**Noname manuscript No.**
(will be inserted by the editor)

# The optimal cost of a singular LQR problem, and fast/slow subspaces of the Hamiltonian system

**Chayan Bhawal · Imrul Qais ·**
**Debasattam Pal · Jan Heiland**

the date of receipt and acceptance should be inserted later

**Abstract** Weakly unobservable and strongly reachable subspaces play a crucial role in singular linear quadratic regulator (LQR) problems. Existing works in the literature use these notions on the system to compute optimal (open loop) solutions of a singular LQR problem. In this paper, we show that these notions applied to the corresponding Hamiltonian system, instead of the plant, provides new insights on the solutions of the singular LQR control problem. We identify the weakly unobservable and strongly reachable subspaces of the (differential-algebraic) Hamiltonian system and characterize these subspaces using a related Rosenbrock system matrix. As a result of this characterization, we derive an algorithm to compute the rank minimizing solutions of an LQR LMI corresponding to a single-input singular LQR problem. This algorithm is a generalization of the classical Hamiltonian matrix approach of solving regular LQR problems. Finally we show that this new approach of solving the singular LQR LMI leads to optimal solutions to singular LQR problems in proportional-derivative feedback form.

**Keywords** Hamiltonian systems, LMI, Rosenbrock system matrix, Weakly unobservable and strongly reachable subspaces.

D. Pal · I. Qais
Department of Electrical Engineering, Indian Institute of Technology (IIT) Bombay, India,
E-mail: debasattam@ee.iitb.ac.in, imrul@ee.iitb.ac.in

C. Bhawal · J. Heiland
Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany,
E-mail: bhawal@mpi-magdeburg.mpg.de, heiland@mpi-magdeburg.mpg.de

## 1 Introduction

In this paper, we deal with the linear matrix inequality (LMI) that arises in the well-known infinite-horizon linear quadratic regulator (LQR) problem.

**Problem 1.1 (Infinite-horizon LQR problem)** *Consider a controllable system $\Sigma$ with state-space dynamics $\frac{d}{dt}x = Ax + Bu$, where $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$, $B \in \mathbb{R}^{\mathtt{n}\times\mathtt{m}}$. Then, for every initial condition $x_0$, find an input $u$ that minimizes the functional*

$$J(x_0, u) := \int_0^{\infty} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt, \tag{1.1}$$

*with $\lim_{t\to\infty} x(t) = 0$, where $\left[\begin{smallmatrix} Q & S \\ S^T & R \end{smallmatrix}\right] \geqslant 0$ and $R \geqslant 0$.*

It is known that for regular LQR problems, i.e., LQR problems with $R > 0$, the input $u$ that minimizes $J(x_0, u)$ in equation (1.1) can be obtained using a static state-feedback constructed using the maximal solution of the algebraic Riccati equation (ARE):

$$A^T K + KA + Q + (KB + S)R^{-1}(B^T K + S^T) = 0. \tag{1.2}$$

However, for *singular* LQR problems, i.e., LQR problems with $R$ singular, the ARE does not exist due to non-invertibility of $R$. All LQR problems, irrespective of being regular or singular, admit LMIs of the form:

$$\begin{bmatrix} A^T K + KA + Q & KB + S \\ B^T K + S^T & R \end{bmatrix} \geqslant 0. \tag{1.3}$$

We call inequality (1.3) the *LQR LMI*. Interestingly, it has been established in [19] that for any LQR problem, the optimal cost is given by $x_0^T K_{\mathtt{max}} x_0$, where $K_{\mathtt{max}}$ is the maximal rank-minimizing solution of the LQR LMI (1.3). Hence, in order to compute the optimal cost of an LQR problem, it is imperative that the maximal rank-minimizing solution of the LQR LMI (1.3) be computed. For a regular LQR problem, the maximal rank-minimizing solution of the LQR LMI is given by the maximal solution of the corresponding ARE. There are numerous methods to compute the maximal solution of an ARE: see [1], [4] for different methods. However, these methods cannot be used to compute the maximal rank-minimizing solution of an LQR LMI for the singular case primarily due to the singularity of $R$ matrix. In this paper, we show that for single-input systems, one of the methods to compute the maximal rank-minimizing solution of an LQR LMI for the regular case (Proposition 2.5) can be extended to the singular case. This method, for the regular case, is based on computing a suitable eigenspace of the corresponding *Hamiltonian system* [12, Chapter 5]. A direct extension of this method to the singular case fails, since the dimension of the suitable eigenspace of the Hamiltonian system in such a case is less than what is required to compute the maximal rank-minimizing solution of the LQR LMI (see Example 2.1). We show in this paper that the Hamiltonian system based method for the regular case can indeed be extended to the singular case by substituting the role of the eigenspace of the Hamiltonian system in the regular case by the subspaces namely *weakly unobservable (slow)* and *strongly reachable (fast) subspaces* of the Hamiltonian system.

The idea of weakly unobservable and strongly reachable subspaces have been known to be crucial in singular LQR problems (see [10], [11], [20], [21]). In these works, the weakly unobservable and strongly reachable subspaces of a system, on which the singular LQR problem is posed, have been characterized. Recursive algorithms, to compute such subspaces for a system, have also been provided in these works. We, however, apply these notions not to the system itself, but to the corresponding Hamiltonian system that one may obtain directly by applying Pontryagin's maximum principle (PMP) to the problem (notwithstanding the fact that the impulsive nature of the optimal control for singular problems makes application of PMP inappropriate). The singularity of $R$ (and hence of the LQR problem) manifests itself in causing the Hamiltonian system to be given by a system of differential algebraic equations (DAEs), as opposed to a system of differential equations in state-space form for the regular case. The DAEs of the Hamiltonian system naturally give rise to its weakly unobservable and strongly reachable subspaces. These subspaces ultimately lead us to an algorithm to construct maximal rank-minimizing solution of the LQR LMI for a single-input system (Theorem 4.1).

In order to arrive at this algorithm, we first use the recursive algorithms to characterize these special subspaces in terms of a suitable matrix pencil known as the *Rosenbrock system matrix* (see Definition 2.2). These are the first two main results of this paper; we develop them in Section 3 (Theorem 3.1 and Theorem 3.2). The primary take away from the results in Section 3 is the relation between the relative degree of the transfer function of a system and the dimensions of its weakly unobservable and strongly reachable subspaces. We make use of this relation and the fact that for autonomous systems the weakly unobservable and strongly reachable subspaces are the direct summands of the state-space to develop an algorithm to compute the maximal rank-minimizing solution of the LQR LMI for the singular case. This is the third main result of this paper (Theorem 4.1), which we present in Section 4. Another important result crucial to the derivation of Theorem 4.1 is the disconjugacy property of a certain eigenspace of a suitable matrix pencil called the Hamiltonian matrix pencil. This is the fourth main result of this paper (Theorem 4.2) presented in Section 4.

Application of the notion of slow and fast subspaces to the Hamiltonian system not only leads to a method to compute the maximal rank-minimizing solution of the LQR LMI, but also corroborates some of the findings in the literature (see Corollary 4.3). Hence, the primary contribution of this paper is the fact that, unlike the conventional approach ([10], [20], [21]) of applying the notion of slow and fast subspaces to the system, application of these notions to the Hamiltonian system brings out further insight into the singular optimal problem. Such an approach results in a methodology for designing feedback controllers that solve the singular LQR problem: see [2]. In order to make this paper self-contained, we present this result of obtaining such feedback controllers in the form of an algorithm (Algorithm 5.1).

Another school of thought in the theory of singular LQR problems is based on the notion of deflating subspaces; see [15], [17]. Numerical algorithms for the computation of these deflating subspaces have been developed in [5], [16], etc. The fundamentally different setup of our analysis does not allow for a general comparison of our results with the results in [15] and [17]. Our results carry forward the classical ideas developed by Silverman, Hautus, Willems, and Kitapçi in [10], [20], and [21]. However, in Section 5.2, we provide a number of examples that illustrate that our approach not only advances the theory developed in [15] and [17] but also leads to practical improvements in cases where the theory of [15] and [17] applies.

## 2 Notation and Preliminaries

### 2.1 Notation

The symbols $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{N}$ are used for the sets of real numbers, complex numbers, and natural numbers, respectively. We use the symbol $\mathbb{R}_+$ and $\mathbb{C}_-$ for the set of positive real numbers and the set of complex numbers with negative real parts, respectively. The symbol $\mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ denotes the set of $\mathtt{n}\times\mathtt{p}$ matrices with elements from $\mathbb{R}$. We use $\bullet$ when a dimension need not be specified: for example, $\mathbb{R}^{\mathtt{w}\times\bullet}$ denotes the set of real constant matrices having $\mathtt{w}$ rows. We use the symbol $I_{\mathtt{n}}$ for an $\mathtt{n}\times\mathtt{n}$ identity matrix and the symbol $0_{\mathtt{n},\mathtt{m}}$ for an $\mathtt{n}\times\mathtt{m}$ matrix with all entries zero. The symbol $\{0\}$ is used to denote the zero subspace. Symbol $\mathtt{col}(B_1, B_2,\ldots,B_{\mathtt{n}})$ represents a matrix of the form $[B_1^T\ B_2^T\ \cdots\ B_{\mathtt{n}}^T]^T$. The symbol $\mathtt{det}(A)$ represents the determinant of a square matrix $A$. Symbol $\mathtt{rank}\,A$ denotes the rank of a matrix $A$. We use the symbol $\mathtt{roots}(p(s))$ to denote the set of roots (over complex numbers) of a polynomial $p(s)$ with real or complex coefficients (counted with multiplicity). Symbol $\mathtt{num}(p(s))$ is used to denote the numerator of a rational function $p(s)$. The symbol $\sigma(\Gamma)$ denotes the set of eigenvalues of a square matrix $\Gamma$ (counted with multiplicity). The symbol $|\Gamma|$ denotes the cardinality of a set $\Gamma$ (counted with multiplicity). We use the symbol $\sigma(A|_{\mathcal{S}})$ to represent the set of eigenvalues of $A$ restricted to a space $\mathcal{S}$. We use the symbol $\mathtt{dim}\,(\mathcal{S})$ to denote the dimension of a space $\mathcal{S}$. The symbol $\mathtt{img}\,A$ and $\mathtt{ker}\,A$ denote the image and nullspace of a matrix $A$, respectively. The space of all infinitely differentiable functions and locally square-integrable functions from $\mathbb{R}$ to $\mathbb{R}^{\mathtt{n}}$ are represented by the symbol $\mathfrak{C}^\infty(\mathbb{R},\mathbb{R}^{\mathtt{n}})$ and $\mathcal{L}^2_{\mathtt{loc}}(\mathbb{R},\mathbb{R}^{\mathtt{n}})$, respectively. We use the symbol $\mathfrak{C}^\infty(\mathbb{R},\mathbb{R}^{\mathtt{n}})|_{\mathbb{R}_+}$

to represent the set of all functions from $\mathbb{R}_+$ to $\mathbb{R}^{\mathtt{n}}$ that are restrictions of $\mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{n}})$ functions to $\mathbb{R}_+$. The symbol $\delta$ represents the Dirac delta impulse function and $\delta^{(i)}$ represents the $i$-th distributional derivative of $\delta$ with respect to $t$.

## 2.2 Regular matrix pencils and their canonical form

Linear matrix pencils and their eigenvectors are crucially used throughout this paper. Hence, we define eigenvalues and eigenvectors corresponding to linear matrix pencils next.

**Definition 2.1** *Consider a regular matrix pencil* $(sU_1 - U_2) \in \mathbb{R}[s]^{\mathtt{n}\times\mathtt{n}}$*, i.e.,* $\mathtt{det}(sU_1 - U_2) \neq 0$*. Let* $\lambda \in \mathtt{roots}\,(\mathtt{det}(sU_1 - U_2))$*. Then* $\lambda$ *is called an* eigenvalue *of* $(U_1, U_2)$ *and every nonzero vector* $v \in \mathtt{ker}\,(\lambda U_1 - U_2)$ *is called an* eigenvector *of the matrix pair* $(U_1, U_2)$ *corresponding to the eigenvalue* $\lambda$*. Further, every nonzero vector* $\tilde{v} \in \mathtt{ker}\,(\lambda U_1 - U_2)^i$*, where* $i \in \{2, 3, \ldots\}$*, is called a* generalized eigenvector *of the matrix pair* $(U_1, U_2)$ *corresponding to the eigenvalue* $\lambda$*.*

We use the symbol $\sigma(U_1, U_2)$ to denote the set of eigenvalues of $(U_1, U_2)$ (with $\lambda \in \sigma(U_1, U_2)$ included in the set as many times as its algebraic multiplicity).

In this paper, we extensively use one of the canonical forms of a linear matrix pencil (see [6] for more on different canonical forms). We review the result that leads to such a canonical form next [6, Lemma 1-2.2].

**Proposition 2.1** *A matrix pair* $(U_1, U_2)$ *is regular, i.e.,* $\mathtt{det}(sU_1 - U_2) \neq 0$ *if and only if there exist nonsingular matrices* $Z_1$ *and* $Z_2$ *such that* $Z_1 U_1 Z_2 = \mathtt{diag}(I_{\mathtt{n}_1}, Y)$ *and* $Z_1 U_2 Z_2 = \mathtt{diag}(U, I_{\mathtt{n}_2})$*, where* $\mathtt{n}_1 + \mathtt{n}_2 = \mathtt{n}, U \in \mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_1}$*, and* $Y \in \mathbb{R}^{\mathtt{n}_2\times\mathtt{n}_2}$ *is nilpotent.*

A matrix pair $(U_1, U_2)$ in the form $\left(\begin{bmatrix} I_{\mathtt{n}_1} & \\ & Y \end{bmatrix}, \begin{bmatrix} U & \\ & I_{\mathtt{n}_2} \end{bmatrix}\right)$ is said to be in a canonical form. Further, note that $\mathtt{det}(sU_1 - U_2) = k \times \mathtt{det}(sI_{\mathtt{n}_1} - U)$, where $k \in \mathbb{R} \setminus \{0\}$.

In the sequel, we also use the notion of Rosenbrock system matrix. We define this next [18].

**Definition 2.2** *Consider a system with an input-state-output (i/s/o) representation of the form* $\frac{d}{dt}x = Ax + Bu$ *and* $y = Cx + Du$*. Then, the matrix* $\begin{bmatrix} sI-A & -B \\ -C & -D \end{bmatrix}$ *is called the Rosenbrock system matrix.*

Note that the Rosenbrock system matrix can also be written as $s\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A & B \\ C & D \end{bmatrix}$. For the ease of exposition, we call the matrix pair $(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} A & B \\ C & D \end{bmatrix})$ the *Rosenbrock matrix pair.*

### 2.3 $(A, B)$-invariant subspace and controllability subspace

The notions of $(A, B)$-invariant subspace and controllability subspace are essential for this paper. We briefly review these notions next (see [22, Chapters 4 and 5] for more on these subspaces).

**Definition 2.3** *Consider $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$ and $B \in \mathbb{R}^{\mathtt{n}\times\mathtt{m}}$. A subspace $\mathcal{S} \subseteq \mathbb{R}^{\mathtt{n}}$ is said to be* $(A, B)$-invariant *if there exists a matrix $F \in \mathbb{R}^{\mathtt{m}\times\mathtt{n}}$ such that $(A+BF)\mathcal{S} \subseteq \mathcal{S}$.*

Following the notation in [22], we use the symbol $\mathfrak{I}(A, B)$ for the family of $(A, B)$-invariant subspaces. The notation $\mathbf{F}(\mathcal{S})$ is used for the collection of matrices $F \in \mathbb{R}^{\mathtt{m}\times\mathtt{n}}$ such that $(A + BF)\mathcal{S} \subseteq \mathcal{S}$. The next proposition provides a test for determining whether a given subspace is $(A, B)$-invariant [22, Lemma 4.2]. We use this test throughout this paper.

**Proposition 2.2** *A subspace $\mathcal{S} \subseteq \mathbb{R}^{\mathtt{n}}$ is $(A, B)$-invariant if and only if $A\mathcal{S} \subseteq \mathcal{S} + \mathtt{img}\, B$.*

The notation $\mathfrak{I}(A, B; \mathtt{ker}\, C)$ denotes the family of $(A, B)$-invariant subspaces that are contained in $\mathtt{ker}\, C$, where $C \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$. Importantly, it is known in the literature that the set $\mathfrak{I}(A, B; \mathtt{ker}\, C)$ admits a supremal element [22, Lemma 4.4], and we represent this supremal element by the symbol $\mathtt{sup}\, \mathfrak{I}(A, B; \mathtt{ker}\, C)$. Formally this means that for all $\mathcal{S} \in \mathfrak{I}(A, B; \mathtt{ker}\, C)$, we must have $\mathcal{S} \subseteq \mathtt{sup}\, \mathfrak{I}(A, B; \mathtt{ker}\, C)$. Such a subspace is of importance to us in this paper.

**Definition 2.4** *Consider $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$ and $B \in \mathbb{R}^{\mathtt{n}\times\mathtt{m}}$. A subspace $\mathscr{R} \subseteq \mathbb{R}^{\mathtt{n}}$ is a* controllability subspace *of the pair $(A, B)$ if there exist $F \in \mathbb{R}^{\mathtt{m}\times\mathtt{n}}$ and $G \in \mathbb{R}^{\mathtt{m}\times\mathtt{m}}$, such that $\mathscr{R}$ is the reachable subspace of the pair $(A + BF, BG)$, i.e.*

$$\mathscr{R} = \mathtt{img}\, \begin{bmatrix} BG & (A + BF)BG & (A + BF)^2 BG & \cdots & (A + BF)^{\mathtt{n}-1} BG \end{bmatrix}.$$

We use the symbol $\mathfrak{C}(A, B)$ for the family of controllability subspaces of $(A, B)$. The notation $\mathfrak{C}(A, B; \mathtt{ker}\, C)$ denotes the family of controllability subspaces that are contained in $\mathtt{ker}\, C$. Similar to $\mathfrak{I}(A, B; \mathtt{ker}\, C)$, the set $\mathfrak{C}(A, B; \mathtt{ker}\, C)$ also admits a supremal element [22, Theorem 5.4]. We represent this element as $\mathtt{sup}\, \mathfrak{C}(A, B; \mathtt{ker}\, C)$.

Define

$$\mathcal{B} := \{\mathcal{S} \in \mathfrak{I}(A, B, \mathtt{ker}\, C) \mid \exists\, F \in \mathbf{F}(\mathcal{S}) \text{ such that } \sigma\left((A + BF)|_{\mathcal{S}}\right) \subsetneq \mathbb{C}_-\}.$$

We call any subspace in $\mathcal{B}$ a *good $(A, B)$-invariant subspace inside* $\mathtt{ker}\, C$. As shown in [22, Lemma 5.8], the set $\mathcal{B}$ admits a supremal element defined as $\mathcal{S}^*_{\mathtt{g}} := \mathtt{sup}\, \mathcal{B}$, i.e., for all elements $\mathcal{S} \in \mathcal{B}, \mathcal{S} \subseteq \mathcal{S}^*_{\mathtt{g}}$. Hence, $\mathcal{S}^*_{\mathtt{g}}$ is called the *largest good $(A, B)$-invariant subspace inside* $\mathtt{ker}\, C$.

Let $\mathcal{S}^* := \sup \mathfrak{I}(A, B; \ker C)$ and $\mathscr{R}^* := \sup \mathfrak{C}(A, B; \ker C)$. Further, let $F \in \mathbf{F}(\mathcal{S}^*)$. Clearly, $\mathscr{R}^* \subseteq \mathcal{S}^*$, hence $\mathcal{S}^*$ admits factoring by the subspace $\mathscr{R}^*$. Let $\overline{(A+BF)|_{\mathcal{S}^*}}$ denote the map induced by $(A+BF)|_{\mathcal{S}^*}$ on the factor space $\mathcal{S}^*/\mathscr{R}^*$. Then, it is known that the set of eigenvalues $\sigma\left(\overline{(A+BF)|_{\mathcal{S}^*}}\right)$ remains invariant for all $F \in \mathbf{F}(\mathcal{S}^*)$. For a system with an i/s/o representation $\frac{d}{dt}x = Ax + Bu$ and $y = Cx$, the complex numbers $\sigma\left(\overline{(A+BF)|_{\mathcal{S}^*}}\right)$ are known as the *transmission zeros* of the system. Note importantly that, for a single-input controllable system, we have $\mathscr{R}^* = \{0\}$. Consequently, $\mathcal{S}^*/\mathscr{R}^* = \mathcal{S}^*$, and $\overline{(A+BF)|_{\mathcal{S}^*}} = (A+BF)|_{\mathcal{S}^*}$. This means that for single-input systems, $\sigma((A+BF)|_{\mathcal{S}^*})$ is the set of the transmission zeros. In other words, the set $\sigma((A+BF)|_{\mathcal{S}^*})$ remains invariant for all $F \in \mathbf{F}(\mathcal{S}^*)$. Further, it can also be shown that for a controllable and observable SISO system, $\sigma((A+BF)|_{\mathcal{S}^*}) = \texttt{roots}\,\texttt{num}(G(s))$, where $G(s) = C(sI_{\mathtt{n}} - A)^{-1}B$ (see [22, Section 5.5]). This property of single-input systems is essential for the development of the theory in Section 3 and Section 4.

## 2.4 Weakly unobservable and strongly reachable subspaces

Consider the system $\Sigma$ with an i/s/o representation $\frac{d}{dt}x = Ax + Bu$ and $y = Cx$, where $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}, B \in \mathbb{R}^{\mathtt{n}\times\mathtt{m}}$ and $C \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$. Associated with such a system are two important subspaces called the weakly unobservable subspace and the strongly reachable subspace. We briefly review the properties of these subspaces next (see [10] for more on these spaces). Before we delve into the definitions of these subspaces, we need to define the space of impulsive-smooth distributions (see [10], [21]).

**Definition 2.5** *The set of impulsive-smooth distributions $\mathfrak{C}^{\mathtt{w}}_{\rm imp}$ is defined as:*

$$\mathfrak{C}^{\mathtt{w}}_{\rm imp} := \Big\{ f = f_{\mathtt{reg}} + f_{\mathtt{imp}} \mid f_{\mathtt{reg}} \in \mathfrak{C}^{\infty}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})|_{\mathbb{R}_+} \text{ and } f_{\mathtt{imp}} = \sum_{i=0}^{k} a_i \delta^{(i)}, \text{ with } a_i \in \mathbb{R}^{\mathtt{w}}, k \in \mathbb{N} \Big\}.$$

In what follows, we denote the state-trajectory $x$ and output-trajectory $y$ of the system $\Sigma$, that result from initial condition $x_0$ and input $u$, using the symbols $x(x_0, u)$ and $y(x_0, u)$, respectively. The symbol $x(0^+; x_0, u)$ denotes the value of the state-trajectory that can be reached from $x_0$ instantaneously on application of the input $u$ at $t = 0$.

