MASTER

On the existence of identifiable reparametrizations for linear compartment models

Baaijens, J.A.

Award date:
2014

Technische Universiteit
Eindhoven
University of Technology

Department of Mathematics and Computer Science

On the existence of identifiable reparametrizations for linear compartment models

Author:
Jasmijn A. Baaijens

Supervisor:
Dr. Jan Draisma

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science

Eindhoven, August 2014

Abstract

Linear compartment models describe the transport of a given material through a set of compartments and can be represented by a directed graph. The parameters of a linear compartment model correspond to the transport rates between different compartments. These parameters are usually estimated from experimental input-output data, by solving the parameter estimation problem. A fundamental prerequisite for parameter identification is identifiability: this concerns the question whether the experimental data correspond to a unique set of parameter values. In this thesis we consider generic local identifiability, meaning that a sufficiently general set of parameter values is at least locally a unique solution to the parameter estimation problem.

A problem that arises is what do to when a model is unidentifiable. A common approach is to search for an identifiable reparametrization of the model: a map which reduces the number of parameters, such that the reduced model is identifiable. Of particular interest in biological applications are rational scaling reparametrizations, which correspond to a rational scaling of the state variables.

We study a specific class of models, which are known to be unidentifiable. Earlier studies presented a criterion to decide whether a model has an identifiable scaling reparametrization, based on the dimension of the input-output map. Using algebraic geometry and graph theory, we translate this criterion to a criterion based on the rank of a bi-adjacency matrix. This new criterion can be evaluated by a randomized algorithm with asymptotic complexity of $O(n^6)$ operations. Furthermore, we present several new constructions to obtain a graph with an identifiable scaling reparametrization. Using these constructions, a large subclass of graphs which have an identifiable scaling reparametrization is obtained. This leads to a procedure of subdividing or deleting edges to ensure that a model has an identifiable scaling reparametrization.

Preface

Six months ago, I started with this graduation project, not knowing what to expect or where it would lead me. During these past months, there were moments of joy as well as moments of desperation, but overall it was a great experience. This thesis presents the results of my work, I hope you'll enjoy reading it.

I would like to use this opportunity to thank my supervisor, Jan Draisma, for finding me such a nice subject and guiding me through the entire process. I have learned so much during the writing of this thesis, you were a great inspiration. Thank you for all your time and enthusiasm during this project.

I would also like to express my gratitude towards Aart Blokhuis, Alexander Schönhuth and Georg Prokert for being part of my graduation committee. Thank you for your interest in my graduation project and for reading an earlier version of this thesis.

Furthermore, a special thanks to my fellow students Bart Post and Fiona Sloothaak for commenting on previous versions of this thesis. I believe your remarks have been a great improvement to my work. I am also very grateful to Guus Bollen, with whom I shared an office, for our useful discussions and reassuring talks.

Last but not least, I would like to thank everyone who has supported me these past months, especially my family and my boyfriend. I appreciate it very much.

Contents

List of notations

CHAPTER 1

Introduction

Linear compartment models are used to describe the transport of material between different compartments of a system and appear widely in the fields of systems biology and pharmacokinetics. These models can be given by a directed graph, where the edges represent the flow of material from one compartment to another. The rate of transport from i to j is assumed to be time-invariant and linear in the amount of material in compartment i. We will study identifiability of a particular class of models and the existence of identifiable scaling reparametrizations, following and extending the ideas of Meshkat and Sullivant [**MS14**].

1.1. Problem description

The parameters corresponding to a compartment model are often unknown and are therefore estimated from experimental data. An important step in the modeling process is to check *before* experimenting whether several or even infinitely many parameter sets could yield the same data. If this is the case, it is impossible to tell which parameter values are correct, hence the parameter estimates could lead to wrong predictions.

We assume that the experimental data consists of input-output values: the input corresponding to the amount of material that was added to the system in certain input compartments, and the output corresponding to the amount or concentration of material measured in the output compartments. A model is called *identifiable* if we can recover the parameter values from the (noiseless) input-output data of this experiment. If there is a finite number of possible parameter values corresponding to given input-output data, then we can indeed recover the parameter values, at least locally. More details on identifiability can be found in Chapter 2.

A compartment model can be described by a directed simple graph $G = (V, E)$, i.e. a directed graph without loops or multiple edges. Throughout this thesis, a graph G is assumed to be directed unless stated otherwise, with $n = |V|$ the number of vertices in G and $m = |E|$ the number of edges in G. Let $[k]$ denote the set $\{1, \ldots, k\}$ for given $k \in \mathbb{N}$. We associate to G the $n \times n$ *parameter matrix* $A(G)$ defined by

$$A(G)_{ij} = \begin{cases} a_{ii} & \text{if } i = j \\ a_{ij} & \text{if } j \to i \in E \\ 0 & \text{otherwise,} \end{cases} \tag{1.1}$$

where the a_{ij} $(i, j \in [n], i \neq j)$ are independent real parameters representing the exchange rate from compartment j to compartment i. Possible outflow of material to the exterior is taken into account: each compartment is allowed to have a leak, given by a_{0i}, which represents the exchange rate of material from compartment i to some compartment outside the system (the environment). The diagonal entries of $A(G)$ are defined as $a_{ii} = -a_{0i} - \sum_{j \neq i} a_{ji}$, the negative total flow out of compartment i. Observe that the parameter matrix $A(G)$ uniquely determines G and vice versa.

The *parameter space* of a compartment model given by G consists of all matrices of the form $A(G)$. This space will be denoted by $\Theta_G \subseteq \mathbb{R}^{n\times n}$, to emphasize that the parameter space depends on the graph G. The elements of Θ_G are $n \times n$ matrices which have zeros on positions (i, j) with $i \neq j$ such that $j \to i$ is not an edge in G. In particular, the elements of Θ_G have $n + m$ nonzero positions which we can choose freely; hence the dimension of Θ_G equals $n + m$.

REMARK 1.1.1. In this thesis almost every statement involves a matrix A corresponding to the given graph G: either $A = A(G)$ or $A \in \Theta_G$. When we write $A = A(G)$, we mean the symbolic matrix defined in equation (1.1). On the other hand, by $A \in \Theta_G$ we mean a matrix with the zero pattern of $A(G)$ and parameter values substituted for the symbolic entries a_{ij}, i.e. an element of $\mathbb{R}^{n\times n}$.

A linear compartment model described by a graph G gives rise to a system of linear differential equations. Let $x \in \mathbb{R}^n$ be the state variable representing the concentration of material in each compartment, let $u \in \mathbb{R}^n$ be the input vector corresponding to the input data of the experiment, and let $y \in \mathbb{R}^n$ be the output vector representing the measurement data. Furthermore, let $A = A(G)$ be the parameter matrix corresponding to G, and let $B \in \mathbb{R}^{n\times n}$ be a matrix that indicates from which compartments the output is obtained. Then the transport of material through the compartments can be described by a parametrized system:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + u(t) \\ y(t) &= Bx(t). \end{aligned} \tag{1.2}$$

Note that the matrices A and B do not depend on the time t, since we assume the model to be time-invariant.

As in [**MS14**], we only consider a specific class of linear compartment models, namely the models that satisfy the following three assumptions:

ASSUMPTION 1.1.2. *The input and output take place only in compartment 1.*

This implies that the input vector is of the form $u = (u_1, 0, \ldots, 0)^T \in \mathbb{R}^n$ and that the output vector y is of the form $(x_1, 0, \ldots, 0)^T$. Therefore, system (1.2) can be simplified to system (1.3).

$$\begin{aligned} \dot{x}(t) &= Ax(t) + u(t) \\ y(t) &= x_1(t). \end{aligned} \tag{1.3}$$

The output y is no longer a vector in $\mathbb{R}^n$, but just a value in $\mathbb{R}$. Because of this assumption, we do not need to indicate the input and output in the graph representation of a given model.

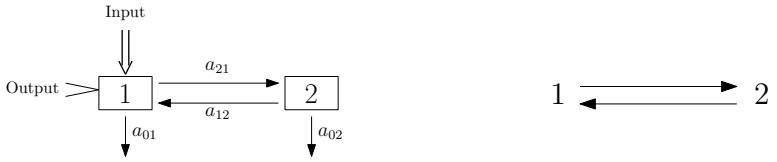
ASSUMPTION 1.1.3. *The graph G is strongly connected.*

In other words, there is a directed path between any pair of vertices in G. A path will be denoted by a sequence of vertices: the sequence $\{v_0, v_1, \ldots, v_k\}$ represents the path from v_0 to v_k using the edges $v_0 \to v_1$, $v_1 \to v_2, \ldots, v_{k-1} \to v_k$. Why we impose Assumption 1.1.3 will become clear in Section 2.2.

ASSUMPTION 1.1.4. *Every compartment has a leak.*

This assumption ensures that all parameters in the model are independent; the a_{ij} were already assumed to be independent and the a_{ii} are defined as

$$a_{ii} = -a_{0i} - \sum_{j\neq i} a_{ji}$$



(A) Linear compartment model (B) Graph representation

FIGURE 1.1. General 2-compartment model

hence $a_{0i} \neq 0$ for all $i \in [n]$ implies that all parameters are independent. A leak at compartment i would correspond to an edge from i to the environment, but these edges are not included in G.

EXAMPLE 1.1.5. Consider the general 2-compartment model and its graph representation in Figure 1.1. This model can be described by the following ODE system:

$$\begin{aligned} \dot{x_1}(t) &= -(a_{01} + a_{21})x_1(t) + a_{12}x_2(t) + u(t) \\ \dot{x_2}(t) &= a_{21}x_1(t) + -(a_{12} + a_{02})x_2(t) \\ y(t) &= x_1(t). \end{aligned}$$

Define $a_{11} = -a_{01} - a_{21}$ and $a_{22} = -a_{12} - a_{02}$, then the equations for $\dot{x_1}, \dot{x_2}$ can be written as

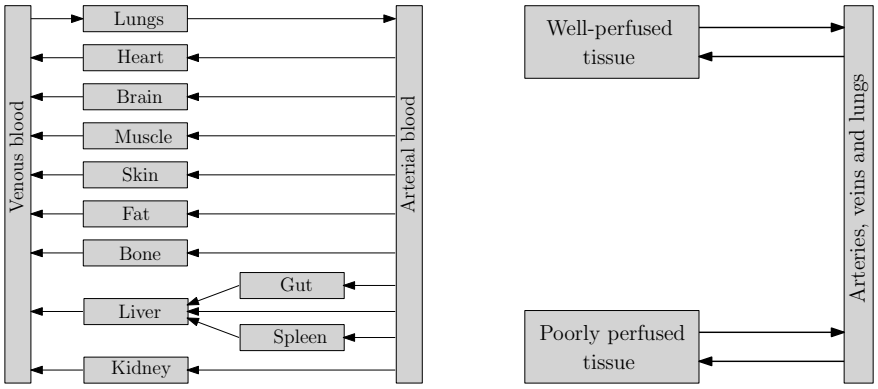
$$\begin{bmatrix} \dot{x_1}(t) \\ \dot{x_2}(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} u_1(t) \\ 0 \end{bmatrix},$$

which bring the system of equations in the form of system (1.3).

When considering model identifiability, a problem that arises is what to do with unidentifiable systems. As we will see in Chapter 2, a model satisfying Assumptions 1.1.2-1.1.4 can never be identifiable unless $n = 1$. We follow the approach of [**MS14**] by searching for identifiable combinations of parameters and using these to find *identifiable scaling reparametrizations* of the original model. An identifiable reparametrization is a map which transforms the model into a lower dimensional model which *is* identifiable. We restrict ourselves to rational scaling reparametrizations; these correspond to a rational scaling of the state variables. The advantage of a scaling reparametrization is that it has a relatively simple connection to the original model. Although the parameters of the reparametrized model do not allow us to estimate the original parameter values, they correspond to certain combinations of the original parameters and hence we can predict relative size changes of these parameters. The aim of this thesis is to find a complete classification of models (satisfying the above assumptions) for which there exists an identifiable scaling reparametrization.

1.2. Applications

Compartment models appear widely in the fields of pharmacokinetics and ecology, but also in engineering. When applying a compartment model, it is always assumed that within each compartment the material is distributed homogeneously. In a pharmacokinetic model the compartments correspond to different organs or regions of the body, and the transport of a certain drug through the body is being modeled. Identifiability of various pharmacokinetic models has been studied, see e.g. [**Yat06, SJH97**]. In ecology, the compartments represent the different species in an ecosystem. The transported material could be any compound, such as calcium or phosphorus [**Hal79**]. In engineering, compartment models may be used to model a distillation column [**EL78**] or the products of chemical reactions [**CP08**].



(A) Full body model (B) Reduced model

FIGURE 1.2. Representations of the human body

From all these applications, pharmacokinetic models correspond best to our assumptions. Indeed, such models are often assumed to be linear. Furthermore, an input-output experiment can be performed by injecting a known amount of drug directly into the blood and at the same time measuring the concentration of this drug in the blood. In this experiment, the input and output both take place in the same compartment; hence if we take compartment 1 to contain the blood, Assumption 1.1.2 is satisfied.

An example of a pharmacokinetic representation of the human body that was taken from [**JRY13**] is presented in Figure 1.2a. Since any compartment lies on a cycle, the corresponding graph is strongly connected. The input-output compartment will be either the arteries or the veins. Many pharmacokinetic models are obtained by lumping together certain compartments of this model. For example, the lungs, arteries and veins are often lumped together to form a new input-output compartment. Another common simplification is to distinguish between well-perfused and poorly perfused tissues, and group these together. This is illustrated in Figure 1.2b. Note that any model obtained by lumping together compartments of a strongly connected model will be strongly connected again.

Finally, consider the assumption that every compartment must have a leak. In a pharmacokinetic model, a leak usually corresponds to the elimination of the drug from the corresponding compartment. Elimination can occur in the liver and kidneys, but also in secondary eliminating tissues such as the skin and the lungs. Hence it is certainly possible to design a model in which also Assumption 1.1.4 is satisfied. In Chapter 5 we will come back to pharmacokinetic models.

REMARK 1.2.1. In a biological setting, the parameters a_{ij} with $i \neq j$ must be nonnegative, or they would correspond to a negative flow. Combining this with the assumption that every compartment has a leak, it follows that the parameters a_{ii} defined as $-a_{0i} - \sum_{j \neq i} a_{ji}$ must be strictly negative. These constraints are not accounted for in our identifiability analysis, but they may help to recover the correct parameters when a model is only locally identifiable or even unidentifiable.

1.3. Previous work

The concept of identifiability of dynamical systems was introduced by Bellman and Åström [**BÅ70**] in 1970, and has been studied extensively since. Godfrey [**God83**] gives a thorough description of compartment models and applications, also treating the concept of identifiability. There have been different approaches to

determining whether a system is identifiable or not, where one has to distinguish between local and global identifiability, as defined in [**BÅ70**]. These methods include a Taylor series expansion approach [**Poh78**], a Laplace transform approach [**BÅ70**], a similarity transformation approach [**CG85**] also known as exhaustive modeling, and a differential algebra approach [**LG94, MED09**]. Identifiability of biological and pharmacokinetic systems has been studied in [**SJH97, Yat06**]. A comparison of different algorithms for parameter identifiability analysis of biological systems is given in [**RKS$^+$14**].

We consider the question what to do with unidentifiable systems. In this case there are too many parameters (unknowns) compared to the amount of information obtained from the experiment, so the parameter space has to be constrained somehow. This can be done using a reparametrization of the original system, which reduces it to a lower dimensional identifiable system. A procedure for finding such a reparametrization has been discussed for the differential algebra approach [**LG94, MED09, MAD11**], for the Taylor series approach [**EC00**], and for the similarity transformation approach [**CG98**]. However, as we only consider a specific class of models, the problem of finding identifiable reparametrizations becomes easier compared to the general setting discussed in these references.

The problem setting of this thesis was taken from [**MS14**], where the same class of models is considered and several results for the existence of identifiable reparametrizations are derived. Most results of [**MS14**] will be discussed and for some also a sketch of the proof is given. Meshkat and Sullivant also present an algorithm to find an identifiable scaling reparametrization if one exists, which was used to obtain Example 2.4.2.

1.4. Outline of the thesis

The main goal of this thesis is to obtain a classification of graphs (satisfying our assumptions) for which there exists an identifiable scaling reparametrization. We continue on the work of Meshkat and Sullivant, presented in [**MS14**]. In order to be self-contained, Chapter 2 contains a brief summary of the definitions and results given in [**MS14**]. Most results are stated without proof, as these can be found in the corresponding article. In Section 2.1 it is defined what it means to be identifiable. An identifiability analysis in Section 2.2 shows that the models satisfying our assumptions can never be identifiable, hence in Section 2.3 the concept of an identifiable reparametrization is defined. Subsequently, a criterion for the existence of an identifiable scaling reparametrization is presented in Section 2.4. This criterion is referred to as the *dimension criterion*, since it depends on the dimension of the image of a certain map. If a graph has an identifiable scaling reparametrization, we say it has the *expected dimension*. Chapter 2 relies heavily upon [**MS14**].

In Chapter 3, we consider the following question:

Question 1.4.1. *For which graphs satisfying our assumptions does there exist an identifiable scaling reparametrization?*

To answer this question, we will derive a reformulation of the dimension criterion from [**MS14**]. In Section 3.1 we analyze the kernel of the differential of the input-output map that was defined in Section 2.2. Then we consider the preimage of this kernel under the commutator map $[\cdot, A]$, which we define in Section 3.2. In Section 3.3 this leads to a necessary condition for a graph to have the expected dimension, i.e. for an identifiable scaling reparametrization to exist. Next, in Section 3.4 we obtain a criterion based on bipartite graphs, the main result of this thesis.

THEOREM 1.4.2. *Let $G = (V, E)$ be a graph satisfying Assumptions 1.1.2-1.1.4. Then G has an identifiable scaling reparametrization if and only if the matrix $B(G)$ defined by*

$$B(G)_{(k,l),(i,j)} = \begin{cases} -a_{jl} & \text{if } i = k,\ j \neq l \text{ and } l \to j \in E \\ a_{ki} & \text{if } i \neq k,\ j = l \text{ and } i \to k \in E \\ a_{kk} - a_{ll} & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise.} \end{cases}$$

has full column rank.

The final section of Chapter 2 is devoted to a complexity analysis, comparing an algorithm based on the dimension criterion from [**MS14**] with an algorithm based on our newly obtained criterion.

In Chapter 4, the following question is considered:

QUESTION 1.4.3. *What constructions can we apply to a given graph, such that the resulting graph has an identifiable scaling reparametrization?*

The first two sections of this chapter present definitions and results from [**MS14**]. In Section 4.3 several new constructions are derived, which allow us to construct many graphs which have the expected dimension. This is analyzed in Section 4.4, where we discuss the concept of an ear decomposition of a graph. An ear is defined to be a directed path and an ear is said to be trivial if it is a path of length one. Moreover, an ear decomposition of G is a partition of the edges of G into a sequence ears, where the first ear is a cycle and the endpoints of each ear belong to earlier ears in the sequence, while the internal vertices do not. We obtain the following result:

THEOREM 1.4.4. *Let G be a graph that has a correct ear decomposition, i.e. an ear decomposition without trivial ears, whose initial cycle contains vertex 1. Then G has the expected dimension.*

In other words, if G has a correct ear decomposition, then there exists an identifiable scaling reparametrization for G. Chapter 4 is concluded with some computational results (obtained using MATHEMATICA) in Section 4.5. For graphs with few vertices, we present the size of certain classes of graphs which have the expected dimension.

The final chapter, Chapter 5, gives a short summary of the previous chapters and the obtained results, which are discussed from a pharmacokinetic point of view as well. Also, a few words will be spent on remaining questions and future research.

At the start of each chapter, we introduce the necessary notation and definitions. All notations used in this thesis can also be found in the list of notations on page vii. Furthermore, some of the proofs in Chapter 2 rely on algebraic theorems. A short overview of the necessary definitions and results can be found in Appendix A.

CHAPTER 2

Identifiability and reparametrizations

In this chapter identifiability is defined more precisely and we shall see that a linear compartment model satisfying Assumptions 1.1.2-1.1.4 can never be identifiable. Therefore, we define what it means to have an identifiable scaling reparametrization. These definitions can be found in Sections 2.1-2.3, which are based on [**MS14**]. In the final section we consider a criterion for the existence of an identifiable scaling reparametrization that was found by Meshkat and Sullivant.

