COSOR-memorandum 87-14

All optimal controls for the singular linear-quadratic problem without stability; a new interpretation of the optimal cost

by

A.H.W. Geerts
All optimal controls for the singular linear-quadratic problem without stability; a new interpretation of the optimal cost

ABSTRACT

The singular linear-quadratic control problem without stability is solved by means of a generalized dual structure algorithm in order to generate all optimal inputs. Furthermore it is shown that the optimal cost can be interpreted as the smallest non-negative rank minimizing solution of a certain matrix inequality, the so-called dissipation inequality.

April 1987

Research supported by the Netherlands Organization for the Advancement of Pure Scientific Research (Z.W.O.).
1. Introduction

In this paper we shall consider semi-definite linear-quadratic control problems for continuous-time systems in which the cost functional is not positive definite w.r.t. the control. In [2] these so-called singular problems were studied in depth and it was stated there that the optimal control is generally not unique. Whereas this feature of singular control problems has been long recognized ([3], also [5]), to the author's knowledge no straightforward calculation of all optimal controls is known up till now.

The present paper should be considered as an extension of [1], in which for the first time distributions were introduced in the class of allowed inputs for the linear-quadratic problem. A 'right structure algorithm' ([1, Sec. 4]), then, characterized several notions from geometric theory which play a large role in singular control problems ([1], [2]).

Here, we will define a modified structure algorithm, following the approach in [1]. This modified structure algorithm will prove to be useful in determining all inputs within the class of impulsive-smooth distributions ([1, Sec. 3]) that are optimal for the singular problem we consider. In fact, the algorithm enables one to compute the linear manifolds on which the optimal trajectories lie for positive times as well as the initial impulsive inputs which let the initial state value jump instantaneously onto these manifolds. Indeed, the smooth part of the state trajectory will be shown to consist of components which follow uniquely from a reduced order Riccati equation together with components that introduce non-uniqueness of optimal controls.

For reasons of surveyability, we will concentrate on infinite horizon problems only. Also, we will discuss in this article only the case where no endpoint conditions are imposed on the state trajectory. We will elaborate on problems with stability (problems where the state should vanish as time goes to infinity) in a forthcoming paper.

A second contribution to be presented here concerns the rank minimizing problem of the dissipation matrix ([17], [18]). In [17] it was shown that the symmetric matrix, that defines the optimal cost for the linear-quadratic problem with stability, can be found as the largest element in the set of matrices that satisfy both the dissipation inequality and minimize the rank of the dissipation matrix. Here, we will give a complete characterization of all rank minimizing solutions of the dissipation inequality by means of the Riccati equation mentioned before. Thus it is shown in particular that the optimal cost for the problem without stability also may be interpreted as a rank minimizing solution of the dissipation inequality and is, in fact, the smallest non-negative one.
2. Outline

In Section 3 the problem is stated and the distributional setup from [1] is, in short, memorized. Also some geometric concepts and a few properties coming along with them are mentioned. In Section 4 we will display the construction of the dual structure algorithm in full detail since it plays a central part in things to come. In addition, several relationships between the algorithm and subspaces of importance are revealed. The full solution of the infinite horizon singular control problem without stability, then, is stated in Section 5. There, a suitable state space decomposition is introduced in order to separate those parts where non-uniqueness in optimal control occurs from those components which are to be chosen uniquely. Finally, in the last Section, the dissipation matrix rank minimizing interpretation of the optimal cost is discussed.
3. Problem statement and some geometric concepts

Since our paper follows the conceptual setup of [1], we will only mention the main features of that approach here and refer for the remaining details to [1].

We will consider the finite-dimensional linear time-invariant system \( \Sigma \):

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0, \\
y(t) &= Cx(t) + Du(t),
\end{align*}
\]

(3.1a) (3.1b)

together with the quadratic cost-functional

\[
J(x_0, u) = \int_0^\infty \|y(t)\|^2 \, dt.
\]

(3.1c)

Here, \( u(t) \in \mathbb{R}^m \), \( x(t) \in \mathbb{R}^n \), \( y(t) \in \mathbb{R}^r \), \( \| \cdot \| \) denotes the Euclidean norm and, without loss of generality, we assume that the mappings \( \begin{bmatrix} B \\ D \end{bmatrix} \), \( \begin{bmatrix} C \\ D \end{bmatrix} \) are injective and surjective, respectively.

The linear-quadratic control problem associated with \( \Sigma \) (LQCP) now is: Find the infimum of \( J(x_0, u) \) with respect to a certain class of inputs (chosen once and for all) and try to compute, if it exists, an optimal control.

The problem is called regular if the matrix \( D \) in (3.1b) is left invertible and singular if it is not. It is well known that the optimal controls will be smooth ([1, Sec. 3]) in regular problems and that in singular problems the optimal inputs in general will be distributions ([1], [2], [5]).

Since regular problems are understood completely ([12], [13], [18]), it will be our standing assumption from now on that \( D \) is not left invertible.

Therefore we have to decide on the class of allowed distributional inputs and, as in [1], we will restrict ourselves to the class of impulsive-smooth distributions \( C_{\text{imp}} \):

**Definition 3.1.**

\[
C_{\text{imp}} := \{ d \in D' \mid d = d_1 + d_2, \; d_1 \text{ impulsive, } d_2 \text{ smooth} \}
\]

where \( D' \) is the set of distributions on \( \mathbb{R} \) with support on \([0, \infty)\) ([1], [5]), smooth elements of \( D' \) are regular distributions that are smooth on \([0, \infty)\) and impulsive elements of \( D' \) are linear combinations of the Dirac distribution \( \delta \) and its derivatives (See for details on distributions [7], also [5]).
We recall the following crucial property of $C_{\text{imp}}$:

**Proposition 3.2.** (11, [6]).

$C_{\text{imp}}$ is closed under convolution.

To simplify notation, we denote convolution by juxtaposition, the $\delta$ distribution by 1 and its derivative by $p$. Thus, an impulsive distribution can be written as $\sum_{i=0}^{k} a_i p^i$, where $a_i \in \mathbb{R}$ for $i = 0, 1, \ldots, k$ and where $p^0$ is understood to be the $\delta$ distribution $1$.

Using straightforward extensions of distributional concepts to vectors and matrices, we are thus led to the distributional interpretation of (3.1a):

$$
px = Ax + Bu + x_0
$$

(3.2a)

where $x_0 = x_0 \cdot 1 = x_0 \delta$ and $u \in C_{\text{imp}}^\bullet$ ([1, Sec. 3]). The solution of (3.2a) within $D^s_+$ is unique, namely

$$
x = (pl - A)^{-1} [Bu + x_0] .
$$

Thus, $x$ is in $C_{\text{imp}}^\bullet$ and therefore

$$
y = Cx + Du
$$

(3.2b)

$$
= T(p)u + C(pl - A)^{-1} x_0
$$

is in $C_{\text{imp}}^\bullet$ with

$$
T(s) := D + C (sl - A)^{-1} B ,
$$

(3.3)

the transfer function. Observe that $T(p)$ is the matrix-valued distribution obtained by setting $s = p$ and interpreting $(pl - A)^{-1}$ to be $e^{At}$ ($t \geq 0$), see [1].

In order to stress dependence of $x, y$ on $x_0$ and $u$, we will write

$$
x(x_0, u) = (pl - A)^{-1} Bu + (pl - A)^{-1} x_0
$$

(3.4a)

and

$$
y(x_0, u) = Cx(x_0, u) + Du .
$$

(3.4b)

Now define formally
\[ J : \mathbb{R}^n \times C^m_{\text{imp}} \rightarrow \mathbb{R}^+ \]

by

\[ J(x_0, u) := \int_0^\infty \|y(x_0, u)\|^2 \, dt \quad (3.5) \]

where we define \( J(x_0, u) := +\infty \) if \( u \) is such that \( y(x_0, u) \notin L_2^2(\mathbb{R}^+), \) the space of all \( r \)-vectors whose components are square-integrable over \( \mathbb{R}^+. \)

Also, for \( u = u_1 + u_2, \) \( u_1 \) impulsive, \( u_2 \) smooth and consequently \( x = x_1 + x_2, \) \( x_1 \) impulsive, \( x_2 \) smooth, set

\[
\begin{align*}
  u(\infty) &:= u_2(\infty), \text{ if existent,} \\
  x(\infty) &= x(x_0, u)(\infty) := x_2(\infty), \text{ if existent, and} \\
  y(\infty) &= y(x_0, u)(\infty) := Cx(\infty) + Du(\infty). 
\end{align*} \quad (3.6)
\]

In addition, \( u(0^+) := u_2(0^+) = \lim_{t \downarrow 0} u_2(t), \) \( x(0^+) := x_2(0^+), \) etc.

Finally, define

\[ J(x_0) := \inf_{u \in C^m_{\text{imp}}} J(x_0, u). \quad (3.7) \]

Thus, we may state the linear-quadratic control problem (LQCP) without stability:

Given the system (3.2), find \( J(x_0) = \inf_{u \in C^m_{\text{imp}}} J(x_0, u) \) and calculate, if they exist, all optimal inputs.

The LQCP with stability, to be discussed in a future article, may be stated as follows:

Given the system (3.2), find \( J(x_0) \) under the side condition \( x(\infty) = 0 \) and calculate, if they exist, all optimal controls.

Since we are only interested in those inputs \( u \) for which \( y(x_0, u) \) is regular, we will call these controls admissible (11), and the space of admissible inputs, which is system dependent, is denoted by \( U_\Sigma. \) The structure algorithm in Section 4 enables one to characterize \( U_\Sigma \) completely, as will be illustrated later on. However, before doing so, we will recall some geometric aspects of singular control first.
Definition 3.3.

A state \( x \) is called strongly reachable from the origin if there exists an impulsive input \( u \in U_\Sigma \) such that for the corresponding state trajectory we have \( x(0, u)(0^+) = x_1 \). The space of strongly reachable states is denoted \( W = W(\Sigma) \).

Lemma 3.4.

\[
W(\Sigma) = \{ x_0 \mid \exists u \in U_\Sigma : x(x_0, u)(0^+) = 0 \} \\
= \{ x_0 \mid \exists u \in C_{\text{imp}} : x(x_0, u)(0^+) = 0, y(x_0, u) = 0 \}.
\]

Proof. Follows from the discussion in [1, Sec. 3] and from the observation that \( x(0^+) \) only depends on the impulsive part of \( u \). Note that if \( u \in U_\Sigma \) gives \( x(0^+) = x_1 \), then \( x(x_0, u)(0^+) = x_0 + x_1 \).

Lemma 3.4 immediately leads to a partial solution of the LQCP:

Lemma 3.5.

\[ \forall x_0 \in W(\Sigma) : J(x_0) = 0. \]

Observe that the optimal cost for the LQCP with stability equals the optimal cost for the LQCP without stability when \( x_0 \in W(\Sigma) \).

The dual concept of \( W(\Sigma) \) is the subspace of weakly unobservable states \( V(\Sigma) \):

Definition 3.6.

A state \( x_0 \) is weakly unobservable if there exists a regular input on \( [0, \infty) \) such that \( y(x_0, u) = 0 \) on \( \mathbb{R}^+ \). The space of weakly unobservable states is denoted \( V = V(\Sigma) \).

For details on \( V \) and \( W \) we refer the reader to [1], [2], [15]. Here, we will primarily be interested in their sum and their intersection.
Proposition 3.7. (1))

\[ x_0 \in V + W \iff \exists u \in U : y(x_0, u)(t) = 0, \ t > 0 \]
\[ \iff \exists u \in \mathbb{C}^{m \times n}_{\text{imp}} : y(x_0, u) = 0 . \]

Because of this result we will call \( V_d := V + W \) the subspace of distributionally weakly unobservable states.

