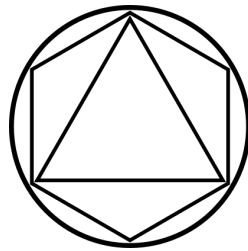




Technische Universität München

Department of Mathematics



Master's Thesis

# Goodness-of-fit tests for non-Gaussian linear causal models

Daniela Schkoda

Supervisor: Prof. Mathias Drton

Submission Date: 31.10.2022

I hereby declare that this thesis is my own work and that no other sources have been used except those clearly indicated and referenced. This thesis was not previously presented to another examination board and has not been published.

Munich, 31.10.2022

# Zusammenfassung auf Deutsch

Das Rückschließen auf kausale Zusammenhänge zwischen Variablen nur auf Grundlage von Beobachtungsdaten ist ein zentrales Problem in vielen wissenschaftlichen Bereichen. Es wurden zahlreiche Algorithmen für diese Aufgabe entwickelt, die von verschiedenen zugrundeliegende Annahmen Gebrauch machen. Ein prominentes Beispiel ist die Annahme, dass die gemeinsame Verteilung der beobachteten Variablen einem linearen nicht-gaußschen Strukturgleichungsmodell folgt.

In dieser Arbeit zeigen wir eine Charakterisierung dieser Linearitätsannahme. Die Annahme ist äquivalent zu einer Rangbedingung für eine Matrix, die aus zweiten und dritten Momenten gebildet wird, zusammen mit der Anforderung, dass der Tensor aller dritten Momente einen bestimmten symmetrischen Rang hat. Diese beiden Rangbedingungen liefern somit einen neuen Ansatz zur Validierung der Hypothese, dass die datenerzeugende Verteilung einem linearen Strukturgleichungsmodell folgt.

Für beide Bedingungen wird untersucht, wie sie in statistische Tests umgesetzt werden können. Um einen Test für die erste Bedingung zu implementieren, betrachten wir eine Multiplikator-Bootstrap-Methode, die unvollständige  $U$ -Statistiken verwendet, um Minoren der Matrix zu schätzen. Außerdem werden Methoden untersucht, die auf dem asymptotischen Verhalten der Singulärwerte basieren. Für die zweite Rangbedingung, die den Rang des symmetrischen Tensors betrifft, werden Ergebnisse aus der algebraischen Geometrie genutzt, um den Rang des symmetrischen Tensors mit Polynomgleichungen in den Einträgen des Tensors in Beziehung zu setzen. Zum Testen dieser Polynomgleichungen wird wieder die unvollständige  $U$ -Statistik verwendet.

Die Methoden werden für die Tübingen Sammlung von Benchmark-Datensätzen aus Ursache-Wirkungs-Paaren und für synthetische Daten illustriert.

# Zusammenfassung auf Englisch

Inferring causal relationships between variables solely from observational data is a central question in many scientific fields. Various algorithms have been developed to tackle this problem by leveraging different types of a priori assumptions. One prominent example is the assumption that the joint distribution of the observed variables follows a linear non-Gaussian structural equation model.

In this thesis, we derive a characterization of this linearity assumption. Specifically, the assumption is shown to be equivalent to a rank condition on a matrix formed from second and third moments together with the requirement that the tensor of all third moments has a certain symmetric rank. Testing those two rank conditions thus offers a new approach to test the hypothesis that the data-generating distribution belongs to a linear structural equation model.

For both conditions, we examine how to turn them into statistical tests. In order to implement a test for the first constraint, we consider a multiplier bootstrap method that uses incomplete  $U$ -statistics to estimate subdeterminants of the matrix. In addition, methods that employ asymptotic approximations of the null distribution of singular values are studied. For the second constraint involving the symmetric tensor rank, results from algebraic geometry are leveraged that relate the symmetric tensor rank to polynomial equations in the entries of the tensor. For testing these polynomial equations, the incomplete  $U$ -statistic is applied again.

The methods are illustrated for the Tübingen collection of benchmark data sets on cause-effect pairs as well as for synthetic data.

## **Acknowledgements**

I want to thank my supervisor Mathias Drton for giving me the opportunity to work on such an interesting topic, for his input and constant support throughout the thesis. I am grateful for all the discussions and the substantial amount of time he invested.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Identifiability for Linear Non-Gaussian Acyclic Models</b>	<b>3</b>
<b>3</b>	<b>Graphical Model associated with the Complete Graph</b>	<b>6</b>
3.1	Second and Third Moments . . . . .	6
3.2	Symmetric Tensor Rank . . . . .	14
3.3	Higher Moments . . . . .	21
3.4	Conditions for one Missing Edge . . . . .	22
<b>4</b>	<b>Statistical Tests</b>	<b>24</b>
4.1	Singular Value Based Approaches . . . . .	25
4.1.1	Kleibergen-Paap Test . . . . .	26
4.1.2	Bootstrap Statistic . . . . .	29
4.2	Tests Using Minors . . . . .	33
4.2.1	Incomplete $U$ -Statistic . . . . .	33
4.2.2	Independent Test . . . . .	39
<b>5</b>	<b>Numerical Experiments</b>	<b>40</b>
5.1	Synthetical Data . . . . .	40
5.1.1	Null Hypothesis Setup Non-Degenerate Case . . . . .	40
5.1.2	Null Hypothesis Setup Degenerate Case . . . . .	45
5.1.3	Power Analysis . . . . .	47
5.2	Tübingen Dataset . . . . .	49
<b>6</b>	<b>Conclusion and Outlook</b>	<b>55</b>

# 1 Introduction

In various scientific fields, it is of interest to answer questions regarding causation between multiple observed attributes. The standard approach to tackle such questions is to use controlled interventional experiments. However, in many situations, controlled experiments are too expensive, unethical, or not feasible at all. Therefore, it is not surprising that in recent years, many algorithms have been developed to infer causal relationships from observational data.

A common way how to depict such causal relations is to use a directed graph  $\mathcal{G} = (V, E)$  where  $V$  is the set of nodes and  $E \subseteq V \times V$  is the set of directed edges. Before explaining how a graph can be used to represent causal relations, some basic concepts of graph theory are introduced. We write  $v \rightarrow w$  for an edge leading  $v$  to  $w$ . Furthermore, we say that  $w$  is a parent of  $v$  if the graph contains the edge  $v \rightarrow w$  and denote the set of all parents of a node  $v$  by  $\mathbf{PA}_v$ . A directed cycle is a sequence of nodes  $(v_1, v_2, \dots, v_k)$  with  $v_k = v_1$  such that for each  $j = 1, \dots, k-1$ ,  $v_j$  and  $v_{j+1}$  are connected by the edge  $v_j \rightarrow v_{j+1}$ . Finally,  $\mathcal{G}$  is called acyclic if it contains no directed cycles. From now on, for a graph with  $p$  nodes, we use the set  $V = [p] := \{1, \dots, p\}$  to label the nodes.

To establish a connection with cause-effect relationships, each node of the graph is associated with one component of a random vector. In other words, one considers a random vector  $\mathbf{X} = (X_j)_{j \in [p]}$  indexed by the set of nodes. We denote the distribution of  $\mathbf{X}$  as  $P^{\mathbf{X}}$ . An edge  $v \rightarrow w$  describes a direct causal effect of  $X_v$  on  $X_w$ . One then aims to learn which graph depicts the true causal structure of the distribution. There are several ways for formally defining *direct causal effects*. Many algorithms rely on a definition that involves structural equation models. The distribution  $P^{\mathbf{X}}$  is said to belong to the *structural equation model* (SEM) associated with the graph  $\mathcal{G}$  if each component of  $\mathbf{X}$  can be written as a function of its parents plus an error term. Formally, this means there are functions  $f_j$  and independent random variables  $\epsilon_j$  such that each  $X_j$  can be written as

$$X_j = f_j(\mathbf{PA}_j) + \epsilon_j.$$

Many algorithms require the additional assumption that  $\mathcal{G}$  is acyclic.

If one wants to learn the associated graph of a distribution, the question arises if the graph is uniquely determined. First of all, if  $P^{\mathbf{X}}$  lies in the SEM of a graph  $\mathcal{G}$ , it also lies in the SEM of every supergraph of  $\mathcal{G}$ . This can be easily resolved by aiming to learn the smallest graph such that  $P^{\mathbf{X}}$  is contained in the associated model. So, it remains to examine whether  $P^{\mathbf{X}}$  could belong to the models of two different graphs  $\mathcal{G}_1, \mathcal{G}_2$  that are not subgraphs of each other. It turns out that under some restrictions on the functions and error terms this is not possible. For example, if the functions are restricted to be linear, the error terms are assumed to follow a non-Gaussian distribution, and the graph is assumed to be acyclic, it has been shown that it is not possible for two equations of the form to hold at the same time [26]. This model is abbreviated as LiNGAM, which stands for linear non-Gaussian acyclic model. Another possible restriction is the post-nonlinear causal model where one assumes

$$X_j = f_{j,1}(f_{j,2}(\mathbf{PA}_j) + \epsilon_j)$$

with  $f_{j,1}$  non-linear and invertible, and  $f_{j,2}$  non-linear for an acyclic graph  $\mathcal{G}$ . In this case, identifiability holds except for five specific choices of the functions and the noise terms [32].

Various algorithms have been developed that are based on one of these identifiability results. They generally work in a way that they initially assume the existence of some directed acyclic graph such that  $P^X$  lies in the model associated with this graph. Then they infer the graph such that the data fits best to the respective graphical model. Consequently, if the initial assumption that there is some appropriate graph is wrong, the results of the algorithm might be misleading.

Especially, the LiNGAM assumption is quite restrictive. Nonetheless, algorithms based on this model, as for example, ICA-LiNGAM [26], Direct LiNGAM [27], or the modified version of Direct LiNGAM for the high dimensional setting [31], are very widely used. This is the motivation for this thesis: to develop a test for this assumption. More specifically, a test for the linearity assumption is created. Hence, we allow the noise term to be Gaussian. There already exist tests specifically designed for testing the Gaussianity assumption as can be found for instance in [25].

**Outline.** The outline of the thesis is as follows. In Section 2, the identifiability result for the LiNGAM model is proven. In Section 3, the theoretical foundation for the testing methods is laid out. More precisely, the algebraic structure of the moments of distributions contained in a linear SEM is examined. We derive a characterization of the second and third moments such distributions can attain. As this characterization involves the concept of a symmetric tensor rank, results regarding this concept are presented subsequently. In the following section, these theoretical results are leveraged to develop algorithms to test if a distribution belongs to a linear SEM. Four different test statistics and their asymptotic behavior are illustrated. In the last section, the practical performance of the methods is examined using synthetic and real-world data.

**Notation.** In this thesis, random vectors are always denoted with bold letters, whereas random variables are denoted with non-bold letters.  $P^X$  denotes the joint distribution of a random vector  $\mathbf{X}$ . Furthermore, we write

$$(\mathbb{R}^p)^{\otimes k} = \{T = (t_{i_1 \dots i_k})_{i_1, \dots, i_k \in [p]}\}$$

for the  $k$ -fold tensor product of  $\mathbb{R}^p$  and  $\text{Sym}_k(\mathbb{R}^p)$  for the subspace of symmetric tensors

$$\{T \in (\mathbb{R}^p)^{\otimes k} : t_{i_1 \dots i_k} = t_{\pi(i_1) \dots \pi(i_k)} \text{ for all permutations } \pi : [p] \rightarrow [p]\}.$$



## 2 Identifiability for Linear Non-Gaussian Acyclic Models

In this section, we formalize and prove the identifiability result for linear SEMs with non-Gaussian noise. Let  $\mathcal{G} = ([p], E)$  be a directed graph with  $p$  nodes and denote by

$$\mathbb{R}^E = \{\Lambda \in \mathbb{R}^{p \times p} : (I - \Lambda) \text{ is invertible and } \lambda_{ji} \neq 0 \text{ only if } (i, j) \in E\}$$

the set of matrices whose non-zero entries correspond to the edges of  $\mathcal{G}$ . Note that the requirement that  $I - \Lambda$  is invertible is obsolete if  $\mathcal{G}$  is acyclic. In this case, it is possible to reorder the nodes  $1, \dots, p$  in a way such that  $I - \Lambda$  is an upper triangular matrix with ones on the diagonal. Reordering the nodes only permutes the rows and columns of  $I - \Lambda$ , which is an operation that retains invertibility.

The linear structural equation model (SEM) for  $\mathcal{G}$  is the set of all probability distributions  $P^X$  on  $\mathbb{R}^p$  that arise as joint distributions of random vectors  $\mathbf{X} = (X_1, \dots, X_p) \sim P^X$  that solve a linear structural equation

$$\mathbf{X} = \Lambda^T \mathbf{X} + \boldsymbol{\epsilon} \quad (2.1)$$

for a choice of  $\Lambda \in \mathbb{R}^E$  and a random vector  $\boldsymbol{\epsilon}$  with independent components. Furthermore, the model is called linear non-Gaussian acyclic model, or abbreviated LiNGAM, if  $\mathcal{G}$  is acyclic, and the distribution of all components of  $\boldsymbol{\epsilon}$  is non-Gaussian. The entries of  $\Lambda$  can be thought of as edge weights. The larger  $\lambda_{ji}$ , the higher is the effect of  $X_i$  on  $X_j$ . For the remainder of this thesis, we assume that the noise vector  $\boldsymbol{\epsilon}$  has mean zero. This assumption makes computations easier without restricting generally since one can replace  $\mathbf{X}$  with its centralized version.

**Theorem 2.1** (Identifiability under LiNGAM assumption, [26]). *If  $P^X$  belongs to the LiNGAM associated with a directed acyclic graph  $\mathcal{G}_0$ , and  $\mathcal{G}_0$  is minimal with this property, then  $\mathcal{G}_0$  is identifiable in the sense that there is no other acyclic graph with this property.*

In order to get a feeling for graphical models and the concept of identifiability, the result is proven for the case  $p = 2$ . For that purpose, the theorem of Darrois-Skitovich is used.

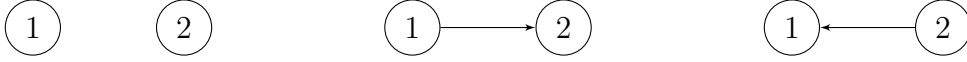
**Theorem 2.2** (Darrois-Skitovich, [8], [28]). *Assume that  $(Z_i)_{i=1, \dots, n}$  are independent random variables and that it is possible to form two linear combinations*

$$\sum_i \alpha_i Z_i, \text{ and } \sum_i \beta_i Z_i,$$

*that are independent of each other. Then if  $\alpha_i \beta_i = 0$ ,  $Z_i$  is Gaussian.*

The proof of Theorem 2.1 now reads as follows.

*Proof.* To obtain a contradiction, assume that there exist a distribution  $P^X$ , and two directed acyclic graphs  $\mathcal{G}_1, \mathcal{G}_2$  such that  $P^X$  lies in the LiNGAM of each of the graphs. Furthermore, we assume that both graphs are minimal with this property. There are only three directed acyclic graphs with 2 nodes.



If one of the graphs  $\mathcal{G}_1, \mathcal{G}_2$  would be the empty graph, this would contradict the minimality assumption. Hence, we can assume without loss of generality that  $\mathcal{G}_1$  is the graph with edge (1, 2) and  $\mathcal{G}_2$  the graph with edge (2, 1). In particular,  $\mathbf{X}$  satisfies a system of structural equations for  $\mathcal{G}_1$ . This yields the existence of a real number  $\alpha \neq 0$  and independent noise terms  $\epsilon_1, \epsilon_2$  such that

$$X_1 = \epsilon_1, \quad X_2 = \alpha X_1 + \epsilon_2.$$

At the same time, a system for  $\mathcal{G}_2$  is fulfilled. Hence, we obtain the equations

$$X_1 = \beta X_2 + \eta_1, \quad X_2 = \eta_2$$

for a real number  $\beta \neq 0$  and independent noise terms  $\eta_1, \eta_2$ . When inserting the equations into one another, one obtains

$$\eta_1 = X_1 - \beta X_2 = \epsilon_1 - \beta(\alpha\epsilon_1 + \epsilon_2) = (1 - \beta\alpha)\epsilon_1 - \beta\epsilon_2 \quad \text{and} \quad \eta_2 = X_2 = \alpha\epsilon_1 + \epsilon_2.$$

Darmois-Skitovich then implies that  $\epsilon_2$  is Gaussian in contradiction to the assumption.  $\square$

**Remark 2.3.** In the case where the distribution of  $\mathbf{X}$  has finite second and third moments, the result can also be proven more directly. Under the model assumption of  $\mathcal{G}_1$ , one knows that

$$\begin{aligned} \mathbb{E}(X_1^2) &= \mathbb{E}(\epsilon_1^2), \\ \mathbb{E}(X_1 X_2) &= \mathbb{E}(\epsilon_1(\alpha\epsilon_1 + \epsilon_2)) = \alpha\mathbb{E}(\epsilon_1^2), \end{aligned}$$

where the independence of the noise terms is used in the last equivalence. Similarly,

$$\begin{aligned} \mathbb{E}(X_1^3) &= \mathbb{E}(\epsilon_1^3), \\ \mathbb{E}(X_1^2 X_2) &= \mathbb{E}(\epsilon_1^2(\alpha\epsilon_1 + \epsilon_2)) = \alpha\mathbb{E}(\epsilon_1^3). \end{aligned}$$

Combining all four equations yields

$$\mathbb{E}(X_1^2)\mathbb{E}(X_1^2 X_2) - \mathbb{E}(X_1 X_2)\mathbb{E}(X_1^3) = 0.$$

In other words, the determinant of

$$\begin{pmatrix} \mathbb{E}(X_1^2) & \mathbb{E}(X_1 X_2) \\ \mathbb{E}(X_1^3) & \mathbb{E}(X_1^2 X_2) \end{pmatrix} \tag{2.2}$$

is zero. If the distribution of  $\mathbf{X}$  belongs to the model associated with  $\mathcal{G}_2$ , in the same manner, one can calculate

$$\begin{aligned}\mathbb{E}(X_1^2) &= \mathbb{E}((\beta\eta_2 + \eta_1)^2) = \beta^2\mathbb{E}(\eta_2^2) + \mathbb{E}(\eta_1^2), \\ \mathbb{E}(X_1X_2) &= \mathbb{E}((\beta\eta_2 + \eta_1)\eta_2) = \beta\mathbb{E}(\eta_2^2), \\ \mathbb{E}(X_1^3) &= \mathbb{E}((\beta\eta_2 + \eta_1)^3) = \beta^3\mathbb{E}(\eta_2^3) + \mathbb{E}(\eta_1^3), \\ \mathbb{E}(X_1^2X_2) &= \mathbb{E}((\beta\eta_2 + \eta_1)^2\eta_2) = \beta^2\mathbb{E}(\eta_2^3).\end{aligned}$$

Hence, (2.2) has determinant zero only for specific choices of  $\boldsymbol{\eta}$  and  $\beta$ . More precisely, one obtains

$$\begin{pmatrix} \mathbb{E}(X_1^2) & \mathbb{E}(X_1X_2) \\ \mathbb{E}(X_1^3) & \mathbb{E}(X_1^2X_2) \end{pmatrix} = \begin{pmatrix} \mathbb{E}(\eta_1^2) & \mathbb{E}(\eta_2^2) \\ \mathbb{E}(\eta_1^3) & \beta\mathbb{E}(\eta_2^3) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \beta^2 & \beta \end{pmatrix}.$$

The second factor is invertible. Therefore, the determinant of (2.2) vanishes only in the special situation where first factor is non-invertible. In all other cases, the determinant gives a possibility to distinguish between the two models. In alignment with Theorem 2.1, Gaussian distributions are one example of the exceptional situation in that identifiability can not be concluded from this argument since, for Gaussian distributions, all third moments vanish.

### 3 Graphical Model associated with the Complete Graph

The objective of this section is to develop a theoretical foundation that can be used for developing the desired tests of the linearity assumption. In the first subsection, the second and third moments of distributions that lie in any linear structural equation model are examined. In particular, in our analysis, we allow graphs containing cycles. We derive a characterization of second and third moments realizable under the model assumption. This characterization involves the notion of *symmetric tensor rank*. Therefore, the following subsection is devoted to a more detailed study of symmetric tensors and the symmetric tensor rank. After this, we focus on higher moments. In the last subsection, it is investigated which constraints hold if a certain edge is not allowed to be part of the graph.

#### 3.1 Second and Third Moments

To slightly simplify notation, we from now on use the linear structural equation (2.1) in its equivalent form

$$\mathbf{X} = B^{-T} \boldsymbol{\epsilon} \quad (3.1)$$

with  $B = I - \Lambda$ . The entries of  $B^{-1}$  are denoted by  $B^{-1} = \left( \tilde{\beta}_{ij} \right)_{i,j \in [p]}$ .

Furthermore, we denote by  $s_{ij}$  the covariance of  $X_i$  and  $X_j$ , by  $t_{ijk}$  the third moment of  $X_i, X_j$  and  $X_k$ , and denote by  $S = (s_{ij})_{i,j \in [p]}$  and  $T = (t_{ijk})_{i,j,t \in [p]}$  the matrix of the second moments and the tensor containing the third moments, respectively. Similarly, let  $\Omega^{(2)}$  the matrix of second moments of  $\boldsymbol{\epsilon}$  and  $\Omega^{(3)}$  the tensor of third moments of  $\boldsymbol{\epsilon}$ . Note that both are diagonal because the components of  $\boldsymbol{\epsilon}$  are independent and have mean zero. As we are interested in the structure of the second and third moments, we also introduce the so-called *third-order moment model* for a graph  $\mathcal{G}$  which is defined as

$$\mathcal{M}^{\leq 3}(\mathcal{G}) = \{(S, T) : P^{\mathbf{X}} \text{ is realizable under a linear SEM associated with } \mathcal{G}\}.$$

For the remainder of this section, we assume that  $\boldsymbol{\epsilon}$  has finite second and third moments. Furthermore, unless specified otherwise,  $\mathcal{G}$  indicates the complete graph, which is the graph containing all edges. Note that considering the linear SEM associated with the complete graph is the same as permitting the distribution to belong to a linear SEM associated with any graph since the model for the complete graph contains all other models.

The main result of this section is a characterization of this third-order moment model for the complete graph by algebraic constraints. The first theorem states a necessary condition.

**Theorem 3.1.** *If  $(S, T)$  lies in  $\mathcal{M}^{\leq 3}(\mathcal{G})$ , then*

a) the matrix

$$M = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} & s_{22} & s_{23} & \cdots & s_{pp} \\ t_{111} & t_{112} & \cdots & t_{11n} & t_{122} & t_{123} & \cdots & t_{1pp} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{11p} & t_{12p} & \cdots & t_{1pp} & t_{22p} & t_{23p} & \cdots & t_{ppp} \end{pmatrix}$$

has at most rank  $p$ .

b) Moreover,  $T$  can be written as

$$T = C \circ C \circ C = \left( \sum_{i=1}^p \gamma_{ia} \gamma_{ib} \gamma_{ic} \right)_{a,b,c \in [p]}, \quad (3.2)$$

for a matrix  $C \in \mathbb{R}^{p \times p}$ .

We first prove part b) of the theorem, which is rather easy to spell out. Then two alternative proofs for part a) are presented. They reveal the structure of the moments in two different ways: The first proof shows that the rows of  $M$  are spanned by  $p$  vectors, while the second proof derives linear equations the rows of  $M$  fulfill. A central tool used in all three proofs is the following lemma stating that the second and third moments of  $\mathbf{X}$  can be written in terms of  $B$  and the moments of  $\epsilon$ .