First recall some basic notation that will be used throughout this thesis. The $n \times n$ identity matrix will be denoted I_n, or simply I when the dimension is clear from the context. Moreover, $e_1, \ldots, e_n \in \mathbb{R}^n$ represent the standard unit vectors, and $E_{ij} = e_i e_j^T \in \mathbb{R}^{n \times n}$ denotes the matrix with a 1 at position (i, j) and zeros elsewhere.

The rank of a matrix M is denoted $\mathrm{rk}(M)$, and equals the greatest number of linearly independent columns of M. When M has size $k \times l$, we say that M has full rank if $\mathrm{rk}(M) = \min\{k, l\}$. The kernel of M is denoted $\ker(M)$.

Furthermore, let $f : X \to Y$ be a mapping, where X and Y are varieties. The *differential* (or *derivative*, or *Jacobian*) of f at the point p is denoted $d_p f$ and contains all first-order partial derivatives of f evaluated at p. The chain rule applies when calculating the differential of the composition $g \circ f$:

$$d_p(g \circ f) = (d_{f(p)} g)(d_p f).$$

The differential $d_p f$ maps the *tangent space* $T_p(X)$ of X at p to the tangent space $T_{f(p)}(Y)$. When f is a surjective regular map, the differential $d_p f$ is generically surjective, in which case the rank of $d_p f$ equals the dimension of $T_p(Y)$. See also Appendix A.

2.1. Identifiability and genericity of parameters

There have been different approaches to the problem of system identifiability, as described in Section 1.3. Here we follow the approach of [**MS14**], which means forming an *input-output equation* to relate the observed data to the (unknown) parameters:

$$\psi(y, u, A) = 0.$$

This equation depends only on the parameters A, input u, and output y, but it may also contain derivatives $\dot{y}, \ddot{y}, \ldots, \dot{u}, \ddot{u}, \ldots$. The input-output equation is obtained by eliminating the state variable x from the parametrized system (1.2). In general, this can be done via Ritt's pseudodivision algorithm, as described in [**BSAD07**]. However, for the class of models corresponding to system (1.3) the problem is less complicated, as we shall see in Section 2.2.

An input-output equation gives rise to a coefficient map c that maps a parameter matrix $A \in \Theta_G$ to the coefficient vector of the input-output equation. This vector contains the coefficients of $y, \dot{y}, \ddot{y}, \ldots, u, \dot{u}, \ddot{u}, \ldots$ in terms of the parameter values a_{ij}. A model is called identifiable when it is possible to recover the parameter values from observing the input-output behaviour. This behaviour is given by the input-output equation and hence by the coefficient map c, so using c we

can define identifiability in mathematical terms. Suppose two distinct parameter matrices A, A' yield the input-output data, i.e. $c(A) = c(A')$. Then it is impossible to tell from only observing the relations among input and output whether the parameter values corresponding to the model should be those of A or of A'. This shows that identifiability of the model can be stated as an injectivity condition on the coefficient map.

DEFINITION 2.1.1. Let $c : \Theta_G \to \mathbb{R}^k$ be a function on the parameter space Θ_G. The parameters are said to be *globally identifiable* from c if and only if the map c is injective. The parameters are *locally identifiable* from c if and only if c is finite-to-one, i.e. for any A there exists an open neighborhood $\mathcal{N}_A$ such that c is injective on $\mathcal{N}_A$. Otherwise, the parameters are *unidentifiable* from c.

We say that a model is globally (resp. locally) identifiable if its parameters are globally (resp. locally) identifiable from some function c. However, it may occur that the parameters are identifiable almost everywhere, except for some small subset of the parameter space. According to Definition 2.1.1 such a model is unidentifiable, even though it is almost identifiable. Therefore, the concept of *generic identifiability* is introduced:

DEFINITION 2.1.2. Let $c : \Theta_G \to \mathbb{R}^k$ be a function on the parameter space Θ_G. The parameters are said to be *generically globally identifiable* from c if and only if the map c is injective on some dense open subset $U \subseteq \Theta_G$. Similarly, the parameters are *generically locally identifiable* from c if they are locally identifiable on some dense open subset of U, and *generically unidentifiable* from c otherwise.

Clearly, global identifiability implies local identifiability. In this thesis only generic local identifiability and generic unidentifiability are considered, so from now on identifiable means generically locally identifiable, and unidentifiable means generically unidentifiable. In other words, G is identifiable if and only if there is a map $c : \Theta_G \to \mathbb{R}^k$ that is locally injective on an open dense subset $U \subseteq \Theta_G$. This set U is called a *generic* set, and the points $A \in U$ are called *sufficiently general*.

The above injectivity conditions can be translated to a condition on the dimensions of the parameter space and the image of c, as stated in the following proposition:

PROPOSITION 2.1.3 ([**MS14**, Prop. 2.6]). *Let $c : \Theta_G \to \mathbb{R}^k$ be a rational map. Then the model is identifiable from c if and only if the dimension of the image of c equals the dimension of the parameter space Θ_G.*

In the next section, the input-output equation and the corresponding coefficient map are derived for the ODE system (1.3).

2.2. The double characteristic polynomial map

Given a matrix M, let M_1 denote the submatrix obtained by deleting the first row and column of M. For linear ODE models satisfying our assumptions we obtain the following input-output equation:

THEOREM 2.2.1 ([**MS14**, Thm. 2.2]). *Let G be a graph satisfying Assumptions 1.1.2-1.1.4 and let $A = A(G)$. Then the input-output equation corresponding to* (1.3) *becomes*

$$y^{(n)} + c_1 y^{(n-1)} + \ldots + c_n y = u_1^{(n-1)} + d_1 u_1^{(n-1)} + \ldots + d_{n-1} u_1 \tag{2.1}$$

where $c_1, \ldots, c_n$ and $d_1, \ldots, d_{n-1}$ are the coefficients of the characteristic polynomial of A and A_1, respectively.

Because of this theorem, the matrices A and A_1 will play a major role in the rest of this thesis. A sketch of the proof is given, more details can be found in [**MS14**].

Proof sketch.
Let ∂ denote the differential operator $\frac{d}{dt}$, then we can rewrite the ODE system (1.3) as $(\partial I_n - A)x = u$. Applying Cramer's rule we get $x_1 = \det(A')/\det(\partial I_n - A)$ where A' is the matrix $\partial I_n - A$ with the first column replaced by u. Since only the first entry of u is nonzero, we have $\det(A') = u_1 \det(\partial I_{n-1} - A_1)$. After replacing x_1 by y we obtain $\det(\partial I_n - A)y = \det(\partial I_{n-1} - A_1)u_1$, which is an equation purely in terms of A, u and y. Let f, f_1 denote the characteristic polynomial of A, A_1, respectively, then this equality can be written as $f(\frac{d}{dt})y = f_1(\frac{d}{dt})u_1$. Meshkat and Sullivant show that f and f_1 are relatively prime if and only if G is strongly connected. This can be seen by rearranging the matrix A such that it becomes a block upper triangular matrix, each block corresponding to a strongly connected component. Then the characteristic polynomial of A factors through the characteristic polynomials of the diagonal blocks, so the strongly connected components not containing vertex 1 cause common factors in f and f_1. If G is strongly connected no factors cancel out, hence we obtain equation (2.1). □

This proof shows that if a graph is not strongly connected, the input-output equation will result from taking the strongly connected component containing vertex 1. Thus, for such models there are parameters which do not appear in the input-output equation. In this case the dimension of Θ_G will be too large compared to the dimension of the image of c. In other words, when G is not strongly connected, it is certainly not identifiable. This motivates why we only consider strongly connected graphs when characterizing whether G is identifiable (Assumption 1.1.3).

From the input-output equation we obtain the coefficient map $c : \Theta_G \to \mathbb{R}^{2n-1}$ given by

$$c(A) := (c_1, \ldots, c_n, d_1, \ldots, d_{n-1}),$$

where $c_1, \ldots, c_n, d_1, \ldots, d_{n-1}$ are the coefficients of equation (2.1). This map is called the *double characteristic polynomial map*.

Example 2.2.2. Consider the general 2-compartment model in Figure 1.1. We derived the ODE system corresponding to this model in Example 1.1.5, and here we compute its double characteristic polynomial map from $\det(tI - A)$ and $\det(tI - A_1)$:

$$\det \begin{bmatrix} t - a_{11} & -a_{12} \\ -a_{21} & t - a_{22} \end{bmatrix} = t^2 - (a_{11} + a_{22})t + a_{11}a_{22} - a_{21}a_{12},$$
$$\det[t - a_{22}] = t - a_{22}.$$

We obtain $c : \Theta_G \to \mathbb{R}^3$ defined by

$$c(A) = (-a_{11} - a_{22},\ a_{11}a_{22} - a_{12}a_{21},\ -a_{22}).$$

There are four parameters to determine, while $c(A)$ only gives three parameter combinations, so we can never recover all parameters. Indeed, from the entries of $c(A)$ it can be seen that a_{11} and a_{22} are identifiable, but for a_{12} and a_{21} we only learn the value of their product. We conclude that G is unidentifiable.

In the previous example, there were too many parameters compared to the amount of information one obtains from $c(A)$. As shown in [**MS14**], this is no coincidence: there will always be too many parameters to recover all of them. Due to our assumptions on the model there are $m + n$ independent parameters, i.e. the parameter space Θ_G has dimension $m + n$. Proposition 2.1.3 implies that G is identifiable if and only the dimension of the image of the double characteristic

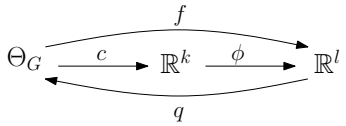


FIGURE 2.1. A reparametrization q and identifiable map f

polynomial map c equals $m + n$. However, if G is strongly connected and $n > 1$ we must have $m \geq n$, which implies that $m + n \geq 2n > 2n - 1 \geq \dim \operatorname{im} c$. This shows that a linear compartment model satisfying our assumptions can never be identifiable, except for the trivial case where $n = 1$.

2.3. Identifiable scaling reparametrizations

One approach when dealing with unidentifiable models is to restrict the parameter space to a lower-dimensional space, such that identifiability is guaranteed. We look for identifiable combinations of parameters, which may be used to find a *reparametrization* of the model. An identifiable reparametrization reduces the number of parameters, such that the model becomes identifiable.

DEFINITION 2.3.1. A *reparametrization* of the input-output equation is a map $q : \mathbb{R}^l \to \Theta_G$ such that $\operatorname{im}(c \circ q) = \operatorname{im} c$, where q is called identifiable if the parameters in $\mathbb{R}^l$ are identifiable from $c \circ q$.

We also introduce the notion of an identifiable function:

DEFINITION 2.3.2. Let $c : \Theta_G \to \mathbb{R}^k$ be a function on the parameter space Θ_G. A function $f : \Theta_G \to \mathbb{R}^l$ is *globally identifiable* from c if there exists a function $\phi : \mathbb{R}^k \to \mathbb{R}^l$ such that $\phi \circ c = f$. The function f is *locally identifiable* from c if there exists a function $\phi : \mathbb{R}^k \to \mathbb{R}^l$ that is finitely multivalued, such that $\phi \circ c = f$.

Similar to Definition 2.1.2, generic identifiability of a function means that it is identifiable on some open dense subset of Θ_G. We shall always consider generic local identifiability, hence from now on a function is called identifiable if it is generically locally identifiable.

According to Meshkat and Sullivant, finding an identifiable reparametrization is equivalent to finding a map $\phi : \operatorname{im} c \to \mathbb{R}^l$ such that $f := \phi \circ c$ consists of functions that are identifiable from c. (Figure 2.1)

Linear compartment models have many applications in biological systems. For these applications a rational scaling reparametrization is preferred over more complicated reparametrizations. Such a reparametrization corresponds to a rational scaling of the state variables and therefore it respects the biological properties of the model. If an identifiable scaling parametrization exists, it can always be made rational [**MS14**]. To obtain an identifiable scaling reparametrization, we search for functions $f_i : \Theta_G \to \mathbb{R}$, $i = 1, \ldots, n$, such that the model with scaled state variables $X_1, \ldots, X_n$ defined by

$$X_i = f_i(A)x_i$$

is identifiable. The parameters b_{ij} of the scaled system are given by

$$b_{ij} = a_{ij} f_i(A)/f_j(A).$$

The corresponding reparametrization is identifiable if and only if the new parameters b_{ij} are identifiable functions from c.

Note that we need the reparametrized system to correspond to the input-output data obtained from the original system. Therefore, we must have $X_1 = x_1$ and

hence $f_1(A) = 1$. Moreover, the diagonal entries of the parameter matrix remain the same:

$$b_{ii} = a_{ii} f_i(A)/f_i(A) = a_{ii}.$$

Once an identifiable scaling reparametrization is found, we can recover the corresponding scaled parameter values b_{ij}. However, from the reparametrized model we cannot recover the original parameter values. The scaling has cancelled out a subset of the parameters, and we learn nothing about their values. We do know the values of certain combinations of parameters, and using these we can still predict relative size changes.

LEMMA 2.3.3 ([**MS14**, Cor. 2.13]). *If G has an identifiable scaling reparametrization, then G has at most $2n-2$ edges.*

Because of Lemma 2.3.3 we only need to consider graphs which have at most $2n-2$ edges, so from now on we will assume $m \leq 2n-2$.

2.4. The dimension criterion

The main result of Meshkat and Sullivant gives a criterion to decide whether an identifiable scaling reparametrization exists.

THEOREM 2.4.1 ([**MS14**, Thm. 1.2]). *A graph G has an identifiable scaling reparametrization if and only if the dimension of the image of the double characteristic polynomial map is $m+1$.*

We will refer to Theorem 2.4.1 as the *dimension criterion*. This criterion reduces the problem of deciding whether or not an identifiable scaling reparametrization exists to calculating the dimension of the double characteristic polynomial map. Note that the dimension criterion allows us to check whether G has an identifiable scaling reparametrization by calculating the rank of the differential $d_A c$ of the map c at a sufficiently general point $A \in \Theta_G$. This follows from the fact c is a polynomial map, surjective on $\operatorname{im} c$, hence for A in an open dense subset of Θ_G the rank of $d_A c$ equals the dimension of the image of c (Appendix A).

EXAMPLE 2.4.2. Consider the graph G given in Figure 2.2, which represents the general 3-compartment model. Then G is not identifiable, but calculating $d_A c$ shows that $\dim \operatorname{im} c = \operatorname{rk}(d_A c) = 5 = m+1$ for sufficiently general A. So according to the dimension criterion, G has an identifiable scaling reparametrization. Using the algorithm presented in [**MS14**], we obtain the following scaling functions:

$$f_1(A) = 1, \qquad f_2(A) = 1/a_{21}, \qquad f_3(A) = 1/a_{31}.$$

We replace the original parameters a_{ij} by the scaled parameters

$$b_{ij} = a_{ij} f_i(A)/f_j(A),$$

such that the parameter matrix A is replaced by the scaled matrix A':

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & 0 \\ a_{31} & 0 & a_{33} \end{bmatrix} \qquad \rightsquigarrow \qquad A' = \begin{bmatrix} a_{11} & a_{12}a_{21} & a_{13}a_{31} \\ 1 & a_{22} & 0 \\ 1 & 0 & a_{33} \end{bmatrix}.$$

The scaled state variables become $X_1 = x_1$, $X_2 = \frac{1}{a_{21}} x_2$ and $X_3 = \frac{1}{a_{31}} x_3$, hence the reparametrized model can be written as:

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ 1 & b_{22} & 0 \\ 1 & 0 & b_{33} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} u_1 \\ 0 \\ 0 \end{bmatrix}$$
$$y = X_1.$$

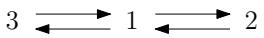


FIGURE 2.2. General 3-compartment model (Example 2.4.2)

Note that in this example, the reparametrized system has only 5 parameters while the original system had 7. Let Θ'_G be the reduced parameter space. Because of the decrease in the number of parameters, we have $\dim \Theta'_G = 5 = m+1 = \dim \operatorname{im} c$, hence the system has become identifiable according to Proposition 2.1.3.

DEFINITION 2.4.3. We say that G *has the expected dimension* if the dimension of the image of the double characteristic polynomial map equals $m + 1$.

In other words, an identifiable scaling reparametrization exists if and only if G has the expected dimension. If this is the case, the scaling reparametrization can be found using the algorithm presented in [**MS14**]. Some known properties of graphs with the expected dimension will be discussed in Section 4.2. However, a full classification of graphs with the expected dimension remains to be found. The goal of this thesis is to find such a classification; our first step is to derive an alternative criterion to test whether a given graph has the expected dimension, which can be found in the next chapter.

CHAPTER 3

Reformulating the dimension criterion

In the previous chapter we discussed all relevant definitions and earlier results. From now on we take a new approach, starting with a reformulation of the dimension criterion that was given in Theorem 2.4.1. This leads us to our main result: an alternative criterion to test whether a given graph has the expected dimension. In other words, we determine whether an identifiable scaling reparametrization exists. This criterion can be verified in probabilistic polynomial time by a randomized algorithm presented in Section 3.5.

So far we have been working over the real numbers, since all parameters are assumed to be real. However, $\mathbb{R}$ lies dense in $\mathbb{C}$ and the dimension of the image of the double characteristic polynomial map c is determined by the rank of its Jacobian at a sufficiently general point, so we might as well work over the complex numbers to determine the dimension of the image of c. From now on let $\Theta_G \subseteq \mathbb{C}^{n\times n}$ and let $c : \Theta_G \to \mathbb{C}^{2n-1}$. Working over the complex numbers simplifies issues concerning diagonalizability, which we shall be using later on.

In this chapter we will consider the matrix group $\mathrm{GL}_n(\mathbb{C})$, the *general linear group*, consisting of all invertible $n \times n$ matrices over $\mathbb{C}$. We shall be working over the complex numbers, hence for simplicity GL_n is written instead of $\mathrm{GL}_n(\mathbb{C})$. The tangent space of GL_n at the identity is its *Lie algebra* $\mathfrak{gl}_n$. This space consists of all $n \times n$ complex matrices, with the commutator serving as the Lie bracket: $[X, A] := XA - AX$. We write $\mathfrak{gl}_n$ instead of $\mathbb{C}^{n\times n}$ to emphasize that it arises as the tangent space of GL_n.

Furthermore, given a matrix $A \in \mathbb{C}^{n\times n}$, the *centralizer* of A in $\mathfrak{gl}_n$ is denoted $Z_{\mathfrak{gl}_n}(A)$. It contains all $X \in \mathfrak{gl}_n$ that commute with A, i.e. $[X, A] = 0$.

3.1. The kernel of the differential map

For sufficiently general $A \in \Theta_G$ we have the following chain of equalities:

$$\dim \operatorname{im} c = \operatorname{rk}(d_A c) = m + n - \dim \ker(d_A c), \tag{3.1}$$

where $d_A c$ denotes the differential (or Jacobian) of c at the point A. The first equality was already mentioned in the previous section; it holds since $c : \Theta_G \to \operatorname{im} c$ is a surjective polynomial map in characteristic zero. The second equality follows directly from the rank-nullity theorem. Thus we have shown the following lemma.

Lemma 3.1.1. *The dimension of the image of the double characteristic polynomial map c equals $m+1$ if and only if the dimension of the kernel of the differential $d_A c$ equals $n-1$ for sufficiently general $A \in \Theta_G$.*

Using this result, we can determine whether a given model has the expected dimension by calculating the rank of the differential $d_A c$. In order to classify which models have the expected dimension, we need to know what the kernel of $d_A c$ looks like. By definition of the double characteristic polynomial map c, the kernel of $d_A c$ is equal to the intersection of the two kernels corresponding to the differentials of the characteristic polynomials of A and A_1. Using this observation we will derive the form of $\ker(d_A c)$.

