Homomorphism Indistinguishability

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Tim Frederik Seppelt, Master of Advanced Study
aus Braunschweig

Berichter: Universitätsprofessor Dr. Martin Grohe
Universitätsprofessor Dr. Holger Dell

Tag der mündlichen Prüfung: 29. November 2024

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

Abstract

Two graphs G and H are *homomorphism indistinguishable* over a class of graphs $\mathcal{F}$ if, for all graphs $F \in \mathcal{F}$, the number of homomorphisms from F to G is equal to the number of homomorphisms from F to H. In 1967, Lovász showed that two graphs are isomorphic if, and only if, they are homomorphism indistinguishable over the class of all graphs. Subsequently, many graph isomorphism relaxations such as quantum isomorphism, spectral, and logical equivalences have been characterised as homomorphism indistinguishability relations over certain graph classes. Thereby, homomorphism indistinguishability connects seemingly disparate fields such as quantum information, finite model theory, and machine learning.

This thesis explores three themes: We first review the plenitude of **characterisations** of graph isomorphism relaxations as a homomorphism indistinguishability relation. Focusing on integer programming relaxations for graph isomorphism, we prove that the feasibility of each level of the Sherali–Adams and Lasserre hierarchies is characterised as homomorphism indistinguishability relations. These results, which are derived using (bi)labelled graphs and homomorphism tensors, shed light on the distinguishing power of these hierarchies. In particular, we determine the precise number of Sherali–Adams levels necessary such that their feasibility guarantees the feasibility of a given Lasserre level.

Abstracting from the wealth of homomorphism indistinguishability characterisations, we embark on a more principled study of homomorphism indistinguishability investigating the distinguishing power and the complexity of homomorphism indistinguishability relations over minor-closed graph class.

The homomorphism distinguishing **closure** $\text{cl}(\mathcal{F})$ of a graph class $\mathcal{F}$ is the central notion for studying the distinguishing power of homomorphism indistinguishability relations. It is defined as the maximal graph class whose homomorphism indistinguishability relation coincides with the one of $\mathcal{F}$. Roberson conjectured that every minor-closed union-closed graph class $\mathcal{F}$ is homomorphism distinguishing closed, i.e. $\text{cl}(\mathcal{F}) = \mathcal{F}$. We confirm Roberson's conjecture, which is generally wide open, for further graphs classes and prove unconditionally that if $\mathcal{F}$ is minor-closed then so is $\text{cl}(\mathcal{F})$.

Lastly, we investigate the **complexity** of deciding whether two graphs are homomorphism indistinguishable over a fixed graph class. For infinite graph classes, this problem is a priori not even decidable. In stark contrast to this, we show that, over every minor-closed graph class of bounded treewidth, homomorphism indistinguishability can be decided in randomised polynomial time.

Zusammenfassung

Zwei Graphen G und H heißen *Homomorphismen-ununterscheidbar* über einer Graphenklasse $\mathcal{F}$, wenn für all Graphen $F \in \mathcal{F}$ gilt, dass die Zahl der Homomorphismen von F nach G gleich der Zahl der Homomorphismen von F nach H ist.

Lovász zeigte 1967, dass zwei Graphen genau dann isomorph sind, wenn sie Homomorphismen-ununterscheidbar über der Klasse aller Graphen sind. Im Anschluss daran wurden zahlreiche Graphenisomorphie-Relaxationen wie Quantenisomorphie, sowie spektrale und logische Äquivalenzen als Homomorphismen-Ununterscheidbarkeits-Relationen charakterisiert. Dadurch verbindet Homomorphismen-Ununterscheidbarkeit scheinbar disparate Forschungsfelder wie Quanteninformationstheorie, Endliche Modelltheorie und Maschinelles Lernen.

Diese Dissertation untersucht drei Themen: Zu Beginn geben wir einen Überblick über die Vielzahl von **Charakterisierungen** von Graphenisomorphie-Relaxationen als Homomorphismen-Ununterscheidbarkeits-Relationen, insbesondere von Relaxationen von ganzzahligen Programmen für Graphenisomorphie. Wir zeigen mittels bimarkierter Graphen und Homomorphismentensoren, dass die Lasserre-Hierarchie solche Charakterisierungen besitzt. Als Korollar bestimmen wir die Unterscheidungskraft dieser Hierarchie im Vergleich zur Sherali-Adams-Hierarchie.

Mit dem Ziel, von der Fülle obiger Charakterisierungen zu abstrahieren, beginnen wir eine prinzipiellere Untersuchung von Homomorphismen-Ununterscheidbarkeit, insbesondere von Unterscheidungsstärke und Komplexität von Homomorphismen-Ununterscheidbarkeits-Relationen über Minoren-abgeschlossenen Graphenklassen.

Der Homomorphismen-Unterscheidungs-**Abschluss** $\text{cl}(\mathcal{F})$ einer Graphenklasse $\mathcal{F}$ ist der zentrale Begriff mit Hinblick auf die Unterscheidungsstärke von Homomorphismen-Ununterscheidbarkeits-Relationen. Er ist definiert als die maximale Graphenklasse, deren Homomorphismen-Ununterscheidbarkeits-Relation mit der von $\mathcal{F}$ zusammenfällt. Roberson vermutete, dass jede Graphenklasse, die unter Minorenbildung und disjunkter Vereinigung abgeschlossen ist, *Homomorphismen-Unterscheidungs-abgeschlossen* ist, d.h. es gilt, dass $\text{cl}(\mathcal{F}) = \mathcal{F}$. Wir bestätigen die Robersonsche Vermutung, die allgemein offen ist, für weitere Graphenklassen. Außerdem zeigen wir, ohne Annahme unbewiesener Aussagen, dass $\text{cl}(\mathcal{F})$ Minoren-abgeschlossen ist, wenn selbiges für $\mathcal{F}$ gilt.

Abschließend betrachten wir die **Komplexität** von Homomorphismen-Ununterscheidbarkeit. Wir zeigen, dass zwei Graphen in randomisierter Polynomialzeit auf Homomorphismen-Ununterscheidbarkeit über jeder Minoren-abgeschlossenen Graphenklasse beschränkter Baumweite geprüft werden können.

Acknowledgements

I would like to thank the many people who have guided me, supported me, and believed in me:

- Martin, for your supervision and guidance. During my time at your chair, I enjoyed every imaginable freedom and profited extensively from the advice you offered me on ideas, writing, and career plans.
- my second supervisors Gerhard and Michael, for helping me gain perspective and not lose sight of the bigger picture.
- Holger, for reviewing this thesis.
- my co-authors Benedikt, Christoph, Daniel, David, Eva, Gaurav, Gian Luca, Jan, Louis, Martin, Moritz, and Nina for our fruitful collaborations. It is the exchange with you and the discussions we shared that made research both productive and very enjoyable.
- David, Radu, Mikołaj, and Anuj for inviting me to research stays in Copenhagen, Warsaw, and Cambridge in 2022–2024. The discussions we had allowed me to broaden my understanding of computer science, more importantly, to connect with others. Thank you for providing a welcoming and warm environment.
- the DFG Research Training Group 2236/2 UnRAVeL and the European Union (ERC, SymSim, 101054974), for their generous support. I am very grateful for the many valuable UnRAVeL events I got to attend and for being part of the UnRAVeL community. Thank you, Birgit and Helen, for putting your heart into UnRAVeL.
- Gaurav and Christoph, for helping me to grow into my life as a PhD student in autumn 2020 when I did not know anyone at our chair. Louis, for being a great office mate and for always having advice, perspective, and exotic complexity classes to offer.
- Martin, Eva, Moritz, Jan, Vincent, Komal, Moritz, Christoph, Matthias, Louis, Olivia, and Dora for proofreading parts of this thesis and for having a vigilant eye for typos, punctuation, inaccuracies, and unintended puzzles.

Lastly, I would like to thank my friends, my family, and foremost Dora. Without your unconditional support, this thesis would have never materialised.

Contents

1 Introduction

Graphs are versatile data structures, which are used in theory and applications to model e.g. chemical compounds, social interactions, trophic networks, and program executions. When maintaining graph data, it is essential to decide whether two graphs are isomorphic, i.e. structurally equivalent [88]. The limited complexity-theoretic understanding of graph isomorphism [83] has led to the study of computationally tractable *graph isomorphism relaxations*, i.e. isomorphism-invariant equivalence relations between graphs. Graph isomorphism relaxations are also practically relevant [80, 75]: In applications, it may be inconsequential whether two graphs are isomorphic but whether they encode e.g. molecules which share desired properties.

In 1967, Lovász [114] proved that two graphs G and H are isomorphic if, and only if, for every graph F, the number of homomorphisms from F to G equals the number of homomorphisms from F to H. A *homomorphism* from F to G is a map $h\colon V(F) \to V(G)$ such that if $uv \in E(F)$, then $h(u)h(v) \in E(G)$. Lovász [114] inspired the notion which is now known as homomorphism indistinguishability: Two graphs G and H are *homomorphism indistinguishable* over a graph class $\mathcal{F}$ if, for every $F \in \mathcal{F}$, the number of homomorphisms from F to G is equal to the number of homomorphisms from F to H.

Not only graph isomorphism but also many of its relaxations can be characterised as homomorphism indistinguishability relations. A substantial body of research characterised graph isomorphism relaxations from diverse fields such as quantum information [124, 15], finite model theory [63, 79, 68], optimisation [61, 86, 151], algebraic graph theory [61, 86], machine learning [130, 177, 81, 180], and category theory [57, 3, 129] as homomorphism indistinguishability relations.

This thesis explores three themes: **Characterisations** are results equating a homomorphism indistinguishability relation with a known graph isomorphism relaxation from e.g. logic or algebraic graph theory. The wealth of such results motivates a more principled study of homomorphism indistinguishability [14, 150]. Here, two fundamental properties of homomorphism indistinguishability relations are investigated: their distinguish power and their complexity. The homomorphism distinguishing **closure** introduced in [150] is the central tool for understanding the distinguishing power of homomorphism indistinguishability relations. **Complexity** is concerned with computationally determining whether two input graphs are homomorphism indistinguishable over a fixed graph class.

Graph Class	Homomorphism Indistinguishability Relation	Reference
All graphs	isomorphism	Theorem 3.1.1
Bipartite graphs	isomorphic bipartite double covers	Corollary 3.1.3
Planar graphs	quantum isomorphism	[124]
Cycles	cospectral adjacency matrices	Theorem 3.3.2
Cycles and paths	cospectral adjacency matrices and cospectral complement adjacency matrices	Corollary 3.3.3
Paths	fractional isomorphism without non-negativity constraints	Corollary 4.2.3
Stars	equal degree sequence	Theorem 3.3.1
Outerplanar graphs	level-1 Lasserre relaxation	Corollary 5.2.22
Treewidth $< k$	C^k-equivalence	Theorem 3.4.4
Treedepth $\leq q$	C_q-equivalence	Theorem 3.4.5
Pathwidth $< k$	level-k Sherali–Adams relaxation without non-negativity constraints	Theorem 4.0.1

Table 1.1: Graph classes and their homomorphism indistinguishability relations.

1.1 Characterisations

The two results which popularised homomorphism indistinguishability subsequent to Lovász [114] provide characterisations of graph isomorphism relaxations as homomorphism indistinguishability relations: Dvořák [63] showed that two graphs satisfy the same sentences in k-variable first-order logic with counting quantifiers if, and only if, they are homomorphism indistinguishable over all graphs of treewidth less than k. In a well received recent paper, Mančinska & Roberson [124] proved that two graphs are quantum isomorphic if, and only if, they are homomorphism indistinguishable over all planar graphs.

Subsequently, many other graph isomorphism relaxation have been characterised as homomorphism indistinguishability relations over various graph classes. Table 1.1 gives an overview of these results.[1] Constitutive for the beauty of homomorphism indistinguishability is the surprising fact these characterisations involve arguably natural graph classes. In fact, all but two graph classes in Table 1.1 are minor-closed and thus rich in structure.

The characterisations in Table 1.1 fall into roughly two groups whose prototypical representatives are the results by Dvořák [63] and Mančinska & Roberson [124]: They equate homomorphism indistinguishability relations with *logical* or *equational* graph isomorphism relaxations.

[1] Citations for results by others which are explicitly stated in this thesis can be found where stated.

1.1.1 Logical Characterisations

Borrowing terminology from database theory or finite model theory, counting homomorphisms from a graph F to a graph G can be thought of as evaluating a query on the graph G. Instead of evaluating to true or false, such a query can evaluate to any non-negative integer. It turns out that when counting is built into the logic, then logical sentences and homomorphism counts can simulate each other. More precisely, Dvořák [63] showed that homomorphism counts are equally expressive as sentences in first-order logic with counting quantifiers.

First-order logic with counting quantifiers C is the extension of first-order logic by quantifiers $\exists^{\geq n} x$ with the semantics 'there exist at least n many values for x'. Its k-variable fragment C^k has received much attention in finite model theory, cf. [76]. Immerman & Lander [93] and Cai, Fürer, & Immerman [37] showed that two graphs are *C^k-equivalent*, i.e. they satisfy the same C^k-sentences, if, and only if, they are not distinguished by the k-dimensional Weisfeiler–Leman algorithm [174]. Recently, connections between the Weisfeiler–Leman algorithm and graph neural networks [177, 130] led to a substantial interest of machine learning specialists in this algorithm. See [139, 100, 81] for further background on the Weisfeiler–Leman algorithm and the logic C.

A connection between counting logic and homomorphism indistinguishability was established independently by Dvořák [63] and Dell, Grohe, & Rattan [61]. They showed that two graphs are C^k-equivalent if, and only if, they are homomorphism indistinguishable over all graphs of treewidth less than k. Thereby, the wealth of connections to other areas possessed by C^k-equivalence is inherited by homomorphism indistinguishability. Furthermore, as this thesis demonstrates, homomorphism indistinguishability is a versatile framework for proving such reformulations.

The characterisation of C^k-equivalence was later paralleled by a characterisation of C_q-equivalence in terms of homomorphism indistinguishability. Here, C_q is the fragment of C of formulas of quantifier-depth at most q. Grohe [79] showed that two graphs are C_q-equivalent if, and only if, they are homomorphism indistinguishable over all graphs of treedepth at most q.

The results [61, 63, 79] were proven using elementary and linear-algebraic techniques. Recently, they have been reproven and extended using a category-theoretic formalism. The starting point of this line of research is the Ehrenfeucht–Fraïssé game for C^k introduced by Hella [91], namely the *bijective k-pebble game*, which is played by two players called Spoiler and Duplicator. Duplicator wins the bijective k-pebble game on two graphs if, and only if, they are C^k-equivalent [37]. Abramsky, Dawar, & Wang [2] cast this game in category-theoretic terms by constructing the *pebbling comonad* $\mathbb{P}_k$. The pebbling comonad is a functor which maps graphs to graphs and satisfies some additional properties. It is designed such that facets of the bijective k-pebble game correspond to its category-theoretic properties: For example, two graphs are C^k-equivalent if, and only if, they are isomorphic in the coKleisli category of $\mathbb{P}_k$. Moreover, a graph has treewidth less than k if, and only if, it admits a $\mathbb{P}_k$-coalgebra.

Notably, homomorphism indistinguishability ties in well with comonads: Dawar, Jakl, & Reggio [57] showed that coKleisli isomorphism with respect to any comonad which sends finite structures to finite structures is a homomorphism indistinguishability relation. Reggio [148] later proved that this holds more generally for every *finite-rank* comonad. Instantiated with the pebbling comonad and variants of it, these results imply those of Dvořák [63] and Grohe [79]. Conversely, Abramsky, Jakl, & Paine [3] demonstrated that every homomorphism indistinguishability relation over a graph class with mild closure properties can be characterised as coKleisli isomorphism with respect to some finite-rank comonad. Besides these general result, the comonadic perspective led to new logical characterisations of homomorphism indistinguishability relations, e.g. for homomorphism indistinguishability over graphs of bounded pathwidth [129].

Contributions. Logical characterisations play only a minor role in thesis. The following contributions were made in preparation of this thesis but are mostly omitted here.

In [68], using elementary means, we reconciled the results by Dvořák [63] and Grohe [79] characterising C^k-equivalence and C_q-equivalence by proving a characterisation of $(\mathsf{C}^k \cap \mathsf{C}_q)$-equivalence in terms of homomorphism indistinguishability over a class of graphs which admit tree decompositions whose width and depth is simultaneously bounded. This result had been previously obtained in [57] using the pebbling comonad. Moreover, we prove homomorphism indistinguishability characterisations for equivalence in guarded fragments of C.

In [111], we showed that the k-variable fragment of linear-algebraic logic does not admit a characterisation in terms of homomorphism indistinguishability (Section 7.2.2). Hence, it does not admit a characterisation in terms of coKleisli isomorphisms with respect to any finite-rank comonad [148]. Thereby, we answered a question raised by Ó Conghaile & Dawar [137] negatively. Linear-algebraic logic [56] is an extension of first-order logic which was studied as part of the quest for logic capturing the complexity class polynomial time. See [55, 108] for further background.

1.1.2 Equational Characterisations

The combinatorial problem of deciding whether there exists a graph isomorphism between two graphs can be reformulated as an integer linear program: Two graphs G and H are isomorphic if, and only if, there exists a permutation matrix $P \in \{0,1\}^{V(G)\times V(H)}$ such that $\boldsymbol{A}_G P = P \boldsymbol{A}_H$. Here, $\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ denote the adjacency matrices of G and H, respectively. A *permutation matrix* is a $\{0,1\}$-matrix all whose rows and columns contain a single 1. The constraints of this system of equations can

Graph Class	Matrix Property	Reference
All graphs	permutation	Theorem 3.1.1
Planar graphs	quantum permutation	[124, 120]
Trees	doubly stochastic	Corollary 4.2.4
Paths	pseudo-stochastic	Corollary 4.2.3
Cycles	orthogonal	Corollary 4.2.1
Cycles and paths	orthogonal pseudo-stochastic	Corollary 3.3.3 and [53, 97]

Table 1.2: Two graphs G and H are homomorphism indistinguishable over a graph class in the first column if, and only if, there exists a matrix X satisfying the corresponding matrix property such that $\boldsymbol{A}_G X = X \boldsymbol{A}_H$.

be relaxed in order to yield potentially polynomial-time *equational graph isomorphism relaxations*.

For example, instead of looking for a permutation matrix P satisfying $\boldsymbol{A}_G P = P \boldsymbol{A}_H$, one may search for a matrix X satisfying $\boldsymbol{A}_G X = X \boldsymbol{A}_H$ which has non-negative rational entries and whose rows and columns sum to 1. Such a matrix is called *doubly stochastic* and two graphs G and H are said to be *fractionally isomorphic* if there exists a doubly stochastic X satisfying $\boldsymbol{A}_G X = X \boldsymbol{A}_H$, cf. [172, 159]. Amounting to a system of linear inequalities over the rationals, fractional isomorphism can be decided in polynomial time using standard algorithms.

Tinhofer [172] showed that two graphs are fractionally isomorphic if, and only if, they are not distinguished by the one-dimensional Weisfeiler–Leman algorithm. Thus, by [61], two graphs are fractionally isomorphic if, and only if, they are homomorphism indistinguishable over all trees. Surprisingly, the feasibility of $\boldsymbol{A}_G X = X \boldsymbol{A}_H$ for matrices X satisfying many other properties can be characterised as homomorphism indistinguishability relations. These results are summarised in Table 1.2.

Fractional isomorphism is a rather coarse graph isomorphism relaxation. By introducing additional constraints, the solutions to $\boldsymbol{A}_G X = \boldsymbol{A}_H X$ can be forced to resemble encodings of graph isomorphisms more closely while maintaining tractability. It is this idea that lies behind the Sherali–Adams [169], Lovász–Schrijver [117], and Lasserre [106] hierarchies, cf. [107]. Applied to the graph isomorphism integer linear or quadratic program, these hierarchies yield sequence of linear or semidefinite programs, e.g. the level-t Sherali–Adams relaxation for $t \geq 1$, which are infeasible for more and more non-isomorphic graphs as the level t grows. The program for each level can be solved in polynomial time using standard algorithms such as the ellipsoid method (at least up to some fixed precision, cf. [13]). However, in order to distinguish all graphs on n vertices, one must solve these programs at level $\Omega(n)$ [16, 122, 138, 43]. The Sherali–Adams and Lasserre hierarchy admit connections to proof complexity, cf. [21, 13]. Whereas the Sherali–Adams hierarchy corresponds to the monomial polynomial calculus proof system and algebraic approaches to graph isomorphism based on Gröbner bases [23, 85], the Lasserre hierarchy corresponds to sum-of-square proofs.

Atserias & Maneva [16] observed that, with respect to their distinguishing power, the levels of the Sherali–Adams hierarchy and the dimensions of the Weisfeiler–Leman algorithm interleave. Independently, Malkin [122] showed that the levels of the Sherali–Adams hierarchy of the graph isomorphism quadratic program precisely match the Weisfeiler–Leman dimensions in distinguishing power. An analogous statement was later shown by Grohe & Otto [84] for a modification of the Sherali–Adams hierarchy of the graph isomorphism linear program. Thus, the feasibility of these systems can be characterised as homomorphism indistinguishability relations.

The one result in Table 1.2 which overshadows all the others in sophistication and impact is the characterisation by Mančinska & Roberson [124] of quantum isomorphism as homomorphism indistinguishability over the class of planar graphs. Quantum isomorphism is a notion from quantum information theory, which was introduced by Atserias, Mančinska, Roberson, Šámal, Severini, & Varvitsiotis [15]. Two graphs are *quantum isomorphic* if players of a certain non-local game have an entanglement-assisted strategy to convince a referee that the graphs are indistinguishable. Quantum isomorphism can also be reformulated as a system of equations: By [120], two graphs G and H are quantum isomorphic if, and only if, there exists a quantum permutation matrix $\mathcal{U}$ such that $\boldsymbol{A}_G\mathcal{U} = \mathcal{U}\boldsymbol{A}_H$. A *quantum permutation matrix* is a matrix whose entries are projections in some possibly infinite-dimensional unital C^*-algebra and whose rows and columns sum to the identity.

The result of Mančinska & Roberson [124] is remarkable for two reasons: First, it establishes a connection between homomorphism indistinguishability and the fields of quantum information and quantum group theory. It does so by using heavy mathematical machinery. Second, it implies that homomorphism indistinguishability over planar graphs is undecidable, a repercussions which will be further discussed in Section 1.3.

Contributions. The result of Mančinska & Roberson [124] was proven using (bi)labelled graphs and homomorphism tensors. These objects, which are the protagonists of a fruitful interplay of combinatorics and linear algebra, are also central to the results presented in this thesis. They are introduced in Chapter 3.

Matrix Equations from Homomorphism Indistinguishability. In Chapter 4, we construct matrix equations whose feasibility characterises homomorphism indistinguishability over the classes of graphs of bounded treewidth, bounded pathwidth, and bounded treedepth. Thereby, we lift the equational characterisations of homomorphism indistinguishability over trees and paths from Table 1.2 to the classes of graphs of bounded treewidth and bounded pathwidth.

Reproving a result of Malkin [122], we recognise our matrix equations as the Sherali–Adams relaxations of the graph isomorphism quadratic program. In Theorem 4.0.1, we show that, for two graphs, the level-k Sherali–Adams relaxation has an arbitrary rational solution if, and only if, they are homomorphism indistinguishable over all graphs of pathwidth less than k. This answers a question raised by Dell, Grohe, & Rattan [61]. Furthermore, we introduce in Theorem 4.0.2 a variant of the Sherali–Adams hierarchy corresponding to homomorphism indistinguishability over graphs of bounded treedepth.

The linear-algebraic core of these results comprises two variants of a theorem of Specht [170] and Wiegmann [175] giving a criterion for simultaneous similarity via an orthogonal matrix. Two families of square matrices $A_1, \dots, A_n$ and $B_1, \dots, B_n$ are *simultaneous similar via a matrix X* if $A_i X = X B_i$ for all $i \in [n]$. In Theorems 4.1.4 and 4.1.10, we give criteria for simultaneous similarity via doubly stochastic and pseudo-stochastic matrices.

Homomorphism Indistinguishability from Matrix Equations. In Chapter 5, we consider a well-established equational graph isomorphism relaxation, namely the Lasserre semidefinite programming hierarchy [106] for the graph isomorphism quadratic program. Starting with these system of equations, we construct graph classes $\mathcal{L}_t$ such that the level-t Lasserre relaxation is feasible for two graphs if, and only if, they are homomorphism indistinguishable over $\mathcal{L}_t$.

As a corollary, we precisely determine the distinguishing power of the Lasserre hierarchy compared to the Sherali–Adams hierarchy. Previously, Atserias & Fijalkow [13] showed that there exists some constant c such that feasibility of the level-ct Sherali–Adams relaxation implies the feasibility of the level-t Lasserre relaxation. They thereby showed that, up to a linear offset in the number of levels, the Sherali–Adams and Lasserre hierarchy have the same distinguishing power. The constant c, however, remained unknown and dependent on implementation details. We show in Theorem 5.0.1 that c can be taken to be 3 and that this is best possible. We do so by conducting a graph-theoretic analysis of the graph classes $\mathcal{L}_t$, thus reducing the semantic question on distinguishing power to a syntactic analysis of the treewidth of the graphs in $\mathcal{L}_t$.

1.2 Closure

In the 1950s, it was believed that a graph is determined up to isomorphism by its *spectrum*, i.e. the multiset of eigenvalues of its adjacency matrix [90, footnote 16]. This belief was refuted by the construction of cospectral non-isomorphic graphs by Collatz & Sinogowitz [44, footnote 11], cf. [54]. Forty years later, it still seemed plausible that a much stronger graph invariant, namely the output of the k-dimensional Weisfeiler–Leman algorithm for some fixed k, would determine a graph up to isomorphism. This would have placed graph isomorphism in polynomial time. Dashing this hope, Cai, Fürer, & Immerman [37] showed in a seminal work that, for every $k \in \mathbb{N}$, there exist non-isomorphic graphs on $O(k)$ vertices which the k-dimensional Weisfeiler–Leman algorithm fails to distinguish.

Both cospectrality and indistinguishability under the k-dimensional Weisfeiler–Leman algorithm are homomorphism indistinguishability relations. Thereby, the initially mentioned results in fact concern the distinguishing power of homomorphism indistinguishability relations.

The central object for studying the distinguishing power of homomorphism indistinguishability relations was recently introduced by Roberson [150]: The *homomorphism distinguishing closure* $\mathrm{cl}(\mathcal{F})$ of a graph class $\mathcal{F}$ is the class of all graphs F such that all graphs G and H which are homomorphism indistinguishable over $\mathcal{F}$ admit the same number of homomorphisms from F. In other words, it is the unique maximal graph class whose homomorphism indistinguishability relation coincides with the one of $\mathcal{F}$. A graph class $\mathcal{F}$ is *homomorphism distinguishing closed* if $\mathrm{cl}(\mathcal{F}) = \mathcal{F}$, i.e. if adding even a single graph to $\mathcal{F}$ would change its homomorphism indistinguishability relation.

Determining the distinguishing power of the homomorphism indistinguishability relation of $\mathcal{F}$ amounts to determining the homomorphism distinguishing closure of $\mathcal{F}$. For example, the aforementioned result of Cai, Fürer, & Immerman [37] asserts that the homomorphism distinguishing closure of the class of graphs of treewidth at most k does not contain all graphs. In contrast, it is not generally the case that the homomorphism distinguishing closure of a proper graph classes is itself a proper graph class: For instance, Dvořák [63] showed that the homomorphism distinguishing closure of the class of 2-degenerate graphs contains all graphs.

Towards systematically understanding these phenomena, Roberson [150] conjectured that *every graph class which is closed under taking minors and disjoint unions is homomorphism distinguishing closed*. This conjecture is wide open and has been confirmed only for the class of planar graphs [150] and the class of graphs of treewidth at most k [134]. Beyond minor-closed graph classes, Roberson [150] showed that the class of graphs of vertex degree $\leq d$ is homomorphism distinguishing closed. Notably, all three of these results were obtained by analysing the graphs constructed by Cai, Fürer, & Immerman [37].

The homomorphism distinguishing closure is entwined with notions in counting and descriptive complexity: Building on the observation that graph motif counts such as (induced) subgraph counts can be written as linear combinations of homomorphism counts, Curticapean, Dell, & Marx [49] showed that the coefficients of a graph motif count with respect to the basis of homomorphism counts are what determines the complexity of the graph motif count. As observed in Lemma 7.1.2 and [134], this holds true also for the descriptive complexity of graph motif counts. Indeed, the result by Neuen [134] that the class of graphs of treewidth at most k is homomorphism distinguishing closed implies a full classifications of the graphs whose subgraph counts can be expressed in C^k, answering a question from [10]. Consequences of this classification were recently used for proving lower bounds in symmetric circuit complexity [59, 60].

Contributions. In Chapter 6, we survey the currently available results on the homomorphism indistinguishability closure. We extend the list of known homomorphism distinguishing closed graph classes by showing that the classes of graphs of bounded treedepth [68] and bounded pathwidth (Theorem 6.4.6), for $h \geq 3$, the class of graphs of treewidth at most two which do not contain the complete bipartite graph $K_{2,h}$ as a minor (Theorem 6.3.15), and the class of disjoint unions of cycles (Theorem 7.1.4) are homomorphism distinguishing closed.

Moreover, we prove Roberson's conjecture for all graph classes which are in a sense finite. Since every homomorphism distinguishing closed graph class is closed under disjoint unions, infinite graph classes arise inevitably when studying homomorphism indistinguishability over finite graph classes. We introduce the notions of *essentially finite* and *essentially profinite* graph classes (Definition 6.5.1) in order to capture the limited behaviour of graph classes arising from finite graph classes. Examples for essentially profinite graph classes include the class of all minors of a fixed graph and the class of cluster graphs, i.e. disjoint unions of arbitrarily large complete graphs. In Theorem 6.5.2, the essentially profinite graph classes which are homomorphism distinguishing closed are fully classified. Thereby, the realm of available examples of homomorphism distinguishing closed graph classes is drastically enlarged.

In Chapter 7, we provide further evidence for Roberson's conjecture by showing that if $\mathcal{F}$ is a minor-closed graph class, then $\mathrm{cl}(\mathcal{F})$ is minor-closed. We do so by establishing a connection between closure properties of graph classes and preservation properties of their homomorphism indistinguishability relations (Table 7.1). As a corollary, we show in Theorem 7.2.2 that equivalence over a self-complementary logic is a homomorphism indistinguishability relation over a minor-closed graph class if it can be characterised as homomorphism indistinguishability relation at all. This result allows to apply deep results from graph minor theory to the study of the distinguishing power of such logics (Theorem 7.2.6). Roughly speaking, a logic is *self-complementary* (Definition 7.2.1) if atomic subformulas Exy can be replaced by $\neg Exy \land (x \neq y)$. Most well-studied logics possess this property.

1.3 Complexity

The interest in graph isomorphism relaxations and their homomorphism indistinguishability characterisations originated in graph isomorphism and its unclear complexity-theoretic status. Notwithstanding considerable efforts, graph isomorphism has thus far resisted a complexity-theoretic classification since Karp [99] included it in their 1972 list of problems in NP which are not known to be NP-complete, cf. [83]. Neither known to be in polynomial time, the fastest known algorithm for graph isomorphism, due to Babai [17], runs in quasi-polynomial time. In the final part of the thesis, we return to computational complexity by considering the decision problem HomInd($\mathcal{F}$): For a fixed graph class $\mathcal{F}$, it asks to decide whether two input graphs are homomorphism indistinguishable over $\mathcal{F}$.

In virtue of the many homomorphism indistinguishability characterisations presented in Section 1.1, the problems HomInd($\mathcal{F}$) subsume deciding graph isomorphism relaxations from seemingly disparate areas. Typically, the graph classes $\mathcal{F}$ featured in these characterisations are infinite. Thus, the trivial approach to HomInd($\mathcal{F}$) of checking whether the input graphs have the same number of homomorphisms from every graph in $\mathcal{F}$ does not even render HomInd($\mathcal{F}$) decidable. Böker, Chen, Grohe, & Rattan [33] constructed esoteric graph classes for which deciding homomorphism indistinguishability is arbitrary hard. But even for minor-closed graph classes, the complexity-theoretic landscape of these problems is rather diverse:

For the class $\mathcal{G}$ of all graphs, HomInd($\mathcal{G}$) is graph isomorphism [114]. Only known to be in quasi-polynomial time [17], the precise complexity of HomInd($\mathcal{G}$) is unknown. For the class $\mathcal{P}$ of all planar graphs, HomInd($\mathcal{P}$) is quantum isomorphism and undecidable [124, 15]. For the class $\mathcal{TW}_k$ of graphs of treewidth at most k, HomInd($\mathcal{TW}_k$) can be solved with the k-dimensional Weisfeiler–Leman algorithm in time $n^{O(k)}$ for n-vertex graphs [93]. Whether this runtime is optimal, is a long standing open question [20, 22].

Although HomInd($\mathcal{TW}_k$) is in polynomial time for every k, there are infinitely many minor-closed graph classes $\mathcal{F}$ of bounded treewidth, e.g. the classes of k-outerplanar graphs, for which HomInd($\mathcal{F}$) could yet be undecidable.

Contributions. In Chapter 9, we show that HomInd($\mathcal{F}$) is in randomised polynomial time for every minor-closed graph class $\mathcal{F}$ of bounded treewidth (Theorem 9.0.1). Furthermore, for minor-closed graph classes of bounded pathwidth, we do not require randomness and give a deterministic polynomial-time algorithm for HomInd($\mathcal{F}$) (Theorem 9.0.2). As a concrete application, we show that the exact feasibility of the Lasserre semidefinite programming hierarchy for graph isomorphism can be decided in randomised polynomial time (Theorems 9.3.1 and 9.3.2).

Finally, we accompany these tractability results with various hardness result, proving, for example, that it is coNP-hard to decide whether two graphs are not distinguished by the k-dimensional Weisfeiler–Leman algorithm when k is part of the input (Theorem 9.4.1). We conclude by conjecturing a trichotomy for the complexity of HomInd($\mathcal{F}$) for minor-closed graph classes $\mathcal{F}$ (Conjecture 9.6.1).

2 Preliminaries

Write $\mathbb{N} = \{0, 1, 2, \ldots\}$ for the set of non-negative integers. For $n \in \mathbb{N}$, write $[n] = \{1, \ldots, n\}$ if $n \geq 1$ and $[0] = \emptyset$ otherwise. For $n \in \mathbb{N}$, write (n) for the tuple $(1, \ldots, n)$ if $n \geq 1$ and $(0) = ()$ for the empty tuple. Write $\mathfrak{S}_n$ for the symmetric group acting on n letters.

Let X and Y be sets. We write $Z = X \sqcup Y$ to indicate that Z is the *disjoint union* of X and Y, i.e. $X \cup Y = Z$ and $X \cap Y = \emptyset$. Slightly abusing notation, write $X \sqcup Y$ to denote a set which is the disjoint union of X and Y, i.e. a set with suitable renamed elements. Here, the elements of $X \sqcup Y$ are identified with those of X and Y. That is, X and Y are thought of as being disjoint sets in the first place.

Let X be a set. Write $2^X := \{Y \mid Y \subseteq X\}$ for the *power set* of X. For $\ell \in \mathbb{N}$, write $\binom{X}{\ell} := \{Y \subseteq X \mid |Y| = \ell\}$ for the set of subsets of X of size ℓ. Write $\binom{X}{\leq \ell} := \bigcup_{0 \leq k \leq \ell} \binom{X}{k}$. Write $\Pi(X)$ for the set of all partitions of X.

We write $\{\!\{\ldots\}\!\}$ to denote multisets. Formally, a multiset is a map $m \colon X \to \mathbb{N}$ for some set X. For $x \in X$, $m(x)$ is the multiplicity of x. For example, $\{\!\{x, x, x, y\}\!\}$ denotes the map $m \colon \{x, y\} \to \mathbb{N}$ with $x \mapsto 3$ and $y \mapsto 1$.

Let $\ell \geq 1$. We denote tuples $\boldsymbol{x} \in X^\ell$ by boldface letters. Their entries are denoted by $x_1, \ldots, x_\ell$. A map $f \colon X \to Y$ induces a map $f \colon X^\ell \to Y^\ell$ via $f(\boldsymbol{x})_i := f(x_i)$ for all $i \in [\ell]$. For two tuples $\boldsymbol{x} \in X^k$ and $\boldsymbol{y} \in X^\ell$ and $k, \ell \geq 1$, write $\boldsymbol{xy} = x_1 \ldots x_k y_1 \ldots y_\ell \in X^{k+\ell}$ for their concatenation.

We write log for the logarithm to base 2 and ln for the natural logarithm whose base is Euler's number e.

2.1 Graph Theory

All graphs in this thesis are finite, undirected, and without multiple edges. Formally, a *graph* is pair $G = (V, E)$ of a finite set V of *vertices* and a set $E \subseteq \binom{V}{2} \cup V$ of *edges*. A *simple* graph is a graph $G = (V, E)$ where $E \subseteq \binom{V}{2}$, i.e. a graph without loops. For a graph G, write $V(G)$ for its set of vertices and $E(G)$ for its set of edges. Instead of writing edges as $\{u, v\} \in E(G)$, we write $uv \in E(G)$. Here, it is not necessarily $u \neq v$.

For $n \in \mathbb{N}$, write K_n, C_n, and P_n for the *complete graph*, *cycle*, and *path*, on n vertices. That is, $V(K_n) = V(C_n) = V(P_n) = [n]$ and $E(K_n) = \binom{[n]}{2}$, $E(P_n) = \{i(i+1) \mid i \in [n-1]\}$, and $E(C_n) = E(P_n)$ for $n = 1$ and $E(C_n) = E(P_n) \cup \{1n\}$ otherwise. Note that $K_n = C_n = P_n$ for $n \in \{0, 1, 2\}$. For $n, m \in \mathbb{N}$, write $K_{n,m}$ for the *complete bipartite graph* with $V(K_{n,m}) := \{(0,i) \mid i \in [n]\} \cup \{(1,j) \mid j \in [m]\}$ and $E(K_{n,m}) := \{(0,i)(1,j) \mid i \in [n], j \in [m]\}$. A complete bipartite graph $K_{1,n}$ for $n \in \mathbb{N}$ is called a *star*.

For a graph G and $v \in V(G)$, write $N_G(v) := \{w \in V(G) \mid vw \in E(G)\}$ for the set of *neighbours* of G. Note that if v carries a loop, then $v \in N_G(v)$. A vertex $v \in V(G)$ is *isolated* if $N_G(v) \subseteq \{v\}$. Write $\deg_G(v) := |N_G(v)|$ for the *degree* of v. Write $\Delta(G) := \max_{v \in V(G)} \deg_G(v)$ for the *maximum degree* of G.

A *walk* in a graph G is a sequence of vertices $v_1, \dots, v_n \in V(G)$ such that $v_i v_{i+1} \in E(G)$ for all $i \in [n-1]$. We say that the walk is *between* the vertices v_1 and v_n. A walk is a *path* if the vertices $v_1, \dots, v_n$ are distinct. A *cycle* is a walk $v_1, \dots, v_n$ with $n \geq 3$ whose first and last vertex coincide, i.e. $v_1 = v_n$. A graph is *cyclic* if it contains a cycle and *acyclic* otherwise. A graph G is *Eulerian* if it contains a cycle $v_1, \dots, v_n \in V(G)$ which visits every edge precisely once, i.e. for every $e \in E(G)$ there exists a unique $i \in [n-1]$ such that $e = v_i v_{i+1}$.

For a graph G and a set $P \subseteq V(G)$, write $G[P]$ for the *subgraph induced by P*, i.e. the graph with vertex set P and edges uv if $u, v \in P$ and $uv \in E(G)$. For a vertex $v \in V(G)$ or a set of vertices $Q \subseteq V(G)$, write $G - Q := G[V(G) \setminus Q]$ and $G - v := G - \{v\}$. A graph G' is a *subgraph* of G, in symbols $G' \subseteq G$, if $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$.

A graph G is *connected* if there exists a walk between every pair of vertices. A *connected component* of a graph G is a set $C \subseteq V(G)$ such that $G[C]$ is connected and there exists no set $C \subsetneq C' \subseteq V(G)$ such that $G[C']$ is connected. A *cut vertex* is a vertex $v \in V(G)$ such that $G - v$ has more connected components than G.

A graph G is a *forest* if there exists a unique path between every pair of distinct vertices. A connected forest is called a *tree*.

A *rooted tree* (T, r) is a pair of a tree T and a vertex $r \in V(T)$ called the *root* of (T, r). The *out-degree* of a vertex $v \in V(T)$ in (T, r) is the number of neighbours of $u \in N_T(v)$ such that the unique path between u and r contains v. The *depth* of (T, r) is the maximal number of vertices on any path starting in r. That is, the depth of the singleton rooted tree is 1.

2.1.1 Graph Operations

For graphs G and H, write $G + H$ for their *disjoint union*. For disjoint unions of more multiple graphs $G_1, \dots, G_n$, write $\coprod_{i=1}^{n} G_i := G_1 + \dots + G_n$. For an integer $\ell \geq 1$ and a graph G, write $\ell G := \coprod_{i=1}^{\ell} G$. The *categorical product* of two graphs G and H is the graph $G \times H$ with vertex set $V(G \times H) := V(G) \times V(H)$ and edges between $vw, v'w' \in V(G) \times V(H)$ if, and only if, $vv' \in E(G)$ and $ww' \in E(H)$. The *lexicographic product* of two graphs G and H is the graph $G \cdot H$ with vertex set $V(G) \times V(H)$ and edges between $vw, v'w' \in V(G) \times V(H)$ if, and only if, $v = v'$ and $ww' \in E(H)$ or $vv' \in E(G)$.

The *complement* of a simple graph F is the simple graph $\overline{F}$ with $V(\overline{F}) = V(F)$

and $E(\overline{F}) = \binom{V(F)}{2} \setminus E(F)$. Observe that $\overline{\overline{F}} = F$ for all simple graphs F. The *full complement* of a graph G is the graph $\widehat{G}$ obtained from G by replacing every edge with a non-edge and every loop with a non-loop, and vice-versa, i.e. $E(\widehat{G}) := \left(\binom{V(G)}{2} \cup V(G)\right) \setminus E(G)$.

The *quotient* $F/\mathcal{P}$ of a simple graph F by a partition $\mathcal{P}$ of $V(F)$ is the simple graph with vertex set $\mathcal{P}$ and edges PQ for $P \neq Q$ if, and only if, there exist vertices $p \in P$ and $q \in Q$ such that $pq \in E(F)$. A graph F' can be obtained from a simple graph F by *contracting edges* if there is a partition $\mathcal{P} \in \Pi(V(F))$ such that $F[P]$ is connected for all $P \in \mathcal{P}$ and $F' = F/\mathcal{P}$. A *minor* of a simple graph F is a subgraph of a graph which can be obtained from F by contracting edges.

Let F be a simple graph with edge $uv \in E(F)$. The graph obtained from F by *subdividing* uv is the graph F' with vertex set $V(F') := V(F) \sqcup \{w\}$ and $E(F') := (E(F) \setminus \{uv\}) \cup \{uw, wv\}$. A graph F' is a *subdivision* of a simple graph F if it can be obtained from F by repeatedly subdividing edges.

2.1.2 Homomorphisms and Isomorphisms

A *homomorphism* from a graph F to a graph G is a map $h\colon V(F) \to V(G)$ such that $h(u)h(v) \in E(G)$ whenever $uv \in E(F)$. In particular, a homomorphism maps vertices carrying a loop to vertices carrying a loop. Write $\mathrm{Hom}(F, G)$ for the set of all homomorphisms from F to G and $\hom(F, G) := |\mathrm{Hom}(F, G)|$ for the number of homomorphisms from F to G. Especially in the context of Section 3.2, we write $F_G := \hom(F, G)$.

For graphs F and K, F is said to be *K-colourable* if there is a homomorphisms $F \to K$. A graph F is *n-colourable* for $n \in \mathbb{N}$ if it is K_n-colourable. The *chromatic number* $\chi(F)$ of a graph F is the least $n \in \mathbb{N}$ such that F is n-colourable. A graph F is *bipartite* if $\chi(F) \leq 2$. The graphs F and K are *homomorphically equivalent* if there exist homomorphisms $F \to K$ and $K \to F$.

It is well-known, cf. e.g. [116, (5.28)–(5.30)], that for all graphs F_1, F_2, G_1, G_2, and all connected graphs K, the following equalities hold:

$$\hom(F_1 + F_2, G) = \hom(F_1, G) \hom(F_2, G), \tag{2.1}$$

$$\hom(F, G_1 \times G_2) = \hom(F, G_1) \hom(F, G_2), \text{ and} \tag{2.2}$$

$$\hom(K, G_1 + G_2) = \hom(K, G_1) + \hom(K, G_2). \tag{2.3}$$

Conventionally, we distinguish between graphs occurring as domains and codomains of homomorphisms, cf. [116, p. 7]. The letters F, K, L will denote graphs which are the domains of homomorphisms, i.e. *left hand-side graphs*. The letters G and H will denote graphs which are the codomains of homomorphisms, i.e. *right hand-side graphs*. That is, a typical homomorphism maps F to G.

Let F and G be graphs. A *strong homomorphism* from F to G is a map $h\colon V(F) \to V(G)$ such that $uv \in E(F)$ if, and only if, $h(u)h(v) \in E(G)$.

Let G and H be a graphs. An *isomorphism* from G to H is a bijection $\varphi\colon V(G) \to V(H)$ such that $uv \in E(G)$ if, and only if, $\varphi(u)\varphi(v) \in E(H)$. Two graphs G and H are *isomorphic*, in symbols $G \cong H$, if there exists an isomorphism $\varphi\colon G \to H$. Conventionally, isomorphic graphs are regarded as equal. An *automorphism* of a graph G is an isomorphism $\varphi\colon G \to G$. The set of automorphisms of G forms a group under composition, the *automorphism group* $\operatorname{Aut}(G)$ of G.

A *graph class* is a set of graphs $\mathcal{F}$ which is closed under isomorphism, i.e. if $F \in \mathcal{F}$ and $F' \cong F$, then $F' \in \mathcal{F}$. Let $\mathcal{G}$ denote the class of all graphs. For a graph class $\mathcal{F}$, an integer $\ell \in \mathbb{N}$, and a graph K, write $\mathcal{F}_{\leq \ell} := \{F \in \mathcal{F} \mid |V(F)| \leq \ell\}$, $\mathcal{F}_{\geq \ell} := \{F \in \mathcal{F} \mid |V(F)| \geq \ell\}$, and $\mathcal{F}_K := \{F \in \mathcal{F} \mid \hom(F, K) > 0\}$. For example, $\mathcal{G}_{K_2}$ is the class of all bipartite graphs.

2.1.3 Tree, Path, and Cycle Decompositions

Treewidth measures how similar a graph is to a tree. We recall the well-known notion of a tree decomposition in the following slightly more general form.

Definition 2.1.1. Let F be a graph. An *F-decomposition* of a graph G is a pair (F, β) where β is a map $V(F) \to 2^{V(G)}$ such that

1. the union of the $\beta(v)$ for $v \in V(F)$ is equal to $V(G)$,
2. for every edge $e \in E(G)$, there exists $v \in V(F)$ such that $e \subseteq \beta(v)$,
3. for every vertex $u \in V(G)$, the set of vertices $v \in V(F)$ such that $u \in \beta(v)$ is connected in F.

The sets $\beta(v)$ for $v \in V(F)$ are called the *bags of* (F, β). The *width* of (F, β) is the maximum over all $|\beta(v)| - 1$ for $v \in V(F)$. An F-decomposition is called a *tree decomposition* if F is a tree, a *path decomposition* if F is a path, and a *cycle decomposition* if F is a cycle. The *treewidth* $\operatorname{tw}(G)$, *pathwidth*, and *cyclewidth* of a graph G is the minimum width of a tree, path, and cycle decomposition of G, respectively. For $t \geq 0$, write $\mathcal{TW}_t$ and $\mathcal{PW}_t$ for the classes of all graphs of treewidth and, respectively, pathwidth at most t. The following Lemma 2.1.2 generalises the standard [26, Lemma 8].

Lemma 2.1.2. *Let $k \geq 1$ and F be a connected graph. If a graph G has an F-decomposition of width at most $k-1$ and $|V(G)| \geq k$, then there is an F'-decomposition $\beta\colon V(F') \to 2^{V(G)}$ of G such that*

1. *$|\beta(t)| = k$ for all $t \in V(F')$, and*
2. *$|\beta(s) \cap \beta(t)| = k - 1$ for all $st \in E(F')$.*

The graph F' can be obtained from F by contracting and/or subdividing edges.

Proof. If $|V(G)| = k$, then F' can be taken to be the single vertex graph. If $|V(G)| > k$, then F must contain at least one edge. Let $\beta\colon V(F) \to 2^{V(G)}$ be the F-decomposition of width at most $k-1$. We repeatedly apply the following steps:

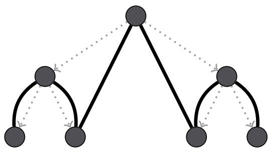


Figure 2.1: The treedepth of the path P_n is $\lceil \log(n+1) \rceil$, cf. [133, Equation (6.2)]. The grey dotted arrows $\rightarrow$ indicate the forest order $\leq$.

- If $st \in E(F)$ is such that $\beta(s) \subseteq \beta(t)$ or $\beta(t) \subseteq \beta(s)$, then the edge st in F can be contracted and the set $\beta(s) \cup \beta(t)$ can be taken to be the bag at the vertex obtained by contraction.
- If $st \in E(F)$ and $|\beta(s)| < k$ and $\beta(t) \not\subseteq \beta(s)$, then $\beta(s)$ can be enlarged by a vertex $v \in \beta(t) \setminus \beta(s)$.
- If $st \in E(F)$ and $|\beta(s)| = |\beta(t)| = k$ and $|\beta(s) \cap \beta(t)| < k - 1$, then subdivide the edge st in F by introducing a fresh vertex r. Choose vertices $v \in \beta(s) \setminus \beta(t)$ and $w \in \beta(t) \setminus \beta(s)$ and let $\beta(r) := (\beta(s) \setminus \{v\}) \cup \{w\}$.

If none of these operations can be applied, the decomposition is as desired. □

Since the class of trees (paths) is closed under contracting and subdividing edges, tree (path) decompositions satisfying the assertions of Lemma 2.1.2 can be found for all graphs of bounded treewidth (pathwidth).

2.1.4 Treedepth and Pebble Forest Covers

Treedepth measures how similar a graph is to a star. This notion was introduced by Nešetřil & Ossona de Mendez [132]. See also [133, Chapter 6] for further context.

A *forest order* on a set X is a partial order $\leq$ on X such that, for every $x \in X$, the set $\{y \in X \mid y \leq x\}$ is finite and totally ordered [57, Section II.C]. Let F be a graph. An *elimination forest* or *forest cover* of F is a forest order $\leq$ on $V(F)$ such that, for every edge $uv \in E(F)$, the vertices u and v are *comparable* with respect to to $\leq$, i.e. it holds that $u \leq v$ or $v \leq u$. The *depth* of $\leq$ is the size of the largest subset of $V(F)$ which is totally ordered under $\leq$. The *treedepth* $\mathrm{td}(F)$ of a graph F is the minimum depth of an elimination forest of F. See Figure 2.1 for an example.

A *leaf* of an elimination forest $\leq$ on F is a vertex $v \in V(F)$ such that, for all $w \in V(F)$, it holds that $v \leq w$ implies that $v = w$. Dually, a *root* is a vertex $v \in V(F)$ such that, for all $w \in V(F)$, it holds that $w \leq v$ implies that $v = w$.

In [2], k-pebble forest covers were introduced under the name *k-traversal*. They account for treedepth and -width simultaneously.

Definition 2.1.3 ([2, Section 2.3])**.** Let F be a graph and $k \geq 1$. A *k-pebble forest cover* of F is a pair of a forest cover $\leq$ of F and a function $p\colon V(F) \to [k]$ such that, for all $uv \in E(F)$ with $u \leq v$, it holds that $p(u) \neq p(w)$ for all $u < w \leq v$. Write $\mathcal{T}_q^k$ for the class of graphs which admit a k-pebble forest cover of depth at most q.

Clearly, every graph admitting a k-pebble forest cover of depth at most q has treedepth at most q. By [68, Theorem 14], a graph admits a k-pebble forest cover of depth at most q if, and only if, it admits a tree decomposition (T, β) of width at most $k-1$ containing a vertex $r \in V(T)$ such that, for every path P in T from r to a leaf in $\leq$, it holds that $\left|\bigcup_{v \in V(P)} \beta(v)\right| \leq q$. In particular, every graph admitting a k-pebble forest cover has treewidth at most $k-1$. See [68] for further graph-theoretic properties of $\mathcal{T}_q^k$.

2.2 Finite Model Theory

A *logic on graphs* [65] is a pair $(\mathsf{L}, \models)$ of a class L and a relation $\models \subseteq \mathcal{G} \times \mathsf{L}$ comparing simple graphs and elements of L which is isomorphism-invariant, i.e. satisfies that, for all $\varphi \in \mathsf{L}$ and simple graphs G and H such that $G \cong H$, it holds that $G \models \varphi$ if, and only if, $H \models \varphi$. When convenient, the reference to $\models$ is omitted and $(\mathsf{L}, \models)$ is denoted by L. Two simple graphs G and H are L-*equivalent*, in symbols $G \equiv_{\mathsf{L}} H$, if, for every $\varphi \in \mathsf{L}$, it holds that $G \models \varphi$ if, and only if, $H \models \varphi$. One may think of a logic on graphs as a mere collection of isomorphism-invariant graph properties. Every $\varphi \in \mathsf{L}$ defines such property.

2.2.1 First-Order Logic

We recall first-order logic FO following [78, 139]. *First-order logic* FO comprises variables $x, y, z, \dots$ which range over vertices of a graph. The formulas $x = y$ and Exy are *atomic* FO-formulas. More complicated FO-formulas can be formed using negation $\neg$, conjunction $\wedge$, disjunction $\vee$, existential quantification $\exists x$, and universal quantification $\forall x$. A variable x is *free* in an FO-formula φ if it appears at least once in φ outside the scope of any quantifier. An FO-*sentence* is an FO-formula without free variables.

An FO-formula $\varphi(x_1, \dots, x_k)$ with free variables $x_1, \dots, x_k$ is evaluated on a graph G by assigning vertices $v_1, \dots, v_k \in V(G)$ to the free variables. We write $G \models \varphi(v_1, \dots, v_k)$ if the formula evaluates to true.

For example, the FO-sentence $\delta := \exists x \exists y \exists z Exy \wedge Eyz \wedge Exz$ is true on a graph G if, and only if, it contains K_3 as a subgraph.

The *quantifier depth* $\mathrm{qd}(\varphi)$ of an FO-formula φ is inductively defined as follows: If φ is atomic, then $\mathrm{qd}(\varphi) := 0$. If φ is obtained from $\psi_1, \dots, \psi_\ell$ using the Boolean connectives $\neg$, $\wedge$, or $\vee$, then $\mathrm{qd}(\varphi) := \max\{\mathrm{qd}(\psi_1), \dots, \mathrm{qd}(\psi_\ell)\}$. If $\varphi = \exists x \psi$ or $\varphi = \forall x \psi$, then $\mathrm{qd}(\varphi) := \mathrm{qd}(\psi) + 1$.

2.2.2 First-Order Logic with Counting Quantifiers

First-order logic with counting quantifiers C is the extension of FO by quantifiers of the form $\exists^{\geq n} x$ for $n \in \mathbb{N}$ with the semantics that 'there exists at least n distinct assignments of vertices to x'. The counting quantifiers $\exists^{\geq n} x \varphi(x)$ can be simulated in FO via $\exists x_1 \ldots \exists x_n \left(\bigwedge_{1 \leq i < j \leq n} x_i \neq x_j \land \bigwedge_{1 \leq i \leq n} \varphi(x_i) \right)$. However, the resulting formulas are much less concise.

Write C^k for the k-variable fragment of C, i.e. the set of all formulas which contain at most k distinct variables. Note that reusing variables is permitted. For example, $\exists x \exists y (Exy \land \exists x Eyx)$ is a C^2-sentence, which expresses that a graph admits a homomorphism from the path P_3.

Write C_q for the quantifier-rank-q fragment of C, i.e. the set of all formulas $\varphi \in \mathsf{C}$ such that $\mathrm{qd}(\varphi) \leq q$. Here, the definition of the quantifier rank for FO-formulas extends to C-formulas by letting $\mathrm{qd}(\varphi) := \mathrm{qd}(\psi) + 1$ if $\varphi = \exists^{\geq n} x \psi$. Write $\mathsf{C}^k_q := \mathsf{C}^k \cap \mathsf{C}_q$.

2.2.3 Monadic Second-Order Logic

Monadic second-order logic is the name of several extensions of first-order logic. We adopt the definitions and examples from [48, Sections 1.3 and 5.2.6].

The logic MSO_1 is the extension of FO by monadic variables $X, Y, Z, \ldots$ which range over sets of vertices. In addition to the atomic FO-formulas, $x \in X$ is an atomic MSO_1-formula for a first-order variable x and a set variable X. For example, the MSO_1-sentence

$$\exists X_1 \exists X_2 \exists X_3 \, \mathrm{partition}(X_1, X_2, X_3) \land \forall x \forall y (Exy \to \neg \bigvee_{i=1}^{3} (x \in X_i \land y \in X_i))$$

expresses that a simple graph is 3-colourable. Here, the subformulas are defined as

$$\mathrm{partition}(X_1, X_2, X_3) := \forall x \bigvee_{i=1}^{3} (x \in X_i) \land \bigwedge_{1 \leq i < j \leq 3} \mathrm{disjoint}(X_i, X_j),$$
$$\mathrm{disjoint}(X, Y) := \neg \exists x (x \in X \land x \in Y).$$

Another example is the MSO_1-formula

$$\mathrm{conn}(X) := \neg \exists Y \, (\emptyset \subsetneq Y \subsetneq X \land \forall x \forall y (x \in X \land y \in X \land Exy \to (x \in Y \leftrightarrow y \in Y)))$$

which expresses that the subgraph induced by the set of vertices assigned to X is connected. Here, $\to$, $\leftarrow$, $\emptyset$, $\subsetneq$ abbreviate MSO_1-formulas which can clearly be constructed using the available expressions. Moreover, for every simple graph H, there exists a MSO_1-sentence which defines the class of all graphs which contain H as a minor. Exemplarily, with $H = K_3$, the following MSO_1-sentence defines the class of graphs containing K_3 as a minor:

$$\exists X_1 \exists X_2 \exists X_3 \bigwedge_{i=1}^{3} \mathrm{conn}(X_i) \land \bigwedge_{1 \leq i < j \leq 3} \mathrm{disjoint}(X_i, X_j) \land \exists x \exists y (x \in X_i \land y \in X_j \land Exy).$$

The logic MSO_2 is the extension of MSO_1 by first-order and monadic variables which range over edges and sets of edges. It also contains atomic formulas $\mathrm{inc}(x, y)$ where x is a first-order vertex variable and y is a first-order edge variable. The formula $\mathrm{inc}(x, y)$ evaluates to true for a vertex v and an edge e if v is *incident* to e, i.e. $v \in e$. For example, in MSO_2 one can express that a graph contains a Hamiltonian cycle.

Finally, the logics CMSO_1 and CMSO_2 are the extensions of MSO_1 and MSO_2 respectively by atomic formulas $\mathrm{card}_{p,q}(X)$ for set variables X and integers $q \geq 2$ and $q > p \geq 0$ with the semantics that the cardinality $|X|$ of X is p modulo q.

2.3 Weisfeiler–Leman Algorithm

The Weisfeiler–Leman algorithm [174] is an important graph isomorphism heuristic with direct connections to various applied and theoretic domains. See [81, Section V] and [100, 101] for background references.

Let G be a graph and $\boldsymbol{v} \in V(G)^\ell$ a tuple of vertices. Define the *atomic type* $\mathrm{atp}(G, \boldsymbol{v})$ as the equivalence class of $(G, \boldsymbol{v})$ under the equivalence relation given by $(G, \boldsymbol{v}) \sim (H, \boldsymbol{w})$ if, and only if, $v_i v_j \in E(G) \Leftrightarrow w_i w_j \in E(H)$ and $v_i = v_j \Leftrightarrow w_i = w_j$ for all $i, j \in [\ell]$.

For a tuple $\boldsymbol{v} \in V(G)^\ell$, $i \in [\ell]$, and $v \in V(G)$, write $\boldsymbol{v}[i/v] \in V(G)^\ell$ for the tuple obtained from $\boldsymbol{v}$ by replacing the i-th entry by v.

Definition 2.3.1. Let $k \geq 1$. Let G be a graph and $\boldsymbol{v} \in V(G)^k$. Define $\mathrm{wl}_k^0(G, \boldsymbol{v}) := \mathrm{atp}(G, \boldsymbol{v})$ and

$$\mathrm{wl}_k^{i+1}(G, \boldsymbol{v}) := \Big(\mathrm{wl}_k^i(G, \boldsymbol{v}), \\ \Big\{\!\!\Big\{ \Big(\mathrm{atp}(G, \boldsymbol{v}v), \mathrm{wl}_k^i(G, \boldsymbol{v}[1/v]), \dots, \mathrm{wl}_k^i(G, \boldsymbol{v}[k/v])\Big) \;\Big|\; v \in V(G) \Big\}\!\!\Big\} \Big)$$

for $i \geq 0$.

Note that the term $\mathrm{atp}(G, \boldsymbol{v}v)$ can be omitted when $k \geq 2$.

If $\mathrm{wl}_k^{i+1}(G, \boldsymbol{v}) = \mathrm{wl}_k^{i+1}(G, \boldsymbol{u})$, then $\mathrm{wl}_k^i(G, \boldsymbol{v}) = \mathrm{wl}_k^i(G, \boldsymbol{u})$. Thus, whenever the colouring of $V(G)^k$ computed by wl_k^{i+1} has more distinct colours than wl_k^i, then it is a proper refinement of the colouring computed by wl_k^i. Hence, the colouring stabilises after at most $n^k - 1$ iterations where n denotes the number of vertices in G. We denote the stable colouring by wl_k^∞ and write

$$\mathrm{wl}_k^\infty(G) := \Big\{\!\!\Big\{ \mathrm{wl}_k^\infty(G, \boldsymbol{v}) \;\Big|\; \boldsymbol{v} \in V(G)^k \Big\}\!\!\Big\}.$$

Two graphs G and H are *distinguished by* wl_k if $\mathrm{wl}_k^\infty(G) \neq \mathrm{wl}_k^\infty(H)$. By [93, Theorem 4.9.5], the partition induced by the stable colouring can be computed in time $O(k^2 n^{k+1} \log n)$ for $n := |V(G)|$.

By the following Theorem 2.3.2, the k-dimensional Weisfeiler–Leman algorithm decides whether two graphs are C^{k+1}-equivalent.

Theorem 2.3.2 ([93, Theorem 4.9.6] and [81, Corollary V.9]). *Let $k \geq 1$. For graphs G and H with $\boldsymbol{v} \in V(G)^k$ and $\boldsymbol{w} \in V(H)^k$, the following are equivalent:*

1. $\mathrm{wl}_k^\infty(G, \boldsymbol{v}) = \mathrm{wl}_k^\infty(H, \boldsymbol{w})$,
2. *for every C^{k+1}-formula $\varphi(x_1, \dots, x_k)$ with k free variables, $G \models \varphi(\boldsymbol{v})$ if, and only if, $H \models \varphi(\boldsymbol{w})$.*

In particular, $\mathrm{wl}_k^\infty(G) = \mathrm{wl}_k^\infty(H)$ if, and only if, G and H are C^{k+1}-equivalent.

It is sometimes more convenient to work with the classical *Colour Refinement* cr algorithm, as defined below, than with the 1-dimensional Weisfeiler–Leman algorithm wl_1. By [81, Proposition V.4], two graphs are distinguished by cr if, and only if, they are distinguished by wl_1.

Definition 2.3.3. Let G be a graph and $v \in V(G)$. Define $\mathrm{cr}^0(G, v) = 1$ and

$$\mathrm{cr}^{i+1}(G, v) := \left(\mathrm{cr}^i(G, v), \left\{\!\!\left\{\mathrm{cr}^i(G, v') \;\middle|\; v' \in N_G(v)\right\}\!\!\right\}\right).$$

Write $\mathrm{cr}^\infty(G, v)$ for the finest such colouring and $\mathrm{cr}^\infty(G) := \{\!\{\mathrm{cr}^\infty(G, v) \mid v \in V(G)\}\!\}$. Two graphs G and H are *distinguished by* cr if $\mathrm{cr}^\infty(G) \neq \mathrm{cr}^\infty(H)$.

2.4 Linear Algebra

We consider finite-dimensional vector spaces over the rationals $\mathbb{Q}$, the reals $\mathbb{R}$, the complex numbers $\mathbb{C}$, or over finite fields $\mathbb{F}_p$ for primes p. For the purpose of this section, let $\mathbb{K}$ denote any such field. See [104, 105] for basics of linear algebra.

For a vector space V, write $\mathrm{End}(V)$ for the vector space of *endomorphisms* of V, that is the set of linear maps $V \to V$. Let $\mathrm{id}_V \in \mathrm{End}(V)$ denote the identity map.

For a finite-dimensional $\mathbb{K}$-vector space V, write $\dim(V)$ for its *dimension*. If U is a subspace of V, we write $U \leq V$. For a vector space V and a set $X \subseteq V$, write $\mathrm{span}(X) := \bigcap_{X \subseteq U \leq V} U$ for the vector space *spanned by* X. Here, U ranges over subspaces of V.

2.4.1 Inner-Product Spaces and Adjoints

In this section, we consider vector spaces without explicit basis. Therefore, we work with abstract vectors rather than with arrays of numbers.

Let V be a $\mathbb{K}$-vector space for some field $\mathbb{K}$. A *non-degenerate inner-product* [105, Section V.1] on V is a map $\langle -, - \rangle : V \times V \to \mathbb{K}$ such that

1. $\langle v, w \rangle = \langle w, v \rangle$ for all $v, w \in V$,
2. $\langle v, \alpha w + \beta x \rangle = \alpha \langle v, w \rangle + \beta \langle v, x \rangle$ for all $v, w, x \in V$ and $\alpha, \beta \in \mathbb{K}$,
3. for every $w \in V$, if $\langle v, w \rangle = 0$ for all $v \in V$, then $w = 0$.

We require the following fact to define the adjoint of a linear map between finite-dimensional spaces with non-degenerate inner-products. Its proof is standard, cf. e.g. [105, VII, §1, Lemma 1.1].

Fact 2.4.1. *Let $\mathbb{K}$ be a field. Let V and W be finite-dimensional $\mathbb{K}$-vector spaces with non-degenerate inner-products $\langle -,-\rangle_V$ and $\langle -,-\rangle_W$. For every linear map $f\colon V \to W$, there exists a unique linear map $g\colon W \to V$ such that $\langle v, g(w)\rangle_V = \langle f(v), w\rangle_W$ for all $v \in V$ and $w \in W$.*

On complex vector spaces, we consider Hermitian products. Let V be a $\mathbb{C}$-vector space. A *positive definite Hermitian product* [105, p. 108] is a map $\langle -,-\rangle : V \times V \to \mathbb{C}$ such that the following hold:

1. $\langle v,w\rangle = \overline{\langle w,v\rangle}$ for all $v,w \in V$,
2. $\langle v, \alpha w + \beta x\rangle = \alpha \langle v,w\rangle + \beta \langle v,x\rangle$ for all $v,w,x \in V$ and $\alpha,\beta \in \mathbb{C}$,
3. for every $v \in V$, $\langle v,v\rangle$ is a non-negative real and $\langle v,v\rangle = 0$ if, and only if, $v = 0$.

That is, a positive definite Hermitian product is linear in the second component. The following analogue of Fact 2.4.1 holds:

Fact 2.4.2. *Let V and W be finite-dimensional $\mathbb{C}$-vector spaces with positive definite Hermitian products $\langle -,-\rangle_V$ and $\langle -,-\rangle_W$. For every linear map $f\colon V \to W$, there exists a unique linear map $g\colon W \to V$ such that $\langle v, g(w)\rangle_V = \langle f(v), w\rangle_W$ for all $v \in V$ and $w \in W$.*

Subsequently, we do not make a distinction between $\mathbb{K}$-vector spaces over an arbitrary field $\mathbb{K}$ with non-degenerate inner-product and $\mathbb{C}$-vector spaces with positive definite Hermitian products. We call such a vector space *inner-product space*.

Let V be an inner-product space with inner product $\langle -,-\rangle$. Two vectors $v,w \in V$ are *orthogonal* if $\langle v,w\rangle = 0$. In this case, write $v \perp w$. For two sets $U,W \subseteq V$, write $U \perp W$ if $u \perp w$ for all $u \in U$ and $w \in W$. Let $U \leq V$ be a subspace. The *orthogonal complement* $U^\perp := \{v \in V \mid u \perp v \ \forall u \in U\}$ of U is the vector space of all vectors orthogonal to all vectors in U.

Let V and W be inner-product spaces over $\mathbb{K}$ with inner-products $\langle -,-\rangle_V$ and $\langle -,-\rangle_W$. Let $X\colon V \to W$ be a linear map. By Facts 2.4.1 and 2.4.2, there exists a unique map $X^*\colon W \to V$ such that $\langle Xv, w\rangle_W = \langle v, X^*w\rangle_V$ for all $v \in V$ and $w \in W$. The map X^* is the *adjoint* of X.

If X corresponds to a matrix $A = (a_{ij}) \in \mathbb{K}^{n\times m}$ after choosing suitable bases and $\mathbb{K} \neq \mathbb{C}$, then X^* corresponds to the *transpose* $A^T := (a_{ji}) \in \mathbb{K}^{m\times n}$ of A. If V and W are over $\mathbb{C}$, then X^* corresponds to the *conjugate transpose* $A^* := \overline{A}^T = (\overline{a_{ji}})$ of A.

The map X is *orthogonal* if $X^*X = \mathrm{id}_V$ and $XX^* = \mathrm{id}_W$ for $X^*\colon W \to V$. Over $\mathbb{C}$, we refer to orthogonal matrices as *unitary*. Note that X is orthogonal if, and only if, $\langle Xv, Xv'\rangle_W = \langle v,v'\rangle_V$ for all $v,v' \in V$.

Let V be an inner-product space. An endomorphism $X \in \mathrm{End}(V)$ is *symmetric* or *Hermitian* if $X^* = X$. A symmetric endomorphism X is *positive semidefinite*, in symbols $X \succeq 0$, if $\langle v, Xv \rangle \geq 0$ for all $v \in V$.

Gram–Schmidt Orthogonalisation

The Gram–Schmidt orthogonalisation procedure is a well-known method in linear algebra used to transform a collection of vectors into a orthonormal basis [105, V, §2, Theorem 2.1]. For the purpose of Section 4.1, a variant of it is required to construct unitary linear maps between vector spaces spanned by possibly infinite sequences of vectors.

Lemma 2.4.3. *Let V and W be finite-dimensional inner-product spaces over a field $\mathbb{K}$. Let $\langle -, - \rangle_V$ and $\langle -, - \rangle_W$ denote their respective inner-products. Let I be a possibly infinite set. Let $(v_i)_{i \in I}$ and $(w_i)_{i \in I}$ be two sequences of vectors such that $v_i \in V$ and $w_i \in W$ for all $i \in I$. Suppose that*

1. *the v_i for $i \in I$ span V, the w_i for $i \in I$ span W, and*
2. *$\langle v_i, v_j \rangle_V = \langle w_i, w_j \rangle_W$ for all $i, j \in I$.*

Then there exists a unitary linear map $\Phi \colon V \to W$ such that $\Phi(v_i) = w_i$ for all $i \in I$.

Proof. Since the v_i for $i \in I$ span V, there exists a finite set $I' \subseteq I$ such that the v_i for $i \in I'$ are a basis for V. We define Φ on this basis by $v_i \mapsto w_i$ for $i \in I'$.

The map Φ is such that $\Phi(v_k) = w_k$ for all $k \in I$. Indeed, let $k \in I$ be arbitrary. Then there exist $\alpha_i \in \mathbb{K}$, $i \in I'$, such that $v_k = \sum_{i \in I'} \alpha_i v_i$. Then for arbitrary $j \in I$,

$$\left\langle w_j, \Phi(v_k) \right\rangle_W = \sum_{i \in I'} \alpha_i \left\langle w_j, \Phi(v_i) \right\rangle_W = \sum_{i \in I'} \alpha_i \left\langle w_j, w_i \right\rangle_W = \left\langle v_j, v_k \right\rangle_V = \left\langle w_j, w_k \right\rangle_W .$$

Thus, $\Phi(v_k) = w_k$. It follows that Φ is unitary. Indeed, $\left\langle \Phi v_k, \Phi v_j \right\rangle_W = \left\langle w_k, w_j \right\rangle_W = \left\langle v_k, v_j \right\rangle_V$ for all $k, j \in I$. □

2.4.2 Tensors

Let $\mathbb{K}$ be a field. Contrasting Section 2.4.1, the vector spaces in this section are of the form $\mathbb{K}^I$ for some finite set I. Therefore, they have an implicit basis given by the standard basis vectors e_i for every $i \in I$.

Note that the index sets I considered in this thesis are not ordered. Typically, the set I will be taken to be the vertex set of some graph, which is not assumed to be ordered, since graph isomorphism and its relaxations are trivial for ordered graphs.

Definition 2.4.4. Let $k, \ell \in \mathbb{N}$. Let I be a finite set. A (k, ℓ)*-shaped tensor* is an element $\varphi \in \mathbb{K}^{I^k \times I^\ell}$. Write $\mathfrak{T}_I(k, \ell) = \mathbb{K}^{I^k \times I^\ell}$ for the $\mathbb{K}$-vector space of all (k, ℓ)-shaped tensors.

For example, $\mathfrak{T}_I(k,\ell)$ contains the tensor $\mathbf{1}$ which is one everywhere. We abbreviate $\mathfrak{T}_I(k) := \mathfrak{T}_I(k,0)$. Strictly speaking, $\mathfrak{T}_I(k,\ell)$ also depends on the field $\mathbb{K}$. To improve legibility, we choose to suppress this from our notation. We consider the following operations on tensors.

Definition 2.4.5. Let $k,\ell,m,n \in \mathbb{N}$. Let I be a finite set. Let $\varphi,\psi \in \mathfrak{T}_I(k,\ell)$, $\chi \in \mathfrak{T}_I(\ell,m)$, $\xi \in \mathfrak{T}_I(m,n)$, $\zeta \in \mathfrak{T}_I(k,k)$, and $\mu,\nu \in \mathfrak{T}_I(k,0)$. Define

1. the *sum-of-entries* $\operatorname{soe}(\varphi) := \sum_{\boldsymbol{x}\in I^k,\boldsymbol{y}\in I^\ell} \varphi(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{K}$,
2. the *trace* $\operatorname{tr}(\zeta) := \sum_{\boldsymbol{x}\in I^k} \zeta(\boldsymbol{x},\boldsymbol{x}) \in \mathbb{K}$,
3. the *matrix product* $\varphi\cdot\chi \in \mathfrak{T}_I(k,m)$ via $(\varphi\cdot\chi)(\boldsymbol{x},\boldsymbol{z}) := \sum_{\boldsymbol{y}\in I^\ell} \varphi(\boldsymbol{x},\boldsymbol{y})\chi(\boldsymbol{y},\boldsymbol{z})$ for all $\boldsymbol{x}\in I^k$ and $\boldsymbol{z}\in I^m$,
4. the *tensor product* $\varphi\otimes\xi \in \mathfrak{T}_I(k+m,\ell+n)$ via
$$(\varphi\otimes\xi)(\boldsymbol{x},\boldsymbol{y}) := \varphi(x_1\dots x_k, y_1\dots y_\ell)\xi(x_{k+1}\dots x_{k+m}, y_{\ell+1}\dots y_{\ell+n})$$
for all $\boldsymbol{x}\in I^{k+m}$ and $\boldsymbol{y}\in I^{\ell+n}$,
5. the *Schur product* $\varphi\odot\psi \in \mathfrak{T}_I(k,\ell)$ via $(\varphi\odot\psi)(\boldsymbol{x},\boldsymbol{y}) := \varphi(\boldsymbol{x},\boldsymbol{y})\psi(\boldsymbol{x},\boldsymbol{y})$ for all $\boldsymbol{x}\in I^k$ and $\boldsymbol{y}\in I^\ell$,
6. the *iterated Schur product* $\varphi^{\odot 0} := \mathbf{1}$ and $\varphi^{\odot r} := \varphi\odot\varphi^{\odot r-1}$ for all integers $r\geq 1$,
7. for $\sigma\in\mathfrak{S}_{k+\ell}$, the *permuted tensor* $\varphi^\sigma \in \mathfrak{T}_I(k,\ell)$ via $\varphi^\sigma(\boldsymbol{x},\boldsymbol{y}) := \varphi(\boldsymbol{u},\boldsymbol{v})$ where $\boldsymbol{u}_i := (\boldsymbol{xy})_{\sigma^{-1}(i)}$ and $\boldsymbol{v}_{j-\ell} := (\boldsymbol{xy})_{\sigma^{-1}(j)}$ for all $1\leq i\leq \ell < j \leq \ell+k$ and $\boldsymbol{x}\in I^k$, $\boldsymbol{y}\in I^\ell$,
8. the *adjoint tensor* $\varphi^* \in \mathfrak{T}_I(\ell,k)$ via $\varphi^*(\boldsymbol{x},\boldsymbol{y}) := \varphi(\boldsymbol{y},\boldsymbol{x})$ for all $\boldsymbol{x}\in I^\ell$ and $\boldsymbol{y}\in I^k$. If the tensor is over $\mathbb{C}$, then $\varphi^*(\boldsymbol{x},\boldsymbol{y}) := \overline{\varphi(\boldsymbol{y},\boldsymbol{x})}$ for all $\boldsymbol{x}\in I^\ell$ and $\boldsymbol{y}\in I^k$, and
9. the *inner-product* $\langle\mu,\nu\rangle := \mu^*\cdot\nu \in \mathbb{K} = \mathfrak{T}_I(0,0)$.

Vandermonde Interpolation

At several points, we will encounter vector spaces $V\leq\mathbb{K}^I$ for some field $\mathbb{K}$ and a finite set I which are *closed under Schur products*, i.e. if $v,w\in V$, then also $v\odot w\in V$. Such spaces admit bases which can be constructed using iterated Schur products.

Fact 2.4.6 ([105, p. 155])**.** *Let $\mathbb{K}$ be a field and I be a finite set of size n. If $v\in\mathbb{K}^I$ has distinct entries, then $\mathbf{1}, v, v^{\odot 2},\dots,v^{\odot(n-1)}$ is a basis of $\mathbb{K}^I$.*

Fact 2.4.6 will be applied in the following form:

Lemma 2.4.7 (Vandermonde Interpolation)**.** *Let $\mathbb{K}$ be a field and I be a finite set. Let $V\leq\mathbb{K}^I$ be a vector space closed under Schur products that contains $\mathbf{1}$. Consider the equivalence relation $\sim$ on I given by $x\sim x'$ if, and only if, $v(x)=v(x')$ for all $v\in V$. Then the indicator vectors $\mathbf{1}_{C_1},\dots,\mathbf{1}_{C_r}$ of the equivalence classes of $\sim$ form a basis of V.*

Proof. Write W for the space spanned by the indicator vectors $\mathbf{1}_{C_1},\dots,\mathbf{1}_{C_r}$. Clearly, $V\leq W$. Conversely, fix distinct equivalence classes $C,C'\subseteq I$. There exists a vector $v\in V$ such that $v(i)\neq v(j)$ for some $i\in C$ and $j\in C'$. Furthermore, v is constant on

C and C'. By Fact 2.4.6, $\operatorname{span}\{\mathbf{1}, v\} \leq V$ contains a vector $v_{C,C'}$ such that $v_{C,C'}(i) = 1$ for all $i \in C$ and $v_{C,C'}(i) = 0$ for all $i \in C'$. The vector $v_{C,C'}$ may assume arbitrary values on $I \setminus (C \cup C')$. By iterating this construction for all equivalence classes, one may write the indicator vector $\mathbf{1}_C$ as $\bigodot_{C' \neq C} v_{C,C'} \in V$ where the Schur product ranges over all equivalence classes C' of $\sim$ other than C. Hence $W \leq V$. □

Pseudo-Stochastic and Doubly-Stochastic Matrices

Let $\mathbb{K}$ be a field. Let I and J be finite sets. Fix inner-product spaces $V \leq \mathbb{K}^I$ and $W \leq \mathbb{K}^J$ containing the all-ones vectors, i.e. $\mathbf{1}_I \in V$ and $\mathbf{1}_J \in W$. Then a map $X \colon V \to W$ is *pseudo-stochastic* if $X\mathbf{1}_I = \mathbf{1}_J$ and $X^*\mathbf{1}_J = \mathbf{1}_I$ for X^* the adjoint of X. A matrix $X \in \mathbb{K}^{I \times J}$ is *pseudo-stochastic* if $X\mathbf{1}_I = \mathbf{1}_J$ and $X^*\mathbf{1}_J = \mathbf{1}_I$. A matrix $X \in \mathbb{R}^{I \times J}$ is *doubly stochastic* if it is pseudo-stochastic and its entries are non-negative reals.

For the sake of completeness, we give a proof of the following lemma on doubly stochastic matrices.

Lemma 2.4.8 ([172, Lemma 1])**.** *Let I and J be finite sets. Let $v \in \mathbb{R}^I$ and $w \in \mathbb{R}^J$. If $X \in \mathbb{R}^{I \times J}$ is doubly stochastic such that $Xw = v$ and $X^T v = w$, then $v(i) = w(j)$ for all $i \in I$ and $j \in J$ such that $X(i,j) > 0$.*

Proof. By [128, Theorem 1a], the assumptions $Xw = v$ and $X^T v = w$ imply that the vectors v and w have the same multisets of entries.

Let $A \subseteq I$, respectively $B \subseteq J$, denote the set of indices on which v, respectively w, assumes its least value $r \in \mathbb{R}$. It is claimed that if $i \in I \setminus A$ and $j \in B$ or if $i \in A$ and $j \in J \setminus B$, then $X(i,j) = 0$. The claim follows by induction on the number of different values in the vectors v and w. First observe that, for $j \in B$,

$$r = v(i) = (Xw)(i) = r \sum_{j \in B} X(i,j) + \sum_{j \in J \setminus B} X(i,j) w(j) \geq r \sum_{j \in J} X(i,j) = r(X\mathbf{1})(i) = r.$$

Hence, equality holds throughout. This implies that $\sum_{j \in J \setminus B} X(i,j) = 0$ as $w(j) > r$ for all $j \in J \setminus B$. It follows that $X(i,j) = 0$ for $i \in A$ and $j \in J \setminus B$. The same holds when $i \in I \setminus A$ and $j \in B$. □

2.4.3 Completely Positive Maps and Choi Matrices

Let $\mathcal{PSD}$ denote the family of real positive semidefinite matrices. By the Spectral Decomposition Theorem [105, VIII, §4, Theorem 4.3], a matrix $M \in \mathbb{R}^{n \times n}$ is positive semidefinite if, and only if, $M_{ij} = v_i^T v_j$ for some vectors $v_1, \dots, v_n$ and all $i, j \in [n]$. The vectors $v_1, \dots, v_n$ are the *Gram vectors* of M. Let $\mathcal{DNN}$ denote the family of *doubly non-negative matrices*, i.e. of entry-wise non-negative positive semidefinite matrices.

Let $m, n \in \mathbb{N}$. Let $\mathcal{K}$ be a family of matrices. A linear map $\Phi \colon \mathbb{C}^{m \times m} \to \mathbb{C}^{n \times n}$ is

- *trace-preserving* if $\operatorname{tr}(\Phi(X)) = \operatorname{tr}(X)$ for all $X \in \mathbb{C}^{m \times m}$,

- *unital* if $\Phi(\mathrm{id}_m) = \mathrm{id}_n$,
- *$\mathcal{K}$-preserving* if $\Phi(K) \in \mathcal{K}$ for all $K \in \mathcal{K}$,
- *completely $\mathcal{K}$-preserving* if $\mathrm{id}_r \otimes \Phi$ is $\mathcal{K}$-preserving for all $r \in \mathbb{N}$,
- *positive* if it is $\mathcal{PSD}$-preserving, and
- *completely positive* if it is completely $\mathcal{PSD}$-preserving.

The *Choi matrix* of Φ is $C_\Phi = \sum_{i,j=1}^{m} E_{ij} \otimes \Phi(E_{ij}) \in \mathbb{C}^{mn \times mn}$. Here, $E_{ij} \in \mathbb{C}^{m \times m}$ for $i, j \in [m]$ denotes the matrix whose (i, j)-th entry is one and all whose other entries are zero.

We recall the following lemmas from [126]. The statement of Lemma 2.4.9 for $\mathcal{PSD}$ is the well-known Choi's Theorem for Completely Positive Maps [42].

Lemma 2.4.9 ([126, Lemma 4.4]). *Consider a family of matrices $\mathcal{K} \in \{\mathcal{DNN}, \mathcal{PSD}\}$ and a linear map $\Phi\colon \mathbb{C}^{m \times m} \to \mathbb{C}^{n \times n}$ for $m, n \in \mathbb{N}$. The following are equivalent:*
1. *the map $\mathrm{id}_m \otimes \Phi$ is $\mathcal{K}$-preserving,*
2. *the Choi matrix C_Φ of Φ lies in $\mathcal{K}$,*
3. *Φ is completely $\mathcal{K}$-preserving.*

Lemma 2.4.10 ([126, Lemma 4.10]). *Let $\Phi\colon \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ for $m, n \in \mathbb{N}$ be a linear map which is completely positive, trace-preserving, and unital. Then, for every matrix $X \in \mathbb{C}^{n \times n}$ such that $\Phi^*(\Phi(X)) = X$, it holds that $\Phi(XW) = \Phi(X)\Phi(W)$ and $\Phi(WX) = \Phi(W)\Phi(X)$ for all $W \in \mathbb{C}^{n \times n}$.*

A vector space $\mathcal{A} \subseteq \mathbb{C}^{n \times n}$ is an *algebra* if it is closed under matrix multiplication. It is *unital* if it contains the identity matrix id_n. It is *self-adjoint* if it is closed under taking conjugate transposes.

Lemma 2.4.11 ([126, Lemma 5.1]). *Let $\mathcal{A}$ and $\mathcal{B}$ be self-adjoint unital subalgebras of $\mathbb{C}^{n \times n}$ for $n \in \mathbb{N}$. Let $\varphi\colon \mathcal{A} \to \mathcal{B}$ be a trace-preserving isomorphism such that $\varphi(X^*) = \varphi(X)^*$ for all $X \in \mathcal{A}$. Then there exists a unitary $U \in \mathbb{C}^{n \times n}$ such that $\varphi(X) = UXU^*$ for all $X \in \mathcal{A}$.*

Lemma 2.4.12 ([126, Lemma 4.5]). *Let $m, n \in \mathbb{N}$. Let $D \in \mathbb{C}^{m \times n}$ be a matrix and let $u \in \mathbb{C}^n$ and $v \in \mathbb{C}^m$. Then the following are equivalent:*
1. *$D(u \odot w) = v \odot (Dw)$ for all $w \in \mathbb{C}^n$,*
2. *$D_{ij} = 0$ for all $i \in [m]$ and $j \in [n]$ such that $v_i \neq u_j$,*
3. *$D^*(v \odot z) = u \odot (D^* z)$ for all $z \in \mathbb{C}^m$.*

2.5 Representation Theory of Involution Monoids

A *monoid* Γ is a possibly infinite set equipped with an associative binary operation and an identity element denoted by 1_Γ. An example for a monoid is the *endomorphism monoid* $\mathrm{End}(V)$ for a vector space V over a field $\mathbb{K}$ with composition as binary operation and id_V as identity element. A *monoid representation* of Γ is a map

$\varphi\colon \Gamma \to \mathrm{End}(V)$ such that $\varphi(1_\Gamma) = \mathrm{id}_V$ and $\varphi(gh) = \varphi(g)\varphi(h)$ for all $g, h \in \Gamma$. The representation is *finite-dimensional* if V is finite-dimensional. For every monoid Γ, there exists a representation, for example the *trivial representation* $\Gamma \to \mathrm{End}(\{0\})$ given by $g \mapsto \mathrm{id}_{\{0\}}$ for all $g \in \Gamma$.

Let $\varphi\colon \Gamma \to \mathrm{End}(V)$ and $\psi\colon \Gamma \to \mathrm{End}(W)$ be two representations of Γ over $\mathbb{K}$. Then φ and ψ are *equivalent* if there exists a $\mathbb{K}$-vector space isomorphism $X\colon V \to W$ such that $X\varphi(g) = \psi(g)X$ for all $g \in \Gamma$. Moreover, φ is a *subrepresentation* of ψ if $V \leq W$ and $\psi(g)$ restricted to V equals $\varphi(g)$ for all $g \in \Gamma$. A representation φ is *simple* if its only subrepresentations are the trivial representation and φ itself. The *direct sum* of φ and ψ denoted by $\varphi \oplus \psi\colon \Gamma \to \mathrm{End}(V \oplus W)$ is the representation that maps $g \in \Gamma$ to $\varphi(g) \oplus \psi(g) \in \mathrm{End}(V) \oplus \mathrm{End}(W) \leq \mathrm{End}(V \oplus W)$. A representation φ is *semisimple* if it is the direct sum of simple representations.

Let $\varphi\colon \Gamma \to \mathrm{End}(V)$ be a representation with subrepresentations $\psi'\colon \Gamma \to \mathrm{End}(V')$ and $\psi''\colon \Gamma \to \mathrm{End}(V'')$. Then the restriction of φ to $V' \cap V''$ is a representation as well, called the *intersection of ψ' and ψ''*. For a set $S \subseteq V$, define the *subrepresentation of φ generated by S* as the intersection of all subrepresentations $\psi'\colon \Gamma \to \mathrm{End}(V')$ of φ such that $S \subseteq V'$.

The *character* of a representation φ is the map $\chi_\varphi\colon \Gamma \to \mathbb{K}$ defined as $g \mapsto \mathrm{tr}(\varphi(g))$. Its significance stems from the following theorem, which can be traced back to Frobenius & Schur [70]. For a contemporary proof, consult [104, Theorem 7.19].

Theorem 2.5.1 (Frobenius & Schur [70])**.** *Let $\mathbb{K}$ be a field of characteristic zero. Let Γ be a monoid. Let $\varphi\colon \Gamma \to \mathrm{End}(V)$ and $\psi\colon \Gamma \to \mathrm{End}(W)$ be finite-dimensional semisimple representations over $\mathbb{K}$. Then φ and ψ are equivalent if, and only if, $\chi_\varphi = \chi_\psi$.*

The monoids studied in this work are equipped with an additional structure which ensures that their finite-dimensional representations are always semisimple: An *involution monoid*[2] is a monoid Γ with a unary operation ${}^*\colon \Gamma \to \Gamma$ such that

$$(gh)^* = h^* g^* \quad \text{and} \quad (g^*)^* = g \quad \text{for all } g, h \in \Gamma. \tag{2.4}$$

Note that if V is an inner-product space, then $\mathrm{End}(V)$ is an involution monoid with the adjoint operation $X \mapsto X^*$. Representations of involution monoids must preserve the involution operations. That is, an *involution monoid representation* $\varphi\colon \Gamma \to \mathrm{End}(V)$ is a monoid representation such that $\varphi(g^*) = \varphi(g)^*$ for all $g \in \Gamma$.

Lemma 2.5.2. *Let $\mathbb{K}$ be a field and V an inner-product space over $\mathbb{K}$. Every finite-dimensional representation $\varphi\colon \Gamma \to \mathrm{End}(V)$ of an involution monoid Γ is semisimple.*

[2] Involution monoids correspond to $*$-algebras: A $*$-algebra is an involution monoid with additional structure. Conversely, an involution monoid Γ gives rise to the $*$-algebra $\mathbb{C}\Gamma$ of formal finite $\mathbb{C}$-linear combinations of elements in Γ. Although $*$-algebras are much more studied than involution monoids, we choose to work with the latter in order to maintain a clear separation between algebraic and combinatorial arguments in Sections 4.1 and 4.3, respectively.

Proof. Let $\varphi\colon \Gamma \to \operatorname{End}(V)$ be a finite-dimensional representation of Γ. It suffices to show that, for every subrepresentation $\psi\colon \Gamma \to \operatorname{End}(W)$ of φ, there exists a subrepresentation $\psi'\colon \Gamma \to \operatorname{End}(W')$ of φ such that $\varphi = \psi \oplus \psi'$, i.e. φ acts as ψ on W and as ψ' on W'. Set W' to be the orthogonal complement of W in V. It has to be shown that $\varphi(g) \in \operatorname{End}(V)$ for every $g \in \Gamma$ can be restricted to an endomorphism of W'. Let $w \in W$ and $w' \in W'$ be arbitrary. Then $\langle \varphi(g)w', w\rangle = \langle w', \varphi(g)^* w\rangle = \langle w', \varphi(g^*)w\rangle = 0$ since $\varphi(g^*)$ maps $W \to W$ and $W \perp W'$. Hence, the image of W' under $\varphi(g)$ is contained in the orthogonal complement of W, which equals W'. Clearly, $\varphi = \psi \oplus \psi'$. □

2.6 Integer Programming for Graph Isomorphism

In this section, two integer programming approaches to graph isomorphism are introduced. The first one uses linear programs while the second one is based on quadratic programs.

2.6.1 Linear Programming

Two graphs G and H are isomorphic if, and only if, there exists a permutation matrix $X \in \{0,1\}^{V(G)\times V(H)}$ such that $\boldsymbol{A}_G X = X \boldsymbol{A}_H$ where $\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ denote the adjacency matrices of G and H respectively. A *permutation matrix* is a $\{0,1\}$-matrix whose rows and columns each contain precisely one 1. This system can be rewritten as follows:

Definition 2.6.1. For simple graphs G and H, consider the linear program $\operatorname{LP}(G, H)$ with variables X_{vw} for $vw \in V(G) \times V(H)$ and the following equations for all $v \in V(G)$ and $w \in V(H)$

$$\sum_{v' \in V(G)} X_{v'w} = 1, \tag{2.5}$$

$$\sum_{w' \in V(H)} X_{vw'} = 1, \tag{2.6}$$

$$\sum_{v' \in V(G)} \boldsymbol{A}_G(v, v') X_{v'w} = \sum_{w' \in V(H)} X_{vw'} \boldsymbol{A}_H(w', w). \tag{2.7}$$

The system in Definition 2.6.1 has a solution with values in $\{0,1\}$ if, and only if, G and H are isomorphic. The standard LP relaxation of Definition 2.6.1 was studied by Tinhofer [172]. It is commonly referred to as fractional isomorphism, cf. [144] and the monograph [159].

Theorem 2.6.2 ([172, Theorem 1])**.** *Two simple graphs G and H are* fractionally isomorphic, *i.e.* $\operatorname{LP}(G, H)$ *has a non-negative rational solution, if, and only if, they are not distinguished by the Colour Refinement algorithm.*

Sherali–Adams Relaxation of $\mathrm{LP}(G, H)$

The notion of fractional isomorphism is strengthened by the Sherali–Adams hierarchy. This hierarchy of linear programming relaxations was introduced in [169]. Applied to $\mathrm{LP}(G, H)$, it yields the following hierarchy of systems of linear equations. See [122] for details on how it is derived from $\mathrm{LP}(G, H)$ using the definition in [169].

In this thesis, we will mostly consider the Sherali–Adams hierarchy of the graph isomorphism integer quadratic program as defined in Definition 2.6.7. The following definitions and results are stated in order to avoid confusion about the variants of the Sherali–Adams hierarchy.

Definition 2.6.3. Let $k \geq 1$. For simple graphs G and H, the *level-k Sherali–Adams relaxation of* $\mathrm{LP}(G, H)$ is the system of equations with variables X_π for $\pi \in \binom{V(G)\times V(H)}{\leq k}$ and the following equations for all $\pi \in \binom{V(G)\times V(H)}{\leq k-1}$ and $v \in V(G)$, $w \in V(H)$:

$$\sum_{v' \in V(G)} X_{\pi \cup \{v'w\}} = X_\pi, \tag{2.8}$$

$$\sum_{w' \in V(H)} X_{\pi \cup \{vw'\}} = X_\pi, \tag{2.9}$$

$$X_\emptyset = 1, \tag{2.10}$$

$$\sum_{v' \in V(G)} \boldsymbol{A}_G(v, v') X_{\pi \cup \{v'w\}} = \sum_{w' \in V(H)} X_{\pi \cup \{vw'\}} \boldsymbol{A}_H(w', w). \tag{2.11}$$

It was shown independently by Atserias & Maneva [16] and Malkin [122] that the hierarchy of Sherali–Adams relaxations of $\mathrm{LP}(G, H)$ interleaves with the hierarchy of dimensions of the Weisfeiler–Leman algorithm, cf. Theorem 2.3.2:

Theorem 2.6.4 ([16, Theorem 1], [122, Theorems 1.1 and 1.2, Lemma 1.3])**.** *For simple graphs G and H and $k \geq 1$, consider the following assertions:*

1. *the level-$(k+1)$ Sherali–Adams relaxation of* $\mathrm{LP}(G, H)$ *has a non-negative rational solution,*
2. *G and H are C^{k+1}-equivalent, and*
3. *the level-k Sherali–Adams relaxation of* $\mathrm{LP}(G, H)$ *has a non-negative rational solution.*

Then the first assertion implies the second and the second assertion implies the third.

Subsequently, Grohe & Otto [84] introduced a variant of the Sherali–Adams relaxation of $\mathrm{LP}(G, H)$ called $\mathrm{ISO}(k + 1/2)$ which has a non-negative solution if, and only if, G and H are not distinguished by the k-dimensional Weisfeiler–Leman algorithm. Note that the only difference between the Sherali–Adams relaxation of $\mathrm{LP}(G, H)$ from Definition 2.6.3 and the system in Theorem 2.6.5 is the scope of Equation (2.11).

Theorem 2.6.5 ([84, Theorem 5.9])**.** *Let $k \geq 1$. Two simple graphs G and H are C^k-equivalent if, and only if, there exists a non-negative rational solution to the system of equations with variables X_π for $\pi \in \binom{V(G)\times V(H)}{\leq k}$ and*

1. *Equations* (2.8) *to* (2.10) *for all* $\pi \in \binom{V(G) \times V(H)}{\leq k-1}$ *and* $v \in V(G)$, $w \in V(H)$ *and*
2. *Equation* (2.11) *for all* $\pi \in \binom{V(G) \times V(H)}{\leq k-2}$ *and* $v \in V(G)$, $w \in V(H)$.

2.6.2 Quadratic Programming

Alternatively, one may formulate graph isomorphism as a quadratic program. Recall the definition of the atomic type atp from Section 2.3.

Definition 2.6.6. For simple graphs G and H, consider the quadratic program $\mathrm{QP}(G, H)$ with variables X_{vw} for $vw \in V(G) \times V(H)$ and equations

$$\sum_{v \in V(G)} X_{vw} = 1 \qquad \text{for all } w \in V(H), \tag{2.12}$$

$$\sum_{w \in V(H)} X_{vw} = 1 \qquad \text{for all } v \in V(G), \tag{2.13}$$

$$X_{vw} X_{v'w'} = 0 \qquad \text{for all } vw, v'w' \in V(G) \times V(H) \text{ s.t. } \mathrm{atp}_G(vv') \neq \mathrm{atp}_H(ww'). \tag{2.14}$$

The system $\mathrm{QP}(G, H)$ has a $\{0, 1\}$-solution if, and only if, G and H are isomorphic.

Sherali–Adams Relaxation of $\mathrm{QP}(G, H)$

Using the Sherali–Adams hierarchy, the system $\mathrm{QP}(G, H)$ can be turned in a linear system of equations. See [122] for details on this construction. A set $\pi = \{v_1 w_1, \dots, v_\ell w_\ell\} \subseteq V(G) \times V(H)$ is a *local isomorphism* if it holds that $\mathrm{atp}_G(v_1, \dots, v_\ell) = \mathrm{atp}_H(w_1, \dots, w_\ell)$. It is called a *local strong homomorphism* if $v_i v_j \in E(G)$ if, and only if, $w_i w_j \in E(H)$ for all $i, j \in [\ell]$.

Definition 2.6.7. Let $k \geq 1$. For simple graphs G and H, the *level-k Sherali–Adams relaxation of* $\mathrm{QP}(G, H)$ is the system $\mathrm{SA}^k(G, H)$ with variables X_π for $\pi \in \binom{V(G) \times V(H)}{\leq k}$ and equations

$$\sum_{v \in V(G)} X_{\pi \cup \{vw\}} = X_\pi \qquad \text{for all } \pi \in \binom{V(G) \times V(H)}{\leq k-1} \text{ and } w \in V(H), \tag{2.15}$$

$$\sum_{w \in V(H)} X_{\pi \cup \{vw\}} = X_\pi \qquad \text{for all } \pi \in \binom{V(G) \times V(H)}{\leq k-1} \text{ and } v \in V(G), \tag{2.16}$$

$$X_\emptyset = 1, \tag{2.17}$$

$$X_\pi = 0 \qquad \text{for all } \pi \in \binom{V(G) \times V(H)}{\leq k} \text{ such that } \pi \text{ is not a local isomorphism.} \tag{2.18}$$

Furthermore, consider the system $\mathrm{SA}^k_{\leftrightarrow}(G, H)$ with the same variables as $\mathrm{SA}^k(G, H)$ and Equations (2.15) to (2.17) and

$$X_\pi = 0 \qquad \text{for all } \pi \in \binom{V(G) \times V(H)}{\leq k} \text{ such that } \pi \text{ is not a local strong homomorphism.} \tag{2.19}$$

Following [84], Equations (2.15) to (2.17) are called the *continuity equations* and Equations (2.18) and (2.19) the *compatibility equations*. Malkin [122] showed that $\mathrm{SA}^k(G,H)$ is related to the Weisfeiler–Leman algorithm. The system $\mathrm{SA}^k(G,H)$ was also studied in [61] where it is called $\mathsf{L}^k_{\mathrm{iso}}(G,H)$. We will reprove Theorem 2.6.8 in Corollary 4.4.1. See also [61, Theorem 3].

