Technische Universität München
TUM School of Computation, Information and Technology

Identifiability and Statistical Inference in Latent Variable Modeling

Nils B. Sturma

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz:
Prof. Dr. Christoph Knochenhauer

Prüfende der Dissertation:
1. Prof. Mathias Drton, Ph.D.
2. Prof. Niels Richard Hansen, Ph.D.
3. Assoc. Prof. Piotr Zwiernik, Ph.D.

Die Dissertation wurde am 17.06.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 06.09.2024 angenommen.

Abstract

In this publication-based dissertation, we study latent variable models in parametric settings. Since latent variable models are families of marginal distributions, they generally feature a complicated geometry that may lead to identifiability issues and failures of standard inference methods. For example, the models often contain irregular points like algebraic singularities, where well-known methods such as the likelihood ratio test or Wald test are no longer valid. One contribution of this thesis is to develop a testing methodology that is valid even if the underlying model contains irregular points. The other focus of this thesis is the investigation of geometry and identifiability in certain types of linear structural equation models with latent variables. Many popular latent variable models are special types of linear structural equation models. Examples are factor analysis models, linear causal models or models assumed in causal representation learning and independent component analysis. Since structural equation models correspond to directed graphs and feature polynomial or rational parameterizations, they allow for a combinatorial study using algebraic methods. In particular, by adopting this algebraic point of view, we give conditions for structure identifiability of causal representations in a multi-domain setup, we provide a criterion for parameter identifiability in linear causal models, and we investigate the geometry of sparse factor analysis models.

Zusammenfassung

In dieser publikationsbasierten Dissertation untersuchen wir Modelle mit latenten Variablen in parametrischen Settings. Da Modelle mit latenten Variablen Familien von Randverteilungen sind, weisen sie im Allgemeinen eine komplizierte Geometrie auf, die zu Identifizierbarkeitsproblemen und zum Versagen von Standard-Inferenzmethoden führt. Zum Beispiel können die Modelle irreguläre Punkte wie algebraische Singularitäten enthalten, an denen bekannte Methoden wie der Likelihood-Ratio-Test oder der Wald-Test nicht mehr gültig sind. Ein Beitrag dieser Arbeit ist die Entwicklung einer Testmethodik, die auch dann gültig ist, wenn das zugrunde liegende Modell unregelmäßige Punkte enthält. Der andere Schwerpunkt dieser Arbeit ist die Untersuchung der Geometrie und der Identifizierbarkeit in bestimmten Typen von linearen Strukturgleichungsmodellen mit latenten Variablen. Viele gängige Modelle mit latenten Variablen sind spezielle Typen von linearen Strukturgleichungsmodellen. Beispiele hierfür sind Modelle in der Faktorenanalyse, lineare kausale Modelle oder Modelle, die beim kausalen Repräsentationslernen und der Unabhängigkeitsanalyse angenommen werden. Da Strukturgleichungsmodelle durch gerichtete Graphen dargestellt werden können und polynomiale oder rationale Parametrisierungen aufweisen, ermöglichen sie eine kombinatorische Untersuchung mit algebraischen Methoden. Insbesondere geben wir unter dieser algebraischen Sichweise Bedingungen für die Strukturidentifizierbarkeit von kausalen Repräsentationen in einem Multi-domain Setup an, wir beweisen ein Kriterium für die Parameteridentifizierbarkeit in linearen kausalen Modellen, und wir untersuchen die Geometrie von dünnbesetzten Modellen in der Faktorenanalyse.

Acknowledgments

I would like to thank many people for helping me along this journey. First and foremost, I am very grateful to my advisor, Mathias Drton. He has been an excellent mentor and has given me very valuable feedback on papers, presentations and applications. His support and guidance really helped me to present my work in the best possible way. Mathias always encouraged me to go to conferences, workshops or visit other researchers so that I could establish myself in our academic community. I feel privileged to have been given the opportunity to grow under his influence as a statistician, and there is no doubt that working with Mathias has been my best decision during my academic career so far.

I would also like to thank my outstanding collaborators. Especially, I want to thank Caroline Uhler for hosting me for a research visit at MIT and Broad Institute during fall 2022. I really appreciate all our inspiring discussions that enabled me to discover new research directions. I would also like to thank my collaborators Alexandros Grosdos, Chandler Squires, Dennis Leung, Irem Portakal, Jane Ivy Coons, Rina Foygel Barber, Luca Weihs and Yulia Alexandr. It is a great pleasure to work with you and I am constantly learning from you.

There is a great atmosphere in the Mathematical Statistics group at TUM, which allows for interesting and entertaining discussions, for example during the many lunch and coffee breaks. I would like to thank every PhD student and postdoc who was part of the group at some point during my doctorate. All of them were always helpful and willing to share their knowledge, be it on mathematical and statistical topics or on technical issues. I would also like to thank Andrea Grant for her tireless support with administrative matters.

I am thankful for the financial support by the Munich Data Science Institute via the Linde/MDSI PhD Fellowship program, which allowed me to focus entirely on my research. In addition, the Munich Data Science Institute made it easy for me to network with other doctoral students, which resulted in valuable contacts and a very interesting exchange.

I am very grateful for the love and support of my family. I especially want to thank my parents for always supporting me during my years of study and during my PhD. Their confidence and encouragement allowed me to enjoy so many opportunities. Thank you.

Finally, I would like to thank Melanie Hirdes, who has been an amazing partner and friend since we first met. She sometimes reminded me that there are more important issues than PhDs, and she was always there with love and support.

List of Contributed Articles

Core publications as main author:

1. Sturma, N., Leung, D., Drton, M., (2024). Testing many constraints in possibly irregular models using incomplete U-statistics. J. R. Stat. Soc. Ser. B. Stat. Methodol., 86(4):987-1012. https://doi.org/10.1093/jrsssb/qkae022.

2. Sturma, N., Squires, C., Drton, M., and Uhler, C. (2023). Unpaired multi-domain causal representation learning. In Oh, A., Neumann, T., Globerson, A., Saenko, K., Hardt, M., and Levine, S., editors, Advances in Neural Information Processing Systems, volume 36, pages 34465–34492. Curran Associates, Inc. https://papers.nips.cc/paper_files/paper/2023/hash/6c8985579293e0209bdaa4f21bb1d237-Abstract-Conference.html.

Additional publications as co-author:

3. Barber, R. F., Drton, M., Sturma, N., and Weihs, L. (2022). Half-trek criterion for identifiability of latent variable models. Ann. Statist., 50(6):3174–3196. https://doi.org/10.1214/22-AOS2221.

4. Drton, M., Grosdos, A., Portakal, I., and Sturma, N. (2024). Algebraic sparse factor analysis. arXiv preprint arXiv:2312.14762. https://arxiv.org/abs/2312.14762.

I, Nils Sturma, am the main author of articles 1 and 2, and I am a co-author of articles 3 and 4. Details of my individual contributions can be found in Appendices A and B, following the respective article summaries.

Contents

1 Introduction

Investigating a system of random variables and inferring certain properties of the system is a main task of data analysts and researchers. Usually, only finitely many observed data samples are available and statistical models are employed to draw valid conclusions. However, a crucial challenge in real world applications is the ubiquitous presence of latent (i.e., unobserved) variables. For example, in a medical study, certain quantities like blood pressure or body weight are measured and therefore observed, but other variables like genetic factors may remain unobserved. More generally, there are a number of reasons why variables are unobserved. It may just be too expensive to carry out the measurements, or the variables may be theoretically hypothesized to exist but impossible to measure directly. The following example is from Drton (2018) and it is a good introduction to latent variable models and the challenges that can arise. It is the simplest example of the types of models we consider in this thesis.

Example 1.1. Does smoking during pregnancy has an effect on the baby's birth weight? To answer this question, suppose that we record in a study the level of maternal smoking during pregnancy (X_1) and an infant's birth weight (X_2) from a sample of n mothers. Hypothesizing that there is a causal effect from the level of smoking on the infant's birth weight, we aim to quantify this effect. Assuming a linear model

$$\begin{aligned} X_1 &= \lambda_{01} \qquad\qquad + \varepsilon_1, \\ X_2 &= \lambda_{02} + \lambda_{12} X_1 + \varepsilon_2, \end{aligned}$$

where ε_1 and ε_2 noise variables, this means that we are interested in statistical inference about the effect λ_{12}. If ε_1 and ε_2 are independent with mean zero and finite variance, it holds that $\mathrm{Cov}(X_1, X_2) = \lambda_{12}\mathrm{Var}(X_1)$, and inference of λ_{12} may be based on the ratio $\mathrm{Cov}(X_1, X_2)/\mathrm{Var}(X_1)$. For example, we obtain an estimator of λ_{12} by dividing the sample covariance of X_1 and X_2 by the sample variance of X_1. However, the equation $\mathrm{Cov}(X_1, X_2) = \lambda_{12}\mathrm{Var}(X_1)$ only holds true if there is no latent variable that has an effect on both X_1 and X_2. Cigarette companies might therefore argue that it is very likely that such a latent variable exists. As an example, there may be genetic or socio-economic factors that have an influence on the smoking behaviour as well as on the baby's birth weight. Then, inference based on the ratio $\mathrm{Cov}(X_1, X_2)/\mathrm{Var}(X_1)$ is not valid anymore. In fact, it becomes impossible to uniquely recover the effect λ_{12} from the observed data even if the number of participants of our study tends to infinity.

One way out of this dilemma is to add a so called "instrumental variable" that only has a direct effect on the level of smoking but not on the birth weight, and that is not influenced by the latent variable (Bowden and Turkington, 1984). As in Evans and Ringel (1999), we could record the tax rate on cigarettes (X_3) of

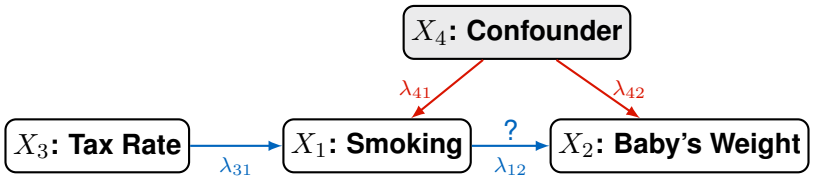


Figure 1.1 Directed graph for the instrumental variable model with latent variable X_4.

the state where each mother is living in as such an instrumental variable. The linear equations defining the model with the latent variable X_4 are then given as

$$\begin{aligned} X_1 &= \lambda_{01} + \lambda_{31}X_3 + \lambda_{41}X_4 + \varepsilon_1, \\ X_2 &= \lambda_{02} + \lambda_{12}X_1 + \lambda_{42}X_4 + \varepsilon_2, \\ X_3 &= \lambda_{03} \qquad\qquad\quad + \varepsilon_3, \\ X_4 &= \lambda_{04} \qquad\qquad\quad + \varepsilon_4, \end{aligned} \tag{1.1}$$

where $\varepsilon_1, \ldots, \varepsilon_4$ are pairwise independent; also see the graphical illustration in Figure 1.1. It follows that the equation $\text{Cov}(X_2, X_3) = \lambda_{12}\text{Cov}(X_1, X_3)$ holds in this model, and as long as $\text{Cov}(X_1, X_3) \neq 0$ we can therefore infer properties of λ_{12} based on the ratio $\text{Cov}(X_2, X_3)/\text{Cov}(X_1, X_3)$. We refer to Didelez et al. (2010) for more background on instrumental variable models.

Example 1.1 shows that the inclusion of latent variables may be necessary to avoid false inference, but also that new challenges may arise in statistical models after the inclusion of latent variables. The example also illustrates a central problem in statistics, namely the identification and quantification of *causal* relationships. In an ideal world, we would set up a randomized controlled experiment to physically intervene into the system and learn about the causal effects. In the context of Example 1.1, to learn about the causal effect from the level of maternal smoking to the infant's birth weight in the presence of latent confounding, this would require us to force pregnant women to smoke. This shows that intervening in a system is often infeasible due to ethical or technical constraints or costs. One of the main topics in causality research is to investigate whether and how causal relationships can be inferred under latent confounding on the basis of observational data only (Pearl, 2009; Peters et al., 2017).

In addition to models that allow for an investigation of causality, latent variable models are either explicitly or implicitly assumed in many other areas of statistics. For example, in cluster analysis it is the goal to group variables into subsets or "clusters", such that those within each cluster are more similar (Hastie et al., 2009). It is hence implicitly assumed that there exist one or multiple latent random variables that determine the membership to a cluster. Moreover, in machine learning it is often the goal to learn a so called "representation" of data, which is then used to tackle downstream tasks. A representation is just another terminology for a set of latent variables that should explain the observed measurements and their variation as well as possible (Bengio et al., 2013). Latent variables are also assumed if we aim to reduce dimensionality by techniques such as principal component analysis or factor analysis. In the latter, correlations among observed random variables are explained using a smaller number of latent variables, also known as factors (Drton et al., 2007).

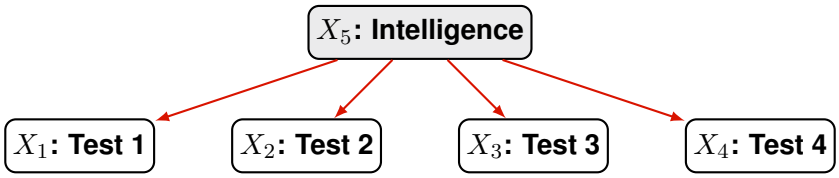


Figure 1.2 Spearman's one-factor model with latent variable X_5.

Example 1.2. Factor analysis first appeared in Spearman (1904). In his work, he studied the correlations of data corresponding to a study where participants took several mental tests. By identifying the results of the mental test i with a random variable X_i and denoting the covariance of X_i and X_j by $\sigma_{ij} = \mathrm{Cov}(X_i, X_j)$, Spearman observed that the fundamental equations, called "tetrads",

$$\sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{jk} = 0$$

hold true for all sets of four different tests i, j, k, l. In other words, all off-diagonal 2×2 minors of the covariance matrix vanish, which means that the off-diagonal entries of the covariance matrix have a rank-one structure. Therefore, assuming a factor analysis model, Spearman argues that there must be a *single* latent variable such as "intelligence" that explains the variability among the tests. See Figure 1.2 for an illustration of the one-factor model.

The topic of this thesis are latent variable models in parametric settings, where the models are defined as families of distribution $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ with finite dimensional parameter space Θ. Latent variable models are often complicated objects, where the parameter space is non-smooth and the models are not identifiable, that is, the map $\theta \mapsto P_\theta$ is not injective. Hence, latent variable models may be irregular and standard inferential methods become invalid. If the parameter space Θ is a semi-algebraic set, i.e. defined by a set of polynomial equality and inequality constraints, it becomes useful to analyze statistical models from an algebraic point of view and many basic statistical questions can be transformed into algebraic problems.

The thesis is comprised of four articles. Three of the articles (including one of the core articles and both additional articles) consider special types of linear structural equation models with latent variables. This includes the previously mentioned factor analysis model, linear causal models such as the instrumental variable model, and linear models in a causal representation learning setup. By taking an algebraic perspective, we investigate parameter and model identifiability and study fundamental questions regarding the model geometry.

The focus of the second core article is on statistical inference. We consider testing the goodness-of-fit of parametric models, where the parameter space is defined by a set of equality and inequalities constraints. We propose a new method that leverages this implicit description and is, in particular, proved to be valid in settings where the description involves a very large amount of constraints and where the model contains irregular points, as it is typically the case in latent variable models.

Organization of the thesis. In Chapter 2, we introduce preliminaries such as the connection to algebraic geometry and the concepts of parameter and model identifiability. We also introduce linear structural

equation models with latent variables as an overarching model class for the models discussed in this thesis. In Chapter 3, we discuss testing null hypotheses defined by equality and inequality constraints and we highlight the contributions of the first core article (Sturma et al., 2024). In Chapter 4, we introduce the problem of causal representation learning from unpaired multi-domain observations and we present our contributions regarding model identifiability given in the second core article (Sturma et al., 2023). Chapter 5 is about the first additional article (Barber et al., 2022), where we study parameter identifiability in latent factor models, a special kind of linear causal models. Finally, we discuss the contributions of the second additional article (Drton et al., 2024) in Chapter 6, in which we investigate the model geometry of sparse factor analysis models.

Notation. Let $\mathbb{N}$ be the set of nonnegative integers. For positive $n \in \mathbb{N}$, we define $[n] = \{1, \dots, n\}$. For a tuple $x = (x_1, \dots, x_b)$ and a set $B \subseteq [b]$, we denote by x_B the tuple only containing the entries indexed by B. Similarly, for a matrix $M \in \mathbb{R}^{a \times b}$, we denote by $M_{A,B}$ the submatrix containing the rows indexed by $A \subseteq [a]$ and the columns indexed by $B \subseteq [b]$. To simplify notation, we write $M_{A,B}^\top$ as a short-hand for $(M_{A,B})^\top$, i.e., we always first take the submatrix and then the transpose of the submatrix. For real numbers $x_1, \dots, x_a$, we denote by $\mathrm{diag}(x_1, \dots, x_a)$ the $a \times a$ diagonal matrix with diagonal entries equal to $x_1, \dots, x_a$. We denote by $N_p(\mu, \Sigma)$ the p-variate normal distribution with mean $\mu \in \mathbb{R}^p$ and positive semidefinite covariance matrix $\Sigma \in \mathbb{R}^p$, and we denote by χ^2_k the central chi-square distribution with k degrees of freedom. Finally, for a sequence of random variables $X_1, X_2, \dots$ we write $X_n \to_d X$ if the sequence converges in distribution to another random variable X for $n \to \infty$.

2 Preliminaries

In this chapter, we formally define latent variable models and lay out the connection to algebraic geometry. We define concepts such as model and parameter identifiability and we introduce linear structural equation models with latent variables.

2.1 Latent Variable Models

A statistical model is given by a triple $(\mathcal{S}, \mathcal{F}, \mathcal{P})$, where $\mathcal{S}$ is a sample space, $\mathcal{F}$ is a suitable σ-algebra and $\mathcal{P}$ is a family of probability measures on $(\mathcal{S}, \mathcal{F})$. In many cases, the sample space $\mathcal{S}$ is equal to $\mathbb{R}^p$, the σ-algebra $\mathcal{F}$ is the Borel σ-algebra, and each element $P \in \mathcal{P}$ is a joint probability distribution of p real-valued random variables. We are then defining a statistical model by only stating the family of distributions $\mathcal{P}$ and assuming the form of $\mathcal{S}$ and $\mathcal{F}$ implicitly, cf. Casella and Berger (1990).

Latent variable models $\mathcal{P}$ are now obtained if only subsets of these p variables are observed, that is, the models are obtained as families of marginal distributions of other models $\mathcal{P}'$. Said differently, latent variable models correspond to projections to lower dimensional subspaces. For some subset of indices $\mathcal{O} \subseteq \{1, \dots, p\}$ that corresponds to the observed variables, define

$$\begin{aligned} \pi_{\mathcal{O}} : \mathbb{R}^p &\longrightarrow \mathbb{R}^{|\mathcal{O}|}, \\ x &\longmapsto (x_i)_{i \in \mathcal{O}} \end{aligned}$$

to be the projection map onto the observed sample space. Moreover, denote by $\pi_{\mathcal{O}} \# P$ the push-forward measure under the projection $\pi_{\mathcal{O}}$, that is, $\pi_{\mathcal{O}} \# P(A) = P(\pi_{\mathcal{O}}^{-1}(A))$ for all Borel measurable sets $A \subseteq \mathbb{R}^{|\mathcal{O}|}$. Then, a latent variable model is given by the family of *observed distributions* $\mathcal{P} = \{\pi_{\mathcal{O}} \# P : P \in \mathcal{P}'\}$.

Example 2.1. Consider again the instrumental variable model from Example 1.1. Setting the mean vector $(\lambda_{01}, \dots, \lambda_{04})^\top$ to zero and rewriting equation system (1.1) in vector form, we have

$$\underbrace{\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix}}_{=: X} = \underbrace{\begin{pmatrix} 0 & \lambda_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 \\ \lambda_{41} & \lambda_{42} & 0 & 0 \end{pmatrix}^\top}_{=: \Lambda} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}}_{=: \varepsilon}.$$

Solving for X yields $X = (I - \Lambda)^{-\top} \varepsilon$. If the noise vector ε is normally distributed with $\varepsilon \sim N(0, \Omega)$, where Ω is a diagonal matrix, then X is also normally distributed with mean zero and covariance matrix

$$\mathrm{Cov}(X) = (I - \Lambda)^{-\top} \Omega (I - \Lambda)^{-1}. \tag{2.1}$$

Now, the fully observed statistical model is the family of all multivariate normal distributions $N(0, \Sigma)$, where the covariance matrix Σ is of the form (2.1). The instrumental variable model is thus the family of marginal distributions over the observed random vector $X_{\mathcal{O}} = (X_i)_{i \in \mathcal{O}}$, where $\mathcal{O} = \{1, 2, 3\}$. It consists of all normal distributions with mean zero and covariance matrix given by the projection

$$\begin{aligned} \mathsf{Cov}(X_{\mathcal{O}}) = \mathsf{Cov}(X)_{\mathcal{O},\mathcal{O}} &= [(I - \Lambda)^{-\top} \Omega (I - \Lambda)^{-1}]_{\mathcal{O},\mathcal{O}} \\ &= \begin{pmatrix} \sigma_{11} & \omega_{44}\lambda_{41}\lambda_{42} + \lambda_{12}\sigma_{11} & \omega_{33}\lambda_{31} \\ \cdot & \omega_{22} + \omega_{44}\lambda_{42}^2 + 2\omega_{44}\lambda_{41}\lambda_{42}\lambda_{12} + \lambda_{12}^2\sigma_{11} & \omega_{33}\lambda_{31}\lambda_{12} \\ \cdot & \cdot & \omega_{33} \end{pmatrix}, \end{aligned} \tag{2.2}$$

where $\omega_{11}, \ldots, \omega_{44}$ are the diagonal entries of the matrix Ω and $\sigma_{11} = \mathsf{Var}(X_1) = \omega_{11} + \omega_{44}\lambda_{41}^2 + \omega_{33}\lambda_{31}^2$ is the variance of the variable X_1.

2.2 Algebraic Approach to Parametric Models

Often, the distributions in a model are indexed by some parameter, so that $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ for a parameter space $\Theta \in \mathbb{R}^k$; compare e.g. (Lehmann and Casella, 1998, Chapter 1). More generally, *parametric models* are defined as the image of a map from a finite dimensional parameter space $\Theta \subseteq \mathbb{R}^k$ to the space of probability distributions, i.e.,

$$\begin{aligned} \Theta &\longrightarrow \Delta(\mathbb{R}^p), \\ \theta &\longmapsto P_\theta, \end{aligned}$$

where $\Delta(\mathbb{R}^p)$ is the space of probability distributions on $\mathbb{R}^p$. As we have seen in Example 2.1, the instrumental variable model is a parametric latent variable model with $k = 8$ parameters given by $\lambda_{12}, \lambda_{31}, \gamma_1, \gamma_2$ and $\omega_{11}, \ldots, \omega_{44}$.

In many cases, the parametrization map is polynomial or rational. It is then possible to apply techniques from algebraic geometry to shed light on statistical problems that would otherwise be difficult to address. In order to define this well and to simplify the presentation, we assume in this section that the model $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ can be identified with a subset $\mathcal{M} \subseteq \mathbb{R}^d$. That is, every probability distribution $P \in \mathcal{P}$ is assumed to be in one-to-one correspondence with a point $p \in \mathcal{M}$. When we talk about a *model*, we hence refer to a set $\mathcal{M} = \mathsf{Im}(\tau)$ defined as the image of the polynomial or rational map

$$\tau : \Theta \longrightarrow \mathbb{R}^d, \tag{2.3}$$

that is defined everywhere on Θ. We have already seen in Example 2.1 that instrumental variable models can be identified with subsets of covariance matrices $\mathcal{M} \subseteq \mathbb{R}^{3\times 3}$. More generally, any Gaussian model with mean zero can be identified with a set of covariance matrices. In the following example, we denote by $\mathit{PD}(m) \subseteq \mathbb{R}^{m\times m}$ the cone of positive definite symmetric $m \times m$ matrices and by $\mathsf{diag}_m^+ \subseteq \mathit{PD}(m)$ the subset of diagonal matrices with positive diagonal entries.

Example 2.2. Recall the graph in Figure 1.2 that corresponds to a one-factor analysis model with four observed variables. In general, we may have m observed variables indexed by $\mathcal{O} = \{1, \dots, m\}$ and one latent variable indexed by $m+1$. Factor analysis then assumes that the variables are linearly related as

$$\begin{aligned} X_{\mathcal{O}} &= \Lambda_{m+1,\mathcal{O}}^{\top} X_{m+1} + \varepsilon_{\mathcal{O}}, \\ X_{m+1} &= \varepsilon_{m+1}, \end{aligned}$$

where $\varepsilon = (\varepsilon_1, \dots, \varepsilon_{m+1})^{\top}$ is the vector of noise variables and $\Lambda_{m+1,\mathcal{O}}^{\top} = (\lambda_{m+1,1}, \dots, \lambda_{m+1,m})^{\top}$ is the vector of parameters. As in Example 2.1, we assume that $\varepsilon = (\varepsilon_1, \dots, \varepsilon_{m+1})^{\top}$ is normally distributed with $\varepsilon \sim N(0, \Omega)$ and $\Omega = \mathsf{diag}(\omega_{11}, \dots, \omega_{m+1,m+1})$ diagonal. Hence, $X_{\mathcal{O}}$ is also normally distributed and the one-factor analysis model can be identified with the image of the polynomial map

$$\begin{aligned} \tau_{1F} : \mathbb{R}^m \times \mathsf{diag}_{m+1}^{+} &\longrightarrow \mathit{PD}(m), \\ (\Lambda_{m+1,\mathcal{O}}, \Omega) &\longmapsto \omega_{m+1,m+1} \Lambda_{m+1,\mathcal{O}}^{\top} \Lambda_{m+1,\mathcal{O}} + \Omega_{\mathcal{O},\mathcal{O}}. \end{aligned}$$

Besides Gaussian models corresponding to sets of covariance matrices, one might also study models defined as sets of higher-order moment tensors corresponding to non-Gaussian distributions, see e.g. Ardiyansyah and Sodomaco (2023), Améndola et al. (2023) or Robeva and Seby (2021). Another large class of models that can be identified with subsets $\mathcal{M} \subseteq \mathbb{R}^d$ are discrete models with d possible outcomes. In this case, a probability is assigned to each outcome and models are subsets of the $d-1$ dimensional probability simplex $\Delta_d = \{x \in \mathbb{R}^d : x_i \geq 0 \text{ for all } i \text{ and } x_1 + \cdots + x_d = 1\}$. See Sullivant (2018), Zwiernik (2016) and Drton et al. (2009) for background on discrete statistical models.

It is most useful to analyze statistical models from an algebraic point of view if, in addition to the parametrization being polynomial or rational, the parameter space Θ is a semialgebraic set defined by polynomial equality and inequality constraints.

Definition 2.3. Let $\mathbb{R}[x_1, \dots, x_k]$ be the ring of polynomials in the indeterminates $x_1, \dots x_k$ with real coefficients. A *basic semi-algebraic set* is a a subset of $\mathbb{R}^k$ that is of the form

$$\Theta = \{\theta \in \mathbb{R}^k : f(\theta) = 0 \text{ for all } f \in F,\ h(\theta) < 0 \text{ for all } h \in H\},$$

where $F, H \subseteq \mathbb{R}[x_1, \dots, x_k]$ are finite (possibly empty) collections of polynomials. A *semi-algebraic set* is a finite union of basic semi-algebraic sets.

We refer to Bochnak et al. (1998), Basu et al. (2006) and Benedetti and Risler (1990) for background on semi-algebraic sets. Importantly, by the Tarski-Seidenberg theorem (Benedetti and Risler, 1990, Theorem 2.3.4), images of semi-algebraic sets are again semi-algebraic under polynomial and rational mappings. That is, given a semi-algebraic set $\Theta \subseteq \mathbb{R}^k$ and a polynomial or rational map $\tau : \mathbb{R}^k \to \mathbb{R}^d$ defined everywhere on Θ, the image $\tau(\Theta)$ is again a semialgebraic set. Many statistical models are images of semi-algebraic sets; see e.g. Sullivant (2018) and Drton et al. (2009).

Example 2.4. Consider again the instrumental variable model that is illustrated in Figure 1.1. As we have seen in Example 2.1, the model is identified with the image of the polynomial map

$$\tau_{IV} : \mathbb{R}^4 \times \mathsf{diag}_4^{+} \longrightarrow \mathit{PD}(3),$$

that maps the parameter vector $(\lambda_{12}, \lambda_{31}, \gamma_1, \gamma_2)^\top$ and the diagonal matrix $\Omega = \text{diag}(\omega_{11}, \ldots, \omega_{44})$ to the covariance matrix Σ as in (2.2). Note that the domain $\Gamma = \mathbb{R}^4 \times \text{diag}_4^+$ is semi-algebraic since diag_4^+ is isomorphic to the set $\{\omega \in \mathbb{R}^4 : \omega_i > 0 \text{ for all } i = 1, \ldots, 4\}$ and cartesian products of semi-algebraic sets are semi-algebraic. Hence, the image $\text{Im}(\tau_{IV})$ is also a semi-algebraic set since τ_{IV} is polynomial. Similarly, the one-factor analysis model given as the image of the map τ_{1F} defined in Example 2.2 is also a semi-algebraic set.

Assuming that Θ is semi-algebraic and that τ is polynomial or rational, fundamental statistical questions can be translated into algebraic problems. For different model classes that involve latent variables, this thesis is concerned with the following topics:

(i) Model Geometry. The geometry of semi-algebraic sets can be studied using methods from algebraic geometry. For statistical applications it is of interest to provide an *implicit description* of the model. That is, we are interested in the polynomial equalities and inequalities that define the semi-algebraic model $\mathcal{M}$. This illustrated in the next example.

Example 2.5. We have claimed in Example 1.2 that the one-factor model is within the zero locus of all tetrads, i.e., the tetrads are part of the implicit description. Indeed, recalling the parametrization map in Example 2.2, we note that for any covariance matrix $\Sigma \in \text{Im}(\tau_{1F})$ the off-diagonal entries σ_{ij} for $i \neq j$ are given by $\lambda_{m+1,i}\lambda_{m+1,j}$. Hence, we have for four pairwise different indices i, j, k, l the identity

$$\sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{jk} = \lambda_{m+1,i}\lambda_{m+1,j}\lambda_{m+1,k}\lambda_{m+1,l} - \lambda_{m+1,i}\lambda_{m+1,l}\lambda_{m+1,j}\lambda_{m+1,k} = 0.$$

However, as shown in Bekker and de Leeuw (1987, Theorem 1), the full implicit description of the one-factor analysis model not only involves equality constraints given by tetrads but also inequality constraints given by different polynomials.

Implicit descriptions may be used for testing the goodness-of-fit of a model. For example, given an estimator $\widehat{p}$ of a point $p \in \mathbb{R}^d$, we can check whether or not $p \in \mathcal{M}$ by plugging $\widehat{p}$ into the polynomial equalities and inequalities; we refer to Chapter 3 for a detailed discussion.

Another geometric feature that is interesting if one is interested in model selection is the *dimension* of the model. As described, for example, in Drton (2009, Section 3.1), any semi-algebraic set can be written as a disjoint union of finitely many smooth manifolds, and the dimension of the semi-algebraic set can then be defined as the largest dimension of any of these manifolds. Knowing the dimension is crucial, for example, if we want to test the goodness-of-fit of a model using the well known likelihood ratio test. The test statistic is approximately chi-square distributed under certain regularity conditions, and the degrees of freedom depend on the dimension of the model (Drton, 2009). Thus, to calibrate the test correctly and to get asymptotically valid critical values, we need to know the dimension.

(ii) Parameter Identifiability. A parametric model is identifiable if for each $p \in \mathcal{M}$ we can uniquely recover the parameter $\theta \in \Theta$ that was mapped to p. This is a fundamental question if the parameters have a special meaning and one is interested in statistical inference of the parameters. The following definition

can be found, for example, in Lehmann and Casella (1998, Section 1.5) or in Sullivant (2018, Chapter 16), but it is tailored to our setup.

Definition 2.6. Let $\tau : \Theta \to \mathbb{R}^d$ be as in (2.3), with model $\mathcal{M} = \text{Im}(\tau)$. Then $\mathcal{M}$ is said to be *parameter identifiable* if the map τ is injective on Θ.

In practice, the true probability distribution, i.e. the true point $p \in \mathcal{M}$, is never known, we only get to see n samples from this distribution. However, knowing the true underlying probability distribution can be seen as observing an infinite number of samples. Hence, if parameters are identifiable, it is theoretical possible to learn the correct parameters in the infinite data limit.

Example 2.7. Parameter identifiability in the instrumental variable model holds if, given a covariance matrix $\Sigma \in \text{Im}(\tau_{IV})$, we can uniquely recover the parameters that are mapped to Σ. That is, we study whether we can solve the equations

$$\begin{aligned}\Sigma &= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \cdot & \sigma_{22} & \sigma_{23} \\ \cdot & \cdot & \sigma_{33} \end{pmatrix} \\ &= \begin{pmatrix} \omega_{11} + \omega_{44}\lambda_{41}^2 + \omega_{33}\lambda_{31}^2 & \omega_{44}\lambda_{41}\lambda_{42} + \lambda_{12}\sigma_{11} & \omega_{33}\lambda_{31} \\ \cdot & \omega_{22} + \omega_{44}\lambda_{42}^2 + 2\omega_{44}\lambda_{41}\lambda_{42}\lambda_{12} + \lambda_{12}^2\sigma_{11} & \omega_{33}\lambda_{31}\lambda_{12} \\ \cdot & \cdot & \omega_{33} \end{pmatrix}\end{aligned}$$

uniquely for the parameters $(\lambda_{12}, \lambda_{31}, \gamma_1, \gamma_2)^\top$ and $\omega_{11}, \ldots, \omega_{44}$.

Requiring injectivity on the whole domain Θ is usually not necessary and there exists a variety of distinct definitions for weaker forms of identifiability; see Sullivant (2018, Definition 16.1.1). For example, *generic identifiability* refers to injectivity on the complement $\Theta \setminus A$ of a set A that has Lebesgue measure zero; we refer to Chapter 5 for a precise definition.

(iii) Model Identifiability. Often, we are not only interested in one specific model but we consider a family of models, and it becomes of interest to select the correct parsimonious model, or rather, a model most appropriate for a downstream task. A natural fundamental question that has to be solved beforehand is whether it is possible to uniquely recover the true model if we are given the true underlying probability distribution. The following definition is a modified version of Drton et al. (2023, Definition 2.3).

Definition 2.8. Let $\{\mathcal{M}_i\}_{i=1}^p$ be a family of models in $\mathbb{R}^d$. Then the family is said to be *model identifiable* if $\mathcal{M}_i \cap \mathcal{M}_j = \emptyset$ for all pairs of distinct indices $i, j \in \{1, \ldots, p\}$.

If a family is model identifiable and we are given a point $p \in \mathbb{R}^d$ that is assumed to lie in one of the models, it is hence possible to uniquely recover the true model. Model identifiability is often also called *model distinguishability*. Similarly, as with parameter identifiability, in practice we never get to now $p \in \mathbb{R}^d$ exactly, but we approximate it as the number of observed samples n tends to infinity. Moreover, there also exist weaker forms of model identifiability where the intersection $\mathcal{M}_i \cap \mathcal{M}_j$ is, for example, allowed to be

a lower dimensional subset; see Drton et al. (2023).

In this thesis, we will study fundamental questions regarding the model geometry as well as parameter and model identifiability for different types of latent variable models. The models we consider are all special types of linear structural equation models that are introduced in full generality in the next section. However, we are not only interested in solving fundamental theoretical questions, but also in leveraging the semi-algebraic structure of models to derive new statistical methods. In particular, we study testing the goodness-of-fit of models that feature an implicit description; see Chapter 3.

2.3 Linear Structural Equation Models with Latent Variables

In this section, we introduce linear structural equation models with latent variables in a very general way. The presentation is an expanded and broader version of the introduction in the article by Barber et al. (2022) that contributes to this thesis. Many models considered in statistics can be seen as a special type of linear equation models with latent variables, examples are factor analysis models (Drton et al., 2007; Anderson and Rubin, 1956), linear causal models (Spirtes et al., 2000; Pearl, 2009; Peters et al., 2017) or linear independent component analysis models (Comon and Jutten, 2010; Hyvärinen, 2013; Eriksson and Koivunen, 2004); we discuss them in Examples 2.9-2.11. Due to their tractability and easy interpretation, linear structural equation models are popular in many different applied sciences; we refer to Bollen (1989) for background.

Let $X = (X_v)_{v \in V}$ be a collection of random variables indexed by a finite set V. In our context, we simply let $V = [p]$ and we assume that the set V is a disjoint union $V = \mathcal{O} \sqcup \mathcal{L}$ of a set $\mathcal{O} \subseteq [p]$ indexing the observed variables and a set $\mathcal{L} \subseteq [p]$ indexing the latent variables. All variables are linearly related via

$$X_v = \lambda_{0v} + \sum_{w \neq v} \lambda_{wv} X_w + \varepsilon_v, \qquad v \in V,$$

where λ_{wv} are real-valued parameters. The ε_v are independent random variables that model noise. Viewing $X = (X_v)_{v \in V}$ and $\varepsilon = (\varepsilon_v)_{v \in V}$ as vectors, the above equation system can be represented in the form

$$X = \lambda_0 + \Lambda^\top X + \varepsilon, \tag{2.4}$$

where the parameters are given by a p-dimensional real-valued vector λ_0 and a $p \times p$ matrix $\Lambda = (\lambda_{wv})$ that has zeros on the diagonal. Now, specific models are derived from (2.4) by assuming specific sparsity patterns in Λ. To make this precise, it is useful to adopt a graphical perspective. Consider a directed graph $G = (\mathcal{O} \sqcup \mathcal{L}, D)$ with nodes $V = \mathcal{O} \sqcup \mathcal{L}$ equivalent to the indexing set of the observed and latent variables, and directed edges $D \subseteq V \times V$. We will denote edges $(w, v) \in D$ more intuitively by $w \to v \in D$ and we assume that the edge set is free of self-loops, so $v \to v \notin D$ for all $v \in D$. Every directed graph defines a different set of linear equations by requiring that the parameter λ_{wv} may be nonzero only if the edge $w \to v$ is contained in the edge set D. We denote by $\mathbb{R}^D$ the set of real $p \times p$ matrices $\Lambda = (\lambda_{wv})$

with support D, that is, $\lambda_{wv} = 0$ if $w \to v \notin D$. Then the structural equations corresponding to a graph $G = (\mathcal{O} \sqcup \mathcal{L}, D)$ are given by

$$X = \lambda_0 + \Lambda^\top X + \varepsilon, \qquad \Lambda \in \mathbb{R}^D. \tag{2.5}$$

Many well known latent variable models are now obtained by restricting the type of edges allowed in the graph or by restricting the distributions of the noise ε. The model then consists of all distributions that the observed random vector $X_\mathcal{O}$ can obtain for a given graph and a given family of distributions of ε. In the remainder of this chapter, we assume for simplicity that $\lambda_0 = 0$.

Example 2.9. Consider a graph $G = (\mathcal{O} \sqcup \mathcal{L}, D)$ such that $D = \mathcal{L} \times \mathcal{O}$, that is, the graph exactly contains all possible edges from the latent to the observed variables but no other edge. If we further restrict the distributions of the noise variables ε to be Gaussian with $\varepsilon \sim N(0, \Omega)$, where Ω is a positive definite diagonal matrix, then the model is called a factor analysis model and Equation (2.5) reduces to $X_\mathcal{O} = \Lambda_{\mathcal{L},\mathcal{O}}^\top X_\mathcal{L} + \varepsilon_\mathcal{O}$, where $X_\mathcal{L} = \varepsilon_\mathcal{L}$. We have already seen the example of an *one-factor* analysis model with $|\mathcal{L}| = 1$, recall Figure 1.2. If the edge set D does not contain all possible edges $\mathcal{L} \times \mathcal{O}$ but only a subset, then we obtain a *sparse* factor analysis model; see Chapter 6.

Example 2.10. If we have the same graphical structure as in sparse factor analysis, i.e., $G = (\mathcal{O} \sqcup \mathcal{L}, D)$ such that $D \subseteq \mathcal{L} \times \mathcal{O}$, but we allow the distributions of the latent noise $(\varepsilon_v)_{v \in \mathcal{L}}$ to be arbitrary, then the setup is known as linear independent component analysis (Comon and Jutten, 2010; Hyvärinen, 2013; Eriksson and Koivunen, 2004). The observed and latent random variables also satisfy the equations $X_\mathcal{O} = \Lambda_{\mathcal{L},\mathcal{O}}^\top X_\mathcal{L} + \varepsilon_\mathcal{O}$ and $X_\mathcal{L} = \varepsilon_\mathcal{L}$. The case of non-Gaussian distributions of the variables $(\varepsilon_v)_{v \in \mathcal{O} \sqcup \mathcal{L}}$ is interesting because it is then possible to identify the parameter matrix $\Lambda_{\mathcal{L},\mathcal{O}}^\top$ up to singed permutations of the columns when the distribution of $X_\mathcal{O}$ is given. If the observed random vector does not feature noise, i.e. $X_\mathcal{O} = \Lambda_{\mathcal{L},\mathcal{O}}^\top X_\mathcal{L}$, a proof and precise conditions for identifiability are given by Comon (1994). The identifiability result follows from the Darmois-Skitovich theorem (Darmois, 1953; Skitovich, 1953), but it can also be shown in a more algebraic way by considering higher-order moments and applying uniqueness results from tensor decomposition; see (Comon and Jutten, 2010, Section 5). If the noise of the observed vector is nonzero and $X_\mathcal{O} = \Lambda_{\mathcal{L},\mathcal{O}}^\top X_\mathcal{L} + \varepsilon_\mathcal{O}$, the problem of identifying $\Lambda_{\mathcal{L},\mathcal{O}}^\top$ is an instance of *overcomplete* independent component analysis; see Wang and Seigal (2024) and Eriksson and Koivunen (2004).

Example 2.11. Linear structural equation models are equipped with an intuitive causal interpretation. The parameters λ_{wv} are also known as *direct causal effects* of X_w on X_v. An interesting problem is the question of causal effect identification, i.e., whether or not the causal effects λ_{wv} are identifiable. In this context, linear structural equation models are also called *linear causal models* (Pearl, 2009). In our work in Barber et al. (2022) that contributes to this thesis, we study identifiability in linear causal models assuming that latent variables are independent and external. The graph $G = (\mathcal{O} \sqcup \mathcal{L}, D)$ then only has edges between observed variables and from latent variables to observed variables, i.e., $D \subseteq (\mathcal{O} \times \mathcal{O}) \cup (\mathcal{L} \times \mathcal{O})$, and Equation (2.5) reduces to $X_\mathcal{O} = \Lambda_{\mathcal{O},\mathcal{O}}^\top X_\mathcal{O} + \Lambda_{\mathcal{L},\mathcal{O}}^\top X_\mathcal{L} + \varepsilon_\mathcal{O}$, where $X_\mathcal{L} = \varepsilon_\mathcal{L}$. We call this model a *latent-factor model*; see Chapter 5.

If the noise vector ε in a linear structural equation model is restricted to be Gaussian, we obtain *covariance models*. For the following definition, we only assume that each of the independent noise variables ε_v

has finite variance $\omega_{vv} > 0$. Moreover, recall that diag_p^+ is defined as the set of $p \times p$ diagonal matrices with positive diagonal entries. We also write $\mathbb{R}^D_{\mathsf{reg}} \subseteq \mathbb{R}^D$ for the subset of matrices $\Lambda \in \mathbb{R}^D$ with $I - \Lambda$ invertible.

Definition 2.12. Let $G = (V, D)$ be a directed graph with $V = [p] = \mathcal{O} \sqcup \mathcal{L}$. The *covariance model* given by G is the image $\mathsf{Im}(\tau_G)$ of the parametrization map

$$\begin{aligned} \tau_G : \mathbb{R}^D_{\mathsf{reg}} \times \mathsf{diag}_p^+ &\longrightarrow \mathit{PD}(|\mathcal{O}|), \\ (\Lambda, \Omega) &\longmapsto [(I - \Lambda)^{-\top} \Omega (I - \Lambda)^{-1}]_{\mathcal{O},\mathcal{O}}. \end{aligned}$$

Since the map τ_G is in the worst case rational and the set $\mathbb{R}^D_{\mathsf{reg}} \times \mathsf{diag}_p^+$ is semi-algebraic, we can study covariance models using algebraic tools, recall Section 2.2. If all variables are observed, i.e., $\mathcal{L} = \emptyset$, and the graph is acyclic, then many fundamental problems are essentially solved. The parametrization of φ_G is injective on the whole domain $\mathbb{R}^D_{\mathsf{reg}} \times \mathsf{diag}_p^+$, and the image is a smooth manifold that admits an implicit description in terms of conditional independence statements that can be read-off from the graph. Also, it can be checked efficiently whether two implicit descriptions are the same, that is, it can be checked efficiently whether models can be distinguished. These results are standard in graphical modeling; see for example Lauritzen (1996) and Maathuis et al. (2019).

On the other hand, if $\mathcal{L}$ is nonempty, then we only observe the subvector $X_{\mathcal{O}}$ and the covariance model becomes a set of projections of "large" covariance matrices. All of the above properties might no longer be true and many questions regarding model geometry and identifiability are still open. In particular, the parametrization map needs not to be injective on the whole domain and the images are in general not smooth.

3 Goodness-of-fit Tests in Irregular Models

In this chapter, we discuss challenges that arise if we are interested in testing the goodness-of-fit of a parametric model and how the first core article (Sturma et al., 2024) contributes towards solving these challenges. Consider a parametric family $\{P_\theta : \theta \in \Theta\}$ with parameter space $\Theta \subseteq \mathbb{R}^k$. Given i.i.d. samples $X_1, \dots, X_n$ from an unknown distribution P_θ with parameter $\theta \in \Theta$, we are interested in how well a model with parameter space $\Theta_0 \subseteq \Theta$ fits to the observed samples. That is, we are interested in testing

$$H_0 : \theta \in \Theta_0 \quad \text{vs.} \quad H_1 : \theta \in \Theta \setminus \Theta_0. \tag{3.1}$$

In our work in Sturma et al. (2024), we consider the situation where the null hypothesis Θ_0 is defined by constraints, that is,

$$\Theta_0 = \{\theta \in \Theta : \ f_j(\theta) \leq 0 \text{ for all } j = 1, \dots, p\},$$

where each constraint f_j is a function $f_j : \mathbb{R}^k \to \mathbb{R}$. Even though the constraints considered can be arbitrary functions, our main motivations comes from models where the parameter space Θ is a semi-algebraic set and hence the functions f_j are polynomials. Note that the description of the null hypothesis also allows for equality constraints since $f_j(\theta) = 0$ can be equivalently described by $f_j(\theta) \leq 0$ and $-f_j(\theta) \leq 0$.

Example 3.1. Consider the example of testing tetrad constraints that is of particular relevance in factor analysis (Bollen and Ting, 2000; Spirtes et al., 2000; Hipp and Bollen, 2003; Drton and Xiao, 2016; Leung et al., 2016). As we have seen in Example 2.5, tetrads are off-diagonal 2×2 sub-determinants of a symmetric matrix $\Sigma = (\sigma_{uv})$, an example is $f_j(\Sigma) = \sigma_{uv}\sigma_{wz} - \sigma_{uz}\sigma_{vw}$ with four different indices u, v, w, z. If a multivariate normal distribution $N(0, \Sigma)$ follows a one-factor analysis model, then all tetrads that can be formed from the covariance matrix Σ vanish. Thus, for given i.i.d. samples $X_1, \dots, X_n \sim N(0, \Sigma)$, one might be interested in testing whether all tetrads vanish simultaneously.

Apart from testing tetrads in the one-factor model, polynomial hypothesis are ubiquitous for example in graphical modeling (Sullivant et al., 2010; Chen et al., 2014; Shiers et al., 2016; Chen et al., 2017), testing of causal effects (Spirtes et al., 2000; Steyer, 2005; Pearl, 2009; Strieder et al., 2021), and in constraint-based causal discovery algorithms (Pearl and Verma, 1995; Spirtes et al., 2000; Claassen and Heskes, 2012).

A standard approach for testing hypotheses like in (3.1) is the likelihood ratio test that measures the difference between the maximum of the likelihood function restricted to the subspace Θ_0 compared to the one of the unconstrained space Θ. If P_θ has density p_θ, the test statistic is

$$\lambda_n = -2 \log \left(\frac{\sup_{\theta \in \Theta_0} \mathcal{L}_n(\theta)}{\sup_{\theta \in \Theta} \mathcal{L}_n(\theta)} \right),$$

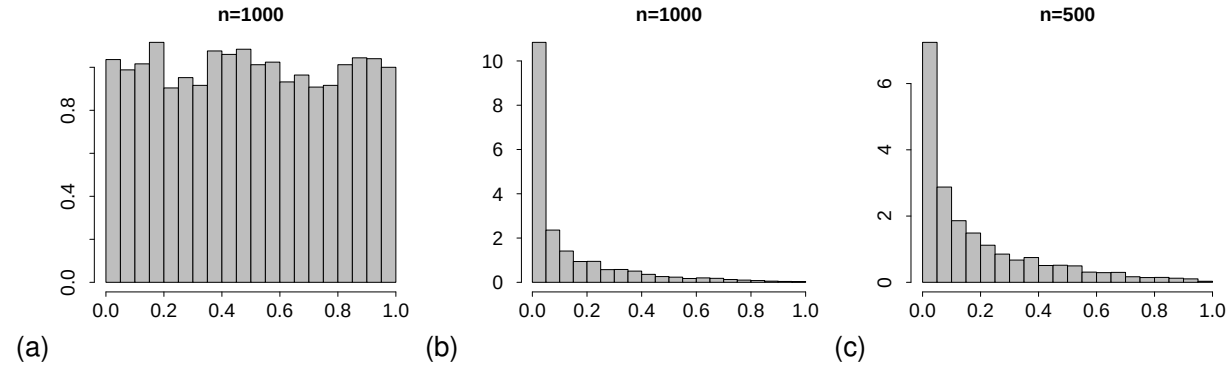


Figure 3.1 Histograms of 5000 simulated p-values for testing the one-factor model using the likelihood ratio test. Setups (a) and (b) are with 15 observed variables, where the true covariance matrix is a regular point in (a) and close to a singular point in (b). Setup (c) is regular but high-dimensional with $k = 200$ observed variables. The exact parameter values are the same as for the simulations in Sturma et al. (2024, Section 5.2). Note that we used the likelihood ratio test with a Bartlett correction for better asymptotic approximation.

where $\mathcal{L}_n(\theta) = \prod_{i=1}^n p_\theta(X_i)$ is the likelihood function (Van der Vaart, 1998, Chapter 16). Let $\theta \in \Theta_0$ be a "true" parameter point in the null hypothesis, that is, $X_1, \dots, X_n \sim P_\theta$. If $\Theta \subseteq \mathbb{R}^k$ is an open set and certain regularity conditions are satisfies, the distribution of λ_n converges to the distribution of the Mahalanobis distance between a random draw from a multivariate normal distribution and the tangent cone of the set Θ_0 at the point θ; see Drton (2009, Theorem 2.6) and Van der Vaart (1998, Theorem 16.7). If the tangent cone at θ is an ℓ-dimensional linear space, the limiting distribution becomes a chi-square distributions with degrees of freedom equal to $k - \ell$. However, the set Θ_0 may contain algebraic singularities at which the rank of the Jacobian of the constraints being tested drops. Then, the tangent cone may have a different dimension or it is not a linear space. Calibrating the likelihood ratio test based on the $\chi^2_{k-\ell}$ distribution is thus not valid anymore. We refer to Drton (2009) for a proof and precise definitions of the *tangent cone* and *singularities*. Note that convergence of the likelihood ratio statistic to the chi-square distribution can already be very slow even if the true parameter point θ is only close to a singularity.

A second challenging scenario for the application of the likelihood ratio test is when the dimension of the parameter space is comparable to the number of samples n, as high dimensionality also has an impact on the speed of convergence. Both issues are illustrated in Figure 3.1. If the likelihood ratio statistic is close to a $\chi^2_{k-\ell}$ distribution, we expect uniformly distributed p-values. We see that the simulated p-values are far from being uniformly distributed in the singular and high-dimensional setups.

Another frequently used approach to test hypotheses like in (3.1) are Wald-type test that make use of the implicit description of the null hypothesis Θ_0 (Lehmann and Romano, 2022, Chapter 14.4). To simplify the discussion, consider the case where the null hypothesis is only defined by one equality constraint $f_1 \in \mathbb{R}[x_1, \dots, x_k]$. Denoting by ∇f_1 the gradient of f_1, the Wald-type test statistic is given as

$$T_{f_1} = \frac{n f_1(\widehat{\theta})^2}{(\nabla f_1(\widehat{\theta}))^\top V(\widehat{\theta}) \nabla f_1(\widehat{\theta})},$$

where $\widehat{\theta}$ is an asymptotically normal estimator such that $\sqrt{n}(\widehat{\theta} - \theta)$ converges in distribution to $N(0, V(\theta))$, and $V(\widehat{\theta})$ is an consistent estimator of $V(\theta)$. If the true parameter θ is regular, T_{f_1} converges to the χ^2_1 distribution (Drton and Weihs, 2016). However, Drton and Xiao (2016) and Dufour et al. (2024) showed

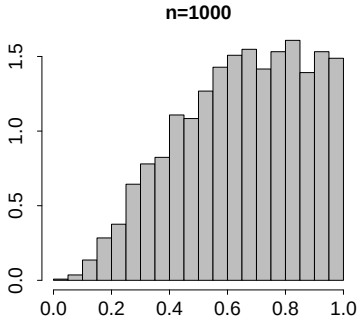


Figure 3.2 Histogram of 5000 simulated p-values for testing a single tetrad constraint using the Wald test. The true covariance matrix $\Sigma = (\sigma_{vw}) \in \mathbb{R}^{4\times 4}$ is close to a singular point, where the covariances σ_{vw} are equivalently generated as the covariances of the "leaves" $L \setminus \{1, 2\}$ in setup (b) in Sturma et al. (2024, Section 5.2).

that the limiting distribution is not chi-square at singularities, where the gradient $\nabla f_1(\theta)$ becomes zero; see Figure 3.2 for an illustration. The Wald statistic can also be defined for multiple constraints but the number of constraints has to be smaller than the dimension of the ambient space k. Hence, the Wald test suffers from similar problems as the likelihood ratio test: It may have a different asymptotic behaviour at singularities and testing many constraints is infeasible.

In our article, we propose a testing strategy that aims to circumvent these difficulties. In particular, the asymptotic approximation remains valid even if the (unknown) parameter θ is singular. One of our key observation is that Wald statistics of polynomial constraints are related to U-statistics. We are now explaining our intuition in the case of testing a single tetrad constraint in a Gaussian setup. In our work in Leung and Sturma (2024) that is not part of this thesis, we discuss testing *arbitrary* polynomial constraints in a Gaussian setup, whereas in the core article Sturma et al. (2024) we take a more general perspective that also allows for *non-polynomial* constraints and *non-Gaussian* distributions.

As in Example 3.1, we consider i.i.d. samples $X_1, \dots, X_n$ from a multivariate normal distribution $N(0, \Sigma)$, and a tetrad $f_1(\Sigma) = \sigma_{13}\sigma_{24} - \sigma_{23}\sigma_{14}$. The sample covariance matrix $\widehat{\Sigma} = \frac{1}{n}\sum_{i=1}^n X_i X_i^\top$ is an asymptotically normal estimator of the covariance matrix Σ; cf. Drton and Xiao (2016, Section 4). Now, one may verify with an easy calculation that the scaled plug-in estimator of the tetrad $\widehat{f_1} = \frac{n}{n-1} f_1(\widehat{\Sigma})$ can be written as a U-statistic $\widehat{f_1} = \frac{1}{\binom{n}{2}} \sum_{i<j} h_1(X_i, X_j)$ with kernel

$$h_1(X_i, X_j) = \frac{1}{2}\{(X_{i1}X_{i3}X_{j2}X_{j4} - X_{i2}X_{i3}X_{j1}X_{j4}) + (X_{j1}X_{j3}X_{i2}X_{i4} - X_{j2}X_{j3}X_{i1}X_{i4})\}. \tag{3.2}$$

Classical U-statistics theory (Koroljuk and Borovskich, 1994, Theorem 4.2.1) gives the Gaussian approximation

$$\sqrt{n}(\widehat{f_1} - f_1(\Sigma)) \longrightarrow_d N(0, m^2\sigma_{g_1}^2), \tag{3.3}$$

where m is the degree of the kernel, i.e. $m = 2$ for the kernel in (3.2), and $\sigma_{g_1}^2$ is the variance of the Hájek projection given by

$$g_1(X_i) = \mathbb{E}[h_1(X_i, X_j)|X_i] = \frac{1}{2}\left\{(X_{i1}X_{i3}\sigma_{24} - X_{i2}X_{i3}\sigma_{14}) + (\sigma_{13}X_{i2}X_{i4} - \sigma_{23}X_{i1}X_{i4})\right\}.$$

It is easy to see that $m^2\sigma_{g_1}^2 = (\nabla f_1(\theta))^\top V(\theta) \nabla f_1(\theta)$; we refer to Leung and Sturma (2024, Lemma B.1) for a proof. Together with Slutsky's theorem and the continuous mapping theorem, the Gaussian

approximation in (3.3) thus explains the chi-square approximation of the Wald statistic of a single tetrad. Crucially, the Gaussian approximation requires non-degeneracy of the kernel, that is, the variance $\sigma_{g_1}^2$ of the Hájek projection has to be strictly larger than zero. If $\sigma_{g_1}^2 = 0$, the U-statistic is said to be degenerate and the Gaussian approximation in (3.3) is not correct any more (Van der Vaart, 1998, Section 12.3). In the case of tetrads, Σ is a singular point, i.e. $\nabla f_1(\Sigma) = 0$, if and only if $\sigma_{g_1}^2 = 0$. Said differently, the Gaussian approximation of the U-statistic is invalid at singularities. We emphasize that close-to-singular singular scenarios with $\sigma_{g_1}^2 \approx 0$ are also problematic in practice since the approximation in (3.3) is very slow. This can be seen by the classical Berry-Esseen bound for non-degenerate U-statistics that measures the accuracy of the Gaussian approximation (Chen et al., 2011, Theorem 10.3). Since the bound depends on $\sigma_{g_1}^{-1}$, it blows up for small values of $\sigma_{g_1}^2$ and the approximation becomes inaccurate.

To accommodate singular settings, we observe that summing over less tuples from the set $I_{n,2} = \{(i_1, i_2) : 1 \leq i_1 < i_2 \leq n\}$ than summing over all tuples as in the case of U-statistics may be beneficial. In the extreme case, consider the average of independent sums

$$S_{\lfloor n/2 \rfloor} = \frac{1}{\lfloor n/2 \rfloor} \sum_{i=1}^{\lfloor n/2 \rfloor} h_1(X_{2i-1}, X_{2i}),$$

where h_1 is the same kernel as in (3.2). By the classical central limit theorem (Klenke, 2014, Theorem 15.37), we obtain that $\sqrt{\lfloor n/2 \rfloor}\ S_{\lfloor n/2 \rfloor} \to_d N(0, \sigma_{h_1}^2)$, where $\sigma_{h_1}^2$ is the variance of the kernel h_1. If the covariance matrix Σ is positive definite, it must be that $\sigma_{h_1}^2 > 0$; see Lemma B.1 in Leung and Sturma (2024). Hence, the Gaussian approximation should also be true if Σ is singular. However, despite being well-suited in singular settings, the independent sum $S_{\lfloor n/2 \rfloor}$ pays the price of dividing down the sample size by $m = 2$. For larger degree kernels with $m > 2$ this implies an even larger loss in statistical efficiency. On the other hand, the complete U-statistic is aggressive in terms of power by summing over all the $\binom{n}{2}$ sample indices in $I_{n,2}$, but it is not suitable for singular settings as explained above.

To find a balance between guarding against singularities and statistical efficiency, we propose to use *randomized incomplete U-statistics* instead of the complete U-statistic $\widehat{f}_1 = \frac{1}{\binom{n}{2}} \sum_{i<j} h_1(X_i, X_j)$. That is, for a *computational budget parameter* $N \leq \binom{n}{2}$, we randomly choose on average N indices from $I_{n,2} = \{(i_1, i_2) : 1 \leq i_1 < i_2 \leq n\}$. For i.i.d. Bernoulli random variables $\{Z_\iota : \iota \in I_{n,2}\}$ with success probability $\rho_n = N/\binom{n}{2}$, the randomized incomplete U-statistic is defined by

$$U'_{n,N} = \frac{1}{\widehat{N}} \sum_{\iota \in I_{n,2}} Z_\iota h_1(X_\iota),$$

where $\widehat{N} = \sum_{\iota \in I_{n,2}} Z_\iota$ is the number of successes and $X_\iota = (X_{i_1}, X_{i_2})$ for $\iota = (i_1, i_2)$. Intuitively, one should expect, by summing over only a random sample of indices in $I_{n,2}$ with an appropriately chosen computational budget N, that $U_{n,N}$ finds a balanced middle ground between $S_{\lfloor n/2 \rfloor}$ and the complete U-statistic. Indeed, our proposal is motivated by a weak convergence result in Janson (1984, Corollary 1). Assuming that N and n tend to infinity such that the ratio $\alpha_n := n/N$ converges to a positive constant, the result implies that

$$\frac{\sqrt{n}\, U'_{n,N}}{\sigma_1} \longrightarrow_d N(0, 1),$$

where $\sigma_1^2 = m^2 \sigma_{g_1}^2 + \alpha_n \sigma_{h_1}^2$ and $m = 2$ if the kernel is as in (3.2). Importantly, the standardizing factor σ_1 is always positive since $\sigma_{h_1}^2 > 0$.

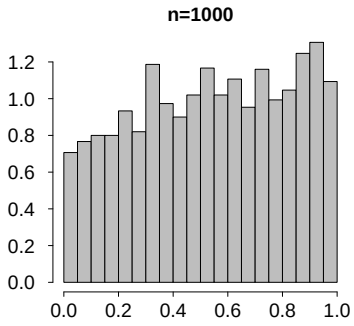


Figure 3.3 Histogram of 5000 simulated p-values using our methodology based on incomplete U-statistics for simultaneously testing 2730 tetrads constraints implied by the one-factor model with 15 observed variables. The computational budget parameter for the incomplete U-statistic is $N = 2n$ and the true covariance matrix is close to a singular point, for exact parameter values see Sturma et al. (2024, Section 5.2), setup (b).

This is our key motivation to propose a testing strategy based on incomplete U-statistics in Sturma et al. (2024). It allows for equality and inequality constraints and it is valid in singular scenarios while still being as efficient as possible. Moreover, our testing methodology allows for a potentially very large number of constraints p. To define the test statistic, we assume that $f(\theta) = (f_1(\theta), \dots, f_p(\theta))^\top$ is U-estimable. That is, for some integer $m \geq 2$ there exists a $\mathbb{R}^p$-valued measurable symmetric function $h(x_1, \dots, x_m)$ such that $\mathbb{E}_\theta[h(X_1, \dots, X_m)] = f(\theta)$ for all $\theta \in \Theta$, when $X_1, \dots, X_m$ are i.i.d. with distribution P_θ. Using the same notation as above but replacing 2 with $m \geq 2$ and the one-dimensional kernel $h_1(x_1, x_2)$ with a p-dimensional kernel $h(x_1, \dots, x_m)$, the randomized incomplete U-statistic based on Bernoulli sampling is defined by

$$U'_{n,N} = \frac{1}{\widehat{N}} \sum_{\iota \in I_{n,m}} Z_\iota h(X_\iota).$$

Due to the form of the null hypothesis Θ_0, we define the test statistic as the maximum of the studentized incomplete U-statistic, i.e.

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

where $\widehat{\sigma}_j^2$ is an estimator of the variance $\sigma_j^2 = m^2 \sigma_{g,j}^2 + \alpha_n \sigma_{h,j}^2$ of the j-th coordinate of the approximating Gaussian distribution. Similar as above, $\sigma_{h,j}^2$ is the variance of the j-th coordinate of the kernel $h(X_1, \dots, X_m)$ and $\sigma_{g,j}^2$ is the variance of the j-th coordinate of the Hájek projection $g(X_1) = \mathbb{E}_\theta[h(X_1, \dots, X_m)|X_1]$. In practice, a data driven estimator $\widehat{\sigma}_j^2$ can be formed with a "divide-and-conquer" strategy and critical values are derived via a data-dependent Gaussian multiplier bootstrap; we refer to Sturma et al. (2024) for the details.

The main theoretical contribution of our work is to show that the Gaussian approximation of high-dimensional incomplete U-statistics is valid under *mixed degeneracy* when we choose the computational budget parameter appropriately, typically of the same order as the sample size. In simple words, mixed degeneracy implies that for each index j, the variance of the Hájek projection $\sigma_{g,j}^2$ is allowed to take more or less arbitrary values, including zero. Based on the Gaussian approximation, we prove that our test asymptotically controls type I error if N is chosen appropriately, even if the true parameter θ is singular or almost singular. Since our results rely on the seminal work in high-dimensional Gaussian approximation

of Chernozhukov et al. (2013) and Chernozhukov et al. (2017), the Gaussian approximation also remains valid in settings where the number of constraints p can be much larger than the sample size n. Finally, we emphasize that there is no need to maximize a possibly multi-modal function as one would do for a likelihood ratio test. As an illustration, Figure 3.3 shows a histogram of p-values using our testing strategy in a singular setting. We see that the distribution of the p-values is significantly closer to a uniform distribution than when using the Wald or likelihood ratio test.

4 Structure Identifiability in Causal Representation Learning

In this chapter, we introduce the problem of causal representation learning and we lay out the setting and the main contributions of the second core article (Sturma et al., 2023). In machine learning, the goal of representation learning is to find a description of data that is interpretable, useful for reasoning, and generalizable (Bengio et al., 2013; Schölkopf et al., 2021). Such a representation synthesizes available measurements into latent variables that serve to tackle downstream tasks. Formally, this can be expressed as follows. We assume that there is a collection of real-valued latent random variables $(X_\ell)_{\ell\in\mathcal{L}}$ that are indexed by a finite set $\mathcal{L}$ and jointly distributed according to $P_\mathcal{L}$. The observed variables $(X_v)_{v\in\mathcal{O}}$, indexed by a finite set $\mathcal{O}$, are now generated via

$$X_\mathcal{O} = g(X_\mathcal{L}),$$

where $g : \mathbb{R}^{|\mathcal{L}|} \to \mathbb{R}^{|\mathcal{O}|}$ is an injective *mixing function*. In representation learning, one is interested in estimating the function g as well as the distribution $P_\mathcal{L}$ from observed samples of the variables $X_\mathcal{O}$. Traditionally, the latent variables are assumed to be "conceptually distinct" (Bengio et al., 2013; Squires et al., 2023). For example, if the observations are images, the latent variables might be abstract objects like the presence of clouds or the presence of trees. Conceptual distinctness of variables is often translated to statistical independence, i.e., the latent variables X_ℓ for $\ell \in \mathcal{L}$ are assumed to be pairwise independent.

However, as argued in Schölkopf et al. (2021), the assumption of independence can be too stringent and a poor match to reality. For example, the presence of clouds and the presence of wet roads in an image may be dependent, since clouds may cause rain which may in turn cause wet roads. Thus, it is natural to seek a *causal representation*, i.e., it is assumed that the distribution $P_\mathcal{L}$ is determined by an unknown structural causal model among the latent variables $X_\mathcal{L}$. General structural causal models are discussed for example in Pearl (2009) and Peters et al. (2017). We will focus here on the case where the causal model among the latent variables is linear and the mixing function g is also linear. In this setup, causal representations are modeled by linear structural equation models corresponding to graphs $G = (\mathcal{O}\sqcup\mathcal{L}, D)$ that only have edges between the latent variables and from the observed to the latent variables, that is, $D \subseteq (\mathcal{L}\times\mathcal{L}) \cup (\mathcal{L}\times\mathcal{O})$. Hence, the linear structural equations are given by

$$\begin{aligned} X_\mathcal{O} &= \Lambda_{\mathcal{L},\mathcal{O}}^\top X_\mathcal{L} + \varepsilon_\mathcal{O}, \\ X_\mathcal{L} &= \Lambda_{\mathcal{L},\mathcal{L}}^\top X_\mathcal{L} + \varepsilon_\mathcal{L}, \end{aligned}$$

where $\Lambda \in \mathbb{R}^D$ is sparse with support corresponding to the edges of the graph G, recall Section 2.3. Since $X_\mathcal{L} = (I - \Lambda_{\mathcal{L},\mathcal{L}})^{-\top}\varepsilon_\mathcal{L}$, the observed random vector satisfies in particular the equation

$$X_\mathcal{O} = \Lambda_{\mathcal{L},\mathcal{O}}^\top (I - \Lambda_{\mathcal{L},\mathcal{L}})^{-\top}\varepsilon_\mathcal{L} + \varepsilon_\mathcal{O}. \tag{4.1}$$

The main interest in causal representation learning is to learn the induced graph on the latent nodes $G[\mathcal{L}] = (\mathcal{L}, D_{\mathcal{L}})$ with edges $D_{\mathcal{L}} = D \cap (\mathcal{L} \times \mathcal{L})$ from observed samples of the random vector $X_{\mathcal{O}}$. Structure identifiability in causal representation learning thus refers to the question under which conditions the recovery of the latent graph is possible in the case of infinite data, i.e., when the distribution of $X_{\mathcal{O}}$ is known. Note that this is a question of model identifiability, which we introduced in Section 2.2. Without further assumptions, the problem is ill-posed and it is in general impossible to recover the latent graph $G[\mathcal{L}]$; we refer to Appendix C in Sturma et al. (2023) for a discussion. In the literature, there are mainly two types of assumptions under which identifiability was shown to be possible. The first line of work studies identifiability while assuming that interventional data is available, see e.g. Liu et al. (2023), Squires et al. (2023) and Ahuja et al. (2023). This means that, in addition to the observed distribution, we also have access to interventional distributions of $X_{\mathcal{O}}$. The interventional distributions are generated equivalently as in (4.1), but the structural equations of one or more latent variables X_ℓ are changed. For example, a "perfect intervention" removes the dependency of X_ℓ on its parents while still allowing for stochasticity; the structural equation then becomes $X_\ell = \widetilde{\varepsilon}_\ell$ for a different noise variable $\widetilde{\varepsilon}_\ell$. Another line of work shows that the graph is identified under certain sparsity assumptions on the mixing matrices $\Lambda_{\mathcal{L},\mathcal{O}}$; we refer to Xie et al. (2020), Chen et al. (2022), Xie et al. (2022) and Huang et al. (2022). Identifiability by assuming either access to interventions or sparsity in the mixing function is also studied in nonlinear setups; see e.g. von Kügelgen et al. (2023), Khemakhem et al. (2020) or Buchholz et al. (2022) for recent advances.

On the other hand, one may also leverage identifiability results from independent component analysis to recover the graph $G[\mathcal{L}]$. For simplicity, assume that the observed random vector $X_{\mathcal{O}}$ does not feature noise, that is, $\varepsilon_{\mathcal{O}} \equiv 0$. This is to avoid the overcomplete setting, recall Example 2.10. If $\Lambda_{\mathcal{L},\mathcal{O}}^\top (I - \Lambda_{\mathcal{L},\mathcal{L}})^{-\top}$ has full column rank, at most one of the distributions of the noise variables ε_ℓ for $\ell \in \mathcal{L}$ is Gaussian, and each distribution of ε_ℓ is assumed to have unit variance, then the matrix $\Lambda_{\mathcal{L},\mathcal{O}}^\top (I - \Lambda_{\mathcal{L},\mathcal{L}})^{-\top}$ is identified up to a signed permutation of its columns (Comon, 1994). That is, we obtain the matrix

$$\Lambda_{\mathcal{L},\mathcal{O}}^\top (I - \Lambda_{\mathcal{L},\mathcal{L}})^{-\top} \Psi, \tag{4.2}$$

where Ψ is a $|\mathcal{L}| \times |\mathcal{L}|$ signed permutation matrix. Adams et al. (2021) derive a sufficient and necessary condition with respect to the sparsity of $\Lambda_{\mathcal{L},\mathcal{O}}$ under which it is possible to identify the graph $G[\mathcal{L}]$, i.e., to identify the support of $\Lambda_{\mathcal{L},\mathcal{L}}$.

In Sturma et al. (2023), we study identifiability in *multi-domain* causal representation learning. In many applications, researchers have access to unpaired data, possibly with different modalities such as image data or text, and the researcher is interested in learning a *shared* causal representation. Our main motivation to study multi-domain causal representation learning comes from the analysis of single-cell gene expression data in biology. Given a population of cells, different technologies such as imaging and sequencing provide different views on same the cell population. Crucially, since these technologies destroy the cells, the observations are unpaired, i.e., a specific cell may either be used for imaging or sequencing but not both; see e.g. Butler et al. (2018), Stuart et al. (2019), Liu et al. (2019) or Yang et al. (2021). The aim of this work is to investigate conditions under which it is possible to integrate the different types of data in each domain in order to understand the underlying common causal mechanisms that determine the

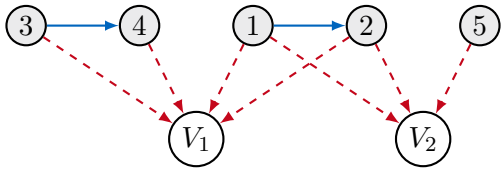


Figure 4.1 Compact version of a 2-domain graph $G = (\mathcal{L} \sqcup V_1 \sqcup V_2)$ with five latent nodes $\mathcal{L} = \{1, 2, 3, 4, 5\}$ and two domains V_1 and V_2. All observed nodes in each domain are represented by a single white node. We draw a dashed red edge from latent node $\ell \in \mathcal{L}$ to domain $V_e \subseteq \mathcal{O}$ if $\ell \in S_e$. The latent nodes $\mathcal{H} = \{1, 2\}$ are shared and $I_1 = \{3, 4\}$ and $I_2 = \{5\}$ are domain-specific. Importantly, the random vectors X_{V_1}, X_{V_2} associated to the two domains are uncoupled, that is, we do not observe their joint distribution.

observed features. Unpaired multi-domain data also appears in many applications other than single-cell biology; see Sturma et al. (2023) for a discussion.

In a linear setting, we define unpaired multi-domain causal representation learning as follows. The observed nodes $\mathcal{O}$ of the graph are a union of pairwise disjoint indexing sets $V_1, \dots, V_m$, where each indexing set corresponds to one observed domain. Importantly, the cardinality of the indexing sets may be different to reflect possibly different data modalities. In each domain, the observed random vector X_{V_e} is a linear function of a subset of the latent variables, that is, X_{V_e} is a linear function of X_{S_e}, where $S_e \subseteq \mathcal{L}$. We assume that there is a *shared* latent set of nodes $\mathcal{H} \subseteq \mathcal{L}$ such that in each domain $e \in \{1, \dots, m\}$, the latent variables are given as a disjoint union $S_e = \mathcal{H} \cup I_e$, and we say that $I_e = S_e \setminus \mathcal{H}$ are the *domain-specific* latent nodes. Recall that we are motivated by settings where the shared latent variables $X_{\mathcal{H}}$ capture the key causal relations and the different domains are able to give us combined information about these relations. Likewise, we may think about the domain-specific latent variables X_{I_e} as "noise" in each domain. A priori, we do not assume that we know the shared latent nodes $\mathcal{H}$, we do not even assume to know their cardinality $|\mathcal{H}|$. Finally, we assume hat the observed data in each domain is *unpaired*, that is, we assume to know only the marginal distributions of X_{V_e} in each domain, but none of the joint distributions over pairs X_{V_e}, X_{V_f} for $e \neq f$. In our core article we rigorously introduce this setup by defining *m-domain graphs*, such that each m-domain graph defines a *multi-domain causal representation model* via the corresponding structural equations; see Figure 4.1 for an illustration. As argued in the introduction of the article, our setup constitutes a generalization of the setups considers in previous works, for example the works of Adams et al. (2021) and Zeng et al. (2021).

The main challenge in proving rigorous identifiability results in the multi-domain setup is that we cannot apply existing results from the single-domain setup in each domain separately. Even if the causal structure of the latent variables in a single domain is identifiable, it remains unclear how to combine multiple causal structures, i.e., in which way latent variables are shared. We circumvent this problem via a two-step approach: First, we extend the identifiability of linear independent component analysis to the multi-domain setup, which allows us to identify the joint distribution and distinguish between shared and domain-specific latent variables. Moreover, we identify an "overall mixing matrix", similar to the one in Equation (4.2) and, in a second step, exploit sparsity constraints in this matrix to identify the causal structure among the shared latent variables. Hence, our main contributions are two-fold: First, we lay out sufficient conditions under which we can identify the joint distribution of $X_{V_1}, \dots, X_{V_e}$. Second, we give additional conditions

under which we are able to identify the graph of the shared latent variables $G[\mathcal{H}] = (\mathcal{H}, D_{\mathcal{H}})$ with $D_{\mathcal{H}} = D \cap (\mathcal{H} \times \mathcal{H})$. As discussed in detail in our paper, our conditions are mostly necessary. However, it remains an open problem to find an “if and only if” condition in the linear multi-domain setup.

5 Parameter Identifiability in Linear Causal Models

In this chapter, we discuss the problem of parameter identifiability in linear causal models and explain the contributions in Barber et al. (2022), which are part of this thesis. As already mentioned in the introduction of this thesis, the gold standard to draw causal conclusions would be to set up a controlled experiment to physically intervene in a system of random variables and learn about the causal effects. If this is infeasible due to costs or ethical and technical constraints, it may still be possible to hypothesize or determine the direction of causal relationships. Since causal relationships in a collection of random variables are naturally represented by directed graphs, this corresponds to setups where the graphical structure is assumed to be known. It is then of interest to investigate whether causal effects are uniquely determined based on the observed data only, i.e., when there is no data from controlled experiments available. This is the problem of *causal effect identifiability* (Pearl, 2009) and is especially challenging in the presence of latent variables. If the causal relations are assumed to be linear, we are hence interested in identifying the coefficients in linear structural equation models with latent variables. In the context of causal effect identifiability, the linear coefficients are also known as *direct causal effects*.

Example 5.1. Assume that we want to quantify the effect of smoking on the baby's birth weight as in Example 1.1. It is reasonable to hypothesize the causal relationships as in the instrumental variable model given by the graph in Figure 1.1. Identifying the direct causal effect from smoking on the baby's birth weight then corresponds to identifying the coefficient λ_{12} appearing in the linear structural equations (1.1). Given the covariance matrix $\Sigma = (\sigma_{vw})$ of the observed random variables, we aim to find a formula in the entries σ_{vw} that identifies λ_{12}. We have already seen in Example 1.1 that the equation $\sigma_{23} = \lambda_{12}\sigma_{13}$ holds in the model. Hence, we obtain the identification formula $\lambda_{12} = \sigma_{23}/\sigma_{13}$ that is valid as long as σ_{13} is nonzero. Since $\sigma_{13} = \omega_{33}\lambda_{31}$, it is nonzero for 'almost all' choices of the parameters ω_{33} and λ_{31} (with respect to an absolutely continuous distribution).

Identifying the direct causal effects from the observable covariance matrix has a long history in structural equation modeling. Most prior research focused on developing methods to decide identifiability of direct effects in a latent projection framework, in which the confounding effects of the latent variables are represented by correlation among noise terms.

Example 5.2. Consider again the instrumental variable model given by the graph in Figure 1.1. In the latent projection framework, the latent variable X_4 is not explicitly modeled, but the effect of the unobserved confounder is absorbed in a possibly nonzero correlation between the noise terms ε_1 and ε_2. Now, the linear structural equations become

$$X_1 = \lambda_{01} + \lambda_{31}X_3 + \varepsilon_1, \qquad X_2 = \lambda_{02} + \lambda_{12}X_1 + \varepsilon_2, \qquad X_3 = \varepsilon_3,$$

Figure 5.1 Mixed graph for the instrumental variable model.

where the error vector ε has mean zero and covariance matrix

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} & 0 \\ \omega_{12} & \omega_{22} & 0 \\ 0 & 0 & \omega_{33} \end{pmatrix}.$$

The projected latent variable model corresponds to the graph displayed in Figure 5.1, where the bidirected edge indicates the possibly nonzero correlation between the noise terms ε_1 and ε_2. Note that the identifying formula $\lambda_{12} = \sigma_{23}/\sigma_{13}$ still holds in the projected model.

In the latent projection framework, models correspond to mixed graphs $G = (\mathcal{O}, D, B)$ that contain the observed nodes $\mathcal{O}$ and additional bidirectional edges collected in the set $B \subseteq \mathcal{O} \times \mathcal{O}$. We denote the edges in B by $v \leftrightarrow w$, and they have no orientation, that is $v \leftrightarrow w \in B$ if and only if $w \leftrightarrow v \in B$. Moreover, we exclude self-loops, i.e., $v \leftrightarrow v \notin B$ for all $v \in \mathcal{O}$. As before, the set D contains directed edges and is also assumed to be free of self-loops.

Definition 5.3 (Foygel et al., 2012, Definition 1)**.** Let $G = (\mathcal{O}, D, B)$ be a mixed graph with $\mathcal{O} = [m]$. The *covariance model* given by G is the image $\mathrm{Im}(\tau_G)$ of the parametrization map

$$\begin{aligned} \tau_G : \mathbb{R}^D_{\text{reg}} \times \mathit{PD}(B) &\longrightarrow \mathit{PD}(|\mathcal{O}|), \\ (\Lambda, \Omega) &\longmapsto (I - \Lambda)^{-\top} \Omega (I - \Lambda)^{-1}, \end{aligned}$$

where $\mathit{PD}(B) \subseteq \mathit{PD}(|\mathcal{O}|)$ is the subcone of matrices with support B, that is, $\omega_{vw} = 0$ if $v \neq w$ and $v \leftrightarrow w \notin B$.

Identifiability of a model holds, if all possibly nonzero direct causal effects λ_{vw} can be uniquely recovered from a given observable covariance matrix $\Sigma \in \mathrm{Im}(\tau_G)$. That is, identifiability holds if the parametrization map is injective. In practice, however, it is usually enough to require injectivity only on a suitably large subset of the domain $\Theta := \mathbb{R}^D_{\text{reg}} \times \mathit{PD}(B)$. This motivates the following definition.

Definition 5.4 (Foygel et al., 2012, Definition 3)**.** The mixed graph $G = (\mathcal{O}, D, B)$ is said to be *generically identifiable* if there exists a proper algebraic subset $A \subset \Theta$ such that τ_G is injective on $\Theta \setminus A$.

An algebraic set is a semi-algebraic set that is defined by equality constraints only; we refer to Cox et al. (2015) or Shafarevich (2013) for background on algebraic geometry. The set $\Theta \setminus A$ is suitably large in statistical applications since proper algebraic subsets $A \subset \Theta$ have Lebesgue measure zero; see, for example, the lemma in Okamoto (1973). In all examples we know, the inverse mapping, also known as the identification formula, is a rational map; see Foygel et al. (2012) and Barber et al. (2022) for a discussion. Hence, we are interested in *rational identifiability* of mixed graphs, which is defined as follows.

Figure 5.2 Left: Latent-factor graph with one latent node. Right: Corresponding mixed graph.

Definition 5.5 (Foygel et al., 2012, Definition 3)**.** The mixed graph $G = (\mathcal{O}, D, B)$ is said to be *rationally identifiable* if there exists a proper algebraic subset $A \subset \Theta$ and a rational map ψ such that $\psi \circ \tau_G(\Lambda, \Omega) = (\Lambda, \Omega)$ for all $(\Lambda, \Omega) \in \Theta \setminus A$.

In principle, the problem of deciding rational identifiability can be solved via Gröbner basis computations; see Garcia-Puente et al. (2010) and Cox et al. (2015). But the complexity of these computations can, in the worst case, be double exponential in the size of the graph (Bardet et al., 2015). Hence they are infeasible even for relatively small graphs and it becomes of great value to develop criteria that can be efficiently checked.

Examples of graphical criteria for mixed graphs include instrumental variables (Bowden and Turkington, 1984), conditional instruments (Brito and Pearl, 2002), the G-criterion (Brito and Pearl, 2006), the half-trek criterion (Foygel et al., 2012), the generalized half-trek criterion (Weihs et al., 2017), auxiliary variables (Chen et al., 2017) and auxiliary cutsets (Kumor et al., 2020); see Barber et al. (2022) for a more comprehensive list of references. Most of these criteria are sufficient conditions for rational identifiability, the only necessary condition is given in Foygel et al. (2012). To our knowledge, it is still an open problem to find an "if and only if" condition for rational identifiability of mixed graphs that can be verified in polynomial time, or to prove fundamentally that such a criterion cannot exist.

In contrast, in our new work in Barber et al. (2022), we consider unprojected models where latent variables are explicitly modeled. Our main observation is that the approach using mixed graphs in the latent projection framework is only effective if the confounding is sparse, that is, if there are only few bidirected edges. A mixed graph can not be rationally identifiable if the dimension of the domain $\Theta = \mathbb{R}^D_{\text{reg}} \times \mathit{PD}(B)$ is larger than the dimension of the ambient space $\mathit{PD}(|\mathcal{O}|)$. Hence, rational identifiability is impossible if the total number of edges $|D| + |B|$ is strictly larger than $\binom{|\mathcal{O}|}{2}$; see Foygel et al. (2012, Proposition 2). In particular, if the number of bidirected edges is large, possibly even equal to $\binom{|\mathcal{O}|}{2}$, then a graph can never be rationally identifiable even if it contains only a small number of directed edges.

Example 5.6. Consider the graph in the left of Figure 5.2. It entails that all pairs of observed variables are confounded by a single latent variable. Hence, the corresponding mixed graph displayed on the right contains all possible bidirected edges, i.e., $|B| = \binom{5}{2} = 10$. Since there are also 3 directed edges, the mixed graph is not rationally identifiable by our considerations above. However, the graph on the left is rationally identifiable, that is, we can find a rational identifying formula for all directed edges. For example, we have the formula $\lambda_{23} = (\sigma_{13}\sigma_{24} - \sigma_{14}\sigma_{23})/(\sigma_{12}\sigma_{24} - \sigma_{14}\sigma_{22})$. Even though the latent confounding is dense and effects all observed variables, identification is possible since the confounding is caused by only one latent variable, i.e. there is structure in the confounding.

In our work in Barber et al. (2022), we restrict ourselves to graphs where all latent variables are independent "factors", that is, there are only edges from latent variables to observed variables and not viceversa, and there are also no edges between the latent variables. We say that a directed graph $G = (\mathcal{O} \sqcup \mathcal{L}, D)$ is a *latent-factor graph* if $D \subseteq (\mathcal{O} \times \mathcal{O}) \cup (\mathcal{L} \times \mathcal{O})$. The covariance model of a latent-factor graph is then given as the image of the parametrization map in Definition 2.12. In particular, each covariance matrix Σ in the model is parametrized as

$$\begin{aligned}\Sigma &= [(I-\Lambda)^{-\top}\Omega(I-\Lambda)^{-1}]_{\mathcal{O},\mathcal{O}} \\ &= (I-\Lambda_{\mathcal{O},\mathcal{O}})^{-\top}(\Omega_{\mathcal{O},\mathcal{O}} + \Lambda_{\mathcal{L},\mathcal{O}}^{\top}\Omega_{\mathcal{L},\mathcal{L}}\Lambda_{\mathcal{L},\mathcal{O}})(I-\Lambda_{\mathcal{O},\mathcal{O}})^{-1},\end{aligned}$$

where $\Lambda \in \mathbb{R}^{D}_{\text{reg}}$ and $\Omega \in \text{diag}^{+}_{|V|}$ with $V = \mathcal{O} \sqcup \mathcal{L}$. The second equality can easily be verified by taking into account the restricted set of edges in latent-factor graphs. Rational identifiability of latent-factor graphs then holds if we can find a rational identifying formula for all nonzero entries in $\Lambda_{\mathcal{O},\mathcal{O}}$; see Barber et al. (2022, Definition 2.2) for a precise definition.

Our key idea is to exploit the low rank structure in the "latent covariance matrix" $\Omega_{\mathcal{O},\mathcal{O}} + \Lambda_{\mathcal{L},\mathcal{O}}^{\top}\Omega_{\mathcal{L},\mathcal{L}}\Lambda_{\mathcal{L},\mathcal{O}}$ if the number of latent variables is locally small. That is, if the number of latent parents $\text{pa}_{\mathcal{L}}(U) = \{\ell \in \mathcal{L} : \ell \to u \in D \text{ for some } u \in U\}$ of a set of observed nodes $U \subseteq \mathcal{O}$ is small, say $|\text{pa}_{\mathcal{L}}(U)| = k$, then the rank of the matrix $[\Lambda_{\mathcal{L},\mathcal{O}}^{\top}\Omega_{\mathcal{L},\mathcal{L}}\Lambda_{\mathcal{L},\mathcal{O}}]_{U,U}$ is at most k. In our work in Barber et al. (2022) we leverage this low-rank structure to derive a graphical criterion, the *latent-factor half-trek criterion*, that is an effective sufficient condition for rational identifiability. To our knowledge, it is the first criterion that is applicable to general latent-factor graphs. Previous works only consider restricted settings, for example only one latent variable is considered (Stanghellini and Wermuth, 2005; Leung et al., 2016) or each latent variable is restricted to have an effect on only a few observed variables (Van Der Zander et al., 2015). We refer to the introduction in Barber et al. (2022) for a detailed discussion of related literature. Importantly, our graphical criterion can be checked in polynomial time in the size of the graph using max-flow computations (Cormen et al., 2009) if we search only over subsets of latent nodes of bounded size. Our algorithm is implemented in the R-package `SEMID` as of version 0.4.0 (Drton et al., 2022), which is available on `CRAN`, the Comprehensive R Archive Network.

In the discussion section in Barber et al. (2022), we propose various interesting research questions that arise from our work. Besides improving the latent-factor half-trek criterion, it is also an open problem to find a necessary condition for latent factor graphs. Moreover, it would be interesting to generalize our criterion to graphs that do not restrict the edges, i.e., to allow edges between latent variables and from observed to latent variables. Finally, note that the latent covariance matrix $\Omega_{\mathcal{O},\mathcal{O}} + \Lambda_{\mathcal{L},\mathcal{O}}^{\top}\Omega_{\mathcal{L},\mathcal{L}}\Lambda_{\mathcal{L},\mathcal{O}}$ corresponds to a covariance matrix in a sparse factor analysis model. Since this is the key object that allows to derive identifying formulas, it is important to better understand the geometry of sparse factor analysis models. This is the topic of the next chapter.

6 Geometry of Sparse Factor Analysis Models

In this chapter, we discuss the contributions in the last article (Drton et al., 2024), where we study the model geometry in sparse factor analysis. Factor analysis is a statistical technique that explains correlation among observed random variables with the help of a small number of latent factors, recall Examples 1.2 and 2.2, and the illustration of the one-factor analysis model in Figure 1.2. It is a popular tool in many applied sciences, including psychology (Horn, 1965; Reise et al., 2000; Caprara et al., 1993), econometrics (Fan et al., 2008; Aßmann et al., 2016), education (Schreiber et al., 2006; Beavers et al., 2013), and epidemiology (Martínez et al., 1998).

Formally, factor analysis models are defined as the covariance models corresponding to graphs $G = (\mathcal{O} \sqcup \mathcal{L}, D)$ where only edges from latent to observed nodes are allowed, i.e., $D \subseteq (\mathcal{L} \times \mathcal{O})$, recall Definition 2.12. Without loss of generality we can fix the scale of the latent variables $\omega_{\ell\ell} = \mathsf{Var}(\varepsilon_\ell) = 1$ for all $\ell \in \mathcal{L}$ since this does not change the image of the parametrization map τ_G. The parametrization of a covariance matrix Σ in a sparse factor model then takes the form

$$\Sigma = \Omega_{\mathcal{O},\mathcal{O}} + \Lambda_{\mathcal{L},\mathcal{O}}^\top \Lambda_{\mathcal{L},\mathcal{O}},$$

where $\Lambda \in \mathbb{R}^D_{\mathsf{reg}}$ and $\Omega_{\mathcal{O},\mathcal{O}} \in \mathsf{diag}^+_{|\mathcal{O}|}$. Traditionally, the term "factor analysis" refers to models that correspond to graphs with all possible edges $D = \mathcal{L} \times \mathcal{O}$, recall Example 2.9. The one-factor model was first studied by Spearman (1904), and the geometry of more general models was first rigorously studied in the work of Anderson and Rubin (1956). The paper by Drton et al. (2007), that can be seen as a predecessor of our work in Drton et al. (2024), first studies the geometry under an algebraic perspective by realizing that the entries of the covariance matrix are polynomial functions in the parameters.

In contrast, we are interested in *sparse* factor analysis models, where the edge set $D \subset \mathcal{L} \times \mathcal{O}$ is a proper subset. As discussed in our article, sparse factor models naturally appear in many applications. We have already seen in Section 5 that they are the building block for linear causal models with latent variables. Examples of different research topics where sparse factor models appear include work on correlation thresholding (Kim and Zhou, 2023), l_1-penalization (Lan et al., 2014; Trendafilov et al., 2017), and Bayesian approaches (Frühwirth-Schnatter et al., 2024; Ohn et al., 2023).

The main contribution of our work is twofold: First, we study the dimension of sparse factor models. We give a general upper bound on the dimension which reveals that sparse factor models can be defective. That is, they may not have expected dimension which is equal to counting the number of parameters. This is a difference to full factor analysis models that are always of expected dimension after an orthogonal transformation of the coefficient matrix $\Lambda_{\mathcal{L},\mathcal{O}}$; see Drton et al. (2007) for details. We also provide a lower bound for the dimension of all models that satisfy a mild condition on the sparsity pattern. In many cases

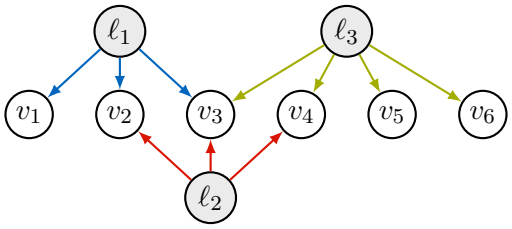


Figure 6.1 Graph corresponding to a sparse factor analysis model with 3 latent variables and 6 observed variables.

we show that the upper and lower bounds coincide, so that we obtain a combinatorial formula for the dimension in these cases.

Example 6.1. Consider the graph in Figure 6.1. The expected dimension of the corresponding model is equal to the number of parameters given by $|\mathcal{O}| + |D| = 16$. However, as we verify in Drton et al. (2024, Corollary 2.13), the dimension of the model is 15.

