), we can set $\hat{x} = P_0 v$ for the corresponding nonnegative (positive) eigenvector. Then, we obtain

$$\begin{aligned} E_1^{-1}AP_0\hat{x} &= \hat{\lambda}x \\ \Leftrightarrow AP_0\hat{x} &= \hat{\lambda}E_1\hat{x} \\ \Leftrightarrow AP_0\hat{x} &= \hat{\lambda}(E - AQ_0)\hat{x} \\ \Leftrightarrow AP_0P_0v &= \hat{\lambda}EP_0v - AQ_0P_0v \\ \Leftrightarrow A(P_0v + Q_0v) &= \hat{\lambda}Ev \\ \Leftrightarrow Av &= \hat{\lambda}Ev, \end{aligned}$$

which is the generalised eigenvalue problem (1.1). Hence, $\rho(E_1^{-1}AP_0) = \rho_f(E, A)$ and if $\rho_f(E, A) \neq 0$, there exists a corresponding nonnegative eigenvector. This completes the proof. □

Note that in the index 1 case, the projector P_0 is the unique spectral projector onto the right finite deflating subspace of (E, A) defined in (1.5).

Example 2.3 Consider the pair (E, A) given by

$$E = \begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

We have $\operatorname{ind}(E, A) = 1$ and the pair has only one finite eigenvalue $\lambda = 0.5$ with eigenvector $v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$, where we may normalise the eigenvector so that $v_1 > 0$.

To check the sufficient condition (2.2) of Theorem 2.2, we choose a projector $\tilde{Q}_0$ onto $\ker E_0$, e.g.,

$$\tilde{Q}_0 = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix},$$

and get

$$\tilde{E}_1 = E_0 - A_0\tilde{Q}_0 = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}.$$

For the inverse we obtain

$$\tilde{E}_1^{-1} = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix},$$

and a projector Q_0 onto $\ker E$ along S_0 is given by

$$Q_0 = -\tilde{Q}_0 \tilde{E}_1^{-1} A_0 = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}.$$

We then have

$$E_1 = E_0 - A_0 Q_0 = \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix}, \; E_1^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{3}{2} \\ 0 & -1 \end{bmatrix},$$

and we set $P_0 = I - Q_0$. Condition (2.2) then reads

$$E_1^{-1} A_1 = E_1^{-1} A P_0 = \begin{bmatrix} \frac{1}{2} & \frac{3}{2} \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix} \geq 0,$$

and we can apply Theorem 2.2.

For this example, the theories in [9] and [87] cannot be applied, since $(E - A)^{-1} A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \ngeq 0$ and there exists a vector, e.g., $y = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ such that $E^T y \geq 0$ but $A^T y \ngeq 0$.

Example 2.4 Consider a pair (E, A) with $\operatorname{ind}(E, A) = 1$ and $E = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix}$, where E_{11} is nonsingular and $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is partitioned accordingly. For a pencil in this form, $\operatorname{ind}(E, A) = 1$ is equivalent to A_{22} being nonsingular, see, e.g., [74]. We choose any projector $\tilde{Q}_0$ onto $\ker E$, e.g.

$$\tilde{Q}_0 = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix},$$

and compute $\tilde{E}_1$ and $\tilde{E}_1^{-1}$. We obtain

$$\tilde{E}_1 = E - A\tilde{Q}_0 = \begin{bmatrix} E_{11} & -A_{12} \\ 0 & -A_{22} \end{bmatrix}, \; \tilde{E}_1^{-1} = \begin{bmatrix} E_{11}^{-1} & -E_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & -A_{22}^{-1} \end{bmatrix}.$$

Then, we determine a projector Q_0 onto $\ker E$ along $S_0 = \{y \in \mathbb{R}^n : Ay \in \operatorname{im} E\}$ as

$$Q_0 = -\tilde{Q}_0 E_1^{-1} A = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} E_{11}^{-1} & -E_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & -A_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ A_{22}^{-1} A_{21} & I \end{bmatrix}.$$

Furthermore, we get $P_0 = \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & 0 \end{bmatrix}$ and then compute E_1 and E_1^{-1}. We obtain

$$\begin{aligned} E_1 &= E - AQ_0 = \begin{bmatrix} E_{11} - A_{12}A_{22}^{-1}A_{21} & -A_{12} \\ -A_{21} & -A_{22} \end{bmatrix}, \\ E_1^{-1} &= \begin{bmatrix} E_{11}^{-1} & -E_{11}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}E_{11}^{-1} & A_{22}^{-1}A_{21}E_{11}^{-1}A_{12}A_{22}^{-1} - A_{22}^{-1} \end{bmatrix}. \end{aligned}$$

Condition (2.2) then reads as

$$\begin{aligned} E_1^{-1}AP_0 &= \begin{bmatrix} E_{11}^{-1} & -E_{11}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}E_{11}^{-1} & A_{22}^{-1}A_{21}E_{11}^{-1}A_{12}A_{22}^{-1} - A_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & 0 \end{bmatrix} \\ &= \begin{bmatrix} E_{11}^{-1}A_S & 0 \\ -A_{22}^{-1}A_{21}E_{11}^{-1}A_S & 0 \end{bmatrix} \geq 0, \end{aligned}$$

where $A_S = A_{11} - A_{12}A_{22}^{-1}A_{21}$.

Consider again the eigenvalue problem

$$(\lambda E - A)v = 0.$$

In our case we obtain

$$\begin{bmatrix} \lambda E_{11} - A_{11} & -A_{12} \\ -A_{21} & -A_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0.$$

Since E_{11} is nonsingular, we can rewrite this system as

$$\begin{cases} (\lambda I - E_{11}^{-1}A_{11})v_1 - E_{11}^{-1}A_{12}v_2 &= 0, \\ -A_{21}v_1 - A_{22}v_2 &= 0, \end{cases}$$

which is equivalent to

$$\begin{cases} (\lambda I - E_{11}^{-1}A_S)v_1 &= 0, \\ v_2 &= -A_{22}^{-1}A_{21}v_1, \end{cases} \tag{2.5}$$

where $A_S = A_{11} - A_{12}A_{22}^{-1}A_{21}$. Condition (2.2) gives $E_{11}^{-1}A_S \geq 0$ and, by the Perron-Frobenius Theorem, we obtain from the first equation of (2.5) that $\rho(E_{11}^{-1}A_S) =: \hat{\lambda}$ is an eigenvalue with a corresponding eigenvector $v_1 \geq 0$. Using this, from the second equation of (2.5) we obtain

$$v_2 = -A_{22}^{-1}A_{21}v_1 = -\lambda^{-1}A_{22}^{-1}A_{21}\lambda v_1 = -\lambda^{-1}A_{22}^{-1}A_{21}E_{11}^{-1}A_S v_1 \geq 0,$$

since $-A_{22}^{-1}A_{21}E_{11}^{-1}A_S \geq 0$ by (2.2) and we have $\hat{\lambda} \geq 0$ and $v_1 \geq 0$ from the first equation of (2.5).

Remark 2.5 1. Considering the case $E = I$ in Theorem 2.2, we have $P_0 = I$, and the condition $E_1^{-1}A_1 \geq 0$ reduces to the condition $A \geq 0$ of the classical Perron-Frobenius theorem.
2. Condition $E_1^{-1}A_1 \geq 0$, written out, reads as

$$(E - A(I - P_0))^{-1}AP_0 \geq 0,$$

which, without the projectors, would be the condition in [9]:

$$(E - A)^{-1}A \geq 0.$$

Yet, whereas $(E - A(I - P_0))$ is nonsingular by construction, the matrix $E - A$ is not necessarily invertible. Hence, the new condition finds a much broader applicability.
3. Consider the case $\sigma_f(E, A) \neq \emptyset$ and $\rho_f(E, A) = 0$. If $E_1^{-1}A_1 \geq 0$, then we obtain that $\rho_f(E, A) = 0$ is an eigenvalue of (E, A), however, there is not necessarily a corresponding nonnegative eigenvector, as the following Example 2.6 shows.

Example 2.6 *Consider the matrices*

$$\begin{aligned} E &:= T^{-1}\tilde{E}T = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix}, \\ A &:= T^{-1}\tilde{A}T = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

Note that by (1.5) we have that $E = P_r$ is the spectral projector onto the right finite deflating subspace of (E, A). Hence, we obtain

$$E_1^{-1}A_1 = E_1^{-1}AE = E_1^{-1} \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} = 0,$$

and $\sigma_f(E, A) \neq \emptyset$. Therefore, $\rho_f(E, A) = 0$ is an eigenvalue of (E, A). However, the eigenpairs of (E, A) are $(0, \begin{bmatrix} 1 & -0.5 \end{bmatrix}^T)$ and $(\infty, \begin{bmatrix} 1 & -1 \end{bmatrix}^T)$ and there does not exists a nonnegative eigenvector corresponding to $\rho_f(E, A)$.

2.2.3 Regular matrix pairs of general index

In this section we consider a regular matrix pair (E, A) of $\operatorname{ind}(E, A) = \nu$. For $\nu > 1$ we need to define the matrix chain in (1.6) with specific projectors. The following Lemma 2.7 guarantees the existence of projectors with the required property. Canonical projectors as defined in [89] fulfil the condition of Lemma 2.7. An alternative way to construct these projectors along with some examples is presented in Section 2.2.4.

Lemma 2.7 *Let (E, A), with $E, A \in \mathbb{R}^{n\times n}$, be a regular matrix pair of $\operatorname{ind}(E, A) = \nu$. Then, a matrix chain as in (1.6) can be constructed with specific projectors Q_i, P_i such that $Q_i v = 0$ holds for all $v \in S_f^{def}$ and for all $0 \le i < \nu$.*

Proof. From [89] we know that for a regular matrix pair (E, A), we have that

$$\ker E_i \cap \ker A_i = \{0\} \tag{2.6}$$

holds for all $0 \le i < \nu$. Furthermore, from (2.6) or, e.g., from [89] we obtain that

$$\ker E_i \cap \ker E_{i+1} = \{0\} \tag{2.7}$$

for all $0 \le i < \nu - 1$. We now show by induction that we can construct projectors Q_i such that $Q_i v = 0$ holds for all $v \in S_f^{def}$ and for all $0 \le i < \nu$. For the existence of such a Q_0, we have to show that $\ker E_0 \cap S_f^{def} = \{0\}$. Suppose that $x \in \ker E_0 \cap S_f^{def}$. Then from $E_0 x = 0$ we obtain that $x = 0$, since otherwise, by definition, x would be an eigenvector of (E, A) corresponding to the eigenvalue ∞. Thus, we can choose the projector Q_0 onto $\ker E_0$ along some subspace M_0 that completely contains S_f^{def}. This ensures $Q_0 v = 0$ for all $v \in S_f^{def}$. Now, suppose that $Q_i v = 0$ holds for all $v \in S_f^{def}$ and for all $0 \le i \le k$ for some $k < \nu - 1$. Note that for the complementary projectors $P_i = I - Q_i$, this implies that $P_i v = v$ for all $v \in S_f^{def}$. To construct a projector Q_{k+1} onto $\ker E_{k+1}$ such that $Q_{k+1} v = 0$ holds for all $v \in S_f^{def}$, we have to show that $\ker E_{k+1} \cap S_f^{def} = \{0\}$. For this, suppose that $0 \ne x \in \ker E_{k+1} \cap S_f^{def}$. Then, by using the assumption, we obtain

$$0 \;=\; E_{k+1}x = (E_0 - A_0 Q_0 - \ldots - A_k Q_k)x = E_0 x,$$

from which we again conclude that $x = 0$, since otherwise, by definition, x would be an eigenvector of (E, A) corresponding to the eigenvalue ∞. Thus, we can choose the projector Q_{k+1} onto $\ker E_{k+1}$ along some subspace M_{k+1} that completely contains S_f^{def}. This ensures $Q_{k+1} v = 0$ for all $v \in S_f^{def}$ and completes the proof. □

Note that for $\nu = 1$, condition $Q_0 v = 0$ holds automatically for all $v \in S_f^{def}$ and in particular for all eigenvectors, see (2.4).

The following Theorem 2.8 states our main result. We prove a new, broadly applicable Perron-Frobenius-type condition for matrix pairs (E, A) in the general index case.

Theorem 2.8 *Let (E, A), with $E, A \in \mathbb{R}^{n\times n}$, be a regular matrix pair of $\operatorname{ind}(E, A) = \nu$. Let a matrix chain as in (1.6) be constructed with projectors Q_i as in Lemma 2.7. If*

$$E_\nu^{-1} A_\nu \ge 0, \tag{2.8}$$

and $\sigma_f(E,A) \neq \emptyset$, then the finite spectral radius $\rho_f(E,A)$ is an eigenvalue of (E,A) and if $\rho_f(E,A) > 0$, then there exists a corresponding nonnegative eigenvector $v \geq 0$. If $E_\nu^{-1}A_\nu$ is, in addition, irreducible, then we have that $\rho_f(E,A)$ is simple and $v > 0$ is unique up to a scalar multiple.

Proof. Consider the generalised eigenvalue problem (1.1)

$$(\lambda E - A)v = 0.$$

If v is an eigenvector corresponding to a finite eigenvalue λ, i.e., $v \in S_f^{def}$, then we have $Q_i v = 0$ for all $0 \leq i < \nu$ and we can equivalently express (1.1) as

$$\begin{aligned} (\lambda(E - A_0Q_0 - A_1Q_1 - \ldots - A_{\nu-1}Q_{\nu-1}) - A)v &= 0 \\ \Leftrightarrow (\lambda E_\nu - A)v &= 0 \\ \Leftrightarrow (\lambda I - E_\nu^{-1}A)v &= 0. \end{aligned} \tag{2.9}$$

By construction, we have that $v = P_0 \cdots P_{\nu-1}v$ and we obtain that (2.9) is equivalent to

$$\begin{aligned} (\lambda I - E_\nu^{-1}A)P_0 \cdots P_{\nu-1}v &= 0 \\ \Leftrightarrow (\lambda I - E_\nu^{-1}AP_0 \cdots P_{\nu-1})P_0 \cdots P_{\nu-1}v &= 0 \\ \Leftrightarrow (\lambda I - E_\nu^{-1}A_\nu)v &= 0. \end{aligned} \tag{2.10}$$

Note, that in this way, we have shown that any finite eigenpair of (E,A) is an eigenpair of $E_\nu^{-1}A_\nu$. Conversely, by (2.10), we have that any eigenpair (λ, v) of $E_\nu^{-1}A_\nu$ with $\lambda \neq 0$ is a finite eigenpair of (E,A). Since $E_\nu^{-1}A_\nu \geq 0$, by the classical Perron-Frobenius Theorem we obtain that $\rho(E_\nu^{-1}A_\nu)$ is an eigenvalue of $E_\nu^{-1}A_\nu$ and there exists a corresponding eigenvector $v \geq 0$. Since we have assumed that $\sigma_f(E,A) \neq \emptyset$, we have that $\rho(E_\nu^{-1}A_\nu) = \rho_f(E,A)$ is also an eigenvalue of (E,A). If $\rho(E_\nu^{-1}A_\nu) > 0$, then there exists a corresponding nonnegative eigenvector. $\square$

Remark 2.9 *In Theorem 2.8 it is shown that any eigenpair (λ, v) of $E_\nu^{-1}A_\nu$ with $\lambda \neq 0$ is a finite eigenpair of (E,A). However, this is not necessarily the case if $\lambda = 0$, since an eigenvalue 0 of $E_\nu^{-1}A_\nu$ can correspond either to the eigenvalue 0 of (E,A) or to the eigenvalue ∞ of (E,A). One can see this by considering an eigenvector w corresponding to an infinite eigenvalue of (E,A), i.e., $Ew = 0$. Then, we obtain $E_\nu^{-1}A_\nu w = 0$, since $P_0 \cdots P_{\nu-1}w = 0$. Since we have assumed that $\sigma_f(E,A) \neq \emptyset$, we have that $\rho(E_\nu^{-1}A_\nu) = \rho_f(E,A) = 0$ is an eigenvalue of (E,A). However, a corresponding nonnegative eigenvector does not necessarily exist as Example 2.6 shows.*

Corollary 2.10 *Let P_r be the spectral projector of the matrix pair (E, A) onto the right finite deflating subspace S_f^{def} defined in (1.5), and let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19. Under the assumptions of Theorem 2.8, each of the conditions*

$$\begin{aligned} P_r E_\nu^{-1} A &\geq 0, & (2.11)\\ E_\nu^{-1} A \hat{E}^D \hat{E} &\geq 0 & (2.12)\\ \hat{E}^D \hat{A} &\geq 0 & (2.13) \end{aligned}$$

is equivalent to condition (2.8), respectively.

Proof. From [89, Theorem 3.1, Section 4] we obtain that for projectors as in Lemma 2.7, we have $P_0 \ldots P_{\nu-1} = P_r = \hat{E}^D \hat{E}$ and

$$E_\nu^{-1} A_\nu = E_\nu^{-1} A P_r = P_r E_\nu^{-1} A = \hat{E}^D \hat{A}. \tag{2.14}$$

□

Remark 2.11 Condition

$$E_\nu^{-1} A \geq 0, \tag{2.15}$$

can also be proved to be sufficient in Theorem 2.8, see Equation (2.9), yet there is no evidence for it to ever hold.

2.2.4 Construction of projectors

In Section 2.2.3, Lemma 2.7, we have proved the existence of specific projectors for constructing the matrix chain in (1.6) in order to establish a sufficient condition in Theorem 2.8 for $\rho_f(E, A)$ to be an eigenvalue with a corresponding nonnegative eigenvector. In [89], projectors with properties as in Lemma 2.7 are called canonical. Note that canonical projectors are not unique for $\mathrm{ind}(E, A) > 1$. In [89], motivated by a decoupling procedure of differential-algebraic equations, specific canonical projectors, the so called *completely decoupling projectors*, are defined by the property $Q_i = -Q_i P_{i+1} \cdots P_{\nu-1} E_{\nu-1}^{-1} A_i$ for all $i = 0, \ldots, \nu-1$. It is shown in [89, Theorem 2.2] that such projectors exist and a constructive proof is given. However, to keep the present work self-contained, we provide an alternative procedure to construct canonical projectors with properties as in Lemma 2.7.

First, we will formulate the construction procedure in the general case and give a proof by induction. Then, in Section 2.2.5, we will exemplarily show how this procedure works in the index $\nu = 2$ case and give two examples for $\nu = 2$.

Consider a regular matrix pair (E, A) of $\operatorname{ind}(E, A) = \nu$. We make the following observations:

1. For fixed projectors $Q_0, \ldots, Q_{\nu-2}$, the projector $Q_{\nu-1}$ is uniquely defined by being a projector onto $\ker E_{\nu-1}$ along $S_{\nu-1}$, see [89].
2. Consider the sets S_i as defined in (1.7). We have that $S_f^{def} \subseteq S_0$, since for any $v \in S_f^{def}$ there exists a $w \in S_f^{def}$ such that $Av = Ew$, and hence, $Av \in \operatorname{im} E$, i.e., $v \in S_0$. Furthermore, we have that $S_0 \subseteq S_1 \subseteq \ldots \subseteq S_{\nu-1}$, see [88], and therefore, $S_f^{def} \subseteq S_{\nu-1}$. From this we conclude that $Q_{\nu-1}v = 0$ holds for all $v \in S_f^{def}$.

In the following recursive constructions of matrix and projector chains, we denote by $E_j^{(i)}, A_j^{(i)}, Q_j^{(i)}, P_j^{(i)}$ the i-th iterate of E_j, A_j, Q_j, P_j in the recursive construction.

With the basic construction of projectors Q_i for $i = 1, \ldots, \nu - 1$ as in [89], i.e., $Q_jQ_i = 0$ for $j > i$, we construct the chain in (1.6) and set $E_j^{(1)} = E_j$, $A_j^{(1)} = A_j$, $Q_j^{(1)} = Q_j$, and $P_j^{(1)} = P_j$. Now, to obtain projectors with properties as in Lemma 2.7, we redefine the initial projectors by the procedure in Algorithm 1:

Algorithm 1: Construction of completely decoupling projectors

Input: projectors $Q_i^{(1)}$ for $i = 1, \ldots, \nu - 1$ such that $Q_jQ_i = 0$ for $j > i$
Output: projectors $Q_i^{(2^{i+1})}$ such that $Q_i^{(2^{i+1})}v = 0$ for all $v \in S_f^{def}$

for $i = 0, \ldots, \nu - 1$ **do**
  for $j = 1, \ldots, \nu - i$ **do**
    $Q_{\nu-j}^{(new)} = -Q_{\nu-j}^{(old)}(E_\nu^{(old)})^{-1}A_{\nu-j}^{(old)}$, where old denotes an appropriate index;
    redefine the last part of the chain using $Q_{\nu-j}^{(new)}$;
    apply Algorithm 1 to $Q_{\nu-j+1}, \ldots, Q_{\nu-1}$ in order to regain the completely decoupling property, i.e. $Q_{\nu-j+1}v = \ldots = Q_{\nu-1}v = 0$ for all $v \in S_f^{def}$.

Theorem 2.12 *For the projectors $Q_i^{(2^{i+1})}$ computed in Algorithm 1 we have $Q_i^{(2^{i+1})}v = 0$ for all $v \in S_f^{def}$.*

Proof. We perform an induction over the length $k = \nu - i$ of the chain in (1.6), where i is the index variable in Algorithm 1.

Let $k = 1$. Without loss of generality, we can consider the index $\nu = 1$ (and $i = 0$) case. We take any projector $Q_0^{(1)}$ onto $\ker E$ and having computed $E_1^{(1)}$ we set $Q_0^{(2)} = -Q_0^{(1)}(E_1^{(1)})^{-1}A$, which by Lemma 1.4 fulfils $Q_0^{(2)}v = 0$ for all $v \in S_f^{def}$.

Suppose that for some chain of length $k > 1$ we can construct completely decoupling projectors and consider a chain of length $k+1$. Without loss of generality we consider the index $\nu = k+1$ case, i.e., we have an initial chain with projectors $Q_0^{(1)}, \ldots, Q_{\nu-1}^{(1)}$, such that $Q_j^{(1)} Q_i^{(1)} = 0$ holds for $j > i$ and start Algorithm 1. Note, that this is also true for any intermediate chain of length $k+1$ in a general index $\nu > k+1$ case due to Lemma 2.7.

Now, we have to subsequently redefine projectors $Q_{\nu-j}$ for $j = 1, \ldots, \nu - i$ and have to show that the redefined projectors are completely decoupling. Therefore, we perform an induction over j. Let $j = 1$. We set $Q_{\nu-1}^{(2)} = -Q_{\nu-1}^{(1)} (E_\nu^{(1)})^{-1} A_{\nu-1}^{(1)}$ that by Lemma 1.4 fulfils $Q_{\nu-1}^{(2)} v = 0$ for all $v \in S_f^{def}$.

Suppose, we have completely decoupling projectors $Q_{\nu-1}, \ldots, Q_{\nu-j}$ for some $1 < j < \nu - i$.

Set $Q_{\nu-j-1}^{(2)} = -Q_{\nu-j-1}^{(1)} (E_\nu^{(k)})^{-1} A_{\nu-j-1}^{(1)}$, where k is an appropriate index. By Lemma 1.5, we have that $Q_{\nu-j-1}^{(2)}$ is a projector. By the definition of deflating subspace, we have that for all $v \in S_f^{def}$ there exists $w \in S_f^{def}$ with $Av = Ew$. Therefore, we obtain $Q_{\nu-j-1}^{(2)} v = 0$ for all $v \in S_f^{def}$, since

$$\begin{aligned}
Q_{\nu-j-1}^{(2)} v &= -Q_{\nu-j-1}^{(1)} (E_\nu^{(k)})^{-1} A_{\nu-j-1}^{(1)} v = -Q_{\nu-j-1}^{(1)} (E_\nu^{(k)})^{-1} A_{\nu-j-2}^{(1)} P_{\nu-j-2}^{(1)} v = \\
&= -Q_{\nu-j-1}^{(1)} (E_\nu^{(k)})^{-1} A_{\nu-j-2}^{(1)} (I - Q_{\nu-j-2}^{(1)}) v = \\
&= -Q_{\nu-j-1}^{(1)} (E_\nu^{(k)})^{-1} A_{\nu-j-2}^{(1)} v - Q_{\nu-j-1}^{(1)} Q_{\nu-j-2}^{(1)} v = \\
&= -Q_{\nu-j-1}^{(1)} (E_\nu^{(k)})^{-1} A_{\nu-j-3}^{(1)} P_{\nu-j-3}^{(1)} v = \ldots = -Q_{\nu-j-1}^{(1)} (E_\nu^{(k)})^{-1} A_0 v = \\
&= -Q_{\nu-2}^{(1)} (E_\nu^{(k)})^{-1} E_0 w = \\
&= -Q_{\nu-j-1}^{(1)} (I - Q_0^{(1)} - \ldots - Q_{\nu-j-1}^{(1)} - Q_{\nu-j} - \ldots - Q_{\nu-1}) w,
\end{aligned}$$

where $Q_{\nu-j} w = \ldots = Q_{\nu-1} w = 0$, since $Q_{\nu-j}, \ldots, Q_{\nu-1}$ are completely decoupling. Furthermore, we have $Q_{\nu-j-1}^{(1)} Q_{i_1}^{(1)} = 0$ for $i_1 = 0, \ldots, \nu-j-2$ and $Q_{\nu-j-1}^{(1)} (I - Q_{\nu-j-1}^{(1)}) = 0$. Hence, we obtain

$$Q_{\nu-j-1}^{(2)} v = 0.$$

This completes the induction over k and we have shown that we can construct a $Q_0^{(2)}$ such that $Q_0^{(2)} v = 0$ for all $v \in S_f^{def}$.

We redefine the chain starting from $Q_0^{(2)}$ and consider the chain starting from Q_1. The new chain has length k and we can construct completely decoupling projectors by applying the induction assumption. This completes the proof. □

In total, we have to make $\sum_{i=0}^{\nu-1} (2^{i+1} - 1)$ updates of the projectors Q_i. The sufficient condition of Theorem 2.8 is then checked using $E_{\nu-1}^{(2^\nu)}$ instead of $E_{\nu-1}$ and reads

$$\begin{aligned} E_{\nu-1}^{(2^\nu)} A P_0^{(2)} P_1^{(4)} \cdots P_{\nu-1}^{(2^\nu)} &\geq 0 \\ \Leftrightarrow E_{\nu-1}^{(2^\nu)} A_{\nu-1}^{(2^{\nu-1})} P_{\nu-1}^{(2^\nu)} &\geq 0. \end{aligned}$$

So far the described procedure is merely of theoretical value. For a discussion of how to apply this procedure numerically, see [78].

2.2.5 Examples

We now show how the projectors are constructed in Algorithm 1 for the index $\nu = 2$ case and give two examples.

We start by choosing any projectors $Q_0^{(1)}, Q_1^{(1)}$ onto $\ker E_0^{(1)}, \ker E_1^{(1)}$, respectively. We then determine $E_2^{(1)}$ and set $Q_1^{(2)} = -Q_1^{(1)}(E_2^{(1)})^{-1}A_1^{(1)}$. Then we have $Q_1^{(2)}v = 0$ for all $v \in S_f^{def}$. By using $Q_1^{(2)}$ we compute $E_2^{(2)}$ and proceed by setting $Q_0^{(2)} = -Q_0^{(1)}(E_2^{(2)})^{-1}A_0^{(1)}$, which is a projector by Lemma 1.5. For any $v \in S_f^{def}$ we have $w \in S_f^{def}$ such that

$$Q_0^{(2)}v = -Q_0^{(1)}(E_2^{(2)})^{-1}Av = -Q_0^{(1)}(E_2^{(2)})^{-1}Ew = -Q_0^{(1)}(I - Q_0^{(1)} - Q_1^{(2)})w = 0,$$

since $Q_1^{(2)}w = 0$. Here we have used the properties $(E_2^{(2)})^{-1}A_i^{(1)}Q_i^{(1)} = -Q_i^{(1)}$ for $i = 0, 1$ and

$$\begin{aligned} E_2^{(2)} &= E_0^{(1)} - A_0^{(1)}Q_0^{(1)} - A_1^{(1)}Q_1^{(2)} \\ \Leftrightarrow I &= (E_2^{(2)})^{-1}E_0^{(1)} - (E_2^{(2)})^{-1}A_0^{(1)}Q_0^{(1)} - (E_2^{(2)})^{-1}A_1^{(1)}Q_1^{(2)} \\ \Leftrightarrow I &= (E_2^{(2)})^{-1}E_0^{(1)} + Q_0^{(1)} + Q_1^{(2)}. \end{aligned}$$

By using $Q_0^{(2)}$ we compute $E_1^{(2)}$ and $A_1^{(2)}$. Now, we proceed as in the case $\nu = 1$ to define $Q_1^{(3)}$ as a projector onto $\ker E_1^{(2)}$. To ensure that it projects along S_1 we again compute $E_2^{(3)}$, set $Q_1^{(4)} = -Q_1^{(3)}(E_2^{(3)})^{-1}A_1^{(2)}$ and obtain that $Q_1^{(4)}v = 0$ for all $v \in S_f^{def}$. Finally, we compute $E_2^{(4)}$. The sufficient condition of Theorem 2.8 is then checked with $E_2^{(4)}$ instead of E_2 and reads $E_2^{(4)}AP_0^{(2)}P_1^{(4)} \geq 0$. For an illustration of the recursive construction of the projectors in the index 2 case with the properties required in Theorem 2.8, see Figure 2.1.

We now present two index $\nu = 2$ examples, where condition (2.8) of Theorem 2.8 holds, whereas the conditions in [9] and [87] do not hold.

Example 2.13 Consider the matrix pair (E, A) with

$$E = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 1 \end{array}\right] \quad \text{and} \quad A = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 2 \end{array}\right].$$

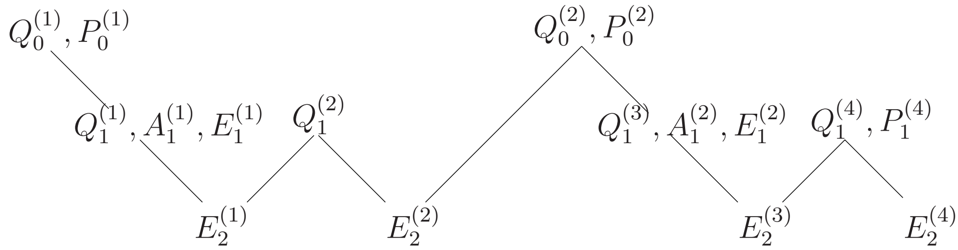


Figure 2.1: Illustration of the recursive construction of projectors in the index 2 case. Top down, we have the chain matrices in increasing order. From left to right, we have the successive calculation of these.

We have that (E, A) is regular with $\operatorname{ind}(E, A) = 2$ and there is one finite eigenvalue $\rho_f(E, A) = 2$ and a corresponding eigenvector $\begin{bmatrix} 0 & 0 & v_3 \end{bmatrix}^T$, which can be chosen so that $v_3 > 0$.

We compute the matrix chain by setting, e.g.,

$$Q_0^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E_1^{(1)} = E - AQ_0^{(1)} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_1^{(1)} = A_0 P_0^{(1)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

We choose, e.g.,

$$Q_1^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad P_1^{(1)} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and compute

$$E_2^{(1)} = E_1^{(1)} - A_1^{(1)} Q_1^{(1)} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad (E_2^{(1)})^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then, we compute the projector onto $\ker E_1^{(1)}$ along S_1 by setting

$$Q_1^{(2)} = -Q_1^{(1)} (E_2^{(1)})^{-1} A_1^{(1)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

and determine

$$E_2^{(2)} = E_1^{(1)} - A_1^{(1)} Q_1^{(2)} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad (E_2^{(2)})^{-1} = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We set

$$Q_0^{(2)} = -Q_0^{(1)}(E_2^{(2)})^{-1}A_0 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad P_0^{(2)} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and compute

$$E_1^{(2)} = E_0 - A_0 Q_0^{(2)} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad A_1^{(2)} = A_0 P_0^{(2)} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Choosing $Q_1^{(3)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, we determine

$$E_2^{(3)} = E_1^{(2)} - A_1^{(2)} Q_1^{(3)} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E_2^{(2)} \quad \text{and} \quad (E_2^{(3)})^{-1} = (E_2^{(2)})^{-1},$$

and verify that $Q_1^{(4)} = -Q_1^{(3)}(E_2^{(3)})^{-1}A_1^{(2)} = Q_1^{(3)}$. We finally set $P^{(4)} = I - Q_1^{(4)}$. The sufficient condition (2.8) of Theorem 2.8 then holds, since

$$(E_2^{(4)})^{-1} A P_0^{(2)} P_1^{(4)} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \geq 0.$$

The condition in [9], however, is not satisfied, since

$$(E - A)^{-1} A = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \not\geq 0.$$

Also the condition in [87] does not hold, since, e.g., for $y = \begin{bmatrix} -1 & 1 & 1 \end{bmatrix}^T$ we have $Ey \geq 0$ but $Ay \not\geq 0$. Note, that we have $P_r = P_0^{(2)} P_1^{(4)}$, yet, condition (2.15) does not hold, since $(E_2^{(4)})^{-1} A \not\geq 0$.

