A tractable second-order cone certificate for external positivity with application to model order reduction

Christian Grussler and Anders Rantzer

Abstract—For linear time-invariant systems, a tractable certificate of external positivity based on second-order cones is presented. Further, we show how balanced truncation can be modified to preserve second-order cone invariance, which together with our certificate can be made into an external positivity preserving model reduction method. This method consistently yields better reduced models than approaches that intend to preserve internal positivity.

Index Terms—positive systems, external positivity, balanced truncation, model order reduction, p-dominance, semi-definite programming, second-order cones

I. INTRODUCTION

Since the emergence of the famous Perron-Frobenius theorem [1, 2], positive operators, i.e., mappings that leave a cone invariant, have attracted much interest [3–10]. For dynamical systems, the importance of cone-invariance has been early recognized by Luenberger [11], but only in the recent years received considerable attention [12–22]. On the one hand, this interest is based on the frequently appearing compartmental network structures, e.g., in bio-medicine, economics and data networks [11, 12, 23–26]. On the other hand, these systems offer a simplified analyses through their dominant dynamics [10, 13, 16–18, 27–29]. Among linear time-invariant systems

\[ \dot{x}(t) = Ax(t) + Bu(t), \]
\[ y(t) = Cx(t) + Du(t), \]

with state \( x \in \mathbb{R}^n \), input \( u \in \mathbb{R}^m \) and output \( y \in \mathbb{R}^k \), the convex cone of externally positive systems, i.e., systems that map nonnegative inputs to nonnegative outputs, are the most prominent representatives of cone-invariant systems, because many physical quantities are by definition nonnegative. For example, \( u \) could be the inflow of a substance into a chemical reactor and \( y \) the concentration of the resulting product. If in addition, the state \( x \) obeys the nonnegativity constraint, the system is usually referred to as internally positive [10–12].

Unfortunately, only for few operations, e.g., serial, parallel and positive feedback interconnections, it is easy to verify that the resulting system is externally positive. For many other operations, where external positivity may not always be preserved, this can be a difficult task. Examples include common model order reduction techniques such as balanced truncation [30], system identification [31] or the interconnection with non-positive systems as for the compound system in strong unimodality certification [19]. Thus, a certificate of external positivity that is both computationally and theoretically tractable is highly desirable.

Note that such a certificate can only be sufficient, because the problem is generally NP-hard [32]. In fact, for single-input-single-output (SISO) systems \((m = k = 1)\), external positivity is equivalent to the state remaining within some convex cone for all nonnegative inputs and \( C \) lying in the corresponding dual cone [33]. Hence, certifying externally positivity is as difficult as finding such a cone. One completely characterized approach that seeks such a convex cone is the determination of an invariant polyhedral cone leading to an internally positive realization [34–36]. A drawback of polyhedral cones, however, is the fact that this may require an arbitrarily large number of extreme rays [12, 35, 36], and as not all externally positive systems omit an internally positive realization, this procedure may not terminate.

The main idea of this work is to seek invariant second-order (ellipsoidal) cones. This has several advantages: Firstly, as the invariance of such cones has been comprehensively studied [4, 6, 37, 38], we are able to derive a simple, tractable certificate, which is representable as a semi-definite programming (SDP) and thus solvable with standard convex optimization software [39, 40]. Secondly, because of the simplicity of this certificate, it is easy to combine it with other methods that can be represented through linear matrix inequalities (LMIs). Here, we demonstrate this by modifying balanced truncation [41] to preserve second-order cone invariance and simultaneously external positivity and/or additional LMI-representable properties such as passivity [42, 43]. This is important, because it ensures that the approximated model does not violate basic physical constraints such as the nonnegativity of quantities. Numerical experiments indicate then that the error compared to standard balanced truncation is fairly small, which consistently yields better results than existing internal positivity preserving model reduction techniques [30, 44–47]. Thirdly, it is easy to construct examples, where the only possible invariant cone is of second order, thus proving the necessity for a certificate based on second-order cones.

Finally note that this paper is an extension of the authors work [48]. Due to the increased interest in second-order cone invariance and external positivity [19, 27, 31, 49–52] since then, this paper puts more focus on the certificate aspect and provides a fully elaborated investigation of our approaches, including the following aspects: (i) We discuss benefits and restrictiveness. (ii) We address the numerical tractability of our approaches and outline first order optimization methods, which makes it possible to apply our results to systems of higher order. (iii) Our modifications to balanced truncation...
are extended to generalized balanced truncation by Lyapunov inequalities [53, 54], which admits the incorporation of additional LMI-representable properties.

The paper is organized as follows. First, we introduce some basic notations and preliminaries on convex cones. Subsequently, we discuss cone-invariant systems, including positive systems. Then we are set to present and discuss our first main result, the SDP-formulation of our certificate. This result is accompanied by a discussion of its necessity, restrictiveness and computational tractability. Then, we give our second main result on the modification of balanced truncation such that our certificate remains intact. Finally, based on several examples, the success of our model reduction approaches is demonstrated and a conclusion is drawn. Proofs are stated in the appendix.

II. Preliminaries & Background

A. Notations

Throughout this paper, we use the following notations for real matrices and vectors $X = (x_{ij}) \in \mathbb{R}^{m \times n}$. The entry-wise absolute value of $X$ is given by $|X| = (|x_{ij}|)$ and the set of entry-wise nonnegative matrices by $\mathbb{R}_{\geq 0}^{m \times n}$. For nonnegative real-valued mappings $u : \mathbb{R}_{\geq 0} \to \mathbb{R}^{m}$, we employ the same notation and write $u(t) \in \mathbb{R}_{\geq 0}^{m}$.

Submatrices of $X$ are denoted by $X_{(p,q,s,t)} := (x_{ij})_{p \leq i, j \leq q}$ and accordingly $X_{(p,:)} := X_{(p,q,1:n)}$ and $X(:,:,t) := X_{(1,m,s,t)}$. $I_{n}$ stands for the identity matrix in $\mathbb{R}^{n \times n}$ and $e_{i}$ for the $i$-th canonical unit-vector in $\mathbb{R}^{n}$. For the spectrum of $X \in \mathbb{R}^{m \times n}$ we write $\sigma(X)$, whose elements $\lambda_{1}(X), \ldots, \lambda_{n}(X)$, the eigenvalues of $X$, are sorted by decreasing real part $\Re(\lambda_{i}(X))$ and subsorted by increasing imaginary part $\Im(\lambda_{i}(X))$. If $X = X^{T}$, we write $X \succ (\succeq \geq 0)$ for $X$ being positive (semi-)definite, i.e., $\sigma(X) \subset [0, \infty]$. We also use these notations to describe the relation between two matrices, e.g., $A \succeq B$ defines $A - B \succeq 0$. The inertia $i(X) = (i_{p}, i_{z}, i_{n})$ of $X$ is defined by the number of positive $i_{p}$, zero $i_{z}$ and negative $i_{n}$ real-parts of $\sigma(X)$.

For $\mathcal{F} \subset \mathbb{R}^{m}$, we denote its interior, boundary and closure by $\text{int}(\mathcal{F})$, $\partial \mathcal{F}$ and $\text{cl}(\mathcal{F})$, respectively. Further, we write $A \mathcal{F} := \{Ax : x \in \mathcal{F}\}$ for its image under an $A \in \mathbb{R}^{m \times m}$, $\text{conv}(\mathcal{F})$ and $\text{cone}(\mathcal{F})$ for its convex hull and convex conic hull. The indicator function $I_{\mathcal{F}} : \mathbb{R}^{m} \to \{0, \infty\}$ is defined as

$$I_{\mathcal{F}}(x) := \begin{cases} 0 & x \in \mathcal{F} \\ \infty & \text{else} \end{cases}$$

Finally, the $H_{m}$ norm of a transfer function $G(s)$ is denoted by $\|G\|_{H_{m}}$.

B. Polyhedral vs. second-order cones

In the following let $\mathcal{K} \subset \mathbb{R}^{n}$ be a convex cone. $\mathcal{K}$ is called solid if $\text{int}(\mathcal{K}) \neq \emptyset$ and pointed if $\mathcal{K} \cap -\mathcal{K} = \{0\}$. If it closed, solid and pointed, then $\mathcal{K}$ is referred to as proper.