**Definition 2.6** *A state $x_0 \in \mathbb{R}^{\mathtt{n}}$ is called* weakly unobservable *if there exists an input $u \in \mathfrak{C}^{\infty}(\mathbb{R}, \mathbb{R}^{\mathtt{m}})|_{\mathbb{R}_+}$ such that $y(t; x_0, u) \equiv 0$ for all $t \geqslant 0$. The collection of all such weakly unobservable states is called the* weakly unobservable subspace *of the state-space and is denoted by $\mathcal{O}_w$.*

The following property of the weakly unobservable subspace is crucially used in this paper (see [10, Theorem 3.10]).

**Proposition 2.3** *The weakly unobservable subspace $\mathcal{O}_w$ is the largest $(A,B)$-invariant subspace inside the kernel of $C$, i.e., $\mathcal{O}_w = \sup \mathfrak{I}(A,B;\ker C)$.*

The other space that we are interested in, is the space of strongly reachable states (see [10]).

**Definition 2.7** *A state $x_1 \in \mathbb{R}^{\mathtt{n}}$ is called* strongly reachable *(from the origin) if there exists an input $u \in \mathfrak{C}^{\mathtt{m}}_{\text{imp}}$ such that $x(0^+;0,u) = x_1$ and $y(0,u) \in \mathfrak{C}^{\infty}(\mathbb{R},\mathbb{R}^{\mathtt{p}})|_{\mathbb{R}_+}$. The collection of all such strongly reachable states is called the* strongly reachable subspace *of the state-space and is denoted by $\mathcal{R}_s$.*

A method to compute the space $\mathcal{R}_s$ is given by the following recursion

$$\mathcal{R}_0 := \{0\} \subsetneq \mathbb{R}^{\mathtt{n}}, \text{ and } \mathcal{R}_{i+1} := \begin{bmatrix} A & B \end{bmatrix} \left\{ (\mathscr{W}_i \oplus \mathscr{P}) \cap \ker \begin{bmatrix} C & 0_{\mathtt{p},\mathtt{m}} \end{bmatrix} \right\} \subseteq \mathcal{R}_s, \tag{2.1}$$

where $\mathscr{W}_i := \{ \left[\begin{smallmatrix} w \\ 0 \end{smallmatrix}\right] \in \mathbb{R}^{\mathtt{n}+\mathtt{m}} \mid w \in \mathcal{R}_i \}$ and $\mathscr{P} := \{ \left[\begin{smallmatrix} 0 \\ \alpha \end{smallmatrix}\right] \in \mathbb{R}^{\mathtt{n}+\mathtt{m}} \mid \alpha \in \mathbb{R}^{\mathtt{m}} \}$. In Section 3.1 we use this recursive algorithm to characterize the strongly reachable subspace of a single-input system in terms of the Rosenbrock system matrix (see [10] for more on the recursive algorithm).

Since the subspace $\mathcal{O}_w$ deals with inputs from the space of infinitely differentiable functions, we call $\mathcal{O}_w$ the *slow subspace* of the system. Further, note that since $\mathcal{O}_w$ is the largest $(A,B)$-invariant subspace inside the kernel of $C$, such a subspace also admits largest good $(A,B)$-invariant subspace inside the kernel of $C$. We call such a space the *good slow subspace* of the system. On the other hand, since the space $\mathcal{R}_s$ admits impulsive inputs, we call $\mathcal{R}_s$ the *fast subspace* of the system.

In the sequel, we use the notion of autonomy of a system and its relation with the spaces $\mathcal{O}_w$ and $\mathcal{R}_s$. Hence, we define autonomy of a system first and then review the result [11, Lemma 3.3] that establishes a noteworthy property of $\mathcal{O}_w$ and $\mathcal{R}_s$ for autonomous systems.

**Definition 2.8** *A system with an output-nulling representation of the form $\frac{d}{dt}x = Ax + Bu$ and $0 = Cx$ is called* autonomous *if for every initial condition $x_0 \in \mathcal{O}_w$ the system has a unique solution $(x,u)$.*

**Proposition 2.4** *Consider the system $\frac{d}{dt}x = Ax + Bu$ and $0 = Cx$. Then the following are equivalent:*

1. *The system is autonomous.*
2. *$G(s) := C(sI_{\mathtt{n}} - A)^{-1}B$ is invertible as a rational matrix.*
3. *$\mathcal{O}_w \oplus \mathcal{R}_s = \mathbb{R}^{\mathtt{n}}$ and $\ker \left[\begin{smallmatrix} B \\ 0_{\mathtt{m},\mathtt{m}} \end{smallmatrix}\right] = \{0\}$.*

Throughout the paper, we consider the matrix $B$ to be of full column-rank without loss of generality. Thus the second condition of Statement *(3)* is assumed to hold true throughout the paper unless stated otherwise.

### 2.5 ARE and Hamiltonian systems

One of the widely used methods to compute the maximal solution of the ARE (1.2) is to use the basis of a suitable eigenspace of the matrix pair $(E, H)$, where

$$E := \begin{bmatrix} I_{\mathtt{n}} & 0 & 0 \\ 0 & I_{\mathtt{n}} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ and } H := \begin{bmatrix} A & 0 & B \\ -Q & -A^T & -S \\ S^T & B^T & R \end{bmatrix}. \tag{2.2}$$

We call the matrix pair $(E, H)$ the *Hamiltonian matrix pair* and the matrix pencil $(sE - H)$ the *Hamiltonian pencil*. The suitable eigenspace used to compute the maximal solution of the ARE (1.2) corresponds to certain choice of eigenvalues of $(E, H)$. In order to understand this choice of eigenvalues the notion of Lambda-sets is essential and hence we define Lambda-sets next ([13]).

**Definition 2.9** *Let $p(s)$ be an even-degree polynomial with $\mathtt{roots}\,(p(s)) \cap j\mathbb{R} = \emptyset$. A set of complex numbers $\Lambda \subsetneq \mathtt{roots}\,(p(s))$ is called a Lambda-set of $p(s)$ if it satisfies the following properties:*

1. $\Lambda = \bar{\Lambda}$.
2. $\Lambda \cap (-\Lambda) = \emptyset$.
3. $\Lambda \cup (-\Lambda) = \mathtt{roots}\,(p(s))$ *(counted with multiplicity).*

Now that we have the definition for Lambda-sets, we review the method to compute the maximal solution of the ARE (1.2) (see [12] for more). Recall that the maximal solution of an ARE is the maximal rank-minimizing solution of the corresponding LMI (1.3).

**Proposition 2.5** *Consider the LQR Problem 1.1 with $R > 0$. Let the corresponding Hamiltonian matrix pair $(E, H)$ be as defined in equation (2.2). Assume $\sigma(E, H) \cap j\mathbb{R} = \emptyset$. Let $\Lambda$ be a Lambda-set of $\mathtt{det}(sE - H)$ with cardinality $\mathtt{n}$ and $\Lambda \subsetneq \mathbb{C}_-$. Let $V_{1\Lambda}, V_{2\Lambda} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$ and $V_{3\Lambda} \in \mathbb{R}^{\mathtt{m}\times\mathtt{n}}$ be such that the columns of $V_{\mathtt{e}\Lambda} = \mathtt{col}(V_{1\Lambda}, V_{2\Lambda}, V_{3\Lambda})$ form a basis of the $\mathtt{n}$-dimensional eigenspace of $(E, H)$ corresponding to the eigenvalues of $(E, H)$ in $\Lambda$. Then, the following statements hold:*

*(1)* $V_{1\Lambda}$ *is invertible.*

*(2)* $K_{\mathtt{max}} := V_{2\Lambda} V_{1\Lambda}^{-1}$ *is symmetric.*

*(3)* $K_{\mathtt{max}}$ *is the maximal solution of the ARE* (1.2).

(4) *$K_{\mathtt{max}}$ is the maximal rank-minimizing solution of the corresponding LQR LMI* (1.3).

(5) $K_{\mathtt{max}} \geqslant 0$.

Although Proposition 2.5 does not explicitly use invertibility of $R$ while finding the maximal rank-minimizing solution of the LQR LMI, yet the proposition cannot be used to compute such a solution for the LQR LMI corresponding to a singular LQR problem. We motivate the reason for this using an example.

*Example 2.1* Consider a system with state-space dynamics

$$\frac{d}{dt}x = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} u.$$

For every initial condition $x_0$, find an input $u$ that minimizes the functional

$$\int_0^\infty x(t)^T Q x(t)\ dt, \text{ where } Q := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

On construction of the Hamiltonian pencil pair $(E, H)$ using $A, B, Q$, it can be verified that $\mathtt{det}(sE - H) = 1 - s^2$. Hence, $\Lambda = \{-1\}$. The eigenvector of $(E, H)$ corresponding to $-1$, is $\left[\begin{smallmatrix} 1 & 1 & -2 & 2 & 0 & 0 & 0 \end{smallmatrix}\right]^T$. Therefore, $V_{1\Lambda} = \left[\begin{smallmatrix} 1 & 1 & -2 \end{smallmatrix}\right]^T$ and $V_{2\Lambda} = \left[\begin{smallmatrix} 2 & 0 & 0 \end{smallmatrix}\right]^T$. But $V_{1\Lambda}$ is not a square matrix. Thus, Proposition 2.5 cannot be used to solve singular LQR problems.

From Example 2.1, it is clear that Proposition 2.5 fails in case of singular LQR problems because the degree of $\mathtt{det}(sE - H)$ is strictly less than $2\mathtt{n}$. This fall in the degree causes a deficit in the cardinality of possible Lambda-sets of $\mathtt{det}(sE - H)$. Indeed, a Lambda set of $\mathtt{det}(sE - H)$ can now have cardinality only $\mathtt{n_s}$, which is strictly less than $\mathtt{n}$. Consequently, the eigenspace of $(E, H)$ corresponding to such a Lambda-set would also show a deficit in its dimension from being $\mathtt{n}$. This deficit in the dimension of the eigenspace is required to be compensated by $(\mathtt{n} - \mathtt{n_s})$ suitable vectors. Of course, this compensation cannot be done by arbitrary vectors. Our main result, Theorem 4.1, shows exactly how this compensation is done.

Since we deal with the singular LQR problem for single-input systems, we rewrite the LQR Problem 1.1 for the single-input case as follows:

**Problem 2.1 (Single-input singular LQR problem)** *Consider a controllable system $\Sigma$ with state-space dynamics $\frac{d}{dt}x = Ax + bu$, where $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$ and $b \in \mathbb{R}^{\mathtt{n}\times 1}$. Then, for every initial condition $x_0$, find an input $u$ that minimizes the functional*

$$J(x_0, u) := \int_0^\infty x(t)^T Q x(t)\ dt, \text{ where } Q \geqslant 0. \tag{2.3}$$

Note that the LQR LMI (1.3) with respect to Problem 2.1 takes the following form:

$$\begin{bmatrix} A^T K + KA + Q & Kb \\ b^T K & 0 \end{bmatrix} \geqslant 0 \ \Leftrightarrow \ \begin{cases} A^T K + KA + Q \geqslant 0, \\ Kb = 0. \end{cases} \tag{2.4}$$

Further for a single-input singular LQR problem as defined in Problem 2.1, the Hamiltonian matrix pair in equation (2.2) takes the following form:

$$E := \begin{bmatrix} I_{\mathtt{n}} & 0 & 0 \\ 0 & I_{\mathtt{n}} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ and } H := \begin{bmatrix} A & 0 & b \\ -Q & -A^T & 0 \\ 0 & b^T & 0 \end{bmatrix}. \tag{2.5}$$

Interestingly, the Hamiltonian matrix pencil $(E, H)$ in equation (2.5) can be associated with a differential algebraic system as given below:

$$\underbrace{\begin{bmatrix} I_{\mathtt{n}} & 0 & 0 \\ 0 & I_{\mathtt{n}} & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{E} \frac{d}{dt} \begin{bmatrix} x \\ z \\ u \end{bmatrix} = \underbrace{\begin{bmatrix} A & 0 & b \\ -Q & -A^T & 0 \\ 0 & b^T & 0 \end{bmatrix}}_{H} \begin{bmatrix} x \\ z \\ u \end{bmatrix}. \tag{2.6}$$

The system represented by this first order representation (2.6) is called the *Hamiltonian system*; we use $\Sigma_{\text{Ham}}$ to denote this system (see [12] for more on Hamiltonian systems).

Another representation of the Hamiltonian system (2.6) extensively used in this paper is the *output-nulling representation* as given below:

$$\frac{d}{dt} \begin{bmatrix} x \\ z \end{bmatrix} = \widehat{A} \begin{bmatrix} x \\ z \end{bmatrix} + \widehat{b} u, \quad 0 = \widehat{c} \begin{bmatrix} x \\ z \end{bmatrix}, \tag{2.7}$$

where $\widehat{A} := \begin{bmatrix} A & 0 \\ -Q & -A^T \end{bmatrix}$, $\widehat{b} := \begin{bmatrix} b \\ 0 \end{bmatrix}$ and $\widehat{c} := \begin{bmatrix} 0 & b^T \end{bmatrix}$. Note that the Hamiltonian matrix pair $(E, H)$ is indeed the Rosenbrock matrix pair for the system $\Sigma_{\text{Ham}}$. In what follows, we shall need the notion of disconjugacy related to the Hamiltonian matrix pair [12, Definition 6.1.5]. We review this next.

**Definition 2.10** *Let $\mathcal{E}$ be an eigenspace of $(E, H)$, where $(E, H)$ is as defined in equation (2.2). Assume the columns of a matrix $V_e$ to be the basis of $\mathcal{E}$. Conforming to the partition of $H$, let $V_e$ be partitioned as* $\mathtt{col}(V_1, V_2, V_3)$. *Then, $\mathcal{E}$ is called* disconjugate *if $V_1$ is full column-rank.*

## 3 Characterization of slow and fast subspaces in terms of Rosenbrock system matrix

Consider $\Sigma_{\mathtt{P}}$ to be a system with an output-nulling representation of the form:

$$\frac{d}{dt}x = Px + Lu, \text{ and } 0 = Mx, \text{ where } P \in \mathbb{R}^{\mathtt{N}\times\mathtt{N}}, L, M^T \in \mathbb{R}^{\mathtt{N}\times 1} \setminus \{0\}. \quad (3.1)$$

Define the matrix pair

$$U_1 := \begin{bmatrix} I_{\mathtt{N}} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{(\mathtt{N}+1)\times(\mathtt{N}+1)} \text{ and } U_2 := \begin{bmatrix} P & L \\ M & 0 \end{bmatrix} \in \mathbb{R}^{(\mathtt{N}+1)\times(\mathtt{N}+1)}. \quad (3.2)$$

Note that $(sU_1 - U_2)$ is the Rosenbrock system matrix for the system $\Sigma_{\mathtt{P}}$ in equation (3.1) and $(U_1, U_2)$ is the Rosenbrock matrix pair. In this section we characterize the slow subspace $\mathcal{O}_w$ and fast subspace $\mathcal{R}_s$ of the system $\Sigma_{\mathtt{P}}$ in terms of the matrix pair $(U_1, U_2)$. Further, we also characterize the good slow subspace of $\Sigma_{\mathtt{P}}$ in terms of the eigenspace of $(U_1, U_2)$. Hence, we have divided this section into three subsections; the first being characterization of the fast subspace of $\Sigma_{\mathtt{P}}$. In the second and third subsection we characterize the slow and good slow subspaces of $\Sigma_{\mathtt{P}}$, respectively in terms of the eigenspace of the Rosenbrock matrix pair $(U_1, U_2)$.

### 3.1 Characterization of fast subspace in terms of the Rosenbrock matrix pair

In order to characterize the fast subspace, we need certain identities related to the Markov parameters of the system $\Sigma_{\mathtt{P}}$. We present this in the next lemma and follow it up with the first main result of this section.

**Lemma 3.1** *Consider the system $\Sigma_{\mathtt{P}}$ as defined in equation* (3.1)*, and the corresponding Rosenbrock matrix pair $(U_1, U_2)$ as defined in equation* (3.2)*. Assume $\mathtt{det}(sU_1 - U_2) \neq 0$. Define $\mathtt{deg}\,\mathtt{det}(sU_1 - U_2) =: \mathtt{N_s}$ and $\mathtt{N_f} := \mathtt{N} - \mathtt{N_s}$. Then, $MP^kL = 0$, for $k \in \{0, 1, \ldots, \mathtt{N_f} - 2\}$ and $MP^{\mathtt{N_f}-1}L \neq 0$.*

*Proof:* Define $G(s) := M(sI_{\mathtt{N}} - P)^{-1}L$. Using the notion of Schur complement we have $\mathtt{det}(sU_1 - U_2) = -M(sI_{\mathtt{N}} - P)^{-1}L \times \mathtt{det}(sI_{\mathtt{N}} - P)$. Since $\mathtt{degdet}(sU_1 - U_2) =: \mathtt{N_s}$ and $\mathtt{degdet}(sI_{\mathtt{N}} - P) = \mathtt{N}$, the relative degree of $G(s)$ must be $\mathtt{N} - \mathtt{N_s} = \mathtt{N_f}$. Expanding $(sI_{\mathtt{N}} - P)^{-1}$ in a Taylor series about $s = \infty$, we have $G(s) = \frac{ML}{s} + \frac{MPL}{s^2} + \frac{MP^2L}{s^3} + \cdots$. Since the relative degree of the rational function $G(s)$ is $\mathtt{N_f}$. Hence, we can infer from the Taylor expansion of $G(s)$ that $\lim_{s\to\infty} s^{k+1}G(s) = 0 = MP^kL$ for $k \in \{0, 1, \ldots, \mathtt{N_f} - 2\}$. Further since relative degree of $G(s)$ is $\mathtt{N_f}$, $\lim_{s\to\infty} s^{\mathtt{N_f}}G(s) \neq 0$. Hence, $MP^{\mathtt{N_f}-1}L \neq 0$. □

**Theorem 3.1** *Consider the system $\Sigma_{\texttt{P}}$ as defined in equation* (3.1)*, and the corresponding Rosenbrock matrix pair* $(U_1, U_2)$ *as defined in equation* (3.2)*. Assume* $\texttt{det}(sU_1 - U_2) \neq 0$. *Define* $\texttt{degdet}(sU_1 - U_2) =: \texttt{N}_{\texttt{s}}$ *and* $\texttt{N}_{\texttt{f}} := \texttt{N} - \texttt{N}_{\texttt{s}}$. *Let $\mathcal{R}_s$ be the fast subspace of $\Sigma_{\texttt{P}}$. Then, the following statements are true:*

1. $\mathcal{R}_s = \texttt{img}\begin{bmatrix} L \; PL \cdots \; P^{\texttt{N}_{\texttt{f}}-1}L \end{bmatrix}$.
2. $\texttt{dim}(\mathcal{R}_s) = \texttt{N}_{\texttt{f}}$.

*Proof:* *(1)*: From equation (2.1) in Section 2.3, the recursive algorithm to compute the fast subspace of $\Sigma_{\texttt{P}}$ is given by:

$$\mathcal{R}_0 := \{0\} \subsetneq \mathbb{R}^{\texttt{N}}, \text{ and } \mathcal{R}_{i+1} := \begin{bmatrix} P \; L \end{bmatrix} \left\{ (\mathscr{W}_i \oplus \mathscr{P}) \cap \texttt{ker} \begin{bmatrix} M \; 0 \end{bmatrix} \right\} \subseteq \mathcal{R}_s, \quad (3.3)$$

where $\mathscr{W}_i := \left\{ \left[\begin{smallmatrix} w \\ 0 \end{smallmatrix}\right] \in \mathbb{R}^{\texttt{N}+1} \,|\, w \in \mathcal{R}_i \right\}$ and $\mathscr{P} := \left\{ \left[\begin{smallmatrix} 0 \\ \alpha \end{smallmatrix}\right] \in \mathbb{R}^{\texttt{N}+1} \,|\, \alpha \in \mathbb{R} \right\}$. Note that since $\mathscr{P} \cap \texttt{ker} \begin{bmatrix} M \; 0 \end{bmatrix} = \mathscr{P}$, the recursion in equation (3.3) can be rewritten as

$$\mathcal{R}_0 = \{0\} \subsetneq \mathbb{R}^{\texttt{N}} \text{ and } \mathcal{R}_{i+1} := \begin{bmatrix} P \; L \end{bmatrix} \left\{ (\mathscr{W}_i \cap \texttt{ker} \begin{bmatrix} M \; 0 \end{bmatrix}) \oplus \mathscr{P} \right\} \subseteq \mathcal{R}_s. \quad (3.4)$$

Now, we claim that $\mathcal{R}_k = \texttt{img}\, L + \texttt{img}\,(PL) + \cdots + \texttt{img}\,(P^{k-1}L)$ for $k \in \{1, 2, 3, \ldots, \texttt{N}_{\texttt{f}}\}$. To prove this we use the principle of mathematical induction along with Lemma 3.1.
**Base case:** ($k = 1$) Note that since $(\mathscr{W}_0 \cap \texttt{ker} \begin{bmatrix} M \; 0 \end{bmatrix}) = \{0\} \subsetneq \mathbb{R}^{\texttt{N}+1}$, we must have from equation (3.4), $\mathcal{R}_1 = \begin{bmatrix} P \; L \end{bmatrix} \{\{0\} \oplus \mathscr{P}\} = \texttt{img}\, L$.
**Induction step:** Assume $\mathcal{R}_k = \texttt{img}\, L + \texttt{img}\,(PL) + \cdots + \texttt{img}\,(P^{k-1}L)$ for $k < \texttt{N}_{\texttt{f}}$. We prove that $\mathcal{R}_{k+1} = \texttt{img}\, L + \texttt{img}\,(PL) + \cdots + \texttt{img}\,(P^kL)$.

From equation (3.4), we have

$$\begin{aligned} \mathcal{R}_{k+1} &= \begin{bmatrix} P \; L \end{bmatrix} \left\{ (\mathscr{W}_k \cap \texttt{ker} \begin{bmatrix} M \; 0 \end{bmatrix}) \oplus \mathscr{P} \right\} \\ &= \begin{bmatrix} P \; L \end{bmatrix} \left\{ \left( \left( \sum_{i=0}^{k-1} \texttt{img} \left[\begin{smallmatrix} P^iL \\ 0 \end{smallmatrix}\right] \right) \cap \texttt{ker} \begin{bmatrix} M \; 0 \end{bmatrix} \right) \oplus \mathscr{P} \right\} \\ &= \begin{bmatrix} P \; L \end{bmatrix} \left\{ \sum_{i=0}^{k-1} \left( \texttt{img} \left[\begin{smallmatrix} P^iL \\ 0 \end{smallmatrix}\right] \cap \texttt{ker} \left[\begin{smallmatrix} M & 0 \end{smallmatrix}\right] \right) \oplus \mathscr{P} \right\}. \end{aligned} \quad (3.5)$$

Since $\left[\begin{smallmatrix} M & 0 \end{smallmatrix}\right] \left[\begin{smallmatrix} P^iL \\ 0 \end{smallmatrix}\right] = MP^iL = 0$ for $i < \texttt{N}_{\texttt{f}} - 1$ (from Lemma 3.1), we must have $\sum_{i=0}^{k-1} \left( \texttt{img} \left[\begin{smallmatrix} P^iL \\ 0 \end{smallmatrix}\right] \cap \texttt{ker} \left[\begin{smallmatrix} M & 0 \end{smallmatrix}\right] \right) = \sum_{i=0}^{k-1} \left( \texttt{img} \left[\begin{smallmatrix} P^iL \\ 0 \end{smallmatrix}\right] \right)$. Thus, from equation (3.5) we have

$$\begin{aligned} \mathcal{R}_{k+1} &= \begin{bmatrix} P \; L \end{bmatrix} \left\{ \sum_{i=0}^{k-1} \left( \texttt{img} \begin{bmatrix} P^iL \\ 0 \end{bmatrix} \right) \oplus \mathscr{P} \right\} \\ &= \texttt{img}\, L + \texttt{img}\,(PL) + \cdots + \texttt{img}\,(P^kL). \end{aligned}$$

By the principle of mathematical induction, we conclude that

$$\mathcal{R}_k = \texttt{img}\, L + \texttt{img}\,(PL) + \cdots + \texttt{img}\,(P^{k-1}L) \text{ for } k \in \{1, 2, 3, \ldots, \texttt{N}_{\texttt{f}}\}. \quad (3.6)$$

This proves our claim.