Example 2.14 Consider the regular matrix pair (E, A) of $\operatorname{ind}(E, A) = 2$, where

$$E = \begin{bmatrix} E_{11} & E_{12} & 0 & 0 \\ E_{21} & E_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 0 & 0 & 0 & A_{14} \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & A_{33} & 0 \\ A_{41} & 0 & 0 & 0 \end{bmatrix}.$$

Note, that every regular matrix pair of index 2 can be equivalently transformed into such a form, where $A_{14}, A_{41}, A_{33}, E_{22}$ are square regular matrices, see [74]. We choose

$$Q_0^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix},$$

and compute

$$P_0^{(1)} = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad E_1^{(1)} = \begin{bmatrix} E_{11} & E_{12} & 0 & -A_{14} \\ E_{21} & E_{22} & 0 & 0 \\ 0 & 0 & -A_{33} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Choosing

$$Q_1^{(1)} = \begin{bmatrix} I & 0 & 0 & 0 \\ -E_{22}^{-1}E_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ A_{14}^{-1}\tilde{E}_{11} & 0 & 0 & 0 \end{bmatrix},$$

where $\tilde{E}_{11} = E_{11} - E_{12}E_{22}^{-1}E_{21}$, we obtain

$$P_1^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ E_{22}^{-1}E_{21} & I & 0 & 0 \\ 0 & 0 & I & 0 \\ -A_{14}^{-1}\tilde{E}_{11} & 0 & 0 & I \end{bmatrix}, \quad A_1^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ A_{41} & 0 & 0 & 0 \end{bmatrix}, \quad \text{and}$$

$$E_2^{(1)} = \begin{bmatrix} E_{11} & E_{12} & 0 & -A_{14} \\ E_{21} + A_{22}E_{22}^{-1}E_{21} & E_{22} & 0 & 0 \\ 0 & 0 & -A_{33} & 0 \\ -A_{41} & 0 & 0 & 0 \end{bmatrix},$$

$$(E_2^{(1)})^{-1} = \begin{bmatrix} 0 & 0 & 0 & -A_{41}^{-1} \\ 0 & E_{22}^{-1} & 0 & E_{22}^{-1}(E_{21} + A_{22}E_{22}^{-1}E_{21})A_{41}^{-1} \\ 0 & 0 & -A_{33}^{-1} & 0 \\ -A_{14}^{-1} & A_{14}^{-1}E_{12}E_{22}^{-1} & 0 & -A_{14}^{-1}(\tilde{E}_{11} - E_{12}E_{22}^{-1}A_{22}E_{22}^{-1}E_{21})A_{41}^{-1} \end{bmatrix}.$$

We verify that $Q_1^{(2)} = -Q_1^{(1)}(E_2^{(1)})^{-1}A_1^{(1)} = Q_1^{(1)}$ and, hence, $P_1^{(2)} = P_1^{(1)}$, $A_1^{(2)} = A_1^{(1)}$, $E_2^{(2)} = E_2^{(1)}$ and $(E_2^{(2)})^{-1} = (E_2^{(1)})^{-1}$. Setting

$$Q_0^{(2)} = -Q_0^{(1)}(E_2^{(2)})^{-1}A_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ A_{14}(\tilde{E}_{11} - E_{12}E_{22}^{-1}A_{22}E_{22}^{-1}E_{21}) & -A_{14}E_{12}E_{22}^{-1}A_{22} & 0 & I \end{bmatrix},$$

we compute

$$
\begin{aligned}
P_0^{(2)} &= \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -A_{14}(\tilde{E}_{11} - E_{12}E_{22}^{-1}A_{22}E_{22}^{-1}E_{21}) & A_{14}E_{12}E_{22}^{-1}A_{22} & 0 & 0 \end{bmatrix}, \\
E_1^{(2)} &= E - AQ_0^{(2)} = \begin{bmatrix} E_{12}(I + E_{22}^{-1}A_{22})E_{22}^{-1}E_{21} & E_{12}(I + E_{22}^{-1}A_{22}) & 0 & -A_{14} \\ E_{21} & E_{22} & 0 & 0 \\ 0 & 0 & -A_{33} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\
A_1^{(2)} &= AP_0^{(2)} = \begin{bmatrix} -E_{11} + E_{12}(I + E_{22}^{-1}A_{22})E_{22}^{-1}E_{21} & E_{12}E_{22}^{-1}A_{22} & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ A_{41} & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

Choosing

$$
Q_1^{(3)} = \begin{bmatrix} I & 0 & 0 & 0 \\ -E_{22}^{-1}E_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},
$$

we determine

$$
\begin{aligned}
P_1^{(3)} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ E_{22}^{-1}E_{21} & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}, \\
E_2^{(3)} &= \begin{bmatrix} E_{11} + E_{12}E_{22}^{-1}A_{22}E_{22}^{-1}E_{21} & E_{12} + E_{12}E_{22}^{-1}A_{22} & 0 & -A_{14} \\ E_{21} + A_{22}E_{22}^{-1}E_{21} & E_{22} & 0 & 0 \\ 0 & 0 & -A_{33} & 0 \\ -A_{41} & 0 & 0 & 0 \end{bmatrix},
\end{aligned}
$$

and

$$
(E_2^{(3)})^{-1} = \begin{bmatrix} 0 & 0 & 0 & -A_{41}^{-1} \\ 0 & E_{22}^{-1} & 0 & \check{E}_{24} \\ 0 & 0 & -A_{33}^{-1} & 0 \\ -A_{14}^{-1} & \check{E}_{42} & 0 & \check{E}_{44} \end{bmatrix},
$$

where

$$
\begin{aligned}
&\check{E}_{24} = E_{22}^{-1}(E_{21} + A_{22}E_{22}^{-1}E_{21})A_{41}^{-1}, \\
&\check{E}_{42} = A_{14}^{-1}E_{12}(I + E_{22}^{-1}A_{22})E_{22}^{-1}, \\
&\check{E}_{44} = -A_{14}^{-1}(\tilde{E}_{11} - E_{12}E_{22}^{-1}A_{22}(I + E_{22}^{-1}A_{22})E_{22}^{-1}E_{21})A_{41}^{-1}.
\end{aligned}
$$

We verify that $Q_1^{(4)} = -Q_1^{(3)}(E_2^{(3)})^{-1}A_1^{(2)} = Q_1^{(3)}$ and, hence, $P_1^{(4)} = P_1^{(3)}$, $E_2^{(4)} = E_2^{(3)}$ and $(E_2^{(4)})^{-1} = (E_2^{(3)})^{-1}$. The sufficient condition (2.8) of Theorem 2.8 then reads as

$$(E_2^{(4)})^{-1}A_1^{(2)}P_1^{(4)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ E_{22}^{-1}A_{22}E_{22}^{-1}E_{21} & E_{22}^{-1}A_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ A_{14}^{-1}E_{12}E_{22}^{-1}A_{22}E_{22}^{-1}A_{22}E_{22}^{-1}E_{21} & A_{14}^{-1}E_{12}E_{22}^{-1}A_{22}E_{22}^{-1}A_{22} & 0 & 0 \end{bmatrix} \geq 0.$$

Consider again the eigenvalue problem

$$(\lambda E - A)v = 0.$$

For the given matrices E and A, we obtain

$$\begin{bmatrix} \lambda E_{11} & \lambda E_{12} & 0 & -A_{14} \\ \lambda E_{21} & \lambda E_{22} - A_{22} & 0 & 0 \\ 0 & 0 & -A_{33} & 0 \\ -A_{14} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = 0.$$

Since A_{41} and A_{33} are nonsingular, we obtain $v_1 = v_3 = 0$ and the following system of equations

$$\begin{cases} \lambda E_{12}v_2 - A_{14}v_4 &= 0, \\ (\lambda E_{22} - A_{22})v_2 &= 0, \end{cases}$$

which is equivalent to

$$\begin{cases} (\lambda I - E_{22}^{-1}A_{22})v_2 &= 0, \\ v_4 &= \lambda A_{14}^{-1}E_{12}v_2. \end{cases}$$

Condition (2.8) gives $E_{22}^{-1}A_{22} \geq 0$ and, hence, we obtain from the first equation that $\rho(E_{22}^{-1}A_{22}) =: \lambda$ is an eigenvalue and there exists a corresponding eigenvector $v_2 \geq 0$. By using this, we obtain from the second equation that

$$v_4 = \lambda A_{14}^{-1}E_{12}v_2 = A_{14}^{-1}E_{12}E_{22}^{-1}A_{22}v_2 = -\lambda^{-1}A_{14}^{-1}E_{12}E_{22}^{-1}A_{22}E_{22}^{-1}A_{22}v_2 \geq 0,$$

since $A_{14}^{-1}E_{12}E_{22}^{-1}A_{22}E_{22}^{-1}A_{22} \geq 0$ by (2.8) and $\lambda \geq 0$, $v_2 \geq 0$ from the first equation.

The condition in [9], however, is not necessarily applicable, since $E - A$ may not be invertible if $E_{22} - A_{22}$ is not. Also the condition in [87] does not hold, since we may choose y_1, y_2 in $y = \begin{bmatrix} y_1 & y_2 & y_3 & y_4 \end{bmatrix}^T$ such that $Ey \geq 0$ and choose y_3, y_4 such that $Ay \ngeq 0$.

Summary

In this section, we have proposed a new generalisation of the well-known Perron-Frobenius theory to matrix pairs (E, A) that, unlike previous such generalisations, is suitable for the analysis of positive descriptor systems. We have presented several examples where the new condition holds whereas previous generalisations do not.

2.3 Nonnegativity of the Drazin inverse

The results of this short section give characterisations of positivity of the Drazin inverse, which will be useful, for instance, since the explicit solution representation given in 1.6 is stated in terms of the Drazin inverse.

The following Lemma 2.15 is given as an exercise in [17].

Lemma 2.15 (Positivity of the Drazin inverse) *Let $A \in \mathbb{R}^{n\times n}$ with $\operatorname{ind}(A) = k$. Then, $A^D \geq 0$ if and only if from $Ax \in \mathbb{R}^n_+ + \ker(A^k)$ and $x \in \operatorname{im}(A^k)$, it follows that $x \geq 0$.*

Proof. "$\Rightarrow$" Let $A^D \geq 0$. Furthermore, let $x \in \operatorname{im}(A^k)$, i.e., there exists $y \in \mathbb{R}$ such that $A^k y = x$, and let $Ax = u + v$, where $u \geq 0$ and $A^k v = 0$. Then, by using the properties of the Drazin inverse in Definition 1.16 we obtain

$$\begin{aligned} x &= A^k y = A^D A^{k+1} y = A^D A x = A^D u + A^D v = A^D u + A^D A A^D v = \\ &= A^D u + (A^D)^{k+1} \underbrace{A^k v}_{=0} = A^D u. \end{aligned}$$

Since $u \geq 0$ and $A^D \geq 0$, we have $x \geq 0$.
"$\Leftarrow$" To show $A^D \geq 0$, let $w \in \mathbb{R}^n_+$ be arbitrarily chosen.
Step 1. Show $A^D w \in \operatorname{im}(A^k)$ and $AA^D w \in \mathbb{R}^n_+ + \ker(A^k)$.
For this purpose, we decompose w into $w = AA^D w + (I - AA^D)w =: u + v$, where $u \in \operatorname{im}(A^k)$, since $u = AA^D w = A^k (A^D)^k w$ and $v \in \ker(A^k)$, since by the properties of the Drazin inverse in Definition 1.16 we have

$$A^k v = A^k (I - AA^D) w = A^{k+1} A^D (I - AA^D) w = A^k (AA^D - AA^D) w = 0.$$

Thus, we obtain:

$$AA^D w = u = w - v \in \mathbb{R}^n_+ + \ker(A^k).$$

Furthermore,

$$A^D w = A^D A A^D w = A^k (A^D)^{k+1} w,$$

from which we conclude that $A^D w \in \mathrm{im}(A^k)$.
Step 2. Set $x := A^D w$. Since from $Ax \in \mathbb{R}^n_+ + \ker(A^k)$ and $x \in \mathrm{im}(A^k)$, it follows that $x \geq 0$. We obtain $x = A^D w \geq 0$. As $w \geq 0$ was arbitrarily chosen, this completes the proof. □

The Drazin inverse may be written in terms of canonical projectors, which we can use for an alternative sufficient condition for positivity of the Drazin inverse in Corollary 2.17.

Lemma 2.16 *Let $E \in \mathbb{R}^{n \times n}$ with $\mathrm{ind}(E) = \nu$. For the matrix pair (E, I), let a matrix chain as in (1.6) be constructed with canonical projectors $P_i, Q_i, i = 0, \dots, \nu - 1$ as in Lemma 2.7. Then, $E^D = E_\nu^{-1} P_0 \dots P_{\nu-1}$.*

Proof. Consider the matrix pair (E, I). We have $\mathrm{ind}(E, I) = \mathrm{ind}(E) = \nu$ and since E commutes with the identity matrix, we conclude from Corollary 2.10 that $E^D = E_\nu^{-1} P_0 \dots P_{\nu-1}$. □

Corollary 2.17 *Let $E \in \mathbb{R}^{n \times n}$ with $\mathrm{ind}(E) = \nu$. For the matrix pair (E, I), let a matrix chain as in (1.6) be constructed with canonical projectors $P_i,\ Q_i,\ i = 0, \dots, \nu - 1$ as in Lemma 2.7. Then, if E_ν is an M-matrix and $P_0 \dots P_{\nu-1} \geq 0$, we have $E^D \geq 0$.*

Proof. From Lemma 2.16, we have that $E^D = E_\nu^{-1} P_0 \dots P_{\nu-1}$. If E_ν is an M-matrix, then $E_\nu^{-1} \geq 0$ and since also $P_0 \dots P_{\nu-1} \geq 0$ we obtain $E^D \geq 0$. □

2.4 Nonnegativity of Schur complements of nonnegative projectors

A main issue in the analysis of descriptor systems is the choice of the right projector onto the deflating subspace that corresponds to the finite eigenvalues of the matrix pair (E, A) [89]. In the following Chapter 3 we will see that nonnegative projectors play an important role in the context of positive descriptor systems. Note that in the linear algebra literature, projectors are also referred to as idempotent matrices. Therefore, these terms are used here interchangeably.

Schur complements constitute a fundamental tool in applications [129], in particular such as algebraic multigrid methods [123] or model reduction [82]. However, one has to ensure that essential properties are preserved. The results of this section become important in Chapter 6 where we discuss positivity preserving model reduction for positive descriptor systems. There, the Schur complement of the nonnegative spectral projector will be required to be again nonnegative.

We now introduce the problem setting. Let $\alpha = \{\alpha_1, \ldots, \alpha_l\} \subsetneq \langle n \rangle$ be a nonempty set and define by $\alpha^c := \langle n \rangle \backslash \alpha$ the nonempty set that is complementary to α.

If $A[\alpha, \alpha]$ is invertible, then the Schur complement of A corresponding to α is given by

$$A(\alpha) := A[\alpha^c, \alpha^c] - A[\alpha^c, \alpha]A[\alpha, \alpha]^{-1}A[\alpha, \alpha^c]. \tag{2.16}$$

If $A[\alpha, \alpha]$ is not invertible, then the Moore-Penrose Schur complement of A corresponding to α is defined by

$$A_\dagger(\alpha) := A[\alpha^c, \alpha^c] - A[\alpha^c, \alpha]A[\alpha, \alpha]^\dagger A[\alpha, \alpha^c], \tag{2.17}$$

where $A[\alpha, \alpha]^\dagger$ is the Moore-Penrose inverse of $A[\alpha, \alpha]$, see Definition 1.12. Furthermore, assuming that $(I - A[\alpha, \alpha])$ is invertible, we consider a shifted Schur complement defined by

$$\tilde{A}_\dagger(\alpha) := A[\alpha^c, \alpha^c] + A[\alpha^c, \alpha](I - A[\alpha, \alpha])^{-1}A[\alpha, \alpha^c], \tag{2.18}$$

which becomes important in this work when considering descriptor systems in discrete-time, see Section 6.3. This construct is used, for instance, in model reduction of discrete-time systems in the singular perturbation balanced truncation method [82].

Properties of generalised Schur complements of projectors were discussed in [8]. We assume that A is a nonnegative projector and show that if $A[\alpha, \alpha]$ has a positive diagonal then $A_\dagger(\alpha)$ is a nonnegative projector. We provide an example for the case that $A[\alpha, \alpha]$ has a zero on its main diagonal, where $A_\dagger(\alpha)$ fails to be nonnegative. The results of this section were published in [42].

For our main result we need the following simplification of the canonical form for nonnegative projectors given in Theorem 1.6

Lemma 2.18 *Let $B \in \mathbb{R}_+^{n \times n}$ be a nonzero nonnegative projector of rank k. Then, there exists a permutation matrix Π such that*

$$P := \Pi B \Pi^T = \begin{bmatrix} J & JG & 0 \\ 0 & 0 & 0 \\ FJ & FJG & 0 \end{bmatrix}, \; J \in \mathbb{R}_+^{n_1 \times n_1}, G \in \mathbb{R}_+^{n_1 \times n_2}, F \in \mathbb{R}_+^{n_3 \times n_1}, \tag{2.19}$$

where $n = n_1 + n_2 + n_3, 1 \le n_1, 0 \le n_2, 0 \le n_3$, F, G are arbitrary nonnegative matrices, and J is a direct sum of $k \ge 1$ positive projectors $J_i \in \mathbb{R}_+^{l_i \times l_i}$ of rank 1, i.e.,

$$J = \oplus_{i=1}^k J_i, \; J_i = u_i v_i^T, 0 < u_i, v_i \in \mathbb{R}_+^{l_i}, \; v_i^T u_i = 1, \; i = 1, \ldots, k. \tag{2.20}$$

Proof. Theorem 1.8 states that B is permutationally similar to the following block matrix [39]

$$C := \begin{bmatrix} J & JG_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ F_1J & F_1JG_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Here, $J \in \mathbb{R}_+^{n_1 \times n_1}$ is of the form (2.20), $G_1 \in \mathbb{R}_+^{n_1 \times m_2}, F_1 \in \mathbb{R}_+^{n_3 \times n_1}$ are arbitrary nonnegative matrices, and the last m_4 rows and columns of C are zero. Hence, $n_1 + m_2 + n_3 + m_4 = n$ and $0 \leq m_2, n_3, m_4$. If $m_4 = 0$ then C is of the form (2.19). It remains to show that C is permutationally similar to P if $m_4 > 0$.

Interchanging the last row and column of C with the $(n_1 + m_2 + 1)$-st row and column of C we obtain a matrix C_1. Then, we interchange the $(n-1)$-st row and column of C_1 with the (n_1+m_2+2)-nd row and column of C_1. We continue this process until we obtain the idempotent matrix P with $n_2 = m_2+m_4$ zero rows located at the rows $n_1+1, \ldots, n_1+n_2$. It follows that P is of the form

$$P := \begin{bmatrix} J & G & 0 \\ 0 & 0 & 0 \\ F & H & 0 \end{bmatrix}, G \in \mathbb{R}_+^{n_1 \times n_2}, F \in \mathbb{R}_+^{n_3 \times n_1}, H \in \mathbb{R}_+^{n_3 \times n_3}.$$

Since $P^2 = P$ we have that

$$G = JG,\ F = FJ,\ H = FG = (FJ)(JG) = FJG.$$

Hence, P is of the form (2.19).

The consequence of Lemma 2.18 is that for nonnegative projectors, without loss of generality, we may assume the following block structure:

$$B = \begin{bmatrix} B_{11} & 0 & \ldots & 0 & B_{1k} & 0 \\ 0 & B_{22} & \ddots & \vdots & B_{2k} & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \vdots \\ \vdots & & \ddots & B_{k,k} & B_{k,k+1} & 0 \\ 0 & \ldots & \ldots & 0 & 0 & 0 \\ B_{k+2,1} & \ldots & \ldots & \ldots & B_{k+2,k+1} & 0 \end{bmatrix}, \tag{2.21}$$

where $B_{ij} \in \mathbb{R}_+^{m_i \times m_j}$ for all $i, j = 1, \ldots, k+2$.

2.4.1 Nonnegativity of the Moore-Penrose inverse Schur complement

In this subsection, in Theorem 2.19, assuming that $A[\alpha,\alpha]$ does not have zero diagonal entries, we show that the Schur complement constructed via the Moore-Penrose inverse as defined in (2.17) is again a nonnegative projector. Note that this includes the case when $A[\alpha,\alpha]$ is invertible. However, this result is false for the general case of the Moore-Penrose Schur complement. A counterexample is given in Example 2.25.

Theorem 2.19 *Let $A \in \mathbb{R}_+^{n\times n}$ be a nonnegative projector. We assume that for $\emptyset \neq \alpha \subsetneqq \langle n\rangle$, the submatrix $A[\alpha,\alpha]$ has a positive diagonal. Then $A_\dagger(\alpha)$ is a nonnegative projector. Furthermore,*

$$\operatorname{rank} A_\dagger(\alpha) = \operatorname{rank} A - \operatorname{rank} A[\alpha,\alpha]. \tag{2.22}$$

Proof. Without loss of generality we may assume that A is of the form (2.19). Since $A[\alpha,\alpha]$ has a positive diagonal, we have that $A[\alpha,\alpha]$ is a submatrix of J. First we consider the special case $A[\alpha,\alpha] = J$. Using the identity $JJ^\dagger J = J$, we obtain that $A_\dagger(\alpha) = 0$. Since $\operatorname{rank} A = \operatorname{rank} J$, also the equality in (2.22) holds.

Let J, F, G be defined as in (2.19) and assume now that $A[\alpha,\alpha]$ is a strict submatrix of J. In the following, for an integer j we write $j + \langle m\rangle$ for the index set $\{j+1,\ldots,j+m\}$. Let $\alpha' := \langle n_1\rangle\backslash\alpha$, $\beta := n_1 + \langle n_2\rangle$ and $\gamma := n_1 + n_2 + \langle n_3\rangle$. Then,

$$\begin{aligned} A[\alpha^c,\alpha]A[\alpha,\alpha]^\dagger A[\alpha,\alpha^c] &= \begin{bmatrix} J[\alpha',\alpha] \\ 0 \\ (FJ)[\gamma,\alpha] \end{bmatrix} J[\alpha,\alpha]^\dagger \begin{bmatrix} J[\alpha,\alpha'] & (JG)[\alpha,\beta] & 0 \end{bmatrix} \\ &= \begin{bmatrix} J[\alpha',\alpha]J[\alpha,\alpha]^\dagger J[\alpha,\alpha'] & J[\alpha',\alpha]J[\alpha,\alpha]^\dagger (JG)[\alpha,\beta] & 0 \\ 0 & 0 & 0 \\ (FJ)[\gamma,\alpha]J[\alpha,\alpha]^\dagger J[\alpha,\alpha'] & (FJ)[\gamma,\alpha]J[\alpha,\alpha]^\dagger (JG)[\alpha,\beta] & 0 \end{bmatrix}. \end{aligned}$$

On the other hand, we have

$$A[\alpha^c,\alpha^c] = \begin{bmatrix} J[\alpha',\alpha'] & (JG)[\alpha',\beta] & 0 \\ 0 & 0 & 0 \\ (FJ)[\gamma,\alpha'] & FJG & 0 \end{bmatrix}.$$

Thus, the nonnegativity of $A_\dagger(\alpha)$ is equivalent to the following entrywise inequalities

$$J[\alpha',\alpha'] \geq J[\alpha',\alpha]J[\alpha,\alpha]^\dagger J[\alpha,\alpha'], \tag{2.23}$$
$$(JG)[\alpha',\beta] \geq J[\alpha',\alpha]J[\alpha,\alpha]^\dagger (JG)[\alpha,\beta], \tag{2.24}$$
$$(FJ)[\gamma,\alpha'] \geq (FJ)[\gamma,\alpha]J[\alpha,\alpha]^\dagger J[\alpha,\alpha'], \tag{2.25}$$
$$FJG \geq (FJ)[\gamma,\alpha]J[\alpha,\alpha]^\dagger (JG)[\alpha,\beta]. \tag{2.26}$$

Without loss of generality, we may assume that J is permuted such that the indices of the first q blocks J_i are contained in α^c, the indices of the following blocks J_i for $i = q+1,\ldots,q+p$ are split between α and α^c and the indices of the blocks J_i for $i = q+p+1,\ldots,q+p+\ell = k$ are contained in α. Partitioning the vectors u_i and v_i in (2.20) according to α and α^c as

$$u_i^T = (a_i^T, x_i^T), \quad v_i^T = (b_i^T, y_i^T), \quad i = q+1,\ldots,q+p,$$

we obtain that

$$J[\alpha',\alpha'] = (\oplus_{i=1}^{q} J_i) \oplus_{i=q+1}^{q+p} a_i b_i^T, \; J[\alpha,\alpha] = (\oplus_{j=q+1}^{q+p} x_i y_i^T) \oplus_{i=q+p+1}^{q+p+\ell} J_i.$$

Note that

$$q = \operatorname{rank} J - \operatorname{rank} A[\alpha,\alpha] = \operatorname{rank} A - \operatorname{rank} A[\alpha,\alpha]. \tag{2.27}$$

We will only consider the case $q,p,\ell > 0$, as other cases follow similarly. We have

$$J[\alpha,\alpha]^\dagger = (\oplus_{i=q+1}^{q+p} \frac{1}{(x_i^T x_i)(y_i^T y_i)} y_i x_i^T) \oplus_{i=q+p+1}^{q+p+\ell} \frac{1}{(u_i^T u_i)(v_i^T v_i)} v_i u_i^T, \tag{2.28}$$

$$J[\alpha,\alpha'] = \begin{bmatrix} 0 & \oplus_{i=q+1}^{q+p} x_i b_i^T \\ 0 & 0 \end{bmatrix}, \; J[\alpha',\alpha] = \begin{bmatrix} 0 & 0 \\ \oplus_{i=q+1}^{q+p} a_i y_i^T & 0 \end{bmatrix}, \tag{2.29}$$

and hence,

$$J[\alpha',\alpha]J[\alpha,\alpha]^\dagger = \begin{bmatrix} 0 & 0 \\ \oplus_{i=q+1}^{q+p} \frac{1}{x_i^T x_i} a_i x_i^T & 0 \end{bmatrix}, \tag{2.30}$$

$$J[\alpha,\alpha]^\dagger J[\alpha,\alpha'] = \begin{bmatrix} 0 & \oplus_{i=q+1}^{q+p} \frac{1}{y_i^T y_i} y_i b_i^T \\ 0 & 0 \end{bmatrix}, \tag{2.31}$$

$$J[\alpha',\alpha]J[\alpha,\alpha]^\dagger J[\alpha,\alpha'] = \begin{bmatrix} 0 & 0 \\ 0 & \oplus_{i=q+1}^{q+p} a_i b_i^T \end{bmatrix}. \tag{2.32}$$

Therefore, we obtain

$$J[\alpha',\alpha'] - J[\alpha',\alpha]J[\alpha,\alpha]^\dagger J[\alpha,\alpha'] = \begin{bmatrix} \oplus_{i=1}^{q} J_i & 0 \\ 0 & 0 \end{bmatrix} \geq 0, \tag{2.33}$$

which proves (2.23).

We now show the inequalities (2.24) and (2.25). First, we observe that JG and FJ have the following block form

$$JG = \begin{bmatrix} u_1 g_1^T \\ \vdots \\ u_k g_k^T \end{bmatrix},\ FJ = \begin{bmatrix} f_1 v_1^T & \cdots & f_k v_k^T \end{bmatrix},\ g_i \in \mathbb{R}_+^{n_2}, f_i \in \mathbb{R}_+^{n_3} \text{ for } i = 1, \ldots, k. \tag{2.34}$$

Hence, we obtain

$$(JG)[\alpha, \beta] = \begin{bmatrix} x_{q+1} g_{q+1}^T \\ \vdots \\ x_{q+p} g_{q+p}^T \\ u_{q+p+1} g_{q+p+1}^T \\ \vdots \\ u_k g_k^T \end{bmatrix}, \tag{2.35}$$

$$(JG)[\alpha', \beta] = \begin{bmatrix} u_1 g_1^T \\ \vdots \\ u_q g_q^T \\ a_{q+1} g_{q+1}^T \\ \vdots \\ u_{q+p} g_{q+p}^T \end{bmatrix}, \tag{2.36}$$

$$(FJ)[\gamma, \alpha] = \begin{bmatrix} f_{q+1} y_{q+1}^T & \cdots & f_{q+p} y_{q+p}^T & f_{q+p+1} v_{q+p+1}^T & \cdots & f_k v_k^T \end{bmatrix}, \tag{2.37}$$

$$(FJ)[\gamma, \alpha'] = \begin{bmatrix} f_1 v_1^T & \cdots & f_q v_q^T & f_{q+1} b_{q+1}^T & \cdots & f_{q+p} b_{q+p}^T \end{bmatrix}. \tag{2.38}$$

Using (2.31), we obtain that

$$(FJ)[\gamma, \alpha] J[\alpha, \alpha]^\dagger J[\alpha, \alpha'] = \begin{bmatrix} 0 & \cdots & 0 & f_{q+1} b_{q+1}^T & \cdots & f_{q+p} b_{q+p}^T \end{bmatrix}.$$

Therefore, we have

$$(FJ)[\gamma, \alpha'] - (FJ)[\gamma, \alpha] J[\alpha, \alpha]^\dagger J[\alpha, \alpha'] = \begin{bmatrix} f_1 v_1^T & \cdots & f_q v_q^T & 0 & \cdots & 0 \end{bmatrix}. \tag{2.39}$$

Similarly, using (2.30), we obtain

$$(JG)[\alpha', \beta] - J[\alpha', \alpha]J[\alpha, \alpha]^{\dagger}(JG)[\alpha, \beta] = \begin{bmatrix} u_1 g_1^T \\ \vdots \\ u_q g_q^T \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2.40}$$

Hence, the inequalities (2.24) and (2.25) hold.

We now show the last inequality (2.26). To this end, we observe that

$$FJG = (FJ)(JG) = \sum_{i=1}^{k} f_i g_i^T. \tag{2.41}$$

Multiplying (2.28), (2.35) and (2.37) we obtain that

$$(FJ)[\gamma, \alpha]J[\alpha, \alpha]^{\dagger}(JG)[\alpha, \beta] = \sum_{i=q+1}^{k} f_i g_i^T.$$

Hence,

$$FJG - (FJ)[\gamma, \alpha]J[\alpha, \alpha]^{\dagger}(JG)[\alpha, \beta] = \sum_{i=1}^{q} f_i g_i^T \geq 0. \tag{2.42}$$

In particular, this proves that (2.26) holds.

It is left to show that $A_{\dagger}(\alpha)$ is a projector. If $q = 0$ then $A_{\dagger}(\alpha) = 0$ and, thus, $A_{\dagger}(\alpha)$ is a trivial projector and (2.27) yields (2.22).

Assuming finally that $q > 0$, it follows that $A_{\dagger}(\alpha)$ has the block form (2.19) with $J = \oplus_{i=1}^{q} J_i \oplus 0$. Hence, $A_{\dagger}(\alpha)$ is a projector whose rank is q and (2.27) yields (2.22).

Corollary 2.20 *Let $A \in \mathbb{R}_{+}^{n \times n}, A \neq 0$ be idempotent. If $\alpha \subsetneqq \langle n \rangle$ is chosen such that $A[\alpha, \alpha]$ is an invertible matrix, then $A[\alpha, \alpha]$ is diagonal.*

Proof. Note that the number ℓ in the proof of Theorem 2.19 is either zero or the corresponding blocks J_i are positive 1×1 matrices for $i = q+p+1, \ldots, q+p+\ell$. Furthermore, for the split blocks, we also have that $x_i y_i^T \in \mathbb{R}^{1 \times 1}$, for $i = q+1, \ldots, q+p$, since $x_i y_i^T$ is of rank 1. Therefore, $A[\alpha, \alpha]$ is diagonal. □

Corollary 2.21 *Let $A \in \mathbb{R}_{+}^{n \times n}, A \neq 0$ be idempotent. If $\alpha \subsetneqq \langle n \rangle$ is chosen such that $A[\alpha, \alpha]$ is a regular matrix, then the standard Schur complement (2.16) is nonnegative.*

Corollary 2.22 *Let $A \in \mathbb{R}_+^{n\times n}, A \neq 0$ be idempotent. Choose $\alpha \subsetneqq \langle n\rangle$, such that $I - A[\alpha,\alpha]$ is regular. Then, $\tilde{A}(\alpha)$ defined in (2.18) is a nonnegative idempotent matrix.*

To prove this Corollary 2.22 we need the following fact for idempotent matrices, which is probably known.

Lemma 2.23 *Let $A \in \mathbb{R}^{n\times n}, A \neq 0$ be idempotent given as a 2×2 block matrix $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$. Assume that $(I - A_{22}) \in \mathbb{R}^{n-m}$ is regular. Then, $B := A_{11} + A_{12}(I - A_{22})^{-1}A_{21}$ is idempotent.*

Proof. Let

$$E = (I - A_{22})^{-1}A_{21},\ D = A_{21} + A_{22}E,\ z = \begin{bmatrix} x \\ Ex \end{bmatrix} \in \mathbb{R}^n,\ x \in \mathbb{R}^m.$$

Note that $Az = \begin{bmatrix} Bx \\ Dx \end{bmatrix}$. As $A^2z = Az$ and x is an arbitrary vector, we obtain the equalities

$$A_{11}B + A_{12}D = B, \quad A_{21}B + A_{22}D = D. \tag{2.43}$$

From the second equality of (2.43) we obtain $D = EB$. Substituting this equality into the first equality of (2.43) we obtain that $B^2 = B$. □

Proof of Corollary 2.22. The assumption that $I - A[\alpha,\alpha]$ is regular implies that $A[\alpha,\alpha]$ does not have an eigenvalue 1, i.e., $\rho(A[\alpha,\alpha]) < 1$. Hence, $I - A[\alpha,\alpha]$ is an M-matrix [17] and $(I - A[\alpha,\alpha])^{-1} \geq 0$. The assertion of Corollary 2.22 now follows using Lemma 2.23. □

2.4.2 An example

In this subsection we assume that the nonnegative idempotent matrix A is of the special form

$$A := \begin{bmatrix} J & JG \\ 0 & 0 \end{bmatrix}. \tag{2.44}$$

Furthermore, we assume that $A[\alpha,\alpha]$ has a zero on its main diagonal. We give an example where $A_\dagger(\alpha)$ may fail to be nonnegative. To this end, we first start with the following known result.

Lemma 2.24 *Let $A \in \mathbb{R}^{n\times n}$ be a singular matrix of the following form*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0_{(n-p)\times p} & 0_{(n-p)\times(n-p)} \end{bmatrix},\ A_{11} \in \mathbb{R}^{p\times p}, A_{12} \in \mathbb{R}^{p\times(n-p)},\ \textit{for some } 1 \le p < n.$$

Then $(A^\dagger)^T$ has the same block form as A.

Proof. Let $r = \operatorname{rank} A$, where $r \le p$. Then, the reduced singular value decomposition of A is of the form $U_r\Sigma_r V_r^T$, where $U_r, V_r \in \mathbb{R}^{n\times r}, U_r^T U_r = V_r V_r^T = I_r$ and Σ_r is a diagonal matrix, whose diagonal entries are the positive singular values of A.
Hence, $AA^T = \begin{bmatrix} A_{11}A_{11}^T + A_{12}A_{12}^T & 0 \\ 0 & 0 \end{bmatrix}$ and all eigenvectors of AA^T corresponding to positive eigenvalues are of the form $(x^T, 0^T)^T, x \in \mathbb{R}^p$. Thus, $U_r^T = [U_{r1}^T\ 0_{r\times(n-p)}]$ where $U_{r1} \in \mathbb{R}^{p\times r}$. Since $A^\dagger = V_r\Sigma_r^{-1}U_r^T$, the above form of U_r establishes the lemma. □

In the following example we permute some rows and columns of A, in order to find the Schur complement of the right lower block.

Example 2.25 *Consider a nonnegative idempotent matrix in the block form*

$$B = \left[\begin{array}{ccc|cc} u_1v_1^T & 0 & u_1s_1^T & u_1t_1^T & 0 \\ 0 & a_2b_2^T & a_2s_2^T & a_1t_2^T & a_2y_2^T \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & x_2b_2^T & x_2s_2^T & x_1t_2^T & x_2y_2^T \end{array}\right].$$

Then,

$$B[\alpha,\alpha] = \begin{bmatrix} 0 & 0 \\ x_1t_2^T & x_2y_2^T \end{bmatrix},\ B[\alpha,\alpha]^\dagger = \begin{bmatrix} 0 & \frac{t_2x_2^T}{(x_2^Tx_2)(t_2^Tt_2+y_2^Ty_2)} \\ 0 & \frac{y_2x_2^T}{(x_2^Tx_2)(t_2^Tt_2+y_2^Ty_2)} \end{bmatrix},$$

and

$$B[\alpha^c,\alpha]B[\alpha,\alpha]^\dagger B[\alpha,\alpha^c] = \begin{bmatrix} 0 & \frac{t_1^Tt_2u_1b_2^T}{t_2^Tt_2+y_2^Ty_2} & \frac{t_1^Tt_2u_1s_2^T}{t_2^Tt_2+y_2^Ty_2} \\ 0 & a_2b_2^T & a_2s_2^T \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence $B_\dagger(\alpha)_{11} > 0,\ B_\dagger(\alpha)_{12} \le 0$ and the Moore-Penrose inverse Schur complement is neither nonnegative nor nonpositive if $t_1^Tt_2 > 0$.

□

Summary

In this section we have shown that for a nonnegative projector, i.e. idempotent matrix, we have that the Schur complement constructed via the Moore-Penrose inverse is again a nonnegative projector, if the diagonal of $A[\alpha, \alpha]$ is strictly positive. In particular the nonnegativity also holds for the standard Schur complement if the corresponding part of the matrix is invertible. Also for a shifted Schur complement defined in (2.18), the nonnegativity was proved.

In Chapter 6, Section 6.3 we will propose and discuss a positivity preserving model reduction technique that is based on singular perturbation balanced truncation. In the descriptor case, the choice of the right projector onto the deflating subspace that corresponds to the finite eigenvalues of the matrix pair (E, A) is important for the analysis. In Chapter 3, we assume the nonnegativity of this spectral projector in order to characterise positivity in the descriptor case. Therefore, starting with a nonnegative spectral projector, the reduced order projector should also be nonnegative. It turns out that the reduced order projectors can be constructed via the standard Schur complement defined in (2.16) in the continuous-time case and via the shifted Schur complement defined in (2.18) in the discrete-time case. The nonnegativity of these is essential to ensure the positivity of the reduced order model and this is where the results of this section will be deployed.

Chapter 3

Positive systems and their characterisation

The theory of positive systems is deep and elegant -
and yet pleasantly consistent with intuition.
- David G. Luenberger

Positive systems arise in many applications. An important branch is systems biology [67], [68], [104], where metabolic networks, gene regulatory networks or signalling pathways are models that constitute positive descriptor systems. The state variables represent nonnegative quantities and the dynamics of the system are constrained via mass conservation laws. There are compartmental systems [3], [21], [47], such as models of pollution in connected water reservoirs, epidemic models, heat exchangers but also models in pharmacokinetics [86], where for instance the rates of absorption, distribution, metabolism and excretion of a drug substance are nonnegative quantities. In population dynamics [4], [7], [117], for instance, when modelling a chemostat or predator-prey interactions, the models result in positive systems. Also advection-diffusion-reaction systems [61], used for modelling in atmospheric chemistry, pollutant transport chemistry or for chemo-taxis problems are positive systems. However, positive systems arise not only in applications in life sciences. One also has to mention various applications in economy [77], [103], where mathematical modelling is used for predicting prices and productions.

Some of the models mentioned above, in fact, appear in a linear form as in (1) in continuous-time or in (2) in discrete-time. These are, for instance, models of age-structured population or certain models of connected water reservoirs, see also Chapter 6, Section 6.4 for examples. However, most models in systems biology such as metabolic networks or signalling pathways are highly nonlinear, see, e.g., [67], [104], [126], and appear in a general form

$$F(t, x, \dot{x}, u) = 0,\ x(t_0) = x_0,$$
$$y = G(x, u).$$

Nonlinear positive systems were studied in [50], [51], [70], [71], [72], [99]. Linearisation along constant trajectories leads to the here considered linear time-invariant systems in (1) or (2). An algorithm that would preserve the positivity property of the nonlinear system, however, is not available up to now and, thus, poses an open problem, although the application of the classical procedure sometimes leads to the desired outcome [59]. A more realistic approximation may be obtained by linearising along nonconstant trajectories, which leads to linear time-varying systems [28]. Yet, the analysis of positive time-varying systems is beyond the scope of this thesis but constitutes an interesting and promising research topic as an extension of the present work.

In the literature, an extensive amount of research exists that deals with positive systems in specific applications. Even so, on the theoretical side, this topic only recently has become a popular research area. Previously and even now, when dealing with certain systems theoretical or control theoretical issues of positive systems, the positivity property was and is being neglected in order to scoop the extensive toolbox of the well-developed systems theory and control theory for unconstrained systems. At the same time, there are many examples that show that it is worth the effort to consider problem specific properties and develop structured algorithms that preserve these.