PROPOSITION 3.1.2. *For sufficiently general $A \in \Theta$, the kernel of the differential map $d_A c : \Theta_G \to \mathbb{C}^{2n-1}$ is given by*

$$\left\{ C \in \Theta_G \;\middle|\; \begin{array}{lll} \exists\, X \in \mathfrak{gl}_n : & [X, A] & = C \\ \exists\, Y \in \mathfrak{gl}_{n-1} : & [Y, A_1] & = C_1 \end{array} \right\} \tag{3.2}$$

where A_1, C_1 denote the matrices obtained by removing the first row and the first column from A, C, respectively.

PROOF. We begin by writing $c(A) = [c_0(A)|c_1(A)]$, where c_0, c_1 are the coefficient maps corresponding to the characteristic polynomials of A, A_1, respectively. Let $d_A c_0 \in \mathbb{C}^{n \times (n+m)}$ and $d_A c_1 \in \mathbb{C}^{(n-1)\times(n+m)}$ be the differential maps of c_0, c_1, respectively, containing the partial derivatives with respect to the model parameters. The differential of c can be written as

$$d_A c = \left[\begin{array}{c} (d_A c_0)^T \\ \hline (d_A c_1)^T \end{array} \right],$$

which shows that X lies in the kernel of $d_A c$ if and only if X lies in both the kernel of $d_A c_0$ and $d_A c_1$.

Consider the map $c_0 : \mathbb{C}^{n\times n} \to \mathbb{C}^n$ and define the map $\psi : \mathrm{GL}_n \to \mathbb{C}^{n\times n}$ that sends g to gAg^{-1} for some fixed $A \in \mathbb{C}^{n\times n}$. Define the composition $\phi := c_0 \circ \psi$, which is of the form

$$\begin{array}{ccccc} \mathrm{GL}_n & \xrightarrow{\psi} & \mathbb{C}^{n\times n} & \xrightarrow{c_0} & \mathbb{C}^n \\ g & \longmapsto & gAg^{-1} & \longmapsto & (a_1, \ldots, a_n) \end{array}$$

where $a_1, \ldots, a_n$ are the coefficients of the characteristic polynomial of gAg^{-1}. The characteristic polynomial of A is invariant under conjugation by an element of GL_n, hence $a_1, \ldots, a_n$ are equal to the coefficients of the characteristic polynomial of A. This implies that the composition $c_0 \circ \psi$ is in fact a constant map sending each $g \in \mathrm{GL}_n$ to the fixed point $(a_1, \ldots, a_n) \in \mathbb{C}^n$. Therefore, the differential $d\phi$ is identically zero; in particular $d_I \phi = 0$ for the $n \times n$-identity matrix I.

According to the chain rule for differentiation, $d_I \phi = (d_A c_0)(d_I \psi)$. From the observation that $d_I \phi = 0$ it follows that the image of $d_I \psi$ is contained in the kernel of $d_A c_0$. Note that $d_I \psi$ works on the tangent space $T_I(\mathrm{GL}_n) = \mathfrak{gl}_n$. We can find the image of $d_I \psi$ by determining the coefficient of ε in $\psi(I+\varepsilon X) = (I+\varepsilon X)A(I+\varepsilon X)^{-1}$ modulo ε^2, as described in Appendix A. We find

$$\begin{aligned} (I + \varepsilon X)A(I + \varepsilon X)^{-1} &= (I + \varepsilon X)A(I - \varepsilon X) \mod \varepsilon^2 \\ &= A + \varepsilon(XA - AX) \mod \varepsilon^2, \end{aligned}$$

so the coefficient of ε in the above expression equals the Lie bracket (or commutator) $[X, A] := XA - AX$. This shows that for $X \in \mathfrak{gl}_n$ we have $(d_I \psi)(X) = [X, A]$, hence the image of $d_I \psi$ equals $[\mathfrak{gl}_n, A]$. We conclude that $[\mathfrak{gl}_n, A] \subseteq \ker(d_A c_0)$.

On the other hand, since c_0 is a surjective polynomial map from $\mathbb{C}^{n\times n}$ to $\mathbb{C}^n$, the dimension of the kernel of $d_A c_0$ is generically equal to $n^2 - n$ (Proposition A.4). The kernel of the Lie bracket $[\cdot, A]$ is precisely the centralizer of A in $\mathfrak{gl}_n$, which has dimension n according to Lemma 3.1.3 below. This shows that the dimension of $[\mathfrak{gl}_n, A]$ equals $n^2 - n$, and because the kernel is a linear subspace of Θ_G we conclude that $\ker(d_A c_0) = [\mathfrak{gl}_n, A]$.

The same argument applies to the map $c_1 : \mathbb{C}^{n\times n} \to \mathbb{C}^{n-1}$, showing that $\ker(d_A c_1) = \{C \mid C_1 \in [\mathfrak{gl}_{n-1}, B]\}$ for any $A \in \Theta_G$. Combining these results, we see that C lies in the kernel of $d_A c$ if and only if C lies in $[\mathfrak{gl}_n, A]$ and C_1 lies in $[\mathfrak{gl}_{n-1}, A_1]$. In other words, there exist $X \in \mathfrak{gl}_n$ and $Y \in \mathfrak{gl}_{n-1}$ such that $[X, A] = C$ and $[Y, A_1] = C_1$. □

LEMMA 3.1.3. *For sufficiently general $A \in \mathbb{C}^{n\times n}$, the centralizer of A in $\mathfrak{gl}_n$, denoted by $Z_{\mathfrak{gl}_n}(A)$, has dimension n.*

PROOF. Let $A \in \mathbb{C}^{n\times n}$ be such that it has n distinct, nonzero eigenvalues. The centralizer $Z_{\mathfrak{gl}_n}(A)$ consists of all matrices $B \in \mathfrak{gl}_n$ that commute with A. Such an element B must leave the eigenspaces of A invariant: if v is an eigenvector of A with eigenvalue λ, then $BA = AB$ implies $\lambda(Bv) = A(Bv)$, hence Bv is in the same eigenspace. Therefore the elements of $Z_{\mathfrak{gl}_n}(A)$ must be diagonalized by the same basis that diagonalizes A, and such elements are determined by their eigenvalues on this basis. This leaves us n choices, i.e. the centralizer has dimension n. □

3.2. The preimage of the kernel

In the previous section we saw that G has the expected dimension if and only if the kernel of the differential of the double characteristic polynomial map has dimension $n-1$. So far, we have determined this kernel, but what can we say about its dimension?

For given $A \in \Theta_G$, the kernel of $d_A c$ equals the image of the commutator map $X \mapsto [X, A]$ restricted to the linear subspace $V_A \subseteq \mathfrak{gl}_n$ defined by

$$V_A := \{X \in \mathfrak{gl}_n \mid [X, A] \in \ker(d_A c)\},$$

which is the preimage of $\ker(d_A c)$ under the commutator map. From the fact that $[\cdot, A]$ is a linear map it follows that V_A is a linear subspace of $\mathfrak{gl}_n$. Furthermore, any X that commutes with A is contained in V_A, since $[X, A] = 0$ and $[X, A]_1 = 0 = [Y, A_1]$ for any $Y \in Z_{\mathfrak{gl}_{n-1}}(A_1)$. From Lemma 3.1.3 we know that the kernel of the commutator map has dimension n, which implies that

$$\dim \ker(d_A c) = n - 1 \quad \Leftrightarrow \quad \dim V_A = 2n - 1. \tag{3.3}$$

In words, G has the expected dimension if and only if the dimension of V_A is $2n-1$. Therefore, we will examine the structure of V_A for a given graph G. Besides the centralizer of A, V_A will always contain the space $\mathcal{D}_n$ of all $n \times n$ diagonal matrices with entries in $\mathbb{C}$. Indeed, by computing $DA - AD$ for $D \in \mathcal{D}_n$ we find

$$(DA - AD)_{ij} = (d_{ii} - d_{jj})a_{ij} \quad \text{for } i, j = 1, \ldots, n. \tag{3.4}$$

If position (i, j) of A is zero, it follows that position (i, j) of $[D, A]$ is zero as well. This shows that $[D, A]$ has the correct zero pattern, i.e. $[D, A] \in \Theta_G$. Moreover, one can check that

$$[D_1, A_1] = [D, A]_1$$

so the second constraint for being in the kernel of $d_A c$ is also satisfied. The space of $n \times n$ diagonal matrices $\mathcal{D}_n$ is again n-dimensional, hence we already have two n-dimensional subspaces of V_A. However, these two subspaces have a nontrivial intersection, as the next lemma shows.

LEMMA 3.2.1. *$Z_{\mathfrak{gl}_n}(A) \cap \mathcal{D}_n = \mathbb{C}I_n$ for sufficiently general $A \in \Theta_G$ and G strongly connected.*

PROOF. Suppose that $X = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in Z_{\mathfrak{gl}_n}(A) \cap \mathcal{D}_n$, then by definition of $Z_{\mathfrak{gl}_n}(A)$, X satisfies $XA = AX$. Combining this equality with equation (3.4) shows that for $a_{ij} \neq 0$ this equality implies that $\lambda_i a_{ij} = \lambda_j a_{ij}$ and hence $\lambda_i = \lambda_j$. Since G is strongly connected, starting from vertex 1 we can get to any other vertex j along some path $\{1, i_1, \ldots, i_k, j\}$. The corresponding entries $a_{i_1 1}, a_{i_2 i_1}, \ldots, a_{j i_k}$ are nonzero for sufficiently general A, and by the previous observation it follows that $\lambda_1 = \lambda_{i_1} = \ldots = \lambda_j$. But we can find such a path for any vertex $j \in [n]$, so we conclude that $\lambda_1 = \ldots = \lambda_n$ and therefore X must be of the form cI_n, $c \in \mathbb{C}$. □

What we have seen so far is that $Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n \subseteq V_A$ for any G. According to Lemma 3.2.1 this is a subspace of dimension $2n-1$, so the dimension of V_A is at least $2n-1$. Combining this with equation (3.3), it follows that $V_A = Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n$ if and only if G has the expected dimension.

COROLLARY 3.2.2. *G has the expected dimension if and only if*

$$V_A / (Z_{\mathfrak{gl}_n} + \mathcal{D}_n) = \{0\}.$$

We shall now derive several restrictions on the form of elements of the quotient space in the above corollary. An important tool will be the following lemma:

LEMMA 3.2.3. *Let G be a graph, not necessarily strongly connected, and let $A \in \Theta_G$ sufficiently general. Suppose $v = (v_1, \ldots, v_n)$ is an eigenvector of A. If $v_i \neq 0$ and there exists a path from i to j in G, then also $v_j \neq 0$.*

PROOF. Let $v = (v_1, \ldots, v_n) \in \mathbb{C}^n$ such that $Av = \lambda v$. Partition the indices $1, \ldots, n$ into two sets, $[n] = I \sqcup J$, such that $v_i = 0 \ \forall i \in I$ and $v_j \neq 0 \ \forall j \in J$. Construct the $|J| \times |J|$ matrix A' by removing the rows and columns of A indexed by elements of I. Similarly, let v' be the vector obtained from v by removing its zero entries. Then we have $A'v' = \lambda v'$ and for sufficiently general A this determines the vector $v' \in \mathbb{C}^{|J|}$ up to multiplication by a scalar. Since v is obtained from v' by adding zero entries at positions indexed by I, also v has been determined up to scalar multiplication. However, for $Av = \lambda v$ to hold, v must satisfy a system of n linear equations of the form

$$\sum_{j \in J} a_{ij} v_j = \lambda v_i, \qquad i \in [n].$$

We know that v must be a solution of the subset of these equations corresponding to $i \in J$, since $Av' = \lambda v'$. The equations that remain to be satisfied are of the form

$$\sum_{j \in J} a_{ij} v_j = 0, \qquad i \in I. \tag{3.5}$$

The entries v_j with $j \in J$ are already fixed and only depend on the matrix A', so for sufficiently general A the v_j are completely independent of the entries a_{ij} with $i \in I$. Therefore, if $v_j \neq 0$ and $a_{ij} \neq 0$, the nonzero term $a_{ij} v_j$ cannot be cancelled from equation (3.5). So for v to satisfy $Av = \lambda v$, one must have $v_i \neq 0$ whenever there exists $j \in [n]$ such that $j \to i$ is an edge in G and $v_j \neq 0$.

Now suppose $v_i \neq 0$ and there exists a path $\{i, s_1, \ldots, s_t, j\}$ in G. Then by our previous observation, we have

$$v_i \neq 0 \ \Rightarrow \ v_{s_1} \neq 0 \ \Rightarrow \ \ldots \ \Rightarrow \ v_{s_t} \neq 0 \ \Rightarrow \ v_j \neq 0.$$

□

This lemma implies that v is either all zero or all nonzero on a strongly connected component of G. In other words, the support of v consists of the union of vertex sets of strongly connected components of G. In particular, if G is strongly connected, then v does not have any zero entries.

PROPOSITION 3.2.4. *Let G be strongly connected and $A \in \Theta$ sufficiently general. Then any class $[X] \in V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n\big)$ has a representative $x = (x_{ij}) \in V_A$ whose first row, first column and diagonal are all zero, i.e. $x_{i1} = x_{1i} = x_{ii} = 0$ for all $i \in [n]$.*

PROOF. Let $[X] \in V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n\big)$. First we show that there exists a representative x of $[X]$ whose first row and the diagonal are zero, then we use these facts to show that also the first column must be zero.

We claim that projecting $M \in Z_{\mathfrak{gl}_n}(A)$ onto its first row yields a bijection between $Z_{\mathfrak{gl}_n}(A)$ and the space of matrices with only the first row nonzero. Note that indeed both spaces are n-dimensional. The set of diagonalizable matrices is dense in $\mathbb{C}^{n\times n}$, so a sufficiently general $A \in \Theta_G$ is diagonalizable. Let $A = PDP^{-1}$ be the eigendecomposition of A, the columns of P forming a basis of eigenvectors. If A is diagonalizable, then $MA = AM$ if and only if M is diagonalized by the same basis of eigenvectors as A. Since G is strongly connected, Lemma 3.2.3 implies that P contains no zeros. Hence if M is nonzero, then $M = PD'P^{-1}$ implies that the first row of M has at least one nonzero position. Therefore projection onto the first row is injective and as both spaces have dimension n, it must also be surjective.

Now choose $M \in Z_{\mathfrak{gl}_n}(A)$ such that its first row equals the first row of X and choose a diagonal matrix $D \in \mathcal{D}_n$ whose diagonal equals the diagonal of $X - M$. Then $M + D \in Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n$ and $[X] = [X - (M + D)] \in V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n\big)$, hence $x = X - (M + D)$ is a representative of $[X]$ satisfying $x_{1i} = x_{ii} = 0$ for all $i \in [n]$. Moreover, if x is an element of V_A of this form, then it gives rise to a nonzero class $[x] \in V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n\big)$.

What remains to be shown, is that for $X \in V_A$ which has its first row and diagonal all zero, also the first column of X must be zero. Write both X and A as block matrices in the following way:

$$X = \left[\begin{array}{c|c} 0 & 0^T \\ \hline x_1 & X_1 \end{array}\right] \qquad A = \left[\begin{array}{c|c} a_{11} & a_1^T \\ \hline a_2 & A_1 \end{array}\right]$$

Here x_1, a_1 and a_2 are vectors in $\mathbb{C}^{n-1}$ and X_1, A_1 are matrices in $\mathbb{C}^{(n-1)\times(n-1)}$. Multiplying these matrices to obtain $XA - AX$, we see that

$$[X, A]_1 = X_1A_1 - A_1X_1 + x_1a_1^T.$$

For X to lie in V_A there must exist $Y \in \mathfrak{gl}_{n-1}$ such that $[X, A]_1 = [Y, A_1]$, so we obtain

$$x_1a_1^T = (Y - X_1)A_1 - A_1(Y - X_1) = [Y - X_1, A_1].$$

We need to show that for sufficiently general A this implies $x_1 = 0$, i.e.

$$\big\{x_1a_1^T \mid x_1 \in \mathbb{C}^{n-1}\big\} \cap [\mathfrak{gl}_{n-1}, A_1] = \{0\}.$$

Observe that the first space has dimension $n-1$, while the dimension of the second space equals $\dim \mathfrak{gl}_{n-1} - \dim Z_{\mathfrak{gl}_{n-1}}(A_1) = (n-1)^2 - (n-1)$. This suggests that their intersection is indeed trivial.

Let $B = x_1a_1^T \in [\mathfrak{gl}_{n-1}, A_1]$ en let $v_1^T, \ldots, v_{n-1}^T$ the row eigenvectors of A_1, where $v_1, \ldots, v_{n-1} \in \mathbb{C}^{n-1}$. We know that a_1 has at least one nonzero position, otherwise G would not be strongly connected. Therefore $B = 0$ implies $x_1 = 0$, in which case we are done. From now on, suppose B is not the zero matrix. Then

$$v_i^TB \in \bigoplus_{j\neq i} \mathbb{C}v_j^T \qquad \text{for } i = 1, \ldots, n-1$$

which can be seen from the following. Suppose $B = [C, A_1]$ for some $C \in \mathfrak{gl}_{n-1}$, then

$$v_i^TB = v_i^TCA_1 - \lambda_iv_i^TC = v_i^T(CA_1 - \lambda_iC) = v_i^TC(A_1 - \lambda_iI).$$

For sufficiently general A, the eigenvectors $v_1^T, \ldots, v_{n-1}^T$ of A_1 are linearly independent, hence they form a basis of $\mathbb{C}^{n-1}$. We can write $v_i^TC = \sum_{j=1}^{n-1} \alpha_jv_j^T$ with coefficients $\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{C}$. Observe that $v_i^T(A_1 - \lambda_iI) = 0$, and therefore

$$v_i^TB = \sum_{j=1}^{n-1} \alpha_jv_j^T(A_1 - \lambda_iI) = \sum_{j\neq i} \alpha_j(\lambda_j - \lambda_i)v_j^T \in \bigoplus_{j\neq i} \mathbb{C}v_j^T.$$

Since B has rank one, the row space of B is one-dimensional. Say it is spanned by a vector $w \in \mathbb{C}^{n-1}$, then for each left eigenvector v_i^T of B either $v_i^T B = 0$ or $v_i^T B = cw^T$ for some $c \in \mathbb{C}$. We know that the eigenvectors $v_1^T, \ldots, v_{n-1}^T$ of A_1 are linearly independent, so w can be written as a linear combination of these vectors:

$$w = \sum_{i=1}^{n-1} c_i v_i, \qquad c_1, \ldots, c_{n-1} \in \mathbb{C}^{n-1}.$$

From the observation that $v_i^T B \in \bigoplus_{j \neq i} \mathbb{C} v_j^T$, it follows that if $v_i^T B \neq 0$, then $c_i = 0$. Hence there exists a proper partition $I \sqcup J = [n-1]$ $(I, J \neq \emptyset)$ such that $w = \sum_{j \in J} c_j v_j$ $(c_j \in \mathbb{C}^*)$ and $v_j B = 0$ for all $j \in J$. Since $B = x_1 a_1^T$, it follows that a_1^T must be a scalar multiple of w, i.e. it is a linear combination of the v_j, $j \in J$. Let P be the matrix whose rows are $v_1^T, \ldots, v_{n-1}^T$, then we can write $a_1^T = c^T P$ for some $c \in \mathbb{C}^{n-1}$. If we can show that $c^T = a_1^T P^{-1}$ has no zero entries, we have arrived at a contradiction, because this would imply $I = \emptyset$ while we assumed I to be nonempty.

Note that P^{-1} is the matrix whose columns are the column eigenvectors of A_1. If the graph corresponding to A_1 is strongly connected, we know from Lemma 3.2.3 that P^{-1} contains no zeros. For sufficiently general A, the first row vector a_1 is independent of A_1 and hence of P and P^{-1}. Therefore, if P^{-1} contains no zeros and a_1 has at least one nonzero position, $a_1^T P^{-1}$ has no zeros.