Next we claim that $\mathcal{R}_{\mathtt{N_f}+1} = \mathcal{R}_{\mathtt{N_f}}$. From equation (3.4) and equation (3.6), we have

$$\begin{aligned}
\mathcal{R}_{\mathtt{N_f}+1} &= \begin{bmatrix} P & L \end{bmatrix} \left\{ \left( \mathscr{W}_{\mathtt{N_f}} \cap \mathtt{ker} \begin{bmatrix} M & 0 \end{bmatrix} \right) \oplus \mathscr{P} \right\} \\
&= \begin{bmatrix} P & L \end{bmatrix} \left\{ \sum_{i=0}^{\mathtt{N_f}-1} \left( \mathtt{img} \begin{bmatrix} P^i L \\ 0 \end{bmatrix} \cap \mathtt{ker} \begin{bmatrix} M & 0 \end{bmatrix} \right) \oplus \mathscr{P} \right\} \\
&= \begin{bmatrix} P & L \end{bmatrix} \left\{ \sum_{i=0}^{\mathtt{N_f}-2} \left( \mathtt{img} \begin{bmatrix} P^i L \\ 0 \end{bmatrix} \cap \mathtt{ker} \begin{bmatrix} M & 0 \end{bmatrix} \right) \right. \\
&\qquad \left. + \left( \mathtt{img} \begin{bmatrix} P^{\mathtt{N_f}-1} L \\ 0 \end{bmatrix} \cap \mathtt{ker} \begin{bmatrix} M & 0 \end{bmatrix} \right) \oplus \mathscr{P} \right\}. \qquad (3.7)
\end{aligned}$$

From Lemma 3.1, it is evident that $MP^{\mathtt{N_f}-1}L \neq 0$. Hence, $\mathtt{img} \left[ \begin{smallmatrix} P^{\mathtt{N_f}-1}L \\ 0 \end{smallmatrix} \right] \cap \mathtt{ker} \left[ \begin{smallmatrix} M & 0 \end{smallmatrix} \right] = 0$. Hence, from equation (3.6) and equation (3.7) we have $\mathcal{R}_{\mathtt{N_f}+1} = \mathcal{R}_{\mathtt{N_f}}$. Thus, from [10] (see discussion after equation 3.22), we infer that $\mathcal{R}_{\mathtt{N_f}}$ characterized in equation (3.6) is the fast subspace $\mathcal{R}_s$ of $\Sigma_{\mathtt{P}}$, i.e., $\mathcal{R}_{\mathtt{N_f}} = \mathcal{R}_s$. From equation (3.6), Statement *(1)* of the lemma directly follows.

*(2):* Define $W := \begin{bmatrix} L & PL & \cdots & P^{\mathtt{N_f}-1}L \end{bmatrix}$. We want to show that $W$ is full column-rank. To the contrary, let us assume that there exists a nontrivial vector $w \in \mathbb{R}^{\mathtt{N_f} \times 1}$ such that $Ww = 0$. Conforming to the partition of $W$ let $w := \mathtt{col}(w_0, w_1, \ldots, w_{\mathtt{N_f}-1})$.

Now, we pre-multiply $W$ with $M$ in the equation $Ww = 0$ and use the fact that $MP^kL = 0$ for $k \in \{0, 1, \ldots, \mathtt{N_f} - 2\}$ from Lemma 3.1:

$$\begin{aligned}
\begin{bmatrix} ML & MPL & \cdots & MP^{\mathtt{N_f}-1}L \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{\mathtt{N_f}-1} \end{bmatrix} = 0 &\Rightarrow MP^{\mathtt{N_f}-1}Lw_{\mathtt{N_f}-1} = 0 \\
&\Rightarrow w_{\mathtt{N_f}-1} = 0 \, (\text{since } MP^{\mathtt{N_f}-1}L \neq 0).
\end{aligned}$$

Next, we pre-multiply $W$ with $MP$ in the equation $Ww = 0$ and use Lemma 3.1 with the fact that $w_{\mathtt{N_f}-1} = 0$:

$$\begin{aligned}
\begin{bmatrix} MPL & MP^2L & \cdots MP^{\mathtt{N_f}-1}L & MP^{\mathtt{N_f}}L \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{\mathtt{N_f}-2} \\ 0 \end{bmatrix} = 0 &\Rightarrow MP^{\mathtt{N_f}-1}Lw_{\mathtt{N_f}-2} = 0 \\
&\Rightarrow w_{\mathtt{N_f}-2} = 0.
\end{aligned}$$

Continuing in the same manner, it follows that $w_i = 0$ for $i \in \{0, 1, \ldots, \mathtt{N_f} - 1\}$. However, this is a contradiction since we assume $w$ to be nonzero. Therefore, there exists no nontrivial vector in the kernel of $W$, i.e., $W$ is full column-rank. Hence, from Statement *(1)* of the lemma, it directly follows that $\mathtt{dim}(\mathcal{R}_s) = \mathtt{N_f}$.
□

The fact that the fast subspace of a system is spanned by the columns of a truncated controllability matrix has been alluded to in [20]. However, the important takeaway from Theorem 3.1 is the relation between the degree of the determinant of the Rosenbrock system matrix $(sU_1 - U_2)$, and the dimension of the fast subspace. We use this relation crucially in Section 4 to compute the maximal rank-minimizing solution of an LQR LMI.

### 3.2 Characterization of slow subspace as an eigenspace of the Rosenbrock matrix pair

As motivated in Section 2.4, let $\mathcal{O}_w$ be the slow subspace of the system $\Sigma_{\mathtt{P}}$ defined in equation (3.1). In the next lemma we establish that $\mathcal{O}_w$ can be characterized by the eigenvectors of $(U_1, U_2)$.

**Theorem 3.2** *Consider the system $\Sigma_{\mathtt{P}}$ as defined in equation* (3.1) *and the corresponding Rosenbrock matrix pair $(U_1, U_2)$ as defined in equation* (3.2)*. Assume $\mathtt{det}(sU_1 - U_2) \neq 0$ and $\mathtt{degdet}(sU_1 - U_2) =: \mathtt{N_s}$. Consider $\mathcal{O}_w$ to be the slow subspace of $\Sigma_{\mathtt{P}}$. Let $\widehat{V}_1 \in \mathbb{R}^{\mathtt{N}\times\mathtt{N_s}}$ and $\widehat{V}_2 \in \mathbb{R}^{1\times\mathtt{N_s}}$ be such that $\mathtt{col}(\widehat{V}_1, \widehat{V}_2)$ is full column-rank and*

$$\underbrace{\begin{bmatrix} P & L \\ M & 0 \end{bmatrix}}_{U_2} \begin{bmatrix} \widehat{V}_1 \\ \widehat{V}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} I_{\mathtt{N}} & 0 \\ 0 & 0 \end{bmatrix}}_{U_1} \begin{bmatrix} \widehat{V}_1 \\ \widehat{V}_2 \end{bmatrix} J, \tag{3.8}$$

*where $J \in \mathbb{R}^{\mathtt{N_s}\times\mathtt{N_s}}$ and $\sigma(J) = \mathtt{roots}\left(\mathtt{det}(sU_1 - U_2)\right)$. Then, the following statements are true:*

1. *$\mathcal{O}_w = \mathtt{img}\,\widehat{V}_1$.*
2. *$\mathtt{dim}(\mathcal{O}_w) = \mathtt{N_s}$.*
3. *$\widehat{V}_1$ is full column-rank.*

*Proof: (1):* From equation (3.8), it is clear that $P\widehat{V}_1 + L\widehat{V}_2 = \widehat{V}_1 J$. Hence, by Proposition 2.2, $\mathtt{img}\,\widehat{V}_1$ is a $(P, L)$-invariant subspace. Further, from equation (3.8), $M\widehat{V}_1 = 0$. Therefore, $\mathtt{img}\,\widehat{V}_1 \in \mathfrak{I}\,(P, L; \mathtt{ker}\,M)$. We claim that $\mathtt{img}\,\widehat{V}_1 = \mathtt{sup}\,\mathfrak{I}\,(P, L; \mathtt{ker}\,M) = \mathcal{O}_w$ (Proposition 2.3).

Let us assume to the contrary that $\mathtt{img}\,\widehat{V}_1$ is not the largest $(P, L)$-invariant subspace inside $\mathtt{ker}\,M$. Then, there exists a nontrivial subspace $\mathcal{V}_e$ such that the space $\mathtt{img}\,\widehat{V}_1 \oplus \mathcal{V}_e = \mathcal{O}_w$, where $\mathtt{dim}(\mathcal{V}_e) =: \ell$. Let $\mathcal{V}_e = \mathtt{img}\,\widehat{V}_e$, where $\widehat{V}_e \in \mathbb{R}^{\mathtt{N}\times\ell}$ is a full column-rank matrix. Since $\mathtt{img}\,\widehat{V}_1 \oplus \mathcal{V}_e = \mathcal{O}_w$ and $\mathcal{O}_w$ is $(P, L)$-invariant inside $\mathtt{ker}\,M$, $P\mathcal{V}_e \subseteq \mathcal{O}_w + \mathtt{img}\,L$ (by Proposition 2.2) and

$M\mathcal{V}_e = \{0\}$. Therefore, there exist $T_1 \in \mathbb{R}^{1\times\ell}$, $T_2 \in \mathbb{R}^{\mathtt{N_s}\times\ell}$, and $T_3 \in \mathbb{R}^{\ell\times\ell}$ such that

$$P\widehat{V}_e = LT_1 + \begin{bmatrix}\widehat{V}_1 & \widehat{V}_e\end{bmatrix}\begin{bmatrix}T_2\\ T_3\end{bmatrix} \text{ and } M\widehat{V}_e = 0. \tag{3.9}$$

Therefore, writing equation (3.8) and equation (3.9) together we have

$$\underbrace{\begin{bmatrix}P & L\\ M & 0\end{bmatrix}}_{U_1}\begin{bmatrix}\widehat{V}_1 & \widehat{V}_e\\ \widehat{V}_2 & -T_1\end{bmatrix} = \underbrace{\begin{bmatrix}I_{\mathtt{N}} & 0\\ 0 & 0\end{bmatrix}}_{U_2}\begin{bmatrix}\widehat{V}_1 & \widehat{V}_e\\ \widehat{V}_2 & -T_1\end{bmatrix}\begin{bmatrix}J & T_2\\ 0 & T_3\end{bmatrix}. \tag{3.10}$$

Since $(sU_1 - U_2)$ is a regular matrix pencil, we can rewrite $(U_1, U_2)$ in the canonical form as motivated in Section 2.2. Therefore there exist nonsingular matrices $Z_1, Z_2 \in \mathbb{R}^{(\mathtt{N}+1)\times(\mathtt{N}+1)}$ such that $U_1 = Z_1 \left[\begin{smallmatrix}I & 0\\ 0 & Y\end{smallmatrix}\right] Z_2$ and $U_2 = Z_1 \left[\begin{smallmatrix}J & 0\\ 0 & I\end{smallmatrix}\right] Z_2$, where $Y \in \mathbb{R}^{(\mathtt{N}+1-\mathtt{N_s})\times(\mathtt{N}+1-\mathtt{N_s})}$ is a nilpotent matrix. Define $\widehat{U}_1 := \left[\begin{smallmatrix}I & 0\\ 0 & Y\end{smallmatrix}\right]$ and $\widehat{U}_2 := \left[\begin{smallmatrix}J & 0\\ 0 & I\end{smallmatrix}\right]$. Using this in equation (3.10), we have

$$Z_1\begin{bmatrix}J & 0\\ 0 & I\end{bmatrix}Z_2\begin{bmatrix}\widehat{V}_1 & \widehat{V}_e\\ \widehat{V}_2 & -T_1\end{bmatrix} = Z_1\begin{bmatrix}I & 0\\ 0 & Y\end{bmatrix}Z_2\begin{bmatrix}\widehat{V}_1 & \widehat{V}_e\\ \widehat{V}_2 & -T_1\end{bmatrix}\begin{bmatrix}J & T_2\\ 0 & T_3\end{bmatrix}. \tag{3.11}$$

From equation (3.11) it is clear that $\mathtt{img}\left(Z_2\left[\begin{smallmatrix}\widehat{V}_1\\ \widehat{V}_2\end{smallmatrix}\right]\right)$ is the eigenspace of the matrix pair $(\widehat{U}_1, \widehat{U}_2)$. Note that any eigenvector (or generalized eigenvector) of the matrix pair $(\widehat{U}_1, \widehat{U}_2)$ will be of the form $\mathtt{col}(w, 0) \in \mathbb{R}^{(\mathtt{N}+1)\times 1}$, where $w \in \mathbb{R}^{\mathtt{N_s}\times 1}$ is an eigenvector (or generalized eigenvector) of $J$. Therefore, there must exist a nonsingular matrix $T_{\mathtt{N_s}} \in \mathbb{R}^{\mathtt{N_s}\times\mathtt{N_s}}$ such that $Z_2\left[\begin{smallmatrix}\widehat{V}_1\\ \widehat{V}_2\end{smallmatrix}\right] = \left[\begin{smallmatrix}T_{\mathtt{N_s}}\\ 0\end{smallmatrix}\right] \in \mathbb{R}^{(\mathtt{N}+1)\times\mathtt{N_s}}$. Define $Z_2\left[\begin{smallmatrix}\widehat{V}_e\\ -T_1\end{smallmatrix}\right] =: \left[\begin{smallmatrix}\Upsilon_1\\ \Upsilon_2\end{smallmatrix}\right]$, $\Upsilon_1 \in \mathbb{R}^{\mathtt{N_s}\times\ell}$ and $\Upsilon_2 \in \mathbb{R}^{(\mathtt{N}+1-\mathtt{N_s})\times\ell}$. Thus, from equation (3.11) we have

$$\begin{bmatrix}J & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}T_{\mathtt{N_s}} & \Upsilon_1\\ 0 & \Upsilon_2\end{bmatrix} = \begin{bmatrix}I & 0\\ 0 & Y\end{bmatrix}\begin{bmatrix}T_{\mathtt{N_s}} & \Upsilon_1\\ 0 & \Upsilon_2\end{bmatrix}\begin{bmatrix}J & T_2\\ 0 & T_3\end{bmatrix}. \tag{3.12}$$

Thus, we have $\Upsilon_2 = Y\Upsilon_2 T_3 \Rightarrow Y\Upsilon_2 T_3 = Y^2\Upsilon_2 T_3^2 = \Upsilon_2$. Using this line of reasoning, it is evident that $Y^k\Upsilon_2 T_3^k = \Upsilon_2$ for all $k \in \mathbb{N}$. Since $Y$ is a nilpotent matrix, there exists an $i \in \mathbb{N}$ such that $Y^i = 0$. Therefore, we must have $\Upsilon_2 = 0$. Since, $T_{\mathtt{N_s}}$ is a nonsingular matrix, $\mathtt{img}\,\Upsilon_1 \subsetneq T_{\mathtt{N_s}}$. Thus, we have

$$\mathtt{img}\left[\begin{smallmatrix}\Upsilon_1\\ \Upsilon_2\end{smallmatrix}\right] = \mathtt{img}\left[\begin{smallmatrix}\Upsilon_1\\ 0\end{smallmatrix}\right] \subsetneq \mathtt{img}\left[\begin{smallmatrix}T_{\mathtt{N_s}}\\ 0\end{smallmatrix}\right] \Rightarrow \mathtt{img}\begin{bmatrix}\widehat{V}_e\\ -T_1\end{bmatrix} \subsetneq \mathtt{img}\begin{bmatrix}\widehat{V}_1\\ \widehat{V}_2\end{bmatrix} \Rightarrow \mathtt{img}\,\widehat{V}_e \subsetneq \mathtt{img}\,\widehat{V}_1.$$

Therefore, there does not exist any nontrivial subspace $\mathcal{V}_e$ such that $\mathtt{img}\,\widehat{V}_1 \oplus \mathcal{V}_e = \mathcal{O}_w$. This is a contradiction to $\mathtt{img}\,\widehat{V}_1 \neq \mathtt{sup}\,\mathfrak{I}(P, L; \mathtt{ker}\,M)$. Hence, $\mathtt{img}\,\widehat{V}_1 = \mathcal{O}_w$.

*(2):* Define $G(s) := M(sI_{\mathtt{N}} - P)^{-1}L$. Further, note that $\mathtt{det}(sU_1 - U_2) = \mathtt{det}(-M(sI_{\mathtt{N}} - P)^{-1}L) \times \mathtt{det}(sI_{\mathtt{N}} - P)$. Since $\mathtt{det}(sU_1 - U_2) \neq 0$, $\mathtt{det}(M(sI_{\mathtt{N}} -$

$P)^{-1}L) \neq 0$. Hence, $G(s)$ is nonzero rational function. Therefore from Proposition 2.4 we have $\mathcal{O}_w \oplus \mathcal{R}_s = \mathbb{R}^{\mathtt{N}}$. From Statement *(2)* of Lemma 3.1, we know that $\mathtt{dim}(\mathcal{R}_s) = \mathtt{N} - \mathtt{N_s}$. Therefore, $\mathtt{dim}(\mathcal{O}_w) = \mathtt{N_s}$.
*(3)*: From Statements *(1)* and *(2)* of this theorem, it follows that $\mathtt{dim}(\mathcal{O}_w) = \mathtt{dim}\left(\mathtt{img}\, \widehat{V}_1\right) = \mathtt{N_s}$. Therefore, $\widehat{V}_1$ is full column-rank. □

Next we characterize the good slow subspace of the system $\Sigma_{\mathtt{P}}$ in terms of the eigenspace of the Rosenbrock matrix pair $(U_1, U_2)$. From Theorem 3.2 it is clear that the columns of $\widehat{V}_1$ is the basis of $\mathcal{O}_w$. Further, from equation (3.8) we know that

$$\begin{bmatrix} P & L \\ M & 0 \end{bmatrix} \begin{bmatrix} \widehat{V}_1 \\ \widehat{V}_2 \end{bmatrix} = \begin{bmatrix} I_{\mathtt{N}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \widehat{V}_1 \\ \widehat{V}_2 \end{bmatrix} J. \tag{3.13}$$

Assuming that $\sigma(J) \cap j\mathbb{R} = \emptyset$, it is clear that $\sigma(J)$ can be partitioned as $\sigma(J) = \sigma_{\mathtt{g}}(J) \uplus \sigma_{\mathtt{b}}(J)$, where $\sigma_{\mathtt{g}}(J) \subsetneq \mathbb{C}_-$, $\sigma_{\mathtt{b}}(J) \subsetneq \mathbb{C}^+$. Therefore, there exists a nonsingular matrix $T$ such that $T^{-1}JT = \begin{bmatrix} J_{\mathtt{g}} & 0 \\ 0 & J_{\mathtt{b}} \end{bmatrix}$, where $\sigma(J_{\mathtt{g}}) = \sigma_{\mathtt{g}}(J)$ and $\sigma(J_{\mathtt{b}}) = \sigma_{\mathtt{b}}(J)$. Define $\begin{bmatrix} \widehat{V}_1 \\ \widehat{V}_2 \end{bmatrix} T = \begin{bmatrix} \widehat{V}_{1\mathtt{g}} & \widehat{V}_{1\mathtt{b}} \\ \widehat{V}_{2\mathtt{g}} & \widehat{V}_{2\mathtt{b}} \end{bmatrix}$ where the partitioning is done conforming to the partition in $T^{-1}JT$. Then, equation (3.13) takes the following form:

$$\begin{aligned} \begin{bmatrix} P & L \\ M & 0 \end{bmatrix} \begin{bmatrix} \widehat{V}_1 \\ \widehat{V}_2 \end{bmatrix} T &= \begin{bmatrix} I_{\mathtt{N}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \widehat{V}_1 \\ \widehat{V}_2 \end{bmatrix} TT^{-1}JT \\ \Rightarrow \begin{bmatrix} P & L \\ M & 0 \end{bmatrix} \begin{bmatrix} \widehat{V}_{1\mathtt{g}} & \widehat{V}_{1\mathtt{b}} \\ \widehat{V}_{2\mathtt{g}} & \widehat{V}_{2\mathtt{b}} \end{bmatrix} &= \begin{bmatrix} I_{\mathtt{N}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \widehat{V}_{1\mathtt{g}} & \widehat{V}_{1\mathtt{b}} \\ \widehat{V}_{2\mathtt{g}} & \widehat{V}_{2\mathtt{b}} \end{bmatrix} \begin{bmatrix} J_{\mathtt{g}} & 0 \\ 0 & J_{\mathtt{b}} \end{bmatrix}. \end{aligned} \tag{3.14}$$

Assume $|\sigma_{\mathtt{g}}(J)| = |\sigma(U_1, U_2) \cap \mathbb{C}_-| =: \mathtt{N_g}$. Obviously, from the discussion above $\widehat{V}_{1\mathtt{g}} \in \mathbb{R}^{\mathtt{N} \times \mathtt{N_g}}$. We claim in the next lemma that the good slow subspace of the system $\Sigma_{\mathtt{P}}$ is given by $\mathtt{img}\, \widehat{V}_{1\mathtt{g}}$.

**Lemma 3.2** *Consider the system* $\Sigma_{\mathtt{P}}$ *as defined in equation* (3.1) *and the corresponding Rosenbrock matrix pair* $(U_1, U_2)$ *as defined in equation* (3.2). *Assume* $\mathtt{det}(sU_1 - U_2) \neq 0$ *and* $\sigma(U_1, U_2) \cap j\mathbb{R} = \emptyset$. *Define the family of subspaces:*

$$\mathcal{B} := \{\mathcal{S} \in \Im(P, L, \mathtt{ker}(M) | \exists\, F \in \mathbf{F}(\mathcal{S}) \text{ such that } \sigma\left((P + LF)|_{\mathcal{S}}\right) \subsetneq \mathbb{C}_-\}.$$

*Let* $\mathcal{O}_{wg} := \mathtt{sup}\, \mathcal{B}$. *Consider* $\widehat{V}_{1\mathtt{g}} \in \mathbb{R}^{\mathtt{N} \times \mathtt{N_g}}$ *as defined in equation* (3.14), *where* $|\sigma(U_1, U_2) \cap \mathbb{C}_-| =: \mathtt{N_g}$. *Then, the following statements are true:*

*(1)* $\widehat{V}_{1\mathtt{g}}$ *is full column-rank.*

*(2)* $\mathtt{img}\, \widehat{V}_{1\mathtt{g}} = \mathcal{O}_{wg}$.