The subspace \( V_d \) allows one to decide on the right invertibility of the system \( \Sigma \):

Definition 3.8.

The system \( \Sigma \) is right invertible if for every \( y \in C^\rho_{\text{imp}} \) there exists a \( u \in C^{m \times n}_{\text{imp}} \) such that \( y(0, u) = y \).

Proposition 3.9. (1))

The following statements are equivalent:

(i) \( \Sigma \) is right invertible,

(ii) \( V_d = \mathbb{R}^n, \ \text{im}[C, D] = \mathbb{R}^r \),

(iii) The transfer function \( T(s) \) (see (3.3)) is right invertible as a rational matrix.

Combination of Propositions 3.7, 3.9 leads to an answer to one of the questions in the LQCP without stability for right invertible systems.

Lemma 3.10.

\( \Sigma \) right invertible \( \iff \forall x_0 \in \mathbb{R}^n : J(x_0) = 0. \)

The intersection of \( V \) and \( W, V \cap W := \mathbb{R} \), turns out to be strongly related to the notion of left invertibility:
Definition 3.11.

The system $\Sigma$ is left invertible if for all nonzero $u \in C_{\text{imp}}^n$ we have that $y(0, u) \neq 0$.

Proposition 3.12. ([1])

The following statements are equivalent:

(i) $\Sigma$ is left invertible,

(ii) $R = \{0\}$, $\ker \begin{bmatrix} B \\ D \end{bmatrix} = \{0\}$,

(iii) The transfer function $T(s)$ is left invertible as a rational matrix.

Remarks

1. For a left invertible system, the set of optimal controls for the $LQCP$, if not empty, always contains at most one element.

2. Note that if $R \neq \{0\}$, then there are for every $x_0 \in R$ at least two optimal controls for the $LQCP$ without stability. This follows from Lemma 3.4 and Definition 3.6. In Section 5 we shall see that non-uniqueness in optimal control always occurs when $R \neq \{0\}$.

3. If a system $\Sigma$ is both left invertible and right invertible, then the transfer function is square and invertible (and conversely). Such systems are called invertible ([4]).
4. The generalized dual structure algorithm

In [1] the notion of 'dual structure algorithm' was introduced and applied to study the linear-quadratic problem for left invertible systems. Here, we propose an approach somewhat different from the one in [1, Sec. 4] in order to analyse linear systems which are not necessarily left invertible. Although the construction of the algorithm is rather lengthy and notationally involved, we would like to stress the method’s significance in transforming the linear-quadratic control problem under consideration into a related control problem which is immediately solvable.

Now consider the system \( \Sigma \):

\[
px = Ax + Bu + x_0, \tag{4.1a}
\]

\[
y = Cx + Du \tag{4.1b}
\]

with \( D \) not left invertible, \( \ker \begin{bmatrix} B \\ D \end{bmatrix} = \{0\} \), \( \text{im}[C, D] = \mathbb{R}^r \).

**Step 0.** Assume that \( \text{rank}(D) = q_0 < m \).

Then there exists a permutation matrix \( R_0 = [\tilde{R}_0, \tilde{R}_1] \), \( \text{rank}(\tilde{R}_0) = q_0 \), such that \( D_0 := D\tilde{R}_0 \) is left invertible with \( \text{rank} q_0 \) and \( \text{im}(D\tilde{R}_0) \subseteq \text{im}(D) = \text{im}(D) \). Therefore \( D\tilde{R}_0 = D_0\tilde{K}_0^* \) for some \( q_0 \times (m - q_0) \) matrix \( \tilde{K}_0^* \). If \( R_0^* := (-\tilde{R}_0\tilde{K}_0^* + \tilde{R}_0) \), then it is easily seen that \( S_0 := [\tilde{R}_0, R_0^*] \) is invertible and that

\[
DS_0 = [D_0, 0]. \tag{4.2}
\]

Defining

\[
BS_0 := [\tilde{B}_0, \tilde{B}_1] \tag{4.3}
\]

and

\[
u = S_0 \begin{bmatrix} \bar{w}_0 \\ \bar{v}_0 \end{bmatrix}, \tag{4.4}
\]

then yields the following description for \( \Sigma_0 := \Sigma \):

\[
px = Ax + \tilde{B}_0\bar{w}_0 + \tilde{B}_0\bar{v}_0 + x_0, \tag{4.5a}
\]

\[
y = Cx + D_0\bar{v}_0 \tag{4.5b}
\]

where \( \tilde{B}_0 \) is left invertible since \( \begin{bmatrix} B \\ D \end{bmatrix} \) is.
Note that in case of left invertibility of $D$, it would not have been possible to make the above separation.

It follows from (4.5) that $y$ will be regular if $\bar{w}_0$ is regular and $\bar{w}_0$ is the derivative of a regular function. This suggests the substitution

$$\bar{w}_0 = p\bar{v}_0$$  \hspace{1cm} (4.6)

in (4.5a). If we next define

$$x_1 := x - \bar{B}_0\bar{v}_0$$  \hspace{1cm} (4.7)

(compare the transformations in [8] to [11]), we obtain the system $\Sigma_1$ given by:

$$px_1 = Ax_1 + \bar{B}_0\bar{w}_0 + A\bar{B}_0\bar{v}_0 + x_0,$$  \hspace{1cm} (4.8a)

$$y = Cx_1 + D_0\bar{w}_0 + C\bar{B}_0\bar{v}_0.$$  \hspace{1cm} (4.8b)

Observe that $\text{rank} (D_0, C\bar{B}_0) \geq \text{rank} (D_0)$ and that

$$u = H_0(p) \begin{bmatrix} \bar{w}_0 \\ \bar{v}_0 \end{bmatrix}$$  \hspace{1cm} (4.9a)

where

$$H_0(s) = \begin{bmatrix} I_{r_0} & 0 \\ 0 & sI_{r_0} \end{bmatrix}$$  \hspace{1cm} (4.9b)

with $r_0 = m - q_0$.

All of the following steps that occur in the algorithm in fact consist of three separate column selection procedures. We will indicate their underlying objectives below.

**Step 1**

Part 1. Let $\text{rank} (D_0, C\bar{B}_0)$ be $q_0 + q_1 \leq m$. Then there exists a permutation matrix $R_1 = [\bar{R}_1, \bar{R}_1]$, rank $(\bar{R}_1) = q_1$, such that $C\bar{B}_0\bar{R}_1 := D_1$ is left invertible, has rank $q_1$, and is independent of $D_0$, whereas $\text{im}(C\bar{B}_0\bar{R}_1) = \text{im}(D_0, D_1)$. Thus, for some $q_0 \times (r_0 - q_1)$, $q_1 \times (r_0 - q_1)$ matrices $K^*_1$ and $K^*_2$ it holds that $C\bar{B}_0\bar{R}_1 = D_0K^*_1 + D_1K^*_2$. If $R^*_1 = (-\bar{R}_1K^*_2 + \bar{R}_1)$ then $S^*_1 := [\bar{R}_1, R^*_1]$ is invertible and with the transformation
it is found that
\[ [D_0, C\bar{B}_0] S_1 = [D_0, D_1, 0] \quad (4.11) \]
Defining, as in step 0,
\[ [\bar{B}_0, A\bar{B}_0] S_1 =: [\bar{B}_0, \bar{B}_1, \bar{B}_*] \quad (4.12) \]
and, in addition,
\[ \begin{bmatrix} \bar{w}_0 \\ \bar{v}_0 \end{bmatrix} := \begin{bmatrix} \bar{w}_{10} \\ \bar{w}_{11} \end{bmatrix}, \quad (4.13) \]

(4.8) transforms into
\[
px_1 = Ax_1 + [\bar{B}_0, \bar{B}_1] \begin{bmatrix} \bar{w}_{10} \\ \bar{w}_{11} \end{bmatrix} + \bar{B}_1 \bar{u}_1 + x_0 , \quad (4.14a) \\
y = Cx_1 + [D_0, D_1] \begin{bmatrix} \bar{w}_{10} \\ \bar{w}_{11} \end{bmatrix} , \quad (4.14b) 
\]
with the \( n \times (r_0 - q_1) \) matrix \( \bar{B}_1 \) not necessarily of full column rank.

Note that in part 1 we have tried to 'regularize' the system \( \Sigma_1 \): If we would have found \( \bar{K}_1 = I_{r_0} \) then \( [D_0, C\bar{B}_0] \) would have been of full column rank and hence the usual theory of optimal regulators ([12], [13], [18], [19]) could have been applied to the system \( \Sigma_1 \).

Part 2. Since \( \bar{B}_1 \) is not necessarily left invertible, we may apply a transformation which selects only the independent columns of \( \bar{B}_1 \). To be more specific, assume that the invertible matrix \( \bar{P}_1 = [\bar{P}_1, P_1^*] \) is such that
\[ \bar{B}_1 \bar{P}_1 = [\bar{B}_1 \bar{P}_1, 0] \quad (4.15) \]
with rank \( (\bar{B}_1 \bar{P}_1) = \text{rank } (\bar{P}_1) =: p_1, \text{rank } (P_1^*) =: \sigma_1 =: (r_0 - q_1 - p_1) \) and set
\[ \bar{u}_1 =: \bar{P}_1 \begin{bmatrix} \bar{w}_{10} \\ \bar{w}_{11} \end{bmatrix} . \quad (4.16) \]

Thus, the system equation for \( x_1 \) becomes
\[ px_1 = Ax_1 + \begin{bmatrix} \bar{\mathbf{B}}_0 & \bar{\mathbf{F}}_1 \end{bmatrix} \begin{bmatrix} \bar{w}_{10} \\ \bar{w}_{11} \end{bmatrix} + \bar{B}_1 \bar{F}_1 \bar{w}_1 + x_0 \quad (4.17) \]

Part 3. This part actually sets apart those columns of \( \bar{B}_1 \bar{F}_1 \) which \textit{a priori} cannot enlarge rank \((\mathcal{D}_0, \mathcal{D}_1, C\bar{B}_1 \bar{F}_1)\) w.r.t. rank \((\mathcal{D}_0, \mathcal{D}_1)\); this being the objective in the first part of step 2 (compare step 1, part 1).

Therefore, let \( V_1 = [\bar{V}_1, \bar{V}_2] \) be a permutation matrix, rank \((\bar{V}_1) = r_1, \) rank \((\bar{V}_2) = p_1 = p_1 - r_1, \) such that \( \bar{B}_1 \bar{F}_1 \bar{V}_1 \) is left invertible, independent of \( \bar{B}_0, \) whereas \( \text{im}(\bar{B}_1 \bar{F}_1 \bar{V}_1) \subset \text{im}(\bar{B}_0). \)

Then for some \( r_0 \times p_1 \) matrix \( N_t^* : \bar{B}_1 \bar{F}_1 \bar{V}_1 = \bar{B}_0 N_t^*. \)

Now substitute into (4.17)

\[ \bar{w}_1 = V_1 \begin{bmatrix} \bar{w}_1 \\ \bar{w}_{1c} \end{bmatrix} \quad (4.18) \]

which yields

\[ px_1 = Ax_1 + \begin{bmatrix} \bar{\mathbf{B}}_0 & \bar{\mathbf{F}}_1 \end{bmatrix} \begin{bmatrix} \bar{w}_{10} \\ \bar{w}_{11} \end{bmatrix} + \bar{B}_1 \bar{F}_1 \bar{V}_1 \bar{w}_1 + \bar{B}_1 \bar{F}_1 \bar{V}_1 \bar{w}_{1c} + x_0 \quad (4.19) \]

At this point the fundamental difference between the algorithm in [1, Sec. 4] and our method becomes apparent.