Second, we study implicit descriptions of sparse factor models, where we focus on equality constraints. That is, for a given model $\mathcal{M} \subseteq PD(|\mathcal{O}|)$, we are interested in the *ideal of invariants*

$$I(\mathcal{M}) = \{f \in \mathbb{R}[\sigma_{ij}, i \leq j] : f(\Sigma) = 0 \text{ for all } \Sigma \in \mathcal{M}\}$$

that contains the polynomials that vanish on all points in the model. Since, for a symmetric positive definite matrix $\Sigma \in \mathbb{R}^{|\mathcal{O}|\times|\mathcal{O}|}$, membership in a sparse factor analysis model $\mathcal{M} = \text{Im}(\tau_G)$ only depends on the off-diagonal entries of Σ, we can regard the ideal of invariants of $\text{Im}(\tau_G)$ as an ideal in the subring $\mathbb{R}[\sigma_{ij},\ i < j]$. Our main contribution is to give an explicit description for the generators of Gröbner bases with respect to any circular term order for a subclass of sparse two-factor analysis models. We refer to Cox et al. (2015) for background on Gröbner bases and to de Loera et al. (1995) for background on circular term orders. The advantage of an explicit combinatorial description for the Gröbner basis is that we can avoid computing generators of the ideal by using Buchberger's Algorithm (Buchberger, 2006), which in the worst case has a double exponential complexity in the number of variables (Bardet et al., 2015). Algebraically, the main challenge in studying ideals of sparse factor models in comparison to studying ideals of full factor models is that they corresponds to joins instead of secants; see Sturmfels and Sullivant (2006) for background of join ideals.

Example 6.2. Consider the graph in Figure 6.2. We show in Drton et al. (2024, Theorem 3.13) that the following collection of polynomials defines a Gröbner basis of the ideal of invariants of the corresponding sparse factor analysis model with respect to a so-called circular term order.

- Degree one monomials: $\sigma_{16}, \sigma_{17}, \sigma_{26}, \sigma_{27}, \sigma_{36}, \sigma_{37}$.
- Degree two binomials (tetrads): $\sigma_{47}\sigma_{56} - \sigma_{57}\sigma_{46}$, $\sigma_{12}\sigma_{34} - \sigma_{13}\sigma_{24}$ $\sigma_{14}\sigma_{23} - \sigma_{13}\sigma_{24}$, $\sigma_{12}\sigma_{35} - \sigma_{13}\sigma_{25}$, $\sigma_{15}\sigma_{23} - \sigma_{13}\sigma_{25}$, $\sigma_{15}\sigma_{24} - \sigma_{14}\sigma_{25}$, $\sigma_{15}\sigma_{34} - \sigma_{14}\sigma_{35}$, $\sigma_{25}\sigma_{34} - \sigma_{24}\sigma_{35}$.
- Degree three trinomials: $\sigma_{67}\sigma_{12}\sigma_{45} - \sigma_{67}\sigma_{24}\sigma_{15} - \sigma_{12}\sigma_{47}\sigma_{56}$, $\sigma_{67}\sigma_{13}\sigma_{45} - \sigma_{67}\sigma_{34}\sigma_{15} - \sigma_{13}\sigma_{47}\sigma_{56}$, $\sigma_{67}\sigma_{23}\sigma_{45} - \sigma_{67}\sigma_{34}\sigma_{25} - \sigma_{23}\sigma_{47}\sigma_{56}$.

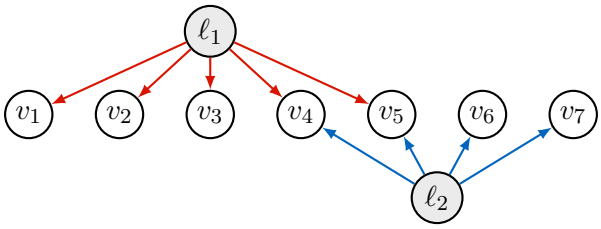


Figure 6.2 Graph corresponding to a sparse factor analysis model with 2 latent variables and 7 observed variables.

In Section 4 of Drton et al. (2024), we outline possible future research directions and open questions. In particular, it remains an open question to find a Gröbner basis for sparse factor analysis models with more than two latent factors. Beyond the discussion in our article, it would be interesting to study parameter identifiability of sparse factor analysis models, recall Definition 5.4. The results on the dimension in our article provide a necessary condition. That is, if the dimension is smaller than the number of parameters, then the parameters are certainly not identifiable. Finding a sufficient condition for (generic) parameter identifiability that can be efficiently checked remains an open problem. Parameter identifiability of *full* factor analysis models was, among others, studied in Bekker and ten Berge (1997): If the number of parameters is strictly smaller than the dimension of the ambient space after an orthogonal transformation, then the parameters in full factor analysis models are always generically identifiable up to certain trivial indeterminacies. However, this result does not translate to sparse models since they consist of non-generic points in full-factor models.

7 Conclusion

Many applications require models with latent variables that can have a complicated parameterization and geometry. Besides the advances in identifiability and inference of latent variable models presented in this thesis, there are still many open problems. Even for linear structural equation models, parameter and model identifiability are not completely understood as we discussed in the previous chapters. Moreover, we consider the study of linear models as the basis for any subsequent study of nonlinear models.

When fundamental questions regarding geometry and identifiability are solved, a set of more statistical questions arises. Besides hypotheses testing as discussed in this thesis, it is for example interesting to study estimation if parameters are certified to be identifiable. Criteria for parameter identifiability directly suggest a strategy for estimation by making use of the identification formula. Exploring different estimators and finding an optimal one is an interesting future direction. Similarly, if the structure in causal representation learning is identifiable, we still need to develop consistent and efficient methods to estimate the structure if only finitely many samples are available.

Bibliography

Adams, J., Hansen, N., and Zhang, K. (2021). Identification of partially observed linear causal models: Graphical conditions for the non-gaussian and heterogeneous cases. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W., editors, *Advances in Neural Information Processing Systems*, volume 34, pages 22822–22833. Curran Associates, Inc.

Ahuja, K., Mahajan, D., Wang, Y., and Bengio, Y. (2023). Interventional causal representation learning. In Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., and Scarlett, J., editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 372–407. PMLR.

Améndola, C., Drton, M., Grosdos, A., Homs, R., and Robeva, E. (2023). Third-order moment varieties of linear non-Gaussian graphical models. *Inf. Inference*, 12(3):1405–1436.

Anderson, T. W. and Rubin, H. (1956). Statistical inference in factor analysis. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. V*, pages 111–150. Univ. California Press, Berkeley-Los Angeles, Calif.

Ardiyansyah, M. and Sodomaco, L. (2023). Dimensions of higher order factor analysis models. *Algebr. Stat.*, 14(1):91–108.

Aßmann, C., Boysen-Hogrefe, J., and Pape, M. (2016). Bayesian analysis of static and dynamic factor models: An ex-post approach towards the rotation problem. *J. Econometrics*, 192(1):190–206.

Barber, R. F., Drton, M., Sturma, N., and Weihs, L. (2022). Half-trek criterion for identifiability of latent variable models. *Ann. Statist.*, 50(6):3174–3196.

Bardet, M., Faugère, J.-C., and Salvy, B. (2015). On the complexity of the f5 gröbner basis algorithm. *J. Symbolic Comput.*, 70:49–70.

Basu, S., Pollack, R., and Roy, M.-F. (2006). *Algorithms in real algebraic geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, second edition.

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., and Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Pract. Assess. Res. Eval.*, 18(6):1–13.

Bekker, P. A. and de Leeuw, J. (1987). The rank of reduced dispersion matrices. *Psychometrika*, 52(1):125–135.

Bekker, P. A. and ten Berge, J. M. F. (1997). Generic global indentification in factor analysis. *Linear Algebra Appl.*, 264:255–263.

Benedetti, R. and Risler, J.-J. (1990). *Real algebraic and semi-algebraic sets*. Actualités Mathématiques. [Current Mathematical Topics]. Hermann, Paris.

Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(8):1798–1828.

Bochnak, J., Coste, M., and Roy, M.-F. (1998). *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin. Translated from the 1987 French original, Revised by the authors.

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York. A Wiley-Interscience Publication.

Bollen, K. A. and Ting, K.-F. (2000). A tetrad test for causal indicators. *Psychol. Methods*, 5(1):3–22.

Bowden, R. J. and Turkington, D. A. (1984). *Instrumental variables*, volume 8 of *Econometric Society Monographs in Quantitative Economics*. Cambridge University Press, Cambridge.

Brito, C. and Pearl, J. (2002). Generalized instrumental variables. In *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence*, pages 85–93. Morgan Kaufmann Publishers Inc.

Brito, C. and Pearl, J. (2006). Graphical condition for identification in recursive SEM. In *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence*, pages 47–54. AUAI Press.

Buchberger, B. (2006). An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal. *J. Symbolic Comput.*, 41(3-4):475–511. Translated from the 1965 German original by Michael P. Abramson.

Buchholz, S., Besserve, M., and Schölkopf, B. (2022). Function classes for identifiable nonlinear independent component analysis. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A., editors, *Advances in Neural Information Processing Systems*, volume 35, pages 16946–16961. Curran Associates, Inc.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.*, 36(5):411–420.

Caprara, G. V., Barbaranelli, C., Borgogni, L., and Perugini, M. (1993). The “big five questionnaire”: A new questionnaire to assess the five factor model. *Pers. Individ. Dif.*, 15(3):281–288.

Casella, G. and Berger, R. L. (1990). *Statistical inference*. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA.

Chen, B., Kumor, D., and Bareinboim, E. (2017). Identification and model testing in linear structural equation models using auxiliary variables. In Precup, D. and Teh, Y. W., editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 757–766. PMLR.

Chen, B., Tian, J., and Pearl, J. (2014). Testable implications of linear structural equation models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, pages 2424–2430. Association for the Advancement of Artificial Intelligence (AAAI).

Chen, L. H. Y., Goldstein, L., and Shao, Q.-M. (2011). *Normal approximation by Stein's method*. Probability and its Applications (New York). Springer, Heidelberg.

Chen, Z., Xie, F., Qiao, J., Hao, Z., Zhang, K., and Cai, R. (2022). Identification of linear latent variable model with arbitrary distribution. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 6350–6357. Association for the Advancement of Artificial Intelligence (AAAI).

Chernozhukov, V., Chetverikov, D., and Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.*, 41(6):2786–2819.

Chernozhukov, V., Chetverikov, D., and Kato, K. (2017). Central limit theorems and bootstrap in high dimensions. *Ann. Probab.*, 45(4):2309–2352.

Claassen, T. and Heskes, T. (2012). A Bayesian approach to constraint based causal inference. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence*, pages 207–216. AUAI Press.

Comon, P. (1994). Independent component analysis, a new concept? *Signal Process.*, 36(3):287–314.

Comon, P. and Jutten, C. (2010). *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Elsevier.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2009). *Introduction to algorithms*. MIT Press, Cambridge, MA, third edition.

Cox, D. A., Little, J., and O'Shea, D. (2015). *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer, Cham, fourth edition. An introduction to computational algebraic geometry and commutative algebra.

Darmois, G. (1953). Analyse générale des liaisons stochastiques. Etude particulière de l'analyse factorielle linéaire. *Rev. Inst. Internat. Statist.*, 21:2–8.

de Loera, J. A., Sturmfels, B., and Thomas, R. R. (1995). Gröbner bases and triangulations of the second hypersimplex. *Combinatorica*, 15(3):409–424.

Didelez, V., Meng, S., and Sheehan, N. A. (2010). Assumptions of IV Methods for Observational Epidemiology. *Statist. Sci.*, 25(1):22–40.

Drton, M. (2009). Likelihood ratio tests and singularities. *Ann. Statist.*, 37(2):979–1012.

Drton, M. (2018). Algebraic problems in structural equation modeling. In *The 50th anniversary of Gröbner bases*, volume 77 of *Adv. Stud. Pure Math.*, pages 35–86. Math. Soc. Japan, Tokyo.

Drton, M., Foygel Barber, R., Sturma, N., and Weihs, L. (2022). *SEMID: Identifiability of linear structural equation models*. R package version 0.4.0 available on CRAN.

Drton, M., Grosdos, A., Portakal, I., and Sturma, N. (2024). Algebraic sparse factor analysis. *arXiv preprint arXiv:2312.14762*.

Drton, M., Hollering, B., and Wu, J. (2023). Identifiability of homoscedastic linear structural equation models using algebraic matroids. *arXiv preprint arXiv:2308.01821*.

Drton, M., Sturmfels, B., and Sullivant, S. (2007). Algebraic factor analysis: tetrads, pentads and beyond. *Probab. Theory Related Fields*, 138(3-4):463–493.

Drton, M., Sturmfels, B., and Sullivant, S. (2009). *Lectures on algebraic statistics*, volume 39 of *Oberwolfach Seminars*. Birkhäuser Verlag, Basel.

Drton, M. and Weihs, L. (2016). Generic identifiability of linear structural equation models by ancestor decomposition. *Scand. J. Stat.*, 43(4):1035–1045.

Drton, M. and Xiao, H. (2016). Wald tests of singular hypotheses. *Bernoulli*, 22(1):38–59.

Dufour, J.-M., Renault, E., and Zinde-Walsh, V. (2024+). Wald tests when restrictions are locally singular. *Ann. Statist.* (forthcoming), available at https://www.e-publications.org/ims/submission/AOS/user/submissionFile/53530?confirm=d0110e54.

Eriksson, J. and Koivunen, V. (2004). Identifiability, separability, and uniqueness of linear ICA models. *IEEE Signal Process. Lett.*, 11(7):601–604.

Evans, W. N. and Ringel, J. S. (1999). Can higher cigarette taxes improve birth outcomes? *J. Public Econ.*, 72(1):135–154.

Fan, J., Fan, Y., and Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. *J. Econometrics*, 147(1):186–197.

Foygel, R., Draisma, J., and Drton, M. (2012). Half-trek criterion for generic identifiability of linear structural equation models. *Ann. Statist.*, 40(3):1682–1713.

Frühwirth-Schnatter, S., Hosszejni, D., and Lopes, H. F. (2024). Sparse Bayesian Factor Analysis When the Number of Factors Is Unknown. *Bayesian Anal. (Advance Publication)*, pages 1–31.

Garcia-Puente, L. D., Spielvogel, S., and Sullivant, S. (2010). Identifying causal effects with computer algebra. In *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence*, pages 193–200. AUAI Press.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning*. Springer Series in Statistics. Springer, New York, second edition. Data mining, inference, and prediction.

Hipp, J. R. and Bollen, K. A. (2003). Model fit in structural equation models with censored, ordinal, and dichotomous variables: Testing vanishing tetrads. *Sociol. Methodol.*, 33(1):267–305.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185.

Huang, B., Low, C. J. H., Xie, F., Glymour, C., and Zhang, K. (2022). Latent hierarchical causal structure discovery with rank constraints. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A., editors, *Advances in Neural Information Processing Systems*, volume 35, pages 5549–5561. Curran Associates, Inc.

Hyvärinen, A. (2013). Independent component analysis: recent advances. *Philos. Trans. Roy. Soc. A*, 371(1984):2011.0534.

Janson, S. (1984). The asymptotic distributions of incomplete U-statistics. *Z. Wahrsch. Verw. Gebiete*, 66(4):495–505.

Khemakhem, I., Kingma, D., Monti, R., and Hyvärinen, A. (2020). Variational autoencoders and nonlinear ICA: A unifying framework. In Chiappa, S. and Calandra, R., editors, *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 2207–2217. PMLR.

Kim, D. S. and Zhou, Q. (2023). Structure learning of latent factors via clique search on correlation thresholded graphs. *arXiv preprint arXiv:2203.01471*.

Klenke, A. (2014). *Probability theory*. Universitext. Springer, London, second edition. A comprehensive course.

Koroljuk, V. S. and Borovskich, Y. V. (1994). *Theory of U-statistics*, volume 273 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht. Translated from the 1989 Russian original by P. V. Malyshev and D. V. Malyshev and revised by the authors.

Kumor, D., Cinelli, C., and Bareinboim, E. (2020). Efficient identification in linear structural causal models with auxiliary cutsets. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 5501–5510. PMLR.

Lan, A. S., Waters, A. E., Studer, C., and Baraniuk, R. G. (2014). Sparse factor analysis for learning and content analytics. *J. Mach. Learn. Res.*, 15:1959–2008.

Lauritzen, S. L. (1996). *Graphical Models*, volume 17 of *Oxford Statistical Science Series*. The Clarendon Press, Oxford University Press, New York. Oxford Science Publications.

Lehmann, E. L. and Casella, G. (1998). *Theory of point estimation*. Springer Texts in Statistics. Springer-Verlag, New York, second edition.

Lehmann, E. L. and Romano, J. P. (2022). *Testing statistical hypotheses*. Springer Texts in Statistics. Springer, Cham, fourth edition.

Leung, D., Drton, M., and Hara, H. (2016). Identifiability of directed Gaussian graphical models with one latent source. *Electron. J. Stat.*, 10(1):394–422.

Leung, D. and Sturma, N. (2024). Singularity-agnostic incomplete U-statistics for testing polynomial constraints in Gaussian covariance matrices. *arXiv preprint arXiv:2401.02112*.

Liu, J., Huang, Y., Singh, R., Vert, J.-P., and Noble, W. S. (2019). Jointly embedding multiple single-cell omics measurements. In Huber, K. T. and Gusfield, D., editors, *19th International Workshop on Algorithms in Bioinformatic*, volume 143 of *Leibniz International Proceedings in Informatics*, pages 10:1–10:13. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Liu, Y., Zhang, Z., Gong, D., Gong, M., Huang, B., van den Hengel, A., Zhang, K., and Shi, J. Q. (2023). Identifying weight-variant latent causal models.

Maathuis, M., Drton, M., Lauritzen, S., and Wainwright, M., editors (2019). *Handbook of graphical models*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, Boca Raton, FL.

Martínez, M. E., Marshall, J. R., and Sechrest, L. (1998). Invited commentary: Factor analysis and the search for objectivity. *Am. J. Epidemiol.*, 148(1):17–19.

Ohn, I., Lin, L., and Kim, Y. (2023). A Bayesian Sparse Factor Model with Adaptive Posterior Concentration. *Bayesian Anal. (Advance Publication)*, pages 1 – 25.

Okamoto, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.*, 1:763–765.

Pearl, J. (2009). *Causality*. Cambridge University Press, Cambridge, second edition. Models, reasoning, and inference.

Pearl, J. and Verma, T. S. (1995). A theory of inferred causation. In Prawitz, D., Skyrms, B., and Westerståhl, D., editors, *Logic, Methodology and Philosophy of Science IX*, volume 134 of *Studies in Logic and the Foundations of Mathematics*, pages 789–811. Elsevier.

Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of causal inference*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA. Foundations and learning algorithms.

Reise, S. P., Waller, N. G., and Comrey, A. L. (2000). Factor analysis and scale revision. *Psychol. Assess.*, 12(3):287–297.

Robeva, E. and Seby, J.-B. (2021). Multi-trek separation in linear structural equation models. *SIAM J. Appl. Algebra Geom.*, 5(2):278–303.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., and King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *J. Educ. Res.*, 99(6):323–338.

Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., and Bengio, Y. (2021). Toward causal representation learning. *Proc. IEEE*, 109(5):612–634.

Shafarevich, I. R. (2013). *Basic Algebraic Geometry 1*. Springer Berlin Heidelberg, third edition. Varieties in projective space.

Shiers, N., Zwiernik, P., Aston, J. A. D., and Smith, J. (2016). The correlation space of Gaussian latent tree models and model selection without fitting. *Biometrika*, 103(3):531–545.

Skitovich, V. (1953). On a property of the normal distribution. *Dokl. Akad. Nauk SSSR (N.S.)*, 89:217–219.

Spearman, C. (1904). "general intelligence," objectively determined and measured. *Am. J. Psychol.*, 15(2):201–292.

Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, prediction, and search*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, second edition. With additional material by David Heckerman, Christopher Meek, Gregory F. Cooper and Thomas Richardson, A Bradford Book.

Squires, C., Seigal, A., Bhate, S. S., and Uhler, C. (2023). Linear causal disentanglement via interventions. In Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., and Scarlett, J., editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 32540–32560. PMLR.

Stanghellini, E. and Wermuth, N. (2005). On the identification of path analysis models with one hidden variable. *Biometrika*, 92(2):337–350.

Steyer, R. (2005). Analyzing individual and average causal effects via structural equation models. *Methodology*, 1(1):39–54.

Strieder, D., Freidling, T., Haffner, S., and Drton, M. (2021). Confidence in causal discovery with linear causal models. In de Campos, C. and Maathuis, M. H., editors, *Proceedings of the 37th Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pages 1217–1226. PMLR.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. *Cell*, 177(7):1888–1902.

Sturma, N. (2021). Testing algebraic constraints on statistical parameters. Master's thesis, Technische Universität München. Available at `https://mediatum.ub.tum.de/download/1601409/1601409.pdf`.

Sturma, N., Drton, M., and Leung, D. (2024). Testing many constraints in possibly irregular models using incomplete U-statistics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 86(4):987–1012.

Sturma, N., Squires, C., Drton, M., and Uhler, C. (2023). Unpaired multi-domain causal representation learning. In Oh, A., Neumann, T., Globerson, A., Saenko, K., Hardt, M., and Levine, S., editors, *Advances in Neural Information Processing Systems*, volume 36, pages 34465–34492. Curran Associates, Inc.

Sturmfels, B. and Sullivant, S. (2006). Combinatorial secant varieties. *Pure Appl. Math. Q.*, 2(3):867–891.

Sullivant, S. (2018). *Algebraic statistics*, volume 194 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI.

Sullivant, S., Talaska, K., and Draisma, J. (2010). Trek separation for Gaussian graphical models. *Ann. Statist.*, 38(3):1665–1685.

Trendafilov, N. T., Fontanella, S., and Adachi, K. (2017). Sparse exploratory factor analysis. *Psychometrika*, 82(3):778–794.

Van der Vaart, A. W. (1998). *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.

Van Der Zander, B., Textor, J., and Liskiewicz, M. (2015). Efficiently finding conditional instruments for causal inference. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 3243–3249. AAAI Press.

von Kügelgen, J., Besserve, M., Wendong, L., Gresele, L., Kekić, A., Bareinboim, E., Blei, D., and Schölkopf, B. (2023). Nonparametric identifiability of causal representations from unknown interventions. In Oh, A., Naumann, T., Globerson, A., Saenko, K., Hardt, M., and Levine, S., editors, *Advances in Neural Information Processing Systems*, volume 36, pages 48603–48638. Curran Associates, Inc.

Wang, K. and Seigal, A. (2024). Identifiability of overcomplete independent component analysis. *arXiv preprint arXiv:2401.14709*.

Weihs, L., Robinson, B., Dufresne, E., Kenkel, J., Kubjas Reginald McGee II, K., Reginald, M. I., Nguyen, N., Robeva, E., and Drton, M. (2017). Determinantal generalizations of instrumental variables. *J. Causal Inference*, 6(1):2017.0009.

Xie, F., Cai, R., Huang, B., Glymour, C., Hao, Z., and Zhang, K. (2020). Generalized independent noise condition for estimating latent variable causal graphs. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H., editors, *Advances in Neural Information Processing Systems*, volume 33, pages 14891–14902. Curran Associates, Inc.

Xie, F., Huang, B., Chen, Z., He, Y., Geng, Z., and Zhang, K. (2022). Identification of linear non-Gaussian latent hierarchical structure. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., and Sabato, S., editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 24370–24387. PMLR.

Yang, K. D., Belyaeva, A., Venkatachalapathy, S., Damodaran, K., Katcoff, A., Radhakrishnan, A., Shivashankar, G. V., and Uhler, C. (2021). Multi-domain translation between single-cell imaging and sequencing data using autoencoders. *Nat. Comm.*, 12(31):1–10.

Zeng, Y., Shimizu, S., Cai, R., Xie, F., Yamamoto, M., and Hao, Z. (2021). Causal discovery with multi-domain lingam for latent factors. In Zhou, Z.-H., editor, *Proceedings of the 30th International Joint*

Conference on Artificial Intelligence, pages 2097–2103. International Joint Conferences on Artificial Intelligence Organization.

Zwiernik, P. (2016). *Semialgebraic statistics and latent tree models*, volume 146 of *Monographs on Statistics and Applied Probability*. Chapman & Hall/CRC, Boca Raton, FL.

A Core Publications

A.1 Testing Many Constraints in Possibly Irregular Models Using Incomplete U-Statistics

Summary

In this article, we consider the problem of testing a null hypothesis defined by equality and inequality constraints on a statistical parameter. As explained in Chapter 3 of this thesis, testing such hypotheses can be challenging because the number of relevant constraints may be on the same order or even larger than the number of observed samples. Moreover, standard distributional approximations may be invalid due to irregularities in the null hypothesis. We propose a general testing methodology that aims to circumvent these difficulties. The constraints are estimated by incomplete U-statistics, and we derive critical values by Gaussian multiplier bootstrap. The main result is a Berry-Esseen type bound on the Gaussian approximation of high-dimensional incomplete U-statistic under a mild conditions on the status of degeneracy of the kernel that we call "mixed degeneracy". Importantly, the bound implies that the bootstrap approximation of incomplete U-statistics is valid when the computational budget parameter used to compute the incomplete U-statistic is of the same order as the sample size. It follows that our test controls type I error even in irregular settings. Since the bootstrap approximation covers high-dimensional settings, it also makes our testing strategy suitable for problems with many constraints.

The article is structured as follows. In the introduction, we informally explain the intuition behind the proposal of our test statistic. In Section 2, we first define the concept of mixed degeneracy. Then, we give a non-asymptotic Berry–Esseen-type bound for the high-dimensional Gaussian approximation of incomplete U-statistics. Moreover, we show that the limiting Gaussian distribution can be further approximated via a data-dependent Gaussian multiplier bootstrap, and we incorporate studentization of the incomplete U-statistic. In Section 3, we formally propose our testing methodology by defining the test statistic and showing how to derive critical values. Our results on incomplete U-statistics yield that the test is asymptotically valid and consistent even in irregular settings. In Section 4, we show that our methodology is applicable, in particular, when the constraints to be tested are polynomials. We explain a general method for constructing a suitable kernel such that the incomplete U-statistic becomes an unbiased estimator of the polynomial constraints. In Section 5, we then apply our strategy for testing the goodness-of-fit of latent tree models, which are of particular relevance in phylogenetics. In numerical experiments, we compare our strategy for different computational budget parameters with the likelihood-ratio test in terms of size and power.

Individual contributions

As the lead author of this article, I formulated the necessary concepts, developed all proofs, wrote software implementations, conducted simulation studies, and drafted the manuscript. Mathias Drton proposed to investigate test statistics based on incomplete U-statistics. Dennis Leung proposed to use the kernel in Example 1.1 of the article to test tetrad constraints with so-called "m-dependent sums" in a preliminary conference paper (Leung et al., 2016). Importantly, the contents of this article are very different from those of the preliminary conference paper. In particular, this article considers incomplete U-statistics instead of m-dependent sums, generalizes the kernel to accommodate arbitrary constraints and provides an entirely new theory. Both co-authors made helpful suggestions regarding both the content and presentation of the article during regular discussions.

The article originated from my master thesis (Sturma, 2021). In the master thesis, I empirically compared three approaches for testing many and possibly irregular polynomial constraints: Constructing the test statistics via independent sums, via m-dependent sums, and via incomplete U-statistics. The theory part in the master thesis is mainly about extending the known theory that allows for sub-Exponential kernels to sub-Weibull kernels. In contrast, the new article is only about incomplete U-statistics, and it contains rigorous proofs that mathematize the intuition that incomplete U-statistics allow for correctly calibrated statistical decisions also at irregular points. The proofs and necessary concepts were developed after completing the master thesis and were written specifically for this article. None of the proven theoretical results in the main paper appear in the master thesis, and only Appendix C in the supplement on sub-Weibull random variables exists similarly. The implementation of our testing method for the example of Gaussian latent tree models was, however, already done during the master thesis. On the other hand, the simulations were carried out specifically for this article. For my dissertation, to explicitly prevent double counting, only those parts of this article that go beyond my master thesis are to be accredited.

Journal of the Royal Statistical Society Series B: Statistical Methodology, 2024, **00**, 1–26
https://doi.org/10.1093/jrsssb/qkae022

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS
Journal of the Statistics Society
Series B
Statistical Methodology

Original Article

Testing many constraints in possibly irregular models using incomplete *U*-statistics

Nils Sturma[1], Mathias Drton[1] and Dennis Leung[2]

[1]Munich Center for Machine Learning and Department of Mathematics, School of Computation, Information and Technology, Technical University of Munich, Munich, Germany
[2]School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia

Address for correspondence: Nils Sturma, Department of Mathematics, School of Computation, Information and Technology, Technical University of Munich, Boltzmannstr. 3, 85748 Garching, Germany. Email: nils.sturma@tum.de

Abstract

We consider the problem of testing a null hypothesis defined by equality and inequality constraints on a statistical parameter. Testing such hypotheses can be challenging because the number of relevant constraints may be on the same order or even larger than the number of observed samples. Moreover, standard distributional approximations may be invalid due to irregularities in the null hypothesis. We propose a general testing methodology that aims to circumvent these difficulties. The constraints are estimated by incomplete *U*-statistics, and we derive critical values by Gaussian multiplier bootstrap. We show that the bootstrap approximation of incomplete *U*-statistics is valid for kernels that we call mixed degenerate when the number of combinations used to compute the incomplete *U*-statistic is of the same order as the sample size. It follows that our test controls type I error even in irregular settings. Furthermore, the bootstrap approximation covers high-dimensional settings making our testing strategy applicable for problems with many constraints. The methodology is applicable, in particular, when the constraints to be tested are polynomials in U-estimable parameters. As an application, we consider goodness-of-fit tests of latent-tree models for multivariate data.

Keywords: Gaussian approximation, high dimensions, incomplete *U*-statistics, latent-tree model, multiplier bootstrap, non-asymptotic bound
JEL codes: 62F03, 62R01, 62E17

1 Introduction

Let $\{P_\theta : \theta \in \Theta\}$ be a statistical model with parameter space $\Theta \subseteq \mathbb{R}^d$. Given i.i.d. samples $X_1, \ldots, X_n$ from an unknown distribution P_θ, we are interested in testing

$$H_0 : \theta \in \Theta_0 \quad \text{vs.} \quad H_1 : \theta \in \Theta \setminus \Theta_0 \tag{1.1}$$

for a subset $\Theta_0 \subseteq \Theta$. In this paper, we consider the situation where the null hypothesis Θ_0 is defined by constraints, that is,

$$\Theta_0 = \{\theta \in \Theta : \ f_j(\theta) \leq 0 \text{ for all } j = 1, \ldots, p\}, \tag{1.2}$$

where each constraint f_j is a function $f_j : \mathbb{R}^d \to \mathbb{R}$. The description of the null hypothesis also allows for equality constraints since $f_j(\theta) = 0$ can be equivalently described by $f_j(\theta) \leq 0$ and $-f_j(\theta) \leq 0$.

Received: July 18, 2022. **Revised:** January 17, 2024. **Accepted:** February 9, 2024

These types of hypotheses are an important and general class and appear in a variety of statistical problems. Our work is in particular motivated by polynomial hypotheses where each f_j belongs to the ring $\mathbb{R}[\theta_1, \ldots, \theta_d]$ of polynomials in the indeterminates $\theta_1, \ldots, \theta_d$ with real coefficients. Examples of polynomial hypotheses feature in graphical modelling (Chen et al., 2017, 2014; Shiers et al., 2016; Sullivant et al., 2010), testing causal effects (Pearl, 2009; Spirtes et al., 2000; Steyer, 2005; Strieder et al., 2021), testing sub-determinants, tetrads, pentads and more in factor analysis models (Bollen & Ting, 2000; Drton et al., 2008, 2007; Drton & Xiao, 2016; Dufour et al., 2013; Gaffke et al., 2002; Leung & Drton, 2018; Silva et al., 2006), and in constraint-based causal discovery algorithms (Claassen & Heskes, 2012; Pearl & Verma, 1995; Spirtes et al., 2000).

The standard method for dealing with testing problems like equation (1.1) is the likelihood ratio (LR) test. However, a likelihood function may be multi-modal and difficult to maximize. If the LR test is not suitable, then one might instead make use of the implicit characterization of Θ_0 given by equation (1.2). In Wald-type tests, for example, the strategy is to form estimates of the involved functions $f_1, \ldots, f_p$ and aggregate them in a test statistic. Standard Wald tests require the number of restrictions p to be smaller than or equal to the dimension d. However, in many of the above examples, this might not be the case.

Another challenge for the classical LR and Wald test is that the null hypothesis Θ_0 may contain irregular points. For example, when the hypothesis is polynomial, it may contain singularities; a rigorous definition of singularities can be found in Drton (2009, Section 4.1) or Cox et al. (2015, Section 9). At singularities the rank of the Jacobian of the constraints being tested drops and the asymptotic behaviour of the LR test and the Wald-type test can be different than at regular points, resulting in an (also asymptotically) invalid test (Drton, 2009; Drton & Xiao, 2016; Dufour et al., 2013; Gaffke et al., 2002, 1999). In practice, it is unknown whether the true parameter θ is an irregular point, and it is therefore desirable to construct a test statistic for which one can give asymptotic approximations that accommodate and remain valid in irregular settings.

In this work, we propose a testing strategy that aims to cover set-ups where the number of restrictions p can be much larger than the sample size n and where the true parameter may be an irregular point. A precise definition of regular and irregular points is given later. Our method incorporates estimating the constraints $f_1, \ldots, f_p$ by an *incomplete* U-statistic. By first considering the commonly used *complete* U-statistic, we now give intuition for how this allows for high dimensionality and irregular points. Complete U-statistics provide an efficient method for unbiased estimation of $f := (f_1, \ldots, f_p)$. We assume that $f(\theta)$ is (U-)estimable, i.e. for some integer m there exists a $\mathbb{R}^p$-valued measurable symmetric function $h(x_1, \ldots, x_m)$ such that

$$\mathbb{E}_\theta[h(X_1, \ldots, X_m)] = f(\theta) \quad \text{for all } \theta \in \Theta,$$

when $X_1, \ldots, X_m$ are i.i.d. with distribution P_θ. The U-statistic with kernel h is the average of $h(X_{i_1}, \ldots, X_{i_m})$ over all distinct m-tuples $(i_1, \ldots, i_m)$ from $\{1, \ldots, n\}$, in formulas

$$U_n = \frac{1}{|I_{n,m}|} \sum_{(i_1,\ldots,i_m)\in I_{n,m}} h(X_{i_1}, \ldots, X_{i_m}), \tag{1.3}$$

where $I_{n,m} = \{(i_1, \ldots, i_m) : 1 \le i_1 < \ldots < i_m \le n\}$. Due to the form of the null hypothesis Θ_0, it is natural to define the test statistic as the maximum of the studentized U-statistic, i.e.

$$\max_{1\le j\le p} \sqrt{n}\, U_{n,j}/\hat{\sigma}_j \tag{1.4}$$

and reject H_0 for 'large' values of it. Here, $U_{n,j}$ refers to the jth coordinate of U_n for all $j = 1, \ldots, p$ and $\hat{\sigma}_j^2$ is a 'good' estimator of the asymptotic variance of $U_{n,j}$.

Example 1.1 As a leading example we consider testing of so-called 'tetrad constraints', a problem of particular relevance in factor analysis (Bollen & Ting, 2000; Drton & Xiao, 2016; Hipp & Bollen, 2003; Leung et al., 2016;

Spirtes et al., 2000) that can be traced back to Spearman (1904) and Wishart (1928). For a given symmetric matrix $\Sigma = (\sigma_{uv})$, a tetrad is an off-diagonal 2×2 sub-determinant. An example is $f_j(\Sigma) = \sigma_{uv}\sigma_{wz} - \sigma_{uz}\sigma_{vw}$ with four different indices u, v, w, z. If an l-dimensional normal distribution $N_l(0, \Sigma)$ follows a one-factor analysis model, where it is assumed that all variables are independent conditioned on one hidden factor, then all tetrads that can be formed from the covariance matrix Σ vanish. Thus, for given i.i.d. samples $X_1, \ldots, X_n \sim N_l(0, \Sigma)$, one might be interested in testing whether all tetrads vanish simultaneously. It is easy to see that a tetrad $f_j(\Sigma) = \sigma_{uv}\sigma_{wz} - \sigma_{uz}\sigma_{vw}$ is estimable by the kernel

$$h_j(X_1, X_2) = \frac{1}{2}\{(X_{1u}X_{1v}X_{2w}X_{2z} - X_{1u}X_{1z}X_{2v}X_{2w}) + (X_{2u}X_{2v}X_{1w}X_{1z} - X_{2u}X_{2z}X_{1v}X_{1w})\}.$$

Interestingly, the resulting U-statistic corresponds to the 'plug-in' estimate, i.e. $U_{n,j} = \frac{n}{n-1} f_j(S)$, where $S = \frac{1}{n}\sum_{i=1}^n X_i X_i^\top$ is the sample covariance. The plug-in estimate is considered in previous work on testing tetrads; see for example Shiers et al. (2016).

Critical values for the test statistic (1.4) can be derived by bootstrap methods that approximate the sampling distribution of U_n. Crucial to the validity of the bootstrap is the approximation by a Gaussian distribution. Recent progress in high-dimensional central limit theory yields valid Gaussian approximation of U-statistics in settings where $p \gg n$ is allowed. In particular, Chen (2018) and Chen and Kato (2020) derive finite-sample Berry–Esseen-type bounds on the Gaussian approximation in a two-step procedure: In the first step, the centred U-statistic is approximated by the linear component in the Hoeffding decomposition (a.k.a. the Hájek projection)

$$\frac{m}{n}\sum_{i=1}^n (g(X_i) - f(\theta)), \tag{1.5}$$

where $g(x) = \mathbb{E}_\theta[h(x, X_2, \ldots, X_m)]$ and the expectation is taken with respect to the unknown distribution P_θ. In the second step, the linear term (1.5) is further approximated by a Gaussian random vector using the 'classical' central limit theorem for high-dimensional independent sums (Chernozhukov et al., 2013, 2017a). Crucially, this procedure assumes that the U-statistic is *non-degenerate*, that is, the individual variances $\sigma^2_{g,\theta,j} = \mathrm{Var}_\theta[g_j(X_1)]$ of the linear component do not vanish. Hence, the Berry–Esseen-type bound on the approximation relies on the standard assumption that the minimum $\underline{\sigma}_{g,\theta} = \min_{1\le j\le p} \sigma^2_{g,\theta,j}$ is bounded away from zero. However, the minimum $\underline{\sigma}_{g,\theta}$ depends on the unknown parameter θ, and there may be certain points in Θ_0 where we have $\underline{\sigma}_{g,\theta} = 0$.

Definition 1.2 We say that a point $\theta \in \Theta$ is *regular* with respect to the kernel h if $\underline{\sigma}_{g,\theta} = \min_{1\le j\le p} \sigma^2_{g,\theta,j} > 0$. Otherwise, we say that θ is an *irregular* point.

If θ is an irregular point, then the Gaussian approximation of the U-statistic is not valid any more. This is illustrated in Example 1.3. Even if the parameter θ is only 'close' to an irregular point, the minimum $\underline{\sigma}_{g,\theta}$ can be very small. In this case, a very large sample size may be required for the Gaussian limiting distribution to provide a good approximation of the U-statistic since convergence is not uniform and the rate depends on the minimum $\underline{\sigma}_{g,\theta}$.

Example 1.3 Recall the kernel h_j of the tetrad $f_j(\Sigma) = \sigma_{uv}\sigma_{wz} - \sigma_{uz}\sigma_{vw}$ from Example 1.1. The corresponding random variable $g_j(X_1)$ in the Hájek projection (1.5) is given by

$$g_j(X_1) = \frac{1}{2}\left\{(X_{1u}X_{1v}\sigma_{wz} - X_{1u}X_{1z}\sigma_{vw}) + (\sigma_{uv}X_{1w}X_{1z} - \sigma_{uz}X_{1v}X_{1w})\right\}. \tag{1.6}$$

By inspecting equation (1.6), we see that $g_j(X_1)$ is degenerate if $\sigma_{wz} = \sigma_{vw} = \sigma_{uv} = \sigma_{uz} = 0$ and, in general, non-degenerate if at least one of the covariances is non-zero. Thus, the covariance matrices $\Sigma = (\sigma_{uv})$ in the one-factor analysis model that have $\sigma_{wz} = \sigma_{vw} = \sigma_{uv} = \sigma_{uz} = 0$ correspond to irregular points. Moreover, when the covariance matrix is only close to irregular, then the variance $\sigma^2_{g,\theta,j}$ might already be very small.

To accommodate irregularity, we propose to use *randomized incomplete* U-*statistics* instead of the usual, complete *U*-statistic from equation (1.3). That is, for a *computational budget parameter* $N \leq |I_{n,m}|$, we randomly choose on average N indices from $I_{n,m}$. Then the incomplete *U*-statistic is defined as the sample average of $h(X_{i_1}, \ldots, X_{i_m})$ taken only over the subset of chosen indices. The test statistic is then formed as in equation (1.4) by replacing U_n with the incomplete counterpart.

The main theoretical contribution of this work is to show that the high-dimensional Gaussian approximation of incomplete *U*-statistics remains valid under *mixed degeneracy*, that is, for each index j, the variance of the Hájek projection $\sigma^2_{g,\theta,j}$ is allowed to take more or less arbitrary values, including zero. Therefore, our result covers testing hypotheses as in equation (1.2) when the underlying true parameter may be irregular or close to irregular. The setting of mixed degeneracy constitutes a further development of prior results in the literature which prove validity of the Gaussian approximation of incomplete *U*-statistics in either the fully non-degenerate case or the fully degenerate case (Chen & Kato, 2019; Song et al., 2019). That is, either *all* variances $\sigma^2_{g,\theta,j}$ are bounded away from zero or *all* of them are equal to zero. Our result is intermediate since we allow a different status of degeneracy for each index j. The crucial fact we exploit is that the asymptotic variance in the Gaussian approximation of the incomplete *U*-statistic is a weighted sum of the variance of the Hájek projection and the variance of the kernel itself. Hence, vanishing of the variance of the Hájek projection need not cause degeneracy of the asymptotic distribution. Indeed, the approximation is valid when we choose the computational budget parameter appropriately, typically of the same order as the sample size, i.e. $N = \mathcal{O}(n)$.

To derive critical values of our test statistic, we further approximate the limiting Gaussian distribution via a data-dependent Gaussian multiplier bootstrap as proposed in Chen and Kato (2019). We show that when choosing N appropriately, the bootstrap approximation remains trustworthy under mixed degeneracy and yields asymptotically valid critical values. The bootstrap is computationally feasible even for a large number of constraints p, as incomplete *U*-statistics also offer computational advantages over complete *U*-statistics. The computation of the complete *U*-statistic (1.3) requires $\mathcal{O}(n^m p)$ operations, which can be challenging for $m \geq 3$, while the incomplete *U*-statistic only requires $\mathcal{O}(Np)$ operations. We would like to highlight that these computational advantages were the main motivation for consideration of incomplete *U*-statistics in prior literature; in contrast, our work raises statistical advantages.

Example 1.4 Figure 1 shows histograms of simulated p-values for testing a large number of tetrad constraints in the one-factor analysis model when the true parameter matrix is close to an irregular point. In addition to our proposed strategy of using incomplete *U*-statistics, we include two other strategies: one that is based on complete *U*-statistics and another based on the LR test. The method using incomplete *U*-statistics yields p-values that are only slightly conservative (nearing a uniform distribution), while the other two methods fail drastically.

Our strategy can be applied for general hypotheses (1.2) as long as the functions $f_j(\theta)$ are estimable using a kernel function h. This is an advancement of Leung and Drton (2018), where only tetrads are considered. Leung and Drton (2018) also used a kernel function to estimate the tetrads, but the test statistic is an m-dependent average over the kernels instead of an incomplete *U*-statistic. The randomized incomplete *U*-statistic proposed in this work is much more flexible, supports broad application, and yields better results in simulations. Moreover, we show useful theoretical guarantees of incomplete *U*-statistics that yield an asymptotically valid test even at irregularities.

Remark 1.5 (Non-parametric set-ups). The setting we consider is formulated as pertaining to a parametric model with d-dimensional parameter space. However,

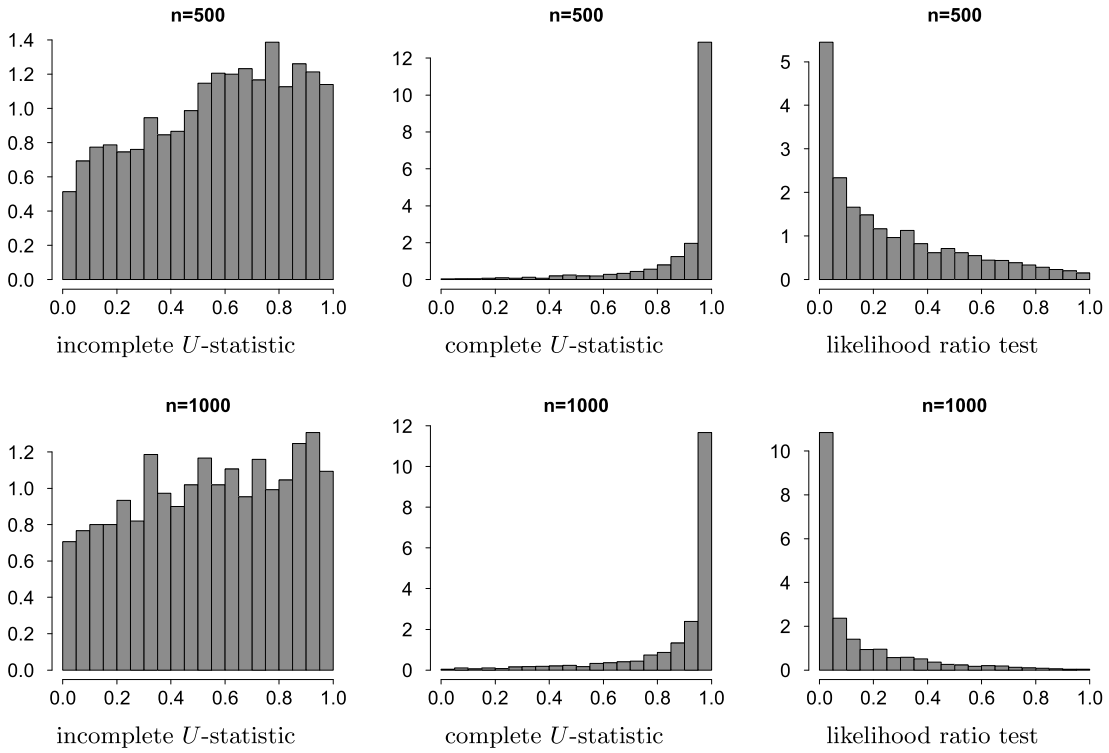


Figure 1. Histograms of 5,000 simulated p-values for simultaneously testing 2,730 tetrad constraints implied by the one-factor model with $l = 15$ observed variables. The computational budget parameter for the incomplete U-statistic is $N = 2n$ and the true covariance matrix is close to an irregular point; for exact parameter values, see Section 5, set-up (b).

our testing methodology applies without change to settings where we test constraints on a d-dimensional parameter $\theta(P)$ of the distributions in a non-parametric model. For example, one could consider testing tetrad constraints in covariance matrices of non-Gaussian distributions.

Remark 1.6 (Comparison to literature on shape restrictions). It is natural to compare our work to recent progress in the literature on shape restrictions that also considers testing equality and inequality constraints; see Chetverikov et al. (2018) for a review. In this line of work, the restricted parameter space $\Theta_0 \subseteq \Theta$ is usually considered to be infinite dimensional, so that more general parameters such as entire function classes are covered. On the other hand, to the best of our knowledge, the methods do not explicitly focus on set-ups with a large amount of constrains and possibly irregular points. For example, in Chernozhukov et al. (2023b), the authors consider a test statistic that minimizes a generalized method of moments objective function over the restricted and the whole parameter space and compares the difference. It is the main goal to study the behaviour of the test statistic in regions near the boundary of the restricted parameter space Θ_0. The different focus is reflected in the conditions the authors of Chernozhukov et al. (2023b) assume for their theoretical analysis. In particular, they assume that the Jacobian matrix of the equality constraints has full row rank in a neighbourhood of the true parameter. This implies that the maximal number of equality constraints is smaller than the dimension of the parameter space and that algebraic singularities are excluded since the Jacobian is not allowed to drop rank. Thus, the conditions do not allow for *many* equality constraints and *irregular* points. Moreover, the test statistic requires an optimization over the restricted and the whole parameter space which requires extra assumptions such as

convexity and compactness and can be difficult to implement in practice. In contrast, our method is optimization free and does not require any assumptions on the parameter space.

1.1 Organization of the paper

In Section 2, we give non-asymptotic Berry–Esseen-type bounds for high-dimensional Gaussian and bootstrap approximation of incomplete U-statistics. Importantly, the incomplete U-statistic is assumed to be of mixed degeneracy. In Section 3, we propose our testing methodology by formally defining the test statistic and showing how to derive critical values. Our results on incomplete U-statistics yield that the test is asymptotically valid and consistent even in irregular settings. In Section 4, we show that our test is applicable to polynomial hypotheses by explaining a general method for constructing a kernel h. In Section 5, we then apply our strategy for testing the goodness-of-fit of latent tree models of which the one-factor analysis model is a special case. In numerical experiments we compare our strategy with the LR test. The online supplementary material contains additional material such as all technical proofs (Appendix A, online supplementary material), additional lemmas (Appendix B, online supplementary material), properties of sub-Weibull random variables (Appendix C, online supplementary material), and additional simulation results for Gaussian latent-tree models (Appendix D, online supplementary material). Moreover, we provide a second application of our methodology in Appendix E, online supplementary material, where we test minors in two-factor analysis models.

1.2 Notation

Given $\beta \in (0, \infty)$, we define the function $\psi_\beta(x) = \exp(x^\beta) - 1$ for $x > 0$. A real random variable Y is said to be *sub-Weibull of order* β if $\|Y\|_{\psi_\beta} := \inf\{t > 0 : \mathbb{E}[\psi_\beta(|Y|/t)] \leq 1\}$ is finite. We use the usual convention that $\inf\{\emptyset\} = \infty$. For $\beta \geq 1$ we have that $\|Y\|_{\psi_\beta}$ is a norm, while for $\beta \in (0, 1)$ it is only a quasi-norm. If $\|Y\|_{\psi_1} < \infty$, then Y is called a *sub-Exponential* random variable and if $\|Y\|_{\psi_2} < \infty$, then Y is called *sub-Gaussian*. For a random element Y, let $P|_Y(\cdot)$ and $\mathbb{E}|_Y[\,\cdot\,]$ denote the conditional probability and expectation given Y. We denote a sequence of random variables $Y_i, \ldots, Y_{i'}$ by $Y_i^{i'}$ for $i \leq i'$ and for a tuple of indices $\imath = (i_1, \ldots, i_m)$ we write $Y_\imath = (Y_{i_1}, \ldots, Y_{i_m})$.

For $a, b \in \mathbb{R}$, define $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$. For $a, b \in \mathbb{R}^p$, we write $a \leq b$ if $a_j \leq b_j$ for all $j = 1, \ldots, p$, and we write $[a, b]$ for the hyper-rectangle $\prod_{j=1}^p [a_j, b_j]$. If $a \leq b$, then the hyper-rectangle $[a, b]$ is non-empty but if there is at least one index $j \in \{1, \ldots, p\}$ such that $a_j > b_j$, then the hyper-rectangle is equal to the empty set. The class of hyper-rectangles in $\mathbb{R}^p$ is denoted by $\mathbb{R}^p_{\mathrm{re}} = \{\prod_{j=1}^p [a_j, b_j] : a_j, b_j \in \mathbb{R} \cup \{-\infty, \infty\}\}$. For a vector $a \in \mathbb{R}^p$ and $r, t \in \mathbb{R}$, we write $ra + t$ for the vector in $\mathbb{R}^p$ with jth component $ra_j + t$. Finally, for a vector $a \in \mathbb{R}^p$ and two integers p_1, p_2 such that $p_1 + p_2 = p$ we write $a = (a^{(1)}, a^{(2)})$, where $a_j^{(1)} = a_j$ for all $j = 1, \ldots, p_1$ and $a_j^{(2)} = a_{p_1+j}$ for all $j = 1, \ldots, p_2$. Let $\|A\|_\infty = \max_{i,j} |a_{ij}|$ be the element-wise maximum norm of a matrix $A = (a_{ij})$.

2 Incomplete *U*-statistics under mixed degeneracy

Suppose we are given i.i.d. samples $X_1, \ldots, X_n$ from an unknown distribution P in a statistical model. For some integer m, let $h(x_1, \ldots, x_m)$ be a fixed $\mathbb{R}^p$-valued measurable function that is symmetric in its arguments. In this section, we consider the general case of inference on the mean vector $\mathbb{E}[h(X_1, \ldots, X_m)] = (\mu_1, \ldots, \mu_p)^\top = \mu$, where μ is an arbitrary estimable parameter of the underlying distribution P. In our set-up, where we want to test hypotheses characterized by constraints, the distribution P depends on θ and μ is given by the constraints $f(\theta) = (f_1(\theta), \ldots, f_p(\theta))$.

We begin with the formal definition of randomized incomplete U-statistics using similar notation as in Chen and Kato (2019) and Song et al. (2019). Let $N \leq \binom{n}{m}$ be a computational budget parameter and generate i.i.d. Bernoulli random variables $\{Z_\imath : \imath \in I_{n,m}\}$ with success probability $\rho_n = N \big/ \binom{n}{m}$. Then the incomplete U-statistics based on Bernoulli sampling is defined by

$$U'_{n,N} = \frac{1}{\hat{N}} \sum_{\imath \in I_{n,m}} Z_\imath h(X_\imath), \tag{2.1}$$

where $\hat{N} = \sum_{\iota \in I_{n,m}} Z_\iota$ is the number of successes. The variable $\hat{N}$ follows a Binomial distribution with parameters $(|I_{n,m}|, \rho_n)$. Therefore, $\mathbb{E}(\hat{N}) = |I_{n,m}|\rho_n = N$, and the incomplete U-statistic is on average a sum over N objects. Thus, we may view the computational budget N as a sparsity parameter for the incomplete U-statistic. We denote by $\sigma^2_{h,j} = \mathbb{E}[(h_j(X_1^m) - \mu_j)^2]$ the variance of the jth coordinate of the kernel and by $\sigma^2_{g,j} = \mathbb{E}[(g_j(X_1) - \mu_j)^2]$ the variance of $g_j(X_1)$; recall that the Hájek projection is given by $g(x) = \mathbb{E}[h(x, X_2, \ldots, X_m)]$. Note that $\sigma^2_{h,j}$ and $\sigma^2_{g,j}$ depend on the underlying distribution P as emphasized in the introduction; however, in the rest of the paper, we omit the explicit dependence to simplify notation.

2.1 Gaussian approximation

We will derive non-asymptotic Gaussian approximation error bounds for the incomplete U-statistic $U'_{n,N}$ that allow for mixed degenerate kernels h when choosing the computational budget parameter N appropriately. To state the formal approximation results, we assume $2 \leq m \leq \sqrt{n}$, $n \geq 4$, $p \geq 3$ and $\rho_n = N/|I_{n,m}| < 1/2$. We start by making assumptions on the moment structure of the kernel h before formally introducing mixed degeneracy. Let $\beta \in (0, 1]$ and suppose there exists a constant $D_n \geq 1$ such that:

(C1) $\mathbb{E}[|h_j(X_1^m) - \mu_j|^{2+l}] \leq \sigma^2_{h,j} D_n^l$ for all $j = 1, \ldots, p$ and $l = 1, 2$.
(C2) $\|h_j(X_1^m) - \mu_j\|_{\psi_\beta} \leq D_n$ for all $j = 1, \ldots, p$.
(C3) There exists $\underline{\sigma}^2_{\mathrm{h}} > 0$ such that $\underline{\sigma}^2_{\mathrm{h}} \leq \min_{1 \leq \mathrm{j} \leq \mathrm{p}} \sigma^2_{\mathrm{h,j}}$.

Assuming conditions similar to (C1)–(C3) is standard in high-dimensional Gaussian approximation theory. Condition (C2) assumes that the kernel h is sub-Weibull. In prior work on Gaussian approximation of high-dimensional U-statistics (Chen, 2018; Chen & Kato, 2019), the authors usually consider sub-Exponential kernels with $\beta = 1$. However, the kernel we propose in Section 4 for testing polynomial hypotheses will typically be sub-Weibull and not sub-Exponential, also see Example 4.2. We discuss important properties of sub-Weibull random variables in Appendix C, online supplementary material. Note that we allow the bound D_n to depend on the sample size n since, in the high-dimensional setting, the distribution P may depend on n. Condition (C1) is of a more technical nature and serves for clear presentation. In principle, it would be possible to omit this assumption, but the resulting error bound would be more complicated. Finally, Condition (C3) requires that the minimal variance of the individual kernels h_j is bounded away from zero, even for large p. Put differently, we only considers kernels h_j such that $h_j(X_1^m)$ is not almost surely constant. It remains to make assumptions with respect to the degeneracy of the Hájek projection, i.e. we formally define mixed degeneracy.

Definition 2.1 Let $p_1, p_2 \in \mathbb{N}_{\geq 0}$ such that $p_1 + p_2 = p$. We say that the kernel h, or also simply the incomplete U-statistic, is *mixed degenerate* for distribution P if the following two conditions are satisfied:

(C4) There exists $\underline{\sigma}^2_{\mathrm{g}^{(1)}} > 0$ such that $\underline{\sigma}^2_{\mathrm{g}^{(1)}} \leq \min_{1 \leq \mathrm{j} \leq \mathrm{p}_1} \sigma^2_{\mathrm{g,j}}$.
(C5) There exists $k > 0$ such that $\|g_j(X_1) - \mu_j\|_{\psi_\beta} \leq n^{-k} D_n$ for all $j = p_1 + 1, \ldots, p$.

In other words, mixed degeneracy of a kernel requires that each index $j = 1, \ldots, p$ either satisfies condition (C4) or (C5). By rearranging the indices, we then find that the first p_1 indices satisfy Condition (C4) and the remaining indices satisfy Condition (C5). Note that whether a kernel is mixed degenerate or not depends on the underlying distribution P from which the samples $X_1, \ldots, X_n$ are drawn.

Assuming mixed degenerate kernels is the main difference in comparison to the existing literature on Gaussian approximation of high-dimensional incomplete U-statistics (Chen & Kato, 2019; Song et al., 2019). Usually, either the non-degenerate case where $p_2 = 0$ or the fully degenerate case where $p_1 = 0$ and $\sigma^2_{g,j} = 0$ for all $j = 1, \ldots, p$ are treated. In contrast, we allow for a different status of degeneracy of the Hájek projection for each index j. In particular, Condition (C5) covers degenerate cases where $\sigma^2_{g,j} = 0$ since zero variance implies that $\|g_j(X_1) - \mu_j\|_{\psi_\beta} = 0$ almost

surely. But our assumptions even allow for more flexibility. For example, it may be the case that the variance $\sigma^2_{g,j}$ decreases with the sample size n. In this case, Condition (C5) is an assumption on the rate of convergence to degeneracy, i.e. the rate is polynomial in n.

Remark 2.2 (Discussion on mixed degeneracy). The notion of mixed degeneracy is particularly interesting for kernels h, where we do not know whether the individual components h_j are degenerate or non-degenerate, for example when the underlying distribution is unknown. If the number of constraints p does *not* grow with the sample size, then mixed degeneracy holds for any distribution P. Indeed, by letting $\underline{\sigma}^2_{g^{(1)}}$ be the minimum of the non-zero variances $\sigma^2_{g,j}$, we see that Condition (C4) is satisfied. All other indices j have $\sigma^2_{g,j} = 0$ which implies that Condition (C5) is also satisfied. However, we emphasize that mixed degeneracy is a more subtle condition if the number of constraints p is growing.

Remark 2.3 (Parametric families and irregular points). In our testing problem (1.1), we are considering a parametric family of distributions $\{P_\theta, \theta \in \Theta\}$ but the true parameter θ is unknown. In this case, the variances $\sigma^2_{g,j}$ of the individual Hájek projections depend on θ as outlined in the introduction. If the number of constraints p does *not* grow with the sample size, then mixed degeneracy holds uniformly over the whole parameter space Θ, as we have seen in Remark 2.2. However, the rate of convergence of the incomplete U-statistic depends on the minimum $\underline{\sigma}^2_{g^{(1)}}$. If $\underline{\sigma}^2_{g^{(1)}}$ is really small, the index corresponding to the minimum may already satisfy (C5) for large sample sizes n, so $\underline{\sigma}^2_{g^{(1)}}$ can in fact be chosen larger. Therefore, mixed degeneracy is also suitable for points $\theta \in \Theta$ that are close to irregular, see Corollary 2.7 for a precise statement.

In principle, it would be possible to extend our results on more general sequences γ_n converging to zero instead of n^{-k} in Condition (C5), but this would result in more involved error bounds. For simplicity, we also assume $p_1, p_2 \geq 3$, even though one could specify the bounds for arbitrary p_1 and p_2. Our last technical assumption is similar to Condition (C1) and is also necessary for the sake of clear presentation:

(C6) $\mathbb{E}[|g_j(X_1) - \mu_j|^{2+l}] \leq \sigma^2_{g,j} D_n^l$ for all $j = 1, \ldots, p$ and $l = 1, 2$.

Now, we state our main result that specifies a non-asymptotic error bound on the Gaussian approximation of incomplete U-statistics. We define $\alpha_n = n/N$, $\Gamma_h = \mathrm{Cov}[h(X_1^m)]$ and $\Gamma_g = \mathrm{Cov}[g(X_1)]$. For notational convenience, we further define the quantities

$$\omega_{n,1} = \left(\frac{m^{2/\beta} D_n^2 \log(pn)^{1+6/\beta}}{(\underline{\sigma}^2_{g^{(1)}} \wedge \underline{\sigma}^2_{h} \wedge 1)\ (n \wedge N)} \right)^{1/6}, \quad \omega_{n,2} = \frac{N^{1/2} m^2 D_n \log(pn)^{1/2+2/\beta}}{\underline{\sigma}_h n^{\min\{1/2+k,\ 5/6\}}},$$

$$\omega_{n,3} = \left(\frac{N m^2 D_n^2 \log(p)^2}{(\underline{\sigma}^2_{h} \wedge 1) n^{\min\{1+k,m\}}} \right)^{1/3}.$$

Theorem 2.4 Assume Conditions (C1)–(C6) hold. Then there is a constant $C_\beta > 0$ only depending on β such that

$$\sup_{R \in \mathbb{R}^p_{re}} |P(\sqrt{n}(U'_{n,N} - \mu) \in R) - P(Y \in R)| \leq C_\beta \{\omega_{n,1} + \omega_{n,2} + \omega_{n,3}\},$$

where $Y \sim N_p(0, m^2 \Gamma_g + \alpha_n \Gamma_h)$.

Theorem 2.4 shows that the distribution of $\sqrt{n}(U'_{n,N} - \mu)$ can be approximated by the Gaussian distribution $Y \sim N_p(0, m^2\Gamma_g + \alpha_n\Gamma_h)$ under mixed degeneracy. Since the computational budget parameter N occurs in the numerator of the bound, one has to choose it in proportion to the sample size such that the bound vanishes. In particular, we note that under the regime $N = \mathcal{O}(n)$ the bound vanishes when treating other quantities as constants.

Example 2.5 Let N be of the same order as the sample size. Then α_n can be viewed as a constant and each coordinate of Y is asymptotically non-degenerate. In this case, when we assume $k \geq 1/3$ and treat m, $\underline{\sigma}_{g^{(1)}}$, $\underline{\sigma}_h$ and D_n as fixed constants, the bound $C_\beta\{\omega_{n,1} + \omega_{n,2} + \omega_{n,3}\}$ vanishes asymptotically under Conditions (C1)–(C6) if the dimension p satisfies $\log(pn)^{3/2+6/\beta} = \mathcal{O}(n)$. On the other hand, one can also choose $\mathrm{N} = \mathrm{n}^{1+\varepsilon}$ for small enough $\varepsilon > 0$. Then, $\alpha_n \to 0$ *as* $\mathrm{n} \to \infty$ and the distribution of Y may become asymptotically degenerate. Nevertheless, the bound still vanishes if we fix all other quantities as before and if the dimension p satisfies $\log(pn)^{\frac{3/2+6/\beta}{1/3-\varepsilon/2}} = \mathcal{O}(n)$.

Remark 2.6 (Order of the kernel). The Berry–Esseen-type bound in Theorem 2.4 depends explicitly on the order of the kernel h. In particular, the result also allows for kernels of diverging order, i.e. increasing m. However, larger m imply worse performance of the Gaussian approximation in terms of the required sample size. Also, if m is increasing with n one has to be careful in choosing the computational budget N; it has to be chosen smaller to achieve convergence.

The proof of Theorem 2.4 relies on the seminal papers of Chernozhukov et al. (2013) and Chernozhukov et al. (2017a) on Gaussian approximation of high-dimensional independent sums, and it extends the results of Chen and Kato (2019) and Song et al. (2019) on incomplete U-statistics to the mixed degenerate case. Obtaining sharper bounds for the high-dimensional approximation of independent sums than in the original papers is an ongoing area of research, see for example Fang and Koike (2021), Chernozhukov et al. (2022), Lopes (2022) and Chernozhukov et al. (2023a).

The bound in Theorem 2.4 is stated as general as possible. Importantly, it entails that the incomplete U-statistic can be approximated by $N_p(0, m^2\Gamma_g + \alpha_n\Gamma_h)$ even in irregular set-ups of our testing problem (1.1) as long as $\underline{\sigma}^2_{g^{(1)}} > 0$. However, it might be difficult to read off precise rates for the speed of convergence in the *close to irregular* scenarios. For large n, it is possible that there are close to irregular points such that some indices $j \in \{1, \ldots, p\}$ only satisfy mixed degeneracy if one chooses $\underline{\sigma}^2_{g^{(1)}}$ relatively small or k small. On the other hand, under further distributional assumptions, we manage to improve the bound in Theorem 2.4 so that the speed of convergence is completely independent to the irregularity status of the hypothesis.

Corollary 2.7 Assume that $X_1, \ldots, X_n$ are i.i.d. samples of a Gaussian distribution and assume that each individual kernel $h_1, \ldots, h_p$ is a non-constant polynomial of degree at most $2s$. Suppose that $\mu = \mathbb{E}[h(X_1^m)] = 0$ and that there exists $\underline{\sigma}^2_h > 0$ such that $\underline{\sigma}^2_h \leq \min_{1\leq j\leq p} \sigma^2_{h,j}$. Then there exist $\beta \in (0, 1]$ and $D_n \geq 1$ such that (C1), (C2) and (C6) are satisfied and the kernel h is mixed degenerate. If, additionally, $n \leq N \leq Cn$ for some constant $C > 0$, then it holds that

$$\sup_{R\in\mathbb{R}^p_{\mathrm{re}}} |P(\sqrt{n}U'_{n,N} \in R) - P(Y \in R)| \leq C_{s,m} \frac{(\overline{\sigma}^2_h \vee 1)\log(pn)^{1/2+2s}}{(\underline{\sigma}^2_h \wedge 1)n^{1/9}},$$

where $Y \sim N_p(0, m^2\Gamma_g + \alpha_n\Gamma_h)$, $\overline{\sigma}^2_h = \max_{1\leq j\leq p} \sigma^2_{h,j}$ and $C_{s,m} > 0$ is a constant only depending on s and m.

In the bound in Corollary 2.7, there is no $\underline{\sigma}^2_{g^{(1)}}$ and no k showing up anymore, so the convergent speed shows no dependence on how close to irregular the points in the null hypothesis really are. Even though the bound might not be optimal, it completely mathematizes our intuition why incomplete U-statistics are a good choice for a test statistic to guard against irregular points.

Example 2.8 Each coordinate h_j of the proposed kernel in Example 1.1 is a polynomial of degree 4 in Gaussian variables, which implies that Corollary 2.7 is applicable with $s = 2$. See Section 4 for a general method to construct polynomial kernels.

2.2 Bootstrap approximation

Since the covariance matrix $m^2\Gamma_g + \alpha_n\Gamma_h$ of the approximating Gaussian distribution in Theorem 2.4 is typically unknown in statistical applications, we apply a Gaussian multiplier bootstrap. The procedure is exactly the same as in Chen and Kato (2019) and Song et al. (2019) but their error bounds on the approximation require non-degeneracy, that is, there is a constant $c > 0$ such that $\min_{1\le j\le p}\sigma^2_{g,j} \ge c$. Under mixed degeneracy, this is no longer the case. However, when choosing the computational budget appropriately, we prove that the bootstrap approximation still holds.

The bootstrap is based on the fact that the random vector $Y \sim N_p(0, m^2\Gamma_g + \alpha_n\Gamma_h)$ is a weighted sum of the two independent random vectors $Y_g \sim N_p(0, \Gamma_g)$ and $Y_h \sim N_p(0, \Gamma_h)$, i.e. $Y = mY_g + \sqrt{\alpha_n}Y_h$. Let $\mathcal{D}_n = \{X_1, \ldots, X_n\} \cup \{Z_\iota : \iota \in I_{n,m}\}$ be the data involved in the definition of the incomplete U-statistic $U'_{n,N}$. We will construct data-dependent random vectors $U^{\#}_{n_1,g}$ and $U^{\#}_{n,h}$ such that, given the data $\mathcal{D}_n$, both vectors are independent and approximate Y_g and Y_h.

To approximate the distribution of Y_h take a collection $\{\xi'_\iota : \iota \in I_{n,m}\}$ of independent $N(0, 1)$ random variables that are also independent of $\mathcal{D}_n$. Define the multiplier bootstrap

$$U^{\#}_{n,h} = \frac{1}{\sqrt{\widehat{N}}}\sum_{\iota\in I_{n,m}} \xi'_\iota\sqrt{Z_\iota}(h(X_\iota) - U'_{n,N}) \tag{2.2}$$

and observe that, conditioned on the data $\mathcal{D}_n$, the distribution of $U^{\#}_{n,h}$ is Gaussian with mean zero and covariance matrix $\widehat{N}^{-1}\sum_{\iota\in I_{n,m}} Z_\iota(h(X_\iota) - U'_{n,N})(h(X_\iota) - U'_{n,N})^\top$. Intuitively, this covariance matrix should be a good estimator of the true covariance matrix Γ_h and therefore the distribution of $U^{\#}_{n,h}$ should be 'close' to the distribution of Y_h.

Approximating $Y_g \sim N_p(0, \Gamma_g)$ is more involved since the Hájek projection (1.5) is in general unknown. Thus, we first construct estimates G_{i_1} of $g(X_{i_1})$ for each i_1 in a chosen subset $S_1 \subseteq \{1, \ldots, n\}$ with cardinality $n_1 = |S_1|$. Then, we consider the multiplier bootstrap distribution

$$U^{\#}_{n_1,g} = \frac{1}{\sqrt{n_1}}\sum_{i_1\in S_1} \xi_{i_1}(G_{i_1} - \overline{G}), \tag{2.3}$$

where $\{\xi_{i_1} : i_1 \in S_1\}$ is a collection of independent $N(0, 1)$ random variables that is independent of $\mathcal{D}_n$ and $\{\xi'_\iota : \iota \in I_{n,m}\}$. Here, $\overline{G} = n_1^{-1}\sum_{i_1\in S_1} G_{i_1}$ denotes the average of the constructed estimates. The exact form of the estimates G_{i_1} is specified later. Similar as above, the distribution of $U^{\#}_{n_1,g}$ given the data $\mathcal{D}_n$ should be 'close' to the distribution of Y_g. Combining $U^{\#}_{n,h}$ and $U^{\#}_{n_1,g}$, we obtain the multiplier bootstrap $U^{\#}_{n,n_1} = mU^{\#}_{n_1,g} + \sqrt{\alpha_n}U^{\#}_{n,h}$. It approximates the distribution of Y even under mixed degeneracy, as we verify in our next Lemma. The approximation of $U^{\#}_{n,n_1}$ will depend on the quality of the estimator G_{i_1} that we measure by the quantity

$$\widehat{\Delta}_{g,1} = \max_{1\le j\le p}\frac{1}{n_1}\sum_{i_1\in S_1}(G_{i_1,j} - g_j(X_{i_1}))^2. \tag{2.4}$$

Moreover, the approximation depends on conditions involving the quantity

$$A_n := A_n(\underline{\sigma}_h, m, \beta, N) = \frac{m^{\max\{2/\beta,4\}}}{\underline{\sigma}_h^4 \wedge 1}\max\{(N/n)^2, 1\},$$

which is required to be small. In particular, if m does not grow with n and the computational budget N is chosen appropriately, then Assumption (C3) ensures that A_n does not inflate.

Lemma 2.9 Assume the Conditions (C1)–(C3) hold. If

$$\frac{A_n D_n^4 \log(pn)^{3+4/\beta}}{n_1 \wedge N} \leq C_1 n^{-\zeta_1} \tag{2.5}$$

and

$$P\left(A_n D_n^2 \widehat{\Delta}_{g,1} \log(p)^4 > C_1 n^{-\zeta_2}\right) \leq \frac{C_1}{n} \tag{2.6}$$

for some constants $C_1 > 0$ and $\zeta_1, \zeta_2 \in (0, 1)$, then there exists a constant $C > 0$ depending only on β, ζ_1 and C_1 such that with probability at least $1 - C/n$,

$$\sup_{R \in \mathbb{R}^p_{\mathrm{re}}} \left| P|_{\mathcal{D}_n}(U^{\#}_{n,n_1} \in R) - P(Y \in R) \right| \leq C n^{-(\zeta_1 \wedge \zeta_2)/6}.$$

Note that Lemma 2.9 formally does not require mixed degeneracy, that is, it is completely independent of the status of degeneracy of the kernel. Now, we specify G_{i_1} to be a special case of the divide-and-conquer estimator $\hat{g}_{i_1}$ from Chen and Kato (2019) and Song et al. (2019) that is defined as follows. For each index $i_1 \in S_1$, we partition the remaining indices, $\{1, \ldots, n\} \setminus \{i_1\}$, into disjoint subsets $\{S^{(i_1)}_{2,k}, k = 1, \ldots, K\}$, each of size $m - 1$, where $K = \lfloor (n-1)/(m-1) \rfloor$. Then, we define for each $i_1 \in S_1$ the estimator

$$\hat{g}_{i_1} = \frac{1}{K} \sum_{k=1}^{K} h(X_{i_1}, X_{S^{(i_1)}_{2,k}}). \tag{2.7}$$

Thus, from now on, we again refer by $U^{\#}_{n_1,g}$ to the statistics defined in equation (2.3) but with the specialized estimator $G_{i_1} = \hat{g}_{i_1}$ as defined in equation (2.7). The next theorem builds on Lemma 2.9 and verifies that the bootstrap approximation is valid for this specialized estimator.

Theorem 2.10 Assume the Conditions (C1)–(C3) hold and

$$\frac{A_n D_n^4 \log(pn)^{3+4/\beta}}{n_1 \wedge N} \leq C_1 n^{-\zeta} \tag{2.8}$$

for some constants $C_1 > 0$, $\zeta \in (0, 1)$. Then, for any $\nu \in (\max\{7/6, 1/\zeta\}, \infty)$, there exists a constant $C > 0$ depending only on β, ν, ζ and C_1 such that with probability at least $1 - C/n$,

$$\sup_{R \in \mathbb{R}^p_{\mathrm{re}}} \left| P|_{\mathcal{D}_n}(U^{\#}_{n,n_1} \in R) - P(Y \in R) \right| \leq C n^{-(\zeta - 1/\nu)/6}.$$

Theorem 2.10 says that we can approximate the asymptotic Gaussian distribution on the hyperrectangles via the multiplier bootstrap $U^{\#}_{n,n_1} = m U^{\#}_{n_1,g} + \sqrt{\alpha_n} U^{\#}_{n,b}$. As long as the computational budget N is chosen appropriately, for example $N = \mathcal{O}(n)$, this holds independently of the status of degeneracy since we do not require Conditions (C4)–(C6). Crucially, we are able to simulate the distribution of $U^{\#}_{n,n_1}$ given the data by generating independent $N(0, 1)$ random variables. The complexity of the bootstrap procedure, and therefore the complexity of our testing

methodology, is discussed in Remark 3.2. In a corollary, we combine the Gaussian approximation with the bootstrap approximation.

Corollary 2.11 Assume the Conditions (C1)–(C6) hold. Further assume that for some constants $C_1 > 0$, $\zeta \in (0, 1)$, Condition (2.8) holds and $\omega_{n,1} + \omega_{n,2} + \omega_{n,3} \leq C_1 n^{-\zeta/7}$. Then there exists a constant $C > 0$ depending only on β, ζ and C_1 such that with probability at least $1 - C/n$,

$$\sup_{R \in \mathbb{R}^p_{\text{re}}} \left| P(\sqrt{n}(U'_{n,N} - \mu) \in R) - P|_{\mathcal{D}_n}(U^{\#}_{n,n_1} \in R) \right| \leq C n^{-\zeta/7}.$$

Proof. This follows from Theorems 2.4 and 2.10 with $v = 7/\zeta$. □

2.3 Studentization

Often, the approximate variances of the coordinates of $U'_{n,N}$ are heterogeneous, and it is therefore desirable to studentize the incomplete U-statistic. For $j = 1, \ldots, p$, we denote by $\sigma_j^2 = m^2 \sigma_{g,j}^2 + \alpha_n \sigma_{h,j}^2$ the variance of the jth coordinate of the approximating Gaussian random vector Y, that is, σ_j^2 is equal to the diagonal element $m^2 \Gamma_{g,jj} + \alpha_n \Gamma_{h,jj}$ of the approximating covariance matrix. In line with this, we define the empirical variances $\hat{\sigma}_j^2 = m^2 \hat{\sigma}_{g,j}^2 + \alpha_n \hat{\sigma}_{h,j}^2$ to be the diagonal elements of the conditional covariance matrix of the bootstrap distribution $U^{\#}_{n,n_1}$ given the data. Therefore, $\hat{\sigma}_{g,j}^2$ and $\hat{\sigma}_{h,j}^2$ are given by

$$\hat{\sigma}_{g,j}^2 = \frac{1}{n_1} \sum_{i_1 \in S_1} (\hat{g}_{i_1,j} - \overline{g}_j)^2 \quad \text{and} \quad \hat{\sigma}_{h,j}^2 = \frac{1}{\widehat{N}} \sum_{\iota \in I_{n,m}} Z_\iota (h_j(X_\iota) - U'_{n,N,j})^2.$$

Moreover, we define a $p \times p$ diagonal matrix $\widehat{\Lambda}$ with diagonal elements $\widehat{\Lambda}_{jj} = m^2 \hat{\sigma}_{g,j}^2 + \alpha_n \hat{\sigma}_{h,j}^2$ for all $j = 1, \ldots, p$.

Corollary 2.12 Assume the conditions in Corollary 2.11. Then there exists a constant $C > 0$ depending only on β, ζ and C_1 such that with probability at least $1 - C/n$,

$$\sup_{R \in \mathbb{R}^p_{\text{re}}} \left| P(\sqrt{n}\, \widehat{\Lambda}^{-1/2}(U'_{n,N} - \mu) \in R) - P|_{\mathcal{D}_n}(\widehat{\Lambda}^{-1/2} U^{\#}_{n,n_1} \in R) \right| \leq C n^{-\zeta/7}.$$

Corollary 2.12 allows us to construct a test for hypotheses of the form (1.2) that asymptotically controls type I error and has power against alternatives outside a small neighbourhood of the null hypothesis.

3 Testing methodology

In this section, we propose our test based on incomplete U-statistics. Recall that $X_1, \ldots, X_n$ are i.i.d. samples from a distribution P_θ with parameter $\theta \in \Theta \subseteq \mathbb{R}^d$.

3.1 Test statistic

We want to test null hypotheses $\Theta_0 \subseteq \Theta$ defined by constraints $f(\theta) = (f_1(\theta), \ldots, f_p(\theta))$ as specified in equation (1.2). For now, we assume that an $\mathbb{R}^p$-valued, measurable and symmetric function $h(x_1, \ldots, x_m)$ exists such that $\mathbb{E}[h(X_1^m)] = f(\theta)$. For the case of polynomial hypotheses we show a general construction of kernels in Section 4. We define the test statistic $\mathcal{T}$ to be the maximum

of a studentized incomplete U-statistic, that is,

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

where $\hat{\sigma}_j^2$ is the empirical approximate variance of the jth coordinate of the incomplete U-statistic; recall the definition in Section 2.3. Large values of $\mathcal{T}$ indicate that the null hypothesis is violated and thus it is natural to reject H_0 when $\mathcal{T}$ exceeds a certain critical value.

3.2 Critical value

The idea for construction of a critical value relies on the observation

$$\mathcal{T} \le \max_{1\le j\le p} \sqrt{n}(U'_{n,N,j} - f_j(\theta))/\hat{\sigma}_j, \tag{3.1}$$

whenever θ is a point in the null hypothesis Θ_0. Hence, to make the test size less or equal than α, it is enough to choose the critical value as the $(1-\alpha)$-quantile of the distribution of $\max_{1\le j\le p} \sqrt{n}(U'_{n,N,j} - f_j(\theta))/\hat{\sigma}_j$, which is a centred version of our test statistic $\mathcal{T}$. Since the distribution of $\max_{1\le j\le p} \sqrt{n}(U'_{n,N,j} - f_j(\theta))/\hat{\sigma}_j$ is unknown, we approximate it using the Gaussian multiplier bootstrap introduced in Section 2. The distribution function $P(\max_{1\le j\le p} \sqrt{n}\ (U'_{n,N,j} - f_j(\theta))/\hat{\sigma}_j \le t)$ for $t \in \mathbb{R}$ corresponds to the probabilities $P(\sqrt{n}\ \widehat{\Lambda}^{-1/2}(U'_{n,N} - f(\theta)) \in R_t)$, where $R_t = \{x \in \mathbb{R}^p : -\infty \le x_j \le t \text{ for all } j = 1, \ldots, p\}$ are t-dependent hyper-rectangles. Hence, by Corollary 2.12, the maximum $\max_{1\le j\le p} \sqrt{n}(U'_{n,N,j} - f_j(\theta))/\hat{\sigma}_j$ can be approximated by the maximum of the studentized Gaussian multiplier statistic $W := \max_{1\le j\le p} U^{\#}_{n,n_1,j}/\hat{\sigma}_j$. In particular, the quantiles of W approximate the quantiles of $\max_{1\le j\le p} \sqrt{n}\ (U'_{n,N,j} - f_j(\theta))/\hat{\sigma}_j$ as we verify in the next corollary. For $\alpha \in (0, 1)$ we denote by $c_W(1-\alpha)$ the conditional $(1-\alpha)$-quantile of W given the data $\mathcal{D}_n$.

Corollary 3.1 Assume the conditions in Corollary 2.11. Then there exists a constant $C > 0$ depending only on β, ζ and C_1 such that

$$\sup_{\alpha\in(0,1)} \left| P\left(\max_{1\le j\le p} \sqrt{n}\ (U'_{n,N,j} - f_j(\theta))/\hat{\sigma}_j > c_W(1-\alpha) \right) - \alpha \right| \le C n^{-\zeta/7}.$$

Corollary 3.1 says that, under mild regularity conditions involving mixed degeneracy, using $c_W(1-\alpha)$ as a critical value gives an asymptotically valid test, that is, it asymptotically controls type I error for the significance level α. When the null hypothesis is only defined by equality constraints $f_j(\theta) = 0$, then we have equality in equation (3.1) and Corollary 3.1 implies that our test has asymptotic type I error exactly equal to the chosen level α. When we have $f_j(\theta) < 0$ for certain indices $j = 1, \ldots, p$, then the type I error is smaller or equal to the chosen level. As discussed in Remark 2.3, mixed degeneracy also accommodates irregular settings. The computational budget parameter has to be chosen appropriately such that the conditions in Corollary 2.11 are satisfied, typically $N = \mathcal{O}(n)$ is a good choice as we elaborate in Remark 3.3. In practice, we use the empirical version of $c_W(1-\alpha)$ as a critical value. It is obtained in the following procedure:

(i) Generate many, say $A = 1{,}000$, sets of standard normal random variables $\{\xi'_\iota : \iota \in I_{n,m}\} \cup \{\xi_{i_1} : i_1 \in S_1\}$.
(ii) Evaluate W for each of these A sets.
(iii) Take $\hat{c}_W(1-\alpha)$ to be the $(1-\alpha)$-quantile of the resulting A numbers.
(iv) Reject H_0 if $\mathcal{T} > \hat{c}_W(1-\alpha)$.

Remark 3.2 (Practical considerations and complexity). There is no need to generate $|I_{n,m}| \approx n^m$ Bernoulli random variables to construct the incomplete U-statistic $U'_{n,N}$, nor to generate the same amount of standard normal random variables ξ'_ι. As explained in Chen and Kato (2019, Section 2.1), one

can equivalently generate $\hat{N} \sim \text{Bin}(|I_{n,m}|, \rho_n)$ once and then choose indices $\iota_1, \ldots, \iota_{\hat{N}}$ without replacement from $I_{n,m}$. Then we can compute the incomplete U-statistic as

$$U'_{n,N} = \frac{1}{\hat{N}} \sum_{i=1}^{\hat{N}} h(X_{\iota_i}).$$

Similarly, one only needs to generate $\hat{N}$ versions of ξ'_ι to construct $U^{\#}_{n,h}$. As another practical guidance, we suggest to compute $n_1 = |S_1| = n$ divide-and-conquer estimators. Choosing n_1 smaller only decreases the accuracy of the bootstrap, which is only useful if the computational cost is otherwise too high. With $n_1 = n$, the overall computational cost of the proposed procedure is $\mathcal{O}(n^2 p + A(N+n)p)$ as the estimation step for g can be done outside the bootstrap.

Remark 3.3 (Computational budget parameter N). If one is certain that the underlying parameter θ is regular and *all* random variables $g_j(X_1)$, $j = 1, \ldots, p$ are asymptotically non-degenerate, then it may be appealing to choose the computational budget $0 < N < \binom{n}{m}$ arbitrarily since the bootstrap approximation still holds as shown in Chen and Kato (2019, Theorem 3.1) and further refined in Song et al. (2019, Theorem 2.4). However, the rate of convergence depends on $\underline{\sigma}^2_{g,\theta} = \min_{1 \le j \le p} \sigma^2_{g,\theta,j}$ as outlined in the introduction. Thus, if θ is close to an irregular point, where $\underline{\sigma}^2_{g,\theta}$ is small, the rate of convergence can be very slow. In contrast, a lower computational budget parameter N may imply convergence with a reasonable rate, even when the parameter θ is close to irregular. If mixed degeneracy is satisfied, the rate of convergence only depends on $\underline{\sigma}^2_{g^{(1)}}$, which is a fixed constant large enough to achieve fast convergence; recall Remark 2.3 and Corollary 2.7.

On the other hand, consider an underlying parameter θ such that *all* variances $\sigma^2_{g,\theta,j}$ of the Hájek projection are zero, which in particular means that θ is an irregular point. Then, the incomplete U-statistic is fully degenerate and a bootstrap approximation also holds true for larger N (Chen & Kato, 2019, Theorem 3.3). However, the limiting Gaussian distribution becomes $N_p(0, \Gamma_h)$ under a suitable scaling of the incomplete U-statistic. Choosing the computational budget N lower allows for computing the bootstrap based on the Gaussian approximation $N_p(0, m^2\Gamma_g + \alpha_n \Gamma_h)$, regardless of any potential degeneracy. In particular, we can compute valid critical values without knowing a priori whether or not a point is irregular. Moreover, the degenerate approximation to $N_p(0, \Gamma_h)$ only holds if *all* variances $\sigma^2_{g,\theta,j}$ are zero and not for more general irregular points where only some of them are zero.

To summarize, our approximation results of Section 2 are very useful if it is unknown whether or not the true point is irregular or close to irregular. Mixed degeneracy allows for such set-ups when choosing the computational budget N appropriately. In particular, we recommend to choose the computational budget parameter N of the same order as the sample size n to guard against irregularities. However, this still allows some flexibility in practice. In Section 5, we compare test size and power for different choices of N in numerical experiments.

Remark 3.4 (High dimensionality). We emphasize that bootstrap approximation holds in settings where the number of polynomials p may be much larger than the sample size n, i.e. p may be as large as $\exp(n^c)$ for a constant $c > 0$.

Thus, we can test a very large number of polynomial restrictions simultaneously. Moreover, we do not require any restriction on the correlation structure among the polynomials.

3.3 Power

The following result pertains to the power of the proposed test.

Proposition 3.5 Assume the conditions in Corollary 2.11, and let $\alpha \in (0, 1/2)$. Then there exists a constant $C > 0$ depending only on β, ζ and C_1 such that for every $\epsilon > 0$, whenever

$$\max_{1\le j\le p} (f_j(\theta)/\sigma_j) \ge (1+\varepsilon)\Big(1 + Cn^{-3\zeta/7}\log(p)^{-2}\Big) \times \frac{\sqrt{2\log(p)} + \sqrt{2\log(1/\alpha)}}{\sqrt{n}}, \tag{3.2}$$

we have

$$P(\mathcal{T} > c_{\mathbb{W}}(1-\alpha)) \ge 1 - \exp\Big(-\frac{1}{C}\varepsilon^2\log(p/\alpha)\Big) - Cn^{-\zeta/7}.$$

Proposition 3.5 shows that our test is consistent against all alternatives where at least one constraint, normalized by its approximate variance, is violated by a small amount. In other words, the test is consistent against all alternatives excluding the ones in a small neighbourhood of the null set Θ_0. The size of the neighbourhood shrinks with rate $\sqrt{\log(p)/n}$ as long as $p \to \infty$ as $n \to \infty$. A similar result on power is derived in Chernozhukov et al. (2019) for the special case on independent sums, and we refer to their discussion on the rate the neighbourhood converges to zero.

4 Polynomial hypotheses

In this section, we assume that the constraints defining the null hypothesis Θ_0 in equation (1.2) are polynomial. That is, each constraint $f_j \in \mathbb{R}[\theta_1, \ldots, \theta_d]$ is a polynomial in the indeterminates $\theta_1, \ldots, \theta_d$. We propose a general procedure to find a kernel h such that $h_j(X_1^m)$ is an unbiased estimator of $f_j(\theta)$. For now, let $f \in \mathbb{R}[\theta_1, \ldots, \theta_d]$ be a *single* polynomial of total degree s and write

$$f(\theta) = a_0 + \sum_{r=1}^{s} \sum_{\substack{(i_1,\ldots,i_r)\\ i_l\in\{1,\ldots,d\}}} a_{(i_1,\ldots,i_r)}\theta_{i_1}\cdots\theta_{i_r}$$

with $a_{(i_1,\ldots,i_r)} \in \mathbb{R}$ for all multi-indices $(i_1, \ldots, i_r)$. Note that this notation of multivariate polynomials is somewhat inefficient in comparison to the usual multi-index notation since there may appear various indices repeatedly in $(i_1, \ldots, i_r)$. However, the representation is useful to define an estimator of the polynomial. We construct $h(X_1^m)$ by the following three steps:

1. For a fixed integer $\eta \ge 1$, find $\mathbb{R}$-valued functions $\widehat{\theta}_i$ such that $\widehat{\theta}_i(X_1^\eta)$ is an unbiased estimator of θ_i for all $i = 1, \ldots, d$.
2. Let $m = \eta s$ be the order of the kernel, and define the $\mathbb{R}$-valued function $\check{h}$ via

$$x_1^m \mapsto a_0 + \sum_{r=1}^{s} \sum_{\substack{(i_1,\ldots,i_r)\\ i_l\in\{1,\ldots,d\}}} a_{(i_1,\ldots,i_r)}\widehat{\theta}_{i_1}(x_1^\eta)\widehat{\theta}_{i_2}(x_{\eta+1}^{2\eta})\cdots\widehat{\theta}_{i_r}(x_{(r-1)\eta+1}^{r\eta}),$$

where $x_k^l = (x_k, \ldots, x_l)$ for $k < l$. Note that $\check{h}(X_1^m)$ is an unbiased estimator of $f(\theta)$ by the linearity of the expectation and the independence of the samples $X_1, \ldots, X_m$.

3. To get a symmetric kernel h, average over all permutations $\pi \in S_m$ of the set $\{1, \ldots, m\}$, that is,

$$h(x_1, \ldots, x_m) = \frac{1}{m!} \sum_{\pi \in S_m} \check{h}(x_{\pi(1)}, \ldots, x_{\pi(m)}).$$

The construction works for any polynomial as long as we can construct an unbiased estimator of the parameter θ that only involves a small number η of samples. In this case polynomial hypotheses are estimable and our proposed testing methodology is applicable.

Example 4.1 The kernel for estimating the tetrad $f(\Sigma) = \sigma_{uv}\sigma_{wz} - \sigma_{uz}\sigma_{vw}$ in Example 1.1 is also constructed by steps 1) to 3). The total degree of f is $s = 2$ and to estimate an entry σ_{uv} in the covariance matrix Σ we only need $\eta = 1$ sample, i.e. an unbiased estimator of σ_{uv} is given by $\widehat{\sigma}_{uv}(X_1) = X_{1u}X_{1v}$. Thus, the order of the kernel is $m = 2$ and we obtain

$$\check{h}(X_1, X_2) = X_{1u}X_{1v}X_{2w}X_{2z} - X_{1u}X_{1z}X_{2v}X_{2w}.$$

The symmetric kernel is given by

$$\begin{aligned} h(X_1, X_2) = \frac{1}{2}\{&(X_{1u}X_{1v}X_{2w}X_{2z} - X_{1u}X_{1z}X_{2v}X_{2w}) \\ &+ (X_{2u}X_{2v}X_{1w}X_{1z} - X_{2u}X_{2z}X_{1v}X_{1w})\}. \end{aligned}$$

The bootstrap approximation established in Section 2 requires that the individual estimators $h_j(X_1^m)$ are sub-Weibull of order $\beta > 0$ for all $j = 1, \ldots, p$; recall Condition (C2). If one is able to check that all estimators $\widehat{\theta}_i(X_1^\eta)$, $i = 1, \ldots, d$ are sub-Weibull of order γ, then we obtain by Lemma C.3, online supplementary material that the estimator $h_j(X_1, \ldots, X_m)$ is sub-Weibull of order $\beta = \gamma/s$, where s is the total degree of the polynomial $f_j(\theta)$. For estimating tetrads, this is illustrated in the next example.

Example 4.2 For a Gaussian random vector $X_i \sim N_l(0, \Sigma)$, it is easy to check that each component X_{iu}, $u = 1, \ldots, l$ is sub-Gaussian with $\|X_{iu}\|_{\psi_2} \leq \sqrt{8\sigma_{uu}/3}$. Therefore, $\widehat{\sigma}_{uv}(X_1) = X_{1u}X_{1v}$ is sub-exponential by Lemma C.2, online supplementary material and hence the kernel $h(X_1, X_2)$ for estimating a tetrad in Example 4.1 is sub-Weibull of order $\beta = 1/2$.