There are various different definitions of positivity in systems theory that can be found in the literature. The definition of internally positive systems as discussed in this work goes back to Luenberger who presented some first theoretical results in [83]. There are also external positivity that is due to [38] and weak positivity introduced in [64], which we briefly discuss in the following. The extension of positivity to general cone invariance was studied by many authors, see, e.g., [10], [17], [18], [121]. In [69] the positivity concept is extended to positive operators.

There are mainly two books devoted to positive linear systems [38], [64]. In [38] linear time-invariant single input single output standard positive systems are treated. Certain control theoretical issues such as stability, reachability, observability are discussed and many examples from applications are provided. These include compartmental systems, Markov chains, queueing systems, or in particular the Leslie model of age-structured population or the Leontief input-output model used in economy. In [64] linear 1D and 2D systems are treated. The author also introduces descriptor systems discussing different positivity concepts. These are weak positivity, in which all matrices are assumed to be nonnegative, except for the matrix A in the continuous-time case that is assumed to be a $-Z$-matrix, and also external and internal positivity. However, the characterisation of internal positivity is given only for a special case of index 1 systems.

In Section 3.1 we briefly review the positivity concepts introduced in the literature and in Section 3.2.1 we present a new extension of the definition and characterisation of positivity for continuous-time and discrete-time descriptor systems [124]. The definition is based on consistent initial values of the descriptor system. In Section 3.2.2, we analyse and specify the special structure of the system matrices induced by the characterisation of positivity given in Section 3.2.1. Furthermore, in Section 3.3 we present a reduction technique by means of the Schur complement that allows to reduce special index 1 systems to a standard positive system.

3.1 Standard positive systems

In this section we briefly discuss several concepts of positivity that are encountered in the literature. Then, we state the well-known characterisation of (internal) positivity that is generalised to descriptor systems in the next section.

Internal positivity seems to be the most natural definition from the mathematical point of view. It goes back to Luenberger [83] and reads as follows.

Definition 3.1 (Internal positivity) *The continuous-time system (1) with $E = I$ is called* internally positive *if for any input function $u \in C^0$ such that $u(t) \geq 0$ for all $t \geq 0$ and any initial condition $x_0 \geq 0$ we have $x(t) \geq 0$ and $y(t) \geq 0$ for all $t \geq 0$.*
The discrete-time system (2) with $E = I$ is called internally positive *if for any input sequence $u(t) \geq 0$ for $t \geq 0$ and any initial condition $x_0 \geq 0$ we have $x(t) \geq 0$ and $y(t) \geq 0$ for all $t \geq 0$.*

In [38] Farina and Rinaldi distinguish between internal and external positivity. External positivity is defined as follows.

Definition 3.2 (External positivity) *The continuous-time system (1) with $E = I$ is called* externally positive *if for any input function $u \in C^0$ such that $u(t) \geq 0$ for all $t \geq 0$ and $x_0 = 0$ we have $y(t) \geq 0$ for all $t \geq 0$.*
The discrete-time system (2) with $E = I$ is called externally positive *if for any input sequence $u(t) \geq 0$ for $t \geq 0$ and $x_0 = 0$ we have $y(t) \geq 0$ for all $t \geq 0$.*

Internal positivity implies external positivity but conversely this is not necessarily the case. Moreover, in [38] it is shown that there exist externally positive systems that cannot be made internally positive through any change of basis of the state space. External positivity can be characterised by merely a property of the impulse response or equivalently the transfer function of the system. Therefore, the differentiation between internal and external positivity becomes useful, for instance, in the realisation problem. There, the external positivity condition is used for characterising positively realisable transfer functions [38], i.e. transfer functions that can be realised as an internally positive system. A definition and characterisation of external positivity in the descriptor case are given in [64] and is analogous to the one in the standard case.

The following type of positivity is defined only for descriptor systems since for standard systems it is equivalent to internal positivity. It was introduced by Kaczorek in [64], motivated by applications such as electrical circuits composed of resistances, inductances and voltage sources or of resistances capacitances and voltage sources, where such conditions hold.

Definition 3.3 (Weak positivity) *The continuous-time system (1) is called* weakly positive *if $E, B, C, D \geq 0$ and A is a $-Z$-matrix.*
The discrete-time system (2) is called weakly positive *if $E, A, B, C, D \geq 0$.*

The following theorem states a well-known characterisation of (internally) positive systems in the standard case that is easy to check, see, e.g., [38, 64].

Theorem 3.4 *The continuous-time system (1) with $E = I$ is (internally) positive if and only if A is a $-Z$-matrix and $B, C, D \geq 0$. The discrete-time system (2) with $E = I$ is positive if and only if $A, B, C, D \geq 0$.*

In the following section we generalise Theorem 3.4 to the descriptor case.

3.2 Positive descriptor systems

In this section, we define (internal) positivity for descriptor systems and provide a characterisation in the continuous-time and in the discrete-time case. Furthermore, we discuss the special structure induced by the required conditions.

First, we give a definition of positivity in the continuous-time as well as in the discrete-time case.

Definition 3.5 (Positivity) *We call the continuous-time system (1) with* $\mathrm{ind}(E, A) = \nu$ positive *if for all* $t \in \mathbb{R}_+$ *we have* $x(t) \geq 0$ *and* $y(t) \geq 0$ *for any input function* $u \in C^\nu$ *such that* $u^{(i)}(\tau) \geq 0$ *for* $i = 0, \ldots, \nu - 1$ *and* $0 \leq \tau \leq t$ *and any consistent initial value* $x_0 \geq 0$*. The discrete-time system (2) with* $\mathrm{ind}(E, A) = \nu$ *is called* positive *if for all* $t \in \mathbb{N}$ *we have* $x(t) \geq 0$ *and* $y(t) \geq 0$ *for any input sequence* $u(\tau) \geq 0$ *for* $0 \leq \tau \leq t + \nu - 1$ *and any consistent initial value* $x_0 \geq 0$*.*

3.2.1 Characterisation of positivity

For characterising positive systems in the descriptor case, we consider systems (1) and (2) with $D = 0$. Note, that adding a $D \geq 0$ to a positive system will obviously not spoil positivity, since only nonnegative input functions are allowed. However, for $D \neq 0$ one does not obtain the same only if condition for the matrix C that we prove in this section.

To formulate a characterisation of positivity in the continuous-time case we need the following Lemma.

Lemma 3.6 *For a regular matrix pair* (E, A) *let* $\hat{E}, \hat{A}$ *be defined as in Lemma 1.19. If for all* $v \geq 0$ *we have* $e^{\hat{E}^D \hat{A} t} \hat{E}^D \hat{E} v \geq 0$ *for all* $t \geq 0$*, then there exists* $\alpha \geq 0$ *such that*

$$\hat{E}^D \hat{A} + \alpha \hat{E}^D \hat{E} \geq 0.$$

Proof. By assumption, we have that

$$e^{\hat{E}^D \hat{A} t} \hat{E}^D \hat{E} \geq 0 \quad \text{for all} \quad t \geq 0. \tag{3.1}$$

We now show that from this we obtain that $\hat{E}^D \hat{E} \geq 0$ and $[\hat{E}^D \hat{A}]_{ij} \geq 0$ for all pairs (i, j) such that $[\hat{E}^D \hat{E}]_{ij} = 0$. Suppose that there exists a pair of subscripts (i, j) such that $[\hat{E}^D \hat{E}]_{ij} < 0$ or $[\hat{E}^D \hat{E}]_{ij} = 0$ and $[\hat{E}^D \hat{A}]_{ij} < 0$, then for $t > 0$ small enough, we would obtain

$$[e^{\hat{E}^D \hat{A} t} \hat{E}^D \hat{E}]_{ij} = [\hat{E}^D \hat{E}]_{ij} + [\hat{E}^D \hat{A}]_{ij} t + O(t^2) < 0,$$

which contradicts equation (3.1). Here, we have used the property that $\hat{E}^D \hat{A} \hat{E}^D \hat{E} = \hat{E}^D \hat{A}$, which follows from the properties of the Drazin inverse in Definition 1.16 and from (1.13b). Since $\hat{E}^D \hat{E} \geq 0$, setting

$$\alpha \geq \left| \min_{(i,j):[\hat{E}^D \hat{E}]_{ij} \neq 0} \frac{[\hat{E}^D \hat{A}]_{ij}}{[\hat{E}^D \hat{E}]_{ij}} \right|,$$

we obtain $\hat{E}^D \hat{A} + \alpha \hat{E}^D \hat{E} \geq 0$. $\square$

Remark 3.7 *The important implication of Lemma 3.6 is that we can shift the finite spectrum of the matrix pair (E, A) as in the standard case, see, e.g., [38, p.38], so that the shifted matrix pair $(E, A + \alpha E)$ fulfils the assumptions of Theorem 2.8 and its finite spectral radius is an eigenvalue. For any finite eigenvalue μ of $(E, A + \alpha E)$ we have that $\lambda = \mu - \alpha$ is a finite eigenvalue of (E, A). The eigenvectors and eigenspaces of (E, A) and $(E, A + \alpha E)$ are the same. In particular, the eigenspace that corresponds to the eigenvalue ∞ remains unchanged. Note that we can choose α large enough such that $\rho_f(E, A + \alpha E) > 0$ and, therefore, we always have a corresponding nonnegative eigenvector in this case.*

In addition to the implication of Lemma 3.6 in Remark 3.7 note the following. The proof of Lemma 3.6 implies that if the assumption of Lemma 3.6 holds then we have $\hat{E}^D\hat{E} \geq 0$. Hence, if we require that the homogeneous system $E\dot{x} = Ax$ has a nonnegative solution for any initial value $x_0 \geq 0$ (instead of any consistent initial value $x_0 = \hat{E}^D\hat{E}v \geq 0$), then $\hat{E}^D\hat{E} \geq 0$ turns out to be a necessary condition. Therefore, it seems plausible to have $\hat{E}^D\hat{E} \geq 0$ as an assumption for a characterisation of positivity, which is the case in the following theorem. Moreover, from the point of view of applications it makes more sense to prescribe an initial condition that is just nonnegative instead of one that is nonnegative on some special subspace.

Now we state a characterisation of positivity in the continuous-time case.

Theorem 3.8 *Let E, A, B, C be the matrices in system (1) with (E, A) regular of $\operatorname{ind}(E, A) = \nu$. Let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19 and $\hat{B}$ as in (1.12). Furthermore, assume that*

(i) $(I - \hat{E}^D\hat{E})(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B} \leq 0$ *for* $i = 0, \ldots, \nu - 1$,

(ii) $\hat{E}^D\hat{E} \geq 0$.

Then, the continuous-time system (1) is positive if and only if the following conditions hold

1. *there exists a scalar $\alpha \geq 0$ such that the matrix*

$$\bar{M} := -\alpha I + (\hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E})$$

   is a $-Z$-matrix,

2. $\hat{E}^D\hat{B} \geq 0$,

3. *C is nonnegative on the subspace $\mathcal{X}$ defined by*

$$\mathcal{X} := \mathrm{im}_+[\hat{E}^D\hat{E}, -(I-\hat{E}^D\hat{E})\hat{A}^D\hat{B}, \ldots, -(I-\hat{E}^D\hat{E})(\hat{E}\hat{A}^D)^{\nu-1}\hat{A}^D\hat{B}], \tag{3.2}$$

where for a matrix $W \in \mathbb{R}^{n\times q}$ we define

$$\mathrm{im}_+ W := \{w_1 \in \mathbb{R}^n \mid \exists w_2 \in \mathbb{R}^q_+ : Ww_2 = w_1\}.$$

Proof. "$\Rightarrow$" Let the system in (1) be positive. By definition, for all $t \geq 0$ we have $x(t) \geq 0$ and $y(t) \geq 0$ for every vector function $u \in C^\nu$ that satisfies $u^{(i)}(\tau) \geq 0$ for $i = 0, \ldots, \nu-1$ and $0 \leq \tau \leq t$ and for every consistent $x_0 \geq 0$.

1. Choose $u \equiv 0$, then for any $v \geq 0$ we have that $x_0 = \hat{E}^D\hat{E}v \geq 0$ is a consistent initial condition. Hence, for all $v \geq 0$, from (1.14) we obtain that

$$x(t) = e^{\hat{E}^D\hat{A}t}\hat{E}^D\hat{E}v \geq 0, \quad \text{for all } t \geq 0. \tag{3.3}$$

Then, by Lemma 3.6, there exists a scalar $\alpha \geq 0$ such that $\hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E} \geq 0$. Hence, the matrix $\bar{M} = -\alpha I + (\hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E})$ has nonnegative off-diagonal entries, i.e. $\bar{M}$ is a $-Z$-matrix.

2. Choose now $u(\tau) = \xi\tau^\nu$ for some $\xi \in \mathbb{R}^m_+$. We have that $u^{(i)}(\tau) \geq 0$ for $i = 0, \ldots, \nu-1$ and $0 \leq \tau \leq t$. Furthermore, we have $u^{(i)}(0) = 0$ for $i = 0, \ldots, \nu-1$. Therefore, for some $v \in \ker \hat{E}^D\hat{E}$, we have that $x_0 = \hat{E}^D\hat{E}v = 0$ is a consistent initial condition. Thus, from (1.14) we obtain that for all $t \geq 0$ we have

$$x(t) = \int_0^t e^{\hat{E}^D\hat{A}(t-\tau)}\hat{E}^D\hat{B}u(\tau)d\tau - (I-\hat{E}^D\hat{E})\sum_{i=0}^{\nu-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}u^{(i)}(t) \geq 0. \tag{3.4}$$

Since $\hat{E}^D\hat{E} \geq 0$, we can premultiply the inequality (3.4) by $\hat{E}^D\hat{E}$ and obtain

$$\hat{E}^D\hat{E}x(t) = \int_0^t e^{\hat{E}^D\hat{A}(t-\tau)}\hat{E}^D\hat{B}\xi\tau^\nu d\tau \geq 0. \tag{3.5}$$

We now show that $\hat{E}^D\hat{B} \geq 0$. Suppose that this is not the case, i.e. there exist some indices i, j with $[\hat{E}^D\hat{B}]_{ij} < 0$. Then, for $\xi = e_j$, the j-th unit vector, and for $t > 0$ small enough, we would obtain

$$\begin{aligned} [\hat{E}^D\hat{E}x(t)]_i &= \int_0^t [(I + \hat{E}^D\hat{A}(t-\tau) + O((t-\tau)^2))\hat{E}^D\hat{B}u(\tau)]_i d\tau \\ &= \int_0^t ([\hat{E}^D\hat{B}]_{ij} + O(t-\tau))\tau^\nu d\tau < 0, \end{aligned}$$

which contradicts (3.5). Therefore, we conclude that $\hat{E}^D\hat{B} \geq 0$.

3. Note that by Assumptions (i) and (ii) the subspace $\mathcal{X}$ contains only nonnegative vectors. Let $v \in \operatorname{im}[\hat{E}^D\hat{E}],\ v \geq 0$. For $u \equiv 0$, we have that $x_0 = \hat{E}^D\hat{E}v \geq 0$ is consistent with u. Since the system is positive, we have

$$y(0) = Cx_0 = C\hat{E}^D\hat{E}v \geq 0. \tag{3.6}$$

Since $\hat{E}^D\hat{E}$ is a projector, we have $\hat{E}^D\hat{E}v = v$ and hence, by (3.6), C is nonnegative on $\operatorname{im}[\hat{E}^D\hat{E}]$.

Let now $w_0 \in \operatorname{im}_+[-(I-\hat{E}^D\hat{E})\hat{A}^D\hat{B}]$, then there exists $\xi_0 \geq 0$ such that

$$-(I-\hat{E}^D\hat{E})\hat{A}^D\hat{B}\xi_0 = w_0.$$

Choose $u_0(\tau) \equiv \xi_0$. Then, we have $u_0(0) = \xi_0$ and $u_0^{(i)}(0) = 0$ for $i = 1, \ldots, \nu-1$. The initial condition $x_0 = -(I-\hat{E}^D\hat{E})\hat{A}^D\hat{B}\xi_0$ is nonnegative by Assumption (i) and consistent with u_0 for some $v \in \ker \hat{E}^D\hat{E}$. Since the system is positive, we obtain

$$y(0) = Cx_0 = -C(I-\hat{E}^D\hat{E})\hat{A}^D\hat{B}\xi_0 = Cw_0 \geq 0. \tag{3.7}$$

We have shown that for all $w_0 \in \operatorname{im}_+[-(I-\hat{E}^D\hat{E})\hat{A}^D\hat{B}]$ we have $Cw_0 \geq 0$, i.e. C is nonnegative on $\operatorname{im}_+[-(I-\hat{E}^D\hat{E})\hat{A}^D\hat{B}]$.

Let $w_1 \in \operatorname{im}_+[-(I-\hat{E}^D\hat{E})(\hat{E}\hat{A}^D)\hat{A}^D\hat{B}]$, then there exists $\xi_1 \geq 0$ such that $-(I-\hat{E}^D\hat{E})(\hat{E}\hat{A}^D)\hat{A}^D\hat{B}\xi_1 = w_1$. Set $u_1(\tau) = \xi_1\tau$. Then, we have $u_1(0) = 0$, $u_1'(0) = \xi_1$ and $u_1^{(i)}(0) = 0,\ i = 2, \ldots, \nu-1$. The initial condition $x_0 = -(I-\hat{E}^D\hat{E})(\hat{E}\hat{A}^D)\hat{A}^D\hat{B}\xi_1$ is nonnegative by Assumption (i) and consistent with u_1. Since the system is positive, we have

$$y(0) = Cx_0 = -C(I-\hat{E}^D\hat{E})(\hat{E}\hat{A}^D)\hat{A}^D\hat{B}\xi_1 = Cw_1 \geq 0,$$

and hence, C is nonnegative on $\operatorname{im}_+[-(I-\hat{E}^D\hat{E})(\hat{E}\hat{A}^D)\hat{A}^D\hat{B}]$.

We now proceed in the same manner. By subsequently letting

$$w_i \in \operatorname{im}_+[-(I-\hat{E}^D\hat{E})(\hat{E}\hat{A}^D)^{(i)}\hat{A}^D\hat{B}],$$

for $i = 2, \ldots, \nu-1$, finding the corresponding nonnegative preimage ξ_i, setting $u_i(\tau) = \xi_i\tau^i$ and using the same argument as above we obtain that C is nonnegative on $\operatorname{im}_+[-(I-\hat{E}^D\hat{E})(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}]$ for $i = 2, \ldots, \nu-1$. In total, we have shown that C is nonnegative on $\mathcal{X}$ as in (3.2).

"$\Leftarrow$" Let (i), (ii) and 1.-3. hold. We have to show that the system in (1) is positive, i.e. for all $t \geq 0$ and for every vector function $u \in C^\nu$ such that $u^{(i)}(\tau) \geq 0$ for $i = 0, \ldots, \nu-1$

and $0 \leq \tau \leq t$ and for any consistent $x_0 \geq 0$, we get $x(t) \geq 0$ and $y(t) \geq 0$. The solution at time $t \geq 0$ is given by

$$x(t) = e^{\hat{E}^D\hat{A}t}\hat{E}^D\hat{E}x_0 + \int_0^t e^{\hat{E}^D\hat{A}(t-\tau)}\hat{E}^D\hat{B}u(\tau)d\tau - (I - \hat{E}^D\hat{E})\sum_{i=0}^{\nu-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}u^{(i)}(t), \quad (3.8)$$

and any consistent x_0 satisfies

$$x_0 = \hat{E}^D\hat{E}v - (I - \hat{E}^D\hat{E})\sum_{i=0}^{\nu-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}u^{(i)}(0),$$

for some $v \in \mathbb{R}^n$. We now subsequently show that the three summands in (3.8) are nonnegative.

1) Since $\hat{E}^D\hat{E} \geq 0$, for any consistent $x_0 \geq 0$ we get that $\hat{E}^D\hat{E}x_0 \geq 0$. Note, that for any $v \in S_f^{def}$ we have $\hat{E}^D\hat{E}v = v$ and

$$\bar{M}v = (-\alpha I + \hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E})v = \hat{E}^D\hat{A}v. \quad (3.9)$$

Since $\hat{E}^D\hat{E}$ is a projector onto S_f^{def}, we also have

$$e^{\hat{E}^D\hat{A}t}\hat{E}^D\hat{E} = e^{\bar{M}t}\hat{E}^D\hat{E}. \quad (3.10)$$

and $e^{\bar{M}t} \geq 0$, since $\bar{M}$ is a $-Z$-matrix. Hence, the first term of (3.8) is nonnegative.

2) For the second term we have that $\hat{E}^D\hat{B} \geq 0$ and therefore

$$e^{\hat{E}^D\hat{A}(t-\tau)}\hat{E}^D\hat{B}u(\tau) \geq 0,$$

for all $0 \leq \tau \leq t$. Since integration is monotone, the second term is nonnegative.

3) We have $-(I - \hat{E}^D\hat{E})(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B} \geq 0$ for $i = 0, \ldots, \nu - 1$ and therefore the third term is also nonnegative for any vector function $u \in C^\nu$ such that $u^{(i)}(t) \geq 0$ for $i = 0, \ldots, \nu-1$ and $0 \leq \tau \leq t$.

Thus, $x(t) \geq 0$. From $y(t) = Cx(t)$ with C nonnegative on $\mathcal{X}$ and $x(t) \in \mathcal{X}$ for all t, we also conclude that $y(t) \geq 0$. □

Corollary 3.9 *Let E, A, B, C be the matrices in the system in (1) with (E, A) regular of $\operatorname{ind}(E, A) = \nu$. Let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19 and $\hat{B}$ as in (1.12). Furthermore, we assume that $(I - \hat{E}^D\hat{E})(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B} \leq 0$ for $i = 0, \ldots, \nu - 1$. If the matrix $\hat{E}^D\hat{A}$ is a $-Z$-matrix and $\hat{E}^D\hat{B}, C \geq 0$, then the continuous-time system in (1) is positive.*

Proof. If $\hat{E}^D\hat{A}$ is a $-Z$-matrix, this implies that $\bar{M}$ is a $-Z$-matrix for $\alpha = 0$. Internal positivity follows from Theorem 3.8. $\square$

The first of the following two examples demonstrates that the property that $\hat{E}^D\hat{A}$ is a $-Z$-matrix is not necessary for the system in (1) to be positive. The second example is a system that is not positive.

Example 3.10 *Consider the system*

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \dot{x} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u.$$

Since the matrices E and A commute, we can directly compute

$$E^D A = \begin{bmatrix} -1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E^D E = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E^D B = 0.$$

Note that $E^D A$ is not a $-Z$-matrix. For the state vector, we obtain

$$\begin{aligned} x(t) &= e^{E^D At} E^D E v - (I - E^D E) A^D B u(t) = \\ &= \begin{bmatrix} e^{-t} & e^{-t} - 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 + v_2 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} u(t). \end{aligned}$$

Hence, the system is positive, although $E^D A$ is not a $-Z$-matrix.

Example 3.11 *Consider the system*

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \dot{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u.$$

The matrices E and A commute and we can compute

$$E^D A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E^D E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E^D B = 0.$$

For the solution, we obtain

$$x(t) = e^{E^D At} E^D E v - (I - E^D E) A^D B u(t) = \begin{bmatrix} e^t & -te^t & 0 \\ 0 & e^t & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} u(t).$$

The system is not positive, since the first component of x may become negative.

In [24], the following characterisation of positivity in the case of discrete-time systems was given. Note, that in [24] the proof is given without the consistency requirement on x_0, thus, referring to a somewhat different solution concept. However, with a minor modification of the proof, the characterisation is also valid for positivity as in Definition 1.22, i.e., only for consistent initial values. Furthermore, we add the condition on the matrix C for completeness.

Theorem 3.12 *Let E, A, B, C be the system matrices in (2) with (E, A) regular of $\operatorname{ind}(E, A) = \nu$. Let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19 and $\hat{B}$ as in (1.12). If $\hat{E}^D\hat{E} \geq 0$, then the discrete-time system in (2) is positive if and only if $\hat{E}^D\hat{A}, \hat{E}^D\hat{B} \geq 0$, $(I - \hat{E}^D\hat{E})(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B} \leq 0$ for $i = 0, \ldots, \nu - 1$ and C is nonnegative on $\mathcal{X}$ as defined in (3.2).*

3.2.2 A special structure induced by the characterisations in Theorem 3.8 and Theorem 3.12

The conditions of Theorem 3.8 or Theorem 3.12 impose a very special structure on the system matrices that will be important in the following. We analyse this structure for the continuous-time case. However, the same results hold for discrete-time positive systems with properties as in Theorem 3.12.

Consider the initial continuous-time system in (1a). Since we consider regular matrix pairs (E, A), we have that the matrix $R = EP_r + A(I - P_r)$ is regular and, hence, we may scale the system in (1a) by R^{-1}. Considering (E, A) in Weierstraß canonical form as in (1.4), we have that $R^{-1} = T^{-1}W^{-1}$. We obtain the scaled system

$$R^{-1}E\dot{x} = R^{-1}Ax + R^{-1}Bu, \tag{3.11}$$

that in Weierstraß canonical form is given by

$$T^{-1}\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T\dot{x} = T^{-1}\begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} Tx + T^{-1}\begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix} u.$$

Note that the matrices $R^{-1}E$ and $R^{-1}A$ commute. System (3.11) is equivalent to the system of two equations

$$\begin{cases} \quad P_rR^{-1}E\dot{x} &= P_rR^{-1}Ax + P_rR^{-1}Bu, \\ (I - P_r)R^{-1}E\dot{x} &= (I - P_r)R^{-1}Ax + (I - P_r)R^{-1}Bu. \end{cases},$$

that in Weierstraß canonical form is given by

$$\begin{cases} T^{-1}\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T\dot{x} & = T^{-1}\begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} Tx + T^{-1}\begin{bmatrix} \tilde{B}_1 \\ \tilde{0} \end{bmatrix} u \\ T^{-1}\begin{bmatrix} 0 & 0 \\ 0 & N \end{bmatrix} T\dot{x} & = T^{-1}\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} Tx + T^{-1}\begin{bmatrix} 0 \\ \tilde{B}_2 \end{bmatrix} u \end{cases},$$

which by using the results in Section 1.6 is equal to

$$\begin{cases} P_r\dot{x} & = \hat{E}^D\hat{A}x + \hat{E}^D\hat{B}u \\ (I-P_r)\hat{E}\hat{A}^D\dot{x} & = (I-P_r)x + (I-P_r)\hat{A}^D\hat{B}u \end{cases}. \tag{3.12}$$

We assume that system (3.12) is positive and fulfils the conditions of Theorem 3.8. Note that symmetric permutations of the matrices do not change the matrix properties in Theorem 3.8. Therefore, without loss of generality, we may assume that P_r is in canonical form as in (2.21), i.e.,

$$P_r = \begin{bmatrix} \pi_{11} & 0 & \dots & 0 & \pi_{1,k+1} & 0 \\ 0 & \pi_{22} & \ddots & \vdots & \pi_{2,k+1} & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots & \vdots \\ \vdots & & \ddots & \pi_{kk} & \pi_{k,k+1} & 0 \\ 0 & \dots & \dots & 0 & 0 & 0 \\ \pi_{k+2,1} & \dots & \dots & \pi_{k+2,k} & \pi_{k+2,k+1} & 0 \end{bmatrix}, \tag{3.13}$$

where $\pi_{ij} \in \mathbb{R}_+^{m_i \times m_j}$, $i,j = 1,\dots,k+2$, and $\pi_{11},\dots,\pi_{kk} > 0$ are irreducible with $\rho(\pi_{11}) = \dots = \rho(\pi_{kk}) = 1$. Note that the irreducible diagonal blocks are of rank 1. Partition the matrices $\hat{E}^D\hat{A}, \hat{E}^D\hat{B}, \hat{E}\hat{A}^D, \hat{A}^D B$ accordingly, i.e.

$$\hat{E}^D\hat{A} = \begin{bmatrix} [\hat{E}^D\hat{A}]_{11} & \dots & [\hat{E}^D\hat{A}]_{1,k+2} \\ \vdots & & \vdots \\ [\hat{E}^D\hat{A}]_{k+2,1} & \dots & [\hat{E}^D\hat{A}]_{k+2,k+2} \end{bmatrix}, \quad \hat{E}^D\hat{B} = \begin{bmatrix} [\hat{E}^D\hat{B}]_1 \\ \vdots \\ [\hat{E}^D\hat{B}]_{k+2} \end{bmatrix},$$

$$\hat{E}\hat{A}^D = \begin{bmatrix} [\hat{E}\hat{A}^D]_{11} & \dots & [\hat{E}\hat{A}^D]_{1,k+2} \\ \vdots & & \vdots \\ [\hat{E}\hat{A}^D]_{k+2,1} & \dots & [\hat{E}\hat{A}^D]_{k+2,k+2} \end{bmatrix}, \quad \hat{A}^D\hat{B} = \begin{bmatrix} [\hat{A}^D\hat{B}]_1 \\ \vdots \\ [\hat{A}^D\hat{B}]_{k+2} \end{bmatrix}.$$

Firstly, from condition $P_r\hat{E}^D\hat{A} = \hat{E}^D\hat{A}$, we have that $[\hat{E}^D\hat{A}]_{k+1,i} = 0$ for $i = 1,\dots,k+2$

and from $\hat{E}^D\hat{A}P_r = \hat{E}^D\hat{A}$, we have $[\hat{E}^D\hat{A}]_{i,k+2} = 0$ for $i = 1,\ldots,k+2$. Therefore,

$$\hat{E}^D\hat{A} = \begin{bmatrix} [\hat{E}^D\hat{A}]_{11} & \ldots & [\hat{E}^D\hat{A}]_{1,k+1} & 0 \\ \vdots & & \vdots & \vdots \\ [\hat{E}^D\hat{A}]_{k,1} & \ldots & [\hat{E}^D\hat{A}]_{k,k+1} & 0 \\ 0 & \ldots & 0 & 0 \\ [\hat{E}^D\hat{A}]_{k+2,1} & \ldots & [\hat{E}^D\hat{A}]_{k+2,k+1} & 0 \end{bmatrix}. \tag{3.14}$$

Since by Lemma 3.6 there exists $\alpha \geq 0$ such that $\hat{E}^D\hat{A} + \alpha P_r \geq 0$, additionally, we have that $[\hat{E}^D\hat{A}]_{ij} \geq 0$ for $i,j = 1,\ldots,k$ and $i \neq j$.

Secondly, since we have $P_r\hat{E}^D\hat{B} = \hat{E}^D\hat{B}$, we obtain that

$$\hat{E}^D\hat{B} = \begin{bmatrix} [\hat{E}^D\hat{B}]_1 \\ \vdots \\ [\hat{E}^D\hat{B}]_k \\ 0 \\ * \end{bmatrix}, \tag{3.15}$$

where by $*$ we denote a block entry that we do not need to specify for our purpose.

Furthermore, from Theorem 3.8, we have that $(I - P_r)\hat{A}^D\hat{B} \leq 0$ holds for $\operatorname{ind}(E,A) \geq 1$, i.e.,

$$(I-P_r)\hat{A}^D\hat{B} = \begin{bmatrix} I-\pi_{11} & 0 & \ldots & 0 & -\pi_{1,k+1} & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \vdots \\ \vdots & & \ddots & I-\pi_{kk} & -\pi_{k,k+1} & 0 \\ 0 & \ldots & \ldots & 0 & I & 0 \\ -\pi_{k+2,1} & \ldots & \ldots & -\pi_{k+2,k} & -\pi_{k+2,k+1} & I \end{bmatrix} \begin{bmatrix} [\hat{A}^D\hat{B}]_1 \\ \vdots \\ \vdots \\ [\hat{A}^D\hat{B}]_k \\ [\hat{A}^D\hat{B}]_{k+1} \\ [\hat{A}^D\hat{B}]_{k+2} \end{bmatrix} \leq 0. \tag{3.16}$$

Hence, from the second last row of (3.16), we have $[\hat{A}^D\hat{B}]_{k+1} \leq 0$. Furthermore, we have

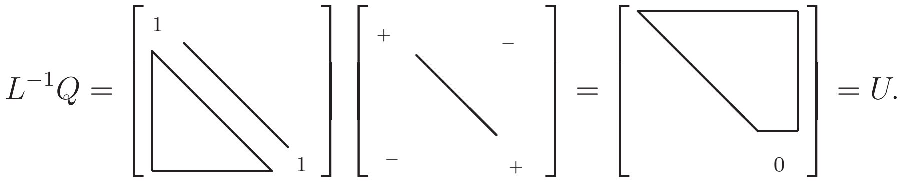
$$(I-\pi_{ii})[\hat{A}^D\hat{B}]_i - \pi_{i,k+1}[\hat{A}^D\hat{B}]_{k+1} \leq 0, \tag{3.17}$$

for $i = 1,\ldots,k$. We have that $\pi_{i,k+1} \geq 0$ and $[\hat{A}^D\hat{B}]_{k+1} \leq 0$ and therefore, $-\pi_{i,k+1}[\hat{A}^D\hat{B}]_{k+1} \geq 0$. Thus, $(I-\pi_{ii})[\hat{A}^D\hat{B}]_i$ must be less than or equal to 0 for $i = 1,\ldots,k$. Now we show that this implies $(I-\pi_{ii})[\hat{A}^D\hat{B}]_i = 0$ for $i = 1,\ldots,k$.

To this end, we use the property that π_{ii} is a positive (irreducible) projector of rank 1 with $\rho(\pi_{ii}) = 1$, see Theorem 1.6. Hence, $I - \pi_{ii}$ is a singular M-matrix. In [123], for a singular M-matrix $I - T$, where T is stochastic, it was shown that for the LU-decomposition of $I - T$ as in (1.9), we have that the last row of L^{-1} is the vector of all ones. For a more general discussion on LU-factorisations of singular M-matrices, see [94] and the references therein. In our case, we show that the last row of L^{-1} has only positive entries.

Lemma 3.13 *Let $P = uv^T$, $0 < u, v \in \mathbb{R}_+^{m_i}$ be a projector. Then, $Q := I - P$ is a singular M-matrix that has an LU-decomposition as in (1.9), where the last row of L^{-1} has only positive entries.*

Proof. We have $Q = LU$, where L is a regular and U is a singular M-matrix, or equivalently



Furthermore, we have that $L^{-1} \geq 0$ and the last row of U is zero. Since $P > 0$, we have $Q_{m_i j} < 0$ for $j = 1, \ldots, m_i - 1$. Since $[L^{-1}]_{m_i m_i} = 1$ we have, for instance, $[L^{-1}]_{m_i m_i} Q_{m_i 1} < 0$ and, therefore, we must have $[L^{-1}]_{m_i,1} > 0$, since otherwise $e_{m_i}^T L^{-1} Q = 0$, where e_{m_i} denotes the e_{m_i}-th unit vector, will not hold. Analogously, we obtain $[L^{-1}]_{m_i j} > 0$ for $j = 1, \ldots, m_i - 1$. $\square$

We now show that $(I - \pi_{ii})[\hat{A}^D \hat{B}]_i \leq 0$ implies $(I - \pi_{ii})[\hat{A}^D \hat{B}]_i = 0$ for $i = 1, \ldots, k$. Since $L^{-1} \geq 0$, we obtain that

$$(I - \pi_{ii})[\hat{A}^D \hat{B}]_i = LU[\hat{A}^D \hat{B}]_i \leq 0,$$

for $i = 1, \ldots, k$, is equivalent to

$$L^{-1}(I - \pi_{ii})[\hat{A}^D \hat{B}]_i = U[\hat{A}^D \hat{B}]_i \leq 0.$$

Furthermore, we have

$$e_{m_i}^T L^{-1}(I - \pi_{ii})[\hat{A}^D \hat{B}]_i = 0.$$

By Lemma 3.13, we have $e_{m_i}^T L^{-1} > 0$ and, since $(I - \pi_{ii})[\hat{A}^D \hat{B}]_i \leq 0$, we conclude that $(I - \pi_{ii})[\hat{A}^D \hat{B}]_i = 0$.

With this, from (3.17), we obtain that $\pi_{i,k+1}[\hat{A}^D \hat{B}]_{k+1} = 0$ for $i = 1, \ldots, k$. From the last row of (3.16), we now obtain

$$-\pi_{k+2,k+1}[\hat{A}^D \hat{B}]_{k+1} + [\hat{A}^D \hat{B}]_{k+2} \leq 0,$$

and since $-\pi_{k+2,k+1}[\hat{A}^D\hat{B}]_{k+1} \geq 0$, we have $[\hat{A}^D\hat{B}]_{k+2} \leq 0$.

Thus, we have

$$(I - P_r)\hat{A}^D\hat{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ [\hat{A}^D\hat{B}]_{k+1} \\ * \end{bmatrix} \leq 0. \tag{3.18}$$

In general, we obtain the following structure.

