

Data-Driven Stabilization Using Prior Knowledge on Stabilizability and Controllability

Amir Shakouri, Henk J. van Waarde, Tren M.J.T. Baltussen, W.P.M.H. (Maurice) Heemels

Abstract—In this work, we study data-driven stabilization of linear time-invariant systems using prior knowledge of system-theoretic properties, specifically stabilizability and controllability. To formalize this, we extend the concept of data informativity by requiring the existence of a controller that stabilizes all systems consistent with the data *and* the prior knowledge. We show that if the system is controllable, then incorporating this as prior knowledge does not relax the conditions required for data-driven stabilization. Remarkably, however, we show that if the system is stabilizable, then using this as prior knowledge leads to necessary and sufficient conditions that are *weaker* than those for data-driven stabilization without prior knowledge. In other words, data-driven stabilization is easier if one knows that the underlying system is stabilizable. We also provide new data-driven control design methods in terms of linear matrix inequalities that complement the conditions for informativity.

Index Terms—Data-driven control, stabilization, prior knowledge, controllability, stabilizability

I. INTRODUCTION

In recent years, it has been shown that stabilizing feedback laws can be directly obtained from measured data, as opposed to the classical approach of using a model of the system (see [1, Ch. 1.2] for a historical account). This idea can be motivated by the argument that bypassing the modeling procedure may reduce the total amount of computations since it does not require system identification as an intermediate step, cf. [2]. In addition, finding a stabilizing feedback directly from the data might be feasible even when the data do not contain sufficient information for accurate modeling of the system, see [3, Ex. 19].

For linear time-invariant (LTI) systems, direct data-driven stabilization has been extensively studied in the literature. It was shown in [4] that a stabilizing state-feedback can be directly obtained from input-state data by solving a linear matrix inequality (LMI). The method provided in [4] requires the input data to be persistently exciting of a certain order, see [5, p. 327]. This condition implies that the system can be uniquely identified. Soon after, it was shown in [3] that persistency of excitation is not necessary. In fact, the necessary and sufficient conditions studied in [3] make it possible to obtain a data-driven stabilizing feedback with minimal requirements on the data. Interestingly, such conditions may hold even if unique identification is not feasible. Data-driven stabilization in the presence of process and measurement noise has been studied in [6], [7]. Apart from computing stabilizing feedback gains, it has been shown that trajectory simulation [8] and construction of predictive controllers [9], [10] can be performed directly from time series data. In addition,

The work of Henk van Waarde was supported by the Dutch Research Council under the NWO Talent Programme Veni Agreement (VI.Veni.22.335). The work of Maurice Heemels was supported by the European Research Council (ERC) under the Advanced ERC grant agreement PROACTHIS, no. 101055384.

Amir Shakouri and Henk van Waarde are with the Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, The Netherlands (e-mail: a.shakouri@rug.nl; h.j.van.waarde@rug.nl).

Tren Baltussen and Maurice Heemels are with the Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands (e-mail: t.m.j.t.baltussen@tue.nl; m.heemels@tue.nl).

data-driven predictive control using frequency-domain data has also been studied in [11].

The majority of papers on data-driven control work in the setting where the parameters of the system are completely unknown, which is interesting in its own right. However, for most physical systems, this is a rather conservative modeling framework. In fact, we often have access to some *prior knowledge* on the system parameters, for example, because they represent physical quantities such as mass, spring constant, or conductance that are between given upper and lower bounds. For such cases, using both prior knowledge and data can lead to design methods that are less conservative than using data alone. For instance, in case computing a stabilizing feedback gain solely from the data is not feasible, prior knowledge could be used in conjunction with the data to enable such feedback design. This motivates developing methods that synthesize feedback laws by leveraging both the collected data and prior knowledge. Existing works on direct data-driven control that incorporate prior knowledge are rather scarce. So far, only prior knowledge in the form of bounds on the system parameters [12]–[14] and exact knowledge of some parameters [15] have been studied in the literature. In particular, it was shown in [12] that if the prior knowledge admits a linear fractional representation, one can combine such knowledge with the data to design a feedback law by solving LMIs. Compared to direct data-driven control, the use of prior knowledge in system identification has a richer history. For instance, subspace identification using the system's stability as prior knowledge has been studied in [16], [17]. This has been extended to incorporating eigenvalue constraints in [18]. Prior knowledge on other system-theoretic properties, such as positivity [19] and passivity [20], [21], is also among the investigated topics. The reader can refer to [22]–[24] and the references therein for other types of prior knowledge that have been used in system identification.

This note studies data-driven stabilization of LTI systems using prior knowledge on stabilizability and controllability. The incorporation of such system-theoretic properties has not yet received attention in data-driven control, and it poses significant technical challenges. This is, among others, due to the fact that sets of stabilizable and controllable systems are not convex, in contrast to the system sets considered in previous works [12]–[14]. Nevertheless, the inclusion of this new type of prior knowledge is highly relevant because, in many cases, it is known a priori that the system is either controllable or stabilizable. This information can, for instance, be deduced from the structure of the system matrices, which has been studied in detail in the literature on *structural controllability and stabilizability analysis* (see, e.g., [25] and the references therein).

In this work, we extend the data informativity framework of [3] to include prior knowledge by requiring the existence of a controller that stabilizes all systems consistent with the data *and* the prior knowledge. Our main results are twofold. First, we show that data-driven stabilization using controllability as prior knowledge is equivalent to data-driven stabilization without prior knowledge (Theorem 5). Therefore, if it is known that the true system is controllable, then this knowledge does not help in relaxing the conditions needed for data-driven stabilization. Next, we show that stabilizability as prior knowledge leads to necessary and sufficient

conditions that are weaker than those of data-driven stabilization without prior knowledge (Theorems 14 and 15). In addition, we provide a tractable method (Proposition 16), by which a stabilizing feedback gain can be computed from data while incorporating prior knowledge of stabilizability.

Notation: Let $\mathbb{Z}$, $\mathbb{Z}_+$, and $\mathbb{N}$ denote the sets of integers, non-negative integers, and positive integers, respectively. Let $M \in \mathbb{R}^{n \times m}$. The Moore-Penrose pseudoinverse of M is denoted by $M^\dagger$. The spectral norm of M is denoted by $\|M\|$. For a set $\mathcal{X} \subseteq \mathbb{R}^m$, we define $M\mathcal{X} := \{Mx \mid x \in \mathcal{X}\}$. In case M is square, we say it is *positive definite*, denoted by $M > 0$, if it is symmetric and all its eigenvalues are positive. We denote the reachable subspace of a pair $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ by $\mathcal{R}(A, B) := \text{im} \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$.

II. PROBLEM FORMULATION

Let $n, m \in \mathbb{N}$. Consider the LTI system

$$x(t+1) = A_{\text{true}}x(t) + B_{\text{true}}u(t), \quad (1)$$

referred to as the *true system*, where $t \in \mathbb{Z}_+$ denoted the time, $x(t) \in \mathbb{R}^n$ is the state, and $u(t) \in \mathbb{R}^m$ is the input. The system matrices

$$(A_{\text{true}}, B_{\text{true}}) \in \mathcal{M} := \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$$

are assumed to be unknown. However, we have access to input-state data of the form

$$\mathcal{D} := \left(\begin{bmatrix} u(0) & \cdots & u(T-1) \end{bmatrix}, \begin{bmatrix} x(0) & \cdots & x(T) \end{bmatrix} \right)$$

collected from (1) within the time horizon $T \in \mathbb{N}$. Given $\mathcal{D}$, we define the matrices

$$\begin{aligned} U_- &:= \begin{bmatrix} u(0) & \cdots & u(T-1) \end{bmatrix}, \\ X_- &:= \begin{bmatrix} x(0) & \cdots & x(T-1) \end{bmatrix}, \text{ and} \\ X_+ &:= \begin{bmatrix} x(1) & \cdots & x(T) \end{bmatrix}. \end{aligned}$$

A. Recap of Data-driven stabilization

Roughly speaking, data-driven stabilization aims at solving the following problem:

Given the data $\mathcal{D}$, find a K such that $A_{\text{true}} + B_{\text{true}}K$ is Schur.

Based on the collected input-state data, the true system satisfies

$$X_+ = A_{\text{true}}X_- + B_{\text{true}}U_-.$$

However, the true system may not be the only one that satisfies this identity. Therefore, a feedback gain that guarantees the stability of the true system must stabilize *all* systems (A, B) satisfying $X_+ = AX_- + BU_-$. Therefore, we define the set of *data-consistent systems* as

$$\Sigma_{\mathcal{D}} := \{(A, B) \in \mathcal{M} \mid X_+ = AX_- + BU_-\},$$

and we sharpen the data-driven stabilization problem as follows:

Given the data $\mathcal{D}$, find a K such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}}$.

