Structured Linear Systems - Realization and Control

J Zeelie

orcid.org/0000-0002-8853-2997

Thesis accepted in fulfilment of the requirements for the degree *Doctor of Philosophy in Applied Mathematics* at the North-West University

Promoter: Prof S Ter Horst

Graduation May 2023

24698245

This thesis is dedicated to my father Ben Zeelie who passed away on 1 April 2017.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

The financial support of the National Graduate Academy for Mathematical and Statistical Sciences (NGA-MASS) for two time buy outs through their University Capacity Development Programme (UCDP) is also hereby acknowledged.

Acknowledgements

The journey towards the completion of this thesis has been one of the most difficult I have ever undertaken. There were challenges and difficult times along the way. I certainly would not have been able to finish the race without the help and support of certain key individuals. As such I would like to extend my sincere gratitude toward those to whom it is due.

Firstly to my supervisor Prof. Sanne ter Horst. You are certainly one of the most hard working and dedicated people I have ever met. Thank for the example you have set for me in this regard. Thank you for encouraging me to keep on trying when I was discouraged. Thank you for your patience in times when progress was slow. Thank you for your leadership and mentoring and developing me as a mathematician and researcher.

I am deeply thankful toward my dear mother for her encouragement, friendship and support, especially during the hard times. To all my friends who supported me during my journey and helped me keep my sense of humour: Ben, Arnoh, Tiaan, Guan, Anansa, Boeshie, Karnu, Job, P.J., Pierre, Alma, Madelein, Nicojan, Gideon, René. A special word of thanks is also due to oom Gert who always welcomed me to their farm when I needed to take break and regain my strength.

Finally I want to thank my God and heavenly Father. Every breath is a gift and I consider my very existence something to be grateful for. I want to exclaim with the psalmist that "The heavens are telling of the glory of God; And their expanse is declaring the work of His hands" (Psalm 19:1). I am also reminded of the words of saint Paul: "His eternal power and divine nature, have been clearly seen, being understood through what has been made". I praise the Creator for endowing being with beauty, order and goodness. I consider it a great privilege to have the opportunity to investigate the mathematical structure and beauty of creation and for that I am deeply grateful.

Abstract

We consider decentralized linear time-invariant (LTI) systems which consist of several interconnected subsystems. The communication structure between the subsystems is reflected by a block-zero pattern in the system matrices. Such block zero patterns are determined by binary relations on the set $P = \{1, \ldots, p\}$ where p is the number of subsystems. Structured linear systems are linear systems that have an underlying binary relation that determines the block zero pattern of the system matrices. Specifically, poset-causal systems, introduced by Shah and Parrilo [42], have an underlying partial order that determines their structure.

Centralized LTI systems with proper real rational transfer functions can be represented via state space realizations $\widehat{G}(\lambda) = C(\lambda I - A)^{-1}B + D$. A state space realization (A, B, C, D) is minimal if its state space dimension is as small as possible. The dual concepts of controllability and observability play a fundamental role in systems theory. A realization is minimal if and only if it is controllable and observable. Given a non-minimal realization, it can be reduced to a minimal realization using its Kalman decomposition.

For a real rational transfer function whose block-zero structure is determined by a pre-order, we employ a so-called block canonical shuffle to obtain a structured realization of the transfer function. We investigate several notions of controllability and observability for poset-causal systems, and their relation under duality. These new notions extend concepts of controllability and observability that were introduced in [23] for a special class of poset-causal systems known as coordinated linear systems. We show that the class of poset-causal systems is closed under duality, which is not the case for coordinated linear systems, and that duality relations between the various notions of observability and controllability exist. With these notions, we define a Kalman type reduction which compresses the poset-causal system to a poset-causal system with a much smaller state space dimension.

The optimal control of structured linear systems proves to be challenging due to the additional structural constraint. In the classical centralized setting for the $\mathcal{H}_2$-control problem, all stabilizing controllers of a plant are given by the Youla parametrization. A reparameterized problem may then be solved in the Youla parameter from which an optimal controller can be recovered. We consider the $\mathcal{H}_2$-control problem for poset-causal systems. In the state feedback case, a related reparameterized problem reduces to several local classical problems which can readily be solved via a spectral factorization approach as is done in [39]. The output feedback case is considerably more difficult. We consider various solution strategies to the output feedback $\mathcal{H}_2$-control problem for poset-causal systems based on a detailed analysis of various approaches to the unstructured output feedback control problem. These approaches aim to reduce the structured control problem to multiple unstructured local problems. Although an optimal structured solution is difficult to obtain, various optimality considerations are taken into account in the construction of feasible controllers for the output feedback case.

Keywords: Decentralized systems, Controllability and observability, Minimality, Kalman decomposition, Optimal control

Contents

Chapter 1

Introduction

In the wake of the fourth industrial revolution, control systems have become increasingly complex. In many practical applications, the systems under consideration are large scale and consist of spatially distributed, but interconnected subsystems. Information flow in such systems occur in a distributed manner. Such systems lend themselves well to decentralized control strategies rather then classical centralized control strategies. Examples of such practical applications include: large scale irrigation systems [6, 26, 32, 36, 51], hydroelectricity plants [10, 9] inland navigation networks [33, 37], civil building systems [14, 28, 31] and electric power systems [29].

Interconnected systems are often approached using graph theoretic techniques where nodes represent subsystems and directed edges represent communication flow (see for example [43, 46]). An equivalent approach was introduced by Shah and Parrilo in [42] using partial orders, leading to the notion of poset-causal systems, which were further studied in the papers [41, 40, 39] and the PhD thesis [38]. Following the lead from Shah and Parrilo, we model communication structures in interconnected systems with underlying binary relations, which are not necessarily posets. Suppose an interconnected system consists of p subsystems labeled $1, \ldots, p$ and let $P = \{1, \ldots, p\}$. A binary relation is a pair $\mathcal{T} = (P, T)$, where $T \subseteq P \times P$. Communication is modelled by the binary relation in the following way. If a subsystem $i \in P$ cannot communicate with another subsystem $j \in P$, then $(i, j) \notin T$.

Partitioning the system matrices of a decentralized system into $p \times p$ blocks, the binary relation also imposes a block zero pattern in the resulting block matrix. Given a system matrix $M = [M_{ij}]$ where M_{ij} is the block entry in the i^{th} row and j^{th} column, $M_{ij} = 0$ if $(j, i) \notin T$. One can then define the set $\mathcal{I}_{\mathcal{T}}$ of all appropriately sized block matrices that have a block zero pattern determined by a binary relation $\mathcal{T}$. Such a set is clearly closed under addition and scalar multiplication, but not necessarily under matrix multiplication. We then consider various properties of the binary relation and investigate the effect it has on the communication structure of the interconnected system. A binary relation $\mathcal{T} = (P, T)$ is

(i) reflexive if $(i, i) \in T$ for all $i \in P$;

(ii) transitive if $(i, j) \in T$ and $(j, k) \in T$, then (i, k) for all $i, j, k \in P$;

(iii) anti-symmetric if (i, j) and (j, i), then $i = j$ for all $i, j \in P$.

It turns out that $\mathcal{I}_{\mathcal{T}}$ is closed under matrix multiplication, provided the sizes correspond appropriately, if and only if $\mathcal{T}$ is transitive. A binary relation satisfying all three conditions above is called a poset. In this case we write $i \succeq j$ if and only if $(i, j) \in T$. Poset-causal systems are interconnected linear systems whose communication structure is determined by an underlying poset. Equivalently, it is a linear system whose system matrices are all in $\mathcal{I}_{\mathcal{P}}$ for some poset $\mathcal{P} = (P, \succeq)$. An important sub-class of poset-causal systems are so-called coordinated linear systems, which were defined in [34]. The underlying poset $\mathcal{P} = (P, \succeq)$ of a coordinated linear system satisfies an additional property, namely that of in-ultra transitivity. The corresponding underlying graphs of coordinated linear systems are out-trees, that is, rooted trees with all directed edges pointing away from a single node called the root.

Poset-causal systems will be the main structured linear system that we will study. As such we now give a more in depth description of poset-causal systems. Consider a decentralized system with p interconnected subsystems labeled $1, 2, \ldots, p$. Let $\succeq$ be a partial order on $P = \{1, \ldots, p\}$. In this setting, subsystem j can 'influence' subsystem i in case $j \succeq i$. We assume that each subsystem is locally given by an input-state-output model with an input space $\mathcal{U}_i = \mathbb{R}^{m_i}$, a state space $\mathcal{X}_i = \mathbb{R}^{n_i}$ and an output space $\mathcal{Y}_i = \mathbb{R}^{r_i}$, with $m_i, n_i, r_i \in \mathbb{Z}_+$. The local outputs $y_i(t) \in \mathcal{Y}_i$ and local states $x_i(t) \in \mathcal{X}_i$ are determined by the states $x_j(t) \in \mathcal{X}_j$ and inputs $u_j(t) \in \mathcal{U}_j$ of subsystems j that can 'influence' subsystem i via interconnected state space system equations

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j \in \uparrow i} A_{ij} x_j(t) + \sum_{j \in \uparrow i} B_{ij} u_j(t), \quad x_i(0) = x_{i,0}, \\ y_i(t) &= \sum_{j \in \uparrow i} C_{ij} x_j(t) + \sum_{j \in \uparrow i} D_{ij} u_j(t), \quad t \geq 0, \end{aligned} \tag{1.1}$$

where $\uparrow i = \{j \in P : j \succeq i\}$ is the set of subsystems j that are upstream of subsystem i in the communication network. That is, if $j \in \uparrow i$, then subsystem j can 'influence' subsystem i. Here $x_{i,0} \in \mathbb{R}^{n_i}$ is the initial state of subsystem i and $A_{ij} \in \mathbb{R}^{n_i \times n_j}$, $B_{ij} \in \mathbb{R}^{n_i \times m_j}$, $C_{ij} \in \mathbb{R}^{r_i \times n_j}$ and $D_{ij} \in \mathbb{R}^{r_i \times m_j}$ are given matrices whenever $j \succeq i$. In case $j \not\succeq i$, set A_{ij}, B_{ij}, C_{ij} and D_{ij} equal to zero matrices of appropriate sizes and define

$$A = [A_{ij}]_{i,j=1}^p, \quad B = [B_{ij}]_{i,j=1}^p, \quad C = [C_{ij}]_{i,j=1}^p, \quad D = [D_{ij}]_{i,j=1}^p. \tag{1.2}$$

Then the combined input, state and output signals

$$\begin{aligned} u(t) &= (u_1(t), \ldots, u_p(t))^\intercal \in \mathcal{U} := \bigoplus_{i=1}^p \mathcal{U}_i, \\ x(t) &= (x_1(t), \ldots, x_p(t))^\intercal \in \mathcal{X} := \bigoplus_{i=1}^p \mathcal{X}_i, \\ y(t) &= (y_1(t), \ldots, y_p(t))^\intercal \in \mathcal{Y} := \bigoplus_{i=1}^p \mathcal{Y}_i \end{aligned}$$

satisfy

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0 \in \mathcal{X}, \\ y(t) &= Cx(t) + Du(t), \quad t \geq 0, \end{aligned} \tag{1.3}$$

where $x_0 := (x_{1,0}, \ldots, x_{p,0})^\intercal$. Hence the decentralized system (1.1) can be written as a classical state space system (1.3) with the communication structure embedded in a prescribed block zero pattern of the system matrices determined by the underlying partial order. In particular, the state and output signals can be represented in terms of the input and initial state by the classical integral formulas

$$\begin{aligned} x(t) &= x(x_0, u, t) = e^{At} x_0 + \int_0^t e^{A(t-\tau)} Bu(\tau) \, \mathrm{d}\tau, \\ y(t) &= y(x_0, u, t) = Ce^{At} x_0 + \int_0^t Ce^{A(t-\tau)} Bu(t) \, \mathrm{d}\tau + Du(t). \end{aligned}$$

There are two main parts to this thesis. We develop a systems theory for poset-causal systems and secondly, we consider an $\mathcal{H}_2$-optimization problem for poset-causal system. In order to give a sketch of the theory in the structured setting, we first consider the classical unstructured setting. In light of the classical results, we then discuss some of the results that are obtained in the structured setting.

Consider a continuous time causal linear time invariant (LTI) input-output system Σ with input u and output y. Such a system Σ has an input-output map $G_\Sigma : u \mapsto y$. The system has a state space realization if there exists a state $x : [0, \infty) \to \mathcal{X}$ and matrices A, B, C and D such that the solution of the system of equations

$$\Sigma \begin{cases} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{cases} \qquad x(0) = x_0 \qquad t \geq 0,$$

is given by $y(t) = (G_\Sigma(u))(t)$ for $t \geq 0$. In that case, we write $\Sigma \sim (A, B, C, D)$. Taking Laplace transforms, the transfer function of the system is $\widehat{G} : \widehat{u} \mapsto \widehat{y}$. If $\Sigma \sim (A, B, C, D)$, then

$$\widehat{G}(\lambda) = C(\lambda I - A)^{-1}B + D =: \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

provided $\lambda \in \mathbb{C}$ is not an eigenvalue of A. A causal LTI system with a proper real rational transfer function always has a state space realization. Such realizations are not unique. Given that there are many possible state space realizations of a given system, we are interested in finding the realizations which are optimal in some sense. In this regard the concepts of controllability and observability play a crucial role. These concepts were first defined by R.E. Kalman in the 1960 paper [18]. Controllability is related to the state equation

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0, \qquad t \geq 0$$

of the system $\Sigma \sim (A, B, C, D)$. Let $\mathcal{X} = \mathbb{R}^n$ be the state space. For any given initial state x_0, the state x at a final time $t > 0$ and any valid input $u : [0, t] \to \mathbb{R}^m$ can be solved by the classical integral formula

$$x(x_0, u, t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\, \mathrm{d}\tau.$$

The system is *controllable* if for any initial state x_0 and any final time $t > 0$ and any final state $\xi \in \mathbb{R}^n$, there exists a valid input $u : [0, t] \mapsto \mathbb{R}^m$ such that $x(x_0, u, t) = \xi$. The pair of matrices (A, B) is said to be controllable if the corresponding linear system is controllable. Related to the pair (A, B) is the controllability matrix

$$\mathcal{C}(A, B) = \begin{bmatrix} B & AB & A^2B & \ldots & A^{n-1}B \end{bmatrix}$$

and the reachable subspace

$$\mathcal{R}(A, B) = \operatorname{Im} \begin{bmatrix} B & AB & A^2B & \ldots A^{n-1}B \end{bmatrix},$$

the image of the controllability matrix. It can be shown that (A, B) is controllable if and only if the entire state space is reachable, that is, $\mathcal{R}(A, B) = \mathcal{X}$.

On the other hand, observability is related to the equations

$$\begin{aligned} \dot{x}(t) &= Ax(t), \qquad x(0) = x_0 \\ y(t) &= Cx(t) \end{aligned}$$

where the input is taken to be zero. The pair (C, A) is said to be *observable* if at any given time $T > 0$, it is possible to determine the initial state x_0 from the outputs $y(t)$ over the time interval $[0, T]$. Related to the pair (C, A), is the observability matrix

$$\mathcal{O}(C, A) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

and the unobservable space

$$\mathcal{N}(C, A) = \ker \mathcal{O}(C, A).$$

It can be shown that (C, A) is observable if and only if $\ker \mathcal{O}(C, A) = \{0\}$. It is well known that controllability and observability are dual notions in the following sense.

Theorem 1.0.1.
The pair (C, A) is observable if and only if $(A^\intercal, C^\intercal)$ is controllable.

Two realizations (A, B, C, D) and $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ are equivalent realizations if they give the same input-output map, that is,

$$\int_0^t Ce^{A(t-\tau)}Bu(\tau)\ \mathrm{d}\tau + Du(t) = \int_0^t \widetilde{C}e^{\widetilde{A}(t-\tau)}\widetilde{B}u(\tau)\ \mathrm{d}\tau + \widetilde{D}u(t)$$

for all inputs u and all times $t > 0$. The *dynamic order* of a realization (A, B, C, D) is defined to be the size of the square matrix A, that is if $A \in \mathbb{R}^{n\times n}$, then the dynamic order of (A, B, C, D) is n. A realization $\Sigma \sim (A, B, C, D)$ with dynamic order n is said to be *minimal* if $n \leq \tilde{n}$ for all other equivalent realizations $\Sigma \sim (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$. The following well-known theorem relates minimality to controllability and observability.

Theorem 1.0.2.
A realization (A, B, C, D) is minimal if and only if (A, B) is controllable and (C, A) is observable.

If a realization (A, B, C, D) with state space $\mathcal{X}$ is not minimal, one way of obtaining a minimal system that has the same input-output map goes through the famous *Kalman decomposition*, first developed in [19]. Since controllability and observability play critical roles in minimality, the state space $\mathcal{X}$ is partitioned in terms of the subspaces $\mathcal{R} := \mathcal{R}(A, B)$ and $\mathcal{N} := \mathcal{N}(C, A)$. Define the subspaces

$$\mathcal{X}_{co} := \mathcal{R} \ominus (\mathcal{R} \cap \mathcal{N}), \quad \mathcal{X}_{c\overline{o}} := \mathcal{R} \cap \mathcal{N}, \quad \mathcal{X}_{\overline{c}o} := (\mathcal{R} + \mathcal{N})^\perp \quad \text{and} \quad \mathcal{X}_{\overline{co}} := \mathcal{N} \ominus (\mathcal{R} \cap \mathcal{N}).$$

Then

$$\begin{aligned} \mathcal{X}_{co} \oplus \mathcal{X}_{c\overline{o}} &= \mathcal{R}, \\ \mathcal{X}_{c\overline{o}} \oplus \mathcal{X}_{\overline{co}} &= \mathcal{N}, \\ \mathcal{X}_{co} \oplus \mathcal{X}_{c\overline{o}} \oplus \mathcal{X}_{\overline{co}} &= \mathcal{R} + \mathcal{N} \\ \mathcal{X}_{co} \oplus \mathcal{X}_{c\overline{o}} \oplus \mathcal{X}_{\overline{c}o} \oplus \mathcal{X}_{\overline{co}} &= \mathcal{X}. \end{aligned}$$

It can then be shown that there exists a similarity transform T such that A, B and C partition into the *Kalman decomposition of* (A, B, C):

$$TAT^{-1} = \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix}, \quad TB = \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix}, \quad CT^{-1} = \begin{bmatrix} C_1 & 0 & C_3 & 0 \end{bmatrix}.$$

With the above decomposition, it holds that (A_{11}, B_1) is controllable, (C_1, A_{11}) is observable and that (A, B, C, D) and (A_{11}, B_1, C_1, D) are equivalent realizations. Thus (A_{11}, B_1, C_1, D) is minimal realization of the system. The realization (A_{11}, B_1, C_1, D) is called the *Kalman reduction* of (A, B, C, D).

In the structured setting, the realization problem is the following: given a proper real rational transfer function $\widehat{G}$ of some linear system where the structure of the transfer function is determined by a underlying binary relation $\mathcal{T}$, find a realization (A, B, C, D) such that each system matrix has the structure determined by $\mathcal{T}$. This problem is considered in the paper [24], for so-called systems over graphs. It can be shown that systems over graphs are essentially the same as poset-causal systems. In the paper [24], the triple (A, B, C) is also required to be stabilizable and detectable. In the absence of the stabilizability and detectability requirements, we consider the structured realization problem for a system whose transfer function has a structured determined by a pre-order, that is a binary relation $\mathcal{T}$ that is reflexive and transitive. Using a so-called block canonical shuffle, we show that given a proper rational transfer function $\widehat{G}$ of a causal LTI system such that $\widehat{G}$ has a structure determined by $\mathcal{T}$, a structured realization (A, B, C, D) may always be obtained. Since pre-orders are more general than partial orders, this result holds for poset-causal systems as well.

Concerning the systems theory of structured linear systems, much of the research on the controllability of decentralized systems take a graph-theoretic approach where leaders are to be chosen from multiple agents in a communication topology in such a way that renders the system controllable. In [47], the controllability of an interconnected system by choosing a single leader among first-order subsystems, is considered. This was extended to multiple leaders in [16]. A graphical characterization of controllability in this context was developed in [15]. In the paper [5], the concepts of formation controllability and complete controllability are studied for multi-agent swarm systems where all the agents are LTI systems and have the same order. The controllability of networked systems is studied in [50].

Several concepts of controllability and observability of coordinated linear systems with one leader and two followers were studied in [23]. The aim in [23] was to develop concepts of controllability and observability that respect the communication structure of the system. Our aim is to study controllability and observability for the larger class of poset-causal systems. Due to the more intricate structures of such systems, our notions are not based in leader-follower concepts, but rather use upstream and downstream subsystems from the perspective of the given subsystems. These lead to notions of controllability that respect the poset-causal structure, in the sense that the space of controllable state vectors decompose in direct sums of subspaces of the local state spaces corresponding to the subsystems of the poset-causal system.

Since the poset-causal system (1.1) can be represented as a classical state space system (1.3) with structured system matrices, all the notions, results and constructions from classical state space theory apply. However, most of these do not preserve the block zero-pattern. For example, the reachable subspace $\mathcal{R}(A, B)$ and unobservable subspace $\mathcal{N}(C, A)$ of the state space $\mathcal{X}$ cannot, in general, be written as direct sums of subspaces of the local state spaces $\mathcal{X}_j$ for $j \in P$, so that the compression to a minimal system obtained from the Kalman decomposition will in general not have the appropriate zero-pattern, destroying the poset-causal structure. As a result, it is not clear how to determine if a poset-causal system is minimal, that is, whether there does not exist a poset-causal system which generates the same input-output map but has smaller state space dimensions. Neither is it clear whether there exists a Kalman-type decomposition if it is also required that the block zero-pattern be preserved. New notions of controllability and observability are required that do preserve the communication structure and still preserve some of the features of the classical notions.

For the subclass of coordinated linear systems defined in [34] various new concepts of controllability and observability that preserve the underlying communication structure were studied in [23], see also [21]. Most of the controllability and observability notions in [23] are defined for a coordinated linear systems that consists of a single coordinator and two followers. Here, the coordinator can communicate with the followers, but the followers cannot communicate with each other or with the coordinator.

We introduce various notions of controllability and observability for poset-causal systems related to upstream and downstream systems associated with the subsystems. When specialized to the setting of coordinated linear systems, these reduce to notions of controllability and observability studied in [23], however, the notions introduced here do not rely on the roles of "coordinator" or "follower" a subsystem may have. While it is possible to study controllability and observability for systems defined on graphs from the perspective of assigning controllers, see e.g. [30], we do not try to assign specific roles to the subsystems. The class of coordinated linear systems is not closed under duality. However, it turns out that this is the case for the class of poset-causal systems, and we prove duality relations between the controllability and observability notions defined in this thesis, as it occurs in the classical case. One notion, independently controllable, is stronger than classical controllability, in the sense that it implies classical controllability, and one notion, weak local controllability, is weaker, in the sense that it is implied by classical controllability. In fact, in a way (see Theorem 6.2.3) weak local controllability is the strongest notion of controllability implied by controllability that preserves the poset-causal structure. Similar notions exist for observability for poset-causal systems, and it turns out that they are related through duality, as in the classical setting.

Finally, we investigate how our notions of controllability and observability can be used to perform a Kalman type reduction of the system in a way that preserves the poset-causal structure. Most of the notions of controllability and observability introduced here are based on variations on the classical reachable subspace and unobservable subspace, with the difference that they can be written as (orthogonal) direct sums of subspaces of the local state spaces of the subsystems. As a result of this, we are able to present a variation on the Kalman reduction formula. Despite that some of the controllability and observability notions introduced here are optimal, this does not carry over to the Kalman-type reduction. Hence, although it may compress the poset-causal system to a poset-causal system with a much smaller state space dimension, it need not be the minimal poset-causal system with the same input-output map.

We now move on to $\mathcal{H}_2$-optimal control. Again, we give an overview of the classical setting before commenting on the structured case. An input-output system (a plant) may be controlled by connecting a second linear system, called the controller, to the plant. As illustrated in Figure 1.1, the key idea is that of feedback. The measured output y of the plant is fed to the controller as its control input. The controller then feeds back its output into the control input u of the plant. Connecting a controller to a plant has the effect of closing the feedback loop and results in a new linear system whose transfer function is the closed loop transfer function of the plant-controller interconnection. This closed loop transfer function maps the disturbance input ω to the external output z. The aim of $\mathcal{H}_2$-optimal control is to synthesize a controller that minimizes the $\mathcal{H}_2$-norm of the closed loop transfer function. The set-up is illustrated in the figure below

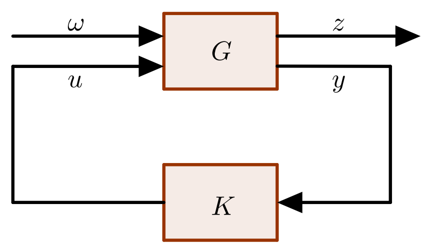


Figure 1.1: $\mathcal{H}_2$-control feedback loop

We assume the plant G is a linear system with state space realization

$$\widehat{G} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix} \tag{1.4}$$

and seek a controller in the form of a linear system with state space realization

$$\widehat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right].$$

It can be shown that the closed loop transfer function after connecting the controller to the plant, is given by the lower linear fractional transformation (LFT), $\underline{\mathcal{F}}(\widehat{G}, \widehat{K})$, of $\widehat{G}$ and $\widehat{K}$, given by

$$\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{K}(I - \widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21} =: \left[\begin{array}{c|c} \underline{A} & \underline{B} \\ \hline \underline{C} & \underline{D} \end{array}\right].$$

A controller K *stabilizes* the plant G if the resulting closed loop matrix $\underline{A}$ above is stable, that is, if $\mathrm{Re}(\lambda) < 0$ for all λ in the spectrum of $\underline{A}$. The matrix pair (A, B) is stabilizable if there exists a matrix E such that $A + BE$ is stable and (C, A) is detectable if there exists M such that $A + MC$ is stable. If (A, B_2, C_2) is stabilizable and detectable, the famous Youla parametrization characterizes all stabilizing controllers of the plant $\widehat{G}$ in terms of some canonical system $\widehat{J}$. Specifically, if

$$\widehat{J} = \left[\begin{array}{c|cc} A + B_2E + MC_2 & -M & B_2 \\ \hline E & 0 & I \\ -C_2 & I & 0 \end{array}\right],$$

then all stabilizing controllers of $\widehat{G}$ are given by

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}, \widehat{R}), \qquad \text{for some} \qquad \widehat{R} \in \mathcal{RH}_\infty$$

and the closed loop transfer function is given by

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}, \qquad \text{where} \qquad \widehat{H} = \widehat{G} \star \widehat{J}.$$

Here $\widehat{G}\star\widehat{J}$ is the so-called Redheffer star product of $\widehat{G}$ and $\widehat{J}$. The $\mathcal{H}_2$-optimal control problem is to construct a controller $\widehat{K}$ such that the closed loop transfer function $\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right)$ is stable, strictly proper and such that its $\mathcal{H}_2$-norm is minimized. This can be stated as follows

$$\begin{aligned} \text{minimize} \qquad & \|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2 \\ \text{subject to} \qquad & \widehat{K} \in \mathcal{RH}_\infty, \quad \underline{\mathcal{F}}(\widehat{G}, \widehat{K}) \in \mathcal{RH}_2. \end{aligned}$$

A discussion of the Hardy spaces $\mathcal{H}_2$, $\mathcal{RH}_2$ and $\mathcal{RH}_\infty$ can be found in Section 2.4. Equivalently, one may consider the reparameterized problem

$$\begin{aligned} \text{minimize} \qquad & \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2 \\ \text{subject to} \qquad & \widehat{R} \in \mathcal{RH}_2. \end{aligned}$$

The classical solution of the $\mathcal{H}_2$-control problem contains stabilizing solutions of Riccati equations related to the realization of $\widehat{G}$ (see Theorem 4.3.4). The conditions in the theorem are sufficient for such stabilizing solutions $X = \text{Ric}(A, B_2, C_1, D_{12})$ and $Y = \text{Ric}(A^\intercal, C_2^\intercal, B_1^\intercal, D_{21}^\intercal)$ to the Riccati equations

$$\begin{aligned} A^\intercal X + XA + C_1^\intercal C_1 - (XB_2 + C_1^\intercal D_{12})R_1^{-1}(B_2^\intercal X + D_{12}^\intercal C_1) = 0 \\ AY + YA^\intercal + B_1B_1^\intercal - (YC_2^\intercal + B_1D_{21}^\intercal)R_2^{-1}(C_2Y + D_{21}B_1^\intercal) = 0 \end{aligned}$$

to exist. For more on Riccati equations refer to Section 3.7. Defining

$$F = -(D_{12}^\intercal D_{12})^{-1}(B_2^\intercal X + D_{12}^\intercal C_1) \quad \text{and} \quad L = -(YC_2^\intercal + B_1D_{21}^\intercal)(D_{21}D_{21}^\intercal)^{-1},$$

the optimal solution is then given by

$$\widehat{K}_{opt} = \left[\begin{array}{c|c} A + B_2F + LC_2 & -L \\ \hline F & 0 \end{array}\right].$$

A special case of the $\mathcal{H}_2$-control problem is the state feedback case. In this case, $\widehat{G}$ has the simpler realization

$$\widehat{G} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C & 0 & D \\ I & 0 & 0 \end{array}\right] = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix}.$$

The state feedback $\mathcal{H}_2$-control problem is closely related to the LQ-control problem and the optimal solution is given by

$$\widehat{K} = F = -(D^\intercal D)^{-1}(B_2^\intercal X + D^\intercal C) \quad \text{where} \quad X = \text{Ric}(A, B, C, D).$$

Optimal control of structured linear systems proves challenging due to the additional requirement that the controller must respect the communication structure of the decentralized system. For example the matrix F appearing in the solution of both state and output feedback case in the classical setting, contains a solution X of a Riccati equation. Riccati equations do not in general preserve the structure imposed on the system matrices of a structured linear system. Hence alternative approaches which avoid solving Riccati equations which contain structured matrices, have to be considered.

The LQ-control problem was studied in [22] using a recursive bottom up optimization approach. More generally, the state-feedback $\mathcal{H}_2$-control problem for poset-causal systems was solved in [39] using a spectral factorization approach. The output feedback case was considered in [25], but only for the simple two player problem. We consider the output feedback $\mathcal{H}_2$-control problem for more general poset-causal systems. In this case, $\widehat{G}$ is a poset-causal plant, that is, the system matrices in (1.4) all have the structure determined by some poset $\mathcal{P}$. Since the set of structured matrices determined by a poset is closed under multiplication and inversion, it follows that if a controller has the poset-causal structure, it preserves the structure in the closed loop transfer function. Hence the aim is to construct a controller $\widehat{K} \in \mathcal{RH}_\infty$

$$\begin{aligned} &\text{minimize} && \|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\| \\ &\text{subject to} && \underline{\mathcal{F}}(\widehat{G}, \widehat{K}) \in \mathcal{RH}_2, \quad \widehat{K} \in \mathcal{I}_{\mathcal{P}}. \end{aligned}$$

Here, the requirement $\widehat{K} \in \mathcal{I}_{\mathcal{P}}$ enforces the poset-causal structure on the controller.

In the state feedback $\mathcal{H}_2$-control problem for poset-causal systems considered in [39], the approach considers the reparameterized problem

$$\begin{aligned} &\text{minimize} && \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\| \\ &\text{subject to} && \widehat{Q} \in \mathcal{RH}_2, \quad \widehat{Q} \in \mathcal{I}_{\mathcal{P}}, \end{aligned} \tag{1.5}$$

where $\widehat{Q} = \widehat{R}\widehat{H}_{21}$. Due to the special form of the realization (1.4) in the state feedback case, $\widehat{Q}$ and $\widehat{R}$ determine each other uniquely. In this form it can then be shown that the problem decomposes into p local unstructured $\mathcal{H}_2$-control problems. Each of these problems may be solved using a spectral factorization technique to obtain an optimal solution $\widehat{Q}_j$ for $j = 1, \ldots, p$. These local solutions can in turn be used to construct a controller $\widehat{Q}$ which satisfies the poset-causal structure and which is optimal for the reparameterized problem. Again due to the special form of the matrices in the state feedback case a stable controller $\widehat{K}$ can then be recovered. This controller $\widehat{K}$ is optimal for the original $\mathcal{H}_2$-control problem.

The output feedback case is considerably more difficult. As in the state feedback, an optimal structured solution to the reparameterized problem (1.5) may be obtained via decomposing to unstructured problems and using the spectral factorization technique, but unlike the state feedback case, it is not possible to factor $\widehat{Q}$ in terms of $\widehat{H}_{21}$. Hence it is not clear how to recover a feasible structured controller $\widehat{K}$ from $\widehat{Q}$. We consider various different approaches to the output feedback problem, in each case taking into account optimality considerations at certain stages of the construction of a feasible structured controller $\widehat{K}$. Though not optimal, this gives options for constructing feasible solutions and in some cases it is possible to show that these controllers are better performing than some of the alternatives.

We end this introduction with an overview of the structure of the thesis.

In Chapter 2, we review some mathematical preliminaries that are required in the rest of the thesis. In Section 2.1, we consider binary relations, their properties and their relation to directed graphs. In Section 2.2, we give a summary of some concepts and results from linear algebra that will be important in the rest of the thesis. We will have to work with block matrices frequently. In Section 2.3, we define some concepts and notation that is required to effectively convey results regarding block matrices and their block zero structure. Important in this section is the *block incidence set* and its closure under multiplication. The section also ends with a description of the block canonical shuffle which plays an important role in our realization result. Section 2.4 is a brief summary of Hardy spaces and its subspaces of real rational matrix functions.

Chapter 3 is a summary of realization and systems theory in the classical unstructured setting. Section 3.1 is a brief overview of rational matrix functions. Section 3.2 is about transfer functions and state space realizations. In Section 3.3, the classical concepts of controllability, observability and their relation under duality are summarized. We review equivalent realizations, minimal realizations and the Kalman reduction of a realization in Section 3.4. Section 3.5 contains some basic operations on realizations. The section also contains some more involved identities regarding realizations. The results in this section will be utilized frequently in the thesis. In Section 3.6, we give a review the relation of realizations to the Hardy spaces $\mathcal{RH}_2$ and $\mathcal{H}_\infty$. Section 3.7 is an important section on the algebraic Riccati equation. The section also reviews inner-outer factorizations. Both Riccati equations and inner-outer factorizations play a fundamental role in $\mathcal{H}_2$-optimization.

In Chapter 4, we review the $\mathcal{H}_2$-control problem in the classical unstructured setting. We also consider some alternative approaches to standard techniques. Section 4.1 reviews the interconnections of linear systems, when such interconnections are well posed and state space realizations of such interconnections. In Section 4.2, we describe the famous Youla parametrization which gives a characterization of all stabilizing controllers of a plant as well as a reparametrization of the closed loop transfer function. In Section 4.3 the classical $\mathcal{H}_2$-control problem is stated and solved via a standard technique involving Riccati equations and inner functions. In Section 4.4, the state feedback case and its solution is presented. An alternative approach to the state feedback case via spectral factorization is presented in Section 4.5. An alternative approach to the output feedback case avoiding the use of Riccati equations in the Youla parametrization is considered in Section 4.6 and finally a one-sided reparameterized problem for the output feedback case is described in Section 4.7.

In Chapter 5, we define and describe structured linear systems. In Section 5.1, we define general structured linear systems with an underlying binary relation and consider the effect of various properties of binary relations on the structure of such systems. A brief discussion regarding quadratic invariance is also included here. Section 5.2 develops the important subclass of poset-causal systems and also considers various derived systems. It is also shown that systems over graphs and poset-causal systems are equivalent. Section 5.3 is a brief section on coordinated linear systems and hierarchical systems.

In Chapter 6, we develop realization and systems theory for poset-causal systems. Section 6.1 describes the results on the structured realization problem for structured linear systems with an underlying pre-order. In Section 6.2 we develop various structured notions of controllability for poset-causal systems and show how these concepts relate to one another. In Section 6.3 we describe similar results for observability. In Section 6.4 it is shown that these concepts are related under duality via so-called dual poset-causal systems. In Section 6.5, we use the various subspaces defined in the previous sections to develop a structured Kalman-type decomposition.

Chapter 7 considers the $\mathcal{H}_2$-control problem for poset-causal systems. In Section 7.1 we describe the spectral factorization approach to the state feedback case and illustrate it via examples. In Section 7.2, we consider the difficult problem of output feedback control for poset-causal systems. We describe some of the difficulties that arise and consider various approaches to constructing feasible solutions.

Chapter 2

Mathematical Preliminaries

In this chapter we review some mathematical concepts that will be used frequently throughout this thesis. We also establish notation that will be employed subsequently in later chapters.

2.1 Order Structures

A *partially ordered set*, or *poset*, is a pair $\mathcal{P} = (P, \succeq)$ with P a set and $\succeq$ a partial order on P. That is, $\succeq$ is a binary relation on P which is

(i) *reflexive*: $i \succeq i$ for all $i \in P$;

(ii) *transitive*: if $i \succeq j$ and $j \succeq k$, then $i \succeq k$ for all $i, j, k \in P$;

(iii) *anti-symmetric*: if $i \succeq j$ and $j \succeq i$, then $i = j$ for all $i, j \in P$.

If a binary relation only satisfies (i) and (ii), then it is a *pre-order*. For $i, j \in P$ we write $i \succ j$ if $i \succeq j$ and $i \neq j$. Also, $i \preceq j$ and $i \prec j$ is equivalent to $j \succeq i$ and $j \succ i$, respectively. In the sequel we will only consider finite posets, usually with $P = \{1, 2, \dots, p\}$ for some positive integer p. Given a subset $R \subseteq P$ of a poset $\mathcal{P} = (P, \succeq)$ we define its *downstream set* $\downarrow R$ and its *upstream set* $\uparrow R$ as

$$\downarrow R := \{i \in P \colon j \succeq i \text{ for some } j \in R\} \qquad \text{and} \qquad \uparrow R := \{i \in P \colon i \succeq j \text{ for some } j \in R\}$$

respectively. In the case that R is a singleton, say $R = \{i\}$, we simply write $\downarrow i$ and $\uparrow i$. By reflexivity it follows that $R \subseteq \uparrow R$ and $R \subseteq \downarrow R$. Transitivity implies that $\uparrow(\uparrow R) \subseteq \uparrow R$ and $\downarrow(\downarrow R) \subseteq \downarrow R$. Taking the above two facts together gives $\uparrow(\uparrow R) = \uparrow R$ and $\downarrow(\downarrow R) = \downarrow R$. In addition, we define the *strict downstream set* $\text{⤈}R$ and *strict upstream set* $\text{⤉}R$ of R as

$$\text{⤈}R := \{i \in \downarrow R \colon i \notin R\} \qquad \text{and} \qquad \text{⤉}R := \{i \in \uparrow R \colon i \notin R\}.$$

Again, these are abbreviated to $\text{⤈}i$ and $\text{⤉}i$, respectively, when $R = \{i\}$.

Any poset $\mathcal{P} = (P, \succeq)$ can be represented by a digraph $\mathcal{G}_{\mathcal{P}} = (P, E_{\succeq})$ where the nodes in $\mathcal{G}_{\mathcal{P}}$ are the elements of P and $E_{\succeq} = \{(i, j) \in P \times P : i \succeq j\}$ is the set of directed edges in $\mathcal{G}_{\mathcal{P}}$. The *Hasse diagram* of a poset $\mathcal{P} = (P, \succeq)$ can be identified with the digraph $\mathcal{G}^{\downarrow}_{\mathcal{P}} = (P, E^{\downarrow}_{\succeq})$, where

$$E^{\downarrow}_{\succeq} := \{(i, j) \in P \times P : i \succ j \text{ and there is no } k \in P \text{ such that } i \succ k \succ j\}.$$

The digraph $\mathcal{G}^{\downarrow}_{\mathcal{P}}$ omits all the directed edges that correspond to reflexivity and transitivity. This is illustrated in the following example.

Example 2.1.1 (Posets and Hasse diagrams)**.**
Consider the poset $\mathcal{P} = (P, \succeq)$ with $P = \{1, 2, 3, 4\}$ determined by the relations $1 \succeq 2$, $3 \succeq 2$ and $2 \succeq 4$ (along with the loops $1 \succeq 1$, $2 \succeq 2$, $3 \succeq 3$ and $4 \succeq 4$, induced by reflexivity, and the edges $1 \succeq 4$ and $3 \succeq 4$, induced by transitivity). The digraph $\mathcal{G}_{\mathcal{P}} = (P, E_{\succeq})$ and Hasse diagram $\mathcal{G}^{\downarrow}_{\mathcal{P}} = (P, E^{\downarrow}_{\succeq})$ of $\mathcal{P}$ are shown below:

We also illustrate some upstream and downstream sets:

$$\downarrow 1 = \{1, 2, 4\}, \quad \mathrel{\rlap{\,-}{\downarrow}} 1 = \{2, 4\}, \quad \uparrow 4 = \{1, 2, 3, 4\} = P \qquad \text{and} \qquad \mathrel{\rlap{\,-}{\uparrow}} 2 = \{1, 3\}.$$

In [4, 3, 2] stronger notions of transitivity are studied, which are defined as follows.

Definition 2.1.2 (Ultra transitive relations)**.**
A binary relation $\mathcal{T} = (P, T)$ is said to be

(i) *in-ultra transitive* if for all $i, j, k \in P$, $(i, j) \in T$ and $(k, j) \in T$ implies that $(i, k) \in T$ or $(k, i) \in T$, and

(ii) *out-ultra transitive* if for all $i, j, k \in P$, $(i, j) \in T$ and $(i, k) \in T$ implies that $(j, k) \in T$ or $(k, j) \in T$.

Recall that a directed tree is a connected acyclic digraph. An *out-tree* with a *root* r (see for example p.999 of [3]) is a directed tree with all edges directed away from the single node r. An *in-tree* is obtained by reversing the direction of the edges. An *out-forest* is the disjoint union of out-trees and an *in-forest* is the disjoint union of in-trees. By Theorem 4.1 in [4] out-tree forests correspond to posets which are in-ultra transitive.

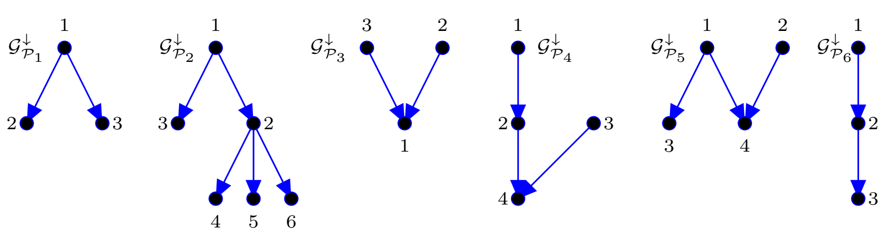
Definition 2.1.3 (Dual poset)**.**
For a poset $\mathcal{P} = (P, \succeq)$, we define its *dual poset* as $\mathcal{P}_d = (P, \succeq_d)$ where

$$j \succeq_d i \iff i \succeq j, \qquad \text{for each } j, i \in P.$$

In the sequel we will use the symbols $\downarrow_d$ and $\uparrow_d$ to indicate downstream and upstream sets, respectively, associated with the dual poset $\mathcal{P}_d$. Clearly we have $\uparrow_d R = \downarrow R$ and $\downarrow_d R = \uparrow R$, for any $R \subseteq P$. The following lemma follows directly from the definitions.

Lemma 2.1.4.
The dual of an in-ultra transitive relation is an out-ultra transitive relation.

Example 2.1.5 (Examples of posets)**.**
Consider posets $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_3$, $\mathcal{P}_4$, $\mathcal{P}_5$ and $\mathcal{P}_6$ with Hasse diagrams given by:



Then $\mathcal{G}^{\downarrow}_{\mathcal{P}_1}$ is the Hasse diagram of poset that is in-ultra transitive. $\mathcal{G}^{\downarrow}_{\mathcal{P}_2}$ is an out-tree corresponding to a more intricate poset that is in-ultra transitive. $\mathcal{G}^{\downarrow}_{\mathcal{P}_3}$ is the dual of $\mathcal{G}^{\downarrow}_{\mathcal{P}_1}$ and hence corresponds to an out-ultra transitive ordering (in-tree). $\mathcal{G}^{\downarrow}_{\mathcal{P}_4}$ is also an in-tree. $\mathcal{G}^{\downarrow}_{\mathcal{P}_5}$ is the Hasse diagram of a poset which is neither in-ultra transitive nor out-ultra transitive. Lastly, $\mathcal{G}_{\mathcal{P}_6}$ corresponds to a complete order.

2.2 Linear Algebra

In this section we summarize some concepts and results from linear algebra that are required in the rest of the thesis. We also establish notation that will be employed in the remainder of the thesis.

2.2.1 Eigenvalues and the Companion Matrix

We briefly review concepts related to eigenvalues for the sake of convenience. The companion matrix plays a prominent role in section 6.1 and hence we briefly summarize it here.

The *characteristic polynomial* of a square matrix $G \in \mathbb{R}^{n\times n}$ is defined to be

$$p_G(\lambda) = \det(\lambda I_n - G), \qquad \lambda \in \mathbb{C},$$

where I_n is the $n \times n$ identity matrix. The *eigenvalues* of G are the n (not necessarily distinct) roots of the characteristic polynomial p_G. The *spectrum* of G is the (multi)set of eigenvalues of G and is denoted by $\sigma(G)$. The *resolvent set* $\rho(G)$ is the complement of $\sigma(G)$ in $\mathbb{C}$. A matrix G is said to be *Hurwitz* or *stable* if each eigenvalue of G has a negative real part, that is, $\mathrm{Re}(\lambda) < 0$ for each $\lambda \in \sigma(G)$. The matrix G is symmetric if $G^\intercal = G$. A real symmetric matrix has real eigenvalues and is said to be *positive definite*, written $G > 0$, if $\lambda > 0$ for each $\lambda \in \sigma(G)$ and G is *positive semi-definite*, written $G \geq 0$, if $\lambda \geq 0$ for each $\lambda \in \sigma(G)$.

Definition 2.2.1 (Companion matrix)**.**
Given a real monic polynomial p of degree n

$$p(\lambda) = \lambda^n + g_{n-1}\lambda^{n-1} + \ldots + g_1\lambda + g_0,$$

with $\lambda \in \mathbb{C}$ and $g_0, \ldots, g_{n-1} \in \mathbb{R}$, the *companion matrix* of p is defined to be the following square matrix $G \in \mathbb{R}^{n\times n}$:

$$G = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -g_0 & -g_1 & -g_2 & \ldots & -g_{n-1} \end{bmatrix}.$$

The key feature of the companion matrix of p is that its characteristic polynomial is equal to p, that is, $p_G(\lambda) = p(\lambda)$ for all $\lambda \in \mathbb{C}$.

2.2.2 Schur Complement and Block Matrix Inversion

We will frequently work with block matrices in this thesis. In this subsection, we consider the inverses of block matrices. For this purpose, the Schur complement of a 2×2 block matrix will be a useful tool.

Definition 2.2.2 (Schur complement)**.**
Suppose $n = \ell + m$ and consider a square block matrix $A \in \mathbb{R}^{n\times n}$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{2.1}$$

where $A_{11} \in \mathbb{R}^{\ell\times\ell}$ and $A_{22} \in \mathbb{R}^{m\times m}$ are also square matrices. If A_{11} is non-singular, then

$$\Delta_1 := A_{22} - A_{21}A_{11}^{-1}A_{12} \tag{2.2}$$

is called the *Schur complement of* A_{11} in A. Similarly, if A_{22} is non-singular, then

$$\Delta_2 := A_{11} - A_{12}A_{22}^{-1}A_{21}$$

is called the *Schur complement of* A_{22} in A.

The following lemma shows how a 2×2 block matrix as in (2.1) can be factored using the Schur complement of A_{11} or A_{22} (if they exist). Furthermore, it shows that the invertibility of the block matrix A is intimately related to the invertibility of its diagonal entries and Schur-complements and gives inversion formulas for A in cases where the complements are invertible.

Lemma 2.2.3 (cf. page 13-14 of [56]).
Let A be a square 2×2 block matrix as in (2.1). If A_{11} is non-singular, then

$$A = \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & \Delta_1 \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix},$$

where Δ_1 is the Schur complement of A_{11} in A. In this case A is non-singular if and only if Δ_1 is non-singular and in that case

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\Delta_1^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\Delta_1^{-1} \\ -\Delta_1^{-1}A_{21}A_{11}^{-1} & \Delta_1^{-1} \end{bmatrix}.$$

Similarly, if A_{22} is non-singular, then

$$A = \begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Delta_2 & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{bmatrix},$$

where Δ_2 is the Schur complement of A_{22} in A. In this case A is non-singular if and only if Δ_2 is non-singular and in that case

$$A^{-1} = \begin{bmatrix} \Delta_2^{-1} & -\Delta_2^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}\Delta_2^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}\Delta_2^{-1}A_{12}A_{22}^{-1} \end{bmatrix}.$$

As special cases, one can consider upper block triangular matrices (where $A_{21} = 0$ in (2.1)) and lower block triangular matrices (where $A_{12} = 0$ in (2.1)). In those cases, the Schur complements reduce to $\Delta_1 = A_{22}$ and $\Delta_2 = A_{11}$ respectively. One then obtains the following inversion formula for block upper and block lower triangular 2×2 matrices.

Corollary 2.2.4.
If A_{11} and A_{22} are invertible, then

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix} \quad \text{and} \tag{2.3}$$

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{bmatrix}. \tag{2.4}$$

Inductively applying (2.3) and (2.4) gives inversion formulas for non-singular triangular $n \times n$ block matrices, in particular the inverse of non-singular 3×3 block triangular matrices is given in the following Lemma.

Corollary 2.2.5.
If A_{11}, A_{22} and A_{33} are invertible, then

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} & A_{11}^{-1}(A_{12}A_{22}^{-1}A_{23} - A_{13})A_{33}^{-1} \\ 0 & A_{22}^{-1} & -A_{22}^{-1}A_{23}A_{33}^{-1} \\ 0 & 0 & A_{33}^{-1} \end{bmatrix} \quad \text{and} \tag{2.5}$$

$$\begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & A_{32} & A_{33} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} & 0 \\ A_{33}^{-1}(A_{32}A_{22}^{-1}A_{21} - A_{31})A_{11}^{-1} & -A_{33}^{-1}A_{32}A_{22}^{-1} & A_{33}^{-1} \end{bmatrix}. \tag{2.6}$$

Similarly, the inverse of non-singular 4×4 block triangular matrices is given in the following Corollary. This result can be obtained by combining Corollaries 2.2.4 and 2.2.5

Corollary 2.2.6.
If A_{11}, A_{22}, A_{33} and A_{44} are invertible, then

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ 0 & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & A_{34} \\ 0 & 0 & 0 & A_{44} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} & A_{11}^{-1}B_{13}A_{33}^{-1} & A_{11}^{-1}B_{14}A_{44}^{-1} \\ 0 & A_{22}^{-1} & -A_{22}^{-1}A_{23}A_{33}^{-1} & A_{22}^{-1}B_{24}A_{44}^{-1} \\ 0 & 0 & A_{33}^{-1} & -A_{33}^{-1}A_{34}A_{44}^{-1} \\ 0 & 0 & 0 & A_{44}^{-1} \end{bmatrix}. \tag{2.7}$$

where

$$B_{13} = A_{12}A_{22}^{-1}A_{23} - A_{13}, \quad B_{14} = A_{12}A_{22}^{-1}A_{24} - A_{12}A_{22}^{-1}A_{23}A_{33}^{-1}A_{34} + A_{13}A_{33}^{-1}A_{34} - A_{14} \quad \text{and}$$
$$B_{24} = A_{23}A_{33}^{-1}A_{34} - A_{24}.$$

Similarly,

$$\begin{bmatrix} A_{11} & 0 & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 \\ A_{31} & A_{32} & A_{33} & 0 \\ A_{41} & A_{42} & A_{43} & A_{44} \end{bmatrix} = \begin{bmatrix} A_{11}^{-1} & 0 & 0 & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} & 0 & 0 \\ A_{33}^{-1}B_{31}A_{11}^{-1} & -A_{33}^{-1}A_{32}A_{22}^{-1} & A_{33}^{-1} & 0 \\ A_{44}^{-1}B_{41}A_{11}^{-1} & A_{44}^{-1}B_{42}A_{22}^{-1} & -A_{44}^{-1}A_{43}A_{33}^{-1} & A_{44}^{-1} \end{bmatrix}$$

where

$$B_{31} = A_{32}A_{22}^{-1}A_{21} - A_{31}, \quad B_{41} \quad = A_{42}A_{22}^{-1}A_{21} - A_{43}A_{33}^{-1}A_{32}A_{22}^{-1}A_{21} + A_{43}A_{33}^{-1}A_{31} - A_{41} \quad \text{and}$$
$$B_{42} = A_{43}A_{33}^{-1}A_{32} - A_{42}.$$

Lastly, the following lemma will sometimes be useful.

Lemma 2.2.7.
If $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times n}$ are matrices such that $I_n + AB$ and $I_m + BA$ are non-singular, then

$$(I_n + AB)^{-1}A = A(I_m + BA)^{-1}.$$

Proof.
Note that

$$(I_n + AB)A = A + ABA = A(I_m + BA).$$

Hence multiplying on the left by $(I_n + AB)^{-1}$ and on the right by $(I_m + BA)^{-1}$ yields the desired result. □

2.2.3 Linear Fractional Transformations

In this section we review linear fractional transformations (LFT's) and the Redheffer star product. These concepts play prominent roles in the connection of controllers to linear systems.

Linear fractional transformations or Möbius functions are well know in complex function theory. A complex function $f : \mathbb{C} \mapsto \mathbb{C}$ is a LFT if it is of the form

$$f(\lambda) = \frac{a + b\lambda}{c + d\lambda}, \qquad a, b, c, d \in \mathbb{C}, \qquad ad - bc \neq 0.$$

The above transformation can be generalized to the matrix case as follows.

Definition 2.2.8 (Lower and upper linear fractional transformations (LFTs))**.**
Given a block matrix M and matrices L and U:

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \in \mathbb{R}^{(n_1+n_2)\times(m_1+m_2)}, \quad L \in \mathbb{R}^{m_2\times n_2}, \quad U \in \mathbb{R}^{m_1\times n_1},$$

such that $I_{n_2} - M_{22}L$ and $I_{n_1} - M_{11}U$ are invertible, the *lower linear fractional transformation (lower LFT)* of M and L is defined to be the map

$$\underline{\mathcal{F}}(M,\cdot) : \mathbb{C}^{m_2\times n_2} \mapsto \mathbb{C}^{n_1\times m_1}, \qquad \underline{\mathcal{F}}(M,L) = M_{11} + M_{12}L(I - M_{22}L)^{-1}M_{21}.$$

Similarly, the *upper linear fractional transformation (upper LFT)* of M and U is defined to be the map

$$\overline{\mathcal{F}}(M,\cdot) : \mathbb{C}^{m_1\times n_1} \mapsto \mathbb{C}^{n_2\times m_2}, \qquad \overline{\mathcal{F}}(M,U) = M_{22} + M_{21}U(I - M_{11}U)^{-1}M_{12}.$$

In Chapter 4, LFT's will allow us to give representations of the connections of controllers to linear systems. We now define the Redheffer star product which will allow us to give representations of the interconnection of two linear systems (see page 52 in Section 4.1).

Definition 2.2.9 (Redheffer star product)**.**
Given two block matrices M and N with compatible dimensions

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \in \mathbb{R}^{(n_1+n_2)\times(m_1+m_2)}, \qquad N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \in \mathbb{R}^{(m_2+m_1)\times(n_2+n_1)},$$

such that $I_{n_2} - M_{22}N_{11}$ and $I_{n_1} - N_{11}M_{22}$ are invertible, the *Redheffer star product* of M and N is defined as

$$\begin{aligned} M \star N &:= \begin{bmatrix} M_{11} + M_{12}N_{11}(I_{n_2} - M_{22}N_{11})^{-1}M_{21} & M_{12}(I_{m_2} - N_{11}M_{22})^{-1}N_{12} \\ N_{21}(I_{n_2} - M_{22}N_{11})^{-1}M_{21} & N_{22} + N_{21}M_{22}(I_{m_2} - N_{11}M_{22})^{-1}M_{12} \end{bmatrix} \\ &= \begin{bmatrix} \underline{\mathcal{F}}(M, N_{11}) & M_{12}(I_{m_2} - N_{11}M_{22})^{-1}N_{12} \\ N_{21}(I_{n_2} - M_{22}N_{11})^{-1}M_{21} & \overline{\mathcal{F}}(N, M_{22}) \end{bmatrix}. \end{aligned}$$

2.3 Block Matrices and Prescribed Zero-Patterns

In this section we are interested in classes of block matrices with prescribed (block) zero-patterns, which are not necessarily square, but are closed under (block) matrix multiplication, provided the sizes of the blocks are compatible. In order to ease notation, we will define sets of block matrices and super-block matrices using so-called partitions and sub-partitions.

Definition 2.3.1 (Partition of n)**.**
Given some $n \in \mathbb{Z}_+$, we will say $\underline{n} = (n_1, n_2, \ldots, n_p) \in \mathbb{Z}_+^p$ is a *partition* of n if $|\underline{n}| := n_1 + n_2 + \ldots + n_p = n$.

Consider partitions $\underline{n} = (n_1, n_2, \ldots, n_p) \in \mathbb{Z}_+^p$ and $\underline{m} = (m_1, m_2, \ldots, m_q) \in \mathbb{Z}_+^q$ of $n = |\underline{n}|$ and $m = |\underline{m}|$ respectively. With $G = [G_{ij}] \in \mathbb{R}^{\underline{n}\times\underline{m}}$, we mean that G is an $n \times m$ matrix and a $p \times q$ block matrix whose ij^{th} block G_{ij} is a $n_i \times m_j$ matrix, that is,

$$\mathbb{R}^{\underline{n}\times\underline{m}} := \{G = [G_{ij}] \in \mathbb{R}^{n\times m} : G_{ij} \in \mathbb{R}^{n_i\times m_j}\}.$$

If $\underline{r} \in \mathbb{Z}_+^p$ is another partition, then the block matrices $G \in \mathbb{R}^{\underline{n}\times\underline{m}}$ and $H \in \mathbb{R}^{\underline{r}\times\underline{n}}$ are said to be *compatible for block matrix multiplication* HG and in that case the product HG is in $\mathbb{R}^{\underline{r}\times\underline{m}}$. These ideas are illustrated in the following example.

Example 2.3.2.
Consider the partitions $\underline{n} = (2,3)$, $\underline{m} = (1,3,2)$ and $\underline{r} = (1,1)$ and the block matrices $G = [G_{ij}] \in \mathbb{R}^{\underline{n}\times\underline{m}}$ and $H = [H_{ki}] \in \mathbb{R}^{\underline{r}\times\underline{n}}$ given by:

$$G = \left[\begin{array}{c|ccc|cc} 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{array}\right] \quad \text{and} \quad H = \left[\begin{array}{cc|ccc} 1 & 1 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 \end{array}\right].$$

Here, for example, the blocks G_{13} and H_{21} are given by

$$G_{13} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad H_{21} = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

The block matrices G and H are compatible for block matrix multiplication and their product is given by:

$$HG = \left[\begin{array}{c|ccc|cc} 3 & 0 & 3 & 1 & 3 & 1 \\ \hline 1 & 1 & 0 & 1 & 1 & 0 \end{array}\right] \in \mathbb{R}^{\underline{r}\times\underline{m}}.$$

Definition 2.3.3 (Sub-partition of n)**.**
Given some $n \in \mathbb{Z}_+$, we will say $\overline{n} = (\underline{n}_1, \underline{n}_2, \dots, \underline{n}_p) \in \mathbb{Z}_+^{rp}$ is a *sub-partition* of n if $(n_1, n_2, \dots, n_p) \in \mathbb{Z}_+^p$ is a partition of n and for $i = 1, 2, \dots, p$, each $\underline{n}_i = (n_i^1, n_i^2, \dots, n_i^r) \in \mathbb{Z}_+^r$ is a partition of n_i into r parts.

Note that a sub-partition $\overline{n}$ is itself also a partition of n and that $|\overline{n}| := n_1^1 + \ldots + n_p^r = n$. Let $\overline{n} = (\underline{n}_1, \underline{n}_2, \dots, \underline{n}_p) \in \mathbb{Z}_+^{rp}$ and $\overline{m} = (\underline{m}_1, \underline{m}_2, \dots, \underline{m}_q) \in \mathbb{Z}_+^{sq}$ be two given sub-partitions. We will write $G = [[G_{ij}^{uv}]] \in \mathbb{R}^{\overline{n}\times\overline{m}}$ in which case it is to be understood that G is a *super-block matrix* with blocks $G_{ij} \in \mathbb{R}^{\underline{r}\times\underline{s}}$ each of which is itself a block matrix $G_{ij} = [G_{ij}^{uv}] \in \mathbb{R}^{\underline{n}_i\times\underline{m}_j}$ with blocks $G_{ij}^{uv} \in \mathbb{R}^{n_i^u\times m_j^v}$, that is,

$$\mathbb{R}^{\overline{n}\times\overline{m}} := \{G = [[G_{ij}^{uv}]] \in \mathbb{R}^{n\times m} : G_{ij} = [G_{ij}^{uv}] \in \mathbb{R}^{n_i\times m_j},\ G_{ij}^{uv} \in \mathbb{R}^{n_i^u\times m_j^v}\}.$$

Again, we illustrate these ideas with an example.

Example 2.3.4.
Consider the sub-partitions $\overline{n} = (\underline{n}_1, \underline{n}_2)$ and $\overline{m} = (\underline{m}_1, \underline{m}_2, \underline{m}_3)$, where $\underline{n}_1 = (2,1)$, $\underline{n}_2 = (1,2,1)$, $\underline{m}_1 = (1,1)$, $\underline{m}_2 = (2,1)$ and $\underline{m}_3 = (1,2)$. Then an example of a super-block matrix G in $\mathbb{R}^{\overline{n}\times\overline{m}}$ is given by:

$$G = \left[\begin{array}{c:c|cc:c|c:cc} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ \hdashline 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hdashline 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ \hdashline 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right].$$

Here, for example, the block G_{23} and the sub-block G_{23}^{22} are given by

$$G_{23} = \begin{bmatrix} G_{23}^{11} & G_{23}^{12} \\ G_{23}^{21} & G_{23}^{22} \\ G_{23}^{31} & G_{23}^{32} \end{bmatrix} = \left[\begin{array}{c:cc} 0 & 0 & 1 \\ \hdashline 0 & 1 & 0 \\ 1 & 0 & 1 \\ \hdashline 1 & 0 & 0 \end{array}\right] \quad \text{and} \quad G_{13}^{22} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Next, we are concerned with the (block)zero structure of (block)matrices as determined by a binary relation. Given a binary relation $\mathcal{T} = (P, T)$ with $P = \{1, \dots, p\}$ for some positive integer p, define the set of matrices

$$\mathcal{I}_{\mathcal{T}} := \{G = [g_{ij}] \in \mathbb{R}^{p\times p} \colon g_{ij} = 0 \text{ if } \{k \in P \colon (j,k), (k,i) \in T\} = \emptyset\}.$$

By the theorem on page 258 of [7] and the subsequent remark on page 259 it follows that the set $\mathcal{I}_{\mathcal{T}}$ forms a subalgebra of $\mathbb{R}^{p\times p}$ if and only if the relation $\mathcal{T}$ is transitive. It is then referred to as the *incidence algebra* associated with $\mathcal{T}$. If $\mathcal{P} = (P, \succeq)$ is a poset, then $\mathcal{I}_{\mathcal{P}}$ is a unital matrix algebra which can also be written as

$$\mathcal{I}_{\mathcal{P}} = \{G = [g_{ij}] \in \mathbb{R}^{p\times p} \colon g_{ij} = 0 \text{ if } j \not\succeq i\}.$$

By analogy of the incidence algebras defined above, we define the following set of block matrices with a block-zero-pattern prescribed by a partial order.

Definition 2.3.5 (Block incidence vector space)**.**
Given a pair $\mathcal{T} = (P, T)$, with $P = \{1, \ldots, p\}$ and T a relation on P and partitions $\underline{n}, \underline{m} \in \mathbb{Z}_+^p$, we define the *block incidence vector space* $\mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}} \subseteq \mathbb{R}^{\underline{n}\times\underline{m}}$ as the subspace

$$\mathcal{I}_{\mathcal{T}}^{\underline{n}\times\underline{m}} := \{G = [G_{ij}] \in \mathbb{R}^{\underline{n}\times\underline{m}} \colon G_{ij} = 0 \text{ if } \{k \in P \colon (j,k),(k,i) \in T\} = \emptyset\}.$$

If the binary relation T is reflexive and transitive (that is, a pre-order), then

$$\mathcal{I}_{\mathcal{T}}^{\underline{n}\times\underline{m}} := \{G = [G_{ij}] \in \mathbb{R}^{\underline{n}\times\underline{m}} \colon G_{ij} = 0 \text{ if } (j,i) \notin T\}.$$

Similarly, if the binary relation T is also anti-symmetric, so that we have a poset $\mathcal{P} = (P, \succeq)$, then

$$\mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}} := \{G = [G_{ij}] \in \mathbb{R}^{\underline{n}\times\underline{m}} \colon G_{ij} = 0 \text{ if } j \not\succeq i\}.$$

The fact that the set $P = \{1, 2, \ldots, p\}$ is also ordered is not relevant to us, the choice to indicate P in this way is just to clarify the relation to the columns and rows of the block matrices. Furthermore, since we only consider finite sets, P can always be take in this form.

Example 2.3.6.
Consider the posets $\mathcal{P}_3$ and $\mathcal{P}_5$ given in Example 2.1.5 and partitions $\underline{n}, \underline{m} \in \mathbb{Z}_+^p$. The matrices G and H given below exhibit the block zero structures of matrices in the incidences spaces $\mathcal{I}_{\mathcal{P}_3}^{\underline{n}\times\underline{m}}$ and $\mathcal{I}_{\mathcal{P}_5}^{\underline{n}\times\underline{m}}$ respectively:

$$G = \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ 0 & G_{22} & 0 \\ 0 & 0 & G_{33} \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} H_{11} & 0 & 0 & 0 \\ 0 & H_{22} & 0 & 0 \\ H_{31} & 0 & H_{33} & 0 \\ H_{41} & H_{42} & 0 & H_{44} \end{bmatrix}.$$

The numbering of nodes of a poset $\mathcal{P}$ can always be done in such a way that the matrices in the corresponding incidence spaces are block lower triangular (for example, for the poset $\mathcal{P}_3$ in Example 2.3.6 if we reorder $(1, 2, 3) \mapsto (3, 2, 1)$, then G will be lower triangular). This is not necessarily the case for pre-orders as anti-symmetry is required to guarantee for this to happen. Note that this also means that the Hasse diagrams corresponding to posets, have no directed cycles (see also Section 5.1).

By arguments similar to those in [7] it follows that if T is reflexive and transitive, then the block zero structure is preserved under block matrix multiplication, provided the block matrices are compatible for block matrix multiplication. The block zero structure is also invariant under inversion (since the inverse of an invertible matrix A is contained in its double commutant $\{A\}''$). These facts are summarized in the following proposition.

Proposition 2.3.7.
Let $\mathcal{T} = (P, T)$ be a pre-order with p elements and let $\underline{n}, \underline{m}, \underline{r} \in \mathbb{Z}_+^p$ be given partitions. If $G \in \mathcal{I}_{\mathcal{T}}^{\underline{r}\times\underline{n}}$ and $H \in \mathcal{I}_{\mathcal{T}}^{\underline{n}\times\underline{m}}$, then $GH \in \mathcal{I}_{\mathcal{T}}^{\underline{r}\times\underline{m}}$. If $G \in \mathcal{I}_{\mathcal{T}}^{\underline{n}\times\underline{n}}$ and $\det G \neq 0$, then $G^{-1} \in \mathcal{I}_{\mathcal{T}}^{\underline{n}\times\underline{n}}$.

Throughout this thesis we work with block compressions associated with subsets of P. Note that we have defined partitions in such a way that zero entries are permitted. It will be convenient to define block compressions by simply setting some of the entries in the partitions equal to zero.

Definition 2.3.8 (Compressions of block matrices)**.**
Let $P = \{1, \ldots, p\}$ and let $R, S \subseteq P$. Let $G \in \mathbb{R}^{\underline{n} \times \underline{m}}$ for partitions $\underline{n}, \underline{m} \in \mathbb{Z}_+^p$. Then $G(R, S)$ denotes the block matrix in $\mathbb{R}^{\underline{n}_R \times \underline{m}_S}$ where

$$\underline{n}_R = (n_{1,R}, \ldots, n_{p,R}) \in \mathbb{Z}_+^p, \qquad \text{with} \qquad n_{j,R} = \begin{cases} 0 & \text{if } j \notin R \\ n_j & \text{if } j \in R \end{cases}$$

$$\underline{m}_S = (m_{1,S}, \ldots, m_{q,S}) \in \mathbb{Z}_+^q, \qquad \text{with} \qquad m_{j,S} = \begin{cases} 0 & \text{if } j \notin S \\ m_j & \text{if } j \in S \end{cases}$$

and where $G(R, S) = [\widetilde{G}_{ij}]_{i,j=1,\ldots,p}$ is defined by

$$\widetilde{G}_{ij} = G_{ij} \quad \text{if } i \in R \text{ and } j \in S \qquad \text{and} \qquad \widetilde{G}_{ij} \text{ vacuous} \quad \text{if } i \notin R \text{ or } j \notin S.$$

If R is a singleton, say $R = \{i\}$, we write $G(i, S)$ and likewise if $S = \{j\}$, we write $G(R, j)$. For one-sided compressions, we follow MATLAB® notation, and write $G(:, S)$ in case $R = P$, or $G(R, :)$ in case $S = P$.

The following example illustrates the ideas in Proposition 2.3.7 and Definition 2.3.8.

Example 2.3.9.
Consider the poset $\mathcal{P}_1$ in Example 2.1.5 and partitions $\underline{n}, \underline{m}, \underline{r} \in \mathbb{Z}_+^p$ as well as matrices $G \in \mathcal{I}_{\mathcal{P}_1}^{\underline{n} \times \underline{m}}$, $H \in \mathcal{I}_{\mathcal{P}_1}^{\underline{m} \times \underline{r}}$ and $K \in \mathcal{I}_{\mathcal{P}_1}^{\underline{n} \times \underline{n}}$. Looking at the product GH, we have

$$\begin{bmatrix} G_{11} & 0 & 0 \\ G_{21} & G_{22} & 0 \\ G_{31} & 0 & G_{33} \end{bmatrix} \begin{bmatrix} H_{11} & 0 & 0 \\ H_{21} & H_{22} & 0 \\ H_{31} & 0 & H_{33} \end{bmatrix} = \begin{bmatrix} G_{11}H_{11} & 0 & 0 \\ G_{21}H_{11} + G_{22}H_{21} & G_{22}H_{22} & 0 \\ G_{31}H_{11} + G_{33}H_{31} & 0 & G_{33}H_{33} \end{bmatrix}$$

which illustrates that $GH \in \mathcal{I}_{\mathcal{P}_1}^{\underline{n} \times \underline{r}}$. Similarly, if the inverse of K exists, then by equation (2.6)

$$\begin{bmatrix} K_{11} & 0 & 0 \\ K_{21} & K_{22} & 0 \\ K_{31} & 0 & K_{33} \end{bmatrix}^{-1} = \begin{bmatrix} K_{11}^{-1} & 0 & 0 \\ -K_{22}^{-1}K_{21}K_{11}^{-1} & K_{22}^{-1} & 0 \\ -K_{33}^{-1}K_{31}K_{11}^{-1} & 0 & K_{33}^{-1} \end{bmatrix}$$

and again, we see that $K^{-1} \in \mathcal{I}_{\mathcal{P}_1}^{\underline{n} \times \underline{n}}$.

Consider the poset $\mathcal{P}_5 = (P_5, \succeq)$ in Example 2.1.5 and partitions $\underline{n}, \underline{m} \in \mathbb{Z}_+^p$ as well as a matrix $H \in \mathcal{I}_{\mathcal{P}_5}^{\underline{n} \times \underline{m}}$ such as in Example 2.3.6. For $\mathcal{P}_5$, we have $\downarrow 1 = \{1, 3, 4\}$, $\uparrow 3 = \{1, 3\}$. Thus

$$H(\downarrow 1, 1) = \begin{bmatrix} H_{11} \\ H_{31} \\ H_{41} \end{bmatrix}, \quad H(\uparrow 3, :) = \begin{bmatrix} H_{11} & 0 & 0 & 0 \\ H_{31} & 0 & H_{33} & 0 \end{bmatrix} \quad \text{and} \quad H(4, 2) = H_{42}.$$

By standard matrix multiplication it follows for matrices $G \in \mathbb{R}^{\underline{r} \times \underline{n}}$ and $H \in \mathbb{R}^{\underline{n} \times \underline{m}}$ that $(GH)(i, j) = G(i, :)H(:, j)$. The following theorem is an important result on block compressions of products of block matrices where the matrix on the right is in a block incidence space. In particular it follows that if $G \in \mathbb{R}^{\underline{r} \times \underline{n}}$ and $H \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{m}}$, then $(GH)(i, j) = G(i, \downarrow j)H(\downarrow j, j)$, because if $k \notin \downarrow j$, then $H(k, j) = 0$ and hence has no influence on the entry $(GH)(i, j)$.

Theorem 2.3.10.
Consider a poset $\mathcal{P} = (P, \succeq)$ with $P = \{1, 2, \ldots, p\}$, partitions $\underline{n}, \underline{m}, \underline{r} \in \mathbb{Z}_+^p$ and subsets $Q, S \subseteq P$. For any block matrices $G \in \mathbb{R}^{\underline{r} \times \underline{n}}$ and $H \in \mathbb{R}^{\underline{n} \times \underline{m}}$ we have

1. *If $H \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{m}}$, then $(GH)(Q, S) = G(Q, R)H(R, S)$ for any subset $R \subseteq P$ with $\downarrow S \subseteq R$.*
2. *If $G \in \mathcal{I}_{\mathcal{P}}^{\underline{r} \times \underline{n}}$, then $(H^\intercal G^\intercal)(Q, S) = H^\intercal(Q, R)G^\intercal(R, S)$ for any subset $R \subseteq P$ with $\uparrow S \subseteq R$.*

In particular, $(GH)(Q, S) = G(Q, \downarrow S)H(\downarrow S, S)$ and $(H^\intercal G^\intercal)(Q, S) = H^\intercal(Q, \uparrow S)G^\intercal(\uparrow S, S)$.

Proof.

1. Assume that $H \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}$. The block matrices $(GH)(Q,S)$ and $G(Q,R)H(R,S)$ are equally sized. Therefore, they are equal if the corresponding block entries $(GH)(i,j)$ and $G(i,R)H(R,j)$ are equal for each $i \in Q$ and each $j \in S$. For $j,k \in P$, if $k \notin \downarrow j$, then $j \not\succeq k$ and thus $H_{kj} = H(k,j) = 0$. Let $i \in Q$ and $j \in S$ be arbitrary. Assume that $\downarrow S \subseteq R \subseteq P$. Now $\downarrow j \subseteq \downarrow S \subseteq R$, so if $k \notin R$, then $k \notin \downarrow j$. Hence $H(k,j) = 0$ if $k \notin R$. With this in mind, we now consider the $(i,j)^{\text{th}}$ block entry of the block matrix GH:

$$\begin{aligned}(GH)(i,j) = G(i,P)H(P,j) = \sum_{k=1}^{p} G(i,k)H(k,j) &= \sum_{k\notin R} G(i,k)H(k,j) + \sum_{k\in R} G(i,k)H(k,j)\\ &= 0 + \sum_{k\in R} G(i,k)H(k,j) = G(i,R)H(R,j),\end{aligned}$$

which is what we had to show.

2. Assume that $G \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{n}}$. As with the proof of 1, we need to show that $(H^\intercal G^\intercal)(i,j) = H^\intercal(i,R)G^\intercal(R,j)$ for each $i \in Q$ and each $j \in S$. For $j,k \in P$, if $k \notin \uparrow j$, then $k \not\succeq j$ and thus $G_{jk} = G(j,k) = G^\intercal(k,j) = 0$. Let $i \in Q$ and $j \in S$ be arbitrary. Assume that $\uparrow S \subseteq R \subseteq P$. Now $\uparrow j \subseteq \uparrow S \subseteq R$, so if $k \notin R$, then $k \notin \uparrow j$. Hence $G^\intercal(k,j) = 0$ if $k \notin R$. Thus

$$\begin{aligned}(H^\intercal G^\intercal)(i,j) = H^\intercal(i,P)G^\intercal(P,j) &= \sum_{k=1}^{p} H^\intercal(i,k)G^\intercal(k,j)\\ &= \sum_{k\in R} H^\intercal(i,k)G^\intercal(k,j) + \sum_{k\notin R} H^\intercal(i,k)G^\intercal(k,j)\\ &= \sum_{k\in R} H^\intercal(i,k)G^\intercal(k,j) + 0 = H^\intercal(Q,R)G^\intercal(R,S),\end{aligned}$$

which is what we had to show. □

Finally, we define the block identity matrix $I_{\underline{n}} \in \mathbb{R}^{\underline{n}\times\underline{n}}$ with respect to a partition $\underline{n} = (n_1,\ldots,n_p) \in \mathbb{Z}_+^p$ as the block diagonal matrix in $\mathbb{R}^{\underline{n}\times\underline{n}}$ with $n_i \times n_i$ identity matrices as diagonal blocks. Then, for any $S \subseteq P$, the matrix $I_{\underline{n}}(:,S)$ can be viewed as the embedding of $\mathbb{R}^{\underline{n}_S}$ into $\mathbb{R}^{\underline{n}}$ and $I_{\underline{n}}(S,:)$ as the projection from $\mathbb{R}^{\underline{n}}$ onto $\mathbb{R}^{\underline{n}_S}$.

The following result about the inverse of a block compression of an invertible matrix now follows easily from Theorem 2.3.10.

Corollary 2.3.11.
If $A \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{n}}$ is invertible, then

$$A(\downarrow j,\downarrow j)^{-1} = A^{-1}(\downarrow j,\downarrow j)$$

for each $j \in P$.

Proof.
Since $\downarrow(\downarrow j) = \downarrow j$, it follows from Theorem 2.3.10 that for any $j \in P$,

$$\begin{aligned}I_{\underline{n}}(\downarrow j,\downarrow j) &= (AA^{-1})(\downarrow j,\downarrow j)\\ &= A(\downarrow j,\downarrow(\downarrow j))A^{-1}(\downarrow(\downarrow j),\downarrow j)\\ &= A(\downarrow j,\downarrow j)A^{-1}(\downarrow j,\downarrow j).\end{aligned}$$

Similarly, $I_{\underline{n}}(\downarrow j,\downarrow j) = A^{-1}(\downarrow j,\downarrow j)A(\downarrow j,\downarrow j)$. Thus $A(\downarrow j,\downarrow j)^{-1} = A^{-1}(\downarrow j,\downarrow j)$. □

Corollary 2.3.12.
Let $H \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}$. For any $S \subseteq P$, we have $H(\mathbb{R}^{\underline{m}_{\downarrow S}}) \subseteq \mathbb{R}^{\underline{n}_{\downarrow S}}$. In particular, if $\underline{m} = \underline{n}$, then $\mathbb{R}^{\underline{m}_{\downarrow S}}$ is an H-invariant subspace.

Proof.
Apply Theorem 2.3.10 with $G = I_{\underline{n}}$, $Q = P$ and $S = \downarrow S$. This gives

$$H(:, \downarrow S) = (I_{\underline{n}} H)(:, \downarrow S) = I_{\underline{n}}(:, \downarrow S) H(\downarrow S, \downarrow S).$$

Therefore, we have

$$H(\mathbb{R}^{\underline{m}_{\downarrow S}}) = \mathrm{Im} B(:, \downarrow S) = \mathrm{Im} I_{\underline{n}}(:, \downarrow S) H(\downarrow S, \downarrow S) \subseteq \mathrm{Im} I_{\underline{n}}(:, \downarrow S) = \mathbb{R}^{\underline{n}_{\downarrow S}}.$$ □

Definition 2.3.13 (Canonical shuffle)**.**
Consider the set $\{1, 2, \dots, kp\}$ for $k, p \in \mathbb{Z}_+$ and the numbers

$$\ell_i^s := (i-1)p + s \quad \text{and} \quad q_s^i = (s-1)k + i \tag{2.8}$$

for $i = 1, \dots, k$ and $s = 1, \dots, p$. Then the *canonical shuffle* γ is the permutation on $\{1, 2, \dots, kp\}$ defined by $\gamma(\ell_i^s) = q_s^i$.

Example 2.3.14.
Suppose $k = 2$ and $p = 3$, then the following table illustrates the canonical shuffle $\gamma : \ell_i^s \mapsto q_s^i$ on the set $\{1, 2, \dots, 6\}$ with $i = 1, 2$ and $s = 1, 2, 3$:

i	s	$\ell_i^s = (i-1)p + s$	$\gamma(\ell_i^s) = q_s^i = (s-1)k + i$
1	1	1	1
1	2	2	3
1	3	3	5
2	1	4	2
2	2	5	4
2	3	6	6

Definition 2.3.15 (Block canonical shuffle matrix)**.**
Given a sub-partition $\overline{n} = ((n_1^1, \dots, n_1^p), \dots, (n_k^1, \dots, n_k^p))$, define $\widetilde{\underline{n}} = ((n_1^1, \dots, n_k^1), \dots, (n_1^p, \dots, n_k^p))$ and let ℓ_i^s and q_s^i be as in (2.8). We define the *block canonical shuffle matrix* associated with $\overline{n}$ to be the block matrix $\Gamma_{\overline{n}} \in \mathbb{R}^{\widetilde{\underline{n}} \times \overline{n}}$ given by

$$\Gamma_{\overline{n}} = \begin{bmatrix} E_1^\intercal & E_{k+1}^\intercal & \cdots & E_{(s-1)k+i}^\intercal & \cdots & E_{pk}^\intercal \end{bmatrix},$$

where $E_{q_s^i} \in \mathbb{R}^{n_i^s \times \overline{n}}$ is the $1 \times kp$ block row matrix, whose $(q_s^i)^{\text{th}}$ block entry is the identity matrix $I_{n_i^s}$. Thus $\Gamma_{\overline{n}}(:, \ell_i^s) = E_{q_s^i}^\intercal$ and $\Gamma_{\overline{n}}(q_s^i, :) = E_{\ell_i^s}$ for $i = 1, 2, \dots, k$ and $s = 1, 2, \dots, p$ or equivalently, $\Gamma_{\overline{n}}(q_s^i, \ell_i^s) = I_{n_i^s}$.

Example 2.3.16.
Suppose $\overline{n} = ((n_1^1, n_1^2, n_1^3), (n_2^1, n_2^2, n_2^3))$ is a sub-partition of n. Let $\gamma : \ell_i^s \to q_s^i$ be the canonical shuffle as in example 2.3.14. Then the block canonical shuffle matrix associated with $\overline{n}$ is given by

$$\Gamma_{\overline{n}} = \left[\begin{array}{ccc|ccc} I_{n_1^1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{n_2^1} & 0 & 0 \\ \hline 0 & I_{n_1^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{n_2^2} & 0 \\ \hline 0 & 0 & I_{n_1^3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{n_2^3} \end{array}\right] \in \mathbb{R}^{\widetilde{\underline{n}} \times \overline{n}}$$

where $\widetilde{\underline{n}} = ((n_1^1, n_2^1), (n_1^2, n_2^2), (n_1^3, n_2^3))$.

By inspection, it is clear that $\Gamma_{\overline{n}}^{\mathsf{T}}\Gamma_{\overline{n}} = I_n$ and $\Gamma_{\overline{n}}\Gamma_{\overline{n}}^{\mathsf{T}} = I_n$ so that $\Gamma_{\overline{n}}^{\mathsf{T}} = \Gamma_{\overline{n}}^{-1} = \Gamma_{\widetilde{n}}$. This also follows from the next lemma, which shows that left-multiplying by $\Gamma_{\overline{n}}$ shuffles the block rows of a super-block matrix according to the canonical shuffle and that right-multiplying a super block matrix by $\Gamma_{\overline{n}}^{\mathsf{T}}$ shuffles its block columns.