Recall our testing problem (1.2) where the null hypothesis $\Theta_0 \subseteq \Theta$ is defined by polynomial constraints. Suppose we have a kernel h such that every point $\theta \in \Theta$ is regular with respect to h and $\underline{\sigma}_{g,\theta} = \min_{1 \leq j \leq p} \sigma^2_{g,\theta,j}$ is not too small, that is, every point is far away from irregular. Then it is favourable to consider critical values based on the Gaussian approximation for non-degenerate incomplete U-statistics as discussed in Remark 3.3. On the other hand, if all points are irregular with *every* individual Hájek projection being degenerate, then one should consider critical values based on the limiting distribution of degenerate incomplete U-statistics. For the proposed kernel, we will show in our next result, that parameter spaces typically contain a measure zero subset of irregular points.

Proposition 4.3 Let $f \in \mathbb{R}[\theta_1, \ldots, \theta_d]$ be a polynomial of total degree $s \geq 2$ and define the set of indices $D(f) = \{j \in [d] : \theta_j \text{ appears in } f\}$. Further, define $\hat{\theta}_{i,l}(x) = \mathbb{E}[\hat{\theta}_i(X_1, \ldots, X_{l-1}, x, X_{l+1}, \ldots, X_\eta)]$ for all $i = 1, \ldots, d$ and $l = 1, \ldots, \eta$ and denote the random vector

$$\hat{\theta}_f(X_1) = \left(\sum_{l=1}^{\eta} \hat{\theta}_{i,l}(X_1) \right)_{i \in D(f)}.$$

If the covariance matrix $Cov[\hat{\theta}_f(X_1)]$ is positive definite, then the irregular points $\theta \in \Theta \subseteq \mathbb{R}^d$ with respect to the kernel h form a measure zero set with respect to the Lebesgue measure on $\mathbb{R}^d$.

Proposition 4.3 is a sufficient condition for checking that almost all points in the parameter space are regular. By inspecting the proof, we see that the set of irregular points is the zero set of a certain non-zero polynomial and thus has Lebesgue measure zero. However, it will typically be non-empty. As we emphasized in the introduction, the Gaussian approximation of complete U-statistics, or incomplete U-statistics with high computational budget, may already be very slow if the true parameter is close to an irregular point. Since the null hypothesis Θ_0 can be defined by a very large amount of constraints p, the sets of irregular points is the union of irregular points given by p sets with Lebesgue measure zero. In other words, there might be many hyper-surfaces with irregular points, such that it is likely that the true parameter is close to an irregular point. Thus, assuming mixed degeneracy is crucial for considering kernels as constructed above.

Example 4.4 Consider again a tetrad $f(\Sigma) = \sigma_{uv}\sigma_{wz} - \sigma_{uz}\sigma_{vw}$ and the corresponding kernel h given in Example 4.2. Since $\hat{\sigma}_{uv}(X_1) = X_{1u}X_{1v}$, we have that $\hat{\theta}_f(X_1) = (X_{1u}X_{1v}, X_{1w}X_{1z}, X_{1u}X_{1z}, X_{1v}X_{1w})^\top$. For Gaussian $X_1 \sim N_l(0, \Sigma)$, it is easy to see that $\mathrm{Cov}[\hat{\theta}_f(X_1)]$ is positive definite since no coordinate of $\hat{\theta}_f(X_1)$ is a linear function of the remaining coordinates. It follows by Proposition 4.3 that almost all covariance matrices Σ are regular points in the parameter space given by the cone of positive definite matrices. Covariance matrices with some entries equal to zero correspond to irregular points as we have seen in Example 1.3. Thus, all covariance matrices with those entries having small absolute values are close to irregular points.

5 Testing Gaussian latent-tree models

In this section, we apply our test for assessing the goodness-of-fit of Gaussian latent-tree models. These models are of particular relevance in phylogenetics (Semple & Steel, 2003; Zwiernik, 2016) and the problem of model selection is, for example, considered in Shiers et al. (2016) and Leung and Drton (2018); for a survey see Sung (2009) and Junker and Schreiber (2011). Our methodology is applicable since a full semi-algebraic description of the model is known, that is, all polynomial equalities and inequalities that fully describe the distributions corresponding to a given tree are known. The constraints include tetrads as well as higher order polynomial constraints. However, as we also point out in the introduction, it is challenging to test the large number of constraints simultaneously. For example, in Shiers et al. (2016) only small trees and a subset of constraints is tested. Typical approaches such as the Wald test can only handle $p \leq n$ constraints and the maximum likelihood function is difficult to optimize. For the latter, we implemented an expectation-maximization algorithm, but it is unclear whether it obtains the global maximum; see Section 5.2. Moreover, Drton (2009), Dufour et al. (2013), and Drton and Xiao (2016) show that irregular points lead to different limiting distributions in Wald and LR tests when testing tetrads. Hence, our testing strategy is an approach that targets both challenges in testing Gaussian latent-tree models: a large number of constraints and irregular points in the null hypothesis.

5.1 Model and constraints

We begin by briefly introducing Gaussian latent-tree models. Let $T = (V, E)$ be an undirected tree where V is the set of nodes and E is the set of edges. For $L \subseteq V$ denoting the set of leaves, we assume that every inner node in $V \setminus L$ has minimal degree 3. By Zwiernik (2016, Chapter 8) we have the following parametric representation of Gaussian latent-tree models. Let $PD(l)$ be the cone of symmetric positive definite $l \times l$ matrices.

Proposition 5.1 The Gaussian latent-tree model of a tree $T = (V, E)$ with leaves $L = \{1, \ldots, l\} \subseteq V$ is the family of Gaussian distributions $N_l(0, \Sigma)$ such

that Σ is in the set

$$\mathcal{M}(T) = \left\{ \Sigma = (\sigma_{uv}) \in PD(l) : \sigma_{uv} = \sqrt{\omega_u}\sqrt{\omega_v} \prod_{e \in ph_T(u,v)} \rho_e \text{ for } v \neq u,\ \sigma_{vv} = \omega_v \right\},$$

where $ph_T(u, v)$ denotes the set of *edges* on the unique path from u to v, the vector $(\omega_v)_{v \in L} \in \mathbb{R}^l_{>0}$ contains the variances and $(\rho_e)_{e \in E}$ is an $|E|$-dimensional vector with $|\rho_e| \in (0, 1)$.

For details on how this parametrization arises from the paradigm of graphical modelling, see Zwiernik (2016) or Drton et al. (2017). We may identify a Gaussian latent tree model with its set of covariance matrices $\mathcal{M}(T)$ and we will simply refer to $\mathcal{M}(T)$ as the model.

From a geometric point of view, the parameter space, i.e. the set of covariance matrices $\mathcal{M}(T)$, is fully understood. It is a semi-algebraic set described by polynomial equalities and inequalities in the entries of the covariance matrix. To state the polynomial constraints, we need the following notation borrowed from Leung and Drton (2018). Recall that in a tree any two nodes are connected by precisely one path. For four distinct leaves $\{u, v, w, z\} \subseteq L$, there are three possible partitions into two subsets of equal size, namely $\{u, v\}|\{w, z\}$, $\{u, w\}|\{v, z\}$, and $\{u, z\}|\{v, w\}$. These three partitions correspond to the three intersection of path pairs

$$ph_T(u, v) \cap ph_T(w, z), \quad ph_T(u, w) \cap ph_T(v, z) \quad \text{and} \quad ph_T(u, z) \cap ph_T(v, w). \tag{5.1}$$

By the structure of a tree, either all of the intersections give an empty set or exactly one is empty while the others are not; cf. Leung and Drton (2018). We let $\mathcal{Q} \subseteq \{\{u, v, w, z\} \subseteq L\}$ be the collections of subsets of size four such that $\{u, v, w, z\} \in \mathcal{Q}$ if exactly one of the intersections is empty. Given $\{u, v, w, z\} \in \mathcal{Q}$, we write $\{u, v\}|\{w, z\} \in \mathcal{Q}$ to indicate that $\{u, v, w, z\}$ belongs to $\mathcal{Q}$ and the paths $ph_T(u, v)$ and $ph_T(w, z)$ have empty intersection.

Proposition 5.2 (Leung & Drton, 2018). Let $\Sigma = (\sigma_{uv}) \in PD(l)$ be a covariance matrix with no zero entries. Then $\Sigma \in \mathcal{M}(T)$ if and only if Σ satisfies the following constraints:

1. *Inequality constraints:*
   (a) For any $\{u, v, w\} \subseteq L$, $-\sigma_{uv}\sigma_{uw}\sigma_{vw} \leq 0$.
   (b) For any $\{u, v, w\} \subseteq L$, $\sigma^2_{uv}\sigma^2_{vw} - \sigma^2_{vv}\sigma^2_{uw} \leq 0$, $\sigma^2_{uw}\sigma^2_{vw} - \sigma^2_{ww}\sigma^2_{uv} \leq 0$ and $\sigma^2_{uv}\sigma^2_{uw} - \sigma^2_{uu}\sigma^2_{vw} \leq 0$.
   (c) For any $\{u, v\}|\{w, z\} \in \mathcal{Q}$, $\sigma^2_{uw}\sigma^2_{vz} - \sigma^2_{uv}\sigma^2_{wz} \leq 0$.
2. *Equality constraints (tetrads):*
   (a) For any $\{u, v\}|\{w, z\} \in \mathcal{Q}$, $\sigma_{uw}\sigma_{vz} - \sigma_{uz}\sigma_{vw} = 0$.
   (b) For any $\{u, v, w, z\} \notin \mathcal{Q}$, $\sigma_{uz}\sigma_{vw} - \sigma_{uw}\sigma_{vz} = \sigma_{uv}\sigma_{wz} - \sigma_{uw}\sigma_{vz} = 0$.

We want to test a Gaussian latent-tree model against the saturated alternative, that is,

$$H_0 : \Sigma \in \mathcal{M}(T) \quad \text{vs.} \quad H_1 : \Sigma \in PD(l) \setminus \mathcal{M}(T)$$

based on i.i.d. samples $X_1, \ldots, X_n$ taken from a Gaussian distribution $N_l(0, \Sigma)$. With our methodology from Sections 3 and 4, we can simultaneously test all constraints defining $\mathcal{M}(T)$ given in Proposition 5.2. The number of constraints is $p = 2\binom{l}{4} + 4\binom{l}{3}$ which is of order $\mathcal{O}(l^4)$. Even for moderate values of l, say $l = 15$, this is a large number of constraints since $p = 4{,}550$.

5.2 Simulations

We have published an implementation of our tests for Gaussian latent-tree models in the `R` package `TestGGM`, available at https://github.com/NilsSturma/TestGGM. In the implementation we use $A = 1{,}000$ sets of Gaussian multipliers to compute the critical value $\hat{c}_W(1 - \alpha)$. The kernel for

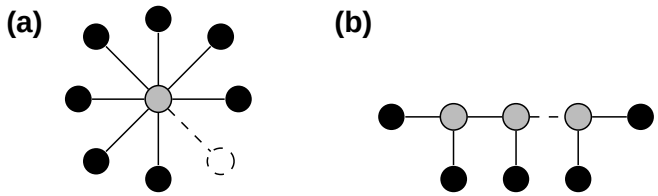


Figure 2. Graphical representation of (a) the star tree and (b) the binary caterpillar tree. Solid black dots correspond to leaves (observed variables).

estimating the polynomials is constructed as demonstrated in Example 4.1, where we treated the special case of tetrads. The package `TestGGM` also provides a routine for determining the specific constraints for a given latent-tree input, i.e. for determining the set $\mathcal{Q}$. For finding a path $ph_T(u, v)$ we use the function `shortest_paths` in the `igraph` package (Csardi & Nepusz, 2006).

We compare our test with the LR test. By Shiers et al. (2016) the dimension of a Gaussian latent-tree model is $|E| + |L|$. Thus, under regularity conditions, the LR test statistic approximates a chi-square distribution with $\binom{l+1}{2} - (|E| + |L|)$ degrees of freedom for $l \geq 4$ (Van der Vaart, 1998). For the one-factor model, the LR test is implemented by the function `factanal` in the `base` library of `R` (R Core Team, 2020). Note that it employs a Bartlett correction for better asymptotic approximation. For general trees the implementation shows some of the challenges of the likelihood ratio test. Optimizing the likelihood function on the subspace $\mathcal{M}(T)$ is difficult using the common non-convex methods due to the large number of constraints defining the subspace and the required positive definiteness of the covariance matrices. Therefore we implement an expectation maximization algorithm which was first proposed by Dempster et al. (1977) and is described in Friedman et al. (2002) and Mourad et al. (2013) for the special case of latent-tree models. A drawback of this method is that global optimization is not guaranteed.

In our numerical experiments we consider two different tree structures. One is the star tree, where the model is equal to the one-factor analysis model. Second, we consider the binary caterpillar tree which is extreme in a sense that it has the longest paths possible between leaves. In Figure 2, both structures are illustrated. We fix dimension $l = 15$, sample size $n = 500$ and generate data from both tree structures in three experimental set-ups:

(a) Star tree: All edge parameters ρ_e are equal to $\sqrt{0.5}$, while we take all variances $\omega_v = 2$.
(b) Star tree: Let h be the unique, unobserved inner node. The parameters for the edges $\{h, 1\}$ and $\{h, 2\}$ are taken to be 0.998, while all other edge parameters ρ_e are independently generated from $N(0, 0.1)$. The marginal variances are taken to be $\omega_v = 100$ for $v = 1, 2$ and $\omega_v = 1$ for all other leaves $v \in L$.
(c) Caterpillar tree: The edge parameters ρ_e are taken to be 0.998 except for all edges incident to four selected inner nodes in $V \setminus L$, where the edge parameters are independently generated from $N(0, 0.1)$. All variances ω_v are taken to be 2.

Set-up (a) is a regular problem where the LR test is the gold standard. In contrast, set-ups (b) and (c) are designed to be irregular problems with small covariances; compare with Example 1.3 in the introduction. In set-up (b), parameters are chosen such that the covariance matrix has exactly one off-diagonal entry which is far away from zero while all the remaining off-diagonal entries are close to zero. In this case, the parameters are also close to an algebraic singularity of the parameter space of the star tree (Drton et al., 2007, Proposition 32). The LR test should fail since the LR test statistic does not follow a chi-square distribution at singularities (Drton, 2009). Similarly, parameters in set-up (c) are close to an algebraic singularity since some of the edge parameters are almost zero (Drton et al., 2017) and thus we also expect the LR test to fail. Moreover, set-up (c) is considered to emphasize that the proposed methodology allows for testing arbitrary Gaussian latent tree models.

In Figures 3 and 4, we compare empirical test sizes of our test with the LR test for different, fixed significance levels $\alpha \in (0, 1)$ that we call the 'nominal level'. In addition, we consider different

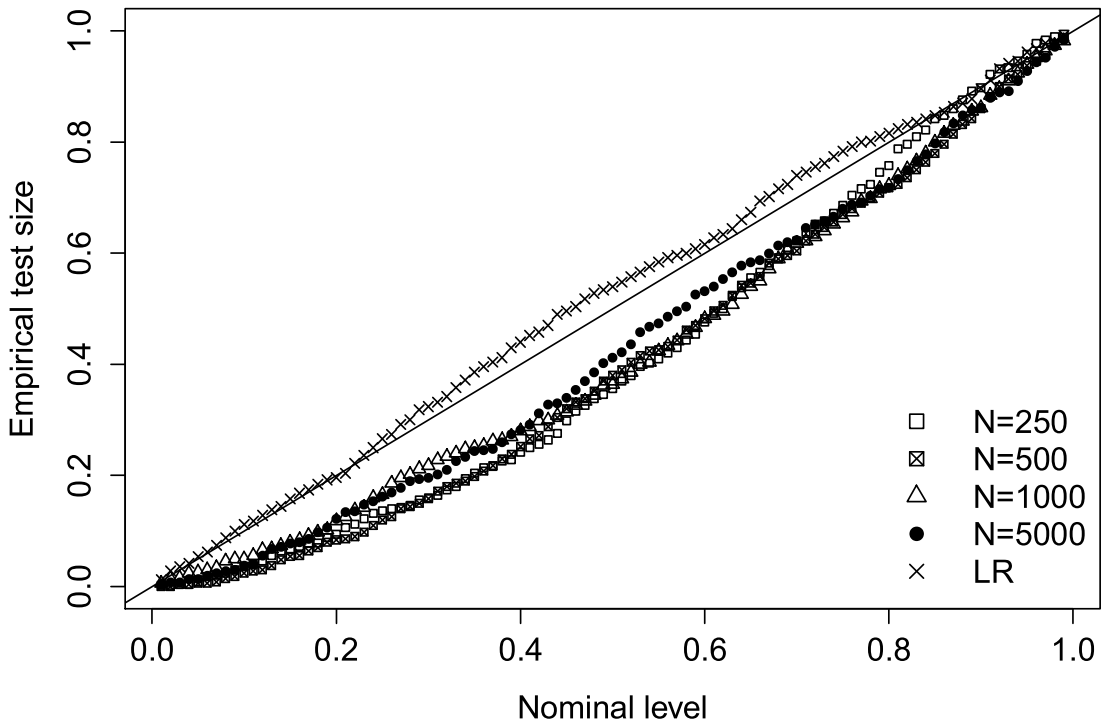


Figure 3. Empirical sizes vs. nominal levels for testing tetrads based on 500 experiments. The computational budget parameter N is varied as indicated and empirical sizes of the LR test are also shown. Data are generated from set-up (a) with $(l, n) = (15, 500)$.

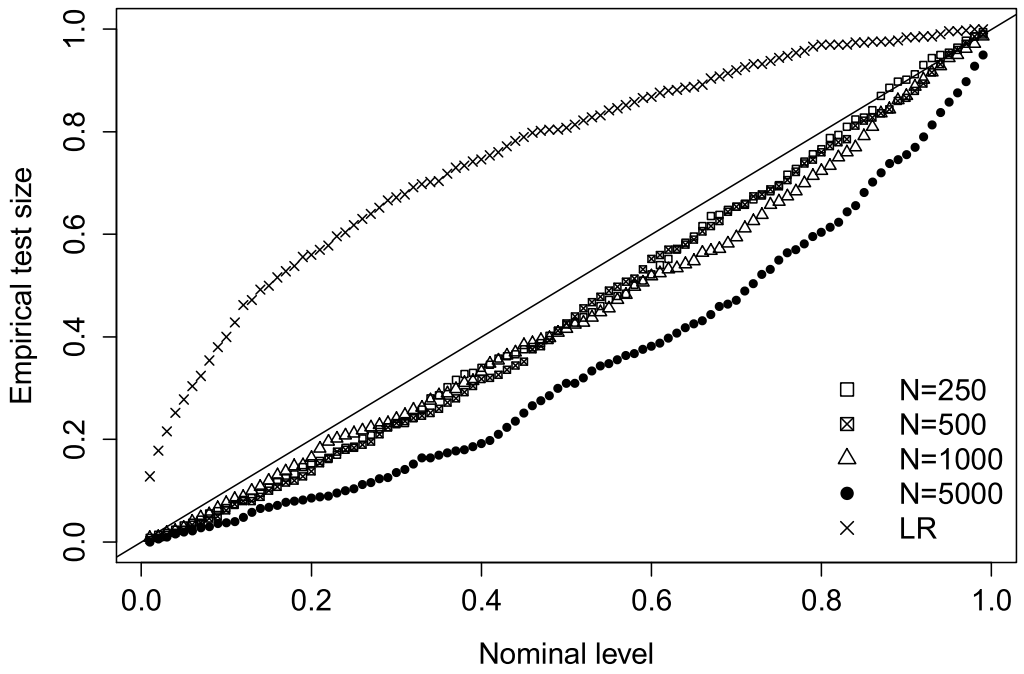
computational budget parameters N to see its influence on the behaviour of our test. To better compare test sizes we only consider testing the equality constraints (tetrads) in Proposition 5.2. This simplifies our analysis since in this case we can expect that the test size is equal to the level α. A statistical test that controls type I error is expected to have empirical test sizes below the 45 degree line in the plots. The lower the line, the more 'conservative' the test is. On the other hand, a test does not control type I error if the empirical sizes are above the 45 degree line. In this case the true null hypothesis is rejected too often.

The irregular set-ups (b) and (c) show the advantage of the proposed testing method because the empirical test sizes of the LR test do not hold the nominal levels, that is, the empirical sizes are above the 45 degree line. In comparison, our testing method holds level for all choices of the computational budget parameter N. However, there is an interesting difference in the behaviour of the empirical test sizes between the regular set-up (a) and set-ups (b) and (c) where parameters are close to an irregular point. In the regular set-up, a higher computational budget parameter N tends to yield empirical sizes closer to the nominal level α while in the irregular set-ups this is the other way round. The behaviour coincides with the theoretical guarantees discussed in Section 2 and Remark 3.3, suggesting that the bootstrap approximation holds under mixed degeneracy only when choosing the computational budget N appropriately, in this case $N = \mathcal{O}(n)$.

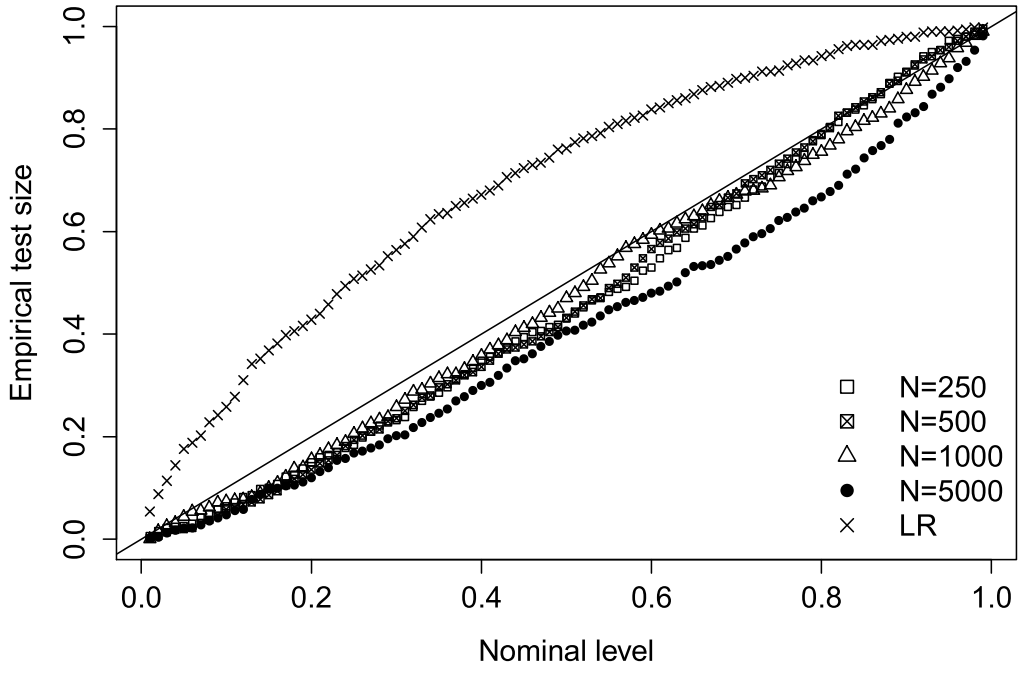
Besides size, we also study the empirical power of our test. For $\Sigma \in \mathcal{M}(T)$, we consider local alternatives

$$\widetilde{\Sigma} = \Sigma + \gamma\gamma^{\top} \frac{h}{\sqrt{n}}, \tag{5.2}$$

where $\gamma = (0, \ldots, 0, 1, 1) \in \mathbb{R}^l$ is fixed and $h > 0$ varies. According to Proposition 3.5, our test is consistent against alternatives outside of a neighbourhood of size $\sqrt{\log(p)/n}$ up to an unknown constant. Thus, it is natural to consider alternatives as in equation (5.2) that also depend on the sample size through $\sqrt{n}$. Moreover, all entries of the alternative covariance matrix $\widetilde{\Sigma}$ are equal to the entries of Σ except for four entries, which implies that only few constraints in Proposition 5.2 are violated. However, since our test statistic is the maximum of an incomplete U-statistic, we expect that our test is consistent even for such sparse signals.



Setup (b)



Setup (c)

Figure 4. Empirical sizes vs. nominal levels for testing tetrads based on 500 experiments. The computational budget parameter N is varied as indicated and empirical sizes of the LR test are also shown. Data are generated from set-ups (b) and (c) with $(l, n) = (15, 500)$.

For the regular set-up (a), the values of h are plotted against the empirical power in Figure 5. This time we considered all constraints in Proposition 5.2, equalities and inequalities. The power increases with h for all choices of the computational budget N. A higher budget empirically yields

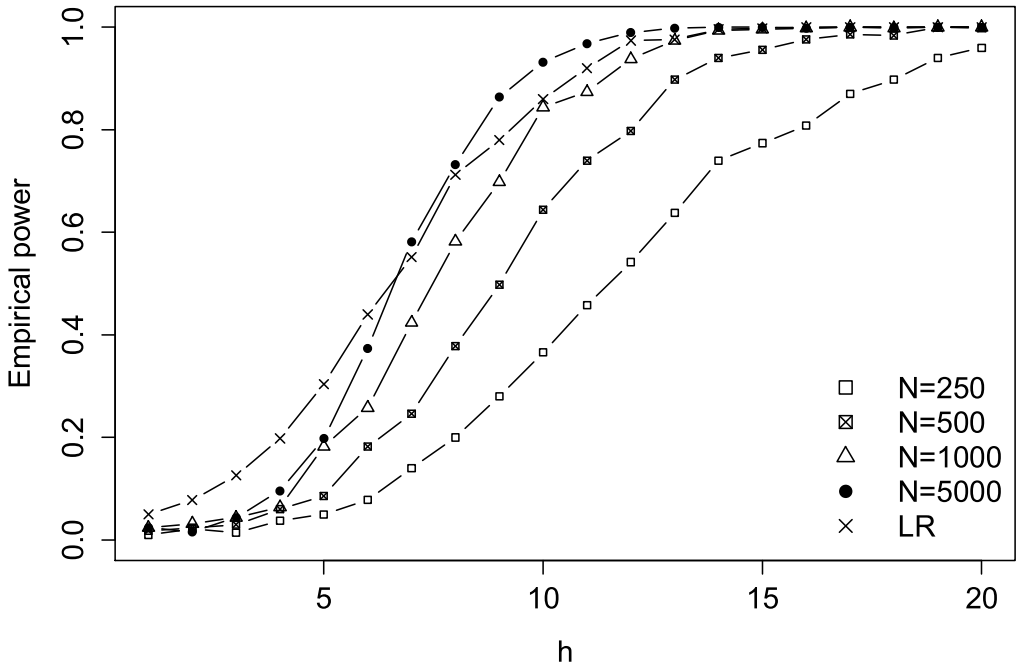


Figure 5. Empirical power for different local alternatives based on 500 experiments. The computational budget parameter N is varied as indicated and empirical power of the LR test is also shown. Local alternatives are generated as described in the text for set-up (a) with $(l, n) = (15, 500)$ and level $\alpha = 0.05$.

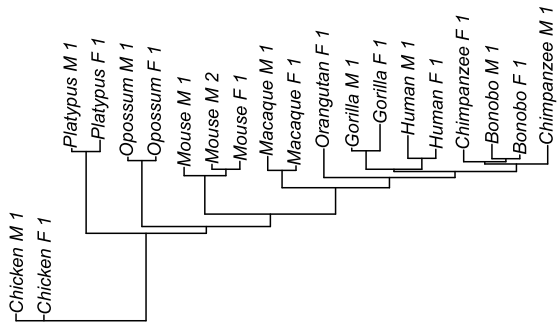


Figure 6. Mammalian gene expression phylogenies for cerebellum.

better power which is reasonable due to more precise estimation of the individual polynomials. For large computational budgets the power on the local alternatives is comparable to the LR test. However, as discussed before, one should be sceptical about the bootstrap approximation for large computational budgets N when the underlying true parameter is close to singular. The above experiments suggest that the choice $N = 2n$ is a good trade-off between guarding against irregularities and statistical efficiency. A similar plot for the singular set-ups can be found in Appendix D, online supplementary material. Moreover, we provide another simulation study in Appendix E, online supplementary material, where we apply our methodology to test minors in the two-factor analysis model.

5.3 Application to gene expression data

In this section, we analyze data from Brawand et al. (2011), where they obtained gene expression levels from sequencing of polyadenylated RNA from the cerebellum across different species such as gorillas, opossums, or chickens. We focus on data of 20 individuals from different species, where we have 5,636 gene expression levels for each individual. Based on the assumption that species evolved from a common ancestor, a phylogenetic tree describes the evolutionary history among a collection of species (Semple & Steel, 2003). In the work of Brawand et al. (2011), the authors reconstruct the phylogenetic tree in Figure 6 by neighbour-joining based on pair-wise

distance matrices. If we log-transform the original gene expression data, it is approximately normal distributed and it becomes a natural question to ask whether the data follows a Gaussian latent-tree model where the underlying tree structure is the same.

In our analysis to check tree compatibility we remove rows containing a zero such that we are left with $N = 4{,}615$ observations. Since the tree has 20 leaves, we have to test $p = 14250$ constraints that define the null hypothesis; recall Proposition 5.2. We compute p-values using the LR test and our test strategy with computational budget $N = 2n$. Both tests indicate overwhelming evidence against the Gaussian latent-tree model since we obtain a p-value of zero in both cases. To check whether the assumption of a Gaussian latent-tree model might hold locally, we also run the same analysis on all substructures of the tree induced by 6 neighbouring leaves. This implies that $p = 110$. Again, we only observe p-values equal to zero, indicating that the Gaussian latent-tree assumption is also too strong for the considered subsets of variables. We summarize that our approach is able to reach the same conclusion as the LR test, but our strategy is optimization-free and safe with respect to possible irregularities.

6 Discussion

We proposed a general methodology to simultaneously test many constraints on a statistical parameter. The test statistic is given by a studentized maximum of an incomplete U-statistic and critical values are approximated by Gaussian multiplier bootstrap. Our method is applicable to all hypotheses that can be defined by estimable equality and inequality constraints. If the constraints are polynomial, we presented a general procedure to construct kernels. In the suggested methodology, there is no need to maximize a possibly multi-modal function as one would do for a LR test. Moreover, our method allows for many constraints, that is, the number of constraints can be much larger than the sample size. If the computational budget N for constructing the incomplete U-statistic is appropriately chosen, typically of the same order as the sample size n, then the test asymptotically controls type I error even if the true parameter is irregular or close to irregular.

The latter fact is due to non-asymptotic Berry–Esseen-type error bounds on the high-dimensional Gaussian and bootstrap approximation of incomplete U-statistics. Compared to previous work, we have shown that the approximation is also valid under mixed degeneracy if we choose a suitable computational budget N. This yields control of type I error as well as consistency of our testing methodology even in irregular settings.

In practice, the requirement $N = \mathcal{O}(n)$ allows some flexibility in choosing the computational budget N. As illustrated in our simulations, a higher computational budget parameter yields more efficient estimates of the constraints and therefore a more powerful test. However, the computations become more involved for higher order kernels and one should be sceptical about the theoretical guarantees of the bootstrap approximation if N is large and the underlying true parameter is close to an irregular point. Thus, there is a trade-off between efficiency and guarding against irregularities. In our simulations we found that choosing the computational budget parameter as $N = 2n$ is a reasonable choice.

If the inequality constraints defining the null hypothesis are strict, then there is a decrease in power. One may improve power by using a two-step approach for testing inequalities that was initially introduced in Romano et al. (2014) and then further developed by Bai et al. (2021) for the high-dimensional set-up. In the first step, a confidence region for the constraints of interest is constructed and in the second step, this set is used to provide information about which constraints are negative. In future work, it would be of interest to improve our strategy using the mentioned two-step approach.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 883818), and from the German Federal Ministry of Education and Research and the Bavarian State Ministry for Science and the Arts. Nils Sturma acknowledges support by the Munich Data Science Institute (MDSI) at Technical University of Munich via the Linde/MDSI PhD Fellowship programme.

Conflict of interests: None declared.

Data availability

We have published an implementation of our tests for Gaussian latent-tree models in the R package TestGGM, available at https://github.com/NilsSturma/TestGGM.

Supplementary material

Supplementary material is available online at *Journal of the Royal Statistical Society: Series B*.

References

Bai Y., Santos A., & Shaikh A. M. (2021). A two-step method for testing many moment inequalities. *Journal of Business & Economic Statistics*, *40*(3), 1–11. https://doi.org/10.1080/07350015.2021.1897016

Bollen K. A., & Ting K.-F. (2000). A tetrad test for causal indicators. *Psychological Methods*, *5*(1), 3–22. https://doi.org/10.1037/1082-989X.5.1.3

Brawand D., Soumillon M., Necsulea A., Julien P., Csárdi G., Harrigan P., Weier M., Liechti A., Aximu-Petri A., Kircher M., Albert F. W., Zeller U., Khaitovich P., Grützner F., Bergmann S., Nielsen R., Pääbo S., & Kaessmann H. (2011). The evolution of gene expression levels in mammalian organs. *Nature*, *478*(7369), 343–348. https://doi.org/10.1038/nature10532

Chen B., Kumor D., & Bareinboim E. (2017). Identification and model testing in linear structural equation models using auxiliary variables. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, ICML'17 (pp. 757–766). JMLR.

Chen B., Tian J., & Pearl J. (2014). Testable implications of linear structural equation models. In *Proceedings of the 20th National Conference on Artificial Intelligence (AAAI)* (Vol. 28). AAAI Press.

Chen X. (2018). Gaussian and bootstrap approximations for high-dimensional U-statistics and their applications. *Annals of Statistics*, *46*(2), 642–678. https://doi.org/10.1214/17-AOS1563

Chen X., & Kato K. (2019). Randomized incomplete *U*-statistics in high dimensions. *Annals of Statistics*, *47*(6), 3127–3156. https://doi.org/10.1214/18-AOS1773

Chen X., & Kato K. (2020). Jackknife multiplier bootstrap: Finite sample approximations to the *U*-process supremum with applications. *Probability Theory and Related Fields*, *176*(3–4), 1097–1163. https://doi.org/10.1007/s00440-019-00936-y

Chernozhukov V., Chetverikov D., & Kato K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics*, *41*(6), 2786–2819. https://doi.org/10.1214/13-AOS1161

Chernozhukov V., Chetverikov D., & Kato K. (2017a). Central limit theorems and bootstrap in high dimensions. *The Annals of Probability*, *45*(4), 2309–2352. https://doi.org/10.1214/16-AOP1113

Chernozhukov V., Chetverikov D., & Kato K. (2017b). 'Detailed proof of Nazarov's inequality', arXiv, arXiv:1711.10696, preprint: not peer reviewed.

Chernozhukov V., Chetverikov D., & Kato K. (2019). Inference on causal and structural parameters using many moment inequalities. *The Review of Economic Studies*, *86*(5), 1867–1900. https://doi.org/10.1093/restud/rdy065

Chernozhukov V., Chetverikov D., Kato K., & Koike Y. (2022). Improved central limit theorem and bootstrap approximations in high dimensions. *Annals of Statistics*, *50*(5), 2562–2586. https://doi.org/10.1214/22-AOS2193

Chernozhukov V., Chetverikov D., & Koike Y. (2023a). Nearly optimal central limit theorem and bootstrap approximations in high dimensions. *The Annals of Applied Probability*, *33*(3), 2374–2425. https://doi.org/10.1214/22-AAP1870

Chernozhukov V., Newey W. K., & Santos A. (2023b). Constrained conditional moment restriction models. *Econometrica*, *91*(2), 709–736. https://doi.org/10.3982/ECTA13830

Chetverikov D., Santos A., & Shaikh A. M. (2018). The econometrics of shape restrictions. *Annual Review of Economics*, *10*(1), 31–63. https://doi.org/10.1146/economics.2018.10.issue-1

Claassen T., & Heskes T. (2012). A Bayesian approach to constraint based causal inference. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence (UAI)*, UAI'12 (pp. 207–216). AUAI Press.

Cox D. A., Little J., & O'Shea D. (2015). *Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra*. Undergraduate Texts in Mathematics (4th ed.). Springer.

Csardi G., & Nepusz T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9. https://doi.org/10.5281/zenodo.7682609

de la Peña V. H., & Giné E. (1999). *Decoupling: From dependence to independence, randomly stopped processes. U-statistics and processes. Martingales and beyond*. Probability and its Applications. Springer-Verlag.

Dempster A. P., Laird N. M., & Rubin D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22. (With discussion). https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Drton M. (2009). Likelihood ratio tests and singularities. *The Annals of Statistics*, *37*(2), 979–1012. https://doi.org/10.1214/07-AOS571

Drton M., Lin S., Weihs L., & Zwiernik P. (2017). Marginal likelihood and model selection for Gaussian latent tree and forest models. *Bernoulli*, *23*(2), 1202–1232. https://doi.org/10.3150/15-BEJ775

Drton M., Massam H., & Olkin I. (2008). Moments of minors of Wishart matrices. *The Annals of Statistics*, *36*(5), 2261–2283. https://doi.org/10.1214/07-AOS522

Drton M., Sturmfels B., & Sullivant S. (2007). Algebraic factor analysis: Tetrads, pentads and beyond. *Probability Theory and Related Fields*, *138*(3–4), 463–493. https://doi.org/10.1007/s00440-006-0033-2

Drton M., & Xiao H. (2016). Wald tests of singular hypotheses. *Bernoulli*, *22*(1), 38–59. https://doi.org/10.3150/14-BEJ620

Dufour J.-M., Renault E., & Zinde-Walsh V. (2013). 'Wald tests when restrictions are locally singular', arXiv, arXiv:1312.0569, preprint: not peer reviewed.

Fang X., & Koike Y. (2021). High-dimensional central limit theorems by Stein's method. *The Annals of Applied Probability*, *31*(4), 1660–1686. https://doi.org/10.1214/20-AAP1629

Friedman N., Ninio M., Pe'er I., & Pupko T. (2002). A structural EM algorithm for phylogenetic inference. *Journal of Computational Biology*, *9*(2), 331–353. https://doi.org/10.1089/10665270252935494

Gaffke N., Heiligers B., & Offinger R. (2002). On the asymptotic null-distribution of the Wald statistic at singular parameter points. *Statistics & Risk Modeling*, *20*(4), 379–398. https://doi.org/10.1524/strm.2002.20.14.379

Gaffke N., Steyer R., & von Davier A. A. (1999). On the asymptotic null-distribution of the Wald statistic at singular parameter points. *Statistics & Risk Modeling*, *17*(4), 339–358. https://doi.org/10.1524/strm.1999.17.4.339

Götze F., Sambale H., & Sinulis A. (2021). Concentration inequalities for polynomials in α-sub-exponential random variables. *Electronic Journal of Probability*, *26*(48), 1–22. https://doi.org/10.1214/21-EJP606

Hipp J. R., & Bollen K. A. (2003). Model fit in structural equation models with censored, ordinal, and dichotomous variables: Testing vanishing tetrads. *Sociological Methodology*, *33*(1), 267–305. https://doi.org/10.1111/j.0081-1750.2003.00133.x

Junker B. H., & Schreiber F. (2011). *Analysis of biological networks*. John Wiley & Sons.

Kuchibhotla A. K., & Chakrabortty A. (2022). Moving beyond sub-Gaussianity in high-dimensional statistics: Applications in covariance estimation and linear regression. *Information and Inference: A Journal of the IMA*, *11*(4), 1389–1456. https://doi.org/10.1093/imaiai/iaac012

Leung D., & Drton M. (2018). Algebraic tests of general Gaussian latent tree models. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)* (Vol. 31, pp. 6301–6310). Neural Information Processing Systems Foundation.

Leung D., Drton M., & Hara H. (2016). Identifiability of directed Gaussian graphical models with one latent source. *Electronic Journal of Statistics*, *10*(1), 394–422. https://doi.org/10.1214/16-EJS1111

Leung D., & Sturma N. (2024). 'Singularity-agnostic incomplete U-statistics for testing polynomial constraints in Gaussian covariance matrices', arXiv, arXiv:2401.02112, preprint: not peer reviewed.

Lopes M. E. (2022). Central limit theorem and bootstrap approximation in high dimensions: Near $1/\sqrt{n}$ rates via implicit smoothing. *The Annals of Statistics*, *50*(5), 2492–2513. https://doi.org/10.1214/22-AOS2184

Mourad R., Sinoquet C., Zhang N. L., Liu T., & Leray P. (2013). A survey on latent tree models and applications. *Journal of Artificial Intelligence Research*, *47*, 157–203. https://doi.org/10.1613/jair.3879

Okamoto M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *The Annals of Statistics*, *1*(4), 763–765. https://doi.org/10.1214/aos/1176342472

Pearl J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge University Press.

Pearl J., & Verma T. S. (1995). A theory of inferred causation. In D. Prawitz, B. Skyrms, & D. Westerståhl (Eds.), *Logic, Methodology and Philosophy of Science IX*. Studies in Logic and the Foundations of Mathematics (Vol. 134, pp. 789–811). Elsevier.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Romano J. P., Shaikh A. M., & Wolf M. (2014). A practical two-step method for testing moment inequalities. *Econometrica*, *82*(5), 1979–2002. https://doi.org/10.3982/ECTA11011

Semple C., & Steel M. (2003). *Phylogenetics*. Oxford Lecture Series in Mathematics and its Applications (Vol. *24*). Oxford University Press.

Shiers N., Zwiernik P., Aston J. A. D., & Smith J. (2016). The correlation space of Gaussian latent tree models and model selection without fitting. *Biometrika*, *103*(3), 531–545. https://doi.org/10.1093/biomet/asw032

Silva R., Scheines R., Glymour C., & Spirtes P. (2006). Learning the structure of linear latent variable models. *Journal of Machine Learning Research*, *7*, 191–246. http://jmlr.org/papers/v7/silva06a.html

Song Y., Chen X., & Kato K. (2019). Approximating high-dimensional infinite-order *U*-statistics: Statistical and computational guarantees. *Electronic Journal of Statistics*, *13*(2), 4794–4848. https://doi.org/10.1214/19-EJS1643

Spearman C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201–292. https://doi.org/10.2307/1412107

Spirtes P., Glymour C., & Scheines R. (2000). *Causation, prediction, and search*. Adaptive Computation and Machine Learning (2nd ed.). MIT Press. (With additional material by David Heckerman, Christopher Meek, Gregory F. Cooper and Thomas Richardson, A Bradford Book).

Steyer R. (2005). Analyzing individual and average causal effects via structural equation models. *Methodology*, *1*(1), 39–54. https://doi.org/10.1027/1614-1881.1.1.39

Strieder D., Freidling T., Haffner S., & Drton M. (2021). Confidence in causal discovery with linear causal models. In C. de Campos & M. H. Maathuis (Eds.). *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence of Proceedings of Machine Learning Research* (Vol. 161, pp. 1217–1226). PMLR.

Sullivant S., Talaska K., & Draisma J. (2010). Trek separation for Gaussian graphical models. *The Annals of Statistics*, *38*(3), 1665–1685. https://doi.org/10.1214/09-AOS760

Sung W.-K. (2009). *Algorithms in bioinformatics: A practical introduction*. Chapman & Hall/CRC.

Van der Vaart A. W. (1998). *Asymptotic statistics*. Cambridge Series in Statistical and Probabilistic Mathematics (Vol. *3*). Cambridge University Press.

Vershynin R. (2018). *High-dimensional probability: An introduction with applications in data science*. Cambridge Series in Statistical and Probabilistic Mathematics (Vol. *47*). Cambridge University Press. (With a foreword by Sara van de Geer).

Wishart J. (1928). Sampling errors in the theory of two factors. *British Journal of Psychology*, *19*(2), 180–187. https://doi.org/10.1111/j.2044-8295.1928.tb00508.x

Zwiernik P. (2016). *Semialgebraic statistics and latent tree models*. Monographs on Statistics and Applied Probability (Vol. *146*). Chapman & Hall/CRC.

Supplement to "Testing Many Constraints in Possibly Irregular Models Using Incomplete U-Statistics"

This supplement contains additional material such as all technical proofs (Appendix A), additional lemmas (Appendix B), properties of sub-Weibull random variables (Appendix C), additional simulation results for Gaussian latent tree models (Appendix D) and another application of our methodology, where we test minors in two-factor analysis models (Appendix E).

Appendix A. Proofs

Since we only consider centered incomplete U-statistics, we assume without loss of generality $\mu = 0$ in all our proofs of the results in Section 2.

A.1. **Proof of Theorem 2.4.** In this section we prove the main theorem. Let $R = [a, b] \in \mathbb{R}^p_{\text{re}}$ be a hyperrectangle. For any partition $p = p_1 + p_2$ with $p_1, p_2 \geq 0$, we write $R = (R^{(1)}, R^{(2)})$, where $R^{(1)}$ and $R^{(2)}$ are hyperrectangles in $\mathbb{R}^{p_1}$ and $\mathbb{R}^{p_2}$. Moreover, for a given hyperrectangle and $t > 0$, we denote the following t-dependent quantities

$$R_t = [a - t, b + t], \quad R_{-t} = [a + t, b - t] \quad \text{and} \quad R_{\text{env},t} = R_t \setminus R_{-t} = [a - t, a + t) \cup (b - t, b + t].$$

We begin with stating an abstract bound on *complete* U-statistics; recall the definition $U_n = \frac{1}{|I_{n,m}|} \sum_{\iota \in I_{n,m}} h(X_\iota)$. Let $\Gamma_g^{(1)} = \text{Cov}[g^{(1)}(X_1)]$ be the covariance matrix of the first p_1 indices of $g(X_1)$ and denote the symmetric $p \times p$ block matrix

$$\overline{\Gamma}_g = \begin{pmatrix} \Gamma_g^{(1)} & 0 \\ 0 & 0 \end{pmatrix}.$$

Lemma A.1. *Assume (C2) and (C4) - (C6) hold and let $R \in \mathbb{R}^p_{\text{re}}$. Then there is a constant $C_\beta > 0$ only depending on β such that*

$$|P(\sqrt{n}(U_n - \mu) \in R) - P(m\overline{Y}_g \in R)| \leq C_\beta \left(\frac{m^2 D_n^2 \log(p_1 n)^{1+6/\beta}}{(\underline{\sigma}^2_{g^{(1)}} \wedge 1)\, n} \right)^{1/6} + \mathbb{1}_{\{0 \in R^{(2)}_{env,t}\}}$$

with $t = C_\beta n^{-\min\{k,1/3\}} m^2 D_n \log(p_2 n)^{2/\beta}$ and $\overline{Y}_g \sim N_p(0, \overline{\Gamma}_g)$.

The bound in Lemma A.1 is not uniform for all hyperrectangles and is therefore of limited use on its own. However, it will be the main tool to prove Theorem 2.4.

Proof of Lemma A.1. Let C_β be a constant that depends only on β and that varies its value from place to place. For any other constant $c > 0$ we assume without loss of generality

$$\frac{m^2 \log(p_2)}{n} \leq c \tag{A.1}$$

since otherwise the statement from Lemma A.1 becomes trivial by choosing C_β large enough. First, we denote preliminary observations which we will use throughout the proof. Let X be a random vector in $\mathbb{R}^{p_1}$ and c be a non-random vector in $\mathbb{R}^{p_2}$.

(i) $P((X, c) \in R) = P(\{X \in R^{(1)}\} \cap \{c \in R^{(2)}\}) = P(X \in R^{(1)}) \mathbb{1}_{\{c \in R^{(2)}\}}$.
(ii) Let Y be another random vector in $\mathbb{R}^{p_1}$. Then

$$\begin{aligned} |P((Y, c) \in R) - P((X, c) \in R)| &= |P(Y \in R^{(1)}) - P(X \in R^{(1)})| \mathbb{1}_{\{c \in R^{(2)}\}} \\ &\leq |P(Y \in R^{(1)}) - P(X \in R^{(1)})|. \end{aligned}$$

(iii) Suppose we are given two hyperrectangles $R = (R^{(1)}, R^{(2)})$ and $\tilde{R} = (R^{(1)}, \tilde{R}^{(2)})$ having identical first part $R^{(1)}$. Then

$$\begin{aligned}|P((X,c) \in \tilde{R}) - P((X,c) \in R)| &= |P(X \in R^{(1)})\mathbb{1}_{\{c \in \tilde{R}^{(2)}\}} - P(X \in R^{(1)})\mathbb{1}_{\{c \in R^{(2)}\}}| \\ &= P(X \in R^{(1)})|\mathbb{1}_{\{c \in \tilde{R}^{(2)}\}} - \mathbb{1}_{\{c \in R^{(2)}\}}| \\ &\leq |\mathbb{1}_{\{c \in \tilde{R}^{(2)}\}} - \mathbb{1}_{\{c \in R^{(2)}\}}|.\end{aligned}$$

Now, let $A_n = \sqrt{n}\, U_n$. Recall the notation of splitting the vector $A_n = (A_n^{(1)}, A_n^{(2)})$ according to the partition $p = p_1 + p_2$. Let $R = [a,b] \in \mathbb{R}^p_{\mathrm{re}}$ be a hyperrectangle and denote

$$\omega_n = \left(\frac{m^2 D_n^2 \log(p_1 n)^{1+6/\beta}}{(\sigma^2_{g^{(1)}} \wedge 1)\, n} \right)^{1/6}.$$

For any $t > 0$ we have

$$\begin{aligned}P(A_n \in R) &\leq P(\{A_n \in R\} \cap \{\|A_n^{(2)} - 0\|_\infty \leq t\}) + P(\|A_n^{(2)} - 0\|_\infty > t) \\ &\leq P(\{A_n^{(1)} \in R^{(1)}\} \cap \{0 \in R_t^{(2)}\}) + P(\|A_n^{(2)} - 0\|_\infty > t) \\ &\leq P(\{mY_g^{(1)} \in R^{(1)}\} \cap \{0 \in R_t^{(2)}\}) + C_\beta \omega_n + P(\|A_n^{(2)}\|_\infty > t) \\ &\leq P(m(Y_g^{(1)}, 0) \in R) + \mathbb{1}_{\{0 \in R_t^{(2)} \setminus R^{(2)}\}} + C_\beta \omega_n + P(\|A_n^{(2)}\|_\infty > t),\end{aligned}$$

where the second to last inequality follows from observation (ii) and Lemma B.2 that is applicable due to Conditions (C2), (C4) and (C6). The last inequality follows from observation (iii). Observe that $(Y_g^{(1)}, 0) \sim \overline{Y}_g$ and $\mathbb{1}_{\{0 \in R_t^{(2)} \setminus R^{(2)}\}} \leq \mathbb{1}_{\{0 \in R^{(2)}_{\mathrm{env},t}\}}$. Thus, it is left to bound $P(\|A_n^{(2)}\|_\infty > t)$. We have

$$\begin{aligned}P(\|A_n^{(2)}\|_\infty > t) &\leq P\left(\left\| A_n^{(2)} - \frac{m}{\sqrt{n}} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty + \left\| \frac{m}{\sqrt{n}} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > t \right) \\ &\leq P\left(\left\| A_n^{(2)} - \frac{m}{\sqrt{n}} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > \frac{t}{2} \right) + P\left(\left\| \frac{m}{\sqrt{n}} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > \frac{t}{2} \right).\end{aligned}$$

Due to (A.1) we may apply Theorem 5.1 in Song et al. (2019) to bound the first summand. Together with the Markov inequality we obtain

$$\begin{aligned}P\left(\left\| A_n^{(2)} - \frac{m}{\sqrt{n}} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > \frac{t}{2} \right) &= P\left(\left\| \frac{A_n^{(2)}}{\sqrt{n}} - \frac{m}{n} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > \frac{t}{2\sqrt{n}} \right) \\ &\leq C_\beta \frac{m^2 D_n \log(p_2)^{1+1/\beta}}{t\sqrt{n}}. \qquad \text{(A.2)}\end{aligned}$$

To bound the second summand we apply Lemma A.2 in Song et al. (2019) which yields

$$P\left(\left\| \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > C_\beta(\sqrt{a_n} \log(p_2 n)^{1/2} + u_n \log(p_2 n)^{2/\beta}) \right) \leq \frac{4}{n},$$

where $a_n = n \max_{p_1+1 \leq j \leq p} \|g_j(X_1)\|_2^2$ and $u_n = \max_{p_1+1 \leq j \leq p} \|g_j(X_j)\|_{\psi_\beta}$. By Assumption (C5) we have $u_n \leq n^{-k} D_n$ and due to Lemma C.4 it holds $a_n \leq C_\beta n \|g_j(X_i)\|^2_{\psi_\beta} \leq C_\beta n^{1-2k} D_n^2$. Hence,

$$P\left(\left\| \frac{m}{\sqrt{n}} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > C_\beta \left(\frac{m D_n \log(p_2 n)^{1/2}}{n^k} + \frac{m D_n \log(p_2 n)^{2/\beta}}{n^{k+1/2}} \right) \right) \leq \frac{4}{n}.$$

Now, choosing $t = C_\beta n^{-\min\{k,1/3\}} m^2 D_n \log(p_2 n)^{2/\beta}$ we get

$$P\left(\left\| \frac{m}{\sqrt{n}} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > \frac{t}{2}\right) \leq \frac{4}{n} \leq \omega_n$$

and due to (A.2) we have

$$P\left(\left\| A_n^{(2)} - \frac{m}{\sqrt{n}} \sum_{i=1}^n g^{(2)}(X_i) \right\|_\infty > \frac{t}{2}\right) \leq C_\beta \frac{n^{\min\{k,1/3\}}}{n^{1/2}} \leq C_\beta n^{-1/6} \leq C_\beta \omega_n.$$

Therefore, we conclude that $P(\|A_n^{(2)}\|_\infty > t) \leq C_\beta \omega_n$ and we have shown that

$$P(A_n \in R) \leq P(m\overline{Y}_g \in R) + C_\beta \omega_n + \mathbb{1}_{\{0 \in R^{(2)}_{\mathrm{env},t}\}}.$$

Likewise, for the other direction it holds

$$\begin{aligned} &P(A_n \in R) + P(\|A_n^{(2)} - 0\|_\infty > t) = P(\{A_n^{(1)} \in R^{(1)}\} \cap \{A_n^{(2)} \in R^{(2)}\}) + P(\|A_n^{(2)} - 0\|_\infty > t) \\ &\geq P(\{A_n^{(1)} \in R^{(1)}\} \cap [\{A_n^{(2)} \in R^{(2)}\} \cup \{|A_n^{(2)} - 0\|_\infty > t\}]) \geq P(\{A_n^{(1)} \in R^{(1)}\} \cap \{0 \in R^{(2)}_{-t}\}). \end{aligned}$$

Thus, we have by Lemma B.2 and observations (ii) and (iii),

$$\begin{aligned} P(A_n \in R) &\geq P(\{A_n^{(1)} \in R^{(1)}\} \cap \{0 \in R^{(2)}_{-t}\}) - P(\|A_n^{(2)} - 0\|_\infty > t) \\ &\geq P(\{mY_g^{(1)} \in R^{(1)}\} \cap \{0 \in R^{(2)}_{-t}\}) - C_\beta \omega_n - P(\|A_n^{(2)}\|_\infty > t) \\ &\geq P(m(Y_g^{(1)}, 0) \in R) - \mathbb{1}_{\{0 \in R^{(2)} \setminus R^{(2)}_{-t}\}} - C_\beta \omega_n - P(\|A_n^{(2)}\|_\infty > t). \end{aligned}$$

Since $\mathbb{1}_{\{0 \in R^{(2)} \setminus R^{(2)}_{-t}\}} \leq \mathbb{1}_{\{0 \in R^{(2)}_{\mathrm{env},t}\}}$ and $P(\|A_n^{(2)}\|_\infty > t) \leq C_\beta \omega_n$, the proof is complete. □

If the bounds $a^{(2)}, b^{(2)}$ of the hyperrectangle $R^{(2)} = [a^{(2)}, b^{(2)}]$ in Lemma A.1 depend on a (Gaussian) random vector, then one may use anti-concentration inequalities to bound $\mathbb{E}[\mathbb{1}_{\{0 \in R^{(2)}_{\mathrm{env},t}\}}] = P(0 \in R^{(2)}_{\mathrm{env},t})$. We use this strategy to prove Theorem 2.4. The proof is split into two further propositions, where we first prove the approximation $\sqrt{n}(U'_{n,N}) - \mu \approx N_p(0, m^2, \overline{\Gamma}_g + \alpha_n \Gamma_h)$ and then n $N_p(0, m^2, \overline{\Gamma}_g + \alpha_n \Gamma_h) \approx N_p(0, m^2, \Gamma_g + \alpha_n \Gamma_h)$ on the hyperrectangles.

Proposition A.2. *Assume (C1) - (C6) hold. Then there is a constant $C_\beta > 0$ only depending on β such that*

$$\sup_{R \in \mathbb{R}^p_{\mathrm{re}}} |P(\sqrt{n}(U'_{n,N} - \mu) \in R) - P(\overline{Y} \in R)| \leq C_\beta \{\omega_{n,1} + \omega_{n,2} + \omega_{n,3}\} \tag{A.3}$$

with $\overline{Y} \sim N_p(0, m^2 \overline{\Gamma}_g + \alpha_n \Gamma_h)$.

Proof of Proposition A.2. Let $C_\beta > 0$ be a constant that depends only on β and that varies its value from place to place. As in the proofs of Theorem 3.1 in Chen and Kato (2019) and Theorem 2.4 in Song et al. (2019) we write $U'_{n,N} = (N/\widehat{N})(U_n + \sqrt{1-\rho_n} B_n)$ with $B_n = \frac{1}{N} \sum_{\iota \in I_{n,r}} \frac{Z_\iota - \rho_n}{\sqrt{1-\rho_n}} h(X_\iota)$. For any hyperrectangle $R = [a, b] \in \mathbb{R}^p_{\mathrm{re}}$ we have

$$\begin{aligned} &P(\sqrt{n}(U_n + \sqrt{1-\rho_n} B_n) \in R) \\ &= \mathbb{E}\left[P|_{X_1^n} \left(\sqrt{N} B_n \in \left(\frac{1}{\sqrt{\alpha_n(1-\rho_n)}} R - \sqrt{\frac{N}{1-\rho_n}} U_n \right) \right) \right] \\ &\leq \mathbb{E}\left[P|_{X_1^n} \left(Y_h \in \left(\frac{1}{\sqrt{\alpha_n(1-\rho_n)}} R - \sqrt{\frac{N}{1-\rho_n}} U_n \right) \right) \right] + C_\beta \omega_{n,1} \end{aligned}$$

$$
\begin{aligned}
&= P(\sqrt{n}U_n \in [R - \sqrt{\alpha_n(1-\rho_n)}Y_h]) + C_\beta \omega_{n,1} \\
&= \mathbb{E}[P|_{Y_h}(\sqrt{n}U_n \in [R - \sqrt{\alpha_n(1-\rho_n)}Y_h])] + C_\beta \omega_{n,1} \\
&\leq \mathbb{E}[P|_{Y_h}(m\overline{Y}_g \in [R - \sqrt{\alpha_n(1-\rho_n)}Y_h]) + \mathbb{1}_{\{0 \in \tilde{R}^{(2)}_{\mathrm{env},t}\}}] + C_\beta \omega_{n,1} \\
&= P(m\overline{Y}_g \in [R - \sqrt{\alpha_n(1-\rho_n)}Y_h]) + \mathbb{E}[\mathbb{1}_{\{0 \in \tilde{R}^{(2)}_{\mathrm{env},t}\}}] + C_\beta \omega_{n,1},
\end{aligned}
$$

where $\tilde{R} = R - \sqrt{\alpha_n(1-\rho_n)}Y_h$ and $t = C_\beta n^{-\min\{k,1/3\}} m^2 D_n \log(p_2 n)^{2/\beta}$. The first inequality follows from Lemma B.1 since B_n is an independent sum given the samples $X_1, \ldots, X_n$ and Conditions (C1) - (C3) are satisfied; for a detailed proof see Song et al. (2019, Lemma A.14). The second inequality follows from Lemma A.1. To bound the expectation appearing on the right-hand side we use an anti-concentration inequality due to Nazarov. It is stated in Lemma A.1 in Chernozhukov et al. (2017a) and a detailed proof is given in the note Chernozhukov et al. (2017b). Let $\tilde{a} = a - \sqrt{\alpha_n(1-\rho_n)}Y_h$ and $\tilde{b} = b - \sqrt{\alpha_n(1-\rho_n)}Y_h$ be random vectors in $\mathbb{R}^p$ such that the rectangle $\tilde{R}$ is given by $\tilde{R} = [\tilde{a}, \tilde{b}]$. Then we have

$$
\begin{aligned}
&\mathbb{E}[\mathbb{1}_{\{0 \in \tilde{R}^{(2)}_{\mathrm{env},t}\}}] = P(0 \in \tilde{R}^{(2)}_{\mathrm{env},t}) = P(\{\tilde{a}^{(2)} - t \leq 0 < \tilde{a}^{(2)} + t\} \cup \{\tilde{b}^{(2)} - t < 0 \leq \tilde{b}^{(2)} + t\}) \\
&\quad \leq P(\tilde{a}^{(2)} - t \leq 0 < \tilde{a}^{(2)} + t) + P(\tilde{b}^{(2)} - t < 0 \leq \tilde{b}^{(2)} + t) \\
&\quad = P(a^{(2)} - t \leq \sqrt{\alpha_n(1-\rho_n)}Y^{(2)}_h < a^{(2)} + t) + P(b^{(2)} - t < \sqrt{\alpha_n(1-\rho_n)}Y^{(2)}_h \leq b^{(2)} + t) \\
&\quad = P\Big((\frac{a^{(2)} - t}{\sqrt{\alpha_n(1-\rho_n)}} \leq Y^{(2)}_h < \frac{a^{(2)} + t}{\sqrt{\alpha_n(1-\rho_n)}}\Big) + P\Big(\frac{b^{(2)} - t}{\sqrt{\alpha_n(1-\rho_n)}} < Y^{(2)}_h \leq \frac{b^{(2)} + t}{\sqrt{\alpha_n(1-\rho_n)}}\Big) \\
&\quad \leq C_\beta \ \frac{t}{\underline{\sigma}_h \sqrt{\alpha_n(1-\rho_n)}} \sqrt{\log(p_2)} \leq C_\beta \ \frac{t}{\underline{\sigma}_h \sqrt{\alpha_n}} \sqrt{\log(p_2)} \\
&\quad \leq C_\beta \ \frac{N^{1/2} m^2 D_n \log(p_2 n)^{1/2+2/\beta}}{\underline{\sigma}_h n^{\min\{1/2+k,\, 5/6\}}} = C_\beta \omega_{n,2}.
\end{aligned}
$$

Importantly, Nazarov's inequality is applicable because $\mathbb{E}[Y^{(2)}_{h,j}] \geq \underline{\sigma}_h$ for all $j = 1, \ldots, p$ and $\underline{\sigma}_h > 0$ due to Condition (C3). The second to last inequality follows from the fact that $1 - \rho_n \geq 1/2$. We have

$$
\begin{aligned}
&P(\sqrt{n}(U_n + \sqrt{1-\rho_n}B_n) \in R) \leq P(m\overline{Y}_g \in [R - \sqrt{\alpha_n(1-\rho_n)}Y_h]) + C_\beta\{\omega_{n,1} + \omega_{n,2}\} \\
&\quad = P(m\overline{Y}_g + \sqrt{\alpha_n(1-\rho_n)}Y_h \in R) + C_\beta\{\omega_{n,1} + \omega_{n,2}\} \\
&\quad = P\left(\alpha_n^{-1/2}\Lambda_h^{-1/2} m\overline{Y}_g + \sqrt{1-\rho_n}\Lambda_h^{-1/2} Y_h \in \alpha_n^{-1/2}\Lambda_h^{-1/2} R\right) + C_\beta\{\omega_{n,1} + \omega_{n,2}\}.
\end{aligned}
$$

Observe that $\mathbb{E}[(\sqrt{1-\rho_n}\sigma^{-1}_{h,j}Y_{h,j})^2] = 1 - \rho_n \geq 1/2$ for any $1 \leq j \leq p$ and $\|\Lambda_h^{-1/2}\Gamma_h\Lambda_h^{-1/2}\|_\infty = 1$ by definition of Λ_h. Since $\rho_n = N/\binom{n}{m} \leq C_\beta N/n^m$ we have by the Gaussian comparison inequality (Chen and Kato, 2019, Lemma C.5)

$$
\begin{aligned}
&P(\sqrt{n}(U_n + \sqrt{1-\rho_n}B_n) \in R) \\
&\leq P\left(\alpha_n^{-1/2}\Lambda_h^{-1/2}(m\overline{Y}_g + \sqrt{\alpha_n}Y_h) \in \alpha_n^{-1/2}\Lambda_h^{-1/2} R\right) + C_\beta\{\omega_{n,1} + \omega_{n,2}\} + C_\beta\left(\frac{N\log(p)^2}{n^m}\right)^{1/3} \\
&\leq P(m\overline{Y}_g + \sqrt{\alpha_n}Y_h \in R) + C_\beta\{\omega_{n,1} + \omega_{n,2} + \omega_{n,3}\}
\end{aligned}
$$

Likewise, we can show

$$
P(\sqrt{n}(U_n + \sqrt{1-\rho_n}B_n) \in R) \geq P(m\overline{Y}_g + \sqrt{\alpha_n}Y_h \in R) - C_\beta\{\omega_{n,1} + \omega_{n,2} + \omega_{n,3}\}.
$$

It is left to show that the latter two inequalities hold with $\sqrt{n}(U_n + \sqrt{1-\rho_n}B_n)$ replaced by $\sqrt{n}U'_{n,N}$. This part of the proof is similar to step 5 in the proof of Theorem 3.1 in Chen and

Kato (2019) and steps 2 and 3 in the proof of Theorem 2.4 in Song et al. (2019) and therefore omitted. □

Proposition A.3. *Assume (C2), (C3) and (C5) hold. Then, for a constant $C_\beta > 0$ only depending on β, it holds* $\sup_{R\in\mathbb{R}^p_{\mathrm{re}}} |P(\overline{Y} \in R) - P(Y \in R)| \leq C_\beta \omega_{n,3}$.

For the proof of Proposition A.3, we define a $p \times p$ diagonal matrix Λ_h such that $\Lambda_{h,jj} = \sigma^2_{h,j}$ for $1 \leq j \leq p$.

Proof of Proposition A.3. The proof is an application of the Gaussian comparison inequality (Chen and Kato, 2019, Lemma C.5) with appropriate normalizing. Let $C_\beta > 0$ be a constant that depends only on β and that varies its value from place to place. Observe that $\mathbb{E}[(\alpha_n^{-1/2}\sigma_{h,j}^{-1}\overline{Y}_j)^2] \geq \mathbb{E}[(\sigma_{h,j}^{-1}Y_{h,j})^2] = 1$ for any $1 \leq j \leq p$ due to independence of $\overline{Y}_g$ and Y_h and Condition (C3). Moreover,

$$\begin{aligned}
&\|\alpha_n^{-1}\Lambda_h^{-1/2}\mathrm{Cov}[\overline{Y}]\Lambda_h^{-1/2} - \alpha_n^{-1}\Lambda_h^{-1/2}\mathrm{Cov}[Y]\Lambda_h^{-1/2}\|_\infty = \|\alpha_n^{-1}m^2\Lambda_h^{-1/2}(\overline{\Gamma}_g - \Gamma_g)\Lambda_h^{-1/2}\|_\infty \\
&\leq \frac{Nm^2}{\underline{\sigma}_h^2 n}\|\overline{\Gamma}_g - \Gamma_g\|_\infty \leq \frac{Nm^2}{\underline{\sigma}_h^2 n}\max\left\{\max_{p_1+1\leq j\leq p}\sigma^2_{g,j}, \left(\max_{1\leq j\leq p_1}\sigma_{g,j}\right)\cdot\left(\max_{p_1+1\leq j\leq p}\sigma_{g,j}\right)\right\} \\
&\leq C_\beta\frac{Nm^2}{\underline{\sigma}_h^2 n}\max\{n^{-2k}D_n^2, n^{-k}D_n^2\} \leq C_\beta\frac{Nm^2D_n^2}{\underline{\sigma}_h^2 n^{1+k}},
\end{aligned}$$

where we used that $\max_{1\leq j\leq p_1}\sigma_{g,j} \leq \max_{1\leq j\leq p_1}\sigma_{h,j} \leq C_\beta D_n$ and $\max_{p_1+1\leq j\leq p}\sigma_{g,j} \leq C_\beta n^{-k}D_n$ due to Conditions (C2) and (C5) and Lemma C.4. Note that Assumption (C2) also implies $\|g_j(X_1) - \mu_j\|_{\psi_\beta} \leq D_n$ for all $j = 1, \ldots, p$. Thus, we have by the Gaussian comparison inequality for any $R \in \mathbb{R}^p_{\mathrm{re}}$

$$\begin{aligned}
P(\overline{Y} \in R) &= P(\alpha_n^{-1/2}\Lambda_h^{-1/2}\overline{Y} \in \alpha_n^{-1/2}\Lambda_h^{-1/2}R) \\
&\leq P(\alpha_n^{-1/2}\Lambda_h^{-1/2}Y \in \alpha_n^{-1/2}\Lambda_h^{-1/2}R) + C_\beta\left(\frac{Nm^2D_n^2\log(p)^2}{\underline{\sigma}_h^2 n^{1+k}}\right)^{1/3} \\
&\leq P(Y \in R) + C_\beta\omega_{n,3}.
\end{aligned}$$

Similarly, we can show $P(\overline{Y} \in R) \geq P(Y \in R) - C_\beta\omega_{n,3}$, which concludes the proof. □

Proof of Theorem 2.4. The theorem follows from Proposition A.2 and Proposition A.3. □

A.2. Other Proofs of Section 2.

Proof of Corollary 2.7. Let $C_{s,m} > 0$ be a constant that depends only on s and m and that may change its value in different occurences. Since the polynomials h_j are not constant, we have that $s \geq 1$. Define $\beta = 1/s \in (0, 1]$ and $D_n = C_{s,m}\overline{\sigma}_h$. Since $\overline{\sigma}_h \geq \underline{\sigma}_h$, it holds that $D_n \geq 1$ by potentially increasing $C_{s,m}$. We will first prove that Assumptions (C1), (C2) and (C6) as well as mixed degeneracy are satisfied, where we fix an arbitrary $k \in (0, 1)$. This implies that the Berry-Esseen type bound of Theorem 2.4 holds. We then optimize this bound over all $k \in (0, 1)$ to obtain the desired result.

To show that (C1) holds, it is enough to note that $\|h_j(X_1^m)\|_{2+l} \leq C_{s,m}\sigma_{h,j}$ holds for $l = 1, 2$ due to the hypercontractivity property of polynomials in Gaussian random variables (Theorem 3.2.10 in de la Peña and Giné, 1999 and Lemma 2.2 in Leung and Sturma, 2024). To show (C6), observe that each g_j is also a polynomial in Gaussian variables of degree at most $2s$. Hence, it also holds that $\|g_j(X_1)\|_{2+l} \leq C_{s,m}\sigma_{g,j}$ for $l = 1, 2$, which implies that (C6) is also satisfied since $\sigma^2_{g,j} \leq \sigma^2_{h,j}$ by Jensen's inequality. Assumption (C2) follows from Lemma C.5.

It remains to show mixed degeneracy, that is, Assumptions (C4) and (C5). Fix an arbitrary $k \in (0, 1)$ and define $\underline{\sigma}^2_{g^{(1)}} = C_{s,m}n^{-2k}\overline{\sigma}^2_h$. Since each g_j is a polynomial in Gaussian variables of

degree at most $2s$, we have by Lemma C.5 that $\|g_j(X_1)\|_{\psi_{1/s}} \leq C_{s,m}\sigma_{g,j}$. Hence, if $\sigma^2_{g,j} < \underline{\sigma}^2_{g^{(1)}}$ for some index $j \in \{1,\dots,p\}$, then

$$\|g_j(X_1)\|_{\psi_{1/s}} \leq C_{s,m}\underline{\sigma}_{g^{(1)}} \leq C_{s,m}n^{-k}\,\overline{\sigma}_h.$$

That is, if (C4) is not satisfied, then (C5) must be satisfied for this index j, and we conclude that mixed degeneracy holds.

Now, Theorem 2.4 implies that

$$\sup_{R\in\mathbb{R}^p_{\mathrm{re}}} |P(\sqrt{n}U'_{n,N} \in R) - P(Y \in R)| \leq C_\beta\{\omega_{n,1} + \omega_{n,2} + \omega_{n,3}\}.$$

Since we have $\beta = 1/s$, $D_n = C_{s,m}\overline{\sigma}_h$ and $\underline{\sigma}^2_{g^{(1)}} = C_{s,m}n^{-2k}\,\overline{\sigma}^2_h$ with $k \in (0,1)$, we can further bound $\omega_{n,1}$, $\omega_{n,2}$ and $\omega_{n,3}$ as follows:

$$\begin{aligned}
\omega_{n,1} &\leq C_{s,m}\left(\frac{(\overline{\sigma}^2_h \vee 1)\log(pn)^{1+6s}}{(\underline{\sigma}^2_h \wedge 1)\,n^{1-k}}\right)^{1/6} \leq C_{s,m}\frac{(\overline{\sigma}^2_h \vee 1)\log(pn)^{1/6+s}}{(\underline{\sigma}^2_h \wedge 1)\,n^{(1-k)/6}},\\
\omega_{n,2} &\leq C_{s,m}\frac{(\overline{\sigma}^2_h \vee 1)\log(pn)^{1/2+2s}}{(\underline{\sigma}^2_h \wedge 1)\,n^{\min\{k,1/3\}}}, \text{ and}\\
\omega_{n,3} &\leq C_{s,m}\frac{(\overline{\sigma}^2_h \vee 1)\log(p)^{2/3}}{(\underline{\sigma}^2_h \wedge 1)\,n^{k/3}},
\end{aligned}$$

where we have also used that $n \leq N \leq C_{s,m}n$. Since $k \in (0,1)$, this implies the overall bound

$$\omega_{n,1} + \omega_{n,2} + \omega_{n,3} \leq C_{s,m}\frac{(\overline{\sigma}^2_h \vee 1)\log(p)^{1/2+2s}}{(\underline{\sigma}^2_h \wedge 1)\,n^{\min\{(1-k)/6,k/3\}}}.$$

It remains to choose $k \in (0,1)$ such that $\min\{(1-k)/6, k/3\}$ is maximal. This yields $k = 1/3$ and $\min\{(1-k)/6, k/3\} = 1/9$. □

Recall the definition of the diagonal matrix Λ_h, that is, $\Lambda_{h,jj} = \sigma^2_{h,j}$ for $1 \leq j \leq p$. We define the standardized kernel $\tilde{h}(x_1,\dots,x_m) = \Lambda_h^{-1/2}h(x_1,\dots,x_m)$. In accordance, let $\Gamma_{\tilde{h}} = \mathrm{Cov}[\tilde{h}(X_1^m)]$ be the covariance matrix of the standardized kernel and let $\tilde{U}'_{n,N} = \Lambda_h^{-1/2}U'_{n,N}$ be the standardized incomplete U-statistic.

Proof of Lemma 2.9. Throughout the proof let $C > 0$ be a constant only depending on β, ζ_1 and C_1 and that varies its value from place to place. It will be useful to note that

$$\max\left\{\frac{m^{2/\beta}}{\underline{\sigma}^4_h}, \frac{N^2m^4}{\underline{\sigma}^4_h n^2}, \frac{Nm^2}{\underline{\sigma}^2_h n}\right\} \leq A_n.$$

We start by observing $\mathbb{E}[(\alpha_n^{-1/2}\sigma_{h,j}^{-1}Y_j)^2] \geq \mathbb{E}[(\sigma_{h,j}^{-1}Y_{h,j})^2] \geq 1$ due to the independence of Y_g and Y_h and Condition (C3). Thus, by the Gaussian comparison inequality (Chen and Kato, 2019, Lemma C.5) we have

$$\begin{aligned}
&\sup_{R\in\mathbb{R}^p_{\mathrm{re}}} \left|P|_{\mathcal{D}_n}(U^{\#}_{n,n_1} \in R) - P(Y \in R)\right|\\
&\leq \sup_{R\in\mathbb{R}^p_{\mathrm{re}}} \left|P|_{\mathcal{D}_n}(U^{\#}_{n,n_1} \in R) - P|_{\mathcal{D}_n}(mU^{\#}_{n_1,g} + \sqrt{\alpha_n}Y_h \in R)\right|\\
&\quad + \sup_{R\in\mathbb{R}^p_{\mathrm{re}}} \left|P|_{\mathcal{D}_n}(mU^{\#}_{n_1,g} + \sqrt{\alpha_n}Y_h \in R) - P(Y \in R)\right|\\
&= \sup_{R\in\mathbb{R}^p_{\mathrm{re}}} \left|P|_{\mathcal{D}_n}(\alpha_n^{-1/2}\Lambda_h^{-1/2}U^{\#}_{n,n_1} \in R) - P|_{\mathcal{D}_n}\left(\Lambda_h^{-1/2}(\alpha_n^{-1/2}mU^{\#}_{n_1,g} + Y_h) \in R\right)\right|\\
&\quad + \sup_{R\in\mathbb{R}^p_{\mathrm{re}}} \left|P|_{\mathcal{D}_n}\left(\Lambda_h^{-1/2}(\alpha_n^{-1/2}mU^{\#}_{n_1,g} + Y_h) \in R\right) - P(\alpha_n^{-1/2}\Lambda_h^{-1/2}Y \in R)\right|
\end{aligned}$$

$$\leq C\left(\widehat{\Delta}_{\tilde{h}}\log(p)^2\right)^{1/3} + C\left(\frac{Nm^2}{\underline{\sigma}_{\tilde{h}}^2 n}\widehat{\Delta}_g\log(p)^2\right)^{1/3},$$

where

$$\widehat{\Delta}_{\tilde{h}} = \left\| \frac{1}{\widehat{N}} \sum_{\iota\in I_{n,m}} Z_\iota(\tilde{h}(X_\iota) - \tilde{U}'_{n,N})(\tilde{h}(X_\iota) - \tilde{U}'_{n,N})^\top - \Gamma_{\tilde{h}} \right\|_\infty$$

and

$$\widehat{\Delta}_g = \left\| \frac{1}{n_1} \sum_{i_1\in S_1} (G_{i_1} - \overline{G})(G_{i_1} - \overline{G})^\top - \Gamma_g \right\|_\infty .$$

Now, it is enough to show for each of the two summands involving $\widehat{\Delta}_{\tilde{h}}$ and $\widehat{\Delta}_g$ that the probability of being greater than $Cn^{-(\zeta_1\wedge\zeta_2)/6}$ is at most C/n. We start with the first summand.

Step 1: Bounding $\widehat{\Delta}_{\tilde{h}}$. This is similar to the proof of Theorem 4.1 in Chen and Kato (2019) and to the proof of Theorem 3.1 in Song et al. (2019). However, we give a full proof for completeness verifying that everything remains true under our assumptions even though some parts will be identical. In particular, we track the constants β, m and $\underline{\sigma}_{\tilde{h}}^2$. Observe that for any integer l, there exists some constant $\tilde{C}$ that depends only on l and ζ_1 such that

$$\frac{\log(n)^l}{n^{\zeta_1}} \leq \tilde{C} \tag{A.4}$$

for all $n \geq 1$. Define

$$\widehat{\Delta}_{\tilde{h},1} = \left\| \frac{1}{N} \sum_{\iota\in I_{n,m}} (Z_\iota - \rho_n)\tilde{h}(X_\iota)\tilde{h}(X_\iota)^\top \right\|_\infty, \quad \widehat{\Delta}_{\tilde{h},2} = \left\| \widehat{\Gamma}_{\tilde{h}} - \Gamma_{\tilde{h}} \right\|_\infty,$$

$$\widehat{\Delta}_{\tilde{h},3} = |N/\widehat{N} - 1|\|\Gamma_{\tilde{h}}\|_\infty \quad \text{and} \quad \widehat{\Delta}_{\tilde{h},4} = \left\| \frac{1}{N} \sum_{\iota\in I_{n,m}} Z_\iota\tilde{h}(X_\iota) \right\|_\infty^2,$$

where $\widehat{\Gamma}_{\tilde{h}} = |I_{n,m}|^{-1}\sum_{\iota\in I_{n,m}} \tilde{h}(X_\iota)\tilde{h}(X_\iota)^\top$ and observe that

$$\widehat{\Delta}_{\tilde{h}} \leq |N/\widehat{N}|\left(\widehat{\Delta}_{\tilde{h},1} + \widehat{\Delta}_{\tilde{h},2}\right) + \widehat{\Delta}_{\tilde{h},3} + |N/\widehat{N}|^2\widehat{\Delta}_{\tilde{h},4}.$$

First, we consider the quantity $|N/\widehat{N}|$. We can assume without loss of generality $C_1 n^{-\zeta_1} \leq 1/16$ since otherwise we may always take C to be large enough and the statement of Lemma 2.9 becomes trivial. Hence, by Condition (2.5), we have $\sqrt{\log(n)/N} \leq (C_1 n^{-\zeta_1})^{1/2} \leq 1/4$ and thus it follows by Lemma A.12 in Song et al. (2019) that $P(|N/\widehat{N} - 1| > C) \leq 2/n$. Since $|N/\widehat{N} - 1| \geq |N/\widehat{N}| - 1$ we also have $P(|N/\widehat{N}| > C) \leq 2/n$. Therefore it suffices to show

$$P\left(\widehat{\Delta}_{\tilde{h},i}\log(p)^2 > Cn^{-\zeta_1/2}\right) \leq \frac{C}{n}$$

for all $i = 1, \ldots, 4$, which we naturally divide into four sub-steps.

Step 1.1: Bounding $\widehat{\Delta}_{\tilde{h},1}$. Conditioned on $\{X_1, \ldots, X_n\}$ we have by Lemma A.3 in Song et al. (2019) the inequality

$$P|_{X_1^n}\left(N\widehat{\Delta}_{\tilde{h},1} > C(\sqrt{NV_n\log(pn)} + M_1\log(pn))\right) \leq \frac{C}{n}, \tag{A.5}$$

where $V_n = \max_{1\leq j,l\leq p}|I_{n,m}|^{-1}\sum_{\iota\in I_{n,m}}\tilde{h}_j(X_\iota)^2\tilde{h}_l(X_\iota)^2$ and $M_1 = \max_{\iota\in I_{n,m}}\max_{1\leq j\leq p}\tilde{h}_j(X_\iota)^2$. Next we shall find bounds such that the probability for V_n and M_1 being larger than

the bounds is at most C/n. First, by Lemma C.6, Lemma C.2 and due to Assumption (C2),

$$\begin{aligned}\|M_1\|_{\psi_{\beta/2}} &\leq \log(p|I_{n,m}|)^{2/\beta} \max_{\iota\in I_{n,m}} \max_{1\leq j\leq p} \|\tilde{h}_j(X_1^m)^2\|_{\psi_{\beta/2}} \\ &\leq C\underline{\sigma}_h^{-2} \log(pn^m)^{2/\beta} \max_{\iota\in I_{n,m}} \max_{1\leq j\leq p} \|h_j(X_1^m)\|_{\psi_\beta}^2 \\ &\leq C\underline{\sigma}_h^{-2} m^{2/\beta} \log(pn)^{2/\beta} D_n^2.\end{aligned}$$

Thus, by the definition of $\|\cdot\|_{\psi_{\beta/2}}$ together with the Markov inequality,

$$P(M_1 > C\underline{\sigma}_h^{-2} m^{2/\beta} D_n^2 \log(pn)^{2/\beta} \log(n)^{2/\beta}) \leq 2/n.$$

Second, we apply Lemma A.6 in Song et al. (2019) to bound V_n. Due to Assumption (C1) we have

$$\max_{1\leq j,l\leq p} \mathbb{E}[\tilde{h}_j(X_1^m)^2 \tilde{h}_l(X_1^m)^2] \leq \underline{\sigma}_h^{-2} D_n^2$$

and due to Assumption (C2) and Lemma C.2 it holds

$$\max_{1\leq j,l\leq p} \|\tilde{h}_j(X_1^m)^2 \tilde{h}_l(X_1^m)^2\|_{\psi_{\beta/4}} \leq \underline{\sigma}_h^{-4} \max_{1\leq j\leq p} \|h_j(X_1^m)\|_{\psi_\beta}^4 \leq \underline{\sigma}_h^{-4} D_n^4.$$

By Condition (2.5) and Observation (A.4) we have

$$\begin{aligned}&\frac{m D_n^2 \log(pn)^{1+4/\beta} \log(n)^{4/\beta-1}}{\underline{\sigma}_h^2 n} + \frac{m^{2/\beta} D_n^2 \log(pn)^{1+8/\beta} \log(n)^{8/\beta-1}}{\underline{\sigma}_h^2 n^2} \\ &\leq C_1 n^{-\zeta_1} \log(n)^{4/\beta-1} + C_1 n^{-2\zeta_1} \log(n)^{8/\beta-1} \leq C\end{aligned}$$

and we obtain the bound $P(V_n > C\underline{\sigma}_h^{-2} D_n^2) \leq \frac{8}{n}$ from Song et al. (2019, Lemma A.6). Here, we used that $|I_{n,m}|^{-1} \leq Cn^{-2}m^{-1/\beta}$; cf. Song et al. (2019, Lemma A.9). Putting the results together and recalling (A.5) we have by Fubini

$$P\left(\widehat{\Delta}_{\tilde{h},1} > C(N^{-1/2}\underline{\sigma}_h^{-1} D_n \log(pn)^{1/2} + N^{-1}\underline{\sigma}_h^{-2} m^{2/\beta} D_n^2 \log(pn)^{1+4/\beta})\right) \leq \frac{C}{n}$$

Now, we use again Condition (2.5) to observe

$$\begin{aligned}&\log(p)^2 C\left(\frac{D_n \log(pn)^{1/2}}{\underline{\sigma}_h N^{1/2}} + \frac{m^{2/\beta} D_n^2 \log(pn)^{1+4/\beta}}{\underline{\sigma}_h^2 N}\right) \\ &\leq C\left(\left(\frac{D_n^2 \log(pn)^5}{\underline{\sigma}_h^2 N}\right)^{1/2} + \frac{m^{2/\beta} D_n^2 \log(pn)^{3+4/\beta}}{\underline{\sigma}_h^2 N}\right) \\ &\leq C\left(n^{-\zeta_1/2} + n^{-\zeta_1}\right) \leq Cn^{-\zeta_1/2},\end{aligned}$$

which implies

$$P\left(\widehat{\Delta}_{\tilde{h},1} \log(p)^2 > Cn^{-\zeta_1/2}\right) \leq \frac{C}{n}.$$

Step 1.2: Bounding $\widehat{\Delta}_{\tilde{h},2}$. By Lemma A.5 in Song et al. (2019) we have

$$P\left(\widehat{\Delta}_{\tilde{h},2} > C((n^{-1}m\kappa^2 \log(pn))^{1/2} + n^{-1}m z_n \log(pn)^{4/\beta})\right) \leq \frac{4}{n},$$

where

$$\begin{aligned}\kappa^2 &= \max_{1\leq j,l\leq p} \mathbb{E}[\{\tilde{h}_j(X_1^m)\tilde{h}_l(X_1^m) - \Gamma_{\tilde{h},jl}\}^2], \\ z_n &= \max_{1\leq j,l\leq p} \|\tilde{h}_j(X_1^m)\tilde{h}_l(X_1^m) - \Gamma_{\tilde{h},jl}\|_{\psi_{\beta/2}}.\end{aligned}$$

Due to Condition (C1) we have the bound $\kappa^2 \leq \max_{1\leq j,l\leq p} \mathbb{E}[\tilde{h}_j(X_1^m)^2\tilde{h}_l(X_1^m)^2] \leq \underline{\sigma}_h^{-2} D_n^2$ and due to Condition (C2) and Lemmas C.7 and C.2 we have $z_n \leq C\underline{\sigma}_h^{-2} D_n^2$. Thus,

$$P\left(\widehat{\Delta}_{\tilde{h},2} > C\left(\left(\frac{mD_n^2\log(pn)}{\underline{\sigma}_h^2 n}\right)^{1/2} + \frac{mD_n^2\log(pn)^{4/\beta}}{\underline{\sigma}_h^2 n}\right)\right) \leq \frac{4}{n}.$$

Using Condition (2.5) we conclude by the same arguments as in Step 1.1,

$$P(\widehat{\Delta}_{\tilde{h},2}\log(p)^2 > Cn^{-\zeta_1/2}) \leq C/n.$$

Step 1.3: Bounding $\widehat{\Delta}_{\tilde{h},3}$. By Lemma A.12 in Song et al. (2019) we have

$$P\left(|N/\widehat{N} - 1| > C\sqrt{\log(n)/N}\right) \leq \frac{2}{n}.$$

Since $\|\Gamma_{\tilde{h}}\|_\infty = 1$ this yields

$$P\left(\widehat{\Delta}_{\tilde{h},3}\log(p)^2 > C\left(\frac{\log(n)\log(p)^4}{N}\right)^{1/2}\right) \leq \frac{2}{n}.$$

Observing $\frac{\log(n)\log(p)^4}{N} \leq C_1 n^{-\zeta_1}$ by Condition (2.5) we conclude

$$P\left(\widehat{\Delta}_{\tilde{h},3}\log(p)^2 > Cn^{-\zeta_1/2}\right) \leq \frac{C}{n}.$$

Step 1.4: Bounding $\widehat{\Delta}_{\tilde{h},4}$. Observe that $\widehat{\Delta}_{\tilde{h},4} \leq 2(\widehat{\Delta}_{\tilde{h},5}^2 + \widehat{\Delta}_{\tilde{h},6}^2)$, where

$$\widehat{\Delta}_{\tilde{h},5} = \left\|\frac{1}{N}\sum_{\iota\in I_{n,m}}(Z_\iota - \rho_n)\tilde{h}(X_\iota)\right\|_\infty \quad \text{and} \quad \widehat{\Delta}_{\tilde{h},6} = \left\|\frac{1}{|I_{n,m}|}\sum_{\iota\in I_{n,m}}\tilde{h}(X_\iota)\right\|_\infty$$

such that it is enough to bound the two terms $\widehat{\Delta}_{\tilde{h},5}^2$ and $\widehat{\Delta}_{\tilde{h},6}^2$ separately. But bounding $\widehat{\Delta}_{\tilde{h},5}$ is similar to bounding $\widehat{\Delta}_{\tilde{h},1}$ by applying Song et al. (2019, Lemma A.3). Moreover, bounding $\widehat{\Delta}_{\tilde{h},6}$ is similar to bounding $\widehat{\Delta}_{\tilde{h},2}$, where we applied Song et al. (2019, Lemma A.5). Therefore, we omit the details.

Step 2: Bounding $\widehat{\Delta}_g$. By the same argument as in the proof of Theorem 4.2 in Chen and Kato (2019) and using $\max_{1\leq j\leq p}\sigma_{g,j} \leq CD_n$, which holds true due to Condition (C2) and Lemma C.4, we have

$$\widehat{\Delta}_g \leq C(D_n\widehat{\Delta}_{g,1}^{1/2} + \widehat{\Delta}_{g,1} + \widehat{\Delta}_{g,2} + \widehat{\Delta}_{g,3}^2),$$

where $\widehat{\Delta}_{g,1}$ is defined in (2.4) and

$$\widehat{\Delta}_{g,2} = \left\|\frac{1}{n_1}\sum_{i_1\in S_1}\{g(X_{i_1})g(X_{i_1})^\top - \Gamma_g\}\right\|_\infty, \quad \widehat{\Delta}_{g,3} = \left\|\frac{1}{n_1}\sum_{i_1\in S_1} g(X_{i_1})\right\|_\infty.$$

Thus, similar to Step 1, we divide the proof into three sub-steps.

Step 2.1: Bounding $\widehat{\Delta}_{g,1}$. By Condition (2.6) we have

$$P\left(\frac{Nm^2}{\underline{\sigma}_h^2 n}D_n\widehat{\Delta}_{g,1}^{1/2}\log(p)^2 > C_1^{1/2}n^{-\zeta_2/2}\right) \leq \frac{C}{n}$$

and

$$P\left(\frac{Nm^2}{\underline{\sigma}_h^2 n}\widehat{\Delta}_{g,1}\log(p)^2 > C_1 n^{-\zeta_2}\right) \leq \frac{C}{n},$$

where we used $D_n \geq 1$ and $p \geq 3$ for the second bound.