The feasibility of this problem depends on the given data. To elaborate on this, we recap the following notion of data informativity.

Definition 1 ([3, Def. 12]): The data $\mathcal{D}$ are called *informative for stabilization* if there exists a $K \in \mathbb{R}^{m \times n}$ such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}}$.

The following result provides a necessary and sufficient LMI condition for the informativity of the data for stabilization.

Proposition 2 ([3, Thm. 17]): The data $\mathcal{D}$ are informative for stabilization if and only if there exists $\Theta \in \mathbb{R}^{T \times n}$ such that

$$X_- \Theta = \Theta^\top X_-^\top \quad \text{and} \quad \begin{bmatrix} X_- \Theta & X_- \Theta \\ \Theta^\top X_-^\top & X_- \Theta \end{bmatrix} > 0. \quad (2)$$

Moreover, $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}}$ if and only if $K = U_- \Theta (X_- \Theta)^{-1}$ for some Θ satisfying (2).

Based on Proposition 2, a necessary condition for the informativity of the data for stabilization is that $\text{rank } X_- = n$. This condition requires the number of data samples to satisfy $T \geq n$.

B. Data-driven stabilization using prior knowledge

Let $\Sigma_{\text{pk}} \subseteq \mathcal{M}$ be a set capturing our prior knowledge of the true system, i.e.,

$$(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\text{pk}}.$$

Using this prior knowledge, we extend the data-driven stabilization problem to the following.

Given the data $\mathcal{D}$ and the set of prior knowledge Σ_{pk} , find a K such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{pk}}$.

The feasibility of this problem depends on the given data and the prior knowledge. To study this problem, we extend the notion of data informativity in Definition 1 to the following, which takes the prior knowledge into account.

Definition 3: The data $\mathcal{D}$ are called Σ_{pk} -*informative for stabilization* if there exists a $K \in \mathbb{R}^{m \times n}$ such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{pk}}$.

We note that $\mathcal{M}$ -informativity for stabilization is equivalent to informativity for stabilization in the sense of Definition 1. In this work, we are interested in two important sets of prior knowledge that capture the *stabilizability* and *controllability* of the true system. Denote the sets of controllable and stabilizable systems, respectively, by

$$\Sigma_{\text{cont}} := \{(A, B) \in \mathcal{M} \mid (A, B) \text{ is controllable}\}, \text{ and}$$

$$\Sigma_{\text{stab}} := \{(A, B) \in \mathcal{M} \mid (A, B) \text{ is stabilizable}\}.$$

The following example demonstrates that, in case

$$\Sigma_{\text{pk}} = \Sigma_{\text{stab}},$$

the conditions for Σ_{pk} -informativity for stabilization are in general *weaker* than those for informativity for stabilization.

Example 4: Consider the input data $u(0) = 1$, $u(1) = 2$, and $u(2) = -1$, and the state data $x(0) = \begin{bmatrix} 1 & 0 \end{bmatrix}^\top$, $x(1) = \begin{bmatrix} 2 & 0 \end{bmatrix}^\top$, $x(2) = \begin{bmatrix} 4 & 0 \end{bmatrix}^\top$, and $x(3) = \begin{bmatrix} 3 & 0 \end{bmatrix}^\top$. The set of data-consistent systems reads

$$\Sigma_{\mathcal{D}} = \left\{ \left(\begin{bmatrix} 1 & \alpha \\ 0 & \beta \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \mid \alpha, \beta \in \mathbb{R} \right\}.$$

It follows from Proposition 2 that the data are not informative for stabilization since $X_- = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix}$ does not have full row rank. However, we have

$$\Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}} = \left\{ \left(\begin{bmatrix} 1 & \alpha \\ 0 & \beta \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \mid \alpha \in \mathbb{R}, |\beta| < 1 \right\}.$$

It is evident that $K = \begin{bmatrix} -1 & 0 \end{bmatrix}$ is a stabilizing feedback gain for all the systems in $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$. Therefore, the data are Σ_{stab} -informative for stabilization.

Example 4 shows that data-driven stabilization using stabilizability as prior knowledge may be possible in case Proposition 2 fails to provide a feedback gain for all data-consistent systems. This

motivates the study of stabilizability as prior knowledge for data-driven stabilization. Another closely related prior knowledge that is studied in this paper is controllability. Formally, we thus consider the following problem.

Problem 1: Find necessary and sufficient conditions under which the data $\mathcal{D}$ are (i) Σ_{cont} -informative for stabilization; (ii) Σ_{stab} -informative for stabilization.

III. CONTROLLABILITY AS PRIOR KNOWLEDGE

The main result of this section is the following theorem presenting the solution for Problem 1(i). This theorem shows that controllability as prior knowledge is *not* useful, i.e., the data-driven stabilization using this prior knowledge is equivalent to data-driven stabilization without any prior knowledge.

Theorem 5: Suppose that $(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\text{cont}}$. Then, the following statements are equivalent:

- (a) The data $\mathcal{D}$ are Σ_{cont} -informative for stabilization.
- (b) The data $\mathcal{D}$ are informative for stabilization.

Moreover, if K is such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{cont}}$, then $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}}$.

The following example provides an intuition on the fact that controllability as prior knowledge does not affect the conditions required for data-driven stabilization.

Example 6: For the sake of illustration, consider a system with $n = 1$ and $m = 2$ as follows:

$$x(t+1) = ax(t) + [b_1 \quad b_2] u(t).$$

Starting from $x(0) = -1$, we apply $u(0) = [1 \quad -1]^\top$ and we measure $x(1) = -1$. Given these data, the set $\Sigma_{\mathcal{D}}$ consists of all systems with parameters a , b_1 , and b_2 lying on the plane shown in Fig. 1. The only uncontrollable system on this plane corresponds to $a = 1$ and $b_1 = b_2 = 0$, which is shown by the red dot. Now, Theorem 5 states that if a feedback gain stabilizes all systems on the plane excluding the one shown in red, then it also stabilizes the red point.

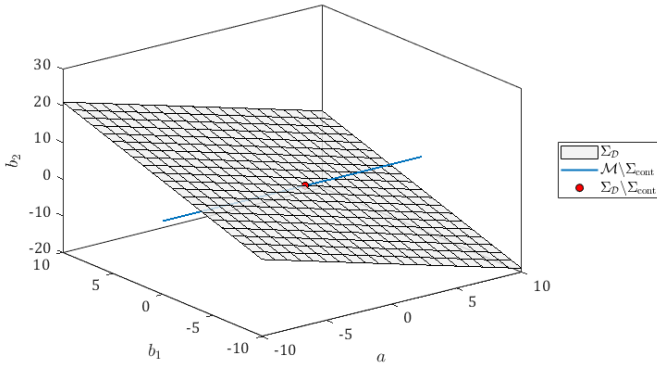


Fig. 1. Set of data-consistent systems, set of uncontrollable systems, and their intersection for Example 6.

To prove Theorem 5, we need three auxiliary results. The first result is the following lemma, showing that if a parameterized family of systems is controllable at a single point, then it is controllable at almost all points.

Lemma 7: Let $(M, N) \in \Sigma_{\text{cont}}$, $M_0 \in \mathbb{R}^{n \times n}$, and $N_0 \in \mathbb{R}^{n \times m}$. Then, the pair $(M + \alpha M_0, N + \alpha N_0)$ is controllable for all but at most n^2 values of $\alpha \in \mathbb{R}$.

Proof: Denote the Kalman controllability matrix of the pair $(M + \alpha M_0, N + \alpha N_0)$ by

$$\mathcal{C}(\alpha) := [N + \alpha N_0 \quad \cdots \quad (M + \alpha M_0)^{n-1}(N + \alpha N_0)].$$

Let $\bar{\mathcal{C}}(\alpha) \in \mathbb{R}^{n \times n}$ be a square submatrix of $\mathcal{C}(\alpha)$ such that $\bar{\mathcal{C}}(0)$ is nonsingular. Observe that the entries of $\bar{\mathcal{C}}(\alpha)$ are polynomials of α of degree at most n . It can be concluded from, e.g., Leibniz's formula for determinants, that $\det(\bar{\mathcal{C}}(\alpha))$ is a polynomial of α of degree at most n^2 . This polynomial has at most n^2 roots counting multiplicity, which implies that $\text{rank } \bar{\mathcal{C}}(\alpha) < n$ for at most n^2 distinct values of α . Therefore, $\mathcal{C}(\alpha)$ is rank deficient for at most n^2 distinct values of α . ■

The second auxiliary result is the following lemma, providing a necessary and sufficient condition for two matrices to share the same eigenvalues.

