university of groningen

faculty of science and engineering

Strong Structural Properties of Structured Linear Systems

Master Project Applied Mathematics

August 2019

Student: B.M. Shali

First supervisor: Prof.dr. M.K. Camlibel

Second supervisor: Prof.dr. H.L. Trentelman

Abstract

This master thesis deals with structured systems and their strong structural properties. The structured systems studied in this thesis describe a family of linear systems in which the entries of the system matrices are fixed zeros, nonzeros or arbitrary real numbers. The definition of a structured system is formalized with the help of pattern matrices, and strong structural properties are characterized as rank properties of these pattern matrices. In parallel to these algebraic characterizations, we provide equivalent graph-theoretic characterizations based on the so-called system graph. In particular, we provide a sufficient condition for strong structural output controllability, necessary and sufficient conditions for strong structural input-state observability, and three mutually unrelated sufficient conditions for strong structural left invertibility. We also discuss the limitations of the approach taken in this thesis, which are most evident in the search for necessary conditions for strong structural output controllability and left invertibility. Finally, we extend the already existing necessary and sufficient conditions for strong structural controllability of linear systems to linear descriptor systems.

Contents

Chapter 1

Introduction

This thesis deals with properties of structured linear systems. The concept of structure for linear systems was introduced by Lin in [1] more than 40 years ago in order to obtain more realistic models of physical systems. He considers a single-input linear time-invariant system of the form

$$\dot{x} = Ax + Bu,$$

where x is the state, $\dot{x}$ is the derivative of the state, u is the input signal, and A and B are a real square matrix and a real vector, respectively. For ease of notation, we will denote this system as (A, B). The novel idea in [1] is that the entries of A and B are not known precisely, but are known to be fixed zeros or arbitrary real numbers. This pattern of fixed zero entries gives the system (A, B) its *structure*, and is what makes it a *structured system*. More precisely, another system $(\bar{A}, \bar{B})$ has the same structure as (A, B) if the fixed zero entries of A and B are also fixed zero entries of $\bar{A}$ and $\bar{B}$. Based on this concept of structure, a system (A, B) is said to be *(weakly) structurally controllable* if there exists at least one controllable system $(\bar{A}, \bar{B})$ with the same structure as (A, B). In [1], Lin characterizes (weak) structural controllability as a graph-theoretic property by associating a graph to the structured system (A, B).

Following this, there have been a number of papers that deal with structured systems and their properties. The results on (weak) structural controllability are extended to multi-input systems in [2] and are shown to be generic in [3]. In other words, it is shown that a system (A, B) being structurally controllable implies that *almost all* systems with the same structure as (A, B) are controllable. Although the uncontrollable systems with the same structure as a structurally controllable system (A, B) are atypical, in some cases the existence of such systems is not permitted. Consequently, the concept of *strong structural controllability* is introduced in [4]. A system (A, B) is said to be strongly structurally controllable if *all* systems with the same structure as (A, B) are controllable. Most of the research on strong structural controllability is based on the assumption that the entries of A and B are fixed zeros or nonzeros. However, this assumption is violated in several relevant applications, where the free choice of parameters, or lack of concrete knowledge about the model, can lead to entries in the system matrices that can be either zero or nonzero. For this reason, strong structural controllability is completely characterized in [5] for systems whose entries are allowed to be fixed zeros, nonzeros or arbitrary real numbers. There, as well as in most literature on the subject, the characterization of strong structural controllability is presented in two ways: algebraic and graph-theoretic.

The concept of structure has important applications in a number of areas. One such application is in networked systems, which can often be described by a structured system (see [6], [7] and [8]). In particular, if one considers a network of scalar linear systems, then the network can be described by a structured system, where the absence of communication between two systems in the network corresponds to a fixed zero in the system matrices of the structured system. Similarly, the presence of communication between two systems corresponds to a fixed nonzero in the system matrices. Although the links in the network are usually known, there are cases in which obtaining a detailed description of the network is impractical or even impossible. Some notable examples are social networks, world wide web, biological networks and ecological systems (see [9], [10], [11], [12] and [13]). Alternatively, it might happen that a detailed description of the network is available but the existence of some links cannot be guaranteed, whether due to possible malfunctions or a malicious attack on the network (see [14] and [15]). In both of these cases, the structured system describing the network must include entries that can be either zero or nonzero. Structured systems also arise in engineering applications when the value of the parameters describing a physical system are unknown or are allowed to vary. Although this usually leads to system matrices with fixed zero and nonzero entries only, it is not hard to imagine a situation in which some entries can be either zero or nonzero (see Example 5.1).

In light of this, we will consider structured systems in which the entries of the system matrices are fixed zeros, nonzeros or arbitrary real numbers. Furthermore, we will focus on strong structural properties of such systems, as opposed to weak structural properties. In particular, we will investigate strong structural output controllability, input-state observability and left invertibility. There has been some research into all three of these properties. Firstly, a special case of strong structural output controllability has been studied in [6] and [7] under the name of strong targeted controllability. There again, the system matrices have only fixed zero and nonzero entries. Strong targeted controllability and strong structural output controllability have obvious applications in the area of networked systems, as described in [6] and [7]. Secondly, strong structural input-state observability has been studied in [8] for linear time-varying systems, while weak structural input-state observability has been studied in [16] for linear time-invariant systems. The research into strong structural input-state observability in [8] has been motivated by its possible application in fault detection and isolation for systems with unknown or varying parameters. Thirdly, weak structural left invertibility has been studied in [17], but there does not seem to be any literature on strong structural left invertibility. In any case, the existing literature on these topics is either restricted to certain types of structured systems or takes a rather unnecessarily complicated approach. Instead, we will attempt to characterize the aforementioned strong structural properties within the general framework established in [5]. More precisely, we will describe structured systems with the help of pattern matrices and will characterize their strong structural properties as properties of these pattern matrices. Will will also relate pattern matrices and their algebraic properties to graphs and their graph-theoretic properties. This will allow us to state our results in two ways: algebraically and graph-theoretically.

This thesis is arranged as follows. The notation and some preliminary definitions are introduced in Chapter 2, after which linear systems and their relevant properties are described in Chapter 3. In Chapter 4, we introduce the concept of a pattern matrix, which is then used to formally define structured systems and strong structural properties in Chapter 5. A precise problem formulation is also included in the same chapter. Then Chapter 6, Chapter 7 and Chapter 8 deal with strong structural output controllability, input-state observability and left invertibility, respectively. Finally, we provide an extension of the results on strong structural controllability to linear descriptor systems in Chapter 9, followed by a conclusion in Chapter 10.

Chapter 2

Notation and preliminaries

In this chapter we will introduce the notation and some definitions that are used throughout the rest of the thesis. To begin with, the field of real numbers is denoted by $\mathbb{R}$ and the field of complex numbers by $\mathbb{C}$. The ring of polynomials with real coefficients and indeterminate s is denoted by $\mathbb{R}[s]$, and the field of rational functions with real coefficients and indeterminate s is denoted by $\mathbb{R}(s)$. For a given set S, the set of $n \times m$ matrices with entries in S is denoted by $S^{n\times m}$. Therefore, $\mathbb{R}^{n\times m}$, $\mathbb{R}[s]^{n\times m}$ and $\mathbb{R}(s)^{n\times m}$ denote the space of real matrices, polynomial matrices and rational matrices of size $n \times m$, respectively. Furthermore, S^n denotes vectors of size n and entries in the set S, that is, $S^n = S^{n\times 1}$. The Euclidean norm of a vector $x \in \mathbb{R}^n$ is denoted by $\|x\|_2$.

The identity matrix and the zero matrix are denoted by I and 0, respectively. For a given $q \times r$ matrix M, and sets $Y \in \{1, \dots, q\}$ and $U \subset \{1, \dots, r\}$, $M_{Y,U}$ denotes the submatrix of M containing the rows indexed by Y and the columns indexed by U. For ease of notation, if M is square, we will write $M^k_{Y,U}$ instead of $(M^k)_{Y,U}$. Furthermore, the k-th column of M is denoted by $M_{\bullet k}$, and the k-th row by $M_{k\bullet}$. A rational function $f(s) \in \mathbb{R}(s)$ is called *proper* if the degree of the numerator does not exceed the degree of the denominator, and it is called *strictly proper* if the degree of the numerator is less than the degree of the denominator. Then a rational matrix $M(s) \in \mathbb{R}(s)^{n\times m}$ is (strictly) proper if it contains only (strictly) proper rational functions. The normal rank of a rational matrix is defined as $\operatorname{rank} M(s) = \max_{\lambda\in\mathbb{C}} \operatorname{rank} M(\lambda)$. The rational matrix $M(s)$ is left invertible if there exists another rational matrix $M_L(s) \in \mathbb{R}(s)^{m\times n}$ such that $M_L(s)M(s) = I$. It can be shown that $M(s)$ is left invertible if and only if $M(\lambda)$ has full column rank for some $\lambda \in \mathbb{C}$, i.e., $\operatorname{rank} M(s) = m$. Moreover, the latter holds if and only if $M(\lambda)$ has full column rank for all except finitely many $\lambda \in \mathbb{C}$ (see [18]).

Throughout this thesis we will often make use of graphs. A graph is defined as an ordered pair of sets (V, E), where V is the *set of vertices* and $E \subset V \times V$ is the *set of edges*. An element of V is called a *vertex* or a *node*, and an element of E is called an *edge* or an *arc*. The set of edges consists of ordered pairs of vertices, indicating the vertices the edge connects as well as the direction of the edge. For example, if $i, j \in V$, then $(i, j) \in E$ if and only if there is an edge from i to j. Often we will not explicitly define the set of edges, but will only describe the condition under which there is an edge from i to j. If there is an edge from j to i, then we say that j is an *in-neighbour* of i, and i is an *out-neighbour* of j. The set of all in-neighbours of i is denoted by N_i^-, and the set of all out-neighbors of i is denoted by N_j^+. More formally, we define $N_i^- = \{j \in V \mid (j, i) \in E\}$ and $N_j^+ = \{i \in V \mid (j, i) \in E\}$. Given a graph (V, E), a *walk* is a finite sequence of vertices $(v_1, v_2, \dots, v_{k+1})$ where $(v_i, v_{i+1}) \in E$ for all $i \in \{1, \dots, k\}$. The *length* of a walk is one less than the number of vertices in the sequence, and a walk of length k will be referred to as a k-walk. Finally, a *path* is a walk in which the vertices in the sequence are distinct, and the distance from node i to node j is the length of the shortest path that starts at i and ends at j.

Chapter 3

Linear systems

In this chapter, we will introduce some concepts from the theory of linear systems that will be used throughout the rest of the thesis. Consider the linear time-invariant system

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{3.1a}$$
$$y(t) = Cx(t) + Du(t), \tag{3.1b}$$

where $t \geq 0$ represents time, $x(t) \in \mathbb{R}^n$ is the state, $\dot{x}(t) \in \mathbb{R}^n$ is the derivative of the state, $u(t) \in \mathbb{R}^m$ is the input, and $y(t) \in \mathbb{R}^p$ is the output. The system in (3.1) is completely characterized by the matrices A, B, C and D, hence we will denote it by (A, B, C, D). If the system does not include an output equation, then we will use the shorthand notation (A, B) to denote a system of the form (3.1a). Throughout this thesis we assume that the input signal u is infinitely differentiable, which we write as $u \in C^\infty$. We do this to avoid unnecessary technicalities and note that the results in this section can be derived for locally integrable functions as well (see Theorem 2.4.13 in [19] and the remark thereafter). For a given initial state $x(0) = x_0 \in \mathbb{R}^n$ and input signal $u \in C^\infty$, the corresponding state and output trajectories will be denoted by $x(t; x_0, u)$ and $y(t; x_0, u)$, respectively. We can write these trajectories explicitly by solving the differential equation in (3.1), which results in

$$x(t; x_0, u) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau \quad \text{and} \quad y(t; x_0, u) = Cx(t; x_0, u) + Du(t).$$

We will write $T(s) \in \mathbb{R}(s)^{p\times m}$ to denote the *transfer matrix* from u to y given by

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

and $P(s) \in \mathbb{R}[s]^{p\times m}$ to denote the *system matrix*

$$P(s) = \begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}.$$

For completeness, we will introduce a few definitions of properties of linear systems and will derive conditions under which a system (A, B, C, D) possesses each of these properties. To begin with, consider the following definition.

Definition 3.1. The system (A, B, C, D) is called *output controllable* if for any $y_1 \in \mathbb{R}^p$ and $x_0 \in \mathbb{R}^n$, there exists an input signal $u \in C^\infty$ and a time $T \geq 0$ such that $y(T; x_0, u) = y_1$.

The following theorem provides a well-known condition for output-controllability.

Theorem 3.1. *The system (A, B, C, D) is output controllable if and only if the matrix*

$$\begin{bmatrix} D & CB & CAB & \cdots & CA^{n-1}B \end{bmatrix}$$

has full row rank.

Proof. ($\Leftarrow$) Suppose that the matrix

$$\begin{bmatrix} D & CB & CAB & \cdots & CA^{n-1}B \end{bmatrix}$$

has full row rank, and note that the system is output controllable if and only if

$$\mathcal{Y} = \{y(T; x_0, u) \mid x_0 \in \mathbb{R}^n,\ u \in C^\infty,\ T \geq 0\} = \mathbb{R}^p.$$

It follows from the linearity of (A, B, C, D) that $\mathcal{Y}$ is a subspace of $\mathbb{R}^p$, hence it is equal to $\mathbb{R}^p$ if and only if its orthogonal complement contains only the zero vector. With this in mind, let $\eta \in \mathbb{R}^p$ be orthogonal to $\mathcal{Y}$. In other words, using the formula for a general output trajectory, we have that

$$\eta^\top \left(Ce^{AT}x_0 + C\int_0^T e^{A(T-\tau)}Bu(\tau)\,\mathrm{d}\tau + Du(T) \right) = 0, \tag{3.2}$$

for all $x_0 \in \mathbb{R}^n$, $u \in C^\infty$, and $T \geq 0$. If we fix $x_0 = 0$ and $T = 0$, then (3.2) implies that $\eta^\top Du(0) = 0$ for all $u(0) \in \mathbb{R}^m$, and hence $\eta^\top D = 0$. Consequently, if we fix $x_0 = 0$ and $T > 0$, and take $u(t) = B^\top e^{A^\top(T-t)}C^\top\eta$, equation (3.2) reduces to

$$\int_0^T \left\|\eta^\top Ce^{A(T-\tau)}B\right\|_2^2 \mathrm{d}\tau = 0.$$

The integrand is non-negative and continuous, hence the integral is zero if and only if $\eta^\top Ce^{At}B = 0$ for all $t \in [0, T]$. In particular, the function $\eta^\top Ce^{At}B$ and its derivatives must all be zero at $t = 0$, thus $\eta^\top CA^iB = 0$ for all non-negative integers i. But this implies that

$$\eta^\top \begin{bmatrix} D & CB & CAB & \cdots & CA^{n-1}B \end{bmatrix} = 0,$$

hence that $\eta = 0$. Therefore, the orthogonal complement of $\mathcal{Y}$ contains only the zero vector, so $\mathcal{Y} = \mathbb{R}^p$ and we conclude that (A, B, C, D) is output controllable.

($\Rightarrow$) Suppose that (A, B, C, D) is output controllable. For the sake of contradiction, suppose that the matrix

$$\begin{bmatrix} D & CB & CAB & \cdots & CA^{n-1}B \end{bmatrix}$$

does not have full row rank. In other words, there exists a nonzero vector $\eta \in \mathbb{R}^p$ such that

$$\eta^\top \begin{bmatrix} D & CB & CAB & \cdots & CA^{n-1}B \end{bmatrix} = 0.$$

As a consequence of the Cayley-Hamilton theorem, we have that $\eta^\top CA^iB = 0$ for all non-negative integers i. Using the definition of the matrix exponential, this implies that $\eta^\top Ce^{At}B = 0$ for all $t \geq 0$, hence $\eta^\top y(T; 0, u) = 0$ for all $u \in C^\infty$ and $T \geq 0$. Now let $y_1 \in \mathbb{R}^p$ be such that $\eta^\top y_1 \neq 0$, which is possible since $\eta \neq 0$. Given that (A, B, C, D) is output controllable, there exists an input signal $u \in C^\infty$ and time $T \geq 0$ such that $y(T; 0, u) = y_1$. But then $\eta^\top y(T; 0, u) = 0$ while $\eta^\top y_1 \neq 0$, which is a contradiction. Therefore, it must be the case that the matrix

$$\begin{bmatrix} D & CB & CAB & \cdots & CA^{n-1}B \end{bmatrix}$$

has full row rank. □

Next, we define the concept of left-invertiblity.

Definition 3.2. The system (A, B, C, D) is called *left invertible* if for all $x_0 \in \mathbb{R}^n$ and $u_1, u_2 \in C^\infty$ we have

$$y(t; x_0, u_1) = y(t; x_0, u_2) \text{ for all } t \geq 0 \quad \implies \quad u_1 = u_2.$$

In other words, for any given initial state, no pair of distinct input signals gives rise to the same output trajectory. Because of linearity, (A, B, C, D) is left invertible if and only if for all $u \in C^\infty$ we have

$$y(t; 0, u) = 0 \text{ for all } t \geq 0 \quad \implies \quad u = 0.$$

Left-invertibility of (A, B, C, D) can be characterized in different ways, only some of which are useful in the context of this thesis. These are included in the following theorem.

Theorem 3.2. *The following statements are equivalent:*

(i) (A, B, C, D) *is left invertible.*

(ii) $T(s)$ *is left invertible as a rational matrix.*

(iii) $P(s)$ *is left invertible as a rational matrix.*

(iv) $P(\lambda)$ *has full column rank for all except finitely many* $\lambda \in \mathbb{C}$.

Proof. We will first prove that the last three statements are equivalent. For (iii) and (iv), we know that a rational matrix is left invertible if and only if it has full column rank for almost all points in its domain. For (ii) and (iii), note that since $A - sI$ is invertible as a rational matrix, we can write

$$P(s) = \begin{bmatrix} A - sI & B \\ C & D \end{bmatrix} = \begin{bmatrix} A - sI & 0 \\ C & I \end{bmatrix} \begin{bmatrix} I & (A - sI)^{-1}B \\ 0 & T(s) \end{bmatrix}.$$

which implies that $P(s)$ is left invertible if and only if $T(s)$ is left invertible. Next, we will show that (i) implies (iv) and (ii) implies (i).

$\big($(i) $\Rightarrow$ (iv)$\big)$ Suppose that (A, B, C, D) is left invertible and that $\operatorname{rank} P(\lambda) < n + m$ for infinitely many $\lambda \in \mathbb{C}$. In that case, we can pick distinct scalars $\lambda_1, \ldots, \lambda_k$, where k is arbitrarily large, such that

$$\begin{bmatrix} A - \lambda_i I & B \\ C & D \end{bmatrix} \begin{bmatrix} x_{0i} \\ u_{0i} \end{bmatrix} = 0, \qquad \text{for all } i \in \{1, \ldots, k\}, \tag{3.3}$$

where $x_{0i} \in \mathbb{R}^n$ and $u_{0i} \in \mathbb{R}^m$ are not both zero. For large enough k, the vectors $x_{01}, \ldots, x_{0k}$, must be linearly dependent, that is, there exist scalars $\alpha_1, \ldots, \alpha_k$, not all zero, such that

$$\sum_{i=1}^{k} \alpha_i x_{0i} = 0.$$

Moreover, at least one of $u_{01}, \ldots, u_{0k}$ must be nonzero, otherwise we will have that $Ax_{0i} = \lambda_i x_{0i}$ for infinitely many distinct λ_i. Using (3.3), it is easy to verify that the input signal $u(t) = \sum_{i=1}^{k} \alpha_i e^{\lambda_i t} u_{0i} \neq 0$ results in the state trajectory $x(t; 0, u) = \sum_{i=1}^{k} \alpha_i e^{\lambda_i t} x_{0i}$ and the corresponding output is given by

$$y(t; 0, u) = Cx(t; 0, u) + Du(t) = \sum_{i=1}^{k} \alpha_i e^{\lambda_i t} (Cx_{0i} + Du_{0i}) = 0.$$

Therefore, $y(t; 0, u) = y(t; 0, 0)$ for all $t \geq 0$, while $u \neq 0$, which contradicts the assumption that (A, B, C, D) is left invertible and proves. In other words, it must be the case that $P(\lambda)$ has full column rank for all except finitely many $\lambda \in \mathbb{C}$.

$\big($(ii) $\Rightarrow$ (i)$\big)$ Suppose that $T(s)$ is left invertible, equivalently, that there exists a rational matrix $T_L(s) \in \mathbb{R}(s)^{m\times p}$ such that $T_L(s)T(s) = I$. We can obtain a *strictly proper* rational matrix from $T_L(s)$ by dividing all entries by s^k for a large enough integer k. This means that we can redefine T_L to be a strictly proper rational matrix such that $T_L(s)T(s) = Is^{-k}$, for some positive integer $k > 1$. Consequently, $T_L(s)$ can be written as the transfer matrix of a linear time-invariant system $(A_L, B_L, C_L, 0)$ for certain matrices A_L, B_L and C_L of appropriate sizes (see Theorem 3.3 in [20]). In other words, we have that $T_L(s) = C_L(sI - A_L)^{-1}B_L$, hence rewriting the equation $T_L(s)T(s) = Is^{-k}$ yields

$$C_L(sI - A_L)^{-1}B_L\left(C(sI - A)^{-1}B + D\right) = Is^{-k}.$$

Using the fact that $(sI - A)^{-1} = \sum_{i=0}^{\infty} A^i s^{-i-1}$, we can expand the latter to obtain

$$C_LB_LDs^{-1} + (C_LB_LCB + C_LA_LB_LD)\,s^{-2} + \cdots = Is^{-k}.$$

or, more precisely,

$$C_LB_LDs^{-1} + \sum_{l=2}^{\infty}\left(C_LA_L^{l-1}B_LD + \sum_{i=1}^{l-1} C_LA_L^{i-1}B_LCA^{l-i}B\right)s^{-l} = Is^{-k}. \tag{3.4}$$

Given that $k > 1$, the equation above implies that $C_LB_LD = 0$ and

$$C_LA_L^{l-1}B_LD + \sum_{i=1}^{l-1} C_LA_L^{i-1}B_LCA^{l-i}B = \begin{cases} I & \text{if } l = k, \\ 0 & \text{otherwise.}\end{cases} \tag{3.5}$$

Now, suppose that the output of (A, B, C, D) is fed as input to $(A_L, B_L, C_L, 0)$, that is, the systems are connected via a series interconnection. Then $(A_L, B_L, C_L, 0)$ can be written as

$$\begin{aligned}\dot{x}_L &= A_Lx_L + B_Ly,\\ v &= C_Lx_L,\end{aligned}$$

where x_L is the state, v is the output, and the time variable t has been omitted in order to ease the notation. The dynamics of the interconnection is given by

$$\begin{bmatrix}\dot{x}\\ \dot{x}_L\end{bmatrix} = \begin{bmatrix}A & 0\\ B_LC & A_L\end{bmatrix}\begin{bmatrix}x\\ x_L\end{bmatrix} + \begin{bmatrix}B\\ B_LD\end{bmatrix}u, \quad v = \begin{bmatrix}0 & C_L\end{bmatrix}\begin{bmatrix}x\\ x_L\end{bmatrix}.$$

Using the general solution, we find that the output $v(t; 0, u)$ is given by

$$v(t; 0, u) = \begin{bmatrix}0 & C_L\end{bmatrix}\int_0^t \exp\left(\begin{bmatrix}A & 0\\ B_LC & A_L\end{bmatrix}(t-\tau)\right)\begin{bmatrix}B\\ B_LD\end{bmatrix}u(\tau)\,\mathrm{d}\tau$$

It turns out that we can drastically simplify this expression for $v(t; 0, u)$. To this end, note that

$$\begin{bmatrix}A & 0\\ B_LC & A_L\end{bmatrix}^{l-1} = \begin{bmatrix}A^{l-1} & 0\\ \sum_{i=1}^{l-1} A_L^{i-1}B_LCA^{l-i} & A_L^{l-1}\end{bmatrix},$$

which implies that

$$\begin{bmatrix}0 & C_L\end{bmatrix}\begin{bmatrix}A & 0\\ B_LC & A_L\end{bmatrix}^{l-1}\begin{bmatrix}B\\ B_LD\end{bmatrix} = C_LA_L^{l-1}B_LD + \sum_{i=1}^{l-1} C_LA_L^{i-1}B_LCA^{l-i}B.$$

Using (3.5), we find that

$$\begin{bmatrix}0 & C_L\end{bmatrix}\begin{bmatrix}A & 0\\ B_LC & A_L\end{bmatrix}^{l-1}\begin{bmatrix}B\\ B_LD\end{bmatrix} = \begin{cases} I & \text{if } l = k, \\ 0 & \text{otherwise.}\end{cases} \tag{3.6}$$

Then substituting the expression for the matrix exponential into the formula for the output yields

$$v(t;0,u) = \int_0^t \frac{(t-\tau)^{k-1}}{(k-1)!} u(\tau)\,\mathrm{d}\tau.$$

Differentiating both sides k times results in $v^{(k)}(t;0,u) = u(t)$, where $v^{(k)}$ indicates the k-th derivative of v. Finally, suppose that the input signal $u \in C^\infty$ is such that $y(t;0,u) = 0$ for all $t \geq 0$. Feeding this into $(A_L, B_L, C_L, 0)$ results in $v(t;0,y) = 0$ for all $t \geq 0$. However, in the context of the interconnection, this means that $v(t;0,u) = 0$ for all $t \geq 0$, hence $v^{(k)}(t;0,u) = u(t) = 0$ for all $t \geq 0$, which shows that (A,B,C,D) is left invertible. □

Next, we consider a notion that is slightly stronger than that of left invertibility, namely, input-state observability.