Further, the corresponding dual cone and its interior are given by [5]

$$\mathcal{K}^{\ast} := \{y : y^{T}x \geq 0 \text{ for all } x \in \mathcal{K}\}. \quad (1a)$$

$$\text{int}(\mathcal{K}^{\ast}) = \{y : y^{T}x > 0 \text{ for all } x \in \text{cl}(\mathcal{K}) \setminus \{0\}\}. \quad (1b)$$

$\mathcal{K}$ is a polyhedral cone if

$$\mathcal{K} = \mathcal{D}_{N} := \mathbb{R}_{\geq 0}^{m}$$

for some $N \in \mathbb{R}^{n \times m}$ and a second-order/ellipsoidal cone if

$$\mathcal{K} = \{x : \|Px\| \leq c^{T}x\}. \quad (3)$$

for some $P \in \mathbb{R}^{m \times m}$, $c \in \mathbb{R}^{m}$ and $\| \cdot \|$ denoting the Euclidean norm. By letting $K := P^{T}P - cc^{T}$, it is easy to see that every second-order cone can alternatively be represented as

$$\mathcal{K} = \mathcal{K}_{K,c} := \{x : x^{T}Kx \leq 0, \ c^{T}x \geq 0\} \quad (4)$$

which reveals its construction by a double-cone

$$\mathcal{K} := \{x : x^{T}Kx \leq 0\} = \mathcal{K}_{K,c} \cup \mathcal{K}_{K,c} = \mathcal{K}_{K,c} \cup \mathcal{K}_{K,c}$$

that is separated through a hyperplane with normal $c$ (see Figure 1). In this work, we are mostly interested in proper cones $\mathcal{K}_{K,c}$, which implies that $t(K) = (n - 1, 0, 1)$ and $c$ is strictly separating, i.e.,

$$\{x : c^{T}x \geq 0\} \cap \mathcal{K}_{K,c} = \{0\} \quad (6)$$

Lemma 1. Let $\mathcal{K}_{K,p}$ be a proper second-order cone. Then the following are equivalent:

1. $\mathcal{K}_{K,c} = \mathcal{K}_{K,p}$
2. $\forall p^{\ast} \in \text{int}(\mathcal{K}_{K,p}) : c \in \text{int}(\mathcal{K}_{K,p}^{\ast}) = \text{int}(\mathcal{K}_{K,-1,p^{\ast}})$
3. $\exists x \in \mathcal{K}_{K,p} : c^{T}x > 0$ and $c^{T}K^{-1}c < 0$
4. $\exists x \in \mathcal{K}_{K,p}, \tau > 0 : c^{T}x > 0$ and $K + \tau cc^{T} \succ 0$

A proof of Lemma 1 is given in appendix A.

C. Cone-invariance

Definition 1 (A-invariance). Let $\mathcal{K} \subset \mathbb{R}^{n}$ and $A \in \mathbb{R}^{n \times n}$. $\mathcal{K}$ is called $A$-invariant if and only if $A\mathcal{K} \subset \mathcal{K}$. $\mathcal{K}$ is exponentially $A$-invariant if and only if $\mathcal{K}$ is $e^{At}$-invariant for all $t \geq 0$.

Remark 1. A necessary condition for the existence of a proper convex $e^{At}$-invariant cone $\mathcal{K}$ is $\lambda_{1}(A) \in \mathbb{R} \setminus \{0, \infty\}$.

By [33], a polyhedral cone $\mathcal{D}_{N}$ is exponentially $A$-invariant with $c \in \mathcal{K}_{N}^{\ast}$ if and only if

$$\exists \gamma \geq 0, P \in \mathbb{R}_{\geq 0}^{m \times m} : (A + \gamma I)N = NP, NP^{T} \in \mathbb{R}_{\geq 0}^{m}.$$ \quad (7)

A similar formulation can be derived for a proper second-order cone $\mathcal{K}_{K,c}$.

Lemma 2. Let $A \in \mathbb{R}^{n \times n}$ and $\mathcal{K}_{K,c} \subset \mathbb{R}^{n}$ be a proper second-order cone. $\mathcal{K}_{K,c}$ is exponentially $A$-invariant if and only if

$$\exists \gamma, \tau \in \mathbb{R} : A^{T}K + KA + 2\gamma K \preceq 0, K + \tau cc^{T} \succ 0$$ \quad (8)

The first part in (8) states that $\mathcal{K}_{K}$ is $e^{At}$-invariant [37], whereas the second part is from Lemma 1. The following result, which is proven in appendix B, shows that sometimes there only exist second-order $e^{At}$-invariant cones.

Lemma 3. Let $A \in \mathbb{R}^{3 \times 3}$ with $\sigma(A) = \{\alpha, \alpha \pm i\beta\}$ where $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$. Then, $\mathcal{K}$ is proper, convex $e^{At}$-invariant cone if and only if $\mathcal{K} = \mathcal{K}_{K,c}$ for some $c \in \mathbb{R}^{3}$ and $K \in \mathbb{R}^{3 \times 3}$ with $t(K) = (2, 0, 1)$. 
Remark 2. Assuming that \( \lambda_1(A) \neq \mathbb{R}(\lambda_2(A)) \), there exists both, \( e^{At} \)-invariant polyhedral \([35]\) and second-order cones. In fact, if \((A,A^T K + KA + 2\gamma K)\) is controllable, e.g., by requiring strictness in \((8)\), it follows that \((A + \gamma I)\) \([55]\). Therefore, for given \( A \) and \( c \) with \( \lambda_1(A) \neq \mathbb{R}(\lambda_2(A)) \), one only needs to solve \((8)\) for some fixed \( \gamma \in (-\mathbb{R}(\lambda_2(A)), -\lambda_1(A)) \) in order to find a solution \((K,\tau)\) with desired inertia. This can be done by semi-definite programming \([56]\).

In contrast, solving \((7)\) is significantly more involving, because even for fixed \( \gamma \), the size of \( N \) is a priori unknown and \( N \) and \( P \) are coupled in a non-convex fashion.

\section{D. Positive systems}

Next we discuss cone-invariance of linear time-invariant systems

\[
\begin{align*}
    x(t) &= Ax(t) + Bu(t), \\
    y(t) &= Cx(t) + Du(t),
\end{align*}
\]

with \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \), \( C \in \mathbb{R}^{k \times n} \) and \( D \in \mathbb{R}^{k \times m} \). In particular, we require that for all \( u(t) \in \mathbb{R}^m \) and \( x(0) \in \mathcal{K} \) it follows that \( x(t) \in \mathcal{K} \) for all \( t \geq 0 \). For convenience, we will often refer to \((A,B,C,D)\) as a system, meaning that the transfer function \( G(s) = C(sI - A)^{-1}B + D \) is realized by \((9)\). If \( D = 0 \), we also write \((A,B,C)\).

Definition 2 \((A,B)\)-invariance. Let \( \mathcal{K} \subset \mathbb{R}^n \). Then \( \mathcal{K} \) is called \((A,B)\)-invariant if \( B_i(\cdot) \in \mathcal{K} \), \( 1 \leq j \leq m \) and \( \mathcal{K} \) is exponentially A-invariant.

If \( \mathcal{K} \) is proper convex cone, then \((A,B)\)-invariance is equivalent to \( x(t) \in \mathcal{K} \) for \( t \geq 0 \), if \( u(t) \in \mathbb{R}^m \) and \( x(0) \in \mathcal{K} \). The smallest \((A,B)\)-invariant proper convex cone is then give by

\[
\mathcal{K}_e(A,B) := \text{cl}(\text{cone} \bigcup_{j=1}^m \{ e^t B_{i(\cdot)} : t \geq 0 \}),
\]

the so-called reachable cone \([33]\). One of the most frequently appearing classes of systems with \((A,B)\)-invariant proper convex cones are externally and internally positive systems.

Definition 3 \((A,B,C)\)-invariance. A linear system \((9)\) is called externally positive if

\[
\forall u \in \mathbb{R}_+^m : x(0) = 0 \implies y(t) \in \mathbb{R}^k_+.
\]

Internal positivity thus requires that the nonnegative orthant \( \mathbb{R}^n_+ \) is \( e^{At} \)-invariant, which is the case if and only if \( A \) is Metzler \([5]\).