Lemma 8 ([26, 2.4.P10]): Let $M, N \in \mathbb{R}^{n \times n}$. Then, M and N share the same eigenvalues with the same algebraic multiplicities if and only if for every $k \in \{1, \dots, n\}$ we have $\text{tr}(M^k) = \text{tr}(N^k)$.

The third auxiliary result discusses the stability of matrix pencils, which is presented in the following lemma.

Lemma 9: Let $\varepsilon \in \mathbb{R}$, $\mathcal{F} \subset \mathbb{R}$ be a finite set, and $M, N \in \mathbb{R}^{n \times n}$ be such that $M + \delta N$ is Schur for all $\delta \in [\varepsilon, \infty) \setminus \mathcal{F}$. Then, N is nilpotent and $M + \delta N$ is Schur for all $\delta \in \mathbb{R}$.

Proof: Let $\bar{\delta} > 0$ be sufficiently large such that $[\bar{\delta}, \infty) \cap \mathcal{F} = \emptyset$. We note that the trace of a matrix is the sum of its eigenvalues. Since $M + \delta N$ is Schur for all $\delta \in [\bar{\delta}, \infty)$, we have

$$|\text{tr}((M + \delta N)^k)| \leq n,$$

for all $\delta \in [\bar{\delta}, \infty)$ and all $k \in \mathbb{N}$. Note that for every $k \in \mathbb{N}$,

$$p_k(\delta) = \text{tr}((M + \delta N)^k)$$

is a polynomial of degree at most k . The boundedness of $p_k(\delta)$ in the interval $[\bar{\delta}, \infty)$ implies that $p_k(\delta)$ is constant, i.e., the coefficients of δ^i are zero for all $i \in \{1, \dots, k\}$. In particular, the coefficient corresponding to δ^k , i.e., $\text{tr}(N^k)$, is equal to zero. Thus, $\text{tr}(N^k) = 0$ for all $k \in \mathbb{N}$. Based on Lemma 8, this implies that all eigenvalues of N are equal to zero. Thus, N is nilpotent. In addition, since $p_k(\delta)$ is constant for all $k \in \mathbb{N}$, we have $p_k(\delta) = p_k(0)$, thus, $\text{tr}(M^k) = \text{tr}((M + \delta N)^k)$ for all $k \in \mathbb{N}$. Now, we use Lemma 8 to conclude that M and $M + \delta N$ share the same eigenvalues for all $\delta \in \mathbb{R}$. Therefore, $M + \delta N$ is Schur for all $\delta \in \mathbb{R}$. ■

The proof of Theorem 5 now follows from Lemmas 7 and 9. To facilitate the proof, we introduce the following notation:

$$\Sigma_{\mathcal{D}}^0 := \{(A_0, B_0) \in \mathcal{M} \mid A_0 X_- + B_0 U_- = 0\}.$$

It is evident that the set $\Sigma_{\mathcal{D}}^0$ satisfies $\Sigma_{\mathcal{D}}^0 + \Sigma_{\mathcal{D}} = \Sigma_{\mathcal{D}}$.

Proof of Theorem 5: (b) $\Rightarrow$ (a): This implication is evident since if $K \in \mathbb{R}^{m \times n}$ is such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}}$, then $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{cont}}$.

(a) $\Rightarrow$ (b): This implication obviously holds if $\Sigma_{\mathcal{D}} \subseteq \Sigma_{\text{cont}}$. Now, assume that there exists $(\bar{A}, \bar{B}) \in \Sigma_{\mathcal{D}}$ that is not controllable. Let K be such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{cont}}$. It suffices to show that $\bar{A} + \bar{B}K$ is Schur. To that end, note that the pair $(\bar{A}, \bar{B})$ satisfies

$$X_+ = \bar{A}X_- + \bar{B}U_-.$$

Since the true system also satisfies $X_+ = A_{\text{true}}X_- + B_{\text{true}}U_-$, we have

$$\bar{A} = A_{\text{true}} + A_0 \quad \text{and} \quad \bar{B} = B_{\text{true}} + B_0,$$

for some $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$. We note that

$$(A_{\text{true}} + \alpha A_0, B_{\text{true}} + \alpha B_0) \in \Sigma_{\mathcal{D}} \quad \text{for all } \alpha \in \mathbb{R}.$$

Based on Lemma 7, since $(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\text{cont}}$, we have $(A_{\text{true}} + \alpha A_0, B_{\text{true}} + \alpha B_0) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{cont}}$ for all but a finite number of $\alpha \in \mathbb{R}$. Thus, $A_{\text{true}} + B_{\text{true}}K + \alpha(A_0 + B_0K)$ is Schur for all but a finite number of $\alpha \in \mathbb{R}$. Based on Lemma 9, this implies that

$A_{\text{true}} + B_{\text{true}}K + \alpha(A_0 + B_0K)$ is Schur for all $\alpha \in \mathbb{R}$. We take $\alpha = 1$ and conclude that $\bar{A} + \bar{B}K$ is Schur. ■

IV. STABILIZABILITY AS PRIOR KNOWLEDGE

In this section, we provide the solution for Problem 1(ii). First, we study necessary conditions for Σ_{stab} -informativity of the data for stabilization in Section IV-A. Next, we provide necessary and sufficient conditions for cases $\text{rank } X_- = n$ and $\text{rank } X_- < n$ in Sections IV-B and IV-C, respectively.

A. Necessary conditions

The following theorem presents four necessary conditions for Σ_{stab} -informativity for stabilization.

Theorem 10: Suppose that $(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\text{stab}}$ and the data $\mathcal{D}$ are Σ_{stab} -informative for stabilization. Let K be such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$. Then, the following statements hold:

- (a) $(A_0 + B_0K)\mathcal{R}(A, B) = \{0\}$ for all $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$ and all $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$.
- (b) If $\text{rank } X_- < n$, then $\text{im} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = \text{im } X_- \times \mathbb{R}^m$.
- (c) $\text{im } X_+ \subseteq \text{im } X_-$.
- (d) $\text{im } X_-$ is A -invariant and contains $\text{im } B$ for all $(A, B) \in \Sigma_{\mathcal{D}}$.

The following remark makes a comparison between the necessary conditions in Theorem 10 and the necessary conditions for data informativity without prior knowledge.

Remark 11: It was shown in [3, Lem. 15] that if the data $\mathcal{D}$ are informative for stabilization and K is a stabilizing gain for all system in $\Sigma_{\mathcal{D}}$, then $A_0 + B_0K = 0$ for all $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$. Here, we see that if the data are Σ_{stab} -informative for stabilization, then this condition is relaxed to that of statement (a). Now, one can observe that if $\Sigma_{\mathcal{D}}$ contains a controllable system, then statement (a) implies $A_0 + B_0K = 0$, which agrees with the result of Theorem 5. Moreover, we recall from Proposition 2 that, without using prior knowledge, a necessary condition for the informativity of the data for stabilization is that $\text{rank } X_- = n$. In that case, statements (c) and (d) obviously hold since $\text{im } X_- = \mathbb{R}^n$. However, these statements, along with (b), are nontrivial in case the data are not informative for stabilization, but Σ_{stab} -informative for stabilization.

To prove Theorem 10, we need some intermediate results presented next.

Lemma 12: Let $\mathcal{V} \subset \mathbb{R}^n$ be a proper subspace. Let $\hat{N} \in \mathbb{R}^{n \times m}$, $N_0 \in \mathbb{R}^{r \times m}$, and $\varepsilon > 0$. Define

$$\mathcal{N} := \left\{ \hat{N} + YN_0 \mid Y \in \mathbb{R}^{n \times r}, \|Y\| \leq \varepsilon \right\}.$$

Then, $\text{im } N \subseteq \mathcal{V}$ for all $N \in \mathcal{N}$ if and only if $\text{im } \hat{N} \subseteq \mathcal{V}$ and $N_0 = 0$.

Proof: The “if” part is obvious. To prove the “only if” part, we assume that $\text{im } N \subseteq \mathcal{V}$ for all $N \in \mathcal{N}$, i.e.,

$$\text{im}(\hat{N} + YN_0) \subseteq \mathcal{V} \quad (3)$$

for all $Y \in \mathbb{R}^{n \times r}$ with $\|Y\| \leq \varepsilon$. Taking $Y = 0$ shows that $\text{im } \hat{N} \subseteq \mathcal{V}$. This, together with (3), implies that

$$Y \text{im } N_0 \subseteq \mathcal{V} \quad (4)$$

for all $Y \in \mathbb{R}^{n \times r}$ with $\|Y\| \leq \varepsilon$. Let $\eta \in \mathbb{R}^r$. We take $Y = \xi\eta^\top$ with nonzero $\xi \in \mathbb{R}^n$ satisfying $\xi^\top \mathcal{V} = \{0\}$ and $\|\xi\eta^\top\| \leq \varepsilon$. We substitute this in (4) and premultiply by $\xi^\top$ to have

$$\|\xi\|^2 \eta^\top \text{im } N_0 \subseteq \xi^\top \mathcal{V} = \{0\}.$$

Since η was arbitrary, we have $\eta^\top N_0 = 0$ for all $\eta \in \mathbb{R}^r$, which implies $N_0 = 0$. ■

Lemma 13: Let $\mathcal{V} \subseteq \mathbb{R}^n$ be a subspace of dimension $r \leq n$. Let $M \in \mathbb{R}^{n \times n}$ and $v \in \mathcal{V}$. If $M^k v \in \mathcal{V}$ for all $k \in [1, r]$, then $M^k v \in \mathcal{V}$ for all $k \in \mathbb{N}$.