Definition 3.3. The system (A,B,C,D) is called *input-state observable* if for all $x_{01}, x_{02} \in \mathbb{R}^n$ and $u_1, u_2 \in C^\infty$ we have

$$y(t;x_{01},u_1) = y(t;x_{02},u_2) \text{ for all } t \geq 0 \quad \Longrightarrow \quad x_{01} = x_{02} \text{ and } u_1 = u_2.$$

In other words, the input signal and initial state that produce a given output trajectory are unique. Because of linearity, (A,B,C,D) is input-state observable if and only if for all $x_0 \in \mathbb{R}^n$ and $u \in C^\infty$ we have

$$y(t;x_0,u) = 0 \text{ for all } t \geq 0 \quad \Longrightarrow \quad x_0 = 0 \text{ and } u = 0.$$

Naturally, the condition for input-state observability is very similar to that for left invertibility.

Theorem 3.3. *The system (A,B,C,D) is input-state observable if and only if $P(\lambda)$ has full column rank for all $\lambda \in \mathbb{C}$.*

In order to prove this result, we need to introduce the so-called weakly unobservable subspace of (A,B,C,D) and use one of its properties.

Definition 3.4. A point $x_0 \in \mathbb{R}^n$ is called *weakly unobservable* if there exists an input signal $u \in C^\infty$ such that $y(t;x_0,u) = 0$ for all $t \geq 0$. The set of all weakly unobservable points is denoted by $\mathcal{V}$ and is called the *weakly unobservable subspace* of (A,B,C,D).

The fact that $\mathcal{V}$ is a subspace follows from the linearity of (A,B,C,D). One can also show that $\mathcal{V}$ is controlled invariant, that is, for all $x_0 \in \mathcal{V}$ and input signal $u \in C^\infty$ such that $y(t;x_0,u) = 0$ for all $t \geq 0$, we have that $x(T;x_0,u) \in \mathcal{V}$ for all $T \geq 0$. Indeed, if $x_T = x(T;x_0,u)$ and $u_T(t) = u(t+T)$, then

$$y(t;x_T,u_T) = y(t+T;x_0,u) = 0 \text{ for all } t \geq 0,$$

hence $x_T \in \mathcal{V}$. This leads to the following result.

Lemma 3.1. *There exists a matrix $F \in \mathbb{R}^{m\times n}$ such that*

$$(A+BF)\mathcal{V} \subset \mathcal{V} \quad \textit{and} \quad (C+DF)\mathcal{V} = \{0\}.$$

Proof. Let $\{x_1,\ldots,x_r\}$ be a basis for $\mathcal{V}$. Then for all $i \in \{1,\ldots,r\}$ there exists an input signal $u_i \in C^\infty$ such that $x(t;x_i,u_i) \in \mathcal{V}$ and $y(t;x_i,u_i) = 0$ for all $t \geq 0$. In particular, as $\mathcal{V}$ is closed, it follows that

$$\dot{x}(0^+;x_i,u_i) = Ax_i + Bu_i(0) = \lim_{t\downarrow 0} \frac{x(t;x_i,u_i) - x_i}{t} \in \mathcal{V},$$

while $Cx_i + Du_i(0) = 0$. Now, let $F \in \mathbb{R}^{m\times n}$ be any matrix such that $Fx_i = u_i(0)$, and note that $(A + BF)x_i \in \mathcal{V}$ and $(C + DF)x_i = 0$ for all $i \in \{1\dots, r\}$. Given that $\{x_1, \dots, x_r\}$ is a basis for $\mathcal{V}$, we conclude that

$$(A + BF)\mathcal{V} \subset \mathcal{V} \quad \text{and} \quad (C + DF)\mathcal{V} = \{0\}.$$

□

We can now prove Theorem 3.3.

Proof of Theorem 3.3. (⇒) Suppose that (A, B, C, D) is input-state observable and that

$$\operatorname{rank} \begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix} < n + m,$$

for some $\lambda \in \mathbb{C}$. Then we can find $x_0 \in \mathbb{R}^n$ and $u_0 \in \mathbb{R}^m$, not both zero, such that

$$\begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix} \begin{bmatrix} x_0 \\ u_0 \end{bmatrix} = 0.$$

Using the latter, we can easily verify that the input signal $u(t) = e^{\lambda t}u_0$ and initial state x_0 yield the state trajectory $x(t; x_0, u) = e^{\lambda t}x_0$ and the corresponding output is given by

$$y(t; x_0, u) = Cx(t; x_0, u) + Du(t) = e^{\lambda t}(Cx_0 + Du_0) = 0.$$

Therefore, $y(t; x_0, u) = 0$ for all $t \geq 0$, while $x_0 \neq 0$ or $u \neq 0$, which contradicts the assumption that (A, B, C, D) is input-state observable.

(⇐) Suppose that the matrix

$$\begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix}$$

has full column rank for all $\lambda \in \mathbb{C}$. Using Theorem 3.2, this implies that (A, B, C, D) is left invertible, hence that

$$y(t; 0, u) = 0 \text{ for all } t \geq 0 \quad \Longrightarrow \quad u = 0.$$

Now, let the initial condition $x_0 \in \mathbb{R}^n$ and input signal $u \in C^\infty$ be such that $y(t; x_0, u) = 0$ for all $t \geq 0$. If we show that $x_0 = 0$, then the proof would follow from left invertibility of (A, B, C, D). In other words, our goal is to show that $\mathcal{V} = \{0\}$. Suppose, on the contrary, that $\mathcal{V} \neq \{0\}$. From Lemma 3.1, we know that there exists a matrix $F \in \mathbb{R}^{m\times n}$ such that

$$(A + BF)\mathcal{V} \subset \mathcal{V} \quad \text{and} \quad (C + DF)\mathcal{V} = \{0\}.$$

Since $\mathcal{V} \neq \{0\}$ is $(A + BF)$-invariant, it must contain an eigenvector of $A + BF$. Therefore, there exists a nonzero vector $v \in \mathcal{V}$ such that $Av + BFv = \lambda v$ and $Cv + DFv = 0$ for some $\lambda \in \mathbb{C}$. Then we have

$$\begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix} \begin{bmatrix} v \\ Fv \end{bmatrix} = 0,$$

which is only possible if $v = 0$. This leads to a contradiction and we conclude that $\mathcal{V} = \{0\}$. Consequently, $x_0 = 0$ and $y(t; 0, u) = 0$ for all $t \geq 0$, hence $u = 0$ because (A, B, C, D) is left invertible. □

Chapter 4

Pattern matrices

In this chapter, we will introduce the concept of pattern matrix in order to formalize the idea of matrices whose entries are not known but are known to be zeros, nonzeros or arbitrary real numbers. To this end, consider the symbols $0, *$ and $?$. These will be used to represent matrix entries whose values are zeros (0), nonzeros ($*$) or arbitrary real numbers (?). In particular, let $\{0, *, ?\}^{q\times r}$ denote the set of all $q \times r$ matrices with entries from the set $\{0, *, ?\}$. We will refer to matrices belonging to this set as *pattern matrices*. The connection to real matrices and the meaning of the symbols is captured in the following definition.

Definition 4.1. Let $\mathcal{A} \in \{0, *, ?\}^{q\times r}$. The *pattern class* of $\mathcal{A}$, denoted by $\mathcal{P}(\mathcal{A})$, is defined as

$$\mathcal{P}(\mathcal{A}) = \{A \in \mathbb{R}^{q\times r} \mid A_{ij} = 0 \text{ if } \mathcal{A}_{ij} = 0 \text{ and } A_{ij} \neq 0 \text{ if } \mathcal{A}_{ij} = *\}.$$

In other words, the pattern class of a $q\times r$ pattern matrix is the set of all $q\times r$ real matrices whose entries are zeros, nonzeros or arbitrary real numbers, depending on the corresponding symbol in the pattern matrix. It can easily be verified that the pattern class of a matrix is invariant under nonzero scaling of rows and columns.

We can define properties of pattern matrices in terms of properties of the real matrices within their pattern classes. For example, we say that a pattern matrix $\mathcal{A}$ has full column (row) rank if A has full column (row) rank for all $A \in \mathcal{P}(\mathcal{A})$. Rank properties are a central theme in this thesis because most system properties, like controllability, observability, invertibility, etc., are characterized in terms of a rank condition on some matrix associated with the system. It can be shown that a pattern matrix has full row rank if and only if it has a specific structure. To make this more precise, consider the following definition.

Definition 4.2. Let $\mathcal{A} \in \{0, *, ?\}^{q\times r}$ with $q \leq r$. We say that $\mathcal{A}$ is of *Form III* if we can permute the columns and rows of $\mathcal{A}$ in such a way that the resulting matrix is of the form

$$\begin{bmatrix} \otimes & \cdots & \otimes & * & 0 & \cdots & 0 \\ \otimes & \cdots & \otimes & \otimes & * & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \otimes & \cdots & \otimes & \otimes & \otimes & \cdots & * \end{bmatrix},$$

where $\otimes$ indicates an entry that can be either 0, $*$ or ?.

The term Form III is introduced in [21], where it is defined for pattern matrices containing 0's and $*$'s only. The generalized definition provided here can be found in [5]. The following lemma can also be found in [5] (see Lemma 19).

Lemma 4.1. *Let $\mathcal{A} \in \{0, *, ?\}^{q \times r}$ with $q \leq r$. Then $\mathcal{A}$ has full row rank if and only if $\mathcal{A}$ is of Form III.*

In other words, a pattern matrix $\mathcal{A} \in \{0, *, ?\}^{q \times r}$ with $q \leq r$ has full row rank if we can permute the rows and columns of $\mathcal{A}$ in such a way that we obtain a $q \times q$ lower triangular submatrix with $*$'s on the diagonal. Although Lemma 4.1 gives a condition under which a pattern matrix has full row rank, we can also use it as a condition under which a pattern matrix has full column rank. In particular, if we define the transpose of a pattern matrix in the obvious way, then it immediately follows that a pattern matrix $\mathcal{A} \in \{0, *, ?\}^{q \times r}$ with $q \geq r$ has full column rank if and only if $\mathcal{A}^\top$ is of Form III.

In practice, we will be working with several unknown matrices that belong to the pattern classes of some known pattern matrices. This will naturally lead to expressions involving sums and products. In order to gain understanding of how the results of such expressions look like, we will define a sensible way of adding and multiplying pattern matrices. Here sensible means that the result of adding and multiplying pattern matrices gives us some useful information on the result of adding and multiplying matrices belonging to their pattern classes.

To this end, we will define addition for a pair of pattern matrices in such a way that the sum of any pair of real matrices belonging to their pattern class is contained in the pattern class of the sum of the pattern matrices. We know that the sum of zero and any number is just the number itself, while the sum of two nonzero numbers can be anything. Motivated by this, we define addition for the set $\{0, *, ?\}$ as shown in the table below.

$+$	0	$*$	$?$
0	0	$*$	$?$
$*$	$*$	$?$	$?$
$?$	$?$	$?$	$?$

Then adding pattern matrices is defined in the usual way, i.e., element-wise.

Example 4.1. Consider the pattern matrices

$$\mathcal{A} = \begin{bmatrix} 0 & * & 0 \\ ? & * & * \\ 0 & ? & * \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 0 & 0 & * \\ * & ? & * \\ ? & ? & 0 \end{bmatrix}.$$

Using the table above, we find that their sum is

$$\mathcal{A} + \mathcal{B} = \begin{bmatrix} 0 & * & * \\ ? & ? & ? \\ ? & ? & * \end{bmatrix}.$$

By definition, we have that for pattern matrices $\mathcal{A}$ and $\mathcal{B}$ of the same dimensions, it holds that $\mathcal{P}(\mathcal{A} + \mathcal{B}) \supset \mathcal{P}(\mathcal{A}) + \mathcal{P}(\mathcal{B})$, where $\mathcal{P}(\mathcal{A}) + \mathcal{P}(\mathcal{B}) = \{A + B \mid A \in \mathcal{P}(\mathcal{A}) \text{ and } B \in \mathcal{P}(\mathcal{B})\}$. It turns out that the converse is true as well.

Proposition 4.1. *For pattern matrices $\mathcal{A}$ and $\mathcal{B}$ of the same dimensions, it holds that*

$$\mathcal{P}(\mathcal{A} + \mathcal{B}) = \mathcal{P}(\mathcal{A}) + \mathcal{P}(\mathcal{B}).$$

Proof. The inclusion $\mathcal{P}(\mathcal{A} + \mathcal{B}) \supset \mathcal{P}(\mathcal{A}) + \mathcal{P}(\mathcal{B})$ follows directly from the definition of $\mathcal{A} + \mathcal{B}$. For the other inclusion, let $M \in \mathcal{P}(\mathcal{A} + \mathcal{B})$ and consider an element M_{ij}. The goal is to show that there exist entries $A_{ij} \in \mathcal{P}(\mathcal{A}_{ij})$ and $B_{ij} \in \mathcal{P}(\mathcal{B}_{ij})$ such that $M_{ij} = A_{ij} + B_{ij}$. We will consider the cases $M_{ij} = 0$ and $M_{ij} \neq 0$ separately.

Suppose that $M_{ij} = 0$. Then either $(\mathcal{A} + \mathcal{B})_{ij} = 0$ or $(\mathcal{A} + \mathcal{B})_{ij} = ?$. In the former, we must have that $\mathcal{A}_{ij} = 0$ and $\mathcal{B}_{ij} = 0$, hence $A_{ij} = 0$ and $B_{ij} = 0$ would work. In the latter, there are several possibilities whose solutions are listed in the table below.

$\mathcal{A}_{ij}$	$\mathcal{B}_{ij}$	A_{ij}	B_{ij}
$*, ?$	$*, ?$	1	-1
0	?	0	0
?	0	0	0

Suppose that $M_{ij} \neq 0$. Then either $(\mathcal{A} + \mathcal{B})_{ij} = *$ or $(\mathcal{A} + \mathcal{B})_{ij} = ?$. In the former, exactly one of $\mathcal{A}_{ij}$ and $\mathcal{B}_{ij}$ is $*$ and the other one is 0, hence we can pick either $A_{ij} = M_{ij}$ and $B_{ij} = 0$, or $A_{ij} = 0$ and $B_{ij} = M_{ij}$. In the latter, there are several possibilities whose solutions are listed in the table below.

$\mathcal{A}_{ij}$	$\mathcal{B}_{ij}$	A_{ij}	B_{ij}
$*, ?$	$*, ?$	$\frac{M_{ij}}{2}$	$\frac{M_{ij}}{2}$
0	?	0	M_{ij}
?	0	M_{ij}	0

The element M_{ij} was chosen arbitrarily, hence we can always find matrices $A \in \mathcal{P}(\mathcal{A})$ and $B \in \mathcal{P}(\mathcal{B})$ such that and $A + B = M$. This implies that $\mathcal{P}(\mathcal{A} + \mathcal{B}) \subset \mathcal{P}(\mathcal{A}) + \mathcal{P}(\mathcal{B})$, which concludes the proof. □

In the same vein, we now turn to the definition of a product of pattern matrices. First, we note that the product of zero and any number is just zero, while the product of two nonzero numbers is always a nonzero number. This motivates the definition of multiplication for the set $\{0, *, ?\}$ shown in the table below.

$\cdot$	0	$*$	?
0	0	0	0
$*$	0	$*$	?
?	0	?	?

.

Then we can define pattern matrix multiplication in the usual way, that is, for $\mathcal{A} \in \mathbb{R}^{q \times r}$ and $\mathcal{B} \in \mathbb{R}^{r \times l}$ we have

$$(\mathcal{A}\mathcal{B})_{ij} = \sum_{k=1}^{r} \mathcal{A}_{ik}\mathcal{B}_{kj},$$

for all $i \in \{1, \ldots, q\}$ and $j \in \{1, \ldots, l\}$.

Example 4.2. The product of the pattern matrices

$$\mathcal{A} = \begin{bmatrix} * & 0 & ? \\ 0 & * & * \\ ? & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{B} = \begin{bmatrix} * & 0 \\ ? & * \\ 0 & * \end{bmatrix}$$

is given by

$$\mathcal{AB} = \begin{bmatrix} * & ? \\ ? & ? \\ ? & 0 \end{bmatrix}.$$

By definition, the product of real matrices belonging to the pattern classes of some pattern matrices is contained in the pattern class of the product of the pattern matrices. In other words, it holds that $\mathcal{P}(\mathcal{A})\mathcal{P}(\mathcal{B}) \subset \mathcal{P}(\mathcal{AB})$, where $\mathcal{P}(\mathcal{A})\mathcal{P}(\mathcal{B}) = \{AB \mid A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}$. However, the converse is generally not true. When multiplying matrices with at least two rows or columns, we usually create dependencies between the entries of the product. These dependencies cannot be captured by the symbolic operations with pattern matrices.

Example 4.3. Consider the pattern vectors

$$\mathcal{A} = \begin{bmatrix} * \\ * \end{bmatrix} \quad \text{and} \quad \mathcal{B} = \begin{bmatrix} * & * \end{bmatrix}.$$

Their product is easily computed as

$$\mathcal{AB} = \begin{bmatrix} * & * \\ * & * \end{bmatrix},$$

whose pattern class contains the matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \in \mathcal{P}(\mathcal{AB}).$$

Note that the latter is a matrix of rank 2, thus it cannot be written as the product of two vectors. In other words, the fact that the columns (or rows) of AB, where $A \in \mathcal{P}(\mathcal{A})$ and $B \in \mathcal{P}(\mathcal{B})$, are always linearly dependent cannot be inferred from the product $\mathcal{AB}$.

We will conclude this section with a short discussion on powers of pattern matrices and their entries. In particular, we will introduce a graph-theoretic procedure to compute the entries of a pattern matrix raised to some power. To begin with, a pattern matrix can be associated to a graph, where the existence and the type of the edges connecting the vertices is determined by the entries of the pattern matrix. In some sense, the pattern matrix serves as an adjacency matrix. To make this more precise, consider a pattern matrix $\mathcal{M} \in \{0, *, ?\}^{q \times r}$. Then the associated graph will have $\max(q, r)$ vertices and there will be an edge from vertex j to vertex i if and only if $\mathcal{M}_{ij} \neq 0$. The edge will be *solid* if $\mathcal{M}_{ij} = *$ and *dashed* (or *non-solid*) if $\mathcal{M}_{ij} =?$. The latter should indicate that the existence of the edge is not guaranteed, in parallel to the fact that M_{ij} for $M \in \mathcal{P}(\mathcal{M})$ could be either zero or nonzero.

Definition 4.3. Let $\mathcal{M} \in \{0, *, ?\}^{q \times r}$. The *graph associated to* $\mathcal{M}$ is the graph $(V, E_* \cup E_?)$, where $V = \{1, \dots, \max(q, r)\}$ is the vertex set, while

$$E_* = \{(j, i) \in V \times V \mid \mathcal{M}_{ij} = *\} \quad \text{and} \quad E_? = \{(j, i) \in V \times V \mid \mathcal{M}_{ij} = ?\},$$

are the sets of solid and dashed edges, respectively. This graph will be denoted by $\mathscr{G}(\mathcal{M})$.

We will distinguish two types of walks in such a graph, solid and non-solid. The walk $(v_1, \dots, v_k)$ is called solid if $(v_i, v_{i+1}) \in E_*$ for all $i \in \{1, \dots, k-1\}$, i.e., if the edges connecting consecutive vertices in the walk are all solid. The walk is non-solid if it is not solid, i.e., at least one of the edges connecting consecutive vertices in the walk is dashed.

Using the definition of a graph associated to a pattern matrix, we obtain the following result on the entries of a pattern matrix raised to some power.

Lemma 4.2. *Let $\mathcal{M} \in \{0, *, ?\}^{q\times q}$ and let $\mathscr{G}(\mathcal{M})$ be the graph associated to $\mathcal{M}$. For $\mathcal{M}^k$ we have*

$$\begin{aligned}
(\mathcal{M}^k)_{ij} = 0 \quad &\Leftrightarrow \quad \textit{there is no } k\textit{-walk from } j \textit{ to } i,\\
(\mathcal{M}^k)_{ij} = * \quad &\Leftrightarrow \quad \textit{there is a unique solid } k\textit{-walk from } j \textit{ to } i,\\
(\mathcal{M}^k)_{ij} = ? \quad &\Leftrightarrow \quad \textit{there is a non-unique or non-solid } k\textit{-walk from } j \textit{ to } i.
\end{aligned}$$

Proof. We will prove this by induction on k. The statement is true for $k = 1$ by definition of $\mathscr{G}(\mathcal{M})$. Suppose it is true for some positive integer k, and consider $\mathcal{M}^{k+1}$, which is given by

$$(\mathcal{M}^{k+1})_{ij} = \sum_{l=1}^{n} \mathcal{M}_{il}(\mathcal{M}^k)_{lj}.$$

Note that $(\mathcal{M}^{k+1})_{ij} = 0$ if and only if all terms in the above sum are equal to 0. A term $\mathcal{M}_{il}(\mathcal{M}^k)_{lj}$ is equal to 0 if and only if $\mathcal{M}_{il} = 0$ or $(\mathcal{M}^k)_{lj} = 0$, equivalently, if there is no edge from l to i or if there is no k-walk from j to l. Therefore, $(\mathcal{M}^{k+1})_{ij} = 0$ if and only if there is no k-walk from j to any of the in-neighbors of i, i.e., there is no $(k+1)$-walk from j to i.