Step 2.2: Bounding $\widehat{\Delta}_{g,2}$. By Lemma A.2 in Song et al. (2019) we have

$$P\left(\widehat{\Delta}_{g,2} > C(n_1^{-1}\nu\log(pn)^{1/2} + n_1^{-1}u_n\log(pn)^{4/\beta})\right) \le \frac{4}{n},$$

where

$$\nu^2 = \max_{1\le j,l\le p}\sum_{i_1\in S_1}\mathbb{E}[\{g_j(X_{i_1})g_l(X_{i_1}) - \Gamma_{g,jl}\}^2],$$
$$u_n = \max_{i_1\in S_1}\max_{1\le j,l\le p}\|g_j(X_{i_1})g_l(X_{i_1}) - \Gamma_{g,jl}\|_{\psi_{\beta/2}}.$$

Due to Condition (C2) and Lemmas C.2, C.4 and C.7 we have the bounds $\nu^2 \le Cn_1D_n^4$ and $u_n \le CD_n^2$. Thus,

$$P\left(\widehat{\Delta}_{g,2} > C(n_1^{-1/2}D_n^2\log(pn)^{1/2} + n_1^{-1}D_n^2\log(pn)^{4/\beta})\right) \le \frac{4}{n}.$$

Now, we use Condition (2.5) to observe

$$\begin{aligned}
&\frac{Nm^2}{\underline{\sigma}_h^2 n}\log(p)^2 C\left(\frac{D_n^2\log(pn)^{1/2}}{n_1^{1/2}} + \frac{D_n^2\log(pn)^{4/\beta}}{n_1}\right)\\
&\le C\left(\left(\frac{N^2m^4D_n^4\log(pn)^5}{\underline{\sigma}_h^4 n^2 n_1}\right)^{1/2} + \frac{Nm^2D_n^2\log(pn)^{2+4/\beta}}{\underline{\sigma}_h^2 n n_1}\right)\\
&\le C\left(n^{-\zeta_1/2} + n^{-\zeta_1}\right) \le Cn^{-\zeta_1/2}.
\end{aligned}$$

Thus we conclude

$$P\left(\frac{Nm^2}{\underline{\sigma}_h^2 n}\widehat{\Delta}_{g,2}\log(p)^2 > Cn^{-\zeta_1/2}\right) \le \frac{4}{n} \le \frac{C}{n}.$$

Step 2.3: Bounding $\widehat{\Delta}_{g,3}$. As in Step 2.2 we apply Lemma A.2 in Song et al. (2019) once again and obtain

$$P\left(\widehat{\Delta}_{g,3} > C(n_1^{-1}\tilde{\nu}\log(pn)^{1/2} + n_1^{-1}\tilde{u}_n\log(pn)^{2/\beta})\right) \le \frac{4}{n},$$

where $\tilde{\nu}^2 = \max_{1\le j\le p}\sum_{i_1\in S_1}\mathbb{E}[g_j(X_{i_1})^2]$ and $\tilde{u}_n = \max_{i_1\in S_1}\max_{1\le j\le p}\|g_j(X_{i_1})\|_{\psi_\beta}$. Due to Assumption (C2) and Lemma C.4 and Lemma C.7, we have the bounds $\tilde{\nu}^2 \le Cn_1D_n^2$ and $\tilde{u}_n \le CD_n$ giving

$$P\left(\widehat{\Delta}_{g,3}^2 > C(n_1^{-1/2}D_n\log(pn)^{1/2} + n_1^{-1}D_n\log(pn)^{2/\beta})\right) \le \frac{4}{n}.$$

By the same arguments as in Step 2.2 we conclude

$$P\left(\frac{Nm^2}{\underline{\sigma}_h^2 n}\widehat{\Delta}_{g,3}^2\log(p)^2 > Cn^{-\zeta_1/2}\right) \le \frac{C}{n}. \qquad \square$$

Proof of Theorem 2.10. Throughout the proof let $C > 0$ be a constant only depending on β, ν, ζ and C_1 and that varies its value from place to place. Let $\zeta_1 = \zeta$ and $\zeta_2 = \zeta - 1/\nu$. Then, due to Lemma 2.9, it suffices to show that (2.6) holds. By Step 2 of the proof of Theorem 3.3 in Song et al. (2019) we have

$$\mathbb{E}[\widehat{\Delta}_{g,1}^\nu] \le C\left(\frac{mD_n^2\log(p)^{1+2/\beta}}{n}\right)^\nu.$$

Then we have by the Markov inequality

$$P(A_nD_n^2\widehat{\Delta}_{g,1}\log(p)^4 > C_1n^{-(\zeta-1/\nu)})$$

$$\leq CA_n^{\nu} D_n^{2\nu} \log(p)^{4\nu} n^{(\zeta-1/\nu)\nu} \left(\frac{mD_n^2 \log(p)^{1+2/\beta}}{n} \right)^{\nu}$$
$$= Cn^{-1} \left(\frac{n^{\zeta} A_n D_n^4 \log(p)^{5+2/\beta}}{n} \right)^{\nu}$$
$$\leq Cn^{-1} \left(n^{\zeta} C_1 n^{-\zeta} \right)^{\nu} = \frac{C}{n},$$

where the last inequality follows from Condition (2.8). □

Proof of Corollary 2.12. Throughout the proof let $C > 0$ be a constant only depending on β, ζ and C_1 and that varies its value from place to place. We begin by bounding the quantity $\max_{1\leq j\leq p} |\hat{\sigma}_j^2/\sigma_j^2 - 1|$. Note that

$$\left| \frac{\hat{\sigma}_j^2}{\sigma_j^2} - 1 \right| = \left| \frac{m^2(\hat{\sigma}_{g,j}^2 - \sigma_{g,j}^2) + \alpha_n(\hat{\sigma}_{h,j}^2 - \sigma_{h,j}^2)}{m^2\sigma_{g,j}^2 + \alpha_n\sigma_{h,j}^2} \right|$$
$$\leq \frac{m^2|\hat{\sigma}_{g,j}^2 - \sigma_{g,j}^2|}{m^2\sigma_{g,j}^2 + \alpha_n\sigma_{h,j}^2} + \frac{\alpha_n|\hat{\sigma}_{h,j}^2 - \sigma_{h,j}^2|}{m^2\sigma_{g,j}^2 + \alpha_n\sigma_{h,j}^2}$$
$$\leq \frac{m^2}{\alpha_n\sigma_{h,j}^2}|\hat{\sigma}_{g,j}^2 - \sigma_{g,j}^2| + \frac{1}{\sigma_{h,j}^2}|\hat{\sigma}_{h,j}^2 - \sigma_{h,j}^2|.$$

Taking the maximum we get

$$\max_{1\leq j\leq p} \left| \frac{\hat{\sigma}_j^2}{\sigma_j^2} - 1 \right| \leq \frac{Nm^2}{\varrho_h^2 n} \max_{1\leq j\leq p} |\hat{\sigma}_{g,j}^2 - \sigma_{g,j}^2| + \max_{1\leq j\leq p} \frac{1}{\sigma_{h,j}^2} |\hat{\sigma}_{h,j}^2 - \sigma_{h,j}^2| \leq \frac{Nm^2}{\varrho_h^2 n} \widehat{\Delta}_g + \widehat{\Delta}_{\tilde{h}},$$

where $\widehat{\Delta}_g$ and $\widehat{\Delta}_{\tilde{h}}$ were defined in the proof of Lemma 2.9. We have shown in Step 1 in the proof of Lemma 2.9 that

$$P\left(\widehat{\Delta}_{\tilde{h}} \log(p)^2 > Cn^{-\zeta/2} \right) \leq \frac{C}{n}.$$

Moreover, if we take $\nu = 7/\zeta$ and $\zeta_2 = \zeta - 1/\nu$, then in the proofs of Lemma 2.9 and Theorem 2.10 we have shown that

$$P\left(\frac{Nm^2}{\varrho_h^{-2} n} \widehat{\Delta}_g \log(p)^2 > Cn^{-3\zeta/7} \right) \leq \frac{C}{n}.$$

Hence, we have

$$P\left(\max_{1\leq j\leq p} |\hat{\sigma}_j^2/\sigma_j^2 - 1| \log(p)^2 > Cn^{-3\zeta/7} \right) \leq \frac{C}{n}. \tag{A.6}$$

The rest of the proof is the same as the proof for Corollary A.1 in Chen and Kato (2019) and thus omitted. □

A.3. **Proofs of Section 3.**

Proof of Corollary 3.1. Throughout the proof let $C > 0$ be a constant only depending on β, ζ and C_1 and that varies its value from place to place. Moreover, we write $\mu_j = f_j(\theta)$ and for $t \in \mathbb{R}$ we consider hyperrectangles of the form $H_t = \{x \in \mathbb{R}^p : -\infty \leq x_j \leq t \text{ for all } j = 1, \ldots, p\}$. The last steps in the proof of Corollary 2.12 are equal to the last steps in the proof of Chen and Kato (2019, Corollary A.1) and they show that we have the bound

$$\sup_{t\in\mathbb{R}} \left| P\left(\max_{1\leq j\leq p} \sqrt{n}\, (U'_{n,N,j} - \mu_j)/\hat{\sigma}_j \leq t \right) - P(\max_{1\leq j\leq p} Y_j/\sigma_j \leq t) \right| \leq Cn^{-\zeta/7}, \tag{A.7}$$

where $Y \sim N_p(m^2\Gamma_g + \alpha_n\Gamma_h)$. Moreover, with probability at least $1 - C/n$ we have

$$\sup_{t\in\mathbb{R}} \left| P(\max_{1\le j\le p} Y_j/\sigma_j \le t) - P|_{\mathcal{D}_n}(W \le t) \right| \le Cn^{-\zeta/7}. \tag{A.8}$$

Now, let $\alpha \in (0,1)$. To simplify notation we write $T_0 = \max_{1\le j\le p} \sqrt{n}\ (U'_{n,N,j} - \mu_j)/\hat{\sigma}_j$ and $Y_0 = \max_{1\le j\le p} Y_j/\sigma_j$. Moreover, we denote by $c_{Y_0}(1-\alpha)$ the $(1-\alpha)$-quantile of Y_0. To begin, we use (A.8) to establish a relation between the quantiles of Y_0 and W. With probability at least $1 - C/n$ we have

$$\begin{aligned} P|_{\mathcal{D}_n}(W \le c_{Y_0}(1-\alpha-Cn^{-\zeta/7})) &\le P(Y_0 \le c_{Y_0}(1-\alpha-Cn^{-\zeta/7})) + Cn^{-\zeta/7} \\ &= 1-\alpha-Cn^{-\zeta/7} + Cn^{-\zeta/7} \\ &= 1-\alpha. \end{aligned}$$

Thus, by definition of a quantile, we have the relation

$$P(c_W(1-\alpha) \ge c_{Y_0}(1-\alpha-Cn^{-\zeta/7})) \ge 1 - C/n. \tag{A.9}$$

Similarly, we can show $P|_{\mathcal{D}_n}(W \le c_{Y_0}(1-\alpha+Cn^{-\zeta/7})) \ge 1-\alpha$ and therefore we have

$$P(c_W(1-\alpha) \le c_{Y_0}(1-\alpha+Cn^{-\zeta/7})) \ge 1 - C/n. \tag{A.10}$$

Now, we apply (A.9) and (A.7) to conclude

$$\begin{aligned} P(T_0 > c_W(1-\alpha)) &\le P(T_0 > c_{Y_0}(1-\alpha-Cn^{-\zeta/7})) + C/n \\ &\le P(Y_0 > c_{Y_0}(1-\alpha-Cn^{-\zeta/7})) + Cn^{-\zeta/7} + C/n \\ &= \alpha + Cn^{-\zeta/7} + Cn^{-\zeta/7} + C/n \\ &= \alpha + Cn^{-\zeta/7}, \end{aligned}$$

where we used that Y_0 is a continuous distribution with no point mass. The constant C may have changed its value in the last step. Equivalently, by using the relation (A.10), we get the other direction

$$\begin{aligned} P(T_0 > c_W(1-\alpha)) &\ge P(T_0 > c_{Y_0}(1-\alpha+Cn^{-\zeta/7})) - C/n \\ &\ge P(Y_0 > c_{Y_0}(1-\alpha+Cn^{-\zeta/7})) - Cn^{-\zeta/7} - C/n \\ &= \alpha - Cn^{-\zeta/7} - Cn^{-\zeta/7} - C/n \\ &= \alpha - Cn^{-\zeta/7}. \end{aligned}$$

Noting that both inequalities hold uniformly over all $\alpha \in (0,1)$ finishes the proof. □

Proof of Proposition 3.5. Throughout the proof let $C > 0$ be a constant only depending on β, ζ and C_1, that varies its value from place to place. Write $\mu_j = f_j(\theta)$ and let $j^* \in \{1 \le j \le p\}$ be any index such that $\mu_{j^*}/\sigma_{j^*} = \max_{1\le j\le p}(\mu_j/\sigma_j)$. To simplify notation, we also define $\kappa_{p,\alpha} = \sqrt{2\log(p)} + \sqrt{2\log(1/\alpha)}$ and $\delta = Cn^{-3\zeta/7}\log(p)^{-2}$ such that Inequality (3.2) is equivalent to

$$\mu_{j^*}/\sigma_{j^*} \ge (1+\varepsilon)(1+\delta)\frac{\kappa_{p,\alpha}}{\sqrt{n}}.$$

We assume without loss of generality $\delta \le 1/2$, since otherwise the statement from Corollary 3.5 becomes trivial by choosing C large enough. Now, let A be the event such that $|\widehat{\sigma}_{j^*}/\sigma_{j^*} - 1| \le \delta$. Then, we have

$$\begin{aligned} &P\left(\mathcal{T} > c_W(1-\alpha)\right) \\ &= P\left(\sqrt{n}\, U_{n,N,j^*}/\widehat{\sigma}_{j^*} > c_W(1-\alpha)\right) \\ &= P\left(\sqrt{n}\,\mu_{j^*}/\widehat{\sigma}_{j^*} + \sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\widehat{\sigma}_{j^*} > c_W(1-\alpha)\right) \end{aligned}$$

$$
\begin{aligned}
&= P\left((\sigma_{j^*}/\widehat{\sigma}_{j^*})\sqrt{n}\,\mu_{j^*}/\sigma_{j^*} + \sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\widehat{\sigma}_{j^*} > c_W(1-\alpha)\right)\\
&\geq P\left((1/(1+\delta))\sqrt{n}\,\mu_{j^*}/\sigma_{j^*} + \sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\widehat{\sigma}_{j^*} > c_W(1-\alpha)\right) - P\left(A^c\right)\\
&\geq P\left((1+\varepsilon)\kappa_{p,\alpha} + \sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\widehat{\sigma}_{j^*} > c_W(1-\alpha)\right) - P\left(A^c\right),
\end{aligned}
$$

where we used that $c_W(1-\alpha) \geq 0$, since $\alpha \in (0, 1/2)$. Furthermore, by Chernozhukov et al. (2019, Lemma D.4) we have the bound $c_W(1-\alpha) \leq \kappa_{p,\alpha}$. Hence, it follows that

$$
\begin{aligned}
P\left(\mathcal{T} > c_W(1-\alpha)\right) &\geq P\left(\sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\widehat{\sigma}_{j^*} > -\varepsilon\,\kappa_{p,\alpha}\right) - P\left(A^c\right)\\
&= P\left((\sigma_{j^*}/\widehat{\sigma}_{j^*})\sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\sigma_{j^*} > -\varepsilon\,\kappa_{p,\alpha}\right) - P\left(A^c\right)\\
&\geq P\left(\sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\sigma_{j^*} > -(1-\delta)\varepsilon\,\kappa_{p,\alpha}\right) - P\left(A^c\right)\\
&= 1 - P\left(\sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\sigma_{j^*} \leq -(1-\delta)\varepsilon\,\kappa_{p,\alpha}\right) - P\left(A^c\right).
\end{aligned}
$$

By Theorem 2.4, we have

$$
\begin{aligned}
&P\left(\sqrt{n}\,(U_{n,N,j^*} - \mu_{j^*})/\sigma_{j^*} \leq -(1-\delta)\varepsilon\,\kappa_{p,\alpha}\right)\\
&\leq P\left(Y_{j^*}/\sigma_{j^*} \leq -(1-\delta)\varepsilon\,\kappa_{p,\alpha}\right) + C\{\omega_{n,1} + \omega_{n,2} + \omega_{n,3}\}\\
&\leq P\left(Y_{j^*}/\sigma_{j^*} \leq -(1-\delta)\varepsilon\,\kappa_{p,\alpha}\right) + Cn^{-\zeta/7}.
\end{aligned}
$$

Moreover, by Equation A.6 in the proof of Corollary 2.12, we have

$$
P\left(A^c\right) = P\left(|\widehat{\sigma}_{j^*}/\sigma_{j^*} - 1| > \delta\right) \leq \frac{C}{n} \leq Cn^{-\zeta/7}.
$$

This implies that

$$
P\left(\mathcal{T} > c_W(1-\alpha)\right) \geq 1 - P\left(Y_{j^*}/\sigma_{j^*} \geq (1-\delta)\varepsilon\,\kappa_{p,\alpha}\right) - Cn^{-\zeta/7}.
$$

Finally, by the Chernoff bound for univariate normal distributions, we have

$$
\begin{aligned}
P\left(Y_{j^*}/\sigma_{j^*} \geq (1-\delta)\varepsilon\,\kappa_{p,\alpha}\right) &\leq \exp\left(-\frac{1}{2}(1-\delta)^2\varepsilon^2\,\kappa_{p,\alpha}^2\right)\\
&\leq \exp\left(-\frac{1}{2}(1-\delta)^2\varepsilon^2\,(2\log(p) + 2\log(1/\alpha))\right)\\
&= \exp\left(-(1-\delta)^2\varepsilon^2\,\log(p/\alpha)\right)\\
&\leq \exp\left(-\frac{1}{C}\,\varepsilon^2\,\log(p/\alpha)\right),
\end{aligned}
$$

where we used $\delta \leq 1/2$ in the last step. □

A.4. **Proofs of Section 4.**

Proof of Proposition 4.3. By Step 3) in the construction of the kernel h, we have that

$$
\begin{aligned}
h(X_1^m) &= \frac{1}{m!}\sum_{\pi\in S_m} \breve{h}(X_{\pi(1)}, \ldots, X_{\pi(m)})\\
&= \frac{1}{m}\sum_{k=1}^{m} \underbrace{\frac{1}{(m-1)!}\sum_{\pi^*\in S_{m-1}^*} \breve{h}(X_{\pi^*(2)}, \ldots, X_{\pi^*(k-1)}, X_1, X_{\pi^*(k)}, \ldots, X_{\pi^*(m)})}_{=:\breve{h}^{(k)}(X_1^m)}\\
&= \frac{1}{m}\sum_{k=1}^{m} \breve{h}^{(k)}(X_1^m),
\end{aligned}
$$

where S_{m-1}^* denotes the group of all permutations of the set $\{2,\dots,m\}$. Now, let $k \leq m$ be a fixed integer and define $K = \lceil k/\eta \rceil$. Then we can write $k = (K-1)\eta + l$ for $1 \leq l \leq \eta$. Applying the definition of $\breve{h}$ in step 2) of the construction, we see that

$$\begin{aligned}
\mathbb{E}[\breve{h}^{(k)}(X_1^m)|X_1] &= \underbrace{a_0 + \sum_{r=1}^{K-1} \sum_{\substack{(i_1,\dots,i_r) \\ i_j \in \{1,\dots,d\}}} a_{(i_1,\dots,i_r)} \theta_{i_1} \cdots \theta_{i_r}}_{=:R^{(K-1)}(\theta)} \\
&\quad + \sum_{r=K}^{s} \sum_{\substack{(i_1,\dots,i_r) \\ i_j \in \{1,\dots,d\}}} a_{(i_1,\dots,i_r)} \theta_{i_1} \cdots \theta_{i_{K-1}} \hat{\theta}_{i_K,l}(X_1) \theta_{i_{K+1}} \cdots \theta_{i_r} \\
&= R^{(K-1)}(\theta) + \sum_{i_K \in D(f)} \hat{\theta}_{i_K,l}(X_1) \tilde{g}_{i_K}^{(K)}(\theta),
\end{aligned}$$

where $R^{(K-1)}$ and $\tilde{g}_{i_K}^{(K)}$ are polynomials in $\mathbb{R}[\theta_1,\dots,\theta_d]$ and $\tilde{g}_{i_K}^{(K)}$ is defined by

$$\tilde{g}_{i_K}^{(K)}(\theta) = \sum_{r=K}^{s} \sum_{\substack{(i_1,\dots,i_{K-1},i_{K+1},\dots,i_r) \\ i_j \in \{1,\dots,d\}}} a_{(i_1,\dots,i_r)} \theta_{i_1} \cdots \theta_{i_{K-1}} \theta_{i_{K+1}} \cdots \theta_{i_r}.$$

Now, recall that $m = \eta s$ and observe that we can write $g(X_1)$ as follows.

$$\begin{aligned}
g(X_1) &= \mathbb{E}[h(X_1^m)|X_1] = \frac{1}{m} \sum_{k=1}^{m} \mathbb{E}[\breve{h}^{(k)}(X_1^m)|X_1] \\
&= \frac{1}{m} \sum_{K=1}^{s} \sum_{l=1}^{\eta} \left(R^{(K-1)}(\theta) + \sum_{i \in D(f)} \hat{\theta}_{i,l}(X_1) \tilde{g}_i^{(K)}(\theta) \right) \\
&= \frac{\eta}{m} \sum_{K=1}^{s} R^{(K-1)}(\theta) + \frac{1}{m} \sum_{K=1}^{s} \sum_{l=1}^{\eta} \sum_{i \in D(f)} \hat{\theta}_{i,l}(X_1) \tilde{g}_i^{(K)}(\theta) \\
&= \underbrace{\frac{\eta}{m} \sum_{K=1}^{s} R^{(K-1)}(\theta)}_{=:R(\theta)} + \sum_{i \in D(f)} \underbrace{\sum_{l=1}^{\eta} \hat{\theta}_{i,l}(X_1)}_{=\hat{\theta}_f(X_1)_i} \underbrace{\left(\frac{1}{m} \sum_{K=1}^{s} \tilde{g}_i^{(K)}(\theta) \right)}_{=:\tilde{g}_i(\theta)} \\
&= R(\theta) + \sum_{i \in D(f)} \hat{\theta}_f(X_1)_i \; \tilde{g}_i(\theta) \\
&= R(\theta) + \tilde{g}(\theta)^\top \hat{\theta}_f(X_1),
\end{aligned}$$

where $\tilde{g}(\theta)$ is the vector $(\tilde{g}_i(\theta))_{i \in D(f)}$. We observe that the variance of $g(X_1)$ is given by

$$\mathrm{Var}_\theta[g(X_1)] = \tilde{g}(\theta)^\top \mathrm{Cov}[\hat{\theta}_f(X_1)] \, \tilde{g}(\theta).$$

Note that this is a polynomial, i.e., $\mathrm{Var}_\theta[g(X_1)] \in \mathbb{R}[\theta_1,\dots,\theta_d]$. We will now argue that $\mathrm{Var}_\theta[g(X_1)]$ is not the zero polynomial. It is easy to see by the above derivations that $\mathbb{E}_\theta[h(X_1^m)] = f(\theta) = R(\theta)$ if all components of $\tilde{g}(\theta)$ are identically to the zero polynomial. But this is a contradiction since the degree of the polynomial $R(\theta)$ is at most $s-1$. Hence, there is at least one component in $\tilde{g}(\theta)$ that is not zero. Since $\mathrm{Cov}[\hat{\theta}_f(X_1)]$ is positive definite, we conclude that $\mathrm{Var}_\theta[g(X_1)]$ is also not identical to the zero polynomial. The proof is finished

by recalling that the zero set of a real polynomial that is not the zero polynomial has Lebesgue measure zero; see the lemma of Okamoto (1973). □

Appendix B. Additional Lemmas

The following Lemma concerns Gaussian approximation for independent sums. It is a generalization of Proposition 2.1 in Chernozhukov et al. (2017a) and implicitly proved in the work of Song et al. (2019). However, we state an explicit version here since our proofs rely on this Lemma.

Lemma B.1. *Let $X_1, \dots, X_n$ be i.i.d. centered $\mathbb{R}^p$-valued random vectors. Let $\beta \in (0,1]$ be an absolute constant and assume there exists $\underline{\sigma}^2 > 0$ and $D_n \geq 1$ such that, for all $i = 1, \dots, n$,*

$$\mathbb{E}[X_{ij}^2] \geq \underline{\sigma}^2, \qquad \mathbb{E}[X_{ij}^{2+l}] \leq \mathbb{E}[X_{ij}^2] D_n^l \quad \text{for} \quad j = 1, \dots, p, \quad l = 1, 2,$$
$$\|X_{ij}\|_{\psi_\beta} \leq D_n \text{ for } j = 1, \dots, p.$$

Then there is a constant $C_\beta > 0$ only depending on β such that

$$\sup_{R \in \mathbb{R}^p_{\mathrm{re}}} \left| P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i \in R \right) - P(Y \in R) \right| \leq C_\beta \left(\frac{D_n^2 \log(pn)^{1+6/\beta}}{\underline{\sigma}^2 \, n} \right)^{1/6},$$

where $Y \sim N_p(0, \Sigma)$ with $\Sigma = \mathbb{E}[X_i X_i^\top]$.

Proof. Essentially, this Lemma is obtained from the high dimensional central limit theorem (Chernozhukov et al., 2017a, Proposition 2.1) with generalizing for $\beta \in (0,1]$ and suitable normalization such that the constant $\underline{\sigma}^2$ is explicit in the bound. Generalizing for $\beta \in (0,1]$ is established in Song et al. (2019, Lemma A.8) and the proof with suitable normalization is implicit in the proof of Corollary 2.2 in Song et al. (2019). □

The next Lemma verifies Gaussian approximation of complete U-statistics. It is proved in Song et al. (2019). Since the Lemma is essential for the proof of our main theorem, we state it here.

Lemma B.2 (Song et al., 2019, Corollary 2.2)**.** *Assume the sub-Weibull condition (C2) is satisfied. Moreover, assume (C4) holds with $p_1 = p$ and (C6) holds. Then there is a constant $C_\beta > 0$ only depending on β such that*

$$\sup_{R \in \mathbb{R}^p_{\mathrm{re}}} \left| P\left(\sqrt{n}(U_n - \mu) \in R \right) - P(mY_g \in R) \right| \leq C_\beta \left(\frac{m^2 D_n^2 \log(pn)^{1+6/\beta}}{(\underline{\sigma}_g^2 \wedge 1) \, n} \right)^{1/6},$$

where $Y_g \sim N_p(0, \Gamma_g)$.

Appendix C. Sub-Weibull Random Variables

In this section we collect important properties of sub-Weibull random variables. For $0 < \beta < 1$ we only have that $\|\cdot\|_{\psi_\beta}$ is a quasinorm, i.e. the triangle inequality does not hold. Nevertheless, we have the following result which is a substitute.

Lemma C.1. *For any $0 < \beta < 1$ and any random variables $X_1, \dots, X_n$, we have*

$$\left\| \sum_{i=1}^n X_i \right\|_{\psi_\beta} \leq n^{\frac{1}{\beta}} \sum_{i=1}^n \|X_i\|_{\psi_\beta}.$$

Proof. For $n = 2$ the claim is equal to Lemma A.3 in Götze et al. (2021). We generalize the statement for arbitrary $n \in \mathbb{N}$ using similar arguments. Let $L_i = \|X_i\|_{\psi_\beta}$ and define $t := n^{\frac{1}{\beta}} \sum_{i=1}^n L_i$. We have

$$\begin{aligned}
& \mathbb{E}\left[\exp\left(\frac{|\sum_{i=1}^n X_i|^\beta}{t^\beta}\right)\right] \\
\leq\ & \mathbb{E}\left[\exp\left(\frac{\left(\sum_{i=1}^n |X_i|\right)^\beta}{t^\beta}\right)\right] \overset{(1)}{\leq} \mathbb{E}\left[\exp\left(\frac{\sum_{i=1}^n |X_i|^\beta}{n\left(\sum_{i=1}^n L_i\right)^\beta}\right)\right] \\
=\ & \mathbb{E}\left[\prod_{i=1}^n \exp\left(\frac{|X_i|^\beta}{n\left(\sum_{i=1}^n L_i\right)^\beta}\right)\right] \\
\leq\ & \mathbb{E}\left[\prod_{i=1}^n \exp\left(\frac{|X_i|^\beta}{nL_i^\beta}\right)\right] = \mathbb{E}\left[\prod_{i=1}^n \exp\left(\frac{|X_i|^\beta}{L_i^\beta}\right)^{\frac{1}{n}}\right] \\
\overset{(2)}{\leq}\ & \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \exp\left(\frac{|X_i|^\beta}{L_i^\beta}\right)\right] = \frac{1}{n}\sum_{i=1}^n \underbrace{\mathbb{E}\left[\exp\left(\frac{|X_i|^\beta}{L_i^\beta}\right)\right]}_{\leq 2 \text{ for all } i} \leq 2.
\end{aligned}$$

Here, (1) follows from the inequality $\left(\sum_{i=1}^n x_i\right)^\beta \leq \sum_{i=1}^n x_i^\beta$ valid for all $x_1, \ldots, x_n \geq 0$ and $\beta \in [0, 1]$, and (2) from $\prod_{i=1}^n a_i^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^n a_i$ for all $a_1, \ldots, a_n \geq 0$. The latter inequality is known as the inequality of geometric and arithmetic means. By the calculation and the definition of the sub-Weibull norm we have

$$\left\|\sum_{i=1}^n X_i\right\|_{\psi_\beta} \leq t = n^{\frac{1}{\beta}} \sum_{i=1}^n \|X_i\|_{\psi_\beta}.$$

□

For products of sub-Weibull random variables we recall the following useful result from Kuchibhotla and Chakrabortty (2022).

Lemma C.2. *Let $X_1, \ldots, X_n$ be random variables satisfying $\|X_i\|_{\psi_{\beta_i}} < \infty$ for some $\beta_i > 0$, $i = 1, \ldots, n$. Then*

$$\left\|\prod_{i=1}^n X_i\right\|_{\psi_\beta} \leq \prod_{i=1}^n \|X_i\|_{\psi_{\beta_i}} \quad \text{where } \frac{1}{\beta} = \sum_{i=1}^n \frac{1}{\beta_i}.$$

Proof. See Kuchibhotla and Chakrabortty (2022, Proposition D.2). □

Combining Lemma C.1 and Lemma C.2 we get a bound on polynomial functions in random variables. This is especially useful for showing that the estimators $h_j(X_1^m)$, $j = 1, \ldots, p$ are sub-Weibull.

Lemma C.3. *Let $\beta > 0$ and let $X_1, \ldots, X_n$ be (possibly dependent) random variables satisfying $\|X_i\|_{\psi_\beta} \leq C$ for some constant $C > 0$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of total degree s with t terms. Then we have*

$$\|f(X_1, \ldots, X_n)\|_{\psi_{\frac{\beta}{s}}} \leq A \begin{cases} t\, C^s & \frac{\beta}{s} \geq 1 \\ t^{\frac{s}{\beta}+1}\, C^s & \frac{\beta}{s} < 1, \end{cases}$$

where A is the maximum over all absolute values of the coefficients in the polynomial f.

Proof. For random variables $X_1, \dots, X_s$ we have by Lemma C.2

$$\left\| \prod_{i=1}^{s} X_i \right\|_{\psi_{\frac{\beta}{s}}} \leq \|X_i\|_{\psi_\beta}^s.$$

Hence, for $f(X_1, \dots, X_n)$ being a monomial of degree s, it follows the inequality

$$\|f(X_1, \dots, X_n)\|_{\psi_{\frac{\beta}{s}}} \leq A\, C^s.$$

To prove the general case where $f(X_1, \dots, X_n)$ is a polynomial with t terms we make a case distinction. For $\frac{\beta}{s} \geq 1$ we have that $\|\cdot\|_{\psi_{\frac{\beta}{s}}}$ is a norm and the triangle inequality holds. For $\frac{\beta}{s} < 1$ the claim follows immediately from Lemma C.1. □

The next well-known Lemma states that sub-Weibull random variables satisfy a stronger moment condition than the existence of all finite q-th moments.

Lemma C.4. *For any $\beta > 0$ and any random variable X we have*

$$d_\beta \sup_{q \geq 1} \frac{\|X\|_q}{q^{\frac{1}{\beta}}} \leq \|X\|_{\psi_\beta} \leq D_\beta \sup_{q \geq 1} \frac{\|X\|_q}{q^{\frac{1}{\beta}}},$$

where $d_\beta = \begin{cases} \frac{1}{2} & \beta \geq 1 \\ \frac{(\beta e)^{1/\beta}}{2} & \beta < 1 \end{cases}$ *and* $D_\beta = \begin{cases} 2e & \beta \geq 1 \\ 2e^{1/\beta} & \beta < 1. \end{cases}$

Proof. See for example Götze et al. (2021, Lemma A.2). □

According to Lemma C.4, the parameter β measures how fast $\|X\|_q$ increases with q. For small β, the norms $\|X\|_q$ are allowed to increase fast and viceversa. In particular, Lemma C.4 implies that for all $q \geq 1$ there is a constant $C_\beta > 0$ only depending on β such that $\|X\|_q \leq C_\beta \|X\|_{\psi_\beta}$. In the special case of polynomials in Gaussian variables, the sub-Weibull norm is always bounded by the standard deviation. This follows from the hypercontractivity property of polynomials in Gaussian random variables.

Lemma C.5. *Let $\mathbf{X} = (X_1, \dots, X_r) \sim N_r(0, \Sigma)$ be an r-variate centered Gaussian random vector, and let $f : \mathbb{R}^r \to \mathbb{R}$ be a polynomial of total degree s. If $0 < \beta \leq 2/s$, then*

$$\|f(\mathbf{X})]\|_{\psi_\beta} \leq C_{\beta,s}\, \|f(\mathbf{X})\|_2,$$

where $C_{\beta,s}$ is a constant depending only on β and s.

Proof. By Lemma C.4, we have the inequality

$$\|f(\mathbf{X})\|_{\psi_\beta} \leq C_\beta \sup_{q \geq 1} \frac{\|f(\mathbf{X})\|_q}{q^{1/\beta}}. \tag{C.1}$$

Now, for all $q \geq 2$, by the hypercontractivity property of polynomials in Gaussian random variables (Theorem 3.2.10 in de la Peña and Giné, 1999 and Lemma 2.2 in Leung and Sturma, 2024),

$$\|f(\mathbf{X})\|_q \leq C_s\, q^{s/2}\, \|f(\mathbf{X})\|_2. \tag{C.2}$$

Since $\|f(\mathbf{X})\|_1 \leq \|f(\mathbf{X})\|_2$, we obtain by combining (C.1) and (C.2) that

$$\|f(\mathbf{X})\|_{\psi_\beta} \leq C_{\beta,s}\, \|f(\mathbf{X})\|_2 \sup_{q \geq 2} q^{\frac{s}{2} - \frac{1}{\beta}}.$$

To conclude the proof, we note that $s/2 \leq 1/\beta$ holds due to our assumptions, which implies that the supremum is achieved for $q = 2$. □

Next, we cite a result from Chen and Kato (2019) that gives a sub-Weibull bound on the maximum of multiple random variables in relation to the bound on the individual variables.

Lemma C.6. *Let $\beta > 0$ and let $X_1, \dots, X_n$ be (possibly dependent) random variables such that $\|X_i\|_{\psi_\beta} < \infty$ for all $i = 1, \dots, n$. Then for $n \geq 2$*

$$\left\| \max_{1 \leq i \leq n} |X_i| \right\|_{\psi_\beta} \leq C_\beta \log(n)^{\frac{1}{\beta}} \max_{1 \leq i \leq n} \|X_i\|_{\psi_\beta},$$

where C_β is a constant depending only on β.

Proof. See Chen and Kato (2019, Lemma C.1). □

The important conclusion of Lemma C.6 is that the bound on the individual variables $\|X_i\|_{\psi_\beta}$ yields a bound on the rate of growth on $\| \max_{1 \leq i \leq n} |X_i| \|_{\psi_\beta}$. It is determined by the slowly growing function $\log(n)^{1/\beta}$, i.e. at most logarithmic in n. The last Lemma is a result considering centered random variables.

Lemma C.7. *For any $\beta > 0$ and any random variable X with $\mathbb{E}[X] < \infty$ we have*

$$\|X - \mathbb{E}[X]\|_{\psi_\beta} \leq C_\beta \|X\|_{\psi_\beta},$$

where C_β is a constant depending only on β.

Proof. The proof is similar to the proof of Vershynin (2018, Lemma 2.6.8) which treats the special case $\beta = 2$. Let C_β be a constant only depending on β but the value can change from place to place. For $\beta \geq 1$ recall that $\| \cdot \|_{\psi_\beta}$ is a norm. Thus we can use the triangle inequality and get

$$\|X - \mathbb{E}[X]\|_{\psi_\beta} \leq \|X\|_{\psi_\beta} + \|\mathbb{E}[X]\|_{\psi_\beta}.$$

If $0 < \beta < 1$ we use Lemma C.1 with $n = 2$ and get

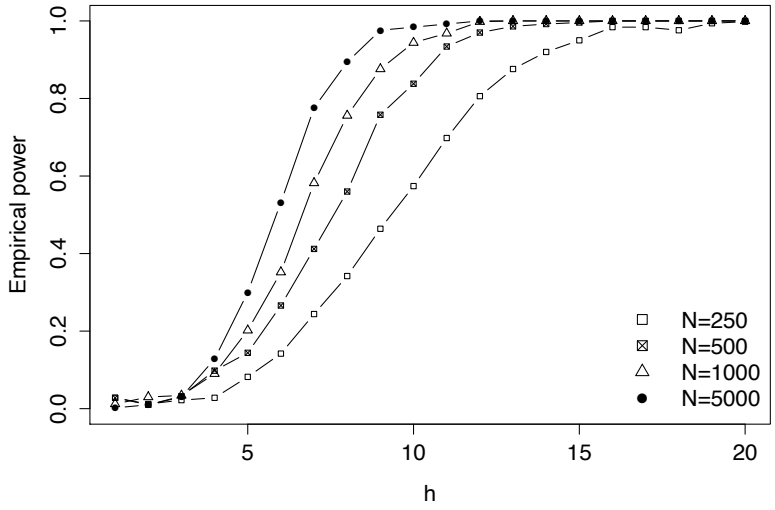
$$\|X - \mathbb{E}[X]\|_{\psi_\beta} \leq 2^{\frac{1}{\beta}} \left(\|X\|_{\psi_\beta} + \|\mathbb{E}[X]\|_{\psi_\beta} \right).$$

We only have to bound the second term of both inequalities. Note that, for any constant a, we trivially have $\|a\|_{\psi_\beta} \leq C_\beta |a|$ by the Definition of $\| \cdot \|_{\psi_\beta}$. Using this for $a = \mathbb{E}[X]$, we get

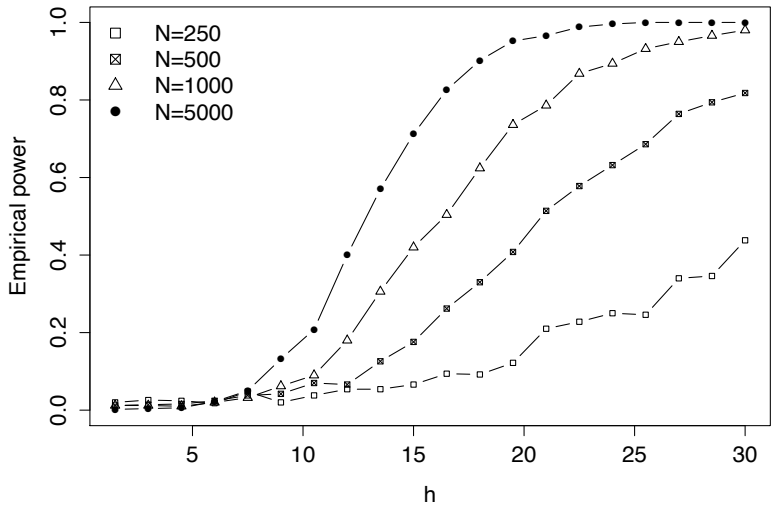
$$\begin{aligned}
\|\mathbb{E}[X]\|_{\psi_\beta} &\leq C_\beta |\mathbb{E}[X]| \\
&\leq C_\beta \mathbb{E}[|X|] \quad \text{(by Jensen's inequality)} \\
&= C_\beta \|X\|_1 \\
&\leq C_\beta \|X\|_{\psi_\beta} \quad \text{(by Lemma C.4 with } q = 1).
\end{aligned}$$

□

Appendix D. Additional Simulation Results for Gaussian Latent Tree Models



Setup (b)



Setup (c)

Figure 7. Empirical power for different local alternatives based on 500 experiments. The computational budget parameters N is varied as indicated. Local alternatives are generated as described in Section 5 for setups (b) and (c) with $(l, n) = (15, 500)$ and level $\alpha = 0.05$.

Appendix E. Testing Minors in Two-Factor Analysis Models

In this section, we consider testing model invariants of two-factor analysis models as another application of our testing methodology. By Drton et al. (2007) we have the following parametric representation of two-factor analysis models.

Proposition E.1. *The two-factor analysis model is the family of Gaussian distributions $N_l(\mu, \Sigma)$ on $\mathbb{R}^l$ whose mean vector μ is an arbitrary vector in $\mathbb{R}^l$ and whose covariance matrix Σ is in the set*

$$F_{l,2} = \{\Lambda + \Psi \in \mathbb{R}^{l\times l} : \Lambda > 0 \text{ diagonal, } \Psi \geq 0 \text{ symmetric, } \mathrm{rank}(\Psi) \leq 2\}.$$

Here the notation $B > 0$ means that B is a positive definite matrix, and $B \geq 0$ means that B is positive semidefinite. Given a covariance matrix $\Sigma \in F_{l,2}$ in the two-factor analysis model, all off-diagonal 3×3 minors are vanishing (Drton et al., 2007). We are interested in testing all of these model invariants simultaneously based on i.i.d. samples $X_1, \ldots, X_n \sim N(0, \Sigma)$, where the covariance matrix Σ in unknown. Up to sign, there are $p = 10\binom{l}{6}$ off-diagonal (3×3)-minors, that is, the number of minors grows very fast with the dimension l. Since minors of $\Sigma = (\sigma_{uv})$ are polynomials in the entries σ_{uv}, we can apply the kernel proposed in Section 4. We consider two experimental setups:

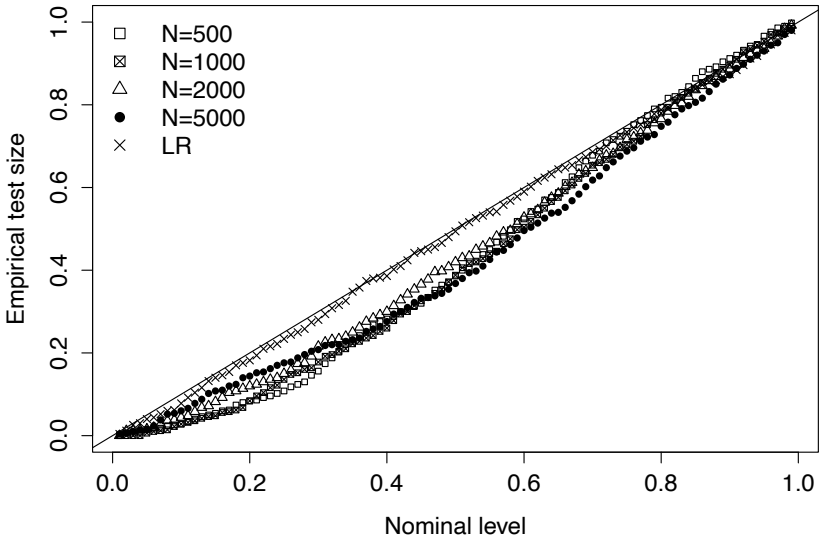
(*Reg*) $\Psi = \Gamma\Gamma^\top$ with $\Gamma \in \mathbb{R}^{l\times 2}$ and all entries of Γ are independently generated based on a standard normal distribution. All diagonal entries Λ_{vv} are taken to be 1.

(*Irreg*) $\Psi = \beta_1\beta_1^\top + \beta_2\beta_2^\top$ with $\beta_1, \beta_2 \in \mathbb{R}^l$. All entries of β_1 are equal to 1. The first two entries of β_2, i.e., β_{21} and β_{22}, are taken to be 10 while all other entries of β_2 are independently generated based on a normal distribution with mean 0 and variance 0.2. All diagonal entries Λ_{vv} are taken to be $1/3$.

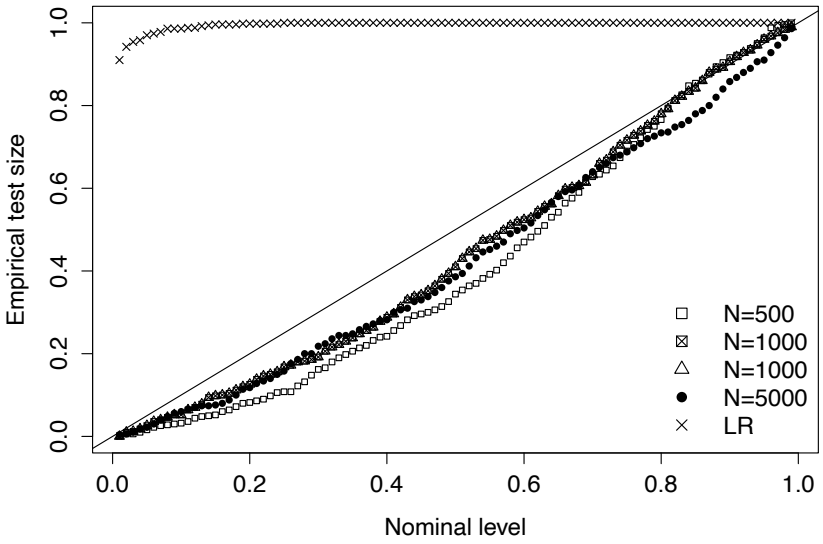
Setup *Reg* is designed to be a regular setup while setup *Irreg* is irregular. Moreover, the parameters in the irregular setup are close to an algebraic singularity (Drton et al., 2007, Example 33) such that we expect that the likelihood ratio test fails to control test size.

In the implementation we use $A = 1000$ sets of Gaussian multipliers to compute the critical value $c_W(1 - \alpha)$. We apply a more general version of the divide-and-conquer bootstrap discussed in Section 2.2, where we use a block-diagonal sampling scheme with a block size of $L = 25$; see Chen and Kato (2019). However, the theoretical bound for the asymptotic approximation remains the same. As before, we compare our methodology with the likelihood ratio test implemented by the `factanal` function in `R` (R Core Team, 2020).

In Figure 8 we compare empirical test sizes for different, fixed nominal levels $\alpha \in (0, 1)$. Once again, we see the advantage of our test methodology in the irregular setup. While the likelihood ratio test fails to control test size, our test has lower empirical size than nominal level for all computational budgets N, i.e., our test controls type I error, albeit conservatively. Moreover, we compare the empirical power in Figure 9, where we construct the local alternatives equivalently to the alternatives considered in Section 5. That is, for $\Sigma \in F_{l,2}$, the alternatives are of the form $\widetilde{\Sigma} = \Sigma + \gamma\gamma^\top h/\sqrt{n}$, where $\gamma = (0, \ldots, 0, 1, 1) \in \mathbb{R}^l$ and $h > 0$ varies. As before, we observe that the empirical power is better for larger computational budgets N.

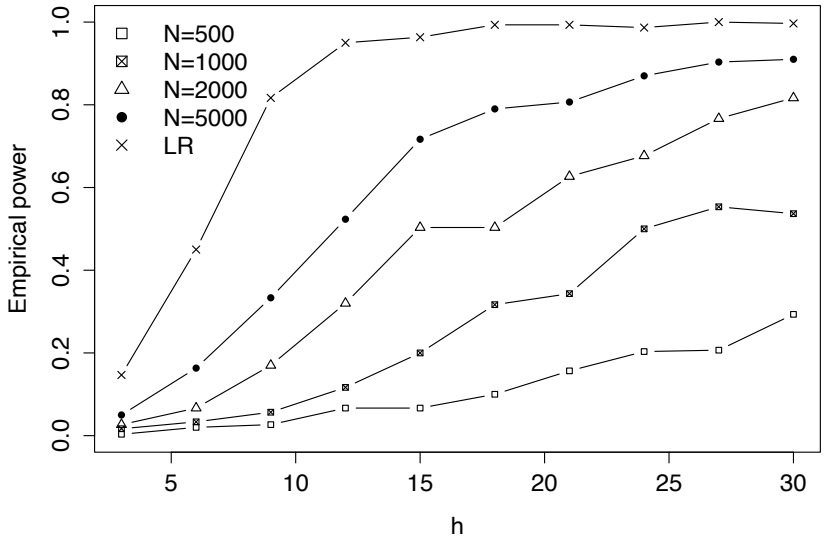


Regular setup

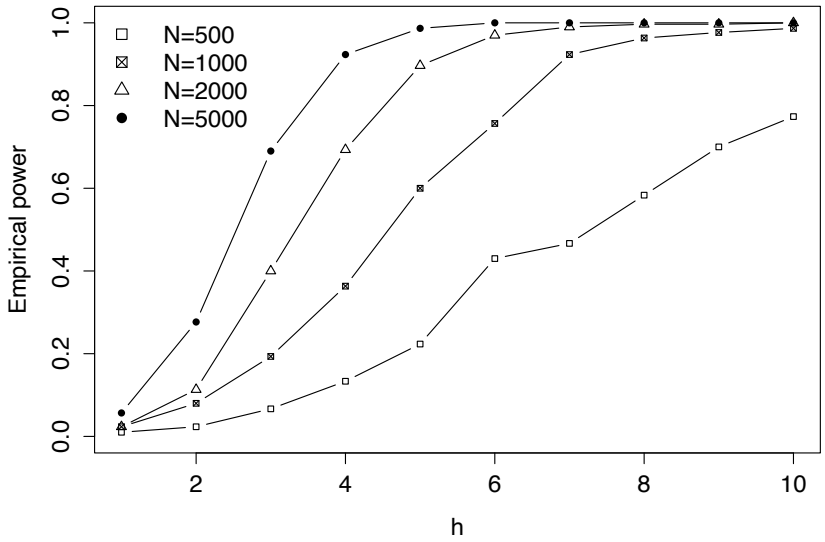


Irregular setup

FIGURE 8. Empirical sizes vs. nominal levels for testing (3×3)-minors based on 500 experiments. The computational budget parameter N is varied as indicated and empirical sizes of the LR test are also shown. Data is generated from regular and singular setups with $(l, n) = (10, 500)$.



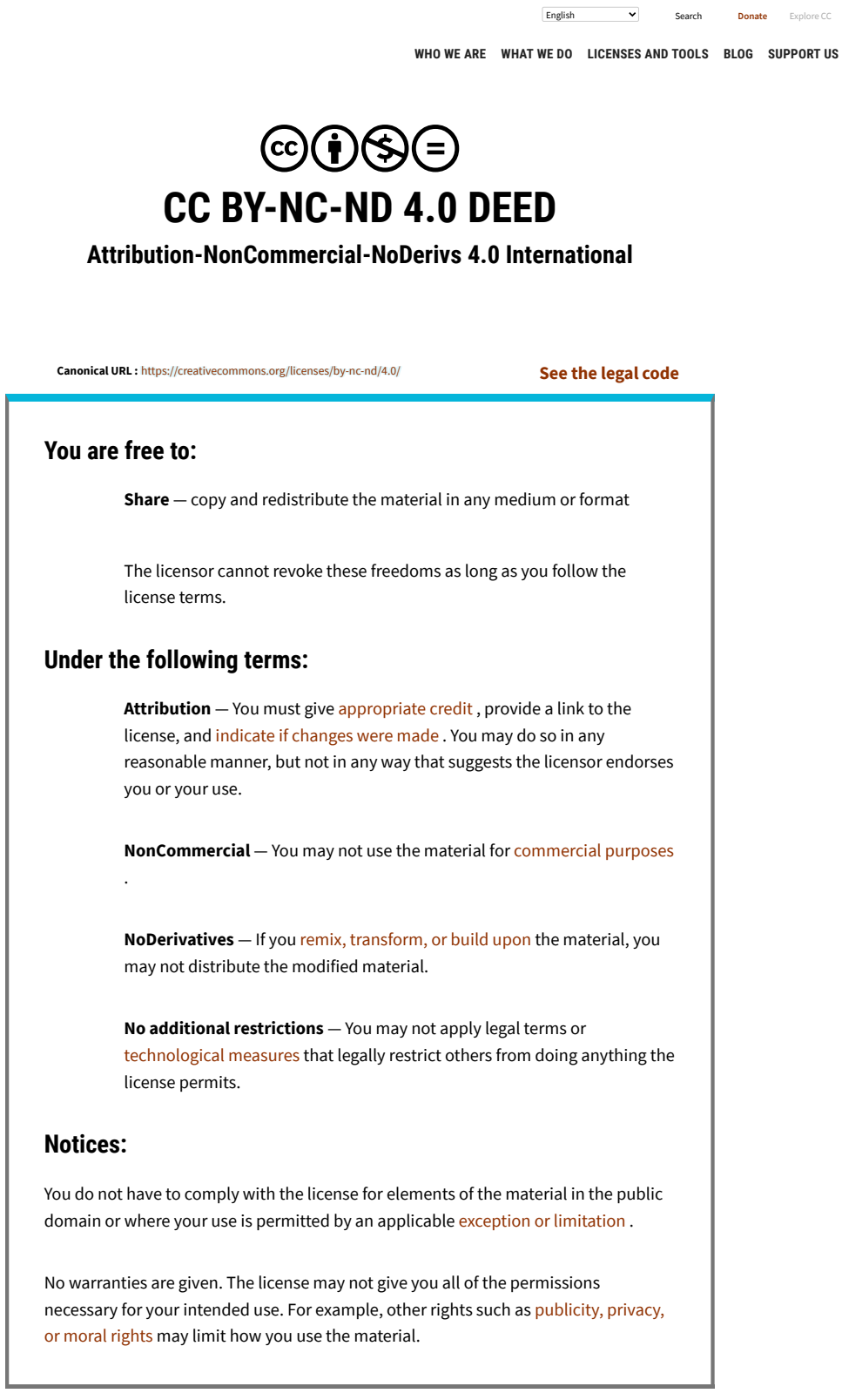
Regular setup



Irregular setup

FIGURE 9. Empirical power in testing (3×3)-minors for different local alternatives based on 300 experiments. The computational budget parameters N is varied as indicated. Local alternatives are generated as described in the text for regular and irregular setups with $(l, n) = (10, 500)$ and level $\alpha = 0.05$.

Permission to include the article



English Search Donate Explore CC

WHO WE ARE WHAT WE DO LICENSES AND TOOLS BLOG SUPPORT US

CC BY-NC-ND 4.0 DEED

Attribution-NonCommercial-NoDerivs 4.0 International

Canonical URL : https://creativecommons.org/licenses/by-nc-nd/4.0/ **See the legal code**

You are free to:

Share — copy and redistribute the material in any medium or format

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit , provide a link to the license, and indicate if changes were made . You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes .

NoDerivatives — If you remix, transform, or build upon the material, you may not distribute the modified material.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation .

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Notice

This deed highlights only some of the key features and terms of the actual license. It is not a license and has no legal value. You should carefully review all of the terms

and conditions of the actual license before using the licensed material.

Creative Commons is not a law firm and does not provide legal services. Distributing, displaying, or linking to this deed or the license that it summarizes does not create a lawyer-client or any other relationship.

Creative Commons is the nonprofit behind the open licenses and other legal tools that allow creators to share their work. Our legal tools are free to use.

- Learn more about our work
- **Learn more about CC Licensing**
- Support our work
- Use the license for your own material.
- Licenses List
- Public Domain List

Footnotes

appropriate credit — If supplied, you must provide the name of the creator and attribution parties, a copyright notice, a license notice, a disclaimer notice, and a link to the material. CC licenses prior to Version 4.0 also require you to provide the title of the material if supplied, and may have other slight differences.
- More info

indicate if changes were made — In 4.0, you must indicate if you modified the material and retain an indication of previous modifications. In 3.0 and earlier license versions, the indication of changes is only required if you create a derivative.
- Marking guide
- More info

commercial purposes — A commercial use is one primarily intended for commercial advantage or monetary compensation.
- More info

remix, transform, or build upon — Merely changing the format never creates a derivative.
- More info

technological measures — The license prohibits application of effective technological measures, defined with reference to Article 11 of the WIPO Copyright Treaty.
- More info

exception or limitation — The rights of users under exceptions and limitations, such as fair use and fair dealing, are not affected by the CC licenses.
- More info

publicity, privacy, or moral rights — You may need to get additional permissions before using the material as you intend.
- More info

Publication rights Navigation

Publication rights

What is our policy?

For the majority of journals published by Oxford University Press (OUP), we have a policy of acquiring a sole and exclusive licence for all published content, rather than asking authors to transfer ownership of their copyright, which has been common practice in the past. We believe this policy more carefully balances the interests of our authors with our need to maintain the viability and reputation of the journals through which our authors are accorded status, recognition, and widespread distribution. In developing this policy we have been guided by the following principles:

- As a university press and not-for-profit academic publisher, we rely heavily on the good relationships we have with our authors. Having a licensing policy which enables an author to be identified as the owner of the copyright in an article is one of the key ways of demonstrating how highly we value these relationships.
- An exclusive licence enables the centralised and efficient management of permissions and licencing, ensuring the widest dissemination of the content through intermediaries;
- Exclusive rights also enable OUP to take measures on behalf of our authors against infringement, inappropriate use of an article, libel or plagiarism;
- At the same time, by maintaining exclusive rights, in all media for all published content, we can monitor and uphold the integrity of an article once refereed and accepted for publication to be maintained;

Please note - Although the majority of the journals published by OUP operate exclusive copyright agreements, some of our society journal may require an assignment of copyright. For further information, please refer to the Authors Guidelines sections of the Journal's homepage.

Copyright agreements

OUP cannot publish your article until a completed copyright agreement has been received. You should receive a link to our online licencing system or a hardcopy form as soon as your article is accepted for publication.

Government employees

- If you are or were a UK Crown servant and the article has been written in that capacity, we have an arrangement with HMSO to enable us to publish it while acknowledging that it is Crown Copyright. Please inform the Editorial office or OUP at the time of acceptance or as soon as possible that the article is Crown Copyright, so that we can ensure the appropriate acknowledgement and copyright line are used, as required by our arrangement with HMSO.
- If you are a US Government employee and the article has been written in that capacity, we acknowledge that the Licence to Publish applies only to the extent allowable by US law.

Re-use of third party content as part of your Oxford journal article

- As part of your article, you may wish to reuse material sourced from third parties such as other publishers, authors, museums, art galleries etc. To assist with this process, we have

a Permission Request form and accompanying Guidelines that specifies the rights required in order for third party material to be published as part of your Article. A copy of the form can be provided by our Production Team.

- Responsibility for clearing these third party permissions must be borne by the Author, and this process completed as soon as possible - preferably before acceptance of the manuscript, but if not possible, before the Article reaches the Production stage of the process.

Rights retained by ALL Oxford journal authors

- The right, after publication by Oxford journals, to use all or part of the Article and abstract, for their own personal use, including their own classroom teaching purposes;
- The right, after publication by Oxford journals, to use all or part of the Article and abstract, in the preparation of derivative works, extension of the article into book-length or in other works, provided that a full acknowledgement is made to the original publication in the journal;
- The right to include the article in full or in part in a thesis or dissertation, provided that this is not published commercially;

For the uses specified here, please note that there is no need for you to apply for written permission from OUP in advance. Please go ahead with the use ensuring that a full acknowledgment is made to the original source of the material including the journal name, volume, issue, page numbers, year of publication, title of article and to OUP and/or the learned society.

The only exception to this is for the re-use of material for commercial purposes, e.g. republication or distribution of an article by a for-profit publisher or medical communications company etc. Permission for this kind of re-use is required and can be obtained by using Rightslink.

With Copyright Clearance Center's Rightslink ® service it's faster and easier than ever before to secure permission from OUP titles to be republished in a coursepack, book, CD-ROM/DVD, brochure or pamphlet, journal or magazine, newsletter, newspaper, make a photocopy, or translate.

- Simply visit the required journal on Oxford Academic and locate your desired content.
- Click on (Order Permissions) within the table of contents and/ or at the bottom article's abstract to open the following page:
- Select the way you would like to reuse the content
- Create an account or login to your existing account
- Accept the terms and conditions and permission is granted

For questions about using the Rightslink service, please contact our journals permissions team, or email Rightslink customer care: customercare@copyright.com.

Author self-archiving policies

Journals from Oxford also offer a range of author self-archiving polices, permitting authors to share and distribute various versions of their articles. For full details of our self-archiving policies please see our Author self-archiving page.

Permissions

- All requests to reuse the article, in whole or in part, in another publication will be handled by Oxford journals. Unless otherwise stated, any permission fees will be retained by the Journal concerned. Where possible, any requests to reproduce substantial parts of the article (including in other OUP publications) will be subject to your approval (which is deemed to be given if we have not heard from you within 4 weeks of the permission being granted).
- If copyright of the article is held by someone other than the Author, e.g. the Author's employer, Oxford journals requires non-exclusive permission to administer any requests from third parties.
- The Journal is registered with the Copyright Licensing Agency (London) and the Copyright Clearance Center (Danvers, Massachusetts), and other Reproduction Rights Organizations. These are non-profit organizations which offer centralised licensing arrangements for photocopying on behalf of publishers such as OUP.
- Please forward requests to re-use all or part or your article, or to use figures contained within it, to the journals permissions team.

Related information

Journals A-Z

Rightslink

Author self-archiving policy

A.2 Unpaired Multi-Domain Causal Representation Learning

Summary

In this article, we introduce the problem of causal representation learning from multi-domain observations. We assume that multiple domains provide complementary information about a set of shared latent variables that are the causal quantities of primary interest, recall our discussion in Chapter 4 of this thesis. Crucially, observations in different domains are assumed to be unpaired, that is, we only observe the marginal distribution in each domain but not their joint distribution. For this problem, we lay out a setting in which we can provably identify the causal relations among the shared latent variables. Our main results are sufficient conditions for identifiability of the joint distribution and the shared causal graph in a linear setup. Identifiability holds if we can uniquely recover the joint distribution and the shared causal graph from the marginal distributions in each domain.

The article is structured as follows. In the introduction, we motivate and introduce the setup, and we extensively discuss related work. In Section 2, we provide precise definitions for the linear model by taking a graphical perspective. Then, we proceed in a two-step approach to identify the desired causal structure: In Section 3, we first give a sufficient condition for identifiability of the joint distribution. We extend identifiability results of linear ICA to the unpaired multi-domain setup by applying it in each domain separately and matching the recovered error distributions. In particular, this allows to identify the shared latent nodes and an "overall mixing matrix", which together imply identifiability of the joint distribution of the domains. Our conditions are also mostly necessary as we discuss in Appendix C. In a second step, we prove a sufficient condition for identifiability of the shared latent causal graph in Section 4. Our condition requires that the joint distribution was already certified to be identifiable and we make use of rank deficiencies in the overall mixing matrix. The rank deficiencies are translated into a graphical condition that involves a certain type of "pure children" of the latent nodes.

Although we emphasize that our focus in this article is on identifiability, our proofs also suggest methods to learn the joint distribution as well as the shared causal graph from finite samples. We provide algorithms for the noisy setting and, moreover, we analyze how the number of domains reduce uncertainty with respect to the learned representation. Due to space constraints, the detailed discussion of the finite sample setting is deferred to the Appendix. In the main text of the article, we deduct a small simulation study as a proof of concept for the finite sample setting in Section 5.

Individual contributions

I am the main author of this article. I formulated the specific research problem, developed all proofs, wrote software implementations, conducted simulation studies, and drafted the manuscript. The idea of investigating identifiability for causal representation learning in multiple domains was proposed by Caroline Uhler when I visited her research group. All co-authors made helpful suggestions regarding both content and presentation of the article during regular discussions.

Unpaired Multi-Domain Causal Representation Learning

Nils Sturma
Technical University of Munich
Munich Center for Machine Learning
nils.sturma@tum.de

Chandler Squires
LIDS, Massachusetts Institute of Technology
Broad Institute of MIT and Harvard
csquires@mit.edu

Mathias Drton
Technical University of Munich
Munich Center for Machine Learning
mathias.drton@tum.de

Caroline Uhler
LIDS, Massachusetts Institute of Technology
Broad Institute of MIT and Harvard
cuhler@mit.edu

Abstract

The goal of causal representation learning is to find a representation of data that consists of causally related latent variables. We consider a setup where one has access to data from multiple domains that potentially share a causal representation. Crucially, observations in different domains are assumed to be unpaired, that is, we only observe the marginal distribution in each domain but not their joint distribution. In this paper, we give sufficient conditions for identifiability of the joint distribution and the shared causal graph in a linear setup. Identifiability holds if we can uniquely recover the joint distribution and the shared causal representation from the marginal distributions in each domain. We transform our results into a practical method to recover the shared latent causal graph.

1 Introduction

An important challenge in machine learning is the integration and translation of data across multiple domains (Zhu et al., 2017; Zhuang et al., 2021). Researchers often have access to large amounts of unpaired data from several domains, e.g., images and text. It is then desirable to learn a probabilistic coupling between the observed marginal distributions that captures the relationship between the domains. One approach to tackle this problem is to assume that there is a latent representation that is invariant across the different domains (Bengio et al., 2013; Ericsson et al., 2022). Finding a probabilistic coupling then boils down to learning such a latent representation, that is, learning high-level, latent variables that explain the variation of the data within each domain as well as similarities across domains.

In traditional representation learning, the latent variables are assumed to be statistically independent, see for example the literature on independent component analysis (Hyvärinen and Oja, 2000; Comon and Jutten, 2010; Khemakhem et al., 2020). However, the assumption of independence can be too stringent and a poor match to reality. For example, the presence of clouds and the presence of wet roads in an image may be dependent, since clouds may cause rain which may in turn cause wet roads. Thus, it is natural to seek a *causal representation*, that is, a set of high-level *causal* variables and relations among them (Schölkopf et al., 2021; Yang et al., 2021b). Figure 1 illustrates the setup of multi-domain causal representation learning, where multiple domains provide different views on a shared causal representation.

Our motivation to study multi-domain causal representations comes, in particular, from single-cell data in biology. Given a population of cells, different technologies such as imaging and sequencing

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

provide different views on the population. Crucially, since these technologies destroy the cells, the observations are uncoupled, i.e., a specific cell may either be used for imaging or sequencing but not both. The aim is to integrate the different views on the population to study the underlying causal mechanisms determining the observed features in various domains (Butler et al., 2018; Stuart et al., 2019; Liu et al., 2019; Yang et al., 2021a; Lopez et al., 2022; Gossi et al., 2023; Cao et al., 2022). Unpaired multi-domain data also appears in many applications other than single-cell biology. For example, images of similar objects are captured in different environments (Beery et al., 2018), large biomedical and neuroimaging data sets are collected in different domains (Miller et al., 2016; Essen et al., 2013; Shafto et al., 2014; Wang et al., 2003), or stocks are traded in different markets.

In this paper, we study identifiability of the shared causal representation, that is, its uniqueness in the infinite data limit. Taking on the same perspective as, for example, in Schölkopf et al. (2021) and Squires et al. (2023), we assume that observed data is generated in two steps. First, the latent variables $Z = (Z_i)_{i \in \mathcal{H}}$ are sampled from a distribution P_Z, where P_Z is determined by an unknown structural causal model among the latent variables. Then, in each domain $e \in \{1, \dots, m\}$, the observed vector $X^e \in \mathbb{R}^{d_e}$ is the image of a subset of the latent variables under a domain-specific, injective mixing function g_e. That is,

$$X^e = g_e(Z_{S_e}),$$

where $S_e \subseteq \mathcal{H}$ is a subset of indices. A priori, it is unknown whether a latent variable Z_i with $i \in S_e$ is shared across domains or domain-specific. Even the number of latent variables which are shared across domains is unknown. Moreover, we only observe the marginal distribution of each random vector X^e, but none of the joint distributions over pairs X^e, X^f for $e \neq f$. Said differently, observations across domains are unpaired. Assuming that the structural causal model among the latent variables as well as the mixing functions are linear, our main contributions are:

1. We lay out conditions under which we can identify the joint distribution of $X^1, \dots, X^m$.
2. We give additional conditions under which we are able to identify the causal structure among the shared latent variables.

In particular, identifiability of the joint distribution across domains enables data translation. That is, given observation x in domain e, translation to domain f can be achieved by computing $\mathbb{E}[X_f | X_e = x]$. Furthermore, identifying the causal structure among the shared latent variables lets us study the effect of interventions on the different domains.

The main challenge in proving rigorous identifiability results for multi-domain data is that we cannot apply existing results for single-domain data in each domain separately. Even if the causal structure of the latent variables in a single domain is identifiable, it remains unclear how to combine multiple causal structures, i.e., in which way latent variables are shared. We circumvent this problem via a two-step approach: First, we extend the identifiability of linear independent component analysis (Comon, 1994; Hyvärinen and Oja, 2000; Eriksson and Koivunen, 2004; Mesters and Zwiernik, 2022) to the multi-domain setup, which allows us to identify the joint distribution and distinguish between shared and domain-specific latent variables. Moreover, we identify an "overall mixing matrix" and, in a second step, exploit sparsity constraints in this matrix to identify the causal structure among the shared latent variables. This leverages recent results on causal discovery under measurement error in single domains that also exploit sparsity (Xie et al., 2020; Chen et al., 2022; Xie et al., 2022; Huang et al., 2022). Although we emphasize that our focus in this paper is on identifiability, our proofs also suggest methods to learn the joint distribution as well as the shared causal graph from finite samples. We provide algorithms for the noisy setting and, moreover, we analyze how the number of domains reduce uncertainty with respect to the learned representation.

The paper is organized as follows. In the next paragraphs we discuss further related work. Section 2 provides a precise definition of the considered setup. In Section 3 we consider identifiability of the joint distribution. Using these results, we study identifiability of the causal graph in Section 4. We conclude with a small simulation study as a proof of concept for the finite sample setting in Section 5. Due to space constraints, the detailed discussion of the finite sample setting is deferred to the Appendix. Moreover, the Appendix contains all proofs, discussions on the necessity of our assumptions, and additional examples and simulation results.

Multi-domain Integration. Motivated by technological developments for measuring different modalities at single-cell resolution, several methods have been proposed recently for domain

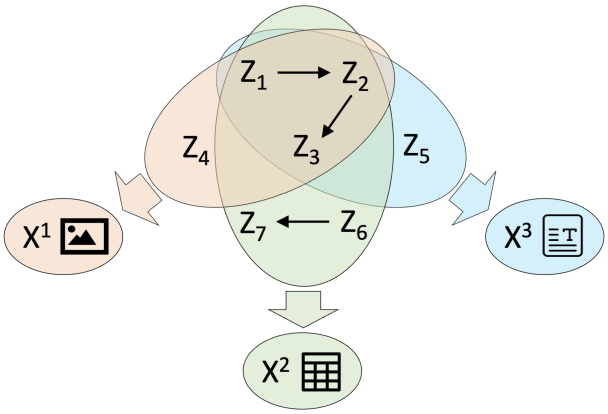


Figure 1: **Setup**. A latent causal representation where multiple domains X^e provide different "views" on subsets of the latent variables Z_i. The domains may correspond to different data modalities such as images, text or numerical data. Crucially, the observations across domains are unpaired, i.e., they arise from different states of the latent causal model.

translation between *unpaired* data. The proposed methods rely on a variety of techniques, including manifold alignment (Welch et al., 2017; Amodio and Krishnaswamy, 2018; Liu et al., 2019), matrix factorization (Duren et al., 2018), correlation analysis (Barkas et al., 2019; Stuart et al., 2019), coupled autoencoders (Yang and Uhler, 2019), optimal transport (Cao et al., 2022), regression analysis (Yuan and Duren, 2022), and semisupervised learning (Lin et al., 2022). Implicitly, these methods presume the existence of a *shared* latent space where the different modalities either completely align or at least overlap. However, to the best of our knowledge, none of these methods have rigorous *identifiability* guarantees, i.e., the methods are not guaranteed to recover a correct domain translation mapping even for infinite data. Our work advances the theoretical understanding of multi-domain integration by providing identifiability guarantees on recovering the shared latent space.

Group Independent Component Analysis. The primary tool that we use for identifiability is linear independent component analysis (ICA) (Comon, 1994; Eriksson and Koivunen, 2004). Many works extend ICA to the multi-domain setting. These methods primarily come from computational neuroscience, where different domains correspond to different subjects or studies. However, to the best of our knowledge, all prior works require pairing between samples. These works can be categorized based on whether the samples are assumed to be voxels (Calhoun et al., 2001; Esposito et al., 2005), time points (Svensén et al., 2002; Varoquaux et al., 2009; Hyvärinen and Ramkumar, 2013), or either (Beckmann and Smith, 2005; Sui et al., 2009). For reviews, see Calhoun et al. (2009) and Chabriel et al. (2014). Related are methods for *independent vector analysis* (Kim et al., 2006; Anderson et al., 2014; Bhinge et al., 2019) and multiset canonical correlation analysis (Nielsen, 2002; Li et al., 2011; Klami et al., 2014), which allow the latent variables to take on different values in each domain but still require sample pairing. Most of the mentioned methods lack identifiability guarantees, only newer work (Richard et al., 2021) provides sufficient conditions for identifiability. Furthermore, all mentioned methods assume that every latent variable is shared across all domains, while our setup allows for shared and domain-specific latent variables. Some methods, e.g., Lukic et al. (2002), Maneshi et al. (2016), and Pandeva and Forré (2023), permit both shared and domain-specific components, but only consider the paired setting. In this paper, we extend these results to the *unpaired* setting.

Latent Structure Discovery. Learning causal structure between latent variables has a long history, e.g., in measurement models (Silva et al., 2006). One recent line of work studies the problem under the assumption of access to interventional data (e.g., Liu et al., 2022; Squires et al., 2023). In particular, Squires et al. (2023) show that the latent graph is identifiable if the interventions are sufficiently diverse. Another line of work, closer to ours and not based on interventional data, shows that the graph is identified under certain sparsity assumptions on the mixing functions (Xie et al., 2020; Chen et al., 2022; Xie et al., 2022; Huang et al., 2022). However, these methods are not suitable in our setup since they require paired data in a single domain. One cannot apply them in each domain separately since it would be unclear how to combine the multiple latent causal graphs, that is, which of the latent variables are shared. In this work, we lay out sparsity assumptions on the mixing

functions that are tailored to the *unpaired multi-domain* setup. The works of Adams et al. (2021) and Zeng et al. (2021) may be considered closest to ours as they also treat a setting with multiple domains and unpaired data. However, our setup and results are more general. Adams et al. (2021) assume that the number of observed variables are the same in each domain, whereas we consider domains of *different dimensions* corresponding to the fact that observations may be of very different nature. Further, we allow for *shared and domain-specific* latent variables, where the number of shared latent variables is unknown, while in Adams et al. (2021) it is assumed that all latent variables are shared. Compared to Zeng et al. (2021), we consider a general but fixed number of observed variables, while Zeng et al. (2021) only show identifiability of the full model in a setup where the number of observed variables in each domain increases to infinity. On a more technical level, the conditions in Zeng et al. (2021) require two pure children to identify the shared latent graph, while we prove identifiability under the weaker assumption of two partial pure children; see Section 4 for precise definitions.

Notation. Let $\mathbb{N}$ be the set of nonnegative integers. For positive $n \in \mathbb{N}$, we define $[n] = \{1, \dots, n\}$. For a matrix $M \in \mathbb{R}^{a \times b}$, we denote by $M_{A,B}$ the submatrix containing the rows indexed by $A \subseteq [a]$ and the columns indexed by $B \subseteq [b]$. Moreover, we write M_B for the submatrix containing all rows but only the subset of columns indexed by B. Similarly, for a tuple $x = (x_1, \dots, x_b)$, we denote by x_B the tuple only containing the entries indexed by B. A matrix $Q = Q_\sigma \in \mathbb{R}^{p \times p}$ is a *signed permutation matrix* if it can be written as the product of a diagonal matrix D with entries $D_{ii} \in \{\pm 1\}$ and a permutation matrix $\widetilde{Q}_\sigma$ with entries $(\widetilde{Q}_\sigma)_{ij} = \mathbb{1}_{j=\sigma(i)}$, where σ is a permutation on p elements. Let P be a p-dimensional joint probability measure of a collection of random variables $Y_1, \dots, Y_p$. Then we denote by P_i the marginal probability measure such that $Y_i \sim P_i$. We say that P has *independent marginals* if the random variables Y_i are mutually independent. Moreover, we denote by $M\#P$ the d-dimensional *push-forward measure* under the linear map defined by the matrix $M \in \mathbb{R}^{d \times p}$. If Q is a signed permutation matrix and the probability measure P has independent marginals, then $Q\#P$ also has independent marginals. For univariate probability measures we use the shorthand $(-1)\#P = -P$.

2 Setup

Let $\mathcal{H} = [h]$ for $h \geq 1$, and let $Z = (Z_1, \dots, Z_h)$ be latent random variables that follow a linear structural equation model. That is, the variables are related by a linear equation

$$Z = AZ + \varepsilon, \tag{1}$$

with $h \times h$ parameter matrix $A = (a_{ij})$ and zero-mean, independent random variables $\varepsilon = (\varepsilon_1, \dots, \varepsilon_h)$ that represent stochastic errors. Assume that we have observed random vectors $X^e \in \mathbb{R}^{d_e}$ in multiple domains of interest $e \in [m]$, where the dimension d_e may vary across domains. Each random vector is the image under a linear function of a subset of the latent variables. In particular, we assume that there is a subset $\mathcal{L} \subseteq \mathcal{H}$ representing the shared latent space such that each X^e is generated via the mechanism

$$X^e = G^e \cdot \begin{pmatrix} Z_{\mathcal{L}} \\ Z_{I_e} \end{pmatrix}, \tag{2}$$

where $I_e \subseteq \mathcal{H} \setminus \mathcal{L}$. We say that the latent variable Z_{I_e} are *domain-specific* for domain $e \in [m]$ while the latent variables $Z_{\mathcal{L}}$ are *shared* across all domains. As already noted, we are motivated by settings where the shared latent variables $Z_{\mathcal{L}}$ capture the key causal relations and the different domains are able to give us combined information about these relations. Likewise, we may think about the domain-specific latent variables Z_{I_e} as "noise" in each domain, independent of the shared latent variables. Specific models are now derived from (1)-(2) by assuming specific (but unknown) sparsity patterns in A and G^e. Each model is given by a "large" directed acyclic graph (DAG) that encodes the multi-domain setup. To formalize this, we introduce pairwise disjoint index sets $V_1, \dots, V_m$, where V_e indexes the coordinates of X^e, i.e., $X^e = (X_v : v \in V_e)$ and $|V_e| = d_e$. Then $V = V_1 \cup \dots \cup V_m$ indexes all observed random variables. We define an m-domain graph such that the latent nodes are the only parents of observed nodes and there are no edges between shared and domain-specific latent nodes.

Definition 2.1. Let $\mathcal{G}$ be a DAG whose node set is the disjoint union $\mathcal{H} \cup V = \mathcal{H} \cup V_1 \cup \dots \cup V_m$. Let D be the edge set of $\mathcal{G}$. Then $\mathcal{G}$ is an *m-domain graph* with *shared latent nodes* $\mathcal{L} = [\ell] \subseteq \mathcal{H}$ if the following is satisfied:

1. All parent sets contain only latent variables, i.e., $\text{pa}(v) = \{w : w \to v \in D\} \subseteq \mathcal{H}$ for all $v \in \mathcal{H} \cup V$.

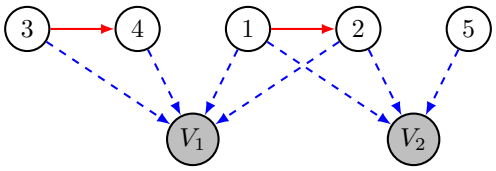


Figure 2: **Compact version of a 2-domain graph** $\mathcal{G}_2$ with five latent nodes and two domains V_1 and V_2. All observed nodes in each domain are represented by a single grey node. We draw a dashed blue edge from latent node $h \in \mathcal{H}$ to domain $V_e \subseteq V$ if $h \in S_e = \text{pa}(V_e)$. The random vectors associated to the two domains are uncoupled, that is, we do not observe their joint distribution.

2. The set $\mathcal{L}$ consists of the common parents of variables in all different domains, i.e., $u \in \mathcal{L}$ if and only if $u \in \text{pa}(v) \cap \text{pa}(w)$ for $v \in V_e$, $w \in V_f$ with $e \neq f$.

3. Let $I_e = S_e \setminus \mathcal{L}$ be the domain-specific latent nodes, where $S_e := \text{pa}(V_e) = \cup_{v \in V_e} \text{pa}(v) \subseteq \mathcal{H}$. Then there are no edges in D that connect a node in $\mathcal{L}$ and a node $\cup_{e=1}^m I_e$ or that connect a node in I_e and a node in I_f for any $e \neq f$.

To emphasize that a given DAG is an m-domain graph we write $\mathcal{G}_m$ instead of $\mathcal{G}$. We also say that S_e is the set of *latent parents* in domain e and denote its cardinality by $s_e = |S_e|$. Note that the third condition in Definition 2.1 does not exclude causal relations between the domain-specific latent variables, that is, there may be edges $v \to w$ for $v, w \in I_e$. Since the sets I_e satisfy $I_e \cap I_f = \emptyset$ for $e \neq f$, we specify w.l.o.g. the indexing convention $I_e = \{\ell + 1 + \sum_{i=1}^{e-1} |I_i|, \ldots, \ell + \sum_{i=1}^{e} |I_i|\}$ and $h = \ell + \sum_{e=1}^{m} |I_e|$.

Example 2.2. Consider the compact version of a 2-domain graph in Figure 2. There are $h = 5$ latent nodes where $\mathcal{L} = \{1, 2\}$ are shared and $I_1 = \{3, 4\}$ and $I_2 = \{5\}$ are domain-specific. A full m-domain graph is given in Appendix B.

Each m-domain graph postulates a statistical model that corresponds to the structural equation model in (1) and the mechanisms in (2), with potentially sparse matrices A and G^e. For two sets of nodes $W, Y \subseteq \mathcal{H} \cup V$, we denote by D_{WY} the subset of edges $D_{WY} = \{y \to w \in D : w \in W, y \in Y\}$. Moreover, let $\mathbb{R}^{D_{WY}}$ be the set of real $|W| \times |Y|$ matrices $M = (m_{wy})$ with rows indexed by W and columns indexed by Y, such that the support is given by D_{WY}, that is, $m_{wy} = 0$ if $y \to w \notin D_{WY}$.

Definition 2.3. Let $\mathcal{G}_m = (\mathcal{H} \cup V, D)$ be an m-domain graph. Define the map

$$\begin{aligned} \phi_{\mathcal{G}_m} : \mathbb{R}^{D_{V\mathcal{H}}} \times \mathbb{R}^{D_{\mathcal{H}\mathcal{H}}} &\longrightarrow \mathbb{R}^{|V| \times |\mathcal{H}|} \\ (G, A) &\longmapsto G \cdot (I - A)^{-1}. \end{aligned}$$

Then the *multi-domain causal representation (MDCR) model* $\mathcal{M}(\mathcal{G}_m)$ is given by the set of probability measures $P_X = B\#P$, where $B \in \text{Im}(\phi_{\mathcal{G}_m})$ and P is an h-dimensional probability measure with independent, mean-zero marginals $P_i, i \in \mathcal{H}$. We say that the pair (B, P) is a *representation* of $P_X \in \mathcal{M}(\mathcal{G}_m)$.

Definition 2.3 corresponds to the model defined in Equations (1) and (2). If $P_X \in \mathcal{M}(\mathcal{G}_m)$ with representation (B, P), then P_X is the joint distribution of the observed domains $X = (X^1, \ldots, X^m)$. The distribution of the random variables ε_i in Equation (1) is given by the marginals P_i. Moreover, for any matrix $G \in \mathbb{R}^{D_{V\mathcal{H}}}$, we denote the submatrix $G^e = G_{V_e, S_e} \in \mathbb{R}^{d_e \times s_e}$ which coincides with the matrix G^e from Equation (2). For the graph in Figure 2, we compute a concrete example of the matrix B in Example B.1 in the Appendix. Importantly, in the rest of the paper we assume to only observe the marginal distribution P_{X^e} in each domain but not the joint distribution P_X.

Ultimately, we are interested in recovering the graph $G_{\mathcal{L}} = (\mathcal{L}, D_{\mathcal{L}\mathcal{L}})$ among the shared latent nodes. We proceed by a two-step approach: In Section 3 we recover the representation (B, P) of the joint distribution P_X. To be precise, we recover a matrix $\widehat{B}$ that is equal to B up to certain permutations of the columns. Then we use the matrix $\widehat{B}$ to recover the shared latent graph $G_{\mathcal{L}}$ in Section 4 and show that recovery is possible up to trivial relabeling of latent nodes that appear in the same position of the causal order.

3 Joint Distribution

To identify the joint distribution P_X, we apply identifiability results from linear ICA in each domain separately and match the recovered probability measures P_i for identifying which of them are shared, that is, whether or not $i \in \mathcal{L}$. Let $\mathcal{G}_m$ be an m-domain graph with shared latent nodes $\mathcal{L}$, and let $P_X \in \mathcal{M}(\mathcal{G}_m)$ with representation (B, P). Recall that $B = G(I - A)^{-1}$ with $G \in \mathbb{R}^{D_{V\mathcal{H}}}$ and $A \in \mathbb{R}^{D_{\mathcal{H}\mathcal{H}}}$. We make the following technical assumptions.

(C1) (Different error distributions.) The marginal distributions $P_i, i \in \mathcal{H}$ are non-degenerate, non-symmetric and have unit variance. Moreover, the measures are pairwise different to each other and to the flipped versions, that is, $P_i \neq P_j$ and $P_i \neq -P_j$ for all $i, j \in \mathcal{H}$ with $i \neq j$. Subsequently, we let d be a distance on the set of univariate Borel probability measures such that $d(P_i, P_j) \neq 0$ and $d(P_i, -P_j) \neq 0$ for $i \neq j$.

(C2) (Full rank of mixing.) For each $e \in [m]$, the matrix $G^e = G_{V_e, S_e} \in \mathbb{R}^{d_e \times s_e}$ is of full column rank.

By not allowing symmetric distributions in Condition (C1), we assume in particular that the distributions of the errors are non-Gaussian. Non-Gaussianity together with the assumptions of pairwise different and non-symmetric error distributions allow us to extend the results on identifiability of linear ICA to the unpaired multi-domain setup and to identify the joint distribution. In particular, the assumption of pairwise different error distributions allows for "matching" the distributions across domains to identify the ones corresponding to the shared latent space. Non-symmetry accounts for the sign-indeterminacy of linear ICA when matching the distributions. We discuss the necessity of these assumptions in Remark 3.2 and, in more detail, in Appendix C. Note that Condition (C1) is always satisfied in a generic sense, that is, randomly chosen probability distributions on the real line are pairwise different and non-symmetric with probability one. Finally, Condition (C2) requires in particular that for each shared latent node $k \in \mathcal{L}$ there is at least one node $v \in V_e$ in every domain $e \in [m]$ such that $k \in \text{pa}(v)$.

Under Conditions (C1) and (C2) we are able to derive a sufficient condition for identifiability of the joint distribution. Let $SP(p)$ be the set of $p \times p$ signed permutation matrices. We define the set of signed permutation *block matrices*:

$$\Pi = \{\text{diag}(\Psi_{\mathcal{L}}, \Psi_{I_1}, \ldots, \Psi_{I_m}) : \Psi_{\mathcal{L}} \in SP(\ell) \text{ and } \Psi_{I_e} \in SP(|I_e|)\}.$$

Our main result is the following.

Theorem 3.1. *Let $\mathcal{G}_m$ be an m-domain graph with shared latent nodes $\mathcal{L} = [\ell]$, and let $P_X \in \mathcal{M}(\mathcal{G}_m)$ with representation (B, P). Suppose that $m \geq 2$ and that Conditions (C1) and (C2) are satisfied. Let $(\widehat{\ell}, \widehat{B}, \widehat{P})$ be the output of Algorithm 1. Then $\widehat{\ell} = \ell$ and*

$$\widehat{B} = B \cdot \Psi \quad \textit{and} \quad \widehat{P} = \Psi^\top \# P,$$

for a signed permutation block matrix $\Psi \in \Pi$.

Theorem 3.1 says that the matrix B is identifiable up to signed block permutations of the columns. Under the assumptions of Theorem 3.1 it holds that $\widehat{B}\#\widehat{P}$ is equal to P_X. That is, the joint distribution of the domains is identifiable.

Remark 3.2. While Theorem 3.1 is a sufficient condition for identifiability of the joint distribution, we emphasize that pairwise different error distributions are in most cases also necessary; we state the exact necessary condition in Proposition C.1 in the Appendix. Said differently, if one is willing to assume that conceptually different latent variables also follow a different distribution, then identification of these variables is possible, and otherwise (in most cases) not. Apart from pairwise different error distributions, non-symmetry is then required to fully identify the joint distribution whose dependency structure is determined by the shared latent variables. If the additional assumption on non-symmetry is not satisfied, then it is still possible to identify the shared, conceptually different latent variables, which becomes clear by inspecting the proofs of Theorem 3.1 and Proposition C.1. The non-identifiability of the joint distribution would only result in sign indeterminacy, that is, entries of the matrix $\widehat{B}$ could have a flipped sign.

Remark 3.3. By checking the proof of Theorem 3.1, the careful reader may notice that the statement of the theorem still holds true when we relax the third condition in the definition of an m-domain

Algorithm 1 IdentifyJointDistribution

1: **Input:** Probability measures P_{X^e} for all $e \in [m]$.
2: **Output:** Number of shared latent variables $\widehat{\ell}$, matrix $\widehat{B}$ and probability measure $\widehat{P}$ with independent marginals.
3: **for** $e \in [m]$ **do**
4: Linear ICA: Find the smallest value $\widehat{s}_e$ such that $P_{X^e} = \widehat{B}^e \# P^e$ for a matrix $\widehat{B}^e \in \mathbb{R}^{d_e \times \widehat{s}_e}$ and an $\widehat{s}_e$-dimensional probability measure P^e with independent, mean-zero and unit-variance marginals P^e_i.
5: **end for**
6: Matching: Let $\widehat{\ell}$ be the maximal number such that there are signed permutation matrices $\{Q^e\}_{e\in[m]}$ satisfying

$$d([(Q^e)^\top \# P^e]_i, [(Q^f)^\top \# P^f]_i) = 0$$

for all $i = 1, \dots, \widehat{\ell}$ and for all $f \neq e$. Let $\widehat{\mathcal{L}} = \{1, \dots, \widehat{\ell}\}$.
7: Construct the matrix $\widehat{B}$ and the tuple of probability measures $\widehat{P}$ given by

$$\widehat{B} = \left(\begin{array}{c|ccc} [\widehat{B}^1 Q^1]_{\widehat{\mathcal{L}}} & [\widehat{B}^1 Q^1]_{[\widehat{s}_1]\setminus\widehat{\mathcal{L}}} & & \\ \vdots & & \ddots & \\ [\widehat{B}^m Q^m]_{\widehat{\mathcal{L}}} & & & [\widehat{B}^m Q^m]_{[\widehat{s}_m]\setminus\widehat{\mathcal{L}}} \end{array}\right) \text{ and } \widehat{P} = \begin{pmatrix} [(Q^1)^\top \# P^1]_{\widehat{\mathcal{L}}} \\ [(Q^1)^\top \# P^1]_{[\widehat{s}_1]\setminus\widehat{\mathcal{L}}} \\ \vdots \\ [(Q^m)^\top \# P^m]_{[\widehat{s}_m]\setminus\widehat{\mathcal{L}}} \end{pmatrix}.$$

8: **return** $(\widehat{\ell}, \widehat{B}, \widehat{P})$.

graph. That is, one may allow directed paths from shared to domain-specific latent nodes but not vice versa. For example, an additional edge $1 \to 4$ between the latent nodes 1 and 4 would be allowed in the graph in Figure 2. In this case, the dependency structure of the domains is still determined by the shared latent space. However, the structural assumption that there are no edges between shared and domain-specific latent nodes is made for identifiability of the shared latent graph in Section 4.

Remark 3.4. The computational complexity of Algorithm 1 depends on the complexity of the chosen linear ICA algorithm, to which we make m calls. Otherwise, the dominant part is the matching in Line 6 with worst case complexity $\mathcal{O}(m \cdot \max_{e\in[m]} d_e^2)$, where we recall that m is the number of domains and d_e is the dimension of domain e.

In Appendix D we state a complete version of Algorithm 1 for the finite sample setting. In particular, we provide a method for the matching in Line 6 based on the two-sample Kolmogorov-Smirnov test. For finite samples, there might occur false discoveries, that is, distributions are matched that are actually not the same. With our method, we show that the probability of falsely discovering shared nodes shrinks exponentially with the number of domains.

4 Identifiability of the Causal Graph

We return to our goal of identifying the causal graph $\mathcal{G}_\mathcal{L} = (\mathcal{L}, D_{\mathcal{L}\mathcal{L}})$ among the shared latent nodes. By Theorem 3.1, we can identify the representation (B, P) of $P_X \in \mathcal{M}(\mathcal{G}_m)$ from the marginal distributions. In particular, we recover the matrix $\widehat{B} = B\Psi$ for a signed permutation block matrix $\Psi \in \Pi$. Moreover, we know which columns correspond to the shared latent nodes. That is, we know that the submatrix $\widehat{B}_\mathcal{L}$ obtained by only considering the columns indexed by $\mathcal{L} = \widehat{\mathcal{L}} = [\ell]$ is equal to $B_\mathcal{L}\Psi_\mathcal{L}$, where $\Psi_\mathcal{L} \in SP(\ell)$.

Problem 4.1. Let $B \in \mathrm{Im}(\phi_{\mathcal{G}_m})$ for an m-domain graph $\mathcal{G}_m$ with shared latent nodes $\mathcal{L}$. Given $\widehat{B}_\mathcal{L} = B_\mathcal{L}\Psi_\mathcal{L}$ with $\Psi_\mathcal{L}$ a signed permutation matrix, when is it possible to identify the graph $\mathcal{G}_\mathcal{L}$?

Recently, Xie et al. (2022) and Dai et al. (2022) show that, in the one-domain setting with independent additive noise, the latent graph can be identified if each latent variable has at least two pure children. We obtain a comparable result tailored to the multi-domain setup.

Definition 4.2. Let $\mathcal{G}_m = (\mathcal{H} \cup V, D)$ be an m-domain graph with shared latent nodes $\mathcal{L} \subseteq \mathcal{H}$. For $k \in \mathcal{L}$, we say that an observed node $v \in V$ is a *partial pure child* of k if $\mathrm{pa}(v) \cap \mathcal{L} = \{k\}$.

Algorithm 2 IdentifySharedGraph

Input: Matrix $B^* \in \mathbb{R}^{|V| \times \ell}$.
Output: Parameter matrix $\widehat{A} \in \mathbb{R}^{\ell \times \ell}$.
Remove rows $B^*_{i,\mathcal{L}}$ from the matrix B^* that are completely zero.
Find tuples $(i_k, j_k)_{k \in \mathcal{L}}$ with $i_k \neq j_k$ such that
    (i) $\text{rank}(B^*_{\{i_k, j_k\},\mathcal{L}}) = 1$ for all $k \in \mathcal{L}$ and
    (ii) $\text{rank}(B^*_{\{i_k, i_q\},\mathcal{L}}) = 2$ for all $k, q \in \mathcal{L}$ with $k \neq q$.
Let $I = \{i_1, \ldots, i_\ell\}$ and consider $B^*_{I,\mathcal{L}} \in \mathbb{R}^{\ell \times \ell}$.
Find two permutation matrices R_1 and R_2 such that $W = R_1 B^*_{I,\mathcal{L}} R_2$ is lower triangular.
Multiply each column of W by the sign of its corresponding diagonal element. This yields a new matrix $\widetilde{W}$ with all diagonal elements positive.
Divide each row of $\widetilde{W}$ by its corresponding diagonal element. This yields a new matrix $\widetilde{W}'$ with all diagonal elements equal to one.
Compute $\widehat{A} = I - (\widetilde{W}')^{-1}$.
return $\widehat{A}$.

For a partial pure child $v \in V$, there may still be domain-specific latent nodes that are parents of v. Definition 4.2 only requires that there is exactly one parent that is in the set $\mathcal{L}$. This explains the name *partial* pure child; see Example B.2 in the Appendix for further elaboration.

W.l.o.g. we assume in this section that the shared latent nodes are *topologically ordered* such that $i \to j \in D_{\mathcal{L}\mathcal{L}}$ implies $i < j$ for all $i, j \in \mathcal{L}$. We further assume:

(C3) (Two partial pure children across domains.) For each shared latent node $k \in \mathcal{L}$, there exist two partial pure children.