If the graph G_1 corresponding to A_1 is not strongly connected, let $C_1, \ldots, C_l$ be its strongly connected components. Let u be a column of P^{-1}, then u is a column eigenvector of A_1. From Lemma 3.2.3 it follows that for each component C_i, the entries of u corresponding to the vertices of C_i are either all zero or all nonzero. The eigenvector u must be nonzero on at least one component C_i, and if this component has an edge to vertex 1 in the original graph G, then $a_1^T u \neq 0$ since u does not depend on a_1. If C_i does not have an edge to vertex 1 in G, there must be a path in G_1 from C_i to some component C_j which does have an edge to 1 in G, as G is strongly connected. But from Lemma 3.2.3 it follows that for every vertex k on this path we have $u_k \neq 0$. In particular, all entries of u corresponding to vertices of C_j are nonzero, and since C_j has an edge to 1 in G, again we obtain $a_1^T u \neq 0$. Hence $c^T = a_1^T P^{-1}$ has no zeros, which contradicts the fact that $|J| < n-1$. This shows that $x_1 = 0$ and therefore the first column of X must be zero. □

This proposition implies that when looking for $X \in V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n\big)$, it suffices to search for X whose first row, first column and diagonal are all zero. By definition, V_A contains all $X \in \mathfrak{gl}_n$ for which $[X, A]$ lies in the kernel of the differential $d_A c$. From Proposition 3.1.2 we know that this implies that $[X, A] \in \Theta_G$ *and* there must exist $Y \in \mathfrak{gl}_{n-1}$ such that $[Y, A_1] = [X, A]_1$. For $X \in \mathfrak{gl}_n$ whose first row, first column and diagonal are all zero, we can just take $Y = X_1$, since

$$\left[\begin{array}{c|c} 0 & 0^T \\ \hline 0 & X_1 \end{array}\right] \left[\begin{array}{c|c} a_{11} & a_1^T \\ \hline a_2 & A_1 \end{array}\right] - \left[\begin{array}{c|c} a_{11} & a_1^T \\ \hline a_2 & A_1 \end{array}\right] \left[\begin{array}{c|c} 0 & 0^T \\ \hline 0 & X_1 \end{array}\right] = \left[\begin{array}{c|c} 0 & a_1^T X_1 \\ \hline X_1 a_2 & [X_1, A_1] \end{array}\right]$$

Hence for X of this form we have

$$X \in V_A \quad \Leftrightarrow \quad [X, A] \in \Theta_G.$$

Combining this observation with Corollary 3.2.2 and Proposition 3.2.4, we obtain the following corollary.

Corollary 3.2.5. *G has the expected dimension if and only if, for sufficiently general $A \in \Theta_G$, there does not exist $X \in \mathfrak{gl}_n$ of the form $X_{1i} = X_{i1} = X_{ii} = 0$ for all $i \in [n]$, $X \neq 0$, such that the commutator $[X, A]$ lies in the parameter space Θ_G.*

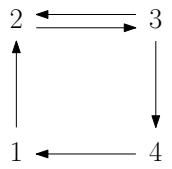


FIGURE 3.1. Example 3.3.2

Thus, to determine whether a graph has the expected dimension, we need to check whether there exists $X \in \mathfrak{gl}_n$ satisfying the properties of Corollary 3.2.5.

3.3. A necessary condition to have the expected dimension

In this section we derive some sufficient conditions on a graph G, such that there exists $X \in \mathfrak{gl}_n$ satisfying Corollary 3.2.5. This implies that G has the expected dimension, i.e. an identifiable scaling reparametrization of the model exists. These sufficient conditions for G *not* to have the expected dimension will lead us to a necessary condition for G to have the expected dimension.

CONDITION 3.3.1. There exists an ordered pair (i, j) with $i, j \in \{2, \dots, n\}$, $i \neq j$, such that the support of the j-th row is contained in the support of the i-th row of A and the support of the i-th column is contained in the support of the j-th column of A.

For a strongly connected graph G, the matrix $A = A(G)$ satisfies the above condition whenever there exist vertices $i, j \neq 1$ such that for all $k \in [n]$ the following holds: for any edge $k \to j$ there is also an edge $k \to i$ and for any edge $i \to k$ there is also an edge $j \to k$. Also the nonzero entries a_{ii} and a_{jj} of A should be taken into account, which implies that both a_{ij} and a_{ji} are nonzero, i.e. i and j form a 2-cycle in G.

EXAMPLE 3.3.2. Consider the graph G in Figure 3.1 and its parameter matrix

$$A(G) = \begin{bmatrix} a_{11} & 0 & 0 & a_{14} \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & 0 \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix}$$

Observe that the pair $(2, 3)$ satisfies Condition 3.3.1. Let $X = E_{23}$ be the matrix with a 1 at position $(2, 3)$ and zeros elsewhere, then $[X, A]$ has the correct zero pattern:

$$[X, A] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} - a_{22} & 0 \\ 0 & 0 & -a_{32} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

This shows that E_{23} represents a nontrivial element of $V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + D_n\big)$ and hence G does not have the expected dimension.

As illustrated by this example, Condition 3.3.1 is a sufficient condition for $V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + D_n\big)$ to be nontrivial. If this condition holds for sufficiently general $A \in \Theta_G$, then G cannot have the expected dimension. This will be shown in the next lemma.

LEMMA 3.3.3. *Let G be a strongly connected graph such that the pair (i, j) satisfies Condition 3.3.1 with $A \in \Theta_G$. Then the matrix $X = E_{ij} \in \mathfrak{gl}_n$ yields a nontrivial class $[X] \in V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + D_n\big)$.*

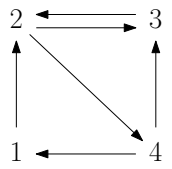


FIGURE 3.2. Example 3.3.5

PROOF. From $i, j \in \{2, \ldots, n\}$ and $i \neq j$, it follows that X is of the correct form: it has its first row, first column and diagonal all zero. According to Corollary 3.2.5 we only need to show that $[X, A] \in \Theta_G$. Consider the two terms of the Lie bracket: XA has its i-th row equal to the j-th row of A and zeros elsewhere, while AX has its j-th column equal to the i-th column of A and zeros elsewhere. Clearly A itself must have the correct zero pattern, so from Condition 3.3.1 we immediately see that $XA - AX$ must be in Θ_G. □

Each $[X] \in V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + D_n\big)$ spans a one-dimensional subspace, since $[cX]$ is an element of $V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + D_n\big)$ for any $c \in \mathbb{C}$.

However, Condition 3.3.1 is not a necessary condition. Figure 3.2 gives an example of a graph which does not satisfy Condition 3.3.1 and it does not have the expected dimension either. This case is covered by the following more general condition:

CONDITION 3.3.4. There exists a set S containing k distinct pairs (i, j) that satisfy Condition 3.3.1, except for at most $k - 1$ positions of A where the supports are not contained in one another. These positions are zero while they should be nonzero for A to satisfy Condition 3.3.1 and will be called *exceptions*.

EXAMPLE 3.3.5. Consider the graph G in Figure 3.2 and its parameter matrix

$$A(G) = \begin{bmatrix} a_{11} & 0 & 0 & a_{14} \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & a_{42} & 0 & a_{44} \end{bmatrix}$$

Then the set $S = \{(2, 4), (3, 4)\}$ satisfies Condition 3.3.4, since only one exception occurs. Indeed, let

$$X = x_{24}E_{24} + x_{34}E_{34}$$

then $[X, A]$ has the correct zero pattern except for position $(2, 4)$:

$$[X, A] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & a_{42}x_{24} & 0 & (a_{44} - a_{22})x_{24} - a_{23}x_{34} \\ 0 & a_{42}x_{34} & 0 & (a_{44} - a_{33})x_{34} - a_{32}x_{24} \\ 0 & 0 & 0 & -a_{42}x_{24} \end{bmatrix}$$

In order to have $[X, A] \in \Theta_G$ we must solve

$$(a_{44} - a_{22})x_{24} - a_{23}x_{34} = 0$$

which yields

$$x_{34} = \frac{a_{44} - a_{22}}{a_{23}} x_{24}.$$

Thus, we have found a one-dimensional subspace of $V_A \,/\, \big(Z_{\mathfrak{gl}_n}(A) + D_n\big)$.

In general, if $A \in \Theta_G$ satisfies Condition 3.3.4 for some set S, define

$$X = \sum_{(i,j) \in S} x_{ij} E_{ij}.$$

Each exception (as defined in Condition 3.3.4) corresponds to a position in $[X, A]$ which is nonzero while it should be zero in order to have the correct zero pattern. Each of these positions contains a linear expression in x_{ij} and a_{ij}, and to ensure that these entries become zero we must solve a system of linear equations. If Condition 3.3.4 holds, we obtain a system of $k-1$ equations in k variables. Setting $x_{ij} = 0$ for all $i, j \in [n]$ certainly is a solution, so the system is solvable and hence there must be a solution space of dimension (at least) 1. In the next section, we will explicitly construct this system of equations.

PROPOSITION 3.3.6. *Let G be a strongly connected graph and S a set such that Condition 3.3.4 holds for sufficiently general $A \in \Theta_G$. Then there exists a nonzero matrix*

$$X = \sum_{(i,j)\in S} x_{ij} E_{ij} \in \mathfrak{gl}_n$$

such that $[X, A] \in \Theta_G$.

In other words, if Condition 3.3.4 holds for sufficiently general $A \in \Theta_G$, then G does not to have the expected dimension. One can use this condition to conclude directly from the parameter matrix $A(G)$ that G does not have the expected dimension, without having to do any calculations.

In conclusion, we have seen that a necessary condition for G to have the expected dimension is that for sufficiently general $A \in \Theta_G$ Condition 3.3.4 does *not* hold. In the next section this will be translated to a criterion on a bipartite graph to decide whether a given graph has the expected dimension.

3.4. A rank criterion based on bipartite graphs

Recall that a bipartite graph consists of two disjoint sets of vertices L, R, such that every edge connects a vertex in L to a vertex in R. A matching $M \subseteq E$ is a subset of the edges, such that no two edges are incident to the same vertex. The matching is called *perfect* if every vertex in the graph is incident to some edge in M. When $|R| \neq |L|$ no perfect matching exists, but a matching is called *L-saturating* if it is perfect with respect to L.

For a given (directed) graph G on n vertices we define a bipartite (undirected) graph $H(G) = (L \cup R, E)$ as follows:

$$\begin{aligned} L &= \{(i,j) \mid i,j \in \{2,\dots,n\} \text{ and } i \neq j\} \\ R &= \{(k,l) \mid k,l \in [n],\ k \neq l \text{ and } l \to k \text{ is not an edge of } G\} \\ E &= \{((i,j),(k,l)) \mid [E_{ij}, A] \text{ is nonzero on position } (k,l)\}. \end{aligned} \tag{3.6}$$

Note that L corresponds to all positions of $X \in V_A$ that are outside the first row, first column and the diagonal. Also note that R corresponds to all zero positions of $A(G)$, hence all positions of $[X, A]$ where an exception might occur. Furthermore, there is an edge between $(i,j) \in L$ and $(k,l) \in R$ if and only if E_{ij} causes an exception at position (k,l). Now recall Hall's marriage theorem:

Hall's Theorem. *Let $G = (L \cup R, E)$ be a bipartite graph, then there exists an L-saturating matching in G if and only if for every subset $S \subseteq L$ we have*

$$|S| \leq |N(S)|$$

where $N(S)$ denotes the set of neighbors of S in G.

Observe that Condition 3.3.4 is satisfied if and only if there exists a subset $S \subseteq L$ such that $|N(S)| < |S|$. But if this is the case, then Hall's theorem implies that there cannot be an L-saturating matching in $H(G)$.

PROPOSITION 3.4.1. *Let $G = (V, E)$ be a graph satisfying Assumptions 1.1.2-1.1.4 and let $H = H(G)$ be the corresponding bipartite graph. If G has an identifiable scaling reparametrization, then there exists an L-saturating matching in H.*

This shows that an L-saturating matching in $H(G)$ is a necessary condition for G to have the expected dimension. But is it also a sufficient condition?
Let $X \in \mathfrak{gl}_n$ have its first row, first column and diagonal all zero and write

$$X = \sum_{(i,j)\in L} x_{ij} E_{ij}.$$

As discussed in the previous section, the constraint $[X, A] \in \Theta_G$ gives rise to a system of linear equations in the entries of X and A. To see what this expression looks like, consider the two terms of the Lie bracket $[E_{ij}, A]$. The product $E_{ij}A$ has its i-th row equal to the j-th row of A and zeros elsewhere, while AE_{ij} has its j-th column equal to the i-th column of A and zeros elsewhere. Hence E_{ij} adds a nonzero term to position (k, l) of $[X, A]$ only in the following three cases:

$$\begin{aligned} i = k,\ j \neq l \text{ and } l \to j \in G &\quad\rightsquigarrow\quad -a_{jl} \\ i \neq k,\ j = l \text{ and } i \to k \in G &\quad\rightsquigarrow\quad a_{ki} \\ i = k \text{ and } j = l &\quad\rightsquigarrow\quad a_{kk} - a_{ll} \end{aligned}$$

Define the matrix $B(G) \in \mathbb{C}^{|R|\times|L|}$ as

$$B(G)_{(k,l),(i,j)} = \begin{cases} -a_{jl} & \text{if } i = k,\ j \neq l \text{ and } l \to j \in E \\ a_{ki} & \text{if } i \neq k,\ j = l \text{ and } i \to k \in E \\ a_{kk} - a_{ll} & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

From our previous observations, it follows that $B(G)$ is the coefficient matrix corresponding to the system of equations obtained from $[X, A] \in \Theta_G$. On the other hand, if we define edge weights for the edges in the bipartite graph $H(G)$ by

$$w((i, j), (k, l)) = [E_{ij}, A]_{kl}$$

with $(i, j) \in L$ and $(k, l) \in R$, then $B(G)$ is the weighted bi-adjacency matrix corresponding to $H(G)$.

Let $x \in \mathbb{C}^{|L|}$ be the vector of coefficients x_{ij}, $(i, j) \in L$, then the linear system corresponding to $[X, A] \in \Theta_G$ is given by

$$B(G)x = 0.$$

It follows that each solution $x \in \mathbb{C}^{|L|}$ gives rise to a class $[X] \in V_A/(Z_{\mathfrak{gl}_n}(A) + \mathcal{D}_n)$ and vice versa. Furthermore, $x = 0$ if and only if $X = 0$. Combining these observations with the fact that $B(G)$ has a nontrivial kernel if and only if its rank is less than $|L|$, we obtain the following theorem.

THEOREM 1.4.2. *Let G be a graph satisfying Assumptions 1.1.2-1.1.4 and let $L \cup R$ be the vertex set of the bipartite graph $H(G)$. Then G has the expected dimension if and only if the $|R| \times |L|$ bi-adjacency matrix $B = B(G)$ has full column rank, i.e.* $\operatorname{rk}(B) = |L|$.

If $|R| < |L|$, then $B(G)$ certainly has rank smaller than $|L|$. However, this implies that the number of zero positions of A is less than $(n-1)(n-2)$. Since the number of zero positions in A equals $n^2 - (n + m)$, we obtain

$$n^2 - (n + m) < (n - 1)(n - 2)$$

and hence $m > 2(n - 1)$. This is equivalent to Lemma 2.3.3, which stated that if $m > 2n - 2$ then G does not have the expected dimension.

Now what can we say about the existence of an L-saturating matching in $H(G)$ and the rank of $B(G)$? First of all, $B(G)$ has rank $|L|$ if and only if there exists some $|L| \times |L|$ submatrix B' which has nonzero determinant. This submatrix corresponds to an induced subgraph H' of $H(G)$, containing L entirely but only a subset of R of size $|L|$. The determinant of B' can be computed as a sum over all permutations σ in $S_{|L|}$:

$$\det(B') = \sum_{\sigma \in S_{|L|}} \operatorname{sgn}(\sigma) \prod_{i=1}^{|L|} (B')_{i,\sigma_i}.$$

Each nonzero term in this sum corresponds to a perfect matching in H'. Therefore, if $\det(B') \neq 0$ then there exists a perfect matching in H' and hence an L-saturating matching in $H(G)$. This is exactly what Proposition 3.4.1 says, but does the converse hold as well? In other words, if there exists an L-saturating matching in $H(G)$, will the corresponding submatrix B' of $B(H)$ have nonzero determinant?

If all entries of B' are linearly independent, then no two matchings can yield the same term in the determinant, hence a perfect matching in H' exists if and only if B' is nonsingular. However, this is not the case: each parameter may appear at multiple positions of B'. Indeed, Example 3.4.2 shows that although an L-saturating matching in H exists, the matrix $B(H)$ does not have rank $|L|$. Calculations in MATHEMATICA have shown that these cases are rare: there is no graph on four vertices which has a perfect matching but not the expected dimension, and for graphs on five vertices there are only ten such cases.

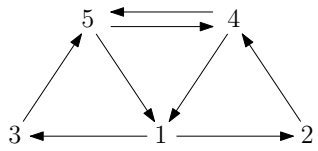
EXAMPLE 3.4.2. Let G be the graph given in Figure 3.3a with its corresponding bipartite graph $H(G)$ in Figure 3.3b, where the red edges represent an L-saturating matching M. Note that $|L| = |R|$, so this is actually a perfect matching. However, M is not the only perfect matching. The matrix $B(G)$ is of the following form: One can check that the last six columns of this matrix are linearly dependent, hence $B(G)$ has rank 11 while $|L| = 12$. In other words, G does not have the expected dimension.

3.5. Complexity analysis

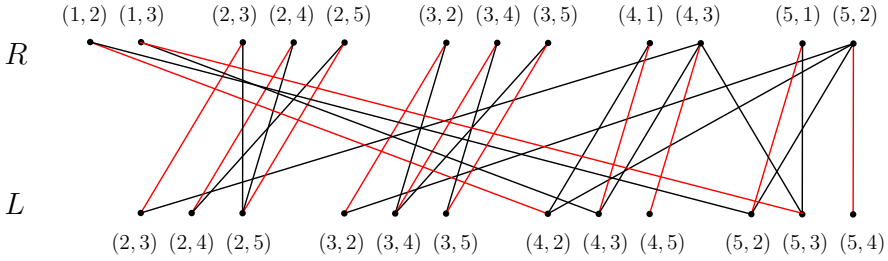
In this section we consider the theoretical complexity of determining whether a graph has the expected dimension or not. Two algorithms are compared: the first based on the dimension criterion from [**MS14**], given in Theorem 2.4.1, and the second based on the main result of this chapter, Theorem 1.4.2.

Recall that the dimension criterion states that G has the expected dimension if and only if the dimension of the image of c equals $m + 1$. This is equivalent to the rank of the differential $d_A c$ being equal to $m + 1$ for sufficiently general $A \in \Theta_G$. Thus, a naive way to check whether G has the expected dimension would be to construct $A = A(G)$ and compute the characteristic polynomials of A and A_1 symbolically in order to obtain the double characteristic polynomial map c. Then we need to differentiate the coordinates of c with respect to the parameters a_{ij} to obtain the differential (or Jacobian) $d_A c$. Finally, the generic rank of $d_A c$ has to be computed, i.e. the rank of $d_A c$ in sufficiently general A.