On the other hand, note that $(\mathcal{M}^{k+1})_{ij} = *$ if and only if exactly one term in the sum is equal to $*$ and the rest are 0. A term $\mathcal{M}_{il}(\mathcal{M}^k)_{lj}$ is equal to $*$ if and only if both $\mathcal{M}_{il}$ and $(\mathcal{M}^k)_{lj}$ are equal to $*$, equivalently, if there is a unique solid k-walk from j to l and there is a solid edge from l to i. Therefore, $(\mathcal{M}^{k+1})_{ij} = *$ if and only if i has only one in-neighbor to which there is a k-walk from j, this k-walk is unique and solid, and the edge connecting that in-neighbor to i is solid. In other words, $(\mathcal{M}^{k+1})_{ij} = *$ if and only if there is a unique solid $(k+1)$-walk from j to i.

Finally, suppose $(\mathcal{M}^{k+1})_{ij} =?$. Since $(\mathcal{M}^{k+1})_{ij} \neq 0$, we know that there is a $(k+1)$-walk from j to i, but since $(\mathcal{M}^{k+1})_{ij} \neq *$, this $(k+1)$-walk is not unique or not solid. For the contrary, suppose there is a non-unique or non-solid $(k+1)$-walk from j to i. Since there is a $(k+1)$-walk, we know that $(\mathcal{M}^{k+1})_{ij} \neq 0$, and since that walk is not unique or not solid (or both), we also know that $(\mathcal{M}^{k+1})_{ij} \neq *$. Therefore it must be the case that $(\mathcal{M}^{k+1})_{ij} =?$. $\square$

We finish this section with an example of a graph associated to a pattern matrix and the use of Lemma 4.2 when computing the powers of the pattern matrix.

Example 4.4. Consider the pattern matrix

$$\mathcal{M} = \begin{bmatrix} * & 0 & ? & 0\\ * & * & ? & 0\\ 0 & 0 & * & ?\\ 0 & 0 & * & * \end{bmatrix}.$$

The graph $\mathscr{G}(\mathcal{M})$ associated to $\mathcal{M}$ is depicted in Figure 4.1. To determine the first column of

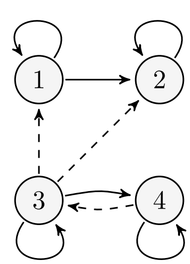


Figure 4.1: The graph $\mathscr{G}(\mathcal{M})$.

$\mathcal{M}^2$, we can look at 2-walks from node 1 to the rest of the nodes. The only 2-walk from 1 to itself is the one where we stay at 1, i.e., $(1, 1, 1)$. Since the edge from 1 to 1 is solid, it follows that this 2-walk is unique and solid, hence $\mathcal{M}^2_{11} = *$. On the other hand, there are two 2-walks from 1 to 2, namely, $(1, 1, 2)$ and $(1, 2, 2)$, hence $\mathcal{M}^2_{12} =?$. Finally, there is no 2-walk from 1 to neither 3 nor 4, hence $\mathcal{M}^2_{13} = \mathcal{M}^2_{14} = 0$. The rest of the entries of $\mathcal{M}^2$ can be found in the same fashion.

Chapter 5

Structured systems and strong structural properties

In this chapter, we will define structured systems and strong structural properties of such systems. This will allow us to write a precise problem formulation, the partial solution of which will be the topic of the remainder of this thesis. To this end, consider the linear time-invariant system

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{5.1a}$$
$$y(t) = Cx(t) + Du(t), \tag{5.1b}$$

which we have been denoting by (A, B, C, D). Suppose that the entries of the matrices A, B, C and D are not known precisely, but are known to be zeros, nonzeros or arbitrary real numbers. In other words, we know that $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ for some known pattern matrices $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$. This naturally leads to a family of systems (A, B, C, D) as A, B, C and D range over $\mathcal{P}(\mathcal{A}), \mathcal{P}(\mathcal{B}), \mathcal{P}(\mathcal{C})$ and $\mathcal{P}(\mathcal{D})$, respectively. This family of systems is completely characterized by the pattern matrices $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$, which is why we will denote it by $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. We will refer to $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ as a *structured system*, or just the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. If the structured system is obtained from a system of the form (5.1a), then we will simply denote it by $(\mathcal{A}, \mathcal{B})$.

Example 5.1. Consider the circuit depicted in Figure 5.1. It consists of an external voltage

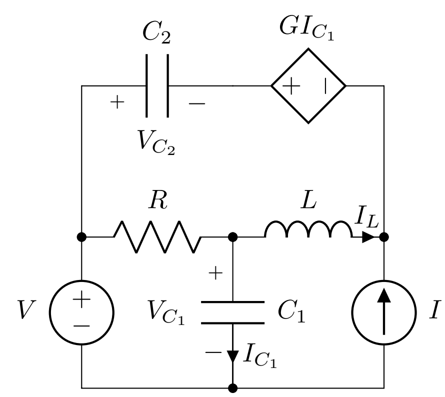


Figure 5.1: Example of an electrical circuit.

source V, an external current source I, two capacitors with capacitance C_1 and C_2, a resistor with resistance R, an inductor with inductance L, and a current controlled voltage source GI_{C_1}, where I_{C_1} is the current through the first capacitor and G is the gain. Define the state vector as $x = \begin{bmatrix} V_{C_1} & V_{C_2} & I_L \end{bmatrix}^\top$, where V_{C_1} and V_{C_2} are the voltages across the capacitors, and I_L is the current through the inductor. Furthermore, define the input vector as $u = \begin{bmatrix} V & I \end{bmatrix}^\top$, and assume that only V_{C_2} and the current through the first capacitor I_{C_1} are available for measurement, i.e., define the output vector as $y = \begin{bmatrix} V_{C_2} & I_{C_1} \end{bmatrix}^\top$. Applying Kirchhoff's current and voltage laws results in the linear time-invariant system

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

where A, B, C and D are given by

$$\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right] = \left[\begin{array}{ccc:cc} -\frac{1}{RC_1} & 0 & -\frac{1}{C_1} & \frac{1}{RC_1} & 0 \\ 0 & 0 & -\frac{1}{C_2} & 0 & -\frac{1}{C_2} \\ \frac{R-G}{RL} & \frac{1}{L} & -\frac{G}{L} & \frac{G-R}{RL} & 0 \\ \hdashline 0 & 1 & 0 & 0 & 0 \\ -\frac{1}{R} & 0 & -1 & \frac{1}{R} & 0 \end{array}\right]$$

Note that some entries in the system matrices are zero for any choice of parameters. Since the constants C_1, C_2, R, L and G are nonzero, entries such as $-1/RC_1$ and $-G/L$ are nonzero for any choice of parameters. However, the entries $(R-G)/RL$ and $(G-R)/RL$ can be either zero or nonzero depending on whether $R = G$ or not. With this in mind, define the pattern matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ as

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{ccc:cc} * & 0 & * & * & 0 \\ 0 & 0 & * & 0 & * \\ ? & * & * & ? & 0 \\ \hdashline 0 & * & 0 & 0 & 0 \\ * & 0 & * & * & 0 \end{array}\right].$$

Then it is clear that $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ for any choice of parameters.

We are interested in properties of (A, B, C, D) that stem solely from the fact that $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. As such, these are more appropriately seen as properties of the structured system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. To make this more precise, let P be a system-theoretic property and consider the following definition.

Definition 5.1. The structured system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is said to have a strong structural property P if the system (A, B, C, D) has the property P for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$.

We have already discussed one example of a strong structural property in the introduction, namely, strong structural controllability. Using the definition above, the system $(\mathcal{A}, \mathcal{B})$ is strongly structurally controllable if (A, B) is controllable for all $A \in \mathcal{P}(\mathcal{A})$ and $B \in \mathcal{P}(\mathcal{B})$. Conditions for strong structural controllability can be found in [5], where it is shown that $(\mathcal{A}, \mathcal{B})$ is strongly structurally controllable if and only if the pattern matrices

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \bar{\mathcal{A}} & \mathcal{B} \end{bmatrix}$$

have full row rank, where $\bar{\mathcal{A}}$ is obtained from $\mathcal{A}$ by modifying the diagonal entries as

$$\bar{\mathcal{A}}_{kk} = \begin{cases} * & \text{if } \mathcal{A}_{kk} = 0, \\ ? & \text{otherwise.} \end{cases}$$

Two necessary and sufficient conditions under which a pattern matrix has full row rank are also given in [5], one algebraic and one graph-theoretic. The algebraic condition was already discussed in the previous section (see Lemma 4.1), but that condition does not provide a systematic way of checking whether a given pattern matrix has full row rank. For this reason, we provide the graph-theoretic condition as well. This has to do with the outcome of the repeated application of a *color change rule* for the graph associated to a pattern matrix.

Color change rule.

Let $\mathcal{M} \in \{0, *, ?\}^{q\times r}$ and consider the corresponding graph $\mathscr{G}(\mathcal{M})$. Suppose that each node of $\mathscr{G}(\mathcal{M})$ is colored either black or white. If node i has exactly one white out-neighbor j, and the edge from i to j is solid, then color j black. We say that node i colors node j black.

Suppose that initially all nodes are colored white. The derived set $\mathscr{D}_{\text{ccr}}$ is the set of all black nodes after repeated application of the color change rule until no more changes are possible. We say that a node is *colorable* if it is contained in $\mathscr{D}_{\text{ccr}}$. and we say that $\mathscr{G}(\mathcal{M})$ is *colorable* if $\mathscr{D}_{\text{ccr}}$ contains q nodes. The graph-theoretic condition is given in the following lemma (Theorem 11 in [5]).

Lemma 5.1. *Let $\mathcal{M} \in \{0, *, ?\}^{q\times r}$ with $q \leq r$. Then $\mathcal{M}$ has full row rank if and only if $\mathscr{G}(\mathcal{M})$ is colorable.*

Inspired by these results, we will attempt to solve the following problem.

Problem formulation.

Find algebraic and graph-theoretic necessary and sufficient conditions for strong structural output controllability, left invertibility and input-state observability of $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

Strong structural output controllability has been studied with some restrictions under the name of strong targeted controllability (see [6] and [7]). There, the focus is on a structured system that arises from a network of scalar dynamical systems, only some of which can be influenced by an input signal. Then the problem is to find a topological characterization of the set of nodes whose states can be controlled. A topological characterization is one that depends solely on the structure of the network, i.e., the graph representing the network, and not on the specific dynamics of the individual systems. It turns out that this is effectively a strong structural output controllability problem with some restrictions on the matrices describing the system (A, B, C, D) and the corresponding structured system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. On the one hand, C and B are restricted to submatrices of the identity matrix, i.e., the rows of C are distinct rows of the the $n \times n$ identity matrix, and the columns of B are distinct columns of the $n \times n$ identity matrix. On the other hand, $\mathcal{D} = 0$ and the entries of $\mathcal{A}$ are restricted to $*$'s and 0's only. We will not make any such restrictions and will devote Chapter 6 to strong structural output controllability within the general context provided here.

Strong structural input-state observability has been studied in [8] for linear time-varying systems, while weak structural input-state observability for linear time-invariant systems has been studied in [16]. The approach taken in [8] appears unnecessarily complicated in the context of linear time-invariant systems. At the same time, the results in [16] are purely graph-theoretic and do not appear to be useful within the framework that we are developing in this thesis. For these reasons, we will take an algebraic approach based on the results on strong structural controllability in [5], which will lead to a relatively simple graph-theoretic test for strong structural input-state observability. The relevant discussion can be found in Chapter 7.

Strong structural left invertibility does not seem to have been studied in such a general context. However, *weak* structural left invertibility has been studied in [17], that is, conditions under which (A, B, C, D) is left invertible for *some* $(A, B, C, D) \in (\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. There again, the entries of the pattern matrices are restricted, this time to 0's and ?'s only. In view of this, we find it worthwhile to investigate strong structural left invertibility within the framework provided here. In fact, this will be the main topic of this thesis, the discussion of which can be found in Chapter 8.

Chapter 6

Output controllability

In this section, we will derive a sufficient condition for strong structural output-controllability of $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, and we will discuss the difficulties in obtaining necessary conditions. In view of Theorem 3.1 and the properties of addition and multiplication of pattern matrices, the following result should not be surprising.

Theorem 6.1. *The system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally output controllable if the pattern matrix*

$$\begin{bmatrix} \mathcal{D} & \mathcal{C}\mathcal{B} & \mathcal{C}\mathcal{A}\mathcal{B} & \cdots & \mathcal{C}\mathcal{A}^{n-1}\mathcal{B} \end{bmatrix}$$

has full row rank.

Proof. Let $A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. Recall that $\mathcal{P}(\mathcal{C})\mathcal{P}(\mathcal{B}) \subset \mathcal{P}(\mathcal{C}\mathcal{B})$, that is, $CB \in \mathcal{P}(\mathcal{C}\mathcal{B})$ for all $C \in \mathcal{P}(\mathcal{C})$ and $B \in \mathcal{P}(\mathcal{B})$. Similarly, we find that

$$\mathcal{P}(\mathcal{C})\mathcal{P}(\mathcal{A})\mathcal{P}(\mathcal{B}) \subset \mathcal{P}(\mathcal{C}\mathcal{A})\mathcal{P}(\mathcal{B}) \subset \mathcal{P}(\mathcal{C}\mathcal{A}\mathcal{B}),$$

and, more generally, that $\mathcal{P}(\mathcal{C})\mathcal{P}(\mathcal{A})^i\mathcal{P}(\mathcal{B}) \subset \mathcal{P}(\mathcal{C}\mathcal{A}^i\mathcal{B})$ for all positive integers i. In other words, we have that

$$\begin{bmatrix} D & CB & CAB & \cdots & CA^{n-1}B \end{bmatrix} \subset \mathcal{P}(\begin{bmatrix} \mathcal{D} & \mathcal{C}\mathcal{B} & \mathcal{C}\mathcal{A}\mathcal{B} & \cdots & \mathcal{C}\mathcal{A}^{n-1}\mathcal{B} \end{bmatrix}).$$

hence, using Theorem 3.1, we conclude that (A, B, C, D) is output-controllable for all $(A, B, C, D) \in A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, equivalently, $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally output controllable. □

One can check for strong structural output controllability by first computing the pattern matrix

$$\begin{bmatrix} \mathcal{D} & \mathcal{C}\mathcal{B} & \mathcal{C}\mathcal{A}\mathcal{B} & \cdots & \mathcal{C}\mathcal{A}^{n-1}\mathcal{B} \end{bmatrix},$$

then using the color change rule to determine whether it has full row rank. However, we can also provide a graph-theoretic procedure that circumvents the computation of this pattern matrix. Although this does not necessarily provide an easier algorithm for checking strong structural output controllability, it does provide some intuition behind the condition of Theorem 6.1. To this end, we will define the so-called system graph associated to the structured system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. The system graph has already been used in [17] to characterize weak structural left invertibility.

Definition 6.1. The system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ associated to $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is defined as the graph associated to the square $(n+m+p) \times (n+m+p)$ pattern matrix

$$\mathcal{M} = \begin{bmatrix} \mathcal{A} & \mathcal{B} & 0 \\ 0 & 0 & 0 \\ \mathcal{C} & \mathcal{D} & 0 \end{bmatrix}.$$

Note that the nodes of $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ can be partitioned into the sets $X = \{1,\ldots,n\}$, $U = \{n+1,\ldots,n+m\}$ and $Y = \{n+m+1,\ldots,n+m+p\}$. The nodes in these sets will be referred to as state nodes, input nodes and output nodes, respectively. Then we have that

- there is an edge from the j-th input node to the i-th state node if $\mathcal{B}_{ij} \neq 0$,
- there is an edge from the j-th state node to the i-th state node if $\mathcal{A}_{ij} \neq 0$,
- there is an edge from the j-th state node to the i-th output node if $\mathcal{C}_{ij} \neq 0$,
- there is an edge from the j-th input node to the i-th output node if $\mathcal{D}_{ij} \neq 0$.

Therefore, the system graph captures how states are influenced by other states and inputs, and how outputs are influenced by states and inputs. For example, if $\mathcal{B}_{31} = 0$, i.e., there is no edge from the first input node to the third state node, then the input signal u_1 does not directly influence the state x_3 in any system (A,B,C,D) with $B \in \mathcal{P}(\mathcal{B})$. In order to emphasize this, as well as for convenience, we will label the i-th state node as x_i, the i-th input node as u_i, and the i-th output node as y_i. To see how this is relevant in the context of output controllability, note that

$$\mathcal{M}^k = \begin{bmatrix} \mathcal{A}^k & \mathcal{A}^{k-1}\mathcal{B} & 0 \\ 0 & 0 & 0 \\ \mathcal{C}\mathcal{A}^{k-1} & \mathcal{C}\mathcal{A}^{k-2}\mathcal{B} & 0 \end{bmatrix}$$

for $k \geq 2$. This implies that

$$\mathcal{M}^k_{Y,U} = \begin{cases} \mathcal{D} & \text{if } k = 1, \\ \mathcal{C}\mathcal{A}^{k-2}\mathcal{B} & \text{if } k \geq 2. \end{cases} \tag{6.1}$$

In light of Lemma 4.2, the entries of $\mathcal{M}^k_{Y,U}$ represent k-walks from input nodes to output nodes. In particular, we have that

$$(\mathcal{M}^k_{Y,U})_{ij} = \begin{cases} 0 & \text{if there is no } k\text{-walk from } u_j \text{ to } y_i, \\ * & \text{if there is a unique solid } k\text{-walk from } u_j \text{ to } y_i, \\ ? & \text{otherwise.} \end{cases} \tag{6.2}$$

We can use (6.1) and (6.2) to determine the entries of

$$\begin{bmatrix} \mathcal{D} & \mathcal{C}\mathcal{B} & \mathcal{C}\mathcal{A}\mathcal{B} & \cdots & \mathcal{C}\mathcal{A}^{n-1}\mathcal{B} \end{bmatrix}$$

directly from the system graph. With this in mind, consider the following *output color change rule* for the system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$.

Output color change rule.

Suppose that each output node of $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is colored either black or white. If there is a k-walk from u_j to exactly one white output node y_i, and that k-walk is unique and solid, then color y_i black. We say that u_j colors y_i black.

Suppose that initially all output nodes in $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ are colored white. The derived set $\mathscr{D}_{\text{occr}}$ is the set of all output nodes that are colored black after repeated application of the output color change rule until no more changes are possible. If $\mathscr{D}_{\text{occr}}$ contains all output nodes, then we say that $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is output colorable.

With the output color change rule defined, we obtain the following sufficient condition for strong structural output controllability

Theorem 6.2. *The system $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is strongly structurally output controllable if the system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is output colorable.*

Proof. We will first show that the pattern matrix

$$\mathcal{O}_k = \begin{bmatrix} \mathcal{D} & \mathcal{C}\mathcal{B} & \mathcal{C}\mathcal{A}\mathcal{B} & \dots & \mathcal{C}\mathcal{A}^{k-2}\mathcal{B} \end{bmatrix}$$

has full row rank for some appropriately chosen positive integer k. To avoid any confusion, we define $\mathcal{O}_1$ to be equal to $\mathcal{D}$. Note that relabeling the output nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ corresponds to permuting the rows of $\mathcal{D}$ and $\mathcal{C}$, which in turn corresponds to permuting the rows of $\mathcal{O}_k$. Since the rank of a pattern matrix does not change after a permutation of the rows, we can assume that $\mathcal{D}$ and $\mathcal{C}$ are permuted in such a way that the output nodes are colored black in order from y_p to y_1. As y_p is colored black first, there exist positive integers k_p, $j_p \leq m$ such that there is a unique solid k_p-walk from u_{j_p} to y_p, and there is no k_p-walk to any other output node. Using (6.2), this implies that

$$(\mathcal{M}^{k_p}_{Y,U})_{\bullet j_p} = \begin{bmatrix} 0 & \cdots & 0 & 0 & * \end{bmatrix}^\top.$$

Then, as y_{p-1} is colored next, there exist positive integers k_{p-1}, $j_{p-1} \leq m$ such that there is a unique solid k_{p-1}-walk from $u_{j_{p-1}}$ to y_{p-1}, and there is no k_{p-1}-walk to any other white output node. Since the only black output node is y_p, this means that

$$(\mathcal{M}^{k_{p-1}}_{Y,U})_{\bullet j_{p-1}} = \begin{bmatrix} 0 & \cdots & 0 & * & \otimes \end{bmatrix}^\top.$$

Repeating this argument, we find positive integers $k_1, \dots, k_p$ and $j_1, \dots, j_p$ such that

$$\begin{bmatrix} (\mathcal{M}^{k_1}_{Y,U})_{\bullet j_1} & \cdots & (\mathcal{M}^{k_p}_{Y,U})_{\bullet j_p} \end{bmatrix} = \begin{bmatrix} * & 0 & \cdots & 0 \\ \otimes & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \otimes & \otimes & \cdots & * \end{bmatrix}.$$

Define $k = \max\{k_1, \dots, k_p\}$. Using (6.1), we find that $(\mathcal{M}^{k_1}_{Y,U})_{\bullet j_1}, \dots, (\mathcal{M}^{k_p}_{Y,U})_{\bullet j_p}$ are columns of $\mathcal{O}_k$. Therefore, any permutation that puts them as the last p columns would make $\mathcal{O}_k$ take the form

$$\begin{bmatrix} \otimes & \cdots & \otimes & * & 0 & \cdots & 0 \\ \otimes & \cdots & \otimes & \otimes & * & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \otimes & \cdots & \otimes & \otimes & \otimes & \cdots & * \end{bmatrix}.$$

In other words, $\mathcal{O}_k$ is of Form III, hence, by Lemma 4.1, it has full row rank.

We can now show that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally output controllable. To this end, let $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. Since $\mathcal{O}_k$ has full row rank and $\mathcal{P}(\mathcal{C})\mathcal{P}(\mathcal{A})^i\mathcal{P}(\mathcal{B}) \subset \mathcal{P}(\mathcal{C}\mathcal{A}^i\mathcal{B})$ for all positive integers i, it follows that the matrix

$$O_k = \begin{bmatrix} D & CB & CAB & \cdots & CA^{k-2}B \end{bmatrix}$$

has full row rank. If $k \leq n+1$, then O_{n+1} certainly has full row rank too. Otherwise, if $k > n+1$, then O_{n+1} has full row rank because of the Cayley-Hamilton theorem. But O_{n+1} is given by

$$\begin{bmatrix} D & CB & CAB & \cdots & CA^{n-1}B \end{bmatrix},$$

hence, by Theorem 3.1, the system (A, B, C, D) is output controllable. Given that A, B, C and D were chosen arbitrarily, we conclude that the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally output controllable. □

The simple example below should illustrate the definition of the system graph as well as the output color change rule and its use in checking strong structural output controllability.