Lemma 2.3.17.
Let $\underline{\overline{n}} = ((n_1^1,\ldots,n_1^p),\ldots,(n_k^1,\ldots,n_k^p))$ be a given sub-partition and let $\underline{\widetilde{n}} = ((n_1^1,\ldots,n_k^1),\ldots,(n_1^p,\ldots,n_k^p))$. Let $\gamma : \ell_i^s \to q_s^i$ be the canonical shuffle with ℓ_i^s and q_s^i as in (2.8). Let $\Gamma_{\underline{\overline{n}}}$ be the block canonical shuffle matrix associated with $\underline{\overline{n}}$. If $\underline{r} = (r_1,\ldots,r_p)$ and $\underline{m} = (m_1,\ldots,m_p)$ are partitions and we have matrices $G \in \mathbb{R}^{\underline{\overline{n}}\times\underline{m}}$ and $H \in \mathbb{R}^{\underline{r}\times\underline{\overline{n}}}$, then

$$(\Gamma_{\underline{\overline{n}}}G)(q_s^i,:) = G(\ell_i^s,:) \quad \textit{and} \quad (H\Gamma_{\underline{\overline{n}}}^{\mathsf{T}})(:,\ q_s^i) = H(:,\ell_i^s) \quad \textit{and hence}$$
$$(\Gamma_{\underline{\overline{n}}}G) \in \mathbb{R}^{\underline{\widetilde{n}}\times\underline{m}} \quad \textit{and} \quad (H\Gamma_{\underline{\overline{n}}}^{\mathsf{T}}) \in \mathbb{R}^{\underline{r}\times\underline{\widetilde{n}}}.$$

In particular, if $K \in \mathbb{R}^{\underline{\overline{n}}\times\underline{\overline{n}}}$, then $(\Gamma_{\underline{\overline{n}}}K\Gamma_{\underline{\overline{n}}}^{\mathsf{T}}) \in \mathbb{R}^{\underline{\widetilde{n}}\times\underline{\widetilde{n}}}$ and $(\Gamma_{\underline{\overline{n}}}K\Gamma_{\underline{\overline{n}}}^{\mathsf{T}})(q_s^i,q_t^j) = K(\ell_i^s,\ell_j^t)$ for $i,j = 1,\ldots,k$ and $s,t = 1,\ldots,p$.

Proof.
Suppose $G \in \mathbb{R}^{\underline{\overline{n}}\times\underline{m}}$. Then by definition of the block canonical shuffle matrix $\Gamma_{\underline{\overline{n}}}$ and the block row matrices $E_{\ell_i^s}$ and by theorem 2.3.10, we have that

$$\begin{aligned}
(\Gamma_{\underline{\overline{n}}}G)(q_s^i,:) &= \begin{bmatrix}\Gamma_{\underline{\overline{n}}}(q_s^i,:)G(:,1) & \ldots & \Gamma_{\underline{\overline{n}}}(q_s^i,:)G(:,p)\end{bmatrix} \\
&= \begin{bmatrix}E_{\ell_i^s}G(:,1) & \ldots & E_{\ell_i^s}G(:,p)\end{bmatrix} \\
&= \begin{bmatrix}G(\ell_i^s,1) & \ldots & G(\ell_i^s,p)\end{bmatrix} \\
&= G(\ell_i^s,:).
\end{aligned}$$

The proofs for $(H\Gamma_{\underline{\overline{n}}}^{\mathsf{T}})(:,q_s^i) = H(:,\ell_i^s)$ and $(\Gamma_{\underline{\overline{n}}}K\Gamma_{\underline{\overline{n}}}^{\mathsf{T}})(q_s^i,q_t^j) = K(\ell_i^s,\ell_j^t)$ follow similarly. □

The following example illustrates the idea captured in Lemma 2.3.17. Multiplying on the left and right with the block canonical shuffle matrix and its transpose respectively, shuffles the block entries of a super block matrix in a natural way.

Example 2.3.18.
Let $\underline{\overline{n}}$ and $\Gamma_{\underline{\overline{n}}}$ be the sub-partition and associated block canonical shuffle given in Example 2.3.16. If $A = [[A_{ij}^{uv}]] \in \mathbb{R}^{\underline{\overline{n}}\times\underline{\overline{n}}}$ is given by

$$A = \left[\begin{array}{ccc|ccc}
A_{11}^{11} & A_{11}^{12} & A_{11}^{13} & A_{12}^{11} & A_{12}^{12} & A_{12}^{13} \\
A_{11}^{21} & A_{11}^{22} & A_{11}^{23} & A_{12}^{21} & A_{12}^{22} & A_{12}^{23} \\
A_{11}^{31} & A_{11}^{32} & A_{11}^{33} & A_{12}^{31} & A_{12}^{32} & A_{12}^{33} \\
\hline
A_{21}^{11} & A_{21}^{12} & A_{21}^{13} & A_{22}^{11} & A_{22}^{12} & A_{22}^{13} \\
A_{21}^{21} & A_{21}^{22} & A_{21}^{23} & A_{22}^{21} & A_{22}^{22} & A_{22}^{23} \\
A_{21}^{31} & A_{21}^{32} & A_{21}^{33} & A_{22}^{31} & A_{22}^{32} & A_{22}^{33}
\end{array}\right] \in \mathbb{R}^{\underline{\overline{n}}\times\underline{\overline{n}}},$$

then

$$\Gamma_{\underline{\overline{n}}}A\Gamma_{\underline{\overline{n}}}^{\mathsf{T}} = \left[\begin{array}{cc|cc|cc}
A_{11}^{11} & A_{12}^{11} & A_{11}^{12} & A_{12}^{12} & A_{11}^{13} & A_{12}^{13} \\
A_{21}^{11} & A_{22}^{11} & A_{21}^{12} & A_{22}^{12} & A_{21}^{13} & A_{22}^{13} \\
\hline
A_{11}^{21} & A_{12}^{21} & A_{11}^{22} & A_{12}^{22} & A_{11}^{23} & A_{12}^{23} \\
A_{21}^{21} & A_{22}^{21} & A_{21}^{22} & A_{22}^{22} & A_{21}^{23} & A_{22}^{23} \\
\hline
A_{11}^{31} & A_{12}^{31} & A_{11}^{32} & A_{12}^{32} & A_{11}^{33} & A_{12}^{33} \\
A_{21}^{31} & A_{22}^{31} & A_{21}^{32} & A_{22}^{32} & A_{21}^{33} & A_{22}^{33}
\end{array}\right] \in \mathbb{R}^{\underline{\widetilde{n}}\times\underline{\widetilde{n}}}$$

where $\underline{\widetilde{n}} = ((n_1^1,n_2^1),(n_1^2,n_2^2),(n_1^3,n_2^3))$.

The following result now follows from Lemma 2.3.17 and shows that if a super-block matrix has block entries that all have the same zero-pattern, then left multiplying by the block canonical shuffle and right multiplying by its transpose, transfers the zero-pattern the super-block level.

Corollary 2.3.19.
Suppose $\overline{\underline{n}} = (\underline{n}_1, \dots, \underline{n}_k) \in \mathbb{Z}^{kp}$ *where* $\underline{n}_i = (n_i^1, \dots, n_i^p) \in \mathbb{Z}^p$ *and* $\underline{r} = (r_1, \dots, r_p) \in \mathbb{Z}^p$ *and* $\underline{m} = (m_1, \dots, m_p) \in \mathbb{Z}^p$ *are sub-partitions and partitions of* n, r, m. *If* $\widetilde{n} = ((n_1^1 + \dots + n_k^1), \dots, (n_1^p + \dots + n_k^p)) \in \mathbb{Z}^p$ *and*

$$A = [[A_{ij}^{st}]] \in \mathbb{R}^{\overline{\underline{n}} \times \overline{\underline{n}}}, \qquad B = [B_i^{st}] \in \mathbb{R}^{\overline{\underline{n}} \times \underline{m}} \qquad \textit{and} \qquad C = [[C_j^{st}]] \in \mathbb{R}^{\underline{r} \times \overline{\underline{n}}}, \ \textit{with}$$
$$A_{ij} = [A_{ij}^{st}] \in \mathcal{I}_{\mathcal{T}}^{\underline{n}_i \times \underline{n}_j}, \qquad B_i = [B_i^{st}] \in \mathcal{I}_{\mathcal{T}}^{\underline{n}_i \times \underline{m}} \qquad \textit{and} \qquad C_j = [C_j^{st}] \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{n}_j}$$

for $i, j = 1, \dots, k$ *and some relation* $\mathcal{T} = (P, T)$ *with* $P = \{1, \dots, p\}$, *then*

$$(\Gamma_{\overline{\underline{n}}} A \Gamma_{\overline{\underline{n}}}^{\intercal}) \in \mathcal{I}_{\mathcal{T}}^{\widetilde{n} \times \widetilde{n}}, \quad (\Gamma_{\overline{\underline{n}}} B) \in \mathcal{I}_{\mathcal{T}}^{\widetilde{n} \times \underline{m}} \quad \textit{and} \quad (C \Gamma_{\overline{\underline{n}}}^{\intercal}) \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \widetilde{n}}.$$

Proof.
Define $\check{\underline{n}} = ((n_1^1, \dots, n_k^1), \dots, (n_1^p, \dots, n_k^p))$ and let $\gamma : \ell_i^s \to q_s^i$ be the canonical shuffle for $i = 1, \dots, k$ and $s = 1, \dots, p$. Then $\check{\underline{n}}$ is a sub-partition of $\widetilde{n}$ and by Lemma 2.3.17, $(\Gamma_{\overline{\underline{n}}} A \Gamma_{\overline{\underline{n}}}^{\intercal}) =: \widetilde{A} = [\widetilde{A}_{st}] = [[\widetilde{A}_{st}^{ij}]] \in \mathbb{R}^{\check{\underline{n}} \times \check{\underline{n}}} \subset \mathbb{R}^{\widetilde{n} \times \widetilde{n}}$ with

$$\widetilde{A}_{st}^{ij} = \widetilde{A}(q_s^i, k_t^j) = A(\ell_i^s, \ell_j^t) = A_{ij}^{st}$$

for $i, j = 1, \dots, k$ and $s, t = 1, \dots, p$. Suppose $A_{ij} = [A_{ij}^{st}] \in \mathcal{I}_{\mathcal{T}}^{\underline{n}_i \times \underline{n}_j}$ for $i, j = 1, \dots, k$. Then $\widetilde{A}_{st}^{ij} = A_{ij}^{st} = 0$ if $\{r \in P \colon (t, r), (r, s) \in T\} = \emptyset$ for $i, j = 1, \dots, k$ and $s, t = 1, \dots, p$. And hence $\widetilde{A}_{st} = [A_{st}^{ij}] = 0$ if $\{r \in P \colon (t, r), (r, s) \in T\} = \emptyset$. This shows that $\widetilde{A} = [\widetilde{A}_{st}] \in \mathcal{I}_{\mathcal{T}}^{\widetilde{n} \times \widetilde{n}}$. The proofs that $(\Gamma_{\overline{\underline{n}}} B) \in \mathcal{I}_{\mathcal{T}}^{\widetilde{n} \times \underline{m}}$ and $(C \Gamma_{\overline{\underline{n}}}^{\intercal}) \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \widetilde{n}}$ follow similarly. □

Example 2.3.20.
Consider the binary relation $\mathcal{T} = (P, T)$ where $P = \{1, \dots, 6\}$ and $T = \{(1,1), (2,2), (2,3), (3,1)\}$ and let $\overline{\underline{n}} = ((n_1^1, n_1^2, n_1^3), (n_2^1, n_2^2, n_2^3))$. Suppose $A = [[A_{ij}^{st}]] \in \mathbb{R}^{\overline{\underline{n}} \times \overline{\underline{n}}}$ is given by

$$A = \left[\begin{array}{ccc|ccc}
A_{11}^{11} & 0 & A_{11}^{13} & A_{12}^{11} & 0 & A_{12}^{13} \\
0 & A_{11}^{22} & 0 & 0 & A_{12}^{22} & 0 \\
0 & A_{11}^{32} & 0 & 0 & A_{12}^{32} & 0 \\ \hline
A_{21}^{11} & 0 & A_{21}^{13} & A_{22}^{11} & 0 & A_{22}^{13} \\
0 & A_{21}^{22} & 0 & 0 & A_{22}^{22} & 0 \\
0 & A_{21}^{32} & 0 & 0 & A_{22}^{32} & 0
\end{array}\right].$$

Then $A_{ij} = [A_{ij}^{st}] \in \mathcal{I}_{\mathcal{T}}^{\underline{n}_i \times \underline{n}_j}$ for $i, j = 1, 2$. Let $\Gamma_{\overline{\underline{n}}}$ be the block canonical shuffle matrix associated with $\overline{\underline{n}}$. Then, when we compute $\Gamma_{\overline{\underline{n}}} A \Gamma_{\overline{\underline{n}}}^{\intercal}$, we see that

$$\Gamma_{\overline{\underline{n}}} A \Gamma_{\overline{\underline{n}}}^{\intercal} = \left[\begin{array}{cc|cc|cc}
A_{11}^{11} & A_{12}^{11} & 0 & 0 & A_{11}^{13} & A_{12}^{13} \\
A_{21}^{11} & A_{22}^{11} & 0 & 0 & A_{21}^{13} & A_{22}^{13} \\ \hline
0 & 0 & A_{11}^{22} & A_{12}^{22} & 0 & 0 \\
0 & 0 & A_{21}^{22} & A_{22}^{22} & 0 & 0 \\ \hline
0 & 0 & A_{11}^{32} & A_{12}^{32} & 0 & 0 \\
0 & 0 & A_{21}^{32} & A_{22}^{32} & 0 & 0
\end{array}\right] \in \mathcal{I}_{\mathcal{T}}^{\widetilde{n} \times \widetilde{n}},$$

where $\widetilde{n} = ((n_1^1 + n_2^1), (n_1^2 + n_2^2), (n_1^3 + n_2^3))$.

2.4 Hardy Spaces

This subsection gives a brief review of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ spaces. This summary is based on chapter 4 of [56]. As is well known, the transfer functions of causal LTI systems correspond with proper rational matrix functions, which, in the case that the functions are also stable, form a subspace of $\mathcal{H}_\infty$ or $\mathcal{H}_2$ (in the strictly proper case). We make use of the $\|\cdot\|_2$-norm in a later chapter on optimal control (Chapter 4).

The Hilbert space $\mathcal{L}_2^{n\times m}(i\mathbb{R})$ is defined to be the space of Lebesgue measurable matrix-valued functions $\widehat{F} : i\mathbb{R} \mapsto \mathbb{C}^{n\times m}$ such that the following integral on the imaginary axis is bounded:

$$\int_{-\infty}^{\infty} \operatorname{trace}[\widehat{F}(i\omega)^*\widehat{F}(i\omega)]\ \mathrm{d}\omega < \infty,$$

where * is the conjugate transpose of a complex matrix. We will write $\mathcal{L}_2^{n\times m}$ instead of $\mathcal{L}_2^{n\times m}(i\mathbb{R})$ and if the size of the matrix function is clear from the context, we simply write $\mathcal{L}_2$, and similarly for other classes of matrix functions, which we will define shortly. The inner product on $\mathcal{L}_2$ is defined as

$$\langle \widehat{F}, \widehat{G}\rangle := \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[\widehat{F}(i\omega)^*\widehat{G}(i\omega)]\ \mathrm{d}\omega.$$

We note that

$$\begin{aligned}
\langle \widehat{F}, \widehat{G}\widehat{H}\rangle &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[\widehat{F}(i\omega)^*\widehat{G}(i\omega)\widehat{H}(i\omega)]\ \mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[(\widehat{G}(i\omega)^*\widehat{F}(i\omega))^*\widehat{H}(i\omega)]\ \mathrm{d}\omega \\
&= \langle \widehat{G}^*\widehat{F}, \widehat{H}\rangle
\end{aligned}$$

The above inner product induces the following 2-norm on $\mathcal{L}_2^{n\times m}$:

$$\|\widehat{F}\|_2 := \sqrt{\langle \widehat{F}, \widehat{F}\rangle} = \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[\widehat{F}(i\omega)^*\widehat{F}(i\omega)]\ \mathrm{d}\omega}. \tag{2.9}$$

By the properties of the trace it follows that

$$\begin{aligned}
\|\widehat{F}^\intercal\|_2 &= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[((\widehat{F}(i\omega))^\intercal)^*(\widehat{F}(i\omega))^\intercal]\ \mathrm{d}\omega} \\
&= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[(\widehat{F}(i\omega)(\widehat{F}(i\omega))^*)^\intercal]\ \mathrm{d}\omega} \\
&= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[\widehat{F}(i\omega)^*\widehat{F}(i\omega)]\ \mathrm{d}\omega} \\
&= \|\widehat{F}\|_2
\end{aligned} \tag{2.10}$$

The (closed) subspace $\mathcal{H}_2^{n\times m} \subset \mathcal{L}_2^{n\times m}$ consists of matrix-valued $F : \mathbb{C} \mapsto \mathbb{C}^{n\times m}$ that are analytic on the open right-half plane $\mathbb{C}^+ = \{\sigma + i\omega \in \mathbb{C} : \sigma > 0\}$, with corresponding norm

$$\|\widehat{F}\|_2^2 = \sup_{\sigma>0}\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[\widehat{F}^*(\sigma+i\omega)\widehat{F}(\sigma+i\omega)]\ \mathrm{d}\omega\right\}.$$

It can be shown that the $\mathcal{H}_2$-norm is the same as the $\mathcal{L}_2$-norm, so that

$$\|\widehat{F}\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[\widehat{F}(i\omega)^*\widehat{F}(i\omega)]\ \mathrm{d}\omega \qquad \text{for } \widehat{F} \in \mathcal{H}_2.$$

The space $(\mathcal{H}_2^{n\times m})^\perp$ is the orthogonal complement of $\mathcal{H}_2^{n\times m}$ in $\mathcal{L}_2^{n\times m}$. In fact, $\mathcal{H}_2^\perp$ is the closed subspace of $\mathcal{L}_2$ that consists of matrix valued functions $\widehat{F} : \mathbb{C} \mapsto \mathbb{C}^{n\times m}$ that are analytic on the open left-half plane $\mathbb{C}^- = \{\sigma + i\omega \in \mathbb{C} : \sigma < 0\}$.

The set $\mathcal{L}_\infty^{n\times m}(i\mathbb{R})$ is the Banach space of matrix-valued functions $\widehat{F} : \mathbb{C} \mapsto \mathbb{C}^{n\times m}$ that are essentially bounded on the imaginary axis $i\mathbb{R}$, with norm

$$\|\widehat{F}\|_\infty := \operatorname*{ess\,sup}_{\omega\in\mathbb{R}} \Big(\text{trace}\,[\widehat{F}(i\omega)^*\widehat{F}(i\omega)]\Big)\,.$$

The closed subspace $\mathcal{H}_\infty \subset \mathcal{L}_\infty(i\mathbb{R})$ consists of functions that are analytic and bounded on the open right half-plane $\mathbb{C}^+$ with norm

$$\|\widehat{F}\|_\infty := \sup_{\text{Re}(\lambda)>0} \Big(\text{trace}\,[\widehat{F}(\lambda)^*\widehat{F}(\lambda)]\Big) = \operatorname*{ess\,sup}_{\omega\in\mathbb{R}} \Big(\text{trace}\,[\widehat{F}(i\omega)^*\widehat{F}(i\omega)]\Big)\,.$$

Similarly, the closed subspace $\mathcal{H}_\infty^- \subset \mathcal{L}_\infty(i\mathbb{R})$ consists of functions that are analytic and bounded on the open left-half plane $\mathbb{C}^-$ with norm

$$\|\widehat{F}\|_\infty := \sup_{\text{Re}(\lambda)<0} \Big(\text{trace}\,[\widehat{F}(\lambda)^*\widehat{F}(\lambda)]\Big) = \operatorname*{ess\,sup}_{\omega\in\mathbb{R}} \Big(\text{trace}\,[\widehat{F}(i\omega)^*\widehat{F}(i\omega)]\Big)\,.$$

We will use the prefix $\mathcal{R}$ to indicate the subspace of real rational matrix functions, so that

- $\mathcal{RH}_2$ is the subspace of $\mathcal{H}_2$ that consists of *strictly* proper real rational stable matrix functions (that is, all the poles are in the open left-half plane).
- $\mathcal{RH}_2^\perp$ is the subspace of $\mathcal{H}_2$ that consists of strictly proper real rational anti-stable matrix functions (that is, all the poles are in the open right-half plane).
- $\mathcal{RH}_\infty$ is the subspace of $\mathcal{H}_2$ that consists of (not necessarily strictly) proper real rational stable matrix functions. So $\mathcal{RH}_2 \subseteq \mathcal{RH}_\infty$.
- $\mathcal{RH}_\infty^-$ is the subspace of $\mathcal{H}_2$ that consists of proper real rational anti-stable matrix functions.

Chapter 3

Linear Systems

In this chapter, we give a review of some concepts and results related to causal linear time invariant (LTI) systems. Although it is difficult to find a single source that summarizes all the results, more information can be found in [8], [13] and [56]. We include some proofs for the sake of clarity and completeness or because the proof illustrates an idea that will be utilized later on. In the section on operations on systems (Section 3.5), we include some results which will be used in later sections in computations. Specifically, Lemmas 3.5.1, 3.5.5 and 3.5.6 are stated with proofs and are useful for computations later in the thesis.

3.1 Rational Matrix Functions

Recall that a polynomial matrix $A(\lambda)$ of degree n and size m is a function $A : \mathbb{C} \mapsto \mathbb{C}^{m\times m}$ of the form

$$A(\lambda) = A_0 + A_1\lambda + \ldots + A_n\lambda^n,$$

where $A_j \in \mathbb{C}^{m\times m}$ for each $j = 1, \ldots, n$. In Definition 2.2.1, we defined the companion matrix associated with a real monic polynomial p. The characteristic polynomial of the companion matrix of p is exactly the polynomial p. By analogy of Definition 2.2.1, given a real monic polynomial matrix

$$L(\lambda) = \lambda^n I_m + \lambda^{n-1}L_{n-1} + \ldots + \lambda L_1 + L_0 = \lambda^n I_n + \sum_{j=0}^{n-1} \lambda^j L_j,$$

where $L_j \in \mathbb{R}^{m\times m}$ for $j = 0, \ldots, n-1$, its *block companion matrix* is defined to be the matrix $G \in \mathbb{R}^{mn\times mn}$ given by:

$$G = \begin{bmatrix} 0 & I_m & 0 & \ldots & 0 \\ 0 & 0 & I_m & \ldots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & I_m \\ -L_0 & -L_1 & -L_2 & \ldots & -L_{n-1} \end{bmatrix}.$$

As in the scalar case, it can be shown that

$$p_G(\lambda) = \det(L(\lambda)), \qquad \lambda \in \mathbb{C},$$

in particular, the zeroes of the scalar polynomial $\det(L(\lambda))$ coincide with the eigenvalues of G.

A *real rational function* $w : \mathbb{C} \to \mathbb{R}$ is the quotient of two real polynomials p (of degree m) and $q \neq 0$ (of degree n):

$$w(\lambda) = \frac{p(\lambda)}{q(\lambda)} = \frac{p_m\lambda^m + p_{m-1}\lambda^{m-1} + \ldots + p_1\lambda + p_0}{q_n\lambda^n + q_{n-1}\lambda^{n-1} + \ldots + q_1\lambda + q_0}. \tag{3.1}$$

In equation (3.1), it is assumed that p and q have no common factors. The rational function w has a *zero* at λ_0 if $p(\lambda_0) = 0$ and a *pole* at λ_0 if $q(\lambda_0) = 0$. If $w(\infty) := \lim_{|\lambda|\to\infty} w(\lambda)$ exists (equivalently if $m \leq n$), then w is a *proper* rational function. If $w(\infty) = 0$ (equivalently $m < n$), then w is a *strictly proper* rational function. A proper rational function w can be written as

$$w(\lambda) = \frac{p_1(\lambda)}{q_1(\lambda)} + w(\infty),$$

where the degree of the polynomial p_1 is strictly less than that of the polynomial q_1.

Similarly, a matrix function $W = [w_{ij}] : \lambda \to [w_{ij}(\lambda)]$ is *real rational* matrix function if each of its entries w_{ij} is a real rational function and we say W has a pole at λ_0 if any w_{ij} has a pole at λ_0. Furthermore, W is *proper* if $\lim_{|\lambda|\to\infty} W(\lambda) =: W(\infty)$ exists and *strictly proper* if $W(\infty) = 0$. As in the scalar case, a proper rational matrix function can be written in terms of matrix polynomials and the limit at infinity.

Proposition 3.1.1 (cf. the proof of Theorem 2.3.3 in [13])**.**
If W is a proper real rational matrix function, then

$$W(\lambda) = H(\lambda)L(\lambda)^{-1} + W(\infty)$$

for a matrix polynomial H and a square monic matrix polynomial L of degree strictly greater than H.

3.2 Transfer Functions and State Space Realizations

The following summary is based mainly based on [13] and [8]. More details on transfer functions and state space realizations can be found in these sources. A continuous time (that is $t \in [0, \infty)$) causal linear time-invariant (LTI) input-output system Σ with inputs $u(t) \in \mathbb{R}^m$ and outputs $y(t) \in \mathbb{R}^r$ and with input-output map $G_\Sigma : u \mapsto y$ has a *state space realization* if there exists a positive integer $n \in \mathbb{Z}_+$, a *state* $x : [0, \infty) \mapsto \mathbb{R}^n$ and matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{r\times n}$ and $D^{r\times m}$ such that the solution of the system of equations

$$\Sigma \begin{cases} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{cases} \qquad x(0) = x_0 \qquad t \geq 0, \tag{3.2}$$

gives $y(t) = (G_\Sigma(u))(t)$ for $t \geq 0$. In this case we write $\Sigma \sim (A, B, C, D)$ to denote that (3.2) is the state space realization of the system Σ. If $\lambda \in \mathbb{C}$ is not an eigenvalue of A, that is $\lambda \in \rho(A)$, then $(\lambda I - A)^{-1}$ exists and the system (3.2) has a transfer function $\widehat{W}$ given by

$$\widehat{W}(\lambda) = C(\lambda I - A)^{-1}B + D \qquad \text{for } \lambda \in \rho(A). \tag{3.3}$$

Taking the Laplace transforms $\widehat{u}$ and $\widehat{y}$ of the input u and output y, the transfer function $\widehat{W}$, gives the input-output relationship $\widehat{y} = \widehat{W}\widehat{u}$ in the special case that $x(0) = 0$. Consequently, an input-output system Σ has a state space realization determined by (A, B, C, D) if its transfer function $\widehat{W}$ is given by equation (3.3). By Cramer's rule,

$$(\lambda I - A)^{-1} = \frac{\text{adj}((\lambda I - A))}{\det(\lambda I - A)},$$

where $\text{adj}((\lambda I - A))$ is the adjugate of $(\lambda I - A)$. Recall that the *adjugate* of an $n \times n$ matrix is the transpose of its cofactor matrix and that it has order less than n. On the other hand, $\det(\lambda I - A)$ is the characteristic polynomial of A, which has order n. As a consequence, the transfer function $\widehat{W}$ in (3.3) of a system $\Sigma \sim (A, B, C, D)$ is a proper real rational matrix function (and $\widehat{W}$ is strictly proper if $D = 0$). The following well-known theorem asserts that the converse also holds.

Theorem 3.2.1 (cf. Theorem 2.3.3 in [13])**.**
If $\widehat{W}$ is the transfer function of causal LTI system Σ, then Σ has a state space realization if and only if $\widehat{W}$ is a proper real rational matrix function.

3.3 Controllability and Observability

Consider an LTI system $\Sigma \sim (A, B, C, D)$ as in (3.2). The state equation of the system is

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0, \qquad t \geq 0. \tag{3.4}$$

For any given initial state x_0, the state x at a final time $t > 0$ and any valid input $u : [0, t] \to \mathbb{R}^m$ can be obtained by the classical integral formula

$$x(x_0, u, t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\ \mathrm{d}\tau.$$

The system Σ, or equivalently the pair (A, B), is controllable if it is possible to steer any initial state to any other state in any given non-zero time by the appropriate choice of input.

Definition 3.3.1 (Controllable (cf. p.28 in [56]))**.**
The system $\Sigma \sim (A, B, C, D)$, or the pair (A, B), is said to be *controllable* if for any initial state $x(0) = x_0$, final time $t_1 > 0$ and final state $\xi \in \mathbb{R}^n$, there exists a valid input $u : [0, t] \mapsto \mathbb{R}^m$ such that $x(x_0, u, t_1) = \xi$.

The related concept of reachability is concerned with the same question, except that the initial state is taken to be zero, that is, $x_0 = 0$. In the case of causal continuous time LTI systems, it turns out that controllability and reachability coincide.

Definition 3.3.2 (Reachable, reachable set (cf. p.63 in [8]))**.**
A state $\xi \in \mathbb{R}^n$ is said to be *reachable* at time $t > 0$ if there exists a valid input function u such that $x(0, u, t) = \xi$. For any fixed time t, the *reachable set* $\mathcal{R}_t$ is defined as

$$\mathcal{R}_t = \{\xi \in \mathbb{R}^n : x(0, u, t) = \xi \text{ for some valid input function } u : [0, \infty) \mapsto \mathbb{R}^m\}.$$

Due to the linearity of the state equation, the reachable set $\mathcal{R}_t$ is a subspace of the state space.

Definition 3.3.3 (Controllability matrix, controllable subspace)**.**
The *controllability matrix* $\mathcal{C}(A, B)$ associated with the matrix pair (A, B) is defined to be

$$\mathcal{C}(A, B) = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix} \in \mathbb{R}^{n \times nm}.$$

The *reachable subspace* $\mathcal{R}(A, B)$ is defined to the image of the controllability matrix $\mathcal{C}(A, B)$, that is,

$$\mathcal{R}(A, B) = \operatorname{Im} \begin{bmatrix} B & AB & A^2B & \dots A^{n-1}B \end{bmatrix}.$$

Yet another concept associated with the state equation is the *controllability gramian*.

Definition 3.3.4 (Controllability gramian)**.**
For each $t > 0$, the (time dependent) controllability gramian of a matrix pair (A, B) is the $n \times n$ matrix

$$X_c^t = \int_0^t e^{A\tau}BB^\intercal e^{A^\intercal \tau}\ \mathrm{d}\tau$$

and the *controllability gramian* of a matrix pair (A, B) is the $n \times n$ matrix

$$X_c = \int_0^\infty e^{At}BB^\intercal e^{A^\intercal t}\ \mathrm{d}t.$$

These Gramians are also solutions to the Lyapunov equations $AX + XA^\intercal + BB^\intercal = 0$ and $A^\intercal Y + YA + C^\intercal C = 0$ (see Theorem 3.6.3). It turns out that the previous three concepts are intimately connected. Indeed, for each $t > 0$, it holds that (see for example Theorem 2.2 in [8]),

$$\mathcal{R}_t = \mathcal{R}(A, B) = \operatorname{Im} X_c^t = \operatorname{Im} X_c.$$

Due to this result, there is no ambiguity in referring to $\mathcal{R}(A, B)$ as the reachable subspace. The following theorem gives equivalent conditions for controllability. In particular, it follows that controllability and reachability coincide, that is if you can reach any state from zero in any given finite time, then you can also reach that state in the same time from any other initial state.

Theorem 3.3.5 (cf. example Theorem 3.1 in [56])**.**
The following statements are equivalent

1. *The pair (A, B) is controllable;*
2. $\mathcal{R}(A, B) = \mathbb{R}^n$*;*
3. *The controllability matrix $\mathcal{C}(A, B)$ has full row rank n;*
4. $X_c^t > 0$ *for any* $t > 0$*;*
5. $X_c > 0$.

Since the reachable space $\mathcal{R}(A, B)$ is the image of the controllability matrix, this gives a simple way to determine whether a system is controllable or not. The following theorem provides an alternative method for checking the controllability of a system.

Theorem 3.3.6 (Popov-Belevitch-Hautus (PBH) test (cf. Theorem 2.14 in [8]))**.**
The pair (A, B) is controllable if and only if for each $\lambda \in \mathbb{C}$ it holds that

$$\operatorname{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n.$$

In light of the previous theorem, the following definition is made.

Definition 3.3.7 (Controllable eigenvalue)**.**
Suppose $A \in \mathbb{R}^{n \times n}$. An eigenvalue λ of A is said to be a *controllable eigenvalue* of A if

$$\operatorname{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n.$$

If λ does not satisfy the above condition, then it is said to be an *uncontrollable eigenvalue*.

Hence, the PBH test can be restated as follows: the pair (A, B) is controllable if and only if each eigenvalue of A is controllable.

If we apply a static state feedback $u(t) = Fx(t)$ for some $F \in \mathbb{R}^{m \times n}$, then the state equation (3.4) becomes

$$\dot{x}(t) = (A + BF)x(t), \qquad x(0) = x_0.$$

The question of *eigenvalue assignment* is whether we can specify the eigenvalues of $A+BF$ exactly by choice of F. That is, given values $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ (with complex numbers coming in conjugate pairs), is it possible to find a $F \in \mathbb{R}^{m \times n}$ such that $\sigma(A + BF) = \{\lambda_1, \dots, \lambda_n\}$? The answer is that this is possible exactly when (A, B) is controllable.

Theorem 3.3.8 (cf. Theorem 2.19 in [8] and Theorem 3.1 in [56])**.**
The eigenvalues of $A + BF$ are freely assignable by choice of F if and only if (A, B) is controllable.

In particular, we may want to choose F in such a way that all the eigenvalues of $A + BF$ have negative real parts, that is, such that $A + BF$ is Hurwitz.

Definition 3.3.9 (Stabilizable)**.**
The pair (A, B) is said to be stabilizable if there exists a $F \in \mathbb{R}^{m \times n}$ such that $A + BF$ is Hurwitz.

As a corollary to the PBH test, we have the following result. This shows, that in order to test for controllability, we only have to apply the PBH test for eigenvalues with non-negative real parts.

Corollary 3.3.10 (cf. Corollary 2.15 in [8])**.**
The pair (A, B) is stabilizable if and only if

$$\operatorname{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n, \qquad \textit{for all } \lambda \in \overline{\mathbb{C}}^{+}.$$

We now turn to the dual notions of observability and detectability. In contrast to controllability, where we considered a system (3.4) without an output $y(t)$, we now consider a system without an input $u(t)$:

$$\begin{aligned} \dot{x}(t) &= Ax(t), \qquad x(0) = x_0 \\ y(t) &= Cx(t). \end{aligned}$$

The second fundamental question in this section, that of observability, is concerned with whether we can determine the value of the initial state x_0 by measuring the output $y(t)$ over some time interval. Clearly, the output at any given time $t > 0$ can be solved by $y(t) = Ce^{At}x_0$.

Definition 3.3.11 (Observable)**.**
A system $\Sigma \sim (A, B, C, D)$, or equivalently, the pair (C, A) is said to be *observable* if at any given time $T > 0$, it is possible to determine the initial state x_0 from the outputs $y(t)$ over the time interval $[0, T]$.

Definition 3.3.12 (Observability matrix, unobservable subspace)**.**
The *observability matrix* associated with the pair (C, A) is defined to be

$$\mathcal{O}(C, A) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}.$$

The *unobservable subspace* $\mathcal{N}(C, A)$ is defined to be the null space of the observability matrix, that is

$$\mathcal{N}(C, A) = \ker \mathcal{O}(C, A).$$

Definition 3.3.13 (Observability gramian)**.**
For each $t > 0$, the (time dependent) observability gramian of a matrix pair (C, A) is the $n \times n$ matrix

$$Y_o^t = \int_0^t e^{A^\mathsf{T}\tau} C^\mathsf{T} C e^{A\tau} \, \mathrm{d}\tau$$

and the *observability gramian* of a matrix pair (C, A) is the $n \times n$ matrix

$$Y_o = \int_0^\infty e^{A^\mathsf{T}t} C^\mathsf{T} C e^{At} \, \mathrm{d}t.$$

As is well known, the concepts of controllability and observability are dual notions in the following sense.

Theorem 3.3.14 (cf. Proposition 2.21 in [8])**.**
The pair (C, A) is observable if and only if $(A^\mathsf{T}, C^\mathsf{T})$ is controllable.

As a consequence, via duality, we obtain results about observability from those about controllability.

Theorem 3.3.15 (cf. Proposition 2.22 in [8] or Theorem 3.3 in [56])**.**
The following are equivalent:

1. *(C, A) is observable.*
2. *$Y_o^t > 0$ for any $t > 0$.*
3. *$\ker \mathcal{O}(C, A) = \{0\}$.*
4. *$\operatorname{rank} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n$ for all $\lambda \in \mathbb{C}$ (PBH test).*
5. *The eigenvalues of $A + LC$ can be freely assigned by choice of L.*

Similarly, the dual notion of stabilizability is detectability.

Definition 3.3.16 (Detectable).
The pair (C, A) is said to be *detectable* if there exists an $L \in \mathbb{R}^{n\times r}$ such that $A + LC$ is stable.

Theorem 3.3.17 (cf. Theorem 3.4 in [56]).
The following are equivalent

1. (C, A) *is detectable.*
2. $\operatorname{rank}\begin{bmatrix} A-\lambda I \\ C \end{bmatrix} = n$ *for all* $\lambda \in \overline{\mathbb{C}}^{+}$.
3. $(A^{\mathsf{T}}, C^{\mathsf{T}})$ *is stabilizable.*

3.4 Equivalent Realizations, Minimality and the Kalman Decomposition

This section is based mainly on [56]. The concepts studied in this section were first developed by R.E. Kalman in [19]. We consider a continuous time causal LTI system Σ with input-output map $G_\Sigma : u \mapsto y$ with $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^r$ for all $t > 0$. Recall that a state space system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0 \\ y(t) &= Cx(t) + Du(t), \quad t \geq 0, \end{aligned}$$

with $x(t) \in \mathbb{R}^{n\times n}$ for $t > 0$, may be solved to obtain

$$y(x_0, u, t) = Ce^{At}x_0 + \int_0^t Ce^{A(t-\tau)}Bu(\tau)\ \mathrm{d}\tau + Du(t).$$

Also recall that (A, B, C, D) is a realization of G_Σ, written $\Sigma \sim (A, B, C, D)$, if

$$G_\Sigma(u(t)) = y(x_0, u, t)$$

for each initial state x_0, each input function u and each time $t > 0$. We now consider the case where $x_0 = 0$ and define equivalent realizations as realizations which give the same input-output map.

Definition 3.4.1 (Equivalent realizations).
Suppose that $A \in \mathbb{R}^{n\times n}$, $\widetilde{A} \in \mathbb{R}^{n_1\times n_1}$, $B \in \mathbb{R}^{n\times m}$, $\widetilde{B} \in \mathbb{R}^{n_1\times m}$, $C \in \mathbb{R}^{r\times n}$, $\widetilde{C} \in \mathbb{R}^{r\times n_1}$ and $D, \widetilde{D} \in \mathbb{R}^{r\times n}$ for some n and n_1. We say that (A, B, C, D) and $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ are equivalent realizations if

$$\int_0^t Ce^{A(t-\tau)}Bu(\tau)\ \mathrm{d}\tau + Du(t) = \int_0^t \widetilde{C}e^{\widetilde{A}(t-\tau)}\widetilde{B}u(\tau)\ \mathrm{d}\tau + \widetilde{D}u(t)$$

for all inputs $u : [0, \infty) \mapsto \mathbb{R}^m$ and all times $t > 0$.

Due to the matrix exponential appearing in the above integrals, an alternative characterization of equivalent realizations may obtained.

Lemma 3.4.2 (cf. Lemma 2.25 and Lemma 2.26 in [8]).
Two realizations (A, B, C, D) *and* $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ *are equivalent if and only if* $D = \widetilde{D}$ *and*

$$Ce^{At}B = \widetilde{C}e^{\widetilde{A}t}\widetilde{B} \qquad \text{for all } t \geq 0$$

if and only if $D = \widetilde{D}$ *and*

$$CA^kB = \widetilde{C}\widetilde{A}^k\widetilde{B} \qquad \text{for all } k \geq 0.$$

Among all equivalent realizations we are interested in finding those which are optimal in some sense. One sense in which a system may be considered to be optimal, is that the system matrices be of as small as possible dimension. The *dynamic order* of a realization (A, B, C, D) is defined to be the dimension of the square matrix A, that is if $A \in \mathbb{R}^{n\times n}$, then the dynamic order of (A, B, C, D) is n.

Definition 3.4.3 (Minimal realization)**.**
A realization $\Sigma \sim (A, B, C, D)$ with dynamic order n is said to be *minimal* if $n \leq \widetilde{n}$ for all other equivalent realizations $\Sigma \sim (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$.

We are interested in finding among all equivalent realizations those which are minimal. For this we will make use of state space similarities.

Definition 3.4.4 (State space similar)**.**
Two realizations (A, B, C, D) and $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ are state space similar if there exists a non-singular matrix $T \in \mathbb{R}^{n\times n}$ such that

$$\widetilde{A} = TAT^{-1}, \qquad \widetilde{B} = TB, \qquad \widetilde{C} = CT^{-1} \quad \text{and} \quad \widetilde{D} = D.$$

In that case the matrix T is called a state space similarity.

Since $CA^kB = CT^{-1}(TAT^{-1})^kTB$, it follows from Lemma 3.4.2, that all state space similar realizations are equivalent realizations. Furthermore, controllability and observability are preserved under state space similarity.

Proposition 3.4.5 (cf. Proposition 2.10 in [8])**.**
If $T \in \mathbb{R}^{n\times n}$ is non-singular, then

1. *(A, B) is controllable if and only if (TAT^{-1}, TB) is controllable.*
2. *(C, A) is observable if and only if (CT^{-1}, TAT^{-1}) is observable.*

Recall that a subspace $\mathcal{V} \subset \mathbb{R}^n$ is A-invariant if $A\mathcal{V} \subset \mathcal{V}$. If $\mathcal{V}$ is also r-dimensional with $r < n$, then it is possible to find a non singular matrix $T \in \mathbb{R}^{n\times n}$ such that

$$T\mathcal{V} = \operatorname{Im}\begin{bmatrix} * \\ 0 \end{bmatrix} \quad \text{and} \quad TAT^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

with $A_{11} \in \mathbb{R}^{r\times r}$.

Proposition 3.4.6 (cf. Proposition 2.11 in [8])**.**
The reachable subspace $\mathcal{R}(A, B)$ is the smallest A-invariant subspace that contains $\operatorname{Im} B$ and the unobservable subspace $\mathcal{N}(C, A)$ is the smallest A-invariant subspace that contains $\operatorname{Im}(C^\perp)$.

Theorem 3.4.7 (cf. Theorem 2.12 in [8] and Proposition 2.22 in [56])**.**

1. *If $\dim \mathcal{R}(A, B) = q$, then there exists a non-singular $T \in \mathbb{R}^{n\times n}$ such that*

$$TAT^{-1} = \begin{bmatrix} A_c & A_{c\bar{c}} \\ 0 & A_{\bar{c}} \end{bmatrix} \quad \text{and} \quad TB = \begin{bmatrix} B_c \\ 0 \end{bmatrix} \qquad \text{where } A_c \in \mathbb{R}^{q\times q} \text{ and } B_c \in \mathbb{R}^{q\times m} \tag{3.5}$$

   and the pair (A_c, B_c) is controllable.

2. *If $\dim \mathcal{N}(A, B) = n - s$, then there exists a non-singular $T \in \mathbb{R}^{n\times n}$ such that*

$$TAT^{-1} = \begin{bmatrix} A_o & 0 \\ A_{\bar{o}o} & A_{\bar{o}} \end{bmatrix} \quad \text{and} \quad CT^{-1} = \begin{bmatrix} C_o & 0 \end{bmatrix} \qquad \text{where } A_o \in \mathbb{R}^{s\times s} \text{ and } C_o \in \mathbb{R}^{r\times s} \tag{3.6}$$

   and the pair (C_o, A_o) is observable.

If the pair (A, B) is transformed into the form (3.5), then we say that (A, B) is in *controllability form*. Similarly, if the pair (C, A) is transformed into the form (3.6), then we say that (C, A) is in *observability form*. Suppose some realization (A, B, C, D) is transformed into controllability form as in (3.5) by a non-singular matrix T and that

$$CT^{-1} = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

Then (A, B, C, D) and $(TAT^{-1}, TB, CT^{-1}, D)$ are equivalent realizations and for all $k \geq 0$ it holds that

$$CT^{-1}(TAT^{-1})^k TB = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} A_c & A_{c\bar{c}} \\ 0 & A_{\bar{c}} \end{bmatrix}^k \begin{bmatrix} B_c \\ 0 \end{bmatrix} = C_1 A_c^k B_c.$$

Thus (A, B, C, D) and (A_c, B_c, C_1, D) are equivalent realizations and the dynamic order of (A_c, B_c, C_1, D) is less than that of (A, B, C, D) if (A, B) is not controllable. Similarly it is possible to find a realization of lower dynamic order than that of (A, B, C, D) if (C, A) is not observable. This suggests that the necessary and sufficient conditions for the minimality of a realization is controllability and observability.

Theorem 3.4.8 (cf. Theorem 2.28 and Proposition 2.27 in [8])**.**
A realization (A, B, C, D) is minimal if and only if (A, B) is controllable and (C, A) is observable.

If a realization (A, B, C, D) is not minimal, one way of obtaining a minimal system that has the same input-output map goes through the famous *Kalman decomposition*. The Kalman decomposition is a generalization of the controllability and observability forms in (3.5) and (3.6). Before continuing with the Kalman decomposition, we state the following result which gives some general identities in finite dimensional inner product spaces.

Lemma 3.4.9.
Let $\mathcal{Y}_1, \mathcal{Y}_2, \ldots, \mathcal{Y}_n$ be subspaces of a finite dimensional inner product space $\mathcal{Y}$. Then

$$\left(\bigcap_{i=1}^{n} \mathcal{Y}_i \right)^{\perp} = \sum_{i=1}^{n} \mathcal{Y}_i^{\perp}. \tag{3.7}$$

Similarly

$$\bigcap_{i=1}^{n} \mathcal{Y}_i^{\perp} = \left(\sum_{i=1}^{n} \mathcal{Y}_i \right)^{\perp}. \tag{3.8}$$

Lastly, if $P_{\mathcal{Y}_2}$ is the orthogonal projection in $\mathcal{Y}$ onto $\mathcal{Y}_2$, then

$$P_{\mathcal{Y}_2}(\mathcal{Y}_1^{\perp}) = \mathcal{Y}_2 \ominus (\mathcal{Y}_1 \cap \mathcal{Y}_2). \tag{3.9}$$

Proof.
The first identity (3.7) follows from extending the well known and easily proved identity $(\mathcal{Y}_1 \cap \mathcal{Y}_2)^{\perp} = \mathcal{Y}_1^{\perp} + \mathcal{Y}_2^{\perp}$. Since we work in finite dimensional spaces, we have $(\mathcal{Y}_i^{\perp})^{\perp} = \mathcal{Y}_i$, and thus (3.8) also follows from (3.7). For the last identity (3.9), we note that from the first identity (3.7), it follows that

$$\mathcal{Y}_1^{\perp} + \mathcal{Y}_2^{\perp} = (\mathcal{Y}_1 \cap \mathcal{Y}_2)^{\perp} = \mathcal{Y}_2 \ominus (\mathcal{Y}_1 \cap \mathcal{Y}_2) \oplus \mathcal{Y}_2^{\perp}.$$

Projecting onto $\mathcal{Y}_2$ on both sides yields

$$P_{\mathcal{Y}_2}(\mathcal{Y}_1^{\perp}) = P_{\mathcal{Y}_2}(\mathcal{Y}_2 \ominus (\mathcal{Y}_1 \cap \mathcal{Y}_2) \oplus \mathcal{Y}_2^{\perp}) = \mathcal{Y}_2 \ominus (\mathcal{Y}_1 \cap \mathcal{Y}_2),$$

as claimed. □

Since controllability and observability play critical roles in minimality, the state space is partitioned in terms of the subspaces $\mathcal{R} = \mathcal{R}(A, B)$ and $\mathcal{N} = \mathcal{N}(C, A)$. Define the subspaces

$$\mathcal{X}_{co} := \mathcal{R} \ominus (\mathcal{R} \cap \mathcal{N}), \quad \mathcal{X}_{c\bar{o}} := \mathcal{R} \cap \mathcal{N}, \quad \mathcal{X}_{\bar{c}o} := (\mathcal{R} + \mathcal{N})^{\perp} \quad \text{and} \quad \mathcal{X}_{\bar{c}\bar{o}} := \mathcal{N} \ominus (\mathcal{R} \cap \mathcal{N}).$$

Alternative formulas $\mathcal{X}_{co} = P_{\mathcal{R}}(\mathcal{N}^{\perp})$ and $\mathcal{X}_{\overline{co}} = P_{\mathcal{N}}(\mathcal{R}^{\perp})$ also follow from (3.9) in Lemma 3.4.9. With the above subspaces of $\mathcal{X}$ we obtain the following orthogonal sum decompositions:

$$\begin{aligned} \mathcal{X}_{co} \oplus \mathcal{X}_{c\overline{o}} &= \mathcal{R}, \\ \mathcal{X}_{c\overline{o}} \oplus \mathcal{X}_{\overline{co}} &= \mathcal{N}, \\ \mathcal{X}_{co} \oplus \mathcal{X}_{c\overline{o}} \oplus \mathcal{X}_{\overline{co}} &= \mathcal{R} + \mathcal{N} \\ \mathcal{X}_{co} \oplus \mathcal{X}_{c\overline{o}} \oplus \mathcal{X}_{\overline{c}o} \oplus \mathcal{X}_{\overline{co}} &= \mathcal{X} \end{aligned}$$

Since $\mathcal{R}$ and $\mathcal{N}$ are the smallest invariant subspaces of A that contain $\mathrm{Im}B$ and $(\mathrm{Im}C)^{\perp}$, respectively, it follows that with respect to this decomposition of the state space, there exists an invertible matrix T which decomposes the system matrices A, B and C into the *Kalman decomposition of* (A, B, C):

$$TAT^{-1} = \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix}, \quad TB = \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix}, \quad CT^{-1} = \begin{bmatrix} C_1 & 0 & C_3 & 0 \end{bmatrix}.$$

Furthermore it holds that

1. (A_{11}, B_1) is controllable and (C_1, A_{11}) is observable;
2. $\left(\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}\right)$ is controllable and $\left(\begin{bmatrix} C_1 & C_2 \end{bmatrix}, \begin{bmatrix} A_{11} & A_{13} \\ 0 & A_{33} \end{bmatrix}\right)$ is observable.

The *Kalman reduction* of (A, B, C, D) is the realization (A_{11}, B_1, C_1, D). Now it can be seen that

$$CA^kB = CT^{-1}(TAT^{-1})^kTB = C_1A_{11}^kB_1, \quad k = 0, 1, \ldots$$

Hence (A, B, C, D) and (A_{11}, B_1, C_1, D) are equivalent realizations. Furthermore, since (A_{11}, B_1) is controllable and (C_1, A_{11}) is observable, the Kalman reduction of (A, B, C, D) is minimal.

In Section 6.5, we will consider systems with additional structure. We will have to consider subspaces that are larger or smaller than $\mathcal{R}$ and $\mathcal{N}$ to maintain the structure. In the setting of the Kalman reduction, one can compress the system to a subspace of $\mathbb{R}^n$ which contains $\mathcal{X}_{co}$ in such a way that the moments are maintained. The next lemma provides a suggestion for such a subspace.

Lemma 3.4.10.
Consider a state space system $\Sigma \sim (A, B, C, D)$ with reachable space $\mathcal{R}$ and unobservable space $\mathcal{N}$. Suppose $\mathcal{R}' \subseteq \mathcal{R} \subseteq \mathcal{R}''$ and $\mathcal{N}' \subseteq \mathcal{N}$ are subspaces. Define $\mathcal{X}_1' := \mathcal{R}'' \ominus (\mathcal{R}' \cap \mathcal{N}')$. Then $\mathcal{X}_{co} \subseteq \mathcal{X}_1'$. Furthermore, if A', B' and C' are the compressions of A, B and C to $\mathcal{X}_1'$, then (A, B, C, D) and (A', B', C', D) are equivalent realizations.

Proof.
By definition, $\mathcal{X}_{co} = \mathcal{R} \ominus (\mathcal{R} \cap \mathcal{N})$. Since, $\mathcal{R} \subseteq \mathcal{R}''$ and $(\mathcal{R}' \cap \mathcal{N}') \subseteq (\mathcal{R} \cap \mathcal{N})$, it follows that $\mathcal{X}_{co} \subseteq \mathcal{X}'$, which is the first part.

Next, we prove the second part. By analogy of the Kalman decomposition, define

$$\begin{aligned} \mathcal{X}_1' &:= \mathcal{R}'' \ominus (\mathcal{R}' \cap \mathcal{N}'), \\ \mathcal{X}_2' &:= \mathcal{R}' \cap \mathcal{N}', \\ \mathcal{X}_3' &:= (\mathcal{R}'' + \mathcal{N}')^{\perp}, \\ \mathcal{X}_4' &:= \mathcal{N}' \ominus (\mathcal{R}' \cap \mathcal{N}') = P_{\mathcal{N}'}((\mathcal{R}')^{\perp}). \end{aligned}$$

Then we have the following orthogonal sum decompositions:

$$\begin{aligned} \mathcal{X}_1' \oplus \mathcal{X}_2' &= \mathcal{R}'', \\ \mathcal{X}_2 \oplus \mathcal{X}_4 &= \mathcal{N}', \\ \mathcal{X}_1' \oplus \mathcal{X}_2' \oplus \mathcal{X}_4' &= \mathcal{R}'' + \mathcal{N}', \\ \mathcal{X}_1' \oplus \mathcal{X}_2' \oplus \mathcal{X}_3' \oplus \mathcal{X}_4' &= \mathbb{R}^n. \end{aligned}$$

Since $\mathcal{R}' \subseteq \mathcal{R}$ and $\mathcal{N}' \subseteq \mathcal{N}$, we have $\mathcal{X}_1' \subseteq \mathcal{X}_{c\bar{o}}$. Also $\mathcal{R} \subseteq \mathcal{R}''$. Thus

$$\mathcal{R}'' = \mathcal{R} \oplus \mathcal{Z}_1 \quad \text{and} \quad \mathcal{X}_{c\bar{o}} = \mathcal{X}_2' \oplus \mathcal{Z}_2$$

where $\mathcal{Z}_1 = \mathcal{R}'' \ominus \mathcal{R}$ and $\mathcal{Z}_2 = \mathcal{X}_{c\bar{o}} \ominus \mathcal{X}_2'$. This leads to the following decomposition of $\mathbb{R}^n$:

$$\begin{aligned}\mathbb{R}^n &= \mathcal{R}'' \oplus \mathcal{R}''^{\perp} \\ &= (\mathcal{R} \oplus \mathcal{Z}_1) \oplus \mathcal{R}''^{\perp} \\ &= (\mathcal{X}_{co} \oplus \mathcal{X}_{c\bar{o}}) \oplus \mathcal{Z}_1 \oplus \mathcal{R}''^{\perp} \\ &= \mathcal{X}_{co} \oplus \mathcal{X}_2' \oplus \mathcal{Z}_2 \oplus \mathcal{Z}_1 \oplus \mathcal{R}''^{\perp}.\end{aligned}$$

With respect to this decomposition of $\mathcal{X}$, the matrices A, B and C decompose as:

$$A = \left[\begin{array}{ccc|cc} A_{11} & 0 & 0 & A_{14} & A_{15} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} \\ \hline 0 & 0 & 0 & A_{44} & A_{45} \\ 0 & 0 & 0 & A_{54} & A_{55} \end{array}\right], \quad B = \left[\begin{array}{c} B_1 \\ B_2 \\ B_3 \\ \hline 0 \\ 0 \end{array}\right], \qquad C = \left[\begin{array}{ccc|cc} C_1 & 0 & 0 & C_4 & C_5 \end{array}\right].$$

The left bottom zero block in A and the zeroes in B are due to the fact that $\mathcal{R} = \mathcal{X}_{co} \oplus \mathcal{X}_2' \oplus \mathcal{Z}_2$ is an A-invariant subspace of $\mathbb{R}^n$ that contains $\mathrm{Im}B$. The two zeroes in the left upper block of A and the zeroes in C are due to the fact that $\mathcal{X}_{c\bar{o}} = \mathcal{X}_2' \oplus \mathcal{Z}_2 \subseteq \mathcal{N}$, which is an A-invariant subspace of $\mathcal{X}$ that contains $(\mathrm{Im}C)^{\perp}$.

Now we compress A, B and C to the subspace $\mathcal{X}_1'$, which is given by

$$\begin{aligned}\mathcal{X}_1' &= \mathcal{R}'' \ominus (\mathcal{R}' \cap \mathcal{N}') \\ &= (\mathcal{R} \oplus \mathcal{Z}_1) \ominus \mathcal{X}_2' \\ &= (\mathcal{X}_{co} \oplus \mathcal{X}_2' \oplus \mathcal{Z}_2 \oplus \mathcal{Z}_1) \ominus \mathcal{X}_2' \\ &= \mathcal{X}_{co} \oplus \mathcal{Z}_2 \oplus \mathcal{Z}_1,\end{aligned}$$

which yields the matrices A', B' and C' given by:

$$A' = \begin{bmatrix} A_{11} & 0 & A_{14} \\ A_{31} & A_{33} & A_{34} \\ 0 & 0 & A_{44} \end{bmatrix}, \quad B' = \begin{bmatrix} B_1 \\ B_3 \\ 0 \end{bmatrix}, \quad C' = \begin{bmatrix} C_1 & 0 & C_4 \end{bmatrix}.$$

It follows that A'^k has the form

$$A'^k = \begin{bmatrix} A_{11}^k & 0 & * \\ * & A_{33}^k & * \\ 0 & 0 & A_{44}^k \end{bmatrix},$$

with $*$ indicating unspecified entries. Therefore, we now see that

$$CA^kB = C_1 A_{11}^k B_1 = C'A'^k B', \quad k = 0, 1, \ldots$$

Thus by Lemma 3.4.2, (A, B, C, D) and (A', B', C', D) are equivalent realizations. □

Remark 3.4.11.

In Lemma 3.4.10 we take $\mathcal{R}'' \ominus (\mathcal{R}' \cap \mathcal{N}')$, for subspaces $\mathcal{R}' \subseteq \mathcal{R} \subseteq \mathcal{R}''$ and $\mathcal{N}' \subseteq \mathcal{N}$, as an upper bound for $\mathcal{X}_{co} = \mathcal{R} \ominus (\mathcal{R} \cap \mathcal{N})$. The alternative formula $\mathcal{X}_{co} = P_{\mathcal{R}}(\mathcal{N}^{\perp})$ suggests we could also consider the subspace $P_{\mathcal{R}''}(\mathcal{N}'^{\perp}) = \mathcal{R}'' \ominus (\mathcal{R}'' \cap \mathcal{N}')$. However, for this choice, the inclusion $\mathcal{X}_{co} \subseteq P_{\mathcal{R}''}(\mathcal{N}'^{\perp})$ need not hold. For instance, one can construct a system with state space $\mathbb{R}^3$, reachable space $\mathcal{R} = \mathrm{span}\{e_1 + e_2 + e_3\}$ and unobservable space $\mathcal{N} = \mathrm{span}\{e_1, e_2 - e_3\}$, in which case $\mathcal{N}^{\perp} = \{e_2 + e_3\}$ and $\mathcal{X}_{co} = P_{\mathcal{R}}(\mathcal{N}^{\perp}) = \mathcal{R}$. Taking $\mathcal{R}'' = \mathrm{span}\{e_1 + e_2, e_3\}$ and $\mathcal{N}' = \mathcal{N}$, we find that $P_{\mathcal{R}''}(\mathcal{N}'^{\perp}) = \mathrm{span}\{\frac{1}{2}\sqrt{2}(e_1 + e_2) + e_3\}$ which does not contain $\mathcal{X}_{co}$.

3.5 Basic Operations on Realizations

In this section we consider some basic operations on state space realizations of systems. Any LTI system with a real rational transfer function $\widehat{W}$ has a state space realization (A, B, C, D) (see Theorem 3.2.1). In that case, it holds that

$$\widehat{W}(\lambda) = C(\lambda I - A)^{-1}B + D, \qquad \lambda \in \rho(A) \subset \mathbb{C}.$$

We will employ the following notation for the resolvent of a square matrix $A \in \mathbb{R}^{n\times n}$:

$$\Phi_A(\lambda) := (\lambda I - A)^{-1}. \tag{3.10}$$

Here I is the $n \times n$ identity matrix and $\lambda \in \mathbb{C}$ is a complex number for which the above inverse exists. The following lemma relays some basic identities involving resolvents. These identities will be used frequently in subsequent proofs and computations.

Lemma 3.5.1.
Let $A, X \in \mathbb{R}^{n\times n}$. For the resolvent Φ_A of A, it holds that

1. $X = \Phi_A^{-1} - \Phi_{A+X}^{-1}$ *(or equivalently* $X = \Phi_{A-X}^{-1} - \Phi_A^{-1}$*);*
2. $\Phi_A X \Phi_{A+X} = \Phi_{A+X} - \Phi_A$ *(or equivalently* $\Phi_{A-X} X \Phi_A = \Phi_A - \Phi_{A-X}$*).*

Proof.
The proofs follow simply from the definition of Φ_A in (3.10):

1. $\Phi_A^{-1} - \Phi_{A+X}^{-1} = (\lambda I - A) - (\lambda I - (A + X)) = X$.
2. By 1, it follows that

$$\begin{aligned}\Phi_A X \Phi_{A+X} &= [\Phi_A(\Phi_A^{-1} - \Phi_{A+X}^{-1})]\Phi_{A+X}\\ &= [I - \Phi_A\Phi_{A+X}^{-1}]\Phi_{A+X}\\ &= \Phi_{A+X} - \Phi_A.\end{aligned}$$

The identities in brackets follow by replacing A with $A - X$. □

As in [56] and [8], the following standard notation will be used for state space realizations of proper real rational matrix functions $\widehat{W}$:

$$\widehat{W}(\lambda) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right](\lambda) = C(\lambda I - A)^{-1}B + D = C\Phi_A(\lambda)B + D.$$

We will often omit λ and simply write $\widehat{W} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = C\Phi_A B + D$.

Definition 3.5.2 (Conjugate (Definition 3.8 [56]))**.**
The conjugate $\widehat{W}^*(\lambda)$ of a proper real rational matrix function $\widehat{W}(\lambda)$ is defined as

$$W^*(\lambda) = \big(W(-\overline{\lambda})\big)^*$$

where the star operation on the right is the conjugate transpose of the matrix $W(-\overline{\lambda})$.

We note that in taking the conjugate of Φ_A, we get

$$\Phi_A^*(\lambda) = \Phi_A(-\overline{\lambda})^* = (-\overline{\lambda}I - A)^{-*} = ((-\overline{\lambda}I - A)^*)^{-1} = -(\lambda I + A^\intercal)^{-1} = -\Phi_{-A^\intercal}(\lambda). \tag{3.11}$$

Again, we will usually omit λ and just write Φ_A instead of $\Phi_A(\lambda)$, so that $\Phi_A^* = -\Phi_{-A^\intercal}$.

The following proposition is the main result of this section and shows how we can apply standard operations on state space realizations of systems.

Proposition 3.5.3 (cf. pages 34-35 in [56] and page 51 in [55]).
Given state space realizations

$$\widehat{W} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad \widehat{W}_1 = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \quad \text{and} \quad \widehat{W}_2 = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right],$$

the following holds, assuming that the matrices have the appropriate sizes in each case:

$$\widehat{W}_1 + \widehat{W}_2 = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array}\right] = D_1 + D_2 + C_1\Phi_{A_1}B_1 + C_2\Phi_{A_2}B_2; \tag{3.12}$$

$$\begin{aligned} \widehat{W}_1\widehat{W}_2 &= \left[\begin{array}{cc|c} A_1 & B_1C_2 & B_1D_2 \\ 0 & A_2 & B_2 \\ \hline C_1 & D_1C_2 & D_1D_2 \end{array}\right] = \left[\begin{array}{cc|c} A_2 & 0 & B_2 \\ B_1C_2 & A_1 & B_1D_2 \\ \hline D_1C_2 & C_1 & D_1D_2 \end{array}\right] \\ &= D_1D_2 + D_1C_2\Phi_{A_2}B_2 + C_1\Phi_{A_1}B_1D_2 + C_1\Phi_{A_1}B_1C_2\Phi_{A_2}B_2; \end{aligned} \tag{3.13}$$

$$\widehat{W}^{-1} = \left[\begin{array}{c|c} A - BD^{-1}C & -BD^{-1} \\ \hline D^{-1}C & D^{-1} \end{array}\right] = D^{-1} - D^{-1}C\Phi_{(A-BD^{-1}C)}BD^{-1} \tag{3.14}$$

provided that D is invertible;

$$\widehat{W}^\intercal = \left[\begin{array}{c|c} A^\intercal & C^\intercal \\ \hline B^\intercal & D^\intercal \end{array}\right] = D^\intercal + B^\intercal\Phi_{A^\intercal}C^\intercal$$

$$\widehat{W}^* = \left[\begin{array}{c|c} -A^\intercal & -C^\intercal \\ \hline B^\intercal & D^\intercal \end{array}\right] = D^\intercal - B^\intercal\Phi_{-A^\intercal}C^\intercal = D^\intercal + B^\intercal\Phi_A^*C^\intercal \tag{3.15}$$

$$\begin{bmatrix} \widehat{W}_1 & \widehat{W}_2 \end{bmatrix} = \left[\begin{array}{cc|cc} A_1 & 0 & B_1 & 0 \\ 0 & A_2 & 0 & B_2 \\ \hline C_1 & C_2 & D_1 & D_2 \end{array}\right]. \tag{3.16}$$

Proof.
We only prove (3.14), the other identities follow easily. Let $\widehat{V}$ be given by the formula in (3.14). We show that $\widehat{V}$ is the inverse of $\widehat{W}$. By applying formula (3.13), it follows that

$$\begin{aligned} \widehat{W}\widehat{V} &= \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \left[\begin{array}{c|c} A - BD^{-1}C & -BD^{-1} \\ \hline D^{-1}C & D^{-1} \end{array}\right] \\ &= \left[\begin{array}{cc|c} A & BD^{-1}C & BD^{-1} \\ 0 & A - BD^{-1}C & -BD^{-1} \\ \hline C & DD^{-1}C & DD^{-1} \end{array}\right] \\ &= \begin{bmatrix} C & C \end{bmatrix} \begin{bmatrix} \Phi_A & \Phi_A(BD^{-1}C)\Phi_{A-BD^{-1}C} \\ 0 & \Phi_{A-BD^{-1}C} \end{bmatrix} \begin{bmatrix} BD^{-1} \\ -BD^{-1} \end{bmatrix} + I \end{aligned}$$

and by Lemma 3.5.1 (2), we then get

$$\begin{aligned} \widehat{W}\widehat{V} &= \begin{bmatrix} C & C \end{bmatrix} \begin{bmatrix} \Phi_A & \Phi_A - \Phi_{A-BD^{-1}C} \\ 0 & \Phi_{A-BD^{-1}C} \end{bmatrix} \begin{bmatrix} BD^{-1} \\ -BD^{-1} \end{bmatrix} + I \\ &= C\Phi_ABD^{-1} - C\Phi_ABD^{-1} + C\Phi_{A-BD^{-1}C}BD^{-1} - C\Phi_{A-BD^{-1}C}BD^{-1} + I \\ &= I. \end{aligned}$$

$\widehat{V}\widehat{W} = I$ follows similarly. Hence $\widehat{W}$ is invertible and $\widehat{W}^{-1} = \widehat{V}$. □

Of course, these operations can also be used in combination, for example

$$\widehat{W}^*\widehat{W} = \left[\begin{array}{cc|c} -A^\intercal & -C^\intercal C & -C^\intercal D \\ 0 & A & B \\ \hline B^\intercal & D^\intercal C & D^\intercal D \end{array}\right] = D^\intercal D + D^\intercal C\Phi_A B + B^\intercal \Phi_A^* C^\intercal D + B^\intercal \Phi_A^* C^\intercal C \Phi_A B \tag{3.17}$$

and we can prove that $(\widehat{W}_1\widehat{W}_2)^\intercal = \widehat{W}_2^\intercal \widehat{W}_1^\intercal$ and $(\widehat{W}_1\widehat{W}_2)^* = \widehat{W}_2^*\widehat{W}_1^*$.

The following lemma shows that left and right inverses can also be determined. The proof follows analogously to the proof of given in Proposition 3.5.3.

Lemma 3.5.4.
Suppose $D_1^\intercal D_1 > 0$ and $D_2 D_2^\intercal > 0$ and that

$$\widehat{W}_1 = \left[\begin{array}{c|c} A & B \\ \hline C & D_1 \end{array}\right] \quad \text{and} \quad \widehat{W}_2 = \left[\begin{array}{c|c} A & B \\ \hline C & D_2 \end{array}\right].$$

Define

$$\widehat{W}_1^\dagger = \left[\begin{array}{c|c} A - BD_1^\dagger C & -BD_1^\dagger \\ \hline D_1^\dagger C & D_1^\dagger \end{array}\right] \quad \text{and} \quad \widehat{W}_2^\ddagger = \left[\begin{array}{c|c} A - BD_2^\ddagger C & -BD_2^\ddagger \\ \hline D_2^\ddagger C & D_2^\ddagger \end{array}\right]$$

where

$$D_1^\dagger = (D_1^\intercal D_1)^{-1} D_1^\intercal \quad \text{and} \quad D_2^\ddagger = D_2^\intercal (D_2 D_2^\intercal)^{-1}.$$

Then $\widehat{W}_1^\dagger \widehat{W}_1 = I$ and $\widehat{W}_2 \widehat{W}_2^\ddagger = I$.

The following lemma will be used often during proofs. It shows how, in a special case where the entries take on a specific form, a state space realization may be written as the sum of two smaller state space realizations.

Lemma 3.5.5.
If the matrices below have the appropriate compatible sizes, then

$$\left[\begin{array}{cc|c} A_1 & XA_1 - A_2X & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array}\right] = \left[\begin{array}{c|c} A_1 & B_1 + XB_2 \\ \hline C_1 & D_1 \end{array}\right] + \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 - C_1X & D_2 \end{array}\right] \quad \text{and} \tag{3.18}$$

$$\left[\begin{array}{cc|c} A_1 & XA_2 - A_1X & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array}\right] = \left[\begin{array}{c|c} A_1 & B_1 - XB_2 \\ \hline C_1 & D_1 \end{array}\right] + \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 + C_1X & D_2 \end{array}\right]. \tag{3.19}$$

Similarly

$$\left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ XA_1 - A_2X & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array}\right] = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 + C_2X & D_1 \end{array}\right] + \left[\begin{array}{c|c} A_2 & B_2 - XB_1 \\ \hline C_2 & D_2 \end{array}\right] \quad \text{and} \tag{3.20}$$

$$\left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ XA_2 - A_1X & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array}\right] = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 - C_2X & D_1 \end{array}\right] + \left[\begin{array}{c|c} A_2 & B_2 + XB_1 \\ \hline C_2 & D_2 \end{array}\right]. \tag{3.21}$$

Proof.
First, the resolvent is determined by applying equation (2.3):

$$\begin{aligned}
\left(\lambda I - \begin{bmatrix} A_1 & XA_2 - A_1X \\ 0 & A_2 \end{bmatrix}\right)^{-1} &= \begin{bmatrix} \Phi_{A_1} & \Phi_{A_1}(XA_2 - A_1X)\Phi_{A_2} \\ 0 & \Phi_{A_2} \end{bmatrix} \\
&= \begin{bmatrix} \Phi_{A_1} & \Phi_{A_1}(XA_2 - A_1X + \lambda X - \lambda X)\Phi_{A_2} \\ 0 & \Phi_{A_2} \end{bmatrix} \\
&= \begin{bmatrix} \Phi_{A_1} & \Phi_{A_1}(\Phi_{A_1}^{-1}X - X\Phi_{A_2}^{-1})\Phi_{A_2} \\ 0 & \Phi_{A_2} \end{bmatrix} \\
&= \begin{bmatrix} \Phi_{A_1} & X\Phi_{A_2} - \Phi_{A_1}X \\ 0 & \Phi_{A_2} \end{bmatrix} \\
&= \begin{bmatrix} X \\ I \end{bmatrix} \Phi_{A_2} \begin{bmatrix} 0 & I \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} \Phi_{A_1} \begin{bmatrix} I & -X \end{bmatrix}.
\end{aligned}$$

Thus

$$\begin{aligned}
\left[\begin{array}{cc|c} A_1 & XA_2 - A_1X & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array}\right] &= \begin{bmatrix} C_1 & C_2 \end{bmatrix} \left(\begin{bmatrix} X \\ I \end{bmatrix} \Phi_{A_2} \begin{bmatrix} 0 & I \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} \Phi_{A_1} \begin{bmatrix} I & -X \end{bmatrix} \right) \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + (D_1 + D_2) \\
&= C_1\Phi_{A_1}B_1 - C_1\Phi_{A_1}XB_2 + C_1X\Phi_{A_2}B_1 + C_2\Phi_{A_2}B_2 + D_1 + D_2 \\
&= C_1\Phi_{A_1}(B_1 - XB_2) + D_1 + (C_2 + C_1X)\Phi_{A_2}B_2 + D_2 \\
&= \left[\begin{array}{c|c} A_1 & B_1 - XB_2 \\ \hline C_1 & D_1 \end{array}\right] + \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 + C_1X & D_2 \end{array}\right].
\end{aligned}$$

The equality (3.19) follows from the first (3.18) by replacing X with $-X$. The identities (3.20) and (3.21) follow in a similar manner as (3.19) and (3.18). □

The final lemma of this section considers a special case of the sum of two state space representations where the two representations have certain matrices in common. The identity is needed in a later proof.

Lemma 3.5.6.
For appropriately sized matrices, it follows that

$$\left[\begin{array}{ccc|c} A_1 & A_{12} & A_{13} & B_1 \\ 0 & A_{22} & A_{23} & B_2 \\ 0 & 0 & A_3 & B_3 \\ \hline C_1 & C_2 & C_3 & D_1 \end{array}\right] \pm \left[\begin{array}{ccc|c} A_1 & \widetilde{A}_{12} & \widetilde{A}_{13} & \widetilde{B}_1 \\ 0 & \widetilde{A}_{22} & \widetilde{A}_{23} & \widetilde{B}_2 \\ 0 & 0 & A_3 & B_3 \\ \hline C_1 & \widetilde{C}_2 & \widetilde{C}_3 & D_2 \end{array}\right] = \left[\begin{array}{cccc|c} A_1 & A_{12} & \widetilde{A}_{12} & A_{13} \pm \widetilde{A}_{13} & B_1 \pm \widetilde{B}_1 \\ 0 & A_{22} & 0 & A_{23} & B_2 \\ 0 & 0 & \widetilde{A}_{22} & \pm\widetilde{A}_{23} & \pm\widetilde{B}_2 \\ 0 & 0 & 0 & A_3 & B_3 \\ \hline C_1 & C_2 & \widetilde{C}_2 & C_3 \pm \widetilde{C}_3 & D_1 \pm D_2 \end{array}\right]. \tag{3.22}$$

In particular,

$$\left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \pm \left[\begin{array}{cc|c} A_1 & \widetilde{A}_{12} & \widetilde{B}_1 \\ 0 & \widetilde{A}_{22} & \widetilde{B}_2 \\ \hline C_1 & \widetilde{C}_2 & D_2 \end{array}\right] = \left[\begin{array}{cc|c} A_1 & 0 \pm \widetilde{A}_{12} & B_1 \pm \widetilde{B}_1 \\ 0 & \widetilde{A}_{22} & \pm\widetilde{B}_2 \\ \hline C_1 & 0 \pm \widetilde{C}_2 & D_1 \pm D_2 \end{array}\right].$$

Proof.
We prove the sum case. The difference case then follows in a similar manner. Applying (2.5), yields

$$\left[\begin{array}{ccc|c} A_1 & A_{12} & A_{13} & B_1 \\ 0 & A_{22} & A_{23} & B_2 \\ 0 & 0 & A_3 & B_3 \\ \hline C_1 & C_2 & C_3 & D_1 \end{array}\right] + \left[\begin{array}{ccc|c} A_1 & \widetilde{A}_{12} & \widetilde{A}_{13} & \widetilde{B}_1 \\ 0 & \widetilde{A}_{22} & \widetilde{A}_{23} & \widetilde{B}_2 \\ 0 & 0 & A_3 & B_3 \\ \hline C_1 & \widetilde{C}_2 & \widetilde{C}_3 & D_2 \end{array}\right]$$

$$= \begin{bmatrix} C_1 & C_2 & C_3 \end{bmatrix} \begin{bmatrix} \Phi_{A_1} & \Phi_{A_1} A_{12} \Phi_{A_{22}} & \Phi_{A_1} A_{12} \Phi_{A_{22}} A_{23} \Phi_{A_3} + \Phi_{A_1} A_{13} \Phi_{A_3} \\ 0 & \Phi_{A_{22}} & \Phi_{A_{22}} A_{23} \Phi_{A_3} \\ 0 & 0 & \Phi_{A_3} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} + D_1$$

$$+ \begin{bmatrix} C_1 & \widetilde{C}_2 & \widetilde{C}_3 \end{bmatrix} \begin{bmatrix} \Phi_{A_1} & \Phi_{A_1} \widetilde{A}_{12} \Phi_{\widetilde{A}_{22}} & \Phi_{A_1} \widetilde{A}_{12} \Phi_{\widetilde{A}_{22}} \widetilde{A}_{23} \Phi_{A_3} + \Phi_{A_1} \widetilde{A}_{13} \Phi_{A_3} \\ 0 & \Phi_{\widetilde{A}_{22}} & \Phi_{\widetilde{A}_{22}} \widetilde{A}_{23} \Phi_{A_3} \\ 0 & 0 & \Phi_{A_3} \end{bmatrix} \begin{bmatrix} \widetilde{B}_1 \\ \widetilde{B}_2 \\ B_3 \end{bmatrix} + D_2$$

$$= D_1 + C_1 \Phi_{A_1} B_1 + C_1 \Phi_{A_1} A_{12} \Phi_{A_{22}} B_2 + C_2 \Phi_{A_{22}} B_2 + C_1 \Phi_{A_1} A_{12} \Phi_{A_{22}} A_{23} \Phi_{A_3} B_3 + C_1 \Phi_{A_1} A_{13} \Phi_{A_3} B_3$$
$$+ C_2 \Phi_{A_{22}} A_{23} \Phi_{A_3} B_3 + C_3 \Phi_{A_3} B_3 + D_2 + C_1 \Phi_{A_1} \widetilde{B}_1 + C_1 \Phi_{A_1} \widetilde{A}_{12} \Phi_{\widetilde{A}_{22}} \widetilde{B}_2 + \widetilde{C}_2 \Phi_{\widetilde{A}_{22}} \widetilde{B}_2$$
$$+ C_1 \Phi_{A_1} \widetilde{A}_{12} \Phi_{\widetilde{A}_{22}} \widetilde{A}_{23} \Phi_{A_3} B_3 + C_1 \Phi_{A_1} \widetilde{A}_{13} \Phi_{A_3} B_3 + \widetilde{C}_2 \Phi_{\widetilde{A}_{22}} \widetilde{A}_{23} \Phi_{A_3} B_3 + \widetilde{C}_3 \Phi_{A_3} B_3$$
$$= (D_1 + D_2) + C_1 \Phi_{A_1} (A_{13} + \widetilde{A}_{13}) \Phi_{A_3} B_3 + (C_3 + \widetilde{C}_3) \Phi_{A_3} B_3 + C_1 \Phi_{A_1} (B_1 + \widetilde{B}_1) + C_1 \Phi_{A_1} A_{12} \Phi_{A_{22}} B_2$$
$$+ C_2 \Phi_{A_{22}} B_2 + C_1 \Phi_{A_1} A_{12} \Phi_{A_{22}} A_{23} \Phi_{A_3} B_3 + C_2 \Phi_{A_{22}} A_{23} \Phi_{A_3} B_3 + C_1 \Phi_{A_1} \widetilde{A}_{12} \Phi_{\widetilde{A}_{22}} \widetilde{B}_2 + \widetilde{C}_2 \Phi_{\widetilde{A}_{22}} \widetilde{B}_2$$
$$+ C_1 \Phi_{A_1} \widetilde{A}_{12} \Phi_{\widetilde{A}_{22}} \widetilde{A}_{23} \Phi_{A_3} B_3 + \widetilde{C}_2 \Phi_{\widetilde{A}_{22}} \widetilde{A}_{23} \Phi_{A_3} B_3$$
$$= (D_1 + D_2) + \begin{bmatrix} C_1 & C_2 & \widetilde{C}_2 & (C_3 + \widetilde{C}_3) \end{bmatrix} \times$$
$$\begin{bmatrix} \Phi_{A_1} & \Phi_{A_1} A_{12} \Phi_{A_{22}} & \Phi_{A_1} \widetilde{A}_{12} \Phi_{\widetilde{A}_{22}} & \Phi_{A_1} (A_{12} \Phi_{A_{22}} A_{23} + \widetilde{A}_{12} \Phi_{\widetilde{A}_{22}} \widetilde{A}_{23} + (A_{13} + \widetilde{A}_{13})) \Phi_{A_3} \\ 0 & \Phi_{A_{22}} & 0 & \Phi_{A_{22}} A_{23} \Phi_{A_3} \\ 0 & 0 & \Phi_{\widetilde{A}_{22}} & \Phi_{\widetilde{A}_{22}} \widetilde{A}_{23} \Phi_{A_3} \\ 0 & 0 & 0 & \Phi_{A_3} \end{bmatrix}$$
$$\times \begin{bmatrix} (B_1 + \widetilde{B}_1) \\ B_2 \\ \widetilde{B}_2 \\ B_3 \end{bmatrix}$$
$$= \left[\begin{array}{cccc|c} A_1 & A_{12} & \widetilde{A}_{12} & A_{13} + \widetilde{A}_{13} & B_1 + \widetilde{B}_1 \\ 0 & A_{22} & 0 & A_{23} & B_2 \\ 0 & 0 & \widetilde{A}_{22} & \widetilde{A}_{23} & \widetilde{B}_2 \\ 0 & 0 & 0 & A_3 & B_3 \\ \hline C_1 & C_2 & \widetilde{C}_2 & C_3 + \widetilde{C}_3 & D_1 + D_2 \end{array}\right]$$

where the last equality follows by equation (2.7). In order to prove the difference case, it is noted that

$$-\left[\begin{array}{ccc|c} A_1 & \widetilde{A}_{12} & \widetilde{A}_{13} & \widetilde{B}_1 \\ 0 & \widetilde{A}_{22} & \widetilde{A}_{23} & \widetilde{B}_2 \\ 0 & 0 & A_3 & B_3 \\ \hline C_1 & \widetilde{C}_2 & \widetilde{C}_3 & D_2 \end{array}\right] = \left[\begin{array}{ccc|c} A_1 & \widetilde{A}_{12} & \widetilde{A}_{13} & -\widetilde{B}_1 \\ 0 & \widetilde{A}_{22} & \widetilde{A}_{23} & -\widetilde{B}_2 \\ 0 & 0 & A_3 & -B_3 \\ \hline C_1 & \widetilde{C}_2 & \widetilde{C}_3 & -D_2 \end{array}\right] = \left[\begin{array}{ccc|c} A_1 & \widetilde{A}_{12} & -\widetilde{A}_{13} & -\widetilde{B}_1 \\ 0 & \widetilde{A}_{22} & -\widetilde{A}_{23} & -\widetilde{B}_2 \\ 0 & 0 & A_3 & B_3 \\ \hline C_1 & \widetilde{C}_2 & -\widetilde{C}_3 & -D_2 \end{array}\right].$$

The difference case now follows by applying the above identity to the sum case. This completes the proof. □

3.6 Realizations and Hardy Spaces

According to Theorem 3.2.1, proper real rational matrix functions can be identified with transfer functions of causal LTI systems. Let $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{r\times n}$ and $D \in \mathbb{R}^{r\times m}$. Then

$$\widehat{G} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty^{r\times m} \qquad \text{and} \qquad \widehat{H} = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \in \mathcal{RH}_2^{r\times m}$$

if A is stable.

Proposition 3.6.1 (cf. p.2313 in [46]).
If $\widehat{G} \in \mathcal{H}_2$, then $\widehat{G}^* \in \mathcal{H}_2^\perp$. Furthermore

1. *if $\widehat{G} \in \mathcal{H}_\infty$, then $\widehat{G}\mathcal{H}_2 \subseteq \mathcal{H}_2$;*
2. *if $\widehat{G} \in \mathcal{H}_\infty^-$, then $\widehat{G}\mathcal{H}_2^\perp \subseteq \mathcal{H}_2^\perp$.*

The following lemma gives a way to compute the $\mathcal{H}_2$-norm of a transfer function $\widehat{G} \in \mathcal{RH}_2$ in terms of the controllability gramian X_c of (A, B) or the observability gramian Y_o of (C, A).

Lemma 3.6.2 (Lemma 4.4 in [56]).
If

$$\widehat{G} = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \in \mathcal{RH}_2,$$

then $\|\widehat{G}\|_2^2 = \text{trace}\,(B^\intercal Y_o B) = \text{trace}\,(C X_c C^\intercal)$, where

$$X_c = \int_0^\infty e^{At} B B^\intercal e^{A^\intercal t}\ \mathrm{d}t \quad \textit{and} \quad Y_o = \int_0^\infty e^{A^\intercal t} C^\intercal C e^{At}\ \mathrm{d}t \tag{3.23}$$

are the controllability and observability gramians of (A, B) and (C, A) respectively.

The following theorem gives a way to compute X_c and Y_o, without computing the integrals in (3.23), by solving Lyapunov equations. These equations can be solved numerically by standard techniques.

Theorem 3.6.3 (Theorem 4.1 [8]).
If A is a stable matrix, then X_c and Y_o as in (3.23) are the unique solutions of the Lyapunov equations

$$AX + XA^\intercal + BB^\intercal = 0 \quad \textit{and} \quad A^\intercal Y + YA + C^\intercal C = 0 \tag{3.24}$$

respectively.