Lemma 3.14 *Consider the system matrices in (3.12) and let ν be the index of nilpotency of $(I - P_r)(\hat{E}\hat{A}^D)$. Then, assuming the block structure induced by P_r in (3.13), for $i = 0, \dots, \nu - 1$, we have*

$$(I - P_r)(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ *_1 \\ *_2 \end{bmatrix} \leq 0, \tag{3.19}$$

where $*_1, *_2$ denote some unspecified entries.

Proof. Note first that for $i \geq \nu$, we have $(I - P_r)(\hat{E}\hat{A}^D)^i = 0$. We perform an induction over the index i. Let $i = 0$, then by Equation (3.18) we have the desired form. Suppose that for some $i > 0$ we have the structure in (3.19). Then, for $i + 1$ we obtain

$$\begin{aligned}
(I - P_r)(\hat{E}\hat{A}^D)^{i+1}\hat{A}^D\hat{B} =& (I - P_r)\hat{E}\hat{A}^D\left((I - P_r)(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}\right) = (I - P_r)\hat{E}\hat{A}^D \begin{bmatrix} 0 \\ \vdots \\ 0 \\ *_1 \\ *_2 \end{bmatrix} \\
=& \hat{E}\hat{A}^D\left((I - P_r)(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}\right) \\
=& \begin{bmatrix} * & \dots & * & [\hat{E}\hat{A}^D]_{1,k+1} & [\hat{E}\hat{A}^D]_{1,k+2} \\ \vdots & & \vdots & \vdots & \vdots \\ * & \dots & * & [\hat{E}\hat{A}^D]_{k,k+1} & [\hat{E}\hat{A}^D]_{k,k+2} \\ * & \dots & * & [\hat{E}\hat{A}^D]_{k+1,k+1} & [\hat{E}\hat{A}^D]_{k+1,k+2} \\ * & \dots & * & * & * \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ *_1 \\ *_2 \end{bmatrix} \leq 0,
\end{aligned} \tag{3.20}$$

where from the third equality of (3.20) we have that

$$\begin{aligned}
[\hat{E}\hat{A}^D]_{j,k+1} &= (I-\pi_{jj})[\hat{E}\hat{A}^D]_{j,k+1} - \pi_{j,k+1}[\hat{E}\hat{A}^D]_{k+1,k+1} \\
[\hat{E}\hat{A}^D]_{j,k+2} &= (I-\pi_{jj})[\hat{E}\hat{A}^D]_{j,k+2} - \pi_{j,k+1}[\hat{E}\hat{A}^D]_{k+1,k+2}
\end{aligned} \tag{3.21}$$

for $j = 1, \ldots, k$. For the second last entry on the left hand side of (3.20) we obtain

$$[(I-P_r)(\hat{E}\hat{A}^D)^{i+1}\hat{A}^D\hat{B}]_{k+1} = [\hat{E}\hat{A}^D]_{k+1,k+1}\cdot(*_1) + [\hat{E}\hat{A}^D]_{k+1,k+2}\cdot(*_2) \leq 0. \tag{3.22}$$

Furthermore, for the entries $j = 1, \ldots, k$, using (3.21), we obtain

$$\begin{aligned}
[\hat{E}\hat{A}^D]_{j,k+1}\cdot(*_1) + [\hat{E}\hat{A}^D]_{j,k+2}(*_2) &= \Big((I-\pi_{jj})[\hat{E}\hat{A}^D]_{j,k+1} - \pi_{j,k+1}[\hat{E}\hat{A}^D]_{k+1,k+1}\Big)\cdot(*_1)+ \\
&\quad + \Big((I-\pi_{jj})[\hat{E}\hat{A}^D]_{j,k+2} - \pi_{j,k+1}[\hat{E}\hat{A}^D]_{k+1,k+2}\Big)\cdot(*_2) \\
&=(I-\pi_{jj})([\hat{E}\hat{A}^D]_{j,k+1}\cdot(*_1) + [\hat{E}\hat{A}^D]_{j,k+2}(*_2))- \\
&\quad - \pi_{j,k+1}[(I-P_r)(\hat{E}\hat{A}^D)^{i+1}\hat{A}^D\hat{B}]_{k+1} \leq 0.
\end{aligned}$$

Since $\pi_{j,k+1} \geq 0$ for $j = 1, \ldots, k$ and $[(I-P_r)(\hat{E}\hat{A}^D)^{i+1}\hat{A}^D\hat{B}]_{k+1} \leq 0$ by (3.20), we have that $(I-\pi_{jj})([\hat{E}\hat{A}^D]_{j,k+1}\cdot(*_1) + [\hat{E}\hat{A}^D]_{j,k+2}(*_2)) \leq 0$. Using Lemma 3.13 and the same argument as before, we obtain

$$(I-\pi_{jj})([\hat{E}\hat{A}^D]_{j,k+1}\cdot(*_1) + [\hat{E}\hat{A}^D]_{j,k+2}(*_2)) = 0,$$

for $j = 1, \ldots, k$ and, hence, $-\pi_{j,k+1}\cdot[(I-P_r)(\hat{E}\hat{A}^D)^{i+1}\hat{A}^D\hat{B}]_{k+1} = 0$ for $j = 1, \ldots, k$. Thus, we have obtained the desired structure, which completes the proof. ∎

By using the property in (3.19), and also the properties in (3.14) and in (3.15), we deduce that the system in (3.11) is equivalent to

$$\tilde{E}\dot{x} = \tilde{A}x + \tilde{B}u, \tag{3.23}$$

where

$$\tilde{A} = \begin{bmatrix}
\tilde{A}_{11} & \tilde{A}_{12} & \ldots & \tilde{A}_{1,k} & \tilde{A}_{1,k+1} & 0 \\
\tilde{A}_{21} & \tilde{A}_{22} & \ddots & \vdots & \tilde{A}_{2,k+1} & \vdots \\
\vdots & \ddots & \ddots & \tilde{A}_{k-1,k} & \vdots & \vdots \\
\tilde{A}_{k1} & \ldots & \tilde{A}_{k,k-1} & \tilde{A}_{kk} & \tilde{A}_{k,k+1} & \vdots \\
0 & \ldots & \ldots & 0 & I & 0 \\
\tilde{A}_{k+2,1} & \ldots & \ldots & \tilde{A}_{k+2,k} & \tilde{A}_{k+2,k+1} & I
\end{bmatrix}, \quad \tilde{B} = \begin{bmatrix}
[\hat{E}^D\hat{B}]_1 \\ \vdots \\ \vdots \\ [\hat{E}^D\hat{B}]_k \\ [\hat{A}^D\hat{B}]_{k+1} \\ *
\end{bmatrix}$$

$$\tilde{E} = P_r + (I-P_r)\hat{E}\hat{A}^D,$$

with $\tilde{A}_{ii} = [\hat{E}^D\hat{A}]_{ii} + (I-\pi_{ii})$, for $i = 1, \ldots, k$, $\tilde{A}_{ij} = [\hat{E}^D\hat{A}]_{ij}$ for $i, j = 1, \ldots, k$ and $i \neq j$ and $\tilde{A}_{ij} = [\hat{E}^D\hat{A}]_{ij} - \pi_{ij}$ for $i > k$ or $j > k$.

Note that the form in (3.23) displays several spectral properties of the system. The last two block rows correspond to the infinite eigenvalues that are responsible for the possibly higher index of the system. The blocks π_{ii}, for $i = 1, \ldots, k$ are positive irreducible projectors, i.e., the associated block rows correspond to one finite eigenvalue and $m_i - 1$ infinite eigenvalues, where m_i is the size of the block π_{ii}. In particular, this means that the system has k finite eigenvalues. Since any projector is similar to a matrix $\left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right]$, the corresponding infinite eigenvalues can only have Jordan chains of a maximum length 1, which means that they do not contribute to a possibly higher index of the system.

3.3 Special case: index 1 systems

The aim of this section is twofold. On the one hand the index 1 case exemplarily verifies the general results established in the previous Section 3.2. On the other hand we present a reduction technique by means of the Schur complement that allows to reduce certain positive index 1 system to the positive standard case.

Continuous-time

Consider a system of the form (1), where $E = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix}$ with $E_{11} \in \mathbb{R}^{r \times r}$ regular and A, B are partitioned accordingly:

$$\begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u. \tag{3.24}$$

Assume that A_{22} is invertible, then we can reduce the descriptor system to a standard system by the following procedure. We premultiply the system (3.24) by the matrix $\begin{bmatrix} E_{11}^{-1} & -E_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & I_{n-r} \end{bmatrix}$ and obtain:

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} E_{11}^{-1} A_S & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} E_{11}^{-1} B_S \\ B_2 \end{bmatrix} u,$$

where $A_S := A_{11} - A_{12} A_{22}^{-1} A_{21}$ and $B_S := B_1 - A_{12} A_{22}^{-1} B_2 \geq 0$. For the solution of the transformed system we obtain

$$\begin{aligned} x_1(t) &= e^{E_{11}^{-1} A_S t} x_{10} + \int_0^t e^{E_{11}^{-1} A_S (t-\tau)} E_{11}^{-1} B_S u(\tau) d\tau \\ x_2(t) &= -A_{22}^{-1} (A_{21} x_1(t) + B_2 u(t)). \end{aligned}$$

If we assume that $E_{11}^{-1}A_S$ is a $-Z$-matrix, $E_{11}^{-1}B_S \geq 0$ and $A_{22}^{-1}B_2 \leq 0$, $A_{22}^{-1}A_{21} \leq 0$, then for any input function $u \geq 0$ we obtain $x(t) \geq 0$ for all $t \geq 0$, i.e., the system is internally positive. Note that x_{20} does not appear in the solution. This is due to the fact that $\hat{E}^D\hat{E}$ projects onto the first component.

For regular matrix pairs (E, A) with $\operatorname{ind}(E, A) = 1$ we have the property that the corresponding system can always be equivalently transformed into a system of the form (3.24), see, e.g. [74].

Lemma 3.15 *Consider a system of the form (1), where (E, A) is a regular matrix pair with $\operatorname{ind}(E, A) = 1$. Then there exist regular matrices P, Q, R such that with $\tilde{E} = PEQ$, $\tilde{A} = PAQ$ and $\tilde{B} = PBR$ we obtain a system of the form (3.24).*

Note that the condition $\operatorname{ind}(E, A) = 1$ is equivalent to E_{11} and A_{22} being regular in the form (3.24). We thus have proved the following theorem that states that every index 1 system that can be equivalently transformed into a system of the form as in (3.24) can be reduced to a standard positive system by means of the Schur complement.

Theorem 3.16 *Consider a system of the form (1), where (E, A) is a regular matrix pair of $\operatorname{ind}(E, A) = 1$. If there exist regular matrices P, Q, R with $Q, R \geq 0$ such that with $\tilde{E} = PEQ$, $\tilde{A} = PAQ$ and $\tilde{B} = PBR$ we obtain a system of the form (3.24) and, if we additionally assume that $E_{11}^{-1}A_S$ is a $-Z$-matrix, $E_{11}^{-1}B_S \geq 0$ and $A_{22}^{-1}B_2, A_{22}^{-1}A_{21} \leq 0$, then the system in (1) is internally positive and can be reduced to an internally positive standard system.*

Note that the transformation matrices P, Q, R in Lemma 3.15 are not necessarily nonnegative. In Theorem 3.16, we assume that only the matrices Q, R are nonnegative, since transformations of the equations from the left with P only scale the system but do not change the solution. The following Corollary 3.17, therefore, states the most realistic case, when such a Schur complement reduction in the index 1 case to an ODE case is possible.

Corollary 3.17 *Consider a system of the form (1), where (E, A) is a regular matrix pair of $\operatorname{ind}(E, A) = 1$. If there exists a regular matrix P such that with $\tilde{E} = PE$, $\tilde{A} = PA$ and $\tilde{B} = PB$ we obtain a system of the form (3.24) and, if we additionally assume that $E_{11}^{-1}A_S$ is a $-Z$-matrix, $E_{11}^{-1}B_S \geq 0$ and $A_{22}^{-1}B_2, A_{22}^{-1}A_{21} \leq 0$, then the system in (1) is internally positive and can be reduced to an internally positive standard system.*

It is now of interest to compare this to the results we obtain by using Theorem 3.8. Assuming $\mathrm{ind}(E, A) = 1$, for the system (3.24) we have

$$\begin{aligned} \hat{E}^D\hat{E} &= \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & 0 \end{bmatrix}, & \hat{E}^D\hat{A} &= \begin{bmatrix} E_{11}^{-1}A_S & 0 \\ -A_{22}^{-1}A_{21}E_{11}^{-1}A_S & 0 \end{bmatrix}, \\ \hat{E}^D\hat{B} &= \begin{bmatrix} E_{11}^{-1}B_S \\ -A_{22}^{-1}A_{21}E_{11}^{-1}B_S \end{bmatrix}, & (I - \hat{E}^D\hat{E})\hat{A}^D\hat{B} &= \begin{bmatrix} 0 \\ A_{22}^{-1}B_2 \end{bmatrix}. \end{aligned} \tag{3.25}$$

Theorem 3.8 restated in our context now means the following. If

i) $A_{22}^{-1}B_2 \leq 0$ and

ii) $-A_{22}^{-1}A_{21} \geq 0$,

then the system in (3.24) is positive if and only if the following conditions hold

1. there exists a scalar $\alpha \geq 0$ such that the matrix
$$\bar{M} := -\alpha I + (\hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E})$$
is a $-Z$-matrix,
2. $\hat{E}^D\hat{B} \geq 0$.

We have that
$$\bar{M} = \begin{bmatrix} E_{11}^{-1}A_S & 0 \\ -A_{22}^{-1}A_{21}(E_{11}^{-1}A_S + \alpha I) & -\alpha I \end{bmatrix}$$
is a $-Z$-matrix if and only if $E_{11}^{-1}A_S$ is a $-Z$-matrix. Note that in this case $\alpha \geq 0$ can be chosen such that $E_{11}^{-1}A_S + \alpha I \geq 0$. Furthermore, we have $\hat{E}^D\hat{B} \geq 0$ if and only if $E_{11}^{-1}B_S \geq 0$. Hence, we see that the intuitive conditions for positivity in this index 1 example are exactly reflected in the corresponding conditions of Theorem 3.8.

Discrete-time

Consider a system of the form (2), where $E = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix}$ with $E_{11} \in \mathbb{R}^{r\times r}$ regular and A, B are partitioned accordingly:

$$\begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t). \tag{3.26}$$

Assume that A_{22} is invertible. Then we can reduce the descriptor system to a standard system by the same procedure as in the continuous-time case. We premultiply the system (3.24) by the matrix $\begin{bmatrix} E_{11}^{-1} & -E_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & I_{n-r} \end{bmatrix}$ and obtain:

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} E_{11}^{-1}A_S & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} E_{11}^{-1}B_S \\ B_2 \end{bmatrix} u(t),$$

where $A_S = (A_{11} - A_{12}A_{22}^{-1}A_{21})$ and $B_S = (B_1 - A_{12}A_{22}^{-1}B_2)$. For the solution of the transformed system we obtain

$$\begin{aligned} x_1(t) &= (E_{11}^{-1}A_S)^t x_{10} + \sum_{\tau=0}^{t-1} (E_{11}^{-1}A_S)^{(t-1-\tau)} E_{11}^{-1} B_S u(\tau) \\ x_2(t) &= -A_{22}^{-1}(A_{21}x_1(t) + B_2 u(t)). \end{aligned}$$

If we assume that $E_{11}^{-1}A_S \geq 0$, $E_{11}^{-1}B_S \geq 0$ and $A_{22}^{-1}B_2, A_{22}^{-1}A_{21} \leq 0$, then for any input function $u \geq 0$ we obtain $x(t) \geq 0$ for all $t \geq 0$, i.e., the system is internally positive.

Theorem 3.18 *Consider a system of the form (2), where (E, A) is a regular matrix pair of $\operatorname{ind}(E, A) = 1$. If there exist regular matrices P, Q, R with $Q, R \geq 0$ such that with $\tilde{E} = PEQ$, $\tilde{A} = PAQ$ and $\tilde{B} = PBR$ we obtain a system of the form (3.26) and, if we additionally assume that $E_{11}^{-1}A_S \geq 0$, $E_{11}^{-1}B_S \geq 0$ and $A_{22}^{-1}B_2, A_{22}^{-1}A_{21} \leq 0$, then the system in (2) is internally positive and can be reduced to an internally positive standard system.*

Note that as in the continuous-time case the transformation matrices Q, R have to be nonnegative whereas the matrix P can be chosen arbitrarily, since it only scales the system without changing the solution.

For completeness, we compare this to the results of Theorem 3.12. Assuming $\operatorname{ind}(E, A) = 1$, for system (3.26) we obtain the same matrices as in (3.25). Theorem 3.12 now states that if $-A_{22}^{-1}A_{21} \geq 0$ then system (3.26) is positive if and only if $E_{11}^{-1}A_S \geq 0$, $E_{11}^{-1}B_S \geq 0$ and $A_{22}^{-1}B_2 \leq 0$. Hence, the intuitive assumptions on the index 1 system that we made here are reflected in the conditions of Theorem 3.12.

Summary

In this chapter we have reviewed the definition and the well-known characterisations of positivity in the standard case. We have generalised the definition of positivity to the descriptor case and have provided a characterisation of positivity that corresponds to the characterisation of standard positive systems. The special structure of the system matrices imposed by this characterisation has been specified. Finally, we have used a Schur complement decoupling approach to illustrate the obtained conditions by an example in a special index 1 situation.

Chapter 4

Stability of positive systems

True stability results when presumed order and presumed disorder are balanced. A truly stable system expects the unexpected, is prepared to be disrupted, waits to be transformed.

- Tom Robbins, "Even Cowgirls Get the Blues"

In the course of this section, we discuss asymptotic stability properties of positive systems. To this end, we consider linear homogeneous time-invariant systems:

- in continuous-time:
$$E\dot{x}(t) = Ax(t),\ x(0) = x_0, \tag{4.1}$$

- or in discrete-time:
$$Ex(t+1) = Ax(t),\ x(0) = x_0. \tag{4.2}$$

The following two definitions describe Lyapunov and asymptotic stability of descriptor systems in the continuous-time case as well as in the discrete-time case [115].

Definition 4.1 (Lyapunov stability) *The trivial solution $x(t) \equiv 0$ of the systems in (4.1) or in (4.2), respectively, is called Lyapunov stable, if for all $\epsilon > 0$ there exists $\delta > 0$, so that $\|x(t, x_0)\| < \epsilon$ for all $t \geq 0$ and for all $x_0 \in \operatorname{im} \hat{E}^D\hat{E}$ with $\|x_0\| < \delta$.*

Definition 4.2 (Asymptotic stability) *The trivial solution $x(t) \equiv 0$ of system (4.1) or (4.2), respectively, is called asymptotically stable, if*

1. *it is Lyapunov stable and*

2. *there exists $\delta > 0$, such that for all $x_0 \in \operatorname{im} \hat{E}^D \hat{E}$ with $\|x_0\| < \delta$ we have that $x(t, x_0) \to 0$ as $t \to \infty$.*

The following Theorem 4.3 is a well-known characterisation of asymptotically stable continuous-time and discrete-time systems in terms of the spectral properties of the corresponding matrix pair (E, A) [34], [52].

Theorem 4.3 *Let (E, A) be a regular matrix pair. The trivial solution $x(t) \equiv 0$ of the system in (4.1)*

(i) is asymptotically stable, if and only if all finite eigenvalues of (E, A) have negative real part.

(ii) is Lyapunov stable, if and only if all finite eigenvalues of (E, A) have nonpositive real part and the eigenvalues with zero real part have the same algebraic and geometric multiplicities.

The trivial solution $x(t) \equiv 0$ of the system in (4.2)

(i) is asymptotically stable, if and only if all finite eigenvalues of (E, A) are of modulus less than 1.

(ii) is Lyapunov stable, if and only if all finite eigenvalues of (E, A) are of modulus less than or equal to 1 and the eigenvalues with modulus equal to 1 have the same algebraic and geometric multiplicities.

Since stability properties of linear time-invariant systems depend only on the spectral properties of the matrix pair (E, A), the following definition is useful.

Definition 4.4 (c-/d-stable matrix pair) *A regular matrix pair (E, A) is called* c-stable *if $\sigma_f(E, A) \in \mathbb{C}_-$. A regular matrix pair (E, A) is called* d-stable *if $\rho_f(E, A) < 1$.*

Note that Definition 4.4 generalises the usual stability definition for matrices, i.e., a matrix A is called c-stable (d-stable) if (I, A) or equivalently A is c-stable (d-stable). The following theorem states the Lyapunov characterisation of stability, see [105], [115].

Theorem 4.5 *Let (E, A) be a regular matrix pair. The pair (E, A) is c-stable if and only if there exists a positive definite matrix X such that*

$$Q := E^T X A + A^T X E \preceq 0,$$

and Q is negative definite on S_f^{def}.
The pair (E, A) is d-stable if and only if there exists a positive definite matrix X such that

$$Q := A^T XA - E^T XE \preceq 0,$$

and Q is negative definite on S_f^{def}.

In the case of positive systems, classical stability criteria take a simple form. Such criteria for standard positive systems can be found in [38], [64] and are presented in Section 4.1. The main tool for standard positive systems that allows this simplification is the classical Perron-Frobenius Theorem 2.1. In Section 2.2, we have established a new generalisation of the Perron-Frobenius theory to matrix pairs that is applicable in the descriptor case. We show that this theory allows the same simplifications of standard stability criteria for positive systems in the descriptor case. These are presented in Section 4.2.

Definition 4.6 (c-/d-positive matrix pair) *We call a matrix pair (E, A)* c-positive *if system (4.1) is positive. We call a matrix pair (E, A)* d-positive *if system (4.2) is positive.*

Remark 4.7 *Note that by Theorem 3.8, if $\hat{E}^D\hat{E} \geq 0$, then (E, A) is c-positive if and only if there exists $\alpha \geq 0$ such that $\hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E} \geq 0$. By Theorem 3.12, if $\hat{E}^D\hat{E} \geq 0$, then (E, A) is d-positive if and only if $\hat{E}^D\hat{A} \geq 0$.*

Stability conditions for positive systems are closely related to and can be characterised by the so called dominant eigenvalue(s) of the system.

Definition 4.8 (c-/d-dominant eigenvalue) *For linear continuous-time systems (4.1), we call a finite eigenvalue λ of the matrix pair (E, A)* c-dominant *if its real part is greater than or equal to the real part of any other eigenvalue of the matrix pair (E, A), i.e. $\Re(\lambda) \geq \Re(\lambda_i)$ for all $\lambda_i \in \sigma_f(E, A)$.*
For linear discrete-time systems (4.2), we call a finite eigenvalue of the matrix pair (E, A) d-dominant *if it is greater than or equal in modulus to any other eigenvalue of the matrix pair (E, A), i.e. $|\lambda| \geq |\lambda_i|$ for all $\lambda_i \in \sigma_f(E, A)$.*

4.1 Standard positive systems

In this subsection we summarise the main stability conditions for standard positive systems. These conditions take a simple form compared to those for unconstrained systems.

The result of the following Theorem 4.9, see e.g. [38], allows to relax the condition of Theorem 4.3 that all eigenvalues have to be in the open left complex half-plane in the continuous-time case to considering only the real eigenvalues. The same applies to the discrete-time case, where the following result ensures that it is sufficient to check that the real eigenvalues are in modulus less than 1.

Theorem 4.9 *For a continuous-time standard positive system, i.e., system (4.1) with $E = I$, the c-dominant eigenvalue is real and unique. There exists a corresponding nonnegative eigenvector.*
For a discrete-time standard positive system, i.e., system (4.2) with $E = I$, $\rho(A)$ is a d-dominant eigenvalue and there exists a corresponding nonnegative eigenvector.

The next result relaxes the Lyapunov condition in Theorem 4.5, [38]. Instead of positive definite Lyapunov functions, here positive definite matrices, for stability of positive systems it is enough to consider diagonal matrices with a positive diagonal.

Theorem 4.10 *The matrix A is c-stable if and only if there exists a positive definite diagonal matrix X such that the matrix $A^T X + XA$ is negative definite.*
The matrix A is d-stable if and only if there exists a positive definite diagonal matrix X such that the matrix $A^T XA - X$ is negative definite.

Note that this relaxation is possible due to the diagonal stability property of M-matrices [6], [17]. For matrix diagonal stability in a more general context see [66] and the references therein. Furthermore, for positive systems, additional stability conditions are given in [38] in terms of certain M-matrix properties of the matrix $-A$ in the continuous-time case and of the matrix $I - A$ in the discrete-time case.

Theorem 4.11 *The matrix A is c-stable if and only if one of the following conditions holds*

1. *all principal minors of the matrix $-A$ are positive;*
2. *the coefficients of the characteristic polynomial of the matrix $-A$ are negative.*

The matrix A is d-stable if and only if one of the above conditions holds for $I - A$.

4.2 Positive descriptor systems

In this subsection we generalise the stability conditions for positive systems from the standard case in Section 4.1 to the descriptor case [124].

In the following Theorem 4.12, we generalise the result on dominant eigenvalues in Theorem 4.9 to descriptor systems.

Theorem 4.12 *Let (E, A) be a regular matrix pair. Consider the positive continuous-time system (4.1). If $\sigma_f(E, A) \neq \emptyset$ and $\hat{E}^D \hat{E} \geq 0$, where $\hat{E}$ is defined as in Lemma 1.19, then the c-dominant eigenvalue λ of the system is real and unique. Furthermore, there exists a nonnegative eigenvector corresponding to λ.*
Consider the positive discrete-time system (4.2). If $\sigma_f(E, A) \neq \emptyset$ and $\hat{E}^D \hat{E} \geq 0$, then $\rho_f(E, A)$ is a d-dominant eigenvalue and there exists a corresponding nonnegative eigenvector.

Proof. In the continuous-time case, since $\hat{E}^D \hat{E} \geq 0$, by Remark 4.7 and Remark 3.7 we have that there exists a scalar $\alpha > 0$ such that for the shifted matrix pair $(E, A + \alpha E)$, by the generalised Perron-Frobenius Theorem 2.8, the finite spectral radius $\rho_f(E, A + \alpha E) =: \mu$ is an eigenvalue. Hence, $\lambda = \mu - \alpha$ is an eigenvalue of (E, A) and it is the eigenvalue with the largest real part, i.e., the c-dominant eigenvalue of the positive system (4.1). Hence, the c-dominant eigenvalue λ is real and unique. Figure 4.1 depicts the situation. By Remark 3.7 there exists a corresponding nonnegative eigenvector.

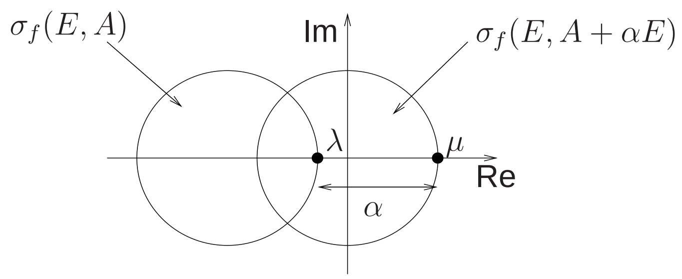


Figure 4.1: The c-dominant eigenvalue λ of (E, A) is real and unique.

For a positive discrete-time system (2), by Remark 4.7, if $\hat{E}^D \hat{E} \geq 0$, we have that $\hat{E}^D \hat{A} \geq 0$. Hence, by the generalised Perron-Frobenius Theorem 2.8 and using the identity in (2.14), the finite spectral radius of (E, A) is an eigenvalue and, by Remark 3.7, there exists a corresponding nonnegative eigenvector. □

Theorem 4.12 implies that a c-positive matrix pair is c-stable if and only if all of its real eigenvalues have negative real part. Analogously, a d-positive matrix pair is d-stable if and only if all of its real eigenvalues are in modulus less than 1.

Example 4.13 *Let $E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. Since E and A commute, we have*

$E^DE = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ *and* $E^DA = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$*. Hence, the system (4.1) for this choice of* (E, A) *is positive, since*

$$e^{E^DAt}E^DEv = \begin{bmatrix} e^{-t} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} e^{-t}v_1 \\ 0 \end{bmatrix} \geq 0,$$

for all $v_1 \geq 0$*. Choosing* $\alpha = 1$*, we obtain*

$$E^DA + \alpha E^DE = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \geq 0.$$

Hence, $\mu := \rho(E^DA + \alpha E^DE) = 0$ *is an eigenvalue and the corresponding c-dominant eigenvalue of* (E, A) *is* $\lambda = \mu - \alpha = -1$*. This means that* (E, A) *is also c-stable. Note that although* $\mu = 0$*, due to the fact that* $E^DE \geq 0$*, we have a nonnegative eigenvector corresponding to* μ *and, hence, to* λ*, see Remark 3.7.*

For a c-stable matrix pair, the following Lemma provides an *associated c-stable matrix* that has all finite eigenvalues of (E, A) as eigenvalues and an additional stable eigenvalue $-\alpha$, where $\alpha > 0$ may be chosen arbitrarily, that corresponds to the eigenvalue ∞ of (E, A). In the case of positive systems, this associated c-stable matrix is in addition a $-M$-matrix and plays an essential role in the generalisation of properties of positive systems from the standard to the descriptor case.

Lemma 4.14 *Let* (E, A) *be a regular c-stable matrix pair. Then, for any* $\alpha > 0$ *we have that*

$$\bar{M} := -\alpha I + \hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E},$$

is a stable (regular) matrix. If, in addition, the matrix pair (E, A) *is c-positive and* $\hat{E}^D\hat{E} \geq 0$*, then there exists* $\alpha > 0$ *such that* $\bar{M}$ *is a* $-M$*-matrix.*

Proof. All finite eigenvalues of (E, A) are also eigenvalues of $\hat{E}^D\hat{A}$ and the eigenvalue ∞ of (E, A) is mapped to the eigenvalue 0 of $\hat{E}^D\hat{A}$, see Section 2.2. For any finite eigenpair (λ, v) of (E, A), we have

$$\bar{M}v = \hat{E}^D\hat{A}v = \lambda v.$$

Therefore, all stable finite eigenvalues of the pair (E, A) are stable eigenvalues of $\bar{M}$. For any eigenvector w corresponding to the eigenvalue ∞ of (E, A), i.e., $Ew = 0$, we have by the properties of $\hat{E}, \hat{A}$ in Lemma 1.19 and Equations (1.13) that

$$\hat{E}^D\hat{A}w = \hat{E}^D\hat{A}\hat{E}^D\hat{E}w = \hat{E}^D\hat{A}\hat{E}^D(\lambda E - A)^{-1}Ew = 0,$$

and hence,

$$\bar{M}w = -\alpha w.$$

Thus, w is now an eigenvector corresponding to a negative eigenvalue $-\alpha$. Hence, all eigenvalues of $\bar{M}$ have negative real parts and therefore $\bar{M}$ is stable. If, in addition, the matrix pair (E, A) is c-positive and $\hat{E}^D\hat{E} \geq 0$, then by Remark 4.7 we have that there exists $\alpha > 0$ such that

$$T := \hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E} \geq 0.$$

By the generalised Perron-Frobenius Theorem 2.8 we have that $\rho(T)$ is an eigenvalue of T and $\rho(T) - \alpha$ is the finite eigenvalue of (E, A) with the largest real part and it is negative, since (E, A) is c-stable. Therefore, we have $\alpha > \rho(T)$ and

$$\bar{M} = -(\alpha I - T)$$

is a $-M$-matrix. □

In the following we generalise a Lyapunov-type stability condition from the standard case in Theorem 4.10 to the descriptor case.

Theorem 4.15 *Let the matrix pair (E, A) be regular and let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19. If (E, A) is c-positive and $\hat{E}^D\hat{E} \geq 0$, then the pair (E, A) is c-stable if and only if there exists a positive definite diagonal matrix Y such that*

$$(\hat{E}^D\hat{A})^TY + Y(\hat{E}^D\hat{A}) \preceq 0,$$

and $(\hat{E}^D\hat{A})^TY + Y(\hat{E}^D\hat{A})$ is negative definite on S_f^{def}.
If (E, A) is d-positive and $\hat{E}^D\hat{E} \geq 0$, then (E, A) is d-stable if and only if there exists a positive definite diagonal matrix Y such that

$$(\hat{E}^D\hat{A})^TY(\hat{E}^D\hat{A}) - Y \prec 0.$$

Proof. Continuous-time case:
"⇒"By Lemma 4.14, we have that there exists $\alpha > 0$ such that the matrix

$$M := \alpha I - (\hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E}),$$

is a regular M-matrix. For all $v \in S_f^{def}$ we have by (3.9) that

$$v^T(\hat{E}^D\hat{A})^TYv + v^TY(\hat{E}^D\hat{A})v = v^T(-M)^TYv + v^TY(-M)v.$$

It is well known that for an M-matrix M there exists a positive definite diagonal matrix Y so that the matrix $-(M^TY + YM)$ is negative definite, see, e.g., [6], [11], [17], [38].

Hence, Y is a positive definite diagonal matrix such that $(\hat{E}^D\hat{A})^TY+Y(\hat{E}^D\hat{A})$ is negative definite on S_f^{def}. For any $w \in \mathbb{R}^n \setminus S_f^{def}$, we have $\hat{E}^D\hat{A}w = 0$ and hence, $(\hat{E}^D\hat{A})^TY + Y(\hat{E}^D\hat{A})$ is negative semidefinite on $\mathbb{R}^n$.
"$\Leftarrow$" We have to show that all finite eigenvalues of (E, A) have negative real part. If $\sigma_f(E, A) = \emptyset$, there is nothing to prove. Therefore, assume that $\sigma_f(E, A) \neq \emptyset$. Then, by Theorem 4.12, we have that the c-dominant eigenvalue λ of (E, A) is real and unique. Hence, it suffices to show that λ is negative. Let v be an eigenvector corresponding to λ. Since the eigenpair (λ, v) is also an eigenpair of $\hat{E}^D\hat{A}$, see Section 2.2, we obtain

$$v^T(\hat{E}^D\hat{A})^TYv + v^TY(\hat{E}^D\hat{A})v = v^T\lambda Yv + v^TY\lambda v = 2\lambda v^TYv < 0,$$

whereas $v^TYv > 0$. Hence, $\lambda < 0$.

Discrete-time case:
"$\Rightarrow$" If $\hat{E}^D\hat{E} \geq 0$, for a positive system we also have $\hat{E}^D\hat{A} \geq 0$, see Remark 4.7. Since the matrix pair (E, A) is d-stable, we have $\rho_f(E, A) < 1$ and hence, the matrix

$$M := I - \hat{E}^D\hat{A},$$

is a regular M-matrix. Therefore, there exists a diagonal positive definite matrix Y so that the matrix $(\hat{E}^D\hat{A})^TY(\hat{E}^D\hat{A}) - Y$ is negative definite, see, e.g. [6], [38].
"$\Leftarrow$" As in the continuous-time case, we assume that $\sigma_f(E, A) \neq \emptyset$. Then, by Theorem 4.12, we have that there exists a d-dominant eigenvalue λ of (E, A) that is nonnegative and real. Hence, it suffices to show that λ is less than 1. Let v be an eigenvector corresponding to λ. Since the eigenpair (λ, v) is also an eigenpair of $\hat{E}^D\hat{A}$, see Section 2.2, we obtain

$$v^T(\hat{E}^D\hat{A})^TY(\hat{E}^D\hat{A})v - v^TYv = \lambda^2v^TYv - v^TYv = (\lambda^2 - 1)v^TYv < 0,$$

whereas $v^TYv > 0$. Since λ is nonnegative, we have $\lambda < 1$. $\square$

The following corollary restates the result of Theorem 4.15 in terms of the continuous-time and discrete-time generalised projected Lyapunov operators, as introduced in [115] for descriptor systems, that are used in Theorem 4.5 for characterising stability properties of unconstrained descriptor systems. To adhere the condition of the existence of a diagonal Lyapunov function, however, an additional condition is needed.

Corollary 4.16 *Let the matrix pair (E, A) be regular of $\operatorname{ind}(E, A) = \nu$ and let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19. Assume that (E, A) is c-positive with $P_r = \hat{E}^D\hat{E} \geq 0$ and that $P_rE_\nu^{-1}$ is diagonal, where E_ν is defined as in (1.6). Then, the pair (E, A) is c-stable if and only if there exists a positive definite diagonal matrix X such that*

$$E^TXA + A^TXE \preceq 0,$$

and $E^TXA + A^TXE$ is negative definite on S_f^{def}.
For the discrete-time case, assume that (E, A) is d-positive with $P_r \geq 0$ and that $P_rE_\nu^{-1}$ is diagonal. Then (E, A) is d-stable if and only if there exists a positive definite diagonal matrix X such that

$$A^TXA - E^TXE \preceq 0,$$

and $A^TXA - E^TXE$ is negative definite on S_f^{def}.