This last step can be performed in polynomial time using a randomized algorithm that substitutes random parameter values from a finite set S. It brings the risk of decreasing the rank of the input matrix, but the rank cannot increase. In other words, substituting random parameter values yields a true-biased Monte-Carlo algorithm: a positive output is always correct, but a negative output may be false with a certain probability. However, the probability of this event can be bounded from above using the Schwartz-Zippel Lemma [**Sch80, Zip79**]:



(A) G



(B) $H(G)$

	(2,3)	(2,4)	(2,5)	(3,2)	(3,4)	(3,5)	(4,2)	(4,3)	(4,5)	(5,2)	(5,3)	(5,4)
(1,2)	0	0	0	0	0	0	a_{14}	0	0	a_{15}	0	0
(1,3)	0	0	0	0	0	0	0	a_{14}	0	0	a_{15}	0
(2,3)	$a_{22}-a_{33}$	0	$-a_{53}$	0	0	0	0	0	0	0	0	0
(2,4)	0	$a_{22}-a_{44}$	$-a_{54}$	0	0	0	0	0	0	0	0	0
(2,5)	0	$-a_{45}$	$a_{22}-a_{55}$	0	0	0	0	0	0	0	0	0
(3,2)	0	0	0	$a_{33}-a_{22}$	$-a_{42}$	0	0	0	0	0	0	0
(3,4)	0	0	0	0	$a_{33}-a_{44}$	$-a_{54}$	0	0	0	0	0	0
(3,5)	0	0	0	0	$-a_{45}$	$a_{33}-a_{55}$	0	0	0	0	0	0
(4,1)	0	0	0	0	0	0	$-a_{21}$	$-a_{31}$	0	0	0	0
(4,3)	a_{42}	0	0	0	0	0	0	$a_{44}-a_{33}$	a_{53}	0	a_{45}	0
(5,1)	0	0	0	0	0	0	0	0	0	$-a_{21}$	$-a_{31}$	0
(5,2)	0	0	0	a_{53}	0	0	a_{54}	0	0	$a_{55}-a_{22}$	0	$-a_{42}$

(C) $B(G)$

FIGURE 3.3. Example 3.4.2

LEMMA 3.5.1 (Schwartz, Zippel). *Let $f \in \mathbb{K}[x_1, \ldots, x_n]$ be a non-zero polynomial of degree $d \geq 0$ over a field $\mathbb{K}$. Let S be a finite subset of $\mathbb{K}$ and choose $r_1, \ldots, r_n$ at random independently and uniformly from S. Then*

$$\Pr[f(r_1, \ldots, r_n) = 0] \leq \frac{d}{|S|}.$$

By taking the domain S large enough, we can make the probability of a false negative as small as desired. This bound depends on the degree of the polynomial being evaluated. In order to apply the Schwartz-Zippel lemma, we must express the event that $d_A c$ has rank $m+1$ as a certain polynomial being nonzero. As described in [**CKL13**], we compress the matrix into a $(m+1) \times (m+1)$ matrix by introducing new variables and taking linear combinations of the rows and columns of $d_A c$ in terms of these new variables. Then the generic rank of D' equals the generic rank of $d_A c$, so we can apply the Schwartz-Zippel lemma to the function $\det(D')$. Note that the entries of D' are polynomials of degree $O(n)$, hence the determinant of D' has degree $O(m)O(n) = O(n^2)$, since we assume $n \leq m \leq 2n-2$. In practice, it is not necessary to perform this step, but it shows that we can indeed apply the Schwartz-Zippel lemma.

Although the rank can be determined efficiently, computing the characteristic polynomials of A and A_1 symbolically is very expensive. Therefore, we derive a

more efficient way to evaluate the dimension criterion. We would like to substitute the random parameter values at an earlier point in the algorithm, to reduce the number of symbolic computations and hence reduce the computation costs.

Observe that for an $n \times n$ matrix B whose entries are polynomials in a_{ij}, $i, j \in [n]$, the following equality holds:

$$\frac{\partial \det(B)}{\partial a_{ij}} = \sum_{k=1}^{n} \det(B_k) \tag{3.8}$$

where B_k is the matrix obtained from B by differentiating its k-th column with respect to the parameter a_{ij}. This equality follows from the product rule for differentiation and the fact that each term in the determinant is the product of entries of B from different columns. Since the characteristic polynomials of A and A_1 are equal to $\det(tI_n - A)$ and $\det(tI_{n-1} - A_1)$, equation (3.8) allows us to differentiate terms of $tI_n - A$ instead of its determinant. When differentiating with respect to a_{ij}, the only position of $tI_n - A$ that remains nonzero is position (i, j). For $tI_{n-1} - A_1$ the same holds, except when $i = 1$ or $j = 1$, in which case differentiating with respect to a_{ij} always yields a zero. Let $B = tI_n - A$, $B_1 = tI_{n-1} - A_1$ and let $B(i, j)$ denote the matrix obtained from B by removing the i-th row and the j-th column. Then

$$\frac{\partial \det(B)}{\partial a_{ij}} = -\det(B(i,j))$$

$$\frac{\partial \det(B_1)}{\partial a_{ij}} = \begin{cases} -\det(B_1(i,j)) & \text{if } i, j \neq 1 \\ 0 & \text{otherwise.} \end{cases}$$

The right hand side expressions do not have to be differentiated anymore, so the random parameter values can be substituted before calculating the determinants of $B(i, j)$. However, $B(i, j)$ still contains the variable t. We can solve this by calculating $\det(B(i, j))$ for $t = 1, 2, \ldots, n$ and then find the interpolating polynomial of this set of points. Since $\det(B(i, j))$ is a polynomial of degree $n-1$, it is uniquely determined by its value in n different points. Similarly, the determinant of $B_1(i, j)$ is a polynomial of degree $n - 2$ and hence it is uniquely determined by its value in $n - 1$ different points.

These observations lead to Algorithm 3.1, which allows us to compute the differential map $d_A c$ in a random point $A^* \in \Theta_G$ without doing symbolic determinant calculations.

We will compare Algorithm 3.1 with Algorithm 3.2, which is based on Theorem 1.4.2: it constructs the matrix $B(G)$ and determines its rank for sufficiently general A. This algorithm avoids the costly procedure of computing the characteristic polynomials; its most expensive step is computing the determinant of $B(G)$. Again, the most efficient way to do this is using a randomized algorithm, thus we substitute random parameter values from a finite domain S into $B(G)$.

Note that the matrix $B(G)$ need not be square, but similar to the case of $d_A c$, we can compress $B(G)$ into a matrix B' of size $(n-1)(n-2) \times (n-1)(n-2)$ such that B' has nonzero determinant if and only if $B(G)$ has full rank. Then we can apply the Schwartz-Zippel lemma to the determinant of B'. The entries of B' are polynomials of degree $O(1)$, so its determinant has degree $O(n^2)$. This is (asymptotically) the same degree as for the previous algorithm, hence using the same set S yields (asymptotically) the same bound for the probability of a false negative.

The computation times of the two algorithms have been compared using the computer algebra package MATHEMATICA, the results of which are presented in

Algorithm 3.1 Jacobian rank test

Input: Graph $G = (V, E)$, finite set S
Output: TRUE if G has the expected dimension, FALSE otherwise

$(n, m) \leftarrow (|V|, |E|)$
$A \leftarrow A(G)$
Construct A_1 by deleting the first row and column of A
$B_0 \leftarrow tI_n - A$
$B_1 \leftarrow tI_{n-1} - A_1$
Choose A^* uniformly at random from S
for $l = 0, 1$ **do**
    $M \leftarrow B_l$
    $n' \leftarrow n - l$
    for (i, j) with $j \to i \in E$ or $i = j$ **do**
        for $t^* = 1, 2, \ldots, n'$ **do**
            Substitute $t = t^*$ into M
            if a_{ij} appears in M **then**
                Construct $M(i, j)$ from M by removing row i and column j
            else
                $M(i, j) \leftarrow 0$
            end if
            Substitute the parameter values A^* into $M(i, j)$
            $x_{t^*} \leftarrow \det(M(i, j))$
        end for
        Let $p_{ij}(t)$ be the interpolating polynomial through $x_1, \ldots, x_{n'}$
        $c_{ij} \leftarrow$ coefficient vector of $1, t, \ldots, t^{n'-1}$ in $p_{ij}(t)$
    end for
    $d_{A^*} c_l \leftarrow [c_{11}| \cdots |c_{ij}| \cdots |c_{n'n'}]^T$
end for
$d_{A^*} c \leftarrow [d_{A^*} c_0 \,|\, d_{A^*} c_1]^T$
if $\mathrm{rk}(d_{A^*} c) = m + 1$ **then**
    return TRUE
else
    return FALSE
end if

Table 3.1. These are average computation times (seconds per graph) of 50 computations on random strongly connected graphs, with a fixed amount of vertices and edges. Both implementations use the same set $S = \{z \in \mathbb{Z} | -1000 \leq z \leq 1000\}$ and the rank of $d_A c$ and $B(G)$ is evaluated in two random points in S. Thus, the probability of a false negative has an asymptotic upper bound of $\frac{1}{4} O(n^4) \cdot 10^{-6}$.

A theoretical upper bound on the computational complexity of Algorithm 3.1 is determined by calculating $\det(B(i, j))$ for $t = 1, 2, \ldots, n$. This can be done using Gaussian elimination taking $O(n^3)$ operations. Computing the interpolating polynomial can also be done using Gaussian elimination and takes another $O(n^3)$ operations. We obtain a total estimate of the number of operations of

$$O((n + m)(n^3 \cdot n + n^3)) = O(n^5),$$

which is a theoretical upper bound for computing the rank of the Jacobian. This algorithm could be optimized by using more efficient techniques than interpolation to calculate the symbolic determinant.

For Algorithm 3.2 the computational complexity is determined by computing the rank of $B(G)$ at a random point A^*. This matrix has $n^2 - (n + m)$ rows and

Algorithm 3.2 Rank test of $B(G)$

Input: Graph $G = (V, E)$, finite set S
Output: TRUE if G has the expected dimension, FALSE otherwise

1: Construct the sets L, R defined by

$$L = \{(i, j) \mid i, j \in \{2, \ldots, n\} \text{ and } i \neq j\}$$
$$R = \{(k, l) \mid k, l \in [n],\ k \neq l \text{ and } l \to k \notin E\}$$

2: Construct the symbolic $|R| \times |L|$ matrix $B(G)$ with entries

$$B(G)_{(k,l),(i,j)} = \begin{cases} -a_{jl} & \text{if } i = k,\ j \neq l \text{ and } l \to j \in E \\ a_{ki} & \text{if } i \neq k,\ j = l \text{ and } i \to k \in E \\ a_{kk} - a_{ll} & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise.} \end{cases}$$

3: Choose A^* uniformly at random from S
4: Calculate the generic rank of $B(G)$ by substituting A^* into $B(G)$
5: **if** $\mathrm{rk}(B(G)) = |L|$ **then**
6: **return** TRUE
7: **else**
8: **return** FALSE
9: **end if**

$(n-1)(n-2)$ columns, hence calculating its rank can be done in $O(n^6)$ operations using Gaussian elimination. Note that the matrix $B(G)$ is much larger than $d_A c$, but also much sparser; one could use this sparseness to improve Algorithm 3.2.

These asymptotic complexities imply that for n large enough, Algorithm 3.2 is less efficient than Algorithm 3.1. However, Table 3.1 shows that this is not the case for small values of n. This might be because MATHEMATICA recognizes $B(G)$ as a sparse matrix and computes its rank more efficiently than using just Gaussian elimination.

(n, m)	Algorithm 3.1	Algorithm 3.2
(5,8)	0.024	0.002
(10,18)	0.484	0.070
(15,28)	3.747	0.540
(20,38)	17.323	2.575
(25,48)	57.880	8.549

TABLE 3.1. Computation times (seconds per graph)

CHAPTER 4

Properties and constructions

In the previous chapter we have seen how to check efficiently whether a given graph has the expected dimension, i.e. whether there exists an identifiable scaling reparametrization. In this chapter, we consider the question how we can extend a given graph with the expected dimension by adding vertices and edges, such that the resulting graph has the expected dimension as well. Some constructions satisfying this property were already presented in [**MS14**], but using Theorem 1.4.2 we can derive stronger results. First, some properties and constructions will be defined, followed by a short summary of constructions proved by Meshkat and Sullivant. In Section 4.3 we derive two new constructions, using the results of the previous chapter. These constructions allow us to build many graphs with the expected dimension. In Section 4.4 we discuss ear decompositions of a graph. This leads us to a procedure of deleting or subdividing edges of a given graph, such that the resulting graph has the expected dimension. The chapter is concluded with computational results for graphs on four and five vertices.

4.1. Definitions

We are interested in extensions of graphs with the expected dimension, such that the resulting graphs again have the expected dimension. These operations we call valid:

Definition 4.1.1. An operation on a graph G is called *valid* if the resulting graph G' has the expected dimension whenever G has the expected dimension. It is said to be *strongly valid* if the inverse implication holds as well, i.e. G' has the expected dimension if and only if G has the expected dimension.

Before we derive several valid operations, we need some definitions. Given two graphs, one can construct a new graph by taking the union:

Definition 4.1.2. Given $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, define the union $G_1 \cup G_2$ to be the graph with vertex set $V_1 \cup V_2$ and edge set $E_1 \cup E_2$.

Of course, one can also start with a single graph, and add or remove vertices or edges from this graph. Such operations are valid only if the graph satisfies certain properties.

Definition 4.1.3. A graph G is said to have an *exchange with* i if there exists a vertex $i \in \{2, \ldots, n\}$ such that both $1 \to i$ and $i \to 1$ are edges in G. More general, a graph has an *exchange* if there exists $i \in V$ such that G has an exchange with i.

If a graph has an exchange, one of the operations that we can apply is the collapse of two vertices:

Definition 4.1.4. Given a graph $G = (V, E)$ that has an exchange with i, the *collapsed graph* $G' = (V', E')$ is the graph in which vertex 1 and i have been identified, with $V' = V \setminus \{i\}$. An edge $u \to v$ appears in G' if $u \to v$ appears in G, or if $v = 1$ and $u \to i$ is an edge in G, or if $u = 1$ and $i \to v$ is an edge in G.

FIGURE 4.1. Left: Exchange (blue) and line segment (red) in G.
Right: The graph G' obtained by collapsing the exchange.

Figure 4.1 illustrates an exchange with vertex 2 in G and the collapsed graph G'. Observe that if G is strongly connected, then the collapsed graph G' is again strongly connected. When collapsing two arbitrary vertices, it is hard to tell whether the resulting graph will have the expected dimension or not. In some special cases where G has an exchange with i and G' is obtained by collapsing the exchange, i.e. vertices 1 and i are identified, we *can* predict whether or not the collapsed graph will have the expected dimension. This will be discussed in the next sections.

Identifying two vertices reduces the number of vertices by one. Conversely, we can also increase the number of vertices, for example by subdividing an edge.

DEFINITION 4.1.5. Let $G = (V, E)$ be a graph on $n-1$ vertices and let $i \to j$ be an edge in G. The graph $G' = (V', E')$ obtained by *subdividing the edge* $i \to j$ has vertex set $V' = V \cup \{n\}$ and edges $E' = (E \setminus \{i \to j\}) \cup \{i \to n, n \to j\}$.

Another way to increase the number of vertices is by adding a line segment to G: choose two vertices k, l of G, add new vertices $n_1, \ldots, n_s$, and add the edges of the path $\{k, n_1, n_2, \ldots, n_s, l\}$. This is called a line segment, as defined below.

DEFINITION 4.1.6. A *line segment* of length $k \geq 2$ in a graph G is a tuple $(v_0, v_1, \ldots, v_k)$ such that $v_0 \to v_1, \ldots, v_{k-1} \to v_k$ are edges in G and these are the only edges incident to $v_1, \ldots, v_{k-1}$.

Note that given an edge $i \to j$ in G, subdiving this edge creates a line segment of length two, since the new vertex n is incident to i and j but no other vertices. Figure 4.1 illustrates a line segment of length two in the graph G.

DEFINITION 4.1.7 ([**MS14**], Def. 5.6). A *chain of cycles* in G is a sequence of directed cycles $C_1, C_2, \ldots, C_k$, such that for $i = 1, \ldots, k$ the cycle C_i is attached to C_{i+1} by joining at a vertex.

Other properties that we shall be using involve specific forms of strongly connectedness:

DEFINITION 4.1.8. A graph $G = (V, E)$ is *minimally strongly connected* if it is strongly connected and for each edge $e \in E$ the graph $(V, E \setminus \{e\})$ is no longer strongly connected. G is said to be *inductively strongly connected* if there exists some ordering of vertices of the n vertices, say $1, \ldots, n$, such that for each $i \in [n]$ the induced subgraph $G_{\{1,\ldots,i\}}$ containing vertices $1, \ldots, i$ is strongly connected.

In Section 2.3 we have seen that if G has the expected dimension, then the number of edges is at most $2n - 2$.

DEFINITION 4.1.9. A graph G is *maximal* if it contains $2n - 2$ edges.

Observe that if G is inductively strongly connected, then it must have at least $2n - 2$ edges. Hence any inductively strongly connected graph that satisfies the bound on the number of edges ($m \leq 2n - 2$) is maximal.

4.2. Earlier results and conjectures

Meshkat and Sullivant have already proven two constructions to be (strongly) valid and derived some properties of graphs with the expected dimension. These results will be considered in this section. The proofs can be found in [**MS14**], and in some cases an alternative proof is given based on the results of Sections 4.3-4.4. Also a conjecture that was formulated by Meshkat and Sullivant will be discussed.

PROPOSITION 4.2.1 ([**MS14**, Prop. 5.3]). *Let G be a strongly connected maximal graph that has the expected dimension. Then G has an exchange.*

If a graph is not maximal, then an exchange is not a necessary condition for a graph to have the expected dimension. For example, a directed cycle has the expected dimension:

LEMMA 4.2.2 ([**MS14**, Prop. 5.4]). *The cycle of length n has the expected dimension for any $n \geq 1$.*

PROOF. This lemma follows immediately from Corollary 4.4.5 and the fact that a cycle is minimally strongly connected. □

The first valid construction that we consider is to add an exchange to a given graph. The proof in [**MS14**] considers the characteristic polynomials of the corresponding parameter matrices, but this proposition is also an immediate consequence of the results in Section 4.3.

PROPOSITION 4.2.3 ([**MS14**, Prop. 5.5]). *Let G be a graph on n vertices and construct G' from G by adding a new vertex $1'$ and an exchange $1 \to 1', 1' \to 1$. Then the resulting graph G' with input-output node $1'$ has the expected dimension if and only if G has the expected dimension.*

PROOF. This follows directly from Proposition 4.3.1 and Lemma 4.2.2, by writing G' as the union of G and a 2-cycle. □

The next proposition shows that adding a line segment of length two is a valid operation, if G has a *chain of cycles* containing both vertex 1 and the line segment.

PROPOSITION 4.2.4 ([**MS14**, Thm. 5.7]). *Let G' be a graph that has the expected dimension with $n-1$ vertices. Let G be a new graph obtained from G' by adding a new vertex n and two edges $k \to n$ and $n \to l$ and such that G has a chain of cycles containing both 1 and n. Then G has the expected dimension.*

PROOF. This is a weaker version of Proposition 4.3.3. □

Recall that an inductively strongly connected graph can be constructed by adding the vertices one by one, while in each step the corresponding subgraph is strongly connected. Combining this fact with Proposition 4.3.3, one can derive the following corollary by induction on the number of vertices.

COROLLARY 4.2.5 ([**MS14**, Thm. 5.13]). *If G is inductively strongly connected with at most $2n-2$ edges, then G has the expected dimension.*

Meshkat and Sullivant have also formulated a conjecture:

CONJECTURE 4.2.6 ([**MS14**, Conj. 6.6]). *Let G be a graph with n vertices, $2n-2$ edges, and an exchange with i. Let the collapsed graph G' be the graph where 1 and i have been identified. If G' has $2n-4$ edges with an exchange, then G has the expected dimension if and only if G' has the expected dimension.*

(A) Graph G (B) Collapsed graph G'

FIGURE 4.2. Counterexample to Conjecture 4.2.6

We have constructed a counterexample, showing that this conjecture certainly does not hold in both directions. Consider the graph G in Figure 4.2a, which is strongly connected, has an exchange, and satisfies $m = 2n-2$. Its parameter matrix is given by

$$A(G) = \begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & a_{26} \\ 0 & a_{32} & a_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} & a_{45} & 0 \\ 0 & 0 & a_{53} & 0 & a_{55} & 0 \\ 0 & 0 & 0 & a_{64} & a_{65} & a_{66} \end{bmatrix}.$$

One can check that G has the expected dimension using MATHEMATICA and the algorithm based on Theorem 1.4.2. After collapsing the exchange with 2, we obtain the graph G' given in Figure 4.2b. This graph has parameter matrix

$$A(G') = \begin{bmatrix} a_{11} & a_{13} & a_{14} & 0 & a_{16} \\ a_{31} & a_{33} & 0 & 0 & 0 \\ 0 & 0 & a_{44} & a_{45} & 0 \\ 0 & a_{53} & 0 & a_{55} & 0 \\ 0 & 0 & a_{64} & a_{65} & a_{66} \end{bmatrix}.$$

We see that G' is again strongly connected, has an exchange and satisfies $m = 2n-4$. However, G' does not have the expected dimension. This follows from the fact that $A(G')$ satisfies Condition 3.3.1: the support of the column corresponding to vertex 6 is contained in the column corresponding to vertex 4, and for the rows vice versa. So the fact that G has the expected dimension does not imply that G' has the expected dimension as well.