Next we consider a class of functions that preserve the $\mathcal{H}_2$-norm under multiplication. These so-called inner functions play an important role in the $\mathcal{H}_2$-control problem.

Definition 3.6.4 (Inner function, co-inner function).
A proper real rational matrix function $\widehat{G}$ is said to be *inner* if $\widehat{G} \in \mathcal{RH}_\infty$ and $\widehat{G}^*\widehat{G} = I$ and *co-inner* if $\widehat{G} \in \mathcal{RH}_\infty$ and $\widehat{G}\widehat{G}^* = I$.

Since $\widehat{G}^\intercal(\widehat{G}^\intercal)^* = \widehat{G}^\intercal(\widehat{G}^*)^\intercal = (\widehat{G}^*\widehat{G})^\intercal$ it follows that $\widehat{G}$ is inner if and only if $\widehat{G}^\intercal$ is co-inner. Hence they are dual notions and the properties of co-inner functions follow from the properties of inner functions by duality. Furthermore, if $\widehat{G}$ is inner and $\widehat{x} \in \mathcal{L}_2$, then it follows from (2.9) and (2.10), that

$$\|\widehat{x}^\intercal\widehat{G}^\intercal\|_2 = \|\widehat{G}\widehat{x}\|_2 = \|\widehat{x}\|_2 = \|\widehat{x}^\intercal\|_2,$$

because $\widehat{G}^*\widehat{G} = I = \widehat{G}^\intercal(\widehat{G}^\intercal)^*$. It is due to this norm-preserving property that inner and co-inner functions will play an important role in $\mathcal{H}_2$-optimal control in a subsequent section.

3.7 Algebraic Riccati Equations

This section is chiefly based on Section 6.2 of [8] and Section 12.1 of [56]. Algebraic Riccati equations (AREs) play a central role in $\mathcal{H}_2$-control. Associated with AREs are Hamiltonian matrices, which we will consider first.

Definition 3.7.1 (Hamiltonian matrix)**.**
A matrix $H \in \mathbb{R}^{2n\times 2n}$ is a called a *Hamiltonian matrix* if

$$JH = (JH)^\intercal = H^\intercal J^\intercal \quad \text{for} \quad J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}.$$

The matrix J is an orthogonal skew symmetric matrix ($J^\intercal = J^{-1} = -J$), so it can be seen that if H is Hamiltonian then $JHJ^{-1} = -H^\intercal$. That is, H and $-H^\intercal$ are similar. But since any matrix is similar to its transpose, it follows that H and $-H$ are also similar. This shows that if $\lambda \in \sigma(H)$, then $-\overline{\lambda} \in \sigma(H)$. Thus if $\lambda \in \mathbb{C}$ is an eigenvalue of H, then $-\lambda$ is an eigenvalue of H of the same multiplicity. It is possible to show that every Hamiltonian matrix $H \in \mathbb{R}^{2n\times 2n}$ has the form

$$H = \begin{bmatrix} A & R \\ -Q & -A^\intercal \end{bmatrix} \tag{3.25}$$

where $A, R, Q \in \mathbb{R}^{n\times n}$ with $R = R^\intercal$ and $Q = Q^\intercal$ and every matrix of the above form is a Hamiltonian matrix. Recall that a matrix A is said to be Hurwitz if $\mathrm{Re}(\lambda) < 0$ for each eigenvalue λ of A. Given a Hamiltonian matrix (3.25), we say H *is in the domain of the Riccati operator* and we write $H \in \mathrm{dom}(\mathrm{Ric})$ if there exists a Hurwitz matrix H_- and there exists a matrix $X \in \mathbb{R}^{n\times n}$ such that

$$\begin{bmatrix} I \\ X \end{bmatrix} H_- = H \begin{bmatrix} I \\ X \end{bmatrix}. \tag{3.26}$$

In this case the matrix X turns out to be uniquely determined, so we can define a function

$$\mathrm{Ric} : \mathrm{dom}(\mathrm{Ric}) \subset \mathbb{R}^{2n\times 2n} \to \mathbb{R}^{n\times n} \quad \text{by} \qquad \mathrm{Ric}(H) = X,$$

where H and X are as in (3.25) and (3.26) respectively.

The following theorem connects the Hamiltonian matrix (3.25) with the associated continuous time algebraic Riccati equation (CARE)

$$A^\intercal X + XA + Q - XRX = 0. \tag{3.27}$$

Theorem 3.7.2 (cf. Theorem 6.3 in [8] or Theorem 12.1 in [56])**.**
Given a Hamiltonian matrix H as in (3.25)*, if $H \in \mathrm{dom}(\mathrm{Ric})$ and $X = \mathrm{Ric}(H)$, then*

1. *$X = X^\intercal$;*
2. *X solves the algebraic Riccati equation* (3.27)*;*
3. *$-X$ stabilizes the pair (A, R) (that is, the matrix $A - RX$ is Hurwitz).*

The previous theorem guarantees the existence of a symmetric stabilizing solution to the CARE (3.27) associated with the Hamiltonian matrix (3.25) in the case that the Hamiltonian matrix is in the domain of the Riccati operator. The following theorem gives necessary and sufficient conditions under which this occurs.

Theorem 3.7.3 (cf. Theorem 12.2 in [56])**.**
Let H be a Hamiltonian matrix as in (3.25) *that has no purely imaginary eigenvalues. If $R \geq 0$ or $R \leq 0$, then*

$$H \in \mathrm{dom}(\mathrm{Ric}) \qquad \Longleftrightarrow \qquad (A, R) \textit{ is stabilizable.}$$

In the next theorem, we consider the case where the sub-matrices Q and R in the Hamiltonian matrix (3.25) are positive semi-definite and are given by $Q = C^\intercal C$ and $R = BB^\intercal$ for some matrices B and C. The theorem gives necessary and sufficient conditions for such a Hamiltonian matrix to be in the domain of the Riccati operator. Furthermore, it shows that, in this case, $X = \mathrm{Ric}(H)$ is positive semi-definite.

Theorem 3.7.4 (cf. Theorem 12.4 in [56]).
If H is a Hamiltonian matrix given by

$$H = \begin{bmatrix} A & BB^\intercal \\ -C^\intercal C & -A^\intercal \end{bmatrix},$$

then $H \in \mathrm{dom(Ric)}$ if and only if (A, B) is stabilizable and (C, A) has no unobservable modes on the imaginary axis. Furthermore, if $H \in \mathrm{dom(Ric)}$, then $X = \mathrm{Ric}(H) \geq 0$.

Corollary 3.7.5 (cf. Corollary 6.6 [8] or Corollary 12.5 in [56]).
If the triple (A, B, C) is stabilizable and detectable, then there exists a unique positive semi-definite symmetric solution X of the CARE

$$A^\intercal X + XA + C^\intercal C - XBB^\intercal X = 0 \tag{3.28}$$

such that $A - BB^\intercal X$ is stable.

Given a realization (A, B, C, D), we will later be concerned with the CARE

$$A^\intercal X + XA + Q - (XB + S^\intercal)R^{-1}(B^\intercal X + S) = 0, \tag{3.29}$$

where $Q = C^\intercal C$, $R = D^\intercal D$ and $S = D^\intercal C$. If the CARE (3.29) has a unique positive semi-definite symmetric stabilizing solution X, then we write $X = \mathrm{Ric}(A, B, C, D)$. The inverse of R exists if and only if $D^\intercal D > 0$. Realizations for which this holds will be called *regular* .

Definition 3.7.6 (Regular, in standard form).

1. A realization (A, B, C, D) is said to be *regular* if $D^\intercal D > 0$.
2. A realization (A, B, C, D) is said to be *in standard form* if $D^\intercal C = 0$ and $D^\intercal D = I$.

Note that if (A, B, C, D) is in standard form, then the Riccati equations (3.29) and (3.28) are identical. The next theorem shows that we can always transform a regular realization into a realization in standard form by a feedback transformation.

Theorem 3.7.7 (cf. page 213 in [49]).
Any system $\Sigma \sim (A, B, C, D)$ can be transformed into a system $\widetilde{\Sigma} \sim (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ in standard form, where

$$\widetilde{A} = A - BR^{-1}S, \quad \widetilde{B} = -BR^{-\frac{1}{2}}, \quad \widetilde{C} = C - DR^{-1}S \quad \text{and} \quad \widetilde{D} = -DR^{-\frac{1}{2}} \tag{3.30}$$

by means of the feedback transformation

$$u(t) = -R^{-1}Sx(t) - R^{-\frac{1}{2}}v(t),$$

where $R = D^\intercal D$, $S = D^\intercal C$ and v is the input variable of $\widetilde{\Sigma}$. Furthermore

1. *$(\widetilde{A}, \widetilde{B})$ is stabilizable if and only if (A, B) is stabilizable;*
2. *$(\widetilde{C}, \widetilde{A})$ is detectable if and only if $\begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix}$ has full column rank for all $\lambda \in \overline{\mathbb{C}}^+$.*
3. *X solves the CARE (3.28) in $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ if and only if X solves the CARE (3.29) in (A, B, C, D).*

Proof.
Consider a regular LTI system $\Sigma \sim (A, B, C, D)$:

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t)\\ y(t) &= Cx(t) + Du(t).\end{aligned}$$

Then by definition $R = D^{\intercal}D > 0$. Hence the feedback transformation $u(t) = -R^{-1}Sx(t) - R^{-\frac{1}{2}}v(t)$ is well defined. Applying this feedback to Σ transforms it to the the system $\widetilde{\Sigma}$, given by:

$$\begin{aligned}\dot{x}(t) &= Ax(t) + B(-R^{-1}Sx(t) - R^{-\frac{1}{2}}v(t)) = \widetilde{A}x(t) + \widetilde{B}v(t)\\ y(t) &= Cx(t) + D(-R^{-1}Sx(t) - R^{-\frac{1}{2}}v(t)) = \widetilde{C}x(t) + \widetilde{D}v(t),\end{aligned}$$

with $\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}$ as in (3.30). Now

$$\begin{aligned}\widetilde{D}^{\intercal}\widetilde{C} &= -R^{-\frac{1}{2}}D^{\intercal}(C - DR^{-1}S) = -R^{-\frac{1}{2}}S + R^{-\frac{1}{2}}S = 0\\ \text{and} \qquad \tilde{D}^{\intercal}\tilde{D} &= -R^{-\frac{1}{2}}D^{\intercal}(-DR^{-\frac{1}{2}}) \qquad = R^{-\frac{1}{2}}RR^{-\frac{1}{2}} \qquad = I,\end{aligned}$$

which shows that $\widetilde{\Sigma} \sim (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ is in standard form.

1. For $F \in \mathbb{R}^{m\times n}$, set $G = -R^{-1}S - R^{-\frac{1}{2}}F \in \mathbb{R}^{m\times n}$. Then $F = -R^{-\frac{1}{2}}S - R^{\frac{1}{2}}G$ and hence

   $$\widetilde{A} + \widetilde{B}F = A - BR^{-1}S - BR^{-\frac{1}{2}}F = A + B(-R^{-1}S - R^{-\frac{1}{2}}F) = A + BG.$$

   Which shows that $(\widetilde{A}, \widetilde{B})$ is stabilizable if and only if (A, B) is stabilizable.

2. By the PBH test $(\widetilde{C}, \widetilde{A})$ is detectable if and only if

   $$\begin{bmatrix}\widetilde{A} - \lambda I\\ \widetilde{C}\end{bmatrix} = \begin{bmatrix}(A - \lambda I) - BR^{-1}S\\ C - DR^{-1}\end{bmatrix} = \begin{bmatrix}A - \lambda I\\ C\end{bmatrix} - \begin{bmatrix}B\\ D\end{bmatrix} R^{-1}S \tag{3.31}$$

   has full column rank (or equivalently has zero nullity) for all $\lambda \in \overline{\mathbb{C}}^{+}$. It can be seen that the above is the Schur complement (see equation (2.2)) of the matrix

   $$\begin{bmatrix}A - \lambda I & B\\ C & D\\ S & R\end{bmatrix} = \begin{bmatrix}A - \lambda I & B\\ C & D\\ D^{\intercal}C & D^{\intercal}D\end{bmatrix}. \tag{3.32}$$

   Form the factorizations in Lemma 2.2.3, it follows that the Schur complement (3.31) has zero kernel if and only if its associated matrix (3.32) has zero kernel. However, since the second and third rows of the matrix (3.32) are dependent, it follows that

   $$\ker\begin{bmatrix}A - \lambda I & B\\ C & D\\ D^{\intercal}C & D^{\intercal}D\end{bmatrix} = \ker\begin{bmatrix}A - \lambda I & B\\ C & D\end{bmatrix} = \{0\}$$

   if and only if $(\widetilde{C}, \widetilde{A})$ is detectable.

3. Substituting in the values of $\widetilde{A}, \widetilde{B}, \widetilde{C}$ in (3.30) into the left hand side of equation (3.28) gives

   $$\begin{aligned}&\widetilde{A}^{\intercal}X + X\widetilde{A} + \widetilde{C}^{\intercal}\widetilde{C} - X\widetilde{B}\widetilde{B}^{\intercal}X\\ &= (A^{\intercal} - S^{\intercal}R^{-1}B^{\intercal})X + X(A - BR^{-1}S) + (C^{\intercal} - S^{\intercal}R^{-1}D^{\intercal})(C - DR^{-1}S)\\ &\qquad - X(-BR^{-\frac{1}{2}})(-R^{-\frac{1}{2}}B^{\intercal})X\\ &= A^{\intercal}X + XA + C^{\intercal}C - [XBR^{-1}S + S^{\intercal}R^{-1}B^{\intercal}X + S^{\intercal}R^{-1}S + S^{\intercal}R^{-1}S\\ &\qquad - S^{\intercal}R^{-1}RR^{-1}S + XBR^{-1}B^{\intercal}X]\\ &= A^{\intercal}X + XA + C^{\intercal}C - (XB + S^{\intercal})R^{-1}(B^{\intercal}X + S),\end{aligned}$$

   which is the left hand side equation (3.29). □

Theorem 3.7.7 and Corollary 3.7.5 show that if the triple $(\widetilde{A}, \widetilde{B}, \widetilde{C})$ is stabilizable and detectable then there exists a solution $X = \mathrm{Ric}(\widetilde{A}, \widetilde{B}, \widetilde{C}, D)$ and in that case it is also the solution $X = \mathrm{Ric}(A, B, C, D)$. The converse is not true in general, see for example Remark 12.3 on page 244 of [56]. We also note that if $\begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix}$ has full column rank for all $\lambda \in \overline{\mathbb{C}}^+$, then in particular $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}$ also has full column rank. So in that case, (C, A) is also detectable by the PBH test.

Theorem 3.7.8 (cf. Corollary 12.7 in [56])**.**
If $D^\intercal D > 0$, then there exists a $X = \mathrm{Ric}(A, B, C, D)$ if and only if

1. *(A, B) is stabilizable*
2. *$\begin{bmatrix} A - i\omega I & B \\ C & D \end{bmatrix}$ has full column rank for all $\omega \in \mathbb{R}$.*

We note that if E is a matrix of appropriate size and we substitute $A + BE$ for A and $C + DE$ for C in the left hand side of the Riccati equation (3.29), this yields

$$\begin{aligned} &(A + BE)^\intercal X + X(A + BE) + (C + DE)^\intercal (C + DE) \\ &\quad - (XB + (C + DE)^\intercal D)(D^\intercal D)^{-1}(B^\intercal X + D^\intercal (C + DE)) \\ &= A^\intercal X + \cancel{E^\intercal B^\intercal X} + XA + \cancel{XBE} + C^\intercal C + \cancel{C^\intercal DE} + \cancel{E^\intercal D^\intercal C} + \cancel{E^\intercal D^\intercal DE} \\ &\quad - (XB(D^\intercal D)^{-1} + C^\intercal D(D^\intercal D)^{-1} + \cancel{E^\intercal})(B^\intercal X + D^\intercal C + \cancel{D^\intercal DE}) \\ &= A^\intercal X + XA + C^\intercal C - (XB + C^\intercal D)(D^\intercal D)^{-1}(B^\intercal X + D^\intercal C), \end{aligned} \tag{3.33}$$

which is the left hand side of equation (3.29). The above equality enables us to prove the following two lemmas.

Lemma 3.7.9.
Suppose $D^\intercal D > 0$ and $X = \mathrm{Ric}(A, B, C, D)$. Let $F = -(D^\intercal D)^{-1}(B^\intercal X + D^\intercal C)$ and set $A_F = A + BF$ and $C_F = C + DF$. Then

1. *A_F is stable;*
2. *$C_F^\intercal C_F = C^\intercal C - C^\intercal D(D^\intercal D)^{-1}D^\intercal C + XB(D^\intercal D)^{-1}B^\intercal X$;*
3. *$B^\intercal X + D^\intercal C_F = 0$ and*
4. *$A_F^\intercal X + XA_F + C_F^\intercal C_F = A^\intercal X + XA + C^\intercal C - (XB + C^\intercal D)(D^\intercal D)^{-1}(B^\intercal X + D^\intercal C)$.*

Proof.

1. By Theorem 3.7.7, it follows that X is a unique positive semi-definite symmetric solution of the CARE

$$\widetilde{A}^\intercal X + X\widetilde{A} + \widetilde{C}^\intercal \widetilde{C} - X\widetilde{B}\widetilde{B}^\intercal X = 0$$

where

$$\widetilde{A} = A - B(D^\intercal D)^{-1}D^\intercal C, \quad \widetilde{B} = -B(D^\intercal D)^{-\frac{1}{2}} \quad \text{and} \quad \widetilde{C} = C - D(D^\intercal D)^{-1}D^\intercal C.$$

By Corollary 3.7.5, $\widetilde{A} - \widetilde{B}\widetilde{B}^\intercal X$ is stable. But

$$\begin{aligned} \widetilde{A} - \widetilde{B}\widetilde{B}^\intercal X &= (A - B(D^\intercal D)^{-1}D^\intercal C) - (-B(D^\intercal D)^{-\frac{1}{2}})(-(D^\intercal D)^{-\frac{1}{2}}B^\intercal)X \\ &= A + B[-(D^\intercal D)^{-1}(D^\intercal C + B^\intercal X)] \\ &= A + BF \end{aligned}$$

and hence A_F is stable.

2. The claim follows by computing $(C+DF)^\intercal(C+DF)$:

$$\begin{aligned}(C+DF)^\intercal(C+DF) &= (C^\intercal-(XB+C^\intercal D)(D^\intercal D)^{-1}D^\intercal)(C-D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C))\\ &= C^\intercal C - C^\intercal D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)-(XB+C^\intercal D)(D^\intercal D)^{-1}D^\intercal C\\ &\quad +(XB+C^\intercal D)(D^\intercal D)^{-1}D^\intercal D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)\\ &= C^\intercal C - C^\intercal D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)-(XB+C^\intercal D)(D^\intercal D)^{-1}D^\intercal C\\ &\quad +(XB+C^\intercal D)(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)\\ &= C^\intercal C+(XB+C^\intercal D)[(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)-(D^\intercal D)^{-1}D^\intercal C]\\ &\quad - C^\intercal D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)\\ &= C^\intercal C+(XB+C^\intercal D)[(D^\intercal D)^{-1}B^\intercal X]-C^\intercal D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)\\ &= C^\intercal C+XB(D^\intercal D)^{-1}B^\intercal X-C^\intercal D(D^\intercal D)^{-1}D^\intercal C.\end{aligned}$$

3. With $C_F = C+DF$ and $F=-(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)$, it follows that

$$\begin{aligned}B^\intercal X+D^\intercal C_F &= B^\intercal X+D^\intercal(C-D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C))\\ &= B^\intercal X+D^\intercal C-(B^\intercal X+D^\intercal C)=0.\end{aligned}$$

4. By substituting $F=-(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)$ for E in equation (3.33), it can be seen that

$$\begin{aligned}&A_F^\intercal X+XA_F+C_F^\intercal C_F\\ &= A_F^\intercal X+XA_F+C_F^\intercal C_F-(XB+(C-D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C))^\intercal D)(D^\intercal D)^{-1}(0)\\ &= A_F^\intercal X+XA_F+C_F^\intercal C_F-(XB+(C-D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C))^\intercal D)(D^\intercal D)^{-1}\\ &\quad \times((B^\intercal X+D^\intercal(C-D(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)))\\ &= (A+BF)^\intercal X+X(A+BF)+(C+DF)^\intercal(C+DF)\\ &\quad -(XB+(C+DF)^\intercal D)(D^\intercal D)^{-1}(B^\intercal X+D^\intercal(C+DF))\\ &= A^\intercal X+XA+C^\intercal C-(XB+C^\intercal D)(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C).\end{aligned}$$

This completes the proof. □

In the above theorem, we saw that

$$A_F^\intercal X+XA_F+C_F^\intercal C_F = A^\intercal X+XA+C^\intercal C-(XB+C^\intercal D)(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)$$

and we note that the left-hand side has the form of a Lyapunov equation as in (3.24). Theorem 3.6.3 gives solutions of such Lyapunov equations. Lastly equation (3.33) showed that replacing A with $A+BE$ and C with $C+DE$ in the Riccati equation (3.29), results in the exact same equation. Taken together, this gives the following result.

Lemma 3.7.10.
Suppose $D^\intercal D>0$. There exists a unique solution $X=\mathrm{Ric}(A,B,C,D)$ if and only if there exists a unique solution $X_E=\mathrm{Ric}(A+BE,B,C+DE,D)$ for all appropriately sized matrices E. Furthermore, if such solutions exist, then $X=X_E=X_o$, where

$$X_o=\int_0^\infty e^{(A+BF)^\intercal t}(C+DF)^\intercal(C+DF)e^{(A+BF)t}\,\mathrm{d}t,\qquad F=-(D^\intercal D)^{-1}(B^\intercal X+D^\intercal C)$$

is the observability Gramian of the pair $(A+BF,C+DF)$.

Proof.
Suppose $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. Let $E \in \mathbb{R}^{m\times n}$ be arbitrary. By assumption $D^\intercal D > 0$.

Secondly, if (A, B) is stabilizable, then there exists a matrix K such that $A+BK = (A+BE)+B(K-E)$ is stable and thus $(A + BE, B)$ is stabilizable. Conversely, if $(A + BE, B)$ is stabilizable, then there exists a K such that $(A + BE) + BK = A + B(E + K)$ is stable and thus (A, B) is stabilizable.

Thirdly, since

$$\begin{bmatrix} I & 0 \\ E & I \end{bmatrix}$$

is invertible, the matrix

$$\begin{bmatrix} A - i\omega I & B \\ C & D \end{bmatrix} \qquad \text{has full column rank for all } \omega \in \mathbb{R}$$

if and only if

$$\begin{bmatrix} A - i\omega I & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ E & I \end{bmatrix} = \begin{bmatrix} (A + BE) - i\omega I & B \\ (C + DE) & D \end{bmatrix} \qquad \text{has full column rank for all } \omega \in \mathbb{R}.$$

Thus, by Theorem 3.7.8, there exists a unique solution $X = \mathrm{Ric}(A, B, C, D)$ if and only if there exists a unique solution $X_E = \mathrm{Ric}(A + BE, B, C + DE, D)$.
By Lemma 3.7.9 (4) and equation (3.33), it follows that

$$\begin{aligned} A_F^\intercal X + XA_F + C_F^\intercal C_F &= A^\intercal X + XA + C^\intercal C - (XB + C^\intercal D)(D^\intercal D)^{-1}(B^\intercal X + D^\intercal C) \\ &= (A + BE)^\intercal X + X(A + BE) + (C + DE)^\intercal (C + DE) \\ &\quad - (XB + (C + DE)^\intercal D)(D^\intercal D)^{-1}(B^\intercal X + D^\intercal (C + DE)), \end{aligned} \tag{3.34}$$

where $A_F = A + BF$ and $C_F = C + DF$. By Theorem 3.6.3, the solution to the Lyaponuv equation $A_F^\intercal X + XA_F + C_F^\intercal C_F = 0$ is given by X_o, the observability gramian of the pair (C_F, A_F). But since $X = \mathrm{Ric}(A, B, C, D)$ and $X = \mathrm{Ric}(A + BE, B, C + DE, D)$, it follows from equation (3.34) that these solutions are all equal to each other, that is, $X = X_E = X_o$. □

Recall that a function $\widehat{U} \in \mathcal{RH}_\infty$ is said to be inner if $\widehat{U}^*\widehat{U} = I$. As in Section 17.6 of [1], we now also define invertible outer functions and inner-outer factorizations.

Definition 3.7.11 (Invertible outer function)**.**
A square proper real rational matrix function $\widehat{L}$ is said to be *invertible outer* if $\widehat{L} \in \mathcal{RH}_\infty$ is invertible and $\widehat{L}^{-1} \in \mathcal{RH}_\infty$.

Definition 3.7.12 (Inner-outer factorization)**.**
Given some proper real rational matrix function $\widehat{G}$, a factorization

$$\widehat{G} = \widehat{U}\widehat{L}$$

is said to be an *inner-outer* factorization of $\widehat{G}$ if $\widehat{U}$ is an inner-function and $\widehat{L}$ is an invertible outer function.

The following theorem gives an inner-outer factorization of stable proper rational matrix functions. In this factorization, both the inner and invertible outer factors contain the stabilizing matrix F which depends on the solution of the Riccati equation related to the realization of the rational matrix function. As the formulation differs slightly from the one given in [1] and the result is given as an exercise in [56], we provide a proof for the sake of completeness.

Theorem 3.7.13 (cf. Theorem 17.26 in [1] and p.248 in [56])**.**
Suppose $\widehat{G} \in \mathcal{RH}_\infty$ has the following realization

$$\widehat{G} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

If

1. *A is stable;*
2. *$R = D^\intercal D > 0$;*
3. *$\begin{bmatrix} A - i\omega I & B \\ C & D \end{bmatrix}$ has full column rank for all $\omega \in \mathbb{R}$,*

then $\widehat{G}$ admits an inner-outer factorization $\widehat{G} = \widehat{U}\widehat{L}$, where

$$\widehat{U} = \left[\begin{array}{c|c} A + BF & BR^{-\frac{1}{2}} \\ \hline C + DF & DR^{-\frac{1}{2}} \end{array}\right] \quad \textit{and} \quad \widehat{L} = \left[\begin{array}{c|c} A & B \\ \hline -R^{\frac{1}{2}}F & R^{\frac{1}{2}} \end{array}\right]$$

and where $F = -R^{-1}(B^\intercal X + D^\intercal C)$ and $X = \mathrm{Ric}(A, B, C, D)$.

Proof.
Since the matrix A is stable, the pair (A, B) is stabilizable. Together with the other two conditions in the Theorem, (A, B, C, D) satisfies the premises of Theorem 3.7.8. Thus there exists a unique stabilizing $X = \mathrm{Ric}(A, B, C, D)$. Since both $A+BF$ and A are stable, $\widehat{U} \in \mathcal{RH}_\infty$ and $\widehat{L} \in \mathcal{RH}_\infty$. Clearly $\widehat{L}$ is invertible (because $R^{\frac{1}{2}}$ is invertible). Let $A_F = A + BF$ and $C_F = C + DF$. Then by equation (3.17),

$$\widehat{U}^*\widehat{U} = \left[\begin{array}{cc|c} -A_F^\intercal & -C_F^\intercal C_F & -C_F^\intercal DR^{-\frac{1}{2}} \\ 0 & A_F & BR^{-\frac{1}{2}} \\ \hline R^{-\frac{1}{2}}B^\intercal & R^{-\frac{1}{2}}D^\intercal C_F & R^{-\frac{1}{2}}(D^\intercal D)R^{-\frac{1}{2}} \end{array}\right].$$

Now by Lemma 3.7.9 (4), $A_F^\intercal X + XA_F + C_F^\intercal C_F = 0$, because $X = \mathrm{Ric}(A, B, C, D)$. Hence, $-C_F^\intercal C_F = A_F^\intercal X + XA_F$. So with $X = \mathrm{Ric}(A, B, C, D)$, $A_1 = -A_F^\intercal$ and $A_2 = A_F$ in (3.19) Lemma 3.5.5, it follows that

$$\begin{aligned}
\widehat{U}^*\widehat{U} &= \left[\begin{array}{c|c} -A_F^\intercal & -C_F^\intercal DR^{-\frac{1}{2}} - XBR^{-\frac{1}{2}} \\ \hline R^{-\frac{1}{2}}B^\intercal & 0 \end{array}\right] + \left[\begin{array}{c|c} A_F & BR^{-\frac{1}{2}} \\ \hline R^{-\frac{1}{2}}D^\intercal C_F + R^{-\frac{1}{2}}B^\intercal X & I \end{array}\right] \\
&= \left[\begin{array}{c|c} -A_F^\intercal & -(D^\intercal C_F + B^\intercal X)^\intercal R^{-\frac{1}{2}} \\ \hline R^{-\frac{1}{2}}B^\intercal & 0 \end{array}\right] + \left[\begin{array}{c|c} A_F & BR^{-\frac{1}{2}} \\ \hline R^{-\frac{1}{2}}(D^\intercal C_F + B^\intercal X) & I \end{array}\right] \\
&= \left[\begin{array}{c|c} -A_F^\intercal & 0 \\ \hline R^{-\frac{1}{2}}B^\intercal & 0 \end{array}\right] + \left[\begin{array}{c|c} A_F & BR^{-\frac{1}{2}} \\ \hline 0 & I \end{array}\right] = I,
\end{aligned}$$

where $D^\intercal C_F + B^\intercal X = 0$ by Lemma 3.7.9 (3). This shows that $\widehat{U}$ is inner. Lastly, we confirm that $\widehat{G}$ can be factorized as $\widehat{U}\widehat{L}$. By equation (3.13) and Lemma 3.5.5, it follows that

$$\begin{aligned}
\widehat{U}\widehat{L} &= \left[\begin{array}{cc|c} A + BF & -BF & B \\ 0 & A & B \\ \hline C + DF & -DF & D \end{array}\right] \\
&= \left[\begin{array}{c|c} A + BF & B - B \\ \hline C + DF & 0 \end{array}\right] + \left[\begin{array}{c|c} A & B \\ \hline C + DF - DF & D \end{array}\right] \\
&= \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].
\end{aligned}$$

This completes the proof. $\square$

Chapter 4

$\mathcal{H}_2$-Optimal Control

In this chapter, we consider the problem of synthesizing a controller which minimizes the $\mathcal{H}_2$-norm of the closed loop transfer function obtained after connecting the controller to a plant. This review is mainly based on chapters 5 and 6 of [8] and chapters 5, 11 and 13 of [56]. We first consider the connection of a controller to a plant, feasibility conditions for such a connection and state space formulas for the closed loop system. It is required that the controller stabilizes the plant. The famous Youla parametrization gives the general form of all stabilizing controllers and plays a prominent role in the solution of the $\mathcal{H}_2$ control problem. The general $\mathcal{H}_2$-optimal control problem and its solution is presented. We consider the special case of state feedback control, which is closely related to the LQ-control problem. An alternative approach via spectral factorization is considered for the state feedback case, because such an approach is required for the control of structured linear systems in Chapter 7. We then consider an alternative solution strategy to the $\mathcal{H}_2$-control problem for the output feedback case which does not require the solutions of Riccati equations in the Youla parameter. Lastly we consider the reparameterized control problem for the output feedback case and compare the results to those of the preceding section.

4.1 Connecting a Controller to a Plant

In this review section, various interconnections between linear systems are considered. In particular, we consider the lower linear fractional transformation and the Redheffer star product of linear systems. Such interconnections have to be well-posed. For the $\mathcal{H}_2$-control problem, we consider the following set-up with a controller K connected to an LTI plant G in the following manner (see for example pages 195-196 in [55]):

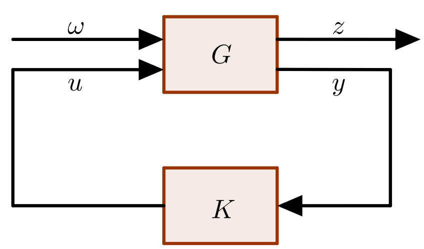


Figure 4.1: Connection of a controller to a plant

Here ω is an exogenous disturbance input which may include reference signals, noise and disturbances. The variable u is the control input and is determined by the controller K. The external output of the plant is represented by z and y is the measured output of the plant and is also the input of the controller K. The aim is to minimize the effect of the disturbance ω on the external output z. The plant also has a state variable

x and the controller has a state variable x_K. In the above set-up, the plant is a state space system with the following realization:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1\omega(t) + B_2u(t), & x(0) &= x_0 \\ z(t) &= C_1x(t) + D_{11}\omega(t) + D_{12}u(t), & t &\geq 0 \\ y(t) &= C_2x(t) + D_{21}\omega(t) + D_{22}u(t). \end{aligned} \tag{4.1}$$

This is a plant $G \sim (A, B, C, D)$ with states $x(t) \in \mathbb{R}^n$, combined inputs $\begin{bmatrix}\omega(t) & u(t)\end{bmatrix}^\intercal \in \mathbb{R}^\ell \oplus \mathbb{R}^m$, combined outputs $\begin{bmatrix}z(t) & y(t)\end{bmatrix}^\intercal \in \mathbb{R}^r \oplus \mathbb{R}^s$ and

$$B = \begin{bmatrix}B_1 & B_2\end{bmatrix}, \quad C = \begin{bmatrix}C_1 \\ C_2\end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix}D_{11} & D_{12} \\ D_{21} & D_{22}\end{bmatrix},$$

where the input is a combined input consisting of the control input u and the disturbance ω and where the output is a combined output consisting of the external output z and the measured output y. The controller is itself a LTI system with state space realization

$$\begin{aligned} \dot{x}_K(t) &= A_Kx_K(t) + B_Ky(t), & x_K(0) &= x_{K,0} \\ u(t) &= C_Kx_K(t) + D_Ky(t), & t &\geq 0, \end{aligned} \tag{4.2}$$

where x_K is the state variable of the controller. We note that the measured output y of the plant, is the input of the controller and the output of the controller is the control input u of the plant. Connecting the controller to the plant in this manner gives the following set of equations:

$$\begin{aligned} \begin{bmatrix}\dot{x}(t) \\ \dot{x}_K(t)\end{bmatrix} &= \begin{bmatrix}A & 0 \\ 0 & A_K\end{bmatrix}\begin{bmatrix}x(t) \\ x_K(t)\end{bmatrix} + \begin{bmatrix}B_2 & 0 \\ 0 & B_K\end{bmatrix}\begin{bmatrix}u(t) \\ y(k)\end{bmatrix} + \begin{bmatrix}B_1 \\ 0\end{bmatrix}\omega(t), \\ \begin{bmatrix}I & -D_K \\ -D_{22} & I\end{bmatrix}\begin{bmatrix}u(t) \\ y(t)\end{bmatrix} &= \begin{bmatrix}0 & C_K \\ C_2 & 0\end{bmatrix}\begin{bmatrix}x(t) \\ x_K(t)\end{bmatrix} + \begin{bmatrix}0 \\ D_{21}\end{bmatrix}\omega(t). \end{aligned}$$

The connection between the plant P and the controller K is said to be *well-posed* if unique solutions exist for $x(t)$, $x_K(t)$, $y(t)$ and $u(t)$ for all initial conditions $x(0)$ and $x_K(0)$ and all disturbance inputs $\omega(t)$ (see page 174 in [8]). If the coefficient matrix of $\begin{bmatrix}u(t) & y(t)\end{bmatrix}^\intercal$ in the second equation above is non-singular, then we can solve for $\begin{bmatrix}u(t) & y(t)\end{bmatrix}^\intercal$ uniquely. Using Schur complements, Lemma 2.2.3 gives the following equivalent condition for well-posedness.

Proposition 4.1.1 (cf. Proposition 5.1 in [8] or Lemma 5.1 in [56])**.**
The connection between a plant $G \sim (A, B, C, D)$ as in (4.1) and a controller $K \sim (A_K, B_K, C_K, D_K)$ as in (4.2) is well posed if and only if $I - D_{22}D_K$ (or equivalently $I - D_KD_{22}$) is non-singular. In particular, the connection between G and K is well posed if $D_{22} = 0$ or $D_K = 0$.

Next, we consider state space formulas for the closed loop system after connecting a controller to the plant. This closed loop transfer function is given by the linear fractional transformation (see subsection 2.2.3) of the transfer functions of the plant and the controller. Passing to the frequency domain, if $\widehat{G}$ and $\widehat{K}$ are the transfer functions of the plant and the controller respectively, then

$$\widehat{G} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] = \begin{bmatrix}C_1\Phi_AB_1 + D_{11} & C_1\Phi_AB_2 + D_{12} \\ C_2\Phi_AB_1 + D_{21} & C_2\Phi_AB_2 + D_{22}\end{bmatrix} =: \begin{bmatrix}\widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22}\end{bmatrix} \quad \text{and} \tag{4.3}$$

$$\widehat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right] = C_K\Phi_{A_K}B_K + D_K. \tag{4.4}$$

Hence it follows that

$$\begin{bmatrix}\widehat{z} \\ \widehat{y}\end{bmatrix} = \begin{bmatrix}\widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22}\end{bmatrix}\begin{bmatrix}\widehat{\omega} \\ \widehat{u}\end{bmatrix} \quad \text{and} \quad \widehat{u} = \widehat{K}\widehat{y}.$$

It can then be shown that closed loop transfer function from $\widehat{\omega}$ to $\widehat{z}$ is

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{K}(I-\widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21},$$

where $\underline{\mathcal{F}}(\widehat{G},\widehat{K})$ is the lower linear fractional transformation (LFT) of $\widehat{G}$ and $\widehat{K}$. If the connection between the plant P and the controller K is well posed, then the LFT between their transfer functions $\widehat{G}$ and $\widehat{K}$ is well defined and proper (see page 196 in [8]). Since the control problem is to minimize the effect of the disturbance ω on the external output z, the aim is to construct a viable controller which minimizes the $\mathcal{H}_2$-norm of the closed loop transfer function $\underline{\mathcal{F}}(\widehat{G},\widehat{K})$ from $\widehat{\omega}$ to $\widehat{z}$. Given realizations of $\widehat{G}$ and $\widehat{K}$ as in (4.3) and (4.4) respectively, the following theorem provides a realization of $\underline{\mathcal{F}}(\widehat{G},\widehat{K})$ in terms of the realizations of $\widehat{G}$ and $\widehat{K}$. Although these formulas are well known, we did not find a proof. As these formulas will appear frequently, we provide a proof for the sake of completeness.

Theorem 4.1.2 (cf. p.224 in [8]).
If the connection between the plant $\widehat{G}$ as in (4.3) and the controller $\widehat{K}$ as in (4.4) is well posed, then the closed loop transfer function after connecting the controller $\widehat{K}$ to the system $\widehat{G}$, has realization $(\underline{A},\underline{B},\underline{C},\underline{D})$, that is,

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{K}(I-\widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21} = \left[\begin{array}{c|c} \underline{A} & \underline{B} \\ \hline \underline{C} & \underline{D} \end{array}\right], \tag{4.5}$$

where

$$\underline{A} = \begin{bmatrix} A & 0 \\ 0 & A_K \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix},$$

$$\underline{B} = \begin{bmatrix} B_2 & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ D_{21} \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix},$$

$$\underline{C} = \begin{bmatrix} C_1 & 0 \end{bmatrix} + D_{12}\begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \qquad \text{and}$$

$$\underline{D} = D_{11} + D_{12}\begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I \end{bmatrix} D_{21}.$$

In particular, if $D_{22} = 0$, then

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \left[\begin{array}{c|c} \underline{A} & \underline{B} \\ \hline \underline{C} & \underline{D} \end{array}\right] = \left[\begin{array}{cc|c} A + B_2D_KC_2 & B_2C_K & B_1 + B_2D_KD_{21} \\ B_KC_2 & A_K & B_KD_{21} \\ \hline C_1 + D_{12}D_KC_2 & D_{12}C_K & D_{11} + D_{12}D_KD_{21} \end{array}\right]. \tag{4.6}$$

Proof.
We will prove the special case, where $D_{22} = 0$. First consider $(I-\widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21}$, which is computed by applying the basic operations of system multiplication (3.13) and system inversion (3.14):

$$\begin{aligned}
(I-\widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21} &= \left(I - \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & 0 \end{array}\right]\left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]\right)^{-1} \left[\begin{array}{c|c} A & B_1 \\ \hline C_2 & D_{21} \end{array}\right] \\
&= \left[\begin{array}{cc|c} A & B_2C_K & B_2D_K \\ 0 & A_K & B_K \\ \hline -C_2 & 0 & I \end{array}\right]^{-1} \left[\begin{array}{c|c} A & B_1 \\ \hline C_2 & D_{21} \end{array}\right] \\
&= \left[\begin{array}{cc|c} A + B_2D_KC_2 & B_2C_K & B_2D_K \\ B_KC_2 & A_K & B_K \\ \hline C_2 & 0 & I \end{array}\right] \left[\begin{array}{c|c} A & B_1 \\ \hline C_2 & D_{21} \end{array}\right] \\
&= \left[\begin{array}{ccc|c} A + B_2D_KC_2 & B_2C_K & B_2D_KC_2 & B_2D_KD_{21} \\ B_KC_2 & A_K & B_KC_2 & B_KD_{21} \\ 0 & 0 & A & B_1 \\ \hline C_2 & 0 & C_2 & D_{21} \end{array}\right].
\end{aligned}$$

Here

$$\begin{bmatrix} B_2 D_K C_2 \\ B_K C_2 \end{bmatrix} = \begin{bmatrix} A + B_2 D_K C_2 & B_2 C_K \\ B_K C_2 & A_K \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix} - \begin{bmatrix} I \\ 0 \end{bmatrix} A.$$

Applying (3.18) with $X = \begin{bmatrix} I & 0 \end{bmatrix}^\intercal$, gives

$$\left[\begin{array}{ccc|c} A + B_2 D_K C_2 & B_2 C_K & B_2 D_K C_2 & B_2 D_K D_{21} \\ B_K C_2 & A_K & B_K C_2 & B_K D_{21} \\ 0 & 0 & A & B_1 \\ \hline C_2 & 0 & C_2 & D_{21} \end{array}\right]$$
$$= \left[\begin{array}{cc|c} A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ B_K C_2 & A_K & B_K D_{21} \\ \hline C_2 & 0 & D_{21} \end{array}\right] + \left[\begin{array}{c|c} A & B_1 \\ \hline 0 & 0 \end{array}\right]$$

which leaves only the first term since the second term is zero. We now compute

$$\begin{aligned} \widehat{K}(I - \widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21} &= \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right] \left[\begin{array}{cc|c} A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ B_K C_2 & A_K & B_K D_{21} \\ \hline C_2 & 0 & D_{21} \end{array}\right] \\ &= \left[\begin{array}{ccc|c} A_K & B_K C_2 & 0 & B_K D_{21} \\ 0 & A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ 0 & B_K C_2 & A_K & B_K D_{21} \\ \hline C_K & D_K C_2 & 0 & D_K D_{21} \end{array}\right] \\ &= \left[\begin{array}{c|c} A_K & 0 \\ \hline C_K & 0 \end{array}\right] + \left[\begin{array}{cc|c} A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ B_K C_2 & A_K & B_K D_{21} \\ \hline D_K C_2 & C_K & D_K D_{21} \end{array}\right] \\ &= \left[\begin{array}{cc|c} A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ B_K C_2 & A_K & B_K D_{21} \\ \hline D_K C_2 & C_K & D_K D_{21} \end{array}\right] \end{aligned}$$

by application of (3.19) with $X = \begin{bmatrix} 0 & I \end{bmatrix}$ in the third step, because

$$\begin{bmatrix} B_K C_2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} A + B_2 D_K C_2 & B_2 C_K \\ B_K C_2 & A_K \end{bmatrix} - \begin{bmatrix} 0 & I \end{bmatrix} A_K.$$

Then

$$\begin{aligned} \widehat{G}_{12}\widehat{K}(I - \widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21} &= \left[\begin{array}{c|c} A & B_2 \\ \hline C_1 & D_{12} \end{array}\right] \left[\begin{array}{cc|c} A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ B_K C_2 & A_K & B_K D_{21} \\ \hline D_K C_2 & C_K & D_K D_{21} \end{array}\right] \\ &= \left[\begin{array}{ccc|c} A & B_2 D_K C_2 & B_2 C_K & B_2 D_K D_{21} \\ 0 & A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ 0 & B_K C_2 & A_K & B_K D_{21} \\ \hline C_1 & D_{12} D_K C_2 & D_{12} C_K & D_{12} D_K D_{21} \end{array}\right] \\ &= \left[\begin{array}{c|c} A & -B_1 \\ \hline C_1 & 0 \end{array}\right] + \left[\begin{array}{cc|c} A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ B_K C_2 & A_K & B_K D_{21} \\ \hline C_1 + D_{12} D_K C_2 & D_{12} C_K & D_{12} D_K D_{21} \end{array}\right] \end{aligned}$$

by application of (3.19) with $X = \begin{bmatrix} I & 0 \end{bmatrix}$ in the third step, because

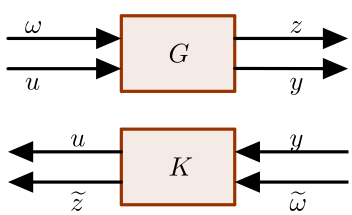
$$\begin{bmatrix} B_2 D_K C_2 & B_2 C_K \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} A + B_2 D_K C_2 & B_2 C_K \\ B_K C_2 & A_K \end{bmatrix} - \begin{bmatrix} I & 0 \end{bmatrix} A.$$

So finally,

$$\begin{aligned}
\underline{\mathcal{F}}(\widehat{G},\widehat{K}) &= \widehat{G}_{11}+\widehat{G}_{12}\widehat{K}(I-\widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21}\\
&= \left[\begin{array}{c|c} A & B_1 \\ \hline C_1 & D_{11}\end{array}\right] + \left[\begin{array}{c|c} A & -B_1 \\ \hline C_1 & 0\end{array}\right] + \left[\begin{array}{cc|c} A+B_2D_KC_2 & B_2C_K & B_1+B_2D_KD_{21} \\ B_KC_2 & A_K & B_KD_{21} \\ \hline C_1+D_{12}D_KC_2 & D_{12}C_K & D_{12}D_KD_{21}\end{array}\right]\\
&= \left[\begin{array}{cc|c} A+B_2D_KC_2 & B_2C_K & B_1+B_2D_KD_{21} \\ B_KC_2 & A_K & B_KD_{21} \\ \hline C_1+D_{12}D_KC_2 & D_{12}C_K & D_{11}+D_{12}D_KD_{21}\end{array}\right],
\end{aligned}$$

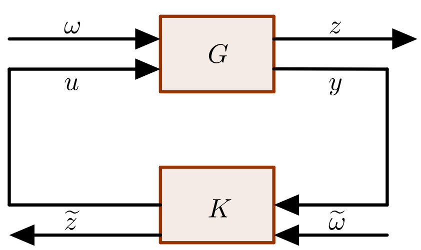
which is what we had to show. □

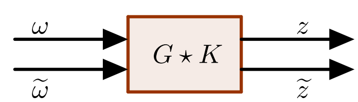
In the preceding discussion an LTI plant G with two inputs u and ω and two outputs z and y, was connected to a controller K with input y and output u. That resulted in a closed loop system whose transfer function is $\underline{\mathcal{F}}(\widehat{G},\widehat{K})$ with input ω and output z. More generally, the controller K itself can also be an LTI system with two inputs y and $\widetilde{\omega}$ and two outputs u and $\widetilde{z}$. In this case the measured output y of G is the control input of K and the measured output u of K is the control input of G. In such a case G and K are said to be *compatible*. Under certain conditions, G and K can be interconnected and this interconnection results in a new system with two inputs and two outputs. The transfer function of this new system is given by the Redheffer star product of $\widehat{G}$ and $\widehat{K}$ (see Section 2.2.3, Definition 2.2.9). In section 4.3, we will see that it is reasonable to assume that both $D_{11}=0$ and $D_{22}=0$. Given two compatible LTI systems G and K with realizations $\widehat{G}$ and $\widehat{K}$ as below



$$\begin{aligned}
\widehat{G} &= \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0\end{array}\right] = \begin{bmatrix}\widehat{G}_{11} & \widehat{G}_{12}\\ \widehat{G}_{21} & \widehat{G}_{22}\end{bmatrix}\\
\widehat{K} &= \left[\begin{array}{c|cc} A_K & B_{K1} & B_{K2} \\ \hline C_{K1} & 0 & D_{K12} \\ C_{K2} & D_{K21} & 0\end{array}\right] = \begin{bmatrix}\widehat{K}_{11} & \widehat{K}_{12}\\ \widehat{K}_{21} & \widehat{K}_{22}\end{bmatrix}
\end{aligned} \tag{4.7}$$

the Redheffer star product of $\widehat{G}$ and $\widehat{K}$ represents the interconnection of the two systems as shown below.





The last result of this section gives a realization of $\widehat{G}\star\widehat{K}$ in terms of the realizations of $\widehat{G}$ and $\widehat{K}$.

Theorem 4.1.3 (cf. pages 179-180 in [56])**.**
If $\widehat{G}$ and $\widehat{K}$ have compatible realizations as in (4.7), then

$$\begin{aligned}
\widehat{G}*\widehat{K} &= \begin{bmatrix} \underline{\mathcal{F}}(\widehat{G},\widehat{K}_{11}) & \widehat{G}_{12}(I-\widehat{K}_{11}\widehat{G}_{22})^{-1}\widehat{K}_{12} \\ \widehat{K}_{21}(I-\widehat{G}_{22}\widehat{K}_{11})^{-1}\widehat{G}_{21} & \overline{\mathcal{F}}(\widehat{K},\widehat{G}_{22})\end{bmatrix}\\
&= \left[\begin{array}{cc|cc} A & B_2C_{K_1} & B_1 & B_2D_{K_{12}} \\ B_{K_1}C_2 & A_K & B_{K_1}D_{21} & B_{K_2} \\ \hline C_1 & D_{12}C_{K_1} & 0 & D_{12}D_{K_{12}} \\ D_{K_{21}}C_2 & C_{K_2} & D_{K_{21}}D_{21} & 0\end{array}\right].
\end{aligned}$$

4.2 Stabilizing Controllers and the Youla Parametrization

In this section, we consider a plant $\widehat{G}$ and a controller $\widehat{K}$ as given below in equation (4.8). It is assumed that $D_{11} = 0$ and $D_{22} = 0$. Again, we mention that these assumptions are reasonable in the context of the $\mathcal{H}_2$-control problem, as will be justified in the following section. Let

$$\widehat{G} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix} \quad \text{and} \quad \widehat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]. \tag{4.8}$$

The connection between the plant G and K is well-posed, because $D_{22} = 0$. A controller K *stabilizes* the plant G if the resulting closed loop matrix $\underline{A}$ in (4.6) is Hurwitz, that is, if $\mathrm{Re}(\lambda) < 0$ for all $\lambda \in \sigma(\underline{A})$. Hence saying that K stabilizes G is the same as saying $\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) \in \mathcal{RH}_\infty$. The $\mathcal{H}_2$-control problem, which we will consider in the following section, is to construct a controller which stabilizes the plant and minimizes the $\mathcal{H}_2$-norm of the resulting closed loop transfer function $\underline{\mathcal{F}}(\widehat{G}, \widehat{K})$. This section is concerned with the first aspect of $\mathcal{H}_2$-control - the existence of controllers that stabilize the plant. In fact, a parametrization of all such stabilizing controllers is provided. This parametrization was first introduced by Youla et al. in [53] and [54] using coprime factorization techniques. We will, however, consider the state space approach as in [56].

The first task is to establish under which conditions stabilizing controllers of a plant such as in (4.8) exist. The following well-known proposition gives both a necessary and sufficient condition for the existence of stabilizing controllers.

Proposition 4.2.1 (cf. Lemma 11.1 in [56] or Proposition 5.6 in [8]).
A plant $\widehat{G}$ as in (4.8) has a stabilizing controller $\widehat{K}$ if and only if (A, B_2, C_2) is stabilizable and detectable. One such controller is

$$\widehat{K} = \left[\begin{array}{c|c} A + B_2E + MC_2 & -M \\ \hline E & 0 \end{array}\right],$$

where E and M are matrices such that $A + B_2E$ and $A + MC_2$ are stable.

The above proposition guarantees the existence of a stabilizing controller for a plant with a stabilizable and detectable realization. In optimal control problems we want to obtain a controller that is not only stabilizing, but which gives the optimal performance of the closed loop system. To this end, we are interested in characterizing all possible stabilizing controllers, after which an optimal controller may be determined. In the case that the plant is already stable (that is $\widehat{G} \in \mathcal{RH}_\infty$), stabilizing controllers are simply controllers that preserve the stability of the plant. In this case the parametrization of stabilizing controllers is easy.

Theorem 4.2.2 (cf. Theorem 11.3 in [56]).
If a plant $\widehat{G}$ as in (4.8) is stable (that is $\widehat{G} \in \mathcal{RH}_\infty$), then the set of all stabilizing controllers of $\widehat{G}$ is

$$\left\{ \widehat{K} = \widehat{R}(I + \widehat{G}_{22}\widehat{R})^{-1} : \widehat{R} \in \mathcal{RH}_\infty \right\}.$$

The previous theorem shows that if $\widehat{G}$ is stable and $\widehat{K}$ is a stabilizing controller for $\widehat{G}$, then $\widehat{K} = \widehat{R}(I + G_{22}\widehat{R})^{-1}$ for some $\widehat{R} \in \mathcal{RH}_\infty$. However, in the case that $\widehat{G}$ is not necessarily stable, it is much more complicated to parametrize the set of stabilizing controllers of $\widehat{G}$.

The famous Youla parametrization shows that all stabilizing controllers $\widehat{K}$ can obtained via the LFT of some canonical system $\widehat{J}$ and a arbitrary stable controller $\widehat{R} \in \mathcal{RH}_\infty$. Furthermore, the closed loop transfer function after connecting the controller $\widehat{K}$ to $\widehat{G}$ is the same as the closed loop transfer function after connecting $\widehat{R}$ to the plant $\widehat{H}$, where $\widehat{H}$ is a stable plant obtained by taking the Redheffer star product of the original plant $\widehat{G}$ and the canonical plant $\widehat{J}$.

Theorem 4.2.3 (The Youla Parametrization (cf. Theorem 11.4 in [56])).
Consider a plant $\widehat{G}$ *as in* (4.8) *with* (A, B_2, C_2) *stabilizable and detectable and let* E *and* M *be matrices such that* $A + B_2E$ *and* $A + MC_2$ *are stable. Define*

$$\widehat{J} := \left[\begin{array}{c|cc} A + B_2E + MC_2 & -M & B_2 \\ \hline E & 0 & I \\ -C_2 & I & 0 \end{array}\right]. \tag{4.9}$$

Then the set of stabilizing controllers for $\widehat{G}$ *is given by*

$$\mathcal{S} = \left\{ \widehat{K} = \underline{\mathcal{F}}(\widehat{J}, \widehat{R}) : \widehat{R} \in \mathcal{RH}_\infty \right\}.$$

Furthermore, the closed loop transfer function after connecting a stabilizing controller

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}, \widehat{R}) \quad \text{where} \quad \widehat{R} = \left[\begin{array}{c|c} A_R & B_R \\ \hline C_R & D_R \end{array}\right] \in \mathcal{RH}_\infty,$$

is given by

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \underline{\mathcal{F}}\left(\widehat{G}, \underline{\mathcal{F}}(\widehat{J}, \widehat{R})\right) = \underline{\mathcal{F}}\left(\widehat{G} \star \widehat{J}, \widehat{R}\right) = \underline{\mathcal{F}}\left(\widehat{H}, \widehat{R}\right) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21} \tag{4.10}$$

$$= \left[\begin{array}{ccc|c} A + B_2D_RC_2 & B_2E - B_2D_RC_2 & B_2C_R & B_1 + B_2D_RD_{21} \\ B_2D_RC_2 - MC_2 & A + B_2E + MC_2 - B_2D_RC_2 & B_2C_R & B_2D_RD_{21} - MD_{21} \\ B_RC_2 & -B_RC_2 & A_R & B_RD_{21} \\ \hline C_1 + D_{12}D_RC_2 & D_{12}E - D_{12}D_RC_2 & D_{12}C_R & D_{12}D_RD_{21} \end{array}\right],$$

where $\widehat{H}$ *is given by the Redheffer star product of* $\widehat{G}$ *and* $\widehat{J}$*:*

$$\widehat{H} = \widehat{G} \star \widehat{J} = \begin{bmatrix} \widehat{H}_{11} & \widehat{H}_{12} \\ \widehat{H}_{21} & \widehat{H}_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} A & B_2E & B_1 & B_2 \\ -MC_2 & A + B_2E + MC_2 & -MD_{21} & B_2 \\ \hline C_1 & D_{12}E & 0 & D_{12} \\ C_2 & -C_2 & D_{21} & 0 \end{array}\right]. \tag{4.11}$$

Finally, with $\widehat{H}$ *as in* (4.11), *we have* $\widehat{H}_{22} = 0$.

The following illustration shows how G, J and R are interconnected to give

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \underline{\mathcal{F}}\left(\widehat{H}, \widehat{R}\right), \quad \text{where} \quad \widehat{H} = \widehat{G} \star \widehat{J} \quad \text{and} \quad \widehat{K} = \underline{\mathcal{F}}(\widehat{J}, \widehat{R}):$$

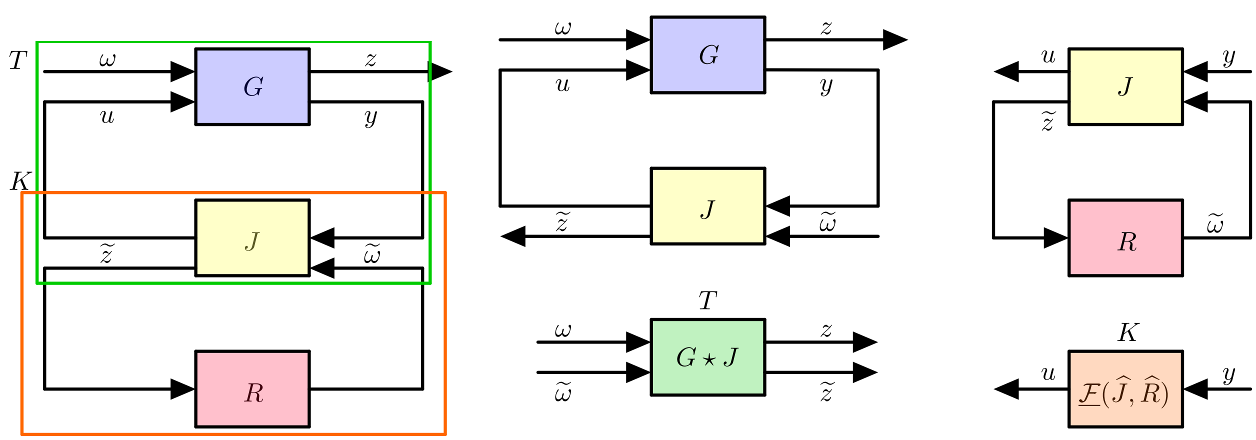


Figure 4.2: Youla parametrization

Note that that the realization for $\widehat{H} = \widehat{G} \star \widehat{J}$ in (4.11) is state space similar to

$$\left[\begin{array}{cc|cc} A+B_2E & -B_2E & B_1 & B_2 \\ 0 & A+MC_2 & B_1+MD_{21} & 0 \\ \hline C_1+D_{12}E & -D_{12}E & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right],$$

with state space similarity $S = \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}$. It is then easy to see that

$$\begin{aligned} \widehat{H}_{11} &= \left[\begin{array}{cc|c} A+B_2E & -B_2E & B_1 \\ 0 & A+MC_2 & B_1+MD_{21} \\ \hline C_1+D_{12}E & -D_{12}E & 0 \end{array}\right], \\ \widehat{H}_{12} &= \left[\begin{array}{c|c} A+B_2E & B_2 \\ \hline C_1+D_{12}E & D_{12} \end{array}\right], \quad \widehat{H}_{21} = \left[\begin{array}{c|c} A+MC_2 & B_1+MD_{21} \\ \hline C_2 & D_{21} \end{array}\right] \quad \text{and} \quad \widehat{H}_{22} = 0. \end{aligned} \tag{4.12}$$

Since $A + B_2E$ and $A + MC_2$ are stable, the plant $\widehat{H}$ is stable. Hence by Theorem 4.2.2, any $\widehat{R} \in \mathcal{RH}_\infty$ stabilizes $\widehat{H}$. And so it follows from (4.10) that $\widehat{K} = \underline{\mathcal{F}}(\widehat{J}, \widehat{R})$ stabilizes $\widehat{G}$.

The Youla parametrization decomposes the closed loop transfer function into two terms:

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}.$$

The first term is independent of $\widehat{R}$ and the second depends on $\widehat{R}$. The last result of this section shows that the $\widehat{R}$-independent term can further be decomposed into the sum of two terms.

Lemma 4.2.4.
Consider a plant $\widehat{G}$ as in (4.8) with (A, B_2, C_2) stabilizable and detectable and let E and M be matrices such that $A + B_2E$ and $A + MC_2$ are stable. Let $\widehat{H}$ be as in (4.11) in the Youla parametrization. Then

$$\widehat{H}_{11} = \widehat{G}_c - \widehat{H}_{12}\widehat{G}_f \quad \text{and hence} \quad \underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \widehat{G}_c - \widehat{H}_{12}\widehat{G}_f + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}.$$

where

$$\widehat{G}_c = \left[\begin{array}{c|c} A+B_2E & B_1 \\ \hline C_1+D_{12}E & 0 \end{array}\right] \quad \text{and} \quad \widehat{G}_f = \left[\begin{array}{c|c} A+MC_2 & B_1+MD_{21} \\ \hline E & 0 \end{array}\right].$$

Proof.
By equation (3.13),

$$\widehat{H}_{12}\widehat{G}_f = \left[\begin{array}{cc|c} A+B_2E & B_2E & 0 \\ 0 & A+MC_2 & B_1+MD_{21} \\ \hline C_1+D_{12}E & D_{12}E & 0 \end{array}\right].$$

By a special case of (3.22), it follows that

$$\begin{aligned} \widehat{G}_c - \widehat{H}_{12}\widehat{G}_f &= \left[\begin{array}{c|c} A+B_2E & B_1 \\ \hline C_1+D_{12}E & 0 \end{array}\right] - \left[\begin{array}{cc|c} A+B_2E & B_2E & 0 \\ 0 & A+MC_2 & B_1+MD_{21} \\ \hline C_1+D_{12}E & D_{12}E & 0 \end{array}\right] \\ &= \left[\begin{array}{cc|c} A+B_2E & 0-B_2E & B_1-0 \\ 0 & A+MC_2 & B_1+MD_{21} \\ \hline C_1+D_{12}E & 0-D_{12}E & 0-0 \end{array}\right] \\ &= \widehat{H}_{11} \end{aligned}$$

where the last equality follows by (4.12). □

4.3 $\mathcal{H}_2$-Control Problem Statement and Solution

In this review section, the $\mathcal{H}_2$-optimal control problem is stated and its solution is provided. The solution strategy makes use of the Youla parametrization (see Section 4.2) and inner functions (see Section 3.6). This review is based on Chapter 6 of [8] and Chapter 13 of [56].

Definition 4.3.1 ($\mathcal{H}_2$-control problem)**.**
For a LTI plant G with state space realization

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1 w(t) + B_2 u(t), && x(0) = x_0 \\ z(t) &= C_1 x(t) + D_{11} w(t) + D_{12} u(t), && t \geq 0 \\ y(t) &= C_2 x(t) + D_{21} w(t) + D_{22} u(t), && \end{aligned} \tag{4.13}$$

the $\mathcal{H}_2$*-control problem* is to synthesize a controller K:

$$\begin{aligned} \dot{x}_K(t) &= A_K x_K(t) + B_K y(t), && x_K(0) = x_{K,0} \\ u(t) &= C_K x_K(t) + D_K y(t), && t \geq 0, \end{aligned} \tag{4.14}$$

satisfying the following properties:

1. the connection between the plant G and the controller K is well posed;
2. $\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) \in \mathcal{RH}_2$ (that is, $\underline{A}$ is stable and $\underline{D} = 0$ in (4.6)),

such that the closed loop performance of the system $\|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2^2$ is minimized.

Before considering a solution strategy for the $\mathcal{H}_2-$control problem, we consider under which assumptions on the system, conditions 1. and 2. in the above definition are satisfied. Recall that by Proposition 4.1.1, the connection between the plant G and the controller K is well posed if and only if $I - D_K D_{22}$ is non-singular. In particular, if it is assumed that $D_{22} = 0$, then the connection between the plant and the controller is well posed. On the other hand if the connection is well posed but $D_{22} \neq 0$, then $I - D_K D_{22}$ is non-singular. Assuming this is the case, define $\widetilde{y} = y - D_{22} u$. Then $\widetilde{y}(t) = C_2 x(t) + D_{21} \omega(t)$ (with no $D_{22} u(t)$ term) and $y = \widetilde{y} + D_{22} u$. Hence, the output equation of the controller in Definition 4.3.1 becomes

$$\begin{aligned} u(t) = C_K x_K(t) + D_K y(t) &= C_K x_K(t) + D_K(\widetilde{y}(t) + D_{22} u(t)) \\ &= C_K x_K(t) + D_K \widetilde{y}(t) + D_K D_{22} u(t). \end{aligned}$$

Rearranging the above equation gives

$$(I - D_K D_{22}) u(t) = C_K x_K(t) + D_K \widetilde{y}(t).$$

Since $(I - D_K D_{22})$ is non-singular, we can solve for $u(t)$ to obtain

$$u(t) = (I - D_K D_{22})^{-1} (C_K x_K(t) + D_K \widetilde{y}(t)).$$

Substituting this into the state equation of the controller yields

$$\begin{aligned} \dot{x}_K(t) &= A_K x_K(t) + B_K y(t) \\ &= A_K x_K(t) + B_K(\widetilde{y}(t) + D_{22} u(t)) \\ &= A_K x_K(t) + B_K \widetilde{y}(t) + B_K D_{22} (I - D_K D_{22})^{-1} (C_K x_K(t) + D_K \widetilde{y}(t)) \\ &= (A_K + B_K D_{22} (I - D_K D_{22})^{-1} C_K) x_K(t) + (B_K + B_K D_{22} (I - D_K D_{22})^{-1} D_K) \widetilde{y}(t). \end{aligned}$$

This yields the following controller:

$$\widetilde{K} = \left[\begin{array}{c|c} \widetilde{A}_K & \widetilde{B}_K \\ \hline \widetilde{C}_K & \widetilde{D}_K \end{array}\right] = \left[\begin{array}{c|c} A_K + B_K D_{22} (I - D_K D_{22})^{-1} C_K & B_K + B_K D_{22} (I - D_K D_{22})^{-1} D_K \\ \hline (I - D_K D_{22})^{-1} C_K & (I - D_K D_{22})^{-1} D_K \end{array}\right].$$

The following Lemma shows that for a well posed connection between a plant and a controller, we can assume without loss of generality that $D_{22} = 0$. This is because connecting a controller $\widehat{K}$ as in (4.14) to a plant with $D_{22} \neq 0$ gives the same closed-loop transfer function as connecting $\widetilde{K}$ to a plant with $D_{22} = 0$.

Lemma 4.3.2.
Suppose the connection between a plant $\widehat{G}$ and a controller $\widehat{K}$ given by

$$\widehat{G} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \quad \textit{and} \quad \widehat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$$

is well posed. Let

$$\widetilde{G} = \begin{bmatrix} \widetilde{G}_{11} & \widetilde{G}_{12} \\ \widetilde{G}_{21} & \widetilde{G}_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] \quad \textit{and}$$

$$\widetilde{K} = \left[\begin{array}{c|c} A_K + D_{22}(I - D_K D_{22})^{-1} C_K & B_K + D_{22}(I - D_K D_{22})^{-1} D_K \\ \hline (I - D_K D_{22})^{-1} C_K & (I - D_K D_{22})^{-1} D_K \end{array}\right] = (I - \widehat{K} D_{22})^{-1} \widehat{K}.$$

Then $\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \underline{\mathcal{F}}(\widetilde{G}, \widetilde{K})$.

Proof.
By Proposition 4.1.1, $I - D_K D_{22}$ is non-singular and thus $\widetilde{K}$ is well defined. Now by equation (3.14) it follows that

$$\begin{aligned}(I - \widehat{K} D_{22})^{-1} &= \left[\begin{array}{c|c} A_K & -B_K D_{22} \\ \hline C_K & I - D_K D_{22} \end{array}\right]^{-1} \\ &= \left[\begin{array}{c|c} A_K + B_K D_{22}(I - D_K D_{22})^{-1} C_K & B_K D_{22}(I - D_K D_{22})^{-1} \\ \hline (I - D_K D_{22})^{-1} C_K & (I - D_K D_{22})^{-1} \end{array}\right]\end{aligned}$$

and hence by equation (3.13) we get that

$$\begin{aligned}&(I - \widehat{K} D_{22})^{-1} \widehat{K} \\ &= \left[\begin{array}{cc|c} A_K + B_K D_{22}(I - D_K D_{22})^{-1} C_K & B_K D_{22}(I - D_K D_{22})^{-1} C_K & B_K D_{22}(I - D_K D_{22})^{-1} D_K \\ 0 & A_K & B_K \\ \hline (I - D_K D_{22})^{-1} C_K & (I - D_K D_{22})^{-1} C_K & (I - D_K D_{22})^{-1} D_K \end{array}\right] \\ &= \left[\begin{array}{c|c} A_K + B_K D_{22}(I - D_K D_{22})^{-1} C_K & B_K + B_K D_{22}(I - D_K D_{22})^{-1} D_K \\ \hline (I - D_K D_{22})^{-1} C_K & (I - D_K D_{22})^{-1} D_K \end{array}\right] + \left[\begin{array}{c|c} A_K & B_K \\ \hline 0 & 0 \end{array}\right] \\ &= \widetilde{K} + 0,\end{aligned}$$

where we applied (3.18) with $X = I$ in the second equality. This shows that $\widetilde{K} = (I - \widehat{K} D_{22})^{-1} \widehat{K}$. Furthermore, since

$$\widetilde{G}_{11} = \widehat{G}_{11}, \quad \widetilde{G}_{12} = \widehat{G}_{12}, \quad \widetilde{G}_{21} = \widehat{G}_{21} \quad \text{and} \quad \widetilde{G}_{22} = \widehat{G}_{22} - D_{22},$$

it follows by (4.5) that

$$\begin{aligned}\underline{\mathcal{F}}(\widetilde{G}, \widetilde{K}) &= \widetilde{G}_{11} + \widetilde{G}_{12} \widetilde{K} (I - \widetilde{G}_{22} \widetilde{K})^{-1} \widetilde{G}_{21} \\ &= \widehat{G}_{11} + \widehat{G}_{12} (I - \widehat{K} D_{22})^{-1} \widehat{K} (I - (\widehat{G}_{22} - D_{22})(I - \widehat{K} D_{22})^{-1} \widehat{K})^{-1} \widehat{G}_{21}.\end{aligned}$$

By Lemma 2.2.7 it follows that $(I - \widehat{K} D_{22})^{-1} \widehat{K} = \widehat{K}(I - D_{22} \widehat{K})^{-1}$. Hence

$$\begin{aligned}\underline{\mathcal{F}}(\widetilde{G}, \widetilde{K}) &= \widehat{G}_{11} + \widehat{G}_{12} \widehat{K} (I - D_{22} \widehat{K})^{-1} (I - (\widehat{G}_{22} - D_{22}) \widehat{K} (I - D_{22} \widehat{K})^{-1})^{-1} \widehat{G}_{21} \\ &= \widehat{G}_{11} + \widehat{G}_{12} \widehat{K} [(I - \widehat{G}_{22} \widehat{K} (I - D_{22} \widehat{K})^{-1} + D_{22} \widehat{K} (I - D_{22} \widehat{K})^{-1})(I - D_{22} \widehat{K})]^{-1} \widehat{G}_{21} \\ &= \widehat{G}_{11} + \widehat{G}_{12} \widehat{K} [((I - D_{22} \widehat{K}) - \widehat{G}_{22} \widehat{K} + D_{22} \widehat{K})]^{-1} \widehat{G}_{21} \\ &= \widehat{G}_{11} + \widehat{G}_{12} \widehat{K} (I - \widehat{G}_{22} \widehat{K})^{-1} \widehat{G}_{21} \\ &= \underline{\mathcal{F}}(\widehat{G}, \widehat{K}),\end{aligned}$$

which is what we had to show. □

The previous lemma concludes our discussion on condition 1 in Definition 4.3.1 and shows that we may without loss of generality consider plants with $D_{22} = 0$. Concerning condition 2 in Definition 4.3.1, we recall from Theorem 4.1.2 that with $D_{22} = 0$, the closed loop transfer function after connecting a controller as in (4.14) to a plant as in (4.13), is given by

$$\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \left[\begin{array}{cc|c} A + B_2 D_K C_2 & B_2 C_K & B_1 + B_2 D_K D_{21} \\ B_K C_2 & A_K & B_K D_{21} \\ \hline C_1 + D_{12} D_K C_2 & D_{12} C_K & D_{11} + D_{12} D_K D_{21} \end{array}\right].$$

Now $\underline{\mathcal{F}}(\widehat{G}, \widehat{K})$ can only be in $\mathcal{RH}_2$ if it is strictly proper, that is, if $D_{11} + D_{12} D_K D_{21} = 0$. This will be the case if $D_{11} = 0$ and $D_K = 0$. This shows that it is reasonable to assume that $D_{11} = 0$ and to restrict to controllers with $D_K = 0$. In order to satisfy conditions 1 and 2 in Definition 4.3.1, we will henceforth assume that $D_{11} = 0$ and $D_{22} = 0$.

We mention here that the LQ-control problem, the Kalman filtering problem and the linear quadratic Gaussian (LQG) control problem can all be framed as $\mathcal{H}_2$-control problems.

Assume that stabilizing solutions $X = \mathrm{Ric}(A, B_2, C_1, D_{12})$ and $Y = \mathrm{Ric}(A^\intercal, C_2^\intercal, B_1^\intercal, D_{21}^\intercal)$ exist and let

$$F := -(D_{12}^\intercal D_{12})^{-1}(B_2^\intercal X + D_{12}^\intercal C_1) \quad \text{and} \quad L := -(Y C_2^\intercal + B_1 D_{21}^\intercal)(D_{21} D_{21}^\intercal)^{-1}, \tag{4.15}$$

then $A + B_2 F$ and $A + L C_2$ are stable (see Lemma 3.7.9). By the Youla parametrization (Theorem 4.2.3), if $\widehat{J}^\circ$ is given by equation (4.9) with F instead of E and L instead of M, then all stabilizing controllers of $\widehat{G}$ are given by

$$\widehat{K} = \underline{\mathcal{F}}\left(\widehat{J}^\circ, \widehat{R}\right) \qquad \text{with } \widehat{R} \in \mathcal{RH}_\infty \text{ arbitrary}$$

and $\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \widehat{H}_{11} + \widehat{H}_{12}^\circ \widehat{R} \widehat{H}_{21}^\circ$ where

$$\widehat{H}^\circ = \begin{bmatrix} \widehat{H}_{11}^\circ & \widehat{H}_{12}^\circ \\ \widehat{H}_{21}^\circ & 0 \end{bmatrix} = \widehat{G} \star \widehat{J}^\circ = \left[\begin{array}{cc|cc} A + B_2 F & -B_2 F & B_1 & B_2 \\ 0 & A + L C_2 & B_1 + L D_{21} & 0 \\ \hline C_1 + D_{12} F & -D_{12} F & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right]. \tag{4.16}$$

Lemma 4.2.4 shows that $\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right)$ can be decomposed into the sum of three terms:

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \widehat{G}_c - \left(\widehat{H}_{12}^\circ R_1^{-\frac{1}{2}}\right)\left(R_1^{\frac{1}{2}} \widehat{G}_f\right) + \left(\widehat{H}_{12}^\circ R_1^{-\frac{1}{2}}\right)\left(R_1^{\frac{1}{2}} \widehat{R} R_2^{\frac{1}{2}}\right)\left(R_2^{-\frac{1}{2}} \widehat{H}_{21}^\circ\right) \tag{4.17}$$

The following lemma shows that if we choose $E = F$ and $M = L$ in the Youla parametrization and Lemma 4.2.4, then transfer functions $\widehat{H}_{12}^\circ R_1^{-\frac{1}{2}}$ and $R_2^{-\frac{1}{2}} \widehat{H}_{21}^\circ$ appearing in the above decomposition (4.17) are inner and co-inner functions respectively and furthermore, terms appearing in this decomposition are orthogonal.

Lemma 4.3.3 (c.f. Lemma 13.6 in [56]).
Assume $R_1 = D_{12}^\intercal D_{12} > 0$ and $R_2 = D_{21} D_{21}^\intercal > 0$ and let F and L be as in (4.15). Set

$$\widehat{G}_c = \left[\begin{array}{c|c} A + B_2 F & B_1 \\ \hline C_1 + D_{12} F & 0 \end{array}\right] \qquad \text{and} \qquad \widehat{G}_f = \left[\begin{array}{c|c} A + L C_2 & B_1 + L D_{21} \\ \hline F & 0 \end{array}\right].$$

Then

1. $\widehat{U} = \widehat{H}_{12}^\circ R_1^{-\frac{1}{2}}$ *is inner and* $\widehat{V} = R_2^{-\frac{1}{2}} \widehat{H}_{21}^\circ$ *is co-inner;*
2. $\widehat{U}^* \widehat{G}_c \in \mathcal{RH}_2^\perp$ *and* $\widehat{G}_f \widehat{V}^* \in \mathcal{RH}_2^\perp$.

Proof.
Let $X = \text{Ric}(A, B_2, C_1, D_{12})$ and let F be as in (4.15). Set $A_F = A + B_2F$ and $C_F = C_1 + D_{12}F$. With $E = F$ in (4.16), it follows that $\widehat{U} = \widehat{H}^{\circ}_{12}R_1^{-\frac{1}{2}}$ is given by

$$\widehat{U} = \left[\begin{array}{cc|c} A + B_2F & -B_2E & B_2 \\ 0 & A + MC_2 & 0 \\ \hline C_1 + D_{12}F & -D_{12}F & D_{12} \end{array}\right] R_1^{-\frac{1}{2}} = \left[\begin{array}{c|c} A + B_2F & B_2 \\ \hline C_1 + D_{12}F & D_{12} \end{array}\right] R_1^{-\frac{1}{2}} = \left[\begin{array}{c|c} A_F & B_2R_1^{-\frac{1}{2}} \\ \hline C_F & D_{12}R_1^{-\frac{1}{2}} \end{array}\right].$$

Using equation (3.15) to calculate $\widehat{U}^*$, we get

$$\widehat{U}^* = \left[\begin{array}{c|c} -A_F^\intercal & -C_F^\intercal \\ \hline R_1^{-\frac{1}{2}}B_2^\intercal & R_1^{-\frac{1}{2}}D_{12}^\intercal \end{array}\right].$$

Then by the multiplication formula (3.13), it follows that

$$\widehat{U}^*\widehat{U} = \left[\begin{array}{cc|c} -A_F^\intercal & -C_F^\intercal C_F & -C_F^\intercal D_{12}R_1^{-\frac{1}{2}} \\ 0 & A_F & B_2R_1^{-\frac{1}{2}} \\ \hline R_1^{-\frac{1}{2}}B_2^\intercal & R_1^{-\frac{1}{2}}D_{12}^\intercal C_F & I \end{array}\right] \quad \text{and} \quad \widehat{U}^*\widehat{G}_c = \left[\begin{array}{cc|c} -A_F^\intercal & -C_F^\intercal C_F & 0 \\ 0 & A_F & I \\ \hline R_1^{-\frac{1}{2}}B_2^\intercal & R_1^{-\frac{1}{2}}D_{12}^\intercal C_F & 0 \end{array}\right].$$

Define the similarity transform T by

$$T = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}.$$

By Lemma 3.7.9 $B_2^\intercal X + D_{12}^\intercal C_F = 0$ and $A_F^\intercal X + XA_F + C_F^\intercal C_F = 0$, because $X = \text{Ric}(A, B_2, C_1, D_{12})$. Applying the similarity transform T and these identities, we see that

$$\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}\begin{bmatrix} -A_F^\intercal & -C_F^\intercal C_F \\ 0 & A_F \end{bmatrix}\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} -A_F^\intercal & -(A_F^\intercal X + XA_F + C_F^\intercal C_F) \\ 0 & A_F \end{bmatrix} = \begin{bmatrix} -A_F^\intercal & 0 \\ 0 & A_F \end{bmatrix}$$

$$\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}\begin{bmatrix} -C_F^\intercal D_{12}R_1^{-\frac{1}{2}} \\ B_2R_1^{-\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} -(XB_2 + C_F^\intercal D_{12})R_1^{-\frac{1}{2}} \\ B_2R_1^{-\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} -(B_2^\intercal X + D_{12}^\intercal C_F)^\intercal R_1^{-\frac{1}{2}} \\ B_2R_1^{-\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ B_2R_1^{-\frac{1}{2}} \end{bmatrix}$$

$$\begin{bmatrix} R_1^{-\frac{1}{2}}B_2^\intercal & R_1^{-\frac{1}{2}}D_{12}^\intercal C_F \end{bmatrix}\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} R_1^{-\frac{1}{2}}B_2^\intercal & R_1^{-\frac{1}{2}}(B_2^\intercal X + D_{12}^\intercal C_F) \end{bmatrix} = \begin{bmatrix} R_1^{-\frac{1}{2}}B_2^\intercal & 0 \end{bmatrix} \quad \text{and}$$

$$\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}\begin{bmatrix} 0 \\ I \end{bmatrix} = \begin{bmatrix} -X \\ I \end{bmatrix}.$$

Now $A_F = A + B_2F$ is stable and

$$\widehat{U}^*\widehat{U} = \left[\begin{array}{cc|c} -A_F^\intercal & 0 & 0 \\ 0 & A_F & B_2R_1^{-\frac{1}{2}} \\ \hline R_1^{-\frac{1}{2}}B_2^\intercal & 0 & I \end{array}\right] = I + \begin{bmatrix} R_1^{-\frac{1}{2}}B_2^\intercal & 0 \end{bmatrix}\begin{bmatrix} \Phi_{-A_F^\intercal} & 0 \\ 0 & \Phi_{A_F} \end{bmatrix}\begin{bmatrix} 0 \\ B_2R_1^{-\frac{1}{2}} \end{bmatrix} = I + 0 = I,$$

which shows that $\widehat{U}$ is inner. Secondly, since $-A_F^\intercal$ is anti-stable and

$$\widehat{U}^*\widehat{G}_c = \left[\begin{array}{cc|c} -A_F^\intercal & 0 & -X \\ 0 & A_F & I \\ \hline R_1^{-\frac{1}{2}}B_2^\intercal & 0 & 0 \end{array}\right] = \begin{bmatrix} R_1^{-\frac{1}{2}}B_2^\intercal & 0 \end{bmatrix}\begin{bmatrix} \Phi_{-A_F^\intercal} & 0 \\ 0 & \Phi_{A_F} \end{bmatrix}\begin{bmatrix} -X \\ I \end{bmatrix} = -R_1^{-\frac{1}{2}}B_2^\intercal\Phi_{-A_F^\intercal}X$$
$$= \left[\begin{array}{c|c} -A_F^\intercal & -X \\ \hline R_1^{-\frac{1}{2}}B_2^\intercal & 0 \end{array}\right].$$

it follows that $\widehat{U}^*\widehat{G}_c \in \mathcal{RH}_2^\perp$. The claims regarding $\widehat{V}$ follow in a similar way as the claims regarding $\widehat{U}$ by taking transposes and noting that $\widehat{G}^\intercal$ is co-inner if and only if $\widehat{G}$ is inner. □

The following theorem gives the solution of the $\mathcal{H}_2$-control problem and is the main result of this section. It follows by taking $E = F$ and $M = L$ in the Youla parametrization and Lemmas 4.2.4 and 4.3.3, where F and L are the stabilizing matrices in equation (4.15). In order to do this one needs that stabilizing solutions to the Riccati equations $X = \text{Ric}(A, B_2, C_1, D_{12})$ and $Y = \text{Ric}(A^\intercal, C_2^\intercal, B_1^\intercal, D_{21}^\intercal)$ exist. The existence of these solutions can be guaranteed by imposing the conditions in Theorem 3.7.8 on the plant $\widehat{G}$. Hence these are exactly the sufficient conditions which guarantee a unique solution to the $\mathcal{H}_2$-control problem.

