Abstract. This paper studies several problems related to quadratic matrix inequalities (QMI’s), i.e., inequalities in the Loewner order involving quadratic functions of matrix variables. In particular, we provide conditions under which the solution set of a QMI is nonempty, convex, bounded, or has nonempty interior. We also provide a parameterization of the solution set of a given QMI. In addition, we state results regarding the image of such sets under linear maps, which characterize a subset of “structured” solutions to a QMI. Thereafter, we derive matrix versions of the classical S-lemma and Finsler’s lemma, that provide conditions under which all solutions to one QMI also satisfy another QMI. The results will be compared to related work in the robust control literature, such as the full block S-procedure and Petersen’s lemma, and it is demonstrated how existing results can be obtained from the results of this paper as special cases. Finally, we show how the various results for QMI’s can be applied to the problem of data-driven stabilization. This involves separating the computation of the Lyapunov function and the controller, and also leads to explicit formulas for data-guided feedback gains.

Key words. Quadratic matrix inequalities, data-driven control, robust control.

AMS subject classifications. 93B51, 93B52, 93D21

1. Introduction. Designing control laws directly using measured data is becoming an increasingly prominent problem in systems and control [1, 9, 10, 12, 13, 17–19, 23–26, 31, 32, 45, 51, 52, 59]. This “one shot” paradigm has several potential advantages over the two-step procedure of system identification combined with model-based controller synthesis. For example, direct data-driven control is simpler from a conceptual point of view and could thus be the preferred choice for practitioners. It has also been argued that stabilizing controllers can be obtained from data that lack the richness assumptions typically imposed for system identification [58]. Despite these advantages, direct data-driven control schemes have their own issues. For example, since system models can be regarded as “condensed” representations of data, control methods that work directly with high-dimensional data matrices often suffer from a higher computational complexity than their model-based counterparts (see, e.g., [30]). Therefore, the problem of reducing computational complexity of direct methods is very relevant. Another important problem is to obtain (non-asymptotic) guarantees of the performance of direct control schemes, working with a finite number of noisy data samples, see e.g. [11, 16].

A fruitful and recent line of work that addresses the latter problem in the context of linear time-invariant (LTI) systems builds on the assumption that the matrix containing the noise samples is the solution to a quadratic matrix inequality (QMI) [2, 5, 8, 15, 28, 46, 47, 53–55]. Such a QMI can capture, for example, energy bounds on the noise or bounds on the noise sample covariance matrix. These bounds give rise to a set of LTI systems that “explain” the measured data, in the sense that each system in this set could have generated the measurements for some admissible noise sequence. Because of the assumption on the noise, the set of explaining systems can also be written as the set of solutions to a QMI. In this setting, direct data-driven control boils down to the robust control problem of finding a single controller that controls...
all systems explaining the data. Different control problems have been addressed in this manner, ranging from stabilization [2,5,15,47,54], to $H_2$ and $H_\infty$ control [2,46,54]. Also data-based analysis problems such as dissipativity have been studied [28, 55], as well as data-driven reduced order modeling [8]. The results are typically phrased in terms of data-dependent linear matrix inequalities (LMI’s). A theorem that is of interest in this context is a matrix version of the classical S-lemma [61], that was developed in [54]. This matrix S-lemma provides a tractable LMI condition under which all solutions to one QMI also satisfy another QMI. By noting that Lyapunov and dissipatin inequalities (with quadratic storage functions) can be regarded as QMI’s in the system matrices, the matrix S-lemma can be used to establish conditions under which all systems explaining the data (i.e., solutions of a QMI) also satisfy the performance specifications (i.e., are also solutions to a second QMI).

Despite this insight, we believe that there are certain aspects of QMI’s that are not yet completely understood. For example, the matrix S-lemma of [54] assumes a so-called generalized Slater condition, which turns out to not be satisfied in the special case that the data are noise-free. Regardless, the paper [53] shows that the LMI condition for stabilization in [54] still holds in the exact data setting. The latter result for exact data was established by deriving a matrix version of Finsler’s lemma, that provides conditions under which all solutions to a quadratic matrix equality satisfy a QMI. The paper [5] makes use of the so-called Petersen’s lemma, which enables the removal of the generalized Slater condition. Nonetheless, the results of [5] are only applicable to full rank data matrices, and to a specific noise model that does not capture all relevant cases (such as cross-covariance bounds, see [47]).

These observations suggest the existence of a general theory that encapsulates all of the above results. The purpose of this paper is to establish this theory. In fact, one of our main contributions is to derive matrix versions of the S-lemma and Finsler’s lemma that do not rely on the generalized Slater condition. We study both the cases of strict and nonstrict QMI’s. In the case of a strict inequality, we even show that the matrix Finsler’s lemma and S-lemma can be unified to obtain a single, more general result. In addition to these highlights, an important contribution of the paper is to give a full account of properties of quadratic matrix inequalities. In particular, we provide conditions under which the set of solutions to a QMI is nonempty, bounded, convex, or has nonempty interior. We study both strict and nonstrict QMI’s, as well as quadratic matrix equalities. Furthermore, we provide parameterizations of all solutions to a QMI. Thereafter, we study the image of solution sets of QMI’s under linear maps. This means that we are interested in matrices of the form $ZW$, where $Z$ satisfies a QMI and $W$ is a given matrix. We will show that, under suitable conditions, the set of all these matrices coincides with the solution set of yet another QMI. This result is relevant for data-driven control whenever the noise is contained in a known subspace of the state-space [54], or if certain prior knowledge of the system matrices is given [3]. Finally, we will study applications of the various results on QMI’s in the context of data-driven control. First, we will provide a general necessary and sufficient condition for data-based quadratic stabilization, using the unified matrix S-lemma and Finsler’s lemma. Subsequently, we apply the results regarding linear maps to reduce the computational complexity of data-based stabilization. In particular, we completely separate the computation of the quadratic Lyapunov function and the controller, which leads to lower-dimensional LMI’s. We also provide an explicit formula for a suitable stabilizing controller, given a (pre-computable) Lyapunov function. In the final part of the paper, we apply our results to resolve more intricate problems. First, we consider the nonlinear class of Lur’e systems, for which we derive stabilization results. After this, we consider the situation of linear systems with Gaussian noise. Our results will lead to conditions on the data which allow us
to design a controller that has a given probability of stabilizing the true system.

Throughout the paper we will compare our work to related results in the literature. For instance, in Section 3.2 we show the relation between [44, Cor. 2.3.6] and the QMI parameterization result. In Section 4.5, we compare the matrix S-lemma of this paper with a matrix S-lemma that can be derived from the full block S-procedure and the literature on LMI relaxations [39,40]. The conclusion is that the latter literature is able to deal with more general uncertainty descriptions than the one in this paper. Nonetheless, in the context of uncertainty described by QMI’s, the matrix S-lemma derived from [39,40] follows from this paper. We also show how the matrix S-lemma of this paper improves that of [54] by removing the generalized Slater condition. Finally, we show that both the strict and nonstrict Petersen’s lemma [35, 36], that were recently applied to data-driven control [5], can be obtained from our results.

The outline is as follows. In Section 2 we motivate our study of QMI’s with the application of data-based stabilization. Next, in Section 3 we study basic properties of solution sets of QMI’s, such as boundedness and convexity, and we establish the parameterization and results on images of solution sets of QMI’s. Section 4 provides matrix versions of the S-lemma and Finsler’s lemma. The results are applied to data-driven stabilization. Next, in Section 3 we study basic properties of solution sets of QMI’s, such as boundedness and convexity, and we establish the application of data-based stabilization. Finally, Section 5 contains our conclusions.

1.1. Notation. The Moore-Penrose pseudo-inverse of a real matrix $A$ is denoted by $A^\dagger$. The set of real symmetric $n \times n$ matrices is denoted by $\mathbb{S}^n$. In what follows, let $A \in \mathbb{S}^n$. The matrix $A$ is called positive semidefinite if $x^\top Ax \geq 0$ for all $x \in \mathbb{R}^n$ and positive definite if $x^\top Ax > 0$ for all nonzero $x \in \mathbb{R}^n$. This is denoted by $A \succeq 0$ and $A > 0$, respectively. Negative semidefiniteness and negative definiteness is defined in a similar way, and denoted by $A \preceq 0$ and $A < 0$, respectively. Throughout the paper, we will only consider definiteness conditions on symmetric matrices. As such, the notation $A \succeq 0$ implies symmetry of $A$. In addition, the notation $A \preceq B$ means that $A - B \preceq 0$ (equivalently, $B - A \succeq 0$), and $A < B$ means that $A - B < 0$. We denote by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ the smallest and largest eigenvalue of $A$, respectively. By [4, Cor. 8.4.2], we have that $\lambda_{\min}(A)I \preceq A \preceq \lambda_{\max}(A)I$. If $A \succeq 0$, there is exactly one positive semidefinite square root of $A$, denoted by $A^{\frac{1}{2}}$, such that $A = A^{\frac{1}{2}}A^{\frac{1}{2}}$.

2. Motivation. To demonstrate how quadratic matrix inequalities arise in the context of data-driven control, this section discusses data-based feedback stabilization. Consider input-state systems with unknown process noise of the form

\begin{equation}
(2.1) \quad x(t + 1) = A_s x(t) + B_s u(t) + w(t),
\end{equation}

where $x, w \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$. The matrices $A_s$ and $B_s$ and the noise $w(t)$ are unknown, but samples of the state and input are measured and collected in the matrices

\begin{equation}
(2.2) \quad X = [x(0) \ x(1) \ \cdots \ x(T)] \quad \text{and} \quad U_\sim = [u(0) \ u(1) \ \cdots \ u(T - 1)].
\end{equation}

We also define $X_\sim = [x(0) \ x(1) \ \cdots \ x(T - 1)]$, $X_\sim = [x(0) \ x(1) \ \cdots \ x(T - 1)]$, and $W_\sim = [w(0) \ w(1) \ \cdots \ w(T - 1)]$. The matrix $W_\sim$ is not known, but assumed to satisfy a quadratic matrix inequality, as explained next. Consider

\begin{equation}
\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \in \mathbb{S}^{n + T}, \quad \text{where} \quad \Phi_{11} \in \mathbb{S}^n, \ \Phi_{12} \in \mathbb{R}^{n \times T}, \ \Phi_{21} \in \mathbb{R}^{T \times n}, \ \Phi_{22} \in \mathbb{S}^T.
\end{equation}

We assume that the unknown noise $W_\sim$ satisfies

\begin{equation}
(2.3) \quad \begin{bmatrix} I \\ W_\sim^\top \end{bmatrix}^\top \Phi \begin{bmatrix} I \\ W_\sim^\top \end{bmatrix} \succeq 0.
\end{equation}

The above noise model captures, for instance,
(i) energy bounds: \( \Phi_{22} = -I \) and \( \Phi_{12} = 0 \) imply \( W_- W_+^\top = \sum_{t=0}^{T-1} w(t) w(t)^\top \leq \Phi_{11} \), which means that the energy of \( w \) on the time interval \([0, T-1]\) is bounded by \( \Phi_{11} \);

(ii) individual noise sample bounds: The noise bound (2.3) with \( \Phi_{22} = -I \), \( \Phi_{12} = 0 \) and \( \Phi_{11} = \varepsilon T I \) is implied by \( \|w(t)\|_2^2 \leq \varepsilon \forall t \). The latter means that the individual noise samples at every time instant are bounded in norm;

(iii) sample covariance bounds: The choices \( \Phi_{22} = \frac{1}{T}(-I + \frac{1}{T} I I^\top) \) and \( \Phi_{12} = 0 \) with \( \mu := \frac{1}{T} \sum_{t=0}^{T-1} w(t) \) lead to \( \frac{1}{T} \sum_{t=0}^{T-1} (w(t) - \mu) (w(t) - \mu)^\top = \frac{1}{T} W_- (I - \frac{1}{T} I I^\top) W_+^\top \leq \Phi_{11} \) where \( I \) denotes the \( T \)-vector of ones. In other words, the sample covariance matrix of \( w \) is bounded by \( \Phi_{11} \);

(iv) bounded noise within a subspace: Let \( E \in \mathbb{R}^{n \times d} \) have full column rank and let \( \hat{\Phi} \in \mathbb{S}^{d+\ell} \). Under suitable conditions, \( W_- = E \hat{W}_- \) for some \( \hat{W}_- \) satisfying

\[
\begin{bmatrix}
I \\
\hat{W}_+^\top
\end{bmatrix} \hat{\Phi} \begin{bmatrix}
I \\
\hat{W}_+^\top
\end{bmatrix} \geq 0 \quad \text{if and only if} \quad W_- \text{ and } \hat{\Phi} := \begin{bmatrix}
E & 0 \\
0 & I
\end{bmatrix} \hat{\Phi} \begin{bmatrix}
E^\top & 0 \\
0 & I
\end{bmatrix}
\]

satisfy (2.3) (see Section 5). This matrix \( \hat{\Phi} \) thus captures the situation that the noise is contained in the subspace \( \text{im} E \), and equal to \( EW_- \), where \( \hat{W}_- \) again satisfies a quadratic matrix inequality;

(v) exact measurements: \( \Phi_{11} = 0, \Phi_{12} = 0, \) and \( \Phi_{22} = -I \) leads to \( W_-=0 \).

The set \( \Sigma \) of all systems compatible with the data is given by

\[
\Sigma = \{(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} | X_+ = AX_- + BU_- + W_- \text{ for some } W_- \text{ satisfying (2.3)}\}.
\]

Since the measurements are obtained from the system (2.1), we clearly have that \((A_0, B_0) \in \Sigma\). Since on the basis of the given data we are not able to distinguish the true system \((A_0, B_0)\) from any other system in \( \Sigma \), a stabilizing controller computed using only the data should necessarily stabilize all systems in \( \Sigma \). This observation leads to the following notion of informativity for quadratic stabilization.

**Definition 2.1.** The data \((U_0, X)\) are called informative for quadratic stabilization if there exists a feedback gain \( K \) and a matrix \( P > 0 \) such that for all \((A, B) \in \Sigma\):

\[
P - (A + BK)^\top P (A + BK)^\top > 0.
\]

Note that \( P > 0 \) satisfies (2.4) if and only \( Q := P^{-1} \) satisfies \( Q - (A + BK)^\top Q (A + BK) > 0 \), which expresses that \( V(x) = x^\top Q x \) is a Lyapunov function for the closed loop system \( x(t + 1) = (A + BK)x(t) \).

Finding a controller that stabilizes all systems in a given set is a problem which has been studied under a number of different names in the literature. One of the most common of these is robust control, where the set of systems takes the form of a nominal system and a set of perturbations. An overview of such methods pertaining to linear systems can be found in [21]. More specifically an LMI-based approach to this problem can be found in [41]. The problem is also studied under the name of simultaneous control, an overview of which can be found in [6, 34]. Apart from being studied on its own merits, results regarding sets of systems have been applied in the study of, for instance linear parameter varying systems [33] or switched systems [29].

Taking inspiration from the aforementioned works, we are interested in quadratic stabilization in the sense that we ask for a common Lyapunov matrix \( P \) for all \((A, B) \in \Sigma\). This assumption is quite customary in data-driven [2, 15, 54] and robust [43] control, and is made mainly because it leads to tractable solutions. This assumption is, however, often conservative in the sense that there may exist a parameter-dependent Lyapunov function in cases that no common Lyapunov function exists. In [56], a first attempt was made to quantify the conservatism of common Lyapunov functions in the context of this paper.