Here, instead of \([\bar{w}_1, \bar{w}_{1c}] = p(\bar{v}_1, \bar{v}_{1c}) \) (See [1]), we propose the substitution

\[ \bar{w}_1 = p\bar{v}_1 \quad , \quad (4.20) \]

(compare (4.6)) and we define in (4.19)

\[ x_2 := x_1 - \bar{B}_1 \bar{F}_1 \bar{V}_1 \bar{v}_1 \quad . \quad (4.21) \]

Thus we arrive at the system \( \Sigma_2 \) described by:

\[ px_2 = Ax_2 + \begin{bmatrix} \bar{\mathbf{B}}_0 & \bar{\mathbf{F}}_1 \end{bmatrix} \begin{bmatrix} \bar{w}_{10} \\ \bar{w}_{11} \end{bmatrix} + A\bar{B}_1 \bar{F}_1 \bar{V}_1 \bar{v}_1 + \]

\[ + \bar{B}_1 \bar{F}_1 \bar{V}_1 \bar{w}_{1c} + x_0 \quad , \quad (4.22a) \]
Note that $\text{rank} \left[ D_0, D_1, CB, \bar{F}, \bar{V} \right] \geq \text{rank} \left[ D_0, D_1 \right]$ and that indeed
$$\text{rank} \left[ D_0, D_1, CB, \bar{F}, \bar{V} \right] = \text{rank} \left[ D_0, D_1 \right].$$

Furthermore, the controls for $\Sigma_1$ and $\Sigma_2$ are linked by $H_1(p)$:

$$\begin{bmatrix}
\bar{w}_0 \\
\bar{v}_0 
\end{bmatrix} = H_1(p)
\begin{bmatrix}
\bar{w}_0 \\
\bar{v}_0 \\
w_{1e} \\
w_{1e}
\end{bmatrix}
\begin{bmatrix}
I_{q_1} \\
0 \\
0 \\
I_{\sigma_1}
\end{bmatrix}
\begin{bmatrix}
I_{q_1} \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}$$

where $(q_1 = q_0 + q_1)$

$$H_1(s) = S_1 
\begin{bmatrix}
I_{q_1} \\
0 \\
0 \\
I_{\sigma_1}
\end{bmatrix}
\begin{bmatrix}
I_{q_1} \\
0 \\
0 \\
I_{\sigma_1}
\end{bmatrix}
\begin{bmatrix}
I_{q_1} \\
0 \\
0 \\
I_{\sigma_1}
\end{bmatrix}.$$

Next, we describe the general iteration step $(k+1)$, $k \geq 0$.

**Step $(k+1)$**

The system $\Sigma_{k+1}$ is given by

$$px_{k+1} = Ax_{k+1} + \bar{B}_k \bar{w}_k + A \bar{B}_k \bar{F}_k \bar{V}_k \bar{v}_k,$$

$$+ \sum_{i=1}^{k} \bar{B}_i \bar{F}_i \bar{V}_i w_i + \sum_{i=1}^{k} 0 \cdot w_i + x_0.$$  \hspace{1cm} (4.25a)

$$y = C x_{k+1} + D_k \bar{w}_k + C \bar{B}_k \bar{F}_k \bar{V}_k \bar{v}_k.$$  \hspace{1cm} (4.25b)

$(\bar{B}_0 \bar{F}_0 \bar{V}_0 := \bar{B}_0)$ with

$$\bar{B}_k = [\bar{B}_0, \bar{B}_1, \ldots, \bar{B}_k],$$

$$D_k = [D_0, D_1, \ldots, D_k].$$  \hspace{1cm} (4.26a)

(4.26b)
Further, for all \( i = 1 \cdots k \) we have that

\[
\tilde{B}_i \tilde{F}_i \tilde{V}_i = W_i N_{i}^*
\]

where \( W_i = [\tilde{B}_0, \tilde{B}_1 \tilde{F}_1 \tilde{V}_1, \cdots, \tilde{B}_{i-1} \tilde{F}_{i-1} \tilde{V}_{i-1}] \) is a left invertible matrix with rank \( L_{i-1}, L_{i-1} = r_0 + r_1 + \cdots + r_{i-1} \) and \( N_i^* \) some \( L_{i-1} \times p_i \) matrix, \( p_i = \text{dim} (w_i) \).

Moreover, \( D_k \) is left invertible, \( \text{rank } q_k = q_0 + q_1 + \cdots + q_k \), \( \text{dim} (\tilde{v}_k) \) = rank \( (\tilde{B}_k \tilde{F}_k \tilde{V}_k) = r_k \) and with

\[
\tilde{\rho}_k := p_1 + \cdots + p_k \quad (\tilde{\rho}_0 := 0) ,
\]

\[
\tilde{\sigma}_k := \sigma_1 + \cdots + \sigma_k \quad (\sigma_0 := 0) \quad \text{dim} (w_i) \quad (i = 1 \cdots k) ,
\]

it holds that

\[
q_k + r_k + \tilde{\rho}_k + \tilde{\sigma}_k = m .
\]

Part 1. Let \( \text{rank } (D_k, CB_k \tilde{F}_k \tilde{V}_k) = q_k + q_{k+1} \leq m \).

Then there exists a permutation matrix \( R_{k+1} = [\tilde{R}_{k+1}, \tilde{R}_{k+1}] \) such that \( D_{k+1} := CB_k \tilde{F}_k \tilde{V}_k \tilde{R}_{k+1} \) is left invertible, with rank equal to \( q_{k+1} \) and independent of \( D_k \). Moreover \( CB_k \tilde{F}_k \tilde{V}_k \tilde{R}_{k+1} = D_k K_{k+1}^* + D_{k+1} K_{k+1}^* \) for certain \( q_k \times (r_k - q_{k+1}), q_{k+1} \times (r_k - q_{k+1}) \) matrices \( K_{k+1}^*, K_{k+1}^* \).

Then

\[
[D_k, CB_k \tilde{F}_k \tilde{V}_k] S_{k+1} = [D_k, D_{k+1}, 0]
\]

where \( S_{k+1} \) is the regular transformation

\[
S_{k+1} = \begin{bmatrix}
1_{\tilde{\rho}_k} & [0, -K_{k+1}^*] \\
0 & S_{k+1}^*
\end{bmatrix}
\]

with \( S_{k+1}^* = [\tilde{R}_{k+1}, R_{k+1}^*], R_{k+1}^* = (-\tilde{R}_{k+1}, K_{k+1}^* + \tilde{R}_{k+1}) \).

Define

\[
[\tilde{B}_k, A\tilde{B}_k \tilde{F}_k \tilde{V}_k] S_{k+1} = [\tilde{B}_k, B_{k+1}, \tilde{B}_{k+1}]
\]

and introduce the new control variables by
\[
\begin{bmatrix}
\bar{w}_k \\
\bar{v}_k
\end{bmatrix} = S_{k+1} \begin{bmatrix}
\bar{w}_{k+1} \\
\bar{u}_{k+1}
\end{bmatrix},
\]

with
\[
\bar{w}_{k+1} = \begin{bmatrix}
\bar{w}_{k+1,0} \\
\bar{w}_{k+1,1} \\
\vdots \\
\bar{w}_{k+1,k}
\end{bmatrix},
\]

then (4.25) becomes
\[
p_{x_{k+1}} = Ax_{k+1} + \bar{E}_{k+1} \bar{w}_{k+1} + \bar{B}_{k+1} \bar{u}_{k+1}
+ \sum_{i=1}^{k} \bar{B}_i \bar{P}_i \bar{V}_i w_{i} + \sum_{i=1}^{k} 0 \cdot w_{i} + x_0
\]
\[
y = Cx_{k+1} + D_{k+1} \bar{w}_{k+1}.
\]

Part 2. Let the regular matrix \( P_{k+1} = [\bar{P}_{k+1}, P_{k+1}^*] \) be such that \( \bar{B}_{k+1} P_{k+1} = [\bar{B}_{k+1} \bar{P}_{k+1}, 0] \) with rank \( \bar{B}_{k+1} \bar{P}_{k+1} = \text{rank} \ (\bar{P}_{k+1}) = p_{k+1} \), rank \( P_{k+1}^* = \sigma_{k+1} \). Then with
\[
\bar{u}_{k+1} = P_{k+1} \begin{bmatrix}
\bar{w}_{k+1,1} \\
\bar{w}_{k+1,2} \\
\vdots \\
\bar{w}_{k+1,k+1}
\end{bmatrix},
\]
the system equation (4.32a) becomes
\[
p_{x_{k+1}} = Ax_{k+1} + \bar{E}_{k+1} \bar{w}_{k+1} + \bar{B}_{k+1} \bar{P}_{k+1} \bar{w}_{k+1}
+ \sum_{i=1}^{k} \bar{B}_i \bar{P}_i \bar{V}_i w_{i} + \sum_{i=1}^{k} 0 \cdot w_{i} + x_0.
\]

Part 3. Assume that the permutation matrix \( V_{k+1} = [\bar{V}_{k+1}, \bar{V}_{k+1}^*] \) is such that rank \( \bar{B}_{k+1} \bar{P}_{k+1} \bar{V}_{k+1} = \text{rank} \ (\bar{V}_{k+1}) = r_{k+1}, \bar{B}_{k+1} \bar{P}_{k+1} \bar{V}_{k+1} \) independent of \( [\bar{B}_0, \bar{B}_1 \bar{P}_1 \bar{V}_1, \ldots, \bar{B}_k \bar{P}_k \bar{V}_k] = w_{k+1} \), whereas for some \( k \times k \) matrix \( N_{k+1}^* \) it holds that \( \bar{B}_{k+1} \bar{P}_{k+1} \bar{V}_{k+1} = w_{k+1} N_{k+1}^* \) with \( \rho_{k+1} = p_{k+1} - r_{k+1} \). As in (4.18), set
\[
\bar{w}_{k+1} = V_{k+1} \begin{bmatrix}
\bar{w}_{k+1,1} \\
\bar{w}_{k+1,2} \\
\vdots \\
\bar{w}_{k+1,k+1}
\end{bmatrix},
\]
Substituting (4.35) into (4.34) then leads to
Finally, we consider $\tilde{w}_{k+1}$ to be the derivative of $\check{v}_{k+1}$, i.e. $\tilde{w}_{k+1} = p\check{v}_{k+1}$, and define

$$x_{k+2} := x_{k+1} - \bar{B}_{k+1}\bar{F}_{k+1}\check{V}_{k+1}\check{v}_{k+1}.$$  \hfill (4.37)

We then obtain the following system that will be called $\Sigma_{k+2}$:

$$p x_{k+2} = Ax_{k+2} + \bar{B}_{k+1}\tilde{w}_{k+1} + A\bar{B}_{k+1}\bar{F}_{k+1}\check{V}_{k+1}\check{v}_{k+1}$$
$$+ \sum_{i=1}^{k+1} \bar{B}_i \bar{F}_i \check{V}_i w_i + \sum_{i=1}^{k+1} 0 \cdot w_i + x_0.$$  \hfill (4.38a)

$$y = C x_{k+2} + D_{k+2}\tilde{w}_{k+1} + C\bar{B}_{k+1}\bar{F}_{k+1}\check{V}_{k+1}\check{v}_{k+1}.$$  \hfill (4.38b)

The controls for $\Sigma_{k+1}$ and $\Sigma_{k+2}$ are related by

$$\begin{bmatrix} \tilde{w}_k \\ \check{v}_k \end{bmatrix} = \tilde{H}_{k+1}(p) \begin{bmatrix} \tilde{w}_{k+1} \\ \check{v}_{k+1} \\ w_{k+1} \\ \check{v}_{k+1} \end{bmatrix}.$$  \hfill (4.39)

with

$$\tilde{H}_{k+1}(s) = S_{k+1} \begin{bmatrix} 1_{4k+1} \\ 0 \\ P_{k+1} \end{bmatrix} \begin{bmatrix} 1_{4k+1} \\ 0 \\ 0 \\ I_{4k+1} \end{bmatrix} \begin{bmatrix} 1_{4k+1} & 0 & 0 & 0 \\ 0 & sI_{4k+1} & 0 & 0 \\ 0 & 0 & 0 & I_{4k+1} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$  \hfill (4.40)

If we denote $(w^0, w^0$ non-existing)

$$w^e = \begin{bmatrix} w^e_1 \\ w^e_2 \\ \vdots \\ w^e_k \end{bmatrix}, \ w^k = \begin{bmatrix} w^k_1 \\ w^k_2 \\ \vdots \\ w^k_k \end{bmatrix},$$  \hfill (4.41)

then it is clear that
Termination of the algorithm

We will agree on terminating the algorithm when for the first time in step \((k + 1)\), part 3, it is found that

\[
 r_{k+1} = 0
\]

(4.44)
i.e., when for the first time

\[
 B_{k+1} F_{k+1} = W_{k+1} \dot{X}_{k+1}^\ast
\]

(4.45)
for some \( \Delta_k \times p_{k+1} \) matrix \( N_{k+1}^\ast \).