Theorem 4.3.4 (Solution of the $\mathcal{H}_2$-optimal control problem (cf. Theorem 13.7 in [56]))**.**
Consider a plant

$$\widehat{G} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] = \begin{bmatrix} C_1\Phi_A B_1 & C_1\Phi_A B_2 + D_{12} \\ C_2\Phi_A B_1 + D_{21} & C_2\Phi_A B_2 \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix}, \tag{4.18}$$

that satisfies the following conditions

1. *the pairs (A, B_1) and (A, B_2) are stabilizable;*
2. *$R_1 = D_{12}^\intercal D_{12} > 0$ and $R_2 = D_{21}D_{21}^\intercal > 0$;*
3. *$\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ and $\begin{bmatrix} A - i\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ have full column and row rank respectively for all $\omega \in \mathbb{R}$.*

Then there exist unique positive semi-definite symmetric stabilizing solutions $X = \text{Ric}(A, B_2, C_1, D_{12})$ and $Y = \text{Ric}(A^\intercal, C_2^\intercal, B_1^\intercal, D_{21}^\intercal)$ to the Riccati equations

$$\begin{aligned} A^\intercal X + XA + C_1^\intercal C_1 - (XB_2 + C_1^\intercal D_{12})R_1^{-1}(B_2^\intercal X + D_{12}^\intercal C_1) = 0 \\ AY + YA^\intercal + B_1B_1^\intercal - (YC_2^\intercal + B_1D_{21}^\intercal)R_2^{-1}(C_2Y + D_{21}B_1^\intercal) = 0 \end{aligned}$$

and subsequently, if the static feedback matrices F and L are defined by

$$F = -R_1^{-1}(B_2^\intercal X + D_{12}^\intercal C_1) \quad \text{and} \quad L = -(YC_2^\intercal + B_1D_{21}^\intercal)R_2^{-1},$$

then the unique optimal solution to the $\mathcal{H}_2$-control problem 4.3.1 is

$$\widehat{K}_{opt} = \left[\begin{array}{c|c} A + B_2F + LC_2 & -L \\ \hline F & 0 \end{array}\right].$$

Furthermore, the closed loop performance achieved by the optimal controller K_2 is

$$\|\underline{\mathcal{F}}(\widehat{G}, \widehat{K}_{opt})\|_2^2 = \|\widehat{G}_c\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f\|_2^2,$$

where

$$\widehat{G}_c = \left[\begin{array}{c|c} A + B_2F & B_1 \\ \hline C_1 + D_{12}F & 0 \end{array}\right] \quad \text{and} \quad \widehat{G}_f = \left[\begin{array}{c|c} A + LC_2 & B_1 + LD_{21} \\ \hline F & 0 \end{array}\right].$$

Proof.
Let $\widehat{J}^\circ$ be given as in the Youla parametrization (Theorem 4.2.3), that is

$$\widehat{J}^\circ := \left[\begin{array}{c|cc} A + B_2F + LC_2 & -L & B_2 \\ \hline F & 0 & I \\ -C_2 & I & 0 \end{array}\right].$$

Then

$$\widehat{H}^\circ = \widehat{G} \star \widehat{J}^\circ = \left[\begin{array}{cc|cc} A + B_2F & -B_2F & B_1 & B_2 \\ 0 & A + LC_2 & B_1 + LD_{21} & 0 \\ \hline C_1 + D_{12}F & -D_{12}F & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right].$$

All stabilizing controllers of $\widehat{G}$ are given by

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}^\circ, \widehat{R}),$$

where $\widehat{R} \in \mathcal{RH}_\infty$ is arbitrary. For such controllers, the closed loop transfer functions after connecting $\widehat{K}$ to $\widehat{G}$ are also given by the Youla parametrization:

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \underline{\mathcal{F}}\left(\widehat{G}, \underline{\mathcal{F}}(\widehat{J}^\circ, \widehat{R})\right) = \underline{\mathcal{F}}\left(\widehat{G} \star \widehat{J}, \widehat{R}\right) = \underline{\mathcal{F}}\left(\widehat{H}^\circ, \widehat{R}\right) = \widehat{H}^\circ_{11} + \widehat{H}^\circ_{12}\widehat{R}\widehat{H}^\circ_{21}.$$

Lemma 4.2.4 gives that

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \underline{\mathcal{F}}\left(\widehat{G}, \underline{\mathcal{F}}(\widehat{J}^\circ, \widehat{R})\right) = \widehat{G}_c - \widehat{U}R_1^{\frac{1}{2}}\widehat{G}_f + \widehat{U}R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\widehat{V}.$$

By Lemma 4.3.3, $\widehat{U}$ and $\widehat{G}_c$ are orthogonal. Thus

$$\begin{aligned} \left\|\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right)\right\|_2^2 &= \|\widehat{G}_c - \widehat{U}(R_1^{\frac{1}{2}}\widehat{G}_f - R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\widehat{V})\|_2^2 \\ &= \|\widehat{G}_c\|_2^2 + \|\widehat{U}(R_1^{\frac{1}{2}}\widehat{G}_f - R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\widehat{V})\|_2^2. \end{aligned}$$

Since $\widehat{U}$ is inner, left multiplying by $\widehat{U}$ preserves the $\mathcal{H}_2$-norm, hence

$$\begin{aligned} \left\|\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right)\right\|_2^2 &= \|\widehat{G}_c\|_2^2 + \|\widehat{U}(R_1^{\frac{1}{2}}\widehat{G}_f - R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\widehat{V})\|_2^2 \\ &= \|\widehat{G}_c\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f - R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\widehat{V}\|_2^2. \end{aligned}$$

By Lemma 4.3.3, $\widehat{V}^\intercal$ and $\widehat{G}_f^\intercal$ are also orthogonal, thus

$$\begin{aligned} \left\|\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right)\right\|_2^2 &= \|\widehat{G}_c\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f - R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\widehat{V}\|_2^2 \\ &= \|\widehat{G}_c\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\widehat{V}\|_2^2. \end{aligned}$$

But $\widehat{V}$ is co-inner, so right multiplying by $\widehat{V}$ also preserves the norm, thus

$$\begin{aligned} \left\|\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right)\right\|_2^2 &= \|\widehat{G}_c\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\widehat{V}\|_2^2 \\ &= \|\widehat{G}_c\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{R}R_2^{\frac{1}{2}}\|_2^2. \end{aligned}$$

Clearly this norm is minimized if the last term is zero, and this attained if $\widehat{R} = 0$. Thus, applying (4.6), the optimal controller is given by

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}^\circ, 0) = \left[\begin{array}{cc|c} A + B_2F + LC_2 & 0 & -L + 0 \\ 0 & 0 & 0 \\ \hline F + 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} A + B_2F + LC_2 & -L \\ \hline F & 0 \end{array}\right],$$

which proves the theorem. ☐

Procedure 4.3.5 (Summary of solution strategy for the general $\mathcal{H}_2$-control problem)**.**
For a system

$$\widehat{G} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right]$$

1. *(A, B_1, C_2) and (A, B_2, C_1) stabilizable and detectable;*
2. *$R_1 = D_{12}^\intercal D_{12} > 0$ and $R_2 = D_{21} D_{21}^\intercal > 0$;*
3. *$\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ and $\begin{bmatrix} A - i\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ full column and row rank respectively.*

Step 1: *Find solutions $X = \mathrm{Ric}(A, B_2, C_1, D_{12})$ and $Y = \mathrm{Ric}(A^\intercal, C_2^\intercal, B_1^\intercal, D_{21}^\intercal)$.*
Step 2: *Set $F = -R_1^{-1}(B_2^\intercal X + D_{12}^\intercal C_1)$ and $L = -(Y C_2^\intercal + B_1 D_{21}^\intercal) R_2^{-1}$.*
Step 3: *Youla parametrization: set*

$$\widehat{J}^\circ := \left[\begin{array}{c|cc} A + B_2F + LC_2 & -L & B_2 \\ \hline F & 0 & I \\ -C_2 & I & 0 \end{array}\right]. \tag{4.19}$$

All stabilizing controllers of $\widehat{G}$ are given by

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}^\circ, \widehat{R})$$

with $\widehat{R} \in \mathcal{RH}_2$ arbitrary. Furthermore,

$$\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \widehat{H}_{11}^\circ + \widehat{H}_{12}^\circ \widehat{R} \widehat{H}_{21}^\circ,$$

where

$$\widehat{H}^\circ = \widehat{G} \star \widehat{J}^\circ = \left[\begin{array}{cc|cc} A + B_2F & -B_2F & B_1 & B_2 \\ 0 & A + LC_2 & B_1 + LD_{21} & 0 \\ \hline C_1 + D_{12}F & -D_{12}F & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right].$$

Step 4: *(Lemmas 4.2.4 and 4.3.3) It follows that*

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \widehat{G}_c - \widehat{U} R_1^{\frac{1}{2}} \widehat{G}_f + \widehat{U} R_1^{\frac{1}{2}} \widehat{R} R_2^{\frac{1}{2}} \widehat{V},$$

where U and V are inner and co-inner functions respectively, specifically

$$\widehat{G}_c = \left[\begin{array}{c|c} A + B_2F & B_1 \\ \hline C_1 + D_{12}F & 0 \end{array}\right], \qquad \widehat{G}_f = \left[\begin{array}{c|c} A + LC_2 & B_1 + LD_{21} \\ \hline F & 0 \end{array}\right],$$

$$\widehat{U} = \left[\begin{array}{c|c} A + B_2F & B_2 R_1^{-\frac{1}{2}} \\ \hline C_1 + D_{12}F & D_{12} R_1^{-\frac{1}{2}} \end{array}\right] \quad \text{and} \quad \widehat{V} = \left[\begin{array}{c|c} A + LC_2 & B_1 + LD_{21} \\ \hline R_2^{-\frac{1}{2}} C_2 & R_2^{-\frac{1}{2}} D_{21} \end{array}\right].$$

Step 5: *It follows (via Lemma 4.3.3) that*

$$\left\| \underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) \right\|_2^2 = \|\widehat{G}_c\|_2^2 + \|R_1^{\frac{1}{2}} \widehat{G}_f\|_2^2 + \|R_1^{\frac{1}{2}} \widehat{R} R_2^{\frac{1}{2}}\|_2^2.$$

Step 6: *Clearly the minimum occurs when $\widehat{R} = 0$. Thus the minimizing controller is given by*

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}^\circ, 0) = \left[\begin{array}{c|c} A + B_2F + LC_2 & -L \\ \hline F & 0 \end{array}\right].$$

4.4 The State Feedback Case

In this section, we consider the special case of the $\mathcal{H}_2$-control problem in Definition 4.3.1 where the controller K has direct access to the state of the plant G. This means that $C_2 = I$ and $D_{21} = 0$ in (4.18), that is,

$$\widehat{G} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C & 0 & D \\ I & 0 & 0 \end{array}\right] = \begin{bmatrix} C\Phi_A & C\Phi_A B + D \\ \Phi_A & \Phi_A \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix} \quad \text{and} \quad x(0) = x_0. \tag{4.20}$$

Note that we cannot apply the solution procedure at the end of the previous section to the state feedback case, because $D_{21} = 0$ and hence the system (4.20) does not satisfy the conditions required for that solution procedure. The following standard result provides sufficient conditions for the state feedback case of the $\mathcal{H}_2$-control problem to have a unique solution and provides the optimal solution.

Theorem 4.4.1 (cf. Theorem 13.3 in [56]).
Consider a system with state feedback as in (4.20). If

1. $D^\intercal D > 0$;
2. (A, B) *is stabilizable;*
3. $\begin{bmatrix} A - i\omega I & B \\ C & D \end{bmatrix}$ *has full column rank for all* $\omega \in \mathbb{R}$,

then there exists a unique optimal controller

$$\widehat{K} = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & F \end{array}\right] = F, \quad \textit{where} \quad F = -(D^\intercal D)^{-1}(B_2^\intercal X + D^\intercal C) \quad \textit{and} \quad X = \mathrm{Ric}(A, B, C, D)$$

with minimum value $\|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2^2 = \mathrm{trace}(B_1^\intercal X B_1)$.

Hence in the state feedback case, the optimal controller is a static state feedback controller. Note that, as in the general output feedback case, the three conditions in the above theorem correspond with the conditions in Theorem 3.7.8 and ensure that $X = \mathrm{Ric}(A, B, C, D)$ exists.

In the next example we consider the LQ regulator problem (see for example Section 13.2 in [56]) and show how it can be rephrased as an $\mathcal{H}_2$-control problem with state feedback.

Example 4.4.2 (LQ optimal control).
The linear quadratic (LQ for short) optimal control problem was formulated and solved by Rudolf Kalman in [17] and is now well known, see for example Chapter 10 in [49]. The LQ control problem considers a linear system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \qquad x(0) = x_0 \\ z(t) &= Cx(t) + Du(t), \qquad\quad t \geq 0. \end{aligned} \tag{4.21}$$

Set

$$Q = C^\intercal C, \quad R = D^\intercal D, \quad S = D^\intercal C.$$

The aim of LQ-control is to keep all components of the output $z(t)$ as small as possible by finding an input u such that the quadratic cost functional

$$J(x_0, u) := \int_0^\infty x(t)^\intercal Q x(t) + u(t)^\intercal R u(t) + 2u(t)^\intercal S x(t) \, \mathrm{d}t \tag{4.22}$$

is minimized. The solution of the LQ control problem is given by the following standard result (see for example Theorem 10.19 in [49]).

For a linear system (4.21), if $D^\intercal D > 0$ and the CARE

$$A^\intercal X + XA + Q - (XB + S^\intercal)R^{-1}(B^\intercal X + S) = 0,$$

has a unique stabilizing solution X, then

$$u^*(t) = Fx(t) \qquad \text{where} \qquad -F = R^{-1}(B^\intercal X + S)$$

minimizes the quadratic cost functional (4.22) and the minimum value of the cost is

$$J^*(x_0) = \inf_{u \in U} J(x_0, u) = x_0^\intercal X x_0.$$

We now show that it is possible to obtain the same result using an $\mathcal{H}_2$-control-setup and solution strategy. By introducing the initial state as an instantaneous disturbance input, the state equation in the plant (4.21) can be rewritten as

$$\dot{x}(t) = Ax(t) + x_0\delta_0(t) + Bu(t),$$

where

$$\delta_0(t) = \begin{cases} 1 & t = 0 \\ 0 & t \neq 0 \end{cases}.$$

Hence, the initial state can be viewed as a momentary disturbance input that occurs at time $t = 0$. As a consequence, the LQ plant (4.21) can be considered to be a special case of $\mathcal{H}_2$-plant in (4.3) with $\omega(t) = \delta_0(t)$ and with

$$B_1 = x_0, \quad B_2 = B, \quad C_1 = C, \quad C_2 = I, \quad D_{11} = 0, \quad D_{12} = D, \quad D_{21} = 0 \quad \text{and} \quad D_{22} = 0,$$

that is

$$\widehat{G} = \left[\begin{array}{c|cc} A & x_0 & B \\ \hline C & 0 & D \\ I & 0 & 0 \end{array}\right] = \begin{bmatrix} C\Phi_A x_0 & C\Phi_A B + D \\ \Phi_A x_0 & \Phi_A B \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix}.$$

This is clearly a special case of the state-feedback $\mathcal{H}_2$-control problem. The optimal solution is given by Theorem 4.4.1:

$$\widehat{K} = F, \quad \text{where} \quad F = -(D^\intercal D)^{-1}(B^\intercal X + D^\intercal C) \quad \text{and} \quad X = \mathrm{Ric}(A, B, C, D).$$

By Proposition 4.1.2,

$$\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \left[\begin{array}{cc|c} A + BF & 0 & x_0 \\ 0 & 0 & 0 \\ \hline C + DF & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} A_F & x_0 \\ \hline C_F & 0 \end{array}\right],$$

where $A_F = A + BF$ and $C_F = C + DF$. By Lemma 3.7.10, $X = X_o$, where X_o is the observability gramian of the pair (C_F, A_F). Hence by Lemma 3.6.2, the optimal closed loop performance is given by

$$\|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2^2 = \mathrm{trace}\,(x_0^\intercal X x_0) = x_0^\intercal X x_0, \quad \text{where} \quad X = \mathrm{Ric}(A, B, C, D).$$

4.5 Spectral Factorization Approach to State Feedback Case

We now summarize the spectral factorization approach to the state feedback case. This approach can be found in [46]. Some of the results in this section will be stated and proved in the more general setting of output feedback in the Section 4.7.

The first step will be to do a reparametrization of the closed loop transfer function $\underline{\mathcal{F}}(\widehat{G},\widehat{K})$. For this, it is required that B_1 is square and invertible. We will assume that $B_1 = I$ (and rename $B_2 = B$) for the sake of simplicity, that is,

$$\widehat{G} = \left[\begin{array}{c|cc} A & I & B \\ \hline C & 0 & D \\ I & 0 & 0 \end{array}\right] = \begin{bmatrix} C\Phi_A & C\Phi_A B + D \\ \Phi_A & \Phi_A B \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix}. \tag{4.23}$$

Furthermore it is assumed that $\widehat{G}$ satisfies the conditions in Theorem 4.4.1, that is

1. $D^\intercal D > 0$;
2. (A,B) is stabilizable;
3. $\begin{bmatrix} A - i\omega I & B \\ C & D \end{bmatrix}$ has full column rank for all $\omega \in \mathbb{R}$.

These conditions guarantee the existence of a solution $X = \mathrm{Ric}(A,B,C,D)$ (see Theorem 3.7.8). By the stabilizability of (A,B), there exists a matrix E such that $A+BE$ is stable. Furthermore, Lemma 3.7.10 guarantees that $X = X_E = \mathrm{Ric}(A+BE, B, C+DE, D)$. By an adaptation of the Youla parametrization (Theorem 4.2.3), we get the following parametrization of stabilizing controllers $\widehat{K}$ of $\widehat{G}$. Set

$$\widehat{J} = \left[\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & E & I \\ 0 & I & 0 \end{array}\right].$$

Then

$$\widehat{H} := \widehat{G} \star \widehat{J} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A+BE & I & B \\ \hline C+DE & 0 & D \\ I & 0 & 0 \end{array}\right]$$

is a stable system and the stabilizing controllers $\widehat{K}$ of $\widehat{G}$ are parametrized as

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J},\widehat{S}) = \widehat{S} + E \quad \text{where} \quad \widehat{S} \in \mathcal{RH}_\infty.$$

Furthermore, the closed loop transfer function can be written as

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \underline{\mathcal{F}}(\widehat{H},\widehat{S}) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{S}(I - \widehat{H}_{22}\widehat{S})^{-1}\widehat{H}_{21}.$$

Since $\widehat{H}_{22}$ is strictly proper, $I - \widehat{H}_{22}\widehat{S}$ is invertible by Proposition 3.5.3. Hence if we define

$$\widehat{R} = \widehat{S}(I - \widehat{H}_{22}\widehat{S})^{-1},$$

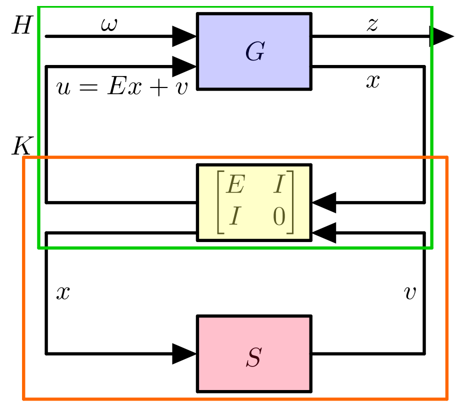
then $I + \widehat{H}_{22}\widehat{R}$ is invertible by Proposition 3.5.3 and $\widehat{S}$ is uniquely determined by $\widehat{R}$ via

$$\widehat{S} = \widehat{R}(I + \widehat{H}_{22}\widehat{R})^{-1}$$

and hence the closed loop transfer function can be written as

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}.$$

This parametrization is illustrated by the following figure:



This gives the following result.

Lemma 4.5.1 (cf. Lemma 9 in [46]).
If (A, B) is stabilizable and E is a matrix such that $A + BE$ is stable, then the set of stabilizing controllers of the plant (4.23) is given by

$$\{\widehat{K} = \widehat{R}(I + \widehat{H}_{22}\widehat{R})^{-1} + E : \widehat{R} \in \mathcal{RH}_\infty\}$$

and the closed loop transfer function is given by

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}, \qquad where \qquad \widehat{H} = \begin{bmatrix} \widehat{H}_{11} & \widehat{H}_{12} \\ \widehat{H}_{21} & \widehat{H}_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A+BE & I & B \\ \hline C+DE & 0 & D \\ I & 0 & 0 \end{array}\right]. \tag{4.24}$$

Now $\widehat{H}_{21} = \Phi_{A+BE}$ in (4.23) is square and invertible. Hence, if we define $\widehat{Q}$ by $\widehat{Q} = \widehat{R}\widehat{H}_{21}$, then the closed loop transfer function $\underline{\mathcal{F}}(\widehat{G}, \widehat{K})$, can be re-written as $\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}$ and clearly $\widehat{Q}$ and $\widehat{R}$ determine each other uniquely, indeed $\widehat{R} = \widehat{Q}\widehat{H}_{21}^{-1}$. Hence, a reparameterized problem of finding a controller $\widehat{Q}$ that minimizes $\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2$ may be considered. However, $\widehat{R}$ and hence $\widehat{K}$ may be improper when recovered from $\widehat{Q}$ via

$$\widehat{K} = \widehat{R}(I + \widehat{H}_{22}\widehat{R})^{-1} + E = \widehat{Q}\widehat{H}_{21}^{-1}(I + \widehat{H}_{22}\widehat{Q}\widehat{H}_{21}^{-1})^{-1} + E \tag{4.25}$$

due to the $\widehat{H}_{21}^{-1} = \lambda I - (A + BE)$ term. Hence, for the stable system (4.24), we seek to synthesize a controller $Q \sim (A_Q, B_Q, C_Q, 0)$ satisfying the following properties

1. $\widehat{R} = \widehat{Q}\widehat{H}_{21}^{-1}$ is proper;
2. A_Q is stable;

such that the closed loop performance $\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2$, is minimized.

Note that $\widehat{H}_{11}, \widehat{H}_{12}, \widehat{H}_{21}, \widehat{H}_{22} \in \mathcal{RH}_\infty$, because $A + BE$ is stable. From the projection theorem (see for example Section 3.3 in [27]), the following equivalent condition for $\widehat{Q}$ to be a minimizer of

$$\min_{\widehat{Q} \in \mathcal{RH}_2} \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 \tag{4.26}$$

can be derived (see for example Lemma 11 in [46]): if $\widehat{H}_{11} \in \mathcal{RH}_2$ and $\widehat{H}_{12} \in \mathcal{RH}_\infty$, then $\widehat{Q} \in \mathcal{RH}_2$ minimizes (4.26) if and only if

$$\widehat{H}_{12}^*\widehat{H}_{11} + \widehat{H}_{12}^*\widehat{H}_{12}\widehat{Q} \in \mathcal{H}_2^\perp. \tag{4.27}$$

The standard spectral factorization theorem (see for example Theorem 13.6 in [1]) gives a factorization $\widehat{H}_{12}^{*}\widehat{H}_{12} = \widehat{L}^{*}\widehat{L}$ where $\widehat{L} \in \mathcal{RH}_\infty$ is square and invertible and $\widehat{L}^{-1} \in \mathcal{RH}_\infty$ (and hence $\widehat{L}^{-*} \in \mathcal{RH}_\infty^{-}$). Here $\widehat{L}$ is given by

$$\widehat{L} = \left[\begin{array}{c|c} A+BE & B \\ \hline -(D^\mathsf{T}D)^{\frac{1}{2}}F_E & (D^\mathsf{T}D)^{\frac{1}{2}} \end{array}\right] \quad \text{with} \quad \begin{aligned} F_E &= -(D^\mathsf{T}D)^{-1}(B^\mathsf{T}X_E + D^\mathsf{T}(C+DE)) \\ X_E &= \mathrm{Ric}(A+BE, B, C+DE, D), \end{aligned} \tag{4.28}$$

where we note that X_E exists under the conditions on $\widehat{G}$ (see Theorem 3.7.8 and Lemma 3.7.10). But by Lemma 3.7.10,

$$X_E = X = \mathrm{Ric}(A, B, C, D).$$

Thus

$$\begin{aligned} F_E &:= -(D^\mathsf{T}D)^{-1}(B^\mathsf{T}X_E + D^\mathsf{T}C_E) = -(D^\mathsf{T}D)^{-1}(B^\mathsf{T}X + D^\mathsf{T}(C+DE)) = F - E \quad \text{where} \\ F &= -(D^\mathsf{T}D)^{-1}(B^\mathsf{T}X + D^\mathsf{T}C) \end{aligned}$$

and so

$$\widehat{L} = \left[\begin{array}{c|c} A+BE & B \\ \hline (D^\mathsf{T}D)^{\frac{1}{2}}(E-F) & (D^\mathsf{T}D)^{\frac{1}{2}} \end{array}\right].$$

Furthermore, it can be shown that (see for example Lemma 13 in [46])

$$\widehat{L}^{-*}\widehat{H}_{12}^{*}\widehat{H}_{11} = (D^\mathsf{T}D)^{\frac{1}{2}}(E-F)\widehat{H}_{21} - (D^\mathsf{T}D)^{-\frac{1}{2}}B^\mathsf{T}\Phi_{-(A+BF)^\mathsf{T}}X. \tag{4.29}$$

In the above equation, the first term on the right hand side is in $\mathcal{RH}_2$ and the second term is in $\mathcal{RH}_2^{\perp}$. By condition (4.27) and the factorization $\widehat{H}_{12}^{*}\widehat{H}_{12} = \widehat{L}^{*}\widehat{L}$, the controller $\widehat{Q}$ is optimal if and only if

$$\widehat{H}_{12}^{*}\widehat{H}_{11} + \widehat{L}^{*}\widehat{L}\widehat{Q} \in \mathcal{H}_2^{\perp}.$$

Since $\widehat{L}^{-*} \in \mathcal{RH}_\infty^{-}$, this is equivalent to

$$\widehat{L}^{-*}\widehat{H}_{12}^{*}\widehat{H}_{11} + \widehat{L}\widehat{Q} \in \mathcal{H}_2^{\perp}.$$

Implementing (4.29) then gives the condition

$$(D^\mathsf{T}D)^{\frac{1}{2}}(E-F)\widehat{H}_{21} - (D^\mathsf{T}D)^{-\frac{1}{2}}B^\mathsf{T}\Phi_{-(A+BF)^\mathsf{T}}X + \widehat{L}\widehat{Q} \in \mathcal{H}_2^{\perp}.$$

Now $\widehat{Q} \in \mathcal{RH}_2$ and $\widehat{L} \in \mathcal{HR}_\infty$, thus $\widehat{L}\widehat{Q} \in \mathcal{RH}_2$ by Proposition 3.6.1. Since the terms $(D^\mathsf{T}D)^{\frac{1}{2}}(E-F)\widehat{H}_{21}$ and $\widehat{L}\widehat{Q}$ are in $\mathcal{RH}_2$ and the term $(D^\mathsf{T}D)^{-\frac{1}{2}}B^\mathsf{T}\Phi_{-(A+BF)^\mathsf{T}}X$ is in $\mathcal{RH}_2^{\perp}$, projecting onto $\mathcal{RH}_2$, gives the following optimality condition

$$(D^\mathsf{T}D)^{\frac{1}{2}}(E-F)\widehat{H}_{21} + \widehat{L}\widehat{Q} = 0.$$

Multiplying by $\widehat{L}^{-1}$ and rearranging then gives

$$\widehat{Q} = -\widehat{L}^{-1}(D^\mathsf{T}D)^{\frac{1}{2}}(E-F)\widehat{H}_{21}.$$

Now by (3.14),

$$\widehat{L}^{-1} = \left[\begin{array}{c|c} A+BF & B(D^\mathsf{T}D)^{-\frac{1}{2}} \\ \hline F-E & (D^\mathsf{T}D)^{-\frac{1}{2}} \end{array}\right] = (F-E)\Phi_{A+BF}B(D^\mathsf{T}D)^{-\frac{1}{2}} + (D^\mathsf{T}D)^{-\frac{1}{2}}.$$

This enables us to determine $\widehat{Q}$ as follows:

$$\begin{aligned}\widehat{Q} &= -\widehat{L}^{-1}(D^\intercal D)^{\frac{1}{2}}(E-F)\widehat{H}_{21} \\ &= -[(F-E)\Phi_{A+BF}B(D^\intercal D)^{-\frac{1}{2}} + (D^\intercal D)^{-\frac{1}{2}}](D^\intercal D)^{\frac{1}{2}}(E-F)\Phi_{A+BE} \\ &= -(F-E)\Phi_{A+BF}B(E-F)\Phi_{A+BE} - (E-F)\Phi_{A+BE}.\end{aligned}$$

Since $(A+BE)-(A+BF) = B(E-F)$, it follows from Lemma 3.5.1 with $X = BE-BF$ and $A = A+BF$ that

$$\Phi_{A+BF}B(E-F)\Phi_{A+BE} = \Phi_{A+BE} - \Phi_{A+BF}.$$

Thus

$$\begin{aligned}\widehat{Q} &= -(F-E)\Phi_{A+BF}B(E-F)\Phi_{A+BE} - (E-F)\Phi_{A+BE} \\ &= (E-F)[\Phi_{A+BE} - \Phi_{A+BF}] - (E-F)\Phi_{A+BE} \\ &= (F-E)\Phi_{A+BF},\end{aligned}$$

that is

$$\widehat{Q} = \left[\begin{array}{c|c} A+BF & I \\ \hline F-E & 0 \end{array}\right]. \tag{4.30}$$

Here $A+BF$ is stable, so it satisfies the second requirement. To show that $\widehat{Q}$ also satisfies the first requirement, note that with $X = BF-BE$ and $A = A+BE$ in Lemma 3.5.1, it follows that

$$BF - BE = \Phi^{-1}_{A+BE} - \Phi^{-1}_{A+BF}, \quad \text{so} \quad \widehat{H}^{-1}_{21} = \Phi^{-1}_{A+BE} = \Phi^{-1}_{A+BF} + BF - BE.$$

We use this fact to determine $\widehat{R} = \widehat{Q}\widehat{H}^{-1}_{21}$:

$$\begin{aligned}\widehat{R} &= (F-E)\Phi_{A+BF}\Phi^{-1}_{A+BE} \\ &= (F-E)\Phi_{A+BF}(\Phi^{-1}_{A+BF} + B(F-E)) \\ &= (F-E) + (F-E)\Phi_{A+BF}B(F-E),\end{aligned}$$

that is

$$\widehat{R} = \left[\begin{array}{c|c} A+BF & B(F-E) \\ \hline F-E & F-E \end{array}\right].$$

In particular, $\widehat{R}$ is proper, so $\widehat{Q}$ satisfies the first requirement. Since $\widehat{Q}$ in (4.30) satisfies both conditions of the reparameterized problem, $\widehat{Q}$ is an optimal solution to the problem (4.26). From this we can recover the optimal controller $\widehat{K}$ for the original state feedback control problem via equation (4.25):

$$\begin{aligned}\widehat{K} &= \widehat{Q}\widehat{H}^{-1}_{21}(I + \widehat{H}_{22}\widehat{Q}\widehat{H}^{-1}_{21})^{-1} + E \\ &= (F-E)\Phi_{A+BF}\Phi^{-1}_{A+BE}(I + \Phi_{A+BE}(BF-BE)\Phi_{A+BF}\Phi^{-1}_{A+BE})^{-1} + E \\ &= (F-E)\Phi_{A+BF}\Phi^{-1}_{A+BE}(I + (\Phi_{A+BF} - \Phi_{A+BE})\Phi^{-1}_{A+BE})^{-1} + E \\ &= (F-E)\Phi_{A+BF}\Phi^{-1}_{A+BE}(\Phi_{A+BF}\Phi^{-1}_{A+BE})^{-1} + E \\ &= F - E + E \\ &= F,\end{aligned}$$

which agrees with the standard result on the state feedback case in Theorem 4.4.1.

In order to calculate the value of the norm

$$\|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2 = \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2,$$

we compute

$$\widehat{H}_{12}\widehat{Q} = \left[\begin{array}{c|c} A+BE & B \\ \hline C+DE & D \end{array}\right] \left[\begin{array}{c|c} A+BF & I \\ \hline F-E & 0 \end{array}\right] = \left[\begin{array}{cc|c} A+BE & BF-BE & 0 \\ 0 & A+BF & I \\ \hline C+DE & DF-DE & 0 \end{array}\right].$$

With $X = I$ in (3.19), it follows by (3.13) that

$$\begin{aligned} \widehat{H}_{12}\widehat{Q} &= \left[\begin{array}{cc|c} A+BE & BF-BE & 0 \\ 0 & A+BF & I \\ \hline C+DE & DF-DE & 0 \end{array}\right] \\ &= \left[\begin{array}{c|c} A+BE & -I \\ \hline C+DE & 0 \end{array}\right] + \left[\begin{array}{c|c} A+BF & I \\ \hline C+DF & 0 \end{array}\right] = -\widehat{H}_{11} + \left[\begin{array}{c|c} A+BF & I \\ \hline C+DF & 0 \end{array}\right]. \end{aligned}$$

Thus

$$\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q} = \left[\begin{array}{c|c} A+BF & I \\ \hline C+DF & 0 \end{array}\right], \quad \text{where} \quad F = -(D^{\intercal}D)^{-1}(B^{\intercal}X + D^{\intercal}C) \quad \text{and} \quad X = \mathrm{Ric}(A,B,C,D).$$

Hence, by Lemma 3.6.2,

$$\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 = \left\| \left[\begin{array}{c|c} A+BF & I \\ \hline C+DF & 0 \end{array}\right] \right\|_2^2 = \mathrm{trace}\,(X_o) = \mathrm{trace}\,(X),$$

where X_o is the controllability gramian of the pair $(A+BF, C+DF)$, but by Lemma 3.7.10, $X_o = X$, so it follows that the result agrees completely with the standard result in Theorem 4.4.1 with $B_1 = I$.

Procedure 4.5.2 (Summary of solution strategy using spectral factorization for the state feedback case)**.** *For a system*

$$\widehat{G} = \left[\begin{array}{c|cc} A & I & B \\ \hline C & 0 & D \\ I & 0 & 0 \end{array}\right] \quad \textit{and} \quad x(0) = x_0$$

1. *(A, B) is stabilizable;*
2. *$R = D^{\intercal}D > 0$;*
3. *$\begin{bmatrix} A - i\omega I & B \\ C & D \end{bmatrix}$ has full column rank.*

Step 1: *Find a matrix E such that $A + BE$ is Hurwitz.*
Step 2: *Youla parametrization (Lemma 4.5.1): set*

$$\widehat{J} := \left[\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & E & I \\ 0 & I & 0 \end{array}\right]. \tag{4.31}$$

Then all stabilizing controllers of $\widehat{G}$ are given by

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J},\widehat{R}) = \widehat{R}(I + \widehat{H}_{22}\widehat{R})^{-1} + E$$

with $\widehat{R} \in \mathcal{RH}_\infty$ arbitrary and

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21},$$

where

$$\widehat{H} = \widehat{G} \star \widehat{J} = \left[\begin{array}{c|cc} A+BE & I & B \\ \hline C+DE & 0 & D \\ I & 0 & 0 \end{array}\right]. \tag{4.32}$$

Step 3: *Set* $\widehat{Q} = \widehat{R}\widehat{H}_{21}$. *Then*

$$\min_{\widehat{K}\in\mathcal{RH}_2} \|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2 = \min_{\widehat{Q}\in\mathcal{RH}_2} \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2, \quad \textit{with} \quad \widehat{K} = \widehat{Q}\widehat{H}_{21}^{-1}(I + \widehat{H}_{22}\widehat{Q}\widehat{H}_{21}^{-1})^{-1} + E.$$

Step 4: *By the projection theorem,*

$$\widehat{Q} \in \mathcal{RH}_2 \textit{ minimizes } \quad \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 \quad \textit{if and only if} \quad \widehat{H}_{12}^*\widehat{H}_{11} + \widehat{H}_{12}^*\widehat{H}_{12}\widehat{Q} \in \mathcal{H}_2^\perp.$$

Step 5: *Using the spectral factorization theorem, construct* $\widehat{L} \in \mathcal{RH}_\infty$ *for which both* $\widehat{L}^{-1} \in \mathcal{RH}_\infty$ *and* $\widehat{L}^*\widehat{L} = \widehat{H}_{12}^*\widehat{H}_{12}$*:*

$$\widehat{L} = \left[\begin{array}{c|c} A+BE & B \\ \hline R^{\frac{1}{2}}(E-F) & R^{\frac{1}{2}} \end{array}\right], \quad \textit{where} \quad X = \mathrm{Ric}(A,B,C,D) \quad \textit{and} \quad F = -(D^\intercal D)^{-1}(B^\intercal X + D^\intercal C).$$

It then follows that

$$\min_{\widehat{Q}\in\mathcal{RH}_2} \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 \quad \textit{if and only if} \quad \widehat{L}^{-*}\widehat{H}_{12}^*\widehat{H}_{11} + \widehat{L}\widehat{Q} \in \mathcal{H}_2^\perp.$$

Step 6: *Projecting onto* $\mathcal{RH}_2$ *gives that*

$$\min_{\widehat{Q}\in\mathcal{RH}_2} \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 \quad \textit{if and only if} \quad (D^\intercal D)^{\frac{1}{2}}(E-F)\widehat{H}_{21} + \widehat{L}\widehat{Q} = 0.$$

Step 7: *Solving for* $\widehat{Q}$ *gives*

$$\widehat{Q} = \left[\begin{array}{c|c} A+BF & I \\ \hline F-E & 0 \end{array}\right].$$

Step 8: *Solving for* $\widehat{K}$ *via* $\widehat{K} = \widehat{Q}\widehat{H}_{21}^{-1}(I + \widehat{H}_{22}\widehat{Q}\widehat{H}_{21}^{-1})^{-1} + E$ *gives* $\widehat{K} = F = -(D^\intercal D)^{-1}(B^\intercal X + D^\intercal C)$.

We note that if we choose $E = F$ in the above solution strategy, then the optimal solution is given by $\widehat{Q} = 0$. So as in the solution strategy of Section 4.3, the optimal controller is obtained when $\widehat{Q} = 0$. In this case, we again obtain the static optimal feedback controller via

$$\widehat{K} = \widehat{Q}\widehat{H}_{21}^{-1}(I + \widehat{H}_{22}\widehat{Q}\widehat{H}_{21}^{-1})^{-1} + F = 0 + F.$$

We also note that with $E = F$, it follows by Lemma 3.6.2, that

$$\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 = \|\widehat{H}_{11}\|_2^2 = \left\| \left[\begin{array}{c|c} A+BF & I \\ \hline C+DF & 0 \end{array}\right] \right\|_2^2 = \mathrm{trace}(X_o),$$

where X_o is the observability gramian of the pair $(A+BF, C+DF)$. But by Lemma 3.7.9, we have $X_o = X = \mathrm{Ric}(A,B,C,D)$, thus

$$\min_{\widehat{K}\in\mathcal{RH}_2} \|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2 = \mathrm{trace}(X).$$

4.6 Alternative Approach to the Output Feedback Case

In section 4.3 the solution to the general output feedback case of the $\mathcal{H}_2$-control problem was obtained in Theorem 4.3.4. Following Procedure 4.3.5, we note that the solution to Riccati equations play an important role in the Youla parametrization. Specifically the stabilizing matrices F and L in (4.19) depend on the solutions of Riccati equations. However, in Chapter 7, we consider the $\mathcal{H}_2$-problem for structured linear systems. For matrices (A, B, C, D) with a specified zero structure, the solution $X = \mathrm{Ric}(A, B, C, D)$ of the associated Riccati equation, if it exists, does not in general have the same zero structure. As such, we consider an alternative approach similar to Procedure 4.3.5, but with arbitrary stabilizing matrices E and M in the Youla parametrization that do not depend on solutions of Riccati equations. This will allow us in Chapter 7 to utilize stabilizing matrices E and M in the Youla parametrization that preserve the structure of the system.

Consider a system as in Theorem 4.3.4, that is

$$\widehat{G} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] = \begin{bmatrix} C_1\Phi_A B_1 & C_1\Phi_A B_2 + D_{12} \\ C_2\Phi_A B_1 + D_{21} & C_2\Phi_A B_2 \end{bmatrix}.$$

As in Theorem 4.3.4, the following assumptions are made about the plant:

1. the triples (A, B_1, C_2) and (A, B_2, C_1) are stabilizable and detectable;
2. $R_1 = D_{12}^{\intercal} D_{12} > 0$ and $R_2 = D_{21} D_{21}^{\intercal} > 0$;
3. $\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ and $\begin{bmatrix} A - i\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ have full column and row rank respectively for all $\omega \in \mathbb{R}$.

By the stabilizability and detectability of the triples (A, B_1, C_2) and (A, B_2, C_1), there exist matrices E and M such that $A + B_2 E$ and $A + MC_2$ are stable. By the Youla Parametrization (Theorem 4.2.3), all stabilizing controllers $\widehat{K}$ of $\widehat{G}$ are of the form

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}, \widehat{R}) = \widehat{J}_{11} + \widehat{J}_{12}\widehat{K}(I - \widehat{J}_{22}\widehat{R})^{-1}\widehat{J}_{21} \qquad \text{for some} \qquad \widehat{R} \in \mathcal{RH}_\infty.$$

Here $\widehat{J}$ is the canonical system

$$\widehat{J} := \left[\begin{array}{c|cc} A + B_2 E + MC_2 & -M & B_2 \\ \hline E & 0 & I \\ -C_2 & I & 0 \end{array}\right]. \tag{4.33}$$

Furthermore, after connecting such a stabilizing controller $\widehat{K}$, the closed loop transfer function is given by

$$\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21},$$

where $\widehat{H} = \widehat{G} \star \widehat{J}$ is the stable system with realization

$$\widehat{H} = \begin{bmatrix} \widehat{H}_{11} & \widehat{H}_{12} \\ \widehat{H}_{21} & \widehat{H}_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} A + B_2 E & -B_2 E & B_1 & B_2 \\ 0 & A + MC_2 & B_1 + MD_{21} & 0 \\ \hline C_1 + D_{12} E & -D_{12} E & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right] = \left[\begin{array}{c|cc} \widetilde{A} & \widetilde{B}_1 & \widetilde{B}_2 \\ \hline \widetilde{C}_1 & 0 & D_{12} \\ \widetilde{C}_2 & D_{21} & 0 \end{array}\right]. \tag{4.34}$$

Here

$$\widetilde{A} = \begin{bmatrix} A + B_2 E & -B_2 E \\ 0 & A + MC_2 \end{bmatrix}, \qquad \widetilde{B}_1 = \begin{bmatrix} B_1 \\ B_1 + MD_{21} \end{bmatrix}, \qquad \widetilde{B}_2 = \begin{bmatrix} B_2 \\ 0 \end{bmatrix}$$

$$\widetilde{C}_1 = \begin{bmatrix} C_1 + D_{12} E & -D_{12} E \end{bmatrix}, \qquad \widetilde{C}_2 = \begin{bmatrix} 0 & C_2 \end{bmatrix}.$$

We note that since $A+B_2E$ and $A+MC_2$ are stable, we have

$$\begin{aligned}
\widehat{H}_{11} &= \left[\begin{array}{cc|c} A+B_2E & -B_2E & B_1 \\ 0 & A+MC_2 & B_1+MD_{21} \\ \hline C_1+D_{12}E & -D_{12}E & 0 \end{array}\right] && \in \mathcal{RH}_2, \\
\widehat{H}_{12} &= \left[\begin{array}{c|c} A+B_2E & B_2 \\ \hline C_1+D_{12}E & D_{12} \end{array}\right] && \in \mathcal{RH}_\infty, \\
\widehat{H}_{21} &= \left[\begin{array}{c|c} A+MC_2 & B_1+MD_{21} \\ \hline C_2 & D_{21} \end{array}\right] && \in \mathcal{RH}_\infty \quad \text{and} \\
\widehat{H}_{22} &= 0.
\end{aligned} \tag{4.35}$$

As in Definition 4.3.1, the goal is to minimize the closed loop performance

$$\|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2 = \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2^2.$$

Since $\widehat{H}_{11} \in \mathcal{RH}_2$, $\widehat{H}_{12} \in \mathcal{RH}_\infty$ and $\widehat{H}_{21} \in \mathcal{RH}_\infty$, it can be seen that $\underline{\mathcal{F}}(\widehat{G},\widehat{K}) \in \mathcal{RH}_2$ if $\widehat{R} \in \mathcal{RH}_2$.

Let $\widehat{H}^\circ$ be as in Procedure 4.3.5, that is,

$$\widehat{H}^\circ = \begin{bmatrix} \widehat{H}^\circ_{11} & \widehat{H}^\circ_{12} \\ \widehat{H}^\circ_{21} & 0 \end{bmatrix} = \left[\begin{array}{cc|cc} A+B_2F & -B_2F & B_1 & B_2 \\ 0 & A+LC_2 & B_1+LD_{21} & 0 \\ \hline C_1+D_{12}F & -D_{12}F & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right], \tag{4.36}$$

where $X = \mathrm{Ric}(A, B_2, C_1, D_{12})$ and $Y = \mathrm{Ric}(A^\intercal, C_2^\intercal, B_1^\intercal, D_{21}^\intercal)$ and

$$F := -(D_{12}^\intercal D_{12})^{-1}(B_2^\intercal X + D_{12}^\intercal C_1) \quad \text{and} \quad L := -(YC_2^\intercal + B_1 D_{21}^\intercal)(D_{21}D_{21}^\intercal)^{-1}. \tag{4.37}$$

We note that $\widehat{H}$ does not contain matrices that depend on solutions of Riccati equations whereas $\widehat{H}^\circ$ does. We now follow the same approach as in Procedure 4.3.5 for $\widehat{H}$ instead of $\widehat{H}^\circ$. Recall that by Lemma 4.3.3,

$$\widehat{H}^\circ_{12} = \widehat{U}R_1^{\frac{1}{2}} \quad \text{and} \quad \widehat{H}^\circ_{21} = R_2^{\frac{1}{2}}\widehat{V}$$

where $\widehat{U}$ and $\widehat{V}$ are inner and co-inner functions respectively. This allows one in step 4 of Procedure 4.3.5, to write the closed loop transfer function in terms of the inner and co-inner functions $\widehat{U}$ and $\widehat{V}$ as

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{H}^\circ_{11} + \widehat{H}^\circ_{12}\widehat{Q}\widehat{H}^\circ_{21} = \widehat{G}_c - \widehat{U}R_1^{\frac{1}{2}}\widehat{G}_f + \widehat{U}R_1^{\frac{1}{2}}\widehat{Q}R_2^{\frac{1}{2}}\widehat{V}.$$

We now obtain an analogous result for $\widehat{H}_{12}$ and $\widehat{H}_{21}$. This is done by applying Theorem 3.7.13 to get inner-outer factorizations of $\widehat{H}_{12}$ and $\widehat{H}_{21}^\intercal$.

Corollary 4.6.1.
For $\widehat{H}_{12}$ and $\widehat{H}_{21}$ as in (4.35), it holds that

$$\widehat{H}_{12} = \widehat{U}\widehat{L} \quad \text{and} \quad \widehat{H}_{21} = \widehat{M}\widehat{V}$$

where $\widehat{U}$ and $\widehat{V}$ are inner and co-inner functions given by

$$\widehat{U} = \left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline C_1+D_{12}F & D_{12} \end{array}\right] R_1^{-\frac{1}{2}} \quad \text{and} \quad \widehat{V} = R_2^{-\frac{1}{2}} \left[\begin{array}{c|c} A+LC_2 & B_1+LD_{21} \\ \hline C_2 & D_{21} \end{array}\right] \tag{4.38}$$

respectively and where $\widehat{L}$ and $\widehat{M}$ are stable invertible outer functions given by

$$\widehat{L} = R_1^{\frac{1}{2}} \left[\begin{array}{c|c} A+B_2E & B_2 \\ \hline E-F & I \end{array}\right] \quad \text{and} \quad \widehat{M} = \left[\begin{array}{c|c} A+MC_2 & M-L \\ \hline C_2 & I \end{array}\right] R_2^{\frac{1}{2}}, \tag{4.39}$$

respectively.

Proof.
We consider the Riccati equations associated with

$$\widehat{H}_{12} = \left[\begin{array}{c|c} A + B_2E & B_2 \\ \hline C_1 + D_{12}E & D_{12} \end{array}\right] \quad \text{and} \quad \widehat{H}_{21}^{\intercal} = \left[\begin{array}{c|c} A^{\intercal} + C_2^{\intercal} M^{\intercal} & C_2^{\intercal} \\ \hline B_1^{\intercal} + D_{21}^{\intercal} M^{\intercal} & D_{21}^{\intercal} \end{array}\right].$$

By Lemma 3.7.10, it follows that

$$\begin{aligned} X &= \operatorname{Ric}(A + B_2E, B_2, C_1 + D_{12}E, D_{12}) = \operatorname{Ric}(A, B_2, C_1, D_{12}) \quad \text{and} \\ Y &= \operatorname{Ric}(A^{\intercal} + C_2^{\intercal} M^{\intercal}, C_2^{\intercal}, B_1^{\intercal} + D_{12}^{\intercal} M^{\intercal}, D_{21}^{\intercal}) = \operatorname{Ric}(A^{\intercal}, C_2^{\intercal}, B_1^{\intercal}, D_{21}^{\intercal}). \end{aligned}$$

Thus with F and L as in (4.37), it holds that

$$\begin{aligned} F_E &:= -(D_{12}^{\intercal} D_{12})^{-1}(B_2^{\intercal} X + D_{12}^{\intercal}(C_1 + D_{12}E)) = F - E \quad \text{and} \\ L_M &:= -(D_{21}D_{21}^{\intercal})^{-1}(C_2Y + D_{21}B_1^{\intercal} + (D_{21}D_{21}^{\intercal})M^{\intercal}) = L^{\intercal} - M^{\intercal}. \end{aligned}$$

Since $A + B_2E$ and $A^{\intercal} + C_2^{\intercal} M^{\intercal}$ are stable and $R_1 = D_{12}^{\intercal} D_{12} > 0$ and $R_2 = (D_{21}^{\intercal})^{\intercal}(D_{12}^{\intercal}) > 0$, it follows by Theorem 3.7.13 that

$$\begin{aligned} \widehat{H}_{12} &= \left[\begin{array}{c|c} A + B_2E + B_2F_E & B_2R_1^{-\frac{1}{2}} \\ \hline C_1 + D_{12}E + D_{12}F_E & D_{12}R_1^{-\frac{1}{2}} \end{array}\right] \left[\begin{array}{c|c} A + B_2E & B_2 \\ \hline -R_1^{\frac{1}{2}} F_E & R_1^{\frac{1}{2}} \end{array}\right] \\ &= \left[\begin{array}{c|c} A + B_2F & B_2R_1^{-\frac{1}{2}} \\ \hline C_1 + D_{12}F & D_{12}R_1^{-\frac{1}{2}} \end{array}\right] \left[\begin{array}{c|c} A + B_2E & B_2 \\ \hline -R_1^{\frac{1}{2}}(F - E) & R_1^{\frac{1}{2}} \end{array}\right] := \widehat{U}\widehat{L} \end{aligned}$$

and

$$\begin{aligned} \widehat{H}_{21}^{\intercal} &= \left[\begin{array}{c|c} A^{\intercal} + C_2^{\intercal} M^{\intercal} + C_2^{\intercal}(L^{\intercal} - M^{\intercal}) & C_2^{\intercal} R_2^{-\frac{1}{2}} \\ \hline B_1^{\intercal} + D_{21}^{\intercal} M^{\intercal} + D_{21}^{\intercal}(L - M) & D_{21}^{\intercal} R_2^{-\frac{1}{2}} \end{array}\right] \left[\begin{array}{c|c} A^{\intercal} + C_2^{\intercal} M^{\intercal} & C_2^{\intercal} \\ \hline -R_2^{\frac{1}{2}} L_M & R_2^{\frac{1}{2}} \end{array}\right] \\ &= \left[\begin{array}{c|c} A^{\intercal} + C_2^{\intercal} L^{\intercal} & C_2^{\intercal} R_2^{-\frac{1}{2}} \\ \hline B_1^{\intercal} + D_{21}^{\intercal} L^{\intercal} & D_{21}^{\intercal} R_2^{-\frac{1}{2}} \end{array}\right] \left[\begin{array}{c|c} A^{\intercal} + C_2^{\intercal} M^{\intercal} & C_2^{\intercal} \\ \hline -R_2^{\frac{1}{2}}(L^{\intercal} - M^{\intercal}) & R_2^{\frac{1}{2}} \end{array}\right] := \widehat{V}^{\intercal}\widehat{M}^{\intercal} \end{aligned}$$

which completes the proof. ☐

Remark 4.6.2. By Lemma 4.3.3, $\widehat{H}_{12}^{\circ} = \widehat{U}R_1^{\frac{1}{2}}$ and $\widehat{H}_{21}^{\circ} = R_2^{\frac{1}{2}}\widehat{V}$ and by Corollary 4.6.1, $\widehat{H}_{12} = \widehat{U}\widehat{L}$ and $\widehat{H}_{21} = \widehat{M}\widehat{V}$. Importantly, $\widehat{H}_{12}$ and $\widehat{H}_{12}^{\circ}$ have the same inner factor $\widehat{U}$ and $\widehat{H}_{21}$ and $\widehat{H}_{21}^{\circ}$ have the same outer factor $\widehat{V}$. Furthermore, if $E = F$ and $M = L$ in Corollary 4.6.1, then $\widehat{L} = R_1^{\frac{1}{2}}$ and $\widehat{M} = R_2^{\frac{1}{2}}$, so that the factorizations in Lemma 4.3.3 are in fact a special case of Corollary 4.6.1.

We now have factorizations of $\widehat{H}_{12}$ and $\widehat{H}_{21}$ in terms of $\widehat{U}$ and $\widehat{V}$. In the next Lemma, we also obtain a decomposition of $\widehat{H}_{11}$ in terms of $\widehat{U}$ and $\widehat{V}$.

Lemma 4.6.3.
Let $\widehat{H}_{11}$ and $\widehat{H}_{11}^{\circ}$ be as in (4.35) and (4.36), respectively, with $\widehat{U}$ and $\widehat{V}$ as in (4.38). Then

$$\widehat{H}_{11} = \widehat{H}_{11}^{\circ} + \widehat{U}\widehat{P}\widehat{V},$$

where

$$\widehat{P} = R_1^{\frac{1}{2}} \left[\begin{array}{cc|c} A + B_2E & B_2E & -L \\ 0 & A + MC_2 & L - M \\ \hline E - F & E & 0 \end{array}\right] R_2^{\frac{1}{2}}.$$

.

Proof.
Let

$$\widehat{G}_c^E = \left[\begin{array}{c|c} A + B_2E & B_1 \\ \hline C_1 + D_{12}E & 0 \end{array}\right], \qquad \widehat{G}_f^M = \left[\begin{array}{c|c} A + MC_2 & B_1 + MD_{21} \\ \hline E & 0 \end{array}\right],$$
$$\widehat{G}_c^F = \left[\begin{array}{c|c} A + B_2F & B_1 \\ \hline C_1 + D_{12}F & 0 \end{array}\right] \quad \text{and} \quad \widehat{G}_f^L = \left[\begin{array}{c|c} A + LC_2 & B_1 + LD_{21} \\ \hline F & 0 \end{array}\right]. \tag{4.40}$$

Then by Lemma 4.2.4

$$\widehat{H}_{11} - \widehat{H}_{11}^\circ = (\widehat{G}_c^E - \widehat{G}_c^F) + (\widehat{H}_{12}^\circ \widehat{G}_f^L - \widehat{H}_{12}\widehat{G}_f^M) = (\widehat{G}_c^E - \widehat{G}_c^F) + \widehat{U}(R_1^{\frac{1}{2}}\widehat{G}_f^L - \widehat{L}\widehat{G}_f^M).$$

Now by equation (3.12)

$$\widehat{G}_c^E - \widehat{G}_c^F = \left[\begin{array}{cc|c} A + B_2F & 0 & -B_1 \\ 0 & A + B_2E & B_1 \\ \hline C_1 + D_{12}F & C_1 + D_{12}E & 0 \end{array}\right].$$

Set

$$\widehat{N} = R_1^{\frac{1}{2}} \left[\begin{array}{c|c} A + B_2E & B_1 \\ \hline E - F & 0 \end{array}\right].$$

Then by equation (3.13),

$$\widehat{U}\widehat{N} = \left[\begin{array}{c|c} A + B_2F & B_2 \\ \hline C_1 + D_{12}F & D_{12} \end{array}\right] R_1^{-\frac{1}{2}} R_1^{\frac{1}{2}} \left[\begin{array}{c|c} A + B_2E & B_1 \\ \hline E - F & 0 \end{array}\right] = \left[\begin{array}{cc|c} A + B_2F & B_2(E-F) & 0 \\ 0 & A + B_2E & B_1 \\ \hline C_1 + D_{12}F & D_{12}(E-F) & 0 \end{array}\right].$$

Applying the state space similarity $T = \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix}$, it follows that

$$\widehat{U}\widehat{N} = \left[\begin{array}{cc|c} A + B_2F & 0 & -B_1 \\ 0 & A + B_2E & B_1 \\ \hline C_1 + D_{12}F & C_1 + D_{12}E & 0 \end{array}\right] = \widehat{G}_c^E - \widehat{G}_c^F$$

Thus

$$\widehat{H}_{11} - \widehat{H}_{11}^\circ = (\widehat{G}_c^E - \widehat{G}_c^F) + \widehat{U}(R_1^{\frac{1}{2}}\widehat{G}_f^L - \widehat{L}\widehat{G}_f^M) = \widehat{U}(\widehat{N} + R_1^{\frac{1}{2}}\widehat{G}_f^L - \widehat{L}\widehat{G}_f^M).$$

We now consider the term $\widehat{N} + \widehat{L}\widehat{G}_f^M - R_1^{\frac{1}{2}}\widehat{G}_f^L$ and show that we can write it as $\widehat{P}\widehat{V}$. By equation (3.13),

$$-\widehat{L}\widehat{G}_f^M = R_1^{\frac{1}{2}} \left[\begin{array}{cc|c} A + B_2E & B_2E & 0 \\ 0 & A + MC_2 & -(B_1 + MD_{21}) \\ \hline E - F & E & 0 \end{array}\right]$$

and by equation (3.12)

$$\widehat{N} + R_1^{\frac{1}{2}}\widehat{G}_f^L = R_1^{\frac{1}{2}} \left[\begin{array}{cc|c} A + B_2E & 0 & B_1 \\ 0 & A + LC_2 & B_1 + LD_{21} \\ \hline E - F & F & 0 \end{array}\right].$$

Thus by equation (3.12)

$$-\widehat{L}\widehat{G}_f^M + \widehat{N} + R_1^{\frac{1}{2}}\widehat{G}_f^L = R_1^{\frac{1}{2}} \left[\begin{array}{cccc|c} A+B_2E & B_2E & 0 & 0 & 0 \\ 0 & A+MC_2 & 0 & 0 & -(B_1+MD_{21}) \\ 0 & 0 & A+B_2E & 0 & B_1 \\ 0 & 0 & 0 & A+LC_2 & B_1+LD_{21} \\ \hline E-F & E & E-F & F & 0 \end{array}\right]$$

$$= R_1^{\frac{1}{2}} \left[\begin{array}{ccc|c} A+B_2E & B_2E & 0 & B_1 \\ 0 & A+MC_2 & 0 & -(B_1+MD_{21}) \\ 0 & 0 & A+LC_2 & B_1+LD_{21} \\ \hline E-F & E & F & 0 \end{array}\right].$$

Applying the similarity $\begin{bmatrix} I & 0 & -I \\ 0 & I & I \\ 0 & 0 & I \end{bmatrix}$, it follows that

$$-\widehat{L}\widehat{G}_f^M + \widehat{N} + R_1^{\frac{1}{2}}\widehat{G}_f^L = R_1^{\frac{1}{2}} \left[\begin{array}{ccc|c} A+B_2E & B_2E & -LC_2 & -LD_{21} \\ 0 & A+MC_2 & (L-M)C_2 & (L-M)D_{21} \\ 0 & 0 & A+LC_2 & B_1+LD_{12} \\ \hline E-F & E & 0 & 0 \end{array}\right]$$

$$= R_1^{\frac{1}{2}} \left[\begin{array}{cc|c} A+B_2E & B_2E & -L \\ 0 & A+MC_2 & L-M \\ \hline E-F & E & 0 \end{array}\right] R_2^{\frac{1}{2}} R_2^{-\frac{1}{2}} \left[\begin{array}{c|c} A+LC_2 & B_1+LD_{21} \\ \hline C_2 & D_{21} \end{array}\right] = \widehat{P}\widehat{V},$$

where the second equality follows by (3.13). Thus

$$\widehat{H}_{11} - \widehat{H}_{11}^{\circ} = \widehat{U}(\widehat{N} + R_1^{\frac{1}{2}}\widehat{G}_f^L - \widehat{L}\widehat{G}_f^M) = \widehat{U}\widehat{P}\widehat{V},$$

which completes the proof. □

By Corollary 4.6.1 and Lemma 4.6.3, we now have a decomposition of the closed loop transfer function in terms of the inner and co-inner functions $\widehat{U}$ and $\widehat{V}$, specifically

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21} = \widehat{H}_{11}^{\circ} + \widehat{U}\widehat{P}\widehat{V} + \widehat{U}\widehat{L}\widehat{R}\widehat{M}\widehat{V} = \widehat{H}_{11}^{\circ} + \widehat{U}(\widehat{P} + \widehat{L}\widehat{R}\widehat{M})\widehat{V}.$$

Using the properties of inner and co-inner functions and Lemma 4.3.3, the norm of the closed loop transfer function can be minimized by choosing $\widehat{R}$ appropriately.

The following theorem, like Theorem 4.3.4, gives the optimal solution to the $\mathcal{H}_2$-control problem in Definition 4.3.1. The minimum norm agrees with Theorem 4.3.4.

Theorem 4.6.4.
The optimal solution to the $\mathcal{H}_2$-control problem

$$\begin{aligned} &\textit{minimize} && \|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2 = \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2^2 \\ &\textit{subject to} && \widehat{R} \in \mathcal{RH}_2 \end{aligned}$$

is given by

$$\widehat{R}_{opt} = \left[\begin{array}{cc|c} A+B_2F & LC_2 & -L \\ 0 & A+LC_2 & M-L \\ \hline F-E & E & 0 \end{array}\right],$$

where F and L are as in (4.37). Furthermore

$$\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}_{opt}\widehat{H}_{21}\|_2^2 = \|\widehat{H}_{11}^{\circ}\|_2^2 = \|\widehat{G}_c^F\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f^L\|_2^2,$$

where $\widehat{H}_{11}^{\circ}$ is as in (4.36) and where $\widehat{G}_c^F$ and $\widehat{G}_f^L$ are as in (4.40).

Proof.
By Corollary 4.6.1 and Lemma 4.6.3, we have

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{H}^{\circ}_{11} + \widehat{U}(\widehat{P} + \widehat{L}\widehat{R}\widehat{M})\widehat{V}.$$

By Lemma 4.2.4 $\widehat{H}^{\circ}_{11} = \widehat{G}^{F}_{c} - \widehat{H}^{\circ}_{12}\widehat{G}^{L}_{f}$. Thus, since $\widehat{H}^{\circ}_{12} = \widehat{U}R_1^{-\frac{1}{2}}$, it follows that

$$\|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2 = \|\widehat{H}^{\circ}_{11} + \widehat{U}(\widehat{P} + \widehat{L}\widehat{R}\widehat{M})\widehat{V}\|_2^2 = \|\widehat{G}^{F}_{c} - \widehat{U}R_1^{\frac{1}{2}}\widehat{G}^{L}_{f} + \widehat{U}(\widehat{P} + \widehat{L}\widehat{R}\widehat{M})\widehat{V}\|_2^2.$$

As in the proof of Theorem 4.3.4, by employing Lemma 4.3.3 and the norm preserving properties of the inner and co-inner functions $\widehat{U}$ and $\widehat{V}$, it follows that

$$\|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2 = \|\widehat{G}^{F}_{c}\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}^{L}_{f}\|_2^2 + \|\widehat{P} + \widehat{L}\widehat{R}\widehat{M}\|_2^2.$$

Clearly this norm attains its minimum when the last term is zero. By Corollary 4.6.1, $\widehat{L}$ and $\widehat{M}$ are invertible outer, thus the minimum is achieved when $\widehat{R} = -\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1}$. Hence we calculate by equations (3.14) and (3.13)

$$\begin{aligned}
\widehat{R} &= -\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} \\
&= -\left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline F-E & I \end{array}\right] R_1^{-\frac{1}{2}} R_1^{\frac{1}{2}} \left[\begin{array}{cc|c} A+B_2E & B_2E & -L \\ 0 & A+MC_2 & L-M \\ \hline E-F & E & 0 \end{array}\right] R_2^{\frac{1}{2}} R_2^{-\frac{1}{2}} \left[\begin{array}{c|c} A+LC_2 & L-M \\ \hline C_2 & I \end{array}\right] \\
&= \left[\begin{array}{cccc|c} A+B_2F & B_2(F-E) & -B_2E & 0 & 0 \\ 0 & A+B_2E & B_2E & -LC_2 & -L \\ 0 & 0 & A+MC_2 & (L-M)C_2 & L-M \\ 0 & 0 & 0 & A+LC_2 & L-M \\ \hline F-E & F-E & -E & 0 & 0 \end{array}\right].
\end{aligned}$$

With similarity transform $T = \begin{bmatrix} I & I & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & -I \\ 0 & 0 & 0 & I \end{bmatrix}$, it can be shown that

$$\begin{aligned}
\widehat{R} &= \left[\begin{array}{cccc|c} A+B_2F & 0 & 0 & -LC_2 & -L \\ 0 & A+B_2E & B_2E & B_2E-LC_2 & -L \\ 0 & 0 & A+MC_2 & 0 & 0 \\ 0 & 0 & 0 & A+LC_2 & L-M \\ \hline F-E & 0 & -E & -E & 0 \end{array}\right] \\
&= (E-F)\Phi_{A+B_2F}L + (F-E)\Phi_{A+B_2F}(LC_2)\Phi_{A+LC_2}(M-L) + E\Phi_{A+LC_2}(M-L) \\
&= \begin{bmatrix} F-E & E \end{bmatrix} \begin{bmatrix} \Phi_{A+B_2F} & \Phi_{A+B_2F}(LC_2)\Phi_{A+LC_2} \\ 0 & \Phi_{A+LC_2} \end{bmatrix} \begin{bmatrix} -L \\ M-L \end{bmatrix} \\
&= \left[\begin{array}{cc|c} A+B_2F & LC_2 & -L \\ 0 & A+LC_2 & M-L \\ \hline F-E & E & 0 \end{array}\right].
\end{aligned}$$

To show the optimal value of the norm, we compute

$$\begin{aligned}
\widehat{R}\widehat{H}_{21} &= \left[\begin{array}{cc|c} A+B_2F & LC_2 & -L \\ 0 & A+LC_2 & M-L \\ \hline F-E & E & 0 \end{array}\right] \left[\begin{array}{c|c} A+MC_2 & B_1+MD_{21} \\ \hline C_2 & D_{21} \end{array}\right] \\
&= \left[\begin{array}{ccc|c} A+B_2F & LC_2 & -LC_2 & -LD_{21} \\ 0 & A+LC_2 & (M-L)C_2 & (M-L)D_{21} \\ 0 & 0 & A+MC_2 & B_1+MD_{21} \\ \hline F-E & E & 0 & 0 \end{array}\right]
\end{aligned}$$

and hence by equation 3.13,

$$\widehat{H}_{12}\widehat{R}\widehat{H}_{21} = \left[\begin{array}{c|c} A+B_2E & B_2 \\ \hline C_1+D_{12}E & D_{12} \end{array}\right] \left[\begin{array}{ccc|c} A+B_2F & LC_2 & -LC_2 & -LD_{21} \\ 0 & A+LC_2 & (M-L)C_2 & (M-L)D_{21} \\ 0 & 0 & A+MC_2 & B_1+MD_{21} \\ \hline F-E & E & 0 & 0 \end{array}\right]$$

$$= \left[\begin{array}{cccc|c} A+B_2E & B_2(F-E) & B_2E & 0 & 0 \\ 0 & A+B_2F & LC_2 & -LC_2 & -LD_{21} \\ 0 & 0 & A+LC_2 & (M-L)C_2 & (M-L)D_{21} \\ 0 & 0 & 0 & A+MC_2 & B_1+MD_{21} \\ \hline C_1+D_{12}E & D_{12}(F-E) & D_{12}E & 0 & 0 \end{array}\right].$$

With similarity transform $T = \begin{bmatrix} 0 & I & -I & I \\ 0 & 0 & -I & I \\ I & -I & I & -I \\ 0 & 0 & 0 & -I \end{bmatrix}$, and employing equation (3.12) it follows that

$$\widehat{H}_{12}\widehat{R}\widehat{H}_{21} = \left[\begin{array}{cccc|c} A+B_2F & -B_2F & 0 & 0 & B_1 \\ 0 & A+LC_2 & 0 & 0 & B_1+LD_{21} \\ 0 & 0 & A+B_2E & -B_2E & -B_1 \\ 0 & 0 & 0 & A+MC_2 & -(B_1+MD_{21}) \\ \hline C_1+D_{12}F & -D_{12}F & C_1+D_{12}E & -D_{12}E & 0 \end{array}\right]$$

$$= \left[\begin{array}{cc|c} A+B_2F & -B_2F & B_1 \\ 0 & A+LC_2 & B_1+LD_{21} \\ \hline C_1+D_{12}F & -D_{12}F & 0 \end{array}\right] - \left[\begin{array}{cc|c} A+B_2E & -B_2E & B_1 \\ 0 & A+MC_2 & B_1+MD_{21} \\ \hline C_1+D_{12}E & -D_{12}E & 0 \end{array}\right]$$

$$= \widehat{H}_{11}^{\circ} - \widehat{H}_{11},$$

which proves that $\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21} = \widehat{H}_{11}^{\circ}$, which completes the proof. □

Corollary 4.6.5.
The optimal solution to the $\mathcal{H}_2$-control problem

$$\begin{array}{ll} \textit{minimize} & \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 \\ \textit{subject to} & \widehat{Q} \in \mathcal{RH}_2 \\ & \widehat{Q} = \widehat{R}\widehat{H}_{21} \textit{ for some } \widehat{R} \in \mathcal{RH}_2, \end{array}$$

is given by

$$\widehat{Q}_R = \left[\begin{array}{ccc|c} A+B_2F & LC_2 & -LC_2 & -LD_{21} \\ 0 & A+LC_2 & (M-L)C_2 & (M-L)D_{21} \\ 0 & 0 & A+MC_2 & B_1+MD_{21} \\ \hline F-E & E & 0 & 0 \end{array}\right] \tag{4.41}$$

and

$$\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}_R\|_2^2 = \|\widehat{H}_{11}^{\circ}\|_2^2 = \|\widehat{G}_c^F\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f^L\|_2^2,$$

where $\widehat{H}_{11}^{\circ}$ is as in (4.36) and where $\widehat{G}_c^F$ and $\widehat{G}_f^L$ are as in (4.40).

We end this section with result similar to Theorem 4.6.4.

Theorem 4.6.6 (cf. Lemma 3 in [25]).
Consider realizations

$$\widehat{F}_{11} = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & 0 \end{array}\right], \quad \widehat{F}_{12} = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right] \quad \text{and} \quad \widehat{F}_{21} = \left[\begin{array}{c|c} A_3 & B_3 \\ \hline C_3 & D_3 \end{array}\right].$$

such that

1. A_1, A_2 *and* A_3 *are stable;*
2. $R_2 = D_2^\intercal D_2 > 0$ *and* $R_3 = D_3 D_3^\intercal > 0$*;*
3. $\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ *and* $\begin{bmatrix} A - i\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ *have full column and row rank respectively for all* $\omega \in \mathbb{R}$.

Define

$$F_2 = -(D_2^\intercal D_2)^{-1}(B_2^\intercal X + D_2^\intercal C_2) \quad \text{and} \quad L_3 = -(Y C_3^\intercal + B_3 D_3^\intercal)(D_3 D_3^\intercal)^{-1},$$

where

$$X = \mathrm{Ric}(A_2, B_2, C_2, D_2) \quad \text{and} \quad Y = \mathrm{Ric}(A_3^\intercal, C_3^\intercal, B_3^\intercal, D_3^\intercal).$$

Then there exists unique solutions Z_1 *and* Z_2 *to the Sylvester equations*

$$\begin{aligned} (A_2 + B_2 F_2)^\intercal Z_1 + Z_1^\intercal A_1 + (C_2 + D_2 F_2)^\intercal C_1 = 0 \quad \text{and} \\ A Z_2 + (A_3 + L_3 C_3)^\intercal Z_2^\intercal + B_1 (B_3 + L_3 D_3)^\intercal = 0. \end{aligned}$$

Furthermore, a solution to the optimization problem

$$\begin{array}{ll} \text{minimize} & \|\widehat{F}_{11} + \widehat{F}_{12}\widehat{R}\widehat{F}_{21}\|_2^2 \\ \text{subject to} & \widehat{R} \in \mathcal{RH}_2 \end{array}$$

is given by

$$\widehat{R}_{opt} = -\widehat{L}^{-1} \left[\begin{array}{c|c} A_1 & B_1 D_3^\intercal + Z_2 C_3^\intercal \\ \hline B_2^\intercal Z_1 + D_2^\intercal C_1 & 0 \end{array}\right] \widehat{M}^{-1},$$

where

$$\widehat{L} = -R_2^{\frac{1}{2}} \left[\begin{array}{c|c} A_2 & B_2 \\ \hline -F_2 & I \end{array}\right] \quad \text{and} \quad \widehat{M} = \left[\begin{array}{c|c} A_3 & -L_3 \\ \hline C_3 & I \end{array}\right] R_3^{\frac{1}{2}}.$$

4.7 Reparameterized $\mathcal{H}_2$-Problem for the Output Feedback Case

We have seen that by the Youla parametrization,

$$\min_{\widehat{K}\in\mathcal{RH}_2} \|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2 = \min_{\widehat{R}\in\mathcal{RH}_2} \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2^2$$

where $\widehat{R} \in \mathcal{RH}_2$ is a parameter, $\widehat{J}$ is a canonical stable system, $\widehat{K} = \underline{\mathcal{F}}(\widehat{J},\widehat{R})$ and $\widehat{H} = \widehat{J} \star \widehat{R}$. Let $\widehat{Q} \in \mathcal{RH}_2$ be another parameter. In Chapter 7, we will consider the $\mathcal{H}_2$-control problem for linear systems that satisfy an additional structure requirement. For such systems it is possible to reduce control problems of the form

$$\begin{aligned} &\text{minimize} && \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 \\ &\text{subject to} && \widehat{Q} \in \mathcal{RH}_2 \end{aligned} \tag{4.42}$$

to the unstructured case. On the other hand this can not necessarily be done when we have the form $\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2^2$. In this section, we solve the control problem (4.42) via a spectral factorization approach similar to Section 4.5. In Section 4.5, the state feedback case of the $\mathcal{H}_2$-control problem also reduced to a problem of the form (4.42). Some of the results related to optimization and spectral factorization are repeated in this section in a more general setting. As in previous sections, we assume we have a plant

$$\widehat{G} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] = \begin{bmatrix} C_1\Phi_A B_1 & C_1\Phi_A B_2 + D_{12} \\ C_2\Phi_A B_1 + D_{21} & C_2\Phi_A B_2 \end{bmatrix} \tag{4.43}$$

such that the following conditions are satisfied.

Conditions 4.7.1.

1. the triples (A, B_1, C_2) and (A, B_2, C_1) are stabilizable and detectable;
2. $R_1 = D_{12}^{\intercal} D_{12} > 0$ and $R_2 = D_{21} D_{21}^{\intercal} > 0$;
3. $\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ and $\begin{bmatrix} A - i\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ have full column and row rank respectively for all $\omega \in \mathbb{R}$.

Let E and M be matrices such that $A + B_2E$ and $A + MC_2$ are stable and let $\widehat{J}$ be as in (4.33). Set $\widehat{H} = \widehat{G} \star \widehat{J}$ as in (4.34), that is

$$\widehat{H} = \begin{bmatrix} \widehat{H}_{11} & \widehat{H}_{12} \\ \widehat{H}_{21} & \widehat{H}_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} A + B_2E & -B_2E & B_1 & B_2 \\ 0 & A + MC_2 & B_1 + MD_{21} & 0 \\ \hline C_1 + D_{12}E & -D_{12}E & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right] = \left[\begin{array}{c|cc} \widetilde{A} & \widetilde{B}_1 & \widetilde{B}_2 \\ \hline \widetilde{C}_1 & 0 & D_{12} \\ \widetilde{C}_2 & D_{21} & 0 \end{array}\right], \tag{4.44}$$

In the following definition of the reparameterized $\mathcal{H}_2$-control problem, we do not require the state space matrices of $\widehat{H}$ to have the specific structure derived from $\widehat{G}$ as above, we only require that $\widetilde{A}$ is stable. We will return to the specific structure as given above after solving the general problem.

Definition 4.7.2 (Reparameterized $\mathcal{H}_2$-control problem)**.**
For

$$\widehat{H}_{11} = \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_1 \\ \hline \widetilde{C}_1 & 0 \end{array}\right] \quad \text{and} \quad \widehat{H}_{12} = \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_2 \\ \hline \widetilde{C}_1 & D_{12} \end{array}\right] \tag{4.45}$$

with $\widetilde{A}$ stable, the *reparameterized $\mathcal{H}_2$-control problem* is to synthesize a controller $\widehat{Q} \in \mathcal{RH}_2$ which minimizes the closed loop performance

$$\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2.$$

Since $\widetilde{A}$ is stable, we have $\widehat{H}_{11} \in \mathcal{RH}_2$ and $\widehat{H}_{12} \in \mathcal{RH}_\infty$. The following lemma is derived from the projection theorem and gives an equivalent condition for $\widehat{Q}$ to be an optimal solution of the reparameterized optimal control problem in Definition 4.7.2.

Lemma 4.7.3 (cf. in Lemma 11, [46])**.**
If $\widehat{H}_{11} \in \mathcal{RH}_2$ and $\widehat{H}_{12} \in \mathcal{RH}_\infty$, then $\widehat{Q}$ is optimal for

$$\begin{aligned} &\textit{minimize} && \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2 \\ &\textit{subject to} && \widehat{Q} \in \mathcal{RH}_2 \end{aligned}$$

if and only if

$$\widehat{H}_{12}^*\widehat{H}_{11} + \widehat{H}_{12}^*\widehat{H}_{12}\widehat{Q} \in \mathcal{H}_2^\perp.$$

In the above optimality condition, we apply spectral factorization to the term $\widehat{H}_{12}^*\widehat{H}_{12}$. In general, $\widehat{H}_{12}$ and $\widehat{H}_{12}^*$ are not invertible since D_{12} will generally not be invertible. Spectral factorization enables us to factorize $\widehat{H}_{12}^*\widehat{H}_{12}$ as $\widehat{L}^*\widehat{L}$ for an invertible outer function $\widehat{L} \in \mathcal{RH}_\infty$ which also has a stable inverse $\widehat{L}^{-1} \in \mathcal{RH}_\infty$ (and hence $\widehat{L}^{-*} \in \mathcal{RH}_2^\perp$ also). This will enable us to take inverses in the optimality condition given in Lemma 4.7.3 and hence solve for an optimal $\widehat{Q}$.

The following standard result on spectral factorization can be found for example in Theorem 13.6 in [1] or Lemma 13 in [46]. We restate the result in our context and also give a proof using our notation.

Theorem 4.7.4 (Spectral Factorization)**.**
Consider the transfer function $\widehat{H}_{12} \in \mathcal{RH}_\infty$ with a realization

$$\widehat{H}_{12} = \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_2 \\ \hline \widetilde{C}_1 & D_{12} \end{array}\right],$$

satisfying the following conditions

1. $R_1 = D_{12}^\intercal D_{12} > 0$
2. $\widetilde{X} = \mathrm{Ric}(\widetilde{A}, \widetilde{B}_2, \widetilde{C}_1, D_{12})$ *exists.*

Define

$$\widehat{L} = \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_2 \\ \hline -R_1^{\frac{1}{2}}\widetilde{F} & R_1^{\frac{1}{2}} \end{array}\right], \quad \textit{where} \quad \widetilde{F} = -R_1^{-1}(\widetilde{B}_2^\intercal\widetilde{X} + D_{12}^\intercal\widetilde{C}_1). \tag{4.46}$$

Then

1. $\widehat{L} \in \mathcal{RH}_\infty$ *(and hence $\widehat{L}^* \in \mathcal{RH}_\infty^-$);*
2. $\widehat{L}^{-1} \in \mathcal{RH}_\infty$ *(and hence $\widehat{L}^{-*} \in \mathcal{RH}_\infty^-$);*
3. $\widehat{H}_{12}^*\widehat{H}_{12} = \widehat{L}^*\widehat{L}$.

In particular, $\widehat{L}$ is an invertible outer function.

Proof.

1. Since $\widetilde{A}$ is stable, $\widehat{L} \in \mathcal{RH}_\infty$ (and hence by Proposition 3.6.1, $\widehat{L}^* \in \mathcal{RH}_\infty^-$).

2. By equation (3.14), it follows that

$$\widehat{L}^{-1} = \left[\begin{array}{c|c} \widetilde{A}+\widetilde{B}_2\widetilde{F} & \widetilde{B}_2 R_1^{-\frac{1}{2}} \\ \hline \widetilde{F} & R_1^{-\frac{1}{2}} \end{array}\right] = \widetilde{F}\Phi_{\widetilde{A}+\widetilde{B}_2\widetilde{F}}\widetilde{B}_2 R_1^{-\frac{1}{2}} + R_1^{-\frac{1}{2}}. \tag{4.47}$$

Since $\widetilde{X} = \mathrm{Ric}(\widetilde{A}, \widetilde{B}_2, \widetilde{C}_1, D_{12})$ is a stabilizing solution, $\widetilde{A}+\widetilde{B}_2\widetilde{F}$ is stable and hence $\widehat{L}^{-1} \in \mathcal{RH}_\infty$.

3. Let $R_1 = (D_{12}^\intercal D_{12})$, $W = \widetilde{B}_2^\intercal \widetilde{X} + D_{12}^\intercal \widetilde{C}_1$ and $\widetilde{F} = -R_1^{-1}W$. Recall that by (3.11)

$$\Phi_{\widetilde{A}} = (\lambda I - \widetilde{A})^{-1} \quad \text{and} \quad \Phi_{\widetilde{A}}^* = -(\lambda I + \widetilde{A}^\intercal)^{-1}.$$

Applying formula (3.17) to $\widehat{H}_{12}$, gives

$$\widehat{H}_{12}^*\widehat{H}_{12} = R_1 + D_{12}^\intercal \widetilde{C}_1 \Phi_{\widetilde{A}}\widetilde{B}_2 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^*(D_{12}^\intercal \widetilde{C}_1)^\intercal + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^* \widetilde{C}_1^\intercal \widetilde{C}_1 \Phi_{\widetilde{A}}\widetilde{B}_2.$$

From the ARE (3.29) it follows that:

$$(\widetilde{X}B_2 + (D_{12}^\intercal \widetilde{C}_1)^\intercal)R_1^{-1}(B^\intercal \widetilde{X} + D_{12}^\intercal \widetilde{C}_1) = \widetilde{A}^\intercal \widetilde{X} + \widetilde{X}\widetilde{A} + \widetilde{C}_1^\intercal \widetilde{C}_1$$

and so

$$\begin{aligned} W^\intercal R_1^{-1} W = \widetilde{A}^\intercal \widetilde{X} + \widetilde{X}\widetilde{A} + \widetilde{C}_1^\intercal \widetilde{C}_1 + \lambda \widetilde{X} - \lambda \widetilde{X} &= \widetilde{C}_1^\intercal \widetilde{C}_1 + \left(\lambda I + \widetilde{A}^\intercal\right)\widetilde{X} - \widetilde{X}\left(\lambda I - \widetilde{A}\right) \\ &= \widetilde{C}_1^\intercal \widetilde{C}_1 - \Phi_{\widetilde{A}}^{-*} X - X\Phi_{\widetilde{A}}^{-1}. \end{aligned}$$

Applying (3.17) to $\widehat{L}$, gives

$$\begin{aligned} \widehat{L}^*\widehat{L} &= R_1 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^* W^\intercal + W\Phi_{\widetilde{A}}\widetilde{B}_2 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^*(W^\intercal R_1^{-1} W)\Phi_{\widetilde{A}}\widetilde{B}_2 \\ &= R_1 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^* W^\intercal + W\Phi_{\widetilde{A}}\widetilde{B}_2 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^*(\widetilde{C}_1^\intercal \widetilde{C}_1 - \Phi_{\widetilde{A}}^{-*}\widetilde{X} - \widetilde{X}\Phi_{\widetilde{A}}^{-1})\Phi_{\widetilde{A}}\widetilde{B}_2 \\ &= R_1 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^* W^\intercal + W\Phi_{\widetilde{A}}\widetilde{B}_2 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^* \widetilde{C}_1^\intercal \widetilde{C}_1 \Phi_{\widetilde{A}}\widetilde{B}_2 - \widetilde{B}_2^\intercal \widetilde{X}\Phi_{\widetilde{A}}\widetilde{B}_2 - \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^* \widetilde{X}\widetilde{B}_2 \\ &= R_1 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^*(W^\intercal - \widetilde{X}\widetilde{B}_2) + (W - \widetilde{B}_2^\intercal \widetilde{X})\Phi_{\widetilde{A}}\widetilde{B}_2 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^* \widetilde{C}_1^\intercal \widetilde{C}_1 \Phi_{\widetilde{A}}\widetilde{B}_2 \\ &= R_1 + (D_{12}^\intercal \widetilde{C}_1)\Phi_{\widetilde{A}}\widetilde{B}_2 + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^*(D_{12}^\intercal \widetilde{C}_1)^\intercal + \widetilde{B}_2^\intercal \Phi_{\widetilde{A}}^* \widetilde{C}_1^\intercal \widetilde{C}_1 \Phi_{\widetilde{A}}\widetilde{B}_2, \end{aligned}$$

which shows that $\widehat{H}_{12}^*\widehat{H}_{12} = \widehat{L}^*\widehat{L}$. ☐

Form Lemma 4.7.3, we have that $\widehat{Q}$ is an optimal solution if and only if

$$\widehat{H}_{12}^*\widehat{H}_{11} + \widehat{H}_{12}^*\widehat{H}_{12}\widehat{Q} \in \mathcal{H}_2^\perp.$$

With $\widehat{L}$ as in equation (4.28), it follows from Theorem 4.7.4 that $\widehat{L}^{-*} \in \mathcal{RH}_\infty^-$ and

$$\widehat{H}_{12}^*\widehat{H}_{11} + \widehat{H}_{12}^*\widehat{H}_{12}\widehat{Q} = \widehat{H}_{12}^*\widehat{H}_{11} + \widehat{L}^*\widehat{L}\widehat{Q}.$$

Thus, by Proposition 3.6.1 (2), $\widehat{Q}$ is an optimal solution if and only if

$$\widehat{L}^{-*}\widehat{H}_{12}^*\widehat{H}_{11} + \widehat{L}\widehat{Q} \in \mathcal{H}_2^\perp.$$

The next lemma shows that the first term $\widehat{L}^{-*}\widehat{H}_{12}^*\widehat{H}_{11}$ in the above optimality condition can be decomposed into the sum of two terms; one of which is in $\mathcal{RH}_2$ and the other in $\mathcal{RH}_2^\perp$.