Proposition 2 \((11)\). The following are equivalent:

1) \( (A,B,C,D) \) is internally positive.
2) \( \exists \alpha \geq 0 : A + \alpha I \in \mathbb{R}^{n \times n}_+ \) and \( B,C,D \) are element-wise nonnegative.
3) \( \mathcal{K}_e(A,B) \subset \mathbb{R}^n_+ \subset \mathcal{K}_o(A,C) \).

In particular, it can be shown that if \((9)\) admits an internally positive realization then

\[
\exists p \in \mathbb{R}, \ N \in \mathbb{R}^{n \times p} : \mathcal{K}_e(A,B) \subset \mathcal{P}_N \subset \mathcal{K}_o(A,C),
\]

\section{III. External positivity certificate}

Equipped with Lemma 2 and Proposition 1, we are ready to state our second-order cone certificate for external positivity.

Theorem 1 \((\text{Certificate for external positivity})\). Let \((A,B,C,D)\) be a linear system and assume that there exist \( K = K^T \in \mathbb{R}^{n \times n} \) with \( t(n-1,0,1) \) and \( \gamma, \tau \in \mathbb{R} \) such that

\[
\begin{align*}
A^TK + KA + 2\gamma K &\leq 0 \quad (15a) \\
B_{(\cdot)}^T KB_{(\cdot)} &\leq 0 \quad \text{for all } j \quad (15b) \\
\lambda_{n-1}(K) &> 0 \quad \lambda_n(K) \quad (15c) \\
K + \tau C_i^T C_i &> 0 \quad \text{for all } i \quad (15d) \\
C^T B, \ D &\in \mathbb{R}^{k \times m}_+ \quad (15e)
\end{align*}
\]

Then \((A,B,C,D)\) is externally positive.

Note that Theorem 1 is simply the modification of \((14)\) to \( \mathcal{K}_o(A,B) \subset \mathcal{K}_{K,C_{(\cdot)}} \subset \ldots \subset \mathcal{K}_{K,C_{(\cdot)}} \subset \mathcal{K}_o(A,C) \), where \( \mathcal{K}_{K,C_{(\cdot)}} \) is \((A,B)\)-invariant. This certificate could be refined by applying Theorem 1 to each subsystem \((A,B_{(\cdot)}), C_{(\cdot)}, D)\), separately. Figure 1 illustrates Theorem 1 in case of a SISO system.

Remark 3. By Lemma 1, the certificate requires that \( C e^{At} B \in \mathbb{R}^{k \times m}_+ \) for all \( t \geq 0 \). This is not a strong restriction, since a numerically decision of the sign of a floating point number can
only be done up to machine precision. In particular, we can use this fact to remove (15c). This is because for sufficiently small $\epsilon > 0$, we have that $e^{-\epsilon X_{K;C}}$ satisfies (16). Thus, under mild assumptions, e.g., $\lambda_1(A)$ is a simple dominant pole, $B_{(i,j)} \in \text{int}(e^{-\epsilon X_{K;C}})$, which allows us to choose (15b) to be strict. Together with (15d), this implies that $\tau(K) = (n - 1, 0, 1)$.

A. Example: Externally but not internally positive

Using Theorem 1 and Lemma 1, we can verify that (9) with

$$A = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & -\beta & \alpha \end{pmatrix}, \quad \beta \neq 0, \quad C = B^T, \quad b_1^2 > b_2^2 + b_3^2$$

(17)
is externally positive with the $e^t$-invariant cone $X_{K;C} = X_{K;C}^* = \{x : x_1^2 \geq x_2^2 + x_3^2\}$. However, by Lemma 3 the system does not admit an internally positive realization.

B. Restrictiveness

Unfortunately, even when restricting ourselves to systems with $Ce^tB \in \mathbb{R}^{3 \times m}$ for all $t \geq 0$, our certificate does not become a necessary condition as the following result shows.

**Proposition 3.** Let $A \in \mathbb{R}^{3 \times 3}$ be such that $\lambda_1(A) \in \mathbb{R}$, $\lambda_4(A) \neq \mathbb{R}(\lambda_2(A))$ and $\emptyset(\lambda_2(A)) = 0$. Then there exist $B, C, \Delta C \in \mathbb{R}^3$ such that

1) $\forall t \geq 0 : Ce^tB \geq 0$ and the only $(A,B)$-invariant cone $X \subset \{x : Cx \geq 0\}$ is $X = X_2(A,B)$. $(A,B)$ is neither polyhedral nor second-order.

2) $\forall t \geq 0 : (C + \Delta C)e^tB > 0$ and no $(A,B)$-invariant cone $X \subset \{x : (C + \Delta C)x \geq 0\}$ is second-order.

A proof to Proposition 3 is stated in appendix C.

**Remark 4.** As pointed out in [12], if $\forall t \geq 0 : (C + \Delta C)e^tB > 0$, then $(A,B,C + \Delta C,D)$ has an internally positive realization. However, as a consequence of Proposition 3, the dimension of such a realization can still be made arbitrarily large by choosing $\Delta C$ sufficiently small. Moreover, this also shows that even with the additional restriction to internally positive systems, our certificate remains only sufficient.

C. First order optimization

Even though the numerical tractability of our certificate is guaranteed by semi-definite programming, conventional SPD-solvers rely on interior point methods, which is why the cost per iteration for checking Theorem 1 grows unfavorably with $n$ [39, 40]. In order to reduce this cost, we will show how proximal splitting methods can be used as alternative solvers [57, 58]. For simplicity, we will restrict this discussion to minimal SISO systems $(A,B,C,D)$, where $\lambda_1(A) \neq \mathbb{R}(\lambda_2(A))$, but it is easy to adopt it to the general case.

We begin by rewriting (15a), (15b), and (15d) as the following optimization problem, where we use $\tau_1 = 1$:

$$\min_{K_1, K_2, R} \underbrace{I_{f_1}(K_1)}_{(K_1)} + \underbrace{I_{g_2}(K_2, R)}_{(K_2, R)} + \underbrace{I_{g_3}(K_1, K_2, R)}_{(K_1, K_2, R)}$$

(18)

with

$\mathcal{C} := \{K_1 : B^TK_1B \leq 0\}$

$\mathcal{Z} := \{(K_2, R_1) : A^TK_2 + K_2A + 2\gamma K_2 + R_1 = 0\}$

$\mathcal{P} := \{(K_1, K_2, R) : K_1 = K_2, K_1 + C^TC \geq \epsilon I_n, R \geq 0\}$

for some small $\epsilon > 0$. As we choose Douglas-Rachford splitting [57, 59] to solve (18), we need to compute the proximal mappings of

$$\text{prox}_{f_i}(Z) := \arg\min_x \left\{ f_i(x) + \frac{1}{2\epsilon}||x - Z||^2_F \right\} \text{ for } i = 1, 2, 3.$$ 

Whereas, $\text{prox}_{f_1}$ amounts to simple averaging with eigenvalue thresholding (see Algorithm 1), we can use Lagrange duality [56] to derive

$$\text{prox}_{f_1}(Z_K) = Z_K - \max \left(0, \frac{B^T Z_K R}{(B^T B)^{\frac{1}{2}}} \right) BB^T,$$

as well as $(Z_K - \bar{A}A - \bar{A}^T, Z_K - \Delta) = \text{prox}_{f_2}(Z_K, Z_K)$, where $\Lambda$ solves the Lyapunov-equation

$$\Lambda + \Lambda^T\bar{A}A\Lambda^T + (\bar{A}A + \bar{A}^T)\Lambda = W$$

with $\bar{A} := A + \gamma I_n$ and $W := \bar{A}^T Z_K + Z_K \bar{A} + Z_K$. The complete algorithm is outlined in Algorithm 1.

If $A$ is diagonal, then (2) can be solved with cost $O(n^2)$. Thus, for diagonalizable $A$, we conclude that each iteration in Algorithm 1 costs at most $O(n^3)$. A fully optimized numerical discussion of (2) is out of the scope of this paper.