Example 6.1. Consider the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cc:c} 0 & 0 & * \\ * & 0 & * \\ \hdashline * & 0 & 0 \\ * & * & 0 \end{array}\right].$$

The corresponding system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is shown in Figure 6.1. Clearly, there is no 1-walk

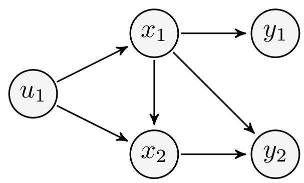


Figure 6.1: System graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

from u_1 to any output node. There is a 2-walk from u_1 to the output nodes y_1 and y_2, but both of them are white, thus non of them can be colored yet. However, there is a 3-walk from u_1 to y_2 only, and that 3-walk is unique and solid, so u_1 colors y_2 black (see Figure 6.2). Then, since

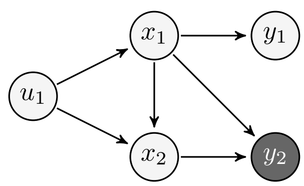


Figure 6.2: System graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ after u_1 has colored y_2 black.

the 2-walk from u_1 to y_1 is unique and solid, and y_1 is the only white output node, it follows that u_1 colors y_1 black (see Figure 6.3). As all output nodes are colored black, we conclude that

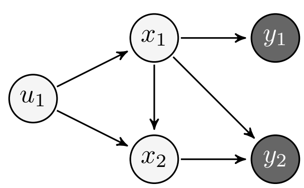


Figure 6.3: System graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ after u_1 has colored y_1 black.

$(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally output controllable.

We will end this section by briefly discussing the difficulties in obtaining necessary conditions. Firstly, the following example shows that the condition in Theorem 6.2 is not necessary.

Example 6.2. Consider the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{ccccc:c} * & 0 & 0 & 0 & 0 & * \\ * & 0 & 0 & 0 & * & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & 0 \\ \hdashline 0 & * & * & 0 & 0 & 0 \end{array}\right].$$

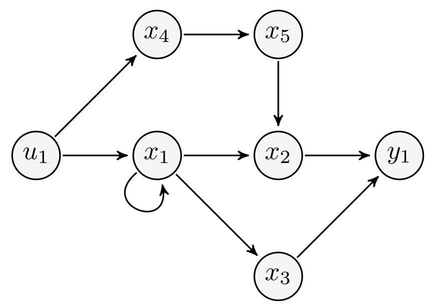


Figure 6.4: The system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$.

The corresponding system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is shown in Figure 6.4. One can check that $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is not output colorable. Indeed, there is no 1-walk or 2-walk from u_1 to y_1, and for every positive integer $k \geq 3$, there is a non-unique k-walk from u_1 to y_1, hence u_1 cannot color y_1 black. Nevertheless, the system $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is strongly structurally output controllable. To show this, let $A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, and write

$$\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right] = \left[\begin{array}{ccccc:c} a_{11} & 0 & 0 & 0 & 0 & b_1 \\ a_{21} & 0 & 0 & 0 & a_{25} & 0 \\ a_{31} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b_4 \\ 0 & 0 & 0 & a_{54} & 0 & 0 \\ \hdashline 0 & c_2 & c_3 & 0 & 0 & 0 \end{array}\right]$$

where $a_{11}, a_{21}, a_{31}, a_{54}, a_{25}, b_1, b_4, c_2$ and c_3 are all nonzero real numbers. As there is only one output, (A, B, C, D) would be output controllable if $CA^iB \neq 0$ for some non-negative integer i. With this in mind, we can compute

$$CAB = c_2a_{21}b_1 + c_3a_{31}b_1, \quad CA^2B = c_2a_{25}a_{54}b_4 + c_2a_{21}a_{11}b_1 + c_3a_{31}a_{11}b_1,$$

where we note that $CA^2B = c_2a_{25}a_{54}b_4 + a_{11}CAB$. This implies that $CA^2B \neq 0$ if $CAB = 0$, hence at least one of CAB an CA^2B is nonzero and the system (A, B, C, D) is output controllable. Since A, B, C and D were chosen arbitrarily, it follows that $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is strongly structurally output controllable.

This example illustrates that the condition in Theorem 6.2 is not necessary because the product of pattern matrices fails to capture dependencies between entries in the product of real matrices belonging to their pattern classes. In particular, the fact that the 3-walks and 4-walks from u_1 to y_1 in the system graph are solid but not unique indicates that CAB and CA^2B can be written as sums of nonzero numbers, which could possibly be zero. However, the symbolic operations with pattern matrices cannot capture the relationship between the values of CAB and CA^2B. Given that there is no real alternative to Theorem 3.1 for checking output controllability, it seems that the only way forward is in trying to develop techniques that avoid this flaw of multiplication of pattern matrices, or at least can capture the relevant dependencies.

Chapter 7

Input-state observability

In this chapter, we will investigate conditions under which the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally input-state observable. As it turns out, by borrowing ideas from [5], we can easily find necessary and sufficient conditions for strong structural input-state observability in the form of a rank condition on a pair of pattern matrices.

Theorem 7.1. *The system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally input-state observable if and only if the pattern matrices*

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} \bar{\mathcal{A}} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

both have full column rank, where $\bar{\mathcal{A}}$ is obtained from $\mathcal{A}$ by modifying the diagonal entries as

$$\bar{\mathcal{A}}_{kk} = \begin{cases} * & \textit{if } \mathcal{A}_{kk} = 0, \\ ? & \textit{otherwise.} \end{cases}$$

Proof. ($\Rightarrow$) Suppose that the matrix

$$\begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix}$$

has full column rank for all $\lambda \in \mathbb{C}$ for all $A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. Substituting $\lambda = 0$ implies that

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank. For the second pattern matrix, let $\bar{A} \in \mathcal{P}(\bar{\mathcal{A}}), B \in \mathcal{P}(\mathcal{B}), C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, and consider the equation

$$\begin{bmatrix} \bar{A} & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0.$$

We will show that the latter has only the trivial solution. To this end, let α be a nonzero real number such that

$$\alpha \notin \{\bar{A}_{kk} \mid k \text{ is such that } \mathcal{A}_{kk} = *\}.$$

In other words, α is such that $(\bar{A} - \alpha I)_{kk} \neq 0$ if $\mathcal{A}_{kk} = *$. Now define the diagonal matrix $\Delta \in \mathbb{R}^{n \times n}$ by

$$\Delta_{kk} = \begin{cases} \alpha / \bar{A}_{kk} & \text{if } \bar{\mathcal{A}}_{kk} = *, \\ 1 & \text{otherwise,} \end{cases}$$

which implies that

$$(\Delta\bar{A})_{kk} = \begin{cases} \alpha & \text{if } \mathcal{A}_{kk} = 0, \\ \bar{A}_{kk} & \text{otherwise,} \end{cases}$$

where we used that $\mathcal{A}_{kk} = 0$ if and only if $\bar{\mathcal{A}}_{kk} = *$. Then $(\Delta\bar{A} - \alpha I)_{kk} = 0$ if $\mathcal{A}_{kk} = 0$, and $(\Delta\bar{A} - \alpha I)_{kk} = (\bar{A} - \alpha I)_{kk} \neq 0$ if $\mathcal{A}_{kk} = *$. This, together with the fact that the pattern class of a pattern matrix is invariant under nonzero scaling of the rows and columns, implies that $\Delta\bar{A} - \alpha I \in \mathcal{P}(\mathcal{A})$. Similarly, $\Delta B \in \mathcal{P}(\mathcal{B})$, hence we can define $\hat{A} = \Delta\bar{A} - \alpha I$ and $\hat{B} = \Delta B$ to obtain

$$\begin{bmatrix} \Delta & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \bar{A} & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \quad \Leftrightarrow \quad \begin{bmatrix} \hat{A} + \alpha I & \hat{B} \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0,$$

where $\hat{A} \in \mathcal{P}(\bar{\mathcal{A}}), \hat{B} \in \mathcal{P}(\mathcal{B}), C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. The matrix on the right has full column rank by assumption, thus $x = 0$ and $y = 0$, which shows that

$$\begin{bmatrix} \bar{\mathcal{A}} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank as well.

($\Leftarrow$) Suppose that the pattern matrices

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \bar{\mathcal{A}} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix},$$

both have full column rank and consider the equation

$$\begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0,$$

where $A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ and $\lambda \in \mathbb{C}$. We would like to show that this equation has only the trivial solution. Note that this follows immediately for $\lambda = 0$ since the pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank. Suppose $\lambda \neq 0$ and let $x = a + bi$ and $y = c + di$, where i is the imaginary unit and a, b, c, d are real vectors. Furthermore, let α be a nonzero real number such that

$$\alpha \notin \left\{ -\frac{a_k}{b_k} \mid b_k \neq 0 \right\} \cup \left\{ -\frac{c_k}{d_k} \mid d_k \neq 0 \right\} \cup \left\{ -\frac{(Aa + Bc)_k}{(Ab + Bd)_k} \mid (Ab + Bd)_k \neq 0 \right\}$$

and define $\hat{x} = a + \alpha b$ and $\hat{y} = c + \alpha d$. The choice of α and the fact that $\lambda \neq 0$ imply that

(i) $\hat{x}_k = 0$ if and only if $x_k = 0$.

(ii) $\hat{y}_k = 0$ if and only if $y_k = 0$.

(iii) $\hat{x}_k = 0$ if and only if $(A\hat{x} + B\hat{y})_k = 0$.

Note that (i) and (ii) follow immediately from the choice of α. For the only if part of (iii), suppose that $\hat{x}_k = 0$, hence that $x_k = 0$ because of (i). Then we have $Ax + By = \lambda x$, which implies that $(Ax + By)_k = 0$, equivalently, that $(Aa + Bc)_k = (Ab + Bd)_k = 0$ and $(A\hat{x} + B\hat{y})_k = 0$. For the if part of (iii), suppose that $(A\hat{x} + B\hat{y})_k = 0$, which, by the choice of α, implies that $(Aa + Bc)_k = (Ab + Bd)_k = 0$, equivalently, that $(Ax + By)_k = 0$. But $(Ax + By)_k = \lambda x_k$ and $\lambda \neq 0$, hence $x_k = 0$ and we conclude that $\hat{x}_k = 0$ because of (i).

With these properties in mind, define the diagonal matrix Δ by

$$\Delta_{kk} = \begin{cases} 1 & \text{if } \hat{x}_k = 0, \\ \dfrac{(A\hat{x} + B\hat{y})_k}{\hat{x}_k} & \text{otherwise,} \end{cases}$$

which is nonsingular because of (iii). Moreover, the definition of Δ is such that $\Delta\hat{x} = A\hat{x} + B\hat{y}$, while $Cx + Dy = 0$ yields $Ca + Dc = Cb + Dd = 0$, thus $C\hat{x} + D\hat{y} = 0$. This allows us to write

$$\begin{bmatrix} A - \Delta & B \\ C & D \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = 0,$$

where $A - \Delta \in \mathcal{P}(\bar{\mathcal{A}})$ since Δ has no zeros on the diagonal. Given that $A - \Delta \in \mathcal{P}(\bar{\mathcal{A}}), B \in \mathcal{P}(\mathcal{B}), C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, we must have that $\hat{x} = 0$ and $\hat{y} = 0$, hence $x = 0$ and $y = 0$ because of (i) and (ii). This concludes the proof. $\square$

Recall that the color change rule can be used to check whether a pattern matrix has full row rank. Moreover, a pattern matrix has full column rank if and only if its transpose has full row rank. Therefore, we can check if the pattern matrices in Theorem 7.1 have full column rank by checking whether the graphs associated to their transposes are colorable. However, just like with strong structural output controllability, we can define a graph-theoretic procedure on the system graph to check whether the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally input-state observable. Although this is not strictly necessary, it does provide some intuition behind the condition of Theorem 7.1. To this end, we will first provide an alternative to the color change rule, which can be used to check whether a pattern matrix has full column rank. Since this corresponds to applying the color change rule to the transpose of the pattern matrix, we will refer to it as the *dual color change rule.*

Dual color change rule.

Let $\mathcal{M} \in \{0, *, ?\}^{q \times r}$ and consider the corresponding graph $\mathscr{G}(\mathcal{M})$. Suppose that each node of $\mathscr{G}(\mathcal{M})$ is colored either black or white. If node i has exactly one white in-neighbor j, and the edge from j to i is solid, then color j black. We say that node i colors node j black.

Suppose that initially all nodes are colored white. The derived set $\mathscr{D}_{\text{dccr}}$ is the set of all black nodes after repeated application of the dual color change rule until no more changes are possible. We say that the nodes $v_1, \ldots, v_k$ are *dual colorable* if they are contained in $\mathscr{D}_{\text{dccr}}$. For a pattern matrix $\mathcal{M} \in \{0, *, ?\}^{q \times r}$, we say that $\mathscr{G}(\mathcal{M})$ is *dual colorable* if $\mathscr{D}_{\text{dccr}}$ contains r nodes. Analogous to Lemma 5.1, we have the following result.

Lemma 7.1. *Let $\mathcal{M} \in \{0, *, ?\}^{q \times r}$ with $q \geq r$. Then $\mathcal{M}$ has full column rank if and only if $\mathscr{G}(\mathcal{M})$ is dual colorable.*

Proof. We know that $\mathcal{M}$ has full column rank if and only if $\mathcal{M}^\top$ has full row rank. Using Lemma 5.1, $\mathcal{M}^\top$ has full row rank if and only if $\mathscr{G}(\mathcal{M}^\top)$ is colorable. In other words, since $\mathcal{M}^\top \in \{0, *, ?\}^{r \times q}$ with $r \leq q$, $\mathcal{M}$ has full column rank if and only if $\mathscr{D}_{\text{ccr}}$ contains r nodes after repeated application of the color change rule to $\mathscr{G}(\mathcal{M}^\top)$. Note that the graphs $\mathscr{G}(\mathcal{M})$ and $\mathscr{G}(\mathcal{M}^\top)$ have the same number of vertices. Since $\mathcal{M}_{ij} = \mathcal{M}^\top_{ji}$, it follows that there is a solid edge from j to i in $\mathscr{G}(\mathcal{M})$ if and only if there is a solid edge from i to j in $\mathscr{G}(\mathcal{M}^\top)$. The same holds for dashed edges, hence the graph $\mathscr{G}(\mathcal{M}^\top)$ can be obtained from $\mathscr{G}(\mathcal{M})$ by flipping the orientation of all edges. Furthermore, the color change rule allows us to color white out-neighbors, while the dual color change rule allows us to color white in-neighbors. This means that applying the color change rule to $\mathscr{G}(\mathcal{M}^\top)$ is equivalent to applying the dual color change rule to $\mathscr{G}(\mathcal{M})$. Therefore, the pattern matrix $\mathcal{M}$ has full column rank if and only if $\mathscr{D}_{\text{dccr}}$ contains r nodes after repeated application of the dual color change rule on $\mathscr{G}(\mathcal{M})$, i.e., the graph $\mathscr{G}(\mathcal{M})$ is dual colorable. $\square$

With the dual color change rule defined, we can now provide the following graph-theoretic condition for strong structural input-state observability.

Theorem 7.2. *The system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally input-state observable if and only if the input and state nodes in the graphs $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ and $\mathscr{G}(\bar{\mathcal{A}}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable, where $\bar{\mathcal{A}}$ is obtained from $\mathcal{A}$ by modifying the diagonal entries as*

$$\bar{\mathcal{A}}_{kk} = \begin{cases} * & \text{if } \mathcal{A}_{kk} = 0, \\ ? & \text{otherwise.} \end{cases}$$

Proof. From Theorem 7.1 we know that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly-structurally input-state observable if and only if the pattern matrices

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \bar{\mathcal{A}} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

have full column rank. We will first prove that the pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank if and only if the input and state nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable. Note that the latter has full column rank if and only if the $(n+m+p) \times (n+m)$ pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ 0 & 0 \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank. Using Lemma 7.1, this is the case if and only if the graph

$$\mathscr{G}\left(\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ 0 & 0 \\ \mathcal{C} & \mathcal{D} \end{bmatrix}\right)$$

is dual colorable. We can recognize this as the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, hence the pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank if and only if $\mathscr{D}_{\text{dccr}}$ contains $n+m$ nodes after repeated application of the dual color change rule to $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ until no more changes are possible. Given that an output node in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is not an in-neighbor of any node, it follows that the output nodes cannot be colored black. This implies that the only way $\mathscr{D}_{\text{dccr}}$ contains $n+m$ nodes is if it contains all input and state nodes, i.e., the input and state nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable. Finally, using precisely the same arguments, we can show that the matrix

$$\begin{bmatrix} \bar{\mathcal{A}} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank if and only if the input and state nodes in $\mathscr{G}(\bar{\mathcal{A}}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable. $\square$

We finish this chapter with an example that illustrates how Theorem 7.2 can be used to check if a structured system is strongly structurally input-state observable.

Example 7.1. Consider the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cc:c} * & 0 & * \\ * & 0 & * \\ \hdashline * & 0 & 0 \\ * & * & 0 \end{array}\right]$$

To check for strong structural input-state observability, we need to check whether the input and state nodes in the system graphs $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ and $\mathscr{G}(\bar{\mathcal{A}}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable.

We start by considering the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, depicted in Figure 7.1, and the repeated application of the dual color change rule to it. Note that x_1 is the only white in-neighbor of y_1 and

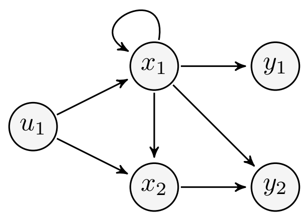
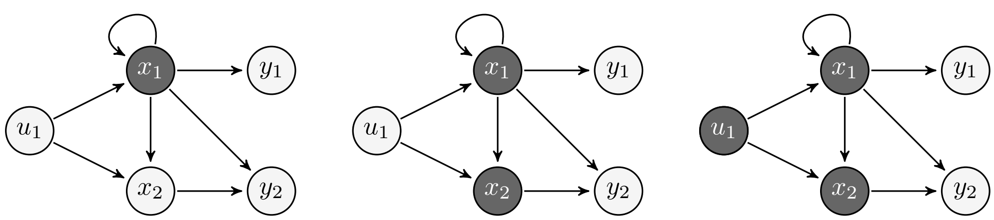


Figure 7.1: The system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

the edge connecting them is solid, hence y_1 colors x_1 black (see Figure 7.2a). With x_1 being black, the only white in-neighbor of y_2 is x_2. Given that the edge connecting them is solid, it follows that y_2 colors x_2 black (see Figure 7.2b). Finally, the only white in-neighbor of x_1 is u_1, and the edge connecting them is solid, hence x_1 colors u_1 black (see Figure 7.2c). As all inputs and states are colored, we conclude that the input and state nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable.



(a) y_1 colors x_1 black. (b) y_2 colors x_2 black. (c) x_2 colors u_1 black.

Figure 7.2: Repeated application of the dual color change rule for the graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

Next, we turn to the system graph $\mathscr{G}(\bar{\mathcal{A}}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, depicted in Figure 7.3, and the repeated application of the dual color change rule to it. As before, x_1 is the only white in-neighbor of y_1 and the edge connecting them is solid, hence y_1 colors x_1 black (see Figure 7.4a). Then x_2 is the only white in-neighbor of y_2 and the edge connecting them is solid, hence y_2 colors x_2 (see Figure 7.4b). Finally, u_1 is the only white in-neighbor of x_2 and the edge connecting them is solid, hence x_2 colors u_1 black (see Figure 7.4c). All input and state nodes are now colored black, so we conclude that the input and state nodes in $\mathscr{G}(\bar{\mathcal{A}}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable. Using Theorem 7.2, this means that the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally input-state observable.

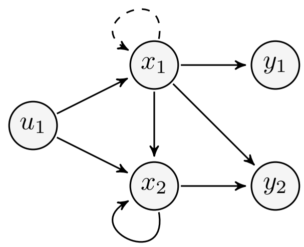
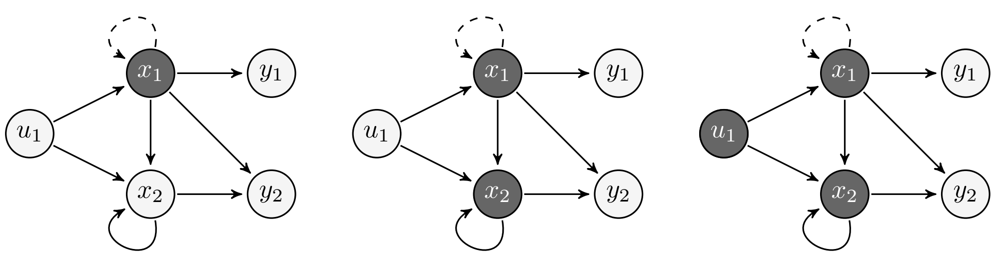


Figure 7.3: The system graph $\mathscr{G}(\bar{\mathcal{A}}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.



(a) y_1 colors x_1 black.

(b) y_2 colors x_2 black.

(c) x_2 colors u_1 black.

Figure 7.4: Repeated application of the dual color change rule on $\mathscr{G}(\bar{\mathcal{A}}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

Chapter 8

Left invertibility

In this chapter, we will investigate conditions under which $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible. Unfortunately, the minor difference in the conditions for left-invertiblity and input-state observability lead to difficulties when investigating necessary and sufficient conditions. In particular, the fact that the system matrix $P(\lambda)$ is allowed to drop rank for some $\lambda \in \mathbb{C}$, and that the values of λ for which this happens are not easy to characterize (see [22]), prevents us from using the same ideas as in the proof of Theorem 7.1 about strong structural input-state observability. Nevertheless, we note that Theorem 7.1 does provide a sufficient condition for strong structural left invertibility, although a rather crude one.

In light of this, we will look into three alternative approaches to obtain sufficient conditions, all of which are based on the result of Theorem 3.2. Moreover, we will provide an additional sufficient condition for strong structural left-invertiblity for the special case of a single-input single-output system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. Thereafter, we will summarize and compare the conditions with the help of a few examples. Finally, we will end with a short discussion on the difficulties in obtaining necessary conditions.

8.1 System matrix

The first approach is based on the system matrix condition in Theorem 3.2. It tells us that (A, B, C, D) is left invertible if and only if the system matrix $P(\lambda)$ has full column rank for all except finitely many $\lambda \in \mathbb{C}$. Therefore, $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if and only if the latter holds for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. The following theorem provides a sufficient condition for that.

Theorem 8.1. *The system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if at least one of the pattern matrices*

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \quad or \quad \begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank, where $\mathcal{A}^*$ is obtained from $\mathcal{A}$ by changing the diagonal entries to $*$'s.

Proof. Suppose that the pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank. This implies that the system matrix $P(\lambda)$ has full column rank for $\lambda = 0$ for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. Recall that $P(\lambda)$ has full column rank for all except finitely many $\lambda \in \mathbb{C}$ if an only if it has full column rank for some $\lambda \in \mathbb{C}$. Consequently, the system matrix $P(\lambda)$ has full column rank for all except finitely many $\lambda \in \mathbb{C}$ for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, hence the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible.

On the other hand, suppose that the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank, and let $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. Note that if $\lambda \neq A_{kk}$, then $(A - \lambda I)_{kk} \neq 0$ and $A - \lambda I \in \mathcal{P}(\mathcal{A}^*)$. This implies that the system matrix $P(\lambda)$ has full column rank for all $\lambda \in \mathbb{C} \setminus \{A_{kk} \mid k = 1, \ldots, n\}$, hence (A, B, C, D) is left invertible. Since A, B, C and D were chosen arbitrarily, it follows that (A, B, C, D) is left invertible for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, equivalently, the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible. □

Remark 8.1. *The condition in Theorem 8.1 is a weaker version of the condition in Theorem 7.1 about strong structural input-state observability. It is obtained after recognizing the fact that the system matrix $P(\lambda)$ is now allowed to drop rank for some $\lambda \in \mathbb{C}$. In particular, we recognize that it is sufficient for only one of the matrices in Theorem 7.1 to have full column rank. Furthermore, the pattern class of the pattern matrix*

$$\begin{bmatrix} \bar{\mathcal{A}} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

contains the pattern class of the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}.$$

Indeed, the only difference between these pattern matrices is in the entries on the diagonals of $\bar{\mathcal{A}}$ and $\mathcal{A}^*$, where $\bar{\mathcal{A}}$ has ?'s instead of $*$'s in some positions. Therefore, if the pattern matrix with $\bar{\mathcal{A}}$ has full column rank, then the pattern matrix with $\mathcal{A}^*$ would also have full column rank. Since the latter is already sufficient, there is no reason to look at the former.