Lemma 4.7.5 (cf. Lemma 13 in [46])**.**
For $\widehat{H}_{11}$ and $\widehat{H}_{12}$ as in (4.45) and $\widehat{L}$ and $\widetilde{F}$ as in (4.46), it holds that

$$\widehat{L}^{-*}\widehat{H}_{12}^*\widehat{H}_{11} = -R_1^{\frac{1}{2}}\widetilde{F}\Phi_{\widetilde{A}}\widetilde{B}_1 - R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal \Phi_{-(\widetilde{A}+\widetilde{B}_2\widetilde{F})^\intercal}\widetilde{X}\widetilde{B}_1, \tag{4.48}$$

with $R_1^{\frac{1}{2}}\widetilde{F}\Phi_{\widetilde{A}}\widetilde{B}_1 \in \mathcal{RH}_2$ and $R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal \Phi_{-(\widetilde{A}+\widetilde{B}_2\widetilde{F})^\intercal}\widetilde{X}\widetilde{B}_1 \in \mathcal{RH}_2^\perp$.

Proof.
Fro m equation (3.15), it follows that $\Phi^{-*}_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})} = -(\lambda I + (\widetilde{A}+\widetilde{B}_2\widetilde{F})^\intercal)$ and from (3.29), it follows that:

$$\widetilde{C}_1^\intercal \widetilde{C}_1 = -\widetilde{F}^\intercal(\widetilde{B}_2^\intercal \widetilde{X} + D_{12}^\intercal \widetilde{C}_1) - \widetilde{A}^\intercal \widetilde{X} - \widetilde{X}\widetilde{A}$$

and hence

$$\widetilde{C}_1^\intercal \widetilde{C} + \widetilde{F}^\intercal D_{12}^\intercal \widetilde{C} = (\lambda I - \widetilde{A})\widetilde{X} - (\lambda I + (\widetilde{A}+\widetilde{B}_2\widetilde{F})^\intercal) = \widetilde{X}\Phi^{-1}_{\widetilde{A}} + \Phi^{-*}_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{X}.$$

The following observation is also in order:

$$\begin{aligned} \widetilde{B}_2\widetilde{F} &= (\lambda I - (\widetilde{A} - +\widetilde{B}_2\widetilde{F})) - (\lambda I - \widetilde{A}) \\ &= \Phi^{-1}_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})} - \Phi^{-1}_{\widetilde{A}}. \end{aligned} \tag{4.49}$$

Applying (3.14), the inverse of $\widehat{L}$ is $\widehat{L}^{-1} = R_1^{-\frac{1}{2}} + \widetilde{F}\Phi_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{B}_2 R_1^{-\frac{1}{2}}$ and $\widehat{H}_{12} = D_{12} + \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_2$, hence by equations (3.13) and (4.49), we get that

$$\begin{aligned} \widehat{H}_{12}\widehat{L}^{-1} &= D_{12}R_1^{-\frac{1}{2}} + D_{12}\widetilde{F}\Phi_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{B}_2 R_1^{-\frac{1}{2}} + \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_2 R_1^{-\frac{1}{2}} + \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_2\widetilde{F}\Phi_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{B}_2 R_1^{-\frac{1}{2}} \\ &= D_{12}R_1^{-\frac{1}{2}} + D_{12}\widetilde{F}\Phi_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{B}_2 R_1^{-\frac{1}{2}} + \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_2 R_1^{-\frac{1}{2}} + \widetilde{C}_1\Phi_{\widetilde{A}}(\Phi^{-1}_{\widetilde{A}} - \Phi^{-1}_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})})\Phi_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{B}_2 R_1^{-\frac{1}{2}} \\ &= D_{12}R_1^{-\frac{1}{2}} + D_{12}\widetilde{F}\Phi_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{B}_2 R_1^{-\frac{1}{2}} + \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_2 R_1^{-\frac{1}{2}} + \widetilde{C}_1\Phi_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{B}_2 R_1^{-\frac{1}{2}} - \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_2 R_1^{-\frac{1}{2}} \\ &= D_{12}R_1^{-\frac{1}{2}} + (\widetilde{C}_1 + D_{12}\widetilde{F})\Phi_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}\widetilde{B}_2 R_1^{-\frac{1}{2}} \end{aligned}$$

Taking the adjoint via equation (3.15) gives

$$\widehat{L}^{-*}\widehat{H}_{12}^* = R_1^{-\frac{1}{2}} D_{12}^\intercal + R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal \Phi^*_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}(\widetilde{C}_1^\intercal + \widetilde{F}^\intercal D_{12}^\intercal).$$

Multiplying on both sides by $\widehat{H}_{11} = \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_1$ and applying (4.49), gives

$$\begin{aligned} \widehat{L}^{-*}\widehat{H}_{12}^*\widehat{H}_{11} &= R_1^{-\frac{1}{2}} D_{12}^\intercal \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_1 + R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal\Phi^*_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}(\widetilde{C}_1^\intercal\widetilde{C}_1 + \widetilde{F}^\intercal D^\intercal \widetilde{C}_1)\Phi_{\widetilde{A}}\widetilde{B}_1 \\ &= R_1^{-\frac{1}{2}} D_{12}^\intercal \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_1 + R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal\Phi^*_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}(X\Phi^{-1} + \Phi^{-*}_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}X)\Phi_{\widetilde{A}}\widetilde{B}_1 \\ &= R_1^{-\frac{1}{2}} D_{12}^\intercal \widetilde{C}_1\Phi_{\widetilde{A}}\widetilde{B}_1 + R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal\Phi^*_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}X\widetilde{B}_1 + R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal X\Phi_{\widetilde{A}}\widetilde{B}_1 \\ &= R_1^{-\frac{1}{2}}(\widetilde{B}_2^\intercal X + D_{12}^\intercal\widetilde{C}_1)\Phi_{\widetilde{A}}\widetilde{B}_1 + R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal\Phi^*_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}X\widetilde{B}_1 \\ &= R_1^{\frac{1}{2}}\widetilde{F}\Phi_{\widetilde{A}}\widetilde{B}_1 + R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal\Phi^*_{(\widetilde{A}+\widetilde{B}_2\widetilde{F})}X\widetilde{B}_1, \end{aligned}$$

which is what we needed to show. Lastly, since $\widetilde{A}$ and $\widetilde{A}+B_2\widetilde{F}$ are stable, it follows that $R_1^{\frac{1}{2}}\widetilde{F}\Phi_{\widetilde{A}}\widetilde{B}_1 \in \mathcal{RH}_2$ and $R_1^{-\frac{1}{2}}\widetilde{B}_2^\intercal\Phi_{-(\widetilde{A}+\widetilde{B}_2\widetilde{F})^\intercal}\widetilde{X}\widetilde{B}_1 \in \mathcal{RH}_2^\perp$. □

The following theorem now gives the optimal solution to the reparameterized $\mathcal{H}_2$-control problem. The result follows quite easily from the previous lemmas. This is the main result of this section. Note that the result agrees with the result in Section 4.5.

Theorem 4.7.6 (cf. in Lemma 4.1 [38]).
Consider realizations

$$\widehat{H}_{11} = \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_1 \\ \hline \widetilde{C}_1 & 0 \end{array}\right] \quad \text{and} \quad \widehat{H}_{12} = \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_2 \\ \hline \widetilde{C}_1 & D_{12} \end{array}\right],$$

satisfying the following conditions

1. $\widetilde{A}$ *is stable;*
2. $R_1 = D_{12}^{\intercal} D_{12} > 0$*;*
3. $\begin{bmatrix} \widetilde{A} - i\omega I & \widetilde{B}_2 \\ \widetilde{C}_1 & D_{12} \end{bmatrix}$ *has full column rank for all* $\omega \in \mathbb{R}$.

Then optimal solution of the $\mathcal{H}_2$*-control problem in Definition 4.7.2, is given by*

$$\widehat{Q} = \left[\begin{array}{c|c} \widetilde{A} + \widetilde{B}_2 \widetilde{F} & \widetilde{B}_1 \\ \hline \widetilde{F} & 0 \end{array}\right], \qquad \text{where } \widetilde{F} = -R_1^{-1}(\widetilde{B}_2^{\intercal} \widetilde{X} + D_{12}^{\intercal} \widetilde{C}_1) \text{ with } \widetilde{X} = Ric\,(\widetilde{A}, \widetilde{B}_2, \widetilde{C}_1, D_{12}).$$

Proof.
By Lemma 4.7.3 $\widehat{Q} \in \mathcal{RH}_2$ minimizes $\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2$ if and only if

$$\widehat{H}_{12}^* \widehat{H}_{11} + \widehat{H}_{12}^* \widehat{H}_{12} \widehat{Q} = \widehat{H}_{12}^* \widehat{H}_{11} + \widehat{L}^* \widehat{L} \widehat{Q} \in \mathcal{H}_2^{\perp}$$

where $\widehat{L}$ is given by (4.46) in Theorem 4.7.4. Since $\widehat{L}^{-*} \in \mathcal{RH}_{\infty}^{-}$ it follows by Proposition 3.6.1 (3) that $\widehat{Q} \in \mathcal{RH}_2$ minimizes $\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\|_2^2$ if and only if

$$\widehat{L}^{-*} \widehat{H}_{12}^* \widehat{H}_{11} + \widehat{L}^{-*} \widehat{L}^* \widehat{L} \widehat{Q} = \widehat{L}^{-*} \widehat{H}_{12}^* \widehat{H}_{11} + \widehat{L} \widehat{Q} \in \mathcal{H}_2^{\perp}.$$

By Theorem 4.7.4, $\widehat{L} \in \mathcal{RH}_{\infty}$ and $\widehat{Q} \in \mathcal{RH}_2$ by assumption, hence $\widehat{L}\widehat{Q} \in \mathcal{RH}_2$. Furthermore, $\widehat{L}^{-*} \widehat{H}_{12}^* \widehat{H}_{11}$ is given by equation (4.48) in Lemma 4.7.5 with one term in $\mathcal{H}_2$ and the other in $\mathcal{H}_2^{\perp}$. Thus, projecting unto $\mathcal{H}_2$, gives

$$P_{\mathcal{H}_2}\left(\widehat{L}^{-*} \widehat{H}_{12}^* \widehat{H}_{11} + \widehat{L}\widehat{Q}\right) = -R_1^{\frac{1}{2}} \widetilde{F} \Phi_{\widetilde{A}} \widetilde{B}_1 + 0 + \widehat{L}\widehat{Q} = 0.$$

This equation can now be solved for $\widehat{Q}$. The inverse of $\widehat{L}$ is given by (4.47). Now $\widetilde{B}_2 \widetilde{F} = \Phi_{\widetilde{A}+\widetilde{B}_2\widetilde{F}}^{-1} - \Phi_{\widetilde{A}}^{-1}$ and so we obtain

$$\begin{aligned}
\widehat{Q} = \widehat{L}^{-1} R_1^{\frac{1}{2}} \widetilde{F} \Phi_{\widetilde{A}} \widetilde{B}_1 &= \left(R_1^{-\frac{1}{2}} + \widetilde{F} \Phi_{\widetilde{A}+\widetilde{B}_2\widetilde{F}} \widetilde{B}_2 R_1^{-\frac{1}{2}}\right) R_1^{\frac{1}{2}} \widetilde{F} \Phi_{\widetilde{A}} \widetilde{B}_1 \\
&= \widetilde{F} \Phi_{\widetilde{A}+\widetilde{B}_2\widetilde{F}} (\widetilde{B}_2 \widetilde{F}) \Phi_{\widetilde{A}} \widetilde{B}_1 + \widetilde{F} \Phi_{\widetilde{A}} \widetilde{B}_1 \\
&= \widetilde{F} \Phi_{\widetilde{A}+\widetilde{B}_2\widetilde{F}} (\Phi_{\widetilde{A}}^{-1} - \Phi_{\widetilde{A}+\widetilde{B}_2\widetilde{F}}^{-1}) \Phi_{\widetilde{A}} \widetilde{B}_1 + \widetilde{F} \Phi_{\widetilde{A}} \widetilde{B}_1 \\
&= \widetilde{F} \Phi_{\widetilde{A}+\widetilde{B}_2\widetilde{F}} \widetilde{B}_1 - \widetilde{F} \Phi_{\widetilde{A}} \widetilde{B}_1 + \widetilde{F} \Phi_{\widetilde{A}} \widetilde{B}_1 \\
&= \left[\begin{array}{c|c} \widetilde{A} + \widetilde{B}_2 \widetilde{F} & \widetilde{B}_1 \\ \hline \widetilde{F} & 0 \end{array}\right],
\end{aligned}$$

which completes the proof. □

In the above theorem, the conditions guarantees the existence of a unique solution $\widetilde{X} = \text{Ric}(\widetilde{A}, \widetilde{B}_2, \widetilde{C}_1, D_{12})$. Recall that for the reparameterized $\mathcal{H}_2$-control problem, the results did not depend on the specific form of the matrices $\widetilde{A}$, $\widetilde{B}_1$, $\widetilde{B}_2$, $\widetilde{C}_1$ and $\widetilde{C}_2$. We now investigate the conditions in Theorem 4.7.6 and the related Riccati equation if the matrices $\widetilde{A}$, $\widetilde{B}_1$, $\widetilde{B}_2$, $\widetilde{C}_1$ and $\widetilde{C}_2$ are as in equation (4.44). The following lemma sheds some light on the form of the solution $\widetilde{X}$ of the Riccati equation and consequently also on the matrix $\widetilde{F}$.

Lemma 4.7.7.
Suppose

1. $D_{12}^\intercal D_{12} > 0$;
2. (A, B_2) *is stabilizable and* (A, C_2) *is detectable;*
3. $\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ *has full column rank for all* $\omega \in \mathbb{R}$,

and that E *and* M *are matrices such that* $A + B_2E$ *and* $A + MC_2$ *are stable. If* $\widetilde{A}$, $\widetilde{B}_1$, $\widetilde{B}_2$, $\widetilde{C}_1$ *and* $\widetilde{C}_2$ *are as in equation* (4.44)*, then there exists a unique solution* $\widetilde{X} = Ric\,(\widetilde{A}, \widetilde{B}_2, \widetilde{C}_1, D_{12})$ *and* $\widetilde{X}$ *and* $\widetilde{F}$ *are given by*

$$\widetilde{X} = \begin{bmatrix} X_1 & 0 \\ 0 & 0 \end{bmatrix} \quad \textit{where} \quad X_1 = \mathrm{Ric}(A, B_2, C_1, D_{12})$$

and

$$\widetilde{F} = \begin{bmatrix} F - E & E \end{bmatrix} \quad \textit{where} \quad F = -(D_{12}^\intercal D_{12})(B_2^\intercal X_1 + D_{12}^\intercal C_1).$$

Proof.
By assumption $D_{12}^\intercal D_{12} > 0$. Secondly, since $A + B_2E$ and $A + MC_2$ are stable, $\widetilde{A}$ is stable and thus $(\widetilde{A}, \widetilde{B}_2)$ is stabilizable. Thirdly, since

$$\begin{bmatrix} I & 0 \\ E & I \end{bmatrix}$$

is invertible,

$$\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix} \begin{bmatrix} I & 0 \\ E & I \end{bmatrix} = \begin{bmatrix} (A + B_2E) - i\omega I & B_2 \\ C_1 + D_{12}E & D_{12} \end{bmatrix}$$

has full column rank for each $\omega \in \mathbb{R}$. But since $A + MC_2$ is stable, $(A + MC_2) - i\omega I$ is invertible, thus

$$\begin{bmatrix} (A + B_2E) - i\omega I & -B_2E & B_2 \\ 0 & (A + MC_2) - i\omega I & 0 \\ C_1 + D_{12}E & -D_{12}E & D_{12} \end{bmatrix}$$

has full column rank for each $\omega \in \mathbb{R}$. Thus, by Theorem 3.7.8, there exists a unique stabilizing solution $\widetilde{X} = \mathrm{Ric}\,(\widetilde{A}, \widetilde{B}_2, \widetilde{C}_1, D_{12})$ to the Riccati equation

$$\widetilde{A}^\intercal \widetilde{X} + \widetilde{X}\widetilde{A} + \widetilde{C}_1^\intercal \widetilde{C}_1 - (\widetilde{X}\widetilde{B}_2 + \widetilde{C}_1^\intercal D_{12})(D_{12}^\intercal D_{12})^{-1}(\widetilde{B}_2^\intercal \widetilde{X} + D_{12}^\intercal \widetilde{C}_1) = 0. \tag{4.50}$$

Since $\widetilde{X}$ is symmetric, we may take

$$\widetilde{X} = \begin{bmatrix} X_1 & X_3 \\ X_3^\intercal & X_2 \end{bmatrix}.$$

With $\widetilde{A}$, $\widetilde{B}_1$, $\widetilde{B}_2$, $\widetilde{C}_1$ and $\widetilde{C}_2$ as in equation (4.44), the left hand side of equation (4.50) is given by

$$\begin{aligned} &\begin{bmatrix} A^\intercal + E^\intercal B_2^\intercal & 0 \\ -E^\intercal B_2^\intercal & A^\intercal + C_2^\intercal M^\intercal \end{bmatrix} \begin{bmatrix} X_1 & X_3 \\ X_3 & X_2 \end{bmatrix} + \begin{bmatrix} X_1 & X_3 \\ X_3 & X_2 \end{bmatrix} \begin{bmatrix} A + B_2E & -B_2E \\ 0 & A + MC_2 \end{bmatrix} \\ &+ \begin{bmatrix} C_1^\intercal + E^\intercal D_{12}^\intercal \\ -E^\intercal D_{12}^\intercal \end{bmatrix} \begin{bmatrix} C_1 + D_{12}E & -D_{12}E \end{bmatrix} \\ &- \left(\begin{bmatrix} X_1 & X_3 \\ X_3 & X_2 \end{bmatrix} \begin{bmatrix} B_2 \\ 0 \end{bmatrix} + \begin{bmatrix} C_1^\intercal + E^\intercal D_{12}^\intercal \\ -E^\intercal D_{12}^\intercal \end{bmatrix} D_{12} \right) R_1^{-1} \left(\begin{bmatrix} B_2^\intercal & 0 \end{bmatrix} \begin{bmatrix} X_1 & X_3 \\ X_3 & X_2 \end{bmatrix} + D_{12}^\intercal \begin{bmatrix} C_1 + D_{12}E & -D_{12}E \end{bmatrix} \right) \end{aligned}$$

which gives

$$
\begin{aligned}
= \quad & \begin{bmatrix} A^\intercal X_1 + E^\intercal B_2^\intercal X_1 & A^\intercal X_3 + E^\intercal B_2^\intercal X_3 \\ A^\intercal X_3 + C_2^\intercal M^\intercal X_3 - E^\intercal B_2^\intercal X_1 & A^\intercal X_2 + C_2^\intercal M^\intercal X_2 - E^\intercal B_2^\intercal X_3 \end{bmatrix} \\
& + \begin{bmatrix} X_1 A + X_1 B_2 E & X_3 A + X_3 M C_2 - X_1 B_2 E \\ X_3 A + X_3 B_2 E & X_2 A + X_2 M C_2 - X_3 B_2 E \end{bmatrix} \\
& + \begin{bmatrix} C_1^\intercal C_1 + E^\intercal D_{12}^\intercal C_1 + C_1^\intercal D_{12} E + E^\intercal R_1 E & -C_1^\intercal D_{12} E - E^\intercal R_1 E \\ -E^\intercal D_{12}^\intercal C_1 - E^\intercal R_1 E & E^\intercal R_1 E \end{bmatrix} \\
& - \begin{bmatrix} (X_1 B_2 + C_1^\intercal D_{12}) + E^\intercal R_1 \\ X_3 B_2 - E^\intercal R_1 \end{bmatrix} R_1^{-1} \begin{bmatrix} (B_2^\intercal X_1 + D_{12}^\intercal C_1) + E & R_1^{-1} B_2^\intercal X_3 - E \end{bmatrix}.
\end{aligned}
$$

Due to restrictions on space, we continue the calculation by considering each block entry separately. The top left block entry gives

$$
\begin{aligned}
& A^\intercal X_1 + E^\intercal B_2^\intercal X_1 + X_1 A + X_1 B_2 E + C_1^\intercal C_1 + E^\intercal D_{12}^\intercal C_1 + C_1^\intercal D_{12} E + E^\intercal R_1 E \\
& -((X_1 B_2 + C_1^\intercal D_{12}) + E^\intercal R_1)(R_1^{-1}(B_2^\intercal X_1 + D_{12}^\intercal C_1) + E) \\
= \quad & A^\intercal X_1 + E^\intercal B_2^\intercal X_1 + X_1 A + X_1 B_2 E + C_1^\intercal C_1 + E^\intercal D_{12}^\intercal C_1 + C_1^\intercal D_{12} E + E^\intercal R_1 E \\
& -(X_1 B_2 + C_1^\intercal D_{12}) R_1^{-1}(B_2^\intercal X_1 + D_{12}^\intercal C_1) - (X_1 B_2 + C_1^\intercal D_{12})E - E^\intercal (B_2^\intercal X_1 + D_{12}^\intercal C_1) - E^\intercal R_1 E \\
= \quad & A^\intercal X_1 + X_1 A + C_1^\intercal C_1 - (X_1 B_2 + C_1^\intercal D_{12}) R_1^{-1}(B_2^\intercal X_1 + D_{12}^\intercal C_1).
\end{aligned}
$$

The top right block entry gives

$$
\begin{aligned}
& A^\intercal X_3 + E^\intercal B_2^\intercal X_3 + X_3 A + X_3 M C_2 - X_1 B_2 E - C_1^\intercal D_{12} E - E^\intercal R_1 E \\
& - ((X_1 B_2 + C_1^\intercal D_{12}) + E^\intercal R_1)(R_1^{-1} B_2^\intercal X_3 - E) \\
= & A^\intercal X_3 + E^\intercal B_2^\intercal X_3 + X_3 A + X_3 M C_2 - X_1 B_2 E - C_1^\intercal D_{12} E - E^\intercal R_1 E \\
& - (X_1 B_2 + C_1^\intercal D_{12}) R_1^{-1} B_2^\intercal X_3 + (X_1 B_2 + C_1^\intercal D_{12}) E - E^\intercal B_2^\intercal X_3 + E^\intercal R_1 E \\
= & A^\intercal X_3 + X_3 A + X_3 M C_2 - (X_1 B_2 + C_1^\intercal D_{12}) R_1^{-1} B_2^\intercal X_3 \\
= & X_3(A + M C_2) + (A^\intercal - (X_1 B_2 + C_1^\intercal D_{12}) R_1^{-1} B_2^\intercal) X_3 \\
= & X_3(A + M C_2) + (A + B_2 F)^\intercal X_3
\end{aligned}
$$

The bottom left block entry is the transpose of the top right entry and thus is given by

$$
(A + M C_2)^\intercal X_3^\intercal + X_3^\intercal (A + B_2 F).
$$

The bottom right block entry gives

$$
\begin{aligned}
& A^\intercal X_2 + C_2^\intercal M^\intercal X_2 - E^\intercal B_2^\intercal X_3 + X_2 A + X_2 M C_2 - X_3 B_2 E + E^\intercal R_1 E - (X_3 B_2 - E^\intercal R_1)(R_1^{-1} B_2^\intercal X_3 - E) \\
= & A^\intercal X_2 + C_2^\intercal M^\intercal X_2 - E^\intercal B_2^\intercal X_3 + X_2 A + X_2 M C_2 - X_3 B_2 E + E^\intercal R_1 E \\
& - X_3 B_2 R_1^{-1} B_2^\intercal X_3 + X_3 B_2 E + E^\intercal B_2^\intercal X_3 - E^\intercal R_1 E \\
= & A^\intercal X_2 + C_2^\intercal M^\intercal X_2 + X_2 A + X_2 M C_2 - X_3 B_2 R_1^{-1} B_2^\intercal X_3 \\
= & (A + M C_2)^\intercal X_2 + X_2(A + M C_2) - X_3 B_2 R_1^{-1} B_2^\intercal X_3.
\end{aligned}
$$

Hence, the Riccati equation (4.50) is equivalent to the following four equations

$$
\begin{aligned}
A^\intercal X_1 + X_1 A + C_1^\intercal C_1 - (X_1 B_2 + C_1^\intercal D_{12}) R_1^{-1}(B_2^\intercal X_1 + D_{12}^\intercal C_1) &= 0, \\
X_3(A + M C_2) + (A + B_2 E)^\intercal X_3 &= 0, \\
(A + M C_2)^\intercal X_3^\intercal + X_3^\intercal (A + B_2 E) &= 0 \quad \text{and} \\
(A + M C_2)^\intercal X_2 + X_2(A + M C_2) - X_3 B_2 R_1^{-1} B_2^\intercal X_3 &= 0.
\end{aligned}
$$

The first equation is a Riccati equation which has a unique solution $X_1 = \text{Ric}(A, B_2, C_1, D_{12})$ by Theorem 3.7.8. Furthermore, $X_2 = 0$ and $X_3 = 0$ solves the other three equations. But, the solution $\widetilde{X} = \text{Ric}\,(\widetilde{A}, \widetilde{B}_2, \widetilde{C}_1, D_{12})$ is unique, hence it is given by

$$\widetilde{X} = \begin{bmatrix} X_1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{where} \quad X_1 = \text{Ric}(A, B_2, C_1, D_{12}).$$

Lastly, computing $\widetilde{F}$ gives

$$\begin{aligned}
\widetilde{F} &= -(D_{12}^{\intercal} D_{12})^{-1} \left(\begin{bmatrix} B_2^{\intercal} & 0 \end{bmatrix} \begin{bmatrix} X_1 & 0 \\ 0 & 0 \end{bmatrix} + D_{12}^{\intercal} \begin{bmatrix} C_1 + D_{12}E & -D_{12}E \end{bmatrix} \right) \\
&= -(D_{12}^{\intercal} D_{12})^{-1} \begin{bmatrix} (B_2^{\intercal} X_1 + D_{12}^{\intercal} C_1) + (D_{12}^{\intercal} D_{12})E & -D_{12}^{\intercal} D_{12}E \end{bmatrix} \\
&= \begin{bmatrix} F - E & E \end{bmatrix} \\
\text{where} \quad F &= -(D_{12}^{\intercal} D_{12})^{-1}(B_2^{\intercal} X_1 + D_{12}^{\intercal} C_1) \quad \text{and} \quad X_1 = \text{Ric}(A, B_2, C_1, D_{12}).
\end{aligned}$$

This completes the proof. □

The following result gives the optimal solution to the reparameterized $\mathcal{H}_2$-control problem where the state space matrices in $\widehat{H}$ are derived from $\widehat{G}$ as in equation (4.44).

Corollary 4.7.8.
Suppose the plant $\widehat{G}$ in (4.43) satisfies the conditions 4.7.1. Let $X_1 = Ric\,(A, B_2, C_1, D_{12})$. Let E and M be matrices such that $A + B_2E$ and $A + MC_2$ are stable and let $\widehat{H}_{11}$ and $\widehat{H}_{12}$ be as in (4.44). Then the optimal solution of the $\mathcal{H}_2$-control problem in Definition 4.7.2 is given by

$$\widehat{Q}_{opt} = \left[\begin{array}{cc|c} A + B_2F & 0 & B_1 \\ 0 & A + MC_2 & B_1 + MD_{21} \\ \hline F - E & E & 0 \end{array}\right], \quad \text{where} \quad F = -R_1^{-1}(B_2^{\intercal} X_1 + D_{12}^{\intercal} C_1). \tag{4.51}$$

Furthermore, the minimum norm is given by

$$\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}_{opt}\|_2 = \|\widehat{G}_c^F\|_2 \qquad \text{where} \qquad \widehat{G}_c^F = \left[\begin{array}{c|c} A + B_2F & B_1 \\ \hline C_1 + D_{12}F & 0 \end{array}\right].$$

Proof.
As shown in the proof of Lemma 4.7.7, the realizations $\widehat{H}_{11}$ and $\widehat{H}_{12}$ satisfy the conditions of Theorem 4.7.6. Thus the optimal solution $\widehat{Q}_{opt}$ is obtained from Theorem 4.7.6.
In order to determine the minimum norm, we compute $\widehat{H}_{12}\widehat{Q}_{\text{opt}}$ by applying equation (3.13) as follows

$$\begin{aligned}
\widehat{H}_{12}\widehat{Q}_{\text{opt}} &= \left[\begin{array}{cc|c} A + B_2E & -B_2E & B_2 \\ 0 & A + MC_2 & 0 \\ \hline C_1 + D_{12}E & -D_{12}E & D_{12} \end{array}\right] \left[\begin{array}{cc|c} A + B_2F & 0 & B_1 \\ 0 & A + MC_2 & B_1 + MD_{21} \\ \hline F - E & E & 0 \end{array}\right] \\
&= \left[\begin{array}{cccc|c} A + B_2E & -B_2E & B_2F - B_2E & B_2E & 0 \\ 0 & A + MC_2 & 0 & 0 & 0 \\ 0 & 0 & A + B_2F & 0 & B_1 \\ 0 & 0 & 0 & A + MC_2 & B_1 + MD_{21} \\ \hline C_1 + D_{12}E & -D_{12}E & D_{12}F - D_{12}E & D_{12}E & 0 \end{array}\right]
\end{aligned}$$

By equation (3.19) in Lemma 3.5.5, it follows that

$$\begin{aligned}
\widehat{H}_{12}\widehat{Q}_{\text{opt}} &= \left[\begin{array}{cc|c} A + B_2E & -B_2E & -B_1 \\ 0 & A + MC_2 & -(B_1 + MD_{21}) \\ \hline C_1 + D_{12}E & -D_{12}E & 0 \end{array}\right] + \left[\begin{array}{cc|c} A + B_2F & 0 & B_1 \\ 0 & A + MC_2 & B_1 + MD_{21} \\ \hline C_1 + D_{12}F & 0 & 0 \end{array}\right] \\
&= -\widehat{H}_{11} + \widehat{G}_c.
\end{aligned}$$

Thus $\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}_{opt}\|_2 = \|\widehat{G}_c^F\|_2$. □

Let $\widehat{Q}_R$ be as in (4.41) and let $\widehat{Q}_{opt}$ be as in equation (4.51) in Corollary 4.7.8. By Corollary 4.6.5, $\widehat{Q}_R$ corresponds to the optimal solution of the original $\mathcal{H}_2$-control problem in Definition 4.3.1 with optimal norm

$$\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}_R\|_2^2 = \|\widehat{H}_{11}^{\circ}\|_2^2 = \|\widehat{G}_c^F\|_2^2 + \|R_1^{\frac{1}{2}}\widehat{G}_f^L\|_2^2 \neq \|\widehat{G}_c^F\|_2 = \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}_{opt}\|_2.$$

This means that the optimal solution $\widehat{Q}_{opt}$ to the reparameterized problem in Definition 4.7.2 does not in general correspond to the optimal solution of the original $\mathcal{H}_2$-control problem in Definition 4.3.1. We note here that in the state feedback case in Section 4.5, the optimal solution to the reparameterized problem indeed also gave the optimal solution to the original $\mathcal{H}_2$-control problem.

The next lemma considers what needs to be added to the solution $\widehat{Q}_{opt}$ of the reparameterized problem in order to give the parameter $\widehat{Q}_R$ which corresponds to the solution of the original $\mathcal{H}_2$-control problem.

Lemma 4.7.9.
Let $\widehat{Q}_R$ and $\widehat{Q}_{opt}$ be given as in Corollaries 4.6.5 and 4.7.8 respectively. Set

$$\check{Q} = \left[\begin{array}{cc|c} A+B_2F & B_2F & 0 \\ 0 & A+LC_2 & -(B_1+LD_{21}) \\ \hline F-E & F & 0 \end{array}\right].$$

Then $\widehat{Q}_{opt} + \check{Q} = \widehat{Q}_R$.

Proof.
By a special case of equation (3.22)

$$\begin{aligned} \widehat{Q}_{opt} + \check{Q} &= \left[\begin{array}{cc|c} A+B_2F & 0 & B_1 \\ 0 & A+MC_2 & B_1+MD_{21} \\ \hline F-E & E & 0 \end{array}\right] + \left[\begin{array}{cc|c} A+B_2F & B_2F & 0 \\ 0 & A+LC_2 & -(B_1+LD_{21}) \\ \hline F-E & F & 0 \end{array}\right] \\ &= \left[\begin{array}{ccc|c} A+B_2F & 0 & B_2F & B_1+0 \\ 0 & A+MC_2 & 0 & B_1+MD_{21} \\ 0 & 0 & A+LC_2 & -(B_1+LD_{21}) \\ \hline F-E & E & F & 0+0 \end{array}\right] \end{aligned}$$

With state space similarity

$$T = \begin{bmatrix} I & 0 & I \\ 0 & I & I \\ 0 & 0 & I \end{bmatrix},$$

it can be seen that

$$\widehat{Q}_{opt} + \check{Q} = \left[\begin{array}{ccc|c} A+B_2F & LC_2 & -LC_2 & -LD_{21} \\ 0 & A+LC_2 & (M-L)C_2 & (M-L)D_{21} \\ 0 & 0 & A+MC_2 & B_1+MD_{21} \\ \hline F-E & E & 0 & 0 \end{array}\right] = \widehat{Q}_R,$$

which completes the proof. □

Chapter 5

Structured Linear Systems

In this chapter, we consider decentralized systems, where decentralization is determined by requiring that the system satisfies a certain sparsity constraint. Often this constraint is expressed by requiring that the system matrices, appearing in a realization of the system, have a specified block-zero structure. We show how binary relations impose such block zero structures on the system matrices. We will call a system a structured linear system if it has an underlying binary relation that determines the block zero structure of its system matrices. We investigate the properties of binary relations and how these properties manifest in the block zero-structure of the system matrices. We show that various sub-classes of decentralized systems found in the literature such as systems over graphs [24], poset-causal systems [38], coordinated linear systems [34] and leader-follower systems [44] (also called two-player systems), correspond with systems that have an underlying binary relation which determines their structure. Since any binary relation has a digraph representation, we show the that the graph-theoretic and binary relation approach to decentralized systems with sparsity constraints are equivalent.

5.1 Structured Linear Systems

Consider a decentralized linear system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0, \\ y(t) &= Cx(t) + Du(t), \quad t \geq 0, \end{aligned} \tag{5.1}$$

consisting of p interconnected subsystems labeled $1, \ldots, p$. Each subsystem has its own local state x_i, local input u_i and local output y_i. Hence the system is given by the interconnected equations

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j=1}^{p} A_{ij} x_j(t) + \sum_{j=1}^{p} B_{ij} u_j(t), \quad x_i(0) = x_{i,0}, \\ y_i(t) &= \sum_{j=1}^{p} C_{ij} x_j(t) + \sum_{j=1}^{p} D_{ij} u_j(t), \quad t \geq 0 \end{aligned} \tag{5.2}$$

and the system matrices partition as

$$A = \begin{bmatrix} A_{11} & \ldots & A_{1p} \\ \vdots & \ddots & \vdots \\ A_{p1} & \ldots & A_{pp} \end{bmatrix}, \; B = \begin{bmatrix} B_{11} & \ldots & B_{1p} \\ \vdots & \ddots & \vdots \\ B_{p1} & \ldots & B_{pp} \end{bmatrix}, \; C = \begin{bmatrix} C_{11} & \ldots & C_{1p} \\ \vdots & \ddots & \vdots \\ C_{p1} & \ldots & C_{pp} \end{bmatrix} \text{ and } D = \begin{bmatrix} D_{11} & \ldots & D_{1p} \\ \vdots & \ddots & \vdots \\ D_{p1} & \ldots & D_{pp} \end{bmatrix}. \tag{5.3}$$

From equation (5.2), it is clear that

$$x_j \text{ influences } x_i \text{ if } A_{ij} \neq 0, \quad u_j \text{ influences } x_i \text{ if } B_{ij} \neq 0,$$
$$x_j \text{ influences } y_i \text{ if } C_{ij} \neq 0, \quad u_j \text{ influences } y_i \text{ if } D_{ij} \neq 0.$$

Hence we see that the location of zeros in the system matrices, that is, the sparsity patterns of the system matrices, reflect how the subsystems influence each other through their local states x_j and local inputs u_j. Thus the sparsity patterns of the system matrices reflect the communication structure of the interconnected subsystems. Let $P = \{1, \dots, p\}$ and let $T \subseteq P \times P$. Then $\mathcal{T} = (P, T)$ is a binary relation on P. Binary relations can model the communication structure of decentralized systems in the following way. For $A = [A_{ij}]$ as in (5.3), set $A_{ij} = 0$ if $\{k \in P \colon (j,k), (k,i) \in T\} = \emptyset$. Then $\mathcal{T}$ determines where the block matrix A has block-zero entries and this reflects how the local states of the subsystems influence each other. Recall that for a binary relation $\mathcal{T}$ and partitions $\underline{n}, \underline{m} \in \mathbb{Z}_+^p$, the block incidence vector space $\mathcal{I}_{\mathcal{T}}^{\underline{n} \times \underline{m}} \subseteq \mathbb{R}^{\underline{n} \times \underline{m}}$ was defined by

$$\mathcal{I}_{\mathcal{T}}^{\underline{n} \times \underline{m}} := \{G = [G_{ij}] \in \mathbb{R}^{\underline{n} \times \underline{m}} \colon G_{ij} = 0 \text{ if } \{k \in P \colon (j,k), (k,i) \in T\} = \emptyset\}.$$

We now define structured linear systems as systems $\Sigma \sim (A, B, C, D)$ for which the block sparsity patterns of A, B, C, D are all determined by the same binary relation T.

Definition 5.1.1 (Structured linear system)**.**
Consider a set $P = \{1, \dots, p\}$, a binary relation $\mathcal{T} = (P, T)$ and partitions $\underline{n}, \underline{m}, \underline{r} \in \mathbb{Z}_+^p$. A linear system (5.1) is called a *structured linear system* if

$$A \in \mathcal{I}_{\mathcal{T}}^{\underline{n} \times \underline{n}}, \quad B \in \mathcal{I}_{\mathcal{T}}^{\underline{n} \times \underline{m}}, \quad C \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{n}}, \quad D \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}.$$

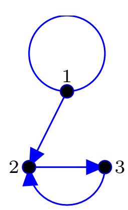
We write $\Sigma_{\mathcal{T}} \sim (A, B, C, D)$ to indicate that the underlying binary relation of the structured linear system is $\mathcal{T}$.

There is a one-to-one correspondence between the binary relation $\mathcal{T} = (P, T)$ and the digraph $\mathcal{G} = (P, T)$ where $P = \{1, \dots, p\}$ is the set of vertices and $T \subseteq \{(i, j) \in P \times P\}$ is the set of directed edges. Here a directed edge (i, j) originates at i and terminates at j. Hence we can illustrate the underlying communication structure of a structured linear system by its associated digraph.

Example 5.1.2.
Consider the binary relation $\mathcal{T} = (P, T)$ with $P = \{1, 2, 3\}$ and

$$T = \{(1,1); (1,2); (2,3); (3,2)\}.$$

The corresponding digraph and block sparsity pattern of a system matrix M of a structured linear system $\Sigma_{\mathcal{T}} \sim (A, B, C, D)$ is shown below.



$$M = \begin{bmatrix} M_{11} & 0 & 0 \\ M_{21} & 0 & M_{23} \\ 0 & M_{32} & 0 \end{bmatrix}.$$

Note that if we multiply two system matrices, say C and A, then the product

$$CA = \begin{bmatrix} C_{11} & 0 & 0 \\ C_{21} & 0 & C_{23} \\ 0 & C_{32} & 0 \end{bmatrix} \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & 0 & A_{23} \\ 0 & A_{32} & 0 \end{bmatrix} = \begin{bmatrix} C_{11}A_{11} & 0 & 0 \\ C_{21}A_{11} & C_{23}A_{32} & 0 \\ C_{32}A_{21} & 0 & C_{32}A_{23} \end{bmatrix}$$

does not necessarily have the same block zero structure as the system matrices.

Example 5.1.2 shows that, in general, the moments

$$CA^kB \quad \text{for} \quad k = 1, \ldots, p$$

of a structured linear system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ do not necessarily have the same block zero pattern as the system matrices.

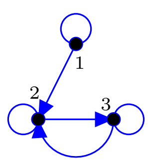
We now investigate various properties of the underlying binary relation $\mathcal{T} = (P, T)$ and investigate the effect on the associated digraph and structure of the system matrices. The various properties considered correspond to various sub-classes of structured linear systems.

Recall from Subsection 2.1 that a binary relation $\mathcal{T} = (P, T)$ is reflexive if $(i, i) \in T$ for each $i \in P$. Structured linear systems $\Sigma_{\mathcal{T}}$ whose underlying binary relation $\mathcal{T}$ is reflexive have entries on the main diagonals of the system matrices. The associated digraph of $\mathcal{T}$ has self loops at each node. Reflexivity implies that the information at each subsystem is communicated to itself (or alternatively, is available to itself). Thus it is natural to assume reflexivity.

Example 5.1.3.
In this example, we make the binary relation $\mathcal{T}$ in Example 5.1.2 reflexive. Hence, let

$$T = \{(1, 1); (1, 2); (2, 2); (2, 3); (3, 2); (3, 3)\}.$$

The corresponding digraph and block sparsity pattern of a system matrix M of a structured linear system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ is shown below.



$$M = \begin{bmatrix} M_{11} & 0 & 0 \\ M_{21} & M_{22} & M_{23} \\ 0 & M_{32} & M_{33} \end{bmatrix}.$$

Again, if we multiply two system matrices, say C and A, then the product

$$CA = \begin{bmatrix} C_{11} & 0 & 0 \\ C_{21} & C_{22} & C_{23} \\ 0 & C_{32} & C_{33} \end{bmatrix} \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} C_{11}A_{11} & 0 & 0 \\ C_{21}A_{11} + C_{22}A_{21} & C_{22}A_{22} + C_{23}A_{32} & C_{22}A_{23} + C_{23}A_{33} \\ C_{32}A_{21} & C_{32}A_{22} + C_{33}A_{32} & C_{32}A_{23} + C_{33}A_{33} \end{bmatrix}$$

does not necessarily have the same block zero structure as the system matrices.

A binary relation $\mathcal{T} = (P, T)$ is transitive if for all $i, j, k \in P$, $(i, j) \in T$ and $(j, k) \in T$, then $(i, k) \in P$. A binary relation that is both reflexive and transitive is a pre-order. Structured linear systems $\Sigma_{\mathcal{T}}$ whose underlying binary relation $\mathcal{T}$ is transitive have directed edges from i to k if there are directed edges from i to j and j to k. Inductively, this then also means that if there is a path from i to j, then there is an edge from i to j. Transitivity implies that no information is lost when a subsystem i communicates with a subsystem k via a subsystem j that lies between i and k. This is also a reasonable assumption to make. As was mentioned in Section 2.3, it can be shown that the set $\mathcal{I}_{\mathcal{T}}$ is closed under multiplication if and only T is transitive (see for example p.258 in [7]). Hence structured linear system $\Sigma_{\mathcal{T}}$ whose underlying binary relation $\mathcal{T}$ is transitive have system matrices whose block zero pattern is preserved under multiplication. Hence if $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ and $\mathcal{T}$ is a pre-order then

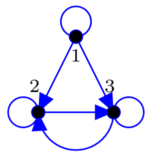
$$CA^kB \in \mathcal{I}_{\mathcal{P}}^{\underline{r} \times \underline{m}},$$

that is, the moments also have the prescribed block zero pattern.

Example 5.1.4.
In this example, we make the binary relation $\mathcal{T}$ in Example 5.1.3 transitive. Note that there $(1, 2) \in T$ and $(2, 3) \in T$, but $(1, 3) \notin T$. In order to make $\mathcal{T}$ transitive, we add $(1, 3)$ to T, that is,

$$T = \{(1, 1); (1, 2), (1, 3); (2, 2); (2, 3); (3, 2); (3, 3)\}.$$

The corresponding digraph and block sparsity pattern of a system matrix M of a structured linear system $\Sigma_{\mathcal{T}} \sim (A, B, C, D)$ is shown below.



$$M = \begin{bmatrix} M_{11} & 0 & 0 \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}.$$

If we now multiply two system matrices, say C and A, then the product

$$\begin{aligned} CA &= \begin{bmatrix} C_{11} & 0 & 0 \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \\ &= \begin{bmatrix} C_{11}A_{11} & 0 & 0 \\ C_{21}A_{11} + C_{22}A_{21} + C_{23}A_{31} & C_{22}A_{22} + C_{23}A_{32} & C_{22}A_{23} + C_{23}A_{33} \\ C_{31}A_{11} + C_{32}A_{21} + C_{33}A_{31} & C_{32}A_{22} + C_{33}A_{32} & C_{32}A_{23} + C_{33}A_{33} \end{bmatrix} \end{aligned}$$

indeed has the same block zero structure as the system matrices.

A binary relation $\mathcal{T} = (P, T)$ is anti-symmetric if $(i, j) \in T$ and $(j, i) \in T$ implies that $i = j$. A binary relation T that is reflexive, transitive and anti-symmetric is a partial order $\succeq$ and in that case we write $i \succeq j$ if $(i, j) \in T$. Structured linear systems $\Sigma_{\mathcal{T}}$ whose underlying binary relation $\mathcal{T}$ is a poset are called poset-causal systems and are studied in the next section. With posets it is customary to consider the Hasse diagram (see Subsection 2.1) associated with the poset rather than the digraph. Recall that the Hasse diagram omits directed edges that correspond to reflexivity and transitivity. If in addition to reflexivity and transitivity, $\mathcal{T}$ is also anti-symmetric (that is if $\mathcal{T}$ is a poset), then the system matrices of $\Sigma_{\mathcal{T}}$ can always be written as block lower triangular matrices and the associated Hasse diagram of $\mathcal{T}$ contains no directed cycles. We note that in Example 5.1.4, the system matrices are not block lower triangular. This was due to the fact that $\mathcal{T}$ was not anti-symmetric ($(2, 3) \in T$ and $(3, 2) \in T$, but $2 \neq 3$).

Example 5.1.5.
In this example, we make the binary relation $\mathcal{T}$ in Example 5.1.4 anti-symmetric. Let $2 = 3$. Then $P = \{1, 2\}$ and

$$T = \{(1, 1); (1, 2), (2, 2)\}.$$

The corresponding Hasse diagram and block sparsity pattern of a system matrix M of a structured linear system $\Sigma_{\mathcal{T}} \sim (A, B, C, D)$ is shown below.

$$M = \begin{bmatrix} M_{11} & 0 \\ M_{21} & M_{22} \end{bmatrix}$$

The system matrices are block lower triangular.

The structured linear system in Example 5.1.5 is a leader-follower system.

Recall from Subsection 2.1 that a partial order $\succeq$ is in-ultra transitive if for all $i, j, k \in P$, $i \succeq j$ and $k \succeq j$ implies that $k \succeq i$ or $k \succeq i$. Structured linear systems $\Sigma_{\mathcal{T}}$ whose underlying binary relation $\mathcal{T}$ is a partial order and in-ultra transitive are called coordinated linear systems and are studied in Section 5.3. As mentioned in Section 2.1, the Hasse diagrams of in-ultra transitive partial orders are out-forests.

5.1.1 Quadratic Invariance

In standard control problems, decentralization manifests by requiring the controller to be designed to be in some constraint set. Thus a canonical decentralized control problem with plant with transfer function $\widehat{G}$ and controller with transfer function $\widehat{K}$, is to minimize some norm of the closed loop transfer function $\underline{\mathcal{F}}(\widehat{G}, \widehat{K})$ subject to a subspace constraint as follows [35]. In our case, we take the $\mathcal{H}_2$-norm:

$$\begin{aligned} \text{minimize} \quad & \|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2 \\ \text{subject to} \quad & \widehat{K} \text{ stabilizes } \widehat{G} \\ & \widehat{K} \in S. \end{aligned}$$

For a general LTI plant and a general subspace S, there is no known tractable algorithm for computing the optimal controller $\widehat{K}$. The paper [35] gives a condition, known as quadratic invariance, under which the above problem may be recast as a convex optimization problem. Quadratic invariance is a algebraic condition which relates the the plant $\widehat{G}$ to the constraint set S. Let $\mathcal{R}_p^{r\times m}$ and $\mathcal{R}_{sp}^{r\times m}$ be the sets of proper and strictly proper rational matrix functions of size $r \times m$ respectively. The sizes are omitted if they are clear from the context or are not of particular importance.

Definition 5.1.6 (Quadratic invariance)**.**
Given a transfer function $\widehat{G} \in \mathcal{R}_{sp}^{r\times m}$ and a subspace $S \subseteq \mathcal{R}_p^{m\times r}$, the set $\mathcal{S}$ is called quadratically invariant under $\widehat{G}$ if

$$\widehat{K}\widehat{G}\widehat{K} \in S \qquad \text{for all} \qquad \widehat{K} \in \mathcal{S}.$$

Define as in [35] the map $h : \mathcal{R}_{sp}^{r\times m} \times \mathcal{R}_p^{m\times r} \to \mathcal{R}_{sp}^{m\times r}$ by

$$h(\widehat{G}, \widehat{K}) = -\widehat{K}(I - \widehat{G}\widehat{K})^{-1}.$$

For a plant

$$\widehat{G} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix} \tag{5.4}$$

it follows that

$$\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{K}(I - \widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21} = \widehat{G}_{11} - \widehat{G}_{12}[h(\widehat{G}_{22}, \widehat{K})]\widehat{G}_{21}.$$

From the Youla parametrization, Rotkowitz and Lall obtain the following result.

Theorem 5.1.7 (cf. Theorem 19 and p.281 in [35])**.**
Given a plant as in (5.4), suppose $\widehat{G}_{22} \in \mathcal{R}_{sp}$ and $S \subseteq \mathcal{R}_p$ is a closed subspace. If S is quadratically invariant under $\widehat{G}_{22}$ and $\widehat{K}_0 \in \mathcal{RH}_\infty \cap S$ stabilizes $\widehat{G}_{22}$, then $\widehat{K}$ is an optimal solution to

$$\begin{aligned} \textit{minimize} \quad & \|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2 \\ \textit{subject to} \quad & \widehat{K} \textit{ stabilizes } \widehat{G} \\ & \widehat{K} \in \mathcal{S} \end{aligned}$$

if and only if

$$\widehat{K} = \widehat{K}_0 - h(h(\widehat{K}_0, \widehat{G}_{22}), \widehat{R})$$

where $\widehat{R}$ is the optimal solution to

$$\begin{aligned} \textit{minimize} \quad & \|\widehat{T}_1 + \widehat{T}_2\widehat{R}\widehat{T}_3\|_2 \\ \textit{subject to} \quad & \widehat{R} \in \mathcal{RH}_\infty \\ & \widehat{R} \in \mathcal{S}, \end{aligned}$$

and where

$$\widehat{T}_1 = \underline{\mathcal{F}}(\widehat{G}, \widehat{K}_0) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{K}_0(I - \widehat{G}_{22}\widehat{K}_0)^{-1}\widehat{G}_{21}$$
$$\widehat{T}_2 = \widehat{G}_{12}\widehat{K}_0(I - \widehat{K}_0\widehat{G}_{22})^{-1} \quad \textit{and}$$
$$\widehat{T}_3 = (I - \widehat{G}_{22}\widehat{K}_0)^{-1}\widehat{G}_{21}.$$

As mentioned in [38], the importance of the above theorem is that the reparameterized optimization problem in the parameter $\widehat{R}$ is a convex optimization problem. We now consider which structured linear systems satisfy a quadratic invariance criteria. Suppose $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ is a structured linear system. Then

$$A \in \mathcal{I}_{\mathcal{T}}^{\underline{n}\times\underline{n}}, \quad B \in \mathcal{I}_{\mathcal{T}}^{\underline{n}\times\underline{m}}, \quad C \in \mathcal{I}_{\mathcal{T}}^{\underline{r}\times\underline{n}}, \quad D \in \mathcal{I}_{\mathcal{T}}^{\underline{r}\times\underline{m}}.$$

If $\mathcal{T}$ is a pre-order, then by Proposition 2.3.7 $\widehat{G}(\lambda) = C(\lambda I - A)^{-1}B + D \in \mathcal{I}_{\mathcal{T}}^{\underline{r}\times\underline{m}}$ for each $\lambda \in \rho(A)$. We will write $\widehat{G} \in \mathcal{I}_{\mathcal{T}}^{\underline{r}\times\underline{m}}$ for short. Again by Proposition 2.3.7, if $\widehat{K} \in \mathcal{T}_{\mathcal{T}}^{\underline{m}\times\underline{r}}$, then

$$\widehat{K}\widehat{G}\widehat{K} \in \mathcal{T}_{\mathcal{T}}^{\underline{m}\times\underline{r}}.$$

Thus if $\mathcal{T}$ is a pre-order and $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ is a structured linear system with transfer function $\widehat{G}$, then $\mathcal{T}_{\mathcal{T}}^{\underline{m}\times\underline{r}}$ is quadratically invariant under $\widehat{G}$.

5.2 Poset-Causal Systems

In this section, we study structured linear systems whose underlying binary relation is a poset. This class of systems was defined by Shah in [38] as poset-causal systems. This will be the class of structured linear systems that we will study in the most detail. We refer the reader back to Chapter 2 for the notation used in this section.

Definition 5.2.1 (Poset-causal system)**.**
Let $\mathcal{P} = (P, \succeq)$ be a poset with $P = \{1, \ldots, p\}$. A poset-causal system (with underlying poset $\mathcal{P}$) is a LTI system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0,$$
$$y(t) = Cx(t) + Du(t), \qquad t \geq 0,$$

with structured system matrices

$$A \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{n}}, \qquad B \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}, \qquad C \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{n}}, \qquad D \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{m}}, \tag{5.5}$$

for some partitions $\underline{n}, \underline{m}, \underline{r} \in \mathbb{Z}_+^p$ and some initial state $x_0 \in \mathcal{X} = \mathbb{R}^{\underline{n}}$.

Equivalently, the poset-causal system $\Sigma_{\mathcal{P}}$ is given by the interconnected equations

$$\dot{x}_i(t) = \sum_{j\in\uparrow i} A_{ij}x_j(t) + \sum_{j\in\uparrow i} B_{ij}u_j(t), \quad x_i(0) = x_{i,0},$$
$$y_i(t) = \sum_{j\in\uparrow i} C_{ij}x_j(t) + \sum_{j\in\uparrow i} D_{ij}u_j(t), \quad t \geq 0,$$

for $i \in P$, determined by the non-zero blocks in the system matrices (5.5) and the components of the initial state and with local input, state and output spaces of dimensions

$$\dim(\mathcal{U}_i) = m_i, \quad \dim(\mathcal{X}_i) = n_i, \quad \dim(\mathcal{Y}_i) = r_i.$$

In short, we will write $\Sigma_{\mathcal{P}} \sim (A, B, C, D, x_0; \mathcal{P})$ to indicate the poset-causal system $\Sigma_{\mathcal{P}}$, or usually, $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ when the poset is clear from the context and the initial state is either clear from the context or unspecified.

Given a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$, we have by definition that $A \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{n}}$. Thus $A(i,j) = 0$ if $j \not\succeq i$ and consequently for the transpose A^T we have $A^\mathsf{T}(j,i) = 0$ if $j \not\succeq i$. Equivalently, in the dual poset, $A^\mathsf{T}(j,i) = 0$ if $i \not\succeq_d j$. Consequently, for a poset-causal poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$, it holds that

$$A^\mathsf{T} \in \mathcal{I}_{\mathcal{P}_d}^{\underline{n} \times \underline{n}}, \quad B^\mathsf{T} \in \mathcal{I}_{\mathcal{P}_d}^{\underline{m} \times \underline{n}}, \quad C^\mathsf{T} \in \mathcal{I}_{\mathcal{P}_d}^{\underline{n} \times \underline{r}}, \quad D^\mathsf{T} \in \mathcal{I}_{\mathcal{P}_d}^{\underline{m} \times \underline{r}}.$$

This observation justifies the following definition.

Definition 5.2.2 (Dual poset-causal system)**.**
For a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$, its *dual system* is defined to be the poset-causal system $\Sigma_{\mathcal{P}_d} \sim (A^\mathsf{T}, C^\mathsf{T}, B^\mathsf{T}, D^\mathsf{T})$.

For a finite dimensional space $\mathcal{V} = \bigoplus_{j \in P} \mathcal{V}_j$ and $S \subset P$ we define

$$\mathcal{V}_S := \bigoplus_{j \in S} \mathcal{V}_j \subseteq \mathcal{V}. \tag{5.6}$$

Specifically, the following spaces will play an important role

$$\mathcal{X}_{\downarrow i} = \bigoplus_{j \in \downarrow i} \mathcal{X}_j, \quad \mathcal{Y}_{\downarrow i} = \bigoplus_{j \in \downarrow i} \mathcal{Y}_j, \quad \mathcal{X}_{\uparrow i} = \bigoplus_{j \in \uparrow i} \mathcal{X}_j \quad \text{and} \quad \mathcal{U}_{\uparrow i} = \bigoplus_{j \in \uparrow i} \mathcal{U}_j.$$

In our analysis of poset-causal systems $\Sigma_{\mathcal{P}}$, various derived systems play a role. Firstly, the *global system* is just the overall classical state space system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0 \in \mathcal{X}, \\ y(t) &= Cx(t) + Du(t), \quad t \geq 0 \end{aligned}$$

with state and output given by

$$\begin{aligned} x(t) &= x(x_0, u, t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\, \mathrm{d}\tau, \\ y(t) &= y(x_0, u, t) = Ce^{At}x_0 + \int_0^t Ce^{A(t-\tau)}Bu(t)\, \mathrm{d}\tau + Du(t). \end{aligned}$$

Other derived systems are determined by a fixed choice of $i \in P$. We define the *i-th local system* where we only consider the impact of the local input u_i on the subsystem i:

$$\begin{aligned} \dot{x}^i(t) &= A_{ii}x^i(t) + B_{ii}u_i(t), \qquad x^i(0) = x_0^i \in \mathcal{X}_i, \\ y^i(t) &= C_{ii}x^i(t) + D_{ii}u_i(t), \qquad t \geq 0. \end{aligned} \tag{5.7}$$

Here, the state x^i and output y^i at some final time $t > 0$ are given by

$$\begin{aligned} x^i(t) &= x^i(x_0^i, u_i, t) = e^{A_{ii}(t)}x_0^i + \int_0^t e^{A_{ii}(t-\tau)}B_{ii}u_i(\tau)\, \mathrm{d}\tau, \\ y^i(t) &= y^i(x_0^i, u_i, t) = C_{ii}e^{A_{ii}(t)}x_0^i + \int_0^t C_{ii}e^{A_{ii}(t-\tau)}B_{ii}u_i(\tau)\, \mathrm{d}\tau + D_{ii}u_i(t). \end{aligned}$$

Next, for the *i-th downstream system* one considers the impact of the i-th input u_i on the local states that are downstream from subsystem i:

$$\begin{aligned} \dot{x}^{\downarrow i}(t) &= A(\downarrow i, \downarrow i)x^{\downarrow i}(t) + B(\downarrow i, i)u_i(t), \qquad x^{\downarrow i}(0) = x_0^{\downarrow i} \in \mathcal{X}_{\downarrow i}, \\ y^{\downarrow i}(t) &= C(\downarrow i, \downarrow i)x^{\downarrow i}(t) + D(\downarrow i, i)u_i(t), \qquad t \geq 0. \end{aligned} \tag{5.8}$$

Note that the state and output signals, $x^{\downarrow i}$ and $y^{\downarrow i}$, take values in the spaces $\mathcal{X}_{\downarrow i}$ and $\mathcal{Y}_{\downarrow i}$, respectively. In this case the state $x^{\downarrow i}$ and output $y^{\downarrow i}$ at some final time $t > 0$ are given by

$$\begin{aligned} x^{\downarrow i}(t) = x^{\downarrow i}(x_0^{\downarrow i}, u_i, t) &= e^{A(\downarrow i,\downarrow i)t}x_0^{\downarrow i} + \int_0^t e^{A(\downarrow i,\downarrow i)(t-\tau)}B(\downarrow i, i)u_i(\tau)\ \mathrm{d}\tau, \\ y^{\downarrow i}(t) = y^{\downarrow i}(x_0^{\downarrow i}, u_i, t) &= C(\downarrow i,\downarrow i)e^{A(\downarrow i,\downarrow i)t}x_0^{\downarrow i} + D(\downarrow i, i)u_i(t) + \\ &\quad + \int_0^t C(\downarrow i,\downarrow i)e^{A(\downarrow i,\downarrow i)(t-\tau)}B(\downarrow i, i)u_i(\tau)\ \mathrm{d}\tau. \end{aligned} \tag{5.9}$$

By Corollary 2.3.12 it follows that the system matrices partition as

$$A = \begin{bmatrix} A(\downarrow i,\downarrow i) & * \\ 0 & * \end{bmatrix},\ B = \begin{bmatrix} B(\downarrow i,\downarrow i) & * \\ 0 & * \end{bmatrix},\ C = \begin{bmatrix} C(\downarrow i,\downarrow i) & * \\ 0 & * \end{bmatrix},\ D = \begin{bmatrix} D(\downarrow i,\downarrow i) & * \\ 0 & * \end{bmatrix},$$

with $*$ indicating unspecified entries.

Lastly, for the *i-th upstream system*, one considers the impact of the system on the i-th output component generated by the subsystems that are upstream of the i-th subsystem:

$$\begin{aligned} \dot{x}^{\uparrow i}(t) &= A(\uparrow i,\uparrow i)x^{\uparrow i}(t) + B(\uparrow i,\uparrow i)u^{\uparrow i}(t), \qquad x^{\uparrow i}(0) = x_0^{\uparrow i} \in \mathcal{X}_{\uparrow i}, \\ y^{\uparrow i}(t) &= C(i,\uparrow i)x^{\uparrow i}(t) + D(i,\uparrow i)u^{\uparrow i}(t), \qquad t \geq 0. \end{aligned} \tag{5.10}$$

where the input and state signals $u^{\uparrow i}$ and $x^{\uparrow i}$ take values in $\mathcal{U}_{\uparrow i}$ and $\mathcal{X}_{\uparrow i}$, respectively, while the output signal $y^{\uparrow i}$ takes values in $\mathcal{Y}_i$. In this case, applying Corollary 2.3.12 to the dual system, it follows that system matrices partition as

$$A = \begin{bmatrix} A(\uparrow i,\uparrow i) & 0 \\ * & * \end{bmatrix},\ B = \begin{bmatrix} B(\uparrow i,\uparrow i) & 0 \\ * & * \end{bmatrix},\ C = \begin{bmatrix} C(\uparrow i,\uparrow i) & 0 \\ * & * \end{bmatrix},\ D = \begin{bmatrix} D(\uparrow i,\uparrow i) & 0 \\ * & * \end{bmatrix},$$

with $*$ indicating unspecified entries. As a consequence we see that the state and output of the i-th upstream system are easily obtained from the global system via:

$$x^{\uparrow i}(x_0^{\uparrow i}, u^{\uparrow i}, t) = I(\uparrow i, :)x(x_0, u, t) \quad \text{and} \quad y^{\uparrow i}(x_0^{\uparrow i}, u^{\uparrow i}, t) = y_i(x_0, u, t),$$

with y_i the i-th component of the global output signal y and where x_0 and u can be any initial state and input satisfying

$$x_0^{\uparrow i} = I(\uparrow i, :)x_0 \quad \text{and} \quad u^{\uparrow i}(t) = I(\uparrow i, :)u(t).$$

The relation between the signals of the i-th downstream system and the global and local systems is less straightforward.

Lemma 5.2.3.
Consider a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$, a given input $u = \bigoplus_{i\in P} u_i$ and initial state $x_0 = \bigoplus_{i\in P} x_{i,0}$. Set $x_0^{\downarrow i} := \bigoplus_{j\in\downarrow i} x_{j,0} \in \mathcal{X}_{\downarrow i}$ and $\widetilde{x}_{0,i} := I_{\underline{n}}(\downarrow i, i)x_{i,0} \in \mathcal{X}_{\downarrow i}$ for all $i \in P$. Then

$$\begin{aligned} x(x_0, u, t) &= \sum_{i\in P} I_{\underline{n}}(:, \downarrow i)x^{\downarrow i}(\widetilde{x}_{0,i}, u_i, t) \quad \text{and} \\ y(x_0, u, t) &= \sum_{i\in P} I_{\underline{n}}(:, \downarrow i)y^{\downarrow i}(\widetilde{x}_{0,i}, u_i, t), \quad t \geq 0. \end{aligned} \tag{5.11}$$

Furthermore, for all $i \in P$ we have

$$x_i^{\downarrow i}(x_0^{\downarrow i}, u_i, t) = x^i(x_{i,0}, u_i, t) \quad \text{and} \quad y_i^{\downarrow i}(u_i, t) = y^i(x_0^i, u_i, t), \quad t \geq 0, \tag{5.12}$$

where $x_j^{\downarrow i}(x_0^{\downarrow i}, u_i, t)$ is the component of $x^{\downarrow i}(x_0^{\downarrow i}, u_i, t)$ taking values in $\mathcal{X}_j$ and where $y_j^{\downarrow i}(x_0^{\downarrow i}, u_i, t)$ is the component of $y^{\downarrow i}(x_0^{\downarrow i}, u_i, t)$ taking values in $\mathcal{Y}_j$. In particular, for all $i \in P$ we have

$$\begin{aligned} x_i(x_0, u, t) &= x^i(x_{i,0}, u_i, t) + \sum_{j \in \not\uparrow i} x_i^{\downarrow j}(\widetilde{x}_{j,0}, u_j, t), \\ y_i(x_0, u, t) &= y^i(x_{i,0}, u_i, t) + \sum_{j \in \not\uparrow i} y_i^{\downarrow j}(\widetilde{x}_{j,0}, u_j, t), \quad t \geq 0. \end{aligned} \tag{5.13}$$

Proof.
Firstly we note that for any $i \in P$, we have $\downarrow i = \{i\} \cup \not\downarrow i$ and that if $j \in \not\downarrow i$, then $j \not\succeq i$. In particular, for $j \in \not\downarrow i$ we have $A_{ij} = 0$, $B_{ij} = 0$, $C_{ij} = 0$, $D_{ij} = 0$ and $e^A(i, j) = 0$. This implies that

$$I_{\underline{n}}(i, \downarrow i)A(\downarrow i, \downarrow i) = A_{ii} I_{\underline{n}}(i, \downarrow i) \quad \text{and} \quad I_{\underline{n}}(i, \downarrow i)C(\downarrow i, \downarrow i) = C_{ii} I_{\underline{n}}(i, \downarrow i). \tag{5.14}$$

The first identity yields $I_{\underline{n}}(i, \downarrow i)e^{A(\downarrow i, \downarrow i)t} = e^{A_{ii}t} I_{\underline{n}}(i, \downarrow i)$. Furthermore, by Theorem 2.3.10 we have

$$I_{\underline{n}}(i, \downarrow i)B(\downarrow i, i) = B(i, i) = B_{ii} \quad \text{and} \quad I_{\underline{n}}(i, \downarrow i)D(\downarrow i, i) = D(i, i) = D_{ii},$$

while Corollary 2.3.12 implies that $e^{At} I_{\underline{n}}(:, \downarrow i) = I_{\underline{n}}(:, \downarrow i)e^{A(\downarrow i, \downarrow i)t}$. Next observe that

$$\begin{aligned} x(x_0, u, t) &= e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\ \mathrm{d}\tau \\ &= e^{At} \sum_{i \in P} I_{\underline{n}}(:, i)x_{i,0} + \int_0^t e^{A(t-\tau)} \sum_{i \in P} B(:, i)u_i(\tau)\ \mathrm{d}\tau \\ &= \sum_{i \in P} e^{At} I_{\underline{n}}(:, i)x_{i,0} + \int_0^t e^{A(t-\tau)}B(:, i)u_i(\tau)\ \mathrm{d}\tau \\ &= \sum_{i \in P} e^{At} I_{\underline{n}}(:, \downarrow i)I_{\underline{n}}(\downarrow i, i)x_{i,0} + \int_0^t e^{A(t-\tau)} I_{\underline{n}}(:, \downarrow i)B(\downarrow i, i)u_i(\tau)\ \mathrm{d}\tau \\ &= \sum_{i \in P} I_{\underline{n}}(:, \downarrow i)e^{A(\downarrow i, \downarrow i)t}\widetilde{x}_{i,0} + \int_0^t I_{\underline{n}}(:, \downarrow i)e^{A(\downarrow i, \downarrow i)(t-\tau)}B(\downarrow i, i)u_i(\tau)\ \mathrm{d}\tau \\ &= \sum_{i \in P} I_{\underline{n}}(:, \downarrow i)\Big(e^{A(\downarrow i, \downarrow i)t}\widetilde{x}_{i,0} + \int_0^t e^{A(\downarrow i, \downarrow i)(t-\tau)}B(\downarrow i, i)u_i(\tau)\ \mathrm{d}\tau\Big) \\ &= \sum_{i \in P} I_{\underline{n}}(:, \downarrow i)x^{\downarrow i}(\widetilde{x}_{0,i}, u_i, t). \end{aligned}$$

Hence the identity for $x(x_0, u, t)$ in (5.11) holds. A similar argument also gives the identity for $y(x_0, u, t)$ in (5.11).

In order to prove the two identities in (5.12), we consider the i-th components of the solutions given in (5.9):

$$\begin{aligned} x_i^{\downarrow i}(x_0^{\downarrow i}, u_i, t) &= I_{\underline{n}}(i, \downarrow i)x^{\downarrow i}(x_0^{\downarrow i}, u_i, t) \\ &= I_{\underline{n}}(i, \downarrow i)e^{A(\downarrow i, \downarrow i)t}x_0^{\downarrow i} + I_{\underline{n}}(i, \downarrow i) \int_0^t e^{A(\downarrow i, \downarrow i)(t-\tau)}B(\downarrow i, i)u_i(\tau)\ \mathrm{d}\tau \\ &= e^{A_{ii}t}x_0^i + \int_0^t e^{A_{ii}(t-\tau)}B_{ii}u_i(\tau)\ \mathrm{d}\tau = x^i(u_i, t). \end{aligned}$$

A similar computation gives the identity for $y_i^{\downarrow i}(x_0^{\downarrow i}, u_i, t)$. The two identities in (5.13) follow by combining (5.11) and (5.12) noting that in (5.12) only the i-th component of the initial state $x_0^{\downarrow i}$ is relevant, so that $x_0^{\downarrow i}$ in the left hand sides of both equations may be replaced by $\widetilde{x}_{i,0}$. □

5.2.1 Systems over Graphs

We now briefly describe systems over graphs as in [24] and show that they are poset-causal systems. Systems over graphs are defined to be continuous times LTI system satisfying an additional structure condition related to a digraph. Let $\mathcal{R}_p$ be the set of proper rational matrix functions. Consider a digraph $\mathcal{G} = (P, E)$ where $P = \{1, \ldots, p\}$ is the set of vertices (nodes) and $E \subseteq P \times P$ is the set of edges. Write $i \to j$ if $(i, j) \in E$. An index set n is defined to be a p-tuple $(n_1, \ldots, n_p)$ of non-negative numbers. For index sets r and m and a commutative ring F, the set $\mathcal{S}_\mathcal{G}(F, n, m)$ is defined to be the set of matrices

$$A = \begin{bmatrix} A_{11} & \ldots & A_{1p} \\ \vdots & \ddots & \vdots \\ A_{p1} & \ldots & A_{pp} \end{bmatrix}$$

such that $A_{ij} \in F^{r_i \times m_j}$ if $j \to i$ and $A_{ij} = 0$ otherwise. Given a digraph $\mathcal{G}$, index sets r and m, a system over the graph $\mathcal{G}$ is a continuous time LTI system with a proper transfer function $\widehat{G} \in \mathcal{S}_\mathcal{G}(\mathcal{R}_p, r, m)$. The following assumptions are made regarding the digraph $\mathcal{G}$:

1. $i \to i$ for all $i \in P$;
2. if $i \to j$ and $j \to k$, then $i \to k$;
3. there are no directed cycles of length 2 or greater.

Define a binary relation $\succeq \subseteq P \times P$ by

$$i \succeq j \quad \text{if and only if} \quad i \to j. \tag{5.15}$$

Then $\succeq$ is a partial order on P:

1. $i \succeq i$ for all $i \in P$, so $\succeq$ is reflexive;
2. if $i \succeq j$ and $j \succeq k$, then $i \succeq k$, so $\succeq$ is transitive;
3. if there is no directed cycle of length 2 or greater, then it cannot hold that $i \to j$ and $j \to i$, except if $i = j$, thus if $i \succeq j$ and $j \succeq i$, then $i = j$, that is $\succeq$ is anti-symmetric.

Thus a digraph satisfying the given assumptions corresponds exactly with the poset $\mathcal{P} = (P, \succeq)$ with $\succeq$ defined as in (5.15). The index sets r and m correspond to what we have called partitions $\underline{r} = (r_1, \ldots, r_p)$ and $\underline{m} = (m_1, \ldots, r_m)$. The set $\mathcal{S}_\mathcal{G}(\mathcal{R}_p, r, m)$ corresponds to our block incidence set $\mathcal{I}_\mathcal{P}^{\underline{r} \times \underline{m}}$. With these correspondences, it is clear that systems over graphs are poset-causal systems.

5.3 Coordinated Linear Systems

Coordinated linear systems are studied in the thesis [20] and the articles [23, 22, 21]. In this section, we show that coordinated linear systems, in the general sense, are poset-causal systems in which the partial order, in addition to being reflexive, transitive and anti-symmetric, satisfies an additional property.

In the thesis [20] a coordinated linear system with one coordinator and two subsystems is defined as follows. Consider a continuous-time linear-time-invariant (LTI) system Σ

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ z(t) &= Cx(t) \end{aligned} \tag{5.16}$$

Let the state space, input and output spaces have the decompositions

$$X = X_c \dot{+} X_1 \dot{+} X_2, \qquad U = U_c \dot{+} U_1 \dot{+} U_2, \qquad Z = Z_c \dot{+} Z_1 \dot{+} Z_2.$$

The state-, input- and output spaces of the coordinator are denoted by X_c, U_c and Y_c respectively and their dimensions are n_c, m_c and r_c respectively. The state-, input- and output spaces of subsystem 1 are denoted by X_1, U_1 and Y_1 and the state, input and output spaces of subsystem 2 are denoted by X_2, U_2 and Y_2 respectively and they have dimensions n_i, m_i and r_i for $i = 1, 2$. The overall state-, input- and output vectors are of the form

$$x = \begin{bmatrix} x_c \\ x_1 \\ x_2 \end{bmatrix}, \qquad u = \begin{bmatrix} u_c \\ u_1 \\ u_2 \end{bmatrix}, \qquad \text{and} \qquad z = \begin{bmatrix} z_c \\ z_1 \\ z_2 \end{bmatrix}. \tag{5.17}$$

This then also partitions the system matrices A, B and C as follows

$$A = \begin{bmatrix} A_{cc} & A_{c1} & A_{c2} \\ A_{1c} & A_{11} & A_{12} \\ A_{2c} & A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{cc} & B_{c1} & B_{c2} \\ B_{1c} & B_{11} & B_{12} \\ B_{2c} & B_{21} & B_{22} \end{bmatrix}, \quad C = \begin{bmatrix} C_{cc} & C_{c1} & C_{c2} \\ C_{1c} & C_{11} & C_{12} \\ C_{2c} & C_{21} & C_{22} \end{bmatrix}. \tag{5.18}$$

Replacing the decompositions (5.17) and (5.18) into the overall state space equations (5.16), we get the following state space equations at subsystem level: for subsystem 1

$$\begin{aligned} \dot{x}_1(t) &= A_{1c}x_c(t) + A_{11}x_1(t) + A_{12}x_2(t) + B_{1c}u_c(t) + B_{11}u_1(t) + B_{12}u_2(t) \\ z_1(t) &= C_{1c}x_c(t) + C_{11}x_1(t) + C_{12}x_2(t), \end{aligned}$$

for subsystem 2

$$\begin{aligned} \dot{x}_2(t) &= A_{2c}x_c(t) + A_{21}x_1(t) + A_{22}x_2(t) + B_{2c}u_c(t) + B_{21}u_1(t) + B_{22}u_2(t) \\ z_2(t) &= C_{2c}x_c(t) + C_{21}x_1(t) + C_{22}x_2(t) \end{aligned}$$

and for the coordinator

$$\begin{aligned} \dot{x}_c(t) &= A_{cc}x_c(t) + A_{c1}x_1(t) + A_{c2}x_2(t) + B_{cc}u_c(t) + B_{c1}u_1(t) + B_{c2}u_2(t) \\ z_c(t) &= C_{cc}x_c(t) + C_{c1}x_1(t) + C_{c2}x_2(t). \end{aligned}$$

The system Σ is called a *coordinated linear system* if the following conditions hold:

1. $AX_1 \subseteq X_1$ and $AX_2 \subseteq X_2$ (X_1 and X_2 are A-invariant),
2. $BU_1 \subseteq X_1$ and $BU_2 \subseteq X_2$,
3. $CX_1 \subseteq Z_1$ and $CX_2 \subseteq Z_2$.

According to requirement 1 above, we must have that $A_{21} = 0$, $A_{c1} = 0$, $A_{12} = 0$ and $A_{c2} = 0$. By requirement 2, $B_{21} = 0$, $B_{c1} = 0$, $B_{12} = 0$ and $B_{c2} = 0$ and by requirement 3, $C_{21} = 0$, $C_{c1} = 0$, $C_{12} = 0$ and $C_{c2} = 0$. Thus a coordinated linear system with one coordinator and two subsystems is a system of the form

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} A_{cc} & 0 & 0 \\ A_{1c} & A_{11} & 0 \\ A_{2c} & 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_c(t) \end{bmatrix} + \begin{bmatrix} B_{cc} & 0 & 0 \\ B_{1c} & B_{11} & 0 \\ B_{2c} & 0 & B_{22} \end{bmatrix} \begin{bmatrix} u_c(t) \\ u_1(t) \\ u_2(t) \end{bmatrix}, \\ z(t) &= \begin{bmatrix} C_{cc} & 0 & 0 \\ C_{1c} & C_{11} & 0 \\ C_{2c} & 0 & C_{22} \end{bmatrix} \begin{bmatrix} x_c(t) \\ x_1(t) \\ x_2(t) \end{bmatrix}. \end{aligned}$$

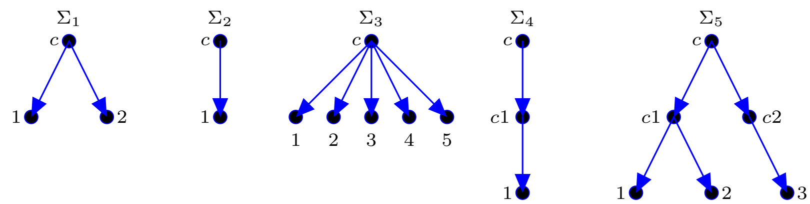
We can also consider coordinated linear systems as poset-causal systems with underlying poset $\mathcal{P} = (P, \succeq)$ where $P = \{c, 1, 2\}$. The requirement $AX_1 \subseteq X_1$ is equivalent to $A_{c1} = 0$ and $A_{21} = 0$ which is equivalent to requiring that $A \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{n}}$ with $1 \not\succeq c$ and $1 \not\succeq 2$. Similarly, the requirement $AX_2 \subseteq X_2$ is equivalent to the requirement that $A \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{n}}$ with $2 \not\succeq c$ and $2 \not\succeq 1$. The same holds for the requirements on B and C.

We can build linear systems with a more complicated structure using coordinated linear systems as building blocks in the following ways

1. change the number of subsystems;
2. nest subsystems by using another coordinated linear system as a subsystem.

Such more intricate linear systems are sometimes called hierarchical systems. Some examples of such systems are given in the next example.

Example 5.3.1.
In the figure below, Σ_1 is the canonical example of a coordinated linear system. In Σ_2 the number of followers is reduced to one whereas in Σ_3 the number of followers is increased to five. System Σ_4 is obtained from Σ_2 by replacing the follower 1 by Σ_2. System Σ_5 is derived from Σ_1 by replacing the follower subsystem 1 by Σ_1 and replacing the follower subsystem 2 by Σ_2.



Remark 5.3.2. Recall that the *in-degree* of a vertex i in a digraph is the number of directed edges that terminate at vertex i. We note that in the above graphs of hierarchical systems, the maximum in-degree of all the vertices is one. Thus the Hasse diagrams of hierarchical systems correspond to out-trees (or out-forests more generally). As was mentioned in Subsection 2.1, the Hasse diagram of a poset $\mathcal{P} = (P, \succeq)$ is an out-forest if and only if the partial order $\succeq$ is in-ultra transitive (see Theorem 4.1 in [3]). This shows that hierarchical systems are poset-causal systems where the underlying poset is in-ultra transitive.

Remark 5.3.3. By Lemma 2.1.4, the dual system of a coordinated linear system is a poset-causal system with an underlying partial order that is out-ultra transitive, which correspond to Hasse diagrams that are in-tree forest. In particular, the dual of a coordinated linear system is not a coordinated linear system unless when the partial order happens to be a total order.

Chapter 6

Systems Theory of Structured Linear Systems

In this chapter, we consider systems theory for structured linear systems and particularly poset-causal systems. In Section 6.1, we consider whether we can obtain a structured realization of a linear system whose transfer function satisfies the structure requirements determined by a pre-order. Of course these results then also apply to transfer functions whose structure is determined by a poset (poset-causal systems). In Sections 6.2 and 6.3, we develop concepts of controllability and observability for poset-causal systems that respect the structure determined by the given poset. In Section 6.4, we show that, as in the classic case, there exist duality relationships between these concepts of controllability and observability. Lastly, in Section 6.5, we use subspaces defined in Sections 6.2 and 6.3 to derive a Kalman-type reduction of a realization of a poset-causal system which respects the structure. The results of Sections 6.2 to 6.5 were published in [48].

6.1 Structured Transfer Functions and Realizations

In this section we consider proper real rational transfer functions W such that W satisfies a prescribed block zero-pattern as determined by a pre-order. Our aim is to construct a realization (A, B, C, D) such that these system matrices have the same block zero-pattern determined by the pre-order.

Recall that if $\mathcal{T} = (P, T)$ is a pre-order on $P = \{1, \dots, p\}$ and $\underline{r}, \underline{m} \in \mathbb{Z}^p_+$ are given partitions, then the set $\mathcal{I}^{\underline{r}\times\underline{m}}_{\mathcal{T}}$ consists of $\underline{r} \times \underline{m}$ block matrices whose block zero-pattern is determined by the pre-order $\mathcal{T}$. Given a transfer function W and a pre-order $\mathcal{T}$ such that $W(\lambda) \in \mathcal{I}^{\underline{r}\times\underline{m}}_{\mathcal{T}}$ for each λ in its domain, we investigate whether we can find a realization (A, B, C, D) such that the systems matrices are in the block incidence spaces determined by $\mathcal{T}$, that is,

$$A \in \mathcal{I}^{\underline{n}\times\underline{n}}_{\mathcal{T}}, \quad B \in \mathcal{I}^{\underline{n}\times\underline{m}}_{\mathcal{T}}, \quad C \in \mathcal{I}^{\underline{r}\times\underline{n}}_{\mathcal{T}} \quad \text{and} \quad D \in \mathcal{I}^{\underline{r}\times\underline{m}}_{\mathcal{T}},$$

for some partition $\underline{n} \in \mathbb{Z}^p_+$.

