M 2014

U.PORTO

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

OUTPUT SELECTION FOR LARGE-SCALE STRUCTURAL SYSTEMS

A MATROID THEORY APPROACH

PEDRO MANUEL SABINO ROCHA

DISSERTAÇÃO DE MESTRADO APRESENTADA
À FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO EM
ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

MIEEC - MESTRADO INTEGRADO EM ENGENHARIA ELETROTÉCNICA E DE COMPUTADORES **2013/2014**

A Dissertação intitulada

"Output Selection for Large Scale Structural Systems: A matroid theory approach"

foi aprovada em provas realizadas em 23-07-2014

o júri

Presidente Professor Doutor Fernando Manuel Ferreira Lobo Pereira
Professor Catedrático do Departamento de Engenharia Eletrotécnica e de Computadores da Faculdade de Engenharia da Universidade do Porto

Professor Doutor Ricardo Jorge Aparício Gonçalves Pereira
Professor Auxiliar do Departamento de Matemática da Universidade de Aveiro

Professor Doutor António Pedro Rodrigues Aguiar
Professor Associado do Departamento de Engenharia Eletrotécnica e de Computadores da Faculdade de Engenharia da Universidade do Porto

O autor declara que a presente dissertação (ou relatório de projeto) é da sua exclusiva autoria e foi escrita sem qualquer apoio externo não explicitamente autorizado. Os resultados, ideias, parágrafos, ou outros extratos tomados de ou inspirados em trabalhos de outros autores, e demais referências bibliográficas usadas, são corretamente citados.

Pedro Manuel Sabino Rocha
Autor – Pedro Manuel Sabino Rocha

Faculdade de Engenharia da Universidade do Porto

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Output Selection for Large-Scale Structural Systems: a Matroid Theory Approach

Pedro Manuel Sabino Rocha

Mestrado Integrado em Engenharia Electrotécnica e de Computadores

Supervisor: António Pedro Aguiar (Associate Professor)

Co-supervisor: Paula Rocha Malonek (Full Professor)

July 31, 2014

Abstract

The question of how many outputs are needed and where to place them in order to fully observe a linear system has been a fundamental and challenging problem in control theory and applications of control systems. Structural system theory combined with graph theoretic tools has provided an efficient framework to answer this question for an equivalent class of systems, where properties are explored based on the structure of the system that corresponds to the location of zero and non-zero values. This approach deals well with applications to large-scale systems, where the systems' dimension has to be taken into account. Within this framework, for a given structure of the state matrix of a large scale linear system, a fundamental problem is the design issue of the structure of the output matrix such that overall the system is structurally observable with the restriction that each output be dedicated, i.e., it can only measure directly a single state variable. In the present work, we propose a novel approach to solve that problem based on matroid theory. Further, we make the connection of the obtained results with graph theory and consequently provide an efficient solution to find a minimum number of dedicated outputs that ensures structural observability. Moreover, we demonstrate that if we additionally impose a performance restriction given by the systems' generic observability index, then the problem falls in the NP-complete class. Several examples are presented that illustrate the derived results.

Resumo

O problema relacionado com a seleção das variáveis de saída, quantas são necessárias e onde devem ser colocadas, de forma a garantir observabilidade para um dado sistema linear continua a ser um problema fundamental e desafiante em teoria de controlo, bem como em aplicações a sistemas de controlo. A teoria estrutural aliada a ferramentas de teoria de grafos tem proporcionado um enquadramento eficiente para dar uma resposta adequada a esta questão quando se considera uma classe de sistemas, em que a propriedade de observabilidade é explorada tendo em consideração apenas a estrutura do sistema, isto é, a localização de zeros e não-zeros. Esta abordagem permite lidar com aplicações para sistemas de grande-escala, onde a dimensão do sistema é um factor a ter em conta. Dentro deste contexto e para uma dada estrutura da matriz de estado de um sistema linear de grande-escala, um problema fundamental consiste em projetar a estrutura da matriz de saída de tal forma que o sistema resultante seja estruturalmente observável, com a restrição de cada variável de saída ser dedicada, isto é, cada variável de saída pode medir apenas uma variável de estado. No presente trabalho, é proposta uma nova solução baseada em teoria de matróides para o problema mencionado. De seguida, é estabelecida a conexão entre os resultados obtidos e a teoria dos grafos e, consequentemente, é proposto um algoritmo eficiente para encontrar uma configuração de tamanho mínimo de variáveis de saída dedicadas que garantem observabilidade estrutural. Uma outra contribuição é que é demonstrado que se se impuser uma restrição de performance adicional, baseada no índice genérico de observabilidade, o problema torna-se NP-completo. Vários exemplos que ilustram os resultados obtidos são apresentados.

Agradecimentos

Gostaria de deixar um breve e, por isso, injusto agradecimento às seguintes pessoas. Ao professor António Pedro Aguiar, que sendo a força motriz deste trabalho, não colocou nenhum entrave às minhas deambulações intelectuais. Pelo contrário, sempre as acolheu, conseguindo extraír resultados de interessantíssima aplicabilidade. À professora Paula Rocha Malonek, que sempre participou com entusiasmo, brilhantismo e dedicação nas várias actividades deste trabalho. Convidou-me para o mundo da investigação, quando numa das suas aulas me entregou um livro de texto para ler em casa. Por isso, ficar-lhe-ei para sempre reconhecido. Ao aluno de doutoramento Sérgio Pequito, uma ajuda importantíssima sem a qual este trabalho não teria sido possível. Ensinou-me, com inúmeras sugestões e preciosos conselhos, a saber investigar.

Ao meus dois grandes amigos Sofia Araújo e Fábio Carneiro. À primeira, pelo seu eterno zelo; tanto dá, sem nada pedir em troca. Ao segundo, por estar sempre pronto a ouvir, tornando esta tarefa bastante mais prazenteira. Finalmente, à minha família sem a qual nada disto faria sentido. À minha mãe pela sua louvável determinação; o ser mais querido com o qual tive a oportunidade de ser educado.

Pedro Manuel Sabino Rocha

"I love deadlines.
I love the whooshing noise they make as they go by."

Douglas Adams

Contents

List of Figures

Abbreviations and Notation

FDOC	Feasible Dedicated Output Configuration
LTI	Linear Time-Invariant
SCC	Strong Connected Component
SCP	Set Covering Problem

$A \cup B$	Resulting set from the union of sets A and B
$A \cap B$	Resulting set from the intersection of sets A and B
$A \setminus B$	Resulting set from the difference between sets A and B
$A \triangle B$	Resulting set from the symmetric difference between sets A and B
$\|A\|$	Cardinality of set A
$\mathscr{X}$	Set of state variables
$\mathscr{Y}$	Set of output variables
$S_{\mathscr{O}}$	Set of state variables where dedicated outputs are placed
$\bar{M}$	Structural pattern of matrix M
$[\bar{M}]$	Classe of matrices that have the same dimension and structure of M
I_n	Identity matrix of size n
$grank(\bar{M})$	Generic rank of structural matrix $\bar{M}$

Chapter 1

Introduction

Over the last few decades, mankind has been experiencing an intensive and limitless desire to build and to control ever larger and more sophisticated systems. These systems of large-scale arise naturally in a variety of fields like biology [1], traffic systems [2] and wireless networks [3] to name a few. Common to the design of these systems is the question of observability and its degree, which provides a measure of how well internal states of a system can be inferred by knowledge of its external outputs. When designing the system's outputs, there are mainly two aspects to take into account. First, one needs to consider the cost that is associated with the output placement. Typically, we would like to use the lowest possible number of outputs to ensure observability. Secondly, consideration has to be paid to the system performance, which may depend on how fast the outputs can effectively observe the system. In this thesis, we will focus in those two problems.

When dealing with large scale systems, some aspects have to be taken into account. Since the dimension of the system is large, approaches that in general lead to an increase of the complexity of analysis and design have to be avoided. For instance, the verification that a given system is observable leads to some known numerical issues [4]. Additionally, the problem of finding the minimum number of inputs needed to control a linear system and, by duality, the minimum number of outputs needed to guarantee observability, has recently been shown to be an NP-hard problem [5].

In order to cope with the aforementioned difficulties, we will rely on structural systems theory which briefly consists in explore the system properties based only on the sparsity pattern, i.e, the zero/non-zero pattern of the state space representation matrices. The study of structural systems has started with Lin [6] when he first introduced the concepts of *structure* and *structural controllability* and provided the necessary and sufficient conditions to verify structural controllability of single-input linear systems. Later on, Shields and Pearson [7] and Glover and Silverman [8] extended Lin's results to multi-input linear systems. Since, within this theory, the only information kept is the zero/non-zero pattern, various system properties can be characterized in terms of quite simple properties of graph theory. Analysis using structural systems provides system-theoretic guarantees that hold for almost all values of the free parameters (the non-zeros) except for a manifold of zero Lebesgue measure [9], which may be further characterized algebraically.

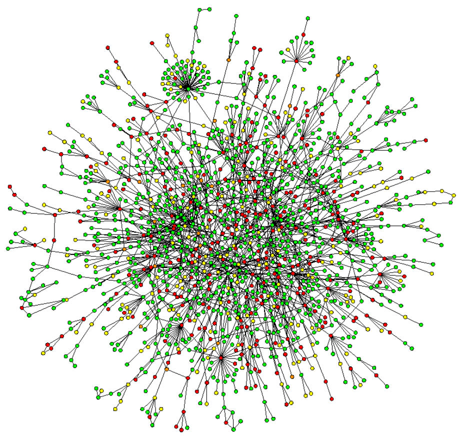


Figure 1.1: A representation of a large-scale network. There are various entities (colored dots) that interact with each other (represented by a black line connecting the dots). Source:http://www.nas.ewi.tudelft.nl/people/Dajie/images/randomgraph.gif

Furthermore, since many of the properties of the system can be translated into simple graph conditions, the computational burden is low and allows to deal with large-scale systems, specially if they are sparse.

Besides structural systems analysis, research work has also been made concerning system design. Systematic approaches to structured systems based design were investigated in different application scenarios as in [10] and [11]. In this work, we study the constrained output placement problem in which the outputs are dedicated, i.e, one output can only measure a single state variable. In [12] a similar study is carried out when the authors investigate the minimal actuator placement problem that arises from the control of biological complex networks. One of the problems addressed in that work is the question of finding the minimum number of dedicated inputs needed to ensure structural controllability (equivalent, by duality, to the problem studied in this thesis). However, the results in [12] hold only for the case in which the structural directed graph representation is strongly connected. In [13], a methodology that incurs polynomial complexity in the number of state variables to find the minimum number of dedicated outputs is proposed. Those results are extended in [14] where it is considered that there exists a cost associated with placing an output to a determined state variable and where the goal is to find the minimum number of dedicated outputs for structural observability.

In this thesis, we propose a novel way to solve the constrained output placement problem based on matroid theory. The concept of matroid was introduced by Whitney in 1935 to try to capture abstractly the essence of independence [15]. This abstraction embraces a diversity of combinatorial structures like graphs and zero-one matrices. Furthermore, matroid theory arises naturally in combinatorial optimization since matroids are precisely the structures for which the greedy algorithm works [16]. Within structural systems theory, the matroid abstraction was used by Murota [17] and Clark [18]. The former when dealing with state space representation matrices that may contain three types of entries, namely fixed zeros and nonzeros divided into free parameters and fixed nonzero constants. The latter when demonstrating that some problems within structural systems can be posed within a matroid optimization framework. Structural systems analysis lies on both graph and matrices concepts. Due to its combinatorics nature, it seems plausible to try to apply matroid theory to solve some problems in an efficiently manner. In the present work, the constrained output placement problem is reformulated as a matroid intersection problem. Within this framework, the generalization to the case where there are costs associated with the output placement is straightforward as the generalization of the matroid intersection algorithm to the weighted matroid intersection.

Within structural systems theory, some questions related with the generic observability indices remained open as noticed in [19] and [20]. Those questions are directly related with the notion of *how fast* the initial state vector is recovered for a given network and it has enormous importance, particularly when we would like to build systems with the ability of performing real-time monitoring. In [21], the structural counterpart of the observability indices are introduced. In [22], a new methodology is proposed for the characterization and computation of the controllability indices of structured systems. An improved upper bound of the generic observability index is given in [23] based on graph representations. However, there is a lack of methods concerning system design when it is imposed a performance restriction translated by the systems' generic observability index. In this work, we show that the problem of finding a minimum dedicated output configuration such that the generic observability index is less than some given constant falls in the NP-complete class.

Main contributions

The main thesis contributions are summarized below.

- We propose a novel way to solve the optimal constrained output placement problem for large-scale structural linear systems based on matroid theory.
- We make the connection of the obtained results with graph theory and consequently simplify the complexity of the matroid intersection algorithm.
- We show that the output placement problem with generic observability index constraint is NP-complete.
- We illustrate the derived results through several examples: a 6-node network example, a wireless sensor application, and simulations with 100-node random networks.
- We provide alternative proofs to some well-known results.

Organization of the Dissertation

The thesis is organized as follows. In chapter 2, we briefly review some fundamental concepts from linear systems theory after which some important results are introduced. We also present an algorithm to test structural observability and to compute the generic dimension of the unobservable subspace when the system is not structurally observable.

In chapter 3, the main problem is formulated. Then, we introduce some tools and concepts from matroid theory that will be applied to the main problem. Next, we reformulate our problem as an a intersection of two matroids. Finally, a new algorithm to find a minimum dedicated output configuration that ensures structural observability is presented.

In chapter 4, the constrained output placement with generic observability index is precisely formulated. After, we demonstrate that this problem is NP-complete.

Finally, in chapter 5, we start by illustrating the concepts regarding matroid theory with an example. After that, the algorithm developed in chapter 3 to find a minimum-cost feasible dedicated output configuration is applied to a spatially distributed sensor network. In the last section, some simulation results are presented.

Chapter 2

Structural Systems Background

This chapter starts with a review of some fundamental concepts from linear systems theory taken from [24] and [25], after which some important results on structural systems theory from [19] are introduced. Only our own alternative proofs are presented.

2.1 Linear Systems

In many applications, dynamical systems can be modeled by a state space representation where an intermediate state variable is introduced in the description of the relation between the input and output variables. In what follows, attention will be paid to discrete linear time-invariant systems (LTI), i.e, systems that can be represented by a discrete-time linear state space model of the form

$$x(t+1) = Ax(t) + Bu(t), \tag{2.1}$$

$$y(t) = Cx(t) + Du(t), \tag{2.2}$$

where the signals $u \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, and $y \in \mathbb{R}^p$ are the *input*, the *state* and the *output* of the system, respectively, and $t \in \mathbb{N}$ is the *iteration instant*. The first equation expresses a relation between the input and the state and it is called the *state equation*, while the second one is called the *output equation*.

Following our motivation, attention will be paid to the scenario where there are no input signals affecting the system. In that case, the matrices B and D are identically zero, yielding the pair of equations

$$x(t+1) = Ax(t), \tag{2.3}$$

$$y(t) = Cx(t). \tag{2.4}$$

Matrix A from (2.3) is usually referred as the *state matrix* whereas matrix C from (2.4) is commonly referred as the *output matrix*.

Lemma 1 (Solution of the system (2.3)-(2.4)) *The solution $y(t)$ of the system described by (2.3)-(2.4) with initial condition $x(0) \in \mathbb{R}^n$ is given by*

$$y(t) = CA^t x(0), \quad \forall t \in \mathbb{N}. \tag{2.5}$$

$\diamond$

Proof By induction one easily verifies that $x(t) = A^t x(0)$. Thus, $y(t) = Cx(t) = CA^t x(0)$.

Considering the system described by the equations (2.3)-(2.4), it is sometimes desirable to obtain the initial state vector, $x(0)$, from the output vector values $y(t)$, for $t \in \mathbb{N}$. In other words, we would like to solve (2.5) for the unknown $x(0) \in \mathbb{R}^n$. Initial state vectors that cannot be obtained in this way are called unobservable. This is formalized in the following definition.

Definition 1 (Unobservable Subspace) *Given a time instant $t_1 \in \mathbb{N}$, the **unobservable subspace** at t_1, $\mathscr{U}_{\mathscr{O}}(t_1)$, consists of all states $x(0) \in \mathbb{R}^n$ for which*

$$CA^t x(0) = 0, \quad \forall t = 0, 1, \ldots, t_1.$$

$\diamond$

Having all the output vectors from $t = 0$ until $t = t_1$ one can write

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(t_1) \end{bmatrix} = \mathscr{O}(t_1)x(0),$$

where $\mathscr{O}(t_1) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{t_1-1} \end{bmatrix}$ is called the *observability matrix* at time instant t_1. This provides us with an alternative way of expressing the unobservable subspace since

$$\mathscr{U}_{\mathscr{O}}(t) = ker\ \mathscr{O}(t), \tag{2.6}$$

where, given $M \in \mathbb{R}^{n\times m}$, $ker\ M$ denotes the *kernel* of M, i.e, $ker\ M = \{v \in \mathbb{R}^m \mid Mv = 0\}$.

It is possible to notice that the unobservable subspace for iteration instants $t \geq n$, where n is the size of the square matrix A, is equal to the unobservable subspace at iteration instant n. In order to perceive this fact, one needs the following theorem.

Theorem 1 (Cayley-Hamilton Theorem) *Let $M \in \mathbb{R}^{n\times n}$ be a square matrix, I_n be the identity matrix of size n, and $p(\lambda) = det(\lambda I_n - M)$ the characteristic polynomial of M, where det is the* *determinant operator. Then*

$$p(M) = \mathbf{O}.$$

◇

Since $p(\lambda) = \sum_{i=0}^{n} p_i \lambda^i$ for suitable values of $p_i \in \mathbb{R}$, $i = 0, \ldots, n$ with $p_n = 1$, $p(M) = \mathbf{0}$ means that $M^n = -\sum_{i=0}^{n-1} p_i M^i$. This implies that the k-th power of a matrix $M \in \mathbb{R}^{n\times n}$, with $k \geq n$, can be expressed as a linear combination of $M^0, \ldots, M^{n-1}$. Thus, it follows that *ker* $\mathscr{O}(t) =$ ker $\mathscr{O}(n)$, $\forall t \geq n$. This motivates the following definition.

Definition 2 (Observable System) *The system described by* (2.3)-(2.4) *or, equivalently, the pair* (A, C) *is said to be* ***observable*** *if and only if the unobservable subspace at iteration instant n is composed only of the zero vector, i.e.,*

$$\mathscr{U}_{\mathscr{O}}(n) = \{\mathbf{0}\}. \tag{2.7}$$

◇

If the system is observable, then there is no ambiguity in determining $x(0)$ from the output values. This is equivalent to say that *ker* $\mathscr{O}(n) = \{\mathbf{0}\}$ and that *rank* $\mathscr{O}(n) = n$. Matrix $\mathscr{O}(n)$ is called the *observability matrix* and it is simply denoted by $\mathscr{O}$. Therefore, we can compute the rank of the observability matrix in order to test whether or not the system is observable. However, there are other ways to test observability.

Theorem 2 (Popov-Belevitch-Hautus test for Observability) *The system described by* (2.3)-(2.4) *is observable if and only if*

$$rank\left(\begin{bmatrix} A - \lambda I_n \\ C \end{bmatrix}\right) = n, \quad \forall \lambda \in \mathbb{C}. \tag{2.8}$$

◇

Example 1 *Consider the following LTI system:*

$$\begin{aligned} x(t+1) &= \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 1 & 0 \end{bmatrix} x(t), \\ y(t) &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} x(t), \end{aligned}$$

where $x(t) \in \mathbb{R}^3$ *and* $y(t) \in \mathbb{R}$. *The observability matrix is given by*

$$\mathscr{O} = \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 1 & 0 \end{bmatrix},$$

and hence we have that $\text{rank}(\mathscr{O}) = 2$*, which allow us to verify that the system is not observable. Furthermore, the unobservable subspace is*

$$\mathscr{U}_{\mathscr{O}} = \begin{bmatrix} 0 \\ 0 \\ a \end{bmatrix}, \forall a \in \mathbb{R}.$$

With the Popolov-Belevitch-Hautus (PBH) test, the same conclusions can be drawn. In fact, for $\lambda = 1$

$$rank\left(\begin{bmatrix} A - \lambda I_n \\ C \end{bmatrix}\right) = rank\left(\begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}\right) = 2,$$

which is less than $n = 3$*.*

Assuming, now, that the system described by (2.3)-(2.4) is observable, it is sometimes pertinent knowing how many iterations one has to wait until it is possible to recover the initial state vector $x(0)$ from the outputs. This question is directly related with the rank of the observability matrix.

Definition 3 (Observability Index) *Given a system described by* (2.3)-(2.4)*, let* μ *be the smallest iteration instant for which the observability matrix has full rank, i.e.,*

$$\mu = min\ \{k \in \mathbb{N} \mid rank[\mathscr{O}(k)] = n\}. \tag{2.9}$$

Integer μ *is called the* ***observability index*** *of the system.* ◇

Assuming that the pair (A, C) is observable, one may also assume that the matrix C has rank equal to p (full row rank). If it were not the case, then some row of C could be written as a linear combination of other rows and by deleting that row (i.e, the effect of that output) observability would not be affected.

If the observability matrix $\mathscr{O}$ has rank n, then it has n linearly independent rows. Let c_i be the

i-th row associated with output number i, $i = 1, \ldots, p$. Then $\mathcal{O}$ can be written explicitly as

$$\mathcal{O} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \\ \cdots \\ \vdots \\ \cdots \\ c_1 A^{n-1} \\ \vdots \\ c_p A^{n-1} \end{bmatrix},$$

and the search of the linearly independent rows can be performed from top to bottom. Because of the pattern of $\mathcal{O}$, once $c_i A^k$ depends on its top rows, the same will happen with $c_i A^j$, for j greater than k. This means that once a row associated with c_i becomes linearly dependent from the previous ones, then all rows associated with c_i thereafter are linearly dependent from the previous ones. The number μ_i of linearly independent rows associated with c_i in $\mathcal{O}$ is called the i-th observability index. Since $\mathcal{O}$ has rank n, one has that

$$\sum_{i=1}^{p} \mu_i = n.$$

The set $\{\mu_1, \ldots, \mu_p\}$ is called *observability indices* set and one can easily verify that $\mu = max(\mu_1, \ldots, \mu_p)$ is the *observability index*. Therefore, the shortest possible μ occurs when $\mu_1 = \mu_2 = \ldots = \mu_p$ while the largest one occurs when all μ_i except one equals one. With that in mind, it is easy to conclude that

$$\left\lceil \frac{n}{p} \right\rceil \leq \mu \leq n - p + 1, \tag{2.10}$$

where, for $x \in \mathbb{R}$, $\lceil x \rceil$ is the smallest integer not less than x.

Example 2 *Consider again Example 1. Another output variable can be added in order to guarantee observability of the system. For instance, adding the row* $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$ *to the matrix C, which means that now also the third state component is measured, yields*

$$C_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Under the PBH criteria for observability (2), it can be concluded that (A, C_2) *is now observable.*

In fact

$$rank\left(\begin{bmatrix} A-\lambda I_3 \\ C_2 \end{bmatrix}\right) = rank\left(\begin{bmatrix} -\lambda & 0 & 0 \\ \frac{1}{2} & 1-\lambda & 0 \\ \frac{1}{2} & 1 & -\lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 3,\ \forall \lambda \in \mathbb{C},$$

since rows 3*,* 4 *and* 5 *are linearly independent for any choice of* λ*.*

From inequality (2.10) *one may conclude that for the pair* (A, C_2) *the observability index* μ *has to be* 1 *or* 2*. Since* $rank(C_2) = 2 < 3$*, the state* $x(0)$ *cannot be recovered by the new output* $y(0)$*. Thus,* $\mu \neq 1$ *and hence* $\mu = 2$*.*

2.2 Structural Systems

As it was seen in the previous section, the methods presented that deal with the state-space model assume full knowledge of the entries of the matrices that describe the relationship between the variables. However, because of the nature of physical systems, experience shows that there is uncertainty and variability when someone is faced with the task of modeling these systems. Additionally, in the case of large-scale systems, the order of the matrices increases rapidly and, obviously, this increase may lead to a high computationally burden.

The study of structural systems, started with Lin [6], tries to overcome the aforementioned difficulties. It captures the generic properties of the linear systems because it no longer pays attention to the numerical values of the entries of the matrices A, B, C and D in the state-space model (2.1) -(2.2). Instead, each entry is either a fixed zero or a free parameter. It is a fixed zero if there is no relationship between the corresponding variables and a free parameter, otherwise. That is, for any real matrix $M \in \mathbb{R}^{n\times m}$, one can abstract from their numerical entries and consider only the zero/non-zero pattern.

Definition 4 *Given a matrix* $M \in \mathbb{R}^{n\times m}$ *its* ***structural counterpart*** $\bar{M} \in \{0,1\}^{n\times m}$ *is the zero/one matrix that verifies*

$$\bar{M}_{i,j} = 1 \text{ if and only if } M_{i,j} \neq 0,\ \forall i, j.$$

Another matrix $M_2 \in \mathbb{R}^{n\times m}$ *is said to be* ***admissible realization*** *to* M *if both belong to the same equivalence class* $[\bar{M}]$ *defined as*

$$[\bar{M}] = \{M \in \mathbb{R}^{n\times m} \mid M_{i,j} = 0 \text{ if } \bar{M}_{i,j} = 0,\ \forall i, j\}. \tag{2.11}$$

◇

In order to evaluate the qualitative properties of a LTI system it will be useful to consider only the zero/non-zero pattern of the real matrices A and C from the pair of equations (2.3)-(2.4).