**Lemma 3.2** ([1, Proposition 2.1]). If  $(S, T)$  is contained in  $\mathcal{M}^{\leq 3}(\mathcal{G})$ , then

$$S = B^{-T} \Omega^{(2)} B^{-1} \quad \text{and} \quad T = \Omega^{(3)} \bullet B^{-1} \bullet B^{-1} \bullet B^{-1}, \quad (3.3)$$

where  $\bullet$  denotes the *Tucker product* which is defined as

$$(U \bullet D \bullet D \bullet D)_{ijk} = \sum_{a,b,c=1}^p u_{abc} d_{ai} d_{bj} d_{ck} \quad \text{for } i, j, k \in [p]$$

for a tensor  $U \in \mathbb{R}^p \otimes \mathbb{R}^p \otimes \mathbb{R}^p$  and a matrix  $D \in \mathbb{R}^{p \times p}$ .

*Proof.* The lemma can be proven by calculating

$$\begin{aligned} s_{ab} &= \mathbb{E}(X_a X_b) \\ &= \mathbb{E}((B^{-T} \epsilon)_a \cdot (B^{-T} \epsilon)_b) \\ &= \mathbb{E} \left( \left( \sum_{i=1}^p \tilde{\beta}_{ia} \epsilon_i \right) \left( \sum_{i=1}^p \tilde{\beta}_{ib} \epsilon_i \right) \right) \\ &= \sum_{i=1}^p \sum_{j=1}^p \tilde{\beta}_{ia} \tilde{\beta}_{jb} \mathbb{E}(\epsilon_i \epsilon_j) \\ &= B^{-T} \Omega^{(2)} B^{-1}. \end{aligned}$$

A similar calculation shows that the equation for  $T$  holds. □

The other way around, given a diagonal positive definite matrix  $\Omega^{(2)}$ , and a diagonal tensor  $\Omega^{(3)}$ , there always exists a random vector  $\epsilon$  with independent components and these second and third moments. Consequently, all pairs of matrices and tensors of the form given in the lemma arise as matrices of seconds and tensors of third moments of distributions contained in the linear SEM. Taken together, we can rewrite  $\mathcal{M}^{\leq 3}(G)$  as

$$\mathcal{M}^{\leq 3}(G) = \{(B^{-T}\Omega^{(2)}B^{-T}, \Omega^{(3)} \bullet B^{-T} \bullet B^{-T} \bullet B^{-T}) : \Omega^{(2)} \in \mathbb{R}^{p \times p} \text{ positive definite and diagonal, } \Omega^{(3)} \in \text{Sym}_3(V) \text{ diagonal, } B = I - \Lambda \text{ for a } \Lambda \in \mathbb{R}^E\}.$$

Part b) of the theorem is an immediate consequence of this lemma.

*First proof of Theorem 3.1 b).* From Lemma 3.2, we know that

$$T = \Omega^{(3)} \bullet B^{-1} \bullet B^{-1} \bullet B^{-1}.$$

As  $\Omega^{(3)}$  is diagonal,

$$\Omega^{(3)} \bullet B^{-1} \bullet B^{-1} \bullet B^{-1} = \sum_{i=1}^p \omega_{iii}^{(3)} \tilde{\beta}_{ci} (\tilde{\beta}_{ai} \tilde{\beta}_{bi}).$$

Thus,  $T$  can be expressed as in (3.2) with  $C$  defined by  $\gamma_{ia} = \sqrt[3]{\omega_{iii}^{(3)}} \tilde{\beta}_{ai}$ . □

Now, we turn to the first proof of part a).

*First proof of Theorem 3.1 a).* Expanding the equations from Lemma 3.2, one obtains

$$s_{ab} = \sum_{i=1}^p \omega_{ii}^{(2)} (\tilde{\beta}_{ai} \tilde{\beta}_{bi}),$$

$$t_{abc} = \sum_{i=1}^p \omega_{iii}^{(3)} \tilde{\beta}_{ci} (\tilde{\beta}_{ai} \tilde{\beta}_{bi})$$

for all  $a, b, c \in [p]$ . Hence, all rows of  $M$  are contained in

$$\text{span} \left( \left\{ \left( \begin{array}{c} \tilde{\beta}_{1i} \tilde{\beta}_{1i} \\ \tilde{\beta}_{1i} \tilde{\beta}_{2i} \\ \vdots \\ \tilde{\beta}_{pi} \tilde{\beta}_{pi} \end{array} \right) : i \in [p] \right\} \right).$$

□

For the second variant of proving the theorem, we need some more notation. Let  $s = (s_{ab})_{a \leq b}$  the vector consisting of all distinct second moments,  $t = (t_{abc})_{a \leq b \leq c}$  the vector containing all distinct third moments, and  $t^{(\nu)}$  the subvector of  $t$  consisting of all  $t_{abc}$  where  $a, b$  or  $c$  is  $\nu$ . So,  $t^{(\nu)T}$  is the  $(\nu + 1)$ th row of  $M$ .

The proof is again based on the relation

$$S = B^{-T} \Omega^{(2)} B^{-1} \quad \text{and} \quad T = \Omega^{(3)} \bullet B^{-1} \bullet B^{-1} \bullet B^{-1}.$$

The diagonality of  $\Omega^{(2)}$  and  $\Omega^{(3)}$  is used in order to derive linear equations of the form

$$A^{(2)} \cdot s = 0 \quad \text{and} \quad A^{(3)} \cdot t = 0.$$

in the second and third moments of  $\mathbf{X}$ . Subsequently, we modify the equations  $A^{(3)} \cdot t = 0$  until we obtain

$$A^{(2)} t^{(\nu)} = 0$$

for all  $\nu \in [p]$ . Hence, for all rows  $a^T$  of  $A^{(2)}$  and all rows  $x$  of  $M$ ,

$$a^T x = 0,$$

which proves the statement.

We begin with inferring the linear equations:

**Lemma 3.3.** With

$$A_{ij,ab}^{(2)} = \sum_{\substack{(\pi_1, \pi_2) \text{ permutation} \\ \text{of } (a,b)}} \beta_{i\pi_1} \beta_{j\pi_2}, \quad i < j, a \leq b$$

and

$$A_{ijk,abc}^{(3)} = \sum_{\substack{(\pi_1, \pi_2, \pi_3) \text{ permutation} \\ \text{of } (a,b,c)}} \beta_{i\pi_1} \beta_{j\pi_2} \beta_{k\pi_3}, \quad i \leq j \leq k \text{ and not } i = j = k, a \leq b \leq c$$

the equation systems

$$A^{(2)} \cdot s = 0 \quad \text{and} \quad A^{(3)} \cdot t = 0 \tag{3.4}$$

hold true.

*Proof.* As the error terms  $\epsilon_i$  are assumed to be independent,  $\Omega^{(2)} = B^T \cdot S \cdot B$  is diagonal. Thus,

$$\omega_{ij}^{(2)} = \sum_{a,b=1}^p \beta_{ia} \beta_{jb} s_{ab} = 0$$

for all  $i \neq j$ . Similarly,

$$\omega_{ijk}^{(3)} = \sum_{a,b,c=1}^p \beta_{ia} \beta_{jb} \beta_{kc} t_{abc} = 0$$

for all  $i, j, k$  such that not  $i = j = k$ . Using the symmetry of the  $s_{ab}$  and  $t_{abc}$ , one obtains

$$\sum_{a \leq b} \sum_{\substack{(\pi_1, \pi_2) \text{ permutation} \\ \text{of } (a, b)}} \beta_{i\pi_1} \beta_{j\pi_2} s_{ab} = 0$$

for all  $i < j$  and

$$\sum_{a \leq b \leq c} \sum_{\substack{(\pi_1, \pi_2, \pi_3) \text{ permutation} \\ \text{of } (a, b, c)}} \beta_{i\pi_1} \beta_{j\pi_2} \beta_{k\pi_3} t_{abc} = 0$$

for all  $i \leq j \leq k$  such that not  $i = j = k$ . This gives exactly the equations (3.4).  $\square$

The next lemma demonstrates how one can combine the equations  $A^{(3)}t = 0$  to obtain that  $a^T t^{(\nu)} = 0$  for all rows  $a^T$  of  $A^{(2)}$  and all  $\nu \in [p]$ .

**Lemma 3.4.** For each  $l \leq m$  and each  $\nu \in [p]$  there exists a  $y^{(l, m, \nu)} \in \mathbb{R}^p$  such that

$$((y^{(l, m, \nu)})^T \cdot A^{(3)})_{ijk} = \begin{cases} a_{jk}^{(l, m)} & \text{if } i = \nu, \\ a_{ik}^{(l, m)} & \text{if } j = \nu, \\ a_{ij}^{(l, m)} & \text{if } k = \nu, \\ 0 & \text{otherwise} \end{cases},$$

where  $a^{(l, m)}$  is the row indexed by  $(l, m)$  of  $A^{(2)}$ .

Note that  $\tilde{a}^{(l, m, \nu)} = (y^{(l, m, \nu)})^T \cdot A^{(3)}$  is zero at all positions  $i, j, k$  not containing  $\nu$ . Consequently, the lemma yields  $a^{(l, m)} t^{(\nu)} = \tilde{a}^{(l, m, \nu)} t = 0$  for all  $\nu$  and all  $l \leq m$ .

*Proof of Lemma 3.4.* Define  $y^{(l, m, \nu)}$  as

$$y_{abc}^{(l, m, \nu)} = \begin{cases} \tilde{\beta}_{\nu k} & \text{if } \{a, b, c\} = \{m, l, k\} \text{ for some } k \in [p], \\ 0 & \text{otherwise} \end{cases}$$

For better readability, the indices  $l, m$ , and  $\nu$  in  $y^{(l, m, \nu)}$  are omitted in the following. We



compute

$$\begin{aligned}
y^T \cdot (A^{(3)})^{(abc)} &= \sum_{i,j,k=1}^p y_{ijk} \cdot A_{ijk,abc}^{(3)} \\
&= \sum_{k=1}^p y_{mlk} \cdot A_{mlk,abc}^{(3)} \\
&= \sum_{k=1}^p y_{mlk} \sum_{\substack{(\pi_1, \pi_2, \pi_3) \text{ permutation} \\ \text{of } (a,b,c)}} \beta_{i\pi_1} \beta_{j\pi_2} \beta_{k\pi_3} \\
&= \sum_{\substack{(\pi_1, \pi_2, \pi_3) \text{ permutation} \\ \text{of } (a,b,c)}} \beta_{i\pi_1} \beta_{j\pi_2} \sum_{k=1}^p \beta_{k\pi_3} y_{mlk} \\
&= \sum_{\substack{(\pi_1, \pi_2, \pi_3) \text{ permutation} \\ \text{of } (a,b,c)}} \beta_{i\pi_1} \beta_{j\pi_2} \sum_{k=1}^p \beta_{k\pi_3} \tilde{\beta}_{\nu k} \\
&= \sum_{\substack{(\pi_1, \pi_2, \pi_3) \text{ permutation} \\ \text{of } (a,b,c)}} \beta_{i\pi_1} \beta_{j\pi_2} \mathbb{1}_{\pi_3=\nu} \\
&= \sum_{\substack{(\pi_1, \pi_2) \text{ permutation} \\ \text{of } (a,b)}} \beta_{i\pi_1} \beta_{j\pi_2},
\end{aligned}$$

which concludes the proof.  $\square$

With this in place, Theorem 3.1 can be shown:

*Second Proof of Theorem 3.1.* If  $(S, T)$  lies in  $\mathcal{M}^{\leq 3}(\mathcal{G})$ , then  $S$  is the matrix of second moments and  $T$  the tensor of third moments of a random vector  $\mathbf{X}$  satisfying  $\mathbf{X} = B^{-T} \boldsymbol{\epsilon}$  for a  $B \in \mathcal{B} := \{B \in \mathbb{R}^{p \times p} : B = I - \Lambda \text{ for a } \Lambda \in \mathbb{R}^E\}$  and a random vector  $\boldsymbol{\epsilon}$  whose components are independent of each other.

We first consider the situation where  $A^{(2)}$  has full rank. In this case, Lemma 3.4 yields that  $M$  has at most rank

$$\# \text{columns of } M - \# \text{rows of } A^{(2)} = \frac{p(p-1)}{2} + p - \frac{p(p-1)}{2} = p.$$

$A^{(2)}$  has full rank, for example for  $B = I$ , where

$$A_{ij,ab}^{(2)} = \begin{cases} 1 & \text{if } (i, j) = (a, b), \\ 0 & \text{otherwise.} \end{cases} \quad (3.5)$$

This generalizes to arbitrary  $(S, T) \in \mathcal{M}^{\leq 3}(\mathcal{G})$  by a continuity argument. Note that  $A^{(2)}$  only depends on  $B$  and not on  $\boldsymbol{\epsilon}$ . Moreover,

$$\{B \in \mathcal{B} : A^{(2)} \text{ has lower rank}\} \quad (3.6)$$

is the intersection of the kernels of the full minors of  $A^{(2)}$ . Each of these minors is a polynomial in the entries of  $B$  and is not the zero polynomial on  $\mathcal{B}$  as shown by the example (3.5) above. Furthermore, the kernels of non-zero polynomials are lower dimensional. This result follows with induction from the fact that a non-zero polynomial in one variable has finitely many zeros [21].

Consequently, the set (3.6) is lower dimensional. This yields the existence of a sequence  $B^{(l)} \rightarrow B$  with all  $B^{(l)}$  having full rank. Now let

$$\mathbf{X}^{(l)} = (B^{(l)})^{-T} \boldsymbol{\epsilon},$$

and  $M^{(l)}$  be the respective matrix of moments. For  $M^{(l)}$ , we know from the first case that it has rank at most  $p$  and consequently all its full minors are zero. Furthermore, the moments of  $\mathbf{X}^{(l)}$  depend continuously on the moments of  $\boldsymbol{\epsilon}$  and the entries of  $B^{(l)}$ . Hence,  $M^{(l)} \rightarrow M$ , which proves that also  $M$  has at most rank  $p$ .  $\square$

Next, we turn to a sufficient condition for belonging to the third-order-moment model.

**Theorem 3.5.** *If  $M$  has rank at most  $p$ , and  $T$  can be expressed as*

$$T = C \circ C \circ C \tag{3.7}$$

for an invertible matrix  $C \in \mathbb{R}^{p \times p}$ , then  $(S, T)$  belongs to the third-order moment model  $\mathcal{M}^{\leq 3}(\mathcal{G})$ .

*Proof.* First note that if  $T$  can be written as in (3.7), it follows that  $T$  can also be written as

$$T = \Omega^{(3)} \bullet B^{-1} \bullet B^{-1} \bullet B^{-1} \tag{3.8}$$

for an  $\Omega^{(3)} \in \text{Sym}_3(V)$  diagonal and a  $B \in \mathbb{R}^{p \times p}$  with ones on the diagonal. To derive such a tensor  $\Omega^{(3)}$  and a matrix  $B$ , we make the ansatz

$$\omega_{iii}^{(3)} \tilde{\beta}_{ai} \tilde{\beta}_{bi} \tilde{\beta}_{ci} = \gamma_{ia} \gamma_{ib} \gamma_{ic} \text{ for all } i = 1, \dots, p. \tag{3.9}$$

A solution is given by defining  $B$  by  $B^T = D \cdot C^{-1}$  where  $D$  is the diagonal matrix with the same entries as  $C^{-1}$  on the diagonal. If further  $\omega_{iii}^{(3)}$  is chosen as  $\omega_{iii}^{(3)} = d_{ii}^3$ , (3.9) holds.

Hence, it remains to show that  $S$  can be written as  $S = B^{-T} \Omega^{(2)} B^{-1}$ . For inferring this relation, we want to write  $s$  a linear combination of the lower  $p$  rows of  $M$ . Since  $M$  has rank  $p$ , the first row is certainly a linear combination of the remaining ones as long as there are no linear dependencies amongst the remaining rows. Denote by  $M_2$  the matrix consisting of the lower  $p$  rows of  $M$ . Equation (3.8) yields that

$$t_{abc} = \sum_{i=1}^p \omega_{iii}^{(3)} \tilde{\beta}_{ci} (\tilde{\beta}_{ai} \tilde{\beta}_{bi}).$$

Thus,  $M_2$  can be computed as

$$M_2 = B^{-T} \Omega^{(3)} \begin{pmatrix} \tilde{\beta}_{11} \tilde{\beta}_{11} & \tilde{\beta}_{11} \tilde{\beta}_{21} & \cdots & \tilde{\beta}_{p1} \tilde{\beta}_{p1} \\ \tilde{\beta}_{12} \tilde{\beta}_{12} & \tilde{\beta}_{12} \tilde{\beta}_{22} & \cdots & \tilde{\beta}_{p2} \tilde{\beta}_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\beta}_{1p} \tilde{\beta}_{1p} & \tilde{\beta}_{1p} \tilde{\beta}_{2p} & \cdots & \tilde{\beta}_{pp} \tilde{\beta}_{pp} \end{pmatrix}.$$

The first two factors are invertible by their definition. To show that the last factor has full rank as well assume that there are  $\alpha_i$  such that

$$\alpha_1 \tilde{\beta}_{a1} \tilde{\beta}_{b1} + \cdots + \alpha_p \tilde{\beta}_{ap} \tilde{\beta}_{bp} = 0 \text{ for all } a \leq b.$$

Setting  $a = 1$ , we derive

$$\left( \alpha_1 \tilde{\beta}_{11} \right) \tilde{\beta}_{b1} + \cdots + \left( \alpha_p \tilde{\beta}_{1p} \right) \tilde{\beta}_{bp} = 0 \text{ for all } b \in [p].$$

The linear independence of the columns of  $B^{-1}$  yields that for each  $i \in [p]$ , either  $\alpha_i$  or  $\tilde{\beta}_{1i}$  is zero. Similarly, one can conclude that for each  $a \in [p]$  and for each  $i \in [p]$ ,  $\alpha_i$  is zero or  $\tilde{\beta}_{ai}$  is zero. Combined, we obtain for that each  $i \in [p]$

$$\alpha_i = 0 \quad \text{or} \quad \tilde{\beta}_{ai} = 0 \text{ for all } a \in [p].$$

The second option can not occur, as it again would contradict the invertibility of  $B^{-1}$ . Hence, all  $\alpha_i$  are zero, which shows linear independence.

Consequently,  $M_2$  has full rank. Thus, the assumption that  $M$  has rank at most  $p$ , yields the existence of a  $\zeta \in \mathbb{R}^p$  fulfilling

$$s_{ab} = \sum_{\nu=1}^p \zeta_{\nu} t_{ab\nu} = \sum_{\nu=1}^p \zeta_{\nu} \sum_{i=1}^p \omega_{iii} \tilde{\beta}_{ai} \tilde{\beta}_{bi} \tilde{\beta}_{\nu i}.$$

Therefore, with  $\omega_{ii}^{(2)} = \sum_{\nu} \omega_{iii} \tilde{\beta}_{\nu i} \zeta_{\nu}$ ,

$$S = B^{-T} \cdot \Omega^{(2)} \cdot B^{-1}$$

and we are done.  $\square$

Overall, we have derived a necessary as well as a sufficient condition related to the model under inspection. While similar, the conditions slightly differ. Specifically, in the first theorem, the matrix  $C$  was arbitrary, whereas it was required to be invertible in the second one. This prevents us from directly obtaining a characterization from the above theorems. However, if one restricts to noise terms that have non-zero third moments, an equivalence holds.

**Corollary 3.6** (Characterization of the third-order-moment model).  $(S, T)$  are realizable under a linear SEM with a noise term with non-zero third moments if and only if  $M$  has at most rank  $p$  and  $T$  can be calculated as

$$T = C \circ C \circ C \tag{3.10}$$

for an invertible matrix  $C \in \mathbb{R}^{p \times p}$ .

*Proof.* For the first direction, the only thing that is left to show is the invertibility of  $C$  under the stronger assumption of non-vanishing third moments. Recall that  $C$  was chosen as  $\gamma_{ia} = \sqrt[3]{\omega_{iii}^{(3)}} \tilde{\beta}_{ai}$  in the proof of Theorem 3.1. Since we assumed the third moments of  $\epsilon$  to be non-zero,  $\Omega^{(3)}$  is invertible. Thus,  $C$  is invertible as a product of the invertible matrices  $\Omega^{(3)}$  and  $B^{-T}$ .

For the other direction, note that in the proof of the previous theorem  $\Omega^{(3)}$  was defined in a way such that all the diagonal elements are positive. This shows the statement.  $\square$

## 3.2 Symmetric Tensor Rank

If  $T$  fulfills the condition

$$T = C \circ C \circ C \tag{3.11}$$

for a  $C \in \mathbb{R}^{p \times p}$ , which appears in Theorem 3.1, we speak of  $T$  being a tensor with real symmetric rank at most  $p$ . There is a lot of research on the inference of the symmetric rank of a tensor [15], [3]. In this subsection we want to derive polynomial equations such tensors fulfill in order to develop a test of this condition later on. For this aim it makes sense to replace (3.11) by a weaker condition. As we are interested in deriving polynomial equations, it is advantageous to work over an algebraically closed field. Therefore, we allow  $C$  to be a complex matrix. Additionally, we consider the closure of tensors fulfilling (3.11) instead of the set itself. Taken together we obtain the relaxed condition that  $T$  is in the closure of all  $T'$  that can be expressed as

$$T' = C \circ C \circ C \tag{3.12}$$

for a  $C \in \mathbb{C}^{p \times p}$ . This property is also called  $T$  has *complex symmetric border rank* at most  $p$ . Note that a polynomial that vanishes for  $T$  fulfilling the condition (3.11) also vanishes for all  $T$  fulfilling the weaker condition (3.12) since  $\mathbb{R} \subseteq \mathbb{C}$  is infinite and kernels of polynomials are always closed.

Before turning to the actual results on the symmetric border rank, we give a short introduction to basic concepts related to tensor spaces. For that we mostly follow [15]. We start with some notation. Throughout the whole section, we work over the field  $\mathbb{C}$ . In particular, vector spaces are always assumed to be complex vector spaces. Let  $V$  be a finite-dimensional vector space. Then denote by

- $V^*$  the dual of  $V$ ,
- $v^\vee \in V^*$  the associated dual element of a  $v \in V$ ,
- $\mathfrak{S}_p = \{\pi \text{ permutation on the set } [p]\}$  the symmetric group of order  $p$ ,
- $\text{sgn}(\pi)$  the sign of an element  $\pi \in \mathfrak{S}_p$ ,
- $\text{End}(V)$  the space of endomorphisms of  $V$ .

Up until now, we have viewed tensors simply as multidimensional arrays. To study the symmetric border rank in more detail it is helpful to look at tensor spaces also from a more abstract perspective. Specifically, we use an alternative definition of tensors as multilinear forms.

**Definition 3.7** (Multilinear form). Let  $V_1, \dots, V_k$  be vector spaces. A map  $Q : V_1 \times \dots \times V_k \rightarrow \mathbb{C}$  is called multilinear if it is linear in each of its arguments.

**Definition 3.8** (Tensor product). The tensor product of finite-dimensional of vector spaces  $V_1, \dots, V_k$  is defined as

$$V_1 \otimes \dots \otimes V_k = \{Q : V_1 \times \dots \times V_k \rightarrow \mathbb{C} : Q \text{ is multilinear}\}.$$

The tensor product of a vector space  $V$  taken  $k$ -times with itself is abbreviated by  $V^{\otimes k}$ . Moreover, for vectors  $v_1 \in V_1, v_2 \in V_2, \dots, v_k \in V_k$ , define their tensor product  $v_1 \otimes v_2 \otimes \dots \otimes v_k$  by

$$(v_1 \otimes v_2 \otimes \dots \otimes v_k)(w_1, \dots, w_k) = \prod_{i=1}^k v_i^\vee(w_i).$$

An element  $v_1 \otimes v_2 \otimes \dots \otimes v_k \in V_1 \otimes \dots \otimes V_k$  is called elementary tensor.