The same problem is considered in the paper [24], for so-called systems over graphs. These systems, as is mentioned in Subsection 5.2.1, are essentially the same as poset-causal systems. In the paper [24], the triple (A, B, C) is also required to be stabilizable and detectable. However in the absence of stabilizability and detectability, the authors obtain a structured realization for a transfer function $W : D_W \to \mathcal{I}^{\underline{r}\times\underline{m}}_{\mathcal{P}}$, for some poset $\mathcal{P}$, some domain $D_W \subset \mathbb{C}$ and some partitions $\underline{r} = (r_1, \dots, r_p)$ and $\underline{m} = (m_1, \dots, m_p)$. This is done in the following manner:

Suppose W is a transfer function that satisfies the structured determined by a poset $\mathcal{P}$, that is, $W : D_W \to \mathcal{I}^{\underline{r}\times\underline{m}}_{\mathcal{P}}$. Partition W according to its block columns:

$$W(\lambda) = \begin{bmatrix} W(:,1)(\lambda) & \dots & W(:,p)(\lambda) \end{bmatrix}$$

for $\lambda \in D_W$. Let (A_i, B_i, C_i, D_i) be a minimal realization for the transfer function $W(:, i)$ for $i = 1, \dots, p$. Suppose $A_i \in \mathbb{R}^{n_i \times n_i}$ for $i = 1, \dots, p$ and define the partition $\underline{n} = (n_1, \dots, n_p)$. For these realizations, we have $C_i \in \mathbb{R}^{\underline{r} \times n_i}$, $B_i \in \mathbb{R}^{n_i \times m_i}$ and $D_i \in \mathbb{R}^{\underline{r} \times m_i}$ with $C_j(i, 1) = 0$ and $D_j(i, 1) = 0$ if $j \not\succeq i$, because $W(i, j) = 0$ if $j \not\succeq i$. We have

$$\left[\left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \quad \cdots \quad \left[\begin{array}{c|c} A_p & B_p \\ \hline C_p & D_p \end{array}\right]\right] = \left[\begin{array}{c|c} \begin{bmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_p \end{bmatrix} & \begin{bmatrix} B_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & B_p \end{bmatrix} \\ \hline \begin{bmatrix} C_1 & \dots & C_p \end{bmatrix} & \begin{bmatrix} D_1 & \dots & D_p \end{bmatrix} \end{array}\right] = \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B} \\ \hline \widetilde{C} & \widetilde{D} \end{array}\right].$$

Clearly we have $\widetilde{A} \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{n}}$ and $\widetilde{B} \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{m}}$, because any block diagonal matrix is in a block incidence space determined by $\mathcal{P}$. Lastly $\widetilde{C} \in \mathcal{I}_{\mathcal{P}}^{\underline{r} \times \underline{n}}$ and $\widetilde{D} \in \mathcal{I}_{\mathcal{P}}^{\underline{r} \times \underline{m}}$, because $C_j(i, 1) = \widetilde{C}(i, j) = 0$ and $D_j(i, 1) = \widetilde{D}(i, j) = 0$ if $j \not\succeq i$.

We now give an alternative method for determining structured realizations. We do this for zero-structures determined by a pre-order, not necessarily a poset. It is also possible to generalize the method in [24], illustrated above, to this more general case.

The next Lemma is the structured version of Proposition 3.1.1 and shows that a structured rational matrix function as determined by a pre-order can be written in terms of structured matrix polynomials as determined by the same pre-order. The proof follows the same idea as Theorem 2.3.3 in [13] for the classical case.

Lemma 6.1.1.
If $\mathcal{T}$ is a pre-order and $W : D_W \to \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$ is a proper rational matrix function, then there exists a matrix polynomial $L : \mathbb{C} \to \mathcal{I}_{\mathcal{T}}^{\underline{m} \times \underline{m}}$ of degree k, a matrix polynomial $H : \mathbb{C} \to \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$ of degree less than k and a matrix $W(\infty) \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$ such that

$$W(\lambda) = H(\lambda)L(\lambda)^{-1} + W(\infty), \quad \textit{for all } \lambda \in D_W \textit{ such that } L(\lambda) \neq 0.$$

Proof.
Suppose the domain of W is $D_W \subset \mathbb{C}$ and that $W(\lambda) \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$ for each $\lambda \in D_W$. Since W is proper, $W(\infty) = \lim_{|\lambda| \to \infty} W(\lambda)$ exists and $W(\infty) \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$. Now the matrix function defined by

$$V(\lambda) = W(\lambda) - W(\infty)$$

for $\lambda \in D_W$, is a strictly proper rational matrix function and $V(\lambda) \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$ for each $\lambda \in D_W$. Let $v_{ij}(\lambda)$ be the scalar entries of $V(\lambda)$. Then

$$v_{ij}(\lambda) = \frac{p_{ij}(\lambda)}{q_{ij}(\lambda)}$$

for polynomials $p_{ij} : \mathbb{C} \to \mathbb{R}$ and $q_{ij} : \mathbb{C} \to \mathbb{R}$, where we take the denominator polynomials to be monic. Since each v_{ij} is a strictly rational function, the degree of each polynomial p_{ij} is strictly less than the degree of the corresponding polynomial q_{ij}. Define the monic polynomial $r : \mathbb{C} \to \mathbb{R}$ to be the product of the monic polynomials q_{ij} and suppose r has degree k. Then $D_W = \{\lambda \in \mathbb{C} : r(\lambda) \neq 0\}$. Define $H : \mathbb{C} \to \mathbb{R}^{\underline{r} \times \underline{m}}$ by

$$H(\lambda) = r(\lambda)V(\lambda)$$

for $\lambda \in D_W$. Since $V(\lambda) \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$, we also have $H(\lambda) \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$ for each $\lambda \in D_W$. Furthermore $L(\lambda) := r(\lambda)I_m \in \mathcal{I}_{\mathcal{T}}^{\underline{m} \times \underline{m}}$ for each $\lambda \in \mathbb{C}$ and $L(\lambda) = 0$ if and only if $r(\lambda) = 0$. Thus $D_W = \{\lambda \in \mathbb{C} : L(\lambda) \neq 0\}$. It follows that

$$W(\lambda) = H(\lambda)L(\lambda)^{-1} + W(\infty)$$

for $L(\lambda) \neq 0$. □

The following theorem illustrates an alternative method for obtaining a structured realization using the companion matrix of L and the matrix function H in the previous lemma. These matrix functions are super-block matrices whose sub-blocks have the required structure. We then use the block canonical shuffle (see Definition 2.3.15) to transfer this structure to the block level of the super-block matrices.

Theorem 6.1.2.
If $\mathcal{T}$ is a pre-order and $\widehat{W} : D_W \subset \mathbb{C} \to \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$ is a proper rational transfer function of the causal LTI system Σ, then there exists a partition $\underline{n}$ such that Σ has a state space realization (A, B, C, D) with

$$A \in \mathcal{I}_{\mathcal{T}}^{\underline{n} \times \underline{n}}, \quad B \in \mathcal{I}_{\mathcal{T}}^{\underline{n} \times \underline{m}}, \quad C \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{n}} \quad \text{and} \quad D \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}.$$

Proof.
By Lemma 6.1.1 and its proof, we know that $W(\lambda) = H(\lambda)L(\lambda)^{-1} + W(\infty)$ for each λ for which $L(\lambda) \neq 0$, where for some k it holds that

$$H(\lambda) = \sum_{j=0}^{k-1} \lambda^j H_j \quad \text{and} \quad L(\lambda) = r(\lambda) I_m = \lambda^k I_m + \sum_{j=0}^{k-1} \lambda^j q_j I_m$$

with $H_j \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{m}}$ for each $j = 1, \ldots, p$ and some real monic polynomial $r(\lambda) = \lambda^k + q_{k-1}\lambda^{k-1} + \ldots + q_1\lambda + q_0$. Define $n = km$ and $\overline{\underline{n}} = (\underline{m}, \underline{m}, \ldots, \underline{m})$ and let $\Gamma_{\underline{n}}$ be the block canonical shuffle matrix associated with the sub-partition $\overline{\underline{n}}$. Define the matrices $\widetilde{A}, \widetilde{B}, \widetilde{C}$ and D by

$$\widetilde{A} = \begin{bmatrix} 0 & I_m & 0 & \ldots & 0 \\ 0 & 0 & I_m & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & I_m \\ -q_0 I_m & -q_1 I_m & -q_2 I_m & \ldots & -q_{k-1} I_m \end{bmatrix}, \quad \widetilde{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I_m \end{bmatrix} \tag{6.1}$$
$$\widetilde{C} = \begin{bmatrix} H_0 & H_1 & H_2 & \ldots & H_{k-1} \end{bmatrix} \quad \text{and} \quad D = W(\infty).$$

By Theorem 3.2.1, $\Sigma \sim (\widetilde{A}, \widetilde{B}, \widetilde{C}, D)$, but the matrices $\widetilde{A}, \widetilde{B}, \widetilde{C}$ do not necessarily have the prescribed zero structured determined by $\mathcal{T}$. However, the block entries of the matrices $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{C}$ do indeed have the prescribed structures. Hence, in order to impose the correct block zero structures on the system matrices, we employ the block canonical shuffle $\Gamma_{\overline{\underline{n}}}$ given in Definition 2.3.15. Define

$$A = \Gamma_{\overline{\underline{n}}} \widetilde{A} \Gamma_{\overline{\underline{n}}}^{\intercal}, \quad B = \Gamma_{\overline{\underline{n}}} \widetilde{B} \quad \text{and} \quad C = \widetilde{C} \Gamma_{\overline{\underline{n}}}^{\intercal}.$$

Then $A \in \mathcal{I}_{\mathcal{T}}^{\underline{n} \times \underline{n}}$, $B \in \mathcal{I}_{\mathcal{T}}^{\underline{n} \times \underline{m}}$, $C \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{n}}$ and $D \in \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{n}}$ by Corollary 2.3.19. Furthermore, we note that (A, B, C, D) and $(\widetilde{A}, \widetilde{B}, \widetilde{C}, D)$ are equivalent realizations:

$$CA^j B = (\widetilde{C} \Gamma_{\overline{\underline{n}}}^{\intercal})(\Gamma_{\overline{\underline{n}}} \widetilde{A} \Gamma_{\overline{\underline{n}}}^{\intercal})^j (\Gamma_{\overline{\underline{n}}} \widetilde{B}) = \widetilde{C} \widetilde{A}^j \widetilde{B}$$

for $j = 0, 1, \ldots$. Hence $W(\lambda) = C(\lambda I - A)^{-1} B + D$ and A, B, C and D have the block zero structure determined by the pre-order $\mathcal{T}$. $\square$

We now follow the proofs of Lemma 6.1.1 and Theorem 6.1.2 to extract the following procedure for obtaining a structured realization of a structured transfer function. We note that the realization that is obtained is highly non-minimal.

Procedure 6.1.3.
Given a pre-order $\mathcal{T}$ and a transfer function $\widehat{W} : D_W \mapsto \mathcal{I}_{\mathcal{T}}^{\underline{r} \times \underline{n}}$ of a causal LTI system Σ, we construct a structured realization of $\widehat{W}$ as follows:

1. *Compute $\widehat{W}(\infty) = \lim_{|\lambda| \to \infty} \widehat{W}(\lambda)$.*

2. Determine

$$\widehat{V}(\lambda) = \widehat{W}(\lambda) - \widehat{W}(\infty)$$

and the polynomials r_{ij} appearing in the rational matrix function

$$\widehat{V}(\lambda) = \left[\frac{p_{ij}(\lambda)}{r_{ij}(\lambda)}\right].$$

3. Determine the polynomial

$$r(\lambda) = \prod_{i,j} r_{ij}(\lambda) = \lambda^k + \sum_{j=0}^{k-1} q_j \lambda^j,$$

its degree k and the coefficients $q_0, \ldots, q_{k-1}$.

4. Set

$$H(\lambda) = r(\lambda)V(\lambda) = \sum_{j=0}^{k-1} \lambda^j H_j \quad \text{and} \quad L(\lambda) = r(\lambda)I_m = \lambda^k I_m + \sum_{j=0}^{k-1} \lambda^j q_j I_m.$$

and determine the matrices H_j. Then $\widehat{W}(\lambda) = H(\lambda)L(\lambda)^{-1} + \widehat{W}(\infty)$.

5. Define $n = km$ and $\underline{\overline{n}}$ as in Definition 2.3.15.

6. Define the matrices $\widetilde{A} \in \mathbb{R}^{\underline{\overline{n}} \times \underline{\overline{n}}}$, $\widetilde{B} \in \mathbb{R}^{\underline{\overline{n}} \times \underline{m}}$, $\widetilde{C} \in \mathbb{R}^{\underline{r} \times \underline{\overline{n}}}$ and $D \in \mathbb{R}^{\underline{r} \times \underline{m}}$ using equation (6.1):

$$\widetilde{A} = \begin{bmatrix} 0 & I_m & 0 & \ldots & 0 \\ 0 & 0 & I_m & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & I_m \\ -q_0 I_m & -q_1 I_m & -q_2 I_m & \ldots & -q_{k-1} I_m \end{bmatrix}, \qquad \widetilde{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I_m \end{bmatrix}$$

$$\widetilde{C} = \begin{bmatrix} H_0 & H_1 & H_2 & \ldots & H_{k-1} \end{bmatrix} \quad \text{and} \qquad D = W(\infty).$$

7. For the given p and k, construct the canonical shuffle γ as in Definition 2.3.13.

8. Construct the sub-partition $\underline{\widetilde{n}}$ and the block canonical shuffle $\Gamma_{\underline{\overline{n}}} \in \mathbb{R}^{\underline{\widetilde{n}} \times \underline{\overline{n}}}$ as in Definition 2.3.15.

9. Define $A = \Gamma_{\underline{\overline{n}}} \widetilde{A} \Gamma_{\underline{\overline{n}}}^{\intercal}$, $\quad B = \Gamma_{\underline{\overline{n}}} \widetilde{B}$ and $C = \widetilde{C} \Gamma_{\underline{\overline{n}}}^{\intercal}$.

Then

$$C(\lambda I_n - A)^{-1} B + D = \widehat{W}(\lambda)$$

and A, B, C and D have the sparsity structure determined by the pre-order $\mathcal{T}$.

The following example is considered in [24]. We apply the above procedure for this example.

Example 6.1.4.

Given partitions $\underline{r} = (1,1)$, $\underline{m} = (1,1)$ and the transfer function

$$\widehat{W}(\lambda) = \begin{bmatrix} \frac{1}{\lambda+1} & 0 \\ \frac{1}{\lambda-1} & \frac{1}{\lambda+1} \end{bmatrix} \in \mathcal{T}_{\mathcal{P}}^{\underline{r} \times \underline{m}}$$

of a leader- follower system with underlying poset $\mathcal{P} = (\{1,2\}, \succeq)$ with $1 \succeq 2$, we construct a structured realization as follows:

1. We compute $\widehat{W}(\infty)$:

$$\widehat{W}(\infty) = \lim_{|\lambda|\to\infty} \widehat{W}(\lambda) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

2. Determine $\widehat{V}(\lambda) = \widehat{W}(\lambda) - \widehat{W}(\infty)$ and the polynomials r_{ij}:

$$\widehat{V}(\lambda) = \begin{bmatrix} \frac{1}{\lambda+1} & 0 \\ \frac{1}{\lambda-1} & \frac{1}{\lambda+1} \end{bmatrix} = \begin{bmatrix} \frac{p_{11}(\lambda)}{r_{11}(\lambda)} & 0 \\ \frac{p_{21}(\lambda)}{r_{21}(\lambda)} & \frac{p_{22}(\lambda)}{r_{22}(\lambda)} \end{bmatrix}$$

3. Determine the polynomial r, its degree k and coefficients $q_0, \ldots, q_{k-1}$:

$$r(\lambda) = \prod_{i,j} q_{ij}(\lambda) = r_{11}(\lambda) r_{21}(\lambda) r_{22}(\lambda) = (\lambda+1)(\lambda-1)(\lambda+1) = \lambda^3 + \lambda^2 - \lambda - 1$$

   and r has degree $k = 3$. The coefficients in r are $q_0 = -1$, $q_1 = -1$ and $q_2 = 1$.

4. Set $H(\lambda) = r(\lambda)V(\lambda)$ and $L(\lambda) = r(\lambda)I_{\underline{m}}$ and determine the matrices H_j:

$$\begin{aligned} H(\lambda) &= \begin{bmatrix} (\lambda+1)(\lambda-1) & 0 \\ (\lambda+1)^2 & (\lambda+1)(\lambda-1) \end{bmatrix} \\ &= \lambda^2 \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix} = \lambda^2 H_2 + \lambda H_1 + H_0 \quad \text{and} \\ L(\lambda) &= \begin{bmatrix} \lambda^3 + \lambda^2 - \lambda - 1 & 0 \\ 0 & \lambda^3 + \lambda^2 - \lambda - 1 \end{bmatrix}. \end{aligned}$$

   Then $\widehat{W}(\lambda) = H(\lambda)L(\lambda)^{-1}$.

5. Define $n = km = 3 \cdot 2 = 6$ and $\overline{n} = ((1,1),(1,1),(1,1))$.

6. Define the matrices $\widetilde{A} \in \mathbb{R}^{\overline{n}\times\overline{n}}$, $\widetilde{B} \in \mathbb{R}^{\overline{n}\times\underline{m}}$, $\widetilde{C} \in \mathbb{R}^{\underline{r}\times\overline{n}}$ using equation (6.1):

$$\widetilde{A} = \begin{bmatrix} 0 & I_2 & 0 \\ 0 & 0 & I_2 \\ -q_0 I_2 & -q_1 I_2 & -q_2 I_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \end{bmatrix} \qquad \widetilde{B} = \begin{bmatrix} 0 \\ 0 \\ I_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\widetilde{C} = \begin{bmatrix} H_0 & H_1 & H_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 2 & 0 & 1 & 1 \end{bmatrix}$$

7. For $k = 3$ and $p = 2$, construct the canonical shuffle γ as in Definition 2.3.13:

i	s	$\ell_i^s = (i-1)p + s$	$\gamma(\ell_i^s) = q_s^i = (s-1)k + i$
1	1	1	1
1	2	2	4
2	1	3	2
2	2	4	5
3	1	5	3
3	2	6	6

8. Construct the sub-partition $\widetilde{\underline{n}}$ and the block canonical shuffle $\Gamma_{\overline{n}} \in \mathbb{R}^{\widetilde{n}\times\overline{n}}$ as in Definition 2.3.15:

$$\widetilde{\underline{n}} = ((1,1,1),(1,1,1))$$

and

$$\Gamma_{\overline{n}} = \begin{bmatrix} 1&0&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&0&1&0\\ 0&1&0&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&0&1 \end{bmatrix} = \left[\begin{array}{cc|cc|cc} 1&0&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&0&1&0\\ \hline 0&1&0&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&0&1 \end{array}\right] \in \mathbb{R}^{\widetilde{n}\times\overline{n}}.$$

9. Define $A = \Gamma_{\overline{n}}\widetilde{A}\Gamma_{\overline{n}}^{\intercal}$, $\quad B = \Gamma_{\overline{n}}\widetilde{B}$ and $C = \widetilde{C}\Gamma_{\overline{n}}^{\intercal}$:

$$A = \begin{bmatrix} 1&0&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&0&1&0\\ 0&1&0&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&0&1 \end{bmatrix} \begin{bmatrix} 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 1&0&1&0&-1&0\\ 0&1&0&1&0&-1 \end{bmatrix} \begin{bmatrix} 1&0&0&0&0&0\\ 0&0&0&1&0&0\\ 0&1&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&1&0&0&0\\ 0&0&0&0&0&1 \end{bmatrix}$$

$$= \left[\begin{array}{ccc|ccc} 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 1&1&-1&0&0&0\\ \hline 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 0&0&0&1&1&-1 \end{array}\right]$$

$$B = \begin{bmatrix} 1&0&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&0&1&0\\ 0&1&0&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&0&1 \end{bmatrix} \begin{bmatrix} 0&0\\ 0&0\\ 0&0\\ 0&0\\ 1&0\\ 0&1 \end{bmatrix} = \left[\begin{array}{c|c} 0&0\\ 0&0\\ 1&0\\ \hline 0&0\\ 0&0\\ 0&1 \end{array}\right]$$

$$C = \begin{bmatrix} -1&0&0&0&1&0\\ 1&-1&2&0&1&1 \end{bmatrix} \begin{bmatrix} 1&0&0&0&0&0\\ 0&0&0&1&0&0\\ 0&1&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&1&0&0&0\\ 0&0&0&0&0&1 \end{bmatrix} = \left[\begin{array}{ccc|ccc} -1&0&1&0&0&0\\ \hline 1&2&1&-1&0&1 \end{array}\right]$$

It can be seen that A, B and C have the sparsity structure

$$\begin{bmatrix} * & 0\\ * & * \end{bmatrix}$$

determined by the poset $\mathcal{P}$ and it can be confirmed that

$$C(\lambda I_6 - A)^{-1}B = \begin{bmatrix} \frac{1}{\lambda+1} & 0\\ \frac{1}{\lambda-1} & \frac{1}{\lambda+1} \end{bmatrix} = \widehat{W}(\lambda).$$

6.2 Downstream Reachable States and Upstream Controllability

In this section we investigate various notions of controllability for poset-causal systems that respect the partitioning of the state space and the associated block zero-pattern of the system matrices. Two of these concepts are generalizations of controllability notions that were defined in [23] for coordinated linear systems with one leader system and two follower subsystems ($\mathcal{P}_1$ in Example 2.1.5). In our approach to poset-causal systems, we do not identify leaders and followers, but rather make use of the concept of downstream reachable states.

6.2.1 Downstream Reachable States

Let $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ be a poset-causal system. For $i \in P$, the *i-downstream reachable set* $\mathcal{R}_i(A, B)$ consists of the states that are reachable in the i-th downstream system (5.8), i.e., vectors in the subspace $\mathcal{X}_{\downarrow i}$ that are reachable in the i-th downstream input-state system

$$\dot{x}^{\downarrow i}(t) = A(\downarrow i, \downarrow i)x^{\downarrow i}(t) + B(\downarrow i, i)u_i(t), \quad x^{\downarrow i}(0) = 0.$$

Thus, vectors ξ in $\mathcal{R}_i(A, B)$ are given by the integral formula

$$\xi = x^{\downarrow i}(0, u_i, t) = \int_0^t e^{A(\downarrow i, \downarrow i)(t-\tau)} B(\downarrow i, i)u_i(\tau)\ \mathrm{d}\tau.$$

Equivalently, the i-downstream reachable set is given by

$$\mathcal{R}_i(A, B) = \mathcal{R}(A(\downarrow i, \downarrow i), B(\downarrow i, i)) = \mathrm{Im}\ \mathcal{C}(A(\downarrow i, \downarrow i), B(\downarrow i, i)).$$

If $\xi \in \mathcal{R}_i(A, B)$, then we say that ξ is *i-downstream reachable*. We note that $\mathcal{R}_i(A, B)$ is the smallest $A(\downarrow i, \downarrow i)$-invariant subspace of $\mathcal{X}_{\downarrow i}$ that contains $\mathrm{Im} B(\downarrow i, i)$. In the sequel, when no confusion can arise we will omit A and B in the notation, and simply write $\mathcal{R}_i$ for $\mathcal{R}_i(A, B)$, and apply similar relaxations of the notation for derived subspaces defined below.

Lemma 6.2.1.
For a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$, we have

$$\mathcal{R}(A, B) = \sum_{i=1}^{p} I_{\underline{n}}(:, \downarrow i)\mathcal{R}_i.$$

Proof.
Fix some final time $t > 0$. Then $\xi \in \mathcal{R}(A, B)$ if and only if there exists some input $u = \bigoplus_{i=1}^{p} u_i$ such that $\xi = x(0, u, t)$. Now $x^{\downarrow i}(0, u_i, t) \in \mathcal{R}_i$ for all $i \in P$ and by (5.11), we have

$$\xi = x(0, u, t) = \sum_{i=1}^{p} I_{\underline{n}}(:, \downarrow i)x^{\downarrow i}(0, u_i, t).$$

Hence $\xi \in \mathcal{R}(A, B)$ if and only if $\xi \in \sum_{i=1}^{p} I_{\underline{n}}(:, \downarrow i)\mathcal{R}_i$. □

6.2.2 Upstream Controllability

For each $j \in P$ and $i \in \downarrow j$ we define the following subspaces of $\mathcal{X}_i$:

$$\overline{\mathcal{R}}_i^j = \overline{\mathcal{R}}_i^j(A, B) := \mathcal{X}_i \cap \mathcal{R}_j(A, B) \quad \text{and} \quad \widetilde{\mathcal{R}}_i^j = \widetilde{\mathcal{R}}_i^j(A, B) := P_{\mathcal{X}_i}\mathcal{R}_j(A, B).$$

Here $P_{\mathcal{X}_i}$ is the orthogonal projection onto $\mathcal{X}_i$. One can view $\overline{\mathcal{R}}_i^j$ as the set of local states $x_i \in \mathcal{X}_i$ that can be reached from a local input u_j in such a way that the other states downstream from j remain unaffected.

The subspace $\widetilde{\mathcal{R}}_i^j$, on the other hand, is the set of local states $x_i \in \mathcal{X}_i$ that can be reached from a local input u_j while the other states downstream from subsystem j may also be affected. From the definitions of the subspaces $\overline{\mathcal{R}}_i^j$ and $\widetilde{\mathcal{R}}_i^j$, we directly get the following inclusions:

$$\bigoplus_{i\in\downarrow j} \overline{\mathcal{R}}_i^j \subseteq \mathcal{R}_j \subseteq \bigoplus_{i\in\downarrow j} \widetilde{\mathcal{R}}_i^j. \tag{6.2}$$

Next we define subspaces, $\overline{\mathcal{R}}$, $\mathcal{R}^\circ$ and $\widetilde{\mathcal{R}}$, of the state space $\mathcal{X}$ which respect the structure imposed by the poset $\mathcal{P}$:

$$\overline{\mathcal{R}} := \bigoplus_{j\in P} \overline{\mathcal{R}}_j, \quad \text{where} \quad \overline{\mathcal{R}}_j := \sum_{i\in\uparrow j} \overline{\mathcal{R}}_j^i \quad \text{and} \quad \widetilde{\mathcal{R}} := \bigoplus_{j\in P} \widetilde{\mathcal{R}}_j, \quad \text{where} \quad \widetilde{\mathcal{R}}_j := \sum_{i\in\uparrow j} \widetilde{\mathcal{R}}_j^i. \tag{6.3}$$

Note that the sums in (6.3) are over upstream sets, while the direct sums in (6.2) were over downstream sets. An example illustrating these subspaces in given on page 109.

Definition 6.2.2.
We call a poset-causal system $\Sigma_{\mathcal{P}}$ *independently controllable* if $\overline{\mathcal{R}} = \mathcal{X}$, and *weakly upstream controllable* if $\widetilde{\mathcal{R}} = \mathcal{X}$.

In the context of coordinated linear systems, what we define as independent controllability above goes by the same name in Definition 3.16 in [23]. Weak upstream controllability does not appear to have been studied for coordinated linear systems yet, however, the subspaces $\widetilde{\mathcal{R}}_j$, play a role in Lemma 3.15 of [23].

The main reason for studying the spaces $\overline{\mathcal{R}}$ and $\widetilde{\mathcal{R}}$ instead of $\mathcal{R}$, is that they are structured as direct sums of subspaces of the local state spaces $\mathcal{X}_j$. Hence compressions, restrictions and projections of the system matrices to these subspaces exhibit the same poset-causal structure as the original system matrices. Such subspaces will be called *structured*.

Via the observation above, the subspace $\widetilde{\mathcal{R}}_j$ can be interpreted as the states in $\mathcal{X}_j$ that can be reached from inputs u_i in the subsystems that are upstream from the j-th subsystem while states in the other subsystems (that is, states x_i with $i \neq j$) are allowed to be affected. For $\overline{\mathcal{R}}_j$ only states in $\mathcal{X}_j$ are included in case they can be reached from an input u_i of an upstream subsystem (i.e., $i \in \uparrow j$) such that no states x_l in local subspaces other than $\mathcal{X}_j$ are effected.

For theoretical purposes we also introduce the structured subspace $\mathcal{R}^\circ$ of $\mathcal{X}$ defined by

$$\mathcal{R}^\circ := \bigoplus_{j\in P} \mathcal{R}_j^\circ, \quad \text{where} \quad \mathcal{R}_j^\circ := \mathcal{X}_j \cap \mathcal{R}. \tag{6.4}$$

There does not appear to be a clear interpretation of $\mathcal{R}^\circ$ in terms of the communication structure of the poset-causal system. Its relevance becomes clear from the following theorem, which is the main result of this section.

Theorem 6.2.3.
For a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$, we have

$$\overline{\mathcal{R}} \subseteq \mathcal{R}^\circ \subseteq \mathcal{R} \subseteq \widetilde{\mathcal{R}}. \tag{6.5}$$

In particular, if $\Sigma_{\mathcal{P}}$ is independently controllable, then $\Sigma_{\mathcal{P}}$ is controllable and if $\Sigma_{\mathcal{P}}$ is controllable, then $\Sigma_{\mathcal{P}}$ is weakly upstream controllable. Furthermore, if

$$\mathcal{Q} = \bigoplus_{j\in P} \mathcal{Q}_j \quad \textit{and} \quad \mathcal{S} = \bigoplus_{j\in P} \mathcal{S}_j \quad \textit{such that} \quad \mathcal{Q} \subseteq \mathcal{R} \subseteq \mathcal{S},$$

where $\mathcal{Q}_j \subseteq \mathcal{X}_j$ and $\mathcal{S}_j \subseteq \mathcal{X}_j$ for each $j \in P$, then $\mathcal{Q} \subseteq \mathcal{R}^\circ$ and $\widetilde{\mathcal{R}} \subseteq \mathcal{S}$.

The last claim of the above theorem can be interpreted as saying that among all structured subspaces of the state space $\mathcal{X}$, the subspace $\mathcal{R}^\circ$ is the largest included in the controllable subspace $\mathcal{R}$ and $\widetilde{\mathcal{R}}$ is the smallest structured subspace that includes $\mathcal{R}$. In the context of coordinated linear systems (with underlying structure $\mathcal{P}_1$ in Example 2.1.5), the inclusions $\overline{\mathcal{R}} \subseteq \mathcal{R} \subseteq \widetilde{\mathcal{R}}$ were obtained in the proof of Lemma 3.15 in [23]. That $\overline{\mathcal{R}}$ is not an maximal structured lower subspace of $\mathcal{R}$ results from the fact that for subspaces A, B and C the subspaces $(A \cap B) + (A \cap C)$ and $A \cap (B + C)$ need not coincide. The observation regarding $\mathcal{R}^\circ$ leads to the observation that $\overline{\mathcal{R}}$ is the maximal structured lower subspace of $\mathcal{R}$ if and only if $\overline{\mathcal{R}}_j = \mathcal{X}_j \cap \mathcal{R}$ for all $j \in P$. A geometric approach to LTI system was first considered in the book [52]. The proof of Theorem 6.2.3 is given later in this subsection, after we have proved the following two intermediate lemmas.

Lemma 6.2.4.
We have the following inclusions:

$$\overline{\mathcal{R}} \subseteq \mathcal{R}^\circ \subseteq \mathcal{R} \subseteq \widetilde{\mathcal{R}}. \tag{6.6}$$

Proof.
For the inclusion $\overline{\mathcal{R}} \subseteq \mathcal{R}^\circ$, note that for all $j \in P$ we have by Lemma 6.2.1 that

$$\begin{aligned}\overline{\mathcal{R}}_j = \sum_{i\in\uparrow j} \overline{\mathcal{R}}_j^i = \sum_{i\in\uparrow j} (\mathcal{X}_j \cap \mathcal{R}_i) \subseteq \mathcal{X}_j \cap \Big(\sum_{i\in\uparrow j} I_{\underline{n}}(:,\downarrow i)\mathcal{R}_i\Big)\\ \subseteq \mathcal{X}_j \cap \Big(\sum_{i\in P} I_{\underline{n}}(:,\downarrow i)\mathcal{R}_i\Big) = \mathcal{X}_j \cap \mathcal{R} = \mathcal{R}_j^\circ.\end{aligned}$$

To prove the second inclusion, $\mathcal{R}^\circ \subseteq \mathcal{R}$, we see that

$$\mathcal{R}^\circ = \bigoplus_{j\in P} \mathcal{R}_j^\circ = \bigoplus_{j\in P} (\mathcal{X}_j \cap \mathcal{R}) \subseteq \mathcal{R} \cap \bigoplus_{j\in P} \mathcal{X}_j = \mathcal{R} \cap \mathcal{X} = \mathcal{R}.$$

For the final inclusion $\mathcal{R} \subseteq \widetilde{\mathcal{R}}$, define $\Phi_1 = \{(i,j) : j \in P,\ i \in \uparrow j\}$ and $\Phi_2 = \{(i,j) : i \in P,\ j \in \downarrow i\}$. Then $\Phi_1 = \Phi_2$, because $i \in \uparrow j$ if and only if $j \in \downarrow i$. By Lemma 6.2.1 and the second inclusion in (6.2),

$$\begin{aligned}\mathcal{R} &= \sum_{i\in P} I_{\underline{n}}(:,\downarrow i)\mathcal{R}_i \subseteq \sum_{i\in P} I_{\underline{n}}(:,\downarrow i) \bigoplus_{j\in\downarrow i} \widetilde{\mathcal{R}}_j^i = \sum_{i\in P} I_{\underline{n}}(:,\downarrow i) \sum_{j\in\downarrow i} I_{\underline{n}}(\downarrow i, j)\widetilde{\mathcal{R}}_j^i\\ &= \sum_{(i,j)\in\Phi_2} I_{\underline{n}}(:,\downarrow i) I_{\underline{n}}(\downarrow i, j)\widetilde{\mathcal{R}}_j^i = \sum_{(i,j)\in\Phi_2} I_{\underline{n}}(:,j)\widetilde{\mathcal{R}}_j^i = \sum_{(i,j)\in\Phi_1} I_{\underline{n}}(:,j)\widetilde{\mathcal{R}}_j^i\\ &= \sum_{j\in P} I_{\underline{n}}(:,j) \sum_{i\in\uparrow j} \widetilde{\mathcal{R}}_j^i = \bigoplus_{j\in P} \widetilde{\mathcal{R}}_j = \widetilde{\mathcal{R}}.\end{aligned}$$

This completes the proof. □

Lemma 6.2.5.
For each $j \in P$, we have that

$$\mathcal{R}_j^\circ = \mathcal{X}_j \cap \mathcal{R} \qquad \textit{and} \qquad \widetilde{\mathcal{R}}_j = P_{\mathcal{X}_j}\mathcal{R}.$$

Proof.
There is nothing to prove for the first identity. For the identity $\widetilde{\mathcal{R}}_j = P_{\mathcal{X}_j}\mathcal{R}$ we have by definition that $\widetilde{\mathcal{R}}_j^i = P_{\mathcal{X}_j}\mathcal{R}_i$ and we have $\mathcal{X}_j \perp \mathcal{R}_i$ if $i \notin \uparrow j$. Thus, by the linearity of the projection $P_{\mathcal{X}_j}$, we get that

$$\begin{aligned}\widetilde{\mathcal{R}}_j &= \sum_{i\in\uparrow j} \widetilde{\mathcal{R}}_j^i = \sum_{i\in\uparrow j} P_{\mathcal{X}_j}\mathcal{R}_i = \sum_{i\in\uparrow j} P_{\mathcal{X}_j}\Big(I_{\underline{n}}(:,\downarrow i)\mathcal{R}_i\Big)\\ &= P_{\mathcal{X}_j} \sum_{i\in\uparrow j} I_{\underline{n}}(:,\downarrow i)\mathcal{R}_i + \{0\} \qquad = P_{\mathcal{X}_j} \sum_{i\in\uparrow j} I_{\underline{n}}(:,\downarrow i)\mathcal{R}_i + P_{\mathcal{X}_j} \sum_{i\notin\uparrow j} I_{\underline{n}}(:,\downarrow i)\mathcal{R}_i\\ &= P_{\mathcal{X}_j} \sum_{i\in P} I_{\underline{n}}(:,\downarrow i)\mathcal{R}_i = P_{\mathcal{X}_j}\mathcal{R},\end{aligned}$$

where we have applied Lemma 6.2.1 in the last step. □

Proof of Theorem 6.2.3.

The inclusion (6.5) was proved in Lemma 6.2.4, and the relations between controllability, independent controllability, upstream controllability and weak upstream controllability are a direct consequence.

To see that $\mathcal{R}^\circ$ and $\widetilde{\mathcal{R}}$ are optimal, assume that

$$\mathcal{Q} = \bigoplus_{j\in P} \mathcal{Q}_j \quad \text{and} \quad \mathcal{S} = \bigoplus_{j\in P} \mathcal{S}_j \quad \text{such that} \quad \mathcal{Q} \subseteq \mathcal{R} \subseteq \mathcal{S},$$

with $\mathcal{Q}_j \subseteq \mathcal{X}_j$ and $\mathcal{S}_j \subseteq \mathcal{X}_j$ for each $j \in P$. Then, by Lemma 6.2.5, we have

$$\mathcal{Q}_j = \mathcal{X}_j \cap \mathcal{Q} \subseteq \mathcal{X}_j \cap \mathcal{R} = \mathcal{R}_j^\circ \quad \text{and} \quad \widetilde{\mathcal{R}}_j = P_{\mathcal{X}_j}\mathcal{R} \subseteq P_{\mathcal{X}_j}\mathcal{S} = \mathcal{S}_j$$

for all $j \in P$. Therefore, by (6.3), we have $\mathcal{Q} \subseteq \mathcal{R}^\circ$ and $\widetilde{\mathcal{R}} \subseteq \mathcal{S}$. □

6.2.3 A Few Examples

We illustrate the above results with two examples. The first example shows in particular that all inclusions in (6.6) can be strict.

Example 6.2.6.

Let $\mathcal{P}_4$ be the poset given in Example 2.1.5. Consider the poset-causal system $\Sigma_{\mathcal{P}_4} \sim (A, B, 0, 0)$ with $A \in \mathcal{I}_{\mathcal{P}_4}^{\underline{n}\times\underline{n}}$ and $B \in \mathcal{I}_{\mathcal{P}_4}^{\underline{n}\times\underline{m}}$, with $\underline{n} = (2, 2, 3, 4)$ and $\underline{m} = (2, 1, 2, 2)$, given by

$$A = \left[\begin{array}{cc|cc|ccc|cccc}
1 & 0 & & & & & & & & & \\
0 & 0 & & & & & & & & & \\
\hline
1 & 0 & 1 & 0 & & & & & & & \\
0 & 0 & 0 & 0 & & & & & & & \\
\hline
 & & & & 0 & 0 & 1 & & & & \\
 & & & & 0 & 0 & 0 & & & & \\
 & & & & 0 & 0 & 0 & & & & \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1
\end{array}\right], \quad B = \left[\begin{array}{cc|c|cc|cc}
1 & 0 & & & & & \\
0 & 0 & & & & & \\
\hline
1 & 0 & 0 & & & & \\
0 & 1 & 1 & & & & \\
\hline
 & & & 1 & 0 & & \\
 & & & 0 & 1 & & \\
 & & & 0 & 0 & & \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1
\end{array}\right].$$

(The open white spaces in A and B represent appropriately sized zero matrices.) Then $\mathcal{X}_1 = \operatorname{span}\{e_1, e_2\}$, $\mathcal{X}_2 = \operatorname{span}\{e_3, e_4\}$, $\mathcal{X}_3 = \operatorname{span}\{e_5, e_6, e_7\}$ and $\mathcal{X}_4 = \operatorname{span}\{e_8, e_9, e_{10}, e_{11}\}$. So that $\mathcal{X} = \operatorname{span}\{e_1, \cdots, e_{11}\} = \mathbb{R}^{11}$. Here e_i is the i-th standard basis vector in $\mathbb{R}^{11}$. Firstly, we note that

$$\downarrow\!1 = \{1, 2, 4\}, \quad \downarrow\!2 = \{2, 4\}, \quad \downarrow\!3 = \{3, 4\}, \quad \downarrow\!4 = \{4\}.$$

Using this, the space of reachable states $\mathcal{R} = \operatorname{Im}\mathcal{C}(A, B)$ as well as the downstream reachable sets $\mathcal{R}_i = \operatorname{Im}\mathcal{C}(A(\downarrow\!i, \downarrow\!i), B(\downarrow\!i, i))$ for $i = 1, 2, 3, 4$ can be determined as:

$$\begin{aligned}
\mathcal{R} &= \operatorname{span}\{e_1, e_3, e_4, (e_5 + e_{10}), e_6, e_8, e_9, e_{11}\} \subsetneq \mathcal{X} \\
\mathcal{R}_1 &= \operatorname{span}\{e_1, e_3, (e_4 + e_8), e_9\} \subsetneq \mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \mathcal{X}_4 \\
\mathcal{R}_2 &= \operatorname{span}\{e_4, e_9\} \subsetneq \mathcal{X}_2 \oplus \mathcal{X}_4 \\
\mathcal{R}_3 &= \operatorname{span}\{(e_5 + e_{10}), (e_6 + e_{11})\} \subsetneq \mathcal{X}_3 \oplus \mathcal{X}_4 \\
\mathcal{R}_4 &= \operatorname{span}\{e_9, e_{11}\} \subsetneq \mathcal{X}_4.
\end{aligned}$$

Next we note that the spaces $\overline{\mathcal{R}}_j^i$ for $i \in P$ and $j \in \downarrow\!i$, are given by:

$$\begin{array}{lll}
\overline{\mathcal{R}}_1^1 = \operatorname{span}\{e_1\} \subsetneq \mathcal{X}_1, & \overline{\mathcal{R}}_2^1 = \operatorname{span}\{e_3\} \subsetneq \mathcal{X}_2, & \overline{\mathcal{R}}_4^1 = \operatorname{span}\{e_9\} \subsetneq \mathcal{X}_4, \\
\overline{\mathcal{R}}_2^2 = \operatorname{span}\{e_4\} \subsetneq \mathcal{X}_2, & \overline{\mathcal{R}}_4^2 = \operatorname{span}\{e_9\} \subsetneq \mathcal{X}_4, & \overline{\mathcal{R}}_3^3 = \{0\} \subsetneq \mathcal{X}_3, \\
\overline{\mathcal{R}}_4^3 = \{0\} \subsetneq \mathcal{X}_4, & \overline{\mathcal{R}}_4^4 = \operatorname{span}\{e_9, e_{11}\} \subsetneq \mathcal{X}_4. &
\end{array}$$

From these the spaces $\overline{\mathcal{R}}_j$ and $\mathcal{R}_j^\circ$ can be computed using (6.3) and $\widetilde{\mathcal{R}}_j$ can be computed using Lemma 6.2.5:

$$\begin{aligned}
&\overline{\mathcal{R}}_1 = \operatorname{span}\{e_1\}, && \mathcal{R}_1^\circ = \operatorname{span}\{e_1\}, && \widetilde{\mathcal{R}}_1 = \operatorname{span}\{e_1\},\\
&\overline{\mathcal{R}}_2 = \operatorname{span}\{e_3, e_4\}, && \mathcal{R}_2^\circ = \operatorname{span}\{e_3, e_4\}, && \widetilde{\mathcal{R}}_2 = \operatorname{span}\{e_3, e_4\},\\
&\overline{\mathcal{R}}_3 = \{0\}, && \mathcal{R}_3^\circ = \{e_6\} && \widetilde{\mathcal{R}}_3 = \operatorname{span}\{e_5, e_6\},\\
&\overline{\mathcal{R}}_4 = \operatorname{span}\{e_9, e_{11}\}, && \mathcal{R}_4^\circ = \operatorname{span}\{e_8, e_9, e_{11}\}, && \widetilde{\mathcal{R}}_4 = \operatorname{span}\{e_8, e_9, e_{10}, e_{11}\}.
\end{aligned}$$

Finally, we can calculate $\overline{\mathcal{R}}$, $\mathcal{R}^\circ$ and $\widetilde{\mathcal{R}}$ using (6.3)

$$\begin{aligned}
\overline{\mathcal{R}} &= \operatorname{span}\{e_1, e_3, e_4, e_9, e_{11}\},\\
\mathcal{R}^\circ &= \operatorname{span}\{e_1, e_3, e_4, e_6, e_8, e_9, e_{11}\},\\
\widetilde{\mathcal{R}} &= \operatorname{span}\{e_1, e_3, e_4, e_5, e_6, e_8, e_9, e_{10}, e_{11}\}.
\end{aligned}$$

This shows that the following inclusions are all strict:

$$\{0\} \subsetneq \overline{\mathcal{R}} \subsetneq \mathcal{R}^\circ \subsetneq \mathcal{R} \subsetneq \widetilde{\mathcal{R}} \subsetneq \mathcal{X}.$$

In particular, $\Sigma_{\mathcal{P}_4}$ is not controllable, neither is it independently or weakly upstream controllable. Note also that no structured subspaces of $\mathcal{X}$ can be strictly included in between $\mathcal{R}^\circ$ and $\mathcal{R}$ and in between $\mathcal{R}$ and $\widetilde{\mathcal{R}}$, confirming the optimality claim of Theorem 6.2.3 for this example.

Finally, note that for this example we have

$$\begin{aligned}
A\overline{\mathcal{R}} = A\mathcal{R}^\circ = A\mathcal{R} &= \operatorname{span}\{e_1, e_3, e_9, e_{11}\} \qquad \text{and}\\
A\widetilde{\mathcal{R}} &= \operatorname{span}\{e_1, e_3, e_9, e_{11}, e_5 + e_{10}\}.
\end{aligned}$$

Hence $\overline{\mathcal{R}}$, $\mathcal{R}^\circ$, $\mathcal{R}$ and $\widetilde{\mathcal{R}}$ are all invariant subspaces of A. For $\mathcal{R}$ this is true in general, but for the other three this need not always happen, as illustrated in the next example.

Example 6.2.7.
Now we consider an example where $\overline{\mathcal{R}}$, $\mathcal{R}^\circ$, and $\widetilde{\mathcal{R}}$ are not invariant under A. In the context of coordinated linear systems (with poset $\mathcal{P}_1$ in Example 2.1.5), for $\overline{\mathcal{R}}$ and $\widetilde{\mathcal{R}}$ this follows from the controllability decompositions in [23]. Consider a poset-causal system with poset $\mathcal{P}_6$ in Example 2.1.5, where $\underline{n} = (1, 1, 2)$, $\underline{m} = (1, 1, 1)$,

$$A = \left[\begin{array}{c|c|cc} 1 & 0 & 0 & 0\\ \hline 1 & 0 & 0 & 0\\ \hline 1 & 0 & 0 & 0\\ 0 & 1 & -1 & 0 \end{array}\right] \quad \text{and} \quad B = \left[\begin{array}{c|c|c} 1 & 0 & 0\\ \hline 0 & 0 & 0\\ \hline 0 & 0 & 0\\ 0 & 0 & 0 \end{array}\right].$$

In this case we have

$$\begin{aligned}
\overline{\mathcal{R}} = \mathcal{R}^\circ &= \operatorname{span}\{e_1\},\\
\mathcal{R} &= \operatorname{span}\{e_1, (e_2 + e_3)\} \quad \text{and}\\
\widetilde{\mathcal{R}} &= \operatorname{span}\{e_1, e_2, e_3\},
\end{aligned}$$

so that

$$A\overline{\mathcal{R}} = A\mathcal{R}^\circ = \operatorname{span}\{e_1 + e_2 + e_3\} \subsetneq \overline{\mathcal{R}} = \mathcal{R}^\circ \quad \text{and} \quad A\widetilde{\mathcal{R}} = \mathbb{R}^4 \subsetneq \widetilde{\mathcal{R}}.$$

6.2.4 Weak Local Controllability

We conclude this section with the study of a third controllability notion for poset-causal systems.

Definition 6.2.8.
We call a poset-causal system $\Sigma_{\mathcal{P}}$ *weakly locally controllable* if

$$\widetilde{\mathcal{R}}_i^i = P_{\mathcal{X}_i}\mathcal{R}_i = \mathcal{X}_i \quad \text{for each } i \in P.$$

Weak local controllability implies that each subsystem of $\Sigma_{\mathcal{P}}$, without external influences, seen as a system in its own right, is a controllable system. For coordinated linear systems it corresponds to Definition 3.10 [23].

Lemma 6.2.9.
A poset-causal system $\Sigma_{\mathcal{P}}$ is weak locally controllable if and only if each local pair (A_{ii}, B_{ii}) is controllable, that is, if and only if all local subsystems (5.7) are controllable.

Proof.
Using (5.14) and the fact that $I_{\underline{n}}(i, \downarrow i)B(\downarrow i, i) = B_{ii}$ it follows for all integers $k \geq 0$ that $I_{\underline{n}}(i, \downarrow i)A(\downarrow i, \downarrow i)^k B(\downarrow i, i) = A_{ii}^k B_{ii}$. Hence

$$\begin{aligned}\widetilde{\mathcal{R}}_i^i &= I_{\underline{n}}(i, \downarrow i)\mathcal{R}_i = \operatorname{Im} I_{\underline{n}}(i, \downarrow i)\, \mathcal{C}(A(\downarrow i, \downarrow i), B(\downarrow i, i)) \\ &= \operatorname{Im}\begin{bmatrix} B_{ii} & A_{ii}B_{ii} & \cdots & A_{ii}^{n-1}B_{ii}\end{bmatrix} \\ &= \operatorname{Im}\begin{bmatrix} B_{ii} & A_{ii}B_{ii} & \cdots & A_{ii}^{n_i-1}B_{ii}\end{bmatrix} \\ &= \operatorname{Im} \mathcal{C}(A_{ii}, B_{ii}) = \mathcal{R}(A_{ii}, B_{ii}).\end{aligned}$$

It follows that $\widetilde{\mathcal{R}}_i^i = \mathcal{X}_i$ if and only if (A_{ii}, B_{ii}) is a controllable pair. □

We next show that weak local controllability also implies controllability of $\Sigma_{\mathcal{P}}$.

Theorem 6.2.10.
If a poset-causal system $\Sigma_{\mathcal{P}}$ is weakly locally controllable, then it is controllable.

Proof.
Assume that $\Sigma_{\mathcal{P}}$ is weakly locally controllable. We show that $\mathcal{X} = \mathcal{R}$. Fix a $t > 0$. Let $\xi = \bigoplus_{j \in P} \xi_j \in \mathcal{X}$ with $\xi_j \in \mathcal{X}_j$. We seek an input $u = \bigoplus_{j \in P} u_j$ with u_j taking values in $\mathcal{U}_j$ so that $\xi = x(0, u, t)$. For $k = 1, 2, \ldots, p$, set

$$L_k := \{j \in P : |\uparrow j| \leq k\} \tag{6.7}$$

and note that $P = L_p = \bigcup_{k=1}^p L_k$ and $L_k \subseteq L_l$ if $k \leq \ell$. We prove by induction that for $k = 1, 2, \ldots, p$ there exist an input u so that $\xi_j = x_j(0, u, t)$ for all $j \in L_k$.

For $k = 1$, if $i \in L_1$, then $\bar{\uparrow} i = \emptyset$. Thus by (5.13), for any input $u = \bigoplus_{j \in P} u_j$ we have $x_i(0, u, t) = x^i(0, u_i, t)$ with x_i the state of the i-th subsystem (1.1) and x^i the state of the i-th local system (5.7). Hence x_i depends only on u_i. Since $\Sigma_{\mathcal{P}}$ is weakly locally controllable, for $i \in L_1$ there exist inputs u_i so that $x_i(0, u, t) = x^i(0, u_i, t) = \xi_i$. Set $u_j = 0$ for $j \notin L_1$. Then u is an input with the required property.

Now let $k \geq 1$ and assume we have an input $\widetilde{u} = \bigoplus_{j \in P} \widetilde{u}_j$ so that $x_j(0, \widetilde{u}, t) = \xi_j$ for all $j \in L_k$. If $k = p$ then we are done. Otherwise, set $u_j = \widetilde{u}_j$ for $j \notin R_k := \{j \in L_{k+1} : j \notin L_k\}$ and $R_p := \emptyset$. For $i \in L_k$ we have $\uparrow i \subseteq L_k$ so that $\xi_i = x_i(0, \widetilde{u}, t) = x_i(0, u, t)$, irrespectively of the choice of the inputs u_j for $j \in R_k$. It remains to select u_i for $i \in R_k$ so that also $\xi_i = x_i(0, u, t)$. Let $i \in R_k$. In that case $\bar{\uparrow} i \subseteq L_k$. Hence, for all $j \in \bar{\uparrow} i$, the input u_j is fixed. By (5.13) in Lemma 5.2.3, we have that for any input u_i

$$x_i(0, u, t) = x^i(0, u_i, t) + \sum_{j \in \bar{\uparrow} i} x_i^{\downarrow j}(0, u_j, t),$$

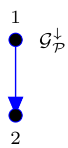
independent of the choice of the inputs u_j for $j \in R_k$, $j \neq i$. By assumption, the local system (5.7) is controllable. Hence there exists an input u_i so that

$$x^i(0, u_i, t) = \xi_i - \sum_{j \in \uparrow i} x_i^{\downarrow j}(0, u_j, t),$$

noting that the right hand side is fixed by our selection of inputs u_j for $j \in L_k$. As observed above, we can select u_i independently of the choice of the inputs u_j for $j \in L_k$ with $j \neq i$. This gives us a way to select the remaining inputs u_i for $i \in R_k$ so that $x_j(0, u, t) = \xi_j$ for all $j \in L_{k+1}$. By proceeding inductively we obtain an input u so that $x_j(0, u, t) = \xi_j$ for all $j \in L_p = P$, which proves our claim. $\square$

For weak local controllability, we only show that it implies controllability, but no inclusion of subspaces. Define $\widehat{\mathcal{R}} := \oplus \widetilde{\mathcal{R}}_i^i$. By Theorem 6.2.10, if $\widehat{\mathcal{R}} = \mathcal{X}$, then $\mathcal{R} = \mathcal{X}$. In view of Theorem 6.2.3, a natural question is whether $\widehat{\mathcal{R}} \subseteq \mathcal{R}$ holds also if $\mathcal{R} \neq \mathcal{X}$. This turns out not to be the case, as shown in the next example.

Example 6.2.11.
Let $\mathcal{P} = (P, \preceq)$ with $P = \{1, 2\}$ and $1 \preceq 2$. Take $\underline{n} = (1, 1)$ and $\underline{m} = (1, 1)$ and let $\Sigma_{\mathcal{P}} \sim (A, B, 0, 0)$ be the leader-follower system with



$$A = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & 0 \end{array}\right] \quad \text{and} \quad B = \left[\begin{array}{c|c} 1 & 0 \\ \hline 1 & 0 \end{array}\right].$$

Then $\mathcal{R} = \text{span}\{e_1 + e_2\}$, $\mathcal{R}_1 = \text{span}\{e_1 + e_2\}$ and $\mathcal{R}_2 = \{0\}$. Hence $\widetilde{\mathcal{R}}_1^1 = P_{\mathcal{X}_1} \mathcal{R}_1 = \text{span}\{e_1\}$ and $\widetilde{\mathcal{R}}_2^2 = \{0\}$ and so $\widehat{\mathcal{R}} = \text{span}\{e_1\}$. This shows that $\widehat{\mathcal{R}} \not\subseteq \mathcal{R}$.

It was pointed out in [23] that, for coordinated linear systems, weak local controllability is necessary and sufficient for pole placement. We now show this is also the case for poset-causal systems. We shall first prove the following lemma. Recall that p_X denotes the characteristic polynomial of a square matrix X. The following Lemma shows that the characteristic polynomial of a matrix in $\mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{n}}$ is the product of the characteristic polynomials of its main diagonal blocks.

Lemma 6.2.12.
If $A = [A_{ij}] \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{n}}$, then

$$p_A(\lambda) = \prod_{i \in \mathcal{P}} p_{A_{ii}}(\lambda), \quad \textit{so that} \quad \sigma(A) = \bigcup_{i \in \mathcal{P}} \sigma(A_{ii}).$$

Proof.
For $k = 1, \ldots, p$ define L_k as in (6.7) and set $R_k := \{j \in L_{k+1} \colon j \notin L_k\} = L_{k+1}/L_k$ and $R_p := \emptyset$ as in the proof of Theorem 6.2.10 and recall that for $i \in R_k$ we have $\uparrow i \subseteq L_k$. For $k = 1, 2, \ldots, p$, set

$$\widehat{A}_k = A(L_k, L_k) \quad \text{and} \quad \widetilde{A}_k = A(R_k, R_k).$$

Since $\uparrow i \subseteq L_k$ for all $i \in R_k$ and $R_k \cap L_k = \emptyset$, we have

$$\widehat{A}_{k+1} = \begin{bmatrix} \widehat{A}_k & 0 \\ * & \widetilde{A}_k \end{bmatrix} \quad \text{and} \quad \widetilde{A}_k = \bigoplus_{i \in R_k} A_{ii},$$

with $*$ indicating an unspecified matrix. It now follows recursively that

$$p_{\widehat{A}_k}(\lambda) = \prod_{i \in L_k} p_{A_{ii}}(\lambda), \quad \text{so that} \quad \sigma(\widehat{A}_k) = \bigcup_{i \in L_k} \sigma(A_{ii}), \quad k \in \mathcal{P}.$$

This proves out claim, since $L_p = \mathcal{P}$ and $A = \widehat{A}_p$. $\square$

Proposition 6.2.13.
A poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ is weakly locally controllable if and only if for any monic polynomial p of degree $n = n_1 + \cdots + n_p$ there exists a matrix $F \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$ such that $\det(\lambda I_{\underline{n}} - (A + BF)) = p(\lambda)$.

Proof.
Note that the observations about the structure of A with respect to the subspaces associated with L_k and R_k also apply to B and to any matrix $F \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$. As a consequence, it follows from Proposition 2.3.7 and Lemma 6.2.12 that

$$p_{(A+BF)}(\lambda) = \prod_{i\in\mathcal{P}} p_{(A_{ii}+B_{ii}F_{ii})}(\lambda). \tag{6.8}$$

In case $\Sigma_{\mathcal{P}}$ is weakly locally controllable, by the standard pole placement theorem (see Theorem 2.19 in [8]), for all monic polynomials p_i for $i \in \mathcal{P}$, with $\deg(p_i) = n_i$ we can find matrices F_{ii} so that $p_{(A_{ii}+B_{ii}F_{ii})}(\lambda) = p_i(\lambda)$. Now factor $p(\lambda) = \prod_{i\in\mathcal{P}} p_i(\lambda)$ with p_i monic and $\deg(p_i) = n_i$, and let F_{ii} be as above. Then the block diagonal matrix $F = \operatorname{diag}_{i\in\mathcal{P}}(F_{ii})$ is in $\mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$ and our claim follows by (6.8).

Conversely, assume $\Sigma_{\mathcal{P}}$ is not weakly locally controllable. Then by Lemma 6.2.9, there is a $i \in \mathcal{P}$ such that the pair (A_{ii}, B_{ii}) is not controllable. This means that A_{ii} has an uncontrollable eigenvalue, say λ_0. But then λ_0 is an eigenvalue of $A_{ii} + B_{ii}F_{ii}$ for all matrices $F_{ii} \in \mathbb{R}^{m_i\times n_i}$. Hence by (6.8), λ_0 is an eigenvalue of $A + BF$ for all matrices $F \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$. Thus, any monic polynomial p with degree n which does not have λ_0 as a root cannot appear as the characteristic polynomial of $A + BF$. □

Proposition 6.2.13 shows that weak local controllability corresponds to pole placement via a structured feedback matrix F. In case a poset-causal system is controllable but not weakly locally controllable, it follows that pole placement is still possible, but not always via a structured feedback matrix. We illustrate this in the following example, where we, in fact, show that state feedback stabilizability of the global system (in the classical sense) need not imply that state feedback stabilizability can be achieved by a structured feedback matrix.

Example 6.2.14.
Consider a poset-causal system $\Sigma_{\mathcal{P}_6} \sim (A, B, 0, 0)$ with $\mathcal{P}_6$ as in Example 2.1.5, $\underline{n} = (2, 2, 1)$ and $\underline{m} = (2, 1, 1)$ and with $A \in \mathcal{I}_{\mathcal{P}_6}^{\underline{n}\times\underline{n}}$ and $B \in \mathcal{I}_{\mathcal{P}_6}^{\underline{n}\times\underline{m}}$ given by:

$$A = \left[\begin{array}{cc|cc|c} 1 & 0 & & & \\ 0 & 0 & & & \\ \hline 1 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 0 & \\ \hline 1 & 0 & -1 & 0 & 1 \end{array}\right] \quad \text{and} \quad B = \left[\begin{array}{cc|c|c} 1 & 0 & & \\ 0 & 1 & & \\ \hline 1 & 0 & 1 & \\ 0 & 1 & 0 & \\ \hline 0 & 1 & 1 & 1 \end{array}\right].$$

We have $\mathcal{X}_1 = \operatorname{span}\{e_1, e_2\}$, $\mathcal{X}_2 = \operatorname{span}\{e_3, e_4\}$ and $\mathcal{X}_3 = \operatorname{span}\{e_5\}$. So that $\mathcal{X} = \operatorname{span}\{e_1, e_2, e_3, e_4, e_5\} = \mathbb{R}^5$. We note that $\downarrow 1 = \{1, 2, 3\}$, $\downarrow 2 = \{1, 2\}$ and $\downarrow 3 = \{3\}$. Using this, we determine the reachable set $\mathcal{R} = \operatorname{Im}\mathcal{C}(A, B)$ as well as the downstream reachable sets $\mathcal{R}_i = \mathcal{C}(A(\downarrow i, \downarrow i), B(\downarrow i, i))$ for $i = 1, 2, 3$:

$$\mathcal{R} = \operatorname{span}\{e_1, e_2, e_3, e_4, e_5\} = \mathcal{X},$$
$$\mathcal{R}_1 = \operatorname{span}\{(e_1 + e_3), e_2, (e_4 + e_5)\} \subsetneq \mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \mathcal{X}_3,$$
$$\mathcal{R}_2 = \operatorname{span}\{(e_3 + e_5)\} \subsetneq \mathcal{X}_2 \oplus \mathcal{X}_3, \quad \mathcal{R}_3 = \operatorname{span}\{e_5\} = \mathcal{X}_3.$$

Next we compute the spaces $\overline{\mathcal{R}}_i^i = X_i \cap \mathcal{R}_i$:

$$\widetilde{\mathcal{R}}_1^1 = \operatorname{span}\{e_1, e_2\} = \mathcal{X}_1, \quad \widetilde{\mathcal{R}}_2^2 = \operatorname{span}\{e_3\} \subsetneq \mathcal{X}_2, \quad \widetilde{\mathcal{R}}_3^3 = \operatorname{span}\{e_5\} = \mathcal{X}_3.$$

Since $\mathcal{R} = \mathcal{X}$, the system $\Sigma_{\mathcal{P}_6}$ is controllable, and hence A can be stabilized via state feedback: There exists a matrix $F \in \mathbb{R}^{4\times 5}$ so that $A+BF$ has eigenvalues only in the open left hand plane $\mathbb{C}^- := \{z \in \mathbb{C} \colon \operatorname{Re}(z) < 0\}$.

However, $\Sigma_{\mathcal{P}_6}$ is not weakly locally controllable, because $\widetilde{\mathcal{R}}_2^2 = \text{span}\{e_3\} \neq \mathcal{X}_2$. Hence there should not exist a matrix $F \in \mathcal{I}_{\mathcal{P}}^{\underline{m} \times \underline{n}}$ so that $A - BF$ has eigenvalues only in $\mathbb{C}^-$. Indeed, for $F = [f_{ij}] \in \mathcal{I}_{\mathcal{P}}^{\underline{m} \times \underline{n}}$ we have

$$A + BF = \left[\begin{array}{cc|cc|c} (1+f_{11}) & f_{12} & & & \\ f_{21} & f_{22} & & & \\ \hline * & * & (1+f_{33}) & f_{34} & \\ * & * & 0 & 0 & \\ \hline * & * & * & * & (1+f_{45}) \end{array}\right],$$

and it follows that 0 will necessarily be an eigenvalue of $A + BF$.

6.3 Upstream Indistinguishable States and Downstream Observability

In this section we define notions of distinguishability and observability for poset-causal systems that are dual to the notions of reachability and controllability considered in the previous section. We give the definitions and main results, but without proofs. The results follow directly from duality relations determined in the next section.

For a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ and a $i \in P$, in correspondence with (5.6), define

$$\mathcal{X}_{\uparrow i} := \bigoplus_{j \in \uparrow i} \mathcal{X}_j \qquad \text{and} \qquad \mathcal{X}_{P \setminus \uparrow i} := \bigoplus_{j \notin \uparrow i} \mathcal{X}_j.$$

The i-upstream indistinguishable set $\mathcal{N}_i(C, A)$ consists of the initial states $x_0^{\uparrow i} \in X_{\uparrow i}$ that cannot be distinguished from 0 using the output of subsystem i only. It follows that $\mathcal{N}_i(C, A)$ is contained in $\mathcal{X}_{\uparrow i}$ and consists of the states $\xi \in \mathcal{X}_{\uparrow i}$ that are indistinguishable from 0 in the system

$$\begin{aligned} \dot{x}^{\uparrow i}(t) &= A(\uparrow i, \uparrow i) x^{\uparrow i}(t), \qquad x^{\uparrow i}(0) = \xi \\ y^{\uparrow i}(t) &= C(i, \uparrow i) x^{\uparrow i}(t), \end{aligned}$$

that is, the i-th upstream system (5.10) with zero inputs. In this case we say that ξ is *i-upstream indistinguishable*. It follows that

$$\mathcal{N}_i(C, A) = \mathcal{N}(C(i, \uparrow i), A(\uparrow i, \uparrow i)) = \ker \mathcal{O}(C(i, \uparrow i), A(\uparrow i, \uparrow i)).$$

Also here we usually write $\mathcal{N}_i$ rather than $\mathcal{N}_i(C, A)$ if this does not cause confusion.

The following result is the analogue of lemma 6.2.1 for upstream indistinguishable sets. In the context of the coordinated linear systems this results corresponds to Lemma 4.2 in [23].

Lemma 6.3.1.
For a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ we have

$$\mathcal{N} = \bigcap_{i \in P} \left(\mathcal{N}_i \oplus \mathcal{X}_{P \setminus \uparrow i} \right).$$

Recall that $\mathcal{N}_i \subseteq \mathcal{X}_{\uparrow i}$ and that $\mathcal{X}_j \subseteq \mathcal{X}_{\uparrow i}$ if $j \in \uparrow i$. For each $j \in \uparrow i$, we define

$$\overline{\mathcal{N}}_i^j = \overline{\mathcal{N}}_i^j(C, A) := \mathcal{N}_i(C, A) \cap \mathcal{X}_j \quad \text{and} \quad \widetilde{\mathcal{N}}_i^j = \widetilde{\mathcal{N}}_i^j(C, A) := P_{\mathcal{X}_j} \mathcal{N}_i(C, A).$$

From these definitions, we immediately get the following inclusions:

$$\bigoplus_{j \in \uparrow i} \overline{\mathcal{N}}_i^j \subseteq \mathcal{N}_i \subseteq \bigoplus_{j \in \uparrow i} \widetilde{\mathcal{N}}_i^j,$$

In analogy with (6.3) and (6.4) we define the following structured subspaces of $\mathcal{X}$:

$$\begin{aligned} \overline{\mathcal{N}} &:= \bigoplus_{j\in P} \overline{\mathcal{N}}^j, & \mathcal{N}^\circ &:= \bigoplus_{j\in P} \mathcal{N}^{\circ j}, & \widetilde{\mathcal{N}} &:= \bigoplus_{j\in P} \widetilde{\mathcal{N}}^j, \quad \text{where} \\ \overline{\mathcal{N}}^j &:= \bigcap_{i\in\downarrow j} \overline{\mathcal{N}}^j_i, & \mathcal{N}^{\circ j} &:= P_{\mathcal{X}_j}\mathcal{N}, & \widetilde{\mathcal{N}}^j &:= \bigcap_{i\in\downarrow j} \widetilde{\mathcal{N}}^j_i. \end{aligned} \tag{6.9}$$

Definition 6.3.2.
We call a poset-causal system $\Sigma_\mathcal{P}$ *independently observable* if $\widetilde{\mathcal{N}} = \{0\}$, and *weakly downstream observable* if $\overline{\mathcal{N}} = \{0\}$.

In the context of coordinated linear systems, what we define as independent observability, goes by the same name in Definition 4.17 in [23]. Downstream observability and weak downstream observability does not appear to have been studied for coordinated linear systems yet, but the subspaces $\widetilde{\mathcal{N}}^j$ play an important role in Lemma 4.16 in [23].

The space $\overline{\mathcal{N}}^j$ may be interpreted as the states in $\mathcal{X}_j$ that are indistinguishable from each other when observing outputs that are downstream from subsystem j (that is, outputs y_i with $i \in \downarrow j$), while not being indistinguishable from states in other subsystems (that is x_i with $i \neq j$). The space $\widetilde{\mathcal{N}}^j$ consists of states in $\mathcal{X}_j$ that are indistinguishable from each other when observing outputs that are downstream from subsystem j (that is, outputs y_i with $i \in \downarrow j$), while in this case these states are also allowed to be indistinguishable from other states x_i with $i \neq j$. There does not seem to be a clear interpretation of the states in the space $\mathcal{N}^{\circ j}$ in terms of the communication structure of the poset-causal system. Its importance is due to the fact that it turns out to be the optimal structured subspace containing $\mathcal{N}$, as is shown in the following theorem - the main result of this section.

Theorem 6.3.3.
For a poset-causal system $\Sigma_\mathcal{P} \sim (A, B, C, D)$, we have

$$\overline{\mathcal{N}} \subseteq \mathcal{N} \subseteq \mathcal{N}^\circ \subseteq \widetilde{\mathcal{N}} \quad \textit{so that} \quad \widetilde{\mathcal{N}}^\perp \subseteq \mathcal{N}^{\circ\perp} \subseteq \mathcal{N}^\perp \subseteq \overline{\mathcal{N}}^\perp \tag{6.10}$$

and

$$\overline{\mathcal{N}}^j = \mathcal{X}_j \cap \mathcal{N} \quad \textit{so that} \quad P_{\mathcal{X}_j}\mathcal{N}^\perp = \mathcal{X}_j \ominus \overline{\mathcal{N}}^j.$$

In particular, if $\Sigma_\mathcal{P}$ is independently observable, then $\Sigma_\mathcal{P}$ is observable and if $\Sigma_\mathcal{P}$ is observable, then $\Sigma_\mathcal{P}$ is weakly downstream observable. Furthermore, if

$$\mathcal{Q} = \bigoplus_{j\in P} \mathcal{Q}_j, \quad \textit{and} \quad \mathcal{S} = \bigoplus_{j\in P} \mathcal{S}_j \quad \textit{such that} \quad \mathcal{Q} \subseteq \mathcal{N} \subseteq \mathcal{S},$$

where $\mathcal{Q}_j \subseteq \mathcal{X}_j$ and $\mathcal{S}_j \subseteq \mathcal{X}_j$ for each $j \in P$, then $\mathcal{Q} \subseteq \overline{\mathcal{N}}$ and $\mathcal{N}^\circ \subseteq \mathcal{S}$.

The above theorem shows that $\overline{\mathcal{N}}$ is the largest structured subspace of $\mathcal{X}$ that is contained in $\mathcal{N}$ and that $\mathcal{N}^\circ$ is the smallest structured subspace of $\mathcal{X}$ which contains $\mathcal{N}$. We conclude this section with the analogue of weak local controllability.

Definition 6.3.4.
The poset-causal system $\Sigma_\mathcal{P}$ is called *weakly locally observable* if

$$\overline{\mathcal{N}}^i_i = \{0\} \text{ for each } i \in P.$$

The analogues of Lemma 6.2.9 and Theorem 6.2.10 are collected in the following result.

Theorem 6.3.5.
The poset-causal system $\Sigma_\mathcal{P}$ is weakly locally observable if and only if each local pair (C_{ii}, A_{ii}) is observable, that is, if and only if all local systems (5.7) are observable. If $\Sigma_\mathcal{P}$ is weakly locally observable, then it is observable.

All inclusions in (6.10) can be strict and it need not be the case that $\bigoplus_{j\in P}\mathcal{N}_j^j$ contains $\mathcal{N}$. Examples that prove these claims can be obtained from the examples in the previous section and the duality relations explained in the next section. We present here an extension of Example 6.2.6 that will be useful in the sequel.

Example 6.3.6.
Consider the poset $\mathcal{P}_4$ given in Example 2.1.5 and the poset-causal system $\Sigma_{\mathcal{P}_4} \sim (A, 0, C, 0)$ with $A \in \mathcal{I}_{\mathcal{P}_4}^{\underline{n}\times\underline{n}}$ and $\underline{n}$ as in Example 6.2.6, $\underline{r} = (1, 1, 1, 1)$ and $C \in \mathcal{I}_{\mathcal{P}_4}^{\underline{r}\times\underline{n}}$, given by

$$C = \left[\begin{array}{cc|cc|ccc|cccc} 1 & 0 & & & & & & & & & \\ \hline 0 & 1 & 0 & 1 & & & & & & & \\ \hline & & & & 0 & 1 & 0 & & & & \\ \hline 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0. \end{array}\right].$$

In this case

$$\mathcal{N} = \operatorname{span}\{(-e_2 + e_4), (-e_5 + e_{10}), e_8, e_9, e_{11}\}$$

and the upstream indistinguishable sets are given by

$$\begin{aligned}
\mathcal{N}_1 &= \operatorname{span}\{e_2\} \subsetneq \mathcal{X}_1,\\
\mathcal{N}_2 &= \operatorname{span}\{e_1, (-e_2 + e_4), e_3\} \subsetneq \mathcal{X}_1 \oplus \mathcal{X}_2,\\
\mathcal{N}_3 &= \operatorname{span}\{e_5, e_7\} \subsetneq \mathcal{X}_3,\\
\mathcal{N}_4 &= \operatorname{span}\{e_2, e_4, (-e_5 + e_{10}), e_6, e_8, e_9, e_{11}\} \subsetneq \mathcal{X}.
\end{aligned}$$

One can further compute that

$$\begin{aligned}
\overline{\mathcal{N}}^1 &= \overline{\mathcal{N}}^2 = \overline{\mathcal{N}}^3 = \{0\},\\
\overline{\mathcal{N}}^4 &= \operatorname{span}\{e_8, e_9, e_{11}\},\\
\widetilde{\mathcal{N}}^1 &= \operatorname{span}\{e_2\} = \mathcal{N}^{\circ 1},\\
\widetilde{\mathcal{N}}^2 &= \operatorname{span}\{e_4\} = \mathcal{N}^{\circ 2},\\
\widetilde{\mathcal{N}}^3 &= \operatorname{span}\{e_5\} = \mathcal{N}^{\circ 3},\\
\widetilde{\mathcal{N}}^4 &= \operatorname{span}\{e_8, e_9, e_{10}, e_{11}\} = \mathcal{N}^{\circ 4},
\end{aligned}$$

from which it follows that

$$\begin{aligned}
\overline{\mathcal{N}} &= \operatorname{span}\{e_8, e_9, e_{11}\},\\
\widetilde{\mathcal{N}} &= \operatorname{span}\{e_2, e_4, e_5, e_8, e_9, e_{10}, e_{11}\} = \mathcal{N}^\circ.
\end{aligned}$$

This shows that

$$\{0\} \subsetneq \overline{\mathcal{N}} \subsetneq \mathcal{N} \subsetneq \mathcal{N}^\circ = \widetilde{\mathcal{N}} \subsetneq \mathcal{X}.$$

Hence the system is not observable, neither is it independently nor weakly upstream observable. Furthermore, no structured subspace can be strictly include between $\overline{\mathcal{N}}$ and $\mathcal{N}$ or between $\mathcal{N}$ and $\mathcal{N}^\circ = \widetilde{\mathcal{N}}$. In particular, unlike in Example 6.2.6, here the two subspaces $\overline{\mathcal{N}}$ and $\widetilde{\mathcal{N}}$ of $\mathcal{X}$ associated with the poset-causal system are the optimal structured subspaces that are included in $\mathcal{N}$ and include $\mathcal{N}$, respectively.

6.4 Duality

For classical centralized systems, controllability and observability are related through the duality identities

$$\mathcal{R}^d = \mathcal{N}^\perp \qquad \text{and} \qquad \mathcal{N}^d = \mathcal{R}^\perp.$$

Here $\mathcal{R}^d = \mathcal{R}(A^\intercal, C^\intercal)$ and $\mathcal{N}^d = \mathcal{N}(B^\intercal, A^\intercal)$ are the spaces of reachable and indistinguishable states, respectively, of the dual system. In this section we show that there are similar duality relations for the various notions of controllability and observability introduced in this paper. Such observations were not made in [23], since the subclass of poset-causal systems considered there is not closed under duality of the underlying posets.

The following theorem is the main result of this section.

Theorem 6.4.1.
Let $\Sigma_\mathcal{P} \sim (A, B, C, D)$ be a poset-causal system, with dual system $\Sigma_{\mathcal{P}_d} \sim (A_d, B_d, C_d, D_d)$. Define $\overline{\mathcal{R}}$, $\mathcal{R}^\circ$, $\widetilde{\mathcal{R}}$ as in (6.3) and (6.4) and $\overline{\mathcal{N}}$, $\mathcal{N}^\circ$, $\widetilde{\mathcal{N}}$ as in (6.9), and define $(\overline{\mathcal{R}})^d$, $(\mathcal{R}^\circ)^d$, $(\widetilde{\mathcal{R}})^d$, $(\overline{\mathcal{N}})^d$, $(\mathcal{N}^\circ)^d$, $(\widetilde{\mathcal{N}})^d$ analogously for $\Sigma_{\mathcal{P}_d}$. Then

$$(\overline{\mathcal{R}})^d = \widetilde{\mathcal{N}},^\perp \;\; (\mathcal{R}^\circ)^d = \mathcal{N},^{\circ\perp} \;\; (\widetilde{\mathcal{R}})^d = \overline{\mathcal{N}},^\perp \;\; (\overline{\mathcal{N}})^d = \widetilde{\mathcal{R}},^\perp \;\; (\mathcal{N}^\circ)^d = \mathcal{R},^{\circ\perp} \;\; (\widetilde{\mathcal{N}})^d = \overline{\mathcal{R}}.^\perp$$

In particular, the following equivalences hold:

(i) $\Sigma_\mathcal{P}$ is upstream controllable if and only if $\Sigma_{\mathcal{P}_d}$ is downstream observable.

(ii) $\Sigma_\mathcal{P}$ is weakly locally controllable if and only if $\Sigma_{\mathcal{P}_d}$ is weakly locally observable.

The identities in Theorem 6.4.1 will be proved via several intermediate steps.

An essential role in our definitions of controllability and observability is played by the downstream reachable and upstream unobservable sets $\mathcal{R}_i$ and $\mathcal{N}_i$ respectively. The next lemma explains the relation of the two sets under duality. Here, we denote the downstream reachable and upstream unobservable sets of the dual system $\Sigma_{\mathcal{P}_d}$ by $(\mathcal{R}_i)^d$ and $(\mathcal{N}_i)^d$ respectively.

Lemma 6.4.2.
Let $\Sigma_\mathcal{P} \sim (A, B, C, D)$ be a poset-causal system, with dual system $\Sigma_{\mathcal{P}_d} \sim (A_d, B_d, C_d, D_d)$. Then

$$\mathcal{X}_{\uparrow i} \ominus (\mathcal{R}_i)^d = \mathcal{N}_i \qquad \text{and} \qquad \mathcal{X}_{\downarrow i} \ominus (\mathcal{N}_i)^d = \mathcal{R}_i \quad \text{for each} \quad i \in P.$$

Proof.
Fix a $i \in P$. Note that $\mathcal{N}_i \subseteq \mathcal{X}_{\uparrow i}$ and $(\mathcal{R}_i)^d \subseteq \mathcal{X}_{\downarrow_d i} = \mathcal{X}_{\uparrow i}$. Now

$$\begin{aligned}(\mathcal{R}_i)^d &= \mathcal{R}(A_d(\downarrow_d i, \downarrow_d i), B_d(\downarrow_d i, i)) = \mathcal{R}(A^\intercal(\uparrow i, \uparrow i), C^\intercal(\uparrow i, i)) \\ &= \mathcal{R}(A(\uparrow i, \uparrow i)^\intercal, C(i, \uparrow i)^\intercal).\end{aligned}$$

By the standard duality identity, we have

$$\mathcal{X}_{\uparrow i} \ominus (\mathcal{R}_i)^d = \mathcal{X}_{\uparrow i} \ominus \mathcal{R}(A(\uparrow i, \uparrow i)^\intercal, C(i, \uparrow i)^\intercal) = \mathcal{N}(C(i, \uparrow i), A(\uparrow i, \uparrow i)) = \mathcal{N}_i.$$

The identity $\mathcal{X}_{\downarrow i} \ominus (\mathcal{N}_i)^d = \mathcal{R}_i$ follows similarly. □

The relations between the subspaces $\overline{\mathcal{R}}_i^j$, $\widetilde{\mathcal{R}}_i^j$, $\overline{\mathcal{N}}_j^i$, $\widetilde{\mathcal{N}}_j^i$ and the related subspaces for the dual system, denoted $(\overline{\mathcal{R}}_i^j)^d$, $(\widetilde{\mathcal{R}}_i^j)^d$, $(\overline{\mathcal{N}}_j^i)^d$, $(\widetilde{\mathcal{N}}_j^i)^d$, respectively, is less straightforward. They are listed in Lemma 6.4.3, the proof of which relies on some general identities in finite dimensional inner product spaces given in Lemma 3.4.9.

Lemma 6.4.3.
Let $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ be a poset-causal system, with dual system $\Sigma_{\mathcal{P}_d} \sim (A_d, B_d, C_d, D_d)$. Then for all $i, j \in P$ we have

$$\big(\widetilde{\mathcal{N}}_j^i\big)^d = \mathcal{X}_i \ominus \overline{\mathcal{R}}_i^j, \quad \big(\widetilde{\mathcal{R}}_i^j\big)^d = \mathcal{X}_i \ominus \overline{\mathcal{N}}_j^i, \quad \big(\overline{\mathcal{N}}_j^i\big)^d = \mathcal{X}_i \ominus \widetilde{\mathcal{R}}_i^j, \quad \big(\overline{\mathcal{R}}_i^j\big)^d = \mathcal{X}_i \ominus \widetilde{\mathcal{N}}_j^i.$$

Proof.
Using Lemma 6.4.2 along with (3.9) with $\mathcal{Y} = \mathcal{X}_{\downarrow j}$, $\mathcal{Y}_1 = \mathcal{R}_j$ and $\mathcal{Y}_2 = \mathcal{X}_i$ yields

$$\big(\widetilde{\mathcal{N}}_j^i\big)^d = P_{\mathcal{X}_i}(\mathcal{N}_j)^d = P_{\mathcal{X}_i}(\mathcal{X}_{\downarrow j} \ominus \mathcal{R}_j) = \mathcal{X}_i \ominus (\mathcal{X}_i \cap \mathcal{R}_j) = \mathcal{X}_i \ominus \overline{\mathcal{R}}_i^j.$$

A similar argument proves the identity $\big(\widetilde{\mathcal{R}}_i^j\big)^d = \mathcal{X}_i \ominus \overline{\mathcal{N}}_j^i$. Using the identity (3.8), with $n = 2$, $\mathcal{Y} = \mathcal{X}_{\downarrow j}$, $\mathcal{Y}_1 = (\mathcal{N}_j)^d$ and $\mathcal{Y}_2 = \mathcal{X}_i$, gives

$$\begin{aligned}\big(\overline{\mathcal{N}}_j^i\big)^d &= \mathcal{X}_i \cap (\mathcal{N}_j)^d = (\mathcal{X}_{\downarrow j} \ominus \mathcal{X}_{\downarrow j\setminus i}) \cap (\mathcal{X}_{\downarrow j} \ominus \mathcal{R}_j) = \mathcal{X}_{\downarrow j} \ominus (\mathcal{X}_{\downarrow j\setminus i} + \mathcal{R}_j)\\ &= \mathcal{X}_{\downarrow j} \ominus (P_{\mathcal{X}_i}\mathcal{R}_j \oplus \mathcal{X}_{\downarrow j\setminus i}) = \mathcal{X}_i \ominus (P_{\mathcal{X}_i}\mathcal{R}_j) = \mathcal{X}_i \ominus \widetilde{\mathcal{R}}_i^j.\end{aligned}$$

The identity $\big(\overline{\mathcal{R}}_i^j\big)^d = \mathcal{X}_i \ominus \widetilde{\mathcal{N}}_j^i$ can be derived analogously. □

Corollary 6.4.4.
The poset-causal system $\Sigma_{\mathcal{P}}$ is weakly locally controllable (weakly locally observable) if and only if the dual system $\Sigma_{\mathcal{P}_d}$ is weakly locally observable (weakly locally controllable).