In this case, \( \dot{V}_{k+1} = I_{p_{k+1}} \), and hence

\[
 \dot{w}_{k+1}^1 = w_{k+1}^1
\]

which leads to the final system description for \( \Sigma_{k+1} \):

\[
 p x_{k+1} = A x_{k+1} + \vec{B}_{k+1} \vec{w}_{k+1} + \sum_{i=1}^{k+1} \vec{B}_i \vec{F}_i \dot{V}_i w_{i_e} + \sum_{i=1}^{k+1} 0 \cdot w_{i_e} + x_0
\]

(4.46a)
The relation of the algorithm with the subspaces from Sec. 3.

Let $\alpha = (k + 1) \geq 1$ be the first integer for which in step $\alpha$, part 3, it holds that $r_\alpha = 0$. Note that indeed $\alpha \geq 1$ since $r_0 = m - q_0 > 0$.

Then

$$\alpha \leq (n + 1 - r_0) \quad (4.47)$$

since

$$W_\alpha = [\tilde{B}_0, \tilde{B}_1 \tilde{V}_1, \ldots, \tilde{B}_{\alpha-1} \tilde{V}_{\alpha-1}] \quad (4.48)$$

is left invertible with rank $\mathcal{L}_{\alpha-1}$ and thus $n \geq \mathcal{L}_{\alpha-1} \geq r_0 + (\alpha - 1)$.

Further, let $\alpha_D$ be the first integer $l$ for which

$$\xi_l = \xi_\alpha \quad (4.49)$$

i.e. let $\alpha_D$ be the first integer $k$ for which in step $l, l = (k + 1) \cdots \alpha$, $\text{im} (C \tilde{B}_{l-1} \tilde{V}_{l-1}) \subset \text{im} (D_k)$. Then, by definition, $\xi_{\alpha_D} = \xi_{\alpha_D+1} = \cdots = \xi_\alpha$.

$\alpha \geq \alpha_D \geq 0 \quad (4.50)$

and thus

$$D_{\alpha_D} = D_{\alpha_D+1} = \cdots = D_\alpha \quad \xi_{\alpha_D} = \xi_{\alpha_D+1} = \cdots = \xi_\alpha \quad (4.51)$$

with

$$D_{\alpha_D} = [D_0, C \tilde{B}_0 \tilde{R}_1, \ldots, C \tilde{B}_{\alpha_D-1} \tilde{R}_{\alpha_D-1} \tilde{V}_{\alpha_D-1} \tilde{R}_{\alpha_D}] \quad \xi_{\alpha_D} = [\tilde{B}_0, A \tilde{B}_0 \tilde{R}_1, \ldots, A \tilde{B}_{\alpha_D-1} \tilde{R}_{\alpha_D-1} \tilde{V}_{\alpha_D-1} \tilde{R}_{\alpha_D}] \quad (4.52)$$

(see (4.11), (4.12), (4.28), (4.30)).

These observations yield the next system equations for $\Sigma_\alpha$:

$$px_\alpha = A x_\alpha + \xi_{\alpha_D} \tilde{W}_\alpha + B x_\alpha + x_0 \quad (4.53a)$$
with
\[ B_a^* = W_\alpha \tilde{H}_1, \]  
where \( \tilde{H}_1 \) is a given upper block triangular matrix (see (4.46), (4.51), (4.41), (4.26d), (4.48)).

Moreover,
\[ x_\alpha = x - W_\alpha \bar{z}_{\alpha-1}, \]  
(4.55)

where
\[ \bar{z}_{\alpha-1} = \begin{bmatrix} \bar{\phi}_0 \\ \bar{\phi}_1 \\ \vdots \\ \bar{\phi}_{\alpha-1} \end{bmatrix}, \]  
(4.56)

and
\[ u = \overline{H}_\alpha(p) \begin{bmatrix} \bar{w}_\alpha \\ \bar{w}_\alpha^c \\ \bar{w}_\alpha^* \end{bmatrix}, \]  
(4.57)

where
\[ \overline{H}_\alpha(s) = H_0(s) H_1(s) \cdots H_\alpha(s), \]  
(4.58)

from (4.7), (4.37), (4.9), (4.42), (4.43).

In order to exploit these results, we need some information on \( W(\Sigma) \) and \( V_d(\Sigma_\omega) \) first (here \( V_d(\Sigma_\omega) \) denotes the weakly unobservable subspace associated with \( \Sigma_\omega \)).

**Proposition 4.1**

Let \( 0 \leq k \leq (\alpha - 1) \).

Then
(i) \( W(\Sigma_k) = \text{im} \left[ \hat{B}_k \tilde{F}_k \tilde{V}_k \right] + W(\Sigma_{k+1}). \)
(ii) \( V_d(\Sigma_k) = V_d(\Sigma_{k+1}) \).

Here \( W(\Sigma) \) denotes the strongly reachable subspace for \( \Sigma_k \).

Proof. To start, we agree on working with the system description (4.25) for \( \Sigma_{k+1} \) and (4.1) for \( \Sigma_0 = \Sigma \). Note that the strongly reachable subspaces for (4.25), (4.32), (4.34) and (4.36) are equal, and so are the distributionally weakly unobservable subspaces.

We now examine \( k = 0 \); the proof for \( 0 < k \leq (\alpha - 1) \) runs analogously.

(i) Assume \( \bar{x} \in W(\Sigma) \). Then there is a \( u \in U_\Sigma \) such that \( \bar{x} = x(0, u)(0^*) \) where \( x = (pl - A)^{-1}Bu \).

Hence \( x = (pl - A)^{-1}[\tilde{B}_0\tilde{w}_0 + \tilde{B}_0\tilde{v}_0] = (pl - A)^{-1} [\tilde{B}_0\tilde{w}_0 + A\tilde{B}_0\tilde{v}_0] + \tilde{B}_0\tilde{v}_0 \) (See step 0 of the algorithm). From (4.8a), with \( x_0 = 0, x_1 = (pl - A)^{-1} [\tilde{B}_0\tilde{w}_0 + A\tilde{B}_0\tilde{v}_0] \), thus

\[
\bar{x} = x_1(0, \begin{bmatrix} \tilde{w}_0 \\ \tilde{v}_0 \end{bmatrix})(0^*) + \tilde{B}_0\tilde{v}_0(0^*) \text{ for some } [\tilde{w}_0^T, \tilde{v}_0^T]^T, \text{i.e.}
\]

\[
W(\Sigma) \subseteq \text{im} (\tilde{B}_0) + W(\Sigma) .
\]

The converse inclusion is obvious, see also [1, Prop. 4.17, (ii)]. Note that in fact (4.7) is used here.

(ii) If \( T_i(s) \) denotes the transfer function for \( \Sigma_i \) \((l = 0, 1, \cdots, \alpha)\) with \( T_0 = T \) ((3.3)), then it can easily be shown that \( T_{k+1}(s) = T_k(s) H_k(s), k = 0, 1, \cdots, (\alpha - 1) \), see (4.42), and thus, in particular, \( T_1(s) = T(s) H_0(s) \) ((4.9)). Note further that \( T_\alpha(s) = T(s) H_\alpha(s) \) ((4.58)).

From Prop. 3.7: \( x_0 \in V_d(\Sigma) \) if and only if there is a \( u \in C_{\text{imp}}^m \) such that ((3.2b))

\[
T(p)u + C(pl - A)^{-1}x_0 = 0 \text{ and } x_0 \in V_d(\Sigma) \text{ iff there is a } [\tilde{w}_0^T, \tilde{v}_0^T]^T \text{ such that}
\]

\[
T_1(p) \begin{bmatrix} \tilde{w}_0 \\ \tilde{v}_0 \end{bmatrix} + C(pl - A)^{-1}x_0 = 0. \text{ The claim now follows from the observation that}
\]

\[
u \in C_{\text{imp}}^m \iff \begin{bmatrix} \tilde{w}_0 \\ \tilde{v}_0 \end{bmatrix} = H_\alpha^{-1}(p)u \in C_{\text{imp}}^m .
\]

Remarks

1. In [1, Prop. 4.17] similar relationships between subspaces of \( \Sigma_0 \) and \( \Sigma_i \) were claimed. Nevertheless we believe that a new proof was necessary since our system \( \Sigma \) is not assumed to be left invertible.
2. One may also show \( V(\Sigma) \subseteq V(\Sigma^{+}) \), see Def. 3.6. Compare with [1. Prop. 4.17, (i)].

We return to (4.53).

Since \( D_{ao} \) is left invertible and has rank \( q_{ao} \), we can write

\[
D_{ao} = U_{ao} G_{ao}
\]

(4.59)

where \( U_{ao}^T U_{ao} = I_{q_{ao}} \), \( G_{ao} \) is invertible.

Let \( U_{c} \) be such that \( U_{c}^T U_{c} = I_{r-q_{ao}} \) and such that \( U := [U_{ao}, U_{c}] \) is invertible, \( U^{-1} = U^T \).

Then for \( y_1 = U_{ao}^T y \) and \( y_2 = U_{c}^T y \) it follows immediately that

\[
y_1 = U_{ao}^T C x_\alpha + G_{ao} \bar{w}_\alpha
\]

(4.60a)

\[
y_2 = U_{c}^T C x_\alpha
\]

(4.60b)

but also that

\[ \| y \|^2 = \| y_1 \|^2 + \| y_2 \|^2 \]

(4.60c)

Applying a preliminary feedback

\[ \bar{w}_\alpha = G_{ao}^{-1} [-U_{ao}^T C x_\alpha + \hat{w}_\alpha] \]

(4.61)

then transforms (4.53a) and (4.60) into

\[
p x_\alpha = A_{ao} x_\alpha + \bar{B}_{ao} G_{ao}^{-1} \bar{w}_\alpha + B_{ao}^c \omega^c + x_0
\]

(4.62a)

\[ y_1 = \hat{w}_\alpha \]

(4.62b)

\[ y_2 = U_{c}^T C x_\alpha \]

(4.62c)

where

\[
A_{ao} = A - \bar{B}_{ao} G_{ao}^{-1} U_{ao}^T C
\]

\[ = A - \bar{B}_{ao} G_{ao}^{-1} \]

(4.63)

with

\[
D_{ao}^{+} = (D_{ao}^T D_{ao})^{-1} D_{ao}^T = (G_{ao}^T G_{ao})^{-1} G_{ao}^T U_{ao}^T
\]

\[ = G_{ao}^{-1} U_{ao}^T \]

(4.64)

Now both \( W(\Sigma) \) and \( V_d(\Sigma) \) turn out to be invariant w.r.t. \( A_{ao} \), the 'preliminary closed-loop' matrix. These invariances, that will be proven in the last two lemmas of this section, will
show their value in the development of Section 5.