(C4) (Rank faithfulness.) For any two subsets $Y \subseteq V$ and $W \subseteq \mathcal{L}$, we assume that

$$\text{rank}(B_{Y,W}) = \max_{B' \in \text{Im}(\phi_{\mathcal{G}_m})} \text{rank}(B'_{Y,W}).$$

The two partial pure children required in Condition (C3) may either be in distinct domains or in a single domain. This is a sparsity condition on the large mixing matrix G. In Appendix C we discuss that the identification of the joint latent graph is impossible without any sparsity assumptions. We conjecture that two partial pure children are not necessary, but we leave it open for future work to find a non-trivial necessary condition. Roughly speaking, we assume in Condition (C4) that no configuration of edge parameters coincidentally yields low rank. The set of matrices $B \in \text{Im}(\phi_{\mathcal{G}_m})$ that violates (C4) is a subset of measure zero of $\text{Im}(\phi_{\mathcal{G}_m})$ with respect to the Lebesgue measure. Note that our conditions do not impose constraints on the graph $\mathcal{G}_\mathcal{L}$. Our main tool to tackle Problem 4.1 will be the following lemma.

Lemma 4.3. *Let $B \in Im(\phi_{\mathcal{G}_m})$ for an m-domain graph $\mathcal{G}_m$. Suppose that Condition (C4) is satisfied and that there are no zero-rows in $B_\mathcal{L}$. Let $v, w \in V$. Then rank$(B_{\{v,w\},\mathcal{L}}) = 1$ if and only if there is a node $k \in \mathcal{L}$ such that both v and w are partial pure children of k.*

The condition on no zero-rows in Lemma 4.3 is needed since we always have $\text{rank}(B_{\{v,w\},\mathcal{L}}) \leq 1$ if one of the two rows is zero. However, this is no additional structural assumption since we allow zero-rows when identifying the latent graph; c.f. Algorithm 2. The lemma allows us to find partial pure children by testing ranks on the matrix $\widehat{B}_\mathcal{L}$. If (i_1, j_1) and (i_2, j_2) are partial pure children of two nodes in $\mathcal{L}$, we make sure that these two nodes are different by checking that $\text{rank}(B_{\{i_1,i_2\},\mathcal{L}}) = 2$.

For a DAG $G = (V, D)$, we define $\mathcal{S}(G)$ to be the set of permutations on $|V|$ elements that are consistent with the DAG, i.e., $\sigma \in \mathcal{S}(G)$ if and only if $\sigma(i) < \sigma(j)$ for all edges $i \to j \in D$. The following result is the main result of this section.

Theorem 4.4. *Let $\widehat{B} = B\Psi$ with $B \in Im(\phi_{\mathcal{G}_m})$ and $\Psi \in \Pi$, and define $B^* = \widehat{B}_\mathcal{L}$ to be the input of Algorithm 2. Assume that Conditions (C3) and (C4) are satisfied, and let $\widehat{A}$ be the output of Algorithm 2. Then $\widehat{A} = Q_\sigma^\top A_{\mathcal{L},\mathcal{L}} Q_\sigma$ for a signed permutation matrix Q_σ with $\sigma \in \mathcal{S}(\mathcal{G}_\mathcal{L})$. Moreover, if $G_{vk} > 0$ for $G \in \mathbb{R}^{D_V \mathcal{H}}$ whenever v is a pure child of k, then Q_σ is a permutation matrix.*

Theorem 4.4 says that the graph $\mathcal{G}_{\mathcal{L}}$ can be recovered up to a permutation of the nodes that preserves the property that $i \to j$ implies $i < j$; see Remark 4.5. Since the columns of the matrix $\widehat{B}$ are not only permuted but also of different signs, we solve the sign indeterminacy column-wise in Line 7 before removing the scaling indeterminacy row-wise in Line 8. In case the coefficients of partial pure children are positive, this ensures that Q_σ is a *permutation matrix* and we have no sign indeterminacy. In Appendix D we adapt Algorithm 2 for the empirical data setting, where we only have $\widehat{B}_{\mathcal{L}} \approx B_{\mathcal{L}}\psi_{\mathcal{L}}$.

Remark 4.5. Let $\widehat{A}$ be the output of Alg. 2. Then we construct the graph $\widehat{G}_{\mathcal{L}} = (\mathcal{L}, \widehat{D}_{\mathcal{L}\mathcal{L}})$ as the graph with edges $j \to i \in \widehat{D}_{\mathcal{L}\mathcal{L}}$ if and only if $\widehat{A}_{ij} \neq 0$. Condition (C4) ensures that $\widehat{G}_{\mathcal{L}}$ is equivalent to $\mathcal{G}_{\mathcal{L}}$ in the sense that there is a permutation $\sigma \in \mathcal{S}(\mathcal{G}_{\mathcal{L}})$ such that $\widehat{D}_{\mathcal{L}\mathcal{L}} = \{\sigma(i) \to \sigma(j) : i \to j \in D_{\mathcal{L}\mathcal{L}}\}$.

Example 4.6. As highlighted in the introduction, the unpaired multi-domain setup is motivated by applications from single-cell biology. For example, consider the domains of (i) gene expression and (ii) high-level phenotypic features extracted from imaging assays (e.g. McQuin et al., 2018). We argue that the requirement of two partial pure children is justifiable on such data as follows. The condition requires, for example, that for each shared latent variable, (i) the expression of some gene depends only upon that shared latent variable plus domain-specific latent variables, and (ii) one of the high-level phenotypic features depends only on the same latent feature plus domain-specific latent variables. Many genes have highly specialized functions, so (i) is realistic, and similarly many phenotypic features are primarily controlled by specific pathways, so (ii) is justified.

Remark 4.7. In Algorithm 2, we determine the rank of a matrix by Singular Value Decomposition, which has worst case complexity $\mathcal{O}(mn\min\{n,m\})$ for an $m \times n$ matrix. Since Line 4 is the dominant part, we conclude that the worst case complexity of Algorithm 2 is $\mathcal{O}(|V|^2 \cdot \ell)$.

5 Simulations

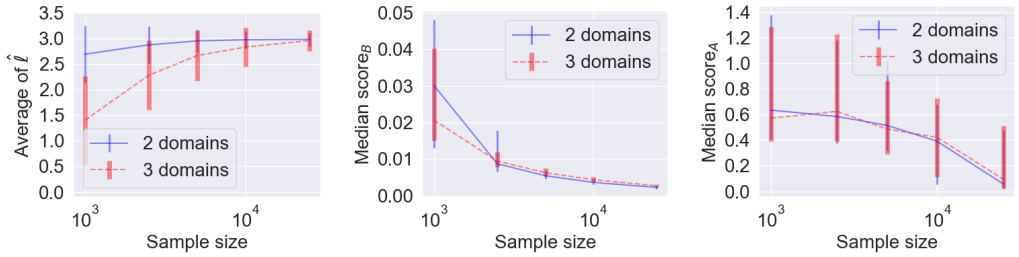
In this section we report on a small simulation study to illustrate the validity of our adapted algorithms for finite samples (detailed in Appendix D). We emphasize that this should only serve as a proof of concept as the focus of our work lies on identifiability. In future work one may develop more sophisticated methods; c.f. Appendix G. The adapted algorithms have a hyperparameter γ, which is a threshold on singular values to determine the rank of a matrix. In our simulations we use $\gamma = 0.2$.

Data Generation. In each experiment we generate 1000 random models with $\ell = 3$ shared latent nodes. We consider different numbers of domains $m \in \{2, 3\}$ and assume that there are $|I_e| = 2$ domain-specific latent nodes for each domain. The dimensions are given by $d_e = d/m$ for all $e \in [m]$ and $d = 30$. We sample the m-domain graph $\mathcal{G}_m$ on the shared latent nodes as follows. First, we sample the graph $\mathcal{G}_{\mathcal{L}}$ from an Erdős-Rényi model with edge probability 0.75 and assume that there are no edges between other latent nodes, that is, between $\mathcal{L}$ and $\mathcal{H} \setminus \mathcal{L}$ and within $\mathcal{H} \setminus \mathcal{L}$. Then we fix two partial pure children for each shared latent node $k \in \mathcal{L}$ and collect them in the set W. The remaining edges from $\mathcal{L}$ to $V \setminus W$ and from $\mathcal{H}$ to V are sampled from an Erdős-Rényi model with edge probability 0.9. Finally, the (nonzero) entries of G and A are sampled from Unif$(\pm[0.25, 1])$. The distributions of the error variables are specified in Appendix E. For simplicity, we assume that the sample sizes coincide, that is, $n_e = n$ for all $e \in [m]$, and consider $n \in \{1000, 2500, 5000, 10000, 25000\}$.

Results. First, we plot the average number of shared nodes $\widehat{\ell}$ in our experiments in Figure 3 (a). Especially for low sample sizes, we see that fewer shared nodes are detected with more domains. However, by inspecting the error bars we also see that the probability of detecting too many nodes $\widehat{\ell} > \ell$ decreases drastically when considering 3 instead of 2 domains. This suggests that the number of falsely detected shared nodes is very low, as expected by Theorem D.3. Our findings show that more domains lead to a more conservative discovery of shared nodes, but whenever a shared node is determined this is more certain. Moreover, we measure the error in estimating $\widehat{B}_{\widehat{\mathcal{L}}}$ in Figure 3 (b), that is, the error in the "shared" columns. We take

$$\text{score}_B(\widehat{B}_{\widehat{\mathcal{L}}}) = \begin{cases} \min_{\Psi \in SP(\ell)} \beta_{\ell,\widehat{\ell}}^{-1/2} \|\widehat{B}_{\widehat{\mathcal{L}}} - [B_{\mathcal{L}}\Psi]_{\widehat{\mathcal{L}}}\|_{\mathcal{F}} & \text{if } \widehat{\ell} \leq \ell, \\ \min_{\Psi \in SP(\widehat{\ell})} \beta_{\ell,\widehat{\ell}}^{-1/2} \|[\widehat{B}_{\widehat{\mathcal{L}}}\Psi]_{\mathcal{L}} - B_{\mathcal{L}}\|_{\mathcal{F}} & \text{if } \widehat{\ell} > \ell, \end{cases}$$

where $\|\cdot\|_{\mathcal{F}}$ denotes the Frobenius norm and $\beta_{\ell,\widehat{\ell}} = \min\{\ell, \widehat{\ell}\} \cdot \sum_{e=1}^{m} d_e$ denotes the number of entries of the matrix over which the norm is taken. In the cases $\ell = \widehat{\ell}$, we also measure the



(a) (b) (c)

Figure 3: **Results.** Logarithmic scale on the x-axis. Error bars in (a) are one standard deviation of the mean and in (b) and (c) they are the interquartile range.

performance of recovering the shared latent graph $\mathcal{G}_\mathcal{L}$ in Figure 3 (c) by taking

$$\text{score}_A(\widehat{A}) = \min_{Q_\sigma \in SP(\ell) \text{ s.t. } \sigma \in S(\mathcal{G}_\mathcal{L})} \frac{1}{\ell} \|Q_\sigma^\top \widehat{A} Q_\sigma - A_{\mathcal{L},\mathcal{L}}\|_\mathcal{F}.$$

As expected, the median estimation errors for $B_\mathcal{L}$ and $A_{\mathcal{L},\mathcal{L}}$ decrease with increasing sample size. In Appendix F we provide additional simulations with larger ℓ. Moreover, we consider setups where we violate specific assumptions, such as pairwise different distributions (C1) and two partial pure children (C3). The results emphasize that the conditions are necessary for the algorithms provided. The computations were performed on a single thread of an Intel Xeon Gold 6242R processor (3.1 GHz), with a total computation time of 12 hours for all simulations presented in this paper (including Appendix).

6 Discussion

This work introduces the problem of causal representation learning from *unpaired* multi-domain observations, in which multiple domains provide complementary information about a set of shared latent nodes that are the causal quantities of primary interest. For this problem, we laid out a setting in which we can provably identify the causal relations among the shared latent nodes. To identify the desired causal structure, we proposed a two-step approach where we first make use of linear ICA in each domain separately and match the recovered error distributions to identify shared nodes and the joint distribution of the domains. In the second step, we identify the causal structure among the shared latent variables by testing rank deficiencies in the "overall mixing matrix" B. To the best of our knowledge, our guarantees are the first principled identifiability results for shared causal representations in a general, unpaired multi-domain setting.

We proposed algorithms for recovering the joint distribution and the shared latent space making our proofs constructive. While our focus is on identifiability guarantees, we show in Appendix D how our proofs give rise to algorithms for the finite sample setting. Moreover, we propose a method to match approximate error distributions and show that the probability of falsely discovering shared nodes decreases exponentially in the number of domains. Our work opens up numerous directions for future work as we discuss in Appendix G.

Acknowledgments and Disclosure of Funding

This project was initiated while the first author was a visitor at the Eric and Wendy Schmidt Center of the Broad Institute of MIT and Harvard. The project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 883818), NCCIH/NIH (1DP2AT012345), ONR (N00014-22-1-2116), DOE-ASCR (DE-SC0023187), NSF (DMS-1651995), the MIT-IBM Watson AI Lab, and a Simons Investigator Award. Nils Sturma acknowledges support by the Munich Data Science Institute at the Technical University of Munich via the Linde/MDSI PhD Fellowship program. Chandler Squires was partially supported by an NSF Graduate Research Fellowship.

References

Adams, J., Hansen, N., and Zhang, K. (2021). Identification of partially observed linear causal models: Graphical conditions for the non-gaussian and heterogeneous cases. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W., editors, *Advances in Neural Information Processing Systems*, volume 34, pages 22822–22833. Curran Associates, Inc.

Ahuja, K., Mahajan, D., Wang, Y., and Bengio, Y. (2023). Interventional causal representation learning. In Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., and Scarlett, J., editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 372–407. PMLR.

Amodio, M. and Krishnaswamy, S. (2018). MAGAN: Aligning biological manifolds. In Dy, J. and Krause, A., editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 215–223. PMLR.

Anderson, M., Fu, G.-S., Phlypo, R., and Adalı, T. (2014). Independent vector analysis: Identification conditions and performance bounds. *IEEE Trans. Signal Process.*, 62(17):4399–4410.

Bach, F. R. and Jordan, M. I. (2003). Kernel independent component analysis. *J. Mach. Learn. Res.*, 3(1):1–48.

Barkas, N., Petukhov, V., Nikolaeva, D., Lozinsky, Y., Demharter, S., Khodosevich, K., and Kharchenko, P. V. (2019). Joint analysis of heterogeneous single-cell RNA-seq dataset collections. *Nat. Methods*, 16(8):695–698.

Beckmann, C. F. and Smith, S. M. (2005). Tensorial extensions of independent component analysis for multisubject fMRI analysis. *Neuroimage*, 25(1):294–311.

Beery, S., Van Horn, G., and Perona, P. (2018). Recognition in terra incognita. In Ferrari, V., Hebert, M., Sminchisescu, C., and Weiss, Y., editors, *Computer Vision – ECCV 2018*, pages 472–489, Cham. Springer International Publishing.

Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(8):1798–1828.

Bhinge, S., Mowakeaa, R., Calhoun, V. D., and Adalı, T. (2019). Extraction of time-varying spatiotemporal networks using parameter-tuned constrained IVA. *IEEE Trans. Med. Imaging*, 38(7):1715–1725.

Buchholz, S., Besserve, M., and Schölkopf, B. (2022). Function classes for identifiable nonlinear independent component analysis. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A., editors, *Advances in Neural Information Processing Systems*, volume 35, pages 16946–16961. Curran Associates, Inc.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.*, 36(5):411–420.

Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.*, 14(3):140–151.

Calhoun, V. D., Liu, J., and Adalı, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage*, 45(1):163–172.

Cao, K., Gong, Q., Hong, Y., and Wan, L. (2022). A unified computational framework for single-cell data integration with optimal transport. *Nat. Comm.*, 13(1).

Cardoso, J. and Souloumiac, A. (1993). Blind beamforming for non-gaussian signals. *IEE Proceedings F Radar and Signal Processing*, 140(6):362.

Chabriel, G., Kleinsteuber, M., Moreau, E., Shen, H., Tichavsky, P., and Yeredor, A. (2014). Joint matrices decompositions and blind source separation: A survey of methods, identification, and applications. *IEEE Signal Process. Mag.*, 31(3):34–43.

Chen, Z., Xie, F., Qiao, J., Hao, Z., Zhang, K., and Cai, R. (2022). Identification of linear latent variable model with arbitrary distribution. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(6):6350–6357.

Comon, P. (1994). Independent component analysis, a new concept? *Signal Process.*, 36(3):287–314.

Comon, P. and Jutten, C. (2010). *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Elsevier.

Dai, H., Spirtes, P., and Zhang, K. (2022). Independence testing-based approach to causal discovery under measurement error and linear non-gaussian models. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A., editors, *Advances in Neural Information Processing Systems*, volume 35, pages 27524–27536. Curran Associates, Inc.

Duren, Z., Chen, X., Zamanighomi, M., Zeng, W., Satpathy, A. T., Chang, H. Y., Wang, Y., and Wong, W. H. (2018). Integrative analysis of single-cell genomics data by coupled nonnegative matrix factorizations. *Proc. Natl. Acad. Sci.*, 115(30):7723–7728.

Ericsson, L., Gouk, H., Loy, C. C., and Hospedales, T. M. (2022). Self-supervised representation learning: Introduction, advances, and challenges. *IEEE Signal Process. Mag.*, 39(3):42–62.

Eriksson, J. and Koivunen, V. (2004). Identifiability, separability, and uniqueness of linear ICA models. *IEEE Signal Process. Lett.*, 11(7):601–604.

Esposito, F., Scarabino, T., Hyvärinen, A., Himberg, J., Formisano, E., Comani, S., Tedeschi, G., Goebel, R., Seifritz, E., and Di Salle, F. (2005). Independent component analysis of fMRI group studies by self-organizing clustering. *Neuroimage*, 25(1):193–205.

Essen, D. C. V., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., and Ugurbil, K. (2013). The WU-minn human connectome project: An overview. *Neuroimage*, 80:62–79.

Gentle, J. E. (1998). *Numerical linear algebra for applications in statistics*. Statistics and Computing. Springer-Verlag, New York.

Gessel, I. and Viennot, G. (1985). Binomial determinants, paths, and hook length formulae. *Adv. in Math.*, 58(3):300–321.

Gossi, F., Pati, P., Chouvardas, P., Martinelli, A. L., Kruithof-de Julio, M., and Rapsomaniki, M. A. (2023). Matching single cells across modalities with contrastive learning and optimal transport. *Brief. Bioinform.*, 24(3):bbad130.

Huang, B., Low, C. J. H., Xie, F., Glymour, C., and Zhang, K. (2022). Latent hierarchical causal structure discovery with rank constraints. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A., editors, *Advances in Neural Information Processing Systems*, volume 35, pages 5549–5561. Curran Associates, Inc.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.*, 10(3):626–634.

Hyvärinen, A. and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Netw.*, 13(4):411–430.

Hyvärinen, A. and Ramkumar, P. (2013). Testing independent component patterns by inter-subject or inter-session consistency. *Front. Hum. Neurosci.*, 7:94.

Khemakhem, I., Kingma, D., Monti, R., and Hyvärinen, A. (2020). Variational autoencoders and nonlinear ICA: A unifying framework. In Chiappa, S. and Calandra, R., editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 2207–2217. PMLR.

Kim, T., Lee, I., and Lee, T.-W. (2006). Independent vector analysis: definition and algorithms. In *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, pages 1393–1396. IEEE.

Klami, A., Virtanen, S., Leppäaho, E., and Kaski, S. (2014). Group factor analysis. *IEEE Trans. Neural Netw. Learn. Syst.*, 26(9):2136–2147.

Li, X.-L., Adalı, T., and Anderson, M. (2011). Joint blind source separation by generalized joint diagonalization of cumulant matrices. *Signal Process.*, 91(10):2314–2322.

Lin, Y., Wu, T.-Y., Wan, S., Yang, J. Y., Wong, W. H., and Wang, Y. (2022). scJoint integrates atlas-scale single-cell RNA-seq and ATAC-seq data with transfer learning. *Nat. Biotechnol.*, 40(5):703–710.

Lindström, B. (1973). On the vector representations of induced matroids. *Bull. London Math. Soc.*, 5:85–90.

Liu, J., Huang, Y., Singh, R., Vert, J.-P., and Noble, W. S. (2019). Jointly embedding multiple single-cell omics measurements. In Huber, K. T. and Gusfield, D., editors, *19th International Workshop on Algorithms in Bioinformatics (WABI 2019)*, volume 143 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:13, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Liu, Y., Zhang, Z., Gong, D., Gong, M., Huang, B., Hengel, A. v. d., Zhang, K., and Shi, J. Q. (2022). Identifying weight-variant latent causal models. *arXiv preprint arXiv:2208.14153*.

Lopez, R., Tagasovska, N., Ra, S., Cho, K., Pritchard, J., and Regev, A. (2022). Learning causal representations of single cells via sparse mechanism shift modeling. In *NeurIPS 2022 Workshop on Causality for Real-world Impact*.

Lukic, A. S., Wernick, M. N., Hansen, L. K., and Strother, S. C. (2002). An ICA algorithm for analyzing multiple data sets. In *Proceedings. International Conference on Image Processing*, volume 2, pages 821–824. IEEE.

Lyche, T. (2020). *Numerical linear algebra and matrix factorizations*, volume 22 of *Texts in Computational Science and Engineering*. Springer, Cham. With a foreword by Geir Dahl.

Maneshi, M., Vahdat, S., Gotman, J., and Grova, C. (2016). Validation of shared and specific independent component analysis (SSICA) for between-group comparisons in fMRI. *Front. Neurosci.*, 10:417.

McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B. A., Karhohs, K. W., Doan, M., Ding, L., Rafelski, S. M., Thirstrup, D., Wiegraebe, W., Singh, S., Becker, T., Caicedo, J. C., and Carpenter, A. E. (2018). CellProfiler 3.0: Next-generation image processing for biology. *PLOS Biol.*, 16(7):e2005970.

Mesters, G. and Zwiernik, P. (2022). Non-independent components analysis. *arXiv preprint arXiv:2206.13668*.

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., Bartsch, A. J., Jbabdi, S., Sotiropoulos, S. N., Andersson, J. L. R., Griffanti, L., Douaud, G., Okell, T. W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P. M., and Smith, S. M. (2016). Multimodal population brain imaging in the UK biobank prospective epidemiological study. *Nat. Neurosci.*, 19(11):1523–1536.

Nielsen, A. A. (2002). Multiset canonical correlations analysis and multispectral, truly multitemporal remote sensing data. *IEEE Trans. Image Process.*, 11(3):293–305.

Pandeva, T. and Forré, P. (2023). Multi-view independent component analysis with shared and individual sources. In Evans, R. J. and Shpitser, I., editors, *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pages 1639–1650. PMLR.

Richard, H., Ablin, P., Thirion, B., Gramfort, A., and Hyvärinen, A. (2021). Shared independent component analysis for multi-subject neuroimaging. *Adv. Neural Inf. Process. Syst.*, 34:29962–29971.

Roeder, G., Metz, L., and Kingma, D. (2021). On linear identifiability of learned representations. In Meila, M. and Zhang, T., editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 9030–9039. PMLR.

Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., and Bengio, Y. (2021). Toward causal representation learning. *Proc. IEEE*, 109(5):612–634.

Shafto, M. A., Tyler, L. K., Dixon, M., Taylor, J. R., Rowe, J. B., Cusack, R., Calder, A. J., Marslen-Wilson, W. D., Duncan, J., Dalgleish, T., Henson, R. N., Brayne, C., and Matthews, F. E. (2014). The cambridge centre for ageing and neuroscience (cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurol.*, 14(1).

Shimizu, S., Inazumi, T., Sogawa, Y., Hyvärinen, A., Kawahara, Y., Washio, T., Hoyer, P. O., and Bollen, K. (2011). DirectLiNGAM: a direct method for learning a linear non-Gaussian structural equation model. *J. Mach. Learn. Res.*, 12:1225–1248.

Silva, R., Scheine, R., Glymour, C., and Spirtes, P. (2006). Learning the structure of linear latent variable models. *J. Mach. Learn. Res.*, 7(8):191–246.

Squires, C., Seigal, A., Bhate, S. S., and Uhler, C. (2023). Linear causal disentanglement via interventions. In Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., and Scarlett, J., editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 32540–32560. PMLR.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. *Cell*, 177(7):1888–1902.

Sui, J., Adali, T., Pearlson, G. D., and Calhoun, V. D. (2009). An ICA-based method for the identification of optimal FMRI features and components using combined group-discriminative techniques. *Neuroimage*, 46(1):73–86.

Svensén, M., Kruggel, F., and Benali, H. (2002). ICA of fMRI group study data. *Neuroimage*, 16(3):551–563.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York. With applications to statistics.

Varoquaux, G., Sadaghiani, S., Poline, J. B., and Thirion, B. (2009). CanICA: Model-based extraction of reproducible group-level ICA patterns from fMRI time series. In *Medical Image Computing and Computer Aided Intervention*, page 1.

Wang, X., Hutchinson, R., and Mitchell, T. M. (2003). Training fmri classifiers to detect cognitive states across multiple human subjects. In Thrun, S., Saul, L., and Schölkopf, B., editors, *Proceedings of the 16th International Conference on Neural Information Processing Systems*, volume 16, pages 709–716. MIT Press.

Wang, Y. S. and Drton, M. (2020). High-dimensional causal discovery under non-Gaussianity. *Biometrika*, 107(1):41–59.

Welch, J. D., Hartemink, A. J., and Prins, J. F. (2017). MATCHER: manifold alignment reveals correspondence between single cell transcriptome and epigenome dynamics. *Genome Biol.*, 18(1):1–19.

Xie, F., Cai, R., Huang, B., Glymour, C., Hao, Z., and Zhang, K. (2020). Generalized independent noise condition for estimating latent variable causal graphs. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H., editors, *Adv. in Neural Inf. Process. Syst.*, volume 33, pages 14891–14902. Curran Associates, Inc.

Xie, F., Huang, B., Chen, Z., He, Y., Geng, Z., and Zhang, K. (2022). Identification of linear non-Gaussian latent hierarchical structure. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., and Sabato, S., editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 24370–24387. PMLR.

Yang, K. D., Belyaeva, A., Venkatachalapathy, S., Damodaran, K., Katcoff, A., Radhakrishnan, A., Shivashankar, G. V., and Uhler, C. (2021a). Multi-domain translation between single-cell imaging and sequencing data using autoencoders. *Nat. Comm.*, 12(1).

Yang, K. D. and Uhler, C. (2019). Multi-domain translation by learning uncoupled autoencoders. *Computational Biology Workshop, International Conference on Machine Learning*.

Yang, M., Liu, F., Chen, Z., Shen, X., Hao, J., and Wang, J. (2021b). Causalvae: Disentangled representation learning via neural structural causal models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9593–9602.

Yuan, Q. and Duren, Z. (2022). Integration of single-cell multi-omics data by regression analysis on unpaired observations. *Genome Biol.*, 23(1):1–19.

Zeng, Y., Shimizu, S., Cai, R., Xie, F., Yamamoto, M., and Hao, Z. (2021). Causal discovery with multi-domain lingam for latent factors. In Zhou, Z.-H., editor, *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, pages 2097–2103. International Joint Conferences on Artificial Intelligence Organization. Main Track.

Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2242–2251.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., and He, Q. (2021). A comprehensive survey on transfer learning. *Proc. IEEE*, 109(1):43–76.

A Proofs

Proof of Theorem 3.1. Let $P_X \in \mathcal{M}(\mathcal{G}_m)$ for an m-domain graph $\mathcal{G}_m = (\mathcal{H} \cup V, D)$ with shared latent nodes $\mathcal{L} = [\ell]$ and representation (B, P). By Condition (C1) the measure P has independent, non-degenerate, non-symmetric marginals P_i, $i \in \mathcal{H}$ with mean zero and variance one. Moreover, since $B \in \mathrm{Im}(\phi_{\mathcal{G}_m})$, we have $B = G(I - A)^{-1}$ for matrices $G \in \mathbb{R}^{D_{V\mathcal{H}}}$ and $A \in \mathbb{R}^{D_{\mathcal{H}\mathcal{H}}}$.

Fix one domain $e \in [m]$. Recall that we denote by $S_e = \mathrm{pa}(V_e) = \mathcal{L} \cup I_e$ the set of latent parents in domain e. Define the matrix

$$B^e := G_{V_e,S_e}[(I - A)^{-1}]_{S_e,S_e} = G^e[(I - A)^{-1}]_{S_e,S_e},$$

and observe that we can write $P_{X^e} = B_{V_e,\mathcal{H}} \# P = B^e \# P_{S_e}$. This is due to the fact that $G_{V_e,\mathcal{H}\setminus S_e} = 0$ and $[(I - A)^{-1}]_{S_e,\mathcal{H}\setminus S_e} = 0$ by the definition of an m-domain graph.

In particular, the equality $P_{X^e} = B^e \# P_{S_e}$ shows that the representation in Line 4 of Algorithm 1 exists. Now, we show that it is unique up to signed permutation by applying results on identifiability of linear ICA. Since G^e has full column rank by Condition (C2) and $[(I-A)^{-1}]_{S_e,S_e}$ is invertible, the matrix B^e also has full column rank. Let $P_{X^e} = \widehat{B}^e \# P^e$ be any representation, where $\widehat{B}^e \in \mathbb{R}^{d_e \times \widehat{s}_e}$ and P^e is an $\widehat{s}_e$-dimensional probability measure with independent, non-degenerate marginals P^e_i. Due to Condition (C1), all probability measures P_i are non-Gaussian and non-degenerate and therefore we have by Eriksson and Koivunen (2004, Theorem 3 and 4) the identities

$$\widehat{B}^e = B^e R^e \Lambda^e \quad \text{and} \quad P^e = \Lambda^e (R^e)^\top \# P_{S_e}, \tag{3}$$

where Λ^e is an $s_e \times s_e$ diagonal matrix with nonzero entries and R^e is an $s_e \times s_e$ permutation matrix. In particular, we have $\widehat{s}_e = s_e$, which means that $\widehat{B}^e \in \mathbb{R}^{d_e \times s_e}$ and that P^e is an s_e-dimensional probability measure. Line 4 also requires that each marginal P^e_i has unit variance. This removes the scaling indeterminacy in (3) and we have

$$\widehat{B}^e = B^e R^e D^e \quad \text{and} \quad P^e = D^e (R^e)^\top \# P_{S_e},$$

where D^e is a diagonal matrix with entries $D^e_{ii} \in \{\pm 1\}$. In particular, this means that the distributions P^e and P_{S_e} coincide up to permutation and sign of the marginals.

The matching in Line 6 identifies which components of P^e are shared. By Condition (C1), two components of different domains P^e_i and P^f_j are shared if and only if they coincide up to sign, that is, if and only if $d(P^e_i, P^f_j) = 0$ or $d(P^e_i, -P^f_j) = 0$. If their distribution coincide up to sign, than either $d(P^e_i, P^f_j) = 0$ or $d(P^e_i, -P^f_j) = 0$ but not both since Condition (C1) requires the distribution of the error variables to be non-symmetric. We conclude that in each domain $e \in [m]$ there exists an $s_e \times s_e$ signed permutation matrix Q^e such that

$$d([(Q^e)^\top \# P^e]_i, [(Q^f)^\top \# P^f]_i) = 0 \tag{4}$$

for all $i = 1, \ldots, \widehat{\ell}$ and for all $f \neq e$. In particular, $\widehat{\ell} = \ell$ and $\widehat{\mathcal{L}} = \mathcal{L}$.

It remains to show that $\widehat{B} = B\Psi$ and $\widehat{P} = \Psi^\top \# P$ for a signed permutation block matrix $\Psi \in \Pi$. By Equation (4), the distributions $[(Q^e)^\top \# P^e]_{\mathcal{L}}$ and $[(Q^e)^\top \# P^e]_{\mathcal{L}}$ coincide, which means that

$$(Q^e)^\top \# P^e = (Q^e)^\top D^e (R^e)^\top \# P_{S_e} = \begin{pmatrix} \Psi_{\mathcal{L}}^\top & 0 \\ 0 & \Psi_{I_e}^\top \end{pmatrix} \# \begin{pmatrix} P_{\mathcal{L}} \\ P_{I_e} \end{pmatrix}, \tag{5}$$

where $\Psi_{\mathcal{L}}$ is an $\ell \times \ell$ signed permutation matrix and Ψ_{I_e} is an $|I_e| \times |I_e|$ signed permutation matrix. Importantly, the matrix $\Psi_{\mathcal{L}}^\top$ does not depend on the domain $e \in [m]$. Hence, the matrix $\Phi^e := R^e D^e Q^e$ is a signed permutation matrix with block structure as in Equation (5). Moreover, we have

$$\widehat{B}^e Q^e = B^e R^e D^e Q^e = B^e \Phi^e = \left(B^e_{\mathcal{L}} \Psi_{\mathcal{L}} \;\middle|\; B^e_{[s_e]\setminus\mathcal{L}} \Psi_{I_e}\right),$$

which means that the matrix $\widehat{B}$ can be factorized as

$$\begin{aligned}\widehat{B} &= \left(\begin{array}{c|ccc} [\widehat{B}^1 Q^1]_{\widehat{\mathcal{L}}} & [\widehat{B}^1 Q^1]_{[\widehat{s}_1]\setminus\widehat{\mathcal{L}}} & & \\ \vdots & & \ddots & \\ [\widehat{B}^m Q^m]_{\widehat{\mathcal{L}}} & & & [\widehat{B}^m Q^m]_{[\widehat{s}_m]\setminus\widehat{\mathcal{L}}} \end{array}\right) \\ &= \left(\begin{array}{c|ccc} B^1_{\mathcal{L}}\Psi_{\mathcal{L}} & B^1_{[s_1]\setminus\mathcal{L}}\Psi_{I_1} & & \\ \vdots & & \ddots & \\ B^m_{\mathcal{L}}\Psi_{\mathcal{L}} & & & B^m_{[s_m]\setminus\mathcal{L}}\Psi_{I_m} \end{array}\right) \\ &= \left(\begin{array}{c|ccc} B^1_{\mathcal{L}} & B^1_{[s_1]\setminus\mathcal{L}} & & \\ \vdots & & \ddots & \\ B^m_{\mathcal{L}} & & & B^m_{[s_m]\setminus\mathcal{L}} \end{array}\right) \cdot \begin{pmatrix} \Psi_{\mathcal{L}} & & & \\ & \Psi_{I_1} & & \\ & & \ddots & \\ & & & \Psi_{I_m} \end{pmatrix} = B \cdot \Psi,\end{aligned}$$

where $\Psi \in \Pi$. Similarly, we have for all $e \in [m]$,

$$(Q^e)^\top \# P^e = (\Phi^e)^\top \# P_{S_e} = \begin{pmatrix} (\Psi_{\mathcal{L}})^\top \# P_{\mathcal{L}} \\ (\Psi_{I_e})^\top \# P_{I_e} \end{pmatrix}.$$

We conclude that

$$\begin{aligned}\widehat{P} &= \begin{pmatrix} [(Q^1)^\top \# P^1]_{\widehat{\mathcal{L}}} \\ [(Q^1)^\top \# P^1]_{[\widehat{s}_1]\setminus\widehat{\mathcal{L}}} \\ \vdots \\ [(Q^m)^\top \# P^m]_{[\widehat{s}_m]\setminus\widehat{\mathcal{L}}} \end{pmatrix} = \begin{pmatrix} (\Psi_{\mathcal{L}})^\top \# P_{\mathcal{L}} \\ (\Psi_{I_1})^\top \# P_{I_1} \\ \vdots \\ (\Psi_{I_m})^\top \# P_{I_m} \end{pmatrix} \\ &= \begin{pmatrix} \Psi_{\mathcal{L}} & & & \\ & \Psi_{I_1} & & \\ & & \ddots & \\ & & & \Psi_{I_m} \end{pmatrix}^\top \# \begin{pmatrix} P_{\mathcal{L}} \\ P_{I_1} \\ \vdots \\ P_{I_m} \end{pmatrix} = \Psi^\top \# P.\end{aligned}$$

□

Before proving Lemma 4.3 and Theorem 4.4 we fix some notation. Let $\mathcal{G}_m = (V \cup \mathcal{H}, D)$ be an m-domain graph. We denote by $\text{anc}(v) = \{k \in \mathcal{H} : \text{there is a directed path } k \to \cdots \to v \text{ in } \mathcal{G}_m\}$ the ancestors of a node $v \in V$. For subsets $W \subseteq V$, we denote $\text{anc}(W) = \bigcup_{v \in W} \text{anc}(w)$. Moreover, for $\mathcal{L} \subseteq \mathcal{H}$ and $v \in V$, we write shortly $\text{pa}_{\mathcal{L}}(v) = \text{pa}(v) \cap \mathcal{L}$.

Proof of Lemma 4.3. Let $B \in \text{Im}(\phi_{\mathcal{G}_m})$. Then we can write $B = G \cdot (I - A)^{-1}$ with

$$G = \left(\begin{array}{c|ccc} G_{V_1,\mathcal{L}} & G_{V_1,I_1} & & \\ \vdots & & \ddots & \\ G_{V_m,\mathcal{L}} & & & G_{V_m,I_m} \end{array}\right).$$

Moreover, observe that, by the definition of an m-domain-graph, the matrix $B_{V,\mathcal{L}}$ factorizes as

$$B_{V,\mathcal{L}} = G_{V,\mathcal{L}}[(I - A)^{-1}]_{\mathcal{L},\mathcal{L}}.$$

Now, suppose that i and j are partial pure children of a fixed node $k \in \mathcal{L}$. Then $\text{pa}_{\mathcal{L}}(i) = \{k\} = \text{pa}_{\mathcal{L}}(j)$. In particular, the only entry that may be nonzero in the row $G_{i,\mathcal{L}}$ is given by G_{ik} and the only entry that may be nonzero in the row $G_{j,\mathcal{L}}$ is given by G_{jk}. Thus, we have

$$B_{i,\mathcal{L}} = \sum_{q \in \mathcal{L}} G_{iq}[(I - A)^{-1}]_{q,\mathcal{L}} = G_{ik}[(I - A)^{-1}]_{k,\mathcal{L}}.$$

Similarly, it follows that $B_{j,\mathcal{L}} = G_{jk}[(I - A)^{-1}]_{k,\mathcal{L}}$. This means that the row $B_{j,\mathcal{L}}$ is a multiple of the row $B_{i,\mathcal{L}}$ and we conclude that $\text{rank}(B_{\{i,j\},\mathcal{L}}) \leq 1$. Equality holds due to the faithfulness condition (C4) which implies that $B_{ik} \neq 0$ and $B_{jk} \neq 0$, i.e., $B_{\{i,j\},\mathcal{L}}$ is not the null matrix.

For the other direction suppose that $\text{rank}(B_{\{i,j\},\mathcal{L}}) = 1$. By applying the Lindström-Gessel-Viennot Lemma (Gessel and Viennot, 1985; Lindström, 1973) equivalently as in Dai et al. (2022, Theorem 1 and 2), it can be seen that

$$\text{rank}(B_{\{i,j\},\mathcal{L}}) \leq \min\{|S| : S \text{ is a vertex cut from anc}(\mathcal{L}) \text{ to } \{i,j\}\}, \tag{6}$$

where S is a vertex cut from $\text{anc}(\mathcal{L})$ to $\{i,j\}$ if and only if there exists no directed path in $\mathcal{G}_m$ from $\text{anc}(\mathcal{L})$ to $\{i,j\}$ without passing through S. Moreover, equality holds in (6) for generic (almost all) choices of parameters. Since we assumed rank faithfulness in Condition (C4) we exclude cases where the inequality is strict and therefore have equality. By the definition of an m-domain graph we have that $\text{anc}(\mathcal{L}) = \mathcal{L}$. Thus, if $\text{rank}(B_{\{i,j\},\mathcal{L}}) = 1$, there must be a single node $k \in \mathcal{L}$ such that $\{k\}$ is a vertex cut from $\mathcal{L}$ to $\{i,j\}$. But then it follows that i and j have to be partial pure children of k by the definition of an m-domain graph and by using the assumption that there are no zero-rows in $B_{\mathcal{L}}$. □

To prove Theorem 4.4 we need the following auxiliary lemma.

Lemma A.1. *Let $G = (V, D)$ be a DAG with topologically ordered nodes $V = [p]$ and let M be a lower triangular matrix with entries $M_{ii} \neq 0$ for all $i = 1, \ldots, p$ and $M_{ij} \neq 0$ if and only if there is a directed path $j \to \cdots \to i$ in G. Let Q_{σ_1} and Q_{σ_2} be permutation matrices. Then the matrix $Q_{\sigma_1} M Q_{\sigma_2}$ is lower triangular if and only if $\sigma_2 = \sigma_1^{-1}$ and $\sigma_2 \in \mathcal{S}(G)$.*

Proof of Lemma A.1. By the definition of a permutation matrix, we have

$$[Q_{\sigma_1} M Q_{\sigma_2}]_{ij} = M_{\sigma_1(i)\sigma_2^{-1}(j)} \text{ or, equivalently, } [Q_{\sigma_1} M Q_{\sigma_2}]_{\sigma_1^{-1}(i)\sigma_2(j)} = M_{ij}. \tag{7}$$

First, suppose that $\sigma_2 = \sigma_1^{-1}$ and $\sigma_2 \in \mathcal{S}(G)$, and let $i, j \in [p]$ such that $\sigma_2(i) < \sigma_2(j)$. Then, by the definition of $\mathcal{S}(G)$, there is no directed path $j \to \cdots \to i$ in the graph G and therefore we have $M_{ij} = 0$. But this means that $[Q_{\sigma_1} M Q_{\sigma_2}]_{\sigma_2(i)\sigma_2(j)} = 0$ and we conclude that the matrix $Q_{\sigma_1} M Q_{\sigma_2}$ is lower triangular.

Now, assume that $Q_{\sigma_1} M Q_{\sigma_2}$ is lower triangular, where σ_1 and σ_2 are arbitrary permutations on the set $[p]$. Since M has no zeros on the diagonal, we have $M_{ii} = [Q_{\sigma_1} M Q_{\sigma_2}]_{\sigma_1^{-1}(i)\sigma_2(i)} \neq 0$ for all $i = 1, \ldots, p$. It follows that $\sigma_1^{-1}(i) \geq \sigma_2(i)$ for all $i = 1, \ldots, p$ because $Q_{\sigma_1} M Q_{\sigma_2}$ is lower triangular. But this is only possible if the permutations coincide on all elements, i.e., we have $\sigma_2 = \sigma_1^{-1}$. It remains to show that $\sigma_2 = \sigma_1^{-1} \in \mathcal{S}(G)$. For any edge $j \to i \in D$ we have that $M_{ij} \neq 0$. Recalling Equation (7) this means that $[Q_{\sigma_1} M Q_{\sigma_2}]_{\sigma_2(i)\sigma_2(j)} \neq 0$. But since $Q_{\sigma_1} M Q_{\sigma_2}$ is lower triangular this can only be the case if $\sigma_2(j) < \sigma_2(i)$ which proves that $\sigma_2 \in \mathcal{S}(G)$. □

Proof of Theorem 4.4. Each latent node in $\mathcal{L}$ has two partial pure children by Condition (C3). After removing zero-rows in Line 3 of Algorithm 2 it holds by Lemma 4.3 that $\text{rank}(B^*_{\{i,j\},\mathcal{L}}) = 1$ if and only if there is a latent node in $\mathcal{L}$ such that i and j are both partial pure children of that latent node. Hence, each tuple $(i_k, j_k)_{k \in \mathcal{L}}$ in Line 4 of Algorithm 2 consists of two partial pure children of a certain latent node. The requirement $\text{rank}(B^*_{\{i_k,i_q\},\mathcal{L}}) = 2$ ensures that each pair of partial pure children has a different parent.

By the definition of an m-domain-graph and the fact that $B^* = \widehat{B}_{\mathcal{L}}$, for $I = \{i_1, \ldots, i_\ell\}$, we have the factorization

$$B^*_{I,\mathcal{L}} = B_{I,\mathcal{L}} \Psi_{\mathcal{L}} = G_{I,\mathcal{L}} (I - A)^{-1}_{\mathcal{L},\mathcal{L}} \Psi_{\mathcal{L}} = G_{I,\mathcal{L}} (I - A_{\mathcal{L},\mathcal{L}})^{-1} \Psi_{\mathcal{L}}, \tag{8}$$

where $G \in \mathbb{R}^{D_{V\mathcal{H}}}$, $A \in \mathbb{R}^{D_{\mathcal{H}\mathcal{H}}}$ and $\Psi_{\mathcal{L}}$ is a signed permutation matrix. Let Q_1 and Q_2 be permutation matrices and let Λ be a diagonal matrix with non-zero diagonal elements and let D be a diagonal matrix with entries in $\{\pm 1\}$. Then we can rewrite Equation (8) as

$$B^*_{I,\mathcal{L}} = Q_1 \underbrace{\Lambda (I - A_{\mathcal{L},\mathcal{L}})^{-1} D}_{=:M} Q_2.$$

Now, we apply Lemma A.1. Since we assume throughout Section 4 that the nodes $\mathcal{L}$ are topologically ordered, the matrix M is lower triangular with no zeros on the diagonal. Moreover, by Condition (C4) we have $M_{ij} \neq 0$ if and only if there is a directed path $j \to \cdots \to i$ in $\mathcal{G}_{\mathcal{L}}$. In Line 6 we find other permutation matrices R_1 and R_2 such that

$$W = R_1 B^*_{I,\mathcal{L}} R_2 = (R_1 Q_1) M (Q_2 R_2)$$

is lower triangular. Now, define the permutation matrices $Q_{\sigma_1} = R_1 Q_1$ and $Q_{\sigma_2} = Q_2 R_2$. Then we have by Lemma A.1 that $Q_{\sigma_1} = Q_{\sigma_2}^\top$ and that $\sigma_2 \in \mathcal{S}(\mathcal{G}_\mathcal{L})$. Hence, the matrix W factorizes as

$$W = Q_{\sigma_2}^\top M Q_{\sigma_2} = Q_{\sigma_2}^\top \Lambda (I - A_{\mathcal{L},\mathcal{L}})^{-1} D Q_{\sigma_2} = \widetilde{\Lambda} Q_{\sigma_2}^\top (I - A_{\mathcal{L},\mathcal{L}})^{-1} Q_{\sigma_2} \widetilde{D},$$

where $\widetilde{\Lambda}$ and $\widetilde{D}$ are diagonal matrices with the entries given by permutations of the entries of Λ and D. Lines 7 and 8 address the scaling and sign matrices $\widetilde{\Lambda}$ and $\widetilde{D}$. In particular, we have that $\widetilde{W}' = D' Q_{\sigma_2}^\top (I - A_{\mathcal{L},\mathcal{L}})^{-1} Q_{\sigma_2} D'$ for another diagonal matrix D' with entries in $\{\pm 1\}$, since each entry on the diagonal of $\widetilde{W}'$ is equal to 1. Thus, we have

$$\begin{aligned}
\widehat{A} &= I - (\widetilde{W}')^{-1} \\
&= I - (D' Q_{\sigma_2}^\top (I - A_{\mathcal{L},\mathcal{L}})^{-1} Q_{\sigma_2} D')^{-1} \\
&= I - D' Q_{\sigma_2}^\top (I - A_{\mathcal{L},\mathcal{L}}) Q_{\sigma_2 D'} \\
&= D' Q_{\sigma_2}^\top A_{\mathcal{L},\mathcal{L}} Q_{\sigma_2} D'.
\end{aligned}$$

Since $Q_{\sigma_2} D'$ is a signed permutation matrix with $\sigma_2 \in \mathcal{S}(\mathcal{G}_\mathcal{L})$, the first part of the theorem is proved. If $G_{vk} > 0$ whenever v is a pure child of k, the matrix $\widetilde{\Lambda}$ only has positive entries which means that D' is equal to the identity matrix. This proves the second part. □

B Additional Examples

The graph in Figure 4 is an m-domain graph corresponding to the compact version in Figure 2 in the main paper.

Example B.1. Consider the m-domain graph in Figure 4. The linear structural equation model among the latent variables is determined by lower triangular matrices of the form

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{43} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Moreover, the domain-specific mixing matrices G^e are of the form

$$G^1 = \begin{pmatrix} g_{11}^1 & g_{12}^1 & g_{13}^1 & 0 \\ g_{21}^1 & 0 & g_{23}^1 & 0 \\ 0 & g_{32}^1 & g_{33}^1 & g_{34}^1 \\ g_{41}^1 & g_{42}^1 & g_{43}^1 & g_{44}^1 \end{pmatrix} \quad \text{and} \quad G^2 = \begin{pmatrix} g_{11}^2 & 0 & g_{13}^2 \\ g_{21}^2 & g_{22}^2 & g_{23}^2 \\ 0 & g_{32}^2 & 0 \\ g_{41}^2 & g_{42}^2 & g_{43}^2 \\ g_{51}^2 & g_{52}^2 & 0 \end{pmatrix}.$$

Since the shared latent nodes are given by $\mathcal{L} = \{1, 2\}$, we have

$$G = \begin{pmatrix} g_{11}^1 & g_{12}^1 & g_{13}^1 & 0 & 0 \\ g_{21}^1 & 0 & g_{23}^1 & 0 & 0 \\ 0 & g_{32}^1 & g_{33}^1 & g_{34}^1 & 0 \\ g_{41}^1 & g_{42}^1 & g_{43}^1 & g_{44}^1 & 0 \\ g_{11}^2 & 0 & 0 & 0 & g_{13}^2 \\ g_{21}^2 & g_{22}^2 & 0 & 0 & g_{23}^2 \\ 0 & g_{32}^2 & 0 & 0 & 0 \\ g_{41}^2 & g_{42}^2 & 0 & 0 & g_{43}^2 \\ g_{51}^2 & g_{52}^2 & 0 & 0 & 0 \end{pmatrix}$$

and

$$B = G \cdot (I - A)^{-1} = \begin{pmatrix} a_{21} g_{12}^1 + g_{11}^1 & g_{12}^1 & g_{13}^1 & 0 & 0 \\ g_{21}^1 & 0 & g_{23}^1 & 0 & 0 \\ a_{21} g_{32}^1 & g_{32}^1 & a_{43} g_{34}^1 + g_{33}^1 & g_{34}^1 & 0 \\ a_{21} g_{42}^1 + g_{41}^1 & g_{42}^1 & a_{43} g_{44}^1 + g_{43}^1 & g_{44}^1 & 0 \\ g_{11}^2 & 0 & 0 & 0 & g_{13}^2 \\ a_{21} g_{22}^2 + g_{21}^2 & g_{22}^2 & 0 & 0 & g_{23}^2 \\ a_{21} g_{32}^2 & g_{32}^2 & 0 & 0 & 0 \\ a_{21} g_{42}^2 + g_{41}^2 & g_{42}^2 & 0 & 0 & g_{43}^2 \\ a_{21} g_{52}^2 + g_{51}^2 & g_{52}^2 & 0 & 0 & 0 \end{pmatrix}.$$

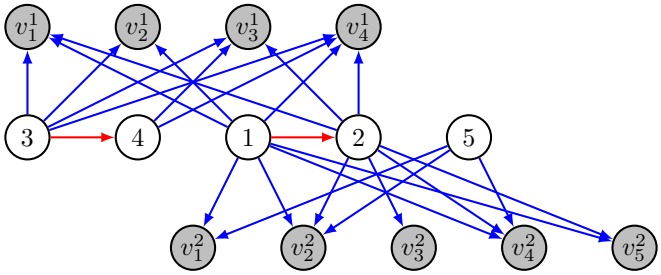


Figure 4: A 2-domain graph with 5 latent nodes and dimensions of the observed domains given by $|V_1| = d_1 = 4$ and $|V_2| = d_2 = 5$. We denote $V_e = \{v_1^e, \dots, v_{d_e}^e\}$, that is, the superscript indicates the domain a node belongs to.

Example B.2. Consider the m-domain graph in Figure 4. The partial pure children of node $1 \in \mathcal{L}$ are given by $\{v_2^1, v_1^2\}$ and the partial pure children of $2 \in \mathcal{L}$ are given by $\{v_3^1, v_3^2\}$. Moreover, by continuing Example B.1, we have that the matrix $B_{\mathcal{L}}$ is given by

$$B_{\mathcal{L}} = \begin{pmatrix} a_{21}g_{12}^1 + g_{11}^1 & g_{12}^1 \\ g_{21}^1 & 0 \\ a_{21}g_{32}^1 & g_{32}^1 \\ a_{21}g_{42}^1 + g_{41}^1 & g_{42}^1 \\ g_{11}^2 & 0 \\ a_{21}g_{22}^2 + g_{21}^2 & g_{22}^2 \\ a_{21}g_{32}^2 & g_{32}^2 \\ a_{21}g_{42}^2 + g_{41}^2 & g_{42}^2 \\ a_{21}g_{52}^2 + g_{51}^2 & g_{52}^2 \end{pmatrix}$$

It is easy to see that the two submatrices

$$\begin{pmatrix} g_{21}^1 & 0 \\ g_{11}^2 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a_{21}g_{32}^1 & g_{32}^1 \\ a_{21}g_{32}^2 & g_{32}^2 \end{pmatrix}$$

have rank one. The first matrix corresponds to the partial pure children $\{v_2^1, v_1^2\}$ in the graph in Figure 4 while the second matrix correspond to the partial pure children $\{v_3^1, v_3^2\}$. Note that the rank of any other 2×2 submatrix is generically (i.e., almost surely) equal to 2.

C Discussion of the Assumptions

In this section, we discuss aspects of Conditions (C1)-(C3) that allow for identifiability. In particular, we discuss the necessity of pairwise different and non-Gaussian error distributions if one is not willing to make further assumptions. Moreover, we elaborate on the sparsity conditions on the mixing matrix and explain why *some* sparsity assumption is necessary.

Pairwise Different Error Distributions. Given any two potentially different m-domain graphs $\mathcal{G}_m = (\mathcal{H} \cup V, D)$ and $\widetilde{\mathcal{G}}_m = (\widetilde{\mathcal{H}} \cup V, \widetilde{D})$, identifiability of the joint distribution in multi-domain causal representation models means that

$$B_{V_e,S_e} \# P_{S_e} = \widetilde{B}_{V_e,\widetilde{S}_e} \# \widetilde{P}_{\widetilde{S}_e} \text{ for all } e \in [m] \implies B \# P = \widetilde{B} \# \widetilde{P} \tag{9}$$

for any representation (B, P) of a distribution in $\mathcal{M}(\mathcal{G}_m)$ and any representation $(\widetilde{B}, \widetilde{P})$ of a distribution in $\mathcal{M}(\widetilde{\mathcal{G}}_m)$, where the matrices B_{V_e,S_e} and $\widetilde{B}_{V_e,\widetilde{S}_e}$ have full column rank for all $e \in [m]$. The left-hand side says that the marginal distributions in each domain are equal, while the right-hand side says that the joint distributions are equal. If there are m-domain graphs, such that the left-hand sides holds but the right-hand side is violated, then we say that the joint distribution is *not identifiable*.

We assume in this section that the marginal error distributions $P_i, i \in \mathcal{H}$ are non-Gaussian and have unit variance, but are not necessarily pairwise different or non-symmetric. Then the right-hand side holds if and only if the number of shared latent nodes in each graph is equal, i.e., $\ell = \widetilde{\ell}$, and there is a signed permutation matrix Ψ such that $B = \widetilde{B}\Psi$ and $P = \Psi^\top \# \widetilde{P}$. Here, the matrix Ψ does

not necessarily have a block structure. The equivalence is implied by the identifiability of the usual, one-domain linear ICA (see, e.g., Buchholz et al. (2022)) together with the fact that for $\ell \neq \widetilde{\ell}$, we have $|\mathcal{H}| \neq |\widetilde{\mathcal{H}}|$ and, therefore, the distributions on the right-hand have support over different dimensional subspaces.

Theorem 3.1 shows that assumptions (C1) and (C2) are sufficient for identifiability of the joint distribution. In particular, we show that they imply identifiability in a stronger sense, namely, that it follows from the left-hand side that $\ell = \widetilde{\ell}$ and $B = \widetilde{B}\Psi$ and $P = \Psi^\top \# \widetilde{P}$ for a signed permutation $\Psi \in \Pi$ with block structure. The next proposition reveals necessary conditions for identifiability.

Proposition C.1. *Let $\mathcal{G}_m$ be an m-domain graph with shared latent nodes $\mathcal{L} = [\ell]$, and let $P_X \in \mathcal{M}(\mathcal{G}_m)$ with representation (B, P). Suppose that $m \geq 2$ and that everything except the assumption about pairwise different error distributions in Conditions (C1) and (C2) is satisfied. Then, the joint distribution is not identifiable if one of the following holds:*

(i) There is $i, j \in \mathcal{L}$ such that $P_i = P_j$ or $P_i = -P_j$.

(ii) There is $i \in \mathcal{L}$ and $j \in I_e$ for some $e \in [m]$ such that $P_i = P_j$ or $P_i = -P_j$.

(iii) For all $e \in [m]$ there is $i_e \in I_e$ such that $P_{i_e} = P_{j_f}$ or $P_{i_e} = -P_{i_f}$ for all $e \neq f$.

Proof. For each of the three cases, we will construct another m-domain graph $\mathcal{G}_m = (\mathcal{H} \cup V, D)$ such that for suitable representations $(\widetilde{B}, \widetilde{P})$ of distributions in $\mathcal{M}(\widetilde{\mathcal{G}}_m)$, the left-hand side of (9) holds, but the right-hand side is violated.

To prove the statement for case (i), let $i, j \in \mathcal{L}$ and assume that $P_i = P_j$. We define the m-domain graph $\widetilde{\mathcal{G}}_m = (\widetilde{\mathcal{H}} \cup V, \widetilde{D})$ to be the almost same graph as $\mathcal{G}_m = (\mathcal{H} \cup V, D)$, we only "swap" the roles of the latent nodes i and j on an arbitrary domain $e \in [m]$. That is, for each $v \in V_e$, if there was an edge $i \to v$ in D, we remove that edge from $\widetilde{D}$ and add the edge $j \to v$ instead, and vice versa. Otherwise, the graph $\widetilde{\mathcal{G}}_m$ has the same structure as $\mathcal{G}_m$. Now, let $\widetilde{P} = P$ and define a the matrix $\widetilde{B}$ to be the same matrix as B, except for the subcolumns $\widetilde{B}_{V_e,i} := B_{V_e,j}$ and $\widetilde{B}_{V_e,j} := B_{V_e,i}$, that is, we swapped $B_{V_e,i}$ and $B_{V_e,j}$. Then the pair $(\widetilde{B}, \widetilde{P})$ is a representation of some distribution in $\mathcal{M}(\widetilde{\mathcal{G}}_m)$. Recall from the proof of Theorem 3.1 that Condition (C2) implies that the matrix B_{V_e,S_e} has full column rank. Since we only swapped columns in $\widetilde{B}_{V_e,\widetilde{S}_e}$, it still has full column rank. Moreover, observe that the left hand side of (9) is satisfied since $P_i = P_j$, that is, the marginal distributions on the single domains coincide.

However, now consider another domain $f \in [m]$ and the submatrices

$$B_{V_e \cup V_f, \{i,j\}} = \begin{pmatrix} B_{V_e,i} & B_{V_e,j} \\ B_{V_f,i} & B_{V_f,j} \end{pmatrix} \quad \text{and} \quad \widetilde{B}_{V_e \cup V_f, \{i,j\}} = \begin{pmatrix} B_{V_e,j} & B_{V_e,i} \\ B_{V_f,i} & B_{V_f,j} \end{pmatrix}.$$

Since all of the four subcolumns are nonzero and neither $B_{V_e,j}$ is equal to $B_{V_e,i}$ nor $B_{V_f,j}$ is equal to $B_{V_f,i}$, there is no signed permutation matrix Ω such that $B_{V_e \cup V_f, \{i,j\}} = \widetilde{B}_{V_e \cup V_f, \{i,j\}}\Omega$. Hence, there is also no larger signed permutation matrix Ψ such that $B = \widetilde{B}\Psi$. We conclude that the right-hand side of (9) is violated and the joint distribution is not identifiable. Finally, note that the above arguments also hold if $P_i = -P_j$ by adding "$-$" signs in appropriate places.

The proof for case (ii) works with exactly the same construction. That is, for $i \in \mathcal{L}$ and $j \in I_e$ we swap the roles of i and j on the domain e. Then, for any other domain $f \in [m]$, we obtain the submatrices

$$B_{V_e \cup V_f, \{i,j\}} = \begin{pmatrix} B_{V_e,i} & B_{V_e,j} \\ B_{V_f,i} & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{B}_{V_e \cup V_f, \{i,j\}} = \begin{pmatrix} B_{V_e,j} & B_{V_e,i} \\ B_{V_f,i} & 0 \end{pmatrix}.$$

By the same arguments as before, this shows that there is no signed permutation matrix Ψ such that $B = \widetilde{B}\Psi$ and, hence, the joint distribution is not identifiable.

To prove case (iii), we consider a slightly different construction. However, we also assume that $P_{i_e} = P_{i_f}$ for all $e \neq f$, since for $P_{i_e} = -P_{i_f}$ we only have to add some "$-$" signs in the following. We define the m-domain graph $\widetilde{\mathcal{G}}_m = (\widetilde{\mathcal{H}} \cup V, \widetilde{D})$ by identifying the nodes $i_e, e \in [m]$ with a new

node k. That is, $\widetilde{\mathcal{L}} = \mathcal{L} \cup \{k\}$ and $\widetilde{\mathcal{H}} = (\bigcup_{e\in[m]} I_e \setminus \{i_e\}) \cup \widetilde{\mathcal{L}}$. For $i \in \widetilde{\mathcal{H}} \setminus \{k\}$ and $v \in V$, the edge set $\widetilde{D}$ contains an edge $i \to v$ if and only if the edge $i \to v$ is in D. For the node $k \in \widetilde{\mathcal{H}}$ and $v \in V$, we put an edge $k \to v$ in $\widetilde{D}$ if and only if there is an edge $i_e \to v$ in D for some $e \in [m]$.

Now, define the matrix $\widetilde{B}$ such that $\widetilde{B}_{V,\widetilde{\mathcal{H}}\setminus\{k\}} := B_{V,\widetilde{\mathcal{H}}\setminus\{k\}}$ and $\widetilde{B}_{V_e,k} := B_{V_e,i_e}$ for all $e \in [m]$. Then the pair $(\widetilde{B}, \widetilde{P})$ is a representation of some distribution in $\mathcal{M}(\widetilde{\mathcal{G}}_m)$. Moreover, each submatrix $\widetilde{B}_{V_e,\widetilde{S}_e}$ is equal to B_{V_e,S_e} up to relabeling of the columns. That is, the column that is labeled by i_e in B_{V_e,S_e} is now labeled by k in $\widetilde{B}_{V_e,\widetilde{S}_e}$. We define the measure $\widetilde{P}$ such that $\widetilde{P}_{\widetilde{\mathcal{H}}\setminus\{k\}} = P_{\widetilde{\mathcal{H}}\setminus\{k\}}$ and $\widetilde{P}_k = P_{i_e}$ for all $e \in [m]$. Then the pair $(\widetilde{B}, \widetilde{P})$ is a representation of some distribution in $\mathcal{M}(\widetilde{\mathcal{G}}_m)$ and, in particular, the left hand side of (9) is satisfied. That is, the marginal distributions coincide on each domain. However, the number of shared latent variables in both m-domain graphs is different since we have $\widetilde{\ell} = \ell + 1$. We conclude that the joint distribution is not identifiable. □

The proposition states that it is in most cases necessary that error distributions are pairwise different. However, in two cases the same error distributions still lead to identifiability. First, if $i, j \in I_e$, then the corresponding error distributions may be the same and the joint distribution is still identifiable. Similarly, if there are latent nodes i_e in a few domains $e \in [m]$ such that the corresponding error distributions coincide, but there is at least one domain $f \in [m]$ where there is no latent node with the same error distribution, then the joint distribution is also identifiable. Both can be seen by taking the the proofs of Theorem 3.1 and Proposition C.1 together.

Gaussian Errors. Without additional assumptions to those in Section 3, it is impossible to recover the joint distribution if the distributions of the errors ε_i of the latent structural equation model in Equation (1) are Gaussian. In this case, the distribution of Z as well as the distribution of each observed random vector X^e is determined by the covariance matrix only. The observed covariance matrix in domain $e \in [m]$ is given by $\Sigma^e = G^e \mathrm{Cov}[Z_{\mathcal{L}\cup I_e}](G^e)^\top$. However, knowing Σ^e gives no information about $Z_{\mathcal{L}\cup I_e}$ other than $\mathrm{rank}(\Sigma^e) = |\mathcal{L}| + |I_e|$, that is, we cannot distinguish which latent variables are shared and which ones are domain-specific. This is formalized in the following lemma.

Lemma C.2. *Let Σ be any $d \times d$ symmetric positive semidefinite matrix of rank p and let Ξ be another arbitrary $p \times p$ symmetric positive definite matrix. Then there is $G \in \mathbb{R}^{d\times p}$ such that $\Sigma = G\Xi G^\top$.*

Proof. Let Σ be a $d\times d$ symmetric positive semidefinite matrix of rank p. Then, Σ has a decomposition similar to the Cholesky decomposition; see e.g. Gentle (1998, Section 3.2.2). That is, there exists a unique matrix T, such that $A = TT^\top$, where T is a lower triangular matrix with p positive diagonal elements and $d - p$ columns containing all zeros. Define $\widetilde{T}$ to be the $d \times p$ matrix containing only the non-zero columns of T.

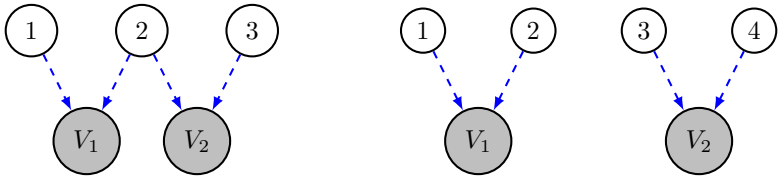
On the other hand, let Ξ be a symmetric positive definite $p \times p$ matrix. By the usual Cholesky decomposition (Lyche, 2020, Section 4.2.1), there exists a unique $p \times p$ lower triangular matrix L with positive diagonal elements such that $\Xi = LL^\top$. Now, define $G := \widetilde{T}L^{-1} \in \mathbb{R}^{d\times p}$. Then,

$$\Sigma = \widetilde{T}\widetilde{T}^\top = \widetilde{T}L^{-1}LL^\top L^{-\top}\widetilde{T}^\top = G\Xi G^\top. \qquad \square$$

Due to Lemma C.2 it is necessary to consider non-Gaussian distributions to obtain identifiability of the joint distribution.

Example C.3. In the Gaussian case we cannot distinguish whether the two observed domains in Figure 5 share a latent variable or not. Said differently, the observed marginal distributions may either be generated by the mechanism defined by graph (a) or graph (b) and there is no way to distinguish from observational distributions only.

Sparsity Assumptions. Let $B \in \mathrm{Im}(\phi_{\mathcal{G}_m})$ for an m-domain graph $\mathcal{G}_m = (\mathcal{H} \cup V, D)$ and suppose that we are given the matrix $B_{\mathcal{L}} = G_{V,\mathcal{L}}(I - A_{\mathcal{L},\mathcal{L}})^{-1}$, that is, we are given the submatrix with columns indexed by the shared latent nodes. Now, assume that the graph does not impose any sparsity restrictions on $G_{V,\mathcal{L}}$, which means that the set $\mathbb{R}^{D_{V\mathcal{L}}}$ of possible matrices $G_{V,\mathcal{L}}$ is equal to $\mathbb{R}^{|V|\times|\mathcal{L}|}$. Then, the set of possible matrices $B_{\mathcal{L}}$ is also unrestricted, that is, $B_{\mathcal{L}}$ can be any matrix in $\mathbb{R}^{|V|\times|\mathcal{L}|}$ no matter the form of the matrix $A_{\mathcal{L},\mathcal{L}} \in \mathbb{R}^{D_{\mathcal{L}\mathcal{L}}}$. In other words, for arbitrary shared latent graphs



(a) (b)

Figure 5: Compact versions of two 2-domain graphs. In both graphs, both domains have two latent parents. In setup (a) there is a shared latent parent while in setup (b) there is not.

$\mathcal{G}_{\mathcal{L}} = (\mathcal{L}, D_{\mathcal{L}\mathcal{L}})$ and arbitrary corresponding parameter matrices $A_{\mathcal{L},\mathcal{L}} \in \mathbb{R}^{D_{\mathcal{L}\mathcal{L}}}$, we don't get any restrictions on the matrix $B_{\mathcal{L}}$. Therefore, it is impossible to infer $A_{\mathcal{L},\mathcal{L}}$ from $B_{\mathcal{L}}$.

Condition (C3) requires that there are two partial pure children for every shared latent node $k \in \mathcal{L}$, which implies that there are $2|\mathcal{L}|$ rows in $G_{V,\mathcal{L}}$ in which only one entry may be nonzero. While we show in Theorem 4.4 that this condition is sufficient for identifiability of $A_{\mathcal{L},\mathcal{L}}$, we leave it open for future work to find a necessary condition.

D Algorithms for Finite Samples

We adjust Algorithm 1 such that it is applicable in the empirical data setting. That is, rather than the exact distribution P_{X^e}, we have a matrix of observations $\boldsymbol{X}^e \in \mathbb{R}^{d_e \times n_e}$ in each domain $e \in [m]$. The sample size n_e might be different across domains. We denote $n_{\min} = \min_{e\in[m]} n_e$ and $n_{\max} = \max_{e\in[m]} n_e$. For implementing linear ICA on finite samples, multiple well developed algorithms are available, e.g., FastICA (Hyvärinen, 1999; Hyvärinen and Oja, 2000), Kernel ICA (Bach and Jordan, 2003) or JADE (Cardoso and Souloumiac, 1993). Applying them, we obtain a measure $\widehat{P}_i^e$ which is an estimator of the true measure P_i^e in Algorithm 1, Line 4.

The remaining challenge is the matching in Line 6 of Algorithm 1. For finite samples, the distance between empirical distributions is almost surely not zero although the true underlying distributions might be equal. In this section, we provide a matching strategy based on the two-sample Kolmogorov-Smirnov test (van der Vaart and Wellner, 1996, Section 3.7). We match two distributions if they are not significantly different. During this process, there might occur false discoveries, that is, distributions are matched that are actually not the same. We show that the probability of falsely discovering shared nodes shrinks exponentially with the number of domains.

For two univariate Borel probability measures P_i, P_j, with corresponding cumulative distribution functions F_i, F_j, the Kolmogorov-Smirnov distance is given by the L^∞-distance

$$d_{\text{KS}}(P_i, P_j) = \|F_i - F_j\|_\infty = \sup_{x\in\mathbb{R}} |F_i(x) - F_j(x)|.$$

The two-sample Kolmogorov-Smirnov test statistic for the null hypothesis $H_0 : d_{\text{KS}}(P_i^e, P_j^f) = 0$ is given by

$$T(\widehat{P}_i^e, \widehat{P}_j^f) = \sqrt{\frac{n_e n_f}{n_e + n_f}} d_{\text{KS}}(\widehat{P}_i^e, \widehat{P}_j^f). \tag{10}$$

It is important to note that $\widehat{P}_i^e$ is not an empirical measure in the classical sense since it is not obtained from data sampled directly from the true distribution P_i^e. In addition to the sampling error there is the uncertainty of the ICA algorithm. However, in the analysis we present here, we will neglect this error and treat $\widehat{P}_i^e$ as an empirical measure. In this case, under H_0, the test statistic $T(\widehat{P}_i^e, \widehat{P}_j^f)$ converges in distribution to $\|B\|_\infty$, where B is a Brownian bridge from 0 to 1 (van der Vaart and Wellner, 1996, Section 2.1). For a given level $\alpha \in (0, 1)$, we choose the critical value as $c_\alpha = \inf\{t : P(\|B\|_\infty > t) \leq \alpha\}$ and reject H_0 if $T(\widehat{P}_i^e, \widehat{P}_j^f) > c_\alpha$.

Definition D.1. Let $\alpha \in (0, 1)$ and suppose the distributions $\{\widehat{P}_1^e, \ldots, \widehat{P}_{\widehat{s}_e}^e\}$ and $\{\widehat{P}_1^f, \ldots, \widehat{P}_{\widehat{s}_f}^f\}$ are given for two domains $e, f \in [m]$. Define

$$\Omega_\alpha(\widehat{P}_i^e, \widehat{P}_j^f) = \begin{cases} 1 & \text{if } T_{ij}^{ef} \leq c_\alpha \text{ and } T_{ij}^{ef} = \min\{\min_{k\in[\widehat{s}_f]} T_{ik}^{ef}, \min_{k\in[\widehat{s}_e]} T_{kj}^{ef}\}, \\ 0 & \text{else,} \end{cases}$$

where $T_{ij}^{ef} = \min\{T(\widehat{P}_i^e, \widehat{P}_j^f), T(\widehat{P}_i^e, -\widehat{P}_j^f)\}$. We say that $\widehat{P}_i^e$, $\widehat{P}_j^f$ are *matched* if $\Omega_\alpha(\widehat{P}_i^e, \widehat{P}_j^f) = 1$.

Definition D.1 essentially states that two measures are matched if the test statistic (10) is not significantly large and the null hypothesis cannot be rejected. Taking the minimum of $T(\widehat{P}_i^e, \widehat{P}_j^f)$ and $T(\widehat{P}_i^e, -\widehat{P}_j^f)$ accounts for the sign indeterminacy of linear ICA. For two fixed domains $e, f \in [m]$, if it happens that the statistic T_{ij}^{ef} for multiple pairs (i, j) is small enough, then the pair with the minimal value of the statistic is matched. Note that one may use any other test than the Kolmogorov-Smirnov test to define a matching as in Definition D.1. We discover a shared latent node if it is matched consistently across domains.

Definition D.2. Let $C = (i_1, \ldots, i_m)$ be a tuple with m elements such that $i_e \in [\widehat{s}_e]$. Then we say that C determines a *shared node* if $\Omega_\alpha(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) = 1$ for all $i_e, i_f \in C$.

Inferring the existence of a shared node which does not actually exist may be considered a more serious error than inferring a shared node determined by a set C, where only some components of C are wrongly matched. In the following theorem we show that the probability of falsely discovering shared nodes shrinks exponentially with the number of wrongly matched components.

Theorem D.3. *Let $C = (i_1, \ldots, i_m)$ be a tuple with m elements such that $i_e \in [\widehat{s}_e]$. Let $g : \mathbb{N} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be a function that is monotonically decreasing in $n \in \mathbb{N}$ and assume the following:*

(i) $P(d_{KS}(\widehat{P}_{i_e}^e, P_{i_e}^e) > x) \leq g(n_e, x)$ for all $e \in [m]$ and for all $x \geq 0$.

(ii) There is $E \subseteq [m]$ with $|E| \geq 2$ and a constant $\kappa > 0$ such that $d_{KS}(P_{i_e}^e, P_{i_f}^f) \geq \kappa$ and $d_{KS}(P_{i_e}^e, -P_{i_f}^f) \geq \kappa$ for all $e \in E$, $f \in [m]$ with $e \neq f$.

Then

$$P(C \text{ determines a shared node}) \leq g\left(n_{\min},\ \max\left\{\frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha, 0\right\}\right)^{|E|-1}.$$

Proof of Theorem D.3. Let $C = (i_1, \ldots, i_m)$ and g be as in the statement of the theorem. W.l.o.g. we assume that $E = \{1, \ldots, |E|\}$ and that $T(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) \leq T(\widehat{P}_{i_e}^e, -\widehat{P}_{i_f}^f)$. Observe that C determines a shared node if and only if

$$\sum_{e<f} \Omega_\alpha(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) = \binom{m}{2}.$$

Now, we have

$$\begin{aligned}
P\left(\sum_{e<f} \Omega_\alpha(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) = \binom{m}{2}\right) &= P\left(\bigcap_{e<f} \left\{\Omega_\alpha(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) = 1\right\}\right) \\
&\leq P\left(\bigcap_{e<f} \left\{T_{ij}^{ef} \leq c_\alpha\right\}\right) \\
&= P\left(\bigcap_{e<f} \left\{T(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) \leq c_\alpha\right\} \cup \left\{T(\widehat{P}_{i_e}^e, -\widehat{P}_{i_f}^f) \leq c_\alpha\right\}\right) \\
&= P\left(\bigcap_{e<f} \left\{T(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) \leq c_\alpha\right\}\right) \\
&\leq P\left(\bigcap_{\substack{e \in E, f \in [m] \\ e<f}} \left\{T(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) \leq c_\alpha\right\}\right). \qquad (11)
\end{aligned}$$

By the triangle inequality we have

$$d_{\text{KS}}(P_{i_e}^e, P_{i_f}^f) \leq d_{\text{KS}}(P_{i_e}^e, \widehat{P}_{i_e}^e) + d_{\text{KS}}(\widehat{P}_{i_e}^e, \widehat{P}_{i_f}^f) + d_{\text{KS}}(\widehat{P}_{i_f}^f, P_{i_f}^f). \qquad (12)$$

Moreover, if $e \in E$ and $f \in [m]$, then we have by Condition (ii)

$$d_{\mathrm{KS}}(P^e_{i_e}, P^f_{i_f}) \geq \kappa. \tag{13}$$

Using (12) and (13) together with the fact $\sqrt{\frac{n_e n_f}{n_e+n_f}} \geq \frac{n_{\min}}{\sqrt{2n_{\max}}}$, we obtain the following chain of implications for $e \in E$ and $f \in [m]$:

$$\begin{aligned}
& T(\widehat{P}^e_{i_e}, \widehat{P}^f_{i_f}) \leq c_\alpha \\
\iff & \sqrt{\frac{n_e n_f}{n_e+n_f}} d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, \widehat{P}^f_{i_f}) \leq c_\alpha \\
\implies & \frac{n_{\min}}{\sqrt{2n_{\max}}} \Big(d_{\mathrm{KS}}(P^e_{i_e}, P^f_{i_f}) - d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) - d_{\mathrm{KS}}(\widehat{P}^f_{i_f}, P^f_{i_f}) \Big) \leq c_\alpha \\
\implies & \frac{n_{\min}}{\sqrt{2n_{\max}}} \Big(\kappa - d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) - d_{\mathrm{KS}}(\widehat{P}^f_{i_f}, P^f_{i_f}) \Big) \leq c_\alpha \\
\iff & d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) + d_{\mathrm{KS}}(\widehat{P}^f_{i_f}, P^f_{i_f}) \geq \kappa - \frac{\sqrt{2n_{\max}}}{n_{\min}} c_\alpha \\
\implies & \left\{ d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) \geq \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha \right\} \text{ or } \left\{ d_{\mathrm{KS}}(\widehat{P}^f_{i_f}, P^f_{i_f}) \geq \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha \right\}.
\end{aligned}$$

Now, consider the event $\bigcap_{\substack{e \in E, f \in [m] \\ e<f}} \{T(\widehat{P}^e_{i_e}, \widehat{P}^f_{i_f}) \leq c_\alpha\}$. On this event, there cannot be two elements $e, f \in E$ such that both

$$\left\{ d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) < \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha \right\} \text{ and } \left\{ d_{\mathrm{KS}}(\widehat{P}^f_{i_f}, P^f_{i_f}) < \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha \right\}.$$

To see this recall that $E \subseteq [m]$. We conclude that it must hold $\{d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) \geq \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha\}$ for all but at most one element of E. We denote this exceptional element by $e^* \in E$. Taking up (11), we get the following:

$$\begin{aligned}
& P\left(\sum_{e<f} \Omega_\alpha(\widehat{P}^e_{i_e}, \widehat{P}^f_{i_f}) = \binom{m}{2} \right) \\
& \leq P\left(\bigcap_{\substack{e \in E, f \in [m] \\ e<f}} \left\{ T(\widehat{P}^e_{i_e}, \widehat{P}^f_{i_f}) \leq c_\alpha \right\} \right) \\
& \leq P\left(\bigcap_{e \in E \setminus \{e^*\}} \left\{ d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) \geq \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha \right\} \right) \\
& = \prod_{e \in E \setminus \{e^*\}} P\left(d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) \geq \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha \right) && (14) \\
& = \prod_{e \in E \setminus \{e^*\}} P\left(d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) \geq \max\left\{ \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha,\ 0 \right\} \right) && (15) \\
& \leq g\left(n_{\min},\ \max\left\{ \frac{\kappa}{2} - \frac{\sqrt{n_{\max}}}{\sqrt{2}\, n_{\min}} c_\alpha, 0 \right\} \right)^{|E|-1}. && (16)
\end{aligned}$$

The last three steps need more explanation: Equality (14) follows from the fact that domains are unpaired. That is, the distances $d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e})$ and $d_{\mathrm{KS}}(\widehat{P}^f_{i_f}, P^f_{i_f})$ are pairwise independent for different domains $e, f \in [m]$. Equality (15) is trivial since $d_{\mathrm{KS}}(\widehat{P}^e_{i_e}, P^e_{i_e}) \geq 0$. Finally, Inequality (16) follows from Condition (i) and that the function g is monotonically decreasing in n. We also used the fact that $|E \setminus \{e^*\}| = |E| - 1$. □

If $\widehat{P}^e_i$ were an empirical measure in the classical sense, then Condition (i) in Theorem D.3 translates to the well-known Dvoretzky–Kiefer–Wolfowitz inequality, that is, the function g is given by $g(n, x) =$

Algorithm 3 IdentifyJointDistributionEmpirical

Hyperparameters: $\gamma > 0$, $\alpha \in (0, 1)$.
Input: Matrix of observations $\boldsymbol{X}^e \in \mathbb{R}^{d_e \times n_e}$ for all $e \in [m]$.
Output: Number of shared latent variables $\widehat{\ell}$, matrix $\widehat{B}$ and probability measure $\widehat{P}$.
for $e \in [m]$ **do**
  Linear ICA: Use any linear ICA algorithm to obtain a mixing matrix $\widetilde{B}^e \in \mathbb{R}^{d_e \times \widehat{s}_e}$, where $\widehat{s}_e = \text{rank}_\gamma(\boldsymbol{X}^e(\boldsymbol{X}^e)^\top)$. Compute the matrix $\tilde{\boldsymbol{\eta}}^e = (\widetilde{B}^e)^\dagger \boldsymbol{X}^e \in \mathbb{R}^{\widehat{s}_e \times n_e}$, where $(\widetilde{B}^e)^\dagger$ is the Moore-Penrose pseudoinverse of $\widetilde{B}^e$.
  Scaling: Let Δ^e be a $\widehat{s}_e \times \widehat{s}_e$ diagonal matrix with entries $\Delta^e_{ii} = \frac{1}{n_e}[\tilde{\boldsymbol{\eta}}^e(\tilde{\boldsymbol{\eta}}^e)^\top]_{ii}$. Define $\widehat{B}^e = \widetilde{B}^e(\Delta^e)^{-1/2}$ and $\boldsymbol{\eta}^e = (\Delta^e)^{-1/2}\tilde{\boldsymbol{\eta}}^e$.
  Let $\widehat{P}^e$ be the estimated probability measure with independent marginals such that $\widehat{P}^e_i$ is the empirical measure of the row $\boldsymbol{\eta}^e_{i,*}$.
end for
Matching: Let $\widehat{\ell}$ be the maximal number such that there is a signed permutation matrix Q^e in each domain $e \in [m]$ such that

$$\Omega_{\alpha_t}([(Q^e)^\top \# \widehat{P}^e]_i, [(Q^f)^\top \# \widehat{P}^f]_i) = 1$$

for all $i = 1, \ldots, \widehat{\ell}$ and all $f \neq e$, where $\alpha_t = \alpha/t$ with $t = 2\sum_{e<f} \widehat{s}_e \widehat{s}_f$. Let $\widehat{\mathcal{L}} = \{1, \ldots, \widehat{\ell}\}$.

Construct the matrix $\widehat{B}$ and the tuple of probability measures $\widehat{P}$ given by

$$\widehat{B} = \left(\begin{array}{c|ccc} [\widehat{B}^1 Q^1]_{\widehat{\mathcal{L}}} & [\widehat{B}^1 Q^1]_{[\widehat{s}_1]\setminus\widehat{\mathcal{L}}} & & \\ \vdots & & \ddots & \\ [\widehat{B}^m Q^m]_{\widehat{\mathcal{L}}} & & & [\widehat{B}^m Q^m]_{[\widehat{s}_m]\setminus\widehat{\mathcal{L}}} \end{array}\right) \text{ and } \widehat{P} = \begin{pmatrix} [(Q^1)^\top \# \widehat{P}^1]_{\widehat{\mathcal{L}}} \\ [(Q^1)^\top \# \widehat{P}^1]_{[\widehat{s}_1]\setminus\widehat{\mathcal{L}}} \\ \vdots \\ [(Q^m)^\top \# \widehat{P}^m]_{[\widehat{s}_m]\setminus\widehat{\mathcal{L}}} \end{pmatrix}.$$

return $(\widehat{\ell}, \widehat{B}, \widehat{P})$.

Algorithm 4 IdentifySharedGraphEmpirical

Hyperparameters: $\gamma > 0$.
Input: Matrix $B^* \in \mathbb{R}^{|V| \times \ell}$.
Output: Parameter matrix $\widehat{A} \in \mathbb{R}^{\ell \times \ell}$.
Remove rows $B^*_{i,\mathcal{L}}$ with $\|B^*_{i,\mathcal{L}}\|_2 \leq \gamma$ from the matrix B^*.
Find tuples $(i_k, j_k)_{k \in \mathcal{L}}$ with the smallest possible scores $\sigma_{\min}(B^*_{\{i_k, j_k\}, \mathcal{L}})$ such that
  (i) $i_k \neq j_k$ for all $k \in \mathcal{L}$ and $\{i_k, j_k\} \cap \{i_q, j_q\} = \emptyset$ for all $k, q \in \mathcal{L}$ such that $k \neq q$ and
  (ii) $\sigma_{\min}(B^*_{\{i_k, i_q\}, \mathcal{L}})| > \gamma$ for all $k, q \in \mathcal{L}$ such that $k \neq q$.
Let $I = \{i_1, \ldots, i_\ell\}$ and consider the matrix $B^*_{I,\mathcal{L}} \in \mathbb{R}^{\ell \times \ell}$.
Find two permutation matrices R_1 and R_2 such that $W = R_1 B^*_{I,\mathcal{L}} R_2$ is as close as possible to lower triangular. This can be measured, for example, by using $\sum_{i<j} W^2_{ij}$.
Multiply each column of W by the sign of its corresponding diagonal element. This yields a new matrix $\widetilde{W}$ with all diagonal elements positive.
Divide each row of $\widetilde{W}$ by its corresponding diagonal element. This yields a new matrix $\widetilde{W}'$ with all diagonal elements equal to one.
Compute $\widehat{A} = I - (\widetilde{W}')^{-1}$.
return $\widehat{A}$.

$2\exp(-2nx^2)$. Given a tuple $C = (i_1, \ldots, i_m)$ that defines a shared node, Condition (ii) is an assumption on the number of wrongly matched components. The most extreme case is when the shared node does not actually exist and all components are wrongly matched. That is, the measures $\widehat{P}^e_{i_e}$ and $\widehat{P}^f_{i_f}$ are matched even though $d_{\text{KS}}(P^e_{i_e}, P^f_{i_f}) \neq 0$ and $d_{\text{KS}}(P^e_{i_e}, -P^f_{i_f}) \neq 0$ *for all* $e, f \in [m]$. On the other hand, if $|E| \ll m$, then C determines a shared node where the majority of the components are correctly matched.

If $g(n, x) \to 0$ for $n \to \infty$ and $x > 0$, the statement of the theorem becomes meaningful under the constraint $\sqrt{n_{\max}}/n_{\min} \to 0$. In this case, the probability that a given tuple C with wrong components E determines a shared node goes to zero for large sample sizes $n_{\min}$. As noted, the probability of falsely discovering a shared node decreases exponentially with the number of wrongly matched components $|E|$. In the extreme case, this means that the probability of falsely discovering shared nodes with all components wrongly matched, i.e., $E = [m]$, decreases exponentially with the number of domains m.

Theorem D.3 also tells us that the probability of falsely matching two measures $\widehat{P}^e_i$ and $\widehat{P}^f_j$ becomes zero if the sample size grows to infinity and the linear ICA algorithm is consistent. However, with finite samples we might fail to match two measures where the underlying true measures are actually the same, i.e., we falsely reject the true null hypothesis H_0. Thus, we might be overly conservative in detecting shared nodes due to a high family-wise error rate caused by multiple testing. We suggest to correct the level α to account for the amount of tests carried out. One possibility is to apply a Bonferroni-type correction. The total number of tests is given by $t = 2\sum_{e<f} \widehat{s}_e \widehat{s}_f$. This means that an adjusted level is given by $\alpha_t = \alpha/t$ and instead of the critical value c_α we consider the adjusted critical value c_{α_t}.

Algorithm 3 is the finite sample version of Algorithm 1 with the matching Ω_α defined in Definition D.1. To determine the number of independent components for the linear ICA step in each domain, we need to check the $\text{rank}(\boldsymbol{X}^e(\boldsymbol{X}^e)^\top)$. We specify the rank of a matrix M as number of singular values which are larger than a certain threshold γ and denote it by $\text{rank}_\gamma(M)$.

In Algorithm 4 we also provide a finite sample version of Algorithm 2 where we only have the approximation $B^\star = \widehat{B}_\mathcal{L} \approx B_\mathcal{L}\Psi_\mathcal{L}$ for a signed permutation matrix $\Psi_\mathcal{L}$. For a matrix M, we denote by $\sigma_{\min}(M)$ the smallest singular value.

E Error Distributions in Simulations

We specify $\mathcal{L} = \{1, 2, 3\}$, $I_1 = \{4, 5\}$, $I_2 = \{6, 7\}$ and $I_3 = \{8, 9\}$. Note that the set I_3 does not exist if the number of domains is $m = 2$. The error distributions in all simulations are specified as follows if not stated otherwise.

$$
\begin{aligned}
\mathcal{L}&: \ \varepsilon_1 \sim \overline{\text{Beta}}(2,3),\ \varepsilon_2 \sim \overline{\text{Beta}}(2,5),\ \varepsilon_3 \sim \overline{\chi}^2_4,\\
I_1&: \ \varepsilon_4 \sim \overline{\text{Gumbel}}(0,1),\ \varepsilon_5 \sim \overline{\text{LogNormal}}(0,1),\\
I_2&: \ \varepsilon_6 \sim \overline{\text{Weibull}}(1,2),\ \varepsilon_7 \sim \overline{\text{Exp}}(0.1),\\
I_3&: \ \varepsilon_8 \sim \overline{\text{SkewNormal}}(6),\ \varepsilon_9 \sim \overline{\text{SkewNormal}}(12),
\end{aligned}
$$

where the overline means that each distribution is standardized to have mean 0 and variance 1. Figure 6 shows histograms of the empirical distributions.

F Additional Simulation Results

In this section, we make additional experiments. First, we consider another setup where all our assumptions are satisfied but we have more domains and more shared latent variables. Then, we also consider two setups where some of our assumptions are not satisfied.

Different Setup. We make additional experiments on a similar scale as in Section 5, but with more shared nodes and less domain specific nodes. This time, we consider $\ell = 5$ shared latent nodes and $|I_e| = 1$ domain-specific latent node in each domain. Moreover, we also consider $m = 4$ domains. The dimensions are given by $d_e = d/m$ for all $e \in [m]$ and $d = 48$. The graphs and edge weights

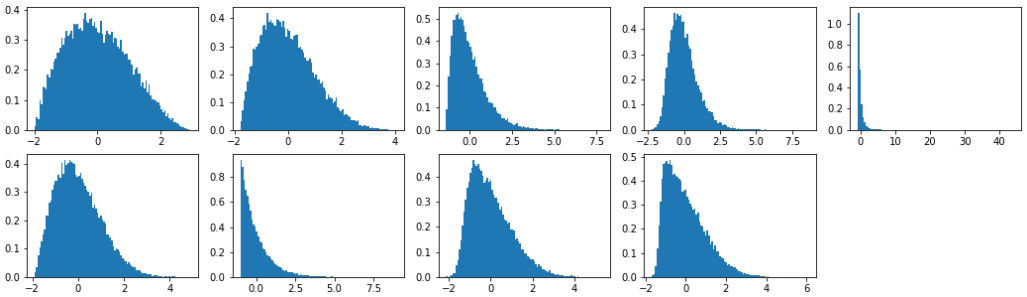
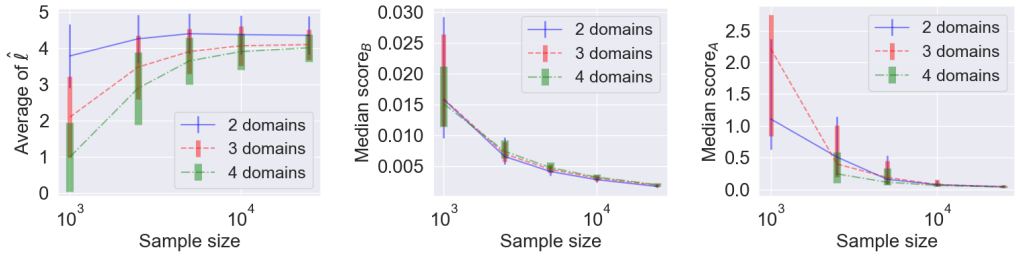


Figure 6: Histograms showing the frequency of 25000 values sampled from the random variables ε_i with distribution as specified in Appendix E. Each distribution has mean zero and variance one. The first row shows the empirical distributions from ε_1 to ε_5 and the second row from ε_6 to ε_9.



(a) (b) (c)

Figure 7: **Simulation results for $\ell = 5$ shared latent nodes.** Logarithmic scale on the x-axis. Error bars in (a) are one standard deviation of the mean and in (b) and (c) they are the interquartile range.

are sampled equivalently as in Section 5 in the main paper. We also consider the same distribution of the error variables as specified in Appendix E, where we specify $\mathcal{L} = \{1, 2, 3, 4, 5\}$, $I_1 = \{6\}$, $I_2 = \{7\}$, $I_3 = \{8\}$ and $I_4 = \{9\}$.

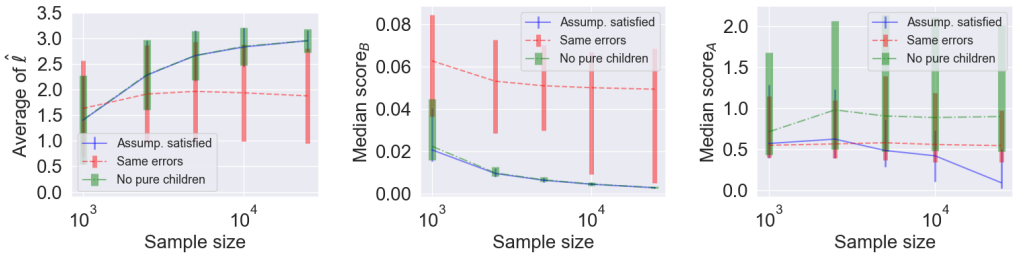
Figure 7 shows the results where the scores are equivalent as in the main paper. Once again, we see that the estimation error for the matrices $B_{\mathcal{L}}$ and $A_{\mathcal{L},\mathcal{L}}$ decreases with increasing sample size. This supports our proof of concept and shows that the adapted algorithms are consistent for recovering $B_{\mathcal{L}}$ and $A_{\mathcal{L},\mathcal{L}}$ from finite samples.