Theorem 2.6.8 ([122, Theorems 1.2])**.** *Let $k \geq 2$. Two simple graphs G and H are not distinguished by wl_k if, and only if, $\mathrm{SA}^{k+1}(G,H)$ has a non-negative rational solution.*

Lasserre Relaxation of $\mathrm{QP}(G,H)$

The Lasserre relaxation of an integer program was introduced in [106]. We define the Lasserre relaxation of $\mathrm{QP}(G,H)$ following [126, Section 10]. See [152, Appendix A] for a comparison to the version used in [13]. See [107] for a comparison of the Lasserre and Sherali–Adams hierarchies.

Definition 2.6.9. Let $t \geq 1$. For simple graphs G and H, the *level-t Lasserre relaxation of* $\mathrm{QP}(G,H)$ is the system $\mathrm{L}^t(G,H)$ with real variables y_I for $I \in \binom{V(G)\times V(H)}{\leq 2t}$ and constraints

$$M_t(y) := (y_{I\cup J})_{I,J\in\binom{V(G)\times V(H)}{\leq t}} \succeq 0, \tag{2.20}$$

$$\sum_{v\in V(G)} y_{I\cup\{vw\}} = y_I \quad \text{for all } I \in \binom{V(G)\times V(H)}{\leq 2t-2} \text{ and } w \in V(H), \tag{2.21}$$

$$\sum_{w\in V(H)} y_{I\cup\{vw\}} = y_I \quad \text{for all } I \in \binom{V(G)\times V(H)}{\leq 2t-2} \text{ and } v \in V(G), \tag{2.22}$$

$$y_\emptyset = 1, \tag{2.23}$$

$$y_I = 0 \quad \text{for all } I \in \binom{V(G)\times V(H)}{\leq 2t} \text{ such that } I \text{ is not a local isomorphism.} \tag{2.24}$$

The Lasserre relaxation of $\mathrm{QP}(G,H)$ was studied in [43, 138], where it is shown that in order to distinguish all non-isomorphic graphs on n vertices one must consider the level-$\Omega(n)$ Lasserre relaxation.

If $\mathrm{L}^t(G,H)$ has a real solution, then $\mathrm{SA}^t(G,H)$ has a non-negative real solution. This is because Equation (2.20) implies that $y_I \geq 0$ for all $I \in \binom{V(G)\times V(H)}{\leq t}$ since the diagonal entries of a positive semidefinite matrix are non-negative.

In Theorem 5.1.8, we show that the $2t-2$ in Equations (2.21) and (2.22) can be replaced by $2t-1$ without loss of generality. That is, the system $\mathrm{L}^t(G,H)$ has a (non-negative) real solution if, and only if, the system obtained from $\mathrm{L}^t(G,H)$ by replacing Equations (2.21) and (2.22) with Equations (5.11) and (5.12) has a (non-negative) real solution. The $2t-2$ in Definition 2.6.9 is an artefact of the construction of $\mathrm{L}^t(G,H)$ from $\mathrm{QP}(G,H)$, cf. [126, Section 10] and [43, Equations (2d)–(2e)].

3 Homomorphism Indistinguishability Characterisations: A Primer

Lovász [114] showed that two graphs G and H are isomorphic if, and only if, they admit the same number of homomorphisms from every graph F. This result has given rise to what is now known as homomorphism indistinguishability.

Definition 3.0.1. Let $\mathcal{F}$ be a graph class. Two simple graphs G and H are *homomorphism indistinguishable* over $\mathcal{F}$, in symbols $G \equiv_{\mathcal{F}} H$, if $\hom(F, G) = \hom(F, H)$ for all $F \in \mathcal{F}$.

In this thesis' language, Lovász's result asserts that two graphs G and H are isomorphic if, and only if, they are homomorphism indistinguishable over all graphs. By restricting the graph class $\mathcal{F}$, one potentially obtains other *homomorphism indistinguishability relations* $\equiv_{\mathcal{F}}$. For every graph class $\mathcal{F}$, the relation $\equiv_{\mathcal{F}}$ is a *graph isomorphism relaxation*, i.e. an equivalence relation $\approx$ comparing simple graphs such that if two simple graphs G and H are isomorphic, then $G \approx H$.

Example 3.0.2. The graphs G and H depicted by Figure 3.1 are homomorphism indistinguishable over the graph class $\{K_1, P_3\}$ where K_1 is the one-vertex graph and P_3 is the path graph on 3 vertices. Indeed, $\hom(K_1, G)$ is the number of vertices in G. Hence, $\hom(K_1, G) = 7 = \hom(K_1, H)$. To compute the number of homomorphisms from P_3 to G and H, observe that, for every possible image $v \in V(G)$ of the central vertex in P_3, there are $\deg_G(v)^2$ many choices for the images of the two adjacent vertices; they can be mapped to either of the neighbours of v. It follows that $\hom(P_3, G) = \sum_{v \in V(G)} \deg_G(v)^2 = 6 \cdot 2^2 + 0^2 = 24 = 3 \cdot 1^2 + 3 \cdot 2^2 + 3^2 = \sum_{w \in V(H)} \deg_H(w)^2 = \hom(P_3, H)$.

(a) G. **(b)** H.

Figure 3.1: A pair of graphs.

This chapter and the subsequent Chapters 4 and 5 are concerned with characterisations of homomorphism indistinguishability relations. Lovász's result characterising the homomorphism indistinguishability relation of the class of all graphs as isomorphism serves as a starting point. In this chapter, such characterisations will be proven for the classes of bipartite graphs, cycles, stars, and trees in terms of combinatorial, spectral, or logical graph isomorphism relaxations. These results are the historically first results on homomorphism indistinguishability and, with few exceptions, folklore.

The main tools for deriving such characterisations are (bi)labelled graphs and homomorphism tensors. Both are introduced in this chapter. (Bi)labelled graphs are combinatorial objects which allow to construct complicated graphs from simple building blocks. Homomorphism tensors are algebraic in nature and used to keep track of homomorphism counts when assembling (bi)labelled graphs.

Homomorphism indistinguishability has been considered in a variety of other settings, such as directed graphs [115], relational structures and relational databases [57, 129, 36, 103], hypergraphs [32, 158, 157], and in general in *locally finite* categories, i.e. in a category in which the morphisms between any two objects form a finite set [57, 148, 94, 143]. The goal of this thesis is to present homomorphism indistinguishability as a rich and well-behaved phenomenon. Seeking connections to classical graph theory, this thesis focuses on simple graphs, which represent the most elementary yet versatile setting for studying homomorphism indistinguishability.

Instead of counting homomorphisms into the graphs of interest G and H, one may also count homomorphisms from G and H. This notion, which is in a sense dual to Definition 3.0.1, was studied under the name *right profile*. Despite being a natural, right profiles have the disadvantage of admitting less characterisations in terms of e.g. logic [14, 38]. Also counting homomorphism into a graph is often computationally easier than counting homomorphisms from a graph [52, 64]. For these reasons, right profiles will not be discussed in this thesis.

Chapter Outline. Beginning with Lovász's seminal result in Section 3.1, characterisations for homomorphism indistinguishability over bipartite graphs, stars, cycles, and trees are proven in Sections 3.1, 3.3, and 3.4. These sections also showcase some of the linear-algebraic techniques that are repeatedly used in subsequent arguments. In Section 3.2, (bi)labelled graphs and homomorphism tensors are introduced. The chapter is concluded in Section 3.4 by stating results of Dvořák [63], Grohe [79], and Dawar, Jakl, & Reggio [57] relating homomorphism indistinguishability to counting logic equivalences.

3.1 Lovász's Theorem

Lovász's Theorem 3.1.1 is the first result on homomorphism indistinguishability.

Theorem 3.1.1 (Lovász [114])**.** *For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over all graphs,*
2. *G and H are homomorphism indistinguishable over all graphs on at most n vertices,*
3. *G and H are isomorphic.*

The central ingredient for the proof of Theorem 3.1.1 is the following Lemma 3.1.2. The lemma has been reproven and generalised many times, cf. [14, 57].

Lemma 3.1.2 (Lovász [114])**.** *Let $n \in \mathbb{N}$. Let $\mathcal{I}$ be a collection of pairwise non-isomorphic simple graphs on at most n vertices such that, for every simple graph G on at most n vertices, there is $H \in \mathcal{I}$ such that $G \cong H$. Then the matrix $(\hom(F, G))_{F,G \in \mathcal{I}}$ is invertible.*

Proof. The proof follows [80, Theorem 4.2]. First observe that every homomorphism $h \colon F \to G$ can be written as $h = f \circ g$ where

- $g \colon F \to F'$ is surjective, i.e. for every vertex $u' \in V(F')$, there is $u \in V(F)$ such that $g(u) = u'$ and, for every edge $u'v' \in E(F')$, there is an edge $uv \in E(F)$ such that $g(uv) = u'v'$,
- $f \colon F' \to G$ is injective, i.e. if $u' \neq v'$ for $u', v' \in V(F')$, then $f(u') \neq f(v')$.

For every fixed homomorphism $h \colon F \to G$, the graph F' as above is unique. Nevertheless, there are multiple factorisations $h = f \circ g$. Indeed, if $h = f \circ g$ is as above and $\tau \in \operatorname{Aut}(F')$ is an automorphism of F', then $h = (f \circ \tau) \circ (\tau^{-1} \circ g)$ is another such factorisations. In fact, the number of factorisations $h = f \circ g$ equals $\operatorname{aut}(F') := |\operatorname{Aut}(F')|$.

Write $\operatorname{surj}(F, F')$ for the number of surjective homomorphisms $F \to F'$ and $\operatorname{inj}(F', G)$ for the number of injective homomorphisms $F' \to G$. With this notation, it holds for $F, G \in \mathcal{I}$ that

$$\hom(F, G) = \sum_{F' \in \mathcal{I}} \operatorname{surj}(F, F') \frac{1}{\operatorname{aut}(F')} \operatorname{inj}(F', G).$$

Consider the following partial order on $\mathcal{I}$. Let $F \leq F'$ for $F, F' \in \mathcal{I}$ if, and only if, $|V(F)| \leq |V(F')|$ and $|E(F)| \leq |E(F')|$. Under this ordering, the matrices $(\operatorname{surj}(F, G))_{F,G \in \mathcal{I}}$ and $(\operatorname{inj}(F, G))_{F,G \in \mathcal{I}}$ are respectively lower and upper triangular with non-zero entries on the diagonal. The middle term $(\operatorname{aut}(F'))^{-1}$ can be thought of as a diagonal matrix with non-zero diagonal entries. Thus, $(\hom(F, G))_{F,G \in \mathcal{I}}$ is the product of three invertible matrices and thus invertible itself. □

Lovász's Theorem 3.1.1 is easily deduced from Lemma 3.1.2.

Proof of Theorem 3.1.1. The only non-trivial implication is from the second to the last assertion. The following argument is by contraposition. By Lemma 3.1.2, the matrix $(\hom(F, G))_{F,G \in \mathcal{I}}$ is invertible. If G and H are non-isomorphic, then the columns $\hom(-, G)$ and $\hom(-, H)$ are linearly independent and in particular non-equal. Thus, there exists a graph F on at most n vertices such that $\hom(F, G) \neq \hom(F, H)$. □

Figure 3.2: Two graphs which are homomorphism indistinguishable over all bipartite graphs. The bipartite double covers of both graphs are isomorphic to the disjoint union of two 6-cycles.

As a corollary, a characterisation of homomorphism indistinguishability over the class of all bipartite graphs is derived in the following Corollary 3.1.3. Corollary 3.1.3 is well-known, cf. [116, Section 5.4.2], [31], [3, Footnote 9]. Here, the *bipartite double cover* of a graph G is the categorical product $G \times K_2$ of G and the complete graph K_2 on 2 vertices.

Corollary 3.1.3. *For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over all bipartite graphs,*
2. *G and H are homomorphism indistinguishable over all bipartite graphs on at most $2n$ vertices,*
3. *G and H have isomorphic bipartite double covers, i.e. $G \times K_2 \cong H \times K_2$.*

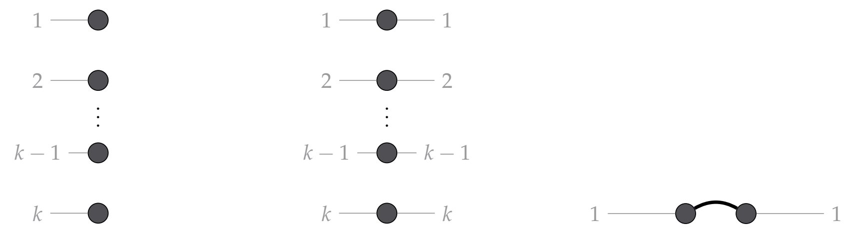
Proof. It is first argued that Item 3 implies Item 1. If $G \times K_2 \cong H \times K_2$, then $G \times K_2$ and $H \times K_2$ are homomorphism indistinguishable over all graphs. For every bipartite graph F, it holds that $\hom(F, K_2) \neq 0$. Hence, by Equation (2.2),

$$\hom(F, G) = \frac{\hom(F, G \times K_2)}{\hom(F, K_2)} = \frac{\hom(F, H \times K_2)}{\hom(F, K_2)} = \hom(F, H).$$

Thus, G and H are homomorphism indistinguishable over all bipartite graphs.

Clearly, Item 1 implies Item 2. For the remaining implication, observe that, if G and H are homomorphism indistinguishable over all bipartite graphs on at most $2n$ vertices, then $G \times K_2$ and $G \times K_2$ are homomorphism indistinguishable over all graphs on at most $2n$ vertices. Indeed, if F is non-bipartite, then $\hom(F, G \times K_2) = 0 = \hom(F, H \times K_2)$ since the bipartite double covers are bipartite. Thus, $G \times K_2 \cong G \times K_2$ follows from Theorem 3.1.1 observing that $G \times K_2$ and $H \times K_2$ have at most $2n$ vertices. □

See Figure 3.2 for an example of two non-isomorphic graphs which are homomorphism indistinguishable over all bipartite graphs. Homomorphism indistinguishability over bipartite graphs is related to the notion of cancellation, which is discussed in detail in Section 7.3. By [144, Theorem 4.2], two graphs G and H have isomorphic bipartite doubly covers if, and only if, they are *semi-isomorphic*, i.e. there



(a) $\mathbf{1} \in \mathcal{G}(k)$. **(b)** $\boldsymbol{I} \in \mathcal{G}(k, k)$. **(c)** $\boldsymbol{A} \in \mathcal{G}(1, 1)$.

Figure 3.3: The (bi)labelled graphs from Example 3.2.3 in wire notation of [124]: A vertex carries in-label (out-label) i if it is connected to the number i on the left (right) by a wire. Actual edges and vertices of the graph are depicted by black.

exists doubly stochastic matrices X and Y such that $\boldsymbol{A}_G = X\boldsymbol{A}_H Y$. Here, $\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ denote the adjacency matrices of G and H, respectively.

3.2 (Bi)Labelled Graphs and Homomorphism Tensors

In this section, the main tool of this thesis is introduced. Lying at the interface of algebra and combinatorics, (bi)labelled graphs and their homomorphism tensors allow to manipulate graphs while keeping track of their homomorphism counts. This makes them very suitable for proving characterisations of homomorphism indistinguishability relations in terms of systems of equations.

Definition 3.2.1. Let $\ell \in \mathbb{N}$. An *ℓ-labelled* graph is a tuple $\boldsymbol{F} = (F, \boldsymbol{v})$ where F is a graph and $\boldsymbol{v} \in V(F)^\ell$. Write $\mathcal{G}(\ell)$ for the class of all ℓ-labelled graphs.

For $i \in [\ell]$, the vertex v_i is said to *carry the i-th label* of $\boldsymbol{F}$. Note that the vertices in $\boldsymbol{v}$ are not necessarily distinct, i.e. vertices may have several labels.[3] Bilabelled graphs are defined similarly:

Definition 3.2.2. Let $k, \ell \in \mathbb{N}$. A *(k, ℓ)-bilabelled graph* is a tuple $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$ for $\boldsymbol{u} \in V(F)^k$ and $\boldsymbol{v} \in V(F)^\ell$. Write $\mathcal{G}(k, \ell)$ for the class of all (k, ℓ)-bilabelled graphs.

For $i \in [k]$, the vertex u_i is said to *carry the i-th in-label* of $\boldsymbol{F}$. Similarly, for $j \in [\ell]$, the vertex v_j is said to *carry the j-th out-label* of $\boldsymbol{F}$. As above, the labelled vertices do not need to be distinct. When convenient, ℓ-labelled graphs are thought of as $(\ell, 0)$-bilabelled graphs, i.e. $\mathcal{G}(\ell, 0) = \mathcal{G}(\ell)$. Unlabelled graphs are just $(0, 0)$-bilabelled graphs, i.e. $\mathcal{G} = \mathcal{G}(0) = \mathcal{G}(0, 0)$. The following example (bi)labelled graphs are depicted by Figure 3.3.

[3]Note that Definition 3.2.1 is dual to what is known as a *graph labelling* [73], where labels are assigned to vertices of a graph, rather than vertices to labels as it is the case here.

Example 3.2.3. For $k \geq 1$, let $\mathbf{1}^k \in \mathcal{G}(k)$ denote the k-labelled graph consisting of k isolated vertices with distinct labels $(1, \dots, k)$, i.e. $\mathbf{1}^k = (I, (1, \dots, k))$ with $V(I) = [k]$ and $E(I) = \emptyset$. Write $\mathbf{1} = \mathbf{1}^k$ if k is implicit. Let $\boldsymbol{I} \in \mathcal{G}(k,k)$ denote the (k,k)-bilabelled graph with $\boldsymbol{I} = (I, (1, \dots, k), (1, \dots, k))$. Let $\boldsymbol{A}$ denote the $(1,1)$-bilabelled graph consisting of a single edge whose endpoints each carry one label, i.e. $\boldsymbol{A} = (K_2, 1, 2)$ for $V(K_2) = \{1, 2\}$ and $E(K_2) = \{12\}$.

Remark 3.2.4. Let $k, \ell \in \mathbb{N}$. Two (k,ℓ)-bilabelled graphs $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$ and $\boldsymbol{F}' = (F', \boldsymbol{u}', \boldsymbol{v}')$ are *isomorphic* if there exists an isomorphism $\varphi \colon F \to F'$ such that $\varphi(\boldsymbol{u}) = \boldsymbol{v}$ and $\varphi(\boldsymbol{u}') = \boldsymbol{v}'$. Isomorphic bilabelled graphs are regarded as equal.

3.2.1 Combinatorial Operations on (Bi)Labelled Graphs

In this section, various combinatorial operations on labelled and bilabelled graphs are considered. Even when applying these operations to (bi)labelled graphs whose underlying unlabelled graph is simple, it might happen that the resulting graph has multiedges or loops. Since homomorphisms will be counted from these graphs into simple graphs, cf. Definition 3.0.1, multiedges can be disregarded. Loops need to be retained in order for Lemmas 3.2.13 and 3.2.14 and Corollary 3.2.15 to hold.

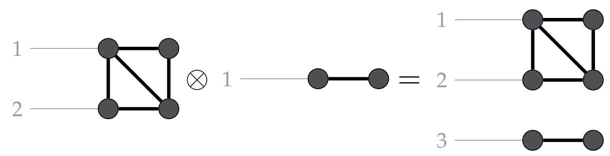
Definition 3.2.5. Let $k, \ell \in \mathbb{N}$. Let $\boldsymbol{F} = (F, \boldsymbol{u}) \in \mathcal{G}(\ell)$ and $\boldsymbol{F}' = (F', \boldsymbol{u}') \in \mathcal{G}(\ell)$. Let $\boldsymbol{F}'' = (F'', \boldsymbol{u}'') \in \mathcal{G}(k)$.

1. The graph $\operatorname{soe}(\boldsymbol{F}) := F$ is the *underlying unlabelled graph* of $\boldsymbol{F}$.
2. The *gluing product* $\boldsymbol{F} \odot \boldsymbol{F}'$ of $\boldsymbol{F}$ and $\boldsymbol{F}'$ is the ℓ-labelled graph obtained by taking the disjoint union of F and F', pairwise identifying the vertices u_i and u'_i for $i \in [\ell]$, and removing multiedges but retaining loops. The i-th label of $\boldsymbol{F} \odot \boldsymbol{F}'$ is carried by the vertex obtained from u_i and u'_i.
3. The *disjoint union* $\boldsymbol{F} \otimes \boldsymbol{F}''$ of $\boldsymbol{F}$ and $\boldsymbol{F}''$ is the $(k+\ell)$-labelled graph $(F + F'', \boldsymbol{u}\boldsymbol{u}'')$.
4. For a permutation $\sigma \in \mathfrak{S}_\ell$, the ℓ-labelled graph $\boldsymbol{F}^\sigma$ obtained from $\boldsymbol{F}$ by *permuting labels* is $\boldsymbol{F}^\sigma = (F, \sigma(\boldsymbol{u}))$ where $\sigma(\boldsymbol{u})_i := u_{\sigma(i)}$ for $i \in [\ell]$.

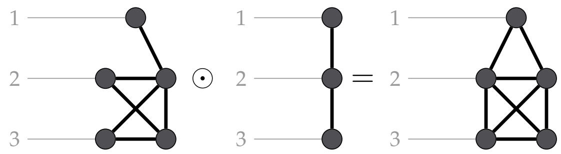
See Figures 3.4 and 3.6a for examples of these operations. On bilabelled graphs, similar operations can be defined.

Definition 3.2.6. Let $k, \ell, m, n \in \mathbb{N}$. Let $\boldsymbol{K} = (K, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{G}(k, \ell)$, $\boldsymbol{K}' = (K', \boldsymbol{u}', \boldsymbol{v}') \in \mathcal{G}(\ell, m)$, $\boldsymbol{K}'' = (K'', \boldsymbol{u}'', \boldsymbol{v}'') \in \mathcal{G}(k, \ell)$, $\boldsymbol{L} = (L, \boldsymbol{x}, \boldsymbol{y}) \in \mathcal{G}(m, n)$, and $\boldsymbol{F} = (F, \boldsymbol{u}') \in \mathcal{G}(\ell)$.

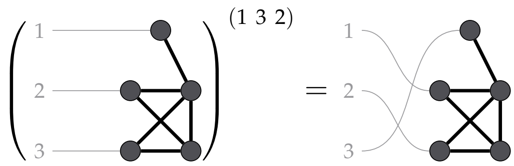
1. The graph $\operatorname{soe}(\boldsymbol{K}) := F$ is the *underlying unlabelled graph* of $\boldsymbol{K}$.
2. The *reverse* of $\boldsymbol{K}$ is $\boldsymbol{K}^* := (K, \boldsymbol{v}, \boldsymbol{u}) \in \mathcal{G}(\ell, k)$.
3. The *disjoint union* $\boldsymbol{K} \otimes \boldsymbol{L}$ of $\boldsymbol{K}$ and $\boldsymbol{L}$ is the $(k+m, \ell+n)$-bilabelled graph $(K + L, \boldsymbol{u}\boldsymbol{x}, \boldsymbol{v}\boldsymbol{y})$.
4. The *series composition* of $\boldsymbol{K}$ and $\boldsymbol{K}'$ is the (k,m)-bilabelled graph $\boldsymbol{K} \cdot \boldsymbol{K}'$ obtained by taking the disjoint union of K and K' and identifying for all $i \in [\ell]$ the vertices v_i and u'_i. In this process, multiedges are removed and loops retained. The in-labels of $\boldsymbol{K} \cdot \boldsymbol{K}'$ lie on $\boldsymbol{u}$ while its out-labels are positioned on $\boldsymbol{v}'$.



(a) Disjoint union of a 2-labelled graph with a 1-labelled graph.



(b) Gluing product of two 3-labelled graphs.



(c) Permutation of labels of a 3-labelled graph.

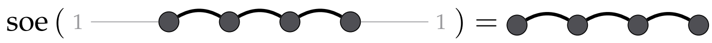
Figure 3.4: Operations on labelled graphs.

5. The *series composition* of $\boldsymbol{K}$ and $\boldsymbol{F}$ is the k-labelled graph $\boldsymbol{K} \cdot \boldsymbol{F}$ obtained by taking the disjoint union of K and F and identifying for all $i \in [\ell]$ the vertices v_i and u'_i. In this process, multiedges are removed and loops retained. The in-labels of $\boldsymbol{K} \cdot \boldsymbol{F}$ lie on $\boldsymbol{u}$.
6. The *parallel composition* of $\boldsymbol{K}$ and $\boldsymbol{K}''$ is the (k, ℓ)-bilabelled graph $\boldsymbol{K} \odot \boldsymbol{K}''$ obtained by taking the disjoint union of K and K' and identifying for all $i \in [k]$ and $j \in [\ell]$ the vertices u_i with u''_i and v_j with v''_j. In this process, multiedges are removed and loops retained. The in-labels of $\boldsymbol{K} \odot \boldsymbol{K}''$ lie on $\boldsymbol{u}$ while its out-labels are positioned on $\boldsymbol{v}$.
7. For a permutation $\sigma \in \mathfrak{S}_{k+\ell}$, write $\boldsymbol{K}^\sigma := (K, \boldsymbol{x}, \boldsymbol{y}) \in \mathcal{G}(k, \ell)$ where $x_i := (\boldsymbol{uv})_{\sigma(i)}$ and $y_{j-k} := (\boldsymbol{uv})_{\sigma(j)}$ for all $1 \leq i \leq k < j \leq k + \ell$, for the graph which is obtained from $\boldsymbol{K}$ by *permuting labels* according to σ.

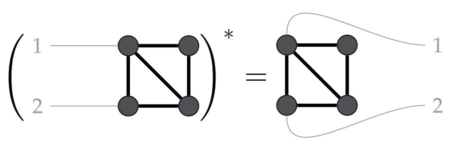
Examples for these operations are depicted by Figure 3.5. When convenient, series composition is written as juxtaposition, i.e. $\boldsymbol{K} \cdot \boldsymbol{K}' = \boldsymbol{KK}'$ and $\boldsymbol{K} \cdot \boldsymbol{F} = \boldsymbol{KF}$. The series composition of several bilabelled graphs is denoted by $\prod_{i=1}^{n} \boldsymbol{F}_i := \boldsymbol{F}_1 \cdot \dots \cdot \boldsymbol{F}_n$. Using the operations in Definitions 3.2.5 and 3.2.6, the following derived operations can be defined. See Figure 3.6 for examples.

Definition 3.2.7. Let $k \in \mathbb{N}$. Let $\boldsymbol{F}, \boldsymbol{F}' \in \mathcal{G}(k)$ and $\boldsymbol{K} \in \mathcal{G}(k, k)$.

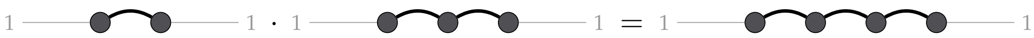
1. The *inner-product* of $\boldsymbol{F}$ and $\boldsymbol{F}'$ is the unlabelled graph $\langle \boldsymbol{F}, \boldsymbol{F}' \rangle := \operatorname{soe}(\boldsymbol{F} \odot \boldsymbol{F}')$.
2. The *trace* of $\boldsymbol{K}$ is the unlabelled graph $\operatorname{tr}(\boldsymbol{F}) := \operatorname{soe}(\boldsymbol{I} \odot \boldsymbol{K})$ for $\boldsymbol{I} \in \mathcal{G}(k, k)$ as



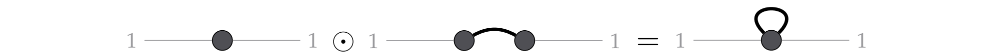
(a) Unlabelling a $(1, 1)$-bilabelled graph.



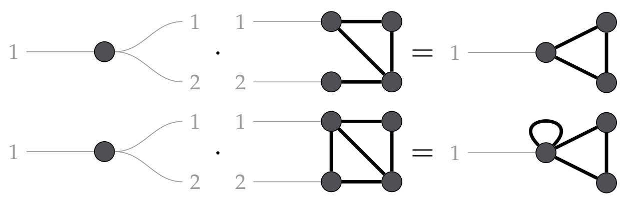
(b) Swapping labels of a $(2, 0)$-bilabelled graph. The result is a $(0, 2)$-bilabelled graph



(c) Series compositions of $(1, 1)$-bilabelled graphs.

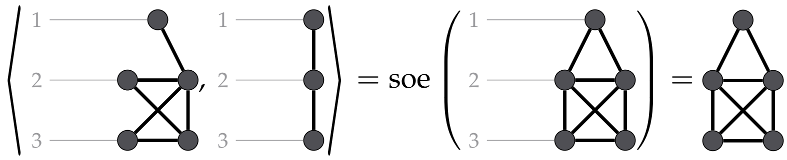


(d) Parallel compositions of $(1, 1)$-bilabelled graphs.

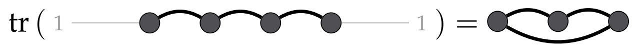


(e) Series composition of a $(1, 2)$-bilabelled graph with a 1-labelled graph. The result is a 2-labelled graphs. A multiedge is dropped and while the loop is retained.

Figure 3.5: Operations on bilabelled graphs.



(a) Inner-product of two 3-labelled graph.



(b) Trace of a $(1,1)$-bilabelled graph.

Figure 3.6: Derived operations on labelled and bilabelled graphs.

defined in Example 3.2.3.

It may appear unusual to give combinatorial operations as those in Definition 3.2.7 linear-algebraic names. The reason for this will become clear in Section 3.2.4, where combinatorial operations are proven to correspond to linear-algebraic operations.

3.2.2 Homomorphism Tensors

In plain terms, homomorphism tensors can be thought of as possibly higher dimensional arrays of numbers counting homomorphisms from (bi)labelled graphs into some fixed graph.

Definition 3.2.8. Let $k, \ell \in \mathbb{N}$ and $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{G}(k, \ell)$. For a graph G, define the *homomorphism tensor* $\boldsymbol{F}_G \in \mathbb{N}^{V(G)^k \times V(G)^\ell}$ of $\boldsymbol{F}$ with respect to G by letting $\boldsymbol{F}_G(\boldsymbol{x}, \boldsymbol{y})$ for $\boldsymbol{x} \in V(G)^k$ and $\boldsymbol{y} \in V(G)^\ell$ be the number of homomorphisms $h \colon F \to G$ such that $h(\boldsymbol{u}) = \boldsymbol{x}$ and $h(\boldsymbol{v}) = \boldsymbol{y}$.

Here, $h(\boldsymbol{u}) = \boldsymbol{x}$ means that $h(u_i) = x_i$ for all $i \in [k]$. As above, k-labelled graphs are regarded as $(k, 0)$-bilabelled graphs and Definition 3.2.8 is extended as follows: If $\boldsymbol{F} \in \mathcal{G}(k)$ is a k-labelled graph, then its homomorphism tensor with respect to G is $\boldsymbol{F}_G \in \mathbb{N}^{V(G)^k}$.

Conventionally, homomorphism tensors of labelled graphs are called *homomorphism vectors* and homomorphism tensors of bilabelled graphs *homomorphism matrices*. Since homomorphism tensors take values over $\mathbb{N}$, they can be regarded as vectors or matrices in vector spaces over any field.

Example 3.2.9. Let G be a simple graph and $x, y \in V(G)$ and recall the (bi)labelled graphs from Example 3.2.3 and Figure 3.3.

1. $\boldsymbol{1} \in \mathcal{G}(1)$ is the 1-labelled edgeless graph with a single vertex. Write I for its underlying unlabelled graph, i.e. $V(I) = \{u\}$ and $E(I) = \emptyset$. By definition, $\boldsymbol{1}_G(v)$ is the number of homomorphism from I to G which map u to x. There

is precisely one such homomorphisms. Hence, $\mathbf{1}_G \in \mathbb{N}^{V(G)}$ is the vector which is 1 everywhere.

2. $\boldsymbol{I} \in \mathcal{G}(1,1)$ is the $(1,1)$-bilabelled edgeless graph with a single vertex. Note that the in-label and the out-label of $\boldsymbol{I}$ reside on the same vertex u. By definition, $\boldsymbol{I}_G(u,v)$ is the number of homomorphisms from I to G which map u to x and v to y. Such a homomorphism exists if, and only if, $x = y$, in which case there is exactly one such homomorphism. Hence, $\boldsymbol{I}_G \in \mathbb{N}^{V(G)\times V(G)}$ is the identity matrix, i.e. the matrix which is 0 everywhere, except on the diagonal where it is 1 everywhere.
3. $\boldsymbol{A} \in \mathcal{G}(1,1)$ is the $(1,1)$-bilabelled graph on two vertices connected by an edge such that the in- and out-label reside on distinct vertices. Its underlying unlabelled graph is the 2-vertex complete graph K_2. Write $V(K_2) = \{u,v\}$ and $\boldsymbol{A} = (K_2,u,v)$. Thus, $\boldsymbol{A}_G(x,y)$ is the number of homomorphisms from K_2 to G which map u to x and v to y. Such a homomorphism exists if, and only if, $xy \in E(G)$, in which case there is exactly one such homomorphism. Thus, $\boldsymbol{A}_G \in \mathbb{N}^{V(G)\times V(G)}$ is the adjacency matrix of G.

Example 3.2.9 featured (bi)labelled graphs all whose vertices are labelled. Such (bi)labelled graphs are called atomic.

Definition 3.2.10. Let $k, \ell \in \mathbb{N}$. A (k,ℓ)-bilabelled graph $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$ is *atomic* if all its vertices are labelled, i.e. $V(F) = \{u_1, \dots, u_k, v_1, \dots, v_\ell\}$. Write $\mathcal{A}(k,\ell) \subseteq \mathcal{G}(k,\ell)$ for the class of atomic (k,ℓ)-bilabelled graphs.

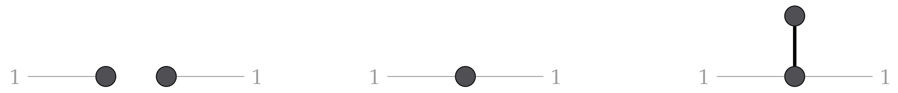
The homomorphism tensors of the atomic graphs in Example 3.2.9 take values from $\{0,1\}$. This is not a mere coincide.

Observation 3.2.11. *Let $k, \ell \in \mathbb{N}$ and $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{G}(k,\ell)$. Let G be a graph. If $\boldsymbol{F}$ is atomic, then $\boldsymbol{F}_G \in \{0,1\}^{V(G)^k \times V(G)^\ell}$.*

Proof. Let $\boldsymbol{x} \in V(G)^k$ and $\boldsymbol{y} \in V(G)^\ell$. By Definition 3.2.8, $\boldsymbol{F}_G(\boldsymbol{x},\boldsymbol{y})$ is the number of homomorphisms $h\colon F \to G$ such that $h(\boldsymbol{u}) = \boldsymbol{x}$ and $h(\boldsymbol{v}) = \boldsymbol{y}$. Since every vertex of F appears in $\boldsymbol{u}$ or $\boldsymbol{v}$, a map $h\colon V(F) \to V(G)$ is uniquely determined by the conditions $h(\boldsymbol{u}) = \boldsymbol{x}$ and $h(\boldsymbol{v}) = \boldsymbol{y}$. The map h is either a homomorphism, in which case $\boldsymbol{F}_G(\boldsymbol{x},\boldsymbol{y}) = 1$, or it is not a homomorphism, in which case $\boldsymbol{F}_G(\boldsymbol{x},\boldsymbol{y}) = 0$. ☐

3.2.3 Formal Linear Combinations of (Bi)Labelled Graphs

Finite formal linear combinations of (bi)labelled graphs $\boldsymbol{q} = \sum \alpha_i \boldsymbol{F}^i$ for $\boldsymbol{F}^i \in \mathcal{G}(k,\ell)$ and coefficients α_i from some field $\mathbb{K}$ are a notational device which allows to state some arguments more succinctly. The $\boldsymbol{F}^i$ are called the *constituents* of $\boldsymbol{q}$. The set of finite formal linear combinations of (k,ℓ)-bilabelled graphs with coefficients in $\mathbb{K}$ is denoted by $\mathbb{K}\mathcal{G}(k,\ell)$. Its elements are denoted by lowercase boldface letters. The operations from Section 3.2.1 extend linearly to operations on formal linear



(a) $\boldsymbol{J} \in \mathcal{G}(1,1)$. **(b)** $\boldsymbol{I} \in \mathcal{G}(1,1)$. **(c)** $\boldsymbol{D} \in \mathcal{G}(1,1)$.

Figure 3.7: Bilabelled graphs from Example 3.2.12.

combinations of (bi)labelled graphs. The homomorphism tensors of (bi)labelled graphs from Definition 3.2.8 can be linearly extended to $\mathbb{K}\mathcal{G}(k,\ell)$, i.e. for a graph G and $\boldsymbol{q} = \sum \alpha_i \boldsymbol{F}^i \in \mathbb{K}\mathcal{G}(k,\ell)$, let $\boldsymbol{q}_G := \sum \alpha_i \boldsymbol{F}^i_G \in \mathbb{K}^{V(G)^k \times V(G)^\ell}$.

In the literature [119, 63], linear combinations of (bi)labelled graphs are sometimes called *quantum graphs*. Given the rise of quantum graph theory, this name seems inappropriate. For example, many graph matrices studied in spectral graph theory are linear combinations of homomorphism matrices [54].

Example 3.2.12. Let $\boldsymbol{J}, \boldsymbol{I}, \boldsymbol{D}, \boldsymbol{A} \in \mathcal{G}(1,1)$ be as in Figures 3.3c and 3.7. Let G be a graph.

1. The matrix $\boldsymbol{D}_G$ is the *degree matrix* of G, i.e. for $x, y \in V(G)$, $\boldsymbol{D}_G(x,y) = \deg_G(x)$ if $x = y$ and zero otherwise.
2. The matrix $\boldsymbol{D}_G - \boldsymbol{A}_G$ is the *Laplacian matrix* of G.
3. The matrix $\boldsymbol{A}_{\overline{G}} = \boldsymbol{J}_G - \boldsymbol{I}_G - \boldsymbol{A}_G$ is the adjacency matrix of the complement $\overline{G}$ of G.

3.2.4 Correspondences between Algebraic and Combinatorial Operations

In this section, it is shown that the combinatorial operations on (bi)labelled graphs defined in Section 3.2.1 correspond to algebraic operations on homomorphism tensors, cf. Definition 2.4.5. Early references for this observation include [116, 119, 63].

Lemma 3.2.13. *Let $k, \ell \in \mathbb{N}$. Let $\boldsymbol{F} = (F, \boldsymbol{u})$ and $\boldsymbol{F}' = (F', \boldsymbol{u}')$ be ℓ-labelled graphs. Let $\boldsymbol{F}'' = (F'', \boldsymbol{u}'') \in \mathcal{G}(k)$. Let G be a graph.*

1. *The number of homomorphisms* $\hom(\operatorname{soe}(\boldsymbol{F}), G) = \operatorname{soe}(\boldsymbol{F}_G)$ *from the unlabelled graph* $\operatorname{soe}(\boldsymbol{F})$ *to G is the sum-of-entries* $\operatorname{soe}(\boldsymbol{F}_G)$ *of the homomorphism vector $\boldsymbol{F}_G$.*
2. *The homomorphism vector $(\boldsymbol{F} \odot \boldsymbol{F}')_G$ of the gluing product $\boldsymbol{F} \odot \boldsymbol{F}'$ is equal to the Schur product $\boldsymbol{F}_G \odot \boldsymbol{F}'_G$.*
3. *The homomorphism vector $(\boldsymbol{F} \otimes \boldsymbol{F}'')_G$ of the disjoint union $\boldsymbol{F} \otimes \boldsymbol{F}''$ is equal to the tensor product $\boldsymbol{F}_G \otimes \boldsymbol{F}''_G$.*
4. *For a permutation $\sigma \in \mathfrak{S}_\ell$, the homomorphism vector $(\boldsymbol{F}^\sigma)_G$ of the permuted $\boldsymbol{F}^\sigma$ is equal to the tensor $(\boldsymbol{F}_G)^\sigma$ with permuted axes.*

Proof. For $\boldsymbol{x} \in V(G)^\ell$, $\boldsymbol{F}_G(\boldsymbol{x})$ is the number of homomorphisms $h \colon F \to G$ such that $h(\boldsymbol{u}) = \boldsymbol{x}$. Thus $\operatorname{soe}(\boldsymbol{F}_G) = \sum_{\boldsymbol{x} \in V(G)^\ell} \boldsymbol{F}_G(\boldsymbol{x})$ is the total number of homomorphisms from F to G.

In order to verify that $(\boldsymbol{F} \odot \boldsymbol{F}')_G(\boldsymbol{x}) = \boldsymbol{F}_G(\boldsymbol{x})\boldsymbol{F}'_G(\boldsymbol{x})$, write $\boldsymbol{F} \odot \boldsymbol{F}' = (F'', \boldsymbol{u}'')$. A bijection between the following sets of (pairs of) homomorphisms is established.

$$\{h'' \colon F'' \to G \mid h''(\boldsymbol{u}'') = \boldsymbol{x}\} \to \{h \colon F \to G \mid h(\boldsymbol{u}) = \boldsymbol{x}\} \times \{h' \colon F' \to G \mid h'(\boldsymbol{u}') = \boldsymbol{x}\}.$$

Let $\pi \colon F \to F''$ and $\pi' \colon F' \to F''$ denote the maps that associate to a vertex of $\boldsymbol{F}$ and $\boldsymbol{F}'$ the vertex in $\boldsymbol{F} \odot \boldsymbol{F}'$ that it corresponds to under the gluing operation. Both π and π' are homomorphisms. It is claimed that $h'' \mapsto (h'' \circ \pi, h'' \circ \pi')$ is the desired bijection. This map is surjective since two homomorphisms h and h' as above can be pieced together to yield a homomorphisms h''. Then, $h'' \circ \pi = h$ and $h'' \circ \pi' = h'$. The map is injective since every vertex $v \in V(F'')$ is in the image of π or in the image of π'.

For the disjoint union, a similar argument yields the claim. For permutations, the claim is purely syntactical. □

Similar correspondences holds for the operations on bilabelled graphs from Definition 3.2.6. The proof of the following lemma is analogous to the proof of Lemma 3.2.13.

Lemma 3.2.14. *Let $k, \ell, m, n \in \mathbb{N}$. Let $\boldsymbol{K} = (K, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{G}(k, \ell)$, $\boldsymbol{K}' = (K', \boldsymbol{u}', \boldsymbol{v}') \in \mathcal{G}(\ell, m)$, $\boldsymbol{K}'' = (K'', \boldsymbol{u}'', \boldsymbol{v}'') \in \mathcal{G}(k, \ell)$, $\boldsymbol{L} = (L, \boldsymbol{x}, \boldsymbol{y}) \in \mathcal{G}(m, n)$, and $\boldsymbol{F} = (F, \boldsymbol{u}') \in \mathcal{G}(\ell)$. Let G be a graph.*

1. $(\operatorname{soe}(\boldsymbol{K}))_G = \operatorname{soe}(\boldsymbol{K}_G)$,
2. $(\boldsymbol{K}^*)_G = (\boldsymbol{K}_G)^*$,
3. $(\boldsymbol{K} \otimes \boldsymbol{L})_G = \boldsymbol{K}_G \otimes \boldsymbol{L}_G$,
4. $(\boldsymbol{K} \cdot \boldsymbol{K}')_G = \boldsymbol{K}_G \cdot \boldsymbol{K}'_G$,
5. $(\boldsymbol{K} \cdot \boldsymbol{F})_G = \boldsymbol{K}_G \cdot \boldsymbol{F}_G$,
6. $(\boldsymbol{K} \odot \boldsymbol{K}'')_G = \boldsymbol{K}_G \odot \boldsymbol{K}''_G$,
7. *For a permutation $\sigma \in \mathfrak{S}_{k+\ell}$, $(\boldsymbol{K}^\sigma)_G = (\boldsymbol{K}_G)^\sigma$.*

The following Corollary 3.2.15 justifies the choices of the operations' names in Definition 3.2.7.

Corollary 3.2.15. *Let $k \in \mathbb{N}$. Let $\boldsymbol{F}, \boldsymbol{F}' \in \mathcal{G}(k)$ and $\boldsymbol{K} \in \mathcal{G}(k, k)$. Let G be a graph.*

1. $\langle \boldsymbol{F}, \boldsymbol{F}' \rangle_G = \langle \boldsymbol{F}_G, \boldsymbol{F}'_G \rangle$,
2. $(\operatorname{tr}(\boldsymbol{F}))_G = \operatorname{tr}(\boldsymbol{F}_G)$.

The following example illustrates why loops must be retained in the operations defined in Section 3.2.1.

Example 3.2.16. In the second identity in Figure 3.5d, a loop arose as the result of a parallel composition. Write this identity as $\boldsymbol{I} \odot \boldsymbol{A} = \boldsymbol{S}$, i.e. $\boldsymbol{I}$ and $\boldsymbol{A}$ are as in Example 3.2.3 and $\boldsymbol{S}$ is the $(1, 1)$-bilabelled graph with a single vertex carrying a loop. Let G be a simple graph. In Example 3.2.9, it was argued that the homomorphism matrices $\boldsymbol{I}_G$ and $\boldsymbol{A}_G$ are the identity matrix and the adjacency matrix, respectively.

Since G is simple, its adjacency matrix has only zeros on the diagonal. Hence, the Schur product $\boldsymbol{I}_G \odot \boldsymbol{A}_G$ is the zero matrix. In accordance with Lemma 3.2.14, the zero matrix is equal to the homomorphism matrix $\boldsymbol{S}_G$. Indeed, as a looped graph, $\operatorname{soe}(\boldsymbol{S})$ does not admit any homomorphism to the simple graph G.

If loops had been removed after taking the parallel composition $\boldsymbol{I} \odot \boldsymbol{A}$, the resulting bilabelled graphs would have been $\boldsymbol{I}$. However, since $\boldsymbol{I}_G$ is the identity matrix, this would have violated the desired identity from Lemma 3.2.14.

3.3 Cycles, Paths, and Stars

In this section, it is illustrated how (bi)labelled graphs and their homomorphism tensors can be used to prove homomorphism indistinguishability characterisations. The graph classes considered are the class of cycles, paths, and stars. All results are folklore.

For a simple graph G and $i \in \mathbb{N}$, write $s_G(i) := \left|\{v \in V(G) \mid \deg_G(v) = i\}\right|$ for the number of vertices of G of degree i.

Theorem 3.3.1. *For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over stars,*
2. *G and H are homomorphism indistinguishable over stars on at most n vertices,*
3. *G and H have the same degree sequence, i.e. $s_G(i) = s_H(i)$ for all $0 \leq i \leq n-1$.*

Proof. For $d \geq 0$, write $\boldsymbol{S}^d$ for the 1-labelled star graph with d leaves and label at the central vertex. For every graph G with vertex $v \in V(G)$, it holds that $\boldsymbol{S}^d_G(v) = (\deg_G(v))^d$. The left hand-side is the number of homomorphisms from $\boldsymbol{S}^d$ to G which map the labelled vertex to v. For each leaf of $\boldsymbol{S}^d$ one may choose an image among the neighbours of v. Thus, there are $(\deg_G(v))^d$ many such homomorphisms.

For the implication from Item 3 to Item 1, let $S^d := \operatorname{soe}(\boldsymbol{S}^d)$ denote the unlabelled star with d leaves. Then, by Lemma 3.2.13,

$$\hom(S^d, G) = \operatorname{soe}(\boldsymbol{S}^d_G) = \sum_{v \in V(G)} (\deg_G(v))^d = \sum_{i=0}^{n-1} i^d s_G(i).$$

Thus, if G and H have the same degree sequence, then Item 1 holds. Item 1 trivially implies Item 2. For the remaining implication, one may view the above equation as a matrix-vector identity as follows: Letting d range from 0 to $n-1$, the left hand-side yields the vector $(\hom(S^d, G))_{0 \leq d \leq n-1} = (\hom(S^d, H))_{0 \leq d \leq n-1}$. The right hand-side yields the product of the matrix $(i^d)_{0 \leq i,d \leq n-1}$ with the vector $(s_G(i))_{0 \leq i \leq n-1}$. By Fact 2.4.6, the matrix, which is the Vandermonde matrix for the values $0 \leq i \leq n-1$, is invertible. Thus, $s_G(i) = s_H(i)$ for all $0 \leq i \leq n-1$, as desired. □

Figure 3.8: Two graphs which are homomorphism indistinguishable over all stars. Both graphs contain four vertices of degree one and two vertices of degree two.

See Figures 3.2 and 3.8 for examples of graphs which are homomorphism indistinguishable over all stars. In contrast, the graphs G and H depicted by Figure 3.1 are distinguished by homomorphism counts of stars, e.g. $\hom(K_{1,3}, G) = 6 \cdot 2^3 + 0^3 = 48 \neq 54 = 3 \cdot 1^3 + 3 \cdot 2^3 + 3^3 = \hom(K_{1,3}, H)$.

Although the statement of Theorem 3.3.1 does not appear linear-algebraic, the core argument in its proof is based on the Vandermonde determinant. In the case of the following Theorem 3.3.2, both statement and proof are linear-algebraic. Again, the statement is well-known, cf. e.g. [80, Theorem 4.3]. For background references on graph spectra, see [54]. See Figures 3.1 and 7.2 for example pairs of cospectral non-isomorphic graphs. Note that, in Theorem 3.3.2, the complete graphs K_1 and K_2 are also considered to be cycles, cf. Section 2.1.

Theorem 3.3.2. *For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over cycles,*
2. *G and H are homomorphism indistinguishable over cycles on at most n vertices,*
3. $\operatorname{tr}(\boldsymbol{A}_G^i) = \operatorname{tr}(\boldsymbol{A}_H^i)$ *for all* $0 \leq i \leq n$*,*
4. *$\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ have the same characteristic polynomial,*
5. *$\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ are* cospectral*, i.e. they have the same multiset*[4] *of eigenvalues.*

Proof. The first step is to rephrase Item 3. For $i \in \mathbb{N}$, the matrices $\boldsymbol{A}_G^i$ and $\boldsymbol{A}_H^i$ are the homomorphism matrices of the $(1,1)$-bilabelled path on $i+1$ vertices where one of the labels is positioned on each of the vertices of degree one, cf. Figure 3.5c. For $i = 0$, the path consists only of one vertex carrying both labels. Taking the traces of these matrices amounts to identifying the labelled vertices and dropping the labels, cf. Figure 3.6b. Thus, for $i \geq 3$, $\operatorname{tr}(\boldsymbol{A}_G^i)$ is the number of homomorphisms from the cycle on i vertices to G. Furthermore, $\operatorname{tr}(\boldsymbol{A}_G^0) = \operatorname{tr}(\boldsymbol{I}_G)$ counts the number of vertices in G, i.e. the number of cycles of length 0. Moreover, $\operatorname{tr}(\boldsymbol{A}_G^1)$ is zero since G is simple, cf. Figure 3.5d, while $\operatorname{tr}(\boldsymbol{A}_G^2)$ is twice the number of edges in G, i.e. the

[4] Note that a priori the geometric and algebraic multiplicities of eigenvalues may differ. The former are the dimensions of the eigenspaces while the latter are the multiplicities of the eigenvalues as roots of the characteristic polynomial. Since G and H are simple undirected graphs, their adjacency matrices $\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ are symmetric. By the Spectral Decomposition Theorem [105, VIII, §4, Theorem 4.3], the geometric and algebraic multiplicities of eigenvalues of symmetric matrices coincide.

number of homomorphisms from $C_2 = K_2$ to G. Thus, Items 2 and 3 are equivalent and implied by Item 1.

The remainder of the proof follows [58, Proposition 1]. By the Spectral Decomposition Theorem [105, VIII, §4, Theorem 4.3], the trace of the i-th power of a symmetric matrix equals the sum of the i-th powers of its eigenvalues. Thus, Item 5 implies that $\operatorname{tr}(\boldsymbol{A}_G^i) = \operatorname{tr}(\boldsymbol{A}_H^i)$ for all $i \in \mathbb{N}$, which yields Item 1 by the initial observations. Clearly, Items 4 and 5 are equivalent.

Thus, it remains to deduce Item 4 from Item 3. Let $\lambda_1, \dots, \lambda_m$ for $m := |V(G)|$ denote the eigenvalues of $\boldsymbol{A}_G$ with multiplicities. The characteristic polynomial of $\boldsymbol{A}_G$ is

$$\det(xI - \boldsymbol{A}_G) = \prod_{i=1}^{m}(x - \lambda_i) = \sum_{d=0}^{m}(-1)^{m+d} e_{m-d}(\lambda_1, \dots, \lambda_m) x^d$$

where $e_k(\lambda_1, \dots, \lambda_m) := \sum_{i \leq j_1 < \cdots < j_k \leq m} \lambda_{j_1}\lambda_{j_2} \cdots \lambda_{j_k}$. Here, $\det(xI - \boldsymbol{A}_G)$ denotes the determinant of the matrix $xI - \boldsymbol{A}_G$. By Newton's identities, cf. [178],

$$k e_k(\lambda_1, \dots, \lambda_m) = \sum_{i=1}^{k}(-1)^{i-1} e_{k-i}(\lambda_1, \dots, \lambda_m) s_i(\lambda_1, \dots, \lambda_m)$$

for all $1 \leq k \leq m$ and $s_k(\lambda_1, \dots, \lambda_m) := \sum_{i=1}^{m} \lambda_i^k = \operatorname{tr}(\boldsymbol{A}_G^k)$. Thus, the characteristic polynomial of $\boldsymbol{A}_G$ is determined by the values $\operatorname{tr}(\boldsymbol{A}_G^k)$ for $0 \leq k \leq m$. The roots of this polynomial are the eigenvalues of $\boldsymbol{A}_G$. Hence, Item 3 implies Item 4. □

The following characterisation of homomorphism indistinguishability over paths and cycles is implied by Theorem 3.3.2. Note that another characterisation for homomorphism indistinguishability over paths was derived by Dell, Grohe, & Rattan [61]. This characterisation is reproven in Corollary 4.2.3. Recall from Section 2.1 that $\overline{G}$ denotes the complement of the graph G. The spectrum of the adjacency matrix of the complement of a graph is well studied in spectral graph theory [54]. Equivalent characterisations for cospectrality of the adjacency matrices of a pair of graphs and their complements are listed in [53, Theorem 3].

Corollary 3.3.3. *For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over cycles and paths,*
2. *G and H are homomorphism indistinguishable over cycles and paths on at most n vertices,*
3. *$\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ are cospectral and $\boldsymbol{A}_{\overline{G}}$ and $\boldsymbol{A}_{\overline{H}}$ are cospectral.*

Proof. Clearly, Item 1 implies Item 2. Assuming Item 2, it follows from Theorem 3.3.2 that $\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ are cospectral. By Example 3.2.12, $\boldsymbol{A}_{\overline{G}} = \boldsymbol{J}_G - \boldsymbol{I}_G - \boldsymbol{A}_G$ where $\boldsymbol{I}$ and $\boldsymbol{J}$ are the $(1,1)$-bilabelled graphs depicted by Figure 3.7. In order to conclude that $\boldsymbol{A}_{\overline{G}}$ and $\boldsymbol{A}_{\overline{H}}$ are cospectral, it is argued that $\operatorname{tr}\left((\boldsymbol{J}_G - \boldsymbol{I}_G - \boldsymbol{A}_G)^\ell\right) =$

$\operatorname{tr}\left((\boldsymbol{J}_H - \boldsymbol{I}_H - \boldsymbol{A}_H)^\ell\right)$ for all $0 \leq \ell \leq n$. By expanding this expression using Definitions 3.2.6 and 3.2.7, it follows that $\operatorname{tr}\left((\boldsymbol{J} - \boldsymbol{I} - \boldsymbol{A})^\ell\right)$ is a linear combination of unlabelled graphs on at most $\ell \leq n$ vertices which are cycles or disjoint unions of paths. By Lemma 3.2.14, Corollary 3.2.15, and Equation (2.1), it follows that $\operatorname{tr}\left((\boldsymbol{J}_G - \boldsymbol{I}_G - \boldsymbol{A}_G)^\ell\right) = \operatorname{tr}\left((\boldsymbol{J}_H - \boldsymbol{I}_H - \boldsymbol{A}_H)^\ell\right)$ for all $0 \leq \ell \leq n$. This yields Item 3 via Theorem 3.3.2 applied to $\overline{G}$ and $\overline{H}$.

Finally, it is argued that Item 3 implies Item 1. To that end, observe that Theorem 3.3.2 yields that G and H are homomorphism indistinguishable over all cycles. Thus, it remains to show that G and H are homomorphism indistinguishable over all paths. Consider the following claim:

Claim 3.3.3a. For all $\ell \geq 1$, if $\operatorname{tr}\left((\boldsymbol{J}_G - \boldsymbol{I}_G - \boldsymbol{A}_G)^i\right) = \operatorname{tr}\left((\boldsymbol{J}_H - \boldsymbol{I}_H - \boldsymbol{A}_H)^i\right)$ for all $0 \leq i \leq \ell$, then G and H are homomorphism indistinguishable over all paths on at most ℓ vertices.

Proof of Claim. For $\ell = 1$, the claim follows since $\operatorname{tr}((\boldsymbol{J} - \boldsymbol{I} - \boldsymbol{A})^0) = \operatorname{tr}(\boldsymbol{I})$ is the one-vertex path. For $\ell \geq 2$, assume that the inductive hypothesis holds for $\ell - 1$ and suppose that $\operatorname{tr}\left((\boldsymbol{J}_G - \boldsymbol{I}_G - \boldsymbol{A}_G)^i\right) = \operatorname{tr}\left((\boldsymbol{J}_H - \boldsymbol{I}_H - \boldsymbol{A}_H)^i\right)$ for all $0 \leq i \leq \ell$. By the inductive hypothesis, G and H are homomorphism indistinguishable over all paths on at most $\ell - 1$ vertices. Thus, it remains to show that $\hom(P_\ell, G) = \hom(P_\ell, H)$ for the ℓ-vertex path P_ℓ.

Write $\Sigma_\ell := \{\boldsymbol{J}, \boldsymbol{I}, \boldsymbol{A}\}^\ell$ for the set of formal length-ℓ words over the alphabet comprised of the letters $\boldsymbol{J}$, $\boldsymbol{I}$, and $\boldsymbol{A}$. For $w \in \Sigma_\ell$, write $p(w) \in \mathbb{N}$ for the number of occurrences of $\boldsymbol{I}$ and $\boldsymbol{A}$ in w. Moreover, let $\boldsymbol{W}^w$ be the $(1,1)$-bilabelled graph obtained by taking the series composition of the $(1,1)$-bilabelled graphs $\boldsymbol{J}$, $\boldsymbol{I}$, and $\boldsymbol{A}$, as listed in w. Expanding the product and applying linearity of the trace, it follows that

$$\operatorname{tr}\left((\boldsymbol{J}_G - \boldsymbol{I}_G - \boldsymbol{A}_G)^\ell\right) = \sum_{w \in \Sigma_\ell} (-1)^{p(w)} \operatorname{tr}(\boldsymbol{W}^w_G) \tag{3.1}$$

and the same expression holds when G is replaced by H throughout. For each $w \in \Sigma_\ell$, $\operatorname{tr}(\boldsymbol{W}^w_G)$ is the number of homomorphisms from some cycle or disjoint union of paths to G as described by the following distinction of cases.

- If w does not contain $\boldsymbol{J}$, then $\operatorname{tr}(\boldsymbol{W}^w)$ is a cycle and it is $\operatorname{tr}(\boldsymbol{W}^w_G) = \operatorname{tr}(\boldsymbol{W}^w_H)$ by Theorem 3.3.2.
- If w contains $\boldsymbol{J}$ at least twice, then $\operatorname{tr}(\boldsymbol{W}^w)$ is a disjoint union of paths on at most $\ell - 1$ vertices. By the inductive hypothesis and Equation (2.1), $\operatorname{tr}(\boldsymbol{W}^w_G) = \operatorname{tr}(\boldsymbol{W}^w_H)$.
- If w contains $\boldsymbol{J}$ precisely once and $\boldsymbol{I}$ at least once, then $\operatorname{tr}(\boldsymbol{W}^w)$ is a path on at most $\ell - 1$ vertices. By the inductive hypothesis, $\operatorname{tr}(\boldsymbol{W}^w_G) = \operatorname{tr}(\boldsymbol{W}^w_H)$.

Write $\Gamma_\ell \subseteq \Sigma_\ell$ for the set of words which contain $\boldsymbol{J}$ precisely once and do not contain $\boldsymbol{I}$. The set Γ_ℓ comprises precisely the words in Σ_ℓ which are not covered by the case distinction above. Hence,

$$\sum_{w\in\Sigma_\ell\setminus\Gamma_\ell}(-1)^{p(w)}\operatorname{tr}(\boldsymbol{W}_G^w) = \sum_{w\in\Sigma_\ell\setminus\Gamma_\ell}(-1)^{p(w)}\operatorname{tr}(\boldsymbol{W}_H^w). \tag{3.2}$$

Note that, for every $w \in \Gamma_\ell$, the unlabelled graph $\operatorname{tr}(\boldsymbol{W}^w) \cong \operatorname{tr}(\boldsymbol{J}\boldsymbol{A}^{\ell-1})$ is the path on ℓ vertices. Moreover, $|\Gamma_\ell| = \ell$. By subtracting Equation (3.2) from Equation (3.1),

$$\hom(P_\ell, G) = \frac{1}{\ell}\sum_{w\in\Gamma_\ell}(-1)^{p(w)}\operatorname{tr}(\boldsymbol{W}_G^w) = \frac{1}{\ell}\sum_{w\in\Gamma_\ell}(-1)^{p(w)}\operatorname{tr}(\boldsymbol{W}_H^w) = \hom(P_\ell, H). \quad \lhd$$

If $\boldsymbol{A}_{\overline{G}}$ and $\boldsymbol{A}_{\overline{H}}$ are cospectral, then $\operatorname{tr}(\boldsymbol{A}_{\overline{G}}^\ell) = \operatorname{tr}(\boldsymbol{A}_{\overline{H}}^\ell)$, for all $\ell \in \mathbb{N}$, by the Spectral Decomposition Theorem [105, VIII, §4, Theorem 4.3]. Hence, Claim 3.3.3a implies that G and H are homomorphism indistinguishable over all paths. □

An example for a pair of graphs which are homomorphism indistinguishable over cycles and paths is given in [53, Figure 1] and depicted by Figure 3.1. It shows that homomorphism counts from cycles and paths do not determine whether a graph is connected. This answers a question raised by Dvořák [63, Section 5] negatively. In contrast, being disjoint unions of paths, the graphs depicted by Figure 3.8 are distinguished by homomorphism counts from paths [14, Theorem 3].

3.4 Trees, Bounded Treewidth, and Bounded Treedepth

Dell, Grohe, & Rattan [61] showed that two graphs are homomorphism indistinguishable over trees if, and only if, they are not distinguished by the Colour Refinement algorithm, cf. Definition 2.3.3. This result is given in Theorem 3.4.1 as a final introductory example.

Theorem 3.4.1 ([61, Theorem 1]). *For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of trees,*
2. *G and H are not distinguished by Colour Refinement.*

For example, the graphs depicted by Figure 3.2 are not distinguished by Colour Refinement. In contrast, Colour Refinement distinguishes the pairs of graphs in Figures 3.1 and 3.8.

Towards the proof of Theorem 3.4.1, let $\mathcal{T} \subseteq \mathcal{G}(1)$ denote the family of 1-labelled trees, i.e. the set of all $\boldsymbol{T} = (T, v)$ where T is a tree and $v \in V(T)$. The following Lemma 3.4.2 will imply Theorem 3.4.1.

Lemma 3.4.2. *For simple graphs G and H with $v \in V(G)$ and $w \in V(H)$, the following are equivalent:*

1. *$\boldsymbol{T}_G(v) = \boldsymbol{T}_H(w)$ for all $\boldsymbol{T} \in \mathcal{T}$,*
2. *$\operatorname{cr}^\infty(G, v) = \operatorname{cr}^\infty(H, w)$.*

Proof. First the backward direction is shown by induction on the depth of $\boldsymbol{T} \in \mathcal{T}$ where $\boldsymbol{T}$ is regarded as rooted tree with root at the labelled vertex. The only tree of depth one is $\mathbf{1} \in \mathcal{T}$, cf. Figure 3.3a. Clearly, $\mathbf{1}_G(v) = 1 = \mathbf{1}_H(w)$. Otherwise, there are two cases to consider:

If the labelled vertex of $\boldsymbol{T}$ is of degree one, then $\boldsymbol{T} = \boldsymbol{A} \cdot \boldsymbol{S}$ for some tree $\boldsymbol{S} \in \mathcal{T}$ of lesser depth. Then, by Lemma 3.2.14, $\boldsymbol{T}_G(v) = \sum_{v' \in V(G)} \boldsymbol{A}_G(vv')\boldsymbol{S}_G(v') = \sum_{v' \in N_G(v)} \boldsymbol{S}_G(v')$. By assumption, there exists a bijection $\pi \colon N_G(v) \to N_H(w)$ such that $\mathrm{cr}^\infty(G, v') = \mathrm{cr}^\infty(H, \pi(v'))$ for every $v' \in N_G(v)$. Thus, by the inductive hypothesis, $\boldsymbol{T}_G(v) = \sum_{v' \in N_G(v)} \boldsymbol{S}_G(v') = \sum_{v' \in N_G(v)} \boldsymbol{S}_H(\pi(v')) = \boldsymbol{T}_H(w)$.

If the labelled vertex of $\boldsymbol{T}$ has degree at least two, then $\boldsymbol{T} = \boldsymbol{T}^1 \odot \cdots \odot \boldsymbol{T}^\ell$ for some $\boldsymbol{T}^1, \dots, \boldsymbol{T}^\ell \in \mathcal{T}$ whose labelled vertices have degree one. The case above applies to these labelled trees. Thus, $\boldsymbol{T}_G(v) = \boldsymbol{T}^1_G(v) \cdots \boldsymbol{T}^\ell_G(v) = \boldsymbol{T}_H(w)$ by Lemma 3.2.13.

The converse direction is proven by induction on the iterations of Colour Refinement. By Definition 2.3.3, $\mathrm{cr}^0(G, v) = \mathrm{cr}^0(H, w)$. For the inductive step, recall that

$$\mathrm{cr}^{i+1}(G, v) = \left(\mathrm{cr}^i(G, v), \left\{\!\!\left\{\mathrm{cr}^i(G, v') \;\middle|\; v' \in N_G(v)\right\}\!\!\right\}\right).$$

Construct for every colour $\mathrm{cr}^i(G, v')$ a formal linear combinations of elements in $\mathcal{T}$ which counts the number of neighbours $v' \in N_G(v)$ of this colour. By the induction hypothesis, for every other colour $c := \mathrm{cr}^i(F, x) \neq \mathrm{cr}^i(G, v')$ where $F \in \{G, H\}$ and $x \in V(F)$, there exists $\boldsymbol{T}^c$ such that $n_1 := \boldsymbol{T}^c_F(x) \neq \boldsymbol{T}^c_G(v') =: n_2$. Consider the formal linear combination $\boldsymbol{t}^c := (n_2 - n_1)^{-1}(\boldsymbol{T}^c - n_1\mathbf{1})$. Then $\boldsymbol{t}^c_G(v') = 1$ and $\boldsymbol{t}^c_F(x) = 0$. Thus, the formal linear combination $\boldsymbol{q} := \bigodot_c \boldsymbol{t}^c$ where c ranges over all colours other than $\mathrm{cr}^i(G, v')$ is such that $\boldsymbol{q}_F(x) = 1$ if, and only if, $\mathrm{cr}^i(F, x) = \mathrm{cr}^i(G, v')$ and zero otherwise. Hence, $(\boldsymbol{A} \cdot \boldsymbol{q})_G(v)$ counts the number of neighbours of colour $\mathrm{cr}^i(G, v')$. By assumption, $(\boldsymbol{A} \cdot \boldsymbol{q})_G(v) = (\boldsymbol{A} \cdot \boldsymbol{q})_H(w)$ since all constituents of these linear combinations are homomorphism vectors of elements of $\mathcal{T}$. It follows that v and w have the same number of neighbours of colour $\mathrm{cr}^i(G, v')$. This implies that $\mathrm{cr}^{i+1}(G, v) = \mathrm{cr}^{i+1}(H, w)$. □

Theorem 3.4.1 can be deduced from Lemma 3.4.2 as follows.

Proof of Theorem 3.4.1. For the backward direction, observe that if $\mathrm{cr}^\infty(G) = \mathrm{cr}^\infty(H)$, then there exists a bijection $\pi \colon V(G) \to V(H)$ such that $\mathrm{cr}^\infty(G, v) = \mathrm{cr}^\infty(G, \pi(v))$ for all $v \in V(G)$. Let T be a tree with $u \in V(T)$ arbitrary and write $\boldsymbol{T} := (T, u) \in \mathcal{T}$. By Lemma 3.4.2, $\boldsymbol{T}_G(v) = \boldsymbol{T}_H(\pi(v))$ for all $v \in V(G)$. Thus,

$$\hom(T, G) = \mathrm{soe}(\boldsymbol{T}_G) = \sum_{v \in V(G)} \boldsymbol{T}_G(v) = \sum_{v \in V(G)} \boldsymbol{T}_H(\pi(v)) = \hom(T, H).$$

Conversely, one may apply the same interpolation argument as in the proof of Lemma 3.4.2. Indeed, for every colour c occurring among the $\mathrm{cr}^\infty(G, v)$ and $\mathrm{cr}^\infty(H, w)$ for $v \in V(G)$ and $w \in V(H)$, there exists a formal linear combination $\boldsymbol{q}^c$ of elements in $\mathcal{T}$ such that $\boldsymbol{q}^c_F(x) = 1$ if, and only if, $\mathrm{cr}^\infty(F, x) = c$ for

$F \in \{G, H\}$ and $x \in V(F)$ and zero otherwise. Since all constituents of $\boldsymbol{q}^c$ are in $\mathcal{T}$, it holds that $\text{soe}(\boldsymbol{q}^c_G) = \text{soe}(\boldsymbol{q}^c_H)$ and this quantity is equal to the number of vertices in G and H, respectively, of colour c. It follows that $\text{cr}^\infty(G) = \text{cr}^\infty(H)$. □

In [61], Theorem 3.4.1 was extended to all dimensions of the Weisfeiler–Leman algorithm. The proofs of Theorem 3.4.1 and the following Theorem 3.4.3 are conceptually similar despite that the latter theorem requires a heavier technical set-up akin to the one developed in Chapter 4.

Theorem 3.4.3 ([61, Theorem 3])**.** *Let $k \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the graphs of treewidth at most k,*
2. *G and H are not distinguished by the k-dimensional Weisfeiler–Leman algorithm.*

Via Theorems 2.3.2 and 2.6.8, the above Theorem 3.4.3 relates homomorphism indistinguishability over graphs of treewidth at most k to C^{k+1}-equivalence and to feasibility of the Sherali–Adams relaxation $\text{SA}^{k+1}(G, H)$. Historically, these connections were first established by the following Theorem 3.4.4 due to Dvořák [63], which predates Theorem 3.4.3. The proof of Theorem 3.4.4 reveals much about the connection between homomorphism indistinguishability and counting logic. Theorems 8.3.2 and 9.3.8 are proven using similar techniques.

Theorem 3.4.4 ([63, Theorem 7])**.** *Let $k \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the graphs of treewidth at most k,*
2. *G and H are C^{k+1}-equivalent.*

A similar result involving the C-fragment of formulas of bounded quantifier rank was shown by Grohe [79].

Theorem 3.4.5 ([79, Theorem 1.1])**.** *Let $k \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the graphs of treedepth at most k,*
2. *G and H are C_k-equivalent.*

The reader is referred to [57, 68], where Theorems 3.4.4 and 3.4.5 were reproven using uniform proof strategies based on category theory and the tools from Section 3.2, respectively. Both of these works also contain the following theorem, which was first proven by Dawar, Jakl, & Reggio [57]. Recall the definition of the class $\mathcal{T}^k_q$ of graphs admitting a k-pebble forest cover of depth at most q from Definition 2.1.3.

Theorem 3.4.6 ([57, Section IV.B])**.** *Let $k, q \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over $\mathcal{T}^k_q$,*
2. *G and H are C^k_q-equivalent.*

4 Matrix Equations from Homomorphism Indistinguishability

The observation that two graphs G and H are isomorphic if, and only if, there exists a permutation matrix $P \in \{0,1\}^{V(G)\times V(H)}$ such that $\boldsymbol{A}_G P = P\boldsymbol{A}_H$ motivates the study of equational graph isomorphism relaxations. An example for such a relaxation is *fractional isomorphism*, which is defined as the feasibility of the system $\boldsymbol{A}_G X = X\boldsymbol{A}_H$ where $X \in \mathbb{Q}^{V(G)\times V(H)}$ is a doubly stochastic matrix. By Theorems 3.4.1 and 2.6.2, two graphs are fractionally isomorphic if, and only if, they are homomorphism indistinguishable over all trees.

Characterisations of homomorphism indistinguishability relations in terms of matrix equations are the subject of this chapter, cf. Table 1.2 for an overview. Using (bi)labelled graphs and homomorphisms tensors, cf. Section 3.2, such characterisations are established for the classes of graphs of bounded treewidth, bounded pathwidth, bounded cyclewidth, and bounded treedepth. In this way, known equational graph isomorphism relaxations such as the Sherali–Adams linear programming hierarchy are recovered.

By Theorem 2.6.8, the level-k Sherali–Adams relaxation $\mathrm{SA}^k(G,H)$ of the graph isomorphism quadratic program $\mathrm{QP}(G,H)$ has a non-negative rational solution if, and only if, G and H are homomorphism indistinguishable over the class of graphs of treewidth at most $k-1$. Dell, Grohe, & Rattan [61] showed that if $\mathrm{SA}^k(G,H)$ has an arbitrary rational solution, then G and H are homomorphism indistinguishable over all graphs of pathwidth at most $k-1$. We show in Theorem 4.0.1 that the converse holds and thereby answer a question raised by Dell, Grohe, & Rattan [61] affirmatively. In the process, we define the linear programs $\mathrm{PW}^k(G,H)$ and $\mathrm{SA}^k_{\leftrightarrow}(G,H)$, which are motivated by homomorphism indistinguishability. Furthermore, we give an explanation for the rather miraculous appearance of non-negativity constraints when moving from homomorphism indistinguishability over graphs of bounded pathwidth to treewidth.

Theorem 4.0.1. *Let $k \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of graphs of pathwidth at most $k-1$,*
2. *$\mathrm{PW}^k(G,H)$ has a rational solution,*
3. *$\mathrm{SA}^k_{\leftrightarrow}(G,H)$ has a rational solution,*
4. *$\mathrm{SA}^k(G,H)$ has a rational solution.*

Moreover, we consider homomorphism indistinguishability over the class of graphs of bounded treedepth. We introduce the linear programming relaxations $\text{TD}^k(G,H)$ and $\text{TD}^k_{\leftrightarrow}(G,H)$ which can be regarded as variants of $\text{SA}^k(G,H)$ and $\text{SA}^k_{\leftrightarrow}(G,H)$ whose variables are indexed by ordered tuples instead of sets. In contrast to Theorems 4.0.1 and 2.6.8, adding non-negativity constraints does not affect the feasibility of these systems.

Theorem 4.0.2. *Let $k \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the graphs of treedepth at most k,*
2. *$\text{TD}^k(G,H)$ has a non-negative rational solution,*
3. *$\text{TD}^k(G,H)$ has a rational solution,*
4. *$\text{TD}^k_{\leftrightarrow}(G,H)$ has a non-negative rational solution,*
5. *$\text{TD}^k_{\leftrightarrow}(G,H)$ has a rational solution.*

The systems $\text{TD}^k_{\leftrightarrow}(G,H)$ and $\text{SA}^k_{\leftrightarrow}(G,H)$ are derived from the systems $\text{TD}^k(G,H)$ and $\text{SA}^k(G,H)$, respectively, by dropping equality constraints. Thereby, Theorems 4.0.1 and 4.0.2 can be thought of as equality elimination results paralleling a similar result [57, Theorem 32] for counting logic.

In order to construct a matrix equation characterising homomorphism indistinguishability over some graph class $\mathcal{F}$, we follow the strategy which is outlined subsequently. The starting point in each case is a class of unlabelled graphs $\mathcal{F}$ for which we desire a characterisation of the homomorphism indistinguishability relation $\equiv_{\mathcal{F}}$ in terms of a matrix equation.

First, we construct a family of labelled or bilabelled graphs $\mathcal{L}$ from $\mathcal{F}$. Usually, we want the underlying unlabelled graphs of the $\boldsymbol{L} \in \mathcal{L}$ to be exactly the graphs in $\mathcal{F}$. Note that, when placing labels on unlabelled graphs, there are inevitably several choices to consider. How to make this choice depends on the two subsequent steps.

Second, we define combinatorial operations on the family $\mathcal{L}$. These may include e.g. series and parallel composition, and gluing, cf. Definitions 3.2.5 and 3.2.6. The placement of labels is crucial. For example, when dealing with the class of graphs of treewidth at most $k-1$, we place for every graph in this class k labels in the same bag of a tree decomposition. In this way, the resulting class of k-labelled graphs is closed under gluing.

Third, we show that $\mathcal{L}$ is generated by finitely many generators under the operations defined above. This is desirable because we want the resulting matrix equation to have finitely many constraints.

Fourth, we derive a matrix equation from the combinatorial data described above by invoking representation-theoretic results. These results, stated in Theorems 4.1.2, 4.1.4, and 4.1.10, give criteria for simultaneous similarity of sequences of matrices in terms of character-like functions. The matrices to which these theorems are applied are the homomorphism matrices of the (bi)labelled graphs in $\mathcal{L}$.

Chapter Outline. In Section 4.1, we begin with laying the representation-theoretic foundations for the fourth step. In Section 4.2, we give basic examples for the proof strategy outlined above. In Section 4.3, we consider more complex instances such as the graph classes of bounded treewidth. In Section 4.4, we relate our matrix equations to the Sherali–Adams hierarchy and prove Theorems 4.0.1 and 4.0.2. In Section 4.5, we extend our combinatorial and algebraic techniques in order to design systems of equations capturing homomorphism indistinguishable over the class of graphs admitting pebble forest covers of bounded depth.

The material in this chapter has been previously published as [86, 87, 145, 146] and is joint work with Martin Grohe and Gaurav Rattan.

4.1 Three Variants of a Theorem by Specht and Wiegmann

In this section, two novel variants of a classical theorem by Specht [170] and Wiegmann [175] are derived. Two sequences of square matrices $A_1, \dots, A_n$ and $B_1, \dots, B_n$ are *simultaneously similar via a matrix X* if $A_i X = X B_i$ for all $i \in [n]$. Inspired by the classical Specht–Wiegmann Theorem, which establishes a criterion for simultaneous similarity via a unitary matrix, Theorems 4.1.4 and 4.1.10 give criteria for simultaneous similarity via a pseudo-stochastic or doubly stochastic matrix. All results give criteria for simultaneous similarity of sequences of matrices in terms of character-like functions.

To state the result formally, fix throughout a possibly infinite set of indices M. Let Γ_M denote the set of all finite words over the alphabet $\{x_i, x_i^* \mid i \in M\}$. The set Γ_M forms a monoid under concatenation with the empty word ε as unit element, i.e. concatenation is an associative binary operation. Furthermore, the monoid Γ_M can be endowed with an involution defined by extending $x_i \mapsto x_i^*$ to Γ_M in accordance with Equation (2.4). In this way, Γ_M can be thought of as a *free involution monoid*.

Let $\mathsf{A} = (A_i)_{i \in M}$ be a sequence of matrices in $\mathbb{K}^{I \times I}$ for some field $\mathbb{K}$ and some finite set I. For a word $w \in \Gamma_M$, let w_{A} denote the matrix obtained by substituting $x_i \mapsto A_i$ and $x_i^* \mapsto A_i^*$ for all $i \in M$ and evaluating the matrix product. Furthermore, ε_{A} is set to be the identity matrix in $\mathbb{K}^{I \times I}$. Crucially, the words in Γ_M are finite despite that the underlying alphabet is infinite. Hence, this map is well-defined. The substitution $w \mapsto w_{\mathsf{A}}$ is an involution monoid representation of Γ_M.

Example 4.1.1. Let $M = \{1, 2\}$ and $w := x_1 x_2^* x_2 \in \Gamma_M$. Then $w_{\mathsf{A}} = A_1 A_2^* A_2 \in \mathbb{K}^{I \times I}$.

4.1.1 Unitary and Orthogonal Similarity

We first recall the classical Specht–Wiegmann Theorem [170, 175]. It gives a criterion for simultaneous unitary or orthogonal similarity of two sequences of matrices in terms of traces. See also [96, 72] for more recent accounts.

Theorem 4.1.2 (Specht [170] and Wiegmann [175]). *Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Let I and J be finite sets. Let M be an arbitrary set. Let $\mathsf{A} = (A_i)_{i \in M}$ and $\mathsf{B} = (B_i)_{i \in M}$ be two sequences of matrices such that $A_i \in \mathbb{K}^{I \times I}$ and $B_i \in \mathbb{K}^{J \times J}$ for all $i \in M$. Then the following are equivalent:*

1. *there exists a unitary matrix $U \in \mathbb{K}^{I \times J}$ such that $A_i U = U B_i$ and $A_i^* U = U B_i^*$ for every $i \in M$,*
2. *there exists an invertible matrix $X \in \mathbb{K}^{I \times J}$ such that $A_i X = X B_i$ and $A_i^* X = X B_i^*$ for every $i \in M$,*
3. *for every word $w \in \Gamma_M$, $\operatorname{tr}(w_{\mathsf{A}}) = \operatorname{tr}(w_{\mathsf{B}})$.*

Note that if $\mathbb{K} = \mathbb{R}$, then the first assertion reduces to the existence of an orthogonal matrix $U \in \mathbb{R}^{I \times J}$ such that $A_i U = U B_i$ and $A_i^T U = U B_i^T$ for every $i \in M$. Following the conventions of Section 2.4.1, we do not need to distinguish this case explicitly. We will see in Remark 4.2.2 that Theorem 4.1.2 does not hold when $\mathbb{K}$ is taken to be $\mathbb{Q}$. If $\mathbb{K} = \mathbb{Q}$, then only Items 2 and 3 are equivalent.

Proof of Theorem 4.1.2. Since all unitary matrices are invertible, Item 1 implies Item 2.

That Items 2 and 3 are equivalent follows from Theorem 2.5.1 and Lemma 2.5.2. To that end, first consider the backward forward direction. If $X \in \mathbb{K}^{I \times J}$ is an invertible matrix such that $A_i X = X B_i$ and $A_i^* X = X B_i^*$ for every $i \in M$, then, for every $w \in \Gamma_M$, it holds that $\operatorname{tr}(w_{\mathsf{A}}) = \operatorname{tr}(w_{\mathsf{A}} X X^{-1}) = \operatorname{tr}(X w_{\mathsf{B}} X^{-1}) = \operatorname{tr}(X^{-1} X w_{\mathsf{B}}) = \operatorname{tr}(w_{\mathsf{B}})$ by cyclicity of the trace.

Conversely, as observed above, the maps $w \mapsto w_{\mathsf{A}}$ and $w \mapsto w_{\mathsf{B}}$ yield two representations of the involution monoid Γ_M. By Lemma 2.5.2, these representations are semisimple. If $\operatorname{tr}(w_{\mathsf{A}}) = \operatorname{tr}(w_{\mathsf{B}})$ for every word $w \in \Gamma_M$, these two representations have the same character, and hence, by Theorem 2.5.1, they are equivalent. Therefore, there exists an invertible matrix $X \in \mathbb{K}^{J \times I}$ such that $X w_{\mathsf{B}} X^{-1} = w_{\mathsf{A}}$ for every word $w \in \Gamma_M$. This yields Item 2.

For the remaining implication, we argue that the unitary matrix U can be obtained from the polar decomposition of the invertible matrix X. By [92, Theorem 7.3.1], write $X = UP$ where $U \in \mathbb{K}^{I \times J}$ is unitary and $P \in \mathbb{K}^{J \times J}$ is positive definite. It is $P^2 = X^* X$ and therefore $P^2 A_i = X^* B_i X = A_i X^* X = A_i P^2$ for all $i \in M$. Similarly, $P^2 A_i^* = A_i^* P^2$ for all $i \in M$. By [92, Theorem 7.2.6b], it holds that $P A_i = A_i P$ and$P A_i^* = A_i^* P$ for all $i \in M$. By multiplying from both sides with P^{-1}, it follows that $P^{-1} A_i = A_i P^{-1}$ and $P^{-1} A_i^* = A_i^* P^{-1}$ for all $i \in M$. By construction, $U = X P^{-1}$ and thus U is such that $U A_i = B_i U$ and $U A_i^* = B_i^* U$ for all $i \in M$. □

Words of Polynomial Length

The condition in Item 3 of Theorem 4.1.2 involves traces of words of arbitrary length. Since the considered matrices are of finite dimension, it suffices to compare traces from words in Γ_M of polynomial length. The following result is due to Pearcy [142]. Tighter bounds are known [140], a linear bound was conjectured by Paz [141].

Theorem 4.1.3 ([142, Theorem 1]). *Writing $n := \max\{|I|, |J|\}$, the conditions of Theorem 4.1.2 are equivalent to the following: For every word $w \in \Gamma_M$ of length at most $2n^2 - 1$, $\mathrm{tr}(w_\mathsf{A}) = \mathrm{tr}(w_\mathsf{B})$.*

Proof. To ease notation, suppose without loss of generality that I and J are disjoint. For a word $w \in \Gamma_M$, define the block matrix $w_{\mathsf{A}\oplus\mathsf{B}} := \begin{pmatrix} w_\mathsf{A} & 0 \\ 0 & w_\mathsf{B} \end{pmatrix} \in \mathbb{K}^{(I\cup J)\times(I\cup J)}$. Observe that $w \mapsto w_{\mathsf{A}\oplus\mathsf{B}}$ is an involution monoid representation. In particular, $(xy)_{\mathsf{A}\oplus\mathsf{B}} = x_{\mathsf{A}\oplus\mathsf{B}} y_{\mathsf{A}\oplus\mathsf{B}}$ and $(x^*)_{\mathsf{A}\oplus\mathsf{B}} = (x_{\mathsf{A}\oplus\mathsf{B}})^*$ for all $x, y \in \Gamma_M$.

Let $S_{\mathsf{A}\oplus\mathsf{B}} \leq \mathbb{K}^{(I\cup J)\times(I\cup J)}$ denote the vector space spanned by the $w_{\mathsf{A}\oplus\mathsf{B}}$ for $w \in \Gamma_M$. Clearly, $S_{\mathsf{A}\oplus\mathsf{B}}$ has dimension at most $2n^2$. Furthermore, for $\ell \geq 0$, write $S^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}} \leq S_{\mathsf{A}\oplus\mathsf{B}}$ for the vector space spanned by the $w_{\mathsf{A}\oplus\mathsf{B}}$ for words $w \in \Gamma_M$ of length at most ℓ. The space $S^{\leq 0}_{\mathsf{A}\oplus\mathsf{B}}$ containing the identity matrix, has dimension 1.

Claim 4.1.3a. If $S^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}} = S^{\leq \ell+1}_{\mathsf{A}\oplus\mathsf{B}}$ for some $\ell \in \mathbb{N}$, then $S_{\mathsf{A}\oplus\mathsf{B}} = S^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}}$. In particular, $S^{\leq 2n^2-1}_{\mathsf{A}\oplus\mathsf{B}} = S_{\mathsf{A}\oplus\mathsf{B}}$.

Proof of Claim. By induction on $j \geq 1$, it is shown that $S^{\leq \ell+j}_{\mathsf{A}\oplus\mathsf{B}} \leq S^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}}$. The base case $j = 1$ holds by assumption. Let $x \in \Gamma_M$ be a word of length $\ell + j + 1$. Let y denote the first letter of x and write z for the length-$(\ell + j)$ suffix of x, i.e. $x = yz$. By assumption, there exist words $z^1, \dots, z^r$ of length at most ℓ and coefficients $\alpha_1, \dots, \alpha_r \in \mathbb{K}$ such that $z_{\mathsf{A}\oplus\mathsf{B}} = \sum_{i=1}^r \alpha_i z^i_{\mathsf{A}\oplus\mathsf{B}}$. Hence,

$$x_{\mathsf{A}\oplus\mathsf{B}} = y_{\mathsf{A}\oplus\mathsf{B}} z_{\mathsf{A}\oplus\mathsf{B}} = \sum_{i=1}^r \alpha_i y_{\mathsf{A}\oplus\mathsf{B}} z^i_{\mathsf{A}\oplus\mathsf{B}} = \sum_{i=1}^r \alpha_i (yz^i)_{\mathsf{A}\oplus\mathsf{B}} \in S^{\leq \ell+1}_{\mathsf{A}\oplus\mathsf{B}}.$$

Thus, $S^{\leq \ell+j+1}_{\mathsf{A}\oplus\mathsf{B}} \leq S^{\leq \ell+1}_{\mathsf{A}\oplus\mathsf{B}} \leq S^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}}$, as desired. ◁

Equipped with Claim 4.1.3a, we prove the main claim. For a matrix $C \in \mathbb{K}^{(I\cup J)\times(I\cup J)}$, write $\mathrm{tr}_\mathsf{A}(C) := \sum_{i\in I} C(i,i)$ and analogously $\mathrm{tr}_\mathsf{B}(C) := \sum_{j\in J} C(j,j)$. Let $w \in \Gamma_M$ be an arbitrary word. By Claim 4.1.3a, there exist words $w^1, \dots, w^r \in \Gamma_M$ of length at most $2n^2 - 1$ and coefficients $\alpha_1, \dots, \alpha_r \in \mathbb{K}$ such that $w_{\mathsf{A}\oplus\mathsf{B}} = \sum_{i=1}^r \alpha_i w^i_{\mathsf{A}\oplus\mathsf{B}}$. Hence,

$$\begin{aligned}\mathrm{tr}(w_\mathsf{A}) &= \mathrm{tr}_\mathsf{A}(w_{\mathsf{A}\oplus\mathsf{B}}) = \sum \alpha_i \,\mathrm{tr}_\mathsf{A}(w^i_{\mathsf{A}\oplus\mathsf{B}}) = \sum \alpha_i \,\mathrm{tr}(w^i_\mathsf{A}) = \sum \alpha_i \,\mathrm{tr}(w^i_\mathsf{B}) \\ &= \mathrm{tr}_\mathsf{B}(w_{\mathsf{A}\oplus\mathsf{B}}) = \mathrm{tr}(w_\mathsf{B}).\end{aligned}$$

Thus, the assertion in Theorem 4.1.3 implies Item 3 of Theorem 4.1.2. □

4.1.2 Pseudo-Stochastic Similarity

Our first variant of Theorem 4.1.2 establishes a criterion for simultaneous similarity with respect to a pseudo-stochastic matrix. In this case, instead of traces, sums-of-entries have to be considered. Theorem 4.1.4 applies to any field $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$.

Theorem 4.1.4. *Let $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Let I and J be finite sets. Let M be an arbitrary set. Let $\mathsf{A} = (A_i)_{i \in M}$ and $\mathsf{B} = (B_i)_{i \in M}$ be two sequences of matrices such that $A_i \in \mathbb{K}^{I \times I}$ and $B_i \in \mathbb{K}^{J \times J}$ for all $i \in M$. Then the following are equivalent:*

1. *there exists a pseudo-stochastic matrix $X \in \mathbb{K}^{I \times J}$ such that $A_i X = X B_i$ and $A_i^* X = X B_i^*$ for all $i \in M$,*
2. *for every word $w \in \Gamma_M$, $\operatorname{soe}(w_\mathsf{A}) = \operatorname{soe}(w_\mathsf{B})$.*

Theorem 4.1.4 is implied by Lemma 4.1.5, which provides a sum-of-entries analogue of Theorem 2.5.1. As it establishes a character-theoretic interpretation of the function soe, it may be of independent interest. In the statement of Lemma 4.1.5, $\mathbf{1}_I \in \mathbb{K}^I$ denotes the vector all whose entries are one.

Lemma 4.1.5. *Let $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Let Γ be an involution monoid. Let I and J be finite sets. Let $\varphi \colon \Gamma \to \mathbb{K}^{I \times I}$ and $\psi \colon \Gamma \to \mathbb{K}^{J \times J}$ be representations of Γ. Let $\varphi' \colon \Gamma \to \operatorname{End}(V)$ and $\psi' \colon \Gamma \to \operatorname{End}(W)$ denote the subrepresentations of φ and of ψ generated by $\mathbf{1}_I$ and $\mathbf{1}_J$, respectively. Then the following are equivalent:*

1. *there exists a unitary pseudo-stochastic map $U \colon W \to V$ such that $\varphi'(g)U = U\psi'(g)$ for all $g \in \Gamma$,*
2. *there exists a pseudo-stochastic matrix $X \in \mathbb{K}^{I \times J}$ such that $\varphi(g)X = X\psi(g)$ for all $g \in \Gamma$,*
3. *for all $g \in \Gamma$, $\operatorname{soe}(\varphi(g)) = \operatorname{soe}(\psi(g))$.*

Proof. Suppose that Item 3 holds. The space V is spanned by the vectors $\varphi(g)\mathbf{1}_I$ for $g \in \Gamma$ while W is spanned by the $\psi(g)\mathbf{1}_J$ for $g \in \Gamma$. For $g, h \in \Gamma$, it holds that

$$\begin{aligned}\langle \varphi(g)\mathbf{1}_I, \varphi(h)\mathbf{1}_I \rangle &= \langle \mathbf{1}_I, \varphi(g)^* \varphi(h) \mathbf{1}_I \rangle = \langle \mathbf{1}_I, \varphi(g^* h) \mathbf{1}_I \rangle \\ &= \operatorname{soe}(\varphi(g^*h)) = \operatorname{soe}(\psi(g^*h)) = \langle \psi(g)\mathbf{1}_J, \psi(h)\mathbf{1}_J \rangle .\end{aligned}$$

Hence, V and W are spanned by vectors whose pairwise inner-products are respectively the same. Thus, by Lemma 2.4.3, there exists a unitary $U \colon W \to V$ such that $U\psi(g)\mathbf{1}_J = \varphi(g)\mathbf{1}_I$ for all $g \in \Gamma$. This implies that $U\psi'(g) = \varphi'(g)U$ for $g \in \Gamma$. Furthermore, $U\mathbf{1}_J = U\psi(1_\Gamma)\mathbf{1}_J = \varphi(1_\Gamma)\mathbf{1}_I = \mathbf{1}_I$ and $U^*\mathbf{1}_I = \mathbf{1}_J$ since U is unitary. Thus, Item 1 holds.

Suppose now that Item 1 holds. By Lemma 2.5.2, write $\varphi = \varphi' \oplus \varphi''$ and $\psi = \psi' \oplus \psi''$ where $\varphi'' \colon \Gamma \to \operatorname{End}(V^\perp)$ and $\psi'' \colon \Gamma \to \operatorname{End}(W^\perp)$ are involution monoid representations of Γ. By assumption, there exists a unitary map $U \colon W \to V$ such that $U\psi'(g) = \varphi'(g)U$ for all $g \in \Gamma$. Extend U to X by letting it annihilate $W^\perp$. Then $X\psi(g) = (U \oplus 0)(\psi' \oplus \psi'')(g) = U\psi'(g) \oplus 0 = \varphi'(g)U \oplus 0 = \varphi(g)X$ for all $g \in \Gamma$. Since U is pseudo-stochastic and $\mathbf{1}_I \in V$ and $\mathbf{1}_J \in W$, X is pseudo-stochastic as well. Hence, Item 2 holds.

For the remaining implication, let $X \in \mathbb{K}^{I \times J}$ be a pseudo-stochastic matrix such that $\varphi(g)X = X\psi(g)$ for all $g \in \Gamma$. Then, for every $g \in \Gamma$,

$$\operatorname{soe}(\varphi(g)) = \mathbf{1}_I^T \varphi(g) \mathbf{1}_I = \mathbf{1}_I^T \varphi(g) X \mathbf{1}_J = \mathbf{1}_I^T X \psi(g) \mathbf{1}_J = \mathbf{1}_J^T \psi(g) \mathbf{1}_J = \operatorname{soe}(\psi(g)).$$

Hence, Item 2 implies Item 3. □

Words of Polynomial Length

The following Theorem 4.1.6 parallels the polynomial bound from Theorem 4.1.3 on the length of the words which need to be inspected. Theorem 9.1.4 will imply that the bound in Theorem 4.1.6 is asymptotically tight up to a constant.

Theorem 4.1.6. *Writing $n := \max\{|I|, |J|\}$, the conditions of Theorem 4.1.4 are equivalent to the following: For every word $w \in \Gamma_M$ of length at most $2n-1$, $\mathrm{soe}(w_{\mathsf{A}}) = \mathrm{soe}(w_{\mathsf{B}})$.*

Proof. Suppose without loss of generality that I and J are disjoint. As in the proof of Theorem 4.1.3, write $w_{\mathsf{A}\oplus\mathsf{B}} := \begin{pmatrix} w_{\mathsf{A}} & 0 \\ 0 & w_{\mathsf{B}} \end{pmatrix} \in \mathbb{K}^{(I\cup J)\times(I\cup J)}$ for $w \in \Gamma_M$. Furthermore, write $\mathbf{1}$ for the all-ones vector in $\mathbb{K}^{I\cup J}$. Write $V_{\mathsf{A}\oplus\mathsf{B}} \leq \mathbb{K}^{I\cup J}$ for the space spanned by the vectors $w_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1}$ for all words $w \in \Gamma_M$. Write $V^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}} \leq V_{\mathsf{A}\oplus\mathsf{B}}$ for $\ell \geq 0$ for the subspace spanned by the vectors $w_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1}$ for words $w \in \Gamma_M$ of length $\leq \ell$. The space $V^{\leq 0}_{\mathsf{A}\oplus\mathsf{B}}$, containing $\mathbf{1}$, is one-dimensional. The space $V_{\mathsf{A}\oplus\mathsf{B}}$ is at most $2n$-dimensional.

Claim 4.1.6a. If $V^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}} = V^{\leq \ell+1}_{\mathsf{A}\oplus\mathsf{B}}$ for some $\ell \in \mathbb{N}$, then $V_{\mathsf{A}\oplus\mathsf{B}} = V^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}}$. In particular, $V^{\leq 2n-1}_{\mathsf{A}\oplus\mathsf{B}} = V_{\mathsf{A}\oplus\mathsf{B}}$.

Proof of Claim. By induction on $j \geq 1$, we show that $V^{\leq \ell+j}_{\mathsf{A}\oplus\mathsf{B}} \leq V^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}}$. The base case $j = 1$ holds by assumption. Let $x \in \Gamma_M$ be a word of length $\ell + j + 1$. Let y denote the first character of x and write z for the length-$(\ell + j)$ suffix of x, i.e. $x = yz$. By assumption, there exist words $z^1, \dots, z^r$ of length at most ℓ and coefficients $\alpha_1, \dots, \alpha_r \in \mathbb{K}$ such that $z_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1} = \sum_{i=1}^r \alpha_i z^i_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1}$. Hence, $x_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1} = y_{\mathsf{A}\oplus\mathsf{B}} z_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1} = \sum_{i=1}^r \alpha_i y_{\mathsf{A}\oplus\mathsf{B}} z^i_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1} = \sum_{i=1}^r \alpha_i (yz^i)_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1} \in V^{\leq \ell+1}_{\mathsf{A}\oplus\mathsf{B}}$. Thus, $V^{\leq \ell+j+1}_{\mathsf{A}\oplus\mathsf{B}} \leq V^{\leq \ell+1}_{\mathsf{A}\oplus\mathsf{B}} \leq V^{\leq \ell}_{\mathsf{A}\oplus\mathsf{B}}$, as desired. ◁

Let $w \in \Gamma_M$ be arbitrary. By Claim 4.1.6a, there exist words $w^i \in \Gamma_M$ of length at most $2n-1$ and coefficients $\alpha_i \in \mathbb{K}$ such that $w_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1} = \sum \alpha_i w^i_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1}$. Hence, writing $\mathbf{1}_A$ and $\mathbf{1}_B$ for the indicator vectors on I and J in $\mathbb{K}^{I\cup J}$,

$$\mathrm{soe}(w_{\mathsf{A}}) = \mathbf{1}_{\mathsf{A}}^T w_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1} = \sum \alpha_i \mathbf{1}_{\mathsf{A}}^T w^i_{\mathsf{A}\oplus\mathsf{B}}\mathbf{1} = \sum \alpha_i \,\mathrm{soe}(w^i_{\mathsf{A}}) = \sum \alpha_i \,\mathrm{soe}(w^i_{\mathsf{B}}) = \mathrm{soe}(w_{\mathsf{B}}),$$

as desired. □

4.1.3 Doubly Stochastic Similarity

Our second variant of the Specht–Wiegmann Theorem gives a criterion for simultaneous doubly stochastic similarity. Recall that a real matrix is doubly stochastic if it is pseudo-stochastic and has non-negative entries. Since double stochasticity makes no sense over complex numbers, we restrict our attention to representations of involution monoids over $\mathbb{R}$ and $\mathbb{Q}$. In contrast to Theorems 4.1.2 and 4.1.4, the criterion derived in this section does not involve words over some set of matrices but trees, defined as follows:

Definition 4.1.7 (Trees over a monoid)**.** Let Γ be a monoid. A *tree over* Γ is a tuple $t = (T, r, e)$ where T is a finite tree, $r \in V(T)$, and $e\colon E(T) \to \Gamma$ is a map which assigns an element of Γ to every edge of T.

Two trees $t = (T, r, e)$ and $t' = (T', r', e')$ over Γ are *isomorphic* if there is a graph isomorphism $h\colon T \to T'$ such that $h(r) = r'$ and $e'(h(u)h(v))) = e(uv)$ for all $uv \in E(T)$. We tacitly identify isomorphic trees over Γ and write $\mathcal{T}(\Gamma)$ for the set of (isomorphism types) of trees over Γ. We consider the following operations on $\mathcal{T}(\Gamma)$.