Recall that \((A, B) \in \Sigma \) if and only if \( W_- = X_+ - AX_- - BU_- \), where \( W_- \) is such that (2.3) holds. Combining this yields that \( (A, B) \in \Sigma \) if and only if \((A, B) \) satisfies

**This manuscript is for review purposes only.**
3. Sets induced by quadratic matrix inequalities. In this section, we are interested in properties of sets defined in terms of quadratic matrix inequalities. We begin with sets of the form

\[ Z \in \mathbb{R}^{r \times q} : \begin{bmatrix} I_q & Z \end{bmatrix}^T \begin{bmatrix} I_q & Z \end{bmatrix} \geq 0, \]

where \( I_q \) is the identity matrix of size \( q \times q \). This set is denoted as \( \mathcal{Z}_r(\Pi) \):

\[ \mathcal{Z}_r(\Pi) := \left\{ Z \in \mathbb{R}^{r \times q} : \begin{bmatrix} I_q & Z \end{bmatrix}^T \begin{bmatrix} I_q & Z \end{bmatrix} \geq 0 \right\}. \]
for $\Pi \in S^{q+r}$. The very first question one may ask is under what conditions on $\Pi$ the set $Z_r(\Pi)$ is nonempty. An immediate necessary condition is that $\Pi$ must have at least $q$ nonnegative eigenvalues. Clearly, this is not sufficient in general. What is necessary and sufficient for $Z \in Z_r(\Pi)$ is that the matrix

$$
(3.1) \Pi = \begin{bmatrix}
\Pi_{11} & \Pi_{12} & -Z^T \\
\Pi_{21} & \Pi_{22} & I \\
-Z & I & 0
\end{bmatrix}, \quad \text{where} \quad \Pi = \begin{bmatrix}
\Pi_{11} & \Pi_{12} \\
\Pi_{21} & \Pi_{22}
\end{bmatrix}
$$

has exactly $r$ negative eigenvalues (see [4, Fact 8.15.28]). From now on, whenever we partition a matrix $\Pi \in S^{q+r}$, we assume that $\Pi_{11}$ is $q \times q$ and $\Pi_{22}$ is $r \times r$. The condition on the eigenvalues of (3.1) does not translate to an easily verifiable condition on $\Pi$ for nonemptiness of $Z_r(\Pi)$. Nevertheless, it leads to a noteworthy dualization result. To state this result, for given $S \subseteq \mathbb{R}^{r \times q}$, we define $S^\top := \{Z^\top : Z \in S\}$.

**Proposition 3.1.** Let $\Pi \in S^{q+r}$ be nonsingular and $Z_r(\Pi) \neq \emptyset$. Then,

$$(Z_\Pi(\Pi))^\top = Z_q(\Pi_{r,q}^\dagger), \text{ where } \Pi_{r,q}^\dagger := \begin{bmatrix} 0 & -I_r \\ I_q & 0 \end{bmatrix} \Pi^{-1} \begin{bmatrix} 0 & -I_q \\ I_r & 0 \end{bmatrix}.$$

A proof can be found in [43, Lem. 4.9]. An alternative proof was given in [55, Lem. 3] based on the Schur complements of (3.1) with respect to $\Pi$ and $\Pi_{22}$.

It turns out that for particular matrices $\Pi$, a Schur complement argument on the matrix $\Pi$ itself leads to a simple characterization of nonemptiness of the set $Z_r(\Pi)$. Specifically, suppose that $\Pi_{22} \preceq 0$ and $\ker \Pi_{22} \subseteq \ker \Pi_{12}$. Since the latter condition is equivalent to $\Pi_{12} \Pi_{22} \Pi_{22}^\dagger = \Pi_{12}$, we have that

$$
(3.2) \Pi = \begin{bmatrix}
\Pi_{11} & \Pi_{12} \\
\Pi_{21} & \Pi_{22}
\end{bmatrix} = \begin{bmatrix} I_q & \Pi_{12} \Pi_{22}^\dagger \\ 0 & \Pi_{22} \end{bmatrix} \begin{bmatrix} \Pi_{11} & 0 \\ 0 & \Pi_{22} \end{bmatrix} \begin{bmatrix} I_q & 0 \\ \Pi_{12}^\dagger \Pi_{21} & I_r \end{bmatrix},
$$

where $\Pi |_{\Pi_{22}} := \Pi_{11} - \Pi_{12} \Pi_{22}^\dagger \Pi_{21}$ is the (generalized) Schur complement of $\Pi$ with respect to $\Pi_{22}$. This results in

$$
(3.3) \Pi |_{\Pi_{22}} = \begin{bmatrix} I_q & \Pi_{12} \Pi_{22}^\dagger \\ 0 & \Pi_{22}\end{bmatrix} \begin{bmatrix} I_q \\ -\Pi_{22}^\dagger \Pi_{21} \end{bmatrix} = \Pi |_{\Pi_{22}} + (Z + \Pi_{22}^\dagger \Pi_{21}) \Pi_{22} (Z + \Pi_{22}^\dagger \Pi_{21})
$$

because $(\Pi_{22}^\top)^\dagger = (\Pi_{22}^\top)$ by symmetry of $\Pi_{22}$ and, since $\Pi_{22} \preceq 0$,

$$
(3.4) \Pi |_{\Pi_{22}} = \begin{bmatrix} I_q & \Pi_{12} \Pi_{22}^\dagger \\ 0 & -\Pi_{22}^\dagger \Pi_{21} \end{bmatrix} = \begin{bmatrix} I_q & \Pi_{12} \Pi_{22}^\dagger \\ 0 & -\Pi_{22}^\dagger \Pi_{21} \end{bmatrix} \begin{bmatrix} I_q \\ -\Pi_{22}^\dagger \Pi_{21} \end{bmatrix} \Pi_{22} \begin{bmatrix} I_q \\ -\Pi_{22}^\dagger \Pi_{21} \end{bmatrix} \Pi_{22} \begin{bmatrix} I_q \\ -\Pi_{22}^\dagger \Pi_{21} \end{bmatrix} Z \Pi_{22} \Pi_{22}^\dagger \Pi_{21}
$$

for any $Z \in \mathbb{R}^{r \times q}$. An immediate consequence of this inequality is that $Z_r(\Pi) \neq \emptyset$ if and only if $\Pi |_{\Pi_{22}} \succeq 0$. Motivated by this observation, we define the set

$$
(3.5) \Pi_{q,r} = \left\{ \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix} \in S^{q+r} : \Pi_{22} \preceq 0, \Pi |_{\Pi_{22}} \succeq 0 \text{ and } \ker \Pi_{22} \subseteq \ker \Pi_{12} \right\}.
$$

All the noise models mentioned after (2.3) can be captured by elements of $\Pi_{q,r}$.

Next, for $\Pi \in S^{q+r}$ we will investigate properties of $Z_r(\Pi)$ and the sets

$$
Z_r^+(\Pi) := \left\{ Z \in \mathbb{R}^{r \times q} : \begin{bmatrix} I_q \\ Z \end{bmatrix} \Pi \begin{bmatrix} I_q \\ Z \end{bmatrix} > 0 \right\}, \quad Z_r^0(\Pi) := \left\{ Z \in \mathbb{R}^{r \times q} : \begin{bmatrix} I_q \\ Z \end{bmatrix} \Pi \begin{bmatrix} I_q \\ Z \end{bmatrix} = 0 \right\}.
$$

**3.1. Basic properties.** In the following theorem, we study nonemptiness, convexity and boundedness of the sets induced by QMI’s.

**Theorem 3.2.** Let $\Pi \in \Pi_{q,r}$. Then, $Z_r(\Pi)$

(a) is nonempty and convex.
with \[\|\text{equation}\] for all \(D\) matrix. Therefore, either \(\Pi\) particular, consider \(\xi\) such that \(Z\) − \(\text{rank} \, \Pi\) of a product of matrices is less than or equal to the ranks of individual matrices, where the last inequality follows from the facts that \(\text{rank}(\Pi)\) proves nonemptiness whereas convexity readily follows from (3.3) that \(\Pi\) is bounded if and only if \(\Pi\mid \Pi_{\text{ii}} > 0\). Let \(\Pi\mid \Pi_{\text{ii}} = 0\). Then, it follows from (3.3) that \((Z + \Pi_{\text{ii}}\Pi_{\text{ii}})\) nonempty. Let \(\Pi\mid \Pi_{\text{ii}} > 0\). Then, it follows from (3.3) that \(Z_{\text{ii}}^+ (\Pi)\) is nonempty. The proof is completed.\(\Box\)\n
Proof. (a): Since \(\Pi\mid \Pi_{\text{ii}} \geq 0\), it follows from (3.4) that \(-\Pi_{\text{ii}}\Pi_{\text{ii}} \in Z_{\text{r}}(\Pi)\). This proves nonemptiness whereas convexity readily follows from (3.3) that \(\Pi\) is bounded if and only if \(\Pi\mid \Pi_{\text{ii}} > 0\). Let \(\Pi\mid \Pi_{\text{ii}} = 0\). Then, it follows from (3.3) that \((Z + \Pi_{\text{ii}}\Pi_{\text{ii}})\) nonempty. Let \(\Pi\mid \Pi_{\text{ii}} > 0\). Then, it follows from (3.3) that \(Z_{\text{ii}}^+ (\Pi)\) is nonempty. The proof is completed.\(\Box\)\n
(b): We first prove the ‘if’ part. Let \(Z \in Z_{\text{r}}(\Pi)\). Then, it follows from (3.3) that \((Z + \Pi_{\text{ii}}\Pi_{\text{ii}})\) nonempty. Let \(\Pi\mid \Pi_{\text{ii}} > 0\). Then, it follows from (3.3) that \(Z_{\text{ii}}^+ (\Pi)\) is nonempty. The proof is completed.\(\Box\)
3.2. Parameterization of $Z_r(\Pi)$ and $Z_r^+(\Pi)$. It turns out that one can parameterize all solutions of a given QMI associated with $\Pi \in \Pi_{q,r}$.

**Theorem 3.3.** Let $\Pi \in \Pi_{q,r}$. The following statements hold:
(a) $Z \in Z_r(\Pi)$ if and only if there exist $S, T \in \mathbb{R}^{r \times q}$ with $S^T S \leq I$ such that

$$
Z = -\Pi_2^\dagger \Pi_{21} + ((-\Pi_2^\dagger))^{\dagger} S (\Pi \Pi_{22})^{\dagger} + (I - \Pi_2^\dagger \Pi_{22}) T. \tag{3.7}
$$

(b) Assume that $Z_r^+(\Pi)$ is nonempty, equivalently, $\Pi \Pi_{22} > 0$. Then, $Z \in Z_r^+(\Pi)$ if and only if $Z$ is of the form (3.7) for some $S, T \in \mathbb{R}^{r \times q}$ with $S^T S < I$.

We note that, in the special case that $\Pi_{22} < 0$, Theorem 3.3.(b) also follows from [44, Corollary 2.3.6] by taking $A = \Pi_{12} \Pi_2^\dagger$, $B = I$, $Q = \Pi \Pi_{22}$, $R = -\Pi_2$ and $X = Z^T$ in that result. Here we provide a proof in the case that $\Pi_{22}$ is not necessarily negative definite and hence $Z_r(\Pi)$ is not necessarily bounded.

**Proof.** We first prove (b). From (3.3) we have that $Z \in Z_r^+(\Pi)$ if and only if

$$
(Z + \Pi_2^\dagger \Pi_{21})^T (\Pi_2^\dagger) (Z + \Pi_2^\dagger \Pi_{21}) < \Pi^T \Pi_{22}. \tag{3.8}
$$

By Lemma A.1(b), we then have that $Z \in Z_r^+(\Pi)$ if and only if there exists a matrix $S$ such that $S^T S < I$ and $(-\Pi_2^\dagger)^T (Z + \Pi_2^\dagger \Pi_{21}) = S (\Pi \Pi_{22})^{\dagger}$. Using the fact that $\ker(-\Pi_2^\dagger) = \ker \Pi_{22}$, and by exploiting Lemma A.2, we see that this is equivalent to $Z + \Pi_2^\dagger \Pi_{21} = ((-\Pi_2^\dagger)\Pi \Pi_{22})^{\dagger} S (\Pi \Pi_{22})^{\dagger} + (I - \Pi_2^\dagger \Pi_{22}) T$ for some $T \in \mathbb{R}^{r \times q}$. This proves (b). The proof of (a) follows the same arguments but instead invokes Lemma A.1(a).

3.3. Image of $Z_r(\Pi)$ and $Z_r^+(\Pi)$ under linear maps. Let $W \in \mathbb{R}^{p \times r}$. For $S \subseteq \mathbb{R}^{r \times q}$, we define $SW := \{SW : S \in S\}$. Also, for $\Pi \in S^{p \times r}$ we define

$$
\Pi_W := \begin{bmatrix}
W^T & 0 \\
0 & I_r
\end{bmatrix} \Pi \begin{bmatrix}
W & 0 \\
0 & I_r
\end{bmatrix} = \begin{bmatrix}
W^T \Pi_{11} W & W^T \Pi_{12} \\
\Pi_{21} W & \Pi_{22}
\end{bmatrix} \in S^{p \times r}. \tag{3.9}
$$

Note that $\Pi_W \in \Pi_{p,r}$ if $\Pi \in \Pi_{q,r}$. Next, we will study the images $Z_r(\Pi)$ and $Z_r(\Pi_W)$.

**Theorem 3.4.** Let $\Pi \in \Pi_{q,r}$ and $W \in \mathbb{R}^{p \times r}$. We have that $Z_r(\Pi)W \subseteq Z_r(\Pi_W)$. Assume, in addition, that either $W$ has full column rank or $\Pi_{22}$ is nonsingular. Then, $Z_r(\Pi)W = Z_r(\Pi_W)$.

**Proof.** First we prove that $Z_r(\Pi)W \subseteq Z_r(\Pi_W)$. Let $Z' \in Z_r(\Pi)W$. Then, $Z' = ZW$ where $Z \in Z_r(\Pi)$. Note that

$$
\begin{bmatrix}
I_q \\
Z
\end{bmatrix}^T \Pi \begin{bmatrix}
I_q \\
Z
\end{bmatrix} \geq 0, \quad \text{implies that } W^T \begin{bmatrix}
I_q \\
Z
\end{bmatrix}^T \Pi \begin{bmatrix}
I_q \\
Z
\end{bmatrix} W = \begin{bmatrix}
I_p \\
Z'
\end{bmatrix}^T \Pi_W \begin{bmatrix}
I_p \\
Z'
\end{bmatrix} \geq 0.
$$

This means that $Z' \in Z_r(\Pi_W)$ and hence $Z_r(\Pi)W \subseteq Z_r(\Pi_W)$.

Now, we assume that at least one of the conditions on $W$ and $\Pi_{22}$ hold. We claim that $Z_r(\Pi_W) \subseteq Z_r(\Pi)W$. Let $Z' \in Z_r(\Pi_W)$. Note that $\Pi_W \Pi_{22} = W^T (\Pi \Pi_{22}) W$. From (3.9) and Theorem 3.3(a), we see that

$$
Z' = -\Pi_2^\dagger \Pi_{21} W + ((-\Pi_2^\dagger))^{\dagger} S (W^T (\Pi \Pi_{22}) W)^{\dagger} + (I - \Pi_2^\dagger \Pi_{22}) V
$$

where $V, S \in \mathbb{R}^{p \times r}$ with $S^T S \leq I_p$. Since $(W^T (\Pi \Pi_{22}) W)^{\dagger} (W^T (\Pi \Pi_{22}) W)^{\dagger} = W^T (\Pi \Pi_{22})^{\dagger} (\Pi \Pi_{22})^{\dagger} W$, we have that $(W^T (\Pi \Pi_{22}) W)^{\dagger} = T (\Pi \Pi_{22})^{\dagger} W$ where $T \in \mathbb{R}^{p \times q}$ is such that $T^T T \leq I_q$ due to Lemma A.1. If $W$ has full column rank then (3.10) results in $Z' = ZW$ where

$$
Z := -\Pi_2^\dagger \Pi_{21} + ((-\Pi_2^\dagger))^{\dagger} S T (\Pi \Pi_{22})^{\dagger} + (I - \Pi_2^\dagger \Pi_{22}) V(W^T W)^{-1} W^T.
$$

On the other hand, if $\Pi_{22}$ is nonsingular then $I - \Pi_2^\dagger \Pi_{22} = 0$ and $Z' = ZW$ with
In either of these two cases, we observe that \( T^T S^T ST \leq T^T T \leq I_q \). Therefore, Theorem 3.3(a) implies that \( Z \in Z_r(II) \). Consequently, we see that \( Z' = ZW \) for some \( Z \in Z_r(II) \) and thus \( Z_r(II_W) \subseteq Z_r(II)W \). This proves the theorem. □

A similar result holds for the sets \( Z^+_r(II) \) and \( Z^+_r(II_W) \), as shown next.