Definition 5 (Structural System) *Given a linear system as in* (2.3)-(2.4)*, the corresponding* ***structural system*** *is obtained by replacing the matrices A and C by their structural counterpart, i.e by* $\bar{A}$ *and* $\bar{C}$*, respectively.* ◇

Example 3 *Consider the LTI system given in Example 1, its structural counterpart is given by*

$$\bar{A} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \text{ and } \bar{C} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}.$$

Due to their nature, the properties of structural systems can be efficiently studied under a combinatorics framework. In fact, the system information can be visualized with a directed graph and many of the system properties can be translated into graph conditions.

Definition 6 (Directed Graph) *A* ***directed graph*** *or* ***digraph*** $\mathscr{D} = (\mathscr{V}, \mathscr{E})$ *consists of a finite set* $\mathscr{V}$ *of elements called vertices and a set* $\mathscr{E} \subseteq \mathscr{V} \times \mathscr{V}$ *of ordered pairs called arcs.* ◇

Definition 7 (Structural Directed Graph Representation) *For a system described by (2.3)-(2.4) and whose structural counterpart is given by* $\bar{A} \in \{0,1\}^{n\times n}$ *and* $\bar{C} \in \{0,1\}^{p\times n}$*, the associated* ***structural directed graph representation*** *is obtained as*

$$\mathscr{D}(\bar{A}, \bar{C}) = (\mathscr{X} \cup \mathscr{Y}, \mathscr{E}_{\mathscr{X},\mathscr{X}} \cup \mathscr{E}_{\mathscr{X},\mathscr{Y}}),$$

where $\mathscr{X} = \{x_1, \ldots, x_n\}$ *and* $\mathscr{Y} = \{y_1, \ldots, y_p\}$ *are the set of the state and output vertices, respectively. The arc set is the union of the following sets:*

$$\mathscr{E}_{\mathscr{X},\mathscr{X}} = \{(x_j, x_i) \mid \bar{A}_{i,j} = 1, \forall i, j = 1, \ldots, n\},$$
$$\mathscr{E}_{\mathscr{X},\mathscr{Y}} = \{(x_j, y_i) \mid \bar{C}_{i,j} = 1, \forall j = 1, \ldots, n, \forall i = 1, \ldots, p\}.$$

◇

Notice that there is a one-to-one correspondence between the structural system and its directed graph representation. The edges set $\mathscr{E}_{\mathscr{X},\mathscr{X}}$ and $\mathscr{E}_{\mathscr{X},\mathscr{Y}}$ contain all the information about matrices $\bar{A}$ and $\bar{C}$, respectively.

Example 4 *Consider the structural system with* $\bar{A}$ *and* $\bar{C}$ *given by*

$$\bar{A} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \text{ and } \bar{C} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

The structural directed graph representation $\mathscr{D}(\bar{A}, \bar{C})$ *can be seen in Figure 2.1.*

Structural systems theory provides a useful framework to characterize generic properties of linear systems. A property is said to hold *structurally* within the same equivalence class (see

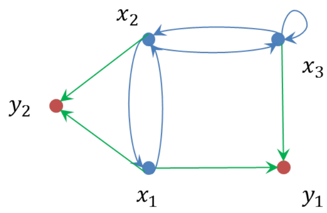


Figure 2.1: A structural directed graph representation. The state vertices are represented in blue while the output vertices are represented in red. The blue arcs belong to $\mathscr{E}_{\mathscr{X},\mathscr{X}}$ while the green ones belong to $\mathscr{E}_{\mathscr{X},\mathscr{Y}}$.

Definition 4) if it holds numerically for *almost all* admissible numerical realizations, i.e, if $\mathscr{P}$ is the property and $\bar{M}$ a structural matrix that satisfies $\mathscr{P}$, then the set $\{M \in [\bar{M}] \mid M$ does not satisfy $\mathscr{P}\}$ has zero Lebesgue measure [9]. To give an example, the concept of generic rank is introduced next.

Definition 8 (Generic Rank) *Given any structural matrix $\bar{M} \in \{0,1\}^{n\times m}$, the **generic rank** is defined as*

$$grank(\bar{M}) = \max_{M\in[\bar{M}]} rank(M). \tag{2.12}$$

$\diamond$

Example 5 *For*

$$\bar{M} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix},$$

it can be concluded that $grank(\bar{M}) = 2$. In fact, the admissible realization

$$M = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 10 & 0 \end{bmatrix}$$

has rank 2 and it is impossible to have a higher rank because of the null pattern of the first row. All admissible realizations concerning the structure of $\bar{M}$ can be written in the form:

$$M = \begin{bmatrix} 0 & 0 & 0 \\ a & b & 0 \\ c & d & 0 \end{bmatrix},$$

with a, b, c and $d \in \mathbb{R}$. Thus, it is possible to conclude that the realizations that do not have rank equal to 2 lie on a proper variety of $\mathbb{R}^{3\times 3}$ defined by the condition $ad = bc$.

Since attention had been paid only to the structure of the systems, it would be important to introduce the structural counterpart of some of the concepts presented in the previous section. One of these concepts, observability, has an intuitive generalization.

Definition 9 (Structural Observability) *The pair* $(\bar{A},\bar{C})$, $\bar{A} \in \{0,1\}^{n\times n}$, $\bar{C} \in \{0,1\}^{p\times n}$, *is said to be* ***structurally observable*** *if and only if there exists an admissible pair* (A,C), *with* $A \in [\bar{A}]$ *and* $C \in [\bar{C}]$, *such that* (A,C) *is observable.*

Conditions for structural observability can be stated in terms of the matrix pair $(\bar{A},\bar{C})$. To formulate these conditions, the following definitions introduced by Lin [6] have to be presented.

Definition 10 (Form I) *The structural pair* $(\bar{A},\bar{C})$,*with* $\bar{A} \in \{0,1\}^{n\times n}$ *and* $\bar{C} \in \{0,1\}^{p\times n}$, *is said to be reducible or to be in form I if there exists a permutation*[1] *matrix P such that*

$$P^T\bar{A}P = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{bmatrix} \text{ and } \bar{C}P = \begin{bmatrix} 0 & \bar{C}_2 \end{bmatrix}, \tag{2.13}$$

where $\bar{A}_{ij}$ *is an* $n_i \times n_j$ *matrix for appropriate* $i, j = 1,2$, *with* $0 < n_1 \leq n$ *and* $n_1 + n_2 = n$, *where* $\bar{C}_2$ *is an* $p \times n_2$ *matrix.* ◇

Definition 11 (Form II) *The structural pair* $(\bar{A},\bar{C})$, *with* $\bar{A} \in \{0,1\}^{n\times n}$ *and* $\bar{C} \in \{0,1\}^{p\times n}$, *is said to be in form II if the inequality holds*

$$grank\left(\begin{bmatrix} \bar{A} \\ \bar{C} \end{bmatrix}\right) < n. \tag{2.14}$$

◇

Example 6 *Consider the structural pair*

$$\bar{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \bar{C} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}.$$

Let, for instance, P be the following permutation matrix:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then, we have that

$$P^T\bar{A}P = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \text{ and } \bar{C}P = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}.$$

[1] A permutation matrix is a square zero-one matrix that has exactly one non-zero entry in each row and each column and zeros elsewhere.

Therefore, the pair $(\bar{A}, \bar{C})$ *can be reduced to form I. On the other hand,*

$$grank\left(\begin{bmatrix}\bar{A}\\ \bar{C}\end{bmatrix}\right) = 4 = n,$$

and, therefore, the pair is not in form II.

If the pair $(\bar{A}, \bar{C})$, after a suitable permutation, can be lead to one of the previous forms, one may conclude that that pair is not structurally observable. In fact, sufficiency also holds [6], [7] and [8].

Theorem 3 *A pair* $(\bar{A}, \bar{C})$ *is structurally observable if and only if it has neither form I nor form II.* ◇

One may verify that forms I and II are purely structural forms concerning the distribution of zeros and ones of the pair $(\bar{A}, \bar{C})$. Thus, these forms can be translated into graph conditions. This helps to visualize the concepts regarding structural observability. Before stating the graph criteria that will allow us to infer about the presence or absence of structural observability, it will be useful to introduce some definitions.

Definition 12 (Path) *Let* $\mathscr{D} = (\mathscr{V}, \mathscr{E})$ *be a directed graph. A* ***path*** P *is a sequence of vertices* $P = v_1, \ldots, v_k$*, with* $k \geq 2$*, such that* $(v_i, v_{i+1}) \in \mathscr{E}$*,* $\forall i = 1, \ldots, k-1$*. Vertex* v_1 *is called the root of the path, while* v_k *is called the tip. If there exists a path between* v_i *and* v_j*, this will be represented by* $v_i \to v_j$*. The size of the path is defined as the number of arcs used. If there exists a path of size* k *between* v_i *and* v_j*, this is written as* $v_i \xrightarrow{k} v_j$*.* ◇

Definition 13 (Elementary Path) *An* ***elementary path*** *is a path where the vertices are all distinct.* ◇

Definition 14 (Cycle) *A* ***cycle*** *is path where all the vertices are distinct except for the first one and the last one, which coincide.* ◇

Definition 15 *The pair* $(\bar{A}, \bar{C})$ *is said to be* ***output-connected*** *if in the associated digraph* $\mathscr{D}(\bar{A}, \bar{C})$ *there exists a path from every* $x_i \in \mathscr{X}$ *to some* $y_j \in \mathscr{Y}$*.* ◇

The previous definition is directly related with form I (Definition 10) by the following proposition.

Proposition 1 *The pair* $(\bar{A}, \bar{C})$ *is output-connected if and only if it is not reducible, i.e., is not in form I.* ◇

Proof If the pair $(\bar{A}, \bar{C})$ is not output-connected, then the set $\mathscr{X}$ can be written as the union of two disjoint sets

$$\mathscr{X} = \mathscr{X}_1 \cup \mathscr{X}_2, \text{ with } \mathscr{X}_1 \cap \mathscr{X}_2 = \emptyset,$$

(a) digraph with a contraction. (b) digraph without contractions.

Figure 2.2: A directed graph with a) a contraction and b) with that contraction removed.

where for every $x_i \in \mathscr{X}_1$ it is impossible to have a path for any $y_j \in \mathscr{Y}$, $\forall j = 1,\dots,p$. After an appropriate permutation, we may assume that $\mathscr{X}_1 = \{x_1,\dots,x_{n_1}\}$ and that $\mathscr{X}_2 = \{x_{n_1+1},\dots,x_n\}$, with $n_1 > 0$. Considering the definition of $\mathscr{X}_1$: $(x_i,y_j) \notin \mathscr{E}$, $\forall x_i \in \mathscr{X}_1$, $\forall y_j \in \mathscr{Y}$. Thus, if we write $\bar{C}$ as $\bar{C} = \begin{bmatrix}\bar{C}_1 & \bar{C}_2\end{bmatrix}$, with $\bar{C}_1 \in \{0,1\}^{p\times n_1}$ and $\bar{C}_2 \in \{0,1\}^{p\times n_2}$, for $n_2 = n - n_1$, it must be that $\bar{C}_1 = \mathbf{0}$. At the same time: $(x_i,x_j) \notin \mathscr{E}$, $\forall x_i \in \mathscr{X}_1$, $\forall x_j \in \mathscr{X}_2$. Otherwise, since there is a path from any $x_j \in \mathscr{X}_2$ to some $y_m \in \mathscr{Y}$, there would exist a path from $x_i \in \mathscr{X}_1$ to y_m (by transitivity) contradicting our assumption about $\mathscr{X}_1$. Thus, $\bar{A}$ can be written as:

$$\bar{A} = \begin{bmatrix}\bar{A}_{11} & \bar{A}_{12}\\ \bar{A}_{21} & \bar{A}_{22}\end{bmatrix},$$

where $\bar{A}_{ij} \in \{0,1\}^{n_i\times n_j}$ and it must be that $\bar{A}_{21} = \mathbf{0}$. Therefore, it is possible to reduce the system to form I.

If the system is in form I, then with similar considerations it is possible to conclude that there exists at least one $x_i \in \mathscr{X}$ such that there is no path between x_i and y_j, $\forall y_j \in \mathscr{Y}$ and, therefore, the system is not output-connected. □

While the fact that the system cannot be reduced to form I means that there must exist a path from each state variable to some output variable, form II deals with the distribution of the non-zero entries in the observability matrix. In order to attain full rank, it must be possible to obtain one non-zero entry per each column in a different row of the observability matrix. This is directly related with the following definitions.

Definition 16 (Target Set) *Considering the digraph $\mathscr{D} = (\mathscr{V},\mathscr{E})$, the **target set** of a vertex v_i is defined as*

$$T(v_i) = \{v_j \in \mathscr{V} \mid (v_i,v_j) \in \mathscr{E}\}.$$

The definition can be extended to a set of vertices $\mathscr{V}_i \subseteq \mathscr{V}$ in the following manner:

$$T(\mathscr{V}_i) = \bigcup_{v_i\in\mathscr{V}_i} T(v_i).$$

◇

Definition 17 (Contraction) *Let $\mathscr{D}(\bar{A},\bar{C})$ be the digraph associated with the pair $(\bar{A},\bar{C})$. Then, $\mathscr{D}(\bar{A},\bar{C})$ is said to have a **contraction** if there exists $\mathscr{X}_s \subseteq \mathscr{X}$ such that*

$$|T(\mathscr{X}_s)| < |\mathscr{X}_s|. \tag{2.15}$$

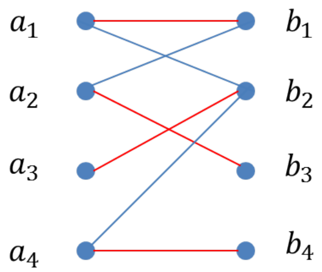


Figure 2.3: A bipartite graph $\mathscr{B} = (\mathscr{V}_1, \mathscr{V}_2, \mathscr{E})$ where $\mathscr{V}_1 = \{a_1, a_2, a_3, a_4\}$ and $\mathscr{V}_2 = \{b_1, b_2, b_3, b_4\}$. The edges are represented in blue and red. The red ones belong to a maximal matching.

$\diamond$

Example 7 *Figure 2.2 illustrates the contraction concept. The digraph represented on the left has a contraction since $|T(\{x_1, x_2\})| = |\{y_1\}| = 1$. Notice, also, that*

$$grank\left(\begin{bmatrix}\bar{A}\\ \bar{C}\end{bmatrix}\right) = \left(\begin{bmatrix}0 & 0\\ 0 & 0\\ 1 & 1\end{bmatrix}\right) = 1,$$

which is less than $n = 2$. On the other and, if, for instance, a self-loop is added to the state vertex x_1, the contraction is removed. Notice that now the cardinality of the target set is $|T(\{x_1, x_2\})| = |\{x_1, y_1\}| = 2$ and

$$grank\left(\begin{bmatrix}\bar{A}\\ \bar{C}\end{bmatrix}\right) = \left(\begin{bmatrix}1 & 0\\ 0 & 0\\ 1 & 1\end{bmatrix}\right) = 2.$$

In order to establish the relationship between contractions and the form II from Definition 11, it will be useful to introduce some concepts regarding bipartite graphs and matching theory.

Definition 18 (Bipartite Graph) *A* ***bipartite graph*** *$\mathscr{B} = (\mathscr{V}_1, \mathscr{V}_2, E)$[2] is an undirected graph whose vertex set can be partitioned into two disjoint sets, $\mathscr{V}_1$ and $\mathscr{V}_2$, such that $E \subseteq \{(v_1, v_2) \mid v_1 \in \mathscr{V}_1,\ v_2 \in \mathscr{V}_2\}$.* $\diamond$

Definition 19 (Matching) *Considering a bipartite graph $\mathscr{B} = (\mathscr{V}_1, \mathscr{V}_2, E)$, a* ***matching*** *$M$ is a subset of the edge set such that every edge in M shares no vertex with any other edge in M. A matching M is* ***maximal*** *if for every other matching M', we have $|M| \geq |M'|$. A matching M is said to* ***cover*** *$\mathscr{V}_1$ (or $\mathscr{V}_2$) if for every $v_1 \in \mathscr{V}_1$ (respectively, $v_2 \in \mathscr{V}_2$) there is an edge $e \in M$ such that $e = (v_1, v_2)$ for some $v_2 \in \mathscr{V}_2$ (respectively, for some $v_1 \in \mathscr{V}_1$).* $\diamond$

When dealing with matchings on bipartite graphs, a pertinent question that may arise is if a given bipartite graph $\mathscr{B} = (\mathscr{V}_1, \mathscr{V}_2, E)$ possesses or not matching that covers $\mathscr{V}_1$ or $\mathscr{V}_2$. As

[2]The variable $\mathscr{E}$ will be used to denote a set of arcs, i.e, oriented pairs of vertices, whereas the variable E will be used to note a set of edges, i.e, non-oriented pairs of vertices.

an example, it can be seen that the bipartite graph depicted in Figure 2.3 has a matching $M = \{(a_1,b_1),(a_2,b_3),(a_3,b_2),(a_4,b_4)\}$ that covers both $\mathscr{V}_1 = \{a_1,a_2,a_3,a_4\}$ and $\mathscr{V}_2 = \{b_1,b_2,b_3,b_4\}$. The following theorem provides necessary and sufficient conditions to guarantee that a bipartite graph $\mathscr{B} = (\mathscr{V}_1,\mathscr{V}_2,E)$ possesses a matching that covers $\mathscr{V}_1$.

Theorem 4 (Hall Marriage Theorem) *Consider a bipartite graph $\mathscr{B} = (\mathscr{V}_1,\mathscr{V}_2,E)$ and $A \subseteq \mathscr{V}_1$. Let*

$$J(A) = \{v_2 \in \mathscr{V}_2 \mid (v_1,v_2) \in E \text{ for some } v_1 \in A\}.$$

Then, a matching that covers $\mathscr{V}_1$ exists if and only if, $\forall A \subseteq \mathscr{V}_1$:

$$|A| \leq |J(A)|. \tag{2.16}$$

◇

Similar to the structural directed graph representation, one can build a bipartite graph associated with the structural pair $(\bar{A},\bar{C})$.

Definition 20 (Structural Bipartite Graph Representation) *For a system described by (2.3)-(2.4) and whose structural counterpart is given by $\bar{A} \in \{0,1\}^{n\times n}$ and $\bar{C} \in \{0,1\}^{p\times n}$, the associated* ***structural bipartite graph representation*** *is the bipartite graph $\mathscr{B}(\bar{A},\bar{C}) = (\mathscr{X}^-,\mathscr{X}^+ \cup \mathscr{Y}, E_{\mathscr{X}^-,\mathscr{X}^+} \cup E_{\mathscr{X}^-,\mathscr{Y}})$ where $\mathscr{X}^- = \{x_1^-,\ldots,x_n^-\}$, $\mathscr{X}^+ = \{x_1^+,\ldots,x_n^+\}$ and $\mathscr{Y} = \{y_1,\ldots,y_p\}$. The set of edges is the union of the following sets:*

$$E_{\mathscr{X}^-,\mathscr{X}^+} = \{(x_j^-,x_i^+) \mid \bar{A}_{ij} = 1,\ \forall i,j = 1,\ldots,n\},$$
$$E_{\mathscr{X}^-,\mathscr{Y}} = \{(x_j^-,y_i) \mid \bar{C}_{ij} = 1,\ \forall j = 1,\ldots,n, \forall i = 1,\ldots,p\}.$$

◇

Now, we are ready to present the graph criteria equivalent to form II (Definition 11).

Proposition 2 *The directed graph $\mathscr{D}(\bar{A},\bar{C})$ is free of contractions if and only if the pair $(\bar{A},\bar{C})$ is not in form II.* ◇

Proof In order to prove this claim, consider the bipartite graph $\mathscr{B}(\bar{A},\bar{C})$ associated with the structural pair $(\bar{A},\bar{C})$ and the structural directed graph representation $\mathscr{D}(\bar{A},\bar{C})$. If the digraph $\mathscr{D}(\bar{A},\bar{C})$ is free of contractions, then $\forall \mathscr{X}_s \subseteq \mathscr{X}^-$: $|\mathscr{X}_s| \leq |T(\mathscr{X}_s)|$. Considering the bipartite graph and the Definition 16 of target set, one may write that $\forall A \subseteq \mathscr{V}_1$: $|A| \leq |J(A)|$ with $J(A)$ defined as in Theorem 4. Thus, with Hall Marriage Theorem in mind, it is possible to conclude that there exists a matching in $\mathscr{B}(\bar{A},\bar{C})$ that covers $\mathscr{X}^-$. On the other hand, the existence of that matching tell us that it is possible to find one non-zero entry per column in different rows of the composite

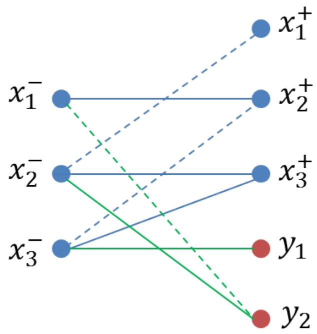


Figure 2.4: A bipartite graph representation associated with the digraph represented in Figure 2.1. The edges that belong to the maxmimal matching are represented by dashed lines.

matrix $M = \begin{bmatrix} \bar{A} \\ \bar{C} \end{bmatrix}$. To see that, notice that if (x_j^-, x_i^+) or (x_j^-, y_k) belong to the maximal matching, then $\bar{A}_{ij} = 1$ or $\bar{C}_{kj} = 1$, respectively. Since, by definition of that matching, the whole set $\mathscr{X}^-$ is covered, for each column of M it is possible to find one non-zero entry such that two different non-zero entries lie on different rows. Thus, it can be concluded that $grank(M) = n$ and, therefore, $(\bar{A}, \bar{C})$ is not in form II.

With analogous arguments, if $(\bar{A}, \bar{C})$ is not in form II, it is possible to see that $\mathscr{B}(\bar{A}, \bar{C})$ has a maximal matching and, again, with Hall Marriage Theorem, one concludes that the digraph associated is contraction-free. □

With the previous results, the sufficient and necessary conditions for structural observability can be stated with respect to the associated digraph.

Theorem 5 *Consider the pair $(\bar{A}, \bar{C})$ and its directed graph representation, $\mathscr{D}(\bar{A}, \bar{C})$. The pair $(\bar{A}, \bar{C})$ is structurally observable if and only if $\mathscr{D}(\bar{A}, \bar{C})$ is output connected and free of contractions.* ◇

Proof The proof follows immediately from Proposition 1, Proposition 2 and Theorem 3. □

Example 8 *Consider again the structural system whose directed graph representation is depicted in Figure 2.1. It can be seen that there is a path from every state vertex to some output vertex since $(x_1, y_2), (x_2, y_2)$ and (x_3, y_1) are edges of the digraph. To check whether the system is contraction-free, one may construct its bipartite graph depicted in Figure 2.4. Since there exist a maximal matching that covers $\mathscr{X}^-$, the system is free of contractions. In fact, only output y_2 is needed to guarantee structural observability. Notice that the edges that belong to the maximal matching do not include y_1 and there is a path from x_3 to y_2: $x_3 \to x_2 \to y_2$.*

Structural Observability Test

The direct correspondence between the necessary and sufficient conditions to guarantee structural observability and graph-theoretic conditions (see Theorem 5) makes suitable the design of an algorithm to check whether a given system is structurally observable or not. First, output-connectivity has to be verified. Second, one must conclude about the presence or absence of contractions.

Definition 21 (Out-Connected Sets) *Let $\mathscr{D}(\bar{A},\bar{C}) = (\mathscr{X} \cup \mathscr{Y}, \mathscr{E}_{\mathscr{X},\mathscr{X}} \cup \mathscr{E}_{\mathscr{X},\mathscr{Y}})$ be the directed graph associated with a structural system given by the pair $(\bar{A},\bar{C})$. Then, for each output component $y_j \in \mathscr{Y}$ and k, $1 \le k \le n$, the corresponding* ***out-connected*** *set $\mathscr{O}_j^k$ is defined as:*

$$\mathscr{O}_j^k = \{x_i \in \mathscr{X} \mid x_i \xrightarrow{p} y_j,\ p \le k\}. \tag{2.17}$$

◇

It is straightforward to see that a structural system is output-connected if and only if the following equality holds:

$$\bigcup_{y_j \in \mathscr{Y}} \mathscr{O}_j^n = \mathscr{X}. \tag{2.18}$$

Furthermore, the definition of out-connected sets is suitable for a recursive computation, since for every $y_j \in \mathscr{Y}$, we have

$$\mathscr{O}_j^1 = \{x_i \in \mathscr{X} \mid (x_i, y_j) \in \mathscr{E}_{\mathscr{X},\mathscr{Y}}\}, \tag{2.19}$$

$$\mathscr{O}_j^k = \{x_i \in \mathscr{X} \mid (x_i, x_l) \in \mathscr{E}_{\mathscr{X},\mathscr{X}} \text{ with } x_l \in \mathscr{O}_j^{k-1}\} \cup \mathscr{O}_j^{k-1},\ 1 < k \le n. \tag{2.20}$$

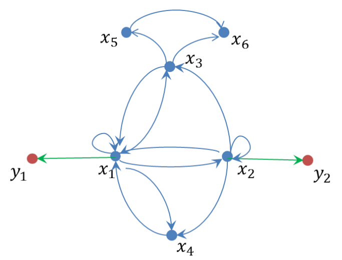
In order to test for the existence of contractions, one constructs the associated bipartite graph $\mathscr{B}(\bar{A},\bar{C}) = (\mathscr{X}^-, \mathscr{X}^+ \cup \mathscr{Y}, E_{\mathscr{X}^-,\mathscr{X}^+} \cup E_{\mathscr{X}^-,\mathscr{Y}})$ and seeks for a matching that covers $\mathscr{X}^-$. This can be done efficiently with the Hopcroft-Karp algorithm that takes as input a bipartite graph and produces as output a maximum matching [26]. If that maximum matching includes all the vertices in $\mathscr{X}^- = \{x_1^-, \dots, x_n^-\}$, then the system is contraction-free. Otherwise, a contraction must exist.