Definition 4.1.8. Let $t = (T, r, e)$ and $t' = (T', r', e')$ be elements of $\mathcal{T}(\Gamma)$. Let $g \in \Gamma$.

1. Define $t \odot t' := (T'', r'', e'') \in \mathcal{T}(\Gamma)$ where T'' is the tree obtained from taking the disjoint union of T and T' and identifying the vertices r and r'. Furthermore, r'' is the identified vertex and $e''\colon E(T'') \to \Gamma$ is such that $e''|_{E(T)} = e$ and $e''|_{E(T')} = e'$.
2. Define $gt := (T'', r'', e'') \in \mathcal{T}(\Gamma)$ where $V(T'') := V(T) \sqcup \{r''\}$, $E(T'') := E(T) \sqcup \{rr''\}$, and $e''\colon E(T'') \to \Gamma$ is such that $e''|_{E(T)} = e$ and $e''(rr'') = g$.

The elements of $\mathcal{T}(\Gamma)$ can be constructed from the tree over Γ with only one vertex by the operations in Definition 4.1.8, i.e. by gluing and by attaching a new root s to a tree $t = (T, r, e)$ and picking an element $g \in \Gamma$ to associate with the new edge sr.

The set $\mathcal{T}(\Gamma)$ forms a monoid under the operation $\odot$ of gluing two of its elements together at their roots. Since isomorphic trees are identified in $\mathcal{T}(\Gamma)$, the monoid is commutative. Its unique neutral element is the one-vertex tree. Endowing trees over monoids with a monoid structure is not a novel idea, cf. e.g. [30].

A representation of a monoid Γ induces a representation[5] of the monoid $\mathcal{T}(\Gamma)$.

Definition 4.1.9. Let $\mathbb{K}$ be a field. Let Γ be a monoid and let I be a finite set. A monoid representation $\varphi\colon \Gamma \to \mathbb{K}^{I \times I}$ of Γ induces a monoid representation $\widehat{\varphi}\colon \mathcal{T}(\Gamma) \to \mathbb{K}^{I}$ of $\mathcal{T}(\Gamma)$ defined inductively as follows:

1. $\widehat{\varphi}(t) := \boldsymbol{1}_I$ if t has only one vertex,
2. $\widehat{\varphi}(t) := \widehat{\varphi}(t') \odot \widehat{\varphi}(t'')$ if $t := t' \odot t''$ for $t', t'' \in \mathcal{T}(\Gamma)$ on more than one vertex,
3. $\widehat{\varphi}(t) := \varphi(g) \cdot \widehat{\varphi}(t')$ if $t = gt'$ for $t' \in \mathcal{T}(\Gamma)$ and $g \in \Gamma$.

For the involution monoid Γ_M and a sequence of matrices $\mathsf{A} = (A_i)_{i \in M}$, the representation of $\mathcal{T}(\Gamma_M)$ induced by $w \mapsto w_{\mathsf{A}}$ is abbreviated as $t \mapsto t_{\mathsf{A}}$. See Figure 4.1 for an example. The main result of this section is the following.

[5] In Definition 4.1.9, $\mathbb{K}^I$ is understood as the monoid whose binary operation is the Schur product $\odot$ and whose neutral element is the all-ones vector $\boldsymbol{1}_I$. Crucially, $\mathbb{K}^I$ is commutative. The map $\widehat{\varphi}\colon \mathcal{T}(\Gamma) \to \mathbb{K}^I$ defined in Definition 4.1.9 is strictly speaking only a monoid homomorphism and not a monoid representation as $\mathbb{K}^I$ is not an endomorphism monoid. However, $\mathbb{K}^I$ can be identified with the set of diagonal matrices in $\mathbb{K}^{I \times I}$ in which case the Schur product in $\mathbb{K}^I$ corresponds to the standard matrix product in $\mathbb{K}^{I \times I}$. Thus, calling $\widehat{\varphi}$ a monoid representation is only slightly abusive. It allows us to make a distinction between matrices and vectors, which will in Section 4.3 amount to a distinction between bilabelled and labelled graphs.

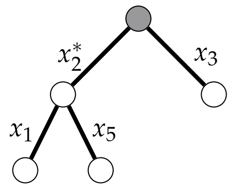


Figure 4.1: Example for a tree $t \in \Gamma_M$ for $M = [5]$. The root is depicted in grey. Here, $t_{\mathsf{A}} = (A_2^*((A_1\mathbf{1}) \odot (A_5\mathbf{1}))) \odot (A_3\mathbf{1})$.

Theorem 4.1.10. *Let $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}\}$. Let I and J be finite sets. Let M be an arbitrary set. Let $\mathsf{A} = (A_i)_{i \in M}$ and $\mathsf{B} = (B_i)_{i \in M}$ be two sequences of matrices such that $A_i \in \mathbb{K}^{I \times I}$ and $B_i \in \mathbb{K}^{J \times J}$ for all $i \in M$. Then the following are equivalent:*

1. *there exists a doubly stochastic matrix $X \in \mathbb{K}^{I \times J}$ such that $A_i X = X B_i$ and $A_i^* X = X B_i^*$ for all $i \in M$.*
2. *for every $t \in T(\Gamma_M)$, $\operatorname{soe}(t_{\mathsf{A}}) = \operatorname{soe}(t_{\mathsf{B}})$.*

Theorem 4.1.10 is implied by the following Lemmas 4.1.11 and 4.1.12.

Lemma 4.1.11. *Let $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}\}$. Let Γ be a monoid. Let I and J be finite sets. Let $\varphi \colon \Gamma \to \mathbb{K}^I$ and $\psi \colon \Gamma \to \mathbb{K}^J$ be representations of Γ where $\mathbb{K}^I$ and $\mathbb{K}^J$ are regarded as monoids whose binary operation is the Schur product and whose unit element is the all-ones vector. Then the following are equivalent:*

1. *there exists a pseudo-stochastic matrix $X \in \mathbb{K}^{I \times J}$ such that $X\psi(g) = \varphi(g)$ for all $g \in \Gamma$,*
2. *there exists a doubly stochastic matrix $X \in \mathbb{K}^{I \times J}$ such that $X\psi(g) = \varphi(g)$ for all $g \in \Gamma$,*
3. *there exists a bijection $\pi \colon J \to I$ such that $\psi(g)_j = \varphi(g)_{\pi(j)}$ for all $j \in J$ and $g \in \Gamma$,*
4. *for all $g \in \Gamma$, $\operatorname{soe}(\varphi(g)) = \operatorname{soe}(\psi(g))$.*

Proof. Each of Items 1 to 3 implies Item 4. Indeed, assuming Item 1, for all $g \in \Gamma$,

$$\operatorname{soe}(\psi(g)) = \mathbf{1}_J^T \psi(g) = \mathbf{1}_I^T X \psi(g) = \mathbf{1}_I^T \varphi(g) = \operatorname{soe}(\varphi(g)). \tag{4.1}$$

Since every doubly stochastic matrix is pseudo-stochastic, Item 2 implies Item 1. Assuming Item 3, for all $g \in \Gamma$,

$$\operatorname{soe}(\psi(g)) = \sum_{j \in J} \psi(g)_j = \sum_{j \in J} \varphi(g)_{\pi(j)} = \sum_{i \in I} \varphi(g)_i = \operatorname{soe}(\varphi(g)).$$

Assuming Item 4, we separately prove each of the converse implications. To that end, write $V \leq \mathbb{K}^I$ and $W \leq \mathbb{K}^J$ for the spaces spanned by the $\varphi(g)$ and $\psi(g)$, respectively, for $g \in \Gamma$. Suppose that $\operatorname{soe}(\varphi(g)) = \operatorname{soe}(\psi(g))$ for all $g \in \Gamma$. Then, for all $g, g' \in \Gamma$,

$$\langle \varphi(g), \varphi(g') \rangle = \operatorname{soe}(\varphi(g) \odot \varphi(g')) = \operatorname{soe}(\varphi(gg')) = \operatorname{soe}(\psi(gg')) = \langle \psi(g), \psi(g') \rangle .$$

Hence, by Lemma 2.4.3, there exists a unitary map $U\colon W \to V$ such that $U\psi(g) = \varphi(g)$ for all $g \in \Gamma$. Since U is unitary, also $\psi(g) = U^*\varphi(g)$ for all $g \in \Gamma$. Extend U to map $X\colon \mathbb{K}^J \to \mathbb{K}^I$ annihilating $W^\perp$. Then X is a pseudo-stochastic map satisfying $X\psi(g) = \varphi(g)$ and $\psi(g) = X^*\varphi(g)$ for all $g \in \Gamma$. Hence, Item 1 holds.

For Item 2, it remains to argue that X is doubly stochastic. To that end, consider the equivalence relation $\sim_I$ on I induced by V via $i \sim i'$ if, and only if, $v(i) = v(i')$ for all $v \in V$. Define an equivalence relation $\sim_J$ on J analogously.

Claim 4.1.11a. Let $X \in \mathbb{K}^{I\times J}$ be a matrix such that $X\psi(g) = \varphi(g)$ for all $g \in \Gamma$. Then X maps indicator vectors of $\sim_J$-classes to vectors whose entries are from $\{0, 1\}$.

Proof of Claim. Let $D \subseteq J$ be a $\sim_J$-class and write $\mathbf{1}_D \in \mathbb{K}^J$ for its indicator vector. Write $\mathbf{1}_D = \sum_r \alpha_r \psi(g_r)$ for some $g_r \in \Gamma$ and coefficients $\alpha_r \in \mathbb{K}$. Then

$$
\begin{aligned}
(X\mathbf{1}_D) \odot (X\mathbf{1}_D) &= \sum_{r,s} \alpha_r\alpha_s (X\psi(g_r)) \odot (X\psi(g_s)) \\
&= \sum \alpha_r\alpha_s \varphi(g_r) \odot \varphi(g_s) \\
&= \sum \alpha_r\alpha_s \varphi(g_r \odot g_s) \\
&= \sum \alpha_r\alpha_s X(\psi(g_r \odot g_s)) \\
&= \sum \alpha_r\alpha_s X(\psi(g_r) \odot \psi(g_s)) \\
&= X(\mathbf{1}_D \odot \mathbf{1}_D) \\
&= X\mathbf{1}_D
\end{aligned}
$$

Hence, $X\mathbf{1}_D$ has entries from $\{0, 1\}$. $\lhd$

Given Claim 4.1.11a, it remains to argue that the entries of X are non-negative. To that end, let $j \in J$ and let $D \subseteq J$ denote its equivalence class under $\sim_J$.

Write $P \in \mathbb{K}^{J\times J}$ for the projection onto W. By construction, W is closed under Schur products. Thus, by Lemma 2.4.7, the indicator vectors of the $\sim_J$-classes are an orthogonal basis of W. Hence, $Pe_j = |D|^{-1}\mathbf{1}_D$ where $e_j \in \mathbb{K}^J$ denotes the standard basis vector corresponding to j. By Claim 4.1.11a,

$$
X(i, j) = e_i^T X e_j = e_i^T X P e_j = |D|^{-1} e_i^T X \mathbf{1}_D
$$

is non-negative for every $i \in I$. Hence, X is doubly stochastic and Item 2 holds.

For Item 3, consider the following Claim 4.1.11b:

Claim 4.1.11b. Let $X \in \mathbb{K}^{I\times J}$ be a matrix such that $X\psi(g) = \varphi(g)$ and $\psi(g) = X^*\varphi(g)$ for all $g \in \Gamma$. Then X maps indicator vectors of $\sim_J$-classes to indicator vectors of $\sim_I$-classes.

Proof of Claim. First observe that X descends to an isomorphism $W \to V$. Indeed, the assumptions imply that the maps XX^* and X^*X restrict to the identity maps on V and W, respectively.

Let $D \subseteq J$ denote a $\sim_J$-class. By Claim 4.1.11a, $X\mathbf{1}_D$ is a vectors whose entries are from $\{0, 1\}$. Furthermore, $X\mathbf{1}_D \neq 0$ since X descends to an isomorphism from W to V.

Therefore, $X\mathbf{1}_D = \mathbf{1}_{C_1} + \cdots + \mathbf{1}_{C_\ell}$ for some $\sim_I$-classes $C_1, \ldots, C_\ell \subseteq I$ and $\ell \geq 1$. Dually, $X^*\mathbf{1}_{C_i} = \mathbf{1}_{D_{i1}} + \ldots \mathbf{1}_{D_{ik_i}}$ for some $\sim_J$-classes $D_{i1}, \ldots, D_{ik_i} \subseteq J$ with $k_i \geq 1$. Hence,

$$\mathbf{1}_D = X^*X\mathbf{1}_D = X^*\mathbf{1}_{C_1} + \cdots + X^*\mathbf{1}_{C_\ell} = \mathbf{1}_{D_{i1}} + \cdots + \mathbf{1}_{D_{\ell k_\ell}}.$$

This implies that $\ell = k_i = 1$ for all i and thereby the claim. ◁

By Claim 4.1.11b, X maps $\sim_J$-class vectors to $\sim_I$-class vectors. Furthermore, $\mathbf{1}^T X\mathbf{1}_D = \mathbf{1}^T\mathbf{1}_D = |D|$ since X is pseudo-stochastic. Thus corresponding classes have the same size. Finally, since X descends to an isomorphism $W \to V$, the map X induces a bijection σ between the $\sim_J$-classes and $\sim_J$-classes. Hence, there exists a bijection $\pi\colon J \to I$ such that $X(\pi(j), j) > 0$ for all $j \in J$. Now, Lemma 2.4.8 implies Item 3. □

Lemma 4.1.11 is applied in the proof of Lemma 4.1.12, which yields Theorem 4.1.10.

Lemma 4.1.12. *Let $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}\}$. Let Γ be an involution monoid. Let I and J be finite sets. Let $\varphi\colon \Gamma \to \mathbb{K}^{I\times I}$ and $\psi\colon \Gamma \to \mathbb{K}^{J\times J}$ be representations of Γ. Let $\widehat{\varphi}$ and $\widehat{\psi}$ be the induced representations of $T(\Gamma)$. Then the following are equivalent:*

1. *there exists a doubly stochastic matrix $X \in \mathbb{K}^{I\times J}$ such that $\varphi(g)X = X\psi(g)$ for all $g \in \Gamma$,*
2. *there exists a pseudo-stochastic matrix $X \in \mathbb{K}^{I\times J}$ such that $\varphi(g)X = X\psi(g)$ for all $g \in \Gamma$ and $X\widehat{\psi}(t) = \widehat{\varphi}(t)$ for all $t \in T(\Gamma)$,*
3. *for all $t \in T(\Gamma)$, $\operatorname{soe}(\widehat{\varphi}(t)) = \operatorname{soe}(\widehat{\psi}(t))$.*

Proof. Item 2 readily implies Item 3, cf. Equation (4.1). Item 3 implies Item 1: Apply Lemma 4.1.11 to the monoid $T(\Gamma)$ and its representations $\widehat{\varphi}$ and $\widehat{\psi}$. The lemma yields a doubly stochastic matrix $X \in \mathbb{K}^{I\times J}$ such that $X\widehat{\psi}(t) = \widehat{\varphi}(t)$ for all $t \in T(\Gamma)$. Let $P \in \mathbb{K}^{J\times J}$ denote the projection onto the space $W \leq \mathbb{K}^J$ spanned by the $\widehat{\psi}(t)$ for $t \in T(\Gamma)$. Then $P\widehat{\psi}(t) = \widehat{\psi}(t)$ for all $t \in T(\Gamma)$ and thus XP satisfies the same identities as X, i.e. $XP\widehat{\psi}(t) = \widehat{\varphi}(t)$ for all $t \in T(\Gamma)$.

It holds that $P\psi(g) = \psi(g)P$ for all $g \in \Gamma$. This follows from the following observation: if $w \in W^\perp$, then $\langle \widehat{\psi}(t), \psi(g)w \rangle = \langle \widehat{\psi}(g^*t), w \rangle = 0$ for all $t \in T(\Gamma)$ and $g \in \Gamma$.

Hence, it suffices to verify the identity $\varphi(g)XP = XP\psi(g)$ for $g \in \Gamma$ on W. Let $t \in T(\Gamma)$. Then $XP\psi(g)\widehat{\psi}(t) = X\widehat{\psi}(gt) = \widehat{\varphi}(gt) = \varphi(g)\widehat{\varphi}(t) = \varphi(g)XP\widehat{\varphi}(t)$.

By Lemma 2.4.7, P is doubly stochastic. Thus XP is doubly stochastic, being the product of two doubly stochastic matrices. Hence, XP is as desired.

Item 1 implies Item 2: Let X be as in Item 1. It has to be shown that $X\widehat{\psi}(t) = \widehat{\varphi}(t)$ for all $t \in T(\Gamma)$. The proof of the following slightly stronger claim is guided by [172, Lemma 1].

Claim 4.1.12a. Let $t \in T(\Gamma)$, $i \in I$, and $j \in J$. If $X(i,j) > 0$, then $\widehat{\psi}(t)(i) = \widehat{\varphi}(t)(j)$.

Claim 4.1.12a implies that $X\widehat{\psi}(t) = \widehat{\varphi}(t)$ for all $t \in T(\Gamma)$. Indeed, for $t \in T(\Gamma)$ and $i \in I$,

$$(X\widehat{\psi}(t))(i) = \sum_{\substack{j \in J, \\ X(i,j)>0}} X(i,j)\widehat{\psi}(t)(j) = \sum_{\substack{j \in J, \\ X(i,j)>0}} X(i,j)\widehat{\varphi}(t)(i) = \widehat{\psi}(t)(i)(X\mathbf{1}_J)(i)$$
$$= \widehat{\psi}(t)(i). \tag{4.2}$$

Recalling the convention that $X^* = X^T$ for real-valued matrices, one may similarly see that $X^*\widehat{\varphi}(t) = \widehat{\psi}(t)$ for all $t \in T(\Gamma)$. This consequence is applied in the inductive proof of Claim 4.1.12a.

Proof of Claim 4.1.12a. The proof is by induction on the structure of the elements of $T(\Gamma)$, cf. Definition 4.1.9. For the single-vertex tree, the claim is vacuous.

For the inductive step, two means of constructing more complex elements $t = (T, r, e) \in T(\Gamma)$ are considered. If $t = t' \odot t''$ for two non-trivial $t', t'' \in T(\Gamma)$, the claim is readily verified. It remains to consider the case in which r has a unique child s in T. Write $t' = (T - r, s, e|_{E(T-r)}) \in T(\Gamma)$ for the subtree rooted at s. Let $g := e(rs) \in \Gamma$. That is, $t = gt'$.

The vectors $\widehat{\psi}(t)$ and $\widehat{\varphi}(t)$ satisfy the assumptions of Lemma 2.4.8. Indeed, by Item 1 and Equation (4.2),

$$X\widehat{\psi}(t) = X\psi(g)\widehat{\psi}(t') = \varphi(g)X\widehat{\psi}(t') = \varphi(g)\widehat{\varphi}(t') = \widehat{\varphi}(t), \tag{4.3}$$

and, alluding to the assumption that Γ is an involution monoid,

$$X^*\widehat{\varphi}(t) = X^*\varphi(g)\widehat{\varphi}(t') = (\varphi(g^*)X)^*\widehat{\varphi}(t') = (X\psi(g^*))^*\widehat{\varphi}(t')$$
$$= \psi(g)X^*\widehat{\varphi}(t') = \psi(g)\widehat{\psi}(t') = \widehat{\psi}(t). \tag{4.4}$$

Equations (4.3) and (4.4) in conjunction with Lemma 2.4.8 imply the claim. ◁

As argued above, Claim 4.1.12a implies Item 2. □

Trees of Polynomial Depth and Degree

In contrast to Theorems 4.1.3 and 4.1.6, an exponential bound on the size of the trees which need to be a considered in Theorem 4.1.10 is proven. Lower bounds are discussed in Remark 4.2.5. First the depth of the trees is bounded. The *depth* of a tree $t = (T, r, e) \in T(\Gamma_M)$ is defined as the depth of the rooted tree (T, r), i.e. as the maximal number of vertices of any path on T starting in r. For example, the tree in Figure 4.1 is of depth three.

Lemma 4.1.13. *Writing $n := \max\{|I|, |J|\}$, the conditions of Theorem 4.1.10 are equivalent to the following: For every tree $t \in T(\Gamma_M)$ of depth at most $n+1$, $\mathrm{soe}(t_\mathsf{A}) = \mathrm{soe}(t_\mathsf{B})$.*

Proof. The following argument is inspired by the proof of [142, Theorem 1]. For $d \geq 1$, write $T_{\mathsf{A}}^{\leq d} \leq \mathbb{K}^I$ for the vector space spanned by the t_{A} for all $t \in T(\Gamma_M)$ of depth $\leq d$. Write T_{A} for the space spanned by the vectors t_{A} for all trees $t \in T(\Gamma_M)$. Clearly, T_{A} is at most n-dimensional. The space $T_{\mathsf{A}}^{\leq 0}$, containing the all-ones vector, is one-dimensional.

Claim 4.1.13a. If $T_{\mathsf{A}}^{\leq d} = T_{\mathsf{A}}^{\leq d+1}$ for some $d \in \mathbb{N}$, then $T_{\mathsf{A}} = T_{\mathsf{A}}^{\leq d}$. In particular, $T_{\mathsf{A}}^{\leq n} = T_{\mathsf{A}}$.

Proof of Claim. By induction on $j \geq 1$, it is shown that $T_{\mathsf{A}}^{\leq d+j} \leq T_{\mathsf{A}}^{\leq d}$. The base case $j = 1$ holds by assumption. Let $t \in T(\Gamma_M)$ be a tree of depth $d + j + 1$. If t can be written as $t = t^1 \odot \cdots \odot t^\ell$ for some trees $t^1, \dots, t^\ell \in T(\Gamma_M)$ whose roots have degree one, then it suffices to show that $t^i_{\mathsf{A}} \in T_{\mathsf{A}}^{\leq d}$ for all $i \in [\ell]$ since $T_{\mathsf{A}}^{\leq d}$ is closed under Schur products.

Hence, it may be supposed that the root r of t has a single child s. Write t' for the subtree of t rooted at s and $g \in \Gamma_M$ for the element associated with the edge rs. By assumption, there exist trees $x^1, \dots, x^m \in T(\Gamma_M)$ of depth at most d and coefficients $\alpha_1, \dots, \alpha_r \in \mathbb{K}$ such that $t'_{\mathsf{A}} = \sum_{i=1}^m \alpha_i x^i_{\mathsf{A}}$. Then $t_{\mathsf{A}} = \sum_{i=1}^m \alpha_i g_{\mathsf{A}} x^i_{\mathsf{A}} \in T_{\mathsf{A}}^{\leq d+1}$. Thus, $T_{\mathsf{A}}^{\leq d+j+1} \leq T_{\mathsf{A}}^{\leq d+1} \leq T_{\mathsf{A}}^{\leq d}$, as desired. ◁

By Claim 4.1.13a, every $t_{\mathsf{A}} \in T_{\mathsf{A}}$ can be written as a linear combination of some $t'_{\mathsf{A}} \in T_{\mathsf{A}}^{\leq n}$. It remains to show that the coefficients in this linear combination are the same for A and B.

Claim 4.1.13b. For every $t \in T(\Gamma_M)$, there exist coefficients $\alpha_1, \dots, \alpha_m \in \mathbb{K}$ and trees $t^1, \dots, t^m \in T(\Gamma_M)$ of depth at most n such that

$$t_{\mathsf{A}} = \sum_i \alpha_i t^i_{\mathsf{A}} \quad \text{and} \quad t_{\mathsf{B}} = \sum_i \alpha_i t^i_{\mathsf{B}}. \tag{4.5}$$

Proof of Claim. By induction on the structure of t. If t is the single vertex tree, then the claim is vacuously true. If t has a unique child, write s for the subtree rooted at this child and $g \in \Gamma_M$ for the element associated to the edge ts. Observe that $t_{\mathsf{A}} = g_{\mathsf{A}} s_{\mathsf{A}}$ and $t_{\mathsf{B}} = g_{\mathsf{B}} s_{\mathsf{B}}$. The inductive hypothesis applies to s yielding coefficients $\alpha_1, \dots, \alpha_m \in \mathbb{K}$ and trees $s^1, \dots, s^m \in T(\Gamma_M)$ of depth at most n such that Equation (4.5) holds. By Claim 4.1.13a, for every $i \in [m]$, there exist coefficients $\beta_{ij} \in \mathbb{K}$ and trees $r^{ij} \in T(\Gamma_M)$ of depth at most n such that $g_{\mathsf{A}} s^i_{\mathsf{A}} = \sum_j \beta_{ij} r^{ij}_{\mathsf{A}}$. In order to conclude that the same identity holds for B, observe that the tree represented by gs^i has depth at most $n + 1$. The same holds for all other trees occurring in the following calculation. For every $i \in [m]$,

$$\begin{aligned}
&\left\langle g_{\mathsf{B}} s^i_{\mathsf{B}} - \sum_j \beta_{ij} r^{ij}_{\mathsf{B}}, g_{\mathsf{B}} s^i_{\mathsf{B}} - \sum_j \beta_{ij} r^{ij}_{\mathsf{B}} \right\rangle \\
&= \operatorname{soe}(g_{\mathsf{B}} s^i_{\mathsf{B}} \odot g_{\mathsf{B}} s^i_{\mathsf{B}}) - 2 \sum_j \beta_{ij} \operatorname{soe}(g_{\mathsf{B}} s^i_{\mathsf{B}} \odot r^{ij}_{\mathsf{B}}) + \sum_{j,k} \beta_{ij} \beta_{ik} \operatorname{soe}(r^{ij}_{\mathsf{B}} \odot r^{ik}_{\mathsf{B}})
\end{aligned}$$

$$= \operatorname{soe}(g_{\mathsf{A}} s^i_{\mathsf{A}} \odot g_{\mathsf{A}} s^i_{\mathsf{A}}) - 2 \sum_j \beta_{ij} \operatorname{soe}(g_{\mathsf{A}} s^i_{\mathsf{A}} \odot r^{ij}_{\mathsf{A}}) + \sum_{j,k} \beta_{ij}\beta_{ik} \operatorname{soe}(r^{ij}_{\mathsf{A}} \odot r^{ik}_{\mathsf{A}})$$
$$= \left\langle g_{\mathsf{A}} s^i_{\mathsf{A}} - \sum_j \beta_{ij} r^{ij}_{\mathsf{A}}, g_{\mathsf{A}} s^i_{\mathsf{A}} - \sum_j \beta_{ij} r^{ij}_{\mathsf{A}} \right\rangle$$
$$= 0.$$

Thus, $g_{\mathsf{B}} s^i_{\mathsf{B}} = \sum_j \beta_{ij} r^{ij}_{\mathsf{B}}$, as desired. If t is of the form $t^1 \odot \cdots \odot t^r$ for some trees $t^1, \dots, t^r$ whose roots have degree one, then the first case applies to each of these subtrees. The claim follows readily. ◁

Finally, for every tree $t \in T(\Gamma_M)$ of arbitrary depth, let $\alpha_1, \dots, \alpha_m \in \mathbb{K}$ and $t^1, \dots, t^m \in T(\Gamma_M)$ be as in Claim 4.1.13b. Then

$$\operatorname{soe}(t_{\mathsf{A}}) = \sum \alpha_i \operatorname{soe}(t^i_{\mathsf{A}}) = \sum \alpha_i \operatorname{soe}(t^i_{\mathsf{B}}) = \operatorname{soe}(t_{\mathsf{B}}). \qquad \square$$

We conclude this section by bounding the degree of the trees which need to be considered in Theorem 4.1.10.

Theorem 4.1.14. *Writing $n := \max\{|I|, |J|\}$, the conditions of Theorem 4.1.10 are equivalent to the following: For every tree $t \in T(\Gamma_M)$ of depth at most $n+1$ and out-degree at most $2n-1$, $\operatorname{soe}(t_{\mathsf{A}}) = \operatorname{soe}(t_{\mathsf{B}})$.*

In particular, it suffices to consider trees on at most $\sum_{d=1}^{n+1} (2n-1)^{d-1} \leq (2n)^{n+1}$ vertices. We will comment on the tightness of the bounds in Theorem 4.1.14 in Remark 4.2.5.

Proof of Theorem 4.1.14. Suppose that $\operatorname{soe}(t_{\mathsf{A}}) = \operatorname{soe}(t_{\mathsf{B}})$ for all trees $t \in T(\Gamma_M)$ of depth at most $n+1$ and out-degree at most $2n-1$. Given Lemma 4.1.13, it suffices to show that $\operatorname{soe}(t_{\mathsf{A}}) = \operatorname{soe}(t_{\mathsf{B}})$ for all trees $t \in T(\Gamma_M)$ of depth at most $n+1$. To ease notation, suppose without loss of generality that I and J are disjoint. Similarly to the set-up of the proof of Theorem 4.1.3, define, for a tree $t \in T(\Gamma_M)$, the block vector $t_{\mathsf{A}\oplus\mathsf{B}} := \binom{t_{\mathsf{A}}}{t_{\mathsf{B}}} \in \mathbb{K}^{I \cup J}$. We construct, for every $d \geq 1$, a spanning set of vectors for the space $T^{\leq d}_{\mathsf{A}\oplus\mathsf{B}}$ spanned by the vectors $t_{\mathsf{A}\oplus\mathsf{B}}$ for all trees $t \in T(\Gamma_M)$ of depth $\leq d$ and out-degree $\leq 2n-1$. For a tree $t \in T(\Gamma_M)$ and $i \geq 1$, write $t^{\odot i}$ for the tree obtained by gluing i copies of t together. Furthermore, let $t^{\odot 0} := \mathbf{1}$ be the one-vertex tree.

Claim 4.1.14a. For every $d \geq 1$, there exist a set $T \subseteq T(\Gamma_M)$ of trees of depth $\leq d$, whose roots have out-degree ≤ 1, and all whose out-degrees are $\leq 2n-1$ such that

$$T^{\leq d}_{\mathsf{A}\oplus\mathsf{B}} = \operatorname{span}\{t^{\odot i}_{\mathsf{A}\oplus\mathsf{B}} \mid t \in T, 0 \leq i \leq 2n-1\}.$$

Proof of Claim. For $d = 1$, the singleton containing the one-vertex tree is as desired. For $d \geq 2$, consider the following equivalence relation on $I \cup J$: Let $i \sim_d j$ if, and only if, $t_{\mathsf{A}\oplus\mathsf{B}}(i) = t_{\mathsf{A}\oplus\mathsf{B}}(j)$ for all $t \in T(\Gamma_M)$ of depth $\leq d$. Observe that if $i \not\sim_d j$, then there exists $s \in T(\Gamma_M)$ of depth $\leq d$ and with root of degree 1 such that $s_{\mathsf{A}\oplus\mathsf{B}}(i) \neq s_{\mathsf{A}\oplus\mathsf{B}}(j)$. Indeed, if the root of a tree t such that $t_{\mathsf{A}\oplus\mathsf{B}}(i) \neq t_{\mathsf{A}\oplus\mathsf{B}}(j)$ has higher degree, then t is the gluing product of multiple trees with root of degree one, and one of these factors is as desired.

Let $S \subseteq T(\Gamma_M)$ be a set of trees of depth at most d and with roots of degree 1 such that $i \sim_d j$ if, and only if, $s_{\mathsf{A}\oplus\mathsf{B}}(i) = s_{\mathsf{A}\oplus\mathsf{B}}(j)$ for all $s \in S$. By Fact 2.4.6,

$$T^{\leq d}_{\mathsf{A}\oplus\mathsf{B}} = \operatorname{span}\{s^{\odot i}_{\mathsf{A}\oplus\mathsf{B}} \mid s \in S, 0 \leq i \leq 2n-1\}.$$

For $s \in S$, write s' for the tree rooted at the unique child of the root of s. Moreover, write $g^s \in \Gamma_M$ for the element associated to the edge incident to the root of s. The tree s' is of depth at most $d-1$. Write $Q \subseteq T(\Gamma_M)$ for the set of trees of depth at most $d-1$ with roots of out-degree ≤ 1, and all whose out-degrees are $\leq 2n-1$, which is guaranteed to exist by induction. Then, by linearity,

$$\begin{aligned} T^{\leq d}_{\mathsf{A}\oplus\mathsf{B}} &= \operatorname{span}\{s^{\odot i}_{\mathsf{A}\oplus\mathsf{B}} \mid s \in S, 0 \leq i \leq 2n-1\} \\ &= \operatorname{span}\{(g^s s')^{\odot i}_{\mathsf{A}\oplus\mathsf{B}} \mid s \in S, 0 \leq i \leq 2n-1\} \\ &\leq \operatorname{span}\{(g(q^{\odot j}))^{\odot i}_{\mathsf{A}\oplus\mathsf{B}} \mid q \in Q, g \in \Gamma_M, 0 \leq i,j \leq 2n-1\} \\ &\leq T^{\leq d}_{\mathsf{A}\oplus\mathsf{B}}. \end{aligned}$$

Note that all trees $g(q^{\odot j})$ appearing in the final set are of depth $\leq d$, have roots of out-degree ≤ 1, and only vertices of out-degree $\leq 2n-1$. Hence, the set of trees $\{g(q^{\odot j}) \mid q \in Q, g \in \Gamma_M, 0 \leq j \leq 2n-1\}$ is as desired. ◁

For a vector $v \in \mathbb{K}^{I \cup J}$, write $\operatorname{soe}_{\mathsf{A}}(v) := \sum_{i \in I} v(i)$ and analogously $\operatorname{soe}_{\mathsf{B}}(v) := \sum_{j \in J} v(j)$. Let $t \in T(\Gamma_M)$ be a tree of depth at most $n+1$. By Claim 4.1.14a, there exist trees $t^1, \dots, t^r \in \Gamma_M$ of depth at most $n+1$ and out-degrees at most $2n-1$ and coefficients $\alpha_1, \dots, \alpha_r \in \mathbb{K}$ such that $t_{\mathsf{A}\oplus\mathsf{B}} = \sum_{i=1}^r \alpha_i t^i_{\mathsf{A}\oplus\mathsf{B}}$. Hence,

$$\operatorname{soe}(t_{\mathsf{A}}) = \operatorname{soe}_{\mathsf{A}}(t_{\mathsf{A}\oplus\mathsf{B}}) = \sum \alpha_i \operatorname{soe}_{\mathsf{A}}(t^i_{\mathsf{A}\oplus\mathsf{B}}) = \sum \alpha_i \operatorname{soe}(t^i_{\mathsf{A}}) = \sum \alpha_i \operatorname{soe}(t^i_{\mathsf{B}}) = \operatorname{soe}(t_{\mathsf{B}}).$$

Thus, the assertion in Theorem 4.1.14 implies the assertion in Lemma 4.1.13. □

4.2 Cycles, Paths, and Trees

In this section, Theorems 4.1.2, 4.1.4, and 4.1.10 are showcased by reproving homomorphism indistinguishability characterisations of three well-studied equational graph isomorphism relaxations. The equational relaxations we consider are feasibility of $\mathrm{LP}(G, H)$, i.e. the system $\boldsymbol{A}_G X = X \boldsymbol{A}_H$, for matrices X which are doubly stochastic, pseudo-stochastic, or orthogonal, cf. Table 1.2. By Theorems 2.6.2

and 3.4.3, the existence of such a doubly stochastic matrix is equivalent to homomorphism indistinguishability over trees. By [61, Theorem 2], such a pseudo-stochastic matrix exists if, and only if, the two graphs are homomorphism indistinguishable over paths. It is folklore, cf. Theorem 3.3.2, that homomorphism indistinguishability over cycles is equivalent to the existence of such an orthogonal matrix.

Given the correspondence between combinatorial operations on (bi)labelled graphs and combinatorial operations on homomorphism tensors introduced in Section 3.2, the following Corollaries 4.2.1, 4.2.3, and 4.2.4 are proven by applying Theorems 4.1.2, 4.1.4, and 4.1.10 to the homomorphism tensors of bilabelled paths.

Recall $\boldsymbol{A} \in \mathcal{G}(1,1)$, the $(1,1)$-bilabelled edge, cf. Example 3.2.3 and Figure 3.3c. The homomorphism matrix $\boldsymbol{A}_G$ is the adjacency matrix of the graph G, cf. Example 3.2.9.

Corollary 4.2.1. *For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of cycles,*
2. *G and H are homomorphism indistinguishable over the class of cycles on at most $2n^2 - 1$ vertices,*
3. *there exists an orthogonal matrix $X \in \mathbb{R}^{V(G) \times V(H)}$ such that $\boldsymbol{A}_G X = X \boldsymbol{A}_H$.*

Proof. Apply Theorem 4.1.2 with $I := V(G)$ and $J := V(H)$, and $\mathsf{A} := (\boldsymbol{A}_G)$ and $\mathsf{B} := (\boldsymbol{A}_H)$. Series composition of $\boldsymbol{A}$ with itself yields precisely the $(1,1)$-bilabelled paths with labels at the vertices of degree ≤ 1, cf. Figure 3.5c. Taking the traces of their homomorphism matrices amounts to identifying the labels of these paths and thus counting homomorphisms from cycles into G and H, cf. Figure 3.6b.

By Theorem 4.1.3, it suffices to consider words in $\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ of length at most $2n^2 - 1$. Each letter corresponds to an edge. Thus, homomorphism counts from cycles on at most $2n^2 - 1$ vertices suffice. □

By Theorem 3.3.2, considering cycles on at most n vertices suffices. The bound in Corollary 4.2.1 is suboptimal as it is derived from the more general Theorem 4.1.3, which, contrary to Newton's identities, gives a criterion of simultaneous orthogonal similarity of multiple matrices. Note that the orthogonal matrix in Corollary 4.2.1 and hence in Theorem 4.1.2 cannot be demanded to have rational entries despite that the adjacency matrices of G and H are rational.

Remark 4.2.2 ([113, Theorem 7, Example 8])**.** There exist simple graphs G and H which are homomorphism indistinguishable over the class of cycles such that there exists no rational orthogonal matrix $X \in \mathbb{Q}^{V(G) \times V(H)}$ satisfying $\boldsymbol{A}_G X = X \boldsymbol{A}_H$.

For paths, we obtain a result analogous to Corollary 4.2.1. The equivalence of the first and last assertion was first shown by Dell, Grohe, & Rattan [61, Theorem 2].

Corollary 4.2.3. *For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of paths,*
2. *G and H are homomorphism indistinguishable over the class of paths on at most $2n$ vertices,*
3. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G) \times V(H)}$ such that $\boldsymbol{A}_G X = X \boldsymbol{A}_H$.*

Proof. Recall the proof of Corollary 4.2.1 and invoke Theorem 4.1.4. Taking sums-of-entries of homomorphism matrices of $(1,1)$-bilabelled paths amounts to counting homomorphisms from the underlying unlabelled paths into G and H, cf. Figure 3.5a. By Theorem 4.1.6, it suffices to consider words in $\boldsymbol{A}_G$ and $\boldsymbol{A}_H$ of length at most $2n-1$. Each letter corresponds to an edge. Thus, homomorphism counts from paths on at most $2n$ vertices suffice. □

The characterisation from [61, Theorem 1] of homomorphism indistinguishability over trees involves a non-negativity condition on the matrix X. While such an assumption appears natural from the viewpoint of solving the system of equations for fractional isomorphism, it lacks an algebraic or combinatorial interpretation. Using Theorem 4.1.10, we reprove this known characterisation and give an alternative description that emphasises its graph-theoretic origin. Recall the family of 1-labelled trees $\mathcal{T} \subseteq \mathcal{G}(1)$ defined in Section 3.4.

Corollary 4.2.4. *For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of trees,*
2. *G and H are homomorphism indistinguishable over all trees T for which there exists a vertex $r \in V(T)$ such that the rooted (T,r) is of depth at most $n+1$ and maximum out-degree at most $2n-1$,*
3. *G and H are homomorphism indistinguishable over all trees on at most $(2n)^{n+1}$ vertices,*
4. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G) \times V(H)}$ satisfying $\boldsymbol{A}_G X = X \boldsymbol{A}_H$ and one of the following equivalent conditions:*
   a) *all entries of X are non-negative,*
   b) *$X\boldsymbol{T}_H = \boldsymbol{T}_G$ for all 1-labelled trees $\boldsymbol{T} \in \mathcal{T}$,*
   c) *X preserves the Schur product on $\mathbb{R}\mathcal{T}_H$, the space spanned by the $\boldsymbol{T}_H$ for $\boldsymbol{T} \in \mathcal{T}$, i.e. $X(u \odot v) = (Xu) \odot (Xv)$ for all $u, v \in \mathbb{R}\mathcal{T}_H$.*

Proof. By Theorem 4.1.14, the first two assertions are equivalent. A tree as in the second assertion has at most $\sum_{d=1}^{n+1} (2n-1)^{d-1} \leq (2n)^{n+1}$ vertices. Hence, the third assertion implies the second. Clearly, the first assertion implies the third.

For the last assertion, consider the following argument: The equivalence of Items 4a and 4b is immediate from Lemma 4.1.11. Assuming Item 4b, Item 4c follows since $X(\boldsymbol{T}_H \odot \boldsymbol{S}_H) = X((\boldsymbol{T} \odot \boldsymbol{S})_H) = (\boldsymbol{T} \odot \boldsymbol{S})_G = \boldsymbol{T}_G \odot \boldsymbol{S}_G$ for all $\boldsymbol{T}, \boldsymbol{S} \in \mathcal{T}$.

Conversely, by induction on the structure of $\boldsymbol{T} \in \mathcal{T}$, if $\boldsymbol{T} = \boldsymbol{A} \cdot \boldsymbol{S}$ for some $\boldsymbol{S} \in \mathcal{T}$, then $X\boldsymbol{T}_H = \boldsymbol{A}_G X \boldsymbol{S}_H = \boldsymbol{T}_G$ by the assumption $X\boldsymbol{A}_H = \boldsymbol{A}_G X$. If $\boldsymbol{T} = \boldsymbol{R} \odot \boldsymbol{S}$ for some $\boldsymbol{R}, \boldsymbol{S} \in \mathcal{T}$, then the claim follows immediately from Item 4c. □

We finally comment on the optimality of the bounds in Corollary 4.2.4.

Remark 4.2.5. By Corollary 6.3.10, homomorphism counts of constant degree trees do not have the same distinguishing power as homomorphism counts from all trees. In particular, the bound on the maximum degree in Corollary 4.2.4 cannot be replaced with a constant. Furthermore, by [71], there exist graphs G and H on n vertices which are distinguished by Colour Refinement but only in $\Theta(n)$ iterations. Thus, by [61], cf. Theorem 3.4.1, these graphs are homomorphism indistinguishable over all trees T for which there exists $r \in V(T)$ such that the rooted tree (T, r) is of depth $\Theta(n)$. Thereby, the bound in Corollary 4.2.4 on the depth of the trees is asymptotically tight.

4.3 Cyclewidth, Pathwidth, Treewidth, and Treedepth

In Corollaries 4.2.1, 4.2.3, and 4.2.4, the machinery from Section 4.1 was applied to involution monoids which are generated by a single non-trivial generator, namely the bilabelled edge, cf. Figure 3.3c. In this section, we consider involution monoids which are generated by more than one non-trivial generator. In the language of Theorems 4.1.2, 4.1.4, and 4.1.10, this amounts to considering multiple matrices. As before, the matrices are homomorphism matrices of bilabelled graphs. Using multiple such graphs permits the treatment of more complicated graph classes such the classes of graphs of bounded cycle-, path-, treewidth, and treedepth.

The following subsections feature four different algebro-combinatorial setups, which are summarised in Figure 4.2. The algebraic structure of the considered class of (bi)labelled graphs determines the domain of the matrix variables in the matrix equations whose feasibility is equivalent to homomorphism indistinguishability over the family of underlying unlabelled graphs. Domains covered by our results are orthogonal matrices, pseudo-stochastic matrices, and doubly stochastic matrices. In some cases, feasibility over two of these possible domains coincides (Section 4.3.4).

4.3.1 Cycle- and Pathwidth: Generators for Involution Monoids

The families of bilabelled graphs considered in this section all are involution monoids in the following sense. For $k \geq 1$, the class $\mathcal{G}(k,k)$ of all (k,k)-bilabelled graphs forms an involution monoid whose binary operation is series composition, whose involution operation is reversal, and whose neutral element is the identity graph $\boldsymbol{I} = (I, (1, \dots, k), (1, \dots, k))$ with $V(I) = [k]$ and $E(I) = \emptyset$, cf. Figure 3.3b. Here it is crucial to to regard isomorphic bilabelled graphs as equal, cf. Remark 3.2.4, since every monoid necessarily has a unique neutral element.

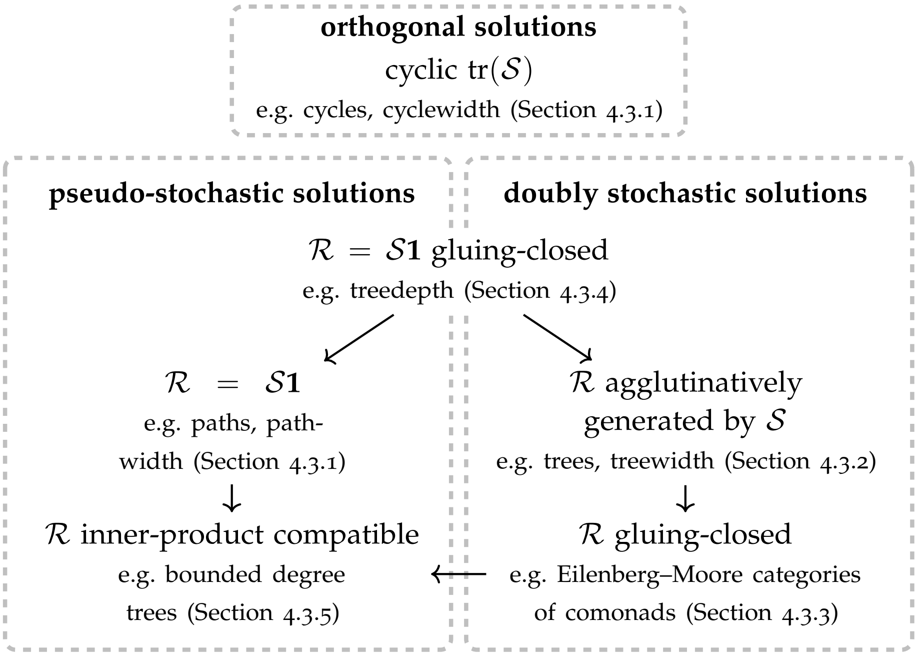


Figure 4.2: Interplay of a family of labelled graphs $\mathcal{R} \subseteq \mathcal{G}(k)$ and a family of bilabelled graphs $\mathcal{S} \subseteq \mathcal{G}(k, k)$ yielding matrix equations with orthogonal, pseudo-stochastic, or doubly stochastic solutions. Arrows indicate implications, e.g. every gluing-closed family of labelled graphs $\mathcal{R}$ is inner-product compatible.

This section features subclasses $\mathcal{S} \subseteq \mathcal{G}(k,k)$ which also form involution monoids. That is, they satisfy the following properties:

1. $\boldsymbol{I} \in \mathcal{S}$,
2. $\boldsymbol{S}^* \in \mathcal{S}$ for all $\boldsymbol{S} \in \mathcal{S}$,
3. $\boldsymbol{S} \cdot \boldsymbol{S}' \in \mathcal{S}$ for all $\boldsymbol{S}, \boldsymbol{S}' \in \mathcal{S}$.

An example of an involution monoid of $(1,1)$-bilabelled graphs is the *path monoid* of all $(1,1)$-bilabelled paths with labels at opposing ends, cf. Figure 3.5c. In order to derive systems of equations with finitely many equations, we consider finite generating sets of involution monoids:

Definition 4.3.1. Let $\mathcal{S}$ be an involution monoid. A set $\mathcal{B} \subseteq \mathcal{S}$ *generates* $\mathcal{S}$ if

1. $\boldsymbol{I} \in \mathcal{B}$,
2. $\boldsymbol{B}^* \in \mathcal{B}$ for all $\boldsymbol{B} \in \mathcal{B}$,
3. for all $\boldsymbol{S} \in \mathcal{S}$ there exist $\boldsymbol{B}^1, \dots, \boldsymbol{B}^r \in \mathcal{B}$ such that $\boldsymbol{S} = \boldsymbol{B}^1 \cdots \boldsymbol{B}^r$.

For example, the path monoid is generated by the $(1,1)$-bilabelled graph $\boldsymbol{A}$ depicted by Figure 3.3c and the identity graph $\boldsymbol{I}$, cf. Figure 3.5c.

Recall that $\mathcal{F}_{\leq \ell} := \{F \in \mathcal{F} \mid |V(F)| \leq \ell\}$ for a graph class $\mathcal{F}$ and $\ell \in \mathbb{N}$. For a class $\mathcal{S} \subseteq \mathcal{G}(k,k)$ of (k,k)-bilabelled graphs where $k \geq 1$, write $\operatorname{soe}(\mathcal{S}) := \{\operatorname{soe}(\boldsymbol{S}) \mid \boldsymbol{S} \in \mathcal{S}\}$ and $\operatorname{tr}(\mathcal{S}) := \{\operatorname{tr}(\boldsymbol{S}) \mid \boldsymbol{S} \in \mathcal{S}\}$. Both $\operatorname{soe}(\mathcal{S})$ and $\operatorname{tr}(\mathcal{S})$ are classes of unlabelled graphs. The following Theorem 4.3.2 is immediate from Theorems 4.1.2 to 4.1.4 and 4.1.6.

Theorem 4.3.2. *Let $k \geq 1$. Let $\mathcal{S} \subseteq \mathcal{G}(k,k)$ be an involution monoid generated by $\mathcal{B} \subseteq \mathcal{S}$. Let G and H be simple graphs on at most n vertices. Suppose that every graph in $\mathcal{B}$ has at most $b \in \mathbb{N} \cup \{\infty\}$ vertices and let $N_1 := 2n^k b \in \mathbb{N} \cup \{\infty\}$ and $N_2 := 2n^{2k} b \in \mathbb{N} \cup \{\infty\}$. Then the following are equivalent:*

1. *G and H are homomorphism indistinguishable over* $\operatorname{soe}(\mathcal{S})$,
2. *G and H are homomorphism indistinguishable over* $(\operatorname{soe}(\mathcal{S}))_{\leq N_1}$,
3. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{B}$.*

Furthermore, the following are equivalent:

1. *G and H are homomorphism indistinguishable over* $\operatorname{tr}(\mathcal{S})$,
2. *G and H are homomorphism indistinguishable over* $(\operatorname{tr}(\mathcal{S}))_{\leq N_2}$,
3. *there exists an orthogonal matrix $U \in \mathbb{R}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G U = U \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{B}$.*

Proof. In the set-up of Theorems 4.1.2 and 4.1.4, let $I := V(G)^k$, $J := V(H)^k$, and $M := \mathcal{B}$. Furthermore, let A (respectively, B) be the sequence of homomorphism tensors $\boldsymbol{B}_G$ (respectively, $\boldsymbol{B}_H$) for $\boldsymbol{B} \in \mathcal{B}$. Words $w \in \Gamma_M$ corresponds to bilabelled graphs from $\mathcal{S}$ and vice-versa. The matrices w_{A} and w_{B} are homomorphism tensors of such a bilabelled graph. With these observations, Theorem 4.3.2 is immediate from Theorems 4.1.2 to 4.1.4 and 4.1.6. □

We remark that we are only interested in the order of magnitude of the parameters N_1 and N_2. In order to state Theorem 4.3.2 more clearly, we chose to be slightly wasteful compared to Theorems 4.1.3 and 4.1.6. See also Section 9.1.

The remainder of this section features an application of Theorem 4.3.2 to homomorphism indistinguishability over graphs of bounded pathwidth and cyclewidth. The prototypical example of an involution monoid is the family of graphs of pathwidth at most k.

Definition 4.3.3. Let $k \geq 1$. Let $\mathcal{PW}(k,k)$ denote the family of all (k,k)-bilabelled graphs $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$ such that F admits a path decomposition (P, β) with vertices $u, v \in V(P)$ satisfying

1. $\beta(u) = \{u_1, \dots, u_k\}$ and $\beta(v) = \{v_1, \dots, v_k\}$,
2. if $u \neq v$, then $\deg_P(u) = \deg_P(v) = 1$; if $u = v$, then $\deg_P(u) = \deg_P(v) = 0$,
3. $|\beta(s)| = k$ for all $s \in V(P)$ and $|\beta(s) \cap \beta(t)| = k - 1$ for all $st \in E(P)$.

The first two axioms of Definition 4.3.3 prescribe where in a path decomposition the labelled vertices have to be placed. The last axiom makes subsequent arguments easier and does not constitute a loss of generality, cf. Lemma 2.1.2. We first show that $\mathcal{PW}(k,k)$ is an involution monoid.

Lemma 4.3.4. *For $k \geq 1$, $\mathcal{PW}(k,k)$ is an involution monoid.*

Proof. Clearly, $\boldsymbol{I} \in \mathcal{PW}(k,k)$ and $\mathcal{PW}(k,k)$ is closed under reversal. In order to show that $\mathcal{PW}(k,k)$ is closed under series composition, let $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$, and $\boldsymbol{F}' = (F', \boldsymbol{u}', \boldsymbol{v}')$ be elements of $\mathcal{PW}(k,k)$ and let (P, β), $u, v \in V(P)$ and (P', β'), $u', v' \in V(P')$ be for $\boldsymbol{F}$ and $\boldsymbol{F}'$ as in Definition 4.3.3, respectively. The graph $\boldsymbol{F}'' := \boldsymbol{F} \cdot \boldsymbol{F}'$ is obtained by taking the disjoint union of F and F' and identifying v_i with u'_i for all $i \in [k]$. A suitable path decomposition (P'', β'') for $\boldsymbol{F}''$ can be constructed by taking the disjoint union of P and P' and identifying v with u'. The map $\beta'' \colon V(P'') \to 2^{V(F'')}$ is defined to coincide with β on $V(P)$ and with β' on $V(P')$. On v and u', where the two domains coincide, the maps β and β' have the same image in $V(F'')$. Let $u'' := u$ and $v'' := v'$. Then (P'', β''), u'' and v'' are for $\boldsymbol{F}''$ as in Definition 4.3.3. □

The following Lemma 4.3.5 describes the unlabelled graphs which can be obtained from the bilabelled graphs in $\mathcal{PW}(k,k)$ by unlabelling and taking traces.

Lemma 4.3.5. *Let $k \geq 1$.*

1. *The class $\operatorname{soe}(\mathcal{PW}(k,k))$ is the class of all graphs of pathwidth at most $k-1$ on at least k vertices.*
2. *The class $\operatorname{tr}(\mathcal{PW}(k,k))$ is the class of all graphs of cyclewidth at most $k-1$ on at least k vertices.*

Proof. By Definition 4.3.3, every $F \in \operatorname{soe}(\mathcal{PW}(k,k))$ has pathwidth at most $k-1$ and at least k vertices. Conversely, if F has pathwidth at most $k-1$ and at least k vertices, then, by Lemma 2.1.2, there exists a path decomposition (P, β) of F satisfying Item 3 of Definition 4.3.3. The labels can be arbitrarily placed on F in accordance with Definition 4.3.3.

For the second claim, let $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{PW}(k, k)$ be a bilabelled graph with path decomposition (P, β) and $u, v \in V(P)$ as in Definition 4.3.3. Let C denote the cycle obtained from P by making a fresh vertex w adjacent to u and v. Extend β to C by letting $\beta(w) := \beta(u) \cup \beta(v)$. In the unlabelled graph $\operatorname{tr}(\boldsymbol{F})$, the vertices u_i and v_i for $i \in [k]$ are respectively identified. Hence, (C, β) gives rise to a cycle decomposition of $\operatorname{tr}(\boldsymbol{F})$ of width at most k.

Conversely, by Lemma 2.1.2, if F has cyclewidth at most $k-1$ and at least k vertices, then there is a cycle decomposition (C, β) of F satisfying Item 3 of Definition 4.3.3. Pick any vertex $w \in V(C)$. Let P denote the path obtained from C by replacing w with two fresh vertices u and v adjacent to the two neighbours of w respectively. Write $X \subseteq \beta(w)$ for the set of all vertices $x \in \beta(w)$ which do not appear in all bags of (C, β), i.e. $x \notin \beta(y)$ for some $y \in V(C)$. For $x \in X$, let $c_1, \dots, c_r$ and $d_1, \dots, d_s$ denote the vertices of C whose bags contain x. Suppose that $d_s d_{s-1} \dots d_1 w c_1 \dots c_r$ is a path in C and that u was made adjacent to c_1, and v to d_1. Construct a graph F' from F by replacing every $x \in X$ by two vertices x' and x'' and making x' adjacent to all neighbours of x in $\beta(c_1) \cup \dots \cup \beta(c_r)$ and x'' adjacent to all neighbours of x in $\beta(d_1) \cup \dots \cup \beta(d_s)$. Define a path decomposition (P, γ) of F' by letting $\gamma(u) := (\beta(w) \setminus X) \cup \{x' \mid x \in X\}$ and $\gamma(v) := (\beta(w) \setminus X) \cup \{x'' \mid x \in X\}$. Every other bag $\gamma(z)$ for $z \in V(P) \setminus \{u, v\}$ is obtained from $\beta(z)$ by replacing $x \in X$ by x' or x'' depending on whether z is among the $c_1, \dots, c_r$ or the $d_1, \dots, d_s$.

Let $\boldsymbol{u}, \boldsymbol{v} \in V(F')^k$ be tuples comprised of the vertices of $\gamma(u)$ and $\gamma(v)$ such that if $u_i = x'$ for some $x \in X$ and $i \in [k]$, then $v_i = x''$. Also, if $u_i \in \beta(w) \setminus X$ for $i \in [k]$, then we require that $u_i = v_i$. Let $\boldsymbol{F'} := (F', \boldsymbol{u}, \boldsymbol{v})$. Taking the trace of $\boldsymbol{F'}$ has the effect that, for every $x \in X$, the copies x' and x'' are identified. Hence, $\operatorname{tr}(\boldsymbol{F'}) \cong F$, as desired. ☐

To apply Theorem 4.3.2, it remains to give a set of generators for $\mathcal{PW}(k, k)$, cf. Figures 3.3b and 4.3.

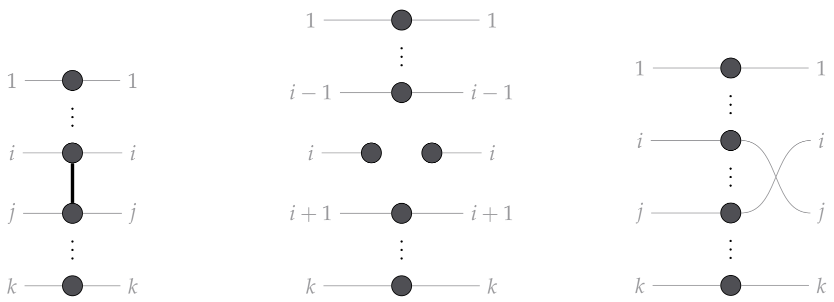
Lemma 4.3.6. *Let $\mathcal{B}(k, k)$ denote the set of the following (k, k)-bilabelled graphs. For $1 \leq i, j \leq k$,*

- *the* identity graph $\boldsymbol{I} = (I, (1, \dots, k), (1, \dots, k))$ *with* $V(I) = [k]$, $E(I) = \emptyset$,
- *the* adjacency graphs $\boldsymbol{A}^{ij} = (A^{ij}, (k), (k))$ *with* $V(A^{ij}) = [k]$ *and* $E(A) = \{ij\}$,
- *the* forgetting graphs

$$\boldsymbol{J}^i = (J^i, (1, \dots, k), (1, \dots, i-1, i', i+1, \dots, k))$$

with $V(J^i) = [k] \cup \{i'\}$ *and* $E(J^i) = \emptyset$,
- *the* swap graphs

$$\boldsymbol{S}^{ij} = (S^{ij}, (1, \dots, k), (1, \dots, i-1, j, i+1, \dots, j-1, i, j+1, \dots, k))$$



(a) adjacency graph $\boldsymbol{A}^{ij}$. **(b)** forgetting graph $\boldsymbol{J}^{i}$. **(c)** swap graph $\boldsymbol{S}^{ij}$.

Figure 4.3: Bilabelled graphs in $\mathcal{B}(k,k)$ as defined in Lemma 4.3.6.

with $V(S^{ij}) = [k]$ and $E(S^{ij}) = \emptyset$.
Then $\mathcal{B}(k,k)$ generates $\mathcal{PW}(k,k)$.

Proof. Clearly, $\mathcal{B}(k,k) \subseteq \mathcal{PW}(k,k)$. Items 1 and 2 of Definition 4.3.1 are immediate. It remains to verify Item 3: To that end, let F be an arbitrary graph with a path decomposition (P,β) with vertices $u,v \in V(P)$ and $\boldsymbol{u},\boldsymbol{v} \in V(F)^k$ as in Definition 4.3.3. The proof is by induction on $|V(P)|$.

If $|V(P)| = 1$, then $u = v$ and $\{u_1,\dots,u_k\} = \{v_1,\dots,v_k\}$. Hence, there exists a permutation $\sigma\colon [k] \to [k]$ such that $u_i = v_{\sigma(i)}$ for all $i \in [k]$. Write $\sigma = \tau_1 \cdots \tau_r$ as product of transpositions. Then $\boldsymbol{F} = (F,\boldsymbol{u},\boldsymbol{v})$ is equal to

$$\boldsymbol{S}^{\tau_1} \cdots \boldsymbol{S}^{\tau_r} \cdot \prod_{i,j\in[k] \text{ s.t. } v_iv_j \in E(F)} \boldsymbol{A}^{ij},$$

a product of graphs in $\mathcal{B}(k,k)$.

If $|V(P)| \geq 2$, let $w \in V(P)$ denote the unique neighbour of u. Let $P' := P - u$. The subgraph F' of F induced by $\bigcup_{p\in V(P')} \beta(p)$ satisfies Definition 4.3.3 with the path decomposition $(P',\beta|_{V(P')})$, the vertices w,v, the tuple $\boldsymbol{v}$ and some tuple $\boldsymbol{w} \in V(F')^k$ such that $\beta(w) = \{w_1,\dots,w_k\}$ and $w_i = u_i$ for all $i \in [k] \setminus \{\ell\}$ for some $\ell \in [k]$. Let $\boldsymbol{F}' := (F',\boldsymbol{w},\boldsymbol{v})$. Then

$$\boldsymbol{F} = \prod_{i,j\in[k] \text{ s.t. } u_iu_j \in E(F)} \boldsymbol{A}^{ij} \cdot \boldsymbol{J}^{\ell} \cdot \boldsymbol{F}'.$$

The claim follows inductively. □

In order to avoid the technicalities of working with small graphs, we record the following Lemma 4.3.7, which describes a padding trick.

Lemma 4.3.7. *Let $\mathcal{F}' \subseteq \mathcal{F}$ be graph classes such that $nK_1 \in \mathcal{F}'$ for some $n \geq 1$. Suppose that, for all graphs $F \in \mathcal{F} \setminus \mathcal{F}'$, it holds that $F + \ell K_1 \in \mathcal{F}'$ for some $\ell \geq 1$. Then, for all simple graphs G and H, it holds that $G \equiv_{\mathcal{F}} H$ if, and only if, $G \equiv_{\mathcal{F}'} H$.*

Proof. Since $\mathcal{F}' \subseteq \mathcal{F}$, it suffices to argue that the backward implication holds. Suppose that G and H are homomorphism indistinguishable over $\mathcal{F}'$. Since $nK_1 \in \mathcal{F}'$ for some $n \geq 1$, it holds that $|V(G)|^n = \hom(nK_1, G) = \hom(nK_1, H) = |V(H)|^n$. Hence, G and H have the same number of vertices. Suppose without loss of generality that this number is non-zero. Let $F \in \mathcal{F} \setminus \mathcal{F}'$. Then $F + \ell K_1 \in \mathcal{F}'$ for some $\ell \geq 1$. Hence, by Equation (2.1),

$$\hom(F, G) = \frac{\hom(F + \ell K_1, G)}{\hom(\ell K_1, G)} = \frac{\hom(F + \ell K_1, H)}{\hom(\ell K_1, H)} = \hom(F, H),$$

which implies that G and H are homomorphism indistinguishable over $\mathcal{F}$. □

For the case of the classes of graphs of pathwidth at most k or cyclewidth at most k, we apply Lemma 4.3.7 with $\mathcal{F}$ being the respective graph class and $\mathcal{F}_{\geq k} := \{F \in \mathcal{F} \mid |V(F)| \geq k\}$ assuming the role of $\mathcal{F}'$. With this choice of $\mathcal{F}'$, the somewhat cumbersome assumptions of Lemma 4.3.7 can be alleviated for graph classes which are minor-closed and closed under disjoint unions. For example, by Lemma 4.3.5, it holds that two graphs G and H are homomorphism indistinguishable over the class of graphs of pathwidth at most k if, and only if, $G \equiv_{\mathrm{soe}(\mathcal{PW}(k,k))} H$. In contrast, the class of graphs of cyclewidth at most k is not closed under disjoint unions but satisfies the weaker assumptions of Lemma 4.3.7. Indeed, if a graph F has at most k vertices, then it has cycle decomposition with a single bag and all graphs $F + nK_1$ for $n \in \mathbb{N}$ are also of cyclewidth at most k. Thus Lemmas 4.3.5 and 4.3.7 yield that two graphs G and H are homomorphism indistinguishable over the class of graphs of cyclewidth at most k if, and only if, $G \equiv_{\mathrm{tr}(\mathcal{PW}(k,k))} H$.

This concludes the preparation for obtaining a system of matrix equations characterising homomorphism indistinguishability over graphs of bounded pathwidth and cyclewidth via Theorem 4.3.2. For later reference in Section 4.4.1, we denote the system of linear equations in Item 3 of Theorem 4.3.8 by $\mathrm{PW}^k(G, H)$.

Theorem 4.3.8. *Let $k \geq 1$. For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over graphs of pathwidth at most $k-1$,*
2. *G and H are homomorphism indistinguishable over graphs of pathwidth at most $k-1$ on at most $2n^k(k+1)$ vertices,*
3. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{B}(k,k)$.*

Proof. In virtue of Lemma 4.3.6, we apply Theorem 4.3.2 to the involution monoid $\mathcal{PW}(k,k)$ with generating set $\mathcal{B}(k,k)$. Write $N := 2n^k(k+1)$. Consider the following additional assertions:

4. G and H are homomorphism indistinguishable over graphs of pathwidth at most $k-1$ on at least k vertices,

5. G and H are homomorphism indistinguishable over graphs of pathwidth at most $k-1$ on at least k vertices and at most N vertices.

By Theorem 4.3.2 and Lemma 4.3.5, Items 3 to 5 are equivalent. By containment of the respective graph classes, Item 1 implies Item 2, which implies Item 5. By Lemma 4.3.7, Items 1 and 4 are equivalent. This closes a cycle of implications. □

Analogously, we obtain the following theorem:

Theorem 4.3.9. *Let $k \geq 1$. For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over graphs of cyclewidth at most $k-1$,*
2. *G and H are homomorphism indistinguishable over graphs of cyclewidth at most $k-1$ on at most $2n^{2k}(k+1)$ vertices,*
3. *there exists an orthogonal matrix $U \in \mathbb{R}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G U = U \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{B}(k,k)$.*

4.3.2 Treewidth: Agglutinative Generation

The class of graphs of treewidth at most $k-1$ is generated by the same generators as $\mathcal{PW}(k,k)$, cf. Lemma 4.3.6. However, instead of only considering series composition, we require the gluing product as well. To that end, recall the k-labelled version $\mathbf{1}$ of the (k,k)-bilabelled graph $\boldsymbol{I}$ from Example 3.2.3 and Figure 3.3a.

Definition 4.3.10. Let $k \geq 1$. Let $\mathcal{S} \subseteq \mathcal{G}(k,k)$ be an involution monoid. The set $\mathcal{X} \subseteq \mathcal{G}(k)$ of graphs *agglutinatively generated by* $\mathcal{S}$ is inductively defined as follows:

1. $\mathbf{1} \in \mathcal{X}$,
2. $\boldsymbol{S} \cdot \boldsymbol{X} \in \mathcal{X}$ for $\boldsymbol{S} \in \mathcal{S}$ and $\boldsymbol{X} \in \mathcal{X}$,
3. $\boldsymbol{X} \odot \boldsymbol{X}' \in \mathcal{X}$ for $\boldsymbol{X}, \boldsymbol{X}' \in \mathcal{X}$.

For a class $\mathcal{R} \subseteq \mathcal{G}(k)$ where $k \geq 1$, write $\operatorname{soe}(\mathcal{R}) := \{\operatorname{soe}(\boldsymbol{R}) \mid \boldsymbol{R} \in \mathcal{R}\}$. For agglutinatively generated graph classes, the following general theorem follows from Theorems 4.1.10 and 4.1.14:

Theorem 4.3.11. *Let $k \geq 1$. Let $\mathcal{S} \subseteq \mathcal{G}(k,k)$ be an involution monoid generated by $\mathcal{B} \subseteq \mathcal{S}$. Let $\mathcal{X} \subseteq \mathcal{G}(k)$ be the class agglutinatively generated by $\mathcal{S}$. Let G and H be simple graphs on at most n vertices. Suppose that every graph in $\mathcal{B}$ has at most $b \in \mathbb{N} \cup \{\infty\}$ vertices and let $N := (2n^k)^{n^k+1} b \in \mathbb{N} \cup \{\infty\}$. Then the following are equivalent:*

1. *G and H are homomorphism indistinguishable over $\operatorname{soe}(\mathcal{X})$,*
2. *G and H are homomorphism indistinguishable over $(\operatorname{soe}(\mathcal{X}))_{\leq N}$,*
3. *there exists a doubly stochastic matrix $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{B}$.*

Equipped with this general theorem, we now turn to establishing that the class of graphs of bounded treewidth is subject to it. To that end, we first define a suitable class of labelled graphs. Definition 4.3.12 parallels the definition of $\mathcal{PW}(k,k)$ from Definition 4.3.3.

Definition 4.3.12. Let $k \geq 1$. Define $\mathcal{TW}(k) \subseteq \mathcal{G}(k)$ as the class of k-labelled graphs $\boldsymbol{F} = (F, \boldsymbol{u})$ such that there exists a tree decomposition (T, β) of F and a vertex $r \in V(T)$ such that

1. $\beta(r) = \{u_1, \dots, u_k\}$ and
2. $|\beta(s)| = k$ for all $s \in V(T)$ and $|\beta(s) \cap \beta(t)| = k - 1$ for all $st \in E(T)$.

We show that $\mathcal{TW}(k)$ is the class which is agglutinatively generated by $\mathcal{PW}(k,k)$.

Lemma 4.3.13. *Let $k \geq 1$. The class* $\operatorname{soe}(\mathcal{TW}(k))$ *is the class of all graphs of treewidth at most $k-1$ on at least k vertices. Furthermore, $\mathcal{TW}(k)$ is agglutinatively generated by $\mathcal{PW}(k,k)$.*

Proof. The first assertions is immediate from Definition 4.3.12 and Lemma 2.1.2. For the second assertion, write $\mathcal{X}$ for the class agglutinatively generate by $\mathcal{PW}(k,k)$.

We first show that all graphs in $\boldsymbol{F} = (F, \boldsymbol{u}) \in \mathcal{X}$ admit a tree decomposition (T, β) of width at most $k-1$ such that there is a vertex $r \in V(T)$ with $\beta(r) = \{u_1, \dots, u_k\}$. We call the vertex r the *root* of the decomposition. The hypothesis clearly holds for $\boldsymbol{1}$.

By structural induction, suppose that $\boldsymbol{F} = \boldsymbol{S} \cdot \boldsymbol{X}$ for $\boldsymbol{X} = (X, \boldsymbol{x}) \in \mathcal{X}$ of lesser complexity and $\boldsymbol{S} = (S, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{PW}(k,k)$. Let (P, β) denote the path decomposition of $\boldsymbol{S}$ with vertices $u, v \in V(P)$ as stipulated in Definition 4.3.3. Let (T, γ) denote the tree decomposition of $\boldsymbol{X}$ with root r whose existence is guaranteed by the inductive hypothesis. Define a tree Q by taking the disjoint union of P and T and identifying v and r. Define $\alpha \colon V(Q) \to 2^{V(F)}$ via

$$\alpha(q) = \begin{cases} \beta(q), & \text{if } q \in V(P), \\ \gamma(q), & \text{if } q \in V(T). \end{cases}$$

Since $\beta(v) = \gamma(r)$ implicitly, the map α is a tree decomposition of width at most k of F. By construction, all labelled vertices in $\boldsymbol{F}$ lie in the same bag.

If $\boldsymbol{F} = \boldsymbol{X} \odot \boldsymbol{X}'$, a tree decomposition for $\boldsymbol{F}$ can be constructed from the tree decompositions of $\boldsymbol{X}$ and $\boldsymbol{X}'$ by taking the disjoint union of the decomposition trees and identifying the roots.

Thus, every $\boldsymbol{F} = (F, \boldsymbol{u}) \in \mathcal{X}$ is such that F admits a tree decomposition (T, β) of width at most $k-1$ with a vertex $r \in V(T)$ such that $\beta(r) = \{u_1, \dots, u_k\}$. By the proof of Lemma 2.1.2, this tree decomposition can be modified while maintaining $\beta(r)$ such that it satisfies the second property of Definition 4.3.12. Hence, $\mathcal{X} \subseteq \mathcal{TW}(k)$.

Conversely, we argue by induction on the size of decomposition tree T that every $\boldsymbol{F} \in \mathcal{TW}(k)$ with tree decomposition (T, β) is in $\mathcal{X}$.

If $|V(T)| = 1$, then T is a path. Clearly, $\boldsymbol{F}' := (F, \boldsymbol{u}, \boldsymbol{u}) \in \mathcal{PW}(k,k)$ and $\boldsymbol{F} = \boldsymbol{F}'\boldsymbol{1} \in \mathcal{X}$, as desired.

If $|V(T)| > 1$, distinguish two cases: First suppose that r has only one neighbour r'. Define T' as the tree obtained from T by deleting r and write F' for the subgraph of F induced by $\bigcup_{t' \in V(T')} \beta(t')$. Let β' denote the restriction of β to $V(T')$. By Lemma 2.1.2, there is a unique index $i \in [k]$ such that $u_i \in \beta(r) \setminus \beta(r')$. Write x for the unique vertex in $\beta(r') \setminus \beta(r)$. Then the inductive hypothesis applies to F', (T', β'), r' and $\boldsymbol{v} := u_1 \dots u_{i-1} x u_{i+1} \dots u_k \in V(F')^k$. Then

$$\boldsymbol{F} = \boldsymbol{A} \cdot \boldsymbol{J}^i \cdot (F', \boldsymbol{v}) \in \mathcal{X} \qquad \text{for } \boldsymbol{A} := \prod_{\ell, j \in [k] \text{ s.t. } u_\ell u_j \in E(F)} \boldsymbol{A}^{\ell j}.$$

Finally, suppose that r has multiple neighbours $r'_1, \dots, r'_m$. For $i \in [m]$, write T_i for the connected component of r'_i in the forest obtained from T by deleting $r'_1, \dots, r'_{i-1}, r'_{i+1}, \dots, r'_m$. Write F'_i for the subgraph of F induced by $\bigcup_{t' \in V(T_i)} \beta(t')$ and β'_i for the restriction of β to $V(T'_i)$. Observe that r is of degree one in all graphs $T_1, \dots, T_m$. By the previous case, $(F'_i, \boldsymbol{u}) \in \mathcal{X}$. Clearly, $\boldsymbol{F} = \bigodot_{i=1}^m (F'_i, \boldsymbol{u}) \in \mathcal{X}$. □

This concludes the preparation for the proof of the main theorem of this section.

Theorem 4.3.14. *Let $k \geq 1$. For simple graphs G and H on at most n vertices, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of graphs of treewidth at most $k-1$,*
2. *G and H are homomorphism indistinguishable over the class of graphs of treewidth at most $k-1$ on at most $(2n^k)^{n^k+1}(k+1)$ vertices,*
3. *there exists a doubly stochastic $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $X\boldsymbol{B}_G = \boldsymbol{B}_H X$ for all $\boldsymbol{B} \in \mathcal{B}(k,k)$.*

Proof. In virtue of Lemma 4.3.6, we apply Theorem 4.3.11 to $\mathcal{TW}(k)$, which is agglutinatively generated by the involution monoid $\mathcal{PW}(k,k)$ with generating set $\mathcal{B}(k,k)$. Write $N := (2n^k)^{n^k+1}(k+1)$. Consider the following additional assertions:

4. G and H are homomorphism indistinguishable over the class of graphs of treewidth at most $k-1$ on at least k vertices,
5. G and H are homomorphism indistinguishable over the class of graphs of treewidth at most $k-1$ on at least k vertices and at most N vertices.

By Theorem 4.3.11 and Lemma 4.3.13, Items 3 to 5 are equivalent. By containment of the respective graph classes, Item 1 implies Item 2, which implies Item 5. By Lemma 4.3.7, Items 1 and 4 are equivalent. This closes a cycle of implications. □

4.3.3 Gluing-Closed Graph Classes

In preparation for Sections 4.3.4 and 6.4.3, we consider classes of labelled graphs satisfying the following property:

Definition 4.3.15. Let $k \geq 1$. A set $\mathcal{R} \subseteq \mathcal{G}(k)$ is *gluing-closed* if $\boldsymbol{R} \odot \boldsymbol{R}' \in \mathcal{R}$ for all $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{R}$.

Every agglutinatively generated graph class is gluing-closed. Other examples of gluing-closed graph classes arise from the Eilenberg–Moore categories of comonads from [57]. The following Theorem 4.3.16 characterises homomorphism indistinguishability over gluing-closed graph classes.

Theorem 4.3.16. *Let $k \geq 1$ and $\mathcal{R} \subseteq \mathcal{G}(k)$. Suppose that $\mathcal{R}$ is gluing-closed and contains $\mathbf{1}$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over $\operatorname{soe}(\mathcal{R})$,*
2. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G) \times V(H)}$ such that $X\boldsymbol{R}_H = \boldsymbol{R}_G$ for all $\boldsymbol{R} \in \mathcal{R}$,*
3. *there exists a doubly stochastic matrix $X \in \mathbb{Q}^{V(G) \times V(H)}$ such that $X\boldsymbol{R}_H = \boldsymbol{R}_G$ for all $\boldsymbol{R} \in \mathcal{R}$,*
4. *there exists a bijection $\pi \colon V(G)^k \to V(H)^k$ such that $\boldsymbol{R}_G(\boldsymbol{v}) = \boldsymbol{R}_H(\pi(\boldsymbol{v}))$ for all $\boldsymbol{v} \in V(G)^k$ and $\boldsymbol{R} \in \mathcal{R}$.*

Proof. The theorem follows from Lemma 4.1.11. To that end, take Γ to be $\mathcal{R}$. Since $\mathcal{R}$ is gluing-closed and contains $\mathbf{1}$, it forms a monoid under the gluing product. Take $I := V(G)^k$ and $J := V(H)^k$. The representations φ and ψ are given by $\boldsymbol{R} \mapsto \boldsymbol{R}_G$ and $\boldsymbol{R} \mapsto \boldsymbol{R}_H$, respectively. □

The bijection in Theorem 4.3.16 should be thought of as preserving the colouring of the tuples $\boldsymbol{v} \in V(G)^k$ and $\boldsymbol{w} \in V(H)^k$ by the values of $\boldsymbol{R}_G(\boldsymbol{v})$ for $\boldsymbol{R}_H(\boldsymbol{w})$ for $\boldsymbol{R} \in \mathcal{R}$, cf. Theorem 3.4.1.

4.3.4 Treedepth: Generation by Involution Monoids and Closure under Gluing

By Theorem 3.4.5, homomorphism indistinguishability over graphs of treedepth at most k corresponds to equivalence over the quantifier-rank-k fragment of first-order logic with counting quantifiers C_k. We extend this characterisation by proposing a linear system of equations very similar to the one for bounded pathwidth. In order to infer such a system of equations, we consider another type of interaction between labelled and bilabelled graphs. Let $\mathcal{S} \subseteq \mathcal{G}(k,k)$ be an involution monoid. Write $\mathcal{S}\mathbf{1} := \{\boldsymbol{S} \cdot \mathbf{1} \mid \boldsymbol{S} \in \mathcal{S}\} \subseteq \mathcal{G}(k)$. This section is concerned with the case when $\mathcal{S}\mathbf{1}$ is gluing-closed. If $\mathcal{S}\mathbf{1}$ is gluing-closed, then pseudo-stochastic solutions exist if, and only if, doubly stochastic solutions exist:

Theorem 4.3.17. *Let $k \geq 1$. Let $\mathcal{S} \subseteq \mathcal{G}(k,k)$ be an involution monoid generated by $\mathcal{B} \subseteq \mathcal{S}$. Suppose that $\mathcal{S}\mathbf{1}$ is gluing-closed. Let G and H be simple graphs on at most n vertices. Suppose that every graph in $\mathcal{B}$ has at most $b \in \mathbb{N} \cup \{\infty\}$ vertices and let $N := 2n^k b \in \mathbb{N} \cup \{\infty\}$. Then the following are equivalent:*

1. *G and H are homomorphism indistinguishable over $\operatorname{soe}(\mathcal{S})$,*
2. *G and H are homomorphism indistinguishable over $(\operatorname{soe}(\mathcal{S}))_{\leq N}$,*

3. *there exists a pseudo-stochastic matrix* $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ *such that* $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ *for all* $\boldsymbol{B} \in \mathcal{B}$,
4. *there exists a doubly stochastic matrix* $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ *such that* $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ *for all* $\boldsymbol{B} \in \mathcal{B}$.

Proof. By Theorem 4.3.2, Items 1 to 3 are equivalent. Let $\mathcal{X} \subseteq \mathcal{G}(k)$ denote the class agglutinatively generated by $\mathcal{S}$, cf. Definition 4.3.10. Since $\mathcal{S}\mathbf{1}$ is gluing-closed, it follows inductively that $\mathcal{X} \subseteq \mathcal{S}\mathbf{1}$. Conversely, $\mathcal{S}\mathbf{1} \subseteq \mathcal{X}$ since taking product of bilabelled graphs in $\mathcal{S}$ is one of the operations listed in Definition 4.3.10. Hence, $\mathcal{S}\mathbf{1} = \mathcal{X}$. Theorem 4.3.11 yields that Items 1 and 4 are equivalent. □

In order to apply Theorem 4.3.17, to the class of graphs of bounded treedepth, we define the following class of bilabelled graphs.

Definition 4.3.18. Let $k \geq 1$. The class $\mathcal{TD}(k,k)$ is the class of all (k,k)-bilabelled graph $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$ such that there exists a forest order $\leq$ on F satisfying

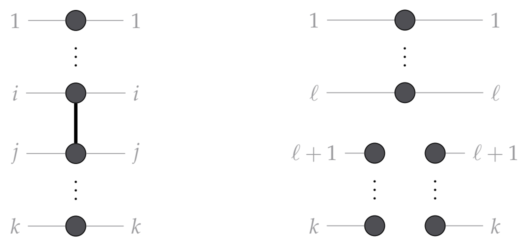
1. every edge $uv \in E(F)$ is such that $u \leq v$ or $v \leq u$,
2. for every leaf $x \in V(F)$ of $\leq$, the set $\{z \in V(F) \mid z \leq x\}$ has size k,
3. $\boldsymbol{u}$ and $\boldsymbol{v}$ are paths from a root to a leaf, i.e. $u_1 < u_2 < \cdots < u_k$ and $v_1 < v_2 < \cdots < v_k$.
4. the leaves u_k and v_k have the least number of common ancestors among all pairs of leaves of F, i.e. writing
$$\operatorname{ca}(x,y) := |\{z \in V(F) \mid z \leq x \land z \leq y\}| \tag{4.6}$$
for $x, y \in V(F)$, it holds that $\operatorname{ca}(x,y) \geq \operatorname{ca}(u_k, v_k)$ for all pairs of leaves $x, y \in V(F)$.

Items 1 and 2 ensure that the underlying unlabelled graph F is of treedepth at most k, cf. Section 2.1.4. Item 3 guarantees that this property is preserved under series composition. The remaining Item 4 helps to establish that $\mathcal{TD}(k,k)$ is finitely generated.

Lemma 4.3.19. *For* $k \geq 1$, $\mathcal{TD}(k,k)$ *is an involution monoid.*

Proof. Clearly, $\mathcal{TD}(k,k)$ is closed under taking reverses and contains $\boldsymbol{I}$. Given $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$ and $\boldsymbol{F}' = (F', \boldsymbol{u}', \boldsymbol{v}')$ with forest orders $\leq$ and $\leq'$, define a forest order $\leq''$ on $\boldsymbol{F} \cdot \boldsymbol{F}'$ by letting $x \leq'' y$ if $x, y \in V(F)$ and $x \leq y$ or if $x, y \in V(F')$ and $x \leq' y$. Since $\boldsymbol{v}$ and $\boldsymbol{u}'$ form paths from roots to leaves, this is a well-defined forest order. Also, in $\leq''$, $\boldsymbol{u}$ and $\boldsymbol{v}'$ are paths from roots to leaves.

It remains to check that Item 4 is satisfied. First observe that $\operatorname{ca}(u_k, v_k), \operatorname{ca}(u'_k, v'_k) \geq \operatorname{ca}(u_k, v'_k)$. Indeed, any common ancestor of u_k and v'_k must be an ancestor of v_k, which is identified with u'_k. Let z denote the maximal vertex such that $z \leq'' u_k, v'_k$. Note that $z \leq'' v_k, u'_k$. Every leaf x of $\boldsymbol{F}$ satisfies $z \leq x$. Indeed, if z and x were incomparable with respect to $\leq$, then $\operatorname{ca}(x, u_k) < \operatorname{ca}(u_k, v'_k)$ which contradicts the



(a) adjacency graph $\boldsymbol{A}^{ij}$. **(b)** join graph $\boldsymbol{J}^{\ell}$.

Figure 4.4: Bilabelled graphs in $\mathcal{TDB}(k,k)$ as defined in Lemma 4.3.21.

previous observation since $\operatorname{ca}(x, u_k) \geq \operatorname{ca}(u_k, v_k)$. The same argument applies to leaves of $\boldsymbol{F}'$. It follows that $\operatorname{ca}(x, y) \geq \operatorname{ca}(u_k, v'_k)$ for every pair of leaves x, y in $\boldsymbol{F} \cdot \boldsymbol{F}'$. $\square$

The following Lemma 4.3.20 establishes another assumption of Theorem 4.3.17:

Lemma 4.3.20. *For $k \geq 1$, $\mathcal{TD}(k,k)\mathbf{1}$ is gluing-closed.*

Proof. Given $\boldsymbol{F} = (F, \boldsymbol{u})$ and $\boldsymbol{F}' = (F', \boldsymbol{u}')$ with forest order $\leq$ and $\leq'$, define a forest order $\leq''$ on $\boldsymbol{F} \odot \boldsymbol{F}'$ by letting $x \leq'' y$ if $x, y \in V(F)$ and $x \leq y$ or $x, y \in V(F')$ and $x \leq' y$. Since $\boldsymbol{u}$ and $\boldsymbol{u}'$ form paths from roots to leaves, this is a well-defined forest order. The other conditions in Definition 4.3.18 are easily verified. $\square$

It remains to define generators for $\mathcal{TD}(k,k)$. The graphs featured in Lemma 4.3.21 are depicted by Figures 3.3b and 4.4. See Figure 4.5 for an example.

Lemma 4.3.21. *Let $k \geq 1$. Define $\mathcal{TDB}(k,k)$ as the set of the following (k,k)-bilabelled graphs:*

- *the* identity graph $\boldsymbol{I} = (I, (k), (k))$ *with* $V(I) = [k]$ *and* $E(I) = \emptyset$,
- *for* $i, j \in [k]$, *the* adjacency graphs $\boldsymbol{A}^{ij} = (A^{ij}, (k), (k))$ *with* $V(A^{ij}) = [k]$ *and* $E(A^{ij}) = \{ij\}$,
- *for* $0 \leq \ell < k$, *the* join graphs $\boldsymbol{J}^{\ell} = (J^{\ell}, (k), (1, \dots, \ell, (\ell+1)', \dots, k'))$ *with* $V(J^{\ell}) = \{1, \dots, k, (\ell+1)', \dots, k'\}$ *and* $E(J^{\ell}) = \emptyset$.

Then $\mathcal{TDB}(k,k)$ generates $\mathcal{TD}(k,k)$.

Proof. Clearly, $\mathcal{TDB}(k,k) \subseteq \mathcal{TD}(k,k)$ is closed under taking reverses and contains $\boldsymbol{I}$. For the second assertions, let $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{TD}(k,k)$. Let $\leq$ denote a forest order for F as in Definition 4.3.18. The proof is by induction on the number of leaves in $\leq$.

If there is only one leaf, then $\boldsymbol{F}$ is the series composition of the $\boldsymbol{A}^{ij}$ such that $i, j \in [k]$ and $u_i u_j \in E(F)$.

Now suppose that there are at least two leaves in $\leq$. Write X for the set of all leaves x of $\leq$ other than u_k such that $\operatorname{ca}(u_k, x)$ is maximal, i.e. such that $\operatorname{ca}(u_k, x) \geq \operatorname{ca}(u_k, y)$ for all leaves $y \neq u_k$. Let $x \in X$ be such that $\operatorname{ca}(x, v_k)$ is minimal, i.e.

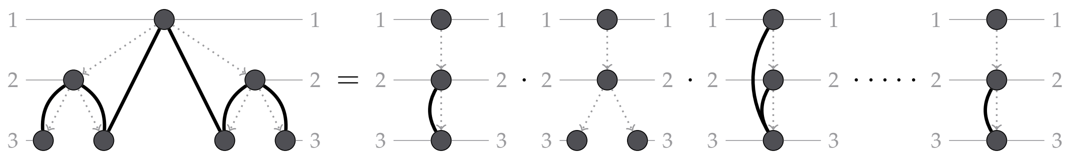


Figure 4.5: Example of a decomposition of bilabelled graph from $\mathcal{TD}(3,3)$ into generators from $\mathcal{TDB}(3,3)$. The grey dotted arrows $\rightarrow$ indicate the forest order $\leq$.

$\mathrm{ca}(x, v_k) \leq \mathrm{ca}(y, v_k)$ for all $y \in X$. Write D for the set of vertices $z \in V(F)$ such that $z \leq u_k$ and z and x are incomparable. Note that D forms a chain in $\leq$.

Let F' be the graph obtained from F by deleting all vertices in D. The forest order $\leq$ restricts to a forest order $\leq'$ of F'. Let F'' be the subgraph of F induced by the vertices $z \leq u_k$. Let $\boldsymbol{w} \in V(F')^k$ be the tuple satisfying $w_1 < w_2 < \cdots < w_k = x$. Define $\boldsymbol{F}' := (F', \boldsymbol{w}, \boldsymbol{v})$ and $\boldsymbol{F}'' := (F'', \boldsymbol{u}, \boldsymbol{u})$. Clearly, $\boldsymbol{F}'' \in \mathcal{TD}(k,k)$.

Towards showing that $\boldsymbol{F}' \in \mathcal{TD}(k,k)$, only Item 4 of Definition 4.3.18 is not straightforward. Distinguish cases.

- If $\mathrm{ca}(u_k, x) > \mathrm{ca}(u_k, v_k)$, then there is a vertex y such that y is a common ancestor of u_k and x but y is not an ancestor of v_k. This implies that every common ancestor of v_k and x is comparable with y and hence $\mathrm{ca}(x, v_k) = \mathrm{ca}(u_k, v_k)$. Hence, $\mathrm{ca}(a,b) \geq \mathrm{ca}(u_k, v_k) = \mathrm{ca}(x, v_k)$ for every pair of leaves a, b in F'.
- If $\mathrm{ca}(u_k, x) = \mathrm{ca}(u_k, v_k)$, then all leaves a of F' are in fact in X and the claim follows readily.

The bilabelled graphs $\boldsymbol{F}', \boldsymbol{F}'' \in \mathcal{TD}(k,k)$ have less leaves than $\boldsymbol{F}$. The claim follows inductively, observing that $\boldsymbol{F} = \boldsymbol{F}'' \cdot \boldsymbol{J}^\ell \cdot \boldsymbol{F}'$ for $0 \leq \ell < k$ minimal such that $w_{\ell+1} \not\in \{u_1, \dots, u_k\}$. ☐

The above observations yield the following Theorem 4.3.22.

Theorem 4.3.22. *Let $k \geq 1$. Let G and H be simple graphs on at most n vertices. Then the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of graphs of treedepth at most k,*
2. *G and H are homomorphism indistinguishable over the class of graphs of treedepth at most k on at most $4kn^k$ vertices,*
3. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{TDB}(k,k)$,*
4. *there exists a doubly stochastic matrix $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{TDB}(k,k)$.*

Proof. In virtue of Lemmas 4.3.19 to 4.3.21, we apply Theorem 4.3.17 to the gluing-closed class of labelled graphs $\mathcal{TD}(k,k)\mathbf{1}$ and the generating set $\mathcal{TDB}(k,k)$. Write $N := 4kn^k$. Consider the following additional assertions:

5. G and H are homomorphism indistinguishable over the class of graphs which admit a forest order satisfying Items 1 and 2 of Definition 4.3.18,
6. G and H are homomorphism indistinguishable over the class of graphs on at most N vertices which admit a forest order satisfying Items 1 and 2 of Definition 4.3.18.

Observe that, for every graph F admitting a forest order satisfying Items 1 and 2 of Definition 4.3.18, one can pick $\boldsymbol{u}, \boldsymbol{v} \in V(F)^k$ such that Items 3 and 4 of Definition 4.3.18 are satisfied as well. Hence, by Theorem 4.3.17, Items 3 to 6 are equivalent. By containment of the respective graph classes, Item 1 implies Item 2, which implies Item 6.

By adding isolated vertices, any graph of treedepth at most k can be turned into a graph with forest order satisfying Items 1 and 2 of Definition 4.3.18. Thus, by Lemma 4.3.7, Items 1 and 5 are equivalent. This closes a cycle of implications. □

4.3.5 Bounded Degree Trees: Inner-Product Compatibility

In this subsection, we turn to a class of labelled graphs which is endowed with less algebraic structure than the graph classes considered before. The class of suitably labelled trees of bounded degree trees is inner-product compatible. Inner-product compatibility is the most general property considered here, cf. Figure 4.2.

Definition 4.3.23. Let $k \geq 1$. A set $\mathcal{R} \subseteq \mathcal{G}(k)$ is *inner-product compatible* if, for all $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{R}$, there exists a labelled graph $\boldsymbol{R}'' \in \mathcal{R}$ such that $\langle \boldsymbol{R}, \boldsymbol{R}' \rangle = \operatorname{soe}(\boldsymbol{R}'')$.

Since $\langle \boldsymbol{R}, \boldsymbol{R}' \rangle = \operatorname{soe}(\boldsymbol{R} \odot \boldsymbol{R}')$ for all $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{G}(k)$, all gluing-closed classes of labelled graphs are also inner-product compatible. Similarly, if $\mathcal{R} = \mathcal{S}\mathbf{1}$ for some involution monoid $\mathcal{S} \subseteq \mathcal{G}(k,k)$, then $\mathcal{R}$ is inner-product compatible. This holds since $\langle \boldsymbol{S}_1\mathbf{1}, \boldsymbol{S}_2\mathbf{1} \rangle = \langle \mathbf{1}, \boldsymbol{S}_1^*\boldsymbol{S}_2\mathbf{1} \rangle = \operatorname{soe}(\boldsymbol{S}_1^*\boldsymbol{S}_2\mathbf{1})$ for $\boldsymbol{S}_1, \boldsymbol{S}_2 \in \mathcal{S}$.

In the following Theorem 4.3.24, we consider the spaces $\mathbb{Q}\mathcal{R}_G \leq \mathbb{Q}^{V(G)^k}$ and $\mathbb{Q}\mathcal{R}_H \leq \mathbb{Q}^{V(H)^k}$ spanned by the $\boldsymbol{R}_G$ and $\boldsymbol{R}_H$ respectively for $\boldsymbol{R} \in \mathcal{R} \subseteq \mathcal{G}(k)$.

Theorem 4.3.24. *Let $k \geq 1$. Let $\mathcal{R} \subseteq \mathcal{G}(k)$ be inner-product compatible containing $\mathbf{1}$. Then for simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over* $\operatorname{soe}(\mathcal{R})$,
2. *there exists an orthogonal pseudo-stochastic map $U \colon \mathbb{Q}\mathcal{R}_H \to \mathbb{Q}\mathcal{R}_G$ such that $U\boldsymbol{R}_H = \boldsymbol{R}_G$ for all $\boldsymbol{R} \in \mathcal{R}$,*
3. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $X\boldsymbol{R}_H = \boldsymbol{R}_G$ for all $\boldsymbol{R} \in \mathcal{R}$.*

Proof. Clearly, Item 3 implies Item 1. Assuming Item 1, we show that Item 2 holds. By inner-product compatibility of $\mathcal{R}$, for any $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{R}$ there exists $\boldsymbol{R}'' \in \mathcal{R}$ such that $\langle \boldsymbol{R}_G, \boldsymbol{R}'_G \rangle = \operatorname{soe}(\boldsymbol{R}''_G) = \operatorname{soe}(\boldsymbol{R}''_H) = \langle \boldsymbol{R}_H, \boldsymbol{R}'_H \rangle$. By Lemma 2.4.3, there exists an orthogonal map $U \colon \mathbb{Q}\mathcal{R}_H \to \mathbb{Q}\mathcal{R}_G$ such that $U\boldsymbol{R}_H = \boldsymbol{R}_G$ for all $\boldsymbol{R} \in \mathcal{R}$. In particular, $U\boldsymbol{1}_H = \boldsymbol{1}_G$ and $U^*\boldsymbol{1}_G = \boldsymbol{1}_H$. Hence, U is pseudo-stochastic and Item 2 holds.

For the remaining implication from Item 2 to Item 3, define $X \colon \mathbb{Q}^{V(G)^k} \to \mathbb{Q}^{V(H)^k}$ as the map coinciding with U on $\mathbb{Q}\mathcal{R}_H$ and annihilating $(\mathbb{Q}\mathcal{R}_H)^\perp$. Then $X\boldsymbol{R}_H = U\boldsymbol{R}_H = \boldsymbol{R}_G$ for all $\boldsymbol{R} \in \mathcal{R}$ and in particular $X\boldsymbol{1}_H = \boldsymbol{1}_G$. Furthermore, for all $v \in \mathbb{Q}\mathcal{R}_H$, $\langle v, X^*\boldsymbol{1}_G \rangle = \langle Uv, \boldsymbol{1}_G \rangle = \langle v, U^*\boldsymbol{1}_G \rangle = \langle v, \boldsymbol{1}_H \rangle$. For all $v \in (\mathbb{Q}\mathcal{R}_H)^\perp$, $\langle v, X^*\boldsymbol{1}_G \rangle = 0 = \langle v, \boldsymbol{1}_H \rangle$. Hence, $X^*\boldsymbol{1}_G = \boldsymbol{1}_H$. Thus, Item 3 holds. □

As noted above, all graph classes considered in previous sections enjoy properties stronger than inner-product compatibility. The class of bounded degree trees, however, does not possess any of the aforementioned properties.

Example 4.3.25. A *d-ary tree* is a tree whose vertices have degree at most $d+1$. For $d \geq 1$, the family of 1-labelled d-ary trees $\mathcal{T}^d$ with label at a vertex of degree at most one is inner-product compatible.

The set $\mathcal{T}^d$ is closed under *guarded Schur products*, i.e. under the d-ary operation

$$\begin{aligned} \circledast^d \colon \mathcal{G}(1) \times \cdots \times \mathcal{G}(1) &\to \mathcal{G}(1) \\ (\boldsymbol{R}^1, \dots, \boldsymbol{R}^d) &\mapsto \boldsymbol{A} \cdot (\boldsymbol{R}^1 \odot \cdots \odot \boldsymbol{R}^d) \end{aligned}$$

The operation $\circledast^d$ induces a d-ary multilinear map on $\mathbb{Q}^{V(G)}$ for every graph G, i.e. $\circledast^d_G(u_1, \dots, u_d) \coloneqq \boldsymbol{A}_G(u_1 \odot \cdots \odot u_d)$ for $u_1, \dots, u_d \in \mathbb{Q}^{V(G)}$.

Theorem 4.3.26. *Let $d \geq 1$. For graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of d-ary trees,*
2. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G) \times V(H)}$ such that $X\boldsymbol{T}_H = \boldsymbol{T}_G$ for all $\boldsymbol{T} \in \mathcal{T}^d$,*
3. *there exists a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G) \times V(H)}$ such that X preserves $\circledast^d$ on $\mathbb{Q}\mathcal{T}^d_H$, i.e. $X(\circledast^d_H(u_1, \dots, u_d)) = \circledast^d_G(Xu_1, \dots, Xu_d)$ for all $u_1, \dots, u_d \in \mathbb{Q}\mathcal{T}^d_H$.*

Proof. That Items 1 and 2 are equivalent follows directly from Theorem 4.3.24.