The other direction remains a conjecture, although a partial result follows from Proposition 4.4.6.

4.3. New constructions

In this section several new properties and valid constructions of graphs with the expected dimension are derived. All proofs rely on Theorem 1.4.2, so the matrix $B = B(G)$ plays an important role in this section. Recall that the rows of B are indexed by pairs (i, j) corresponding to the zero positions of $A(G)$, and the columns are indexed by pairs (k, l) with $k, l \in [n], k \neq l$ and $k, l \neq 1$. These column indices correspond to the possible nonzero entries of $X \in V_A$ whose first row, first column and diagonal are zero. For $G = (V, E)$ the entries of B are given by

$$B(G)_{(k,l),(i,j)} = \begin{cases} -a_{jl} & \text{if } i = k,\ j \neq l \text{ and } l \to j \in E \\ a_{ki} & \text{if } i \neq k,\ j = l \text{ and } i \to k \in E \\ a_{kk} - a_{ll} & \text{if } i = k \text{ and } j = l \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

We will refer to the entry $B(G)_{(k,l),(i,j)}$ as the entry (or position) indexed by $(k,l),(i,j)$, where (k,l) is the row index and (i,j) the column index. We start with some basic observations on the structure of $B = B(G)$. For an entry to be nonzero, the two pairs representing the row and column index must have at least one coordinate in common. Entries indexed by $(i,\cdot),(\cdot,i)$ or $(\cdot,i),(i,\cdot)$ are zero, since neither the rows nor the columns of B have indices (i,i). Furthermore, the column indices have no coordinate equal to 1, hence a nonzero entry in the row indexed by $(i,1)$ must be of the form a_{j1}. Similarly, a nonzero entry in the row indexed by $(1,i)$ must be of the form a_{1j}. Also note that every entry in a given row or column of B contains a different parameter.

PROPOSITION 4.3.1. *Let G be of the form $(V' \cup V'', E' \cup E'')$ for some graphs $G' = (V', E'), G'' = (V'', E'')$, such that $V' \cap V'' = \{v\}$, $E' \cap E'' = \emptyset$ and $1 \in V'$. Let 1 be the input-output compartment of G and G', while G'' has input-output compartment v. Then G has the expected dimension if both G' and G'' have the expected dimension. Conversely, if G'' does not have the expected dimension, then neither does G.*

PROOF. Let $A = A(G)$, $A' = A(G')$ and $A'' = A(G'')$. The input-output compartment of G'' is vertex v, so if we order the vertices of G' such that the last row and column of A' correspond to vertex v, then A is of the following form:

$$A = \begin{bmatrix} A' & 0 \\ 0 & A'' \end{bmatrix}$$

The matrices A', A'' intersect at only one position, which is the entry containing a_{vv}. Let $B = B(G)$, $B' = B(G')$ and $B'' = B(G'')$. First, we will derive that B has full rank whenever both B', B'' have full rank, thus proving the first part of the proposition. To do so, we partition the matrix B into blocks, such that some of these blocks are equal to B', B''. Recall that the rows of B are indexed by the zero entries of A, and the columns of B are indexed by the pairs (i,j) with $i,j \neq 1$ and $i \neq j$. These pairs correspond to the (possibly) nonzero positions of $X \in V_A$ whose first row, first column and diagonal are all zero. We find a block partition of B by partitioning both A and X, since this gives us a partition of the row and column indices. Let A', A'', A_3, A_4 be blocks of A and let X', X'', X_3, X_4 be blocks of X as given below:

$$A = \begin{array}{c} \begin{array}{ccc} V'\setminus\{v\} & v & V''\setminus\{v\} \end{array} \\ \left[\begin{array}{c:c:c} A' & & A_3 \\ \hdashline & & \\ \hdashline A_4 & & A'' \end{array}\right] \begin{array}{l} V'\setminus\{v\} \\ v \\ V''\setminus\{v\} \end{array} \end{array} \qquad X = \begin{array}{c} \begin{array}{ccc} V'\setminus\{v\} & v & V''\setminus\{v\} \end{array} \\ \left[\begin{array}{c:c:c} X' & & X_3 \\ \hdashline & & \\ \hdashline X_4 & & X'' \end{array}\right] \begin{array}{l} V'\setminus\{v\} \\ v \\ V''\setminus\{v\} \end{array} \end{array}$$

The solid lines indicate the partitioning and the dotted lines indicate the position of the row and column indexed by vertex v. We obtain a partition of the rows and columns of B by distinguishing between the four blocks of A and X, respectively. Note that the blocks of A do not form a partition of the matrix, because A' and A'' intersect. However, they intersect at a nonzero position, so this position does not appear as a row index of B. Therefore the blocks of A induce a well-defined partition of the row indices of B. We obtain the following block matrix:

$$B = \begin{array}{c} \\ A' \\ A'' \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{cccc} X' & X'' & X_3 & X_4 \end{array} \\ \left[\begin{array}{c|c|c|c} B' & 0 & C_1 & C_2 \\ \hline 0 & B'' & 0 & 0 \\ \hline C_3 & 0 & D_1 & 0 \\ \hline C_4 & 0 & 0 & D_2 \end{array}\right] \end{array}$$

The zero positions of the matrices A', A'' are exactly the row indices of B', B'', respectively. Furthermore, the positions of X', X'' which are outside the first row, first column and diagonal of X yield exactly the column indices of B', B'', respectively. Since each edge of G' also appears in G, it follows from equation (4.1) that the block indexed by A' and X' is indeed the matrix B' corresponding to G', and similarly, the block indexed by A'' and X'' is exactly the matrix B'' corresponding to G''. Now consider the block indexed by A' and X''; if (k,l) is a zero position of A' and (i,j) is a nonzero position of X'', then these two pairs only have a coordinate in common is when $j = l = v$ or $i = k = v$. Then the corresponding entry of B is of the form a_{ki} or $-a_{jl}$, respectively, with $i, j \in V'' \setminus \{v\}$ and $k, l \in V'$. However, there are no such edges $j \to l$ or $k \to i$ in G, because these would correspond to edges between G' and G'' not incident to v. Hence the corresponding entry of B is zero, and therefore the entire block indexed by A', X'' is zero. A similar analysis shows that each of the blocks of B denoted with a zero indeed is a zero matrix.

Next, we analyze the blocks C_1, C_2, C_3 and C_4. For C_1 to have a nonzero entry, we need a position (k,l) of A' and a position (i,j) of X_3 to have a coordinate in common. From the way that A and X have been partitioned, we see that the only option is $k = i$. This gives the entry $-a_{jl}$, where $j \in V'' \setminus \{v\}$ and $l \in V$. The only parameters of this form are a_{jv}, with $v \to j$ an edge in G''. Therefore, the nonzero entries of C_1 are indexed by $(i,v), (i,j)$ such that $v \to j$ is an edge in G, and the corresponding entry is of the form $-a_{jv}$. Similarly, the nonzero entries of C_2 are indexed by $(v,j), (i,j)$ such that $i \to v$ is an edge in G'', and the corresponding entry is of the form a_{vi}. For the block C_3 the same analysis shows that all nonzero entries are of the form $-a_{vi}$, while the nonzero entries of C_4 are of the form a_{iv}, $i \in V'' \setminus \{v\}$. The exact form of these blocks is not important for our further analysis, all we need is that the only nonzero entries are either $\pm a_{vi}$ or $\pm a_{iv}$ with $i \in V'' \setminus \{v\}$.

Finally, consider the block D_1, which is indexed by A_3, X_3. The block A_3 has size $(|V'| - 1)(|V''| - 1)$ and consists entirely of zeros, so the number of rows of D_1 equals $(|V'| - 1)(|V''| - 1)$. The columns of D_1 are indexed by the block X_3 of size $|V'|(|V''| - 1)$. Since the first row of X must be zero, X_3 gives only $(|V'| - 1)(|V''| - 1)$ column indices. We conclude that D_1 is square, hence we can calculate its determinant to see whether it has full rank. Observe that A_3 yields pairs (i,j) with $i \in V' \setminus \{v\}$ and $j \in V'' \setminus \{v\}$, while the pairs corresponding to X_3 are of the form (i,j) with $i \in V' \setminus \{1\}$ and $j \in V'' \setminus \{v\}$. Let $\{v, i_1, \ldots, i_t, 1\}$ be a path in G_1 from v to 1, and let $r_1, \ldots, r_k$ be the vertices of G' not appearing in this path. We use this path to order the row indices of D_1 by their first coordinate:

$$(i_1, \cdot), (i_2, \cdot), \ldots, (i_t, \cdot), (1, \cdot), (r_1, \cdot), \ldots, (r_k, \cdot).$$

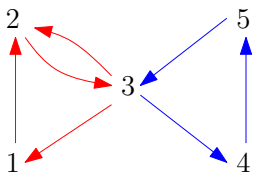


FIGURE 4.3. $G' \cup G''$

We order the column indices in a similar fashion:

$$(v,\cdot),(i_1,\cdot),\dots,(i_t,\cdot),(r_1,\cdot),\dots,(r_k,\cdot).$$

Next, each subset of the form $(i,\cdot)$ is ordered by increasing value of the second coordinate. Note that the second coordinate of both the row and the column indices runs through all elements of $V''\setminus\{v\}$. We choose this ordering because it makes all diagonal entries nonzero: the entries indexed by $(i_s,k),(i_{s-1},k)$ with $k\in V''\setminus\{v\}$ have value $a_{i_si_{s-1}}$, which is nonzero because $i_{s-1}\to i_s$ is an edge in the path from v to 1. The remaining diagonal entries are indexed by $(r,j),(r,j)$ for $r\in\{r_1,\dots,r_k\}$, $j\in V''\setminus\{v\}$. These entries are of the form $a_{rr}-a_{jj}$, which is certainly nonzero as it contains independent diagonal entries of A. Hence the entire diagonal of D_1 has become nonzero and, more importantly, entries of the form $\pm a_{i_si_{s-1}}$ cannot appear outside the diagonal of D_1. This is derived from the fact that there are no indices $(\cdot,i_{s-1}),(\cdot,i_s)$, because $i_s\in V_1$ while the second coordinates lie in V_2. Entries of the form $\pm(a_{rr}-a_{jj})$ can only appear at two positions of M, namely $(j,r),(j,r)$ and $(r,j),(r,j)$, but only the second position lies in D_1. This shows that the diagonal entries are independent of the other entries of D_1, hence the determinant of D_1 is generically nonzero.
A very similar argument shows that also D_2 has nonzero determinant.

Now suppose that both B' and B'' have full rank. These matrices do not need to be square, since the number of rows may be larger than the number of columns. However, being full rank means that there exists a subset of the rows such that the corresponding matrix is square and invertible. Let $\widehat{B'},\widehat{B''}$ be such square submatrices with nonzero determinant, and let $\widehat{B}$ be the corresponding square submatrix of B. Then from the structure of B, we see that the determinant $\det(\widehat{B})$ contains a term

$$\det(\widehat{B'})\det(\widehat{B''})\det(D_1)\det(D_2).$$

Moreover, the determinant of $\widehat{B}$ contains a factor $\det(\widehat{B''})$, because all other entries in the corresponding rows and columns are zero. The nonzero off-diagonal blocks only contain entries of the form a_{vj} and a_{jv} with $j>v$, but these entries do not appear in D_1,D_2 or B'. Therefore the term above can never vanish, i.e. $\widehat{B}$ has nonzero determinant.

The second part of the proposition follows directly from the fact that the determinant of $\widehat{B}$ contains a factor $\det(\widehat{B''})$: if $\det(\widehat{B})$ is nonzero, then $\det(\widehat{B''})$ must also be nonzero. □

REMARK 4.3.2. Let G,G',G'' be as in Proposition 4.3.1. We have just seen that if both G',G'' have the expected dimension, then so does G. Conversely, if G' does not have the expected dimension, this does not necessarily imply that G does not have the expected dimension. For example, the graph in Figure 4.3 has the expected dimension, while its subgraph G' does not. However, if $V'\cap V''=\{1\}$, then applying the proposition twice shows that G has the expected dimension if and only if both G' and G'' have the expected dimension.

For our next result, recall Proposition 4.2.4 of the previous section; it states that if G' is a graph on $n-1$ vertices which has the expected dimension and we construct G from G' by adding a new vertex n and two edges $k \to n$ and $n \to l$ such that G has a chain of cycles containing both 1 and n, then G has the expected dimension as well. Using Theorem 1.4.2 we prove a stronger version of this theorem:

PROPOSITION 4.3.3. *If $G = (V, E)$ on $n-1$ vertices has the expected dimension and we construct G' from G by adding a new vertex n and edges $k \to n$ and $n \to l$ for some $k, l \in V$, then G' has the expected dimension.*

PROOF. Let $A = A(G)$ and $A' = A(G')$ be the parameter matrices corresponding to G and G', respectively. Observe that A' is constructed from A by adding an extra row and column, so A is a submatrix of A'. Consider the coefficient matrices $B = B(G)$ and $B' = B(G')$ corresponding to G and G', respectively. Since G is assumed to have the expected dimension, we know from Theorem 1.4.2 that B has full rank. We will use this fact to show that B' has full rank as well.

From the fact that B has full rank, it follows that B has a square submatrix $\widehat{B}$ of size $(n-3)(n-2)$ which has nonzero determinant. This matrix is obtained from B by deleting a subset of the rows. Because A is a submatrix of A', all row indices of B are also row indices of B', so we can construct $\widehat{B}'$ from B' by removing the same subset of the rows as we removed from B.

Rearrange the rows of $\widehat{B}'$ such that the last $2n-4$ rows are indexed by pairs of the form (n, i) or (j, n), where $i, j \in \{1, \ldots, n-1\}$, $i \neq k$ and $j \neq l$. Similarly, rearrange the columns of B' such that the last $2n-4$ columns are indexed by pairs of the form (n, i) or (j, n), where $i, j \in \{2, \ldots, n-1\}$. If we would remove these rows and columns, the remaining matrix would be precisely the matrix $\widehat{B}$. So we write $\widehat{B}'$ as a block matrix,

$$\widehat{B}' = \left[\begin{array}{c|c} \widehat{B} & C_1 \\ \hline C_2 & D \end{array}\right],$$

where the block D has size $(2n-4) \times (2n-4)$.

Next, consider the blocks C_1, C_2 and D. According to equation (4.1), blocks C_1 and C_2 have entries of the form $\pm a_{in}$ and $\pm a_{nj}$ with $i, j \in \{1, \ldots, n-1\}$, but the only nonzero parameters of this form are a_{nk} and a_{ln}, since these are the only edges to and from vertex n. Block D can again be written as a block matrix, by distinguishing between entries of the form $(n, \cdot)$ and $(\cdot, n)$. Note that $B_{(n,i),(j,n)} = 0$ for all $i, j \neq n$. We claim that D will be block diagonal:

$$D = \begin{array}{c} \\ (n,\cdot) \\ \\ (\cdot,n) \end{array} \begin{array}{c} \begin{array}{cc} (n,\cdot) & \quad (\cdot,n) \end{array} \\ \left[\begin{array}{c|c} -(A^T)_{k,1,n} & \mathbf{0} \\ \hline \mathbf{0} & A_{l,1,n} \end{array}\right] \end{array},$$

where the diagonal blocks $A_{l,1,n}, -(A^T)_{k,1,n}$ are submatrices of A, A^T, respectively, defined as follows. Let $(A^T)_{k,1,n}$ be the submatrix of A^T obtained by removing row k and column 1, and replacing a_{ii} by $a_{ii} - a_{nn}$ for $i \in \{2, \ldots, n-1\}$. Let $A_{l,1,n}$ be the submatrix of A obtained by removing row l and column 1, and replacing a_{ii} by $a_{ii} - a_{nn}$. Indeed, for the entries of block D that are indexed by $((i, n), (j, n))$ with $i \neq l, n$ and $j \neq 1, n$, we obtain the matrix A except for column 1 and row l, with diagonal entries $a_{ii} - a_{nn}$. Similarly, it follows that the other diagonal block of D equals $-(A^T)_{k,1,n}$. For sufficiently general A, the matrices $A_{k,1,n}$ and $A_{l,1,n}$ have full rank (Lemma 4.3.4) hence D will have nonzero determinant.

Neither D nor $\widehat{B}$ can contain the parameters a_{nk}, a_{ln}, which are the only nonzero entries of C_1, C_2. Therefore, we can set $a_{nk} = a_{ln} = 0$, such that $\widehat{B}'$

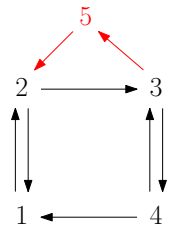


FIGURE 4.4. Graph G'; obtained from G (black) by adding a line segment (red).

becomes a block diagonal matrix with invertible diagonal blocks. This proves that $\widehat{B}'$ has full rank on an open dense subset of $\Theta_{G'}$, so B' has full column rank for sufficiently general A'. We conclude that G' has the expected dimension. □

LEMMA 4.3.4. *Let G be a strongly connected graph on $n-1$ vertices and let $A \in \Theta_G$. For $l \in [n-1]$ define $A_{k,l,n}$ to be the submatrix of A obtained by replacing the diagonal entries a_{ii} by $a_{ii}-a_{nn}$ for all $i \in [n-1]$, and removing row k and column l. Then for sufficiently general A, the matrix $A_{k,l,n}$ has nonzero determinant.*

PROOF. Since G is strongly connected, there exists a path p from k to l, say

$$p = \{k = v_1, v_2, \ldots, v_{r-1}, v_r = l\}.$$

Let $v_{r+1}, \ldots, v_{n-1}$ be the vertices of G that do not appear in p. Rearrange the rows and columns of $A_{k,l,n}$ such that the row indices are ordered as

$$v_2, v_3, \ldots, v_r, v_{r+1}, \ldots, v_{n-1}$$

and the column indices are ordered as

$$v_1, v_2 \ldots, v_{r-1}, v_{r+1}, \ldots, v_{n-1}.$$

Then $A_{k,l,n}$ has diagonal

$$(a_{v_2v_1}, a_{v_3v_2}, \ldots, a_{v_rv_{r-1}}, a_{v_{r+1}v_{r+1}} - a_{nn}, \ldots, a_{v_{n-1}v_{n-1}} - a_{nn})$$

whose entries are nonzero for sufficiently general A. All entries of $A_{k,l,n}$ correspond to different parameters, so taking the diagonal entries large enough will make the determinant of $A_{k,l,n}$ nonzero. Having full rank is a Zariski open condition on the parameters (Appendix A) so it follows that $A_{k,l,n}$ has full rank for sufficiently general A. □

Observe that the proof of Proposition 4.3.3 does not need any restrictions on $k, l \in V$, so we can add a cycle by choosing $k = l$.

The converse of Proposition 4.3.3 does not hold; if G does not have the expected dimension, then G' might still have the expected dimension. For example, consider the graph G' in Figure 4.4 which is obtained from the graph G by adding vertex 5 and edges $3 \to 5$ and $5 \to 2$. The graphs G, G' have parameter matrices A, A', respectively:

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} \\ a_{21} & a_{22} & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix} \qquad A' = \begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} & 0 \\ a_{21} & a_{22} & 0 & 0 & a_{25} \\ 0 & a_{32} & a_{33} & a_{34} & 0 \\ 0 & 0 & a_{43} & a_{44} & 0 \\ 0 & 0 & a_{53} & 0 & a_{55} \end{bmatrix}$$

From the structure of A we see that G does not have the expected dimension, since the pair $(3, 4)$ satisfies Condition 3.3.1. On the other hand, one can check that G' *does* have the expected dimension using Algorithm 3.2.