**Theorem 3.5.** Let \( \Pi \in \Pi_{q,r} \) and \( W \in \mathbb{R}^{q \times p} \). Assume that \( W \) has full column rank and \( Z^+_r(II) \) is nonempty. Then, \( Z^+_r(II)W = Z^+_r(II_W) \).

The proof of Theorem 3.5 is similar to that of Theorem 3.4, but applies Theorem 3.3(b) instead of Theorem 3.3(a). The following two corollaries follow from Theorems 3.4 and 3.5 and provide conditions under which there exists a “structured” matrix in \( Z_r(II) \) (respectively, \( Z^+_r(II) \)) that satisfies a linear equation.

**Corollary 3.6.** Let \( \Pi \in S^{n+r} \) with \( \Pi_{22} \leq 0 \) and \( \ker \Pi_{22} \subseteq \ker \Pi_{21} \), \( W \in \mathbb{R}^{q \times p} \), and \( Y \in \mathbb{R}^{r \times p} \). Suppose that either \( W \) has full column rank or \( \Pi_{22} \) is nonsingular. Then there exists a \( Z \in Z_r(II) \) such that \( ZW = Y \) if and only if \( \Pi \in \Pi_{q,r} \) and \( Y \in Z_r(II_W) \).

**Proof.** To prove the ‘if’ statement, suppose that \( \Pi \in \Pi_{q,r} \) and \( Y \in Z_r(II_W) \). By Theorem 3.4 there exists a \( Z \in Z_r(II) \) such that \( ZW = Y \).

To prove the ‘only if’ statement, suppose that there exists a \( Z \in Z_r(II) \) satisfying \( ZW = Y \). Therefore, \( Z_r(II) \) is nonempty and (3.3) implies that \( \Pi \mid \Pi_{22} = 0 \). Consequently, \( \Pi \in \Pi_{q,r} \). Finally, \( Y \in Z_r(II_W) \) follows directly from multiplying the defining quadratic matrix inequality from left by \( W^T \) and right by \( W \). □

**Corollary 3.7.** Let \( \Pi \in S^{n+r} \) with \( \Pi_{22} \leq 0 \) and \( \ker \Pi_{22} \subseteq \ker \Pi_{21} \). Consider \( W \in \mathbb{R}^{q \times p} \) and \( Y \in \mathbb{R}^{r \times p} \). Assume that \( W \) has full column rank. Then there exists a matrix \( Z \in Z^+_r(II) \) satisfying \( ZW = Y \) if and only if \( \Pi \mid \Pi_{22} > 0 \) and \( Y \in Z^+_r(II_W) \).

The proof of Corollary 3.7 follows the same lines as that of Corollary 3.6, but applies Theorem 3.5 rather than Theorem 3.4. It is therefore omitted. Corollary 3.7 is intimately related to the so-called elimination lemma [22,39]. In fact, in the case that \( \Pi \) is nonsingular and has \( r \) negative and \( q \) positive eigenvalues, Corollary 3.7 can also be obtained from [39, Lem. A.2] by taking \( P = -\Pi, A = I, B = W^\perp \) and \( C = YW^\dag \) where \( W^\perp \in \mathbb{R}^{(q-r)\times q} \) is any full row rank matrix such that \( W^\perp W = 0 \).

### 4. Matrix versions of Finsler’s lemma and Yakubovich’s S-lemma.

In this section we deal with the question under what conditions all solutions to one quadratic matrix inequality also satisfy another QMI. We aim at finding necessary and sufficient conditions for the inclusion \( Z_r(N) \subseteq Z_r(M) \), where \( M, N \in S^{n+r} \). We will also consider \( Z^+_r(N) \) instead of \( Z_r(N) \), and \( Z^+_r(M) \) replacing \( Z_r(M) \). This leads to non-strict and strict versions of Yakubovich’s S-lemma and Finsler’s lemma.

#### 4.1. Recap of standard S-lemma and Finsler’s lemma.

For future reference, we will start with a brief recap of “standard” (vector-valued) S-lemmas and Finsler’s lemma. The idea behind all of these results is that certain implications involving quadratic inequalities and equalities can be reformulated as linear matrix inequalities. The following statement is the S-lemma for non-strict inequalities, which was first proven by Yakubovich in [61], see also the survey paper [37].

**Lemma 4.1 (S-lemma).** Let \( M, N \in S^n \) and suppose that \( N \) has at least one positive eigenvalue. Then \( x^T Mx \geq 0 \) for all \( x \in \mathbb{R}^n \) satisfying \( x^T Nx \geq 0 \) if and only if there exists a real number \( \alpha \geq 0 \) such that \( M - \alpha N \geq 0 \).

Next, we recall a variant with a strict inequality on \( x^T Mx \), see [61] and [7, p. 24].

**Lemma 4.2 (Strict S-lemma).** Let \( M, N \in S^n \) and suppose that \( N \) has at least one positive eigenvalue. Then \( x^T Mx > 0 \) for all nonzero \( x \in \mathbb{R}^n \) satisfying \( x^T Nx \geq 0 \) if and only if there exists a real number \( \alpha \geq 0 \) such that \( M - \alpha N > 0 \).
The strict S-lemma has also been generalized [27] from the set \( \{ \alpha N : \alpha \geq 0 \} \) of multiplier matrices to a more general set satisfying the so-called losslessness property. Finally, we recall Finsler’s lemma [20] which involves an equality \( x^\top N x = 0 \). We state the result for a strict inequality on \( x^\top M x \). We note that also a non-strict version of the result exists (see e.g., [62]), but this will not be used in this paper.

**Lemma 4.3** (Finsler’s lemma). Let \( M, N \in S^n \). Then \( x^\top M x > 0 \) for all nonzero \( x \in \mathbb{R}^n \) satisfying \( x^\top N x = 0 \) if and only if there exists a real number \( \alpha \in \mathbb{R} \) such that \( M - \alpha N > 0 \).

### 4.2. Reduction of the matrix case to the vector case.
Throughout this section, we will consider matrices \( M, N \in S^{n \times n} \) partitioned as

\[
M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}.
\]

We will provide conditions under which the inclusion \( Z_r(N) \subseteq Z_r(M) \) is equivalent to the vector-valued implication \( x^\top N x \geq 0 \implies x^\top M x \geq 0 \). This will provide an important building block in obtaining matrix versions of the S-lemma. To proceed, we will need the following lemma.

**Lemma 4.4.** Let \( S \in S^n, S \geq 0 \). Given a nonzero vector \( x \in \mathbb{R}^n \), there exists a matrix \( \bar{X} \in \mathbb{R}^{n \times (n-1)} \) such that \( x^\top S \bar{X} = 0 \) and \( [x \quad \bar{X}] \) is nonsingular.

**Proof.** If \( x^\top S = 0 \) the statement is immediate. Thus, assume that \( x^\top S \neq 0 \). Let \( \bar{X} \in \mathbb{R}^{n \times (n-1)} \) be a matrix whose columns form a basis for \( \ker x^\top S \). If \( [x \quad \bar{X}] \) is singular, then \( x \in \text{im} \bar{X} \) and, hence, \( x^\top S x = 0 \). However, since \( S \) is symmetric and positive semidefinite, this implies that \( x^\top S = 0 \). This yields a contradiction, and we conclude that \( [x \quad \bar{X}] \) is nonsingular. This proves the lemma.

Next, we state the following instrumental lemma that is an extension of [40, Lemma A.2] to the set of matrices \( \Pi_{q,r} \). Note that we do not require the matrix \( N \) to satisfy \( N_{22} < 0 \) and \( N_{11} - N_{12}N_{22}^{-1}N_{21} > 0 \).

**Lemma 4.5.** Let \( N \in \Pi_{q,r} \). Let \( x \in \mathbb{R}^q \) and \( y \in \mathbb{R}^r \) be vectors, with \( x \) nonzero, such that \( \begin{bmatrix} x \\ y \end{bmatrix} N \begin{bmatrix} x \\ y \end{bmatrix} \geq 0 \). Then there exists a matrix \( Z \in Z_r(N) \) such that \( y = Zx \).

**Proof.** Since \( x \) is nonzero and \( N \mid N_{22} \geq 0 \), we conclude from Lemma 4.4 that there exists a matrix \( \bar{X} \in \mathbb{R}^{q \times (q-1)} \) such that \( x^\top (N \mid N_{22}) \bar{X} = 0 \) and \( [x \quad \bar{X}] \) is nonsingular. Define the matrix \( \bar{Y} := -N_{22}^{-1}N_{21} \bar{X} \). Note that

\[
\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} = \begin{bmatrix} I & N_{12}N_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} N | N_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & N_{22} \end{bmatrix} \begin{bmatrix} I \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}.
\]

Therefore, we have

\[
\begin{bmatrix} x \\ y \end{bmatrix} N \begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} N \begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} = \begin{bmatrix} \bar{X}^\top (N | N_{22}) \bar{X} \end{bmatrix} \geq 0
\]

since \( N \mid N_{22} \geq 0 \). The latter two results imply that

\[
\begin{bmatrix} x \\ \bar{X} \end{bmatrix}^\top N \begin{bmatrix} x \\ \bar{X} \end{bmatrix} = \begin{bmatrix} x^\top N x & 0 \\ 0 & \bar{X}^\top N \bar{X} \end{bmatrix} \geq 0.
\]

Recall that \( [x \quad \bar{X}] \) is nonsingular. Thus, the matrix \( Z := \begin{bmatrix} y & \bar{Y} \end{bmatrix} [x \quad \bar{X}]^{-1} \) is a member of \( Z_r(N) \). In addition, note that \( Zx = y \). This proves the lemma.

---

This manuscript is for review purposes only.
The following theorem provides conditions under which \( Z_r(N) \subseteq Z_r(M) \) and \( Z_r(N) \subseteq Z_r^+(M) \) are equivalent to their respective vector-valued implications.

**Theorem 4.6.** Let \( M, N \in S_{q+r}^+ \) with \( N \in \Pi_{q,r}^+ \).

(a) Assume that \( N \) has at least one positive eigenvalue. Then the following two statements are equivalent:

(i) \( Z_r(N) \subseteq Z_r(M) \),

(ii) \( z^T M z \geq 0 \) for all \( z \in \mathbb{R}^{q+r} \) satisfying \( z^T N z \geq 0 \).

(b) Assume that \( N_{22} < 0 \). Then the following two statements are equivalent:

(i) \( Z_r(N) \subseteq Z_r^+(M) \),

(ii) \( z^T M z > 0 \) for all nonzero \( z \in \mathbb{R}^{q+r} \) satisfying \( z^T N z \geq 0 \).

**Proof.** We first prove that (a).i) implies (a).ii). Assume that (a).i) holds but, on the contrary, (a).ii) does not hold. This implies that there exist vectors \( x \in \mathbb{R}^q \) and \( y \in \mathbb{R}^r \), not both zero, such that

\[
\begin{bmatrix}
  x \\
  y \\
\end{bmatrix}^T \begin{bmatrix}
  x \\
  y \\
\end{bmatrix} \geq 0 \quad \text{and} \quad \begin{bmatrix}
  x \\
  y \\
\end{bmatrix}^T M \begin{bmatrix}
  x \\
  y \\
\end{bmatrix} < 0.
\]

We claim that there exists a pair \((x, y)\) satisfying (4.2) with \( x \neq 0 \).

To see this, suppose that \( x = 0 \) and \( y \) satisfy (4.2). We will use these vectors to construct a new pair \((\bar{x}, \bar{y})\) satisfying (4.2) with \( \bar{x} \neq 0 \). By the hypothesis that \( N_{22} \leq 0 \), this implies that \( N_{22} y = 0 \). In addition, since \( \ker N_{22} \subseteq \ker N_{12} \) we have that \( N \begin{bmatrix}
  0 \\
  y \\
\end{bmatrix} = 0 \). Let \( \begin{bmatrix}
  \bar{x}^T \\
  \bar{y}^T \\
\end{bmatrix} \) be an eigenvector of \( N \) corresponding to a positive eigenvalue \( \lambda \). Note that \( \bar{x} \neq 0 \) because \( N_{22} \leq 0 \). The previous implies that

\[
\begin{bmatrix}
  0 \\
  y \\
\end{bmatrix}^T + \epsilon \begin{bmatrix}
  \bar{x} \\
  \bar{y} \\
\end{bmatrix}^T N \left( \begin{bmatrix}
  0 \\
  y \\
\end{bmatrix} + \epsilon \begin{bmatrix}
  \bar{x} \\
  \bar{y} \\
\end{bmatrix} \right) - \epsilon^2 \lambda \begin{bmatrix}
  \bar{x}^T \\
  \bar{y}^T \\
\end{bmatrix} \begin{bmatrix}
  \bar{x} \\
  \bar{y} \\
\end{bmatrix} \geq 0
\]

for any \( \epsilon \in \mathbb{R} \). In addition,

\[
\begin{bmatrix}
  0 \\
  y \\
\end{bmatrix}^T + \epsilon \begin{bmatrix}
  \bar{x} \\
  \bar{y} \\
\end{bmatrix}^T M \left( \begin{bmatrix}
  0 \\
  y \\
\end{bmatrix} + \epsilon \begin{bmatrix}
  \bar{x} \\
  \bar{y} \\
\end{bmatrix} \right) < 0
\]

if \( \epsilon \) is sufficiently small. Therefore, for sufficiently small \( \epsilon \neq 0 \), the pair \((\epsilon \bar{x}, y + \epsilon \bar{y})\) satisfies (4.2). As \( \bar{x} \neq 0 \), there exists a pair \((\bar{x}, \bar{y}) := (\epsilon \bar{x}, y + \epsilon \bar{y})\) satisfying (4.2) with \( \bar{x} \neq 0 \). By Lemma 4.5 there exists a matrix \( Z \in Z_r(N) \) such that \( y = Z \bar{x} \). By (4.3) we see that \( \bar{x}^T \begin{bmatrix}
  I \\
  Z \\
\end{bmatrix} M \begin{bmatrix}
  I \\
  Z \\
\end{bmatrix} \bar{x} < 0 \), that is, \( Z \not\subseteq Z_r(M) \). This, however, contradicts the assumption that \( Z_r(N) \subseteq Z_r(M) \). Therefore, we conclude that (a).ii) holds.

Next, we prove that (b).i) implies (b).ii). Therefore, assume that (b).i) holds but, on the contrary, (b).ii) does not hold. This implies that there exist vectors \( x \in \mathbb{R}^q \) and \( y \in \mathbb{R}^r \), not both zero, such that \( x^T N x \geq 0 \) and \( x^T M x \leq 0 \). This implies that \( x \neq 0 \). Indeed, if \( x = 0 \) then also \( y = 0 \) by the hypothesis that \( N_{22} < 0 \). Thus, by Lemma 4.5 there exists a matrix \( Z \in Z_r(N) \) such that \( y = Z x \). Combining the previous yields \( x^T \begin{bmatrix}
  I \\
  Z \\
\end{bmatrix} M \begin{bmatrix}
  I \\
  Z \\
\end{bmatrix} x \leq 0 \), that is, \( Z \not\subseteq Z_r^+(M) \). This contradicts the hypothesis that \( Z_r(N) \subseteq Z_r^+(M) \). This shows that (b).ii) holds.

Next, we prove that (b).ii) implies (a).i). Suppose that \( Z \in Z_r(N) \). Then:

\[
y^T \begin{bmatrix}
  I \\
  Z \\
\end{bmatrix} N \begin{bmatrix}
  I \\
  Z \\
\end{bmatrix} y \geq 0, \quad \text{and} \quad y^T \begin{bmatrix}
  I \\
  Z \\
\end{bmatrix} M \begin{bmatrix}
  I \\
  Z \\
\end{bmatrix} y \geq 0
\]

for all \( y \in \mathbb{R}^q \). In other words, \( Z \in Z_r(M) \). The proof that (b).ii) implies (b).i) is analogous and therefore omitted. This proves the theorem. \( \square \)
4.3. Matrix versions of the S-lemma and Finsler’s lemma for non-strict inequalities. In the following theorem we apply the results of the previous section to establish a matrix version of the S-lemma.

Theorem 4.7 (Matrix S-lemma). Let $M, N \in \mathbb{S}_{q+r}$. If there exists a real $\alpha \geq 0$ such that $M - \alpha N \geq 0$, then $Z_r(N) \subseteq Z_r(M)$.