*Proof:* *(1):* Since by construction $\widehat{V}_1 T = \begin{bmatrix}\widehat{V}_{1\mathtt{g}} \ \widehat{V}_{1\mathtt{g}}\end{bmatrix}$ with $T$ being nonsingular and $\widehat{V}_1$ full column-rank (by Theorem 3.2), we must have $\widehat{V}_{1\mathtt{g}}$ full column-rank, as well.

*(2):* Using Proposition 2.2, from equation (3.14) we can infer that $\mathtt{img}\,\widehat{V}_{1\mathtt{g}} \in \mathfrak{I}(P, L; \mathtt{ker}\, M)$. Further note that since $\widehat{V}_{1\mathtt{g}}$ is full column-rank (from Statement *(1)* of this lemma), there exists $F \in \mathbb{R}^{1\times \mathtt{N}}$ such that $\widehat{V}_{2\mathtt{g}} = F\widehat{V}_{1\mathtt{g}}$. Therefore, from equation (3.14) it further follows that $(P + LF)\widehat{V}_{1\mathtt{g}} = \widehat{V}_{1\mathtt{g}} J_{\mathtt{g}} \Rightarrow \sigma\left((P + LF)|_{\mathtt{img}\,\widehat{V}_{1\mathtt{g}}}\right) = \sigma(J_{\mathtt{g}}) \subsetneq \mathbb{C}_{-}$ and $F \in \mathbf{F}(\mathtt{img}\,\widehat{V}_{1\mathtt{g}})$. Therefore, $\mathtt{img}\,\widehat{V}_{1\mathtt{g}} \in \mathcal{B}$. Let us assume to the contrary that $\mathtt{img}\,\widehat{V}_{1\mathtt{g}} \subsetneq \mathcal{O}_{wg}$. Then there exists a nontrivial subspace $\widetilde{\mathcal{V}}$ such that $\mathtt{img}\,\widehat{V}_{1\mathtt{g}} \oplus \widetilde{\mathcal{V}} = \mathcal{O}_{wg}$. Define $\mathtt{dim}(\widetilde{\mathcal{V}}) =: \mathtt{N}_\ell$. Let $\widetilde{\mathcal{V}} =: \mathtt{img}\,\widetilde{V}$, where $\widetilde{V} \in \mathbb{R}^{\mathtt{N}\times \mathtt{N}_\ell}$ is full column-rank. Following the same line of argument as in the proof of Statement *(1)* of Theorem 3.2, there exist $\widehat{T}_1 \in \mathbb{R}^{1\times \mathtt{N}_\ell}, \widehat{T}_2 \in \mathbb{R}^{\mathtt{N}_\mathtt{g}\times \mathtt{N}_\ell}$ and $\widehat{T}_3 \in \mathbb{R}^{\mathtt{N}_\ell\times \mathtt{N}_\ell}$ such that

$$P\widetilde{V} = L\widehat{T}_1 + \begin{bmatrix}\widehat{V}_{1\mathtt{g}} \ \widetilde{V}\end{bmatrix}\begin{bmatrix}\widehat{T}_2\\ \widehat{T}_3\end{bmatrix}, M\widetilde{V} = 0 \text{ and } \sigma(\widehat{T}_3) \subsetneq \mathbb{C}_{-}. \quad (3.15)$$

Therefore, from equation (3.14) and equation (3.15) we have

$$\underbrace{\begin{bmatrix}P & L\\ M & 0\end{bmatrix}}_{U_2}\begin{bmatrix}\widehat{V}_{1\mathtt{g}} & \widetilde{V}\\ \widehat{V}_{2\mathtt{g}} & -\widehat{T}_1\end{bmatrix} = \underbrace{\begin{bmatrix}I_\mathtt{N} & 0\\ 0 & 0\end{bmatrix}}_{U_1}\begin{bmatrix}\widehat{V}_{1\mathtt{g}} & \widetilde{V}\\ \widehat{V}_{2\mathtt{g}} & -\widehat{T}_1\end{bmatrix}\begin{bmatrix}J_\mathtt{g} & \widehat{T}_2\\ 0 & \widehat{T}_3\end{bmatrix} \text{ and } \sigma(\widehat{T}_3) \cup \sigma(J_\mathtt{g}) \subsetneq \mathbb{C}_{-}. \quad (3.16)$$

Now there exist nonsingular matrices $Z_1, Z_2 \in \mathbb{R}^{(\mathtt{N}+1)\times(\mathtt{N}+1)}$ such that $U_1 = Z_1\left[\begin{smallmatrix}I & 0\\ 0 & Y\end{smallmatrix}\right]Z_2$ and $U_2 = Z_1\left[\begin{smallmatrix}J & 0\\ 0 & I\end{smallmatrix}\right]Z_2$. Therefore, equation (3.16) takes the following form:

$$Z_1\underbrace{\begin{bmatrix}J & 0\\ 0 & I\end{bmatrix}}_{\widehat{U}_2}Z_2\begin{bmatrix}\widehat{V}_{1\mathtt{g}} & \widetilde{V}\\ \widehat{V}_{2\mathtt{g}} & -\widehat{T}_1\end{bmatrix} = Z_1\underbrace{\begin{bmatrix}I & 0\\ 0 & Y\end{bmatrix}}_{\widehat{U}_1}Z_2\begin{bmatrix}\widehat{V}_{1\mathtt{g}} & \widetilde{V}\\ \widehat{V}_{2\mathtt{g}} & -\widehat{T}_1\end{bmatrix}\begin{bmatrix}J_\mathtt{g} & \widehat{T}_2\\ 0 & \widehat{T}_3\end{bmatrix}. \quad (3.17)$$

From equation (3.17) it is clear that $\mathtt{img}\left(Z_2\left[\begin{smallmatrix}\widehat{V}_{1\mathtt{g}}\\ \widehat{V}_{2\mathtt{g}}\end{smallmatrix}\right]\right)$ is a subspace of the eigenspace of the matrix pair $(\widehat{U}_1, \widehat{U}_2)$. Note that any eigenvector (or generalized eigenvector) of the matrix pair $(\widehat{U}_1, \widehat{U}_2)$ will be of the form $\mathtt{col}(w, 0) \in \mathbb{R}^{(\mathtt{N}+1)\times 1}$, where $w \in \mathbb{R}^{\mathtt{N}_\mathtt{s}\times 1}$ is an eigenvector (or generalized eigenvector) of $J_\mathtt{g}$. Thus, there exists a full column-rank matrix $T_{\mathtt{N}_\mathtt{g}} \in \mathbb{R}^{\mathtt{N}_\mathtt{s}\times\mathtt{N}_\mathtt{g}}$ such that $Z_2\left[\begin{smallmatrix}\widehat{V}_{1\mathtt{g}}\\ \widehat{V}_{2\mathtt{g}}\end{smallmatrix}\right] = \left[\begin{smallmatrix}T_{\mathtt{N}_\mathtt{g}}\\ 0\end{smallmatrix}\right] \in \mathbb{R}^{(\mathtt{N}+1)\times\mathtt{N}_\mathtt{g}}$. Define $Z_2\left[\begin{smallmatrix}\widetilde{V}\\ -\widehat{T}_1\end{smallmatrix}\right] =: \left[\begin{smallmatrix}\widehat{\Upsilon}_1\\ \widehat{\Upsilon}_2\end{smallmatrix}\right]$, where $\widehat{\Upsilon}_1 \in \mathbb{R}^{\mathtt{N}_\mathtt{s}\times\mathtt{N}_\mathtt{g}}$ and $\widehat{\Upsilon}_2 \in \mathbb{R}^{(\mathtt{N}+1-\mathtt{N}_\mathtt{s})\times\mathtt{N}_\mathtt{g}}$. Thus, from equation (3.17) we have

$$\begin{bmatrix}J & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}T_{\mathtt{N}_\mathtt{g}} & \widehat{\Upsilon}_1\\ 0 & \widehat{\Upsilon}_2\end{bmatrix} = \begin{bmatrix}I & 0\\ 0 & Y\end{bmatrix}\begin{bmatrix}T_{\mathtt{N}_\mathtt{g}} & \widehat{\Upsilon}_1\\ 0 & \widehat{\Upsilon}_2\end{bmatrix}\begin{bmatrix}J_\mathtt{g} & \widehat{T}_2\\ 0 & \widehat{T}_3\end{bmatrix}. \quad (3.18)$$

From equation (3.18), we have $\widehat{\Upsilon}_2 = Y\widehat{\Upsilon}_2\widehat{T}_3$. Using the fact that $Y$ is nilpotent it can be shown, similar to the proof of Statement *(1)* of Theorem 3.2, that $\widehat{\Upsilon}_2 = 0$. Hence, equation (3.18) can be rewritten as

$$J\begin{bmatrix} T_{\mathtt{N_g}} & \widehat{\Upsilon}_1 \end{bmatrix} = \begin{bmatrix} T_{\mathtt{N_g}} & \widehat{\Upsilon}_1 \end{bmatrix}\begin{bmatrix} J_{\mathtt{g}} & \widehat{T}_2 \\ 0 & \widehat{T}_3 \end{bmatrix}. \tag{3.19}$$

It follows from equation (3.19) that $\sigma(J_{\mathtt{g}}) \cup \sigma(\widehat{T}_3) \subseteq \sigma(J)$. But, recall that $\sigma(J) \cap \mathbb{C}_- = \sigma(J_{\mathtt{g}})$, and $\sigma(J) \cap j\mathbb{R} = \emptyset$. Therefore, we must have $\sigma(\widehat{T}_3) \subsetneq \mathbb{C}^+$. However this is a contradiction to the fact that $\sigma(\widehat{T}_3) \subsetneq \mathbb{C}_-$ (see equation (3.15)). Therefore, there exists no nontrivial subspace $\widetilde{\mathcal{V}}$ such that $\mathtt{img}\,\widehat{V}_{1\mathtt{g}} \oplus \widetilde{\mathcal{V}} = \mathcal{O}_{wg}$. Hence, $\widehat{V}_{1\mathtt{g}} = \mathcal{O}_{wg}$. □

## 4 Constructive solutions of the LQR LMI for single-input systems

At the very outset of this section, we present the first main result of this section that leads to a method to compute the maximal rank-minimizing solution of an LQR LMI. As motivated in Section 1, the next theorem also provides a method to compute the optimal cost for a singular LQR problem.

**Theorem 4.1** *Consider Problem 2.1 with the corresponding Hamiltonian matrix pair $(E, H)$ as defined in equation (2.6). Assume $\sigma(E, H) \cap j\mathbb{R} = \emptyset$ and $\mathtt{det}(sE - H) \neq 0$. Define $\mathtt{deg}\,\mathtt{det}(sE - H) =: 2\mathtt{n_s}$. Let $\Lambda$ be a Lambda-set of $\mathtt{det}(sE - H)$ with cardinality $\mathtt{n_s} < \mathtt{n}$ such that $\Lambda \subsetneq \mathbb{C}_-$. Let $V_{1\Lambda}, V_{2\Lambda} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n_s}}$ and $V_{3\Lambda} \in \mathbb{R}^{1\times\mathtt{n_s}}$ be such that the columns of $V_{\mathtt{e}\Lambda} = \mathtt{col}(V_{1\Lambda}, V_{2\Lambda}, V_{3\Lambda})$ form a basis of the $\mathtt{n_s}$-dimensional eigenspace of $(E, H)$ corresponding to the eigenvalues of $(E, H)$ in $\Lambda$, i.e.,*

$$\begin{bmatrix} A & 0 & b \\ -Q & -A^T & 0 \\ 0 & b^T & 0 \end{bmatrix}\begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \\ V_{3\Lambda} \end{bmatrix} = \begin{bmatrix} I_{\mathtt{n}} & 0 & 0 \\ 0 & I_{\mathtt{n}} & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \\ V_{3\Lambda} \end{bmatrix}\Gamma, \tag{4.1}$$

*where $\sigma(\Gamma) = \Lambda$. Construct $V_\Lambda := \mathtt{col}\,(V_{1\Lambda}, V_{2\Lambda})$ and assume $\mathtt{n_f} := \mathtt{n} - \mathtt{n_s}$. Define $W := \begin{bmatrix} \widehat{b} & \widehat{A}\widehat{b} & \cdots & \widehat{A}^{\mathtt{n_f}-1}\widehat{b} \end{bmatrix} \in \mathbb{R}^{2\mathtt{n}\times\mathtt{n_f}}$, where $\widehat{A}, \widehat{b}$ and $\widehat{c}$ are as defined in equation (2.7). Let $X_\Lambda := \begin{bmatrix} V_\Lambda & W \end{bmatrix} =: \left[\begin{smallmatrix} X_{1\Lambda} \\ X_{2\Lambda} \end{smallmatrix}\right]$, where $X_{1\Lambda}, X_{2\Lambda} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$. Then, the following statements hold:*

*(1)* *$X_{1\Lambda}$ is invertible.*

*(2)* *$K_{\mathtt{max}} := X_{2\Lambda}X_{1\Lambda}^{-1}$ is symmetric.*

*(3)* *$K_{\mathtt{max}}$ is a rank-minimizing solution of LMI (2.4).*

*(4)* *For any other solution $K$ of LMI (2.4), $K \leqslant K_{\mathtt{max}}$.*

*(5)* $K_{\mathtt{max}} \geqslant 0$.

We defer the proof of this theorem till the development of a few auxiliary results. Note the close parallel between Proposition 2.5 and Theorem 4.1. For the case when $\mathtt{n_f} = 0$, i.e. the regular LQR case, Theorem 4.1 is indeed equivalent to Proposition 2.5. Thus, Theorem 4.1 is a generalization to Proposition 2.5. To demonstrate that Theorem 4.1 finds the maximal rank-minimizing solution of the LQR LMI (1.3), we revisit Example 2.1 that we have previously failed to solve using Proposition 2.5.

*Example 4.1* Note that in Example 2.1, we have $\mathtt{n} = 3$ and $\mathtt{n_s} = 1$. Thus, $\mathtt{n_f} = \mathtt{n} - \mathtt{n_s} = 2$. Therefore, using Theorem 4.1, we have

$$[V_\Lambda \; W] = \begin{bmatrix} V_{1\Lambda} & b & Ab \\ V_{2\Lambda} & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow K_{\mathtt{max}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

It can be verified that LQR LMI (2.4) evaluated at $K_{\mathtt{max}}$ turns out to be $\begin{bmatrix} 4 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \geqslant 0$. Further, $\mathtt{rank}(K_{\mathtt{max}}) = 1$. This is the minimum rank achievable by the LQR LMI (2.4) (see proof of Statement *(3)* of Theorem 4.1 in Section 4.3 for the justification of the LQR LMI's minimum rank being 1 in this case). Further, $K_{\mathtt{max}}$ is also the maximal solution of the LQR LMI (2.4) (see proof of Statement *(4)* of Theorem 4.1 in Section 4.3 for a justification of this claim). Thus, from the example it is clear that Theorem 4.1 indeed provides a method to compute the maximal rank-minimizing solution of an LQR LMI corresponding to a singular LQR problem.

Now we relate the results in Section 3 with the Hamiltonian system $\Sigma_{\text{Ham}}$ defined in Section 2.5. Using the parallel between the output-nulling representations of $\Sigma_{\mathtt{P}}$ (in equation (3.1)) and $\Sigma_{\text{Ham}}$ (in equation (2.7)), we define $P := \widehat{A}$, $L := \widehat{b}$, $M := \widehat{c}$, $U_1 := E$, and $U_2 := H$. Further, we have $\mathtt{degdet}(sE - H) = 2\mathtt{n_s}$. Therefore, $\mathtt{N_s} = 2\mathtt{n_s}$ and $\mathtt{N_f} = \mathtt{N} - \mathtt{N_s} = 2\mathtt{n} - 2\mathtt{n_s} = 2\mathtt{n_f}$. Hence, Theorem 3.1, Theorem 3.2, and Lemma 3.2 can be directly applied to the system $\Sigma_{\text{Ham}}$.

Note that since $|\sigma(\Gamma)| = \mathtt{n_s}$ and $\sigma(E, H) \cap \mathbb{C}_- = \sigma(\Gamma)$, in terms of Lemma 3.2 we have $\mathtt{N_g} = \mathtt{n_s}$. Further, since $\left[\begin{smallmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{smallmatrix}\right] \in \mathbb{R}^{\mathtt{n} \times \mathtt{n_s}}$, we must have from Statement *(1)* of Lemma 3.2 that $\left[\begin{smallmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{smallmatrix}\right]$ is full column-rank. From Lemma 3.2 it is also evident that $\mathtt{img} \left[\begin{smallmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{smallmatrix}\right]$ is the largest good $(\widehat{A}, \widehat{b})$-invariant subspace inside the kernel of $\widehat{c}$. Hence, the good slow subspace of $\Sigma_{\text{Ham}}$ is given by $\mathcal{O}_{wg} = \mathtt{img} \left[\begin{smallmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{smallmatrix}\right]$. Further, using Theorem 3.1, it is also evident that $\mathtt{img}\, W \subsetneq \mathcal{R}_s$, where $W$ is as defined in Theorem 4.1 and $\mathcal{R}_s$ is the fast subspace of $\Sigma_{\text{Ham}}$. For easy reference in the sequel, we formally state these implications as a lemma next.

**Lemma 4.1** *Let $V_{1\Lambda}, V_{2\Lambda}, W$ be as defined in Theorem 4.1. Let $\mathcal{O}_{wg}$ and $\mathcal{R}_s$ be the good slow subspace and fast subspace of $\Sigma_{\text{Ham}}$, respectively. Then, the following statements are true:*

*(1)* $\left[\begin{smallmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{smallmatrix}\right]$ *is full column-rank.*

*(2)* $\mathcal{O}_{wg} = \texttt{img}\left[\begin{smallmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{smallmatrix}\right]$.

*(3)* $\texttt{img}\, W \subsetneq \mathcal{R}_s$.

Before we start developing the results required for the proof of Theorem 4.1, we review a result that establishes the relation between the basis vectors of the left- and right-eigenspaces of the Hamiltonian matrix pair (see [12] for more on these properties).

**Proposition 4.1** *Let the columns of* $V_{\mathtt{e}\Lambda} = \texttt{col}(V_{1\Lambda}, V_{2\Lambda}, V_{3\Lambda})$ *be the eigenbasis of* $(E, H)$ *corresponding to the eigenvalues in* $\Lambda$*, where* $E, H, V_{1\Lambda}, V_{2\Lambda}, V_{3\Lambda}$*, and* $\Lambda$ *are as defined in Theorem 4.1. Then, the following statements are true:*

1. *Rows of* $\left[V_{2\Lambda}^T \; -V_{1\Lambda}^T \; V_{3\Lambda}^T\right]$ *form a basis of the left eigenspace of* $(E, H)$ *corresponding to eigenvalues in* $-\Lambda$.
2. $V_{1\Lambda}^T V_{2\Lambda} = V_{2\Lambda}^T V_{1\Lambda}$.

These properties of the eigenspaces of $(E, H)$ is crucially used in the sequel. Now we develop the results required for the proof of Theorem 4.1. The first step in the proof of Theorem 4.1 is the following theorem:

**Theorem 4.2** *Let* $V_{1\Lambda}$ *be as defined in Theorem 4.1. Then,* $V_{1\Lambda}$ *is full column-rank.*

Since the columns of $V_{\mathtt{e}\Lambda}$ form a basis of an eigenspace of $(E, H)$, in terms of Definition 2.10, Theorem 4.2 establishes that the subspace $\texttt{img}\, V_{\mathtt{e}\Lambda}$ is disconjugate. We develop the proof for the disconjugacy of $\texttt{img}\, V_{\mathtt{e}\Lambda}$ in the next section.

## 4.1 Disconjugacy of $\texttt{img}\, V_{\mathtt{e}\Lambda}$

We prove Theorem 4.2 using a few auxiliary results that we present next.

**Lemma 4.2** *Define the following family*

$$\mathcal{B}_\Sigma := \big\{\mathcal{V} \subsetneq \mathbb{R}^{\mathtt{n}} \,|\, \exists F \in \mathbb{R}^{1\times \mathtt{n}} \text{ such that } (A + bF)\mathcal{V} \subseteq \mathcal{V}, Q\mathcal{V} = 0, \\ \sigma((A + bF)|_{\mathcal{V}}) \subsetneq \mathbb{C}_-\big\}.$$

*Then,* $\mathcal{B}_\Sigma$ *has a unique supremal element.*

*Proof:* It directly follows from [22, Lemma 5.8]. □

Note that the unique supremal element of $\mathcal{B}_\Sigma$ is indeed the largest good $(A, b)$-invariant subspace inside the kernel of $Q$. In the next lemma we establish the relation between this subspace and the subspace $\mathcal{O}_{wg} = \texttt{img}\left[\begin{smallmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{smallmatrix}\right]$ of $\Sigma_{\text{Ham}}$.

**Lemma 4.3** *Let* $\mathcal{V}_{\mathtt{g}} := \sup \mathcal{B}_{\Sigma}$. *Suppose* $V_{\mathtt{g}} \in \mathbb{R}^{\mathtt{n}\times \mathtt{g}}$ *be such that* $V_{\mathtt{g}}$ *is full column-rank and* $\mathtt{img}\, V_{\mathtt{g}} = \mathcal{V}_{\mathtt{g}}$. *Define* $\mathcal{V}_{\mathtt{gHam}} := \mathtt{img} \begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{g}} \end{bmatrix}$. *Let* $V_{1\Lambda}, V_{2\Lambda} \in \mathbb{R}^{\mathtt{n}\times \mathtt{n_s}}$ *be as defined in Theorem 4.1. Then,* $\mathcal{V}_{\mathtt{gHam}} \subseteq \mathtt{img} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$. *In particular, if* $\mathtt{g} = \mathtt{dim}\,(\mathcal{V}_{\mathtt{g}}) = \mathtt{n_s}$ *then* $\mathcal{V}_{\mathtt{gHam}} = \mathtt{img} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$.