Assuming Item 2, let $X \in \mathbb{Q}^{V(G) \times V(H)}$ be pseudo-stochastic such that $X\boldsymbol{T}_H = \boldsymbol{T}_G$ for all $\boldsymbol{T} \in \mathcal{T}^d$. Then for all $\boldsymbol{T}^1, \dots, \boldsymbol{T}^d \in \mathcal{T}^d$,

$$\begin{aligned} X(\circledast^d_H(\boldsymbol{T}^1_H, \dots, \boldsymbol{T}^d_H)) &= X(\circledast^d(\boldsymbol{T}^1, \dots, \boldsymbol{T}^d))_H \\ &= (\circledast^d(\boldsymbol{T}^1, \dots, \boldsymbol{T}^d))_G \\ &= \circledast^d_G(\boldsymbol{T}^1_G, \dots, \boldsymbol{T}^d_G) \\ &= \circledast^d_G(X\boldsymbol{T}^1_H, \dots, X\boldsymbol{T}^d_H). \end{aligned}$$

Finally, Item 1 follows inductively from Item 3 observing that every $\boldsymbol{1} \neq \boldsymbol{T} \in \mathcal{T}^d$ can be written as $\boldsymbol{T} = \circledast^d(\boldsymbol{S}^1, \dots, \boldsymbol{S}^d)$ for some $\boldsymbol{S}^1, \dots, \boldsymbol{S}^d \in \mathcal{T}^d$ of lower depth. □

4.4 Comparison to Known Systems of Equations

In this section, we relate the matrix equations from Section 4.3 to the Sherali–Adams linear programming hierarchy for graph isomorphism as defined in Section 2.6.1. The goal is two-fold: Firstly, we confirm a conjecture of Dell, Grohe, & Rattan [61] by showing that feasibility of $\mathrm{SA}^k(G, H)$ over the rationals is equivalent to the homomorphism indistinguishability of G and H over the class of graphs of pathwidth at most $k-1$. Secondly, we establish that an ordered variant of $\mathrm{SA}^k(G, H)$ corresponds to homomorphism indistinguishability over the class of graphs of treedepth at most k.

For both instances, we establish an equality elimination result similar to [57, Theorem 32]. They showed that every C^k-sentence is equivalent to one without equality. We show that one of the systems $\mathrm{SA}^k(G, H)$ and $\mathrm{SA}^k_{\leftrightarrow}(G, H)$ is feasible over the (non-negative) rationals if, and only if, the other is. The systems $\mathrm{SA}^k(G, H)$ and $\mathrm{SA}^k_{\leftrightarrow}(G, H)$ differ in that $\mathrm{SA}^k_{\leftrightarrow}(G, H)$ does not impose the equality constraints of $\mathrm{SA}^k(G, H)$. Hence, the equality constraints from the classical Sherali–Adams system can be relaxed without curtailing its distinguishing power.

4.4.1 Sherali–Adams without Non-Negativity Constraints

We resolve the aforementioned conjecture by Dell, Grohe, & Rattan [61] by showing that the system of equations $\mathrm{SA}^k(G, H)$ is has a rational solution if, and only if, the system of equations $\mathrm{PW}^k(G, H)$ stated in Theorem 4.3.8 is feasible. The proof repeatedly makes use of the observation that the equations in $\mathrm{SA}^k(G, H)$ can be viewed as equations in $\mathrm{PW}^k(G, H)$ where certain (k, k)-bilabelled graphs model the continuity and compatibility equations of $\mathrm{SA}^k(G, H)$. Building on Theorem 4.3.8, we obtain the following Theorem 4.0.1. The systems $\mathrm{SA}^k(G, H)$ and $\mathrm{SA}^k_{\leftrightarrow}(G, H)$ are defined in Definition 2.6.7. The system $\mathrm{PW}^k(G, H)$ is defined in Theorem 4.3.8.

Theorem 4.0.1. *Let $k \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of graphs of pathwidth at most $k-1$,*
2. *$\mathrm{PW}^k(G, H)$ has a rational solution,*
3. *$\mathrm{SA}^k_{\leftrightarrow}(G, H)$ has a rational solution,*
4. *$\mathrm{SA}^k(G, H)$ has a rational solution.*

Proof. By Theorem 4.3.8, Items 1 and 2 are equivalent. Clearly, Item 4 implies Item 3.

We first show that Item 2 implies Item 3. Since X already denotes the solution to the system $\mathrm{PW}^k(G, H)$ in Theorem 4.3.8, the variables of $\mathrm{SA}^k(G, H)$ will be denoted by Y_π instead of X_π. Let $\mathfrak{S}_k$ denote the symmetric group acting on k letters. For a vector $\boldsymbol{v} \in V(G)^k$ and $\sigma \in \mathfrak{S}_k$, write $\sigma(\boldsymbol{v})$ for the vector $v_{\sigma(1)} \dots v_{\sigma(k)}$.

Claim 4.4.0a. Let $k \geq 2$. Let X denote a solution to $\mathrm{PW}^k(G, H)$. Then

$$X(\sigma(\boldsymbol{v}), \sigma(\boldsymbol{w})) = X(\boldsymbol{v}, \boldsymbol{w})$$

for all $\boldsymbol{v} \in V(G)^k$, $\boldsymbol{w} \in V(H)^k$, and $\sigma \in \mathfrak{S}_k$.

Proof of Claim. Recall the swap graph $\boldsymbol{S}^{ij}$ from Lemma 4.3.6. Let τ denote the transposition $(ij) \in \mathfrak{S}_k$. The equation $\boldsymbol{S}^{ij}_G X = X \boldsymbol{S}^{ij}_H$ of $\mathrm{PW}^k(G, H)$ is equivalent to $X(\tau(\boldsymbol{v}), \boldsymbol{w}) = X(\boldsymbol{v}, \tau(\boldsymbol{w}))$ for all $\boldsymbol{v} \in V(G)^k$ and $\boldsymbol{w} \in V(H)^k$. Hence, writing $\sigma \in \mathfrak{S}_k$ as product of transpositions $\sigma = \tau_1 \cdots \tau_r$, it holds that

$$\begin{aligned} X(\sigma(\boldsymbol{v}), \boldsymbol{w}) &= X(\tau_1 \cdots \tau_r(\boldsymbol{v}), \boldsymbol{w}) = X(\tau_2 \cdots \tau_r(\boldsymbol{v}), \tau_1(\boldsymbol{w})) \\ &= \cdots = X(\boldsymbol{v}, \tau_r \cdots \tau_1(\boldsymbol{w})) = X(\boldsymbol{v}, \sigma^{-1}(\boldsymbol{w})), \end{aligned}$$

as desired. ◁

The following Claim 4.4.0b shows that Equations (2.15) and (2.16) hold.

Claim 4.4.0b. Let $\ell \geq 1$. Let $X \in \mathbb{Q}^{V(G)^{\ell+1} \times V(G)^{\ell+1}}$ be a solution to $\mathrm{PW}^{\ell+1}(G, H)$. Then

$$\sum_{v' \in V(G)} X(\boldsymbol{v}v, \boldsymbol{w}w') = \sum_{w' \in V(H)} X(\boldsymbol{v}v', \boldsymbol{w}w) =: \check{X}(\boldsymbol{v}, \boldsymbol{w}) \tag{4.7}$$

for all $\boldsymbol{v} \in V(G)^\ell$, $\boldsymbol{w} \in V(H)^\ell$, $v \in V(G)$, and $w \in V(H)$. Furthermore, the matrix $\check{X} \in \mathbb{Q}^{V(G)^\ell \times V(G)^\ell}$ defined in Equation (4.7) is a solution to $\mathrm{PW}^\ell(G, H)$.

Proof of Claim. The first equality in Equation (4.7) is equivalent to $\boldsymbol{J}^{\ell+1}_G X = X \boldsymbol{J}^{\ell+1}_H$ where $\boldsymbol{J}^{\ell+1}$ denotes the forgetting graph from Lemma 4.3.6 and Figure 4.3b. Clearly, $\check{X}$ is pseudo-stochastic. It remains to argue that $\check{X}$ is a solution to $\mathrm{PW}^\ell(G, H)$.

Let $\boldsymbol{J}$ denote the $(1, 1)$-bilabelled edgeless 2-vertex graph whose labels reside on distinct vertices, cf. Figure 3.7a. For every $\boldsymbol{B} \in \mathcal{B}(\ell, \ell)$, the $(\ell + 1, \ell + 1)$-bilabelled graph $\boldsymbol{B} \otimes \boldsymbol{J}$ can be written as series composition of graphs in $\mathcal{B}(\ell + 1, \ell + 1)$. Hence, for arbitrary $w \in V(H)$ and $v \in V(G)$,

$$(\check{X}\boldsymbol{B}_H)(\boldsymbol{v}, \boldsymbol{w}) = (X(\boldsymbol{B}_H \otimes \boldsymbol{J}_H))(\boldsymbol{v}v, \boldsymbol{w}w) = ((\boldsymbol{B}_G \otimes \boldsymbol{J}_G)X)(\boldsymbol{v}v, \boldsymbol{w}w) = (\boldsymbol{B}_G\check{X})(\boldsymbol{v}, \boldsymbol{w}).$$

This concludes the proof. ◁

It remains to consider Equation (2.17). To that end, Claim 4.4.0c is shown. Note that the claim's assertion is well-defined by Claim 4.4.0a.

Claim 4.4.0c. Let $\ell \geq 1$. Let X be a solution to $\mathrm{PW}^\ell(G, H)$. Let $\boldsymbol{v} \in V(G)^\ell$ and $\boldsymbol{w} \in V(H)^\ell$. If $\pi = \{v_1w_1, \ldots, v_\ell w_\ell\}$ is not a local strong homomorphism, then $X(\boldsymbol{v}, \boldsymbol{w}) = 0$.

Proof of Claim. If π is not a local strong homomorphism, then there exists $i \neq j \in [\ell]$ such that without loss of generality $v_iv_j \in E(G)$ but $w_iw_j \notin E(H)$. By Observation 3.2.11, this implies that $\boldsymbol{A}^{ij}_G(\boldsymbol{v}, \boldsymbol{v}) = 1$ while $\boldsymbol{A}^{ij}_H(\boldsymbol{w}, \boldsymbol{w}) = 0$. Moreover, $X(\boldsymbol{v}, \boldsymbol{w})\boldsymbol{A}^{ij}_H(\boldsymbol{w}, \boldsymbol{w}) = (X\boldsymbol{A}^{ij}_H)(\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{A}^{ij}_G X)(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{A}^{ij}_G(\boldsymbol{v}, \boldsymbol{v})X(\boldsymbol{v}, \boldsymbol{w})$ since the in- and out-labels of $\boldsymbol{A}^{ij}$ coincide. Hence, $X(\boldsymbol{v}, \boldsymbol{w}) = 0$. ◁

This concludes the preparations for proving that Item 2 implies Item 3. Let X denote a solution of $\mathrm{PW}^k(G,H)$. Construct via Claim 4.4.0b solutions X^ℓ of $\mathrm{PW}^\ell(G,H)$ for $1 \leq \ell \leq k$ satisfying Equation (4.7). Define $Y_\emptyset := 1$ and, for every non-empty $\pi = \{v_1w_1, \dots, v_\ell w_\ell\} \in \binom{V(G)\times V(H)}{\leq k}$, let $Y_\pi := X^{|\pi|}(v_1 \dots v_\ell, w_1 \dots w_\ell)$. By Claim 4.4.0a, the entries of X do not depend on the ordering of the indices. Thus, Y is well-defined. Then Equations (2.15) and (2.16) hold by Equation (4.7). Equation (2.18) holds by Claim 4.4.0c, and Equation (2.17) by definition. Thus, Item 3 holds.

That Item 4 implies Item 1 was shown in [61, Theorem 4]. For the sake of completeness, we prove that Item 3 implies Item 2. To that end, let Y denote a solution of $\mathrm{SA}^k(G,H)$. Define a candidate solution X for $\mathrm{PW}^k(G,H)$ by letting $X(\boldsymbol{v},\boldsymbol{w}) := Y_{\{v_1w_1,\dots,v_kw_k\}}$ for $\boldsymbol{v} \in V(G)^k$ and $\boldsymbol{w} \in V(H)^k$. By repeatedly applying Equations (2.15) and (2.16), it follows that X is pseudo-stochastic. For commutation with the graphs from Lemma 4.3.6, three cases have to be considered: Let $\boldsymbol{v} \in V(G)^k$ and $\boldsymbol{w} \in V(H)^k$.

1. Let $\boldsymbol{A}^{ij}$ be an adjacency graph. If $(\boldsymbol{v},\boldsymbol{w})$ is a local strong homomorphism, then $\boldsymbol{A}^{ij}_G(\boldsymbol{v},\boldsymbol{v}) = \boldsymbol{A}^{ij}_H(\boldsymbol{w},\boldsymbol{w})$ and thus X and $\boldsymbol{A}^{ij}$ commute. If $(\boldsymbol{v},\boldsymbol{w})$ is not a local strong homomorphism, then $X(\boldsymbol{v},\boldsymbol{w}) = 0$ and the tensors commute as well.
2. Let $\boldsymbol{J}^\ell$ be a forgetting graph. To ease notation, suppose $\ell = k$. Then by Equations (2.15) and (2.16),

$$\begin{aligned}(X\boldsymbol{J}^k_H)(\boldsymbol{v},\boldsymbol{w}) &= \sum_{w\in V(H)} X(\boldsymbol{v}, w_1 \dots w_{k-1}w) = \sum_{w \in V(H)} Y_{\{v_1w_1,\dots,v_{k-1}w_{k-1},v_kw\}} \\ &\overset{(2.15)}{=} Y_{\{v_1w_1,\dots,v_{k-1}w_{k-1}\}} \overset{(2.16)}{=} \sum_{v\in V(G)} Y_{\{v_1w_1,\dots,v_{k-1}w_{k-1},vw_k\}} \\ &= \sum_{v \in V(G)} X(v_1 \dots v_{k-1}v, \boldsymbol{w}) = (\boldsymbol{J}^k_G X)(\boldsymbol{v},\boldsymbol{w}).\end{aligned}$$

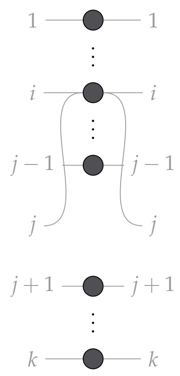
3. For the swap graphs, the statement follows as in Claim 4.4.0a since the value of X does not depend on the ordering of the indices.

This shows that Item 3 implies Item 2.

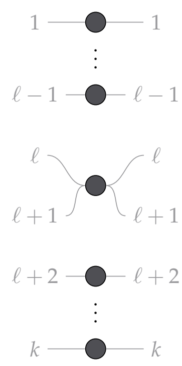
We establish the remaining implication by showing that Item 1 implies Item 4. To that end, we consider additional (k,k)-bilabelled graphs. For $i,j \in [k]$, $i \neq j$, the *identification graph* $\boldsymbol{I}^{ij}$ is obtained from $\boldsymbol{A}^{ij}$, cf. Lemma 4.3.6, by contracting the edge between vertices i and j, cf. Figure 4.6a. Write $\mathcal{B}'(k,k) := \mathcal{B}(k,k) \cup \{\boldsymbol{I}^{ij} \mid i,j \in [k], i \neq j\}$. Write $\mathcal{S}$ for the involution monoid generated by $\mathcal{B}'(k,k)$. By structural induction on $\boldsymbol{S} \in \mathcal{S}$, one may observe that, for every $\boldsymbol{S} = (S,\boldsymbol{u},\boldsymbol{v}) \in \mathcal{S}$, there exists a path decomposition (P,β) of S of width at most $k-1$ with vertices $u, v \in V(P)$ such that

1. $\beta(u) \supseteq \{u_1,\dots,u_k\}$ and $\beta(v) \supseteq \{v_1,\dots,v_k\}$, and
2. if $u \neq v$, then $\deg_P(u) = \deg_P(v) = 1$; if $u = v$, then $\deg_P(u) = \deg_P(v) = 0$.

In particular, all graphs in $\mathrm{soe}(\mathcal{S})$ are of pathwidth at most $k-1$. By Theorem 4.3.2, if G and H are homomorphism indistinguishable over all graphs of pathwidth



(a) $\boldsymbol{I}^{ij}$ from Theorem 4.0.1.



(b) $\boldsymbol{I}^{\ell}$ from Theorem 4.0.2.

Figure 4.6: Bilabelled identification graphs from Theorems 4.0.1 and 4.0.2.

at most $k-1$, then there exists a pseudo-stochastic $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{B}'(k,k)$. Supplying this matrix X to the argument that yielded that Item 2 implies Item 3 and adapting Claim 4.4.0c, shows that X can be transformed into a solution to $\mathrm{SA}^k(G, H)$. □

As a corollary, we show that $\mathrm{PW}^k(G, H)$ has a non-negative rational solution if, and only if, $\mathrm{SA}^k(G, H)$ has a non-negative rational solution. Consequently, the system of linear equations $\mathrm{PW}^k(G, H)$ has a non-negative rational solution if, and only if, G and H are homomorphism indistinguishable over graphs of treewidth at most k. Hence, the systems of equations $\mathrm{PW}^k(G, H)$, for $k \in \mathbb{N}$, form an alternative well-motivated hierarchy of linear programming relaxations of the graph isomorphism problem.

Corollary 4.4.1. *Let $k \geq 1$. Let G and H be two simple graphs. Then the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of graphs of treewidth at most $k-1$,*
2. *$\mathrm{PW}^k(G, H)$ has a non-negative rational solution,*
3. *$\mathrm{SA}^k_{\leftrightarrow}(G, H)$ has a non-negative rational solution,*
4. *$\mathrm{SA}^k(G, H)$ has a non-negative rational solution.*

Proof. The equivalence of Items 1 and 2 follows from Theorem 4.3.14. Theorem 2.6.8 in conjunction with Theorems 3.4.3 and 3.4.4 yields that Items 1 and 4 are equivalent. Alternatively, this follows by adapting the proof of Theorem 4.0.1 observing that the transformations devised in Claims 4.4.0a and 4.4.0b preserve non-negativity. The remaining implications follow as in the proof of Theorem 4.0.1. □

4.4.2 An Ordered Variant of Sherali–Adams

In this section, we reinterpret the systems of equations in Theorem 4.3.22 characterising homomorphism indistinguishability over graphs of bounded treedepth as an ordered variant of $\mathsf{SA}^k(G, H)$. In contrast to $\mathsf{SA}^k(G, H)$, it is equivalent for this system to have an arbitrary rational solution and a non-negative rational solution.

In line with Section 2.6.1, a pair $(\boldsymbol{v}, \boldsymbol{w}) \in V(G)^\ell \times V(H)^\ell$ for $\ell \geq 1$ is said to be a *local strong homomorphism* if $v_i v_j \in E(G) \Leftrightarrow w_i w_j \in E(H)$ for all $i, j \in [\ell]$. A pair $(\boldsymbol{v}, \boldsymbol{w}) \in V(G)^\ell \times V(H)^\ell$ is said to be a *local pseudo-isomorphism* if it is a local strong homomorphism and $v_i = v_{i+1} \Leftrightarrow w_i = w_{i+1}$ for all $i \in [\ell - 1]$.

Definition 4.4.2. Let $k \geq 1$. For simple graphs G and H, consider the system $\mathsf{TD}^k(G, H)$ with variables $X(\boldsymbol{v}, \boldsymbol{w})$ for every pair of tuples $\boldsymbol{v} \in V(G)^\ell$ and $\boldsymbol{w} \in V(H)^\ell$ for $0 \leq \ell \leq k$.

$$\sum_{v' \in V(G)} X(\boldsymbol{v}v', \boldsymbol{w}w) = X(\boldsymbol{v}, \boldsymbol{w}) \qquad \text{for all } w \in V(H) \text{ and } \boldsymbol{v} \in V(G)^\ell,\ \boldsymbol{w} \in V(H)^\ell \text{ where } 0 \leq \ell < k, \tag{4.8}$$

$$\sum_{w' \in V(H)} X(\boldsymbol{v}v, \boldsymbol{w}w') = X(\boldsymbol{v}, \boldsymbol{w}) \qquad \text{for all } v \in V(G) \text{ and } \boldsymbol{v} \in V(G)^\ell,\ \boldsymbol{w} \in V(H)^\ell \text{ where } 0 \leq \ell < k, \tag{4.9}$$

$$X((), ()) = 1, \tag{4.10}$$

$$X(\boldsymbol{v}, \boldsymbol{w}) = 0 \qquad \text{whenever } (\boldsymbol{v}, \boldsymbol{w}) \in V(G)^\ell \times V(H)^\ell \text{ for } 1 \leq \ell \leq k \text{ is not a local pseudo-isomorphism.} \tag{4.11}$$

Furthermore, consider the system $\mathsf{TD}^k_{\leftrightarrow}(G, H)$ with the same variables as $\mathsf{TD}^k(G, H)$ and Equations (4.8) to (4.10) and

$$X(\boldsymbol{v}, \boldsymbol{w}) = 0 \qquad \text{whenever } (\boldsymbol{v}, \boldsymbol{w}) \in V(G)^\ell \times V(H)^\ell \text{ for } 1 \leq \ell \leq k \text{ is not a local strong homomorphism.} \tag{4.12}$$

The main result of this section is the following:

Theorem 4.0.2. *Let $k \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the graphs of treedepth at most k,*
2. *$\mathsf{TD}^k(G, H)$ has a non-negative rational solution,*
3. *$\mathsf{TD}^k(G, H)$ has a rational solution,*
4. *$\mathsf{TD}^k_{\leftrightarrow}(G, H)$ has a non-negative rational solution,*
5. *$\mathsf{TD}^k_{\leftrightarrow}(G, H)$ has a rational solution.*

Proof. We first establish that Items 1, 4, and 5 are equivalent. To that end, given Theorem 4.3.22, it suffices to show that $\mathsf{TD}^k_{\leftrightarrow}(G, H)$ has a (non-negative) rational solution if, and only if, there is a (doubly stochastic) pseudo-stochastic matrix X such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{TDB}(k, k)$.

We first consider the implication from Item 1 to Item 4. Given a doubly stochastic $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{TDB}(k,k)$, observe that if $(\boldsymbol{v}, \boldsymbol{w}) \in V(G)^k \times V(H)^k$ is not a local strong homomorphism, then $X(\boldsymbol{v}, \boldsymbol{w}) = 0$. Indeed, in this case there exist $1 \leq i \neq j \leq k$ such that $w_i w_j \in E(F) \not\Leftrightarrow v_i v_j \in E(F)$. In particular, precisely one of $\boldsymbol{A}^{ij}_G(\boldsymbol{v}, \boldsymbol{v})$ and $\boldsymbol{A}^{ij}_H(\boldsymbol{w}, \boldsymbol{w})$ is zero. Then $\boldsymbol{A}^{ij}_G(\boldsymbol{v}, \boldsymbol{v}) X(\boldsymbol{v}, \boldsymbol{w}) = X(\boldsymbol{v}, \boldsymbol{w}) \boldsymbol{A}^{ij}_H(\boldsymbol{w}, \boldsymbol{w})$ implies that $X(\boldsymbol{v}, \boldsymbol{w}) = 0$. Hence, Equation (4.11) holds for $\ell = k$. With this observation at hand, define a solution to $\mathrm{TD}^k_{\leftrightarrow}(G, H)$ by invoking the following claim whose proof is analogous to the proof of Claim 4.4.0b.

Claim 4.4.2a. Let $\ell \geq 1$. If $X \in \mathbb{Q}^{V(G)^{\ell+1} \times V(H)^{\ell+1}}$ is doubly stochastic and such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{TDB}(\ell+1, \ell+1)$, then

$$\sum_{v' \in V(G)} X(\boldsymbol{v}v, \boldsymbol{w}w') = \sum_{w' \in V(H)} X(\boldsymbol{v}v', \boldsymbol{w}w) =: \check{X}(\boldsymbol{v}, \boldsymbol{w})$$

for all $v \in V(G)$, $w \in V(H)$, and $\boldsymbol{v} \in V(G)^\ell$, $\boldsymbol{w} \in V(H)^\ell$. Furthermore, the matrix $\check{X} \in \mathbb{Q}^{V(G)^\ell \times V(H)^\ell}$ is doubly stochastic and such that $\boldsymbol{B}_G \check{X} = \check{X} \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{TDB}(\ell, \ell)$.

A solution to $\mathrm{TD}^k_{\leftrightarrow}(G, H)$ can now be defined inductively invoking Claim 4.4.2a. Equations (4.8) and (4.9) are immediate from Claim 4.4.2a. Equation (4.10) follows since double stochasticity is preserved throughout the induction. Equation (4.11) follows as observed initially.

For the implication from Item 5 to Item 1, define a pseudo-stochastic matrix $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ by extracting the values on $V(G)^k \times V(H)^k$ from a rational solution Y to $\mathrm{TD}^k_{\leftrightarrow}(G, H)$. By Equations (4.8) to (4.10), X is pseudo-stochastic. It remains to consider commutation with the graphs from Lemma 4.3.21. To that end, let $(\boldsymbol{v}, \boldsymbol{w}) \in V(G)^k \times V(H)^k$.

1. If $(\boldsymbol{v}, \boldsymbol{w})$ is a local strong homomorphism, then $\boldsymbol{A}^{ij}_G(\boldsymbol{v}, \boldsymbol{v}) = \boldsymbol{A}^{ij}_H(\boldsymbol{w}, \boldsymbol{w})$ for all $1 \leq i \neq j \leq k$. Hence, $\boldsymbol{A}^{ij}_G(\boldsymbol{v}, \boldsymbol{v}) X(\boldsymbol{v}, \boldsymbol{w}) = X(\boldsymbol{v}, \boldsymbol{w}) \boldsymbol{A}^{ij}_H(\boldsymbol{w}, \boldsymbol{w})$. If $(\boldsymbol{v}, \boldsymbol{w})$ is not a local strong homomorphism, then the same assertions follows readily from Equation (4.11).
2. For $0 \leq \ell < k$, by repeatedly applying Equations (4.8) and (4.9),

$$\begin{aligned}(X \boldsymbol{J}^\ell_H)(\boldsymbol{v}, \boldsymbol{w}) &= \sum_{w_{\ell+1}, \dots, w_k \in V(H)} X(\boldsymbol{v}, w_1 \dots w_\ell w_{\ell+1} \dots w_k) \\ &= Y(v_1 \dots v_\ell, w_1 \dots w_\ell) \\ &= (\boldsymbol{J}^\ell_G X)(\boldsymbol{v}, \boldsymbol{w}).\end{aligned}$$

Since the system $\mathrm{TD}^k(G, H)$ is more restrictive than the system $\mathrm{TD}^k_{\leftrightarrow}(G, H)$, Item 2 implies Items 3 and 4, which respectively imply Item 5. Thus it remains to argue that Item 1 implies Item 2. To that end, consider the following additional (k,k)-bilabelled graphs. For $\ell \in [k-1]$, the *identification graph* $\boldsymbol{I}^\ell$ is obtained from $\boldsymbol{A}^{\ell, \ell+1}$ by contradicting the edge between the vertices labelled ℓ and $\ell+1$, cf. Figure 4.6b.

Write $\mathcal{TDB}'(k,k) := \mathcal{TDB}(k,k) \cup \{\boldsymbol{I}^\ell \mid \ell \in [k-1]\}$ and $\mathcal{S}$ for the involution monoid generated by $\mathcal{TDB}'(k,k)$. By structural induction on the elements of $\mathcal{S}$, it is easy to see that every $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{S}$ is such that there exists a forest cover $\leq$ of F of depth at most k such that $u_1 \leq u_2 \leq \cdots \leq u_k$ and $v_1 \leq v_2 \leq \cdots \leq v_k$. In particular, every graph in $\operatorname{soe}(\mathcal{S})$ has treedepth at most k. Furthermore, $\mathcal{S}\mathbf{1}$ is gluing-closed. By Theorem 4.3.17, assuming Item 1, there exists a doubly stochastic $X \in \mathbb{Q}^{V(G)^k \times V(H)^k}$ such that $\boldsymbol{B}_G X = X \boldsymbol{B}_H$ for all $\boldsymbol{B} \in \mathcal{TDB}'(k,k)$. As above, it follows that a solution of $\mathrm{TD}^k_{\leftrightarrow}(G,H)$ can be constructed from X. This solution satisfies additionally Equation (4.11). □

4.5 Matrix Equations from Augmented Homomorphism Tensors

In this section, we construct a system of equations whose feasibility characterises homomorphism indistinguishability over the class $\mathcal{T}^k_q$ of graphs which admit a k-pebble forest cover of depth $\leq d$, cf. Definition 2.1.3. By Theorem 3.4.6, two graphs are homomorphism indistinguishable over $\mathcal{T}^k_q$ if, and only if, they are C^k_q-equivalent.

The main difficulty arising when applying the strategy utilised in Sections 4.2 and 4.3 to $\mathcal{T}^k_q$ is that there is no obvious way to define labelling and operations such that the resulting class of labelled graphs is closed under these. We overcome this obstacle by promoting the labels from mere distinguished vertices to objects encoding the role the labelled vertices play in the associated pebble forest cover. Subsequently, the gluing operation can be restricted to pairs of labelled graphs whose labels play the same role. While this resolves the combinatorial problems regarding labelling, operations, and finite generation, alterations have to be made also on the algebraic side. For this purpose, we introduce augmented homomorphism tensors of labelled graphs which not only encode the homomorphism counts but also the role of the labelled vertex in the decomposition.

Although this approach might appear to be tailored to $\mathcal{T}^k_q$, it is in fact inspired by categorical principles laid out in [2, 4, 57]. Their framework of comonads on the category of relational structures has given rise to a categorical language for capturing natural graph classes and decompositions leading moreover to results in homomorphism indistinguishability [57, 129, 3]. We introduce bilabelled graphs augmented by additional information accompanied by corresponding representations and operations. In [146, Appendix C], we argue that this approach can be viewed as an instantiation of a comonadic strategy for homomorphism indistinguishability in the case of the pebbling comonad [2, 57].

The material in this section is based on joint work with Gaurav Rattan and has first been published in [145]. The main result of this paper [145, Theorem 1.3] introduces a linear system of equations whose feasibility characterises indistinguishability of two graphs under q iterations of the k-dimensional Weisfeiler–Leman algorithm, for

$k \geq 1$, $q \geq 0$. In this section, we prove a similar result characterising C^k_q-equivalence in terms of the feasibility of a linear system of equations. Although being much alike, C^k_q-equivalence and indistinguishability after q iterations of wl_k are not quite the same.[6] For the purpose of this thesis, we choose to cover only C^k_q-equivalence. Thereby, Theorem 3.4.6 can be used as a black box without deviating from the main subject of this thesis.

First, we deviate from Section 4.3 by augmenting the labels with information about the role they play in a fixed pebble forest cover. Recall the definition of $\mathcal{T}^k_q$ from Definition 2.1.3. As observed in Section 2.1.4, every graph in $\mathcal{T}^k_q$ has treedepth at most q. Thus, when introducing bilabelled graphs whose underlying unlabelled graphs are those in $\mathcal{T}^k_q$ in the following Definition 4.5.1, we build on the definition of $\mathcal{TD}(q,q)$ from Definition 4.3.18.

Definition 4.5.1. Let $k, q \geq 1$. Let $\widehat{\mathcal{T}}^k(q,q)$ denote the set of all tuples $\widehat{\boldsymbol{F}} = (\boldsymbol{F}, p_{\text{in}}, p_{\text{out}})$ such that $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{TD}(q,q)$ is a (q,q)-bilabelled graph and there exists a k-pebble forest cover $\leq$ for F with pebbling function $p\colon V(F) \to [k]$ such that $p_{\text{in}}, p_{\text{out}}\colon [q] \to [k]$ are functions such that $p_{\text{in}}(i) = p(u_i)$ and $p_{\text{out}}(i) = p(v_i)$ for all $i \in [q]$. The set $\widehat{\mathcal{T}}^k(q,q)$ furthermore contains the symbol $\perp$.

Recording not only the graph and its in- and out-labels but also the value of the pebbling function at the labels allows us to introduce well-defined operations in the next section. The symbol $\perp$ will be used to define the result of an operation which is not permitted. We first augment the homomorphism tensors from Definition 3.2.8 by information about the pebbling functions.

Definition 4.5.2. Let $k, q \geq 1$ and $\widehat{\boldsymbol{F}} = (\boldsymbol{F}, p_{\text{in}}, p_{\text{out}}) \in \widehat{\mathcal{T}}^k(q,q)$. For a graph G, define the *augmented homomorphism tensor* $\widehat{\boldsymbol{F}}_G \in \mathbb{N}^{V(G)^k \times V(G)^\ell}$ of $\boldsymbol{F}$ with respect to G by letting $\widehat{\boldsymbol{F}}_G \coloneqq \boldsymbol{F}_G \otimes e_{p_{\text{in}}} e^T_{p_{\text{out}}} \in \mathbb{N}^{V(G)^q \times V(G)^q} \otimes \mathbb{N}^{[k]^{[q]} \times [k]^{[q]}}$. The augmented homomorphism tensor of $\perp$ with respect to any graph is the zero tensor.

Here, $e_p \in \mathbb{N}^{[k]^{[q]}}$ for $p\colon [q] \to [k]$ denotes the vector which is 1 at the entry indexed by p and 0 at all other entries. If G has n vertices, then $\widehat{\boldsymbol{F}}_G$ has dimension $(nk)^{2q}$. In the application scenario envisaged in [145], k and q are considered to be constants. Under these assumptions, the augmented homomorphism tensors have polynomial dimension. Given Definitions 4.5.1 and 4.5.2, we can formally state the theorem which we desire to prove in this section. It provides a linear system of equations whose feasibility characterises C^k_q-equivalence.

Theorem 4.5.3. *Let $k, q \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are C^k_q-equivalent,*
2. *G and H are homomorphism indistinguishable over $\mathcal{T}^k_q$,*

[6] After one iteration of wl_k, any two non-isomorphic graphs on at most k vertices are distinguished. In C^k_1, one can express only unary properties and e.g. $2K_1$ and K_2 are not distinguished.

3. *there exists a pseudo-stochastic* $X \in \mathbb{Q}^{V(G)^q \times V(H)^q} \otimes \mathbb{Q}^{[k]^{[q]} \times [k]^{[q]}}$ *such that* $\widehat{\boldsymbol{B}}_G X = X \widehat{\boldsymbol{B}}_H$ *for all* $\widehat{\boldsymbol{B}} \in \widehat{\mathcal{B}}^k(q,q)$,
4. *there exists a doubly stochastic* $X \in \mathbb{Q}^{V(G)^q \times V(H)^q} \otimes \mathbb{Q}^{[k]^{[q]} \times [k]^{[q]}}$ *such that* $\widehat{\boldsymbol{B}}_G X = X \widehat{\boldsymbol{B}}_H$ *for all* $\widehat{\boldsymbol{B}} \in \widehat{\mathcal{B}}^k(q,q)$.

The second step towards Theorem 4.5.3 is to define combinatorial operations on $\widehat{\mathcal{T}}^k(q,q)$ and accompanying algebraic operations respecting the augmented homomorphism representation. Due to the correspondence between these operations, which is established in Lemma 4.5.6, we abusively use the same notation for both. The following combinatorial operations are adaptions of operations on bilabelled graphs from Definition 3.2.6.

Definition 4.5.4. Let $k, q \geq 1$. Let $\widehat{\boldsymbol{F}} = (\boldsymbol{F}, p_{\text{in}}, p_{\text{out}}), \widehat{\boldsymbol{F}}' = (\boldsymbol{F}', p'_{\text{in}}, p'_{\text{out}}) \in \widehat{\mathcal{T}}^k(q,q)$.

1. The *unlabelling* of $\widehat{\boldsymbol{F}} \neq \bot$, denoted by $\operatorname{soe}(\widehat{\boldsymbol{F}})$, is the underlying unlabelled graph of $\boldsymbol{F}$. The unlabelling of $\bot$ is undefined.
2. The *reverse* of $\widehat{\boldsymbol{F}} := (\boldsymbol{F}, p_{\text{in}}, p_{\text{out}})$ is $\widehat{\boldsymbol{F}}^* = (\boldsymbol{F}^*, p_{\text{out}}, p_{\text{in}})$. The reverse of $\bot$ is $\bot$.
3. The *series composition* of $\widehat{\boldsymbol{F}}$ and $\widehat{\boldsymbol{F}}'$ is defined to be $\bot$ if $p_{\text{out}} \neq p'_{\text{in}}$ and otherwise to be $(\boldsymbol{F} \cdot \boldsymbol{F}', p_{\text{in}}, p'_{\text{out}})$. The series composition of $\bot$ with any element of $\widehat{\mathcal{T}}^k(q,q)$ on either side is $\bot$.

Unlabelling establishes a connection to unlabelled graphs. Reversal is needed for algebraic purposes, cf. Theorem 4.1.4, while series composition is the operation under which finite generation will be proven in Lemma 4.5.8. Recall the definition of $\mathcal{T}^k_q$ from Definition 2.1.3.

Lemma 4.5.5. *Let* $k, q \geq 1$. $\widehat{\mathcal{T}}^k(q,q)$ *is closed under reversal and series composition. The graphs obtained from elements of* $\widehat{\mathcal{T}}^k(q,q) \setminus \{\bot\}$ *by unlabelling are in* $\mathcal{T}^k_q$.

Proof. The closure of $\widehat{\mathcal{T}}^k(q,q)$ under reversal and the assertion that graphs obtained from elements of $\widehat{\mathcal{T}}^k(q,q)$ by unlabelling are in $\mathcal{T}^k_q$ follow immediately from Definition 4.5.1.

It remains to show that $\widehat{\mathcal{T}}^k(q,q)$ is closed under series composition. To that end, let $\widehat{\boldsymbol{F}} = (\boldsymbol{F}, p_{\text{in}}, p_{\text{out}})$ and $\widehat{\boldsymbol{F}}' = (\boldsymbol{F}', p'_{\text{in}}, p'_{\text{out}})$ be elements of $\widehat{\mathcal{T}}^k(q,q)$, without loss of generality different from $\bot$. If $p_{\text{out}} \neq p'_{\text{in}}$, then the result of the series composition is $\bot \in \widehat{\mathcal{T}}^k(q,q)$.

Otherwise, it follows from Lemma 4.3.19 that $\boldsymbol{F} \cdot \boldsymbol{F}' \in \mathcal{TD}(k,k)$. In the proof of Lemma 4.3.19 a forest cover of $\boldsymbol{F} \cdot \boldsymbol{F}'$ is constructed by taking the disjoint union of the forest covers of $\boldsymbol{F}$ and $\boldsymbol{F}'$ and identifying the out-labelled vertices of $\boldsymbol{F}$ with the in-labelled vertices of $\boldsymbol{F}'$. Since $p_{\text{out}} = p'_{\text{in}}$, a suitable pebbling function for this forest cover can be obtained in the same way from the pebbling functions of $\boldsymbol{F}$ and $\boldsymbol{F}'$. □

The following Lemma 4.5.6 extends Lemma 3.2.14 to augmented homomorphism tensors.

Lemma 4.5.6. *Let* $\widehat{\boldsymbol{F}} = (\boldsymbol{F}, p_{\text{in}}, p_{\text{out}}), \widehat{\boldsymbol{F}}' = (\boldsymbol{F}', p'_{\text{in}}, p'_{\text{out}}) \in \widehat{\mathcal{T}}^k(q,q)$*. Let G be a graph.*

1. *Unlabelling corresponds to sum-of-entries, i.e.* $\hom(\operatorname{soe}(\widehat{\boldsymbol{F}}), G) = \operatorname{soe}(\widehat{\boldsymbol{F}}_G)$ *when* $\widehat{\boldsymbol{F}} \neq \bot$.
2. *Reversal corresponds to transposition, i.e.* $(\widehat{\boldsymbol{F}}^*)_G = (\widehat{\boldsymbol{F}}_G)^*$.
3. *Series composition corresponds to matrix product, i.e.* $(\widehat{\boldsymbol{F}} \cdot \widehat{\boldsymbol{F}}')_G = \widehat{\boldsymbol{F}}_G \cdot \widehat{\boldsymbol{F}}'_G$.

Proof. The correspondence between reversal and transposition is purely syntactical. For the correspondence of unlabelling and sum-of-entries, recall Lemma 3.2.14 and observe that

$$\begin{aligned}
\hom(\operatorname{soe}(\widehat{\boldsymbol{F}}), G) &= \hom(\operatorname{soe}(\boldsymbol{F}), G) \\
&= \sum_{\boldsymbol{v},\boldsymbol{v}' \in V(G)^q} \boldsymbol{F}_G(\boldsymbol{v}, \boldsymbol{v}') \\
&= \sum_{\substack{\boldsymbol{v},\boldsymbol{v}' \in V(G)^q, \\ p,p' \colon [q] \to [k]}} \boldsymbol{F}_G(\boldsymbol{v}, \boldsymbol{v}') e_p^T e_{p_{\text{in}}} e_{p_{\text{out}}}^T e_{p'} \\
&= \operatorname{soe}(\widehat{\boldsymbol{F}}_G).
\end{aligned}$$

Finally, consider series composition and matrix product. If one of the factors is $\bot$, whose augmented homomorphism tensor is the zero tensor, the statement is readily verified. Otherwise, let $\boldsymbol{u}, \boldsymbol{u}' \in V(G)^q$ and $p, p' \colon [q] \to [k]$ be arbitrary.

$$\begin{aligned}
(\widehat{\boldsymbol{F}}_G \cdot \widehat{\boldsymbol{F}}'_G)(\boldsymbol{u}, p; \boldsymbol{u}, p') &= \sum_{\substack{\boldsymbol{v} \in V(G)^q, \\ r \colon [q] \to [k]}} \widehat{\boldsymbol{F}}_G(\boldsymbol{u}, p; \boldsymbol{v}, r) \widehat{\boldsymbol{F}}'_G(\boldsymbol{v}, r; \boldsymbol{u}', p') \\
&= \sum_{\substack{\boldsymbol{v} \in V(G)^q, \\ r \colon [q] \to [k]}} \boldsymbol{F}_G(\boldsymbol{u}, \boldsymbol{v}) \boldsymbol{F}'_G(\boldsymbol{v}, \boldsymbol{u}') \delta_{p = p_{\text{in}}} \delta_{p_{\text{out}} = r = p'_{\text{in}}} \delta_{p'_{\text{out}} = p'} \\
&= \sum_{\boldsymbol{v} \in V(G)^q} \boldsymbol{F}_G(\boldsymbol{u}, \boldsymbol{v}) \boldsymbol{F}'_G(\boldsymbol{v}, \boldsymbol{u}') \delta_{p = p_{\text{in}}} \delta_{p_{\text{out}} = p'_{\text{in}}} \delta_{p'_{\text{out}} = p'} \\
&= \begin{cases} (\boldsymbol{F} \cdot \boldsymbol{F}')_G(\boldsymbol{u}, \boldsymbol{u}') \delta_{p = p_{\text{in}}} \delta_{p'_{\text{out}} = p'}, & \text{if } p_{\text{out}} = p'_{\text{in}}, \\ 0, & \text{otherwise}, \end{cases} \\
&= (\widehat{\boldsymbol{F}} \cdot \widehat{\boldsymbol{F}}')_G(\boldsymbol{u}, p; \boldsymbol{u}, p').
\end{aligned}$$

The last equality holds by Lemma 3.2.14 and Definition 4.5.4. □

To prove that $\widehat{\mathcal{T}}^k(q,q)$ is finitely generated under series composition, we define the set $\widehat{\mathcal{B}}^k(q,q)$ extending $\mathcal{TDB}(q,q)$ from Lemma 4.3.21.

Definition 4.5.7. Let $k, q \geq 1$. The set $\widehat{\mathcal{B}}^k(q,q)$ is the subset of $\widehat{\mathcal{T}}^k(q,q)$ containing the following elements:

- the *identity graph* $(\boldsymbol{I}, p, p)$, where $\boldsymbol{I} \in \mathcal{TDB}(q,q)$ for every $p \colon [q] \to [k]$.
- the *adjacency graph* $(\boldsymbol{A}^{ij}, p, p)$ for $1 \leq i, j \leq q$, where $\boldsymbol{A}^{ij} \in \mathcal{TDB}(q,q)$ for every $p \colon [q] \to [k]$ such that $p(i) \neq p(\ell)$ for all $i < \ell \leq j$.

- the *join graph* $(\boldsymbol{J}^\ell, p_{\text{in}}, p_{\text{out}})$ for $0 \leq \ell < q$, where $\boldsymbol{J}^\ell \in \mathcal{TDB}(q,q)$ for every $p_{\text{in}}, p_{\text{out}} \colon [q] \to [k]$ such that $p_{\text{in}}|_{[\ell]} = p_{\text{out}}|_{[\ell]}$.

Observe that $\boldsymbol{I}$, $\boldsymbol{A}^{ij}$, and $\boldsymbol{J}^{ij}$ admit forest covers compatible with all stipulated pebbling functions.

Lemma 4.5.8. *Let $k, q \geq 1$. For every $\widehat{\boldsymbol{F}} \in \widehat{\mathcal{T}}^k(q,q)$, there exist $\widehat{\boldsymbol{B}}^1, \dots, \widehat{\boldsymbol{B}}^r \in \widehat{\mathcal{B}}^k(q,q)$ for some $r \in \mathbb{N}$ such that $\widehat{\boldsymbol{F}}$ is the series composition of $\widehat{\boldsymbol{B}}^1, \dots, \widehat{\boldsymbol{B}}^r$, i.e. $\widehat{\boldsymbol{F}} = \widehat{\boldsymbol{B}}^1 \cdot \dots \cdot \widehat{\boldsymbol{B}}^r$.*

Proof. Let $\widehat{\boldsymbol{F}} = (\boldsymbol{F}, p_{\text{in}}, p_{\text{out}})$. By Lemma 4.3.21, there exist $\boldsymbol{B}^1, \dots, \boldsymbol{B}^r \in \mathcal{TDB}(q,q)$ such that $\boldsymbol{F} = \boldsymbol{B}^1 \cdot \dots \cdot \boldsymbol{B}^r$. Each of the $\operatorname{soe}(\boldsymbol{B}^i)$ can be regarded as subgraph of $\operatorname{soe} \boldsymbol{F}$. Thus, the $\boldsymbol{B}^i$ inherit pebbling functions from the pebbling function of $\widehat{\boldsymbol{F}}$. □

As a final step, we apply Theorems 4.1.4 and 4.1.10 to the augmented homomorphism tensors of the elements of $\widehat{\mathcal{T}}^k(q,q)$.

Proof of Theorem 4.5.3. The equivalence of Items 1 and 2 follows from Theorem 3.4.6. The implication from Item 2 to Items 3 and 4 follows from Theorems 4.1.4 and 4.1.10, respectively, combined with Lemmas 4.5.5, 4.5.6, and 4.5.8. See also the proof of Theorem 4.3.17.

Clearly, Item 4 implies Item 3. For the remaining implication from Item 3 to Item 2 observe as in the proof of Theorem 4.3.22 that the graphs in $\operatorname{soe}(\widehat{\mathcal{T}}^k(q,q))$ are precisely those which admit a k-pebble forest covers all whose maximal totally ordered subsets have size exactly q. By invoking Lemma 4.3.7, it follows that homomorphism indistinguishability over this graph class is equivalent to homomorphism indistinguishability over $\mathcal{T}_q^k$. □

5 Homomorphism Indistinguishability from Matrix Equations

In Chapter 4, we demonstrated how to characterise a given homomorphism indistinguishability relation as feasibility of some matrix equation. For example, starting from the class of graphs of bounded treedepth, we constructed a system of linear equations in Definition 4.4.2, which is feasible for two graphs if, and only if, they are homomorphism indistinguishable over this graph class. In this chapter, we take the opposite approach: We start with a well-known system of equations, namely the level-t Lasserre relaxation $\mathrm{L}^t(G, H)$ of the graph isomorphism quadratic program $\mathrm{QP}(G, H)$, cf. Definition 2.6.9, and construct a graph class $\mathcal{L}_t$ such that G and H are homomorphism indistinguishable over $\mathcal{L}_t$ if, and only if, $\mathrm{L}^t(G, H)$ is feasible. Remarkably, our characterisation of feasibility of $\mathrm{L}^t(G, H)$ as homomorphism indistinguishable relations allows to precisely measure the Lasserre hierarchy's distinguishing power compared to the Sherali–Adams hierarchy.

In [43, 138, 23], it was shown that only the level-$\Omega(n)$ Lasserre system of equations can distinguish all non-isomorphic n-vertex graphs. In general and not only for graph isomorphism, feasibility of the level-t Lasserre relaxation of an integer program implies feasibility of its level-t Sherali–Adams relaxation [107]. For graph isomorphism, it was shown by Atserias & Fijalkow [13] that the converse holds up to multiplicative offset in the number of levels. Thus, perhaps surprisingly, the Lasserre hierarchy is not more powerful than the Sherali–Adams hierarchies when applied to graph isomorphism. More precisely, by [13, Corollary 6.7], there exists a constant c such that if the level-ct Sherali–Adams relaxations $\mathrm{SA}^{ct}(G, H)$ is feasible over the non-negative reals for two graphs G and H, then the level-t Lasserre relaxation $\mathrm{L}^t(G, H)$ is feasible over the reals. However, this constant c is not explicit in [13] and depends on the implementation details of an algorithm developed in that paper. This chapter's main result asserts that c can be taken to be three and that this constant is best possible.[7]

[7]The constant c in [13, Theorem 6.3] depends on the implementation details of the algorithm that yields [13, Corollary 5.1], which in turn depends on the precise version of the Lasserre system of equations used there. As discussed in [152, Appendix A], our Lasserre system of equations is defined slightly differently. In Theorem 5.0.1, we abstract from these details by proving a statement that involves only the feasibility of $\mathrm{L}^t(G, H)$ and $\mathrm{SA}^t(G, H)$. Since our Lasserre formulation and the one in [13] distinguish the same pairs of graphs [152, Lemma A.1], Theorem 5.0.1 yields that c in [13, Theorem 6.3] can be taken to be three (and that this is best possible). Theorem 5.0.1 does not imply bounds on the complexity of the algorithm yielding [13, Corollary 5.1].

Theorem 5.0.1. *For simple graphs G and H and every $t \geq 1$, consider the following assertions:*

1. $\mathrm{SA}^{3t}(G, H)$ *has a non-negative rational solution,*
2. $\mathrm{L}^{t}(G, H)$ *has a real solution,*
3. $\mathrm{SA}^{t}(G, H)$ *has a non-negative rational solution.*

The implications 1 ⇒ 2 ⇒ 3 hold.

Furthermore, for every $t \geq 1$, there exist simple graphs G and H such that $\mathrm{SA}^{3t-1}(G, H)$ *has a non-negative rational solution but* $\mathrm{L}^{t}(G, H)$ *has no real solution.*

We prove Theorem 5.0.1 by constructing graph classes $\mathcal{L}_t$ for $t \geq 1$ such that $\mathrm{L}^{t}(G, H)$ has a real solution if, and only if, G and H are homomorphism indistinguishable over $\mathcal{L}_t$. To that end, we build on the correspondence between bilabelled graphs and their homomorphism tensors from Section 3.2, which was also used in Chapter 4. As a further proof ingredient, we generalise linear-algebraic techniques developed by Mančinska, Roberson, Šámal, Severini, & Varvitsiotis [125] and Mančinska, Roberson, & Varvitsiotis [126].

By Corollary 4.4.1, the system $\mathrm{SA}^{t}(G, H)$ has a non-negative real solution if, and only if, G and H are homomorphism indistinguishable over all graphs of treewidth at most $t - 1$. We obtain Theorem 5.0.1 by analysing the treewidth of the graphs in $\mathcal{L}_t$. More precisely, we show that all graphs in $\mathcal{L}_t$ have treewidth at most $3t - 1$. This yields the upper bound in Theorem 5.0.1 via Theorems 3.4.3 and 2.6.8. The lower bound is deferred to Corollary 6.4.2, as it requires the homomorphism distinguishing closure, which is introduced in Definition 6.0.1.

Our techniques extend to a stronger version of the Lasserre hierarchy which imposes non-negativity constraints on all variables. For every $t \geq 1$, we construct a graph class $\mathcal{L}_t^{+}$ such that $\mathrm{L}^{t}(G, H)$ has a non-negative real solution if, and only if, G and H are homomorphism indistinguishable over $\mathcal{L}_t^{+}$. By Theorem 4.0.1, the system $\mathrm{SA}^{t}(G, H)$ has a rational solution if, and only if, G and H are homomorphism indistinguishable over the class of graphs of pathwidth at most $t - 1$. Given these result, we conduct a detailed study of the relationship between the class of graphs of bounded treewidth, pathwidth, and the classes $\mathcal{L}_t$ and $\mathcal{L}_t^{+}$. The results, depicted by Figure 5.1, yield independent proofs of the known relations between feasibility of the Lasserre relaxation with and without non-negativity constraints and the Sherali–Adams relaxation with and without non-negativity constraints [23, 13] using homomorphism indistinguishability.

Chapter Outline. In Section 5.1, we derive several linear-algebraic reformulations of the system $\mathrm{L}^{t}(G, H)$ generalising [125, 126]. Ultimately, we obtain an equivalent system of equations involving homomorphism tensors of graphs from $\mathcal{L}_t$ and $\mathcal{L}_t^{+}$. In Section 5.2, we study the graph-theoretic properties of the graph classes $\mathcal{L}_t$ and $\mathcal{L}_t^{+}$. The results in this chapter are joint work with David E. Roberson and were previously published in [151, 152].

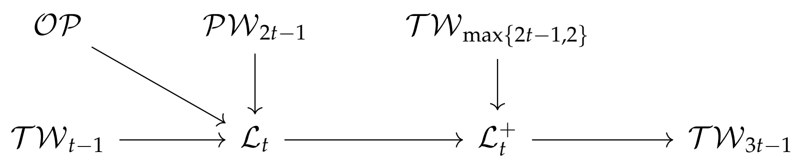


Figure 5.1: Relationships between $\mathcal{L}_t$, $\mathcal{L}_t^+$, the classes of graphs of bounded treewidth, bounded pathwidth, and the class $\mathcal{OP}$ of outerplanar graphs for $t \geq 1$. An arrow $\mathcal{A} \to \mathcal{B}$ indicates that $\mathcal{A} \subseteq \mathcal{B}$ and thus that $G \equiv_{\mathcal{B}} H$ implies $G \equiv_{\mathcal{A}} H$ for all simple graphs G and H. For formal statements, see Sections 5.2.1, 5.2.2, and 5.2.4.

5.1 From Lasserre to Homomorphism Tensors

In this section, the tools are developed which will be used to translate a solution to the level-t Lasserre relaxation into a statement on homomorphism indistinguishability. For this purpose, three equivalent characterisations of the feasibility of $\mathrm{L}^t(G, H)$ over the reals and the non-negative reals are introduced. Theorems 5.1.1 and 5.1.2 summarise our results. The notions in Items 2–4 and the graph classes $\mathcal{L}_t$ and $\mathcal{L}_t^+$ are defined in Sections 5.1.1, 5.1.2, 5.1.4, and 5.2, respectively. Most of the proofs are of a linear-algebraic nature. Graph-theoretic repercussions are discussed in Section 5.2.

Theorem 5.1.1. *Let $t \geq 1$. For simple graphs G and H, the following are equivalent:*

1. $\mathrm{L}^t(G, H)$ *has a real solution,*
2. *G and H are level-t $\mathcal{PSD}$-isomorphic, cf. Definition 5.1.3,*
3. *there is a level-t $\mathcal{PSD}$-isomorphism map from G to H, cf. Theorem 5.1.6,*
4. *G and H are partially t-equivalent, cf. Definition 5.1.10,*
5. *G and H are homomorphism indistinguishable over $\mathcal{L}_t$, cf. Definition 5.2.1.*

Theorem 5.1.2. *Let $t \geq 1$. For simple graphs G and H, the following are equivalent:*

1. $\mathrm{L}^t(G, H)$ *has a non-negative real solution,*
2. *G and H are level-t $\mathcal{DNN}$-isomorphic, cf. Definition 5.1.3,*
3. *there is a level-t $\mathcal{DNN}$-isomorphism map from G to H, cf. Theorem 5.1.6,*
4. *G and H are t-equivalent, cf. Definition 5.1.12,*
5. *G and H are homomorphism indistinguishable over $\mathcal{L}_t^+$, cf. Definition 5.2.1.*

Variants of the notions in Items 2–4 have already been defined for the case $t = 1$ in [126]. Our contribution amounts to extending these definitions to the entire Lasserre hierarchy. A recurring theme in this context is accounting for additional symmetries. The variables y_I of $\mathrm{L}^t(G, H)$, cf. Definition 2.6.9, are indexed by sets of vertex pairs rather than by tuples of such. Hence, when passing from such variables to tuple-indexed matrices, one must impose the additional symmetries arising this way. This is formalised at various points using an action of the symmetric group on the axes of the matrices. In the case $t = 1$, such a set-up is not necessary since indices I are of size at most two and all occurring matrices can be taken to be invariant under transposition.

In the subsequent sections, Theorems 5.1.1 and 5.1.2 will be proven in parallel. The equivalence of Items 1 and 2, 2 and 3, and 3 and 4 are established in Section 5.1.3, Section 5.1.2, and Section 5.1.4, respectively. The statements on homomorphism indistinguishability are proven in Section 5.2.

5.1.1 Isomorphism Relaxations via Matrix Families

In this section, as a first step towards proving Theorems 5.1.1 and 5.1.2, the notion of level-t $\mathcal{K}$-isomorphic graphs for arbitrary families of matrices $\mathcal{K}$ is introduced. In [126], level-1 $\mathcal{K}$-isomorphic graphs were studied for various families of matrices $\mathcal{K}$. In this work, the main interest lies on the family $\mathcal{PSD}$ of positive semidefinite matrices and the family $\mathcal{DNN}$ of entry-wise non-negative positive semidefinite matrices. Level-t isomorphism for these families is proven to correspond to feasibility of $\mathrm{L}^t(G, H)$ over the reals and the non-negative reals, respectively, cf. Theorems 5.1.8 and 5.1.9.

Definition 5.1.3. Let $\mathcal{K}$ be a family of matrices and $t \geq 1$. Two simple graphs G and H are *level-t $\mathcal{K}$-isomorphic* if there exists a matrix $M \in \mathcal{K}$ with rows and columns indexed by $(V(G) \times V(H))^t$ such that, for every $v_1w_1 \dots v_tw_t, v_{t+1}w_{t+1} \dots v_{2t}w_{2t} \in (V(G) \times V(H))^t$, the following equations hold:

For every $i \in [2t]$,

$$\sum_{v_i \in V(G)} M_{v_1w_1\dots v_tw_t, v_{t+1}w_{t+1}\dots v_{2t}w_{2t}} = \sum_{w_i \in V(H)} M_{v_1w_1\dots v_tw_t, v_{t+1}w_{t+1}\dots v_{2t}w_{2t}}, \tag{5.1}$$

$$\sum_{v_1,\dots,v_{2t} \in V(G)} M_{v_1w_1\dots v_tw_t, v_{t+1}w_{t+1}\dots v_{2t}w_{2t}} = 1 = \sum_{w_1,\dots,w_{2t} \in V(H)} M_{v_1w_1\dots v_tw_t, v_{t+1}w_{t+1}\dots v_{2t}w_{2t}}. \tag{5.2}$$

If $\mathrm{atp}_G(v_1, \dots, v_{2t}) \neq \mathrm{atp}_H(w_1, \dots, w_{2t})$, then

$$M_{v_1w_1\dots v_tw_t, v_{t+1}w_{t+1}\dots v_{2t}w_{2t}} = 0. \tag{5.3}$$

For all $\sigma \in \mathfrak{S}_{2t}$,

$$M_{v_1w_1\dots v_tw_t, v_{t+1}w_{t+1}\dots v_{2t}w_{2t}} = M_{v_{\sigma(1)}w_{\sigma(1)}\dots v_{\sigma(t)}w_{\sigma(t)}, v_{\sigma(t+1)}w_{\sigma(t+1)}\dots v_{\sigma(2t)}w_{\sigma(2t)}}. \tag{5.4}$$

In Equation (5.3), atp denotes the atomic type, cf. Section 2.3. Note that, for $t = 1$, Equation (5.4) asserts that M is symmetric, i.e. that $M = M^T$. Thus, Definition 5.1.3 generalises [126, Equations (3)–(6)].

The graph classes $\mathcal{L}_t$ and $\mathcal{L}_t^+$ that will be constructed in Definition 5.2.1 are closed under taking minors, cf. Theorem 5.2.12. That is is not a mere coincidence is demonstrated by the following Remark 5.1.4. By Theorem 7.0.1, every homomorphism indistinguishability relation which is preserved under complements is over a minor-closed graph class.

Remark 5.1.4. Let $t \geq 1$. For simple graphs G and H and $v_1, \dots, v_{2t} \in V(G)$, $w_1, \dots, w_{2t} \in V(H)$, it holds that

$$\operatorname{atp}_G(v_1, \dots, v_{2t}) = \operatorname{atp}_H(w_1, \dots, w_{2t}) \iff \operatorname{atp}_{\overline{G}}(v_1, \dots, v_{2t}) = \operatorname{atp}_{\overline{H}}(w_1, \dots, w_{2t}).$$

Thus, for all families of matrices $\mathcal{K}$ and $t \geq 1$, G and H are level-t $\mathcal{K}$-isomorphic if, and only if, $\overline{G}$ and $\overline{H}$ are level-t $\mathcal{K}$-isomorphic.

5.1.2 Choi Matrices and Isomorphism Maps

In this section, an alternative characterisation for level-t $\mathcal{K}$-isomorphism is given. Intuitively, the indices of the matrix $M \in \mathbb{C}^{(V(G)\times V(H))^t \times (V(G)\times V(H))^t}$ from Definition 5.1.3 are regrouped yielding a linear map $\Phi \colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(H)^t \times V(H)^t}$. In linear-algebraic terms, M is the Choi matrix of Φ. The map Φ will later be interpreted as a function sending homomorphism tensors of (t,t)-bilabelled graphs $\boldsymbol{F}_G \in \mathbb{C}^{V(G)^t \times V(G)^t}$ with respect to G to their counterparts $\boldsymbol{F}_H$ for H.

The most basic bilabelled graphs, so-called atomic graphs, make their first appearance in Theorem 5.1.6. Recall from Definition 3.2.10 that a bilabelled graph $\boldsymbol{F} \in \mathcal{G}(t,t)$ is atomic if all its vertices are labelled. These graphs are used to reformulate Equations (2.24) and (5.3). The atomic graphs are also the graphs which the sets $\mathcal{L}_t$ and $\mathcal{L}_t^+$ are generated by, cf. Definition 5.2.1. Examples are depicted by Figures 5.2 and 5.4.

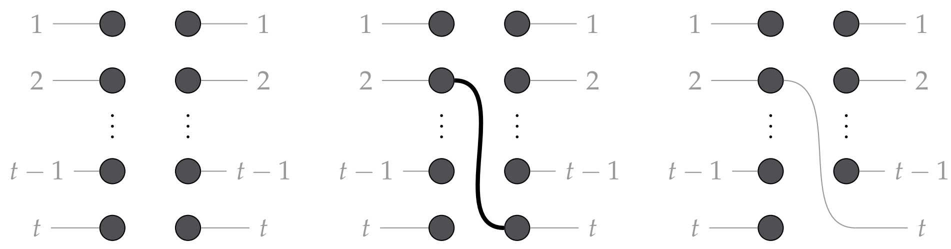
Observation 5.1.5. *Let $t \geq 1$. The set of atomic graphs $\mathcal{A}(t,t)$ is generated under parallel composition by the graphs*

- $\boldsymbol{J} := (J, (1, \dots, t), (t+1, \dots, 2t))$ *with* $V(J) = [2t]$, $E(J) = \emptyset$,
- $\boldsymbol{A}^{ij} := (A^{ij}, (1, \dots, t), (t+1, \dots, 2t))$ *with* $V(A^{ij}) = [2t]$, $E(A^{ij}) = \{ij\}$ *for* $1 \leq i < j \leq 2t$,
- $\boldsymbol{I}^{ij}$ *for* $1 \leq i < j \leq 2t$ *which is obtained from* $\boldsymbol{A}^{ij}$ *by contracting and removing the edge* ij.

The following Theorem 5.1.6 relates the properties of Φ and M. In Equation (5.7), J denotes the all-ones matrix of appropriate dimension. See Section 2.4.3 for definitions of the linear-algebraic notions used in Theorem 5.1.6.

Theorem 5.1.6. *Let $t \geq 1$. Let G and H be simple graphs and $\mathcal{K} \in \{\mathcal{DNN}, \mathcal{PSD}\}$ be a family of matrices. Let $\Phi \colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(H)^t \times V(H)^t}$ be a linear map. Then the following are equivalent.*

1. *the Choi matrix C_Φ of Φ satisfies Equations* (5.1) *to* (5.4) *and $C_\Phi \in \mathcal{K}$,*



(a) $\boldsymbol{J}$. **(b)** $\boldsymbol{A}^{2,2t}$. **(c)** $\boldsymbol{I}^{2,2t}$.

Figure 5.2: Examples of the atomic graphs from Observation 5.1.5.

2. Φ *is a* level-t $\mathcal{K}$-isomorphism map from G to H*, i.e. it satisfies that*

$$\Phi \text{ is completely } \mathcal{K}\text{-preserving}, \tag{5.5}$$
$$\Phi(\boldsymbol{A}_G \odot X) = \boldsymbol{A}_H \odot \Phi(X) \text{ for all } \boldsymbol{A} \in \mathcal{A}(t,t) \text{ and } X \in \mathbb{C}^{V(G)^t \times V(G)^t}, \tag{5.6}$$
$$\Phi(J) = J = \Phi^*(J), \tag{5.7}$$
$$\Phi(X^\sigma) = \Phi(X)^\sigma \text{ for all } \sigma \in \mathfrak{S}_{2t} \text{ and all } X \in \mathbb{C}^{V(G)^t \times V(G)^t}, \text{ and} \tag{5.8}$$

3. Φ^* *is a level-t $\mathcal{K}$-isomorphism map from H to G.*

We remark that Theorem 5.1.6, and in particular its Equations (5.6) and (5.7), have brought us closer to interpreting the Lasserre system of equation from the perspective of homomorphism indistinguishability. As argued in Remark 5.1.7, the map Φ, which will be understood as mapping homomorphism tensors $\boldsymbol{F}_G$ to $\boldsymbol{F}_H$, is sum-preserving. Since the sum of the entries of these tensors equals the number of homomorphisms from their underlying unlabelled graphs to G and H, respectively, this is relevant for establishing a connection between $\mathcal{K}$-isomorphism maps and homomorphism indistinguishability.

Remark 5.1.7. If a linear map $\Phi\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{m\times m}$ is such that $J = \Phi^*(J)$, then it is *sum-preserving*, i.e. $\operatorname{soe}(X) = \operatorname{soe}(\Phi(X))$ for all $X \in \mathbb{C}^{n\times n}$. Indeed, $\operatorname{soe}(X) = \langle X, J\rangle = \langle X, \Phi^*(J)\rangle = \langle \Phi(X), J\rangle = \operatorname{soe}(\Phi(X))$ where $\langle A, B\rangle := \operatorname{tr}(A^*B)$. In particular, if there is Φ satisfying Equations (5.6) and (5.7) for graphs G and H, then $|V(G)| = |V(H)|$.

Equipped with Remark 5.1.7, we conduct the proof of Theorem 5.1.6.

Proof of Theorem 5.1.6. The equivalence of Items 2 and 3 follows from Lemmas 2.4.9 and 2.4.12 and [126, Lemma 4.2i]. That $\Phi(X^\sigma) = \Phi(X)^\sigma$ for all $X \in \mathbb{C}^{V(G)^t \times V(G)^t}$ and $\sigma \in \mathfrak{S}_{2t}$ if, and only if, $\Phi^*(Y^\sigma) = \Phi^*(Y)^\sigma$ for all $Y \in \mathbb{C}^{V(H)^t \times V(H)^t}$ and $\sigma \in \mathfrak{S}_{2t}$, follows by observing that $(\Lambda_\sigma)^* = \Lambda_{\sigma^{-1}}$ for the map $\Lambda_\sigma\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(G)^t \times V(G)^t}$ defined via $X \mapsto X^\sigma$.

For the equivalence of Items 1 and 2, first note that, by Lemma 2.4.9, $C_\Phi \in \mathcal{K}$ if, and only if, Property (5.5) holds. Moreover, for $v_1, \dots, v_{2t} \in V(G)$ and $w_1, \dots, w_{2t} \in$

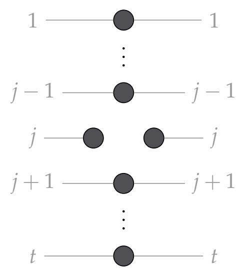


Figure 5.3: The atomic graph $\boldsymbol{K}^j$ as defined in Equation (5.9).

$V(H)$, the assertions

$$\forall \boldsymbol{A} \in \mathcal{A}(t,t), \quad \boldsymbol{A}_G(v_1 \dots v_t, v_{t+1} \dots v_{2t}) = \boldsymbol{A}_H(w_1 \dots w_t, w_{t+1} \dots w_{2t})$$

and $\operatorname{atp}_G(v_1, \dots, v_{2t}) = \operatorname{atp}_H(w_1, \dots, w_{2t})$ are equivalent. By Lemma 2.4.12, Equations (5.3) and (5.6) are equivalent. Furthermore, Equations (5.4) and (5.8) and Equations (5.2) and (5.7) are respectively equivalent.

Finally, we argue that Items 2 and 3 imply Equation (5.1). To that end, consider the atomic graph $\boldsymbol{K}^j \in \mathcal{A}(t,t)$ for $j \in [t]$ as defined in Equation (5.9) and depicted by Figure 5.3.

$$\boldsymbol{K}^j := \boldsymbol{I}^{1,t+1} \odot \dots \odot \boldsymbol{I}^{j-1,t+j-1} \odot \boldsymbol{I}^{j+1,t+j+1} \odot \dots \odot \boldsymbol{I}^{t,2t}. \tag{5.9}$$

In order to apply Lemma 2.4.10, we first argue that Φ is trace-preserving. By Equation (5.7) and Remark 5.1.7, Φ is sum-preserving. Hence, for every $X \in \mathbb{C}^{V(G)^t \times V(G)^t}$, with $\boldsymbol{I} \in \mathcal{A}(t,t)$ as in Figure 3.3b,

$$\operatorname{tr}(\Phi(X)) = \operatorname{soe}(\boldsymbol{I}_H \odot \Phi(X)) \overset{(5.6)}{=} \operatorname{soe}(\Phi(\boldsymbol{I}_G \odot X)) = \operatorname{soe}(\boldsymbol{I}_G \odot X) = \operatorname{tr}(X).$$

Thus, Lemma 2.4.10 and Equation (5.6) yield that, for all $j \in [t]$ and all $X \in \mathbb{C}^{V(G)^t \times V(G)^t}$,

$$\Phi(\boldsymbol{K}^j_G X) = \Phi(\boldsymbol{K}^j_G)\Phi(X) \quad \text{and} \quad \Phi(X\boldsymbol{K}^j_G) = \Phi(X)\Phi(\boldsymbol{K}^j_G). \tag{5.10}$$

Next, standard basis elements are substituted for X in Equation (5.10). For vertices $v_1, \dots, v_{2t} \in V(G)$, write $E^{v_1 \dots v_{2t}} \in \mathbb{C}^{V(G)^t \times V(G)^t}$ for the corresponding standard basis vector. To ease notation, Equation (5.1) is verified for $i = 1$. For all vertices $v_1, \dots, v_{2t} \in V(G)$ and $w_1, \dots, w_{2t} \in V(H)$,

$$\begin{aligned}
\sum_{v \in V(G)} M_{vw_1v_2w_2\dots v_tw_t, v_{t+1}w_{t+1}\dots v_{2t}w_{2t}} &= \sum_{v \in V(G)} \Phi_{w_1\dots w_{2t}, vv_2\dots v_{2t}} \\
&= \sum_{v \in V(G)} \Phi(E^{vv_2\dots v_{2t}})_{w_1\dots w_{2t}} \\
&= \Phi(\boldsymbol{K}^1_G E^{v_1\dots v_{2t}})_{w_1\dots w_{2t}}
\end{aligned}$$

$$
\begin{aligned}
&\overset{(5.10)}{=} (\Phi(\boldsymbol{K}_G^1)\Phi(E^{v_1\dots v_{2t}}))_{w_1\dots w_{2t}} \\
&\overset{(5.6)}{=} (\boldsymbol{K}_H^1\Phi(E^{v_1\dots v_{2t}}))_{w_1\dots w_{2t}} \\
&= \sum_{w\in V(H)} \Phi_{ww_2\dots w_{2t},v_1\dots v_{2t}} \\
&= \sum_{w\in V(G)} M_{v_1wv_2w_2\dots v_tw_t,v_{t+1}w_{t+1}\dots v_{2t}w_{2t}},
\end{aligned}
$$

as desired. □

5.1.3 From $\mathcal{K}$-Isomorphism Maps to the Lasserre Hierarchy

By the following Theorems 5.1.8 and 5.1.9, the notions introduced in Definition 5.1.3 and Theorem 5.1.6 are equivalent to the object of our main interest, namely the feasibility of the level-t Lasserre relaxation $\mathrm{L}^t(G,H)$ with and without non-negativity constraints. Our results extend those of [126, Lemma 10.1] to the entire Lasserre hierarchy.

Theorem 5.1.8. *Let $t \geq 1$. Two simple graphs G and H are level-t $\mathcal{PSD}$-isomorphic if, and only if, $\mathrm{L}^t(G,H)$ has a real solution.*

Proof. Suppose that $(y_I)_{I\in\binom{V(G)\times V(H)}{\leq 2t}}$ is a real solution to $\mathrm{L}^t(G,H)$, i.e. it satisfies Equations (2.20) to (2.24). We argue that the matrix defined via

$$
M_{v_1w_1\dots v_tw_t,v_{t+1}w_{t+1}\dots v_{2t}w_{2t}} := y_{\{v_1w_1,\dots,v_{2t}w_{2t}\}}
$$

satisfies Equations (5.1) to (5.4). Equation (5.3) follows directly from Equation (2.24). Equation (5.4) is immediate from the definition.

By Equation (2.20) and Section 2.4.3, there exist vectors x_I for every $I \in \binom{V(G)\times V(H)}{\leq t}$ such that $y_{I\cup J} = \langle x_I, x_J\rangle$ for $I, J \in \binom{V(G)\times V(H)}{\leq t}$. Then

$$
M_{v_1w_1\dots v_tw_t,v_{t+1}w_{t+1}} = y_{\{v_1w_1,\dots,v_{2t}w_{2t}\}} = \left\langle x_{\{v_1w_1,\dots,v_tw_t\}}, x_{\{v_{t+1}w_{t+1},\dots,v_{2t}w_{2t}\}}\right\rangle .
$$

Thus, M is positive semidefinite. It remains to verify Equations (5.1) and (5.2).

Claim 5.1.8a. Every $I \in \binom{V(G)\times V(H)}{\leq t-1}$ satisfies $\sum_{v\in V(G)} x_{I\cup\{vw\}} = x_I = \sum_{w\in V(H)} x_{I\cup\{vw\}}$.

Proof of Claim. Recall that $y_{I\cup J} = \langle x_I, x_J\rangle$ for $I, J \in \binom{V(G)\times V(H)}{\leq t}$. By Equations (2.21) and (2.24),

$$
\begin{aligned}
\left\langle \sum_{v\in V(G)} x_{I\cup\{vw\}}, \sum_{v\in V(G)} x_{I\cup\{vw\}} \right\rangle &= \sum_{v,v'\in V(G)} \left\langle x_{I\cup\{vw\}}, x_{I\cup\{v'w\}}\right\rangle \\
&= \sum_{v,v'\in V(G)} y_{I\cup\{vw\}\cup\{v'w\}}
\end{aligned}
$$

$$\overset{(2.24)}{=} \sum_{v \in V(G)} y_{I \cup \{vw\}} \overset{(2.21)}{=} y_I.$$

Observe that Equation (2.21) is indeed applicable since $t - 1 \leq 2t - 2$ for all $t \geq 1$. Moreover,

$$\left\langle x_I, \sum_{v \in V(G)} x_{I \cup \{vw\}} \right\rangle = \sum_{v \in V(G)} y_{I \cup \{vw\}} \overset{(2.21)}{=} y_I.$$

Hence, combining the above equations, $\left\| x_I - \sum_{v \in V(G)} x_{I \cup \{vw\}} \right\|^2 = y_I - 2y_I + y_I = 0$. The second part of the claim, i.e. $x_I = \sum_{w \in V(H)} x_{I \cup \{vw\}}$, is proven analogously. $\lhd$

Claim 5.1.8a implies Equation (5.2). Indeed,

$$\begin{aligned}
\sum_{v_1 \dots v_{2t} \in V(G)} M_{v_1 w_1 \dots v_t w_t, v_{t+1} w_{t+1} \dots v_{2t} w_{2t}} &= \sum_{v_1 \dots v_{2t} \in V(G)} y_{\{v_1 w_1, \dots, v_t w_t\} \cup \{v_{t+1} w_{t+1} \dots v_{2t} w_{2t}\}} \\
&= \sum_{v_1 \dots v_{2t} \in V(G)} \left\langle x_{\{v_1 w_1, \dots, v_t w_t\}}, x_{\{v_{t+1} w_{t+1} \dots v_{2t} w_{2t}\}} \right\rangle \\
&= \langle x_\emptyset, x_\emptyset \rangle \\
&= y_\emptyset \\
&\overset{(2.24)}{=} 1.
\end{aligned}$$

Moreover, for Equation (5.1), letting $i = 1$ to ease notation, by Claim 5.1.8a,

$$\begin{aligned}
\sum_{v_1 \in V(G)} M_{v_1 w_1 \dots v_t w_t, v_{t+1} w_{t+1} \dots v_{2t} w_{2t}} &= \sum_{v_1 \in V(G)} y_{\{v_1 w_1\} \cup \{v_2 w_2 \dots v_t w_t\} \cup \{v_{t+1} w_{t+1} \dots v_{2t} w_{2t}\}} \\
&= \sum_{v_1 \in V(G)} \left\langle x_{\{v_1 w_1\} \cup \{v_2 w_2 \dots v_t w_t\}}, x_{\{v_{t+1} w_{t+1} \dots v_{2t} w_{2t}\}} \right\rangle \\
&= \sum_{w_1 \in V(G)} \left\langle x_{\{v_1 w_1\} \cup \{v_2 w_2 \dots v_t w_t\}}, x_{\{v_{t+1} w_{t+1} \dots v_{2t} w_{2t}\}} \right\rangle \\
&= \sum_{w_1 \in V(G)} M_{v_1 w_1 \dots v_t w_t, v_{t+1} w_{t+1} \dots v_{2t} w_{2t}}.
\end{aligned}$$

This concludes the proof that M satisfies Equations (5.1) to (5.4).

Conversely, let $x_{\vec{I}}$ for $\vec{I} \in (V(G) \times V(H))^t$ denote the Gram vectors of a matrix M satisfying Equations (5.1) to (5.4). For $I \in \binom{V(G) \times V(H)}{t}$, define $x_I := x_{\vec{I}}$ for any ordering $\vec{I}$ of I. By Equation (5.4), x_I is well-defined. Let furthermore

$$x_I^{v_{i+1} \dots v_t} := \sum_{w_{i+1}, \dots, w_t \in V(H)} x_{\vec{I} v_{i+1} w_{i+1} \dots v_t w_t}$$

for $I \in \binom{V(G) \times V(H)}{i}$ and $v_{i+1} \dots v_t \in V(G)^{t-i}$ for $0 \leq i < t$. Define $x_I^{w_{i+1} \dots w_t}$ analogously.

Claim 5.1.8b. Let $0 \leq i < t$. For all $v_{i+1} \dots v_t, v'_{i+1} \dots v'_t \in V(G)^{t-i}$, it holds that $x_I^{v_{i+1} \dots v_t} = x_I^{v'_{i+1} \dots v'_t}$.

Proof of Claim. By definition, the term $\left\| x_I^{v_{i+1} \dots v_t} - x_I^{v'_{i+1} \dots v'_t} \right\|^2$ is equal to the sum over the vertices $w_{i+1}, \dots, w_t \in V(H)$ and $w'_{i+1}, \dots, w'_t \in V(H)$ of

$$M_{\vec{I}v_{i+1}w_{i+1} \dots v_t w_t, \vec{I}v_{i+1}w'_{i+1} \dots v_t w'_t} - 2M_{\vec{I}v_{i+1}w_{i+1} \dots v_t w_t, \vec{I}v'_{i+1}w'_{i+1} \dots v'_t w'_t} + M_{\vec{I}v'_{i+1}w_{i+1} \dots v'_t w_t, \vec{I}v'_{i+1}w'_{i+1} \dots v'_t w'_t}.$$

By Equation (5.1), this expression is zero. ◁

By Claim 5.1.8b, the reference to $v_{i+1} \dots v_t$ in $x_I^{v_{i+1} \dots v_t}$ can be dropped, yielding vectors x_I^G and x_I^H for every $I \in \binom{V(G) \times V(H)}{\leq t}$. It follows that

$$\begin{aligned} |V(G)|^{t-i} x_I^G &= \sum_{v_{i+1} \dots v_t \in V(G)^{t-i}} x_I^{v_{i+1} \dots v_t} = \sum_{\substack{v_{i+1} \dots v_t \in V(G)^{t-i} \\ w_{i+1} \dots w_t \in V(H)^{t-i}}} x_{\vec{I}v_{i+1}w_{i+1} \dots v_t w_t} \\ &= \sum_{w_{i+1} \dots w_t \in V(H)^{t-i}} x_I^{w_{i+1} \dots w_t} = |V(H)|^{t-i} x_I^H. \end{aligned}$$

This implies that $x_I^G = x_I^H$ since G and H have the same number of vertices, cf. Remark 5.1.7. Let $x_I \coloneqq x_I^G = x_I^H$. The following Claim 5.1.8c is immediate from Equation (5.4):

Claim 5.1.8c. If $I \cup J = I' \cup J'$ for $I, I', J, J' \in \binom{V(G) \times V(H)}{\leq t}$, then $\langle x_I, x_J \rangle = \langle x_{I'}, x_{J'} \rangle$.

Hence, the variable y_I, for $I \in \binom{V(G) \times V(H)}{\leq 2t}$, can be set to $\langle x_{I'}, x_{I''} \rangle$ for any $I', I'' \in \binom{V(G) \times V(H)}{\leq t}$ such that $I = I' \cup I''$. Then Equations (2.20) to (2.22) hold by construction. In fact, it follows that Equations (5.11) and (5.12) below, which imply Equations (2.21) and (2.22), hold:

$$\sum_{v \in V(G)} y_{I \cup \{vw\}} = y_I \qquad \text{for all } I \in \binom{V(G) \times V(H)}{\leq 2t-1} \text{ and } w \in V(H), \tag{5.11}$$

$$\sum_{w \in V(H)} y_{I \cup \{vw\}} = y_I \qquad \text{for all } I \in \binom{V(G) \times V(H)}{\leq 2t-1} \text{ and } v \in V(G). \tag{5.12}$$

Equation (2.24) follows from Equation (5.3). □

The following Theorem 5.1.9 is proven analogously, observing that the construction in the proof of Theorem 5.1.8 preserves non-negativity in both directions.

Theorem 5.1.9. *Let $t \geq 1$. Two simple graphs G and H are level-t $\mathcal{DNN}$-isomorphic if, and only if, $\mathrm{L}^t(G, H)$ has a non-negative real solution.*

5.1.4 Isomorphisms between Matrix Algebras

To the two reformulations of feasibility of $\mathrm{L}^t(G, H)$ over the reals and the non-negative reals from the previous sections, a third characterisation is added in this section. It is shown that two graphs are level-t $\mathcal{PSD}$-isomorphic ($\mathcal{DNN}$-isomorphic) if, and only if, certain matrix algebras associated to them are isomorphic. These algebras will be identified as the algebras of homomorphism tensors for graphs from the families $\mathcal{L}_t$ and $\mathcal{L}_t^+$. The so-called (partially) coherent algebras considered in this section are natural generalisations of the coherent algebras which are well-studied in the context of the 2-dimensional Weisfeiler–Leman algorithm [40].

Partially Coherent Algebras and $\mathcal{PSD}$-Isomorphism Maps

Let $n, t \geq 1$ and $S \subseteq \mathbb{C}^{n^t \times n^t}$. A matrix algebra $\mathcal{A} \subseteq \mathbb{C}^{n^t \times n^t}$ is *S-partially coherent* if it is unital, i.e. $I \in \mathcal{A}$, self-adjoint, i.e. $A^* \in \mathcal{A}$ for all $A \in \mathcal{A}$, contains the all-ones matrix J, and is closed under Schur products with any matrix in S, i.e. if $A \in \mathcal{A}$ and $B \in S$, then $A \odot B \in \mathcal{S}$.

A matrix algebra $\mathcal{A} \subseteq \mathbb{C}^{n^t \times n^t}$ is *self-symmetrical* if, for every $A \in \mathcal{A}$ and $\sigma \in \mathfrak{S}_{2t}$, also $A^\sigma \in \mathcal{A}$. Note that, for $t = 1$, an algebra $\mathcal{A}$ is self-symmetrical if, and only if, for all $A \in \mathcal{A}$, also $A^T \in \mathcal{A}$ where A^T is the transpose of A.

Definition 5.1.10. Let $t \geq 1$. The *t-partially coherent algebra* $\widehat{\mathcal{A}}^t_G$ of a simple graph G is the minimal[8] self-symmetrical S-partially coherent algebra where S is the set of homomorphism tensors of (t, t)-bilabelled atomic graphs for G.

Two simple graphs G and H are *partially t-equivalent* if there is a *partial t-equivalence*, i.e. a vector space isomorphism $\varphi\colon \widehat{\mathcal{A}}^t_G \to \widehat{\mathcal{A}}^t_H$ such that

1. $\varphi(M^*) = \varphi(M)^*$ for all $M \in \widehat{\mathcal{A}}^t_G$,
2. $\varphi(MN) = \varphi(M)\varphi(N)$ for all $M, N \in \widehat{\mathcal{A}}^t_G$,
3. $\varphi(I) = I$, $\varphi(\boldsymbol{A}_G) = \boldsymbol{A}_H$ for all $\boldsymbol{A} \in \mathcal{A}(t, t)$, and $\varphi(J) = J$,
4. $\varphi(\boldsymbol{A}_G \odot M) = \boldsymbol{A}_H \odot \varphi(M)$ for all $\boldsymbol{A} \in \mathcal{A}(t, t)$ and any $M \in \widehat{\mathcal{A}}^t_G$.
5. $\varphi(M^\sigma) = \varphi(M)^\sigma$ for all $M \in \widehat{\mathcal{A}}^t_G$ and all permutations $\sigma \in \mathfrak{S}_{2t}$.

The following Theorem 5.1.11 extends [126, Theorem 6.2].

Theorem 5.1.11. *Let $t \geq 1$. Two simple graphs G and H are partially t-equivalent if, and only if, there is a level-t $\mathcal{PSD}$-isomorphism map from G to H.*

Proof. Let $\Phi\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(H)^t \times V(H)^t}$ be a level-t $\mathcal{PSD}$-isomorphism map from G to H, i.e. it satisfies Equations (5.5) to (5.8). By Remark 5.1.7 and Equations (5.6) and (5.7), $\Phi(\boldsymbol{A}_G) = \boldsymbol{A}_H$ for all atomic $\boldsymbol{A} \in \mathcal{A}(t, t)$ and $|V(G)| = |V(H)| =: n$. Similarly, $\Phi^*(\boldsymbol{A}_H) = \boldsymbol{A}_G$ for all atomic $\boldsymbol{A}$ by Theorem 5.1.6. By Equations (5.5)

[8] The intersection of two self-symmetrical S-partially coherent algebras is itself a self-symmetrical S-partially coherent algebra. Hence, $\widehat{\mathcal{A}}^t_G$ is well-defined.

(a) $\boldsymbol{J}$. **(b)** $\boldsymbol{A}^{1,2}$. **(c)** $\boldsymbol{I}^{1,2}$.

Figure 5.4: The three atomic simple graphs in $\mathcal{A}(1,1)$.

and (5.6), Φ is completely positive and unital. By Theorem 5.1.6, $\Phi^*(I) = I$ and thus Φ is trace-preserving [126, Lemma 4.2]. Furthermore,

$$\Phi(\boldsymbol{A}_G) = \boldsymbol{A}_H, \quad \Phi^*(\boldsymbol{A}_H) = \boldsymbol{A}_G, \quad \Phi(J) = J = \Phi^*(J).$$

for all atomic $\boldsymbol{A} \in \mathcal{A}(t,t)$. Thus, Lemma 2.4.10 implies that $\Phi(\boldsymbol{A}_G W) = \boldsymbol{A}_H \Phi(W)$ and $\Phi(W\boldsymbol{A}_G) = \Phi(W)\boldsymbol{A}_H$ for all atomic $\boldsymbol{A} \in \mathcal{A}(t,t)$ and all $W \in \mathbb{C}^{V(G)^t \times V(G)^t}$. Hence, the restriction of Φ to $\widehat{\mathcal{A}}^t_G$ is a partial t-equivalence from G and H.

Conversely, suppose that $\varphi\colon \widehat{\mathcal{A}}^t_G \to \widehat{\mathcal{A}}^t_H$ is as in Definition 5.1.10. By [126, Lemma 5.3], φ is trace-preserving. By Lemma 2.4.11, there exists a unitary matrix $U \in \mathbb{C}^{n^t \times n^t}$ such that $\varphi(X) = UXU^*$ for all $X \in \widehat{\mathcal{A}}^t_G$. Let $\widehat{\varphi}\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(H)^t \times V(H)^t}$ be the map given by $\widehat{\varphi}(X) = UXU^*$. Let $\Pi\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \widehat{\mathcal{A}}^t_G$ be the orthogonal projection onto $\widehat{\mathcal{A}}^t_G$. Define a map $\Phi\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(H)^t \times V(H)^t}$ by $\Phi := \widehat{\varphi} \circ \Pi$. By [126, Lemma 5.3], $\widehat{\varphi}$ is completely positive and trace-preserving. By [126, Lemma 5.4], so is Π and hence their composition Φ. Hence, Equation (5.5) holds.

Furthermore, $\Pi(J) = J$ and hence $\Phi(J) = J = \Phi^*(J)$. So Φ satisfies Equation (5.7).

For Equation (5.8), consider the linear map $\Lambda_\sigma\colon X \mapsto X^\sigma$ for $\sigma \in \mathfrak{S}_{2t}$. Since $\widehat{\mathcal{A}}_G$ is closed under the action of $\mathfrak{S}_{2t}$, it holds that $\Lambda_\sigma \circ \Pi = \Pi \circ \Lambda_\sigma \circ \Pi$. Furthermore, $(\Lambda_\sigma)^* = \Lambda_{\sigma^{-1}}$ and Π is Hermitian, i.e. $\Pi^* = \Pi$. Hence,

$$\Pi \circ \Lambda_\sigma = \Pi^* \circ \Lambda_\sigma = (\Lambda_{\sigma^{-1}} \circ \Pi)^* = (\Pi \circ \Lambda_{\sigma^{-1}} \circ \Pi)^* = \Pi \circ \Lambda_\sigma \circ \Pi = \Lambda_\sigma \circ \Pi.$$

So Π and Λ_σ commute. Hence,

$$\Phi(X^\sigma) = (\widehat{\varphi} \circ \Pi \circ \Lambda_\sigma)(X) = (\widehat{\varphi} \circ \Lambda_\sigma \circ \Pi)(X) = ((\widehat{\varphi} \circ \Pi)(X))^\sigma = \Phi(X)^\sigma.$$

Equation (5.6) follows similarly, cf. the proof of [126, Theorem 6.2]. □

Coherent Algebras and $\mathcal{DNN}$-Isomorphism Maps

Let $n, t \geq 1$. A matrix algebra $\mathcal{A} \subseteq \mathbb{C}^{n^t \times n^t}$ is *coherent* if it is unital, self-adjoint, contains the all-ones matrix and is closed under Schur products.

For $t = 1$, the 1-adjacency algebra as defined below is equal to the well-studied *adjacency algebra* of a graph G, cf. [40]. The latter is the smallest coherent algebra containing the adjacency matrix of the graph. The former is generated by the homomorphism tensors of $(1,1)$-bilabelled atomic graphs. These graphs are depicted by Figure 5.4. Their homomorphism tensors are the all-ones matrix, the adjacency matrix of the graph, and the identity matrix.

Definition 5.1.12. Let $t \geq 1$. The *t-adjacency algebra* $\mathcal{A}_G^t$ of a simple graph G is the minimal self-symmetrical coherent algebra containing the homomorphism tensors of (t, t)-bilabelled atomic graphs for G.

Two simple graphs G and H are *t-equivalent* if there is a *t-equivalence*, i.e. a vector space isomorphism $\varphi\colon \mathcal{A}_G^t \to \mathcal{A}_H^t$ such that

1. $\varphi(M^*) = \varphi(M)^*$ for all $M \in \mathcal{A}_G^t$,
2. $\varphi(MN) = \varphi(M)\varphi(N)$ for all $M, N \in \mathcal{A}_G^t$,
3. $\varphi(I) = I$, $\varphi(\boldsymbol{A}_G) = \boldsymbol{A}_H$ for all $\boldsymbol{A} \in \mathcal{A}(t,t)$, and $\varphi(J) = J$,
4. $\varphi(M \odot N) = \varphi(M) \odot \varphi(N)$ for all $M, N \in \mathcal{A}_G^t$.
5. $\varphi(M^\sigma) = \varphi(M)^\sigma$ for all $M \in \mathcal{A}_G^t$ and all permutations $\sigma \in \mathfrak{S}_{2t}$.

The following Theorem 5.1.13 extends [126, Theorem 7.3].

Theorem 5.1.13. *Let $t \geq 1$. Two simple graphs G and H are t-equivalent if, and only if, there is a level-t $\mathcal{DNN}$-isomorphism map from G to H.*

Proof. Let $\Phi\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(H)^t \times V(H)^t}$ be a level-t $\mathcal{DNN}$-isomorphism map. Let φ be the restriction of Φ to $\mathcal{A}_G^t$. Given the arguments in the proof of Theorem 5.1.11, it suffices to show that $\varphi(M \odot N) = \varphi(M) \odot \varphi(N)$ for all $M, N \in \mathcal{A}_G^t$ and that $\varphi^*(M \odot N) = \varphi^*(M) \odot \varphi^*(N)$ for all $M, N \in \mathcal{A}_H^t$. This follows from [126, Lemma 7.2].

Conversely, suppose that $\varphi\colon \mathcal{A}_G^t \to \mathcal{A}_H^t$ is as in Definition 5.1.12. It follows as in [126, Lemma 5.3] that φ is trace-preserving. By Lemma 2.4.11, there exists a unitary matrix $U \in \mathbb{C}^{n^t \times n^t}$ such that $\varphi(X) = UXU^*$ for all $X \in \mathcal{A}_G^t$. Let $\widehat{\varphi}\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(H)^t \times V(H)^t}$ be the map given by $\widehat{\varphi}(X) = UXU^*$. Let $\Pi\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathcal{A}_G^t$ be the orthogonal projection onto $\mathcal{A}_G^t$. Define a map $\Phi\colon \mathbb{C}^{V(G)^t \times V(G)^t} \to \mathbb{C}^{V(H)^t \times V(H)^t}$ by $\Phi := \widehat{\varphi} \circ \Pi$. Given Theorem 5.1.11, it suffices to argue that the Choi matrix of Φ is entry-wise non-negative. This can be done as in the proof of [126, Theorem 7.3]. $\square$

5.2 Homomorphism Indistinguishability over $\mathcal{L}_t$ and $\mathcal{L}_t^+$

Using techniques from Chapter 4, we finally establish a characterisation of when $\mathrm{L}^t(G, H)$ is feasible in terms of homomorphism indistinguishability of G and H. In order to do so, we introduce the graph classes $\mathcal{L}_t$ and $\mathcal{L}_t^+$. In Sections 5.2.1 and 5.2.2, we compare $\mathcal{L}_t$ and $\mathcal{L}_t^+$ to the classes of graphs of bounded treewidth and pathwidth obtaining the results depicted by Figure 5.1. In Section 5.2.4, $\mathcal{L}_1$ and $\mathcal{L}_1^+$ are identified as the classes of outerplanar graphs and graphs of treewidth at most two, respectively.

Definition 5.2.1. Let $t \geq 1$. Write $\mathcal{L}^+(t,t)$ for the class of (t,t)-bilabelled graphs generated by the set of atomic graphs $\mathcal{A}(t,t)$ under parallel composition, series composition, and the action of $\mathfrak{S}_{2t}$ on the labels. Write $\mathcal{L}_t^+ := \operatorname{soe}(\mathcal{L}^+(t,t))$.

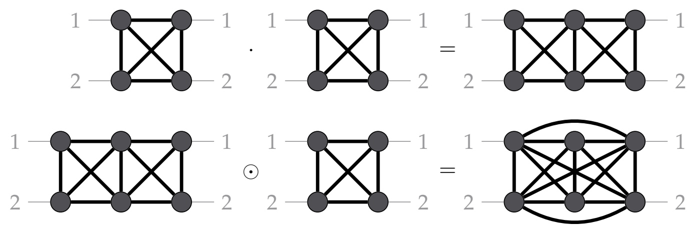


Figure 5.5: The bilabelled graphs in Observation 5.2.2 for $t = 2$.

Write $\mathcal{L}(t,t) \subseteq \mathcal{L}^+(t,t)$ for the class of (t,t)-bilabelled graphs generated by the set of atomic graphs $\mathcal{A}(t,t)$ under parallel composition with graphs from $\mathcal{A}(t,t)$, series composition, and the action of $\mathfrak{S}_{2t}$ on the labels. Write $\mathcal{L}_t := \operatorname{soe}(\mathcal{L}(t,t))$.

Clearly, $\mathcal{L}(t,t) \subseteq \mathcal{L}^+(t,t)$ and thus $\mathcal{L}_t \subseteq \mathcal{L}_t^+$. The only difference between $\mathcal{L}(t,t)$ and $\mathcal{L}^+(t,t)$ is that $\mathcal{L}(t,t)$ is closed under parallel composition with atomic graphs only. This reflects an observation in Corollary 4.2.4 relating the closure under arbitrary gluing products to non-negative solutions to systems of equations characterising homomorphism indistinguishability. Intuitively, one may use arbitrary Schur products, the algebraic counterparts of gluing, for a Vandermonde interpolation argument to obtain non-negative solutions, cf. Section 4.3.2.

The following Observation 5.2.2 illustrates how the operations in Definition 5.2.1 can be used to generate more complicated graphs from the atomic graphs, cf. Figure 5.5.

Observation 5.2.2. *Let $t \geq 1$. The class $\mathcal{L}_t$ contains the $3t$-vertex complete graph K_{3t}.*

Proof. Consider the atomic graph $\boldsymbol{E} := \bigodot_{1 \leq i < j \leq 2t} \boldsymbol{A}^{ij} \in \mathcal{A}(t,t)$. The graph underlying $\boldsymbol{E} \odot (\boldsymbol{E} \cdot \boldsymbol{E})$ is isomorphic to K_{3t}. □

The only missing implications of Theorems 5.1.1 and 5.1.2 follow from the next two theorems:

Theorem 5.2.3. *Let $t \geq 1$. Two simple graphs G and H are homomorphism indistinguishable over $\mathcal{L}_t$ if, and only if, they are partially t-equivalent.*

Theorem 5.2.4. *Let $t \geq 1$. Two simple graphs G and H are homomorphism indistinguishable over $\mathcal{L}_t^+$ if, and only if, they are t-equivalent.*

For the proofs of Theorems 5.2.3 and 5.2.4, we generalise the notion of inner-product compatibility from Definition 4.3.23. Recall that a class of labelled graphs $\mathcal{R}$ is inner-product compatible if for every two labelled graphs $\boldsymbol{R}, \boldsymbol{S} \in \mathcal{R}$ one can write the inner-product of their homomorphism vectors $\boldsymbol{R}_G$ and $\boldsymbol{S}_G$ as the sum-of-entries of some $\boldsymbol{T}_G$ where $\boldsymbol{T} \in \mathcal{R}$ is labelled graph from the class. Due to the correspondence between combinatorial operations on labelled graphs and algebraic operations on their homomorphism vectors, cf. Section 3.2, this is equivalent to the graph-theoretic assumption that $\operatorname{soe}(\boldsymbol{R} \odot \boldsymbol{S}) = \operatorname{soe}(\boldsymbol{T})$, i.e. the unlabelled graph obtained by unlabelling the gluing product of $\boldsymbol{R}$ and $\boldsymbol{S}$ can be labelled such that the resulting labelled graph is in the class. We extend this notion to bilabelled graphs. This definition is inspired by the inner-product on $\mathbb{C}^{n \times n}$ given by $\langle A, B \rangle := \operatorname{tr}(A^* B)$.

Definition 5.2.5. Let $t \geq 1$. A class of (t,t)-bilabelled graphs $\mathcal{S}$ is said to be *inner-product compatible* if, for all $\boldsymbol{R}, \boldsymbol{S} \in \mathcal{S}$, there is a bilabelled graph $\boldsymbol{T} \in \mathcal{S}$ such that $\operatorname{tr}(\boldsymbol{R}^* \cdot \boldsymbol{S}) = \operatorname{soe}(\boldsymbol{T})$.

Lemma 5.2.6. *Let $t \geq 1$. The classes $\mathcal{L}(t,t)$ and $\mathcal{L}^+(t,t)$ are inner-product compatible.*

Proof. Since $\mathcal{L}(t,t)$ is closed under matrix products and taking transposes, it suffices to show that, for every $\boldsymbol{S} \in \mathcal{L}(t,t)$, the graph $\operatorname{tr}(\boldsymbol{S})$ is the underlying unlabelled graph of some element of $\mathcal{L}(t,t)$, i.e. $\operatorname{tr}(\mathcal{L}(t,t)) \subseteq \operatorname{soe}(\mathcal{L}(t,t))$. Indeed, by Definition 3.2.7, for every (t,t)-bilabelled graphs $\boldsymbol{F}$ it holds that $\operatorname{tr}(\boldsymbol{F}) = \operatorname{soe}(\boldsymbol{I}^{1,t+1} \odot \cdots \odot \boldsymbol{I}^{t,2t} \odot \boldsymbol{F})$ where the $\boldsymbol{I}^{ij}$ are as in Observation 5.1.5. Since $\mathcal{L}(t,t)$ is closed under parallel composition with atomic graphs, the claim follows. For $\mathcal{L}^+(t,t)$, an analogous argument yields the claim. □

The following Theorem 5.2.7 is a straightforward generalisation of Theorem 4.3.24. Write $\mathbb{C}\mathcal{S}_G \subseteq \mathbb{C}^{V(G)^t \times V(G)^t}$ for the $\mathbb{C}$-vector space spanned by homomorphism tensors $\boldsymbol{S}_G$ for $\boldsymbol{S} \in \mathcal{S}$.