Proof. The ‘if’ statements are obvious. Now, assume that $Z_r(N) \subseteq Z_r(M)$. By Theorem 4.6, $x^T M x \geq 0$ for all $x \in \mathbb{R}^{q+r}$ satisfying $x^T N x \geq 0$. Finally, by Lemma 4.1, we conclude that there exists a scalar $\alpha \geq 0$ such that $M - \alpha N \geq 0$.

Similar to the ‘standard’ S-lemma (Lemma 4.1), we note that the matrix S-lemma requires $N$ to have at least one positive eigenvalue, an assumption known as the Slater condition. It turns out, however, that under additional assumptions on $M$ and $N$, we can state a theorem analogous to Theorem 4.7 in the case where $N \in \mathbb{S}_{q,r}$ has no positive eigenvalues, equivalently, $N \, | \, N_{22} = 0$. In this special case, $Z_r(N) = \mathbb{Z}_r^0(N)$ which leads to a matrix version of Finsler’s lemma.

Theorem 4.8 (Matrix Finsler’s lemma). Let $M, N \in \mathbb{S}_{q+r}$. If there exists $\alpha \in \mathbb{R}$ such that $M - \alpha N \geq 0$ then $\mathbb{Z}_r^0(N) \subseteq Z_r(M)$. Next, define $\Theta \in \mathbb{S}_{q}$ by

$$\Theta := \begin{bmatrix} I \\ -N_{22}N_{21} \end{bmatrix}^T M \begin{bmatrix} I \\ -N_{22}^T N_{21} \end{bmatrix}.$$

Assume that $M, N \in \mathbb{S}_{q,r}$, $N \, | \, N_{22} = 0$, and $\ker \Theta \subseteq \ker M \, | \, M_{22}$. Then $\mathbb{Z}_r^0(N) \subseteq Z_r(M)$ if and only if there exists $\alpha \geq 0$ such that $M - \alpha N \geq 0$.

Proof. The ‘if’ statements are obvious. Now, assume that $\mathbb{Z}_r^0(N) \subseteq Z_r(M)$. Let $Z \in \mathbb{Z}_r^0(N), \xi \in \ker N_{22}$, and $\eta \in \mathbb{R}^q$ be a nonzero vector. By hypothesis, we have

$$Z + \gamma \xi \eta^T \in Z_r(M)$$

for all $\gamma \in \mathbb{R}$. Recall that $M \in \mathbb{S}_{q,r}$ and therefore $M_{22} \leq 0$. This implies that $M_{22} \xi = 0$, for otherwise there exists a sufficiently large $\gamma \in \mathbb{R}$ that violates (4.5). We have thus proven that $\ker N_{22} \subseteq \ker M_{22}$. Let $T := \begin{bmatrix} I \\ -N_{22}N_{21} \end{bmatrix}$, then

$$T^T N T = \begin{bmatrix} 0 & 0 \\ 0 & N_{22} \end{bmatrix}$$

and $T^T M T = \begin{bmatrix} \Theta^T & M_{12} - N_{12}N_{22}^T M_{22} \\ M_{21} - M_{22}N_{22}^T N_{21} & M_{22} \end{bmatrix}$.

This yields

$$T^T(M - \alpha N) T = \begin{bmatrix} \Theta & M_{12} - N_{12}N_{22}^T M_{22} \\ M_{21} - M_{22}N_{22}^T N_{21} & M_{22} - \alpha N_{22} \end{bmatrix}.$$

Next, note that $\ker M_{22} \subseteq \ker M_{12}$ implies the following two statements:

$$\Theta = M \, | \, M_{22} + (M_{22}^T M_{21} - N_{22}^T N_{21}) M_{22}^T (M_{22}^T M_{21} - N_{22}^T N_{21})$$

and

$$M_{22}^T (M_{22} - N_{22}^T N_{21}) = M_{21} - M_{22}N_{22}^T N_{21}.$$
Since $N_{22} \leq 0$ and $\ker N_{22} \subseteq \ker M_{22} \subseteq \ker M_{12}$, we conclude that there exists a sufficiently large $\alpha \geq 0$ such that (4.8) holds. This implies that there exists an $\alpha \geq 0$ such that $M - \alpha N \geq 0$. This proves the theorem. □

The assumption on the matrix $\Theta$ is required in the sense that Theorem 4.8 is, in general, not valid without it. We illustrate this as follows.

**Example 4.9.** Suppose that

$$N = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. $$

Note that $M, N \in \Pi_{1,1}$ and $N | N_{22} = 0$. In this case, $\Theta = 0$ and $M | M_{22} = 1$ so the assumption on $\Theta$ of Theorem 4.8 does not hold. In addition, we see that $Z^+_1(N) = \{1\} \subseteq Z_1(M)$. Nonetheless, there does not exist an $\alpha \geq 0$ such that $M - \alpha N \geq 0$.

### 4.4. Matrix versions of the S-lemma and Finsler’s lemma for a strict inequality.

Subsequently, we consider strict versions of the above theorems. We focus on the case where the inequality involving $M$ is strict, while the inequality on $N$ is nonstrict. We note that a matrix S-lemma with two strict inequalities was considered in [49]. The following theorem provides a strict matrix S-lemma in case $N_{22}$ is negative definite.

**Theorem 4.10** (Strict matrix S-lemma). Let $M, N \in S^{q+r}$. If there exists a real $\alpha \geq 0$ such that $M - \alpha N > 0$, then $Z_r(N) \subseteq Z^+_r(M)$. Next, assume that $N \in \Pi_{q,r}$ and $N_{22} < 0$. Then $Z_r(N) \subseteq Z^+_r(M)$ if and only if there exists a real $\alpha \geq 0$ such that $M - \alpha N > 0$.

**Proof.** The ‘if’ parts are clear. Therefore, we focus on proving the ‘only if’ part of the second statement. Suppose that $Z_r(N) \subseteq Z^+_r(M)$. By Theorem 4.6, we have that $x^\top M x > 0$ for all nonzero $x \in \mathbb{R}^{q+r}$ satisfying $x^\top N x \geq 0$. We now distinguish two cases. First suppose that $N$ has at least one positive eigenvalue. Then, by Lemma 4.2, there exists a real $\alpha \geq 0$ such that $M - \alpha N > 0$. Next, suppose that $N$ does not have any positive eigenvalues, i.e., $N \leq 0$. We clearly have that $x^\top M x > 0$ for all nonzero $x \in \mathbb{R}^{q+r}$ satisfying $x^\top N x = 0$. Then, by Lemma 4.3, there exists a real $\bar{\alpha} \in \mathbb{R}$ such that $M - \bar{\alpha} N > 0$. If $\bar{\alpha} > 0$ then we have $M - \alpha N > 0$ for $\alpha = \bar{\alpha}$. On the other hand, if $\bar{\alpha} < 0$ then $M > \bar{\alpha} N \geq 0$, so $M - \alpha N > 0$ for $\alpha = 0$. This proves the theorem. □

One can even prove a strict matrix S-lemma in the case that $N_{22}$ is not necessarily negative definite, but under the extra assumptions that $M_{22} \leq 0$ and the Slater condition holds on $N$. It turns out, however, that in that case we need two real numbers $\alpha \geq 0$ and $\beta > 0$ to state a necessary and sufficient condition.

**Theorem 4.11** (Strict matrix S-lemma with $\alpha$ and $\beta$). Let $M, N \in S^{q+r}$. Then we have that $Z_r(N) \subseteq Z^+_r(M)$ if there exist scalars $\alpha \geq 0$ and $\beta > 0$ such that

$$(4.9) \quad M - \alpha N \geq \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix}.$$ 

Assume, in addition, that $N \in \Pi_{q,r}$, $M_{22} \leq 0$ and $N$ has at least one positive eigenvalue. Then $Z_r(N) \subseteq Z^+_r(M)$ if and only if there exist $\alpha \geq 0$ and $\beta > 0$ such that (4.9) holds.

**Proof.** Both ‘if’ statements are clear, so we focus on the ‘only if’ part. Assume that $Z_r(N) \subseteq Z^+_r(M)$. We will first prove that $\ker N_{22} \subseteq \ker M_{22}$ and $\ker N_{22} \subseteq \ker M_{12}$. Let $Z \in Z_r(N)$ and $v \in \ker N_{22}$. In addition, select any nonzero vector $w \in \mathbb{R}^q$ and define $\hat{Z} := vw^\top$. Since $N \in \Pi_{q,r}$, we have that $Z + \gamma \hat{Z} \in Z_r(N)$ for all $\gamma \in \mathbb{R}$. Therefore $Z + \gamma \hat{Z} \in Z^+_r(M)$. We write

$$(4.10) \quad 0 < \begin{bmatrix} I \\ Z + \gamma \hat{Z} \end{bmatrix}^\top M \begin{bmatrix} I \\ Z + \gamma \hat{Z} \end{bmatrix} = \mathcal{L}(\gamma) + \gamma^2 (v^\top M_{22} v) w w^\top,$$

where $\mathcal{L}(\gamma)$ is a matrix that depends affinely on $\gamma$. This implies that $M_{22} v = 0$. Indeed, if $M_{22} v \neq 0$ then $v^\top M_{22} v < 0$ and we can find a sufficiently large $\gamma \in \mathbb{R}$ that
violates (4.10). We conclude that \( \ker N_{22} \subseteq \ker M_{22} \). Next, let \( v \in \ker N_{22} \) and define \( \hat{Z} := -vv^\top M_{21} \). Since \( v \in \ker M_{22} \), we can write

\[
0 < \begin{bmatrix} I \\ Z + \gamma \hat{Z} \end{bmatrix}^\top M \begin{bmatrix} I \\ Z + \gamma \hat{Z} \end{bmatrix} = \begin{bmatrix} I \\ Z \end{bmatrix}^\top M \begin{bmatrix} I \\ Z \end{bmatrix} - 2\gamma M_{12}vv^\top M_{21}.
\]

This implies that \( M_{12}v = 0 \), or otherwise we can select a sufficiently large \( \gamma \in \mathbb{R} \) violating (4.11). Therefore, we conclude that \( \ker N_{22} \subseteq \ker M_{12} \). Subsequently, we claim that there exists a scalar \( \beta > 0 \) such that

\[
\beta > 0.
\]

Suppose on the contrary that this claim is false. Then there exists a sequence \( \{\beta_i\} \) such that \( \beta_i \to 0 \) (\( i \to \infty \)) and for all \( i \) there exists \( Z_i \in Z_r(N) \) such that

\[
Z_i \notin Z_r^+(M - \begin{bmatrix} \beta_i & 0 \\ 0 & 0 \end{bmatrix}).
\]

Define \( V := \{Z \in \mathbb{R}^{r \times q} | N_{22}Z = 0\} \). Write \( Z_i \) as \( Z_i = Z_i^1 + Z_i^2 \) where \( Z_i^1 \in V^\perp \) and \( Z_i^2 \in V \). Since \( \ker N_{22} \subseteq \ker N_{12} \) we see that \( Z_i^1 \in Z_r(N) \) for all \( i \). Next, we claim that \( \{Z_i^1\} \) is bounded. We will prove this by contradiction. Thus, assume that \( \{Z_i^1\} \) is unbounded. Clearly, the sequence \( \{\frac{Z_i^1}{\|Z_i^1\|}\} \) is bounded. By Bolzano-Weierstrass, it thus has a convergent subsequence with limit, say \( Z_* \). Note that

\[
\frac{1}{\|Z_i^1\|^2} (N_{11} + N_{12}Z_i^1 + (N_{12}Z_i^1)^\top + (Z_i^1)^\top N_{22}Z_i^1) \geq 0.
\]

By taking the limit along the subsequence as \( i \to \infty \), we obtain \( Z_*^\top N_{22}Z_* \geq 0 \). Using the fact that \( N_{22} \leq 0 \), we conclude that \( Z_* \in V \). Since \( Z_i^1 \in V^\perp \) for all \( i \), also \( \frac{Z_i^1}{\|Z_i^1\|} \in V^\perp \) and thus \( Z_* \in V^\perp \). Therefore, we conclude that both \( Z_* \in V \) and \( Z_* \in V^\perp \). That is, \( Z_* = 0 \). This is a contradiction as \( \frac{Z_i^1}{\|Z_i^1\|} \) has norm 1 for all \( i \). We conclude that the sequence \( \{Z_i^1\} \) is bounded. It thus contains a convergent subsequence with limit, say \( Z_* \). Note that \( Z_r(N) \) is closed and thus \( Z_* \in Z_r(\hat{N}) \). Since \( \ker N_{22} \subseteq \ker M_{22} \) and \( \ker N_{22} \subseteq \ker M_{12} \), (4.13) implies that

\[
\begin{bmatrix} I \\ Z_i^1 \end{bmatrix}^\top \left(M - \begin{bmatrix} \beta_i & 0 \\ 0 & 0 \end{bmatrix}\right) \begin{bmatrix} I \\ Z_i^1 \end{bmatrix} \neq 0
\]

for all \( i \). We take the limit as \( i \to \infty \) along a subsequence with limit \( Z_* \), which yields \( Z_* \notin Z_r^+(M) \). However, since \( Z_* \in Z_r(\hat{N}) \), this contradicts our hypothesis that \( Z_r(N) \subseteq Z_r^+(M) \). Therefore, we conclude that there exists a \( \beta > 0 \) such that (4.12) holds. In particular, this implies that there exists \( \beta > 0 \) such that

\[
Z_r(N) \subseteq Z_r^+ \left(M - \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix}\right).
\]

Finally, by Theorem 4.7, there exists a scalar \( \alpha \geq 0 \) such that (4.9) holds. \( \square \)

Next, we state a matrix Finsler’s lemma in the case of a strict inequality.

**Theorem 4.12 (Strict matrix Finsler’s lemma).** Let \( M, N \in S^{r \times r} \). Then \( Z_r^0(N) \subseteq Z_r^+(M) \) if there exist scalars \( \alpha \in \mathbb{R} \) and \( \beta > 0 \) such that (4.9) holds. Next, assume that \( N \in \Pi_{q,r} \) \( N_{22} = 0 \) and \( M_{22} \leq 0 \). Then \( Z_r^0(N) \subseteq Z_r^+(M) \) if and only if there exist \( \alpha > 0 \) and \( \beta > 0 \) such that (4.9) holds.

**Proof.** The ‘if’ statements are obvious. To prove the ‘only if’ statement, assume that \( Z_r^0(N) \subseteq Z_r^+(M) \). Since \( N_{22} = 0 \) and \( N_{22} \in \Pi_{q,r} \), we have that \( N \leq 0 \).
This implies that $Z_r^0(N) = Z_r(N)$. Therefore, we also have that $Z_r(N) \subseteq Z_r^+(M)$. We can thus use the same argument as in the proof of Theorem 4.7 to show that $\ker N_{22} \subseteq \ker M_{22}$ and $\ker N_{22} \subseteq \ker M_{12}$. Let $T := \begin{bmatrix} I & 0 \\ -N_{22}^T N_{21} & I \end{bmatrix}$ and $\Theta$ as (4.4), then

$$T^T NT = \begin{bmatrix} 0 & 0 \\ 0 & N_{22} \end{bmatrix} \quad \text{and} \quad T^T MT = \begin{bmatrix} \Theta & M_{12} - N_{12} N_{22}^T M_{22} \\ M_{21} - M_{22} N_{22}^T N_{21} & M_{22} \end{bmatrix}.$$ 

Since $-N_{22}^T N_{21} \in Z_r^0(N) \subseteq Z_r^+(M)$ we have $\Theta > 0$. Then obviously, there exists a real $\beta > 0$ so that $\Theta - \beta I > 0$. We have that

$$T^T \left( M - \alpha N - \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix} \right) T = \begin{bmatrix} \Theta - \beta I & M_{12} - N_{12} N_{22}^T M_{22} \\ M_{21} - M_{22} N_{22}^T N_{21} & M_{22} - \alpha N_{22} \end{bmatrix}.$$ 

Therefore it holds that $T^T \left( M - \alpha N - \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix} \right) T \geq 0$ if and only if

$$M_{22} - \alpha N_{22} - (M_{21} - M_{22} N_{22}^T N_{21}) (\Theta - \beta I)^{-1} (M_{12} - N_{12} N_{22}^T M_{22}) \geq 0. \tag{4.14}$$

Because $N_{22} \leq 0$, $\ker N_{22} \subseteq \ker M_{22}$ and $\ker N_{22} \subseteq \ker M_{12}$, there exists a sufficiently large $\alpha \geq 0$ such that (4.14) holds. This proves the statement.