Proof: Suppose that $M^k v \in \mathcal{V}$ for all $k \in [1, r]$. We use induction to show that $M^k v \in \mathcal{V}$ for all $k \in \mathbb{N}$. Let $j \geq r$ be such that $M^k v \in \mathcal{V}$ for all $k \in [1, j]$. Since the dimension of $\mathcal{V}$ is equal to r , the matrix $[v \quad Mv \quad \cdots \quad M^j v]$ is of rank at most r . Since the number of columns of this matrix is larger than r , there exists $i \in [1, j]$ such that

$$M^i v \in \text{im} [v \quad Mv \quad \cdots \quad M^{i-1} v].$$

Multiply this from left by M^{j-i+1} to have

$$M^{j+1} v \in \text{im} [M^{j-i+1} v \quad M^{j-i+2} v \quad \cdots \quad M^j v] \subseteq \mathcal{V}.$$

Therefore, $M^k v \in \mathcal{V}$ for all $k \in [1, j+1]$, which completes the proof. ■

Lemmas 12 and 13 can now be used to prove Theorem 10.

Proof of Theorem 10: (a) Suppose that $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{cont}}$ is nonempty. Since the data are Σ_{stab} -informative for stabilization, they are also Σ_{cont} -informative for stabilization. Thus, we use Theorem 5 to conclude that the data are informative for stabilization and $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}}$. It follows from [3, Lem. 15] that we have $A_0 + B_0K = 0$ for all $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$. Therefore, statement (a) holds. Now, suppose that $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{cont}}$ is empty. Let $(A, B) \in \Sigma_{\mathcal{D}}$ be stabilizable but not controllable. Let $T \in \mathbb{R}^{n \times n}$ be nonsingular such that

$$TAT^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad \text{and} \quad TB = \begin{bmatrix} B_1 \\ 0 \end{bmatrix},$$

where $(A_{11}, B_1) \in \mathbb{R}^{n_1 \times n_1} \times \mathbb{R}^{n_1 \times m}$ is controllable, $A_{22} \in \mathbb{R}^{n_2 \times n_2}$ is Schur, and $n_1 + n_2 = n$. Also, let $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$ and $Z \in \mathbb{R}^{n_1 \times n}$. Define

$$\mathcal{A}(\alpha) := A + \alpha T^{-1} \begin{bmatrix} Z \\ 0 \end{bmatrix} A_0 \quad \text{and} \quad \mathcal{B}(\alpha) := B + \alpha T^{-1} \begin{bmatrix} Z \\ 0 \end{bmatrix} B_0.$$

First, we claim that

$$(\mathcal{A}(\alpha), \mathcal{B}(\alpha)) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}} \quad (5)$$

for all but at most n_1^2 values of α . To show this, observe that

$$\mathcal{A}(\alpha)X_- + \mathcal{B}(\alpha)U_- = AX_- + BU_- = X_+.$$

This implies that $(\mathcal{A}(\alpha), \mathcal{B}(\alpha)) \in \Sigma_{\mathcal{D}}$ for all $\alpha \in \mathbb{R}$. Now, let the matrices $R_1 \in \mathbb{R}^{n_1 \times n_1}$ and $R_2 \in \mathbb{R}^{n_1 \times n_2}$ be defined by $[R_1 \quad R_2] = ZA_0T^{-1}$. Observe that

$$T\mathcal{A}(\alpha)T^{-1} = \begin{bmatrix} A_{11} + \alpha R_1 & A_{12} + \alpha R_2 \\ 0 & A_{22} \end{bmatrix} \quad \text{and} \\ T\mathcal{B}(\alpha) = \begin{bmatrix} B_1 + \alpha ZB_0 \\ 0 \end{bmatrix}.$$

Since (A_{11}, B_1) is controllable, we use Lemma 7 with $M = A_{11}$, $M_0 = R_1$, $N = B_1$, and $N_0 = ZB_0$ to conclude that

$$(A_{11} + \alpha R_1, B_1 + \alpha ZB_0) \in \Sigma_{\text{cont}}$$

for all but at most n_1^2 values of α . Moreover, since A_{22} is Schur, we have

$$(T\mathcal{A}(\alpha)T^{-1}, T\mathcal{B}(\alpha)) \in \Sigma_{\text{stab}},$$

and thus, $(\mathcal{A}(\alpha), \mathcal{B}(\alpha)) \in \Sigma_{\text{stab}}$ for all but at most n_1^2 values of α .

Next, we use inclusion (5) to show that statement (a) holds. Let $\mathcal{F}$ be the set of all values of α such that $(\mathcal{A}(\alpha), \mathcal{B}(\alpha)) \notin \Sigma_{\text{stab}}$. By the previous discussion, $\mathcal{F}$ is finite. Since K is a stabilizing gain for all systems within $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$, we have that

$$\mathcal{A}(\alpha) + \mathcal{B}(\alpha)K = A + BK + \alpha T^{-1} \begin{bmatrix} Z \\ 0 \end{bmatrix} (A_0 + B_0K)$$

is Schur for all $\alpha \in \mathbb{R} \setminus \mathcal{F}$. Let $N \in \mathbb{R}^{n \times n_1}$ and $M \in \mathbb{R}^{n \times n_2}$ be such that $\begin{bmatrix} N & M \end{bmatrix} = (A_0 + B_0 K)T^{-1}$. Also, define $F \in \mathbb{R}^{m \times n_1}$ and $G \in \mathbb{R}^{m \times n_2}$ by $\begin{bmatrix} F & G \end{bmatrix} = KT^{-1}$. We observe that

$$T(A(\alpha) + B(\alpha)K)T^{-1} = \begin{bmatrix} A_{11} + B_1 F + \alpha ZN & A_{12} + B_1 G + \alpha ZM \\ 0 & A_{22} \end{bmatrix}.$$

Hence, $A_{11} + B_1 F + \alpha ZN$ is Schur for all $\alpha \in \mathbb{R} \setminus \mathcal{F}$. It follows now from Lemma 9 that $A_{11} + B_1 F$ is Schur and ZN is nilpotent. Note that this argument holds for all $Z \in \mathbb{R}^{n_1 \times n}$ and all $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$. Take $Z = N^\top$. Since $N^\top N$ is symmetric and nilpotent, we have $N = 0$. Hence, for any $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$ we have $A_0 + B_0 K = \begin{bmatrix} 0 & M \end{bmatrix} T$. This implies that

$$\begin{aligned} (A_0 + B_0 K) \text{im} \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \\ = \begin{bmatrix} 0 & M \end{bmatrix} T \text{im} \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \\ = \begin{bmatrix} 0 & M \end{bmatrix} \text{im} \begin{bmatrix} B_1 & A_{11}B_1 & \cdots & A_{11}^{n-1}B_1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \{0\}. \end{aligned}$$

Therefore, statement (a) holds.

(b) To prove this part, first, we claim that

$$\mathcal{R}(A, B) \subseteq \text{im } X_- \text{ for all } (A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}. \quad (6)$$

To show this, let $A_0 \in \mathbb{R}^{n \times n}$ be such that $\ker A_0 = \text{im } X_-$. Observe that $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$ with $B_0 = 0$. It follows from part (a) that $A_0 \mathcal{R}(A, B) = \{0\}$. This implies that $\mathcal{R}(A, B) \subseteq \ker A_0 = \text{im } X_-$. Now, let $\xi \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^m$ be such that $\begin{bmatrix} \xi^\top & \eta^\top \end{bmatrix}^\top \in \ker \begin{bmatrix} X_-^\top & U_-^\top \end{bmatrix}$. Since $(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\mathcal{D}}$, for every $Y \in \mathbb{R}^n$ we have $(A_{\text{true}} + Y\xi^\top, B_{\text{true}} + Y\eta^\top) \in \Sigma_{\mathcal{D}}$. Let $\varepsilon > 0$ be small enough such that for every $Y \in \mathbb{R}^n$ satisfying $\|Y\| \leq \varepsilon$ we have $(A_{\text{true}} + Y\xi^\top, B_{\text{true}} + Y\eta^\top) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$. It follows from (6) that $\text{im}(B_{\text{true}} + Y\eta^\top) \subseteq \text{im } X_-$ for all $Y \in \mathbb{R}^n$ satisfying $\|Y\| \leq \varepsilon$. We use Lemma 12 with $\hat{N} = B_{\text{true}}$ and $N_0 = \eta^\top$ to conclude that $\eta = 0$. Therefore, we have

$$\ker \begin{bmatrix} X_- \\ U_- \end{bmatrix}^\top = \ker X_-^\top \times \{0\}, \quad (7)$$

which implies (b).