	V	$(\cdot,n_1)$	$(n_1,\cdot)$	$(\cdot,n_2)$	$(n_2,\cdot)$	$(\cdot,n_3)$	$(n_3,\cdot)$	.....	$(\cdot,n_{s-1})$	$(n_{s-1},\cdot)$	$(\cdot,n_s)$	$(n_s,\cdot)$
V	$\widehat{B}$	a_{n_1k}										a_{ln_s}
$(\cdot,n_1)$		D_1		$a_{n_2n_1}$								a_{ln_s}
$(n_1,\cdot)$	a_{n_1k}		D_1'									
$(\cdot,n_2)$				D_2		$a_{n_3n_2}$						a_{ln_s}
$(n_2,\cdot)$			$a_{n_2n_1}$		D_2'							
$(\cdot,n_3)$						D_3						a_{ln_s}
$(n_3,\cdot)$					$a_{n_3n_2}$		D_3'					
$\vdots$								$\ddots$				
$(\cdot,n_{s-1})$									D_{s-1}		$a_{n_sn_{s-1}}$	a_{ln_s}
$(n_{s-1},\cdot)$										D_{s-1}'		
$(\cdot,n_s)$	a_{ln_s}		a_{ln_s}		a_{ln_s}		a_{ln_s}			a_{ln_s}	D_s	
$(n_s,\cdot)$										$a_{n_sn_{s-1}}$		D_s'

FIGURE 4.5. Block form of $\hat{B}'$

We have just proven that it is a valid operation to add a line segment of length two to a given graph with the expected dimension. The next theorem strengthens this result, saying that it is a valid operation to add a line segment of any length.

PROPOSITION 4.3.5. *Let $G = (V, E)$ on $n-1$ vertices be a graph with the expected dimension. Construct G' from G by adding new vertices $n_1, \ldots, n_s$ and edges $k \to n_1$, $n_s \to l$ and $n_i \to n_{i+1}$ for $i = 1, \ldots, s-1$, where $k, l \in V$ are vertices of G. Then G' has the expected dimension.*

PROOF. Let $A = A(G)$ and $A' = A(G')$ and observe that A is a submatrix of A', since G is a subgraph of G'. Consider the coefficient matrices $B = B(G)$ and $B' = B(G')$ corresponding to G and G', respectively. Similar to the proof of Proposition 4.3.3, we will use the fact that B has full rank to show that also B' has full rank. Let $\widehat{B}$ be obtained from B by deleting a subset of the rows, such that $\widehat{B}$ has nonzero determinant. Let $\widehat{B}'$ be obtained from B' by deleting the same subset of the rows, and additionally deleting the rows indexed by $(1, n_p)$ for all $p \in [s]$. Again, we rearrange the rows and columns of $\widehat{B}'$, but since we add multiple vertices this needs a bit more attention. Distinguish between indices of the form (i, j) with $i, j \in V$ and indices of the form (n_p, r) or (r, n_p) with $p \in [s]$ and $r \in V \cup \{n_q \mid q < p\}$. We claim that this brings $\widehat{B}'$ into the form of Figure 4.5, where the empty blocks are all zero, and the blocks containing a parameter a_{ij} contain both zero entries and entries of the form a_{ij}.

To show that $\widehat{B}'$ is indeed of this form, we consider the different blocks one by one, starting with the upper left block. From the fact that A is a submatrix of A', it follows that the submatrix of $\widehat{B}'$ corresponding to the rows and columns indexed by pairs of the form (i, j) with $i, j \in V$ is exactly the matrix $\widehat{B}$.

The remaining positions in the rows indexed by (i, j) with $i, j \in V$ have columns indexed by (n_p, r) or (r, n_p) with $p \in [s]$ and $r \in V \cup \{n_q \mid q < p\}$. According to equation (4.1), the only nonzero entries occur when $r = j$ or $r = i$, respectively, and this gives the entry a_{n_pi} or a_{jn_p}. However, the only entries of A' of this form are a_{n_1k} and a_{ln_s}. A similar analysis of the positions in the columns indexed by (i, j) with $i, j \in V$ but outside $\widehat{B}$, shows that the only nonzero entries are of the form a_{n_1k} and a_{ln_s} as well.

Next, consider the blocks D_p, where $p \in [s]$. The positions of D_p are indexed by $(i, n_p), (j, n_p)$, hence they have at least their second coordinate in common. If $i \neq j$, the corresponding entry is of the form a_{ij}, and if $i = j$ we obtain $a_{ii} - a_{n_p n_p}$. Because of our ordering of rows and columns, we have $i, j \in V \cup \{n_q \mid q < p\}$, $i \neq 1$, and by definition of $B' = B(G')$ also $j \neq 1$. Therefore the block D_p equals the submatrix of A' obtained by deleting rows and columns indexed by 1 or n_q with $q \geq p$. Furthermore, the diagonal entries a_{ii} of A' have been replaced by entries $a_{ii} - a_{n_p n_p}$. Thus, the block D_p is a block lower diagonal square matrix of the form

$$D_p = \left[\begin{array}{c|l} A_{1,1,n_p} & \\ \hline a_{n_1 k} & a_{n_1 n_1} - a_{n_p n_p} \\ & a_{n_2 n_1} \quad \ddots \\ & \qquad\quad \ddots \quad a_{n_{p-1} n_{p-1}} - a_{n_p n_p} \\ & \qquad\quad a_{n_{p-1} n_{p-2}} \end{array}\right]$$

with determinant

$$\det(D_p) = \begin{cases} \det(A_{1,1,n_p}) & \text{if } p = 1 \\ \det(A_{1,1,n_p}) \prod_{j=1}^{p-1} (a_{n_j n_j} - a_{n_p n_p}) & \text{if } p \geq 2. \end{cases}$$

Analogously, the row and column indices of blocks D'_p all have their first coordinate in common. The positions are indexed by $(n_p, i), (n_p, j)$, which implies that the entries are of the form $-a_{ji}$ or $-(a_{ii} - a_{n_p n_p})$. Note that we obtain $-a_{ji}$ instead of a_{ij}, so if we apply a similar analysis as we did for D_p, we obtain some submatrix of $-(A')^T$. Furthermore, there is no row index of the form (n_p, n_{p-1}), since the edge $n_{p-1} \to n_p$ occurs in G'. Hence $-(D'_p)^T$ equals the submatrix of A' obtained by deleting rows and columns indexed by n_q with $q \geq p$, and also deleting column n_{p-1} and row 1. Again, the diagonal entries a_{ii} of A' have been replaced by entries $a_{ii} - a_{n_p n_p}$. If we separate column k from the rest of the columns indexed by (n_p, i) with $i \in V$, we see that the block D'_p is a block upper diagonal square matrix of the form

$$-(D'_p)^T = \left[\begin{array}{c|c|l} & a_{2k} & \\ A_{1,k,n_p} & \vdots & \\ & a_{(n-1)k} & \\ \hline & a_{n_1 k} & a_{n_1 n_1} - a_{n_p n_p} \\ \hline & & a_{n_2 n_1} \quad \ddots \\ & & \qquad\quad \ddots \quad a_{n_{p-2} n_{p-2}} - a_{n_p n_p} \\ & & \qquad\quad a_{n_{p-1} n_{p-2}} \end{array}\right]$$

and its determinant satisfies

$$\pm \det(D'_p) = \begin{cases} \det(A_{1,k,n_1}) & \text{if } p = 1 \\ \det(A_{1,k,n_2}) a_{n_1 k} & \text{if } p = 2 \\ \det(A_{1,k,n_p}) a_{n_1 k} \prod_{j=1}^{p-2} (a_{n_{j+1} n_j}) & \text{if } p \geq 3. \end{cases}$$

Finally, consider the entries which do not lie in any of the blocks D_p or D'_p or in the rows and columns denoted by V in $\widehat{B}'$. These rows are indexed by $(n_p, \cdot)$ or $(\cdot, n_p)$, while the columns are indexed by $(n_q, \cdot)$ or $(\cdot, n_q)$ with $q \neq p$. If two such pairs have an entry in common, then the two remaining entries determine the value of the corresponding entry of $\widehat{B}'$. Equation (4.2) gives an overview of all possible

pairs of indices having a coordinate in common, the corresponding entry of $\widehat{B}'$ and the conditions for this entry to be nonzero.

$$
\begin{array}{rcll}
(n_p,i),(n_q,i) & \rightsquigarrow & a_{n_pn_q} & \neq 0 \text{ iff } p=q+1 \\
(i,n_p),(i,n_q) & \rightsquigarrow & a_{n_qn_p} & \neq 0 \text{ iff } p=q-1 \\
(n_p,n_q),(i,n_q) & \rightsquigarrow & a_{n_pi} \quad \text{if } p>q & =0 \\
(n_q,n_p),(n_q,i) & \rightsquigarrow & a_{in_p} \quad \text{if } p>q & \neq 0 \text{ iff } p=s, i=l \\
(n_p,i),(n_p,n_q) & \rightsquigarrow & a_{n_qi} \quad \text{if } p<q & =0 \\
(i,n_p),(n_q,n_p) & \rightsquigarrow & a_{in_q} \quad \text{if } p<q & \neq 0 \text{ iff } q=s, i=l
\end{array} \tag{4.2}
$$

We conclude that $\widehat{B}'$ is indeed as claimed. Since the diagonal blocks of $\widehat{B}'$ have full rank, its determinant contains a term

$$m_1 = \det(\widehat{B}) \prod_{p=1}^{s} \det(D_p) \det(D'_p),$$

and we will show that this term cannot be cancelled out.

Recall that it is enough to show that $\det(\widehat{B}')$ is nonzero for A in some Zariski open subset of Θ, since we only consider sufficiently general A (see also Appendix A). Therefore, we can simplify things by setting $a_{ln_s} = 0$ and $a_{n_sn_{s-1}} = 0$. Note that these parameters do not appear in m_1, because none of the D_p, D'_p contain rows or columns of A' that are indexed by n_s, hence setting these parameters to zero does not affect m_1.

Suppose m_2 is a term of the determinant of $\widehat{B}'$ which cancels out m_1, then it must have at least one entry from outside the diagonal blocks of $\widehat{B}'$. From our previous observations, we know that these entries are of the form

$$a_{n_1k}, a_{n_2n_1}, \ldots, a_{n_{s-1}n_{s-2}},$$

since $a_{n_sn_{s-1}} = a_{jn_s} = 0$. As we have seen, these parameters also appear in the determinants of D'_p. More specific, $a_{n_{s-1}n_{s-2}}$ occurs only in $\det(D'_s)$, while $a_{n_{s-2}n_{s-3}}$ occurs in $\det(D'_s)$ and $\det(D'_{s-1})$, continuing up to a_{n_1k} which divides all of $\det(D'_1), \ldots, \det(D'_s)$. Using this observation, we will argue that m_2 can never yield the term m_1.

Consider the block D'_s. From the fact that we set $a_{n_sn_{s-1}}$ to zero, it follows that the rows of $\widehat{B}'$ indexed by $(n_s, \cdot)$ are all zero outside D'_s. Therefore, any term in the determinant of $\widehat{B}'$ contains $\det(D'_s)$. The parameter $a_{n_{s-1}n_{s-2}}$ appears exactly once in m_1, namely in $\det(D'_s)$, so m_2 cannot contain another factor $a_{n_{s-1}n_{s-2}}$.

Now consider the block D'_{s-1} and observe that the only nonzero entries in the rows of $\widehat{B}'$ indexed by $(n_{s-1}, \cdot)$ are of the form $a_{n_{s-1}n_{s-2}}$. From the observation that m_2 cannot contain these entries, it follows that m_2 contains a factor $\det(D'_{s-1})$. The parameter $a_{n_{s-2}n_{s-3}}$ appears exactly twice in m_1, namely in $\det(D'_s)$ and in $\det(D'_{s-1})$, hence m_2 cannot contain a third factor $a_{n_{s-2}n_{s-3}}$. One can repeat this argument, showing step by step dat m_2 cannot contain any entries which are outside the diagonal blocks. However, this implies that m_1 cannot be cancelled out by m_2 in the determinant of $\widehat{B}'$. This shows that $\det(\widehat{B}')$ is nonzero for sufficiently general A', thus proving that G' has the expected dimension. □

One can obtain Proposition 4.3.5 from Proposition 4.3.3 by subdividing the edges of the line segment. We believe that in general, subdividing an edge is a valid operation. It has been verified for all graphs G on four and five vertices using MATHEMATICA, as well as for larger random graphs. Unfortunately, the techniques we used to prove the previous propositions cannot be applied here so easily, because the matrix $A(G)$ is not a submatrix of $A(G')$.

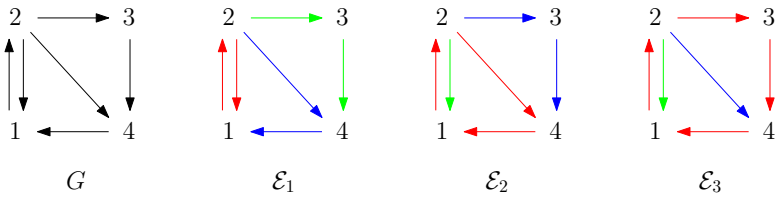


FIGURE 4.6. A graph G with three different ear decompositions $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$.

CONJECTURE 4.3.6. *Let G be a graph on $n-1$ vertices and let $k \to l$ be an edge in G. Construct the graph G' on n vertices subdividing the edge $k \to l$, i.e. by adding vertex n to G and replacing the edge $k \to l$ by two edges $k \to n, n \to l$. Then if G has the expected dimension, G' has the expected dimension as well.*

4.4. Ear decompositions

This section describes how to construct graphs with the expected dimension using Propostion 4.3.5. Starting from a cycle, which has the expected dimension, we can add line segments to obtain new graphs, for example all minimally strongly connected graphs. This gives rise to a procedure to obtain a graph which has the expected dimension from a graph which does not have the expected dimension. An important concept that we shall be using is the ear decomposition of a directed graph, as defined in [**BJG07**]:

DEFINITION 4.4.1. Given a directed graph G, let $\mathcal{E} = \{P_0, P_1, \ldots, P_t\}$ be a sequence of cycles and paths in G, $t \geq 0$, and define $G_i = P_0 \cup P_1 \cup \ldots \cup P_i$. Then $\mathcal{E}$ is an *ear decomposition* of G if P_0 is a cycle, $G_t = G$, and each P_i is a path from G_{i-1} to G_{i-1}, but has no vertices in common with $P_0, \ldots, P_{i-1}$ except for its starting point and its endpoint. The P_i are called the *ears* of $\mathcal{E}$, and if P_i has length one it is called a *trivial ear*.

Note that the graphs $G_0, \ldots, G_t$ are strongly connected, hence if G has an ear decomposition then it must be strongly connected. The converse also holds: if a graph is strongly connected, then it must have an ear decomposition. Indeed, a strongly connected graph G contains at least one cycle, so let this be P_0. Suppose we have found ears $P_0, P_1, \ldots, P_i$ with corresponding graphs $G_0, G_1, \ldots, G_i$ as defined above. If G_i does not contain all vertices of G, choose a vertex y that is not in G_i, such that there exists an edge $x \to y$ in G for some vertex x in G_i. Since G is strongly connected, such a vertex must exist. Furthermore, there must be a path P from y back to G_i, so choose P_{i+1} to consist of $x \to y$ followed by P. We can continue this procedure, until at some point G_j contains all vertices of G. Then all remaining edges must be trivial, so we can add these one by one to complete the ear decomposition. Thus, we have shown the following well-known condition:

LEMMA 4.4.2. *A graph has an ear decomposition if and only if it is strongly connected.*

A graph may have many different ear decompositions, as shown in Figure 4.6. Each of these decompositions has the same number of ears, namely $m-n+1$. This follows from the fact that all ears are edge disjoint, and if the ear P_i contains k edges, then the corresponding graph G_i has $k-1$ new vertices with respect to G_{i-1}. So if a decomposition has t ears, then the number of vertices of G must be equal to $m - t + 1$, where the term "+1" comes from the initial cycle P_0. We conclude that indeed the number of ears in an ear decomposition of a given graph is fixed.

Our purpose is to construct graphs with the expected dimension, using the results of the previous section. Therefore, we define a specific kind of ear decomposition:

DEFINITION 4.4.3. A graph G is said to have a *correct ear decomposition* if it has an ear decomposition without trivial ears, and such that the initial cycle P_0 contains vertex 1.

Consider the graph G given in Figure 4.6, and the three ear decompositions $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$. In each $\mathcal{E}_i$, let P_0, P_1 and P_2 be the red, blue and green ears, respectively. Then the initial cycle P_0 contains vertex 1 in each of the three decompositions. However, $\mathcal{E}_1$ is the only ear decomposition without trivial ears. In other words, $\mathcal{E}_1$ is a correct ear decomposition of G, but $\mathcal{E}_2$ and $\mathcal{E}_3$ are not.

THEOREM 1.4.4. *Let G be a graph that has a correct ear decomposition, then G has the expected dimension.*

PROOF. Let G have a correct ear decomposition $\mathcal{E} = \{P_0, \ldots, P_t\}$. A correct ear decomposition consists of nontrivial ears, and a nontrivial ear corresponds to a line segment of length at least two, as defined in Section 4.1. Since a cycle is known to have the expected dimension, it follows that G_0 has the expected dimension, and we can apply Proposition 4.3.5 t times to conclude that $G_1, G_2, \ldots, G_t$ all have the expected dimension. In other words, if a graph has a correct ear decomposition, then it certainly has the expected dimension. □

Conversely, if a graph has the expected dimension, it does not need to have a correct ear decomposition. For example, the graph in Figure 4.3 has no correct ear decomposition, yet it does have the expected dimension.

PROPOSITION 4.4.4. *A graph G is minimally strongly connected if and only if all its ear decompositions have no trivial ears.*

PROOF. Suppose that G has an ear decomposition with a trivial ear. Then the graph obtained from G by deleting the edge of this trivial ear is strongly connected, because it has an ear decomposition. Hence G is not minimally strongly connected.

Conversely, suppose that G is not minimally strongly connected, then it has an edge e that can be removed, such that the resulting graph G' remains strongly connected. Then G' has an ear decomposition, and adding the trivial ear that contains the edge e results in an ear decomposition of G that has a trivial ear. □

We have seen that a graph which has a correct ear decomposition has the expected dimension, hence an immediate consequence of the above proposition is the following:

COROLLARY 4.4.5. *If G is minimally strongly connected, then it has the expected dimension.*

The converse does not hold, because for G to have the expected dimension it is enough to have only one correct ear decomposition. For example, the graph G in Figure 4.6 has a correct ear decomposition, hence the expected dimension, but it is not minimally strongly connected.

PROPOSITION 4.4.6. *Let G be a graph that contains a 2-cycle $\{i \to j, j \to i\}$, and let G' be the graph where the vertices i and j have been identified. If G has a correct ear decomposition, then so does G'.*

PROOF. Let $\mathcal{E}$ be a correct ear decomposition of G. Any 2-cycle C in G must appear as an ear in $\mathcal{E}$, because if there is an ear that contains only one of the two edges in C, then the other edge can only appear as a trivial ear. After identifying

vertices i and j to obtain G', the 2-cycle no longer exists. A correct ear decomposition of G' is obtained from $\mathcal{E}$ by removing the ear that is equal to the 2-cycle and replacing vertex j by vertex i in the remaining ears. Note that the number of edges in these ears does not change, so they remain nontrivial. □

A special case of this theorem occurs when G has an exchange, and G' is obtained by collapsing the exchange. This shows that Conjecture 4.2.6 holds when G has a correct ear decomposition.

Theorem 1.4.4 gives rise to two options to turn a graph that does not have the expected dimension into a graph that does have the expect dimension. If G does not have the expected dimension, then every ear decomposition of G contains a trivial ear. In order to transform the graph into one that has a correct ear decomposition, start with an arbitrary ear decomposition and either remove the trivial ears, or subdivide the corresponding edges, such that they are no longer trivial. To keep the number of changes as small as possible, one should start with an ear decomposition with the smallest possible number of trivial ears. In Section 5.2 we will briefly discuss what this means in practice.

4.5. Computational results

In the previous section, we have seen two classes of graphs with the expected dimension: the graphs which have a correct ear decomposition, and those which are minimally strongly connected. Moreover, from Corollary 4.2.5 we know that all inductively strongly connected graphs (with at most $2n-2$ vertices) have the expected dimension. Using the computer algebra package MATHEMATICA, the cardinalities of these classes have been calculated for $n = 3, 4, 5$. We will first discuss in short how these results were obtained, the corresponding source code is available upon request.