**Algorithm 1.** Evaluate Theorem 1 by Douglas-Rachford

1: **Input:** Asymptotically stable minimal SISO system $(A,B,C,D)$ with $CB > 0$ and $\gamma \in (-\mathbb{R}(\lambda_2(A)), \lambda_1(A))$

2: Set $e_1 = e_2 = e_3 = 0$

3: $Z_{K_1} = Z_{K_2} = I_n$ and $Z_R = 0$

4: While $e_1 + e_2 + e_3 > -\epsilon$

5: Proximal mappings of $f_1$ and $f_2$:

$$X_{K_1} = \text{prox}_{f_1}(Z_{K_1}), \text{ see (P1)}$$

$$(X_{K_1}, X_R) = (Z_{K_2} - AA - A^T, Z_R - \Lambda), \text{ $\Lambda$ solves (P2)}$$

6: Proximal mapping of $f_3$:

$$S_K = \sum_{i=1}^n 2X_{K_i} + C^TC = \sum_{i=1}^n 2X_{K_i} + \sum_{i=1}^n \lambda_i(S_K) v_i v_i^T$$

$$R_K = \sum_{i=1}^n \max\{\epsilon, \lambda_i(S_K)\} u_i u_i^T - C^TC$$

$$Y_K = \sum_{i=1}^n \max\{0, \lambda_i(S_K)\} u_i u_i^T$$

7: Update $Z_{K_1}$ and $Z_R$:

$$Z_{K_1} = Z_{K_1} + Y_K - X_{K_1}, i = 1, 2$$

$$Z_R = Z_R + Y_K - X_R$$

8: Update $e_i$:

$$e_1 = \max\{0, CK^{-1}C^T + \epsilon\}, e_2 = \max\{0, B^T KB\}, e_3 = \lambda_1(A^TK + KA + 2\gamma K)$$

9: **Output:** $K$
IV. CONE BALANCED TRUNCATION

In general, neither balanced truncation nor other model reduction methods preserve a dominant real pole as required in Remark 1, unless the system is reduced to order one [30]. In the following, we modify generalized balanced truncation (BT) [53] such that exponential invariance with respect to a second-order cone is preserved. This procedure will be independent of the chosen cone, so it will also allow us to preserve \((A,B)\)-invariance, as well as the requirements of our external positivity certificate. Moreover, as for generalized balanced truncation, it is possible to incorporate additional LMI-representable conditions, e.g., to preserve passivity.

We begin by introducing the concepts of cone-balanced realization and truncation. For notational convenience, we restrict ourselves to minimal realization, but the readers should convince themselves that everything can be adopted to a non-minimal setting. Subsequently, we show how to transform a minimal system into cone-balanced form and conclude with an algorithmic discussion.

**Definition 5** (Cone-balanced realization). A minimal linear system realization \((\hat{A}, \hat{B}, \hat{C}, D)\) is called cone-balanced, if there exists diagonal \(K\) with \(\text{int}(K) = (n - 1,0,1)\), diagonal \(\bar{P}, \bar{Q} > 0\) and \(\gamma > 0\) such that

\[
\begin{align*}
\bar{A}^T K + \bar{K} A + 2\gamma K &\preceq 0, \quad \text{(19a)} \\
\bar{A} \bar{P} + P \bar{A}^T &\preceq -\bar{B} B^T, \quad \text{(19b)} \\
\bar{A}^T \bar{Q} + \bar{Q} A &\preceq -\bar{C} C^T, \quad \text{(19c)} \\
\bar{P}_{11} = \bar{Q}_{11} &\preceq \cdots \preceq \bar{P}_{nn} = \bar{Q}_{nn} \quad \text{and} \quad \bar{k}_{11} < 0 \quad \text{(19d)}
\end{align*}
\]

The following result is proven in Appendix D.

**Theorem 2** (Cone-balanced truncation). Suppose \((\hat{A}, \hat{B}, \hat{C}, \hat{D})\) is an asymptotically stable, cone-balanced realization of the transfer function \(G(s)\) with \(K\), \(\gamma\) and

\[
\bar{P} = \text{blkdiag}(\sigma_1, \sigma_2, \ldots, \sigma_p, \rho_p)
\]

as in (19a) and (19b), where \(\sigma_2 > \cdots > \sigma_p\).

Then, for any \(1 \leq r < p\), \((\hat{A}_{(1:R,1:R)}, \hat{B}_{(1:R,)}, \hat{C}_{(1:R)}, \hat{D}_{(1:R,1:R)})\) with \(R := 1 + \sum_{i=1}^{p} l_i + 1\) and transfer function \(G_R(s)\), is an asymptotically stable, controllable, cone-balanced system fulfilling

\[
\|G - G_R\|_\infty \leq 2 \sum_{i=r+1}^{p} \sigma_i. \quad \text{(E)}
\]

Further, the following are preserved:

1) \(\lambda_1(\hat{A}_{(1:R,1:R)}) \leq \gamma\).

2) If \(x_K\) is \((\hat{A}, \hat{B})\)-invariant, then \(x_{\hat{K}_{(1:R,1:R)}}\) is \((\hat{A}_{(1:R,1:R)}, \hat{B}_{(1:R,)}, \hat{C}_{(1:R)}, \hat{D}_{(1:R,1:R)})\)-invariant.

3) If \((\hat{A}, \hat{B}, \hat{C}, \hat{D})\) fulfills Theorem 1 with \(K = \bar{K}\), then the same holds for \((\hat{A}_{(1:R,1:R)}, \hat{B}_{(1:R,)}, \hat{C}_{(1:R)}, \hat{D}_{(1:R,1:R)})\) with \(K = \bar{K}_{(1:R,1:R)}\).

**A. Cone-balancing**

Next we will discuss how to compute a cone-balanced realization. We start with the first step that yields a state-space transformation such that (19a) and (19b) are fulfilled.

**Proposition 4.** Given \((A, B)\) and \(N \geq 0\), assume that there exists \(\gamma > 0\), \(K = K^T\) with \(t(K) = (n - 1,0,1)\) and \(P > 0\) such that

\[
\begin{align*}
A^T K + K A + 2\gamma K &\preceq 0 \quad \text{(20a)} \\
\text{trace}(N K) &\preceq 0 \quad \text{(20b)} \\
A \bar{P} + \bar{P} A^T &\preceq -N \quad \text{(20c)}
\end{align*}
\]

Then there exists \(T \in \mathbb{R}^{n \times n}\) such that

\[
\begin{align*}
\bar{P} &:= T^{-1} PT^{-T} = \text{blkdiag}(\sigma_1, \sigma_2, \ldots, \sigma_p) \\
\bar{K} &:= T^T K T = \text{blkdiag}(-\sigma_1, \sigma_2, \ldots, \sigma_p)
\end{align*}
\]

where \(\sigma_1 \geq \cdots \geq \sigma_p > 0\) and \(l_2 + \cdots + l_p = n - 1\) and

\[
\sigma_i^2 \geq \sum_{l=1}^{n} l_i \sigma_i^2. \quad \text{(21)}
\]

In particular, if \(N \geq BB^T\), then \((\hat{A}, \hat{B}) = (T^{-1}AT, T^{-1}B)\) fulfills (19a) and (19b) with diagonal \(\bar{K}\) and \(\bar{P}\).

**Proposition 5.** Let \((\hat{A}, \hat{B}, \hat{C}, \hat{D})\) and \(\hat{N} \geq 0\) be such that there exist diagonal \(\bar{K}\) and \(\bar{P}\) with \(\bar{P} = |K|\), \(\text{trace}(\hat{N} \bar{K}) < 0\), \(\bar{P} > 0\), \(t(K) = (n - 1,0,1)\), \(x_{\bar{K}}\) being \(e^{\mathbb{C}T}\)-invariant and

\[
\begin{align*}
\bar{A} \bar{P} + \bar{P} A^T &\preceq -\bar{C} C^T
\end{align*}
\]

Then, there exists diagonal \(\Delta > 0\) such that

\[
\begin{align*}
\bar{A}^T \Delta + \Delta A &\preceq -\bar{C} C^T
\end{align*}
\]

In particular, \((T^{-1} \bar{A}^T, T^{-1} \bar{B}, T^{-1} \bar{C})\) is cone-balanced with respect to \(x_{\bar{K}^T \bar{K}^T}\), if \(\hat{N} \geq BB^T\) and \(\hat{k}_{11} < 0\), where

\[
T := \text{blkdiag} \left(1, \frac{\bar{p}_{22}}{\delta_{22}}, \ldots, \frac{\bar{p}_{nn}}{\delta_{nn}}\right)^{\frac{1}{2}} \Pi
\]

and \(\Pi\) is a permutation matrix according to (19d).