*Proof:* Since $\mathcal{V}_{\mathtt{g}} = \mathtt{img}\, V_{\mathtt{g}} \in \mathcal{B}_{\Sigma}$, there exists $F \in \mathbf{F}(\mathcal{V}_{\mathtt{g}})$ such that $(A+bF)V_{\mathtt{g}} = V_{\mathtt{g}} J_{\mathtt{g}}$, where $J_{\mathtt{g}} = (A+bF)|_{\mathcal{V}_{\mathtt{g}}}$ and $QV_{\mathtt{g}} = 0$. On defining $V_{3\mathtt{g}} := FV_{\mathtt{g}}$ we can, therefore, have the following equation:

$$\begin{bmatrix} A & 0 & b \\ -Q & -A^T & 0 \\ 0 & b^T & 0 \end{bmatrix} \begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{g}} \\ V_{3\mathtt{g}} \end{bmatrix} = \begin{bmatrix} I_{\mathtt{n}} & 0 & 0 \\ 0 & I_{\mathtt{n}} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{g}} \\ V_{3\mathtt{g}} \end{bmatrix} J_{\mathtt{g}} \Rightarrow \begin{bmatrix} \widehat{A} & \widehat{b} \\ \widehat{c} & 0 \end{bmatrix} \begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{g}} \\ V_{3\mathtt{g}} \end{bmatrix} = \begin{bmatrix} I_{2\mathtt{n}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{g}} \\ V_{3\mathtt{g}} \end{bmatrix} J_{\mathtt{g}}. \tag{4.2}$$

Thus, $\sigma(J_{\mathtt{g}}) \subsetneq \sigma(E,H)$. Using Proposition 2.2 in equation (4.2) we can infer that $\mathcal{V}_{\mathtt{gHam}} = \mathtt{img} \begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{g}} \end{bmatrix}$ is an $(\widehat{A}, \widehat{b})$-invariant subspace. Further, using the fact that $\widehat{c} \begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{g}} \end{bmatrix} = 0$ in equation (4.2), it is evident that $\mathcal{V}_{\mathtt{gHam}}$ is indeed an $(\widehat{A}, \widehat{b})$-invariant subspace inside $\mathtt{ker}\, \widehat{c}$ with $\sigma(J_{\mathtt{g}}) \subsetneq \mathbb{C}_{-}$. Since $\mathcal{O}_{wg}$ is the largest good $(\widehat{A}, \widehat{b})$-invariant subspace inside $\mathtt{ker}\, \widehat{c}$, we have $\mathcal{V}_{\mathtt{gHam}} \subseteq \mathcal{O}_{wg} = \mathtt{img} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$ (Statement *(2)* of Lemma 4.1).

If $\mathtt{g} = \mathtt{n_s}$, then $\mathcal{V}_{\mathtt{gHam}} = \mathtt{img} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$ is trivially true. □

Since $\mathtt{img} \begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{g}} \end{bmatrix} \subseteq \mathtt{img} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$ and $\begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix} \in \mathbb{R}^{2\mathtt{n}\times \mathtt{n_s}}$ is full column-rank (Statement *(1)* of Lemma 4.1), it is evident that there exists a full column-rank matrix $\begin{bmatrix} V_{1\mathtt{e}} \\ V_{2\mathtt{e}} \end{bmatrix} \in \mathbb{R}^{2\mathtt{n}\times(\mathtt{n_s}-\mathtt{g})}$ such that $\mathtt{img} \begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0_{\mathtt{n},\mathtt{g}} & V_{2\mathtt{e}} \end{bmatrix} = \mathtt{img} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$. The next lemma deals with such an extension.

**Lemma 4.4** *Let* $V_{1\mathtt{e}}, V_{2\mathtt{e}} \in \mathbb{R}^{\mathtt{n}\times(\mathtt{n_s}-\mathtt{g})}$ *be such that* $\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0_{\mathtt{n},\mathtt{g}} & V_{2\mathtt{e}} \end{bmatrix}$ *is full column-rank and* $\mathtt{img} \begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0_{\mathtt{n},\mathtt{g}} & V_{2\mathtt{e}} \end{bmatrix} = \mathcal{O}_{wg}$, *where* $\mathcal{O}_{wg} = \mathtt{img} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$ *with* $V_{1\Lambda}, V_{2\Lambda}$ *as defined in Theorem 4.1 and* $V_{\mathtt{g}}$ *is as defined in Lemma 4.3. Then, the following statements are true:*

*1.* $V_{2\mathtt{e}}$ *is full column-rank.*

*2.* $\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \end{bmatrix}$ *is full column-rank.*

*Proof: (1):* Since $\mathtt{img} \begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0 & V_{2\mathtt{e}} \end{bmatrix}$ is an $(\widehat{A}, \widehat{b})$-invariant subspace inside $\mathtt{ker}\, \widehat{c}$, there exist $V_{3\mathtt{e}} \in \mathbb{R}^{1\times(\mathtt{n_s}-\mathtt{g})}$, $\Gamma_{12} \in \mathbb{R}^{\mathtt{g}\times(\mathtt{n_s}-\mathtt{g})}$ and $\Gamma_{22} \in \mathbb{R}^{(\mathtt{n_s}-\mathtt{g})\times(\mathtt{n_s}-\mathtt{g})}$ such that

$$\widehat{A} \begin{bmatrix} V_{1\mathtt{e}} \\ V_{2\mathtt{e}} \end{bmatrix} = \begin{bmatrix} V_{\mathtt{g}} \\ 0 \end{bmatrix} \Gamma_{12} + \begin{bmatrix} V_{1\mathtt{e}} \\ V_{2\mathtt{e}} \end{bmatrix} \Gamma_{22} - \widehat{b} V_{3\mathtt{e}}, \text{ and } \widehat{c} \begin{bmatrix} V_{1\mathtt{e}} \\ V_{2\mathtt{e}} \end{bmatrix} = 0. \tag{4.3}$$

Now writing equation (4.2) and equation (4.3) together, we have

$$\begin{bmatrix} A & 0 & b \\ -Q & -A^T & 0 \\ 0 & b^T & 0 \end{bmatrix} \begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0 & V_{2\mathtt{e}} \\ V_{3\mathtt{g}} & V_{3\mathtt{e}} \end{bmatrix} = \begin{bmatrix} I_{\mathtt{n}} & 0 & 0 \\ 0 & I_{\mathtt{n}} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0 & V_{2\mathtt{e}} \\ V_{3\mathtt{g}} & V_{3\mathtt{e}} \end{bmatrix} \begin{bmatrix} J_{\mathtt{g}} & \Gamma_{12} \\ 0 & \Gamma_{22} \end{bmatrix}. \tag{4.4}$$

Since $\mathtt{img} \begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0_{\mathtt{n},\mathtt{g}} & V_{2\mathtt{e}} \end{bmatrix} = \mathcal{O}_{wg}$, from equation (4.4) we have $\sigma\left(\begin{bmatrix} J_{\mathtt{g}} & \Gamma_{12} \\ 0 & \Gamma_{22} \end{bmatrix}\right) \subsetneq \mathbb{C}_- \Rightarrow \sigma(\Gamma_{22}) \subsetneq \mathbb{C}_-$. From equation (4.4) we have the following equations:

$$AV_{1\mathtt{e}} + bV_{3\mathtt{e}} = V_{\mathtt{g}}\Gamma_{12} + V_{1\mathtt{e}}\Gamma_{22}, \tag{4.5}$$

$$-QV_{1\mathtt{e}} - A^T V_{2\mathtt{e}} = V_{2\mathtt{e}}\Gamma_{22}, \tag{4.6}$$

$$b^T V_{2\mathtt{e}} = 0. \tag{4.7}$$

From Statement *(2)* of Proposition 4.1, we can infer that

$$\begin{bmatrix} 0 & V_{2\mathtt{e}} \end{bmatrix}^T \begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \end{bmatrix} = \begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \end{bmatrix}^T \begin{bmatrix} 0 & V_{2\mathtt{e}} \end{bmatrix} \Rightarrow \begin{cases} V_{2\mathtt{e}}^T V_{\mathtt{g}} = 0, \\ V_{2\mathtt{e}}^T V_{1\mathtt{e}} = V_{1\mathtt{e}}^T V_{2\mathtt{e}}. \end{cases} \tag{4.8}$$

Now pre-multiplying equations (4.5) and equation (4.6) with $V_{2\mathtt{e}}^T$ and $-V_{1\mathtt{e}}^T$, respectively and adding, we get

$$\begin{aligned} &V_{2\mathtt{e}}^T AV_{1\mathtt{e}} + V_{2\mathtt{e}}^T bV_{3\mathtt{e}} + V_{1\mathtt{e}}^T QV_{1\mathtt{e}} + V_{1\mathtt{e}}^T A^T V_{2\mathtt{e}} \\ &\qquad = V_{2\mathtt{e}}^T V_{\mathtt{g}}\Gamma_{12} + V_{2\mathtt{e}}^T V_{1\mathtt{e}}\Gamma_{22} - V_{1\mathtt{e}}^T V_{2\mathtt{e}}\Gamma_{22}. \end{aligned} \tag{4.9}$$

Using equation (4.7), equation (4.8) in equation (4.9) , we have

$$V_{2\mathtt{e}}^T AV_{1\mathtt{e}} + V_{1\mathtt{e}}^T QV_{1\mathtt{e}} + V_{1\mathtt{e}}^T A^T V_{2\mathtt{e}} = 0. \tag{4.10}$$

Now we prove that $V_{2\mathtt{e}}$ is full column-rank by contradiction. Hence, to the contrary, let us assume $V_{2\mathtt{e}}$ is not full column-rank. Therefore, there exists a nonzero $w \in \mathbb{R}^{(\mathtt{n_s}-\mathtt{g})\times 1}$ such that $V_{2\mathtt{e}}w = 0$. Pre- and post-multiplying equation (4.10) with $w^T$ and $w$, respectively, we get $w^T V_{1\mathtt{e}}^T QV_{1\mathtt{e}}w = 0$. Since $Q \geqslant 0$, we must have

$$QV_{1\mathtt{e}}w = 0 \Rightarrow \mathtt{ker}\, V_{2\mathtt{e}} \subseteq \mathtt{ker}(QV_{1\mathtt{e}}). \tag{4.11}$$

Post-multiplying equation (4.6) with $w$ and using the fact that $V_{2\mathtt{e}}w = 0$ and $QV_{1\mathtt{e}}w = 0$, we have

$$-QV_{1\mathtt{e}}w - A^T V_{2\mathtt{e}}w = V_{2\mathtt{e}}\Gamma_{22}w \Rightarrow 0 = V_{2\mathtt{e}}\Gamma_{22}w \Rightarrow \mathtt{ker}\, V_{2\mathtt{e}} \text{ is } \Gamma_{22}\text{-invariant}. \tag{4.12}$$

Therefore, from equation (4.12) it follows that there exists a full column-rank matrix $\widetilde{T} \in \mathbb{R}^{(\mathtt{n_s}-\mathtt{g})\times \bullet}$ such that $V_{2\mathtt{e}}\widetilde{T} = 0$ and $\Gamma_{22}\widetilde{T} = \widetilde{T}\widetilde{\Gamma}, \sigma(\widetilde{\Gamma}) \subseteq \sigma(\Gamma_{22}) \subsetneq \mathbb{C}_-$. Further, from equation (4.11), we must have $QV_{1\mathtt{e}}\widetilde{T} = 0$. Post-multiplying equation (4.5) by $\widetilde{T}$, we get

$$AV_{1\mathtt{e}}\widetilde{T} + bV_{3\mathtt{e}}\widetilde{T} = V_{\mathtt{g}}\Gamma_{12}\widetilde{T} + V_{1\mathtt{e}}\Gamma_{22}\widetilde{T} \Rightarrow AV_{1\mathtt{e}}\widetilde{T} + bV_{3\mathtt{e}}\widetilde{T} = V_{\mathtt{g}}\Gamma_{12}\widetilde{T} + V_{1\mathtt{e}}\widetilde{T}\widetilde{\Gamma}. \tag{4.13}$$

Using Proposition 2.2 in equation (4.13) combined with the fact that $\mathtt{img}\, V_{\mathtt{g}}$ is a good $(A, b)$-invariant subspace and $\sigma(\widetilde{\Gamma}) \subsetneq \mathbb{C}_-$, we infer that $\mathtt{img}\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}}\widetilde{T}\end{bmatrix}$ is also a good $(A, b)$-invariant subspace. Further, $Q\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}}\widetilde{T}\end{bmatrix} = 0$. Thus, $\mathtt{img}\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}}\widetilde{T}\end{bmatrix} \in \mathcal{B}_\Sigma$, where $\mathcal{B}_\Sigma$ is as defined in Lemma 4.2. Since $\mathcal{V}_{\mathtt{g}} = \mathtt{sup}\, \mathcal{B}_\Sigma$ and $\mathtt{img}\, V_{\mathtt{g}} = \mathcal{V}_{\mathtt{g}}$, we must have $\mathtt{img}\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}}\widetilde{T}\end{bmatrix} = \mathcal{V}_{\mathtt{g}}$. Therefore, there exist $\alpha_1 \in \mathbb{R}^{\mathtt{g}\times 1}$ and a nonzero $\alpha_2 \in \mathbb{R}^{\bullet\times 1}$ such that $V_{\mathtt{g}}\alpha_1 + V_{1\mathtt{e}}\widetilde{T}\alpha_2 = 0$, i.e., $\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0 & V_{2\mathtt{e}}\end{bmatrix}\begin{bmatrix} \alpha_1 \\ \widetilde{T}\alpha_2\end{bmatrix} = \begin{bmatrix} V_{\mathtt{g}}\alpha_1 + V_{1\mathtt{e}}\widetilde{T}\alpha_2 \\ V_{2\mathtt{e}}\widetilde{T}\alpha_2\end{bmatrix} = 0$ (Recall that $V_{2\mathtt{e}}\widetilde{T} = 0$). Note that since $\widetilde{T}$ is full column-rank, $\widetilde{T}\alpha_2 \neq 0$. Thus, we have a nonzero vector $\begin{bmatrix} \alpha_1 \\ \widetilde{T}\alpha_2\end{bmatrix}$ inside $\mathtt{ker}\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0 & V_{2\mathtt{e}}\end{bmatrix}$. This is a contradiction to the fact that $\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0 & V_{2\mathtt{e}}\end{bmatrix}$ is full column-rank. Thus, $V_{2\mathtt{e}}$ must be full column-rank.

*(2):* To the contrary, assume that $\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}}\end{bmatrix}$ is not full column-rank. Then, there exist $\beta_1 \in \mathbb{R}^{\mathtt{g}\times 1}$ and $\beta_2 \in \mathbb{R}^{(\mathtt{n}-\mathtt{g})\times 1}$ such that $\begin{bmatrix} \beta_1 \\ \beta_2\end{bmatrix} \neq 0$ and $V_{\mathtt{g}}\beta_1 + V_{1\mathtt{e}}\beta_2 = 0$.

Now pre-multiplying equation (4.5) with $V_{2\mathtt{e}}^T$ and adding it to the transpose of equation (4.6) post-multiplied with $V_{1\mathtt{e}}$, we have

$$\begin{aligned} V_{2\mathtt{e}}^T A V_{1\mathtt{e}} + V_{2\mathtt{e}}^T b V_{3\mathtt{e}} - V_{1\mathtt{e}}^T Q V_{1\mathtt{e}} - V_{2\mathtt{e}}^T A V_{1\mathtt{e}} \\ = V_{2\mathtt{e}}^T V_{\mathtt{g}} \Gamma_{12} + V_{2\mathtt{e}}^T V_{1\mathtt{e}} \Gamma_{22} + \Gamma_{22}^T V_{2\mathtt{e}}^T V_{1\mathtt{e}}. \end{aligned} \tag{4.14}$$

Using equation (4.7) and equation (4.8) in equation (4.14), we have

$$\Gamma_{22}^T V_{2\mathtt{e}}^T V_{1\mathtt{e}} + V_{2\mathtt{e}}^T V_{1\mathtt{e}} \Gamma_{22} = -V_{1\mathtt{e}}^T Q V_{1\mathtt{e}}. \tag{4.15}$$

Now we assume that there exists a nonzero $y \in \mathtt{ker}\,(V_{2\mathtt{e}}^T V_{1\mathtt{e}})$. Pre- and post-multiplying (4.15) by $y^T$ and $y$, respectively and using equation (4.8) we have

$$\begin{aligned} y^T \Gamma_{22}^T V_{2\mathtt{e}}^T V_{1\mathtt{e}} y + y^T V_{2\mathtt{e}}^T V_{1\mathtt{e}} \Gamma_{22} y = -y^T V_{1\mathtt{e}}^T Q V_{1\mathtt{e}} y \Rightarrow y^T V_{1\mathtt{e}}^T Q V_{1\mathtt{e}} y = 0 \\ \Rightarrow Q V_{1\mathtt{e}} y = 0. \end{aligned} \tag{4.16}$$

Now, post-multiplying equation (4.15) with $y$ and using equation (4.16), we have

$$V_{2\mathtt{e}}^T V_{1\mathtt{e}} \Gamma_{22} y = 0 \Rightarrow \mathtt{ker}(V_{2\mathtt{e}}^T V_{1\mathtt{e}}) \text{ is } \Gamma_{22}\text{-invariant}. \tag{4.17}$$

Therefore, using equation (4.17) and the fact that $\sigma(\Gamma_{22}) \subsetneq \mathbb{C}_-$, we have

$$\exists \text{ a nonzero } q \in \mathbb{C}^{(\mathtt{n}-\mathtt{g})\times 1} \text{ such that } V_{2\mathtt{e}}^T V_{1\mathtt{e}} q = 0 \,\&\, \Gamma_{22} q = \mu q, \text{ where } \mu \in \mathbb{C}_-. \tag{4.18}$$

Post-multiplying equation (4.6), by $q$, we have $-QV_{1\mathtt{e}}q - A^T V_{2\mathtt{e}} q = V_{2\mathtt{e}} \Gamma_{22} q \Rightarrow A^T V_{2\mathtt{e}} q = -\mu V_{2\mathtt{e}} q$. If $V_{2\mathtt{e}} q$ is nonzero, then it is a left-eigenvector of $A$. However, from equation (4.7) we can infer that $(V_{2\mathtt{e}} q)^T b = 0$. This means that the system $(A, b)$ is uncontrollable. This is a contradiction. Therefore, $V_{2\mathtt{e}} q$ must be a zero vector. Now from the fact that $V_{2\mathtt{e}}$ is full column-rank (Statement *(1)* of this lemma), it is evident that $q = 0$, which contradicts equation (4.18). Thus, our initial assumption that there exists a nonzero vector $y \in \mathtt{ker}(V_{2\mathtt{e}}^T V_{1\mathtt{e}})$ is not true. Hence, $\mathtt{ker}(V_{2\mathtt{e}}^T V_{1\mathtt{e}}) = \{0\}$.

Recall that we have assumed $V_{\mathtt{g}}\beta_1 + V_{1\mathtt{e}}\beta_2 = 0$. Pre-multiplying this equation with $V_{2\mathtt{e}}^T$, we have $V_{2\mathtt{e}}^T V_{\mathtt{g}}\beta_1 + V_{2\mathtt{e}}^T V_{1\mathtt{e}}\beta_2 = 0$. Using equation (4.8) and the fact that $\mathtt{ker}(V_{2\mathtt{e}}^T V_{1\mathtt{e}}) = \{0\}$, we have $V_{2\mathtt{e}}^T V_{1\mathtt{e}}\beta_2 = 0 \Rightarrow \beta_2 = 0$. Thus, we have $V_{\mathtt{g}}\beta_1 + V_{1\mathtt{e}}\beta_2 = 0 \Rightarrow V_{\mathtt{g}}\beta_1 = 0$. However, since $V_{\mathtt{g}}$ is full column-rank, we must have $\beta_1 = 0$. This is a contradiction to the fact that $\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \neq 0$. Hence, $\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \end{bmatrix}$ is full column-rank. □

Now using Lemma 4.4, we proceed to prove Theorem 4.2.

*Proof of Theorem 4.2:* Since $\mathtt{img}\begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix} = \mathtt{img}\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \\ 0 & V_{2\mathtt{e}} \end{bmatrix}$, where $V_{\mathtt{g}} \in \mathbb{R}^{\mathtt{n}\times\mathtt{g}}$, and $V_{1\mathtt{e}}, V_{2\mathtt{e}} \in \mathbb{R}^{\mathtt{n}\times(\mathtt{n_s}-\mathtt{g})}$ is as defined in Lemma 4.4, we must have $\mathtt{img}\, V_{1\Lambda} = \mathtt{img}\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \end{bmatrix}$. Note that the number of columns of $V_{1\Lambda}$ and $\begin{bmatrix} V_{\mathtt{g}} & V_{1\mathtt{e}} \end{bmatrix}$ are the same. Therefore, using Statement *(2)* of Lemma 4.4 it follows that $V_{1\Lambda}$ is full column-rank. □

Since $V_{1\Lambda}$ is full column-rank, it follows from Definition 2.10 that $\mathtt{img}\, V_{\mathtt{e}\Lambda}$ is disconjugate. This property of disconjugacy is crucially used to prove Theorem 4.1. Apart from this property, there are a few more identities that are required to prove Theorem 4.1. We present these identities as two lemmas in the next section.

## 4.2 Auxiliary results for the proof of Theorem 4.1

In this section, we present two lemmas that establish a few identities involving the system matrices $(A, b)$, cost matrix $Q$ and a solution $K$ of the LQR LMI (2.4). These identities are crucially used in the proof of Theorem 4.1.

**Lemma 4.5** *Let $(\widehat{A}, \widehat{b}, \widehat{c})$, $Q$ and $\mathtt{n_f}$ be as defined in Theorem 4.1. Then, the following statements are true:*

*(1)* $\widehat{c}\widehat{A}^k\widehat{b} = 0$ *for* $k \in \{0, 1, \ldots, 2(\mathtt{n_f} - 1)\}$.

*(2)* $QA^\ell b = 0$ *for* $\ell \in \{0, 1, \ldots, \mathtt{n_f} - 2\}$.

*(3)* $\widehat{A}^\ell\widehat{b} = \mathtt{col}(A^\ell b, 0)$ *and* $\widehat{c}\widehat{A}^\ell = \begin{bmatrix} 0 & (-1)^\ell (A^\ell b)^T \end{bmatrix}$ *for* $\ell \in \{0, 1, \ldots, \mathtt{n_f} - 1\}$.

*Proof: (1):* We define $P := \widehat{A}$, $L := \widehat{b}$, $M := \widehat{c}$, $U_1 := E$, and $U_2 := H$ in Lemma 3.1. Further we have $\mathtt{degdet}(sE - H) = 2\mathtt{n_s}$. Therefore, $\mathtt{N_s} = 2\mathtt{n_s}$ and $\mathtt{N_f} = \mathtt{N} - \mathtt{N_s} = 2\mathtt{n} - 2\mathtt{n_s} = 2\mathtt{n_f}$. Therefore from Lemma 3.1 Statement *(1)* immediately follows.

*(2)* and *(3)*: Now, we use induction to prove these statements.