Proof.
By Lemma 6.4.3 it follows that

$$\bigoplus_{i\in P} (\widetilde{\mathcal{R}}_i^i)^d = \Big(\bigoplus_{i\in P} \overline{\mathcal{N}}_i^i\Big)^\perp \quad \text{and} \quad \bigoplus_{i\in P} (\overline{\mathcal{N}}_i^i)^d = \Big(\bigoplus_{i\in P} \widetilde{\mathcal{R}}_i^i\Big)^\perp,$$

from which we immediately obtain the result. □

We now prove duality results for $(\overline{\mathcal{R}}_j)^d$, $(\widetilde{\mathcal{R}}_j)^d$, $(\mathcal{R}_j^\circ)^d$, $(\overline{\mathcal{N}}^j)^d$, $(\mathcal{N}^{j\circ})^d$ and $(\widetilde{\mathcal{N}}^j)^d$.

Lemma 6.4.5.
Let $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ be a poset-causal system, with dual system $\Sigma_{\mathcal{P}_d} \sim (A_d, B_d, C_d, D_d)$. Then for all $j \in P$ we have

$$\begin{aligned}&\big(\overline{\mathcal{R}}_j\big)^d = \mathcal{X}_j \ominus \widetilde{\mathcal{N}}^j, && \big(\mathcal{R}_j^\circ\big)^d = \mathcal{X}_j \ominus \mathcal{N}^{\circ j}, && \big(\widetilde{\mathcal{R}}_j\big)^d = \mathcal{X}_j \ominus \overline{\mathcal{N}}^j,\\ &\big(\overline{\mathcal{N}}^j\big)^d = \mathcal{X}_j \ominus \widetilde{\mathcal{R}}_j, && \big(\mathcal{N}^{\circ j}\big)^d = \mathcal{X}_j \ominus \mathcal{R}_j^\circ, && \big(\widetilde{\mathcal{N}}^j\big)^d = \mathcal{X}_j \ominus \overline{\mathcal{R}}_j.\end{aligned}$$

Proof.
By (6.3), Lemma 6.4.3 and (3.7), we have

$$\big(\widetilde{\mathcal{R}}_j\big)^d = \sum_{i\in\uparrow_d j} \big(\widetilde{\mathcal{R}}_j^i\big)^d = \sum_{i\in\downarrow j} \big(\mathcal{X}_j \ominus \overline{\mathcal{N}}_i^j\big) = \mathcal{X}_j \ominus \bigcap_{i\in\downarrow j} \overline{\mathcal{N}}_i^j = \mathcal{X}_j \ominus \overline{\mathcal{N}}^j.$$

The identities for $\big(\overline{\mathcal{R}}_j\big)^d$, $\big(\widetilde{\mathcal{R}}_j\big)^d$ and $\big(\widetilde{\mathcal{N}}^j\big)^d$ follow in a similar manner. The identity for $\big(\mathcal{R}_j^\circ\big)^d$ follows from Lemma 6.4.2 and the identity (3.8):

$$\begin{aligned}\big(\mathcal{R}_j^\circ\big)^d &= \mathcal{X}_j \cap \mathcal{R}^d = \mathcal{X}_{P\setminus j}^\perp \cap \mathcal{N}^\perp = (\mathcal{X}_{P\setminus j} + \mathcal{N})^\perp = (\mathcal{X}_{P\setminus j} \oplus P_{\mathcal{X}_j}\mathcal{N})^\perp\\ &= \mathcal{X}_j \ominus \mathcal{N}^{\circ j}.\end{aligned}$$

The identity for $\big(\mathcal{N}^{j\circ}\big)^d$ follows similarly. □

Theorem 6.4.1 now follows directly from the identities in Lemma 6.4.5. Note that Theorems 6.3.3 and 6.3.5 in Section 6.3 also follow directly from the result obtained in this section.

6.5 Minimality and Kalman Reduction for Poset-Causal Systems

The concepts and theory of minimality for poset-causal systems are problematic due to the additional structure in the form of the prescribed zero-block structure and the state space decomposition. The Kalman decomposition in the classical setting was summarized in Section 3.4. If the system Σ comes with the additional structure of a poset-causal system, i.e., $\Sigma = \Sigma_{\mathcal{P}}$ for some poset $\mathcal{P}$, then, in general, the poset structure is lost when $\Sigma_{\mathcal{P}}$ is compressed to the Kalman reduction $\Sigma_{\min}$, and one may have to compress to a larger structured subspace of the state space in order to preserve the poset structure.

Definition 6.5.1 (Poset-causal reduction)**.**
Consider a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ with state space $\mathcal{X} = \bigoplus_{j\in P} \mathcal{X}_j$ and a subspace $\widetilde{\mathcal{X}} = \bigoplus_{j\in P} \widetilde{\mathcal{X}}_j$ such that $\widetilde{\mathcal{X}}_j \subseteq \mathcal{X}_j$ for each j. If $\widetilde{A}, \widetilde{B}, \widetilde{C}$ are the compressions of A, B, C to $\widetilde{\mathcal{X}}$, respectively, and

$$CA^kB = \widetilde{C}\widetilde{A}^k\widetilde{B} \quad \text{for } k = 0, 1, \ldots,$$

then the realization $\widetilde{\Sigma}_{\mathcal{P}} \sim (\widetilde{A}, \widetilde{B}, \widetilde{C}, D)$ is called a *poset-causal reduction* of $\Sigma_{\mathcal{P}}$ to the subspace $\widetilde{\mathcal{X}}$.

Using Lemma 3.4.10 and the state space sub-spaces that underlie our notions of controllability and observability defined in Sections 6.2 and 6.3, we obtain the following candidate for a poset-causal reduction.

Proposition 6.5.2.
Consider a poset-causal system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$ with associated subspaces $\widetilde{\mathcal{R}}$, $\overline{\mathcal{R}}$ and $\overline{\mathcal{N}}$ as defined in (6.3) and (6.9). Define the subspace $\widetilde{\mathcal{X}} = \widetilde{\mathcal{R}} \ominus (\overline{\mathcal{R}} \cap \overline{\mathcal{N}})$ and let $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{C}$ be the compressions of A, B and C to $\widetilde{\mathcal{X}}$. Then $\widetilde{\Sigma}_{\mathcal{P}} \sim (\widetilde{A}, \widetilde{B}, \widetilde{C}, D)$ is a poset-causal reduction of $\Sigma_{\mathcal{P}}$. Furthermore, if $P_{\mathcal{X}_j}\widetilde{\mathcal{X}} = P_{\mathcal{X}_j}\mathcal{X}_{co}$ for all $j \in P$, and $\widehat{\mathcal{X}} = \bigoplus_{j\in P} \widehat{\mathcal{X}}_j$ is a subspace of $\mathcal{X}$ with $\widehat{\mathcal{X}}_j \subseteq \mathcal{X}_j$ for all $j \in P$ so that $\mathcal{X}_{co} \subseteq \widehat{\mathcal{X}}$, then $\widetilde{\mathcal{X}} \subseteq \widehat{\mathcal{X}}$.

Proof.
Using definitions (6.3) and (6.9), it follows that

$$\widetilde{\mathcal{X}} = \bigoplus_{j=1}^{p} \widetilde{\mathcal{X}}_j, \quad \text{with } \widetilde{\mathcal{X}}_j = \widetilde{\mathcal{R}}_j \ominus (\overline{\mathcal{R}}_j \cap \overline{\mathcal{N}}^j) \subset \mathcal{X}_j,$$

where $\widetilde{\mathcal{X}}_j$, $\overline{\mathcal{R}}_j$ and $\overline{\mathcal{N}}^j$ are also as defined in (6.3) and (6.9). Hence $\widetilde{\mathcal{X}}$ is a structured subspace of the state space. Since $\overline{\mathcal{R}} \subseteq \mathcal{R} \subseteq \widetilde{\mathcal{R}}$ and $\overline{\mathcal{N}} \subseteq \mathcal{N}$ by (6.5) and (6.10), respectively, it follows from Lemma 3.4.10 that $CA^kB = \widetilde{C}\widetilde{A}^k\widetilde{B}$ for $k = 0, 1, \ldots$ and hence $\widetilde{\Sigma}_{\mathcal{P}}$ is a poset-causal reduction of $\Sigma_{\mathcal{P}}$.

For the final claim, assume $\widetilde{\mathcal{X}}_j = P_{\mathcal{X}_j}\widetilde{\mathcal{X}} = P_{\mathcal{X}_j}\mathcal{X}_{co}$ for all $j \in P$ and let $\widehat{\mathcal{X}}$ be as in the proposition, then

$$\widetilde{\mathcal{X}}_j = P_{\mathcal{X}_j}\mathcal{X}_{co} \subseteq P_{\mathcal{X}_j}\widehat{\mathcal{X}} = \widehat{\mathcal{X}}_j, \quad j \in P. \qquad \square$$

In the above proposition we worked with the subspaces $\widetilde{\mathcal{R}}$, $\overline{\mathcal{R}}$ and $\overline{\mathcal{N}}$ since they have a natural interpretation in the context of the poset-causal system $\Sigma_{\mathcal{P}}$ and satisfy the inclusion conditions of Lemma 3.4.10. Furthermore, $\widetilde{\mathcal{R}}$ is the smallest structured subspace of $\mathcal{X}$ that contains $\mathcal{R}$ and $\overline{\mathcal{N}}$ is the largest structured subspace of $\mathcal{X}$ contained in $\mathcal{N}$, but $\overline{\mathcal{R}}$ need not be the largest structured subspace of $\mathcal{X}$ contained in $\mathcal{R}$, unless $\overline{\mathcal{R}} = \mathcal{R}^\circ$. A potentially smaller structured subspace that contains $\mathcal{X}_{co}$ is thus given by $\widetilde{\mathcal{X}}' := \widetilde{\mathcal{R}} \ominus (\mathcal{R}^\circ \cap \overline{\mathcal{N}})$. However, despite the fact that $\mathcal{R}^\circ$ is the largest structured subspace contained in $\mathcal{R}$, it need not be the case that $P_{\mathcal{X}_j}\widetilde{\mathcal{X}}' = P_{\mathcal{X}_j}\mathcal{X}_{co}$ for all $j \in P$, as illustrated in the next example.

Example 6.5.3.
Consider a leader-follower system $\Sigma_{\mathcal{P}} \sim (A, B, C, D)$, where $\mathcal{P} = (P, \preceq)$ is the poset with $P = \{1, 2\}$ and $1 \preceq 2$. Suppose $A \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{n}}$, $B \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}$ and $A \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{n}}$ with $\underline{n} = (2, 2)$, $\underline{m} = (2, 1)$ and $\underline{r} = (1, 1)$, are given by

$$A = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right], \quad B = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right] \quad \text{and} \quad C = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 \end{array}\right].$$

Then $\mathcal{X}_1 = \mathrm{span}\{e_1, e_2\}$ and $\mathcal{X}_2 = \mathrm{span}\{e_3, e_4\}$. Calculating the reachable space $\mathcal{R}$ and the unobservable space $\mathcal{N}$ gives

$$\mathcal{R} = \mathrm{span}\{e_1, e_2 + e_4\} \quad \text{and} \quad \mathcal{N} = \mathrm{span}\{e_2, e_4\}.$$

In this case, we have $\mathcal{X}_{co} = \mathcal{R} \ominus (\mathcal{R} \cap \mathcal{N}) = \mathrm{span}\{e_1\}$ and hence

$$P_{\mathcal{X}_1}\mathcal{X}_{co} = \mathrm{span}\{e_1\} \quad \text{and} \quad P_{\mathcal{X}_2}\mathcal{X}_{co} = \{0\}.$$

Now we consider the structured space $\widetilde{\mathcal{X}}' = \widetilde{\mathcal{R}} \ominus (\mathcal{R}^\circ \cap \overline{\mathcal{N}}) = \bigoplus_{j=1,2} \widetilde{\mathcal{X}}'_j$. Then $\mathcal{X}'_j = \widetilde{\mathcal{X}}_j \ominus (\mathcal{R}^\circ_j \cap \overline{\mathcal{N}}^j)$ and we can compute $\widehat{R}_j$, $\mathcal{R}^\circ_j$ and $\overline{\mathcal{N}}^j$ for $j = 1, 2$ using Lemma 6.2.5 and Theorem 6.3.3:

$$\begin{aligned} &\widetilde{\mathcal{R}}_1 = P_{\mathcal{X}_1}\mathcal{R} = \mathrm{span}\{e_1, e_2\}, && \mathcal{R}^\circ_1 = \mathcal{X}_1 \cap \mathcal{R} = \{e_1\}, && \overline{\mathcal{N}}^1 = \mathcal{X}_1 \cap \mathcal{N} = \mathrm{span}\{e_2\},\\ &\widetilde{\mathcal{R}}_2 = P_{\mathcal{X}_2}\mathcal{R} = \mathrm{span}\{e_4\}, && \mathcal{R}^\circ_2 = \mathcal{X}_2 \cap \mathcal{R} = \{0\}, && \overline{\mathcal{N}}^2 = \mathcal{X}_2 \cap \mathcal{N} = \mathrm{span}\{e_4\}, \end{aligned}$$

which gives

$$\mathcal{X}'_1 = \widetilde{\mathcal{R}}_1 \ominus (\mathcal{R}^\circ_1 \cap \overline{\mathcal{N}}^1) = \mathcal{X}_1 \quad \text{and} \quad \mathcal{X}'_2 = \widetilde{\mathcal{R}}_2 \ominus (\mathcal{R}^\circ_2 \cap \overline{\mathcal{N}}^2) = \mathrm{span}\{e_4\}.$$

For both $j = 1, 2$, we see that $P_{\mathcal{X}_j}\mathcal{X}_{co} \neq \mathcal{X}'_j$. Hence, despite $\mathcal{X}_{co}$ being a structured subspace of $\mathcal{X}$, of dimension 1, our approximation obtained from Proposition 6.5.2 is a structured subspace of dimension 3. One can further check that in this case $\overline{\mathcal{R}} = \mathcal{R}^\circ$.

There are many different choices of state space subspaces to compress the matrices A, B and C to a minimal realization, which may or may not be structured, and when it is not structured, there may or may not be a natural way to embed this subspace in a structured subspace of $\mathcal{X}$ for which compressed matrices preserve the moments. We have chosen to work with the space $\mathcal{X}_{co} = \mathcal{R} \ominus (\mathcal{R} \cap \mathcal{N})$, since it appears naturally in the Kalman decomposition of the system and there are natural structured analogues of the observability and controllability spaces that meet the requirements. Alternatively, using a duality argument, one can also work with $\mathcal{N}^\perp$ and $\mathcal{R}^\perp$ instead of $\mathcal{R}$ and $\mathcal{N}$, respectively. In this case, the (possibly) non-structured subspace becomes $\mathcal{N}^\perp \ominus (\mathcal{N}^\perp \cap \mathcal{R}^\perp)$, which can be embedded in the structured subspace $\widetilde{\mathcal{N}}^\perp \ominus (\overline{\mathcal{N}}^\perp \cap \widetilde{\mathcal{R}}^\perp)$, or in the (possibly) smaller structured subspace $\mathcal{N}^{\circ\perp} \ominus (\overline{\mathcal{N}}^\perp \cap \widetilde{\mathcal{R}}^\perp)$. In the above example, in fact, it turns out that all three subspaces of $\mathcal{X}$ are the same, so that in this case it is better to work with $\mathcal{N}^\perp$ and $\mathcal{R}^\perp$ instead of $\mathcal{R}$ and $\mathcal{N}$.

Example 6.5.4.
Let A, B and C as well as $\mathcal{P}$ be as in Example 6.5.3. In this case one can compute that $\widetilde{\mathcal{N}} = \mathcal{N}^\circ = \mathcal{N} = \mathrm{span}\{e_2, e_4\}$, while it was already observed that $\overline{\mathcal{N}} = \mathcal{N}$, $\mathcal{R} = \mathrm{span}\{e_1, e_2 + e_4\}$ and $\widetilde{\mathcal{R}} = \mathrm{span}\{e_1, e_2, e_4\}$. Hence

$$\overline{\mathcal{N}}^\perp = \widetilde{\mathcal{N}}^\perp = \mathcal{N}^{\circ\perp} = \mathcal{N}^\perp = \mathrm{span}\{e_1, e_3\}, \quad \mathcal{R}^\perp = \mathrm{span}\{e_2 - e_4, e_3\}, \quad \widetilde{\mathcal{R}}^\perp = \mathrm{span}\{e_3\}.$$

From this we obtain that

$$\mathcal{N}^\perp \ominus (\mathcal{N}^\perp \cap \mathcal{R}^\perp) = \mathrm{span}\{e_1\} = \widetilde{\mathcal{N}}^\perp \ominus (\overline{\mathcal{N}}^\perp \cap \widetilde{\mathcal{R}}^\perp).$$

Hence, in this case, when compressing to $\widetilde{\mathcal{N}}^\perp \ominus (\overline{\mathcal{N}}^\perp \cap \widetilde{\mathcal{R}}^\perp)$, a minimal realization is obtained.

Chapter 7

Control of Structured Linear Systems

We now want to investigate the optimal control of structured linear systems. In particular, we will look at the $\mathcal{H}_2$-optimal control problem for poset-causal systems. The solution to the classical (unstructured) $\mathcal{H}_2$-control problem was summarized in Section 4.3. The state feedback case was considered in Section 4.4 and the spectral factorization approach to the state feedback case was explained in Section 4.5. Poset-causal systems were discussed in Section 5.2.

In the first section of this chapter, we consider the state feedback case of $\mathcal{H}_2$-optimal control problem for poset-causal systems using a spectral factorization approach as used by Shah in [38] and Swigart in [45]. We will loosen the state feedback assumption in Section 7.2 where we consider the output feedback case. The output feedback $\mathcal{H}_2$-control problem for poset-causal systems is significantly more difficult than the state feedback case. In the state feedback case, the problem may readily be reduced to local unstructured problems. Moreover a solution may be recovered from these local solutions. This depends essentially on the form of the plant realization under the state feedback restrictions. These conditions do not hold in the output feedback case.

7.1 State Feedback Control

In this section, a summary is given of the poset-causal $\mathcal{H}_2$-control problem and solution strategy considered by Shah in [38]. Similar results were also obtained in [45] by Swigart and Lall for systems defined over graphs. Although the results are not new, we will consider continuous time systems and not discrete time systems as in [38] and [45]. Furthermore, we will rewrite both the problem statement and solution strategy in terms of the notation we have developed in Section 2. We hope that this will make the exposition of the results more precise. Shah and Swigart assumes that full state feedback is available and this plays a critical role in the solution strategy of the current section. In the next section, we will consider the more general problem of output feedback control. We will only provide discussions and sketches of proof of some of the results in this section as similar results will be proved in the more general context of output feedback control in Section 7.2.

7.1.1 Definition of Poset-Causal $\mathcal{H}_2$-Control Problem (State Feedback Case)

Let $\mathcal{P} = (P, \succeq)$, with $P = \{1, \dots, p\}$, be a poset and let $\underline{n}, \underline{m}, \underline{r} \in \mathbb{Z}_+^p$ be partitions. Consider a $\mathcal{P}$-poset-causal plant with state space realization

$$\begin{aligned} \dot{x}(t) &= Ax(t) + \omega(t) + Bu(t), & x(0) &= x_0 \\ z(t) &= Cx(t) + Du(t), & t &\geq 0 \\ y(t) &= x(t), \end{aligned} \tag{7.1}$$

where $A \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{n}}$, $B \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}$, $C \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{n}}$, $D \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{m}}$. The plant in (7.1) is a special case of the general plant in (4.3) with $B_1 = I_{\underline{n}}$, $B_2 = B$, $C_1 = C$, $C_2 = I_{\underline{n}}$, $D_{11} = 0$, $D_{12} = D$, $D_{21} = 0$ and $D_{22} = 0$. Hence, it is given by

$$\widehat{G} = \left[\begin{array}{c|cc} A & I_{\underline{n}} & B \\ \hline C & 0 & D \\ I_{\underline{n}} & 0 & 0 \end{array}\right] = \begin{bmatrix} C\Phi_A & C\Phi_A B + D \\ \Phi_A & \Phi_A B \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix}. \tag{7.2}$$

At the subsystem level, the local subsystems have the following state space equations:

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j\in\uparrow i} A_{ij}x_j(t) + \sum_{j\in\uparrow i} B_{ij}u_j(t), \quad && x_i(0) = x_{i,0}, \\ z_i(t) &= \sum_{j\in\uparrow i} C_{ij}x_j(t) + \sum_{j\in\uparrow i} D_{ij}u_j(t), && t \geq 0, \\ y_i(t) &= x_i(t), && i = 1,\ldots,p. \end{aligned}$$

Communication restrictions are imposed by the underlying poset $\mathcal{P}$, this can be seen in that the sums are taken over upstream sets $\uparrow i$. The following assumptions are made about the realization in (7.2):

1. A is stable. Hence $\widehat{G}_{11}, \widehat{G}_{21}, \widehat{G}_{22} \in \mathcal{RH}_2$ and $\widehat{G}, \widehat{G}_{12} \in \mathcal{RH}_\infty$.
2. $C^\intercal D = 0$ and $D^\intercal D = I$.

The above assumptions will be loosened in Section 7.2, where a more general problem is considered. The assumption that A is stable is made for the sake of simplicity of presentation and can be relaxed to the assumption that the local pairs (A_{ii}, B_{ii}) must be stabilizable for each $i = 1,\ldots,p$. In that case we can apply the Youla parametrization in Lemma 4.5.1 with a block diagonal matrix E whose block diagonal entries E_{ii} are such that $A_{ii} + B_{ii}E_{ii}$ is stable for each $i = 1,\ldots,p$. In that case we obtain a new system $\widehat{H}$ with a stable state matrix $A + BE$, output matrix $C + DE$ and all other matrices the same as in $\widehat{G}$ (see equation (4.24)). We may then proceed with $\widehat{H}$ instead of $\widehat{G}$. Since the system matrices in (7.1) all have the poset-causal structure determined by $\mathcal{P}$, Proposition 2.3.7 guarantees that each one of the constituent transfer functions of $\widehat{G}$ also has the poset-causal structure, that is,

$$\widehat{G}_{11} \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{n}}, \quad \widehat{G}_{12} \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{m}}, \quad \widehat{G}_{21} \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{n}} \quad \text{and} \quad \widehat{G}_{22} \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}.$$

As in the classical $\mathcal{H}_2$-control problem, the goal is to construct a controller that stabilizes the plant and minimizes the $\mathcal{H}_2$-norm of the resulting closed loop transfer function. For structured linear systems it is more complicated since the communication structure of the plant needs to be preserved under feedback. Shah restricts attention to the case where the controller has the same communication structure as the plant, that is $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$ (see p.92 in [38]). In [45], Swigart considers the more general problem where the controller may have a different communication structure than the plant. He then investigates conditions under which such problems are "tractable". The condition under which this occurs is that the transitive closures of the underlying graphs of the plant and the controller must be equal (see Theorem 2 on p.17 in [45]). Since Swigart also restricts attention to graphs that are acyclic, this is equivalent to the requirement that the controller must satisfy the communication structure determined by the underlying poset of the plant, that is $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$. Hence, we also restrict attention to linear time invariant controllers that has the same poset-causal communication structure as the plant.

By Proposition 4.1.2, connecting a controller $\widehat{K}$ to the plant results in the closed loop transfer function

$$\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{K}(I - \widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21} = \left[\begin{array}{c|c} \underline{A} & \underline{B} \\ \hline \underline{C} & 0 \end{array}\right] = \left[\begin{array}{cc|c} A + BD_K & BC_K & I_{\underline{n}} \\ B_K & A_K & 0 \\ \hline C + DD_K & DC_K & 0 \end{array}\right]. \tag{7.3}$$

The classical (unstructured) $\mathcal{H}_2$-control problem was given in Definition 4.3.1. Since $D_{22} = 0$, the connection between the plant and any controller is well-posed (see proposition 4.1.1). Hence the first condition of the classical problem in Definition 4.3.1 is satisfied. Furthermore $\underline{\mathcal{F}}(\widehat{G},\widehat{K})$ in (7.3) is a strictly proper rational matrix function (note that $\underline{D}=0$, because $D_{11}=0$ and $D_{21}=0$). It follows then that $\underline{\mathcal{F}}(\widehat{G},\widehat{K})$ is in $\mathcal{RH}_2$ if and only if $\underline{A}$ is stable. The $\mathcal{H}_2$-control problem for poset-causal systems, has an additional requirement, namely that the controller must preserve the poset-causal structure, that is, $\underline{\mathcal{F}}(\widehat{G},\widehat{K}) \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{n}}$. Theorem 2.3.10 guarantees that this will happen if attention is restricted to controllers $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$ with the same communication structure as the plant. Consider a controller

$$\widehat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]. \tag{7.4}$$

We shall assume that for some $n_K \in \mathbb{Z}_+$ and some partition $\underline{n}_K \in \mathbb{Z}_+^p$ of n_K, we have

$$A_K \in \mathcal{I}_{\mathcal{P}}^{\underline{n}_K\times\underline{n}_K}, \qquad B_K \in \mathcal{I}_{\mathcal{P}}^{\underline{n}_K\times\underline{n}}, \qquad C_K \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}_K}, \qquad D_K \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}} \tag{7.5}$$

Then $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$. At the level of subsystems, we have

$$\begin{aligned} \dot{x}_{K,i}(t) &= \sum_{j\in\uparrow i} A_{K,ij}x_{K,j}(t) + \sum_{j\in\uparrow i} B_{K,ij}x_j(t), \quad x_{K,i}(0) = x_{K,i,0}, \\ u_i(t) &= \sum_{j\in\uparrow i} C_{K,ij}x_{K,j}(t) + \sum_{j\in\uparrow i} D_{K,ij}x_j(t), \quad t \geq 0. \end{aligned}$$

Hence the stipulation $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$ means that the local input u_i depends on states x_j that are upstream from subsystem i, that is, on x_j for $j \in \uparrow i$. While the full state $x(t)$ of the poset-causal plant 7.1 is available to the controller $\widehat{K}$ in (7.4) via its input $u_K(t) = y(t) = x(t)$, the fact that the system matrices A_K, B_K, C_K and D_K of the controller are structured as in (7.5), implies that, at the level of subsystems of the controller, subsystem i of the controller only has access to local states upstream from i, that is to x_j for $j \in \uparrow i$. Having discussed the connection between an input-output framework and overall transfer function framework, we will henceforth only consider the control problem on the transfer function level. The state feedback case of the $\mathcal{H}_2$-control problem for poset-causal systems may now be stated as follows.

Problem 7.1.1 ($\mathcal{H}_2$ state feedback control problem for poset-causal systems (cf. p.93 in [38])). For a $\mathcal{P}$-poset-causal plant as in (7.1), construct a controller

$$\widehat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$$

satisfying

1. $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$;
2. $\underline{\mathcal{F}}(\widehat{G},\widehat{K}) \in \mathcal{RH}_2$;

such that $\|\underline{\mathcal{F}}(\widehat{G},\widehat{K})\|_2^2$ is minimized.

Condition 1 requires that the controller $\widehat{K}$ has the poset-causal structure. Proposition 2.3.7 then guarantees that the poset-causal structure is preserved in the closed loop transfer function, that is, $\underline{\mathcal{F}}(\widehat{G},\widehat{K}) \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{n}}$. Condition 2 requires that the controller is stabilizing, that is, that $\underline{A}$ in (7.3) is a stable matrix.

7.1.2 Reparametrization and Reduction to Local Classical Problems

Since we have the same control set-up as in Section 4.5 with an additional structural requirement, we can use the same parametrization as in Procedure 4.5.2 if care is taken with regard to the poset-causal structure.

Since it is assumed that A is stable, we may take $E = 0$ in (4.31). Hence by the Youla parametrization for the state feedback case (Lemma 4.5.1), all stabilizing controllers of $\widehat{G}$ are given by

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J},\widehat{R}) = \widehat{R}(I + \widehat{G}_{22}\widehat{R})^{-1}$$

with $\widehat{R} \in \mathcal{H}_\infty$ arbitrary (note that $\widehat{H} = \widehat{G}$ in (4.32)). Furthermore, the closed loop transfer function after connecting $\widehat{K}$ to $\widehat{G}$ is

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{R}\widehat{G}_{21}.$$

Let $\widehat{Q} = \widehat{R}\widehat{G}_{21}$. Then $\widehat{Q}$ and $\widehat{R}$ determine each other uniquely, because $\widehat{G}_{21}$ is invertible. Furthermore, since $\widehat{G}_{21} \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{n}}$, Proposition 2.3.7 ensures that $\widehat{Q} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$ if and only if $\widehat{R} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$. Now we have

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{Q}$$

and

$$\widehat{K} = \widehat{Q}\widehat{G}_{21}^{-1}(I + \widehat{G}_{22}\widehat{Q}\widehat{G}_{21}^{-1})^{-1}. \tag{7.6}$$

Again, since $\widehat{G}_{21} \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{n}}$ and $\widehat{G}_{22} \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}$, it follows by Proposition 2.3.7 and equation (7.6) that $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$ if and only if $\widehat{Q} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$, that is, the controller $\widehat{K}$ satisfies the structural requirement imposed by the poset $\mathcal{P}$ if and only if the parameter $\widehat{Q}$ does. Finally, suppose

$$\widehat{Q} = \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & D_Q \end{array}\right].$$

Then by equations (3.13) and (3.12), it follows that

$$\underline{\mathcal{F}}(\widehat{G},\widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{Q} = \left[\begin{array}{c|c} A & I_{\underline{n}} \\ \hline C & 0 \end{array}\right] + \left[\begin{array}{cc|c} A & BC_Q & BD_Q \\ 0 & A_Q & B \\ \hline C & DC_Q & DD_Q \end{array}\right] = \left[\begin{array}{ccc|c} A & 0 & 0 & I_{\underline{n}} \\ 0 & A & BC_Q & BD_Q \\ 0 & 0 & A_Q & B \\ \hline C & C & DC_Q & DD_Q \end{array}\right].$$

Now $\underline{\mathcal{F}}(\widehat{G},\widehat{K}) \in \mathcal{RH}_2$ if the state matrix is stable and $DD_Q = 0$. Since A is stable by assumption, this happens if A_Q is stable and $D_Q = 0$. Due to the above considerations, we obtain the following proposition which shows that Problem 7.1.1 is equivalent to a reparameterized $\mathcal{H}_2$-control problem (cf. p.96 of [38]).

Problem 7.1.2 (Reparameterized $\mathcal{H}_2$-control problem for poset-causal systems for the state feedback case)**.** For $\widehat{G}_{11}$ and $\widehat{G}_{12}$ as in (7.2) find a controller

$$\widehat{Q} = \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & 0 \end{array}\right] \qquad \text{satisfying}$$

1. $\widehat{Q} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$;
2. A_Q is stable;
3. $\widehat{K} = \widehat{Q}\widehat{G}_{21}^{-1}(I + \widehat{G}_{22}\widehat{Q}\widehat{G}_{21}^{-1})^{-1}$ is proper

such that the cost $\|\widehat{G}_{11} + \widehat{G}_{12}\widehat{Q}\|_2^2$ is minimized.

Proposition 7.1.3 (cf. p.96 in [38])**.**
A controller $\widehat{K} = \widehat{Q}\widehat{G}_{21}^{-1}(I+\widehat{G}_{22}\widehat{Q}\widehat{G}_{21}^{-1})^{-1}$ is an optimal solution to the $\mathcal{H}_2$-control problem in Problem 7.1.1 if and only if $\widehat{Q}$ is an optimal solution of the $\mathcal{H}_2$-control Problem 7.1.2.

Next, we explain how we may reduce the above reparameterized control problem to p local control problems. This can be done in such a way that, in contrast to the overall reparameterized control problem, the local control problems have no structural restrictions. Hence the classical results of Chapter 4 can then be applied to solve these local control problems. Due to the column separability of the $\mathcal{H}_2$-norm, it follows that the cost is equal to the sum of p terms

$$\left\|\widehat{G}_{11} + \widehat{G}_{12}\widehat{Q}\right\|_2^2 = \sum_{j=1}^{p} \left\|\left(\widehat{G}_{11} + \widehat{G}_{12}\widehat{Q}\right)(:,j)\right\|_2^2.$$

Applying Theorem 2.3.10 to each of these p terms then gives

$$\left\|\widehat{G}_{11} + \widehat{G}_{12}\widehat{Q}\right\|_2^2 = \sum_{j=1}^{p} \left\|\widehat{G}_{11}(:,j) + \widehat{G}_{12}(:,\downarrow j)\widehat{Q}(\downarrow j,j)\right\|_2^2.$$

By Theorem 2.3.10 and Corollary 2.3.11, it follows that

$$\begin{aligned}\widehat{Q}(\downarrow j, j) &= (C_Q(\lambda I - A_Q)^{-1}B_Q)(\downarrow j, j)\\ &= C_Q(\downarrow j,\downarrow j)(\lambda I(\downarrow j,\downarrow j) - A_Q(\downarrow j,\downarrow j))^{-1}B_Q(\downarrow j, j) = \left[\begin{array}{c|c} A_Q(\downarrow j,\downarrow j) & B_Q(\downarrow j, j)\\ \hline C_Q(\downarrow j,\downarrow j) & 0\end{array}\right].\end{aligned}$$

Similarly,

$$\widehat{G}_{11}(:, j) = \left[\begin{array}{c|c} A(\downarrow j,\downarrow j) & I_{\underline{n}}(\downarrow j, j)\\ \hline C(:,\downarrow j) & 0\end{array}\right] \quad \text{and} \quad \widehat{G}_{12}(:,\downarrow j) = \left[\begin{array}{c|c} A(\downarrow j,\downarrow j) & B(\downarrow j,\downarrow j)\\ \hline C(:,\downarrow j) & D(:,\downarrow j)\end{array}\right]. \tag{7.7}$$

Importantly, $\widehat{Q}(\downarrow j, j)$ does not have a specific zero-structure. Furthermore, A_Q is stable if and only if the local matrices $A_Q(\downarrow j,\downarrow j)$ are stable for $j = 1, \ldots, p$. The previous discussion is a sketch of the proof of following result which shows that solving the reparameterized problem is equivalent to solving p local control problems.

Problem 7.1.4.
For $j \in P$ and for $\widehat{G}_{11}(:, j)$ and $\widehat{G}_{12}(:,\downarrow j)$ as in (7.7), construct a controller

$$\widehat{Q}_j = \left[\begin{array}{c|c} A_Q^j & B_Q^j\\ \hline C_Q^j & 0\end{array}\right] \in \mathcal{RH}_2,$$

such that $\left\|\widehat{G}_{11}(:, j) + \widehat{G}_{12}(:,\downarrow j)\widehat{Q}_j\right\|_2^2$ is minimized.

Proposition 7.1.5 (cf. Theorem 4.2 in [38])**.**
The controller $\widehat{Q} = \begin{bmatrix} I_{\underline{m}}(:,\downarrow 1)\widehat{Q}(\downarrow 1, 1) & \ldots & I_{\underline{m}}(:,\downarrow p)\widehat{Q}(\downarrow p, p)\end{bmatrix}$ *is an optimal solution of Problem 7.1.2 if and only if* $\widehat{Q}(\downarrow j, j) = \widehat{Q}_j$ *for each* $j = 1, \ldots, p$ *where* $\widehat{Q}_j$ *is the optimal solution of the local Problem 7.1.4.*

7.1.3 Solution to the Local Classical Problems

Since A is assumed to be stable, each sub-matrix $A(\downarrow j,\downarrow j)$ is stable. As a consequence, each triple $(A(\downarrow j,\downarrow j), B(\downarrow j,\downarrow j), C(:,\downarrow j))$ is stabilizable and detectable. Furthermore, $D^\intercal C = 0$ and $D^\intercal D = I$ if and only if $D^\intercal(\downarrow j, :)C(:,\downarrow j) = 0$ and $D^\intercal(\downarrow j, :)D(:,\downarrow j) = I(\downarrow j,\downarrow j)$. Hence, each one of the above mentioned control problems is a classical optimization problem of the form in Definition 4.3.1. Set

$$\widehat{H}_{11} = \widehat{G}_{11}(:, j) \in \mathcal{RH}_2 \quad \text{and} \quad \widehat{H}_{12} = \widehat{G}_{12}(:,\downarrow j) \in \mathcal{RH}_\infty$$

as in (7.7). By Corollary 3.7.5, the conditions on $(A(\downarrow j,\downarrow j), B(\downarrow j,\downarrow j), C(:,\downarrow j), D(:,\downarrow j))$ guarantee the existence of a stabilizing solution to the related Riccati equation. Applying Theorem 4.7.6, we can now solve each one of the local $\mathcal{H}_2$-optimization problems in Problem 7.1.4.

Proposition 7.1.6 (cf. Lemma 4.2 in [38])**.**
The optimal controller for Problem 7.1.4 is given by

$$\widehat{Q}_j = \left[\begin{array}{c|c} A(\downarrow j,\downarrow j) + B(\downarrow j,\downarrow j)F_j & I_{\underline{n}}(\downarrow j, j)\\ \hline F_j & 0\end{array}\right], \tag{7.8}$$

where

$$F_j = -(B(\downarrow j,\downarrow j)^\intercal X_j + D(:,\downarrow j)^\intercal C(:,\downarrow j)) \in \mathbb{R}^{\underline{m}_{\downarrow j}\times \underline{n}_{\downarrow j}} \tag{7.9}$$

and

$$X_j = \operatorname{Ric}(A(\downarrow j,\downarrow j), B(\downarrow j,\downarrow j), C(:,\downarrow j), D(:,\downarrow j)) \in \mathbb{R}^{\underline{n}_{\downarrow j}\times \underline{n}_{\downarrow j}}.$$

7.1.4 Notational Preliminaries

We interrupt our solution of the state feedback case with a subsection on notation in order to exhibit the solution more clearly. In this subsection, we establish notation related to partitions, sub-partitions and the sizes associated with them. This will be utilized in the following subsections to indicate the sizes of matrices appearing in a solution of the poset-causal state feedback $\mathcal{H}_2$-control problem.

Consider a poset $\mathcal{P} = (P, \succeq)$ with $P = \{1, \ldots, p\}$. For each $j = 1, \ldots, p$, let q_j be the size of the downstream set $\downarrow j$ and let q be the sum of these sizes, that is

$$q_j = |\downarrow j| \quad \text{for } j = 1, \ldots, p \qquad \text{and} \qquad q = \sum_{j=1}^{p} q_j.$$

Define the numbers $r_0, r_1, \ldots, r_j$ inductively as follows

$$r_0 = 0 \qquad \text{and} \qquad r_j = r_{j-1} + q_j = q_1 + q_2 + \ldots + q_j, \qquad j = 1, \ldots, p.$$

Then in particular $r_1 = q_1$ and $r_p = q$. This enables us to enumerate each downstream set $\downarrow j$ for $j = 1, \ldots, p$ as

$$\downarrow j = \{r_{j-1} + 1,\ r_{j-1} + 2,\ r_{j-1} + 3, \ldots,\ r_{j-1} + (q_j - 1),\ r_j\}.$$

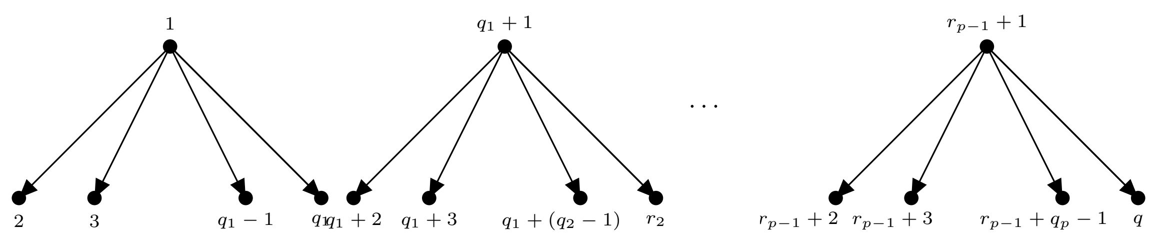
Define the set $P_\downarrow$ as $\{1, \ldots, q\}$ enumerated in the following way:

$$P_\downarrow = \{1, \ldots, q_1, \quad q_1 + 1, \ldots, r_2, \quad r_2 + 1, \ldots, r_3, \quad \ldots \quad , r_{p-1} + 1, \ldots, q\} = \{1, \ldots, q\}$$

and the auxiliary poset $(P_\downarrow, \succeq_\downarrow)$, where

$$1 \succeq_\downarrow 2, \ldots, q_1, \quad q_1 + 1 \succeq_\downarrow q_1 + 2, \ldots,\ r_2, \qquad \ldots, \qquad r_{p-1} + 1 \succeq_\downarrow r_{p-1} + 2, \ldots q.$$

This auxiliary poset $(P_\downarrow, \succeq_\downarrow)$ has the following Hasse diagram:



Furthermore, define the set P_1 as the set of leaders and $P_{\twoheaddownarrow}$ as the set of followers in the above Hasse diagram, that is

$$P_1 = \{1, r_1 + 1, r_2 + 1, \ldots, r_{p-1} + 1\} \qquad \text{and}$$
$$P_{\twoheaddownarrow} = \{2, \ldots, r_1, \quad r_1 + 2, \ldots, r_2, \quad \ldots, \quad r_{p-1}, \quad r_{p-1} + 2, \ldots, q\} = P_\downarrow / P_1.$$

Then the sizes of the sets $P_\downarrow$, P_1 and $P_{\twoheaddownarrow j}$ are $|P_\downarrow| = q$, $|P_1| = p$ and $|P_{\twoheaddownarrow j}| = q - p$ respectively. We note that $P_\downarrow = \downarrow P_1$ in the auxiliary poset $(P_\downarrow, \succeq_\downarrow)$.

Suppose $\underline{n} = (n_1, \ldots, n_p) \in \mathbb{Z}_+^p$. For each $j = 1, \ldots, p$, define

$$N_j = \sum_{k \in \downarrow j} n_k = n_{r_{j-1}+1} + \ n_{r_{j-1}+2} + \ n_{r_{j-1}+3} + \ \ldots + \ n_{r_j} \qquad \text{and let} \qquad N = \sum_{j=1}^{p} N_j = \sum_{j=1}^{p} \sum_{k \in \downarrow j} n_k.$$

Define the sub-partitions

$$\underline{n}_\downarrow = (\underline{n}_{\downarrow 1}, \ldots, \underline{n}_{\downarrow p}) \in \mathbb{Z}_+^{p^2} \qquad \text{and} \qquad \underline{n}_{\twoheaddownarrow} = (\underline{n}_{\twoheaddownarrow 1}, \ldots, \underline{n}_{\twoheaddownarrow p}) \in \mathbb{Z}_+^{p^2},$$

where $\underline{n}_{\downarrow j} \in \mathbb{Z}_+^p$ and $\underline{n}_{⤈ j} \in \mathbb{Z}_+^p$ are defined as in Definition 2.3.8. Recall that the k^{th} entry of $\underline{n}_{\downarrow j}$ and $\underline{n}_{⤈ j}$ is zero if $k \notin \downarrow j$ and $k \notin ⤈ j$ respectively. By omitting the zero-entries in $\underline{n}_{\downarrow j}$ and $\underline{n}_{⤈ j}$, the partitions $\underline{n}_{\downarrow j}$ and $\underline{n}_{⤈ j}$ can also be interpreted as partitions in $\mathbb{Z}_+^{q_j}$ and $\mathbb{Z}_+^{q_j-1}$ respectively. In this case, we can represent the partitions $\underline{n}_{\downarrow j}$ and $\underline{n}_{⤈ j}$ as

$$\underline{n}_{\downarrow j} = (\widetilde{n}_{r_{j-1}+1},\ \widetilde{n}_{r_{j-1}+2}, \dots,\ \widetilde{n}_{r_j}) \in \mathbb{Z}_+^{q_j} \qquad \text{and} \qquad \underline{n}_{⤈ j} = (\widetilde{n}_{q_{j-1}+2},\ \dots,\ \widetilde{n}_{r_j}) \in \mathbb{Z}_+^{q_j-1}$$

respectively, where $\widetilde{n}_1, \dots, \widetilde{n}_q \in \{n_1, n_2, \dots, n_p\}$. Similarly, the sub-partitions $\underline{n}_{\downarrow} \in \mathbb{Z}_+^{p^2}$ and $\underline{n}_{⤈} \in \mathbb{Z}_+^{p^2}$ can also be interpreted as sub-partitions in $\mathbb{Z}_+^q$ and $\mathbb{Z}_+^{q-p}$ respectively, specifically

$$\begin{aligned} \underline{n}_{\downarrow} &= (\widetilde{n}_1,\ \widetilde{n}_2,\ \dots,\ \widetilde{n}_{q_1},\quad \widetilde{n}_{q_1+1},\ \widetilde{n}_{q_1+2},\ \dots,\ \widetilde{n}_{q_1+q_2},\quad \dots,\quad \widetilde{n}_{r_{j-1}+1},\ \widetilde{n}_{r_{j-1}+2},\ \dots,\ \widetilde{n}_q) \in \mathbb{Z}_+^q \quad \text{and} \\ \underline{n}_{⤈} &= (\widetilde{n}_2,\ \widetilde{n}_3,\ \dots,\ \widetilde{n}_{q_1},\quad \widetilde{n}_{q_1+2},\ \widetilde{n}_{q_1+3},\ \dots,\ \widetilde{n}_{q_1+q_2},\quad \dots,\quad \widetilde{n}_{r_{j-1}+2},\ \widetilde{n}_{r_{j-1}+3},\ \dots,\ \widetilde{n}_q) \in \mathbb{Z}_+^{q-p}. \end{aligned} \tag{7.10}$$

Now by definition

$$|\underline{n}_{\downarrow j}| = N_j, \qquad |\underline{n}_{⤈ j}| = N_j - n_j, \qquad |\underline{n}_{\downarrow}| = N \quad \text{and} \quad |\underline{n}_{⤈}| = N - n.$$

Note that $P_1 \subset P_{\downarrow} = \{1, \dots, q\}$ and $P_{⤈} \subset P_{\downarrow} = \{1, \dots, q\}$. Interpreting $\underline{n}_{\downarrow}$ and $\underline{n}_{⤈}$ as vectors (or row matrices) in $\mathbb{Z}_+^q$ and $\mathbb{Z}_+^{q-p}$ respectively as in (7.10) and using the notation of Definition 2.3.8, it follows that

$$\underline{n}_{\downarrow}(1, P_1) = \underline{n} \in \mathbb{Z}_+^q \qquad \text{and} \qquad \underline{n}_{\downarrow}(1, P_{⤈}) = \underline{n}_{⤈} \in \mathbb{Z}_+^{q-p}$$

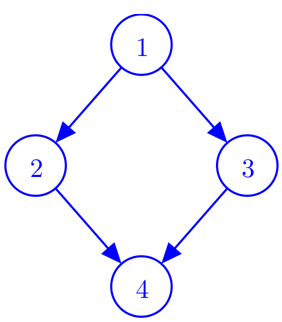
and hence

$$I_{\underline{n}_{\downarrow}}(P_1, P_1) = I_{\underline{n}} \qquad \text{and} \qquad I_{\underline{n}_{\downarrow}}(P_{⤈}, P_{⤈}) = I_{\underline{n}_{⤈}}. \tag{7.11}$$

Given $\underline{m} = (m_1, \dots, m_p) \in \mathbb{Z}_+^p$ and $\underline{r} = (r_1, \dots, r_p) \in \mathbb{Z}_+^p$, define the (sub)-partitions $\underline{m}_{\downarrow}$, $\underline{m}_{⤈}$, $\underline{r}_{\downarrow}$ and $\underline{r}_{⤈}$ in the same manner.

The following example illustrates the use of the notation defined above. In the next section, we will consider an example of the state-feedback $\mathcal{H}_2$-control problem for a poset-causal system given by Shah in Section 4.5.4 of [38]. This example also illustrates the notation used in that example.

Example 7.1.7 (Notation for Shah's example (Section 4.5.4 [38])).
Let $\mathcal{P} = (P, \succeq)$ be the poset with $P = \{1, 2, 3, 4\}$ and $1 \succeq 2$, $1 \succeq 3$, $2 \succeq 4$ and $3 \succeq 4$ as illustrated by the following Hasse diagram



Then

$$\begin{aligned} &\downarrow 1 = \{1,2,3,4\} && \text{and} && q_1 = |\downarrow 1| = 4 \\ &\downarrow 2 = \{2,4\} && \text{and} && q_2 = |\downarrow 2| = 2 \\ &\downarrow 3 = \{3,4\} && \text{and} && q_3 = |\downarrow 3| = 2 \\ &\downarrow 4 = \{4\} && \text{and} && q_4 = |\downarrow 4| = 1 \end{aligned}$$

$$\text{so} \quad q = \sum_{j=1}^{4} q_j = 9.$$

For the given poset $\mathcal{P}$, the sets $P_{\downarrow}$, P_1 and $P_{⤈}$ are given by

$$P_{\downarrow} = \{1,2,3,4,5,6,7,8,9\}, \quad P_1 = \{1,5,7,9\} \quad \text{and} \quad P_{⤈} = \{2,3,4,6,8\} = P_{\downarrow}/P_1$$

and we see that

$$|P_{\downarrow}| = 9 = q, \quad |P_1| = 4 = p \quad \text{and} \quad |P_{⤈}| = 5 = q - p.$$

Consider a partition $\underline{n} = (n_1, n_2, n_3, n_4)$. Then $n = n_1 + n_2 + n_3 + n_4$, $N_1 = n_1 + n_2 + n_3 + n_4 = n$, $N_2 = n_2 + n_4$, $N_3 = n_3 + n_4$, $N_4 = n_4$ and $N = (n_1 + n_2 + n_3 + n_4) + (n_2 + n_4) + (n_3 + n_4) + (n_4)$. By Definition 2.3.8, we have

$$\underline{n}_{\downarrow 1} = (n_1, n_2, n_3, n_4), \quad \underline{n}_{\downarrow 2} = (n_2, n_4), \quad \underline{n}_{\downarrow 3} = (n_3, n_4) \quad \text{and} \quad \underline{n}_{\downarrow 4} = (n_4)$$

and we see that $|\underline{n}_{\downarrow j}| = N_j$ for $j = 1, 2, 3, 4$. The sub-partitions $\underline{n}_{\downarrow}$ and $\underline{n}_{\twoheaddownarrow}$ are given by

$$\underline{n}_{\downarrow} = (n_1, n_2, n_3, n_4, n_2, n_4, n_3, n_4, n_4) \quad \text{and} \quad \underline{n}_{\twoheaddownarrow} = (n_2, n_3, n_4, n_4, n_4).$$

It is noted that

$$\underline{n}_{\downarrow}(1, P_{\twoheaddownarrow}) = \underline{n}_{\downarrow}(1, \{2, 3, 4, 6, 8\}) = (n_2, n_3, n_4, n_4, n_4) = \underline{n}_{\twoheaddownarrow} \qquad \text{and}$$
$$\underline{n}_{\downarrow}(1, P_1) = \underline{n}_{\downarrow}(1, \{1, 5, 7, 9\}) = (n_1, n_2, n_3, n_4) = \underline{n}.$$

Furthermore $|\underline{n}_{\downarrow}| = N$ and $|\underline{n}_{\twoheaddownarrow}| = N - n$.

For example, if $\underline{n} = (1, 1, 1, 1)$, then $n = 4$, $N_1 = 4$, $N_2 = 2$, $N_3 = 2$, $N_4 = 1$ and $N = 9$. By Definition 2.3.8, $\underline{n}_{\downarrow 1} = (1, 1, 1, 1)$, $\underline{n}_{\downarrow 2} = (1, 1)$, $\underline{n}_{\downarrow 3} = (1, 1)$ and $\underline{n}_{\downarrow 4} = (1)$, so we see that $|\underline{n}_{\downarrow j}| = N_j$ for $j = 1, 2, 3, 4$. In this case, the sub-partitions $\underline{n}_{\downarrow}$ and $\underline{n}_{\twoheaddownarrow}$ are simply given by

$$\underline{n}_{\downarrow} = (1, 1, 1, 1, 1, 1, 1, 1, 1) \quad \text{and} \quad \underline{n}_{\twoheaddownarrow} = (1, 1, 1, 1, 1),$$

so that $|\underline{n}_{\downarrow}| = 9$ and $|\underline{n}_{\twoheaddownarrow}| = 5 = 9 - 4$.

7.1.5 Solution to the Reparameterized Problem

By Theorem 2.3.10, it is clear that

$$\widehat{Q} = \begin{bmatrix} \widehat{Q}(:, 1) & \dots & \widehat{Q}(:, p) \end{bmatrix} = \begin{bmatrix} I_{\underline{m}}(:, \downarrow 1)\widehat{Q}(\downarrow 1, 1) & \dots & I_{\underline{m}}(:, \downarrow p)\widehat{Q}(\downarrow p, p) \end{bmatrix}.$$

The optimal solution of the reparameterized problem is now constructed from the local optimal solutions (7.8). Let $F_j \in \mathbb{R}^{\underline{m}_{\downarrow j} \times \underline{n}_{\downarrow j}}$ be as in (7.9) and $\widehat{Q}_j$ as in (7.8). Set $\widehat{Q}(\downarrow j, j) = \widehat{Q}_j$ for each $j = 1, \dots, p$. Then

$$I_{\underline{m}}(:, \downarrow j)\widehat{Q}(\downarrow j, j) = I_{\underline{m}}(:, \downarrow j)\widehat{Q}_j = \left[\begin{array}{c|c} A(\downarrow j, \downarrow j) + B(\downarrow j, \downarrow j)F_j & I_{\underline{n}}(\downarrow j, j) \\ \hline I_{\underline{m}}(:, \downarrow j)F_j & 0 \end{array}\right]$$

for each $j = 1, \dots, p$. Using the row concatenation formula for realizations (3.16), it follows that

$$\begin{aligned} \widehat{Q} &= \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & 0 \end{array}\right] \\ &= \left[\begin{array}{ccc|ccc} A(\downarrow 1, \downarrow 1) + B(\downarrow 1, \downarrow 1)F_1 & \dots & 0 & I_{\underline{n}}(\downarrow 1, 1) & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & A(\downarrow p, \downarrow p) + B(\downarrow p, \downarrow p) & 0 & \dots & I_{\underline{n}}(\downarrow p, p) \\ \hline I_{\underline{m}}(:, \downarrow 1)F_1 & \dots & I_{\underline{m}}(:, \downarrow p)F_p & 0 & \dots & 0 \end{array}\right]. \end{aligned} \tag{7.12}$$

Here $A_Q \in \mathbb{R}^{\underline{n}_{\downarrow} \times \underline{n}_{\downarrow}}$, $B_Q \in \mathbb{R}^{\underline{n}_{\downarrow} \times \underline{n}}$ and $C_Q \in \mathbb{R}^{\underline{m} \times \underline{n}_{\downarrow}}$ using the notation defined in the previous subsection. Recall that in order for $\widehat{Q}$ to be a solution, it has to satisfy the following conditions:

1. $\widehat{Q} \in \mathcal{I}_{\mathcal{P}}^{\underline{m} \times \underline{n}}$;

2. A_Q is stable;

3. $\widehat{K} = \widehat{Q}\widehat{G}_{21}^{-1}(I + \widehat{G}_{22}\widehat{Q}\widehat{G}_{21}^{-1})^{-1}$ is proper.

Since A_Q and B_Q are $p \times p$ block diagonal matrices, it is clearly the case that $A_Q \in \mathcal{I}_{\mathcal{P}}^{\underline{n}_\downarrow \times \underline{n}_\downarrow}$ and $B_Q \in \mathcal{I}_{\mathcal{P}}^{\underline{n}_\downarrow \times \underline{n}}$. Next, note that $F_j \in \mathbb{R}^{\underline{m}_{\downarrow j} \times \underline{n}_{\downarrow j}}$ and that the j-th column of the matrix C_Q is the matrix $I_{\underline{m}}(:, \downarrow j)F_j$ which itself has $|\downarrow j| = q_j$ columns. Thus the ij-th entry of the matrix C_Q is the i-th row of the matrix $I_{\underline{m}}(:, \downarrow j)F_j$, that is,

$$C_Q(i,j) = \begin{bmatrix} I_{\underline{m}}(:, \downarrow 1)F_1 & \dots & I_{\underline{m}}(:, \downarrow p)F_p \end{bmatrix}(i,j) = (I_{\underline{m}}(:, \downarrow j)F_j)(i, \downarrow j).$$

By Theorem 2.3.10 and the fact that $\downarrow\downarrow j = \downarrow j$, it follows that

$$\begin{aligned} C_Q(i,j) &= (I_{\underline{m}}(:, \downarrow j))(i, \downarrow\downarrow j)F_j(\downarrow\downarrow j, \downarrow j) \\ &= I_{\underline{m}}(i, \downarrow j)F_j(\downarrow j, \downarrow j) \\ &= I_{\underline{m}}(i, \downarrow j)F_j. \end{aligned}$$

If $j \not\succeq i$, that is, if $i \notin \downarrow j$, then $I_{\underline{m}}(i, \downarrow j) = 0$. Thus $C_Q(i,j) = 0$ if $j \not\succeq i$, which shows that $C_Q \in \mathcal{I}_{\mathcal{P}}^{\underline{m} \times \underline{n}_\downarrow}$. Since $A_Q \in \mathcal{I}_{\mathcal{P}}^{\underline{n}_\downarrow \times \underline{n}_\downarrow}$, $B_Q \in \mathcal{I}_{\mathcal{P}}^{\underline{n}_\downarrow \times \underline{n}}$ and $C_Q \in \mathcal{I}_{\mathcal{P}}^{\underline{m} \times \underline{n}_\downarrow}$, Proposition 2.3.7 guarantees that $\widehat{Q} \in \mathcal{I}_{\mathcal{P}}^{\underline{m} \times \underline{n}}$, that is $\widehat{Q}$ satisfies the first condition.

Secondly, for each $j = 1, \dots, p$, the matrix $A(\downarrow j, \downarrow j) + B(\downarrow j, \downarrow j)F_j$ is stable. Since it is the j-th block-diagonal entry of the block diagonal matrix A_Q, the matrix A_Q is stable. This shows that $\widehat{Q}$ satisfies the second condition.

The third condition will be verified in the next section, where a solution to the original problem is constructed.

We illustrate the solution to the reparameterized problem for a poset-causal system with underlying poset and partitions given in Example 7.1.7.

Example 7.1.8. (cf. Section 4.5.4 on p.116 of [38])
Let $\mathcal{P} = (P, \succeq)$ be the poset with $P = \{1,2,3,4\}$ and $1 \succeq 2$, $1 \succeq 3$, $2 \succeq 4$ and $3 \succeq 4$ as in Example 7.1.7. Let $\underline{n} = (1,1,1,1)$, $\underline{m} = (1,1,1,1)$ and $\underline{r} = (2,2,2,2)$ be partitions of the state space, input space and output space dimensions of the realization a poset-causal system respectively. From Example 7.1.7, we have that $n = 4$, $m = 4$ and $N = 9$. Consider a poset-causal system of the form (7.1) and with system matrices

$$A = \begin{bmatrix} -0.5 & 0 & 0 & 0 \\ -1 & -0.25 & 0 & 0 \\ -1 & 0 & -0.2 & 0 \\ -1 & -1 & -1 & -0.1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then $A \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{n}}$, $B \in \mathcal{I}_{\mathcal{P}}^{\underline{n} \times \underline{m}}$, $C \in \mathcal{I}_{\mathcal{P}}^{\underline{r} \times \underline{n}}$, $D \in \mathcal{I}_{\mathcal{P}}^{\underline{r} \times \underline{m}}$ and $D^\intercal C = 0$. Let

$$\widehat{G} = \left[\begin{array}{c|cc} A & I & B \\ \hline C & 0 & D \\ I & 0 & 0 \end{array}\right] = \begin{bmatrix} C\Phi_A & C\Phi_A B + D \\ \Phi_A & \Phi_A B \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix}.$$

Suppose we want to solve the optimization problem $\min_{\hat{Q} \in I_{\mathcal{P}}} \|\hat{G}_{11} + \hat{G}_{12}\hat{Q}\|_2^2$. In order to employ the solution (7.12), we solve p local Riccati equations of the form (3.29):

$$X_j = \operatorname{Ric}\left(A(\downarrow j, \downarrow j), B(\downarrow j, \downarrow j), C(:, \downarrow j), D(:, \downarrow j)\right). \tag{7.13}$$

Hence, we calculate

$$A(\downarrow 1,\downarrow 1) = A \qquad A(\downarrow 2,\downarrow 2) = \begin{bmatrix} -0.25 & 0 \\ -1 & -0.1 \end{bmatrix}, \quad A(\downarrow 3,\downarrow 3) = \begin{bmatrix} -0.2 & 0 \\ -1 & -0.1 \end{bmatrix} \quad \text{and}$$
$$A(\downarrow 4,\downarrow 4) = -0.1$$
$$B(\downarrow 1,\downarrow 1) = B \qquad B(\downarrow 2,\downarrow 2) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad B(\downarrow 3,\downarrow 3) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$
$$B(\downarrow 4,\downarrow 4) = 1$$

$$C(:,\downarrow 1) = C, \qquad C(:,\downarrow 2) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad C(:,\downarrow 3) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad C(:,\downarrow 4) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$D(:,\downarrow 1) = D, \qquad D(:,\downarrow 2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad D(:\downarrow 3) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad D(:,\downarrow 4) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

We obtain the p local continuous time Riccati equation solutions X_j (7.13) using the MATLAB command "icare":

$$X_1 = \begin{bmatrix} 3.3490 & -1.6008 & -1.6643 & 0.6337 \\ -1.6008 & 1.7477 & 0.9948 & -0.7902 \\ -1.6643 & 0.9948 & 1.8432 & -0.8121 \\ 0.6337 & -0.7902 & -0.8121 & 0.8935 \end{bmatrix}$$
$$X_2 = \begin{bmatrix} 1.8248 & -0.8011 \\ -0.8011 & 0.9001 \end{bmatrix}, \quad X_3 = \begin{bmatrix} 1.9186 & -0.8226 \\ -0.8226 & 0.9019 \end{bmatrix}, \quad X_4 = 0.905.$$

Using the solutions X_j of the local Riccati equations, we define the static feedback matrices $F_j = -(D(\downarrow j,\downarrow j)^\intercal D(\downarrow j,\downarrow j))^{-1} B(\downarrow j,\downarrow j)^\intercal X_j$. This is also given by the MATLAB command "icare":

$$F_1 = \begin{bmatrix} -0.7175 & -0.3515 & -0.3616 & 0.0751 \\ 0.9671 & -0.9575 & -0.1827 & -0.1033 \\ 1.0306 & -0.2045 & -1.0312 & -0.0814 \\ -0.6337 & 0.7902 & 0.8121 & -0.8935 \end{bmatrix}, \qquad F_2 = \begin{bmatrix} -1.0237 & -0.0990 \\ 0.8011 & -0.9001 \end{bmatrix}$$
$$F_3 = \begin{bmatrix} -1.0960 & -0.0792 \\ 0.8226 & -0.9019 \end{bmatrix} \quad \text{and} \qquad F_4 = -0.905.$$

So now we can compute the state space matrices of the optimal controller $\widetilde{Q}$ using equation (7.12):

$$A_Q = \operatorname{diag}\left(A(\downarrow j,\downarrow j) + B(\downarrow j,\downarrow j)F_j\right) \in \mathbb{R}^{N\times N},$$
$$B_Q = \operatorname{diag}\left(I_{\underline{n}}(\downarrow j, j)\right) \in \mathbb{R}^{N\times n} \quad \text{and}$$
$$C_Q = \begin{bmatrix} I_{\underline{m}}(:,\downarrow 1)F_1 & \dots & I_{\underline{m}}(:,\downarrow p)F_p \end{bmatrix} \in \mathbb{R}^{m\times N}.$$

This gives

$$\widetilde{Q} = \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & 0 \end{array}\right]$$

$$= \left[\begin{array}{ccccccccc|cccc}
-1.2175 & -0.3515 & -0.3616 & 0.0751 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-0.7505 & -1.5589 & -0.5443 & -0.0282 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.6869 & -0.5560 & -1.5927 & -0.0064 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.3535 & -1.7232 & -1.7634 & -1.1032 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1.2737 & -0.0990 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1.2226 & -1.0991 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1.2960 & -0.0792 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1.2733 & -1.0811 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1.005 & 0 & 0 & 0 & 1 \\
\hline
-0.7175 & -0.3515 & -0.3616 & 0.0751 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.9671 & -0.9575 & -0.1827 & -0.1033 & -1.0237 & -0.0990 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1.0306 & -0.2045 & -1.0312 & -0.0814 & 0 & 0 & -1.0960 & -0.0792 & 0 & 0 & 0 & 0 & 0 \\
-0.6337 & 0.7902 & 0.8121 & -0.8935 & 0.8011 & -0.9001 & 0.8226 & -0.9019 & -0.905 & 0 & 0 & 0 & 0
\end{array}\right].$$

We can see that A_Q, B_Q and C_Q have the structure

$$\begin{bmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ * & 0 & * & 0 \\ * & * & * & * \end{bmatrix}$$

as determined by the poset $\mathcal{P}$.

7.1.6 Solution to the Original Problem

In this paragraph,

$$\widehat{Q} = \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & 0 \end{array}\right]$$

as in (7.12) is the solution to the reparameterized problem on page 124. We showcase the approach taken by Shah in [38] to obtain an optimal solution $\widehat{K}$ to the original problem from the solution $\widehat{Q}$. We will employ the notation we have defined in Subsection 7.1.4. By applying Lemma 3.5.5 to (4.25), it is easy to see that

$$\widehat{K} = \left[\begin{array}{c|c} A_Q - B_Q B C_Q & B_Q \\ \hline C_Q & 0 \end{array}\right] \Phi_A^{-1}. \tag{7.14}$$

If it is possible to write

$$\left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & 0 \end{array}\right] = \left[\begin{array}{cc|c} A + BD_K & BC_K & I \\ B_K & A_K & 0 \\ \hline D_K & C_K & 0 \end{array}\right] \tag{7.15}$$

for some A_K, B_K, C_K, D_K, then by applying (7.14) to (7.15) above, it follows that

$$\begin{aligned}
\widehat{K} &= \left[\begin{array}{c|c} \begin{bmatrix} A+BD_K & BC_K \\ B_K & A_K \end{bmatrix} - \begin{bmatrix} B \\ 0 \end{bmatrix} \begin{bmatrix} D_K & C_K \end{bmatrix} & \begin{bmatrix} I \\ 0 \end{bmatrix} \\ \hline \begin{bmatrix} D_K & C_K \end{bmatrix} & 0 \end{array}\right] \Phi_A^{-1} \\
&= \left[\begin{array}{cc|c} A & 0 & I \\ B_K & A_K & 0 \\ \hline D_K & C_K & 0 \end{array}\right] \Phi_A^{-1} \\
&= \begin{bmatrix} D_K & C_K \end{bmatrix} \begin{bmatrix} \Phi_A & 0 \\ \Phi_{A_K} B_K \Phi_A & \Phi_{A_K} \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix} \Phi_A^{-1} \\
&= D_K + C_K \Phi_{A_K} B_K \\
&= \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right].
\end{aligned}$$

In particular, $\widehat{K}$ is proper, so $\widehat{Q}$ in (7.12) satisfies the third condition on page 124. Now (7.15) holds if we can find a state space similarity Λ and matrices A_K, B_K, C_K, D_K such that

$$\Lambda^{-1} A_Q \Lambda = \begin{bmatrix} A+BD_K & BC_K \\ B_K & A_K \end{bmatrix}, \qquad \Lambda^{-1} B_Q = \begin{bmatrix} I_n \\ 0_{(N-n)\times n} \end{bmatrix} \quad \text{and} \quad C_Q \Lambda = \begin{bmatrix} D_K & C_K \end{bmatrix}. \tag{7.16}$$

Hence we need to construct a state space similarity Λ such that (7.16) holds for some A_K, B_K, C_K and D_K. Define as on page 101 of [38], the embeddings

$$\begin{aligned}
\Pi_1 &= \operatorname{diag}_{j\in P}(I_{\underline{n}}(\downarrow j, j)) = I_{\underline{n}_\downarrow}(P_\downarrow, P_1) \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}} \\
\Pi_2 &= \operatorname{diag}_{j\in P}(I_{\underline{n}}(\downarrow j, \twoheaddownarrow j)) = I_{\underline{n}_\downarrow}(P_\downarrow, P_{\twoheaddownarrow}) \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}_{\twoheaddownarrow}},
\end{aligned} \tag{7.17}$$

where the sets $P_\downarrow$, P_1 and $P_{\twoheaddownarrow}$ and the sub-partitions $\underline{n}_\downarrow$ and $\underline{n}_{\twoheaddownarrow}$ are defined on page 126.

Lemma 7.1.9.
The matrix $\begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix}$ *is orthogonal.*

Proof.
Since $\downarrow P_1 = P_\downarrow$ and $\downarrow P_{\twoheaddownarrow} \subset \downarrow P_\downarrow = P_\downarrow$, it follows from Theorem 2.3.10 and equation (7.11) that

$$\begin{aligned}
\Pi_1^\intercal \Pi_1 &= I_{\underline{n}_\downarrow}(P_\downarrow, P_1)^\intercal I_{\underline{n}_\downarrow}(P_\downarrow, P_1) = I_{\underline{n}_\downarrow}(P_1, P_\downarrow) I_{\underline{n}_\downarrow}(P_\downarrow, P_1) = I_{\underline{n}_\downarrow}(P_1, P_1) = I_{\underline{n}}, \\
\Pi_2^\intercal \Pi_2 &= I_{\underline{n}_\downarrow}(P_\downarrow, P_{\twoheaddownarrow})^\intercal I_{\underline{n}_\downarrow}(P_\downarrow, P_{\twoheaddownarrow}) = I_{\underline{n}_\downarrow}(P_{\twoheaddownarrow}, P_\downarrow) I_{\underline{n}_\downarrow}(P_\downarrow, P_{\twoheaddownarrow}) = I_{\underline{n}_\downarrow}(P_{\twoheaddownarrow}, P_{\twoheaddownarrow}) = I_{\underline{n}_{\twoheaddownarrow}} \\
\Pi_1^\intercal \Pi_2 &= I_{\underline{n}_\downarrow}(P_1, P_\downarrow) I_{\underline{n}_\downarrow}(P_\downarrow, P_{\twoheaddownarrow}) = I_{\underline{n}_\downarrow}(P_1, P_{\twoheaddownarrow}) = 0_{n\times(N-n)} \\
\Pi_2^\intercal \Pi_1 &= I_{\underline{n}_\downarrow}(P_{\twoheaddownarrow}, P_\downarrow) I_{\underline{n}_\downarrow}(P_\downarrow, P_1) = I_{\underline{n}_\downarrow}(P_{\twoheaddownarrow}, P_1) = 0_{(N-n)\times n} \\
\Pi_1 \Pi_1^\intercal + \Pi_2 \Pi_2^\intercal &= I_{\underline{n}_\downarrow}(P_\downarrow, P_1) I_{\underline{n}_\downarrow}(P_1, P_\downarrow) + I_{\underline{n}_\downarrow}(P_\downarrow, P_{\twoheaddownarrow}) I_{\underline{n}_\downarrow}(P_{\twoheaddownarrow}, P_\downarrow) = I_{\underline{n}_\downarrow}.
\end{aligned} \tag{7.18}$$

So

$$\begin{aligned}
\begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix} \begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix}^\intercal &= \Pi_1 \Pi_1^\intercal + \Pi_2 \Pi_2^\intercal = I_{\underline{n}_\downarrow} \qquad\qquad \text{and} \\
\begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix}^\intercal \begin{bmatrix} \Pi_1 & \Pi_2 \end{bmatrix} &= \begin{bmatrix} \Pi_1^\intercal \Pi_1 & \Pi_1^\intercal \Pi_2 \\ \Pi_2^\intercal \Pi_1 & \Pi_2^\intercal \Pi_2 \end{bmatrix} = \begin{bmatrix} I_{\underline{n}} & 0 \\ 0 & I_{\underline{n}_{\twoheaddownarrow}} \end{bmatrix}
\end{aligned}$$

and the result follows. ☐

Let A_Q, B_Q, C_Q be as in the solution (7.12), that is

$$
\begin{aligned}
A_Q &= \begin{bmatrix} A(\downarrow 1, \downarrow 1) + B(\downarrow 1, \downarrow 1)F_1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & A(\downarrow p, \downarrow p) + B(\downarrow p, \downarrow p)F_p \end{bmatrix} \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}_\downarrow}, \\
B_Q &= \begin{bmatrix} I_{\underline{n}}(\downarrow 1, 1) & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & I_{\underline{n}}(\downarrow p, p) \end{bmatrix} \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}} \qquad \text{and} \\
C_Q &= \begin{bmatrix} I_{\underline{m}}(:, \downarrow 1)F_1 & \dots & I_{\underline{m}}(:, \downarrow p)F_p \end{bmatrix} \in \mathbb{R}^{\underline{m} \times \underline{n}_\downarrow}.
\end{aligned}
\tag{7.19}
$$

With $\Pi_1 \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}}$ and $\Pi_2 \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}_{\not\downarrow}}$ as in (7.17), define the matrices

$$
\begin{aligned}
R &= \begin{bmatrix} I_{\underline{n}}(:, \downarrow 1) & \dots & I_{\underline{n}}(:, \downarrow p) \end{bmatrix} \in \mathbb{R}^{\underline{n} \times \underline{n}_\downarrow} \\
A_\Pi &= \Pi_2^\intercal A_Q \Pi_2 \in \mathbb{R}^{\underline{n}_{\not\downarrow} \times \underline{n}_{\not\downarrow}} \\
B_\Pi &= \Pi_2^\intercal A_Q \Pi_1 \in \mathbb{R}^{\underline{n}_{\not\downarrow} \times \underline{n}} \\
C_\Pi &= R\Pi_2 \in \mathbb{R}^{\underline{n} \times \underline{n}_{\not\downarrow}}.
\end{aligned}
\tag{7.20}
$$

The following lemma is a technical result that is required in subsequent computations.

Lemma 7.1.10.
With $\Pi_1 \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}}$ and $\Pi_2 \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}_{\not\downarrow}}$ as in (7.17) and $A_Q \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}_\downarrow}$, $B_Q \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}}$ and $C_Q \in \mathbb{R}^{\underline{m} \times \underline{n}_\downarrow}$ as in (7.19), the following identities hold:

$$
\begin{aligned}
(\Pi_1^\intercal + C_\Pi \Pi_2^\intercal) A_Q \Pi_1 &= A + BC_Q\Pi_1 \\
(\Pi_1^\intercal + C_\Pi \Pi_2^\intercal) A_Q (\Pi_2 - \Pi_1 C_\Pi) &= BC_Q(\Pi_2 - \Pi_1 C_\Pi).
\end{aligned}
\tag{7.21}
$$

Proof.
We first compute the factor $(\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)$, because it appears in both identities. Since $\downarrow({\not\downarrow} j) = {\not\downarrow} j \subset \downarrow j$ for each j, it follows by Theorem 2.3.10 that

$$
C_\Pi = R\Pi_2 = \begin{bmatrix} I_{\underline{n}}(:, \downarrow 1) & \dots & I_{\underline{n}}(:, \downarrow p) \end{bmatrix} \begin{bmatrix} I_{\underline{n}}(\downarrow 1, {\not\downarrow} 1) & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & I_{\underline{n}}(\downarrow p, {\not\downarrow} p) \end{bmatrix} = \begin{bmatrix} I_{\underline{n}}(:, {\not\downarrow} 1) & \dots & I_{\underline{n}}(:, {\not\downarrow} p) \end{bmatrix}.
$$

Thus

$$
\begin{aligned}
C_\Pi \Pi_2^\intercal &= \begin{bmatrix} I_{\underline{n}}(:, {\not\downarrow} 1) & \dots & I_{\underline{n}}(:, {\not\downarrow} p) \end{bmatrix} \begin{bmatrix} I_{\underline{n}}(\downarrow 1, {\not\downarrow} 1) & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & I_{\underline{n}}(\downarrow p, {\not\downarrow} p) \end{bmatrix} \\
&= \begin{bmatrix} I_{\underline{n}}(:, {\not\downarrow} 1) I_{\underline{n}}({\not\downarrow} 1, \downarrow 1) & \dots & I_{\underline{n}}(:, {\not\downarrow} p) I_{\underline{n}}({\not\downarrow} p, \downarrow p) \end{bmatrix}.
\end{aligned}
$$

Since $\Pi_1^\intercal = \operatorname{diag}(I_{\underline{n}}(j, \downarrow j)) = \begin{bmatrix} I_{\underline{n}}(:,1)I_{\underline{n}}(1,\downarrow 1) & \dots & I_{\underline{n}}(:,p)I_{\underline{n}}(p,\downarrow p) \end{bmatrix}$, we get that

$$\begin{aligned} \Pi_1^\intercal + C_\Pi \Pi_2^\intercal &= \begin{bmatrix} I_{\underline{n}}(:,1)I_{\underline{n}}(1,\downarrow 1) & \dots & I_{\underline{n}}(:,p)I_{\underline{n}}(p,\downarrow p) \end{bmatrix} + \begin{bmatrix} I_{\underline{n}}(:,\not\downarrow 1)I_{\underline{n}}(\not\downarrow 1,\downarrow 1) & \dots & I_{\underline{n}}(:,\not\downarrow p)I_{\underline{n}}(\not\downarrow p,\downarrow p) \end{bmatrix} \\ &= \begin{bmatrix} I_{\underline{n}}(:,\downarrow 1) & \dots & I_{\underline{n}}(:,\downarrow p) \end{bmatrix}. \end{aligned}$$

Thus

$$\begin{aligned} (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q &= \begin{bmatrix} I_{\underline{n}}(:,\downarrow 1) & \dots & I_{\underline{n}}(:,\downarrow p) \end{bmatrix} \\ &\quad \times \begin{bmatrix} A(\downarrow 1,\downarrow 1) + B(\downarrow 1,\downarrow 1)F_1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & A(\downarrow p,\downarrow p) + B(\downarrow p,\downarrow p)F_p \end{bmatrix} \\ &= \begin{bmatrix} A(:,\downarrow 1) & \dots & A(:,\downarrow p) \end{bmatrix} + \begin{bmatrix} B(:,\downarrow 1)F_1 & \dots & B(:,\downarrow p)F_p \end{bmatrix} \\ &= A\begin{bmatrix} I_{\underline{n}}(:,\downarrow 1) & \dots & I_{\underline{n}}(:,\downarrow p) \end{bmatrix} + B\begin{bmatrix} I_{\underline{n}}(:,\downarrow 1)F_1 & \dots & I_{\underline{n}}(:,\downarrow p)F_p \end{bmatrix} \\ &= A\begin{bmatrix} I_{\underline{n}}(:,\downarrow 1) & \dots & I_{\underline{n}}(:,\downarrow p) \end{bmatrix} + BC_Q. \end{aligned}$$

And so, for the first equality, we compute

$$\begin{aligned} (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q\Pi_1 &= A\begin{bmatrix} I_{\underline{n}}(:,\downarrow 1) & \dots & I_{\underline{n}}(:,\downarrow p) \end{bmatrix} \begin{bmatrix} I_{\underline{n}}(\downarrow 1,1) & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & I_{\underline{n}}(\downarrow p,p) \end{bmatrix} + BC_Q\Pi_1 \\ &= A\begin{bmatrix} I_{\underline{n}}(:,1) & \dots & I_{\underline{n}}(:,p) \end{bmatrix} + BC_Q\Pi_1 \\ &= A + BC_Q\Pi_1. \end{aligned}$$

We now prove the second equality. By the first equality, it follows that

$$\begin{aligned} (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q(\Pi_2 - \Pi_1 C_\Pi) &= (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q\Pi_2 - [(\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q\Pi_1]C_\Pi \\ &= (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q\Pi_2 - [A + BC_Q\Pi_1]C_\Pi \\ &= (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q\Pi_2 - AC_\Pi + BC_Q(-\Pi_1 C_\Pi). \end{aligned}$$

We already computed the values of C_Π and $(\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q$ in the first part of the proof, thus it follows that

$$\begin{aligned} (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q\Pi_2 &= A\begin{bmatrix} I_{\underline{n}}(:,\downarrow 1) & \dots & I_{\underline{n}}(:,\downarrow p) \end{bmatrix} \begin{bmatrix} I_{\underline{n}}(\downarrow 1,\not\downarrow 1) & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \dots & 0 & I_{\underline{n}}(\downarrow p,\not\downarrow p) \end{bmatrix} + BC_Q\Pi_2 \\ &= A\begin{bmatrix} I_{\underline{n}}(:,\not\downarrow 1) & \dots & I_{\underline{n}}(:,\not\downarrow p) \end{bmatrix} + BC_Q\Pi_2 \\ &= AC_\Pi + BC_Q\Pi_2. \end{aligned}$$

Thus

$$\begin{aligned} (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal)A_Q(\Pi_2 - \Pi_1 C_\Pi) &= AC_\Pi + BC_Q\Pi_2 - AC_\Pi + BC_Q(-\Pi_1 C_\Pi) \\ &= BC_Q(\Pi_2 - \Pi_1 C_\Pi), \end{aligned}$$

which completes the proof of the second identity. □

Shah then defines the state space similarity

$$\Lambda = \begin{bmatrix}\Pi_1 & \Pi_2 - \Pi_1 C_\Pi\end{bmatrix} = \begin{bmatrix}\Pi_1 & \Pi_2\end{bmatrix}\begin{bmatrix} I_{\underline{n}} & -C_\Pi \\ 0 & I_{\underline{n}_\downarrow}\end{bmatrix} \in \mathbb{R}^{\underline{n}_\downarrow \times \underline{n}_\downarrow}. \tag{7.22}$$

Since $\begin{bmatrix}\Pi_1 & \Pi_2\end{bmatrix}$ is orthogonal, it follows by equation (2.3) that

$$\Lambda^{-1} = \begin{bmatrix} I_{\underline{n}} & -C_\Pi \\ 0 & I_{\underline{n}_\downarrow}\end{bmatrix}^{-1}\begin{bmatrix}\Pi_1 & \Pi_2\end{bmatrix}^{-1} = \begin{bmatrix} I_{\underline{n}} & C_\Pi \\ 0 & I_{\underline{n}_\downarrow}\end{bmatrix}\begin{bmatrix}\Pi_1^\intercal \\ \Pi_2^\intercal\end{bmatrix} = \begin{bmatrix}\Pi_1^\intercal + C_\Pi \Pi_2^\intercal \\ \Pi_2^\intercal\end{bmatrix}.$$

By definition, $A_\Pi = \Pi_2^\intercal A_Q \Pi_2$ and $B_\Pi = \Pi_2^\intercal A_Q \Pi_1$. Hence

$$\begin{aligned}\Lambda^{-1} A_Q \Lambda &= \begin{bmatrix}\Pi_1^\intercal + C_\Pi \Pi_2^\intercal \\ \Pi_2^\intercal\end{bmatrix} A_Q \begin{bmatrix}\Pi_1 & \Pi_2 - \Pi_1 C_\Pi\end{bmatrix} \\ &= \begin{bmatrix}(\Pi_1^\intercal + C_\Pi \Pi_2^\intercal) A_Q \Pi_1 & (\Pi_1^\intercal + C_\Pi \Pi_2^\intercal) A_Q (\Pi_2 - \Pi_1 C_\Pi) \\ B_\Pi & A_\Pi - B_\Pi C_\Pi\end{bmatrix} \\ &= \begin{bmatrix} A + B C_Q \Pi_1 & B C_Q (\Pi_2 - \Pi_1 C_\Pi) \\ B_\Pi & A_\Pi - B_\Pi C_\Pi\end{bmatrix},\end{aligned}$$

where the last equality follows from (7.21). Secondly, since $B_Q = \Pi_1$, it follows from (7.18) that

$$\Lambda^{-1} B_Q = \begin{bmatrix}\Pi_1^\intercal + C_\Pi \Pi_2^\intercal \\ \Pi_2^\intercal\end{bmatrix}\Pi_1 = \begin{bmatrix}\Pi_1^\intercal \Pi_1 + C_\Pi \Pi_2^\intercal \Pi_1 \\ \Pi_2^\intercal \Pi_1\end{bmatrix} = \begin{bmatrix} I_{\underline{n}} + 0 \\ 0\end{bmatrix}.$$

Finally

$$C_Q \Lambda = \begin{bmatrix} C_Q \Pi_1 & C_Q(\Pi_2 - \Pi_1 C_\Pi)\end{bmatrix}.$$

The above equalities show that the realizations

$$\widetilde{Q} = \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & 0\end{array}\right] \qquad \text{and} \qquad \widehat{Q} = \left[\begin{array}{cc|c} A + B C_Q \Pi_1 & B C_Q(\Pi_2 - \Pi_1 C_\Pi) & I_{\underline{n}} \\ B_\Pi & A_\Pi - B_\Pi C_\Pi & 0 \\ \hline C_Q \Pi_1 & C_Q(\Pi_2 - \Pi_1 C_\Pi) & 0\end{array}\right]$$

are state space similar with state space similarity given by Λ in (7.22). Since the optimal solution $\widetilde{Q}$ of the reparameterized problem is now written in the form (7.15), it is possible to obtain a solution to the original problem by applying (7.14) to $\widehat{Q}$ above. Thus the optimal controller to the original is given by

$$\widetilde{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K\end{array}\right] = \left[\begin{array}{c|c} A_\Pi - B_\Pi C_\Pi & B_\Pi \\ \hline C_Q(\Pi_2 - \Pi_1 C_\Pi) & C_Q \Pi_1\end{array}\right]. \tag{7.23}$$

Example 7.1.11 (Construction of Π_1, Π_2 and R for Shah's example in Section 4.5.4)**.**
Let $\mathcal{P} = (P, \succeq)$ be the poset with $P = \{1, 2, 3, 4\}$ and $1 \succeq 2$, $1 \succeq 3$, $2 \succeq 4$ and $3 \succeq 4$ as in Example 7.1.7. Let $\underline{n} = (1, 1, 1, 1)$, $\underline{m} = (1, 1, 1, 1)$ and $\underline{r} = (2, 2, 2, 2)$ be partitions of the state space, input space and output space dimensions of the realization the poset-causal system in Example 7.1.8 respectively. In Example 7.1.8, the optimal solution $\widehat{Q}$ to the reparameterized state feedback $\mathcal{H}_2$-control problem was constructed. In this example, the optimal solution $\widetilde{K}$ of the original problem is constructed from $\widetilde{Q}$.