Violated Assumptions. We consider the same setup as in the main paper in Section 5 with $l = 3$ shared latent nodes, but we fix the number of domains to $m = 3$. In this experiment, we compare the results where data was generated such that all our assumptions are satisfied with two setups where we violate some of the assumptions. In the first setup, we violate Condition (C1) that requires pairwise different error distributions. We specify the error distributions as follows.

$$
\begin{aligned}
\mathcal{L}&: \ \varepsilon_1 \sim \overline{\text{Beta}}(2,3),\ \varepsilon_2 \sim \overline{\text{Beta}}(2,5),\ \varepsilon_3 \sim \overline{\chi}^2_4,\\
I_1&: \ \varepsilon_4 \sim \overline{\text{Beta}}(2,3),\ \varepsilon_5 \sim \overline{\text{LogNormal}}(0,1),\\
I_2&: \ \varepsilon_6 \sim \overline{\text{Beta}}(2,5),\ \varepsilon_7 \sim \overline{\text{LogNormal}}(0,1),\\
I_3&: \ \varepsilon_8 \sim \overline{\chi}^2_4,\ \varepsilon_9 \sim \overline{\text{LogNormal}}(0,1),
\end{aligned}
$$

where, as before, the overline means that each distribution is standardized to have mean 0 and variance 1. In the second setup, we do not change the error distributions but we violate Condition (C3) that requires two partial pure children per shared latent node. In this experiment, we do not make any sparsity assumptions on the mixing matrix $G_{V,\mathcal{L}}$.

Figure 8 shows the results of our experiments. As expected, we see in (a) and (b) that identifyability of the joint distribution fails if we do not require pairwise different error distributions. Recovering the joint distribution still works well in the second setup where we violate the partial pure children



(a) (b) (c)

Figure 8: **Simulation results where assumptions are not satisfied.** Logarithmic scale on the x-axis. Error bars in (a) are one standard deviation of the mean and in (b) and (c) they are the interquartile range.

conditions. However, identifying the shared latent graph is impossible in this setup, as we explained in Appendix C. This is supported by the experimental results displayed in Figure 8 (c), where we can see that recovery of the shared latent graph does not work when the partial pure children assumption is not satisfied.

G Future Work

We see many directions for future work, which include the following.

- Our work and algorithms rely on linear ICA. It would be interesting to study a more direct approach to recover the joint distribution and the causal graph. This might potentially be done by testing certain constraints implied by the model similar as the developments in the LiNGAM literature; see e.g. Shimizu et al. (2011) and Wang and Drton (2020).
- Our results require non-Gaussianity and that both the latent structural equation model and the mixing functions are linear. We consider the linear setup as a basis for any subsequent study of nonlinear cases. For example, recent advances in non-linear ICA allow identifiability of up to linear transformations, see e.g. Khemakhem et al. (2020), Buchholz et al. (2022) and Roeder et al. (2021). Identifiability of a causal representation might then be obtained from identifiability results for the linear model.
- Our sufficient condition for identifiability of the shared latent graph requires two partial pure children per shared latent node. In this regard, it would also be interesting to study necessary conditions; c.f. our discussion in Appendix C.
- This work focused purely on the observational case. However, considering interventional data can be expected to permit relaxing some conditions in both of the key steps, i.e., recovering the joint distribution and the shared latent graph. For example, recent work shows that interventional data allow for identification of the latent graph without sparsity constraints in a single-domain setup (Squires et al., 2023; Ahuja et al., 2023). Extending this to the multi-domain setup is an interesting problem for future work.
- It would be interesting to study the statistical properties of our setup such as theoretical bounds on the accuracy of recovering the matrices B and $A_{\mathcal{L},\mathcal{L}}$ as well as developing algorithms that meet these bounds. Currently, our adapted algorithms for finite samples determine the rank of a matrix by using a threshold for singular values. The algorithms depend on the choice of this parameter and it would be worth studying optimal choices. Moreover, one might consider different methods for determining the rank of a matrix.
- There might be different matching strategies of the estimated error distributions in the finite sample setting. For example, instead of matching pairwise consistently across all domains as we propose in Appendix D, one might find an optimal matching by solving a linear program.

Permission to include the article

Who holds the Copyright on a NeurIPS paper

According to U.S. Copyright Office's page, **What is a Copyright**, when you create an original work you are the author and the owner and hold the copyright, unless you have an agreement to transfer the copyright to a third party such as the company or school you work for.

Authors do not transfer the copyright of their papers to NeurIPS. Instead, they grant NeurIPS a non-exclusive, perpetual, royalty-free, fully-paid, fully-assignable license to copy, distribute and publicly display all or part of the paper.

The NeurIPS Logo above may be used on presentations. Right-click and choose download. It is a vector graphic and may be used at any scale.

USEFUL LINKS

- Press
- Exhibitor Information

CONTACT

1269 Law St, San Diego CA 92109

Email

NeurIPS Proceedings

B Further Articles

B.1 Half-Trek Criterion for Identifiability of Latent Variable Models

Summary

In this article, we consider linear structural equation models with latent variables and develop a criterion to certify whether the direct causal effects between the observable variables are identifiable based on the observed covariance matrix. As discussed in Chapter 5 of this thesis, prior research has developed a variety of methods to decide identifiability of direct effects in a latent projection framework, in which the confounding effects of the latent variables are represented by correlation among noise terms. This approach is effective when the confounding is sparse and effects only small subsets of the observed variables. In contrast, the new latent-factor half-trek criterion (LF-HTC) we develop in this article operates on the original unprojected latent variable model, where the corresponding directed graph explicitly includes latent nodes. The LF-HTC is a graphical criterion that is able to certify identifiability in settings, where some latent variables may also have dense effects on many or even all of the observables. Moreover, the graph is allowed to be cyclic, the only restriction that we make in this work is that all latent variables are source nodes in the graph. The new criterion can be checked in time that is polynomial in the size of the graph if we restrict the search steps in LF-HTC to only consider subsets of latent nodes of bounded size. The restriction of the search space is necessary since checking the criterion without any restriction is in general NP-hard.

The article is structured as follows. In Section 2, we provide a precise definition of linear structural equation models given by directed graphs and rigorously introduce the concept of rational identifiability for unprojected latent variable models. Moreover, we derive basic necessary conditions for rational identifiability based on dimension arguments. In Section 3, we present our main result, the LF-HTC. In Section 4, we discuss the latent projection framework considered in previous research and compare the new LF-HTC to existing criteria. In particular, we compare the LF-HTC to the original half-trek criterion (Foygel et al., 2012) that is applicable in the latent projection framework. In Section 5, we present an algorithm to check the LF-HTC efficiently. Using this algorithm we systematically check identifiability of certain classes of small latent-factor graphs in Section 6. The restriction to small graphs allows for these checks to be validated using suitably designed Gröbner basis computations. Finally, the proof of the main result is given in Section 7. The supplementary material contains further elements of proofs, the hardness result for checking the LF-HTC without a bound on the cardinality of searched sets of latent variables and an explanation on how to effectively deploy techniques from computational algebraic geometry.

Individual contributions

I am a co-author of this article, which has the authors listed in alphabetical order. Rina Foygel Barber proved Theorem 3.7 and introduced the necessary theoretical concepts in Definitions 3.2, 3.4 and 3.6. Luca Weihs proved Theorem B.1 in the Appendix and wrote a preliminary implementation of Algorithm 1. I proved all of the remaining results, and was responsible for conceptualization and writing of the article. I finalized the software implementation and carried out all simulation studies. Mathias Drton made helpful suggestions regarding both content and presentation of the article during regular discussions. All co-authors contributed further by reviewing and commenting on the final version before submission and during the peer-review process.

The Annals of Statistics
2022, Vol. 50, No. 6, 3174–3196
https://doi.org/10.1214/22-AOS2221

HALF-TREK CRITERION FOR IDENTIFIABILITY OF LATENT VARIABLE MODELS

BY RINA FOYGEL BARBER[1,a], MATHIAS DRTON[2,b], NILS STURMA[2,c] AND LUCA WEIHS[3,d]

[1]*Department of Statistics, University of Chicago,* [a]*rina@uchicago.edu*
[2]*School of Computation, Information and Technology, Technical University of Munich,* [b]*mathias.drton@tum.de,* [c]*nils.sturma@tum.de*
[3]*Allen Institute for AI,* [d]*lucaw@allenai.org*

We consider linear structural equation models with latent variables and develop a criterion to certify whether the direct causal effects between the observable variables are identifiable based on the observed covariance matrix. Linear structural equation models assume that both observed and latent variables solve a linear equation system featuring stochastic noise terms. Each model corresponds to a directed graph whose edges represent the direct effects that appear as coefficients in the equation system. Prior research has developed a variety of methods to decide identifiability of direct effects in a latent projection framework, in which the confounding effects of the latent variables are represented by correlation among noise terms. This approach is effective when the confounding is sparse and effects only small subsets of the observed variables. In contrast, the new latent-factor half-trek criterion (LF-HTC) we develop in this paper operates on the original unprojected latent variable model and is able to certify identifiability in settings, where some latent variables may also have dense effects on many or even all of the observables. Our LF-HTC is an effective sufficient criterion for rational identifiability, under which the direct effects can be uniquely recovered as rational functions of the joint covariance matrix of the observed random variables. When restricting the search steps in LF-HTC to consider subsets of latent variables of bounded size, the criterion can be verified in time that is polynomial in the size of the graph.

1. Introduction. Equipped with an intuitive causal interpretation, structural equation models are very popular tools in a broad range of applied sciences (Spirtes, Glymour and Scheines (2000); Pearl (2009); Peters, Janzing and Schölkopf (2017)). Often, structural equation models involve latent variables, and it becomes a key problem to clarify whether parameters of interest are identifiable from the joint distribution of the observable variables. Many different criteria have been developed to decide such identifiability. The dominant approach in state-of-the-art methods is to project away latent variables, that is, their effects are absorbed into correlations among error terms in the structural equations. In contrast, we here consider models with explicit latent variables and show how the latent dependence structure may be used to certify identifiability even in cases with dense latent confounding, where projection approaches remain inconclusive.

Concretely, we study linear structural equation models with explicit latent variables. The precise setting of interest may be described as follows. Let $X = (X_v)_{v \in V}$ be a collection of $d = |V|$ observed variables, and let $L = (L_h)_{h \in \mathcal{L}}$ be $\ell = |\mathcal{L}|$ latent (unobserved) variables.

Received January 2022; revised June 2022.
MSC2020 subject classifications. 62H22, 62J05, 62R01.
Key words and phrases. Covariance matrix, factor analysis, graphical model, hidden variables, latent variables, parameter identification, structural equation model.

Suppose all variables are related by linear equations as

$$X_v = \sum_{w \neq v} \lambda_{wv} X_w + \sum_{h \in \mathcal{L}} \gamma_{hv} L_h + \varepsilon_v, \quad v \in V,$$

where λ_{wv} and γ_{hv} are real-valued parameters that are also known as *direct causal effects* of X_w on X_v and L_h on X_v, respectively. The ε_v are independent mean zero random variables that model noise. We assume that each ε_v has finite variance $\omega_v > 0$. The latent variables $(L_h)_{h \in \mathcal{L}}$ are assumed to be independent, and also independent of the noise terms $\varepsilon = (\varepsilon_v)_{v \in V}$. Since we are primarily interested in identification of direct causal effects λ_{vw}, we may fix, without loss of generality, the latent scale such that each L_h has mean zero and variance 1. Viewing X, L, and ε as vectors, the above equation system can be presented in the form

$$X = \Lambda^\top X + \Gamma^\top L + \varepsilon \tag{1.1}$$

with $d \times d$ parameter matrix $\Lambda = (\lambda_{wv})$ and $\ell \times d$ parameter matrix $\Gamma = (\gamma_{hv})$. The matrix Λ has zeros along the diagonal. Specific models are now derived from (1.1) by assuming specific sparsity patterns in Λ and Γ. The resulting models assume that all unobserved confounding is caused only by the explicitly modeled, independent latent variables. Thus, the latent structure corresponds to factor analysis models, and we will refer to the latent variables also as *latent factors*.

The models belong to the general framework of structural equation models with latent variables as they are considered, for example, in Bollen (1989). However, where many of the examples in Bollen's book are concerned with measurement models, that is, latent variables are measured through observations and these observations are conditionally independent given the latent variables, our interest here is the setting where we have direct causal effects λ_{wv} between observed variables and the latent variables constitute confounders.

The focus of this paper will be on the covariance structure posited by models derived from (1.1). In particular, we will be interested in determining when sparsity in the matrices Λ and Γ allows one to *identify* (i.e., uniquely recover) the direct effects λ_{wv} from the covariance matrix of the observable random vector X. Solving (1.1), we find

$$X = (I_d - \Lambda)^{-\top}\big(\Gamma^\top L + \varepsilon\big).$$

The vector $\Gamma^\top L + \varepsilon$ follows a latent factor model and has covariance matrix

$$\Omega = \operatorname{Var}[\varepsilon] + \Gamma^\top \operatorname{Var}[L]\Gamma = \Omega_{\mathsf{diag}} + \Gamma^\top \Gamma = \Omega_{\mathsf{diag}} + \sum_{h \in \mathcal{L}} \Gamma_h^\top \Gamma_h, \tag{1.2}$$

where Ω_{diag} is diagonal with entries $\Omega_{\mathsf{diag},vv} = \omega_v$ and Γ_h is the hth row of Γ such that the entries of Γ_h correspond to the causal effects associated to the latent factor L_h. We term the matrix Ω the *latent covariance matrix*. It follows that X has covariance matrix

$$\Sigma = (I_d - \Lambda)^{-\top} \Omega (I_d - \Lambda)^{-1}.$$

In order to study structural equation models it is useful to adopt a graphical perspective. To this end, the zero patterns in Λ and Γ are associated to a directed graph $G = (V \cup \mathcal{L}, D)$, where $D \subset (V \cup \mathcal{L}) \times (V \cup \mathcal{L})$ is a collection of directed edges $w \to v$. For two observed nodes $v, w \in V$, the effect λ_{wv} may be nonzero only if the edge $w \to v$ is contained in the set D. Similarly, for a latent node $h \in \mathcal{L}$ and an observed node $v \in V$, the effect γ_{hv} is possibly nonzero if $h \to v \in D$. In figures, we draw latent nodes h in gray, and we draw edges $h \to v$ dashed for better distinction. This is illustrated in the next example.

EXAMPLE 1.1. We consider an augmented version of an example from Stanghellini and Wermuth ((2005), Section 7), which pertains to the effects of sequential treatments in randomized clinical trials. Suppose that the patients receive two treatment doses in sequence, T_1

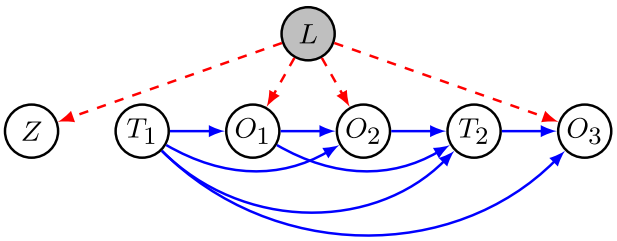


FIG. 1. *Graph corresponding to a randomized clinical trial for sequential administered treatments with a latent factor L.*

and T_2, and at both times the treatment dose is assigned at random. The randomization distribution of the second treatment dose T_2 depends on the previous treatment dose T_1 and on two intermediate outcome measures O_1 and O_2. The intermediate outcome measures are deemed potentially related, that is, O_2 may causally depend on O_1. After the second treatment a final outcome measure O_3 is recorded. Assume now that there is a latent factor L, such as a specific characteristic of a patient, that has effects on all outcomes O_1, O_2, O_3. Finally, as in Stanghellini and Wermuth (2005), we assume that there exists an auxiliary observed variable Z that provides a noisy measurement of L. The direct effects in this setup are depicted in the graph shown in Figure 1.

We aim to characterize those models of the form (1.1) that are rationally identifiable, that is, all possibly nonzero direct causal effects λ_{wv} can be uniquely recovered as rational functions of the entries of the observable covariance matrix Σ. This kind of identifiability has been examined in previous research in the context of latent projections where latent variables are not explicitly modeled. Models then correspond to *mixed graphs* that contain only the observed nodes V, but bidirectional edges in addition to the directed edges. Each bidirected edge represents a possibly nonzero entry in the latent covariance matrix Ω, that is, it implicitly indicates the presence of a confounding latent factor. The starting point for deriving sufficient criteria for rational identifiability are then the equations

$$\big[(I_d - \Lambda)^\top \Sigma (I_d - \Lambda)\big]_{vw} = \Omega_{vw} = 0, \tag{1.3}$$

which hold whenever no confounding latent factor affects both, X_v and X_w with $v \neq w$. The equations (1.3) are then solved to obtain the nonzero effects in Λ. This strategy has been leveraged to formulate graphical criteria applicable to mixed graph representations of latent variable models.

An example of a graphical criterion leveraging the latent projection approach is the half-trek criterion of Foygel, Draisma and Drton (2012), which can be considered as a predecessor and special case of the new results in this paper. But there are also various other graphical criteria on mixed graphs such as instrumental variables (Bowden and Turkington (1984)), conditional instruments (Brito and Pearl (2002)), the G-criterion (Brito and Pearl (2006)), auxiliary variables (Chen, Pearl and Bareinboim (2016)) and Chen, Kumor and Bareinboim (2017)), decomposition techniques (Tian (2005)) and several generalizations and further developments; cf. Tian (2009), Drton and Weihs (2016), Weihs et al. (2018), Kumor, Chen and Bareinboim (2019) and Kumor, Cinelli and Bareinboim (2020).

In contrast, in this work we consider the original, unprojected latent variable model as defined in (1.1), and we allow the latent covariance matrix Ω to be dense with only few or no zero entries. Then the usual approach of exploiting the zero structure in Ω that was highlighted in (1.3) is no longer effective. However, dense confounding of the observed variables may be caused by only a small number of latent factors, in which case the latent covariance

matrix Ω exhibits exploitable structure. Our key observation is that Ω may contain rank-deficient submatrices. For example, let $Y, Z \subseteq V$ be two disjoint sets of observed nodes. Then by (1.2) the submatrix $\Omega_{Y,Z}$ equals

$$\Omega_{Y,Z} = (\Omega_{\text{diag}})_{Y,Z} + \sum_{h \in \mathcal{L}} (\Gamma_h^\top \Gamma_h)_{Y,Z} = \sum_{h \in H} (\Gamma_h^\top \Gamma_h)_{Y,Z},$$

where the subset $H \subseteq \mathcal{L}$ over which we sum on the right-hand side contains exactly those latent factors that have an effect on a node in Y and at the same time also an effect on a node in Z. Since the matrix $\Gamma_h^\top \Gamma_h$ has rank one for each latent node h, the submatrix $\Omega_{Y,Z}$ is not of full column rank if $|H| < |Z|$. Exploiting this low rank structure of the latent covariance matrix Ω yields our main result, which is a sufficient criterion for rational identifiability of the direct causal effects λ_{wv}. We show how to convert the criterion into a graphical condition that can be checked using efficient algorithms under a bound on the considered rank. The graphical criterion is directly applicable to directed graphs $G = (V \cup \mathcal{L}, D)$ that explicitly contain the latent nodes $\mathcal{L}$, that is, the criterion operates on the unprojected latent variable model and allows to explore specific confounding. We refer to it as the *latent-factor half-trek criterion* (LF-HTC).

EXAMPLE 1.2. We take up the earlier example of a randomized clinical trial with sequential treatments, which we summarized in the graph in Figure 1. It is natural to investigate the direct causal effects between the observed variables T_1, O_1, O_2, T_2 and O_3. These direct causal effects correspond to the blue (non-dashed) edges in the figure. Our new latent-factor half-trek criterion will be able to certify that the whole parameter matrix Λ is rationally identifiable and all nonzero effects λ_{vw} can be written as rational formulas in the entries of the observable covariance matrix Σ. For example, the direct effect from the first treatment dose T_1 on the intermediate outcome O_1 is given by $\Sigma_{T_1,O_1} / \Sigma_{T_1,T_1}$; a standard regression coefficient. But remarkably, we can even identify effects corresponding to the edges $T_1 \to O_2$ and $O_1 \to O_2$ by the latent-factor half-trek criterion. We verified that it is impossible to identify the latter two effects in the latent projection framework (cf. Section 4).

While most of the general identification criteria have been developed in the setting of latent projections, some existing work also considers unprojected latent factor models as defined in (1.1). However, this work addresses special types of latent confounding only. For example, Stanghellini and Wermuth (2005) and Leung, Drton and Hara (2016) examine linear latent variable models with one latent variable, and the conditional instrument approach in Van Der Zander, Textor and Liskiewicz (2015) covers scenarios in which no confounding factor has an effect on all observed variables. Another approach requires that latent factors are measured through observed proxy variables and relies on identifying the causal effect between the latent factor and the proxy; see, for example, Kuroki and Pearl (2014), Miao, Geng and Tchetgen Tchetgen (2018) and Lee and Bareinboim (2021), the latter of which deals with the discrete case.

It should be noted that, in principle, rational identifiability is always decidable by computational algebraic geometry (Garcia-Puente, Spielvogel and Sullivant (2010)) involving Gröbner basis computations (Cox, Little and O'Shea (2007)). However, in the worst case, the complexity of these methods can be double exponential in the size of the graph. Thus, they may be infeasible even for relatively small graphs, and more efficient graphical criteria are of great value. To check the new latent-factor half-trek criterion we propose an algorithm based on max-flow computations (Cormen et al. (2009)) that runs in polynomial time in the size of the graph if we confine ourselves to search only over subsets of latent factors of bounded size. We show that the restriction of the search space is necessary since the task of checking the latent-factor half-trek criterion without restrictions is in general NP-complete.

The organization of the paper is as follows. In Section 2, we provide a precise definition of linear structural equation models given by directed graphs and rigorously introduce the concept of rational identifiability. Moreover, we derive basic necessary conditions for rational identifiability based on dimension arguments. In Section 3, we present our main result, the LF-HTC. In Section 4, we discuss the latent projection framework considered in previous research and compare the new LF-HTC to existing criteria. In particular, we compare the LF-HTC to the original half-trek criterion. In Section 5 we present an algorithm to check the LF-HTC efficiently. Using this algorithm we systematically check identifiability of certain classes of small latent-factor graphs in Section 6. The restriction to small graphs allows for these checks to be validated using suitably designed Gröbner basis computations. Finally, the proof of the main result is given in Section 7. Further elements of proofs, a hardness result for checking the LF-HTC without a bound on the cardinality of searched sets of latent variables and an explanation on how to effectively deploy techniques from computational algebraic geometry are deferred to the Supplementary Material (Barber et al. (2022)).

2. Graphical representation and identifiability. Let $G = (V \cup \mathcal{L}, D)$ be a directed graph where V and $\mathcal{L}$ are finite disjoint sets of observed and latent nodes, respectively. We emphasize that G is allowed to contain directed cycles. Let $d = |V|$ and $\ell = |\mathcal{L}|$. The edge set $D \subset (V \cup \mathcal{L}) \times (V \cup \mathcal{L})$ is assumed to be free of self-loops, so $v \to v \notin D$ for all $v \in V \cup \mathcal{L}$. For each vertex $v \in V \cup \mathcal{L}$, define its set of parents as $\mathrm{pa}(v) = \{w \in V \cup \mathcal{L} : w \to v \in D\}$. Throughout the paper, we require $\mathrm{pa}(h) = \varnothing$ for all $h \in \mathcal{L}$, so that all latent nodes are source nodes and the outgoing edges of latent nodes only point to observed nodes. If this condition is satisfied, we call G a *latent-factor graph* and, to emphasize the set of latent variables, write $G^{\mathcal{L}}$ instead of G.

The edge set of a latent-factor graph may be partitioned as $D = D_V \cup D_{\mathcal{L}V}$, where $D_V = D \cap (V \times V)$ is the set of directed edges between observed nodes and $D_{\mathcal{L}V} = D \cap (\mathcal{L} \times V)$ is the set of directed edges that point from latent to observed nodes. Let $\mathbb{R}^{D_V}$ be the set of real $d \times d$ matrices $\Lambda = (\lambda_{wv})$ with support D_V, that is, $\lambda_{wv} = 0$ if $w \to v \notin D_V$. Write $\mathbb{R}^{D_V}_{\mathrm{reg}}$ for the subset of matrices $\Lambda \in \mathbb{R}^{D_V}$ with $I_d - \Lambda$ invertible; recall that we allow $G^{\mathcal{L}}$ to contain directed cycles. Similarly, let $\mathbb{R}^{D_{\mathcal{L}V}}$ be the set of real $\ell \times d$ matrices $\Gamma = (\gamma_{hv})$ with support $D_{\mathcal{L}V}$, that is, $\gamma_{hv} = 0$ if $h \to v \notin D_{\mathcal{L}V}$. Additionally, we write diag^+_d for the set of all $d \times d$ diagonal matrices with a positive diagonal indexed by the elements of V.

Each latent-factor graph postulates a covariance model that corresponds to a linear structural equation model specified via (1.1).

DEFINITION 2.1. The covariance model given by a latent-factor graph $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ with $|V| = d$ and $|\mathcal{L}| = \ell$ is the family of covariance matrices

$$\Sigma = (I_d - \Lambda)^{-\top} \Omega (I_d - \Lambda)^{-1} \tag{2.1}$$

obtained from choices of $\Lambda \in \mathbb{R}^{D_V}_{\mathrm{reg}}$ and Ω in the image of the map

$$\tau : \mathbb{R}^{D_{\mathcal{L}V}} \times \mathrm{diag}^+_d \longrightarrow \mathrm{PD}(d),$$
$$(\Gamma, \Omega_{\mathsf{diag}}) \longmapsto \Omega_{\mathsf{diag}} + \Gamma^\top \Gamma,$$

where $\mathrm{PD}(d)$ is the cone of positive definite symmetric $d \times d$ matrices. We term the image $\mathrm{Im}(\tau) \subseteq \mathrm{PD}(d)$ the *cone of latent covariance matrices*.

We are interested in the question of identifiability, that is, whether the matrix Λ can be uniquely recovered from a given covariance matrix Σ of the form (2.1). If it is possible to recover the whole matrix Λ uniquely, we can determine Ω uniquely by the equation

$$(I_d - \Lambda)^\top \Sigma (I_d - \Lambda) = \Omega, \tag{2.2}$$

since the matrix $I_d - \Lambda$ is assumed to be invertible. Thus, for $\Theta = \mathbb{R}^{D_V}_{\text{reg}} \times \text{Im}(\tau)$, identifiability holds if the parametrization map

$$\begin{aligned} \varphi_{G^{\mathcal{L}}} : \Theta &\longrightarrow \text{PD}(d), \\ (\Lambda, \Omega) &\longmapsto (I_d - \Lambda)^{-\top} \Omega (I_d - \Lambda)^{-1} \end{aligned} \tag{2.3}$$

is injective on Θ, or a suitably large subset. Since identifiability will usually not hold on the whole set Θ, we need to clarify what we mean by a "suitably large" subset. We use terminology from algebraic geometry, background can be found in Cox, Little and O'Shea (2007), Shafarevich (2013) or Hartshorne (1977).

A property on an irreducible algebraic set W is said to be generically true if the property holds on the complement $W \setminus A$ of a proper algebraic subset $A \subseteq W$. Due to irreducibility, the complement $W \setminus A$ is dense in W with respect to the Zariski topology and therefore considered as a "suitably large" subset. When W is an irreducible algebraic set defined over the real numbers, a proper algebraic subset of W has Lebesgue measure zero; see, for example, the lemma in Okamoto (1973).

To connect this terminology to our setup, we observe that the Zariski closure $\overline{\Theta}$, that is, the smallest algebraic subset that contains the domain Θ, is irreducible. This is true because Θ is the polynomial image of an open set. Hence, we say that a property on Θ is generically true if there exists a proper algebraic subset $A \subset \overline{\Theta}$ such that the property holds on the complement $\Theta \setminus A$. Our interest is now in generically identifying the direct causal effects λ_{wv}. Since the parametrization $\varphi_{G^{\mathcal{L}}}$ is rational, the identification formula, in the worst case, is an algebraic function (Garcia-Puente, Spielvogel and Sullivant (2010)). However, in all examples we know, if generic identifiability is possible, then by rational formulas. This motivates the following definition.

DEFINITION 2.2 (Rational identifiability).

(a) The latent-factor graph $G^{\mathcal{L}}$ is said to be rationally identifiable if there exists a proper algebraic subset $A \subset \overline{\Theta}$ and a rational map $\psi : \text{PD}(d) \to \mathbb{R}^{D_V}_{\text{reg}} \times \text{PD}(d)$ such that $\psi \circ \varphi_{G^{\mathcal{L}}}(\Lambda, \Omega) = (\Lambda, \Omega)$ for all $(\Lambda, \Omega) \in \Theta \setminus A$.

(b) The direct causal effect λ_{vw}, or also simply the edge $v \to w \in D_V$, is rationally identifiable if there exists a proper algebraic subset $A \subset \overline{\Theta}$ and a rational map $\psi : \text{PD}(d) \to \mathbb{R}$ such that $\psi \circ \varphi_{G^{\mathcal{L}}}(\Lambda, \Omega) = \lambda_{vw}$ for all $(\Lambda, \Omega) \in \Theta \setminus A$.

Rational identifiability of $G^{\mathcal{L}}$ is equivalent to rational identifiability of all edges in D_V; recall (2.2). If $G^{\mathcal{L}}$ is rationally identifiable, then a (absolutely continuous) random choice of the effects in (Λ, Γ) and the error variances in Ω_{diag} will almost surely yield a covariance matrix for the observable vector X from which Λ can be recovered uniquely by rational formulas. If $G^{\mathcal{L}}$ is not generically identifiable, its parametrization $\varphi_{G^{\mathcal{L}}}$ may be either generically finite-to-one or generically infinite-to-one.

DEFINITION 2.3. Let $f : S \to \mathbb{R}^n$ be a map defined on a subset $S \subseteq \mathbb{R}^m$ such that the Zariski closure $\overline{S}$ is irreducible. Then f is generically finite-to-one if there exists a proper algebraic subset $A \subseteq \overline{S}$ such that the fiber $\mathcal{F}_f(s) = f^{-1}(f(s))$ is finite for all $s \in S \setminus A$. Otherwise, f is said to be generically infinite-to-one.

DEFINITION 2.4. A latent-factor graph $G^{\mathcal{L}}$ is generically finite-to-one if its parametrization $\varphi_{G^{\mathcal{L}}}$ is generically finite-to-one. In this case, we will also say that $G^{\mathcal{L}}$ is finitely identifiable. Otherwise, $G^{\mathcal{L}}$ is said to be generically infinite-to-one.

Note that if a latent-factor graph $G^{\mathcal{L}}$ is rationally identifiable, then the fiber $\mathcal{F}_{\varphi_{G^{\mathcal{L}}}}(\Lambda, \Omega) = \{(\Lambda, \Omega)\}$ for all parameter choices outside of a proper algebraic subset. In particular, a graph that is rationally identifiable is generically finite-to-one. The following lemma is an important tool to check if a rational map is generically finite-to-one. For completeness, we provide a proof in Appendix A in the Supplementary Material (Barber et al. (2022)). Here, we rely on the notion of semialgebraic sets, which are finite unions of sets defined by finitely many polynomial equations and inequalities. For background on semialgebraic sets, we refer to Bochnak, Coste and Roy (1998), Basu, Pollack and Roy (2006) and Benedetti and Risler (1990).

LEMMA 2.5. *Let $S \subseteq \mathbb{R}^m$ be a semialgebraic set such that the Zariski closure $\overline{S}$ is irreducible. Then a rational mapping $f : S \to \mathbb{R}^n$ is generically finite-to-one if and only if* $\dim(f(S)) = \dim(S)$. *In particular, if* $\dim(S) > n$ *then f must be generically infinite-to-one.*

REMARK 2.6. If the rational mapping in Lemma 2.5 is infinite-to-one, then it holds that the fiber is infinite for almost all $s \in S$. This can be seen, in particular, by inspecting the proof of Lemma 2.5.

In our context, the rational mapping of interest is the parametrization map $\varphi_{G^{\mathcal{L}}}$, which maps into the positive definite cone $\mathrm{PD}(d)$. We observe that a latent-factor graph $G^{\mathcal{L}}$ cannot be finite-to-one if the dimension of the domain $\Theta = \mathbb{R}^{D_V}_{\mathrm{reg}} \times \mathrm{Im}(\tau)$ is larger than the dimension of $\mathrm{PD}(d)$. This gives a basic necessary condition.

COROLLARY 2.7. *A latent-factor graph $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ is generically infinite-to-one if* $|D_V| + \dim(\mathrm{Im}(\tau)) > \binom{d+1}{2}$.

PROOF. To apply Lemma 2.5, we have to show that $\Theta = \mathbb{R}^{D_V}_{\mathrm{reg}} \times \mathrm{Im}(\tau)$ is semialgebraic, its closure is irreducible and that the parametrization map $\varphi_{G^{\mathcal{L}}}$ is rational. The first two claims are true since Θ is the polynomial image of an open semialgebraic set. Moreover, the map $\varphi_{G^{\mathcal{L}}}$ is rational due to Cramér's rule.

Now, we study the dimensions of Θ and the image $\varphi_{G^{\mathcal{L}}}(\Theta)$. The dimension of Θ is equal to $|D_V| + \dim(\mathrm{Im}(\tau))$ since the dimension of the product of two semialgebraic sets is the sum of their individual dimensions (Bochnak, Coste and Roy (1998), Proposition 2.8.5). Since the image of $\varphi_{G^{\mathcal{L}}}$ lies in the positive definite cone $\mathrm{PD}(d)$, we have

$$\dim\bigl(\varphi_{G^{\mathcal{L}}}(\Theta)\bigr) \le \dim\bigl(\mathrm{PD}(d)\bigr) = \binom{d+1}{2}.$$

Thus, if $|D_V| + \dim(\mathrm{Im}(\tau)) > \binom{d+1}{2}$, then $\dim(\Theta) > \dim(\varphi_{G^{\mathcal{L}}}(\Theta))$ and by Lemma 2.5 we conclude that $\varphi_{G^{\mathcal{L}}}$ is generically infinite-to-one. □

EXAMPLE 2.8. Consider the graph in Figure 2 where the latent structure is that of a one-factor model. By Theorem 2 in Drton, Sturmfels and Sullivant (2007), we have

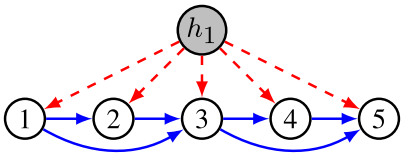


FIG. 2. *Latent-factor graph that is (trivially) generically infinite-to-one.*

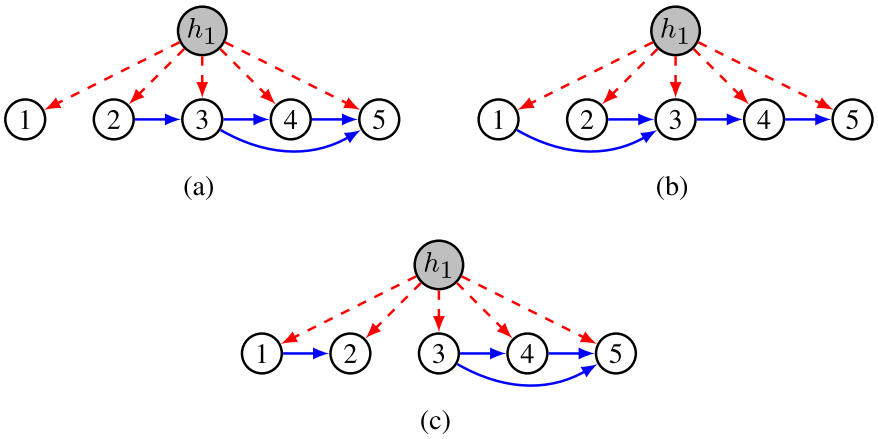


FIG. 3. *Latent-factor graphs with one latent-factor. (a) Rationally identifiable. (b) Generically finite-to-one but not rationally identifiable. (c) Generically infinite-to-one.*

$\dim(\mathrm{Im}(\tau)) = 10$; with only one factor the dimension is equal to the number of edges from the latent node to the observed nodes, $|D_{LV}| = 5$, plus the 5 parameters appearing on the diagonal of the matrix Ω_{diag}. But since the number of observed edges is $|D_V| = 6$, we have that $16 = |D_V| + \dim(\mathrm{Im}(\tau)) > \binom{6}{2} = 15$ and therefore the graph is generically infinite-to-one by Corollary 2.7.

If a latent-factor graph is not trivially infinite-to-one by dimension comparison, then it becomes more difficult to decide whether it is generically infinite-to-one, generically finite-to-one or rationally identifiable. Figure 3 shows latent-factor graphs that only have subtle differences in their structures but each of them has a different status of identifiability.

3. Main identifiability result. The main idea underlying our sufficient condition for rational identifiability is to exploit the low rank structure of the latent covariance matrix

$$\Omega = \Omega_{\mathsf{diag}} + \sum_{h \in \mathcal{L}} \Gamma_h^\top \Gamma_h.$$

Recall that $\Omega_{\mathsf{diag}} \in \mathrm{diag}_d^+$ is diagonal and Γ_h is the hth row of $\Gamma \in \mathbb{R}^{D_{\mathcal{L}V}}$. For a node $v \in V$, denote by $\mathrm{pa}_V(v) = \{w \in V : w \to v \in D_V\}$ the set of *observed parents* and by $\mathrm{pa}_{\mathcal{L}}(v) = \{w \in \mathcal{L} : w \to v \in D_{\mathcal{L}V}\}$ the set of *latent parents*. So, $\mathrm{pa}(v) = \mathrm{pa}_V(v) \cup \mathrm{pa}_{\mathcal{L}}(v)$. Focusing on a fixed node $v \in V$, it is our goal to find linear equations that determine the direct causal effects corresponding to the observed parents, that is, we aim to determine the vector $\Lambda_{\mathrm{pa}_V(v),v}$. Our approach is to find suitable sets of observed nodes $Y, Z \subseteq V \setminus \{v\}$ and a set of latent nodes $H \subseteq \mathcal{L}$ with $|H| = |Z|$ such that the latent covariance matrix contains a submatrix that satisfies

$$\Omega_{Y,Z\cup\{v\}} = \sum_{h \in H} (\Gamma_h^\top \Gamma_h)_{Y,Z\cup\{v\}} \tag{3.1}$$

and fails to have full column rank. The drop in rank means that the entries of the submatrix exhibit algebraic relations, which we may then use to identify the targeted direct causal effects.

The equality in (3.1) holds if (i) $Y \cap (Z \cup \{v\}) = \varnothing$ and (ii) $\mathrm{pa}_{\mathcal{L}}(Y) \cap \mathrm{pa}_{\mathcal{L}}(Z \cup \{v\}) \subseteq H$. Indeed, (i) ensures that $(\Omega_{\mathsf{diag}})_{Y,Z\cup\{v\}} = 0$ because the considered submatrix does not involve any diagonal elements. And by (ii), the set H contains all latent factors that have an effect on a node in Y and at the same time an effect on a node in $Z \cup \{v\}$. Assume there exists a triple of sets (Y, Z, H) with $|H| = |Z|$ and satisfying (i) and (ii) above. Then

$$\mathrm{rank}(\Omega_{Y,Z\cup\{v\}}) = \mathrm{rank}\Big(\sum_{h \in H} (\Gamma_h^\top \Gamma_h)_{Y,Z\cup\{v\}}\Big) \le |H| = |Z|,$$

since the matrix $\Gamma_h^\top \Gamma_h$ has rank one for each $h \in \mathcal{L}$. Hence, the matrix $\Omega_{Y,Z\cup\{v\}}$ does not have full column rank. Moreover, suppose that we are able to ensure that the smaller submatrix $\Omega_{Y,Z}$ is of full column rank $|Z|$. Then, since the column ranks of $\Omega_{Y,Z\cup\{v\}}$ and $\Omega_{Y,Z}$ are equal, the vector $\Omega_{Y,v}$ must be a linear combination of the columns of $\Omega_{Y,Z}$, that is, there exists $\psi \in \mathbb{R}^{|Z|}$ such that $\Omega_{Y,Z} \cdot \psi = \Omega_{Y,v}$. Using the identity $(I_d - \Lambda)^\top \Sigma (I_d - \Lambda) = \Omega$ from (2.2), this is equivalent to

$$\big[(I_d - \Lambda)^\top \Sigma (I_d - \Lambda)\big]_{Y,v} - \big[(I_d - \Lambda)^\top \Sigma (I_d - \Lambda)\big]_{Y,Z} \cdot \psi = 0.$$

Rewriting the matrix on the left, we get the system of equations

$$\begin{aligned} &\Big(\big[(I_d - \Lambda)^\top \Sigma\big]_{Y,\mathrm{pa}_V(v)} \quad \big[(I_d - \Lambda)^\top \Sigma (I_d - \Lambda)\big]_{Y,Z}\Big) \cdot \begin{pmatrix} \Lambda_{\mathrm{pa}_V(v),v} \\ \psi \end{pmatrix} \\ &\quad = \big[(I_d - \Lambda)^\top \Sigma\big]_{Y,v}. \end{aligned} \tag{3.2}$$

Now, if we make sure the matrix on the left-hand side in (3.2) is square and invertible, we can solve the system for the unknown parameters $\Lambda_{\mathrm{pa}_V(v),v}$. However, for this to be useful for parameter identification, suitable entries of Λ must already be known from earlier similar calculations in order to determine the coefficient matrix and the vector on the right-hand side of (3.2).

EXAMPLE 3.1. Consider the graph in Figure 3(a). Since there is one latent factor having dense effect on all observed variables, the parameter matrix Γ is given by the row vector $(\gamma_{11}, \ldots, \gamma_{15})$. Now focus on node $v = 3$ which only has a single observed parent. We aim to recover the effect $\Lambda_{\mathrm{pa}_V(3),3} = \lambda_{23}$ and we claim that the triple $(Y, Z, H) = (\{2, 4\}, \{1\}, \{h_1\})$ satisfies the properties discussed above. Clearly, it holds that $|H| = |Z|$, we have empty intersection $Y \cap (Z \cup \{v\})$, and the only common latent parent of Y and $Z \cup \{v\}$ is h_1, that is, $\mathrm{pa}_{\mathcal{L}}(Y) \cap \mathrm{pa}_{\mathcal{L}}(Z \cup \{v\}) \subseteq H$. By inspecting the rank one submatrix

$$\Omega_{Y,Z\cup\{v\}} = \begin{pmatrix} \gamma_{12} \\ \gamma_{14} \end{pmatrix} \cdot (\gamma_{11} \quad \gamma_{13}) = \begin{pmatrix} \gamma_{12}\gamma_{11} & \gamma_{12}\gamma_{13} \\ \gamma_{14}\gamma_{11} & \gamma_{14}\gamma_{13} \end{pmatrix}$$

we can easily deduce the relation

$$\Omega_{Y,Z} \cdot \frac{\gamma_{13}}{\gamma_{11}} = \Omega_{Y,v}$$

which holds true for generic choices of γ_{11}, that is, for $\gamma_{11} \neq 0$. In other words, the parameter ψ is equal to γ_{13}/γ_{11} and the equation system (3.2) is given by

$$\begin{pmatrix} \sigma_{22} & \sigma_{12} \\ -\lambda_{34}\sigma_{23} + \sigma_{24} & -\lambda_{34}\sigma_{13} + \sigma_{14} \end{pmatrix} \begin{pmatrix} \lambda_{23} \\ \psi \end{pmatrix} = \begin{pmatrix} \sigma_{23} \\ -\lambda_{34}\sigma_{33} + \sigma_{34} \end{pmatrix},$$

where σ_{ij} is the ijth entry of the covariance matrix Σ. If we already knew that the effect λ_{34} is given by a rational function in Σ, then we could also recover the effect λ_{23} by a rational function of Σ since the matrix on the left-hand side is quadratic and generically invertible.

Our main result shows that the above story can be made practical and yields a criterion to recursively identify columns in Λ. Importantly, the imposed conditions can all be translated into combinatorial conditions on the considered latent-factor graph. The resulting method is proven correct in Theorem 3.7 below. Before stating the theorem we define the necessary graphical concepts, which involve special types of paths that we term latent-factor half-treks. Recall that a path from node v to w in a latent-factor graph $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ is a sequence of edges that connects the consecutive nodes in a sequence of nodes beginning in v and ending in w.

DEFINITION 3.2 (Latent-factor half-trek). A path π in the latent-factor graph $G^{\mathcal{L}}$ is a *latent-factor half-trek from source v to target w* if it is a path from $v \in V$ to $w \in V$ in $G^{\mathcal{L}}$ and is of the form

$$v \to x_1 \to \cdots \to x_n \to w$$

or of the form

$$v \leftarrow h \to x_1 \to \cdots \to x_n \to w$$

for $x_1, \ldots, x_n \in V$ and for some $h \in \mathcal{L}$.

The name latent-factor half-trek is inspired by the customary notion of a trek, which is a pair of directed paths (π_1, π_2) that share the same source node. If a latent-factor half-trek is of the first form in Definition 3.2, we say that the left-hand side of π, written $\text{Left}(\pi)$, is the node v and the right-hand side, written $\text{Right}(\pi)$, is the set of nodes $\{v, x_1, \ldots, x_n, w\}$. In the second case $\text{Left}(\pi) = \{v, h\}$ and $\text{Right}(\pi) = \{h, x_1, \ldots, x_n, w\}$. A latent-factor half-trek from v to v may have no edges, in this case $\text{Left}(\pi) = \text{Right}(\pi) = \{v\}$ and the half-trek is called trivial. For a set of n latent-factor half-treks, $\Pi = \{\pi_1, \ldots, \pi_n\}$, let v_i and w_i be the source and the target of π_i. If the sources are all distinct and the targets are all distinct, then we say that Π is a system of latent-factor half-treks from $A = \{v_1, \ldots, v_n\}$ to $B = \{w_1, \ldots, w_n\}$. A set of latent-factor half-treks $\Pi = \{\pi_1, \ldots, \pi_n\}$ has no sided intersection if

$$\text{Left}(\pi_i) \cap \text{Left}(\pi_j) = \varnothing = \text{Right}(\pi_i) \cap \text{Right}(\pi_j) \quad \text{for all } i \neq j.$$

EXAMPLE 3.3. Consider the graph in Figure 3(a). Then the system of latent-factor half-treks

$$\{\pi_1 : 5 \leftarrow h_1 \to 3, \pi_2 : 4 \to 5\}$$

has no sided intersection. On the other hand, the system

$$\{\widetilde{\pi}_1 : 2 \leftarrow h_1 \to 3, \widetilde{\pi}_2 : 3 \to 4 \to 5\}$$

has sided intersection since $\text{Right}(\widetilde{\pi}_1) \cap \text{Right}(\widetilde{\pi}_2) = \{3\}$.

DEFINITION 3.4 (Latent-factor half-trek reachability). Let $v, w \in V$ be two distinct observed nodes in a latent-factor graph $G^{\mathcal{L}}$. Let $H \subseteq \mathcal{L}$ be a set of latent factors. If there exists a latent-factor half-trek from v to w through the latent-factor graph $G^{\mathcal{L}}$, which does not pass through any node in H, then we say that w *is half-trek reachable from v while avoiding H*, and write $w \in \text{htr}_H(v)$. For a set $U \subseteq V$, we write $w \in \text{htr}_H(U)$ if $w \in \text{htr}_H(u)$ for some $u \in U$.

EXAMPLE 3.5. Consider the graph in Figure 3(a), and let $H = \varnothing$. Then $2 \in \text{htr}_H(1)$ since there is the latent-factor half-trek $1 \leftarrow h_1 \to 2$ and $h_1 \notin H$. But if $H = \{h_1\}$, then $\text{htr}_H(1) = \varnothing$ since there is no latent-factor half-trek from node 1 to any other node in the graph while avoiding the node h_1.

DEFINITION 3.6 (Latent-factor half-trek criterion). Given a node $v \in V$, the triple $(Y, Z, H) \in 2^{V \setminus \{v\}} \times 2^{V \setminus \{v\}} \times 2^{\mathcal{L}}$ satisfies the *latent-factor half-trek criterion* (LF-HTC) with respect to v if:

(i) $|Y| = |\text{pa}_V(v)| + |H|$ and $|Z| = |H|$ with $Z \cap \text{pa}_V(v) = \varnothing$,
(ii) $Y \cap (Z \cup \{v\}) = \varnothing$ and $\text{pa}_{\mathcal{L}}(Y) \cap \text{pa}_{\mathcal{L}}(Z \cup \{v\}) \subseteq H$, and

(iii) there exists a system of latent-factor half-treks with no sided intersection from Y to $Z \cup \mathrm{pa}_V(v)$ in $G^{\mathcal{L}}$, such that for each $z \in Z$, the half-trek terminating at z takes the form $y \leftarrow h \rightarrow z$ for some $y \in Y$ and some $h \in H$.

If a triple (Y, Z, H) satisfies the LF-HTC with respect to a node v, then condition (ii) ensures that the submatrix $\Omega_{Y,Z\cup\{v\}}$ of the latent covariance matrix can be written as in (3.1) and, since $|Z| = |H|$, the submatrix does not have full column rank. Moreover, condition (iii) ensures that the matrix on the left-hand side of (3.2) is invertible. The latter claim will be established by means of an application of the Gessel–Viennot–Lindström lemma (Gessel and Viennot (1985); Lindström (1973)). We now state our main result; its proof is deferred to Section 7. For a directed edge $u \rightarrow y \in D$, we say that y is the *head* of the edge.

THEOREM 3.7 (LF-HTC-identifiability). *Suppose triple* $(Y, Z, H) \in 2^{V\setminus\{v\}} \times 2^{V\setminus\{v\}} \times 2^{\mathcal{L}}$ *satisfies the LF-HTC with respect to* $v \in V$. *If all directed edges* $u \rightarrow y \in D_V$ *with head* $y \in Z \cup (Y \cap \mathrm{htr}_H(Z \cup \{v\}))$ *are rationally identifiable, then all directed edges in* D_V *with* v *as a head are rationally identifiable.*

This theorem yields the basis for an efficient algorithm that recursively solves for all direct causal effects corresponding to the edges D_V in a latent-factor graph. That is, we recover the matrix Λ column-by-column. The corresponding algorithm is detailed in Section 5. We refer to a latent-factor graph $G^{\mathcal{L}}$ as *LF-HTC-identifiable* if all columns of Λ may be recovered recursively by Theorem 3.7.

EXAMPLE 3.8. The latent-factor graph in Figure 3(a) is LF-HTC-identifiable. To see this, we recursively check all nodes $v \in V = \{1, 2, 3, 4, 5\}$. That is, for each $v \in V$ we find a triple (Y, Z, H) that satisfies the LF-HTC such that all nodes in $Z \cup (Y \cap \mathrm{htr}_H(Z \cup \{v\}))$ were already checked successfully to satisfy the LF-HTC in the steps before.

$\underline{v = 1, 2}$: The triple $(Y, Z, H) = (\varnothing, \varnothing, \varnothing)$ trivially satisfies the LF-HTC since $\mathrm{pa}_V(v) = \varnothing$.

$\underline{v = 4}$: Let $(Y, Z, H) = (\{2, 3\}, \{1\}, \{h_1\})$. Conditions (i) and (ii) are easily checked and for condition (iii) consider the system of latent-factor half-treks $\{3, 2 \leftarrow h_1 \rightarrow 1\}$ where 3 corresponds to the trivial trek from 3 to 3. Finally, note that we have $Y \cap \mathrm{htr}_H(Z \cup \{v\}) = \{2, 3\} \cap \{4, 5\} = \varnothing$ and that the node $1 \in Z$ was already checked successfully in the last step.

$\underline{v = 3}$: Let $(Y, Z, H) = (\{2, 4\}, \{1\}, \{h_1\})$. Then the system of latent-factor half-treks $\{2, 4 \leftarrow h_1 \rightarrow 1\}$ satisfies (iii) and $Z \cup (Y \cap \mathrm{htr}_H(Z \cup \{v\})) = \{1, 4\}$.

$\underline{v = 5}$: Let $(Y, Z, H) = (\{2, 3, 4\}, \{1\}, \{h_1\})$. Then the system of latent-factor half-treks $\{3, 4, 2 \leftarrow h_1 \rightarrow 1\}$ satisfies (iii) and $Z \cup (Y \cap \mathrm{htr}_H(Z \cup \{v\})) = \{1\}$.

If the observed part (V, D_V) of a latent-factor graph does not contain directed cycles, then the latent-factor graph is said to be *acyclic*. Moreover, we say that a latent-factor graph is *bow-free* if it does not contain any two observed vertices $v, w \in V$ such that there is a directed edge between v and w and, in addition, there is a latent factor $h \in \mathcal{L}$ that has directed edges pointing to both v and w. As a special case of Theorem 3.7, we have the following straightforward observation.

COROLLARY 3.9. *Bow-free acyclic latent-factor graphs are rationally identifiable.*

PROOF. Let $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ be a latent-factor graph. It is easy to see that for every node $v \in V$ the triple $(Y, Z, H) = (\mathrm{pa}_V(v), \varnothing, \varnothing)$ satisfies the LF-HTC with respect to v since v and $\mathrm{pa}_V(v)$ do not have a common latent parent (i.e., $\mathrm{pa}_{\mathcal{L}}(\mathrm{pa}_V(v)) \cap \mathrm{pa}_{\mathcal{L}}(v) = \varnothing$).

The observed part (V, D_V) is a directed aycylic graph (DAG) and therefore induces at least one topological ordering $\prec$ on V, that is, an ordering such that $v \to w \in D_V$ only if $v \prec w$. Importantly, all parents $w \in \mathrm{pa}_V(v)$ are predecessors of v with respect to $\prec$. Thus, by Theorem 3.7, we can determine rational identifiability of all edges in D_V in a step-wise manner according to the ordering $\prec$ and using the triple $(\mathrm{pa}_V(v), \varnothing, \varnothing)$ for each $v \in V$. We conclude that $G^{\mathcal{L}}$ is LF-HTC-identifiable and hence, in particular, rationally identifiable. □

4. Latent projections. As mentioned in the introduction, previous criteria for rational identifiability of direct causal effects operate on mixed graphs obtained by a projection. These projections can be defined for general directed graphs with hidden variables (Maathuis et al. (2019), Chapter 2 and Pearl (2009), Chapter 2), but we treat the special case of latent-factor graphs:

DEFINITION 4.1 (Maathuis et al. (2019), Chapter 2). Let $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ be a latent-factor graph. Define a new graph starting with the induced subgraph $G' = (V, D_V)$ and add edges as follows:

$$\text{Whenever } v \leftarrow h \to w \quad \text{in } G^{\mathcal{L}} \text{ for } h \in \mathcal{L} \text{ and } v, w \in V, \text{ add } v \leftrightarrow w \text{ to } G'.$$

The mixed graph $G' = (V, D_V, B)$ is the *latent projection* of $G^{\mathcal{L}}$, where B is the collection of bidirected edges $v \leftrightarrow w$. They have no orientation, that is, $v \leftrightarrow w \in B$ if and only if $w \leftrightarrow v \in B$.

Every mixed graph defines a covariance model. Denote $\mathrm{PD}(B) \subseteq \mathrm{PD}(d)$ the subcone of matrices with support B, that is, for $\Omega = (\omega_{vw}) \in \mathrm{PD}(B)$ we have $\omega_{vw} = 0$ if $v \neq w$ and $v \leftrightarrow w \notin B$.

DEFINITION 4.2. The covariance model given by a mixed graph $G' = (V, D_V, B)$ with $V = |d|$ is the family of covariance matrices

$$\Sigma = (I_d - \Lambda)^{-\top} \Omega (I_d - \Lambda)^{-1}$$

obtained from choices of $\Lambda \in \mathbb{R}_{\mathrm{reg}}^{D_V}$ and $\Omega \in \mathrm{PD}(B)$.

For any latent-factor graph, the cone of latent covariance matrices $\mathrm{Im}(\tau)$ is clearly a subset of $\mathrm{PD}(B)$, the cone of latent covariance matrices of the latent projection. Thus, a covariance model given by a latent-factor graph is a submodel of the covariance model given by its latent projection. More details on the at times subtle differences between $\mathrm{Im}(\tau)$ and $\mathrm{PD}(B)$ can be found in Drton and Yu (2010).

In the remainder of this section, we focus on the predecessor of the LF-HTC that operates on mixed graphs, namely the original half-trek criterion (HTC) of Foygel, Draisma and Drton (2012). We say that a mixed graph is HTC-identifiable if it is rationally identifiable by this criterion.

At first sight, it appears as if the HTC coincides with the version of the LF-HTC obtained by only allowing $H = Z = \varnothing$; compare Def. 4 in Foygel, Draisma and Drton (2012) with Definition 3.6 here. However, as we will show below there is a subtle difference in the way systems of half-treks with no sided intersection are defined. Indeed, in the setting of the LF-HTC two half-treks may also intersect at latent nodes, whereas in the HTC intersections are only possible at observed nodes. Intuitively, each bidirected edge in a latent projection can amount to confounding induced by a separate latent variable. Before highlighting this subtlety, we first exemplify an application of HTC.

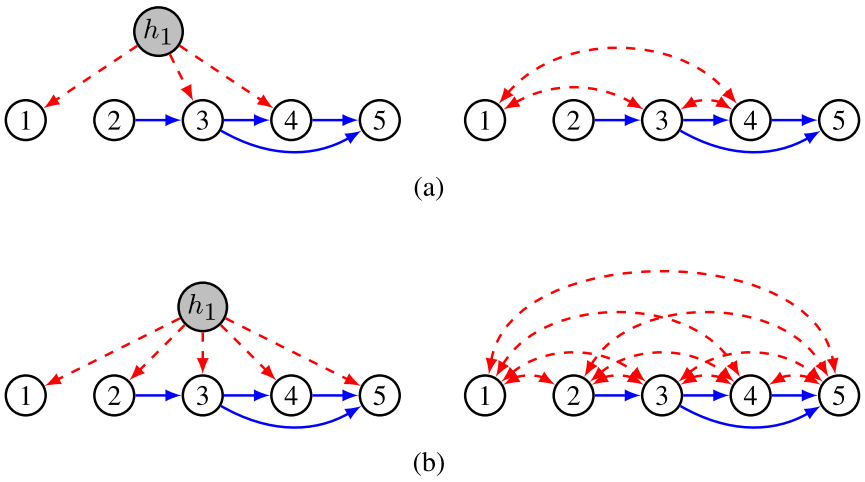


FIG. 4. *Latent-factor graphs and their latent projection.*

EXAMPLE 4.3. Figure 4 shows two latent-factor graphs and their latent projection. Both latent-factor graphs are LF-HTC-identifiable; cf. Example 3.8. But only the latent projection in the upper panel (a) is HTC-identifiable while the latent projection in panel (b) is generically infinite-to-one. The latter is easily seen since the number of model parameters corresponding to the mixed graph is larger than the dimension $\binom{d+1}{2}$ of the space $\mathrm{PD}(d)$; see, for example, Proposition 2 in Foygel, Draisma and Drton (2012).

Comparing the graphs in Figure 4, the latent-factor graphs on the left-hand side assume that all unobserved confounding is caused by a single latent factor. In contrast, for the latent projections on the right-hand side, there may be multiple latent factors that are the sources of confounding represented by bidirected edges. This leads to rational identifiability of the latent-factor graphs while the projection on the mixed graphs may be generically infinite-to-one.

Surprisingly, a mixed graph G' being rationally identifiable does *not* imply that all latent-factor graphs $G^{\mathcal{L}}$ having G' as their latent projection are rationally identifiable. Recall that in the case of rational identifiability of the latent projection there may be a proper algebraic subset A of the Zariski closure of $\mathbb{R}_{\mathrm{reg}}^{D_V} \times \mathrm{PD}(B)$ such that identification is not possible on A. If the dimensionality of the cone of latent covariance matrices $\mathrm{Im}(\tau)$ is strictly smaller than the dimension of $\mathrm{PD}(B)$, it can therefore happen that $\Theta = \mathbb{R}_{\mathrm{reg}}^{D_V} \times \mathrm{Im}(\tau) \subseteq A$ and the latent-factor graph is generically infinite-to-one. As an example, the latent projection in Figure 5 is HTC-identifiable while the latent-factor graph itself is generically infinite-to-one. In this example, $\dim(\mathrm{Im}(\tau)) = 11$ while $\dim(\mathrm{PD}(B)) = 13$. Hence, although the model given by the graph to the left is still a submodel of the one given by the graph to the right, the relevant notion of genericity is different, referring to proper subsets of $\mathrm{PD}(B)$ and of $\mathrm{Im}(\tau)$, respectively.

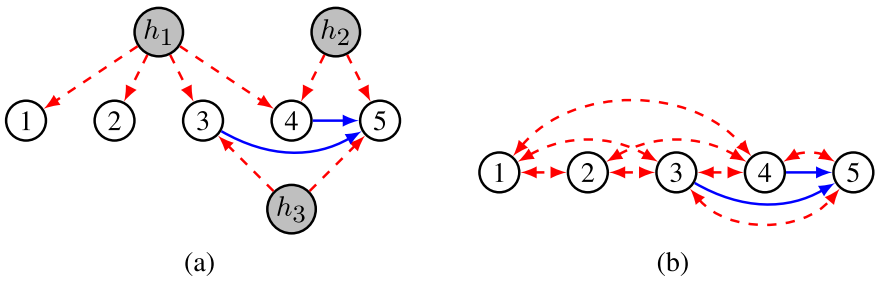


FIG. 5. *Latent-factor graph that is generically infinite-to-one but its latent projection is HTC-identifiable.*

In the experiments in Section 6, we systematically compare LF-HTC-identifiability of latent-factor graphs with HTC-identifiability applied to the corresponding latent projection.

5. Computation. In this section, we propose an efficient algorithm for deciding whether a latent-factor graph is LF-HTC-identifiable. It is similar to the algorithm of the original half-trek criterion in Foygel, Draisma and Drton (2012) and makes use of maximum flows in a special flow graph $G_{\text{flow}} = (V_f, D_f)$ from a designated source node $s \subseteq V_f$ to a target node $t \subseteq V_f$. The standard maximum-flow framework is introduced in Cormen et al. (2009). We highlight that the maximum flow can be computed in polynomial time and the complexity is $\mathcal{O}((|V_f| + r)^3)$ where $r \leq |D_f|/2$ is the number of reciprocal edge pairs in D_f. A reciprocal edge pair is a pair $v \to w$ and $w \to v$ for distinct nodes $v \neq w \in V_f$.

Let $G^{\mathcal{L}}$ be a latent-factor graph, and fix a node $v \in V$. Then we denote by $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ the decision problem whether there exists a triple $(Y, Z, H) \in 2^{V\setminus\{v\}} \times 2^{V\setminus\{v\}} \times 2^{\mathcal{L}}$ satisfying the LF-HTC for $v \in V$ in $G^{\mathcal{L}}$. To solve this problem, we first address a subproblem by assuming that we are given a fixed set $H \subseteq \mathcal{L}$ and a fixed set $Z \subseteq \text{ch}(H) \setminus (\{v\} \cup \text{pa}_V(v))$ such that $|Z| = |H|$. Since the second part of condition (ii) of the LF-HTC is equivalent to $Y \cap \text{ch}(\text{pa}_{\mathcal{L}}(Z \cup \{v\}) \setminus H) = \varnothing$, the set $A = V \setminus (Z \cup \{v\} \cup \text{ch}(\text{pa}_{\mathcal{L}}(Z \cup \{v\}) \setminus H))$ is the set of "allowed" nodes that may contain a set $Y \subseteq A$ such that (Y, Z, H) satisfies the LF-HTC with respect to v. We are able to prove the existence or inexistence of such a set Y efficiently by one maximum flow computation on a suitable flow graph $G_{\text{flow}}(v, A, Z) = (V_f, D_f)$.

The flow graph is defined as follows: Let V' and $\mathcal{L}'$ be copies of the sets V and $\mathcal{L}$. Then the graph contains the nodes $V_f = (A \cup \mathcal{L}) \cup (V' \cup \mathcal{L}') \cup \{s, t\}$, where s is a source node and t is a sink node. The set of edges D_f contains:

(a) $s \to a$ for all $a \in A$,
(b) $a \to w$ if $a \in A$ and $w \to a \in D_{\mathcal{L}V}$,
(c) $w \to w'$ for all $w \in A \cup \mathcal{L}$,
(d) $u' \to w'$ for all $u \to w \in D_{\mathcal{L}V}$ and for all $u \to w \in D_V$ such that $w \notin Z$,
(e) $w' \to t$ for all $w \in \text{pa}_V(v) \cup Z$.

We assign to all edges capacity ∞. The source node s and the target node t have capacity ∞ while all other nodes have capacity 1. Note that, by construction, no flow in $G_{\text{flow}}(v, A, Z)$ can exceed $|\text{pa}_V(v)| + |Z|$ in size, therefore one may replace the infinite capacities with $|\text{pa}_V(v)| + |Z|$ in practice. An example of a flow graph is shown in Figure 6(b).

Let $\texttt{MaxFlow}(G_{\text{flow}}(v, A, Z))$ be the maximum flow from s to t in $G_{\text{flow}}(v, A, Z)$. The following theorem is proven in Appendix A in the Supplementary Material (Barber et al. (2022)).

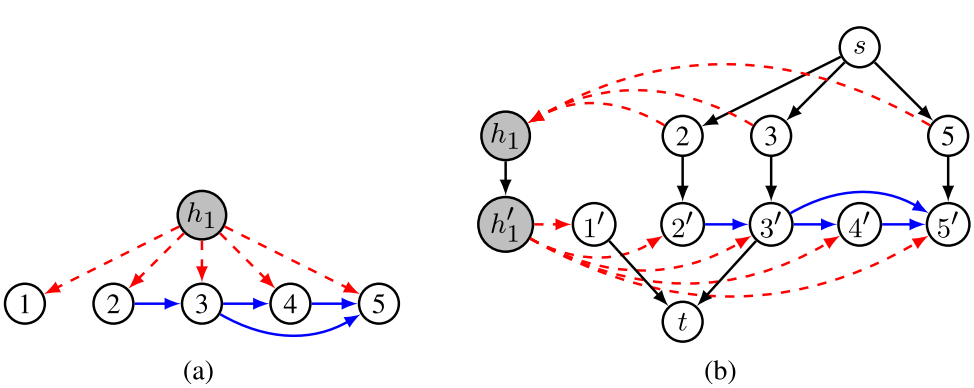


FIG. 6. *Using maximum-flow to find a set $Y \subseteq A$ such that the triple (Y, Z, H) with fixed sets $H = \{h_1\}$ and $Z = \{1\}$ satisfies the LF-HTC with respect to $v = 4$. The set of allowed nodes is $A = \{2, 3, 5\}$. (a) The concerned latent-factor graph. (b) The corresponding flow graph $G_{flow}(v, A, Z)$.*

THEOREM 5.1. *Let $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ be a latent-factor graph, and fix a node $v \in V$, a set $H \subseteq \mathcal{L}$ and a set $Z \subseteq \mathrm{ch}(H) \setminus (\{v\} \cup \mathrm{pa}_V(v))$ such that $|Z| = |H|$. For the set of allowed nodes $A = V \setminus (Z \cup \{v\} \cup \mathrm{ch}(\mathrm{pa}_{\mathcal{L}}(Z \cup \{v\}) \setminus H))$, we have that $\mathtt{MaxFlow}(G_{flow}(v, A, Z)) = |\mathrm{pa}_V(v)| + |Z|$ if and only if there exists $Y \subseteq A$ such that the triple (Y, Z, H) satisfies the LF-HTC for $v \in V$.*

For solving the decision problem $\mathtt{LF\text{-}HTC}(G^{\mathcal{L}}, v)$ we iterate over all suitable sets $H \subseteq \mathcal{L}$ and $Z \subseteq \mathrm{ch}(H) \setminus (\{v\} \cup \mathrm{pa}_V(v))$ such that $|Z| = |H|$ and check for each pair (Z, H) if there is a corresponding set $Y \subseteq A$. In each iteration, we have to compute one maximum flow by Theorem 5.1. It is enough to iterate over subsets $H \subseteq \mathcal{L}_{\geq 4}$ where $\mathcal{L}_{\geq 4} = \{h \in \mathcal{L} : |\mathrm{ch}(h)| \geq 4\}$ contains only those latent nodes with more than four children. Recall that the children of a node $v \in V \cup \mathcal{L}$ are formally defined as $\mathrm{ch}(v) = \{w \in V \cup \mathcal{L} : v \to w \in D\}$. We prove the following fact in Appendix A in the Supplementary Material.

PROPOSITION 5.2. *Let $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ be a latent-factor graph, and fix a node $v \in V$. If the triple (Y, Z, H) satisfies the LF-HTC for $v \in V$ and there is a node $h \in H$ such that $|\mathrm{ch}(h)| \leq 3$, then there are subsets $\widetilde{Y} \subseteq Y$ and $\widetilde{Z} \subseteq Z$ such that the triple $(\widetilde{Y}, \widetilde{Z}, \widetilde{H})$ with $\widetilde{H} = H \setminus \{h\}$ satisfies the LF-HTC for $v \in V$ as well.*

Next, we give an algorithm to determine whether a graph $G^{\mathcal{L}}$ is LF-HTC-identifiable by iterating over all nodes $v \in V$ and solving $\mathtt{LF\text{-}HTC}(G^{\mathcal{L}}, v)$ in each step. Moreover, when solving $\mathtt{LF\text{-}HTC}(G^{\mathcal{L}}, v)$ for a specific node $v \in V$, we have to make sure that, for a possible solution (Y, Z, H), each node $w \in Z \cup (Y \cap \mathrm{htr}_H(Z \cup \{v\}))$ was solved before. This intuition is formalized in Algorithm 1. In Theorem 5.3, we prove that the algorithm correctly determines LF-HTC-identifiability. Our implementation of Algorithm 1 is included in the `R` package `SEMID` as of version 0.4.0 (R Core Team (2020); Foygel Barber et al. (2022)), which is available on `CRAN`, the Comprehensive R Archive Network.

Algorithm 1 Testing LF-HTC-identifiability of a latent-factor graph

Require: Latent-factor graph $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$.
Ensure: Solved nodes $S \leftarrow \{v \in V : \mathrm{pa}_V(v) = \varnothing\}$.

repeat
  for $v \in V \setminus S$ **do**
    for $H \in \mathcal{L}_{\geq 4}$ **do**
      $Z_a \leftarrow (S \cap \mathrm{ch}(H)) \setminus (\{v\} \cup \mathrm{pa}_V(v))$.
      for $Z \subseteq Z_a$ such that $|Z| = |H|$ **do**
        $A \leftarrow V \setminus (Z \cup \{v\} \cup \mathrm{ch}(\mathrm{pa}_{\mathcal{L}}(Z \cup \{v\}) \setminus H) \cup (\mathrm{htr}_H(Z \cup \{v\}) \setminus S))$.
        if $\mathtt{MaxFlow}(G_{\mathrm{flow}}(v, A, Z)) = |\mathrm{pa}_V(v)| + |Z|$ **then**
          $S \leftarrow S \cup \{v\}$
          break
        end if
      end for
      if $v \in S$ **then**
        break
      end if
    end for
  end for
until $S = V$ or no change has occurred in the last iteration.
return "yes" if $S = V$, "no" otherwise.

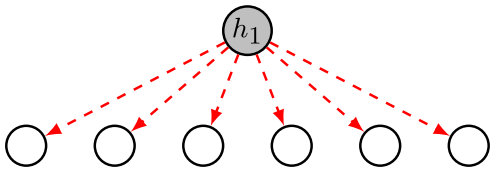


FIG. 7. *Latent structure of unlabeled latent-factor graph with one global latent factor.*

THEOREM 5.3. *A latent-factor graph $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ is LF-HTC-identifiable if and only if Algorithm 1 returns "yes." If we only allow sets H with $|H| \leq k$ in line 3, then the algorithm has complexity at most $\mathcal{O}(|V|^{2+k}|\mathcal{L}|^k(|V| + |\mathcal{L}| + r)^3)$ where $r \leq |D_V|/2$ is the number of reciprocal edge pairs in D_V.*

In Algorithm 1, we iterate over subsets of the power sets of $\mathcal{L}$ and V, and we put effort into iterating over a small subset. Nevertheless, if we allow the cardinality of $|H|$ to be unbounded in line three, then we search over an exponentially large space and, thus, our algorithm will in general take exponential time $\mathcal{O}(2^{|\mathcal{L}|+|V|})$. In fact, there is a fundamental barrier in finding a polynomial time algorithm as we are able to show that $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ is an NP-complete problem.

To see that $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ is NP-complete, first note that $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ is in the NP-complexity class due to Theorem 5.1. Every candidate triple (Y, Z, H) to solve $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ can be checked to be a solution in polynomial time by first checking if (Y, Z, H) satisfies conditions (i) and (ii) of the LF-HTC and then checking if $\texttt{MaxFlow}(G_{\text{flow}}(v, Y, Z)) = |\text{pa}_V(v)| + |Z|$. Moreover, we are able to show NP-hardness of $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ by a reduction from the Boolean satisfiability problem in conjunctive normal form; this result is developed in Appendix B in the Supplementary Material (Barber et al. (2022)).

6. Numerical experiments. This section reports on the results of experiments with small latent-factor graphs, for which the identification problem can be fully solved by techniques from computational algebraic geometry, as we discuss in Appendix C in the Supplementary Material (Barber et al. (2022)). We study acyclic latent-factor graphs with $|V| = 6$ observed nodes.

In the first experimental setup we consider one global latent factor that has an effect on all observed variables, as illustrated in Figure 7. All possible DAGs on 6 nodes are considered for the observed part (V, D_V). Table 1 lists the counts when there are $|D_V| \leq 9$ edges in

TABLE 1
Counts of unlabeled DAGs with $|V| = 6$ observed nodes and one latent node as in Figure 7

No. of obs. edges $\lvert D_V \rvert$	Total	Generically finite-to-one	Rationally identifiable	LF-HTC-identifiable
0	1	1	1	1
1	1	1	1	1
2	4	4	4	4
3	13	13	13	13
4	51	51	51	50
5	163	160	159	134
6	407	401	398	250
7	796	770	747	234
8	1169	1047	956	64
9	1291	896	631	4
Total	3896	3344	2961	755

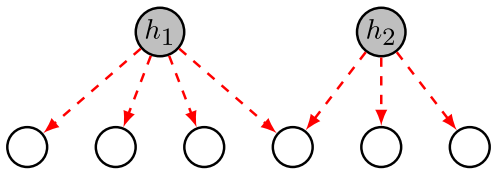


FIG. 8. *Latent structure of unlabeled latent-factor graphs with two latent factors.*

the observed part of the graph. Graphs with $|D_V| > 9$ are trivially generically infinite-to-one by Corollary 2.7. In the counts in Table 1, we treat graphs as unlabeled, that is, we count isomorphism classes of graphs. Formally, two latent-factor graphs $G = (V \cup \mathcal{L}, D)$ and $G' = (V \cup \mathcal{L}, D')$ with the same set of nodes are isomorphic if there is a permutation π of the observed nodes V such that for two nodes $h \in \mathcal{L}$ and $v \in V$ the edge $h \to v \in D$ if and only if $h \to \pi(v) \in D'$ and for two nodes $v, w \in V$ the edge $v \to w \in D$ if and only if $\pi(v) \to \pi(w) \in D'$.

In the second setup, we consider two latent factors, each of them only having influence on some of the observed variables. The precise latent structure is illustrated in Figure 8. Since the number of isomorphism classes is much larger in this case, for computational reasons we only consider graphs with at most $|D_V| = 6$ edges between observed nodes. Up to this constraint, the observed part may be any DAG. Table 2 lists the counts for these graphs, again up to isomorphism. In this setup it is possible that the latent projection is rationally identifiable. Thus, we compare the LF-HTC with the original HTC applied to the projection and the results are counted in an additional column.

In the considered setups, we see that the latent factor-criterion is very successful in certifying the graphs to be rationally identifiable as long as the number of observed edges $|D_V|$ is not too large. It misses more graphs the larger the number of observed edges is. Moreover, in the second setup, the latent-factor half-trek criterion declares about four times more graphs to be rationally identifiable than the original half-trek criterion applied to the latent projection.

7. Proof of main result.

In this section, we prove the main theorem.

PROOF OF THEOREM 3.7. Let $\mathrm{pa}_V(v) = \{p_1, \ldots, p_n\}$, $H \subseteq \mathcal{L}$ with $|H| = r$, $Y = \{y_1, \ldots, y_{n+r}\}$, and $Z = \{z_1, \ldots, z_r\}$ be as in the statement of the theorem. Define matrices

TABLE 2
Counts of unlabeled DAGs with $|V| = 6$ observed nodes and two latent nodes as in Figure 8

No. of obs. edges $\lvert D_V \rvert$	Total	Generically finite-to-one	Rationally identifiable	LF-HTC-identifiable	HTC-identifiable
0	1	1	1	1	1
1	8	6	6	6	4
2	63	45	45	43	24
3	391	255	255	236	104
4	1983	1171	1171	1018	384
5	7570	3907	3898	3028	900
6	21,029	9080	8960	5861	1157
Total	31,045	14,465	14,336	10,193	2574

$A \in \mathbb{R}^{(n+r)\times n}$, $B \in \mathbb{R}^{(n+r)\times r}$ and a vector $c \in \mathbb{R}^{n+r}$ as follows:

$$A_{ij} = \begin{cases} \left[(I_d - \Lambda)^\top \Sigma\right]_{y_i p_j} & \text{if } y_i \in \operatorname{htr}_H(Z \cup \{v\}), \\ \Sigma_{y_i p_j} & \text{if } y_i \notin \operatorname{htr}_H(Z \cup \{v\}), \end{cases}$$

and

$$B_{ij} = \begin{cases} \left[(I_d - \Lambda)^\top \Sigma (I_d - \Lambda)\right]_{y_i z_j} & \text{if } y_i \in \operatorname{htr}_H(Z \cup \{v\}), \\ \left[\Sigma (I_d - \Lambda)\right]_{y_i z_j} & \text{if } y_i \notin \operatorname{htr}_H(Z \cup \{v\}), \end{cases}$$

and

$$c_i = \begin{cases} \left[(I_d - \Lambda)^\top \Sigma\right]_{y_i v} & \text{if } y_i \in \operatorname{htr}_H(Z \cup \{v\}), \\ \Sigma_{y_i v} & \text{if } y_i \notin \operatorname{htr}_H(Z \cup \{v\}). \end{cases}$$

CLAIM 1. *The matrices A and B and the vector c are all rationally identifiable.*

By assumption, all columns of Λ indexed by a vertex in $Z \cup (Y \cap \operatorname{htr}_H(Z \cup \{v\}))$ are rationally identifiable (i.e., rational functions of Σ). Inspecting the above expressions, we observe that only entries from these columns of Λ appear in the definition of A, B, and c. Hence, A, B, and c are rationally identifiable, as claimed.

Next, note that there is a set $Y_Z \subseteq Y$ such that there is a system of latent-factor half-treks with no sided intersection from Y_Z to Z. In this system each half-trek takes the form $y \leftarrow h \rightarrow z$ for $y \in Y$, $z \in Z$ and $h \in H$. Since the system has no sided intersection, it follows from Proposition 3.4 in Sullivant, Talaska and Draisma (2010) that $\det(\Omega_{Y_Z,Z}) \neq 0$, generically. Thus, the matrix $\Omega_{Y,Z}$ has full column rank r because $\Omega_{Y_Z,Z}$ is a submatrix. Using this fact we prove our next claim.

CLAIM 2. *There exists some $\psi \in \mathbb{R}^r$ such that*

$$\begin{pmatrix} A & B \end{pmatrix} \cdot \begin{pmatrix} \Lambda_{\operatorname{pa}_V(v),v} \\ \psi \end{pmatrix} = c.$$

To show this, we implicitly construct ψ. Let $\Omega_h = \Gamma_h^\top \Gamma_h$ for each $h \in \mathcal{L}$, and observe that

$$\Omega_{Y,Z\cup\{v\}} = (\Omega_{\mathsf{diag}})_{Y,Z\cup\{v\}} + \sum_{h\in H} (\Omega_h)_{Y,Z\cup\{v\}} + \sum_{h\in\mathcal{L}\setminus H} (\Omega_h)_{Y,Z\cup\{v\}}.$$

Since $Y \cap (Z \cup \{v\}) = \varnothing$ by definition of the latent-factor half-trek criterion, we have that $(\Omega_{\mathsf{diag}})_{Y,Z\cup\{v\}} = 0$. The definition of the latent-factor half-trek criterion further yields that for any $h \in \mathcal{L} \setminus H$, either $Y \cap \operatorname{ch}(h) = \varnothing$ or $(Z \cup \{v\}) \cap \operatorname{ch}(h) = \varnothing$. Hence, $(\Omega_h)_{Y,Z\cup\{v\}} = 0$. We obtain that

$$\Omega_{Y,Z\cup\{v\}} = \sum_{h\in H} (\Omega_h)_{Y,Z\cup\{v\}} = (\Omega_H)_{Y,Z\cup\{v\}},$$

where $\Omega_H := \sum_{h\in H} \Omega_h$. Note that $\operatorname{rank}(\Omega_H) \leq |H| = r$. Moreover, $(\Omega_H)_{V,Z}$ has full column rank r by assumption (since $\Omega_{Y,Z}$ is a submatrix of this matrix), which proves that

$$(\Omega_H)_{V,Z} \cdot \psi = (\Omega_H)_{V,v} \tag{7.1}$$

for some $\psi \in \mathbb{R}^r$.

Next, consider any index i such that $y_i \in \mathrm{htr}_H(Z \cup \{v\})$. Then

$$\begin{aligned}
&\left[\begin{pmatrix} A & B \end{pmatrix} \cdot \begin{pmatrix} \Lambda_{\mathrm{pa}_V(v),v} \\ \psi \end{pmatrix}\right]_i \\
&\quad = \left[(I_d - \Lambda)^\top \Sigma\right]_{y_i,\mathrm{pa}_V(v)} \cdot \Lambda_{\mathrm{pa}_V(v),v} + \left[(I_d - \Lambda)^\top \Sigma (I_d - \Lambda)\right]_{y_i,Z} \cdot \psi \\
&\quad = \left[(I_d - \Lambda)^\top \Sigma \cdot \Lambda\right]_{y_i v} + [\Omega_{Y,Z} \cdot \psi]_i \qquad (7.2)
\end{aligned}$$

because $\Lambda_{wv} = 0$ unless $w \in \mathrm{pa}_V(v)$ and $(I_d - \Lambda)^\top \Sigma (I_d - \Lambda) = \Omega$. Since $\Omega_{Y,Z\cup\{v\}} = (\Omega_H)_{Y,Z\cup\{v\}}$, it follows from (7.1) that

$$[\Omega_{Y,Z} \cdot \psi]_i = [\Omega_{Y,v}]_i = \Omega_{y_i v}.$$

Hence, we may rewrite (7.2) as

$$\begin{aligned}
\left[\begin{pmatrix} A & B \end{pmatrix} \cdot \begin{pmatrix} \Lambda_{\mathrm{pa}_V(v),v} \\ \psi \end{pmatrix}\right]_i &= \left[(I_d - \Lambda)^\top \Sigma\right]_{y_i v} - \left[(I_d - \Lambda)^\top \Sigma (I_d - \Lambda)\right]_{y_i v} + \Omega_{y_i v} \\
&= \left[(I_d - \Lambda)^\top \Sigma\right]_{y_i v} - \Omega_{y_i v} + \Omega_{y_i v} \\
&= c_i,
\end{aligned}$$

by the definition of c.

To conclude the proof of Claim 2, consider any index i such that $y_i \notin \mathrm{htr}_H(Z \cup \{v\})$. For any such i, any latent-factor half-trek from a node $w \in Z \cup \{v\}$ to y_i must be of the form

$$w \leftarrow h \rightarrow x_1 \rightarrow \cdots \rightarrow x_m \rightarrow y_i$$

for some $h \in H$. This implies that

$$\left[\Omega (I_d - \Lambda)^{-1}\right]_{w y_i} = \left[\Omega_H (I_d - \Lambda)^{-1}\right]_{w y_i} \qquad (7.3)$$

for all $w \in Z \cup \{v\}$. Consequently,

$$\begin{aligned}
\left[\begin{pmatrix} A & B \end{pmatrix} \cdot \begin{pmatrix} \Lambda_{\mathrm{pa}_V(v),v} \\ \psi \end{pmatrix}\right]_i &= \Sigma_{y_i,\mathrm{pa}(v)} \cdot \Lambda_{\mathrm{pa}_V(v),v} + \left[\Sigma (I_d - \Lambda)\right]_{y_i,Z} \cdot \psi \\
&= [\Sigma \Lambda]_{y_i v} + \left[\Sigma (I_d - \Lambda)\right]_{y_i,Z} \cdot \psi \\
&= \Sigma_{y_i v} - \left[\Sigma (I_d - \Lambda)\right]_{y_i v} + \left[\Sigma (I_d - \Lambda)\right]_{y_i,Z} \cdot \psi \\
&= \Sigma_{y_i v} - \left[(I_d - \Lambda)^{-\top} \Omega\right]_{y_i v} + \left[(I_d - \Lambda)^{-\top} \Omega\right]_{y_i,Z} \cdot \psi, \qquad (7.4)
\end{aligned}$$

because $\Omega = (I_d - \Lambda)^\top \Sigma (I_d - \Lambda)$. Applying first (7.3) and then (7.1), we find that

$$\begin{aligned}
&- \left[(I_d - \Lambda)^{-\top} \Omega\right]_{y_i v} + \left[(I_d - \Lambda)^{-\top} \Omega\right]_{y_i,Z} \cdot \psi \\
&\quad = -\left[(I_d - \Lambda)^{-\top} \Omega_H\right]_{y_i v} + \left[(I_d - \Lambda)^{-\top} \Omega_H\right]_{y_i,Z} \cdot \psi \\
&\quad = -\left[(I_d - \Lambda)^{-\top} \Omega_H\right]_{y_i v} + \left[(I_d - \Lambda)^{-\top} \Omega_H\right]_{y_i v} = 0.
\end{aligned}$$

Taking up (7.4) and recalling the definition of c, we conclude that

$$\left[\begin{pmatrix} A & B \end{pmatrix} \cdot \begin{pmatrix} \Lambda_{\mathrm{pa}_V(v),v} \\ \psi \end{pmatrix}\right]_i = \Sigma_{y_i v} = c_i.$$

The theorem is now proven if the equation system exhibited in Claim 2 has a unique solution generically. This is addressed by our last claim.

CLAIM 3. *The matrix $(A\ B)$ is generically invertible.*

To prove Claim 3, we will show that if we set some parameters equal to zero, then the considered matrix is invertible for generic choices of the remaining free parameters, which is sufficient to show that the matrix will be generically invertible with respect to choices of all parameters.

By assumption, the latent-factor graph $G^{\mathcal{L}}$ contains a system of latent-factor half-treks from Y to $Z \cup \mathrm{pa}_V(v)$, where half-treks terminating at any $z \in Z$ are of the form $y_i \leftarrow h \rightarrow z$ for some $h \in H$. For every $z \in Z$, set $\Lambda_{\mathrm{pa}_V(z),z} = 0$. Furthermore, every node $h \in H$ appears in at most one of the latent-factor half-treks in the system. Suppose it appears as $y_i \leftarrow h \rightarrow w$. Then we will define Ω_h to have value $\omega_{y_i w}$ at entries $\{y_i, w\} \times \{y_i, w\}$, and zeros elsewhere.

Consider now a mixed graph $\widehat{G}$ constructed as follows. Starting with the induced subgraph $\widehat{G} = (V, D_V)$, first remove all edges with head in Z. Next, looking at the selected system of latent-factor half-treks from Y to $Z \cup \mathrm{pa}_V(v)$ in the latent-factor graph $G^{\mathcal{L}}$, any time we see a half-trek beginning with $y_i \leftarrow h \rightarrow w$, add a bidirected edge $y_i \leftrightarrow w$ to $\widehat{G}$.

By definition of the new graph $\widehat{G}$, the selected system of latent-factor half-treks from Y to $Z \cup \mathrm{pa}_V(v)$ in $G^{\mathcal{L}}$ has a corresponding system of half-treks in $\widehat{G}$. Here, any latent-factor half-trek that begins with edges $y_i \leftarrow h \rightarrow w$ has these two initial two edges replaced by the bidirected edge $y_i \leftrightarrow w$. The resulting system of half-treks in $\widehat{G}$ has no sided intersection. Let $\widehat{\Lambda}$ and $\widehat{\Omega}$ be the parameter matrices for this graph. Note that $(I - \widehat{\Lambda})_{*,Z} = I_{*,Z}$ because $\widehat{\Lambda}_{*,Z} = 0$ by construction. Therefore, we can write

$$B_{ij} = \begin{cases} \left[(I_d - \Lambda)^\top \Sigma\right]_{y_i z_j} & \text{if } y_i \in \mathrm{htr}_H\big(Z \cup \{v\}\big), \\ \Sigma_{y_i z_j} & \text{if } y_i \notin \mathrm{htr}_H\big(Z \cup \{v\}\big). \end{cases}$$

We now apply Lemma 2 in the original half-trek paper (Foygel, Draisma and Drton (2012)) to conclude that $(A\ B)$ is generically invertible. □

8. Discussion. In this work, we proposed a graphical criterion that provides an effective sufficient condition for rational identifiability in linear structural equation models where latent variables are not projected to correlation among noise terms. To the best of our knowledge, it is the most general graphical criterion to decide identifiability for graphs explicitly including latent nodes. The new criterion can be checked in time that is polynomial in the size of the graph if we search only over subsets of latent nodes of bounded size. The restriction of the search space is necessary since checking the criterion without any restriction is in general NP-hard.

The criterion applies to a wide range of models and allows for presence of multiple latent factors that may even have an effect on many or all of the observed variables. The corresponding directed graph is allowed to be cyclic, the only restriction that we made in this work is that all latent factors are source nodes in the graph.

It is noteworthy that even if a model is not LF-HTC-identifiable, the latent-factor half-trek method can still prove certain columns of Λ to be identifiable. This is the case if the recursive procedure of Algorithm 1 stops early declaring some but not all nodes to satisfy the LF-HTC. In this case, the status of identifiability of the whole graph remains inconclusive but for the nodes v that the method successfully visits, the parameters $\Lambda_{\mathrm{pa}_V(v),v}$ are proven to be rationally identifiable.

Methods for identifiability of latent-factor graphs are useful also as a refinement of methods that operate on mixed graphs in the latent projection framework: Imagine a model that is generically infinite-to-one in the latent projection framework. The main reason for this is often denser confounding, that is, there is confounding between many of the observed variables. There is then the natural question whether the model would be (rationally) identifiable if the confounding originated from a simpler structure, that is, is caused by only a few latent factors. Then the LF-HTC may be applicable and may prove a model rationally identifiable.

On the other hand, if a model is rationally identifiable in the latent projection framework, then the identifiability may be due to the assumption that confounding is caused by multiple different latent factors. As shown in Figure 5, there may be settings where rational identifiability no longer holds when the confounding is in fact caused by fewer factors. Using our method it is possible to check for such identifiability failures.

We would like to emphasize that the LF-HTC is useful also if the goal is model selection. One may then be interested in testing the goodness-of-fit of a particular model, a problem for which it is crucial to know the dimension of the model. The LF-HTC asserting identifiability also means that the model has expected dimension.

An interesting research program emerges from the work presented here. Indeed, one may strive to improve and extend the efficiency of the LF-HTC along similar lines as those that have been applied in previous work that has led to improvements of the original half-trek criterion for mixed graphs. In particular, it would be useful to find a latent-factor modification of the criterion for edgewise identifiability that allows for identification of a subset or even single direct causal effects λ_{wv} instead of only targeting whole columns $\Lambda_{\mathrm{pa}_V(v),v}$; compare to Weihs et al. (2018) and references therein. This extension is of interest when effects between particular variables are the primary targets of investigation, but it may also make the criterion more powerful as a whole. Another way to extend the scope of the LF-HTC would be to apply graph decomposition techniques as proposed by Tian (2005); see also Foygel, Draisma and Drton (2012) and Drton (2018), Section 6.

Furthermore, it would be interesting to generalize the LF-HTC to a version in which we relax the condition that all latent factors are source nodes in the graph. For example, one may consider models where latent nodes are only required to be *upstream*, that is, there may be direct causal effects between latent variables but no effects from observed variables to latent variables. Put differently, in addition to the equation system (1.1) that defines the model, the vector of latent variables $(L_h)_{h\in\mathcal{L}}$ is required to satisfy the equation

$$L = B^T L + \delta,$$

where B is an $\ell \times \ell$ matrix with zeros along the diagonal and the noise terms $\delta = (\delta_h)_{h\in\mathcal{L}}$ are independent with mean zero and variance 1. The latent covariance matrix is now of the form

$$\Omega = \Omega_{\mathsf{diag}} + \Gamma^\top (I_\ell - B)^{-\top} (I_\ell - B)^{-1} \Gamma.$$

Thus the parametrization τ of the cone of latent covariance matrices is rational and depends on the three parameter matrices $(B, \Gamma, \Omega_{\mathsf{diag}})$. The question is how to identify effects between observed variables in this case, or, even more, what can be said in terms of identifying causal effects between latent variables. Note that such a setting cannot be handled by a mixed graph approach which marginalizes out the effects of interest. Hence, our work sets the scene for future developments of identifiability between latent variables.

In Lemma 2.5, we gave a simple necessary condition for the parametrization map to be generically finite-to-one. In future work, we hope to obtain more powerful necessary conditions for generic identifiability in the form of efficient graphical criteria. This will amount to studying the Jacobian matrix of the parametrization $\varphi_{G\mathcal{L}}$, taking into account the algebraic geometry of the cone of latent covariance matrices.

Acknowledgments. Mathias Drton and Nils Sturma are affiliated with the Munich Data Science Institute (MDSI) and the Munich Center for Machine Learning (MCML).

Funding. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 883818).

Nils Sturma acknowledges support by the MDSI via the Linde/MDSI PhD Fellowship program.

Rina Foygel Barber was supported by the U.S. National Science Foundation via grants DMS-1654076 and DMS-2023109, and by the Office of Naval Research via grant N00014-20-1-2337.

SUPPLEMENTARY MATERIAL

Supplement to "Half-trek criterion for identifiability of latent variable models" (DOI: 10.1214/22-AOS2221SUPP; .pdf). The supplement contains additional material such as further elements of proofs, a hardness result for checking the LF-HTC without a bound on the cardinality of searched sets of latent variables, and an explanation on how to effectively deploy techniques from computational algebraic geometry.