A proof of this result can be found in Appendix F. Since for given \(K\), we can always find \(N \geq BB^T\) and \(\hat{N} \geq BB^T\) as in Propositions 4 and 5, we have shown that (20a) is necessary and sufficient for the existence of a cone-balanced realization. Further, if \(x_{\bar{K}}\) is \((\hat{A}, \hat{B})\)-invariant, then

\[
\text{trace}(BB^T K) = \sum_j B_{(:,j)} KB_{(:,j)} \leq 0
\]

i.e., we can choose \(N = BB^T\) and receive equality in (19b).

**Corollary 1.** Let \((A, B, C, D)\) be such that \(x_{\bar{K}}\) is \((\hat{A}, \hat{B})\)-invariant. Then there exists a transformation \(T\) such that \((\hat{A}, \hat{B}, \hat{C}, \hat{D}) := (T^{-1}AT, T^{-1}B, CT, D)\) is cone-balanced with respect to \(x_{\bar{K}^T \bar{K}^T}\) and equality holds in (19b).

Finally note that Propositions 4 and 5 are intentionally presented based on \(N\) and \(\hat{N}\), respectively. In this way, it is easy to see how other LMI-representable properties (see e.g. [42, 43]) can be incorporated.
B. Error-bound minimization

Let us now discuss the question of choosing $\tilde{P}$ and $\Delta$ such that the error-bound (E) is small. We only consider the case where we also want to preserve external positivity through Theorem 1. In this case, Corollary 1 applies and we can fix $\tilde{P}$ to be the controllability Gramian. Indeed, this is the best possible choice, since the eigenvalues of $\tilde{P}$ are always at least as large as those of the controllability Gramian [30]. Then for finding $\Delta$, we can minimize the low-rank promoting nuclear norm [60] of $\tilde{P}\Delta$. Alternatively, any other low-rank promoting norms [61, 62] may also be considered. A summary of the algorithm is outlined in Algorithm 2.

Finally note that since $\tilde{K}$ is not unique, its choice may be of considerable importance. Our experiments indicate that computing $\tilde{K}$ with respect to a balanced realization gives satisfactory results.

Algorithm 2 (Positive) cone balanced truncation

1. **Input:** Asymptotically stable minimal $(A, B, C, D)$ and $K$ that fulfills (15a)-(15c) for cone balanced truncation and additionally (15d) and (15e) for positive cone balanced truncation.
2. Compute $\tilde{P}$ and $T$ in Proposition 4 with $N = BB^T$ and $(\tilde{A}, \tilde{B}, \tilde{C}, D) := (T^{-1}AT, T^{-1}B, CT, D)$.
3. Minimize $\sum_{i>j} \delta_i \delta_j$ subject to $\tilde{A}^T \Delta + \Delta \tilde{A} \preceq -C^T \tilde{C}, \quad \Delta := \text{blkdiag}(\delta_1, \ldots, \delta_m) \succeq 0$.
4. Find a cone-balanced realization $(\tilde{A}, \tilde{B}, \tilde{C}, D)$ as in Proposition 5 with generalized singular values $\sigma_i := \sqrt{\delta_i \delta_i}, \ i > 1$.
5. Choose a reduced order $\bar{R}$ according to (E).
6. **Output:** $(\tilde{A}(1:R,1:R), \tilde{B}(1:R,:), \tilde{C}(:,1:R), D)$.

V. Numerical Examples

In our first example, we will see how our approach overcomes the conservatism of [30], if one only seeks to preserve external positivity. In addition, the second example will analyze how our approach performs, when balanced truncation does not retain a dominant real pole.

Note that by [30, 63] and [47] it follows, that even a reduced model of order one often outperforms the methods in [44–47]. Therefore, it suffices to compare our reduction approaches to symmetric balanced truncation (SBT) [63] as well as balanced truncation (BT).

Our comparison will always start from a minimal realization, which can be considered a pre-reduction. In order to make our solutions comparable, we will add to minimize trace$(Q + \tau C^T C)$ in case of positive cone balanced truncation (PCBT), which turned out to even improve the results. For cone balanced truncation (CBT) we use the same $\gamma$ as determined by PCBT and $K$ is given by

$$ A^T K + KA + 2\gamma K = -C^T C. $$

A. Discretized heat equation

We begin with the two-dimensional heat equation on a square

$$ T = \Delta T = \frac{\partial^2}{\partial x^2} T + \frac{\partial^2}{\partial y^2} T $$

with control of the Dirichlet boundary conditions of the four edges. Discretization on a uniform grid leads to the following normalized linear internally positive system:

$$ \tilde{T} = AT + Bu \quad \text{with} \ u \in \mathbb{R}^4 $$

where $A := (a_{ij}) \in \mathbb{R}^{N^2 \times N^2}$ and $B := (b_{ij}) \in \mathbb{R}^{N^2 \times 4}$ are zero except for

$$ a_{ii} := -4, \quad \text{for} \ i = 1, 2, \ldots, N^2 $$

$$ a_{i,i+1} = a_{i+1,i} := 1, \quad \text{for} \ i = 1, \ldots, N^2 - 1 $$

$$ a_{i,N+i} = a_{N+i,i} := 1, \quad \text{for} \ i = 1, \ldots, N(N-1) $$

and

$$ b_{i1} := 1, \quad \text{for} \ i = 1, 2, \ldots, N $$

$$ b_{i2} := 1, \quad \text{for} \ i = N, 2N, \ldots, N^2 $$

$$ b_{i3} := 1, \quad \text{for} \ i = N(N-1) + 1, N(N-1) + 2, \ldots, N^2 $$

$$ b_{i4} := 1, \quad \text{for} \ i = 1, N+1, \ldots, N(N-1)+1. $$

In our example with $N = 10$, we will use the second and the fourth input, i.e., $u_1 = u_3 = 0$. Furthermore, we split the unit-square into 5 equally spaced vertical stripes and let $y$ represent the average temperature in each of these zones, i.e.

$$ C = \text{blkdiag} \left( I_{x^2}^T, I_{x^2}^T, I_{x^2}^T, I_{x^2}^T, I_{x^2}^T \right), $$

where $1_{x^2} \in \mathbb{R}^{x^2}$ stands for the vector of all ones. Since the minimal balanced system is not symmetric, SBT will return an approximation of order 1 with the same error as BT, but cannot achieve any higher order approximation. In contrast, the normalized errors shown in Figure 2 show that PCBT and CBT perform fairly close to BT.
B. Balanced truncation without dominant real pole

Next, let us have a look at an externally positive system, where we have not computed an internally positive realization and where BT does not preserve a dominant real pole and thus neither external positivity. Our system of order 11 is defined by \((A, B, C) := (A_n - 2I_1, B_a, C_a)\), where \((A_n, B_a, C_a)\) is the MATLAB canonical companion form of

\[
G_a(s) = \frac{(s + 1)^{10}}{(s - 1)\prod_{k=2}^{5}(s^2 - 2\cos(\sqrt{k}\pi) + 1)\prod_{k=4}^{9}(s + \frac{1}{2})}.
\]

Using Theorem 1, we can verify that this system is externally positive with \(\gamma = 1.0033\). However, BT does not preserve the dominant real pole property for the reduced models of order two and four. A comparison of the normalized errors of BT, CBT and PCBT is presented in Figure 3 and we can see that the later two perform generally well. Nevertheless, we can observe a large error of PCBT for order two. Moreover, for larger orders we observe that CBT and PCBT increasingly deviate from BT. However, in these cases we can use Theorem 1 to verify that BT preserves external positivity, which is why we do not have to employ either of these methods here.

Finally note that this example was constructed such that other well established model reduction methods, e.g., [44] and [65], also do not preserve a dominant real pole for this example.

VI. CONCLUSION

In this work, we have formally derived solutions to the following problems:

1) External positivity certificate based on exponentially invariant second-order cones.
2) Modified balanced truncation to preserve a dominant real pole.
3) Modified balanced truncation to preserve external positivity when our certificate applies.

Our certificate has the advantage that it is tractable through semi-definite programming, has fixed computational cost in contrast to computing internal positive realization and allows us to certify external positivity, where no other invariant would work. Nevertheless, this certificate is still only a sufficient test as we constructed systems that do not fulfill its requirements.