We first construct the matrices Π_1, Π_2 and R. By equation (7.17) we have

$$
\Pi_1 = \operatorname{diag}(I_{\underline{n}}(\downarrow j, j)) = \left[\begin{array}{c|c|c|c}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\ \hline
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\ \hline
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 1
\end{array}\right] \in \mathbb{R}^{\underline{n}_{\downarrow} \times \underline{n}}
$$

$$
\Pi_2 = \operatorname{diag}(I_{\underline{n}}(\downarrow j, \twoheaddownarrow j)) = \left[\begin{array}{ccc|c|c}
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\ \hline
0 & 0 & 0 & 0 & 0
\end{array}\right] \in \mathbb{R}^{\underline{n}_{\downarrow} \times \underline{n}_{\twoheaddownarrow}}.
$$

Lastly, by equation (7.20),

$$
R = \begin{bmatrix} I_{\underline{n}}(:, \downarrow 1) & \dots & I_{\underline{n}}(:, \downarrow p) \end{bmatrix} = \left[\begin{array}{cccc|cc|cc|c}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1
\end{array}\right] \in \mathbb{R}^{\underline{n} \times \underline{n}_{\downarrow}}.
$$

With A_Q, B_Q and C_Q as in Example 7.1.8 and using Π_1, Π_2 and R, we can now construct the matrices A_Π, B_Π and C_Π using equation (7.20):

$$
\begin{aligned}
A_\Pi &= \Pi_2^\intercal A_Q \Pi_2 \\
&= \begin{bmatrix}
-1.5589 & -0.5443 & -0.0282 & 0 & 0 \\
-0.5560 & -1.5927 & -0.0064 & 0 & 0 \\
-1.7232 & -1.7634 & -1.1032 & 0 & 0 \\
0 & 0 & 0 & -1.0991 & 0 \\
0 & 0 & 0 & 0 & -1.0811
\end{bmatrix} \\
B_\Pi &= \Pi_2^\intercal A_Q \Pi_1 \\
&= \begin{bmatrix}
-0.7505 & 0 & 0 & 0 \\
-0.6869 & 0 & 0 & 0 \\
-0.3535 & 0 & 0 & 0 \\
0 & -1.2226 & 0 & 0 \\
0 & 0 & -1.2733 & 0
\end{bmatrix} \\
C_\Pi &= R \Pi_2 \\
&= \begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1
\end{bmatrix}
\end{aligned}
$$

By equation (7.23), the optimal controller $\widetilde{K}$ is given by

$$\widetilde{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$$
$$= \left[\begin{array}{ccccc|cccc} -1.5589 & -0.5443 & -0.0282 & 0 & 0 & -0.7505 & 0 & 0 & 0 \\ -0.5560 & -1.5927 & -0.0064 & 0 & 0 & -0.6869 & 0 & 0 & 0 \\ -1.7232 & -1.7634 & -1.1032 & 0 & 0 & -0.3535 & 0 & 0 & 0 \\ 1.2226 & 0 & 0 & -1.0991 & 0 & 0 & -1.2226 & 0 & 0 \\ 0 & 1.2733 & 0 & 0 & -1.0811 & 0 & 0 & -1.2733 & 0 \\ \hline -0.3515 & -0.3616 & 0.0751 & 0 & 0 & -0.7175 & 0 & 0 & 0 \\ 0.0662 & -0.1827 & -0.1033 & -0.0990 & 0 & 0.9671 & -1.0237 & 0 & 0 \\ -0.2045 & 0.0648 & -0.0814 & 0 & -0.0792 & 1.0306 & 0 & -1.0960 & 0 \\ -0.0109 & -0.0106 & 0.0115 & 0.0049 & 0.0031 & -0.6337 & 0.8011 & 0.8226 & -0.9050 \end{array}\right]$$

Note that $A_K \in \mathcal{I}_{\mathcal{P}}^{\widetilde{n}\times\widetilde{n}}$, $B_K \in \mathcal{I}_{\mathcal{P}}^{\widetilde{n}\times\widetilde{r}}$, $A_K \in \mathcal{I}_{\mathcal{P}}^{\widetilde{m}\times\widetilde{n}}$ and $D_K \in \mathcal{I}_{\mathcal{P}}^{\widetilde{m}\times\widetilde{r}}$ with $\widetilde{n} = (3,1,1,0)$, $\widetilde{m} = (1,1,1,1)$ and $\widetilde{r} = (1,1,1,1)$. With these partitions, it can be seen that the system matrices of $\widetilde{K}$ have the poset-causal structure

$$\begin{bmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ * & 0 & * & 0 \\ * & * & * & * \end{bmatrix}.$$

7.2 Output Feedback Control

In this section, we consider the significantly more difficult problem of structured output feedback control. We do not obtain an optimal solution. We will, however, give feasible structured solution strategies that are based on certain optimality considerations.

7.2.1 Definition of the Poset-Causal $\mathcal{H}_2$-Control Problem

In Section 4.3, we reviewed the classical unstructured $\mathcal{H}_2$-control problem for the output feedback case. We saw that it is reasonable to assume that $D_{11} = 0$, $D_{22} = 0$ and $D_K = 0$. For a poset-causal system whose structure is determined by a poset $\mathcal{P}$, we have the additional requirement to the control problem that the controller must also have the structure determined by $\mathcal{P}$, so that the structure of the system is preserved under feedback.

Problem 7.2.1 (Poset-causal $\mathcal{H}_2$-control problem)**.**
Let $\mathcal{P} = (P, \succeq)$ with $P = \{1, \ldots, p\}$ be a poset. Let $\underline{n}, \underline{\ell}, \underline{m}, \underline{r}, \underline{s} \in \mathbb{Z}_+^p$ be partitions. For a $\mathcal{P}$-poset-causal plant

$$\widehat{G} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] = \begin{bmatrix} C_1\Phi_A B_1 & C_1\Phi_A B_2 + D_{12} \\ C_2\Phi_A B_1 + D_{21} & C_2\Phi_A B_2 \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11} & \widehat{G}_{12} \\ \widehat{G}_{21} & \widehat{G}_{22} \end{bmatrix}, \tag{7.24}$$

with

$$A \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{n}}, \quad B_1 \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{\ell}}, \quad B_2 \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}, \quad C_1 \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{n}}, \quad C_2 \in \mathcal{I}_{\mathcal{P}}^{\underline{s}\times\underline{n}}, \quad D_{12} \in \mathcal{I}_{\mathcal{P}}^{\underline{r}\times\underline{m}}, \quad D_{21} \in \mathcal{I}_{\mathcal{P}}^{\underline{s}\times\underline{\ell}},$$

construct a controller

$$\widehat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & 0 \end{array}\right] \qquad \text{satisfying}$$

1. $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}}$;
2. $\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) \in \mathcal{RH}_2$;

such that $\|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2^2$ is minimized.

Note here, that by applying a canonical shuffle as in Corollary 2.3.19, we can get

$$\begin{bmatrix} B_1 & B_2 \end{bmatrix} \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times(\underline{\ell}+\underline{m})}, \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \in \mathcal{I}_{\mathcal{P}}^{(\underline{r}+\underline{s})\times\underline{n}}, \quad \text{and} \quad \begin{bmatrix} 0 & D_{12} \\ D_{21} & 0 \end{bmatrix} \in \mathcal{I}_{\mathcal{P}}^{(\underline{r}+\underline{s})\times(\underline{\ell}+\underline{m})}.$$

Thus $\widehat{G}$ is indeed a poset-causal system as in Definition 5.2.1. By Proposition 4.1.2, we have

$$\underline{\mathcal{F}}(\widehat{G}, \widehat{K}) = \widehat{G}_{11} + \widehat{G}_{12}\widehat{K}(I - \widehat{G}_{22}\widehat{K})^{-1}\widehat{G}_{21} = \left[\begin{array}{cc|c} A & B_2 C_K & B_1 \\ B_K C_2 & A_K & B_K D_{21} \\ \hline C_1 & D_{12} C_K & 0 \end{array}\right].$$

In the unstructured case there exists a stabilizing controller if and only if (A, B_2, C_2) is stabilizable and detectable (Proposition 4.2.1). For the poset-causal system $\Sigma_{\mathcal{P}}$, the system matrices can always be written as block lower triangular matrices. Write $A_{ii} = A(i,i)$ for the i^{th} diagonal entry of A. If E and M are block diagonal matrices, then $A + B_2 E$ and $A + MC_2$ have the poset-causal structure and by Lemma 6.2.12,

$$\sigma(A + B_2 E) = \bigcup_{i\in P} \sigma(A_{ii} + B_2(i,i)E_i) \quad \text{and} \quad \sigma(A + MC_2) = \bigcup_{i\in P} \sigma(A_{ii} + M_i C_2(i,i)),$$

where E_i and M_i are the diagonal blocks of E and M respectively. Analogously to Proposition 4.2.1, we have the following result.

Proposition 7.2.2 (cf. Theorem 4.1 in [38] and Lemma 30 in [45])**.**
Suppose the local pairs $(A_{ii}, B_2(i,i))$ and $(C_2(i,i), A_{ii})$ are stabilizable and detectable for $i = 1, \ldots, p$ respectively. Then the poset-causal plant $\widehat{G}$ in (7.24) has a stabilizing controller $\widehat{K} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{s}}$.

7.2.2 Youla Parametrization and Reformulation

We will henceforth assume that each local triple $(A(i,i), B_2(i,i)), C_2(i,i))$ is stabilizable and detectable for $i \in P$. By Proposition 7.2.2, there exists matrices

$$E \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{n}} \quad \text{and} \quad M \in \mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{s}} \quad \text{such that} \quad A + B_2E \text{ and } A + MC_2 \text{ are stable.} \tag{7.25}$$

Furthermore, $A + B_2E$ and $A + MC_2$ still have the poset-causal structure. We now employ the Youla parametrization (Theorem 4.2.3). Set

$$\widehat{J} := \left[\begin{array}{c|cc} A + B_2E + MC_2 & -M & B_2 \\ \hline E & 0 & I \\ -C_2 & I & 0 \end{array}\right]. \tag{7.26}$$

Let

$$\widehat{H} := \widehat{G} \star \widehat{J} = \begin{bmatrix} \widehat{H}_{11} & \widehat{H}_{12} \\ \widehat{H}_{21} & \widehat{H}_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} A + B_2E & -B_2E & B_1 & B_2 \\ 0 & A + MC_2 & B_1 + MD_{21} & 0 \\ \hline C_1 + D_{12}E & -D_{12}E & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right] =: \left[\begin{array}{c|cc} \widetilde{A} & \widetilde{B}_1 & \widetilde{B}_2 \\ \hline \widetilde{C}_1 & 0 & D_{12} \\ \widetilde{C}_2 & D_{21} & 0 \end{array}\right].$$

Then, as before, applying the appropriate canonical shuffle, the matrices $\widetilde{A}$, $\widetilde{B}_1$, $\widetilde{B}_2$, $\widetilde{C}_1$ and $\widetilde{C}_2$ all have the poset-causal structure and so do the following transfer functions

$$\begin{aligned}
\widehat{H}_{11} &= \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_1 \\ \hline \widetilde{C}_1 & 0 \end{array}\right] = \left[\begin{array}{cc|c} A + B_2E & -B_2E & B_1 \\ 0 & A + MC_2 & B_1 + MD_{21} \\ \hline C_1 + D_{12}E & -D_{12}E & 0 \end{array}\right], \\
\widehat{H}_{12} &= \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_2 \\ \hline \widetilde{C}_1 & D_{12} \end{array}\right] = \left[\begin{array}{cc|c} A + B_2E & -B_2E & B_2 \\ 0 & A + MC_2 & 0 \\ \hline C_1 + D_{12}E & -D_{12}E & D_{12} \end{array}\right] = \left[\begin{array}{c|c} A + B_2E & B_2 \\ \hline C_1 + D_{12}E & D_{12} \end{array}\right], \\
\widehat{H}_{21} &= \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B}_1 \\ \hline \widetilde{C}_2 & D_{21} \end{array}\right] = \left[\begin{array}{cc|c} A + B_2E & -B_2E & B_1 \\ 0 & A + MC_2 & B_1 + MD_{21} \\ \hline 0 & C_2 & D_{21} \end{array}\right] = \left[\begin{array}{c|c} A + MC_2 & B_1 + MD_{21} \\ \hline C_2 & D_{21} \end{array}\right] \quad \text{and} \\
\widehat{H}_{22} &= 0.
\end{aligned} \tag{7.27}$$

All stabilizing controllers of $\widehat{G}$ are given by

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}, \widehat{R}) \qquad \text{with } \widehat{R} \in \mathcal{RH}_2^{\underline{m}\times\underline{s}} \text{ arbitrary}$$

and the closed loop transfer functions are given by

$$\underline{\mathcal{F}}\left(\widehat{G}, \widehat{K}\right) = \underline{\mathcal{F}}\left(\widehat{G}, \underline{\mathcal{F}}(\widehat{J}, \widehat{R})\right) = \underline{\mathcal{F}}\left(\widehat{G} \star \widehat{J}, \widehat{R}\right) = \underline{\mathcal{F}}\left(\widehat{H}, \widehat{R}\right) = \widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}.$$

Since $\widehat{H}_{11}$, $\widehat{H}_{12}$ and $\widehat{H}_{21}$ all have the structure determined by the poset $\mathcal{P}$, the closed loop transfer function will also have it if $\widehat{R} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{s}}$. Since $\mathcal{P}$ is a poset, $\mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{s}}$ is quadratically invariant (see Subsection 5.1.1) under the poset-causal plant $\widehat{G}$. Thus the Youla parametrization gives the following equivalent problem.

Proposition 7.2.3.
Let $\widehat{J}$ be as in (7.26). Then $\widehat{K} = \underline{\mathcal{F}}\left(\widehat{J}, \widehat{R}\right)$ is an optimal solution to the $\mathcal{H}_2$-control Problem 7.2.1 if and only if $\widehat{R}$ is an optimal solution to the $\mathcal{H}_2$-control problem

$$\begin{aligned} &\textit{minimize} && \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2 \\ &\textit{subject to} && \widehat{R} \in \mathcal{RH}_2, \quad \widehat{R} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{s}}. \end{aligned} \tag{7.28}$$

7.2.3 Recursive, Bottom-up Approach

In a bottom up approach, local optimal controllers are determined for subsystems at the end of the communication chain, that is, subsystems that correspond to the nodes in a digraph that have out-degree zero. A solution may then be recursively calculated for the following layers of subsystems in the communication chain until a controller for the overall poset-causal system is determined. Such an approach was followed in [22] for the LQ-optimization problem for coordinated linear systems. A similar approach was followed in [25] for the output feedback $\mathcal{H}_2$-control problem, but only for the 2-player problem. Such an approach may possibly be extended to other hierarchical systems such as those with underlying graph structure given in Example 5.3.1. However, this approach may not be appropriate to more general poset-causal systems such as one with an underlying graph structure given below

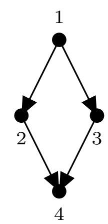


Figure 7.1: Typical poset-causal structure

due to subsystem 4 which has an in-degree greater than one. In the Hasse diagrams of hierarchical systems, there is at most one directed path between any two vertices in the graph. But in the Hasse diagram of a poset-causal system as the one in the above figure, there are two directed paths from 1 to 4.

7.2.4 Construction of Structured Functions

In this approach, we consider the optimal solution obtained in Section 4.6 to the unstructured problem and by analogy of that solution, we construct a (not necessarily optimal) solution in the structured case. With $\widehat{H}_{11}$ and $\widehat{H}_{11}^{\circ}$ as in (4.35) and (4.36) respectively, with $\widehat{U}$ and $\widehat{V}$ inner functions as in (4.38) and with $\widehat{L}$ and $\widehat{M}$ invertible outer functions as in (4.39), it follows from Lemma 4.6.3 that

$$\widehat{H}_{11} = \widehat{H}_{11}^{\circ} + \widehat{U}\widehat{P}\widehat{V} = \widehat{H}_{11}^{\circ} + \widehat{U}\widehat{L}\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1}\widehat{M}\widehat{V}$$

where

$$\widehat{P} = R_1^{\frac{1}{2}} \left[\begin{array}{cc|c} A + B_2E & B_2E & -L \\ 0 & A + MC_2 & L - M \\ \hline E - F & E & 0 \end{array}\right] R_2^{\frac{1}{2}},$$

and where F and L are given by (4.37). By Corollary 4.6.1, $\widehat{H}_{12} = \widehat{U}\widehat{L}$ and $\widehat{H}_{21} = \widehat{M}\widehat{V}$. Thus

$$\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21} = \widehat{H}_{11}^{\circ} + \widehat{U}\widehat{L}\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1}\widehat{M}\widehat{V} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21} = \widehat{H}_{11}^{\circ} + \widehat{H}_{12}(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})\widehat{H}_{21}.$$

As in the proof of Theorem 4.6.4, it follows that

$$\begin{aligned}\|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2^2 &= \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2^2 \\ &= \|\widehat{H}_{11}^{\circ}\|_2^2 + \|\widehat{H}_{12}(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})\widehat{H}_{21}\|_2^2.\end{aligned}$$

Since $\widehat{U}$ and $\widehat{V}$ are inner and co-inner respectively, it follows that

$$\begin{aligned}\|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2^2 &= \|\widehat{H}_{11}^{\circ}\|_2^2 + \|\widehat{H}_{12}(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})\widehat{H}_{21}\|_2^2 \\ &= \|\widehat{H}_{11}^{\circ}\|_2^2 + \|\widehat{U}\widehat{L}(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})\widehat{M}\widehat{V}\|_2^2 \\ &= \|\widehat{H}_{11}^{\circ}\|_2^2 + \|\widehat{L}(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})\widehat{M}\|_2^2.\end{aligned}$$

Note that due to the solutions of Riccati equations appearing in F and L in $\widehat{P}$, $\widehat{L}$ and $\widehat{M}$, these functions do not have the poset-causal structure. In the unstructured case, the optimal $\widehat{R}$ is obtained by simply computing $\widehat{R} = -\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1}$, but in the structured case, this $\widehat{R}$ will not in general have the poset-causal structure. A possibility is to build a structured $\widehat{R} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{s}}$ in such a way that its non-zero blocks are equal to the corresponding blocks in $-\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1}$ and with zero blocks in the appropriate positions to ensure that it has the poset-causal structure. If it is the case that $\widehat{L} = I$ and $\widehat{M} = I$, then

$$\|\underline{\mathcal{F}}(\widehat{G}, \widehat{K})\|_2^2 = \|\widehat{H}_{11}^{\circ}\|_2^2 + \|\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R}\|_2^2.$$

In this case, we solve the structured optimization problem

$$\begin{array}{ll} \text{minimize} & \|\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R}\|_2^2 \\ \text{subject to} & \widehat{R} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{s}}. \end{array}$$

By the column-wise separability of the $\mathcal{H}_2$-norm,

$$\|\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R}\|_2^2 = \sum_{j=1}^{p} \|(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(:, j)\|_2^2$$

and by definition of the $\mathcal{H}_2$-norm (see equation (2.9)),

$$\begin{aligned}
\|(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(:, j)\|_2^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}\Big[\Big((\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(:, j)\Big)(i\omega)^* \Big((\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(:, j)\Big)(i\omega)\Big]\, \mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}\Big[\Big((\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})^*(j, :)\Big)\Big((\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(:, j)\Big)(i\omega)\Big]\, \mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}\Big[\sum_{i=1}^{p}\Big((\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})^*(j, i)\Big)\Big((\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(i, j)\Big)(i\omega)\Big]\, \mathrm{d}\omega \\
&= \sum_{i=1}^{p}\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}\Big[\Big((\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(i, j)\Big)(i\omega)^* \Big((\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(i, j)\Big)(i\omega)\Big]\, \mathrm{d}\omega \\
&= \sum_{i=1}^{p}\|(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})(i, j)\|_2^2
\end{aligned}$$

Thus

$$\|\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R}\|_2^2 = \sum_{j=1}^{p}\sum_{i=1}^{p} \|\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1}(i, j) + \widehat{R}(i, j)\|_2^2.$$

Since it is required that $\widehat{R} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{s}}$, we define

$$\widehat{R}(i, j) = \begin{cases} 0 & \text{if} \quad j \not\succeq i \\ -\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1}(i, j) & \text{if} \quad j \succeq i. \end{cases}$$

Then $\widehat{R}$ is a structured approximation of the optimal solution $-\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1}$. With $\widehat{R}$ as above, define

$$\widehat{K} = \underline{\mathcal{F}}\left(\widehat{J}, \widehat{R}\right).$$

Then $\widehat{K}$ is a feasible controller for the poset-causal $\mathcal{H}_2$-control Problem 7.2.1. However, this approach only works if $\widehat{L} = I$ and $\widehat{M} = I$, that is when $\widehat{H}_{12}$ and $\widehat{H}_{21}$ are inner and co-inner functions respectively (since then $\widehat{H}_{12} = \widehat{U}$ and $\widehat{H}_{21} = \widehat{V}$). Without this condition, even though $\widehat{L}$ and $\widehat{M}^{\intercal}$ are invertible outer functions, they significantly distort the problem since we have to minimize $\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R}$ in the weighted norm $\|\widehat{L}(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})\widehat{M}\|_2^2$.

7.2.5 Adjustment of the One-sided Optimal Solution

Consider the reformulated $\mathcal{H}_2$-control control problem in (7.28). By Proposition 7.2.3, if we can find an optimal solution $\widehat{R}_{opt}$, then this also gives an optimal solution $\widehat{K}_{opt}$ to the original poset-causal $\mathcal{H}_2$-control problem via $\widehat{K} = \underline{\mathcal{F}}\left(\widehat{J}, \widehat{R}\right)$. We follow an approach inspired by the solution to the state feedback case. We will call the related problem

$$\begin{aligned} \text{minimize} \quad & \|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\| \\ \text{subject to} \quad & \widehat{Q} \in \mathcal{RH}_2 \\ & \widehat{Q} \in \mathcal{I}_{\mathcal{P}}^{\underline{m}\times\underline{\ell}}, \end{aligned} \tag{7.29}$$

the *one-sided reparameterized problem*. We first solve this problem and then endeavour to find an $\widehat{R}$ from which we can recover a controller $\widehat{K}$. The solution of the one sided reparameterized problem follows the same approach as the solution of the poset-causal $\mathcal{H}_2$ problem for the state feedback case in Section 7.1.

Due to the column separability of the $\mathcal{H}_2$-norm, it follows from Theorem 2.3.10 that

$$\left\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\right\|_2^2 = \sum_{j=1}^{p} \left\|\left(\widehat{H}_{11} + \widehat{H}_{12}\widehat{Q}\right)(:, j)\right\|_2^2 = \sum_{j=1}^{p} \left\|\widehat{H}_{11}(:, j) + \widehat{H}_{12}(:, \downarrow j)\widehat{Q}(\downarrow j, j)\right\|_2^2.$$

Furthermore,

$$\widehat{Q} = \begin{bmatrix}\widehat{Q}(:, 1) & \dots & \widehat{Q}(:, p)\end{bmatrix} = \begin{bmatrix} I_{\underline{m}}(:, \downarrow 1)\widehat{Q}(\downarrow 1, 1) & \dots & I_{\underline{m}}(:, \downarrow p)\widehat{Q}(\downarrow p, p)\end{bmatrix}. \tag{7.30}$$

Importantly, $\widehat{Q}(\downarrow j, j)$ does not have a specific zero structure. This allows us to solve p local unstructured control problems using classical theory. An optimal solution $\widehat{Q}$ to the reparameterized problem (7.29) can then be reconstructed by applying (7.30).

Lemma 7.2.4 (cf. Theorem 4.2 in [38])**.**
Let the matrices E and M be as in (7.25) and $\widehat{H}_{11}$ and $\widehat{H}_{12}$ as in (7.27). Then $\widehat{Q}$ is the optimal solution to the $\mathcal{H}_2$-optimal control problem in (7.29) if and only if

$$\widehat{Q}(\downarrow j, j) = \widehat{Q}_j \qquad \textit{for } j \in P$$

where $\widehat{Q}_j$ is the optimal solution of

$$\begin{aligned} \textit{minimize} \quad & \left\|\widehat{H}_{11}(:, j) + \widehat{H}_{12}(:, \downarrow j)\widehat{Q}_j\right\|_2^2 \\ \textit{subject to} \quad & \widehat{Q}_j \in \mathcal{RH}_2. \end{aligned} \tag{7.31}$$

We now seek local solutions $\widehat{Q}_j$ to the control problems in (7.31). In order to do this, consider the local transfer functions $\widehat{H}_{11}(:, j)$ and $\widehat{H}_{12}(:, \downarrow j)$. By theorem 2.3.10, we get the following realizations

$$\widehat{H}_{11}(:, j) = \left[\begin{array}{c|c} \widetilde{A}(\downarrow j, \downarrow j) & \widetilde{B}_1(\downarrow j, j) \\ \hline \widetilde{C}_1(:, \downarrow j) & 0 \end{array}\right] \quad \text{and}$$
$$\widehat{H}_{12}(:, \downarrow j) = \left[\begin{array}{c|c} \widetilde{A}(\downarrow j, \downarrow j) & \widetilde{B}_2(\downarrow j, \downarrow j) \\ \hline \widetilde{C}_1(:, \downarrow j) & D_{12}(:, \downarrow j) \end{array}\right].$$

We now to apply Theorem 4.7.6 to each of the local problems (7.31).

Lemma 7.2.5.
Consider a poset-causal plant $\widehat{G}$ as in (7.24) satisfying the following conditions

1. $(A_{jj}, B_2(j,j), C_2(j,j))$ *is stabilizable and detectable for each* $j \in P$*;*
2. $R_1 = D_{12}^{\intercal} D_{12} > 0$*;*
3. $\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ *has full column rank for all* $\omega \in \mathbb{R}$*.*

Let the matrices E and M be as in (7.25) and $\widehat{H}_{11}$ and $\widehat{H}_{12}$ as in (7.27). For each $j \in P$, the optimal solution to the local control problem

$$\begin{aligned} \text{minimize} \quad & \left\| \widehat{H}_{11}(:,j) + \widehat{H}_{12}(:,\downarrow j)\widehat{Q}_j \right\|_2^2 \\ \text{subject to} \quad & \widehat{Q}_j \in \mathcal{RH}_2, \end{aligned}$$

is given by

$$\widehat{Q}_j = \left[\begin{array}{c|c} \widetilde{A}(\downarrow j, \downarrow j) + \widetilde{B}_2(\downarrow j, \downarrow j)\widetilde{F}_j & \widetilde{B}_1(\downarrow j, j) \\ \hline \widetilde{F}_j & 0 \end{array} \right],$$

where

$$\begin{aligned} \widetilde{F}_j &= -R_1(\downarrow j, \downarrow j)^{-1}(\widetilde{B}_2(\downarrow j, \downarrow j)^{\intercal}\widetilde{X}_j + D_{12}(:,\downarrow j)^{\intercal}\widetilde{C}_1(:,\downarrow j)) \quad \text{with} \\ \widetilde{X}_j &= Ric\,(\widetilde{A}(\downarrow j, \downarrow j), \widetilde{B}_2(\downarrow j, \downarrow j), \widetilde{C}_1(:,\downarrow j), D_{12}(:,\downarrow j)). \end{aligned}$$

Proof.
For $j \in P$, we show that $\widetilde{A}(\downarrow j, \downarrow j), \widetilde{B}_2(\downarrow j, \downarrow j), \widetilde{C}_1(:,\downarrow j), D_{12}(:,\downarrow j)$ satisfy the conditions of Theorem 4.7.6:

1. Since $A + B_2E$ and $A + MC_2$ are stable, $\widetilde{A}$ is stable. Thus, by Lemma 6.2.12, $\widetilde{A}(\downarrow j, \downarrow j)$ is stable.
2. By assumption, $D_{12}^{\intercal} D_{12} > 0$, that is, D_{12} has full column rank. But then $D_{12}(:,\downarrow j)$ still has full column rank and thus $D_{12}(:,\downarrow j)^{\intercal} D_{12}(:,\downarrow j) > 0$.
3. As in the proof of Lemma 4.7.7, since $\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all $\omega \in \mathbb{R}$,

$$\begin{bmatrix} \widetilde{A} - i\omega I & \widetilde{B}_2 \\ \widetilde{C}_1 & D_{12} \end{bmatrix} = \begin{bmatrix} (A + B_2E) - i\omega I & -B_2E & B_2 \\ 0 & (A + MC_2) - i\omega I & 0 \\ C_1 + D_{12}E & -D_{12}E & D_{12} \end{bmatrix}$$

has full column rank for each $\omega \in \mathbb{R}$. By Corollary 2.3.12, $\widetilde{A}$ and $\widetilde{B}_2$ have the following block partitioning

$$\widetilde{A} = \begin{bmatrix} \widetilde{A}(\downarrow j, \downarrow j) & * \\ 0 & * \end{bmatrix} \quad \text{and} \quad \widetilde{B}_2 = \begin{bmatrix} \widetilde{B}_2(\downarrow j, \downarrow j) & * \\ 0 & * \end{bmatrix}.$$

Thus

$$\begin{bmatrix} \widetilde{A} - i\omega I & \widetilde{B}_2 \\ \widetilde{C}_1 & D_{12} \end{bmatrix} = \begin{bmatrix} \widetilde{A}(\downarrow j, \downarrow j) - i\omega I & * & \widetilde{B}_2(\downarrow j, \downarrow j) & * \\ 0 & * & 0 & * \\ \widetilde{C}_1(:,\downarrow j) & * & D_{12}(:,\downarrow j) & * \end{bmatrix}.$$

But then the block compression

$$\begin{bmatrix} \widetilde{A}(\downarrow j, \downarrow j) - i\omega I & \widetilde{B}_2(\downarrow j, \downarrow j) \\ \widetilde{C}_1(:,\downarrow j) & D_{12}(:,\downarrow j) \end{bmatrix}$$

must have full column rank for all $\omega \in \mathbb{R}$.

The result now follows from Theorem 4.7.6. □

We now have local optimal solutions $\widehat{Q}_j$ to the control problems 7.31. By Lemma 7.2.4 and the column concatenation formula (3.16), we can construct a solution to the one-sided reparameterized problem from these local optimal solutions.

Theorem 7.2.6.
Given a $\mathcal{P}$-poset-causal plant $\widehat{G}$ as in (7.24) such that

1. $(A_{jj}, B_2(j,j), C_2(j,j))$ *is stabilizable and detectable for each* $j \in P$;
2. $R_1 = D_{12}^{\intercal} D_{12} > 0$;
3. $\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ *has full column rank for all* $\omega \in \mathbb{R}$.

With matrices E and M as in (7.25) and $\widehat{H}_{11}$ and $\widehat{H}_{12}$ as in (7.27), the optimal solution to the reparameterized problem

$$\begin{aligned} \text{minimize} \quad & \left\| \widehat{H}_{11} + \widehat{H}_{12}\widehat{Q} \right\|_2^2 \\ \text{subject to} \quad & \widehat{Q} \in \mathcal{RH}_2, \quad \widehat{Q} \in \mathcal{I}_{\mathcal{P}}^{m \times \underline{\ell}}. \end{aligned}$$

is given by

$$\widehat{Q}_{opt} = \left[\begin{array}{ccc|ccc} \widetilde{A}(\downarrow 1, \downarrow 1) + \widetilde{B}_2(\downarrow 1, \downarrow 1)\widetilde{F}_1 & \dots & 0 & \widetilde{B}_1(\downarrow 1, 1) & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \widetilde{A}(\downarrow p, \downarrow p) + \widetilde{B}_2(\downarrow p, \downarrow p)\widetilde{F}_p & 0 & \dots & \widetilde{B}_1(\downarrow p, p) \\ \hline I_{\underline{m}}(:, \downarrow 1)\widetilde{F}_1 & \dots & I_{\underline{m}}(:, \downarrow p)\widetilde{F}_p & 0 & \dots & 0 \end{array} \right],$$

where for each $j = 1, \dots, p$

$$\begin{aligned} \widetilde{F}_j &= -R_1(\downarrow j, \downarrow j)^{-1}(\widetilde{B}_2(\downarrow j, \downarrow j)^{\intercal} \widetilde{X}_j + D_{12}(:, \downarrow j)^{\intercal} \widetilde{C}_1(:, \downarrow j)) \quad \text{and} \\ \widetilde{X}_j &= \operatorname{Ric}(\widetilde{A}(\downarrow j, \downarrow j), \widetilde{B}_2(\downarrow j, \downarrow j), \widetilde{C}_1(:, \downarrow j), D_{12}(:, \downarrow j)). \end{aligned}$$

This gives an optimal solution $\widehat{Q}_{opt}$ to the one-sided reparameterized problem (7.29). However, we are looking for an optimal solution $\widehat{R}_{opt}$ to the reformulated problem (7.28). The difficulty is that it is not clear that it is at all possible to factorize $\widehat{Q}_{opt}$ as

$$\widehat{Q}_{opt} = \widehat{R}_{opt}\widehat{H}_{21} \quad \text{with} \quad \widehat{R}_{opt} \in \mathcal{I}_{\mathcal{P}}^{\underline{m} \times \underline{s}} \cap \mathcal{RH}_2.$$

Remark 7.2.7. The approach in Subsection 7.1.6 to the structured state feedback $\mathcal{H}_2$-control problem relies heavily on the special form of the matrix B-matrix in (7.12). Now one can construct a state space similarity so that an analogue of the realization (7.15) holds in the output feedback case, but the B-matrix there will not have the form $\begin{bmatrix} I & 0 \end{bmatrix}^{\intercal}$. The difficulty is that this specific form of the B-matrix plays an essential role in obtaining a stable realization for the the controller $\widehat{K}$.

Secondly, we already saw in the unstructured case in Section 4.7 that the solution of the one-sided reparameterized problem for the output feedback case does not factor through $\widehat{H}_{21}$.

Due to the above mentioned reasons, we propose to find some "reasonable" perturbation of $\widehat{Q}_{opt}$ so that it is indeed possible to factorize it as above. That is find structured $\mathcal{RH}_2$ functions $\widehat{Z}, \widehat{R} \in \mathcal{I}_{\mathcal{P}}^{\underline{m} \times \underline{s}}$ such that

$$\widehat{Q}_{opt} + \widehat{Z} = \widehat{R}\widehat{H}_{21}.$$

Equivalently, taking transposes, we have the following problem: find $\widehat{Z}^\mathsf{T}, \widehat{R}^\mathsf{T} \in \mathcal{I}_{\mathcal{P}_d}^{\underline{s} \times \underline{m}}$ such that

$$\widehat{Q}_{opt}^\mathsf{T} + \widehat{Z}^\mathsf{T} = \widehat{H}_{21}^\mathsf{T} \widehat{R}^\mathsf{T}.$$

In the above, $\mathcal{P}_d$ is the dual poset of $\mathcal{P}$. Since $\widehat{Z}^\mathsf{T}, \widehat{R}^\mathsf{T} \in \mathcal{I}_{\mathcal{P}_d}^{\underline{s} \times \underline{m}}$, we can write

$$\begin{aligned} \widehat{Z}^\mathsf{T} &= \begin{bmatrix} I_{\underline{m}}(:, \downarrow_d 1)\widehat{Z}^\mathsf{T}(\downarrow_d 1, 1) & \dots & I_{\underline{m}}(:, \downarrow_d p)\widehat{Z}^\mathsf{T}(\downarrow_d p, p) \end{bmatrix} \quad \text{and} \\ \widehat{R}^\mathsf{T} &= \begin{bmatrix} I_{\underline{m}}(:, \downarrow_d 1)\widehat{R}^\mathsf{T}(\downarrow_d 1, 1) & \dots & I_{\underline{m}}(:, \downarrow_d p)\widehat{R}^\mathsf{T}(\downarrow_d p, p) \end{bmatrix}. \end{aligned}$$

Applying Theorem 2.3.10, this allows us to consider the local equations

$$\begin{aligned} & \widehat{Q}_{opt}^\mathsf{T}(\downarrow_d j, j) + \widehat{Z}_j^\mathsf{T} = \widehat{H}_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j)\widehat{R}_j^\mathsf{T} \\ \text{subject to} \quad & \widehat{Z}_j^\mathsf{T}, \widehat{R}_j^\mathsf{T} \in \mathcal{RH}_2. \end{aligned} \tag{7.32}$$

Here $\widehat{Z}_j^\mathsf{T}$ and $\widehat{R}_j^\mathsf{T}$ do not have a specific zero structure. By analogy of Corollary 4.6.1, we get the following inner-outer factorization of $\widehat{H}_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j)$.

Lemma 7.2.8.
The function $\widehat{H}_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j)$ with realization

$$\widehat{H}_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j) = \left[\begin{array}{c|c} A^\mathsf{T}(\downarrow_d j, \downarrow_d j) + C_2^\mathsf{T}(\downarrow_d j, \downarrow_d j)M^\mathsf{T}(\downarrow_d j, \downarrow_d j) & C_2^\mathsf{T}(\downarrow_d j, \downarrow_d j) \\ \hline B_1^\mathsf{T}(\downarrow_d j, \downarrow_d j) + D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j)M^\mathsf{T}(\downarrow_d j, \downarrow_d j) & D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j) \end{array}\right]$$

has the inner-outer factorization

$$\widehat{H}_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j) = \widehat{V}_j^\mathsf{T} \widehat{M}_j^\mathsf{T}$$

where

$$\begin{aligned} \widehat{V}_j^\mathsf{T} &= \left[\begin{array}{c|c} A^\mathsf{T}(\downarrow_d j, \downarrow_d j) + C_2^\mathsf{T}(\downarrow_d j, \downarrow_d j)L_j^\mathsf{T} & C_2^\mathsf{T}(\downarrow_d j, \downarrow_d j) \\ \hline B_1^\mathsf{T}(\downarrow_d j, \downarrow_d j) + D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j)L_j^\mathsf{T} & D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j) \end{array}\right] (D_{21}(\downarrow_d j, \downarrow_d j)D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j))^{-\frac{1}{2}} \quad \textit{and} \\ \widehat{M}_j^\mathsf{T} &= (D_{21}(\downarrow_d j, \downarrow_d j)D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j))^{\frac{1}{2}} \left[\begin{array}{c|c} A^\mathsf{T}(\downarrow_d j, \downarrow_d j) + C_2^\mathsf{T}(\downarrow_d j, \downarrow_d j)M^\mathsf{T}(\downarrow_d j, \downarrow_d j) & C_2^\mathsf{T}(\downarrow_d j, \downarrow_d j) \\ \hline M^\mathsf{T}(\downarrow_d j, \downarrow_d j) - L_j^\mathsf{T} & I \end{array}\right] \end{aligned} \tag{7.33}$$

where

$$\begin{aligned} L_j^\mathsf{T} &= -(D_{21}(\downarrow_d j, \downarrow_d j)D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j))^{-1}(C_2(\downarrow_d j, \downarrow_d j)Y_j + D_{21}(\downarrow_d j, \downarrow_d j)B_1^\mathsf{T}(\downarrow_d j, \downarrow_d j)) \quad \textit{and} \\ Y_j &= \mathrm{Ric}(A^\mathsf{T}(\downarrow_d j, \downarrow_d j),\ C_2^\mathsf{T}(\downarrow_d j, \downarrow_d j),\ B_1^\mathsf{T}(\downarrow_d j, \downarrow_d j),\ D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j)). \end{aligned}$$

Proof.
We check the conditions of Theorem 3.7.13 for the realization of $\widehat{H}_{21}^\mathsf{T}(j, \downarrow_d j)$.

1. By assumption $A + MC_2$ is stable. Thus its transpose $A^\mathsf{T} + C_2^\mathsf{T} M^\mathsf{T}$ is also stable. By Lemma 6.2.12, $A^\mathsf{T}(\downarrow_d j, \downarrow_d j) + C_2^\mathsf{T}(\downarrow_d j, \downarrow_d j)M^\mathsf{T}(\downarrow_d j, \downarrow_d j)$ is stable.

2. Note that D_{21}^T has the dual poset-causal structure, that is, $D_{21}^\mathsf{T} \in \mathcal{I}_{\mathcal{P}_d}^{\underline{\ell} \times \underline{s}}$. By assumption, D_{21}^T has full column rank, so that $D_{21}^\mathsf{T}(:, \downarrow_d j)$ also has full column rank. From Corollary 2.3.12 we get that $(D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j))^\mathsf{T} D_{21}^\mathsf{T}(\downarrow_d j, \downarrow_d j) = (D_{21}^\mathsf{T}(:, \downarrow_d j))^\mathsf{T} D_{21}^\mathsf{T}(:, \downarrow_d j) > 0$.

3. The matrices appearing in

$$\begin{bmatrix} A^\intercal + C_2^\intercal M^\intercal i\omega I & C_2^\intercal \\ B_1^\intercal + D_{21}^\intercal M^\intercal & D_{21}^\intercal \end{bmatrix}$$

have the dual poset-causal structure. By an argument similar to the one given in Lemma 7.2.5, but utilizing the dual poset, it follows that

$$\begin{bmatrix} A^\intercal(\downarrow_d j, \downarrow_d j) + C_2^\intercal(\downarrow_d j, \downarrow_d j) M^\intercal(\downarrow_d j, \downarrow_d j) - i\omega I & C_2^\intercal(\downarrow_d j, \downarrow_d j) \\ B_1^\intercal(\downarrow_d j, \downarrow_d j) + D_{21}^\intercal(\downarrow_d j, \downarrow_d j) M^\intercal(\downarrow_d j, \downarrow_d j) & D_{21}^\intercal(\downarrow_d j, \downarrow_d j) \end{bmatrix}$$

has full column rank for all $\omega \in \mathbb{R}$.

Thus there exists a unique stabilizing solution to the Riccati equation

$$\begin{aligned} Y_j = \operatorname{Ric}(&A^\intercal(\downarrow_d j, \downarrow_d j) + C_2^\intercal(\downarrow_d j, \downarrow_d j) M^\intercal(\downarrow_d j, \downarrow_d j),\ C_2^\intercal(\downarrow_d j, \downarrow_d j), \\ &B_1^\intercal(\downarrow_d j, \downarrow_d j) + D_{21}^\intercal(\downarrow_d j, \downarrow_d j) M^\intercal(\downarrow_d j, \downarrow_d j),\ D_{21}^\intercal(\downarrow_d j, \downarrow_d j)) \\ = \operatorname{Ric}(&A^\intercal(\downarrow_d j, \downarrow_d j),\ C_2^\intercal(\downarrow_d j, \downarrow_d j),\ B_1^\intercal(\downarrow_d j, \downarrow_d j),\ D_{21}^\intercal(\downarrow_d j, \downarrow_d j)) \end{aligned}$$

and the result follows from Theorem 3.7.13. □

With $\widehat{V}_j^\intercal$ and $\widehat{M}_j^\intercal$ as in equation (7.33), we can find $\widehat{Z}_j^\intercal$ and $\widehat{R}_j^\intercal$ as in (7.32).

Lemma 7.2.9.
For $j \in P$, the functions

$$\widehat{Z}_j^\intercal := \widehat{V}_j^\intercal \mathrm{P}_{\mathcal{H}_2}(\widehat{V}_j^\intercal)^* \widehat{Q}_{opt}^\intercal(\downarrow_d j, j) - \widehat{Q}_{opt}^\intercal(\downarrow_d j, j) \quad \textit{and} \quad \widehat{R}_j^\intercal := (\widehat{M}_j^\intercal)^{-1} \mathrm{P}_{\mathcal{H}_2}(\widehat{V}_j^\intercal)^* \widehat{Q}_{opt}^\intercal(\downarrow_d j, j)$$

satisfy (7.32). Here $\mathrm{P}_{\mathcal{H}_2}$ is the orthogonal projection onto $\mathcal{H}_2$.

Proof.
The function $\widehat{Z}_j^\intercal$ is chosen in such a way that $\widehat{Q}_{opt}^\intercal(\downarrow_d j, j) + \widehat{Z}_j^\intercal$ is the projection of $\widehat{Q}_{opt}^\intercal(\downarrow_d j, j)$ onto the range $\widehat{V}_j^\intercal$, more precisely, onto the range of the Wiener Hopf operator $T_{\widehat{V}_j^\intercal}$ associated with the inner function $\widehat{V}_j^\intercal$ translated back to $\mathcal{H}_2$ functions on the right half plane via the Fourier transform, see [11, 12] for details. Then

$$\widehat{Q}_{opt}^\intercal(\downarrow_d j, j) + \widehat{Z}_j^\intercal = \mathrm{P}_{\mathcal{H}_2} \widehat{V}_j^\intercal \mathrm{P}_{\mathcal{H}_2}(\widehat{V}_j^\intercal)^* \widehat{Q}_{opt}^\intercal(\downarrow_d j, j) = \widehat{V}_j^\intercal \mathrm{P}_{\mathcal{H}_2}(\widehat{V}_j^\intercal)^* \widehat{Q}_{opt}^\intercal(\downarrow_d j, j)$$

where the second projection onto $\mathcal{H}_2$ dissapears because $\widehat{V}_j^\intercal$ is stable. This shows that $\widehat{Z}_j^\intercal$ is given by (7.32) and with this choice of $\widehat{Z}_j^\intercal$, it follows that

$$\begin{aligned} \widehat{Q}_{opt}^\intercal(\downarrow_d j, j) + \widehat{Z}_j^\intercal &= \widehat{V}_j^\intercal \mathrm{P}_{\mathcal{H}_2}(\widehat{V}_j^\intercal)^* \widehat{Q}_{opt}^\intercal(\downarrow_d j, j) \\ &= \widehat{V}_j^\intercal \widehat{M}_j^\intercal (\widehat{M}_j^\intercal)^{-1} \mathrm{P}_{\mathcal{H}_2}(\widehat{V}_j^\intercal)^* \widehat{Q}_{opt}^\intercal(\downarrow_d j, j) \\ &= \widehat{H}_{21}^\intercal(\downarrow_d j, \downarrow_d j) \widehat{R}_j^\intercal, \end{aligned}$$

with $\widehat{R}_j^\intercal$ as in (7.32). Moreover, since $\widehat{M}_j^\intercal$ is invertible outer, $\widehat{R}_j^\intercal \in \mathcal{RH}_2$. □

Define $\widehat{R} \in \mathcal{RH}_2$ and $\widehat{Z} \in \mathcal{RH}_2$ by

$$\widehat{R}^\intercal = \left[I(:, \downarrow_d 1) \widehat{R}_1^\intercal \quad \dots \quad I(:, \downarrow_d p) \widehat{R}_p^\intercal \right] \quad \text{and} \quad \widehat{Z}^\intercal = \left[I(:, \downarrow_d 1) \widehat{Z}_1^\intercal \quad \dots \quad I(:, \downarrow_d p) \widehat{Z}_p^\intercal \right].$$

Then $\widehat{R}$ and $\widehat{Z}$ have the poset-causal structure, that is $\widehat{R}, \widehat{Z} \in \mathcal{I}_{\mathcal{P}}^{s \times s}$ and $\widehat{Q}_{opt} + \widehat{Z} = \widehat{R}\widehat{H}_{21}$. Define

$$\widehat{K} = \underline{\mathcal{F}}\left(\widehat{J}, \widehat{R}\right).$$

Then $\widehat{K}$ is a feasible solution to the poset-causal $\mathcal{H}_2$ control Problem 7.2.1.

7.2.6 Partial Structured Optimization Approach

As in the previous section, we aim to solve the reformulated $\mathcal{H}_2$-control control problem in (7.28). By the column-wise separability of the $\mathcal{H}_2$-norm and Theorem 2.3.10, it follows that

$$\begin{aligned}
\|\widehat{H}_{11}+\widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2^2 &= \sum_{j=1}^{p}\|\widehat{H}_{11}(:,j)+\widehat{H}_{12}\widehat{R}\widehat{H}_{21}(:,j)\|_2^2 \\
&= \sum_{j=1}^{p}\|\widehat{H}_{11}(:,j)+\sum_{i=1}^{p}(\widehat{H}_{12}\widehat{R})(:,i)\widehat{H}_{21}(i,j)\|_2^2 \\
&= \sum_{j=1}^{p}\|\widehat{H}_{11}(:,j)+\sum_{i=1}^{p}\widehat{H}_{12}(:,\downarrow i)\widehat{R}(\downarrow i,i)\widehat{H}_{21}(i,j)\|_2^2 \\
&= \sum_{j=1}^{p}\|\widehat{H}_{11}(:,j)+\sum_{i\in\downarrow j}\widehat{H}_{12}(:,\downarrow i)\widehat{R}(\downarrow i,i)\widehat{H}_{21}(i,j)\|_2^2,
\end{aligned}$$

where the last equality follows because $\widehat{H}_{21}\in\mathcal{I}_{\mathcal{P}}^{s\times\ell}$, thus $\widehat{H}_{21}(i,j)=0$ if $j\not\preceq i$, that is, $\widehat{H}_{21}(i,j)=0$ if $i\notin\downarrow j$. In the above, $\widehat{R}(\downarrow i,i)$ does not have a specified zero-structure. Hence we consider p unstructured local control problems

$$\begin{aligned}
\text{minimize}\qquad & \|\widehat{H}_{11}(:,j)+\sum_{i\in\downarrow j}\widehat{H}_{12}(:,\downarrow i)\widehat{R}_{ji}\widehat{H}_{21}(i,j)\|_2^2 \\
\text{subject to}\qquad & \widehat{R}_{ji}\in\mathcal{RH}_2.
\end{aligned}$$

for $j\in P$ and $i\in\downarrow j$. Consider the case where $\widehat{R}_{ji}=0$ for $i\neq j$, that is, for $i\in\mathrel{\downarrow\!\!\!\!\!\;\cdot}\, j$. Then

$$\|\widehat{H}_{11}(:,j)+\sum_{i\in\downarrow j}\widehat{H}_{12}(:,\downarrow i)\widehat{R}_{ji}\widehat{H}_{21}(i,j)\|_2^2=\|\widehat{H}_{11}(:,j)+\widehat{H}_{12}(:,\downarrow j)\widehat{R}_{jj}\widehat{H}_{21}(j,j)\|_2^2.$$

Referring to the realizations in equation (7.27) and applying Theorem 2.3.10, we have the following compressed realizations

$$\widehat{H}_{11}(:,j)=\left[\begin{array}{c|c}\widetilde{A}(\downarrow j,\downarrow j) & \widetilde{B}_1(\downarrow j,j)\\ \hline \widetilde{C}_1(:,\downarrow j) & 0\end{array}\right]$$

$$=\left[\begin{array}{cc|c}A(\downarrow j,\downarrow j)+B_2(\downarrow j,\downarrow j)E(\downarrow j,\downarrow j) & -B_2(\downarrow j,\downarrow j)E(\downarrow j,\downarrow j) & B_1(\downarrow j,j)\\ 0 & A(\downarrow j,\downarrow j)+M(\downarrow j,\downarrow j)C_2(\downarrow j,\downarrow j) & B_1(\downarrow j,j)+M(\downarrow j,\downarrow j)D_{21}(\downarrow j,j)\\ \hline C_1(:,\downarrow j)+D_{12}(:,\downarrow j)E(\downarrow j,\downarrow j) & -D_{12}(:,\downarrow j)E(\downarrow j,\downarrow j) & 0\end{array}\right]$$

and

$$\widehat{H}_{12}(:,\downarrow j)=\left[\begin{array}{c|c}\widetilde{A}(\downarrow j,\downarrow j) & \widetilde{B}_2(\downarrow j,\downarrow j)\\ \hline \widetilde{C}_1(:,\downarrow j) & D_{12}(:,\downarrow j)\end{array}\right]=\left[\begin{array}{c|c}A(\downarrow j,\downarrow j)+B_2(\downarrow j,\downarrow j)E(\downarrow j,\downarrow j) & B_2(\downarrow j,\downarrow j)\\ \hline C_1(:,\downarrow j)+D_{12}(:,\downarrow j)E(\downarrow j,\downarrow j) & D_{12}(:,\downarrow j)\end{array}\right]$$

$$\begin{aligned}
\widehat{H}_{21}(j,j)&=\left[\begin{array}{c|c}\widetilde{A}(\downarrow j,\downarrow j) & \widetilde{B}_1(\downarrow j,j)\\ \hline \widetilde{C}_2(j,\downarrow j) & D_{21}(j,j)\end{array}\right]\\
&=\left[\begin{array}{c|c}A(\downarrow j,\downarrow j)+M(\downarrow j,\downarrow j)C_2(\downarrow j,\downarrow j) & B_1(\downarrow j,j)+M(\downarrow j,\downarrow j)D_{21}(\downarrow j,j)\\ \hline C_2(j,\downarrow j) & D_{21}(j,j)\end{array}\right].
\end{aligned}$$

We apply Theorem 4.6.6, to get local optimal solutions as follows.

Lemma 7.2.10.
Consider a poset-causal plant $\widehat{G}$ *as in* (7.24) *satisfying the following conditions*

1. $(A(j,j), B_2(j,j)), C_2(j,j))$ *is stabilizable and detectable for* $j \in P$*;*
2. $R_1 = D_{12}^{\intercal} D_{12} > 0$ *and* $R_{2j} = D_{21}(j,j) D_{21}^{\intercal}(j,j) > 0$ *for* $j \in P$*;*
3. $\begin{bmatrix} A - i\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ *has full column rank for all* $\omega \in \mathbb{R}$*;*
4. $\begin{bmatrix} A(\downarrow j, \downarrow j) + M(\downarrow j, \downarrow j) C_2(\downarrow j, \downarrow j) - i\omega I & B_1(\downarrow j, j) + M(\downarrow j, \downarrow j) D_{21}(\downarrow j, j) \\ C_2(j, \downarrow j) & D_{21}(j,j) \end{bmatrix}$ *has full row rank for* $\omega \in \mathbb{R}$ *for each* $j \in P$.

Let E *and* M *be as in* (7.25) *and let* $\widehat{H}_{11}$, $\widehat{H}_{12}$ *and* $\widehat{H}_{21}$ *be as in* (7.27). *For* $j \in P$, *the optimal solution to the* $\mathcal{H}_2$*-control problem*

$$\begin{aligned} &\text{minimize} && \|\widehat{H}_{11}(:, j) + \widehat{H}_{12}(:, \downarrow i) \widehat{R}_{jj} \widehat{H}_{21}(j,j)\|_2^2 \\ &\text{subject to} && \widehat{R}_{jj} \in \mathcal{RH}_2 \end{aligned}$$

is given by

$$\widehat{R}_{jj} = -\widehat{L}_j^{-1} \left[\begin{array}{c|c} A_{jR} & B_{jR} \\ \hline C_{jR} & 0 \end{array} \right] \widehat{M}_j^{-1},$$

where

$$\begin{aligned} A_{jR} &= \begin{bmatrix} A(\downarrow j, \downarrow j) + B_2(\downarrow j, \downarrow j) E(\downarrow j, \downarrow j) & -B_2(\downarrow j, \downarrow j) E(\downarrow j, \downarrow j) \\ 0 & A(\downarrow j, \downarrow j) + M(\downarrow j, \downarrow j) C_2(\downarrow j, \downarrow j) \end{bmatrix} \\ B_{jR} &= \begin{bmatrix} B_1(\downarrow j, j) \\ B_1(\downarrow j, j) + M(\downarrow j, \downarrow j) D_{21}(\downarrow j, j) \end{bmatrix} D_{21}^{\intercal}(j,j) + W_j C_2(j, \downarrow j) \\ C_{jR} &= B_2^{\intercal}(\downarrow j, \downarrow j) V_j + D_{12}^{\intercal}(\downarrow j, :) \begin{bmatrix} C_1(:, \downarrow j) + D_{12}(:, \downarrow j) E(\downarrow j, \downarrow j) & -D_{12}(:, \downarrow j) E(\downarrow j, \downarrow j) \end{bmatrix} \end{aligned}$$

with

$$\begin{aligned} \widehat{L}_j &= (D_{12}^{\intercal}(\downarrow j, :) D_{12}(:, \downarrow j))^{\frac{1}{2}} \left[\begin{array}{c|c} A(\downarrow j, \downarrow j) + B_2(\downarrow j, \downarrow j) E(\downarrow j, \downarrow j) & B_2(\downarrow j, \downarrow j) \\ \hline -F_j & I \end{array} \right], \\ \widehat{M}_j &= \left[\begin{array}{c|c} A(\downarrow j, \downarrow j) + M(\downarrow j, \downarrow j) C_2(\downarrow j, \downarrow j) & -L_j \\ \hline C_2(j, \downarrow j) & I \end{array} \right] (D_{21}(j,j) D_{21}^{\intercal}(j,j))^{\frac{1}{2}} \\ F_j &= -(D_{12}^{\intercal}(\downarrow j, :) D_{12}(:, \downarrow j))^{-1} (B_2^{\intercal}(\downarrow j, \downarrow j) X_j + D_{12}^{\intercal}(\downarrow j, :)(C_1(:, \downarrow j) + D_{12}(:, \downarrow j) E(\downarrow j, \downarrow j))), \\ L_j &= -(Y_j C_2^{\intercal}(\downarrow j, j) + (B_1(\downarrow j, j) + M(\downarrow j, \downarrow j) D_{21}(\downarrow j, j)) D_{21}^{\intercal}(j,j))(D_{21}(j,j) D_{21}^{\intercal}(j,j))^{-1}, \\ X_j &= \mathrm{Ric}(A(\downarrow j, \downarrow j) + B_2(\downarrow j, \downarrow j) E(\downarrow j, \downarrow j),\ B_2(\downarrow j, \downarrow j,\ C_1(:, \downarrow j) + D_{12}(:, \downarrow j) E(\downarrow j, \downarrow j),\ D_{21}(j,j) \quad \textit{and} \\ Y_j &= \mathrm{Ric}(A^{\intercal}(\downarrow j, \downarrow j) + C_2^{\intercal}(\downarrow j, \downarrow j) M^{\intercal}(\downarrow j, \downarrow j), C_2^{\intercal}(\downarrow j, j), B_1^{\intercal}(j, \downarrow j) + D_{21}^{\intercal}(j, \downarrow j) M^{\intercal}(\downarrow j, \downarrow j), D_{21}^{\intercal}(j,j)). \end{aligned}$$

We now define

$$\widehat{R} = \begin{bmatrix} I(:, \downarrow 1) \widehat{R}_{11} & \ldots & I(:, \downarrow p) \widehat{R}_{pp} \end{bmatrix}$$

Then by construction $\widehat{R}$ has the poset-causal structure. Furthermore, since $\widehat{R}_{jj} \in \mathcal{RH}_2$ for each $j \in P$, we have $\widehat{R} \in \mathcal{RH}_2$. Define

$$\widehat{K} = \underline{\mathcal{F}}(\widehat{J}, \widehat{R}).$$

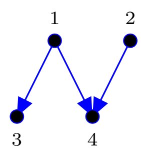
Then $\widehat{K}$ is a feasible controller for the poset-causal control Problem 7.2.1. Though $\widehat{K}$ is not necessarily optimal, we at least know its performance must be better than that of $\widehat{K}_0 = \underline{\mathcal{F}}\left(\widehat{J}, 0\right)$, because each $\widehat{R}_{jj}$ optimizes the norm

$$|\widehat{H}_{11}(:,j) + \widehat{H}_{12}(:,\downarrow j)\widehat{R}_{jj}\widehat{H}_{21}(j,j)\|_2^2.$$

The following example illustrates why we choose $\widehat{R}_{ji} = 0$ for $i \in \not\downarrow j$.

Example 7.2.11.

Consider a poset-causal system with underlying poset whose Hasse diagram is given by



Then

$$\downarrow 1 = \{1,3,4\}, \quad \downarrow 2 = \{2,4\}, \quad \downarrow 3 = \{3\} \quad \text{and} \quad \downarrow 4 = \{4\}.$$

The norm of the closed loop transfer function is equal to

$$\begin{aligned}
&\|\widehat{H}_{11} + \widehat{H}_{12}\widehat{R}\widehat{H}_{21}\|_2^2 \\
&= \sum_{j=1}^{p} \|\widehat{H}_{11}(:,j) + \sum_{i\in\downarrow j} \widehat{H}_{12}(:,\downarrow i)\widehat{R}(\downarrow i, i)\widehat{H}_{21}(i,j)\|_2^2 \\
&= \|\widehat{H}_{11}(:,1) + \widehat{H}_{12}(:,\downarrow 1)\widehat{R}(\downarrow 1,1)\widehat{H}_{21}(1,1) + \widehat{H}_{12}(:,\downarrow 3)\widehat{R}(\downarrow 3,3)\widehat{H}_{21}(3,1) + \widehat{H}_{12}(:,\downarrow 4)\widehat{R}(\downarrow 4,4)\widehat{H}_{21}(4,1)\|_2^2 \\
&\quad + \|\widehat{H}_{11}(:,2) + \widehat{H}_{12}(:,\downarrow 2)\widehat{R}(\downarrow 2,2)\widehat{H}_{21}(2,2) + \widehat{H}_{12}(:,\downarrow 4)\widehat{R}(\downarrow 4,4)\widehat{H}_{21}(4,2)\|_2^2 \\
&\quad + \|\widehat{H}_{11}(:,3) + \widehat{H}_{12}(:,\downarrow 3)\widehat{R}(\downarrow 3,3)\widehat{H}_{21}(3,3)\|_2^2 + \|\widehat{H}_{11}(:,4) + \widehat{H}_{12}(:,\downarrow 4)\widehat{R}(\downarrow 4,4)\widehat{H}_{21}(4,4)\|_2^2.
\end{aligned}$$

Note that $\widehat{R}(\downarrow 4, 4)$ appears three times and $\widehat{R}(\downarrow 3, 3)$ appears two times in the norm. This complicates matters when consider the optimization problems

$$\begin{aligned}
\text{minimize} \quad & \|\widehat{H}_{11}(:,j) + \sum_{i\in\downarrow j} \widehat{H}_{12}(:,\downarrow i)\widehat{R}_{ji}\widehat{H}_{21}(i,j)\|_2^2 \\
\text{subject to} \quad & \widehat{R}_{ji} \in \mathcal{RH}_2.
\end{aligned}$$

But if $\widehat{R}_{ji} = 0$ for $i \in \not\downarrow j$, then

$$\begin{aligned}
&\sum_{j=1}^{p} \|\widehat{H}_{11}(:,j) + \sum_{i\in\downarrow j} \widehat{H}_{12}(:,\downarrow i)\widehat{R}_{ji}\widehat{H}_{21}(i,j)\|_2^2 \\
&= \|\widehat{H}_{11}(:,1) + \widehat{H}_{12}(:,\downarrow 1)\widehat{R}_{11}\widehat{H}_{21}(1,1)\|_2^2 + \|\widehat{H}_{11}(:,2) + \widehat{H}_{12}(:,\downarrow 2)\widehat{R}_{22}\widehat{H}_{21}(2,2)\|_2^2 \\
&\quad + \|\widehat{H}_{11}(:,3) + \widehat{H}_{12}(:,\downarrow 3)\widehat{R}_{33}\widehat{H}_{21}(3,3)\|_2^2 + \|\widehat{H}_{11}(:,4) + \widehat{H}_{12}(:,\downarrow 4)\widehat{R}_{44}\widehat{H}_{21}(4,4)\|_2^2.
\end{aligned}$$

The above example shows that the entries $\widehat{R}_{ji}$ are taken to be zero for $i \in \not\downarrow j$ in order to avoid complications in the reconstruction of $\widehat{R}$ from the entries $\widehat{R}_{ji}$. Furthermore, it is not clear how to optimize the terms

$$\|\widehat{H}_{11}(:,j) + \sum_{i\in\downarrow j} \widehat{H}_{12}(:,\downarrow i)\widehat{R}_{ji}\widehat{H}_{21}(i,j)\|_2^2$$

due to the sum appearing in the norm.

7.2.7 Discussion of Solution Strategies

The recursive, bottom-up approach mentioned in Subsection 7.2.3 may be considered for hierarchical systems with a tree like communication structure, but may be unsuitable for general poset-causal structures as the one in Figure 7.1.

The approach taken in Subsection 7.2.4 aims to approximate a function appearing in the optimal solution of the $\mathcal{H}_2$-control problem considered as classical unstructured problem. This approach may work in the special case where $\widehat{H}_{12}$ and $\widehat{H}_{21}$ are inner and co-inner functions respectively, and, more generally, when the invertible outer factors $\widehat{L}$ and $\widehat{M}$ are block diagonal. If this does not happen, it is not clear how to choose a structured parameter $\widehat{R}$ due to the fact that we have a weighted norm in

$$\|\widehat{L}(\widehat{L}^{-1}\widehat{P}\widehat{M}^{-1} + \widehat{R})\widehat{M}\|_2^2.$$

The approach in Subsection 7.2.5 obtains an optimal solution to the one-sided reparameterized $\mathcal{H}_2$-control problem. There the optimal norm may be computed explicitly in terms of p localized Riccati equations. The Riccati equations are not independent and, using ideas from the bottom up approach, one may be able to get a better description of the Riccati solutions in relation to the Riccati solution from the unstructured problem. In the second part of this approach, we project onto $\mathcal{H}_2$ in determining $\widehat{Z}_j$ and $\widehat{R}_j$ in

$$\widehat{Z}_j^{\intercal} = \widehat{V}_j^{\intercal} \mathrm{P}_{\mathcal{H}_2} (\widehat{V}_j^{\intercal})^* \widehat{Q}_{opt}^{\intercal}(\downarrow_d j, j) - \widehat{Q}_{opt}^{\intercal}(\downarrow_d j, j) \quad \text{and} \quad \widehat{R}_j^{\intercal} = (\widehat{M}_j^{\intercal})^{-1} \mathrm{P}_{\mathcal{H}_2} (\widehat{V}_j^{\intercal})^* \widehat{Q}_{opt}^{\intercal}(\downarrow_d j, j)$$

As with the spectral factorization approach in Lemma 4.7.5, it may be possible to decompose $(\widehat{V}_j^{\intercal})^* \widehat{Q}_{opt}^{\intercal}(\downarrow_d j, j)$ into the sum of two terms, one which is analytic and the other which is anti-analytic. In that case a solution $\widehat{R}$ may be computed more explicitly. Since $\widehat{Q}_{opt}$ is optimal for the one-sided problem, one may hope that the solution obtained in this manner gives good performance.

The partial structured optimization approach of Subsection 7.2.6 is only applicable under additional local constraints

1. $R_{2j} = D_{21}(j,j) D_{21}^{\intercal}(j,j) > 0$ for $j \in P$;
2. $\begin{bmatrix} A(\downarrow j, \downarrow j) + M(\downarrow j, \downarrow j) C_2(\downarrow j, \downarrow j) - i\omega I & B_1(\downarrow j, j) + M(\downarrow j, \downarrow j) D_{21}(\downarrow j, j) \\ C_2(j, \downarrow j) & D_{21}(j,j) \end{bmatrix}$ has full row rank for $\omega \in \mathbb{R}$ for each $j \in P$

on the poset-causal system. However, under those conditions, it may be seen that the performance is at least better than that of the controller $\widehat{K} = \underline{\mathcal{F}}(\widehat{J}, 0)$. The choice of $\widehat{R}_{ji} = 0$ for $i \in \not\downarrow j$ may seem arbitrary at first, but, as is illustrated in Example 7.2.11, this is done in order to ensure that $\widehat{R}$ may be constructed in an unambiguous manner with entries in $\widehat{R}$ being optimally determined under the condition that $\widehat{R}_{ji} = 0$ for $i \in \not\downarrow j$.

Bibliography

[1] H. Bart, I. Gohberg, M.A. Kaashoek and A.C.M. Ran, *A State Space Approach to Canonical Factorization with Applications*, Oper. Theory Adv. Appl. **200** Birkhäuser Verlag, Basel, B2010.

[2] H. Bart, T. Ehrhardt, and B. Silbermann, L-free directed bipartite graphs and echelon-type canonical forms, *Oper. Theory Adv. Appl.* **271** (2018), 75–117.

[3] H. Bart, T. Ehrhardt, and B. Silbermann, Echelon type canonical forms in upper triangular matrix algebras, *Oper. Theory Adv. Appl.* **259** (2017), 79–124.

[4] H. Bart, T. Ehrhardt, and B. Silbermann, Rank decomposition in zero pattern matrix algebras, *Czechoslovak Math. J.* **66(141)** (2016), 987–1005.

[5] N. Cai and Y.S. Zhong, Formation controllability of high-order linear time invariant swarm systems, *IET Control Theory and Applications* **4** (2010), 646–654.

[6] M. Cantoni, E. Weyer, Y. Li, S.K. Ooi and I. Mareels, Control of Large-Scale Irrigation Networks, *Proc. IEEE* **95** (2007), 75–91.

[7] R.L. Davis, Algebras defined by patterns of zeros, *J. Combinatorial Theory* **9** (1970), 257–260.

[8] G.E. Dullerud and F. Paganini, *A Course in Robust Control Theory. A Convex Approach*, Texts in Applied Mathematics **36**, Springer-Verlag, New York, 2000.

[9] J.Z. Flórez, J. Martinez, G. Besancon and D. Faille, Explicit coordination for MPC-based distributed control with application to Hydro-Power Valleys, *50th IEEE Conference on Decision and Control and European Control Conference*, 2011.

[10] J.Z. Flórez, J. Martinez, G. Besancon and D. Faille, Decentralized-coordinated model predictive control for a hydro-power valley, *Mathematics and Computers in Simulation* **91** (2013), 108–118.

[11] A.E. Frazho, A.C.M. Ran, Inner outer factorization of wide rational matrix valued functions on the half plane, *Operator theory, functional analysis and applications* **282** (2021), 219—233.

[12] I. Gohberg, S. Goldberg, M.A. Kaashoek, *Basic classes of linear operators*, Birkhäuser Verlag, Basel, 2003.

[13] C. Heij, A.C.M. Ran, F. van Schagen, *Introduction to mathematical systems theory: linear systems, identification and control*, Birkhäuser Verlag, Basel, 2007.

[14] N. Jain, J.P. Koeln, S. Sundaram and A.G. Alleyne, Partially decentralized control of large-scale variable-refrigerant-flow systems in buildings, *Journal of Process Control*, **24** (2014), 798–819.

[15] Z. Ji, H. Lin and H. Yu, Leaders in multi-agent controllability under consensus algorithm and tree topology, *Systems Control Lett.* **61** (2012), 918–925.

[16] M. Ji and M. Egerstedt, A Graph-Theoretic Characterization of Controllability for Multi-agent Systems, *Proceedings of the 2007 American Control Conference*, 2007.

[17] R.E. Kalman, Contributions to the theory of optimal control, *Boletin Sociedad Matemática Mexicana* **5** (1960), 102–119.

[18] R.E. Kalman, On the general theory of control systems, *IFAC Proceedings Volumes* **1** (1960), 491–502.

[19] R.E. Kalman, Mathematical description of linear dynamical systems, *Journal of the Society for Industrial and Applied Mathematics, Series A: Control* **1(2)** (1963), 152–192.

[20] P.L. Kemper, Coordination control of linear systems, PhD thesis, VU Amsterdam, 2012.

[21] P.L. Kempker, A.C.M. Ran, and J.H. van Schuppen, Construction and minimality of coordinated linear systems, *Linear Algebra Appl.* **452** (2014), 202–236.

[22] P.L. Kempker, A.C.M. Ran, and J.H. van Schuppen, LQ control for coordinated linear systems, *IEEE Trans. Automat. Control* **59** (2014), 851–862.

[23] P.L. Kempker, A.C.M. Ran, and J.H. van Schuppen, Controllability and observability of coordinated linear systems, *Linear Algebra Appl.* **437** (2012), 121–167.

[24] L. Lessard, M. Kristalny, and A. Rantzer, On structured realizability and stabilizability of linear systems, *2013 American Control Conference*, pp. 5784–5790, IEEE, 2013.

[25] L. Lessard and S. Lall, Optimal controller synthesis for the decentralized two-player problem with output feedback, *In 2012 American Control Conference (ACC)*, pp. 6314–6321, IEEE, 2012.

[26] Y. Li, M. Cantoni, Distributed controller design for open water channels, *Proceedings of the 17th World Congress IFAC*, 2008.

[27] D. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.

[28] J.P. Lynch and K.H. Law, Decentralized Control Techniques for Large-scale Civil Structural Systems, *Proceedings of the 20th International Modal Analysis Conference (IMAC XX)*, 2002.

[29] D.K. Molzahn, F. Doërfler, H. Sandberg, S.H. Low, S. Chakrabarti and R. Baldick, A survey of Distributed Optimization and Algorithms for Electric Power Systems, *IEEE Transactions on Smart Grids* **8** (2017) 2941–2962.

[30] N. Monshizadeh, S. Zhang, and M.K. Camlibel, Zero forcing sets and controllability of dynamical systems defined on graphs, *IEEE Trans. Automat. Control* **59** (2014), 2562–2567.

[31] P.D. Moroçan, R. Bourdais, Didier Dumur and J. Buisson, Building temperature regulation using a distributed model predictive control, *Energy and Buildings* **42** (2010), 1445–1452.

[32] R.R. Negenborn, P.J. van Overloop and B. de Schutter, Coordinated Distributed Model Predictive Reach Control of Irrigation Canals, *Proceedings of the European Control Conference (ECC)*, IEEE, pp. 1420–1425, 2009.

[33] L. Rajaoarisoa, K. Horáth, E. Duviella and K. Chuquet, Large-scale system control on decentralized design, Application to Cuichy Fontinette Reach, *IFAC Proceedings* **47**, pp. 11105–11110, 2014.

[34] A.C.M. Ran and J.H. van Schuppen, Control for coordination of linear systems, *Proceedings of the 18th international symposium on the mathematical theory of networks and systems (MTNS 2008)*, 2008.

[35] M. Rotkowitz and S. Lall, A characterization of convex problems in decentralized control, *IEEE Transactions on Automatic Control* **51(2)** (2006), 274–286.

[36] A. Sadowska, P.J. van Overloop, C. Burt and B. De Schutter, Hierarchical Operation of Water Level Controllers: Formal Analysis and Apllication on a Large Scale Irrigation Canal, *Water Resource Management* **28** (2014), 4999–5019.

[37] P. Segovia, L. Rajaoarisoa, F. Nejjari, J. Blesa, V. Puig and E. Duviella, Decentralized Fault-Tolerant Control of Inland Navigation Networks: a Challenge, *Journal of Physics: Conference series* **783** (2017).

[38] P. Shah, A partial order approach to decentralized control, PhD thesis, Massachusetts Institute of Technology, 2011.

[39] P. Shah and P.A. Parrilo, H_2-optimal decentralized control over posets: a state-space solution for state-feedback, *IEEE Trans. Automat. Control* **58** (2013), 3084–3096.

[40] P. Shah and P.A. Parrilo, An optimal controller architecture for poset-causal systems, *50th IEEE Conference on Decision and Control and European Control Conference*, IEEE, 2011.

[41] P. Shah and P.A. Parrilo, A poset framework to model decentralized control problems, *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, IEEE, 2009.

[42] P. Shah and P.A. Parrilo, A partial order approach to decentralized control, *47th IEEE Conference on Decision and Control*, IEEE, 2008.

[43] J. Swigart and S. Lall, A Graph-Theoretic Approach to Distributed Control over Networks, *Joint 48th Conference on Decision and Control and 28th Chinese Control Conference*, pp. 5409–5414, 2009.

[44] J. Swigart and S. Lall, Optimal Synthesis and Explicit State-Space Solution for a Decentralized Two-Player Linear Quadratic Regulator, *49th IEEE Conference on Decision and Control*, pp. 132–137, 2010.

[45] J. Swigart, Optimal controller synthesis for decentralized systems, Stanford University, 2010.

[46] J. Swigart and S. Lall, Optimal Controller Synthesis for Decentralized Systems over graphs via Spectral Factorization, *IEEE Transactions on Automatic Control* **59** (2014), 2311–2323.

[47] H.G. Tanner, On the Controllability of Nearest Neighbor Interconnections, *43rd IEEE Conference on Decision and Control (CDC)*, pp. 2467–2472, 2004.

[48] S. ter Horst and J. Zeelie, Realization theory for poset-causal systems: controllability, observability and duality, *Mathematics of Control, Signals, and Systems* **33** (2021), 197–236.

[49] H.L. Trentelman, A.A. Stoorvogel and M.L.J. Hautus, *Control theory for linear systems*, Springer Science & Business Media, 2012.

[50] L. Wang, G. Chen, X. Wang and W.K.S. Tang, Controllability of networked MIMO systems, *Automatica* **69** (2016), 405–409.

[51] E. Weyer, Control of Irrigation Channels, *IEEE Transactions on Control Systems Technology* **16** (2008), 664–675.

[52] W.M. Wonham, *Linear Multivariable Control A Geometric Approach*, Springer-Verlag New York Berlin Heidelberg Tokyo, 1979.

[53] D.C. Youla, H.A. Jabr and J.J. Bongiorno, Modern Wiener-Hopf Design of Optimal Controllers: Part I, *IEEE Trans. Automat. Contr.* **AC-21** (1976), 3–13.

[54] D.C. Youla, H.A. Jabr and J.J. Bongiorno, Modern Wiener-Hopf Design of Optimal Controllers: Part II, *IEEE Trans. Automat. Contr.* **AC-21** (1976), 319–338.

[55] Q.C. Zhong, *Robust control of time-delay systems*, Springer Science and Media, 2006.

[56] K. Zhou, J. Doyle and K. Glover, *Robust and Optimal Control*, Prentice Hall, 1995.

List of Symbols

Special letters

$\mathbb{C}$	the set of complex numbers
$\mathbb{C}^n$	set of $n \times 1$ vectors with complex entries
$\mathbb{C}^{n\times m}$	set of $n \times m$ matrices with complex entries
$\mathbb{C}^+$	set of complex numbers with positive real part
$\overline{\mathbb{C}}^+$	set of complex numbers with non-negative real part
$\mathbb{C}^-$	set of complex numbers with negative real part
$E_{\succeq}^{\downarrow}$	edge set in Hasse diagram
$\underline{\mathcal{F}}$	lower linear fractional transformation of
$\overline{\overline{\mathcal{F}}}$	upper linear fractional transformation of
$\mathcal{G}_{\mathcal{P}}^{\downarrow}$	Hasse diagram related to a poset $\mathcal{P}$
$\mathcal{H}_2$	Hardy Hilbert space
$\mathcal{H}_\infty$	Hardy space
I_n	$n \times n$ identity matrix
$I_{\underline{n}}$	block $\underline{n} \times \underline{n}$ identity matrix
$\mathcal{I}_{\mathcal{T}}$	incidence algebra related to $\mathcal{T}$
$\mathcal{I}_{\mathcal{P}}^{\underline{n}\times\underline{m}}$	$\underline{n} \times \underline{m}$ block incidence set related to $\mathcal{P}$
$\mathcal{L}_2^{n\times m}$	Hilbert space of Lebesgue measurable $n \times m$ matrix functions
$\mathcal{N}(C, A)$	unobservable subspace
$\mathcal{O}(C, A)$	observability matrix
$\mathcal{P}$	a poset
$\mathbb{R}$	the set of real numbers
$\mathbb{R}^n$	set of $n \times 1$ vectors with real entries
$\mathbb{R}^{n\times m}$	set of $n \times m$ matrices with real entries
$\mathbb{R}^{\underline{n}\times\underline{m}}$	set of $\underline{n} \times \underline{m}$ block matrices
$\mathbb{R}^{\overline{n}\times\overline{m}}$	set of $\overline{n} \times \overline{m}$ super-block matrices
$\mathcal{R}$	indicates a subspace of rational functions
$\mathbb{Z}$	set of integers
$\mathbb{Z}_+$	set of non-negative integers
$\mathbb{Z}^n$	set of vectors with integer entries
$\mathbb{Z}_+^n$	vectors with non-negative integer entries

Arrows

$\downarrow R$	downstream set of (the set) R
$\uparrow R$	upstream set of (the set) R
$\downarrow i = \downarrow\{i\}$	downstream set of a singleton
$\uparrow i = \uparrow\{i\}$	upstream set of a singleton
$\twoheaddownarrow R$	strict downstream set of (the set) R
$\twoheaduparrow R$	strict upstream set of (the set) R
$\twoheaddownarrow i$	strict downstream set of a singleton
$\twoheaduparrow i$	strict upstream set of a singleton

Greek letters

$\delta_0(t)$	Dirac-delta function
Δ_1	Schur complement
Δ_2	Schur complement
γ	the canonical shuffle
$\Gamma_{\overline{n}}$	the block canonical shuffle matrix
$\rho(A)$	resolvent set of a matrix A
$\sigma(A)$	spectrum of a matrix A
Σ	linear system
Φ_A	resolvent of A ($(\lambda I - A)^{-1}$)

Letter operators

$\mathrm{adj}(A)$	adjugate of a matrix S
deg	degree
$\det(A)$	determinant of a matrix A
$\mathrm{diag}_j(A_j)$	diagonal matrix with entries A_j
$\dim(V)$	dimension of linear space V
dom(Ric)	domain of the Riccati operator
ess sup	essential supremum
$\mathrm{Im}(\lambda)$	imaginary part of a complex number λ
$\ker(T)$	kernel of a matrix T
min	minimum or minimize
$\mathrm{rank}(T)$	rank of a matrix T
$\mathrm{Re}(\lambda)$	real part of a complex number λ
Ric	solution to Riccati equation
$\mathrm{span}\{\cdot\}$	linear span of a set of vectors

Superscripts

$S^{\perp}$	orthogonal complement of the set S
$A^{\intercal}$	transpose of a matrix A
A^{-1}	inverse of a matrix A
A^{*}	conjugate transpose of a matrix A
$\widehat{W}^{*}(\lambda)$	conjugate of a proper real rational matrix function
M_{ij}^{st}	sub-block entry in a super-block matrix occurring in s^{th} block row and t^{th} block column of the block entry M_{ij}
p_A	characteristic polynomial of a matrix A

Symbols

$\oplus$	orthogonal direct sum
$\ominus$	orthogonal difference
$\geq$	larger than or equal to
$>$	strictly larger than
$\leq$	less than or equal to
$<$	strictly less than
$\succeq$	partial order
$\succ$	strict partial order ($i \succ j$ if $i \succeq$ and $i \neq j$)
$\preceq$	partial order ($i \preceq j$ equivalent to $j \succeq i$)
$\prec$	strict partial order ($i \prec j$ equivalent to $j \succ i$)
$\not\succeq$	not partially related to
$\not\preceq$	not partially related to
$\{ \}$	set braces
$(\cdot,\cdot)$	ordered pair
$\subseteq$	subset of
$\subset$	proper subset of
$\in$	element of
$\notin$	not an element of
$\emptyset$	the empty set
$\star$	Redheffer star product

Subscripts

$E_{\succeq}$	edge set related to a poset $(P,\succeq)$
G_{Σ}	input-output map of linear system Σ
$\mathcal{G}_{\mathcal{P}}$	digraph related to a poset $\mathcal{P}$
$\mathcal{P}_d$	dual poset of a poset $\mathcal{P}$
$\succeq_d$	dual partial order of a partial order $\succeq$
$\downarrow_d R$	downstream set of (the set) R in dual poset
$\uparrow_d R$	upstream set of (the set) R in dual poset
M_{ij}	block entry in a block matrix occurring in the i^{th} block row and j^{th} block column
Y_o^t	time dependent observability gramian
Y_o	observability gramian

Other

$M(R,S)$	block compression of a block matrix M with only block rows whose indices are in the set R and block columns whose indices are in the set S
$M(:,S)$	block compression of a block matrix M with all block rows, but only block columns whose indices are in the set S
$M(R,:)$	block compression of a block matrix M with all block columns, but only block rows whose indices are in the set R
$M(i,j)$	block entry of block matrix M occurring in the the i^{th} block row and j^{th} block column
$\lvert S\rvert$	cardinality of the set S
$\lvert\underline{n}\rvert$	size of the partition $\underline{n}$
$[M_{ij}]$	block matrix with block entries M_{ij}
$[[M_{ij}]]$	super-block matrix with entries M_{ij}
$\underline{n}_R$	see Definition 2.3.8
$\underline{n}$	a partition of n
$\overline{\underline{n}}$, $\check{\underline{n}}$, $\widetilde{\underline{n}}$	a sub-partition of n
$\widehat{F}$	Laplace transform of F

List of Acronyms

ARE	algebraic Riccati equation
CARE	continuous time algebraic Riccati equation
LFT	linear fractional transformation
LTI	linear time invariant