Finally, we note that it is possible to combine the strict versions of the matrix S-lemma and Finsler’s lemma, Theorems 4.11 and 4.12, into one result. This results in the following corollary. Note the absence of the Slater condition on $N$.

**Corollary 4.13.** Let $M, N \in S^{n+r}$. Then $Z_r(N) \subseteq Z_r^+(M)$ if there exist scalars $\alpha \geq 0$ and $\beta > 0$ such that (4.9) holds. Next, assume that $N \in \Pi_{S_r}$ and $M_{22} \leq 0$. Then $Z_r(N) \subseteq Z_r^+(M)$ if and only if there exist $\alpha \geq 0$ and $\beta > 0$ such that (4.9) holds.

**Proof.** Once again, the ‘if’ parts are clear. To prove the ‘only if’ statement, we distinguish the cases that $N$ has at least one positive eigenvalue, and $N \leq 0$ (equivalently, $N \mid N_{22} = 0$). In the first case, Theorem 4.11 is directly applicable, resulting in the existence of $\alpha \geq 0$ and $\beta > 0$ such that (4.9) holds. In the second case, $Z_r(N) = Z_r^0(N)$ and Theorem 4.12 yields $\alpha \geq 0$ and $\beta > 0$ satisfying (4.9).

**4.5. Comparison with related work.**

**4.5.1. Full block S-procedure and LMI relaxations.** In this section, we point out a relation between Theorem 4.10 and the literature on LMI relaxations in robust control, see [40–42]. First of all, we acknowledge that these references are able to deal with more general uncertainty sets than those defined by quadratic matrix inequalities (i.e., the set $Z_r(N)$). Nevertheless, we can apply the general theory of [40] to obtain a matrix S-lemma with a strict inequality on $M$. To obtain the result, we substitute $A = 0$, $B = I$, $W(x) = -M$ and

$$\Delta = \left\{ \Delta \mid \begin{bmatrix} \Delta \\ I \end{bmatrix}^T N \begin{bmatrix} \Delta \\ I \end{bmatrix} \geq 0 \right\}$$

into Equation (1.2) of [40]. Then, we combine the full block S-procedure (c.f. [39], [40, p. 367]) with the fact that the LMI relaxation of [40] is exact for a single full block [40, Thm. 5.3]. This yields the following result.

**Proposition 4.14.** Let $M, N \in S^{n+r}$. Assume that $N$ is nonsingular, $N_{11} \geq 0$ and $N_{22} < 0$. Then we have that $Z_r(N) \subseteq Z_r^+(M)$ if and only if there exists a real $\alpha \geq 0$ such that $M - \alpha N > 0$.

This manuscript is for review purposes only.
We note that Proposition 4.14 also follows from Theorem 4.10 of this paper. In fact, \( N_{11} \geq 0 \) and \( N_{22} < 0 \) imply that \( N \in \Pi_{q,r} \), making Theorem 4.10 readily applicable. Theorem 4.10, however, is more general in the sense that nonsingularity of \( N \) and \( N_{11} \geq 0 \) are not required in general as long as \( N \mid N_{22} \geq 0 \).

4.5.2. Other matrix S-lemmas and Finsler’s lemmas. In the paper [54], various matrix S-lemmas were presented. The main contribution of Section 4 of this paper is to weaken some of the assumptions of these results, especially the so-called generalized Slater condition. We illustrate this by recalling [54, Thm. 13].

**Proposition 4.15.** Let \( M, N \in S^{q+r} \). Assume that \( \ker N_{22} \subseteq \ker N_{12}, N_{22} \leq 0 \) and \( M_{22} \leq 0 \). Suppose that \( Z_r^+(N) \) is nonempty. Then we have that \( Z_r(N) \subseteq Z_r^+(M) \) if and only if there exist \( \alpha \geq 0 \) and \( \beta > 0 \) such that (4.9) holds.

In [54], the nonemptiness of \( Z_r^+(N) \) was referred to as the generalized Slater condition. This condition, combined with the assumptions that \( N_{22} \leq 0 \) and \( \ker N_{22} \subseteq \ker N_{12}, \) implies that \( N \mid N_{22} > 0 \) (see (3.3)). As such, the conditions of Proposition 4.15 imply that \( N \in \Pi_{q,r} \). Proposition 4.15 thus follows immediately from the more general Corollary 4.13. The relevance of Corollary 4.13 is that the condition \( N \mid N_{22} > 0 \) has been replaced by the less restrictive assumption \( N \mid N_{22} \geq 0 \). In Section 5, we will illustrate the benefits of this weaker condition.

In addition to [54, Thm. 13], also a matrix S-lemma for non-strict inequalities was presented in [54, Thm. 9]. This matrix S-lemma is not directly comparable to Theorem 4.7 of this paper because it operates under different assumptions. In fact, [54, Thm. 9] does not assume that \( N \in \Pi_{q,r} \), but does instead assume the generalized Slater condition. For the application of data-driven control that we have in mind, however, the condition \( N \in \Pi_{q,r} \) is always satisfied (see Section 5). In this case, the relevance of Theorem 4.7 lies in the fact that it only assumes the “ordinary” Slater condition \( N \) has at least one positive eigenvalue), rather than the generalized Slater condition which requires \( N \) to have at least \( q \) positive eigenvalues.

Finally, in [53], a matrix version of Finsler’s lemma was provided. This results, however, operates under rather strong conditions such as \( M_{11} > 0 \) and \( M_{12} = 0 \). It can be shown that [53, Thm. 1], in fact, follows from Theorem 4.8 of this paper.

4.5.3. Petersen’s lemma. In the recent paper [5], the problem of data-driven stabilization is tackled using a result known as Petersen’s lemma, which originates from [35, 36]. In [5], a non-strict and a strict version of Petersen’s lemma are stated. These results are relevant for this paper, because they are intimately related to matrix S-lemmas. In fact, as we will see shortly, they can be regarded as matrix S-lemmas for special choices of the matrices \( M \) and \( N \), given in (4.17).

Petersen’s lemma, as well as the resulting conditions for data-driven stabilization [5], are interesting because they work under different assumptions than the matrix S-lemmas of [54]. Specifically, the results of [54] apply to unbounded sets \( Z_r(N) \) but assume the generalized Slater condition. On the other hand, when considering the results of [5] in the framework of this paper, the particular noise model [5, Eq. 8] and assumption on the data [5, As. 1] imply that the resulting set \( Z_r(N) \) is bounded. However, Petersen’s lemma does not require the generalized Slater condition (only the “ordinary” Slater condition is required in the non-strict Petersen’s lemma, while the strict version does not require a Slater condition).

The results in Section 4 of the current paper can be regarded as the best of both worlds. In fact, we will show in Proposition 4.16 that both the non-strict and the strict Petersen’s lemma can be obtained directly from Theorems 4.7 and 4.10, respectively. With respect to the literature on Petersen’s lemma, we conclude that the results of Section 4 are relevant because i) they are more general (i.e., they recover Petersen’s...
lemmas as special cases, but also apply to unbounded $Z_{r}(N)$, and ii) they appear naturally in data-driven control applications that can often be formulated as implications involving QMI’s. In comparison, the application of Petersen’s lemma requires a less straightforward reformulation of the set (2.5), cf. [5].

**Proposition 4.16** (Petersen’s lemma). Consider $C \in \mathbb{R}^{n \times n}$, $E \in \mathbb{R}^{n \times p}$, $F \in \mathbb{R}^{q \times q}$ and $G \in \mathbb{R}^{n \times n}$ with $C = C^\top$ and $F \succeq 0$, and define $\mathcal{F} := \{ F \in \mathbb{R}^{p \times q} \mid F^\top F \preceq \bar{F} \}$.

(a) The two following statements are equivalent

\begin{align}
(4.15a) & \quad C + EFG + G^\top F^\top E^\top < 0 \quad \text{for all } F \in \mathcal{F}, \\
(4.15b) & \quad \text{there exists } \lambda > 0 \text{ s.t. } C + \lambda EE^\top + \lambda^{-1}G^\top \bar{F}G < 0.
\end{align}

(b) Suppose that $E \neq 0$, $\bar{F} > 0$ and $G \neq 0$. Then the following are equivalent

\begin{align}
(4.16a) & \quad C + EFG + G^\top F^\top E^\top \preceq 0 \quad \text{for all } F \in \mathcal{F}, \\
(4.16b) & \quad \text{there exists } \lambda > 0 \text{ s.t. } C + \lambda EE^\top + \lambda^{-1}G^\top \bar{F}G \preceq 0.
\end{align}

**Proof.** First, note that by applying Lemma A.1 to $A = F$ and $B = \bar{F}^{\frac{1}{2}}$, we get

$$\{ FG \in \mathbb{R}^{p \times n} \mid F \in \mathbb{R}^{p \times q} \text{ and } F^\top F \leq \bar{F} \} = \{ SF^{\frac{1}{2}}G \in \mathbb{R}^{p \times n} \mid S \in \mathbb{R}^{p \times q} \text{ and } S^\top S \leq I \}.$$ 

Applying the same lemma with $A = H$ and $B = \bar{F}^{\frac{1}{2}}G$, this set is equal to $\{ H \in \mathbb{R}^{p \times n} \mid H^\top H \leq G^\top \bar{F}G \}$. Therefore, by defining the matrices in (4.1) as

\begin{equation}
(4.17) \quad M := \begin{bmatrix} -C & -E \\ -E^\top & 0 \end{bmatrix} \quad \text{and} \quad N := \begin{bmatrix} G^\top \bar{F}G & 0 \\ 0 & -I \end{bmatrix},
\end{equation}

we see that (4.15a) is equivalent to the inclusion $Z_{p}(N) \subseteq Z_{p}^+(M)$, while (4.16a) is equivalent to $Z_{p}(N) \subseteq Z_{p}(M)$. Note that $N \in \Pi_{n,p}$ and $N_{22} < 0$.

Now, we prove statement (a). By Theorem 4.10, $Z_{p}(N) \subseteq Z_{p}^+(M)$ is equivalent to the existence of $\alpha > 0$ such that $M - \alpha N \succeq 0$. In fact, $\alpha$ is necessarily positive because $M_{22} = 0$ and $N_{22} < 0$. By using a Schur complement argument, $M - \alpha N \succeq 0$ is equivalent to $-C - \alpha G^\top \bar{F}G - \frac{1}{\alpha} EE^\top \geq 0$, thus (4.15b) holds for $\lambda := \frac{1}{\alpha}$.

The proof of statement (b) is analogous, but uses Theorem 4.7 instead of Theorem 4.10. Note that the conditions of that theorem are satisfied since $\bar{F} > 0$ and $G \neq 0$ imply that $N$ has at least one positive eigenvalue. Moreover, $M - \alpha N \succeq 0$, $E \neq 0$, and $\alpha \geq 0$ imply that $\alpha > 0$. This proves the proposition.

5. **Applications to data-driven control.** In this section we apply the results to data-driven control problems. For brevity, we restrict our attention to data-based stabilization. However, we note that the results of this paper are also relevant for other analysis and design problems such as data-based dissipativity analysis [28,55], $H_\infty$- and $H_\infty$-control [54], and even data-driven reduced order modeling [8]. In Section 5.1 we provide a general necessary and sufficient condition for informativity for quadratic stabilization. Subsequently, in Section 5.2 we provide alternative conditions for quadratic stabilization that are more appealing from a computational point of view. We also provide an explicit formula for the controller.

5.1. **Data-driven quadratic stabilization.** Using the theory of Section 4 we want to understand i) under which conditions $(U_-, X)$ are informative for quadratic stabilization, and ii) how to construct a controller $K$ given informative data. Throughout this section, we will assume that $\Phi \in \Pi_{n,T}$. Define the matrices

---

This manuscript is for review purposes only.
From a design point of view, the matrices $P$ then, it can be chosen to be equal to 1. We summarize this in the following theorem.

Recall from Section 2 and (2.5) that $\Sigma = Z_{n+m}(N)$. Next, suppose that we fix a Lyapunov matrix $P > 0$ and a feedback gain $K$. The inequality (2.4) is equivalent to $[A \ B] \in Z_{n+m}^+(M)$. Then we need to find conditions on the data such that there exist $P > 0$ and $K$ for which the inclusion

$$Z_{n+m}(N) \subseteq Z_{n+m}^+(M)$$

holds. In order to find such conditions, we will apply Corollary 4.13. To do so, we need to verify its assumptions. In particular, we need to verify that $N_{22} \leq 0$, $N \nmid N_{22} \geq 0$, $\ker N_{22} \subseteq \ker N_{12}$ and $M_{22} \leq 0$. Note that $N_{22} = \begin{bmatrix} X_-^\top & \Phi_{12} & X_-^\top \end{bmatrix} \leq 0$ because $\Phi_{22} \leq 0$. This implies that

$$\ker N_{22} = \ker \Phi_{22} \begin{bmatrix} X_-^\top \\ U_-^\top \end{bmatrix}^\top, \text{ and } \ker N_{12} = \ker \begin{bmatrix} (\Phi_{12} + X_+ \Phi_{22}) & X_-^\top \\ U_-^\top \end{bmatrix}^\top.$$

Since $\ker \Phi_{22} \subseteq \ker \Phi_{12}$, also $\ker N_{22} \subseteq \ker N_{12}$. In addition, (3.4) and $(A_s, B_s) \in \Sigma$ imply $N \nmid N_{22} \geq 0$. Since $P > 0$, $M_{22} = -\begin{bmatrix} I \\ K \end{bmatrix} \begin{bmatrix} P \\ K \end{bmatrix}^\top \leq 0$. Now, Corollary 4.13 asserts that (5.3) holds if and only if there exist scalars $\alpha \geq 0$ and $\beta > 0$ such that

$$M - \alpha N \geq \begin{bmatrix} \beta I_n \\ 0 \\ 0_{(n+m) \times (n+m)} \end{bmatrix}.$$

From a design point of view, the matrices $P$ and $K$ that appear in $M$ are not given. However, the idea is now to compute matrices $P$, $K$ and scalars $\alpha$ and $\beta$ such that (5.4) holds. In fact, by the above discussion, the data $(U_-, X)$ are informative for quadratic stabilization if and only if there exists an $n \times n$ matrix $P > 0$, a $K \in \mathbb{R}^{m \times n}$ and two scalars $\alpha \geq 0$ and $\beta > 0$ such that (5.4) holds. We note that (5.4) (in particular, $M$) is not linear in $P$ and $K$. Nonetheless, by a rather standard change of variables and a Schur complement argument, we can transform (5.4) into a linear matrix inequality. Moreover, it turns out that the scalar $\alpha$ is necessarily positive. By a scaling argument then, it can be chosen to be equal to 1. We summarize this in the following theorem.

**Theorem 5.1.** Let $\Phi \in \Pi_{n,T}$ and consider the data $(U_-, X)$, generated by (2.1) with noise model (2.3). The following statements hold.

(a) The data $(U_-, X)$ are informative for quadratic stabilization if and only if there exists an $n \times n$ matrix $P > 0$, an $L \in \mathbb{R}^{m \times n}$ and a scalar $\beta > 0$ satisfying

$$\begin{bmatrix} P - \beta I & 0 & 0 \\ 0 & -P & -L^\top \\ 0 & L & P \end{bmatrix} - \begin{bmatrix} I \\ -X_+^\top \\ \Phi_{21} \end{bmatrix} \begin{bmatrix} \Phi_{11} \\ \Phi_{12} \\ \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ -X_+^\top \\ \Phi_{21} \end{bmatrix}^\top \geq 0.$$

(b) Assume, in addition, that $\Phi_{22} < 0$ and $\text{rank} \begin{bmatrix} X_+^\top \\ U_-^\top \end{bmatrix} = n + m$. Then the data $(U_-, X)$ are informative for quadratic stabilization if and only if there exists an $n \times n$ matrix $P > 0$ and a matrix $L \in \mathbb{R}^{m \times n}$ satisfying
Moreover, if either (5.5) or (5.6) is feasible then \( K := LP^{-1} \) is a stabilizing feedback gain for all \((A, B) \in \Sigma\).