Proof. Consider first the continuous-time case. It is enough to prove the if part, since the only if part follows directly from Theorem 4.5. By Theorem 4.15 we have that there exists a diagonal positive definite matrix Y such that $(\hat{E}^D\hat{A})^TY + Y(\hat{E}^D\hat{A}) \preceq 0$. Setting

$$X := E_\nu^{-T}P_r^TYP_rE_\nu^{-1}, \tag{4.3}$$

we have that X is diagonal positive definite. For all $v \in S_f^{def}$ we have

$$\begin{aligned} v^TE^TXAv + v^TA^TXEv =& v^TE^T(E_\nu^{-T}P_r^TYP_rE_\nu^{-1})Av + v^TA^T(E_\nu^{-T}P_r^TYP_rE_\nu^{-1})Ev \\ =& v^TY\hat{E}^D\hat{A}v + v^T(\hat{E}^D\hat{A})^TYv < 0, \end{aligned}$$

where we have used that $P_rE_\nu^{-1}E = P_r$ and $P_rE_\nu^{-1}A = \hat{E}^D\hat{A}$, see Section 2.2.3. Hence, $E^TXA + A^TXE$ is negative definite on S_f^{def}. Furthermore, for any $w \in \mathbb{R}^n \setminus S_f^{def}$, we have $P_rw = 0$ and hence, $E^TXA + A^TXE$ is negative semidefinite on $\mathbb{R}^n$.
Consider now the discrete-time case. As in the continuous-time case it is enough to prove the if part, since the only if part follows directly from Theorem 4.5. By Theorem 4.15 we have that there exists a diagonal positive definite matrix Y such that $(\hat{E}^D\hat{A})^TY(\hat{E}^D\hat{A}) - Y \prec 0$. Setting X as in (4.3) we obtain that X is diagonal positive definite and for all $v \in S_f^{def}$ we have

$$\begin{aligned} v^TA^TXAv - v^TE^TXEv =& v^TA^T(E_\nu^{-T}P_r^TYP_rE_\nu^{-1})Av - v^TE^T(E_\nu^{-T}P_r^TYP_rE_\nu^{-1})Ev = \\ =& v^T(\hat{E}^D\hat{A})^TY\hat{E}^D\hat{A}v - v^TYv < 0. \end{aligned}$$

Hence, $A^TXA - E^TXE$ is negative definite on S_f^{def}. Furthermore, for any $w \in \mathbb{R}^n \setminus S_f^{def}$, we have $P_rw = 0$ and hence, $A^TXA + E^TXE$ is negative semidefinite on $\mathbb{R}^n$.
□

The following corollary restates the result of Theorem 4.15 in terms of M-matrix properties.

Corollary 4.17 *Let the matrix pair (E, A) be regular and let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19. If (E, A) is c-positive and $\hat{E}^D\hat{E} \geq 0$, then the matrix pair (E, A) is c-stable if and only if there exists a scalar $\alpha > 0$ such that for the matrix $M := \alpha I - (\hat{E}^D\hat{A} + \alpha\hat{E}^D\hat{E})$ one of the following properties holds*

1. *all principal minors of M are positive;*

2. *the coefficients of the characteristic polynomial of M are negative.*

If (E, A) is d-positive and $\hat{E}^D\hat{E} \geq 0$, then the matrix pair (E, A) is d-stable if and only if one of the properties 1.-2. holds for the matrix $M := I - \hat{E}^D\hat{A}$.

Proof. In the continuous-time case, by Lemma 4.14 there exists $\alpha > 0$ such that M is an M-matrix. In the discrete-time case, M is an M-matrix by Theorem 4.15. Therefore, the assertions of this corollary follow directly from the M-matrix properties [17], [38]. □

4.3 Stability of switched positive descriptor systems

The study of stability properties of switched systems is subject to ongoing research, see [112] and the references therein. Especially, in the case of standard positive systems, progress has been made on this subject due to the existence of a diagonal Lyapunov function, see, e.g., [90], [91], and the references therein. The existence of a common diagonal Lyapunov function of two positive systems, i.e. a diagonal positive definite matrix Y such that

$$\begin{aligned} &A_1^TY + YA_1 \quad \text{and} \\ &A_2^TY + YA_2 \end{aligned}$$

are negative definite, guarantees the stability of the switched system under arbitrary switching. In this section, we show how we can use the framework established throughout this chapter in order to generalise these results to positive descriptor systems.

The following sufficient conditions for the existence of a common diagonal Lyapunov function in the standard case can be found, e.g., in [90], [91].

Theorem 4.18 *Let $A_1, A_2 \in \mathbb{R}^{n\times n}$ be $-M$-matrices, i.e., stable $-Z$-matrices. Then, each of the following conditions is sufficient for the existence of a common diagonal Lyapunov function:*

1. *$A_1A_2^{-1}$ and $A_2^{-1}A_1$ are both M-matrices.*

2. *$A_1A_2^{-1}$ and $A_2^{-1}A_1$ are both nonnegative.*

The generalisation to positive descriptor systems uses Theorem 4.15 and is as follows.

Theorem 4.19 *Let $(E_1, A_1), (E_2, A_2)$ be two c-stable matrix pairs and let $\hat{E}_i, \hat{A}_i,\ i = 1, 2$ be defined as in Lemma 1.19 with $\hat{E}_1^D \hat{E}_1 \geq 0$ and $\hat{E}_2^D \hat{E}_2 \geq 0$. Then there exist scalars $\alpha_1, \alpha_2 > 0$ such that*

$$\begin{aligned} M_1 &:= \alpha I - \hat{E}_1^D \hat{A}_1 - \alpha \hat{E}_1^D \hat{E}_1, \quad \text{and} \\ M_2 &:= \alpha I - \hat{E}_2^D \hat{A}_2 - \alpha \hat{E}_2^D \hat{E}_2 \end{aligned}$$

are M-matrices and each of the following conditions is sufficient for the existence of a common diagonal Lyapunov function:

1. *$M_1 M_2^{-1}$ and $M_2^{-1} M_1$ are both M-matrices.*
2. *$M_1 M_2^{-1}$ and $M_2^{-1} M_1$ are both nonnegative.*

Proof. By Lemma 4.14, there exist scalars $\alpha_1, \alpha_2 > 0$ such that M_1, M_2 are M-matrices. The rest follows as in the proof of the standard case in Theorem 4.18. □

Summary

In this chapter we have discussed stability properties of positive descriptor systems in the continuous-time as well as in the discrete-time case. We have reviewed some prevalent stability concepts in the positive standard case along with different stability criteria that take a simple form in the case of positive systems. We have presented generalisations of (internal) stability criteria for the case of positive descriptor systems. It was shown that if the spectral projector onto the finite deflating subspace of the matrix pair (E, A) is nonnegative, then all stability criteria for standard positive systems take a comparably simple form in the positive descriptor case. As an application of the framework established throughout this chapter, we have shown how stability criteria of switched standard positive systems can be extended to the descriptor case.

Chapter 5

Generalised Lyapunov equations for positive systems

An idea is always a generalisation, and generalisation is a property of thinking.
To generalise means to think.
Georg Wilhelm Friedrich Hegel

We consider the following projected generalised Lyapunov equations [115] in continuous-time

$$E^T XA + A^T XE = -P_r^T GP_r, \tag{5.1}$$

or in discrete-time

$$A^T XA - E^T XE = -P_r^T GP_r, \tag{5.2}$$

where $G \in \mathbb{R}^{n \times n}$ and P_r, as defined in (1.5), is the unique spectral projector onto the finite deflating subspace S_f^{def} of the pencil (E, A).

Lyapunov equations are named after Alexander Mikhailovitch Lyapunov, who presented the stability theory for linear and nonlinear systems in 1892 [84].

Lyapunov equations have been studied in many different contexts, especially in applications such as differential and difference equations [45], [48], [105], [114], [115]. Stability properties in systems theory and also many other control theoretical issues such as model reduction methods or the quadratic optimal control problem are also closely related to the solution of Lyapunov equations, see. e.g., [5], [44], [45], [79], [95].

The following Theorem 5.1 gives necessary and sufficient conditions for the existence of solutions of (5.1) [115].

Theorem 5.1 *Let (E, A) be a regular matrix pair and let P_r and P_l be the spectral projectors onto the right and left finite deflating subspaces defined in (1.5). Furthermore, let $\lambda_1, \ldots, \lambda_{n_f}$, where $n_f = \operatorname{rank}(P_r)$ be the finite not necessarily distinct eigenvalues of (E, A). The generalised Lyapunov equation (5.1) has a solution for every matrix G if and only if $\lambda_j + \bar{\lambda}_k \neq 0$ for all $j, k = 1, \ldots, n_f$. If, in addition, the solution X is required to satisfy the condition $X = XP_l$, then it is unique.*

The additional condition $X = XP_l$ corresponds to the requirement that the nonunique part of the solution X is zero in the Weierstraß canonical form. Therefore, in [115], the following system of equations is considered

$$\begin{aligned} E^T XA + A^T XE &= -P_r^T GP_r, \\ X &= XP_l. \end{aligned} \tag{5.3}$$

For the discrete-time case, consider the projected generalised discrete-time Lyapunov equation (5.2). The following Theorem 5.2 gives necessary and sufficient conditions for the existence of solutions of (5.2) [115].

Theorem 5.2 *Let E, A be a regular matrix pair and let P_r and P_l be the spectral projectors onto the right and left finite deflating subspaces defined in (1.5). Furthermore, let $\lambda_1, \ldots, \lambda_{n_f}$, where $n_f = \operatorname{rank}(P_r)$ be the finite not necessarily distinct eigenvalues of (E, A). The generalised Lyapunov equation (5.2) has a solution for every matrix G if and only if $\lambda_j \bar{\lambda}_k \neq 1$ for all $j, k = 1, \ldots, n_f$. If, in addition, the solution X is required to satisfy the condition $P_l^T X = XP_l$, then it is unique.*

As in the continuous-time case, the additional condition $P_l^T X = XP_l$ corresponds to the requirement that the nonunique part of the solution X is zero in the Weierstraß canonical form. Therefore, in [115], the following system of equations is considered

$$\begin{aligned} A^T XA - E^T XE &= -P_r^T GP_r, \\ P_l^T X &= XP_l. \end{aligned} \tag{5.4}$$

In the context of positive systems one is interested not only in positive (semi)definite solutions of such Lyapunov equations but rather in doubly nonnegative solutions, i.e., solutions that are both positive semidefinite and entry-wise nonnegative. Such results for standard Lyapunov equations are well known. We summarise these in the next Section 5.1.

5.1 Doubly nonnegative solutions of projected generalised Lyapunov equations: standard case

The following result is on the nonnegativity of the solution of the standard continuous-time Lyapunov equation.

Theorem 5.3 *Let A be a $-Z$-matrix and let A be c-stable, i.e., $\sigma(A) \subset \mathbb{C}_-$. Then, the solution X of the Lyapunov equation*

$$A^T X + XA = -G \tag{5.5}$$

is given by

$$X = \int_0^\infty e^{A^T t} G e^{At} dt. \tag{5.6}$$

The solution X is positive (semi)definite for any positive (semi)definite matrix G and X is nonnegative for any nonnegative G. Moreover, if $G > 0$ then, $X > 0$.

Proof. For a c-stable matrix A it is well known, that the solution to (5.5) can be explicitly given by (5.6), see, e.g., [60]. Since A is a $-Z$-matrix, we get by Lemma 1.3 that $e^{A^T t} \geq 0$ and $e^{At} \geq 0$ for all $t \geq 0$. Since $G \geq 0$ ($G > 0$) and integration is monotone, we get that $X \geq 0$ ($X > 0$). Furthermore, if $G \succeq 0$ ($G \succ 0$) we have that $X \succeq 0$ ($X \succ 0$). □

Remark 5.4 *Note, that the condition in Theorem 5.3 that A is a c-stable $-Z$-matrix is equivalent to the condition that A is a regular $-M$-matrix, see Lemma 1.8. The proof to Theorem 5.3 can therefore alternatively be accomplished by forming $\operatorname{vec} X$, which stacks the columns $X_1, \ldots, X_n$ of the matrix X into a long vector $\begin{bmatrix} X_1^T & \ldots & X_n^T \end{bmatrix}^T$ and by solving instead of (5.5) the equivalent linear system*

$$(I \otimes A^T + A^T \otimes I) \operatorname{vec} X = -\operatorname{vec} G, \tag{5.7}$$

where $\otimes$ is the Kronecker product. Since A is a regular $-M$-matrix, the matrix $(I \otimes A^T + A^T \otimes I)$ is also a regular $-M$-matrix [55] and hence $(I \otimes A^T + A^T \otimes I)^{-1} \leq 0$. Therefore, we conclude that if $G \geq 0$, we get $X \geq 0$ and if $G > 0$ we get $X > 0$.

The following result is on positivity of the solution of the standard discrete-time Lyapunov equation.

Theorem 5.5 *Let $A \geq 0$ be d-stable, i.e., $\rho(A) < 1$. Then, the solution X of the Lyapunov equation*

$$A^T XA - X = -G \tag{5.8}$$

is given by

$$X = \sum_{k=0}^{\infty} (A^T)^k G A^k. \tag{5.9}$$

The solution X is positive (semi)definite for any positive (semi)definite matrix G and X is nonnegative for any nonnegative G. Moreover, if $A, G > 0$ then, $X > 0$.

Proof. The series in (5.9) is absolutely convergent since the spectral radius of A is less than one and (5.9) is a solution since

$$\begin{aligned} A^T XA - X &= A^T \sum_{k=0}^{\infty} (A^T)^k G A^k A - \sum_{k=0}^{\infty} (A^T)^k G A^k = \\ &= \sum_{k=1}^{\infty} (A^T)^k G A^k - \sum_{k=0}^{\infty} (A^T)^k G A^k = -G. \end{aligned}$$

Now, if $A, G \geq 0$ ($A, G > 0$), we get $X \geq 0$ ($X > 0$). Furthermore, if $G \succeq 0$ ($G \succ 0$) we have that $X \succeq 0$ ($X \succ 0$). □

In the following Section 5.2 we extend the well-known results of this section to the descriptor case [124].

5.2 Doubly nonnegative solutions of projected generalised Lyapunov equations: descriptor case

Consider the projected generalised continuous-time Lyapunov equation given in (5.3) and recall that $P_r = \hat{E}^D \hat{E}$, see Section 1.6. The following Theorem 5.6 gives sufficient conditions for the existence of a doubly nonnegative solution of (5.3).

Theorem 5.6 *Let (E, A) be a regular c-stable matrix pair. Let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19 and assume that $\hat{E}^D \hat{E} \geq 0$. Then, system (5.3) has a unique solution for every matrix G. The solution is given by*

$$X = E_\nu^{-T} \left(\int_0^\infty e^{(\hat{E}^D \hat{A})^T t} P_r^T G P_r e^{(\hat{E}^D \hat{A})t} dt \right) E_\nu^{-1}, \tag{5.10}$$

where E_ν is defined as in the matrix chain in (1.6). If the matrix G is symmetric positive (semi)definite, then X is symmetric positive semidefinite. If, in addition, we have that the matrix pair (E, A) is c-positive, $P_r^T G P_r \geq 0$ and $P_r E_\nu^{-1} \geq 0$, then also $X \geq 0$.

Proof. We first show that X as defined in (5.10) is solution of (5.1). Since (E, A) is c-stable, by Lemma 4.14, we have that for any $\alpha > 0$ the matrix

$$\bar{M} := -\alpha I + \hat{E}^D \hat{A} + \alpha \hat{E}^D \hat{E}$$

is c-stable and $\bar{M} P_r = P_r \bar{M} = \hat{E}^D \hat{A}$. We now use the following properties that can be deduced from the properties of canonical projectors in [89], [96]:

$$E_\nu^{-1} A_i Q_i = -Q_i \quad \text{for all} \quad i = 0, \ldots, \nu - 1, \tag{5.11}$$

where E_ν, A_i are defined as in the matrix chain (1.6) with canonical projectors Q_i. By definition, we have $E_\nu = E - A_0 Q_0 - \ldots - A_{\nu-1} Q_{\nu-1}$ and with the identities in (5.11) we get

$$E_\nu^{-1} E = I - Q_0 - \ldots - Q_{\nu-1}. \tag{5.12}$$

Since $P_r = P_0 \ldots P_{\nu-1}$, where $P_i = I - Q_i$, we have, [89], [96],

$$P_r Q_i = 0, \quad \text{for all} \quad i = 0, \ldots, \nu - 1. \tag{5.13}$$

By using this, we obtain that

$$\begin{aligned} E^T X E \quad &= \quad E^T E_\nu^{-T} \left(\int_0^\infty e^{(\hat{E}^D \hat{A})^T t} P_r^T G P_r e^{(\hat{E}^D \hat{A}) t} dt \right) E_\nu^{-1} E \\ &\overset{(5.12)}{=} \quad E^T E_\nu^{-T} \left(\int_0^\infty e^{(\hat{E}^D \hat{A})^T t} P_r^T G P_r e^{(\hat{E}^D \hat{A}) t} dt \right) (I - Q_0 - \ldots - Q_{\nu-1}) \\ &\overset{(5.13)}{=} \quad \int_0^\infty e^{(\hat{E}^D \hat{A})^T t} P_r^T G P_r e^{(\hat{E}^D \hat{A}) t} dt \\ &\overset{(3.10)}{=} \quad \int_0^\infty e^{\bar{M}^T t} P_r^T G P_r e^{\bar{M} t} dt \end{aligned}$$

is a solution of the standard Lyapunov equation

$$(E^T X E) \bar{M} + \bar{M}^T (E^T X E) = -P_r^T G P_r.$$

On the other hand, by using the identity (2.14), we obtain

$$\begin{aligned} A^T X E &= A^T E_\nu^{-T} \left(\int_0^\infty e^{(\hat{E}^D \hat{A})^T t} P_r^T G P_r e^{(\hat{E}^D \hat{A})t} dt \right) E_\nu^{-1} E = \\ &= P_r^T A^T E_\nu^{-T} \left(\int_0^\infty e^{(\hat{E}^D \hat{A})^T t} P_r^T G P_r e^{(\hat{E}^D \hat{A})t} dt \right) = \\ &= (\hat{E}^D \hat{A})^T \left(\int_0^\infty e^{(\hat{E}^D \hat{A})^T t} P_r^T G P_r e^{(\hat{E}^D \hat{A})t} dt \right) = \\ &= \bar{M}^T \left(\int_0^\infty e^{(\hat{E}^D \hat{A})^T t} P_r^T G P_r e^{(\hat{E}^D \hat{A})t} dt \right) = \\ &= \bar{M}^T (E^T X E), \end{aligned}$$

and analogously $E^T X A = (E^T X E)\bar{M}$. Hence, if we plug X defined in (5.10) into equation (5.1), then we obtain

$$\begin{aligned} E^T X A + A^T X E &= (E^T X E)\bar{M} + \bar{M}^T (E^T X E) = \\ &= -P_r^T G P_r. \end{aligned}$$

To show the uniqueness of the solution we make use of the Weierstraß canonical form. By the construction of E_ν, we have that

$$P_r E_\nu^{-1} = T^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, \tag{5.14}$$

and P_l is defined as in (1.5). Hence, we have that $P_r E_\nu^{-1} P_l = P_r E_\nu^{-1}$ and therefore, the condition $X = X P_l$ holds.

If G is positive (semi)definite, then X is positive semidefinite [115]. If (E, A) is c-positive and $P_r \geq 0$, then $e^{(\hat{E}^D \hat{A})t} P_r \geq 0$. With $P_r^T G P_r \geq 0$ and $P_r E_\nu^{-1} \geq 0$ we obtain $X \geq 0$. □

Remark 5.7 *Note that by (5.14), in the solution representation (5.10), the matrix E_ν can be replaced by any regular matrix R such that R^{-1} in the Weierstraß canonical form has the following structure*

$$R^{-1} = T^{-1} \begin{bmatrix} I & 0 \\ 0 & * \end{bmatrix} W^{-1},$$

where $*$ denotes an arbitrary regular submatrix. For instance, such a matrix could also be $R = (EP_r + A(I - P_r))$.

With such a matrix R, alternatively, the results of Theorem 5.6 may be obtained by *considering the Weierstraß canonical form and verifying that the relations*

$$P_r R^{-1} E = T^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T = P_r,$$
$$P_r R^{-1} A = T^{-1} \begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} T = \hat{E}^D \hat{A},$$

hold.

For the discrete-time case, consider the projected generalised discrete-time Lyapunov equation in (5.4). The following Theorem 5.8 gives sufficient conditions for the existence of a doubly nonnegative solution of (5.4).

Theorem 5.8 *Let (E, A) be a regular d-stable matrix pair. Let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19 and assume $\hat{E}^D \hat{E} \geq 0$. Then system (5.4) has a unique solution for every matrix G. The solution is given by*

$$X = E_\nu^{-T} \left(\sum_{t=0}^{\infty} ((\hat{E}^D \hat{A})^T)^t P_r^T G P_r (\hat{E}^D \hat{A})^t \right) E_\nu^{-1}, \tag{5.15}$$

where E_ν is defined as in the matrix chain in (1.6). If G is symmetric positive (semi)definite, then X is symmetric positive semidefinite. If, in addition, we have that the matrix pair (E, A) is d-positive, $P_r^T G P_r \geq 0$ and $P_r E_\nu^{-1} \geq 0$, then also $X \geq 0$.

Proof. We first show that X as defined in (5.15) is solution of (5.2). For X as defined in (5.15) we have that

$$E^T X E = E^T E_\nu^{-T} \left(\sum_{t=0}^{\infty} ((\hat{E}^D \hat{A})^T)^t G (\hat{E}^D \hat{A})^t \right) E_\nu^{-1} E = \sum_{t=0}^{\infty} ((\hat{E}^D \hat{A})^T)^t P_r^T G P_r (\hat{E}^D \hat{A})^t$$

is a solution of the standard discrete-time Lyapunov equation

$$(\hat{E}^D \hat{A})^T (E^T X E)(\hat{E}^D \hat{A}) - (E^T X E) = -P_r^T G P_r.$$

On the other hand, we have

$$A^T X A = (\hat{E}^D \hat{A})^T \left(\sum_{t=0}^{\infty} ((\hat{E}^D \hat{A})^T)^t G (\hat{E}^D \hat{A})^t \right) (\hat{E}^D \hat{A}).$$

Hence, if we plug X into equation (5.2), then we obtain

$$A^T X A - E^T X E = (\hat{E}^D \hat{A})^T (E^T X E)(\hat{E}^D \hat{A}) - (E^T X E) = -P_r^T G P_r.$$

Condition $P_l^T X = X P_l$ can be shown as in the continuous-time case by considering the Weierstraß canonical form representation in (5.14).

If (E, A) is d-positive and $P_r \geq 0$, then we have that $\hat{E}^D \hat{A} \geq 0$ [24]. With $P_r^T G P_r \geq 0$ and $P_r E_\nu^{-1} \geq 0$ we obtain $X \geq 0$. $\square$

Note that Remark 5.7 is also true in the discrete-time case.

5.3 Special case: index 1 systems

In this section, the Schur complement decoupling technique presented in Section 3.3 is applied to show nonnegativity of the solution of projected generalised Lyapunov equations for special systems of index 1.

In the following theorem we give sufficient conditions for the nonnegativity of the solution X of (5.3) for a special matrix E.

Theorem 5.9 *Let P_r and P_l be the spectral projectors onto the right and left finite deflating subspaces of the matrix pair (E, A) defined in (1.5), where $E = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix}$ with E_{11} regular and $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is partitioned accordingly. Suppose that A_{22} is invertible and $A_S E_{11}^{-1}$ is a regular $-M$-matrix, where $A_S = A_{11} - A_{12} A_{22}^{-1} A_{21}$.*
Moreover, suppose that $A_{12} A_{22}^{-1} \leq 0$. Then, for any G such that

$$P_r^T G P_r =: \begin{bmatrix} \tilde{G}_{11} & \tilde{G}_{12} \\ \tilde{G}_{21} & \tilde{G}_{22} \end{bmatrix}, \tag{5.16}$$

and $E_{11}^{-T} \tilde{G}_{11} E_{11}^{-1} \geq 0$, the unique solution X of Equation (5.3) satisfies $X \geq 0$.

Proof. We have that

$$PEQ = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad PAQ = \begin{bmatrix} A_S E_{11}^{-1} & 0 \\ 0 & A_{22} \end{bmatrix},$$

where $P = \begin{bmatrix} I & -A_{12} A_{22}^{-1} \\ 0 & I \end{bmatrix}$ and $Q = \begin{bmatrix} E_{11}^{-1} & 0 \\ -A_{22}^{-1} A_{21} E_{11}^{-1} & I \end{bmatrix}$ are regular matrices. Hence, we can compute the projectors P_r and P_l by

$$\begin{aligned} P_r &= Q \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} Q^{-1} = \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & 0 \end{bmatrix}, \\ P_l &= P^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} P = \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{5.17}$$

Scaling the first equation in (5.3) with Q^T and Q we obtain the equivalent equation

$$Q^T E^T X A Q + Q^T A^T X E Q = -Q^T P_r^T G P_r Q.$$

We have

$$EQ = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}E_{11}^{-1} & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} =: \tilde{E}, \tag{5.18}$$

and

$$AQ = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} E_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}E_{11}^{-1} & I \end{bmatrix} = \begin{bmatrix} A_S E_{11}^{-1} & A_{12} \\ 0 & A_{22} \end{bmatrix} =: \tilde{A}. \tag{5.19}$$

Since

$$P_r Q = \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & 0 \end{bmatrix} \begin{bmatrix} E_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}E_{11}^{-1} & I \end{bmatrix} = \begin{bmatrix} E_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}E_{11}^{-1} & 0 \end{bmatrix},$$

we obtain

$$\tilde{E}^T \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \tilde{A} + \tilde{A}^T \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \tilde{E} = \begin{bmatrix} -E_{11}^{-T}\tilde{G}_{11}E_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix},$$

where $\tilde{G}_{11} = (G_{11} - A_{21}^T A_{22}^{-T} G_{21} - (G_{12} - A_{21}^T A_{22}^{-T} G_{22}) A_{22}^{-1} A_{21})$. Hence, we have the following decoupled system of equations

$$\left\{ \begin{array}{rcl} X_{11}A_S E_{11}^{-1} + E_{11}^{-T} A_S^T X_{11} &=& -E_{11}^{-T}\tilde{G}_{11}E_{11}^{-1} \\ X_{11}A_{12} + X_{12}A_{22} &=& 0 \\ A_{12}^T X_{11} + A_{22}^T X_{21} &=& 0 \\ 0 &=& 0 \end{array} \right. .$$

By Theorem 5.3, since $A_S E_{11}^{-1}$ is a $-M$-matrix and $E_{11}^{-T}\tilde{G}_{11}E_{11}^{-1} \geq 0$, we obtain that the first equation has the unique solution $X_{11} \geq 0$. We have that A_{22} is invertible and $A_{12}A_{22}^{-1} \leq 0$. Therefore, from the second and third equations, we get

$$\begin{aligned} X_{12} &= -X_{11}A_{12}A_{22}^{-1} \geq 0, \\ X_{21} &= -A_{22}^{-T}A_{12}^T X_{11} \geq 0. \end{aligned}$$

Furthermore, since we required that $X = XP_l$, i.e.,

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} X_{11} & -X_{11}A_{12}A_{22}^{-1} \\ X_{21} & -X_{21}A_{12}A_{22}^{-1} \end{bmatrix},$$

we obtain $X_{22} = -X_{21}A_{12}A_{22}^{-1} \geq 0$. Hence, we conclude that $X \geq 0$. □

Corollary 5.10 *Consider Equation (5.3), where* $E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ *and* $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ *is partitioned accordingly. If* A *is a regular* $-M$*-matrix, then for any* $G \geq 0$ *the unique solution* X *satisfies* $X \geq 0$*.*

Proof. Since A is a regular $-M$-matrix, we have that the Schur complement $A_S = A_{11} - A_{12}A_{22}^{-1}A_{21}$ is a regular $-M$-matrix [85]. Also we know that $-A_{22}^{-1} \geq 0$ and $A_{12} \geq 0$. Hence, $A_{12}A_{22}^{-1} \leq 0$. Finally we obtain $P_r^T G P_r \geq 0$ for any $G \geq 0$ since $\tilde{G}_{11} = G_{11} - A_{21}^T A_{22}^{-T} G_{21} - (G_{12} - A_{21}^T A_{22}^{-T} G_{22})A_{22}^{-1}A_{21} \geq 0$. □

The following Theorem 5.11 is the discrete version of Theorem 5.9 and gives sufficient conditions for nonnegativity of the solution X of (5.2) for a special matrix E.

Theorem 5.11 *Let* P_r *and* P_l *be the spectral projectors onto the right and left finite deflating subspaces of the matrix pair* (E, A) *defined in (1.5), where* $E = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix}$ *with* E_{11} *regular and* $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ *is partitioned accordingly. Suppose that* A_{22} *is invertible and* $A_S E_{11}^{-1}$*, where* $A_S = A_{11} - A_{12}A_{22}^{-1}A_{21}$*, is d-stable, i.e., all eigenvalues are of modulus less than one.*

Moreover, suppose that $A_{12}A_{22}^{-1} \leq 0$*. Then, for any* G *such that* $E_{11}^{-T}\tilde{G}_{11}E_{11}^{-1} \geq 0$*, where* $\tilde{G}_{11}$ *is defined as in (5.16), the unique solution* X *to Equation (5.4) satisfies* $X \geq 0$*.*

Proof. As in the proof to Theorem 5.9, we have

$$PEQ = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \text{ and } PAQ = \begin{bmatrix} A_S E_{11}^{-1} & 0 \\ 0 & A_{22} \end{bmatrix},$$

where $P = \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix}$ and $Q = \begin{bmatrix} E_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}E_{11}^{-1} & I \end{bmatrix}$ are regular matrices. Therefore, we have the same projectors P_r and P_l as in (5.17).

Scaling the first equation in (5.4) with Q^T and Q we obtain the equivalent equation

$$Q^T A^T X A Q - Q^T E^T X E Q = -Q^T P_r^T G P_r Q.$$

As in the continuous-time case, we have $EQ = \tilde{E}$ and $AQ = \tilde{A}$, where $\tilde{E}, \tilde{A}$ are defined as in (5.18) and (5.19), respectively.

Since

$$P_r Q = \begin{bmatrix} I & 0 \\ -A_{22}^{-1} A_{21} & 0 \end{bmatrix} \begin{bmatrix} E_{11}^{-1} & 0 \\ -A_{22}^{-1} A_{21} E_{11}^{-1} & I \end{bmatrix} = \begin{bmatrix} E_{11}^{-1} & 0 \\ -A_{22}^{-1} A_{21} E_{11}^{-1} & 0 \end{bmatrix},$$

we obtain

$$\tilde{A}^T \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \tilde{A} - \tilde{E}^T \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \tilde{E} = \begin{bmatrix} -E_{11}^{-T} \tilde{G}_{11} E_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix},$$

where $\tilde{G}_{11} = (G_{11} - A_{21}^T A_{22}^{-T} G_{21} - (G_{12} - A_{21}^T A_{22}^{-T} G_{22}) A_{22}^{-1} A_{21})$. Hence, we have the following decoupled system of equations

$$\begin{cases} E_{11}^{-T} A_S^T X_{11} A_S E_{11}^{-1} - X_{11} &= -E_{11}^{-T} \tilde{G}_{11} E_{11}^{-1} \\ E_{11}^{-T} A_S^T X_{11} A_{12} + E_{11}^{-T} A_S^T X_{12} A_{22} &= 0 \\ A_{12}^T X_{11} A_S E_{11}^{-1} + A_{22}^T X_{21} A_S E_{11}^{-1} &= 0 \\ A_{12}^T X_{11} A_{12} + A_{12}^T X_{12} A_{22} + A_{22}^T X_{21} A_{12} + A_{22}^T X_{22} A_{22} &= 0 \end{cases}.$$

Since $A_S = (A_{11} - A_{12} A_{22}^{-1} A_{21}) \geq 0$ is d-stable, i.e. all eigenvalues are of modulus less than one, and $E_{11}^{-T} \tilde{G}_{11} E_{11}^{-1} \geq 0$, we obtain from Theorem 5.5 that the first equation has a unique solution $X_{11} \geq 0$.

We have that A_{22} and A_S are invertible and $A_{12} A_{22}^{-1} \leq 0$. Hence from the second and third equations, we get

$$\begin{aligned} X_{12} &= -X_{11} A_{12} A_{22}^{-1} \geq 0, \\ X_{21} &= -A_{22}^{-T} A_{12}^T X_{11} \geq 0. \end{aligned}$$

Furthermore, since we required the condition $P_l^T X = X P_l$, i.e.,

$$\begin{bmatrix} X_{11} & X_{12} \\ -A_{22}^{-T} A_{12}^T X_{11} & -A_{22}^{-T} A_{12}^T X_{12} \end{bmatrix} = \begin{bmatrix} X_{11} & -X_{11} A_{12} A_{22}^{-1} \\ X_{21} & -X_{21} A_{12} A_{22}^{-1} \end{bmatrix},$$

we obtain

$$X_{22} = -A_{22}^{-T} A_{12}^T X_{12} = -X_{21} A_{12} A_{22}^{-1} \geq 0.$$

Therefore, we conclude that $X \geq 0$. $\square$

Summary

In this chapter we have reviewed the solvability of projected generalised Lyapunov equations for descriptor systems and the well-known sufficient conditions that guarantee doubly nonnegative solutions of Lyapunov equations in the positive standard case. We have presented a generalisation of such sufficient condition that guarantee doubly nonnegative solutions of projected generalised Lyapunov equations in the positive descriptor case. Finally, we have used the Schur complement decoupling approach presented in Section 3.3 to deduce such conditions in a special index 1 situation. All results were given in the continuous-time as well as in the discrete-time case.

Chapter 6

Positivity preserving model reduction

All exact science is dominated by the idea of approximation.
- Bertrand Russel

In this chapter, we present a model reduction technique that preserves the positivity of a system in the continuous-time as well as in the discrete-time case. In Section 6.1, we review the methods of standard balanced truncation [46] and singular perturbation balanced truncation [82].

For standard systems the proposed positivity preserving method, which we present in Section 6.2, is based on the existence of a diagonal solution of Lyapunov inequalities that are shown to be feasible. Such solutions may be obtained via LMI solution methods [20]. The reduction is then performed by standard balanced truncation or singular perturbation balanced truncation methods. It is shown that both methods preserve positivity. These results were published in [108].

Furthermore, we generalise this technique to positive descriptor systems. Here, the procedure involves the additive decomposition of the transfer function into a strictly proper and a polynomial part as in [116]. It is shown that the system matrices may also be additively decomposed according to these two parts using the spectral projector. The reduced order model is then obtained via positivity preserving reduction of the strictly proper part, where we apply the reduction technique as proposed for the standard case, whereas the polynomial part remains unchanged. We give a reduced order descriptor system and show that it is positive.

Finally, numerical examples in the continuous-time and in the discrete-time case are provided and illustrate the functionality of the proposed algorithm.

6.1 Balanced truncation

In this section we review the properties of standard balanced truncation and singular perturbation balanced truncation established in [46], [82].

Continuous-time case

Consider the standard continuous-time system (1) with $E = I$ and the transfer function $G(s) = C(sI - A)^{-1}B + D$. Assume that A is c-stable. Let $\mathcal{P}, \mathcal{Q} \succeq 0$ be the solutions of the continuous-time Lyapunov equations

$$\begin{aligned} A\mathcal{P} + \mathcal{P}A^T + BB^T &= 0, \\ A^T\mathcal{Q} + \mathcal{Q}A + C^TC &= 0. \end{aligned} \tag{6.1}$$

System (1) is said to be *balanced* if

$$\mathcal{P} = \begin{bmatrix} \Sigma & & & \\ & \Sigma_c & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \quad \mathcal{Q} = \begin{bmatrix} \Sigma & & & \\ & 0 & & \\ & & \Sigma_o & \\ & & & 0 \end{bmatrix}, \tag{6.2}$$

are partitioned accordingly with square matrices $\Sigma_c \succ 0$, $\Sigma_o \succ 0$ and

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k) \text{ for some } \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > 0. \tag{6.3}$$

The numbers $\sigma_1 \ldots \sigma_k$ are called *Hankel singular values*. Consider a partition of the balanced system

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \tag{6.4}$$

where $A_{11} \in \mathbb{R}^{\ell \times \ell}$ and either $\ell = k$ or $\ell < k$ such that $\sigma_{\ell+1} < \sigma_\ell$. The matrices B and C are partitioned accordingly. By means of balanced realisations, reduced order models

$$\begin{aligned} \dot{x}_\ell(t) &= A_\ell x_\ell(t) + B_\ell u_\ell(t), \\ y_\ell(t) &= C_\ell x_\ell(t) + D_\ell u_\ell(t), \end{aligned} \tag{6.5}$$

can now be constructed. In the method of standard balanced truncation [46], the matrices A_ℓ, B_ℓ, C_ℓ, D_ℓ are defined by

$$A_\ell = A_{11}, \quad B_\ell = B_1, \quad C_\ell = C_1, \quad D_\ell = D. \tag{6.6}$$

An alternative method for the construction of reduced order models is *singular perturbation balanced truncation* [82]. There, the reduced order model is defined by

$$\begin{aligned} A_\ell &= A_{11} - A_{12}A_{22}^{-1}A_{21}, \quad & B_\ell &= B_1 - A_{12}A_{22}^{-1}B_2, \\ C_\ell &= C_1 - C_2A_{22}^{-1}A_{21}, \quad & D_\ell &= D - C_2A_{22}^{-1}B_2. \end{aligned} \tag{6.7}$$

For the reduced order models defined by (6.6) or (6.7), we have the following result on an error bound in the H_∞-norm [82]:

Lemma 6.1 *Let* $(\,A\,,\,B\,,\,C\,,\,D\,)$ *be a realisation of* $G(s)$ *that is c-stable, balanced with* Σ *and partitioned as in (6.4) and let* $(\,A_\ell\,,\,B_\ell\,,\,C_\ell\,,\,D_\ell\,)$ *be the realisation that is either constructed by (6.6) or (6.7). Then, the system in (6.5) is balanced with* $\mathcal{P} = \mathcal{Q} = \operatorname{diag}(\sigma_1, \ldots, \sigma_\ell)$*. For the corresponding transfer function* $\tilde{G}(s) = C_\ell(sI_\ell - A_\ell)^{-1}B_\ell + D_\ell$*, we have*

$$\|G - \tilde{G}\|_{\infty,c} \leq 2 \sum_{i=\ell+1}^{k} \sigma_i.$$

The main difference between the discussed truncation methods is that standard balanced truncation is exact for $s = \infty$ meaning that $G(\infty) = \tilde{G}(\infty)$, whereas singular perturbation balanced truncation is exact at $s = 0$.