In particular, we have seen that there exist systems where our test as well as the attempt of finding an internal positive realization fail. Thus, showing the need for more general cones, which can be tractability shown to be exponentially invariant. Moreover, despite the introduced first order optimization methods, our certificate does not scale well to large systems. Therefore, it would be interesting to see if Theorem 1 can be solved analytically with a sparse \(K\) for a class of sparse \(A\).

Based on the examples, we can see that our modified balanced truncation methods have can yield approximations that are qualitatively close to traditional balanced truncation. Thus, indicating that these approaches only imposes a mild conservatism and outperforms methods that intend to preserve internal positivity [44, 47]. However, it remains to understand how one can systematically choose a second-order cone that gives small truncation errors/error bounds.

Finally note that it is straightforward to extend all results to discrete-time linear systems, see e.g. [4] for invariant second-order cones. Moreover, since our reduction methods preserves the concept of 1-dominance [27], it can also be modified to preserve \(p\)-dominance.

APPENDIX

A. Proof to Lemma 1

Proof. We start with the equivalence of Items 1 and 2. For \(Q_{Q_e} = \mathcal{K}_{Q_e}^p\) to hold true, \(c\) must fulfill (6), which by the definition of the dual cone (1a) is the case if and only if \(c \in \text{int}(\mathcal{K}_{Q_e}^\ast)\). In order to see the equality, note that there exists \(T \in \mathbb{R}^{n \times n}\) such that \(T^T Q_e := \text{diag}(1, \ldots, 1, -1)\) and \(T^{-1}\) maps \(\mathcal{K}_{Q_e}\) onto the self-dual cone \(\mathcal{K}_{Q_e, e_n}\) [56, Example 2.25], where \(e_n\) is the n-th canonical unit vector. Hence,

\[
(T^{-1}\mathcal{K}_{Q_e}) = \mathcal{K}_{Q_e, e_n} = (T^{-1}\mathcal{K}_{Q_e}^\ast)^T = T^T \mathcal{K}_{Q_e}^p,
\]

showing that \(\mathcal{K}_{Q_e}^\ast = \mathcal{K}_{TQ_eT^T, Te_n} = \mathcal{K}_{Q_e}^{-1, Te_n}\). As before, it follows that all the normals to strictly separating hyperplanes of \(\mathcal{K}_{Q_e}^{-1}\) are given by \(\text{int}(\mathcal{K}_{Q_e}) = \text{int}(\mathcal{K}_{Q_e}^\ast)\).

Then Item 3 just makes explicit Item 2. Further, \(c \in \text{int}(\mathcal{K}_{Q_e}^\ast)\) if and only if there exists \(\varepsilon > 0\) such that

\[
\forall x \in \mathcal{K}_{Q_e, e_n} \setminus \{0\} : x^T Q_e + \varepsilon x^T e_n x > 0,
\]

which is equivalent to \(Q + \varepsilon e_n e_n^T > 0\).

B. Proof to Lemma 3

Proof. Without loss of generality, let

\[
A = \begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & \beta \\
0 & -\beta & 0
\end{pmatrix}
\]

Then for all \(b \in \mathbb{R}^3\) with \(b_1 \neq 0\), the set \(\{e^{At}b : t \geq 0\}\) is an ellipse, which implies that \(\mathcal{K}_e(A, b) = \mathcal{K}_{Q_b, e_b}\) with \(Q_b = \text{diag}(-\frac{b_2^2 + b_3^2}{b_1^2}, 1, 1)\) and \(e_b = \text{sign}(b_1)(1, 0, 0)^T\). Hence, any \(e^{At}\)-invariant proper convex cone \(\mathcal{K}\) can be written as

\[
\mathcal{K} = \text{cone}(\bigcup_{b \in \mathcal{K}} \mathcal{K}_{Q_b, e_b}) = \mathcal{K}_{Q_{\beta, \max, e_{\max}}},
\]

(25)
where $b_{\max} = \arg\max_{b \in \mathcal{X}} \frac{b_2^T + b_3^T}{b_1^T}$.

C. Proof to Proposition 3

Proof. Without loss of generality, let
\[
A = \begin{pmatrix}
\alpha & \beta & 0 \\
-\beta & \alpha & 0 \\
0 & 0 & 0
\end{pmatrix} =: \text{blkdiag}(A_1, 0),
\]
where $\alpha < 0$ and $\beta \neq 0$. Further, let $B = (1, 0, 1)^T$ and $x(t) := e^{tB}$. Since $\{(x_1(t), x_2(t)) : t \geq 0\} \subset \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 0\}$ is not a closed contour, there exists a tangent hyperplane $\mathcal{T}^2 := \{x \in \mathbb{R}^2 : c^T x = c_1\}$ to $\mathcal{K}^2 := \text{cl}(\text{conv}(\{(x_1(t), x_2(t)) : t \geq 0\}))$ such that
1. $\mathcal{K}^2 \subset \mathcal{T}^2$
2. $\exists x^* > 0 : c_1 x_1(t^*) + c_2 x_2(t^*) = c_1$
and therefore $T := \{x : c x \geq 0\}$ with $C := (c_1, c_2, -c_1)$ is a tangent hyperplane to $\mathcal{K}_p(A, B) = \text{cl}(\text{cone}(\{1 \times \mathcal{K}^2\}))$. Thus, $C e^B t \geq 0$ for all $t \geq 0$ and $C B \leq C e^B = 0$. In particular, for all $B_\theta \notin \text{conv}(\mathcal{K}^2)$ there exists $\theta \geq 0$ such that $C e^B (B_\theta, 1)^T \in \mathcal{K}^2$. However, since $\text{conv}(\mathcal{K}^2)$ is neither a polygon nor an ellipse, $\mathcal{K}_p(A, B)$ can neither be polyhedral nor second-order.

Finally, note that for arbitrary $\delta > 0$ and $\Delta C := (0, \delta, \epsilon c_1)$, it holds that $(C + \Delta C) e^B t \geq 0$ for all $t \geq 0$. Assume that for all $\delta > 0$, there exists an $(A, B)$-invariant proper second order cone $\mathcal{K}_p(A, B) \subset \{x : \text{conv}(\{1 \times \mathcal{K}^2\}) \cap \{x : C e^B t \geq 0\} \}$ for some $P > 0$ and $k \in \mathbb{R}^2$ with $\text{conv}(\{(x(t^*), (1, 0)^T)\}) \subset \{x : c^T x \leq (1 - \epsilon)c_1\}$. However, as $\epsilon \to 0$, this requires that either $x_1(t) \to \infty$ or $x_3(t) \to 0$. Thus the area of $\delta$ can be made arbitrarily small or large, which either contradicts that $\mathcal{K}_p$ is $(A, B, \epsilon)$-invariant or $\mathcal{K}_p(A, B) \subset \{x : (C + \Delta C) x \geq 0\}$.

D. Proof to Theorem 2

Proof. The first part and the error bound follows as for generalized balanced truncation [53, 54]. Item 2 follows by
\[
\tilde{B}_{(i,j)}^T K \tilde{B}_{(i,j)} \geq B^T_{(i,j)} K_{(i,j)} B_{(i,j)}
\]
which implies that if (15a)–(15c) are fulfilled for $(\tilde{A}, B, C, D)$, $K$ and some $\gamma$, then the same applies to $(\tilde{A}_{(1:1,1:1)}, \tilde{B}_{(1:1,1:1)})$. If additionally (15d) and (15e) hold, as assumed in Item 3, then Lemma 1 yields that
\[
(B^T_{(1,1:1)}, 0)^T \in \mathcal{K}_{K, k_1} = \mathcal{K}_{K, C(k_1)}
\]
for all $i, j$
\[
0 \geq \tilde{C}_{(i,j)} K^{-1} \tilde{C}_{(i,j)}^T \geq \tilde{C}_{(i,j)} K^{-1} C_{(i,j)}^T
\]
for all $i, j$.