REFERENCES

BARBER, R. F, DRTON, M., STURMA, N. and WEIHS, L. (2022). Supplement to "Half-trek criterion for identifiability of latent variable models." https://doi.org/10.1214/22-AOS2221SUPP

BASU, S., POLLACK, R. and ROY, M.-F. (2006). *Algorithms in Real Algebraic Geometry*, 2nd ed. *Algorithms and Computation in Mathematics* **10**. Springer, Berlin. MR2248869

BENEDETTI, R. and RISLER, J.-J. (1990). *Real Algebraic and Semi-Algebraic Sets*. *Actualités Mathématiques*. [*Current Mathematical Topics*]. Hermann, Paris. MR1070358

BOCHNAK, J., COSTE, M. and ROY, M.-F. (1998). *Real Algebraic Geometry*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **36**. Springer, Berlin. Translated from the 1987 French original, Revised by the authors. MR1659509 https://doi.org/10.1007/978-3-662-03718-8

BOLLEN, K. A. (1989). *Structural Equations with Latent Variables*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Wiley, New York. MR0996025 https://doi.org/10.1002/9781118619179

BOWDEN, R. J. and TURKINGTON, D. A. (1984). *Instrumental Variables*. *Econometric Society Monographs in Quantitative Economics* **8**. Cambridge Univ. Press, Cambridge. MR0798790

BRITO, C. and PEARL, J. (2002). Generalized instrumental variables. In *Proceedings of the* 18*th Conference on Uncertainty in Artificial Intelligence* (*UAI*). *UAI'*02 85–93. Morgan Kaufmann Publishers Inc., San Francisco, CA.

BRITO, C. and PEARL, J. (2006). Graphical condition for identification in recursive SEM. In *Proceedings of the* 22*nd Conference on Uncertainty in Artificial Intelligence* (*UAI*) 47–54. AUAI Press, Arlington, VA.

CHEN, B., KUMOR, D. and BAREINBOIM, E. (2017). Identification and model testing in linear structural equation models using auxiliary variables. In *Proceedings of the* 34*th International Conference on Machine Learning* (*ICML*). *ICML'*17 757–766. JMLR, USA.

CHEN, B., PEARL, J. and BAREINBOIM, E. (2016). Incorporating knowledge into structural equation models using auxiliary variables. In *Proceedings of the* 25*th International Joint Conference on Artificial Intelligence* (*IJCAI*). *IJCAI'*16 3577–3583. AAAI Press, Menlo Park, CA.

CORMEN, T. H., LEISERSON, C. E., RIVEST, R. L. and STEIN, C. (2009). *Introduction to Algorithms*, 3rd ed. MIT Press, Cambridge, MA. MR2572804

COX, D., LITTLE, J. and O'SHEA, D. (2007). *Ideals*, *Varieties*, *and Algorithms*, 3rd ed. *Undergraduate Texts in Mathematics*. Springer, New York. An introduction to computational algebraic geometry and commutative algebra. MR2290010 https://doi.org/10.1007/978-0-387-35651-8

DRTON, M. (2018). Algebraic problems in structural equation modeling. In *The* 50*th Anniversary of Gröbner Bases*. *Adv. Stud. Pure Math*. **77** 35–86. Math. Soc. Japan, Tokyo. MR3839705 https://doi.org/10.2969/aspm/07710035

DRTON, M., STURMFELS, B. and SULLIVANT, S. (2007). Algebraic factor analysis: Tetrads, pentads and beyond. *Probab. Theory Related Fields* **138** 463–493. MR2299716 https://doi.org/10.1007/s00440-006-0033-2

DRTON, M. and WEIHS, L. (2016). Generic identifiability of linear structural equation models by ancestor decomposition. *Scand. J. Stat*. **43** 1035–1045. MR3573674 https://doi.org/10.1111/sjos.12227

DRTON, M. and YU, J. (2010). On a parametrization of positive semidefinite matrices with zeros. *SIAM J. Matrix Anal. Appl.* **31** 2665–2680. MR2740626 https://doi.org/10.1137/100783170

FOYGEL, R., DRAISMA, J. and DRTON, M. (2012). Half-trek criterion for generic identifiability of linear structural equation models. *Ann. Statist.* **40** 1682–1713. MR3015040 https://doi.org/10.1214/12-AOS1012

FOYGEL BARBER, R., DRTON, M., STURMA, N. and WEIHS, L. (2022). SEMID: Identifiability of linear structural equation models. R package version 0.4.0.

GARCIA-PUENTE, L. D., SPIELVOGEL, S. and SULLIVANT, S. (2010). Identifying causal effects with computer algebra. In *Proceedings of the* 26*th Conference on Uncertainty in Artificial Intelligence* (*UAI*) AUAI Press, Arlington, VA.

GESSEL, I. and VIENNOT, G. (1985). Binomial determinants, paths, and hook length formulae. *Adv. Math.* **58** 300–321. MR0815360 https://doi.org/10.1016/0001-8708(85)90121-5

HARTSHORNE, R. (1977). *Algebraic Geometry*. *Graduate Texts in Mathematics*, *No*. 52. Springer, New York. MR0463157

KUMOR, D., CHEN, B. and BAREINBOIM, E. (2019). Efficient identification in linear structural causal models with instrumental cutsets. In *Advances in Neural Information Processing Systems* (*NeurIPS*) **32** 12477–12486. Neural Information Processing Systems Foundation. Curran Associates, Red Hook, NY.

KUMOR, D., CINELLI, C. and BAREINBOIM, E. (2020). Efficient identification in linear structural causal models with auxiliary cutsets. In *Proceedings of the* 37*th International Conference on Machine Learning* (*ICML*). *Proceedings of Machine Learning Research* **119** 5501–5510. PMLR, USA.

KUROKI, M. and PEARL, J. (2014). Measurement bias and effect restoration in causal inference. *Biometrika* **101** 423–437. MR3215357 https://doi.org/10.1093/biomet/ast066

LEE, S. and BAREINBOIM, E. (2021). Causal identification with matrix equations. In *Advances in Neural Information Processing Systems* (M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang and J. W. Vaughan, eds.) **34** 9468–9479. Curran Associates, Red Hook, NY.

LEUNG, D., DRTON, M. and HARA, H. (2016). Identifiability of directed Gaussian graphical models with one latent source. *Electron. J. Stat.* **10** 394–422. MR3466188 https://doi.org/10.1214/16-EJS1111

LINDSTRÖM, B. (1973). On the vector representations of induced matroids. *Bull. Lond. Math. Soc.* **5** 85–90. MR0335313 https://doi.org/10.1112/blms/5.1.85

MAATHUIS, M., DRTON, M., LAURITZEN, S. and WAINWRIGHT, M., eds. (2019). *Handbook of Graphical Models*. *Chapman & Hall/CRC Handbooks of Modern Statistical Methods*. CRC Press, Boca Raton, FL. MR3889064

MIAO, W., GENG, Z. and TCHETGEN TCHETGEN, E. J. (2018). Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika* **105** 987–993. MR3877879 https://doi.org/10.1093/biomet/asy038

OKAMOTO, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.* **1** 763–765. MR0331643

PEARL, J. (2009). *Causality*, 2nd ed. Cambridge Univ. Press, Cambridge. Models, reasoning, and inference. MR2548166 https://doi.org/10.1017/CBO9780511803161

PETERS, J., JANZING, D. and SCHÖLKOPF, B. (2017). *Elements of Causal Inference*. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA. Foundations and learning algorithms. MR3822088

R CORE TEAM, (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

SHAFAREVICH, I. R. (2013). *Basic Algebraic Geometry.* 1, 3rd ed. Springer, Heidelberg. MR3100243

SPIRTES, P., GLYMOUR, C. and SCHEINES, R. (2000). *Causation*, *Prediction*, *and Search*, 2nd ed. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA. With additional material by David Heckerman, Christopher Meek, Gregory F. Cooper and Thomas Richardson, A Bradford Book. MR1815675

STANGHELLINI, E. and WERMUTH, N. (2005). On the identification of path analysis models with one hidden variable. *Biometrika* **92** 337–350. MR2201363 https://doi.org/10.1093/biomet/92.2.337

SULLIVANT, S., TALASKA, K. and DRAISMA, J. (2010). Trek separation for Gaussian graphical models. *Ann. Statist.* **38** 1665–1685. MR2662356 https://doi.org/10.1214/09-AOS760

TIAN, J. (2005). Identifying direct causal effects in linear models. In *Proceedings of the* 20*th National Conference on Artificial Intelligence* (*AAAI*). *AAAI'*05 346–352. Association for the Advancement of Artificial Intelligence. AAAI Press, Menlo Park, CA.

TIAN, J. (2009). Parameter identification in a class of linear structural equation models. In *Proceedings of the* 21*st International Joint Conference on Artificial Intelligence* (*IJCAI*). *IJCAI'*09 1970–1975. AAAI Press, Menlo Park, CA.

VAN DER ZANDER, B., TEXTOR, J. and LISKIEWICZ, M. (2015). Efficiently finding conditional instruments for causal inference. In *Proceedings of the* 24*th International Joint Conference on Artificial Intelligence* (*IJCAI*). *IJCAI'*15 3243–3249. AAAI Press, Menlo Park, CA.

WEIHS, L., ROBINSON, B., DUFRESNE, E., KENKEL, J., KUBJAS, K., MCGEE, R. II, NGUYEN, N., ROBEVA, E. and DRTON, M. (2018). Determinantal generalizations of instrumental variables. *J. Causal Inference* **6** 20170009. MR4351483 https://doi.org/10.1515/jci-2017-0009

Submitted to the Annals of Statistics

SUPPLEMENT TO "HALF-TREK CRITERION FOR IDENTIFIABILITY OF LATENT VARIABLE MODELS"

BY RINA FOYGEL BARBER[1], MATHIAS DRTON[2,*], NILS STURMA[2,†] AND LUCA WEIHS[3]

[1]*Department of Statistics, University of Chicago, rina@uchicago.edu*

[2]*Technical University of Munich, Germany; TUM School of Computation, Information and Technology, Department of Mathematics; Munich Data Science Institute (MDSI), *mathias.drton@tum.de; †nils.sturma@tum.de*

[3]*Allen Institute for AI, lucaw@allenai.org*

This supplement contains additional material such as further elements of proofs (Appendix A), a hardness result for checking the LF-HTC without a bound on the cardinality of searched sets of latent variables (Appendix B), and an explanation on how to effectively deploy techniques from computational algebraic geometry (Appendix C).

APPENDIX A: PROOFS

PROOF OF LEMMA 2.5. Throughout the proof we let $p = \dim(S)$. Since f is rational, it is a semialgebraic mapping according to Definition 2.2.5 in Bochnak, Coste and Roy (1998). Images and preimages of semialgebraic sets under semialgebraic mappings are again semialgebraic. Hence, the image $f(S)$ is a semialgebraic set. The rest of the proof is an application of Hardt's triviality theorem (Basu, Pollack and Roy, 2006, Theorem 5.45) which states that there exists a finite partition of $f(S)$ into semialgebraic sets $f(S) = \bigcup_{i=1}^{r} T_i$ such that for each i and for each $y \in T_i$ the product $T_i \times f^{-1}(y)$ is semialgebraically homeomorphic to $f^{-1}(T_i)$. In particular, we have for all $y \in T_i$ the equality

$$\dim(f^{-1}(y)) = \dim(f^{-1}(T_i)) - \dim(T_i). \tag{A.1}$$

Now suppose that $k = \dim(f(S)) < p$. Observe that $S = \bigcup_{i=1}^{r} f^{-1}(T_i)$ is a finite union of semi-algebraic sets. We write C for the union of all preimages $f^{-1}(T_i)$ of dimension strictly less than p. Then for all $x \in S \setminus C$ we have by Equation (A.1)

$$\dim(f^{-1}(f(x))) \geq \dim(S) - \dim(f(S)) = p - k > 0,$$

which means that for all $x \in S \setminus C$ the fiber $f^{-1}(f(x))$ is a semialgebraic subset of S with positive dimension, i.e., it contains infinitely many elements (cf. Theorem 5.19 in Basu, Pollack and Roy (2006)). Moreover, the Zariski closure $\overline{S}$ is equal to the union of Zariski closures $\overline{S \setminus C} \cup \overline{C}$. By Proposition 2.8.5 in Bochnak, Coste and Roy (1998) the dimension of C is strictly less than p, i.e., $\overline{S} \neq \overline{C}$. Since $\overline{S}$ is irreducible, it must be the case $\overline{S} = \overline{S \setminus C}$. Thus there is no proper algebraic subset of $\overline{S}$ that contains $S \setminus C$ and we conclude that f is generically infinite-to-one.

For the other direction, suppose that $k = \dim(f(S)) = p$. Let $I = \{i \in \{1, \ldots, r\} : \dim(T_i) < p\}$ and $B = \bigcup_{i \in I} T_i$. Then the Zariski closure $\overline{B}$ in $\mathbb{R}^n$ has dimension strictly smaller than p. Applying Equation (A.1) we get for all $y \in f(S) \setminus \overline{B}$ that

$$\dim(f^{-1}(y)) \leq \dim(S) - p = p - p = 0.$$

MSC2020 subject classifications: 62H22, 62J05, 62R01.

Keywords and phrases: Covariance matrix, factor analysis, graphical model, hidden variables, latent variables, parameter identification, structural equation model.

Therefore, for all $x \in S \setminus f^{-1}(\overline{B})$ the fiber $f^{-1}(f(x))$ is a zero-dimensional and thus finite semialgebraic set (compare Theorem 5.19 in Basu, Pollack and Roy (2006) again). To finish the proof it remains to show that the Zariski closure of $f^{-1}(\overline{B})$ is a proper subset of $\overline{S}$. As $\overline{S}$ is assumed to be irreducible, it suffices to argue that $\overline{S}$ contains a point outside the Zariski closure of $f^{-1}(\overline{B})$. Using that f is rational, we see that the preimage $f^{-1}(\overline{B})$ is an algebraic subset of S. Since $\dim(\overline{B}) < p$, the set $f(S) \setminus \overline{B}$ is nonempty and therefore $S \setminus f^{-1}(\overline{B})$ is nonempty as well. Now observe that the points in $S \setminus f^{-1}(\overline{B})$ are not contained in the Zariski closure of $f^{-1}(\overline{B})$. We conclude that f is generically finite-to-one. □

PROOF OF THEOREM 5.1. The proof is similar to the proof of Theorem 6 in Foygel, Draisma and Drton (2012a). If $(Y, Z, H) \in 2^{V\setminus\{v\}} \times 2^{V\setminus\{v\}} \times 2^{\mathcal{L}}$ satisfies the LF-HTC with respect to v, then we have a system Π of latent-factor half-treks from Y to $\mathrm{pa}_V(v) \cup Z$ with no sided intersection such that for each $z \in Z$, the half-trek terminating at z takes the form $y \leftarrow h \rightarrow z$ for some $y \in Y$ and some $h \in H$.

For each latent-factor half-trek $\pi_k \in \Pi$ of the form

$$\pi_k : y_k \leftarrow h_k \rightarrow w_k \rightarrow \cdots \rightarrow k, \quad k \in \mathrm{pa}_V(v) \cup Z,$$

add a flow of size 1 along the path

$$\widetilde{\pi}_k : s \rightarrow y_k \rightarrow h_k \rightarrow h'_k \rightarrow w'_k \rightarrow \cdots \rightarrow k' \rightarrow t$$

in the flow graph G_{flow}. Similarly, for each latent-factor half-trek $\pi_k \in \Pi$ of the form

$$\pi_k : y_k \rightarrow w_k \rightarrow \cdots \rightarrow k, \quad k \in \mathrm{pa}_V(v) \cup Z,$$

add a flow of size 1 along the path

$$\widetilde{\pi}_k : s \rightarrow y_k \rightarrow y'_k \rightarrow w'_k \rightarrow \cdots \rightarrow k' \rightarrow t$$

in the flow graph G_{flow}. Let $\widetilde{\Pi} = \{\widetilde{\pi}_k : k \in \mathrm{pa}_V(v) \cup Z\}$ be the system of directed paths that we obtain in the flow graph G_{flow}. Clearly, the total flow size from s to t in the flow graph is $|\mathrm{pa}_V(v)| + |Z|$. It is left to check that no capacity constraint is exceeded. This is trivial for the infinite edge capacities as well as for the infinite capacities of the nodes s and t. For all other nodes that appear in some of the paths of the system $\widetilde{\Pi}$, note that they appear exactly once in the system since the original system of latent-factor half-treks Π has no sided intersection.

Now suppose $\texttt{MaxFlow}(G_{\text{flow}}(v, A, Z)) = |\mathrm{pa}_V(v)| + |Z|$. By the properties of the max-flow problem with integer-valued capacities (Ford and Fulkerson, 1962), this means that there are $|\mathrm{pa}_V(v)| + |Z|$ directed paths from s to t with flow size 1 along each path. We denote the collection of these paths by $\widetilde{\Pi} = \{\widetilde{\pi}_k : k \in \mathrm{pa}_V(v) \cup Z\}$, recall that by assumption $Z \cap \mathrm{pa}_V(v) = \emptyset$. Since all nodes in the flow graph that are not equal to s or t have capacity 1, each node different from s and t can appear at most once in the system of paths $\widetilde{\Pi}$. Consider a specific path $\widetilde{\pi}_k \in \widetilde{\Pi}$. By construction of the graph G_{flow}, it has one of two forms. First, we may have

$$\widetilde{\pi}_k : s \rightarrow y_k \rightarrow h_k \rightarrow h'_k \rightarrow w'_k \rightarrow \cdots \rightarrow k' \rightarrow t$$

with $y_k \in A$, $k \in \mathrm{pa}_V(v) \cup Z$ and $h_k \in \mathcal{L}$. This defines the latent-factor half-trek

$$\pi_k : y_k \leftarrow h_k \rightarrow w_k \rightarrow \cdots \rightarrow k$$

in $G^{\mathcal{L}}$. The other possibility is that the path has the form

$$\widetilde{\pi}_k : s \rightarrow y_k \rightarrow y'_k \rightarrow w'_k \rightarrow \cdots \rightarrow k' \rightarrow t$$

with $y_k \in A$ and $k \in \mathrm{pa}_V(v) \cup Z$. This defines the latent-factor half-trek

$$\pi_k : y_k \rightarrow w_k \rightarrow \cdots \rightarrow k$$

in $G^{\mathcal{L}}$. In this way, we obtain a system of latent-factor half-treks $\Pi = \{\pi_k : k \in \mathrm{pa}_V(v) \cup Z\}$ in $G^{\mathcal{L}}$. Because each node other than s or t appears at most once in the system $\widetilde{\Pi}$ in G_{flow}, the constructed system Π has no sided intersection. Furthermore, if $k \in Z$, we have that $w_k = k$ in the latent-factor half-trek π_k since by construction the flow graph $G_{\text{flow}}(v, A, Z)$ does not contain the edge $w' \to z'$ if $w \in Z$. Moreover, if $k \in Z$, it must be the case that $h_k \in H$. Indeed, if we have $h_k \notin H$, then $y_k \in \mathrm{ch}(\mathrm{pa}_{\mathcal{L}}(Z \cup \{v\}) \setminus H)$ which is impossible by assumption since $y_k \in A$. Thus, Π is a system of latent-factor half-treks with no sided intersection from $Y = \{y_k : k \in \mathrm{pa}_V(v) \cup Z\}$ to $Z \cup \mathrm{pa}_V(v)$ in $G^{\mathcal{L}}$, such that for each $z \in Z$, the half-trek terminating at z takes the form $y \leftarrow h \to z$ for some $y \in Y$ and some $h \in H$. Finally, note that for the triple (Y, Z, H) conditions (i) and (ii) of the LF-HTC are trivially satisfied by construction and the fact that $Y \subseteq A$. □

PROOF OF PROPOSITION 5.2. Suppose the triple (Y, Z, H) satisfies the LF-HTC for $v \in V$ in $G^{\mathcal{L}}$. Recall that there exists a system of latent-factor half-treks Π with no sided intersection from Y to $\mathrm{pa}_V(v) \cup Z$ such that, for each $z \in Z$, the half-trek terminating at z takes the form $y \leftarrow h \to z$ for some $y \in Y$ and some $h \in H$. Since $|Z| = |H|$, it is clearly not possible that there is a node $h \in H$ such that $|\mathrm{ch}(h)| = 1$.

Now let $h \in H$ such that $|\mathrm{ch}(h)| \in \{2, 3\}$. Then there is a unique latent-factor half-trek in Π that has the form $y \leftarrow h \to z$ for some $y \in Y$ and some $z \in Z$. Let $\widetilde{Y} = Y \setminus \{y\}$ and $\widetilde{Z} = Z \setminus \{z\}$. It is clear that the triple $(\widetilde{Y}, \widetilde{Z}, \widetilde{H})$ satisfies conditions (i) and (iii) of the LF-HTC and $\widetilde{Y} \cap (\widetilde{Z} \cup \{v\}) = \emptyset$. Thus it is left to show that $h \notin \mathrm{pa}_{\mathcal{L}}(\widetilde{Y}) \cap \mathrm{pa}_{\mathcal{L}}(\widetilde{Z} \cup \{v\})$.

If $|\mathrm{ch}(h)| = 2$, there are no more children of h other than y and z. Thus, we directly see that $h \notin \mathrm{pa}_{\mathcal{L}}(\widetilde{Y}) \cap \mathrm{pa}_{\mathcal{L}}(\widetilde{Z} \cup \{v\})$. If $|\mathrm{ch}(h)| = 3$, there might be one child $w \in \mathrm{ch}(h) \setminus \{y, z\}$. But then due to $\widetilde{Y} \cap (\widetilde{Z} \cup \{v\}) = \emptyset$, this node w cannot be in both sets $\widetilde{Y}$ and $\widetilde{Z} \cup \{v\}$ at the same time. Thus $h \notin \mathrm{pa}_{\mathcal{L}}(\widetilde{Y}) \cap \mathrm{pa}_{\mathcal{L}}(\widetilde{Z} \cup \{v\})$ as well. We conclude that condition (ii) of the LF-HTC is satisfied by the triple $(\widetilde{Y}, \widetilde{Z}, \widetilde{H})$ and therefore it satisfies the LF-HTC for $v \in V$. □

PROOF OF THEOREM 5.3. The proof works in the same way as the proof of Theorem 7 in Foygel, Draisma and Drton (2012a). We start by analyzing the complexity of the algorithm.

Observe that we run the "inner" algorithm (line 3 to 15) at most $|V|^2$ times. This can be seen by counting the maximal number of repetitions in line 1. Another repetition is only done if a node was added to S in the repetition before, otherwise the algorithm terminates. Thus after $|V|$ repetitions of line 1 either all nodes were added to S or the algorithm terminated before. By investigating line 2 we see that in every pass we also iterate over at most $|V|$ nodes which yields the maximal number of $|V|^2$ runs of the inner algorithm.

In the inner algorithm itself we iterate first through all sets $H \subseteq \mathcal{L}_{\geq 4} \subseteq \mathcal{L}$ with cardinality at most k. The number of subsets of $\mathcal{L}$ with cardinality at most k is

$$\sum_{i=0}^{k} \binom{|\mathcal{L}_{\geq 4}|}{i} = \mathcal{O}(|\mathcal{L}|^k).$$

In line 5 we then iterate over all $Z \subseteq Z_a \subseteq V$ with $|Z| = |H|$. Similarly as before, we see that in the worst case these are $\mathcal{O}(|V|^k)$ iterations. Hence, we compute at most $\mathcal{O}(|V|^2|\mathcal{L}|^k|V|^k)$ maximum flows on a graph with at most $2(|V| + |\mathcal{L}|) + 2$ nodes and $4|V| + |\mathcal{L}| + |D|$ edges and the same number r of reciprocal edge pairs as in D_V. By Cormen et al. (2009, Section 26) each maximum flow computation has complexity at most $\mathcal{O}((|V| + |\mathcal{L}| + r)^3)$. Finally, note that the sets $\mathrm{htr}_H(U)$ for a subset $U \subseteq V$ can be found using breadth first search which has complexity $\mathcal{O}(|V| + |\mathcal{L}| + |D|)$ by Cormen et al. (2009, Section 22.2). Finding parents

and children of nodes is not of higher complexity. Since $|D| \leq |V|^2$, we conclude that the total complexity is $\mathcal{O}(|V|^{2+k}|\mathcal{L}|^k(|V|+|\mathcal{L}|+r)^3)$.

Next we show that the algorithm indeed determines LF-HTC-identifiability. Suppose that $G^{\mathcal{L}}$ is LF-HTC-identifiable. Then by Theorem 3.7 there is a total ordering $\prec$ on V such that $w \prec v$ whenever $w \in Z_v \cup (Y_v \cap \mathrm{htr}_{H_v}(Z_v \cup \{v\}))$ where $(Y_v, Z_v, H_v) \in 2^{V\setminus\{v\}} \times 2^{V\setminus\{v\}} \times 2^{\mathcal{L}}$ is a triple satisfying the LF-HTC with respect to v. Hence, if $G^{\mathcal{L}}$ is LF-HTC-identifiable, we might label the elements $\{v_1, \ldots, v_d\} = V$ such that $v_1 \prec v_2 \prec \cdots \prec v_d$.

Now we claim that after at most $k+1$ passes through the for loop in line 2, all nodes v_i, $i \prec k$, have already been added to the solved nodes S. We prove this by induction. Suppose that all nodes $v_1, \ldots, v_{k-1} \in S$ and we are now testing the k-th node v_k. Let $(Y_{v_k}, Z_{v_k}, H_{v_k})$ be the triple satisfying the LF-HTC with respect to v_k. At one point, we will visit the correct set $H_{v_k} \in \mathcal{L}_{\geq 4}$ in line 3 due to Proposition 5.2. If $z \in Z_{v_k}$, then $z \prec v_k$ and therefore $z \in S$ already. Additionally, $z \in \mathrm{ch}(H_{v_k})$ and $z \notin \{v_k\} \cup \mathrm{pa}_V(v_k)$ by definition of the LF-HTC. Thus, we will visit the correct set $Z_{v_k} \subseteq Z_a$ in line 5. Now take any $y \in Y_{v_k}$. By definition of the LF-HTC, we have that $y \notin Z_{v_k} \cup \{v_k\} \cup \mathrm{ch}(\mathrm{pa}(Z_{z_k} \cup \{v_k\}) \setminus H_{v_k})$. Moreover, if $y \in \mathrm{htr}_{H_{v_k}}(Z_{v_k} \cup \{v_k\})$, then $y \prec v_k$ and thus $y \in S$, which means $y \in A$. If instead $y \notin \mathrm{htr}_{H_{v_k}}(Z_{v_k} \cup \{v_k\})$, then $y \in A$ by definition of A. Therefore, $Y_{v_k} \subseteq A$ and by Theorem 5.1 we will add v_k to S. By induction, we obtain that $S = V$ after at most $|V|$ repetitions of line 2 to 16.

Conversely, suppose the algorithm finds $S = V$, and fix a node $v \in V$. It remains to show that there is a triple $(Y_v, Z_v, H_v) \in 2^{V\setminus\{v\}} \times 2^{V\setminus\{v\}} \times 2^{\mathcal{L}}$ such that all nodes $w \in Z_v \cup (Y_v \cap \mathrm{htr}_{H_v}(Z_v \cup \{v\}))$ were added to S in the steps before. When v was added to S, there must have been sets $H_v \subseteq \mathcal{L}_{\geq 4}$ and $Z_v \subseteq (S \cap \mathrm{ch}(H_v)) \setminus (\{v\} \cup \mathrm{pa}_V(v))$ with $|Z| = |H|$ such that $\texttt{MaxFlow}(G_{\text{flow}}(v, A, Z_v)) = |\mathrm{pa}_V(v)| + |Z_v|$. By Theorem 5.1, this means that there is a set $Y_v \subseteq A$ such that the triple (Y_v, Z_v, H_v) satisfies the LF-HTC with respect to v. By construction, $Z_v \subseteq S$ at this stage of the algorithm. Moreover, we have for all $w \in A$ that either $w \in S$ already or $w \notin \mathrm{htr}_{H_v}(Z_v \cup \{v\})$. Thus, we have as well that $Y_v \cap \mathrm{htr}_{H_v}(Z_v \cup \{v\}) \subseteq S$ at this stage of the algorithm. Applying this reasoning to all $v \in V$, we see that $G^{\mathcal{L}}$ is LF-HTC-identifiable.

□

APPENDIX B: NP-HARDNESS OF THE LF-HTC

In this section we show that the task of deciding $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ for unrestricted graphs is NP-hard. That is, it is at least as hard as the hardest problems in the NP-complexity class. Formally, we have to show that every problem in NP is reducible to $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ in polynomial time. Fortunately, it is enough to show that one arbitrary problem that is known to be NP-hard is reducible to $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ in polynomial time. For this purpose, we choose the Boolean satisfiability problem in conjunctive normal form (CNFSAT). This is the problem of determining whether a Boolean expression in conjunctive normal form is satisfiable. That is, suppose we have Boolean variables $\{x_1, \ldots, x_n\}$, and let

$$C = C_1 \wedge \cdots \wedge C_M = (\ell_1^1 \vee \cdots \vee l_{m_1}^1) \wedge \cdots \wedge (\ell_1^M \vee \cdots \vee l_{m_M}^M)$$

where $\ell_j^i \in \{x_1, \ldots, x_n, \neg x_1, \ldots, \neg x_n\}$ for all $1 \leq i \leq M$ and $1 \leq j \leq m_i$. We call the elements of $\{x_1, \ldots, x_n, \neg x_1, \ldots, \neg x_n\}$ *literals* and $\neg x_i$ the *negation* of x_i. Then $\texttt{CNFSAT}(\{x_1, \ldots, x_n\}, C)$ is the problem of determining if there exist assignments of `True` and `False` to each x_i such that, under this assignment, C is `True`, i.e., satisfied.

THEOREM B.1. *There exists a polynomial time reduction from* CNFSAT *to* LF-HTC$(G^{\mathcal{L}}, v)$ *so that* LF-HTC$(G^{\mathcal{L}}, v)$ *is NP-hard.*

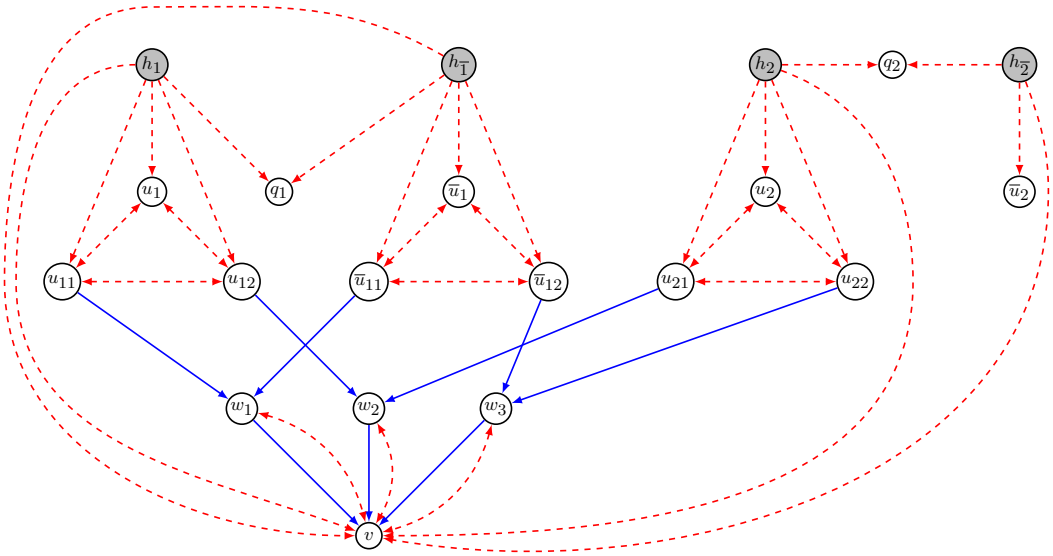


FIG 1. *The graph $G^{\mathcal{L}}$ corresponding to the* CNFSAT *problem with Boolean expression $C = (x_1 \vee \neg x_1) \wedge (x_1 \vee x_2) \wedge (\neg x_1 \vee x_2)$. To not clutter the graph, a red bidirected edge (dashed) corresponds to a latent factor that has only arrows pointing to the two endpoints of the edge, e.g. the red bidirected edge $w_1 \leftrightarrow v$ corresponds to $w_1 \leftarrow h_{w_1 v} \rightarrow v$.*

PROOF. To see that there is a polynomial time reduction from CNFSAT to LF-HTC$(G^{\mathcal{L}}, v)$, let $X = \{x_1, \ldots, x_n\}$ and let C be as above. We now construct a latent-factor graph $G^{\mathcal{L}}$ and show that solving LF-HTC$(G^{\mathcal{L}}, v)$ in $G^{\mathcal{L}}$ solves CNFSAT(X, C). Our construction initializes $G^{\mathcal{L}} = (V \cup \mathcal{L}, D)$ to be empty. In the following when we add nodes to $G^{\mathcal{L}}$ they will implicitly be added to V unless they are labeled h_* for some index $*$, in which case they are to be added to $\mathcal{L}$.

Begin by adding to the graph the nodes $v, w_1, \ldots, w_M, h_{w_1 v}, \ldots, h_{w_m v}$, the edges $w_i \rightarrow v$, and the edges $w_i \leftarrow h_{w_i v} \rightarrow v$ for all $1 \leq i \leq M$. The w_i will correspond to the M disjunctive clauses in C. Now for the ith Boolean variable x_i, let A_i be the number of times x_i (in non-negated form) appears in C, and let B_i be the number of times $\neg x_i$ appears in C. Then add to the graph

(i) the nodes $u_{i1}, \ldots, u_{iA_i}$, $\overline{u}_{i1}, \ldots, \overline{u}_{iB_i}$, $u_i, \overline{u}_i$, $h_i, h_{\overline{i}}$, and q_i,
(ii) $u_{ij} \rightarrow w_k$ if the j-th appearance, from the left, of x_i (in non-negated form) in C occurs in the k-th disjunctive clause of C,
(iii) $\overline{u}_{ij} \rightarrow w_k$ if the j-th appearance, from the left, of $\neg x_i$ in C occurs in the k-th disjunctive clause of C,
(iv) $h_i \rightarrow a$ for each $a \in \{u_{i1}, \ldots, u_{iA_i}, u_i, q_i, v\}$,
(v) $h_{\overline{i}} \rightarrow a$ for each $a \in \{\overline{u}_{i1}, \ldots, \overline{u}_{iB_i}, \overline{u}_i, q_i, v\}$,
(vi) a node $h_{ab} = h_{ba}$ and edges $a \leftarrow h_{ab} \rightarrow b$ for each pair of variables $a, b \in \{u_i, u_{i1}, \ldots, u_{iA_i}\}$, and
(vii) a node $h_{ab} = h_{ba}$ and edges $a \leftarrow h_{ab} \rightarrow b$ for each pair of variables $a, b \in \{\overline{u}_i, \overline{u}_{i1}, \ldots, \overline{u}_{iB_i}\}$.

An example of a graph $G^{\mathcal{L}}$ corresponding to a Boolean expression can be found in Figure 1. Now that we have constructed $G^{\mathcal{L}}$ we claim that every triple (Y, Z, H) satisfying the LF-HTC for $v \in V$ in $G^{\mathcal{L}}$ corresponds to an assignment to X such that C is satisfied under

this assignment and vice versa.

We will now start with the more complicated direction. Suppose that there is a triple (Y, Z, H) satisfying the LF-HTC for v in $G^{\mathcal{L}}$. That is, there exists a latent-factor half-trek system Π from Y to $Z \cup \mathrm{pa}_V(v)$ satisfying the appropriate LF-HTC conditions.

Claim 1**:** No w_i is an element of Y.

Suppose for contradiction that some $w_i \in Y$. Since there exists a node $h_{w_i v}$ whose only edges are $w_i \leftarrow h_{w_i v} \rightarrow v$, condition (ii) of the LF-HTC implies that $h_{w_i v}$ must be in H. But then condition (iii) implies that there must be some $z \in \mathrm{ch}(h_{w_i v}) \cap Z$ for which the latent-factor half-trek $y \leftarrow h_{w_i v} \rightarrow z$ is in Π. By $\mathrm{ch}(h_{w_i v}) = \{w_i, v\}$ we have a contradiction since if $z = v$ we would have $v \in Z$, and if $z = w_i$ we have that $w_i \in Y \cap Z$ so that $Y \cap Z \neq \emptyset$. Hence there is no $w_i \in Y$.

Claim 2**:** If $Y \cap \{u_i,\ u_{i1}, \ldots, u_{iA_i}\} \neq \emptyset$, then $Y \cap \{\overline{u}_i,\ \overline{u}_{i1}, \ldots, \overline{u}_{iB_i}\} = \emptyset$.

Suppose that $u \in Y \cap \{u_i,\ u_{i1}, \ldots, u_{iA_i}\}$. Since $\mathrm{ch}(h_i) = \{v, u_i, q_i, u_{i1}, \ldots, u_{iA_i}\}$, it follows from condition (ii) of the LF-HTC that $h_i \in H$. Hence, by condition (iii), it must be the case that there is some $y, z \in \{q_i, u_i, u_{i1}, \ldots, u_{iA_i}\}$ with $y \neq z$ such $y \in Y$ and $z \in Z$ and the latent-factor half-trek $y \leftarrow h_i \rightarrow z \in \Pi$. There are two cases.

Case 1: $z \neq q_i$. We must have that $z \in \{u_i, u_{i1}, \ldots, u_{iA_i}\} \setminus \{u\}$. Recall, in this case, that there exists h_{uz} whose only children are u and z. By a similar argument as in claim 1, it follows that $u \leftarrow h_{uz} \rightarrow z$ must also be in Π which contradicts the fact that Π must have no sided intersection (since z is already in the right-hand side of the latent-factor half-trek $u \leftarrow h_i \rightarrow z$).

Case 2: $z = q_i$. In this case we must have that the latent-factor half-trek $y \leftarrow h_i \rightarrow q_i \in \Pi$ for some $y \in \{u_i, u_{i1}, \ldots, u_{iA_i}\}$. Now, for essentially identical reasons as above, if we have $\overline{u} \in Y \cap \{\overline{u}_i,\ \overline{u}_{i1}, \ldots, \overline{u}_{iB_i}\}$, there is some $\overline{y} \in \{\overline{u}_i, \overline{u}_{i1}, \ldots, \overline{u}_{iB_i}\} \setminus \{\overline{u}\}$ such that $\overline{y} \leftarrow h_{\overline{y}q_i} \rightarrow q_i$ is in Π. This contradicts the fact that Π has no sided intersection (since q_i is in the right part of two latent-factor half-treks) and thus we must have that $Y \cap \{\overline{u}_i, \overline{u}_{i1}, \ldots, \overline{u}_{iB_i}\} = \emptyset$.

Since the first case results in a contradiction we must be in the second case with $Y \cap \{\overline{u}_i, \overline{u}_{i1}, \ldots, \overline{u}_{iB_i}\} = \emptyset$.

Claim 3**:** If $Y \cap \{\overline{u}_i,\ \overline{u}_{i1}, \ldots, \overline{u}_{iB_i}\} \neq \emptyset$, then $Y \cap \{u_i,\ u_{i1}, \ldots, u_{iA_i}\} = \emptyset$.

The claim follows by symmetry from claim 2.

We can now show that our given triple (Y, Z, H) satisfying the LF-HTC for $v \in V$ in $G^{\mathcal{L}}$ corresponds to an assignment to X such that C is satisfied under this assignment. For each $1 \leq i \leq M$, assign x_i to be `True` if $Y \cap \{u_i,\ u_{i1}, \ldots, u_{iA_i}\} \neq \emptyset$ and `False` otherwise. To see that this satisfies C consider the i-th disjunctive clause $C_i = (\ell_1^i \vee \cdots \vee l_{m_i}^i)$ of C. Since (Y, Z, H) satisfies the LF-HTC for v, there must exist some $y \in Y$ such that there is a latent-factor half-trek from y to w_i in Π. By claim 1 and the proof of claim 2, we must have that $y \in \{u_j, u_{j1}, \ldots, u_{jA_j}\}$ or $y \in \{\overline{u}_j, \overline{u}_{j1}, \ldots, \overline{u}_{jB_j}\}$ for some j. Suppose that $y \in \{u_j, u_{j1}, \ldots, u_{jA_j}\}$. Then, since there must be a half-trek from y to w_i, we have, by the construction of $G^{\mathcal{L}}$ that x_j appears (in non-negated form) in C_i. Since $\{u_j, u_{j1}, \ldots, u_{jA_j}\} \cap Y \neq \emptyset$, we must have set x_j to be `True` and thus C_i is satisfied. If instead $y \in \{\overline{u}_j, \overline{u}_{j1}, \ldots, \overline{u}_{jB_j}\}$, it follows, by the same logic, that $\neg x_j$ must appear in C_i and that we set x_j to be `False`, so again C_j is satisfied. As j was arbitrary, C is satisfied by

the assignment.

Now we wish to show the opposite direction, namely, that if there is an assignment to X such that C is satisfied under this assignment, then there must also be a set (Y, Z, H) satisfying the LF-HTC for v. Suppose we have assigned `True` and `False` to the x_i, so that C is satisfied. Let $1 \leq i \leq M$. For the ith disjunctive clause C_i in C, let l_k be the first literal in the clause, which evaluates to `True` (there must be at least one such literal since $C = \texttt{True}$ implies $C_i = \texttt{True}$). Now l_k must equal x_j or $\neg x_j$ for some j. Suppose that $l_k = x_j$. Then there exists a unique $1 \leq \ell \leq A_j$ such that edge $u_{j\ell} \to w_i$ is in the graph. If $l_k = \neg x_j$ then, similarly, there exists a unique $1 \leq \ell \leq B_j$ such that there exists an edge $\overline{u}_{j\ell} \to w_i$ in the graph. In either case, denote $u_{j\ell}$ or $\overline{u}_{j\ell}$ as y_i.

Now for $M + 1 \leq i \leq M + n$, let $y_i = u_i$ if x_i is `True`, and $y_u = \overline{u}_i$ otherwise. Let $Z = \{q_1, \ldots, q_n\}$, $L = \{\tilde{h}_1, \ldots, \tilde{h}_n\}$ where $\tilde{h}_i = h_i$ if $x_i = \texttt{True}$ and $\tilde{h}_i = h_{\bar{i}}$ if otherwise, and $Y = \{y_1, \ldots, y_{M+n}\}$. By our construction, it holds that

(i) $|Y| = |\mathrm{pa}_V(v)| + |H|$,
(ii) $Y \cap (Z \cup \{v\}) = \emptyset$ and if $y \in Y$ and v (or $z \in Z$) are children of the same latent factor, then that latent factor is some h_i for which $x_i = \texttt{True}$ or some $h_{\bar{i}}$ for which $x_i = \texttt{False}$, and
(iii) the set of latent-factor half-treks Π with elements

$$\begin{aligned} &y_i \to w_i \quad && \text{for } 1 \leq i \leq M, \\ &y_{M+i} \leftarrow h_i \to z_i = q_i \quad && \text{for } 1 \leq i \leq n \text{ if } x_i \text{ is } \texttt{True}, \text{ and} \\ &y_{M+i} \leftarrow h_{\bar{i}} \to z_i = q_i \quad && \text{for } 1 \leq i \leq n \text{ if } x_i \text{ is } \texttt{False} \end{aligned}$$

forms a latent-factor half-trek system from Y to $\mathrm{pa}_V(v) \cup Z$ for which the half-trek to each z_i is of the form $y_{M+i} \leftarrow \tilde{h}_i \to z_i$ and if y_i has a common latent parent with v (or $z \in Z$) then the latent parent must correspond to some $\tilde{h}_j \in H$ and we have that $y_{M+j} \leftarrow h_j \to q_j \in \Pi$.

Note that the above three conditions immediately imply that (Y, Z, H) satisfies the LF-HTC for v. We have thus shown that `CNFSAT` reduces to $\texttt{LF-HTC}(G^{\mathcal{L}}, v)$ in polynomial time. □

APPENDIX C: ALGEBRAIC TECHNIQUES FOR DETERMINING IDENTIFIABILITY

As discussed in Garcia-Puente, Spielvogel and Sullivant (2010), rational identifiability may be decided by techniques from computational algebraic geometry. For the original half-trek criterion, Foygel, Draisma and Drton (2012a) provide an effective algorithm to perform the necessary computations. In this section we show how their approach may be generalized to cover the latent-factor setup from this paper. Our implementation and the code to compute the numerical experiments in Section 6 are available at `https://github.com/NilsSturma/lfhtc`.

Consider a slightly more general setting than before, i.e., let $S \subseteq \mathbb{R}^m$ be an open semialgebraic set, and let

$$\begin{aligned} \tau : S &\longrightarrow PD(d) \\ \Delta &\longmapsto \tau(\Delta) \end{aligned}$$

be a polynomial map that parametrizes the cone of latent covariance matrices $\mathrm{Im}(\tau)$. Together with a directed graph on the observed nodes $G^V = (V, D_V)$ with $V = |d|$, the cone of latent covariance matrices $\mathrm{Im}(\tau)$ postulates a covariance model.

DEFINITION C.1. The covariance model given by a tuple $(G^V, \mathrm{Im}(\tau))$, consisting of a directed graph $G^V = (V, D_V)$ with $V = |d|$ and a cone of latent covariance matrices $\mathrm{Im}(\tau)$, is given by the family of all covariance matrices

$$\Sigma = (I_d - \Lambda)^{-\top} \Omega (I_d - \Lambda)^{-1}$$

for $\Lambda \in \mathbb{R}^{D_V}_{\mathrm{reg}}$ and $\Omega \in \mathrm{Im}(\tau)$.

A covariance matrix $\Sigma \in \boldsymbol{PD}(d)$ is in the covariance model given by a tuple $(G^V, \mathrm{Im}(\tau))$ if and only if $\Sigma = \varphi(\Lambda, \tau(\Delta))$ for $\Lambda \in \mathbb{R}^{D_V}_{\mathrm{reg}}$ and $\Delta \in S$ where the parametrization map φ is given by

$$\begin{aligned} \varphi : \mathbb{R}^{D_V}_{\mathrm{reg}} \times \boldsymbol{PD}(d) &\longrightarrow \boldsymbol{PD}(d) \\ (\Lambda, \Omega) &\longmapsto (I_d - \Lambda)^{-\top} \Omega (I_d - \Lambda)^{-1}. \end{aligned}$$

If we let $S = \mathbb{R}^{D_{\mathcal{L}V}} \times \mathrm{diag}^+_d$ and $\tau : (\Gamma, \Omega_{\mathsf{diag}}) \mapsto \Omega_{\mathsf{diag}} + \Gamma^\top \Gamma$, then Definition C.1 coincides with Definition 2.1 of a covariance model given by a latent-factor graph.

In what follows, let $\lambda = \{\lambda_{ij} : i \to j \in D_V\}$ be variables representing the non-zero entries of $\Lambda \in \mathbb{R}^{D_V}_{\mathrm{reg}}$, and let $\omega = \{\omega_{ij} : 1 \le i \le j \le d\}$ be variables representing the entries of $\Omega \in \boldsymbol{PD}(d)$. Let $d(\lambda) \in \mathbb{R}[\lambda]$ be the polynomial defined by $\det(I_d - \Lambda)$ for $\Lambda \in \mathbb{R}^{D_V}$. Now observe that, for $(\Lambda, \Omega) \in \mathbb{R}^{D_V}_{\mathrm{reg}} \times \boldsymbol{PD}(d)$, we may write the ij-th entry of the matrix $\varphi(\Lambda, \Omega)$ as a rational function

$$\varphi_{ij}(\Lambda, \Omega) = \frac{\widetilde{\varphi}_{ij}(\lambda, \omega)}{d(\lambda)^2}$$

with $\widetilde{\varphi}_{ij}(\lambda, \omega) \in \mathbb{R}[\lambda, \omega]$ due to Cramer's rule. Furthermore, we write $\sigma = \{\sigma_{ij} : 1 \le i \le j \le d\}$ and $\delta = \{\delta_i : i = 1, \ldots, m\}$ for variables representing the entries of $\Sigma \in \boldsymbol{PD}(d)$ and $\Delta \in S$, respectively. Consider the polynomial ring $\mathbb{R}[\lambda, \sigma, \delta, \xi]$ with one additional variable ξ. Then the vanishing ideal of the graph of the parametrization in Definition C.1 is given by

$$\mathcal{J} = \langle \{\sigma_{ij} d(\lambda)^2 - \widetilde{\varphi}_{ij}(\lambda, \tau(\delta)), 1 \le i \le j \le d\} \cup \{\xi d(\lambda) - 1\} \rangle \subseteq \mathbb{R}[\lambda, \sigma, \delta, \xi].$$

The additional variable ξ and the polynomial $\xi d(\lambda) - 1$ are needed to ensure that $d(\lambda)$ never vanishes. Eliminating the variables λ, δ and ξ, we get the vanishing ideal of the image of $\Theta = \mathbb{R}^{D_V}_{\mathrm{reg}} \times \mathrm{Im}(\tau)$ under the parametrization φ, in formulas,

$$I(\varphi(\Theta)) = \mathcal{J} \cap \mathbb{R}[\sigma].$$

Nevertheless, for the purpose of identifying direct causal effects, we are interested in an ideal where λ is not eliminated, i.e., we are interested in

$$\mathcal{I} = \mathcal{J} \cap \mathbb{R}[\lambda, \sigma].$$

By definition, this ideal consists exactly of those polynomials $f(\lambda, \sigma) \in \mathbb{R}[\lambda, \sigma]$ such that $f(\Lambda, \varphi(\Lambda, \Omega)) = 0$ for all $(\Lambda, \Omega) \in \Theta$.

PROPOSITION C.2. *The parameter λ_{ij} is rationally identifiable if and only if $\mathcal{I}$ contains an element of the form $a(\sigma)\lambda_{ij} - b(\sigma)$ with $a, b \in \mathbb{R}[\sigma]$ and $a \notin I(\varphi(\Theta))$.*

PROOF. The proof is similar to the proofs of Lemma 7 in Foygel, Draisma and Drton (2012b) and Proposition 3 in Garcia-Puente, Spielvogel and Sullivant (2010). For completeness, we give the full proof in our notation. Suppose that the parameter λ_{ij} is rationally identifiable. By definition, there is a rational function $b(\sigma)/a(\sigma) \in \mathbb{R}(\sigma)$ such that

$$\frac{b(\varphi(\Lambda, \Omega))}{a(\varphi(\Lambda, \Omega))} = \lambda_{ij}.$$

for all $(\Lambda, \Omega) \in \Theta \setminus A$, where A is a proper algebraic subset of the Zariski closure of Θ. In particular, outside A we must have that $a(\varphi(\Lambda, \Omega)) \neq 0$ and therefore it must be the case that $a \notin I(\varphi(\Theta))$. On the other hand, it is clear that the polynomial $a(\sigma)\lambda_{ij} - b(\sigma)$ is a member of $\mathcal{I}$ since this polynomial vanishes if we substitute σ by $\varphi(\Lambda, \Omega)$ for any $(\Lambda, \Omega) \in \Theta \setminus A$.

Conversely, suppose that a and b satisfy the given conditions. Since $a \notin I(\varphi(\Theta))$, we have $a(\varphi(\Lambda, \Omega)) \neq 0$ for all $(\Lambda, \Omega) \in \Theta \setminus A$, where A is a proper algebraic subset of the Zariski closure of Θ. But then b/a is a rational function identifying λ_{ij} from σ outside the proper algebraic subset A. □

For checking rational identifiability one has to check the membership of polynomials of the form $a(\sigma)\lambda_{ij} - b(\sigma)$ in the ideal $\mathcal{I}$. This can be achieved by computation of a Gröbner basis of $\mathcal{I}$ with eliminating term order using Buchberger's algorithm; see Garcia-Puente, Spielvogel and Sullivant (2010). The Gröbner basis computation can be very challenging in terms of running times, even for small graphs. One reason is that the input polynomials to the computation $\widetilde{\varphi}_{ij}(\lambda, \tau(\delta)), 1 \leq i \leq j \leq n$, may already have large degree. As mentioned in Foygel, Draisma and Drton (2012b), it is easy to construct graphs where the bit-size of those polynomials is already exponential in the size of the graphs. By Definition C.1, we have the equation $\Sigma = (I_d - \Lambda)^{-\top}\tau(\Delta)(I_d - \Lambda)^{-1}$. Since the matrix $(I_d - \Lambda)$ is required to be invertible, this is equivalent to the equation

$$(I_d - \Lambda)^\top \Sigma (I_d - \Lambda) = \tau(\Delta). \tag{C.1}$$

Clearly, the entries of the matrix on the left-hand side are cubic, i.e., the maximal degree of the involved polynomials in σ and λ is 3. We suggest to exploit this fact instead of computing the Gröbner basis for $\mathcal{I}$ directly. The idea is to compute a generating set of the vanishing ideal of the cone of latent covariance matrices $\mathrm{Im}(\tau)$ and then to plug-in the polynomials from the left-hand side. The resulting polynomials then indeed define the same ideal $\mathcal{I}$ and may be much smaller in size. Therefore, the Gröbner basis computation may be faster. This is proved in Proposition C.3 below, but we need to introduce some more notation beforehand.

As usual, we denote $I(\mathrm{Im}(\tau)) = \{f \in \mathbb{R}[\omega] : f(\Omega) = 0 \text{ for all } \Omega \in \mathrm{Im}(\tau)\}$ for the vanishing ideal of $\mathrm{Im}(\tau)$. We will also need the map corresponding to Equation (C.1), i.e.,

$$\begin{aligned} \psi : \mathbb{R}^{D_V}_{\mathrm{reg}} \times \boldsymbol{PD}(d) &\longrightarrow \boldsymbol{PD}(d) \\ (\Lambda, \Sigma) &\longmapsto (I_d - \Lambda)^\top \Sigma (I_d - \Lambda). \end{aligned}$$

Note that ψ may be interpreted as an "inverse" of φ in the sense that $\psi(\Lambda, \varphi(\Lambda, \Omega)) = \Omega$ and $\varphi(\Lambda, \psi(\Lambda, \Sigma)) = \Sigma$. Since ψ and τ are polynomial functions by definition, we write under abuse of notation $\psi(\lambda, \sigma)$ and $\tau(\delta)$ for the collection of polynomials they define. Similarly, we write $\widetilde{\varphi}(\lambda, \omega)$ for the collection of polynomial functions $\widetilde{\varphi}_{ij}(\lambda, \omega)$ for $1 \leq i \leq j \leq d$.

PROPOSITION C.3. *Let* $A_S = \{h \circ \psi \in \mathbb{R}[\lambda, \sigma] : h \in I(\mathit{Im}(\tau))\}$. *Then*

$$\mathcal{I} = \langle A_S, \xi d(\lambda) - 1 \rangle \cap \mathbb{R}[\lambda, \sigma].$$

PROOF. We begin by showing $A_S \subseteq \mathcal{I}$. Thus let $f(\lambda, \sigma) \in A_S$. By definition of A_S, there is $h \in I(\mathrm{Im}(\tau))$ such that $f = h \circ \psi$. Hence, for any point $(\Lambda, \Omega) \in \Theta$, we have

$$f(\Lambda, \varphi(\Lambda, \Omega)) = h(\psi(\Lambda, \varphi(\Lambda, \Omega))) = h(\Omega) = 0$$

since $\Omega \in \mathrm{Im}(\tau)$. This means that $f \in \mathcal{I}$ and therefore $A_S \subseteq \mathcal{I}$. But this yields that $\langle A_S, \xi d(\lambda) - 1 \rangle \subseteq \mathcal{J}$, and by the definition of $\mathcal{I}$ we conclude that $\langle A_S, \xi d(\lambda) - 1 \rangle \cap \mathbb{R}[\lambda, \sigma] \subseteq \mathcal{I}$.

For the other inclusion, let $I(\mathrm{Im}(\tau)) = \langle h_1, \dots, h_r \rangle \subseteq \mathbb{R}[\omega]$ and $f(\lambda, \sigma) \in \mathcal{I}$. Define the polynomial $g(\lambda, \omega, \xi) = f(\lambda, \xi^2 \widetilde{\varphi}(\lambda, \omega))$, which is an element of the polynomial ring $\mathbb{R}[\lambda, \omega, \xi]$. By the definition of $\mathcal{I}$, the polynomial g becomes zero if we plug in $(\lambda, \omega, \xi) = (\Lambda, \Omega, d(\Lambda)^{-1})$ for any $(\Lambda, \Omega) \in \Theta$. Therefore, g lies in the ideal $\langle h_1(\omega), \dots, h_r(\omega), \xi d(\lambda) - 1 \rangle$ interpreted in the ring $\mathbb{R}[\lambda, \omega, \xi]$. Hence, we can write

$$g(\lambda, \omega, \xi) = \sum_{i=1}^{r} g_i(\lambda, \omega, \xi) h_i(\omega) + g_{r+1}(\lambda, \omega, \xi)(\xi d(\lambda) - 1)$$

for suitable coefficient polynomials $g_i(\lambda, \omega, \xi) \in \mathbb{R}[\lambda, \omega, \xi]$. Plugging in $\psi(\lambda, \sigma)$ for ω, we see that

$$\begin{aligned} g(\lambda, \psi(\lambda, \sigma), \xi) = &\sum_{i=1}^{r} g_i(\lambda, \psi(\lambda, \sigma), \xi) h_i(\psi(\lambda, \sigma)) \\ &+ g_{r+1}(\lambda, \psi(\lambda, \sigma), \xi)(\xi d(\lambda) - 1) \end{aligned}$$

is an element of $\langle A_S, \xi d(\lambda) - 1 \rangle$ since each polynomial $h_i(\psi(\lambda, \sigma)) \in A_S$. Moreover, we have the equality

$$g(\lambda, \psi(\lambda, \sigma), \xi) = f(\lambda, \xi^2 \widetilde{\varphi}(\lambda, \psi(\lambda, \sigma))) = f(\lambda, \xi^2 d(\lambda)^2 \sigma)$$

and thus $f(\lambda, \xi^2 d(\lambda)^2 \sigma) \in \langle A_S, \xi d(\lambda) - 1 \rangle$. But by the fact that $\xi d(\lambda) - 1 \in \langle A_S, \xi d(\lambda) - 1 \rangle$, it holds that the difference

$$f(\lambda, \xi^2 d(\lambda)^2 \sigma) - f(\lambda, \sigma) \in \langle A_S, \xi d(\lambda) - 1 \rangle$$

and therefore it must be the case that the polynomial $f(\lambda, \sigma)$ itself is in the ideal $\langle A_S, \xi d(\lambda) - 1 \rangle$ since every ideal is an additive group. We conclude the proof by noting that $f(\lambda, \sigma)$ does not depend on ξ which means that $f(\lambda, \sigma) \in \mathbb{R}[\lambda, \sigma]$ as well. □

Propositions C.2 and C.3 yield Algorithm 1 for checking rational identifiability of a covariance model given by $(G^V, \mathrm{Im}(\tau))$. The proof of the correctness of the algorithm is identical to the proof of Algorithm 1 in Foygel, Draisma and Drton (2012b) and therefore omitted.

With the reduced Gröbner basis obtained in step 3 of Algorithm 1, it is not just possible to determine rational identifiability, but it is straightforward to modify the algorithm to check if a graph is generically finite-to-one (Garcia-Puente, Spielvogel and Sullivant, 2010).

Note that the computation of the ideal $\mathcal{I}$ requires the polynomials $\psi(\lambda, \sigma)$. To speed-up the algorithm for large-scale computational experiments as in Section 6, it is advantageous to replace the variables σ by numerical values obtained from the entries of a randomly chosen matrix Σ in the model. Put differently, we randomly generate parameters $\Lambda_0 \in \mathbb{R}_{\mathrm{reg}}^{D_V}$ and $\Delta_0 \in S$ and then use the polynomials $\psi(\lambda, \varphi(\Lambda_0, \tau(\Delta_0)))$ instead of $\psi(\lambda, \sigma)$. The Gröbner basis then readily yields the dimension and cardinality of the solution set. In practice, we generate (Λ_0, Δ_0) from prime numbers and we repeat the randomized calculation several times for each graph to avoid false conclusions from random draws yielding parameters $(\Lambda_0, \tau(\Delta_0))$ in special constellations.

Algorithm 1 Algebraically checking rational identifiability

1: Compute a Gröbner basis $\langle h_1, \dots, h_r \rangle \subseteq \mathbb{R}[\omega]$ of the vanishing ideal $I(\mathrm{Im}(\tau))$ using elimination theory.
2: Choose a block-monomial order $\geq$ on the monomials in the variables λ, σ with $\lambda > \sigma$.
3: Let $\mathcal{I} = \langle h_1(\psi(\lambda, \sigma)), \dots, h_r(\psi(\lambda, \sigma)), \xi d(\lambda) - 1 \rangle \cap \mathbb{R}[\lambda, \sigma]$ and compute the reduced Gröbner basis T with respect to $\geq$ of the ideal $\mathcal{I}$.
4: The covariance model given by $(G^V, \mathrm{Im}(\tau))$ is rationally identifiable if and only if for each $i \to j \in D_V$ the basis T contains an element whose leading monomial equals a monomial in σ times λ_{ij}.

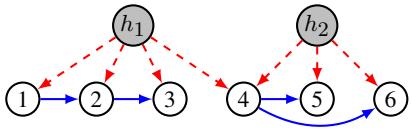


FIG 2. *LF-HTC-identifiable and therefore rationally identifiable.*

EXAMPLE C.4. Consider the latent-factor graph in Figure 2. In this case, the parameter space S is given by $S = \mathbb{R}^{D_{\mathcal{LV}}} \times \mathrm{diag}_d^+$ and we have $\tau : (\Gamma, \Omega_{\mathsf{diag}}) \mapsto \Omega_{\mathsf{diag}} + \Gamma^\top \Gamma$. By implicitization, we find that the Gröbner basis of $I(\mathrm{Im}(\tau))$ is given by the following list of polynomials:

$$\omega_{12}\omega_{34} - \omega_{14}\omega_{23},\ \omega_{13}\omega_{24} - \omega_{14}\omega_{23},\ \omega_{15},\ \omega_{16},\ \omega_{25},\ \omega_{26},\ \omega_{35},\ \omega_{36}.$$

Now, we plug in the relevant polynomials $\psi(\lambda, \sigma)$, which for this graph are given by

$$\begin{aligned}
\psi_{12}(\lambda,\sigma) &= -\lambda_{12}\sigma_{11} + \sigma_{12},\\
\psi_{13}(\lambda,\sigma) &= -\lambda_{23}\sigma_{12} + \sigma_{13},\\
\psi_{14}(\lambda,\sigma) &= \sigma_{14},\\
\psi_{15}(\lambda,\sigma) &= -\lambda_{45}\sigma_{14} + \sigma_{15},\\
\psi_{16}(\lambda,\sigma) &= -\lambda_{46}\sigma_{14} + \sigma_{16},\\
\psi_{23}(\lambda,\sigma) &= \lambda_{12}\lambda_{23}\sigma_{12} - \lambda_{12}\sigma_{13} - \lambda_{23}\sigma_{22} + \sigma_{23},\\
\psi_{24}(\lambda,\sigma) &= -\lambda_{12}\sigma_{14} + \sigma_{24},\\
\psi_{25}(\lambda,\sigma) &= \lambda_{12}\lambda_{45}\sigma_{14} - \lambda_{12}\sigma_{15} - \lambda_{45}\sigma_{24} + \sigma_{25},\\
\psi_{26}(\lambda,\sigma) &= \lambda_{12}\lambda_{46}\sigma_{14} - \lambda_{12}\sigma_{16} - \lambda_{46}\sigma_{24} + \sigma_{26},\\
\psi_{34}(\lambda,\sigma) &= -\lambda_{23}\sigma_{24} + \sigma_{34},\\
\psi_{35}(\lambda,\sigma) &= \lambda_{23}\lambda_{45}\sigma_{24} - \lambda_{23}\sigma_{25} - \lambda_{45}\sigma_{34} + \sigma_{35},\\
\psi_{36}(\lambda,\sigma) &= \lambda_{23}\lambda_{46}\sigma_{24} - \lambda_{23}\sigma_{26} - \lambda_{46}\sigma_{34} + \sigma_{36}.
\end{aligned}$$

As in step 3 of Algorithm 1, we compute the reduced Gröbner basis T of the ideal $\mathcal{I}$. Since T contains the four polynomials

$$\begin{aligned}
&\lambda_{12}\sigma_{11}\sigma_{12}\sigma_{24}\sigma_{34} - \lambda_{12}\sigma_{11}\sigma_{13}\sigma_{24}^2 - \lambda_{12}\sigma_{11}\sigma_{14}\sigma_{22}\sigma_{34}\\
&\quad + \lambda_{12}\sigma_{11}\sigma_{14}\sigma_{23}\sigma_{24} - \lambda_{12}\sigma_{12}\sigma_{14}^2\sigma_{23} + \lambda_{12}\sigma_{13}\sigma_{14}^2\sigma_{22}\\
&\quad - \sigma_{12}^2\sigma_{24}\sigma_{34} + \sigma_{12}\sigma_{13}\sigma_{24}^2 + \sigma_{12}\sigma_{14}\sigma_{22}\sigma_{34} - \sigma_{13}\sigma_{14}\sigma_{22}\sigma_{24},\\
&\lambda_{23}\sigma_{12}\sigma_{24} - \lambda_{23}\sigma_{14}\sigma_{22} - \sigma_{13}\sigma_{24} + \sigma_{14}\sigma_{23},\\
&\lambda_{45}\sigma_{14} - \sigma_{15} \text{ and}\\
&\lambda_{46}\sigma_{14} - \sigma_{16},
\end{aligned}$$

we conclude that the latent-factor graph in Figure 2 is rationally identifiable. It is in fact even LF-HTC-identifiable.

REFERENCES

BASU, S., POLLACK, R. and ROY, M.-F. (2006). *Algorithms in real algebraic geometry*, second ed. *Algorithms and Computation in Mathematics* **10**. Springer-Verlag, Berlin. MR2248869

BOCHNAK, J., COSTE, M. and ROY, M.-F. (1998). *Real algebraic geometry*. *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]* **36**. Springer-Verlag, Berlin Translated from the 1987 French original, Revised by the authors. MR1659509

CORMEN, T. H., LEISERSON, C. E., RIVEST, R. L. and STEIN, C. (2009). *Introduction to algorithms*, Third ed. MIT Press, Cambridge, MA. MR2572804

FORD, L. R. JR. and FULKERSON, D. R. (1962). *Flows in networks*. Princeton University Press, Princeton, N.J. MR0159700

FOYGEL, R., DRAISMA, J. and DRTON, M. (2012a). Half-trek criterion for generic identifiability of linear structural equation models. *Ann. Statist.* **40** 1682–1713. MR3015040

FOYGEL, R., DRAISMA, J. and DRTON, M. (2012b). Supplement to Half-trek criterion for generic identifiability of linear structural equation models. *Ann. Statist.* **40**.

GARCIA-PUENTE, L. D., SPIELVOGEL, S. and SULLIVANT, S. (2010). Identifying Causal Effects with Computer Algebra. In *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence (UAI)*. AUAI Press.

Permission to include the article

COPYRIGHT TRANSFER AGREEMENT

INSTITUTE OF MATHEMATICAL STATISTICS
BRAZILIAN STATISTICAL ASSOCIATION
ASSOCIATION PUBLICATIONS DE L'INSTITUT HENRI POINCARÉ
BERNOULLI SOCIETY FOR MATHEMATICAL STATISTICS & PROBABILITY

This signed Agreement must be received before the manuscript can be scheduled for publication. *Please fill in and sign as indicated. Fax, email or mail to:* ***Mattson Publishing Services, 2002 Holly Neck Road, Baltimore, MD 21221 USA Email: assistant@mattsonpublishing.com, Fax: 410-391-2542***

Transmittal #: AOS2221

Author(s): Rina Foygel Barber, Mathias Drton, Nils Sturma, Luca Weihs

Title: Half-Trek Criterion for Identifiability of Latent Variable Models

Copyright of the above work is hereby transferred to (please check the box next to the journal in which your article will appear).

Institute of Mathematical Statistics
□ Annals of Applied Probability
□ Annals of Probability
□ Annals of Applied Statistics
■ Annals of Statistics
□ Statistical Science

Association Publications de l'Institute Henri Poincaré
□ Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques

Bernoulli Society for Mathematical Statistics & Probability
□ Bernoulli

Brazilian Statistical Association
□ Brazilian Journal of Probability and Statistics

However, **the author(s) reserve the following**:

1. All proprietary rights other than copyright, such as patent rights.
2. The rights to use all or part of this work (with appropriate reference to the journal and article) in future works of their own, such as lectures, press releases, reviews, textbooks, or reprint books.
3. The right to place the final version of this article (exactly as published in the journal) on their own homepage or in a public digital repository, provided there is a link to the official journal site.
4. The right to provide a copy of the final peer-reviewed manuscript to the NIH upon acceptance for publication, for public archiving in PubMed Central as soon as possible but no later than 12 months after publication.

All authors must sign this form. This form may be duplicated for signing by co-authors.
In the case of a "work made for hire," a representative of the employer may send an equivalent form.

[redacted]
Signature

Nils Sturma
Name

11-11-2022
Date

B.2 Algebraic Sparse Factor Analysis

Summary

In this article, we derive novel results on algebro-geometric aspects of sparse factor analysis models by adopting a graphical perspective. We prove upper and lower bounds for the dimension of the models. While the upper bound holds for arbitrary sparse factor analysis models, the lower bound holds for models that satisfy a minimal level of sparsity, which we formalize in the "ZUTA condition". In many cases, upper and lower bounds coincide and one obtains a formula for the dimension. In particular, our study reveals that sparse factor analysis models, unlike full factor analysis models, may not have expected dimension obtained from counting parameters. Moreover, we study the ideal of invariants of sparse factor models with two latent nodes. We present an ideal that cuts out the model and, moreover, we derive a Gröbner basis for models with overlap at most two, i.e., models where at most two observed nodes have more than one latent parent. On a technical level, we extended the "delightful strategy", which was previously applied to secants, to joins of ideals.

The article is structured as follows. In the introduction, we motivate sparsity in factor analysis models by discussing related work, and we formally introduce the setup. In Section 2, we study the dimension of factor analysis models. First, we recall why full factor analysis models are always of expected dimension. Then, we define the necessary combinatorial concepts and prove our upper and lower bounds on the dimension of sparse factor analysis models. Finally, we provide conditions on the sparsity so that the lower and upper bounds match and, moreover, we derive conditions so that sparse factor analysis models have the expected dimension. In Section 3, we study the ideal of invariants of sparse two-factor analysis models. We characterize the Zariski closure of the model, and we give an explicit description for the generators of Gröbner bases with respect to any circular term order for models with overlap at most two. It turns out that the associated initial ideal is the join of the initial edge ideals of complete graphs with isolated vertices and can therefore be realized as the monomial edge ideal of a hypergraph. We conclude the article with Section 4, where we discuss several open questions that arise from our results.

Individual contributions

I am a co-author of this article, which has the authors listed in alphabetical order. I drafted the abstract, the introduction, and the entire Section 2. In particular, I developed the theoretical concepts and the proofs for Theorem 2.9 and Theorem 2.12. Moreover, I developed and drafted the proof of Theorem 3.4. The other results in Section 3 and the writing of Section 3 were carried out by Irem Portakal. Alexandros Grosdos drafted Section 4. Mathias Drton suggested studying algebro-geometric properties of sparse factor models by taking his work on full factor models as a starting point (Drton et al., 2007). All co-authors made helpful suggestions regarding both the content and presentation of the parts for which I am responsible during regular discussions.

ALGEBRAIC SPARSE FACTOR ANALYSIS

MATHIAS DRTON, ALEXANDROS GROSDOS, IREM PORTAKAL, AND NILS STURMA

ABSTRACT. Factor analysis is a statistical technique that explains correlations among observed random variables with the help of a smaller number of unobserved factors. In traditional full factor analysis, each observed variable is influenced by every factor. However, many applications exhibit interesting sparsity patterns, that is, each observed variable only depends on a subset of the factors. In this paper, we study such sparse factor analysis models from an algebro-geometric perspective. Under mild conditions on the sparsity pattern, we examine the dimension of the set of covariance matrices that corresponds to a given model. Moreover, we study algebraic relations among the covariances in sparse two-factor models. In particular, we identify cases in which a Gröbner basis for these relations can be derived via a 2-delightful term order and joins of toric edge ideals.

1. INTRODUCTION

Factor analysis provides powerful statistical tools to analyze complex data by representing a possibly large number of dependent random variables as linear functions of a smaller number of underlying source variables, the factors. Techniques from factor analysis have found widespread application in a variety of fields, including psychology [Hor65, RWC00, CBBP93], econometrics [FFL08, ABHP16], education [SNS+06, BLR+13], and epidemiology [MMS98, dOSGdC+19].

Factor analysis models may be defined as follows. Consider an observed random vector $X = (X_v)_{v\in V}$ that is indexed by a finite set V and a vector $Y = (Y_h)_{h\in\mathcal{H}}$ of unobserved random variables, called *factors*, that is indexed by a finite set $\mathcal{H}$. In applications, the number of factors $|\mathcal{H}|$ is usually smaller than the number of observed variables $|V|$. The factor analysis model postulates that the observed variables are linear functions of the factors and noise, i.e.,

$$X = \Lambda Y + \varepsilon,$$

where $\Lambda = (\lambda_{vh}) \in \mathbb{R}^{|V|\times|\mathcal{H}|}$ is an unknown coefficient matrix, known as *factor loading matrix*. The noise $\varepsilon = (\varepsilon_v)_{v\in V}$ is comprised of independent random variables with mean zero and positive variance; so $\mathbb{E}[\varepsilon_v] = 0$ and $\mathrm{Var}[\varepsilon_v] =: \omega_{vv} \in (0,\infty)$. The latent (unobserved) factors $(Y_h)_{h\in\mathcal{H}}$ are assumed to be mutually independent, and also independent of the noise ε. Without loss of generality, we fix the scale of the latent factors such that each Y_h has mean zero and variance one. The main object of study is now the covariance matrix Σ of the observed random vector X, which is given by

$$\Sigma := \mathrm{Cov}[X] = \Lambda\Lambda^\top + \Omega, \tag{1}$$

where Ω is a diagonal matrix with entries $\omega_{vv} = \mathrm{Cov}[\varepsilon_v]$. In traditional full factor analysis, all coefficients λ_{vh} are nonzero [AR56]. Full factor analysis models were studied from a computational algebraic geometry point of view in [DSS07], where Gröbner bases were used to investigate the *ideal of invariants* that vanish on the space of covariance matrices. The generators emerge from rank conditions on the symmetric covariance matrix under elimination of the diagonal entries.

The journey of this paper extends *beyond* [DSS07], prompting a study of *sparse* factor analysis models under an algebro-geometric perspective. Recently, there has been considerable interest in sparse factor analysis models, which posit that some (or often many) of the coefficients λ_{vh} are equal to zero. Examples of recent research on sparsity include work on correlation thresholding [KZ22], l_1-penalization [LWSB14, TFA17], and Bayesian approaches [FSL18, OLK23]. Moreover,

2020 *Mathematics Subject Classification.* 62H25, 62R01, 13F65, 14M25, 14N07.
Key words and phrases. factor analysis model, dimension, join of ideals, Gröbner basis, toric, edge ideal, hypergraph.

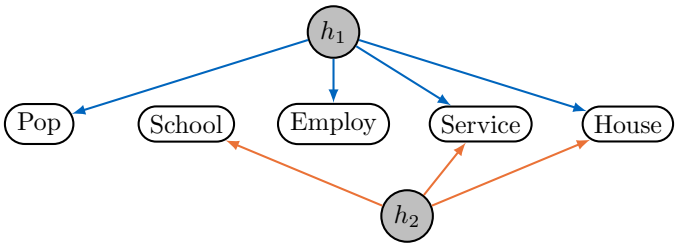


FIGURE 1. Graph corresponding to a sparse factor analysis model for a study on metropolitan districts. Gray nodes correspond to latent nodes.

sparse factor analysis models are the building block for many directed graphical models with latent variables [Bol89, BDSW22] that have applications in causality [Pea00, PJS17]. In this context, the coefficients (or factor loadings) λ_{vh} can be interpreted as causal effects of the latent variables Y_h on the observed variables X_v. To represent sparsity assumptions, it is useful to adopt a graphical perspective and encode a zero pattern in Λ by a directed graph with nodes $V \cup \mathcal{H}$ [MDLW19]. For an observed node $v \in V$ and a latent node $h \in \mathcal{H}$, the coefficient λ_{vh} is allowed to be nonzero only if the edge $h \to v$ is in the graph.

Example 1.1. We revisit a study from [Har76, p. 14] that pertains to five socio-economic variables that are observed in twelve districts in the greater Los Angeles area: Total population, median school years, total employment, miscellaneous professional services, and median house value. Applying l_1-penalization techniques to the data, a model corresponding to the graph G in Figure 1 is found by [TFA17, Table 1, Column 3]. The model imposes, for example, that total population is independent of total employment given only the first latent variable. In this model, the factor loading matrix has the zero pattern

$$\Lambda = \begin{pmatrix} \lambda_{11} & 0 & \lambda_{31} & \lambda_{41} & \lambda_{51} \\ 0 & \lambda_{22} & 0 & \lambda_{42} & \lambda_{52} \end{pmatrix}^\top,$$

and gives rise to the covariance matrix $\Sigma \in F(G)$ of the form

$$\Sigma = (\sigma_{uv}) = \begin{pmatrix} \omega_{11} + \lambda_{11}^2 & 0 & \lambda_{11}\lambda_{31} & \lambda_{11}\lambda_{41} & \lambda_{11}\lambda_{51} \\ 0 & \omega_{22} + \lambda_{22}^2 & 0 & \lambda_{22}\lambda_{42} & \lambda_{22}\lambda_{52} \\ \lambda_{11}\lambda_{31} & 0 & \omega_{33} + \lambda_{31}^2 & \lambda_{31}\lambda_{41} & \lambda_{31}\lambda_{51} \\ \lambda_{11}\lambda_{41} & \lambda_{22}\lambda_{42} & \lambda_{31}\lambda_{41} & \omega_{44} + \lambda_{41}^2 + \lambda_{52}^2 & \lambda_{41}\lambda_{51} + \lambda_{42}\lambda_{52} \\ \lambda_{11}\lambda_{51} & \lambda_{22}\lambda_{52} & \lambda_{31}\lambda_{51} & \lambda_{41}\lambda_{51} + \lambda_{42}\lambda_{52} & \omega_{55} + \lambda_{51}^2 + \lambda_{52}^2 \end{pmatrix}.$$

The factor analysis model $F(G)$ is 12-dimensional which is equal to the expected dimension obtained from counting parameters. However, we will show in this paper that the dimension of sparse factor models is not always equal to the number of parameters. The (toric) ideal of variants $I(G)$ is generated by two monomials and one binomial:

$$\langle \sigma_{12}, \sigma_{23}, \sigma_{15}\sigma_{34} - \sigma_{14}\sigma_{35} \rangle.$$

The binomials of this form are known as *tetrads* in statistics which reflects the fact that the polynomial arises with four observed random variables. They also arise as generators of the ideal of invariants for one-factor analysis models or, equivalently, as the toric edge ideals of complete graphs [Sul09]; compare also [Sul08, Cor. 6.5]. Harman [Har76, p.77] highlighted the absence of knowledge regarding the ideal of invariants for models involving two or more factors. This was subsequently addressed in the context of full factor analysis models in [DSS07]. The ideal of invariants can enhance useful statistics for testing goodness-of-fit; see, e.g., [BT00, SSGS06, DSS07, DX16, SDL22]. In this paper, we address this gap of knowledge for sparse factor analysis models.

The organization and the main results of the paper are as follows: In Section 2, we study the dimension of factor analysis models. First, we give a general upper bound on the dimension in Theorem 2.9, which reveals that sparse factor models can be defective, that is, they may not have expected dimension. This is a difference to full factor analysis models that are always of expected dimension. For models that exhibit a minimal level of sparsity, which we call the zero upper triangular assumption (ZUTA), we also provide a lower bound on the dimension in Theorem 2.12. The assumption ensures that the rows and columns of the matrix Λ can be permuted such that the upper triangle of the matrix is zero. In many cases, the upper and lower bounds coincide, and we obtain a combinatorial formula for the dimension.

In Section 3, we study the ideal of invariants of sparse *two*-factor analysis models. First, we characterize the Zariski closure of the model in Theorem 3.4. Second, we give an explicit description for the generators of Gröbner bases with respect to any circular term order for a subclass of sparse two-factor analysis models in Theorem 3.13. It turns out that the associated initial ideal can be realized as the monomial edge ideal of a hypergraph. The study of this Gröbner basis uses the 2-delightful strategy that was introduced in [SS05] for secant varieties. We generalize this strategy to joins of sparse one-factor analysis models, i.e., we study joins of toric edge ideals of complete graphs with isolated vertices. Supplementary code for our results can be found on the MathRepo page:

https://mathrepo.mis.mpg.de/sparse-factor-analysis

2. Dimension

Let $G = (V \cup \mathcal{H}, D)$ be a directed graph, where V and $\mathcal{H}$ are finite disjoint sets of observed and latent nodes. We assume that the graph G only contains edges that point from latent to observed nodes, that is, $D \subseteq \mathcal{H} \times V$; see Figure 2 for an example with $H = \{h_1, h_2\}$ and $V = \{v_1, \dots, v_7\}$. We refer to such bipartite graphs as *factor analysis graphs*. If $(h, v) \in D$, which we also denote by $h \to v \in D$, then $h \in \mathcal{H}$ is a parent of its child $v \in V$. The respective sets of all parents and children are denoted by $\mathrm{pa}(v) = \{h \in \mathcal{H} : h \to v \in D\}$ and $\mathrm{ch}(h) = \{v \in V : h \to v \in D\}$.

Every factor analysis graph determines a factor analysis model that for our purposes may conveniently be identified with the set of its covariance matrices. For a definition, we let $\mathbb{R}^D$ denote the set of real $|V| \times |\mathcal{H}|$ matrices $\Lambda = (\lambda_{vh})$ with support D, that is, $\lambda_{vh} = 0$ if $h \to v \notin D$. Furthermore, we write $\mathrm{PD}(p)$ for the cone of positive definite $p \times p$ matrices, and $\mathbb{R}^p_{>0} \subset \mathrm{PD}(p)$ for the subset of diagonal positive definite matrices.

Definition 2.1. *Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph with $|V| = p$ and $|\mathcal{H}| = m$. As a model of the covariance matrix, the factor analysis model determined by G is the image $F(G) = \mathrm{Im}(\tau_G)$ of the parametrization map*

$$\begin{aligned} \tau_G : \mathbb{R}^p_{>0} \times \mathbb{R}^D &\longrightarrow \mathrm{PD}(p) \\ (\Omega, \Lambda) &\longmapsto \Omega + \Lambda\Lambda^\top. \end{aligned} \tag{2}$$

The covariance model $F(G)$ is a parameterized subset of the $\binom{p+1}{2}$-dimensional space of symmetric $p \times p$ matrices, and its dimension is the maximal rank of the Jacobian matrix of the map τ_G in Definition 2.1. Naturally, the *expected dimension* of $F(G)$ is equal to $\min\{|V| + |D|, \binom{|V|+1}{2}\}$, the minimum of the number of parameters in (Ω, Λ) and the dimension of the ambient space.

Example 2.2. The graph in Figure 2 corresponds to a sparse model with $|V| = 7$ nodes and $|D| = 9$ edges. To simplify notation, we identify $v_1, \dots, v_7$ with the integers $1, \dots, 7$, and h_1, h_2 with the integers $1, 2$. Then, the sparse factor loading matrix is

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{21} & \lambda_{31} & \lambda_{41} & \lambda_{51} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{42} & \lambda_{52} & \lambda_{62} & \lambda_{72} \end{pmatrix}^\top, \tag{3}$$

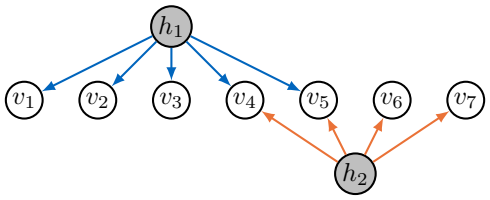


FIGURE 2. Factor analysis graph with 2 latent factors and 7 observed variables.

which gives rise to the covariance matrix $\Sigma \in F(G)$ of the form

$$\Sigma = \begin{pmatrix} \omega_{11}+\lambda_{11}^2 & \lambda_{11}\lambda_{21} & \lambda_{11}\lambda_{31} & \lambda_{11}\lambda_{41} & \lambda_{11}\lambda_{51} & 0 & 0 \\ \lambda_{11}\lambda_{21} & \omega_{22}+\lambda_{21}^2 & \lambda_{21}\lambda_{31} & \lambda_{21}\lambda_{41} & \lambda_{21}\lambda_{51} & 0 & 0 \\ \lambda_{11}\lambda_{31} & \lambda_{21}\lambda_{31} & \omega_{33}+\lambda_{31}^2 & \lambda_{31}\lambda_{41} & \lambda_{31}\lambda_{51} & 0 & 0 \\ \lambda_{11}\lambda_{41} & \lambda_{21}\lambda_{41} & \lambda_{31}\lambda_{41} & \omega_{44}+\lambda_{41}^2+\lambda_{42}^2 & \lambda_{41}\lambda_{51}+\lambda_{42}\lambda_{52} & \lambda_{42}\lambda_{62} & \lambda_{42}\lambda_{72} \\ \lambda_{11}\lambda_{51} & \lambda_{21}\lambda_{51} & \lambda_{31}\lambda_{51} & \lambda_{41}\lambda_{51}+\lambda_{42}\lambda_{52} & \omega_{55}+\lambda_{51}^2+\lambda_{52}^2 & \lambda_{52}\lambda_{62} & \lambda_{52}\lambda_{72} \\ 0 & 0 & 0 & \lambda_{42}\lambda_{62} & \lambda_{52}\lambda_{62} & \omega_{66}+\lambda_{62}^2 & \lambda_{62}\lambda_{72} \\ 0 & 0 & 0 & \lambda_{42}\lambda_{72} & \lambda_{52}\lambda_{72} & \lambda_{62}\lambda_{72} & \omega_{77}+\lambda_{72}^2 \end{pmatrix}.$$

The expected dimension of the corresponding model is equal to $|V| + |D| = 15$, and as we verify in Theorem 2.9, this is indeed the dimension of the model.

If $G = (V \cup \mathcal{H}, D)$ is a factor analysis graph with all possible edges, so $D = \mathcal{H} \times V$, then the corresponding covariance model recovers the *full* factor analysis model [DSS07, AR56]. However, using orthogonal transformations as in the QR decomposition, any covariance matrix Σ in a full factor analysis model can be written as $\Sigma = \Omega + \Lambda\Lambda^\top$ such that the upper triangle of $\Lambda = (\lambda_{vh})$ is zero. Hence, any full factor model is equivalent to a sparse factor analysis model where only the edges corresponding to the upper triangle in Λ are removed from the complete bipartite graph. Said differently, we obtain a graph that belongs to the set of factor analysis graphs satisfying the following assumption.

Assumption (ZUTA). *A factor analysis graph and its associated model satisfy the* Zero Upper Triangular Assumption (ZUTA) *if there exists a relabeling of the latent nodes $\mathcal{H} = \{h_1, \dots, h_m\}$ such that $\mathrm{ch}(h_i)$ is not contained in $\bigcup_{j>i} \mathrm{ch}(h_j)$ for all $i = 1, \dots, m$. In this case, there is then a relabeling of the observed nodes $V = \{v_1, \dots, v_p\}$ such that $v_i \in \mathrm{ch}(h_i)$ and $v_i \notin \bigcup_{j>i} \mathrm{ch}(h_j)$ for all $i = 1, \dots, m$.*

ZUTA ensures that the rows and columns of the factor loading matrix Λ can be permuted such that the upper triangle of the matrix is zero.

Example 2.3. The graph in Figure 2 satisfies ZUTA. The latent nodes h_1 and h_2 are already ordered as desired. A ZUTA labeling of V is obtained if we permute, for example, the labelings of v_2 and v_4. This corresponds to permuting rows 2 and 4 of the parameter matrix Λ in Equation (3).

Note that ZUTA requires that $p \geq m$ and that each latent node has at least one observed child. However, isolated latent nodes need not be considered as they only add a zero column in Λ.

Remark 2.4. In the special case where a factor analysis graph contains an observed node $v \in V$ such that $\mathrm{pa}(v) = \emptyset$, the dimension of the model is by one larger than the dimension of the model corresponding to the smaller graph where this node is removed.

Remark 2.5. ZUTA is more general than the "k-pure-children" condition that is often employed in previous work on structure identifiability of sparse factor analysis models [AGM12, BBNW20, MSWB22, MLAS23]. The k-pure-children condition requires that each latent node $h \in \mathcal{H}$ has at least k (pure) children that have no other parents than h. In particular, the 1-pure child condition implies that there is an upper $m \times m$ matrix inside Λ which is diagonal. Hence, any k-pure children condition

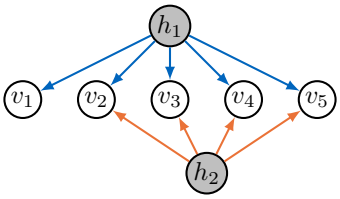


FIGURE 3. A graph whose associated model is equal to the full two-factor analysis model with 5 observed nodes.

with $k \geq 1$ implies ZUTA. Note that ZUTA also requires that each latent node $h \in \mathcal{H}$ has at least 1 child, but the children are allowed to have more parents. For example, after relabeling such that ZUTA is satisfied, node v_3 needs to be a child of h_3, but it could also have h_1 and h_2 as parents. Only the first node v_1 has to be a *pure* child of h_1. Conversely, a factor analysis graph in which there is no latent node that has a pure child does not satisfy ZUTA.

It is proved in [DSS07] that the dimension of full factor analysis models is always equal to the expected dimension obtained by counting parameters. Since full factor models are equivalent to models satisfying ZUTA, the number of edges is equal to $|D| = pm - \binom{m}{2}$, which implies that the expected dimension is given by $\min\{p(m+1) - \binom{m}{2}, \binom{p+1}{2}\}$.

Example 2.6. Consider the full factor analysis model with $m = 2$ latent nodes and $p = 5$ observed nodes. The model is equivalent to the model $F(G)$ corresponding to the graph in Figure 3, and the dimension is equal to the number of parameters, that is, $\dim(F(G)) = 14$.

In Example 2.2, we saw a *sparse* model that is also of expected dimensions. However, sparse factor analysis models differ fundamentally from full factor analysis models in the sense that their dimension is not always equal to the expected dimension. The next example shows models where the dimension drops, that is, the dimension is strictly smaller than the expected dimension.

Example 2.7. Consider the three graphs in Figure 4. The expected dimension of the model $F(G)$ corresponding to graph (a) is $|V| + |D| = 14$. On the other hand, the model is a subset of the space of symmetric matrices that has dimension $\binom{p+1}{2} = 15$, and every covariance matrix $\Sigma = (\sigma_{vw}) \in F(G)$ has three zeros, $\sigma_{v_1v_4} = \sigma_{v_1v_5} = \sigma_{v_2v_5} = 0$. Thus, we obtain $15 - 3 = 12$ as a trivial upper bound for the dimension. It turns out that we have indeed $\dim(F(G)) = 12$. However, the model corresponding to graph (b), obtained by adding one more node, shows that counting zeros in the covariance matrix is not enough. In this case, we have $\binom{p+1}{2} = 21$ and there are five zeros in every covariance matrix in the model, namely $\sigma_{v_1v_4} = \sigma_{v_1v_5} = \sigma_{v_1v_6} = \sigma_{v_2v_5} = \sigma_{v_2v_6} = 0$. Thus we obtain an upper bound of 16 for the dimension that is also equal to the expected dimension $|V| + |D|$. Nevertheless, the true dimension is given by $\dim(F(G)) = 15$. The model corresponding to graph (c) has a similar drop of dimension. In this case there are no zeros in the covariance matrix and the expected dimension is $|V| + |D| = 18$, but the true dimension is $\dim(F(G)) = 17$.

To study the dimension of sparse factor analysis models, we first introduce necessary terminology. Let $C(V, 2) := \{\{v, w\} : v, w \in V, v \neq w\}$ be the set of 2-pairs of V, i.e., the set of all subsets consisting of 2 distinct nodes of V. We write $\mathrm{jpa}(\{u, v\}) = \{h \in \mathcal{H} : h \in \mathrm{pa}(u) \cap \mathrm{pa}(v)\}$ for the set of *joint parents* of a pair $\{u, v\} \in C(V, 2)$. For any latent node $h \in \mathcal{H}$, we let $C(V, 2)_h = \{\{v, w\} \in C(V, 2) : h \in \mathrm{jpa}(\{v, w\})\}$ be the collection of pairs of nodes that have h as a joint parent.

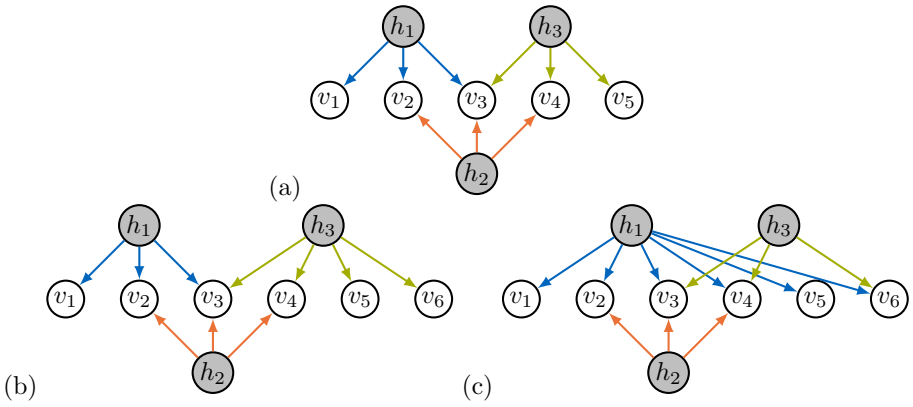


FIGURE 4. Graphs with lower model dimension than number of parameters.