(c) To prove this part, it suffices to show that $x(T) \in \text{im } X_-$. If X_- has full row rank, then this condition is obviously satisfied. Suppose that $\text{rank } X_- = r < n$. It follows from part (b) that $T \geq r + m$. Let $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$. Observe that we have

$$x(t) = A^t x(0) + \sum_{k=0}^{t-1} A^k B u(t-k-1) \quad (8)$$

for all $t \in [1, T]$. It follows from (6) that the last term satisfies

$$\sum_{k=0}^{t-1} A^k B u(t-k-1) \in \mathcal{R}(A, B) \subseteq \text{im } X_- \text{ for all } t \in [1, T].$$

Since $x(t) \in \text{im } X_-$ for all $t \in [1, T-1]$, we have $A^t x(0) \in \text{im } X_-$ for all $t \in [0, T-1]$. As $T-1 \geq r$, we use Lemma 13 with $\mathcal{V} = \text{im } X_-$ and $M = A$ to conclude that $A^T x(0) \in \text{im } X_-$. Now, it follows from (8) with $t = T$ that $x(T) \in \text{im } X_-$.

(d) If X_- has full row rank, then (d) obviously holds. Now, suppose that X_- does not have full row rank. Let $(A, B) \in \Sigma_{\mathcal{D}}$. First, we show that $\text{im } B \subseteq \text{im } X_-$. For this, observe that since both pairs $(A_{\text{true}}, B_{\text{true}})$ and (A, B) belong to $\Sigma_{\mathcal{D}}$, we have

$$(A - A_{\text{true}})X_- + (B - B_{\text{true}})U_- = 0. \quad (9)$$

Based on (7) and (9), we have $B = B_{\text{true}}$. It follows now from (6) that $\text{im } B = \text{im } B_{\text{true}} \subseteq \mathcal{R}(A_{\text{true}}, B_{\text{true}}) \subseteq \text{im } X_-$. To show that $\text{im } X_-$ is A -invariant, we observe from part (c) that

$A \text{im } X_- + B \text{im } U_- = \text{im } X_+ \subseteq \text{im } X_-$. Since $\text{im } B \subseteq \text{im } X_-$, we have $A \text{im } X_- \subseteq \text{im } X_-$, which completes the proof. ■

B. Necessary and sufficient conditions with full rank state data

In case X_- has full row rank, the following theorem shows that data-driven stabilization using stabilizability as prior knowledge is equivalent to data-driven stabilization without prior knowledge.

Theorem 14: Suppose that $(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\text{stab}}$ and $\text{rank } X_- = n$. Then, the following statements are equivalent:

- (a) The data $\mathcal{D}$ are Σ_{stab} -informative for stabilization.
- (b) The data $\mathcal{D}$ are informative for stabilization.

Proof: It is evident that (b) implies (a). To prove that (a) implies (b), first, we assume that $(A_{\text{true}}, B_{\text{true}})$ is controllable. Since the data $\mathcal{D}$ are Σ_{stab} -informative for stabilization, they are also Σ_{cont} -informative for stabilization. Hence, it follows from Theorem 5 that the data are informative for stabilization.

Next, we assume that $(A_{\text{true}}, B_{\text{true}})$ is uncontrollable. Let K be a stabilizing gain for all the systems in $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$. We show that there exists a $\hat{K}$ that stabilizes all the systems in $\Sigma_{\mathcal{D}}$. For this, let $\mathcal{S}$ be a subspace satisfying $\mathcal{S} \oplus \mathcal{R}(A_{\text{true}}, B_{\text{true}}) = \mathbb{R}^n$, where $\oplus$ denotes direct sum. Thus, every $v \in \mathbb{R}^n$ can be written uniquely as $v = v_1 + v_2$ with $v_1 \in \mathcal{R}(A_{\text{true}}, B_{\text{true}})$ and $v_2 \in \mathcal{S}$. We define $\hat{K}$ as the matrix satisfying $\hat{K}v = Kv_1 + U_- X_-^\top v_2$ for all $v \in \mathbb{R}^n$. Let $(A, B) \in \Sigma_{\mathcal{D}}$. Take $(A_0, B_0) \in \Sigma_{\mathcal{D}}^0$ such that $A = A_{\text{true}} + A_0$ and $B = B_{\text{true}} + B_0$. We show that $A_0 + B_0 \hat{K} = 0$. Since X_- has full row rank, we have $A_0 = -B_0 U_- X_-^\top$. Hence, we have

$$(A_0 + B_0 \hat{K})v = A_0 v + B_0 K v_1 + B_0 U_- X_-^\top v_2 = (A_0 + B_0 K)v_1.$$

It follows from Theorem 10(a) that $(A_0 + B_0 K)v_1 = 0$. Thus, we have $(A_0 + B_0 \hat{K})v = 0$ for all $v \in \mathbb{R}^n$, which implies that $A_0 + B_0 \hat{K} = 0$. Therefore, we have

$$A + B \hat{K} = A_{\text{true}} + A_0 + (B_{\text{true}} + B_0) \hat{K} = A_{\text{true}} + B_{\text{true}} \hat{K}.$$

Hence, what remains to be proven is that $A_{\text{true}} + B_{\text{true}} \hat{K}$ is Schur. Let $\hat{T} \in \mathbb{R}^{n \times n}$ be such that

$$\hat{T} A_{\text{true}} \hat{T}^{-1} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix} \quad \text{and} \quad \hat{T} B_{\text{true}} = \begin{bmatrix} \hat{B}_1 \\ 0 \end{bmatrix},$$

where $(\hat{A}_{11}, \hat{B}_1) \in \mathbb{R}^{n_1 \times n_1} \times \mathbb{R}^{n_1 \times m}$ is controllable, $\hat{A}_{22} \in \mathbb{R}^{n_2 \times n_2}$ is Schur, and $n_1 + n_2 = n$. We note that the first n_1 columns of $\hat{T}^{-1}$ span $\mathcal{R}(A_{\text{true}}, B_{\text{true}})$ and the rest of its columns span $\mathcal{S}$. Let $F \in \mathbb{R}^{m \times n_1}$ and $G \in \mathbb{R}^{m \times n_2}$ be defined by $\begin{bmatrix} F & G \end{bmatrix} = K \hat{T}^{-1}$. Since $A_{\text{true}} + B_{\text{true}} K$ is Schur, we have that $\hat{A}_{11} + \hat{B}_1 F$ is Schur. Now, we observe that

$$T(A_{\text{true}} + B_{\text{true}} \hat{K})T^{-1} = \begin{bmatrix} \hat{A}_{11} + \hat{B}_1 F & \hat{A}_{12} + \hat{B}_1 \hat{G} \\ 0 & \hat{A}_{22} \end{bmatrix},$$

where $\hat{G} \in \mathbb{R}^{m \times n_2}$ satisfies $\begin{bmatrix} F & \hat{G} \end{bmatrix} = \hat{K} \hat{T}^{-1}$. Since $\hat{A}_{11} + \hat{B}_1 F$ and $\hat{A}_{22}$ are both Schur, we have that $A_{\text{true}} + B_{\text{true}} \hat{K}$ is Schur. Therefore, $A + B \hat{K}$ is Schur. This implies that the data $\mathcal{D}$ are informative for stabilization. ■

Due to Theorem 14, in case the state data X_- is of full row rank, stabilizability as prior knowledge does not help in relaxing the conditions for data-driven stabilization. Hence, in this case, a stabilizing feedback gain may still be computed using Proposition 2. Nevertheless, if the true system is not controllable, then collecting full rank state data might not be possible. In fact, this depends on the initial condition of the system. For instance, if an uncontrollable system is initially at rest, $x(0) = 0$, then X_- will not have full row rank, no matter what input signal is applied to the system. This motivates studying the case where the state data is rank-deficient, which is the topic of the next section.

C. Necessary and sufficient conditions with rank-deficient state data

In case X_- does not have full row rank, the data are not informative for stabilization, and Proposition 2 fails to provide a stabilizing feedback. Interestingly, it turns out that in this case, one may be able to find a stabilizing feedback from data by incorporating the prior knowledge on stabilizability.