**Base step:** ($\ell = 0$) Using Statement *(1)* of this lemma we have

$$\widehat{c}\widehat{A}\widehat{b} = 0 \Rightarrow \begin{bmatrix} 0 \; b^T \end{bmatrix} \begin{bmatrix} A & 0 \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} b \\ 0 \end{bmatrix} = b^T Q b = 0. \tag{4.19}$$

Since $Q \geqslant 0$, using the property of positive-semidefinite matrices in equation (4.19) we get $Qb = 0$. Further, $\widehat{b} = \mathtt{col}(b, 0)$ and $\widehat{c} = \begin{bmatrix} 0 \; b^T \end{bmatrix}$ by definition.
**Induction step**: Assume

$$QA^{\ell}b = 0, \widehat{A}^{\ell}\widehat{b} = \mathtt{col}(A^{\ell}b, 0), \text{ and } \widehat{c}\widehat{A}^{\ell} = \begin{bmatrix} 0 \; (-1)^{\ell} b^T (A^T)^{\ell} \end{bmatrix}, \text{ where } \ell < \mathtt{n_f} - 2.$$

We prove that

$$QA^{\ell+1}b = 0, \widehat{A}^{\ell+1}\widehat{b} = \mathtt{col}(A^{\ell+1}b, 0), \text{ and } \widehat{c}\widehat{A}^{\ell+1} = \begin{bmatrix} 0 \; (-1)^{\ell+1} b^T (A^T)^{\ell+1} \end{bmatrix}.$$

Note that

$$\begin{aligned} \widehat{A}^{\ell+1}\widehat{b} &= \begin{bmatrix} A & 0 \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} A^{\ell}b \\ 0 \end{bmatrix} = \begin{bmatrix} A^{\ell+1}b \\ -QA^{\ell}b \end{bmatrix} = \begin{bmatrix} A^{\ell+1}b \\ 0 \end{bmatrix}, \\ \widehat{c}\widehat{A}^{\ell+1} &= \begin{bmatrix} 0 (-1)^{\ell} (A^{\ell}b)^T \end{bmatrix} \begin{bmatrix} A & 0 \\ -Q & -A^T \end{bmatrix} \\ &= \begin{bmatrix} (-1)^{\ell} (QA^{\ell}b)^T (-1)^{\ell+1} (A^{\ell+1}b)^T \end{bmatrix} = \begin{bmatrix} 0 (-1)^{\ell+1} (A^{\ell+1}b)^T \end{bmatrix}. \end{aligned}$$

Since $\ell < \mathtt{n_f} - 2 \Rightarrow 2\ell + 3 < 2\mathtt{n_f} - 1$, using Statement *(1)* of this lemma and the induction hypothesis, we have

$$\begin{aligned} \widehat{c}\widehat{A}^{2\ell+3}\widehat{b} = 0 &\Rightarrow (\widehat{c}\widehat{A}^{\ell+1})\widehat{A}(\widehat{A}^{\ell+1}\widehat{b}) = 0 \\ &\Rightarrow \begin{bmatrix} 0 (-1)^{\ell+1} (A^{\ell+1}b)^T \end{bmatrix} \begin{bmatrix} A & 0 \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} A^{\ell+1}b \\ 0 \end{bmatrix} = 0 \\ &\Rightarrow (A^{\ell+1}b)^T Q (A^{\ell+1}b) = 0 \Rightarrow QA^{\ell+1}b = 0 \text{ (Since } Q \geqslant 0). \end{aligned}$$

This completes the proof of Statements *(2)* and *(3)* for $\ell \in \{0, 1, \ldots, \mathtt{n_f} - 2\}$.

In what follows we complete the proof of Statement *(3)* by proving the identity for the case $\ell = \mathtt{n_f} - 1$. $\widehat{A}^{\mathtt{n_f}-1}\widehat{b} = \left[\begin{smallmatrix} A & 0 \\ -Q & -A^T \end{smallmatrix}\right] \left[\begin{smallmatrix} A^{\mathtt{n_f}-2}b \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} A^{\mathtt{n_f}-1}b \\ 0 \end{smallmatrix}\right]$ (Since $QA^{\mathtt{n_f}-2}b = 0$ from Statement *(2)* of this lemma). Similarly,

$$\begin{aligned} \widehat{c}\widehat{A}^{\mathtt{n_f}-1} &= \begin{bmatrix} (-1)^{\mathtt{n_f}-1} (A^{\mathtt{n_f}-2}b)^T Q \; (-1)^{\mathtt{n_f}-1} (A^{\mathtt{n_f}-1}b)^T \end{bmatrix} \\ &= \begin{bmatrix} 0 \; (-1)^{\mathtt{n_f}-1} (A^{\mathtt{n_f}-1}b)^T \end{bmatrix}. \end{aligned}$$

This completes the proof of Statement *(3)* of this lemma. □

**Lemma 4.6** *Let $K$ be any solution of the singular LQR LMI* (2.4) *with* $\mathtt{degdet}(sE - H) = 2\mathtt{n_s}$ *and* $\mathtt{n_f} := \mathtt{n} - \mathtt{n_s}$*, where* $(E, H)$ *are as defined in Theorem 4.1. Then, for any* $\alpha \in \{0, 1, \ldots, \mathtt{n_f} - 1\}$*,* $KA^{\alpha}b = 0$.

*Proof:* We prove this using induction and Lemma 4.5.
**Base case:** ($\alpha = 0$) Since $K$ is a solution of the LQR LMI (2.4), $Kb = 0$ is trivially true.
**Inductive step:** Suppose $\alpha \leqslant \mathtt{n_f} - 1$. Assume $KA^{(\alpha-1)}b = 0$, we show that $KA^{\alpha}b = 0$ . Pre- and post-multiplying $\mathcal{L}(K) := A^TK + KA + Q$ by $(A^{(\alpha-1)}b)^T$ and $A^{(\alpha-1)}b$, respectively, we get $(A^{(\alpha-1)}b)^T\mathcal{L}(K)(A^{(\alpha-1)}b) \geqslant 0$. Opening the brackets and using the inductive hypothesis this inequality becomes $(A^{(\alpha-1)}b)^TQ(A^{(\alpha-1)}b) \geqslant 0$. Further, using Statement *(2)* of Lemma (4.5) in this inequality, we get $\left(A^{(\alpha-1)}b\right)^T Q \left(A^{(\alpha-1)}b\right) = 0 \Rightarrow \mathcal{L}(K)A^{(\alpha-1)}b = 0$ (Since $\mathcal{L}(K) \geqslant 0$). Therefore, $(A^TK+KA+Q)A^{(\alpha-1)}b = 0 \Rightarrow A^TKA^{(\alpha-1)}b + KA^{\alpha}b + QA^{(\alpha-1)}b = 0$. Using inductive hypothesis and Statement *(2)* of Lemma 4.5, we therefore have $KA^{\alpha}b = 0$. □

Now that we have developed all the crucial results required to prove Theorem 4.1, in the ensuing section we prove Theorem 4.1.

### 4.3 Proof of Theorem 4.1

*Proof of Statement* (1) *of Theorem 4.1:* Partition $W =: \left[\begin{smallmatrix} W_1 \\ W_2 \end{smallmatrix}\right]$, where $W_1, W_2 \in \mathbb{R}^{\mathtt{n}\times\mathtt{n_f}}$. Using Statement *(3)* of Lemma 4.5, it is evident that

$$W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} \widehat{b} \ \widehat{A}\widehat{b} \cdots \ \widehat{A}^{\mathtt{n_f}-1}\widehat{b} \end{bmatrix} = \begin{bmatrix} b \ Ab \cdots A^{\mathtt{n_f}-1}b \\ 0 \ 0 \ \cdots \ 0 \end{bmatrix} \Rightarrow \begin{cases} W_1 = \begin{bmatrix} b \ Ab \cdots \ A^{\mathtt{n_f}-1}b \end{bmatrix}, \\ W_2 = 0_{\mathtt{n},\mathtt{n_f}}. \end{cases} \tag{4.20}$$

Therefore, $X_\Lambda = \begin{bmatrix} X_\Lambda \ W \end{bmatrix} = \begin{bmatrix} V_{1\Lambda} & W_1 \\ V_{2\Lambda} & 0_{\mathtt{n},\mathtt{n_f}} \end{bmatrix} = \begin{bmatrix} X_{1\Lambda} \\ X_{2\Lambda} \end{bmatrix}$. Then, we need to prove that $X_{1\Lambda} = \begin{bmatrix} V_{1\Lambda} \ W_1 \end{bmatrix}$ is invertible.

Note that since $V_{1\Lambda}$ is full column-rank (Theorem 4.2), there exists $F \in \mathbb{R}^{1\times\mathtt{n}}$ such that $V_{3\Lambda} = FV_{1\Lambda}$. Thus, from equation (4.1), $(A+bF)V_{1\Lambda} = V_{1\Lambda}\Gamma$. Define $W_{\mathtt{1F}} := \left[ b \ (A+bF)b \ \cdots \ (A+bF)^{\mathtt{n_f}-1}b \right]$. Then, clearly $\mathtt{img}\, W_1 = \mathtt{img}\, W_{\mathtt{1F}}$. Since $\Sigma$ is controllable, $W_1$ is full column-rank $\Leftrightarrow$ $W_{\mathtt{1F}}$ is also full column-rank. Thus, proving $X_{1\Lambda}$ is invertible is equivalent to proving $\widetilde{X}_{1\Lambda} := \left[ V_{1\Lambda} \ W_{\mathtt{1F}} \right]$ is invertible.

Now, we extend the columns of $V_{1\Lambda}$ to form a basis of $\mathbb{R}^{\mathtt{n}}$, say $\mathbb{B}$. Without loss of generality, we assume that the matrices $A, b$ are represented in the basis $\mathbb{B}$. Since $V_{1\Lambda}$ is $(A, b)$-invariant, in the new basis $(A+bF)$ must have the following structure $A + bF = \left[\begin{smallmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{smallmatrix}\right]$, where $\bar{A}_{11} \in \mathbb{R}^{\mathtt{n_s}\times\mathtt{n_s}}$ and $\bar{A}_{22} \in \mathbb{R}^{(\mathtt{n}-\mathtt{n_s})\times(\mathtt{n}-\mathtt{n_s})}$. Conforming to the partition in $A + bF$, we partition $b =: \left[\begin{smallmatrix} \bar{b}_1 \\ \bar{b}_2 \end{smallmatrix}\right]$. Note $V_{1\Lambda}$ in the basis $\mathbb{B}$ is of the form $\left[\begin{smallmatrix} I_{\mathtt{n_s}} \\ 0 \end{smallmatrix}\right]$. Further, $W_{\mathtt{1F}}$ in this new basis $\mathbb{B}$ has the following structure $W_{\mathtt{1F}} = \left[\begin{smallmatrix} \bar{b}_1 & \star & \cdots & \star \\ \bar{b}_2 & \bar{A}_{22}\bar{b}_2 & \cdots & \bar{A}^{\mathtt{n_f}-1}\bar{b}_2 \end{smallmatrix}\right]$, where $\star$ are suitable matrices with elements from $\mathbb{R}$. Since the system is controllable, we have $(A, b)$

controllable $\Leftrightarrow (A+bF, b)$ controllable $\Rightarrow (\bar{A}_{22}, \bar{b}_2)$ is controllable. Therefore, $T := \left[\bar{b}_2 \; \bar{A}_{22}\bar{b}_2 \; \cdots \; \bar{A}^{\mathtt{n_f}-1}\bar{b}_2\right] \in \mathbb{R}^{\mathtt{n_f}\times\mathtt{n_f}}$ is a nonsingular matrix. Now, note that the matrix $\widetilde{X}_{1\Lambda} = \left[V_{1\Lambda} \; W_{1\mathrm{F}}\right]$ in the basis $\mathbb{B}$ takes the form $\left[\begin{smallmatrix} I_{\mathtt{n_s}} & \star \\ 0 & T\end{smallmatrix}\right]$. Thus, $\widetilde{X}_{1\Lambda}$ is a block upper-triangular matrix with the diagonal blocks $I_{\mathtt{n_s}}$ and $T$ being nonsingular. Therefore, $\widetilde{X}_{1\Lambda}$ is invertible and hence $X_{1\Lambda}$ is invertible. □

*Proof of Statement* (2) *of Theorem 4.1:* To prove $X_{2\Lambda}X_{1\Lambda}^{-1} = (X_{2\Lambda}X_{1\Lambda}^{-1})^T$ is equivalent to proving $X_{1\Lambda}^T X_{2\Lambda} = X_{2\Lambda}^T X_{1\Lambda}$. Hence instead of proving $X_{2\Lambda}X_{1\Lambda}^{-1} = (X_{2\Lambda}X_{1\Lambda}^{-1})^T$ we prove that $X_{1\Lambda}^T X_{2\Lambda} - X_{2\Lambda}^T X_{1\Lambda} = 0$. Now, using equation (4.20) to evaluate $X_{1\Lambda}^T X_{2\Lambda} - X_{2\Lambda}^T X_{1\Lambda}$, we get

$$\begin{aligned} X_{1\Lambda}^T X_{2\Lambda} - X_{2\Lambda}^T X_{1\Lambda} &= \begin{bmatrix} V_{1\Lambda}^T \\ W_1^T \end{bmatrix} \left[V_{2\Lambda} \; 0_{\mathtt{n},\mathtt{n_f}}\right] - \begin{bmatrix} V_{2\Lambda}^T \\ 0_{\mathtt{n_f},\mathtt{n}} \end{bmatrix} \left[V_{1\Lambda} \; W_1\right] \\ &= \begin{bmatrix} V_{1\Lambda}^T V_{2\Lambda} - V_{2\Lambda}^T V_{1\Lambda} & -V_{2\Lambda}^T W_1 \\ W_1^T V_{2\Lambda} & 0_{\mathtt{n_f},\mathtt{n_f}} \end{bmatrix}. \end{aligned} \tag{4.21}$$

From Proposition 4.1, we have $V_{1\Lambda}^T V_{2\Lambda} = V_{2\Lambda}^T V_{1\Lambda}$. Hence, to prove $X_{1\Lambda}^T X_{2\Lambda} - X_{2\Lambda}^T X_{1\Lambda} = 0$, we need to prove that $V_{2\Lambda}^T W_1 = 0$. From equation (4.1), we have

$$-QV_{1\Lambda} - A^T V_{2\Lambda} = V_{2\Lambda}\Gamma \Rightarrow V_{1\Lambda}^T Q + V_{2\Lambda}^T A = -\Gamma^T V_{2\Lambda}^T. \tag{4.22}$$

We first prove that $V_{2\Lambda}^T A^k b = 0$ for $k \in \{0, 1, \ldots, \mathtt{n_f} - 1\}$ using mathematical induction.
**Base step:** $(k = 0)$ $V_{2\Lambda}^T b = 0$ follows from equation (4.1).
**Induction step:** Let $V_{2\Lambda}^T A^k b = 0$ for $k < \mathtt{n_f} - 1$. We prove that $V_{2\Lambda}^T A^{k+1} b = 0$. Post-multiplying equation (4.22) with $A^k b$ gives $V_{1\Lambda}^T Q A^k b + V_{2\Lambda}^T A^{k+1} b = -\Gamma^T V_{2\Lambda}^T A^k b$. Since $k < \mathtt{n_f} - 1$, we know that $QA^k b = 0$ (Lemma 4.5). This equation along with the inductive hypothesis imply that $V_{2\Lambda}^T A^{k+1} b = 0$. Hence, by mathematical induction, we have proved that $V_{2\Lambda}^T A^k b = 0$ for $k \in \{0, 1, 2, \ldots, \mathtt{n_f} - 1\}$. In other words, we proved that

$$V_{2\Lambda}^T \left[b \; Ab \cdots \; A^{\mathtt{n_f}-1} b\right] = 0 \Rightarrow V_{2\Lambda}^T W_1 = 0. \tag{4.23}$$

Thus, from equation (4.21), we have $X_{1\Lambda}^T X_{2\Lambda} = X_{2\Lambda}^T X_{1\Lambda}$. Therefore, $X_{2\Lambda}X_{1\Lambda}^{-1}$ is symmetric. □

*Proof of Statement* (3) *of Theorem 4.1:* Define $\mathcal{L}(K_{\mathtt{max}}) := A^T K_{\mathtt{max}} + K_{\mathtt{max}} A + Q$. Evaluating $X_{1\Lambda}^T \mathcal{L}(K_{\mathtt{max}}) X_{1\Lambda}$, we get

$$X_{1\Lambda}^T \mathcal{L}(K_{\mathtt{max}}) X_{1\Lambda} = \begin{bmatrix} V_{1\Lambda}^T \mathcal{L}(K_{\mathtt{max}}) V_{1\Lambda} & V_{1\Lambda}^T \mathcal{L}(K_{\mathtt{max}}) W_1 \\ W_1^T \mathcal{L}(K_{\mathtt{max}}) V_{1\Lambda} & W_1^T \mathcal{L}(K_{\mathtt{max}}) W_1 \end{bmatrix}. \tag{4.24}$$

Note that $K_{\mathtt{max}} V_{1\Lambda} = X_{2\Lambda} X_{1\Lambda}^{-1} V_{1\Lambda} = \left[V_{2\Lambda} \; W_2\right] \left[V_{1\Lambda} \; W_1\right]^{-1} V_{1\Lambda} = V_{2\Lambda}$, and $K_{\mathtt{max}} W_1 = \left[V_{2\Lambda} \; W_2\right] \left[V_{1\Lambda} \; W_1\right]^{-1} W_1 = W_2 = 0$ (From equation (4.20)). Using the fact that $K_{\mathtt{max}}V_{1\Lambda} = V_{2\Lambda}$ and evaluating $V_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})V_{1\Lambda}$ gives

$$\begin{aligned} V_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})V_{1\Lambda} &= V_{1\Lambda}(A^TK_{\mathtt{max}} + K_{\mathtt{max}}A + Q)V_{1\Lambda} \\ &= \begin{bmatrix} V_{2\Lambda}^T & -V_{1\Lambda}^T \end{bmatrix} \begin{bmatrix} A & 0 \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix} \end{aligned} \quad (4.25)$$

Using equation (4.1) and Proposition 4.1 in equation (4.25), we have

$$V_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})V_{1\Lambda} = \begin{bmatrix} V_{2\Lambda}^T & -V_{1\Lambda}^T \end{bmatrix} \left( \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix} \Gamma - \begin{bmatrix} b \\ 0 \end{bmatrix} V_{3\Lambda} \right) = -V_{2\Lambda}^T b V_{3\Lambda} = 0. \quad (4.26)$$

Using the fact that $K_{\mathtt{max}}V_{1\Lambda} = V_{2\Lambda}$ and $K_{\mathtt{max}}W_1 = 0$ to evaluate $V_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})W_1$ gives

$$\begin{aligned} V_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})W_1 &= V_{1\Lambda}^TA^TK_{\mathtt{max}}W_1 + V_{1\Lambda}^TK_{\mathtt{max}}AW_1 + V_{1\Lambda}^TQW_1 \\ &= V_{2\Lambda}^TAW_1 + V_{1\Lambda}^TQW_1. \end{aligned} \quad (4.27)$$

Post-multiplying equation (4.22) by $W_1$ and using it in equation (4.27) gives

$$V_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})W_1 = V_{1\Lambda}^TQW_1 + V_{2\Lambda}^TAW_1 = -\Gamma^TV_{2\Lambda}^TW_1. \quad (4.28)$$

From equation (4.23), we have $V_{2\Lambda}^TW_1 = 0$. Thus, $V_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})W_1 = 0$.

Since $K_{\mathtt{max}}W_1 = 0$, we must have

$$W_1^T\mathcal{L}(K_{\mathtt{max}})W_1 = W_1^TA^TK_{\mathtt{max}}W_1 + W_1^TK_{\mathtt{max}}AW_1 + W_1^TQW_1 = W_1^TQW_1.$$

Now, using Statement *(1)* of Lemma 4.5, we have

$$W_1^TQW_1 = \begin{bmatrix} 0_{(\mathtt{n_f}-1),(\mathtt{n_f}-1)} & 0_{(\mathtt{n_f}-1),1} \\ 0_{1,(\mathtt{n_f}-1)} & (A^{\mathtt{n_f}-1}b)^TQA^{\mathtt{n_f}-1}b \end{bmatrix}.$$

Thus, equation (4.24) becomes

$$X_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})X_{1\Lambda} = \begin{bmatrix} 0_{(\mathtt{n}-1),(\mathtt{n}-1)} & 0_{(\mathtt{n}-1),1} \\ 0_{1,(\mathtt{n}-1)} & (A^{\mathtt{n_f}-1}b)^TQA^{\mathtt{n_f}-1}b \end{bmatrix}. \quad (4.29)$$

Since $Q \geqslant 0$, we have $(A^{\mathtt{n_f}-1}b)^TQA^{\mathtt{n_f}-1}b \geqslant 0$. Therefore, $X_{1\Lambda}^T\mathcal{L}(K_{\mathtt{max}})X_{1\Lambda} \geqslant 0$. Since $X_{1\Lambda}$ is invertible, we must have $\mathcal{L}(K_{\mathtt{max}}) \geqslant 0$. Next using Statement *(2)* of this theorem and the fact that $V_{2\Lambda}^Tb = 0$ from equation (4.1) we have

$$K_{\mathtt{max}}b = X_{2\Lambda}X_{1\Lambda}^{-1}b = (X_{1\Lambda}^{-1})^TX_{2\Lambda}^Tb = (X_{1\Lambda}^{-1})^T \begin{bmatrix} V_{2\Lambda}^T \\ 0 \end{bmatrix} b = 0. \quad (4.30)$$

Thus, $K_{\mathtt{max}}$ is a solution of the LQR LMI (2.4). From equation (4.29), we infer that $\mathtt{rank}(\mathcal{L}(K_{\mathtt{max}}))$ is either 0 or 1.

Note that $\mathtt{rank}(\mathcal{L}(K_{\mathtt{max}})) = 0$ is equivalent to $\mathcal{L}(K_{\mathtt{max}}) = 0$, i.e., $A^TK_{\mathtt{max}} + K_{\mathtt{max}}A + Q = 0$ and $K_{\mathtt{max}}b = 0$. The equations $A^TK + KA + Q = 0$ and $Kb = 0$

are the continuous generalized constrained ARE (CGCARE) corresponding to the LQR Problem 2.1 (see [7], [8], [3] for more on CGCARE). Interestingly, from [3, Corollary 1] it is evident that a necessary condition for solvability of CGCARE is $\mathtt{det}(sE - H) = 0$. Since in this theorem $\mathtt{det}(sE - H) \neq 0$ by assumption, CGCARE is not solvable here. This implies that $\mathcal{L}(K) = 0$, i.e., $\mathtt{rank}(\mathcal{L}(K)) = 0$ is not possible in our case. Therefore, the minimum rank that can be attained by LQR LMI (2.4) is 1 and $\mathcal{L}(K_{\mathtt{max}})$ attains this rank. □

*Proof of Statement* (4) *and* (5) *of Theorem 4.1:* Note that proving Statement *(4)* of this theorem is equivalent to proving that $K - K_{\mathtt{max}} \leqslant 0$ for all $K$ that satisfies the LQR LMI (2.4). We prove this in two steps. First, we prove that for $V_{1\Lambda}$ defined in the theorem, $\Delta := V_{1\Lambda}(K - K_{\mathtt{max}})V_{1\Lambda}$ satisfies a suitable Lyapunov inequality (see equation (4.37) below). Then, using this Lyapunov inequality we finally show that $K - K_{\mathtt{max}} \leqslant 0$ for all $K$ that satisfies the LQR LMI (2.4).