Theorem 5.2.7. *Let $t \geq 1$ and $\mathcal{S}$ be an inner-product compatible class of (t,t)-bilabelled graphs containing $\boldsymbol{J}$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over* $\operatorname{soe}(\mathcal{S})$,
2. *there exists a unitary pseudo-stochastic $\varphi\colon \mathbb{C}\mathcal{S}_G \to \mathbb{C}\mathcal{S}_H$ such that $\varphi(\boldsymbol{S}_G) = \boldsymbol{S}_H$ for all $\boldsymbol{S} \in \mathcal{S}$.*

This completes the preparations for the proof of Theorems 5.2.3 and 5.2.4.

Proof of Theorems 5.2.3 and 5.2.4. First suppose that G and H are homomorphism indistinguishable over $\mathcal{L}_t$. By comparing the operations from Definitions 5.1.10 and 5.2.1, it follows that $\mathbb{C}\mathcal{S}_G = \widehat{\mathcal{A}}^t_G$ for $\mathcal{S} = \mathcal{L}(t,t)$. By Lemma 5.2.6 and Theorem 5.2.7, G and H are homomorphism indistinguishable over $\mathcal{L}_t$ if, and only if, there is a unitary pseudo-stochastic $\varphi\colon \widehat{\mathcal{A}}^t_G \to \widehat{\mathcal{A}}^t_H$ satisfying $\varphi(\boldsymbol{S}_G) = \boldsymbol{S}_H$ for all $\boldsymbol{S} \in \mathcal{L}(t,t)$.

For all atomic $\boldsymbol{A} \in \mathcal{A}(t,t)$, it holds that $\varphi(\boldsymbol{A}_G) = \boldsymbol{A}_H$. Furthermore, since $\mathcal{L}(t,t)$ is closed under the action of $\mathfrak{S}_{2t}$, $\varphi(\boldsymbol{S}^\sigma_G) = \varphi((\boldsymbol{S}^\sigma)_G) = (\boldsymbol{S}^\sigma)_H = \boldsymbol{S}^\sigma_H$ for all permutations $\sigma \in \mathfrak{S}_{2t}$. Finally, for all $\boldsymbol{S}, \boldsymbol{T} \in \mathcal{L}(t,t)$ it holds that $\varphi(\boldsymbol{S}_G \cdot \boldsymbol{T}_G) = \varphi((\boldsymbol{S} \cdot \boldsymbol{T})_G) = \boldsymbol{S}_H \cdot \boldsymbol{T}_H$ and $\varphi(\boldsymbol{S}_G \odot \boldsymbol{T}_G) = \varphi((\boldsymbol{S} \odot \boldsymbol{T})_G) = \boldsymbol{S}_H \odot \boldsymbol{T}_H$. The homomorphism matrices $\boldsymbol{S}_G$ for $\boldsymbol{S} \in \mathcal{L}(t,t)$ span $\mathbb{C}\mathcal{S}_G = \widehat{\mathcal{A}}^t_G$. Hence, φ is a partial t-equivalence.

Conversely, every partial t-equivalence $\varphi\colon \widehat{\mathcal{A}}^t_G \to \widehat{\mathcal{A}}^t_H$ is such that $\varphi(\boldsymbol{S}_G) = \boldsymbol{S}_H$ for all $\boldsymbol{S} \in \mathcal{L}(t,t)$ by definition of $\mathcal{L}(t,t)$. With slight modifications, [126, Lemma 5.3] yields that φ is trace-preserving, which implies with $\varphi(J) = J$ that φ is sum-preserving. Hence, $\text{soe}(\boldsymbol{S}_G) = \text{soe}(\varphi(\boldsymbol{S}_G)) = \text{soe}(\boldsymbol{S}_H)$ for every $\boldsymbol{S} \in \mathcal{L}(t,t)$. Thus, G and H are homomorphism indistinguishable over $\mathcal{L}_t$. The proof of Theorem 5.2.4 is analogous. □

5.2.1 The Classes $\mathcal{L}_t$ and $\mathcal{L}_t^+$ and Graphs of Bounded Treewidth

In this section, the classes $\mathcal{L}_t$ and $\mathcal{L}_t^+$ are compared to the classes of graphs of bounded treewidth. Figure 5.1 depicts the relationships between these classes. The first result, Lemma 5.2.8, gives an upper bound on the treewidth of graphs in $\mathcal{L}_t^+$. Recall that graph classes are assumed to be closed under isomorphism.

Lemma 5.2.8. *For $t \geq 1$, $\mathcal{L}_t^+ \subseteq \mathcal{TW}_{3t-1}$.*

Proof. By structural induction, it is shown that every $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{L}^+(t,t)$ admits a tree decomposition $\beta \colon V(T) \to 2^{V(F)}$ of width at most $3t-1$ such that the labelled vertices $\boldsymbol{u}$ and $\boldsymbol{v}$ lie together in one bag, i.e. there exists $x \in V(T)$ such that $\{u_1, \dots, u_t, v_1, \dots, v_t\} \subseteq \beta(x)$.

In the base case, i.e. if $\boldsymbol{F} \in \mathcal{A}(t,t)$, then $\boldsymbol{F}$ has at most $2t$ vertices, which can all be placed in the single bag of a tree decomposition over the singleton tree.

For the inductive step, let $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$ and $\boldsymbol{F}' = (F', \boldsymbol{u}', \boldsymbol{v}')$ from $\mathcal{L}^+(t,t)$ be given. Suppose there are tree decompositions $\beta \colon V(T) \to 2^{V(F)}$ and $\beta' \colon V(T') \to 2^{V(F')}$ as in the inductive hypothesis. Let $x \in V(T)$ and $x' \in V(T')$ be such that the labelled vertices of $\boldsymbol{F}$ and $\boldsymbol{F}'$ lie in $\beta(x)$ and $\beta'(x')$ respectively. Let S be the tree obtained by taking the disjoint union of T, T', and a fresh vertex y, and connecting x and x' to y.

For the graph $\boldsymbol{F} \cdot \boldsymbol{F}'$, an S-decomposition is given by the function

$$\gamma \colon z \mapsto \begin{cases} \beta(z), & \text{if } z \in V(T), \\ \beta'(z), & \text{if } z \in V(T'), \\ \{u_1, \dots, u_t, v'_1, \dots, v'_t, v_1, \dots, v_t\}, & \text{if } z = y. \end{cases}$$

where one may note that $v_i = u'_i$ for every $i \in [t]$ in $\boldsymbol{F} \cdot \boldsymbol{F}'$. It is easy to check that Definition 2.1.1 is satisfied. The decomposition is of width $3t - 1$.

For the graph $\boldsymbol{F} \odot \boldsymbol{F}'$, an S-decomposition is given by the function

$$\gamma \colon z \mapsto \begin{cases} \beta(z), & \text{if } z \in V(T), \\ \beta'(z), & \text{if } z \in V(T'), \\ \{u_1, \dots, u_t, v_1, \dots, v_t\}, & \text{if } z = y. \end{cases}$$

where one may note that $u_i = u'_i$ and $v_i = v'_i$ for every $i \in [t]$ in $\boldsymbol{F} \odot \boldsymbol{F}'$. Again, it is easy to check that Definition 2.1.1 is satisfied. The decomposition is of width at most $3t - 1$. □

Lemma 5.2.8 in conjunction with Theorem 5.1.1 and Corollary 4.4.1 imply the first assertion of Theorem 5.0.1.

5.2.2 Further Relations between $\mathcal{TW}_t$, $\mathcal{PW}_t$, $\mathcal{L}_t$, and $\mathcal{L}_t^+$

This subsection is dedicated to some further relations between the classes of graphs of bounded treewidth or pathwidth, $\mathcal{L}_t$, and $\mathcal{L}_t^+$. These facts give independent proofs for the correspondence between the feasibility of the level-t Sherali–Adams relaxation (without non-negativity constraints), which corresponds to homomorphism indistinguishability over graphs of treewidth (pathwidth) at most $t-1$, as established in Theorem 4.0.1 and Corollary 4.4.1, and the feasibility of the level-t Lasserre relaxation with and without non-negativity constraints.

First of all, one may drop Equation (2.20) from $\mathrm{L}^t(G,H)$ with non-negativity constraints to obtain $\mathrm{SA}^{2t}(G,H)$ in its original form, i.e. with non-negativity constraints. This is paralleled by Lemma 5.2.9.

Lemma 5.2.9. *For $t \geq 1$, $\mathcal{TW}_{2t-1} \subseteq \mathcal{L}_t^+$.*

Proof. Let $F \in \mathcal{TW}_{2t-1}$. If $|V(F)| \leq 2t$, then there exists an atomic graph $\boldsymbol{F} \in \mathcal{A}(t,t)$ whose underlying unlabelled graph is isomorphic to F. Otherwise, by Lemma 2.1.2, there exists a tree decomposition $\beta\colon V(T) \to 2^{V(F)}$ of F such that $|\beta(v)| = 2t$ for all $v \in V(T)$ and $|\beta(s) \cap \beta(t)| = 2t-1$ for all $st \in E(T)$. It is shown by induction on $|V(T)|$ that, for every $r \in V(T)$, there exist $\boldsymbol{u} \in V(F)^t$, $\boldsymbol{v} \in V(F)^t$ with $\beta(r) = \{u_1, \dots, u_t, v_1, \dots, v_t\}$ such that $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{L}^+(t,t)$. Observe that this implies that the labels of $\boldsymbol{F}$ lie on distinct vertices of F.

In the base case, when $|V(T)| = 1$, the tuples $\boldsymbol{u}$ and $\boldsymbol{v}$ can be chosen arbitrarily subject to the desired condition. In this case, $\boldsymbol{F}$ is an atomic graph.

Let $|V(T)| \geq 2$ and $r \in V(T)$ be arbitrary. Write $s_1, \dots, s_\ell$ for the neighbours of r in T. First a bilabelled graph $\boldsymbol{F}_i \in \mathcal{L}^+(t,t)$ is constructed for each $i \in [\ell]$. Let T_i be the connected component of $T - r$ containing s_i. Let F_i be the induced subgraph of F on $\bigcup_{t \in V(T_i)} \beta(t)$. The restriction of β to $V(T_i)$ is a tree decomposition of F_i with the properties stated in the inductive hypothesis. Hence, there exist $\boldsymbol{u}^i \in V(F_i)^t$, $\boldsymbol{v}^i \in V(F_i)^t$ with $\beta(s_i) = \{u^i_1, \dots, u^i_t, v^i_1, \dots, v^i_t\}$ such that $\boldsymbol{F}_i := (F_i, \boldsymbol{u}^i, \boldsymbol{v}^i) \in \mathcal{L}^+(t,t)$.

Let $x_1, \dots, x_{2t}$ denote the vertices in $\beta(r)$. By permuting labels, it can be guaranteed that, for every $i \in [\ell]$, the tuples $u^i_1 \dots u^i_t v^i_1 \dots v^i_t$ and $x_1 \dots x_{2t}$ differ at precisely one index $j_i \in [2t]$. Recall the bilabelled graphs defined in Observation 5.1.5 and $\boldsymbol{K}^j$ from Equation (5.9) and Figure 5.3. Let $\boldsymbol{F}'_i := \boldsymbol{K}^{j_i} \cdot \boldsymbol{F}_i$ if $j_i \leq t$ and $\boldsymbol{F}'_i := \boldsymbol{F}_i \cdot \boldsymbol{K}^{j_i - t}$ otherwise. Intuitively, the bilabelled graph $\boldsymbol{F}'_i$ is obtained from $\boldsymbol{F}_i$ by adding a fresh vertex and moving the j_i-th label to this vertex. Since $\boldsymbol{F}_i \in \mathcal{L}^+(t,t)$ and $\boldsymbol{K}^{j_i} \in \mathcal{A}(t,t)$, it holds that $\boldsymbol{F}'_i \in \mathcal{L}^+(t,t)$. Finally, let $\boldsymbol{F} = \boldsymbol{F}'_1 \odot \dots \odot \boldsymbol{F}'_\ell \odot \bigodot_{x_i x_j \in E(F)} \boldsymbol{A}^{ij}$. □

Furthermore, it is easy to see that dropping the semidefiniteness constraint Equation (2.20) of $\mathrm{L}^t(G,H)$ turns this system essentially into $\mathrm{SA}^{2t}(G,H)$ without non-negativity constraints as defined in Definition 2.6.7. This is paralleled by Lemma 5.2.10.

Lemma 5.2.10. *For $t \geq 1$, $\mathcal{PW}_{2t-1} \subseteq \mathcal{L}_t$.*

Proof. Let $F \in \mathcal{PW}_{2t-1}$. If $|V(F)| \leq 2t$, then there exists an atomic graph $\boldsymbol{F} \in \mathcal{A}(t,t)$ whose underlying unlabelled graph is isomorphic to F. Otherwise, by Lemma 2.1.2, there exists a path decomposition $\beta\colon V(P) \to 2^{V(F)}$ such that $|\beta(v)| = 2t$ for all $v \in V(P)$ and $|\beta(s) \cap \beta(t)| = 2t-1$ for all $st \in E(P)$.

It is shown by induction on $|V(P)|$ that, for every vertex $r \in V(P)$ of degree at most one, there exist $\boldsymbol{u} \in V(F)^t$, $\boldsymbol{v} \in V(F)^t$ with $\beta(r) = \{u_1, \dots, u_t, v_1, \dots, v_t\}$ such that $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{L}_t$.

The inductive argument is very similar to the one in the proof of Lemma 5.2.9. Indeed, since the vertex r has at most one neighbour, it holds that $\ell \leq 1$ in the proof of Lemma 5.2.9. Thus the construction described there does not require arbitrary parallel compositions. □

Since the diagonal entries of a positive semidefinite matrix are necessarily non-negative, Equation (2.20) implies that any solution (y_I) to the level-t Lasserre system of equations is such that $y_I \geq 0$ for all $I \in \binom{V(G)\times V(H)}{\leq t}$. Hence, such a solution is a solution to $\mathrm{SA}^t(G,H)$ as well. This is paralleled by Lemma 5.2.11.

Lemma 5.2.11. *For $t \geq 1$, $\mathcal{TW}_{t-1} \subseteq \mathcal{L}_t$.*

Proof. Let $F \in \mathcal{TW}_{t-1}$. If $|V(F)| \leq t$, then there exists an atomic graph $\boldsymbol{F} \in \mathcal{A}(t,t)$ whose underlying unlabelled graph is isomorphic to F. Otherwise, by Lemma 2.1.2, there exists a tree decomposition $\beta\colon V(T) \to 2^{V(F)}$ of F such that $|\beta(v)| = t$ for all $v \in V(T)$ and $|\beta(s) \cap \beta(t)| = t-1$ for all $st \in E(T)$. It is shown by induction on $|V(T)|$ that, for every $r \in V(T)$, there exist $\boldsymbol{u} = u_1 \dots u_t \in V(F)^t$ with $\beta(r) = \{u_1, \dots, u_t\}$ such that $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{u}) \in \mathcal{L}(t,t)$.

In the base case, when $|V(T)| = 1$, the tuple $\boldsymbol{u}$ can be chosen arbitrarily and $\boldsymbol{F}$ is an atomic graph.

Let $|V(T)| \geq 2$ and $r \in V(T)$ be arbitrary. Write $s_1, \dots, s_\ell$ for the neighbours of r in T. First a graph $\boldsymbol{F}_i \in \mathcal{L}(t,t)$ is constructed for each $i \in [\ell]$. Let T_i be the connected component of $T - r$ containing s_i. Let F_i be the induced subgraph of F on $\bigcup_{t \in V(T_i)} \beta(t)$. The restriction of β to $V(T_i)$ is a tree decomposition of F_i with the properties listed in the inductive hypothesis. Hence, there exist $\boldsymbol{u}^i = u^i_1 \dots u^i_t \in V(F_i)^t$ with $\beta(s_i) = \{u^i_1, \dots, u^i_t\}$ such that $\boldsymbol{F}_i := (F_i, \boldsymbol{u}^i, \boldsymbol{u}^i) \in \mathcal{L}(t,t)$.

Let $x_1, \dots, x_t$ denote the vertices in $\beta(r)$. By permuting labels, it can be guaranteed that, for every $i \in [\ell]$, the tuples $u^i_1 \dots u^i_t$ and $x_1 \dots x_t$ differ at precisely one index $j_i \in [t]$. Recall the bilabelled graphs defined in Observation 5.1.5 and $\boldsymbol{K}_j$ from Equation (5.9) and Figure 5.3. Let $\boldsymbol{F}'_i := \boldsymbol{I}^{j_i, t+j_i} \odot (\boldsymbol{K}_{j_i} \cdot \boldsymbol{F} \cdot \boldsymbol{K}_{j_i})$. By construction, $\boldsymbol{F}'_i \in \mathcal{L}(t,t)$. The labelled vertices of $\boldsymbol{F}_i$ differ from those of $\boldsymbol{F}_i$ in x_{j_i}. Finally, let

$$\boldsymbol{F} := (\boldsymbol{I}^{1,t+1} \odot \dots \odot \boldsymbol{I}^{t,2t}) \odot (\boldsymbol{F}'_1 \cdot \dots \cdot \boldsymbol{F}'_\ell) \odot \bigodot_{x_i x_j \in E(F)} \boldsymbol{A}^{ij}.$$

This graph is as desired. □

5.2.3 Bilabelled Minors

In this section, we introduce minors of bilabelled graphs. As a corollary, we show in Theorem 5.2.12 that the classes of unlabelled graphs $\mathcal{L}_t$ and $\mathcal{L}_t^+$ are minor-closed. Hence, by the Robertson–Seymour Theorem [155] and [154], membership in $\mathcal{L}_t$ and $\mathcal{L}_t^+$ can be tested in polynomial time for every fixed $t \geq 1$. Furthermore, bilabelled minors will be used in Section 5.2.4 to describe $\mathcal{L}_1$.

Theorem 5.2.12. *For $t \geq 1$, the classes $\mathcal{L}_t$ and $\mathcal{L}_t^+$ are closed undertaking minors and disjoint unions.*

In order to prove Theorem 5.2.12, we introduce bilabelled analogues of graph minors.

Definition 5.2.13. Let $k, \ell \in \mathbb{N}$. Let $\boldsymbol{M}, \boldsymbol{F} \in \mathcal{G}(k, \ell)$. Then $\boldsymbol{M}$ is a *bilabelled minor* of $\boldsymbol{F}$, in symbols $\boldsymbol{M} \leq \boldsymbol{F}$, if it can be obtained from $\boldsymbol{F}$ by applying a sequence of the following *bilabelled minor operations*:

1. edge contraction,
2. edge deletion,
3. deletion of unlabelled vertices,

A family of bilabelled graphs is *minor-closed* if it is closed under taking bilabelled minors.

Note that, for $(0,0)$-bilabelled graphs, i.e. unlabelled graphs, Definition 5.2.13 and the standard definition of graph minors coincide.

Example 5.2.14. Let $t \geq 1$. The class of atomic graphs $\mathcal{A}(t,t)$ as defined in Definition 3.2.10 is minor-closed.

We proceed to prove various lemmas characterising how bilabelled minors behave under the operations applied to bilabelled graphs, namely labelling and unlabelling and series and parallel composition.

Lemma 5.2.15 (Minor Unlabelling Lemma)**.** *Let $k, \ell \in \mathbb{N}$. Let $\boldsymbol{M} \leq \boldsymbol{F} \in \mathcal{G}(k, \ell)$. Then* $\operatorname{soe}(\boldsymbol{M}) \leq \operatorname{soe}(\boldsymbol{F})$.

Proof. It is argued by induction on the number of bilabelled minor operations necessary to transform $\boldsymbol{F}$ into $\boldsymbol{M}$. If $\boldsymbol{M} = \boldsymbol{F}$, then $\operatorname{soe}(\boldsymbol{M}) = \operatorname{soe}(\boldsymbol{F})$, and the claim follows. Suppose that $\boldsymbol{M} \leq \boldsymbol{M}' \leq \boldsymbol{F}$ where $\boldsymbol{M}'$ can be transformed into $\boldsymbol{M}$ by applying a single minor operation and $\boldsymbol{M}'$ is minimal among all such graphs with respect to the number of minor operations necessary to derive it from $\boldsymbol{F}$. By the inductive hypothesis, $\operatorname{soe}(\boldsymbol{M}') \leq \operatorname{soe}(\boldsymbol{F})$. Since bilabelled minor operations are more restrictive than minor operations, any operation of Definition 5.2.13 carried out on $\boldsymbol{M}'$ can be applied to $\operatorname{soe}(\boldsymbol{M}')$. It follows that $\operatorname{soe}(\boldsymbol{M}) \leq \operatorname{soe}(\boldsymbol{F})$. □

Lemma 5.2.16 (Minor Labelling Lemma). *Let $k, \ell \in \mathbb{N}$. Let $\boldsymbol{F} \in \mathcal{G}(k, \ell)$ and $M \in \mathcal{G}$. If $M \leq \operatorname{soe}(\boldsymbol{F})$, then there exists a bilabelled minor $\boldsymbol{M} \leq \boldsymbol{F}$ such that $\operatorname{soe}(\boldsymbol{M})$ is the disjoint union of M and potential isolated vertices which are labelled in $\boldsymbol{M}$.*

Proof. It is argued by induction on the number of minor operations needed to transform $\operatorname{soe}(\boldsymbol{F})$ into M. If $M = \operatorname{soe}(\boldsymbol{F})$, let $\boldsymbol{M} := \boldsymbol{F}$. Now suppose there are $M \leq M' \leq \operatorname{soe}(\boldsymbol{F})$ such that M' can be transformed into M by applying a single minor operation. Then there exists $\boldsymbol{M}' \leq \boldsymbol{F}$ such that $\operatorname{soe}(\boldsymbol{M}')$ is the disjoint union of M' and potential isolated vertices. Distinguish cases:

- M is obtained from M' by deleting or contracting an edge e. Then e has a counterpart in $\boldsymbol{M}'$ since $\operatorname{soe}(\boldsymbol{M}')$ contains M'. Contracting or deleting the edge there yields the desired $\boldsymbol{M}$.
- M is obtained from M' by deleting a vertex v. If v is unlabelled in $\boldsymbol{M}'$, then it can be deleted from $\boldsymbol{M}'$ yielding $\boldsymbol{M}$. If v is labelled in $\boldsymbol{M}'$, remove all edges incident to v and let $\boldsymbol{M}$ be the resulting graph. In this case, $\operatorname{soe}(\boldsymbol{M})$ is the disjoint union of M and an isolated vertex. □

Intuitively, the following Lemmas 5.2.17 and 5.2.18 assert that minor operations commute with bilabelled graph multiplication.

Lemma 5.2.17 (Minor Parallel Composition Lemma). *Let $k, \ell \in \mathbb{N}$ and $\boldsymbol{P}_1, \boldsymbol{P}_2 \in \mathcal{G}(k, \ell)$.*

1. *If $\boldsymbol{M}_1$ is a minor of $\boldsymbol{P}_1$ and $\boldsymbol{M}_2$ is a minor of $\boldsymbol{P}_2$, then $\boldsymbol{M}_1 \odot \boldsymbol{M}_2$ is a minor of $\boldsymbol{P}_1 \odot \boldsymbol{P}_2$.*
2. *If $\boldsymbol{K}$ is a minor of $\boldsymbol{P}_1 \odot \boldsymbol{P}_2$, then there exist $\boldsymbol{M}_1, \boldsymbol{M}_2 \in \mathcal{G}(k, \ell)$ such that $\boldsymbol{K} = \boldsymbol{M}_1 \odot \boldsymbol{M}_2$, $\boldsymbol{M}_1$ is a minor of $\boldsymbol{P}_1$, and $\boldsymbol{M}_2$ is a minor of $\boldsymbol{P}_2$.*

Proof. For the first claim, it is argued by induction on the sum of the number of minor operations applied to transform $\boldsymbol{P}_1$ into $\boldsymbol{M}_1$ and $\boldsymbol{P}_2$ into $\boldsymbol{M}_2$. For the base case, $\boldsymbol{M}_1 = \boldsymbol{P}_1$ and $\boldsymbol{M}_2 = \boldsymbol{P}_2$, and the claim follows trivially.

Now suppose that $\boldsymbol{M}_1$ is obtained from $\boldsymbol{M}_1'$, a minor of $\boldsymbol{P}_1$, by applying a single minor operation. Suppose inductively that $\boldsymbol{M}_1' \odot \boldsymbol{M}_2$ is a minor of $\boldsymbol{P}_1 \odot \boldsymbol{P}_2$. Distinguish cases:

- $\boldsymbol{M}_1$ is obtained from $\boldsymbol{M}_1'$ by contracting an edge e. In $\boldsymbol{M}_1' \odot \boldsymbol{M}_2$, this edge is either a loop or a proper edge. In the former case, it can be deleted, in the latter case, it can be contracted, yielding in both cases $\boldsymbol{M}_1 \odot \boldsymbol{M}_2$.
- $\boldsymbol{M}_1$ is obtained from $\boldsymbol{M}_1'$ by deleting an edge e. In $\boldsymbol{M}_1' \odot \boldsymbol{M}_2$, this edge is either a loop or a proper edge. In both cases, it can be deleted yielding $\boldsymbol{M}_1 \odot \boldsymbol{M}_2$.
- $\boldsymbol{M}_1$ is obtained from $\boldsymbol{M}_1'$ by deleting an unlabelled vertex v. Then v is unlabelled in $\boldsymbol{M}_1' \odot \boldsymbol{M}_2$ and can be deleted. The resulting graph is $\boldsymbol{M}_1 \odot \boldsymbol{M}_2$.

For the second claim, it is argued by induction on the number of minor operations necessary to transform $\boldsymbol{P}_1 \odot \boldsymbol{P}_2$ into $\boldsymbol{K}$. For the base case, if $\boldsymbol{K} = \boldsymbol{P}_1 \odot \boldsymbol{P}_2$, let $\boldsymbol{M} := \boldsymbol{K}$, $\boldsymbol{M}_1 := \boldsymbol{P}_1$, and $\boldsymbol{M}_2 := \boldsymbol{P}_2$.

Now suppose that $\boldsymbol{K}$ is a minor of $\boldsymbol{P}_1 \odot \boldsymbol{P}_2$. Then there exists a bilabelled minor $\boldsymbol{K}'$ of $\boldsymbol{P}_1 \odot \boldsymbol{P}_2$ such that $\boldsymbol{K}$ is obtained from $\boldsymbol{K}'$ by applying a single minor operation.

By the induction hypothesis, there exist $\boldsymbol{M}_1'$ and $\boldsymbol{M}_2'$ such that the assertions of this lemma are satisfied. Distinguish cases:

- $\boldsymbol{K}$ is obtained from $\boldsymbol{K}'$ by deleting or contracting an edge e.
  The edge e may lie in both $\boldsymbol{M}_1'$ and $\boldsymbol{M}_2'$ or in only one of the two graphs. In either case, construct $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ by respectively deleting or contracting the edge in $\boldsymbol{M}_1'$ and $\boldsymbol{M}_2'$ or leaving the graph unchanged if it does not contain the edge.
- $\boldsymbol{K}$ is obtained from $\boldsymbol{K}'$ by deleting an unlabelled vertex v.
  Since no vertex is unlabelled under parallel composition, the vertex v is also unlabelled in the graph $\boldsymbol{M}_1'$ or $\boldsymbol{M}_2'$ which it contains. It follows that v can be deleted from $\boldsymbol{M}_i'$ leaving the other graph untouched. This yields $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$. □

Lemma 5.2.18 (Minor Series Composition Lemma). *Let $k, \ell, j \in \mathbb{N}$. Let $\boldsymbol{P}_1 \in \mathcal{G}(k, \ell)$ and $\boldsymbol{P}_2 \in \mathcal{G}(\ell, j)$.*

1. *If $\boldsymbol{M}_1$ is a minor of $\boldsymbol{P}_1$ and $\boldsymbol{M}_2$ is a minor of $\boldsymbol{P}_2$, then $\boldsymbol{M}_1 \cdot \boldsymbol{M}_2$ is a minor of $\boldsymbol{P}_1 \cdot \boldsymbol{P}_2$.*
2. *If $\boldsymbol{K}$ is a minor of $\boldsymbol{P}_1 \cdot \boldsymbol{P}_2$, then there exists a $\boldsymbol{M} \in \mathcal{G}(k, j)$, $\boldsymbol{M}_1 \in \mathcal{G}(k, \ell)$, and $\boldsymbol{M}_2 \in \mathcal{G}(\ell, j)$ such that*
   a) *$\boldsymbol{M}$ is the disjoint union of $\boldsymbol{K}$ and potential isolated unlabelled vertices, which are labelled in $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$,*
   b) *$\boldsymbol{M} = \boldsymbol{M}_1 \cdot \boldsymbol{M}_2$, and*
   c) *$\boldsymbol{M}_1$ is a minor of $\boldsymbol{P}_1$ and $\boldsymbol{M}_2$ is a minor of $\boldsymbol{P}_2$.*

Proof. The proof of the first claim is analogous to the proof of the first claim of Lemma 5.2.17.

For the second claim, it is argued by induction on the number of minor operations necessary to transform $\boldsymbol{P}_1 \cdot \boldsymbol{P}_2$ into $\boldsymbol{K}$. For the base case, if $\boldsymbol{K} = \boldsymbol{P}_1 \cdot \boldsymbol{P}_2$, let $\boldsymbol{M} := \boldsymbol{K}$, $\boldsymbol{M}_1 := \boldsymbol{P}_1$, and $\boldsymbol{M}_2 := \boldsymbol{P}_2$.

Now suppose that $\boldsymbol{K}$ is a minor of $\boldsymbol{P}_1 \cdot \boldsymbol{P}_2$. Then there exists a (k, j)-bilabelled graph $\boldsymbol{K}'$ such that $\boldsymbol{K}'$ is a minor of $\boldsymbol{P}_1 \cdot \boldsymbol{P}_2$ and $\boldsymbol{K}$ is obtained from $\boldsymbol{K}'$ by applying a single minor operation. By the induction hypothesis, there exist $\boldsymbol{M}'$, $\boldsymbol{M}_1'$, and $\boldsymbol{M}_2'$ such that Items 2a to 2c are satisfied. Distinguish cases:

- $\boldsymbol{K}$ is obtained from $\boldsymbol{K}'$ by deleting or contracting an edge e.
  Define $\boldsymbol{M}$ by deleting/contracting the same edge in $\boldsymbol{M}'$. The edge e may lie in both $\boldsymbol{M}_1'$ and $\boldsymbol{M}_2'$ or only in one of the two graphs. In the first case, both endpoints of e are labelled in both graphs. In either case, construct $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ by respectively deleting or contracting the edge in $\boldsymbol{M}_1'$ and $\boldsymbol{M}_2'$ or leaving the graph unchanged if it does not contain the edge.
- $\boldsymbol{K}$ is obtained from $\boldsymbol{K}'$ by deleting an unlabelled vertex v.
  If v is among the unlabelled vertices of $\boldsymbol{M}_i'$ for $i \in \{1, 2\}$, then define $\boldsymbol{M}$ by deleting v from $\boldsymbol{M}'$. It follows that v can be deleted from $\boldsymbol{M}_i'$ leaving the other graph untouched. This yields $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$.

If otherwise v is among the vertices at which $\boldsymbol{M}'_1$ and $\boldsymbol{M}'_2$ are glued together, then define $\boldsymbol{M}$ as the graph obtained from $\boldsymbol{M}'$ by deleting all edges incident with v but retaining the vertex v. By the inductive hypothesis, $\boldsymbol{M}'$ is the disjoint union of $\boldsymbol{K}'$ and isolated unlabelled vertices and via the aforementioned construction the same holds for $\boldsymbol{M}$ and $\boldsymbol{K}$. Note that v is neither in-labelled in $\boldsymbol{M}'_1$ nor out-labelled in $\boldsymbol{M}'_2$ as it would otherwise be labelled in $\boldsymbol{M}$. Delete all edges incident to v in both $\boldsymbol{M}'_1$ and $\boldsymbol{M}'_2$. The resulting $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ satisfy $\boldsymbol{M} = \boldsymbol{M}_1 \cdot \boldsymbol{M}_2$, as desired. □

With these general facts at hand, we proceed to show the following about the graph classes $\mathcal{L}(t,t)$ and $\mathcal{L}^+(t,t)$.

Theorem 5.2.19. *For $t \geq 1$, the classes $\mathcal{L}(t,t)$ and $\mathcal{L}^+(t,t)$ are closed under taking bilabelled minors.*

Proof. By induction on the structure of elements $\boldsymbol{F} \in \mathcal{L}(t,t)$, it is proven that if $\boldsymbol{K} \leq \boldsymbol{F}$, then also $\boldsymbol{K} \in \mathcal{L}(t,t)$. For $\mathcal{L}^+(t,t)$, the proof is very similar, requiring fewer case distinctions. It is therefore omitted. If $\boldsymbol{F}$ is atomic, then all its minors are atomic by Example 5.2.14. This constitutes the base case of the induction.

If $\boldsymbol{F} = \boldsymbol{F}_1 \odot \boldsymbol{F}_2$ for $\boldsymbol{F}_1 \in \mathcal{A}(t,t)$, $\boldsymbol{F}_2 \in \mathcal{L}(t,t)$ of lesser complexity, and $\boldsymbol{K} \leq \boldsymbol{F}$, then, by Lemma 5.2.17, there exist $\boldsymbol{K}_1 \leq \boldsymbol{F}_1$ and $\boldsymbol{K}_2 \leq \boldsymbol{F}_2$ such that $\boldsymbol{K} = \boldsymbol{K}_1 \odot \boldsymbol{K}_2$. By Example 5.2.14, $\boldsymbol{K}_1$ is atomic and, by the inductive hypothesis, $\boldsymbol{K}_2 \in \mathcal{L}(t,t)$. Hence, $\boldsymbol{K} \in \mathcal{L}(t,t)$.

If $\boldsymbol{F} = \boldsymbol{F}_1 \cdot \boldsymbol{F}_2$ for two $\boldsymbol{F}_1, \boldsymbol{F}_2 \in \mathcal{L}(t,t)$ of lesser complexity and $\boldsymbol{K} \leq \boldsymbol{F}$, then, by Lemma 5.2.18, there exist $\boldsymbol{M}$, $\boldsymbol{M}_1$, $\boldsymbol{M}_2$ such that $\boldsymbol{M}_1 \leq \boldsymbol{F}_1$, $\boldsymbol{M}_2 \leq \boldsymbol{F}_2$, and $\boldsymbol{M} = \boldsymbol{M}_1 \cdot \boldsymbol{M}_2$ is the disjoint union of $\boldsymbol{K}$ and potential isolated unlabelled vertices which are labelled both in $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$. By the inductive hypothesis, $\boldsymbol{M}_1, \boldsymbol{M}_2 \in \mathcal{L}(t,t)$. It remains to remove the isolated vertices. Suppose that the i-th out-label of $\boldsymbol{M}_1$ and the i-th in-label of $\boldsymbol{M}_2$ are carried by isolated vertices. Then graph $(\boldsymbol{I}^{1,t+1+i} \odot \boldsymbol{M}_1) \cdot \boldsymbol{M}_2$ does not contain this isolated vertex, since taking the parallel composition with $\boldsymbol{I}^{1,t+1+i}$ as defined in Observation 5.1.5 amounts to gluing it to the first in-labelled vertex of $\boldsymbol{M}_1$. Observe that $\boldsymbol{I}^{1,t+1+i} \odot \boldsymbol{M}_1 \in \mathcal{L}(t,t)$. Proceeding in this fashion, one can construct $\boldsymbol{K}_1, \boldsymbol{K}_2 \in \mathcal{L}(t,t)$ such that $\boldsymbol{K} = \boldsymbol{K}_1 \cdot \boldsymbol{K}_2 \in \mathcal{L}(t,t)$, as desired.

If $\boldsymbol{F} = \boldsymbol{F}_1^\sigma$ for $\sigma \in \mathfrak{S}_{2t}$ and $\boldsymbol{F}_1 \in \mathcal{L}(t,t)$ of lesser complexity and $\boldsymbol{K} \leq \boldsymbol{F}$, then $\boldsymbol{K}^{\sigma^{-1}} \leq \boldsymbol{F}_1$ and $\boldsymbol{K}^{\sigma^{-1}} \in \mathcal{L}(t,t)$ by the inductive hypothesis. Hence, $\boldsymbol{K} \in \mathcal{L}(t,t)$, as desired. □

This concludes the preparations for the proof of Theorem 5.2.12.

Proof of Theorem 5.2.12. For (t,t)-bilabelled graphs $\boldsymbol{F}$ and $\boldsymbol{F}'$ and $\boldsymbol{J} \in \mathcal{A}(t,t)$ as defined in Observation 5.1.5, the graph underlying $\boldsymbol{F} \cdot \boldsymbol{J} \cdot \boldsymbol{F}'$ is isomorphic to the disjoint union of the graphs underlying $\boldsymbol{F}$ and $\boldsymbol{F}'$. Hence, the graph classes $\mathcal{L}_t$ and $\mathcal{L}_t^+$ are closed under disjoint unions.

It remains to extend Theorem 5.2.19 to the classes of unlabelled graphs $\mathcal{L}_t$ and $\mathcal{L}_t^+$. By Lemma 5.2.16, if an unlabelled graph M is a minor of $\operatorname{soe}(\boldsymbol{F})$ for some $\boldsymbol{F} \in \mathcal{L}(t,t)$, then there exists $\boldsymbol{M} \leq \boldsymbol{F}$ such that $\operatorname{soe}(\boldsymbol{M})$ is the disjoint union of M and potential isolated vertices which are labelled in $\boldsymbol{M}$. By Theorem 5.2.19, $\boldsymbol{M} \in \mathcal{L}(t,t)$. As in the proof of Theorem 5.2.19, the potential isolated vertices can be identified with other vertices labelled in $\boldsymbol{M}$ by taking the parallel composition of this graph with atomic graphs. Hence, it may be assumed that $M = \operatorname{soe}(\boldsymbol{M})$. This yields the claim. ☐

5.2.4 The Classes $\mathcal{L}_1$ and $\mathcal{L}_1^+$

The classes $\mathcal{L}_1$ and $\mathcal{L}_1^+$ can be identified as the class of outerplanar graphs and as the class of graphs of treewidth at most two, respectively. The following Theorem 5.2.20 parallels [126, Result 2, Lemma 10.1] asserting that two graphs G and H are indistinguishable under the 2-WL algorithm if, and only if, $\mathrm{L}^1(G,H)$ has a non-negative real solution.

Theorem 5.2.20. *The classes* $\mathcal{L}_1^+$ *and* $\mathcal{TW}_2$ *coincide.*

Proof. Given Lemmas 5.2.8 and 5.2.9, it suffices to show that if a graph F is such that $\operatorname{tw}(F) = 2$, then there is a graph $\boldsymbol{F} \in \mathcal{L}_1^+$ whose underlying unlabelled graph is isomorphic to F. In order to reduce the technical overhead, we suppose that F is a simple graph. The case when F contains loops can be dealt with similarly.

By Lemma 2.1.2, there exists a tree decomposition $\beta\colon V(T) \to 2^{V(F)}$ of F such that $|\beta(v)| = 3$ for all $v \in V(T)$ and $|\beta(s) \cap \beta(t)| = 2$ for all $st \in E(T)$. It is shown by induction on $|V(T)|$ that, for every $r \in V(T)$ and $x \neq y \in \beta(r)$, the graph $\boldsymbol{F} = (F, x, y)$ is in $\mathcal{L}^+(1,1)$.

If $|V(T)| = 1$, write $\{x,y,z\}$ for the unique bag. Since F has treewidth 2, it is isomorphic to the 3-vertex complete graph which is the underlying unlabelled graph of $\boldsymbol{A}^{12} \odot (\boldsymbol{A}^{12} \cdot \boldsymbol{A}^{12})$, cf. Observation 5.2.2, which is contained in $\mathcal{L}^+(1,1)$ by construction.

Assuming $|V(T)| \geq 2$, let $r \in V(T)$ be arbitrary. Write $\beta(r) = \{x_1, x_2, x_3\}$. Partition the neighbours of r in T in three sets X_1, X_2, X_3 such that $s \in X_i$ if, and only if, $x_i \in \beta(r) \setminus \beta(s)$ for $i \in [3]$.

For every neighbour s of r, let T_s be the connected component of $T - r$ containing s. Let F_s be the induced subgraph of F on $\bigcup_{t \in V(T_s)} \beta(t)$. The restriction of β to $V(T_s)$ is a tree decomposition of F_s with the properties listed in the inductive hypothesis. Hence, for every s, there exists $\boldsymbol{F}_s \in \mathcal{L}^+(1,1)$ as stipulated. By permuting labels, it may be supposed that, for every $s \in X_1$, the labels of $\boldsymbol{F}_s$ lie on x_2x_3, for $\boldsymbol{F}_s$ with

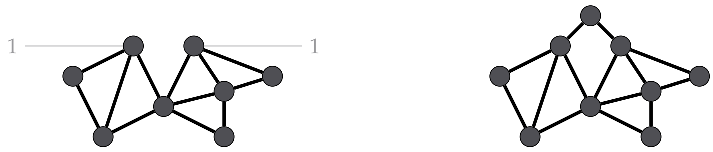


Figure 5.6: A $(1, 1)$-bilabelled outerplanar graph and its expansion.

$s \in X_2$ on x_1x_3, and for $\boldsymbol{F}_s$ with $s \in X_3$ on x_1x_2. For $i \in [3]$, let

$$\boldsymbol{F}_i := \begin{cases} \bigodot_{s \in X_i} \boldsymbol{F}_s, & \text{if } X_i \neq \emptyset, \\ \boldsymbol{A}^{12}, & \text{if } X_i = \emptyset \text{ and the two vertices in } \beta(r) \setminus \{x_i\} \text{ are adjacent}, \\ \boldsymbol{J}, & \text{otherwise.} \end{cases}$$

Finally, let $\boldsymbol{F} := \boldsymbol{F}_2 \odot (\boldsymbol{F}_3 \cdot \boldsymbol{F}_1)$. This graph is as desired if x_1, x_3 are required to be labelled. For other choices of labels, $\boldsymbol{F}_1, \boldsymbol{F}_2, \boldsymbol{F}_3$ can be permuted and if necessary transposed yielding any desired labelling. □

A graph F is *outerplanar* if it does not have K_4 or $K_{2,3}$ as a minor. Equivalent, it is outerplanar if it has a planar drawing such that all its vertices lie on the same face [171], i.e. the *outer face* of the outerplanar drawing. Write $\mathcal{OP}$ for the class of all outerplanar graphs. The main result of this section is the following.

Theorem 5.2.21. *The classes $\mathcal{L}_1$ and $\mathcal{OP}$ coincide.*

Theorem 5.2.21 yields the following Corollary 5.2.22 via Theorem 5.1.1:

Corollary 5.2.22. *Two simple graphs G and H are homomorphism indistinguishable over all outerplanar graphs if, and only if,* $\mathrm{L}^1(G, H)$ *has a real solution.*

Towards proving Theorem 5.2.21, we define a class of $(1, 1)$-bilabelled graphs whose underlying unlabelled graphs are outerplanar.

Definition 5.2.23. The *expansion* of a $(1, 1)$-bilabelled graphs $\boldsymbol{F} = (F, u, v)$ is the graph F' obtained from F by adding a path of length two between u and v, i.e. $V(F') := V(F) \sqcup \{x\}$ and $E(F') := E(F) \sqcup \{ux, xv\}$. Write $\mathcal{OP}(1, 1) \subseteq \mathcal{G}(1, 1)$ for the class of $(1, 1)$-bilabelled graphs $\boldsymbol{F}$ whose expansion is outerplanar.

In terms of the outerplanar drawing of F, Definition 5.2.23 means that the two labelled vertices u and v of $\boldsymbol{F} = (F, u, v)$ must occur consecutively on the outer face of the drawing. See Figure 5.6 for an example.

Note that the above definition implies that, for all $\boldsymbol{F} = (F, u, v) \in \mathcal{OP}(1, 1)$, the underlying unlabelled graph F is outerplanar as it is a minor of the expansion of $\boldsymbol{F}$, i.e. $\operatorname{soe}(\mathcal{OP}(1, 1)) \subseteq \mathcal{OP}$. If the two labels of $\boldsymbol{F}$ coincide, then its expansion is obtained by adding a dangling edge and outerplanar if, and only if, the underlying unlabelled graph of $\boldsymbol{F}$ is outerplanar. Hence, $\mathcal{OP} = \operatorname{soe}(\mathcal{OP}(1, 1))$.

(a) The graph $\boldsymbol{P} = \boldsymbol{A} \cdot \boldsymbol{A}$. **(b)** The graph $\boldsymbol{C}$.

Figure 5.7: Bilabelled graphs from the proof of Lemma 5.2.24.

Write $\boldsymbol{A}$ and $\boldsymbol{I}$ for the $(1,1)$-bilabelled graphs corresponding to the adjacency matrix and the identity matrix respectively. In the notation of Observation 5.1.5, $\boldsymbol{A} = \boldsymbol{A}^{12}$ and $\boldsymbol{I} = \boldsymbol{I}^{12}$. These graphs are depicted by Figure 5.4.

Lemma 5.2.24. *The class $\mathcal{OP}(1,1)$ possesses the following closure properties:*

1. *If $\boldsymbol{F} \in \mathcal{OP}(1,1)$, then $\boldsymbol{F}^* \in \mathcal{OP}(1,1)$.*
2. *If $\boldsymbol{F} \in \mathcal{OP}(1,1)$, then $\boldsymbol{A} \odot \boldsymbol{F} \in \mathcal{OP}(1,1)$ and $\boldsymbol{I} \odot \boldsymbol{F} \in \mathcal{OP}(1,1)$.*
3. *If $\boldsymbol{F}_1, \boldsymbol{F}_2 \in \mathcal{OP}(1,1)$, then $\boldsymbol{F}_1 \cdot \boldsymbol{F}_2 \in \mathcal{OP}(1,1)$.*

Proof. The first claim is purely syntactical. The underlying unlabelled graphs of $\boldsymbol{F}$ and $\boldsymbol{F}^*$ are isomorphic and so are their expansions. Thus, $\boldsymbol{F}^* \in \mathcal{OP}$ if $\boldsymbol{F} \in \mathcal{OP}$.

For the second claim, first consider the case when the labels of $\boldsymbol{F}$ coincide. Then $\boldsymbol{I} \odot \boldsymbol{F} = \boldsymbol{F}$ and $\boldsymbol{A} \odot \boldsymbol{F}$ differs from $\boldsymbol{F}$ only in the loop at the labelled vertex. Hence, $\boldsymbol{A} \odot \boldsymbol{F}$ is in $\mathcal{OP}(1,1)$. Now consider the case when the labelled vertices of $\boldsymbol{F}$ are distinct. Write F for the unlabelled graph underlying $\boldsymbol{F}$ and F' for the expansion of $\boldsymbol{F}$. It can be easily seen that the graphs underlying $\boldsymbol{A} \odot \boldsymbol{F}$ and $\boldsymbol{I} \odot \boldsymbol{F}$ are minors of F' and thus outerplanar. The expansion of $\boldsymbol{I} \odot \boldsymbol{F}$ is a minor of the expansion of $\boldsymbol{A} \odot \boldsymbol{F}$. Thus, it suffices to argue that the expansion of $\boldsymbol{A} \odot \boldsymbol{F}$ is outerplanar.

Write K for the unlabelled graph underlying $\boldsymbol{A} \odot \boldsymbol{F}$ and K' for the expansion of $\boldsymbol{A} \odot \boldsymbol{F}$. Since K and F' are outerplanar, any K_4-minor of K' can be obtained from K' without contracting the triangle induced by the labelled vertices of $\boldsymbol{A} \odot \boldsymbol{F}$ and the vertex added by expansion. This cannot be since the latter vertex is of degree two. By the same argument, since $K_{2,3}$ is triangle-free, the graph K' does not contain any $K_{2,3}$-minor either.

For the third claim, let F denote the graph underlying $\boldsymbol{F}_1 \cdot \boldsymbol{F}_2$. Let y denote the vertex at which $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are glued together and write x, z for the vertices labelled in $\boldsymbol{F}_1 \cdot \boldsymbol{F}_2$. The graph $F - y$ is disconnected. Hence, if K_4 or $K_{2,3}$ is a minor of F, then $\boldsymbol{F}_1$ or $\boldsymbol{F}_2$ are not outerplanar. Hence, F is outerplanar.

Write F' for the expansion of $\boldsymbol{F} := \boldsymbol{F}_1 \cdot \boldsymbol{F}_2$. In symbols, $F' = \operatorname{soe}(\boldsymbol{P} \odot \boldsymbol{F})$ where $\boldsymbol{P}$ is the bilabelled graph in Figure 5.7a. For $K \in \{K_4, K_{2,3}\}$, observe the following: If F' contains K as a minor, then, by Lemma 5.2.16, there exists a bilabelled minor $\boldsymbol{K} \leq \boldsymbol{P} \odot \boldsymbol{F}$ such that $\operatorname{soe}(\boldsymbol{K})$ is the disjoint union of K and potential isolated vertices which are labelled in $\boldsymbol{K}$. By Lemma 5.2.17, $\boldsymbol{K}$ can be written as $\boldsymbol{K} = \boldsymbol{K}_1 \odot \boldsymbol{K}_2$ such that $\boldsymbol{K}_1 \leq \boldsymbol{P}$ and $\boldsymbol{K}_2 \leq \boldsymbol{F}$. The graph $\boldsymbol{P}$ has six bilabelled minors. Distinguish cases:

1. If $\boldsymbol{K}_1 = \boldsymbol{P}$, then $K = K_{2,3}$. The labels of $\boldsymbol{K}_2$ must lie on distinct vertices because $\boldsymbol{K}$ does not contain any vertices of degree one. Furthermore, the labelled

vertices in $\boldsymbol{K}$ must be connected via a path of length two with an intermediate vertex of degree two. Hence, $\boldsymbol{K}_2 = \boldsymbol{C}$ where $\boldsymbol{C}$ is the graph in Figure 5.7b.

2. If $\boldsymbol{K}_1 = \boldsymbol{A}$, then the labels of $\boldsymbol{K}_2$ must lie on distinct vertices because $\boldsymbol{K}$ does not contain any loops. Furthermore, the labelled vertices in $\boldsymbol{K}$ must be adjacent. Hence, $\boldsymbol{K}_2$ is a graph obtained from K_4 or $K_{2,3}$ by labelling two adjacent vertices and potentially removing the edge between them. In any case, $\boldsymbol{C} \leq \boldsymbol{K}_2$.
3. If $\boldsymbol{K}_1 = \boldsymbol{I}$, then $\boldsymbol{K}_2$ is obtained from K_4 or $K_{2,3}$ by either picking one vertex and placing both labels on it or by adding a fresh vertex, placing a label on it, and connecting it to a subset of the neighbours of a chosen original vertex, which receives the other label.
4. If $\boldsymbol{K}_1 = \boldsymbol{J}$, then $\boldsymbol{K}_2 = \boldsymbol{K}$. In particular, $\boldsymbol{K}_2$ is obtained from K_4 or $K_{2,3}$ by the procedure described in Item 3.
5. $\boldsymbol{K}_1$ cannot be any of the two remaining bilabelled minors of $\boldsymbol{P}$ since these contain an unlabelled vertex of degree at most one which is not the case for $\boldsymbol{K}$.

For Items 1 and 2, when $\boldsymbol{C} \leq \boldsymbol{F} = \boldsymbol{F}_1 \cdot \boldsymbol{F}_2$, by Lemma 5.2.18, $\boldsymbol{C} \leq \boldsymbol{F}_1$ or $\boldsymbol{C} \leq \boldsymbol{F}_2$ because the graph $\boldsymbol{C}$ cannot be written as the series composition of two graphs different from $\boldsymbol{I}$. The bilabelled minor $\boldsymbol{C}$ of $\boldsymbol{F}_1$ or $\boldsymbol{F}_2$ gives rise to a $K_{2,3}$-minor in their expansion, contradicting that $\boldsymbol{F}_1, \boldsymbol{F}_2 \in \mathcal{OP}(1,1)$.

For Items 3 and 4, let $\boldsymbol{K}_2$ be the graph described there. This graph can only be written as the series composition of two graphs different from $\boldsymbol{I}$ if the two labels do not coincide. In this case, one of the labelled vertices is adjacent to a subset of neighbours of the other labelled vertex. The graph $\boldsymbol{K}_2$ may be written as series composition of $\boldsymbol{A}$ or $\boldsymbol{J}$ with another graph $\boldsymbol{K}_2'$. The graph $\operatorname{soe}(\boldsymbol{K}_2')$ contains K_4 or $K_{2,3}$ as a minor. By Lemma 5.2.15, one of the factors $\boldsymbol{F}_1$ or $\boldsymbol{F}_2$ is not outerplanar, a contradiction. □

We proceed to prove the following auxiliary lemma:

Lemma 5.2.25. *Let F be an outerplanar simple graph with vertex $u \in V(F)$. If $\deg_F(u) \geq 1$, then there exists a neighbour $v \in N_F(u)$ such that the graph obtained from F by subdividing the edge uv is outerplanar.*

Proof. Take an outerplanar embedding of F which has some face incident to all the vertices, consider some edge incident to u that is incident to this face, and subdivide that edge. Since the vertex created by subdivision is incident to the outer face, the embedding remains outerplanar when the edge is subdivided. Alternatively, one may consider the following argument:

If u is of degree one or two, then any of its neighbours is as desired. If u has degree at least three, observe that $F[N_F(u)]$ cannot contain K_3 as a minor. Indeed, any such minor would give rise to a K_4-minor in F. Hence, $F[N_F(u)]$ is a forest and contains a vertex v of degree at most one in $F[N_F(u)]$. Write F' for the graph obtained from F by subdividing the edge uv. Write w for the vertex added this way.

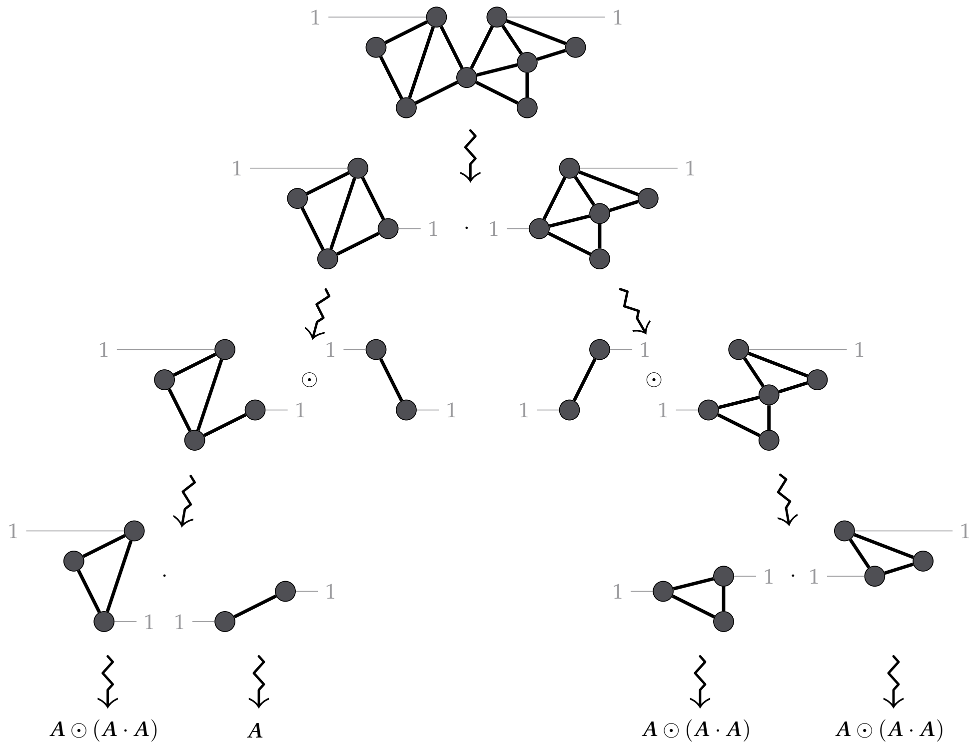


Figure 5.8: Decomposition of a bilabelled outerplanar graph into atomic graphs.

If F' is not outerplanar, then it contains a minor K_4 or $K_{2,3}$ which can be obtained from F' without undoing the subdivision. This minor cannot be K_4 because w is of degree two. Hence, F' contains a $K_{2,3}$-minor which can be obtained from F such that the path uwv is not contracted. This implies that v is adjacent to at least two neighbours of u which cannot be since it was chosen to be of degree one in $F[N_F(u)]$, a contraction. Hence, the graph F' is outerplanar. □

Lemma 5.2.25 facilitates decomposing bilabelled outerplanar graphs into simpler ones. See Figure 5.8 for an example.

Lemma 5.2.26. *Let $\boldsymbol{F} = (F, u, v) \in \mathcal{OP}(1, 1)$ have $n \geq 3$ vertices.*

1. *If $u = v$ and $uv \in E(F)$, then $\boldsymbol{F} = \boldsymbol{A} \odot \boldsymbol{K}$ where $\boldsymbol{K} = (K, x, x) \in \mathcal{OP}(1, 1)$ has at most n vertices and $x \notin E(F)$,*
2. *If $u = v$ and $uv \notin E(F)$, then $\boldsymbol{F} = \boldsymbol{I} \odot (\boldsymbol{K} \cdot \boldsymbol{J})$ or $\boldsymbol{F} = \boldsymbol{I} \odot ((\boldsymbol{A} \odot \boldsymbol{K}) \cdot \boldsymbol{J})$ where $\boldsymbol{K} = (K, x, y) \in \mathcal{OP}(1, 1)$ has at most n vertices and $x \neq y$.*
3. *If $u \neq v$ and $uv \in E(F)$, then $\boldsymbol{F} = \boldsymbol{A} \odot \boldsymbol{K}$ where $\boldsymbol{K} = (K, x, y) \in \mathcal{OP}(1, 1)$ has at most n vertices and satisfies $x \neq y$ and $xy \notin E(K)$,*
4. *If $u \neq v$ and $uv \notin E(F)$, then $\boldsymbol{F} = \boldsymbol{K} \cdot \boldsymbol{L}$ where $\boldsymbol{K}, \boldsymbol{L} \in \mathcal{OP}(1, 1)$ have at least 2 and at most $n - 1$ vertices.*

Proof. For Item 1, $\boldsymbol{K}$ is obtained from $\boldsymbol{F}$ by removing the loop at $u = v$.

For Item 2, distinguishing two cases.

- If $u = v$ is isolated in F, then let $x \in V(F) \setminus \{u\}$ be arbitrary. Define $K := F$. Then $\boldsymbol{K} := (K, u, x)$ is such that $\boldsymbol{F} = \boldsymbol{I} \odot (\boldsymbol{K} \cdot \boldsymbol{J})$. By definition, K is outerplanar. Since u is isolated, the expansion of $\boldsymbol{K}$ differs from $\boldsymbol{K}$ dangling path at u. Hence, $\boldsymbol{K}$ is outerplanar as well.
- If $u = v$ is not isolated, pick a neighbour x in virtue of Lemma 5.2.25, and let K be the graph obtained from F by deleting the edge ux, i.e. $V(K) := V(F)$ and $E(K) := E(F) \setminus \{ux\}$. Let $\boldsymbol{K} := (K, u, x)$. As a subgraph of F, K is outerplanar. The expansion of $\boldsymbol{K}$ is the graph obtained from F by subdividing the edge ux and outerplanar by Lemma 5.2.25. Hence, $\boldsymbol{K} \in \mathcal{OP}(1,1)$. Furthermore, $\boldsymbol{F} = \boldsymbol{I} \odot ((\boldsymbol{A} \odot \boldsymbol{K}) \cdot \boldsymbol{J})$.

For Item 3, define K by removing the edge uv from F, i.e. $V(K) := V(F)$ and $E(K) := E(F) \setminus \{uv\}$. The graph $\boldsymbol{K} := (K, u, v)$ satisfies $\boldsymbol{F} = \boldsymbol{A} \odot \boldsymbol{K}$ and all other stipulated properties. The expansion of $\boldsymbol{K}$ is a minor of the expansion of $\boldsymbol{F}$.

For Item 4, first suppose that u and v lie in the same connected component of F. Observe that there are no two internally vertex-disjoint paths from u to v since a pair of two such paths would give rise to a $K_{2,3}$-minor in the expansion of $\boldsymbol{F}$. By Menger's Theorem, there exists a vertex $x \neq u, v$ meeting all paths from u to v. Thus, removing x from F causes u and v to lie in separate connected components. Let A denote the connected component of $F - x$ containing u, B the connected component of $F - x$ containing v, and C the union of all connected components of $F - x$ containing neither u nor v. By definition, $V(F) = A \sqcup B \sqcup C \sqcup \{x\}$. Define $K := F[A \cup \{x\}]$ as the subgraph of F induced by $A \cup \{x\}$ and similarly $L := F[B \cup C \cup \{x\}]$. Let $\boldsymbol{K} := (K, u, x)$ and $\boldsymbol{L} := (L, x, v)$. Then $\boldsymbol{F} = \boldsymbol{K} \cdot \boldsymbol{L}$, as desired. As they are induced subgraphs of F, the graphs K and L are outerplanar. The expansions of $\boldsymbol{K}$ and $\boldsymbol{L}$ are minors of the expansion of $\boldsymbol{F}$ and thus outerplanar. Observe that $|V(K)| + |V(L)| = n + 1$ and $|V(K)|, |V(L)| \geq 2$, as desired.

Now suppose that u and v lie in separate connected components of F. Let A denote the connected component of F containing u, B the connected component of F containing v, and C the union of all connected components of F containing neither u nor v. Observe that $|A| + |B| + |C| = n \geq 3$. Distinguish cases:

- If $|A| + |C| \geq 2$, let $K := F[A \cup C]$ and $L' := F[B]$. Define $\boldsymbol{K} := (K, u, u)$, $\boldsymbol{L}' := (L', v, v)$, and $\boldsymbol{L} := \boldsymbol{J} \cdot \boldsymbol{L}'$.
- Otherwise, it holds that $|B| \geq 2$ and $C = \emptyset$. Let $K' := F[A \cup B]$ and $L := F[B]$. Define $\boldsymbol{K}' := (K', u, u)$, $\boldsymbol{L} := (L, v, v)$, and $\boldsymbol{K} := \boldsymbol{K}' \cdot \boldsymbol{J}$.

In both cases, $\boldsymbol{F} = \boldsymbol{K} \cdot \boldsymbol{L}$ and $\boldsymbol{K}, \boldsymbol{L} \in \mathcal{OP}(1,1)$. Furthermore, writing K and L for the graphs underlying $\boldsymbol{K}$ and $\boldsymbol{L}$ respectively, it holds that $|V(K)| + |V(L)| = n + 1$ and $|V(K)|, |V(L)| \geq 2$ since multiplication with $\boldsymbol{J}$ amounts to adding a fresh isolated vertex. □

The following Theorem 5.2.27 implies Theorem 5.2.21.

Theorem 5.2.27. *The classes $\mathcal{L}(1,1)$ and $\mathcal{OP}(1,1)$ coincide.*

Proof. For the inclusion $\mathcal{L}(1,1) \subseteq \mathcal{OP}(1,1)$, observe that the atomic graphs in $\mathcal{A}(1,1)$ are $\boldsymbol{A}, \boldsymbol{J}, \boldsymbol{I} \in \mathcal{OP}(1,1)$, cf. Figure 5.4. By Lemma 5.2.24, $\mathcal{OP}(1,1)$ is closed under series composition, parallel composition with atomic graphs, and permutation of labels. It follows inductively that $\mathcal{L}(1,1) \subseteq \mathcal{OP}(1,1)$.

For the inclusion $\mathcal{L}(1,1) \supseteq \mathcal{OP}(1,1)$, it is argued that $\boldsymbol{F} \in \mathcal{L}(1,1)$ if $\boldsymbol{F} \in \mathcal{OP}(1,1)$ by induction on the number of vertices in $\boldsymbol{F}$. If $\boldsymbol{F}$ has at most two vertices, this is clear. Suppose $\boldsymbol{F} = (F, u, v)$ has $n \geq 3$ vertices. By Items 1 to 3 of Lemma 5.2.26 and the closure properties of $\mathcal{L}(1,1)$ from Definition 5.2.1, it may be supposed that $u \neq v$ and $uv \notin E(F)$. In this case, again by Lemma 5.2.26, $\boldsymbol{F} = \boldsymbol{K} \cdot \boldsymbol{L}$ for graphs $\boldsymbol{K}$ and $\boldsymbol{L}$, to which the inductive hypothesis applies. It follows that $\boldsymbol{F} \in \mathcal{L}(1,1)$. □

6 The Homomorphism Distinguishing Closure

The quest for a polynomial-time algorithm for graph isomorphism [88] has led to an interest in various polynomial-time graph isomorphism relaxation. For example, the k-dimensional Weisfeiler–Leman algorithm wl_k runs in polynomial time and used to be a reasonable candidate for a polynomial-time graph isomorphism procedure. Only when the seminal Cai–Fürer–Immerman (CFI) graphs [37] were constructed, it emerged that wl_k does not decide isomorphism for all graphs. In general, separating a given graph isomorphism relaxation from isomorphism is a notoriously hard task and usually requires the subtle construction of separating examples.

This chapter presents a discussion of the perspective that homomorphism indistinguishability offers on the distinguishing power of graph isomorphism relaxations. Via the various characterisations of homomorphism indistinguishability relations in Chapters 3 to 5, the problem of separating homomorphism indistinguishability relations from isomorphism subsumes this task for graph isomorphisms relaxations from finite model theory [68], optimisation (Corollaries 6.4.2 and 6.3.17), or machine learning [180].

In order to separate a homomorphism indistinguishability relation $\equiv_{\mathcal{F}}$ of a graph class classes $\mathcal{F}$ from isomorphism, one must construct two simple graphs G and H such that $G \not\cong H$ and $G \equiv_{\mathcal{F}} H$. By Lovász's Theorem 3.1.1, isomorphism is a homomorphism indistinguishability relation. Thus, the task of separating $\equiv_{\mathcal{F}}$ from isomorphism is subsumed by the following more general problem: Given the homomorphism indistinguishability relations $\equiv_{\mathcal{F}_1}$ and $\equiv_{\mathcal{F}_2}$ of two graph classes $\mathcal{F}_1$ and $\mathcal{F}_2$, when is $\equiv_{\mathcal{F}_1}$ a refinement of $\equiv_{\mathcal{F}_2}$, i.e. when does $G \equiv_{\mathcal{F}_1} H$ imply that $G \equiv_{\mathcal{F}_2} H$ for all simple graphs G and H?

As argued above, this question has a counterpart in logic. For example, while it is clear that the set of formulas C^k in k-variable first-order logic with counting quantifiers forms a proper subset of C^{k+1}, it requires the CFI construction to argue that C^{k+1}-equivalence is a strictly finer graph isomorphism relaxation than C^k-equivalence. Here, $\mathsf{C}^k \subsetneq \mathsf{C}^{k+1}$ is a syntactic assertion whereas the relationship of C^{k+1}-equivalence and C^k-equivalence is semantic.

Analogous to syntax and semantics in logic, the reader is invited to think about the properties of the graph classes $\mathcal{F}_1$ and $\mathcal{F}_2$ as *syntactic* while the properties of the homomorphism indistinguishability relations $\equiv_{\mathcal{F}_1}$ and $\equiv_{\mathcal{F}_1}$ are referred to as *semantic*. Of course, if the syntactic statement $\mathcal{F}_1 \supseteq \mathcal{F}_2$ holds, then $\equiv_{\mathcal{F}_1}$ refines $\equiv_{\mathcal{F}_2}$, which is a semantic assertion. As it turns out, the semantic statement that $\equiv_{\mathcal{F}_1}$ refines $\equiv_{\mathcal{F}_2}$ does not always imply that $\mathcal{F}_1 \supseteq \mathcal{F}_2$ holds syntactically.

The central notion for studying this phenomenon is the homomorphism distinguishing closure. It was introduced by Roberson [150] in the first paper which systematically studied the semantics of homomorphism indistinguishability.

Definition 6.0.1 ([150, p. 3])**.** Let $\mathcal{F}$ be a graph class. The *homomorphism distinguishing closure* of $\mathcal{F}$ is

$$\mathrm{cl}(\mathcal{F}) := \{K \mid \forall G, H. \quad G \equiv_{\mathcal{F}} H \implies \hom(K, G) = \hom(K, H)\}.$$

Here, K, G, and H denote simple graphs. The class $\mathcal{F}$ is *homomorphism distinguishing closed* if $\mathrm{cl}(\mathcal{F}) = \mathcal{F}$.

Intuitively, $\mathrm{cl}(\mathcal{F})$ is the largest graph class whose homomorphism indistinguishability relation coincides with the one of $\mathcal{F}$. In other words, a graph class $\mathcal{F}$ is homomorphism distinguishing closed if, for every $F \notin \mathcal{F}$, there exist graphs G and H such that $G \equiv_{\mathcal{F}} H$ and $\hom(F, G) \neq \hom(F, H)$. Clearly, the class $\mathcal{G}$ of all graphs is homomorphism distinguishing closed. Further basic examples will be given in Theorems 6.1.1 and 6.1.2.

Establishing that a graph class is homomorphism distinguishing closed is a hard task and the list of known examples is limited. The subsequent sections present an almost exhaustive list of the known homomorphism distinguishing closed graph classes and the available proof techniques, which can be categorised as follows:

Oddomorphisms. In this thesis' language, Cai, Fürer, & Immerman [37] constructed their CFI graphs in order to show that the homomorphism distinguishing closure of the class of graphs of treewidth at most k does not contain all graphs. Strengthening this result, Roberson [150] observed that the CFI construction can be employed to show that many graph classes (not only the class of graphs of bounded treewidth) are homomorphism distinguishing closed. Roberson proposed a purely combinatorial criterion that, when satisfied by a graph class, implies that the graph class is homomorphism distinguishing closed: The graph class must be *closed under weak oddomorphisms*. Here, weak oddomorphisms are homomorphisms satisfying certain parity conditions.

The following graph classes were shown to be closed under weak oddomorphisms and thus to be homomorphism distinguishing closed: the class of graphs of bounded degree (Corollary 6.3.10), forests (Corollary 6.3.12), disjoint unions of paths (Corollary 6.3.13), and, for $h \geq 3$, the class of $K_{2,h}$-minor-free graphs of treewidth at most two (Corollary 6.3.16). Moreover, with slight adaptations of this approach, the class of disjoint unions of cycles is shown to be homomorphism distinguishing closed (Theorem 7.1.4).

Games. For graph classes whose homomorphism indistinguishability relations have connections to logic, the above approach can be combined with techniques from model theory and structural graph theory. The key technical challenge is to show that, for given a graph G which does not belong to such a graph class, its CFI graphs G_0 and G_1 are homomorphism indistinguishable over the class. Combinatorial games can be used to rephrase this statement: Graph searching games characterise membership in a graph class while model comparison games characterise logical equivalence and thus the homomorphism indistinguishability relation.

This approach was applied successfully for the classes of graphs of bounded treewidth (Theorem 6.4.1), bounded treedepth (Theorem 6.4.4), bounded pathwidth (Theorem 6.4.6), and the class of graphs with k-pebble forest covers of bounded depth (Theorem 6.4.5).

Finiteness. For graph classes which are essentially finite (Definition 6.5.1), the homomorphism distinguishing closure can be computed explicitly (Theorem 6.5.2).

Theorem 6.5.2 applies e.g. to the set of all minors of a fixed graph and serves as source for examples and counter-examples.

A central result which is not covered here asserts that the class of planar graphs is homomorphism distinguishing closed [150, Lemma 8.2, Theorem 8.3]. It relies on the characterisation of quantum isomorphism as homomorphism indistinguishability over planar graphs due to Mančinska & Roberson [124]. Furthermore, it makes use of the fact that the CFI graphs of a graph G are quantum isomorphic if, and only if, G is non-planar, cf. [15, 8].

It would be advantageous to be able to determine the homomorphism distinguishing closure of a given graph class with as few assumptions as possible. The techniques listed above are insufficient for this task. Roberson [150] suggested to consider minor-closed graph classes, which are rich in structure and general enough to be relevant in applications, e.g. in Chapter 5. They proposed the following Conjecture 6.0.2.

Conjecture 6.0.2 ([150, Conjecture 1])**.** *Every minor-closed and union-closed graph class is homomorphism distinguishing closed.*

While it is not hard to see that every homomorphism distinguishing closed graph class is closed under disjoint unions, cf. Lemma 6.2.2, the relevance of minor-closed graph classes in the context of homomorphism indistinguishability is less clear, cf. Chapter 7.

Conjecture 6.0.2 is wide open. The same holds true for its weaker version given below. The following sections present evidence for both conjectures. In Theorem 6.5.2, Conjecture 6.0.2 is confirmed for all essentially finite graph classes.

Conjecture 6.0.3 ([150, Conjecture 5]). *For a proper minor-closed graph class $\mathcal{F}$, the homomorphism indistinguishability relation $\equiv_{\mathcal{F}}$ is not isomorphism.*

In terms of the homomorphism distinguishing closure, Conjecture 6.0.3 can be rephrased as asserting that, for a proper minor-closed graph class $\mathcal{F}$, $\mathrm{cl}(\mathcal{F})$ does not contain all graphs.

Chapter Outline. In Section 6.1, intuition for the homomorphism distinguishing closure is built by considering first examples. Section 6.2 features some general properties of the homomorphism distinguishing closure. In Sections 6.3 to 6.5, the three main techniques available for proving that a graph class is homomorphism distinguishing closed are presented: oddomorphisms, games, and finiteness. The chapter is concluded in Section 6.6 with an overview of further research directions.

The material in Section 6.3.1 is joint work with Moritz Lichter and Benedikt Pago and was published in [111, 110]. The material in Section 6.3.4 is joint unpublished work with Daniel Neuen. Corollary 6.4.2 and Question 6.4.3 are joint work with David E. Roberson and were published in [151, 152]. The material in Section 6.5 was published in [162, 165].

6.1 First Examples and Non-Examples

This section features an example of a homomorphism distinguishing closed graph class and an example of a proper graph class whose homomorphism distinguishing closure is the class of all graphs. The first one is the class of bipartite graphs, for which this property is established in Theorem 6.1.1. This observation can be traced back to [150, 63]. The second one is the class of 2-degenerate graphs. We repeat an argument from [63] which is based on ideas from [119].

In plain terms, Theorem 6.1.1 asserts that, for every simple non-bipartite graph F, there exist simple graphs G and H such that G and H are homomorphism indistinguishable over all bipartite graphs but $\hom(F, G) \neq \hom(F, H)$.

Theorem 6.1.1 ([150, Lemma 5.8]). *The class $\mathcal{G}_{K_2}$ of all bipartite graphs is homomorphism distinguishing closed.*

Proof. Let F be a non-bipartite simple graph. Define $G := F \times K_2$ and $H := F + F$. We claim that $G \equiv_{\mathcal{G}_{K_2}} H$ and $\hom(F, G) \neq \hom(F, H)$. For the latter claim, observe that $\hom(F, G) = 0$ since G is bipartite. Clearly, $\hom(F, H) \neq 0$ since every graph admits a homomorphism into itself.

For the former claim, first note that, by Equation (2.1), it suffices to show that $\hom(L, G) = \hom(L, H)$ for all connected bipartite graphs L. For such a graph L, $\hom(L, G) = 2 \hom(L, F)$ by Equation (2.2) and since every connected bipartite graph admits exactly two homomorphisms to K_2. Secondly, by Equation (2.3), $\hom(L, H) = \hom(L, F) + \hom(L, F) = 2 \hom(L, F)$. Thus, $G \equiv_{\mathcal{G}_{K_2}} H$. $\square$

We now turn to Theorem 6.1.2 which is due to Dvořák [63]. Let $k \in \mathbb{N}$. A graph F is *k-degenerate* if every subgraph of F contains a vertex of degree at most k. In plain terms, Theorem 6.1.2 asserts that two simple graphs G and H are isomorphic whenever they are homomorphism indistinguishable over all 2-degenerate graphs.

Theorem 6.1.2 ([63, Theorem 12])**.** *The homomorphism distinguishing closure of the class of* 2*-degenerate graphs is the class of all graphs.*

The proof of Theorem 6.1.2 is based on the following Lemma 6.1.3, which is due to [119].

Lemma 6.1.3 ([119, Proof of Theorem 1.4], cf. [116, Proposition 6.21])**.** *For simple graphs G and H, there exist coefficients $\alpha_2, \dots, \alpha_\ell \in \mathbb{R}$ such that $\boldsymbol{A}_G = \sum_{i=2}^{\ell} \alpha_i \boldsymbol{A}_G^i$ and $\boldsymbol{A}_H = \sum_{i=2}^{\ell} \alpha_i \boldsymbol{A}_H^i$.*

Proof. Let $\lambda_1, \dots, \lambda_\ell$ be the non-zero eigenvalues of the adjacency matrix $\boldsymbol{A}_{G+H}$ of the disjoint union of G and H. Consider the polynomial $p(x) = x \prod_{i=1}^{\ell}(1 - x/\lambda_i) \in \mathbb{R}[x]$. By construction, $p(\boldsymbol{A}_{G+H})$ is the zero matrix. The polynomial p does not have a constant term and its linear monomial has non-zero coefficient. Thus, there exist coefficients $\alpha_2, \dots, \alpha_{\ell+1} \in \mathbb{R}$ such that $\boldsymbol{A}_{G+H} = \sum_{i=2}^{\ell+1} \alpha_i \boldsymbol{A}_{G+H}^i$ The adjacency matrix $\boldsymbol{A}_{G+H}$ is of block shape $\boldsymbol{A}_{G+H} = \begin{pmatrix} A_G & 0 \\ 0 & A_H \end{pmatrix}$. Hence, the desired identities hold. □

Lemma 6.1.3 implies the following Lemma 6.1.4.

Lemma 6.1.4 ([63, Lemma 11])**.** *Let F, G, and H be simple graphs. Let $e \in E(F)$. If* $\hom(F, G) \neq \hom(F, H)$*, then there exists a graph F' obtained from F by subdividing the edge e at least once such that* $\hom(F', G) \neq \hom(F', H)$*.*

Proof. Write $u, v \in V(F)$ for the vertices incident to the edge e. Furthermore, let F' be the graph obtained from F by deleting e. Let $\boldsymbol{F}' := (F', u, v) \in \mathcal{G}(1,1)$. Recall the definition of $\boldsymbol{A} \in \mathcal{G}(1,1)$ from Example 3.2.3. Then $F \cong \operatorname{soe}(\boldsymbol{A} \odot \boldsymbol{F}')$. Let $\alpha_2, \dots, \alpha_\ell \in \mathbb{R}$ denote the coefficients from Lemma 6.1.3. Then $\hom(F, G) = \operatorname{soe}(\boldsymbol{A}_G \odot \boldsymbol{F}'_G) = \sum_{i=2}^{\ell} \alpha_i \operatorname{soe}(\boldsymbol{A}_G^i \odot \boldsymbol{F}'_G)$ and the same equality holds for H. Hence, if $\hom(F, G) \neq \hom(F, H)$, then there exists an integer $2 \leq i \leq \ell$ such that $\operatorname{soe}(\boldsymbol{A}_G^i \odot \boldsymbol{F}'_G) \neq \operatorname{soe}(\boldsymbol{A}_H^i \odot \boldsymbol{F}'_H)$. The graph $\operatorname{soe}(\boldsymbol{A}^i \odot \boldsymbol{F}')$ is obtained from F by subdividing the edge e for $i - 1$ times. □

This concludes the preparations for the proof of Theorem 6.1.2.

Proof of Theorem 6.1.2. Let F be an arbitrary simple graph. We show that if two simple graphs G and H are homomorphism indistinguishable over all 2-degenerate graphs, then $\hom(F, G) = \hom(F, H)$. Contrapositively, suppose that $\hom(F, G) \neq \hom(F, H)$. By repeatedly subdividing edges and invoking Lemma 6.1.4, one may construct a 2-degenerate graph F' from F such that $\hom(F', G) \neq \hom(F', H)$. □

6.2 Properties of the Homomorphism Distinguishing Closure

In this section, the homomorphism distinguishing closure is demonstrated to be a robust notion by proving a selection of basic properties. These will be useful in subsequent arguments.

First note that cl is a *closure operator* in the sense that $\mathrm{cl}(\mathcal{F}) \subseteq \mathrm{cl}(\mathcal{F}')$ if $\mathcal{F} \subseteq \mathcal{F}'$ and $\mathrm{cl}(\mathrm{cl}(\mathcal{F})) = \mathrm{cl}(\mathcal{F})$ for all graph classes $\mathcal{F}$ and $\mathcal{F}'$. Moreover, it holds that that $\mathrm{cl}(\emptyset) = \{K_0\}$ where K_0 denotes the empty graph, i.e. $V(K_0) = \emptyset = E(K_0)$. Furthermore, the intersection of homomorphism distinguishing closed graph classes is homomorphism distinguishing closed:

Lemma 6.2.1 ([150, Lemma 6.1])**.** *Let I be an arbitrary set and let $\mathcal{F}_i$ for every $i \in I$ be a homomorphism distinguishing closed graph class. Then $\mathcal{F} := \bigcap_{i\in I} \mathcal{F}_i$ is homomorphism distinguishing closed.*

Proof. For the sake of completeness, the following proof from [150, Lemma 6.1] is included. Let $F \notin \mathcal{F}$ be a simple graph. Then $F \notin \mathcal{F}_i$ for some $i \in I$. Since $\mathcal{F}_i$ is homomorphism distinguishing closed, there exist simple graphs G and H such that $G \equiv_{\mathcal{F}_i} H$ and $\hom(F, G) \neq \hom(F, H)$. Since $\mathcal{F} \subseteq \mathcal{F}_i$, it holds that $G \equiv_{\mathcal{F}} H$ and $\hom(F, G) \neq \hom(F, H)$, as desired. □

The analogue of Lemma 6.2.1 for unions does not holds. This is demonstrated by Example 6.5.11 and hinted at by the following Lemma 6.2.2. It implies for example that no finite graph class is homomorphism distinguishing closed.

Lemma 6.2.2. *Every homomorphism distinguishing closed graph class $\mathcal{F}$ is closed under disjoint unions, i.e. if $F_1, F_2 \in \mathcal{F}$, then $F_1 + F_2 \in \mathcal{F}$.*

Proof. We show that if $F_1, F_2 \in \mathcal{F}$, then $F_1 + F_2 \in \mathrm{cl}(\mathcal{F})$. To that end, let G and H be simple graphs such that $G \equiv_{\mathcal{F}} H$. By Equation (2.1), $\hom(F_1 + F_2, G) = \hom(F_1, G)\hom(F_2, G) = \hom(F_1, H)\hom(F_2, H) = \hom(F_1 + F_2, H)$. Hence, $F_1 + F_2 \in \mathrm{cl}(\mathcal{F})$. □

From Lemma 6.2.1, the following useful Lemma 6.2.3 can be derived. Neither of the inclusions in Lemma 6.2.3 can be reversed in general, cf. Example 6.5.11.

Lemma 6.2.3. *Let I be an arbitrary index set and $\mathcal{F}_i$ for every $i \in I$ be a graph class. Then*

$$\mathrm{cl}\left(\bigcap_{i\in I} \mathcal{F}_i\right) \subseteq \bigcap_{i\in I} \mathrm{cl}(\mathcal{F}_i) \quad \textit{and} \quad \bigcup_{i\in I} \mathrm{cl}(\mathcal{F}_i) \subseteq \mathrm{cl}\left(\bigcup_{i\in I} \mathcal{F}_i\right).$$

Proof. By Lemma 6.2.1, $\bigcap_{i\in I} \mathrm{cl}(\mathcal{F}_i)$ is homomorphism distinguishing closed and it suffices to observe that $\bigcap_{i\in I} \mathcal{F}_i \subseteq \bigcap_{i\in I} \mathrm{cl}(\mathcal{F}_i)$. For the second claim, $\mathcal{F}_j \subseteq \bigcup_{i\in I} \mathcal{F}_i$ and hence $\mathrm{cl}(\mathcal{F}_j) \subseteq \mathrm{cl}(\bigcup_{i\in I} \mathcal{F}_i)$ for every $j \in I$. □

6.3 CFI Graphs and Oddomorphisms

The most versatile tool for proving that a graph class is homomorphism distinguishing closed is the CFI construction. Originally [37], CFI graphs were introduced to separate C^k-equivalence from isomorphism. This was done by constructing, given a graph G, two graphs G_0 and G_1, the even and odd CFI graphs of G. Intuitively, if G is a complicated graph, e.g. if $\operatorname{tw}(G) > k$, then G_0 and G_1 are indistinguishable, e.g. C^{k+1}-equivalent (Theorems 6.4.1 and 3.4.4). Roberson [150] demonstrated that the similarity of G_0 and G_1 can be understood via homomorphism indistinguishability. In this section, these observations are extended to more general CFI graphs.

6.3.1 Homomorphisms into CFI Graphs over Finite Abelian Groups

Roberson [150] studied homomorphisms to CFI graphs constructed over the group $\mathbb{Z}_2$ of order two. Their variant of CFI graphs was introduced in [71]. In [135], the classical CFI construction over $\mathbb{Z}_2$ was generalised to arbitrary finite abelian groups. Combing the construction from [135] with the observations from [150], we consider homomorphisms into CFI graphs over arbitrary finite abelian groups. For an application of such CFI graphs, see [108, 111]. Apart from this section, this thesis is concerned with CFI graphs over $\mathbb{Z}_2$. We state the results in this section more generally for future reference.

Throughout this section we fix a finite abelian group Γ and write it additively. For a simple graph G and a vertex $v \in V(G)$, write $E(v) := \{e \in E(G) \mid v \in e\}$ for the set of edges incident to v. We consider vectors $U \in \Gamma^X$ for finite sets X. For an element $x \in X$, we write $U(x) \in \Gamma$ for the x-th entry of U. When convenient, we write $\gamma x \in \Gamma^X$ for the vector which is γ at the x-th entry and zero everywhere else for $x \in X$ and $\gamma \in \Gamma$. Write $\sum U$ for $\sum_{x \in X} U(x)$ and $0 \in \Gamma^X$ for the function which is zero everywhere.

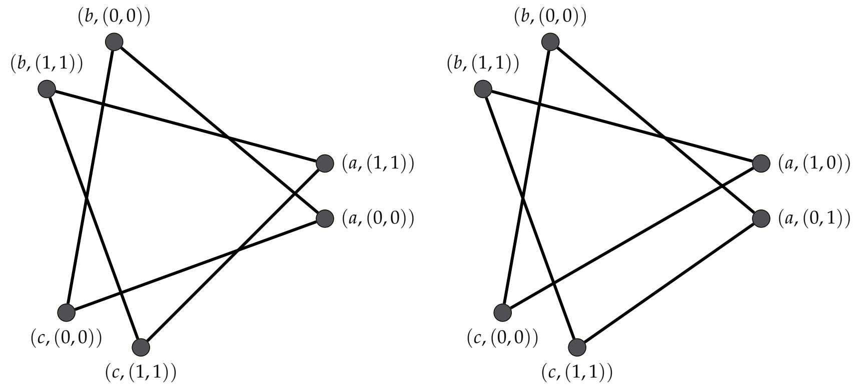
Definition 6.3.1. A *base graph* is a connected simple graph. Let G be a base graph and $U \in \Gamma^{V(G)}$. The graph G_U has vertices (u, S) for every $u \in V(G)$ and $S \in \Gamma^{E(u)}$ such that $\sum S = U(u)$. Two vertices (u, S) and (v, T) are adjacent if $uv \in E(G)$ and $S(uv) + T(uv) = 0$.

Consider the example in Figure 6.1. For further examples, see Figures 3.1 and 3.8, which depict the CFI graphs over $\mathbb{Z}_2$ of the star $K_{1,3}$ and the path P_4, respectively.

The proof of the following Lemma 6.3.2 uses well-known arguments for CFI graphs [71, 135, 109].

Lemma 6.3.2. *Let G be a base graph and $U, U' \in \Gamma^{V(G)}$. If $\sum U = \sum U'$, then $G_U \cong G_{U'}$.*

Proof. Let $uv \in E(G)$. Let $\gamma \in \Gamma$. First consider $U' := U + \gamma u - \gamma v$ where γu and γv denote the vectors in $\Gamma^{V(G)}$ with γ at the u-th and v-th component, respectively,



(a) $G_0 \cong 2K_3$. **(b)** $G_a \cong C_6$.

Figure 6.1: The CFI graphs of K_3 with $V(K_3) = \{a, b, c\}$ over the abelian group $\mathbb{Z}_2$. The entries of the vectors correspond to the neighbours of each vertex in alphabetical order.

and zero otherwise. Define the map $\varphi \colon V(G_U) \to V(G_{U'})$ by

$$\varphi((w, S)) := \begin{cases} (u, S + \gamma uv), & \text{if } w = u, \\ (v, S - \gamma uv), & \text{if } w = v, \\ (w, S), & \text{otherwise,} \end{cases}$$

where γuv denotes the vector in $\Gamma^{E(u)}$ in the first case or in $\Gamma^{E(v)}$ in the second case with γ at the uv-th component and zero otherwise. Observe that

$$\sum_{e \in E(v)} (S - \gamma uv)(e) = U(v) - \gamma = U'(v)$$

and analogously for u. Hence, φ is indeed a well-defined map to $V_{U'}$. Clearly, φ is a bijection. Let $(x, S), (y, T) \in V(G_U)$ be arbitrary vertices of G_U and define $(x, S') := \varphi(x, S)$ and $(y, T') := \varphi(y, T)$. Then $S'(xy) + T'(xy) = S(xy) + T(xy)$. Hence, (x, S) and (y, T) are adjacent in G_U if, and only if, they are adjacent in $G_{U'}$.

Since G is connected, the maps constructed above can be composed to yield $G_U \cong G_{U + \gamma u - \gamma v}$ for every pair of vertices $u, v \in V(G)$ and every $\gamma \in \Gamma$. This yields $G_U \cong G_{U'}$ as desired. $\square$

We proceed by counting homomorphisms into CFI graphs. For a base graph G and $U \in \Gamma^{V(G)}$, consider the *projection map* $\rho \colon G_U \to G$ sending (v, S) to v. Clearly, ρ is a homomorphism. For a graph F and a homomorphism $\psi \colon F \to G$, define

$$\mathrm{Hom}_\psi(F, G_U) := \{\varphi \in \mathrm{Hom}(F, G_U) \mid \rho \circ \varphi = \psi\}.$$

The sets $\mathrm{Hom}_\psi(F, G_U)$ for all $\psi \colon F \to G$ partition the set $\mathrm{Hom}(F, G_U)$ of homomorphisms $F \to G_U$. Write $\hom_\psi(F, G_U)$ for the cardinality of $\mathrm{Hom}_\psi(F, G_U)$. Lemma 6.3.3 generalises [150, Lemma 3.4].

Lemma 6.3.3. *Let F be a simple graph and G be a base graph. Let $U \in \Gamma^{V(G)}$ and fix $\psi \in \mathrm{Hom}(F, G)$. Consider the system of equations $\mathrm{HOM}(F, G, U, \psi)$ with variables x_e^a for all $a \in V(F)$ and $e \in E(\psi(a))$ and equations*

$$\sum_{e \in E(\psi(a))} x_e^a = U(\psi(a)) \qquad \text{for all } a \in V(F), \tag{6.1}$$

$$x_e^a + x_e^b = 0 \qquad \text{for all } ab \in E(F) \text{ and } e = \psi(ab) \in E(G). \tag{6.2}$$

Then the number of solutions to $\mathrm{HOM}(F, G, U, \psi)$ over Γ is equal to $\hom_\psi(F, G_U)$.

Proof. We construct a bijection between the solution set of $\mathrm{HOM}(F, G, U, \psi)$ and $\mathrm{Hom}_\psi(F, G_U)$. Let $\xi = (\xi_e^a)_{a \in V(F), e \in E(\psi(a))}$ be a solution to $\mathrm{HOM}(F, G, U, \psi)$ over Γ. Define a homomorphism $\varphi_\xi \in \mathrm{Hom}_\psi(F, G_U)$ via $\varphi_\xi(a) := \big(\psi(a), (\xi_e^a)_{e \in E(\psi(a))}\big)$. Equation (6.1) guarantees that φ_ξ is indeed a map from the vertices of F to the ones of G_U. If a and b are adjacent in F, then so are $\psi(a)$ and $\psi(b)$ in G. Furthermore, $\xi^a_{\psi(ab)} + \xi^b_{\psi(ab)} = 0$ by Equation (6.2). Hence, $\varphi_\xi(a)$ and $\varphi_\xi(b)$ are adjacent in G_U.

It is easy to see that this construction is injective, i.e. if $\varphi_\xi = \varphi_\zeta$ for solutions ξ and ζ of $\mathrm{HOM}(F, G, U, \psi)$ over Γ, then $\xi = \zeta$. For surjectivity, let $\varphi \in \mathrm{Hom}_\psi(F, G_U)$. For $a \in V(F)$ and $e \in E(\psi(a))$, define ξ_e^a as the second component of $\varphi(a)$, i.e. $\xi_e^a := S_a(e)$ where $\varphi(a) = (\psi(a), S_a)$. Clearly, $\xi := (\xi_e^a)$ is such that $\varphi_\xi = \varphi$. The fact that ξ satisfies Equations (6.1) and (6.2) is easily verified. □

Lemma 6.3.3 is used in the following Theorem 6.3.4 to compare homomorphisms counts into G_0 and G_U. Theorem 6.3.4 generalises [150, Theorem 3.6].

Theorem 6.3.4. *For a base graph G, a vector $U \in \Gamma^{V(G)}$, and $\psi \in \mathrm{Hom}(F, G)$ for some simple graph F, the following hold:*

1. $\hom_\psi(F, G_0) > 0$.
2. *If $\mathrm{HOM}(F, G, U, \psi)$ has a solution, then $\hom_\psi(F, G_0) = \hom_\psi(F, G_U)$.*
3. *If $\mathrm{HOM}(F, G, U, \psi)$ has no solution, then $\hom_\psi(F, G_U) = 0$.*

Proof. For $U = 0$, the system in Lemma 6.3.3 is homogeneous. Hence, it always has a solution, namely the assignment of zero to all variables. This yields the first claim.

For the second claim, let $\xi = (\xi_e^a)_{a \in V(F), e \in E(\psi(a))}$ denote a solution to the system $\mathrm{HOM}(F, G, U, \psi)$. Then the map $x \mapsto x + \xi$ is a bijection between the sets of solutions of $\mathrm{HOM}(F, G, 0, \psi)$ and $\mathrm{HOM}(F, G, U, \psi)$.

The third claim is implied by Lemma 6.3.3. □

Theorem 6.3.4 implies Corollary 6.3.5, which gives a criterion for a CFI graph G_U to satisfy that $\sum U = 0$ in terms of homomorphism counts from G. The condition in Item 3 is relevant for subsequent arguments. Corollary 6.3.5 generalises [150, Corollary 3.7].

Corollary 6.3.5. *Let G be a base graph and $U \in \Gamma^{V(G)}$. Then the following are equivalent:*

1. $\sum U = 0$,

2. $G_U \cong G_0$,
3. $\hom(G, G_U) = \hom(G, G_0)$, *and*
4. $\hom_{\mathrm{id}}(G, G_U) = \hom_{\mathrm{id}}(G, G_0)$, *where* id *is the identity map on* G.

Proof. By Lemma 6.3.2, Item 1 implies Item 2. It is immediate that Item 2 implies Item 3. The fact that Item 3 implies Item 4 follows from Theorem 6.3.4.

It thus remains to prove that Item 4 implies Item 1. By Theorem 6.3.4, let ξ be a solution to $\mathrm{HOM}(G, G, U, \mathrm{id})$. Then,

$$\sum_{a \in V(G)} U(a) \overset{(6.1)}{=} \sum_{a \in V(G)} \sum_{e \in E(a)} \xi_e^a = \sum_{e = ab \in E(G)} \xi_e^a + \xi_e^b \overset{(6.2)}{=} 0.$$

Hence, Item 1 holds. □

In the subsequent sections, we will be interested in CFI graphs over the abelian group $\mathbb{Z}_2$ on two elements. By Lemma 6.3.2 and Corollary 6.3.5, there are, up to isomorphism, precisely two CFI graphs over a base graph G, namely G_0 and $G_1 \cong G_U$ for some $U \in \mathbb{Z}_2^{V(G)}$ with $\sum U = 1$.

Definition 6.3.6. Let G be a base graph. The graphs G_0 and G_1 are the *even* and *odd CFI graphs* of G as defined in Definition 6.3.1 over the group $\mathbb{Z}_2$.

6.3.2 Oddomorphisms

Roberson [150] gave a combinatorial criterion for the CFI graphs G_0 and G_1 of a base graph G to have a different number of homomorphisms from a graph F. This criterion is the existence of a so-called weak oddomorphism from F to G.

We restrict our attention to CFI graphs over $\mathbb{Z}_2$. Note that while Corollary 6.3.5 holds for any finite abelian group, the following Definition 6.3.7 and Theorem 6.3.8 require the CFI graphs to be constructed over the additive group of a finite field. This is because their proofs are based on the Fredholm alternative [160, Corollary 3.1b], which gives a criterion for the infeasibility of a linear system of equations over a field.

Definition 6.3.7 ([150, Definition 3.9])**.** Let F and G be simple graphs and $\psi \colon F \to G$ a homomorphism. A vertex $a \in V(F)$ is *odd/even with respect to* ψ if $|N_F(a) \cap \psi^{-1}(v)|$ is odd/even for every $v \in N_G(\psi(a))$. The homomorphism ψ is an *oddomorphism* if

1. every vertex of F is even or odd with respect to ψ,
2. for every $v \in V(G)$, $\psi^{-1}(v)$ contains an odd number of odd vertices.

The homomorphism $\psi \colon F \to G$ is a *weak oddomorphism* if there is a subgraph F' of F such that $\psi|_{F'}$ is an oddomorphism from F' to G.

If a vertex $a \in V(F)$ is odd or even with respect to ψ, it is referred to as *ψ-odd* or *ψ-even*, respectively. The sets $\psi^{-1}(v) \subseteq V(F)$ for $v \in V(G)$ are called the *fibres* of ψ.

Note that every weak oddomorphism $\psi\colon F \to G$ is surjective on vertices and edges. That is, for every $v \in V(G)$, there exists $a \in V(F)$ such that $\psi(a) = v$ and, for every $uv \in E(G)$, there exists $ab \in E(F)$ such that $\psi(ab) = uv$. In particular, $|V(F)| \geq |V(G)|$ and $|E(F)| \geq |E(G)|$. An example for an oddomorphism is the identity map $\mathrm{id}\colon G \to G$.

The notion of oddomorphisms is featured in the following characterisation of homomorphism indistinguishability of CFI graphs.

Theorem 6.3.8 ([150, Theorem 3.13]). *Let G be a base graph. Let G_0 and G_1 denote the even and odd CFI graphs of G over $\mathbb{Z}_2$. Let F be a simple graph. Then*

$$\hom(F, G_0) \geq \hom(F, G_1)$$

with strict inequality if, and only if, there exists a weak oddomorphism from F to G.

In virtue of Theorem 6.3.8, oddomorphisms can be used to show that a graph class is homomorphism distinguishing closed via Theorem 6.3.9.

Theorem 6.3.9 ([150, Theorem 6.2]). *Let $\mathcal{F}$ be a graph class such that*

1. *$\mathcal{F}$ is* closed under weak oddomorphisms, *i.e. if $F \in \mathcal{F}$ and there exists a weak oddomorphism $F \to G$, then $G \in \mathcal{F}$,*
2. *$\mathcal{F}$ is closed under disjoint unions and taking summands, i.e. $F_1 + F_2 \in \mathcal{F}$ if, and only if, $F_1, F_2 \in \mathcal{F}$ for all graphs F_1, F_2.*

Then $\mathcal{F}$ is homomorphism distinguishing closed.

The conditions of Theorem 6.3.9 can be simplified for graph classes which are closed under taking subgraphs: If a graph class $\mathcal{F}$ is closed under oddomorphisms, taking subgraphs, and disjoint unions, then $\mathcal{F}$ is homomorphism distinguishing closed.

How restrictive are the assumptions of Theorem 6.3.9? By Lemma 6.2.2, the assumption that $\mathcal{F}$ is closed under disjoint unions does not constitute a loss of generality. In contrast, not every homomorphism distinguishing closed graph class is closed under weak oddomorphism. Here, an example is the class of disjoint unions of cycles, cf. Theorem 7.1.4. The last assumption that $\mathcal{F}$ is closed under taking summands also restricts the scope of Theorem 6.3.9 as demonstrated by Example 6.5.11.

Nevertheless, Theorem 6.3.9 applies to most graph classes which are known to be homomorphism distinguishing closed. The graph classes to which it does not apply are rather particular, cf. Theorems 6.5.2 and 7.1.4. Note that in Corollary 7.1.5, using results from Chapter 7, the first assumption in Theorem 6.3.9 will be relaxed.

6.3.3 Graphs of Bounded Degree, Paths, and Trees

In this section, we summarise some elementary examples for the use of Theorem 6.3.9. We start with the titular result of [150] on the classes of graphs of bounded degree.