The following couple of examples illustrate how the condition in Theorem 8.1 can be used to determine if a system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly-structurally left invertible. They also show that the full column rank condition might be satisfied for only one of the pattern matrices in Theorem 8.1.

Example 8.1. Consider the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cc:c} 0 & * & * \\ * & 0 & 0 \\ \hdashline * & 0 & * \end{array}\right].$$

We can easily see that the pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} 0 & * & * \\ * & 0 & 0 \\ * & 0 & * \end{bmatrix}$$

has full column rank, hence the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible. However, the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & 0 \\ * & 0 & * \end{bmatrix}$$

does not have full column rank since the matrix

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

is contained in its pattern class and does not have full column rank.

Example 8.2. Consider the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cc:c} * & 0 & 0 \\ * & 0 & * \\ \hdashline * & 0 & * \end{array}\right].$$

We can easily see that the pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} * & 0 & 0 \\ * & 0 & * \\ * & 0 & * \end{bmatrix}$$

does not have full column rank. However, the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} * & 0 & 0 \\ * & * & * \\ * & 0 & * \end{bmatrix}$$

clearly does have full column rank, hence the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible.

Next, we provide an example of a strongly structurally left invertible system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ for which the condition of Theorem 8.1 does not hold.

Example 8.3. Consider the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cc:c} * & * & * \\ * & 0 & 0 \\ \hdashline 0 & * & * \end{array}\right].$$

The pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

does not have full column rank since the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

belongs to its pattern class and does not have full column rank. Similarly, the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

does not have full column rank since the matrix

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

belongs to its pattern class and does not have full column rank. This implies that the condition of Theorem 8.1 does not hold. Nevertheless, it can be shown that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible. To this end, let $A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ and write

$$\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right] = \left[\begin{array}{cc:c} a_{11} & a_{12} & b_1 \\ a_{21} & 0 & 0 \\ \hdashline 0 & c_2 & d \end{array}\right],$$

where $a_{11}, a_{12}, a_{21}, b_1, c_2$ and d are nonzero real numbers. Then we can compute

$$\det P(\lambda) = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & b_1 \\ a_{21} & -\lambda & 0 \\ 0 & c_2 & d \end{bmatrix} = d\lambda^2 - da_{11}\lambda - da_{12}a_{21} + c_2 a_{21} b_1.$$

Since $\det P(\lambda)$ cannot be the zero polynomial, it follows that $\det P(\lambda) = 0$ only when λ is one of the two roots of the polynomial. This implies that the system matrix $P(\lambda)$ has full column rank for all except finitely many $\lambda \in \mathbb{C}$, hence (A, B, C, D) is left invertible. Given that A, B, C and D were chosen arbitrarily, it follows that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible.

We will end this section by providing a graph-theoretic condition for strong structural left invertibility of $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. Just like with strong structural input-state observability, we can check whether the pattern matrices in Theorem 8.1 have full column rank by checking whether the graphs associated to their transposes are colorable. However, we have also introduced a graph-theoretic procedure on the system graph to check whether matrices of such form have full column rank.

Theorem 8.2. *The system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if the input and state nodes in at least one of the graphs $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ or $\mathscr{G}(\mathcal{A}^*, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is dual colorable, where $\mathcal{A}^*$ is obtained from $\mathcal{A}$ by changing the diagonal entries to $*$'s.*

Proof. From Theorem 8.1 we know that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly-structurally left invertible if at least one of the pattern matrices

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank. Using the same arguments as in the proof of Theorem 7.2, we can show that the pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank if and only if the input and state nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable. Similarly, the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank if and only if the input and state nodes in $\mathscr{G}(\mathcal{A}^*, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable. $\square$

8.2 Transfer function

The second approach is based on the transfer function condition of Theorem 3.2. In particular, we can show that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if we show that the transfer matrix $T(s)$ is left invertible for all $A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. To this end, let $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, and note that the inverse of $sI - A$ can be written as the infinite sum

$$(sI - A)^{-1} = \sum_{i=1}^{\infty} A^{i-1} s^{-i},$$

hence the transfer matrix of (A, B, C, D) can be written as

$$T(s) = D + C \left(\sum_{i=1}^{\infty} A^{i-1} s^{-i} \right) B = D + CBs^{-1} + CABs^{-2} + \cdots .$$

If $(D)_{\bullet k} = 0$, then

$$\begin{aligned} s\,(T(s))_{\bullet k} &= s \left(CBs^{-1} + CABs^{-2} + CA^2Bs^{-1} + \cdots \right)_{\bullet k} \\ &= \left(CB + CA(sI - A)^{-1} B \right)_{\bullet k} , \end{aligned}$$

thus the k-th column of $sT(s)$ contains only proper rational functions of s. If, in addition, we have that $(CB)_{\bullet k} = 0$, then the k-th column of $s^2 T(s)$ contains only proper rational functions of s.

With this in mind, let $d_1, \ldots, d_m$, be positive integers such that $d_k = 1$ if $(D)_{\bullet k} \neq 0$, otherwise, $d_k \geq 2$ is the smallest integer for which $CA^{d_k - 2}B$ is nonzero. In other words, d_k is such that the k-th column of $s^{d_k - 1} T(s)$ contains only proper rational functions of s. Then we can write

$$T(s) = \left(T_0 + T_1 s^{-1} + T_2 s^{-2} + \cdots \right) \begin{bmatrix} s^{-d_1+1} & & \\ & \ddots & \\ & & s^{-d_m+1} \end{bmatrix},$$

where $(T_0 + T_1 s^{-1} + T_2 s^{-2} + \cdots)$ is the Laurent series of a proper rational matrix. Furthermore, we have that

$$(T_0)_{\bullet k} = \begin{cases} (D)_{\bullet k} & \text{if } \delta_k = 1, \\ (CA^{\delta_k - 2}B)_{\bullet k} & \text{if } \delta_k \geq 2, \end{cases}$$

for all $k \in \{1, \ldots, m\}$.

Lemma 8.1. *The transfer matrix $T(s)$ has full column rank if T_0 has full column rank.*

Proof. Suppose that there exists a nonzero rational vector $u(s)$ such that $T(s)u(s) = 0$. This is possible if and only if there exists a nonzero rational vector $v(s)$ such that

$$\left(T_0 + T_1 s^{-1} + T_2 s^{-2} + \cdots \right) v(s) = 0. \tag{8.1}$$

We can make $v(s)$ strictly proper by dividing both sides by s^k for a large enough integer k, hence we can write $v(s)$ in terms of its Laurent series as

$$v(s) = v_0 + v_1 s^{-1} + v_2 s^{-2} + \cdots ,$$

where $v_1, v_2, \ldots,$ are real numbers. The goal is to show that $v_k = 0$ for all non-negative integers k. With this in mind, note that (8.1) can be expanded to

$$T_0 v_0 + (T_0 v_1 + T_1 v_0)\, s^{-1} + (T_0 v_2 + T_1 v_1 + T_2 v_0)\, s^{-2} + \cdots = 0,$$

equivalently,

$$\sum_{k=0}^{\infty} \left(\sum_{i=0}^{k} T_i v_{k-i} \right) s^{-k} = 0,$$

which is satisfied if and only if

$$\sum_{i=0}^{k} T_i v_{k-i} = 0 \tag{8.2}$$

for all non-negative integers k. We will show that this implies that $v_k = 0$ for all non-negative integers k by using strong induction. Firstly, from (8.2) we know that $T_0 v_0 = 0$, hence $v_0 = 0$ because T_0 has full column rank. Now suppose that $v_i = 0$ for all $i \in \{0, \ldots, k\}$ and some non-negative integer k, and consider v_{k+1}. From (8.2) we know that

$$\sum_{i=0}^{k+1} T_i v_{k+1-i} = 0 \qquad \Longrightarrow \qquad T_0 v_{k+1} = 0,$$

hence $v_{k+1} = 0$ as well. Therefore, $v_k = 0$ for all non-negative integers k, which implies that $v(s) = 0$ and we conclude that $T(s)$ has full column rank. □

Using this result, we can derive the following sufficient condition for strong structural left invertibility.

Theorem 8.3. *The system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if the pattern matrix $\mathcal{T}_0 \in \{0, *, ?\}^{p \times m}$ has full column rank, where*

$$(\mathcal{T}_0)_{\bullet k} = \begin{cases} (\mathcal{D})_{\bullet k} & \text{if } \delta_k = 1, \\ (\mathcal{C}\mathcal{A}^{\delta_k - 2}\mathcal{B})_{\bullet k} & \text{if } \delta_k \geq 2, \end{cases}$$

for all $k \in \{1, \ldots, m\}$, and $\delta_k = 1$ if $(\mathcal{D})_{\bullet k} \neq 0$, otherwise, $\delta_k \geq 2$ is the smallest integer for which $(\mathcal{C}\mathcal{A}^{\delta_k - 2}\mathcal{B})_{\bullet k} \neq 0$.

Proof. Let $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. The integers $\delta_1, \ldots, \delta_m$ are defined in such a way that we can write

$$T(s) = \left(T_0 + T_1 s^{-1} + \cdots\right) \begin{bmatrix} s^{-\delta_1 + 1} & & \\ & \ddots & \\ & & s^{-\delta_m + 1} \end{bmatrix},$$

where $(T_0 + T_1 s^{-1} + T_2 s^{-2} + \cdots)$ is the Laurent series of a proper rational matrix, and

$$(T_0)_{\bullet k} = \begin{cases} (D)_{\bullet k} & \text{if } \delta_k = 1, \\ \left(CA^{\delta_k - 2}B\right)_{\bullet k} & \text{if } \delta_k \geq 2. \end{cases}$$

Using Lemma 8.1, the transfer matrix $T(s)$ would be left invertible for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ if T_0 has full column rank for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. With this in mind, define $\mathcal{T}_0$ by

$$(\mathcal{T}_0)_{\bullet k} = \begin{cases} (\mathcal{D})_{\bullet k} & \text{if } \delta_k = 1, \\ (\mathcal{C}\mathcal{A}^{\delta_k - 2}\mathcal{B})_{\bullet k} & \text{if } \delta_k \geq 2. \end{cases}$$

We know that $\mathcal{P}(\mathcal{C})\mathcal{P}(\mathcal{A})^n\mathcal{P}(\mathcal{B}) \subset \mathcal{P}(\mathcal{C}\mathcal{A}^n\mathcal{B})$, hence $T_0 \in \mathcal{P}(\mathcal{T}_0)$. Therefore, T_0 has full column rank for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ if the pattern matrix $\mathcal{T}_0$ has full column rank. □

We can check whether the pattern matrix $\mathcal{T}_0$ has full column rank by checking whether $\mathscr{G}(\mathcal{T}_0^\top)$ is colorable. This requires the computation of $\delta_1, \ldots, \delta_m$ and then the computation of $\mathcal{T}_0$. As an alternative, we can provide a graph-theoretic procedure for the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ to check whether the condition of Theorem 8.3 is satisfied. To this end, recall that the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is the graph associated to the matrix

$$\mathcal{M} = \begin{bmatrix} \mathcal{A} & \mathcal{B} & 0 \\ 0 & 0 & 0 \\ \mathcal{C} & \mathcal{D} & 0 \end{bmatrix},$$

and that

$$\mathcal{M}^k = \begin{bmatrix} \mathcal{A}^k & \mathcal{A}^{k-1}\mathcal{B} & 0 \\ 0 & 0 & 0 \\ \mathcal{C}\mathcal{A}^{k-1} & \mathcal{C}\mathcal{A}^{k-2}\mathcal{B} & 0 \end{bmatrix}.$$

Therefore δ_k is the smallest positive integer for which $(\mathcal{M}^{\delta_k}_{Y,U})_{\bullet k} \neq 0$, where, just like before, $Y = \{n+m+1, \dots, n+m+p\}$ and $U = \{n+1, \dots, n+m\}$ are the sets of output nodes and input nodes, respectively. From Lemma 4.2 we know that the entries of $\mathcal{M}^{\delta_k}_{Y,U}$ are related to the existence and type of δ_k-walks from input nodes to output nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. In particular, $(\mathcal{M}^{\delta_k}_{Y,U})_{\bullet k} \neq 0$ if and only if there is a δ_k-walk from u_k to the one of the output nodes. Consequently, δ_k is the distance from u_k to the set of output nodes, and

$$(\mathcal{T}_0)_{ij} = \begin{cases} 0 & \text{if there is no } \delta_j\text{-walk from } u_j \text{ to } y_i, \\ * & \text{if there is a unique solid } \delta_j\text{-walk from } u_j \text{ to } y_i, \\ ? & \text{otherwise.} \end{cases} \tag{8.3}$$

Note that δ_j being the distance from u_j to the set of output nodes implies that the δ_j-walks in (8.3) are actually paths of length δ_j. Assuming that $\delta_1, \dots, \delta_m$ are known, we can define the following *input color change rule* for the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

Input color change rule.

Suppose that each input node of $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is colored either black or white. If there is a δ_j-walk to y_i from exactly one white input node u_j, and the δ_j-walk from u_j to y_i is unique and solid, then change the color of u_j to black. We say that y_i colors u_j black.

Suppose that initially all input nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are colored white. The derived set $\mathscr{D}_{\text{iccr}}$ is the set of all input nodes that are colored black after repeated application of the input color change rule. If $\mathscr{D}_{\text{iccr}}$ contains all input nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, then we say that the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is input colorable.

Theorem 8.4. *The system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is input colorable.*

Proof. To begin with, note that relabeling the input nodes and output nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ corresponds to permuting the columns of $\mathcal{B}$ and $\mathcal{D}$, and the rows $\mathcal{C}$ and $\mathcal{D}$, respectively. For a system (A, B, C, D) with $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, this corresponds to permuting columns and rows of the transfer matrix $T(s)$. Since the left invertibility of $T(s)$ is not affected by such a permutation, it follows that the left invertibility of (A, B, C, D) is not affected either. Therefore, we can assume that the input nodes and output nodes are labeled in such a way that $u_1, \dots, u_m$ are colored in that order, and u_i is colored by y_i for all $i \in \{1, \dots, m\}$. We claim that $\mathcal{T}_0$ has the form

$$\mathcal{T}_0 = \begin{bmatrix} * & 0 & \cdots & 0 \\ \otimes & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \otimes & \otimes & \cdots & * \\ \otimes & \otimes & \cdots & \otimes \\ \vdots & \vdots & \ddots & \vdots \\ \otimes & \otimes & \cdots & \otimes \end{bmatrix}, \tag{8.4}$$

To show this, note that the input node u_1 is colored by y_1 while all input nodes are white. This means that there is a unique solid δ_1-walk from u_1 to y_1 and there is no δ_i-walk from u_i to y_1 for all $i \in \{2, \dots, m\}$. In view of (8.3), it follows that $\mathcal{T}_{11} = *$ and $\mathcal{T}_{1i} = 0$ for all $i \in \{2, \dots, m\}$, equivalently,

$$(\mathcal{T}_0)_{1\bullet} = \begin{bmatrix} * & 0 & \cdots & 0 \end{bmatrix}.$$

Then the input node u_2 is colored by y_2 while the only black input node is u_1. This means that there is unique solid δ_2-walk from u_2 to y_2 and there is no δ_i-walk from u_i to y_2 for all $i \in \{3, \dots, m\}$. Consequently, we obtain

$$(\mathcal{T}_0)_{2\bullet} = \begin{bmatrix} \otimes & * & \cdots & 0 \end{bmatrix}.$$

Proceeding in the same fashion, we find that $(\mathcal{T}_0)_{ii} = *$ and $(\mathcal{T}_0)_{ij} = 0$ for all $i \in \{1, \dots, m\}$ and $j \in \{i+1, \dots, m\}$. This shows that $\mathcal{T}_0$ has the form in (8.4), that is, it contains a triangular submatrix with $*$'s on the diagonal. Then it is easy to see that $\mathcal{T}_0^\top$ is of Form III, hence it has full row rank, equivalently, $\mathcal{T}_0$ has full column rank. Using Theorem 8.3, we conclude that the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible. $\square$

We finish this subsection with an example that demonstrates the input color change rule and its use in determining whether the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible.

Example 8.4. Consider the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cccc:ccc} 0 & ? & 0 & 0 & * & * & 0 \\ * & 0 & * & 0 & 0 & * & 0 \\ 0 & ? & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & * \\ \hdashline * & ? & 0 & 0 & 0 & 0 & 0 \\ 0 & * & ? & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 & 0 \end{array}\right].$$

The corresponding system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is depicted in Figure 8.1. We can show that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is left invertible by showing that the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is input colorable. To this end, note that the distance from u_1 to the output nodes is 2, the distance from u_2 to the

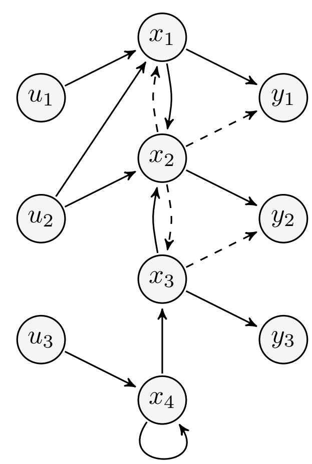


Figure 8.1: The system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

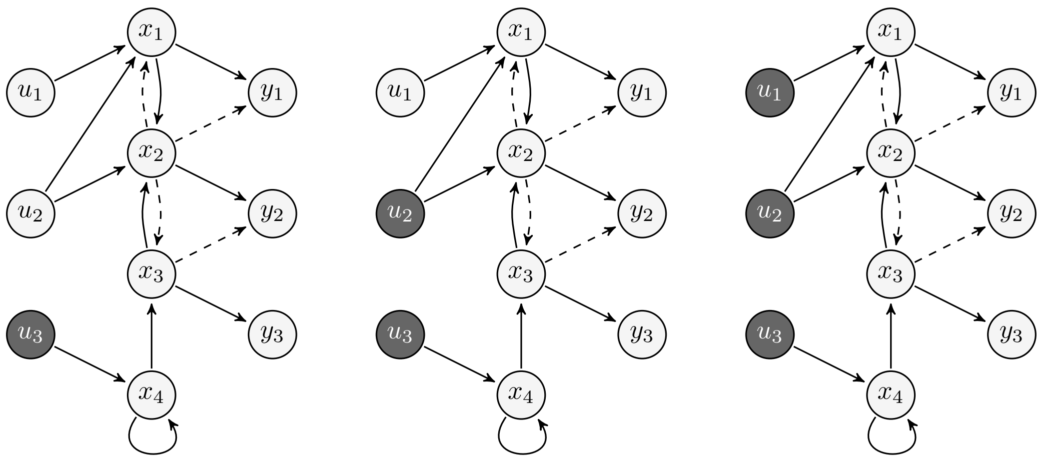
output nodes is also 2, and the distance from u_3 to the output nodes is 3. Therefore, we have that $\delta_1 = 2, \delta_2 = 2$ and $\delta_3 = 3$. There is a unique solid δ_1-walk from u_1 to y_1, but there is also a δ_2-walk from u_2 to y_1, thus y_1 cannot color u_1 black. Similarly, there is a unique solid δ_2-walk from u_2 to y_2, but there is also a δ_3-walk from u_3 to y_2, thus y_2 cannot color u_2 black. However, there is a unique solid δ_3-walk from u_3 to y_3, and there is no δ_i-walk from u_i to y_3 for any $i \in \{1, 2\}$, hence y_3 colors u_3 black. Now that u_3 is black, the δ_3-walk from u_3 to y_2 is irrelevant, and since there is a unique solid δ_2-walk from u_2 to y_2, it follows that y_2 colors u_2 black. Similarly, now that u_2 is black, the δ_2-walk from u_2 to y_1 is irrelevant, and since there is a unique solid δ_1-walk from u_1

to y_1, it follows that y_1 colors u_1 black. Given that all input nodes are colored, we conclude that the graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is input colorable and the system $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is strongly structurally left invertible.

Alternatively, we can use the graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ to find the entries of $\mathcal{T}_0$. For example, the distance from u_2 to the output nodes is 2, there is a non-solid 2-walk from u_2 to y_1, and a unique solid 2-walk from u_2 to y_2. Therefore, the second column of $\mathcal{T}_0$ is given by $\begin{bmatrix} ? & * & 0 \end{bmatrix}^\top$. Proceeding in the same manner, we find that

$$\mathcal{T}_0 = \begin{bmatrix} * & ? & 0 \\ 0 & * & ? \\ 0 & 0 & * \end{bmatrix},$$

which obviously has full column rank, hence the system $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is strongly structurally left invertible.



(a) y_3 colors u_3 black. (b) y_2 colors u_2 black. (c) y_1 colors u_1 black.

Figure 8.2: Repeated application of the input color change rule for the system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$.

8.3 Graph simplification process

The third approach relies on the transfer matrix condition of Theorem 3.2 and the intermediate results obtained in [23]. In order to show the relevance of these results, we will first provide an alternative way to express the transfer matrix $T(s)$. Let $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$, and let M be the $(n+m+p)\times(n+m+p)$ matrix defined by

$$M = \begin{bmatrix} A & B & 0 \\ 0 & 0 & 0 \\ C & D & 0 \end{bmatrix}.$$

The matrix M is such that

$$\left(I - Ms^{-1}\right)^{-1} = \begin{bmatrix} s(sI-A)^{-1} & (sI-A)^{-1}B & 0 \\ 0 & I & 0 \\ C(sI-A)^{-1} & s^{-1}T(s) & I \end{bmatrix}.$$

Here we see that $s^{-1}T(s)$ is a submatrix of the inverse of $I - Ms^{-1}$. In particular, if we define $U = \{n+1, \ldots, n+m\}$ and $Y = \{n+m+1, \ldots, n+m+p\}$, then

$$\left(I - Ms^{-1}\right)^{-1}_{Y,U} = s^{-1}T(s).$$

Therefore $T(s)$ is left invertible if and only if

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = m,$$

hence $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is left invertible if the latter holds for all $M \in \mathcal{P}(\mathcal{M})$, where $\mathcal{M}$ is defined as before, that is,

$$\mathcal{M} = \begin{bmatrix} \mathcal{A} & \mathcal{B} & 0 \\ 0 & 0 & 0 \\ \mathcal{C} & \mathcal{D} & 0 \end{bmatrix}.$$

This closely resembles some of the intermediate results obtained in [23]. To see this, consider the pattern matrix $\bar{\mathcal{M}} \in \{0,*\}^{r\times r}$ with zeros on the diagonal, and let $\mathcal{P}_{\mathrm{R}}(\bar{\mathcal{M}})$ be the collection of rational matrices $\bar{M}(s) \in \mathbb{R}(s)^{r\times r}$ such that

P1. $\bar{M}(s)$ is proper.

P2. $\left(\bar{M}(s)\right)_{ij} \neq 0$ if and only if $\bar{\bar{\mathcal{M}}}_{ij} = *$.