Discrete-time case

Consider a discrete-time system (2) with transfer function $G(z) = C(zI - A)^{-1}B + D$ and assume that A is d-stable. Then there exist matrices $\mathcal{P}, \mathcal{Q} \in \mathbb{R}^{n\times n}$, $\mathcal{P}, \mathcal{Q} \succeq 0$ that solve the the discrete-time Lyapunov equations

$$\begin{aligned} A\mathcal{P}A^T - \mathcal{P} + BB^T &= 0, \\ A^T\mathcal{Q}A - \mathcal{Q} + C^TC &= 0. \end{aligned} \tag{6.8}$$

In accordance with the continuous-time case, the system in (2) is called *balanced* if $\mathcal{P}$ and $\mathcal{Q}$ are defined as in (6.2), where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k$ are the *Hankel singular values*. Consider a partition of the balanced system as in (6.4) and assume that either $\ell = k$ or $\ell < k$ with $\sigma_{\ell+1} < \sigma_\ell$. Standard balanced truncation leads to a reduced order model

$$\begin{aligned} x_\ell(t+1) &= A_\ell x_\ell(t) + B_\ell u_\ell(t), \\ y_\ell(t) &= C_\ell x_\ell(t) + D_\ell u_\ell(t), \end{aligned} \tag{6.9}$$

that is constructed via (6.6). In the singular perturbation balanced truncation technique, the matrices in (6.9) are given by

$$\begin{aligned} A_\ell &= A_{11} + A_{12}(I_{n-r} - A_{22})^{-1}A_{21}, \\ B_\ell &= B_1 + A_{12}(I_{n-r} - A_{22})^{-1}B_2, \\ C_\ell &= C_1 + C_2(I_{n-r} - A_{22})^{-1}A_{21}, \\ D_\ell &= D + C_2(I_{n-r} - A_{22})^{-1}B_2. \end{aligned} \tag{6.10}$$

For the reduced order models, we have an analogous result as in Lemma 6.1 [46], [82].

Lemma 6.2 *Let* $(\, A\,,\, B\,,\, C\,,\, D\,)$ *be a realisation of* $G(z)$ *that is stable, balanced with* Σ *and partitioned as in (6.4) and let* $(\, A_\ell\,,\, B_\ell\,,\, C_\ell\,,\, D_\ell\,)$ *be the realisation that is either constructed by (6.6) or (6.10). Then, the system in (6.9) is balanced with* $\mathcal{P} = \mathcal{Q} = \operatorname{diag}(\sigma_1, \ldots, \sigma_\ell)$*. For the corresponding transfer function* $\tilde{G}(z) = C_\ell(zI_\ell - A_\ell)^{-1}B_\ell + D_\ell$*, we have*

$$\|G - \tilde{G}\|_{\infty,d} \leq 2 \sum_{i=\ell+1}^{k} \sigma_i. \tag{6.11}$$

6.2 Model reduction for standard positive systems

In this section we generalise the model reduction methods reviewed in Section 6.1 to standard positive systems. To this end, note that the results in Lemma 6.1 and Lemma 6.2 still hold if the Lyapunov equations in (6.1) and (6.8) are replaced by Lyapunov inequalities in continuous-time

$$\begin{aligned} A\mathcal{P} + \mathcal{P}A^T + BB^T &\preceq 0, \\ A^T\mathcal{Q} + \mathcal{Q}A + C^TC &\preceq 0, \end{aligned} \tag{6.12}$$

or in discrete-time

$$\begin{aligned} A\mathcal{P}A^T - \mathcal{P} + BB^T &\preceq 0, \\ A^T\mathcal{Q}A - \mathcal{Q} + C^TC &\preceq 0, \end{aligned} \tag{6.13}$$

respectively. The proofs can be found in [81] for continuous-time and in [12] for discrete-time standard balanced truncation. Note that the results for singular perturbation balanced truncation can be deduced analogously. We show that for a transfer function of any positive system there exists a positive realisation with $\mathcal{P}$ and $\mathcal{Q}$ as in (6.2) that solves the inequalities (6.12) or (6.13), respectively.

Theorem 6.3 *Consider the standard c-stable continuous-time positive system (1) with $E = I$. Then, there exists a diagonal positive definite matrix T such that the positive system $(\tilde{A}, \tilde{B}, \tilde{C}, D)$ given by*

$$\tilde{A} = T^{-1}AT, \quad \tilde{B} = T^{-1}B \quad \text{and} \quad \tilde{C} = CT, \tag{6.14}$$

is balanced in the sense that there exist matrices $\mathcal{P} \succeq 0$, $\mathcal{Q} \succeq 0$ as in (6.2) with diagonal and positive definite Σ, such that the following Lyapunov inequalities hold:

$$\begin{aligned} \tilde{A}\mathcal{P} + \mathcal{P}\tilde{A}^T + \tilde{B}\tilde{B}^T &\preceq 0, \\ \tilde{A}^T\mathcal{Q} + \mathcal{Q}\tilde{A} + \tilde{C}^T\tilde{C} &\preceq 0. \end{aligned} \tag{6.15}$$

Proof. It is well known that a $-M$-matrix is diagonally stable, i.e., there exist diagonal positive definite matrices X, Y such that

$$AX + XA^T \prec 0 \quad \text{and} \quad A^TY + YA \prec 0,$$

see, e.g. [6], [17]. In particular, there exist positive semi-definite diagonal matrices X, Y such that

$$AX + XA^T + BB^T \preceq 0, \quad A^TY + YA + C^TC \preceq 0.$$

Take a permutation matrix $\Pi \in \Pi_n$ such that

$$\Pi^T X \Pi = \begin{bmatrix} X_{11} & & & \\ & X_{22} & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \quad \Pi Y \Pi^T = \begin{bmatrix} Y_{11} & & & \\ & 0 & & \\ & & Y_{33} & \\ & & & 0 \end{bmatrix}$$

with the additional property that $X_{11} = \operatorname{diag}(x_1, \ldots, x_k)$ and $Y_{11} = \operatorname{diag}(y_1, \ldots, y_k)$ satisfy

$$x_1y_1 \geq x_2y_2 \geq \ldots \geq x_ky_k > 0.$$

Now defining $\bar{T} = \operatorname{diag}((X_{11}Y_{11}^{-1})^{\frac{1}{4}}, I, I, I)$ and $T = \Pi\bar{T}$, we have that $\mathcal{P} = T^{-1}XT^{-T}$, $\mathcal{Q} = T^TYT$ have the desired form. The transformed system is given by $(\tilde{A}, \tilde{B}, \tilde{C}, D)$ as defined in (6.14). Since $\tilde{A}$ is a $-Z$-matrix and $B, C, D \geq 0$, the transformed system is positive by Theorem 3.4. □

In the discrete-time case, we have an analogous result, which is stated without proof.

Theorem 6.4 *Consider a discrete-time positive system (A, B, C, D), i.e., $A, B, C, D \geq 0$, that is d-stable, i.e., $\rho(A) < 1$. Then, there exists a positive definite diagonal matrix T such that the system $(\tilde{A}, \tilde{B}, \tilde{C}, D)$, where $\tilde{A} = TAT^{-1}$,*

$\tilde{B} = TB$ *and* $\tilde{C} = CT^{-1}$*, is balanced in the sense that there exists a positive definite diagonal matrix* Σ *such that the following Lyapunov inequalities hold:*

$$\begin{aligned} \tilde{A}\Sigma\tilde{A}^T - \Sigma + \tilde{B}\tilde{B}^T &\preceq 0, \\ \tilde{A}^T\Sigma\tilde{A} - \Sigma + \tilde{C}^T\tilde{C} &\preceq 0. \end{aligned} \tag{6.16}$$

Theorem 6.3 and Theorem 6.4 guarantee the existence of a positive balanced realisation. Once we have a positive balanced realisation, standard balanced truncation and singular perturbation balanced truncation can be applied.

The reduced order systems are again positive, which can be verified as follows. In the continuous-time case, the reduced system defined in (6.6) is again a positive system, since $B_\ell \geq 0$, $C_\ell \geq 0$, $D_\ell \geq 0$ and A_ℓ is a $-M$-matrix as a submatrix of a $-M$-matrix. In (6.7), the $-M$-matrix property of A_ℓ is preserved, since it is a Schur complement of A [122]. Furthermore, since A_{22} is also a $-M$-matrix, we have $A_{22}^{-1} \leq 0$, and hence, $B_\ell, C_\ell, D_\ell \geq 0$.
In the discrete-time case, the reduced system defined by (6.6) is positive, since A_ℓ, B_ℓ, C_ℓ, D_ℓ are submatrices of positive matrices. Furthermore, (6.10) is also a positive system, which can be observed as follows. By the stability assumption, we have that $\rho(A_{22}) \leq \rho(A) < 1$. Hence, $I_{n-\ell} - A_{22}$ is an M-matrix and $(I_{n-\ell} - A_{22})^{-1} \geq 0$. Therefore, we obtain $A_\ell, B_\ell, C_\ell, D_\ell \geq 0$.

Note that for the computation of positive reduced order models, there is no need to compute a balanced realisation explicitely. Instead, for diagonal solutions

$$\mathcal{P} = \operatorname{diag}(p_1, \ldots, p_n) \quad \text{and} \quad \mathcal{Q} = \operatorname{diag}(q_1, \ldots, q_n)$$

of (6.1) or (6.8), indices $\{\alpha_1, \ldots, \alpha_n\}$ have to be found such that

$$p_{\alpha_1} q_{\alpha_1} \geq \ldots \geq p_{\alpha_\ell} q_{\alpha_\ell} > p_{\alpha_{\ell+1}} q_{\alpha_{\ell+1}} \geq \ldots \geq p_{\alpha_n} q_{\alpha_n}.$$

Reduced order models (6.5) (or (6.9)) can be obtained in the following way: Let $\alpha = \{\alpha, \ldots, \alpha_\ell\}$, $\beta = \{\alpha_{\ell+1}, \ldots, \alpha_n\}$ and

$$\begin{aligned} \bar{A}_{11} &= A[\alpha, \alpha], & \bar{A}_{12} &= A[\alpha, \beta], \\ \bar{A}_{21} &= A[\beta, \alpha], & \bar{A}_{22} &= A[\beta, \beta], \\ \bar{B}_1 &= B[\alpha, \{1, \ldots, p\}], & \bar{B}_2 &= B[\beta, \{1, \ldots, p\}], \\ \bar{C}_1 &= C[\{1, \ldots, q\}, \alpha], & \bar{C}_2 &= C[\{1, \ldots, q\}, \beta]. \end{aligned}$$

Then, the following properties hold:

(i) The continuous-time (discrete-time) system $(\bar{A}_{11}, \bar{B}_1, \bar{C}_1, D)$ is positive and has the same transfer function as the ℓ-th order system obtained by positive standard balanced truncation in (6.6).

(ii) The continuous-time system

$$(\,\bar{A}_{11}-\bar{A}_{12}\bar{A}_{22}^{-1}\bar{A}_{21}\,,\;\bar{B}_1-\bar{A}_{12}\bar{A}_{22}^{-1}\bar{B}_2\,,\;\bar{C}_1-\bar{C}_2\bar{A}_{22}^{-1}\bar{A}_{21}\,,\;D-\bar{C}_2\bar{A}_{22}^{-1}\bar{B}_2\,)$$

is positive and has the same transfer function as the ℓ-th order system obtained by positive singular perturbation balanced truncation in (6.7).

(iii) The discrete-time system

$$\begin{aligned}(\,\bar{A}_{11}+\bar{A}_{12}(I_{n-r}-\bar{A}_{22})^{-1}\bar{A}_{21}\,,\;\bar{B}_1+\bar{A}_{12}(I_{n-r}-\bar{A}_{22})^{-1}\bar{B}_2\,,\\ \bar{C}_1+\bar{C}_2(I_{n-r}-\bar{A}_{22})^{-1}\bar{A}_{21}\,,\;D+\bar{C}_2(I_{n-r}-\bar{A}_{22})^{-1}\bar{B}_2\,)\end{aligned}$$

is positive and has the same transfer function as the ℓ-th order system obtained by positive singular perturbation balanced truncation in (6.10).

Let us finally give a remark on the Lyapunov inequalities (6.12) and (6.13). It is clear that their solutions are not unique and one should look for solutions $\mathcal{P} = \operatorname{diag}(p_1, \ldots, p_n)$, $\mathcal{Q} = \operatorname{diag}(q_1, \ldots, q_n)$ such that $\sqrt{\mathcal{P}\mathcal{Q}}$ has a large number of small diagonal elements. This yields components of the state which are candidates to truncate. A good heuristic for this is the minimisation of the trace of $\mathcal{P}$ and $\mathcal{Q}$. For getting even sharper bounds, the Lyapunov inequalities can be solved once more while now minimising the sum of those diagonal elements of $\mathcal{P}$ and $\mathcal{Q}$ corresponding to the candidates for truncation.

6.3 Model reduction for positive descriptor systems

In the present section, we generalise the results of Section 6.2 to the descriptor case as characterised in Theorem 3.8.

From [98], we have that the transfer function $G(s) = C(sE-A)^{-1}B+D$ can be additively decomposed as $G(s) = G_{sp}(s)+P(s)$, where $G_{sp}(s)$ is the strictly proper and $P(s)$ is the polynomial part of $G(s)$. By considering the pair (E, A) in Weierstraß canonical form and the Laurent expansion at infinity, it is shown in [98] that

$$G_{sp}(s) = \sum_{k=1}^{\infty} CF_{k-1}Bs^{-k}, \quad \text{where} \quad F_k = T^{-1}\begin{bmatrix} J^k & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, \tag{6.17}$$

and

$$P(s) = \sum_{k=-\nu+1}^{0} CF_{k-1}Bs^{-k} + D, \quad \text{where} \quad F_k = T^{-1}\begin{bmatrix} 0 & 0 \\ 0 & -N^{-k-1} \end{bmatrix} W^{-1}. \tag{6.18}$$

The following Lemma 6.5 reformulates the two parts of the transfer function $G_{sp}(s)$ and $P(s)$ in terms of the matrices that appear in the solution formulas of the continuous-time and discrete-time state equations in Theorem 1.20, respectively. In particular, this shows that the system matrices can also be additively decomposed according to the two parts of the transfer function.

Lemma 6.5 *Let (E, A) be a regular matrix pair and let $\hat{E}, \hat{A}$ be defined as in Lemma 1.19 and $\hat{B}$ as in (1.12). Then, we have that the strictly proper part $G_{sp}(s)$ of the transfer function $G(s)$ can be written as*

$$G_{sp}(s) = \sum_{k=1}^{\infty}(CP_r)(\hat{E}^D\hat{A})^{k-1}(\hat{E}^D\hat{B})s^{-k} = (CP_r)(sI - \hat{E}^D\hat{A})^{-1}(\hat{E}^D\hat{B}), \tag{6.19}$$

and the polynomial part can be written as

$$\begin{aligned} P(s) =& -C(I-P_r)\sum_{k=0}^{\nu-1}(\hat{E}\hat{A}^D)^k(I-P_r)\hat{A}^D\hat{B}s^k + D = \\ =& C(I-P_r)(s(I-P_r)\hat{E}\hat{A}^D - I)^{-1}(I-P_r)\hat{A}^D\hat{B} + D. \end{aligned} \tag{6.20}$$

Proof. By using the Weierstraß canonical form for (E, A), we get the following facts, see Section 1.6. For some $\hat{\lambda}$ chosen such that $\hat{\lambda}E - A$ is nonsingular, we have

$$\begin{aligned} (\hat{\lambda}E - A)^{-1} &= T^{-1}\begin{bmatrix} (\hat{\lambda}I - J)^{-1} & 0 \\ 0 & (\hat{\lambda}N - I)^{-1} \end{bmatrix} W^{-1}, \\ \hat{E}^D &= T^{-1}\begin{bmatrix} \hat{\lambda}I - J & 0 \\ 0 & 0 \end{bmatrix} T, \\ \hat{E}^D\hat{A} &= T^{-1}\begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} T, \end{aligned}$$

where $\hat{E}$ is defined as in Lemma 1.19. Hence, from (6.17) for $k \geq 0$ we obtain

$$\begin{aligned} F_k &= T^{-1}\begin{bmatrix} J^k & 0 \\ 0 & 0 \end{bmatrix} W^{-1} = \\ &= T^{-1}\begin{bmatrix} J^k & 0 \\ 0 & 0 \end{bmatrix} TT^{-1}\begin{bmatrix} \hat{\lambda}I - J & 0 \\ 0 & 0 \end{bmatrix} TT^{-1}\begin{bmatrix} (\hat{\lambda}I - J)^{-1} & 0 \\ 0 & (\hat{\lambda}N - I)^{-1} \end{bmatrix} W^{-1} = \\ &= (\hat{E}^D\hat{A})^k\hat{E}^D(\hat{\lambda}E - A)^{-1}. \end{aligned}$$

Since $\hat{B} = (\hat{\lambda}E - A)^{-1}B$ and $\hat{E}^D = \hat{E}^D\hat{E}\hat{E}^D = P_r\hat{E}^D$, we obtain (6.19).

Analogously, by using the results of Section 1.6, i.e.,

$$\begin{aligned}\hat{E}\hat{A}^D &= T^{-1}\begin{bmatrix} J^D & 0 \\ 0 & N \end{bmatrix} T^{-1}, \\ \hat{A}^D &= T^{-1}\begin{bmatrix} J^D(\hat{\lambda}I - J) & 0 \\ 0 & (\hat{\lambda}N - I) \end{bmatrix} T^{-1},\end{aligned} \tag{6.21}$$

for $k = -1, \ldots, -\nu + 1$, we have

$$F_k = T^{-1}\begin{bmatrix} 0 & 0 \\ 0 & -N^{-k-1} \end{bmatrix} W^{-1} = (I - P_r)(\hat{E}\hat{A}^D)^{-k-1}\hat{A}^D(\hat{\lambda}E - A)^{-1}.$$

By setting k to $-k$, we obtain the first equality of (6.20). The second equality follows due to the fact that $(I - P_r)\hat{E}\hat{A}^D$ is nilpotent. □

6.3.1 Continuous-time case

In this section, for positive continuous-time descriptor systems, we first prove the existence of a positive balanced realisation. Based on this, we define a reduced order system and show that it is positive and that it yields the usual H_∞ error bound in Lemma 6.1.

Consider the continuous-time Lyapunov equations corresponding to G_{sp}

$$(\hat{E}^D\hat{A})\mathcal{P}_c + \mathcal{P}_c(\hat{E}^D\hat{A})^T + \hat{E}^D\hat{B}\hat{B}^T(\hat{E}^D)^T = 0, \tag{6.22a}$$
$$(\hat{E}^D\hat{A})^T\mathcal{Q}_c + \mathcal{Q}_c(\hat{E}^D\hat{A}) + P_r^TC^TCP_r = 0. \tag{6.22b}$$

In the following, we show as in the standard case that there exist diagonal positive definite matrices that fulfil the corresponding Lyapunov inequalities. Note that if $\mathcal{P}_c = P_r\mathcal{P}_cP_r$ holds, as is assumed in [98], then the Lyapunov equations in (6.22) are equivalent to the generalised Lyapunov equations in (5.3), see also [98]. However, for diagonal positive definite solutions of the corresponding Lyapunov inequalities this will not necessarily be the case.

Theorem 6.6 *Consider the positive c-stable continuous-time system* (E, A, B, C, D). *Then, there exists a diagonal positive definite matrix* T *such that the positive system* $(\tilde{E}, \tilde{A}, \tilde{B}, \tilde{C}, D)$*, where*

$$\tilde{E} = T^{-1}ET, \quad \tilde{A} = T^{-1}AT, \quad \tilde{B} = T^{-1}B \quad \text{and} \quad \tilde{C} = CT, \tag{6.23}$$

is positive and balanced in the sense that there exists a diagonal positive definite matrix Σ such that the following Lyapunov inequalities hold

$$
\begin{aligned}
(\hat{\tilde{E}}^D\hat{\tilde{A}})\Sigma + \Sigma(\hat{\tilde{E}}^D\hat{\tilde{A}})^T + \hat{\tilde{E}}^D\hat{\tilde{B}}\hat{\tilde{B}}^T(\hat{\tilde{E}}^D)^T &\preceq 0, && \text{(6.24a)}\\
(\hat{\tilde{E}}^D\hat{\tilde{A}})^T\Sigma + \Sigma(\hat{\tilde{E}}^D\hat{\tilde{A}}) + \tilde{P}_r^T\tilde{C}^T\tilde{C}\tilde{P}_r &\preceq 0, && \text{(6.24b)}
\end{aligned}
$$

where $\hat{\tilde{E}}, \hat{\tilde{A}}, \hat{\tilde{B}}$ are obtained as in Lemma 1.19 and in (1.12) from the matrices $\tilde{E}, \tilde{A}, \tilde{B}$ and $\tilde{P}_r$ is the corresponding spectral projector as in (1.5).

Proof. From Theorem 4.15, we have that for a positive c-stable system, if $P_r \geq 0$, then there exist $X, Y \succ 0$, such that

$$
\begin{aligned}
(\hat{E}^D\hat{A})X + X(\hat{E}^D\hat{A})^T &\preceq 0,\\
(\hat{E}^D\hat{A})^TY + Y(\hat{E}^D\hat{A}) &\preceq 0,
\end{aligned}
$$

and the above inequalities are strict on S_f^{def}, i.e., for any $0 \neq v \in S_f^{def}$, we have

$$
\begin{aligned}
v^T(\hat{E}^D\hat{A})Xv + v^TX(\hat{E}^D\hat{A})^Tv &< 0,\\
v^T(\hat{E}^D\hat{A})^TYv + v^TY(\hat{E}^D\hat{A})v &< 0.
\end{aligned}
$$

Since we have $\hat{E}^D\hat{B}\hat{B}^T(\hat{E}^D)^T \geq 0$ and $P_r^TC^TCP_r \geq 0$ by assumption and since both terms are projected onto S_f^{def}, we obtain the existence of $\tilde{X}, \tilde{Y} \succ 0$ such that

$$
\begin{aligned}
(\hat{E}^D\hat{A})\tilde{X} + \tilde{X}(\hat{E}^D\hat{A})^T + \hat{E}^D\hat{B}\hat{B}^T(\hat{E}^D)^T &\preceq 0,\\
(\hat{E}^D\hat{A})^T\tilde{Y} + \tilde{Y}(\hat{E}^D\hat{A}) + P_r^TC^TCP_r &\preceq 0,
\end{aligned}
$$

by using the same argument as in the proof of Theorem 6.3. Hence, there exist diagonal positive definite matrices Σ and $\tilde{T}$ such that by setting $\tilde{E} = \tilde{T}E\tilde{T}^{-1}$, $\tilde{A} = \tilde{T}A\tilde{T}^{-1}$, $\tilde{B} = \tilde{T}B$ and $\tilde{C} = C\tilde{T}^{-1}$, we obtain

$$
\begin{aligned}
(\hat{\tilde{E}}^D\hat{\tilde{A}})\Sigma + \Sigma(\hat{\tilde{E}}^D\hat{\tilde{A}})^T + \hat{\tilde{E}}^D\hat{\tilde{B}}\hat{\tilde{B}}^T(\hat{\tilde{E}}^D)^T &\preceq 0,\\
(\hat{\tilde{E}}^D\hat{\tilde{A}})^T\Sigma + \Sigma(\hat{\tilde{E}}^D\hat{\tilde{A}}) + \tilde{P}_r^T\tilde{C}^T\tilde{C}\tilde{P}_r &\preceq 0,
\end{aligned}
$$

where $\hat{\tilde{E}}, \hat{\tilde{A}}, \hat{\tilde{B}}$ are obtained as in Lemma 1.19 and in (1.12) from the matrices $\tilde{E}, \tilde{A}, \tilde{B}$ and $\tilde{P}_r = \tilde{T}P_r\tilde{T}^{-1}$ is the corresponding spectral projector as in (1.5). Since $\tilde{T}$ is diagonal with positive diagonal entries, the transformed system is again positive. $\square$

From now on, we consider the balanced system (E, A, B, C, D) in the sense of Theorem 6.6. Scaling the state equation of the system by the regular matrix $R^{-1} = (EP_r + A(I - P_r))^{-1}$ as in Section 3.2.2, we obtain an equivalent system $(\tilde{E}, \tilde{A}, \tilde{B}, C, D)$

with system matrices $\tilde{E} := R^{-1}E$, $\tilde{A} := R^{-1}A$, $\tilde{B} := R^{-1}B$. The multiplication with the spectral projector and its complementary projector, respectively, leads to an equivalent system of two equations as in (3.12).

We now derive a procedure for computing a reduced order system that is again positive having the usual H_∞ error bound as for standard balanced truncation, see Lemma 6.1.

Consider a partitioning as in (6.4) but for the matrices $\hat{E}^D\hat{A}, \hat{E}^D\hat{B}, CP_r$:

$$\hat{E}^D\hat{A} = \begin{bmatrix} [\hat{E}^D\hat{A}]_{11} & [\hat{E}^D\hat{A}]_{12} \\ [\hat{E}^D\hat{A}]_{21} & [\hat{E}^D\hat{A}]_{22} \end{bmatrix}, \; \hat{E}^D\hat{B} = \begin{bmatrix} [\hat{E}^D\hat{B}]_1 \\ [\hat{E}^D\hat{B}]_2 \end{bmatrix}, \; CP_r = \begin{bmatrix} [CP_r]_1 & [CP_r]_2 \end{bmatrix}, \qquad (6.25)$$

where $[\hat{E}^D\hat{A}]_{11} \in \mathbb{R}^{\ell\times\ell}$ and ℓ is chosen as in the standard case in (6.4). The matrices $\hat{E}^D\hat{B}$ and CP_r are partitioned accordingly.

Our aim is to construct a reduction method that allows to obtain an H_∞ error bound as in the standard case. This is possible, for instance, if the polynomial part $P(s)$ of the transfer function $G(s)$ remains unchanged, whereas the strictly proper part $G_{sp}(s)$ is reduced as in the standard case [116]. In this case, the polynomial parts of the original and the reduced transfer functions cancel out in the H_∞ norm and we obtain the usual H_∞ error bound. Note that, since (E, A) was assumed to be c-stable, we have that $\hat{E}^D\hat{A}$ has only stable eigenvalues except for possibly several eigenvalues zero that correspond to the eigenvalue ∞ of (E, A), see Section 2.2.3. To obtain an H_∞ error bound, these must not be reduced and, hence, we have to make sure that the block $[\hat{E}^D\hat{A}]_{22}$ is regular.

We partition the spectral projector P_r and the matrices $\hat{E}\hat{A}^D, \hat{A}^D\hat{B}$ conformably with the partitioning of the matrix $\hat{E}^D\hat{A}$,

$$P_r = \begin{bmatrix} [P_r]_{11} & [P_r]_{12} \\ [P_r]_{21} & [P_r]_{22} \end{bmatrix}, \; \hat{E}\hat{A}^D = \begin{bmatrix} [\hat{E}\hat{A}^D]_{11} & [\hat{E}\hat{A}^D]_{12} \\ [\hat{E}\hat{A}^D]_{21} & [\hat{E}\hat{A}^D]_{22} \end{bmatrix}, \; \hat{A}^D\hat{B} = \begin{bmatrix} [\hat{A}^D\hat{B}]_1 \\ [\hat{A}^D\hat{B}]_2 \end{bmatrix}. \qquad (6.26)$$

The following Lemma 6.7, in particular, states that $[P_r]_{22}$ is regular whenever $[\hat{E}^D\hat{A}]_{22}$ is regular.

Lemma 6.7 *Let the matrix $\hat{E}^D\hat{A}$ and the nonnegative projector P_r be partitioned as in (6.25) and (6.26), respectively, such that $[\hat{E}^D\hat{A}]_{22}$ is regular. Then, $[P_r]_{22}$ is a (regular) diagonal matrix with positive diagonal entries.*

Proof. We have that $P_r \geq 0$ is a projector. Hence, there exists a permutation matrix $Q \in \Pi_n$ such that QP_rQ^T is in canonical form (2.21). We use the permutation matrix Q

to obtain a corresponding permutation of $\hat{E}^D\hat{A}$ and partition it accordingly

$$Q\hat{E}^D\hat{A}Q^T = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1,k+2} \\ A_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ A_{k+2,1} & \dots & & A_{k+2,k+2} \end{bmatrix}. \tag{6.27}$$

Since $P_r\hat{E}^D\hat{A} = \hat{E}^D\hat{A}$ and $\pi_{k+1,k+1} = 0$, we have that $\pi_{k+1,k+1}A_{k+1,i} = A_{k+1,i} = 0$ for $i = 1, \dots, k+2$. Furthermore, since $\hat{E}^D\hat{A}P_r = \hat{E}^D\hat{A}$ and $\pi_{k+2,k+2} = 0$, we have that $A_{i,k+2}\pi_{k+2,k+2} = A_{i,k+2} = 0$ for $i = 1, \dots, k+2$. This implies that whenever we choose a regular part of $\hat{E}^D\hat{A}$, then the corresponding part of P_r will have a positive diagonal by construction. Furthermore, since $\hat{E}^D\hat{A}P_r = \hat{E}^D\hat{A}$ we have that $A_{ii}\pi_{ii} = A_{ii}$ for $i = 1, \dots, k-1$ and hence $\operatorname{rank}(A_{ii}) \le \operatorname{rank}(\pi_{ii}) = 1$. Since $\operatorname{rank}(\hat{E}^D\hat{A}) = \operatorname{rank}(P_r)$ we conclude that $\operatorname{rank}(A_{ii}) = \operatorname{rank}(\pi_{ii}) = 1$. Hence, for a regular part of $\hat{E}^D\hat{A}$, we can pick at most one row/column from each block row/column in (6.27). By construction the corresponding part of P_r will be also regular. Moreover, this part will be diagonal with positive diagonal entries. □

Consider the decoupled system in (3.12) that is equivalent to the system in (1a). The special structure of $(I - P_r)\hat{A}^D\hat{B}$ given in (3.18) and the special structure of $(I - P_r)(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{B}$ given in (3.19) lead to the following facts. Let

$$(I - P_r)\hat{A}^D\hat{B} = \begin{bmatrix} [(I - P_r)\hat{A}^D\hat{B}]_1 \\ [(I - P_r)\hat{A}^D\hat{B}]_2 \end{bmatrix}$$

be partitioned according to $\hat{E}^D\hat{A}$ in (6.25). Since the part $[(I-P_r)\hat{A}^D\hat{B}]_2$ corresponds to the regular part P_{22}, by considering the canonical form (2.21) of P_r and the corresponding form of $(I - P_r)\hat{A}^D\hat{B}$ in (3.18), the term $[(I - P_r)\hat{A}^D\hat{B}]_2$ must be zero. By (3.19), we have that

$$(\hat{E}\hat{A}^D)^i(I - P_r)\hat{A}^D\hat{B} = \begin{bmatrix} [\hat{E}\hat{A}^D]_{11} & [\hat{E}\hat{A}^D]_{12} \\ [\hat{E}\hat{A}^D]_{21} & [\hat{E}\hat{A}^D]_{22} \end{bmatrix} \begin{bmatrix} [(I - P_r)\hat{A}^D\hat{B}]_1 \\ 0 \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix}, \tag{6.28}$$

and we conclude that $[\hat{E}\hat{A}^D]_{21}[(I - P_r)\hat{A}^D\hat{B}]_1 = 0$.

Since $(I - P_r)\hat{E}\hat{A}^D$ is nilpotent, the second equation of (3.12) has the solution

$$(I - P_r)x(t) = -\sum_{i=1}^{\nu-1}(\hat{E}\hat{A}^D)^i(I - P_r)\hat{A}^D\hat{B}u^{(i)}(t),$$

and by (6.28) is equivalent to

$$\begin{bmatrix} (I - P_{11})[\hat{E}\hat{A}^D]_{11} & 0 \\ 0 & 0 \end{bmatrix}\dot{x} = (I - P_r)x + \begin{bmatrix} [(I - P_r)\hat{A}^D\hat{B}]_1 \\ 0 \end{bmatrix}u. \tag{6.29}$$

The system in (3.12) is therefore equivalent to the following decoupled system of two equations

$$\begin{cases} \dot{x}_f = & A_f x_f + B_f u \\ E_\infty \dot{x}_\infty = & x_\infty + B_\infty u \end{cases}, \tag{6.30}$$

where

$$\begin{aligned} x_f :=& P_r x, & A_f :=& \hat{E}^D \hat{A}, & B_f :=& \hat{E}^D \hat{B}, \\ x_\infty :=& (I-P_r)x, & E_\infty :=& \begin{bmatrix} (I-P_{11})[\hat{E}\hat{A}^D]_{11} & 0 \\ 0 & 0 \end{bmatrix}, & B_\infty :=& \begin{bmatrix} [(I-P_r)\hat{A}^D \hat{B}]_1 \\ 0 \end{bmatrix}. \end{aligned} \tag{6.31}$$

Furthermore, we set

$$C_f := CP_r, \quad C_\infty := C(I-P_r). \tag{6.32}$$

Consider again the transfer function $G(s)$. By using (3.12) and Lemma 6.5, we can additively decompose $G(s)$ as

$$\begin{aligned} G(s) =& C(sE-A)^{-1}B + D = C(sR^{-1}E - R^{-1}A)^{-1}R^{-1}B + D = \\ =& (CP_r + C(I-P_r))\left\{ s(P_r + (I-P_r)\hat{E}\hat{A}^D) - (\hat{E}^D\hat{A} + (I-P_r)) \right\}^{-1} \cdot \\ & \cdot (\hat{E}^D\hat{B} + (I-P_r)\hat{A}^D\hat{B}) + D = \\ =& G_{sp}(s) + P(s), \end{aligned}$$

where

$$\begin{aligned} G_{sp}(s) =& (CP_r)(sI - \hat{E}^D\hat{A})^{-1}(\hat{E}^D\hat{B}) = C_f(sI - A_f)^{-1}B_f \\ P(s) =& C(I-P_r)(s(I-P_r)\hat{E}\hat{A}^D - I)^{-1}(I-P_r)\hat{A}^D\hat{B} + D = \\ =& C_\infty(sE_\infty - I)^{-1}B_\infty + D. \end{aligned}$$

The first equation of the system in (6.30) is a standard system on the subspace $\operatorname{im} P_r$. Requiring that $[\hat{E}^D\hat{A}]_{22} = [A_f]_{22}$ is regular, we can apply the reduction scheme in (6.7) to the system $(\, A_f,\, B_f,\, C_f,\, 0\,)$. We obtain the reduced order system

$$(A_{f\ell}, B_{f\ell}, C_{f\ell}, D_{f\ell}) \quad \text{with} \quad \tilde{G}_{sp}(s) = C_{f\ell}(sI - A_{f\ell})^{-1}B_{f\ell} + D_{f\ell}.$$

Since A_f is c-stable on $\operatorname{im} P_r$ and $(\, A_f,\, B_f,\, C_f,\, 0\,)$ is balanced on $\operatorname{im} P_r$, we have that $\tilde{G}_{sp}(s)$ yields the error bound in Lemma 6.1.