Theorem 15: Suppose that $(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\text{stab}}$ and $\text{rank } X_- < n$. Then, the data $\mathcal{D}$ are Σ_{stab} -informative for stabilization if and only if the following conditions hold:

- (a) $\text{im } X_+ \subseteq \text{im } X_-$,
- (b) $\text{im } \begin{bmatrix} X_- \\ U_- \end{bmatrix} = \text{im } X_- \times \mathbb{R}^m$.

Theorem 15 provides a full characterization of Σ_{stab} -informativity for stabilization in case $\text{rank } X_- < n$. Before providing a proof, we first also consider the problem of computing a stabilizing feedback gain from Σ_{stab} -informative data. For this, given data $\mathcal{D}$, let $r := \text{rank } X_-$, $S \in \mathbb{R}^{n \times n}$ be nonsingular, and $\hat{X}_- \in \mathbb{R}^{r \times T}$ be of full row rank such that

$$SX_- = \begin{bmatrix} \hat{X}_- \\ 0 \end{bmatrix}. \quad (10)$$

These matrices S and $\hat{X}_-$ can be computed, for example, using a QR decomposition of X_- . Moreover, let $\hat{X}_+ \in \mathbb{R}^{r \times n}$ be defined as

$$\hat{X}_+ := [I_r \quad 0] SX_+.$$

Proposition 16: Suppose that $(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\text{stab}}$, the data $\mathcal{D}$ are Σ_{stab} -informative for stabilization, and $\text{rank } X_- < n$. Then, the following statements hold:

- (a) There exists $\Theta \in \mathbb{R}^{T \times r}$ such that the following LMI is feasible:

$$\hat{X}_- \Theta = \Theta^\top \hat{X}_-^\top \quad \text{and} \quad \begin{bmatrix} \hat{X}_- \Theta & \hat{X}_+ \Theta \\ \Theta^\top \hat{X}_-^\top & \hat{X}_+ \Theta \end{bmatrix} > 0. \quad (11)$$

- (b) Suppose that Θ satisfies LMI (11). Let $K = \begin{bmatrix} K_1 & K_2 \end{bmatrix} S$, where $K_1 = U_- \Theta (\hat{X}_- \Theta)^{-1}$ and $K_2 \in \mathbb{R}^{m \times (n-r)}$ is arbitrary. Then, $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$.

To prove Theorem 15 and Proposition 16, we need an auxiliary result presented in the following lemma. This lemma shows that if the data satisfy conditions (a) and (b) of Theorem 15, then S can be used as a state transformation to simultaneously factorize all members of $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$ into a stabilizable part and an autonomous, stable part.

Lemma 17: Suppose that the data $\mathcal{D}$ satisfy $\text{rank } X_- < n$. Let $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$. If statements (a) and (b) of Theorem 15 hold, then we have

$$SAS^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad \text{and} \quad SB = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \quad (12)$$

where the pair $(A_{11}, B_1) \in \mathbb{R}^{r \times r} \times \mathbb{R}^{r \times m}$ is stabilizable and $A_{22} \in \mathbb{R}^{(n-r) \times (n-r)}$ is Schur. Moreover, we have

$$\begin{bmatrix} A_{11} & B_1 \end{bmatrix} = \hat{X}_+ \begin{bmatrix} \hat{X}_- \\ U_- \end{bmatrix}^\dagger. \quad (13)$$

Proof: Write $SAS^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ and $SB = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, where $A_{21} \in \mathbb{R}^{(n-r) \times r}$ and $B_2 \in \mathbb{R}^{(n-r) \times m}$. First, we show that A_{21} and B_2 are both equal to zero. Since $\text{im } X_+ \subseteq \text{im } X_-$, we have $SX_+ = \begin{bmatrix} \hat{X}_+ \\ 0 \end{bmatrix}$. It follows from $SX_+ = (SAS^{-1})SX_- + SBU_-$ that $0 = A_{21}\hat{X}_- + B_2U_-$ and $\hat{X}_+ = A_{11}\hat{X}_- + B_1U_-$. Since $\text{im } \begin{bmatrix} \hat{X}_- \\ U_- \end{bmatrix}^\top = \text{im } X_- \times \mathbb{R}^m$, we have $\text{rank } \begin{bmatrix} \hat{X}_-^\top & U_-^\top \end{bmatrix} = r + m$. Thus, $0 = A_{21}\hat{X}_- + B_2U_-$

implies that $A_{21} = 0$ and $B_2 = 0$. Now, we observe that $(A, B) \in \Sigma_{\text{stab}}$ implies $(SAS^{-1}, SB) \in \Sigma_{\text{stab}}$. This, together with $B_2 = 0$ and $A_{21} = 0$, implies that A_{22} is Schur and (A_{11}, B_1) is stabilizable. Finally, the formula in (13) follows immediately from $\hat{X}_+ = \begin{bmatrix} A_{11} & B_1 \end{bmatrix} \begin{bmatrix} \hat{X}_-^\top & U_-^\top \end{bmatrix}^\top$ and the fact that $\begin{bmatrix} \hat{X}_-^\top & U_-^\top \end{bmatrix}$ is of full column rank. ■

We note that, unlike Kalman decomposition, the data-driven decomposition provided by Lemma 17 does not guarantee the pair (A_{11}, B_1) to be controllable. Now, we use Lemma 12 and Theorem 10 to prove Theorem 15.

Proof of Theorem 15: The “only if” part follows immediately from parts (b) and (c) of Theorem 10. To prove the “if” part, assume that conditions (a) and (b) hold. Since $(A_{\text{true}}, B_{\text{true}}) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$, it follows from Lemma 17 that

$$SA_{\text{true}}S^{-1} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix} \quad \text{and} \quad SB_{\text{true}} = \begin{bmatrix} \hat{B}_1 \\ 0 \end{bmatrix}, \quad (14)$$

where $(\hat{A}_{11}, \hat{B}_1) \in \mathbb{R}^{r \times r} \times \mathbb{R}^{r \times m}$ is stabilizable and $\hat{A}_{22} \in \mathbb{R}^{(n-r) \times (n-r)}$ is Schur. Since $(A_{\text{true}}, B_{\text{true}})$ is stabilizable, there exists a K such that $A_{\text{true}} + B_{\text{true}}K$ is Schur. We show that K is a stabilizing feedback gain for all systems belonging to $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$. For this, let $K_1 \in \mathbb{R}^{m \times r}$ and $K_2 \in \mathbb{R}^{m \times (n-r)}$ be defined by $\begin{bmatrix} K_1 & K_2 \end{bmatrix} = KS^{-1}$ and observe that

$$S(A_{\text{true}} + B_{\text{true}}K)S^{-1} = \begin{bmatrix} \hat{A}_{11} + \hat{B}_1K_1 & \hat{A}_{12} + \hat{B}_1K_2 \\ 0 & \hat{A}_{22} \end{bmatrix}$$

is Schur. This implies that $\hat{A}_{11} + \hat{B}_1K_1$ is Schur. Now, let $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$. Define A_{11} , A_{12} , A_{22} , and B_1 as in (12). In view of Lemma 17, A_{22} is Schur and we have

$$\begin{bmatrix} A_{11} & B_1 \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{B}_1 \end{bmatrix} = \hat{X}_+ \begin{bmatrix} \hat{X}_- \\ U_- \end{bmatrix}^\dagger. \quad (15)$$

Therefore, since $\hat{A}_{11} + \hat{B}_1K_1$ is Schur, we have that $A_{11} + B_1K_1$ is Schur. Hence,

$$S(A + BK)S^{-1} = \begin{bmatrix} A_{11} + B_1K_1 & A_{12} + B_1K_2 \\ 0 & A_{22} \end{bmatrix},$$

is Schur. Therefore, $A + BK$ is Schur, which implies that the data $\mathcal{D}$ are Σ_{stab} -informative for stabilization. ■

Proof of Proposition 16. (a) Since the data $\mathcal{D}$ are Σ_{stab} -informative for stabilization, the conditions (a) and (b) in Theorem 15 hold. Let $\hat{A}_{11}$, $\hat{A}_{12}$, $\hat{A}_{22}$, and $\hat{B}_1$ be defined as in (14). It follows from Lemma 17 that $\hat{A}_{22}$ is Schur. Let K be such that $A_{\text{true}} + B_{\text{true}}K$ is Schur. Define $K_1 \in \mathbb{R}^{m \times n_1}$ and $K_2 \in \mathbb{R}^{m \times n_2}$ by $\begin{bmatrix} K_1 & K_2 \end{bmatrix} = KS^{-1}$. Observe that

$$S(A_{\text{true}} + B_{\text{true}}K)S^{-1} = \begin{bmatrix} \hat{A}_{11} + \hat{B}_1K_1 & \hat{A}_{12} + \hat{B}_1K_2 \\ 0 & \hat{A}_{22} \end{bmatrix}.$$