Proof. We first focus on statement (a). To prove the ‘if’ part, suppose that there exist \( P, L \) and \( \beta \) satisfying (5.5). Define \( K := LP^{-1} \). By computing the Schur complement of (5.5) with respect to its fourth diagonal block, we obtain (5.4) with \( \alpha = 1 \). As such, (5.3) holds. We conclude that the data \((U_-, X)\) are informative for quadratic stabilization and \( K = LP^{-1} \) is indeed a stabilizing controller for all \((A, B) \in \Sigma\).

Conversely, to prove the ‘only if’ part, suppose that the data \((U_-, X)\) are informative for quadratic stabilization. This means that there exist \( P > 0 \) and \( K \) such that (5.3) holds. By Corollary 4.13 there exist \( \alpha > 0 \) and \( \beta > 0 \) satisfying (5.4). Then, by defining \( L := KP \) and using a Schur complement argument, we conclude that

\[
\begin{bmatrix}
P - \beta I & 0 & 0 & 0 \\
0 & -P & -L^\top & 0 \\
0 & -L & 0 & L \\
0 & 0 & L^\top & P
\end{bmatrix}
- \alpha
\begin{bmatrix}
I & X_+ \\
0 & -X_-
\end{bmatrix}
\begin{bmatrix}
\Phi_{11} & \Phi_{12} \\
\Phi_{21} & \Phi_{22}
\end{bmatrix}
\begin{bmatrix}
I & X_+ \\
0 & -X_-
\end{bmatrix}^\top
> 0
\]

is feasible. Zooming in on the \((2, 2)\) block, this yields \(-P - \alpha X_\Phi_2 X_\Phi_2^\top > 0\). Since, by assumption, \( \Phi_{22} \leq 0 \), the inequality \( P > 0 \) implies \( \alpha > 0 \). As a consequence, by scaling \( P, L \) and \( \beta \) by \( \frac{\alpha}{\beta} \) we may assume that \( \alpha = 1 \), so the LMI (5.5) is feasible.

The proof of statement (b) follows the same steps as the proof of (a), with the difference that Theorem 4.10 is invoked instead of Corollary 4.13, since the assumptions of (b) imply that

\[
\begin{bmatrix}
X_- \\
U_-
\end{bmatrix}
\Phi_{22}
\begin{bmatrix}
X_- \\
U_-
\end{bmatrix}^\top < 0
\]

This proves the theorem. \( \Box \)

Theorem 5.1 provides linear matrix inequality conditions for informativity for quadratic stabilization. Such feasibility problems have been studied extensively (see e.g. [7]) given their importance in the optimization literature. To be precise, LMI’s arise as the constraints in second-order cone programming, or more generally, in semi-definite programming [7]. Because of this importance, powerful computational tools such as the solvers Sedumi [48] and Mosek have been developed in order to verify LMI feasibility. Theorem 5.1(a) provides a genuine necessary and sufficient condition for quadratic stabilization, in the sense that no additional rank conditions on the data or assumptions on the noise model are needed. From a numerical point of view, however, the strict inequality of Theorem 5.1(b) may be preferred, since LMI solvers are known to be unreliable for LMIs that define feasible sets without interior points. Theorem 5.1 is an significant improvement of [54, Thm. 14]. In fact, the condition (5.7) is exactly the same as in [54], but Theorem 5.1 works under weaker assumptions. Indeed, in [54], the generalized Slater condition was imposed, which is equivalent to \( N \) having \( n \) positive eigenvalues. Since \( N_{22} \leq 0 \), the generalized Slater condition implies that \( N \mid N_{22} > 0 \). In particular, this implies that the conditions of Corollary 4.13 are satisfied. It should be clear, however, that \( N \mid N_{22} > 0 \) is not required in Theorem 5.1.

The weaker assumptions of Theorem 5.1 are especially relevant in the case that \( w \) takes its values in an a priori given subspace of the state-space, which is the case for noise model (iv) in Section 2. Indeed, suppose that \( w(t) = E\hat{w}(t) \) for all \( t = 0, 1, 2, \ldots, T-1 \), where \( \hat{w}(t) \in \mathbb{R}^d \) and \( E \in \mathbb{R}^{n \times d} \) is a given matrix of full column rank. The matrix \( \hat{W}_- = \begin{bmatrix} \hat{w}(0) & \hat{w}(1) & \cdots & \hat{w}(T-1) \end{bmatrix} \) captures the noise. As before, \( \hat{W}_- \) is unknown but is assumed to satisfy \( \hat{W}_-^\top \in \mathbb{Z}_T(\hat{\Phi}) \), where \( \hat{\Phi} \in \Pi_{d,T} \) is such that
\( \Phi_{22} < 0 \). Now, by Theorem 3.4, \( W_+ = E \hat{W}_+ \) for some \( \hat{W}_+ \in \mathcal{Z}_T(\Phi) \) if and only if \( W_+^\top \in \mathcal{Z}_T(\Phi) \), where
\[
(5.8) \quad \Phi := \begin{bmatrix} E\hat{\Phi}_{11}E^\top & E\hat{\Phi}_{12} \\ \hat{\Phi}_{21}E^\top & \hat{\Phi}_{22} \end{bmatrix} \in \Pi_{n,T}.
\]
The conclusion is that Theorem 5.1 is applicable in cases where the noise is constrained to a known subspace, which is captured by the noise bound (2.3) with \( \Phi \) in (5.8). We note that [54, Thm. 14] is generally not applicable in this context. The reason is that the matrix \( \Phi \) in (5.2) can only be positive definite in the case that \( d = n \). By [14, Thm. 3.1], this means that \( N \) in (5.2) always has less than \( n \) positive eigenvalues whenever \( d < n \).

**Remark 5.2.** We note that Theorem 5.1 makes use of state measurements \( X \). It is possible to extend this result and to consider data-driven stabilization of input-output systems in autoregressive form \( P_s(\sigma)y(t) = Q_s(\sigma)u(t) + w(t) \), where \( u \in \mathbb{R}^m \) is the input, \( y \in \mathbb{R}^p \) is the output and \( w \in \mathbb{R}^p \) is the noise. Here \( P_s \) and \( Q_s \) are unknown polynomial matrices and \( \sigma \) denotes the shift operator, i.e., \( (\sigma f)(t) = f(t+1) \). In this case, the data consist of samples of the inputs and outputs, and the goal is to use these data to construct a controller of the form \( G(\sigma)u(t) = F(\sigma)y(t) \), with \( F \) and \( G \) polynomial matrices, that renders the interconnected system stable in the sense that all solutions \((u, y)\) to
\[
\begin{bmatrix}
G(\sigma) & -F(\sigma) \\
-Q(\sigma) & P_s(\sigma)
\end{bmatrix}
\begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix} w(t)
\]
converge to zero whenever \( w = 0 \). This can be done by relying on tools for quadratic matrix inequalities, as developed in this paper, and the notion of quadratic difference forms [60] from behavioral theory. We refer to [57] for a detailed analysis of input-output systems.

### 5.2. Reducing Computational Complexity

The computational complexity of determining feasibility of an LMI depends on the size of the LMI and the number of unknowns. The LMI (5.5), together with the constraints \( P > 0 \) and \( \beta > 0 \), has size \( 4n + m + 1 \) and contains \( \frac{n(n+1)}{2} + nm + 1 \) unknowns. Using Theorem 3.4, we can separate the computation of the Lyapunov matrix \( P \) and the controller \( K \). Below it will be shown that this leads to another LMI with size \( 4n + 1 \) and \( \frac{n(n+1)}{2} + 1 \) unknowns. This result thus has a significant computational advantage over Theorem 5.1.

**Theorem 5.3.** Let \( \Phi \in \Pi_{n,T} \) and consider the data \((U_-, X)\), generated by (2.1) with noise model (2.3). Define the matrix \( \Theta := \Phi_{12} + X_+ \Phi_{22} \). The following statements hold.

(a) The data \((U_-, X)\) are informative for quadratic stabilization if and only if there exists an \( n \times n \) matrix \( P > 0 \) and a scalar \( \beta > 0 \) satisfying
\[
(5.9a) \quad P - \beta I - [I \ X_+ \Phi \left( \begin{array}{c|c} I & X_+^\top \\ \hline & \end{array} \right)] + \Theta \left[ \begin{array}{c} X_- \\ U_- \end{array} \right] \left( \begin{array}{c} X_- \\ U_- \end{array} \right)^\top \left[ \begin{array}{c} X_- \\ U_- \end{array} \right] \Theta^\top \geq 0,
\]
\[
(5.9b) \quad \begin{bmatrix}
P - \beta I & 0 \\ 0 & -P
\end{bmatrix} - \Phi_{11} \Phi_{12} \left[ \begin{array}{c} I \\ X_+
\end{array} \right] \Phi_{21} \Phi_{22} \left[ \begin{array}{c} I \\ X_+
\end{array} \right]^\top \geq 0.
\]

Moreover, if \( P > 0 \) and \( \beta > 0 \) satisfy (5.9a) and (5.9b) then
\[
(5.10) \quad K = (U_-(\Phi_{22} + \Theta^\top \Gamma^\top \Theta)X_+^\top)(X_-(\Phi_{22} + \Theta^\top \Gamma^\top \Theta)X_+^\top)^\top
\]
is a stabilizing gain for all \((A, B) \in \Sigma\), where \( \Gamma = P - \beta I - [I \ X_+] \Phi \left( \begin{array}{c|c} I & X_+ \end{array} \right)^\top \).

(b) Assume, in addition, that \( \Phi_{22} < 0 \) and \( \text{rank} \left[ \begin{array}{c} X_-^\top \\ U_-^\top \end{array} \right] = n + m \). Then the data \((U_-, X)\) are informative for quadratic stabilization if and only if there exists an \( n \times n \) matrix \( P > 0 \) satisfying

\[ This manuscript is for review purposes only.
(5.11a) \[ P - [I \ X_+] \Phi \left[ \begin{array}{c} t \\ X_+^T \end{array} \right] + \Theta \left[ \begin{array}{c} X_- \\ U_- \end{array} \right]^T \left( \left[ \begin{array}{c} X_+ \\ U_- \end{array} \right] \Phi_{12} \left[ \begin{array}{c} X_- \\ U_- \end{array} \right] \right)^{-1} \left[ \begin{array}{c} X_- \\ U_- \end{array} \right] \Theta^T > 0. \]

(5.11b) \[ \begin{bmatrix} P & 0 & 0 \\ 0 & -P & 0 \\ 0 & 0 & -P \end{bmatrix} - \begin{bmatrix} I & X_+ \\ 0 & -X_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ -X_- \end{bmatrix}^T > 0. \]

Moreover, if \( P > 0 \) satisfies (5.11a) and (5.11b) then \( K \) in (5.10) is a stabilizing feedback gain for all systems \((A, B) \in \Sigma\), where \( \Gamma = P - [I \ X_+] \Phi \left[ \begin{array}{c} t \\ X_+^T \end{array} \right] \).

Proof. We first prove (a). Let \( P > 0 \) be an \( n \times n \) matrix and \( \beta > 0 \) be a real number. According to Theorem 5.1, it is enough to show that there exists \( L \in \mathbb{R}^{m \times n} \) satisfying (5.5) if and only if (5.9a) and (5.9b) are satisfied. By taking the Schur complement of the left hand side in (5.5) with respect to \( P \), one can see that there exists \( L \in \mathbb{R}^{m \times n} \) satisfying (5.5) if and only if there exists \( L \in \mathbb{R}^{m \times n} \) satisfying

\[
(5.12) \quad \begin{bmatrix} P - \beta I & 0 & 0 \\ 0 & -P & -L^T \\ 0 & 0 & -L \end{bmatrix} - \begin{bmatrix} I & X_+ \\ 0 & -X_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ -X_- \end{bmatrix}^T > 0.
\]

By direct inspection, one can verify that (5.12) is equivalent to the following QMI:

\[
(5.13) \quad \begin{bmatrix} I_n & 0 & 0 \\ 0 & I_n & 0 \\ 0 & 0 & I_{m,n} \\ 0 & 0 & P^{-1}L^T \end{bmatrix} \Psi_{11} \Psi_{12} \Psi_{21} \Psi_{22} \begin{bmatrix} I_n & 0 & 0 \\ 0 & I_n & 0 \\ 0 & 0 & I_{m,n} \\ 0 & 0 & P^{-1}L^T \end{bmatrix} > 0, \quad \text{where}
\]

\[
\Psi_{11} = \begin{bmatrix} P - \beta I & 0 & 0 \\ 0 & -P & 0 \\ 0 & 0 & -P \end{bmatrix} - \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ -X_- \\ 0 \end{bmatrix}^T, \quad \Psi_{12} = \begin{bmatrix} 0_{n,n} & P \\ 0 & 0 \end{bmatrix},
\]

and \( \Psi_{22} = -P \). Observe that (5.13) is equivalent to

\[
(5.14) \quad [0_{n,2n}, \ P^{-1}L^T] \in Z_n(\Psi).
\]

Now, by defining \( W = [I_{2n}, \ 0_{2n,m}]^T \) and \( Y = 0_{n,2n} \) it follows from Corollary 3.6 that there exists an \( L \) such that (5.14) holds if and only if

\[
(5.15) \quad \Psi \in \Pi_{2n+m,n} \quad \text{and} \quad 0_{n,2n} \in Z_n(\Psi_W)
\]

where \( \Psi_W \) is defined in (3.9). Next, we observe that \( \Psi \in \Pi_{2n+m,n} \) if and only if \( \Psi \mid \Psi_{22} \geq 0 \) since \( \Psi_{22} = -P < 0 \). Note that

\[
\Psi \mid \Psi_{22} = \begin{bmatrix} P - \beta I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ -X_- \\ 0 \end{bmatrix}^T.
\]

By a Schur complement argument, we see that \( \Psi \mid \Psi_{22} \geq 0 \) if and only if (5.9a) holds. Now, observe that \( 0_{n,2n} \in Z_n(\Psi_W) \) if and only if \( W^\top \Psi_{11}W \geq 0 \). Note that

\[
(5.16) \quad W^\top \Psi_{11}W = \begin{bmatrix} P - \beta I & 0 & 0 \\ 0 & -P & 0 \\ 0 & 0 & -P \end{bmatrix} - \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ -X_- \\ 0 \end{bmatrix}^T.
\]

Therefore, \( 0_{n,2n} \in Z_n(\Psi_W) \) if and only if (5.9b) holds. Consequently, the data are informative for quadratic stabilization if and only if there exists an \( n \times n \) matrix \( P > 0 \) and a scalar \( \beta > 0 \) satisfying (5.9a) and (5.9b).

For the construction of the controller, assume that \( P > 0 \) and \( \beta > 0 \) satisfy (5.9a) and (5.9b). By (5.16), this implies that \( W^\top \Psi_{11}W \geq 0 \). This is equivalent to

\[-W^\top \Psi_{12} P^{-1} \Psi_{21} W \leq W^\top (\Psi \mid \Psi_{22}) W.\]
Using Lemma A.1 with $A = -\Psi_{22}^{-1} \Psi_{21} W$ and $B = (\Psi | \Psi_{22}) \frac{1}{2} W$ there exists a matrix $S \in \mathbb{R}^{n \times (2n + m)}$ satisfying $S^\top S \leq I$ and
\begin{equation}
(5.17) \quad -(-\Psi_{22})^{-1} \Psi_{21} W = S(\Psi | \Psi_{22}) \frac{1}{2} W.
\end{equation}

In fact, the matrix $S := -(-\Psi_{22})^{-1} \Psi_{21} W \left( (\Psi | \Psi_{22}) \frac{1}{2} W \right)^\dagger$ works. Since $\Psi_{22} < 0$ and $S^\top S \leq I$, Theorem 3.3 yields $Z := -\Psi_{22}^{-1} \Psi_{21} + (-\Psi_{22})^{-1} S(\Psi | \Psi_{22}) \frac{1}{2} \in \mathbb{Z}_n(\Psi)$. It follows from (5.17) that $ZW = 0$. Now define $K := \begin{bmatrix} 0 & I_m \end{bmatrix} Z^\top$. Then, by (5.14) and Theorem 5.1, $K$ is a stabilizing feedback for all $(A, B) \in \Sigma$. It remains to be shown that $K$ is equal to (5.10).