At first, we state a result for the matrix $W_\alpha$ in (4.48).

**Lemma 4.2.**

$$A_\alpha (\text{im} (W_\alpha)) \subseteq \text{im} (W_\alpha).$$

Proof. See Lemma 2 in Appendix 1.

**Lemma 4.3.**

$$W(\Sigma) = \text{im} (W_\alpha), \dim W(\Sigma) = L_\alpha - 1.$$  

Proof. To start, it is stated in [2] that $W(\Sigma_\alpha) = W(\Sigma_{aux})$, where $\Sigma_{aux}$ is described by $(A_\alpha, B_\alpha^*, U^*_C)$.

From Lemma 3 in Appendix 1 it follows that $W(\Sigma_{aux}) = <A_\alpha, \text{im}(B_\alpha^*)>$, hence, with Lemma 4.2, $W(\Sigma_\alpha) \subseteq \text{im}(W_\alpha)$ since $\text{im}(B_\alpha^*) \subseteq \text{im}(W_\alpha)$ (4.54)).

Finally, iterating the equality in Prop. 4.1,

$$W(\Sigma) = \text{im}(W_\alpha) + W(\Sigma_\alpha) = \text{im}(W_\alpha)$$

and thus $\dim W(\Sigma) = \text{rank} (W_\alpha) = L_\alpha - 1$.

Observe that we can take as "output injection" $G = -\bar{B}_\alpha D_\alpha^+$ in [1, Th. 3.15] since $B + GD = \bar{B}_\alpha \bar{R}_0^T$.

Not only $W(\Sigma)$ is $A_\alpha$-invariant. So is $V_\alpha(\Sigma)$, according to Lemma 4.4.

**Lemma 4.4.**

$$V_\alpha(\Sigma) = <\ker (U^*_I C) \mid A_\alpha>.$$  

Proof. Consider (4.62), and apply Prop. 3.7:

$$x_0 \in V_\alpha(\Sigma) \iff \bar{\psi}_\alpha = 0 \land U^*_I C x_\alpha =$$

$$U^*_I C (pl - A_\alpha)^{-1} (\bar{B}_\alpha D_\alpha^{-1} \bar{\psi}_\alpha + B_\alpha \bar{\psi}_\alpha + x_0) = 0.$$
for some impulsive-smooth distribution $w^\alpha_a$.

However, since $\text{im } (CW_\alpha) \subseteq \text{im}(D_{\alpha_0})$ (see (A 1.5) in Appendix 1), one easily sees that, with Lemma 4.2,

$$U_\alpha^T (pl - A_{\alpha_0})^{-1} W_\alpha = 0.$$ 

Therefore, recalling (4.54), we have that

$$V_d(\Sigma_\alpha) = \{ x_0 \mid U_\alpha^T (pl - A_{\alpha_0})^{-1} x_0 = 0 \} = \langle \ker (U_\alpha^T C) \mid A_{\alpha_0} \rangle ,$$

but also (Prop. 4.1. (ii))

$$V_d(\Sigma) = V_d(\Sigma_\alpha) .$$

This completes the proof. Note that $V_d(\Sigma_\alpha) = V(\Sigma_\alpha)$, the weakly unobservable subspace for $\Sigma_\alpha$.

Remark

For all points in $\langle \ker (U_\alpha^T C) \mid A_{\alpha_0} \rangle$ the optimal cost without stability equals zero. This follows immediately from Prop. 3.7 and Lemma 4.4.

Summary

The generalized dual structure algorithm yields a transformed system (4.62), where $\text{im } (B^\alpha_a) \subseteq \text{im}(W_\alpha)$.

In addition, $A_{\alpha_0}(V_d(\Sigma)) \subseteq V_d(\Sigma)$ but also $A_{\alpha_0}(W(\Sigma)) \subseteq W(\Sigma)$.

These results will enable us to solve the $LQCP$ completely. This will be shown in Section 5.

We conclude this Section with an explicit description for the set of admissible inputs $U_\Sigma$ (Sec. 3), which obviously contains all optimal inputs for the $LQCP$ (if existent).

**Proposition 4.5.**

$$U_\Sigma = \left\{ u \in C_{\text{imp}}^m \mid u = \bar{H}_\alpha(p) \begin{bmatrix} f(\Phi_\alpha, w^\alpha_a) \\ w^\sigma_a \\ w^\bullet_a \end{bmatrix} \right\} .$$
\( \hat{w}_a \in C_{\text{imp}}^{\varepsilon a_p}, \text{ smooth; } \left[ \begin{array}{c} w_\alpha^c \\ w_\alpha^a \end{array} \right] \in C_{\text{imp}}^{m-\varepsilon a_p}, \text{ arbitrary } \)

where \( f(\hat{w}_a, w_\alpha^c) \) denotes a distribution in \( C_{\text{imp}}^{\varepsilon a_p} \), depending on \( \hat{w}_a \) and \( w_\alpha^c \), defined by

\[
\begin{align*}
  f(\hat{w}_a, w_\alpha^c) &= \left( D_{\alpha p}^T D_{\alpha p} \right)^{-1} \left( \hat{f}(p) G_{\alpha p}^{-1} \hat{w}_a - D_{\alpha p}^T C(p I - A_{a p})^{-1} B_{\alpha p} w_\alpha^c \right) \\
  \hat{f}(p) &= I - D_{\alpha p}^T C(p I - A_{a p})^{-1} \bar{B}_{a p}.
\end{align*}
\]

Proof. Immediate form (4.62), (4.61), (4.57) with \( x_0 = 0 \).

Remark

Proposition 4.5 is a generalization of \([1, (5.1), (5.2)]\).
5. Determination of all open-loop controls for the linear-quadratic problem without stability

For the solution of the \textit{LQCP} we start from (4.62):

\[
p x = A x + B x + x_0 \ , \quad (5.1a)
\]

\[
y_1 = w_a \ , \quad (5.1b)
\]

\[
y_2 = U_c C x_a \ , \quad (5.1c)
\]

and recall from (3.1c), (4.60c) that

\[
J(x_0, u) = \int_0^\infty [\|y_1\|^2 + \|y_2\|^2] dt \ . \quad (5.1d)
\]

Now make a direct sum decomposition of the state space as follows: let \( X_1 := W(\Sigma) \), let \( X_2 \) be a subspace such that \( X_1 \oplus X_2 = V_d(\Sigma) \) and let \( X_3 \) be a subspace such that \( X_1 \oplus X_2 \oplus X_3 = \mathbb{R}^n \).

Let \( W_{c_1} \) and \( W_{c_2} \) be left invertible matrices such that

\[
X_2 = \text{im}(W_{c_1}) \ , \quad X_3 = \text{im}(W_{c_2}) \ . \quad (5.2)
\]

Then

\[
\hat{W} := [W_a, W_{c_1}, W_{c_2}] \quad (5.3)
\]

is invertible with inverse

\[
\hat{W}^{-1} = \begin{bmatrix} L_a \\ L_{c_1} \\ L_{c_2} \end{bmatrix} \ . \quad (5.4)
\]

Decompose

\[
x_a = W_a x^a + W_{c_1} x_1^b + W_{c_2} x_2^b \ , \quad (5.5)
\]

i.e.,

\[
x^a = L_a x_a \ , \quad x_i^b = L_{c_i} x_a \quad (i = 1, 2) \quad (5.6)
\]

then (5.1) transforms into

\[
\begin{bmatrix} x^a \\ x_1^b \\ x_2^b \end{bmatrix} = \begin{bmatrix} \bar{A}_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix} \begin{bmatrix} x^a \\ x_1^b \\ x_2^b \end{bmatrix} + \begin{bmatrix} H_{1} \\ H_{2} \\ H_{3} \end{bmatrix} \begin{bmatrix} w_a \\ \hat{w}_a \end{bmatrix} + \begin{bmatrix} H_{1} \\ \hat{H}_1 \end{bmatrix} \begin{bmatrix} x^a_0 \\ x_1^{b0} \end{bmatrix} \quad (5.7a)
\]
\[
y_2 = \begin{bmatrix} 0 & 0 & C_3 \end{bmatrix} \begin{bmatrix} x^* \\ x^*_1 \\ x^*_2 \end{bmatrix}, \tag{5.7b}
\]

and

\[
J(x_0, u) = \int_{0}^{\infty} \left[ ||\hat{w}_d||^2 + ||y_d||^2 \right] dt. \tag{5.7c}
\]

Moreover, \((C_3, A_{33})\) is observable.

To see this, note that the zero blocks in the system matrix appearing in (5.7) follow from Lemmas 4.2 and 4.3. The other zero blocks in (5.7a) are a translation of (4.54).

Finally, the block decomposition for \(y_2\) and the observability of the pair \((C_3, A_{33})\) follow from Lemma 4.4.

Using (5.7), we may establish that the problem of infimizing \(J(x_0, u)\) in fact is determined by the regular subproblem:

Given the subsystem:

\[
\begin{bmatrix} x^*_1 \\ x^*_2 \end{bmatrix} = \begin{bmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{bmatrix} \begin{bmatrix} x^*_1 \\ x^*_2 \end{bmatrix} + \begin{bmatrix} H_2 \\ H_3 \end{bmatrix} \hat{w}_a + \begin{bmatrix} x^{*0}_1 \\ x^{*0}_2 \end{bmatrix}, \tag{5.8a}
\]

find

\[
\inf_{\hat{w}_a} \int_{0}^{\infty} \left[ ||\hat{w}_d||^2 + ||\begin{bmatrix} 0 & C_3 \end{bmatrix} \begin{bmatrix} x^*_1 \\ x^*_2 \end{bmatrix}||^2 \right] dt, \tag{5.8b}
\]

with \(\hat{w}_a \in C_{\text{imp}}^{\infty}\).

The regular linear-quadratic control problems are well established ([2], [12], [13], [18], [19]). It is generally agreed that one should compute the optimal solution of a regular problem by means of the Algebraic Riccati Equation associated with the system involved ([2], [12]). In order to ensure solvability of the problem under consideration, stabilizability of the system is commonly reassumed. Thus, we assume here

**Assumption 5.1.**

The pair \(\begin{bmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{bmatrix}, \begin{bmatrix} H_2 \\ H_3 \end{bmatrix}\) is stabilizable.
Sufficient for Assumption 5.1 to hold is the stabilizability of the pair \((A, B)\) ([14]).

Next, consider the Algebraic Riccati Equation corresponding to the subsystem in (5.8):

\[
0 = \begin{bmatrix} 0 & 0 \\ 0 & C^T C_3 \end{bmatrix} + \begin{bmatrix} A_{22}^T & 0 \\ A_{23}^T & A_{33}^T \end{bmatrix} \begin{bmatrix} \hat{K}_{11} & \hat{K}_{12} \\ \hat{K}_{12}^T & \hat{K}_{22} \end{bmatrix} + \begin{bmatrix} \hat{K}_{11} & \hat{K}_{12} \\ \hat{K}_{12}^T & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{bmatrix}
\]

and let \(\hat{K}^- = \begin{bmatrix} \hat{K}_{11}^- & \hat{K}_{12}^- \\ \hat{K}_{12}^- & \hat{K}_{22}^- \end{bmatrix}\) be the smallest non-negative definite solution of (5.9).

We will now state the main result of this Section.

**Theorem 5.2.**

Consider the LQCP without stability:

\[
\text{"Determine } J(x_0) := \inf_{c_{imp}} \int_0^\infty \|y\|^2 dt
\]

subject to \(px = Ax + Bu + x_0\),

\[
y = Cx + Du
\]

and let Assumption 5.1 hold.