When the system is not structurally observable, the unobservable subspace and its dimension vary as function of the system parameters. However, as Hosoe demonstrated, the unobservable subspace dimension takes some constant value for all but an exceptional set of the free parameters that lie on a proper variety of zero Lebesgue measure [27].

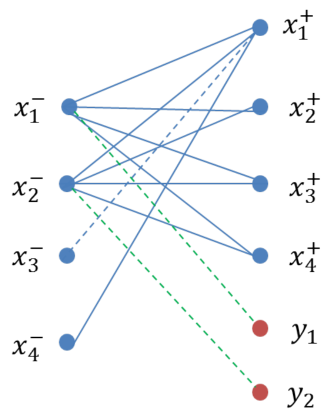
Definition 22 (Generic Dimension the Unobservable Subspace) *Given a structural pair $(\bar{A},\bar{C})$, the* ***generic dimension*** *of the corresponding unobservable subspace, denoted by $d_{\mathscr{U}}$, is defined as*

$$d_{\mathscr{U}} = n - grank\left(\begin{bmatrix} \bar{C} \\ \bar{C}\bar{A} \\ \vdots \\ \bar{C}\bar{A}^{n-1} \end{bmatrix}\right)^{3}, \tag{2.21}$$

[3] For two structural matrices $\bar{A} \in \{0,1\}^{n\times m}$ and $\bar{B} \in \{0,1\}^{m\times p}$, the product $\bar{C} = \bar{A}\bar{B}$ is a structural matrix $\bar{C} \in \{0,1\}^{n\times p}$ such that $\bar{C}_{ij} = \vee_{k=1}^{m}(\bar{A}_{ik} \wedge \bar{B}_{kj})$. The operations $\vee$ and $\wedge$ stand for the usual boolean operations *or* and *and*.



(a) The structural directed graph representation.



(b) The bipartite representation excluding vertices x_5 and x_6 since they do not belong to the out-reachable sets. The dashes lines represents edges that belong to a maximum matching.

Figure 2.5: Example 9: in the left, a structural directed graph representation and, in the right, the bipartite graph representation constructed after STEP 2 of the Algorithm 1.

where n is the size of the square matrix $\bar{A}$. ◇

Hosoe had proved the following theorem concerning the generic dimension of the unobservable subspace [27].

Theorem 6 *Let $d_{\mathscr{U}}$ be the generic dimension of the unobservable subspace associated with the pair $(\bar{A}, \bar{C})$, where $\bar{A}$ has size n. Assuming that $(\bar{A}, \bar{C})$ is not in form I:*

$$d_{\mathscr{U}} = n - grank\left(\begin{bmatrix}\bar{A}\\ \bar{C}\end{bmatrix}\right). \tag{2.22}$$

◇

If the system is in Form I (10), one has:

$$grank\left(\begin{bmatrix}\bar{C}\\ \bar{C}\bar{A}\\ \vdots\\ \bar{C}\bar{A}^{n-1}\end{bmatrix}\right) = grank\left(\begin{bmatrix}\bar{C}_2\\ \bar{C}_2\bar{A}_{22}\\ \vdots\\ \bar{C}\bar{A}_{22}^{n-1}\end{bmatrix}\right), \tag{2.23}$$

which means that the problem of determining $d_{\mathscr{U}}$ for $(\bar{A}, \bar{C})$ reduces to the same problem for the irreducible system $(\bar{A}_{22}, \bar{C}_2)$. With this in mind, it is possible to develop an algorithm that tests if a given system is structurally observable and, if not, returns the generic dimension of the unobservable subspace. It starts with the computation of the out-connected sets $\mathscr{O}_j^n$ for all $y_j \in \mathscr{Y}$. Then, it constructs the system $(\bar{A}_{22}, \bar{C}_2)$ by eliminating the state vertices for which there is not any path to some output variable, i.e, the ones that do not belong to $\bigcup_{y_j \in \mathscr{Y}} \mathscr{O}_j^n$. Finally, it proceeds to the construction of the bipartite graph of the pair $(\bar{A}_{22}, \bar{C}_2)$ and computes a maximum matching. If the size n' of that matching equals the size n of $\bar{A}$, the system is structurally observable. Otherwise, the generic dimension of the unobservable subspace is given by $n - n'$.

Example 9 *As an example to illustrate Algorithm 1, let us consider the structural system depicted in Figure 2.5. After the first step, the computation of the output-connected sets $\mathscr{O}_j^6$, for $j = 1, 2$, allow us to conclude that*

$$\bigcup_{y_j \in \mathscr{Y}} \mathscr{O}_j^n = \{x_1, x_2, x_3, x_4\} \neq \mathscr{X},$$

and, therefore, the system is not structurally observable because it can be reduced to form I. Thus, hereafter we work only with the irreducible system $(\bar{A}_{22}, \bar{C}_2)$ formed by the state vertices x_1, x_2, x_3 and x_4, in order to compute the generic dimension of the unobservable subspace.

The next step consists in verifying the size of the maximum matching in the bipartite graph constituted of the previously mentioned state variables. Figure 2.5 shows a maximum matching with three edges. Thus, the conclusion is that $d_{\mathscr{U}}$ is equal to $6 - 3 = 3$.

ALGORITHM 1: Algorithm for checking Structural Observability

Input: $\mathscr{D}(\bar{A},\bar{C}) = (\mathscr{X} \cup \mathscr{Y}, \mathscr{E}_{\mathscr{X},\mathscr{X}} \cup \mathscr{E}_{\mathscr{X},\mathscr{Y}})$
Output: TRUE if the system is structurally observable; $n - d_{\mathscr{Y}}$, otherwise

STEP 1

for $j = 1, \ldots, p$ **do**
  $\mathscr{O}_j^1 = \{x_i \in \mathscr{X} \mid (x_i, y_j) \in \mathscr{E}_{\mathscr{X},\mathscr{Y}}\}$
end for
for $j = 1, \ldots, p$ **do**
  for $k = 2, \ldots, n$ **do**
    $\mathscr{O}_j^k = \{x_i \in \mathscr{X} \mid (x_i, x_l) \in \mathscr{E}_{\mathscr{X},\mathscr{X}}$ with $x_l \in \mathscr{O}_j^{k-1}\} \cup \mathscr{O}_j^{k-1}$
  end for
end for
$\mathscr{X}^* \leftarrow \bigcup_{y_j \in \mathscr{Y}} \mathscr{O}_j^n$

STEP 2

Construct $\mathscr{B} = (\mathscr{V}_1, \mathscr{V}_2, E)$:

1. $\mathscr{V}_1 = \{x_1^-, \ldots, x_n^-\}$
2. $\mathscr{V}_2 = \{x_1^+, \ldots, x_n^+, y_1, \ldots, y_p\}$
3. $E = E_{\mathscr{X}^-,\mathscr{X}^+} \cup E_{\mathscr{X}^-,\mathscr{Y}}$
4. $\mathscr{E}_{\mathscr{X}^-,\mathscr{X}^+} = \{(x_j^-, x_i^+) \mid \bar{A}_{ij} = 1$ and $x_i, x_j \in \mathscr{X}^*\}$
5. $\mathscr{E}_{\mathscr{X}^-,\mathscr{Y}} = \{(x_j^-, y_i) \mid \bar{C}_{ij} = 1$ and $x_j \in \mathscr{X}^*\}$

Find a maximum matching in $\mathscr{B}$ with the Hopcroft-Karp algorithm. Set $\mathscr{V}_1^M \subseteq \mathscr{V}_1$ to the vertices that belong to the matching.
if $|\mathscr{V}_1^M| = |\mathscr{X}|$ **then**
  STOP and return TRUE
else
  STOP and return $n - |\mathscr{V}_1^M|$

Chapter 3

Output Selection for Structural Observability

In the previous chapter, some fundamental tools to verify structural observability were outlined. In this chapter, we are concerned with the design. More precisely, for a given structural pattern of the state matrix A, the goal is to construct the structural pattern of the output matrix C so that the resulting system is structurally observable.

3.1 Problem Statement

Consider a network of n entities. Each entity, denoted by x_i with $i \in N = \{1,\ldots,n\}$, updates its dynamics at each iteration instant $t \in \mathbb{N}$ according

$$x(t+1) = Ax(t), \tag{3.1}$$

where $x(t) = [x_1(t)\ldots x_n(t)]^T$ is the state vector and $A \in \mathbb{R}^{n\times n}$ is a weighting matrix whose entries are not precisely known. In fact, we just consider that the structural pattern $\bar{A} \in \{0,1\}^{n\times n}$ is known, which, for a network system, is the same as saying that the communication topology is known. This means that if x_j is not allowed to communicate with x_i, then the entry A_{ij} must be zero, otherwise can be any real value.

Since the goal is to obtain structural observability, one has to start placing output variables according to that constraint. Hereafter, we will consider the case where each output variable y_j measures one and only one state variable x_i, i.e, is a *dedicated output variable*. The goal is then to choose $S_{\mathscr{O}} \subset \mathscr{X} = \{x_1,\ldots,x_n\}$ in order to obtain an output equation

$$y(t) = Cx(t), \tag{3.2}$$

where $y \in \mathbb{R}^p$ is the output vector with $p = |S_{\mathscr{O}}|$ and $C \in \mathbb{R}^{p\times n}$, so that the pair (A,C) is observable. Such a dedicated assignment $S_{\mathscr{O}}$ of outputs to state variables will be called a *feasible dedicate output configuration (FDOC)*. As in above, we will concentrate first on the design of

$\bar{C}$, the zero/non-zero pattern of C, whereas the design of C will be performed after. The relation between $\bar{C}$ and $S_{\mathscr{O}}$ is expressed by

$$\bar{C} = I_n^{S_{\mathscr{O}}}, \tag{3.3}$$

where I_n is the identity matrix of size $n \times n$ and $I_n^{S_{\mathscr{O}}}$ consists of the rows $j_1, j_2, \ldots, j_p$ of the identity matrix corresponding to the state components in $S_{\mathscr{O}}$, i.e, such that $S_{\mathscr{O}} = \{x_{j_1}, x_{j_2}, \ldots, x_{j_p}\}$.

Obviously, $S_{\mathscr{O}} = \mathscr{X}$ is always a solution to the problem described. However, in practical applications, there exists sometimes a cost associated with measuring a state component. With that in mind, we would like to obtain structural observability but with the cheapest dedicated output configuration. Having said that, our main problem can be formulated precisely as follows.

Output Selection Problem: Consider the system (3.1) and let $\mathscr{X} = \{x_1, \ldots, x_n\}$ be the set of state variables, $\bar{A} \in \{0,1\}^{n \times n}$ the structural pattern of A and $\mathscr{C} : \mathscr{X} \to \mathbb{R}^+$ a given cost function. The problem is to find efficiently, i.e., with polynomial complexity in n, a set $S_{\mathscr{O}}$, with $S_{\mathscr{O}} \subseteq \mathscr{X}$, that solves the following optimization problem:

$$\begin{aligned} \min \quad & \sum_{x_i \in S_{\mathscr{O}}} \mathscr{C}(x_i) \\ \text{s.t.} \quad & (\bar{A}, I_n^{S_{\mathscr{O}}}) \text{ is structurally observable} \end{aligned} \tag{3.4}$$

where $I_n^{S_{\mathscr{O}}}$ is defined as in (3.3).

Particularly, if $\mathscr{C}(x_i) = 1$ for all $x_i \in \mathscr{X}$, the problem reduces to selecting a minimum number of state variables to be measured in order to guarantee structural observability.

3.2 A New Approach via Matroid Theory

Matroids were introduced by Whitney in 1935 as a common generalization of graphs and matrices [15]. Nowadays, matroids play an important role in combinatorial optimization. With the advent of modern computation, the algorithmic complexity is a key factor. It is not enough to propose a solution to solve a determined problem, we also have to guarantee computational efficiency. It happens that matroids are exactly those structures where the greedy algorithm [16] provides an optimal solution. Furthermore, as Edmonds [28] discovered, a set of maximum cardinality that belongs to the intersection of two matroids can also be computed efficiently.

Structural systems analysis lies on both graph and matrices concepts. Due to its combinatorial nature, it seems plausible to try to apply matroid theory to solve some problems in an efficiently manner. In what follows, we provide a brief review of the fundamental concepts and results of matroid theory based on [29], [30], [31] and [32]. After that, our main problem (3.4) is addressed within the matroid framework.

Matroids Background

Matroid theory is an abstraction that is suitable to deal with the concept of *independence* both in graphs and in matrices. As an abstraction, it is built on some properties; not too many but enough to prove nontrivial statements.

Definition 23 (Matroid) *A pair $(S, \mathscr{I})$ is called a matroid if S is a finite set and $\mathscr{I}$ is a nonempty collection of subsets of S satisfying the following axioms:*

(i) if $I \in \mathscr{I}$ and $J \subseteq I$, then $J \in \mathscr{I}$,

(ii) if $I, J \in \mathscr{I}$ and $|J| > |I|$, then there exists $z \in J \setminus I$ such that $I \cup \{z\} \in \mathscr{I}$.

$\diamond$

The set S is called the *ground set*. For a given matroid $M = (S, \mathscr{I})$ and a set $I \subseteq S$, I is called *independent* if it belongs to $\mathscr{I}$, and *dependent* otherwise. The first axiom from Definition 23 is usually called the *hereditary property* and the second axiom the *augmentation property*. It is possible to notice the equivalence of that axiom system with another one, that will reveal useful in proofs.

Theorem 7 ([30]) *Let S be a finite set and $\mathscr{I}$ a nonempty collection of subsets of S satisfying the hereditary property of Definition 23. Then, the augmentation property is equivalent to the following:*

(ii') if $I, J \in \mathscr{I}$ and $|I \setminus J| = 1$, $|J \setminus I| = 2$, then there exists $z \in J \setminus I$ such that $I \cup \{z\} \in \mathscr{I}$.

$\diamond$

Proof It is straightforward to see that the augmentation property from Definition 23 implies condition (ii'). To prove the other implication, induction will be used on $|I \setminus J|$. Let $I, J \in \mathscr{I}$ be such that $|J| > |I|$. Assume that $|I \setminus J| = 0$. Then, I is a proper subset of J and there exists an element $z \in J \setminus I$ such that $I \cup \{z\} \subseteq J$. Due to the hereditary property, this implies that $I \cup \{z\} \in \mathscr{I}$, and hence (ii) holds in this case.

Assume, now, that (ii) holds for $|I \setminus J| \leq k$ and let I and J be such that $|J| > |I|$ and $|I \setminus J| = k+1$. Then, there exists $x \in I \setminus J$ and $|(I \setminus \{x\}) \setminus J| = k$. By applying the induction hypothesis to $I \setminus \{x\}$ and J, we find that there exists $y \in J \setminus (I \setminus \{x\}) = J \setminus I$ such that $(I \setminus \{x\}) \cup \{y\} \in \mathscr{I}$. Now, $|J \setminus ((I \setminus \{x\}) \cup \{y\})| = |J \setminus (\{y\} \cup I)|$ and the induction hypothesis can be applied to $I \setminus \{x\} \cup \{y\}$ and J to conclude that there exists $y' \in J \setminus ((I \setminus \{x\}) \cup \{y\})$ such that $I \setminus \{x\} \cup \{y\} \cup \{y'\} \in \mathscr{I}$. Finally, using condition (ii') on I and $I \setminus \{x\} \cup \{y\} \cup \{y'\}$, we conclude that $I \cup \{y\} \in \mathscr{I}$ or $I \cup \{y'\} \in \mathscr{I}$, i.e, that axiom (ii) of Definition 23 holds for $|I \setminus J| = k+1$. $\square$

We now provide some examples of matroids.

Proposition 3 *Let S be the set of column labels of a $m \times n$ matrix A over $\mathbb{R}$. For a subset X of S, let A_X denote the submatrix of A consisting of those columns indexed by X. Now, let*

$$\mathscr{I} = \{I \subseteq S \mid rank(A_I) = |I|\}. \tag{3.5}$$

Then, $M = (S, \mathscr{I})$ is a matroid. $\diamond$

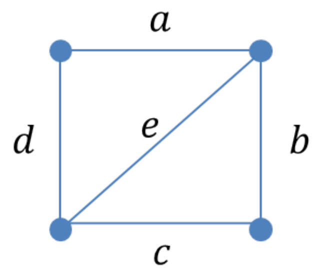


Figure 3.1: An undirected graph where a graphical matroid is defined. The independent sets correspond to the subsets of the edge set that do not contain any cycle.

Proof The hereditary property is trivially verified. In order to prove the augmentation property, a fundamental result of linear algebra will be used. Consider $I,J \in \mathscr{I}$ with $|J| > |I|$. If A_I and A_J have full column rank, their columns span a space of dimension $|I|$ and $|J|$, respectively. Therefore, since $|J| > |I|$, there must exist a column of A_J, say column $j*$, that is not in the span of the columns of A_I. Thus, $rank(A_{I\cup\{j^*\}}) = |I \cup \{j^*\}|$, i.e, there exists $j^* \in J \setminus I$ such that $I \cup \{j^*\} \in \mathscr{I}$. □

Definition 24 (Linear Matroid) *The matroid of Proposition 3 will be called a **linear** matroid.* ◇

Example 10 *Consider the following matrix*

$$M = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix},$$

and let the columns of M be labelled from left to right with $1,2,\ldots,5$*. Define* $S = \{1,2,\ldots,5\}$ *and let* $\mathscr{I}$ *be the collection of subsets of S for which the multiset of columns labelled by I is linearly independent. Then,* $\mathscr{I} = \{\emptyset,\{1\},\{2\},\{4\},\{5\},\{1,2\},\{1,5\},\{2,4\},\{2,5\},\{4,5\}\}$ *and it can be seen that the pair* $(E,\mathscr{I})$ *is a matroid.*

Another important class of matroids that usually arise in combinatorial optimization is the class of graphic matroids, whose independent sets are those subsets of edges which are *forests*, i.e, that do not contain any cycles.

Proposition 4 *Consider an undirected graph* $G = (V,E)$*. Let*

$$\mathscr{I} = \{I \subseteq E \mid (V,I) \text{ does not contain any cycles}\}.$$

Then, $M = (E,\mathscr{I})$ *is a matroid.* ◇

Proof The hereditary property is easily verified since if the edges in I define a forest, surely a subset of these edges will also define a forest. To prove the augmentation property, consider $I,J \in \mathscr{I}$ such that $|J\setminus I| = |\{j_1, j_2\}| = 2$ and $|I\setminus J| = |\{i_1\}| = 1$. Let us assume that $I\cup\{j_1\}$ and $I\cup\{j_2\}$ do not belong to $\mathscr{I}$. If $I\cup\{j_1\}$ contains a cycle C_1, surely that cycle must comprise both i_1 and j_1. If it was not the case, then I or J would contain a cycle contradicting our assumption that

$I,J \in \mathscr{I}$. The same argument applies to $I \cup \{j_2\}$, that must contain a cycle C_2. From C_1 and C_2 and considering that $i_1 = (a,b)$ one can easily construct two distinct paths in $G_s = (V,I)$ between a and b. But, if there are two distinct paths in G_s, then G_s should contain a cycle, contradicting the initial assumption that $I \in \mathscr{I}$. Thus, $I \cup \{j_1\}$ or $I \cup \{j_2\}$ have to belong to $\mathscr{I}$ and the second property follows from Theorem 7. □

Definition 25 (Graphic Matroid) *The matroid of Proposition 4 will be called a* ***graphic*** *matroid.* ◇

Example 11 *Consider the graph $G = (V,E)$ depicted in Figure 3.1. Let S be the edge set associated with G and $\mathscr{I}$ the collection of edge sets for which the corresponding subgraph contain no cycles. Then, $M = (S,\mathscr{I})$ is a matroid. Notice that $\{b,c,d\}$ and $\{a,e\}$ belong to $\mathscr{I}$. One easily concludes that $\{a,e,b\}$ is also independent. This illustrates the situation of the second matroid axiom.*

Finally, there is a class of matroids suitable to deal with bipartite matchings: the *partition* matroid.

Proposition 5 *Consider a set S that is partitioned into l disjoint sets: $S_1,\ldots,S_l$. Let*

$$\mathscr{I} = \{I \subseteq S \mid |I \cap S_i| \leq k_i,\ \forall i = 1,\ldots,l\},$$

for some given parameters k_i, with $k_i \in \mathbb{N}, \forall i = 1,\ldots,n$. Then, $M = (S,\mathscr{I})$ is a matroid. ◇

Proof The hereditary property is trivially verified. Consider $I,J \in \mathscr{I}$. If $|J| > |I|$, then there must exist i such that $|J \cap S_i| > |I \cap S_i|$ which implies that we can add any element x in $(J \setminus I) \cap S_i$ to I such that $I \cup \{x\} \in \mathscr{I}$. Therefore, the augmentation property is also verified. □

Definition 26 (Partition Matroid) *The matroid of Propostion 5 will be called a* ***partition*** *matroid.* ◇

As matroids abstract the concept of independence, it is not surprising that the following results arise naturally.

Definition 27 (Base) *Let $M = (S,\mathscr{I})$ be a matroid. For any $U \subseteq S$, a subset B of U is called a* ***base*** *of U if B is an inclusionwise maximally independent subset of U, i.e., $B \in \mathscr{I}$ and there is no $Z \in \mathscr{I}$ with $B \subset Z \subseteq U$.* ◇

It is not difficult to see that in general a base of a given set U, $U \subseteq S$, is not unique. However, distinct bases have the same size. In fact, let us suppose that B_1 and B_2 are bases of U but with different sizes. Without loss of generality, it may be assumed that $|B_2| > |B_1|$. But, according to axiom (ii) of Definition 23, there exists $z \in B_2 \setminus B_1$ such that $B_1 \cup \{z\} \in \mathscr{I}$. Moreover, since $B_1 \subseteq U$ and $B_2 \subseteq U$, surely it will happen that $B_1 \cup \{z\} \subseteq U$. That contradicts the assumption that B_1 is a base. The common size of the bases of U is called *rank* of U and denoted by $r(U)$. If $U = S$, then a base of S is simply called a base and the common size of all bases is called the *rank* of the matroid.

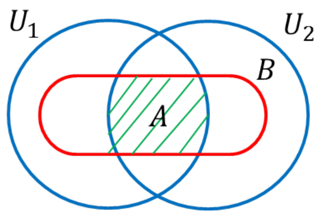


Figure 3.2: The construction given in proof of Theorem 8. The set A is a maximal independent set of $U_1 \cap U_2$ and it was augmented to create the set B, a maximal independent set of $U_1 \cup U_2$.

Theorem 8 ([32]) *Let $M = (S, \mathscr{I})$ be a matroid. Then, for every U_1, $U_2 \subseteq S$, the following holds:*

(i) $0 \leq r(U_1) \leq |U_1|$,

(ii) if $U_1 \subseteq U_2$, *then* $r(U_1) \leq r(U_2)$,

(iii) $r(U_1) + r(U_2) \geq r(U_1 \cup U_2) + r(U_1 \cap U_2)$.

◇

Proof The proof of (i) and (ii) follows from the definition of rank. To prove (iii), consider the construction depicted in Figure 3.2. Let $A \subseteq U_1 \cap U_2$ be a maximally independent subset of $U_1 \cap U_2$. Applying successively the augmentation property to A, one can obtain the subset $B \subseteq U_1 \cup U_2$ such that $A \subseteq B$ and B is a maximal independent subset of $U_1 \cup U_2$. From the figure, it is easy to see that $|B \cap U_1| + |B \cap U_2| = |B| + |A|$ since the elements of A are counted twice. Since B is independent, so it is $B \cap U_1$ and $B \cap U_2$. Further, according to (ii), $r(U_1) \geq |B \cap U_1|$ and $r(U_2) \geq |B \cap U_2|$. Thus, $r(U_1) + r(U_2) \geq |B \cap U_1| + |B \cap U_2| = |B| + |A| = r(U_1 \cup U_2) + r(U_1 \cap U_2)$. □

The third property of Theorem 8 is called *submodularity* and plays an important role in various applications of matroids. Another important property is the association of matroids with the greedy algorithm. As the name indicates, the greedy algorithm is an algorithm that follows the heuristic of making the locally optimal choice at each step. Although this does not work to solve every optimization problem, it will be seen that matroids are exactly the structures for which the algorithm provides an optimal solution.

ALGORITHM 2: Greedy algorithm for selecting the max-weight independent set of a matroid

Input: A matroid $M = (S, \mathscr{I})$, with $S = \{1, 2, \ldots, n\}$, and a weight function $w : S \to \mathbb{R}$
Output: An independent set $I^* \in \mathscr{I}$ such that $w(I^*) = max_{I \in \mathscr{I}} w(I)$, where $w(I) = \sum_{i \in I} w(i)$
Relabel the elements of the matroid so that $w(1) \geq w(2) \geq \ldots \geq w(n)$.
$I^* \leftarrow \emptyset$.
for $i = 1, \ldots, n$ **do**
    if $I^* \cup \{i\} \in \mathscr{I}$ **then**
        $I^* \leftarrow I^* \cup \{i\}$.
end for Return I^*.