First, observe that $Z \begin{bmatrix} 0 \\ I_m \end{bmatrix} = \Psi_{22}^{-1} \Psi_{21} W \left( (\Psi | \Psi_{22}) \frac{1}{2} W \right)^\dagger (\Psi | \Psi_{22}) \frac{1}{2} \begin{bmatrix} 0 \\ I_m \end{bmatrix}$ because $\Psi_{21} \begin{bmatrix} 0 \\ I_m \end{bmatrix} = 0$. It can be shown that $\left( (\Psi | \Psi_{22}) \frac{1}{2} W \right)^\dagger = (\Psi_W | \Psi_{22})^\dagger W^\top (\Psi | \Psi_{22})^\frac{1}{2}$.

Then, using the fact that $\Psi_{22}^{-1} \Psi_{21} W = \begin{bmatrix} 0 & I_n \end{bmatrix}$, this yields
\begin{equation}
(5.18) \quad K = \begin{bmatrix} 0 & I_m \end{bmatrix} (\Psi | \Psi_{22}) W (\Psi_W | \Psi_{22})^\dagger \begin{bmatrix} 0 \\ I_n \end{bmatrix},
\end{equation}

which implies
\begin{equation}
(5.19) \quad \begin{bmatrix} 0 & I_m \end{bmatrix} (\Psi | \Psi_{22}) W = U_- \begin{bmatrix} \Theta^\top & -\Phi_{22} X^- \end{bmatrix}.
\end{equation}

Moreover, note that $\Psi_W | \Psi_{22} = W^\top (\Psi | \Psi_{22}) W = \begin{bmatrix} \Gamma & \Theta X^- \\ X^- \Theta^\top & \Omega \end{bmatrix}$, where $\Omega = -X_- \Phi_{22} X^- \top$. It follows from [50, Thm. 2.10] that
\begin{equation}
(5.20) \quad (\Psi_W | \Psi_{22})^\dagger \begin{bmatrix} 0 \\ I_n \end{bmatrix} = \begin{bmatrix} \Gamma^\top \Theta X^- \\ -I_n \end{bmatrix} (X_- (\Phi_{22} + \Theta^\top \Gamma (\Theta) X^-))^\dagger.
\end{equation}

By substituting (5.19) and (5.20) into (5.18), we see that $K$ is equal to (5.10).

The proof of (b) can be established by following the same steps as the proof of (a), but replaces the term $P - \beta I$ by $P$ and non-strict inequalities by strict ones. Note that in this case, the proof of the if and only if statement relies on Corollary 3.7 rather than Corollary 3.6 and on Theorem 5.1(b) instead of Theorem 5.1(a). Also the construction of the controller builds on the results for strict inequalities in Lemma A.1(b) and Theorem 3.3(b), rather than their non-strict counterparts. This proves the theorem. []

53. Data-driven control of Lur’e systems. In this section, we will apply some of the results of this paper to the control of Lur’e systems. First, we will explain the classical problem of absolute stability for such systems. Consider the Lur’e system
\begin{equation}
(5.21) \quad x(t + 1) = Ax(t) + B u(t) + E \varphi(y(t)) \\
y(t) = C x(t) + D u(t)
\end{equation}
where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input, $y \in \mathbb{R}$ is the output and $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is a (nonlinear) function satisfying the so-called sector condition
\begin{equation}
(5.22) \quad \varphi(y)(\varphi(y) - 2y) \leq 0 \quad \forall y \in \mathbb{R}.
\end{equation}

We note that more general sector conditions can be transformed to (5.22) by means of loop transformations [7]. The real matrices $A, B, C, D$ and $E$ are of appropriate dimensions. Suppose that we apply a state feedback controller $u = K x$ resulting in
\begin{equation}
(5.23) \quad x(t + 1) = (A + BK) x(t) + E \varphi(y(t)) \\
y(t) = (C + DK) x(t).
\end{equation}
For systems of the form (5.23), a problem with a rich history is that of absolute stability, i.e., global asymptotic stability of the equilibrium point 0 for all sector-bounded nonlinearities, see [7] and the references therein. We focus on showing absolute stability of (5.23) by means of a quadratic Lyapunov function $V(x) := x^TPx$ where $P := P^T > 0$. We thus want that $V(x(t+1)) < V(x(t))$ for all sector-bounded nonlinearities $\varphi$ and all nonzero $x(t)$ and resulting $x(t+1)$ satisfying (5.23). We will mimic the continuous-time setting of [7, Ch. 5]. Let $A_K := A + BK$ and $C_K := C + DK$. Then it can be shown [53] that proving absolute stability of (5.23) by a quadratic Lyapunov function boils down to finding $P = P^T > 0$ such that

$$
(5.24) \quad \begin{bmatrix} P - A_K^T P A_K & -A_K^T P E - C_K^T \\ -E^T P A_K - C_K & 1 - E^T P E \end{bmatrix} > 0.
$$

Next, we focus on data-based stabilization of Lur'e systems. Consider the system

$$
(5.25) \quad \begin{align*}
 x(t+1) &= A_s x(t) + B_s u(t) + E \varphi(y(t)) + w(t) \\
 y(t) &= C_s x(t) + D_s u(t) + v(t),
\end{align*}
$$

where $A_s, B_s, C_s$ and $D_s$ are unknown real matrices and the matrix $E$ is known. The signals $w$ and $v$ are process and measurement noise terms that are unknown. We obtain state and input measurements from (5.25), collected in the matrices $X$ and $U$ as in Section 5.1, in addition to corresponding measurements of the form $Y_n = [y(0), y(1), \ldots, y(T-1)]$ and $F_n = [\varphi(y(0)), \varphi(y(1)), \ldots, \varphi(y(T-1))]$. During the experiment, the noise samples

$$
W_n := \begin{bmatrix} w(0) & w(1) & \cdots & w(T-1) \\
 v(0) & v(1) & \cdots & v(T-1) \end{bmatrix}
$$

are assumed to satisfy $W_n^T \in \mathcal{Z}_T(\Phi)$ for some known matrix $\Phi \in \Pi_{n+1,T}$. If we define $X_+$ and $X_-$ before then all systems $(A, B, C, D)$ explaining the data are given by the set $\Sigma$ defined by $\Sigma := \{ (A, B, C, D) : \begin{bmatrix} X_+ - EF_- \\ Y_- \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} \text{ for some } W_n^T \in \mathcal{Z}_T(\Phi) \}$. This leads to the following definition of informative data.

**Definition 5.4.** Let $\Phi \in \Pi_{n+1,T}$. Suppose that the data $(U_-, F_-, X, Y_-)$ have been generated by (5.25) for some noise sequence $W_n^T \in \mathcal{Z}_T(\Phi)$. Then $(U_-, F_-, X, Y_-)$ are called informative for absolute quadratic stabilization if there exists an $n \times n$ matrix $P > 0$ and a $K \in \mathbb{R}^{m \times n}$ such that (5.24) holds for all $(A, B, C, D) \in \Sigma$.

Next, we state the following theorem that gives a necessary and sufficient condition for informativity for absolute quadratic stabilization.

**Theorem 5.5.** Let $\Phi \in \Pi_{n+1,T}$ and consider the data $(U_-, F_-, X, Y_-)$, generated by (5.25) for some $W_n^T \in \mathcal{Z}_T(\Phi)$, and $\alpha \in \mathbb{R}$ and $\beta > 0$ such that

$$
\begin{bmatrix} Q - \beta I & -E & 0 & 0 & 0 \\
-E^T & 1 - \beta & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & Q \\
0 & 0 & 0 & L & 0 \\
0 & Q & L^T & Q \end{bmatrix} + \alpha \begin{bmatrix} I & 0 & X_+ - EF_- \\ 0 & 1 & Y_- \\ 0 & 0 & -X_- \\ 0 & 0 & -U_- \\ 0 & 0 & 0 \end{bmatrix} \Phi \begin{bmatrix} I & 0 & X_+ - EF_- \end{bmatrix}^T \geq 0.
$$

In this case, $K := LQ^{-1}$ is such that (5.23) is absolutely stable for all $(A, B, C, D) \in \Sigma$.

The proof follows similar lines as that of Theorem 5.1. It uses a dualization step [43, Lem. 4.9] on the inequality (5.24) and relies on Corollary 4.13 using the relation $Q = P^{-1}$ between $P$ and $Q$. 

---

This manuscript is for review purposes only.
5.4. The setting of stochastic noise. Consider an unknown system
\[ x(t + 1) = A_s x(t) + B_s u(t) + w(t), \]
where \( x(t), w(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^m \), and the matrices \( A_s, B_s \) have the appropriate dimensions. Assume that \( w \sim \mathcal{N}(0, \sigma^2 I) \), that is, \( w \) is normally distributed with mean 0 and covariance \( \sigma^2 I \). We suppose that \( \sigma \) is known and that we have access to a finite number of state and input measurements. Clearly, any linear system of the given dimensions \( n \) and \( m \) could have generated these measurements, yet not every system is a probable explanation. We are interested in characterizing sets of systems that contain the true system with a given probability. For this, we will follow the method of [51].

Let \( X, U_-, X_+ \), and \( X_+ \) be as before. Clearly, if no noise were present we would have \( X_+ = A_s X_- + B_s U_- \), and the true system \( (A_s, B_s) \) would be a solution to this inhomogeneous linear equation. However, due to the noise term, we consider its least squares solutions. Let \( \| \cdot \|_F \) denote the Frobenius norm of a given matrix and let \((\hat{A}, \hat{B})\) be an ordinary least squares estimate of the matrix pair \((A_s, B_s)\), that is,
\[ (\hat{A}, \hat{B}) \in \arg \min_{(A,B) \in \mathbb{R}^{n \times m}} \| X_+ - [A \ B] [X_- \ U_-] \|_F. \]

We can employ the Moore-Penrose inverse to explicitly write a solution of the previous as \( [\hat{A} \ \hat{B}] = X_+ \left[ \begin{array}{c} X_- \\ U_- \end{array} \right]^{-1} \). This solution is unique if and only if the matrix \( \left[ \begin{array}{c} X_- \\ U_- \end{array} \right] \) has full row rank \( n + m \). Assume this to be the case in the sequel.

Denote the \( \chi^2 \)-squared probability distribution with \( k \) degrees of freedom by \( \chi^2_k \). The paper [51] shows that the vectorization of \( [A \ B] \) is distributed according to \( \chi^2_{n(n+m)} \), with as mean the vectorization of the least square estimate \([\hat{A} \ \hat{B}]\). Using a generalization of the aforementioned confidence intervals for vector variables, we can then formulate confidence intervals for the matrix variable \((A, B)\). Given probability \( 0 < \delta < 1 \), denote the quantile function for probability \( \delta \) of \( \chi^2_{n(n+m)} \) by \( c_\delta \).

Now, define the set
\[ \Theta_\delta := \left\{ (A, B) \mid \left[ \begin{array}{c} \hat{A}^T - A^T \\ \hat{B}^T - B^T \end{array} \right]^T \left[ \begin{array}{c} X_- \\ U_- \end{array} \right]^T \left[ \begin{array}{c} \hat{A}^T - A^T \\ \hat{B}^T - B^T \end{array} \right] \leq \sigma^2 c_\delta I \right\}. \]

The following lemma was proven in [51, Lem. 3.1].

**Lemma 5.6.** Let \( 0 < \delta < 1 \). Then with probability \( 1 - \delta \) we have that \((A_s, B_s) \in \Theta_\delta\).

Recall that we are interested in finding a stabilizing controller for the system \((A_s, B_s)\). Suppose that there exists a gain \( K \) and a matrix \( P > 0 \) such that the Lyapunov inequality (2.4) holds for all \((A, B) \in \Theta_\delta\). Then we can conclude from Lemma 5.6 that, with probability \( 1 - \delta \), this gain also stabilizes the true system \((A_s, B_s)\). This leads to the following definition:

**Definition 5.7.** Given \( 0 < \delta < 1 \), the data \((U_-, X)\) are called informative for quadratic stabilization with probability \( 1 - \delta \) if there exists a feedback gain \( K \) and a matrix \( P > 0 \) such that the Lyapunov inequality (2.4) holds for all \((A, B) \in \Theta_\delta\).

It should be stressed that Definition 5.7 is dependent on the construction of \( \Theta_\delta \). However, the choice of confidence interval is not unique. In general, there are different sets \( \Theta \) that also contain \((A_s, B_s)\) with probability \( 1 - \delta \). Each of such sets could lead to a definition analogous to Definition 5.7. In order to apply the results of this paper, we choose to base our definition on the specific choice of \( \Theta_\delta \). In the following theorem we will apply the framework of this paper to provide necessary and sufficient conditions for this notion of informativity.

This manuscript is for review purposes only.
THEOREM 5.8. Given \(0 < \delta < 1\), and data \((U_-, X)\) such that \(\begin{bmatrix} X_- \\ U_- \end{bmatrix}\) has full row rank.

Then the data \((U_-, X)\) are informative for quadratic stabilization with probability \(1 - \delta\) if and only if there exists an \(n \times n\) matrix \(P > 0\), an \(L \in \mathbb{R}^{m \times n}\) and a scalar \(\beta > 0\) satisfying

\[
\begin{bmatrix}
P - \beta I & 0 & 0 & 0 \\
0 & -P & -L^T & 0 \\
0 & -L & 0 & L \\
0 & 0 & L^T & P
\end{bmatrix}
\begin{bmatrix}
I & X_+ \\
0 & -X_- \\
0 & L \\
0 & U_-
\end{bmatrix}
\begin{bmatrix}
\sigma^2 c_3 I \\
0 \\
- \begin{bmatrix} X_- & X_- \end{bmatrix}^T \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
I & X_+ \\
0 & -X_- \\
0 & L \\
0 & U_-
\end{bmatrix}^T \geq 0.
\] (5.26)

Moreover, if \(P \) and \(L\) satisfy (5.26) then \(K := LP^{-1}\) is a stabilizing feedback gain for \((A_s, B_s)\) with probability at least \(1 - \delta\).

Proof. Let \(M\) be defined as in (5.1). Clearly, the data is informative for quadratic stabilization with probability \(1 - \delta\) if and only if there exists \(P > 0\) and \(K\) such that \(\Theta_\delta \subseteq Z_n^+(M)\). The proof of this theorem will continue as follows. First, we will define a matrix \(\Phi \in \Pi_{n,T}\). After this, we show that by defining \(N\) as in (5.2), we obtain \(\Theta_\delta = Z_{n+m}(N)\). Using this, we will prove the theorem by invoking Theorem 5.1. Let

\[
\Phi := \begin{bmatrix}
\sigma^2 c_3 I \\
0 \\
- \begin{bmatrix} X_- & X_- \end{bmatrix}^T \\
0 & 0 & 0
\end{bmatrix}.
\]

Since \(\begin{bmatrix} X_- \\ U_- \end{bmatrix}\) has full row rank, we see that \(\Phi_{22} < 0\). Moreover, we can immediately see that \(\Phi \Phi_{22} = \sigma^2 c_3 I > 0\). We may conclude that \(\Phi \in \Pi_{n,T}\). Let \(N\) be defined as in (5.2). By performing some routine operations regarding the pseudo-inverse, and employing the definition of \((\hat{A}, \hat{B})\), we can conclude that

\[
N = \begin{bmatrix}
\sigma^2 c_3 I - \begin{bmatrix} A^T \\ B^T \end{bmatrix}^T & \begin{bmatrix} X_- \\ U_- \end{bmatrix}^T & \begin{bmatrix} A^T \\ B^T \end{bmatrix}^T & \begin{bmatrix} X_- \\ U_- \end{bmatrix}^T & \begin{bmatrix} X_- \\ U_- \end{bmatrix}^T \\
\begin{bmatrix} X_- \\ U_- \end{bmatrix}^T & \begin{bmatrix} A^T \\ B^T \end{bmatrix}^T & - \begin{bmatrix} X_- \\ U_- \end{bmatrix}^T & \begin{bmatrix} X_- \\ U_- \end{bmatrix}^T & \begin{bmatrix} X_- \\ U_- \end{bmatrix}^T
\end{bmatrix}.
\]

As such, it can be concluded that \(\Theta_\delta = Z_{n+m}(N)\). This shows that informativity for quadratic stabilization with probability \(1 - \delta\) is equivalent to informativity for quadratic stabilization corresponding to the specific noise model given by \(\Phi\). We can now invoke Theorem 5.1 to finalize the proof.