E. Proof to Proposition 4

Proof. Let $P$ and $K$ be as in the claim. We perform a singular value decomposition $P = U \Sigma R^T$ and define $L := U \Sigma R$. By another singular value decomposition of $L^T K L = \Sigma^2 V^T$, we define $T := LV \Sigma^{-1}$. Thus, $\tilde{P} := L^T P^{-1} T$ and $\tilde{k} := T^T QT$ fulfill
\[
\tilde{P} = \Sigma^2 V T^{-1} L L^T V^T \Sigma^{-1} = \Sigma,
\]
\[
|K| = |\Sigma^2 V T^T L^T Q L V^T \Sigma^{-1}| = \Sigma,
\]
with $\Sigma = \text{blkdiag}(\sigma_i I, \ldots, \sigma_i I)$. $\sigma_i \gg \sigma_i > 0$ and $l_1 + \cdots + l_i = n$. By [66, Theorem 4.58], $t(T^T K T) = t(K)$, which is why $\tilde{P}$ and $K$ are equal up to a sign-change on one of the diagonal entries.

Next, we will see that $\text{trace}(K) < 0$, which is why the sign-change occurs at $\sigma_i$ and $l_i = 1$. To this end, assume without loss of generality that $P = I$ and $|K| = \Sigma^2$, i.e.
\[
\tilde{K} T + K \tilde{A} + 2 \gamma K \leq 0,
\]
\[
\text{trace}(K) \leq 0
\]
\[
\tilde{K} + \tilde{A}^T = -N.
\]
By substitution of $\tilde{A} = -N - \tilde{A}^T$ in (26) we get
\[
-(N + \tilde{A}) K - (N + \tilde{A}^T) - 2 \gamma K \leq -4 \gamma K.
\]

Taking the trace over (29) and using
- $\text{trace}(K) = \text{trace}(K^T)$
- $\text{trace}(K^T + \tilde{A}^T + 2 \gamma K) = \text{trace}(\tilde{A}^T K + \tilde{K} + 2 \gamma K) \leq 0$

yields
\[
2 \gamma \text{trace}(K) \leq \text{trace}(K) \leq 0,
\]
which is why the sign-change occurs at $\sigma_i$ and $l_i = 1$. To this end, assume without loss of generality that $P = I$ and $|K| = \Sigma^2$, i.e.
\[
\tilde{K} T + K \tilde{A} + 2 \gamma K \leq 0,
\]
\[
\text{trace}(K) \leq 0
\]
\[
\tilde{K} + \tilde{A}^T = -N.
\]
By substitution of $\tilde{A} = -N - \tilde{A}^T$ in (26) we get
\[
-(N + \tilde{A}) K - (N + \tilde{A}^T) - 2 \gamma K \leq -4 \gamma K.
\]

F. Proof to Proposition 5

Proof. Let $(\tilde{A}, \tilde{C})$ and $N = LL^T$ be as in the assumptions. Since
\[
\text{trace}(N \tilde{K}) = \sum_j L_{(j, i)}^T \tilde{K}_{(j, i)} < 0,
\]
we can assume without loss of generality that $L_{(j, i)}^T \tilde{K}_{(j, i)} < 0$. Thus, by Lemmas 1 and 2 there exists a sufficiently large $\varepsilon > 0$ such that
\[
\tilde{A}^T \tilde{K} + \tilde{K} \tilde{A} + 2 \gamma \tilde{K} \leq 0,
\]
\[
\tilde{A}^T + \tilde{A} \tilde{K} \leq -L_{(i, i)}^T\tilde{K}_{(i, i)} < 0,
\]
\[
\tilde{K}^{-1} + \varepsilon L_{(i, i)}^T \tilde{K}_{(i, i)} \geq 0,
\]
\[
2 \gamma p_{11} - p_{11}^T > 0.
\]
Multiplying (31a) with $\tilde{K}^{-1}$ from the right and the left yields
\[
\tilde{A} \tilde{K}^{-1} + \tilde{K}^{-1} \tilde{A} T + 2 \gamma \tilde{K} \leq 0,
\]
and multiplying (31b) by $2 \gamma$ gives
\[
2 \gamma \tilde{A} \tilde{P} + 2 \gamma \tilde{P} \tilde{A} \tilde{T} + 2 \gamma \varepsilon L_{(i, i)}^T \tilde{K}_{(i, i)} \leq 0.
\]
Adding up (32) and (33) results in
\[
\tilde{A} \Delta^{-1} + \Delta^{-1} \tilde{A} T + 2 \gamma \left(\tilde{K}^{-1} + \varepsilon L_{(i, i)}^T \tilde{K}_{(i, i)}\right) \leq 0.
\]
with $\Delta := (2\gamma \tilde{P} + \tilde{K}^{-1})^{-1} \succ 0$. Finally, a proper scaling of $\Delta$ gives a diagonal solution to
\[ \tilde{A}^T \Delta + \Delta \tilde{A} \preceq -C^T \tilde{C}. \] (34)

The last implication follows by Proposition 4. \qed

ACKNOWLEDGMENT

This work was supported by the ELLIIT Excellence Center, by the Swedish Research Council through the LCCCC Linnaeus Center and the European Research Council under the Advanced ERC Grant Agreement Switchlet n.670645.

REFERENCES

[1] O. Perron, “Zur theorie der matrices,” Mathematische Annalen, vol. 64, no. 2, pp. 248–263, Jun 1907.
[2] G. Frobenius, “Über Matrizen aus nicht negativen Elementen,” 1912.
[3] G. Birkhoff, “Extensions of Jentzsch’s theorem,” Transactions of the American Mathematical Society, vol. 85, no. 1, pp. 219–227, 1957.
[4] R. J. Stern and H. Wolokowicz, “Invariant ellipsoidal cones,” Linear Algebra and its Applications, vol. 150, pp. 81 – 106, 1991.
[5] A. Berman and R. Plemmons, Nonnegative Matrices in the Mathematical Sciences. SIAM, 1994.
[6] R. Loewy and H. Schneider, “Positive operators on the $n$-dimensional ice cream cone,” Journal of Mathematical Analysis and Applications, vol. 49, no. 2, pp. 375 – 392, 1975.
[7] H. Schneider and M. Vidyasagar, “Cross-positive matrices,” SIAM Journal on Numerical Analysis, vol. 7, no. 4, pp. 508–519, 1970.
[8] C. Mostajeran and R. Sepulchre, “Ordering positive definite matrices,” Information Geometry, vol. 1, no. 2, pp. 287–313, Dec 2018.
[9] T. Damm, “Positive groups on $\mathbb{R}^n$ are completely positive,” Linear Algebra and its Applications, vol. 393, pp. 127 – 137, 2004, special Issue on Positivity in Linear Algebra.
[10] A. Berman, M. Neumann, and R. J. Stern, Nonnegative Matrices in Dynamic Systems. Wiley & Sons, 1989, vol. 3.
[11] D. Luenberger, Introduction to Dynamic Systems: Theory, Models & Applications. John Wiley & Sons, 1979.
[12] L. Farina and S. Rinaldi, Positive Linear Systems: Theory and Applications. John Wiley & Sons, 2011.
[13] A. Rantzer, “Scalable control of positive systems,” European Journal of Control, vol. 24, pp. 72 – 80, 2015.
[14] C. Briat, “Robust stability and stabilization of uncertain linear positive systems via integral linear constraints: $L_1$-gain and $L_\infty$-gain characterization,” pp. 6337–6342, 2011.
[15] T. Tanaka and C. Langbort, “The bounded real lemma for internally positive systems and H-infinity structured static state feedback,” IEEE Transactions on Automatic Control, vol. 56, no. 9, pp. 2218–2223, 2011.
[16] D. Angeli and E. D. Sontag, “Monotone control systems,” IEEE Transactions on Automatic Control, vol. 48, no. 10, pp. 1684–1698, Oct 2003.
[17] H. L. Smith, Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems. American Mathematical Soc., 2008, no. 41.
[18] F. Forni and R. Sepulchre, “Differentially positive systems,” IEEE Transactions on Automatic Control, vol. 61, no. 2, pp. 346–359, 2016.
[19] C. Grussler and R. Sepulchre, “Strongly unimodal systems,” arXiv:1811.03986, 2018.
[20] N. K. Son and D. Hinrichsen, “Robust stability of positive continuous time systems,” Numerical Functional Analysis and Optimization, vol. 17, no. 5-6, pp. 649–659, 1996.
[21] P. D. Leenheer and D. Aeyels, “Stabilization of positive linear systems,” Systems & Control Letters, vol. 44, no. 4, pp. 259 – 271, 2001.
[22] T. Kaczorek, Positive 1D and 2D systems. Springer Science & Business Media, 2012.
[23] R. F. Brown, “Compartmental system analysis: State of the art,” IEEE Transactions on Biomedical Engineering, vol. BME-27, no. 1, pp. 1–11, 1980.
[24] R. Shorten, F. Wirth, and D. Leith, “A positive systems model of TCP-like congestion control: asymptotic results,” IEEE/ACM Transactions on Networking, vol. 14, no. 3, pp. 616–629, 2006.
[25] M. E. Valcher, “Reachability properties of continuous-time positive systems,” IEEE Transactions on Automatic Control, vol. 54, no. 7, pp. 1586–1590, July 2009.
[26] V. Jonsson, A. Rantzer, and R. M. Murray, “A scalable formulation for engineering combination therapies for evolutionary dynamics of disease,” in 2014 American Control Conference, June 2014, pp. 2771–2778.
[27] F. Forni and R. Sepulchre, “Differential dissipativity theory for dominance analysis,” IEEE Transactions on Automatic Control, pp. 1–1, 2018.
[28] C. Mostajeran and R. Sepulchre, “Positivity, monotonicity, and consensus on lie groups,” SIAM Journal on Control and Optimization, 2018.
[29] A. Soothing and A. Mauroy, “Operator-theoretic characterization of eventually monotone systems,” IEEE Control Systems Letters, vol. 2, no. 3, pp. 429–434, 2018.
[30] C. Grussler and T. Damm, “A symmetry approach for balanced truncation of positive linear systems,” in 51st IEEE Conference on Decision and Control (CDC), Maui, HI, 2012, pp. 4308–4313.
[31] C. Grussler, J. Umenberger, and I. R. Manchester, “Identification of externally positive systems,” in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, VIC, Australia, 2017, pp. 6549–6554.
[32] V. D. Blondel and N. Portier, “The presence of a zero in an integer linear recurrent sequence is NP-hard to decide,” Linear Algebra and its Applications, vol. 351, pp. 91 – 98, 2002.
[33] Y. Ohta, H. Maeda, and S. Kodama, “Reachability, observability, and realizability of continuous-time positive systems,” SIAM Journal on Control and Optimization, vol. 22, no. 2, pp. 171–180, 1984.
[34] B. D. O. Anderson, M. Deistler, L. Farina, and L. Benvenuti, “Nonnegative realization of a linear system with
nonnegative impulse response,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 43, no. 2, pp. 134–142, 1996.