The second equation of (6.30) corresponds to the system $(\, E_\infty,\, B_\infty,\, C_\infty,\, D\,)$. Assuming the same partitioning as for $(\, A_f,\, B_f,\, C_f,\, 0\,)$, we can apply the standard balanced truncation reduction scheme in (6.6). We obtain the reduced order system

$$(\, E_{\infty\ell},\, B_{\infty\ell},\, C_{\infty\ell},\, D\,) \quad \text{with} \quad \tilde{P}(s) = C_{\infty\ell}(sE_{\infty\ell} - I)^{-1}B_{\infty\ell} + D = P(s).$$

That $\tilde{P}(s) = P(s)$ can be verified as follows. By Lemma 6.5 we have that

$$P(s) = -C(I - P_r)\sum_{k=0}^{\nu-1}(\hat{E}\hat{A}^D)^k(I - P_r)\hat{A}^D\hat{B}s^k + D.$$

On the other hand, by (3.19) we know that if P_r is in canonical form and the matrices $\hat{E}\hat{A}^D, \hat{A}^D\hat{B}$ are permuted and partitioned accordingly, we have that

$$(\hat{E}\hat{A}^D)^k(I - P_r)\hat{A}^D\hat{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ * \\ * \end{bmatrix}.$$

Therefore, if we consider a partitioning as in (6.29), we obtain that

$$P(s) = -\begin{bmatrix} C_{\infty_1} & C_{\infty_2} \end{bmatrix}\begin{bmatrix} [E_\infty B_\infty]_1 \\ 0 \end{bmatrix} = -C_{\infty_1}[E_\infty B_\infty]_1 = \tilde{P}(s).$$

Note that, in particular, this proves that Assertion (i) of Theorem 3.8 still holds for the reduced order system.

We obtain a corresponding descriptor system $(\, E_\ell,\, A_\ell,\, B_\ell,\, C_\ell,\, D_\ell\,)$ of $\operatorname{ind}(E_\ell, A_\ell) = \operatorname{ind}(E, A) = \nu$ by setting the reduced order spectral projector P_ℓ to

$$P_\ell := [P_r]_{11} - [P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21}, \tag{6.33}$$

and

$$\begin{aligned} E_\ell &:= P_\ell + E_{\infty_\ell}, \\ A_\ell &:= A_{f_\ell} + (I - P_\ell), \\ B_\ell &:= B_{f_\ell} + B_{\infty_\ell}, \\ C_\ell &:= C_{f_\ell} + C_{\infty_\ell}, \\ D_\ell &:= D_{f_\ell} + D. \end{aligned} \tag{6.34}$$

For the reduced order transfer function $\tilde{G}(s)$ we obtain

$$\tilde{G}(s) = \tilde{G}_{sp}(s) + \tilde{P}(s) = C_\ell(sE_\ell - A_\ell)^{-1}B_\ell + D_\ell,$$

and since $P(s) = \tilde{P}(s)$ we have the error bound

$$\|G - \tilde{G}\|_{\infty,c} = \|G_{sp} - \tilde{G}_{sp}\|_{\infty,c} \le 2\sum_{i=\ell+1}^{k}\sigma_i,$$

where $\sigma_{\ell+1}, \ldots, \sigma_k$ are the truncated Hankel singular values.

We still have to show that the thus obtained reduced order system in (6.34) is again positive in the sense of Theorem 3.8.

The matrix P_ℓ is again a nonnegative projector, which is proved in Section 2.4. Furthermore, the projector P_ℓ has the following properties that are essential for the positivity of the reduced order system.

Lemma 6.8 *Let P_ℓ be defined as in (6.33) and the reduced order system matrices as in (6.34). Then, the following relations hold:*

1. $P_\ell A_{f_\ell} = A_{f_\ell} P_\ell = A_{f_\ell}$;
2. $P_\ell B_{f_\ell} = B_{f_\ell}$;
3. $C_{f_\ell} P_\ell = C_{f_\ell}$.

Proof. For Relation 1. we have to take into account that $A_f = \hat{E}^D \hat{A}$ and use the relations for the partitioned block matrices that arise from the property $P_r \hat{E}^D \hat{A} = \hat{E}^D \hat{A} P_r = \hat{E}^D \hat{A}$, i.e.,

$$\begin{aligned}\begin{bmatrix}[A_f]_{11} & [A_f]_{12}\\ [A_f]_{21} & [A_f]_{22}\end{bmatrix} &= \begin{bmatrix}[P_r]_{11}[A_f]_{11} + [P_r]_{12}[A_f]_{21} & [P_r]_{11}[A_f]_{12} + [P_r]_{12}[A_f]_{22}\\ [P_r]_{21}[A_f]_{11} + [P_r]_{22}[A_f]_{21} & [P_r]_{21}[A_f]_{12} + [P_r]_{22}[A_f]_{22}\end{bmatrix}\\ &= \begin{bmatrix}[A_f]_{11}[P_r]_{11} + [A_f]_{12}[P_r]_{21} & [A_f]_{11}[P_r]_{12} + [A_f]_{12}[P_r]_{22}\\ [A_f]_{21}[P_r]_{11} + [A_f]_{22}[P_r]_{21} & [A_f]_{21}[P_r]_{12} + [A_f]_{22}[P_r]_{22}\end{bmatrix}.\end{aligned} \tag{6.35}$$

Exemplarily, we prove the relation $P_\ell A_{f_\ell} = A_{f_\ell}$, since the proof of the relation $A_{f_\ell} P_\ell = A_{f_\ell}$ is completely analogous. We have

$$\begin{aligned}P_\ell A_{f_\ell} =& ([P_r]_{11} - [P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21})([A_f]_{11} - [A_f]_{12}[A_f]_{22}^{-1}[A_f]_{21})\\ =& [P_r]_{11}[A_f]_{11} - [P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21}[A_f]_{11} - [P_r]_{11}[A_f]_{12}[A_f]_{22}^{-1}[A_f]_{21}\\ &+ [P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21}[A_f]_{12}[A_f]_{22}^{-1}[A_f]_{21},\end{aligned}$$

where plugging in the relations

$$\begin{aligned}[P_r]_{21}[A_f]_{11} =& [A_f]_{21} - [P_r]_{22}[A_f]_{21},\\ [P_r]_{11}[A_f]_{12} =& [A_f]_{12} - [P_r]_{12}[A_f]_{22},\\ [P_r]_{21}[A_f]_{12} =& [A_f]_{22} - [P_r]_{22}[A_f]_{22},\end{aligned} \tag{6.36}$$

we obtain

$$\begin{aligned}P_\ell A_{f_\ell} =& [P_r]_{11}[A_f]_{11} - [P_r]_{12}[P_r]_{22}^{-1}[A_f]_{21} + [P_r]_{12}[A_f]_{21} - [A_f]_{12}[A_f]_{22}^{-1}[A_f]_{21} + [P_r]_{12}[A_f]_{21}\\ &+ [P_r]_{12}[P_r]_{22}^{-1}[A_f]_{21} - [P_r]_{12}[A_f]_{21}\\ =& A_{f_\ell},\end{aligned}$$

where for the last equality we have used that $[P_r]_{11}[A_f]_{11} + [P_r]_{12}[A_f]_{21} = [A_f]_{11}$.

Similarly, Relation 2. follows by direct calculation, recalling that $B_{f_\ell} = \hat{E}^D\hat{B}$ and taking into account the block relations that arise from $P_r\hat{E}^D\hat{B} = \hat{E}^D\hat{B}$, i.e.,

$$\begin{bmatrix}[B_f]_1\\ [B_f]_2\end{bmatrix} = \begin{bmatrix}[P_r]_{11}[B_f]_1 + [P_r]_{12}[B_f]_2\\ [P_r]_{21}[B_f]_1 + [P_r]_{22}[B_f]_2\end{bmatrix}. \tag{6.37}$$

We obtain

$$\begin{aligned}P_\ell B_{f_\ell} =&([P_r]_{11} - [P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21})([B_f]_1 - [A_f]_{12}[A_f]_{22}^{-1}[B_f]_2)\\ =&[P_r]_{11}[B_f]_1 - [P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21}[B_f]_1 - [P_r]_{11}[A_f]_{12}[A_f]_{22}^{-1}[B_f]_2\\ &+ [P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21}[A_f]_{12}[A_f]_{22}^{-1}[B_f]_2,\end{aligned}$$

where plugging in the last two relations of (6.36) and $[P_r]_{21}[B_f]_1 = [B_f]_2 - [P_r]_{22}[B_f]_2$, we obtain

$$\begin{aligned}P_\ell B_{f_\ell} =&[P_r]_{11}[B_f]_1 - [P_r]_{12}[P_r]_{22}^{-1}[B_f]_2 + [P_r]_{12}[B_f]_2 - [A_f]_{12}[A_f]_{22}^{-1}[B_f]_2 + [P_r]_{12}[B_f]_2\\ &+ [P_r]_{12}[P_r]_{22}^{-1}[B_f]_2 - [P_r]_{12}[B_f]_2\\ =&B_{f_\ell}.\end{aligned}$$

Here, for the last equality, we have used the first relation of (6.37).

For Relation 3., note that $C_f = CP_r$, i.e.,

$$C_f = \begin{bmatrix}C_1[P_r]_{11} + C_2[P_r]_{21} & C_1[P_r]_{12} + C_2[P_r]_{22}\end{bmatrix}. \tag{6.38}$$

Therefore, we need the block relations that arise from $P_r = P_r^2$, i.e.,

$$\begin{bmatrix}[P_r]_{11} & [P_r]_{12}\\ [P_r]_{21} & [P_r]_{22}\end{bmatrix} = \begin{bmatrix}[P_r]_{11}[P_r]_{11} + [P_r]_{12}[P_r]_{21} & [P_r]_{11}[P_r]_{12} + [P_r]_{12}[P_r]_{22}\\ [P_r]_{21}[P_r]_{11} + [P_r]_{22}[P_r]_{21} & [P_r]_{21}[P_r]_{12} + [P_r]_{22}[P_r]_{22}\end{bmatrix}. \tag{6.39}$$

By using (6.38), we obtain

$$\begin{aligned}C_{f_\ell}P_\ell =&([C_f]_1 - [C_f]_2[A_f]_{22}^{-1}[A_f]_{21})([P_r]_{11} - [P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21})\\ =&[C_f]_1[P_r]_{11} - [C_f]_1[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21} - [C_f]_2[A_f]_{22}^{-1}[A_f]_{21}[P_r]_{11}\\ &+ [C_f]_2[A_f]_{22}^{-1}[A_f]_{21}[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21}\\ =&C_1[P_r]_{11}[P_r]_{11} + C_2[P_r]_{21}[P_r]_{11} - C_1[P_r]_{11}[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21}\\ &- C_2[P_r]_{21}[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21} - C_1[P_r]_{12}[A_f]_{22}^{-1}[A_f]_{21}[P_r]_{11}\\ &- C_2[P_r]_{22}[A_f]_{22}^{-1}[A_f]_{21}[P_r]_{11} + C_1[P_r]_{12}[A_f]_{22}^{-1}[A_f]_{21}[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21}\\ &+ C_2[P_r]_{22}[A_f]_{22}^{-1}[A_f]_{21}[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21},\end{aligned}$$

where plugging in the first and the last relations of (6.36) and using the relations in (6.39), we obtain

$$\begin{aligned}
C_{f_\ell}P_\ell =& C_1[P_r]_{11}[P_r]_{11} + C_2[P_r]_{21}[P_r]_{11} - C_1[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21} + C_1[P_r]_{12}[P_r]_{21} \\
&- C_2[P_r]_{21} + C_2[P_r]_{22}[P_r]_{21} - C_1[P_r]_{12}[A_f]_{22}^{-1}[A_f]_{21} + C_1[P_r]_{12}[P_r]_{21} \\
&- C_2[P_r]_{22}[A_f]_{22}^{-1}[A_f]_{21} + C_2[P_r]_{22}[P_r]_{21} + C_1[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21} - C_1[P_r]_{12}[P_r]_{21} \\
&+ C_2[P_r]_{21} - C_2[P_r]_{22}[P_r]_{21} \\
=& C_1[P_r]_{11} + C_2[P_r]_{21} - (C_1[P_r]_{12} + C_2[P_r]_{22})[A_f]_{22}^{-1}[A_f]_{21} \\
=& C_{f_\ell}.
\end{aligned}$$

□

By Lemma 6.8 we have that the matrices E_ℓ and A_ℓ commute, since $P_\ell E_{\infty_\ell} = E_{\infty_\ell} P_\ell = 0$, which follows from $E_\infty = (I - P_r)\hat{E}\hat{A}^D$ and $P_r E_\infty = 0$. Therefore, by using the properties of P_ℓ we have that

$$E_\ell \dot{\tilde{x}} = A_\ell \tilde{x} + B_\ell u,$$

is equivalent to the decoupled system

$$\begin{cases} P_\ell \dot{\tilde{x}} &= A_{f_\ell}\tilde{x} + B_{f_\ell}u \\ E_{\infty_\ell}\dot{\tilde{x}} &= (I - P_\ell)\tilde{x} + B_{\infty_\ell}u \end{cases}.$$

We have already shown that Assumptions (i)-(ii) of Theorem 3.8 hold for the reduced order system. It remains to show Relations 1.-3. in Theorem 3.8.

To this end, note that by Lemma 6.7, $[P_r]_{22}$ is a diagonal matrix with a strictly positive diagonal. Hence, from the relation

$$A_f + \alpha P_r = \hat{E}^D \hat{A} + \alpha P_r \geq 0$$

we conclude that $[A_f]_{22}$ must be a $-Z$-matrix. Since $A_f = \hat{E}^D\hat{A}$ has only stable eigenvalues except for the eigenvalue 0 that corresponds to the eigenvalue ∞ of (E, A) and since $[A_f]_{22}$ is regular, it must be c-stable and therefore, a $-M$-matrix and we have $[A_f]_{22}^{-1} \leq 0$. By using Lemma 6.8 and the relations $P_\ell[P_r]_{11} = [P_r]_{11}P_\ell = P_\ell$ and $P_\ell[P_r]_{12} = [P_r]_{21}P_\ell = 0$ that can be verified by direct calculation, we obtain

$$\begin{aligned}
A_{f_\ell} + \alpha P_\ell =& P_\ell(A_{f_\ell} + \alpha P_\ell)P_\ell \\
=& P_\ell([A_f]_{11} + \alpha[P_r]_{11} - [A_f]_{12}[A_f]_{22}^{-1}[A_f]_{21} - \alpha[P_r]_{12}[P_r]_{22}^{-1}[P_r]_{21})P_\ell \\
=& P_\ell([A_f]_{11} + \alpha[P_r]_{11} - ([A_f]_{12} + \alpha[P_r]_{12})[A_f]_{22}^{-1}([A_f]_{21} + \alpha[P_r]_{21}))P_\ell \geq 0,
\end{aligned}$$

since $A_f + \alpha P_r \geq 0$.

Next, we show that $B_{f_\ell} \geq 0$. By Lemma 6.8 we know that $P_\ell B_{f_\ell} = B_{f_\ell}$. Then, by using $P_\ell [P_r]_{12} = 0$, we obtain

$$B_{f_\ell} = P_\ell([B_f]_1 - [A_f]_{12}[A_f]_{22}^{-1}[B_f]_2) = P_\ell([B_f]_1 - ([A_f]_{12} + \alpha[P_r]_{12})[A_f]_{22}^{-1}[B_f]_2) \geq 0.$$

Similarly, by Lemma 6.8 and since $[P_r]_{21}P_\ell = 0$ we have that

$$C_{f_\ell} = ([C_f]_1 - X[C_f]_2[A_f]_{22}^{-1}[A_f]_{21})P_\ell = ([C_f]_1 - X[C_f]_2[A_f]_{22}^{-1}([A_f]_{21} + \alpha[P_r]_{21}))P_\ell \geq 0.$$

Finally, $D_\ell = D_{f_\ell} + D \geq 0$ holds, since $D \geq 0$ remains unchanged and

$$D_{f_\ell} = -[C_f]_2[A_f]_{22}^{-1}[B_f]_2 \geq 0.$$

We have shown that the reduced order system as defined in 6.34 is again positive. The strictly proper part $G_{sp}(s)$ of the transfer function $G(s)$ is reduced as in the standard case, whereas the polynomial part $P(s)$ remains unchanged, which leads to the usual H_∞ error bound as for standard balanced truncation in Lemma 6.1.

6.3.2 Discrete-time case

Consider the discrete-time Lyapunov equations corresponding to G_{sp}

$$\begin{aligned}(\hat{E}^D\hat{A})\mathcal{P}_c(\hat{E}^D\hat{A})^T - \mathcal{P}_c + \hat{E}^D\hat{B}\hat{B}^T(\hat{E}^D)^T &= 0,\\ (\hat{E}^D\hat{A})^T\mathcal{Q}_c(\hat{E}^D\hat{A}) - \mathcal{Q}_c + P_r^TC^TCP_r &= 0.\end{aligned}$$

For a positive discrete-time system, by Theorem 3.8, if $P_r \geq 0$ then we have $\hat{E}^D\hat{A} \geq 0$. Since (E, A) is d-stable, we also have that $\rho(\hat{E}^D\hat{A}) < 1$. Hence, as in the standard discrete-time case in Theorem 6.4, there exists a balanced positive realisation that fulfils the corresponding Lyapunov inequalities.

From now on, we consider the balanced system $(\,E,\, A,\, B,\, C,\, D\,)$. Scaling the system as in the continuous-time case by $R^{-1} = (EP_r + A(I - P_r))^{-1}$ and multiplying the state equation with P_r and $(I - P_r)$, respectively leads to the equivalent system of two equations

$$\left\{\begin{aligned} P_r x(t+1) &= \hat{E}^D\hat{A}x(t) + \hat{E}^D\hat{B}u(t)\\ (I-P_r)\hat{E}\hat{A}^D x(t+1) &= (I-P_r)x(t) + (I-P_r)\hat{A}^D\hat{B}u(t)\end{aligned}\right. . \tag{6.40}$$

By using the conditions on $(I - P_r)\hat{A}^D\hat{B}$ and $(I - P_r)(\hat{E}\hat{A}^D)\hat{A}^D\hat{B}$ derived in Section 6.2 for the continuous-time case, we deduce that the system in (6.40) is equivalent to

$$\left\{\begin{aligned} x_f(t+1) &= A_f x_f(t) + B_f u(t)\\ E_\infty x_\infty(t+1) &= x_\infty(t) + B_\infty u(t)\end{aligned}\right. , \tag{6.41}$$

where the systems $(\, A_f,\, B_f,\, C_f,\, 0\,)$ for $G_{sp}(s)$ and $(\, E_\infty,\, B_\infty,\, C_\infty,\, D\,)$ for $P(s)$ are given by the matrices in (6.31) and (6.32).

As in the continuous-time case, we reduce the strictly proper part $G_{sp}(s)$ of $G(s)$ using standard singular perturbation balanced truncation, whereas the polynomial part $P(s)$ remains unchanged. We show that we obtain a reduced positive descriptor system that approximates the original system with the usual H_∞ error bound in Lemma 6.2.

Consider a partitioning as in (6.25) and (6.26). As in the continuous-time case we choose the block $[A_f]_{22}$ regular and therefore, $\rho([A_f]_{22}) < 1$ and we have that $(I - [A_f]_{22})$ is an M-matrix with $(I - [A_f]_{22}) \geq 0$.

The first equation of the system in (6.41) is a standard system on the subspace $\operatorname{im} P_r$. If the block $[P_r]_{22}$ contains ones on the diagonal, we first apply the balanced truncation scheme in (6.6) to the corresponding part of the system. The truncated projector is again a nonnegative projector and also the system is again positive. Therefore, without loss of generality, we may assume that the diagonal entries of $[P_r]_{22}$ are strictly less than 1.

We apply the reduction scheme in (6.10) and obtain a reduced order system

$$(\, A_{f_\ell},\, B_{f_\ell},\, C_{f_\ell},\, D_{f_\ell}\,) \quad \text{with} \quad G_{sp}(s) = C_{f_\ell}(sI - A_{f_\ell})^{-1}B_{f_\ell} + D_{f_\ell}.$$

Since A_f is d-stable on $\operatorname{im} P_r$ and $(\, A_f,\, B_f,\, C_f,\, 0\,)$ is balanced on $\operatorname{im} P_r$ we have that $G_{sp}(s)$ yields the error bound in Lemma 6.2.

As in the continuous-time case, we partition system $(\, E_\infty,\, B_\infty,\, C_\infty,\, D\,)$ according to $(\, A_f,\, B_f,\, C_f,\, 0\,)$ and reduce it by standard balanced truncation in (6.6). We obtain the reduced system

$$(\, E_{\infty_\ell},\, B_{\infty_\ell},\, C_{\infty_\ell},\, D\,) \quad \text{with} \quad \tilde{P}(s) = C_{\infty_\ell}(sE_{\infty_\ell} - I)^{-1}B_{\infty_\ell} + D = P(s).$$

To obtain a corresponding descriptor system we set the reduced order spectral projector to

$$P_\ell := [P_r]_{11} + [P_r]_{12}(I - [P_r]_{22})^{-1}[P_r]_{21}. \tag{6.42}$$

In Section 2.4.1 we have shown that P_ℓ is again a projector and that $P_\ell \geq 0$ if $P_r \geq 0$. The reduced order descriptor system $(\, E_\ell,\, A_\ell,\, B_\ell,\, C_\ell,\, D_\ell\,)$ is then given by the matrices in (6.34).

For the reduced order transfer function $\tilde{G}(s)$ we obtain

$$\tilde{G}(s) = \tilde{G}_{sp}(s) + \tilde{P}(s) = C_\ell(sE_\ell - A_\ell)^{-1}B_\ell + D_\ell,$$

with $\tilde{P}(s) = P(s)$ and

$$\|G - \tilde{G}\|_{\infty,d} = \|G_{sp} - \tilde{G}_{sp}\|_{\infty,d} \leq 2 \sum_{i=\ell+1}^{k} \sigma_i,$$

where $\sigma_{\ell+1}, \ldots, \sigma_k$ are the truncated Hankel singular values.

We still have to show that the thus obtained reduced order system is again positive in the sense of Theorem 3.12. As in the standard case we have that $(I - [A_f]_{22})$ is an M-matrix and hence, $A_{f_\ell}, B_{f_\ell}, C_{f_\ell}, D_{f_\ell} \geq 0$.

For the reduced order discrete-time descriptor system given by the matrices in (6.34), we can obtain an equivalent system of two decoupled equations by using the properties of the projector P_ℓ given in the following Lemma 6.9.

Lemma 6.9 *Let P_ℓ be defined as in (6.42) and A_{f_ℓ}, B_{f_ℓ} be the reduced order matrices from singular perturbation balanced truncation. Then, we have*

1. $P_\ell A_{f_\ell} = A_{f_\ell} P_\ell = A_{f_\ell}$;
2. $P_\ell B_{f_\ell} = B_{f_\ell}$.

Proof. Exemplarily, we show the relation $P_\ell A_{f_\ell} = A_{f_\ell}$. Using the relations arising from $P_r \hat{E}^D \hat{A} = \hat{E}^D \hat{A}$ in (6.35), we obtain

$$\begin{aligned}
P_\ell A_{f_\ell} =& ([P_r]_{11} + [P_r]_{12}(I - [P_r]_{22})^{-1}[P_r]_{21})([A_f]_{11} + [A_f]_{12}(I - [A_f]_{22})^{-1}[A_f]_{21}) \\
=& [P_r]_{11}[A_f]_{11} + [P_r]_{11}[A_f]_{12}(I - [A_f]_{22})^{-1}[A_f]_{21} + [P_r]_{12}(I - [P_r]_{22})^{-1}[P_r]_{21}[A_f]_{11} + \\
& + [P_r]_{12}(I - [P_r]_{22})^{-1}[P_r]_{21}[A_f]_{12}(I - [A_f]_{22})^{-1}[A_f]_{21} \\
=& [P_r]_{11}[A_f]_{11} + ([A_f]_{12} - [P_r]_{12}[A_f]_{22})(I - [A_f]_{22})^{-1}[A_f]_{21} + \\
& + [P_r]_{12}(I - [P_r]_{22})^{-1}(I - [P_r]_{22})[A_f]_{21} + \\
& + [P_r]_{12}(I - [P_r]_{22})^{-1}(I - [P_r]_{22})[A_f]_{22}(I - [A_f]_{22})^{-1}[A_f]_{21} \\
=& ([P_r]_{11}[A_f]_{11} + [P_r]_{12}[A_f]_{21}) + [A_f]_{12}(I - [A_f]_{22})^{-1}[A_f]_{21} = A_{f_\ell}.
\end{aligned}$$

The other relations follow similarly. □

By using Lemma 6.9, we obtain that E_ℓ and A_ℓ commute and the reduced order state equation

$$E_\ell \dot{\tilde{x}}(t+1) = A_\ell \tilde{x}(t) + B_\ell(u),$$

is equivalent to

$$\begin{cases} P_\ell \tilde{x}(t+1) = & A_{f_\ell} \tilde{x}(t) + B_{f_\ell} u(t) \\ E_{\infty_\ell} \tilde{x}(t+1) = & (I - P_\ell)\tilde{x}(t) + B_{\infty_\ell} u(t) \end{cases} .$$

We have shown that the discrete-time reduced order system as defined in 6.34 is again positive. The strictly proper part $G_{sp}(s)$ of the transfer function $G(s)$ is reduced as in the standard case, whereas the polynomial part $P(s)$ remains unchanged, which leads to the usual H_∞ error bound as for standard balanced truncation in Lemma 6.2.

6.4 Examples

In this section we present some numerical examples to demonstrate the properties of the discussed model reduction approaches for positive systems. The numerical tests were run in MATLAB® Version 7.4.0 on a PC with an Intel(R) Pentium(R) 4 CPU 3.20GHz processor.

Example 6.10 (Continuous-time) *Consider a system of n water reservoirs such as schematically shown in Figure 6.1. All reservoirs $R_1, \dots, R_n$ are assumed to be located on the same level. The base area of R_i and its fill level are denoted by a_i and h_i, respectively. The first reservoir R_1 has an inflow u which is the input of the system, and for each $i \in \{1, \dots, n\}$, R_i has an outflow $f_{o,i}$ through a pipe with diameter $d_{o,i}$. The output of the system is assumed to be the sum of all outflows. Furthermore, each R_i and R_j are connected by a pipe with diameter $d_{ij} = d_{ji} \geq 0$. The direct flow from R_i to R_j is denoted by f_{ij}. We assume that the flow depends linearly on the difference between the pressures on both ends. This leads to the equations*

$$f_{ij}(t) = d_{ij}^2 \cdot c \cdot (h_i(t) - h_j(t)), \quad f_{o,i}(t) = d_{o,i}^2 \cdot c \cdot (h_i(t) - h_j(t)),$$

where c is a constant that depends on the viscosity and density of the medium and gravity. The fill level of R_i thus satisfies the following differential equation

$$\dot{h}_i = \frac{c}{a_i}\left(-d_{o,i}^2 h_i(t) + \sum_{j=1}^{n} d_{ij}^2 (h_j(t) - h_i(t))\right) + \frac{1}{a_i}\delta_{1i}u(t),$$

where δ_{1i} denotes the Kronecker symbol, that is $\delta_{1i} = 1$ if $i = 1$ and zero otherwise. Then, we obtain System (1) with $D = 0$ and matrices $A = [a_{ij}]_{i,j=1,\dots,n}$, $B = [b_{i1}]_{i=1,\dots,n}$, $C = [c_{1j}]_{j=1,\dots,n}$ with $b_{i1} = \frac{\delta_{1i}}{a_1}$, $c_{i1} = c \cdot d_{o,i}^2$ and

$$a_{ij} = \frac{c}{a_i} \cdot \begin{cases} -d_{o,i}^2 - \sum_{k=1}^{n} d_{ik}^2 & i = j, \\ d_{ij}^2 & i \neq j, \end{cases}$$

where we define $d_{ii} = 0$.

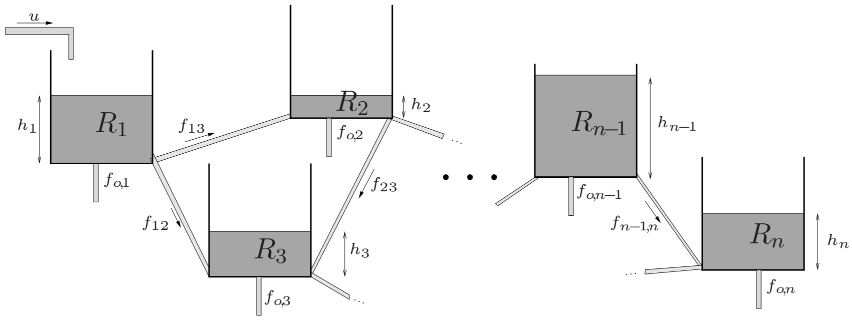


Figure 6.1: System of n water reservoirs

For our illustrative computation, we have constructed the presented compartment model with ten states. We assume that we have two well connected substructures each consisting of five reservoirs, where each reservoir is connected with every other reservoir by a pipe of diameter 1. The substructures are connected with each other by a pipe of diameter 0.01 between reservoirs one and ten. For simplicity reasons, we set all base areas of the reservoirs to 1 and also $c = 1$. The system matrices for this model are as follows,

$$A = \begin{bmatrix} -5.01 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0.01 \\ 1 & -5 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & -5 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & -5 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & -5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -5 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -5 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -5 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -5 & 1 \\ 0.01 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & -5.01 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T,$$

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

With standard balanced truncation the reduced model with five states is again positive *with*

$$A_\ell = \begin{bmatrix} -5.01 & 1.32 & 1.32 & 1.32 & 1.32 \\ 0.76 & -5.00 & 1.00 & 1.00 & 1.00 \\ 0.76 & 1.00 & -5.00 & 1.00 & 1.00 \\ 0.76 & 1.00 & 1.00 & -5.00 & 1.00 \\ 0.76 & 1.00 & 1.00 & 1.00 & -5.00 \end{bmatrix},$$

$$B_\ell = \begin{bmatrix} 0.45 & 0 & 0 & 0 & 0 \end{bmatrix}^T,$$

$$C_\ell = \begin{bmatrix} 2.20 & 2.90 & 2.90 & 2.90 & 2.90 \end{bmatrix}.$$

With singular perturbation balanced truncation the reduced model with five states is

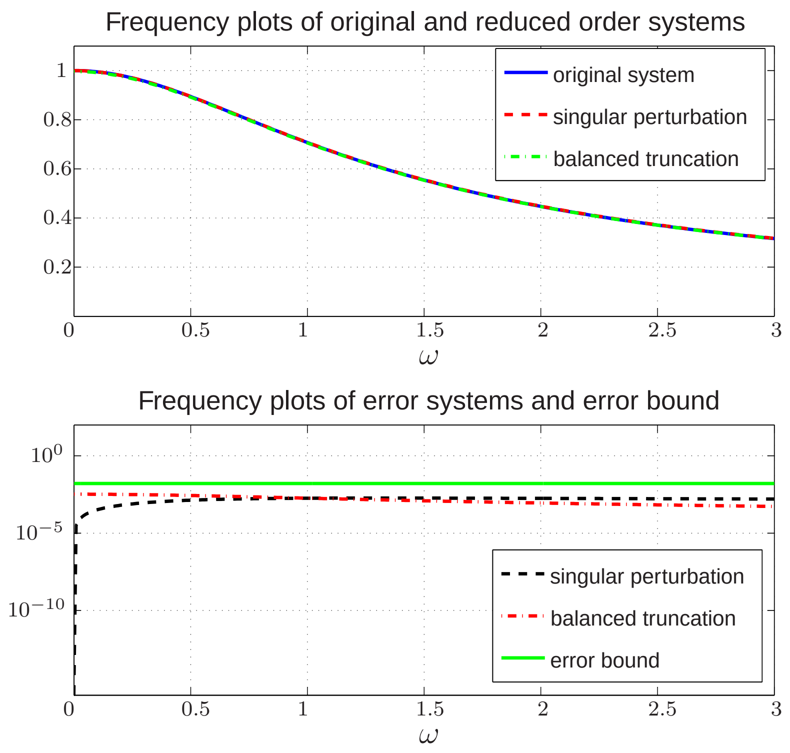


Figure 6.2: Frequency plot showing original and reduced order models.

again positive with

$$
\begin{aligned}
A_\ell &= \begin{bmatrix} -5.01 & 1.35 & 1.35 & 1.35 & 1.35 \\ 0.76 & -5.00 & 1.00 & 1.00 & 1.00 \\ 0.76 & 1.00 & -5.00 & 1.00 & 1.00 \\ 0.76 & 1.00 & 1.00 & -5.00 & 1.00 \\ 0.76 & 1.00 & 1.00 & 1.00 & -5.00 \end{bmatrix}, \\
B_\ell &= \begin{bmatrix} 0.45 & 0 & 0 & 0 & 0 \end{bmatrix}^T, \\
C_\ell &= \begin{bmatrix} 2.22 & 2.90 & 2.90 & 2.90 & 2.90 \end{bmatrix}.
\end{aligned}
$$

The frequency responses, i.e., the transfer function $G(s)$ at values $s = j\omega$, for $\omega \in [0, 3]$, of the original and of the reduced order models are depicted in the upper diagram of Figure 6.2.The lower diagram shows the frequency response of the error systems along with the mutual error bound 0.0162.

As an example in discrete-time, we consider the well-known Leslie model [80], which describes the time evolution of age-structured populations.

Example 6.11 (Discrete-time) *Let the time $t \in \mathbb{N}_0$ describe the reproduction season (year) and let $x_i(t),\ i = 1, \ldots, n$, represent the number of individuals of age i at time t. We assume constant survival rates $s_i,\ i = 0, \ldots, n-1$, i.e., the fraction of individuals of age i that survive for at least one year, and fertility rates $f_i,\ i = 1, \ldots, n$, i.e., the mean number of offspring born from an individual at age i. For purely illustrative purposes of this example, we use the estimated data given in [38, p. 118] for squirrel reproduction. Furthermore, we assume that immigration into the considered tribe can only happen at birth, i.e., the input is a positive multiple of the first unit vector, and as the output we take the total population, i.e., the sum of the population numbers over all ages. Thus, the aging process is described by the following difference equations*

$$x_{i+1}(t+1) = s_i x_i(t), \quad i = 1, \ldots, n-1, \tag{6.43}$$

and the first state equation takes into account reproduction and immigration

$$x_1(t+1) = s_0(f_1 x_1(t) + f_2 x_2(t) + \ldots + f_n x_n(t) + u(t)). \tag{6.44}$$

The system matrices for squirrel reproduction in [38, p. 118] are as follows,

$$A = \begin{bmatrix} 0.24 & 0.48 & 0.76 & 0.76 & 0.76 & 0.76 & 0.76 & 0.76 & 0.72 & 0.64 \\ 0.24 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.34 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.33 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.30 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.28 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.24 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.27 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T,$$

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

With standard balanced truncation the reduced model with five states is again positive with

$$A_\ell = \begin{bmatrix} 0.24 & 1.25 & 3.74 & 5.33 & 7.12 \\ 0.09 & 0 & 0 & 0 & 0 \\ 0 & 0.16 & 0 & 0 & 0 \\ 0 & 0 & 0.23 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0 \end{bmatrix},$$

$$B_\ell = \begin{bmatrix} 0.19 & 0 & 0 & 0 & 0 \end{bmatrix}^T,$$

$$C_\ell = \begin{bmatrix} 2.15 & 5.60 & 10.61 & 15.12 & 20.18 \end{bmatrix}.$$

With singular perturbation balanced truncation the reduced model with five states is again positive with

$$A_\ell = \begin{bmatrix} 0.24 & 1.25 & 3.74 & 5.33 & 10.43 \\ 0.09 & 0 & 0 & 0 & 0 \\ 0 & 0.16 & 0 & 0 & 0 \\ 0 & 0 & 0.23 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0 \end{bmatrix},$$

$$B_\ell = \begin{bmatrix} 0.19 & 0 & 0 & 0 & 0 \end{bmatrix}^T,$$

$$C_\ell = \begin{bmatrix} 2.15 & 5.60 & 10.61 & 15.12 & 29.56 \end{bmatrix}.$$

The frequency responses, i.e., the transfer function $G(z)$ at values $z = e^{j\omega}$, for $\omega \in [0, 2\pi]$, of the original and of the reduced order models are depicted in the upper diagram

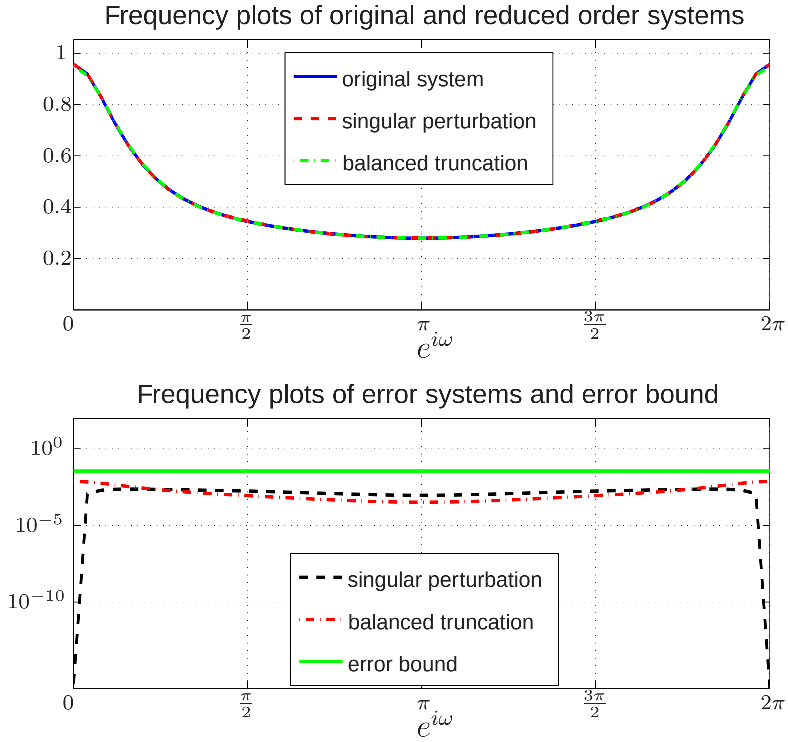


Figure 6.3: Frequency plot showing original and reduced order models.

of Figure 6.3. The lower diagram shows the frequency response of the error systems along with the mutual error bound 0.0357*.*

We now present two examples of descriptor systems in the index 1 case.