Note that not all tensors contained in  $V_1 \otimes \dots \otimes V_k$  are elementary tensors. However, if  $\mathcal{B}_i$  for  $i = 1, \dots, k$  are bases for  $V_i$ , then the set

$$\{b_1 \otimes \dots \otimes b_k : b_i \in \mathcal{B}_i\}$$

gives a basis for  $V_1 \otimes \dots \otimes V_k$  viewed as vector space over  $\mathbb{C}$ . In particular, all elementary elements span the tensor space.

After introducing this alternative definition, it makes sense to see how it aligns with the previous notion of tensors. The space of multilinear maps  $\mathbb{C}^{d_1} \times \dots \times \mathbb{C}^{d_k} \rightarrow \mathbb{C}$  is isomorphic to the space of multidimensional arrays  $\mathbb{C}^{d_1 \times \dots \times d_k}$  by associating to  $Q : \mathbb{C}^{d_1} \times \dots \times \mathbb{C}^{d_k} \rightarrow \mathbb{C}$  the array  $T \in \mathbb{C}^{d_1 \times \dots \times d_k}$  with entries

$$t_{i_1, \dots, i_k} = Q(e_{i_1}, \dots, e_{i_k}),$$

where  $e_i$  denotes the  $i$ th standard basis vector of  $\mathbb{C}^d$ . Conversely, for  $T \in \mathbb{C}^{d_1 \times \dots \times d_k}$ , the respective multilinear map is given by

$$Q(e_{i_1}, \dots, e_{i_k}) = t_{i_1, \dots, i_k}.$$

Next, we are interested in the subspaces of symmetric and skew-symmetric tensors.

**Definition 3.9** (Symmetric tensors). For  $v_1, v_2, \dots, v_k \in \mathbb{C}^p$ , denote

$$v_1 v_2 \dots v_k = \frac{1}{k!} \sum_{\pi \in \mathfrak{S}_k} v_{\pi(1)} \otimes \dots \otimes v_{\pi(k)}$$

and accordingly denote  $v^k = vv \cdots v$ , where  $v$  is multiplied  $k$  times with itself, for  $v \in V$ . Furthermore, define the map  $\pi_S : (\mathbb{C}^p)^{\otimes k} \rightarrow (\mathbb{C}^p)^{\otimes k}$  by

$$\pi_S(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = v_1 v_2 \cdots v_k$$

and extend the definition to  $(\mathbb{C}^p)^{\otimes k}$  by linearity. The subspace of symmetric tensors is defined as the image

$$\text{Sym}_k(\mathbb{C}^p) = \pi_S((\mathbb{C}^p)^{\otimes k}).$$

**Definition 3.10** (Skew-symmetric tensors). Similarly, for  $v_1, v_2, \dots, v_k \in \mathbb{C}^p$ , let

$$v_1 \wedge v_2 \wedge \cdots \wedge v_k = \frac{1}{k!} \sum_{\pi \in \mathfrak{S}_k} \text{sgn}(\pi) v_{\pi(1)} \otimes \cdots \otimes v_{\pi(k)}$$

and define the map  $\pi_\Lambda : (\mathbb{C}^p)^{\otimes k} \rightarrow (\mathbb{C}^p)^{\otimes k}$  by

$$\pi_\Lambda(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = v_1 \wedge v_2 \wedge \cdots \wedge v_k.$$

The subspace of skew-symmetric tensors is given by

$$\Lambda^k(\mathbb{C}^p) = \pi_\Lambda((\mathbb{C}^p)^{\otimes k}).$$

The space of symmetric tensors indeed consists of symmetric tensors in the sense that

$$\text{Sym}_k(\mathbb{C}^p) = \{T : T \circ \pi = T \text{ for all } \pi \in \mathfrak{S}_p\}.$$

This also shows that the new definition of the space of symmetric tensors aligns with the previous one. Likewise, the skew-symmetric tensors are precisely the tensors that are invariant under permutations except for a change in the sign:

$$\Lambda^k(\mathbb{C}^p) = \{T : T \circ \pi = \text{sgn}(\pi)T \text{ for all } \pi \in \mathfrak{S}_p\}.$$

An important result needed for the next theorem is that each non-zero element  $\Omega \in \Lambda^p(\mathbb{C}^p)$  gives rise to an isomorphism  $\Lambda^k(\mathbb{C}^p) \simeq \Lambda^{p-k}(\mathbb{C}^p)$  given by

$$\begin{aligned} \Lambda^k(\mathbb{C}^p) &\rightarrow \Lambda^{p-k}(\mathbb{C}^p) \\ v_1 \wedge \cdots \wedge v_k &\mapsto \Omega(v_1, \dots, v_k, \cdot, \dots, \cdot). \end{aligned} \tag{3.13}$$

The tensor  $\Omega$  is called volume form. Two special cases of this fact are relevant to us. If one chooses  $\Omega = e_1 \wedge e_2 \wedge e_3$ , where  $\{e_1, e_2, e_3\}$  denotes the standard basis of  $\mathbb{C}^3$ , one obtains the isomorphism

$$\begin{aligned} \Lambda^3(\mathbb{C}^3) &\rightarrow \mathbb{C} \\ v_1 \wedge v_2 \wedge v_3 &\mapsto \Omega(v_1, v_2, v_3) = (e_1 \wedge e_2 \wedge e_3)(v_1, v_2, v_3), \end{aligned}$$

where the left expression turns out to be the determinant of the matrix with columns  $v_1$ ,  $v_2$  and  $v_3$ . The same volume form furthermore gives the isomorphism

$$\begin{aligned} \Lambda^2(\mathbb{C}^3) &\rightarrow \mathbb{C}^3 \\ v_1 \wedge v_2 &\mapsto \Omega(v_1, v_2, \cdot) = (e_1 \wedge e_2 \wedge e_3)(v_1, v_2, \cdot) = \det((v_1 \ v_2 \ \cdot)). \end{aligned}$$

Now, we can give a general definition of the symmetric border rank of a tensor.

**Definition 3.11** (Symmetric rank). A symmetric tensor  $T \in \text{Sym}_k(\mathbb{C}^p)$  is said to have rank one if it can be written as  $T = v^d$  for some  $v \in \mathbb{C}^p$ . Further,  $T$  has rank  $r$  if  $r$  is minimal with the property that  $T$  can be written as the sum of  $r$  rank one tensors.

If we would replace  $\mathbb{C}$  by  $\mathbb{R}$  in this definition, the property (3.11) in Theorem 3.1 would be indeed equivalent to  $T$  having symmetric rank at most  $p$ . If  $T$  is a tensor of the form  $T = \sum_{i=1}^p v_i^d$  for  $d$  vectors  $v_i \in \mathbb{C}^p$ , then  $T$  viewed as an array takes the form

$$t_{abc} = \sum_{i=1}^p (v_i)_a (v_i)_b (v_i)_c,$$

where  $(v_i)_j$  denotes the  $j$ th coordinate of  $v_i$ . This proves the claim. The symmetric rank over  $\mathbb{C}$  is in general lower or equal than the symmetric rank over  $\mathbb{R}$ .

Next, the symmetric border rank is introduced.

**Definition 3.12** (Symmetric border rank). The symmetric border rank of a tensor  $T$  is the smallest  $r$  such that there exists a sequence  $(T_l)$  of tensors of symmetric rank  $r$  converging to  $T$ . We denote the set of all tensors with symmetric border rank at most  $r$  by  $\hat{\sigma}_{r, \text{Sym}_k(\mathbb{C}^p)}$ .

One observes that the set of symmetric tensors with symmetric border rank bounded by  $r$  is the closure of all symmetric tensors with rank at most  $r$ . This shows on the one hand that the symmetric border rank is always lower or equal than the symmetric rank. On the other hand, the closedness of the set is also the reason why it is easier to study the symmetric border rank.

Now, we examine under which circumstances a tensor  $T \in \text{Sym}_3(\mathbb{C}^p)$  has symmetric border rank at most  $p$ . We start with  $p = 2$ .

**Proposition 3.13.** Every symmetric tensor  $T \in \text{Sym}_3(\mathbb{C}^2)$  has symmetric border rank at most two.

*Proof.*  $T$  viewed as an array is uniquely determined by the entries  $t_{111}$ ,  $t_{112}$ ,  $t_{122}$ , and  $t_{222}$  due to symmetry. Hence,  $T$  can be written as the sum of two tensors of the form  $v^2$  and  $w^d$  if and only if

$$\begin{aligned} t_{111} &= (v_1)^3 + (w_1)^3, \\ t_{112} &= (v_1)^2 v_2 + (w_1)^2 w_2, \\ t_{122} &= v_1 (v_2)^2 + w_1 (w_2)^2, \\ t_{222} &= (v_2)^3 + (w_2)^3. \end{aligned}$$

This equation system has a solution for all  $\text{Sym}_3(\mathbb{C}^2)$  except for a lower dimensional subset. Consequently, the subset of tensor with symmetric rank at most two is dense in  $\text{Sym}_3(\mathbb{C}^2)$ , which shows that all elements of  $\text{Sym}_3(\mathbb{C}^2)$  have symmetric border rank at most two.  $\square$

For  $p = 3$ , a polynomial equation characterizing the set of tensors with symmetric border rank bounded by three is known.

**Theorem 3.14** (Aronhold invariant, [2]). *A symmetric tensor  $T \in \text{Sym}_3(\mathbb{C}^p)$  has symmetric border rank at most three if and only if the so-called Aronhold invariant*

$$\begin{aligned} & t_{111}t_{222}t_{333}t_{123} - (t_{222}t_{333}t_{112}t_{113} + t_{333}t_{111}t_{122}t_{223} + t_{111}t_{222}t_{133}t_{233}) \\ & - t_{123}(t_{111}t_{223}t_{233} + t_{222}t_{133}t_{113} + t_{333}t_{112}t_{122}) + (t_{111}t_{122}t_{233}^2 + t_{111}t_{133}t_{223}^2 \\ & + t_{222}t_{112}t_{133}^2 + t_{222}t_{233}t_{113}^2 + t_{333}t_{223}t_{112}^2 + t_{333}t_{113}t_{122}^2) - t_{123}^4 \\ & + 2t_{123}^2(t_{122}t_{133} + t_{233}t_{112} + t_{113}t_{223}) - 3t_{123}(t_{112}t_{223}t_{133} + t_{113}t_{122}t_{233}) \\ & - (t_{122}^2t_{133}^2 + t_{233}^2t_{112}^2 + t_{113}^2t_{223}^2) + (t_{233}t_{112}t_{113}t_{223} + t_{113}t_{223}t_{122}t_{133} \\ & + t_{122}t_{133}t_{233}t_{112}) \end{aligned}$$

vanishes for  $T$ .

There are several proofs of this theorem. We present a version that expresses the Aronhold invariant as a so-called *Pfaffian* of a matrix since a similar procedure is used in the next result in this subsection.

More precisely, the proof is accomplished by associating a matrix  $A_T \in \mathbb{C}^{9 \times 9}$  to  $T \in \text{Sym}_3(\mathbb{C}^p)$ . This matrix has rank bounded by six if  $T$  has rank bounded by three. Hence, all  $8 \times 8$ -minors of  $A_T$  vanish for  $T \in \hat{\sigma}_{3, \text{Sym}_3(\mathbb{C}^3)}$ . What is more, skew-symmetric matrices with even numbers of columns and rows have the special property that their determinant is always the square of a polynomial. This polynomial is called the Pfaffian of the matrix. Accordingly, the principal Pfaffians of order  $2m$  of a skew-symmetric matrix are the square roots of the principal minors of order  $2m$ . For  $A_T$ , all those principal 8-Pfaffians coincide up to multiplicity with a scalar and further match the Aronhold invariant up to scale.

*Proof of Theorem 3.14.* (Following [22]) Throughout the whole proof we implicitly use the isomorphisms  $\Lambda^3(\mathbb{C}^3) \simeq \mathbb{C}$  and  $\Lambda^2(\mathbb{C}^3) \simeq \mathbb{C}^3$  as in (3.13) with the fixed volume form  $\Omega = e_1 \wedge e_2 \wedge e_3$ .

First, it is shown that the Aronhold invariant indeed vanishes for tensors of symmetric border rank not larger than three. As already mentioned, we want to define the matrix  $A_T$ . To do that, we proceed in several steps. The assignment  $T \mapsto A_T$  should be linear, so it makes sense to give the definition first for elementary elements  $T = w_1 \otimes w_2 \otimes w_3 \in \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$ .

As a first step, view  $w_1 \otimes w_2 \otimes w_3$  as a linear map  $\mathbb{C}^3 \rightarrow \mathbb{C}^3 \otimes \mathbb{C}^3$ :

$$y \mapsto y^\vee(w_1)w_2 \otimes w_3.$$

Then, tensor this map with the identity  $\text{Id}_{\mathbb{C}^3} : \mathbb{C}^3 \rightarrow \mathbb{C}^3$  and obtain a map

$$\begin{aligned} \mathbb{C}^3 \otimes \mathbb{C}^3 &\rightarrow \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3 \\ x \otimes y &\mapsto y \otimes ((y^\vee(w_1))w_2 \otimes w_3) = y^\vee(w_1)(x \otimes w_2 \otimes w_3). \end{aligned}$$



Finally, the two penultimate factors are skew-symmetrized to combine them into a factor in  $\Lambda^2 \mathbb{C}^p \simeq \mathbb{C}^p$

$$x \otimes y \mapsto y^\vee(w_1) ((x \wedge w_2) \otimes w_3).$$

Thus, we end up with a map  $\mathbb{C}^3 \otimes \mathbb{C}^3 \rightarrow \Lambda^2(\mathbb{C}^3) \otimes \mathbb{C}^3$  which we define as the desired  $A_T$ . For the next steps of the proof, it is convenient to view  $A_T$  as a linear map from  $\text{End}(\mathbb{C}^3)$  to  $\text{End}(\mathbb{C}^3)$ . So, we identify both, the definition and image space, with the space  $\text{End}(\mathbb{C}^3)$ , by viewing  $x \wedge w_2 \in \Lambda^2(\mathbb{C}^3)$  as an element of  $\mathbb{C}^3$  and then using the isomorphism

$$\begin{aligned} \mathbb{C}^3 \otimes \mathbb{C}^3 &\rightarrow \text{End}(\mathbb{C}^3) \\ u_1 \otimes u_2 &\mapsto \phi \text{ where } \phi(z) = u_1^\vee(z)u_2. \end{aligned}$$

Applying the isomorphism first to the definition space yields the formula

$$A_T(C)(v) = (C^T(w_1) \wedge w_2) \otimes w_3,$$

for  $C$  a matrix representing an element in  $\text{End}(\mathbb{C}^3)$ . Subsequently, applying it to the image space gives

$$A_T(C)(v) = (C^T(w_1) \wedge w_2 \wedge v)w_3$$

for  $C$  a matrix representing an element in  $\text{End}(\mathbb{C}^3)$  and  $v \in \mathbb{C}^3$ . In the last step we used that for the element  $C^T(w_1) \wedge w_2 \in \Lambda^2(\mathbb{C}^3)$  the dual element is given by

$$(C^T(w_1) \wedge w_2)^\vee = \Omega(C^T(w_1), w_2, \cdot) = C^T(w_1) \wedge w_2 \wedge \cdot.$$

To verify that  $A_T$  is skew-symmetric we equip the vector space  $\text{End}(\mathbb{C}^3)$  with a scalar product  $\langle \cdot, \cdot \rangle$  so that the skew-symmetry can be verified by checking that

$$\langle A_T(C), D \rangle = -\langle A_T(C), D \rangle$$

for all  $C, D \in \text{End}(\mathbb{C}^3)$ . A scalar product is given by the *Killing scalar product*  $\langle C, D \rangle = \text{tr}(AB^T)$ . With this,

$$\begin{aligned} \text{tr}(A_T(C)D^T) &= \sum_{j=1}^3 e_j^\wedge (A_T(C)(D^T(e_j))) \\ &= \sum_{j=1}^3 e_j^\wedge (A_T(C)(D^T(e_j))) \\ &= \sum_{j=1}^3 e_j^\wedge \left( \sum_{\pi \in \mathfrak{S}_3} (C^T(e_{i_{\pi(1)}}) \wedge e_{i_{\pi(2)}} \wedge D^T(e_j))e_{i_{\pi(3)}} \right) \\ &= \sum_{\pi \in \mathfrak{S}_3} \sum_{j=1}^3 e_j^\wedge \left( (C^T(e_{i_{\pi(1)}}) \wedge e_{i_{\pi(2)}} \wedge D^T(e_j))e_{i_{\pi(3)}} \right) \\ &= \sum_{\pi \in \mathfrak{S}_3} (C^T(e_{i_{\pi(1)}}) \wedge e_{i_{\pi(2)}} \wedge D^T(e_{i_{\pi(3)}})) \\ &= - \sum_{\pi \in \mathfrak{S}_3} (D^T(e_{i_{\pi(1)}}) \wedge e_{i_{\pi(2)}} \wedge D^T(e_{i_{\pi(3)}})) \\ &= -\text{tr}(A_T(D)C^T). \end{aligned}$$

The next step is to show that for a symmetric tensor with rank one,  $\text{rank}(A_T) \leq 2$ . On one hand,

$$\begin{aligned} \text{tr}(A_{e_{i_1}e_{i_2}e_{i_3}}(C)) &= \sum_{j=1}^3 e_j^\wedge \sum_{\pi \in \mathfrak{S}_3} (C^T(e_{i_{\pi(1)}}) \wedge e_{i_{\pi(2)}} \wedge e_j) e_{i_{\pi(3)}} \\ &= \sum_{\pi \in \mathfrak{S}_3} (C^T(e_{i_{\pi(1)}}) \wedge e_{i_{\pi(2)}} \wedge e_{i_{\pi(3)}}) = 0. \end{aligned}$$

Hence, from linearity of  $T \mapsto A_T$ ,  $\text{tr}(A_T(C))$  is zero for all  $T \in \text{Sym}_3(\mathbb{C}^3)$  and all  $C$ . On the other hand, for  $w^3$  a tensor with symmetric rank 1, the matrix  $A_{w^3}$  fulfills

$$A_{w^3}(C)(v) = 6 (C^T(w) \wedge w \wedge v) w \subseteq \text{span}(w)$$

for every  $v$ . Overall,

$$\text{Im}(A_{w^3}) \subseteq \{D : \text{Im}(D) \subseteq \text{span}(w) \text{ and } \text{tr}(D) = 0\},$$

which has dimension two.

Therefore, for  $T$  a sum of three rank one tensors  $w_i^3$ , one obtains

$$\text{rank}(A_T) = \text{rank}\left(\sum_{i=1}^r A_{w_i^3}\right) \leq \sum_{i=1}^r \text{rank}(A_{w_i^3}) \leq 3 \cdot 2 = 6.$$

Due to continuity, the same bound holds for all tensors with symmetric border rank at most 3. As a consequence, all principal Pfaffians of size eight of  $A_T$  vanish.

The final step for showing the necessity of the Aronhold equation is verifying that the Pfaffians of  $A_T$  have indeed the form given in the theorem. For that, one needs to express  $A_T$  as a matrix for some choice of a basis for  $\text{End}(\mathbb{C}^p)$ . For example, for the basis  $M_1 = e_1 \otimes e_1, M_2 = e_2 \otimes e_1, M_3 = e_3 \otimes e_1, M_4 = e_1 \otimes e_2, \dots, M_9 = e_3 \otimes e_3$ , one obtains

$$A_T = \begin{pmatrix} 0 & t_{113} & -t_{112} & 0 & -t_{123} & t_{122} & 0 & t_{133} & -t_{123} \\ -t_{113} & 0 & t_{111} & t_{123} & 0 & -t_{112} & -t_{133} & 0 & t_{113} \\ t_{112} & -t_{111} & 0 & -t_{122} & t_{112} & 0 & t_{123} & -t_{113} & 0 \\ 0 & -t_{123} & t_{122} & 0 & t_{223} & -t_{222} & 0 & -t_{233} & t_{223} \\ t_{123} & 0 & -t_{112} & -t_{223} & 0 & t_{122} & t_{233} & 0 & -t_{123} \\ -t_{122} & t_{112} & 0 & t_{222} & -t_{122} & 0 & -t_{223} & t_{123} & 0 \\ 0 & t_{113} & -t_{123} & 0 & -t_{233} & t_{223} & 0 & t_{333} & -t_{233} \\ -t_{133} & 0 & t_{113} & t_{233} & 0 & -t_{123} & -t_{333} & 0 & t_{133} \\ t_{123} & -t_{113} & 0 & -t_{223} & t_{123} & 0 & t_{233} & -t_{133} & 0 \end{pmatrix}$$

for  $T = (t_{ijk})$ . It turns out that all the principal 8-Pfaffians coincide up to multiplication with a scalar and also match the Aronhold invariant up to scale.

To derive sufficiency, one shows that the codimension of  $\hat{\sigma}_{3, \text{Sym}_3(\mathbb{C}^3)}$  is one, which means that  $\hat{\sigma}_{3, \text{Sym}_3(\mathbb{C}^3)}$  is the kernel of one polynomial. In addition, there is no polynomial with a lower degree as the Aronhold invariant that vanishes on the set in question. This step is for example carried out in Chapter 15 in [15].  $\square$

For higher  $p$ , no equation system characterizing  $\hat{\sigma}_{p, \text{Sym}_3(\mathbb{C}^p)}$  is known. However, there are results about necessary equations tensors contained in this space are fulfilling. As shown in [15, (3.10.1)], those equations can be obtained similarly to the Aronhold invariant. Let  $a = \lfloor \frac{p}{2} \rfloor$ . We again define a matrix  $A_T$  first for an elementary element  $T = w_1 \otimes w_2 \otimes w_3 \in \mathbb{C}^p \otimes \mathbb{C}^p \otimes \mathbb{C}^p$  as follows: View  $w_1 \otimes w_2$  as a map  $\mathbb{C}^p \rightarrow \mathbb{C}^p$ , and use the last factor  $w_3$  to obtain a function  $\Lambda^a(\mathbb{C}^p) \rightarrow \Lambda^{a+1}(\mathbb{C}^p)$  by  $v \mapsto v \wedge w_3$ . Combining the two maps yields the desired element  $A_T : \mathbb{C}^p \otimes \Lambda^a(\mathbb{C}^p) \rightarrow \mathbb{C}^p \otimes \Lambda^{a+1}(\mathbb{C}^p)$ . Similarly as before,  $A_T$  viewed as a matrix has lower rank for  $T \in \hat{\sigma}_{3, \text{Sym}_3(\mathbb{C}^p)}$ .

**Theorem 3.15** ([15, (3.10.1)]). *If  $T$  has symmetric border rank at most  $p$ , then  $A_T \in \mathbb{R}^{p \binom{p}{a} \times p \binom{p}{a}}$  has at most rank  $\binom{p-1}{a} p$ .*

Thus, for every  $p$ , the minors of  $A_T$  vanish on  $\hat{\sigma}_{p, \text{Sym}_3(\mathbb{C}^p)}$ . Moreover, for odd  $p$ ,  $A_T$  is skew-symmetric, so again the Pfaffians of the corresponding size vanish.

In addition to those algebraic results, there exist various algorithms to infer the symmetric rank of a tensor. Examples include the methods in [19], [3] and [4].

### 3.3 Higher Moments

Considering also higher moments, we can find even more linear dependencies. More precisely, we obtain constraints in the *cumulant tensors* of  $\mathbf{X}$ .