[35] L. Farina, “On the existence of a positive realization,” *Systems & Control Letters*, vol. 28, no. 4, pp. 219 – 226, 1996.

[36] L. Benvenuti and L. Farina, “A tutorial on the positive realization problem,” *IEEE Transactions on Automatic Control*, vol. 49, no. 5, pp. 651–664, 2004.

[37] R. J. Stern and H. Wolkowicz, “Exponential nonnegativity on the ice cream cone,” *SIAM Journal on Matrix Analysis and Applications*, vol. 12, no. 1, pp. 160–165, 1991.

[38] R. Hildebrand, “An LMI description for the cone of lorentz-positive maps ii,” *Linear and Multilinear Algebra*, vol. 59, no. 7, pp. 719–731, 2011.

[39] D. Peaucelle, D. Henrion, Y. Labit, and K. Taitz, “User’s guide for SEDUMI INTERFACE 1.04,” 2002, LAAS-CNRS, Toulouse.

[40] K. C. Toh, R. H. Tutuncu, and M. J. Todd, “On the implementation of SDPT3 (version 3.1) – a MATLAB software package for semidefinite-quadratic-linear programming,” in *IEEE International Conference on Robotics and Automation*, 2004, pp. 290–296.

[41] B. Moore, “Principal component analysis in linear systems: Controllability, observability, and model reduction,” *IEEE Transactions on Automatic Control*, vol. 26, no. 1, pp. 17–32, 1981.

[42] X. Du and G. Yang, “h∞ model reduction of linear continuous-time systems over finite-frequency interval,” *IET Control Theory Applications*, vol. 4, no. 3, pp. 499–508, 2010.

[43] S. Gugercin and A. C. Antoulas, “A survey of model reduction by balanced truncation and some new results,” *International Journal of Control*, vol. 77, no. 8, pp. 748–766, 2004.

[44] T. Reis and E. Virnik, “Positivity preserving balanced truncation for descriptor systems,” *SIAM Journal on Control and Optimization*, vol. 48, no. 4, pp. 2600–2619, 2009.

[45] J. Feng, J. Lam, Z. Shu, and Q. Wang, “Internal positivity preserved model reduction,” *Int. Journal of Control*, vol. 83, no. 3, pp. 575–584, 2010.

[46] P. Li, J. Lam, Z. Wang, and P. Date, “Positivity-preserving H∞ model reduction for positive system,” *Automatica*, vol. 47, no. 7, pp. 1504 – 1511, 2011.

[47] A. Sootla and A. Rantzer, “Scalable positivity preserving model reduction using linear energy functions,” in 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), 2012, pp. 4285–4290.

[48] C. Grussler and A. Rantzer, “Modified balanced truncation preserving ellipsoidal cone-invariance,” in 53rd IEEE Conference on Decision and Control (CDC), Los Angeles, CA, 2014, pp. 2365–2370.

[49] J. Zheng, Y. Zhang, and L. Qiu, “Projected spectrahedral cone-invariant realization of an LTI system with nonnegative impulse response,” in 2016 IEEE 55th Conference on Decision and Control (CDC), 2016, pp. 6613–6618.

[50] C. Altafini, “Minimal eventually positive realizations of externally positive systems,” *Automatica*, vol. 68, pp. 140 – 147, 2016.

[51] Y. Ebihara, “H∞ analysis of LTI systems via conversion to externally positive systems,” *IEEE Transactions on Automatic Control*, vol. 63, no. 8, pp. 2566–2572, Aug 2018.

[52] J. Zheng, J. Dong, and L. Xie, “Minimal strongly eventually positive realization for a class of externally positive systems,” *IEEE Transactions on Automatic Control*, pp. 1–1, 2019.

[53] C. L. Beck, J. Doyle, and K. Glover, “Model reduction of multidimensional and uncertain systems,” *IEEE Transactions on Automatic Control*, vol. 41, no. 10, pp. 1466–1477, 1996.

[54] H. Sandberg and A. Rantzer, “Balanced truncation of linear time-varying systems,” *IEEE Transactions on Automatic Control*, vol. 49, no. 2, pp. 217–229, 2004.

[55] A. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2005.

[56] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[57] P. L. Combettes and J.-C. Pesquet, *Proximal Splitting Methods in Signal Processing*. Springer New York, 2011, pp. 185–212.

[58] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, ser. CMS Books in Mathematics. Springer New York, 2011.

[59] J. Douglas and H. H. Rachford, “On the numerical solution of heat conduction problems in two and three space variables,” *Transactions of the American Mathematical Society*, vol. 82, no. 2, pp. 421–439, 1956.

[60] M. Fazel, H. Hindi, and S. P. Boyd, “A rank minimization heuristic with application to minimum order system approximation,” in *Proceedings of the 2001 American Control Conference.*, vol. 6, 2001, pp. 4734–4739.

[61] C. Grussler and P. Giselsson, “Low-rank inducing norms with optimality interpretations,” *SIAM Journal on Optimization*, vol. 28, no. 4, pp. 3057–3078, 2018.

[62] A. M. McDonald, M. Pontil, and D. Stamos, “New perspectives on k-support and cluster norms,” *Journal of Machine Learning Research*, vol. 17, no. 155, pp. 1–38, 2016.

[63] C. Grussler, “Model reduction of positive systems,” M.Sc.Thesis, Lund University, Department of Automatic Control, 2012.

[64] K. Glover, “All optimal Hankel-norm approximations of linear multivariable systems and their L∞-error bounds,” *International Journal of Control*, vol. 39, no. 6, pp. 1115–1193, 1984.

[65] S. Gugercin, A. C. Antoulas, and C. Beattie, “H∞ model reduction for large-scale linear dynamical systems,” *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 609–638, 2008.

[66] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2012.