Corollary 6.3.10 ([150, Lemma 4.1 and Theorem 4.7]). *Let $d \geq 0$. The class of graphs of degree at most d is closed under weak oddomorphisms. In particular, it is homomorphism distinguishing closed.*

Proof. Clearly, the class in question is closed under disjoint unions and restriction to connected components. It remains to argue that it is closed under weak oddomorphisms. Let F be a graph of maximal degree d and $F \to G$ a weak oddomorphism. By Definition 6.3.7, there exists a subgraph F' of F such that there exists an oddomorphism $\psi \colon F' \to G$. Suppose that $v \in V(G)$ has degree greater than d. The fibre $\psi^{-1}(v)$ contains an odd vertex $a \in V(F')$. Since a is odd, it has at least one neighbour in every $\psi^{-1}(w)$ for $w \in N_G(v)$. Hence, $\deg_{F'}(a) > d$, a contradiction. The second claim follows from Theorem 6.3.9. □

Corollary 6.3.10 implies that the class of graphs that are disjoint unions of cycles and paths, i.e. the class of graphs of degree at most two, is homomorphism distinguishing closed. The homomorphism indistinguishability relation of this class was characterised in Corollary 3.3.3 as cospectrality of the adjacency matrices of both the graphs and their complements. In contrast, the class of disjoint unions of cycles, whose homomorphism indistinguishability relation is cospectrality of the adjacency matrices, cf. Theorem 3.3.2, is not closed under weak oddomorphisms. For example, the homomorphism $C_4 \to P_3$ sending $1 \mapsto 1$, $2 \mapsto 2$, $3 \mapsto 3$, $4 \mapsto 2$ is a weak oddomorphism inducing an oddomorphism to P_3 from the induced subgraph of C_4 with vertex set $\{1, 2, 3\}$. Nevertheless, it will be shown in Theorem 7.1.4 that the class of disjoint unions of cycles is homomorphism distinguishing closed.

Corollary 6.3.10 answers a question in [61] negatively: Do homomorphism counts from graphs of bounded degree characterise a graph up to isomorphism? This question was also addressed in [86, Theorem 3], which asserts that homomorphism counts from trees of bounded degree do not characterise a graph up to indistinguishability under Colour Refinement, cf. Theorem 3.4.3. Being superseded by Corollary 6.3.10, this result is omitted here.

In light of Conjecture 6.0.2, minor-closed graph classes are of special interest. By the Robertson–Seymour Theorem [155], every minor-closed graph class can be described by a finite list of forbidden minors. These forbidden minors serve as certificates for a graph not being in the class. Thus, in light of Theorem 6.3.9, when showing that a graph class is closed under weak oddomorphisms, it has to be argued that if $F \to G$ is a weak oddomorphism and G' a minor of G, then F already has G' has a minor. The following Lemma 6.3.11 simplifies such arguments.

Lemma 6.3.11 ([150, Lemma 5.6]). *Let F and G be simple graphs admitting a weak oddomorphism $F \to G$. For every minor G' of G, there exists a minor F' of F such that F' admits an oddomorphism to G'.*

As an example for the use of Lemma 6.3.11, consider the following Corollary 6.3.12.

Homomorphism indistinguishability over the class of forests is characterised by Theorems 3.4.3 and 3.4.4 and Corollary 4.2.4.

Corollary 6.3.12 ([150, Corollaries 5.16 and 6.8])**.** *The class of forests is closed under weak oddomorphisms. In particular, it is homomorphism distinguishing closed.*

Proof. In order to apply Theorem 6.3.9, we argue that the class in question is closed under weak oddomorphisms. Towards a contradiction, let F be a forest and G be graph which is not a forest. By assumption, G contains K_3 as a minor. If there is a weak oddomorphism $F \to G$, then, by Lemma 6.3.11, there exists a minor F' of F and an oddomorphism $\varphi\colon F' \to K_3$. All vertices in K_3 have even degree. Hence, all vertices in F' have even degree.

Indeed, if $a \in V(F')$ is φ-odd, then it has an odd number of neighbours in every fibre $\varphi^{-1}(v)$ for every $v \in N_{K_3}(\varphi(a))$. There are exactly two such fibres. Hence, a has even degree. If a is φ-even, then it has an even number of neighbours in every fibre. Thus, a has even degree. Hence, F' is not a forest, since every forest contains a vertex of degree one. In particular, F has a minor which is not a forest, a contradiction. □

Corollaries 6.3.10 and 6.3.12 imply that the class of disjoint unions of paths is homomorphism distinguishing closed. Their homomorphism indistinguishability relation is characterised by Corollary 4.2.3.

Corollary 6.3.13. *The class of disjoint unions of paths is closed under weak oddomorphisms. In particular, it is homomorphism distinguishing closed.*

Proof. Let F be a disjoint union of paths and $F \to G$ a weak oddomorphism. By Corollary 6.3.12, G is a forest. By Corollary 6.3.10, all vertices of G have degree at most two. Thus, G is a disjoint union of paths. For the second claim, apply Theorem 6.3.9 or Lemma 6.2.3. □

We give a last corollary of Lemma 6.3.11. It will be applied in Section 6.3.4 and strengthened by Theorem 6.4.1. Homomorphism indistinguishability over the class of graphs of treewidth at most two is characterised by Theorem 3.4.3 as indistinguishability under the 2-dimensional Weisfeiler–Leman algorithm. In fact, this result is used in the proof of Corollary 6.3.14.

Corollary 6.3.14 ([150, Corollary 6.9])**.** *The class of graphs of treewidth at most two is closed under weak oddomorphisms. In particular, it is homomorphism distinguishing closed.*

Proof. By [26, Theorem 17], a graph is of treewidth at most two if, and only if, it does not contain K_4 as a minor. By Lemma 6.3.11, it suffices to show that if a graph F admits an oddomorphism to K_4, then F has treewidth greater than two. Let G_0 and G_1 denote the even and odd CFI graphs of K_4. Both of these graphs are strongly regular with parameters $(16, 6, 2, 2)$. They are known as the 4×4 *rook graph* and the

Shrikhande graph, respectively, cf. [81, Figure 2]. As strongly regular graphs with equal parameters, they are not distinguished by the 2-dimensional Weisfeiler–Leman algorithm. Thus, by Theorem 3.4.3, they are homomorphism indistinguishable over all graphs of treewidth at most two. By Theorem 6.3.8, F has treewidth greater than two, as desired. For the second claim, apply Theorem 6.3.9. □

6.3.4 $K_{2,h}$-Minor-Free Graphs of Treewidth at Most Two

In this section, we show, for every $h \geq 3$, that the class of graphs of treewidth at most two that do not have the complete bipartite graph $K_{2,h}$ as a minor is homomorphism distinguishing closed. The best known example of such a graph class is the class of outerplanar graphs, which, by definition or [171], is characterised by the forbidden minors K_4 and $K_{2,3}$. The proof is based on Theorem 6.3.9 and showcases some intricate combinatorial arguments which are needed to establish that a graph class is closed under oddomorphisms.

Theorem 6.3.15. *Let $h \geq 3$. If F is a simple graph of treewidth at most two admitting an oddomorphism $\varphi \colon F \to K_{2,h}$, then $K_{2,h}$ is a minor of F.*

Theorem 6.3.15 and Corollary 6.3.14 yield the following corollary.

Corollary 6.3.16. *Let $h \geq 3$. The class of graphs of treewidth at most two without $K_{2,h}$ as a minor is closed under weak oddomorphisms. In particular, the class of outerplanar graphs is homomorphism distinguishing closed.*

Proof of Corollary 6.3.16 assuming Theorem 6.3.15. In order to apply Theorem 6.3.9, we need to argue that the graph class in question is closed under weak oddomorphisms. Let F be a $K_{2,h}$-minor-free graph of treewidth at most two. Let G be an arbitrary graph admitting a weak oddomorphism $F \to G$. By Corollary 6.3.14, the graph G has treewidth at most two. Thus, we may suppose that G contains $K_{2,h}$ as a minor. By Lemma 6.3.11, there exists a minor F' of F such that F' admits an oddomorphism to $K_{2,h}$. This contradicts Theorem 6.3.15. Hence, G is $K_{2,h}$-minor-free.

A graph is outerplanar if, and only if, it does not contain K_4 and $K_{2,3}$ as a minor. A graph is of treewidth at most two if, and only if, it does not contain K_4 as a minor [26, Theorem 17]. Thus a graph is outerplanar if, and only if, it is of treewidth at most two and does not contain $K_{2,3}$ as a minor. Hence, the second claim follows from the first. □

Corollary 6.3.16 in conjunction with Theorems 5.1.1, 5.1.2, 5.2.20, and 5.2.21 yields the following statement on the feasibility of the level-1 Lasserre semidefinite program. Here, by Theorem 6.3.8, G and H can be taken to be the even and odd CFI graphs of K_4, cf. Corollary 6.3.14.

Corollary 6.3.17. *There exist simple graphs G and H for which $\mathrm{L}^1(G, H)$ has a real solution but no non-negative real solution.*

Vertices of Low Degree Which Do Not Cut

Towards the proof of Theorem 6.3.15, several lemmas are devised which provide vertices of low degree in graphs of bounded treewidth that are not cut vertices. For a tree decomposition (T, β) of some graph F, define $\|(T, \beta)\| := \sum_{t \in V(T)} |\beta(t)|^2$. We will consider tree decompositions which are minimal with respect to this quantity.

Lemma 6.3.18. *Let $F = F_1 + F_2$ be a simple graph with tree decomposition (T, β). Define $\beta_i(t) := \beta(t) \cap V(F_i)$ for $i \in [2]$. Then (T, β_i) is a tree decomposition of F_i for $i \in [2]$ and $\|(T, \beta_1)\| + \|(T, \beta_2)\| \leq \|(T, \beta)\|$. Furthermore, if there exists $r \in V(T)$ such that both $\beta(r) \cap V(F_1)$ and $\beta(r) \cap V(F_2)$ are non-empty, then this inequality is strict.*

Proof. It is easily verified that (T, β_i) is a tree decomposition of F_i. For the second claim,

$$\begin{aligned}
\|(T, \beta_1)\| + \|(T, \beta_2)\| &= \sum_{t \in V(T)} |\beta_1(t)|^2 + \sum_{t \in V(T)} |\beta_1(t)|^2 \\
&= \sum_{t \in V(T)} \left(|\beta(t) \cap V(F_1)|^2 + |\beta(t) \cap V(F_2)|^2\right) \\
&\leq \sum_{t \in V(T)} |\beta(t)|^2.
\end{aligned}$$

If $\beta(r) \cap V(F_1)$ and $\beta(r) \cap V(F_2)$ are non-empty, then the inequality is strict. □

A *cut vertex* $v \in V(F)$ is a vertex such that $F - v$ has more connected components than F. The following Fact 6.3.19 is well-known.

Fact 6.3.19. *Every simple graph F contains a vertex $v \in V(F)$ which is not a cut vertex.*

Proof. Suppose without loss of generality that F is connected. Let $v, w \in V(F)$ be vertices such that their distance $\mathrm{dist}_F(v, w)$ is maximal. As usual, the distance of two vertices is the length of the shortest walk between them. Let $a \in V(F) \setminus \{v, w\}$ be arbitrary. Then there exists a path in F from a to w avoiding v. Hence, $F - v$ is connected. □

The following Lemma 6.3.20 yields vertices of low degree which do not cut.

Lemma 6.3.20. *Let $k \geq 0$. Let F be a simple graph of treewidth at most k. Then there exists a vertex $v \in V(F)$ of degree at most k which is not a cut vertex.*

Proof. Let (T, β) be a tree decomposition of F of width at most k such that $\|(T, \beta)\|$ is minimal. If T has only one vertex, then F has at most $k + 1$ vertices and Fact 6.3.19 yields the desired vertex. Otherwise, let $\ell \in V(T)$ be a leaf of T and let $r \in V(T)$ denote the adjacent vertex.

By minimality, $\beta(\ell) \not\subseteq \beta(r)$ and hence there exists a vertex $v \in \beta(\ell) \setminus \beta(r)$. This vertex has $\deg_F(v) \leq k$. Suppose that v is a cut vertex of F and write $F - v = F_1 + F_2$. Distinguish two cases:

If $V(F_1) \cap \beta(r) = \emptyset$, then F_1 is a subgraph of $F[\beta(\ell)]$. Without loss of generality, we may suppose that F_1 is connected. If $|N_F(v) \cap V(F_1)| \geq 2$, there exists, by Fact 6.3.19, a vertex $w \in V(F_1)$ which is not a cut vertex of F_1. This vertex is not a cut vertex of F since there are two neighbours of v in F_1. If $|N_F(v) \cap V(F_1)| = 1$, then apply the same argument to the vertex $w \in N_F(v) \cap V(F_1)$ and $F_1' := F_1 - w$ instead of F_1.

If $V(F_1) \cap \beta(r) \neq \emptyset$ and $V(F_2) \cap \beta(r) \neq \emptyset$, consider the tree decomposition (T', β') for the subgraph $F - (\beta(\ell) \setminus \beta(r))$ of $F - v$ obtained by deleting the leaf ℓ from (T, β). Write (T', β_1') and (T', β_2') for the tree decomposition for $F_1 - (\beta(\ell) \setminus \beta(r)) \cap V(F_1)$ and $F_2 - (\beta(\ell) \setminus \beta(r)) \cap V(F_2)$ defined for (T', β') as in Lemma 6.3.18. Since $\beta'(r)$ contains elements from both F_1 and F_2, it holds that $\|(T', \beta_1')\| + \|(T', \beta_2')\| < \|(T', \beta')\|$. Let S denote the tree obtained by taking two copies T_1 and T_2 of T' and connecting both copies of r to a new vertex s. Define a new tree decomposition (S, γ) for F by letting

$$\gamma(x) := \begin{cases} \beta(\ell), & \text{if } x = s, \\ \beta_1'(x), & \text{if } x \in V(T_1), \\ \beta_2'(x), & \text{if } x \in V(T_2). \end{cases}$$

By Lemma 6.3.18,

$$\|(S, \gamma)\| = |\beta(\ell)|^2 + \|(T, \beta_1')\| + \|(T, \beta_2')\| < |\beta(\ell)|^2 + \|(T', \beta')\| = \|(T, \beta)\|.$$

This contradicts the minimality of (T, β). □

Reducing Oddomorphisms

In the proof of Theorem 6.3.15, the aim is to find a $K_{2,h}$-minor in F given an oddomorphism from F to $K_{2,h}$. To that end, several lemmas are proven which fall short to yield such a minor in F but succeed to guarantee other useful features of F.

Lemma 6.3.21. *Let F and G be simple graphs admitting an oddomorphism $\varphi\colon F \to G$. If $v \in V(F)$ is an isolated vertex, then $\varphi|_{F-v}\colon F - v \to G$ is an oddomorphism.*

Proof. Clearly, $\psi := \varphi|_{F-v}$ is a homomorphism. Furthermore, for every $a \in V(F - v)$, $N_F(a) \cap \varphi^{-1}(v) = N_{F-v}(a) \cap \psi^{-1}(v)$ for every $v \in N_G(\varphi(a))$. Hence, all vertices in $F - v$ are even or odd with respect to ψ. The vertex v is even with respect to φ. Hence, ψ is an oddomorphism. □

Lemma 6.3.22. *Let F and G be simple graphs admitting an oddomorphism $\varphi\colon F \to G$. If there exist vertices $v \neq w \in V(F)$ such that $\varphi(v) = \varphi(w)$ and $N_F(v) = N_F(w)$, then $\varphi|_{F-\{v,w\}}$ is an oddomorphism from $F - \{v, w\}$ to G.*

Proof. Clearly, $\psi := \varphi|_{F-\{v,w\}}$ is a homomorphism. Furthermore, for every vertex $a \in V(F) \setminus \{v, w\}$ and $x \in N_G(\psi(a))$,

$$N_{F-\{v,w\}}(a) \cap \psi^{-1}(x) = (N_F(a) \cap \varphi^{-1}(x)) \setminus \{v, w\}.$$

The parities of both sets are the same since $v \in N_F(a)$ if, and only if, $w \in N_F(a)$. Hence, if a vertex in $V(F) \setminus \{v, w\}$ is φ-odd (φ-even), then it is ψ-odd (ψ-even). Furthermore, v and w have the same parity with respect to φ. Hence, $\psi^{-1}(\varphi(v))$ contains an odd number of odd vertices. All other fibres are unaffected. □

In the following Lemma 6.3.23, we write $X \triangle Y$ for the symmetric difference of sets X and Y, i.e. $X \triangle Y := (X \setminus Y) \cup (Y \setminus X)$.

Lemma 6.3.23. *Let F and G be simple graphs admitting an oddomorphism $\varphi\colon F \to G$. Let $v, w \in V(F)$ be such that $\varphi(v) = \varphi(w)$. Let C be a connected component of $F - \{v, w\}$. Suppose that the following hold:*

1. *$F[C \cup \{v, w\}]$ is connected,*
2. *$C \cap \varphi^{-1}(x)$ contains an even number of φ-odd vertices for every $x \in V(G)$, and*
3. *$(N_F(v) \triangle N_F(w)) \cap C \cap \varphi^{-1}(x)$ is of even size for every $x \in N_G(\varphi(v))$.*

Then there exists a strict minor F' of F admitting an oddomorphism $\varphi'\colon F' \to G$. Moreover, if $N_F(v) \cap N_F(w) \subseteq C$, then F' can be obtained from F by edge contractions.

Proof. To obtain F' from F, perform the following operations: For all vertices $x \in (N_F(v) \cap N_F(w)) \setminus C$, delete the edges vx and wx. Contract $C \cup \{v, w\}$ into a single vertex v^*. Then F' is a strict minor of F. Define φ' as follows:

$$\varphi'\colon x \mapsto \begin{cases} \varphi(v), & \text{if } x = v^*, \\ \varphi(x), & \text{if } x \in V(F') \setminus \{v^*\}. \end{cases}$$

Clearly, φ' is a homomorphism. To see that every vertex in F' is either φ'-odd or φ'-even, observe first that, for all $x \in N_G(\varphi'(v^*))$, by Item 3,

$$\begin{aligned} &|N_{F'}(v^*) \cap \varphi'^{-1}(x)| \\ &= |((N_F(v) \triangle N_F(w)) \setminus C) \cap \varphi^{-1}(x)| \\ &= |(N_F(v) \triangle N_F(w)) \cap \varphi^{-1}(x)| - |(N_F(v) \triangle N_F(w)) \cap C \cap \varphi^{-1}(x)| \\ &\equiv |(N_F(v) \triangle N_F(w)) \cap \varphi^{-1}(x)| \\ &\equiv |N_F(v) \cap \varphi^{-1}(x)| + |N_F(w) \cap \varphi^{-1}(x)| \mod 2. \end{aligned}$$

Hence, if both v and w are φ-odd or φ-even, then v^* is φ'-even. Otherwise, v^* is φ'-odd. For all other vertices $a \in V(F') \setminus \{v^*\}$ and $x \in N_G(\varphi'(a)) = N_G(\varphi(a))$,

$$N_{F'}(a) \cap \varphi'^{-1}(x) = \varphi^{-1}(x) \cap \begin{cases} N_F(a), & \text{if } a \notin N_F(v) \cup N_F(w), \\ (N_F(a) \setminus \{v\}) \cup \{v^*\}, & \text{if } a \in N_F(v) \setminus N_F(w), \\ (N_F(a) \setminus \{w\}) \cup \{v^*\}, & \text{if } a \in N_F(w) \setminus N_F(v), \\ N_F(a) \setminus \{v, w\}, & \text{otherwise.} \end{cases}$$

In all cases, $|N_{F'}(a) \cap \varphi'^{-1}(x)| \equiv |N_F(a) \cap \varphi^{-1}(x)| \mod 2$. Hence, if a vertex in $V(F')$ is φ-odd (φ-even), then it is φ'-odd (φ'-even). By Item 2, every φ'-fibre contains an odd number of odd vertices. □

Lemma 6.3.23 will be applied in the following form:

Corollary 6.3.24. *Let F and G be simple graphs admitting an oddomorphism $\varphi\colon F \to G$. For every φ-even vertex $x \in V(F)$ of degree two, there exists a strict minor F' of F such that there exists an oddomorphism $\varphi'\colon F' \to G$.*

Proof. It is shown that the assumptions of Lemma 6.3.23 are met. Let $v, w \in V(F)$ denote the two neighbours of x. Since x is φ-even, $\varphi(v) = \varphi(w)$. Let $C = \{x\}$. Clearly, this is a connected component of $F - \{v, w\}$ and $F[\{v, w, x\}]$ is connected. Furthermore, C does not contain any φ-odd vertices. For Item 3, observe that $(N_F(v) \triangle N_F(w)) \cap C = \emptyset$. Hence, Lemma 6.3.23 applies. □

Oddomorphisms to Complete Bipartite Graphs

Equipped with the preceding lemmas, we conduct the proof of Theorem 6.3.15.

Proof of Theorem 6.3.15. Towards a contradiction, let F be a minimal $K_{2,h}$-minor-free simple graph of treewidth at most two admitting an oddomorphism $\varphi\colon F \to K_{2,h}$. Let $V(K_{2,h}) = \{x_1, \dots, x_h, y_1, y_2\}$ and write $X_i := \varphi^{-1}(x_i)$ for $i \in [h]$ and $Y_j := \varphi^{-1}(y_j)$ for $j \in \{1, 2\}$ for the corresponding fibres along φ. Let $X := \bigcup_{i \in [h]} X_i$ and $Y := Y_1 \cup Y_2$.

Claim 6.3.24a. Every vertex in F has degree at least two.

Proof of Claim. By Lemma 6.3.21, all vertices in F have degree at least one. Any vertex v of degree one in F must be φ-odd. Since $\varphi(v)$ is of degree at least two in $K_{2,h}$, the vertex v must have at least two neighbours. ◁

Write $V_2 \subseteq V(F)$ for the set of all vertices of degree two. By Corollary 6.3.24, it can be supposed that all vertices in V_2 are odd with respect to φ. This implies that $V_2 \subseteq X$ because every odd vertex in Y is of degree at least $h \geq 3$. In particular, no two vertices in V_2 are adjacent. Consider the graph F' with $V(F') := V(F) \setminus V_2$ and

$$E(F') := \{uv \mid u, v \in V(F'),\ uv \in E(F) \vee (\exists x \in V_2.\ ux, xv \in E(F))\}. \tag{6.3}$$

Note that F' is a minor of F, i.e. it is obtained from F by contracting, for each $w \in V_2$, one of the edges incident to w. Hence, F' does not contain any isolated vertices.

For every vertex $w \in V(F')$, consider the map $p_w\colon N_F(w) \to N_{F'}(w)$ sending $x \in N_F(w) \cap V_2$ to the unique vertex $y \neq w$ such that $yx \in E(F)$ and $x \in N_F(w) \setminus V_2$ to x. By definition of F', p_w is surjective. Observe that $p_w|_{N_F(w) \setminus V_2}$ is injective. By Lemma 6.3.20, F' contains a vertex v^* with $\deg_{F'}(v^*) \in \{1, 2\}$ which is not a cut vertex in F'. In preparation for a case distinction, consider the following claims.

Claim 6.3.24b. $v^* \in Y$.

Proof of Claim. Since p_{v^*} is surjective, it holds that $\deg_{F'}(v^*) \leq \deg_F(v^*)$. By Claim 6.3.24a and since $v^* \notin V_2$, it holds that $\deg_F(v^*) \geq 3$. Hence, there exist $x \neq y \in N_F(v^*)$ such that $p_{v^*}(x) = p_{v^*}(y)$. Since $p_w|_{N_F(w) \setminus V_2}$ is injective, without loss of generality, $x \in V_2 \subseteq X$. This implies that $v^* \in Y$. ◁

Claim 6.3.24c. If $v \in V(F')$ is a cut vertex of F, then it is a cut vertex of F'. In particular, v^* is not a cut vertex of F.

Proof of Claim. By contraposition, suppose that, for all $a, b \in V(F') \setminus \{v\}$ which are connected in F', there exists a path connecting them in $F' - v$. Let $a, b \in V(F) \setminus \{v\}$ be arbitrary vertices in the same connected component of F. Set a' to a if $a \in V(F')$ and to any $x \in N_F(a)$ such that $x \neq v$ if $a \in V_2$. Define b' analogously. Observe that a' and b' lie in the same connected component of F. Hence, there exist a path connecting them. This path can be turned into a path in F' by shortcutting every vertex in V_2 appearing on it. Hence, there exists a path connecting them in $F' - v$. This path can be transformed into a path in $F - v$ by replacing shortcut edges $wy \in E(F' - v)$ with walks via any associated vertex in $x \in V_2$, i.e. $wx, xy \in E(F - v)$, cf. Equation (6.3). ◁

Claim 6.3.24d. Let $I \subseteq [h]$ be a set of even size. Let $H := F[Y_1 \cup Y_2 \cup \bigcup_{i \in I} X_i]$. Then $\deg_H(v)$ is even for every $v \in V(H)$.

Proof of Claim. For every vertex $v \in V(H)$, the number of neighbours of $\varphi(v)$ among y_1, y_2 and x_i for $i \in I$ is even. ◁

Without loss of generality, assume that $v^* \in Y_1$. Recall that $\deg_{F'}(v^*) \leq 2$. Hence, the image of p_{v^*} is of size at most two. Observe that $p_{v^*}(N_F(v^*) \cap V_2) \subseteq Y$ and $p_{v^*}(N_F(v^*) \setminus V_2) \subseteq X$. By Claim 6.3.24a and since $v^* \notin V_2$, the degree of v^* in F is at least 3. Hence, there must be at least one vertex from X in the image of p_{v^*}. Distinguish cases based on the parity of v^* with respect to φ:

1. v^* is φ-even.
   Distinguish cases:
   a) The image of p_{v^*} has size one.
      Then there exists a fibre X_i containing two neighbours u_1, u_2 of v^*. Since $p_{v^*}(u_1) = p_{v^*}(u_2)$, it holds that $u_1, u_2 \in V_2$ and that they share their second neighbour, i.e. $N_F(u_1) = N_F(u_2)$. Hence, Lemma 6.3.22 applies and F is not minimal.
   b) The image of p_{v^*} contains one vertex in X and one vertex in Y.
      By injectivity of p_{v^*} on $N_F(v^*) \setminus V_2$, all but one neighbour of v^* are in V_2. Since every fibre in X contains an even number of neighbours of v^*, there must be a fibre containing two neighbours u_1, u_2 of v^* from V_2. As in the previous case, $N_F(u_1) = N_F(u_2)$ and F is not minimal by Lemma 6.3.22.

c) The image of p_{v^*} contains two vertices v_1, v_2 in Y.
In this case, $N_F(v^*) \subseteq V_2$. Hence, all neighbours of v^* are φ-odd and in particular $v_1, v_2 \in Y_2$ lie in the same fibre. The set $C := N_F(v^*) \cup \{v^*\}$ is a connected component of $F - \{v_1, v_2\}$. Clearly, $F[C \cup \{v_1, v_2\}]$ is connected. The φ-odd vertices in C are precisely those in $N_F(v^*)$ and every φ-fibre contains an even number of them since v^* is φ-even.
If one of the φ-fibres contains a pair of vertices with the same image under p_{v^*}, then Lemma 6.3.22 applies as above. Hence, it may be supposed that every fibre contains either no neighbour of v^* or exactly one vertex adjacent to v_1 and one vertex adjacent to v_2. This implies that Item 3 of Lemma 6.3.23 is met. F is not minimal.

2. v^* is φ-odd.
The vertex v^* has at least one neighbour $u_i \in X_i$ in every fibre $i \in [h]$. In particular, it is of degree at least h. Distinguish further cases:
   a) The image of p_{v^*} has size one.
   In this case, the vertices $v^*, u_1, \dots, u_h$ together with the vertex in the image of p_{v^*} induce a subgraph $K_{2,h}$ of F.
   b) The image of p_{v^*} contains one vertex in X and one vertex v in Y.
   By injectivity of p_{v^*} on $N_F(v^*) \setminus V_2$, v^* has at least $h-1$ neighbours $u_2, \dots, u_h$ in V_2.
      i. If h is even, then F is Eulerian by Claim 6.3.24d. The same remains true if $u_2, \dots, u_h$ (an odd number of vertices) are deleted and an edge between v and v^* is inserted. Hence, there exists a path in F from v to v^* avoiding the vertices $u_2, \dots, u_h$. This path can be contracted to a path of length two to yield a $K_{2,h}$-minor of F.
      ii. If h is odd, consider the subgraph $H := F[Y_1 \cup Y_2 \cup X_1 \cup \dots \cup X_{h-1}]$. By Claim 6.3.24d, H is Eulerian. Deleting $u_2, \dots, u_{h-1}$ (an odd number of vertices) vertices from H and adding an edge between v and v^* yields another Eulerian graph. Thus, there exists a path in H from v to v^* avoiding $u_2, \dots, u_{h-1}$. As above, this yields a minor $K_{2,h-1}$ in H and, together with the neighbour $u_h \in X_h$ of v^*, a minor $K_{2,h}$ in F.
   c) The image of p_{v^*} contains two vertices v_1, v_2 in Y.
   By Claim 6.3.24c, v^* is not a cut vertex of F. Hence, there exists a path in F from v_1 to v_2 avoiding v^*. This path can be contracted to yield a minor $K_{2,h}$ of F. □

6.4 Games

In this section, we demonstrate how games from structural graph theory and finite model theory can be orchestrated to show that certain graph classes are homomorphism distinguishing closed. As examples serve the classes of graphs of bounded treewidth, bounded treedepth, and $\mathcal{T}_k^q$. A previously unpublished result presented here asserts that the class of graphs of bounded pathwidth is homomorphism distinguishing closed. As a corollary, we complete Theorem 5.0.1.

6.4.1 Bounded Treewidth

In [37], the original CFI graphs were used to separate C^k-equivalence from isomorphism. Once the CFI graphs G_0 and G_1 over some graph G are constructed, it is not hard to prove that they are non-isomorphic, cf. Corollary 6.3.5. The harder assertion to establish is that they are C^{k+1}-equivalent assuming that G is a graph of treewidth greater than k. To that end, two ingredients are crucial:

The first ingredient is a model comparison game which characterises C^k-equivalence [37]. Such games are played by two players called Spoiler and Duplicator on the graphs G_0 and G_1. Spoiler aims to exhibit dissimilarities between G_0 and G_1 while Duplicator tries to hide them. The dissimilarities between G_0 and G_1 lie at the twisted vertex $u \in V(G)$ such that $G_u \cong G_1$. The second ingredient is a node searching game which characterises treewidth [168]. In such a game, a robber tries to evade a number of cops which are positioned on the vertices of a graph. The robber moves along the edges of the graph without traversing the vertices occupied by the cops. Robber wins the game against $k+1$ cops if, and only if, the graph which the game is played on has treewidth greater than k. Intuitively, the position of the robber corresponds to the twisted vertex while the positions of the cops correspond to the pebbles placed by Spoiler. If the robber succeeds to escape from the cops, Duplicator can hide the twist from Spoiler.

The above argument has been used often, cf. e.g. [11, 12]. Combined with the insights from Section 6.3.2, it allows to show that a graph class such as the class of graphs of bounded treewidth is closed under weak oddomorphisms and thus homomorphism distinguishing closed.

Theorem 6.4.1 ([134, Corollary 13 and Theorem 2])**.** *For $k \geq 0$, the class of graphs of treewidth at most k is closed under weak oddomorphisms. In particular, it is homomorphism distinguishing closed.*

Formally, the key step when proving Theorem 6.4.1 is to show that when G is a connected simple graph of treewidth greater than k, then its CFI graphs G_0 and G_1 are homomorphism indistinguishable over all graphs of treewidth at most k. We omit the proof of Theorem 6.4.1 but include a similar argument for the class of graphs of bounded pathwidth in full detail, cf. Theorem 6.4.6.

Due to the many characterisations of homomorphism indistinguishability over graphs of bounded treewidth, cf. Theorems 3.4.3 and 3.4.4 and Corollary 4.4.1, Theorem 6.4.1 is arguably one of the most relevant results on the homomorphism distinguishing closure. In particular, it has the following corollary, which establishes the final claim in Theorem 5.0.1.

Corollary 6.4.2. *For every $t \geq 1$, there exist simple graphs G and H such that the system* $\mathrm{SA}^{3t-1}(G, H)$ *has a non-negative rational solution but the system* $\mathrm{L}^t(G, H)$ *has no real solution.*

Proof. Since $\operatorname{tw}(K_{3t}) = 3t - 1$, there exist, by Theorem 6.4.1, two simple graphs G and H such that $G \equiv_{\mathcal{TW}_{3t-2}} H$ and $\hom(K_{3t}, G) \neq \hom(K_{3t}, H)$. By Corollary 4.4.1, $\mathrm{SA}^{3t-1}(G, H)$ has a non-negative rational solution. By Observation 5.2.2 and Theorem 5.1.1, $\mathrm{L}^t(G, H)$ has no real solution. Note that, by Theorems 6.4.1 and 6.3.9, G and H can be taken to be the odd and even CFI graph over K_{3t}. □

Before we turn to other graph classes to which the proof strategy yielding Theorem 6.4.1 applies, we state the following Question 6.4.3, which concerns the lower bound missing in Theorem 5.0.1. Since $\mathcal{TW}_{t-1} \subseteq \mathcal{L}_t$ by Lemma 5.2.11, feasibility of $\mathrm{L}^t(G, H)$ implies the existence of a non-negative rational solution for $\mathrm{SA}^t(G, H)$. Missing in Theorem 5.0.1 is a tight lower bound on the number of Lasserre levels necessary to ensure feasibility of a given Sherali–Adams level:

Question 6.4.3. *Do there exist, for every $t \geq 3$, simple graphs G and H such that* $\mathrm{L}^{t-1}(G, H)$ *has a real solution but* $\mathrm{SA}^t(G, H)$ *has no non-negative rational solution?*

Note that it is not even clear whether $\mathcal{TW}_{t-1} \not\subseteq \mathcal{L}_{t-1}$ for $t \geq 3$. Moreover, it is open whether $\mathcal{L}_t$ is homomorphism distinguishing closed, as predicted by Conjecture 6.0.2, cf. Theorem 5.2.12.

6.4.2 Bounded Treedepth and Graphs with Pebble Forest Covers of Bounded Depth

The proof strategy that yields Theorem 6.4.1 can be adapted to show that other graph classes are homomorphism distinguishing closed. The graph class in question must be such that its homomorphism indistinguishability relation is characterised by a pebble game and such that membership in the class if characterised by a node searching game. Examples for such classes are the class of graphs of bounded treedepth and $\mathcal{T}_q^k$, cf. Definition 2.1.3. See Theorems 3.4.5 and 3.4.6 for characterisations of their homomorphism indistinguishability relations in terms of counting logic equivalences.

Theorem 6.4.4 ([68, Theorem 3])**.** *For $k \geq 1$, the class of graphs of treedepth at most k is closed under weak oddomorphisms. In particular, it is homomorphism distinguishing closed.*

Theorem 6.4.5 ([5, Theorem 29])**.** *For $k, q \geq 1$, the class $\mathcal{T}_k^q$ is closed under weak oddomorphisms. In particular, it is homomorphism distinguishing closed.*

6.4.3 Bounded Pathwidth

In order to present the proof strategy which yields Theorems 6.4.1, 6.4.4, and 6.4.5, we prove that the classes of graphs of bounded pathwidth are homomorphism distinguishing closed. As above, the proof has two key ingredients: a model comparison game characterising homomorphism indistinguishability over graphs of bounded pathwidth [129] and a node searching game characterising bounded pathwidth [25].

Theorem 6.4.6. *For $k \geq 0$, the class of graphs of pathwidth at most k is closed under weak oddomorphisms. In particular, it is homomorphism distinguishing closed.*

We first recall the following model comparison game introduced by Montacute & Shah [129]. To that end, let X be a finite set and $k, \ell \geq 1$. Let $\boldsymbol{s} \in ([k] \times X)^\ell$ be a sequence of pairs $\boldsymbol{s} = [(p_1, x_1), \dots, (p_\ell, x_\ell)]$. We say that $\boldsymbol{s}$ *contains* $p \in [k]$ if there is $i \in [\ell]$ such that $p_i = p$. For $p \in [k]$, define $\text{last}_p(\boldsymbol{s})$ as the element x_i such that $i \in [\ell]$ is maximal with $p_i = p$. If $\boldsymbol{s}$ does not contain p, then $\text{last}_p(s)$ is undefined. For $i \in [\ell]$, write $\boldsymbol{s}[1, i] := [(p_1, x_1), \dots, (p_i, x_i)]$.

Definition 6.4.7 ([129, Definition 5.7])**.** Let $k \geq 1$ and G and H be simple graphs. The *all-in-one bijective k-pebble game* on G and H is played by Spoiler and Duplicator and consists of a single round:

1. Spoiler chooses an $n \geq 1$ and a sequence of pebbles $\boldsymbol{p} = [p_1, \dots, p_n] \in [k]^n$.
2. Duplicator chooses a bijection $h_{\boldsymbol{p}} \colon V(G)^n \to V(H)^n$.
3. Spoiler chooses a sequence of pebble placements $\boldsymbol{s} = [(p_1, v_1), \dots, (p_n, v_n)] \in ([k] \times V(G))^n$.

Let $\boldsymbol{t} = [(p_1, w_1), \dots, (p_n, w_n)]$ where $(w_1, \dots, w_n) = h_{\boldsymbol{p}}(v_1, \dots, v_n)$. Duplicator wins if, for every $i \in [n]$, the map $v^q \mapsto w^q$ for $q \in [k]$ where $v^q := \text{last}_q(\boldsymbol{s}[1, i])$ and $w^q := \text{last}_q(\boldsymbol{t}[1, i])$ is a local isomorphism.

By the following Theorem 6.4.8, Duplicator wins the above game on the graphs G and H if, and only if, they are homomorphism indistinguishable over the class of graphs of bounded pathwidth.

Theorem 6.4.8. *Let $k \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are homomorphism indistinguishable over the class of graphs of pathwidth at most $k - 1$,*
2. *Duplicator always wins the all-in-one bijective k-pebble game on G and H.*

Proof. The theorem follows from [129, Theorem 5.9, Corollary 5.13]. We sketch a self-contained proof relying on arguments from Chapter 4.

For $\boldsymbol{p} \in [k]^n$, write $\mathcal{W}^{\boldsymbol{p}} \subseteq \mathcal{G}(n)$ for the class of n-labelled graphs $\boldsymbol{F} = (F, \boldsymbol{u})$ such that

1. F has at most n vertices, all of which appear in $\boldsymbol{u} \in V(F)^n$,
2. the relation $u_1 \leq u_2 \leq \dots \leq u_n$ is a total order on $V(F)$, and

3. the map $V(F) \to [k]$ given by $u_i \mapsto p_i$ together with the order $\leq$ form a k-pebble forest cover of F, cf. Definition 2.1.3.

We collect the following claims:

Claim 6.4.8a. For every $n \geq 1$ and $\boldsymbol{p} \in [k]^n$, every graph in $\operatorname{soe}(\mathcal{W}^{\boldsymbol{p}})$ has pathwidth at most $k-1$.

Proof of Claim. The claim essentially follows from the proof of [68, Theorem 14]. More explicitly, let P_n denote the path graph on vertex set $[n]$ and define a path decomposition of $\operatorname{soe}(\boldsymbol{F})$ for $\boldsymbol{F} = (F, \boldsymbol{u}) \in \mathcal{W}^{\boldsymbol{p}}$ by letting for $\ell \in [n]$

$$\beta(\ell) := \{\operatorname{last}_p(\boldsymbol{u}[1, \ell]) \mid p \in [k]\}.$$

By definition of $\mathcal{W}^{\boldsymbol{p}}$, (P_n, β) is a path decomposition of $\operatorname{soe}(\boldsymbol{F})$. ◁

Claim 6.4.8b. For every simple graph F of pathwidth at most $k-1$, there exist $n \geq 1$, $\boldsymbol{p} \in [k]^n$, and $\boldsymbol{u} \in V(F)^n$ such that $(F, \boldsymbol{u}) \in \mathcal{W}^{\boldsymbol{p}}$.

Proof of Claim. Let (P, β) be a path decomposition of F. Let $\leq'$ be a total order on $V(P)$ such that adjacent vertices in P are successors in $\leq'$. Then $\leq'$ induces a total order $\leq$ on $V(F)$ where $u \leq v$ if, and only if, the $\leq'$-least $x, y \in V(P)$ such that $u \in \beta(x)$ and $v \in \beta(y)$ are such that $x \leq' y$. A pebbling function $p \colon V(F) \to [k]$ can be defined vertex by vertex in the order $\leq$ such that p is injective on every bag $\beta(x)$, $x \in V(P)$. From p and $\leq$, the tuples $\boldsymbol{p}$ and $\boldsymbol{u}$ can be constructed, as desired. ◁

Recall Definition 4.3.15 and observe that $\mathcal{W}^{\boldsymbol{p}}$ is gluing-closed and contains $\boldsymbol{1} \in \mathcal{G}(k)$, cf. Figure 3.3a. Thus, Theorem 4.3.16 applies.

First assume that G and H are homomorphism indistinguishable over the class of graphs of pathwidth at most $k-1$. Let $\boldsymbol{p} \in [k]^n$ be arbitrary. By Claim 6.4.8a, G and H are homomorphism indistinguishable over $\operatorname{soe}(\mathcal{W}^{\boldsymbol{p}})$. By Theorem 4.3.16, there exists a bijection $\pi \colon V(G)^n \to V(H)^n$ such that $\boldsymbol{F}_G(\boldsymbol{v}) = \boldsymbol{F}_H(\pi(\boldsymbol{v}))$ for all $\boldsymbol{v} \in V(G)^n$ and $\boldsymbol{F} \in \mathcal{W}^{\boldsymbol{p}}$. We show that this bijection lets Duplicator win the game.

Let $i \in [n]$ and consider the map $v^q \mapsto w^q$ for $q \in [k]$ where $v^q := \operatorname{last}_q(\boldsymbol{s}[1, i])$ and $w^q := \operatorname{last}_q(\boldsymbol{t}[1, i])$. For $q \in [k]$, write $j^q \in [n]$ for the index of v^q in $\boldsymbol{s}$. Consider the n-labelled graph $\boldsymbol{F} \in \mathcal{W}^{\boldsymbol{p}}$ obtained from $\boldsymbol{1}$ with vertex set $[n]$ by (1) merging the vertices j^q and $j^{q'}$ for every $q \neq q' \in [k]$ such that $v^q = v^{q'}$, (2) drawing an edge between j^q and $j^{q'}$ for every for every $q \neq q' \in [k]$ such that v^q are adjacent $v^{q'}$. Since all vertices of $\boldsymbol{F}$ are labelled, its homomorphism vector $\boldsymbol{F}_G$ has entries from $\{0, 1\}$, cf. Observation 3.2.11. It encodes the isomorphism type of the subgraph of G induced by $\{v^1, \dots, v^k\}$. Hence, $\boldsymbol{F}_G(\boldsymbol{v}) = \boldsymbol{F}_H(\pi(\boldsymbol{v}))$ for $\boldsymbol{v} \in V(G)^n$ implies that $v^q \mapsto w^q$ is a local isomorphism.

Conversely, let $\boldsymbol{p} \in [k]^n$ and $\pi \colon V(G)^n \to V(H)^n$ be the bijection which allows Duplicator to win. By Claim 6.4.8b and Theorem 4.3.16, it suffices to show that $\boldsymbol{F}_G(\boldsymbol{v}) = \boldsymbol{F}_H(\pi(\boldsymbol{v}))$ for every $\boldsymbol{F} \in \mathcal{W}^{\boldsymbol{p}}$ and $\boldsymbol{v} \in V(G)^n$. This holds for the elements of $\mathcal{W}^p$ which were considered in the other proof direction. Observing that $\mathcal{W}^p$ is generated by these under gluing products, we conclude the proof. □

Equipped with Theorem 6.4.8, we recall the following definition from [25, p. 282]. Let G be a simple graph. A *search* of G is a sequence $(X_1, \dots, X_m)$ of subsets $X_i \subseteq V(G)$ such that $X_1 = \emptyset$ and $X_{i+1} \subseteq X_i$ or $X_i \subseteq X_{i+1}$ for all $i \in [m-1]$. Define $B_1 := V(G)$ and inductively B_i for $2 \leq i \leq m$ as the set of all vertices $v \in V(G)$ for which there exists a path P between v and some vertex of B_{i-1} with $V(P) \cap X_i = \emptyset$. The search is *successful* if $B_m = \emptyset$.

Theorem 6.4.9 ([25, (5.1)]). *For a simple graph G and an integer $k \geq 1$, the following are equivalent:*

1. *G has pathwidth at most $k-1$,*
2. *there exists a successful search $(X_1, \dots, X_m)$ of G with $|X_i| \leq k$ for all $i \in [m]$.*

The following Lemma 6.4.10 allows to convert an unsuccessful search into a sequence of bijections from which Duplicator's winning strategy for the game in Definition 6.4.7 is constructed.

Lemma 6.4.10 ([134, Lemma 11]). *Let G be a connected graph. Let $u, v \in V(G)$. Let P be a path in G from u to v. Then there exists an isomorphism $\varphi\colon G_{\{u\}} \to G_{\{v\}}$ such that*

1. *$\rho(\varphi(w, S)) = w$ for all $(w, S) \in V(G_{\{u\}})$ and*
2. *$\varphi(w, S) = (w, S)$ for all $(w, S) \in V(G_{\{u\}})$ with $w \in V(G) \setminus P$.*

The key step towards Theorem 6.4.6 is the following Lemma 6.4.11.

Lemma 6.4.11. *Let $k \geq 1$. If G is a connected simple graph of pathwidth at least k, then G_0 and G_1 are homomorphism indistinguishable over the class of graphs of pathwidth at most $k-1$.*

Proof. Let $u_0 \in V(G)$ be a vertex. The proof is by construction of a winning strategy for Duplicator in the all-in-one bijective k-pebble game from Definition 6.4.7 played on G_0 and $G_{u_0} \cong G_1$ using unsuccessful searches from Theorem 6.4.9.

For $\boldsymbol{r} = [(p_1, v_1), \dots, (p_n, v_n)] \in ([k] \times V(G))^n$, define the set of *active* vertices $A(\boldsymbol{r}) := \{\mathrm{last}_p(\boldsymbol{r}) \mid p \in [k]\}$. Define a search $X(\boldsymbol{r}) = (X_1, \dots, X_{2n})$ on G by letting $X_1 := \emptyset$ and, for $i \geq 1$,

$$X_{2i} := A(\boldsymbol{r}[1, i]), \qquad X_{2i+1} := X_{2i} \setminus \{\mathrm{last}_{p_{i+1}}(\boldsymbol{r}[1, i])\}.$$

Clearly, $|X_i| \leq k$ for all $i \in [2n]$. Moreover, $X_{2i-1} \subseteq X_{2i}$ for all $i \in [n]$ and $X_{2i} \supseteq X_{2i+1}$ for all $i \in [n-1]$. Thus, $X(\boldsymbol{r})$ is a search as in Theorem 6.4.9. Write $B(\boldsymbol{r})$ for the set B_{2n} defined for $X(\boldsymbol{r}) = (X_1, \dots, X_{2n})$. By Theorem 6.4.9, $B(\boldsymbol{r}) \neq \emptyset$.

Fix $\boldsymbol{p} = [p_1, \dots, p_n] \in [k]^n$. We inductively construct the bijection $h_{\boldsymbol{p}}\colon V(G_0)^n \to V(G_{u_0})^n$ which lets Duplicator win the game. That is, by induction on $i \in [n]$, we construct bijections $h_i\colon V(G_0)^i \to V(G_{u_0})^i$ satisfying Duplicator's winning condition in Definition 6.4.7. The final bijection $h_{\boldsymbol{p}}$ will be defined to be h_n.

For a tuple $\boldsymbol{s} = [(p_1, (v_1, S_1)), \ldots, (p_i, (v_i, S_i))] \in ([k] \times V(G_0))^i$, define $\rho(\boldsymbol{s}) := [(p_1, v_1), \ldots, (p_i, v_i)] \in ([k] \times V(G))^i$.

As base case, we define a bijection $h_1 \colon V(G_0) \to V(G_{u_0})$. By construction, there exists a vertex $u_1^{\boldsymbol{s}} \in B(\rho(\boldsymbol{s}[1,1]))$ and a path P from $u_1^{\boldsymbol{s}}$ to u_0 that does not traverse v_1. Let $\varphi_1^{\boldsymbol{s}} \colon G_{u_1^{\boldsymbol{s}}} \to G_{u_0}$ denote the isomorphism constructed from P via Lemma 6.4.10. The image of (v_1, S_1) under h is defined as $\varphi_1^{\boldsymbol{s}}(v_1, S_1)$. Since $u_1 \neq v_1$, we can regard $(v_1, S_1) \in V(G_0)$ as a vertex of $V(G_{u_1})$ and the image of (v_1, S_1) under $\varphi_1^{\boldsymbol{s}}$ is well-defined. By using for every vertex of G_0 in $\rho^{-1}(v_1)$ the same map $\varphi_1^{\boldsymbol{s}}$, it follows that h is a bijection.

For subsequent steps, suppose that we have constructed a bijection $h_i \colon V(G_0)^i \to V(G_{u_0})^i$ for some $1 \leq i < n$ which induces a local isomorphism on $A(\rho(\boldsymbol{s}[1,i]))$. In order to define the image of $[(v_1, S_1), \ldots, (v_{i+1}, S_{i+1})]$ under $h_{i+1} \colon V(G_0)^{i+1} \to V(G_{u_0})^{i+1}$, let $u_i^{\boldsymbol{s}} \in B(\rho(\boldsymbol{s}[1,i]))$ and $\varphi_i^{\boldsymbol{s}} \colon G_{u_i} \to G_{u_0}$ denote the previously constructed auxiliary objects.

By construction, there exists some vertex $u_{i+1}^{\boldsymbol{s}} \in B(\rho(\boldsymbol{s}[1,i+1]))$ and a path P from $u_{i+1}^{\boldsymbol{s}}$ to $u_i^{\boldsymbol{s}}$ which does not traverse the vertices in $A(\rho(\boldsymbol{s}[1,i+1]))$. Let $\psi \colon G_{u_{i+1}^{\boldsymbol{s}}} \to G_{u_i^{\boldsymbol{s}}}$ denote the isomorphism constructed from P in Lemma 6.4.10 and let $\varphi_{i+1}^{\boldsymbol{s}} := \varphi_i^{\boldsymbol{s}} \circ \psi \colon G_{u_{i+1}^{\boldsymbol{s}}} \to G_{u_0}$. As above, $v_{i+1} \neq u_{i+1}^{\boldsymbol{s}}$ and $(v_{i+1}, S_{i+1}) \in V(G_0)$ may be regarded as a vertex of $G_{u_{i+1}^{\boldsymbol{s}}}$. Thus, we may define the image of $[(v_1, S_1), \ldots, (v_{i+1}, S_{i+1})]$ as the tuple obtained from $h_i([(v_1, S_1), \ldots, (v_i, S_i)])$ by appending $\varphi_{i+1}^{\boldsymbol{s}}(v_{i+1}, S_{i+1})$. It follows as in the base case that h_{i+1} is a bijection.

It remains to verify that h_{i+1} is a local isomorphism on $A(\rho(\boldsymbol{s}[1,i+1]))$. Let $(v_j, S_j) \in A(\rho(\boldsymbol{s}[1,i+1]))$. Write (v_j, T_j) and (w_{i+1}, T_{i+1}) for the images of (v_j, S_j) and (v_{i+1}, S_{i+1}) under h_{i+1}, respectively. Then v_j and v_{i+1} do not appear on any of the paths used to construct the isomorphisms $\varphi_j^{\boldsymbol{s}}, \ldots, \varphi_{i+1}^{\boldsymbol{s}}$. Hence, $(v_j, T_j) = \varphi_j^{\boldsymbol{s}}(v_j, S_j) = \varphi_{i+1}^{\boldsymbol{s}}(v_j, S_j)$, by Lemma 6.4.10, and $(v_{i+1}, T_{i+1}) = \varphi_{i+1}^{\boldsymbol{s}}(v_{i+1}, S_{i+1})$, by definition. Since $\varphi_{i+1}^{\boldsymbol{s}}$ is an isomorphism, it holds that (v_{i+1}, S_{i+1}) and (v_j, S_j) are adjacent, non-adjacent, or equal if, and only if, (v_{i+1}, T_{i+1}) and (v_j, T_j) are so. □

This concludes the preparations for the proof of Theorem 6.4.6, which is similar to the proof of [134, Theorem 2].

Proof of Theorem 6.4.6. Given Theorem 6.3.9, it remains to argue that the class of graphs of pathwidth at most k is closed under weak oddomorphisms. Let F be a graph of pathwidth at most k and $F \to G$ a weak oddomorphism. Suppose that G has a connected component G' which has pathwidth greater than k. By Lemma 6.3.11, there exists a minor F' of F such that F' admits an oddomorphism to G'. By Theorem 6.3.8, $\hom(F', G'_0) \neq \hom(F', G'_1)$. However, this contradicts Lemma 6.4.11. Hence, G is of pathwidth at most k. □

6.5 Classification of Homomorphism Distinguishing Closed Essentially Profinite Graph Classes

The central result of this section is a classification of the homomorphism distinguishing closed graph classes which are in a sense finite. This classification confirms Conjecture 6.0.2 for all such graph classes. The distinguishing power of homomorphism counts from finitely many graphs is of particular importance in practice. Applications include the design of graph kernels [102], motif counting [7, 127], or machine learning on graphs [19, 136, 80]. A theoretic interest stems for example from database theory where homomorphism counts correspond to results of queries under bag-semantics [39, 103], see also [41, 38].

Since every homomorphism distinguishing closed graph class is closed under disjoint unions, cf. Lemma 6.2.2, infinite graph classes arise naturally when studying the semantic properties of the homomorphism indistinguishability relations of finite graph classes. Nevertheless, the infinite graph classes arising in this way are *essentially finite*, i.e. they exhibit only finitely many distinct connected components. One may generalise this definition slightly by observing that all graphs F admitting a homomorphism into some fixed graph G have chromatic number bounded by the chromatic number of G. Thus, in order to make a graph class $\mathcal{F}$ behave much like an essentially finite class, it suffices to impose a finiteness condition, for every simple graph K, on the subfamily of all K-colourable graphs in $\mathcal{F}$.

Formally, for a graph F, write $\Gamma(F)$ for the set of connected components of F. For a graph class $\mathcal{F}$, define the graph class $\Gamma(\mathcal{F})$ as the union of the $\Gamma(F)$ for every $F \in \mathcal{F}$. Recall that $\mathcal{F}_K := \{F \in \mathcal{F} \mid \hom(F, K) > 0\}$, for a graph class $\mathcal{F}$ and a graph K, denotes the class of K-colourable graphs in $\mathcal{F}$.

Definition 6.5.1. A graph class $\mathcal{F}$ is *essentially finite* if $\Gamma(\mathcal{F})$ is finite. It is *essentially profinite* if $\mathcal{F}_K$ is essentially finite for every simple graph K.

Clearly, every finite graph class is essentially finite and hence essentially profinite. Another example for essentially profinite classes is the class $\mathcal{K}$ of all complete graphs. It represents a special case of the following construction from [150, Theorem 6.16], i.e. $\mathcal{K} = \mathcal{K}^{\mathbb{N}}$: For every $S \subseteq \mathbb{N}$, the family

$$\mathcal{K}^S := \{K_{n_1} + \cdots + K_{n_r} \mid r \in \mathbb{N}, \{n_1, \dots, n_r\} \subseteq S\} \tag{6.4}$$

is essentially profinite. In particular, there are uncountably many such families of graphs. Note that one may replace the sequence of complete graphs $(K_n)_{n \in \mathbb{N}}$ in Equation (6.4) by any other sequence of connected graphs $(F_n)_{n \in \mathbb{N}}$ such that the sequence of chromatic numbers $(\chi(F_n))_{n \in \mathbb{N}}$ takes every value only finitely often.

Every graph F of an essentially finite family $\mathcal{F}$ can be represented uniquely as the vector $\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F})}$ whose C-th entry for $C \in \Gamma(\mathcal{F})$ is the number of occurrences of C a connected component of F. The classification of the homomorphism distinguishing closed essentially profinite graph classes can now be stated as follows.

Theorem 6.5.2. *For an essentially profinite graph class $\mathcal{F}$, the following are equivalent:*

1. *$\mathcal{F}$ is homomorphism distinguishing closed,*
2. *for every simple graph K, if $\vec{K} \in \operatorname{span}\{\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F}_K \cup \{K\})} \mid F \in \mathcal{F}_K\}$, then $K \in \mathcal{F}$,*
3. *$\mathcal{F}_K$ is homomorphism distinguishing closed for every simple graph K.*

The remainder of this section is dedicated to the proof of Theorem 6.5.2. Note that Theorem 6.5.2 directly implies Conjecture 6.0.2 for essentially profinite graph classes. In particular, Corollary 6.5.3 implies that all essentially profinite union-closed minor-closed graph classes are homomorphism distinguishing closed. For example, for every graph G, the union-closure of the class of minors of G is homomorphism distinguishing closed, cf. [150, Question 4].

Corollary 6.5.3. *Let $\mathcal{F}$ be an essentially profinite graph class. If $\mathcal{F}$ is closed under disjoint unions and taking summands, i.e. $F_1, F_2 \in \mathcal{F}$ if, and only if, $F_1 + F_2 \in \mathcal{F}$, then $\mathcal{F}$ is homomorphism distinguishing closed.*

Proof. If $\mathcal{F}$ is union-closed and closed under summands, then $\Gamma(\mathcal{F}) \subseteq \mathcal{F}$ and every graph K such that $\Gamma(K) \subseteq \Gamma(\mathcal{F})$ is itself in $\mathcal{F}$. Thus, Theorem 6.5.2 yields the claim. □

Corollary 6.5.3 also applies to the graph class from Equation (6.4). It illustrates that the world of homomorphism distinguishing closed graph classes is rather complicated. In particular, there exist infinite chains and antichains of homomorphism distinguishing closed graph classes.

Corollary 6.5.4 ([150, Theorem 6.16])**.** *For every $S \subseteq \mathbb{N}$, the graph class $\mathcal{K}^S$ is homomorphism distinguishing closed. In particular, there are uncountably many homomorphism distinguishing closed graph classes.*

Towards proving Theorem 6.5.2, we first make the following general observation: Considering essentially profinite graph classes is very natural in light of the following Lemma 6.5.5. For every graph K, the subset $\mathcal{F}_K$ of $\mathcal{F}$ is the object prescribing whether $K \in \operatorname{cl}(\mathcal{F})$.

Lemma 6.5.5. *Let $\mathcal{F}$ be a graph class and K be a simple graph. Then $K \in \operatorname{cl}(\mathcal{F})$ if, and only if, $K \in \operatorname{cl}(\mathcal{F}_K)$.*

Proof. The backward implication is immediate since $\mathcal{F}_K \subseteq \mathcal{F}$ and thus $\operatorname{cl}(\mathcal{F}_K) \subseteq \operatorname{cl}(\mathcal{F})$. Conversely, suppose that $K \in \operatorname{cl}(\mathcal{F})$. Let G and H be simple graphs such that $G \equiv_{\mathcal{F}_K} H$. Then, by Equation (2.2), $G \times K \equiv_{\mathcal{F}} H \times K$ since $\hom(F, K) = 0$ for all $F \in \mathcal{F} \setminus \mathcal{F}_K$. By assumption, $\hom(K, G \times K) = \hom(K, H \times K)$, and therefore $\hom(K, G) = \hom(K, H)$ since $\hom(K, K) > 0$. Thus, $K \in \operatorname{cl}(\mathcal{F}_K)$. □

The proof of Theorem 6.5.2 is based on a generalisation of a result by Kwiecień, Marcinkowski, & Ostropolski-Nalewaja [103]. They proved the following Theorem 6.5.6 for finite graph classes. The extension to essentially finite graph classes does not require much additional work but might make core ideas appear more transparently.

Theorem 6.5.6 ([103, Lemma 31]). *Let K be a simple graph. Let $\mathcal{F}$ be an essentially finite family of K-colourable graphs. Then*

$$K \in \mathrm{cl}(\mathcal{F}) \iff \vec{K} \in \operatorname{span}\{\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F}\cup\{K\})} \mid F \in \mathcal{F}\}.$$

Proof. Write $\mathcal{C} := \Gamma(\mathcal{F} \cup \{K\})$. For the backward direction, suppose that $\vec{K} \in \operatorname{span}\{\vec{F} \in \mathbb{R}^{\mathcal{C}} \mid F \in \mathcal{F}\}$. Observe that this implies that $\Gamma(\mathcal{F}) = \mathcal{C}$, i.e. all connected components of K appear as connected components of some $F \in \mathcal{F}$. Write $\vec{K} = \sum_{i=1}^{r} \alpha_i \vec{F_i}$ for some $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ and $F_1, \dots, F_r \in \mathcal{F}$. Also write $K = \coprod_{C\in\mathcal{C}} \beta_C C$ for some $\beta_C \in \mathbb{N}$. Observe that $\beta_C = \vec{K}_C = \sum_{i=1}^{r} \alpha_i (\vec{F_i})_C$ for all $C \in \mathcal{C}$.

Let G and H be simple graphs such that $G \equiv_{\mathcal{F}} H$. If $\hom(F, G) = 0 = \hom(F, H)$ for some $F \in \mathcal{F}$, then $\hom(K, G) = 0 = \hom(K, H)$ by the assumption that $\hom(F, K) > 0$ for all $F \in \mathcal{F}$. Hence, it may be supposed that $\hom(F, G) = \hom(F, H) > 0$ for all $F \in \mathcal{F}$. This is crucial for ruling out division by zero in the following argument. Observe that, by Equation (2.1),

$$\hom(F, -) = \prod_{C\in\mathcal{C} \text{ s.t. } \vec{F}_C \geq 1} \hom(C, -)^{\vec{F}_C} \tag{6.5}$$

for all $F \in \mathcal{F} \cup \{K\}$. It follows that $\hom(C, G) > 0$ and $\hom(C, H) > 0$ for all $C \in \Gamma(\mathcal{F}) = \mathcal{C}$. Thus, in every step of the following calculation, the bases of all exponentials are positive integers.

$$\begin{aligned}
\hom(K, G) &\overset{(6.5)}{=} \prod_{C\in\mathcal{C}} \hom(C, G)^{\beta_C} = \prod_{C\in\mathcal{C}} \hom(C, G)^{\sum_{i=1}^{r} \alpha_i (\vec{F_i})_C} \\
&= \prod_{i=1}^{r} \prod_{C\in\mathcal{C}} \hom(C, G)^{\alpha_i (\vec{F_i})_C} \overset{(6.5)}{=} \prod_{i=1}^{r} \hom(F_i, G)^{\alpha_i} \\
&= \prod_{i=1}^{r} \hom(F_i, H)^{\alpha_i} = \hom(K, H).
\end{aligned}$$

Conversely, pick via Lemma 3.1.2 a finite family of graphs $\mathcal{G}$ such that the matrix $M := (\hom)|_{\mathcal{C}\times\mathcal{G}}$ is invertible. Consider the map

$$\begin{aligned}
\Psi \colon \mathbb{R}^{\mathcal{G}} &\to \mathbb{R}^{\mathcal{C}} \\
a &\mapsto \sum_{G\in\mathcal{G}} \hom(\mathcal{C}, G) a_G = Ma.
\end{aligned}$$

We think of $\mathbb{N}^{\mathcal{G}} \subseteq \mathbb{R}^{\mathcal{G}}$ as the space of instructions for constructing graphs as disjoint unions of elements in $\mathcal{G}$. The vector $a \in \mathbb{N}^{\mathcal{G}}$ corresponds to the graph $\coprod_{G\in\mathcal{G}} a_G G$. The map Ψ associates with such a graph its $\mathcal{C}$-homomorphism vector. In this way, $\mathbb{R}^{\mathcal{C}}$ may be thought of as the space of $\mathcal{C}$-homomorphism vectors. As a vector space isomorphism, Ψ is a homeomorphism, i.e. a continuous bijective map whose inverse is continuous.

Contrapositively, suppose that $\vec{K} \notin \operatorname{span}\{\vec{F} \in \mathbb{R}^{\mathcal{C}} \mid F \in \mathcal{F}\}$. Pick an integer vector $z \in \mathbb{Z}^{\mathcal{C}}$ such that $\langle z, \vec{F} \rangle = 0$ for all $F \in \mathcal{F}$ and $\langle z, \vec{K} \rangle \neq 0$. This can be done by Gram–Schmidt orthogonalisation applied to the rational vectors spanning $\operatorname{span}\{\vec{F} \in \mathbb{R}^{\mathcal{C}} \mid F \in \mathcal{F}\}$ and to the rational vector $\vec{K}$. The resulting vector z' with rational entries is orthogonal to all vectors $\vec{F}$ for $F \in \mathcal{F}$ and has non-zero inner-product with the vector $\vec{K}$. The integer vector z is then obtained from z' by clearing denominators.

Pick $p \in \Psi(\mathbb{Q}^{\mathcal{G}}_{>0}) \subseteq \mathbb{Q}^{\mathcal{C}}_{>0}$. In what follows, the vector p is perturbed in a direction depending on z. Technical complications arise when proving that the perturbed vector can again be interpreted as an instruction for constructing a graph, i.e. has a positive rational preimage under Ψ. To that end, consider the continuous function $\varphi\colon \mathbb{R}_{>0} \to \mathbb{R}^{\mathcal{C}}_{>0}$ which maps $t \mapsto (t^{z_C} p_C \mid C \in \mathcal{C})$.

Claim 6.5.6a. There exists a rational $t > 1$ such that $\varphi(t) \in \Psi(\mathbb{Q}^{\mathcal{G}}_{>0})$.

Proof of Claim. As the image of an open set under a homeomorphism, the set $\Psi(\mathbb{R}^{\mathcal{G}}_{>0})$ is open in $\mathbb{R}^{\mathcal{C}}$. Hence, there exists $\varepsilon > 0$ such that $B_\varepsilon(p) \subseteq \Psi(\mathbb{R}^{\mathcal{G}}_{>0})$. Here, $B_\varepsilon(p) \subseteq \mathbb{R}^{\mathcal{G}}$ denotes the radius-ε open ball around p. Since $\varphi^{-1}(B_\varepsilon(p))$ is open, $\varphi(1) = p$, and $\mathbb{Q}_{>0}$ is dense in $\mathbb{R}_{>0}$, there exists a rational $t > 1$ such that $\varphi(t) \in B_\varepsilon(p) \subseteq \Psi(\mathbb{R}^{\mathcal{G}}_{>0})$. Because t and p are rational and z is integral, $\varphi(t) \in \mathbb{Q}^{\mathcal{C}}$. The matrix M from definition of Ψ has integer entries and hence its preimages of rational vectors are rational. This implies that $\varphi(t) \in \Psi(\mathbb{Q}^{\mathcal{G}}_{>0})$. $\triangleleft$

Write $p' := \varphi(t)$ for $t > 1$, the rational whose existence is asserted by Claim 6.5.6a. Let $s, s' \in \mathbb{Q}^{\mathcal{G}}_{>0}$ denote the preimages of p and p' under Ψ, respectively. There exists a natural number $\lambda \geq 1$ such that λs and $\lambda s'$ lie in $\mathbb{N}^{\mathcal{G}}$. Write H and H' for the structures obtained by interpreting λs and $\lambda s'$ as instructions for disjoint unions over elements in $\mathcal{G}$, i.e. $H = \coprod_{G \in \mathcal{G}} (\lambda s)_G G$ and $H' = \coprod_{G \in \mathcal{G}} (\lambda s')_G G$. Then

$$\hom(F, H) \overset{(6.5)}{=} \prod_{C \in \mathcal{C}} \hom(C, H)^{\vec{F}_C} = \prod_{C \in \mathcal{C}} \Psi(\lambda s)_C^{\vec{F}_C} = \prod_{C \in \mathcal{C}} (\lambda \Psi(s)_C)^{\vec{F}_C} = \prod_{C \in \mathcal{C}} (\lambda p_C)^{\vec{F}_C}$$

and, similarly,

$$\hom(F, H') \overset{(6.5)}{=} \prod_{C \in \mathcal{C}} \Psi(\lambda s')_C^{\vec{F}_C} = \prod_{C \in \mathcal{C}} \left(\lambda \Psi(s')_C\right)^{\vec{F}_C} = \prod_{C \in \mathcal{C}} \left(\lambda p'_C\right)^{\vec{F}_C} = t^{\langle z, \vec{F} \rangle} \prod_{C \in \mathcal{C}} (\lambda p_C)^{\vec{F}_C}$$

for all simple graphs F which are disjoint unions of graphs in $\mathcal{C}$. In particular, $\hom(F, H) = \hom(F, H')$ for all $F \in \mathcal{F}$ but $\hom(K, H) \neq \hom(K, H')$. Thus $K \notin \operatorname{cl}(\mathcal{F})$. $\square$

Towards the proof of Theorem 6.5.2, we collect the following lemmas:

Lemma 6.5.7. *For every graph class $\mathcal{F}$, it holds that $\mathrm{cl}(\mathcal{F})_K \subseteq \mathrm{cl}(\mathcal{F}_K)$ for every simple graph K.*

Proof. Consider the following chain of inclusions:

$$\mathrm{cl}(\mathcal{F})_K = \bigcup_{L \in \mathrm{cl}(\mathcal{F})_K} \{L\} \subseteq \bigcup_{L \in \mathrm{cl}(\mathcal{F})_K} \mathrm{cl}(\mathcal{F}_L) \subseteq \mathrm{cl}\left(\bigcup_{L \in \mathrm{cl}(\mathcal{F})_K} \mathcal{F}_L \right) \subseteq \mathrm{cl}(\mathcal{F}_K).$$

The first inclusion follows from Lemma 6.5.5. Indeed, if $L \in \mathrm{cl}(\mathcal{F})$, then $L \in \mathrm{cl}(\mathcal{F}_L)$. The second inclusion is implied by Lemma 6.2.3. The third inclusion holds since, if $F \in \mathcal{F}_L$ for $L \in \mathrm{cl}(\mathcal{F})_K$, then $\hom(F, K) > 0$. □

For essentially profinite graph classes, Theorem 6.5.6 implies that the converse of Lemma 6.5.7 holds.

Lemma 6.5.8. *For every essentially profinite graph class $\mathcal{F}$, it holds that $\mathrm{cl}(\mathcal{F}_K) = \mathrm{cl}(\mathcal{F})_K$ for every simple graph K.*

Proof. That $\mathrm{cl}(\mathcal{F})_K \subseteq \mathrm{cl}(\mathcal{F}_K)$ is the assertion of Lemma 6.5.7. Conversely, $\mathrm{cl}(\mathcal{F}_K) \subseteq \mathrm{cl}(\mathcal{F})$ since $\mathcal{F}_K \subseteq \mathcal{F}$. It remains to argue that every $F \in \mathrm{cl}(\mathcal{F}_K)$ is K-colourable. If $F \in \mathrm{cl}(\mathcal{F}_K)$, then, by Theorem 6.5.6, $\Gamma(F) \subseteq \Gamma(\mathcal{F}_K)$. In other words, all connected components of F are K-colourable. This implies that F is K-colourable. Hence, $\mathrm{cl}(\mathcal{F}_K) \subseteq \mathrm{cl}(\mathcal{F})_K$. □

The concludes the preparations for the proof of Theorem 6.5.2.

Proof of Theorem 6.5.2. Suppose that $\mathcal{F}$ is homomorphism distinguishing closed. Let K be a simple graph such that $\Gamma(K) \subseteq \Gamma(\mathcal{F}_K)$. If $\vec{K} \in \mathrm{span}\{\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F}_K)} \mid F \in \mathcal{F}_K\}$, then Theorem 6.5.6 applies. Hence, $K \in \mathrm{cl}(\mathcal{F})$ and thus $K \in \mathcal{F}$ since $\mathcal{F}$ is homomorphism distinguishing closed.

Assuming Item 2, let $K \notin \mathcal{F}$ be a simple graph. Then, $\vec{K} \notin \mathrm{span}\{\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F}_K \cup \{K\})} \mid F \in \mathcal{F}_K\}$. By Theorem 6.5.6, $K \notin \mathrm{cl}(\mathcal{F}_K)$. By Lemma 6.5.5, $K \notin \mathrm{cl}(\mathcal{F})$. Hence, $\mathcal{F}$ is homomorphism distinguishing closed.

The equivalence of Items 1 and 3 follows from Lemma 6.5.8. Indeed, if $\mathcal{F}$ is homomorphism distinguishing closed, then $\mathrm{cl}(\mathcal{F}_K) = \mathrm{cl}(\mathcal{F})_K = \mathcal{F}_K$ for every simple graph K. Conversely,

$$\mathrm{cl}(\mathcal{F}) = \bigcup_K \mathrm{cl}(\mathcal{F})_K = \bigcup_K \mathrm{cl}(\mathcal{F}_K) = \bigcup_K \mathcal{F}_K = \mathcal{F}.$$

Here, K ranges over all simple graphs. □

Finally, we deduce the following Corollaries 6.5.9 and 6.5.10 from Theorem 6.5.6:

Corollary 6.5.9. *If a graph class $\mathcal{F}$ is essentially finite, then $\mathrm{cl}(\mathcal{F})$ is essentially finite.*

Proof. Let $K \in \mathrm{cl}(\mathcal{F})$. By Lemma 6.5.5, $K \in \mathrm{cl}(\mathcal{F}_K)$. By Theorem 6.5.6, $\vec{K} \in \mathrm{span}\{\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F}_K \cup \{K\})} \mid F \in \mathcal{F}_K\}$. In particular, $\Gamma(K) \subseteq \Gamma(\mathcal{F}_K)$. Hence, $\Gamma(\mathrm{cl}(\mathcal{F})) = \bigcup_{K \in \mathrm{cl}(\mathcal{F})} \Gamma(K) \subseteq \bigcup_{K \in \mathrm{cl}(\mathcal{F})} \Gamma(\mathcal{F}_K) \subseteq \Gamma(\mathcal{F})$. Thus, $\mathrm{cl}(\mathcal{F})$ is essentially finite. □

Since the class of all graphs is not essentially profinite, the following Corollary 6.5.10 implies that no homomorphism indistinguishability relation of an essentially profinite graph class is as fine as isomorphism.

Corollary 6.5.10. *If a graph class $\mathcal{F}$ is essentially profinite, then* $\mathrm{cl}(\mathcal{F})$ *is essentially profinite.*

Proof. It has to be argued that, for every simple graph K, the class $\mathrm{cl}(\mathcal{F})_K$ is essentially finite. By Lemma 6.5.7, $\mathrm{cl}(\mathcal{F})_K \subseteq \mathrm{cl}(\mathcal{F}_K)$ and the right hand-side is an essentially finite graph class by Corollary 6.5.9. □

In order to illustrate the internal functioning of condition Item 2 in Theorem 6.5.2, we consider the following examples. The first example shows that not all homomorphism distinguishing closed graph classes are closed under taking summands, cf. Figure 7.1. The second example answers a question from [150, p. 29] negatively: Is the disjoint union closure of the union of homomorphism distinguishing closed families homomorphism distinguishing closed? Finally, the second and third example illustrate that the inclusions in Lemma 6.2.3 can be proper.

Example 6.5.11. Let F_1 and F_2 be connected non-isomorphic homomorphically equivalent simple graphs.

1. The graph class $\mathcal{F}_1 := \{n(F_1 + F_2) \mid n \in \mathbb{N}\}$ is homomorphism distinguishing closed and not closed under taking summands.
2. For the homomorphism distinguishing closed graph class $\mathcal{F}_2 := \{nF_1 \mid n \in \mathbb{N}\}$, the disjoint union closure of $\mathcal{F}_1 \cup \mathcal{F}_2$ is not homomorphism distinguishing closed.
3. Let $\mathcal{F}_3 := \{n_1F_1 + n_2F_2 \mid n_1, n_2 \in \mathbb{N}, n_1 \geq n_2\}$ and $\mathcal{F}_4 := \{n_1F_1 + n_2F_2 \mid n_1, n_2 \in \mathbb{N}, n_2 \geq n_1\}$. Then $\mathrm{cl}(\mathcal{F}_3 \cap \mathcal{F}_4) \subsetneq \mathrm{cl}(\mathcal{F}_3) \cap \mathrm{cl}(\mathcal{F}_4)$.

Proof. For the first example, observe that clearly $\Gamma(\mathcal{F}_1) = \{F_1, F_2\}$. To verify the assumptions of Theorem 6.5.2, let $K := n_1F_1 + n_2F_2$ for $n_1, n_2 \geq 0$. It holds that $\mathrm{span}\{\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F})} \mid F \in \mathcal{F}\} = \left\{\lambda \binom{1}{1} \mid \lambda \in \mathbb{R}\right\}$ and thus $\vec{K}$ is in this space only if $n_1 = n_2$. In this case, however, $K \in \mathcal{F}$ and thus $\mathcal{F}$ is homomorphism distinguishing closed. It remains to observe that, for example, $F_1 \notin \mathcal{F}$.

In the second example, by Theorem 6.5.2, $\mathcal{F}_2$ is homomorphism distinguishing closed. The disjoint union closure of $\mathcal{F}_1 \cup \mathcal{F}_2$ is $\mathcal{F} := \{n_1F_1 + n_2F_2 \mid n_1, n_2 \in \mathbb{N}, n_1 \geq n_2\}$. Since $\hom(F_1, F_2) > 0$, it holds that $\mathcal{F}_{F_2} = \mathcal{F}$ and hence $\vec{F_2} \in \mathrm{span}\{\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F})} \mid \vec{F} \in \mathcal{F}_{F_2}\} = \mathbb{R}^{\Gamma(\mathcal{F})}$. However, $F_2 \notin \mathcal{F}$. By Theorem 6.5.2, $\mathcal{F}$ is not homomorphism distinguishing closed.

In the third example, $\mathrm{cl}(\mathcal{F}_3) = \mathrm{cl}(\mathcal{F}_4) = \{n_1 F_1 + n_2 F_2 \mid n_1, n_2 \in \mathbb{N}\}$, by Theorem 6.5.2. However, $\mathcal{F}_3 \cap \mathcal{F}_4 = \{n(F_1 + F_2) \mid n \in \mathbb{N}\}$, which is homomorphism distinguishing closed by the first example. □

In light of Conjecture 6.0.3, minor-closed graph classes are of special interest. The following Lemma 6.5.12 shows that considering essentially profinite rather than essentially finite graph classes, in a sense, does not add any value for minor-closed graph classes.

Lemma 6.5.12. *Every essentially profinite minor-closed graph class $\mathcal{F}$ is essentially finite.*

Proof. Since the class of all graphs is not essentially profinite, there exists a number $t \in \mathbb{N}$ such that all $F \in \mathcal{F}$ do not contain the complete graph K_t as a minor. By a weak form of Hadwiger's conjecture [167] proven by Wagner [173], this implies that all $F \in \mathcal{F}$ are 2^{t-1}-colourable. In particular, $\mathcal{F} = \mathcal{F}_{K_{2^{t-1}}}$ is essentially finite. □

As final observation, we remark that Theorems 6.5.2 and 6.3.9 are orthogonal in the sense that there exist graph classes to which only one of the theorems applies. Clearly, there are graph classes which are closed under weak oddomorphism but which are not essentially profinite. One such example is the class of graphs of bounded degree from Corollary 6.3.10. Conversely, the graph class $\{nP_4 \mid n \in \mathbb{N}\}$ is essentially finite and homomorphism distinguishing closed by Corollary 6.5.3 but not closed under weak oddomorphisms. Indeed, the homomorphism $P_4 \to K_2$ is an oddomorphism but K_2 is not in the class.

6.6 Further Directions

The central open problem regarding the homomorphism distinguishing closure is Conjecture 6.0.2. Besides this conjecture, we propose to consider the relative homomorphism distinguishing closure, as defined below.

Definition 6.6.1. Let $\mathcal{F}$ and $\mathcal{H}$ denote classes of simple graphs. Define the *homomorphism distinguishing closure of $\mathcal{F}$ relative to $\mathcal{H}$* as

$$\mathrm{cl}_{\mathcal{H}}(\mathcal{F}) := \{K \mid \forall G, H \in \mathcal{H}.\ G \equiv_{\mathcal{F}} H \implies \hom(K, G) = \hom(K, H)\}.$$

Here, K ranges over all simple graphs.

Few of the techniques developed in this chapter for studying the homomorphism distinguishing closure seem to be applicable to the relative homomorphism distinguishing closure. For example, CFI graphs, which underpin Theorem 6.3.9, are typically rather complicated graphs, even when constructed for rather simple base graphs. Thus, studying the relative homomorphism distinguishing closure requires developing new techniques which might aid addressing Conjecture 6.0.2.

To illustrate the intricacies of Definition 6.6.1, first observe that typically $\mathrm{cl}_{\mathcal{H}}(\mathcal{H})$ is the class of all graphs $\mathcal{G}$. By [14, Theorem 3], which is based on Lemma 3.1.2, this holds for every graph class $\mathcal{H}$ such that $K \in \mathcal{H}$ for every graph K admitting a surjective homomorphisms $H_1 \to K$ and an injective homomorphisms $K \to H_2$ for some $H_1, H_2 \in \mathcal{H}$. All graph classes closed under taking subgraphs have this property.

The relative homomorphism distinguishing closure is related to the Weisfeiler–Leman dimension of a graph class [78, Definition 18.4.2]: The *Weisfeiler–Leman dimension* of a graph class $\mathcal{H}$ is the least $k \in \mathbb{N}$ such that, for all $H \in \mathcal{H}$ and an arbitrary graph G, it holds that if G and H are not distinguished by the k-dimensional Weisfeiler–Leman algorithm, then $G \cong H$. By Theorem 3.4.3, if $\mathcal{H}$ has Weisfeiler–Leman dimension at most k, then $\mathrm{cl}_{\mathcal{H}}(\mathcal{TW}_k)$ is the class of all graphs. Grohe [78, Corollary 18.4.1] showed that every graph class excluding some minor has bounded Weisfeiler–Leman dimension. See [100] for further background on the Weisfeiler–Leman dimension. We conclude with the following concrete questions, cf. [161].

Question 6.6.2. *What is the homomorphism distinguishing closure of the class of cycles relative to the class of trees?*

Question 6.6.3. *Can Colour Refinement on graphs of bounded degree be characterised as homomorphism indistinguishability over trees of bounded degree? In other words, what is the homomorphism distinguishing closure of the class of trees of degree at most d relative to the class of graphs of degree at most d'?*

7 Syntax and Semantics of Homomorphism Indistinguishability

In Chapters 3 to 5, several graph isomorphism relaxations from diverse fields such as logic or optimisation were characterised as homomorphism indistinguishability relations. Notably, most of these characterisations involve minor-closed graph classes. In this chapter, we show that this is not a mere coincide. More precisely, we show the following theorem:

Theorem 7.0.1. *For a graph class $\mathcal{F}$ and the assertions*

1. *$\mathcal{F}$ is closed under edge contraction and deletion,*
2. *$\equiv_{\mathcal{F}}$ is* preserved under taking complements, *i.e., for all simple graphs G and H, it holds that $G \equiv_{\mathcal{F}} H$ if, and only if, $\overline{G} \equiv_{\mathcal{F}} \overline{H}$,*
3. *$\mathrm{cl}(\mathcal{F})$ is minor-closed,*

the implications $1 \Rightarrow 2 \Leftrightarrow 3$ hold.

In particular, Theorem 7.0.1 implies that if a graph isomorphism relaxation $\approx$ is a homomorphism indistinguishability relation and preserved under taking complements, then it is a homomorphism indistinguishability relation over a minor-closed graph class. This necessary condition can be used to rule out that a graph isomorphism relaxation which is preserved under taking complements has a homomorphism indistinguishability characterisations, cf. Corollary 7.1.17 and Theorem 7.2.6.

Graph isomorphism relaxations from logic are typically preserved under taking complements. Thereby, Theorem 7.0.1 implies that such logical equivalences are typically homomorphism indistinguishability relations over minor-closed graph classes if there are homomorphism indistinguishability relations at all. In Section 7.2, we formalise this statement and give examples.

Moreover, Theorem 7.0.1 gives evidence for Conjecture 6.0.2, indicating that minor-closed graph classes play a distinct role in homomorphism indistinguishability.[9] Indeed, while Conjecture 6.0.2 asserts that $\mathrm{cl}(\mathcal{F}) = \mathcal{F}$ for every minor-closed union-closed graph class $\mathcal{F}$, Theorem 7.0.1 implies unconditionally that $\mathrm{cl}(\mathcal{F})$ is a minor-closed union-closed graph class.

[9] A priori it is not clear whether minors have anything to do with homomorphism indistinguishability. Roberson [150, p. 4] remarks that 'our main goal is to understand what types of families are homomorphism distinguishing closed, or to identify some nice family of families of graphs that all give rise to distinct (and preferably nice) homomorphism indistinguishability relations. So "minor-closed" may need to be replaced with some other property in order for the conjecture [Conjecture 6.0.2] to hold.'

Closure property of $\mathcal{F}$	**Preservation property of $\equiv_{\mathcal{F}}$**	**Theorem**
taking minors	complements	Theorem 7.0.1
taking summands	disjoint unions	Theorem 7.1.3
taking subgraphs	full complements	Theorem 7.1.10
taking induced subgraphs	left lexicographic products	Theorem 7.1.11
contracting edges	right lexicographic products	Theorem 7.1.12

Table 7.1: Equivalent properties of a homomorphism distinguishing closed graph class $\mathcal{F}$ and of its homomorphism indistinguishability relation $\equiv_{\mathcal{F}}$.

We think of Theorem 7.0.1 as a step towards understanding the nature of homomorphism indistinguishability relations. Ultimately, we would like to establish sufficient and necessary criteria for a graph isomorphism relaxation to be a homomorphism indistinguishability relation. Theorem 7.0.1 gives such a necessary criterion for homomorphism indistinguishability over minor-closed graph classes. Throughout Section 7.1, we show further results paralleling Theorem 7.0.1 but involving other pairs of closure properties of $\mathcal{F}$ and preservation properties of $\equiv_{\mathcal{F}}$. These results are summarised in Table 7.1.

In Section 7.3, we consider cancellation properties of homomorphism indistinguishability relations. A graph isomorphism relaxation $\approx$ *admits K-cancellation* for some simple graph K if $G \times K \approx H \times K$ implies $G \approx H$ for all simple graphs G and H. For graph classes $\mathcal{F}$ closed under subdivision, we show that $\equiv_{\mathcal{F}}$ admits K-cancellation if, and only if, K is non-bipartite, cf. Lemma 7.3.4 and Theorem 7.3.5. This gives a full picture of the cancellation admitted by the quantum isomorphism relation or counting logic equivalences, cf. [124] and Theorem 3.4.4.

Chapter Outline. In Section 7.1, we prove Theorem 7.0.1 and the other results listed in Table 7.1. In Section 7.2, we discuss repercussions for homomorphism indistinguishability characterisations of logical equivalences. Section 7.3 is concerned with cancellation properties of homomorphism indistinguishability relations. We conclude in Section 6.6 by proposing further directions. The material in this section has been previously published in [162, 165].

7.1 Closure Properties Correspond to Preservation Properties

This section is concerned with the interplay of closure properties of a graph class $\mathcal{F}$ and preservation properties of its homomorphism indistinguishability relation $\equiv_{\mathcal{F}}$. The central results of this section are those listed in Table 7.1. The relevance of the results in Table 7.1 is twofold: On the one hand, they yield that if a graph class $\mathcal{F}$ has a certain closure property, then so does $\mathrm{cl}(\mathcal{F})$. In the case of minor-closed graph families, this provides evidence for Conjecture 6.0.2. On the other hand, they establish that every graph isomorphism relaxation which is preserved under certain operations coincides with the homomorphism indistinguishability relation over a graph class with a certain closure property, if it is a homomorphism indistinguishability relation at all. Further consequences are discussed in Sections 7.1.6 and 7.2.

The closure properties of graph classes in Table 7.1 should be self-explanatory. To build an intuition of what preservation properties are, we consider the following two preservation properties which are shared by all homomorphism indistinguishability relations:

Lemma 7.1.1. *For every graph class $\mathcal{F}$, the following hold:*

1. $\equiv_{\mathcal{F}}$ *is* preserved under categorical products, *i.e., for all simple graphs G, H, and K, if $G \equiv_{\mathcal{F}} H$, then $G \times K \equiv_{\mathcal{F}} H \times K$.*
2. $\equiv_{\mathcal{F}}$ *is* preserved under blow-ups, *i.e., for all simple graphs G and H and $n \geq 1$, if $G \equiv_{\mathcal{F}} H$, then $G \cdot \overline{K_n} \equiv H \cdot \overline{K_n}$.*

Proof. For the first claim, by Equation (2.2), for every $F \in \mathcal{F}$,

$$\begin{aligned}\hom(F, G \times K) &= \hom(F, G) \hom(F, K) \\ &= \hom(F, H) \hom(F, K) = \hom(F, H \times K).\end{aligned}$$

For the second claim, observe that

$$\hom(F, G \cdot \overline{K_n}) = n^{|V(F)|} \hom(F, G). \tag{7.1}$$

With this identity, the second claim follows analogously to the first claim. □

Essential to all proofs in this section is the following Lemma 7.1.2:

Lemma 7.1.2. *Let $\mathcal{F}$ and $\mathcal{L}$ be classes of simple graphs. Suppose $\mathcal{L}$ is finite and that its elements are pairwise non-isomorphic. Let $\alpha \colon \mathcal{L} \to \mathbb{R} \setminus \{0\}$. If, for all simple graphs G and H,*

$$G \equiv_{\mathcal{F}} H \implies \sum_{L \in \mathcal{L}} \alpha_L \hom(L, G) = \sum_{L \in \mathcal{L}} \alpha_L \hom(L, H),$$

then $\mathcal{L} \subseteq \mathrm{cl}(\mathcal{F})$.

Proof. The following argument is due to [49, Lemma 3.6]. Let n be an upper bound on the number of vertices of graphs in $\mathcal{L}$ and let $\mathcal{L}'$ denote a collection of non-isomorphic simple graphs on at most n vertices containing as in Lemma 3.1.2. By Lemma 3.1.2, the matrix $M := (\hom(K, L))_{K,L \in \mathcal{L}'}$ is invertible. Extend α to a function $\alpha' \colon \mathcal{L}' \to \mathbb{R}$ by setting $\alpha'(L) := \alpha(L)$ for all $L \in \mathcal{L}$ and $\alpha'(L') := 0$ for all $L' \in \mathcal{L}' \setminus \mathcal{L}$. By Lemma 7.1.1, if $G \equiv_{\mathcal{F}} H$, then $G \times K \equiv_{\mathcal{F}} H \times K$ for all simple graphs K. Hence, by Equation (2.2),

$$\sum_{L \in \mathcal{L}'} \hom(L, K) \alpha'_L \hom(L, G) = \sum_{L \in \mathcal{L}'} \hom(L, K) \alpha'_L \hom(L, H).$$

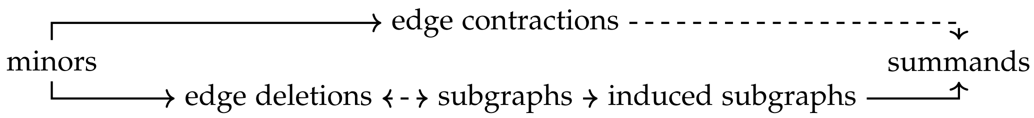


Figure 7.1: Relationships between closure properties of homomorphism distinguishing closed graph classes. The non-obvious implications are dashed and proven in Lemmas 7.1.9 and 7.1.15. No other implications hold in general, cf. Example 7.1.16.

Both sides can be read as the product of the matrix M^T with a vector of the form $(\alpha'_L \hom(L, -))_{L \in \mathcal{L}'}$. By multiplying from the left with the inverse of M^T, it follows that $\alpha'_L \hom(L, G) = \alpha'_L \hom(L, H)$ for all $L \in \mathcal{L}'$, which in turn implies that $\hom(L, G) = \hom(L, H)$ for all $L \in \mathcal{L}$. Thus, $\mathcal{L} \subseteq \mathrm{cl}(\mathcal{F})$. □

In the setting of Lemma 7.1.2, we say that the relation $\equiv_{\mathcal{F}}$ *determines* the linear combination $\sum_{L \in \mathcal{L}} \alpha_L \hom(L, -)$. Note that it is essential for the argument to carry through that the elements of $\mathcal{L}$ are pairwise non-isomorphic and that $\alpha_L \neq 0$ for all L. Efforts will be undertaken to establish this property for certain linear combinations in the subsequent sections.

7.1.1 Taking Summands and Preservation under Disjoint Unions

In this section, the strategy yielding the results in Table 7.1 is presented for the rather simple case of Theorem 7.1.3. This theorem relates the property of a graph class $\mathcal{F}$ to be closed under taking summands to the property of $\equiv_{\mathcal{F}}$ to be preserved under disjoint unions. This closure property is often assumed in the context of homomorphism indistinguishability, cf. Theorem 6.3.9 and [3], and fairly mild. It is the most general property among those studied here, cf. Figure 7.1. Theorem 7.1.3 answers a question from [150, p. 7] affirmatively: Is it true that if $\equiv_{\mathcal{F}}$ is preserved under disjoint unions, then $\mathrm{cl}(\mathcal{F})$ is closed under taking summands?

Theorem 7.1.3. *For a graph class $\mathcal{F}$ and the assertions*

1. $\mathcal{F}$ *is* closed under taking summands, *i.e., if $F_1 + F_2 \in \mathcal{F}$, then $F_1, F_2 \in \mathcal{F}$,*
2. $\equiv_{\mathcal{F}}$ *is* preserved under disjoint unions, *i.e., for all simple graphs G, G', H, and H', if $G \equiv_{\mathcal{F}} G'$ and $H \equiv_{\mathcal{F}} H'$, then $G + H \equiv_{\mathcal{F}} G' + H'$,*
3. $\mathrm{cl}(\mathcal{F})$ *is closed under taking summands,*

the implications 1 ⇒ 2 ⇔ 3 hold.

Proof. The central idea is to write, given simple graphs F, G, and H, the quantity $\hom(F, G + H)$ as expression in $\hom(F', G)$ and $\hom(F', H)$ where F' ranges over summands of F. To this end, write $F = C_1 + \cdots + C_r$ as disjoint union of its connected components. Then,

$$\begin{aligned}\hom(F, G+H) &\overset{(2.1)}{=} \prod_{i=1}^{r} \hom(C_i, G+H) \\ &\overset{(2.3)}{=} \prod_{i=1}^{r} (\hom(C_i, G) + \hom(C_i, H)) \\ &\overset{(2.1)}{=} \sum_{I \subseteq [r]} \hom(\coprod_{i \in I} C_i, G) \hom(\coprod_{i \in [r] \setminus I} C_i, H). \end{aligned} \tag{7.2}$$

In particular, if $\mathcal{F}$ is closed under taking summands, then $\coprod_{i\in I} C_i \in \mathcal{F}$ for all $I \subseteq [r]$. Thus, 1 implies 2.

Assume 2 and let $F \in \mathrm{cl}(\mathcal{F})$. Write as above $F = C_1 + \cdots + C_r$ as disjoint union of its connected components. By the assumption that $\equiv_{\mathcal{F}}$ is preserved under disjoint unions, for all simple graphs G and G', if $G \equiv_{\mathcal{F}} G'$, then $G + F \equiv_{\mathcal{F}} G' + F$ and hence $\hom(F, G+F) = \hom(F, G'+F)$. By Equation (7.2) with $H = F$, the relation $\equiv_{\mathcal{F}}$ determines the linear combination $\sum_{I\subseteq[r]} \hom(\coprod_{i\in I} C_i, -) \hom(\coprod_{i\in[r]\setminus I} C_i, F)$. Note that it might be the case that $\coprod_{i\in I} C_i \cong \coprod_{j \in J} C_j$ for some $I \neq J$. Grouping such summands together and adding their coefficients yields a linear combination satisfying the assumptions of Lemma 7.1.2 since $\hom(C_i, F) > 0$ for all $i \in [r]$. Hence, $\coprod_{i\in I} C_i \in \mathrm{cl}(\mathcal{F})$ for all $I \subseteq [r]$ and 3 follows.

The implication 3 ⇒ 2 follows from 1 ⇒ 2 for $\mathrm{cl}(\mathcal{F})$ since $\equiv_{\mathcal{F}}$ and $\equiv_{\mathrm{cl}(\mathcal{F})}$ coincide on simple graphs. □

The proofs of the other results in Table 7.1 are conceptually similar to the just completed proof. The general idea can be briefly described as follows:

1. Derive a linear expression similar to Equation (7.2) for the number of homomorphisms from F into the graph constructed using the assumed preservation property of $\equiv_{\mathcal{F}}$, e.g. the graph $G + H$ in the case of Theorem 7.1.3. These linear combinations typically involve sums over subsets U of vertices or edges of F, each contributing a summand of the form $\alpha_U \hom(F_U, -)$ where α_U is some coefficient and F_U is a graph constructed from F using U. Hence, if the graph class $\mathcal{F}$ is closed under the construction transforming F to F_U, then $\equiv_{\mathcal{F}}$ has the desired preservation property.
2. In general, it can be that F_U and $F_{U'}$ are isomorphic despite that $U \neq U'$, e.g., in Equation (7.2), if F contains two isomorphic connected components. In order to apply Lemma 7.1.2, one must group the summands $\alpha_U \hom(F_U, -)$ by the isomorphism type F' of the F_U. The coefficient of $\hom(F', -)$ in the new linear combination ranging over pairwise non-isomorphic graphs is the sum of α_U over all U such that $F_U \cong F'$. Once it is established that this coefficient is non-zero, it follows that if $\equiv_{\mathcal{F}}$ has the preservation property, then $\mathrm{cl}(\mathcal{F})$ has the desired closure property.

(a) G. **(b)** H.

Figure 7.2: The cospectral graphs from the proof of Theorem 7.1.4.

7.1.2 Cycles are Homomorphism Distinguishing Closed

In this section, we divert from this section's overall goal of establishing the results in Table 7.1 by proving, as an application of Theorem 7.1.3, that the class of disjoint unions of cycles is homomorphism distinguishing closed. Homomorphism indistinguishability of two graphs over disjoint unions of cycles is characterised by Theorem 3.3.2 as cospectrality of their adjacency matrices.

Theorem 7.1.4. *The class of disjoint unions of cycles is homomorphism distinguishing closed.*

In order to show that a graph class $\mathcal{F}$ is homomorphism distinguishing closed, one must construct, by Definition 6.0.1, for every simple graph $F \notin \mathcal{F}$, two simple graphs G and H such that $G \equiv_{\mathcal{F}} H$ and $\hom(F, G) \neq \hom(F, H)$. The following corollary of Theorem 7.1.3 asserts that this must be done only for *connected* graphs $F \notin \mathcal{F}$ if $\mathcal{F}$ is closed under disjoint unions and taking summands. Thereby, Corollary 7.1.5 strengthens Theorem 6.3.9.

Corollary 7.1.5. *Let $\mathcal{F}$ be a graph class such that*

1. *for every connected simple graph $F \notin \mathcal{F}$, there exist simple graphs G and H such that $G \equiv_{\mathcal{F}} H$ and* $\hom(F, G) \neq \hom(F, H)$ *and*
2. *$\mathcal{F}$ is closed under disjoint unions and taking summands, i.e. $F_1 + F_2 \in \mathcal{F}$ if, and only if, $F_1, F_2 \in \mathcal{F}$ for all graphs F_1, F_2.*

Then $\mathcal{F}$ is homomorphism distinguishing closed.

Proof. Let $F \notin \mathcal{F}$. Since $\mathcal{F}$ is closed under disjoint unions, there exists a connected component F' of F such that $F' \notin \mathcal{F}$. By assumption, $F' \notin \mathrm{cl}(\mathcal{F})$. By Theorem 7.1.3, if $\mathcal{F}$ is closed taking summands, then $\mathrm{cl}(\mathcal{F})$ is closed under taking summands. Hence, $F \notin \mathrm{cl}(\mathcal{F})$. □

By Corollary 6.3.10, the class of graphs of degree at most two is homomorphism distinguishing closed. By Theorem 3.3.2 and Corollary 7.1.5, in order to prove Theorem 7.1.4, it suffices to exhibit two cospectral graphs which admit different numbers of homomorphisms from every path graph. These graphs are depicted by Figure 7.2.

Proof of Theorem 7.1.4. By Theorem 3.3.2 and Corollary 7.1.5, for every connected simple graph F which is not a cycle, we must construct two simple graphs G and H whose adjacency matrices are cospectral but $\hom(F, G) \neq \hom(F, H)$. If F contains a vertex of degree at least three, then Corollary 6.3.10 and Theorem 6.3.8 yield such graphs. Hence, it remains to consider the case when F is a path. Consider the graphs G and H depicted by Figure 7.2. By [50, (2b)], $G = W_2$ and $H = C_4 + P_2$ are cospectral. Hence, they are homomorphism indistinguishable over all cycles by Theorem 3.3.2. We argue that G and H admit a different number of homomorphisms from every path on at least three vertices.

An inductive argument yields for $i \geq 0$ that

$$\boldsymbol{A}_H^{2i+1} = \begin{pmatrix} 0 & 2^{2i} & 0 & 2^{2i} & 0 & 0 \\ 2^{2i} & 0 & 2^{2i} & 0 & 0 & 0 \\ 0 & 2^{2i} & 0 & 2^{2i} & 0 & 0 \\ 2^{2i} & 0 & 2^{2i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{A}_H^{2i+2} = \begin{pmatrix} 2^{2i+1} & 0 & 2^{2i+1} & 0 & 0 & 0 \\ 0 & 2^{2i+1} & 0 & 2^{2i+1} & 0 & 0 \\ 2^{2i+1} & 0 & 2^{2i+1} & 0 & 0 & 0 \\ 0 & 2^{2i+1} & 0 & 2^{2i+1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Hence, the number of homomorphism from the path on $i + 1$ vertices to H is $\operatorname{soe}(\boldsymbol{A}_H^i) = 2^{i+2} + 2$. The first values of this sequence for $i \geq 2$ are 18, 34, 66, and 130.