**Definition 3.16.** The  $k$ -th order cumulant tensor  $\mathcal{C}^{(k)} \in (\mathbb{R}^p)^{\otimes k}$  is given by

$$\mathcal{C}_{(i_1, \dots, i_k)}^{(k)} = \sum_{(A_1, \dots, A_l)} (-1)^{l-1} (l-1)! \mathbb{E} \left[ \prod_{j \in A_1} X_j \right] \cdots \mathbb{E} \left[ \prod_{j \in A_l} X_j \right],$$

where  $(A_1, \dots, A_l)$  is an arbitrary partition of  $\{i_1, \dots, i_k\}$ .

Like for the second and third moments, one can establish a connection between the cumulants of  $\mathbf{X}$  and  $\boldsymbol{\epsilon}$  if  $\mathbf{X}$  solves a linear structural equation system. For  $\mathcal{C}^{(k)}$  the  $k$ -th order cumulant tensor of  $\mathbf{X}$ , and  $\Omega^{(k)}$  the  $k$ -th order cumulant tensor of  $\boldsymbol{\epsilon}$  the following holds.

**Lemma 3.17** ([30, Lemma 2.4]). *If  $P^{\mathbf{X}}$  belongs to the graphical model of the complete graph  $\mathcal{G}$ , then*

$$\mathcal{C}^{(k)} = \Omega^{(k)} \bullet B^{-1} \bullet \dots \bullet B^{-1},$$

where the factor  $B^{-1}$  appears  $k$  times.

The proof works similarly to the case of the second and third moments.

From Lemma 3.17, analogous to the proof of Theorem 3.1, one can show that for each tuple  $(j_1, j_2, \dots, j_{l-2}) \in [p]^{l-2}$  the vector

$$\begin{pmatrix} c_{j_1 j_2 \dots j_{k-2} 11} \\ c_{j_1 j_2 \dots j_{k-2} 12} \\ \vdots \\ c_{j_1 j_2 \dots j_{k-2} pp} \end{pmatrix} \quad (3.14)$$

where  $c_{j_1 j_2 \dots j_k}$  denote the entries of  $\mathcal{C}^{(k)}$  is contained in

$$\text{span} \left( \left\{ \begin{pmatrix} \tilde{\beta}_{1i} \tilde{\beta}_{1i} \\ \tilde{\beta}_{1i} \tilde{\beta}_{2i} \\ \vdots \\ \tilde{\beta}_{pi} \tilde{\beta}_{pi} \end{pmatrix} : i \in [p] \right\} \right).$$

Therefore, if  $\mathcal{C}^{(k)}$  is rearranged into a matrix in a way that each row contains a vector of the form (3.14), and is appended to  $M$ , the rank remains bounded by  $p$ .

### 3.4 Conditions for one Missing Edge

In this section, it is examined for the case  $p = 3$  whether further equations hold if one edge is removed from the graph.

**Theorem 3.18.** *Let  $p = 3$ . If  $P^X$  belongs to a linear SEM with a noise term with non-zero third moments and furthermore one edge weight  $\beta_{j^* i^*}$  is zero, then all full minors of the submatrix of  $M$  with columns indexed by  $(a, b)$  where  $a, b \neq i^*$  are zero.*

**Example 3.19.** For example, if  $\beta_{21}$  is zero, then all full minors of the matrix

$$N = \begin{pmatrix} s_{11} & s_{13} & s_{33} \\ t_{111} & t_{113} & t_{133} \\ t_{211} & t_{213} & t_{233} \\ t_{311} & t_{313} & t_{333} \end{pmatrix}$$

vanish.

*Proof.* Without loss of generality, let  $i^* = 1$  and  $j^* = 2$ . If  $\beta_{21} = 0$ , the minor of  $B^{-T}$  obtained by eliminating the second row and the first column has determinant zero. Hence, there is a  $\zeta$  such that

$$\begin{pmatrix} \tilde{\beta}_{11} \\ \tilde{\beta}_{31} \end{pmatrix} = \zeta \cdot \begin{pmatrix} \tilde{\beta}_{12} \\ \tilde{\beta}_{32} \end{pmatrix}.$$

From these two equations, for all  $(a, b)$  in  $\{(1, 1), (1, 3), (3, 3)\}$ ,

$$\tilde{\beta}_{a1} \tilde{\beta}_{b1} = \zeta^2 \tilde{\beta}_{a2} \tilde{\beta}_{b2}.$$

To show that the last three rows of  $N$  are linearly dependent define  $\alpha_\nu = \sum_{\mu=1, \mu \neq 2}^3 \delta_\mu \omega_{\mu\mu\mu}^{-1} \beta_{\mu\nu}$ , for  $\delta = (1, 0, -\zeta^2)$ . Then

$$\begin{aligned}
\sum_{\nu=1}^3 \alpha_\nu t_{\nu ab} &= \sum_{\nu=1}^3 \sum_{\mu=1, \mu \neq 2}^3 \sum_{i=1}^3 \delta_\mu \beta_{\mu\nu} \omega_{\mu\mu\mu}^{-1} \omega_{iii} \tilde{\beta}_{ai} \tilde{\beta}_{bi} \tilde{\beta}_{\nu i} \\
&= \sum_{\mu=1, \mu \neq 2}^3 \delta_\mu \tilde{\beta}_{a\mu} \tilde{\beta}_{b\mu} \\
&= 0.
\end{aligned} \tag{3.15}$$

Additionally, from the proof of Theorem 3.6, the first row of  $M$  is a linear combination of the other rows. This together with (3.15) proves the statement.  $\square$

## 4 Statistical Tests

In this section, we exploit the results of the previous section to develop a procedure for testing the assumption that a distribution follows a linear SEM. First, some basic notions in the field of statistics are recalled. When developing a test to assess a certain null hypothesis  $H_0$  about a distribution  $P^X$ , one is interested in controlling the *size* of it. The size is the probability of rejecting the null hypothesis although it is true. More precisely, one requires the size to be lower than a fixed *level* or *nominal size*  $\alpha$ , which is typically chosen to be 0.05. Furthermore, the *power* of the test is the likelihood of rejecting the null hypothesis if it is indeed wrong. One aims for a power as high as possible.

Coming back to our aim of assessing the linearity assumption, the question arises which conditions inferred in the previous section to test exactly. Recall that in the previous section we derived

- a) a necessary condition for the general third-order moment model,
- b) a slightly stronger condition that characterizes the restricted third-order moment model where only noise terms with non-zero third moments are allowed.

Thus, testing for condition *b*) could result in a slightly improved power compared to testing for condition *a*). However, distributions following linear SEMs with noise terms whose third moments vanish would be wrongly rejected. Therefore, we validate condition *a*). Recall that condition *a*) consists of two parts:

- $M$  has rank at most  $p$ , and
- $T$  has symmetric rank at most  $p$ .

We present four options for testing the rank condition on  $M$ . There exists a vast collection of methods for that in the literature. For example, algorithms have been developed that are based on the QR, Cholesky, LU, or Singular Value decomposition of the matrix in question. We refer to [24] for an overview of existing tests and their asymptotics.

Most of the existing algorithms consider the null hypothesis

$$H_0 : \text{rank}(\Pi) = r \text{ vs. } H_1 : \text{rank}(\Pi) > r \quad (4.1)$$

and do not take the possibility of an even lower rank into account. This is the main motivation why we choose to consider the four approaches presented. All of them, apart from the first one, have the advantage that they can be used to address null hypothesis

$$H_0 : \text{rank}(\Pi) \leq r \text{ vs. } H_1 : \text{rank}(\Pi) > r.$$

The first test we introduce is used as a sub-procedure in the second algorithm presented. The test on its own can only handle the null hypothesis (4.1). Both algorithms utilize

the singular value decomposition of  $M$ . The other two methods are procedures to test polynomial constraints. Thus, for our objective, they can be used by validating if all full minors of  $M$  are zero.

For testing the rank condition regarding  $T$ , again the two tests assessing polynomial constraints can be used by applying them with the necessary equations for the symmetric border rank of a tensor obtained in section 3.2.

Throughout the whole section,  $n$  denotes the sample size, and  $\mathcal{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)})$  is an independent and identically distributed sample drawn from  $P^X$ . Furthermore, we write  $\text{vec}(A)$  for the column vectorization of a matrix  $A$ , so for  $A = (a_{ij})$ ,

$$\text{vec}(A) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \\ a_{12} \\ \vdots \\ a_{kl} \end{pmatrix}.$$

## 4.1 Singular Value Based Approaches

Before the exact procedures of the tests are explained, we shortly recall the singular value decomposition. Let  $\Pi \in \mathbb{R}^{k \times m}$ . To simplify notation we assume without loss of generality that  $k \leq m$ . The singular value decomposition of  $\Pi$  is defined as a decomposition of the form

$$\Pi = U\Sigma V^T$$

where  $U \in \mathbb{R}^{k \times k}$  and  $V \in \mathbb{R}^{m \times m}$  are required to be orthogonal, and  $\Sigma \in \mathbb{R}^{k \times m}$  takes the form

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k & 0 & \cdots & 0 \end{pmatrix},$$

where the diagonal entries satisfy  $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq 0$ . The  $\sigma_i$  are called singular values of  $\Pi$ . Note that  $U$  and  $V$  are in general not uniquely determined, whereas  $\Sigma$  is uniquely determined. As, both,  $U$  and  $V$  are invertible,  $\Pi$  has rank  $r$  if and only if precisely the last  $k - r$  singular values are zero.

### 4.1.1 Kleibergen-Paap Test

The first method we introduce is the rank test by Kleibergen and Paap [13]. The setup is as follows: One considers the null hypothesis

$$H_0 : \text{rank}(\Pi) = r \text{ vs. } H_1 : \text{rank}(\Pi) > r,$$

where  $\Pi \in \mathbb{R}^{k \times m}$  is an arbitrary matrix of parameters of a distribution. Furthermore, the existence of a consistent estimator  $\hat{\Pi}$  of  $\Pi$  satisfying

$$\sqrt{n} \left( \text{vec}(\Pi) - \text{vec}(\hat{\Pi}) \right) \xrightarrow{d} \mathcal{N}(0, W) \quad (4.2)$$

for some covariance matrix  $W$  is required. The test is based on the above-mentioned equivalence of  $\Pi$  having rank  $r$  and the smallest  $k - r$  singular values of  $\Pi$  being zero. Let

$$\Pi = U \Sigma V^T$$

be the singular value decomposition of  $\Pi$ . As we are interested in the last  $k - r$  singular values, we consider the block decomposition

$$\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \begin{pmatrix} V_{11}^T & V_{21}^T \\ V_{12}^T & V_{22}^T \end{pmatrix} = \begin{pmatrix} U_{11} \\ U_{21} \end{pmatrix} \Sigma_1 (V_{11}^T \ V_{21}^T) + \begin{pmatrix} U_{12} \\ U_{22} \end{pmatrix} \Sigma_2 (V_{12}^T \ V_{22}^T)$$

with  $U_{11} \in \mathbb{R}^{r \times r}$ ,  $U_{12} \in \mathbb{R}^{r \times k-r}$ ,  $U_{21} \in \mathbb{R}^{k-r \times r}$ ,  $U_{22} \in \mathbb{R}^{k-r \times k-r}$ ,  $V_{11} \in \mathbb{R}^{r \times r}$ ,  $V_{12} \in \mathbb{R}^{r \times m-r}$ ,  $V_{21} \in \mathbb{R}^{m-r \times r}$ , and  $V_{22} \in \mathbb{R}^{m-r \times m-r}$ .

Further, we denote  $U_1 = \begin{pmatrix} U_{11} \\ U_{21} \end{pmatrix}$ ,  $V_1 = \begin{pmatrix} V_{11} \\ V_{21} \end{pmatrix}$ ,  $U_2 = \begin{pmatrix} U_{12} \\ U_{22} \end{pmatrix}$ , and  $V_2 = \begin{pmatrix} V_{12} \\ V_{22} \end{pmatrix}$ .

We want to base a test statistic on the matrix  $\hat{\Sigma}_2$  of the last singular values of  $\hat{\Pi}$ . Hence, we need to find out the limiting distribution of the matrix. As singular values are by definition non-negative, the limiting distribution of  $\hat{\Sigma}_2$  is certainly not normal. Therefore, we do not want to use  $\hat{\Sigma}_2$  directly as a test statistic but, first, transform it to obtain a statistic with a normal limiting distribution.

The transformation is carried out as follows: One aims at finding matrices  $A \in \mathbb{R}^{k \times r}$ ,  $B \in \mathbb{R}^{r \times m}$ ,  $\Xi \in \mathbb{R}^{m-r \times k-r}$ ,  $A_\perp \in \mathbb{R}^{k \times k-r}$ , and  $B_\perp \in \mathbb{R}^{m-r \times m}$  such that

$$U_1 \Sigma_1 V_1^T = AB \quad \text{and} \quad U_2 \Sigma_2 V_2^T = A_\perp \Xi B_\perp.$$

At the same time, the matrices are required to fulfill

- $A^T A_\perp = 0$  and  $B_\perp B^T = 0$ ,
- $A_\perp^T A_\perp = I$  and  $B_\perp B_\perp^T = I$ ,
- $A$  and  $B$  have full rank.

With this ansatz the following three properties are fulfilled:



**Proposition 4.1.** For  $\Xi$  obtained as described above the following holds.

- a)  $\Xi$  can be obtained from  $\Sigma_2$  by pre- and post-multiplying with orthogonal matrices,
- b)  $\Xi$  is zero if and only if  $\Sigma_2$  is zero, and
- c)  $\Xi$  can be derived from  $\Pi$  by calculating  $\Xi = A_{\perp}^T \Pi B_{\perp}^T$ .

*Proof.* To show a) first note that

$$\Xi = A_{\perp}^T U_2 \Sigma_2 V_2^T B_{\perp}^T.$$

Moreover,  $A_{\perp}^T U_2$  is orthogonal because  $(A_{\perp}^T U_2)^T A_{\perp}^T U_2 = U_2 A_{\perp} A_{\perp}^T U_2 = U_2 U_2$ . Similarly,  $(V_2^T B_{\perp}^T)$  is orthogonal, which concludes the proof.

To derive the equivalence b) first assume that  $\Pi$  has rank  $r$ . In this case,

$$A_{\perp}^T \Xi B_{\perp}^T = U_2 \Sigma_2 \begin{pmatrix} V_{12}^T & V_{22}^T \end{pmatrix} = 0.$$

As  $A_{\perp}$  and  $B_{\perp}$  are left and right invertible, respectively, it follows that  $\Xi$  is zero. If the other way around  $\Xi = 0$ , then  $\Pi = AB$ . But  $AB$  has rank  $r$  as  $A$  and  $B$  have full rank.

Property c) follows since

$$\begin{aligned} A_{\perp}^T \Pi B_{\perp}^T &= A_{\perp}^T (AB + A_{\perp} \Xi B_{\perp}) B_{\perp}^T \\ &= A_{\perp}^T A B B_{\perp} + A_{\perp}^T A_{\perp} \Xi B_{\perp} B_{\perp}^T \\ &= \Xi. \end{aligned}$$

from the assumptions made on  $A$ ,  $A_{\perp}$ ,  $B$ ,  $B_{\perp}$  and  $\Xi$ . □

The crucial point is to choose a specific orthogonal transformation such that the resulting limiting distribution of  $\Xi$  is normal. Note that this is not the case for all potential choices of  $A_{\perp}$  and  $B_{\perp}$ . It is for example possible to define  $A_{\perp}$  as  $U_2$  and  $B_{\perp}$  as  $V_2$ . Then  $\Xi$  would be  $\Sigma_2$  which does not have a normal limiting distribution.

To illustrate how one can obtain a choice that leads to a normal limiting distribution first consider the special case of testing for a normal matrix  $\Pi \in \mathbb{R}^{k \times k}$  if  $\text{rank}(\Pi) = k - 1$ . As  $\Pi$  is normal, it has an eigendecomposition

$$\Pi = O D O^T$$

with  $D$  diagonal and  $O$  orthogonal. The singular value decomposition is closely related to the eigendecomposition. Namely, the singular values are the absolute values of the eigenvalues. Like the columns of  $O$ , the columns of  $U$  and  $V$  are the eigenvectors of  $\Pi$ . The only difference is that exactly for the negative eigenvalues,  $u_i$  needs to be  $-v_i$ . This aligns with fact that  $\Sigma$  consists of the absolute values of the entries in  $D$ . Thus, if one defines

$$\Xi = \text{sign}(u_{kk}) \sigma_k \text{sign}(v_{kk})$$

one reverses taking the absolute value so that the formula yields exactly the smallest eigenvalue of  $\Pi$ . Contrary to the absolute value, the function that assigns to a matrix its smallest eigenvalue is locally differentiable [18]. Consequently, the delta theorem and the fact that  $\Pi$  has a normal limiting distribution ensure that  $\Xi$  has a normal limiting distribution.

This idea can be generalized by defining  $A_{\perp}$  and  $B_{\perp}$  as

$$A_{\perp} = U_2 U_{22}^{-1} (U_{22} U_{22}^T)^{\frac{1}{2}}, \quad (4.3)$$

$$B_{\perp} = (V_{22} V_{22}^T)^{\frac{1}{2}} V_{22}^{-T} V_2. \quad (4.4)$$

With this choice of  $A_{\perp}$  and  $B_{\perp}$ ,  $\Xi$  takes the form

$$\Xi = (U_{22} U_{22}^T)^{-\frac{1}{2}} U_{22} \Sigma_2 (V_{22} V_{22}^T)^{-\frac{1}{2}} V_{22}.$$

After laying out the theoretical background, we turn to the definition of the actual test statistic. In the same way as  $\Xi$  is calculated from  $\Pi$  we obtain  $\hat{\Xi}$  from  $\hat{\Pi}$  by first calculating the singular value decomposition of  $\hat{\Pi}$ , calculate  $\hat{A}_{\perp}^T$  and  $\hat{B}_{\perp}^T$  according to (4.3) and then define  $\hat{\Xi}$  as

$$\hat{\Xi} = \hat{A}_{\perp}^T \hat{\Pi} \hat{B}_{\perp}^T.$$

For the exact same reason as for  $\Pi$ ,  $\hat{\Xi}$  is zero if and only if  $\hat{\Pi}$  has rank  $r$ . In order to obtain a test procedure, it remains to find the limiting distribution of  $\Xi$ . For that, denote

$$\hat{\xi} = \text{vec}(\hat{\Xi})$$

and

$$W_r = (B_{\perp} \otimes A_{\perp}^T) W (B_{\perp} \otimes A_{\perp}^T).$$

Then the limiting distribution is given by the subsequent formula.

**Theorem 4.2** ([13, Theorem 1]). *If  $W_r$  is non-singular and  $H_0$  holds true, then*

$$\sqrt{n} \hat{\xi} \xrightarrow{d} \mathcal{N}(0, W_r). \quad (4.5)$$

Hence, the test statistic can be defined as

$$\text{rk} = n \hat{\xi}^T W_r^{-1} \hat{\xi}.$$

As a direct consequence of (4.5), one obtains

$$\text{rk} \xrightarrow{d} \chi^2((k-r)(m-r)).$$

Thus, for a given nominal level  $\alpha$  we reject if

$$\text{rk} \geq c_{\alpha},$$

where  $c_{\alpha}$  is the  $(1 - \alpha)$ th quantile of  $\chi^2((k-r)(m-r))$ .

**Fulfilment of assumptions** For the purpose of testing the rank of the matrix  $M$  the requirement (4.2) can be easily fulfilled. The matrix  $M$  consists of sample moments. Hence, the central limit theorem yields

$$\sqrt{n}(\text{vec}(M) - \text{vec}(M)) \xrightarrow{d} \mathcal{N}(0, W)$$

for the matrix

$$W = \begin{pmatrix} \text{Cov}(X_1^2, X_1^2) & \text{Cov}(X_1^2, X_1^3) & \cdots & \text{Cov}(X_1^2, X_n^3) \\ \text{Cov}(X_1^3, X_1^2) & \text{Cov}(X_1^3, X_1^3) & \cdots & \text{Cov}(X_1^3, X_n^3) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n^3, X_1^2) & \text{Cov}(X_n^3, X_1^3) & \cdots & \text{Cov}(X_n^3, X_n^3) \end{pmatrix},$$

which can be estimated for example using sample covariances.

#### 4.1.2 Bootstrap Statistic

An alternative way to handle the non-standard limiting distribution of the singular values is to use bootstrap as proposed by Chen [6]. Bootstrap is a resampling method developed by Bradley Efron in 1979 [10]. The principal idea behind the methodology is to use the sample itself to obtain an estimate of the limiting distribution of the test statistic under  $H_0$  instead of only leveraging theoretical knowledge about the limiting distribution.

For the aim of testing the rank of a matrix  $\Pi \in \mathbb{R}^{k \times m}$ , one can make use of the bootstrap method as follows. We require the existence of a consistent estimator  $\hat{\Pi}$  of  $\Pi$  such that the distribution of the difference converges to some random matrix  $\mathcal{W}$ :

$$\sqrt{n}(\hat{\Pi}_n - \Pi) \xrightarrow{d} \mathcal{W}. \quad (4.6)$$

In contrast to the previous method,  $\mathcal{W}$  does not necessarily need to have a normal distribution. We assume again  $k \leq m$ . Leveraging the same relation of the rank and the singular values as for the Kleibergen-Paap statistic, one might use

$$\phi_r = n \sum_{j=r+1}^k \sigma_j^2(\hat{\Pi}_n)$$

as a test statistic. To be able to calculate the critical value, one needs to approximate the limiting distribution under the null hypothesis. For that, we first look at the limiting distribution the statistic has in theory.

**Theorem 4.3** ([6, Theorem 3.1]). *Under  $H_0 : \text{rank}(\Pi) \leq r$ ,*

$$\phi_r \xrightarrow{d} \sum_{j=r-r_0+1}^{k-r} \sigma_j^2(U_2^T \mathcal{W} V_2),$$

where  $r_0$  indicates the true rank of  $\Pi$ .

The theorem can be proven by invoking a generalization of the delta method. The test statistic  $\phi_r$  is not differentiable in the classical sense. Consequently, the delta method can not be applied. Instead, we use a generalization of the delta method that is based on a relaxed notion of differentiability, namely the *Hadamard directional derivative*.

**Definition 4.4.** Let  $\phi : \mathbb{R}^{k \times m} \rightarrow \mathbb{R}$ . Then  $\phi$  is Hadamard directionally differentiable in  $A \in \mathbb{R}^{k \times m}$  if there is a function  $\phi'_A : \mathbb{R}^{k \times m} \rightarrow \mathbb{R}$  such that

$$\lim_{l \rightarrow \infty} \frac{\phi(A + h_l C_l) - \phi(A)}{h_l} = \phi'_A(C)$$

for all sequences  $C_l$  converging to some  $C \in \mathbb{R}^{k \times m}$  and all positive sequences  $h_l \rightarrow 0$ . In this case,  $\phi'_A$  is called the Hadamard directional derivative of  $\phi$  in  $A$ . If additionally there exists a function  $\phi''_A : \mathbb{R}^{m \times k} \rightarrow \mathbb{R}$  such that

$$\lim_{l \rightarrow \infty} \frac{\phi(A + h_l C_l) - \phi(A) - h_l \phi'_A(A_l)}{h_l^2} = \phi''_A(C)$$

for all sequences  $A_l \rightarrow A$  and all positive sequences  $h_l \rightarrow 0$ ,  $\phi$  is second-order Hadamard directionally differentiable in  $A$ .

With this weaker notion of differentiability, the test statistic  $\phi_r$  becomes differentiable.

**Lemma 4.5.**  $\phi_r$  is Hadamard directionally differentiable on  $\mathbb{R}^{k \times m}$  with

$$\phi'_{r, \Pi}(C) = \min_{U \in \Psi(\Pi)} 2\text{tr}(UC(U\Pi)^T),$$

where  $\Psi(\Pi) = \arg \min_{U \in \mathbb{S}^{(k-r) \times k}} \|U\Pi\|^2$  and  $\mathbb{S} = \{U : U^T U = I\}$ .