**Step 1:** Note that for all $(x, u)$ that satisfies $\frac{d}{dt}x = Ax + bu$, evaluation of $\frac{d}{dt}(x^T Kx) + x^T Qx$ results in the following equation:

$$\begin{aligned}\frac{d}{dt}(x^T Kx) + x^T Qx &= \dot{x}^T Kx + x^T K\dot{x} + x^T Qx \\ &= (Ax + bu)^T Kx + x^T K(Ax + bu) + x^T Qx \\ &= \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^T K + KA + Q & Kb \\ b^T K & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}, \text{ for all } t \in \mathbb{R}.\end{aligned} \tag{4.31}$$

Since $K$ is a solution of the LQR LMI (2.4), using the LMI $\left[\begin{smallmatrix} A^T K+KA+Q & Kb \\ b^T K & 0 \end{smallmatrix}\right] \geqslant 0$ in equation (4.31), we have

$$\frac{d}{dt}(x^T Kx) + x^T Qx = \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^T K + KA + Q & Kb \\ b^T K & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \geqslant 0, \text{ for all } t \in \mathbb{R}. \tag{4.32}$$

From equation (4.1), we know that $AV_{1\Lambda} + bV_{3\Lambda} = V_{1\Lambda}\Gamma$. Further, since $V_{1\Lambda}$ is full column-rank (Theorem 4.2), we infer that there exists $F \in \mathbb{R}^{1\times \mathtt{n_s}}$ such that $FV_{1\Lambda} = V_{3\Lambda}$. Therefore, we have $(A + bF)V_{1\Lambda} = V_{1\Lambda}\Gamma$. Thus, corresponding to an initial condition $x_0 = V_{1\Lambda}\beta$, where $\beta \in \mathbb{R}^{\mathtt{n_s}\times 1}$, $\bar{x}_{\mathtt{s}} := V_{1\Lambda}e^{\Gamma t}\beta$, $\bar{u}_{\mathtt{s}} := FV_{1\Lambda}e^{\Gamma t}\beta$ must satisfy $\frac{d}{dt}x = Ax + bu$. Using $\bar{x}_{\mathtt{s}}$ in equation (4.32), we have

$$\frac{d}{dt}(\bar{x}_{\mathtt{s}}^T K\bar{x}_{\mathtt{s}}) + \bar{x}_{\mathtt{s}}^T Q\bar{x}_{\mathtt{s}} \geqslant 0 \Rightarrow \frac{d}{dt}(\bar{x}_{\mathtt{s}}^T K\bar{x}_{\mathtt{s}}) \geqslant -\bar{x}_{\mathtt{s}}^T Q\bar{x}_{\mathtt{s}}, \text{ for all } t \in \mathbb{R}. \tag{4.33}$$

Note that $\dot{\bar{x}}_{\mathtt{s}} = V_{1\Lambda}\Gamma e^{\Gamma t}\beta = (A + bF)V_{1\Lambda}e^{\Gamma t}\beta$ (Since $(A + bF)V_{1\Lambda} = V_{1\Lambda}\Gamma$). Since $K_{\mathtt{max}}$ is a solution of the LQR LMI (2.4), using the fact that $K_{\mathtt{max}}b = 0$

we have for all $t \in \mathbb{R}$.

$$\begin{aligned}\frac{d}{dt}\left(\bar{x}_{\mathtt{s}}^T K_{\mathtt{max}} \bar{x}_{\mathtt{s}}\right) + \bar{x}_{\mathtt{s}}^T Q \bar{x}_{\mathtt{s}} &= \dot{\bar{x}}_{\mathtt{s}}^T K_{\mathtt{max}} \bar{x}_{\mathtt{s}} + \bar{x}_{\mathtt{s}}^T K_{\mathtt{max}} \dot{\bar{x}}_{\mathtt{s}} + \bar{x}_{\mathtt{s}}^T Q \bar{x}_{\mathtt{s}} \\ &= \beta^T e^{\Gamma^T t} V_{1\Lambda}^T (A+bF)^T K_{\mathtt{max}} V_{1\Lambda} e^{\Gamma t}\beta \\ &\quad + \beta^T e^{\Gamma^T t} V_{1\Lambda}^T K_{\mathtt{max}} (A+bF) V_{1\Lambda} e^{\Gamma^T t}\beta + \bar{x}_{\mathtt{s}}^T Q \bar{x}_{\mathtt{s}} \\ &= \beta^T e^{\Gamma^T t} V_{1\Lambda}^T (A^T K_{\mathtt{max}} + K_{\mathtt{max}} A + Q) V_{1\Lambda} e^{\Gamma t}\beta. \end{aligned} \tag{4.34}$$

From equation (4.26), it is evident that the right hand side of equation (4.34) is 0. Therefore,

$$\frac{d}{dt}\left(\bar{x}_{\mathtt{s}}^T K_{\mathtt{max}} \bar{x}_{\mathtt{s}}\right) = -\bar{x}_{\mathtt{s}}^T Q \bar{x}_{\mathtt{s}}, \text{ for all } t \in \mathbb{R}. \tag{4.35}$$

Subtracting equation (4.35) from inequality (4.33) gives

$$\frac{d}{dt}\left(\bar{x}_{\mathtt{s}}^T (K - K_{\mathtt{max}}) \bar{x}_{\mathtt{s}}\right) = \dot{\bar{x}}_{\mathtt{s}}^T (K - K_{\mathtt{max}}) \bar{x}_{\mathtt{s}} + \bar{x}_{\mathtt{s}}^T (K - K_{\mathtt{max}}) \dot{\bar{x}}_{\mathtt{s}} \geqslant 0, \text{ for all } t \in \mathbb{R}.$$

On using $\bar{x}_{\mathtt{s}} = V_{1\Lambda} e^{\Gamma t}\beta$ and $\dot{\bar{x}}_{\mathtt{s}} = V_{1\Lambda}\Gamma e^{\Gamma t}\beta$ in this inequality, we get

$$(V_{1\Lambda} e^{\Gamma t}\Gamma\beta)^T (K - K_{\mathtt{max}})(V_{1\Lambda} e^{\Gamma t}\beta) + (V_{1\Lambda} e^{\Gamma t}\beta)^T (K - K_{\mathtt{max}})(V_{1\Lambda} e^{\Gamma t}\Gamma\beta) \geqslant 0, \tag{4.36}$$

for all $t \in \mathbb{R}$. Since inequality (4.36) is true for all $t$, evaluating it at $t = 0$, in particular, we get $\beta^T(\Gamma^T V_{1\Lambda}^T (K - K_{\mathtt{max}}) V_{1\Lambda} + V_{1\Lambda}^T (K - K_{\mathtt{max}}) V_{1\Lambda}\Gamma)\beta = \beta^T(\Gamma^T \Delta + \Delta\Gamma)\beta \geqslant 0$, where $\Delta := V_{1\Lambda}^T (K - K_{\mathtt{max}}) V_{1\Lambda}$. Since this inequality is true for all $\beta \in \mathbb{R}^{\mathtt{n_s}\times 1}$, we have

$$\Gamma^T \Delta + \Delta\Gamma \geqslant 0, \text{ where } \Delta = V_{1\Lambda}^T (K - K_{\mathtt{max}}) V_{1\Lambda}. \tag{4.37}$$

This ends the first step of the proof.

**Step 2:** Note that since $X_{1\Lambda}$ is nonsingular, proving $K - K_{\mathtt{max}} \leqslant 0$ is equivalent to proving that $X_{1\Lambda}^T (K - K_{\mathtt{max}}) X_{1\Lambda} \leqslant 0$ (by congruence transformation property). Hence, we prove $X_{1\Lambda}^T (K - K_{\mathtt{max}}) X_{1\Lambda} \leqslant 0$ in the sequel.

Note that $X_{1\Lambda} = \begin{bmatrix} V_{1\Lambda} \; W_1 \end{bmatrix}$, where $W_1$ is as defined in equation (4.20). On evaluating $X_{1\Lambda}^T (K - K_{\mathtt{max}}) X_{1\Lambda}$, we therefore have

$$X_{1\Lambda}^T (K - K_{\mathtt{max}}) X_{1\Lambda} = \begin{bmatrix} V_{1\Lambda}^T (K - K_{\mathtt{max}}) V_{1\Lambda} & V_{1\Lambda}^T (K - K_{\mathtt{max}}) W_1 \\ W_1^T (K - K_{\mathtt{max}}) V_{1\Lambda} & W_1^T (K - K_{\mathtt{max}}) W_1 \end{bmatrix}. \tag{4.38}$$

Since $W_1 = \begin{bmatrix} b \; Ab \cdots A^{\mathtt{n_f}-1} b \end{bmatrix}$ (equation (4.20)), we have from Lemma 4.6, $KW_1 = 0$ and $K_{\mathtt{max}} W_1 = 0$. Therefore, $(K - K_{\mathtt{max}})W_1 = 0$. Thus, from equation (4.38) it follows that

$$X_{1\Lambda}^T (K - K_{\mathtt{max}}) X_{1\Lambda} = \begin{bmatrix} V_{1\Lambda}^T (K - K_{\mathtt{max}}) V_{1\Lambda} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \Delta & 0 \\ 0 & 0 \end{bmatrix}, \tag{4.39}$$

Since $\sigma(\Gamma) \subsetneq \mathbb{C}_-$, from equation (4.37), we have $\Delta \leqslant 0$. Using this negative-semidefiniteness property of $\Delta$ in equation (4.39), we infer that $X_{1\Lambda}^T(K - K_{\texttt{max}})X_{1\Lambda} \leqslant 0 \Leftrightarrow K - K_{\texttt{max}} \leqslant 0$. This completes the proof of Statement *(4)* of the theorem.

Note that 0 is a solution of the LQR LMI (2.4). Thus, from Statement *(4)* of Theorem (4.1) we must have $0 \leqslant K_{\texttt{max}}$. Thus, Statement *(5)* of the theorem is proved. □

Now that we have proved Theorem 4.1, using the main results of this paper, we present a few corollaries next. These corollaries reaffirm some of the well-known findings present in the literature: see [14], [9], [10], [21].

Recall from Statement *(1)* of Theorem 4.1 and equation (4.20) we know that $X_\Lambda = \begin{bmatrix} X_{1\Lambda} \\ X_{2\Lambda} \end{bmatrix} = \begin{bmatrix} V_{1\Lambda} & W_1 \\ V_{2\Lambda} & 0_{\texttt{n},\texttt{n}_\texttt{f}} \end{bmatrix}$. Thus, we have $X_{1\Lambda} = \begin{bmatrix} V_{1\Lambda} \; W_1 \end{bmatrix}$ and $X_{2\Lambda} = \begin{bmatrix} V_{2\Lambda} \; 0 \end{bmatrix}$. Further, from Lemma 4.4, we know that $\texttt{img} \begin{bmatrix} V_{\texttt{g}} & V_{1\texttt{e}} \\ 0_{\texttt{n},\texttt{g}} & V_{2\texttt{e}} \end{bmatrix} = \texttt{img} \begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$ and $\begin{bmatrix} V_{\texttt{g}} \; V_{1\texttt{e}} \end{bmatrix}$ is full column-rank. Hence, the matrix $X_{1\Lambda}$, without loss of generality, is given by $X_{1\Lambda} = \begin{bmatrix} V_{\texttt{g}} \; V_{1\texttt{e}} \; W_1 \end{bmatrix}$ and the corresponding $X_{2\Lambda}$ matrix is then $X_{2\Lambda} = \begin{bmatrix} 0_{\texttt{n},\texttt{g}} \; V_{2\texttt{e}} \; 0_{\texttt{n},\texttt{n}_\texttt{f}} \end{bmatrix}$.

Since $X_{1\Lambda}$ is invertible (Statement *(1)* of Theorem 4.1), it is evident that the columns of $X_{1\Lambda}$ can be assumed to be a basis for $\mathbb{R}^{\texttt{n}}$. Hence, any initial condition $x_0$ of the system $\Sigma$ can be decomposed as

$$x_0 =: x_{\texttt{gs}} + x_{\texttt{es}} + x_{0\texttt{f}}, \text{ where } x_{\texttt{gs}} \in \texttt{img}\, V_{\texttt{g}} =: \mathcal{V}_{\texttt{g}}, x_{\texttt{es}} \in \texttt{img}\, V_{1\texttt{e}} =: \mathcal{V}_{\texttt{e}}, \\ x_{0\texttt{f}} \in \texttt{img}\, W_1 =: \mathcal{W}. \tag{4.40}$$

Now we have the following corollary.

**Corollary 4.1** *Consider the LQR Problem 2.1 and let $K_{\texttt{max}}$ be the maximal rank-minimizing solution of the corresponding LQR LMI* (2.4). *Assume $x_0 =: x_{\texttt{gs}} + x_{\texttt{es}} + x_{0\texttt{f}}$ to be an initial condition of the system $\Sigma$ as defined in equation* (4.40). *Then the following statements hold:*

*(1)* $x_{\texttt{gs}}^T K_{\texttt{max}} x_{\texttt{gs}} = 0$.

*(2)* $x_{0\texttt{f}}^T K_{\texttt{max}} x_{0\texttt{f}} = 0$.

*(3) The optimal cost of the LQR problem is* $x_{\texttt{es}}^T K_{\texttt{max}} x_{\texttt{es}}$.

*Proof: (1):* Let $x_{\texttt{gs}} := V_{\texttt{g}}\alpha$, where $\alpha \in \mathbb{R}^{\texttt{g}\times 1}$. Note that

$$K_{\texttt{max}} x_{\texttt{gs}} = K_{\texttt{max}} V_{\texttt{g}}\alpha = X_{2\Lambda} X_{1\Lambda}^{-1} V_{\texttt{g}}\alpha = \begin{bmatrix} 0_{\texttt{n},\texttt{g}} \; V_{2\texttt{e}} \; 0_{\texttt{n},\texttt{f}} \end{bmatrix} \begin{bmatrix} V_{\texttt{g}} \; V_{1\texttt{e}} \; W_1 \end{bmatrix}^{-1} V_{\texttt{g}}\alpha = 0. \tag{4.41}$$

Therefore, $x_{\mathtt{gs}}^T K_{\mathtt{max}} x_{\mathtt{gs}} = \alpha^T V_{\mathtt{g}}^T K_{\mathtt{max}} V_{\mathtt{g}} \alpha = 0$.
*(2):* Let $x_{0\mathtt{f}} := W_1 \beta$, where $\beta \in \mathbb{R}^{\mathtt{n_f} \times 1}$. Note that

$$K_{\mathtt{max}} x_{0\mathtt{f}} = K_{\mathtt{max}} W_1 \beta = \begin{bmatrix} 0_{\mathtt{n},\mathtt{g}} \; V_{2\mathtt{e}} \; 0_{\mathtt{n},\mathtt{f}} \end{bmatrix} \begin{bmatrix} V_{\mathtt{g}} \; V_{1\mathtt{e}} \; W_1 \end{bmatrix}^{-1} W_1 \beta = 0. \tag{4.42}$$

Therefore, $x_{0\mathtt{f}}^T K_{\mathtt{max}} x_{0\mathtt{f}} = \beta^T W_1^T K_{\mathtt{max}} W_1 \beta = 0$.
*(3):* From [19], it is known that the optimal cost corresponding to the LQR Problem 2.1 is given by $x_0^T K_{\mathtt{max}} x_0$, where $K_{\mathtt{max}}$ is the maximal rank-minimizing solution of the LQR LMI (2.4). Hence, using equations (4.41) and (4.42) and evaluating the optimal cost for the LQR Problem 2.1, we have

$$x_0^T K_{\mathtt{max}} x_0 = (x_{\mathtt{gs}} + x_{\mathtt{es}} + x_{0\mathtt{f}})^T K_{\mathtt{max}} (x_{\mathtt{gs}} + x_{\mathtt{es}} + x_{0\mathtt{f}}) = x_{\mathtt{es}}^T K_{\mathtt{max}} x_{\mathtt{es}}. \tag{4.43}$$

This completes the proof of the corollary. □

From Corollary 4.1 it is evident that if the initial condition of the system is from $\mathcal{W}$ or $\mathcal{V}_{\mathtt{g}}$ then the cost incurred by the system is zero. Thus, the optimal cost of an LQR problem depends only on the maximal rank-minimizing solution of the corresponding LQR LMI and the projection of the initial condition of the system onto the subspace $\mathcal{V}_{\mathtt{e}}$.

Next we look at a special case of LQR Problems when the system admits the zero matrix as the only solution to the corresponding LQR LMI.

**Corollary 4.2** *Consider the singular LQR Problem 2.1 with assumptions as given in Theorem 4.1. Suppose* $\mathtt{dim}\,(\sup \mathcal{B}_{\Sigma}) = \mathtt{n_s}$*, where* $\mathcal{B}_{\Sigma}$ *is as defined in Lemma 4.3. Then,* $K_{\mathtt{max}} = 0_{\mathtt{n},\mathtt{n}}$.

*Proof:* Since $\mathtt{dim}\,(\sup \mathcal{B}_{\Sigma}) = \mathtt{n_s}$ and $\mathtt{dim}\,\mathcal{O}_{wg} = \mathtt{n_s}$, from Lemma 4.4 it is evident that $\mathtt{img}\begin{bmatrix} V_{\mathtt{g}} \\ 0_{\mathtt{n},\mathtt{n_s}} \end{bmatrix} = \mathtt{img}\begin{bmatrix} V_{1\Lambda} \\ V_{2\Lambda} \end{bmatrix}$. Therefore, $V_{2\Lambda} = 0_{\mathtt{n},\mathtt{n_s}}$. Further, from equation (4.20) we have $W_2 = 0$. Therefore, $X_{2\Lambda} = 0$ and hence using Theorem 4.1, we must have $K_{\mathtt{max}} = 0_{\mathtt{n},\mathtt{n}}$. □

The next corollary states that if the transfer function induced by the cost-matrix $Q$ and the system $\Sigma$ is minimum-phase[1], then the optimal cost of the corresponding LQR problem is zero.

**Corollary 4.3** *Consider the singular LQR Problem 2.1 with* $\mathtt{rank}\, Q = 1$ *and* $(Q, A)$ *observable. Let* $c \in \mathbb{R}^{1 \times \mathtt{n}}$ *be such that* $Q = c^T c$*. Define* $G(s) := c(sI_{\mathtt{n}} - A)^{-1} b$*. If the system* $G(s)$ *is minimum-phase, then the optimal cost of the LQR problem is zero.*

*Proof:* Recall $\widehat{A}, \widehat{b}, \widehat{c}$ are as defined in equation (2.7). Define $\mathtt{det}(sI_{\mathtt{n}} - A) =: d(s)$. Therefore, $\mathtt{det}(sI_{2\mathtt{n}} - \widehat{A}) = d(s)d(-s)$. Further, since the system is $(A, b)$ controllable and $(Q, A)$ observable, there exists a real-polynomial $n(s)$ such that $G(s) = \frac{n(s)}{d(s)}$ with $n(s)$ and $d(s)$ are coprime.

[1] A transfer function $G(s) = \frac{n(s)}{d(s)}$ is said to be minimum-phase if $\mathtt{roots}(n(s)) \subsetneq \mathbb{C}_-$.

Note that $\texttt{det}(sE-H) = \texttt{det}\begin{bmatrix} sI_{2\mathtt{n}}-\widehat{A} & -\widehat{b} \\ -\widehat{c} & 0 \end{bmatrix} = \widehat{c}(sI_{2\mathtt{n}}-\widehat{A})^{-1}\widehat{b}\times\texttt{det}(sI_{2\mathtt{n}}-\widehat{A}) =: p(s)$.

Further, by simple multiplication it can be seen that

$$G(-s)G(s) = \widehat{c}(sI_{2\mathtt{n}} - \widehat{A})^{-1}\widehat{b} \Rightarrow \frac{n(-s)n(s)}{d(-s)d(s)} = \frac{p(s)}{d(-s)d(s)},$$

Therefore, $p(s) = n(-s)n(s)$. Since $|\sigma(E,H)| = 2\mathtt{n_s} \Rightarrow |\texttt{roots}\, p(s)| = 2\mathtt{n_s} \Rightarrow |\texttt{roots}\,(n(s))| = \mathtt{n_s}$. Since $G(s)$ is minimum-phase, $\texttt{roots}(n(s)) \subsetneq \mathbb{C}_-$.

Consider the system $\frac{d}{dt}x = Ax + bu$ and $y := cx$. Note that this is a SISO system which is $(A,b)$ controllable and $(Q,A)$ observable $\Rightarrow (c,A)$ observable. Therefore, as discussed in Section 2.3, $\sigma((A+bF)|_{\sup\mathcal{B}_\Sigma}) = \texttt{roots}\,\texttt{num}\,(G(s))$. Therefore, $\texttt{dim}\,(\sup\mathcal{B}_\Sigma) = \mathtt{n_s}$. Hence, by Corollary 4.2 we have $K_{\max} = 0_{\mathtt{n},\mathtt{n}} \Rightarrow$ the optimal cost is zero. □

Note that Corollary 4.3, albeit for single-input systems, corroborates the findings on minimum-phase systems in [9, Theorem 2] and [14].

## 5 Application of the main result and comparison with results in literature

In this section we present an algorithm to design proportional-derivative (PD) state-feedback controllers that solve singular LQR problems and tabulate the optimal trajectories related to Problem 2.1. Further, we also elaborate on the restrictions of the deflating subspace method presented in [17] with the help of suitable examples.

### 5.1 Algorithm to compute PD-controllers for single-input singular LQR problems

Using Theorem 4.1 and the results in [2], we now present an algorithm to design PD state-feedback controllers to solve a singular LQR problem.

**Algorithm 5.1** Algorithm to compute the gain matrices (proportional and derivative) to solve a singular LQR problem

**Input:** $(A,b,Q)$ matrices corresponding to a singular LQR problem.
**Output:** $K = K^T \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$.

1: Construct $(E,H)$ as defined in equation (2.5) and compute $\mathtt{n_s} = \{\texttt{degdet}(sE-H)\}/2$.
2: Use generalized real-Schur decomposition algorithm on $(E,H)$ to compute basis of eigenspace corresponding to all eigenvalues of $(E,H)$ in $\mathbb{C}_-$. Let columns of $V_{\mathtt{e}\Lambda} \in \mathbb{R}^{(2\mathtt{n}+1)\times\mathtt{n_s}}$ be the basis.