P3. Every principal minor of $\lim_{s\to\infty}\left(I - \bar{M}(s)\right)$ is nonzero.

The results in [23] provide a graph-theoretic necessary and sufficient condition for a submatrix of $(I - \bar{M}(s))^{-1}$ to have full column rank for all $\bar{M}(s) \in \mathcal{P}_{\mathrm{R}}(\bar{\mathcal{M}})$. There are two key differences between the context of these results and the context of this thesis. Firstly, we consider $\mathcal{M} \in \{0,*,?\}^{(n+m+p)\times(n+m+p)}$, without any restrictions on the diagonal entries. Secondly, we do not consider all rational matrices $M(s) \in \mathcal{P}_{\mathrm{R}}(\mathcal{M})$ but only ones of the form Ms^{-1} with $M \in \mathcal{P}(\mathcal{M})$. Unfortunately, the latter makes the necessary and sufficient conditions from [23] only sufficient. Nevertheless, we can still use the ideas developed there and adopt the graph-theoretic condition for the purpose of checking for strong structural left invertibility.

The following lemmas are straightforward modifications of the lemmas in [23] to fit the context of this thesis. Since these hold for an arbitrary square pattern matrix $\mathcal{M}$, we will briefly forget about the particular definition of $\mathcal{M}$ introduced in this section. With this in mind, let $\mathcal{M}$ be a square pattern matrix and let $\mathscr{G}(\mathcal{M}) = (V, E_* \cup E_?)$ be the graph associated to it.

Lemma 8.2. *Let $U, Y \subset V$. Moreover, let $\bar{\mathcal{M}}$ be the pattern matrix obtained from $\mathcal{M}$ by setting $\mathcal{M}_{V,Y} = 0$, i.e., the graph $\mathscr{G}(\bar{\mathcal{M}})$ is obtained from $\mathscr{G}(\mathcal{M})$ by removing all outgoing edges of the nodes in Y. Then*

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } M \in \mathcal{P}(\mathcal{M})$$

if and only if

$$\operatorname{rank}\left(I - \bar{M}s^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } \bar{M} \in \mathcal{P}(\bar{\mathcal{M}}).$$

Proof. Consider $M \in \mathcal{P}(\mathcal{M})$ and let P be a permutation matrix such that

$$PMP^\top = \begin{bmatrix} M_{R,R} & M_{R,Y} \\ M_{Y,R} & M_{Y,Y} \end{bmatrix},$$

where $R = V \setminus Y$. In other words, P shifts the rows corresponding to Y to the end. Now define $\bar{M}$ as

$$\bar{M} = P^\top \begin{bmatrix} M_{R,R} & 0 \\ M_{Y,R} & 0 \end{bmatrix} P \quad \implies \quad P\bar{M}P^\top = \begin{bmatrix} M_{R,R} & 0 \\ M_{Y,R} & 0 \end{bmatrix}, \tag{8.5}$$

and note that $\bar{M} \in \mathcal{P}(\bar{\mathcal{M}})$. By Proposition 2.8.7 of [24], the inverse of $I - PMP^\top s^{-1}$ can be written as

$$\left(I - PMP^\top s^{-1}\right)^{-1} = \begin{bmatrix} \bullet & \bullet \\ SM_{Y,R}s^{-1}\left(I - M_{R,R}s^{-1}\right)^{-1} & S \end{bmatrix},$$

where S is the inverse of the Schur complement of the top left block of $PMP^\top s^{-1}$, namely,

$$S = \left(I - M_{Y,Y}s^{-1} - M_{Y,R}s^{-1}\left(I - M_{R,R}s^{-1}\right)^{-1} M_{R,Y}s^{-1}\right)^{-1}.$$

By the same proposition, we also have that

$$\left(I - P\bar{M}P^\top s^{-1}\right)^{-1} = \begin{bmatrix} \bullet & \bullet \\ M_{Y,R}s^{-1}\left(I - M_{R,R}s^{-1}\right)^{-1} & I \end{bmatrix},$$

which shows that the second block row of $\left(I - PGP^\top s^{-1}\right)^{-1}$ is S times the second block row of $\left(I - P\bar{M}P^\top s^{-1}\right)^{-1}$. Moreover,

$$\begin{aligned}\left(I - PMP^\top s^{-1}\right)^{-1} &= P\left(I - Ms^{-1}\right)^{-1}P^\top,\\ \left(I - P\bar{M}P^\top s^{-1}\right)^{-1} &= P\left(I - \bar{M}s^{-1}\right)^{-1}P^\top,\end{aligned}$$

hence, by definition of P,

$$\left(I - Ms^{-1}\right)^{-1}_{Y,V} = S\left(I - Ms^{-1}\right)^{-1}_{Y,V}.$$

As S is invertible, this implies that

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = \operatorname{rank}\left(I - \bar{M}s^{-1}\right)^{-1}_{Y,U}, \tag{8.6}$$

which will be used to prove the lemma.

($\Leftarrow$) Suppose that

$$\operatorname{rank}\left(I - \bar{M}s^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } \bar{M} \in \mathcal{P}(\bar{\mathcal{M}}),$$

and let $M \in \mathcal{P}(\mathcal{M})$. Using M, define the matrix $\bar{M}$ as in (8.5). Since $\bar{M} \in \mathcal{P}(\bar{\mathcal{M}})$, it follows from (8.6) that

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U|,$$

and we conclude that

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } M \in \mathcal{P}(\mathcal{M}).$$

($\Rightarrow$) Suppose that

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } G \in \mathcal{P}(\mathcal{M}),$$

and let $\bar{M} \in \mathcal{P}(\bar{\mathcal{M}})$. Just like before, if P is the permutation matrix that shifts the rows corresponding to Y to the end, then

$$P\bar{M}P^\top = \begin{bmatrix} \bar{M}_{R,R} & 0 \\ \bar{M}_{Y,R} & 0 \end{bmatrix}.$$

Now, define M in terms of its submatrices $M_{Y,Y}, M_{Y,R}, M_{R,Y}, M_{R,R}$, with

$$M_{Y,R} = \bar{M}_{Y,R}, \qquad M_{R,R} = \bar{M}_{R,R},$$

and $M_{Y,Y}, M_{R,Y}$ such that $M \in \mathcal{P}(\mathcal{M})$. With M defined in this manner, we have that (8.5) holds and we can use (8.6) to conclude that

$$\operatorname{rank}\left(I - \bar{M}s^{-1}\right)^{-1}_{Y,U} = |U|,$$

and consequently that

$$\operatorname{rank}\left(I - \bar{M}s^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } \bar{M} \in \mathcal{P}(\bar{\mathcal{M}}).$$

$\square$

The result of this lemma will allow us to modify the pattern matrix $\mathcal{M}$, or equivalently, the graph $\mathscr{G}(\mathcal{M})$, without affecting the full column rank property of the submatrix $(I-Ms^{-1})^{-1}_{Y,U}$. Similarly, the next lemma will allow us to modify the set Y without changing the rank of the submatrix $(I-Ms^{-1})^{-1}_{Y,U}$.

Lemma 8.3. *Let $U, Y \subset V$. Suppose that a node $k \in Y \setminus U$ has exactly one in-neighbor j that is reachable from U. If the edge from k to j is solid, then*

$$\operatorname{rank}\left(I-Ms^{-1}\right)^{-1}_{Y,U} = \operatorname{rank}\left(I-Ms^{-1}\right)^{-1}_{\bar{Y},U} \text{ for all } M \in \mathcal{P}(\mathcal{M}),$$

where $\bar{Y} = (Y \setminus \{k\}) \cup \{j\}$.

Proof. As a consequence of Lemma 8.2, we can assume that $\mathcal{M}_{V,Y} = 0$. Let $M \in \mathcal{M}$ and $v \in U$, and note that

$$\left(I-Ms^{-1}\right)\left(I-Ms^{-1}\right)^{-1} = I \quad \Longrightarrow \quad \sum_{i=1}^{|V|}\left(I-Ms^{-1}\right)_{ki}\left(\left(I-Ms^{-1}\right)^{-1}\right)_{iv} = I_{kv} = 0,$$

where $I_{kv} = 0$ since $k \in Y \setminus U$ and $v \in U$ are distinct. Similarly, $I_{ki} = 0$ for $k \neq i$, and $M_{kk} = 0$ since $k \in Y$ and $M_{V,Y} = 0$. Then it follows that

$$\left(I-Ms^{-1}\right)^{-1}_{kv} = \sum_{i \in N_k^-} M_{ki}s^{-1}\left(I-Ms^{-1}\right)^{-1}_{iv}, \tag{8.7}$$

where N_k^- is the in-neighborhood of node k. Recall that

$$\left(I-Ms^{-1}\right)^{-1} = \sum_{l=0}^{\infty} M^l s^{-l},$$

and that $\mathcal{M}^l_{iv} = 0$ if there is no l-walk from v to i. As $v \in U$ and the only neighbor of k that is reachable from U is j, it follows that

$$\left(M^l\right)_{iv} = 0 \text{ for all } i \in N_k^- \setminus \{j\},\ l \geq 0.$$

Therefore, we can write

$$\left(I-Ms^{-1}\right)^{-1}_{iv} = 0 \text{ for all } i \in N_k^- \setminus \{j\},$$

which reduces (8.7) to

$$\left(I-Ms^{-1}\right)^{-1}_{kv} = M_{kj}s^{-1}\left(I-Ms^{-1}\right)^{-1}_{jv}.$$

Given that $v \in U$ was arbitrary, we obtain

$$\left(I-Ms^{-1}\right)^{-1}_{k,U} = M_{kj}s^{-1}\left(I-Ms^{-1}\right)^{-1}_{j,U}.$$

Furthermore, the edge from j to k is solid, hence $M_{kj} \neq 0$ and we can replace the column $\left(I-Ms^{-1}\right)^{-1}_{k,U}$ of $\left(I-Ms^{-1}\right)^{-1}_{Y,U}$ with the column $\left(I-Ms^{-1}\right)^{-1}_{j,U}$ without affecting the rank of $\left(I-Ms^{-1}\right)^{-1}_{Y,U}$. This implies that

$$\operatorname{rank}\left(I-Ms^{-1}\right)^{-1}_{Y,U} = \operatorname{rank}\left(I-Ms^{-1}\right)^{-1}_{\bar{Y},U}$$

where $\bar{Y} = (Y \setminus \{k\}) \cup \{j\}$, which concludes the proof. $\square$

Finally, the next lemma provides an easy to check condition for the submatrix $(I-Ms^{-1})^{-1}_{Y,U}$ to have full column rank.

Lemma 8.4. *Let $U, Y \subset V$. If $U \subset Y$ then*

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } M \in \mathcal{P}(\mathcal{M})$$

Proof. As a consequence of Lemma 8.2, we can assume that $\mathcal{M}_{V,Y} = 0$. In particular, $\mathcal{M}_{V,U} = 0$, hence we can take $M \in \mathcal{M}$ and a permutation matrix P such that

$$PMP^\top = \begin{bmatrix} 0 & M_{U,R} \\ 0 & M_{R,R} \end{bmatrix},$$

where $R = V \setminus U$ is the complement of U in V. In other words P shifts the rows corresponding to U to the beginning. Then we obtain

$$P\left(I - Ms^{-1}\right)^{-1} P^\top = \left(I - PMP^\top s^{-1}\right)^{-1} = \begin{bmatrix} I & -M_{U,R} \\ 0 & I - M_{R,R} \end{bmatrix}^{-1} = \begin{bmatrix} I & \bullet \\ 0 & \bullet \end{bmatrix},$$

thus, by the definition of P,

$$\left(I - Ms^{-1}\right)^{-1}_{U,U} = I.$$

Therefore,

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{U,U} = |U| \text{ for all } M \in \mathcal{P}(\mathcal{M}),$$

and we conclude that

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } M \in \mathcal{P}(\mathcal{M})$$

□

With the results from these lemmas in mind, we can define a graph simplification process analogous to the one in [23].

Graph simplification process.
Let $\mathcal{M}$ be a square pattern matrix and let $\mathscr{G}(\mathcal{M}) = (V, E_* \cup E_?)$ be the graph associated to it. Suppose $U, Y \subset V$ and consider the following operations on $\mathscr{G}(\mathcal{M})$ and the set Y:

(i) Remove all outgoing edges of nodes in Y from $\mathscr{G}(\mathcal{M})$.

(ii) If $k \in Y \setminus U$ has only one in-neighbour j that is reachable from U and the edge connecting j to k is solid, then replace k with j in Y.

Consecutively apply operations (i) and (ii) until no more changes are possible.

Note that operation (i) changes the graph $\mathscr{G}(\mathcal{M})$, while operation (ii) changes the set Y. The derived set $\mathscr{D}_{\text{gsp}}(Y; U)$ is the set Y obtained after the graph simplification process has terminated. Similarly, the *derived graph* is the graph obtained after the graph simplification process has terminated. This graph corresponds to a pattern matrix, which we refer to as the *derived pattern matrix* and denote it by $\mathscr{D}_{\text{gsp}}(\mathcal{M})$. Therefore, the derived graph is the graph $\mathscr{G}(\mathscr{D}_{\text{gsp}}(\mathcal{M}))$. Using the lemmas above, we can prove the following result.

Proposition 8.1. *Let $\mathcal{M}$ be a square pattern matrix and let $\mathscr{G}(\mathcal{M}) = (V, E_* \cup E_?)$ be the graph associated to it. Moreover, let $U, Y \subset V$ and consider the graph simplification process for the graph $\mathscr{G}(\mathcal{M})$. If $U \subset \mathscr{D}_{\text{gsp}}(Y; U)$, then*

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } M \in \mathcal{P}(\mathcal{M}).$$

Proof. From Lemma 8.4 it follows that

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{\mathscr{D}_{\mathrm{gsp}}(Y;U),U} = |U| \text{ for all } M \in \mathcal{P}(\mathscr{D}_{\mathrm{gsp}}(\mathcal{M})).$$

Applying Lemma 8.2 and Lemma 8.3 for each time operation (i) and operation (ii) from the graph simplification process are applied allows us to conclude that

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } M \in \mathcal{P}(\mathcal{M}).$$

□

We can use the result of this proposition to derive the following graph-theoretic sufficient condition for strong structural left invertibility of $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

Theorem 8.5. *The system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if $U \subset \mathscr{D}_{\mathrm{gsp}}(Y;U)$ when applying the graph simplification process to the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, where U is the set of input nodes and Y is the set of output nodes.*

Proof. Recall that the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if the transfer matrix $T(s)$ is left invertible for all $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$. Let $A \in \mathcal{P}(\mathcal{A})$, $B \in \mathcal{P}(\mathcal{B})$, $C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ and consider the transfer matrix $T(s)$. We have shown that $T(s)$ has full column rank if and only if

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| = m,$$

where $U = \{n+1, \ldots, n+m\}$, $Y = \{n+m+1, \ldots, n+m+p\}$ and

$$M = \begin{bmatrix} A & B & 0 \\ 0 & 0 & 0 \\ C & D & 0 \end{bmatrix}.$$

Therefore, the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| \text{ for all } M \in \mathcal{P}(\mathcal{M}),$$

where

$$\mathcal{M} = \begin{bmatrix} \mathcal{A} & \mathcal{B} & 0 \\ 0 & 0 & 0 \\ \mathcal{C} & \mathcal{D} & 0 \end{bmatrix}.$$

In view of Proposition 8.1, this holds if $\mathscr{D}_{\mathrm{gsp}}(Y;U)$ when applying the graph simplification process to the graph $\mathscr{G}(\mathcal{M})$. However, the graph $\mathscr{G}(\mathcal{M})$ is just the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. Furthermore, the set of input nodes and the set of output nodes in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are precisely the sets of nodes U and Y in $\mathscr{G}(\mathcal{M})$. Therefore, we can rephrase the condition of Proposition 8.1 to conclude that

$$\operatorname{rank}\left(I - Ms^{-1}\right)^{-1}_{Y,U} = |U| = m \text{ for all } M \in \mathcal{P}(\mathcal{M})$$

if $U \subset \mathscr{D}_{\mathrm{gsp}}(Y;U)$ when applying the graph simplification process to the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, where U is the set of input nodes and Y is the set of output nodes. □

Examples of how the graph simplification process can be used to check whether a system is strongly structurally left invertible can be found in Section 8.5.

8.4 Single-input single-output systems

In this section we consider only single-input single-output systems $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. The transfer matrix of such systems is scalar, hence left invertibility of (A, B, C, D) is equivalent to having a nonzero transfer function $T(s) = D + C(sI - A)^{-1}B$. Checking for a nonzero transfer function is much simpler than checking for a full column rank transfer matrix, thus it naturally leads to much simpler conditions for left invertibility of (A, B, C, D), and in turn, for strong structural left invertibility of $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

We have already seen that the transfer matrix $T(s)$ can be written as

$$T(s) = D + \sum_{i=1} CA^{i-1}Bs^{-i}.$$

Therefore, in the single input single output case, the transfer function $T(s)$ is nonzero if and only if $D \neq 0$ or $CA^{i-1}B \neq 0$ for some positive integer i. This also happens to be equivalent to the condition for output controllability given in Theorem 3.1. With the discussion in the previous subsections in mind, the following result should not be surprising.

Theorem 8.6. *The single-input single-output system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible and strongly structurally output controllable if there exists a positive integer k such that there is unique solid k-walk from the input node to the output node in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.*

Proof. Let $(A, B, C, D) \in A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ and let M be the $(n+2) \times (n+2)$ matrix defined by

$$M = \begin{bmatrix} A & B & 0 \\ 0 & 0 & 0 \\ C & D & 0 \end{bmatrix}.$$

Recall that the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is defined as the graph associated to

$$\mathcal{M} = \begin{bmatrix} \mathcal{A} & \mathcal{B} & 0 \\ 0 & 0 & 0 \\ \mathcal{C} & \mathcal{D} & 0 \end{bmatrix}.$$

Let $u = n + 1$ be the input node and $y = n + 2$ be the output node in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. Given that there exists a positive integer k and a unique solid k-walk from u to y, it follows that $\mathcal{M}^k_{yu} = *$, which implies that $M^k_{yu} \neq 0$. Since $M_{yu} = D$ and $M^i_{uy} = CA^{i-1}B$ for all positive integers i, we find that $D \neq 0$ if $k = 1$ and $CA^{k-1}B \neq 0$ if $k \geq 2$. In both cases, this implies that (A, B, C, D) is left invertible and output controllable. As A, B, C and D were chosen arbitrarily, we conclude that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible and strongly structurally output controllable. $\square$

Remark 8.2. *The argument in the proof of Theorem 8.6 can be applied to multi-output systems for strong structural left invertibility, and to multi-input systems for strong structural output controllability. Indeed, in a single-input multi-output system (A, B, C, D), the transfer matrix $T(s)$ is a column vector, hence $T(s)$ is left invertible if and only if it is nonzero. This implies that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally left invertible if there exists a positive integer k such that there is a unique solid k-walk from the input node to at least one of the output nodes. Similarly, in a multi-input single-output system (A, B, C, D), the matrix $\begin{bmatrix} D & CB & \cdots & CA^{n-1}B \end{bmatrix}$ is a row vector, hence it has full row rank if and only if it is nonzero. This implies that $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is strongly structurally output controllable if there exists a positive integer k such that there is a unique solid k-walk from at least one of the input nodes to the output node.*

Note that the there exists a positive integer k such that there is a unique solid k-walk from the input node to the output node in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ if and only if $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is output colorable. In view of this, we have already seen how the condition of Theorem 8.6 is not necessary in Example 6.2.

8.5 Summary and comparison

In this section we summarize the sufficient conditions obtained so far and compare them with the help of a few illustrative examples. To this end, consider the system $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ and let U and Y denote the set of input nodes and output nodes in the system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$, respectively. Furthermore, let $\mathcal{A}^*$ be the pattern matrix obtained from $\mathcal{A}$ by changing the diagonal entries to $*$'s and define $\mathcal{T}_0 \in \{0,*,?\}^{p\times m}$ by

$$(\mathcal{T}_0)_{ij} = \begin{cases} 0 & \text{if there is no } \delta_j\text{-walk from } u_j \text{ to } y_i, \\ * & \text{if there is a unique solid } \delta_j\text{-walk from the } u_j \text{ to } y_i, \\ ? & \text{otherwise,} \end{cases}$$

where δ_j is the distance from u_j to the output nodes in $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$. Then we have the following sufficient conditions for strong structural left invertibility of $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$.

S1. One of the pattern matrices

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

has full column rank. Equivalently, the input and state nodes in one of the system graphs

$$\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D}) \quad \text{or} \quad \mathscr{G}(\mathcal{A}^*,\mathcal{B},\mathcal{C},\mathcal{D})$$

are dual colorable.

S2. The pattern matrix $\mathcal{T}_0 \in \{0,*,?\}^{p\times m}$ has full column rank. Equivalently, the system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is input colorable.

S3. The derived set $\mathscr{D}_{\text{gsp}}(Y;U)$ contains all inputs after applying the graph simplification process to the system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$.

SISO. The system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ contains a single input node, a single output node, and a unique solid k-walk from the input node to the output node for some positive integer k.

We begin by comparing the first three conditions, which are applicable to any system $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$. In particular, we will provide three examples that show how these conditions do not imply each other.

Example 8.5. Consider the system $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cc:cc} * & * & ? & 0 \\ * & 0 & * & 0 \\ \hdashline * & 0 & 0 & 0 \\ * & * & 0 & * \end{array}\right].$$

The corresponding system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is depicted in Figure 8.3. Note that $\delta_1 = 2$ and $\delta_1 = 1$.

S1. The pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} * & * & ? & 0 \\ * & 0 & * & 0 \\ * & 0 & 0 & 0 \\ * & * & 0 & * \end{bmatrix}.$$

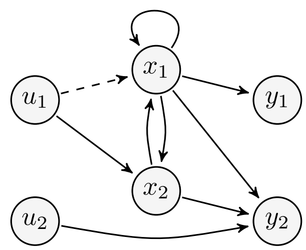


Figure 8.3: The system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$.

has full column rank. Indeed, by swapping the first and third row, and the first and second column, we obtain the pattern matrix

$$\begin{bmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ * & ? & * & 0 \\ * & 0 & * & * \end{bmatrix},$$

which is of Form III, hence **S1** is satisfied.

S2. There is a δ_1-walk from u_1 to both y_1 and y_2. However, the δ_1-walk from u_1 to y_1 is not solid, while the δ_1-walk from u_1 to y_2 is neither unique nor solid. Given that there is a δ_2-walk from u_2 only to y_2, and that δ_2-walk is unique and solid, it follows that

$$\mathcal{T}_0 = \begin{bmatrix} ? & 0 \\ ? & * \end{bmatrix}.$$

Clearly, the pattern matrix $\mathcal{T}_0$ does not have full column rank, thus **S2** is not satisfied.

S3. Note that y_2 has three in-neighbors that are reachable from the input nodes, hence we cannot apply operation (ii) with y_2. On the other hand, x_1 is the only in-neighbor of y_1 that is reachable from the input nodes, hence we replace y_1 with x_1 in Y, which results in $Y = \{x_1, y_2\}$. Now, we can remove all outgoing edges of x_1 to obtain the simplified graph depicted in Figure 8.4, where the current node set Y is colored black. Since both nodes in Y have two in-neighbors that are reachable from the input nodes, and both nodes in Y have no outgoing edges, we conclude that the graph simplification process has terminated. Given that $\mathscr{D}_{\text{gsp}}(Y;U)$ does not contain all input nodes, we conclude that **S3** is not satisfied.