In the case where  $\phi_r(\Pi) = 0$ , the function is additionally second-order Hadamard directionally differentiable, and the derivative is given by

$$\phi''_r(\Pi)(C) = \sum_{j=r-r_0+1}^{k-r_0} \sigma_j^2(U_2^T C V_2).$$

The reason why we are also interested in the second-order derivative is that  $\phi'_{r, \Pi}$  is zero for matrices  $\Pi$  that fulfill the null hypothesis. If  $\text{rank}(\Pi) = 0$ , then

$$U_2^T \Pi = U_2^T (U \Sigma V^T) = U_2^T U_1 \Sigma V^T + U_2^T U_2 \Sigma V^T = 0$$

since  $U$  is orthogonal and the last  $k-r$  singular values are zero. Thus,  $\min_{U \in \mathbb{S}^{(k-r) \times k}} \|U\Pi\|^2 = 0$  and  $\phi'_r(\Pi) \equiv 0$ . Due to this, the delta method would give

$$\sqrt{n} \phi_r(\hat{\Pi}) \xrightarrow{d} 0.$$

Instead, we turn to the second-order delta method. This method can be generalized for Hadamard directionally differentiable functions as shown by [7]:

**Theorem 4.6** (Second-order delta method, [7, Theorem 2.1]). *Let  $(a^{(l)})$  a sequence of real numbers,  $(\mathbf{X}^{(l)})$  a sequence of random vectors, and  $b \in \mathbb{R}$ . If  $a^{(l)} \rightarrow \infty$ ,  $a^{(l)}(\mathbf{X}^{(l)} - b) \xrightarrow{d} \mathbf{Z}$ , and  $g : \mathbb{R}^{k \times m} \rightarrow \mathbb{R}$  is Hadamard directionally differentiable in  $b$ , then*

$$(a^{(l)})^2 (g(\mathbf{X}^{(l)}) - g(b) - g'_b(\mathbf{X}^{(l)} - b)) \xrightarrow{d} g''_b(\mathbf{Z}).$$

With this, the result about the asymptotic distribution of the test statistic can be proven.

*Proof of Theorem 4.3.* Assumption (4.6) ensures that

$$\sqrt{n}(\hat{\Pi}_n - \Pi) \xrightarrow{d} \mathcal{W}$$

for some matrix  $\mathcal{W}$ . Besides that, under  $H_0$ ,  $\phi_r(\Pi)$  is zero. Consequently, from Lemma (4.5),  $\phi_r$  is second-order Hadamard directionally differentiable and the second-order delta method can be applied:

$$n(\phi_r(\hat{\Pi}_n) - \phi_r(\Pi) - \phi'_{r,\Pi}(\hat{\Pi}_n - \Pi)) \xrightarrow{d} \phi''_{r,\Pi}(\mathcal{W}) = \sum_{j=r-r_0+1}^{k-r} \sigma_1^2(U_2^T \mathcal{W} V_2).$$

$\phi_r(\Pi)$ , as well as the function  $\phi'_{r,\Pi}$  are zero. Thus, the statement follows.  $\square$

However,  $\mathcal{W}$  is not known in practice. Here the bootstrap idea comes into place. We randomly sample  $E$  so-called bootstrap samples  $\mathbf{x}_1^*, \dots, \mathbf{x}_E^*$  from  $\mathcal{X}$  with replacement. Each bootstrap sample has the same sample size as  $\mathcal{X}$ . These samples are now used to estimate the distribution of  $\mathcal{W}$ . For that, obtain an estimate of  $\Pi$  from each of the bootstrap samples and denote them by

$$\hat{\Pi}(\mathbf{x}_1^*), \dots, \hat{\Pi}(\mathbf{x}_E^*).$$

With this, obtain  $E$  estimates

$$\mathcal{W}_e^* = \sqrt{n} \left( \hat{\Pi}(\mathbf{x}_e^*) - \hat{\Pi}_n \right), \quad e = 1, \dots, E,$$

of  $\mathcal{W}$ . Hence, the empirical distribution

$$\hat{F}(t) = \frac{1}{E} \sum_{e=1}^E \mathbb{1}_{\mathcal{W}_e^* \leq t}$$

approximates the distribution of  $\mathcal{W}$ . Now, let  $\hat{c}_{n,1-\alpha}$  be the  $(1 - \alpha)$ th quantile of the values

$$\sigma_1^2(U_2^T \mathcal{W}_1^* V_2), \dots, \sigma_1^2(U_2^T \mathcal{W}_E^* V_2)$$

and reject  $H_0$  if

$$n\sigma_{r+1}^2(\hat{\Pi}_n) > \hat{c}_{n,1-\alpha}.$$

**Lower Rank** If we want to test for the null hypothesis

$$H_0 : \text{rank}(\Pi) \leq r,$$

instead of examining if the rank is equal to  $r$ , we have to find a way how to deal with the appearance of  $r_0$  in the formula of the limiting distribution. There are several options for tackling this. The proof of Theorem 4.3 shows that the limiting distribution of the test statistic can alternatively be written as  $\phi''_{r,\Pi}(\mathcal{W})$ . So, one possibility to estimate the limiting distribution is to numerically approximate the derivative  $\phi''_{r,\Pi}$  by computing

$$\hat{\phi}''_{r,\hat{\Pi}}(C) = \frac{\phi(\hat{\Pi} + h_n C) - \phi(\hat{\Pi})}{h_n^2}$$

for some sequence  $h_n \downarrow 0$ . However, in practice, it is quite difficult to choose the sequence  $h_n$ .

An alternative method is to replace  $r_0$  with a consistent estimator  $\hat{r}_0$ . For example, one can iteratively apply the Kleibergen and Paap test to obtain  $\hat{r}_0$ . The exact algorithm then reads as

1. Compute the singular value decomposition of  $\hat{\Pi}$ .
2. Sequentially apply the Kleibergen-Paap algorithm for each  $r' = 0, \dots, r$  with level  $\beta = \alpha/10$ .
3. Define  $\hat{r}_0$  as the smallest  $r'$  for that the Kleibergen-Paap test accepted. If it accepted in none of the iterations, reject the null hypothesis.
4. Compute  $E$  bootstrap samples  $\mathcal{W}_e^*$  of  $\mathcal{W}$ .
5. Based on that compute  $E$  estimates

$$\sum_{j=r-\hat{r}_0+1}^{k-r} \sigma_j^2 (U_2^T \mathcal{W}_1^* V_2), \dots, \sum_{j=r-\hat{r}_0+1}^{k-r} \sigma_j^2 (U_2^T \mathcal{W}_E^* V_2) \quad (4.7)$$

of the limiting distribution.

6. Let  $\hat{c}_{1-\alpha+\beta}$  be the  $(1 - \alpha + \beta)$ th quantile of the values in (4.7).
7. Reject if  $n \sum_{j=r+1}^k \sigma_j^2 (\hat{\Pi}_n) > \hat{c}_{1-\alpha+\beta}$ .

The motivation behind using  $\beta$  and  $\alpha - \beta$  in steps 2. and 7. is that the test procedure should have level  $\alpha$  in total. Choosing level  $\beta$  in step 2., and  $\alpha - \beta$  in the last step, should lead to a rejection rate of approximately  $\beta$  in step 2. and  $\alpha - \beta$  in the last step. Taken together, we should obtain the desired size. A formal proof that the size is indeed bounded by  $\alpha$  can be found in [6, Theorem 3.3].

## 4.2 Tests Using Minors

### 4.2.1 Incomplete $U$ -Statistic

The second approach is to assess whether all full minors of  $M$  are zero. Hence, one needs to test for several polynomials in the moments of  $\mathbf{X}$  if they vanish. For that, the incomplete  $U$ -statistic by Sturman, Drton, and Leung [29] is used. The statistic is a method to test polynomial inequalities in parameters of a multivariate distribution  $\mathbf{X}$ . Here, it is presented how the statistic works for polynomial equality constraints regarding the moments of the distribution since this is the situation relevant for this thesis.

To illustrate the idea behind the incomplete  $U$ -statistic, first an example is examined. Consider a two-dimensional random variable  $\mathbf{X} = (X_1, X_2) \in \mathbb{R}^2$  and the polynomial

$$f(\mathbb{E}(X_1), \mathbb{E}(X_2)) = \mathbb{E}(X_1)\mathbb{E}(X_2).$$

If one wants to test whether  $f$  is zero one can use the sample mean of  $\mathbf{X}$  and define the plug-in statistic

$$T_n(\mathbf{X}) = \bar{X}_1 \cdot \bar{X}_2 = \frac{1}{n^2} \sum_{i,j=1}^n X_1^{(i)} X_2^{(j)}.$$

However, this statistic is biased as

$$\begin{aligned} \mathbb{E} \left( \frac{1}{n^2} \sum_{i,j=1}^n X_1^{(i)} X_2^{(j)} \right) &= \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E} \left( X_1^{(i)} X_2^{(j)} \right) \\ &= \frac{1}{n^2} \sum_{i,j=1, i \neq j}^n \mathbb{E}(X_1^{(i)})\mathbb{E}(X_2^{(j)}) + \frac{1}{n^2} \sum_{i=1}^n \mathbb{E} \left( X_1^{(i)} X_2^{(i)} \right) \\ &= \frac{n^2 - n}{n^2} \mathbb{E}(X_1^{(1)})\mathbb{E}(X_2^{(1)}) + \frac{n}{n^2} \mathbb{E} \left( X_1^{(i)} X_2^{(i)} \right), \end{aligned}$$

which is in general not equal to  $\mathbb{E}(X_1^{(1)})\mathbb{E}(X_2^{(1)})$ . To obtain an unbiased estimator, one can instead use

$$U_n(\mathbf{X}) = \frac{1}{n^2 - n} \sum_{i,j=1, i \neq j}^n X_1^{(i)} X_2^{(j)}. \quad (4.8)$$

This is the so-called  $U$ -statistic, on which the incomplete  $U$ -statistic builds on. The "U" in its name stands for unbiased. To give the definition of the  $U$ -statistic for general polynomials, some notation is introduced. Let

- $\mathcal{M}^\eta$  be the set of all moments with order at most  $\eta$  of  $\mathbf{X}^{(1)}$ ,
- $f = a_0 + \sum_{r=1}^{\deg(f)} \sum_{\mu_1, \dots, \mu_r \in \mathcal{M}^\eta} a(\mu_1, \dots, \mu_r) \mu_1 \cdots \mu_r$  be an arbitrary polynomial in the moments of  $\mathbf{X}^{(1)}$ ,

- $\mathbf{X}^{(l,\dots,m)}$  the subsample of  $(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)})$  consisting of  $(\mathbf{X}^{(l)}, \dots, \mathbf{X}^{(m)})$ ,
- $\hat{\mu}_i(\mathbf{X}^{(l,\dots,m)})$  be the sample moment of  $\mu_i$  obtained from the subsample  $\mathbf{X}^{(l,\dots,m)}$ .

Likewise as for the example (4.8) above we want to define the general  $U$ -statistic as a mean of several unbiased estimators of  $f$  that are only based on a part of the sample. More precisely, we use the estimator  $\check{h}$

$$\check{h}(\mathbf{X}^{(1,\dots,d)}) = a_0 + \sum_{r=1}^d \sum_{\mu_1, \dots, \mu_r \in \mathcal{M}^n} a_{\mu_1, \dots, \mu_r} \hat{\mu}_1(\mathbf{X}^{(1)}) \hat{\mu}_2(\mathbf{X}^{(2)}) \cdots \hat{\mu}_r(\mathbf{X}^{(r)}),$$

with  $d = \deg(f)$ , as a basic building block for the statistic.  $\check{h}$  is an unbiased estimator of  $f$  since to estimate the factors of the monomials appearing in  $f$  independent parts of the sample are used and the sample moments  $\hat{\mu}_i(X_{(l,\dots,d)})$  are unbiased estimators.

To obtain an estimator  $h$  that is additionally symmetric in the sense that it does not depend on the order of the samples  $(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)})$  the average over all permutations is taken:

$$h(\mathbf{X}^{(1,\dots,d)}) = \frac{1}{d!} \sum_{\pi \in S_d} \check{h}_j(\mathbf{X}^{(\pi(1), \dots, \pi(d))}).$$

For the example from above,  $h$  takes the form

$$h(\mathbf{X}^{(1,2)}) = \frac{1}{2} \left( X_1^{(1)} \cdot X_2^{(2)} + X_1^{(2)} \cdot X_2^{(1)} \right).$$

Finally, the  $U$ -statistic is given by

$$U_{n,N} = \frac{1}{N} \sum_{\iota \in I_{n,d}} h(\mathbf{X}^{(\iota)}), \quad (4.9)$$

where

$$I_{n,d} = \{(i_1, \dots, i_d) : 1 \leq i_1 < \dots < i_d \leq n\}$$

is the set containing all ordered subset of size  $d$  of  $[n]$ . This statistic is quite commonly used to test polynomial constraints. However, it possesses two drawbacks. First,  $|I_{n,d}|$  grows exponentially in  $d$  and linear in the number of polynomials. For the purpose of testing all full minors of  $M$  we need to consider  $\binom{\frac{1}{2}p(p-1)+p}{p+1}$  polynomials of degree  $p$  so that the  $U$ -statistic becomes far too costly for higher  $p$ . Secondly, there is no known method how to estimate the limiting distribution of the statistic in the case where  $\text{Var}_{x \sim \mathbf{X}^{(2)}}(\mathbb{E}(h(x, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(d)}))) = 0$ , which might happen in our case. For example, if one tests for the determinant of  $M$  for  $p = 2$  and  $M$  has rank even lower than two, then

$$\begin{aligned} \check{h}(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}^{(3)}) &= \hat{s}_{11}(\mathbf{X}^{(1)}) \hat{t}_{112}(\mathbf{X}^{(2)}) \hat{t}_{222}(\mathbf{X}^{(3)}) - \hat{s}_{11}(\mathbf{X}^{(1)}) \hat{t}_{122}(\mathbf{X}^{(2)}) \hat{t}_{122}(\mathbf{X}^{(3)}) \\ &\quad - \hat{s}_{12}(\mathbf{X}^{(1)}) \hat{t}_{111}(\mathbf{X}^{(2)}) \hat{t}_{222}(\mathbf{X}^{(3)}) + \hat{s}_{12}(\mathbf{X}^{(1)}) \hat{t}_{122}(\mathbf{X}^{(2)}) \hat{t}_{112}(\mathbf{X}^{(3)}) \\ &\quad + \hat{s}_{22}(\mathbf{X}^{(1)}) \hat{t}_{111}(\mathbf{X}^{(2)}) \hat{t}_{122}(\mathbf{X}^{(3)}) - \hat{s}_{22}(\mathbf{X}^{(1)}) \hat{t}_{112}(\mathbf{X}^{(2)}) \hat{t}_{112}(\mathbf{X}^{(3)}). \end{aligned}$$



Consequently,

$$\begin{aligned} \mathbb{E}(h(x, \mathbf{X}^{(2)}, \mathbf{X}^{(3)})) &= \frac{1}{3}(\hat{s}_{11}(x)(t_{112}t_{222} - t_{122}^2) \\ &\quad + \hat{s}_{12}(x)(t_{122}t_{112} - t_{111}t_{222}) \\ &\quad + \hat{s}_{22}(x)(t_{111}t_{122} - t_{112}^2) \\ &\quad + \hat{t}_{111}(x)(s_{12}t_{222} - s_{22}t_{122}) \\ &\quad + \hat{t}_{112}(x)(s_{11}t_{222} - s_{22}t_{121} + s_{12}t_{212} - s_{22}t_{112}) \\ &\quad + \hat{t}_{122}(x)(s_{11}t_{122} - s_{12}t_{121} + s_{11}t_{212} - s_{22}t_{111}) \\ &\quad + \hat{t}_{222}(x)(s_{11}t_{112} - s_{12}t_{111})) = 0 \end{aligned}$$

since all  $2 \times 2$ -minor of  $M$  are zero. Likewise, for testing the full minors of a general matrices  $A$  consisting of moments of the distribution,  $\mathbb{E}(h(x, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(d)}))$  is zero if the minors of one order lower vanish.

These obstacles can be circumvented by randomly choosing a subset of the summands in (4.9) which then gives the incomplete  $U$ -statistic. For the formal definition, let  $N \leq \binom{n}{d}$  and define  $Z_\iota : \iota \in I_{n,d}$  as Bernoulli random variables with success probability  $N/\binom{n}{d}$ . Then the incomplete  $U$ -statistic is given by

$$U'_{n,N} = \frac{1}{\hat{N}} \sum_{\iota \in I_{n,d}} Z_\iota h(\mathbf{X}^{(\iota)}),$$

where  $\hat{N} = \sum_{\iota \in I_{n,d}} Z_\iota$  is the number of successes.

As derived in [29], for this modified  $U$ -statistic, the limiting distribution can be obtained without needing to make assumptions on the variance  $\text{Var}_{x \sim \mathbf{X}^{(2)}}(\mathbb{E}(h(x, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(d)})))$ . Additionally, the lower  $N$  is chosen, the lower the computational effort.

For the choice  $N = \binom{n}{d}$  the success probability is 1, the incomplete  $U$ -statistic is the same as the  $U$ -statistic.

To now test several polynomials  $f_1, \dots, f_q$  at once, so, to test the null hypothesis,

$$H_0 : f_i(\mathbf{X}) = 0 \text{ for all } i = 1, \dots, q,$$

the maximum of the studentized incomplete  $U$ -statistics

$$\mathcal{T} = \max_{1 \leq j \leq q} \sqrt{n} U'_{n,N,j} / \hat{\sigma}_j,$$

where  $\hat{\sigma}_j$  is an estimate of the variance of  $U'_{n,N,j}$ , should be used as the test statistic. Thus, we need to know how to derive the estimates  $\hat{\sigma}_j$  and how to obtain the critical values for the statistic.

First, we look at the limiting distribution of  $U_{n,N,j}$ . For that, we need to impose some conditions on the functions  $h_j$  and  $g_j$ , which is defined as  $g_j(x) = \mathbb{E}(h_j(x, \mathbf{X}^{(2, \dots, m)}))$ , as

well as on the distribution  $P^X$ . Two of them include the so-called sub-Weibull property. A random variable  $X$  is called sub-Weibull of order  $\gamma$  if

$$\|X\|_{\psi_\gamma} = \inf\{t > 0 : \mathbb{E}\left(\psi_\gamma\left(\frac{|X|}{t}\right)\right) \leq 1\} < \infty\},$$

where the function  $\psi_\gamma$  is given by  $\psi_\gamma = \exp(x^\gamma) - 1$  on  $\mathbb{R}_{>0}$ . Note that  $\inf \emptyset = \infty$ .

The conditions read as follows. There need to exist a  $D_n \geq 1$ , and a  $\gamma \in (0, 1]$  such that for all  $h_j$  the following holds:

$$\mathbb{E}\left(|h_j(\mathbf{X}^{(1,\dots,d)}) - \mu_j|^{2+l}\right) \leq \sigma_{h,j}^2 D_n^l \text{ for } l = 1, 2, \quad (\text{C1})$$

$$\|h_j(\mathbf{X}^{(1,\dots,d)}) - \mu_j\|_{\psi_\gamma} \leq D_n, \quad (\text{C2})$$

$$\sigma_{h,j}^2 > 0. \quad (\text{C3})$$

As already mentioned, contrary to the complete  $U$ -statistic, the variances of the  $g_j$  are allowed to vanish. Instead, it suffices that for each  $j = 1, \dots, q$  one of the following bounds hold. Either

$$\sigma_{g,j}^2 > 0, \quad (\text{C4})$$

or,

$$\text{there exist a } \kappa_j \text{ such that } \|g_j(\mathbf{X}^{(1)}) - \mu_j\|_{\psi_\gamma} \leq n^{-\kappa_j} D_n \quad (\text{C5})$$

needs to be true. Further, all  $g_j$  need to satisfy

$$\mathbb{E}\left(|g_j(\mathbf{X}^{(1)}) - \mu_j|^{2+l}\right) \leq \sigma_{g,j}^2 D_n^l \text{ for all } j = 1, \dots, q \text{ and } l = 1, 2. \quad (\text{C6})$$

Finally, as the last requirement we impose

$$N = \mathcal{O}(n). \quad (\text{C7})$$

This last assumption can be violated if the variances of all the  $g_j$  are bounded away from zero. Under these assumptions the following theorem about the limiting distribution of the statistic holds true.

**Theorem 4.7.** *If the conditions (C1) – (C7) are fulfilled, then there exist a  $C \in \mathbb{R}$  and a  $\kappa \in (0, 1)$ , both only depending on  $m, q$  and the constants  $\gamma, \kappa_j, \sigma_{h,j}^2, \sigma_{g,j}^2$  appearing in conditions (C1) – (C7), such that*

$$\sup_{R \in \mathbb{R}_{re}^q} |\mathbb{P}(\sqrt{n}(U'_{n,N} - \mu) \in R) - \mathbb{P}(\mathbf{Y} \in R)| \leq CD_n n^{-\kappa},$$

where  $\mathbb{R}_{re}^q = \{\Pi_{j=1}^q [a_j, b_j] : a_j, b_j \in \mathbb{R} \cup \{-\infty, \infty\}\}$  is the set of all hyperrectangles and  $\mathbf{Y} \sim \mathcal{N}_q(0, d^2 W_g + \alpha_n W_h)$ , with  $\alpha_n = n/N$ ,  $W_h = \text{Cov}(h(\mathbf{X}^{(1,\dots,d)}))$ , and  $W_g = \text{Cov}(g(\mathbf{X}^{(1)}))$ .

The theorem follows from Theorem 3.1. and Proposition 3.3. in [29]. Using it, we can approximate the limiting distribution of the test statistic  $\mathcal{T}$  under the null hypothesis. We know that under the null hypothesis all  $f_j$  vanish. Hence,

$$\mathcal{T} = \max_{1 \leq j \leq q} \sqrt{n}(U'_{n,N,j} - f_j(\theta))/\hat{\sigma}_j.$$

As a consequence of the theorem, the distribution of the left-hand side can be approximated by

$$\max_{1 \leq j \leq q} Y_j/\sigma_j$$

with  $\sigma_j$  indicating the true covariance of  $U_{n,N,j}$ .

In practice, the covariance of  $\mathbf{Y}$ , and consequently also its distribution, are unknown. Therefore, again bootstrap is used. It is first explained how to approximate the covariance matrices  $W_h$  and  $W_g$ . Those approximations are then used to obtain an estimate of the approximation of  $\mathbf{Y}$  with bootstrap.