For the graph G, one may compute that the first values of the sequence $\operatorname{soe}(\boldsymbol{A}_G^i)$ for $i \geq 2$ as 22, 42, 86, and 170. Furthermore, when regarded as matrices modulo 16,

$$\boldsymbol{A}_G^5 \equiv \begin{pmatrix} 0 & 0 & 11 & 0 & 5 & 5 \\ 0 & 0 & 11 & 0 & 5 & 5 \\ 11 & 11 & 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 & 11 & 11 \\ 5 & 5 & 0 & 11 & 0 & 0 \\ 5 & 5 & 0 & 11 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{A}_G^6 \equiv \begin{pmatrix} 11 & 11 & 0 & 5 & 0 & 0 \\ 11 & 11 & 0 & 5 & 0 & 0 \\ 0 & 0 & 11 & 0 & 5 & 5 \\ 5 & 5 & 0 & 11 & 0 & 0 \\ 0 & 0 & 5 & 0 & 11 & 11 \\ 0 & 0 & 5 & 0 & 11 & 11 \end{pmatrix}.$$

It follows that $\boldsymbol{A}_G^{5+2i} \equiv \boldsymbol{A}_G^5 \mod 16$ and $\boldsymbol{A}_G^{6+2i} \equiv \boldsymbol{A}_G^6 \mod 16$ for all $i \geq 0$. Finally, $\operatorname{soe}(\boldsymbol{A}_G^5) \equiv 10 \mod 16$ and $\operatorname{soe}(\boldsymbol{A}_G^6) \equiv 6 \mod 16$. Thus, $\operatorname{soe}(\boldsymbol{A}_G^i) - 2$ is not divisible by 16 for any $i \geq 5$. In particular, $\operatorname{soe}(\boldsymbol{A}_G^i) \neq \operatorname{soe}(\boldsymbol{A}_H^i)$ for all $i \geq 2$. Thus, the graphs G and H admit different numbers of homomorphisms from every path on at least three vertices. □

7.1.3 Taking Minors and Preservation under Complements

The strategy outlined in Section 7.1.1 is now applied to prove Theorem 7.0.1. This answers a question of Roberson [150, Question 8] affirmatively: Is it true that if $\mathcal{F}$ is such that $\equiv_{\mathcal{F}}$ is preserved under taking complements, then there exist a minor-closed graph class $\mathcal{F}'$ such that $\equiv_{\mathcal{F}}$ and $\equiv_{\mathcal{F}'}$ coincide? In fact, it is shown that $\mathcal{F}'$ can be taken to be $\operatorname{cl}(\mathcal{F})$.

Theorem 7.0.1. *For a graph class $\mathcal{F}$ and the assertions*

1. *$\mathcal{F}$ is closed under edge contraction and deletion,*
2. *$\equiv_{\mathcal{F}}$ is* preserved under taking complements, *i.e., for all simple graphs G and H, it holds that $G \equiv_{\mathcal{F}} H$ if, and only if, $\overline{G} \equiv_{\mathcal{F}} \overline{H}$,*
3. *$\mathrm{cl}(\mathcal{F})$ is minor-closed,*

the implications 1 ⇒ 2 ⇔ 3 hold.

The strategy is to write $\hom(F, \overline{G})$ as a linear combination of $\hom(F', G)$ for minors F' of F. This is accomplished in two steps: First, we consider the full complement $\widehat{G}$ of G in which not only every edge of G is replaced by a non-edge (and vice versa) but also every loop is replaced by a non-loop (and vice versa). Secondly, we consider, for a simple graph G, the looped graph G° obtained from G by adding a loop to every vertex. The fact that $\widehat{G^\circ} \cong \overline{G}$ motivates this two-step approach.

Lemma 7.1.6 ([116, Equation (5.23)])**.** *For every graph G and every simple graph F,*

$$\hom(F, \widehat{G}) = \sum_{F' \subseteq F \text{ s.t. } V(F') = V(F)} (-1)^{|E(F')|} \hom(F', G).$$

Proof. Write S for the set of all maps $V(F) \to V(G)$. For $e \in E(F)$, write $A_e \subseteq S$ for set of all maps h such that the image of e under h is an edge or a loop in G. It holds that $h \notin A_e$ if, and only if, e is mapped under h either to two non-adjacent vertices or to a single vertex without a loop. Hence, $\hom(F, \widehat{G}) = \left|\bigcap_{e \in E(F)} \overline{A_e}\right|$ where $\overline{A_e} := S \setminus A_e$. By the Inclusion–Exclusion Principle,

$$\left|\bigcap_{e \in E(F)} \overline{A_e}\right| = |S| + \sum_{\emptyset \neq E' \subseteq E(F)} (-1)^{|E'|} \left|\bigcap_{e \in E'} A_e\right|.$$

Finally, note that $|S| = \hom(|V(F)|K_1, G)$ and $\left|\bigcap_{e \in E(F')} A_e\right|$ for some $F' \subseteq F$ is the number of homomorphisms from the graph F' to G which is obtained by deleting all edges from F which are not in F'. □

In light of Lemma 7.1.6, it suffices to write $\hom(F, G^\circ)$ as a linear combination of $\hom(F', G)$ for minors F' of F. To ease bookkeeping, we consider a particular type of quotient graphs. For a simple graph F and a set of edges $L \subseteq E(F)$, define the *contraction relation* $\sim_L$ on $V(F)$ by declaring $v \sim_L w$ if v and w lie in the same connected component of the subgraph of F with vertex set $V(F)$ and edge set L. Write $[v]_L$ for the classes of $v \in V(F)$ under the equivalence relation $\sim_L$.

The *contraction quotient* $F \oslash L$ is the graph whose vertex set is the set of equivalence classes under $\sim_L$ and with an edge between $[v]_L$ and $[w]_L$ if, and only if, there is an edge $xy \in E(F) \setminus L$ such that $x \sim_L v$ and $y \sim_L w$. In general, $F \oslash L$ may contain loops, cf. Example 7.1.8. However, if it is a simple graph, then it is equal to $F/\mathcal{P}$ where $\mathcal{P}$ is the partition of $V(F)$ into equivalence classes under $\sim_L$, i.e. $\mathcal{P} := \{[v]_L \mid v \in V(F)\}$. In this case, $F \oslash L$ is a graph obtained from F by edge contractions. With this notation, the quantity $\hom(F, G^\circ)$ can be succinctly written as a linear combination.

Theorem 7.1.7. *Let F and G be simple graphs. Then*

$$\hom(F, G^\circ) = \sum_{L \subseteq E(F)} \hom(F \oslash L, G).$$

Proof. We establish a bijection between the set $\mathrm{Hom}(F, G^\circ)$ of homomorphisms $h \colon F \to G^\circ$ and the set X of pairs (L, φ) where $L \subseteq E(F)$ and $\varphi \colon F \oslash L \to G$ is a homomorphism.

Claim 7.1.7a. The map which associates a homomorphism $h \colon F \to G^\circ$ with the pair $(L, \varphi) \in X$ consisting of the set $L := \{uv \in E(F) \mid h(u) = h(v)\}$ and the homomorphism $\varphi \colon F \oslash L \to G$, $[x]_L \mapsto h(x)$ is well-defined.

Proof of Claim. First observe that if vertices $v, w \in V(F)$ are such that $v \sim_L w$, then $h(v) = h(w)$. Indeed, if without loss of generality $v \neq w$, then there exists a walk $v, u_1, u_2, \dots, u_n, w \in L$. By construction, $|h(vu_1)| = |h(u_1u_2)| = \cdots = |h(u_nw)| = 1$ and hence $h(v) = h(w)$. In particular, $F \oslash L$ is a simple graph without any loops. Indeed, if an edge $vw \in E(F)$ is such that $v \sim_L w$, then $h(v) = h(w)$ and hence $vw \in L$.

The initial observation implies that φ is a well-defined map $V(F \oslash L) \to V(G)$. It remains to argue that φ is a homomorphism $F \oslash L \to G$. Indeed, if $[v]_L \neq [w]_L$ are adjacent in $F \oslash L$, then there exist $x \sim_L v$ and $y \sim_L w$ such that $xy \in E(F) \setminus L$. Hence, h maps xy to an edge in G°, rather than to a loop. In particular, φ maps $[v]_L$ and $[w]_L$ to an edge in G. $\lhd$

Claim 7.1.7b. The map $\mathrm{Hom}(F, G^\circ) \to X$ described in Claim 7.1.7a is surjective.

Proof of Claim. Let $L \subseteq E(F)$ and $\varphi \colon F \oslash L \to G$ be a homomorphism. Observe that this implies that $F \oslash L$ is without loops. Let $\pi \colon V(F) \to V(F \oslash L)$ denote the projection $v \mapsto [v]_L$. We claim that $h := \varphi \circ \pi$ is a homomorphism $F \to G^\circ$. Let $xy \in E(F)$. If $x \sim_L y$, then the image of xy under h is a loop since $\pi(x) = \pi(y)$. If $x \not\sim_L y$, then in particular $xy \notin L$ and there is an edge between $[x]_L$ and $[y]_L$ in $F \oslash L$. Since φ is a homomorphism, $h(x)$ and $h(y)$ are in both cases adjacent in G°.

It remains to argue that this h is mapped to (L, φ) under the construction described in Claim 7.1.7a. Write $L' := \{e \in E(F) \mid |h(e)| = 1\}$. Towards concluding that $L = L'$, let $uv \in L$. Then $h(uv) = \varphi(\pi(uv))$ is a singleton since in particular $u \sim_L v$. Hence, $L \subseteq L'$. Conversely, let $uv \in L'$. By assumption, $h(uv) = \varphi(\pi(uv))$ is a singleton and thus it remains to distinguish two cases: If $\pi(u) = \pi(v)$, then $uv \in L$ because otherwise there would be a loop at $[u]_L = [v]_L$ in $F \oslash L$. If $\pi(u) \neq \pi(v)$, then $\varphi([u]_L) = \varphi([v]_L)$ and also $uv \in L$ because otherwise there would be an edge between $[u]_L \neq [v]_L$ in $F \oslash L$ and this cannot happen since φ is a homomorphism into a simple graph.

Finally, write φ' for the homomorphism $F \oslash L \to G$ constructed from h as described in Claim 7.1.7a. Then for every vertex $x \in V(F)$ by definition, $\varphi'([x]_L) = h(x) = \varphi(\pi(x)) = \varphi([x]_L)$ and thus $\varphi = \varphi'$. $\triangleleft$

It is easy to see that the map devised in Claim 7.1.7a is injective. The desired equation follows observing that the map in Claim 7.1.7a provides a bijection between the sets whose elements are counted on the right and left hand-side respectively. $\square$

The following Example 7.1.8 illustrates the above construction and Theorem 7.1.7.

Example 7.1.8. Let K_3 denote the complete graph with vertex set $\{1, 2, 3\}$. Then $K_3 \oslash \emptyset \cong K_3$, $K_3 \oslash \{12\} \cong K_2$, $K_3 \oslash \{12, 23\} \cong K_1^\circ$, and $K_3 \oslash \{12, 23, 13\} \cong K_1$. For every simple graph G, $\hom(K_3, G^\circ) = \hom(K_3, G) + 3\hom(K_2, G) + \hom(K_1, G)$ since $\hom(K_1^\circ, G) = 0$.

The final ingredient for the proof of Theorem 7.0.1 is the following Lemma 7.1.9, which establishes one of the implications in Figure 7.1.

Lemma 7.1.9. *A homomorphism distinguishing closed graph class $\mathcal{F}$ is closed under deleting edges if, and only if, it is closed under taking subgraphs.*

Proof. If $\mathcal{F}$ is closed under taking subgraphs, then it is closed under deleting edges. Conversely, since $\mathcal{F}$ is assumed to be closed under deleting edges, it suffices to show that it is closed under deleting vertices. Let $F \in \mathcal{F}$ be a simple graph on n vertices. Since the graph nK_1 can be obtained from F by deleting all its edges, it holds that $nK_1 \in \mathcal{F}$. Thus, for all simple graphs G and H such that $G \equiv_{\mathcal{F}} H$, by Equation (2.1), $|V(G)|^n = \hom(nK_1, G) = \hom(nK_1, H) = |V(H)|^n$. Hence, $\hom(K_1, G) = \hom(K_1, H)$. Now let F' denote the graph obtained from F by deleting a vertex $v \in V(F)$. By deleting all incident edges, it can be assumed that v is isolated in F and that $F' + K_1 \cong F$. Hence, for all G and H as above, by Equation (2.1),

$$\hom(F', G)\hom(K_1, G) = \hom(F, G) = \hom(F, H) = \hom(F', H)\hom(K_1, H).$$

It follows that $\hom(F', G) = \hom(F', H)$ and $F' \in \mathcal{F}$. $\square$

Proof of Theorem 7.0.1. Assuming 1, let G and H be simple graphs such that $G \equiv_{\mathcal{F}} H$ and let $F \in \mathcal{F}$. By Theorem 7.1.7 and Lemma 7.1.6,

$$\hom(F, \overline{G}) = \hom(F, \widehat{G^\circ}) = \sum_{\substack{F' \subseteq F, \\ V(F') = V(F)}} (-1)^{|E(F')|} \sum_{L \subseteq E(F')} \hom(F' \oslash L, G). \tag{7.3}$$

All simple graphs $F' \oslash L$ appearing in this sum are obtained from F by repeated edge contractions or deletions. Hence, $\overline{G} \equiv_{\mathcal{F}} \overline{H}$.

Assuming 2, it is first shown that $\mathrm{cl}(\mathcal{F})$ is closed under deleting and contracting edges. To that end, let $F \in \mathrm{cl}(\mathcal{F})$. Let K be obtained from F by deleting an edge $e \in E(F)$. Since K has the same number of vertices as F and precisely one edge less, the only summation indices (F', L) in Equation (7.3) such that $F' \oslash L \cong K$ satisfy $|E(F')| = |E(K)|$ and $L = \emptyset$. Hence, the coefficient of $\hom(K, -)$ in the linear combination obtained from Equation (7.3) by grouping isomorphic summation indices is a non-zero multiple of $(-1)^{|E(K)|}$. Hence, $K \in \mathrm{cl}(\mathcal{F})$ by Lemma 7.1.2.

Let now K be obtained from F by contracting a single edge e. By deleting edges, it can be supposed without loss of generality that there are no vertices in F which are adjacent to both end vertices of e. By the following Claim 7.1.9a, $|E(K)| = |E(F)| - 1$.

Claim 7.1.9a. For a simple graph F with $uv \in E(F)$, write $\Delta_F(uv) := \{w \in V(F) \mid vw, uw \in E(F)\}$ for the set of vertices inducing a triangle with the edge uv. Let F be a simple graph and $uv \in E(F)$. Then

$$|\Delta_F(uv)| = |E(F)| - |E(F \oslash \{uv\})| - 1.$$

Proof of Claim. Write $K = F \oslash \{uv\}$. By writing x for the vertex of K obtained by contraction and identifying $V(K) \setminus \{x\} = V(F) \setminus \{u, v\}$, for all $w \in V(K)$,

$$\deg_K(w) = \begin{cases} \deg_F(u) + \deg_F(v) - |\Delta_F(uv)| - 2, & \text{if } w = x, \\ \deg_F(w) - 1, & \text{if } w \in \Delta_F(uv), \\ \deg_F(w), & \text{otherwise.} \end{cases}$$

The desired equation now follows readily from the Handshaking Lemma. ◁

Hence, all summation indices (F', L) in Equation (7.3) such that $F' \oslash L \cong K$ must satisfy the following:

1. $L = \{e'\}$ for some edge $e' \in E(F)$.
   This is immediate from $|V(K)| = |V(F)| - 1$ and $V(F) = V(F')$.
2. $F' = F$.
   Indeed, if $F' \oslash \{e'\} \cong K$ for some $F' \subseteq F$ with $e' \in E(F')$ and $V(F') = V(F)$, then, by Claim 7.1.9a, $|E(F')| - 1 = |E(F' \oslash \{e'\})| + |\Delta_{F'}(e')| \geq |E(F' \oslash \{e'\})| = |E(K)| = |E(F)| - 1$ and thus $E(F') = E(F)$.

Each of these summation indices contributes $(-1)^{|E(F')|} = (-1)^{|E(F)|}$ to the coefficient of $\hom(K, -)$ in the linear combination ranging over non-isomorphic graphs. Hence, this coefficient is non-zero. By Lemma 7.1.2, $K \in \mathrm{cl}(\mathcal{F})$.

Hence, $\mathrm{cl}(\mathcal{F})$ is closed under deleting and contracting edges. By Lemma 7.1.9, it is closed under taking minors. The implication 3 ⇒ 2 follows from 1 ⇒ 2 for $\mathrm{cl}(\mathcal{F})$ since $\equiv_{\mathcal{F}}$ and $\equiv_{\mathrm{cl}(\mathcal{F})}$ coincide on simple graphs. □

7.1.4 Taking Subgraphs and Preservation under Full Complements

Theorem 7.1.10, which relates the property of a graph class $\mathcal{F}$ to be closed under taking subgraphs to the property of $\equiv_{\mathcal{F}}$ to be preserved under taking full complements, can now be extracted from the insights in Section 7.1.3. Since our definition of the homomorphism distinguishing closure involves only simple graphs, Theorem 7.1.10 deviates slightly from the other results in Table 7.1. This is because the relations $\equiv_{\mathcal{F}}$ and $\equiv_{\mathrm{cl}(\mathcal{F})}$ a priori coincide only on simple graphs and not necessarily on all graphs, a crucial point raised by a reviewer.

Theorem 7.1.10. *For a graph class $\mathcal{F}$ and the assertions*

1. *$\mathcal{F}$ is closed under deleting edges,*
2. *$\equiv_{\mathcal{F}}$ is* preserved under taking full complements of arbitrary graphs*, i.e., for all graphs G and H, it holds that $G \equiv_{\mathcal{F}} H$ if, and only if, $\widehat{G} \equiv_{\mathcal{F}} \widehat{H}$,*
3. *$\equiv_{\mathcal{F}}$ is* preserved under taking full complements*, i.e., for all simple graphs G and H, it holds that $G \equiv_{\mathcal{F}} H$ if, and only if, $\widehat{G} \equiv_{\mathcal{F}} \widehat{H}$,*
4. *$\mathrm{cl}(\mathcal{F})$ is closed under deleting edges,*
5. *$\mathrm{cl}(\mathcal{F})$ is* closed under taking subgraphs*, i.e. it is closed under deleting edges and vertices,*
6. *$\equiv_{\mathrm{cl}(\mathcal{F})}$ is preserved under taking full complements of arbitrary graphs,*
7. *$\equiv_{\mathrm{cl}(\mathcal{F})}$ is preserved under taking full complements,*

the implications 1 $\Rightarrow$ 2 $\Rightarrow$ 3 $\Rightarrow$ 4 $\Leftrightarrow$ 5 $\Leftrightarrow$ 6 $\Leftrightarrow$ 7 hold.

Proof. Assuming 1, by Lemma 7.1.6, for all not necessarily simple graphs G and H, if $G \equiv_{\mathcal{F}} H$ and $\mathcal{F}$ is closed under deleting edges it holds that $\hom(F, \widehat{G}) = \hom(F, \widehat{H})$ for all $F \in \mathcal{F}$ as all graphs F' appearing in Lemma 7.1.6 are obtained from F by deleting edges. Hence, 2 holds. Clearly, 2 implies 3.

Suppose that 3 holds. Since Lemma 7.1.2 is not directly applicable since the condition in 3 involves graphs with loops, we first prove the following claim:

Claim 7.1.10a. Let $\mathcal{F}$ be a graph class satisfying 3. For all graphs G and H with loops at every vertex, if $G \equiv_{\mathcal{F}} H$, then $G \equiv_{\mathrm{cl}(\mathcal{F})} H$.

Proof of Claim. Let K be an arbitrary graph with a loop at every vertex. Then $G \times K$ and $H \times K$ are graph with loops at every vertex. Hence, their full complements $\widehat{G \times K}$ and $\widehat{H \times K}$ are simple graphs.

Let $F \in \mathrm{cl}(\mathcal{F})$. By Equation (2.2), if $G \equiv_{\mathcal{F}} H$, then $G \times K \equiv_{\mathcal{F}} H \times K$. By 3, $\widehat{G \times K} \equiv_{\mathcal{F}} \widehat{H \times K}$, which implies that $\hom(F, \widehat{G \times K}) = \hom(F, \widehat{H \times K})$ since $\equiv_{\mathcal{F}}$ and $\equiv_{\mathrm{cl}(\mathcal{F})}$ coincide on simple graphs. By Lemma 7.1.6 and Equation (2.2), this in turn implies that

$$\sum_{F'} (-1)^{|E(F')|} \hom(F', G) \hom(F', K) = \sum_{F'} (-1)^{|E(F')|} \hom(F', H) \hom(F', K) \quad (7.4)$$

where both sums range over $F' \subseteq F$ such that $V(F') = V(F)$.

Let $n := |V(F)|$ and write $\mathcal{L}$ for a collection of non-isomorphic simple graphs on at most n vertices as in Lemma 3.1.2. We claim that the matrix $(\hom(F, G^\circ)))_{F,G \in \mathcal{L}}$ is invertible. Indeed, by Theorem 7.1.7,

$$\hom(F, G^\circ) = \sum_{L \subseteq E(F)} \hom(F \oslash L, G) = \sum_{F' \in \mathcal{L}} |\{L \subseteq E(F) \mid F \oslash L \cong F'\}| \hom(F', G).$$

By Lemma 3.1.2, the matrix $(\hom(F, G))_{F,G \in \mathcal{L}}$ is invertible. When ordering the elements of $\mathcal{L}$ first by number of vertices and then by number of edges, the matrix $(|\{L \subseteq E(F) \mid F \oslash L \cong F'\}|)_{F,F' \in \mathcal{L}}$ is upper triangular and all its diagonal entries are 1 since $F \oslash L \cong F$ if, and only if, $L = \emptyset$. Hence, $(\hom(F, G^\circ))_{F,G \in \mathcal{L}}$ is invertible as the product of two invertible matrices. As in the proof of Lemma 7.1.2, one may multiply Equation (7.4) with the inverse of this matrix to conclude that $\hom(F, G) = \hom(F, H)$. It follows that $G \equiv_{\mathrm{cl}(\mathcal{F})} H$. ◁

By 3, for all simple graphs G and H, the assumption $G \equiv_{\mathcal{F}} H$ implies that $\widehat{G} \equiv_{\mathcal{F}} \widehat{H}$. Hence, by Claim 7.1.10a, $\widehat{G} \equiv_{\mathrm{cl}(\mathcal{F})} \widehat{H}$. Finally, by Lemma 7.1.6, for $F \in \mathrm{cl}(\mathcal{F})$,

$$\sum_{\substack{F' \subseteq F \\ V(F') = V(F)}} (-1)^{|E(F')|} \hom(F', G) = \sum_{\substack{F' \subseteq F \\ V(F') = V(F)}} (-1)^{|E(F')|} \hom(F', H).$$

Since G and H are simple, we are in the setting of Lemma 7.1.2. Let K be any graph obtained from F by deleting edges, i.e. $V(K) = V(F)$ and $E(K) \subseteq E(F)$. Each summation index F' in Lemma 7.1.6 such that $F' \cong K$ contributes $(-1)^{|E(F')|} = (-1)^{|E(K)|}$. Hence, the coefficient of $\hom(K, -)$ in the linear combination obtained from the one above by grouping isomorphic summation indices is non-zero. By Lemma 7.1.2, it holds that $F' \in \mathrm{cl}(\mathcal{F})$.

The implication 4 ⇒ 5 follows from Lemma 7.1.9. The implication 5 ⇒ 6 follows from 1 ⇒ 3. The implication 6 ⇒ 7 is immediate. The final implication 7 ⇒ 4 follows from 3 ⇒ 4 observing that $\mathrm{cl}(\mathrm{cl}(\mathcal{F})) = \mathrm{cl}(\mathcal{F})$. □

7.1.5 Taking Induced Subgraphs, Contracting Edges, and Lexicographic Products

In this section, it is shown that a homomorphism distinguishing closed graph class is closed under taking induced subgraphs (contracting edges) if, and only if, its homomorphism indistinguishability relation is preserved under lexicographic products with a fixed graph from the left (from the right).

Examples for equivalence relations preserved under lexicographic products related to chromatic graph parameters are listed in Corollary 7.1.18. Further examples are of model-theoretic nature. It is for example easy to see that winning strategies of the Duplicator player in Hella's bijective pebble game [91] can be composed along lexicographic products.

A correspondence between taking induced subgraphs and left lexicographic products is established by the following Theorem 7.1.11:

Theorem 7.1.11. *For a graph class $\mathcal{F}$ and the assertions*

1. *$\mathcal{F}$ is closed under taking induced subgraphs,*
2. *$\equiv_{\mathcal{F}}$ is* preserved under left lexicographic products, *i.e., for all simple graphs G, H, and H', if $H \equiv_{\mathcal{F}} H'$, then $G \cdot H \equiv_{\mathcal{F}} G \cdot H'$,*
3. $\mathrm{cl}(\mathcal{F})$ *is closed under taking induced subgraphs.*

the implications $1 \Rightarrow 2 \Leftrightarrow 3$ hold.

An example for a lexicographic product from the right is the *n-blow-up $G \cdot \overline{K_n}$* of a graph G. By Lemma 7.1.1, every homomorphism indistinguishability relation is preserved under blow-ups. Preservation under arbitrary lexicographic products from the right, however, is a non-trivial property corresponding to the associated graph class being closed under edge contractions:

Theorem 7.1.12. *For a graph class $\mathcal{F}$ and the assertions*

1. *$\mathcal{F}$ is closed under edge contractions,*
2. *$\equiv_{\mathcal{F}}$ is* preserved under right lexicographic products, *i.e. for all simple graphs G, G', and H, if $G \equiv_{\mathcal{F}} G'$, then $G \cdot H \equiv_{\mathcal{F}} G' \cdot H$,*
3. $\mathrm{cl}(\mathcal{F})$ *is closed under edge contractions.*

the implications $1 \Rightarrow 2 \Leftrightarrow 3$ hold.

Towards Theorems 7.1.11 and 7.1.12, homomorphism counts $\hom(F, G \cdot H)$ are written as linear combinations of homomorphism counts $\hom(F', H)$ where F' ranges over induced subgraphs of F in the case of Theorem 7.1.11, or as linear combinations of homomorphism counts of $\hom(F'', G)$ where F'' ranges over graphs obtained from F by contracting edges in the case of Theorem 7.1.12. The following succinct formula for counts of homomorphisms into a lexicographic product generalises Equation (7.1).

Theorem 7.1.13. *Let F, G, and H be simple graphs. Then*

$$\hom(F, G \cdot H) = \sum_{\mathcal{R}} \hom(F/\mathcal{R}, G) \hom(\coprod_{R \in \mathcal{R}} F[R], H)$$

where the outer sum ranges over all partitions $\mathcal{R} \in \Pi(V(F))$ such that $F[R]$ is connected for all parts $R \in \mathcal{R}$.

Proof. We construct a bijection between the set of homomorphisms $F \to G \cdot H$ and the set of triples $(\mathcal{R}, g, h)$ where $\mathcal{R} \in \Pi(V(F))$ is such that all $F[R]$, $R \in \mathcal{R}$, are connected, and $g \colon F/\mathcal{R} \to G$ and $h \colon \coprod_{R \in \mathcal{R}} F[R] \to H$ are homomorphisms. To that end, write $\pi_G \colon V(G \cdot H) \to V(G)$ and $\pi_H \colon V(G \cdot H) \to V(H)$ for the projection maps.

Let $f \colon F \to G \cdot H$ be a homomorphism. Define a partition $\mathcal{R}'$ of $V(F)$ with parts $(\pi_G \circ f)^{-1}(v)$ for $v \in V(G)$ and $\mathcal{R} \leq \mathcal{R}'$ as the coarsest partition whose parts $R \in \mathcal{R}$

all induce connected subgraphs $F[R]$, i.e. for every $R' \in \mathcal{R}'$, the partition $\mathcal{R}$ contains one part for every connected component of $F[R']$.

The homomorphism $g\colon F/\mathcal{R} \to G$ is given by the map sending $R \in \mathcal{R}$ to $(\pi_G \circ f)(v)$ for any $v \in R$. By definition of $\mathcal{R}$, this map is well-defined. It is indeed a homomorphism since, if $R_1, R_2 \in \mathcal{R}$ are adjacent in $F/\mathcal{R}$, there exist $x_1 \in R_1$ and $x_2 \in R_2$ such that $x_1x_2 \in E(F)$. Observe that $R_1 \neq R_2$ because $F/\mathcal{R}$ is simple and hence x_1 and x_2 lie in different classes of $\mathcal{R}'$. As $(\pi_G \circ f)(x_1) \neq (\pi_G \circ f)(x_2)$, it holds that $(\pi_G \circ f)(x_1)$ and $(\pi_G \circ f)(x_2)$ are adjacent vertices in G because f is a homomorphism. Hence, $g(R_1)$ and $g(R_2)$ are adjacent in G.

The homomorphism $h\colon \coprod_{R\in\mathcal{R}} F[R] \to H$ is given by $\pi_H \circ f$. This is indeed a homomorphism since if $x_1, x_2 \in V(F)$ are adjacent in $\coprod_{R\in\mathcal{R}} F[R]$, then they lie in the same part of $\mathcal{R}'$ and hence $h(x_1)h(x_2)$ is an edge of H.

For injectivity, suppose that $f, f'\colon F \to G \cdot H$ are both mapped to $(\mathcal{R}, g, h)$. Then, $\pi_H \circ f = h = \pi_H \circ f'$ and furthermore for every $v \in V(G)$ with $v \in R$ for some $R \in \mathcal{R}$, $(\pi_G \circ f)(v) = g(R) = g'(R) = (\pi_G \circ f)(v)$. Hence, $f = f'$.

For surjectivity, let $(\mathcal{R}, g, h)$ be a triple where $\mathcal{R} \in \Pi(V(F))$ is such that all $F[R]$, $R \in \mathcal{R}$, are connected, and $g\colon F/\mathcal{R} \to G$ and $h\colon \coprod_{R\in\mathcal{R}} F[R] \to H$ are homomorphisms. Define $f\colon F \to G \cdot H$ by $v \mapsto ((g \circ \rho)(v), h(v))$ where $\rho\colon F \to F/\mathcal{R}$ is the map sending $v \in V(F)$ to $R \in \mathcal{R}$ such that $v \in R$.

The map f is a homomorphism. Indeed, for $uv \in E(F)$, distinguish cases: If $(g \circ \rho)(u) = (g \circ \rho)(v)$, then $\rho(u) = \rho(v)$ since otherwise $\rho(u)$ and $\rho(v)$ are adjacent in $F/\mathcal{R}$ and $(g \circ \rho)(u) \neq (g \circ \rho)(v)$ as g is a homomorphism into a simple graph. In this case, $u, v \in R$ for some $R \in \mathcal{R}$ and thus $h(u)h(v)$ is an edge of H. If $(g \circ \rho)(u) \neq (g \circ \rho)(v)$, then in particular $\rho(u) \neq \rho(v)$, $\rho(u)\rho(v)$ is an edge of $F/\mathcal{R}$, and hence $(g \circ \rho)(u)$ and $(g \circ \rho)(v)$ are adjacent in G. In any case, $f(u)f(v)$ is an edge of $G \cdot H$.

Write $(\mathcal{R}', g', h')$ for the image of f under the aforementioned construction. It is claimed that $(\mathcal{R}', g', h') = (\mathcal{R}, g, h)$.

To argue that $\mathcal{R} = \mathcal{R}'$, let $x_1, \dots, x_\ell$ be a path in F such that $(\pi_G \circ f)(x_1) = \cdots = (\pi_G \circ f)(x_\ell)$. Then $(g \circ \rho)(x_1) = \cdots = (g \circ \rho)(x_\ell)$ by construction. It has to be shown that $\rho(x_1) = \cdots = \rho(x_\ell)$. If $\rho(x_i) \neq \rho(x_{i+1})$ for some $1 \leq i < \ell$, then $\rho(x_i)$ and $\rho(x_{i+1})$ are adjacent in $F/\mathcal{R}$, which cannot be since G is simple and both of these vertices have the same image under g. This implies that if x, y are in the same part of $\mathcal{R}'$, then they are in the same part of $\mathcal{R}$. Conversely, let $x_1, \dots, x_\ell$ be a path in F such that all $x_1, \dots, x_\ell \in R$ for some $R \in \mathcal{R}$. Then $\rho(x_1) = \cdots = \rho(x_\ell)$ and hence $(g \circ \rho)(x_1) = \cdots = (g \circ \rho)(x_\ell)$. In particular, if x, y are in the same part of $\mathcal{R}$, then they are in the same part under $\mathcal{R}'$.

To argue that $g = g'$, note that, for any $R \in \mathcal{R} = \mathcal{R}'$ with $v \in R$, $g'(R) = (\pi_G \circ f)(v) = (g \circ \rho)(v) = g(R)$. Finally, $h' = \pi_H \circ f = h$. □

Theorem 7.1.13 yields Theorems 7.1.11 and 7.1.12 as follows:

Proof of Theorem 7.1.11. That 1 implies 2 is immediate from Theorem 7.1.13.

Let $F \in \text{cl}(\mathcal{F})$ and $U \subseteq V(F)$. In order to show that $F[U] \in \text{cl}(\mathcal{F})$ assuming 2, let $G = K_n$ where $n := |V(F)|$. Observe that $\hom(F/\mathcal{P}, G) > 0$ for every $\mathcal{P} \in \Pi(V(F))$. For the discrete partition $\mathcal{D}$, $\coprod_{R \in \mathcal{D}} F[R] \cong nK_1$. By Lemma 7.1.2, $K_1 \in \text{cl}(\mathcal{F})$. Write $\mathcal{P}$ for the partition whose parts are formed by the connected components of $F[U]$ and singleton parts otherwise. Then $\coprod_{P \in \mathcal{P}} F[P] \cong F[U] + mK_1$ where $m := |V(F) \setminus U|$. By Lemma 7.1.2, $F[U] + mK_1 \in \text{cl}(\mathcal{F})$ which implies, as for example argued in Lemma 7.1.9, that $F[U] \in \text{cl}(\mathcal{F})$. Hence, 2 implies 3.

The implication 3 $\Rightarrow$ 2 follows from 1 $\Rightarrow$ 2 for $\text{cl}(\mathcal{F})$ since $\equiv_{\mathcal{F}}$ and $\equiv_{\text{cl}(\mathcal{F})}$ coincide on simple graphs. □

Proof of Theorem 7.1.12. That 1 implies 2 is immediate from Theorem 7.1.13.

Let $F \in \text{cl}(\mathcal{F})$. It is to show that $F/\mathcal{P} \in \text{cl}(\mathcal{F})$ for every $\mathcal{P} \in \Pi(V(F))$ such that all $F[P]$ are connected for $P \in \mathcal{P}$. To that end, write $n := |V(F)|$ and let $H = K_n$. Observe that $\hom(F[U], H) > 0$ for every $U \subseteq V(F)$. Hence, the coefficient of $\hom(F/\mathcal{P}, -)$ in the linear combination from Theorem 7.1.13 is non-zero. By Lemma 7.1.2, $F/\mathcal{P} \in \text{cl}(\mathcal{F})$. Hence, 2 implies 3.

The implication 3 $\Rightarrow$ 2 follows from 1 $\Rightarrow$ 2 for $\text{cl}(\mathcal{F})$ since $\equiv_{\mathcal{F}}$ and $\equiv_{\text{cl}(\mathcal{F})}$ coincide on simple graphs. □

Theorems 7.1.11 and 7.1.12 can be combined to yield the following Corollary 7.1.14:

Corollary 7.1.14. *For a graph class $\mathcal{F}$ and the assertions*

1. *$\mathcal{F}$ is closed under taking induced subgraphs and edge contractions,*
2. *$\equiv_{\mathcal{F}}$ is such that, for all simple graphs G, G', H, and H', if $G \equiv_{\mathcal{F}} G'$ and $H \equiv_{\mathcal{F}} H'$, then $G \cdot H \equiv_{\mathcal{F}} G' \cdot H'$,*
3. *$\text{cl}(\mathcal{F})$ is closed under taking induced subgraphs and edge contractions.*

the implications 1 $\Rightarrow$ 2 $\Leftrightarrow$ 3 hold.

Proof. Assuming 1, if $G \equiv_{\mathcal{F}} G'$ and $H \equiv_{\mathcal{F}} H'$, then $G \cdot H \equiv_{\mathcal{F}} G' \cdot H$ by Theorem 7.1.12 and $G' \cdot H \equiv_{\mathcal{F}} G' \cdot H'$ by Theorem 7.1.11. By transitivity, $G \cdot H \equiv_{\mathcal{F}} G' \cdot H'$ and 2 holds. The implication 2 $\Rightarrow$ 3 is immediate from Theorems 7.1.11 and 7.1.12. The implication 3 $\Rightarrow$ 2 follows from 1 $\Rightarrow$ 2 for $\text{cl}(\mathcal{F})$ since $\equiv_{\mathcal{F}}$ and $\equiv_{\text{cl}(\mathcal{F})}$ coincide on simple graphs. □

As a final observation, the following Lemma 7.1.15 relates the property of being closed under edge contractions to the other closure properties in Table 7.1 and Figure 7.1.

Lemma 7.1.15. *If a homomorphism distinguishing closed graph class $\mathcal{F}$ is closed under contracting edges, then it is closed under taking summands.*

Proof. Let $F \in \mathcal{F}$. Since every homomorphism distinguishing closed graph class is closed under disjoint unions, cf. Lemma 6.2.2, it suffices to show that every connected component C of F is in $\mathcal{F}$. Let m denote the number of connected components of F. By contracting all edges, the graph mK_1 can be obtained from F. Hence, as argued in Lemma 7.1.9, $K_1 \in \mathcal{F}$. Moreover, the graph $C + (m-1)K_1$ can be obtained from F by contracting all edges not in C. This implies as in Lemma 7.1.9 that $C \in \mathcal{F}$. □

Finally, it is noted that no other implications hold between the closure properties considered in Figure 7.1.

Example 7.1.16.
1. The graph class $\{nP_3 \mid n \in \mathbb{N}\}$ is homomorphism distinguishing closed (Corollary 6.5.3) and closed under taking summands but neither under edges contractions nor under taking induced subgraphs.
2. The class of disjoint unions of cycles is homomorphism distinguishing closed (Theorem 7.1.4) and closed under edge contractions but not under taking induced subgraphs.
3. The class of disjoint unions of complete graphs is homomorphism distinguishing closed (Corollary 6.5.3) and closed under taking induced subgraphs but not under deleting edges.
4. The class of graphs of bounded degree is homomorphism distinguishing closed (Corollary 6.3.10) and closed under edge deletions but not closed under edge contractions.

7.1.6 Applications

As applications of Theorems 7.0.1 and 7.1.10 to 7.1.12, we conclude, in the spirit of [14], that certain equivalence relations on graphs cannot be homomorphism distinguishing relations.

Corollary 7.1.17. *Let $\mathcal{F}$ be a graph class containing a non-empty graph such that one of the following holds:*
1. *$\equiv_{\mathcal{F}}$ is preserved under complements, cf. Theorem 7.0.1,*
2. *$\equiv_{\mathcal{F}}$ is preserved under full complements, cf. Theorem 7.1.10,*
3. *$\equiv_{\mathcal{F}}$ is preserved under left lexicographic products, cf. Theorem 7.1.11, or*
4. *$\equiv_{\mathcal{F}}$ is preserved under right lexicographic products, cf. Theorem 7.1.12.*

Then $G \equiv_{\mathcal{F}} H$ implies that $|V(G)| = |V(H)|$ for all simple graphs G and H.

Proof. By Theorems 7.0.1 and 7.1.10 to 7.1.12, $\mathcal{F}$ can be chosen to be closed under taking minors, subgraphs, induced subgraphs, or contracting edges. In any case, $K_1 \in \mathcal{F}$ and hence $|V(G)| = \hom(K_1, G) = \hom(K_1, H) = |V(H)|$. □

As concrete examples, consider the following relations. For a simple graph G, $\operatorname{aut}(G)$ denotes the order of the automorphism group of G, $\alpha(G)$ denotes the size of the largest independent set of G, $\omega(G)$ denotes the size of the largest clique of G, and $\chi(G)$ denotes the chromatic number of G.

Corollary 7.1.18. *There is no graph class $\mathcal{F}$ satisfying any of the following assertions for all simple graphs G and H:*

1. $G \equiv_{\mathcal{F}} H$ *if, and only if,* $\operatorname{aut}(G) = \operatorname{aut}(H)$*,*
2. $G \equiv_{\mathcal{F}} H$ *if, and only if,* $\alpha(G) = \alpha(H)$*,*
3. $G \equiv_{\mathcal{F}} H$ *if, and only if,* $\omega(G) = \omega(H)$*,*
4. $G \equiv_{\mathcal{F}} H$ *if, and only if,* $\chi(G) = \chi(H)$*.*

Proof. The relation in Item 1 is preserved under taking complements. By [74, Theorem 1, Corollary p. 90], the relations in Items 2 and 4 are preserved under left lexicographic products. For Item 3, the same follows from [74, Theorem 1] observing that $\overline{G} \cdot \overline{H} = \overline{G \cdot H}$ and $\omega(G) = \alpha(\overline{G})$. In each case, it is easy to exhibit a pair of graphs G and H in the same equivalence class with different number of vertices. By Corollary 7.1.17, none of the equivalence relations is a homomorphism indistinguishability relation. □

7.2 Equivalences over Self-Complementary Logics

In this section, Theorem 7.2.2 is derived from Theorem 7.0.1. The theorem applies to self-complementary logics, of which examples are given subsequently. Finally, a result from graph minor theory is used to relate logics on graphs to quantum isomorphism. The very general definition of logics on graphs from Section 2.2 needs to be strengthened only slightly in order to yield Theorem 7.2.2:

Definition 7.2.1. A logic on graphs $(\mathsf{L}, \models)$ is *self-complementary* if, for all $\varphi \in \mathsf{L}$, there is a $\overline{\varphi} \in \mathsf{L}$ such that, for all simple graphs G, it holds that $G \models \varphi$ if, and only if, $\overline{G} \models \overline{\varphi}$.

Theorem 7.2.2. *Let* $(\mathsf{L}, \models)$ *be a self-complementary logic on graphs for which there exists a graph class $\mathcal{F}$ such that simple graphs are homomorphism indistinguishable over $\mathcal{F}$ if, and only if, they are* L*-equivalent. Then there exists a minor-closed graph class $\mathcal{F}'$ whose homomorphism indistinguishability relation coincides with* L*-equivalence.*

Proof. It is shown that $\equiv_{\mathcal{F}}$ is preserved under taking complements in the sense of Theorem 7.0.1. Suppose that $G \equiv_{\mathcal{F}} H$ for simple graphs G and H. By assumption, for all $\varphi \in \mathsf{L}$, it holds that $G \models \varphi$ if, and only if, $H \models \varphi$ and hence, by self-complementarity,

$$\overline{G} \models \varphi \iff G \models \overline{\varphi} \iff H \models \overline{\varphi} \iff \overline{H} \models \overline{\overline{\varphi}} \iff \overline{H} \models \varphi.$$

Here, the ultimate equivalence holds since $H \models \varphi$ if, and only if, $H \models \overline{\overline{\varphi}}$ for all φ and H by the definition of self-complementarity observing that $\overline{\overline{H}} \cong H$. Thus, $\overline{G} \equiv_{\mathcal{F}} \overline{H}$. By Theorem 7.0.1, $\mathcal{F}' := \operatorname{cl}(\mathcal{F})$ is minor-closed. □

In particular, by Corollary 7.1.17, for a self-complementary logic L, all L-equivalent simple graphs G and H must have the same number of vertices unless L is trivial in the sense that all simple graphs G and H are L-equivalent.

7.2.1 Examples for Self-Complementary Logics

A first example of a self-complementarity logic is first-order logic FO over the signature of graphs $\{E\}$. In order to establish this property, a formula $\overline{\varphi} \in \mathsf{FO}$ has to be constructed for every $\varphi \in \mathsf{FO}$ such that $G \models \varphi$ if, and only if, $\overline{G} \models \overline{\varphi}$ for all graphs G. Only subformulas Exy require non-trivial treatment:

Definition 7.2.3. For every $\varphi \in \mathsf{FO}$, define $\overline{\varphi} \in \mathsf{FO}$ inductively as follows:

1. if $\varphi = Exy$, then $\overline{\varphi} := \neg Exy \wedge (x \neq y)$,
2. if $\varphi = \bot$ or $\varphi = \top$, then $\overline{\varphi} := \varphi$, if $\varphi = (x = y)$, then $\overline{\varphi} := \varphi$, if $\varphi = \neg\psi$, then $\overline{\varphi} := \neg\overline{\psi}$, if $\varphi = \psi \wedge \chi$, then $\overline{\varphi} := \overline{\psi} \wedge \overline{\chi}$, if $\varphi = \psi \vee \chi$, then $\overline{\varphi} := \overline{\psi} \vee \overline{\chi}$, if $\varphi = \exists x\psi$, then $\overline{\varphi} := \exists x\overline{\psi}$, and if $\varphi = \forall x\psi$, then $\overline{\varphi} := \forall x\overline{\psi}$.

Lemma 7.2.4. *Let $\varphi \in \mathsf{FO}$ be a formula with $k \geq 0$ free variables. Then, for all simple graphs G with $\boldsymbol{v} \in V(G)^k$, it holds that*

$$G \models \varphi(\boldsymbol{v}) \iff \overline{G} \models \overline{\varphi}(\boldsymbol{v}).$$

In particular, FO *is self-complementary.*

Proof. The proof is by induction on the structure of φ. If φ is Exy, observe that $v_1v_2 \in E(G)$ if, and only if, $v_1v_2 \notin E(\overline{G})$ and $v_1 \neq v_2$. In all other cases, the claim is purely syntactical. □

Lemma 7.2.4 gives a purely syntactical sufficient criterion for a fragment $\mathsf{L} \subseteq \mathsf{FO}$ to be self-complementary. Indeed, if $\overline{\varphi} \in \mathsf{L}$ as defined in Definition 7.2.3 for all $\varphi \in \mathsf{L}$, then L is self-complementary. Note that the operation in Definition 7.2.3 increases neither the number of variables nor affects the quantifiers in the formula. Thus, Lemma 7.2.4 automatically extends to fragments of FO defined by restricting the number of variables, order or number of quantifiers. For extensions of FO, Definition 7.2.3 can be easily extended. This yields a rich realm of self-complementary logics, of which the following Example 7.2.5 lists only a selection.

Example 7.2.5. The following logics on graphs are self-complementary. For every $k, d \geq 0$,

1. the k-variable and quantifier-depth-d fragments FO^k and FO_d of FO,
2. first-order logic with counting quantifiers C and its k-variable and quantifier-depth-d fragments C^k and C_d,
3. inflationary fixed-point logic IFP, cf. [78],
4. second-order logic SO and its fragments monadic second-order logic MSO_1, existential second-order logic ESO, cf. Section 2.2.3, [47, 66].

Corollary 7.1.17 readily gives an alternative proof of [14, Propositions 1 and 2], which assert that neither FO^k-equivalence nor FO_d-equivalence are characterised by homomorphism indistinguishability relations. By Theorems 3.4.4 and 3.4.5, the logic fragments C^k and C_d are, however, characterised by homomorphism indistinguishability relations.

7.2.2 Applications: Graph Minor Theory and Distinguishing Power

The final result of this section demonstrates how graph minor theory can yield insights into the distinguishing power of logics via Theorem 7.2.2. Subject to it are self-complementary logics which have a homomorphism indistinguishability characterisation and are stronger than C^k for every k, e.g. they are capable of distinguishing CFI graphs. It is shown that equivalence with respect to any such logic is a sufficient condition for quantum isomorphism, cf. [124].

Theorem 7.2.6. *Let $(\mathsf{L}, \models)$ be a self-complementary logic on graphs for which there exists a graph class $\mathcal{F}$ such that simple graphs are homomorphism indistinguishable over $\mathcal{F}$ if, and only if, they are L-equivalent. Suppose that, for all $k \in \mathbb{N}$, there exist simple graphs G and H such that $G \equiv_{\mathsf{C}^k} H$ and $G \not\equiv_{\mathsf{L}} H$. Then all L-equivalent simple graphs are quantum isomorphic.*

The proof of Theorem 7.2.6 is based on the following result from graph minor theory.

Theorem 7.2.7 ([153, (2.1)], cf. [133, Theorem 3.8]). *For a minor-closed graph class $\mathcal{F}$, the following are equivalent:*

1. *$\mathcal{F}$ has bounded treewidth, i.e. there exists $w \in \mathbb{N}$ such that $\operatorname{tw}(F) \leq w$ for all $F \in \mathcal{F}$,*
2. *there exists a planar graph H such that $H \notin \mathcal{F}$.*

Proof of Theorem 7.2.6. By assumption and Theorem 3.4.4, for all $k \in \mathbb{N}$, there exist simple graphs G and H such that $G \equiv_{\mathcal{TW}_k} H$ and $G \not\equiv_{\mathcal{F}} H$. This implies that $\mathcal{F} \not\subseteq \operatorname{cl}(\mathcal{TW}_k)$ for every $k \in \mathbb{N}$. Hence, $\operatorname{cl}(\mathcal{F}) \not\subseteq \mathcal{TW}_k$ for every $k \in \mathbb{N}$. By Theorem 7.0.1, $\operatorname{cl}(\mathcal{F})$ is a minor-closed graph class. Thus, by Theorem 7.2.7, it holds that $\mathcal{P} \subseteq \operatorname{cl}(\mathcal{F})$ where $\mathcal{P}$ denotes the class of all planar graphs. It follows that all L-equivalent graphs are quantum isomorphic by [124]. □

In [111], Theorem 7.2.6 was applied to show that equivalence in a specific logic stronger than first-order logic with counting quantifiers is not a homomorphism indistinguishability relation. This logic is *linear-algebraic logic* LA, as introduced in [56]. By [111, Theorem 1], equivalence in the k-variable fragment of LA is not a homomorphism indistinguishability relation for any $k \geq 6$. In [111], the graphs G and H required by Theorem 7.2.6 were taken to be CFI graphs over a planar base graph constructed with respect to a cyclic group whose order is a power of two. By Theorem 6.3.8, these CFI graphs are not quantum isomorphic. That they are distinguished by k-variable LA was shown in [111] modifying techniques of [109].

7.3 Cancellation Laws

Towards understanding the properties of homomorphism indistinguishability relations, this section is concerned with cancellation of graph isomorphism relaxations.

Definition 7.3.1. Let K be a simple graph. A graph isomorphism relaxation $\approx$ *admits K-cancellation* if, for all simple graphs G and H, it holds that $G \times K \approx H \times K \implies G \approx H$.

Lovász [115] proved that the isomorphism relation $\cong$ admits K-cancellation if, and only if, K is non-bipartite. The following Lemma 7.3.2 shows that K-cancellation of a homomorphism indistinguishability relation $\equiv_{\mathcal{F}}$ depends solely on the homomorphism distinguishing closure of the subclass $\mathcal{F}_K \subseteq \mathcal{F}$ of K-colourable graphs in $\mathcal{F}$. In fact, the proof is based on the observation that the equivalence relation $- \times K \equiv_{\mathcal{F}} - \times K$ coincides with the homomorphism indistinguishability relation $\equiv_{\mathcal{F}_K}$.

Lemma 7.3.2. *For a graph class $\mathcal{F}$ and a simple graph K, the following are equivalent:*

1. *$\equiv_{\mathcal{F}}$ admits K-cancellation,*
2. *$\mathcal{F} \subseteq \mathrm{cl}(\mathcal{F}_K)$,*
3. *$\mathrm{cl}(\mathcal{F}) = \mathrm{cl}(\mathcal{F}_K)$.*

Proof. First observe that, for simple graphs G and H, the conditions $G \times K \equiv_{\mathcal{F}} H \times K$ and $G \equiv_{\mathcal{F}_K} H$ are equivalent. Indeed, if $G \times K \equiv_{\mathcal{F}} H \times K$ and $F \in \mathcal{F}_K$, then, by Equation (2.2), $\hom(F, G) = \hom(F, G \times K) / \hom(F, K) = \hom(F, H)$ and hence $G \equiv_{\mathcal{F}_K} H$. Conversely, if $G \equiv_{\mathcal{F}_K} H$ and $F \in \mathcal{F}$, then, by Equation (2.2),

$$\hom(F, G \times K) = \hom(F, G) \hom(F, K) = \hom(F, H) \hom(F, K) = \hom(F, G \times K)$$

since $\hom(F, G) = \hom(F, H)$ if $\hom(F, K) \neq 0$. Hence, $G \times K \equiv_{\mathcal{F}} H \times K$.

By the initial observation, Item 1 is equivalent to the assertion that $G \equiv_{\mathcal{F}_K} H$ implies $G \equiv_{\mathcal{F}} H$ for all graphs G and H. Hence, Items 1 and 2 are equivalent. Since $\mathcal{F} \subseteq \mathrm{cl}(\mathcal{F})$ and $\mathrm{cl}(\mathcal{F}_K) \subseteq \mathrm{cl}(\mathcal{F})$, assertions Items 2 and 3 are equivalent. □

Essentially profinite graph classes admitting K-cancellation can be succinctly characterised.

Lemma 7.3.3. *Let $\mathcal{F}$ be an essentially profinite graph class and let K be a simple graph. Then $\equiv_{\mathcal{F}}$ admits K-cancellation if, and only if, all graphs in $\mathcal{F}$ are K-colourable.*

Proof. By Lemma 6.5.8, $\mathrm{cl}(\mathcal{F}_K) = \mathrm{cl}(\mathcal{F})_K$. By Lemma 7.3.2, if $\equiv_{\mathcal{F}}$ admits K-cancellation, then $\mathcal{F} \subseteq \mathrm{cl}(\mathcal{F}_K) = \mathrm{cl}(\mathcal{F})_K$. In particular, all graphs in $\mathcal{F}$ are K-colourable. Conversely, if $\mathcal{F} = \mathcal{F}_K$, then $\mathrm{cl}(\mathcal{F}) = \mathrm{cl}(\mathcal{F}_K)$, and $\equiv_{\mathcal{F}}$ admits K-cancellation by Lemma 7.3.2. □

In particular, if an essentially profinite graph class admits K-cancellation for some simple graph K, then it is essentially finite. Beyond essentially profinite graph classes, it is less clear when $\equiv_{\mathcal{F}}$ admits K-cancellation. If K is bipartite, then this depends solely on whether the graphs in $\mathcal{F}$ are K-colourable.

Lemma 7.3.4. *For every graph class $\mathcal{F}$ and every bipartite simple graph K, $\mathrm{cl}(\mathcal{F}_K) = \mathrm{cl}(\mathcal{F})_K$. In particular, $\equiv_{\mathcal{F}}$ admits K-cancellation if, and only if, all graphs in $\mathcal{F}$ are K-colourable.*

Proof. By Theorem 6.1.1, the class $\mathcal{G}_{K_2}$ of all bipartite graphs is homomorphism distinguishing closed. By Theorem 6.5.2, so is $\mathcal{G}_{K_1}$, the family of all graphs which are K_1-colourable, i.e. the edgeless graphs. Finally, the class $\mathcal{G}_{K_0} = \{K_0\}$, the family containing only the empty graph K_0 is homomorphism distinguishing closed. Hence, for $i \in \{0, 1, 2\}$, by Lemma 6.2.3,

$$\mathrm{cl}(\mathcal{F}_{K_i}) = \mathrm{cl}(\mathcal{F} \cap \mathcal{G}_{K_i}) \subseteq \mathrm{cl}(\mathcal{F}) \cap \mathrm{cl}(\mathcal{G}_{K_i}) = \mathrm{cl}(\mathcal{F}) \cap \mathcal{G}_{K_i} = \mathrm{cl}(\mathcal{F})_{K_i}.$$

The converse containment follows from Lemma 6.5.7. Hence, $\mathrm{cl}(\mathcal{F}_{K_i}) = \mathrm{cl}(\mathcal{F})_{K_i}$ for $i \in \{0, 1, 2\}$. It remains to observe that if K is bipartite then $\mathcal{F}_K$ equals either $\mathcal{F}_{K_0}$, $\mathcal{F}_{K_1}$, or $\mathcal{F}_{K_2}$, depending on whether K is edgeless or empty. □

Under the additional assumption that $\mathcal{F}$ is closed under subdivisions, the following Theorem 7.3.5 completes the characterisation of the graphs K for which $\equiv_{\mathcal{F}}$ admits K-cancellation. For example, the quantum isomorphism relation $\equiv_{\mathcal{P}}$ admits K-cancellation if, and only if, K is non-bipartite, cf. [124]. Theorem 7.3.5 strengthens Lovász's result [115] using an argument due to Dvořák [63].

Theorem 7.3.5. *If $\mathcal{F}$ is closed under subdivisions, then $\mathrm{cl}(\mathcal{F}) = \mathrm{cl}(\mathcal{F}_K)$ and $\equiv_{\mathcal{F}}$ admits K-cancellation for every non-bipartite simple graph K.*

Proof. The proof is by showing the contrapositive, i.e. for simple graphs G and H, if $G \not\equiv_{\mathcal{F}} H$, then $G \not\equiv_{\mathcal{F}_K} H$. Since K is non-bipartite, it contains an odd cycle C_{2n+1} for some $n \geq 1$. Let $F \in \mathcal{F}$ be such that $\hom(F, G) \neq \hom(F, H)$. Without loss of generality, F is not K-colourable. Consider the following claim:

Claim 7.3.5a. Let F be a simple graph and $n \in \mathbb{N}$. Let F' be obtained from F by replacing every edge by a path on at least $2n + 1$ vertices. Then F' is C_{2n+1}-colourable.

Proof of Claim. Pick an arbitrary vertex $x \in V(C_{2n+1})$. The C_{2n+1}-colouring of F' is constructed by mapping all vertices from $V(F) \subseteq V(F')$ to x. Let $e \in E(F)$ be an edge. If the number m of vertices introduced in F' by subdividing e is odd, then the resulting path can be mapped to x and one of its neighbours $y \in V(C_{2n+1})$. This is because every path on an odd number of vertices can be mapped to K_2 such that both of its degree-1 vertices are mapped to the same vertex. If m is even, then the path can be mapped to a walk in C_{2n+1} which starts in x, wraps around C_{2n+1} once, and finally alternates between x and one of its neighbours $y \in V(C_{2n+1})$. ◁

By Lemma 6.1.4, there exists a graph $F' \in \mathcal{F}_{C_{2n+1}} \subseteq \mathcal{F}_K$ such that $\hom(F', G) \neq \hom(F', H)$, as desired. □

The following corollary summarises the results on cancellation properties of homomorphism indistinguishability relations over graph classes closed under subdivisions.

Corollary 7.3.6. *Let $\mathcal{F}$ be a graph class which is closed under subdivisions. Let K be a simple graph. Then precisely one of the following holds:*

1. *$\mathcal{F}$ contains a graph which contains a cycle and then $\equiv_{\mathcal{F}}$ admits K-cancellation if, and only if, K is not bipartite.*
2. *$\mathcal{F}$ contains only acyclic graphs and at least one graph which contains an edge and then $\equiv_{\mathcal{F}}$ admits K-cancellation if, and only if, K contains an edge.*
3. *$\mathcal{F}$ contains only edgeless graphs and at least one non-empty graph and then $\equiv_{\mathcal{F}}$ admits K-cancellation if, and only if, K is not empty.*
4. *$\mathcal{F}$ contains only the empty graph and then $\equiv_{\mathcal{F}}$ admits K-cancellation.*

Proof. Since $\mathcal{F}$ is closed under subdivisions, $\equiv_{\mathcal{F}}$ admits K-cancellation for every non-bipartite simple graph K by Theorem 7.3.5. If $\mathcal{F}$ contains only acyclic graphs, then all graphs in $\mathcal{F}$ are bipartite and $\equiv_{\mathcal{F}}$ admits K-cancellation for every simple graph K containing an edge by Lemma 7.3.4. If $\mathcal{F}$ contains only edgeless graphs, then $\equiv_{\mathcal{F}}$ admits K-cancellation for every non-empty bipartite simple graph K by Lemma 7.3.4. If $\mathcal{F}$ contains only the empty graph, then $\equiv_{\mathcal{F}}$ admits K-cancellation for every bipartite simple graph K by Lemma 7.3.4.

If $\mathcal{F}$ contains at least one non-empty graph, then, by Lemma 7.3.4, $\equiv_{\mathcal{F}}$ does not admit K_0-cancellation for the empty graph K_0. If $\mathcal{F}$ contains at least one edge, then $\equiv_{\mathcal{F}}$ does not admit K-cancellation whenever K does not contain an edge by Lemma 7.3.4. If $\mathcal{F}$ contains a graph containing a cycle, then $\mathcal{F}$ contains a non-bipartite graph and hence $\mathcal{F}$ does not admit K-cancellation whenever K is bipartite. □

7.4 Further Directions

The results summarised in Table 7.1 give necessary conditions for an equivalence relation $\approx$ comparing graphs to be a homomorphism indistinguishability relation over a graph class with certain closure properties. It is tempting to view these results as instances of a potentially richer connection between graph-theoretic properties of $\mathcal{F}$ and *polymorphisms* of $\equiv_{\mathcal{F}}$, i.e. isomorphism-invariant[10] maps $\mathfrak{p}$ sending tuples of graphs to graphs such that $\mathfrak{p}(G_1, \dots, G_k) \equiv_{\mathcal{F}} \mathfrak{p}(H_1, \dots, H_k)$ whenever $G_i \equiv_{\mathcal{F}} H_i$ for all $i \in [k]$. Recalling the algebraic approach to constraint satisfaction problems, cf. [18], one may ask what structural insights into $\mathcal{F}$ can be gained by considering polymorphisms of $\equiv_{\mathcal{F}}$.[11] The following questions are more concrete:

[10] That is, if $G_i \cong H_i$ for all $i \in [k]$, then $\mathfrak{p}(G_1, \dots, G_k) \cong \mathfrak{p}(H_1, \dots, H_k)$.

[11] Preservation properties of homomorphism indistinguishability relations appear to be related the categorical preservation properties studied in [95]. Curiously, not all preservation properties studied

Question 7.4.1. *Can closure under topological minors of $\mathcal{F}$ be characterised in terms of some polymorphism of $\equiv_{\mathcal{F}}$?*

Graph exponentiation is an operation on graphs with loops studied in [166, Section 4.2] in the context of graph homomorphism counts.

Question 7.4.2. *Can preservation under graph exponentiation of $\equiv_{\mathcal{F}}$ be characterised in terms of a closure property of $\mathcal{F}$?*

Corollary 7.3.6 and the results in Table 7.1 may be viewed as a step towards an axiomatic characterisation of homomorphism indistinguishability relations, similar to the characterisation [69, 118] of functions $f \colon \mathcal{G} \to \mathbb{N}$ on the set of all graphs $\mathcal{G}$ which are of the form $f = \hom(-, G)$ for some graph G. Such a characterisation should give sufficient and necessary criteria for an equivalence relation $\approx$ to be a homomorphism indistinguishability relation over some graph class. In this context, exploring abstract properties of graph isomorphism relaxations and their connection to homomorphism indistinguishability is imperative. We propose two questions: The first questions aims to generalise [69, 118].

Question 7.4.3. *Let $\mathcal{F}$ be a proper graph class. How are the functions $f \colon \mathcal{F} \to \mathbb{N}$ such that $f = \hom(-, G)$ for some simple graph G characterised?*

By [34, Theorem 3], there are finite graph classes $\mathcal{F}$ for which deciding whether a function $f \colon \mathcal{F} \to \mathbb{N}$ is of the form $f = \hom(-, G)$ for some graph G is NP-hard. Question 7.4.3 might be a stepping stone for the following much vaguer question:

Question 7.4.4. *What are sufficient and necessary conditions for a graph isomorphism relaxation to be a homomorphism indistinguishability relation?*

here are functorial. For example, not every homomorphisms $G \to H$ induces a homomorphism $\overline{G} \to \overline{H}$.

8 Modular Homomorphism Indistinguishability

As an interlude and in preparation for Chapter 9, we consider homomorphism indistinguishability modulo integers $n \geq 1$. For a graph class $\mathcal{F}$, two simple graphs G and H are said to be *homomorphism indistinguishable over $\mathcal{F}$ modulo n*, in symbols $G \equiv^n_{\mathcal{F}} H$, if $\hom(F, G) \equiv \hom(F, H) \mod n$ for every $F \in \mathcal{F}$. Modular homomorphism indistinguishability has first been considered by Faben & Jerrum [67] in the context of the modular counting complexity of graph colouring problems.

The main result of this chapter is Theorem 8.3.2, which characterises homomorphism indistinguishability over graphs of bounded treewidth modulo a prime as equivalence over a version of counting logic with modular counting quantifiers. The result and its proof parallel Dvořák's Theorem 3.4.4.

Chapter Outline. We review the results of Faben & Jerrum [67] in Section 8.1. In Section 8.2, we study connections between modular homomorphism indistinguishability and non-modular homomorphism indistinguishability, i.e. the setting introduced in Chapter 3. In Section 8.3, we prove a modular analogue of Dvořák's Theorem 3.4.4. Material in this chapter has been previously published in [111, 110] as joint work with Moritz Lichter and Benedikt Pago.

8.1 Modular Homomorphism Indistinguishability over All Graphs

In contrast to Lovász's Theorem 3.1.1 asserting that two graphs are homomorphism indistinguishable over all graphs if, and only if, they are isomorphic, homomorphism counts modulo an integer n do not suffice to determine a graph up to isomorphism. To that end, consider the following example.

Example 8.1.1. For $n \geq 1$, the one-vertex graph K_1 and the $(n+1)$-vertex edgeless graph $\overline{K_{n+1}}$ are homomorphism indistinguishable over all graphs modulo n.

Proof. If F is an edgeless graph, then $\hom(F, K_1) = 1 \equiv (n+1)^{|V(F)|} = \hom(F, \overline{K_{n+1}}) \mod n$. If otherwise F contains an edge, then $\hom(F, K_1) = 0 = \hom(F, \overline{K_{n+1}})$. □

In [67], homomorphism indistinguishability over all graphs modulo a prime p was characterised using automorphisms of order p. For a graph F, the automorphism group $\mathrm{Aut}(G)$ of a simple graph G acts on the set of homomorphisms $\mathrm{Hom}(F, G)$ by composition. Thus, for an automorphism $\sigma \in \mathrm{Aut}(G)$ of order p, the set $\mathrm{Hom}(F, G)$ can be partitioned into orbits under the action of the cyclic subgroup of $\mathrm{Aut}(G)$ generated by σ. By the Orbit–Stabiliser Formula, each orbits is either a singleton or of cardinality divisible by p.[12] When counting homomorphisms modulo p, the latter orbits do not contribute. This argument implies that G and $G^\sigma := G[\{v \in V(G) \mid \sigma(v) = v\}]$, the subgraph of G induced by the fixed points of σ, are homomorphism indistinguishable over all graphs modulo p [67, Lemma 3.3].

Hence, order-p automorphism allow to reduce a graph to one of its induced subgraphs while preserving homomorphism counts. Write $G \to_p G'$ for two simple graphs G and G' if there is an automorphism σ of G of order p such that $G^\sigma \cong G'$. Write $G \to_p^* H$ if there is a sequence of graphs $G_1, \dots, G_n$ such that $G \to_p G_1 \to_p G_2 \to_p \cdots \to_p G_n \to_p H$. By [67, Theorem 3.7], for every simple graph G and prime p, there is a graph G_p^*, unique up to isomorphism, such that G_p^* has no automorphisms of order p and $G \to_p^* G_p^*$. Furthermore, by [67, Theorem 3.4], the graphs G and G_p^* are homomorphism indistinguishable over all graphs modulo p. A characterisation of homomorphism indistinguishability over all graphs modulo p can now be stated:

Theorem 8.1.2 ([67, Lemma 3.10])**.** *Let p be a prime. Two simple graphs G and H are homomorphism indistinguishable over all graphs modulo p if, and only if, G_p^* and H_p^* are isomorphic.*

8.2 Connections between Modular and Non-Modular Homomorphism Indistinguishability

We proceed by investigating the connections between modular homomorphism indistinguishability and non-modular homomorphism indistinguishability, i.e. as defined in Definition 3.0.1. The following Lemma 8.2.1 shows that non-trivial modular homomorphism indistinguishability relations cannot be expressed by non-modular homomorphism indistinguishability relations. Furthermore, considering sets of moduli does not yield more relations. We may restrict our attention to homomorphism indistinguishability relations modulo some not necessarily prime integer $n \geq 1$. In the following lemma, we write $G \equiv_{\mathcal{F}}^N H$ for a set N of positive integers if $G \equiv_{\mathcal{F}}^n H$ for every $n \in N$.

Lemma 8.2.1. *Let $\mathcal{F}$ and $\mathcal{K}$ be graph classes. Let $n \geq 1$ and N be a set of positive integers.*

1. *If N is infinite, then $\equiv_{\mathcal{F}}^N$ and $\equiv_{\mathcal{F}}$ coincide.*
2. *If N is finite and m is the least common multiple of the numbers in N, then $\equiv_{\mathcal{F}}^N$ and $\equiv_{\mathcal{F}}^m$ coincide.*

[12] Here, it is crucial that p is prime and not composite.

3. *If $\equiv_{\mathcal{F}}$ and $\equiv^n_{\mathcal{K}}$ coincide, then all pairs of graphs are homomorphism indistinguishable over $\mathcal{F}$, i.e. for all simple graphs G and H it holds that $G \equiv_{\mathcal{F}} H$.*

The proof of Lemma 8.2.1 is based on the Chinese Remainder Theorem:

Fact 8.2.2 (Chinese Remainder Theorem, cf. [131, Theorem 2.10])**.** *Let $k \geq 2$ and $a_1, \dots, a_k \in \mathbb{Z}$. Let $m_1, \dots, m_k \geq 1$ be pairwise coprime. Then there exists $x \in \mathbb{Z}$ such that $x \equiv a_i \mod m_i$ for all $i \in [k]$. Furthermore, if $x, y \in \mathbb{Z}$ are solutions to these congruences, then $x \equiv y \mod m_1 \cdots m_k$.*

Proof of Lemma 8.2.1. For the first claim, let G and H be simple graphs and $F \in \mathcal{F}$. Since N is infinite, there exists $n \in N$ greater than $\hom(F, G)$ and $\hom(F, H)$. Then $\hom(F, G) \equiv \hom(F, H) \mod n$ implies that $\hom(F, G) = \hom(F, H)$.

For the second claim, first observe that $G \equiv^m_{\mathcal{F}} H$ entails $G \equiv^N_{\mathcal{F}} H$ since all $n \in N$ divide m. Conversely, for a prime p, write $\nu(p)$ for the greatest integer $k \geq 0$ such that there is an $n \in N$ that is divisible by p^k. Then $m = \prod_p p^{\nu(p)}$, where the product ranges over all primes. Hence, if $\hom(F, G) \equiv \hom(F, H) \mod n$ for all $n \in N$, then $\hom(F, G) \equiv \hom(F, H) \mod p^{\nu(p)}$ for all primes p appearing as divisors of elements in N, i.e. $\nu(p) > 0$. Hence, by Fact 8.2.2, also $\hom(F, G) \equiv \hom(F, H) \mod m$.

Towards the third claim, we first show the following Claim 8.2.2a: Fix an integer $n \geq 1$. Write ℓ for the maximum integer such that p^ℓ divides n for some prime p. Write $\varphi \colon \mathbb{N} \to \mathbb{N}$ for Euler's totient function [131, Section 2.3] and $G^{\times k}$ for the k-th categorical power of the graph G, cf. Section 2.1.1.

Claim 8.2.2a. For every simple graph G, the graphs $G^{\times(\varphi(n)+\ell)}$ and $G^{\times \ell}$ are homomorphism indistinguishable over all graphs modulo n.

Proof of Claim. We show that $a^\ell(a^{\varphi(n)} - 1) \equiv 0 \mod n$ for every $a \in \mathbb{N}$. By Fact 8.2.2, writing $n = \prod p_i^{\ell_i}$ as product of powers of distinct primes, it suffices to show that this equality holds modulo $p_i^{\ell_i}$ for every i. Distinguish cases: If a and p_i are coprime, by Euler's Theorem [131, Theorem 2.12], $a^{\varphi(p_i^{\ell_i})} \equiv 1 \mod p_i^{\ell_i}$. Since $\varphi(n) = \prod \varphi(p_i^{\ell_i})$ [131, Theorem 2.7], also $a^{\varphi(n)} \equiv 1 \mod p_i^{\ell_i}$. If p_i divides a, then $a^\ell \equiv 0 \mod p_i^{\ell_i}$ as $\ell_i \leq \ell$. Finally, for every graph F, $\hom(F, G^{\times(\varphi(n)+\ell)}) = \hom(F, G)^{\varphi(n)+\ell} \equiv \hom(F, G)^\ell \mod n$ by Equation (2.2). ◁

Towards a contradiction, assume that there exists a non-empty graph $F \in \mathcal{F}$. Let $m := |V(F)|$, and write K_m for the complete graph on m vertices. Then $\hom(F, K_m) > 1$. Define $G := K_m^{\times(\varphi(n)+\ell)}$ and $H := K_m^{\times \ell}$. By Equation (2.2) and $\varphi(n) \geq 1$, it holds that $\hom(F, G) = \hom(F, K_m)^{\varphi(n)+\ell} \neq \hom(F, K_m)^\ell = \hom(F, H)$. Hence, $G \not\equiv_{\mathcal{F}} H$. However, $G \equiv^n_{\mathcal{K}} H$ by Claim 8.2.2a contradicting that $\equiv_{\mathcal{F}}$ and $\equiv^n_{\mathcal{K}}$ coincide. □

8.3 Modular Counting Logic

By Theorem 3.4.4, two graphs are homomorphism indistinguishable over the graphs of treewidth at most $k-1$ if, and only if, they are C^k-equivalent. In this section, the modular analogue of this result is shown: Two graphs are homomorphism indistinguishable over the graphs of treewidth at most $k-1$ modulo a prime p if, and only if, they are equivalent in k-variable modular counting logic $\mathsf{C}^k[p]$.

Definition 8.3.1. Let p be a prime. *Modular counting logic* $\mathsf{C}[p]$ denotes the set of formulas inductively defined as follows:

- for variables x and y, the formulas $x = y$ and Exy are in $\mathsf{C}[p]$,
- if $\varphi, \psi \in \mathsf{C}[p]$, then $\neg\varphi, \varphi \wedge \psi, \varphi \vee \psi \in \mathsf{C}[p]$, and
- if $\varphi(\boldsymbol{x}, y) \in \mathsf{C}[p]$, y is a variable, and $c \in \mathbb{F}_p$, then $\psi(\boldsymbol{x}) := \exists^c y \varphi(\boldsymbol{x}, y)$ is in $\mathsf{C}[p]$.

The semantics of modular counting quantifiers is as expected, e.g. a graph G satisfies a sentence $\exists^c x \varphi(x)$ if the number of distinct $v \in V(G)$ such that $G \models \varphi(v)$ is equal to $c \mod p$. Let $\mathsf{C}^k[p]$ denote the k-variable fragment of $\mathsf{C}[p]$.

This section's main result is the following Theorem 8.3.2.

Theorem 8.3.2. *Let p be a prime and $k \geq 1$. Two simple graphs G and H are homomorphism indistinguishable over all graphs of treewidth at most $k-1$ modulo p if, and only if, G and H are $\mathsf{C}^k[p]$-equivalent.*

The strategy for proving Theorem 8.3.2 is to construct, for every graph F of treewidth $\leq k-1$ and every $m \in \mathbb{F}_p$, a modular counting logic formula with $\leq k$ variables such that a graph satisfies the formula if, and only if, it admits $m \mod p$ many homomorphisms from F. Conversely, counting logic formulas are translated into $\mathbb{F}_p$-linear combinations of graphs of bounded treewidth such that the linear combination of their homomorphism counts in a graph is $1 \mod p$ if, and only if, the formula is satisfied. In this direction, it is crucial that $\mathbb{F}_p$ is a field for an interpolation argument to carry through.

Note that $\mathsf{C}[p]$ is not a syntactical extension of first-order logic as it does not contain the usual quantifiers $\exists$ and $\forall$. By Theorem 8.3.2 and Example 8.1.1, the graphs K_1 and $\overline{K_{p+1}}$ are $\mathsf{C}[p]$-equivalent. As non-isomorphic graphs, they are distinguished by FO. Hence, $\mathsf{C}[p]$ does not semantically extend FO.

Theorem 8.3.2 is implied by the following Theorem 8.3.3. The two directions of the proof are the content of Lemmas 8.3.4 and 8.3.6. Let $k \geq 1$ and $0 \leq \ell \leq k$. Deviating from Definition 4.3.12, write $\mathcal{TW}_k(\ell)$ for the set of ℓ-labelled graphs $\boldsymbol{F} = (F, \boldsymbol{u}) \in \mathcal{G}(\ell)$ admitting a tree decomposition (T, β) of F of width $\leq k-1$ with a vertex $r \in V(T)$ such that $\beta(r) = \{u_1, \dots, u_\ell\}$.

Theorem 8.3.3. *Let p be a prime. Let $k \geq 1$ and $0 \leq \ell \leq k$. Let G and H be simple graphs with $\boldsymbol{v} \in V(G)^\ell$ and $\boldsymbol{w} \in V(H)^\ell$. Then the following are equivalent:*

1. $\boldsymbol{F}_G(\boldsymbol{v}) \equiv \boldsymbol{F}_H(\boldsymbol{w}) \mod p$ *for every* $\boldsymbol{F} \in \mathcal{TW}_k(\ell)$,

2. *for all formulas* $\varphi(x_1, \dots, x_\ell) \in \mathsf{C}^k[p]$ *with* ℓ *free variables, it holds that* $G \models \varphi(\boldsymbol{v})$ *if, and only if,* $H \models \varphi(\boldsymbol{w})$.

The backward implication in Theorem 8.3.3 is proven via the following lemma.

Lemma 8.3.4. *Let* p *be a prime,* $k \geq 1$, *and* $0 \leq \ell \leq k$. *For every* $m \in \mathbb{F}_p$ *and* $\boldsymbol{F} \in \mathcal{TW}_k(\ell)$, *there exists a formula* $\varphi_m(x_1, \dots, x_\ell) \in \mathsf{C}^k[p]$ *such that, for every simple graph* G *and* $\boldsymbol{v} \in V(G)^\ell$, *it holds that* $G \models \varphi_m(\boldsymbol{v})$ *if, and only if,* $\boldsymbol{F}_G(\boldsymbol{v}) \equiv m \mod p$.

Proof. Write $\boldsymbol{F} = (F, \boldsymbol{u})$ and let (T, β) denote a tree decomposition of F of width $\leq k-1$ with a vertex $r \in V(T)$ such that $\beta(r) = \{u_1, \dots, u_\ell\}$. The proof is by induction on the size of the decomposition tree T.

If $|V(T)| = 1$, then all vertices of F are labelled and $\boldsymbol{F}_G(\boldsymbol{v}) \in \{0, 1\}$ for every simple graph G and $\boldsymbol{v} \in V(G)^\ell$ by Observation 3.2.11. If m is neither 0 nor 1, set φ_m to false. If $m = 1$, set

$$\varphi_1(x_1, \dots, x_\ell) := \bigwedge_{\substack{i,j \in [\ell] \\ u_i = u_j}} (x_i = x_j) \wedge \bigwedge_{\substack{i,j \in [\ell] \\ u_i u_j \in E(F)}} E(x_i, x_j).$$

Then $G \models \varphi_1(\boldsymbol{v}) \iff \boldsymbol{F}_G(\boldsymbol{v}) = 1 \iff \boldsymbol{F}_G(\boldsymbol{v}) \equiv 1 \mod p$, where the second equivalence holds since $\boldsymbol{F}_G(\boldsymbol{v}) \in \{0, 1\}$. Finally, set $\varphi_0 := \neg \varphi_1$.

If $|V(T)| \geq 2$, distinguish cases by the number of neighbours of the vertex $r \in V(T)$. If r has unique neighbour $s \in V(T)$, then, by suitably modifying the tree decomposition as in Lemma 2.1.2, we may suppose that $|\beta(s) \setminus \beta(t)| = 1$. Write F' for the subgraph of F induced by $\bigcup_{t \in V(T) \setminus \{r\}} \beta(t)$. Let $\boldsymbol{u}' \in V(F)^{\ell'}$ be such that $\beta(s) = \{u'_1, \dots, u'_{\ell'}\}$ for $\ell' := |\beta(s)|$. To ease notation, suppose that $u'_{\ell'} \notin \beta(t)$. Furthermore, let $R := F[\beta(r)]$ and $\boldsymbol{R} := (R, \boldsymbol{u})$. As graph on at most ℓ vertices, R has a tree decomposition of width $\leq k-1$ with a single bag. By the induction hypothesis, there exist a $\mathsf{C}^k[p]$-formula φ'_m with ℓ' free variables for $\boldsymbol{F}' := (F', \boldsymbol{u}')$ and $m \in \mathbb{F}_p$. Similarly, there exist a $\mathsf{C}^k[p]$-formula χ_m with ℓ free variables for $\boldsymbol{R}$ and $m \in \mathbb{F}_p$. Let $i_1, \dots, i_{\ell'-1} \in [\ell]$ be indices such that $u_{i_j} = u'_j$ for all $j \in [\ell' - 1]$. Then, by Lemmas 3.2.13 and 3.2.14,

$$\boldsymbol{F}_G(\boldsymbol{v}) = \boldsymbol{R}_G(\boldsymbol{v}) \sum_{v' \in V(G)} \boldsymbol{F}'_G(v_{i_1} \dots v_{i_{\ell'-1}} v').$$

For $m \in \mathbb{F}_p$, define $\varphi_m(x_1, \dots, x_\ell)$ as

$$\bigvee_{\substack{m', m'' \in \mathbb{F}_p, \\ m' m'' = m.}} \Big(\chi_{m'}(x_1, \dots, x_\ell) \wedge \bigvee_{\substack{c_1, \dots, c_p \in \mathbb{F}_p, \\ \sum_{j=1}^p j c_j = m''}} \bigwedge_{1 \leq j \leq p} \exists^{c_j} x_\ell \varphi'_j(x_{i_1}, \dots, x_{i_{\ell'-1}}, y) \Big).$$

This formula is as desired. Indeed, if $G \models \varphi_m(\boldsymbol{v})$, then there exist $m', c_1, \dots, c_p \in \mathbb{F}_p$ such that $\boldsymbol{R}_G(\boldsymbol{v}) \equiv m' \mod p$,

$$\left| \left\{ v' \in V(G) \mid \boldsymbol{F}'_G(v_{i_1}, \dots, v_{i_{\ell'-1}}, v') \equiv j \mod p \right\} \right| \equiv c_j \mod p$$

for all $j \in [p]$, and $m' \sum_{j=1}^{p} jc_j = m$. Hence, $\boldsymbol{F}_G(\boldsymbol{v}) \equiv m' \sum_{i=1}^{p} jc_j = m$. The converse is readily verified.

It remains to consider the case when r has multiple neighbours. In this case, $\boldsymbol{F} = \boldsymbol{F}^1 \odot \cdots \odot \boldsymbol{F}^r$ for some graphs $\boldsymbol{F}^1, \dots, \boldsymbol{F}^r$ falling into the case considered above. Let $\varphi^1_m, \dots, \varphi^r_m$ denote the corresponding inductively constructed formulas. Set

$$\varphi_m(x_1, \dots, x_\ell) := \bigvee_{\substack{m_1, \dots, m_r \in \mathbb{F}_p, \\ m_1 \cdots m_r = m.}} \left(\varphi^1_{m_1}(x_1, \dots, x_\ell) \wedge \cdots \wedge \varphi^r_{m_r}(x_1, \dots, x_\ell) \right).$$

Since $\boldsymbol{F}_G(\boldsymbol{v}) = \prod_{i=1}^{r} \boldsymbol{F}^i_G(\boldsymbol{v})$ by Lemma 3.2.13, this formula is as desired. □

Conversely, we construct for every $\mathsf{C}^k[p]$-formula a linear combination of labelled graph of bounded treewidth whose homomorphism counts have the same semantics as the formula. The following auxiliary Lemma 8.3.5 is used in the proof of Lemma 8.3.6. Recall the definition from Section 3.2.3 of the set $\mathbb{F}_p\mathcal{TW}_k(\ell)$ of formal $\mathbb{F}_p$-linear combinations of labelled graphs from $\mathcal{TW}_k(\ell)$.

Lemma 8.3.5. *Let p be a prime, $k \geq 1$, and $0 \leq \ell \leq k$. Let $\boldsymbol{q} \in \mathbb{F}_p\mathcal{TW}_k(\ell)$ and $X_1 \subseteq \mathbb{F}_p$. Then there exists $\boldsymbol{q}' \in \mathbb{F}_p\mathcal{TW}_k(\ell)$ such that, for all simple graphs G with $\boldsymbol{v} \in V(G)^\ell$,*

- *if $\boldsymbol{q}_G(\boldsymbol{v}) \notin X_1$, then $\boldsymbol{q}'_G(\boldsymbol{v}) = 0$,*
- *if $\boldsymbol{q}_G(\boldsymbol{v}) \in X_1$, then $\boldsymbol{q}'_G(\boldsymbol{v}) = 1$.*

Proof. Consider the Lagrange polynomial $p(X) = \sum_{x \in X_1} \prod_{y \in \mathbb{F}_p \setminus \{x\}} \frac{X-y}{x-y} \in \mathbb{F}_p[X]$. Observe that $p(z) = 1$ if $z \in X_1$ and $p(z) = 0$ if $z \notin X_1$. Define $\boldsymbol{q}' := p(\boldsymbol{q})$ by interpreting multiplication in $\mathbb{F}_p[X]$ as gluing product in $\mathcal{TW}_k(\ell)$. For $\ell = 0$, the gluing product degenerates to the disjoint union. By Lemma 3.2.13, it holds that $\boldsymbol{q}'_G(\boldsymbol{v}) = p(\boldsymbol{q})_G(\boldsymbol{v}) = p(\boldsymbol{q}_G(\boldsymbol{v}))$, as desired. □

The remaining ingredient for the proof of Theorem 8.3.3 is Lemma 8.3.6.

Lemma 8.3.6. *Let p be a prime, $k \geq 1$, and $0 \leq \ell \leq k$. For every $\varphi(x_1, \dots, x_\ell) \in \mathsf{C}^k[p]$, there exists a $\boldsymbol{q} \in \mathbb{F}_p\mathcal{TW}_k(\ell)$ such that $\boldsymbol{q}$* models *φ, i.e. for all simple graphs G with $\boldsymbol{v} \in V(G)^\ell$,*

- *if $G \not\models \varphi(\boldsymbol{v})$, then $\boldsymbol{q}_G(\boldsymbol{v}) = 0$,*
- *if $G \models \varphi(\boldsymbol{v})$, then $\boldsymbol{q}_G(\boldsymbol{v}) = 1$.*

Proof. By induction on the structure of φ.

- If $\varphi = (x_1 = x_2)$, then the graph $\boldsymbol{I} = (I, (1,1)) \in \mathcal{TW}_k(2)$ with $V(I) = \{1\}$ and $E(I) = \emptyset$ models φ.
- If $\varphi = E(x_1, x_2)$, then the graph $\boldsymbol{A} = (A, (1,2)) \in \mathcal{TW}_k(2)$ with $V(A) = \{1,2\}$ and $E(A) = \{1,2\}$ models φ.
- If $\varphi = \varphi_1 \wedge \varphi_2$, let $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathbb{F}_p\mathcal{TW}^k$ denote the elements modelling φ_1 and φ_2 respectively. Their gluing product $\boldsymbol{t}_1 \odot \boldsymbol{t}_2$ models φ.

- If $\varphi = \neg\varphi_1$ and $\boldsymbol{t}_1$ is as above, then $\boldsymbol{J} - \boldsymbol{t}_1$ models φ. Here, $\boldsymbol{J} = (J, (1, \dots, \ell)) \in \mathcal{TW}_k(\ell)$ with $V(J) = [\ell]$ and $E(J) = \emptyset$ where ℓ denotes the number of free variables in φ. If $\ell = 0$, then J is the empty graph.
- If $\varphi = \varphi_1 \vee \varphi_2$, then apply the two cases above to $\neg(\neg\varphi_1 \wedge \neg\varphi_2)$.
- If $\varphi(x_1, \dots, x_{\ell-1}) = (\exists^m x_\ell)\psi(x_1, \dots, x_\ell)$, then let $\boldsymbol{t} = \sum \alpha_i \boldsymbol{F}^i \in \mathcal{TW}_k(\ell)$ denote the element modelling ψ. For every $\boldsymbol{F}^i = (F^i, \boldsymbol{v}^i)$, define $\boldsymbol{K}^i = (F^i, v^i_1 \dots v^i_{\ell-1})$. Observe that, for every simple graph G with $\boldsymbol{v} \in V(G)^{\ell-1}$,
$$\boldsymbol{K}^i_G(\boldsymbol{v}) = \sum_{v' \in V(G)} \boldsymbol{F}^i_G(v_1 \dots v_{\ell-1} v').$$
Let $\boldsymbol{q} \coloneqq \sum \alpha_i \boldsymbol{K}^i \in \mathbb{F}_p\mathcal{TW}_k(\ell)$. By induction,
$$\begin{aligned}\boldsymbol{q}_G(\boldsymbol{v}) &\equiv \sum_{v' \in V(G)} \boldsymbol{t}_G(v_1 \dots v_{\ell-1} v') \\ &\equiv |\{v' \in V(G) \mid G \models \psi(v_1, \dots, v_{\ell-1}, v')\}| \mod p.\end{aligned}$$
The desired graph can now be constructed via Lemma 8.3.5. □

8.4 Further Directions

Much less is known about modular homomorphism indistinguishability than about non-modular homomorphism indistinguishability. We conclude this chapter with a collection of questions.

The first question asks to generalise Theorem 8.1.2. A resolution might help to classify the complexity of counting graph colourings modulo prime powers. For prime moduli, the complexity of these problems was recently completely described in [35] building on Theorem 8.1.2.

Question 8.4.1. *What does it mean for two graphs to be homomorphism indistinguishable over all graphs modulo a prime power?*

The second question concerns a generalisation of Theorems 4.0.1 and 2.6.8 on the Sherali–Adams hierarchy $\mathrm{SA}^k(G, H)$ for graph isomorphism, cf. Definition 2.6.7. Since a graph isomorphism amounts to a $\{0, 1\}$-solution to $\mathrm{QP}(G, H)$ and thus also to $\mathrm{SA}^k(G, H)$, the feasibility of $\mathrm{SA}^k(G, H)$ over $\mathbb{F}_p$ or $\mathbb{Z}$ is a graph isomorphism relaxation. Since linear equations over $\mathbb{F}_p$ and $\mathbb{Z}$ can be solved in polynomial time [98], they are polynomial-time graph isomorphism relaxation.

By [23, 24], there is no $k \geq 1$ such that all pairs of non-isomorphic graphs can be distinguished by checking the feasibility of $\mathrm{SA}^k(G, H)$ over $\mathbb{Z}$. No characterisations of this graph isomorphism relaxation are known. Considering $\mathrm{SA}^k(G, H)$ over $\mathbb{F}_p$ might be a pathway to understanding $\mathrm{SA}^k(G, H)$ over $\mathbb{Z}$. Indeed, being a linear system of equations, $\mathrm{SA}^k(G, H)$ is feasible over $\mathbb{Z}$ if, and only if, it is feasible modulo every prime power, cf. [160, p. 51] and Fact 8.2.2.

Question 8.4.2. *Can the feasibility of* $\mathrm{SA}^k(G, H)$ *over* $\mathbb{F}_p$ *be characterised in terms of modular homomorphism indistinguishability? Can the feasibility of* $\mathrm{SA}^k(G, H)$ *over* $\mathbb{Z}$ *be characterised in terms of homomorphism indistinguishability?*

Modular and non-modular homomorphism indistinguishability both count homomorphisms, only over different semirings, cf. [45]. In the first case, homomorphisms are counted in $\mathbb{Z}/n\mathbb{Z}$ for some $n \geq 2$, whereas in the second case, the natural numbers $\mathbb{N}$ are used for counting. A semiring with intimate connections to logic is the Boolean semiring. 'Counting' homomorphisms from F to G over the Boolean semiring amounts to deciding whether there exists a homomorphism $F \to G$. Thus, two graphs G and H are Boolean homomorphism indistinguishable over all graph if, and only if, they are homomorphically equivalent, i.e. there exist homomorphisms $G \to H$ and $H \to G$. Boolean homomorphism indistinguishability was studied in [14, 38]. We propose the following questions entailing Question 7.4.4.

Question 8.4.3. *Which graph isomorphism relaxations are Boolean homomorphism indistinguishability relations? Given a semiring* $\mathbb{S}$*, which graph isomorphism relaxations are homomorphism indistinguishability relations with respect to* $\mathbb{S}$*?*

9 Complexity of Homomorphism Indistinguishability

From a computational perspective, the central question on homomorphism indistinguishability concerns the complexity and computability of the following decision problem for a fixed graph class $\mathcal{F}$ [150, Question 9]:

> HomInd($\mathcal{F}$)
> **Input** Simple graphs G and H.
> **Question** Are G and H homomorphism indistinguishable over $\mathcal{F}$?

The graphs G and H may be arbitrary graphs and do not necessarily have to be in $\mathcal{F}$. Typically, the graph class $\mathcal{F}$ is infinite. Thus, the trivial approach to HomInd($\mathcal{F}$) of checking whether G and H have the same number of homomorphisms from every $F \in \mathcal{F}$ does not even render HomInd($\mathcal{F}$) decidable.

In Section 3.2.1, building on results of Böker, Chen, Grohe, & Rattan [33], esoteric graph classes $\mathcal{F}$ are constructed for which HomInd($\mathcal{F}$) is arbitrarily hard. For natural, e.g. minor-closed, graph classes $\mathcal{F}$, the understanding of the problems HomInd($\mathcal{F}$) is limited to the short list of examples given as Table 9.1.

The results in Table 9.1 illustrate that the complexity of HomInd($\mathcal{F}$) is not monotone. That is, it does not hold that if $\mathcal{F}_1 \subseteq \mathcal{F}_2$, then HomInd($\mathcal{F}_1$) is at most as hard as HomInd($\mathcal{F}_2$). For example, despite that $\mathcal{TW}_2 \subseteq \mathcal{P} \subseteq \mathcal{G}$, deciding homomorphism indistinguishability over $\mathcal{TW}_2$, the class $\mathcal{P}$ of all planar graphs, and the class $\mathcal{G}$ of all graphs is polynomial-time, undecidable, and quasi-polynomial-time, respectively. Furthermore, although HomInd($\mathcal{TW}_k$) is in polynomial time for every k by Theorem 3.4.3, there are infinitely many minor-closed graph classes $\mathcal{F}$ of bounded treewidth, e.g. the classes of k-outerplanar graphs, for which HomInd($\mathcal{F}$) could yet be undecidable. The main result of this chapter shows that this is not the case: HomInd($\mathcal{F}$) is in randomised polynomial time for every minor-closed graph class $\mathcal{F}$ of bounded treewidth.

Theorem 9.0.1. *Let $k \geq 1$. If $\mathcal{F}$ is a k-recognisable class of graphs of treewidth at most $k-1$, then* HomInd($\mathcal{F}$) *is in* coRP.

Spelled out, Theorem 9.0.1 asserts that there exists a randomised algorithm for HomInd($\mathcal{F}$) which always runs in polynomial time, accepts all YES-instances and incorrectly accepts NO-instances with probability less than one half.[13] Recognisability

[13] See [9, Section 7.3] for a formal definition of the complexity class coRP.

Graph Class	**Complexity**	**Reference**
All graphs Bipartite graphs	in quasi-polynomial time	[17] Theorem 9.4.6
Cycles	in polynomial time	Theorem 3.3.2
Treewidth $\leq k$	PTIME-complete	Theorem 3.4.3, [77]
Pathwidth $\leq k$	in polynomial time	Theorem 4.0.1
Treedepth $\leq k$	in polynomial time	Theorem 4.0.2
Recognisable graph class of bounded treewidth	in coRP	Theorem 9.0.1
Recognisable graph class of bounded pathwidth	in polynomial time	Theorem 9.0.2
Complete graphs	$C_=P$-complete	Theorem 9.5.5
Planar graphs	undecidable	[124, 15]

Table 9.1: Complexity of HomInd($\mathcal{F}$) for natural graph classes $\mathcal{F}$.

is a fairly general property that arises in the context of Courcelle's Theorem [46], cf. Definition 9.1.5. Courcelle showed that every graph class definable in counting monadic second-order logic CMSO_2 is recognisable. This subsumes graph classes defined by finitely many forbidden (induced) subgraphs and minors, and by the Robertson–Seymour Theorem [155], all minor-closed graph classes. As a concrete application, we show in Theorems 9.3.1 and 9.3.2 that the exact feasibility of the Lasserre semidefinite programming hierarchy for graph isomorphism can be decided in (randomised) polynomial time.

The proof of Theorem 9.0.1 combines Courcelle's graph algebras [48] with the homomorphism tensors from Section 3.2. Graph algebras comprise labelled graphs and operations on them such as series and parallel composition. Homomorphism tensors keep track of homomorphism counts of labelled graphs. We show that recognisability and bounded treewidth guarantee that homomorphism tensors yield finite-dimensional representations of suitable graph algebras which certify homomorphism indistinguishability and are efficiently computable. The algorithm in Theorem 9.0.1 is randomised as it employs arithmetic modulo random primes to deal with integers which would otherwise grow to exponential size. For graph classes of bounded pathwidth, this issue can be avoided:

Theorem 9.0.2. *Let $k \geq 1$. If $\mathcal{F}$ is a k-recognisable class of graphs of pathwidth at most $k-1$, then* HomInd($\mathcal{F}$) *is in polynomial time.*

In Section 9.4, Theorems 9.0.1 and 9.0.2 are counterpointed by various hardness results. In particular, we show in Theorem 9.4.1 that deciding whether two graphs are not distinguished by the k-dimensional Weisfeiler–Leman algorithm is coNP-hard when k is part of the input. Finally, we conjecture a trichotomy for the complexity of HomInd($\mathcal{F}$) for minor-closed $\mathcal{F}$ in Conjecture 9.6.1.

Chapter Outline. The first four sections are devoted proving Theorems 9.0.1 and 9.0.2. In Section 9.1, the problem HomInd($\mathcal{F}$) is shown to be decidable for every recognisable graph class $\mathcal{F}$ of bounded treewidth and the notion of witness functions is introduced. In Section 9.2, this result is strengthened to yield Theorems 9.0.1 and 9.0.2. In Section 9.3, the exact feasibility of the Lasserre semidefinite programming relaxation for graph isomorphism is shown to be decidable in (randomised) polynomial time. In Section 9.4, various lower bounds on the complex of homomorphism indistinguishability are derived. Section 9.5 features a discussion of the complexity of deciding homomorphism indistinguishability over essentially profinite graph classes. We conclude in Section 9.6 by conjecturing a trichotomy for the complexity of homomorphism indistinguishability over minor-closed graph classes.

Except for Sections 9.3.2 and 9.5, the material in this section was previously published in [164, 163]. Section 9.3.2 is joint work with David E. Roberson and was published in [151, 152]. Section 9.5 was published in [162, 164, 163, 165].

9.1 Decidability

For an infinite graph class $\mathcal{F}$, the problem HomInd($\mathcal{F}$) is a priori not even decidable. This is because, for graphs G and H such that $G \not\equiv_{\mathcal{F}} H$, there is, in general, no bound on the size of the smallest graph F such that $\hom(F, G) \neq \hom(F, H)$. Such a graph F *witnesses* that $G \not\equiv_{\mathcal{F}} H$. The size of witnesses is the subject of Section 9.1.1. In Section 9.1.2, we show that HomInd($\mathcal{F}$) is decidable for every recognisable graph class of bounded treewidth by bounding the size of these witnesses.

9.1.1 Witness Functions

In this section, we introduce witness functions as a combinatorial tool for studying the complexity of homomorphism indistinguishability. Recall that $\mathcal{F}_{\leq \ell} := \{F \in \mathcal{F} \mid |V(F)| \leq \ell\}$ for a graph class $\mathcal{F}$ and $\ell \in \mathbb{N}$.

Definition 9.1.1. Let $\mathcal{F}$ be a graph class. A *witness function* for $\mathcal{F}$ is a function $f \colon \mathbb{N} \to \mathbb{N}$ such that $G \equiv_{\mathcal{F}} H$ if, and only if, $G \equiv_{\mathcal{F}_{\leq f(n)}} H$ for all simple graphs G and H on at most n vertices.

The existence of a computable witness function f for a decidable graph class $\mathcal{F}$ implies the decidability of HomInd($\mathcal{F}$). Indeed, given input graphs G and H on at most n vertices, one can enumerate all graphs $F \in \mathcal{F}$ on up to $f(n)$ vertices and check whether $\hom(F, G) = \hom(F, H)$.

Graph Class	Witness Function $f(n)$	Reference
All graphs	n	Theorem 3.1.1
Bipartite graphs	$2n$	Corollary 3.1.3
Cycles	n	Theorem 3.3.2
Paths	$2n$	Corollary 4.2.3
Stars	n	Theorem 3.3.1
Complete graphs	n	Lemma 9.1.3
Treedepth $\leq k$	$4kn^k$	Theorem 4.3.22
Pathwidth $\leq k-1$	$2n^k + k - 1$	Corollary 9.1.19
k-recognisable graph class of pathwidth $\leq k-1$	$2Cn^k + k - 1$ for k-recognisability index C	Theorem 9.1.7
Trees	$(2n)^{n+1}$	Corollary 4.2.4
Treewidth $\leq k-1$	$\max\{k^{2n^k}, 2n^k\}$	Corollary 9.1.18
k-recognisable graph class of treewidth $\leq k-1$	$\max\{k^{2Cn^k}, 2Cn^k\}$ for k-recognisability index C	Theorem 9.1.6
Planar graphs	not computable	[124, p. 663]

Table 9.2: Known witness functions grouped by order of magnitude.

For example, the identity function $n \mapsto n$ is a witness function for the class of complete graphs $\mathcal{K}$. This is because, if G and H are simple graphs on at most n vertices, then $\hom(K_m, G) = 0 = \hom(K_m, H)$ for all $m > n$. Hence, if $G \not\equiv_{\mathcal{K}} H$, then there exists an $m \leq n$ such that $\hom(K_m, G) \neq \hom(K_m, H)$.