Then

(i) \(J(x_0) = x^T K^- x_0\)

with \(K^- = [L_{c_1}^T, L_{c_2}^T] \hat{K}^- \left[ \begin{array}{c} L_{c_1} \\ L_{c_2} \end{array} \right] \)

\(\hat{K}^-\) being the smallest non-negative definite solution of (5.9).

(ii) If \(U_{\alpha}^{opt}(x_0)\) denotes the set of optimal controls for the LQCP without stability, then

\[
U_{\alpha}^{opt}(x_0) = \left\{ u \in C_{imp}^{m} \mid u = \tilde{H}_0(p) \begin{bmatrix} -g(K^-)B_\alpha w_\alpha + x_0 \\ w^c_\alpha \\ w^*_\alpha \end{bmatrix} \right\},
\]
where $g(K^-)$ is a matrix-valued distribution defined by

$$g(K^-) := (\mathcal{D}_{a_p}^T \mathcal{D}_{a_p})^{-1} (\mathcal{D}_{a_p}^T C + \mathcal{B}_{a_p}^T K^-) (pI - A_{a_p}^{-1})^{-1}$$

with

$$A_{a_p}^{-1} = A_{a_p} - \mathcal{B}_{a_p} (\mathcal{D}_{a_p}^T \mathcal{D}_{a_p})^{-1} \mathcal{B}_{a_p}^T K^-.$$  

Consequently, there is generally more than one optimal trajectory. However, for any optimal trajectory, to be denoted by $x^*$, it turns out that $L_{c_l}x^*$ ($l = 1, 2$) is independent of $x^*$.

(iii) Moreover, $\bar{K}^- = \begin{bmatrix} 0 & 0 \\ 0 & \bar{K}_{22}^- \end{bmatrix}$, with $\bar{K}_{22}^- > 0$.

and, additionally,

$$A_{33}^* = A_{33} - \bar{H}_3 \bar{H}_3^T \bar{K}_{22}^-$$

is asymptotically stable.

Proof. The solution of the subproblem stated in (5.8) is given by ([2]):

$$\inf_{\mathcal{C}_{a_p}} \int_0^\infty [||\dot{\mathcal{Y}}_a_d||^2 + ||C_3 x_3^2||^2 dt] = [x_1^{b \sigma}, x_2^{b \sigma}] \bar{K}^- \begin{bmatrix} x_1^{b \sigma} \\ x_2^{b \sigma} \end{bmatrix}$$

this infimum is achieved by $\dot{\mathcal{Y}}_a = -[H_2^T, H_3^T] \bar{K}^- \begin{bmatrix} x_1^{b} \\ x_2^{b} \end{bmatrix}$.

Hence (5.7)

$$J(x_0) = x_0^T \bar{K}^- x_0,$$

and $w_a^c, w_{a}^*$ may be chosen completely arbitrarily.

Substituting

$$\dot{\mathcal{Y}}_a = -[H_2^T, H_3^T] \bar{K}^- \begin{bmatrix} x_1^{b} \\ x_2^{b} \end{bmatrix}$$
into (5.1a) yields

\[ px_a = (A_{ap} - B_{ap}G_{ap}^{-1}(G_{ap}^{-1})^T B_{ap}^TK^-)x_a + B_{ap}w_c + x_0 \]

\[ = A_{ap}^*x_a + B_{ap}w_c + x_0, \text{ see (4.64)} . \]

Hence

\[ x_a = (pI - A_{ap})^{-1}(B_{ap}w_c + x_0) . \] (5.13)

From (4.61), (5.12), (4.64),

\[ \bar{w}_a = -((D_{ap}^T D_{ap})^{-1}[D_{ap}^T C + B_{ap}^T K^-]) x_a \]

and then (5.10) follows with (4.57), (5.13).

Next, from (4.55), (5.4), (5.6),

\[ \begin{cases} L_a x = x^o + \bar{x}_a \bar{a}^- \\ L_{c_i} x = x^{b_i} (i = 1, 2) \end{cases} . \] (5.14)

Hence, if \( x^{b_i} (i = 1, 2) \) denotes the optimal trajectory for (5.8) obtained by the minimizing feedback law for \( \bar{w}_a \), then \( L_{c_i} x^* = x^{b_i} (i = 1, 2) \) for any optimal trajectory \( x^* \) and therefore is independent of \( x^* \). We will elaborate extensively on the non-uniqueness of \( L_{c_i} x^* \) in the future article announced earlier.

This completes the proof of (i) and (ii).

Finally, from Prop. 3.7,

\[ \mathcal{V}_d(\Sigma) = \{ x_0 = W_d x^0 + W_c x^{b_1} + W_c x^{b_2} \mid \begin{bmatrix} x_1^{b_1} \\ x_2^{b_1} \end{bmatrix} \in \ker \tilde{K}^- \} \]

\[ = \im [W_d, W_c], \text{ whence} \]

\[ \tilde{K}^- = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{K}_{22} \end{bmatrix} , \] (5.15)

with \( \tilde{K}_{22} \) satisfying ((5.9), (5.15))

\[ 0 = C_3^T C_3 + A_{33}^T \tilde{K}_{22} + \tilde{K}_{22} A_{33} - \tilde{K}_{22} H_3 H_3^T \tilde{K}_{22} . \] (5.16)

Due to the observability of \((C_3, A_{33})\), we have that \( \tilde{K}_{22} > 0 \) and furthermore that \( A_{33}^T \) is asymptotically stable.
This proves (iii). Note that $A_{33}^{-1} = L_{c}A_{d}^{*}W_{c}$. 

**Remarks**

1. Assumption 5.1 shows that $(A, B)$-stabilizability is not necessary for solvability of the LQCP, even not for solvability of the LQCP with stability, as will be shown in a future paper. This contradicts [2, Th. 3, (i)].

2. Since $K$ is of the form given in (5.14), we might as well have restricted ourselves to the regular subproblem

$$px_2^b = A_{33}x_2^b + H_3\dot{w}_a + x_2^{b0}$$

and infimization of

$$\int_0^\infty \left[\|\dot{w}_a\|^2 + \|C_3x_2^b\|^2\right] dt$$

which is solvable if $(A_{33}, H_3)$ is stabilizable.

3. From (5.14) we immediately see that necessary for $x(\infty) = 0$ is: $x_l^b(\infty) = 0$ ($l = 1, 2$). Therefore Assumption 5.1 is necessary for solvability of the LQCP with stability (see Remark 1), whereas it is sufficient for solvability of the LQCP treated in the present article (see Remark 2).

4. In the future article mentioned in Remark 1, it will turn out that $x_l^b(\infty) = 0$ ($l = 1, 2$) can be also sufficient for $x(\infty)$ to be zero.

We close this Section with the interpretation of Theorem 5.2 for left and right invertible systems (see Sec. 3). Therefore, let Assumption 5.1 hold.

**Lemma 5.3.**

The transfer function $T(s)$ has rank $g_{\alpha_0}$ over the field of rational functions.

Proof. From (4.53), rank $(T_{\alpha}(s)) = g_{\alpha_0}$ since rank $(Q_{\alpha_0}) = g_{\alpha_0}$ and im $(CB_\alpha) \subseteq$ im $(Q_{\alpha_0})$ (App. 1). Hence ((4.58)) rank $(T(s)) = \text{rank}(T_{\alpha}(s)\tilde{H}_\alpha^\alpha(s)) = \text{rank}(T(s)) = g_{\alpha_0}$.

In [16, Th. 4.3] this result was proven by means of a "left structure algorithm".

In Sec. 3, Remark 1, it was stated that for left invertible systems the optimal control, if existent, is unique. Indeed
**Corollary 5.4.**
\[ \Sigma \text{ left invertible } \iff \varrho_{\alpha_D} = m \iff \# (U_{\tilde{x}_0} (x_0)) = 1. \]

**Proof.** From Prop 3.12, Lemma 5.3, (5.10a): \( \Sigma \) left invertible \( \iff m = \varrho_{\alpha_D} \iff \dim (w_{\alpha_D}^\perp) = 0 \) and \( \dim (w_{\alpha_D}^\perp) = 0 \iff \) optimal control is unique.

**Remarks.**
1. Note that for left invertible systems \( \alpha = \alpha_D \).
2. The dual algorithm in Sec. 4 and the algorithm in [1, Sec. 4] are identical for a left invertible system.

**Corollary 5.5.**
\[ \Sigma \text{ right invertible } \iff \varrho_{\alpha_D} = r \iff K^- = 0. \]

**Proof.** Combine Prop. 3.9, Lemmas 3.10, 5.3 and Theorem 5.2 (i).

**Remark.**
For a right invertible system, \( W_{\varepsilon_2} \) in (5.3) does not appear and \( \hat{L} = \begin{bmatrix} L_{\alpha} \\ L_{\varepsilon_1} \end{bmatrix} \) in (5.4). Therefore \( x_2^\varepsilon \) and \( y_2 \) do not appear either.
6. The linear-quadratic control problem and the dissipation inequality

In [18] it was shown that a necessary condition for the quadratic form \( x^T K x \) to represent \( \inf_{u} J(x_0, u) \) under any conditions on the long-term behaviour of the state is that the real symmetric matrix \( K \) satisfies the dissipation inequality
\[
F(K) \geq 0. \tag{6.1}
\]
Here \( F(K) \) is called the dissipation matrix ([17]), which for any \( n \times n \) matrix \( K \) is defined by
\[
F(K) = \begin{bmatrix}
A^T K + KA + C^T C & KB + C^T D \\
B^T K + D^T C & D^T D
\end{bmatrix}. \tag{6.2}
\]

The dissipation inequality has been a topic of interest in several papers since its introduction, for instance in [18]. There it was noted that for the regular LQCP all rank minimizing solutions of (6.1) are solutions of the Algebraic Riccati Equation
\[
0 = C^T C + A^T K + KA - (KB + C^T D)(D^T D)^{-1}(B^T K + D^T C).
\]
Recently, it was shown in [17] that for the singular LQCP with stability the symmetric matrix defining the optimal cost, denoted by \( K^* \), also is a rank minimizing solution of (6.1).

In this section we generalize the results in [17]. Here, it will be shown that for real symmetric matrices \( K: \mathbb{R}^{np} (\text{rank } F(K)) = \text{rank } (T(s)) ((3.3)) \) and that the rank minimizing solutions of (6.1) are solutions of a specified Algebraic Riccati Equation. Thus, in particular, \( K \) in Th. 5.2 turns out to be the smallest non-negative rank minimizing solution of (6.1).

First, observe that with every system we may associate a dissipation matrix. Now let \( F_k(K) \) be the dissipation matrix belonging to the system \( \Sigma_k \) in Sec. 4, \( k = 0, 1, \cdots, \alpha \). Recalling Prop. 4.1, we work with descriptions (4.1) for \( \Sigma_0 = \Sigma \), (4.25) for \( \Sigma_{l+1}, l = 0, 1, \cdots, (\alpha - 2) \), and (4.53) for \( \Sigma_{\alpha} \).

Then lemma 6.1 expresses the key result.

**Lemma 6.1**

Let \( i = 0, 1, \cdots, (\alpha - 1) \) and consider step \( (i + 1) \), part 1 of the algorithm.

Then
\[
F_i(K) \geq 0 \iff F_{i+1}(K) \geq 0 \text{ and } K\bar{B}_i = 0,
\]
and, additionally,
\[
\text{rank } (F_i(K)) = \text{rank } (F_{i+1}(K)).
\]
Consequently ((4.48)),
\[ F(K) \geq 0 \iff F_\alpha(K) \geq 0 \text{ and } KW_\alpha = 0 \] (6.3)
and
\[ \operatorname{rank} (F(K)) = \operatorname{rank} (F_\alpha(K)) . \] (6.4)

Proof. Appendix 2.