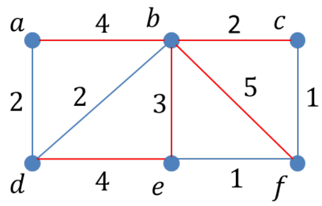


Figure 3.3: An undirected graph whose edges were attributed a cost. The edges in red represent a maximum spanning tree.

Theorem 9 ([30]) *Let $\mathscr{I}$ be a nonempty collection of subsets of S closed under taking subsets. Then, the pair $(S, \mathscr{I})$ is a matroid if and only if for any weight function $w : S \to \mathbb{R}^+$, the greedy algorithm returns a set $I \in \mathscr{I}$ of maximum weight $w(I)$.* ◇

Proof To show sufficiency, let $M = (S, \mathscr{I})$ be a matroid and $w : S \to \mathbb{R}^+$ a weight function. The solution of Algorithm 2 will consist of a maximum weight base. Then, it suffices to show that I is always contained in a maximum weight base. Since I^* starts with the empty set, the desired property holds initially. Now, to prove the induction step, assume that I^* is contained in some maximum weight base, B, and let y be an element in $S \setminus I^*$ with $w(y)$ as large as possible. If $y \in B$, then $I^* \cup \{y\} \subseteq B$. If $y \notin B$, then there exists a base B' such that $I^* \cup \{y\} \subseteq B'$ and $B' \subseteq B$. Thus, there exists $z \in B \setminus I^*$ such that $B' = B \setminus \{z\} \cup \{y\}$. Since $I^* \cup \{z\} \subseteq B$, $I^* \cup \{z\} \in \mathscr{I}$ and since $w(y)$ was chosen maximum, $w(y) \geq w(z)$. Thus, $w(B') \geq w(B)$ and therefore B' is a maximum weight base that contains I^*.

To prove necessity, let us consider that the greedy algorithm leads to a maximum weigh independent set for any weight function $w : S \to \mathbb{R}^+$. Condition (i) of Definition 23 is satisfied by assumption. Let us assume that condition (ii) of Definition 23 does not hold. Thus, there are $I, J \in \mathscr{I}$ with $|J| > |I|$ such that $I \cup \{z\} \notin \mathscr{I}$ for all $z \in J \setminus I$. Let $k = |I|$ and consider the following weight function:

$$w(x) = \begin{cases} k+2 \text{ if } x \in I, \\ k+1 \text{ if } x \in J \setminus I, \\ 0, \text{otherwise.} \end{cases} \tag{3.6}$$

Then, in the first k iterations, it will be $I^* = I$. Since, by assumption, $I \cup \{z\} \notin \mathscr{I}$ for all $z \in J \setminus I$, the weight of the output $w(I^*)$ will be $k(k+2)$. Notice, however, that $w(J) = |J|(k+1) \geq (k+1)(k+1) > k(k+2) = w(I)$, contradicting our initial assumption. □

Example 12 *Consider the undirected graph $G = (V, E)$ depicted in Figure 3.3 and a weighting function $w : E \to \mathbb{R}^+$. A problem that usually arises in computer science is that of finding a maximum weighted tree, where a tree is a connected graph without cycles. Since the original graph G is connected, the problem is equivalent to that of computing the maximum weight forest. If we define $S = E$ and $\mathscr{I} = \{I \subseteq E \mid$ such that (V, I) do not contain any cycles$\}$, it follows from Proposition 4 that $M = (S, \mathscr{I})$ is a matroid. Thus, we can apply the greedy algorithm. It turns out* *that this algorithm is equivalent to Kruskal's algorithm to compute a maximum spanning tree for a connected weighted graph. It starts by sorting the edges by decreasing weight. Then, a edge is added if that addition do not create a cycle.*

It happens that some relevant problems within the scope of combinatorics can be expressed as the intersection of two matroids. Let $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$ be two matroids. The intersection is the collection of sets $I \in \mathscr{I}_1 \cap \mathscr{I}_2$. However it is not generally the case that $(S, \mathscr{I}_1 \cap \mathscr{I}_2)$ is a collection of independent sets of a matroid on E and therefore the greedy algorithm cannot be applied. In spite of this, Edmonds [28] showed that there exist efficient algorithms for the intersection of two matroids. More precisely, he showed that it is possible to find a maximum-weight common independent set in two matroids in strongly polynomial time.

Before describing the algorithm to compute a maximum-size common independent set in two matroids, we first describe a min-max relation that will be used to prove the optimallity of the algorithm. First, notice that given two matroids on the same ground set, $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$, with rank functions r_1 and r_2, respectively, one may write for any $I \in \mathscr{I}_1 \cap \mathscr{I}_2$ and any $U \subseteq S$

$$|I| = |I \cap U| + |I \cap (S \setminus U)| \leq r_1(U) + r_2(S \setminus U),$$

since, by the first axiom, $I \cap U$ and $I \cap S \setminus U$ are both independent in $\mathscr{I}_1$ and $\mathscr{I}_2$. If the maximum over I and the minimum over U is taken, it is possible to write the inequality:

$$\max_{I \in \mathscr{I}_1 \cap \mathscr{I}_2} |I| \leq \min_{U \subseteq S} [r_1(U) + r_2(S \setminus U)] \; .$$

This inequality was proved first by Edmonds [28] and is known as the Matroid Intersection Theorem, stated next.

Theorem 10 (Matroid Intersection) *For any two matroids $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$ with rank functions r_1 and r_2 respectively, the following equality holds:*

$$\max_{I \in \mathscr{I}_1 \cap \mathscr{I}_2} |I| = \min_{U \subseteq S} [r_1(U) + r_2(E \setminus U)]. \tag{3.7}$$

◇

The previous theorem provides the upper limit in which we may stop the algorithm to compute a maximum-size common independent set in two matroids. However, one has to know what are the steps to follow in order to obtain the desired result. Before we present the algorithm, some technical results concerning bipartite graphs are presented.

Lemma 2 ([31]) *Let $G = (V, E)$ be a bipartite graph with bipartition $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$, and suppose that G has a unique matching M that includes all the vertices of V_1. Then there exists an edge $e = (v_1, v_2) \in M$, $v_1 \in V_1$, $v_2 \in V_2$, such that:*

$$(v_1, v_2') \notin E, \forall v_2' \in V_2 \setminus \{v_2\}.$$

◇

For any matroid and an independent set I, one can define a bipartite graph that expresses the changes that can occur in I so that it remains independent as follows.

Definition 28 (Bipartite Exchange Graph) *Let $M = (S, \mathscr{I})$ be a matroid and $I \in \mathscr{I}$ an independent set. Then, the **bipartite exchange graph** $G_M(I)$ is the bipartite graph with partition I and $S \setminus I$ such that for any $y \in I$, $x \in S \setminus I$,*

$$(y,x) \text{ is an edge of } G_M(I) \text{ if and only if } I \setminus \{y\} \cup \{x\} \in \mathscr{I}.$$

◇

The following lemma states that an exchange between two independent sets implies a perfect matching in the bipartite exchange graph.

Lemma 3 ([31]) *Let $M = (S, \mathscr{I})$ be a matroid with $I, J \in \mathscr{I}$ and $|I| = |J|$. Let $G_M(I)$ be the bipartite exchange graph associated with I. Then, $G_M(I)$ contains a perfect matching between $I \setminus J$ and $J \setminus I$.* ◇

A very useful partial converse also holds.

Lemma 4 ([31]) *Let $M = (S, \mathscr{I})$ be a matroid with $I \in \mathscr{I}$ and $G_M(I)$ the bipartite exchange graph. Let $J \subseteq S$ with $|J| = |I|$. If $G_M(I)$ contains a unique perfect matching between $I \setminus J$ and $J \setminus I$, then $J \in \mathscr{I}$.* ◇

Returning to the subject of matroid intersection, a structure similar to the bipartite exchange graph can be constructed for a pair of matroids.

Definition 29 (Bipartite Exchange Digraph) *Given two matroids defined over the same ground set, $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$, and any $I \in \mathscr{I}_1 \cap \mathscr{I}_2$, the exchange graph $\mathscr{D}_{M_1,M_2}(I)$ is the directed graph with bipartition I and $S \setminus I$ such that for any $y \in I$, $x \in S \setminus I$,*

$$(y,x) \text{ is an arc of } \mathscr{D}_{M_1,M_2}(I) \text{ if and only if } I \setminus \{y\} \cup \{x\} \in \mathscr{I}_1,$$
$$(x,y) \text{ is an arc of } \mathscr{D}_{M_1,M_2}(I) \text{ if and only if } I \setminus \{y\} \cup \{x\} \in \mathscr{I}_2.$$

◇

Relying on the previous definition, it would be of interest to construct directed paths along the bipartite exchange digraph that would result in augmenting the number of elements of I. To that purpose, certain vertices in $S \setminus I$ are termed *sources* while others are the *sinks*. A source (respectively, sink) of $\mathscr{D}_{M_1,M_2}(I)$ is a vertex $x \in S \setminus I$ such that $I \cup \{x\} \in \mathscr{I}_1$ (respectively, $I \cup \{x\} \in \mathscr{I}_2$). A *source-sink dipath* is a directed path that begins in a source and ends in a sink and the

(a) Graphical representation of matroid M_1. (b) Graphical representation of matroid M_2.

Figure 3.4: This figure represents two undirected graphs in correspondence with two graphical matroids. The edges in red belong to an independent set that belongs to both matroids.

definition includes the degenerate scenario where the path contains no edges, for which the source and the sink are the same vertex. Then, the path P will have the vertex sequence $x_0, y_1, x_1, \ldots, y_n, x_n$ where for all i, $x_i \in S \setminus I$, $y_i \in I$, and x_0 is a source whereas x_n is a sink. The source-sink dipath is said to be *augmenting* if $I' = I \setminus \{y_1, \ldots, y_n\} \cup \{x_0, x_1, \ldots, x_n\}$ is in $\mathscr{I}_1 \cap \mathscr{I}_2$. Notice that if $V(P)$ is the set that contain the vertices that comprise the path P, then I' can be written as $I' = I \triangle V(P)$. Finally, the following lemma makes possible the task of finding an augmenting dipath.

Lemma 5 ([31]) *Let $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$ be matroids and let $I \in \mathscr{I}_1 \cap \mathscr{I}_2$. If P is a shortest source-sink dipath in $\mathscr{D}_{M_1,M_2}(I)$, then it is augmenting.* $\diamond$

Example 13 *In order illustrate Lemma 5, consider the two graphical matroids depicted in Figure 3.4. The ground set is $S = \{a, b, c, d, e\}$ and let $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$ be the graphical matroids associated with Figure 3.4a and Figure 3.4b, respectively. Since $I = \{b, d\}$ do not contain any cycle both in Figure 3.4a and Figure 3.4b, one must have that $I \in \mathscr{I}_1 \cap \mathscr{I}_2$. With I, M_1 and M_2, one can construct the bipartite exchange digraph $\mathscr{D}_{M_1,M_2}(I)$, which is represented in Figure 3.5. The set of sources is $X_1 = \{x \in S \setminus I \mid I \cup \{x\} \in \mathscr{I}_1\} = \{a\}$ while the the of sinks is $X_2 = \{x \in S \setminus I \mid I \cup \{x\} \in \mathscr{I}_2\} = \{e\}$. If any path between a and e is considered, there is no guarantee that this path will correspond to an augmenting one. Indeed, take the directed path P_1, whose edges are painted red and orange in Figure 3.5. Then $V(P_1) = \{a, b, c, d, e\}$ and $I_1 = I \triangle V(P_1) = \{a, c, e\}$ that although independent in $\mathscr{I}_1$ it is not in $\mathscr{I}_2$. On the other hand, if we take the directed path P_2 whose edges are painted green and orange and that corresponds to a shortest path between a and e, then, by Lemma 5, we know that P_2 is augmenting. In fact, $V(P_2) = \{a, d, e\}$ and $I_2 = I \triangle V(P_1) = \{a, b, e\}$, with $I_2 \in \mathscr{I}_1 \cap \mathscr{I}_2$ and $|I_2| > |I|$.*

With the previous lemma in mind, the following algorithm to compute a maximum-cardinality set that is independent in any given two matroids comes out.

Theorem 11 (Correctness of the Cardinality Matroid-Intersection Algorithm) *Let $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$ be matroids. Then the Cardinality Matroid Intersection Algorithm (3) returns an independent set $I^* \in \mathscr{I}_1 \cap \mathscr{I}_2$ such that $|I^*| \geq |I|$, $\forall I \in \mathscr{I}_1 \cap \mathscr{I}_2$.* $\diamond$

ALGORITHM 3: Cardinality Matroid-Intersection Algorithm

Input: Matroids $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$
Output: An independent set $I^* \in \mathscr{I}_1 \cap \mathscr{I}_2$ such that $|I^*| \geq |I|, \forall I \in \mathscr{I}_1 \cap \mathscr{I}_2$
1:$I^* \leftarrow \emptyset$.
2:Construct the set $X_1 = \{x \in S \setminus I^* \mid I^* \cup \{x\} \in \mathscr{I}_1\}$.
3:Construct the set $X_2 = \{x \in S \setminus I^* \mid I^* \cup \{x\} \in \mathscr{I}_2\}$.
if *$\mathscr{D}_{M_1,M_2}(I^*)$ has an $X_1 - X_2$ path* **then**
| Take the shortest such path P. $I^* \leftarrow I^* \triangle V(P)$. Go to 2.
else
| Return I^*.

Proof Consider that when the algorithm stops it returns the set I. Since I starts to be the empty set, attending to Lemma 5, it is easy to see that in the end I is in $\mathscr{I}_1 \cap \mathscr{I}_2$. Thus, it remains to prove that I has indeed maximum cardinality. To that purpose, we will rely on the following construction. Let $U = \{x \in S \mid$ there is $x - y$ path in $\mathscr{D}_{M_1,M_2}(I)$, where y $\in X_2\}$. Taking into account the definition of U and considering that there is no directed path between X_1 and X_2, one easily verifies that $U \cap X_1 = \emptyset$, $X_2 \subseteq U$ and there is no arc entering U (see Figure 3.6). First, we will show that $r_1(U) \leq |I \cap U|$, where r_1 is the rank function associated with M_1. If $r_1(U) > |I \cap U|$, then there exists $x \in U \setminus I$ such that $(I \cap U) \cup \{x\} \in \mathscr{I}_1$. It must be that $I \cap \{x\} \notin \mathscr{I}_1$, since x is not in X_1. Therefore, there must exist $y \in I \setminus U$ with $I \setminus \{y\} \cap \{x\} \in \mathscr{I}_1$. Nevertheless, by definition of $\mathscr{D}_{M_1,M_2}(I)$, there is an arc from y to x, contradicting the fact that no arc enters U. In a similar manner, it is possible to prove that $r_2(S \setminus U) \leq |S \setminus U|$. Thus

$$|I| = |I \cap U| + |I \cap (S \setminus U)| \geq r_1(U) + r_2(S \setminus U),$$

and by Theorem 10 it is easy to see that I is a maximum-cardinality independent set resulting from the intersection of the two matroids. □

Notice that the algorithm has polynomial complexity provided that the time required to test if a given set I belongs (or not) to M_1 or M_2 is a polynomial function in the size of I. That is, for matroids M_1 and M_2 there must exist a subroutine to test whether or not a set of elements is independent. That subroutine is commonly referred as an *independence oracle*.

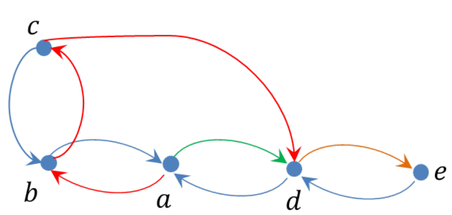


Figure 3.5: A bipartite exchange digraph associated with M_1, M_2 and I as defined in Example 13. The vertex a is the only source whereas vertex e is the only sink. The directed path whose arcs are in green and orange represent a shortest source-sink path.

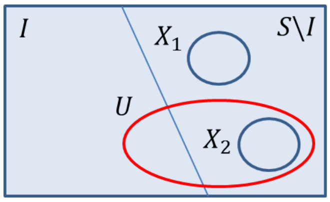


Figure 3.6: This figure depicts the construction of a set U with the conditions described in the proof of Theorem 11.

Example 14 (Bipartite Matching) *Let $\mathscr{B} = (V_1, V_2, E)$ be a bipartite graph. First, we note that matchings in a bipartite graph do not form a matroid because in general the second axiom is not verified. However, matchings can be regarded as an intersection of two matroids. For a given vertex $x \in V_1$, define $\sigma(x)$ as the set of edges incident to x, i.e, $\sigma(x) = \{e \in E \mid$ there exists $y \in V_2$ with $e = (x,y)\}$. Then, E can be partitioned as $E = \bigcup_{x \in V_1} \sigma(x)$. This forms a partition since all edges have precisely one endpoint in V_1. Thus, we can define*

$$\mathscr{I}_1 = \{I \subseteq E \mid |I \cap \sigma(x)| \leq 1\,,\, \forall x \in V_1\}.$$

By Proposition 5, we know that $M_1 = (E, \mathscr{I}_1)$ is a partition matroid. Notice that a set of edges is independent in M_1 if it has at most one edge incident to every vertex in V_1. In a similar fashion, we can define

$$\mathscr{I}_2 = \{I \subseteq E \mid |I \cap \sigma(x)| \leq 1\,,\, \forall x \in V_2\},$$

and construct another matroid $M_2 = (E, \mathscr{I}_2)$. Then, it can be concluded that $I \in \mathscr{I}_1 \cap \mathscr{I}_2$ if and only if I corresponds to a matching in $\mathscr{B}$. Therefore, in order to compute a maximum matching, one can resort to the matroid intersection algorithm.

We will consider now the *weighted* matroid intersection problem. Let $M_1 = (S, \mathscr{I}_1)$ and $M_2 = (S, \mathscr{I}_2)$ be two matroids and let $w : S \to \mathbb{R}^+$ be a weight function. Suppose that we would like to find an independent set $I \in \mathscr{I}_1 \cap \mathscr{I}_2$ with maximum weight. It turns out that Algorithm 3 can be generalized in a straightforward fashion to the weighted case, where the only difference is in finding the shortest source-sink directed path P. In computing P, one assigns weights to each vertex $x \in \mathscr{D}_{M_1,M_2}(I)$ as $w(x)$ if $x \in \mathscr{I}$ and $-w(x)$ when $x \notin \mathscr{I}$. Then, P should have the minimum number of arcs among all the minimum length $X_1 - X_2$ directed paths. The correctness of the algorithm will not be demonstrated here. It can be found in [30].

3.3 The Solution of the Output Selection Problem

In this section, the solution to the Output Selection Problem (3.4) is presented. The key idea is to formulate the problem as the intersection of two matroids, such that the necessary and sufficient conditions of Theorem 5 to guarantee structural observability hold. That is, two matroids will be constructed over the set of the state variables: one to guarantee output-connectivity (cf. Definition 15); and another one to guarantee that the resulting structural pair is free of contractions (cf. Definition 17). Then, we will demonstrate that if dedicated outputs are assigned to the state variables in $S_{\mathscr{O}}$, with $S_{\mathscr{O}} \subseteq \mathscr{X}$, then that assignment leads to structural observability if and only if the set $\mathscr{X} \setminus S_{\mathscr{O}}$ belongs to the intersection of that two matroids. Thus, the matroid intersection algorithm can be applied to obtain a maximum-cardinality independent set I^* that lies in the intersection of the two matroids. Finally, $\mathscr{X} \setminus I^*$ will correspond to a minimum feasible dedicated output configuration.

Initially, attention will be devoted to the particular instance of the problem, when the cost function associated with the output placement is uniform. That is, the problem of finding the minimum number and the corresponding location of outputs to attain structural observability. The generalization to the scenario where the cost function takes different values according to the specific output follows from the generalization of Algorithm 3 to the problem of finding a maximum-weight common independent set in two matroids.

In the previous chapter, the necessary and sufficient conditions to guarantee structural observability were outlined in correspondence to graph criteria (see Theorem 5). Briefly, the pair $(\bar{A},\bar{C})$ is structurally observable if and only if the directed graph representation $\mathscr{D}(\bar{A},\bar{C})$ is *output-connected* (cf. Definition 15) and *free of contractions* (cf. Definition 17). With the structural matrix $\bar{A} \in \{0,1\}^{n\times n}$, one can associate the directed graph $\mathscr{D}(\bar{A}) = (\mathscr{X},\mathscr{E}_{\mathscr{X},\mathscr{X}})$, where $\mathscr{X}$ is the set of the state variables and $\mathscr{E}_{\mathscr{X},\mathscr{X}} = \{(x_j,x_i) \mid \bar{A}_{ij} = 1\}$ describes the relations among them. Considering two sets X_1 and X_2, $X_1,X_2 \subseteq \mathscr{X}$, the set X_1 is said to be *completely connected* to X_2 if for all $x_i \in X_1$ there exists a directed path from x_i to some $x_j \in X_2$. This is denoted by $X_1 \xrightarrow{CC} X_2$.

Proposition 6 *Let $\bar{A} \in \{0,1\}^{n\times n}$ be the structural pattern of some state matrix A and $\mathscr{D}(\bar{A}) = (\mathscr{X},\mathscr{E}_{\mathscr{X},\mathscr{X}})$ its associated directed graph. Let $\mathscr{O}_C = \{I \subseteq \mathscr{X} \mid I \xrightarrow{CC} \mathscr{X} \setminus I\}$. Then, $M = (\mathscr{X},\mathscr{O}_C)$ is a matroid.* ◇

Proof The hereditary property is trivially verified. In order to prove the augmentation property, consider $I,J \in \mathscr{O}_C$ such that $|J| > |I|$. Let $x_i \in J \setminus I$. Since $x_i \in J$, there exists at least one $x_j \in \mathscr{X} \setminus J$ such that there is a directed path from x_i to x_j. Now, there are two possibilities. If $x_j \notin I$, then we can add x_i to I since x_j remaining outside I will guarantee the path from x_i to x_j. On the other hand, if $x_j \in I$, then there exists another $x_k \in \mathscr{X} \setminus I$ such that $x_j \to x_k$ and since $x_i \to x_j$, it must be by transitivity that $x_i \to x_k$. Thus, $I \cup \{x_i\} \in \mathscr{O}_C$ and the proposition is proved. □

Definition 30 (Output-Connected Matroid) *The matroid of Proposition 6 will be referred as the **output-connected matroid** associated with the structural state matrix $\bar{A}$.* ◇

Example 15 *Consider the directed graph representation of a matrix $\bar{A} \in \{0,1\}^{4\times 4}$, depicted in Figure 3.7, and whose vertex set is $\mathscr{X} = \{x_1,x_2,x_3,x_4,x_5\}$. Let $\mathscr{O}_C = \{I \subseteq \mathscr{X} \mid I \xrightarrow{CC} \mathscr{X} \setminus I\}$ be the collection of independent sets of the associated output-connected matroid. Notice that x_2 cannot belong to any $I \in \mathscr{O}_C$, because there is no state variable $x_i \in \mathscr{X}$, with $i \neq 2$, such that there*

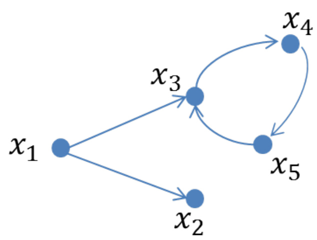


Figure 3.7: A structural directed graph representation associated with an output-connected matroid $M = (\mathscr{X}, \mathscr{O}_C)$ where $\mathscr{O}_C = \{I \subseteq \mathscr{X} \mid I \xrightarrow{CC} \mathscr{X} \setminus I\}$.

exists a directed path from x_2 to x_i. Furthermore, we may write $\mathscr{O}_C$ as $\mathscr{O}_C = \{\{x_1\}\} \cup \{\mathscr{X}_S \subseteq \{x_3, x_4, x_5\} \mid |\mathscr{X}_S| \leq 2\}$.

Another necessary condition to satisfy structural observability is that the directed graph $\mathscr{D}(\bar{A}, \bar{C})$ associated with the system must be free of contractions. Similar to the construction given in Definition 20, one can associate with $\bar{A}$ the bipartite graph $\mathscr{B}(\bar{A}) = (\mathscr{X}, \mathscr{X}^+, E_{\mathscr{X},\mathscr{X}^+})$, where $\mathscr{X} = \{x_1, \ldots, x_n\}$, $\mathscr{X}^+ = \{x_1^+, \ldots, x_n^+\}$ and $E_{\mathscr{X},\mathscr{X}^+} = \{(x_j, x_i^+) \mid \bar{A}_{ij} = 1\}$.