For retrieving an estimate of  $W_h$ , as in the previous case, one could create  $E$  bootstrap copies of the sample  $\mathcal{X}$  by sampling from  $\mathcal{X}$  with replacement and then compute  $E$  copies of estimates of  $h_j(\mathbf{X}^{(1,\dots,m)})$  to obtain an estimate of  $W_h$ . However, in the computation of the incomplete  $U$ -statistic we already evaluated  $h$  at different subsamples of the whole sample  $\mathcal{X}$ . It turns out that we can reuse these estimates of  $h$  to perform the bootstrap and save a lot of computational cost. The exact procedure is to approximate  $W_h$  as

$$\hat{W}_h \approx \frac{1}{\hat{N}} \sum_{\iota \in I_{n,m}} Z_\iota \left( h(\mathbf{X}^{(\iota)}) - U'_{n,N} \right) \left( h(\mathbf{X}^{(\iota)}) - U'_{n,N} \right)^T.$$

Now, we turn to  $W_g$ . Recall that  $g_j$  was defined as  $g_j(x) = \mathbb{E} \left( h_j(x, \mathbf{X}^{(2,\dots,m)}) \right)$ . Estimates of  $g$  are not already computed for obtaining  $U'_{n,N}$ . So, to first of all obtain  $E_1 \leq n$  estimates of  $g_j$  we choose  $E_1$  different indices  $i_e^* \in \{1, \dots, n\}$ . For each of the indices  $i_e^*$ , the remaining set  $\{1, \dots, n\} \setminus \{i_e^*\}$  is split into disjunct subsets  $\{\iota_{e,1}, \dots, \iota_{e,[(n-1)/(m-1)]}\}$  where each set is of the size  $m-1$ . Then, we obtain  $E_1$  estimates of  $g$  by computing

$$\hat{g}_e = \sum_{k=1}^{\lfloor (n-1)/(m-1) \rfloor} h \left( \mathbf{X}^{(i_e^*)}, \mathbf{X}^{(\iota_{e,k})} \right)$$

for each  $i^*$ . Those are then used to derive the empirical covariance matrix

$$\hat{W}_g = \frac{1}{E_1} \sum_{e=1}^{E_1} (\hat{g}_e - \bar{g})(\hat{g}_e - \bar{g})^T,$$

where  $\bar{g}$  denotes the mean of all the  $g_e$ . These two estimates can be used to derive  $\hat{\sigma}_j$  as follows. The diagonal elements of  $W_h$  and  $W_g$  can be used to estimate  $\sigma_{h,j}^2$  and  $\sigma_{g,j}^2$

$$\begin{aligned} \hat{\sigma}_{h,j}^2 &= \frac{1}{\hat{N}} \sum_{\iota \in I_{n,m}} Z_\iota \left( h_j(\mathbf{X}^{(\iota)}) - U'_{n,N,j} \right)^2 \\ \hat{\sigma}_{g,j}^2 &= \frac{1}{E_1} \sum_{e=1}^{E_1} ((\hat{g}_e)_j - \bar{g}_j)^2. \end{aligned}$$

Then the variance of  $U_{n,N,j}$  can be approximated by

$$\hat{\sigma}_j^2 = m^2 \hat{\sigma}_{g,j}^2 + \frac{n}{N} \hat{\sigma}_{h,j}^2.$$

What is still missing now, is to obtain  $E_2$  estimates of  $\mathbf{Y}$  that together form an estimate for its distribution. We utilize the fact that a normally distributed random variable with arbitrary mean and covariance can be obtained as a linear combination of standard normal variables. Specifically, if  $\zeta_\iota, \iota \in I_{n,m}$  are normally distributed, and we define

$$\hat{\mathbf{Y}}_h^* = \frac{1}{\hat{N}} \sum_{\iota \in I_{n,m}} \zeta_\iota \sqrt{Z_\iota} \left( h(\mathbf{X}^{(\iota)}) - U'_{n,N} \right),$$

then  $\hat{\mathbf{Y}}_h^*$  has a standard normal distribution with mean zero and covariance matrix  $\hat{W}_h$ . In the same way, for  $\zeta_{e_1}, e_1 = 1, \dots, E$  independent normally distributed random variables,

$$\hat{\mathbf{Y}}_g^* = \frac{1}{E_1} \sum_{e_1}^{E_1} \zeta_{e_1} (\hat{g}_{e_1} - \bar{g})$$

gives an estimate of  $\mathbf{Y}_g$ . Hence,  $\mathbf{Y}$  can be approximated by

$$\hat{\mathbf{Y}}^* = m^2 \hat{\mathbf{Y}}_g^* + \alpha_n \hat{\mathbf{Y}}_h^*.$$

To obtain multiple estimates of  $\mathbf{Y}$ , the procedure is repeated  $E_2$  times. So,  $E_2$  times, we sample sets of independent standard normal random variables  $\{\zeta_\iota : \iota \in I_{n,m}\} \cup \{\zeta_{e_1} : e_1 \in \{1, \dots, E_1\}\}$ , and with these samples we compute  $E_2$  bootstrap approximations  $\hat{\mathbf{Y}}_{e_2}^*$  of  $\mathbf{Y}$ .

Finally, we define  $\hat{c}_{1-\alpha}$  as the  $(1 - \alpha)$ th quantile of the  $\hat{\mathbf{Y}}_{e_2}^*, e_2 \in \{1, \dots, E_2\}$  and reject  $H_0$  if  $\mathcal{T} > \hat{c}_{1-\alpha}$ .

As a last point regarding the incomplete  $U$ -statistic, we look at what the requirements mean for our specific setup. As pointed out in Remark 2.4 in [29], in practice, usually all the conditions (C1) – (C6) except condition (C2) that the  $h_j$  need to be Sub-Weibull of order  $\gamma$  for some  $\gamma \in (0, 1]$  are satisfied. According to Lemma C.3 in [29], condition (C2) is fulfilled if all the estimators  $\hat{\mu}_i$  appearing in  $h$  are Sub-Weibull. For our case, this breaks then further down to the distributions of the noise terms satisfying the Sub-Weibull requirement. To see that, note that each  $\hat{\mu}_i$  takes either the form

$$X_i^{(1)} X_j^{(1)} \quad \text{or} \quad X_i^{(1)} X_j^{(1)} X_k^{(1)}.$$

Moreover,

$$\mathbf{X}^{(1)} = E^{-T} \boldsymbol{\epsilon}.$$

Hence, each  $\hat{\mu}_i$  consists of a product of a linear combination of components of  $\boldsymbol{\epsilon}$ . But, both, linear combinations as well as products, preserve the Sub-Weibull property according to Lemma C.1 in [29] and Proposition D.2 in [14].

### 4.2.2 Independent Test

While the incomplete  $U$ -statistic is already much faster than its complete counterpart, the computational effort is still quite high and may be too high if  $p$  or the sample size is large. An alternative method that in principle builds on the same idea is the independent statistic. As the incomplete  $U$ -statistic can be thought of as a way to trade some power for a lower computational effort compared to the complete  $U$ -statistic, the independent statistic goes even one step further in that direction.

Again, a subset of the summands of the  $U$ -statistic is chosen. The only difference is that instead of randomly choosing a subset of the indices, the sample is split into several parts of size  $d$  to that the estimator  $h$  is applied. Then, the average of all evaluations of  $h$  forms the independent statistic.

$$I(\mathbf{X}) = \frac{d}{N} \sum_{\iota \in \{(1, \dots, d), (d+1, \dots, 2d), (n-d+1, \dots, n)\}} h(\mathbf{X}^{(\iota)}).$$

As fewer summands are taken into account, the computational effort is lower. On the other hand, the variance of the statistic is higher and therefore power is lower.

The test statistic is then again defined as the studentized maximum of  $I$

$$\mathcal{T} = \max_{1 \leq j \leq q} \sqrt{n} U'_{n,N,j} / \hat{\sigma}_j,$$

where the sample variance is calculated as

$$\hat{\sigma}_j = \frac{d}{N} \sum_{\iota \in \{(1, \dots, d), (d+1, \dots, 2d), (n-d+1, \dots, n)\}} \left( h(\mathbf{X}^{(\iota)}) - I(\mathbf{X})_j \right)^2.$$

For the independent statistic, the limiting distribution is

$$\sqrt{n} (U'_{n,N} - f(\theta)) \xrightarrow{d} \mathcal{N} \left( 0, \text{Cov}(h(\mathbf{X}^{(1, \dots, m)})) \right).$$

So, with  $\hat{W}_h$  defined in a similar manner as for the incomplete  $U$ -statistic, specifically as,

$$\hat{W}_h = \frac{d}{N} \sum_{\iota \in \{(1, \dots, d), (d+1, \dots, 2d), (n-d+1, \dots, n)\}} \left( h(\mathbf{X}^{(\iota)}) - I(\mathbf{X}) \right) \left( h(\mathbf{X}^{(\iota)}) - I(\mathbf{X}) \right)^T,$$

an approximation of the limiting distribution of  $\mathcal{T}$  can be obtained.

## 5 Numerical Experiments

In this section, the performance of the four proposed methods is examined with simulated as well as real-world data.<sup>1</sup> First, the results obtained using simulated data are presented. We look at the distribution of the  $p$ -values, the empirical size, and the empirical power of the four algorithms. Recall that the result of a test always depends on the choice of the level  $\alpha$ . The  $p$ -value is defined as the smallest level  $\alpha$  at that a test rejects for the present sample. Observing the  $p$ -value has the advantage that one can assess the behavior of a test without needing to fix a level. Under the null hypothesis, the  $p$ -values ideally should be uniformly distributed. If the  $p$ -values tend to be low, the test rejects quite often. Hence, the size is not controlled. If they tend to be high, the method is too conservative in the sense that it accepts too often. Moreover, the empirical size as well as the empirical power is the ratio of rejections obtained when repeatedly applying the test. The difference is that one speaks of empirical size when the data is generated in a way that the null hypothesis is fulfilled, while for calculating the empirical power data under an alternative is sampled. We always perform 1000 simulations, for obtaining a distribution of  $p$ -values, as well as for obtaining empirical sizes and powers.

We start with the behavior under the null hypothesis in the non-degenerate case that  $M$  has rank precisely  $p$ . Then, experiments for the degenerate case are executed. Afterwards alternatives are considered.

The last part of the section is devoted to the results that are obtained for the Tübingen benchmark data set.

### 5.1 Synthetical Data

#### 5.1.1 Null Hypothesis Setup Non-Degenerate Case

For examining the properties of the four algorithms under the null hypotheses in the non-degenerate case we sample data from a distribution following the linear SEM associated with the complete graph with two and three nodes, respectively. Specifically, the sample  $\mathbf{X}$  is obtained by

$$\mathbf{X} = B^{-T}\boldsymbol{\epsilon}, \quad (5.1)$$

where  $B$  is a matrix with ones on the diagonal and off-diagonal entries uniformly chosen from the interval  $[-1, 1]$ . For  $\boldsymbol{\epsilon}$ , we consider three choices. It is sampled as

- Beta( $\alpha, \beta$ ) with  $\alpha \in [0.5, 2]$ ,  $\beta \in [2, 10]$ ,
- Gamma( $\alpha, \beta$ ) with  $\alpha \in [1, 3]$ ,  $\beta \in [1, 5]$ , and

---

<sup>1</sup>The code with an implementation in R of all the proposed tests and code to reproduce all the results of the experiments can be found at <https://github.com/DanielaSchkoda/test-linear-SEM-assumption>.

- an overlapping Gaussian distribution, which means that half of the data is sampled as  $\mathcal{N}(-\mu, \sigma_1)$  and the other half as  $\mathcal{N}(\mu, \sigma_2)$  with  $\sigma_1 \in [0.5, 1]$ ,  $\sigma_2 \in [1, 3]$  and  $\mu \in [1.5, 2]$ .

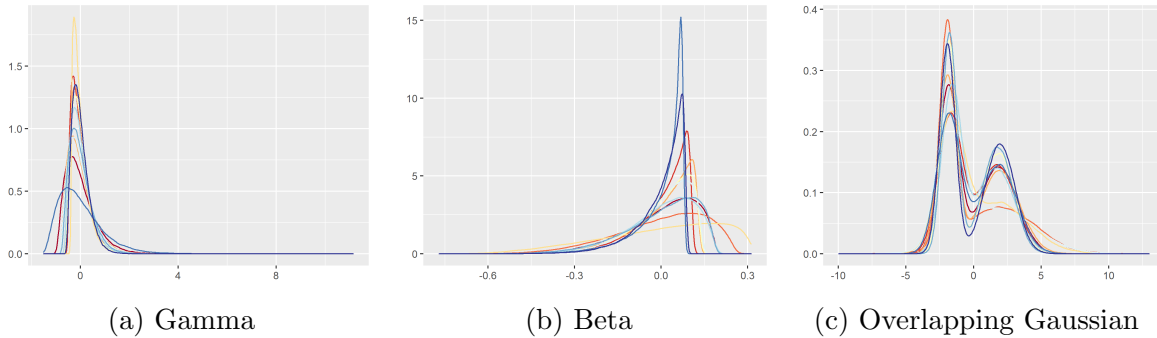


Figure 1: Distributions of the noise terms.

Those three distributions are chosen as they have third moments bounded away from zero so that they yield data for the non-degenerate situation.

Note that all choices satisfy the Sub-Weibull condition required for the incomplete  $U$ -statistic. Gamma distributions are Sub-Weibull for example with  $\gamma = 1$ . If  $X \sim \text{Gamma}(\alpha, \beta)$ , then for  $t = \frac{1}{c\beta}$ ,

$$\begin{aligned}
\mathbb{E} \left( \exp \left( \frac{|X|}{t} \right) \right) &= \int_0^\infty e^{\frac{x}{t}} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx \\
&= \int_0^\infty e^{\frac{x}{t}} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{\frac{x}{t} - \beta x} x^{\alpha-1} dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{\frac{x}{t} - \beta x} x^{\alpha-1} dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-(1-c)\beta x} x^{\alpha-1} dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha)}{((1-c)\beta)^\alpha} = (1-c)^\alpha
\end{aligned}$$

for  $c \leq 1$ . For  $c$  close enough to one, the last expression is certainly smaller than two. Hence, the distribution is Sub-Weibull.

Similarly, one can show for Gaussian distributions that they fulfill the requirement. This generalizes directly to overlapping Gaussian distributions.

For the beta distribution, note that the distribution is supported in  $[0, 1]$ . Hence, for arbitrary  $\gamma$

$$\mathbb{E} \left( \exp \left( \left( \frac{|X|}{t} \right)^\gamma \right) \right) \leq \exp \left( \frac{1}{t} \cdot \gamma \right),$$

which is smaller than two for  $t$  chosen sufficiently large.

We start with examining the distributions of the  $p$ -values for  $p = 2$  and the noise terms gamma-distributed. As one can see in figure 2, already for the sample size  $n = 250$ , the  $p$ -values are rather uniformly distributed. The KP statistic, as well as the Bootstrap SVD statistic, tend to overreject slightly, while the other two statistics are undersized.

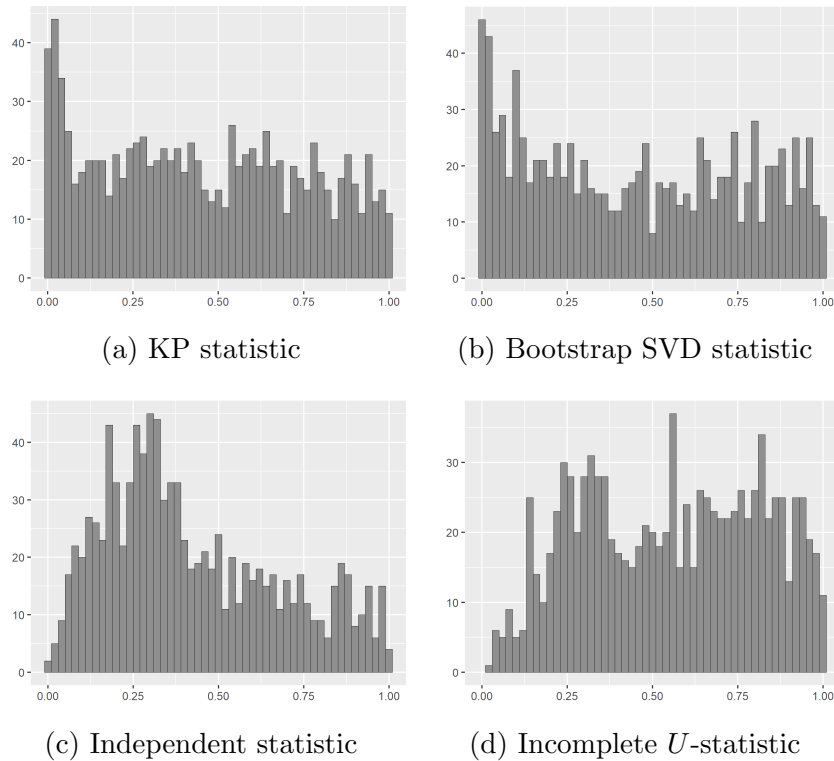


Figure 2:  $p$ -value distributions for  $p = 2$  and  $n = 250$ .

In order to see more clearly how the sizes of the four different procedures compare, the empirical sizes are plotted against the nominal sizes for all four tests in one plot.

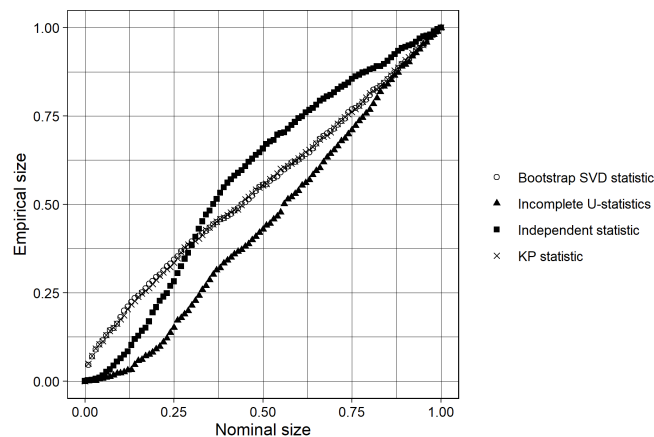


Figure 3: Empirical sizes for  $p = 2$  and  $n = 250$ .



For the higher sample size of 1000, the  $p$ -value are more uniformly distributed. However, the tendency of the first two tests to be oversized, and the tendency of the other two tests to be undersized still remain.

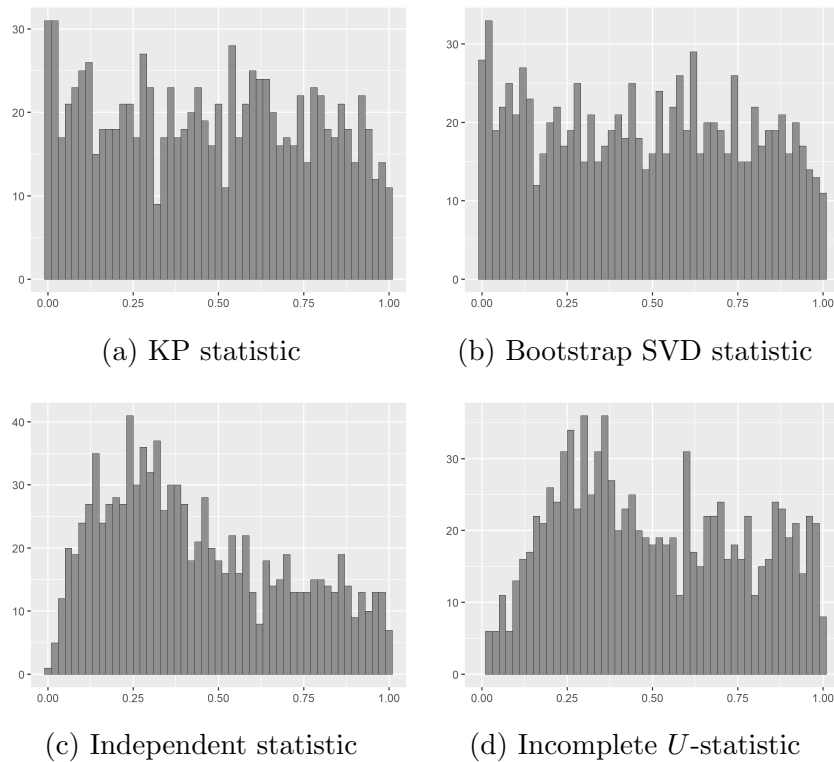


Figure 4:  $p$ -value distributions for  $p = 2$  and  $n = 1000$ .

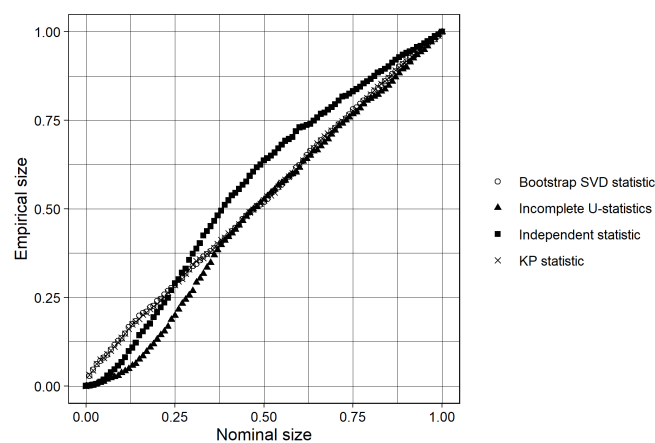


Figure 5: Empirical sizes for  $p = 2$  and  $n = 1000$ .

Next, the beta and the overlapping gaussian distribution are considered for sampling the noise term. In the size plot for  $n = 1000$ , one can see that the tests hold the size and that it does not make a huge difference which distribution is used to generate the noise terms.

We note that likewise for the remaining experiments that were executed, the choice of the noise term did not affect the results too much. Therefore, from now on, only the results for the gamma distribution setup are presented.

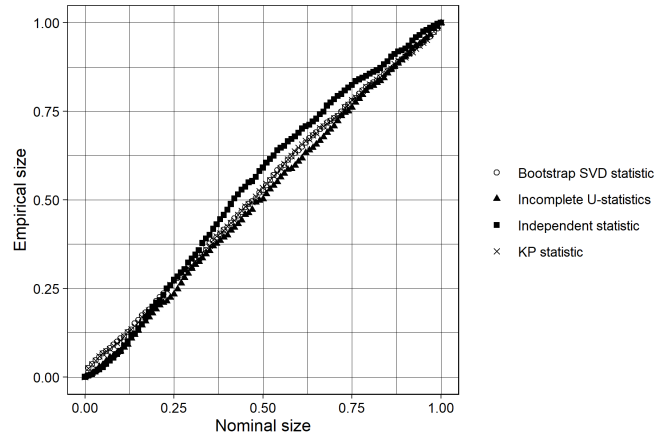


Figure 6: Empirical sizes for beta distribution

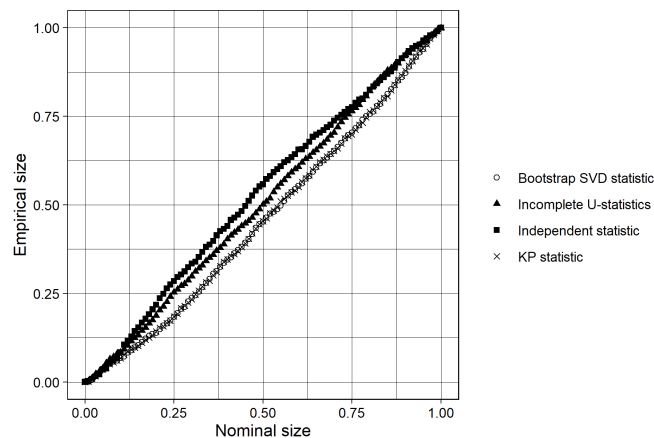
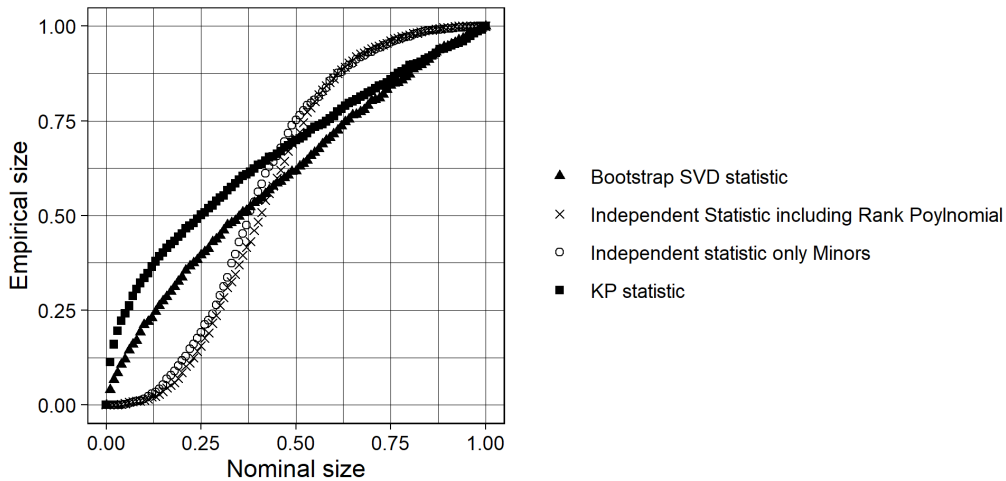


Figure 7: Empirical sizes for Overlapping Gaussian distribution

**Three nodes.** The next experiment aims at analyzing the performance of the algorithms for the graphical model with three nodes. In this case, the computation time of the incomplete  $U$ -statistic is too high. Therefore, only the three other tests are considered. Recall that for  $p = 3$ , under the linear SEM assumption,  $M$  has lower rank and additionally the Aronhold invariant for the tensor of third moments vanishes. To compare whether it makes a difference in practice if one only tests for the rank of  $M$  or also takes the Aronhold invariant into account the independent statistic is applied in two variants. Once we test if all minors of  $M$  are zero. In a second experiment, we examine if all minors are zero and at the same time, the Aronhold invariant vanishes.