3: Partition $V_{\mathtt{e}\Lambda} =: \mathtt{col}(V_{1\Lambda}, V_{2\Lambda}, V_{3\Lambda})$ where $V_{1\Lambda}, V_{2\Lambda} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n_s}}, V_{3\Lambda} \in \mathbb{R}^{\mathtt{n_s}}$ and construct the matrix $V_\Lambda := \mathrm{col}(V_{1\Lambda}, V_{2\Lambda})$.
4: **if** $\mathtt{n_s} \neq \mathtt{n}$ **then**
5: Compute $\mathtt{n_f} = \mathtt{n} - \mathtt{n_s}$ and construct

$$W := \begin{bmatrix} \widehat{b} & \widehat{A}\widehat{b} & \widehat{A}^2\widehat{b} & \cdots & \widehat{A}^{\mathtt{n_f}-1}\widehat{b} \end{bmatrix} \in \mathbb{R}^{2\mathtt{n}\times\mathtt{n_f}}.$$

6: Construct $X := \begin{bmatrix} V_\Lambda & W \end{bmatrix} \in \mathbb{R}^{2\mathtt{n}\times\mathtt{n}}$
7: **else**
8: Construct $X := V_\Lambda \in \mathbb{R}^{2\mathtt{n}\times\mathtt{n}}$
9: **end if**
10: Partition $X$ as $X =: \begin{bmatrix} X_{1\Lambda} \\ X_{2\Lambda} \end{bmatrix}$ where $X_{1\Lambda}, X_{2\Lambda} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$.
11: Compute $K := X_{2\Lambda} X_{1\Lambda}^{-1} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$.
12: Construct

$$F_{\mathtt{p}} := \begin{bmatrix} V_{3\Lambda} & f_0 & f_1 & \cdots & f_{\mathtt{n_f}-1} \end{bmatrix} X_{1\Lambda}^{-1} \qquad F_{\mathtt{d}} := \begin{bmatrix} 0_{1,\mathtt{n_s}} & 1 & -f_0 & \cdots & -f_{\mathtt{n_f}-2} \end{bmatrix} X_{1\Lambda}^{-1},$$

where $f_0, f_1, \cdots, f_{\mathtt{n_f}-1} \in \mathbb{R}$ such that $\mathtt{det}\,(s(I_{\mathtt{n}} - bF_{\mathtt{d}}) - (A + bF_{\mathtt{p}})) \neq 0$.

Step 1 to Step 11 of Algorithm 5.1 is based on Theorem 4.1. The final step (Step 12) of the algorithm involves computation of matrices $F_{\mathtt{p}}$ and $F_{\mathtt{d}}$. On using these gain matrices in a feedback law given by $u = F_{\mathtt{p}} x + F_{\mathtt{d}} \frac{d}{dt} x$, one gets a control law that solves the corresponding singular LQR problem. This result has been proved in [2] (see Theorems 2 and 3 there).

## 5.2 Comparison with deflating subspace method

In this section we present a few instructive examples that demonstrate the restrictions of the results in [17]. For the ease of reference, we present the LQR problem dealt with in [17] next.

**Problem 5.2** *Consider a differential-algebraic system of the form:*

$$E\frac{d}{dt}x = Ax + Bu, \;\; \textit{where } A, E \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}} \textit{ and } B \in \mathbb{R}^{\mathtt{n}\times\mathtt{m}}. \tag{5.1}$$

*Find an input $u$ that minimizes the cost-functional*

$$\int_0^\infty \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt$$

*subject to $(x, u) \in \mathfrak{B}_{[E,A,B]}$ with $Ex(0) = Ex_0$ and $\lim_{t\to\infty} Ex(t) = 0$, where*

$$\mathfrak{B}_{[E,A,B]} := \{(x, u) \in \mathcal{L}^2_{\mathtt{loc}}(\mathbb{R}, \mathbb{R}^{\mathtt{n}}) \times \mathcal{L}^2_{\mathtt{loc}}(\mathbb{R}, \mathbb{R}^{\mathtt{m}}) \mid E\frac{d}{dt}x \in \mathcal{L}^2_{\mathtt{loc}}(\mathbb{R}, \mathbb{R}^{\mathtt{n}}) \textit{ and } (x, u) \textit{ satisfies equation (5.1) at almost all } t \in \mathbb{R}\}.$$

Note that unlike Problem 1.1, the inputs and states in Problem 5.2 are considered to be from the space of locally square-integrable functions. Further Problem 5.2 also imposes restrictions on the initial conditions of the system. Such restrictions, specially the ones on the initial conditions, are not desirable for a system, since a state-space system can always have arbitrary initial conditions. In this paper we do not impose any such restrictions on the function space or initial conditions. In the next few examples we highlight the advantage of using the results in this paper by illustrating the restrictions of the results in [17] and demonstrating how our methods can overcome these restrictions.

In the first example we consider a single-integrator system. Such systems find widespread application in the field of multi-agent systems. The same example had also been used in [10] to demonstrate the importance of $\mathfrak{C}_{\text{imp}}$ functions in singular LQR problems. The next example illustrates that the singular LQR problem admits no solution if the results in [17] are used. However, the same problem does admit a solution using the results in this paper.

*Example 5.3 Consider the system* $\frac{d}{dt}x = u$*. Let the performance index be* $J := \int_0^\infty x^2(t)dt$*.*

*Corresponding to an arbitrary non-zero initial condition* $\alpha \in \mathbb{R}$*, the state of the single-integrator system is given by* $x = \alpha + \int_0^t u(\tau)d\tau$*. It is evident that* $J$ *can be made arbitrarily small using an input* $u \in \mathcal{L}^2_{\texttt{loc}}(\mathbb{R}, \mathbb{R})$*. However,* $J$ *can never be made zero unless we chose* $u = -\alpha\delta \notin \mathcal{L}^2_{\texttt{loc}}(\mathbb{R}, \mathbb{R})$*. Hence, by the theory provided in [17] the problem does not have an optimal solution. However, using the theory developed in this paper one can find the optimal input to be* $-\alpha\delta$*.*

Note that the state-space system in Example 5.3 can trivially be written as a differential-algebraic system as follows.

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{E} \frac{d}{dt} \begin{bmatrix} x \\ p \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}}_{A} \begin{bmatrix} x \\ p \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \tag{5.2}$$

Since all state-space systems can always be rewritten as differential-algebraic systems, from the example above it becomes clear that the consideration of differential-algebraic state-space systems as in [17] does not provide additional tools for solving this particular issue. Here, the restriction of inputs to $\mathcal{L}^2_{\texttt{loc}}(\mathbb{R}, \mathbb{R})$ prevents the theory of [17] from providing an optimal solution. A related and similarly significant restriction is posed by the consistency requirement on parts of the initial conditions of the underlying system [17]. We elaborate on this next.

Note that in Problem 5.2 the initial conditions corresponding to the system are subjected to the constraint $Ex(0) = Ex_0$ and the continuity of $Ex$. In the context of a state-space system with $E = I$, this transfers to the continuity of $x$ at $t = 0$ or, in other words, the initial conditions are restricted to be such that the states of the system do not exhibit any jump at $t = 0^+$. We explain next, with the help of an example, the implications of such a restriction on the initial conditions of a system in the context of an LQR problem.

*Example 5.4 Consider the following system:*

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

*Let the cost-functional be* $\int_0^\infty (x_1^2 + x_2^2)dt$.

*For this example, using Theorem 4.1 we have* $V_{1\Lambda} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. *Further, here* $W_1 = b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

*Using our results in [2, Theorem 1], the optimal input corresponding to an arbitrary initial condition* $x_0 = V_{1\Lambda}\beta + W_1\alpha = \begin{bmatrix} 1 \\ -1 \end{bmatrix}\beta + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\alpha$ *is given by* $u_{\texttt{opt}} = e^{-t}\beta - \delta\alpha$, *where* $\alpha, \beta \in \mathbb{R}$. *The optimal states would then be:* $\begin{bmatrix} x_{1\texttt{opt}} \\ x_{2\texttt{opt}} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-t}\beta$. *(see [2, Theorem 1] for method to compute optimal inputs).*

*For initial conditions with* $\alpha = 0$, *it is evident that* $x(0) = x_0$. *This satisfies the constraints on initial condition as given in Problem 5.2. However, if* $\beta = 0$ *and* $\alpha \neq 0$, *then we have* $x(0) = 0 \neq x_0$. *This is a violation of the constraint* $x(0) = x_0$.

*Note that if one imposes the constraint* $x(0) = x_0$ *as in Problem 5.2, then the optimal input would be* $u = e^{-t}\beta$ *and the corresponding optimal state would be* $\begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-t}\beta$. *This conforms with the results in [17].*

Note that from our results in [2] it follows that whenever the initial conditions of the system are from $\texttt{img}(V_{1\Lambda})$, the optimal inputs/states are exponential in nature. On the other hand, if the initial conditions are from $\texttt{img}(W_1)$, then the optimal inputs/states are impulsive. Thus, the restriction $x(0) = x_0$ in Problem 5.2, when rewritten in terms of the notions developed in this paper is equivalent to $x(0) \in \texttt{img}(V_{1\Lambda})$. Evidently, any singular LQR problem that admits a trivial $\texttt{img}(V_{1\Lambda})$, i.e., $\texttt{img}(V_{1\Lambda}) = \{0\}$, does not admit any optimal input according to the results in [17]. A typical example of such a system is Example 5.3. In case of Example 5.3, it can be verified that $\texttt{det}(sE - H) = 1$ and therefore $\texttt{n}_\texttt{s} = 0$ and $\texttt{n}_\texttt{f} = 1$. Thus, $\texttt{img}(V_{1\Lambda}) = \{0\}$ in Example 5.3. Hence, for such a singular LQR problem, there exists no locally square-integrable input that can attain the minimum cost of 0 for arbitrary initial condition.

Even if a singular LQR problem admits a non-trivial $\mathtt{img}(V_{1\Lambda})$, like that in Example 5.4, it is evident that Problem 5.2 in the state-space setting does not cater to all initial conditions; in particular it does not cater to the initial conditions in the non-trivial subspace $\mathtt{img}(W_1)$. In this respect, the results of this paper cover a problem class that exceeds the theory of [17], since the results in this paper consider *arbitrary* initial condition.

A fundamental advantage of our theory over [17] is evident from an engineering viewpoint. In [17] the authors show that an optimal control law corresponding to a singular LQR problem fulfills an implicit control law of the form $Px+Lu = 0$, where $P$ and $L$ are solutions of the Luré equations involved with the LQR LMI: see [17, Sec. 7] for the details. Thus, even if we allow for restrictions on initial conditions of the system as given in Problem 5.2, the implementability of the control law is a concern. We elaborate on this with the help of Example 5.4. Using Algorithm 5.1 we can compute the maximal stabilizing rank-minimizing solution of the corresponding LQR LMI to be $K_{\mathtt{max}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. The corresponding LQR LMI then becomes

$$\begin{bmatrix} A^T K_{\mathtt{max}} + K_{\mathtt{max}} A + Q & K_{\mathtt{max}} B \\ B^T K_{\mathtt{max}} & R \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$$

Hence, $P = \begin{bmatrix} 1 & 1 \end{bmatrix}$ and $L = 0$ and the control law proposed in [17] becomes $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 0u = 0 \Rightarrow x_1 + x_2 = 0$.

Note that the optimal states in Example 5.4 do satisfy $x_1 + x_2 = 0$. However, from an engineering perspective the implementation of this control law, in its current form, is not possible as no information about the input is divulged in such a law. In other words, this controller is not feedback implementable in its current form. On the other hand, application of Algorithm 5.1 based on Theorem 4.1 provides us with an explicit state-feedback control law of the form $u = F_{\mathtt{p}} x + F_{\mathtt{d}} \frac{d}{dt} x$. In case of Example 5.4 one of the control laws (on choosing $f_0 = -1$ in Step 12 of Algorithm 5.1) would be $u = \begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \end{bmatrix} \frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. This is a PD state-feedback control law and is implementable.

In the above example, an optimal control that fulfills the requirements of the theory [17] exists and such an optimal control can be recovered from the implicit relation together with the state equations. However, in the case of Example 5.3 where it does not, the implicit relation gives $x = 0$ but no further indication on the (non)existence of $u$.

Thus, the theory developed in this paper is suitable from an implementation point of view as well.

Summing up, one of the primary advantages of the method developed in this paper is the use of the notion of strongly reachable subspaces. The definition of a strongly reachable subspace inherently uses the function-space of impulsive-smooth distributions that are excluded in the $\mathcal{L}^2_{\mathtt{loc}}$ theory of [17]. With $\mathcal{L}^2_{\mathtt{loc}}$ restrictions, optimal solutions that feature impulses in the control or jumps/impulses in the states do not qualify as admissible solutions by the theory in [17]. Moreover, since such solutions can be approximated arbitrarily closely by $\mathcal{L}^2_{\mathtt{loc}}$ functions but never reached, the corresponding optimal control problems are not solvable using the theory in [17].

## 6 Conclusion

In this paper, we presented a method to compute the maximal rank-minimizing solution of an LQR LMI corresponding to a single-input system (Theorem 4.1). We have developed this method using the notion of fast subspace (strongly reachable subspace) and slow subspace (weakly unobservable subspace) of the Hamiltonian system. We have shown that augmenting the basis of the good slow subspace of the Hamiltonian system $\Sigma_{\text{Ham}}$ with the basis of a subspace of the fast subspace of $\Sigma_{\text{Ham}}$ is the crucial idea that leads to the method. While developing this method, we also showed that the fast subspace and the slow subspace of a SISO system can be characterized in terms of its Rosenbrock system matrix (Theorem 3.1 and Theorem 3.2). Further, we showed that the good slow subspace of the Hamiltonian system is disconjugate (Theorem 4.2). Using the results in this paper, we also inferred that the optimal cost of an LQR problem depends on the maximal rank-minimizing solution of the corresponding LQR LMI and the projection of the initial condition of the system onto the subspace $\mathcal{V}_{\mathsf{e}}$ only. The theory in this paper finally leads to a method to design PD state-feedback controllers to solve singular LQR problem for single-input systems (Algorithm 5.1).

Although this paper deals with singular LQR problems for single-input systems, the results presented in this paper will form the bedrock for solving such problems for the multi-input case. The key idea of weakly unobservable subspace and strongly reachable subspace used in this paper are valid for the multi-input systems as well and hence it is a matter of extending the concepts developed in this paper for the multi-input case. Since the results for the single-input case itself provide great insights into the working of a single-input singular LQR problem, we present the results for single-input systems only in this paper. An extension of the results of this paper to the multi-input case will be a matter of our forthcoming paper.

Present results in the literature, in particular the ones in [17], consider the case of a singular LQR problem but are not sufficient to treat truly singular control problems with impulses in the input or jumps/impulses in the state. This work provides singular solutions to corresponding LQR problems based on the ideas introduced in [10], [20], [21] that used impulsive-smooth distributions as the function-space for the states and inputs. Such a setting seems particularly advantageous for differential-algebraic systems, since such systems inherently admit impulsive states. Hence, the approach adapted in this paper to solve singular LQR problems for state-space systems have the potential of being generalized to differential-algebraic systems as well. This will be a matter of our future research.

## References

1. Benner, P., Bujanović, Z., Kürschner, P., Saak, J.: A numerical comparison of solvers for large-scale, continuous-time algebraic Riccati equations. arXiv e-prints arXiv:1811.00850v1, Cornell University (Nov. 5, 2018)
2. Bhawal, C., Pal, D.: Almost every single-input LQR optimal control problem admits a PD feedback solution. IEEE Control Systems Letters **3**(2), 452 – 457 (2019)
3. Bhawal, C., Qais, I., Pal, D.: Constrained generalized continuous algebraic Riccati equations (CGCAREs) are generically unsolvable. IEEE Control Systems Letters **3**(1), 192–197 (2019)
4. Bini, D., Iannazzo, B., Meini, B.: Numerical Solution of Algebraic Riccati Equations. SIAM, Philadelphia (2012)
5. Byers, R., Mehrmann, V., Xu, H.: A structured staircase algorithm for skew-symmetric/symmetric pencils. Electronic Transactions on Numerical Analysis **26**, 1–33 (2007)
6. Dai, L.: Singular Control Systems. Springer-Verlag Berlin, Heidelberg (1989)
7. Ferrante, A., Ntogramatzidis, L.: The generalized continuous algebraic Riccati equation and impulse-free continuous-time LQ optimal control. Automatica **50**(4), 1176–1180 (2014)
8. Ferrante, A., Ntogramatzidis, L.: On the reduction of the continuous-time generalized algebraic Riccati equation: An effective procedure for solving the singular LQ problem with smooth solutions. Automatica **93**, 554–558 (2018)
9. Francis, B.: The optimal linear-quadratic time-invariant regulator with cheap control. IEEE Transactions on Automatic Control **24**(4), 616–621 (1979)
10. Hautus, M., Silverman, L.: System structure and singular control. Linear Algebra and its Application **50**, 369–402 (1983)
11. Heemels, W., Schumacher, H., Weiland, S.: Linear complementarity systems. SIAM Journal on Applied Mathematics **60**(4), 1234–1269 (2000)
12. Ionescu, V., Oară, C., Weiss, M.: Generalized Riccati theory and robust control: a Popov function approach. John Wiley (1999)
13. Kučera, V.: Algebraic Riccati equation: Hermitian and definite solutions. In: S. Bittanti, A. Laub, J. Willems (eds.) The Riccati Equation, pp. 53–88. Springer Berlin Heidelberg (1991)
14. Kwakernaak, H., Sivan, R.: The maximally achievable accuracy of linear optimal regulators and linear optimal filters. IEEE Transactions on Automatic Control **17**(1), 79–86 (1972)
15. Reis, T.: Luré equations and even matrix pencils. Linear Algebra and its Applications **434**, 152–173 (2011)
16. Reis, T., Rendel, O.: Numerical solution of matrix equations arising in linear-quadratic optimal control of differential-algebraic equations. Hamburger Beiträge zur Angewandten Mathematik 2016-36, Universität Hamburg, Fachbereich Mathematik (2016)
17. Reis, T., Rendel, O., Voigt, M.: The Kalman-Yakubovich-Popov inequality for differential-algebraic systems. Linear Algebra and its Applications **485**, 153–193 (2015)
18. Rosenbrock, H.H.: Transformation of linear constant system equations. Proceedings of the Institution of Electrical Engineers **114**(4), 541–544 (1967)

19. Schumacher, J.: The role of the dissipation matrix in singular optimal control. Systems & Control Letters **2**(5), 262–266 (1983)
20. Willems, J.: Almost invariant subspaces: an approach to high gain feedback design–part I: almost controlled invariant subspaces. IEEE Transactions on Automatic Control **26**(1), 235–252 (1981)
21. Willems, J., Kitapçi, A., Silverman, L.: Singular optimal control: a geometric approach. SIAM Journal on Control and Optimization **24**(2), 323–337 (1986)
22. Wonham, W.: Linear Multivariable Control: A Geometric Approach. Springer-Verlag, New York (1985)

# Almost every single-input LQR optimal control problem admits a PD feedback solution: An addendum

Chayan Bhawal and Debasattam Pal

**Abstract**

In this technical note we present an example to demonstrate the method to design PD controllers that solve singular LQR problems.

Consider a system with state-space dynamics

$$\frac{d}{dt}x = \begin{bmatrix}1 & 0 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{bmatrix} x + \begin{bmatrix}0\\1\\0\end{bmatrix} u.$$

For every initial condition $x_0$, find an input $u$ that minimizes the functional

$$\int_0^\infty \left(x^T Q x\right) dt, \text{ where } Q := \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

**Step1 (Computation of eigen-basis of $(E,H)$ corresponding to $\Lambda \subset \sigma(E,H) \subset \mathbb{C}_-$):**
The Hamiltonian pencil pair $(E,H)$ for this problem is

$$E := \begin{bmatrix}1 & 0 & 0 & & & & \\ 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & & & & \\ & & & 1 & 0 & 0 & \\ & & & 0 & 1 & 0 & \\ & & & 0 & 0 & 1 & \\ & & & & & & 0\end{bmatrix} \text{ and } H := \begin{bmatrix}1 & 0 & 1 & 0 & 0 & 0 & 0\\ 1 & 0 & 1 & 0 & 0 & 0 & 1\\ 1 & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -1 & -1 & -1 & 0\\ 0 & 0 & 0 & 0 & 0 & -1 & 0\\ 0 & 0 & -1 & -1 & -1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & 0\end{bmatrix}.$$

It can be verified that $\texttt{det}(sE - H) = 1 - s^2$. Hence, $\Lambda = \{-1\}$. The eigenvector of $(E,H)$ corresponding to $-1$, is $\begin{bmatrix}1 & 1 & -2 & 2 & 0 & 0 & 0\end{bmatrix}^T$. Therefore, $V_{1\Lambda} = \begin{bmatrix}1 & 1 & -2\end{bmatrix}^T$, $V_{2\Lambda} = \begin{bmatrix}2 & 0 & 0\end{bmatrix}^T$ and $V_{3\Lambda} = 0$. Here, $\mathtt{n} = 3$ and $\mathtt{s} = 1$. Thus, $\mathtt{f} = \mathtt{n} - \mathtt{s} = 2$. Therefore, we have

$$X_{1\Lambda} = \begin{bmatrix}V_{1\Lambda} & b & Ab\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ -2 & 0 & 1\end{bmatrix}$$

**Step2 (Computation of the controller matrice $F_{\mathtt{p}}$ and $F_{\mathtt{d}}$):**
We compute $F_{\mathtt{p}}$ and $F_{\mathtt{d}}$ using the following equations:

$$F_{\mathtt{p}} := \begin{bmatrix}V_{3\Lambda} & g_0 & g_1\end{bmatrix} X_{1\Lambda}^{-1}, \tag{1}$$

$$F_{\mathtt{d}} := \begin{bmatrix}0_{1,\mathtt{s}} & 1 & -g_0\end{bmatrix} X_{1\Lambda}^{-1}, \tag{2}$$

Assigning $g_0 = 0$ and defining $g_1 =: g$ in equation (1), we have

$$F_{\mathtt{p}} = \begin{bmatrix}0 & 0 & g\end{bmatrix} \begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ -2 & 0 & 1\end{bmatrix}^{-1} = \begin{bmatrix}0 & 0 & g\end{bmatrix} \begin{bmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ 2 & 0 & 1\end{bmatrix} = \begin{bmatrix}2g & 0 & g\end{bmatrix}$$

Similarly, from equation (2), we have

$$F_{\mathtt{d}} = \begin{bmatrix}0 & 1 & 0\end{bmatrix} \begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ -2 & 0 & 1\end{bmatrix}^{-1} = \begin{bmatrix}-1 & 1 & 0\end{bmatrix}$$

Thus,

$$I_3 - BF_{\mathtt{d}} = \begin{bmatrix}1 & 0 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}, \text{ and } A + BF_{\mathtt{p}} = \begin{bmatrix}1 & 0 & 1\\ 1+2g & 0 & 1+g\\ 1 & 1 & 0\end{bmatrix}.$$

Note that $\texttt{det}\left(s(I_3 - BF_{\mathtt{d}}) - (A + BF_{\mathtt{p}})\right) = -g(s+1)$. Thus, if we chose any $g \in \mathbb{R}\backslash 0$ then $\texttt{det}\left(s(I_3 - BF_{\mathtt{d}}) - (A + BF_{\mathtt{p}})\right) \neq 0$. Hence, for any value of $g \in \mathbb{R} \setminus 0$, we have a PD-controller that solves the singular LQR problem. Note that there are uncountable numbers of PD-controllers that solve this singular optimal control problem.

For initial condition $x_0 = \left[\begin{smallmatrix}1\\1\\-2\end{smallmatrix}\right]\beta + \left[\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right]\alpha_0 + \left[\begin{smallmatrix}0\\0\\1\end{smallmatrix}\right]\alpha_1$, the optimal input for this problem is $\bar{u} = -2e^{-t}\beta - \alpha_0\delta - \alpha_1\dot{\delta}$.