Proposition 7 *Let $\bar{A} \in \{0,1\}^{n \times n}$ be the structural pattern of some state matrix A and $\mathscr{B}(\bar{A}) = (\mathscr{X}, \mathscr{X}^+, E_{\mathscr{X},\mathscr{X}^+})$ the associated bipartite graph. Let $\mathscr{C}_F$ be the collection of subsets I of $\mathscr{X}$ for which there exists a matching in $\mathscr{B}(\bar{A})$ that covers I. Then, $M = (\mathscr{X}, \mathscr{C}_F)$ is a matroid.* $\diamond$

Proof Consider that $I \in \mathscr{C}_F$. Then, for every subset J of I there is also a matching that covers J and therefore $J \in \mathscr{C}_F$. In order to prove the augmentation property, consider Theorem 7. Let $I, J \in \mathscr{C}_F$ with $|J \setminus I| = 2$ and $|I \setminus J| = 1$. We will apply induction on $|I|$ to prove that there always exists $x_i \in J \setminus I$ such that $I \cup \{x_i\} \in \mathscr{C}_F$. For the basis, let $I \setminus J = I = \{x_\alpha\}$. Since $I \in \mathscr{C}_F$, there exists an edge $(x_\alpha, x_{\alpha_2}^+)$ for some $\alpha_2 \in \{1, \ldots, n\}$. Similarly, let $J \setminus I = J = \{x_\beta, x_\gamma\}$. Since $J \in \mathscr{C}_F$, there must exist edges $(x_\beta, x_{\beta_2}^+)$ and $(x_\gamma, x_{\gamma_2}^+)$ for appropriate $\beta_2, \gamma_2 \in \{1, \ldots, n\}$ with $\beta_2 \neq \gamma_2$. Notice that $\alpha \neq \beta$ and $\alpha \neq \gamma$. Considering, for instance, that β_2 might be equal to α_2 it is always possible to add the edge $(x_\gamma, x_{\gamma_2}^+)$ to the matching M_I that covered I. Therefore, $I \cup \{x_\gamma\} \in \mathscr{C}_F$. On the other hand, if $\gamma_2 = \alpha_2$, then $I \cup \{x_\beta\}$ must belong to $\mathscr{C}_F$. Finally, if both β_2 and γ_2 are different from α_2, then it is possible to add either $\{x_\beta\}$ or $\{x_\gamma\}$ so that the augmented I remains independent.

For the inductive step, let $|I| = k$ with $k > 1$. Then J is comprised of $k+2$ elements and again we can consider that $I \setminus J = \{x_\alpha\}$, for $\alpha \in \{1, \ldots, n\}$, with $(x_\alpha, x_{\alpha_2}^+)$ an edge of the matching that covers I. Let us take an edge $(x_\beta, x_{\beta_2}^+)$ from the matching M_J that covers J such that $x_\beta \in J \setminus I$ and $x_{\beta_2}^+ \neq x_{\alpha_2}^+$. Notice that such an edge must always exist since, by hypothesis, $|J \setminus I| > |I \setminus J|$. The vertex $x_{\beta_2}^+$ may be covered or not by the matching of I. If $x_{\beta_2}^+$ is not covered, then surely $I \cup \{x_\beta\} \in \mathscr{C}_F$. On the other hand, if $x_{\beta_2}^+$ is covered by M_I, then there exists an edge $(x_i, x_{\beta_2}^+) \in M_I$ with $x_i \in I \cap J$. Build the sets $I' = I \setminus \{x_i\}$ and $J' = J \setminus \{x_\beta\}$ and remove the vertex $x_{\beta_2}^+$ from $\mathscr{X}^+$. Then, $|I'| = k-1$ and $|I' \setminus J'| = 1$ while $|J' \setminus I'| = 2$. Thus, by the induction hypothesis, there must exist $x_j \in J' \setminus I'$ such that $I' \cup \{x_j\} \in \mathscr{C}_F$. Notice that x_j is either x_i or x_γ, where x_γ is the other element of $J \setminus I$. In the first case, we can add x_β to $I' \cup \{x_i\} = I$ and we have $I \cup \{x_\beta\} \in \mathscr{C}_F$. In the second case, we can simply add x_i to $I' \cup \{x_\gamma\}$ and, therefore $I \cup \{x_\gamma\} \in \mathscr{C}_F$. $\square$

Definition 31 (Contraction-free Matroid) *The matroid of Proposition 7 will be referred as the* ***contraction-free matroid*** *associated with the structural state matrix $\bar{A}$.* ⋄

According to the Output Selection Problem (3.4), an appropriate subset of state variables, $S_{\mathscr{O}} \subseteq \mathscr{X}$, has to be selected in order to achieve structural observability. Remember that the output configuration is dedicated. Thus, if $S_{\mathscr{O}} = \{x_{j_1}, \ldots, x_{j_p}\}$ then the matrix $\bar{C}$ can be written as $I_n^{S_{\mathscr{O}}}$ where the latter is the identity matrix of order n but only with rows $j_1, \ldots, j_p$. Following the results from Proposition 6 and Proposition 7, one can obtain the following.

Theorem 12 *Let $\bar{A} \in \{0,1\}^{n\times n}$ be the structural pattern of some state matrix A, $\mathscr{X} = \{x_1, \ldots, x_n\}$ the set of state variables and $S_{\mathscr{O}} \subseteq \mathscr{X}$ a given dedicated output configuration. Additionally, let $M_1 = (\mathscr{X}, \mathscr{O}_C)$ and $M_2 = (\mathscr{X}, \mathscr{C}_F)$ be the output-connected and contraction-free matroids, respectively, associated with $\bar{A}$. Then, the pair $(\bar{A}, I_n^{S_{\mathscr{O}}})$ is structurally observable if and only if*

$$\mathscr{X} \setminus S_{\mathscr{O}} \in \mathscr{O}_C \cap \mathscr{C}_F.$$

⋄

Proof Let $S_{\mathscr{O}} = \{x_{j_1}, \ldots, x_{j_p}\}$ be the dedicated output configuration. According to Theorem 5 the pair $(\bar{A}, I_n^{S_{\mathscr{O}}})$ is structurally observable if and only if the directed graph $\mathscr{D}(\bar{A}, I_n^{S_{\mathscr{O}}})$ is output-connected and free of contractions. Since outputs are placed in the state variables that belong to $S_{\mathscr{O}}$, obviously there is always a directed path between any $x_i \in S_{\mathscr{O}}$ and some output vertex, that is comprised of a single arc. Thus, testing for output-connectivity is restricted to the set $\mathscr{X} \setminus S_{\mathscr{O}}$, which has to belong to $\mathscr{O}_C$ such that the pair $(\bar{A}, I_n^{S_{\mathscr{O}}})$ be output-connected. Regarding the contraction condition, note that as it was seen in Proposition 2, the system is free of contractions if and only if there exists a matching in the bipartite graph $\mathscr{B}(\bar{A}, \bar{C}) = (\mathscr{X}^-, \mathscr{X}^+ \cup \mathscr{Y}, E_{\mathscr{X}^-,\mathscr{X}^+} \cup E_{\mathscr{X}^-,\mathscr{Y}})$ that covers $\mathscr{X}^-$. In order to maintain the previously used notation, relabel each $x_i^- \in \mathscr{X}^-$ to x_i. Since $\bar{C} = I_n^{S_{\mathscr{O}}}$, for any $x_i \in S_{\mathscr{O}}$ there is one and only one y_j in $\mathscr{Y}$ such that $(x_i, y_j) \in E_{\mathscr{X}^-,\mathscr{Y}}$. Thus, x_i can always be covered by the edge (x_i, y_j) and the existence of a matching that covers $\mathscr{X}$ is equivalent to the existence of a matching in $\mathscr{B}(\bar{A})$ that covers the state variables in $\mathscr{X} \setminus S_{\mathscr{O}}$. Therefore, $\mathscr{X} \setminus S_{\mathscr{O}}$ must belong to $\mathscr{C}_F$. Notice that, by similar arguments, if $\mathscr{X} \setminus S_{\mathscr{O}} \in \mathscr{O}_C \cap \mathscr{C}_F$, then we conclude that the pair $(\bar{A}, I_n^{S_{\mathscr{O}}})$ is at the same time output-connected and free of contractions. Therefore, by Theorem 5, the pair $(\bar{A}, I_n^{S_{\mathscr{O}}})$ is structurally observable. □

With the previous theorem, all the feasible (in the sense that the system is structurally observable) dedicated output configurations can be written as the intersection of two matroids. Algorithm 4 and Algorithm 5 describe independence oracles associated with the output-connected and contraction-free matroids, respectively.

At this point we are ready to solve the problem of, given the structural pattern $\bar{A} \in \{0,1\}^{n\times n}$ of some state matrix A, find the minimum number of outputs and where to place them in order to achieve structural observability. The idea is to apply the matroid intersection algorithm (Algorithm 3) to $M_1 = (\mathscr{X}, \mathscr{O}_C)$ and $M_2 = (\mathscr{X}, \mathscr{O}_F)$ and the returned value will be a set $I^* \in \mathscr{O}_C \cap \mathscr{C}_F$ of

ALGORITHM 4: Independence Oracle for output-connected matroid $M_1 = (\mathscr{X}, \mathscr{O}_C)$

Input: Matrix $\bar{A} \in \{0,1\}^{n\times n}$ and the set I, with $I \subseteq \mathscr{X} = \{x_1, \ldots, x_n\}$
Output: TRUE if $I \in \mathscr{O}_C$; FALSE, otherwise
Let $\mathscr{X} \setminus I = \{x_{i_1}, \ldots, x_{i_p}\}$.
for $j = 1, \ldots, p$ **do**
  $\mathscr{O}_j^1 = \{x_k \in \mathscr{X} \mid \bar{A}_{k,i_j} = 1\}$
end for
for $j = 1, \ldots, p$ **do**
  for $l = 1, \ldots, n-1$ **do**
    $\mathscr{O}_j^l = \{x_k \in \mathscr{X} \mid \bar{A}_{k,r} = 1$ with $x_r \in \mathscr{O}_j^{l-1}\} \cup \mathscr{O}_j^{l-1}$
  end for
end for
if $\cup_{1\le j\le p} \mathscr{O}_j^{n-1} = I$ **then**
  Return TRUE.
else
  Return FALSE.

ALGORITHM 5: Independence Oracle for contraction-free matroid $M_2 = (\mathscr{X}, \mathscr{C}_F)$

Input: Matrix $\bar{A} \in \{0,1\}^{n\times n}$ and the set I, with $I \subseteq \mathscr{X} = \{x_1, \ldots, x_n\}$
Output: TRUE if $I \in \mathscr{C}_F$; FALSE, otherwise
Construct $\mathscr{B} = (I, \mathscr{X}^+, E_{I,\mathscr{X}^+})$:

1. $\mathscr{X}^+ = \{x_1^+, \ldots, x_n^+\}$
2. $E_{I,\mathscr{X}^+} = \{(x_j, x_i^+) \mid \bar{A}_{ij} = 1, \forall x_i^+ \in \mathscr{X}^+, \forall x_j \in I\}$

Find a maximum matching in $\mathscr{B}$ with the Hopcroft-Karp algorithm. Let $I^M \subseteq I$ be the vertices that belong to the matching.
if $I^M = I$ **then**
  Return TRUE.
else
  Return False.

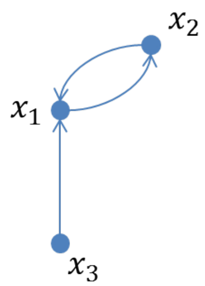


Figure 3.8: Example 16: a directed graph in which the weighted matroid intersection algorithm is applied. The cost of placing an output in x_1, x_2 and x_3 is $1, 3$ and 1, respectively.

maximum-cardinality. Since, by Theorem 12, $S_{\mathscr{O}} \subseteq \mathscr{X}$ is a feasible solution if and only if $\mathscr{X} \setminus S_{\mathscr{O}} \in \mathscr{O}_C \cap \mathscr{C}_F$, then a minimal feasible dedicated output configuration (FDOC) is surely $\mathscr{X} \setminus I^*$. Furthermore, since the subroutines to test for independence in M_1 (Algorithm 4) and M_2 (Algorithm 5) have polynomial complexity, the overall algorithm is polynomial in time.

Corollary 1 *Let $\bar{A} \in \{0,1\}^{n\times n}$ be the structural pattern of some state matrix A and $\mathscr{X} = \{x_1, \ldots, x_n\}$ be the set of state variables. A minimal possible feasible dedicated output configuration $S_{\mathscr{O}} \subseteq \mathscr{X}$ to achieve structural observability can be found in polynomial time.* ◇

Proof The correctness of the solution follows directly from Theorem 12 and Theorem 11. The polynomial time complexity is consequence of the polynomial algorithms Algorithm 4 and Algorithm 5. □

When there are costs associated with the output allocation, we just have to apply the weighted-matroid intersection algorithm to the output-connected matroid $M_1 = (\mathscr{X}, \mathscr{O}_C)$ and the contraction-free matroid $M_2 = (\mathscr{X}, \mathscr{C}_F)$. With similar arguments as the ones stated before, we can build a feasible dedicated output configuration of minimum cost $S_{\mathscr{O}}$ with $S_{\mathscr{O}} = \mathscr{X} \setminus I^*$, where I^* is a maximum-weighted independent set in $\mathscr{O}_C$ and $\mathscr{C}_F$.

Example 16 *Let $\bar{A} \in \{0,1\}^{n\times n}$ be a structural matrix whose directed graph $\mathscr{D}(\bar{A}) = (\mathscr{X}, \mathscr{E}_{\mathscr{X},\mathscr{X}})$ is depicted in Figure 3.8. Consider the problem of finding the minimum-cost feasible dedicated output configuration, $S_{\mathscr{O}} \subseteq \mathscr{X}$, where the cost of placing an output in x_1, x_2 and x_3 is $1, 3$ and 1, respectively. The output-connected matroid $M_1 = (\mathscr{X}, \mathscr{O}_C)$ has a collection of independent sets given by $\mathscr{O}_C = \{\emptyset, \{x_1\}, \{x_2\}, \{x_3\}, \{x_1, x_3\}, \{x_2, x_3\}\}$. It can be seen that all the independent sets of $\mathscr{O}_C$ belong to $\mathscr{C}_F$ except the set $\{x_2, x_3\}$, where $M_2 = (\mathscr{X}, \mathscr{C}_F)$ is the contraction-free matroid. Thus, $\mathscr{O}_C \cap \mathscr{C}_F = \{\emptyset, \{x_1\}, \{x_2\}, \{x_3\}, \{x_1, x_3\}\}$ and the set $I_1 = \{x_1, x_3\}$ is a maximum-cardinality set in $\mathscr{O}_C \cap \mathscr{C}_F$, whereas the set $I_2 = \{x_2\}$ is a maximum-weight set in $\mathscr{O}_C \cap \mathscr{C}_F$. Therefore, $\mathscr{X} \setminus I_1 = \{x_2\}$ is a minimum feasible dedicated output configuration and $\mathscr{X} \setminus I_2 = \{x_1, x_3\}$ is a feasible dedicated output configuration of minimum cost.*

3.4 Interpreting the Matroid Intersection Algorithm for Output Selection

Even though the matroid intersection algorithm (Algorithm 3) has polynomial time, for some cases it can be simplified to reduce the overall complexity. For instance, in Example 14, the problem of finding a maximum matching in a bipartite graph was reduced to the matroid intersection problem. If we study carefully that algorithm when applied to the bipartite matching, it can be seen that some steps can be simplified and the result will be a more efficient procedure to find a maximum matching. In this subsection, a similar analysis will be performed on Algorithm 3 when applied to the output-connected (Proposition 6) and contraction-free (Proposition 7) matroids.

The output-connected matroid deals with the presence of directed paths between vertices of a given directed graph. Problems involving connectivity in a directed graph can be greatly simplified by partitioning that graph into subgraphs with special properties. Within structural systems theory, that decomposition is frequently used when dealing with structural observability [13]. Before we pursue, we first present some useful notions from graph theory.

Definition 32 *Let $\mathscr{D} = (\mathscr{V}, \mathscr{E})$ be a directed graph. For any two vertices, v_i, $v_j \in \mathscr{V}$, we say that $v_i \sim v_j$, i.e., that v_i is equivalent to v_j, if and only if there is a directed path from v_i to v_j and a directed path from v_j to v_i.* ◇

Notice that the previously defined relation is an equivalence relation. More precisely, it satisfies the following three properties: for any $v_i \in \mathscr{V}$, $v_i \sim v_i$ (consider a directed path of zero length). Symmetry also holds, directly from the definition, and it is straightforward to see that, for $v_i, v_j, v_k \in \mathscr{V}$, if $v_i \sim v_j$ and $v_j \sim v_k$, then $v_i \sim v_k$ (transitivity). Furthermore, it is possible to partition $\mathscr{V}$ into equivalence classes, under the defined equivalence relation. If $\mathscr{D} = (\mathscr{V}, \mathscr{E})$ is *strongly connected*, i.e, every vertex in $\mathscr{V}$ is reachable from every other vertex, there will be only one equivalence class, $\mathscr{V}$. Each equivalence class will be called a strong connected component (SCC).

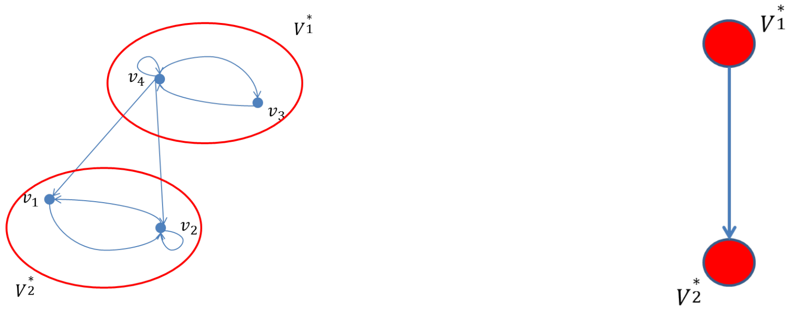
Definition 33 (Strong Connected Component (SCC)) *Let $D = (\mathscr{V}, \mathscr{E})$ be a directed graph. For $\mathscr{V}_s \subseteq \mathscr{V}$, a subgraph $\mathscr{D}_s = (\mathscr{V}_s, (\mathscr{V}_s \times \mathscr{V}_s) \cap \mathscr{E})$ is a **strong connected component** of $\mathscr{D}$ if it is strongly connected and there are no two vertices $v_i \in \mathscr{V}_s$ and $v_j \notin \mathscr{V}_s$ such that $v_i \sim v_j$.* ◇

Notice that the equivalence classes under $\sim$ and the strong connected components are equivalent concepts. Furthermore, they induce a partition in the original digraph.

Definition 34 (Condensation) *Let $\mathscr{D} = (\mathscr{V}, \mathscr{E})$ be a directed graph and let $\{\mathscr{V}_1, \ldots, \mathscr{V}_N\}$ be the strong connected components of $\mathscr{D}$. Define*

$$\begin{aligned} \mathscr{V}^* &= \{\mathscr{V}_i \mid \mathscr{V}_i \text{ is a SCC}\}, \\ \mathscr{E}^* &= \{(\mathscr{V}_j, \mathscr{V}_i) \mid i \neq j \text{ and } \exists v_j \in \mathscr{V}_j, v_i \in \mathscr{V}_i \text{ such that } (v_j, v_i) \in \mathscr{E}\}. \end{aligned}$$

*The digraph $\mathscr{D}^* = (\mathscr{V}^*, \mathscr{E}^*)$ is called the **condensation** of $\mathscr{D}$.* ◇



(a) a directed graph

(b) a condensation.

Figure 3.9: A directed graph in a) and the associated condensation in b). The red circles are the SCCs.

Example 17 *Consider the directed graph $\mathscr{D}$ depicted in Figure 3.9 a), whose vertex set is $\mathscr{V} = \{v_1, v_2, v_3, v_4\}$. Since $v_1 \sim v_2$ and $v_3 \sim v_4$, the strong connected components are $V_1^* = \{v_3, v_4\}$ and $V_2^* = \{v_1, v_2\}$. Notice, for instance, that (v_4, v_1) is an arc of $\mathscr{D}$. Therefore, (V_1^*, V_2^*) is an arc of the condensation of $\mathscr{D}$, whose representation is given in Figure 3.9 b).*

Theorem 13 *Given any directed graph $\mathscr{D} = (\mathscr{V}, \mathscr{E})$, its associated condensation $\mathscr{D}^*$ is acyclic.* $\diamond$

Proof Suppose that there is a cycle in $\mathscr{D}^* = (\mathscr{V}^*, \mathscr{E}^*)$. Then, there must exist $\mathscr{V}_i$, $\mathscr{V}_j \in \mathscr{V}^*$ such that there is a directed path from $\mathscr{V}_i$ to $\mathscr{V}_j$ and a directed path from $\mathscr{V}_j$ to $\mathscr{V}_i$. By definition of $\mathscr{D}^*$, we know that exist $v_{i_1}, v_{i_2} \in \mathscr{V}_i$ and $v_{j_1}, v_{j_2} \in \mathscr{V}_j$ such that $v_{i_1} \to v_{j_1}$ and $v_{j_2} \to v_{i_2}$. But since $v_{i_1} \sim v_{i_2}$ and $v_{j_1} \sim v_{j_2}$, it must be that $v_{j_1} \to v_{i_1}$. Thus, $v_{i_1} \sim v_{j_1}$ contradicting the fact that v_{i_1} and v_{j_1} belong to distinct SCC's. $\square$

The properties of the condensation digraph are well-suited to the study of connectivity in directed graphs. In particular, from a structural directed graph, one can efficiently construct the associated condensation using a depth-first search algorithm [33]. In fact, that condensation can be applied only to the directed graph representation of the structural system matrix $\bar{A} \in \{0,1\}^{n\times n}$. Notice that for each output variable $y_i \in \mathscr{Y}$, the possible arcs are (x_k, y_i) with $x_k \in \mathscr{X}$. Therefore, each $y_i \in \mathscr{Y}$ is a strong connected component.

Theorem 14 *Let $(\bar{A}, \bar{C})$ be a structural pair and $\mathscr{D}(\bar{A}, \bar{C}) = (\mathscr{X} \cup \mathscr{Y}, \mathscr{E}_{\mathscr{X},\mathscr{X}} \cup \mathscr{E}_{\mathscr{X},\mathscr{Y}})$ its associated directed graph representation. The digraph $\mathscr{D}(\bar{A}, \bar{C})$ is output-connected if and only if its condensation $\mathscr{D}^* = (\mathscr{X}^* \cup \mathscr{Y}^*, \mathscr{E}^*_{\mathscr{X}^*,\mathscr{X}^*} \cup \mathscr{E}^*_{\mathscr{X}^*,\mathscr{Y}^*})$ is output-connected.* $\diamond$

Proof If $\mathscr{D}^*$ is output-connected, then for any $X_k \in \mathscr{X}^*$ there must exist some $Y_j \in \mathscr{Y}^*$ such that $X_i \to Y_j$. Then, by definition of condensation, for every $x_i \in X_k$ there exist a directed path to $y_j \in Y_j$. Therefore, $\mathscr{D}(\bar{A}, \bar{C})$ is output-connected.

If $\mathscr{D}^*$ is not output-connected, there must exist at least one $X_i \in \mathscr{X}^*$ such that there is no path from X_i to every $Y_j \in \mathscr{Y}^*$. Therefore, by definition of condensation, no $x_k \in \mathscr{X}_i$ can reach an $y_j \in \mathscr{Y}$ and $\mathscr{D}$ is not output-connected. $\square$

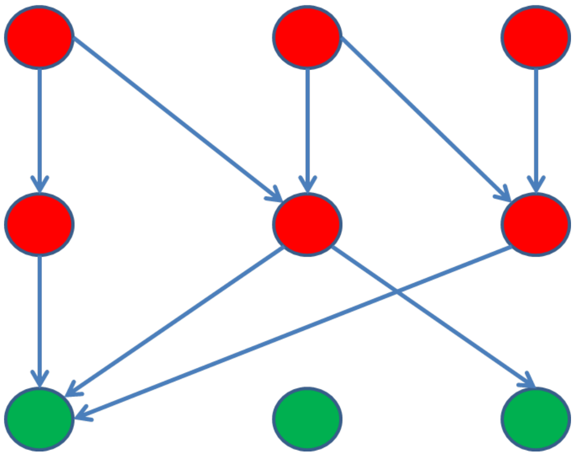


Figure 3.10: A representation of a condensation of some structural directed graph. The circles represent the strong connected components, where the green ones are non-bottom linked SCC's.

With the previously presented tools in mind, one can perceive in a better way the structure of the independent sets of the output-connected matroid $M = (\mathscr{X}, \mathscr{O}_C)$. Let us recall that for a given directed graph $\mathscr{D}(\bar{A}) = (\mathscr{X}, \mathscr{E}_{\mathscr{X},\mathscr{X}})$, the collection of independent sets of the output-connected matroid was defined as $\mathscr{O}_C = \{I \subseteq \mathscr{X} \mid I \overset{CC}{\longrightarrow} \mathscr{X} \setminus I\}$ (Proposition 7). In other words, I is an independent set if it is possible to have a directed path in $\mathscr{D}(\bar{A})$ for each state variable in I to some vertex that does not belong to I. Considering the condensation of $\mathscr{D}(\bar{A})$, the state variables set can be partitioned into strong connected components

$$\mathscr{X} = \cup_{i=1}^{k} \mathscr{X}_i,$$

where $\mathscr{X}_i \cap \mathscr{X}_j = \emptyset$, for $i \neq j$. A SCC without outgoing arcs is usually denoted in literature, see for instance [13], as a *non-bottom linked* SCC, i.e, $\mathscr{X}_i$ is a non-bottom linked SCC if $(\mathscr{X}_i, \mathscr{X}_j) \notin \mathscr{E}^*, \forall \mathscr{X}_j \in \mathscr{X}^*$ (see Figure 3.10).