Example 6.12 (Continuous-time index 1 descriptor system) *For a purely illustrative example of a continuous-time system in the index 1 case, consider Example 6.10 and, furthermore, assume that we have an additional reservoir with fill level equal to the inflow to the system. This results in an additional equation $h_{n+1}(t) = u(t)$. The system matrices are*

$$E = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ & & & 1 & 0 \\ 0 & & \dots & & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} A_{ODE} & 0 \\ 0 & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.4 & 0 & \dots & 0 & -1 \end{bmatrix}^T, \qquad C = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix},$$

where A_{ODE} denotes the system matrix A in Example 6.10. The reduced order system obtained by the procedure as described in Section 6.3 is given by

$$E_\ell = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A_\ell = \begin{bmatrix} -5.01 & 1.32 & 1.32 & 1.32 & 1.32 & 0 \\ 0.76 & -5 & 1 & 1 & 1 & 0 \\ 0.76 & 1 & -5 & 1 & 1 & 0 \\ 0.76 & 1 & 1 & -5 & 1 & 0 \\ 0.76 & 1 & 1 & 1 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$
$$B_\ell = \begin{bmatrix} 0.45 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}^T, \qquad C_\ell = \begin{bmatrix} 2.22 & 2.90 & 2.90 & 2.90 & 2.90 & 1 \end{bmatrix}.$$

The frequency responses, i.e., the transfer function $G(s)$ at values $s = j\omega$, for $\omega \in [0, 3]$, of the original and of the reduced order models are depicted in the upper diagram of Figure 6.4. The lower diagram shows the frequency response of the error systems along with the error bound.

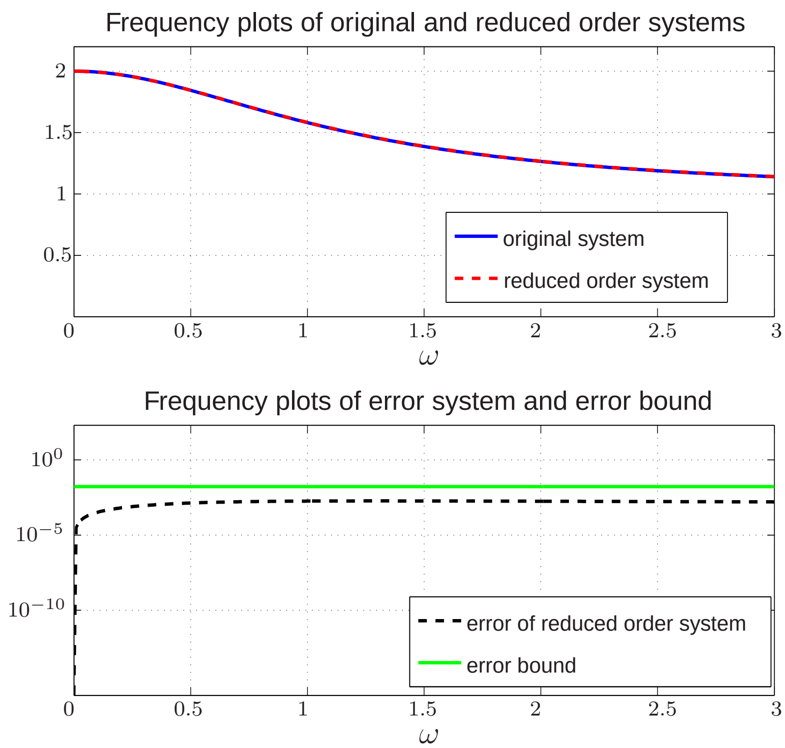


Figure 6.4: Frequency plot showing original and reduced order models.

Example 6.13 (Discrete-time index 1 descriptor system) *For an example of a discrete-time descriptor system in the index 1 case, consider Example 6.11 and additionally, assume that the number of immigrants at each time t into the considered tribe is equal to the number of those who die at the age $n+1$. This results in an additional equation $x_{n+1}(t) = u(t)$. The system matrices are*

$$E = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ & & & 1 & 0 \\ 0 & & \dots & & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} A_{ODE} & 0 \\ 0 & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.4 & 0 & \dots & 0 & -1 \end{bmatrix}^T, \qquad C = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix},$$

where A_{ODE} denotes the system matrix A in Example 6.11.

The reduced order system obtained by the procedure as described in Section 6.3 is given by

$$E_\ell = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A_\ell = \begin{bmatrix} 0.24 & 1.25 & 3.74 & 5.32 & 10.41 & 0 \\ 0.092 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.16 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.23 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$B_\ell = \begin{bmatrix} 0.19 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}^T, \qquad C_\ell = \begin{bmatrix} 2.16 & 5.61 & 10.61 & 15.1 & 29.54 & 1 \end{bmatrix}.$$

The frequency responses, i.e., the transfer function $G(z)$ at values $z = e^{j\omega}$, for $\omega \in [0, 2\pi]$, of the original and of the reduced order models are depicted in Figure 6.5. The lower diagram shows the frequency response of the error systems along with the error bound.

Summary

In this chapter we have presented a model reduction approach that preserves positivity of continuous-time as well as of discrete-time systems in the standard and in the descriptor case. In particular, we have reviewed the basic concept of standard balanced truncation and singular perturbation balanced truncation methods and extended these to preserve positivity of standard systems. The proposed approach is based

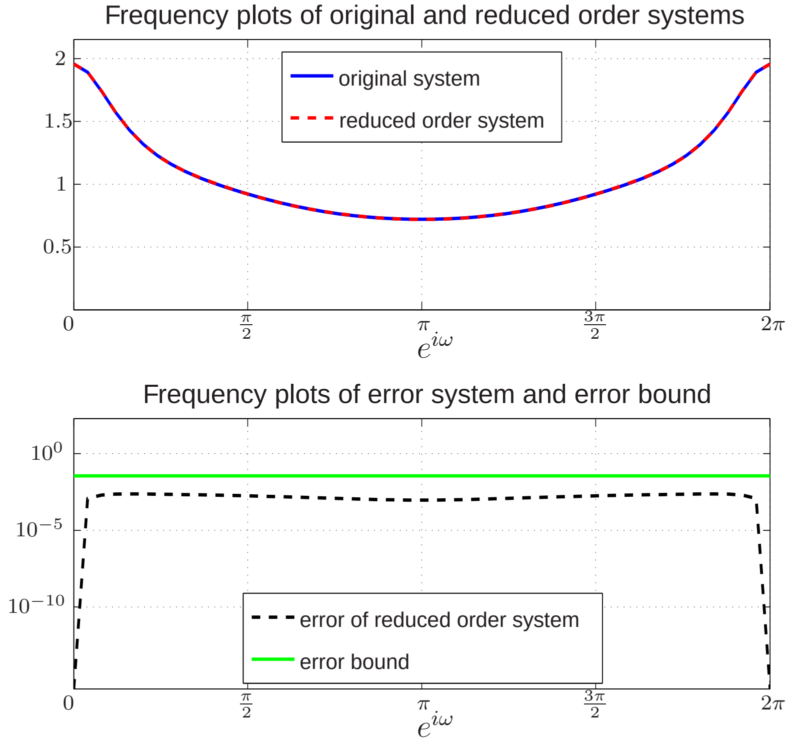


Figure 6.5: Frequency plot showing original and reduced order models.

on the existence of a diagonal solution of Lyapunov inequalities, which may be used instead of Lyapunov equations in the classical approach. In this method, along with positivity, also stability is preserved and an error bound in the H_∞ norm is provided. Furthermore, we have generalised this positivity preserving model reduction technique to positive descriptor systems. The additive decomposition of the transfer function into a strictly proper and a polynomial part allows to use the results established for the standard case. The strictly proper part may be reduced as in the standard case, whereas the polynomial part remains unchanged. This guarantees the same H_∞ error bound as in the standard case. Positivity of the reduced order system has been shown using the projector properties proved in Section 2.4, and the special properties of the system matrices established in Section 3.2.2. The functionality of the proposed method has been illustrated by some numerical examples. Note that the numerical solution of linear matrix inequalities is up to now only possible for small to medium-size problems [20]. However, further research is conducted in this area.

Chapter 7

Conclusions

I think and think for months and years. Ninety-nine times, the conclusion is false. The hundredth time I am right.
- Albert Einstein

In the present thesis we define and characterise (internally) positive descriptor systems and discuss several related topics. We establish two results in nonnegative linear algebra that are fundamental for the analysis of such systems. Furthermore, stability properties, solution of Lyapunov equations and model reduction, which constitute central topics in systems and control theory, are treated in the context of positivity. All result are shown for both, continuous-time and discrete-time systems.

In the following, we briefly summarise the novel contributions of this thesis disclosing own contributions and results that were obtained in collaboration with other authors.

Chapter 2 addresses mainly two important topics in linear algebra that shape up as key results for the analysis of positive descriptor systems. A suitable generalisation of the well-known Perron-Frobenius theory that is presented in Section 2.2 has been developed in joint work with V. Mehrmann and R. Nabben and was published in [96]. The positivity of standard, generalised and shifted Schur complements of a positive projector presented in Section 2.4 was proved in joint work with S. Friedland and was published in [42].

The generalisation of the definition of positivity and the characterisations of positive systems to the descriptor case established in Section 3.2, Chapter 3, were published in [124]. In Section 3.2.2 we analyse and specify the special structure, which the characterisation of positivity for descriptor systems in Section 3.2.1 imposes on the system matrices.

In Chapter 4 we present generalisations of (internal) stability criteria for the case of positive descriptor systems. By using the generalised Perron-Frobenius theory developed in Section 2.2, it is shown that if the spectral projector onto the finite deflating subspace of the matrix pair (E, A) is nonnegative, then all stability criteria for standard positive systems take a comparably simple form in the positive descriptor case. As an application of the framework established throughout this chapter, we show how stability criteria of switched standard positive systems can be extended to the descriptor case. The results of this chapter were published in [124].

In Chapter 5 we present sufficient conditions that guarantee doubly nonnegative solutions of projected generalised Lyapunov equations in the positive descriptor case, which were also published in [124].

Chapter 6 treats the problem of positivity preserving model reduction. The method proposed in Section 6.2 for standard systems has been developed in joint work with T. Reis and was published in [108]. It is based on the existence of a diagonal solution of Lyapunov inequalities combined with standard balanced truncation or singular perturbation balanced truncation methods. Along with positivity, also stability is preserved and an error bound in the H_∞ norm is available.

The generalisation of positivity preserving model reduction to positive descriptor systems proposed in Section 6.3 is based on the additive decomposition of the transfer function into a strictly proper and a polynomial part. The strictly proper part is reduced as in the standard case, whereas the polynomial part remains unchanged, which provides us with the same H_∞ error bound as in the standard case. Positivity of the reduced order system is shown using the projector properties proved in Section 2.4, and the special properties of the system matrices established in Section 3.2.2.

Chapter 8

Outlook and open questions

> [..] but I now see that the whole problem is so intricate that it is safer to leave its solution to the future.
> - Charles Darwin, “The Descent of Man”, 2nd edition, 1874[1]

In the scope of this thesis, we have treated a set of topics and presented several results that, so hopes the author, lay the foundation for further investigation and understanding of positive descriptor systems. However, the domain of theory, simulation and control of positive descriptor systems is still widely open. In the following, we list some related open problems and research directions, which may be of interest in the future.

Simulation

When modelling real world problems, one often encounters the situation that the systems are not easily modelled from first principles or specifications, and therefore a dynamical description of the system is not available but has to be fitted from measured input-output data. This leads to the problem of model or parameter identification. One approach that such data allows is to approximate the transfer function of the system. The problem of finding a realisation of this transfer function of minimal size is mostly understood for unconstrained systems, see, e.g., [5, Section 4.4]. However, for positive systems this is still an open problem where only sufficient existence conditions could be given so far [13]. On the other hand, there are examples of transfer functions, where one can prove that the minimal realisation size for an unconstrained system cannot be achieved when requiring that the realisation is positive [16].

[1] What the quote refers to is the sex ratio problem, whose solution is known as Fischer’s theory (1930) but which in fact was solved only a decade later by C. Düsing (1884).

Discretisation is the first step in solving systems of differential or differential-algebraic equations numerically. A suitable discretisation has to preserve important properties of the original system. Hence, for simulation of positive systems it will be fundamental to find and explore discretisation techniques that would preserve the positivity condition. In, e.g., [61], [111] some discretisation techniques are discussed. In [111] it was shown that for two point boundary value problems special inverse-isotone discretisation methods lead to a nonnegative solution. For spatial methods introduced in [61] the retention of positivity cannot be guaranteed. From these results, one can see that the task of finding the right discretisation technique for positive systems of ordinary differential equations is already problematic. For positive descriptor systems this area is still widely open.

Although there are many different concepts of an index of differential-algebraic equations, all of them aim to make a characterisation or classification of the DAE with respect to its solvability. For instance, only for the differentiation index 1 DAEs, the stability of the numerical solution can be generally guaranteed. For higher index differential-algebraic equations, it can happen that although the problem has a unique solution, the discretised equation is not stably solvable. Therefore, index reduction techniques are used to reduce DAEs of higher index to the index 1 case. Such techniques were discussed, e.g., in [57], [73], [75], [92]. The existing techniques do not take into account positivity, which would be the important property to preserve in our context.

Control

Reachability, controllability, observability and stabilisability are the core notions of the axiomatic framework in control theory. A systematic extension of this framework to the case of positive descriptor systems is therefore essential. For positive systems of ODEs, these concepts were partially studied, e.g., in [22], [30], [40], [62], [64], [109], [127], [128]. For positive systems of DAEs, this problem has not been addressed yet.

One central topic in control theory is that of finding an optimal control u in the sense that it would minimise a certain assigned cost function. Here again, the aim would be to extend the existing theory, see, e.g., [35], [53], [54], [56], [63], to positive descriptor systems. For systems of ordinary and stochastic differential equations this optimisation problem leads to the problem of solving differential or algebraic Riccati equations. Therefore, an obvious idea is to derive a corresponding equivalent Riccati formulation of the problem for positive descriptor systems and, by using results as, e.g., in [53], [56], to prove the existence of an elementwise nonnegative minimal solution.

Time-varying and nonlinear systems

In the scope of this thesis, we have studied positive descriptor systems in the linear time-invariant case. When dealing with real world applications, one recognises that most problems in biology and medicine show nonlinear behaviour [67], [104], [126]. Therefore, the study of positive nonlinear systems is indispensable.

A possible approach to deal with positive nonlinear systems is to develop positivity preserving linearisation techniques. Standard linearisation techniques do not preserve positivity in general, although the application of the classical procedures sometimes leads to the desired outcome [59]. It is possible that linearisation along constant trajectories, which results in linear time-invariant systems, may not produce a positive system or the accuracy of the approximation is not sufficient. Alternatively, the nonlinear system may be linearised along nonconstant trajectories where one obtains linear time-varying systems [28]. In this case, the concepts presented in this thesis would have to be extended to the time-varying case.

Bibliography

[1] D. Aeyels and P. De Leenheer, *Stabilization of positive linear systems*, Systems Control Lett., 44 (2001), pp. 259–271.

[2] B. D. O. Anderson, *New developments in the theory of positive systems*, in Systems and Control in the Twenty-First Century, vol. 22 of Progr. Systems Control Theory, Birkhäuser, Boston, MA, 1997, pp. 17–36.

[3] D. H. Anderson, *Compartmental Modeling and Tracer Kinetics*, vol. 50 of Lecture Notes in Biomathematics, Springer-Verlag, Berlin, 1983.

[4] S. Aniţa, *Analysis and Control of Age-Dependent Population Dynamics*, vol. 11 of Mathematical Modelling: Theory and Applications, Kluwer Academic Publishers, Dordrecht, 2000.

[5] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, vol. 6 of Advances in Design and Control, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[6] M. Araki, *Application of M-matrices to the stability problems of composite dynamical systems*, J. Math. Anal. Appl., 52 (1975), pp. 309–321.

[7] O. Arino, ed., *Mathematical Aspects of Population Dynamics*, University of Zielona Góra, Zielona Góra, 2000. Int. J. Appl. Math. Comput. Sci. **10** (2000), no. 1.

[8] J. K. Baksalary, O. M. Baksalary, and T. Szulc, *Properties of Schur complements in partitioned idempotent matrices*, Linear Algebra Appl., 379 (2004), pp. 303–318.

[9] R. B. Bapat, D. D. Olesky, and P. V. den Driessche, *Perron-Frobenius theory for a generalized eigenproblem*, Linear and Multilinear Algebra, 40 (1995), pp. 141–152.

[10] G. P. Barker, *On matrices having an invariant cone*, Czechoslovak Math. J., 22(97) (1972), pp. 49–68.

[11] G. P. Barker, A. Berman, and R. J. Plemmons, *Positive diagonal solutions to the Lyapunov equations*, Linear and Multilinear Algebra, 5 (1977/78), pp. 249–256.

[12] C. Beck, J. Doyle, and K. Glover, *Model reduction of multidimensional and uncertain systems*, IEEE Trans. Automat. Control, 41 (1996), pp. 1466–1477.

[13] L. Benvenuti, *The positive realization problem: past and future challenges*, in Positive Systems, vol. 341 of Lecture Notes in Control and Inform. Sci., Springer, Berlin, 2006, pp. 11–18.

[14] L. Benvenuti, A. De Santis, and L. Farina, eds., *Positive systems*, vol. 294 of Lecture Notes in Control and Information Sciences, Berlin, 2003, Springer-Verlag.

[15] L. Benvenuti and L. Farina, *Positive and compartmental systems*, IEEE Trans. Automat. Control, 47 (2002), pp. 370–373.

[16] ——, *A tutorial on the positive realization problem*, IEEE Trans. Automat. Control, 49 (2004), pp. 651–664.

[17] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, vol. 9 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994. Second edition.

[18] G. Birkhoff, *Linear transformations with invariant cones*, Amer. Math. Monthly, 74 (1967), pp. 274–276.

[19] G. Birkhoff and R. S. Varga, *Reactor criticality and non-negative matrices*, J. Soc. Indust. Appl. Math., 6 (1958), pp. 354–377.

[20] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15 of SIAM Studies in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[21] R. F. Brown, *Biomedical Systems Analysis via Compartmental Concept*, Cybernetics and Systems Series, Abacus Press, Tunbridge Wells, 1985.

[22] R. Bru, C. Coll, S. Romero, and E. Sánchez, *Reachability indices of positive linear systems*, Electron. J. Linear Algebra, 11 (2004), pp. 88–102.

[23] R. Bru, C. Coll, S. Romero-Vivo, and E. Sánchez, *Some problems about structural properties of positive descriptor systems*, in Positive Systems, vol. 294 of Lecture Notes in Control and Inform. Sci., Springer-Verlag, Berlin, 2003, pp. 233–240.

[24] R. Bru, C. Coll, and E. Sánchez, *About positively discrete-time singular systems*, in System and Control: Theory and Applications, Electr. Comput. Eng. Ser., World Sci. Eng. Soc. Press, Athens, 2000, pp. 44–48.

[25] R. Bru, C. Coll, and E. Sánchez, *Structural properties of positive linear time-invariant difference-algebraic equations*, Linear Algebra Appl., 349 (2002), pp. 1–10.

[26] S. L. Campbell, *Singular Systems of Differential Equations*, vol. 40 of Research Notes in Mathematics, Pitman (Advanced Publishing Program), Boston, Mass., 1980.

[27] ——, *A general form for solvable linear time varying singular systems of differential equations*, SIAM J. Math. Anal., 18 (1987), pp. 1101–1115.

[28] ——, *Linearization of DAEs along trajectories*, Z. Angew. Math. Phys., 46 (1995), pp. 70–84.

[29] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, Dover Publications Inc., New York, 1991. Corrected reprint of the 1979 original.

[30] C. Commault and M. Alamir, *On the reachability in any fixed time for positive continuous-time linear systems*, Systems Control Lett., 56 (2007), pp. 272–276.

[31] C. Commault and N. Marchand, eds., *Positive systems*, vol. 341 of Lecture Notes in Control and Information Sciences, Berlin, 2006, Springer-Verlag.

[32] P. G. Coxson, L. C. Larson, and H. Schneider, *Monomial patterns in the sequence* $A^k b$, Linear Algebra Appl., 94 (1987), pp. 89–101.

[33] P. G. Coxson and H. Shapiro, *Positive input reachability and controllability of positive systems*, Linear Algebra Appl., 94 (1987), pp. 35–53.

[34] L. DAI, *Singular Control Systems*, vol. 118 of Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 1989.

[35] T. DAMM, *Rational Matrix Equations in Stochastic Control*, vol. 297 of Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 2004.

[36] M. P. DRAZIN, *Pseudo-inverses in associative rings and semigroups*, Amer. Math. Monthly, 65 (1958), pp. 506–514.

[37] Y. EBIHARA AND T. HAGIWARA, *On H_∞ model reduction using LMIs*, IEEE Trans. Automat. Control, 49 (2004), pp. 1187–1191.

[38] L. FARINA AND S. RINALDI, *Positive Linear Systems: Theory and its Applications*, John Wiley and Sons Inc., New York, 2000.

[39] P. FLOR, *On groups of non-negative matrices*, Compositio Math., 21 (1969), pp. 376–382.

[40] E. FORNASINI AND M. E. VALCHER, *Controllability and reachability of 2-D positive systems: a graph theoretic approach*, IEEE Trans. Circuits Syst. I Regul. Pap., 52 (2005), pp. 576–585.

[41] K.-H. FÖRSTER AND B. NAGY, *Spectral properties of operator polynomials with nonnegative coefficients*, in Operator Theory and Indefinite Inner Product Spaces, vol. 163 of Oper. Theory Adv. Appl., Birkhäuser, Basel, 2006, pp. 147–162.

[42] S. FRIEDLAND AND E. VIRNIK, *Nonnegativity of Schur complements of nonnegative projectors*, Tech. Report 420, DFG Research Center MATHEON in Berlin, 2008.

[43] G. FROBENIUS, *Über Matrizen aus nicht negativen Elementen*, S.-B. Preuss. Akad. Wiss. (Berlin), (1912), pp. 456–477.

[44] Z. GAJIC AND M. T. J. QURESHI, *Lyapunov Matrix Equation in System Stability and Control*, vol. 195 of Mathematics in Science and Engineering, Academic Press Inc., San Diego, CA, 1995.

[45] F. R. GANTMACHER, *The Theory of Matrices. Vols. 1, 2*, Translated by K. A. Hirsch, Chelsea Publishing Co., New York, 1959.

[46] K. GLOVER, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds*, Internat. J. Control, 39 (1984), pp. 1115–1193.

[47] K. Godfrey, *Compartmental Models and their Application*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1983.

[48] S. K. Godunov, *Ordinary Differential Equations with Constant Coefficients*, vol. 169 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1997. Translated from the 1994 Russian original by Tamara Rozhkovskaya.

[49] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, third ed., 1996.

[50] J.-L. Gouzé, *Global behaviour of polynomial differential systems in the positive orthant*, in International Conference on Differential Equations, Vol. 1, 2 (Barcelona, 1991), World Sci. Publ., River Edge, NJ, 1993, pp. 561–567.

[51] ——, *Positivity, space scale and convergence towards the equilibrium*, J. Biol. Syst., 3 (1995), pp. 613–620.

[52] E. Griepentrog and R. März, *Differential-Algebraic Equations and their Numerical Treatment*, vol. 88 of Teubner-Texte zur Mathematik [Teubner Texts in Mathematics], BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1986.

[53] C.-H. Guo, *Iterative solution of a matrix Riccati equation arising in stochastic control*, in Linear Operators and Matrices, vol. 130 of Oper. Theory Adv. Appl., Birkhäuser, Basel, 2002, pp. 209–221.

[54] ——, *A note on the minimal nonnegative solution of a nonsymmetric algebraic Riccati equation*, Linear Algebra Appl., 357 (2002), pp. 299–302.

[55] C.-H. Guo and A. J. Laub, *On the iterative solution of a class of nonsymmetric algebraic Riccati equations*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 376–391 (electronic).

[56] X.-X. Guo, W.-W. Lin, and S.-F. Xu, *A structure-preserving doubling algorithm for nonsymetric algebraic Riccati equation*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 225–242.

[57] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations. II. Stiff and Differential-Algebraic Problems*, vol. 14 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 1996.

[58] P. R. Halmos, *Finite-Dimensional Vector Spaces*, The University Series in Undergraduate Mathematics, D. Van Nostrand Co., Inc., Princeton-Toronto-New York-London, 1958. 2nd ed.

[59] H. M. Härdin and J. H. van Schuppen, *System reduction of nonlinear positive systems by linearization and truncation*, in Positive Systems, vol. 341 of Lecture Notes in Control and Inform. Sci., Springer, Berlin, 2006, pp. 431–438.

[60] D. Hinrichsen and A. J. Pritchard, *Mathematical Systems Theory. I. Modelling, State Space Analysis, Stability and Robustness*, vol. 48 of Texts in Applied Mathematics, Springer-Verlag, Berlin, 2005.

[61] W. Hundsdorfer and J. G. Verwer, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, vol. 33 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2003.

[62] C. R. Johnson, V. Mehrmann, and D. D. Olesky, *Sign controllability of a nonnegative matrix and a positive vector*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 398–407.

[63] J. Juang and W.-W. Lin, *Existence of algebraic matrix Riccati equations arising in transport theory*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 228–243.

[64] T. Kaczorek, *Positive 1D and 2D Systems*, Springer-Verlag, London, 2002.

[65] B. Kȧgström and A. Ruhe, eds., *Matrix Pencils*, vol. 973 of Lecture Notes in Mathematics, Berlin, 1983, Springer-Verlag.

[66] E. Kaszkurewicz and A. Bhaya, *Matrix Diagonal Stability in Systems and Computation*, Birkhäuser Boston Inc., Boston, MA, 2000.

[67] E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach, *Systems Biology in Practice: Concepts, Implementation and Application*, Wiley-VCH, 2005.

[68] A. K. Konopka, ed., *Systems Biology. Principles, Methods, and Concepts*, CRC Press, Boca Raton, FL, 2007.

[69] M. A. Krasnosel′skij, J. A. Lifshits, and A. V. Sobolev, *Positive Linear Systems. The Method of Positive Operators*, vol. 5 of Sigma Series in Applied Mathematics, Heldermann Verlag, Berlin, 1989. Translated from the Russian by Jürgen Appell.

[70] U. KRAUSE, *Positive nonlinear systems in economics*, in Nonlinear and Convex Analysis in Economic Theory (Tokyo, 1993), vol. 419 of Lecture Notes in Econom. and Math. Systems, Springer, Berlin, 1995, pp. 181–195.

[71] ——, *Positive nonlinear systems: some results and applications*, in World Congress of Nonlinear Analysts ’92, Vol. I–IV (Tampa, FL, 1992), de Gruyter, Berlin, 1996, pp. 1529–1539.

[72] ——, *Positive nonlinear difference equations*, Nonlinear Anal., 30 (1997), pp. 301–308.

[73] P. KUNKEL AND V. MEHRMANN, *Index reduction for differential-algebraic equations by minimal extension*, Z. Angew. Math. Mech., 84 (2004), pp. 579–597.

[74] P. KUNKEL AND V. MEHRMANN, *Differential-Algebraic Equations. Analysis and Numerical Solution*, EMS Textbooks in Mathematics, European Mathematical Society (EMS), Zürich, 2006.

[75] P. KUNKEL, V. MEHRMANN, AND W. RATH, *Analysis and numerical solution of control problems in descriptor form*, Math. Control Signals Systems, 14 (2001), pp. 29–61.

[76] I. W. KUO, *A note on factorizations of singular M-matrices*, Linear Algebra Appl., 16 (1977), pp. 217–220.

[77] S. KUSUOKA AND T. MARUYAMA, eds., *Advances in Mathematical Economics. Vol. 5*, vol. 5 of Advances in Mathematical Economics, Springer-Verlag, Tokyo, 2003.

[78] R. LAMOUR, *Index determination and calculation of consistent initial values for DAEs*, Comput. Math. Appl., 50 (2005), pp. 1125–1140.

[79] P. LANCASTER AND M. TISMENETSKY, *The Theory of Matrices*, Computer Science and Applied Mathematics, Academic Press Inc., Orlando, FL, second ed., 1985.

[80] P. H. LESLIE, *On the use of matrices in certain population mathematics*, Biometrika, 33 (1945), pp. 183–212.

[81] L. LI AND F. PAGANINI, *Structured coprime factor model reduction based on LMIs*, Automatica J. IFAC, 41 (2005), pp. 145–151.

[82] Y. LIU AND B. D. O. ANDERSON, *Singular perturbation approximation of balanced systems*, Internat. J. Control, 50 (1989), pp. 1379–1405.

[83] D. G. LUENBERGER, *Introduction to Dynamic Systems*, John Wiley and Sons Inc., New York, 1979.

[84] A. M. LYAPUNOV, *The general problem of the stability of motion*, Internat. J. Control, 55 (1992), pp. 521–790. Translated from Edouard Davaux's French translation (1907) of the 1892 Russian original.

[85] M. S. LYNN, *On the Schur product of H-matrices and nonnegative matrices and related inequalities*, Proc. Cambridge Philo. Soc., 60 (1964), pp. 425–431.

[86] P. MACHERAS AND A. ILIADIS, *Modeling in Biopharmaceutics, Pharmacokinetics, and Pharmacodynamics. Homogeneous and Heterogeneous Approaches*, vol. 30 of Interdisciplinary Applied Mathematics, Springer, New York, 2006.

[87] O. L. MANGASARIAN, *Perron-Frobenius properties of $Ax - \lambda Bx$*, J. Math. Anal. Appl., 36 (1971), pp. 86–102.

[88] R. MÄRZ, *Numerical methods for differential algebraic equations*, in Acta Numerica, Cambridge University Press, Cambridge, 1992, pp. 141–198.

[89] ——, *Canonical projectors for linear differential algebraic equations*, Comput. Math. Appl., 31 (1996), pp. 121–135.

[90] O. MASON AND R. SHORTEN, *Some results on the stability of positive switched linear systems*, in IEEE Conference on Decision and Control, 2004.

[91] ——, *On the simultaneous diagonal stability of a pair of positive linear systems*, Linear Algebra Appl., 413 (2006), pp. 13–23.

[92] S. E. MATTSON AND G. SÖDERLIND, *Index reduction in differential-algebraic equations using dummy derivatives*, SIAM J. Sci. Comput., 14 (1993), pp. 677–692.

[93] J. J. MCDONALD, D. D. OLESKY, H. SCHNEIDER, M. J. TSATSOMEROS, AND P. VAN DEN DRIESSCHE, *Z-pencils*, Electron. J. Linear Algebra, 4 (1998), pp. 32–38 (electronic).

[94] J. J. MCDONALD AND H. SCHNEIDER, *Block LU factorizations of M-matrices*, Numer. Math., 80 (1998), pp. 109–130.

[95] V. MEHRMANN, *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*, no. 163 in Lecture Notes in Control and Information Sciences, Springer-Verlag, Heidelberg, 1991. http://springerlink.com/content/110315/.

[96] V. MEHRMANN, R. NABBEN, AND E. VIRNIK, *Generalisaton of the Perron-Frobenius theory to matrix pencils*, Linear Algebra Appl., 428 (2008), pp. 20–38.

[97] V. MEHRMANN, D. OLESKY, T. PHAN, AND P. VAN DEN DRIESSCHE, *Relations between Perron-Frobenius results for matrix pencils*, Linear Algebra Appl., 287 (1998), pp. 257–269.

[98] V. MEHRMANN AND T. STYKEL, *Balanced truncation model reduction for large-scale systems in descriptor form*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. Sorensen, eds., vol. 45 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin/Heidelberg, 2005, pp. 83–115.

[99] T. NESEMANN, *Positive nonlinear difference equations: some results and applications*, Nonlinear Anal., 47 (2001), pp. 4707–4717.

[100] J. W. NIEUWENHUIS, *When to call a linear system nonnegative*, Linear Algebra Appl., 281 (1998), pp. 43–58.

[101] G. OBINATA AND B. D. O. ANDERSON, *Model Reduction for Control System Design*, Springer, London, 2001.

[102] Y. OHTA, H. MAEDA, AND S. KODAMA, *Reachability, observability, and realizability of continuous-time positive systems*, SIAM J. Control Optim., 22 (1984), pp. 171–180.

[103] E. A. OK, *Real Analysis with Economic Applications*, Princeton University Press, Princeton, NJ, 2007.

[104] B. G. OLIVIER AND J. L. SNOEP, *Web-based kinetic modelling using JWS Online*, Bioinformatics, 20 (2004), pp. 2143–2144. http://jjj.biochem.sun.ac.za.

[105] D. H. OWENS AND D. DEBELJKOVIC, *Consistency and Liapunov stability of linear descriptor systems: A geometric analysis*, IMA J. Math. Control Inform., 2 (1985), pp. 139–151.

[106] O. PERRON, *Zur Theorie der Matrizen*, Math. Ann., 64 (1907), pp. 248–263.

[107] P. J. Psarrakos and M. J. Tsatsomeros, *A primer of Perron-Frobenius theory for matrix polynomials*, Linear Algebra Appl., 393 (2004), pp. 333–351.

[108] T. Reis and E. Virnik, *Positivity preserving balanced truncation*, Tech. Report 414, DFG Research Center Matheon in Berlin, 2007.

[109] V. Rumchev, *On reachability of positive linear discrete-time systems with scalar controls*, Decision and Control, 2000. Proceedings of the 39th IEEE Conference on, 4 (2000), pp. 3159–3162.

[110] P. Santesso and M. E. Valcher, *On the zero pattern properties and asymptotic behavior of continuous-time positive system trajectories*, Linear Algebra Appl., 425 (2007), pp. 283–302.

[111] U. Schrader, *Invers-isotone Diskretisierungsmethoden für invers-isotone lineare und quasilineare Zwei-Punkt-Randwertaufgaben*, dissertation, Technische Universität Chemnitz, Chemnitz, 2001. In German.

[112] R. N. Shorten, F. Wirth, O. Mason, K. Wulff, and C. King, *Stability theory for switched and hybrid systems*, SIAM Review, 49 (2007), pp. 545–592.

[113] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Computer Science and Scientific Computing, Academic Press Inc., Boston, 1990.

[114] T. Stykel, *On criteria for asymptotic stability of differential-algebraic equations*, ZAMM. Z. Angew. Math. Mech., 82 (2002), pp. 147–158.

[115] ——, *Stability and inertia theorems for generalized Lyapunov equations*, Linear Algebra Appl., 355 (2002), pp. 297–314.

[116] ——, *Gramian-based model reduction for descriptor systems*, Math. Control Signals Systems, 16 (2004), pp. 297–319.

[117] P. Turchin, *Complex Population Dynamics: A Theoretical/Empirical Synthesis*, vol. 35 of Monographs in Population Biology, Princeton University Press, Princeton, NJ, 2003.

[118] M. E. Valcher, *Controllability and reachability criteria for discrete time positive systems*, Internat. J. Control, 65 (1996), pp. 511–536.

[119] ——, *Nonnegative linear systems in the behavioral approach: the autonomous case*, Linear Algebra Appl., 319 (2000), pp. 147–162.

[120] M. E. Valcher and E. Fornasini, *Corrigendum: “State models and asymptotic behaviour of 2D positive systems”*, IMA J. Math. Control Inform., 12 (1995), p. 179.

[121] J. S. Vandergraft, *Spectral properties of matrices which have invariant cones*, SIAM J. Appl. Math., 16 (1968), pp. 1208–1222.

[122] R. S. Varga, *Matrix Iterative Analysis*, vol. 27 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, expanded ed., 2000.

[123] E. Virnik, *An algebraic multigrid preconditioner for a class of singular M-matrices*, SIAM J. Sci. Comput., 29 (2007), pp. 1982–1991.

[124] ——, *Stability analysis of positive descriptor systems*, Linear Algebra Appl., (2008).

[125] D. Watkins, *The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2007.

[126] G. F. Webb, *Theory of Nonlinear Age-dDependent Population Dynamics*, vol. 89 of Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker Inc., New York, 1985.

[127] H. Yoshida and T. Tanaka, *Positive controllability test for continuous-time linear systems*, IEEE Trans. Automat. Control, 52 (2007), pp. 1686–1689.

[128] H. Yoshida, T. Tanaka, and K. Yunokuchi, *Positive controllability test of discrete-time linear systems*, Trans. Soc. Instr. Contr., 34 (1998), pp. 1603–1610.

[129] F. Zhang, ed., *The Schur Complement and its Applications*, vol. 4 of Numerical Methods and Algorithms, Springer-Verlag, New York, 2005.