6. Conclusions. In this paper we have studied properties of quadratic matrix inequalities. We have established conditions under which the solution sets of such QMIs are nonempty, convex, bounded, or have nonempty interior (Theorem 3.2). In addition, we have given parameterizations of the solution set of a given QMI, for both strict and nonstrict inequalities (Theorem 3.3). We have also shown that under suitable conditions all matrices of the form \(ZW\), where \(Z\) is a solution to a QMI and \(W\) is a given matrix, are again solutions to a (different) QMI (Theorems 3.4 and 3.5). Finally, we have established matrix versions of the classical S-lemma and Finsler’s lemma, that provide LMI conditions under which all solutions to one QMI also satisfy another QMI (Theorems 4.7, 4.8, 4.10, 4.11 and 4.12). We have demonstrated that our results generalize previous work, and in particular, do not require the assumption of the generalized Slater condition [54]. In the case of a strict inequality, we have given a combined matrix S-lemma and Finsler’s lemma, that does not even require the “ordinary” Slater condition (Corollary 4.13). It was also shown that both the strict and the nonstrict Petersen’s lemma [35,36] can be obtained from the matrix S-lemmas of this paper (Proposition 4.16).

We have studied applications of the various QMI results in the context of data-driven control. In particular, we have studied the problem of finding a stabilizing
controller of an unknown LTI system influenced by noise from a finite set of input-state samples. To this end, we have assumed that the (unknown) matrix of noise samples satisfies a QMI. This led to a set of solutions to a QMI consisting of LTI systems that “explain” the data. We have applied the combined matrix S-lemma and Finsler’s lemma to establish a necessary and sufficient LMI condition under which all systems explaining the data can be quadratically stabilized by a single controller (Theorem 5.1). Subsequently, we have reduced the computational complexity of this scheme by separating the computation of the Lyapunov function and the controller, leading to lower-dimensional LMI’s (Theorem 5.3). This result also included an explicit formula for a stabilizing controller, given a (pre-computable) Lyapunov function. We then tackled data-based stabilization of a class of Lur’ë systems in Theorem 5.5. Lastly, we applied the methods of this paper in order to derive Theorem 5.8, which provides probabilistic stabilization guarantees in the presence of Gaussian noise.

So far, we have only applied the QMI results to data-driven stabilization. Nonetheless, we believe that our results are also relevant for other data-based analysis and control problems that involve QMI’s. Some examples of this are $H_2$ and $H_\infty$ control [2, 46], data-driven dissipativity analysis [28], as well as data-driven reduced order modeling [8]. In particular, the more general matrix S-lemmas of this paper will lead to weaker conditions on the data for these analysis and design problems. Another topic for future research is to develop dedicated algorithms for solving the various linear matrix inequalities in this paper.

**Appendix A. Facts from matrix theory.** We start with the following:

**Lemma A.1.** Let $A \in \mathbb{R}^{r \times q}$ and $B \in \mathbb{R}^{p \times q}$.

(a) $A^\top A \leq B^\top B$ if and only if there exists $S \in \mathbb{R}^{r \times p}$ such that

$$A = SB \text{ and } S^\top S \leq I.$$  

(b) Assume, in addition, that $B$ has full column rank. Then $A^\top A < B^\top B$ if and only if there exists $S \in \mathbb{R}^{r \times p}$ such that

$$A = SB \text{ and } S^\top S < I.$$  

Moreover, if $A^\top A - B^\top B \leq 0$ (respectively, $< 0$), then $S := AB^\top$ satisfies (A.1) (respectively, (A.2)).

The proof of Lemma A.1 for the case $r = q = p$ is given in [4, Fact 5.10.19] and for the case $r = p$ in [38, Lem. 3]. Here we provide a constructive proof for the case that $r$ and $p$ are not necessarily equal.

**Proof.** To prove the ‘if’ parts of statements (a) and (b), assume that $A = SB$ with $S^\top S \leq I$ (respectively, $< I$). Then $A^\top A = B^\top S^\top S B \leq B^\top B$ (respectively, $< B^\top B$), where we have made use of full column rank of $B$ to prove the strict inequality.

Next, we prove the ‘only if’ part of (a). We thus assume that $A^\top A \leq B^\top B$. Our goal is to show that $S := AB^\top$ satisfies (A.1). First, note that $A^\top A \leq B^\top B$ implies that $\ker B \subseteq \ker A$, equivalently, $\im A^\top \subseteq \im B^\top$. Thus, there exists a matrix $Z \in \mathbb{R}^{r \times q}$ such that $A^\top = B^\top Z^\top$, equivalently, $A = ZB$. Therefore, $SB = AB^\top B = ZBB^\top B = ZB = A$. Moreover, $S^\top S = (B^\top)^\top A^\top AB^\top \leq (B^\top)^\top B^\top BB^\top = BB^\top \leq I$, where the last inequality follows from the fact that $BB^\top$ is an orthogonal projection matrix. This shows that $S$ satisfies (A.1). To prove the ‘only if’ part of statement (b), assume that $B$ has full column rank and $A^\top A < B^\top B$. This implies that there exists an $\epsilon > 0$ such that $(1 + \epsilon)A^\top A < B^\top B$. As such, by statement (a), the matrix $\hat{S} := \sqrt{1 + \epsilon} A^\top AB^\top$ satisfies $\sqrt{1 + \epsilon} = SB$ and $\hat{S}^\top \hat{S} \leq I$. Define $S := \frac{1}{\sqrt{1 + \epsilon}} \hat{S} = AB^\top$. Then $A = SB$ and $S^\top S = \frac{1}{1 + \epsilon} \hat{S}^\top \hat{S} < I$. We conclude that $S$ satisfies (A.2).
The following is a direct consequence of [4, Prop. 6.1.7].

**Lemma A.2.** Let $A \in \mathbb{R}^{n \times p}$ and $B \in \mathbb{R}^{n \times p}$. Then $AM = B$ if and only if $\text{im} B \subseteq \text{im} A$ and there exists $T \in \mathbb{R}^{r \times n}$ such that $M = A^TB + (I_q - A^TA)T$.

**References**

[1] G. Baggio, V. Katewa, and F. Pasqualetti, *Data-driven minimum-energy controls for linear systems*, IEEE Control Systems Letters, 3 (2019), pp. 589–594.

[2] J. Berberich, A. Koch, C. W. Scherer, and F. Allgöwer, *Robust data-driven state-feedback design*, in Proceedings of the American Control Conference, 2020, pp. 1532–1538.

[3] J. Berberich, C. W. Scherer, and F. Allgöwer, *Combining prior knowledge and data for robust controller design*, arxiv.org/abs/2009.05253, (2020).

[4] D. S. Bernstein, *Matrix Mathematics*, Princeton University Press, 2nd ed., 2009.

[5] A. Bisoffi, C. De Persis, and P. Tesi, *Data-driven control via Petersen’s lemma*, https://arxiv.org/abs/2109.12175, (2021).

[6] V. Blondel, *Simultaneous stabilization of linear systems*, vol. 191, Springer, 1994.

[7] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM studies in applied mathematics, 1994.

[8] A. M. Burohman, B. Besselink, J. M. A. Scherpen, and M. K. Camlibel, *From data to reduced-order models via generalized balanced truncation*, https://arxiv.org/abs/2109.11685, (2021).

[9] L. Campestrini, D. Eckhard, A. S. Buzanella, and M. Gevers, *Data-driven model reference control design by prediction error identification*, Journal of the Franklin Institute, 354 (2017), pp. 2628–2647.

[10] M. C. Campi, A. Lecchini, and S. M. Savaresi, *Virtual reference feedback tuning: a direct method for the design of feedback controllers*, Automatica, 38 (2002), pp. 1337–1346.

[11] A. Cohen, T. Koren, and Y. Mansour, *Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret*, in Proceedings of the International Conference on Machine Learning, vol. 97, July 2019, pp. 1300–1309.

[12] J. Coulson, J. Lygeros, and F. Dörfler, *Data-enabled predictive control: In the shallows of the DeePC*, in Proceedings of the European Control Conference, June 2019, pp. 307–312.

[13] T. Dai and M. Sznaier, *A moments based approach to designing MIMO data driven controllers for switched systems*, in Proceedings of the IEEE Conference on Decision and Control, 2018, pp. 5652–5657.

[14] J. Dancis, *A quantitative formulation of Sylvester’s law of inertia III*, Linear Algebra and its Applications, 80 (1986), pp. 141–158.

[15] C. De Persis and P. Tesi, *Formulas for data-driven control: Stabilization, optimality, and robustness*, IEEE Transactions on Automatic Control, 65 (2020), pp. 909–924.

[16] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, *On the sample complexity of the linear quadratic regulator*, Foundations of Computational Mathematics, (2019).

[17] J. Eising and M. K. Camlibel, *Data informativity for analysis of linear systems with convex conic constraints*, IEEE Control Systems Letters, 6 (2022), pp. 391–396.

[18] J. Eising and H. L. Trentelman, *Informativity of noisy data for structural properties of linear systems*, Systems & Control Letters, 158 (2021), p. 105058.

[19] M. Ferizbegovic, J. Umenberger, H. Hjalmarsson, and T. Schönb, *Learning Robust LQ-Controllers Using Application Oriented Exploration*, IEEE Control Systems Letters, 4 (2020), pp. 19–24.

[20] P. Finsler, *Über das vorkommen definiter und semidefiniter formen in scharen quadratischer formen*, Commentarii Mathematici Helvetici, 9 (1936), pp. 188–192.

[21] M. Green and D. J. Limebeer, *Linear robust control*, Courier Corporation, 2012.

[22] A. Helmersson, *Iac synthesis based on inertia constraints*, IFAC Proceedings Volumes, 32 (1999), pp. 3361–3366.

[23] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, *Learning-based model predictive control: Toward safe learning in control*, Annual Review of Control, Robotics, and Autonomous Systems, 3 (2020), pp. 269–296.

[24] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin, *Iterative feedback tuning: theory and applications*, IEEE Control Systems Magazine, 18 (1998), pp. 26–41.

[25] H. Hjalmarsson, S. Gunnarsson, and M. Gevers, *A convergent iterative restricted complexity control design scheme*, in Proceedings of the IEEE Conference on Decision and Control, vol. 2, Dec 1994, pp. 1735–1740.

[26] A. Iannelli and R. S. Smith, *A multiobjective LQR synthesis approach to dual control for uncertain plants*, IEEE Control Systems Letters, 4 (2020), pp. 952–957.

[27] T. Iwasaki, G. Meinsma, and M. Fu, *Generalized S-procedure and finite frequency KYP lemma*, Mathematical Problems in Engineering, 6 (2000), pp. 305–320.
H. J. Van Waarde, M. K. Camlibel, J. Eising, and H. L. Trentelman

[28] A. Koch, J. Berberich, and F. Allgöwer, Provably robust verification of dissipativity properties from data, To appear in IEEE Transactions on Automatic Control, (2021).

[29] H. Lin and P. J. Antsaklis, Switching stabilizability for continuous-time uncertain switched linear systems, IEEE Transactions on Automatic Control, 52 (2007), pp. 633–646.

[30] I. Markovsky and F. Dörfler, Behavioral systems theory in data-driven analysis, signal processing, and control, Annual Reviews in Control, 52 (2021), pp. 42–64.

[31] I. Markovsky and P. Rapisarda, Data-driven simulation and control, International Journal of Control, 81 (2008), pp. 1946–1959.

[32] T. M. Maupong and P. Rapisarda, Data-driven control: A behavioral approach, Systems & Control Letters, 101 (2017), pp. 37–43.

[33] J. Mohammadpour and C. W. Scherer, Control of linear parameter varying systems with applications, Springer Science & Business Media, 2012.

[34] I. R. Petersen, A procedure for simultaneously stabilizing a collection of single input linear systems using non-linear state feedback control, Automatica, 23 (1987), pp. 33–40.

[35] I. R. Petersen, A stabilization algorithm for a class of uncertain linear systems, Systems & Control Letters, 8 (1987), pp. 351–357.

[36] I. R. Petersen and C. V. Hollot, A Riccati equation approach to the stabilization of uncertain linear systems, Automatica, 22 (1986), pp. 397–411.

[37] I. Pólik and T. Terlaky, A survey of the S-lemma, SIAM Review, 49 (2007), pp. 371–418.

[38] A. Rantzer, On the Kalman-Yakubovich-Popov lemma, Systems & Control Letters, 28 (1996), pp. 7–10.

[39] C. W. Scherer, LPV control and full block multipliers, Automatica, 37 (2001), pp. 361–375.

[40] C. W. Scherer, Relaxations for robust linear matrix inequality problems with verifications for exactness, SIAM Journal on Matrix Analysis and Applications, 27 (2005), pp. 365–395.

[41] C. W. Scherer, BMI relaxations in robust control, European Journal of Control, 12 (2006), pp. 3–29.

[42] C. W. Scherer and C. W. J. Hol, Matrix sum-of-squares relaxations for robust semi-definite programs, Mathematical Programming, 107 (2006), pp. 189–211.

[43] C. W. Scherer and S. Weiland, Lecture Notes DISC Course on Linear Matrix Inequalities in Control, 1999.

[44] R. Skelton, T. Iwasaki, and K. Grigoriadis, A Unified Algebraic Approach to Linear Control Design, Taylor and Francis, 1998.

[45] T. Steentjes, M. Lazar, and P. Van den Hof, Controller identification for data-driven model-reference distributed control, in Proc. of the European Control Conference, 2021, pp. 2358–2363.

[46] T. Steentjes, M. Lazar, and P. Van den Hof, H∞ performance analysis and distributed controller synthesis for interconnected linear systems from noisy input-state data, in Proceedings of the IEEE Conference on Decision and Control, 2021, pp. 3723–3728.

[47] T. Steentjes, M. Lazar, and P. Van den Hof, On data-driven control: Informativity of noisy input-output data with cross-covariance bounds, IEEE Control Systems Letters, 6 (2022), pp. 2192–2197.

[48] J. F. Sturm, Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones, Optimization Methods and Software, 11 (1999), pp. 625–653.

[49] Z. Szabó, Z. Bíró, and J. Bokor, Multivariate S-procedure, in Proceedings of the European Control Conference, 2013, pp. 3907–3912.

[50] Y. Tian and Y. Takane, Schur complements and Banachiewicz-Schur forms, Electronic Journal of Linear Algebra, 13 (2005), pp. 405–418.

[51] J. Umenberger, M. Felizzi, T. B. Schön, and H. Hjalmarsson, Robust exploration in linear quadratic reinforcement learning, in Advances in Neural Information Processing Systems, vol. 32, 2019.

[52] H. J. Van Waarde, Beyond persistent excitation: Online experiment design for data-driven modeling and control, IEEE Control Systems Letters, 6 (2022), pp. 319–324.

[53] H. J. Van Waarde and M. K. Camlibel, A matrix Finsler’s lemma with applications to data-driven control, in Proceedings of the IEEE Conference on Decision and Control, 2021, pp. 5777–5782.

[54] H. J. Van Waarde and M. K. Camlibel, A note on persistency of excitation, Systems & Control Letters, 54 (2005), pp. 325–329.
[60] J. C. WILLEMS AND H. L. TRENTELMAN, On quadratic differential forms, SIAM Journal on Control and Optimization, 36 (1998), pp. 1703–1749.
[61] V. A. YAKUBOVICH, S-procedure in nonlinear control theory, Vestnik Leningrad University Mathematics, 4 (1977), pp. 73–93. (English translation).
[62] Y. Zi-ZONG AND G. JIN-HAI, Some equivalent results with Yakubovich’s s-lemma, SIAM Journal on Control and Optimization, 48 (2010), pp. 4474–4480.