Proposition 8 *Let $\mathscr{D}(\bar{A}) = (\mathscr{X}, \mathscr{E}_{\mathscr{X},\mathscr{X}})$ be a structural directed graph representation for some structural state matrix $\bar{A} \in \{0,1\}^{n\times n}$. Additionally, let $\mathscr{D}^* = (\mathscr{X}^*, \mathscr{E}^*_{\mathscr{X}^*,\mathscr{X}^*})$ be its associated condensation digraph and $M = (\mathscr{X}, \mathscr{O}_C)$ the output-connected matroid of $\bar{A}$. Then, for $I \subseteq \mathscr{X}$, I belongs to $\mathscr{O}_C$ if and only if*

$$|I \cap \mathscr{X}_i| < |\mathscr{X}_i|, \tag{3.8}$$

for all $\mathscr{X}_i \in \mathscr{X}^*$ such that $\mathscr{X}_i$ is a non-bottom linked SCC. ◇

Proof Let $I \in \mathscr{O}_C$ and assume that there is some non-bottom linked SCC $\mathscr{X}_k$ such that $\mathscr{X}_k \subseteq I$. By definition of $\mathscr{O}_C$, for any $x_i \in I$, there exists a directed path to some $x_j \in \mathscr{X} \setminus I$. Take $x_i \in \mathscr{X}_k$. Notice that it is impossible to have a path from x_i to another state variable in $\mathscr{X} \setminus I$, since x_i belongs to a non-bottom linked SCC $\mathscr{X}_k$ and the other state variables that might belong to the same SCC are contained in I. This is a contradiction with the fact that $I \in \mathscr{O}_C$. Thus, $I \in \mathscr{O}_C$ implies (3.8).

On the other hand, let us take some I such that (3.8) is satisfied. Suppose that $I \notin \mathscr{O}_C$. Then, there exists some variable $x_i \in I$ such that there is no path from x_i to any $x_j \notin I$. If x_i is in a non-bottom linked SCC $\mathscr{X}_k$, then $|\mathscr{X}_k| \geq 2$ and surely there is another $x_j \in \mathscr{X}_k$, with $x_j \notin I$, such that $x_i \rightarrow x_j$ what contradicts the initial assumption that $I \notin \mathscr{O}_C$. On the other hand, consider that x_i belongs to a SCC, other than a non-bottom linked SCC. Notice that, by definition of non-bottom linked SCC, it is always possible to have a path from any $x_i \in \mathscr{X}$, such that x_i is not in a non-bottom linked SCC, to some $x_j \in \mathscr{X}$, with x_j contained in a non-bottom linked SCC. Therefore, since (3.8) is satisfied, we have again a contradiction. □

ALGORITHM 6: Minimum Sensor Selection for Structural Observability

---

Input: Directed graph $\mathscr{D}(\bar{A}) = (\mathscr{X}, \mathscr{E}_{\mathscr{X},\mathscr{X}})$ associated with a structural state matrix $\bar{A} \in \{0,1\}^{n\times n}$
Output: A minimum-size feasible dedicated output configuration $S_{\mathscr{O}} \subseteq \mathscr{X}$

STEP 1: Compute the condensation of $\mathscr{D}(\bar{A})$

Build the condensation of $\mathscr{D}(\bar{A})$. Let $\mathscr{D}^* = (\mathscr{X}^*, \mathscr{E}_{\mathscr{X}^*,\mathscr{X}^*})$ be that condensation.

STEP 2: Find the non-bottom linked SCCs in $\mathscr{D}^*$
$SCC_nBottom \leftarrow \emptyset$
for $j = 1, \ldots, |\mathscr{X}^*|$ **do**
  if $out_{degree}(\mathscr{X}_j) = 0$ **then**
    $SCC_nBottom \leftarrow \mathscr{X}_j$

STEP 3: Build the weighted bipartite graph

Construct $\mathscr{B} = (\mathscr{V}_1, \mathscr{V}_2, E)$ and a weight function $w : E \rightarrow \mathbb{N}$

1. $\mathscr{V}_1 = \{x_1, \ldots, x_n\}$
2. $\mathscr{V}_2 = \{x_1^+, \ldots, x_n^+, z_1^*, \ldots, z^*_{i_{|SCC_nBottom|}}\}$
3. $E = E_{\mathscr{X},\mathscr{X}^+} \cup E_{\mathscr{X},Z^*}$
4. $\mathscr{E}_{\mathscr{X},\mathscr{X}^+} = \{(x_j, x_i^+) \mid \bar{A}_{ij} = 1\}$
5. $\mathscr{E}_{\mathscr{X},Z^*} = \{(x_i, z_j^*)$ such that x_i belongs to the non-bottom linked SCC $\mathscr{X}_j\}$
6. $w(e) = 1$ if $e \in \mathscr{E}_{\mathscr{X},\mathscr{X}^+}$
7. $w(e) = n+1$ if $e \in \mathscr{E}_{\mathscr{X},Z^*}$

STEP 4: Find a maximum-weight matching

1. Find a maximum-weight matching M_w in $\mathscr{B}$
2. Set $V_M \subseteq \mathscr{V}_1$ to the vertices covered by the matching
3. $V_M \leftarrow V_M \setminus \{x_i \in \mathscr{X} \mid (x_i, z_j^*) \in M_w\}$
4. Return $\mathscr{X} \setminus V_M$

With the previous proposition, an efficient algorithm (see Algorithm 6) to find a maximum-cardinality set that is independent in both the output-connected and contraction-free matroids can be constructed. Recall that a set $I \subseteq \mathscr{X}$ belongs to the collection of independent sets of the contraction-free matroid $M_2 = (\mathscr{X}, \mathscr{C}_F)$ if there exists a matching in $\mathscr{B}(\bar{A}) = (\mathscr{X}, \mathscr{X}^+, E_{\mathscr{X},\mathscr{X}^+})$ that covers I. At the same time, we have to guarantee that I belongs to $\mathscr{O}_C$ where $\mathscr{O}_C$ is the collection of independent sets of the output-connected matroid, which is equivalent to say that at least one state variable per non-bottom SCC is left out of I (Proposition 8).

With the previously considerations in mind, we can construct the following bipartite graph. Let $\mathscr{B}(\bar{A}) = (\mathscr{X}, \mathscr{X}^+, E_{\mathscr{X},\mathscr{X}^+})$ be the bipartite graph associated with $\bar{A}$ and let $\{\mathscr{X}_1, \ldots, \mathscr{X}_k\}$ be the set of the the non-bottom SCCs resulting from the condensation of $\mathscr{D}(\bar{A})$. For each non-bottom SCC $\mathscr{X}_i$, we will introduce in $\mathscr{B}$ a *dummy* variable z_i^*. Hereafter, we will consider the extended bipartite graph $\mathscr{B}^*(\bar{A}) = (\mathscr{X}, \mathscr{X}^+ \cup Z^*, E_{\mathscr{X},\mathscr{X}^+} \cup E_{\mathscr{X},Z^*})$ where $Z^* = \{z_1^*, \ldots, z_k^*\}$ and $E_{\mathscr{X},Z^*} = \{(x_i, z_j^*) \mid x_i \in \mathscr{X}_i\}$. Additionally, we will consider the weight-function $W : E_{\mathscr{X},\mathscr{X}^+} \cup E_{\mathscr{X},Z^*} \to \mathbb{N}$, where $w(e) = 1$ if $e \in E_{\mathscr{X},\mathscr{X}^+}$ and $w(e) = n+1$ if $e \in E_{\mathscr{X},Z^*}$. In order to obtain a maximum-cardinality set I^* in $\mathscr{O}_C \cap \mathscr{C}_F$, one can compute a maximum-weight matching M_W in $\mathscr{B}^*(\bar{A})$. Let $\mathscr{X}_M \subseteq \mathscr{X}$ be the state variables covered by M_W. To obtain I^*, we remove the state variables that are covered by an edge in $E_{\mathscr{X},Z^*}$, i.e, $I^* = \mathscr{X}_M \setminus \{x_i \mid (x_i, z_j^*) \in M_W \text{ for some } z_j^* \in Z^*\}$. In that manner, we always *force* the removal of at least one state variable per non-bottom SCC and we guarantee that the matching obtained is as maximum as possible.

Algorithm 6 can be easily modified in order to solve Problem 3.4, i.e., the problem of finding a minimum-cost feasible dedicated output configuration for a given cost function $\mathscr{C} : \mathscr{X} \to \mathbb{N}^+$. The idea is to change the weight function in the third step. If $e = (x_i, x_j^+)$ for some $x_j^+ \in \mathscr{X}^+$, then $w(e) = \mathscr{C}(x_i)$. On the other hand, $w(e) = \sum_{i=1}^{n} \mathscr{C}(x_i) + 1$ if $e = (x_i, z_j^*)$ for some $z_j^* \in \mathscr{Z}^*$.

Chapter 4

Output Selection Problem with Generic Observability Index Constraint

In this chapter we analyze the output selection problem when we include a performance restriction given by the generic observability index. First, we will formulate the problem. Then, we will show that it is NP-complete.

4.1 Generic Observability Index: Problem Formulation

Consider a network of n entities where each entity (denoted by x_i) updates its data according to the linear dynamics

$$x(t+1) = Ax(t), \tag{4.1}$$

where $A \in \mathbb{R}^{n \times n}$ is the state matrix and $x(t) \in \mathbb{R}^n$ is the state vector. The goal is to design the output matrix C

$$y(t) = Cx(t), \tag{4.2}$$

where $y \in \mathbb{R}^p$ is the output vector such that the pair (A, C) is observable.

In the previous chapter, we developed a polynomial-time algorithm to solve the structural related problem (see Problem (3.4)) that consists of finding a minimum-size dedicated output configuration $S_{\mathscr{O}} = \{x_{i_1}, \ldots, x_{i_p}\}$ in order to guarantee structural observability. We could construct the pattern $\bar{C}$ of the output matrix with $\bar{C} = I_n^{S_{\mathscr{O}}}$, where $I_n^{S_{\mathscr{O}}}$ is the identity matrix of size n with rows $i_1, \ldots, i_p$. Then, a numerical realization could be provided: $A \in [\bar{A}]$ and $C \in [\bar{C}]$ with (A, C) observable, where $[\bar{A}]$ and $[\bar{C}]$ are defined in Definition 4.

Another far more difficult but more practical problem is when we also would like to impose some performance observability constraint. For example, when it is important to recover the initial state vector in the fewest iteration instant $t \in \mathbb{N}$ as possible (e.g, a critical real-time network). As it was seen in Chapter 2, this constraint is directly related with the observability index μ, where μ is defined as

$$\mu(A, C) = min\ \{k \in \mathbb{N} \mid rank[\mathscr{O}(k)] = n\}, \tag{4.3}$$

where $\mathscr{O}(k)$ is the observability matrix at iteration instant k. Analogously, we can define the generic observability index μ_G as

$$\mu_G(\bar{A},\bar{C}) = min\ \{k \in \mathbb{N} \mid grank[\bar{O}(k)] = n\}, \tag{4.4}$$

where

$$\bar{O}(k) = \begin{bmatrix} \bar{C} \\ \bar{C}\bar{A} \\ \vdots \\ \bar{C}\bar{A}^{k-1} \end{bmatrix} \tag{4.5}$$

is the structural counterpart of the observability matrix at iteration instant k. Now, we can reformulate problem (3.4) that includes a performance constraint as follows.

Output Selection Problem with Generic Index Constraint: Consider the system (4.1) and let $\mathscr{X} = \{x_1,\dots,x_n\}$ be the set of state variables and $\bar{A} \in \{0,1\}^{n\times n}$ the structural pattern of A. The problem is to find, for a given k, with $1 \le k \le n$, the set of feasible dedicated output variables $S_{\mathscr{O}}$, with $S_{\mathscr{O}} \subseteq \mathscr{X}$, that solves the following optimization problem:

$$\begin{aligned} \min_{I\subseteq\{1,\dots,n\}} \quad & |S_{\mathscr{O}}| \\ \text{s.t.} \quad & \mu_G(\bar{A}, I_n^{S_{\mathscr{O}}}) \le k \end{aligned} \tag{4.6}$$

where $\mu_G(\bar{A}, I_n^{S_{\mathscr{O}}})$ is defined as in (4.4).

Notice that for $k = n$ (which is the same as for every $k \ge n$) the problem is equivalent to find a minimum feasible dedicated output configuration whereas for the case $k = 1$, the solution $S_{\mathscr{O}}$ will be the all set of state variables, $\mathscr{X}$, which implies that we would have one output connected to each state. However, as we will demonstrate in the next section, problem (4.6) is NP-complete, notwithstanding the fact that some extreme cases are easy to solve.

4.2 Computational Complexity Analysis

A computational problem $\mathscr{P}_1$ is said to be *polynomial-time reducible* to another problem $\mathscr{P}_2$ if there exists a procedure to transform $\mathscr{P}_1$ into $\mathscr{P}_2$ using a polynomial number of operations on the size of its inputs. More formally, $\mathscr{P}_1$ polynomial reduces to $\mathscr{P}_2$, which we denote by $\mathscr{P}_1 \le_P \mathscr{P}_2$, if $\mathscr{P}_1$ can be solved using a polynomial number of standard computational steps and a polynomial number of calls to an oracle that solves problem $\mathscr{P}_2$.

In what follows we will consider for $\mathscr{P}_1$ the set covering problem, which is a classic NP-complete problem in combinatorics and computer science, and is described as follows.

Set Covering Problem: Given a finite set $\mathscr{U} = \{a_1,\dots,a_m\}$ of m elements (called universe) and a collection $S = \{S_1, \dots, S_n\}$ of n subsets of $\mathscr{U}$ such that

$$\bigcup_{i=1}^{n} S_i = \mathscr{U}, \tag{4.7}$$

the set covering problem aims to determine a set of indices $I \subseteq \{1, \dots, n\}$ that solves the following optimization problem:

$$\begin{aligned} \min \quad & |I| \\ \text{s.t.} \quad & \bigcup_{i \in I} S_i = \mathscr{U} \end{aligned} \tag{4.8}$$

The following result will be used in order to demontrate the NP-completeness of problem (4.6).

Lemma 6 ([34]) *If problem $\mathscr{P}_1$ is NP-complete, problem $\mathscr{P}_2$ is in NP and $\mathscr{P}_1 \leq_P \mathscr{P}_2$, then $\mathscr{P}_2$ is NP-complete.* ◇

We will polynomial reduce the set covering problem to problem (4.6) in order to show the NP-completeness of the latter. Mainly, this suffices since problem (4.6) is in NP, given that there exists a polynomial algorithm to ascertain the satisfaction of the conditions on it. For that purpose, consider the following definitions and results.

Definition 35 *Consider an instance $(\mathscr{U}, S)$ of the set covering problem where $\mathscr{U} = \{a_1, \dots, a_m\}$ is the universe and $S = \{S_1, \dots, S_n\}$ is the collection of subsets of $\mathscr{U}$ such that $\bigcup_{i=1}^{n} S_i = \mathscr{U}$. Let $k = \max_i |S_i|$ and consider that $k \geq 2$. With that instance, we define the directed graph $\mathscr{D}_{SCP}(\mathscr{U}, S) = (\mathscr{U} \cup S^* \cup S_k, \mathscr{E}_{\mathscr{U},\mathscr{U}} \cup \mathscr{E}_{\mathscr{U},S^*} \cup \mathscr{E}_{S^*,S_k})$ given by*

$$\begin{aligned} \mathscr{U} &= \{a_1, \dots, a_m\}, \\ S^* &= \{S_1^*, \dots, S_n^*\}, \\ S_k &= \bigcup_{i=1}^{n} \bigcup_{j=1}^{k-1} \{s_{ij}\}, \\ \mathscr{E}_{\mathscr{U},\mathscr{U}} &= \{(a_i, a_i) \mid \forall a_i \in \mathscr{U}\}, \\ \mathscr{E}_{\mathscr{U},S^*} &= \{(a_i, S_j^*) \mid a_i \in S_j,\ \forall a_i \in \mathscr{U}, \forall S_j \in S\}, \\ \mathscr{E}_{S^*,S_k} &= \{(S_i^*, s_{i1}) \mid \forall S_i \in S, \forall s_{i1} \in S_k\} \cup \{(s_{ij}, s_{i(j+1)}) \mid i = 1, \dots, n, j = 1, \dots, k-2\}. \end{aligned}$$

◇

Example 18 *Consider the universe $\mathscr{U} = \{a_1, a_2, a_3, a_4\}$ and the collection of subsets of $\mathscr{U}$ given by $S = \{S_1, S_2, S_3\}$ where $S_1 = \{a_1, a_2, a_3\}$, $S_2 = \{a_2, a_4\}$ and $S_3 = \{a_3, a_4\}$. Notice that $\bigcup_{i=1}^{3} S_i = \mathscr{U}$. The construction given in Definition 35 is depicted in Figure 4.1.*

The following results establishes the relationship between the set covering problem and problem (4.6).

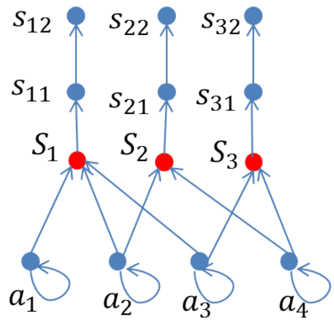


Figure 4.1: A directed graph representation associated with the Set Covering Problem instance given in Example 18.

Lemma 7 *Consider an instance $(\mathscr{U}, S)$ of the set covering problem where $\mathscr{U} = \{a_1, \ldots, a_m\}$ is the universe and $S = \{S_1, \ldots, S_n\}$ is the collection of subsets of $\mathscr{U}$ such that $\bigcup_{i=1}^{n} S_i = \mathscr{U}$. Let $\mathscr{D}_{SCP}(\mathscr{U}, S)$ be its associated directed graph as defined in Definition 35 and let $I = \{i_1, \ldots, i_l\}$ be a set of indices with $I \subseteq \{1, \ldots, n\}$. Additionally, let $S_{\mathscr{O}} = \{s_{11}, \ldots, s_{n1}, S^*_{i_1}, \ldots, S^*_{i_l}\}$. Then, $\bigcup_{i \in I} S_i = \mathscr{U}$ if and only if $\mu_G(\bar{A}, I_n^{S_{\mathscr{O}}}) \leq k$, where $\bar{A}$ is the structural matrix associated with $\mathscr{D}_{SCP}$ and $k = \max_i |S_i|$.* ⋄

Proof Consider the structural directed graph representation $\mathscr{D}_1$ depicted in Figure 4.2. Notice that it suffices to place an output in x_1 to ensure that the generic observability index μ_G is less or equal than k. In fact, the structural pair $(\bar{A}_1, \bar{C}_1)$ associated with $\mathscr{D}_1$ is such that

$$\bar{A}_1 = \begin{bmatrix} \mathbf{0}_{k-1,1} & I_{k-1} \\ 0 & \mathbf{0}_{1,k-1} \end{bmatrix},$$
$$\bar{C}_1 = \begin{bmatrix} 1 & \mathbf{0}_{1,k-1} \end{bmatrix},$$

where $\mathbf{0}_{m,n}$ is the $m \times n$ matrix of zero entries and I_m is the identity matrix of size m. Therefore, it is possible to see that

$$grank\left(\begin{bmatrix} \bar{C}_1 \\ \bar{C}_1\bar{A}_1 \\ \vdots \\ \bar{C}_1\bar{A}_1^{k-1} \end{bmatrix}\right) = grank(I_k),$$

and that $\mu_G(\bar{A}_1, \bar{C}_1) = k$.

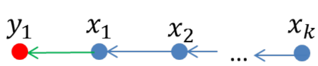


Figure 4.2: A structural directed graph representation $\mathscr{D}_1(\bar{A}, \bar{C})$ for which $\mu_G(\bar{A}, \bar{C}) = k$.

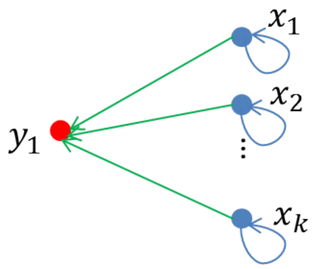


Figure 4.3: A structural directed graph representation $\mathscr{D}_2(\bar{A},\bar{C})$ for which $\mu_G(\bar{A},\bar{C}) = k$.

Consider now the structural directed graph representation $\mathscr{D}_2$ of Figure 4.3 and consider that $(\bar{A}_2,\bar{C}_2)$ is the associated structural pair such that

$$\bar{A}_2 = I_k,$$
$$\bar{C}_2 = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}.$$

Therefore, we have that

$$grank\left(\begin{bmatrix} \bar{C}_2 \\ \bar{C}_2\bar{A}_2 \\ \vdots \\ \bar{C}_2\bar{A}_2^{k-1} \end{bmatrix}\right) = grank(\mathbf{1}_{k,k}),$$

where $\mathbf{1}_{k,k}$ is the $k \times k$ matrix whose entries are all equal to 1, and $\mu_G(\bar{A}_2,\bar{C}_2) = k$.

It is possible to see that any dedicated output configuration that ensures structural observability of $\mathscr{D}_{SCP}$ has to be comprised of the variables $s_{1(k-1)}, s_{2(k-1)}, \dots, s_{n(k-1)}$ since those variables constitute non-bottom linked SCCs. Furthermore, the outputs placed in each of those variables are necessary and sufficient (see Figure 4.2) to ensure structural observability with generic index less than k of the directed graph whose vertex set is $\mathscr{V}_1 = S^* \cup S_k$ and whose arc set is $\mathscr{E}_1 = \mathscr{E}_{S^*,S_k}$.

By comparison with Figure 4.3, notice that to guarantee structural observability of $\mathscr{D}_{SCP}$ with μ_G less than k, one has to start to add outputs in the variables in S^*. Furthermore, it may be easily verified that $\bigcup_{i\in I} S_i = \mathscr{U}$ with $I = \{i_1,\dots,i_l\}$ if and only if the set of variables in S^* indexed by I is sufficient to ensure a generic observability index less than k of the directed graph with vertex set $\mathscr{V}_2 = \mathscr{U} \cup S^*$ and arc set $\mathscr{E}_2 = \mathscr{E}_{\mathscr{U},\mathscr{U}} \cup \mathscr{E}_{\mathscr{U},S^*}$. Since $\mathscr{D}_{SCP} = (\mathscr{V}_1 \cup \mathscr{V}_2, \mathscr{E}_1 \cup \mathscr{E}_2)$, the lemma is proved. □

Theorem 15 *The output selection problem with generic index constraint (problem (4.6)) is NP-complete.* ◇

Proof First, notice that problem (4.6) is in NP since for a structural pair (A,C) there exists a polynomial procedure to compute the generic observability index [22]. Therefore, it remains to show that there is a polynomial procedure that reduces the set covering problem to problem (4.6) and, by Lemma 6, the theorem is proved.

For that, consider the directed graph construction given in Definition 35 (see also Algorithm 7). First, we call a procedure that solves problem (4.6) for the structural state matrix $\bar{A}$ associated with $\mathscr{D}_{SCP}$ and let $S_{\mathscr{O}}$ be a solution. We will show that from $S_{\mathscr{O}}$ it is possible to construct a solution with the same number of outputs and with variables $s_{1(k-1)}, \ldots, s_{n(k-1)}$ and variables in S^*. If there is a $s_{ij} \in S_{\mathscr{O}}$ with $j \neq k-1$, then the set $S_{\mathscr{O}} \setminus \{s_{ij}\} \cup \{S_i^*\}$ is certainly a solution to problem (4.6). Further, for each $a_i \in S_{\mathscr{O}}$, there must exist an $S_j^* \notin S_{\mathscr{O}}$ with $(a_i, S_{\mathscr{O}}) \in \mathscr{E}$ and with $S_{\mathscr{O}} \setminus \{a_i\} \cup \{S_j^*\}$ a solution to problem (4.6). If this was not the case, then the number of outputs would not be minimal, which is a contradiction. Therefore, it is always possible to, given a solution $S_{\mathscr{O}}$, construct another solution $S'_{\mathscr{O}}$ with variables $s_{i(k-1)}$, $i = 1, \ldots, n$, and some variables from S^*. Then, with Lemma 7 in mind, the theorem is proved. □

For instance, notice that in Example 18 if we call the procedure to solve problem (4.6), $S_{\mathscr{O}} = \{s_{13}, s_{23}, s_{33}, S_1^*, a_4\}$ might be the returned value. Since (a_4, S_2) is an arc of the directed graph, $S'_{\mathscr{O}} = \{s_{13}, s_{23}, s_{33}, S_1^*, S_2^*\}$ is also a solution of problem (4.6) and, therefore, $I = \{1, 2\}$ is a solution of the Set Covering Problem.

ALGORITHM 7: Reduction of the Set Covering Problem to Problem 4.6

Input: An instance of problem 4.8: an universe $\mathscr{U} = \{a_1, \dots, a_m\}$ and collection of subsets $S = \{S_1, \dots, S_n\}$ of $\mathscr{U}$ whose union is $\mathscr{U}$.
Output: A set of indices $I \subseteq \{1, \dots, m\}$ that solves problem 4.8.
Let k be the size of the largest $S_i \in S$, with $i = 1, \dots, n$.
Build the directed graph $\mathscr{D}_{SCP} = (\mathscr{U} \cup S^* \cup S_k, \mathscr{E}_{\mathscr{U},\mathscr{U}} \cup \mathscr{E}_{\mathscr{U},S^*} \cup \mathscr{E}_{S^*,S_k})$, where

- $\mathscr{U} = \{a_1, \dots, a_m\}$
- $S^* = \{S_1^*, \dots, S_n^*\}$
- $S_k = \bigcup_{i=1}^{n} \bigcup_{j=1}^{k} s_{ij}$
- $\mathscr{E}_{\mathscr{U},\mathscr{U}} = \{(a_i, a_i) \mid \forall a_i \in \mathscr{U}\}$
- $\mathscr{E}_{\mathscr{U},S^*} = \{(a_i, S_j^*) \mid a_i \in S_j, \forall a_i \in \mathscr{U}, \forall \quad S_j \in S\}$
- $\mathscr{E}_{S^*,S_k} = \{(S_i^*, s_{i1}) \mid \forall \quad S_i \in S, \forall s_{i1} \in S_k\} \cup \{(s_{ij}, s_{i(j+1)} \mid i = 1, \dots, m, j = 1, \dots, k-2\}$

Call a procedure that solves problem 4.6 for $\mathscr{D}_{SCP}$ and k. Let $S_{\mathscr{O}}$ be the solution.
$I \leftarrow \emptyset$
for $s_{i(k-1)} \in S_{\mathscr{O}}$ **do**
  $S_{\mathscr{O}} \leftarrow S_{\mathscr{O}} \setminus \{s_{i(k-1)}\}$
end for
for $s_{ij} \in S_{\mathscr{O}}$ **do**
  $S_{\mathscr{O}} \leftarrow S_{\mathscr{O}} \setminus \{s_{ij}\}$
  $I \leftarrow I \cup \{i\}$
end for
for $S_i^* \in S_{\mathscr{O}}$ **do**
  $S_{\mathscr{O}} \leftarrow S_{\mathscr{O}} \setminus \{S_i^*\}$
  $I \leftarrow I \cup \{i\}$
end for
for $a_i \in S_{\mathscr{O}}$ **do**
  Find $j \notin I$ such that $(a_i, S_j^*) \in \mathscr{E}_{\mathscr{U},S^*}$
  $S_{\mathscr{O}} \leftarrow S_{\mathscr{O}} \setminus \{a_i\}$
  $I \leftarrow I \cup \{j\}$
end for
Return I

Chapter 5

Illustrative Examples

In this chapter, we start by illustrating the concepts regarding matroid theory with an example. After that, the algorithm developed to find a minimum-cost feasible dedicated output configuration is applied to a spatially distributed sensor network. Finally, some simulation results are presented.