For a matrix $\Sigma \in F(G)$, the parametrization of the entries σ_{uv} depends on the joint parents of the pair $\{u, v\}$. In particular, for $\Omega = \mathrm{diag}(\omega_{vv}) \in \mathbb{R}^p_{>0}$ and $\Lambda = (\lambda_{vh}) \in \mathbb{R}^D$, we have

$$\sigma_{uv} = \begin{cases} \sum_{h \in \mathrm{jpa}(\{u,v\})} \lambda_{uh}\lambda_{vh} & \text{if } u \neq v, \\ \omega_{uu} + \sum_{h \in \mathrm{pa}(u)} \lambda_{uh}^2 & \text{if } u = v, \end{cases}$$

where we use the convention that the empty sum is zero. Recall that the dimension of $F(G) = \mathrm{Im}(\tau_G)$ is equal to the maximal rank of the Jacobian of τ_G. Hence, we need to study the Jacobian matrix that has the form

$$J = \begin{matrix} \\ u \\ \{u,v\} \end{matrix} \begin{matrix} \omega & \lambda \\ \end{matrix} \begin{pmatrix} I_p & C \\ 0 & B \end{pmatrix},$$

where the rows in the upper part correspond to the derivatives of σ_{uu} and the rows in the lower part correspond to the derivatives of σ_{uv} for $u \neq v$. In particular, the entries in the unit matrix I_p on the upper left are given by

$$\frac{\partial \sigma_{uu}}{\partial \omega_{vv}} = \begin{cases} 1 & \text{if } u = v, \\ 0 & \text{else.} \end{cases}$$

Thus, the rank of the Jacobian is equal to $p + \mathrm{rank}(B)$; recall that $p = |V|$. The entries of the matrix B are given by

$$\frac{\partial \sigma_{uv}}{\partial \lambda_{zh}} = \begin{cases} \lambda_{vh} & \text{if } z = u \text{ and } h \in \mathrm{jpa}(\{u, v\}), \\ \lambda_{uh} & \text{if } z = v \text{ and } h \in \mathrm{jpa}(\{u, v\}), \\ 0 & \text{else.} \end{cases} \tag{4}$$

Note that the rows of B are indexed by 2-pairs $\{u, v\} \in C(V, 2)$. A necessary condition for a model to have expected dimension is the crucial observation that, for each latent node h, there has to be a *different* set of 2-pairs of children of h that has same cardinality as the number of children of h. Otherwise, the dimension drops accordingly. We formalize the concept of different 2-pairs, a.k.a rows of B, by considering pairwise disjoint collections.

Definition 2.8. *Let $G = (V \cup \mathcal{H})$ be a factor analysis graph and let $\mathcal{A} = (A_h)_{h \in \mathcal{H}}$ be a collection of observed 2-pairs, that is, $A_h \subseteq C(V, 2)$. We say that the collection $\mathcal{A}$ is* valid *if*

(i) *$A_h \subseteq C(V, 2)_h$ with cardinality $|A_h| \leq |\,\mathrm{ch}(h)|$ for all $h \in \mathcal{H}$, and*
(ii) *the collection is pairwise disjoint, i.e., $A_h \cap A_\ell = \emptyset$ for $h \neq \ell$.*

Moreover, we say that $\sum_{h \in \mathcal{H}} |A_h|$ is the sum of cardinalities *of a valid collection.*

The next theorem gives an upper bound on the dimension. It is obtained by choosing a valid collection $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ such that the sum of cardinalities $\sum_{h\in\mathcal{H}} |A_h|$ is maximal. The upper bound holds for all sparse factor models, even if ZUTA is not satisfied.

Theorem 2.9. *Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph. Let $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ be a valid collection of 2-pairs such that the sum of cardinalities $\sum_{h\in\mathcal{H}} |A_h|$ is maximal among all valid collections. Then*

$$\dim(F(G)) \leq |V| + \sum_{h\in\mathcal{H}} |A_h|.$$

Proof. It is enough to show that $\operatorname{rank}(B) \leq \sum_{h\in\mathcal{H}} |A_h|$. Define $\lambda_i := (\lambda_{\operatorname{ch}(h_i),h_i}) \in \mathbb{R}^{|\operatorname{ch}(h_i)|}$ and $\mathcal{A}^{\complement} := C(V,2) \setminus \left(\bigcup_{h\in\mathcal{H}} A_h\right)$. Then the matrix B can be written as

$$B = \begin{array}{c} \\ A_{h_1} \\ \vdots \\ A_{h_m} \\ \mathcal{A}^{\complement} \end{array} \begin{array}{c} \begin{array}{ccc} \lambda_1 & \cdots & \lambda_m \end{array} \\ \left(\begin{array}{ccc} B_{1,1} & \cdots & B_{1,m} \\ \vdots & & \vdots \\ B_{m,1} & \cdots & B_{m,m} \\ \hline B_{\mathcal{A}^{\complement},1} & \cdots & B_{\mathcal{A}^{\complement},m} \end{array}\right) \end{array}.$$

The proof is structured as follows. We first show in Claim 1 and Claim 3 that some submatrices of B are equal to zero. Claim 2 is an intermediate result we need to prove for Claim 3. Then, we restructure the matrix B and show in Claim 4 that the rank of the matrix B can not be larger than $\sum_{h\in\mathcal{H}} |A_h|$. Let $[m] := \{1,\dots,m\}$ and define the index sets $I^{(=)} = \{i \in [m] : |A_{h_i}| = |\operatorname{ch}(h_i)|\}$ and $I^{(<)} = [m] \setminus I^{(=)}$.

Claim 1: If $i \in I^{(<)}$, then $B_{\mathcal{A}^{\complement},i} = 0$.

Consider an index $i \in I^{(<)}$ and a row indexed by $\{u,z\} \in \mathcal{A}^{\complement}$. Observe that we must have $h_i \notin \operatorname{pa}(u) \cap \operatorname{pa}(z)$. Otherwise we could have chosen $\widetilde{A}_{h_i} = A_{h_i} \cup \{\{u,z\}\}$ that has empty intersection with any A_{h_j} for $j \neq i$. But this defines another valid collection $\widetilde{\mathcal{A}} = (A_{h_1},\dots,A_{h_{i-1}},\widetilde{A}_{h_i},A_{h_{i+1}},\dots,A_{h_m})$ such that the sum of cardinalities is greater by one, which contradicts the assumption on the maximality of $\mathcal{A}$. We conclude that the row in $B_{\mathcal{A}^{\complement},i}$ that is indexed by $\{u,z\}$ is equal to zero; recall (4). Since this holds for all rows $\{u,z\} \in \mathcal{A}^{\complement}$, we have that $B_{\mathcal{A}^{\complement},i} = 0$, which proves the claim.

To state Claim 2 we define

$$J^0 = \{i \in I^{(=)} : B_{\mathcal{A}^{\complement},i} \neq 0\} = \{i \in I^{(=)} : h_i \in \operatorname{jpa}(R) \text{ for some } R \in \mathcal{A}^{\complement}\},$$

and, for all $k \geq 1$, we define

$$J^k = \{j \in I^{(=)} : \text{there is } i \in J^{k-1} \text{ such that } h_j \in \operatorname{jpa}(R) \text{ for some } R \in A_{h_i}\}.$$

Since $h_j \in \operatorname{jpa}(R)$ for all $R \in A_{h_j}$, we clearly have that $J^k \subseteq J^{k+1}$ for all $k \geq 0$.

Claim 2: Let $R \in A_{h_j}$ for some $j \in J^k$, $k \geq 0$. Then $\{i \in [m] : h_i \in \operatorname{jpa}(R)\} \subseteq I^{(=)}$.

We first assume $k = 0$. Let $j \in J^0$ and $R \in A_{h_j}$, and suppose there is $h_l \in \operatorname{jpa}(R)$ such that $l \in I^{(<)}$. On the one hand, this means that $|A_{h_l}| < |\operatorname{ch}(h_l)|$. On the other hand, since $j \in J^0$, there has to be a pair $S \in \mathcal{A}^{\complement}$ such that $h_j \in \operatorname{jpa}(S)$. Therefore, we can define a collection $\widetilde{\mathcal{A}} = (\widetilde{A}_h)_{h\in\mathcal{H}}$ such that $\widetilde{A}_{h_j} = (A_{h_j} \setminus \{R\}) \cup \{S\}$, $\widetilde{A}_{h_l} = A_{h_l} \cup \{R\}$ and $\widetilde{A}_h = A_h$ for all $h \notin \{h_j,h_l\}$. Note that the collection $\widetilde{\mathcal{A}}$ is valid but $\sum_{h\in\mathcal{H}} |\widetilde{A}_h| = 1 + \sum_{h\in\mathcal{H}} |A_h|$. This is a contradiction to the maximality assumption on the collection $\mathcal{A}$ and we conclude that we must have $l \in I^{(=)}$.

Now, let $k \geq 1$ and assume that $j \in J^k$ and $R \in A_{h_j}$. If $j \in J^0$, we are done. If $j \in J^k \setminus J^0$, suppose there is $h_l \in \operatorname{jpa}(R)$ such that $l \in I^{(<)}$. Once again, this means that $|A_{h_l}| < \operatorname{ch}(h_l)$. Now,

we recursively choose integers $j_1, \dots, j_n$ and corresponding subsets $S_{j_i} \in A_{h_{j_i}}$ as follows. First, since $j \in J^k$, there has to be a minimal integer $k_1 < k$ such that there is $j_1 \in J^{k_1}$ and $h_j \in \mathrm{jpa}(S_{j_1})$ for some $S_{j_1} \in A_{h_{j_1}}$. Note that $j \notin J^{k_1}$, since otherwise there would exist $\widetilde{j}_1 \in J^{k_1-1}$ such that $h_j \in \mathrm{jpa}(\widetilde{S})$ for some $\widetilde{S} \in A_{h_{\widetilde{j}_1}}$, which is a contradiction on the minimality assumption on k_1. Further, define k_{i+1} as the minimal integer $k_{i+1} < k_i$ such that there is $j_{i+1} \in J^{k_{i+1}}$ and $h_{j_i} \in \mathrm{jpa}(S_{j_{i+1}})$ for some $S_{j_{i+1}} \in A_{h_{j_{i+1}}}$. We stop this procedure as soon as we arrived at some $n \geq 1$ such that $k_n = 0$. It can be seen as before that $j_i \notin J^{k_{i+1}}$ for all $i = 1, \dots, n$. Hence, the integers $j, j_1, \dots, j_n$ are pairwise different by construction, which also implies that the pairs $R, S_{j_1}, \dots, S_{j_n}$ are pairwise different. Moreover, since $j_n \in J^0$, there has to be a pair $S \in \mathcal{A}^{\complement}$ such that $h_{j_n} \in \mathrm{jpa}(S)$. Now, we define a collection $\widetilde{\mathcal{A}} = (\widetilde{A}_h)_{h \in \mathcal{H}}$ as follows:

$$\widetilde{A}_{h_{j_n}} = (A_{h_{j_n}} \setminus \{S_{j_n}\}) \cup \{S\}, \qquad \widetilde{A}_{h_{j_i}} = (A_{h_{j_i}} \setminus \{S_{j_i}\}) \cup \{S_{j_{i+1}}\} \text{ for } i = 1, \dots n-1,$$
$$\widetilde{A}_{h_j} = (A_{h_j} \setminus \{R\}) \cup \{S_{j_1}\}, \qquad \widetilde{A}_{h_l} = A_{h_l} \cup \{R\},$$

and $\widetilde{A}_h = A_h$ for all $h \in \mathcal{H}$ that do not appear above. Since the pairs $R, S_{j_1}, \dots, S_{j_n}$ and S are pairwise different, the collection is valid. However, we have that $\sum_{h \in \mathcal{H}} |\widetilde{A}_h| = 1 + \sum_{h \in \mathcal{H}} |A_h|$, which is a contradiction to the maximality assumption on the collection $\mathcal{A}$. We conclude that we must have $l \in I^{(=)}$, which proves the claim.

Now, observe that there must exist a $k^* \geq 0$ such that the sequence $J^0 \subseteq J^1 \subseteq \dots$ stabilizes, that is, $J^{k^*} = J^{k^*+1} = \dots$. This is true since $J^{k-1} \subseteq J^k$ and $J^k \subseteq I^{(=)}$ for all $k \geq 1$. Define $\overline{J} := J^{k^*}$.

Claim 3: $B_{j,i} = 0$ for all $j \in \overline{J}, i \in [m] \setminus \overline{J}$.

Consider indices $j \in \overline{J}$ and let $R \in A_{h_j}$. It is enough to show that $\{i \in [m] : h_i \in \mathrm{jpa}(R)\} \subseteq \overline{J}$. By Claim 2, we have that $\{i \in [m] : h_i \in \mathrm{jpa}(R)\} \subseteq I^{(=)}$. Now, assume that there is $h_l \in \mathrm{jpa}(R)$ such that $l \notin \overline{J}$. By definition, this means that $l \in J^{k^*+1}$. But this is a contradiction since $\overline{J} = J^{k^*+1}$. We conclude that we must have $l \in \overline{J}$.

Claim 4: The rank of the matrix B cannot exceed $\sum_{h \in \mathcal{H}} |A_h|$.

Without loss of generality, $\overline{J} = [k]$ for some positive integer $k \leq m$. Then, by Claims 1-3, the matrix B has the form

$$B = \begin{array}{c} \\ A_{h_1} \\ \vdots \\ A_{h_k} \\ \mathcal{A}^{\complement} \\ A_{h_{k+1}} \\ \vdots \\ A_{h_m} \end{array} \begin{array}{c} \begin{array}{cccc} \lambda_1 & \cdots & \lambda_k & \end{array} \\ \left(\begin{array}{ccc|ccc} B_{1,1} & \cdots & B_{1,k} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ B_{k,1} & \cdots & B_{k,k} & & & \\ B_{\mathcal{A}^{\complement},1} & \cdots & B_{\mathcal{A}^{\complement},k} & 0 & \cdots & 0 \\ \hline B_{k+1,1} & \cdots & B_{k+1,k} & B_{k+1,k+1} & \cdots & B_{k+1,m} \\ \vdots & & \vdots & \vdots & & \vdots \\ B_{m,1} & \cdots & B_{m,k} & B_{m,k+1} & \cdots & B_{m,m} \end{array}\right) \end{array} \begin{array}{c} \\ \lambda_{k+1} \\ \vdots \\ \lambda_m \end{array}.$$

The rank of this matrix is smaller or equal to the sum of the minimum of the number of rows and columns of the upper left block plus the minimum of the number of rows and columns of the lower right block. The minimum of the number of rows and columns of the upper left block is given by $\min\{\sum_{h \in \overline{J}} |A_h| + |\mathcal{A}^{\complement}|, \sum_{h \in \overline{J}} |\mathrm{ch}(h)|\}$. Since $\sum_{h \in \overline{J}} |\mathrm{ch}(h)| = \sum_{h \in \overline{J}} |A_h|$, this minimum is equal to $\sum_{h \in \overline{J}} |A_h|$. On the other hand, the minimum of the number of rows and columns of the lower right block is given by $\min\{\sum_{h \in \overline{J}^{\complement}} |A_h|, \sum_{h \in \overline{J}^{\complement}} |\mathrm{ch}(h)|\}$. Since $\sum_{h \in \overline{J}^{\complement}} |A_h| \leq \sum_{h \in \overline{J}^{\complement}} |\mathrm{ch}(h)|$, this

minimum is given by $\sum_{h\in\overline{J}^{\complement}} |A_h|$. Thus, the rank of the matrix B cannot be larger than

$$\sum_{h\in\overline{J}} |A_h| + \sum_{h\in\overline{J}^{\complement}} |A_h| = \sum_{h\in\mathcal{H}} |A_h|. \qquad \square$$

Example 2.10. Consider the graph in Figure 4 (b). Then we have

$$\begin{aligned} C(V,2)_{h_1} &= \{\{v_1,v_2\},\{v_1,v_3\},\{v_2,v_3\}\},\\ C(V,2)_{h_2} &= \{\{v_2,v_3\},\{v_2,v_4\},\{v_3,v_4\}\},\\ C(V,2)_{h_3} &= \{\{v_3,v_4\},\{v_3,v_5\},\{v_3,v_6\},\{v_4,v_5\},\{v_4,v_6\},\{v_5,v_6\}\}. \end{aligned}$$

To obtain an upper bound for the dimension, we want to choose the subsets $A_{h_i} \subseteq C(V,2)_{h_i}$ with cardinality as large as possible but not larger than the number of children. However, to obtain a valid, i.e. pairwise disjoint, collection we have to choose either $|A_{h_1}| = 2$ or $|A_{h_2}| = 2$. If both $|A_{h_1}| = 3$ and $|A_{h_2}| = 3$, then we must have that $\{v_2,v_3\} \in A_{h_1} \cap A_{h_2}$, i.e., the collection is not pairwise disjoint. On the other hand, we can choose A_{h_3} with cardinality at most 4, e.g. $\{\{v_3,v_5\},\{v_3,v_6\},\{v_4,v_6\},\{v_5,v_6\}\}$ that does not intersect with any of A_{h_1} and A_{h_2}. Thus, any pairwise disjoint collection $\mathcal{A} = (A_{h_1}, A_{h_2}, A_{h_3})$ with $|A_{h_i}| \leq |\operatorname{ch}(h_i)|$ has a maximal sum of cardinalities equal to $2+3+4 = 9$. Applying the upper bound in Theorem 2.9, we obtain that $\dim(F(G))) \leq 6+9 = 15$ which is strictly less than the expected dimension 16.

While Theorem 2.9 holds for any sparse factor analysis graph, also for graphs that do not satisfy ZUTA, to obtain a lower bound on the dimension, we consider more refined collections of 2-pairs that require ZUTA to be satisfied. If ZUTA is satisfied, we can assume that the latent nodes are labeled as $\mathcal{H} = \{h_1,\dots,h_m\}$ and the observed nodes are labeled as $V = \{v_1,\dots,v_p\}$ such that $v_i \in \operatorname{ch}(h_i)$ and $v_i \notin \bigcup_{j>i} \operatorname{ch}(h_j)$ for all $i = 1,\dots,m$.

Definition 2.11. *Suppose that ZUTA is satisfied. A valid collection $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ of 2-pairs is* ZUTA-conform *if $\{v_i,w\} \in A_{h_i}$ for all $w \in \operatorname{ch}(h_i)\setminus\{v_i\}$ and for all $i \in [m]$.*

Note that a valid, ZUTA-conform collection always exists for a factor analysis graph that satisfies ZUTA. Indeed, one may just choose $A_{h_i} = \{\{v_i,w\} : w \in \operatorname{ch}(h_i)\setminus\{v_i\}\}$. In this collection, the cardinality of each set of 2-pairs A_{h_i} is equal to $|\operatorname{ch}(h_i)|-1$. However, there might exist other valid, ZUTA-conform collections where the components A_{h_i} potentially contain one more 2-pair, that is, A_{h_i} might be chosen such that its cardinality is equal to $|\operatorname{ch}(h_i)|$. Each of these ZUTA-conform collections gives a lower bound on the dimension as we prove in the next theorem.

Theorem 2.12. *Let $G = (V\cup\mathcal{H}, D)$ be a factor analysis graph. Suppose that ZUTA is satisfied and let $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ be a valid collection that is ZUTA-conform. Then,*

$$\dim(F(G)) \geq |V| + \sum_{h\in\mathcal{H}} |A_h|. \tag{5}$$

Proof. It suffices to show that, for generic parameter choices, the rank of B is larger or equal to $r = \sum_{h\in\mathcal{H}} |A_h|$. Let $[m] := \{1,\dots,m\}$ and define the index sets $I^{(=)} = \{i \in [m] : |A_{h_i}| = |\operatorname{ch}(h_i)|\}$ and $I^{(<)} = [m]\setminus I^{(=)}$ as in the proof of Theorem 2.9. Consider the sets $C_i := \{\{v_i,w\} : w \in \operatorname{ch}(h_i)\setminus\{v_i\}\} \subseteq C(V,2)_{h_i}$ that have a cardinality of at most $|\operatorname{ch}(h_i)|-1$ and are pairwise disjoint. By definition, the collection $\mathcal{A}$ is given by

$$A_{h_i} = \begin{cases} C_i \cup \{S_i\} & \text{if } i \in I^{(=)},\\ C_i & \text{if } i \in I^{(<)}, \end{cases} \tag{6}$$

where $S_i \in C(V,2)_h \setminus C_i$. Now, let $\operatorname{ch}(h_i)^- = \operatorname{ch}(h_i)\setminus\{v_i\}$ and $\lambda_i^- = \lambda_{\operatorname{ch}(h_i)^-,h_i} \in \mathbb{R}^{|\operatorname{ch}(h_i)|-1}$. Moreover, we write $\mathcal{S} = \{S_i : i \in I^{(=)}\}$ and $\mathcal{A}^{\complement} = C(V,2)\setminus\left(\bigcup_{h\in\mathcal{H}} A_h\right)$. To see that the matrix B has

generically has at least rank r, we arrange it as

$$B \;=\; \begin{array}{c} \\ C_1 \\ \vdots \\ C_m \\ \mathcal{S} \\ \mathcal{A}^\complement \end{array} \begin{array}{c} \begin{array}{cccc|ccc} \lambda_{v_1,h_1} & \cdots & \lambda_{v_m,h_m} & & \lambda_1^- & \cdots & \lambda_m^- \end{array} \\ \left(\begin{array}{ccc|ccc} \lambda_1^- & & & B_{1,1} & & \\ & \ddots & & \vdots & \ddots & \\ & & \lambda_m^- & B_{m,1} & \cdots & B_{m,m} \\ \hline & & & B_{\mathcal{S},1} & \cdots & B_{\mathcal{S},m} \\ & & & B_{\mathcal{A}^\complement,1} & \cdots & B_{\mathcal{A}^\complement,m} \end{array}\right) \end{array},$$

where void entries are zero; recall ZUTA. Now, we choose a specific matrix $\Lambda^0 = (\lambda^0_{v_i,h_j}) \in \mathbb{R}^D$ such that the rank of $B^0 = B(\Lambda^0)$ is at least $r = \sum_{h\in\mathcal{H}} |A_h| = \sum_{i\in[m]} |C_i| + |\mathcal{S}|$. The existence of such a matrix implies that the rank of B is at least r for a generic choice of Λ. We choose the entries of Λ^0 as follows. For all $i \in [m]$, we set $\lambda^0_{v_i,h_i} = 1$. Since the submatrices $B_{i,i} = \lambda_{v_i,h_i} I_{|\operatorname{ch}(h_i)|-1}$ are diagonal, this implies that the upper right block of B is of full rank $\sum_{h\in\mathcal{H}}(|\operatorname{ch}(h)|-1) = \sum_{i\in[m]} |C_i|$. The remaining non-zero entries of Λ^0 are determined as follows. For any row $S_i = \{u_i, w_i\} \in \mathcal{S}$, we set exactly two entries equal to one, namely $\lambda^0_{u_i,h_i}$ and $\lambda^0_{w_i,h_i}$. All other entries of Λ^0 remain zero. Since the matrix Λ^0 has entries in $\{0, 1\}$, the same holds for the matrix B^0.

Consider a block $B^0_{i,j}$ with $j < i$ of the upper right block of B^0. Fix any row of this block indexed by $\{v_i, z\}$ and assume that $h_j \in \operatorname{jpa}(\{v_i, z\})$. Since v_i does not appear in any pair in $\mathcal{S}$, the entry $\lambda^0_{v_i,h_j}$ must be zero. If $\lambda_{z,h_j} = 1$, we can eliminate this entry by subtracting the row indexed by $\{v_j, v_i\} \in C_j$. The relevant submatrix of B^0 is given by

$$\begin{array}{c} \\ \{v_j, v_i\} \\ \{v_i, z\} \end{array} \begin{array}{c} \begin{array}{cc} \lambda_{v_j,h_j} & \lambda_{v_i,h_j} \end{array} \\ \left(\begin{array}{c|c} \lambda^0_{v_i,h_j} & \lambda^0_{v_j,h_j} \\ 0 & \lambda^0_{z,h_j} \end{array}\right) \end{array} = \begin{array}{c} \\ \{v_j, v_i\} \\ \{v_i, z\} \end{array} \begin{array}{c} \begin{array}{cc} \lambda_{v_j,h_j} & \lambda_{v_i,h_j} \end{array} \\ \left(\begin{array}{c|c} 0 & 1 \\ 0 & 1 \end{array}\right) \end{array}.$$

No fill-in occurs in the entry indexed by $\{v_i, z\}$ and λ_{v_j,h_j} since this would only happen if both $\lambda^0_{v_i,h_j}$ and λ^0_{z,h_j} are equal to one. Hence, no fill-in occurs in the upper left block of B^0. We conclude that the upper right block of B^0 can be transformed into a diagonal matrix of size $\sum_{i\in[m]} |C_i|$, while no fill-in occurs in the upper left block of B^0.

Next, we eliminate the rows in the lower right block indexed by $\mathcal{S}$ by subtraction of rows from the upper half. Fix any row indexed $\{y, z\} \in \mathcal{S}$ and assume that $h_j \in \operatorname{jpa}(\{y, z\})$. Both λ^0_{y,h_j} and λ^0_{z,h_j} are equal to one if and only if $\{y, z\} = S_j$. In this case, eliminating these two entries creates fill-in in the left-hand side of the row indexed by $\{y, z\}$. It occurs at the entry indexed by column λ_{v_j,h_j} and is equal to -2. However, this is the only fill-in that occurs. If there is another joint latent parent $h_i \in \operatorname{jpa}(\{y, z\})$ but $\{y, z\} \neq S_i$, then at least one of λ^0_{y,h_i} and λ^0_{z,h_i} is zero, in which case no fill-in occurs in the left-hand side of the row indexed by $\{y, z\}$. This can be seen similarly to the above.

It follows that, after elimination, the submatrix of B^0 that is indexed by the rows in $\mathcal{S}$ and the columns of the left-half is of full row rank $|\mathcal{S}|$. We conclude that the rank of the matrix B^0 is at least $\sum_{i\in[m]} |C_i| + |\mathcal{S}| = r$ as we have claimed. □

There might be several relabelings of observed and latent nodes such that ZUTA is satisfied. For each of these relabelings, one might potentially obtain a different lower bound in Theorem 2.12. Thus the best lower bound is obtained by maximizing the sum of cardinalities over all ZUTA-conform, valid collections and over all relabelings such that ZUTA is satisfied. If the maximized lower bound from Theorem 2.12 coincides with the upper bound in Theorem 2.9, we obtain a formula for the dimension.

Corollary 2.13. *Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph and suppose that ZUTA is satisfied. If there is a ZUTA-conform, valid collection $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ that has maximal sum of cardinalities $\sum_{h\in\mathcal{H}} |A_h|$ among all valid collections (which are not necessarily ZUTA-conform), then*

$$\dim(F(G)) = |V| + \sum_{h\in\mathcal{H}} |A_h|.$$

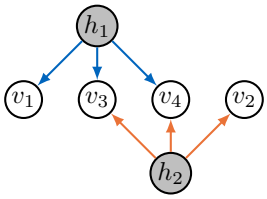


FIGURE 5. A graph where each latent nodes has exactly one pure child.

Example 2.14. Consider the graph in Figure 4 (b). In Example 2.10 we have seen that $\dim(F(G))) \leq 6 + 9 = 15$ by considering the valid collection

$$\begin{aligned} A_{h_1} &= \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}\}, \\ A_{h_2} &= \{\{v_2, v_4\}, \{v_3, v_4\}\}, \\ A_{h_3} &= \{\{v_3, v_5\}, \{v_3, v_6\}, \{v_4, v_6\}, \{v_5, v_6\}\}. \end{aligned}$$

that is of maximal sum of cardinalities. Observe that after swapping the labels of v_3 and v_6, the chosen collection is also ZUTA-conform. Hence, we obtain equality and it holds $\dim(F(G)) = 15$.

If there is one pure child per latent node, the dimension formula from Corollary 2.13 always holds.

Corollary 2.15. *Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph. Suppose that for every latent node $h \in \mathcal{H}$, there is an observed node $v \in V$ such that $\mathrm{pa}(v) = \{h\}$. Let $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ be a valid collection that has maximal sum of cardinalities $\sum_{h\in\mathcal{H}} |A_h|$ among all valid collections. Then*

$$\dim(F(G)) = |V| + \sum_{h\in\mathcal{H}} |A_h|.$$

Proof. Let $\mathcal{H} = \{h_1, \ldots, h_m\}$ and relabel the observed nodes $V = \{v_1, \ldots, v_p\}$ such that v_i is a pure child of h_i, i.e., $\mathrm{pa}(v_i) = \{h_i\}$. To show the claim, it is enough by Corollary 2.13 to define another collection $\widetilde{\mathcal{A}} = (\widetilde{A}_h)_{h\in\mathcal{H}}$ that is also valid and has the same sum of cardinalities as $\mathcal{A}$, but is additionally ZUTA-conform.

As in the proof of Theorem 2.12, define the index sets $I^{(=)} = \{i \in [m] : |A_{h_i}| = |\mathrm{ch}(h_i)|\}$ and $I^{(<)} = [m] \setminus I^{(=)}$. Consider the sets $C_i := \{\{v_i, w\} : w \in \mathrm{ch}(h_i) \setminus \{v_i\}\} \subseteq C(V, 2)_{h_i}$ that have a cardinality of at most $|\mathrm{ch}(h_i)| - 1$. Moreover, they are pairwise disjoint since h_i is the only parent of v_i with the given labeling. If $i \in I^{(=)}$, observe that the intersection $\{\{u, w\} : u, w \in \mathrm{ch}(h_i)\setminus\{v_i\}\} \cap A_{h_i}$ has to be nonempty. For any pair $\{u, w\}$ in this intersection, it must hold that neither u nor w is equal to v_j for all $j \in [m] \setminus \{i\}$, since v_j is a pure child of h_j. Hence, the pair $\{u, w\}$ is not contained in any C_j. Now, we choose a pair $S_i = \{u_i, w_i\}$ from the intersection $\{\{u, w\} : u, w \in \mathrm{ch}(h_i) \setminus \{v_i\}\} \cap A_{h_i}$ for all $i \in I^{(=)}$, and we define $\widetilde{\mathcal{A}} = (\widetilde{A}_h)_{h\in\mathcal{H}}$ to be the collection given by

$$\widetilde{A}_{h_i} = \begin{cases} C_i \cup \{S_i\} & \text{if } i \in I^{(=)}, \\ C_i & \text{if } i \in I^{(<)}. \end{cases}$$

By construction, this collection is valid and ZUTA-conform. In particular, it is pairwise disjoint. Moreover, the sum of cardinalities is unchanged, that is, $\sum_{h\in\mathcal{H}} |\widetilde{A}_h| = \sum_{h\in\mathcal{H}} |A_h|$. □

Example 2.16. Consider the graph in Figure 5, where both latent nodes have exactly one pure child. We have

$$\begin{aligned} C(V, 2)_{h_1} &= \{\{v_1, v_3\}, \{v_1, v_4\}, \{v_3, v_4\}\}, \\ C(V, 2)_{h_2} &= \{\{v_2, v_3\}, \{v_2, v_4\}, \{v_3, v_4\}\}, \end{aligned}$$

and both latent variables have three children. It is easy to see that in any valid collection with maximal sum of cardinalities it must be that either $|A_{h_1}| = 2$ or $|A_{h_2}| = 2$. If both A_{h_1} and A_{h_2} have

cardinality equal to $3 = |\operatorname{ch}(h_i)|$, then we must have that $\{v_3, v_4\} \in A_{h_1} \cap A_{h_2}$, that is, the collection is not pairwise disjoint. By Corollary 2.15, the dimension is therefore given by

$$\dim(F(G)) = |V| + \left(\sum_{h\in\mathcal{H}} |\operatorname{ch}(h)|\right) - 1 = 4 + 6 - 1 = 9,$$

which is one less than the expected dimension.

The next example considers a graph, where our upper and lower bound do not coincide, even after potential relabeling of the nodes.

Example 2.17. Computations using our code on MathRepo show that the dimension of the model corresponding to the graph in Figure 6 is 35, which coincides with the expected dimension from counting parameters. It is easy to find a valid collection $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ that has sum of cardinalities $\sum_{h\in\mathcal{H}} |A_h|$ equal to the total number of children $\sum_{h\in\mathcal{H}} |\operatorname{ch}(h)| = 26$. However, there are no relabelings of the latent and observed nodes such that ZUTA is satisfied and there is a ZUTA-conform valid collection that also has the sum of cardinalities equal to the total number of children. Hence, the lower bound from Theorem 2.12 is different than the upper bound from Theorem 2.9. For example, with the labeling as displayed in Figure 6, any valid, ZUTA-conform collection has sum of cardinalities at most 23. If we permute the labels of the nodes v_5 and v_6 to the end, that is, the nodes v_5 and v_6 become v_8 and v_9, then it is possible to construct a ZUTA-conform collection of cardinalities at most 24, but this is still less than the total number of children.

By Theorem 2.12, a model has expected dimension $|V| + |D|$ if is satisfies ZUTA and there is a ZUTA-conform, valid collection $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ such that $A_h \subseteq C(V, 2)_h$ has cardinality $|\operatorname{ch}(h)|$ for all $h \in \mathcal{H}$. Hence, a trivial necessary condition for expected dimension is that each latent node has at least three children. If a latent node $h \in \operatorname{ch}(h)$ has at most two children, we have that $|C(V, 2)_h| < |\operatorname{ch}(h)|$ and thus we must have that the cardinality of A_h is strictly smaller than the number of children. For a class of factor analysis graphs that satisfy stronger sparsity conditions than ZUTA, we obtain that the dimension is always equal to the expected dimension. Providing a lower bound that also holds for graphs violating ZUTA appears to be challenging, and we have not found a feasible approach that goes beyond case-by-case studies for each graph.

Corollary 2.18. *Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph such that $|\operatorname{ch}(h)| \geq 3$ for all $h \in \mathcal{H}$. Moreover, assume that there exist relabelings of the latent and observed nodes such that $\mathcal{H} = \{h_1, \dots, h_m\}$ and $V = \{v_1, \dots, v_p\}$ and it holds that $v_{2i-1}, v_{2i} \in \operatorname{ch}(h_i)$ and $v_{2i-1}, v_{2i} \notin \bigcup_{j>i} \operatorname{ch}(h_j)$ for all $i = 1, \dots, m$. Then, we have*

$$\dim(F(G)) = |V| + |D|.$$

Proof. For every latent node h_i there are at least three children. Two of them are given by v_{2i-1} and v_{2i} and we denote an arbitrary third child by w_i. Note that the children v_{2i-1} and v_{2i} are different for every $i \in [m]$, that is $\{v_{2i-1}, v_{2i}\} \cap \{v_{2j-1}, v_{2j}\} = \emptyset$ for $i \neq j$, while the third child w_i might also be a child of some other latent node h_j. In particular, it might be that $w_i = w_j$. We define a collection $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ by

$$A_{h_i} = \{\{v_{2i-1}, w\} : w \in \operatorname{ch}(h_i) \setminus \{v_{2i-1}\}\} \cup \{\{v_{2i}, w_i\}\} \subseteq C(V, 2)_{h_i}.$$

Clearly, the collection $\mathcal{A}$ is valid. It is also ZUTA-conform if we relabel the nodes v_{2i-1} to be v_i for all $i \in [m]$. Since $|A_h| = |\operatorname{ch}(h)|$, the sum of cardinalities $\sum_{h\in\mathcal{H}} |A_h|$ is maximal and it is equal to $\sum_{h\in\mathcal{H}} |\operatorname{ch}(h)| = |D|$. □

Note that none of the graphs in Figure 4 satisfies the condition in Corollary 2.18. But the graph in Figure 2 satisfies the condition if we swap, for example, the label of nodes v_3 and v_6.

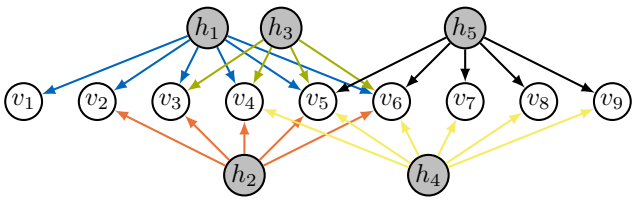


FIGURE 6. A graph where lower and upper bound on the dimension do not coincide.

Remark 2.19. One might be tempted to compare the condition in Corollary 2.18 to the often employed "2-pure-children" condition, recall Remark 2.19. The 2-pure children condition requires that each latent node $h \in \mathcal{H}$ has at least 2 children that have no other parents than h. Similarly, the condition in Corollary 2.18 also requires that each latent node has at least 2 children. However, children are generally allowed to have more than one parent. For example, after relabeling, nodes v_5 and v_6 need to be children of h_3, but h_1 and h_2 could also be parents of v_5 or v_6. The condition only requires that all h_j with $j > 3$ are not parents of v_5 and v_6. Said differently, only v_1 and v_2 are *pure* children of h_1 in the classical sense. Conversely, if there is no latent node that has two pure children in the classical sense, then the factor analysis graph does not satisfy the condition in Corollary 2.18.

3. Algebraic Invariants of the Sparse Two-Factor Model

We are interested in polynomial invariants that hold on a covariance matrix $\Sigma \in F(G)$, where G is a sparse factor analysis graph. For any subset $F \subseteq \mathrm{PD}(p)$, the *ideal of invariants* is defined as

$$I(F) = \{f \in \mathbb{R}[\sigma_{ij},\ i \leq j] : f(\Sigma) = 0 \text{ for all } \Sigma \in F\}.$$

Our object of interest is the ideal of invariants of sparse factor analysis models. Since, for a symmetric positive definite matrix $\Sigma \in \mathbb{R}^{p\times p}$, membership in $F(G)$ only depends on the off-diagonal entries of Σ, we can regard the ideal of invariants of $F(G)$, i.e., $I(G) := I(F(G))$ as an ideal in the subring $\mathbb{R}[\sigma_{ij},\ i < j]$. It is our goal to find a finite set of polynomials that generate $I(G)$. If a factor analysis graph has an edge to every observed node, the model is equivalent to a *full* factor analysis model. In the case of one or two latent nodes, the ideal of invariants is then completely understood, see [DSS07, Theorem 16] and [Sul09]. However, finding a minimal set of generators or a Gröbner basis for the full factor analysis model with three latent nodes is, to the best of our knowledge, still an open problem.

First, we consider the special case where the children sets $\mathrm{ch}(h_i)$ of a factor analysis graph $G = (V \cup \mathcal{H}, D)$ only intersect in at most one node. Our next proposition reveals that in this case the ideal of invariants is a sum of ideals obtained from induced subgraphs corresponding to full *one*-factor analysis models. Note that the ideal of the full one-factor analysis model is toric, that is, it is prime and binomial; see Theorem 3.6. In particular, it is the toric edge ideal of the complete graph on the observable node set V. Thus, we obtain that $I(G)$ is also toric.

Proposition 3.1. *Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph such that for any disjoint pair $(h_i, h_j) \in \mathcal{H}\times\mathcal{H}\backslash\{(h_1, h_m)\}$ of latent nodes, we have that $|ch(h_i)\cap ch(h_j)| \leq 1$ and $|ch(h_1)\cap ch(h_m)| = 0$. Let G_i be the induced subgraph $G[\{h_i\} \cup \mathrm{ch}(h_i)] \subseteq G$ on the vertex set $\{h_i\} \cup \mathrm{ch}(h_i)$, for $i \in [m]$. Then we obtain that $I(G) = I(G_1) + \cdots + I(G_m) + \langle \sigma_{ij} : \mathrm{pa}(i) \cap \mathrm{pa}(j) = \emptyset\rangle$ and it is toric.*

Proof. The off-diagonal entries of the parametrization τ_G given in (2) are monomial and $\sigma_{ij} = 0$, if $\mathrm{pa}(i) \cap \mathrm{pa}(j) = \emptyset$ for $i \neq j$. Moreover, the ideals of invariants of submatrices of Σ that correspond to the covariance models of induced subgraphs $G[\{h_i\} \cup \mathrm{ch}(h_i)] \subseteq G$ are toric [DSS07, Theorem 16]. Finally, since G is a polytree or a union of polytrees, the ideal is toric [ADG+23, Proposition 5.4]. □

In what follows, we focus on $|\mathcal{H}| = 2$, that is, factor analysis graphs G with two latent nodes.

3.1. **Variety.** In this section, we propose generators of an ideal such that the variety corresponding to the ideal is the smallest variety that contains the model. Recall that for any ideal $I \subseteq \mathbb{R}[x_1, \dots, x_n]$, the corresponding variety (over the complex numbers) is defined as $\mathbb{V}(I) = \{x \in \mathbb{C}^n : f(x) = 0 \text{ for all } f \in I\}$. We say that a set of latent nodes $H \subseteq \mathcal{H}$ *separates* two sets of observed nodes $A, B \subseteq V$ if $\mathrm{pa}(A) \cap \mathrm{pa}(B) \subseteq H$, where we use the notation $\mathrm{pa}(A) = \bigcup_{a \in A} \mathrm{pa}(a)$. Importantly, the set H might also be empty, that is, A and B are separated given the empty set if they do not have a joint parent. We denote the submatrix of Σ given by the rows A and columns B by $\Sigma_{A,B}$.

Definition 3.2. *Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph. The ideal $M_{\leq 1}(G) \subseteq \mathbb{R}[\sigma_{ij},\ i \leq j]$ is generated by all minors $\det(\Sigma_{A,B})$, where $A, B \subseteq V$ are two sets of observed nodes with cardinality $|A| = |B| \leq 2$ and there is $H \subseteq \mathcal{H}$ with $|H| < |A|$ such that H separates A and B.*

In other words, the ideal $M_{\leq 1}(G)$ is generated by minors $\det(\Sigma_{A,B})$ such that A and B are separated by at most one latent factor. Note that A and B are not necessarily disjoint.

Example 3.3. Consider the graph from Figure 2. Let $A = \{1\}$, $B = \{7\}$ and $H = \emptyset$. Since the nodes 1 and 7 are separated by the empty set, i.e., $\mathrm{pa}(1) \cap \mathrm{pa}(7) = \emptyset$, we have that the monomial σ_{17} is in the generating set of the ideal $M_{\leq 1}(G)$. The sets $A = \{1, 2\}$ and $B = \{4, 5\}$ are separated by $H = \{h_1\}$ since $\mathrm{pa}(A) \cap \mathrm{pa}(B) = \{h_1\}$. Thus, the minor $\sigma_{14}\sigma_{25} - \sigma_{24}\sigma_{15}$ is a generator of $M_{\leq 1}(G)$. On the other hand, the sets $A = \{1, 4\}$ and $B = \{2, 5\}$ can only be separated by $H = \{h_1, h_2\}$, that is, we need at least two latent factors for separation. This yields that the minor $\sigma_{12}\sigma_{45} - \sigma_{24}\sigma_{15}$ is *not* in the generating set of $M_{\leq 1}(G)$.

Let $M_{p,m} \subseteq \mathbb{R}[\sigma_{ij},\ i \leq j]$ be the ideal that is generated by all $(m+1) \times (m+1)$-minors of a symmetric matrix $\Sigma \in \mathbb{R}^{p \times p}$.

Theorem 3.4. *Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph with $|\mathcal{H}| = 2$ latent factors. Then*

$$\mathbb{V}(I(G)) = \mathbb{V}((M_{p,2} + M_{\leq 1}(G)) \cap \mathbb{R}[\sigma_{ij},\ i < j]),$$

where the varieties $\mathbb{V}(\cdot)$ are understood over the field $\mathbb{C}$ of complex numbers.

Proof. Let $\mathcal{L} = \{\Lambda\Lambda^\top \in \mathbb{R}^{p \times p} : \Lambda \in \mathbb{R}^D\}$. We first prove that $\mathbb{V}(M_{p,2} + M_{\leq 1}(G))$ is equal to the Zariski closure $\overline{\mathcal{L}}$ in $\mathbb{C}^{p \times p}$. For the inclusion $\overline{\mathcal{L}} \subseteq \mathbb{V}(M_{p,2} + M_{\leq 1}(G))$, consider a matrix $\Sigma \in \mathcal{L}$. Then Σ is symmetric and $\Sigma \in \mathbb{V}(M_{p,2})$. Moreover one can check that $\Sigma \in \mathbb{V}(M_{\leq 1}(G))$ by applying trek separation [STD10]. Thus, $\mathcal{L} \subseteq \mathbb{V}(M_{p,2} + M_{\leq 1}(G))$ and since the variety $\mathbb{V}(M_{p,2} + M_{\leq 1}(G))$ is Zariski closed, we obtain that $\overline{\mathcal{L}} \subseteq \mathbb{V}(M_{p,2} + M_{\leq 1}(G))$.

For the other direction, assume that $\Sigma \in \mathbb{V}(M_{p,2} + M_{\leq 1}(G))$. We explicitly construct a matrix $\Lambda \in \mathbb{C}^D$ such that $\Sigma = \Lambda\Lambda^\top$. We assume that there is no node $v \in V$ such that $\mathrm{pa}(v) = \emptyset$, since this case is trivial where the row of Λ that is indexed by v is zero. Let $\mathcal{H} = \{h_1, h_2\}$ and $V = V_1 \dot\cup V_2 \dot\cup V_3$ be a partition of the observed nodes V into three subsets such that $V_1 = \mathrm{ch}(h_1) \backslash \mathrm{ch}(h_2)$, $V_3 = \mathrm{ch}(h_2) \backslash \mathrm{ch}(h_1)$ and $V_2 = V \setminus (V_1 \cup V_3)$. Without loss of generality we assume that there exists a node $v \in V_1$ such that $\sigma_{vv} \neq 0$. We fix this node v and define the matrices

$$\Lambda_1 = \begin{pmatrix} \mathbf{x} & 0 \\ \mathbf{y} & I_{|V_2|+|V_3|} \end{pmatrix} \in \mathbb{C}^{p \times (1+|V_2|+|V_3|)} \quad \text{and} \quad \Sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \in \mathbb{C}^{(1+|V_2|+|V_3|) \times (1+|V_2|+|V_3|)}.$$

The vector $\mathbf{x} \in \mathbb{C}^{V_1}$ is defined by $\mathbf{x}_w = \sigma_{vw}/\sqrt{\sigma_{vv}}$ for $w \neq v$ and by $\mathbf{x}_v = \sqrt{\sigma_{vv}}$ for the node v. The vector $\mathbf{y} \in \mathbb{C}^{V_2 \cup V_3}$ is defined by $\mathbf{y}_w = \sigma_{vw}/\sqrt{\sigma_{vv}}$ if $w \in V_2$ and by $\mathbf{y}_w = 0$ else. Finally, the symmetric matrix $A = (a_{uw}) \in \mathbb{C}^{V_2 \cup V_3, V_2 \cup V_3}$ is defined as $A = \Sigma_{V_2 \cup V_3, V_2 \cup V_3} - \mathbf{y}\mathbf{y}^\top$.

We prove next that $\Lambda_1\Sigma_1\Lambda_1^\top = \Sigma$. The essential step is to exploit that $\Sigma \in M_{\leq 1}(G)$, which implies the three properties:

$$\Sigma_{V_1, V_3} = 0, \quad \mathrm{rank}(\Sigma_{V_1, V_1 \cup V_2}) \leq 1, \quad \text{and} \quad \mathrm{rank}(\Sigma_{V_2 \cup V_3, V_3}) \leq 1. \tag{7}$$

It holds that $[\Lambda_1\Sigma_1\Lambda_1^\top]_{V_1,V_3} = \Sigma_{V_1,V_3}$, since $[\Lambda_1\Sigma_1\Lambda_1^\top]_{V_1,V_3} = \mathbf{x}\mathbf{y}_{V_3}^\top = 0$. Next, we show that $[\Lambda_1\Sigma_1\Lambda_1^\top]_{V_1,V_1\cup V_2} = \Sigma_{V_1,V_1\cup V_2}$. For any node $w \in V_1 \setminus \{v\}$, we have that $[\Lambda_1\Sigma_1\Lambda_1^\top]_{vw} = \mathbf{x}_w\mathbf{x}_v =$

$(\sigma_{vw}/\sqrt{\sigma_{vv}})\sqrt{\sigma_{vv}} = \sigma_{vw}$. Let $k, w \in V_1 \setminus \{v\}$. By (7), it holds that $\det(\Sigma_{\{v,k\},\{v,w\}}) = 0$, i.e. $\sigma_{vv}\sigma_{kw} = \sigma_{vk}\sigma_{vw}$. Hence,

$$[\Lambda_1\Sigma_1\Lambda_1^\top]_{kw} = \mathbf{x}_k\mathbf{x}_w = \frac{\sigma_{vk}\sigma_{vw}}{\sigma_{vv}} = \sigma_{kw}.$$

If w is an element of V_2 instead of V_1, the conclusion follows similarly by replacing $\mathbf{x}_w$ with $\mathbf{y}_w$. We finally observe that the equality $[\Lambda_1\Sigma_1\Lambda_1^\top]_{V_2\cup V_3,V_2\cup V_3} = \Sigma_{V_2\cup V_3,V_2\cup V_3}$ follows directly from the definitions of Λ_1 and Σ_1.

Now, we return to proving that $\Sigma = \Lambda_1\Sigma_1\Lambda_1^\top \in \overline{\mathcal{L}}$. The matrix Λ_1 has full rank equal to $1+|V_2|+|V_3|$ and the rank of Σ is at most 2. By Sylvester's rank inequality, this implies $\operatorname{rank}(\Sigma_1) \leq 2$. In particular, we have that $\operatorname{rank}(A) \leq 1$. Without loss of generality we may assume that there is a node $u \in V_2 \cup V_3$ such that that $a_{uu} \neq 0$. We fix this node u and define the matrix

$$\Lambda_2 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \mathbf{z} \end{array}\right),$$

where $\mathbf{z} \in \mathbb{C}^{V_2\cup V_3}$ is defined by $\mathbf{z}_w = a_{wu}/\sqrt{a_{uu}}$ for $w \neq u$ and by $\mathbf{z}_u = \sqrt{a_{uu}}$ for the node u. Using similar arguments as above, it is easy to see that $\Sigma_1 = \Lambda_2\Lambda_2^\top$. Finally, define $\Lambda = \Lambda_1\Lambda_2$ and observe that

$$\Lambda = \left(\begin{array}{c|c} \mathbf{x} & 0 \\ \hline \mathbf{y}_{V_2} & \mathbf{z}_{V_2} \\ \hline 0 & \mathbf{z}_{V_3} \end{array}\right) \in \mathbb{C}^D.$$

This shows that $\Sigma \in \overline{\mathcal{L}}$ since $\Lambda\Lambda^\top = \Lambda_1\Lambda_2\Lambda_2^\top\Lambda_1^\top = \Lambda_1\Sigma_1\Lambda_1^\top = \Sigma$.

We now prove the statement of the theorem. Consider the projection π of the space of symmetric $p \times p$ matrices onto the space of the off-diagonal entries. We have that

$$I(G) = I(F(G)) = I(\pi(F(G))) = I(\overline{\pi(F(G))}) = I(\overline{\pi(\mathcal{L})}),$$

where the second equality follows from the fact that membership in $I(F(G))$ only depends on the off-diagonal entries, also see [BD11]. Since the Zariski closure of the projection of an arbitrary set is equal to the Zariski closure of the projection of the Zariski closure of the set, it follows that $\mathbb{V}(I(G)) = \overline{\pi(\mathcal{L})}$. Consequently, we have that $\pi(\overline{\mathcal{L}}) = \pi(\mathbb{V}(M_{p,2} + M_{\leq 1}(G)))$ and by [CLO08, §4.4, Theorem 4], the Zariski closure of the projection $\pi(\mathbb{V}(M_{p,2} + M_{\leq 1}(G)))$ is $\mathbb{V}((M_{p,2} + M_{\leq 1}(G)) \cap \mathbb{R}[\sigma_{ij},\ i < j])$. □

It was shown in [BD11] that the ideal $M_{p,2} \cap \mathbb{R}[\sigma_{ij},\ i < j]$ is generated by two types of generators: off-diagonal 3×3-minors and certain polynomials of degree 5 known as *pentads* [Kel35]. Thus, it is natural to conjecture that the ideal $(M_{p,2} + M_{\leq 1}(G)) \cap \mathbb{R}[\sigma_{ij},\ i < j]$ is generated by off-diagonal 3×3-minors, pentads, and the off-diagonal 1×1 and 2×2-minors in $M_{\leq 1}(G)$; see Conjecture 4.1.

Theorem 3.4 implies that the ideal $(M_{p,2} + M_{\leq 1}(G)) \cap \mathbb{R}[\sigma_{ij},\ i < j]$ is included in the ideal of invariants $I(G)$ we are interested in. In the next section, we combinatorially find a Gröbner basis of $I(G)$ for the special case where the two-factor analysis model has overlap two, that is, there are at most two observed nodes that have two latent parents and all other observed nodes have at most one latent parent. We obtain as Corollary 3.14 that we have indeed $I(G) = (M_{p,2} + M_{\leq 1}(G)) \cap \mathbb{R}[\sigma_{ij},\ i < j]$, and that this ideal is generated by the concerned polynomials that vanish on the model.

Our readers are encouraged to use our code on MathRepo to experiment with Gröbner basis computations of $I(G)$. Gröbner bases for the *full* factor analysis model with one and two latent nodes are given in [DLST95] and [Sul09].

3.2. **Gröbner Basis.** We study the join of sparse one-factor models and provide a technique to construct a Gröbner basis for its ideal of invariants with a "2-delightful" approach from [Sul09]. To give a Gröbner basis for sparse one-factor analysis models, we recall the definition for toric edge ideals. Let G be a simple undirected graph on p vertices and consider the following ring homomorphism:

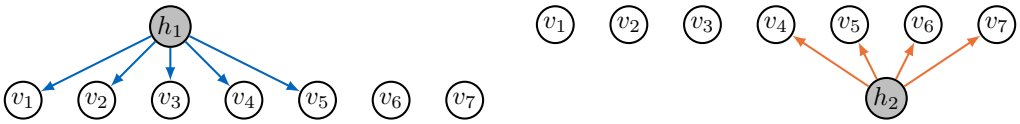


FIGURE 7. Two sparse one-factor analysis graphs where $\{4,5\}$ is the only pair of vertices such that its joint parents are the set of latent nodes $\mathcal{H} = \{h_1, h_2\}$. Identifying these two graphs via the observed nodes $V = \{v_1, \ldots, v_7\}$ concludes the two-factor sparse analysis graph from Figure 2.

$$\Phi_G \colon \mathbb{C}[\sigma_{ij} | 1 \leq i < j \leq p] \longrightarrow \mathbb{C}[t_i | 1 \leq i \leq p]$$
$$\sigma_{ij} \mapsto t_i t_j, \text{ whenever } ij \in E(G).$$

The kernel of Φ_G is called the *toric edge ideal* of G. If $G = K_p$ is the complete graph on p vertices, then $\ker(\Phi_G)$ is the ideal $I_{p,1}$ of the full one-factor analysis model. This immediately follows by the parametrization (2). This ideal is also called the *ideal for the second hypersimplex*.

For a Gröbner basis of $I_{p,1}$, we consider a circular embedding of the complete graph K_p where vertices are presented as the p-th roots of unity in the complex plane. The edges of K_p belong to $\lfloor \frac{p}{2} \rfloor$ the orbits under the action of dihedral group D_p on the roots of unity. The kth class of edges is those that are equivalent to the edge $1k$, for $k \in \{2, \ldots, \lceil \frac{p}{2} \rceil\}$. In other words, the edges that are closer to the boundary of the circular embedding correspond to larger variables in the block ordering.

Definition 3.5. *A circular term order is any block term order such that $\sigma_{i_1 j_1} \succ \sigma_{i_2 j_2}$ whenever the edge $i_1 j_1$ is in a smaller class than the edge $i_2 j_2$.*

The Gröbner basis for the ideal for the second hypersimplex, or equivalently for $I_{p,1}$, with respect to any circular order is studied by De Loera, Sturmfels, and Thomas.

Theorem 3.6 ([DLST95, Theorem 2.1])**.** *The set of square-free quadratic binomials*

$$\{\underline{\sigma_{ij}\sigma_{kl}} - \sigma_{ik}\sigma_{jl}, \underline{\sigma_{il}\sigma_{jk}} - \sigma_{ik}\sigma_{jl} \mid 1 \leq i < j < k < l \leq n\} \tag{8}$$

is a reduced Gröbner basis for the one-factor analysis model $I_{p,1}$ with respect to any circular term order.

These square-free quadratic binomials are known as *tetrads* in the statistics literature. We first adapt this result to sparse one-factor analysis models. Let us consider a sparse one-factor analysis graph where $A \subseteq V$ is the set of children of the latent node and $B = [p] \backslash A$, i.e., the set of isolated vertices. We denote the ideal of invariants of a sparse one-factor analysis model as $I_{A,B,1}$. The ideal $I_{A,B,1}$ is the toric edge ideal of the complete graph $K_{|A|}$ on the vertex set A with the set B of isolated vertices. Thus, one needs to add $|A||B| + \binom{|B|}{2}$ degree-one monomials to the set in Theorem 3.6 to form a reduced Gröbner basis for $I_{A,B,1}$. To simplify the next statement, we relabel the vertices of A as $1, \ldots, |A|$ and the vertices of B as $|A|+1, \ldots, p$.

Proposition 3.7. *The set of degree-one monomials and tetrads*

$$\{\sigma_{ij} \mid i \in B \text{ or } j \in B\} \cup \{\underline{\sigma_{ij}\sigma_{kl}} - \sigma_{ik}\sigma_{jl}, \underline{\sigma_{il}\sigma_{jk}} - \sigma_{ik}\sigma_{jl} \mid 1 \leq i < j < k < l \leq |A|\}.$$

is a reduced Gröbner basis for the sparse one-factor analysis model $I_{A,B,1}$ with respect to any circular term order.

Example 3.8. For the left model in Figure 7 with $A = \{1,2,3,4,5\}$ and $B = \{6,7\}$, the ideal $I_{A,B,1}$ is the toric edge ideal of the graph depicted on the left of Figure 8. There are 11 degree-one monomials

$$\{\sigma_{16}, \sigma_{17}, \sigma_{26}, \sigma_{27}, \sigma_{36}, \sigma_{37}, \sigma_{46}, \sigma_{47}, \sigma_{56}, \sigma_{57}, \sigma_{67}\}.$$

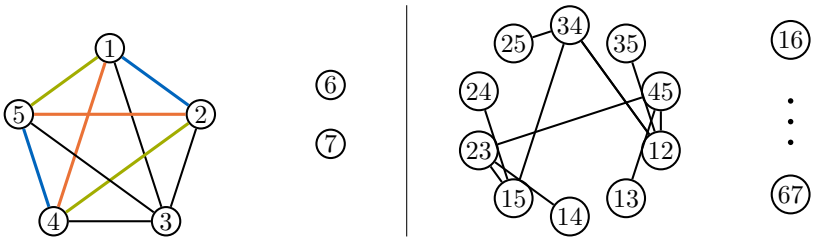


FIGURE 8. The ideal of the sparse one-factor analysis model associated to the left graph in Figure 7 is the toric edge ideal of the graph on the left-hand side consisting of the complete graph $K_{|A|}$ and $|B| = 2$ isolated vertices. The right-hand side pictures the corresponding initial ideal graph $G_{A,B}$ with 11 isolated vertices, where $A = \{1, 2, 3, 4, 5\}$ and $B = \{6, 7\}$.

and $2\binom{|A|}{4} = 10$ tetrads which form a reduced Gröbner basis for $I_{A,B,1}$ with respect to any circular term order. An easy way to construct these tetrads is by looking at the subgraph induced by the vertex set $\{i, j, k, l\}$. We say that a pair of edges ij, kl cross if the line segments in the circular drawing intersect (also at the endpoints) in the circular embedding. The noncrossing pairs (blue and green) of edges correspond to the leading terms of the tetrad generators and the crossing edges (orange) correspond to the remaining monomial of the tetrad. For $\{1, 2, 4, 5\}$, we obtain the tetrads $\underline{\sigma_{12}\sigma_{45}} - \sigma_{14}\sigma_{25}$ and $\underline{\sigma_{24}\sigma_{15}} - \sigma_{14}\sigma_{25}$.

In the case of full-factor analysis, the ideal of invariants is the m-th secant ideal of $I_{p,1}$. However, in the case of sparse-factor analysis models, we need to consider the *join* of the ideals of sparse one-factor analysis models; see, e.g. [SS05, MS05] for a rigorous definition of secants and joins. We consider a sparse k-factor analysis graph as k sparse one-factor analysis graphs that are identified at their observed nodes V. Alternatively, these sparse one-factor analysis graphs can be seen as the induced subgraphs $G[\{h_i\} \cup V]$, where $h_i \in \mathcal{H}$ for $i \in [m]$; see Figure 7.

To construct a Gröbner basis for a sparse two-factor analysis model with respect to any circular term order, we follow an analogous "2-delightful" strategy which was used for the full two-factor analysis models in [Sul09]. For this, we first need to describe the join of the initial ideals of sparse one-factor analysis models. The initial ideal $\mathrm{in}_{\prec}(I_{\{A,B\},1})$ is generated by all noncrossing pairs in the circular embedding of the complete graph $K_{|A|}$ on the vertex set A and the degree-one monomials from Proposition 3.7. Thus it is the monomial edge ideal of a certain graph with isolated vertices.

Definition 3.9. *We define the simple graph called the* initial ideal graph $G_{A,B}$ *whose vertices are labeled as ij where $\{i, j\} \in A \sqcup B = [p]$ and $\{ij, kl\} \in E(G_{A,B})$ whenever (ij, kl) is a noncrossing pair in the circular embedding of the complete graph $K_{|A|}$.*

To avoid confusion, the edges of the initial ideal graph are denoted by $\{\ ,\ \}$, different from the edges of the complete graph $K_{|\,\mathrm{ch}(j_i)|}$ with isolated vertices $\mathrm{ch}(h_i)^{\complement}$. The following definition describes how identifying two sparse one-factor analysis graphs via the observed nodes corresponds to identifying the two associated initial ideal graphs. We focus on the case where the "overlap" is two, that is, $|A_1 \cap A_2| = 2$.

Definition 3.10. *Let G_{A_1,B_1} and G_{A_2,B_2} be two initial ideal graphs with $A_1 \sqcup B_1 = A_2 \sqcup B_2 = V$ and $A_1 \cap A_2 = \{j_1, j_2\}$. We construct a* glued hypergraph identified via V *denoted by $G_{A_1,B_1} \times_V G_{A_2,B_2}$ as follows:*

- *The vertex set is $V(G_{A_1,B_1}) \cup V(G_{A_2,B_2}) \backslash \{$isolated vertices of $V(G_{A_1,\emptyset}) \cup V(G_{A_2,\emptyset})\}$.*
- *The hyperedges of size 2 are all those of G_{A_1,B_1} and G_{A_2,B_2} which do not contain vertex $j_1 j_2$.*
- *The hyperedges of size 3 are those $\{i, j_1 j_2, k\}$, where $i \in V(G_{A_1,\emptyset})$ and and $k \in V(G_{A_2,\emptyset})$.*

Note that the glued hypergraph $G_{A_1,B_1} \times_V G_{A_2,B_2}$ has $|A_1\backslash\{j_1,j_2\}||A_2\backslash\{j_1,j_2\}|$ isolated vertices xy where $x \in A_1\backslash\{j_1,j_2\}$ and $y \in A_2\backslash\{j_1,j_2\}$. In particular, these correspond to the degree-one monomials $M_{\leq 1}(G)$ from Theorem 3.4.

Example 3.11. The initial ideal graph for the sparse one-factor analysis model with $A_1 = \{1,2,3,4,5\}$ and $B_1 = \{6,7\}$ is depicted on the right of Figure 8. Identifying two sparse one-factor analysis graphs from Figure 7 gives rise to the sparse two-factor analysis graph from Figure 2. Here we have $A_2 = \{4,5,6,7\}$ and $B_2 = \{1,2,3\}$. This corresponds to identifying the initial ideal graphs G_{A_1,B_1} and G_{A_2,B_2} via the vertex 45 as in Figure 9. $\{12,45,67\}, \{23,45,67\}, \{13,45,67\}$ are the hyperedges of size 3 and the rest are the hyperedges of size 2 of two initial ideal graphs which do not contain the vertex 45. Since 46 and 57 are crossing edges in the complete graph on 4 vertices, they are not vertices of the glued hypergraph.

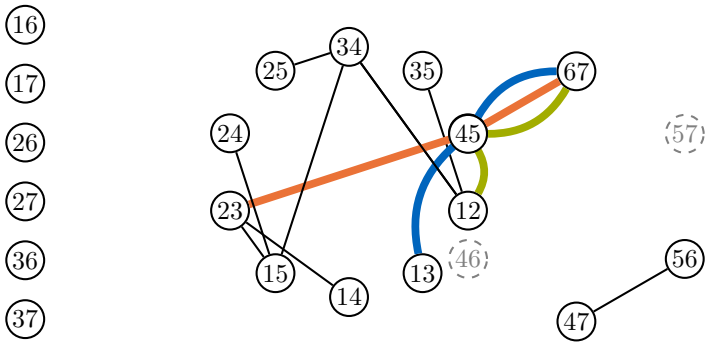


FIGURE 9. Glued hypergraph identified via 7 observable nodes.

Let I_1 and I_2 be two ideals in a polynomial ring $\mathbb{R}[\mathbf{x}] := \mathbb{R}[x_1,\dots,x_n]$. We now recall the definition of join of I_1 and I_2 from [SS05]. Introduce $2n$ new unknowns, grouped in 2 vectors $\mathbf{y}_j = (y_{j1},\dots,y_{jn})$, $j \in \{1,2\}$ and consider the polynomial ring $\mathbb{R}[\mathbf{x},\mathbf{y}]$ in $2n+n$ variables. Moreover, let $I_j(\mathbf{y}_j)$ be the image of the ideal I_j in $\mathbb{R}[\mathbf{x},\mathbf{y}]$ under the map $\mathbf{x} \mapsto \mathbf{y}_j$. Then the *join* $I_1 * I_2$ is the elimination ideal

$$(I_1(\mathbf{y}_1) + I_2(\mathbf{y}_2) + \langle y_{1i} + y_{2i} - x_i \mid 1 \leq i \leq n\rangle) \cap \mathbb{R}[\mathbf{x}].$$

Given a factor analysis graph $G = (V \cup \mathcal{H}, D)$ with $|\mathcal{H}| = 2$ latent nodes, we can identify it with two one-factor analysis graphs. By definition, we have that the ideal of invariants of the two-factor analysis model is equal to the join of the ideals of the one-factor models, that is, $I(G) = I_{A_1,B_1,1} * I_{A_2,B_2,1}$. In this section, we find a Gröbner basis of this join ideal if $|\mathrm{ch}(h_1) \cap \mathrm{ch}(h_2)| = |A_1 \cap A_2| = 2$. We assume that $p \geq 4$, since the ideal of invariants is otherwise empty.

By [SU00, Theorem 2.3], for any term order $\prec$, and any two ideals I_1, I_2, we have that $\mathrm{in}_\prec(I_1 * I_2) \subseteq \mathrm{in}_\prec(I_1) * \mathrm{in}_\prec(I_2)$. Thus, if we find a collection of polynomials $\mathcal{G} \subset I_1 * I_2$ such that $\langle \mathrm{in}_\prec(g) \mid g \in \mathcal{G}\rangle = \mathrm{in}_\prec(I_1) * \mathrm{in}_\prec(I_2)$, then we can deduce that $\mathcal{G}$ is a Gröbner basis with respect to the term order $\prec$ for $I_1 * I_2$. A term order $\prec$ is called *2-delightful* for two ideals I_2 and I_2, when the equality $\mathrm{in}_\prec(I_1 * I_2) = \mathrm{in}_\prec(I_1) * \mathrm{in}_\prec(I_2)$ holds. We next describe the join of the initial ideals of two sparse one-factor analysis models with overlap two with respect to any circular term order.

Lemma 3.12. *Let $I_{A_1,B_1,1}$ and $I_{A_2,B_2,1}$ be the toric ideals of invariants of two sparse one-factor models with $|A_1 \cap A_2| = 2$ and $\prec$ be any circular term order. Then $\mathrm{in}_\prec(I_{A_1,B_1,1}) * \mathrm{in}_\prec(I_{A_2,B_2,1})$ is the monomial edge ideal of the glued hypergraph $G_{A_1,B_2} \times_V G_{A_2,B_2}$.*

Proof. Let $I = \mathrm{in}_\prec(I_{A_1,B_1,1})$ and $J = \mathrm{in}_\prec(I_{A_2,B_2,1})$. We use the Alexander duality formula for the join of monomial ideals from [SS05, Theorem 2.6]. Consider the irreducible component decomposition of $I = \bigcap I_\nu$ and $J = \bigcap J_\mu$. Since I and J are monomial edge ideals, the irreducible components are

the minimal vertex covers of G_{A_1,B_1} and G_{A_2,B_2}, including the isolated vertices ([VT13, Corollary 1.35]). By [SS05, Lemma 2.3], we obtain that the indices of the square-free irreducible monomial ideals $I_\nu * J_\mu$ can be in two forms:

(i) a minimal vertex cover of G_{A_1,B_1}, a minimal vertex cover of G_{A_2,B_2}, $V(G_{V\setminus(A_1\cap A_2),\emptyset})$,
(ii) a minimal vertex cover of G_{A_1,B_1}, a minimal vertex cover of G_{A_2,B_2}, both containing a vertex $ij \in V(G_{A_1\cap A_2,\emptyset})$, $V(G_{V\setminus(A_1\cap A_2),\emptyset})$.

Note that this collection of vertices covers all minimal vertex covers of the glued hypergraph $G_{A_1,B_2}\times_V G_{A_2,B_2}$. By setting up the facets of the associated simplicial complex to be the maximal independent sets of the glued hypergraph, by [HH11, Lemma 1.5.4], we conclude that $\bigcap_{\nu,\mu}(I_\nu * J_\mu)$ is the monomial edge ideal of the glued hypergraph. □

We next construct a Gröbner basis for two factor models with overlap two, that is, $|A_1\cap A_2| = 2$. We find a collection of polynomials $\mathcal{G}$ such that their initial terms with respect to any circular term order are the generators of the monomial edge ideal of the glued hypergraph, and thus this collection forms a Gröbner basis. We discuss the application of the "2-delightful" strategy to cases where $|A_1\cap A_2| \geq 3$ in Section 4.1. We refer to the degree 3 generators below as *hexads* because they are polynomials obtained from six observed random variables, analogous to tetrads, which are degree 2 generators obtained from four observed random variables.

Theorem 3.13. *The generators of a Gröbner basis for $I_{A_1,B_1,1} * I_{A_2,B_2,1}$ with respect to any circular term order for sparse two-factor analysis models where $A_1 \cap A_2 = \{j_1, j_2\}$ comes in three types:*

(1) *Degree-one monomial: σ_{ik} is a generator, where $\mathrm{pa}(i) \cap \mathrm{pa}(j) = \emptyset$.*
(2) *Tetrads: The binomial generators of the Gröbner basis of $I_{A_1,B_1,1}$ and $I_{A_2,B_2,1}$ with respect to any circular order that do not contain $\sigma_{j_1j_2}$.*
(3) *Hexads: Consider $i_1, i_2 \in A_1\setminus\{j_1, j_2\}$ and $k_1, k_2 \in A_2\setminus\{j_1, j_2\}$. Then*

$$\underline{\sigma_{k_1k_2}\sigma_{i_1i_2}\sigma_{j_1j_2}} - \sigma_{k_1k_2}\sigma_{j_1i_2}\sigma_{j_2i_1} - \sigma_{i_1i_2}\sigma_{j_1k_2}\sigma_{j_2k_1},$$

is a degree three generator, where $\{i_1, i_2\}, \{j_1, j_2\}$ and $\{j_1, j_2\}, \{k_1, k_2\}$ are non-crossing edges of the complete graphs on the vertices $A_1\setminus\{j_1, j_2\}$ and $A_2\setminus\{j_1, j_2\}$ respectively.

Proof. This set of polynomials is in the join by the combinatorial definition of join of ideals. And by the previous lemma, their initials are exactly those [SU00, Theorem 2.3]. This concludes the proof. □

The theorem implies that if the set of children of one latent node is strictly contained in the other, then the generators consist of degree-one monomials and tetrads, and thus it is a toric ideal. This means equivalently that the (two) children of exactly one latent node are non-pure.

Corollary 3.14. *Let $|A_1 \cap A_2| = 2$ for a sparse two-factor analysis model graph. Then the ideal $(M_{p,2} + M_{\leq 1}(G)) \cap \mathbb{R}[\sigma_{ij},\ i < j]$ is equal to the join $I(G) = I_{A_1,B_1,1} * I_{A_2,B_2,1}$, and thus prime. In particular, $I(G)$ is generated by all off-diagonal minors of size at most 3 in $M_{p,2}$ and off-diagonal minors of size at most 2 in $M_{\leq 1}(G)$.*

Proof. Let $J = (M_{p,2} + M_{\leq 1}(G)) \cap \mathbb{R}[\sigma_{ij},\ i < j]$. By Theorem 3.4, $J \subseteq I_{A_1,B_1,1} * I_{A_2,B_2,1}$. Thus it is enough to show $I := I_{A_1,B_1,1} * I_{A_2,B_2,1} \subseteq J$. The degree-one and degree-two generators of $I_{A_1,B_1,1} * I_{A_2,B_2,1}$ are in $M_{\leq 1}(G)$. The degree three generators can be described as a 3×3 off-diagonal minor. Since I_1 and I_2 are both prime ideals, I is also prime [SU00, Proposition 1.2]. □

Example 3.15. Consider the sparse two-factor analysis graph G with $A_1\cap A_2 = \{4, 5\}$ from Figure 2. The degree-one monomials of the generators of the Gröbner basis constructed in Theorem 3.13 are $\sigma_{16}, \sigma_{17}, \sigma_{26}, \sigma_{27}, \sigma_{36}, \sigma_{37}$. These are the same as the set $M_{\leq 1}(G)$. The tetrads are all the ones in form (8) that do not contain the σ_{45}. For example, $\underline{\sigma_{12}\sigma_{45}} - \sigma_{14}\sigma_{25}$ from Example 3.8 is not a tetrad

generator of the Gröbner basis for the ideal of the sparse two-factor analysis. Finally, we obtain the following three hexads to form the Gröbner basis:

$$\sigma_{67}\sigma_{12}\sigma_{45} - \sigma_{67}\sigma_{24}\sigma_{15} - \sigma_{12}\sigma_{47}\sigma_{56},$$
$$\sigma_{67}\sigma_{13}\sigma_{45} - \sigma_{67}\sigma_{34}\sigma_{15} - \sigma_{13}\sigma_{47}\sigma_{56},$$
$$\sigma_{67}\sigma_{23}\sigma_{45} - \sigma_{67}\sigma_{34}\sigma_{25} - \sigma_{23}\sigma_{47}\sigma_{56}.$$

4. Conclusion and Open Questions

In this paper, we derived novel results on the algebro-geometric aspects of sparse factor analysis models. We first proved upper and lower bounds for the dimension. While the upper bound holds for arbitrary models, the lower bound holds for models that satisfy a minimal level of sparsity, which we formalize in the ZUTA condition. In many cases, upper and lower bounds coincide and one obtains a formula for the dimension. In particular, our study reveals that sparse factor analysis models, unlike full factor analysis models, may not have expected dimension. Then, we studied the ideal of invariants of sparse factor models with two latent nodes. We presented an ideal that cuts out the model and, moreover, we derived a Gröbner basis for models with at most overlap two, i.e., models where at most two observed nodes have more than one latent parent. On a technical level, we extended the *delightful* strategy, which was previously applied to secants, to joins of ideals.

In what follows, we outline some possible future directions and open questions that arose from our paper. All the examples below can be reproduced by our code on MathRepo.

intersection size	degree	# indeterminates	# monomials
3	3	8	4
(p=7)	5	9	6
4	5	11	6
(p=8)	5	11	8
	5	12	10
	5	12	12

Table 1. Polynomials in Gröbner bases for sparse two-factor models with children sets of overlap 3 or 4. Each line reports the degrees, number of indeterminates, and number of monomials of one type of polynomial in the Gröbner basis.

4.1. **Sparse factor models with larger overlaps.** The circular term order is not always 2-delightful for the ideal of invariants of sparse two-factor analysis models for examples where $|\operatorname{ch}(h_1) \cap \operatorname{ch}(h_2)| = |A_1 \cap A_2| \geq 3$. Consider the sparse two-factor analysis graph from Figure 2 with the additional edge $h_2 \to v_3$. The generators of a Gröbner basis with respect to any circular term order of the join has degrees one, two, three (degree-one monomials, tetrads, and non-hexads) and five, whereas the join of initial ideals is generated by at most degree three generators. In particular, one of the generators has the form

$$\sigma_{45}\sigma_{67}\sigma_{57}\sigma_{14}\sigma_{36} - \sigma_{45}\sigma_{67}\sigma_{15}\sigma_{36}\sigma_{47} + \sigma_{56}\sigma_{67}\sigma_{35}\sigma_{14}\sigma_{47}$$
$$-\sigma_{56}\sigma_{57}\sigma_{14}\sigma_{36}\sigma_{47} - \sigma_{67}\sigma_{35}\sigma_{46}\sigma_{57}\sigma_{14} + \sigma_{46}\sigma_{57}\sigma_{15}\sigma_{36}\sigma_{47}.$$

This is a polynomial of degree 5 in 9 indeterminates involving 6 monomial terms. The monomial terms coincide with monomial terms in a pentad [Kel35, DSS07] although the pentad has twelve terms. The missing monomials of the pentad are reduced by the elements in $M_{\leq 1}(G)$. We list the types of homogeneous polynomials when we compute Gröbner bases for larger intersections among the children sets of the latent variables in Table 1. Our computations support the following conjecture for factor analysis models with two latent variables.

Conjecture 4.1. *The ideal of the sparse two-factor analysis model corresponding to graph G is generated by off-diagonal 3×3-minors, pentads, and the polynomials in $M_{\leq 1}(G)$.*

Although the "delightful strategy" with circular term orders is helpful for Gröbner basis, we observed that it fails for sparse two-factor analysis models with $|\operatorname{ch}(h_1) \cap \operatorname{ch}(h_2)| \geq 3$. However, since by [SS05, Prop. 2.4], the join of monomial ideals is monomial, one may consider constructing the join $\operatorname{in}_\prec(I_{A_1,B_1,1}) * \operatorname{in}_\prec(I_{A_2,B_2,1})$ as the monomial edge ideal of another hypergraph with respect to another term order.

Question 1. *Is there a 2-delightful term order for sparse two-factor analysis models? In other words, is there a term order $\prec$ such that $\operatorname{in}_\prec(I_{A_1,B_1,1} * I_{A_2,B_2,1}) = \operatorname{in}_\prec(I_{A_1,B_1,1}) * \operatorname{in}_\prec(I_{A_2,B_2,1})$?*

4.2. **Sparse factor models with more than two latent factors.** The generators introduced in Theorem 3.13 can be partially used for sparse factor analysis models with more than two latent nodes where we have non-empty intersections only for consecutive intersections $|\operatorname{ch}(h_i) \cap \operatorname{ch}(h_{i+1})| = |A_i \cap A_{i+1}| = 2$ with $i \in [m-1]$. Consider for instance the graph G in Figure 10. Since the induced subgraph $G[\{h_3\} \cap V] \subset G$ gives rise to the toric edge ideal of K_3 and isolated vertices [5], we obtain the generators of a Gröbner basis with respect to any circular term order in three types as follows:

(1) 11 degree-one monomials σ_{ij}, where $\operatorname{pa}(i) \cap \operatorname{pa}(j) = \emptyset$:

$$\sigma_{15}, \sigma_{16}, \sigma_{17}, \sigma_{18}, \sigma_{25}, \sigma_{26}, \sigma_{27}, \sigma_{28}, \sigma_{38}, \sigma_{48}, \sigma_{58},$$

(2) 6 tetrads that do not contain σ_{34} or σ_{67}:

$$\sigma_{47}\sigma_{56} - \sigma_{46}\sigma_{57}, \quad \sigma_{37}\sigma_{56} - \sigma_{36}\sigma_{57}, \quad \sigma_{37}\sigma_{46} - \sigma_{36}\sigma_{47},$$
$$\sigma_{37}\sigma_{45} - \sigma_{35}\sigma_{47}, \quad \sigma_{36}\sigma_{45} - \sigma_{35}\sigma_{46}, \quad \sigma_{14}\sigma_{23} - \sigma_{13}\sigma_{24},$$

(3) and 2 hexads:

$$\sigma_{12}\sigma_{34}\sigma_{57} - \sigma_{12}\sigma_{35}\sigma_{47} - \sigma_{13}\sigma_{24}\sigma_{57}, \quad \sigma_{12}\sigma_{34}\sigma_{56} - \sigma_{12}\sigma_{35}\sigma_{46} - \sigma_{13}\sigma_{24}\sigma_{56}.$$

If we add one more observable node v_9 and the edge $h_3 \to v_9$, we obtain degree-one monomials, tetrads, hexads, and a degree four generator in ten indeterminates which seems to have a combinatorial structure as in the hexad case, e.g.,

$$\sigma_{12}\sigma_{34}\sigma_{67}\sigma_{89} - \sigma_{12}\sigma_{34}\sigma_{68}\sigma_{79} - \sigma_{12}\sigma_{89}\sigma_{36}\sigma_{47} - \sigma_{67}\sigma_{89}\sigma_{13}\sigma_{24} + \sigma_{13}\sigma_{24}\sigma_{68}\sigma_{79}.$$

Moreover, we obtain that the circular term order is 3-delightful, i.e.,

$$\operatorname{in}_\prec(I_{A_1,B_1,1} * I_{A_2,B_2,1} * I_{A_3,B_3,1}) = \operatorname{in}_\prec(I_{A_1,B_1,1}) * \operatorname{in}_\prec(I_{A_2,B_2,1}) * \operatorname{in}_\prec(I_{A_3,B_3,1}).$$

Adding a fourth latent variable h_4 while keeping the cardinality of intersections 2 will also give rise to polynomials of degree 5 in 13 indeterminates with 8 summands, like

$$\sigma_{12}\sigma_{34}\sigma_{67}\sigma_{910}\sigma_{1112} - \sigma_{12}\sigma_{34}\sigma_{67}\sigma_{911}\sigma_{1012} - \sigma_{12}\sigma_{34}\sigma_{1112}\sigma_{69}\sigma_{710} - \sigma_{12}\sigma_{910}\sigma_{1112}\sigma_{36}\sigma_{47} +$$
$$\sigma_{12}\sigma_{911}\sigma_{1012}\sigma_{36}\sigma_{47} - \sigma_{67}\sigma_{910}\sigma_{1112}\sigma_{13}s_{24} + \sigma_{67}\sigma_{13}\sigma_{24}\sigma_{911}\sigma_{1012} + \sigma_{1112}\sigma_{13}\sigma_{24}\sigma_{69}\sigma_{710}.$$

We summarize the findings for the maximal degree polynomials we have computed based on the number of latent nodes in Table 2.

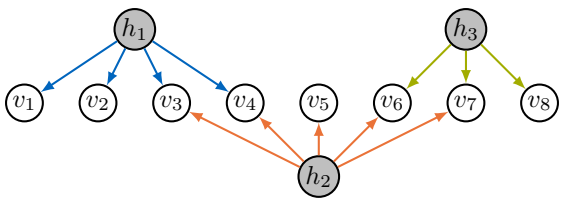


FIGURE 10. A factor analysis graph with 3 latent factors and two overlaps.

# latent nodes	degree	# indeterminates	# monomials
0	1	1	1
1	2	4	2
2	3	7	3
3	4	10	5
4	5	13	8

TABLE 2. Degrees, number of variables and number of terms for different latent nodes when there are at most 2 intersections.

Question 2. *Can we show that this behavior generalizes for more hidden variables with intersection two? That is, in the presence of k latent nodes, is the polynomial of maximal degree a degree $k+1$ polynomial in $3k+1$ variables which has $(k+2)$ Fibonacci number of terms? In fact, using the delightful strategy, can we find a Gröbner basis with respect to any circular term order for sparse k-factor analysis models with more than two latent nodes where $|\operatorname{ch}(h_i) \cap \operatorname{ch}(h_{i+1})| = 2$ for $i \in [m-1]$?*

Describing a Gröbner basis for this case would require a generalization of the glued hypergraph (Definition 3.9) and of the construction of the polynomials as in Theorem 3.13.

ACKNOWLEDGEMENTS

This project was started at a workshop funded by the TUM Global Incentive Fund "Algebraic Methods in Data Science". It received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 883818) and from the German Federal Ministry of Education and Research and the Bavarian State Ministry for Science and the Arts. Nils Sturma acknowledges support from Munich Data Science Institute (MDSI) at the Technical University of Munich through the Linde/MDSI PhD Fellowship program. The authors are grateful for helpful discussions with Danai Deligeorgaki, Alex Markham, Pratik Misra, Lisa Seccia, and Liam Solus.

REFERENCES

[ABHP16] Christian Aßmann, Jens Boysen-Hogrefe, and Markus Pape. Bayesian analysis of static and dynamic factor models: An ex-post approach towards the rotation problem. *Journal of Econometrics*, 192(1):190–206, 2016.

[ADG+23] Carlos Améndola, Mathias Drton, Alexandros Grosdos, Roser Homs, and Elina Robeva. Third-order moment varieties of linear non-Gaussian graphical models. *Information and Inference: A Journal of the IMA*, 12(3):1405–1436, 04 2023.

[AGM12] S. Arora, R. Ge, and A. Moitra. Learning topic models – Going beyond SVD. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 1–10, Los Alamitos, CA, USA, 2012. IEEE Computer Society.

[AR56] T. W. Anderson and Herman Rubin. Statistical inference in factor analysis. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. V*, pages 111–150. Univ. California Press, Berkeley-Los Angeles, Calif., 1956.

[BBNW20] Xin Bing, Florentina Bunea, Yang Ning, and Marten Wegkamp. Adaptive estimation in structured factor models with applications to overlapping clustering. *Ann. Statist.*, 48(4):2055–2081, 2020.

[BD11] Andries E. Brouwer and Jan Draisma. Equivariant Gröbner bases and the Gaussian two-factor model. *Math. Comp.*, 80(274):1123–1133, 2011.

[BDSW22] Rina Foygel Barber, Mathias Drton, Nils Sturma, and Luca Weihs. Half-trek criterion for identifiability of latent variable models. *Ann. Statist.*, 50(6):3174–3196, 2022.

[BLR+13] Amy S. Beavers, John W. Lounsbury, Jennifer K. Richards, Schuyler W. Huck, Gary J. Skolits, and Shelley L. Esquivel. Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, 18(6), 2013.

[Bol89] Kenneth A. Bollen. *Structural equations with latent variables*. John Wiley & Sons, Inc., New York, 1989.

[BT00] Kenneth A Bollen and Kwok-Fai Ting. A tetrad test for causal indicators. *Psychol. Methods*, 5(1):3–22, 2000.

[CBBP93] Gian Vittorio Caprara, Claudio Barbaranelli, Laura Borgogni, and Marco Perugini. The "big five questionnaire": A new questionnaire to assess the five factor model. *Personality and Individual Differences*, 15(3):281–288, 1993.

[CLO08] D.A. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer New York, 2008.

[DLST95] Jesús A. De Loera, Bernd Sturmfels, and Rekha R. Thomas. Gröbner bases and triangulations of the second hypersimplex. *Combinatorica*, 15:409–424, 1995.

[dOSGdC+19] Roberta de Oliveira Santos, Bartira Mendes Gorgulho, Michelle Alessandra de Castro, Regina Mara Fisberg, Dirce Maria Marchioni, and Valéria Troncoso Baltar. Principal component analysis and factor analysis: Differences and similarities in nutritional epidemiology application. *Revista Brasileira de Epidemiologia*, 22, 2019.

[DSS07] Mathias Drton, Bernd Sturmfels, and Seth Sullivant. Algebraic factor analysis: tetrads, pentads and beyond. *Probab. Theory Relat. Fields*, 138(3-4):463–493, 2007.

[DX16] Mathias Drton and Han Xiao. Wald tests of singular hypotheses. *Bernoulli*, 22(1):38–59, 2016.

[FFL08] Jianqing Fan, Yingying Fan, and Jinchi Lv. High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147(1):186–197, 2008.

[FSL18] Sylvia Fruehwirth-Schnatter and Hedibert Freitas Lopes. Sparse Bayesian factor analysis when the number of factors is unknown. *arXiv preprint arXiv:1804.04231*, 2018.

[Har76] Harry H. Harman. *Modern factor analysis*. University of Chicago Press, Chicago, Ill.-London, revised edition, 1976.

[HH11] Jürgen Herzog and Takayuki Hibi. *Monomial ideals*. Springer, 2011.

[Hor65] John L. Horn. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185, June 1965.

[Kel35] T.L. Kelley. *Essential Traits of Mental Life*, volume 26 of *Harvard Studies in Education*. Harvard University Press, 1935.

[KZ22] Dale S. Kim and Qing Zhou. A correlation thresholding algorithm for learning factor analysis models. *arXiv preprint arXiv:2203.01471*, 2022.

[LWSB14] Andrew S. Lan, Andrew E. Waters, Christoph Studer, and Richard G. Baraniuk. Sparse factor analysis for learning and content analytics. *J. Mach. Learn. Res.*, 15:1959–2008, 2014.

[MDLW19] Marloes Maathuis, Mathias Drton, Steffen Lauritzen, and Martin Wainwright, editors. *Handbook of graphical models*. CRC Press, Boca Raton, FL, 2019.

[MLAS23] Alex Markham, Mingyu Liu, Bryon Aragam, and Liam Solus. Neuro-causal factor analysis. *arXiv preprint arXiv:2305.19802*, 2023.

[MMS98] M. E. Martínez, J. R. Marshall, and L. Sechrest. Invited commentary: Factor analysis and the search for objectivity. *American Journal of Epidemiology*, 148(1):17–19, July 1998.

[MS05] Ezra Miller and Bernd Sturmfels. *Combinatorial commutative algebra*, volume 227 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.

[MSWB22] Gemma Elyse Moran, Dhanya Sridhar, Yixin Wang, and David Blei. Identifiable deep generative models via sparse decoding. *Transactions on Machine Learning Research*, 2022.

[OLK23] Ilsang Ohn, Lizhen Lin, and Yongdai Kim. A Bayesian sparse factor model with adaptive posterior concentration. *arXiv preprint arXiv:2305.18488*, 2023.

[Pea00] Judea Pearl. *Causality*. Cambridge University Press, Cambridge, 2000. Models, reasoning, and inference.

[PJS17] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 2017. Foundations and learning algorithms.

[RWC00] Steven P. Reise, Niels G. Waller, and Andrew L. Comrey. Factor analysis and scale revision. *Psychological Assessment*, 12(3):287–297, September 2000.

[SDL22] Nils Sturma, Mathias Drton, and Dennis Leung. Testing many and possibly singular polynomial constraints. *arXic preprint arXiv:2208.11756*, 2022.

[SNS+06] James B. Schreiber, Amaury Nora, Frances K. Stage, Elizabeth A. Barlow, and Jamie King. Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99(6):323–338, 2006.

[SS05] Bernd Sturmfels and Seth Sullivant. Combinatorial secant varieties. *Pure Appl. Math. Q.*, 2, 07 2005.

[SSGS06] Ricardo Silva, Richard Scheines, Clark Glymour, and Peter Spirtes. Learning the structure of linear latent variable models. *J. Mach. Learn. Res.*, 7:191–246, 2006.

[STD10] Seth Sullivant, Kelli Talaska, and Jan Draisma. Trek separation for Gaussian graphical models. *Ann. Statist.*, 38(3):1665–1685, 2010.

[SU00] Aron Simis and Bernd Ulrich. On the ideal of an embedded join. *Journal of Algebra*, 226(1):1–14, 2000.

[Sul08] Seth Sullivant. Algebraic geometry of Gaussian Bayesian networks. *Adv. in Appl. Math.*, 40(4):482–513, 2008.

[Sul09] Seth Sullivant. A Gröbner basis for the secant ideal of the second hypersimplex. *Journal of Commutative Algebra*, 1(2):327 – 338, 2009.

[TFA17] Nickolay T. Trendafilov, Sara Fontanella, and Kohei Adachi. Sparse exploratory factor analysis. *Psychometrika*, 82(3):778–794, 2017.

[VT13] Adam Van Tuyl. A beginner's guide to edge and cover ideals. *Monomial ideals, computations and applications*, pages 63–94, 2013.

Munich Center for Machine Learning (MCML) and Department of Mathematics, School of Computation, Information and Technology, Technical University of Munich
Email address: mathias.drton@tum.de

Institute of Mathematics, University of Augsburg
Email address: alexandros.grosdos@uni-a.de

Max Planck Institute for Mathematics in the Sciences, Leipzig
Email address: mail@irem-portakal.de

Munich Center for Machine Learning (MCML) and Department of Mathematics, School of Computation, Information and Technology, Technical University of Munich
Email address: nils.sturma@tum.de

Permission to include the article

arXiv.org - Non-exclusive license to distribute

The URI http://arxiv.org/licenses/nonexclusive-distrib/1.0/ is used to record the fact that the submitter granted the following license to arXiv.org on submission of an article:

- I grant arXiv.org a perpetual, non-exclusive license to distribute this article.
- I certify that I have the right to grant this license.
- I understand that submissions cannot be completely removed once accepted.
- I understand that arXiv.org reserves the right to reclassify or reject any submission.

Revision history

2004-01-16 - License above introduced as part of arXiv submission process
2007-06-21 - This HTML page created

Contact

We gratefully acknowledge support from the Simons Foundation,
member institutions, and all contributors.

Reuse Requests

This FAQ is an attempt to collect answers to your common questions surrounding reusing content from arXiv in your materials.

Can I reuse figures from an arXiv paper?

The short answer is "it depends". More specifically: - If the license applied to the work allows for remixing or reuse with citation, then yes. - If not, then the version is assigned one of the arXiv perpetual non-exclusive licenses, and you will need to contact the submitter or copyright holder (if published) to determine applicable permissions.

Do I need arXiv's permission to repost the full text?

Note: All e-prints submitted to arXiv are subject to copyright protections. arXiv is not the copyright holder on any of the e-prints in our corpus.

In some cases, submitters have provided permission in advance by submitting their e-print under a permissive Creative Commons license. The overwhelming majority of e-prints are submitted using the arXiv perpetual non-exclusive license, which does not grant further reuse permissions directly. In these cases you will need to contact the author directly with your request.

How can I determine what license the version was assigned?

All arXiv abstract pages indicate an assigned license underneath the "Download:" options.

The link may appear as just the text `(license)`, such as at arXiv:2201.14176. Articles between 1991 and 2003 have an assumed license. These are functionally equivalent to the arXiv non-exclusive license.

If the license applied by the submitter is one of the Creative Commons licenses, then a "CC" logo will appear, such as at arXiv:2201.04182.

I want to include a paper of mine from arXiv in my thesis, do I need specific permission?

If you are the copyright holder of the work, you do not need arXiv's permission to reuse the full text.

I want to include a paper of mine from arXiv in an institutional repository, do I need permission?

You do not need arXiv's permission to deposit arXiv's version of *your* work into an institutional repository. For all other institutional repository cases, see our help page on institutional repositories.

Can I harvest the full text of works?

Plase see our bulk data help page, and the API Terms of Use for specific options. Note that the license for the full text is not a part of the current search API schema. The license is, however, provided within arXiv's output from the OAI-PMH in either `arXiv` or `arXivRaw` formats.