Since $S(A_{\text{true}} + B_{\text{true}}K)S^{-1}$ is Schur, this implies that $\hat{A}_{11} + \hat{B}_1K_1$ is Schur. Now, it follows from (15) that

$$\hat{A}_{11} + \hat{B}_1K_1 = \hat{X}_+ \begin{bmatrix} \hat{X}_- \\ U_- \end{bmatrix}^\dagger \begin{bmatrix} I \\ K_1 \end{bmatrix}.$$

Let $P > 0$ be such that the Lyapunov inequality

$$P - (\hat{A}_{11} + \hat{B}_1K_1)P(\hat{A}_{11} + \hat{B}_1K_1)^\top > 0$$

holds. Now, take $\Theta = \begin{bmatrix} \hat{X}_- \\ U_- \end{bmatrix}^\dagger \begin{bmatrix} I \\ K_1 \end{bmatrix} P$. We show that, with this Θ , the LMI (11) is feasible. To this end, we show that the lower block of (11), $\hat{X}_- \Theta$, and its Schur complement with respect to the lower block, $\hat{X}_- \Theta - \hat{X}_+ \Theta (\hat{X}_- \Theta)^{-1} \Theta^\top \hat{X}_+^\top$, are positive definite. First,

we note that $\hat{X}_-\Theta = P > 0$ and $\hat{X}_+\Theta P^{-1} = \hat{A}_{11} + \hat{B}_1 K_1$. Hence, we have

$$\begin{aligned} \hat{X}_-\Theta - \hat{X}_+\Theta(\hat{X}_-\Theta)^{-1}\Theta^\top\hat{X}_+^\top &= P - \hat{X}_+\Theta P^{-1}\Theta^\top\hat{X}_+^\top \\ &= P - (\hat{A}_{11} + \hat{B}_1 K_1)P(\hat{A}_{11} + \hat{B}_1 K_1)^\top > 0. \end{aligned}$$

Therefore, LMI (11) is feasible.

(b) Assume that Θ satisfies LMI (11). Let $(A, B) \in \Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$, $K_1 = U_- \Theta (\hat{X}_- \Theta)^{-1}$, $K_2 \in \mathbb{R}^{m \times r}$, and $K = [K_1 \ K_2] S$. We show that $A + BK$ is Schur. We note that conditions (a) and (b) of Theorem 15 hold. Hence, let A_{11} , A_{12} , A_{22} , and B_1 be defined as in Lemma 17. Hence,

$$S(A + BK)S^{-1} = \begin{bmatrix} A_{11} + B_1 K_1 & A_{12} + B_1 K_2 \\ 0 & A_{22} \end{bmatrix}.$$

Since A_{22} is Schur, it suffices now to show that $A_{11} + B_1 K_1$ is Schur. Since $\hat{X}_+ = A_{11} + B_1 U_-$, we have

$$\begin{aligned} A_{11} + B_1 K_1 &= A_{11} + B_1 U_- \Theta (\hat{X}_- \Theta)^{-1} \\ &= (A_{11} \hat{X}_- + B_1 U_-) \Theta (\hat{X}_- \Theta)^{-1} = \hat{X}_+ \Theta (\hat{X}_- \Theta)^{-1}. \end{aligned}$$

Now, based on a Schur complement argument, we observe that LMI (11) implies

$$\begin{aligned} \hat{X}_-\Theta - \hat{X}_+\Theta(\hat{X}_-\Theta)^{-1}\Theta^\top\hat{X}_+^\top \\ = \hat{X}_-\Theta - (A_{11} + B_1 K_1)(\hat{X}_-\Theta)(A_{11} + B_1 K_1)^\top > 0. \end{aligned}$$

Since $\hat{X}_-\Theta > 0$, this implies that $A_{11} + B_1 K_1$ is Schur. ■

We close this section with two remarks.

Remark 18: Suppose that the data $\mathcal{D}$ are Σ_{stab} -informative for stabilization, but they are *not* informative for stabilization. In this case, the following facts hold: Theorems 5 implies that $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{cont}}$ is empty, hence, the true system is stabilizable but not controllable; Theorems 14 and 15 imply that the state data satisfy $\text{rank } X_- < n$; Theorem 15(b) implies that B_{true} can be uniquely recovered from the data, i.e., every $(A, B) \in \Sigma_{\mathcal{D}}$ is such that $B = B_{\text{true}}$.

Remark 19: It was shown in [27, Sec. IV] that if there exists a K such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}}$, then there exists a $P > 0$ such that the Lyapunov inequality $P - (A + BK)P(A + BK)^\top > 0$ holds for all $(A, B) \in \Sigma_{\mathcal{D}}$, i.e., the data $\mathcal{D}$ are informative for stabilization *if and only if* they are informative for quadratic stabilization (see [27, Def. 2]). However, we note that this does not necessarily hold in case the data are Σ_{stab} -informative for stabilization. This can be shown using Example 4. In that example, the closed-loop system for every member of $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$ is of the form

$$A + BK = \begin{bmatrix} 0 & \alpha \\ 0 & \beta \end{bmatrix} \text{ for some } \alpha \in \mathbb{R} \text{ and } |\beta| < 1.$$

All such matrices are Schur. However, we show that there does not exist a common P for all these matrices. Aiming for a contradiction, suppose that there exists a positive definite $P := \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ with the property that the Lyapunov inequality,

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} - \begin{bmatrix} 0 & \alpha \\ 0 & \beta \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \alpha & \beta \end{bmatrix} > 0, \quad (16)$$

holds for all $\alpha \in \mathbb{R}$ and $|\beta| < 1$. Multiply (16) from left and right, respectively, by $\begin{bmatrix} 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \end{bmatrix}^\top$ to have $P_{11} - \alpha^2 P_{22} > 0$. Since $P_{22} > 0$, it is evident that there exists a sufficiently large α such that this inequality does not hold. Hence, we reach a contradiction. Therefore, the Lyapunov inequality (16) does not hold for all closed-loop systems consistent with the data and the prior knowledge.

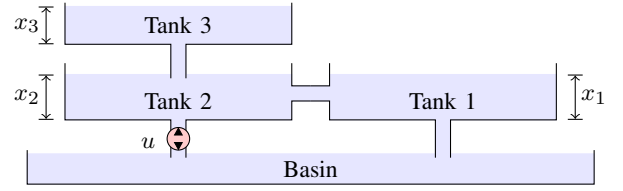


Fig. 2. Schematic representation of the three-tank system.

V. NUMERICAL EXAMPLE

To demonstrate our results, we consider a three-tank system, depicted in Fig. 2, which is described by the continuous-time model:

$$\dot{x}(\tau) = A_c x(\tau) + B_c u(\tau),$$

where $\tau \in \mathbb{R}$, $x(\tau) \in \mathbb{R}^3$ is the state vector, and $u(\tau) \in \mathbb{R}$ is the control input. The system matrices A_c and B_c are given by

$$A_c = \begin{bmatrix} -\frac{k_{01}+k_{12}}{a_1} & \frac{k_{12}}{a_1} & 0 \\ \frac{k_{12}}{a_2} & -\frac{k_{12}}{a_2} & \frac{k_{23}}{a_2} \\ 0 & 0 & -\frac{k_{23}}{a_3} \end{bmatrix} \text{ and } B_c = \begin{bmatrix} 0 \\ \frac{1}{a_2} \\ 0 \end{bmatrix},$$

with tank areas $a_1 = a_2 = a_3 = 1$ and flow coefficients $k_{01} = 0.1$, $k_{12} = 0.5$, and $k_{23} = 0.5$. The i th entry of the state, x_i , is the height of the fluid in tank $i = 1, 2, 3$, and the input u is the flow rate between tank 2 and the basin. Due to the structural properties of this system, the third mode of the system, x_3 , is uncontrollable. This lack of controllability arises from the directional flow from tank 3 to 2. The system is discretized using the zero-order hold method with a sampling time of 0.1, yielding the matrices

$$A_{\text{true}} = \begin{bmatrix} 0.9429 & 0.0473 & 0.0012 \\ 0.0473 & 0.9524 & 0.0476 \\ 0 & 0 & 0.9512 \end{bmatrix} \text{ and } B_{\text{true}} = \begin{bmatrix} 0.0024 \\ 0.0976 \\ 0 \end{bmatrix}.$$

We perform an open-loop experiment of length $T = 5$, during which we collect the following input-state data:

t	0	1	2	3	4	5
$u(t)$	1	0	-1	0	1	
$x(t)$	1	1.04	1.0778	1.1086	1.1334	1.1575
	2	2.0498	2.0015	1.8597	1.8237	1.8881
	0	0	0	0	0	0

Here, X_- has rank $r = 2$. A matrix S , satisfying (10), can be simply taken to be $S = I$. We solve (11) using MATLAB¹ with YALMIP [28] and MOSEK [29], which yields

$$\Theta = \begin{bmatrix} -47.4426 & -30.3733 & -1.5964 & 49.2034 & 36.0139 \\ -0.9001 & 17.7153 & 32.3315 & 20.4120 & -68.9591 \end{bmatrix}^\top.$$

Following Proposition 16, we compute $K_1 = [-2.7728 \ -9.7123]$ and set K_2 to zero, which yields $K = [-2.7728 \ -9.7123 \ 0]$ that is a stabilizing feedback gain for all systems within $\Sigma_{\mathcal{D}} \cap \Sigma_{\text{stab}}$.