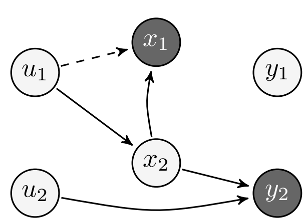


Figure 8.4: The system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ after y_1 has been replaced by x_1.

This example shows that **S1** does not imply **S2** or **S3**.

Example 8.6. Consider the system $(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ with

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cc:cc} * & * & * & 0 \\ * & 0 & * & * \\ \hdashline * & 0 & 0 & 0 \\ * & * & 0 & * \end{array}\right].$$

The corresponding system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$ is depicted in Figure 8.5. Note that $\delta_1 = 2$ and $\delta_1 = 1$.

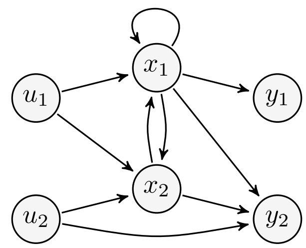


Figure 8.5: The system graph $\mathscr{G}(\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D})$.

S1. The pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} * & * & * & 0 \\ * & 0 & * & * \\ * & 0 & 0 & 0 \\ * & * & 0 & * \end{bmatrix}.$$

does not have full column rank since the matrix

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

is contained in its pattern class and does not have full column rank. Similarly, the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} * & * & * & 0 \\ * & * & * & * \\ * & 0 & 0 & 0 \\ * & * & 0 & * \end{bmatrix}.$$

does not have full column rank since the matrix

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

is contained in its pattern class and does not have full column rank. This implies that **S1** is not satisfied.

S2. There is a δ_1-walk from u_1 to both y_1 and y_2. However, only the δ_1-walk from u_1 to y_1 is unique and solid. Given that there is a δ_2-walk from u_2 only to y_2, and that δ_2-walk is unique and solid, it follows that

$$\mathcal{T}_0 = \begin{bmatrix} * & 0 \\ ? & * \end{bmatrix}.$$

Clearly, the pattern matrix $\mathcal{T}_0$ has full column rank, thus **S2** is satisfied.

S3. Note that y_2 has three in-neighbors that are reachable from the input nodes, hence we cannot apply operation (ii) with y_2. On the other hand, x_1 is the only in-neighbor of y_1 that is reachable from the input nodes, hence we replace y_1 with x_1 in Y, which results in $Y = \{x_1, y_2\}$. Now, we can remove all outgoing edges of x_1 to obtain the simplified graph depicted in Figure 8.6, where the current node set Y is colored black. Since both nodes in Y have two in-neighbors that are reachable from the input nodes, and both nodes in Y have no outgoing edges, we conclude that the graph simplification process has terminated. Given that $\mathscr{D}_{\text{gsp}}(Y;U)$ does not contain all input nodes, we conclude that **S3** is not satisfied.

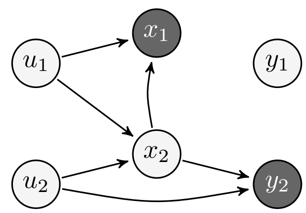


Figure 8.6: The system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ after y_1 has been replaced by x_1.

This example shows that **S2** does not imply **S1** or **S3**.

Example 8.7. Consider the system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ with

$$\left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cccc|cc} * & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & * & * & * \\ 0 & * & * & * & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ \hline * & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & ? & 0 \end{array}\right].$$

The corresponding system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is depicted in Figure 8.7. Note that $\delta_1 = 1$ and $\delta_1 = 2$.

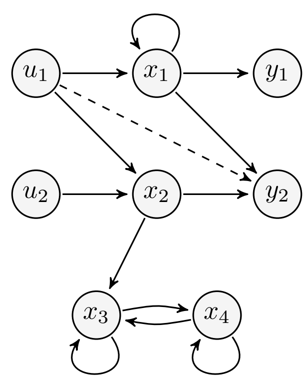


Figure 8.7: The system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

S1. The pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} * & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & * & * & * \\ 0 & * & * & * & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & ? & 0 \end{bmatrix}$$

does not have full column rank since the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

is contained in its pattern class and does not have full column rank. Similarly, the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} * & 0 & 0 & 0 & * & 0 \\ 0 & * & 0 & * & * & * \\ 0 & * & * & * & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & ? & 0 \end{bmatrix}$$

does not have full column rank since the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

is contained in its pattern class and does not have full column rank. This implies that **S1** is not satisfied.

S2. There is a δ_1-walk from u_1 only to y_1, and that δ_1-walk is not solid. Furthermore, there is a δ_2-walk from u_2 only to y_1, and that δ_2-walk is unique and solid. Therefore, we obtain

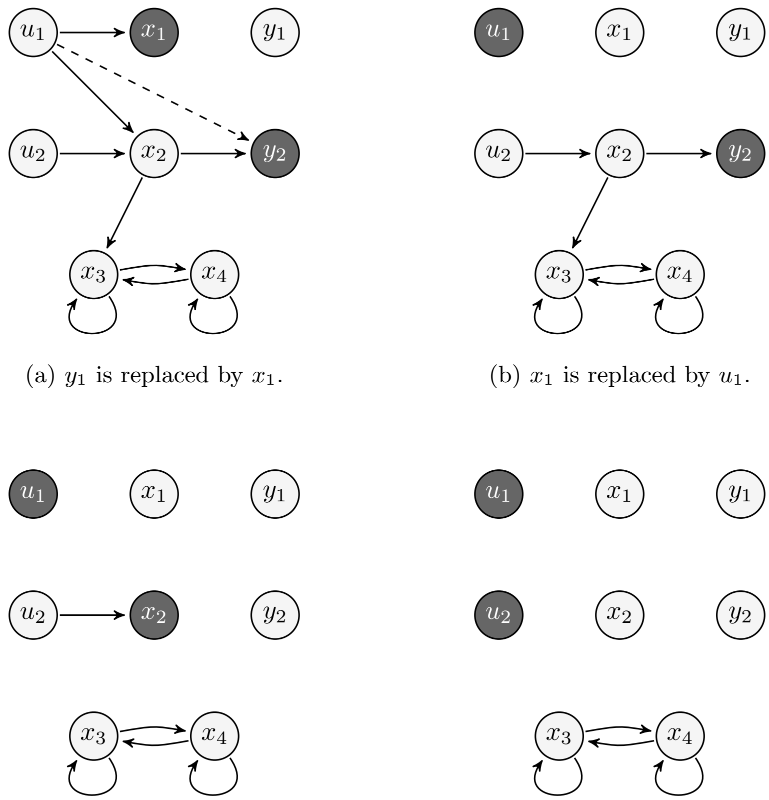
$$\mathcal{T}_0 = \begin{bmatrix} 0 & 0 \\ ? & * \end{bmatrix},$$

which does not have full column rank, thus **S2** is not satisfied.

S3. Note that x_1 is the only in-neighbor of y_1 that is reachable from the input nodes, hence we replace y_1 with x_1 in Y, which results in $Y = \{x_1, y_2\}$. We can remove all outgoing edges of x_1 to obtain the simplified graph depicted in Figure 8.8a, where the current node set Y is colored black. Proceeding in the same fashion, u_1 is the only in-neighbor of x_1 that is reachable from the input nodes, hence we replace x_1 by u_1 to obtain $Y = \{u_1, y_2\}$. Removing all outgoing edges of u_1 results in the simplified graph depicted in Figure 8.8b. Now, x_2 is the only in-neighbor of y_2 that is reachable from the inputs, hence we replace y_2 by x_2 to obtain $Y = \{u_1, x_2\}$. After removing all outgoing edges of x_2, we obtain the simplified graph depicted in Figure 8.8c. Finally, u_2 is the only in-neighbor of x_2 that is reachable from the input nodes, hence we replace x_2 by u_2 to obtain $Y = \{u_1, u_2\}$. The simplified graph after all outgoing edges of u_2 have been removed is depicted in Figure 8.8d. Clearly, $\mathscr{D}_{\text{gsp}}(Y; U)$ contains all input nodes, thus **S3** is satisfied.

This example shows that **S3** does not imply **S1** or **S2**.

In these examples we have seen that none of the first three conditions implies the others, i.e., these conditions are generally independent from each other. However, this is not true in the case of a single-input single-output system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. To see this, note that **S2** holds if and



(a) y_1 is replaced by x_1.

(b) x_1 is replaced by u_1.

(c) y_2 is replaced by x_2.

(d) x_2 is replaced by u_2

Figure 8.8: The graph simplification process applied to the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$.

only if the shortest path from the input node to the output node in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is unique and solid. Moreover, **S3** holds if and only if there is a unique solid path from the input node to the output node in $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$. Indeed, if the path is not unique, then at some point in the graph simplification process, the node in Y will have at least two in-neighbors that are reachable from the input node. Otherwise, if the path is not solid, then at some point in the graph simplification process, the node in Y will have an in-neighbor that is reachable from the input node but the edge connecting that in-neighbor to the node in Y will not be solid. In both cases, the graph simplification process will terminate before $\mathscr{D}_{\text{gsp}}(Y;U)$ contains the input node. With this in mind, we find that in the single-input single-output case, **S3** implies **S2**, which implies **SISO**. This means that one should just check **SISO** instead of **S2** or **S3**.

The role of **S1** is not that clear at the moment. There seems to be no example where **S1** holds but **SISO** does not, which suggests that **S1** also implies **SISO**. Unfortunately, we have not proven this statement, and can only make the following conjecture.

Conjecture 8.1. *The input nodes and state nodes in the system graph $\mathscr{G}(\mathcal{A}^*, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are dual colorable only if $U \in \mathscr{D}_{\text{gsp}}(Y;U)$ when applying the graph simplification process to the system graph $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, where U is the set of input nodes and Y is the set of output nodes.*

Incomplete proof. Consider the dual color change rule for the system graph $\mathscr{G}(\mathcal{A}^*, \mathcal{B}, \mathcal{C}, \mathcal{D})$ and suppose that the input and state nodes are dual colorable. It is clear that in the first step of the application of the dual color change rule, when all nodes are white, we need a node with a single in-neighbor. Furthermore, when that in-neighbor is colored black, we can remove all of its outgoing edges. This is because removing a black in-neighbor has no effect on the next steps of the application of the dual color change rule. If we remove the outgoing edges of each node that gets colored, then we can restate the dual color change rule as follows.

Dual color change rule.

Let $\mathcal{M} \in \{0, *, ?\}^{r\times q}$ and consider the corresponding graph $\mathscr{G}(\mathcal{M})$. Suppose that each node of $\mathscr{G}(\mathcal{M})$ is colored either black or white. If node i has exactly one in-neighbor j, and the edge from j to i is solid, then color j black and remove all of its outgoing edges. We say that node i colors node j black.

In this form, the dual color change rule looks very similar to the consecutive application of operations (i) and (ii) of the graph simplification process. The only difference is that in the graph simplification process we look for in-neighbors that are reachable from the input nodes. With this in mind, we will construct a collection of paths from input nodes to output nodes in $\mathscr{G}(\mathcal{A}^*, \mathcal{B}, \mathcal{C}, \mathcal{D})$, where each node in the path has been colored by the node succeeding it. Note that an input node u_i can be colored only by one of its out-neighbors, i.e., by one of the state nodes or one of the output nodes. If u_i has been colored by an output node, then we are done. Otherwise, since the state nodes have self-loops, a state node could have colored u_i only if it was already black. Furthermore, a node can color only once, hence the state node that has colored u_i could not have colored itself. But then that state node must have been colored by one of the other state nodes or one of the output nodes. Since there are only finitely many state nodes, repeating this argument would eventually result in a path from u_i to an output node, say y_i, where all nodes in the path have been colored black by the next node in the path. Consider one such path and denote it by $(u_i, x_i^1, x_i^2, \ldots, x_i^k, y_i)$ where $k \leq n$ is a positive integer and $x_i^1, \ldots, x_i^k$ are distinct state nodes. Since y_i has colored x_i^k black, we must have had that x_i^k was the only in-neighbor of y_i. Clearly, x_i^k was reachable from the input nodes, hence x_i^k was the only in-neighbor of y_i that is reachable from the input nodes. Given that the only difference between the system graphs $\mathscr{G}(\mathcal{A}^*, \mathcal{B}, \mathcal{C}, \mathcal{D})$ and $\mathscr{G}(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is in the self-loops of the state nodes, the latter is a strong indication that at some point during the graph simplification process, x_i^k was the only in-neighbor of y_i that is reachable from the input nodes. □

If this conjecture is true, then the pattern matrix

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

having full column rank implies that **S3** holds, and consequently, that **SISO** holds. We have already seen in Example 8.5 that the pattern matrix

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$$

having full column rank does not imply that **S3** holds . Nevertheless, it is still possible that **S1** implies **SISO**.

8.6 Necessary condition

In this section, we will briefly discuss the difficulties we encounter when looking for necessary conditions for strong structural left invertibility. We already mentioned that the system matrix approach for left invertibility is not as straightforward as the one for input-state observability mainly because the system matrix $P(\lambda)$ is allowed to drop rank for some $\lambda \in \mathbb{C}$, and the λ's for which it does are not easy to characterize (see [22]). In view of this, we can take an alternative approach and try to device a simplification procedure based on the system matrix and the assumption that it has full column rank for all except finitely many $\lambda \in \mathbb{C}$. To this end, let $A \in \mathcal{P}(\mathcal{A}),\ B \in \mathcal{P}(\mathcal{B}),\ C \in \mathcal{P}(\mathcal{C})$ and $D \in \mathcal{P}(\mathcal{D})$ and consider the corresponding system matrix

$$P(\lambda) = \begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix}.$$

Since $P(\lambda)$ has full column rank for almost all $\lambda \in \mathbb{C}$, we know that the second block column of $P(\lambda)$ must have full column rank, hence we can find a nonsingular $m \times m$ submatrix of $[B^\top \ D^\top]^\top$. For simplicity, suppose that the latter is a submatrix of B and denote it by B_1. Then we can permute the states and partition the system matrix so that

$$P(\lambda) = \begin{bmatrix} A_{11} - \lambda I & A_{12} & B_1 \\ A_{21} & A_{22} - \lambda I & B_2 \\ C_1 & C_2 & D \end{bmatrix},$$

where $B_1 \in \mathbb{R}^{m \times m}$ is nonsingular. We can multiply $P(\lambda)$ from the right by the nonsingular matrix

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -B_1^{-1}(A_{11} - \lambda I) & -B_1^{-1}A_{12} & I \end{bmatrix}$$

to obtain

$$\begin{bmatrix} 0 & 0 & B_1 \\ A_{21} - B_2 B_1^{-1}(A_{11} - \lambda I) & A_{22} - B_2 B_1^{-1} A_{12} - \lambda I & B_2 \\ C_1 - D B_1^{-1}(A_{11} - \lambda I) & C_2 - D B_1^{-1} A_{12} & D \end{bmatrix}.$$

This implies that $P(\lambda)$ has full column rank for all except finitely many $\lambda \in \mathbb{C}$ if and only if $\hat{P}(\lambda)$ has full column rank for all except finitely many $\lambda \in \mathbb{C}$, where $\hat{P}(\lambda)$ is defined as

$$\hat{P}(\lambda) = \begin{bmatrix} A_{21} - B_2 B_1^{-1}(A_{11} - \lambda I) & A_{22} - B_2 B_1^{-1} A_{12} - \lambda I \\ C_1 - D B_1^{-1}(A_{11} - \lambda I) & C_2 - D B_1^{-1} A_{12} \end{bmatrix}.$$

Unfortunately, this does not seem very promising. Even in the single-input single-output case with $D = 0$, we have

$$\hat{P}(\lambda) = \begin{bmatrix} a_{21} - b_1^{-1}(a_{11} - \lambda) b_2 & A_{22} - b_1^{-1} b_2 a_{12}^\top - \lambda I \\ c_1 & c_2^\top \end{bmatrix},$$

where $b_1, a_{11}, c_1 \in \mathbb{R}$, with $b_1 \neq 0$, and $a_{12}, a_{21}, c_2 \in \mathbb{R}^{n-1}$. Here we see that the matrix in the top right block involves a term of the form $b_2 a_{12}^\top$, which is hard to deal with in the context of pattern matrices. This is because the rows and columns of $b_2 a_{12}^\top$ are generally dependent on each other, which implies that we cannot express $A_{22} - b_1^{-1} b_2 a_{12}^\top$ as a general element of the pattern class of some pattern matrix.

The problem with dependencies between entries has already appeared in the context of strong structural output controllability. In Example 6.2, we saw that the product of pattern matrices fails to capture the dependencies between entries in the product of real matrices belonging to their pattern classes. More precisely, we saw that in a single-input single-output system $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ it might happen that $\mathcal{C}\mathcal{A}\mathcal{B} = \mathcal{C}\mathcal{A}^2\mathcal{B} = ?$, while it is never the case that $CAB = CA^2B = 0$ for some $A \in \mathcal{P}(\mathcal{A}), B \in \mathcal{P}(\mathcal{B})$ and $C \in \mathcal{P}(\mathcal{C})$. Given that the transfer matrix $T(s)$ can be expressed as

$$T(s) = D + CBs^{-1} + CABs^{-2} + CA^2Bs^{-3} + \cdots,$$

it should be clear that the same problem appears when taking the transfer matrix approach to strong structural left invertibility.

In conclusion, it seems that the problems in finding necessary conditions for strong structural output controllability and strong structural left invertibility are the same, namely, the currently developed techniques with pattern matrices do not provide a way to capture dependencies in the products of real matrices belonging to their pattern classes. One possible solution might be to modify the definition of a pattern matrix to include dependencies between entries and define addition and multiplication for such pattern matrices just as we did in this thesis. Some work in this direction is already done in [25] for the problem of characterizing strong structural controllability for leader-follower network of linear systems. There, the pattern matrix describing the network dynamics can include nonzero entries that are constrained to be equal. Unfortunately, characterizing strong structural controllability in this case seems to be much more difficult and the results in [25] provide only a sufficient condition. This suggests that one might have to develop a completely new toolbox to deal with this problem, and, consequently, with the problem of characterizing strong structural output controllability and left invertibility for the structured systems studied in this thesis. In any case, the solution seems to be non-trivial and would require further research.

Chapter 9

Controllability of linear descriptor systems

In this section, we will extend the results on strong structural controllability developed in [5] to linear descriptor systems. Descriptor systems, also known as singular systems, are systems in which the time evolution of the state is described by a differential-algebraic equation, as opposed to an ordinary differential equation. We will consider linear descriptor systems of the form

$$E\dot{x}(t) = Ax(t) + Bu(t), \tag{9.1}$$

where $x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^n$ and $E \in \mathbb{R}^{n \times n}$. We will denote the system in (9.1) by (E, A, B). For a given initial state $x(0) = x_0 \in \mathbb{R}^n$ and input signal $u \in C^\infty$, the corresponding state trajectory will be denoted by $x(t; x_0, u)$. Without going into details, we note that the concept of controllability for linear descriptor systems is often broken down into specific kinds of controllability. In this thesis, we will only consider the natural generalization of controllability to linear descriptor systems.

Definition 9.1. The descriptor system (E, A, B) is called *controllable* if for any $x_0, x_1 \in \mathbb{R}^n$, there exists an input signal $u \in C^\infty$ and a time $T \geq 0$ such that $x(T; x_0, u) = x_1$.

Naturally, the characterization of controllability for descriptor systems is very similar to the one for ordinary linear systems.

Theorem 9.1. *The system (E, A, B) is controllable if and only if matrix $\begin{bmatrix} E & B \end{bmatrix}$ has full row rank, and the matrix*

$$\begin{bmatrix} A - \lambda E & B \end{bmatrix}$$

has full row rank for all $\lambda \in \mathbb{C}$.

Proof. For a proof of this theorem, as well as more information on controllability of descriptor systems, we refer to [26]. ☐

Just like before, we will consider a descriptor system (E, A, B), where the entries of the matrices E, A and B are not known precisely, but are known to be fixed zeros, nonzeros or arbitrary real numbers. In other words, we know that $E \in \mathcal{P}(\mathcal{E}), A \in \mathcal{P}(\mathcal{A})$ and $B \in \mathcal{P}(\mathcal{B})$, where $\mathcal{E}, \mathcal{A} \in \{0, *, ?\}^{n \times n}$ and $\mathcal{B} \in \{0, *, ?\}^{n \times m}$ are known pattern matrices. This naturally leads to a family of systems as E, A and B range over $\mathcal{P}(\mathcal{E}), \mathcal{P}(\mathcal{A})$ and $\mathcal{P}(\mathcal{B})$, respectively. We will denote this family of systems by $(\mathcal{E}, \mathcal{A}, \mathcal{B})$, and will refer to it as a *structured descriptor system*, or just the descriptor system $(\mathcal{E}, \mathcal{A}, \mathcal{B})$. We are interested in properties of (E, A, B) that stem solely from the fact that

$E \in \mathcal{P}(\mathcal{E}), A \in \mathcal{P}(\mathcal{A})$ and $B \in \mathcal{P}(\mathcal{B})$. As such, they are more appropriately seen as properties of the structured descriptor system $(\mathcal{E}, \mathcal{A}, \mathcal{B})$. With this in mind, consider the following definition.

Definition 9.2. The descriptor system $(\mathcal{E}, \mathcal{A}, \mathcal{B})$ is strongly structurally controllable if (E, A, B) is controllable for all $E \in \mathcal{P}(\mathcal{E}), A \in \mathcal{P}(\mathcal{A})$ and $B \in \mathcal{P}(\mathcal{B})$.

It turns out that strong structural controllability of linear descriptor systems can be characterized very similarly to strong structural controllability of ordinary linear systems. Using the ideas in [5] and the definition of addition for pattern matrices, we obtain the following result.

Theorem 9.2. *The descriptor system $(\mathcal{E}, \mathcal{A}, \mathcal{B})$ is strongly structurally controllable if and only if the pattern matrices*

$$\begin{bmatrix}\mathcal{E} & \mathcal{B}\end{bmatrix}, \quad \begin{bmatrix}\mathcal{A} & \mathcal{B}\end{bmatrix} \quad \textit{and} \quad \begin{bmatrix}\mathcal{A}+\mathcal{E} & \mathcal{B}\end{bmatrix}$$

have full row rank.

Proof. ($\Rightarrow$) Suppose $(\mathcal{E}, \mathcal{A}, \mathcal{B})$ is strongly structurally controllable. Using Theorem 9.1, this means that $\begin{bmatrix}E & B\end{bmatrix}$ has full row rank for all $E \in \mathcal{P}(\mathcal{E})$ and $B \in \mathcal{P}(\mathcal{B})$, hence $\begin{bmatrix}\mathcal{E} & \mathcal{B}\end{bmatrix}$ has full row rank. Furthermore, the matrix

$$\begin{bmatrix}A - \lambda E & B\end{bmatrix}$$

has full row rank for all $\lambda \in \mathbb{C}$ for all $E \in \mathcal{P}(\mathcal{E}), A \in \mathcal{P}(\mathcal{A})$ and $B \in \mathcal{P}(\mathcal{B})$. Taking $\lambda = 0$ shows that $\begin{bmatrix}\mathcal{A} & \mathcal{B}\end{bmatrix}$ has full row rank. On the other hand, taking $\lambda = -1$ and using Proposition 4.1 shows that $\begin{bmatrix}\mathcal{A}+\mathcal{E} & \mathcal{B}\end{bmatrix}$ has full row rank.