According to lemma 6.1, we can concentrate on the inequality \( F_\alpha(K) \geq 0 \) in order to find the set of solutions for (6.1). Using (6.3), (6.4), it is then immediate that
\[ F(K) \geq 0 \iff F_\alpha(K) \geq 0 \text{ and } KW_\alpha = 0 \] (6.5)
and that
\[ \operatorname{rank} (F(K)) = \operatorname{rank} (F_\alpha(K)) , \] (6.6)
with
\[ F_\alpha(K) = \begin{bmatrix} A^T K + K A + C^T C & K E_{\alpha_\rho} + C^T D_{\alpha_\rho} \\ E_{\alpha_\rho}^T K + D_{\alpha_\rho}^T C & D_{\alpha_\rho}^T D_{\alpha_\rho} \end{bmatrix} . \] (6.7)

Since (Schur's lemma) \( F_\alpha(K) \) is similar to
\[ \begin{bmatrix} \Phi(K) & 0 \\ 0 & D_{\alpha_\rho}^T D_{\alpha_\rho} \end{bmatrix} , \] with
\[ \Phi(K) := A^T K + K A + C^T C - (K E_{\alpha_\rho} + C^T D_{\alpha_\rho}) (D_{\alpha_\rho}^T D_{\alpha_\rho})^{-1} (\overline{E}_{\alpha_\rho} K + \overline{D}_{\alpha_\rho} C) . \] (6.8)
we thus obtain

Theorem 6.2.

Let \( M_\mathbb{R}(n) \) denote the set of real symmetric \( n \times n \) matrices and \( \Gamma := \{ K \in M_\mathbb{R}(n) \mid F(K) \geq 0 \} \). Then
\[ \Gamma = \{ K \in M_\mathbb{R}(n) \mid KW_\alpha = 0 \text{ and } \Phi(K) \geq 0 \} . \] (6.9)
Moreover, for every \( K \in \Gamma \) it holds that
\[ \operatorname{rank} (F(K)) = \alpha_\rho + \operatorname{rank} (\Phi(K)) . \]
\[ = \operatorname{rank} (T(s)) + \operatorname{rank} (\Phi(K)) . \] (6.10)
Hence, if $\Gamma_{\min}$ denotes the subset of $\Gamma$ containing all rank minimizing solutions of the inequality

$$\min_{K \in \mathcal{M}_R(n)} (\text{rank } F(K)) = q_{\alpha_D} = \text{rank } (T(s))$$

and

$$\Gamma_{\min} = \{ K \in \Gamma \mid \Phi(K) = 0 \}$$

$$= \{ K \in \mathcal{M}_R(n) \mid KW_\alpha = 0 \text{ and } \Phi(K) = 0 \}.$$

(6.11)

Proof. Statements (6.9), (6.10) follow from (6.5) - (6.8) and Lemma 5.3. Further, rank $\Phi(K)) \geq 0$ and rank $\Phi(K)) = 0 \iff \Phi(K) = 0$.

The Riccati Equation $\Phi(K) = 0$ can be transformed into (5.9) if $KW_\alpha = 0$:

**Corollary 6.3.**

$$\Gamma_{\min} = \{ K \in \mathcal{M}_R(n) \mid K = [L_{c_1}^T, L_{c_2}^T \ K] \tilde{K} \begin{bmatrix} L_{c_1} \\ L_{c_2} \end{bmatrix}, \tilde{K} \text{satisfies (5.9)} \}.$$

Proof. The Riccati Equation $\Phi(K) = 0$ is equivalent to

$$0 = C^T U_c U_c^T C + A_{\alpha_D}^T K + K A_{\alpha_D} - K B_{\alpha_D} G_{\alpha_D}^{-1} (G_{\alpha_D}^{-1})^T B_{\alpha_D}^T K$$

(see (5.1)). Define $K = [L_{a}^T, L_{c_1}^T, L_{c_2}^T ]\tilde{K} \begin{bmatrix} L_{a} \\ L_{c_1} \\ L_{c_2} \end{bmatrix}$.

Then for all $K \in \mathcal{M}_R(n)$:

$$KW_\alpha = 0 \iff \tilde{K} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \tilde{K}_{11} & \tilde{K}_{12} \\ 0 & \tilde{K}_{12} & \tilde{K}_{22} \end{bmatrix}.$$

Thus $\Phi(K) = 0$ and $KW_\alpha = 0 \iff \tilde{K} = \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} \\ \tilde{K}_{12} & \tilde{K}_{22} \end{bmatrix}$ satisfies (5.9).

Finally, combining Th. 5.2 and the results obtained here, yields
Corollary 6.4
Let $K^-$ denote the matrix in $M_R(n)$ denoting the optimal cost without stability. Then $K^-$ can be characterized as the smallest non-negative definite rank minimizing solution of (6.1).

Comment
Corollary 6.4 is in fact a characterization of the optimal cost for the LQCP without stability directly related to the coefficients of the original system, whereas e.g. Theorem 5.2 implicitly preassumes the knowledge of the system $\Sigma_o$ obtained by the dual algorithm.

On the other hand we emphasize that the rank minimizing procedure actually is equivalent to the column generating process in the algorithm.
Conclusions

The generalized dual structure algorithm is an appropriate instrument to compute all optimal controls for the singular linear-quadratic control problem without stability. Also, it has enabled us to give an elegant characterization of all rank minimizing solutions of the dissipation inequality. In particular we have proven that the optimal cost for the problem considered in this paper can be interpreted as the smallest non-negative rank minimizing solution of the dissipation inequality.

Acknowledgement

The author expresses his gratitude to Prof. M.L.J. Hautus who suggested the subject of the paper and to Dr. H.L. Trentelman who was always available for discussion and who came up with valuable criticism.
Appendix 1

Lemma 1

Consider the dual algorithm. Then we have the following relations:

(i) \[ \forall i = 1 \ldots \alpha : \tilde{B}_i = A_{\alpha D} \tilde{B}_{i-1} \tilde{P}_{i-1} \tilde{V}_{i-1} \tilde{R}_i , \]

(ii) \[ \forall i = 1 \ldots \alpha : \tilde{B}_i \tilde{R}_i^T = A_{\alpha D} \tilde{B}_{i-1} \tilde{P}_{i-1} \tilde{V}_{i-1} \]

(iii) \[ \forall i = 0 \ldots (\alpha_D - 1) : C \tilde{B}_i \tilde{P}_i \tilde{V}_i = Q_{i+1} \begin{bmatrix} K_{i+1}^* \end{bmatrix} \begin{bmatrix} \tilde{R}_{i+1}^T \end{bmatrix} + \begin{bmatrix} 0 \end{bmatrix} , \]

\[ \forall i = (\alpha_D - \alpha) : C \tilde{B}_i \tilde{P}_i \tilde{V}_i = Q_{i+1} K_{i+1}^* , \]

(iv) \[ \forall i = 1 \ldots (\alpha - 1) : \tilde{B}_i = W_{i+1} \begin{bmatrix} 0 \\ \tilde{V}_i \end{bmatrix} + \begin{bmatrix} N_i^* \tilde{V}_i^T \\ 0 \end{bmatrix} Q_i^* , \]

\[ \tilde{B}_\alpha = W_{\alpha} N_{\alpha} Q_{\alpha}^* , \]

with \( W_{i+1} = [\tilde{B}_0, \tilde{B}_1 \tilde{P}_1 \tilde{V}_1, \ldots, \tilde{B}_i \tilde{P}_i \tilde{V}_i] \)

and \( Q_i^* \) is such that

\[ \begin{bmatrix} Q_i^* \\ \tilde{P}_i^T \end{bmatrix} = P_i^{-1} . \]

Proof.

We start with the observation

\[ \forall i = 1 \ldots \alpha_D : \tilde{B}_i = -\tilde{E}_{i} \begin{bmatrix} K_{i}^* \\ K_{i+1}^* \end{bmatrix} + A\tilde{B}_{i-1} \tilde{P}_{i-1} \tilde{V}_{i-1} \tilde{R}_i , \]

\[ \forall i = (\alpha_D + 1) \ldots \alpha : \tilde{B}_i = -\tilde{E}_{\alpha D} K_{i+1}^* + A\tilde{B}_{i-1} \tilde{P}_{i-1} \tilde{V}_{i-1} , \]

which follows from (4.12), (4.30), (4.51). Note that \( \tilde{R}_i = I_{i+1} \) for \( i \geq (\alpha_D + 1) \).

Also, one easily sees from parts 1 in each step that
\[ \forall i = 1 \cdots \alpha_D : {\cal D}_i \begin{bmatrix} K_{1_i}^* \\ K_{2_i}^* \end{bmatrix} = C \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} \bar{R}_i \, , \]

\[ \forall i = (\alpha_D + 1) \cdots \alpha : {\cal D}_{\alpha_D} K_{1_i}^* = C \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} \, . \]  

(A1.2)

Thus, for \( i = (\alpha_D + 1) \cdots \alpha \),

\[ K_i := K_{1_i}^* = D_{\alpha_D}^* C \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} \, , \quad ((4.64)) \]

which, substituted in (A1.1), yields

\[ \tilde{B}_i = (A - \tilde{E}_{\alpha_D} \bar{D}_{\alpha_D}^* C) \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} = A_{\alpha_D} \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} \, , \]

see (4.63).

In addition, (A1.2) can be written for \( i = 1 \cdots \alpha_D \) as

\[ {\cal D}_i \begin{bmatrix} K_{1_i}^* \\ K_{2_i}^* \end{bmatrix} = [D_i, D_{i+1}, \cdots, D_{\alpha_D}] K_i = D_{\alpha_D} K_i \]

\[ = C \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} \bar{R}_i \]

with \( K_i = \begin{bmatrix} K_{1_i}^* \\ K_{2_i}^* \\ 0 \end{bmatrix} \) a \( q_{\alpha_D} \times (r_{i-1} - q_i) \) matrix.

Consequently,

\[ K_i = D_{\alpha_D}^* C \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} \bar{R}_i \]

and hence ((A1.1))

\[ \tilde{B}_i = -\tilde{E}_{\alpha_D} K_i + A \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} \bar{R}_i = A_{\alpha_D} \bar{B}_{i-1} \bar{F}_{i-1} \bar{V}_{i-1} \bar{R}_i \, ; \]

we have proven (i).

Since (i) and (ii) are equal for \( i = (\alpha_D + 1) \cdots \alpha \), we take \( i = \alpha_D \), and rewrite a trivial equality:

\[ \tilde{E}_{\alpha_D} = \tilde{E}_{\alpha_D} D_{\alpha_D}^* D_{\alpha_D} = \tilde{E}_{\alpha_D} D_{\alpha_D}^* [D_{\alpha_D - 1}, D_{\alpha_D}] = \]

\[ = \tilde{E}_{\alpha_D} D_{\alpha_D}^* [D_{\alpha_D - 1}, C \bar{B}_{\alpha_D - 1} \bar{F}_{\alpha_D - 1} \bar{V}_{\alpha_D - 1} \bar{R}_{\alpha_D}] \]

\[ = [\tilde{E}_{\alpha_D} D_{\alpha_D}^* D_{\alpha_D - 1}, \tilde{E}_{\alpha_D} D_{\alpha_D}^* C \bar{B}_{\alpha_D - 1} \bar{F}_{\alpha_D - 1} \bar{V}_{\alpha_D - 1} \bar{R}_{\alpha_D}] \]
Then

\[
(A1.3) \quad \begin{align*}
\overline{e}_{a_0^{-1}} &= \overline{e}_{a_0} D_{a_0} D_{a_0}^{-1}, \\
A_{a_0} \overline{e}_{a_0^{-1}} \overline{v}_{a_0^{-1}} &= 0
\end{align*}
\]

where we have used the second equality from (A1.3) and (i).