5.1 A 6-node Networked Example

In this section, we apply the matroid intersection algorithm (Algorithm 3) in order to obtain a minimum dedicated output configuration that guarantees structural observability. Consider, for instance, the structural state matrix

$$\bar{A} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \tag{5.1}$$

that represents the zero/non-zero pattern of some real-valued matrix $A \in \mathbb{R}^{n\times n}$, with $n = 6$. In Figure (5.1), we provide the structural directed graph representation $\mathscr{D}(\bar{A}) = (\mathscr{X}, \mathscr{E}_{\mathscr{X},\mathscr{X}})$.

When applying the matroid intersection algorithm, we have to consider the output-connected matroid $M_1 = (\mathscr{X}, \mathscr{O}_C)$ and the contraction-free matroid $M_2 = (\mathscr{X}, \mathscr{C}_F)$. Recall that a set $I \subseteq \mathscr{X}$ belongs to $\mathscr{O}_C$ if and only if $I \xrightarrow{CC} \mathscr{X} \setminus I$, i.e, if for any $x_i \in I$, there exists a directed path from x_i to some $x_j \in \mathscr{X} \setminus I$. On the other hand, $I \in \mathscr{C}_F$, if and only if there is a matching in the bipartite graph $\mathscr{B}(\bar{A}) = (\mathscr{X}, \mathscr{X}^+, E_{\mathscr{X},\mathscr{X}^+})$ that covers I.

Let us take, for instance, the set $X_A = \{x_1, x_2, x_3, x_4\}$ of the right-covered vertices of Figure 5.2, where we can see the bipartite graph associated with $\bar{A}$ and a maximum matching. By definition, $X_A \in \mathscr{C}_F$. If X_A also belongs to the collection of independent sets of the output-connected matroid, then surely X_A is a maximum-cardinality set in $\mathscr{O}_C \cap \mathscr{C}_F$ since X_A is covered by a maximum matching. To test that possibility, we will rely on the independence oracle for $M_1 = (\mathscr{X}, \mathscr{O}_C)$ (Algorithm

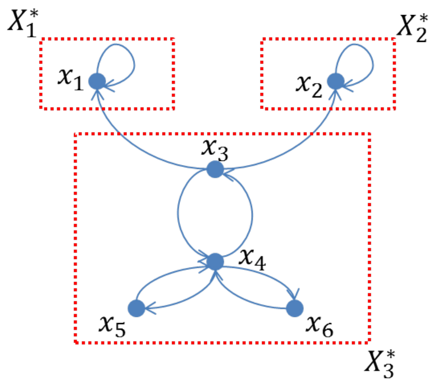


Figure 5.1: The structural directed graph representation of matrix $\bar{A}$ described in (5.1). The strong connected components are represented by red dashed lines.

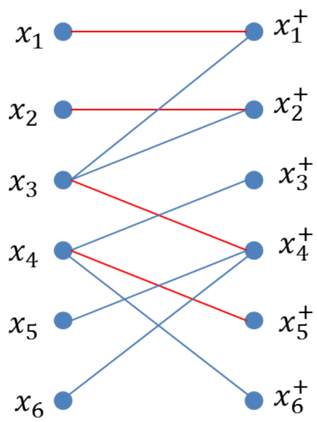


Figure 5.2: The bipartite graph representation of matrix $\bar{A}$ from equation (5.1). A maximal matching is depicted with red-colored edges.

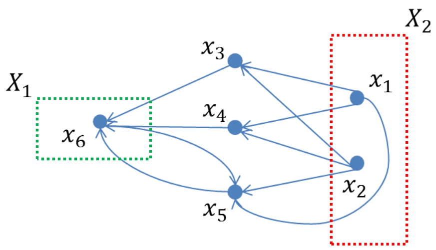


Figure 5.3: The bipartite exchange graph for $I = \{x_1, x_2, x_6\}$.

4). The output-connected sets $\mathscr{O}_j^K$ for the variables in $\mathscr{X} \setminus X_A = \{x_5, x_6\}$ are initialized as

$$\mathscr{O}_5^1 = \mathscr{O}_6^1 = \{x_4\} \ . \tag{5.2}$$

Since

$$\mathscr{O}_j^l = \{x_k \in \mathscr{X} \mid \bar{A}_{k,r} = 1 \text{ with } x_r \in \mathscr{O}_j^{l-1}\} \cup \mathscr{O}_j^{l-1}, \tag{5.3}$$

we have that

$$\mathscr{O}_5^{n-1} = \mathscr{O}_6^{n-1} = \{x_3, x_4\}, \tag{5.4}$$

and therefore $\cup_{j=5,6} \mathscr{O}_j^{n-1} \neq X_A$, which implies that $X_A \notin \mathscr{O}_C$. The same conclusion can be drawn from Proposition 8. Notice, that in the condensation of $\mathscr{D}(\bar{A})$, depicted in Figure (5.1), $\mathscr{X}_1^* = \{x_1\}$ and $\mathscr{X}_2^* = \{x_2\}$ are two non-bottom linked strong connected components. Therefore, x_1 and x_2 do not belong to any independent set of the output-connected matroid. If we remove x_1 and x_2 from X_A, we have $X_B = \{x_3, x_4\}$ that now lies on the intersection of $\mathscr{O}_C$ and $\mathscr{C}_F$. Then, $\mathscr{X} \setminus \mathscr{X}_B$ is a feasible dedicated output configuration.

In order to obtain a minimum feasible dedicated output configuration, one may apply Algorithm 3 to $M_1 = (\mathscr{X}, \mathscr{O}_C)$ and $M_2 = (\mathscr{X}, \mathscr{C}_F)$ with $I = X_B$. Notice that the set of the sources $X_1 = \{x_i \in \mathscr{X} \setminus I \mid I \cup \{x_i\} \in \mathscr{O}_C\}$ is

$$X_1 = \{x_5, x_6\}, \tag{5.5}$$

since x_3, x_4, x_5 and x_6 are all in the same strong connected component. On the other hand, the set of sinks $X_2 = \{x_i \in \mathscr{X} \setminus I \mid I \cup \{x_i\} \in \mathscr{C}_F\}$ is

$$X_2 = \{x_1, x_2, x_5, x_6\}, \tag{5.6}$$

since all those variables (one at a time) can be added to I and we still have a matching in $\mathscr{B}(\bar{A})$ that covers the extended I. Since x_5 and x_6 are simultaneously sources and sinks, the dipaths $P = x_5$ or $P = x_6$ are possible shortest source-sink dipaths. Therefore, $I \cup \{x_5\}$ and $I \cup \{x_6\}$ belong to $\mathscr{O}_C \cap \mathscr{C}_F$ and we can take, for instance, $I = \{x_3, x_4, x_5\}$.

In the second iteration of Algorithm 3, the set of sources is given by

$$X_1 = \{x_6\}, \tag{5.7}$$

whereas the set of sinks is

$$X_2 = \{x_1, x_2\}, \tag{5.8}$$

since x_5 and x_6 cannot belong simultaneously to a set of vertices covered by a matching in $\mathscr{B}(\bar{A})$. The bipartite exchange digraph $\mathscr{D}_{M_1,M_2}(I)$ is represented in Figure 5.3. Since there is not any directed path between x_6 and x_1 or x_2, we conclude that $I = \{x_3, x_4, x_5\}$ is a maximum-cardinality set in $\mathscr{O}_C \cap \mathscr{C}_F$. Therefore, $S_\mathscr{O}$ is a minimum feasible dedicated output configuration where $S_\mathscr{O}$ is

$$S_\mathscr{O} = \mathscr{X} \setminus I = \{x_1, x_2, x_6\}. \tag{5.9}$$

Then, we may write $\bar{C} = I_n^{S_\mathscr{O}}$, where $I_n^{S_\mathscr{O}}$ is the identity matrix of size n but only with rows 1, 2 and 6 and we have that

$$\bar{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.10}$$

which implies that the pair $(\bar{A}, \bar{C})$ is structurally observable. Finally, we can choose randomly the real entries of matrices $A \in [\bar{A}]$ and $C \in [\bar{C}]$. Since structural observability is a *generic* property, there is a high probability that the pair (A, C), with numerical values randomly chosen, is observable. Consider that the non-zero values of matrices A and C are chosen randomly from the uniform distribution on the interval $[0, 1]$. Then, we may have

$$A = \begin{bmatrix} 0.5470 & 0 & 0.4868 & 0 & 0 & 0 \\ 0 & 0.6256 & 0.4359 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.6443 & 0 & 0 \\ 0 & 0 & 0.3063 & 0 & 0.5502 & 0.2305 \\ 0 & 0 & 0 & 0.8116 & 0 & 0 \\ 0 & 0 & 0 & 0.5328 & 0 & 0 \end{bmatrix} \tag{5.11}$$

and

$$C = \begin{bmatrix} 4.302 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.4389 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5079 \end{bmatrix}, \tag{5.12}$$

and it can be seen that the pair (A, C) is observable.

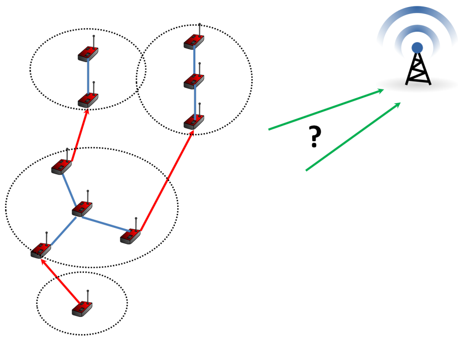


Figure 5.4: A set of spatially distributed wireless sensors. Each sensor belongs to a local area (represented by a dashed black circle). Some sensors can communicate (bidirectional) with each other in the same local area (blue lines). A node can transmit information to different local areas (red arrows). The objective is to choose the set of sensors allowed to change information with the central authority in order to recover the vector of initial measurements.

5.2 Wireless Sensor Application

Consider, as an application, a set of spatially distributed wireless sensors (Figure 5.4). The sensors are divided between local areas and they are equipped with a short-range communication device that allow them to communicate with some other sensors in the same local area in order to update data. Additionally, some sensors can also transmit information to other sensors in different local areas. The objective is to recover the vector of initial sensor measurements at some central authority. To that purpose, a subset of sensors will be equipped with a long-range communication device. However, since the communication cost increases with distance and since the central authority is far apart, it is desirable to install the long-range communication device in the fewest possible sensors.

Each sensor node updates its data according to linear dynamics and maintains a state $x_i(t)$, where t is the iteration instant. The state $x_i(0)$ corresponds to the field measurement collected by the i-th sensor. The sensor communication is described by a dynamic system of the form

$$x(t+1) = Ax(t), \tag{5.13}$$

where $A \in \mathbb{R}^{n\times n}$ describes the state update dynamics. Matrix A has to respect the communication graph structure, i.e, if sensor j is not allowed to transmit information to sensor i, then A_{ij} must be zero. By other words, the structural pattern $\bar{A} \in \{0,1\}^{n\times n}$ of matrix A is given, in the sense that if $\bar{A}_{ij} = 0$, then it must be that $A_{ij} = 0$. For the sensor network depicted in Figure 5.4, the structural directed graph representation $\mathscr{D}(\bar{A})$ is presented in Figure 5.5. Notice that if sensor i can transmit information to sensor j and if they belong to same local are, then it is fair enough to assume that j can also transmit information to i. Furthermore, each sensor has memory which means that $\bar{A}_{ii} = 1$. The cost of place the long-range communication device at sensor i is given by: $\mathscr{C}(x_1) = 3$, $\mathscr{C}(x_2) = 10$, $\mathscr{C}(x_3) = 3$, $\mathscr{C}(x_4) = 15$, $\mathscr{C}(x_5) = 2$, $\mathscr{C}(x_6) = 5$, $\mathscr{C}(x_7) = 10$, $\mathscr{C}(x_8) = 2$, $\mathscr{C}(x_9) = 5$ and $\mathscr{C}(x_{10}) = 1$.

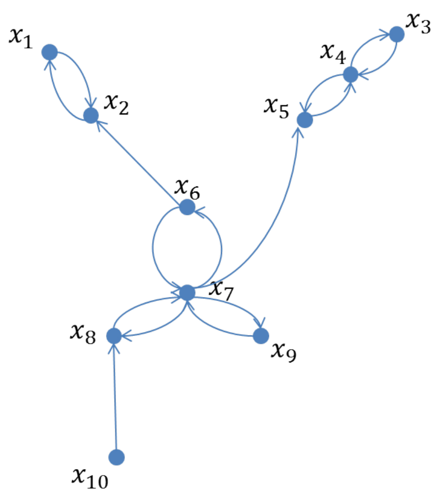


Figure 5.5: A structural directed graph representation of the network depicted in Figure 5.4Although not depicted, to ease the illustration, each node has a self-loop.

Notice that the present goal is to ensure that the central identity can use the dynamics and the collection of measurements to recover the vector of initial sensor states $x(0)$. Therefore, observability has to hold. If additionally we require that the cost of the dedicated output configuration be as less as possible, then we can apply Algorithm 6 to obtain the correct subset of sensors $S_{\mathscr{O}}$ in which to place the long-range communication device.

STEP 1 In this step, the condensation $\mathscr{D}^*$ of the structural directed graph representation is computed using the Tarjan algorithm [33] (Figure 5.6). Notice that there are four strong connected components, respectively $X_1 = \{x_1, x_2\}$, $X_2 = \{x_3, x_4, x_5\}$, $X_3 = \{x_6, x_7, x_8, x_9\}$ and $X_4 = \{x_{10}\}$.

STEP 2 From the digraph $\mathscr{D}^* = (\mathscr{X}^*, \mathscr{E}_{\mathscr{X}^*,\mathscr{X}^*})$ represented in Figure 5.6, it can be seen that X_1 and X_2 are the only SCCs for which there are no edged going out. Therefore, X_1 and X_2 are non-bottom linked SCCs. Then, in the solution, there must exist at least one state variable from $X_1 = \{x_1, x_2\}$ and one state variable $X_2 = \{x_3, x_4, x_5\}$.

STEP 3 In this step, the weighted bipartite graph representation is constructed and represented in Figure 5.7. Since there are two non-bottom linked SCCs, two dummy variables z_1^* and z_2^* are added to the bipartite graph $\mathscr{B}(\bar{A})$. Variable z_1^* is linked to x_1 and x_2 and z_2^* is linked to x_3, x_4 and x_5. The weigh of those edges is $\sum_{i=1}^{10} \mathscr{C}(x_i) + 1 = 57$. Each edge (x_i, x_j^+) has weight $\mathscr{C}(x_i)$.

STEP 4 Finally, a weighted-maximum matching M_w is computed using the Hungarian algorithm [**?**]. In Figure 5.7, the dashed edges belong to M_w. Set $\mathscr{X}_M \subseteq \mathscr{X}$ to the state variables covered by M_w. Then $\mathscr{X}_M = \mathscr{X}$. Next, we have to remove those verices whose edges are incident with z_1^* or z_2^*. Since (x_1, z_1^*) and (x_5, z_2^*), then $\mathscr{X}_M$ is update to $\mathscr{X}_M' = \mathscr{X}_M \setminus \{x_1, x_5\}$. Thus, a feasible dedicated output configuration of minimum cost is $S_{\mathscr{O}} = \mathscr{X} \setminus \mathscr{X}_M' = \{x_1, x_5\}$. The cost of this configuration is $\mathscr{C}(x_1) + \mathscr{C}(x_5) = 3 + 2 = 5$.

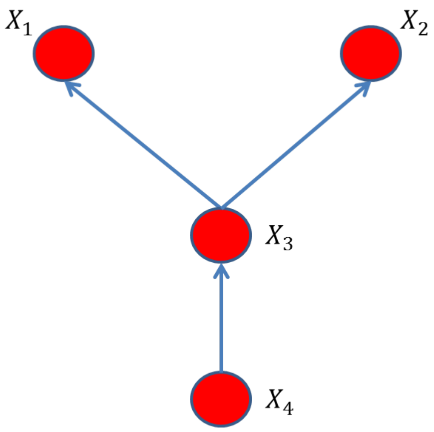


Figure 5.6: Condensation of the structural directed graph representation of Figure 5.5. The strong connected components are $X_1 = \{x_1, x_2\}$, $X_2 = \{x_3, x_4, x_5\}$, $X_3 = \{x_3, x_6, x_7, x_8, x_9\}$ and $X_4 = \{x_{10}\}$.

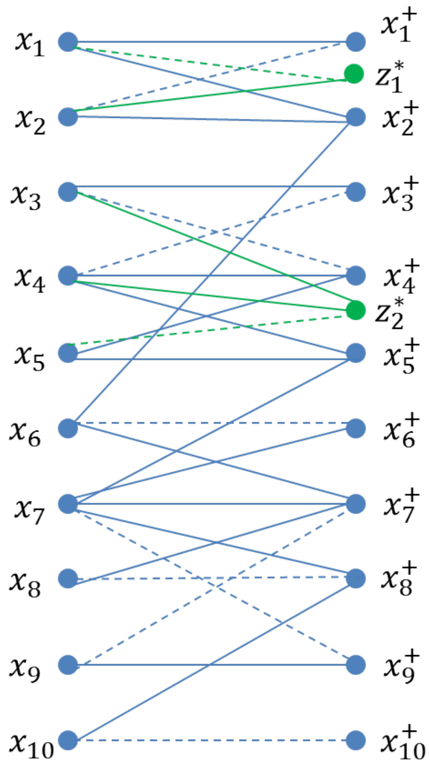


Figure 5.7: The bipartite graph construction of STEP 3. The edges in green correspond to the connections between state variables and the associated non-bottom linked SCCs. The dashed edges belong to a weighted-maximal matching.

5.3 Simulation Results of Random Networks

In this section, we provide results of a set of MATLAB simulations in order to conclude about the following

Q1: *How does the presence/absence of zeroes in the structural pattern of some state matrix affect the minimum number of dedicated outputs to obtain structural observability? Further, how this conclusions change with the dimension of the system?*

Q2: *Is the presence of self-loops an important requirement to reduce the number of required dedicated outputs to obtain structural observability?*

To that purpose, the following experiment was simulated. To each entry of $\bar{A} \in \{0,1\}^{n\times n}$ the value 1 was assigned with probability p, i.e., for each $i, j = 1,\ldots,n$

$$\bar{A}_{ij} = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1-p. \end{cases}$$

Additionally, for each $\bar{A}$ constructed according to the previously manner, it was built two more structural matrices: one, with *forced self-loops*, i.e, with $\bar{A}_{ii} = 1$ for $i = 1,\ldots,n$; and another with zero self-loops, i.e, with $\bar{A}_{ij} = 0$ for $i = 1,\ldots,n$.

For each value of p, with $p \in \{0.001, 0.01, 0.1, 0, 2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, the minimum number of dedicated outputs needed to ensure structural observability was computed for $n = 1,\ldots,100$. The simulation was realized 100 times.

For small values of p (Figure 5.8 and Figure 5.9) the minimum number of dedicated outputs, $n_{\mathscr{O}}$, increases with the size of the structural state matrix, n. For p between 0.1 and 0.8 (Figures 5.10, 5.11, 5.12, 5.13, 5.14, 5.15 and 5.16) there is a peak that corresponds to the maximum value of $n_{\mathscr{O}}$ that occurs for values of n successively small until $p = 0.8, 0.9$ (Figures 5.17 and 5.18) for which that peak value is negligible. There is an abrupt decrease in $n_{\mathscr{O}}$ between $p = 0.001$ and $p = 0.01$ and between $p = 0.01$ and $p = 0.1$, and after that $n_{\mathscr{O}}$ is always less than 6. Notice that for $p \geq 0.2$ and $n \geq 10$, the minimum number of dedicated outputs needed to obtain structural observability is approximately 1, regardless of the size of $\bar{A}$. Furthermore, it can be concluded that the presence of self-loops is significant only for $p = 0.01$ (Figure 5.9), where $n_{\mathscr{O}}$ takes as maximum approximately 45, whereas for the case with forced self-loops the maximum value of $n_{\mathscr{O}}$ is approximately 35. For values of p inferior and superior than 0.01, the blue, red, and green curves almost overlap.

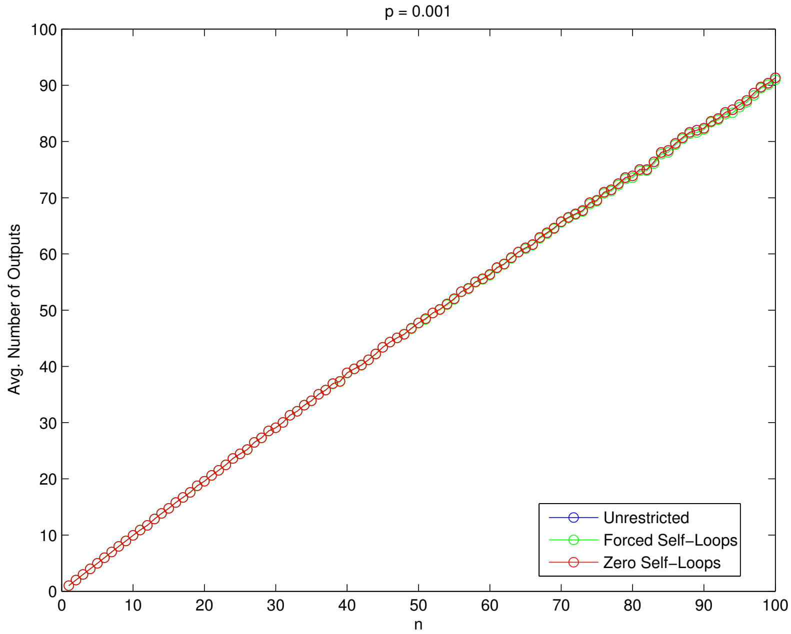


Figure 5.8: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.001$.

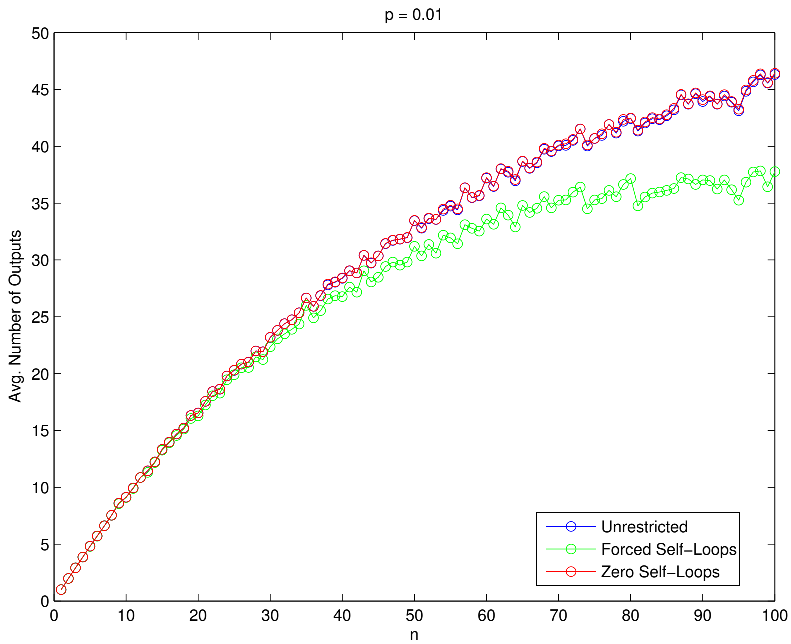


Figure 5.9: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.01$.