($\Leftarrow$) Suppose that the pattern matrices

$$\begin{bmatrix}\mathcal{E} & \mathcal{B}\end{bmatrix}, \quad \begin{bmatrix}\mathcal{A} & \mathcal{B}\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}\mathcal{A}+\mathcal{E} & \mathcal{B}\end{bmatrix}$$

have full row rank. Let $E \in \mathcal{P}(\mathcal{E}), A \in \mathcal{P}(\mathcal{A})$ and $B \in \mathcal{P}(\mathcal{B})$ and note that $\begin{bmatrix}E & B\end{bmatrix}$ has full row rank because $\begin{bmatrix}\mathcal{E} & \mathcal{B}\end{bmatrix}$ has full row rank. Next, consider the equation

$$x^\top \begin{bmatrix}A - \lambda E & B\end{bmatrix} = 0,$$

where $\lambda \in \mathbb{C}$. We would like to show that $x = 0$ for all $\lambda \in \mathbb{C}$. For $\lambda = 0$, the equation boils down to

$$x^\top \begin{bmatrix}A & B\end{bmatrix} = 0,$$

which implies that $x = 0$ because $\begin{bmatrix}\mathcal{A} & \mathcal{B}\end{bmatrix}$ has full row rank. Suppose $\lambda \neq 0$ and write $x = a + bi$, where i is the imaginary unit and a, b are real vectors. Furthermore, let α be a nonzero real number such that

$$\alpha \notin \left\{\frac{a_k}{b_k} \mid b_k \neq 0\right\} \cup \left\{\frac{(a^\top A)_k}{(b^\top A)_k} \mid (b^\top A)_k \neq 0\right\} \cup \left\{\frac{(a^\top E)_k}{(b^\top E)_k} \mid (b^\top E)_k \neq 0\right\}$$

and define $\hat{x} = a + \alpha b$. The choice of α and the assumption that $\lambda \neq 0$ imply that

(i) $\hat{x}_k = 0$ if and only if $x_k = 0$.

(ii) $(\hat{x}^\top A)_k = 0$ if and only if $(\hat{x}^\top E)_k = 0$.

One can easily verify that the choice of α implies (i). For the only if part of (ii), suppose that $(\hat{x}^\top A)_k = 0$, which implies that $(a^\top A)_k + \alpha(b^\top A)_k = 0$. Due to the choice of α, the latter is possible if and only if $(a^\top A)_k = (b^\top A)_k = 0$. But then we have $(x^\top A)_k = 0$, and since x is such that $x^\top A = \lambda x^\top E$, it follows that $(x^\top A)_k = (x^\top E)_k = 0$. Therefore $(a^\top E)_k = 0$ and $(b^\top E)_k = 0$, which shows that $(\hat{x}^\top E)_k = 0$. The if part of (ii) follows the same reasoning.

With this in mind, define the diagonal matrix Δ by

$$\Delta_{kk} = \begin{cases} -1 & \text{if } (\hat{x}^\top E)_k = 0, \\ -\dfrac{(\hat{x}^\top A)_k}{(\hat{x}^\top E)_k} & \text{otherwise,} \end{cases}$$

which is nonsingular because of (ii). This also implies that $E\Delta \in \mathcal{P}(\mathcal{E})$ since the pattern class of a pattern matrix is invariant under nonzero scaling of the columns. Furthermore, Δ is such that $\hat{x}^\top A + \hat{x}^\top E\Delta = 0$, while $x^\top B = 0$ only if $\hat{x}^\top B = 0$. Consequently, we obtain

$$\hat{x}^\top \begin{bmatrix} A + E\Delta & B \end{bmatrix} = 0.$$

However, $A - E\Delta \in \mathcal{P}(\mathcal{A} + \mathcal{E})$ and $\begin{bmatrix} \mathcal{A} + \mathcal{E} & \mathcal{B} \end{bmatrix}$ has full row rank, hence $\hat{x} = 0$. Using (i), this implies that $x = 0$, which concludes the proof. $\square$

We can check whether the pattern matrices in Theorem 9.2 have full row rank by checking whether the graphs associated to them are colorable. This is illustrated in the following example.

Example 9.1. Consider the descriptor system $(\mathcal{E}, \mathcal{A}, \mathcal{B})$ with

$$\mathcal{E} = \begin{bmatrix} * & * & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & * & 0 & * \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{A} = \begin{bmatrix} * & 0 & 0 & 0 \\ 0 & ? & 0 & * \\ * & 0 & * & * \\ 0 & ? & 0 & * \end{bmatrix} \quad \text{and} \quad \mathcal{B} = \begin{bmatrix} * & 0 \\ ? & 0 \\ 0 & 0 \\ 0 & * \end{bmatrix}.$$

The system is controllable if and only if the pattern matrices

$$\begin{bmatrix} \mathcal{E} & \mathcal{B} \end{bmatrix} = \begin{bmatrix} * & * & 0 & 0 & * & 0 \\ 0 & * & 0 & 0 & ? & 0 \\ 0 & * & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix},$$

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \end{bmatrix} = \begin{bmatrix} * & 0 & 0 & 0 & * & 0 \\ 0 & ? & 0 & * & ? & 0 \\ * & 0 & * & * & 0 & 0 \\ 0 & ? & 0 & * & 0 & * \end{bmatrix},$$

$$\begin{bmatrix} \mathcal{A} + \mathcal{E} & \mathcal{B} \end{bmatrix} = \begin{bmatrix} ? & * & 0 & 0 & * & 0 \\ 0 & ? & 0 & * & ? & 0 \\ * & * & * & ? & 0 & 0 \\ 0 & ? & 0 & * & 0 & * \end{bmatrix}$$

have full row rank. We will check this using the color change rule, i.e., we will check whether the graphs $\mathscr{G}((\begin{bmatrix} \mathcal{E} & \mathcal{B} \end{bmatrix}))$, $\mathscr{G}\left(\begin{bmatrix} \mathcal{A} & \mathcal{B} \end{bmatrix}\right)$ and $\mathscr{G}\left(\begin{bmatrix} \mathcal{A} + \mathcal{E} & \mathcal{B} \end{bmatrix}\right)$ are colorable.

To this end, consider the graph $\mathscr{G}((\begin{bmatrix} \mathcal{E} & \mathcal{B} \end{bmatrix}))$ depicted in Figure 9.1. Note that 4 is the only white out-neighbor of 6, hence 6 colors 4 black (see Figure 9.2a). Similarly, 3 is the only white out-neighbor of 4, hence 4 colors 3 black (see Figure 9.2b), and 1 is the only white out-neighbor of 1, hence 1 colors itself black (see Figure 9.2c). With 1 and 3 being black, the only white out-neighbor of 2 is 2, hence 2 colors itself black (see Figure 9.2d). Given that four nodes have been colored black, we conclude that the graph $\mathscr{G}\left(\begin{bmatrix} \mathcal{E} & \mathcal{B} \end{bmatrix}\right)$ is colorable.

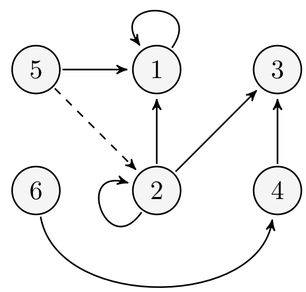
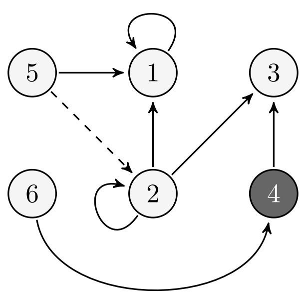
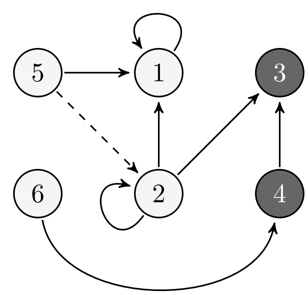


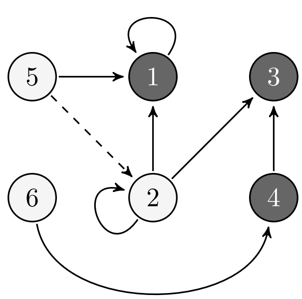
Figure 9.1: The graph $\mathscr{G}\left(\begin{bmatrix}\mathcal{E} & \mathcal{B}\end{bmatrix}\right)$.



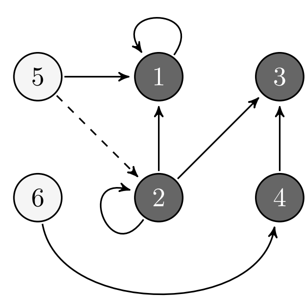
(a) 6 colors 4 black.



(b) 4 colors 3 black.



(c) 1 colors 1 black.



(d) 2 colors 2 black.

Figure 9.2: Repeated application of the color change rule for the graph $\mathscr{G}\left(\begin{bmatrix}\mathcal{E} & \mathcal{B}\end{bmatrix}\right)$.

Next, consider the graph $\mathscr{G}\left(\begin{bmatrix}\mathcal{A} & \mathcal{B}\end{bmatrix}\right)$ depicted in Figure 9.3. Note that 4 is the only white out-neighbor of 6, hence 6 colors 4 black (see Figure 9.4a). Similarly, 3 is the only white out-neighbor of 3, hence 3 colors itself black (see Figure 9.4b). With 3 and 4 being black, the only white out-neighbor of 4 is 2, hence 4 colors 2 black (see Figure 9.4c). Finally, the only white out-neighbor of 5 is 1, hence 5 colors 1 black (see Figure 9.4d). Given that four nodes have been colored black, we conclude that the graph $\mathscr{G}\left(\begin{bmatrix}\mathcal{A} & \mathcal{B}\end{bmatrix}\right)$ is colorable.

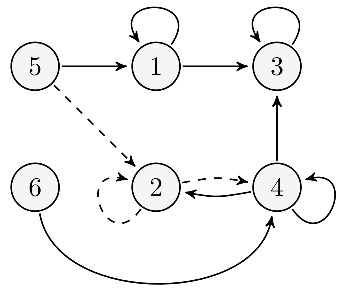
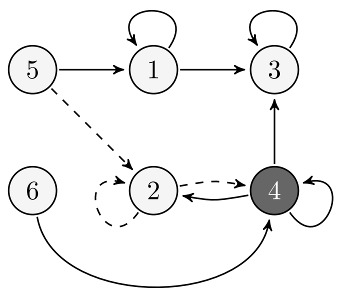
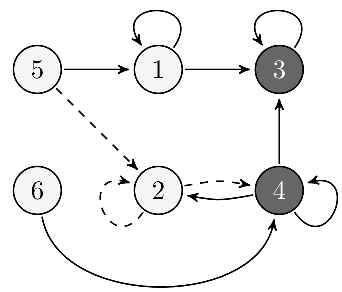


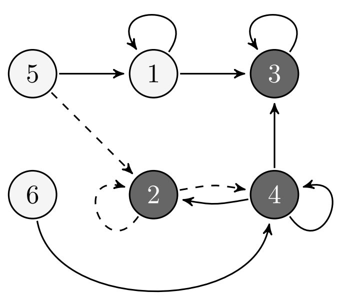
Figure 9.3: The graph $\mathscr{G}\left(\begin{bmatrix}\mathcal{A} & \mathcal{B}\end{bmatrix}\right)$.



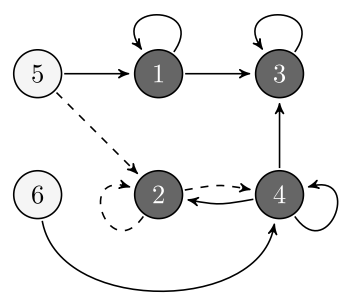
(a) 6 colors 4 black.



(b) 3 colors 3 black.



(c) 4 colors 2 black..



(d) 5 colors 1 black.

Figure 9.4: Repeated application of the color change rule for the graph $\mathscr{G}\left(\begin{bmatrix}\mathcal{A} & \mathcal{B}\end{bmatrix}\right)$.

Finally, consider the graph $\mathscr{G}\left(\begin{bmatrix}\mathcal{A}+\mathcal{E} & \mathcal{B}\end{bmatrix}\right)$ depicted in Figure 9.5. Note that 4 is the only white out-neighbor of 6, hence 6 colors 4 black (see Figure 9.6a). Similarly, 3 is the only white out-neighbor of 3, hence 3 colors itself black. With 3 and 4 being black, the only white out-neighbor of 4 is 2, hence 4 colors 2 black. Finally, the only white out-neighbor of 5 is 1, hence 5 colors 1 black. Given that four nodes have been colored black, we conclude that the graph $\mathscr{G}\left(\begin{bmatrix}\mathcal{A}+\mathcal{E} & \mathcal{B}\end{bmatrix}\right)$ is colorable. This also shows that the descriptor system $(\mathcal{E}, \mathcal{A}, \mathcal{B})$ is strongly structurally controllable.

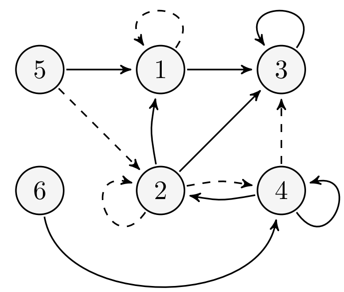
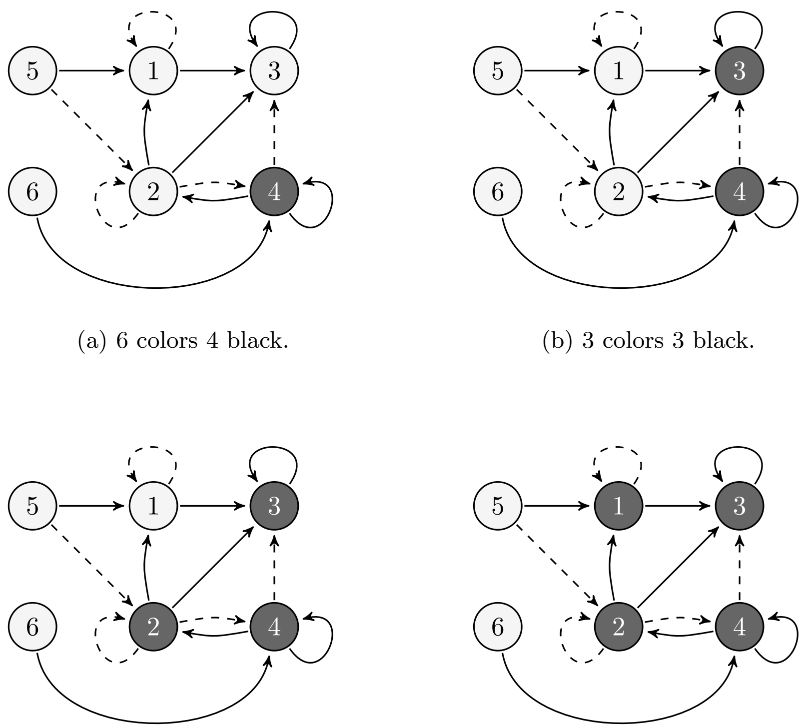


Figure 9.5: The graph $\mathscr{G}\left(\left[\mathcal{A}+\mathcal{E} \quad \mathcal{B}\right]\right)$.



(a) 6 colors 4 black.

(b) 3 colors 3 black.

(c) 4 colors 2 black..

(d) 5 colors 1 black.

Figure 9.6: Repeated application of the color change rule for the graph $\mathscr{G}\left(\left[\mathcal{A}+\mathcal{E} \quad \mathcal{B}\right]\right)$.

Chapter 10

Conclusion

In this thesis, we have studied strong structural output controllability, input-state observability and left invertibility of structured linear systems. Furthermore, we have extended the existing results on strong structural controllability to linear descriptor systems. We have obtained results for structured systems in which the entries of the system matrices are fixed zeros, nonzeros or arbitrary real numbers. This has been formalized by the introduction of pattern matrices and their pattern classes. In order to get an idea of the results of expressions with real matrices belonging to the pattern classes of some known pattern matrices, sensible definitions of addition and multiplication for pattern matrices have been introduced. Furthermore, a pattern matrix has been associated to a graph, where the existence and the type of edges represent the entries of the pattern matrix. Consequently, strong structural properties have been characterized as properties of the pattern matrices describing the structured system, as well as graph-theoretic properties of the graphs associated to those pattern matrices.

Sufficient conditions for strong structural output controllability have been found in terms of a full row rank property of a particular pattern matrix. As opposed to this algebraic characterization, we have also provided a graph-theoretic sufficient condition based on the system graph of the structured system. This has been phrased as a condition on the outcome of an output color change rule for the system graph. Moreover, the limitation of this approach have been discussed and put forward as reasons for our inability to obtain necessary conditions.

Necessary and sufficient conditions for strong structural input-state observability have been found in terms of a full column rank property of a pair of pattern matrices. As with strong structural output controllability, this algebraic characterization has been complemented by a graph-theoretic necessary and sufficient condition based on the system graph and a modified system graph. This has been phrased as a condition on the outcome of a dual color change rule for the system graph and the modified system graph.

Four sufficient conditions for strong structural left invertibility have been found. One of these is based on the system matrix, while the rest are based on the transfer matrix. The one based on the system matrix has been presented as a full column rank property of at least one out of two pattern matrices, as well as a condition on the outcome of a dual color change rule on the system graph or a modified system graph. The first result based on the transfer matrix has been presented as a full column rank property of a pattern matrix that can be obtained using the system graph. This has been complemented by an equivalent condition based on the outcome of an input color change rule for the system graph. The second result based on the transfer matrix is purely graph-theoretic and has been phrased as a condition on the outcome of a graph simplification process for the system graph. The last result based on the transfer matrix is applicable only in the case of a single-input single-output system and has been phrased as a graph-theoretic condition on the system graph. All four conditions have been compared, and it has been concluded that, except in the single-input single-output case, they are independent from each other. Finally, the difficulties in obtaining necessary conditions have been discussed and it has been shown that they have to do with the same limitations that were present in the study of strong structural output controllability.

Finally, necessary and sufficient conditions for strong structural controllability of linear descriptor systems have been provided in terms of a full row rank property of three pattern matrices. The conditions are comparable to the ones for strong structural controllability of ordinary linear systems.

In the future, research into these topics should focus on overcoming the limitations of the approach taken in this thesis. In particular, the definition of multiplication for pattern matrices is very useful in obtaining sufficient conditions but it ultimately fails in the search for necessary conditions. This has been elaborated upon in the chapters on strong structural output controllability and left invertibility. Given that the problem lies in the inability of pattern matrices to capture dependencies between entries, it seems that the only way forward is in redefining pattern matrices, or taking a completely novel approach in characterizing strong structural properties.

Bibliography

[1] C.-T. Lin, "Structural controllability," *IEEE Transactions on Automatic Control*, vol. 19, pp. 201–208, June 1974.

[2] R. Shields and J. Pearson, "Structural controllability of multiinput linear systems," *IEEE Transactions on Automatic Control*, vol. 21, pp. 203–212, Apr. 1976.

[3] K. Glover and L. Silverman, "Characterization of structural controllability," *IEEE Transactions on Automatic Control*, vol. 21, no. 4, pp. 534–537, 1976.

[4] H. Mayeda and T. Yamada, "Strong structural controllability," *SIAM Journal on Control and Optimization*, vol. 17, no. 1, pp. 123–138, 1979.

[5] J. Jia, H. J. van Waarde, H. L. Trentelman, and M. K. Camlibel, "A Unifying Framework for Strong Structural Controllability," *arXiv e-prints*, p. arXiv:1903.03353, 2019.

[6] N. Monshizadeh, M. K. Camlibel, and H. L. Trentelman, "Strong targeted controllability of dynamical networks," in *Proceedings of the 54th IEEE Conference on Decision and Control, December 15-18, 2015, Osaka, Japan*, pp. 4782–4787, Feb. 2015.

[7] H. J. van Waarde, M. K. Camlibel, and H. L. Trentelman, "A distance-based approach to strong target control of dynamical networks," *IEEE Transactions on Automatic Control*, vol. 62, pp. 6266–6277, Dec. 2017.

[8] S. Gracy, F. Garin, and A. Kibangou, "Structural and strongly structural input and state observability of linear network systems," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 2062–2072, 2018.

[9] G. Kossinets, "Effects of missing data in social networks," *Social Networks*, vol. 28, no. 3, pp. 247–268, 2006.

[10] R. West, A. Paranjape, and J. Leskovec, "Mining missing hyperlinks from human navigation traces: A case study of Wikipedia," in *Proceedings of the 24th international conference on World Wide Web* (I. W. W. W. C. S. Committee, ed.), pp. 1242–1252, 2015.

[11] A. Clauset, C. Moore, and M. E. Newman, "Hierarchical structure and the prediction of missing links in networks," *Nature*, vol. 453, no. 7191, p. 98, 2008.

[12] R. Guimerà and M. Sales-Pardo, "Missing and spurious interactions and the reconstruction of complex networks," *Proceedings of the National Academy of Sciences*, vol. 106, no. 52, pp. 22073–22078, 2009.

[13] T. Kuwae, E. Miyoshi, S. Hosokawa, K. Ichimi, J. Hosoya, T. M. T. Amano, M. Kondoh, R. C. Ydenberg, and R. W. Elner, "Variable and complex food web structures revealed by exploring missing trophic links between birds and biofilm," *Ecology Letters*, vol. 15, no. 4, pp. 347–356, 2012.

[14] M. Kim and J. Leskovec, “The network completion problem: Inferring missing nodes and edges in networks,” in *Proceedings of the 2011 SIAM International Conference on Data Mining*, pp. 47–58, SIAM, 2011.

[15] V. Latora and M. Marchiori, “Vulnerability and protection of infrastructure networks,” *Physical Review E*, vol. 71, no. 1-015103, 2005.

[16] T. Boukhobza, F. Hamelin, and S. Martinez-Martinez, “State and input observability for structured linear systems: A graph-theoretic approach.,” *Automatica*, vol. 43, no. 7, pp. 1204–1210, 2007.

[17] J. van der Woude, “A graph-theoretic characterization for the rank of the transfer matrix of a structured system,” *Mathematics of Control, Signals and Systems*, vol. 4, pp. 33–40, Mar. 1991.

[18] T. Kaczorek, *Polynomial and Rational Matrices.* Communications and Control Engineering, Springer, London, 1 ed., 2007.

[19] J. Willems and J. Polderman, *Introduction to Mathematical Systems Theory*, vol. 26 of *Texts in Applied Mathematics.* Springer-Verlag New York, 1 ed., 1998.

[20] P. Antsaklis and A. N. Michel, *Linear Systems.* Birkhäuser Basel, 1 ed., 2006.

[21] K. Reinschke, F. Svaricek, and H.-D. Wend, “On strong structural controllability of linear systems,” in *Proceedings of the 31st IEEE Conference on Decision and Control*, vol. 1, pp. 203–208, 1992.

[22] H. Aling and J. M. Schumacher, “A nine-fold canonical decomposition for linear systems,” *International Journal of Control*, vol. 39, no. 4, pp. 779–805, 1984.

[23] H. J. van Waarde, P. Tesi, and M. K. Camlibel, “Necessary and Sufficient Topological Conditions for Identifiability of Dynamical Networks,” *arXiv e-prints*, p. arXiv:1807.09141, 2018.

[24] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas.* Princeton University Press, 2011.

[25] J. Jia, H. L. Trentelman, W. Baar, and M. K. Camlibel, “Strong Structural Controllability of Systems on Colored Graphs,” *arXiv e-prints*, p. arXiv:1810.05580, 2018.

[26] L. Dai, *Singular Control Systems*, vol. 118 of *Lecture Notes in Control and Information Sciences.* Springer-Verlag Berlin Heidelberg, 1 ed., 1989.