This process is now continued by exploiting the first result of (A1.3) where we set 
\[D_{a_0^{-1}} = [D_{a_0^{-2}}, D_{a_0^{-1}}], \overline{e}_{a_0^{-1}} = [\overline{e}_{a_0^{-2}}, \overline{e}_{a_0^{-1}}], \text{etc.}
\]

It turns out that

\[
\forall i = 1 \cdots a : B_{a_0^{-i}} = \overline{e}_{a_0} D_{a_0} D_{a_0}^{-1}
\]

and therefore (ii).

The proof of (iii) is immediate from (A1.2) and 
\[B_{a_0^{-1}} \overline{v}_{a_0^{-1}} = B_{a_0^{-1}} \overline{v}_{a_0^{-1}} (\overline{v}_{a_0^{-1}} \overline{R}_{a_0^{-1}} + \overline{v}_{a_0^{-1}} \overline{R}_{a_0^{-1}}^T).
\]

Note that (iii) implies that

\[
\text{im} (C W_a) \subseteq \text{im} (D_{a_0}) .
\]

Finally, we show (iv).

With (4.45) we find that

\[
\begin{align*}
\hat{B}_a &= B_{a_0} (\overline{e}_{a_0} Q_{a_0}^* + P_{a_0}^* \hat{P}_{a_0}^T) = W_a N_{a_0} Q_{a_0}^* \quad .
\end{align*}
\]

Next,

\[
\begin{align*}
\hat{B}_{a^{-1}} &= B_{a_0^{-1}} \overline{e}_{a_0^{-1}} Q_{a_0}^* = B_{a_0^{-1}} \overline{e}_{a_0^{-1}} (\overline{v}_{a_0^{-1}} \overline{v}_{a_0^{-1}}^T + \overline{v}_{a_0^{-1}} \overline{v}_{a_0^{-1}}^T) Q_{a_0}^* \\
&= (\hat{B}_{a_0^{-1}} \overline{v}_{a_0^{-1}} \overline{v}_{a_0^{-1}}^T + \hat{N}_{a_0} \overline{v}_{a_0^{-1}}^T) Q_{a_0}^*
\end{align*}
\]

\[
= W_a \begin{bmatrix}
0 \\
\overline{v}_{a_0^{-1}}^T
\end{bmatrix} + \begin{bmatrix}
N_{a_0}^* \overline{v}_{a_0^{-1}}^T \\
0
\end{bmatrix}
\]

and it is clear that in this way (iv) can be proven.
Lemma 2.

\[ A_{\alpha_0} W_{\alpha} = W_{\alpha} \tilde{A}_{11} \]

with

\[
\tilde{A}_{11} = \begin{bmatrix}
N_1^* Z_1 & N_2^* Z_2 & N_{a-1}^* Z_{a-1} & N_{a_1}^* Z_a \\
\Lambda_1 & N_2^* Z_2 & N_{a_2}^* Z_a \\
\Lambda_2 & N_{a_2}^* Z_a \\
\Lambda_{a-1} & N_{a_{a-1}}^* Z_{a-1} & N_{a_{a-1}}^* Z_a \\
\end{bmatrix}
\]

where

\[ Z_i = \tilde{V}_i^T \tilde{Q}_i^* \tilde{R}_i^T, \quad i = 1 \cdots \alpha, \]

\[ \Lambda_i = \tilde{V}_i^T \tilde{Q}_i^* \tilde{R}_i^T, \quad i = 1 \cdots (\alpha - 1) \]

and

\[ N_i^* = \begin{bmatrix}
N_{i_1}^* \\
N_{i_2}^* \\
\vdots \\
N_{i_{\alpha}}^* \\
\end{bmatrix}, \quad 1 = 1 \cdots \alpha, N_{i_j}^* a \ t_j - 1 \times \rho, \text{ matrix } N_{11}^* = N_1^* \]

Proof.

Immediate from (ii), (iv) of Lemma 1.

Lemma 3.

Given \( \Sigma_{aux} \):

\[ px_{aux} = A_{\alpha_0} x_{aux} + B_{\alpha}^c u_{aux} + x_0, \]

\[ y_{aux} = U_0^T C x_{aux}, \]

with \( \text{im}(B_{\alpha}^c) \subseteq \text{im}(W_{\alpha}) \).

Then

\[ W(\Sigma_{aux}) = \langle A_{\alpha_0} | \text{im}(B_{\alpha}^c) \rangle. \]
Proof.

It holds that \( W(\Sigma_{\text{sum}}) = W_n \) where \( W_n \) is defined inductively by

\[
W_0 := \{0\},
\]

\[
W_{i+1} := \text{im} (B_0^\alpha) + A_{\alpha_i}W_i \cap \ker(U_i^TC),
\]

see [1, (3.22)].

Thus \( W_1 = \text{im} (B_0^\alpha) \). Further, since \( U_i^TC(pI - A_{\alpha_i})^{-1}W_a = 0 \), \( W_2 = \text{im} (B_0^\alpha) + A_{\alpha_i} \text{im} (B_0^\alpha) \), etc., and finally,

\[
W_n = \langle A_{\alpha_{i+1}}, \text{im} (B_0^\alpha) \rangle.
\]

Remark

Note that \( U_i^TC(pI - A_{\alpha_i})^{-1}W_a = 0 \) is a consequence of Lemma 1 (i) and Lemma 2.
Appendix 2

Proof of Lemma 6.1.

Let $F_k(K)$ denote the dissipation matrix corresponding with $\Sigma_k$, $F_0(K) := F(K)$. We will only prove the first two equivalencies to indicate the inductive process. Assume that $F(K) \geq 0$. Then premultiplication of $F(K)$ by $\begin{bmatrix} I_n & 0 \\ 0 & S_0 \end{bmatrix}$ and postmultiplication by $\begin{bmatrix} I_n & 0 \\ 0 & S_0 \end{bmatrix}$ yields ((4.2), (4.3))

$$0 \leq \begin{bmatrix} \tilde{F}_0(K) & [KB_0] \\ [B_0K & 0] & 0 \end{bmatrix},$$

whence $KB_0 = 0$ and $\tilde{F}_0(K) \succeq 0$ with

$$\tilde{F}_0(K) = \begin{bmatrix} A^TK + KA + C^TC & KB_0 + C^TD_0 \\ B_0^TK + D_0^TC & D_0^TD_0 \end{bmatrix}.$$  \hfill (A2.1)

Moreover, rank $(F(K)) = \text{rank } (\tilde{F}_0(K))$.

Now also rank $(\tilde{F}_0(K)) = \text{rank } (M_0 \tilde{F}_0(K) M_0)$ for any right invertible matrix; we propose

$$M_0 = \begin{bmatrix} I_n & 0 & \tilde{B}_0 \\ 0 & I_{q_0} & 0 \end{bmatrix}.$$ Then it turns out that ((4.8))

$$0 \leq M_0 \tilde{F}_0(K) M_0 = \begin{bmatrix} \tilde{F}_0(K) & [KAB_0 + C^TCB_0] \\ [B_0^TA^TK + B_0^TC^TC, B_0^TC^TD_0] & [D_0^TCB_0] \end{bmatrix} = F_1(K).$$ \hfill (A2.3)

Thus $F(K) \geq 0 \iff F_1(K) \geq 0$ and $KB_0 = 0$. Also rank $(F(K)) = \text{rank } (F_1(K))$. Conversely, if $F_1(K) \geq 0$ and $KB_0 = 0$ then from (A2.3) immediately $\tilde{F}_0(K) \succeq 0$, hence ((A2.1)) $F(K) \geq 0$.

Next, let $F_1(K) \succeq 0$, i.e.

$$0 \leq \begin{bmatrix} \tilde{F}_1(K) & [KB_1] \\ [B_1^T[0, 0] & 0] & 0 \end{bmatrix},$$

((4.11), (4.12)), then $KB_1 = 0$ and $\tilde{F}_1(K) \succeq 0$, where
\[
\mathcal{F}_1(K) = \begin{bmatrix}
A^T K + KA + C^T C & KB_0 + C^T D_0 & KB_1 + C^T D_1 \\
B_0^T K + D_0^T C & D_0^T D_0 & D_1^T D_1 \\
B_1^T K + D_1^T C & D_1^T D_0 & D_1^T D_1
\end{bmatrix}.
\] (A2.5)

and rank \( (F_1(K)) = \text{rank} (\mathcal{F}_1(K)) \). Obviously, \( KB_1 \mathcal{F}_1 \mathcal{V}_1 = 0 \), \( KB_1 \mathcal{F}_1 \mathcal{V}_1 = 0 \). Also, \( \text{rank} (\mathcal{F}_1(K)) = \text{rank} (M_1^T \mathcal{F}_1(K) M_1) \) with the right invertible matrix

\[
M_1 = \begin{bmatrix}
I_n & 0 & 0 & \tilde{B}_1 \mathcal{F}_1 \mathcal{V}_1 \\
0 & I_{d_0} & 0 & 0 \\
0 & 0 & I_{d_1} & 0
\end{bmatrix},
\]

while \( M_1^T \mathcal{F}_1(K) M_1 = \)

\[
\begin{bmatrix}
\mathcal{F}_1(K) & \mathcal{K} B_1 \mathcal{F}_1 \mathcal{V}_1 + C^T C B_1 \mathcal{F}_1 \mathcal{V}_1 \\
\tilde{V}_1^T \mathcal{F}_1 \mathcal{B}_1 A^T K + \tilde{V}_1^T \mathcal{F}_1 \mathcal{B}_1 C^T C, \tilde{V}_1^T \mathcal{F}_1 \mathcal{B}_1 C^T D_0, \tilde{V}_1^T \mathcal{F}_1 \mathcal{B}_1 C^T D_1 \\
\tilde{V}_1^T \mathcal{F}_1 \mathcal{B}_1 \mathcal{F}_1 \mathcal{V}_1 & \tilde{V}_1^T \mathcal{F}_1 \mathcal{B}_1 \mathcal{F}_1 \mathcal{V}_1
\end{bmatrix}
\] (A2.6)

Since (4.22), (A2.6) \( F_2(K) = \)

\[
\begin{bmatrix}
\mathcal{K} B_1 \mathcal{F}_1 \mathcal{V}_1 \\
0 \\
0 \\
0
\end{bmatrix}
\]

we thus have shown that

\[
F_1(K) \geq 0 \iff F_2(K) \geq 0 \text{ and } KB_1 = 0.
\]

On the other hand, if \( F_2(K) \geq 0 \) and \( KB_1 = 0 \) then from (A2.6) \( F_1(K) \geq 0 \) and hence ((A2.4)) \( F_1(K) \geq 0 \). Furthermore rank \( (F_1(K)) = \text{rank} (F_2(K)) \).

Note also that \( K [\tilde{B}_0, \tilde{B}_1] = 0 \iff K [\tilde{B}_0, \tilde{B}_1 \mathcal{F}_1 \mathcal{V}_1] = 0 \).

Now in general (\( i = 0, 1, \cdots, (\alpha - 1) \))

\[
F_i(K) \geq 0 \iff F_{i+1}(K) \geq 0 \text{ and } KB_i = 0
\]

and therefore

\[
F(K) \geq 0 \iff F_0(K) \geq 0 \text{ and } KW_0 = 0
\]

since
Moreover

\[ K[\bar{B}_0, \bar{B}_1, \ldots, \bar{B}_{\alpha-1}] = 0 \iff KW_\alpha = 0. \]

Moreover

\[ \text{rank } (F(K)) = \text{rank } (F_\alpha(K)) \]
References