To illustrate the advantage of incorporating prior knowledge of stabilizability, we perform Monte Carlo simulations with 1000 scenarios. For each scenario, we simulate the system from $t = 0$ to $t = 100$ with the input sequence and the entries of the initial condition drawn independently at random from a Poisson distribution with parameter $\lambda = 1$. To investigate the effect of the number of samples on the informativity of the data, we use the first T samples for each round of analysis. Here, we take $T = 3, 4, 5, 10$, and finally, the entire dataset $T = 100$, see Table I. We see that for $T = 3$, none of the scenarios yield informative data for system identification because

¹The MATLAB code for this example is available at <https://github.com/TrenBaltussen/Data-Driven-Stabilization>.

TABLE I
INFORMATIVITY OF RANDOMLY GENERATED DATA.

T	Informative for system identification	Σ_{pk} -informative for stabilization	
		$\Sigma_{pk} = \mathcal{M}$	$\Sigma_{pk} = \Sigma_{stab}$
3	0%	8.1%	42%
4	62.4%	63.2%	99.4%
5	62.8%	63.2%	99.8%
10	63.2%	63.2%	100%
100	63.2%	63.2%	100%

$T < n + m$, and thus, $[X^\top \ U^\top]$ does not have full column rank, which is a necessary and sufficient condition for system identification [3, Prop. 6]. Nevertheless, in this case, 8.1% of the datasets are informative for stabilization (without using prior knowledge). At this point, 42% of the datasets are Σ_{stab} -informative for stabilization. By increasing the number of samples, the percentage of datasets that are informative for identification approaches that of stabilization (without using prior knowledge). Both of these numbers eventually reach 63.2% and remain unchanged. Interestingly, with $T \geq 10$, we see that all the datasets are Σ_{stab} -informative for stabilization. This clearly demonstrates the advantage of incorporating stabilizability as prior knowledge in data-driven control.

VI. CONCLUSIONS

In this work, data-driven stabilization using prior knowledge on controllability and stabilizability has been studied. It has been shown that data-driven stabilization using controllability as prior knowledge is equivalent to data-driven stabilization without prior knowledge. It has also been shown that if the state data satisfy a rank condition, then incorporating stabilizability as prior knowledge does not lead to weaker conditions on the data. For the case where the state data are rank deficient, however, it has been shown that data-driven stabilization incorporating stabilizability as prior knowledge requires weaker conditions on the data when compared to data-driven stabilization without prior knowledge. A somewhat curious outcome of the paper is that knowledge of stabilizability may weaken the conditions on the data, while controllability, a stronger property, does not. This is due to the fact that in the former scenario, there exists the possibility that *none* of the data-consistent systems are controllable.

A class of prior knowledge that has not been studied in this work is the one representing an upper bound on the dimension of the reachable subspace of the system. In that case, one expects that obtaining a data-driven feedback gain requires an even weaker condition. In addition, here, we only focused on noise-free data. Data-driven stabilization using such prior knowledge in the presence of noisy data is also an interesting topic that is left as future work.

REFERENCES

- [1] H. J. van Waarde, M. K. Camlibel, and H. L. Trentelman, *Data-Based Linear Systems and Control Theory*, 1st ed. Kindle Direct Publishing, 2025. [Online]. Available: <https://henkvanwaarde.github.io/dblsc>
- [2] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct and indirect data-driven control formulations via regularizations and relaxations," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 883–897, 2022.
- [3] H. J. van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel, "Data informativity: A new perspective on data-driven analysis and control," *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4753–4768, 2020.
- [4] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2019.

- [5] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. M. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.
- [6] H. J. van Waarde, M. K. Camlibel, and M. Mesbahi, "From noisy data to feedback controllers: Nonconservative design via a matrix S-lemma," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 162–175, 2020.
- [7] L. Li, A. Bisoffi, C. De Persis, and N. Monshizadeh, "Controller synthesis from noisy-input noisy-output data," *Automatica*, vol. 183, p. 112545, 2026.
- [8] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.
- [9] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shadows of the DeePC," in *European Control Conference (ECC)*, 2019, pp. 307–312.
- [10] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1702–1717, 2020.
- [11] T. J. Meijer, S. A. N. Nouwens, K. J. A. Scheres, V. S. Dolk, and W. P. M. H. Heemels, "Frequency-domain data-driven predictive control," in *IFAC Conference on Nonlinear Model Predictive Control (NMPC)*, 2024, pp. 86–91.
- [12] J. Berberich, C. W. Scherer, and F. Allgöwer, "Combining prior knowledge and data for robust controller design," *IEEE Transactions on Automatic Control*, vol. 68, no. 8, pp. 4618–4633, 2022.
- [13] W. L. Qi, K. Z. Liu, R. Wang, and C. Y. Xu, "Data-driven tracking control for uncertain linear systems using a dual-system approach," *IEEE Control Systems Letters*, vol. 7, pp. 3331–3336, 2023.
- [14] N. Niknejad, F. A. Yaghmaie, and H. Modares, "Online learning of stabilizing controllers using noisy input-output data and prior knowledge," *IEEE Open Journal of Control Systems*, vol. 4, pp. 156–171, 2025.
- [15] H. Huang, M. K. Camlibel, R. Carloni, and H. J. van Waarde, "Data-driven stabilization of polynomial systems using density functions," *arXiv preprint arXiv:2503.07092*, 2025.
- [16] T. van Gestel, J. A. Suykens, P. van Dooren, and B. De Moor, "Identification of stable models in subspace identification by using regularization," *IEEE Transactions on Automatic Control*, vol. 46, no. 9, pp. 1416–1420, 2002.
- [17] S. L. Lacy and D. S. Bernstein, "Subspace identification with guaranteed stability using constrained optimization," *IEEE Transactions on Automatic Control*, vol. 48, no. 7, pp. 1259–1263, 2003.
- [18] D. N. Miller and R. A. De Callafon, "Subspace identification with eigenvalue constraints," *Automatica*, vol. 49, no. 8, pp. 2468–2473, 2013.
- [19] A. De Santis and L. Farina, "Identification of positive linear systems with poisson output transformation," *Automatica*, vol. 38, no. 5, pp. 861–868, 2002.
- [20] I. Goethals, T. van Gestel, J. Suykens, P. van Dooren, and B. De Moor, "Identification of positive real models in subspace identification by using regularization," *IEEE Transactions on Automatic Control*, vol. 48, no. 10, pp. 1843–1847, 2003.
- [21] B. M. Shali and H. J. van Waarde, "Towards a representer theorem for identification of passive systems," in *Conference on Decision and Control (CDC)*. IEEE, 2024, pp. 8760–8765.
- [22] M. Inoue, "Subspace identification with moment matching," *Automatica*, vol. 99, pp. 22–32, 2019.
- [23] M. Khosravi and R. S. Smith, "Kernel-based identification with frequency domain side-information," *Automatica*, vol. 150, p. 110813, 2023.
- [24] —, "Kernel-based impulse response identification with side-information on steady-state gain," *IEEE Transactions on Automatic Control*, vol. 68, no. 10, pp. 6401–6408, 2023.
- [25] J. Jia, H. J. van Waarde, H. L. Trentelman, and M. K. Camlibel, "A unifying framework for strong structural controllability," *IEEE Transactions on Automatic Control*, vol. 66, no. 1, pp. 391–398, 2020.
- [26] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.
- [27] H. J. van Waarde and M. K. Camlibel, "A matrix Finsler's lemma with applications to data-driven control," in *Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 5777–5782.
- [28] J. Löfberg, "Yalmip: A toolbox for modeling and optimization in matlab," in *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.
- [29] MOSEK ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 11.0*, 2025. [Online]. Available: <https://docs.mosek.com/11.0/toolbox/index.html>