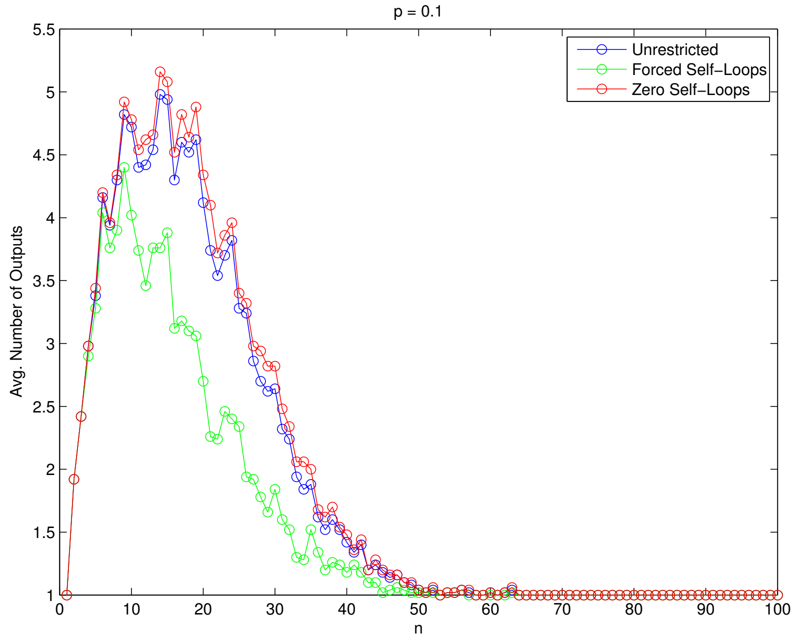


Figure 5.10: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.1$.

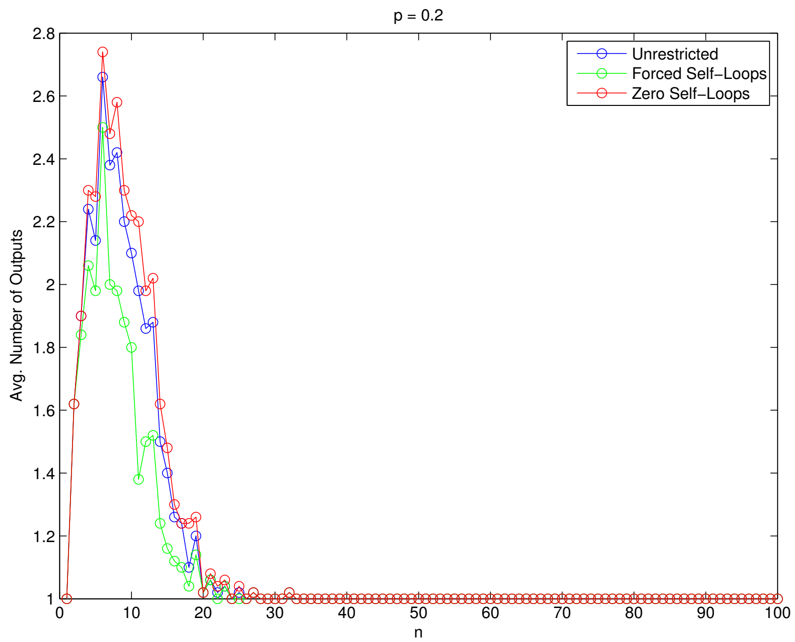


Figure 5.11: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.2$.

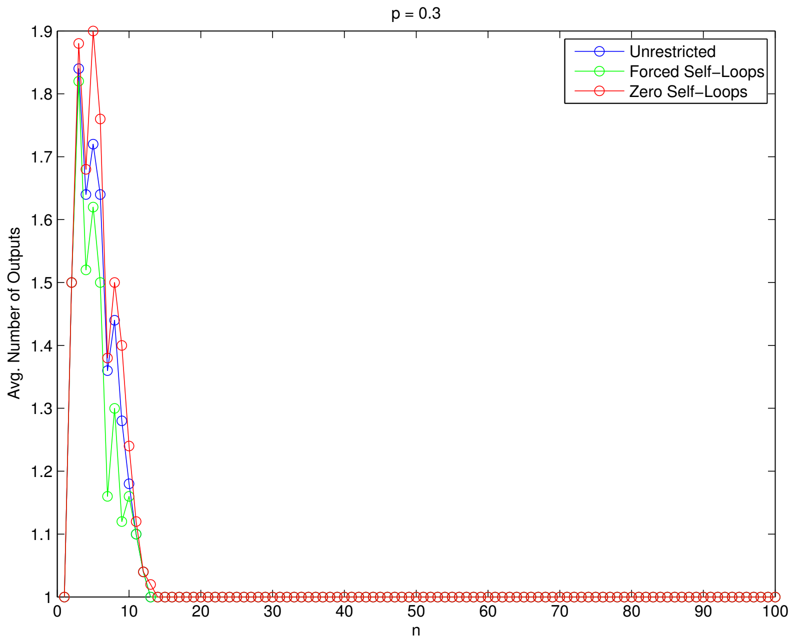


Figure 5.12: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.3$.

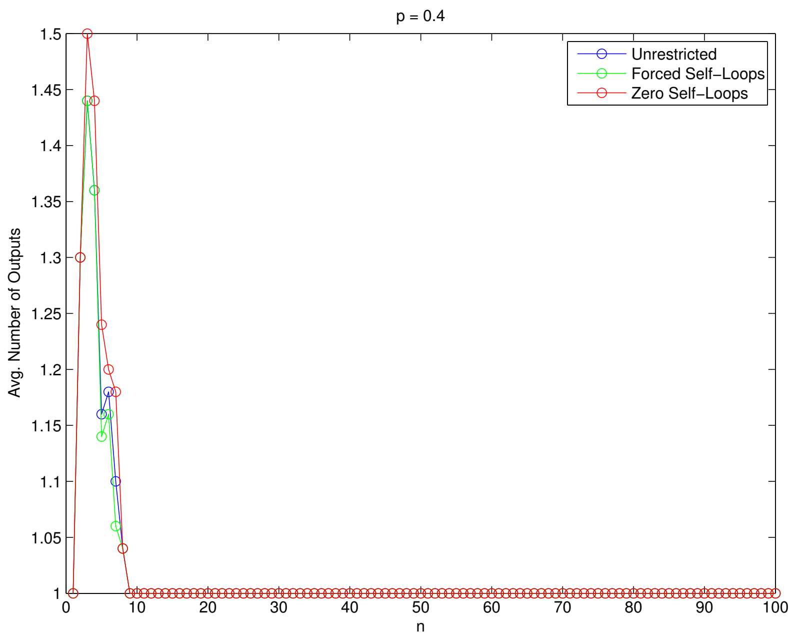


Figure 5.13: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.4$.

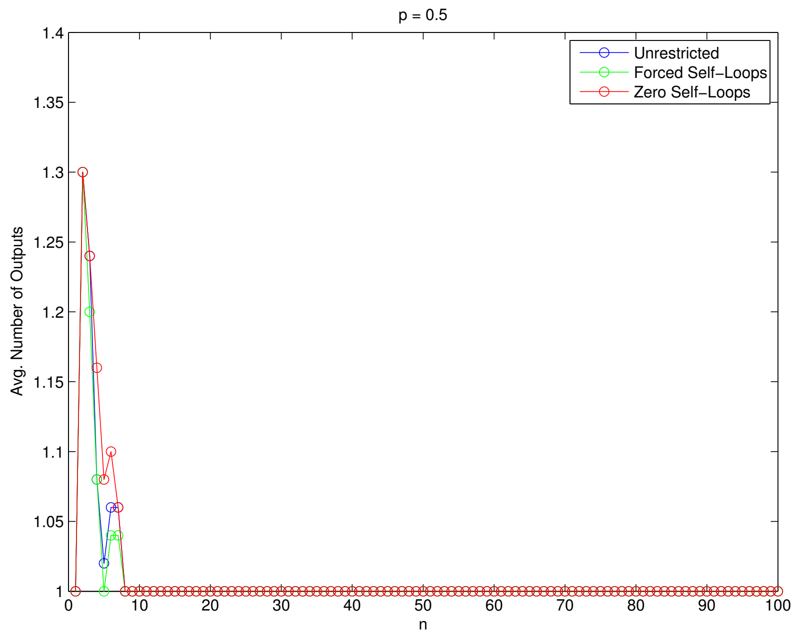


Figure 5.14: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.5$.

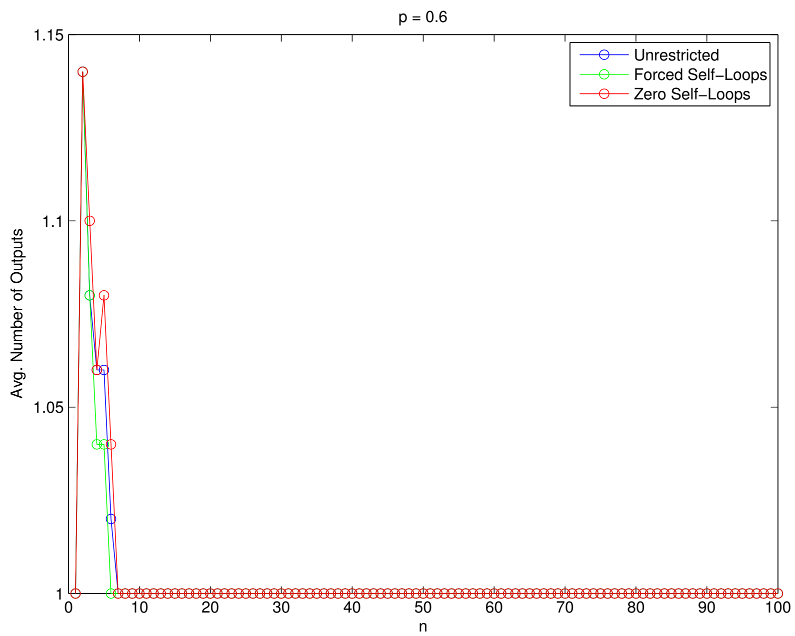


Figure 5.15: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.6$.

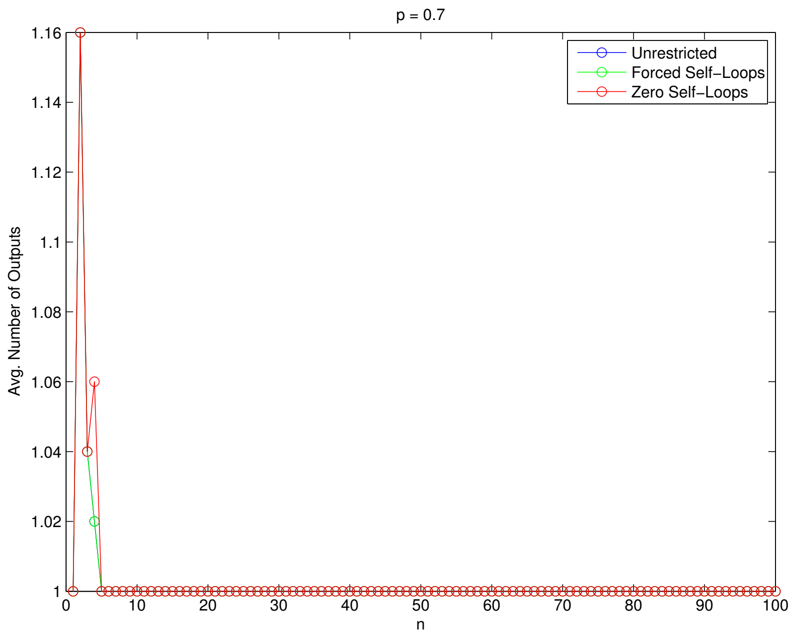


Figure 5.16: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.7$.

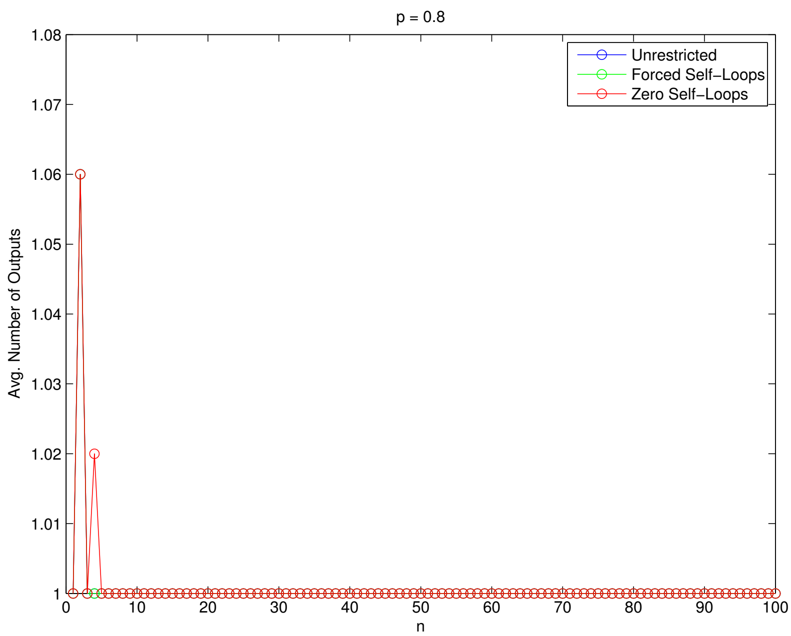


Figure 5.17: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.8$.

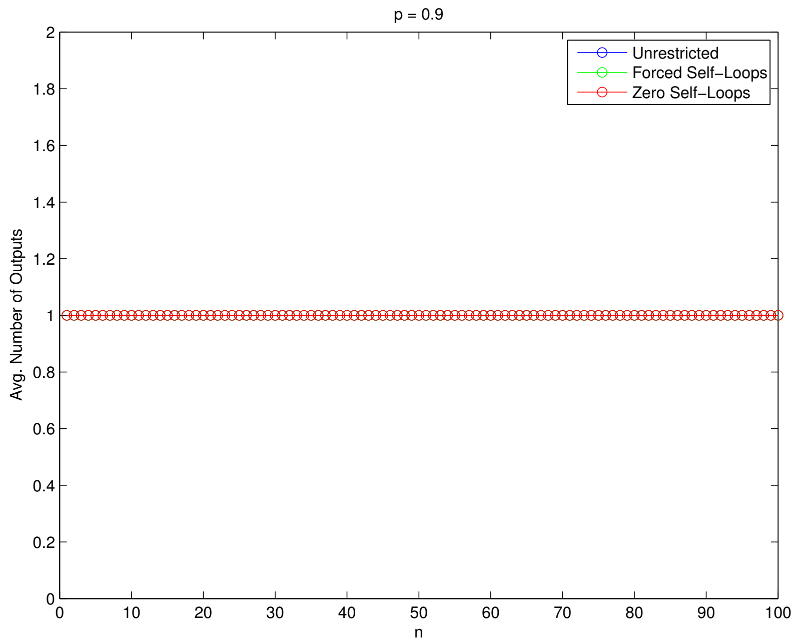


Figure 5.18: The average number of minimum dedicated outputs needed to ensure structural observability for $p = 0.9$.

Chapter 6

Conclusions

Achievements

In this thesis, we reformulated the minimum feasible dedicated output problem (3.4) within a matroid framework. We have shown that it was possible to reduce that problem to an intersection of two matroids. Next, after established the connection of the obtained results to graph-theoretic concepts, we provided a new algorithm to find a minimum feasible dedicated output configuration. These results were easily extended to the scenario where costs associated with the output allocation were considered. We have also proved that the output placement problem with generic observability index constraint (4.6) was NP-complete by a polynomial reduction of the well-known set covering problem to the first. By duality, the derived observability results can be readily extended to the structural controllability and the corresponding input (actuator) design.

Future Work

As part of future research, we believe that the matroid framework should be deeply studied in order to provide new results within structural systems theory. Another important future work is to try to reduce problem (4.6) to a well-known NP-complete problem that have good approximation algorithms. Thus, if the reduction is successful it would be possible to adapt such approximation algorithms to solve the original problem. Since problem (3.4) was formulated as the intersection of two matroids, perhaps problem (4.6) can be rewritten as an intersection of three matroids. It is known that the matroid intersection problem becomes NP-hard when three matroids are involved, instead of only two.

Appendix A

MATLAB implementation of some algorithms

A.1 Maximum-Cardinality Matroid Intersection Algorithm

```
function [ I ] = maxIntersectMatroid(@oracle1, @oracle2, n);
%Inputs:
% @oracle1 - Independence oracle of the first matroid
% @oracle2 - Independece oracle of the second matroid
% n - size of the univeser {1, ..., n}
%Outputs:
% I - A maximum-size independent set that lies on the instersection of the
%     two matroids given by independence oracles @oracle1 and @oracle2

I = zeros(1, n);

while(1)
    %Construct the sets X_1 and X_2
    X_1 = zeros(1, n);
    X_2 = zeros(1, n);
    A = zeros(n, n);

    for i = 1 : n
        aux = zeros(1, n);
        aux(i) = 1;
        if(I(i) == 0)
            X_1(i) = oracle1(I + aux, n);
```

```
            X_2(i) = oracle2(I + aux, n);
        else
            for j = 1 : n
                aux2 = zeros(1,n);
                aux2(j) = 1;
                if(I(j) == 0)
                    A(i,j) = oracle1(I−aux+aux2, n);
                    A(j,i) = oracle2(I−aux+aux2, n);
                end
            end
        end
    end

    B = zeros( n + 2, n + 2);
    B(1,2:n+1) = X_1;
    B(2:n+1, n+2) = X_2';
    B(2:n+1, 2:n+1) = A;

    [d pred] = shortest_paths(sparse(B),1);

    if(d(n+2) == inf)
        return I;
    end

        i = pred(n + 2);
        aux = 1;
        while(1)
            if(i == 1)
                break;
            end
            if(aux == 1)
                I(pred(i)) = 1;
            else
                I(pred(i)) = 0;
            end

            aux = aux * −1;
            i = pred(i);
        end
end
```

```

end
```

A.2 Minimum-Size Output Selection

```
function [ C, p ] = minimumOutput(A)

%Inputs:
%A − Structural pattern of some state matrix
%Outputs:
%p − Number of outputs of the minimum feasible dedicate output
%      configuration associated with A
%C − Structural pattern of the output matrix associated with a minimum
%      feasible dedicate output configuration for A

n = length(A);
%STEP 1 − COMPUTE THE CONDENSATION OF D(A)
SCC = scomponents(A);
ncomp = max(SCC);
R = sparse(1:size(A,1),SCC,1,size(A,1),ncomp);
CG = R'*A*R;
CG = (CG ~= 0);
CG = CG & (~eye(ncomp, ncomp));

%STEP 2 − FIND THE NON−BOTTOM LINKED SCC'S
SCCnBottom = zeros(1, ncomp);
k = 1 : ncomp;
SCCnBottom = sum(CG(:,k)) == 0;

%STEP 3 − BUILD THE WEIGHTED BIPARTITE GRAPH
nBottom = sum(SCCnBottom);
A2 = zeros(n + nBottom, n + nBottom);
A2(1:n, 1:n) = A;
r = 1;
for k = 1 : ncomp
    if(SCCnBottom(k) == 1)
        A2(n + r, find((SCC == k), n)) = (n+1);
        r = r + 1;
    end
end
```

```

%STEP 4 - FIND A MAXIMUM-WEIGHT MATCHING
[val mi mj] = bipartite_matching(A2);
mj = mj(mi <= n);
X_M = zeros(1,n);
X_M(mj) = 1;
S_u = ones(1,n) - X_M(1:n);
p = sum(S_u);
C = zeros(p, n);
S_u = find(S_u, n);
for k = 1: p
    C(k, S_u(k)) = 1;
end

end
```

A.3 Minimum-Cost Output Selection

```
function [ C, p ] = minimumOutputCost(A, cost )
%Inputs:
%A - Structural pattern of some state matrix
%Outputs:
%p - Number of outputs of the minimum-cost feasible dedicate output
%     configuration associated with A
%C - Structural pattern of the output matrix associated with a minimum-cost
%     feasible dedicate output configuration for A

n = length(A);
%STEP 1 - COMPUTE THE CONDENSATION OF D(A)
SCC = scomponents(A);
ncomp = max(SCC);
R = sparse(1:size(A,1),SCC,1,size(A,1),ncomp);
CG = R'*A*R;
CG = (CG ~= 0);
CG = CG & (~eye(ncomp, ncomp));

%STEP 2 - FIND THE NON-BOTTOM LINKED SCC'S
SCCnBottom = zeros(1, ncomp);
k = 1 : ncomp;
```

```
SCCnBottom = sum(CG(:,k)) == 0;

%STEP 3 - BUILD THE WEIGHTED BIPARTITE GRAPH
nBottom = sum(SCCnBottom);
A2 = zeros(n + nBottom, n + nBottom);
totalCost = sum(cost);
for k = 1 : n
  A(find(A(:,k), n),k) = cost(k);
end
A2(1:n, 1:n) = A;
r = 1;
for k = 1 : ncomp
    if(SCCnBottom(k) == 1)
        A2(n + r, find((SCC == k), n)) = totalCost + 1;
        r = r + 1;
    end
end


%STEP 4 - FIND A MAXIMUM-WEIGHT MATCHING
[val mi mj] = bipartite_matching(A2);
mj = mj(mi <= n);
X_M = zeros(1,n);
X_M(mj) = 1;
S_u = ones(1,n) - X_M(1:n);
p = sum(S_u);
C = zeros(p, n);
S_u = find(S_u, n);
for k = 1: p
    C(k, S_u(k)) = 1;
end



end
```

References

[1] Indika Rajapakse, Mark Groudine, and Mehran Mesbahi. What can systems theory of networks offer to biology? *PLoS computational biology*, 8(6):e1002543, 2012.

[2] Bo Chen and Harry H Cheng. A review of the applications of agent technology in traffic and transportation systems. *Intelligent Transportation Systems, IEEE Transactions on*, 11(2):485–497, 2010.

[3] Gregory J Pottie and William J Kaiser. Wireless integrated network sensors. *Communications of the ACM*, 43(5):51–58, 2000.

[4] Chris Paige. Properties of numerical algorithms related to computing controllability. *Automatic Control, IEEE Transactions on*, 26(1):130–138, 1981.

[5] Alex Olshevsky. Minimal controllability problems. *arXiv preprint arXiv:1304.3071*, 2013.

[6] Ching-Tai Lin. Structural controllability. *Automatic Control, IEEE Transactions on*, 19(3):201–208, 1974.

[7] Robert Shields and J. Pearson. Structural controllability of multiinput linear systems. *Automatic Control, IEEE Transactions on*, 21(2):203–212, 1976.

[8] K. Glover and L. Silverman. Characterization of structural controllability. *Automatic Control, IEEE Transactions on*, 21(4):534–537, 1976.

[9] Kurt Johannes Reinschke. Multivariable control: a graph theoretic approach. 1988.

[10] Usman A Khan and Ali Jadbabaie. Coordinated networked estimation strategies using structured systems theory. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 2112–2117. IEEE, 2011.

[11] Usman A Khan and Mohammadreza Doostmohammadian. A sensor placement and network design paradigm for future smart grids. In *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on*, pages 137–140. IEEE, 2011.

[12] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.

[13] Sergio Pequito, Soummya Kar, and A Pedro Aguiar. A structured systems approach for optimal actuator-sensor placement in linear time-invariant systems. In *American Control Conference (ACC), 2013*, pages 6108–6113. IEEE, 2013.

[14] Sérgio Pequito, Soummya Kar, and A Pedro Aguiar. Minimum cost input-output and control configuration selection: A structural systems approach. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 4895–4900. IEEE, 2013.

[15] Hassler Whitney. On the abstract properties of linear dependence. *American Journal of Mathematics*, pages 509–533, 1935.

[16] Jack Edmonds. Matroids and the greedy algorithm. *Mathematical programming*, 1(1):127–136, 1971.

[17] Kazuo Murota. *Matrices and matroids for systems analysis*, volume 20. Springer, 2000.

[18] Andrew Clark, Linda Bushnell, and Radha Poovendran. On leader selection for performance and controllability in multi-agent systems. In *CDC*, pages 86–93, 2012.

[19] Jean-Michel Dion, Christian Commault, and Jacob Van Der Woude. Generic properties and control of linear structured systems: a survey. *Automatica*, 39(7):1125–1144, 2003.

[20] Airlie Chapman. *Semi-Autonomous Networks: Effective Control of Networked Systems through Protocols, Design, and Modeling*. PhD thesis, 2013.

[21] H Mortazavian. On k-controllability and k-observability of linear systems. In *Analysis and Optimization of Systems*, pages 600–612. Springer, 1982.

[22] C Sueur and G Dauphin-Tanguy. Controllability indices for structured systems. *Linear algebra and its applications*, 250:275–287, 1997.

[23] Shreyas Sundaram and Christoforos N Hadjicostis. Structural controllability and observability of linear systems over finite fields with applications to multi-agent systems. *Automatic Control, IEEE Transactions on*, 58(1):60–73, 2013.

[24] Chi-Tsong Chen. *Linear system theory and design*. Oxford University Press, Inc., 1998.

[25] João P. Hespanha. *Linear systems theory*. Princeton university press, 2009.

[26] John E. Hopcroft and Richard M. Karp. An n^5/2 algorithm for maximum matchings in bipartite graphs. *SIAM Journal on computing*, 2(4):225–231, 1973.

[27] Shigeyuki Hosoe. Determination of generic dimensions of controllable subspaces and its application. *Automatic Control, IEEE Transactions on*, 25(6):1192–1196, 1980.

[28] Jack Edmonds. Matroid intersection. *Annals of discrete Mathematics*, 4:39–49, 1979.

[29] James G Oxley. *Matroid theory*, volume 3. Oxford university press, 2006.

[30] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer, 2003.

[31] Jon Lee. *A first course in combinatorial optimization*, volume 36. Cambridge University Press, 2004.

[32] András Recski. *Matroid theory and its applications in electric network theory and in statics*. Springer, 1989.

[33] Robert Tarjan. Depth-first search and linear graph algorithms. *SIAM journal on computing*, 1(2):146–160, 1972.

[34] Michael R Garey and David S Johnson. *Computers and intractability*, volume 29. wh freeman, 2002.

[35] Harold W Kuhn. The hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97, 1955.