Figure 8: Empirical sizes for  $p = 3$  and  $n = 250$ .

The plot shows that the KP statistic overrejects. Also, the empirical sizes of the Bootstrap SVD statistic tends to be too low but not as much as for the KP statistic. Furthermore, the results of the independent statistic do not vary to much between the two different setups. In both cases the tests are undersized for a small nominal size, and oversized for a higher nominal size.

### 5.1.2 Null Hypothesis Setup Degenerate Case

An especially critical point is how the tests perform under the degenerate case that  $M$  has even lower rank than  $p$  or is close to having even lower rank. Recall that under a linear SEM the rank can be strictly smaller than  $p$  only if the  $\epsilon$  has at least one component with vanishing third moment. Nonetheless, also for noise terms with non-zero third moments, the distribution of  $\mathbf{X}$  can be close to the degenerate case. For instance, if  $p = 2$  and  $X_2$  equals  $X_1$  plus a very small error term, then both components of  $\mathbf{X}$  are highly correlated and have very similar marginal distributions. Consequently, all the second moments almost coincide with each other. The same holds for the third moments. Thus, all columns of  $M$  almost coincide with each other.

We execute experiments for both setups. First, it is tested how the algorithms perform if  $\epsilon$  is drawn from a normal distribution, and apart from that  $\mathbf{X}$  is sampled as in the above setup (5.1). For analyzing the behavior in the case where all components have a similar distribution and are highly correlated, we obtain a sample of  $\mathbf{X}$  by defining

$$X_1 = \epsilon_1 \quad \text{and} \quad X_2 = X_1 + 0.001\epsilon_2$$

where  $\epsilon$  is sampled from a gamma distribution with the same parameters as for the non-degenerate setup.

For the Gaussian setup, all four tests hold the size. Only the independent statistic exhibits modest overrejection. Interestingly, even the empirical size of the Kleibergen-Paap statistic is very close to the nominal size even though its theoretical foundation does not take the possibility of a rank strictly lower than  $p$  into account and it was shown by [6] that in some degenerate situations the statistic fails to control the size.

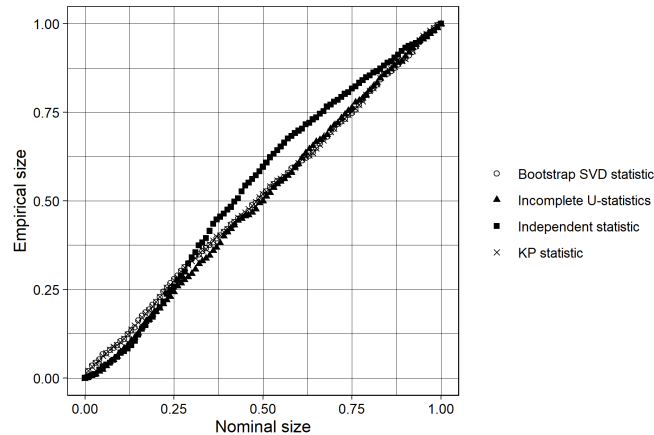


Figure 9: Empirical sizes with Gaussian noise terms for  $p = 2$  and  $n = 1000$ .

The experiment where  $X_2$  is the samples as  $X_1$  plus a small disturbance, is the first experiment with a huge difference between the Kleibergen-Paap statistic and the bootstrap SVD based statistic. The former one is significantly undersized, whereas the latter one is only slightly undersized. The sizes of the incomplete  $U$ -statistic match the nominal sizes quite well and the sizes of independent statistic tend to be a bit too high.

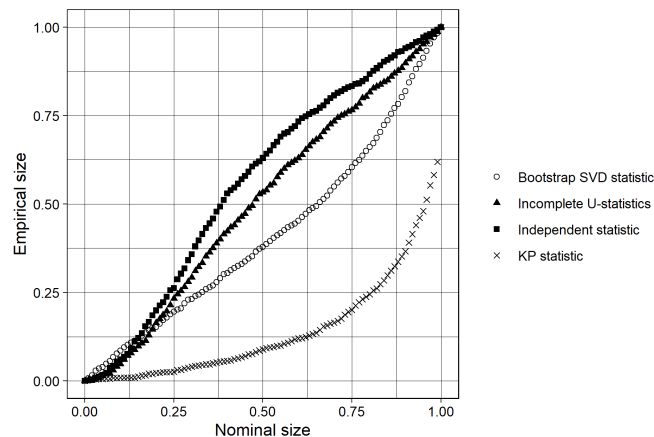


Figure 10: Empirical sizes under high correlation for  $p = 2$  and  $n = 1000$ .

### 5.1.3 Power Analysis

For the alternative hypothesis, we look at the behaviour of the tests for the data sampled as

$$X = \cos(B^{-T}\epsilon).$$

Additionally, we consider the setup

$$X_1 = \epsilon_1 \quad X_2 = \delta(X_1 - 1)^2 + X_1 + \epsilon_2$$

for the parameter  $\delta$  approaching zero, to analyze the properties of the algorithms under local alternatives. The edge weights  $B$  and the noise term  $\epsilon$  are sampled in the same way as for the null hypothesis setting.

For the first setup and sample size  $n = 250$ , the empirical power at level  $\alpha = 0.05$  obtained of the KP statistic and the bootstrap statistic is 0.84. For the incomplete  $U$ -statistic it is 0.34 and for the independent statistic 0.29.

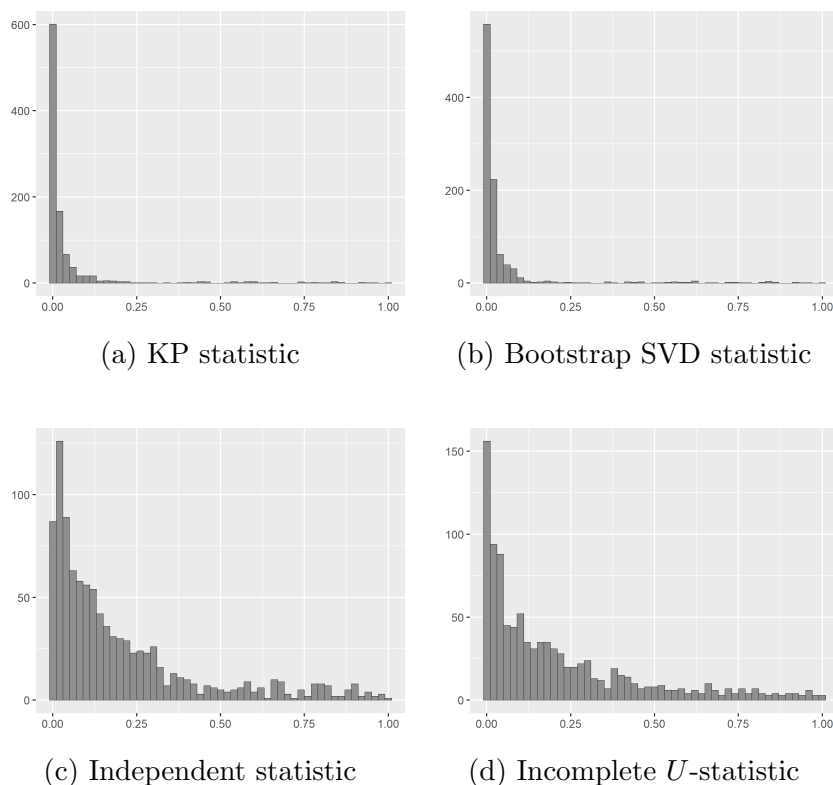
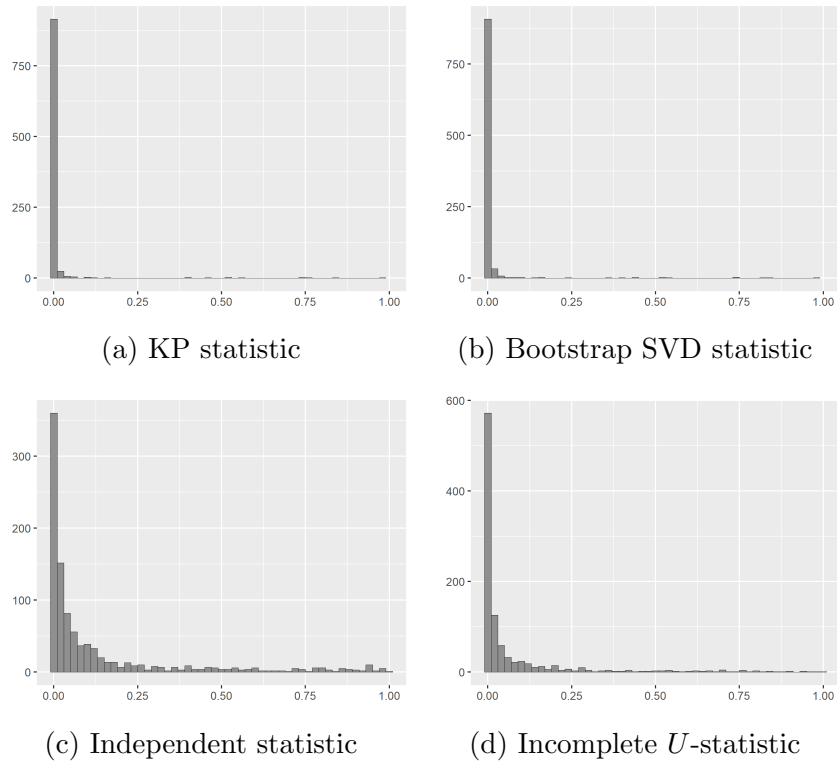
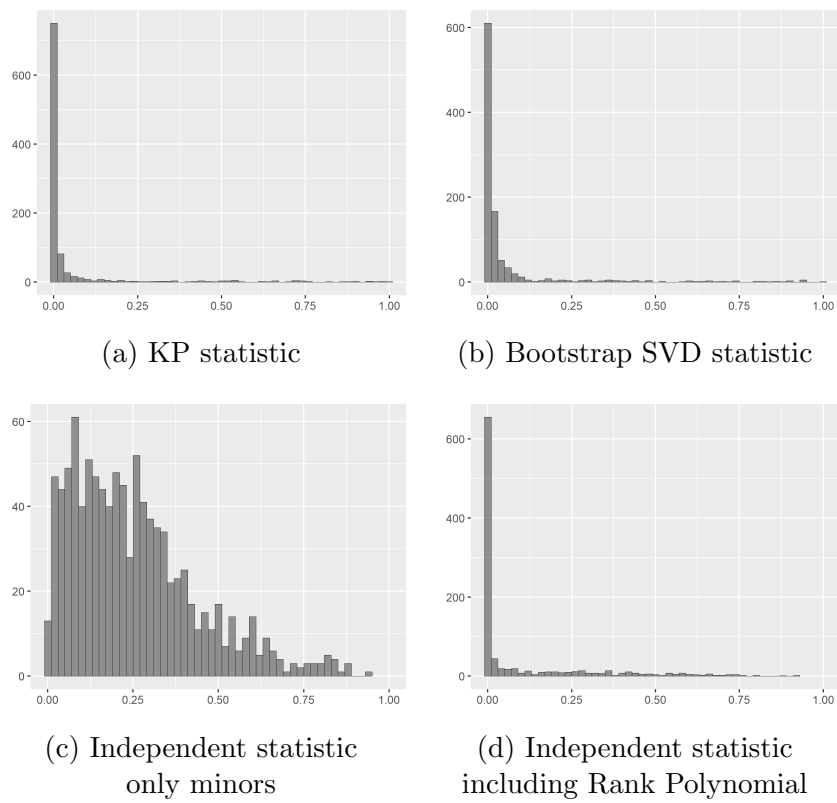


Figure 11:  $p$ -value distributions for  $p = 2$  and  $n = 250$ .

For higher sample size, the power improves for all of the tests: For both singular value based tests the power at level  $\alpha = 0.05$  is 0.95. For the incomplete  $U$ -statistic it improves to 0.75, and for the independent statistic it reaches 0.6.

Figure 12:  $p$ -value distributions for  $p = 2$  and  $n = 1000$ .Figure 13:  $p$ -value distributions for  $p = 3$  and  $n = 250$ .

**Three nodes.** For the case of three nodes, again the singular value based tests, and the independent test are considered. The last-mentioned test is applied once to validate if all minors vanish and once to test if all minors and the Aronhold invariant vanish. As can be seen in figure 13, for  $n = 250$ , the singular value based tests have already decent power. It is visible that the power of the KP, as well as the SVD based bootstrap statistic are quite high. The independent statistic only testing for the minors yields by far the lowest power. Including the Aronhold invariant leads to a significant increase in the power. This procedure has power even higher than the singular value based procedures.

**Local Alternatives.** For the second alternative, we are interested in the power when the parameter  $\delta$  moves away from zero. The higher  $\delta$  is, the further away the data is from fulfilling the null hypothesis. Thus, the power should improve for increasing  $\delta$ . For  $\delta \in [0, 1]$ ,  $p = 2$ ,  $n = 250$ , and a fixed level  $\alpha = 0.05$ , we obtain the following results.

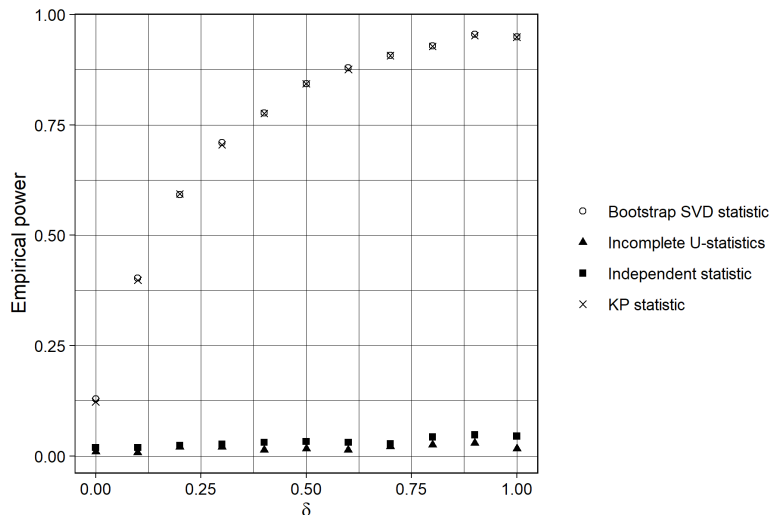


Figure 14: Empirical powers for  $p = 2$  and  $n = 250$ .

In this setting the performance of the singular value tests is much more favorable compared to the minor based test. For both singular value tests, the power increases quite rapidly for  $\delta$  approaching one. The power of the other two tests barely improves for higher  $\delta$ . This is likely caused by the low sample size. For a higher sample size, the power of the minor based tests increases. For instance, for  $n = 1000$ , the power for  $\delta = 1$  is 0.15 for the independent and 0.21 for the incomplete  $U$ -statistic.

## 5.2 Tübingen Dataset

In this section, the performance of the proposed tests is examined with real-world data. Specifically, we are using the Tübingen cause-effect-pair data sets [20]. The data is often used as a benchmark data set to assess the performance of causal inference algorithms including algorithms relying on the LiNGAM assumption [16].

The collection contains 108 data sets, each containing a cause-effect pair for which the ground truth is known. The data stems from different domains including meteorology, biology, engineering, and economy. For example, one pair consists of the day of the year and the mean daily temperature of Furtwangen in Germany for the years 1979 until 2004. Another pair contains the monthly income compared to the age from a study conducted in 1994 and 1995 in the U.S.. It is known that the day of the year has a causal effect on temperature, and the age on the income, but not vice versa.

The sample sizes vary between 72 and 16382. The distributions also take diverse forms. There are data sets with a rather normal distribution, as well as sets that are not normally distributed.

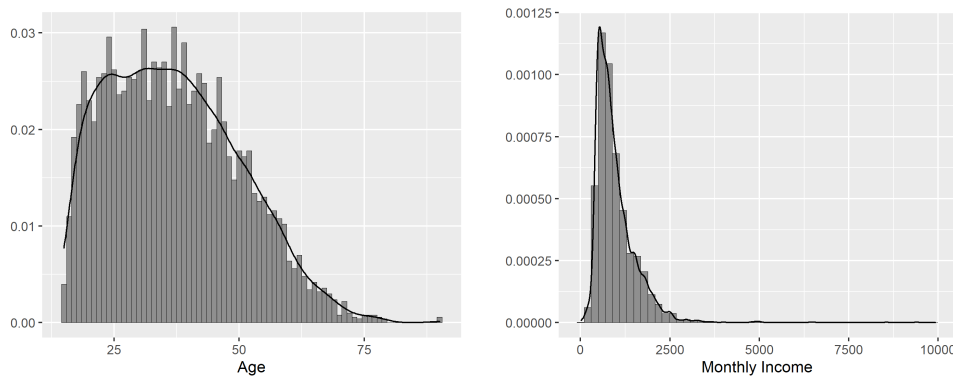


Figure 15: Distributions of the components of pair 12.

For nine data sets, the pair consists of two multivariate samples. We omit those pairs. For the remaining pairs, we apply each of the tests and present if it accepted or rejected the null hypothesis. For that, we fix the level to  $\alpha = 0.05$ .

As depicted in figure 16, the Kleibergen-Paap test rejects 50 of the pairs. While for many of the rejected pairs, the scatter plots suggest a non-linear relationship, it stands out that five pairs are rejected whose scatter plots seem to point to a linear relationship. Those cause-effect-pairs are 43, 44, 45, 46, and 84. All those pairs exhibit a high correlation and additionally both variables have a very similar distribution. Thus, a potential reason for the rejection could be that we are close to the degenerate case where the Kleibergen-Paap might give wrong results. However, in the simulation studies with synthetic data the Kleibergen-Paap statistic exhibited precisely the opposite behavior: It was undersized in the experiment presented for the situation of a high correlation and similar distributions of the components. Another potential explanation is that the scatter plot the distributions looks rather linear, while the distributions are actually not well-described by a linear SEM. This is for example the case if both components of  $\mathbf{X}$  are mostly effected by an unobserved third variable  $X_3$ . Such a hidden variable is also called *confounder* or *latent variable*. More precisely, consider a random vector  $\mathbf{X} \in \mathbb{R}^3$  solving the system of structural equations

$$\begin{aligned} X_1 &= X_3 + \epsilon_1, \\ X_2 &= X_3 + \epsilon_2, \\ X_3 &= \epsilon_3, \end{aligned}$$



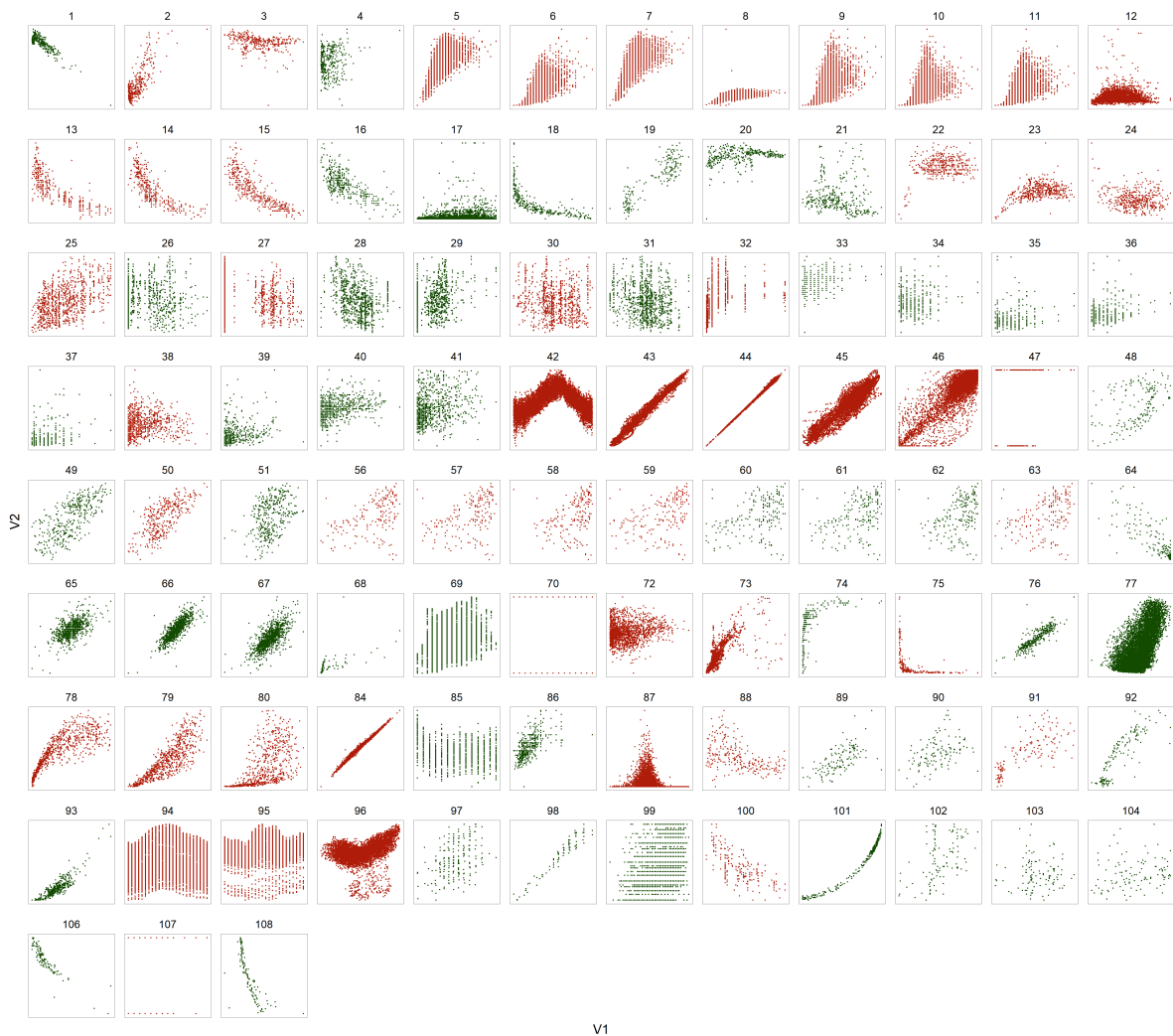
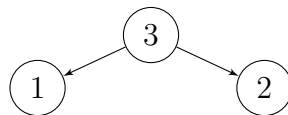


Figure 16: Results for the KP test. The pairs colored in red are rejected the green ones are accepted.

with  $\epsilon_i$  independent random variables, for the following graph:



If  $\epsilon_1$  and  $\epsilon_2$  are small compared to  $\epsilon_3$ , then the points of scatter plot of a sample drawn from the joint distribution of  $(X_1, X_2)$  will be accumulated close to the diagonal as for the scatter plots of the pairs 43, 44, 45, 46, and 84. However, the distribution of such random variables  $(X_1, X_2)$  in general does not belong to a linear SEM with two nodes. Thus, the rejection would be correct. According to [20], for these five pairs it is indeed likely that the components are confounded by a third variable.