Let $\mathcal{G}(n)$ denote the class of strongly connected graphs on n vertices with at most $2n-2$ edges, up to the following equivalence. Since vertex 1 has a special role, graphs are considered to be equivalent if they can be obtained from one another by permuting vertices $2, 3, \ldots, n$. This class can be constructed in MATHEMATICA using the function "`ListGraphs[n, m, Directed]`" from the Combinatorica package. This function generates all directed graphs on n vertices and m edges, up to permuting vertices $1, 2, \ldots, n$. We obtain only a subset of $\mathcal{G}(n)$, hence to each graph generated by MATHEMATICA we apply the permutations $(12), (13), \ldots, (1n)$, and filter the resulting graphs up to permutations of $2, 3, \ldots, n$.

Let $\mathcal{G}^*(n)$ be the subset of $\mathcal{G}(n)$ containing all graphs which have the expected dimension. This class is calculated efficiently by applying Algorithm 3.2 of the previous chapter to the graphs in $\mathcal{G}(n)$.

Define $\mathcal{G}_c(n)$ to be the class of graphs in $\mathcal{G}(n)$ that have a correct ear decomposition. This class can be calculated by doing a breadth-first search, computing all correct ear decompositions of the input graph.

Let $\mathcal{G}_{\mathrm{ISC}}(n)$ consists of all graph in $\mathcal{G}(n)$ which are inductively strongly connected. These graphs can be constructed by adding the vertices one by one, while at each step the resulting graph is strongly connected. Again, this class can be computed by doing a breadth-first search. Note that an inductively strongly connected graph has at least $2n-2$ edges, so for elements of $\mathcal{G}_{\mathrm{ISC}}(n)$ equality holds.

Finally, let $\mathcal{G}_{\mathrm{MSC}}(n)$ be the class containing all minimally strongly connected graphs. To check whether a graph $G = (V, E)$ is in this class, let $E = \{e_1, \ldots, e_m\}$ and construct graphs $G_i = (V, E \setminus \{e_i\})$ for $i = 1, \ldots, m$. For each of these graphs, check whether is is strongly connected; as soon as some G_i is indeed strongly connected, G is not in $\mathcal{G}_{\mathrm{MSC}}(n)$. If there is no such G_i, then G is minimally strongly connected.

n	$\lvert\mathcal{G}(n)\rvert$	$\lvert\mathcal{G}^*(n)\rvert$	$\lvert\mathcal{G}_c(n)\rvert$	$\lvert\mathcal{G}_{\mathrm{ISC}}(n)\rvert$	$\lvert\mathcal{G}_{\mathrm{MSC}}(n)\rvert$
3	6	5	5	4	3
4	71	43	39	26	12
5	1472	628	450	267	57

TABLE 4.1. Computational results

From section 4.4, we know that

$$\mathcal{G}_{\mathrm{MSC}}(n) \subsetneq \mathcal{G}_c(n) \subsetneq \mathcal{G}^*(n) \subsetneq \mathcal{G}(n).$$

The class of inductively strongly connected graphs $\mathcal{G}_{\mathrm{ISC}}(n)$ is also a subset of $\mathcal{G}_c(n)$, but $\mathcal{G}_{\mathrm{MSC}}(n)$ is not contained in $\mathcal{G}_{\mathrm{ISC}}(n)$ or vice versa.

The cardinalities of these classes (for $n = 3, 4, 5$) are presented in Table 4.1. It shows that the class of graphs with a correct ear decomposition is a large subset of $\mathcal{G}^*(n)$, but the ratio $|\mathcal{G}_c(n)|/|\mathcal{G}^*(n)|$ decreases as n grows.

CHAPTER 5

Conclusions

In this chapter we give an overview of the previous chapters and of the results that we obtained. Section 5.2 discusses how these results can be applied to pharmacokinetic models, and we conclude with some suggestions for future research in Section 5.3

5.1. Overview

Identifiability concerns the possibility of uniquely determining the parameters of a given model from input-output data. We considered generic local identifiability, meaning that sufficiently general parameter values are at least locally a unique solution to the parameter estimation problem. A question that arises is what to do when a model is unidentifiable. Therefore, we studied a specific class of linear compartment models, which are known to be unidentifiable [**MS14**]. This class consists of all models that can be represented by a strongly connected graph with a unique input-output compartment, and satisfy the assumption that every compartment has an outflow of material to the environment. An input-output map was presented: the double characteristic polynomial map. We saw that the dimension of the image of this map is smaller than the dimension of the parameter space. Thus, the double characteristic polynomial map is not injective and the model parameters can never be identifiable.

To deal with unidentifiable models, we considered the existence of an identifiable scaling reparametrization. This is a map which reduces the number of parameters by scaling the state variables, such that the model becomes identifiable. If an identifiable scaling reparametrization exists, then the model is said to have the expected dimension. In Chapter 3 we considered the following quesion:

Question 1.4.1. *For which graphs satisfying our assumptions does there exist an identifiable scaling reparametrization?*

A criterion to test whether a graph has the expected dimension was presented in [**MS14**]. This criterion was based on the dimension of the image of the double characteristic polynomial map. Using basic algebraic geometry and graph theory, we translated this criterion to an alternative criterion which is easier and more efficient to evaluate. This was our main result, stated in Theorem 1.4.2. The criterion is based on the rank of the bi-adjacency matrix of the bipartite graph defined in equation (3.6).

The two criteria lead to algorithms to test whether a graph has the expected dimension. Both criteria consider the rank of a matrix, namely the differential $d_A c$ of the double characteristic polynomial map and the bi-adjacency matrix $B(G)$. The differential map is of size $(2n-1) \times (n+m)$ while the matrix $B(G)$ has size $(n^2-n-m) \times (n-1)(n-2)$. On the other hand, the entries of $d_A c$ are polynomials in a_{ij} of degree n, while the entries of $B(G)$ are linear combinations of parameters, hence of degree 1. A complexity analysis showed that both criteria lead to a probabilistic algorithm that runs in polynomial time. The algorithm that considers the differential has an asymptotic complexity of $O(n^5)$ operations, while

the algorithm that considers the bi-adjacency matrix has an asymptotic complexity of $O(n^6)$ operations. Computational results showed that for small n the second algorithm is more efficient, in which case our new criterion reduces the costs of determining whether a graph has the expected dimension.

Thus, we have derived a new criterion to answer Question 1.4.1. The second question we considered is the following:

QUESTION 1.4.3. *What constructions can we apply to a given graph, such that the resulting graph has an identifiable scaling reparametrization?*

This question was treated in Chapter 4. Theorem 1.4.2 led us to some new constructions that can be applied to a graph with the expected dimension, such that the resulting graph has the expected dimension as well. Our first result states that if a graph can be written as a union of two subgraphs with only one vertex in common, then this graph has the expected dimension if both subgraphs have the expected dimension. We saw that this is a sufficient condition, but not necessary. Our second result states that if a graph has the expected dimension, then after adding a line segment of length at least two, the resulting graph still has the expected dimension. This is a very useful construction, because starting from the trivial graph with only one vertex, one can build many graphs simply by adding line segments. We defined what it means for a graph to have a correct ear decomposition, which led to our main result of Chapter 4: if a graph has a correct ear decomposition (i.e. an ear decomposition without trivial ears and whose initial cycle contains vertex 1), then it has the expected dimension. This was stated in Theorem 1.4.4.

Finally, we observed that Theorem 1.4.4 allows us to transform any graph into one that has the expected dimension: find an ear decomposition with the least number of trivial ears, then adjust the graph such that the trivial ears disappear. This can be done either by deleting the edges corresponding to the trivial ears, or by subdividing those edges.

In conclusion, we derived a new criterion for the existence of an identifiable scaling reparametrization and proved several valid constructions. These results bring us a few steps towards our goal to obtain a full classification of graphs which have an identifiable scaling reparametrization. The computational results in Section 4.5 show that the class of graphs which have a correct ear decomposition does not contain all graphs which have the expected dimension. Furthermore, we saw that the ratio of such graphs decreases as n grows. Thus, there are more constructions to be discovered, as discussed in Section 5.3.

5.2. Applications

Section 1.2 covered some applications of linear compartment models, with a focus on pharmacokinetic models. We will briefly revisit these models, and apply the results that were obtained in this thesis.

Pharmacokinetic models often have a mammillary or catenary structure, as depicted in Figure 5.1. For example, the 3-compartment model that was presented in Figure 1.2b has a mammillary structure. From the results in Chapter 4 it follows that these models have the expected dimension. In other words, there exists an identifiable scaling reparametrization. The full body model that was discussed in the introduction of this thesis (Figure 1.2a) can be seen to have a correct ear decomposition and therefore it has the expected dimension. Note that the fact that an identifiable scaling reparametrization exists doest not mean that it is appropriate; there may be situations in which the model as originally defined has to be studied.

We have seen two ways to transform a model which does not have the expected dimension into one that does have the expected dimension: either remove the trivial

(A) Mamillary structure (B) Catenary structure

FIGURE 5.1. Pharmacokinetic model structures

ears, or subdivide the corresponding edge. When constructing a model, the latter is the more natural operation to apply. Indeed, an edge represents the flow of material, and deleting an edge corresponds to ignoring part of the flow, and this would give an incorrect model. On the other hand, subdividing an edge corresponds to adding an extra compartment in between two existing compartments, which is a much more natural way to adapt the model. Of course, the additional compartment should also have a leak, in order to satisfy Assumption 1.1.4.

5.3. Future work

In Chapter 4 it was shown that adding a nontrivial ear to graph is a valid operation, i.e. if the input graph has the expected dimension, then so does the resulting graph. This allows us to construct many graphs which have the expected dimension, but not all. A remaining problem is to find out what distinguishes the graphs which have the expected dimension but no correct ear decomposition, from the graphs which do not have the expected dimension. More valid constructions remain to be discovered, which will eventually lead to a full classification of graphs which have an identifiable scaling reparametrization. In Section 4.3 the following conjecture was formulated, which would be a nice starting point for future research.

CONJECTURE 4.3.6. *Let G be a graph on $n-1$ vertices and let $k \to l$ be an edge in G. Construct the graph G' on n vertices subdividing the edge $k \to l$, i.e. by adding vertex n to G and replacing the edge $k \to l$ by two edges $k \to n, n \to l$. Then if G has the expected dimension, G' has the expected dimension as well.*

In this thesis we only considered a specific class of models, so the next step will be to consider more general models. For example, what can we say about the case where the input and output do not take place in the same compartment? Suppose the input takes place in compartment 1, while the output takes place in compartment 2. This affects the input-output equation we found in Section 2.2 and hence also the double characteristic polynomial map. Similar to Theorem 2.2.1, the input-output equation becomes

$$\det(\partial I_n - A)y = \det(\partial I_{n-1} - A_2)u,$$

where A_2 denotes the matrix obtained from A by removing its first row and its second column. This equation gives rise to a coordinate map c', analogous to the definition of the double characteristic polynomial map c in Section 2.2. It would be interesting to see if we can apply a similar analysis to the coordinate map c' as we did for c.

Another extension would be to consider multiple input-output compartments. In this case we also obtain a different, more complicated input-output equation. The techniques that were used in this thesis cannot be applied directly, but the ideas might still be useful.

A final interesting question is what happens if not every compartment has a leak. In this case the input-output equations remain the same, but the parameters are no longer independent. In other words, the dimension of the parameter space Θ_G is smaller than $m+n$. Therefore, such a model might be identifiable, in which case a reparametrization is no longer required. If it is not identifiable, do the obtained results concerning the existence of an identifiable scaling reparametrization still hold?

APPENDIX A

Varieties and tangent spaces

In this appendix we state some basic concepts from algebraic geometry which are used in the text. The material we present comes from [**CLO07**], [**Har92**] and [**Sha13**].

Let $\mathbb{K}$ be an algebraically closed field, then an *affine variety* is the set of solutions of a system of polynomial equations over $\mathbb{K}$. When $\mathbb{K} = \mathbb{C}$ we speak of a *complex variety*. In this thesis we only consider complex varieties, but the results stated here hold in general.

Definition A.1. Given a set of polynomials $f_1, \ldots, f_s \in \mathbb{K}[x_1, \ldots, x_n]$, the *affine variety* $X = \mathbf{V}(f_1, \ldots, f_s)$ is defined as

$$\mathbf{V}(f_1, \ldots, f_s) = \{(a_1, \ldots, a_n) \in \mathbb{K}^n \mid f_i(a_1, \ldots, a_n) = 0 \text{ for } i = 1, \ldots, s\}.$$

This means that a variety X can be represented by an ideal $I = (f_1, \ldots, f_s) \in \mathbb{K}[x_1, \ldots, x_n]$, which is called the implicit representation of X. Conversely, a variety X defines an ideal $\mathbf{I}(X)$ by

$$\mathbf{I}(X) = \{f \in \mathbb{K}[x_1, \ldots, x_n] \mid f \equiv 0 \text{ on } X\}.$$

It is clear that $I \subseteq \mathbf{I}(\mathbf{V}(I))$, but equality doesn't necessarily hold. If so, I is called a radical ideal.

The topology that is used on algebraic varieties is the *Zariski topology*. On a variety X, the Zariski topology is defined by its closed sets, which are taken to be the common zero loci of polynomials on X. In other words, a set is closed if it is of the form $\mathbf{V}(S)$ for some set S of polynomials in $\mathbb{K}[x_1, \ldots, x_n]$. A set is open if it is the complement of a closed set.

The Zariski closure of a subset $S \subseteq X$ is the smallest Zariski closed set containing S, which is $\mathbf{V}(\mathbf{I}(S))$. A subset $S \subseteq X$ is called *Zariski dense* in X if $X = \overline{S}$, where $\overline{S}$ denotes the closure of S. For $\mathbb{K}$ algebraically closed, any nonempty open set is dense in the Zariski topology. This fact allows us to conclude that if some Zariski open property holds at a certain point on X, then it must hold almost everywhere on X.

The tangent space to an affine variety X at a point p consists of all lines through x which are tangent to X. Given a variety X with ideal $\mathbf{I}(X) = (f_1, \ldots, f_s)$, define the $s \times n$ matrix $M = (\partial f_i / \partial f_j)$, where $i = 1, \ldots, s$ and $j = 1, \ldots, n$. Then the *Zariski tangent space* to X at p, denoted $T_p(X)$, is defined as the kernel of M. In general, the dimension of $T_p(X)$ is at least the dimension of X itself, but it may be larger. When equality holds, X is said to be *smooth* at p. The set of smooth points of a variety is dense, i.e. a variety is smooth almost everywhere.

Definition A.2. Let $U \subset X$ be an open set and $p \in U$ a point. A function f on U is called *regular at* p if in a neighborhood of p it can be written as a quotient g/h of polynomials $g, h \in \mathbb{K}[x_1, \ldots, x_n]$ with $h(p) \neq 0$. Moreover, if f is regular at every point of U, then f is called *regular on* U.

Any regular map $f : X \to Y$ induces a map $df : T_p(X) \to T_{f(p)}(Y)$, the *differential* of f. The differential can be obtained by determining the coefficient of ε in $f(x + \varepsilon v)$. This follows from the fact that the Taylor series of a function f around x is of the form

$$f(x + \varepsilon v) = f(x) + (d_x f)(v)\varepsilon + \dots$$

where $x, v \in X$ and $\varepsilon \in \mathbb{R}$. To simplify calculations, note that it is sufficient to determine the coefficient of ε in $f(x + \varepsilon v)$ modulo ε^2. In Section 3.1 this is used to determine the differential of the conjugation map $\psi : g \mapsto gAg^{-1}$ at the identity.

DEFINITION A.3. A regular map $f : X \to Y$ is called *dominant* if its image is dense in Y.

The following proposition can be used to determine the dimension of the image of f, and plays an important role in Chapter 3.

PROPOSITION A.4. *Let $f : X \to Y$ be a dominant regular map of varieties defined over a field K of characteristic 0. Then there exists a nonempty open subset $U \subset Y$ such that for any smooth point $p \in f^{-1}(U)$ in the inverse image of U, the differential $d_p f$ is surjective. In other words,* $\operatorname{rk}(d_p f) = \dim Y$ *for* $p \in U$.

PROOF. A proof of this proposition can be found in [**Har92**]. □

Bibliography

[BÅ70] R. Bellman and K. Åström, *On structural identifiability*, Math. Biosci. **7** (1970), 329–339.

[BJG07] J. Bang-Jensen and G. Gutin, *Digraphs: Theory, algorithms and applications*, Springer, New York, 2007.

[BSAD07] G. Bellu, M. P. Saccomani, S. Audoly, and L. D'Angiò, *Daisy: a new software tool to test global identifiability of biological and physiological systems*, Comput. Meth. Programs Biomed. **88** (2007), 52–61.

[CG85] M. J. Chapman and K. R. Godfrey, *Some extensions to the exhaustive-modeling approach to structural identifiability*, Math. Biosci. **77** (1985), no. 1-2, 305–323.

[CG98] M. J. Chappell and R. N. Gunn, *A procedure for generating locally identifiable reparameterisations of unidentifiable non-linear systems by the similarity transformation approach*, Math. Biosci. **148** (1998), no. 1, 21–41.

[CKL13] H. Y. Cheung, T. C. Kwok, and L. C. Lau, *Fast matrix rank algorithms and applications*, J. ACM **60** (2013), no. 5, Art. 31, 25.

[CLO07] D. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra*, third ed., Undergraduate Texts in Mathematics, Springer, New York, 2007.

[CP08] G. Craciun and C. Pantea, *Identifiability of chemical reaction networks*, J. Math. Chem. **44** (2008), 244–259.

[EC00] N. D. Evans and M. J. Chappell, *Extensions to a procedure for generating locally identifiable reparameterisations of unidentifiable systems*, Math. Biosci. **168** (2000), no. 2, 137–159.

[EL78] M. Espana and I. D. Landau, *Reduced order bilinear models for distillation columns*, Automatica **14** (1978), 345–355.

[God83] K. R. Godfrey, *Compartmental models and their application*, Academic Press, Inc., London, 1983.

[Hal79] E. Halfon, *Theoretical systems ecology: Advances and case studies*, Academic Press, Inc., New York, 1979.

[Har92] J. Harris, *Algebraic geometry*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1992.

[JRY13] H. M. Jones and K. Rowland-Yeo, *Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development*, CPT: Pharmacometrics & Systems Pharmacology **2** (2013).

[LG94] L. Ljung and T. Glad, *On global identifiability for arbitrary model parametrizations*, Automatica J. IFAC **30** (1994), no. 2, 265–276.

[MAD11] N. Meshkat, C. Anderson, and J. J. DiStefano, III, *Finding identifiable parameter combinations in nonlinear ODE models and the rational reparameterization of their input-output equations*, Math. Biosci. **233** (2011), no. 1, 19–31.

[MED09] N. Meshkat, M. Eisenberg, and J. J. DiStefano, III, *An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner bases*, Math. Biosci. **222** (2009), no. 2, 61–72.

[MS14] N. Meshkat and S. Sullivant, *Identifiable reparametrizations of linear compartment models*, J. Symbolic Comput. **63** (2014), 46–67.

[Poh78] H. Pohjanpalo, *System identifiability based on the power series expansion of the solution*, Math. Biosci. **41** (1978), no. 1-2, 21–33.

[RKS+14] A. Raue, J. Karlsson, M. P. Saccomani, M. Jirstrand, and J. Timmer, *Comparison of approaches for parameter identifiability analysis of biological systems*, Bioinformatics **30** (2014), no. 10, 1440–1448.

[Sch80] J. T. Schwartz, *Fast probabilistic algorithms for verification of polynomial identities*, J. Assoc. Comput. Mach. **27** (1980), no. 4, 701–717.

[Sha13] I. R. Shafarevich, *Basic algebraic geometry. 1*, Springer, Heidelberg, 2013.

[SJH97] W. Slob, P. H. M. Janssen, and J. M. van den Hof, *Structural identifiability of pbpk models: practical consequences for modeling strategies and study designs*, Critical reviews in Toxicology **27** (1997), 261–272.

[Yat06] J. W. T. Yates, *Structural identifiability of pysiologically based pharmacokinetic models*, Journal of Pharmacokinetics and Pharmacodynamics **33** (2006), 421–439.

[Zip79] R. Zippel, *Probabilistic algorithms for sparse polynomials*, Symbolic and algebraic computation (EUROSAM '79, Internat. Sympos., Marseille, 1979), Lecture Notes in Comput. Sci., vol. 72, Springer, Berlin-New York, 1979, pp. 216–226.