For the SVD bootstrap test, additionally to the result for each pair, it is indicated what the estimated true rank  $r_0$  computed in step 3. of the algorithm is.

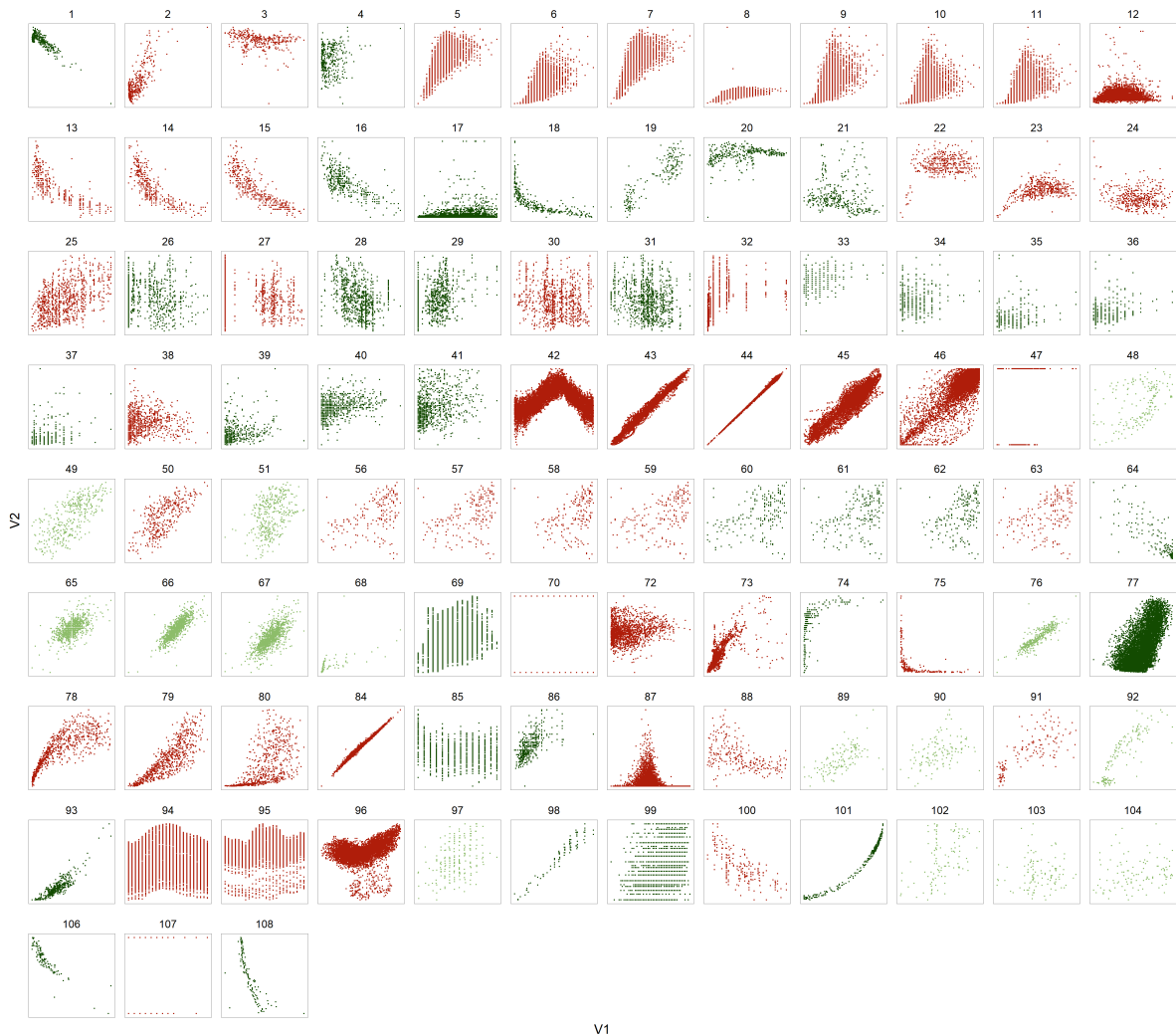


Figure 17: Results for the SVD bootstrap test. For all red pairs, the test rejected, for all green pairs, it was accepted. Light green means that the estimated true rank computed in step 3. is 1, dark green that it is 2.

49 of the pairs were rejected. Amongst them, 38 were rejected in step 3. and the rest in the last step. Again, the null hypothesis is rejected for the data sets 43, 44, 45, 46, and 84.

The independent statistic and the incomplete  $U$ -statistic are applied in two variants. First, the proposed test procedure is used to validate how well the data fits the graphical model for the complete graph.

As already mentioned, the LiNGAM algorithm does specifically assume that the distribution belongs to a graphical model related to an *acyclic* graph. For  $p = 2$  there are, apart from the empty graph, only two acyclic graphs. If  $P^X$  lies in the graphical model belonging to the acyclic graph  $1 \rightarrow 2$ , it is known that the matrix

$$\begin{pmatrix} s_{11} & s_{12} \\ t_{111} & t_{112} \\ t_{112} & t_{122} \end{pmatrix}$$

has rank 1, similarly for the graph  $2 \rightarrow 1$  [1]. So, a distribution belonging to one of the acyclic models necessarily has to fulfill one of the two equations systems. We apply both tests once with the equation system for the model  $1 \rightarrow 2$  and once with the equation system belonging to the graph  $2 \rightarrow 1$ . The maximum of both  $p$ -values is then used as a  $p$ -value for the hypothesis

$$H_0 : P_X \text{ belongs to an acyclic model}$$

The following results are obtained:

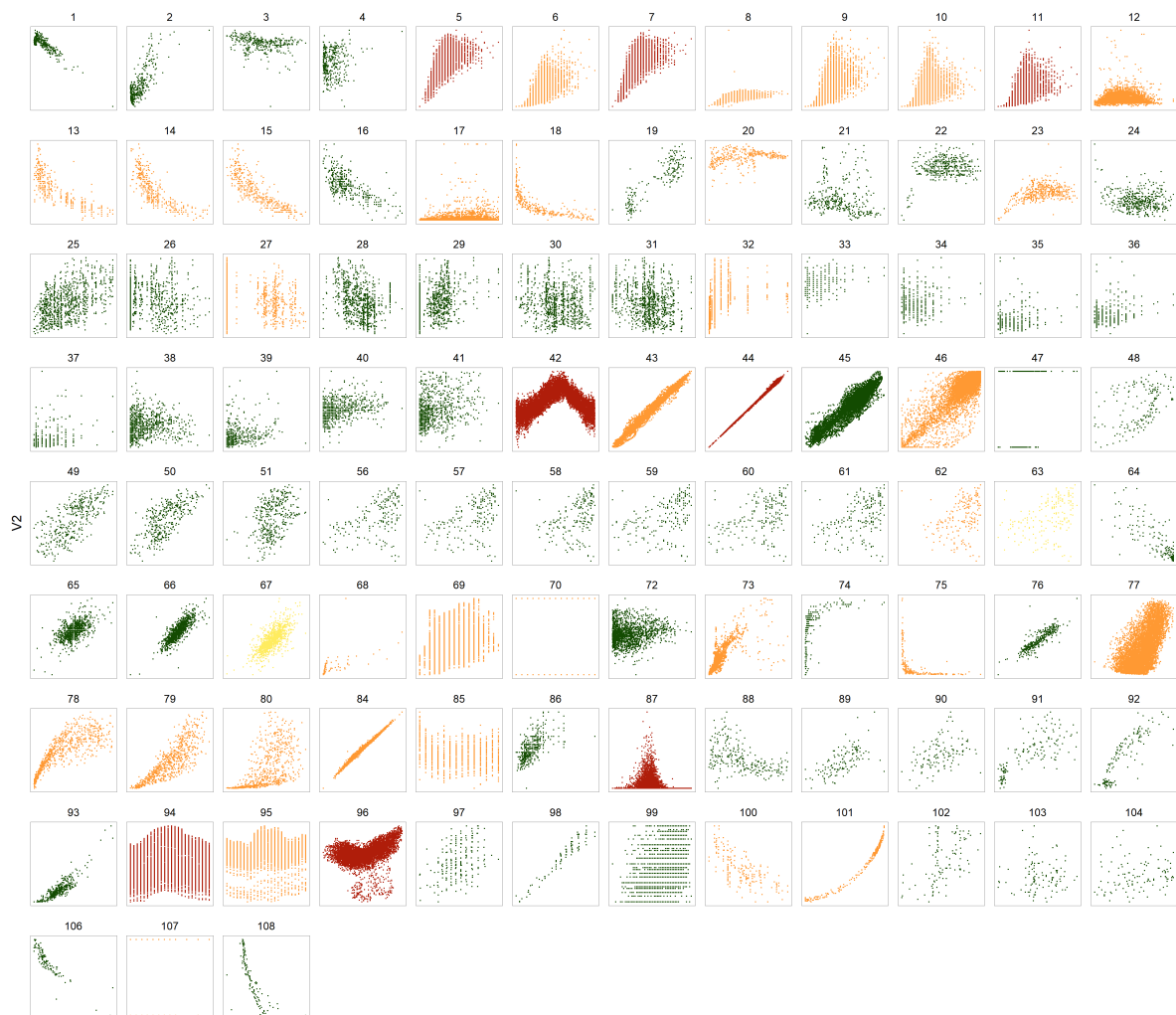


Figure 18: Results for the independent statistic: The red pairs are rejected by both variants of the test, the oranges one only by the test for the acyclic model, the yellow ones only by the test related to the complete graph, and all green pairs are accepted.

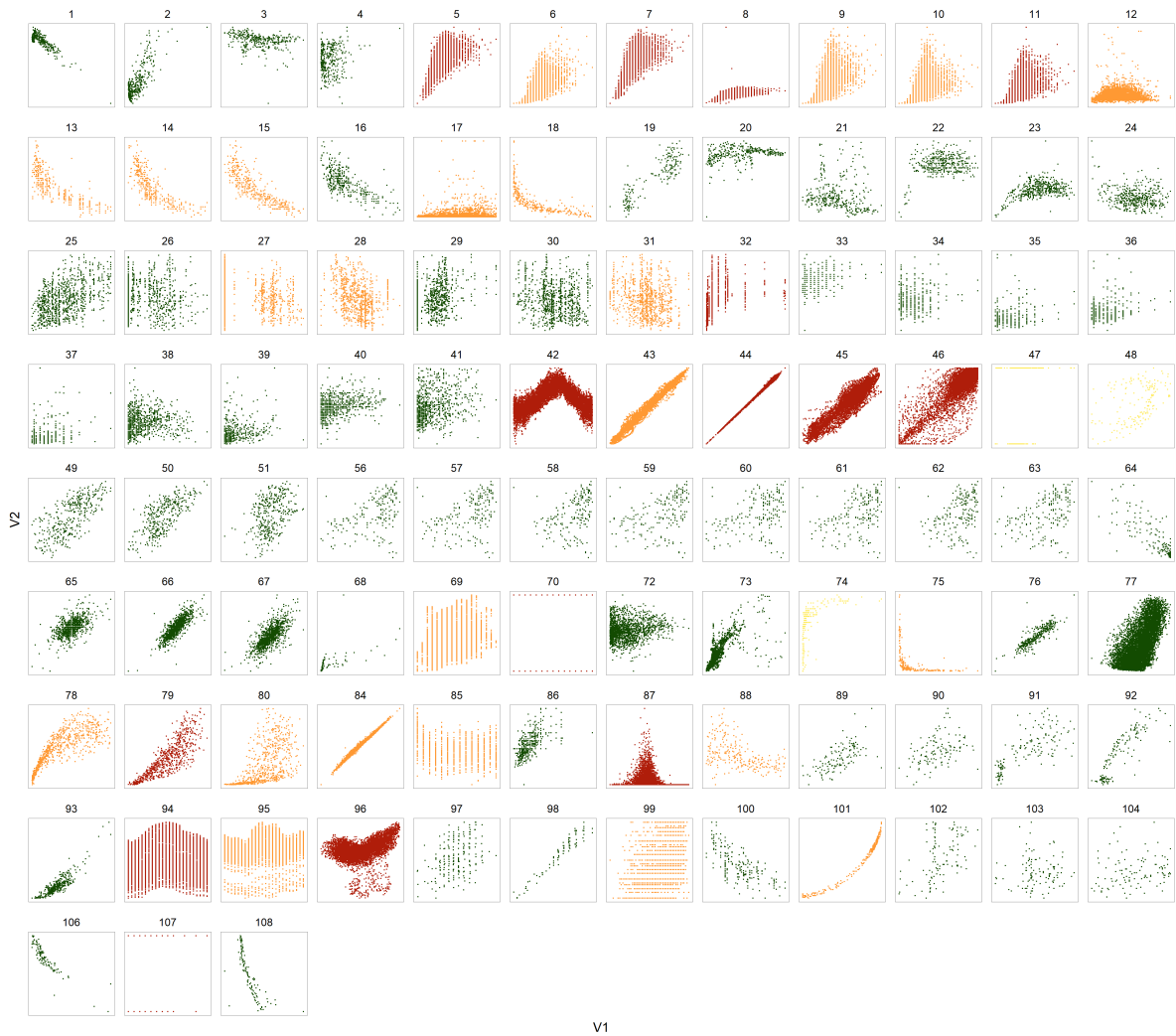


Figure 19: Results for the incomplete  $U$ -statistic: The red pairs are rejected by both variants of the test, the oranges one only by the test for the acyclic model, and all green pairs are accepted.

With both algorithms the null hypothesis that the underlying distribution lies in the model associated with the complete graph is rejected in rather few cases, specifically only nine times for the independent, and 13 times for the incomplete  $U$ -statistic. That aligns with the findings obtained with the synthetical data that the minor based statistics reject less often compared to the singular value based tests.

## 6 Conclusion and Outlook

In this thesis, we aimed at developing a test for assessing the validity of the assumption that a distribution belongs to a linear structural equation model. For that, we first examined the algebraic structure of the moments of such distributions. As a main theoretical result, we showed a characterization of the set of second and third moments realizable under the linearity assumption. Namely, the set is described by two conditions: that a matrix formed of second and third moments does not have full rank and at the same time the tensor of third moments has symmetric tensor rank not higher than the number of nodes  $p$ . For the second condition, we furthermore derived necessary polynomial equations. In the case of three nodes there exists only one polynomial equation, which is known as Aronhold invariant. Subsequently, we proved that also the higher moments of distributions contained in a linear SEM obey a certain structure. Further equations for the case where a fixed edge is not contained in the underlying causal graph were given for the case  $p = 3$ .

After laying out this theoretical basis, we turned to illustrate the design of different methodologies for testing the rank conditions and their asymptotic behavior under the null hypothesis. Namely, we presented the Kleibergen-Paap test and the bootstrap test by Chen, which both make use of the singular value decomposition. Then the incomplete  $U$ -statistic and the independent statistic for testing polynomial constraints were described.

In the last part, we compared the performance of all proposed methods with synthetical data as well as with the Tübingen cause-effect pairs. In the non-degenerate case, we saw that for all four tests the empirical sizes align well with the nominal sizes under the null hypothesis. The same is true for the degenerate case apart from one exception: The Kleibergen-Paap statistic is significantly undersized in the experiment where the components are highly correlated. In the power analysis in most of the experiments for two nodes, the singular value based tests outperform the other two methods. That is especially the case for rather low sample sizes. In contrast to this, for the case of three nodes, the independent statistic yields the most favorable result. The crucial point here is that the power is substantially better when testing for the rank of  $M$  as well as for the Aronhold invariant, which is only possible with the tests for polynomial constraints. For the Tuebingen pairs, both of the singular value based tests reject almost half of the pairs. The minor based tests reject rather few times. An especially outstanding observation was the rejection of some pairs with a seemingly linear relationship. Here more detailed investigations would be interesting to see if the conjecture that both pairs are mostly influenced by a latent variable, and are therefore correctly rejected, is true.

Further research could for example go in the direction of finding out more equations that hold if one restricts to specific graphs, as was done in this thesis for the case  $p = 3$  and one missing edge. While there already exist results for equations for directed graphs [1], there is rather few knowledge concerning cyclic graphs.

Regarding the purpose of assessing the linearity assumption for the complete graph, it would be interesting to study the case of higher  $p$  more detailed. A challenge is how to

deal with the high computational effort of the independent statistic and the incomplete  $U$ -statistic. While for testing the rank condition on  $M$  one can use the computational more efficient Kleibergen-Paap test or the rank test proposed by Chen instead, this is not possible for the polynomial equations resulting from the rank condition on the tensor of third moments.

Furthermore, one could examine if it is possible to derive polynomial inequalities that hold for tensors with real symmetric border rank at most  $p$ . While we noted that there can not exist polynomials that vanish for tensors with real symmetric border rank bounded by  $p$  but not for tensors with complex symmetric border rank bounded by  $p$ , there can exist polynomial inequalities distinguishing the real and complex symmetric border rank. As we saw in the simulation studies for three nodes, testing for the Aronhold invariant immensely improved the power. Thus, including more even conditions resulting from the symmetric tensor rank constraint might again lead to an increased power.

## References

- [1] Carlos Amendola et al. *Third-Order Moment Varieties of Linear Non-Gaussian Graphical Models*. arXiv preprint. 2021. URL: <https://arxiv.org/pdf/2112.10875.pdf>.
- [2] S. Aronhold. “Theorie der homogenen Funktionen dritten Grades von drei Veränderlichen,” in: *Journal für die reine und angewandte Mathematik* 55 (1858), pp. 97–191.
- [3] Alessandra Bernardi, Alessandro Gimigliano, and Monica Idà. “Computing symmetric rank for symmetric tensors”. In: *Journal of Symbolic Computation* 46.1 (Jan. 2011), pp. 34–53.
- [4] Jerome Brachat et al. “Symmetric tensor decomposition”. In: *Linear Algebra and its Applications* 433.11 (2010), pp. 1851–1872.
- [5] Gonzalo Camba-Mendez and George Kapetanios. “Statistical Tests and Estimators of the Rank of a Matrix and Their Applications in Econometric Modelling”. In: *Econometric Reviews* 28 (July 2009), pp. 581–611.
- [6] Qihui Chen and Zheng Fang. “Improved inference on the rank of a matrix”. In: *Quantitative Economics* 10.4 (2019), pp. 1787–1824.
- [7] Qihui Chen and Zheng Fang. “Inference on functionals under first order degeneracy”. In: *Journal of Econometrics* 210 (Mar. 2019).
- [8] G. Darmais. “Analyse générale des liaisons stochastiques: etude particulière de l’analyse factorielle linéaire”. In: *Revue de l’Institut International de Statistique / Review of the International Statistical Institute* 21.1/2 (1953), pp. 2–8.
- [9] Igor Dolgachev and Vassil Kanev. “Polar Covariants of Plane Cubics and Quartics”. In: *Advances in Mathematics* 98 (1993), pp. 216–301.
- [10] B. Efron. “Bootstrap Methods: Another Look at the Jackknife”. In: *The Annals of Statistics* 7.1 (1979), pp. 1–26.
- [11] Zheng Fang and Andres Santos. “Inference on Directionally Differentiable Functions”. In: *Review of Economic Studies* 86 (Apr. 2014).
- [12] Clark Glymour, Kun Zhang, and Peter Spirtes. “Review of Causal Discovery Methods Based on Graphical Models”. In: *Frontiers in Genetics* 10 (2019).
- [13] Frank Kleibergen and Richard Paap. “Generalized reduced rank tests using the singular value decomposition”. In: *Journal of Econometrics* 133.1 (2006), pp. 97–126.
- [14] Arun Kuchibhotla and Abhishek Chakraborty. “Moving Beyond Sub-Gaussianity in High-Dimensional Statistics: Applications in Covariance Estimation and Linear Regression”. In: *Information and Inference A Journal of the IMA* (June 2022).
- [15] J. M. Landsberg. *Tensors: Geometry and Applications*. Vol. 128. Graduate Studies in Mathematics. Providence, Rhode Island: American Mathematical Society, 2012.

- [16] D. Lopez-Paz et al. “Towards a Learning Theory of Cause-Effect Inference”. In: *Proceedings of the 32nd International Conference on Machine Learning*. Vol. 37. JMLR Workshop and Conference Proceedings. JMLR, 2015, pp. 1452–1461.
- [17] Marloes Maathuis et al., eds. *Handbook of graphical models*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, Boca Raton, FL, 2019, pp. xviii+536.
- [18] Jan R. Magnus. “On Differentiating Eigenvalues and Eigenvectors”. In: *Econometric Theory* 1.2 (1985), pp. 179–191.
- [19] Arthur Marmin, Marc Castella, and Jean-Christophe Pesquet. “Detecting the Rank of a Symmetric Tensor”. In: *2019 27th European Signal Processing Conference (EUSIPCO)*. 2019, pp. 1–5.
- [20] Joris M. Mooij et al. “Distinguishing Cause from Effect Using Observational Data: Methods and Benchmarks”. In: *Journal of Machine Learning Research* 17.32 (2016), pp. 1–102.
- [21] Masashi Okamoto. “Distinctness of the Eigenvalues of a Quadratic form in a Multivariate Sample”. In: *The Annals of Statistics* 1.4 (1973), pp. 763–765. (Visited on 10/13/2022).
- [22] Giorgio Ottaviani. “An Invariant Regarding Waring’s Problem for Cubic Polynomials”. In: *Nagoya Mathematical Journal* 193 (2009), pp. 95–110.
- [23] Dimitris N. Politis, Joseph P. Romano, and Michael Wolf. *Subsampling*. Springer Series in Statistics. Springer New York, 1999.
- [24] Majid M. Al-Sadoon. “A unifying theory of tests of rank”. In: *Journal of Econometrics* 199.1 (2017), pp. 49–62.
- [25] Shohei Shimizu. *Statistical Causal Discovery: LiNGAM Approach*. Springer Tokyo, 2022.
- [26] Shohei Shimizu et al. “A linear non-Gaussian acyclic model for causal discovery.” In: *Journal of Machine Learning Research* 7.10 (2006).
- [27] Shohei Shimizu et al. “DirectLiNGAM: A direct method for learning a linear non-Gaussian structural equation model”. In: *The Journal of Machine Learning Research* 12 (2011), pp. 1225–1248.
- [28] V. P. Skitovic. “On a property of the normal distribution”. In: *Doklady Akademii Nauk SSSR* 89 (1953), pp. 217–219.
- [29] Nils Sturma, Mathias Drton, and Dennis Leung. *Testing Many and Possibly Singular Polynomial Constraints*. arXiv preprint. 2022. URL: <https://arxiv.org/abs/2208.11756>.
- [30] Daniele Tramontano, Anthea Monod, and Mathias Drton. “Learning linear non-Gaussian polytree models”. In: *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*. Ed. by James Cussens and Kun Zhang. Vol. 180. Proceedings of Machine Learning Research. PMLR, 2022, pp. 1960–1969.
- [31] Y. Samuel Wang and Mathias Drton. “High-dimensional causal discovery under non-Gaussianity”. In: *Biometrika* 107.1 (2020), pp. 41–59.



- [32] Kun Zhang and Aapo Hyvärinen. “On the Identifiability of the Post-Nonlinear Causal Model”. In: *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. UAI '09. Montreal, Quebec, Canada: AUAI Press, 2009, pp. 647–655.