Throughout Chapters 3 and 5, we derived witness functions for many graph classes. These are listed in Table 9.2.

In the reminder of this section, we collect some general properties of witness functions. Witness functions always exists. This is because there are only finitely many equivalence classes under $\equiv_{\mathcal{F}}$ on simple graphs on at most n vertices.

Lemma 9.1.2. *For every graph class $\mathcal{F}$, there exists a witness function $f\colon \mathbb{N} \to \mathbb{N}$.*

Proof. Let $n \in \mathbb{N}$. For every pair of simple graphs G and H on at most n vertices such that $G \not\equiv_{\mathcal{F}} H$, there exists a graph F such that $\hom(F, G) \neq \hom(F, H)$. There are finitely many such pairs. Define $f(n)$ as the maximum number of vertices of all such graphs F. Then f is a witness function for $\mathcal{F}$. □

Note that Lemma 9.1.2 is non-constructive and gives no indication on the growth of the witness function. Under certain assumptions, we can do better. The following lemma generalises the example above involving the class of complete graphs $\mathcal{K}$. Graph classes to which the lemma applies are necessarily essentially profinite, cf. Lemma 6.5.12 A function $f\colon \mathbb{N} \to \mathbb{N}$ is *non-decreasing* if $n \leq m$ implies that $f(n) \leq f(m)$ for all $n, m \in \mathbb{N}$. Recall that $\chi(F)$ denotes the chromatic number of F.

Lemma 9.1.3. *Let $\mathcal{F}$ be a graph class. Every non-decreasing function $f \colon \mathbb{N} \to \mathbb{N}$ such that $|V(F)| \leq f(\chi(F))$ for all $F \in \mathcal{F}$ is a witness function for $\mathcal{F}$.*

Proof. Let G and H be simple graphs on at most n vertices. If $F \in \mathcal{F}$ has chromatic number greater than n, then it does not admit a homomorphism into G or H. Hence, if $G \not\equiv_{\mathcal{F}} H$, then there exists $F \in \mathcal{F}$ with $\chi(F) \leq n$ such that $\hom(F, G) \neq \hom(F, H)$. By assumption, $|V(F)| \leq f(\chi(F)) \leq f(n)$. Hence, f is a witness function for $\mathcal{F}$. □

Giving lower bounds on witness functions for a graph class $\mathcal{F}$ sheds light on the complexity of the problem HomInd($\mathcal{F}$) bypassing the intricacies of computation. Finding the slowest growing witness function for a graph class $\mathcal{F}$ is a purely combinatorial problem, which may be more approachable than proving hardness of the problem HomInd($\mathcal{F}$). We conclude with the following lower bounds, which are proven using CFI graphs.

Theorem 9.1.4. *Let $f \colon \mathbb{N} \to \mathbb{N}$ be a witness function for a graph class $\mathcal{F}$.*

1. *If $\mathcal{F}$ contains arbitrarily large connected graphs of maximum degree at most $\Delta \in \mathbb{N}$, then there exist infinitely many numbers $n \in \mathbb{N}$ such that $f(n) \geq 2^{1-\Delta} n$.*
2. *If $\mathcal{F}$ contains arbitrarily large connected graphs, then, for every $\varepsilon > 0$, there exist infinitely many numbers $n \in \mathbb{N}$ such that $f(n) \geq (\log n)^{1-\varepsilon}$.*

Proof. For the first claim, let $G \in \mathcal{F}$ be a N-vertex connected graph of maximum degree at most Δ. Consider the CFI graphs G_0 and G_1 over G as defined in Definition 6.3.6. By Corollary 6.3.5, $\hom(G, G_0) \neq \hom(G, G_0)$ and hence $G_0 \not\equiv_{\mathcal{F}} G_1$. By Theorem 6.3.8, for every simple graph F, it holds that $\hom(F, G_0) \neq \hom(F, G_1)$ if, and only if, there is a weak oddomorphism from F to G. Weak oddomorphisms are necessarily surjective on vertices. Hence, G_0 and G_1 are homomorphism indistinguishable over all graphs on at most $N - 1$ vertices. In particular, $G_0 \equiv_{\mathcal{F}_{\leq N-1}} G_1$. It follows that $f(n) \geq N$ for $n := |V(G_0)| = |V(G_1)|$. It holds that $n = \sum_{v \in V(G)} 2^{\deg(v)-1} \leq N 2^{\Delta - 1}$. Hence, $f(n) \geq N \geq 2^{1-\Delta} n$.

For the second claim, fix $\varepsilon > 0$. Let $G \in \mathcal{F}$ be a connected graphs on N vertices for some $N \in \mathbb{N}$ large enough such that $N \log N \leq N^{1+\varepsilon}$. Let as above $n := |V(G_0)| = |V(G_1)|$. Then $n = \sum_{v \in V(G)} 2^{\deg(v)-1} \leq 2^{N \log N} \leq 2^{N^{1+\varepsilon}}$. It follows that $f(n) \geq N \geq (\log n)^{1-\varepsilon}$. □

The first assertion of Theorem 9.1.4 yields a linear lower bound for most of the graph classes featured in Table 9.2. Hence, most of the linear witness functions listed there are asymptotically optimal, cf. Corollary 9.5.4.

9.1.2 Witness Functions for Recognisable Graph Classes of Bounded Treewidth

In this section, we show that every recognisable graph class of bounded treewidth admits an exponential witness function. For recognisable graph classes of bounded pathwidth, we derive polynomial witness functions. Recognisability is a property of a class of unlabelled graphs which is shared by most named graph classes. We consider the following definitions:

Let $k, \ell \in \mathbb{N}$. A (k,ℓ)-bilabelled graph $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v})$ is *distinctly* (k,ℓ)*-bilabelled* if $u_i \neq u_j$ for all $1 \leq i < j \leq k$ and $v_i \neq v_j$ for all $1 \leq i < j \leq \ell$. Note that it might be that $u_i = v_j$ for some $i \in [k]$ and $j \in [\ell]$. Write $\mathcal{D}(k,\ell) \subseteq \mathcal{G}(k,\ell)$ for the class of distinctly (k,ℓ)-bilabelled graphs and $\mathcal{D}(k) \subseteq \mathcal{G}(k)$ for the class of distinctly k-labelled graphs, i.e. $\mathcal{D}(k) = \mathcal{D}(k,0)$. Note that the operations from Definitions 3.2.5 and 3.2.6 restrict to operations on distinctly bilabelled graphs.

Definition 9.1.5 ([29])**.** Let $k \geq 1$. For class of unlabelled graphs $\mathcal{F}$, define the equivalence relation $\sim^k_{\mathcal{F}}$ on the class of distinctly k-labelled graphs $\mathcal{D}(k)$ by letting $\boldsymbol{F}_1 \sim^k_{\mathcal{F}} \boldsymbol{F}_2$ if, and only if, for all $\boldsymbol{K} \in \mathcal{D}(k)$, it holds that

$$\operatorname{soe}(\boldsymbol{K} \odot \boldsymbol{F}_1) \in \mathcal{F} \iff \operatorname{soe}(\boldsymbol{K} \odot \boldsymbol{F}_2) \in \mathcal{F}.$$

The class $\mathcal{F}$ is *k-recognisable* if $\sim^k_{\mathcal{F}}$ has finitely many equivalence classes. The number of classes of $\sim^k_{\mathcal{F}}$ is the *k-recognisability index* of $\mathcal{F}$.

The main results of this section can now be stated as Theorems 9.1.6 and 9.1.7:

Theorem 9.1.6. *Let $k \geq 1$. Let $\mathcal{F}$ be a graph class of treewidth $\leq k-1$ with k-recognisability index C. For all simple graphs G and H on at most n vertices, with $f_{k,C}(n) := \max\{k^{2Cn^k}, 2Cn^k\}$,*

$$G \equiv_{\mathcal{F}} H \iff G \equiv_{\mathcal{F}_{\leq f_{k,C}(n)}} H.$$

Theorem 9.1.7. *Let $k \geq 1$. Let $\mathcal{F}$ be a graph class of pathwidth $\leq k-1$ with k-recognisability index C. For all simple graphs G and H on at most n vertices, with $f_{k,C}(n) := 2Cn^k + k - 1$,*

$$G \equiv_{\mathcal{F}} H \iff G \equiv_{\mathcal{F}_{\leq f_{k,C}(n)}} H.$$

Before we prove Theorems 9.1.6 and 9.1.7, we build some intuition for the notion of recognisability and discuss its connections to counting monadic second-order logic CMSO_2, cf. Section 2.2.3. To parse Definition 9.1.5, first recall that $\boldsymbol{K} \odot \boldsymbol{F}_1$ is the k-labelled graph obtained by gluing $\boldsymbol{K}$ and $\boldsymbol{F}_1$ together at their labelled vertices. The soe-operator drops the labels, yielding unlabelled graphs. Intuitively, $\boldsymbol{F}_1 \sim^k_{\mathcal{F}} \boldsymbol{F}_2$ if, and only if, both or neither of their underlying unlabelled graphs are in $\mathcal{F}$ and the positions of the labels in $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ is equivalent with respect to membership in $\mathcal{F}$. This intuition is made more concrete in the following example:

Example 9.1.8. The class $\mathcal{W}$ of all paths is 1-recognisable. Its 1-recognisability index is 4. The equivalence classes are described by the representatives in Figure 9.1.

(a) $\boldsymbol{1}$. **(b)** $\boldsymbol{P}$. **(c)** $\boldsymbol{Q}$. **(d)** $\boldsymbol{C}$.

Figure 9.1: Representatives for $\sim^1_{\mathcal{W}}$ from Example 9.1.8.

Proof. To show that the labelled graphs in Figure 9.1 cover all equivalence classes, let $\boldsymbol{F} = (F, u) \in \mathcal{D}(1)$ be arbitrary. If F is not a path, then $\boldsymbol{F} \sim^1_{\mathcal{W}} \boldsymbol{C}$. Indeed, for every $\boldsymbol{K} \in \mathcal{D}(1)$, the graph F is a subgraph of $\operatorname{soe}(\boldsymbol{K} \odot \boldsymbol{F})$. Hence, regardless of $\boldsymbol{K}$, both $\operatorname{soe}(\boldsymbol{K} \odot \boldsymbol{F})$ and $\operatorname{soe}(\boldsymbol{K} \odot \boldsymbol{C})$ are not paths. If F is a path, then $\boldsymbol{F}$ and $\boldsymbol{1}$, $\boldsymbol{P}$, or $\boldsymbol{Q}$ are equivalent depending on whether the degree of u is 0, 1, or 2.

To show that the representatives in Figure 9.1 are in distinct classes, observe for example that $\operatorname{soe}(\boldsymbol{P} \odot \boldsymbol{P}) \in \mathcal{W}$ while $\operatorname{soe}(\boldsymbol{P} \odot \boldsymbol{Q}) \notin \mathcal{W}$, thus $\boldsymbol{P} \not\sim^1_{\mathcal{W}} \boldsymbol{Q}$. Similarly, $\operatorname{soe}(\boldsymbol{1} \odot \boldsymbol{Q}) \in \mathcal{W}$ whereas $\operatorname{soe}(\boldsymbol{P} \odot \boldsymbol{Q}) \notin \mathcal{W}$, thus $\boldsymbol{1} \not\sim^1_{\mathcal{W}} \boldsymbol{P}$. □

A more involved example is the following. Analogously, one may argue that every class defined by forbidden minors is recognisable.

Example 9.1.9. Let $\mathcal{F}$ be the family of H-subgraph-free graphs for some graph H. Then $\mathcal{F}$ is k-recognisable for every $k \geq 1$.

Proof. Suppose that H has m vertices. For a distinctly k-labelled graph $\boldsymbol{F} = (F, \boldsymbol{u})$, consider the set $\mathcal{H}(\boldsymbol{F})$ of (isomorphism types of) distinctly k-labelled graphs $\boldsymbol{F}' = (F', \boldsymbol{u})$ where F' is a subgraph of F such that $V(F') \supseteq \{u_1, \dots, u_k\}$ and F' has at most $k + m$ vertices. Clearly, there are only finitely many possible sets $\mathcal{H}(\boldsymbol{F})$. Furthermore, if $\mathcal{H}(\boldsymbol{F}_1) = \mathcal{H}(\boldsymbol{F}_2)$, then $\boldsymbol{F}_1 \sim^k_{\mathcal{F}} \boldsymbol{F}_2$. Indeed, if $\boldsymbol{K} \in \mathcal{D}(k)$ is such that $\operatorname{soe}(\boldsymbol{K} \odot \boldsymbol{F}_1)$ contains H as a subgraph, then so does $\operatorname{soe}(\boldsymbol{K} \odot \boldsymbol{F}_2)$ since $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ contain the same subgraphs on $k + m$ vertices containing their labelled vertices. □

Courcelle [46] proved that every CMSO_2-definable graph class is *recognisable*, i.e. it is k-recognisable for every $k \in \mathbb{N}$. Conversely, Bojańczyk & Pilipczuk [29] proved that if a recognisable class $\mathcal{F}$ has bounded treewidth, then it is CMSO_2-definable. Furthermore, they conjecture that k-recognisability is a sufficient condition for a graph class of treewidth at most $k - 1$ to be CMSO_2-definable.

We choose to work with distinctly k-labelled graphs in Definition 9.1.5 in order to be aligned with [46, 29]. The subsequent arguments also work when $\mathcal{D}(k)$ is replaced with $\mathcal{G}(k)$ or $\mathcal{TW}(k)$, cf. [62, Definition 12.7.4].

Labelled Graphs with Tree Decompositions of Bounded Width and Branching

Building on Lemma 2.1.2, we show in the following Lemma 9.1.10 that every tree decomposition can be rearranged such that each vertex in the decomposition tree has bounded degree. This will be needed in Lemma 9.1.12, where it is proven that

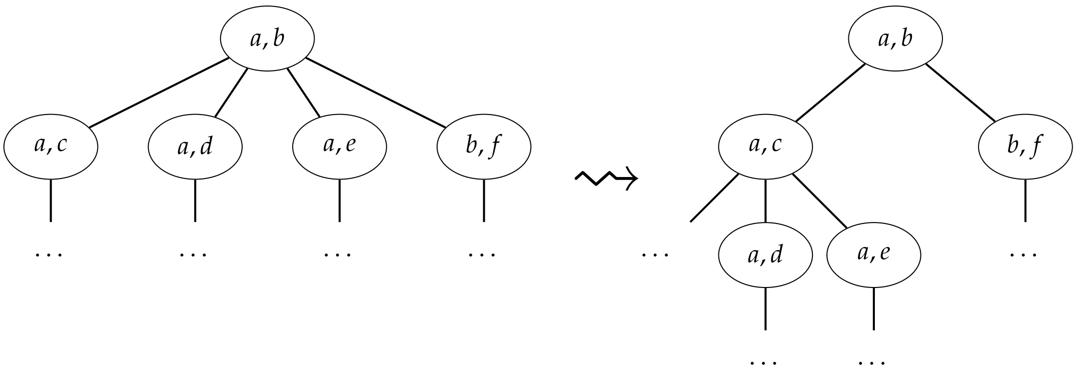


Figure 9.2: Example for one step of the construction in the proof of Lemma 9.1.10 for $k = 2$. The labels of the vertices indicate the contents of the bags of the tree decomposition.

the depth of the decomposition tree gives a bound on the number of vertices in the decomposed graph.

Lemma 9.1.10. *Let $k \geq 1$ and F be a graph such that $\operatorname{tw}(F) \leq k-1$ and $|V(F)| \geq k$. Then F admits tree decomposition (T, β) such that*

1. $|\beta(t)| = k$ *for all* $t \in V(T)$,
2. $|\beta(s) \cap \beta(t)| = k-1$ *for all* $st \in E(T)$,
3. *there exists a vertex $r \in V(T)$ such that the out-degree of every vertex in the rooted tree (T, r) is at most k.*

Proof. By Lemma 2.1.2, there exists a tree decomposition (T, β) of F satisfying the first two assertions. Pick a root $r \in V(T)$ arbitrarily. To ensure that the last property holds, the tree decomposition is modified recursively as follows. See Figure 9.2 for an example.

By merging vertices, it can be ensured that no two children of r carry the same bag, i.e. that there exist no two children $s_1 \neq s_2$ of r such that $\beta(s_1) = \beta(s_2)$.

For every $v \in \beta(r)$, let $C(v)$ denote the set of all children t of r such that $\beta(r) \setminus \beta(t) = \{v\}$, i.e. $C(v)$ is the set of all children of r whose bags do not contain the vertex v. The collection $C(v)$ for $v \in \beta(r)$ is a partition of the children of r in at most k parts.

Note that, for two distinct children $t_1 \neq t_2$ in the same part $C(v)$, it holds that $|\beta(t_1) \cap \beta(t_2)| = k-1$. Rewire the children of r as follows: For every $v \in \beta(r)$ with $C(v) \neq \emptyset$, pick $t \in C(v)$, make t a child of r and all other elements of $C(v)$ children of t. The vertex r now has at most k children and the new tree decomposition still satisfies the first two assertions. Proceed by processing the children of r. □

Inspired by Lemma 9.1.10, we refine Definition 4.3.12 of $\mathcal{TW}(k)$ as follows:

Observe that $\mathcal{TW}(k) \subseteq \mathcal{D}(k)$ and recall that the depth of a rooted tree (T, r) is the maximal number of vertices on any path from r to a leaf.

Definition 9.1.11. Let $k, d \geq 1$. Define $\mathcal{TW}_d(k) \subseteq \mathcal{TW}(k)$ as the class of distinctly k-labelled graphs $\boldsymbol{F} = (F, \boldsymbol{u})$ such that there exists a tree decomposition (T, β) of F and a vertex $r \in V(T)$ such that

1. $\beta(r) = \{u_1, \dots, u_k\}$ and
2. $|\beta(s)| = k$ for all $s \in V(T)$ and $|\beta(s) \cap \beta(t)| = k - 1$ for all $st \in E(T)$,
3. (T, r) is of depth at most d,
4. every vertex in (T, r) has out-degree at most k.

By Lemma 9.1.10, $\bigcup_{d \geq 1} \mathcal{TW}_d(k) = \mathcal{TW}(k)$. Every graph in $\mathcal{TW}_d(k)$ has at least k vertices. Conversely, Definition 9.1.11 permits the following upper bound on the size of the graphs in $\mathcal{TW}_d(k)$ in terms of d and k.

Lemma 9.1.12. *Let $k, d \geq 1$. Every $\boldsymbol{F} \in \mathcal{TW}_d(k)$ has at most* $\max\{k^d, d\}$ *vertices.*

Proof. Let (T, β) and $r \in V(T)$ be as in Definition 9.1.11. If $k = 1$, then every vertex in (T, r) has out-degree 1 and F at most d vertices.

Suppose that $k \geq 2$. The proof is by induction on the depth d of the rooted tree (T, r). If $d = 1$, then T contains only a single vertex and $\boldsymbol{F}$ has at most k vertices.

For the inductive step, let $\boldsymbol{F}$ be of depth $d \geq 2$. If r has only a single neighbour s, then $S := T - r$ is such that (S, s) is of depth $d - 1$. By the inductive hypothesis, $\left|\bigcup_{s \in V(S)} \beta(s)\right| \leq k^{d-1}$. Furthermore, $\left|\bigcup_{t \in V(T)} \beta(t) \setminus \bigcup_{s \in V(S)} \beta(s)\right| = 1$. Hence, $\boldsymbol{F}$ has at most $k^{d-1} + 1 \leq k^d$ many vertices.

If r has multiple neighbours, observe that, due to Lemma 9.1.10, every vertex in $\beta(r)$ is also in $\beta(s)$ for some neighbour s of r. Hence, the number of vertices in $\boldsymbol{F}$ is bounded by the number of vertices covered by the subtrees of $T - r$ rooted in s. Thus, $\boldsymbol{F}$ has at most $k^{d-1} \cdot k \leq k^d$ many vertices. □

The set-up introduced above can be adapted for graphs of bounded pathwidth. The following Definition 9.1.13 refines Definition 4.3.3.

Definition 9.1.13. For $k, d \geq 1$, write $\mathcal{PW}_d(k) \subseteq \mathcal{D}(k)$ for the class of all distinctly k-labelled graphs $\boldsymbol{F} = (F, \boldsymbol{u})$ such that there exists a path decomposition (P, β) of F where

1. there exists a vertex $r \in V(P)$ of degree at most 1 such that $\beta(r) = \{u_1, \dots, u_k\}$,
2. $|\beta(t)| = k$ for all $t \in V(P)$ and $|\beta(s) \cap \beta(t)| = k - 1$ for all $st \in E(P)$,
3. $|V(P)| \leq d$.

Let $\mathcal{PW}(k) := \bigcup_{d \geq 1} \mathcal{PW}_d(k)$.

Note that $\mathcal{PW}(k) = \mathcal{PW}(k, k)\boldsymbol{1}$ with $\mathcal{PW}(k, k)$ as defined in Definition 4.3.3. Analogously to Lemma 9.1.12, we obtain the following bound on the size of the graphs in $\mathcal{PW}_d(k)$:

Lemma 9.1.14. *Let $k, d \geq 1$. Every $\boldsymbol{F} \in \mathcal{PW}_d(k)$ has at most $k + d - 1$ vertices.*

Proof. If $d = 1$, then the path decomposition (P, β) of $\boldsymbol{F}$ as in Definition 9.1.13 has at most one bag and $\boldsymbol{F}$ has at most k vertices. For larger d, observe that, when traversing the decomposition from one end of P to the other, every bag introduces exactly one new vertex. Thus, there are at most $k + d - 1$ vertices in $\boldsymbol{F}$. □

Nested Spaces of Homomorphism Vectors

In this section, we prove Theorems 9.1.6 and 9.1.7. We execute the proof of Theorem 9.1.6 in full detail. Theorem 9.1.7 follows along similar lines.

Fix throughout a graph class $\mathcal{F}$ as in Theorem 9.1.6. In reminiscence of Courcelle's Theorem, we let Q denote the set of equivalence classes of $\sim^k_{\mathcal{F}}$, as defined in Definition 9.1.5, and call them *states*. A state $q \in Q$ is *accepting* if, for an (and equivalently, every) $\boldsymbol{F}$ in q, it holds that $\operatorname{soe}(\boldsymbol{F}) \in \mathcal{F}$. Write $A \subseteq Q$ for the set of all accepting states.

To every state $q \in Q$, we associate a finite-dimensional vector space spanned by the homomorphism tensors of the k-labelled graphs $\boldsymbol{F}$ that belong to state q. We show that these vector spaces certify homomorphism indistinguishability. Using a dimensionality argument, we show that these vector spaces are spanned by homomorphism tensors of graphs whose size is bounded by the function $f_{k,C}$ from Theorem 9.1.6. To that end, we decompose the labelled graphs $\boldsymbol{F} \in \mathcal{TW}(k)$ using the operations considered in Lemma 4.3.6.

Formally, we associate to a state $q \in Q$ and an integer $d \geq 1$ the vector space[14]

$$S_d(q) := \operatorname{span}\{\boldsymbol{F}_G \oplus \boldsymbol{F}_H \mid \boldsymbol{F} \in \mathcal{TW}_d(k) \text{ in state } q\} \subseteq \mathbb{R}^{V(G)^k \cup V(H)^k}. \tag{9.1}$$

Here, $\boldsymbol{F}$ is a k-labelled graph of bounded treewidth in the state q. The vector $\boldsymbol{F}_G \oplus \boldsymbol{F}_H := \binom{\boldsymbol{F}_G}{\boldsymbol{F}_H} \in \mathbb{R}^{V(G)^k \cup V(H)^k}$ is obtained by stacking the homomorphism vectors of $\boldsymbol{F}$ with respect to G and H.

Since $\mathcal{TW}_d(k) \subseteq \mathcal{TW}_{d+1}(q)$, the space $S_d(q)$ is a subspace of $S_{d+1}(q)$ for every $d \geq 1$. Ultimately, we are interested in $S(q) := \bigcup_{d \geq 1} S_d(q)$, i.e. the vector space spanned by the homomorphism vectors of all labelled graphs of treewidth $\leq k - 1$ in state q.

By the following Lemma 9.1.15, the vectors in $S(q)$ for $q \in A$ can be used to infer whether G and H are homomorphism indistinguishable over $\mathcal{F}$. Recall that $\mathcal{F}_{\geq k} := \{F \in \mathcal{F} \mid |V(F)| \geq k\}$ for a graph class $\mathcal{F}$ and an integer $k \in \mathbb{N}$.

Lemma 9.1.15. *Two simple graphs G and H are homomorphism indistinguishable over $\mathcal{F}_{\geq k}$ if, and only if, $\boldsymbol{1}_G^T v = \boldsymbol{1}_H^T v$ for every $q \in A$ and every $v \in S(q)$.*

[14] Without loss of generality, we may suppose that $V(G)$ and $V(H)$ are disjoint.

Proof. For the forward direction, note that $S(q)$ is spanned by the $\boldsymbol{F}_G \oplus \boldsymbol{F}_H$ where $\boldsymbol{F} \in \mathcal{TW}(k)$ is in q. Let $\boldsymbol{F}$ in $q \in A$ be arbitrary. Then $\operatorname{soe}(\boldsymbol{F}) =: F \in \mathcal{F}_{\geq k}$ and it holds that $\mathbf{1}_G^T(\boldsymbol{F}_G \oplus \boldsymbol{F}_H) = \operatorname{soe}(\boldsymbol{F}_G) = \hom(F, G) = \hom(F, H) = \mathbf{1}_H^T(\boldsymbol{F}_G \oplus \boldsymbol{F}_H)$.

Conversely, let $F \in \mathcal{F}_{\geq k}$ be arbitrary. Since $\operatorname{tw}(F) \leq k-1$, by Lemma 9.1.10, there exists $\boldsymbol{u} \in V(F)^k$ such that $\boldsymbol{F} := (F, \boldsymbol{u}) \in \mathcal{TW}(k)$. Furthermore, $\boldsymbol{F}$ belongs to some accepting $q \in Q$. Thus, $\boldsymbol{F}_G \oplus \boldsymbol{F}_H \in S(q)$, and hence $\hom(F, G) = \mathbf{1}_G^T(\boldsymbol{F}_G \oplus \boldsymbol{F}_H) = \mathbf{1}_H^T(\boldsymbol{F}_G \oplus \boldsymbol{F}_H) = \hom(F, H)$. □

By Lemma 9.1.12, the space $S_d(q)$ for $d \geq 1$ is spanned by homomorphism tensors of graphs of size $\max\{k^d, d\}$. Thus, Theorem 9.1.6 follows once we establish that $S_{d'}(q) = S(q)$ for all $q \in Q$ and $d' := 2Cn^k$. This d' arises as an upper bound on the dimension of the space $\bigoplus_{q \in Q} S(q)$. The spaces $\bigoplus_{q \in Q} S_d(q)$ for $d \geq 1$ form a chain of nested subspaces in $\bigoplus_{q \in Q} S(q)$. The following Lemma 9.1.16 shows that, once this chain becomes stationary, then the maximal subspace is reached.

Lemma 9.1.16. *If $S_d(q) = S_{d+1}(q)$ for $d \geq 1$ and all $q \in Q$, then $S_d(q) = S(q)$ for all $q \in Q$. In particular, $S_{2Cn^k}(q) = S(q)$ for all $q \in Q$.*

The proof of Lemma 9.1.16 relies on the properties of the relation $\sim_{\mathcal{F}}^k$. In particular, it uses the fact that series composition and gluing, the operations under which $\mathcal{TW}(k)$ is generated by Lemmas 4.3.6 and 4.3.13, preserve the relation $\sim_{\mathcal{F}}^k$. This is established in the following Lemma 9.1.17:

Lemma 9.1.17. *For $\boldsymbol{F}, \boldsymbol{F}', \boldsymbol{F}_1, \boldsymbol{F}_2, \boldsymbol{F}_1', \boldsymbol{F}_2' \in \mathcal{D}(k)$, $\boldsymbol{L} \in \mathcal{D}(k, k)$,*

1. *if $\boldsymbol{F}_1 \sim_{\mathcal{F}}^k \boldsymbol{F}_1'$ and $\boldsymbol{F}_2 \sim_{\mathcal{F}}^k \boldsymbol{F}_2'$, then $\boldsymbol{F}_1 \odot \boldsymbol{F}_2 \sim_{\mathcal{F}}^k \boldsymbol{F}_1' \odot \boldsymbol{F}_2'$,*
2. *if $\boldsymbol{F} \sim_{\mathcal{F}}^k \boldsymbol{F}'$, then $\boldsymbol{L} \cdot \boldsymbol{F} \sim_{\mathcal{F}}^k \boldsymbol{L} \cdot \boldsymbol{F}'$.*

Proof. Let $\boldsymbol{K} = (K, \boldsymbol{u}) \in \mathcal{D}(k)$ be arbitrary. Then $\operatorname{soe}((\boldsymbol{K} \odot \boldsymbol{F}_1) \odot \boldsymbol{F}_2) \in \mathcal{F} \Leftrightarrow \operatorname{soe}((\boldsymbol{K} \odot \boldsymbol{F}_1) \odot \boldsymbol{F}_2') \in \mathcal{F} \Leftrightarrow \operatorname{soe}(\boldsymbol{K} \odot \boldsymbol{F}_1' \odot \boldsymbol{F}_2') \in \mathcal{F}$.

For the second claim, observe that $\operatorname{soe}(\boldsymbol{K} \odot (\boldsymbol{L} \cdot \boldsymbol{F})) = \operatorname{soe}((\boldsymbol{L}^* \cdot \boldsymbol{K}) \odot \boldsymbol{F})$. Thus the second claim follows from the first. □

The algebraic operations on homomorphism tensors corresponding to series composition and gluing are the matrix-vector product and Schur product. Crucially, these operations are linear and bilinear, respectively. This allows Lemma 9.1.16 to be proven by structural induction along Lemmas 4.3.6 and 4.3.13.

Proof of Lemma 9.1.16. We argue that $S_d(q) \supseteq S_{d+i}(q)$ for all $i \geq 1$ by induction on i. The base case holds by assumption. The space $S_{d+i+1}(q)$ is spanned by the vectors $\boldsymbol{F}_G \oplus \boldsymbol{F}_H$ where $\boldsymbol{F} \in \mathcal{TW}_{d+i+1}(k)$ is in q. Recall the bilabelled graphs $\boldsymbol{J}^\ell$ and $\boldsymbol{A}^{ij}$ from Lemma 4.3.6 and Figure 4.3. For such $\boldsymbol{F} \in \mathcal{TW}_{d+i+1}(k)$ in state q, by Lemmas 4.3.6 and 4.3.13, there exist $A \subseteq \binom{[k]}{2}$, $L \subseteq [k]$, and $\boldsymbol{F}^\ell \in \mathcal{TW}_{d+i}(k)$ for $\ell \in L$ such that

$$\boldsymbol{F} = \prod_{ij \in A} \boldsymbol{A}^{ij} \cdot \bigodot_{\ell \in L} \boldsymbol{J}^\ell \boldsymbol{F}^\ell.$$

Let q_ℓ denote the state of $\boldsymbol{F}^\ell$. By assumption, there exist $\boldsymbol{K}^{\ell m} \in \mathcal{TW}_d(k)$ in state q_ℓ and $\alpha_m \in \mathbb{R}$ such that $\boldsymbol{F}^\ell_G \oplus \boldsymbol{F}^\ell_H = \sum_m \alpha_m \boldsymbol{K}^{\ell m}_G \oplus \boldsymbol{K}^{\ell m}_H$. By Lemma 9.1.17, observing that $\mathcal{B}(k,k) \subseteq \mathcal{D}(k,k)$ by definition in Lemma 4.3.6,

$$\boldsymbol{F} \sim^k_{\mathcal{F}} \prod_{ij \in A} \boldsymbol{A}^{ij} \cdot \bigodot_{\ell \in L} \boldsymbol{J}^\ell \boldsymbol{K}^{\ell m}$$

for all m. Thus, $\boldsymbol{F}_G \oplus \boldsymbol{F}_H$ can be written as a linear combination of vectors in $S_{d+i}(q) \subseteq S_d(q)$, by induction. For the final claim, consider the chain of nested subspaces

$$\bigoplus_{q \in Q} S_1(q) \subseteq \bigoplus_{q \in Q} S_2(q) \subseteq \cdots \subseteq \bigoplus_{q \in Q} S(q).$$

By what was just shown, for every $d \geq 1$, either $\bigoplus_{q \in Q} S_d(q)$ is a proper subspace of $\bigoplus_{q \in Q} S_{d+1}(q)$ or $\bigoplus_{q \in Q} S_d(q) = \bigoplus_{q \in Q} S(q)$. Since the dimension of $\bigoplus_{q \in Q} S(q)$ is at most $2Cn^k$, the chain becomes stationary after at most $2Cn^k$ steps. □

This concludes the preparations for the proof of Theorem 9.1.6.

Proof of Theorem 9.1.6. It suffices to prove the backward implication. Since $k \leq k^{2Cn^k}$, it suffices to show that $G \equiv_{\mathcal{F}_{\geq k}} H$ by verifying the condition in Lemma 9.1.15. By Lemma 9.1.16, $S_d(q) = S(q)$ for $d := 2Cn^k$ and all $q \in Q$. Hence, $S(q)$ is spanned by the $\boldsymbol{F}_G \oplus \boldsymbol{F}_H$ where $\boldsymbol{F} \in \mathcal{TW}_d(k)$ is in state q. By Lemma 9.1.12, these graphs have at most $\max\{k^d, d\} = \max\{k^{2Cn^k}, 2Cn^k\}$ vertices. Thus, $\boldsymbol{1}^T_G(\boldsymbol{F}_G \oplus \boldsymbol{F}_H) = \hom(\operatorname{soe}(\boldsymbol{F}), G) = \hom(\operatorname{soe}(\boldsymbol{F}), H) = \boldsymbol{1}^T_H(\boldsymbol{F}_G \oplus \boldsymbol{F}_H)$, as desired. □

Finally, we adapt the techniques developed so far to prove the following analogue of Theorem 9.1.6 for graph classes of bounded pathwidth. In contrast to Theorem 9.1.6, the function in Theorem 9.1.7 bounding the size of the graphs which need to be considered is polynomial.

Proof of Theorem 9.1.7. The proof of Theorem 9.1.6 needs to be adapted as follows. Define $S_d(q)$ and $S(q)$ as in Equation (9.1) with $\mathcal{PW}_d(k)$ instead of $\mathcal{TW}_d(k)$, i.e.

$$S_d(q) := \operatorname{span}\{\boldsymbol{F}_G \oplus \boldsymbol{F}_H \mid \boldsymbol{F} \in \mathcal{PW}_d(k) \text{ in state } q\} \subseteq \mathbb{R}^{V(G)^k \cup V(H)^k}.$$

and $S(q) := \bigcup_{d \geq 1} S_d(q)$.

By Lemma 4.3.6, observing that, for every $\boldsymbol{F} \in \mathcal{PW}_{d+1}(k)$, there exists $A \subseteq \binom{[k]}{2}$, $\ell \in [k]$, and $\boldsymbol{F}' \in \mathcal{PW}_d(k)$ such that $\boldsymbol{F} = \boldsymbol{A}\boldsymbol{J}^\ell \boldsymbol{F}'$ where $\boldsymbol{A} := \prod_{ij \in A} \boldsymbol{A}^{ij}$, Lemma 9.1.16 follows analogously. Thus, $S_d(q) = S(q)$ for all $q \in Q$ when $d := 2Cn^k$. By Lemma 9.1.14, this space is spanned by homomorphism tensors of graphs of size at most $2Cn^k + k - 1$. The statement follows as in Theorem 9.1.6. □

For the classes $\mathcal{TW}_k$ and $\mathcal{PW}_k$, i.e. the classes of all graphs of treewidth or pathwidth $\leq k$, the proofs of Theorems 9.1.6 and 9.1.7 yield witness functions which slightly improve those derived in Theorems 4.3.8 and 4.3.14. For these classes, we do not need to consider the states from Definition 9.1.5. Thus, all constructions take place in a vector space of maximal dimension $2n^k$. This allows to shave off the factor C from Theorems 9.1.6 and 9.1.7.

Corollary 9.1.18. *Let $k \geq 1$. For simple graphs G and H on at most n vertices, with $f_k(n) := \max\{k^{2n^k}, 2n^k\}$,*

$$G \equiv_{\mathcal{TW}_{k-1}} H \iff G \equiv_{(\mathcal{TW}_{k-1})_{\leq f_k(n)}} H.$$

Proof. Departing from Equation (9.1), define

$$S_d := \operatorname{span}\{\boldsymbol{F}_G \oplus \boldsymbol{F}_H \mid \boldsymbol{F} \in \mathcal{TW}_d(k)\} \subseteq \mathbb{R}^{V(G)^k \cup V(H)^k}.$$

and $S := \bigcup_{d \geq 1} S_d$. Then $\dim(S) \leq 2n^k$. Lemma 9.1.16 implies that $S_d = S$ for some $d \leq 2n^k$. The assertion follows invoking Lemmas 9.1.12 and 9.1.15 as in the proof of Theorem 9.1.6. □

Corollary 9.1.19. *Let $k \geq 1$. For simple graphs G and H on at most n vertices, with $f_k(n) := 2n^k + k - 1$,*

$$G \equiv_{\mathcal{PW}_{k-1}} H \iff G \equiv_{(\mathcal{PW}_{k-1})_{\leq f_k(n)}} H.$$

Proof. Departing from Equation (9.1), define

$$S_d := \operatorname{span}\{\boldsymbol{F}_G \oplus \boldsymbol{F}_H \mid \boldsymbol{F} \in \mathcal{PW}_d(k)\} \subseteq \mathbb{R}^{V(G)^k \cup V(H)^k}.$$

and $S := \bigcup_{d \geq 1} S_d$. Then $\dim(S) \leq 2n^k$. Lemma 9.1.16 implies that $S_d = S$ for some $d \leq 2n^k$. The assertion follows invoking Lemmas 9.1.14 and 9.1.15 as in the proof of Theorem 9.1.6. □

9.2 Algorithmic Meta Theorems for Homomorphism Indistinguishability

In this section, the central results of this chapter, Theorems 9.0.1 and 9.0.2 are proven. First, in Section 9.2.1, a deterministic polynomial-time algorithm for modular homomorphism indistinguishability over every recognisable graph class of bounded treewidth is given. In Section 9.2.2, Theorems 9.0.1 and 9.0.2 are reduced to this result. In Section 9.2.3, arguments due to Courcelle are employed to derive fixed-parameter algorithms for deciding homomorphism indistinguishability over a graph classes specified by a CMSO_2-formula as part of the input.

9.2.1 Modular Homomorphism Indistinguishability in Polynomial Time

The insight that yielded Theorems 9.1.6 and 9.1.7 is that the chain of vector spaces $S_1(q) \subseteq \cdots \subseteq S_d(q) \subseteq S_{d+1}(q) \subseteq \dots$ as defined in Equation (9.1) reaches the fixed point $S(q)$ after polynomially many steps. In this section, we strengthen this result by showing that bases $B(q)$ for the spaces $S(q)$ can be efficiently computed. A technical difficulty arising here is that the numbers produced in the process can be of doubly exponential magnitude. In order to overcome this problem, we first consider homomorphism indistinguishability modulo primes. See Chapter 8, for background on modular homomorphism indistinguishability.

ModHomInd($\mathcal{F}$)
Input Simple graphs G and H, a prime p in binary.
Question Are G and H homomorphism indistinguishable over $\mathcal{F}$ modulo p?

Theorem 9.2.1. *Let $k \geq 1$. If $\mathcal{F}$ is a k-recognisable graph class of treewidth $\leq k-1$, then* ModHomInd($\mathcal{F}$) *is in polynomial time.*

The algorithm yielding Theorem 9.2.1 is formally stated as Algorithm 9.1. The idea is to iteratively compute bases $B(q)$ for the spaces

$$S(q) := \langle \{ \boldsymbol{F}_G \oplus \boldsymbol{F}_H \mid \boldsymbol{F} \in \mathcal{TW}(k) \text{ in state } q \} \rangle \subseteq \mathbb{F}_p^{V(G)^k \cup V(H)^k}.$$

Initially, all $B(q)$ are empty. Only $B(q_0)$ where q_0 is the state of $\mathbf{1} \in \mathcal{TW}(k)$ from Figure 3.3a contains the homomorphism vector $\mathbf{1}_G \oplus \mathbf{1}_H$. Subsequently, the operations from Lemmas 4.3.6 and 4.3.13 are applied to compute new homomorphism vectors. For every new vector belonging to state q, it is checked whether it is a linear combination of the already computed basis vectors in $B(q)$. If not, it is added to $B(q)$. Analogous to Lemma 9.1.16, this process reaches a fixed point after a polynomial number of iterations. At this point, the computed $B(q)$ are bases for the $S(q)$. Finally, Lemma 9.1.15 can be invoked to conclude whether the input graphs are homomorphism indistinguishable over $\mathcal{F}$ modulo p.

Algorithm 9.1 is supplied with a hard-coded description of the graph class $\mathcal{F}$. To that end, consider the following objects. Recall Q, the set of equivalence classes (states) of $\sim^k_{\mathcal{F}}$, and $A \subseteq Q$, the set of accepting states from Section 9.1. Lemma 9.1.17 asserts that the operations generating $\mathcal{TW}(k)$ can be regarded as operations on the equivalence classes of $\sim^k_{\mathcal{F}}$. The algorithm is provided with a table encoding how each of the operations acts on the states.

Write $\pi \colon \mathcal{TW}(k) \to Q$ for the map that associates an $\boldsymbol{F} \in \mathcal{TW}(k)$ to its state $q \in Q$. Write q_0 for the state of $\mathbf{1} \in \mathcal{TW}(k)$. Recall the definition of $\mathcal{B}(k,k)$ from Lemma 4.3.6. Write $g \colon Q \times Q \to Q$ and $b_{\boldsymbol{B}} \colon Q \to Q$ for every $\boldsymbol{B} \in \mathcal{B}(k,k)$ such that

$$g(\pi(\boldsymbol{F}), \pi(\boldsymbol{F}')) = \pi(\boldsymbol{F} \odot \boldsymbol{F}'), \tag{9.2}$$
$$b_{\boldsymbol{B}}(\pi(\boldsymbol{F})) = \pi(\boldsymbol{B} \cdot \boldsymbol{F}). \tag{9.3}$$

for every $\boldsymbol{F}, \boldsymbol{F}' \in \mathcal{TW}(k)$ and $\boldsymbol{B} \in \mathcal{B}(k,k)$. Note that Q, A, g, q_0 and the $b_{\boldsymbol{B}}$, $\boldsymbol{B} \in \mathcal{B}(k,k)$, are finite objects, which can be hard-coded. The map π does not need to be computable and is only needed for analysing the algorithm. Having defined all necessary notions, we can now describe Algorithm 9.1.

Algorithm 9.1: ModHomInd($\mathcal{F}$) for k-recognisable graph classes $\mathcal{F}$ of treewidth $\leq k-1$.

Input: Simple graphs G and H, a prime p in binary.
Data: k, Q, A, q_0, g, $b_{\boldsymbol{B}}$ for $\boldsymbol{B} \in \mathcal{B}(k,k)$.
Output: Whether $G \equiv^p_{\mathcal{F}} H$.

1 with brute force check whether G and H are homomorphism indistinguishable over the finite graph class $\mathcal{F}_{\leq k}$ modulo p and reject if not;
2 for every $\boldsymbol{B} \in \mathcal{B}(k,k)$, compute the homomorphism matrices $\boldsymbol{B}_G \in \mathbb{F}_p^{V(G)^k \times V(G)^k}$ and $\boldsymbol{B}_H \in \mathbb{F}_p^{V(H)^k \times V(H)^k}$;
3 initialise $B(q_0) \leftarrow \{\boldsymbol{1}_G \oplus \boldsymbol{1}_H\} \subseteq \mathbb{F}_p^{V(G)^k \cup V(H)^k}$;
4 initialise $B(q) \leftarrow \emptyset \subseteq \mathbb{F}_p^{V(G)^k \cup V(H)^k}$ for all $q \neq q_0$;
5 **repeat**
6   **foreach** $\boldsymbol{B} \in \mathcal{B}(k,k)$, $q \in Q$, $v \in B(q)$ **do**
7         $w \leftarrow (\boldsymbol{B}_G \oplus \boldsymbol{B}_H)v := \begin{pmatrix} \boldsymbol{B}_G & 0 \\ 0 & \boldsymbol{B}_H \end{pmatrix} v$;
8   **if** $w \notin \operatorname{span}(B(b_{\boldsymbol{B}}(q)))$ **then**
9   add w to $B(b_{\boldsymbol{B}}(q))$;
10   **foreach** $q_1, q_2 \in Q$, $v_1 \in B(q_1)$, $v_2 \in B(q_2)$ **do**
11         $w \leftarrow v_1 \odot v_2$;
12   **if** $w \notin \operatorname{span}(B(g(q_1, q_2)))$ **then**
13   add w to $B(g(q_1, q_2))$;
14 **until** *none of the* $B(q)$, $q \in Q$, *are updated*;
15 **if** $\boldsymbol{1}_G^T v = \boldsymbol{1}_H^T v$ *for all* $q \in A$ *and* $v \in B(q)$ **then**
16   accept;
17 **else**
18   reject;

Lemma 9.2.2. *Write* $n := \max\{|V(G)|, |V(H)|\}$ *and* $C := |Q|$. *There exists a computable function* f *such that Algorithm 9.1 runs in time* $f(k, C)n^{O(k)}(\log p)^{O(1)}$.

Proof. Consider the following individual runtimes:

Counting homomorphisms of a graph on k vertices into a graph on n vertices, can be done in time $O(n^k)$. Hence Line 1 requires time $f(k)n^k$ for some computable f.

Computing the homomorphism tensors in Line 2 requires time $O(k^2 n^{2k})$.

Throughout the execution of Algorithm 9.1, the vectors in each $B(q)$, $q \in Q$, are linearly independent. Thus, $|B(q)| \leq \dim(S(q)) \leq 2n^k$ and $\sum_{q \in Q} |B(q)| \leq 2Cn^k$.

Hence, the body of the loop in Line 5 is entered at most $2Cn^k$ many times.

The loop in Line 6 iterates over at most $O(k^2) \cdot C \cdot 2n^k$ many objects. Computing the vector w takes polynomial time in $2n^k \cdot \log p$. The same holds for checking the condition in Lines 8 and 12, e.g. via Gaussian elimination. The loop in Line 10 iterates over at most $C^2 \cdot (2n^k)^2$ many objects.

Finally, checking the condition in Line 15 takes $C \cdot 2n^k \cdot (\log p)^{O(1)}$ many steps. □

The following Lemma 9.2.3 implies that Algorithm 9.1 is correct.

Lemma 9.2.3. *When Algorithm 9.1 terminates, $B(q)$ spans $S(q)$ for all $q \in Q$.*

Proof. First observe that the invariant $B(q) \subseteq S(q)$ for all $q \in Q$ is preserved throughout Algorithm 9.1. Indeed, for example in Line 6, since $v \in B(q) \subseteq S(q)$, it can be written as a linear combination of $\boldsymbol{F}_G \oplus \boldsymbol{F}_H$ for $\boldsymbol{F} \in \mathcal{TW}(k)$ of state q. Because $\boldsymbol{B} \cdot \boldsymbol{F}$ is in state $b_{\boldsymbol{B}}(q)$ by Equation (9.3), $(\boldsymbol{B}_G \oplus \boldsymbol{B}_H)v$ is in the span of the vectors $\boldsymbol{B}_G\boldsymbol{F}_G \oplus \boldsymbol{B}_H\boldsymbol{F}_H \in S(b_{\boldsymbol{B}}(q))$ for $\boldsymbol{F} \in \mathcal{TW}(k)$ of state q.

Now consider the converse inclusion. The proof is by induction on the structure in Lemmas 4.3.6 and 4.3.13. After Line 3, $\boldsymbol{1}_G \oplus \boldsymbol{1}_H$ is in the span of $B(q_0)$.

For the inductive step, suppose that $\boldsymbol{F} \in \mathcal{TW}(k)$ of state $q \in Q$ is such that $\boldsymbol{F}_G \oplus \boldsymbol{F}_H = \sum_{v \in B(q)} \alpha_v v$ for some coefficients $\alpha_v \in \mathbb{F}_p$. Let $\boldsymbol{B} \in \mathcal{B}(k,k)$ and $\boldsymbol{F}' := \boldsymbol{B} \cdot \boldsymbol{F}$. Then $(\boldsymbol{B}_G \oplus \boldsymbol{B}_H)v$ is in the span of $B(b_{\boldsymbol{B}}(q))$ for all $v \in B(q)$ by the termination condition. Hence, $\boldsymbol{F}'_G \oplus \boldsymbol{F}'_H = \sum_{v \in B(q)} \alpha_v (\boldsymbol{B}_G \oplus \boldsymbol{B}_H)v$ is in the span of $B(b_{\boldsymbol{B}}(q))$.

Let $\boldsymbol{F}^1, \boldsymbol{F}^2 \in \mathcal{TW}(k)$ of states $q_1, q_2 \in Q$ be such that $\boldsymbol{F}^1_G \oplus \boldsymbol{F}^1_H = \sum_{v \in B(q_1)} \alpha_v v$ and $\boldsymbol{F}^2_G \oplus \boldsymbol{F}^2_H = \sum_{w \in B(q_2)} \beta_w w$ for some coefficients $\alpha_v, \beta_w \in \mathbb{F}_p$. Since the algorithm terminated, all $v \odot w$ for $v \in B(q_1)$ and $w \in B(q_2)$ are in the span of $B(g(q_1, q_2))$. Then $(\boldsymbol{F}^1 \odot \boldsymbol{F}^2)_G \oplus (\boldsymbol{F}^1 \odot \boldsymbol{F}^2)_H = (\boldsymbol{F}^1_G \oplus \boldsymbol{F}^1_H) \odot (\boldsymbol{F}^2_G \oplus \boldsymbol{F}^2_H) = \sum_{v \in B(q_1), w \in B(q_2)} \alpha_v \beta_w (v \odot w)$ is in the span of $B(g(q_1, q_2))$. □

This concludes the preparations for the proof of Theorem 9.2.1:

Proof of Theorem 9.2.1. Lemma 9.2.3 implies that the conditions in Lemma 9.1.15 and Line 15 are equivalent. Thus, G and H are homomorphism indistinguishable over $\mathcal{F}_{\geq k}$ modulo p if, and only if, the condition in Line 15 holds. The runtime bound is given in Lemma 9.2.2. □

9.2.2 Homomorphism Indistinguishability in Randomised Polynomial Time

In this section, we give a randomised polynomial-time reduction from HomInd($\mathcal{F}$) to ModHomInd($\mathcal{F}$) for recognisable graph classes $\mathcal{F}$ of bounded treewidth. If $\mathcal{F}$ is of bounded pathwidth, then this reduction can be achieved deterministically. Thereby, we prove Theorems 9.0.1 and 9.0.2.

Theorems 9.1.6 and 9.1.7 give bounds N on the size of the largest graph in $\mathcal{F}$ which needs to be considered in order to conclude whether two graphs on at most

n vertices are homomorphism indistinguishable over $\mathcal{F}$. A graph on at most N vertices may have at most n^N homomorphisms to a graph on n vertices. Thus, for graphs on at most n vertices, homomorphism indistinguishability over $\mathcal{F}$ is the same as homomorphism indistinguishability over $\mathcal{F}$ modulo any number greater than n^N. Equipped with the following Lemma 9.2.4, which is derived from the Chinese Remainder Theorem and the Prime Number Theorem, we show Theorems 9.0.1 and 9.0.2.

Lemma 9.2.4. *Let $\mathcal{F}$ be a graph class with witness function $f \colon \mathbb{N} \to \mathbb{N}$. Let G and H be simple graphs on at most n vertices. Suppose that $f(n) \log n \geq e^{2000}$. If $G \not\equiv_{\mathcal{F}} H$, then the probability that a random prime $f(n) \log n < p \leq (f(n) \log n)^2$ is such that $G \equiv_{\mathcal{F}}^p H$ is at most $\frac{2}{f(n) \log n}$.*

Before proving Lemma 9.2.4, we state the following consequences of (an explicit version of) the Prime Number Theorem [156]. Here, e denotes Euler's number.

Fact 9.2.5. *For $n \geq e^{2000}$, the following hold:*

1. *the number of primes $n < p \leq n^2$ is at least $\frac{n^2}{2 \log n}$,*
2. *the product of all primes $n < p \leq n^2$ is at least is $2^{n^2/2}$, and*
3. *the probability that a random number $n < m \leq n^2$ is prime is at least $\frac{1}{2 \log n}$.*

Proof. Let ln denote the natural logarithm. Let $\pi(n)$ denote the number of primes $\leq n$. By [156, Theorem 30A], for $n \geq e^{2000}$,

$$\frac{n}{\ln(n)} < \pi(n) < \frac{n}{\ln(n) - 2}.$$

Thus the desired quantity $\pi(n^2) - \pi(n)$ is at least

$$\frac{n^2}{2\ln(n)} - \frac{n}{\ln(n) - 2} \geq \frac{n^2}{2\ln(n)} - \frac{2n}{\ln(n)} = \frac{n^2 - 4n}{2\ln(n)} \geq \frac{n^2 \ln 2}{2\ln(n)} = \frac{n^2}{2\log(n)}.$$

For the second claim, the product of the primes is $> n^{\frac{n^2}{2\log n}} = 2^{\frac{n^2}{2\log n} \log n} = 2^{n^2/2}$. For the third claim, $\frac{n^2}{2\log n} \cdot \frac{1}{n^2 - n} \geq \frac{1}{2\log n}$. $\square$

Lemma 9.2.4 follows from Fact 9.2.5 and the Chinese Remainder Theorem (Fact 8.2.2).

Proof of Lemma 9.2.4. Let $N := f(n)$. By Definition 9.1.1, if $G \not\equiv_{\mathcal{F}} H$, there exist $F \in \mathcal{F}_{\leq N}$ such that $\hom(F, G) \neq \hom(F, H)$. Both numbers are non-negative and at most n^N. Thus, $\hom(F, G) \not\equiv \hom(F, H) \mod n^N$.

By Fact 8.2.2, in every set X of at least $\frac{N \log n}{\log(N \log n)}$ many primes $N \log n < p' \leq (N \log n)^2$, there exists a prime $p \in X$ such that $\hom(F, G) \not\equiv \hom(F, H) \mod p$. This is because $\prod_{p' \in X} p' > (N \log n)^{\frac{N \log n}{\log(N \log n)}} = n^N$.

By Fact 9.2.5, the probability that a random prime $N \log n < p \leq (N \log n)^2$ is such that $\hom(F, G) \equiv \hom(F, H) \mod p$ is

$$\leq \frac{N \log n}{\log(N \log n)} \cdot \frac{2 \log(N \log n)}{(N \log n)^2} = \frac{2}{N \log n}.$$

Since $G \equiv^p_{\mathcal{F}} H$ implies that $\hom(F, G) \equiv \hom(F, H) \mod p$, the desired probability is bounded by the same value. □

By Lemma 9.2.4, HomInd($\mathcal{F}$) reduces to ModHomInd($\mathcal{F}$). The complexity of the reduction depends on the growth of the witness function of $\mathcal{F}$. If the witness function is polynomial, then all relevant primes can be enumerated deterministically. If the witness function is exponential, then one may sample primes of polynomial size and invoke ModHomInd($\mathcal{F}$). Lemma 9.2.4 implies that a random prime p of appropriate size certifies that $G \not\equiv_{\mathcal{F}} H$ with high probability.

Lemma 9.2.6. *Let $\mathcal{F}$ be a graph class. Let $q \in \mathbb{N}[x]$ be a polynomial.*

1. *If $n \mapsto q(n)$ is a witness function for $\mathcal{F}$, then there is a deterministic polynomial-time algorithm for* HomInd($\mathcal{F}$) *requiring a* ModHomInd($\mathcal{F}$)*-oracle.*
2. *If $n \mapsto 2^{q(n)}$ is a witness function for $\mathcal{F}$, then there is a randomised algorithm for* HomInd($\mathcal{F}$) *requiring a* ModHomInd($\mathcal{F}$)*-oracle whose runtime is always polynomial. It accepts all* YES*-instances and accepts* NO*-instances with probability less than one half.*

Proof. First suppose that $f \colon n \mapsto q(n)$ is a witness function for $\mathcal{F}$. Let G and H be simple graphs on at most n vertices. We may assume without loss of generality that $f(n) \log n \geq e^{2000}$. By Lemma 9.2.4, $G \equiv_{\mathcal{F}} H$ if, and only if, $G \equiv^p_{\mathcal{F}} H$ for all primes $f(n) \log n < p \leq (f(n) \log n)^2$. Thus, $G \equiv_{\mathcal{F}} H$ can be checked by enumerating all numbers between $f(n) \log n$ and $(f(n) \log n)^2$ and executing Algorithm 9.1 for every prime. This procedure requires a deterministic polynomial-time primality test as given in [6].

For the second claim, suppose that $f \colon n \mapsto 2^{p(n)}$ is a witness function for $\mathcal{F}$. Consider Algorithm 9.2. Let G and H be simple graphs on at most n vertices. We may suppose that $f(n) \log n \geq e^{2000}$. Since $\log f(n) = q(n)$ is polynomial in the input size, Algorithm 9.2 runs in polynomial time, requiring a polynomial-time primality test [6].

For correctness, first observe that if $G \equiv_{\mathcal{F}} H$, then Algorithm 9.1 always accepts. If $G \not\equiv_{\mathcal{F}} H$, then Algorithm 9.1 might accept incorrectly. In each iteration, the probability of not rejecting immediately is at most the probability of not sampling a prime plus the probability of the prime p being such that $G \equiv^p_{\mathcal{F}} H$. By Fact 9.2.5 and Lemma 9.2.4, it is at most

$$1 - \frac{1}{2 \log(f(n) \log n)} + \frac{2}{f(n) \log n} \leq 1 - \frac{1}{4 \log(f(n) \log n)}.$$

By [1, Equation 4.2.30], the total probability of incorrectly accepting if $G \not\equiv_{\mathcal{F}} H$ is at most $e^{-1} < 1/2$. □

Algorithm 9.2: A randomised reduction from HomInd($\mathcal{F}$) to ModHomInd($\mathcal{F}$) for a graph class $\mathcal{F}$ with exponential witness function $f\colon \mathbb{N} \to \mathbb{N}$.

Input: Simple graphs G and H.
Output: Whether $G \equiv_{\mathcal{F}} H$.
$n \leftarrow \max\{|V(G)|, |V(H)|\}$;
for $\lceil 4 \log(f(n) \log n) \rceil$ *times* **do**
  sample a random integer $f(n) \log n < p \leq (f(n) \log n)^2$;
  if *p is a prime and* $G \not\equiv^p_{\mathcal{F}} H$ **then**
    reject;
accept;

Lemma 9.2.6 implies the two main theorems of this chapter.

Theorem 9.0.1. *Let $k \geq 1$. If $\mathcal{F}$ is a k-recognisable class of graphs of treewidth at most $k-1$, then* HomInd($\mathcal{F}$) *is in* coRP.

Proof. Let C denote the k-recognisability index of $\mathcal{F}$. By Theorem 9.1.6, $f\colon n \mapsto \max\{k^{2Cn^k}, 2Cn^k\}$ is a witness function for $\mathcal{F}$. By Theorem 9.2.1, ModHomInd($\mathcal{F}$) is in polynomial time. Hence, Lemma 9.2.6 yields a randomised polynomial-time algorithm for HomInd($\mathcal{F}$) with the desired error probabilities. □

Theorem 9.0.2. *Let $k \geq 1$. If $\mathcal{F}$ is a k-recognisable class of graphs of pathwidth at most $k-1$, then* HomInd($\mathcal{F}$) *is in polynomial time.*

Proof. Let C denote the k-recognisability index of $\mathcal{F}$. By Theorem 9.1.7, $f\colon n \mapsto 2Cn^k + k - 1$ is a witness function for $\mathcal{F}$. By Theorem 9.2.1, ModHomInd($\mathcal{F}$) is in polynomial time. Hence, Lemma 9.2.6 yields a deterministic polynomial-time algorithm for HomInd($\mathcal{F}$). □

9.2.3 Homomorphism Indistinguishability as Parametrised Problem

The connection to Courcelle's Theorem motivates considering the parametrised problem HomInd. Here, the CMSO_2-sentence φ allows the graph class to be specified as part of the input. Using results by Courcelle [46], we generalise Theorem 9.0.1 in Theorem 9.2.7.

HomInd
Input Simple graphs G and H, a CMSO_2-sentence φ, an integer $k \geq 1$.
Parameter $|\varphi| + k$.
Question Are G and H homomorphism indistinguishable over the graph class $\mathcal{F}_{\varphi,k}$ of graphs of treewidth at most $k-1$ satisfying φ?

Theorem 9.2.7. *There exists a computable function $f\colon \mathbb{N} \to \mathbb{N}$ and a randomised algorithm for* HomInd *of runtime $f(|\varphi| + k)n^{O(k)}$ for $n := \max\{|V(G)|, |V(H)|\}$ which accepts all* YES*-instances and accepts* NO*-instances with probability less than one half.*

In light of Theorem 9.0.1, the challenge is to efficiently compute the data describing the graph class $\mathcal{F}_{\varphi,k}$ for Algorithm 9.1 from the CMSO_2-sentence φ and the integer k. That this can be done was proven by Courcelle [46]. More precisely, Courcelle proved that, given a CMSO_2-sentence φ and integer k, one can compute a finite automaton processing expressions which encode (tree decompositions of) graphs of bounded treewidth. It is this automaton from which the data required by Algorithm 9.1 can be derived.

For graph classes of bounded pathwidth, the analogous problem can be decided deterministically in the same time. Theorem 9.2.8 places PWHomInd in the parametrised complexity class XP.

PWHomInd
Input Simple graphs G and H, a CMSO_2-sentence φ, an integer $k \geq 1$.
Parameter $|\varphi| + k$.
Question Are G and H homomorphism indistinguishable over the graph class $\mathcal{F}_{\varphi,k}$ of graphs of pathwidth at most $k-1$ satisfying φ?

Theorem 9.2.8. *There exists a computable function $f\colon \mathbb{N} \to \mathbb{N}$ and a deterministic algorithm for* PWHomInd *of runtime $f(|\varphi| + k)n^{O(k)}$ for $n := \max\{|V(G)|, |V(H)|\}$.*

In order to state the argument yielding Theorems 9.2.7 and 9.2.8 precisely with minimal technical overhead, we introduce the following syntactical counterpart of $\mathcal{TW}(k)$. Write $\mathfrak{TW}(k)$ for the set of *terms* defined inductively as the following formal expressions:

1. $\boldsymbol{1} \in \mathcal{TW}(k)$ is a term,
2. if $t_1, t_2 \in \mathfrak{TW}(k)$, then $t_1 \odot t_2 \in \mathfrak{TW}(k)$,
3. if $t \in \mathfrak{TW}(k)$, then $\boldsymbol{B} \cdot t \in \mathfrak{TW}(k)$ for every $\boldsymbol{B} \in \mathcal{B}(k,k)$.

The key difference between $\mathcal{TW}(k)$ and $\mathfrak{TW}(k)$ is that, for elements of the former, the tree decomposition satisfying Definition 9.1.11 is implicit. For applying Courcelle's techniques, we require this decomposition to be explicit, as in $\mathfrak{TW}(k)$.

There is a mapping $\operatorname{val}\colon \mathfrak{TW}(k) \to \mathcal{TW}(k)$ interpreting the formal expressions above as concrete k-labelled graphs. By Lemmas 4.3.6 and 4.3.13, this mapping is surjective. We require the following Theorem 9.2.9.

Theorem 9.2.9 (Courcelle [46])**.** *Given a CMSO_2-sentence φ and an integer $k \geq 1$, one can compute a finite set Q, a subset $A \subseteq Q$, an element $q_0 \in Q$, and functions $g\colon Q \times Q \to Q$ and $b_{\boldsymbol{B}}\colon Q \to Q$ for every $\boldsymbol{B} \in \mathcal{B}(k,k)$ such that there exists a map $\pi\colon \mathfrak{TW}(k) \to Q$ satisfying*

1. $\pi(\boldsymbol{1}) = q_0$,
2. *for $t \in \mathfrak{TW}(k)$, $\operatorname{soe}(\operatorname{val}(t)) \models \varphi$ if, and only if, $\pi(t) \in A$,*

3. *for all* $t_1, t_2 \in \mathfrak{TW}(k)$, $\pi(t_1 \odot t_2) = g(\pi(t_1), \pi(t_2))$,
4. *for all* $t \in \mathfrak{TW}(k)$ *and* $\boldsymbol{B} \in \mathcal{B}(k,k)$, $\pi(\boldsymbol{B} \cdot t) = b_{\boldsymbol{B}}(\pi(t))$.

Proof. We sketch how the theorem follows from results in [48]. By the proof of [48, Theorem 6.3], one can compute, given φ and k, a finite deterministic automaton recognising whether a so-called $F^{\mathrm{HR}}_{[k]}$-term evaluates to a graph satisfying φ.

$F^{\mathrm{HR}}_{[k]}$ plays the role of $\mathfrak{TW}(k)$. It is a many-sorted signature [48, Definition 2.123] whose sorts correspond to the sources (labels, in our terminology) of sourced (distinctly labelled) graphs of bounded treewidth. The operations of $F^{\mathrm{HR}}_{[k]}$ are gluing, dropping of labels, and renaming of labels [48, Definition 2.32]. Note that the series composition, gluing, and unlabelling from Section 3.2.1 can be derived from these. Moreover, our graph $\mathbf{1} \in \mathcal{TW}(k)$ in Figure 3.3a can be derived from the constants in $F^{\mathrm{HR}}_{[k]}$.

A finite deterministic automaton processing $F^{\mathrm{HR}}_{[k]}$-terms [48, Definition 3.46] comprises a finite set of states Q, a set of accepting states $A \subseteq Q$, and a transition function δ mapping tuples $(q_1, \dots, q_\ell, f)$ to q where $q_1, \dots, q_\ell, q \in Q$ and $f \in F^{\mathrm{HR}}_{[k]}$ is an ℓ-ary operation, $\ell \geq 0$. Furthermore, there is map σ that associates to every $q \in Q$ a sort, that is a subset of $[k]$ of labels.

The automaton processes $F^{\mathrm{HR}}_{[k]}$-terms, which in our language are elements of $\mathfrak{TW}(k)$. In order to derive the desired objects from this automaton, we retain those states for which σ evaluates to $[k]$. The maps g and $b_{\boldsymbol{B}}$ can be computed from δ. We mark a state q as accepting if the transition function δ maps q and soe (an operation derived from those in $F^{\mathrm{HR}}_{[k]}$) to an accepting state. The state q_0 is the state produced by δ on the 0-ary derived function $\mathbf{1}$. The map π (which we do not require to be computable) is defined by the semantics of the automaton. □

Equipped with Theorem 9.2.9, we prove Theorems 9.2.7 and 9.2.8.

Proof of Theorem 9.2.7. Let $\mathcal{F} := \mathcal{F}_{\varphi,k}$. Invoking Theorem 9.2.9, compute Q, A, q_0, g, and $b_{\boldsymbol{B}}$ for $\boldsymbol{B} \in \mathcal{B}(k,k)$ in time only depending on $|\varphi|$ and k. Now invoke Algorithm 9.2, which calls Algorithm 9.1.

Since the data obtained from Theorem 9.2.9 describes an equivalence relation $t_1 \sim t_2$ if, and only if, $\pi(t_1) = \pi(t_2)$ on $\mathfrak{TW}(k)$ rather than on $\mathcal{TW}(k)$ as in Sections 9.1 and 9.2.1, the definitions in these sections need to be slightly adapted.

By Theorem 9.2.9, the following assertions analogous to Lemma 9.1.17 hold: If $\pi(t_1) = \pi(t_2)$ and $\pi(t_1') = \pi(t_2')$ for $t_1, t_2, t_1', t_2' \in \mathfrak{TW}(k)$, then $\pi(t_1 \odot t_1') = g(\pi(t_1), \pi(t_1')) = g(\pi(t_2), \pi(t_2')) = \pi(t_2 \odot t_2')$. Furthermore, for all $\boldsymbol{B} \in \mathcal{B}(k,k)$, $\pi(\boldsymbol{B} \cdot t_1) = b_{\boldsymbol{B}}(\pi(t_1)) = b_{\boldsymbol{B}}(\pi(t_2)) = \pi(\boldsymbol{B} \cdot t_2)$.

Define $S(q)$ for $q \in Q$ as the $\mathbb{F}_p$-vector space spanned by the $\boldsymbol{F}_G \oplus \boldsymbol{F}_H$ where $\boldsymbol{F} = \mathrm{val}(t)$ for some $t \in \mathfrak{TW}(k)$ in state q.

Then the following analogue of Lemma 9.1.15 holds: Two simple graphs G and H are homomorphism indistinguishable over $\mathcal{F}_{\geq k}$ modulo p if, and only if,

$\mathbf{1}_G^T v = \mathbf{1}_H^T v$ for every $v \in S(q)$ and $q \in A$. Indeed, if $F \in \mathcal{F}_{\geq k}$, then there exists by Theorem 9.2.9 a term $t \in \mathfrak{TW}(k)$ such that $\operatorname{soe}(\boldsymbol{F}) \cong F$ for $\boldsymbol{F} := \operatorname{val}(t)$. Then $\hom(F, G) \equiv \mathbf{1}_G^T(\boldsymbol{F}_G \oplus \boldsymbol{F}_H) \mod p$. The claim follows as in Lemma 9.1.15.

The proof of Lemma 9.2.3 goes through analogously, replacing structural induction on $\mathcal{TW}(k)$ via Lemmas 4.3.6 and 4.3.13 by structural induction on the definition of $\mathfrak{TW}(k)$. The same holds for Lemma 9.1.15. Finally, Lemma 9.2.2 gives the desired runtime. □

Proof of Theorem 9.2.8. The desired algorithm can be obtained as described in Theorem 9.2.7 disregarding the gluing operation. By Theorem 9.1.7, the relevant numbers are not too large and all necessary moduli can be computed explicitly as in Theorem 9.0.2. □

9.3 Deciding Exact Feasibility of Lasserre Relaxations in Polynomial Time

In this section, we show that the feasibility of the Lasserre semidefinite program $\mathrm{L}^t(G, H)$ from Definition 2.6.9 can be decided in randomised polynomial time. If non-negativity constraints are imposed, then feasibility can be decided in deterministic polynomial time.

The Lasserre semidefinite program can be solved approximately in polynomial time using e.g. the ellipsoid method. How to decide *exact* feasibility efficiently is generally unknown [13, 89].

By Theorems 5.1.1 and 5.1.2, the system $\mathrm{L}^t(G, H)$ has a (non-negative) real solution if, and only if, G and H are homomorphism indistinguishability over the graph class $\mathcal{L}_t$ (the graph class $\mathcal{L}_t^+$) as defined in Definition 5.2.1. Since $\mathcal{L}_t$ and $\mathcal{L}_t^+$ are minor-closed and of treewidth at most $3t - 1$, cf. Theorem 5.2.12 and Lemma 5.2.8, Theorem 9.0.1 immediately yields a randomised polynomial-time algorithm for each level of these hierarchies. However, it is not clear how to compute the data describing $\mathcal{L}_t$ and $\mathcal{L}_t^+$ given t. The following Theorems 9.3.1 and 9.3.2 overcome this problem by making the dependence on the parameter t effective.

Lasserre **Input** Simple graphs G and H, an integer $t \geq 1$. **Parameter** t. **Question** Does $\mathrm{L}^t(G, H)$ have a real solution?

Theorem 9.3.1. *There exists a randomised algorithm for* Lasserre *of runtime* $t^{O(1)} n^{O(t)}$ *for* $n := \max\{|V(G)|, |V(H)|\}$ *which accepts all* YES*-instances and accepts* NO*-instances with probability less than one half.*

NonNegLasserre **Input** Simple graphs G and H, an integer $t \geq 1$. **Parameter** t. **Question** Does $\mathrm{L}^t(G, H)$ have a non-negative real solution?

Theorem 9.3.2. *There exists a deterministic algorithm for* NonNegLasserre *of runtime* $t^{O(1)} n^{O(t)}$ *for* $n := \max\{|V(G)|, |V(H)|\}$.

Whereas the proof of Theorem 9.3.1 is based the techniques developed for Theorem 9.2.7, the proof of Theorem 9.3.2 deviates from these arguments. Thereby, Theorem 9.3.2 suggests a possibility to avoid randomness which does not rely on polynomial witness functions as in Theorem 9.0.2.

9.3.1 Lasserre without Non-Negativity Constraints

Theorem 9.3.1 is proven by modifying Theorem 9.2.7. By Theorem 5.1.1, deciding Lasserre amounts to deciding homomorphism indistinguishability over the graph class $\mathcal{L}_t$. As in Section 9.2.3, we define a set $\mathfrak{L}_t$ of formal expressions which resembles the set $\mathcal{L}(t, t)$ of (t, t)-bilabelled graphs defined in Definition 5.2.1. The difference is that $\mathfrak{L}_t$ retains information about how the bilabelled graphs in $\mathcal{L}(t, t)$ are constructed.

Write $\mathfrak{L}_t$ for the following inductively defined set of formal expressions. We simultaneously define a complexity measure $\mathrm{depth}\colon \mathfrak{L}_t \to \mathbb{N}$ on this set.

1. if $\boldsymbol{A} \in \mathcal{A}(t, t)$ is atomic, then $\boldsymbol{A} \in \mathfrak{L}_t$ and $\mathrm{depth}(\boldsymbol{A}) := 1$,
2. if $w \in \mathfrak{L}_t$ and $\boldsymbol{A}$ is atomic, then $\boldsymbol{A} \odot w \in \mathfrak{L}_t$ and $\mathrm{depth}(\boldsymbol{A} \odot w) := \mathrm{depth}(w)$,
3. if $w \in \mathfrak{L}_t$ and $\sigma \in \mathfrak{S}_{2t}$, then $w^\sigma \in \mathfrak{L}_t$ and $\mathrm{depth}(w^\sigma) := \mathrm{depth}(w)$,
4. if $w_1, w_2 \in \mathfrak{L}_t$, then $w_1 \cdot w_2 \in \mathfrak{L}_t$ and

$$\mathrm{depth}(w_1 \cdot w_2) := \max\{\mathrm{depth}(w_1), \mathrm{depth}(w_2)\} + 1.$$

Clearly, each expression $w \in \mathfrak{L}_t$ can be associated with the (t, t)-bilabelled graph $\mathrm{val}(w) \in \mathcal{L}(t, t)$ it encodes.

As in Section 9.1, we first give an upper bound on the size of the graphs whose homomorphism counts need to be considered in order to decide Lasserre. In other words, we give a witness function for $\mathcal{L}_t$.

Theorem 9.3.3. *For $t \geq 1$ and simple graphs G and H on at most n vertices with* $f_t(n) := 2t \cdot 4^{n^{2t}}$,

$$G \equiv_{\mathcal{L}_t} H \iff G \equiv_{(\mathcal{L}_t)_{\leq f_t(n)}} H.$$

Towards Theorem 9.3.3, we make the following observations:

Lemma 9.3.4. *Let $t \geq 1$. If $w \in \mathfrak{L}_t$, then $\mathrm{val}(w)$ is a bilabelled graph on at most* $2t \cdot 2^{\mathrm{depth}(w)}$ *vertices.*

Proof. The proof is by induction on the definition of $\mathfrak{L}_t$. If w is atomic, then $\operatorname{val}(w)$ has at most $2t$ vertices and $\operatorname{depth}(w) = 1$.

If $w = \boldsymbol{A} \odot w'$ for some atomic $\boldsymbol{A}$, then $\operatorname{val}(w) = \boldsymbol{A} \odot \operatorname{val}(w')$ has at most as many vertices as $\operatorname{val}(w')$ since all vertices of $\boldsymbol{A}$ are labelled. This implies the claim. The case $w = (w')^\sigma$ for $\sigma \in \mathfrak{S}_{2t}$ is analogous.

If $w = w_1 \cdot w_2$, then the number of vertices in $\operatorname{val}(w)$ is at most the number of vertices in $\operatorname{val}(w_1)$ plus the number of vertices in $\operatorname{val}(w_2)$. This number is by induction at most $2t \cdot (2^{\operatorname{depth}(w_1)} + 2^{\operatorname{depth}(w_2)}) \leq 2t \cdot 2 \cdot 2^{\operatorname{depth}(w)-1} = 2t \cdot 2^{\operatorname{depth}(w)}$. □

For two simple graphs G and H on at most n vertices, define S_d for $d \geq 1$ as the subspace of $\mathbb{R}^{(V(G)^t \cup V(H)^t) \times (V(G)^t \cup V(H)^t)}$ spanned by $\boldsymbol{L}_G \oplus \boldsymbol{L}_H := \begin{pmatrix} \boldsymbol{L}_G & 0 \\ 0 & \boldsymbol{L}_H \end{pmatrix}$ where $\boldsymbol{L} = \operatorname{val}(w)$ for $w \in \mathfrak{L}_t$ with $\operatorname{depth}(w) \leq d$.

Lemma 9.3.5. *If $d \geq 1$ is such that $S_d = S_{d+1}$, then $S_d = S$. In particular, $S_{2n^{2t}} = S$.*

Proof. We show by induction that $S_{d+i} \subseteq S_d$ for all $i \geq 1$. The base case follows by assumption. For the inductive step, we argue by structural induction that, for every $w \in \mathfrak{L}_t$ with $\operatorname{depth}(w) \leq d + i + 1$ and $\boldsymbol{L} := \operatorname{val}(w)$, it holds that $\boldsymbol{L}_G \oplus \boldsymbol{L}_H \in S_d$.

If w is atomic, then this clearly holds. If $w = \boldsymbol{A} \odot w'$, then by the inner induction the homomorphism tensor of $\operatorname{val}(w')$ is in S_d. The Schur product of any matrix in this space with $\boldsymbol{A}_G \oplus \boldsymbol{A}_H$ is in S_d by definition. Thus, the claim follows. The case $w = (w')^\sigma$ for $\sigma \in \mathfrak{S}_{2t}$ is analogous. If $w = w' \cdot w''$, then $\operatorname{depth}(w'), \operatorname{depth}(w'') < \operatorname{depth}(w) \leq d + i + 1$. By the outer inductive hypothesis, $S_{d+i} \subseteq S_d$. Thus, writing $\boldsymbol{L}' := \operatorname{val}(w')$ and $\boldsymbol{L}'' := \operatorname{val}(w'')$, it holds that $\boldsymbol{L}'_G \oplus \boldsymbol{L}'_H, \boldsymbol{L}''_G \oplus \boldsymbol{L}''_H \in S_d$. Thus, writing $\boldsymbol{L} := \operatorname{val}(w)$, $\boldsymbol{L}_G \oplus \boldsymbol{L}_H = (\boldsymbol{L}'\boldsymbol{L}'')_G \oplus (\boldsymbol{L}'\boldsymbol{L}'')_H = (\boldsymbol{L}'_G \oplus \boldsymbol{L}'_H)(\boldsymbol{L}''_G \oplus \boldsymbol{L}''_H) \in S_{d+1} \subseteq S_d$.

For the final claim, observe that the dimension of $S := \bigcup_{d \geq 1} S_d$ is at most $2n^{2t}$. This is because all matrices in S are of shape $\left(\begin{smallmatrix} * & 0 \\ 0 & * \end{smallmatrix}\right)$ with two non-zero blocks of size $n^t \times n^t$. □

This concludes the preparations for the proof of Theorem 9.3.3.

Proof of Theorem 9.3.3. By Lemma 9.3.5, $S_d = S$ for $d := 2n^{2t}$. The space S_d is spanned by $\boldsymbol{L}^i_G \oplus \boldsymbol{L}^i_H$ for some $\boldsymbol{L}^i \in \mathcal{L}_t$ on at most $2t \cdot 2^d$ vertices, by Lemma 9.3.4. For arbitrary $\boldsymbol{F} \in \mathcal{L}(t, t)$, there exist coefficients $\alpha_i \in \mathbb{R}$ such that $\boldsymbol{F}_G \oplus \boldsymbol{F}_H = \sum \alpha_i \boldsymbol{L}^i_G \oplus \boldsymbol{L}^i_H$. Hence,

$$\begin{aligned}
\hom(\operatorname{soe}(\boldsymbol{F}), G) &= \boldsymbol{1}_G^T (\boldsymbol{F}_G \oplus \boldsymbol{F}_H) \boldsymbol{1}_G \\
&= \sum \alpha_i \boldsymbol{1}_G^T (\boldsymbol{L}^i_G \oplus \boldsymbol{L}^i_H) \boldsymbol{1}_G \\
&= \sum \alpha_i \hom(\operatorname{soe}(\boldsymbol{L}^i), G) \\
&= \sum \alpha_i \hom(\operatorname{soe}(\boldsymbol{L}^i), H) \\
&= \hom(\operatorname{soe}(\boldsymbol{F}), H).
\end{aligned}$$ □

We now describe Algorithm 9.3 for Lasserre and prove Theorem 9.3.1.

Algorithm 9.3: Modular LASSERRE.

Input: Simple graphs G and H, an integer $t \geq 1$, a prime p in binary.
Output: Whether $G \equiv_{\mathcal{L}_t}^{p} H$.
for every $\boldsymbol{A} \in \mathcal{A}(t,t)$, compute the homomorphism matrices
  $\boldsymbol{A}_G \in \mathbb{F}_p^{V(G)^t \times V(G)^t}$ and $\boldsymbol{A}_H \in \mathbb{F}_p^{V(H)^t \times V(H)^t}$;
initialise $B \leftarrow \{\boldsymbol{J}_G \oplus \boldsymbol{J}_H\} \subseteq \mathbb{F}_p^{(V(G)^t \cup V(H)^t) \times (V(G)^t \cup V(H)^t)}$;
repeat
  foreach $\boldsymbol{A} \in \mathcal{A}(t,t)$, $v \in B$ **do**
    $w \leftarrow (\boldsymbol{A}_G \oplus \boldsymbol{A}_H) \odot v$;
    if $w \notin \operatorname{span}(B)$ **then**
      add w to B;
  foreach $v_1, v_2 \in B$ **do**
    $w \leftarrow v_1 \cdot v_2$;
    if $w \notin \langle B \rangle$ **then**
      add w to B;
  foreach $1 \leq i < j \leq 2t$, $v \in B$ **do**
    $w \leftarrow v^{(i\ j)}$;
    if $w \notin \operatorname{span}(B)$ **then**
      add w to B;
until *B is not updated*;
if $\boldsymbol{1}_G^T v \boldsymbol{1}_G = \boldsymbol{1}_H^T v \boldsymbol{1}_H$ *for all* $v \in B$ **then**
  accept;
else
  reject;

Proof of Theorem 9.3.1. Correctness follows as in Lemma 9.2.3 for Algorithm 9.1. For the runtime, note that the main loop is entered at most $2n^{2t}$ times. The subloops are entered at most $4t(2t-1)2n^{2t} + 4n^{4t}$ times. The linear-algebraic operations can be performed in time polynomial in $2n^{2t} \log p$. Thus Algorithm 9.3 has runtime $t^{O(1)} n^{O(t)} (\log p)^{O(1)}$. To close the gap between modular homomorphism indistinguishability over $\mathcal{L}_t$ and LASSERRE, invoke Lemma 9.2.6 and Theorem 9.3.3. □

9.3.2 Lasserre with Non-Negativity Constraints

This section is dedicated to proving Theorem 9.3.2. We show that NONNEGLASSERRE lies in the parametrised complexity class XP by showing that it can be decided using a colouring algorithm akin to the Weisfeiler–Leman algorithm, cf. Definition 2.3.1. The *matrix Weisfeiler–Leman* or mwl algorithm has polynomial runtime for every fixed dimension.

The colourings computed by this algorithm should be regarded as bases of spaces of homomorphism matrices as in Algorithm 9.1. Each colour class corresponds to a colour class indicator vector whose values are 0 and 1. Since the graph class $\mathcal{L}^+(t,t)$ is closed under parallel composition, the space of homomorphism matrices of $\mathcal{L}^+(t,t)$ with respect to the input graphs G and H has a basis comprising such indicator vectors, cf. Fact 2.4.6. That parallel composition/gluing allows to interpolate colours from homomorphism matrices was observed by Dvořák [63] and demonstrated already in Theorem 3.4.1. Since colour class indicator vectors occupy only polynomial space, NONNEGLASSERRE can be placed in XP (and thus in polynomial time for every fixed t).

Definition 9.3.6. Let $t \geq 1$. For a simple graph G, $\boldsymbol{r}, \boldsymbol{s} \in V(G)^t$, and an integer $i \geq 1$, define

$$\begin{aligned}
\mathrm{mwl}_t^0(G,\boldsymbol{r},\boldsymbol{s}) &:= \mathrm{atp}(G,\boldsymbol{r},\boldsymbol{s}),\\
\mathrm{mwl}_t^{i-1/2}(G,\boldsymbol{r},\boldsymbol{s}) &:= \left(\mathrm{mwl}_t^{i-1}(G,\sigma(\boldsymbol{r},\boldsymbol{s})) \;\middle|\; \sigma \in \{(i\ j) \mid 1 \leq i \leq j \leq 2t\} \subseteq \mathfrak{S}_{2t}\right),\\
\mathrm{mwl}_t^{i}(G,\boldsymbol{r},\boldsymbol{s}) &:= \Big(\mathrm{mwl}_t^{i-1/2}(G,\boldsymbol{r},\boldsymbol{s}),\\
&\qquad \left\{\!\!\left\{ \left(\mathrm{mwl}_t^{i-1/2}(G,\boldsymbol{r},\boldsymbol{t}), \mathrm{mwl}_t^{i-1/2}(G,\boldsymbol{t},\boldsymbol{s})\right) \;\middle|\; \boldsymbol{t} \in V(G)^t \right\}\!\!\right\} \Big).
\end{aligned}$$

The mwl_t^i for $i \in \mathbb{N}$ define increasingly fine[15] colourings of $V(G)^{2t}$. Let mwl_t^∞ denote the finest such colouring. Two simple graphs G and H are not *distinguished by the t-dimensional* mwl *algorithm* if

$$\left\{\!\!\left\{ \mathrm{mwl}_t^\infty(G,\boldsymbol{r},\boldsymbol{s}) \;\middle|\; \boldsymbol{r},\boldsymbol{s} \in V(G)^t \right\}\!\!\right\} = \left\{\!\!\left\{ \mathrm{mwl}_t^\infty(H,\boldsymbol{u},\boldsymbol{v}) \;\middle|\; \boldsymbol{u},\boldsymbol{v} \in V(H)^t \right\}\!\!\right\}.$$

[15]In the 1/2-step, we think of the tuple as a function from elements of $\mathfrak{S}_{2t}$ to colours. In particular, id $\in \mathfrak{S}_{2t}$ is assigned the original colour. Hence, $\mathrm{mwl}_t^{i-1/2}$ refines mwl_t^{i-1}.

Since the finest colouring mwl_t^∞ is reached in $\leq n^{2t} - 1$ iterations for graphs on n vertices, for fixed t, it can be tested in polynomial time whether two graphs are distinguished by the t-dimensional mwl algorithm. We are about to show that the latter happens if, and only if, the level-t Lasserre relaxation with non-negative constraints is feasible. As a by-product, we obtain a logical characterisation for this equivalence relation akin to Theorem 2.3.2.

Definition 9.3.7. For $t \geq 1$, an M^t*-formula* has $2t$ free variables and is of the following form:

- every quantifier-free FO-formula over the signature $\{E\}$ with $2t$ variables is an M^t-formula,
- if φ, ψ are M^t-formulas with the same free variables, then $\neg\varphi$, $\varphi \wedge \psi$, and $\varphi \vee \psi$ are M^t-formulas,
- if φ, ψ are M^t-formulas and $n \in \mathbb{N}$, then $\exists^{\geq n} \boldsymbol{y} \, (\varphi(\boldsymbol{x}, \boldsymbol{y}) \wedge \psi(\boldsymbol{y}, \boldsymbol{z}))$ is an M^t-formula. Here, the boldface letters $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ denote pairwise disjoint t-tuples of distinct variables.

An M^t*-sentence* is an expression $\exists^{\geq n} \boldsymbol{x} \varphi(\boldsymbol{x})$ where φ is an M^t-formula, $\boldsymbol{x}$ is a tuple of $2t$ distinct variables, and $n \in \mathbb{N}$.

The semantics of the quantifier $\exists^{\geq n} \boldsymbol{y} \varphi(\boldsymbol{y})$ is that there exist at least n many $2t$-tuples of vertices from the graph over which the formula is evaluated that satisfy φ.

Theorem 9.3.8. *Let $t \geq 1$. For simple graphs G and H, the following are equivalent:*

1. *G and H are not distinguished by the t-dimensional* mwl *algorithm,*
2. *G and H are homomorphism indistinguishable over $\mathcal{L}_t^+$,*
3. *G and H satisfy the same M^t-sentences.*

Assuming Theorem 9.3.8, we derive Theorem 9.3.2.

Proof of Theorem 9.3.2 assuming Theorem 9.3.8. By the observation following Definition 9.3.6, the t-dimensional mwl algorithm stabilises after n^{2t} iterations on graphs on n vertices. In step $i + 1/2$, we compute for every $\boldsymbol{r}, \boldsymbol{s} \in V(G)^t$, the tuple

$$\text{mwl}_t^{i-1/2}(G, \boldsymbol{r}, \boldsymbol{s}) := \left(\text{mwl}_t^{i-1}(G, \sigma(\boldsymbol{r}, \boldsymbol{s})) \;\middle|\; \sigma \in \{(i \; j) \mid 1 \leq i \leq j \leq 2t\} \subseteq \mathfrak{S}_{2t} \right).$$

Then, we sort these tuples for every $\boldsymbol{r}, \boldsymbol{s} \in V(G)^t$ and use the indices in this sorted list as new colour $\text{mwl}_t^{i-1/2}(G, \boldsymbol{r}, \boldsymbol{s})$. In total, this takes time $t^{O(1)} n^{O(t)}$. In the step $i + 1$, we proceed similarly, also taking time $t^{O(1)} n^{O(t)}$. □

We note that the mwl algorithm can be implemented more efficiently by adapting efficient implementations of the Weisfeiler–Leman algorithm [93]. These are beyond the scope of this thesis.

The proof of Theorem 9.3.8 is conceptually similar to arguments of [37, 63]. It is implied by the following Theorem 9.3.9:

Theorem 9.3.9. *Let $t \geq 1$. For simple graphs G and H with $\boldsymbol{r}, \boldsymbol{s} \in V(G)^t$ and $\boldsymbol{u}, \boldsymbol{v} \in V(H)^t$, the following are equivalent:*

1. $\mathrm{mwl}^\infty_G(\boldsymbol{r}, \boldsymbol{s}) = \mathrm{mwl}^\infty_H(\boldsymbol{u}, \boldsymbol{v})$,
2. $\boldsymbol{F}_G(\boldsymbol{r}, \boldsymbol{s}) = \boldsymbol{F}_H(\boldsymbol{u}, \boldsymbol{v})$ *for all* $\boldsymbol{F} \in \mathcal{L}^+(t,t)$*, and*
3. $G \models \varphi(\boldsymbol{r}, \boldsymbol{s})$ *if, and only if,* $H \models \varphi(\boldsymbol{u}, \boldsymbol{v})$ *for all* M^t*-formulas* φ.

The proof of Theorem 9.3.9 is based on the following lemma, which is adopted from [63, Lemma 6]. Recall from Section 3.2.3 that $\mathbb{R}\mathcal{L}^+(t,t)$ denotes the set of formal finite $\mathbb{R}$-linear combinations of bilabelled graphs from $\mathcal{L}^+(t,t)$.

Lemma 9.3.10. *Let $t \geq 1$ and $n \in \mathbb{N}$. For every M^t-formula φ, there exists an $\boldsymbol{a} \in \mathbb{R}\mathcal{L}^+(t,t)$ such that, for all simple graphs G on at most n vertices and $\boldsymbol{r}, \boldsymbol{s} \in V(G)^t$,*

- *if $G \models \varphi(\boldsymbol{rs})$, then $\boldsymbol{a}_G(\boldsymbol{r}, \boldsymbol{s}) = 1$, and*
- *if $G \not\models \varphi(\boldsymbol{rs})$, then $\boldsymbol{a}_G(\boldsymbol{r}, \boldsymbol{s}) = 0$.*

Proof. If φ is a quantifier-free formula, then there exists an atomic graph $\boldsymbol{F} \in \mathcal{A}(t,t)$ such that $G \models \varphi(\boldsymbol{r}, \boldsymbol{s})$ if, and only if, $\boldsymbol{F}_G(\boldsymbol{r}, \boldsymbol{s}) = 1$. By Observation 3.2.11, the homomorphism tensor of any atomic graph has entries from $\{0, 1\}$.

If φ is of the form $\neg\psi$, let $\boldsymbol{a}$ denote the element of $\mathbb{R}\mathcal{L}^+(t,t)$ constructed inductively for ψ. Then $\boldsymbol{b} := \boldsymbol{J} - \boldsymbol{a}$ for $\boldsymbol{J}$ as defined in Definition 3.2.10 is as desired.

If φ is of the form $\psi \wedge \chi$ where ψ and χ have the same free variables as φ, let $\boldsymbol{a}$ and $\boldsymbol{b}$ denote the elements of $\mathbb{R}\mathcal{L}^+(t,t)$ constructed inductively for ψ and χ, respectively. Then $\boldsymbol{c} := \boldsymbol{a} \odot \boldsymbol{b} \in \mathbb{R}\mathcal{L}^+(t,t)$ is as desired.

If φ is of the form $\psi \vee \chi$ where ψ and χ have the same free variables as φ, then φ is equivalent to $\neg(\neg\psi \wedge \neg\chi)$ and the two previous cases can be applied jointly.

It remains to consider the case in which φ is of the form $\exists^{\geq \ell} \boldsymbol{y} \psi(\boldsymbol{x}, \boldsymbol{y}) \wedge \chi(\boldsymbol{y}, \boldsymbol{z})$. Let $\boldsymbol{a}$ and $\boldsymbol{b}$ denote the elements of $\mathbb{R}\mathcal{L}^+(t,t)$ constructed inductively for ψ and χ, respectively. Then, by the inductive hypothesis,

$$(\boldsymbol{a} \cdot \boldsymbol{b})_G(\boldsymbol{r}, \boldsymbol{s}) = \sum_{\boldsymbol{t} \in V(G)^t} \boldsymbol{a}_G(\boldsymbol{r}, \boldsymbol{t}) \boldsymbol{b}_G(\boldsymbol{t}, \boldsymbol{s})$$

is equal to the number of elements $\boldsymbol{t} \in V(G)^t$ such that $G \models \psi(\boldsymbol{r}, \boldsymbol{t}) \wedge \chi(\boldsymbol{t}, \boldsymbol{s})$.

Let $P = \sum c_i x^i \in \mathbb{R}[x]$ be a polynomial which evaluates to 0 on $\{0, 1, \dots, \ell - 1\}$ and to 1 on $\{\ell, \ell + 1, \dots, n^t\}$. For instance, P can be taken to be the Lagrange polynomial from Lemma 8.3.5. Consider $\boldsymbol{c} := P(\boldsymbol{a} \cdot \boldsymbol{b}) = \sum c_i (\boldsymbol{a} \cdot \boldsymbol{b})^{\odot i} \in \mathbb{R}\mathcal{L}^+(t,t)$ where $(\boldsymbol{a} \cdot \boldsymbol{b})^{\odot i}$ denotes the parallel composition of i copies of $\boldsymbol{a} \cdot \boldsymbol{b}$. It follows that $\boldsymbol{c}_G(\boldsymbol{r}, \boldsymbol{s})$ equals one if, and only if, the number of elements $\boldsymbol{t} \in V(G)^t$ such that $G \models \psi(\boldsymbol{r}, \boldsymbol{t}) \wedge \chi(\boldsymbol{t}, \boldsymbol{s})$ is in $\{\ell, \ell + 1, \dots, n^t\}$. It is zero otherwise. Thus, $\boldsymbol{c}$ is as desired. □

Proof of Theorem 9.3.9. Supposing Item 1, Item 2 is proven by induction on the structure of $\boldsymbol{F}$. If $\boldsymbol{F}$ is atomic, then the statement follows from $\mathrm{atp}(G, \boldsymbol{r}, \boldsymbol{s}) = \mathrm{atp}(H, \boldsymbol{u}, \boldsymbol{v})$. For $\boldsymbol{F} = \boldsymbol{K} \odot \boldsymbol{L}$ and $\boldsymbol{F} = \boldsymbol{K}^\sigma$ with $\boldsymbol{K}, \boldsymbol{L} \in \mathcal{L}^+(t,t)$ and $\sigma \in \mathfrak{S}_{2t}$, the statement is easily verified. It remains to consider the case $\boldsymbol{F} = \boldsymbol{K} \cdot \boldsymbol{L}$. By definition of mwl, there exists a bijection $\pi \colon V(G)^t \to V(H)^t$ such that

$$\mathrm{mwl}^\infty_G(\boldsymbol{r}, \boldsymbol{t}) = \mathrm{mwl}^\infty_H(\boldsymbol{u}, \pi(\boldsymbol{t})) \quad \text{and} \quad \mathrm{mwl}^\infty_G(\boldsymbol{t}, \boldsymbol{s}) = \mathrm{mwl}^\infty_H(\pi(\boldsymbol{t}), \boldsymbol{v})$$

for all $\boldsymbol{t} \in V(G)^t$. Hence,

$$\boldsymbol{F}_G(\boldsymbol{r}, \boldsymbol{s}) = \sum_{\boldsymbol{t} \in V(G)^t} \boldsymbol{K}_G(\boldsymbol{r}, \boldsymbol{t}) \boldsymbol{L}_G(\boldsymbol{t}, \boldsymbol{s}) = \sum_{\boldsymbol{t} \in V(G)^t} \boldsymbol{K}_H(\boldsymbol{u}, \pi(\boldsymbol{t})) \boldsymbol{L}_H(\pi(\boldsymbol{t}), \boldsymbol{v}) = \boldsymbol{F}_H(\boldsymbol{u}, \boldsymbol{v}).$$

Thus, Item 2 holds.

Now suppose that Item 2 holds. If φ is a M^t-formula such that $G \models \varphi(\boldsymbol{r}, \boldsymbol{s})$ and $H \not\models \varphi(\boldsymbol{u}, \boldsymbol{v})$, then, by Lemma 9.3.10, there exists a graph $\boldsymbol{F} \in \mathcal{L}^+(t, t)$ such that $\boldsymbol{F}_G(\boldsymbol{r}, \boldsymbol{s}) \neq \boldsymbol{F}_H(\boldsymbol{u}, \boldsymbol{v})$. This yields Item 3.

That Item 3 implies Item 1 is proven similarly as [37, Theorem 5.2] by induction on the number of iterations. Since atp can be defined using quantifier-free M^t-formulas, $\mathrm{mwl}^0_G(\boldsymbol{r}, \boldsymbol{s}) = \mathrm{mwl}^0_H(\boldsymbol{u}, \boldsymbol{v})$.

Since M^t is closed under permuting the names of the variables, it holds for all permutations $\sigma \in \mathfrak{S}_{2t}$ that $G \models \varphi(\boldsymbol{r}, \boldsymbol{s}) \iff H \models \varphi(\boldsymbol{u}, \boldsymbol{v})$ for all $\varphi \in \mathsf{M}^t$ with $2t$ free variables if, and only if, $G \models \varphi(\sigma(\boldsymbol{r}, \boldsymbol{s})) \iff H \models \varphi(\sigma(\boldsymbol{u}, \boldsymbol{v}))$ for all $\varphi \in \mathsf{M}^t$ with $2t$ free variables. Hence, if $\mathrm{mwl}^i_t(G, \boldsymbol{r}, \boldsymbol{s}) = \mathrm{mwl}^i_t(H, \boldsymbol{u}, \boldsymbol{v})$ for some $i \in \mathbb{N}$, then also $\mathrm{mwl}^{i+1/2}_t(G, \boldsymbol{r}, \boldsymbol{s}) = \mathrm{mwl}^{i+1/2}_t(H, \boldsymbol{u}, \boldsymbol{v})$.

For the step from $i + 1/2$ to $i + 1$, suppose contrapositively that

$$\mathrm{mwl}^{i+1}_t(G, \boldsymbol{r}, \boldsymbol{s}) \neq \mathrm{mwl}^{i+1}_t(H, \boldsymbol{u}, \boldsymbol{v}).$$

Since the case $\mathrm{mwl}^{i+1/2}_t(G, \boldsymbol{r}, \boldsymbol{s}) \neq \mathrm{mwl}^{i+1/2}_t(H, \boldsymbol{u}, \boldsymbol{v})$ was dealt with in the previous argument, we may suppose that these colours are equal.

Hence, there exists a pair of colours $\mathrm{mwl}^{i+1/2}_t(G, \boldsymbol{r}, \boldsymbol{t})$, $\mathrm{mwl}^{i+1/2}_t(G, \boldsymbol{t}, \boldsymbol{s})$ which appears in the multisets for G and H differently often, without loss of generality more often in G than in H. By the inductive hypothesis, for each pair of distinct $\mathrm{mwl}^{i+1/2}_t$-colours, there exists an M^t-formula φ which is satisfied by all vertex tuples of the first colour and by none of the second colour. By taking the conjunction of several such formulas, a formula can be constructed which holds for a $2t$-tuple of vertices of G or H if, and only if, they have a specified colour in $\mathrm{mwl}^{i+1/2}_t$.

Let φ and ψ be formulas which hold exactly for the $2t$-tuples of vertices of G or H of colours $\mathrm{mwl}^{i+1/2}_t(G, \boldsymbol{r}, \boldsymbol{t})$ and $\mathrm{mwl}^{i+1/2}_t(G, \boldsymbol{t}, \boldsymbol{s})$, respectively. By assumption, there is an $N \in \mathbb{N}$ such that the formula $\chi := \exists^{\geq N} \boldsymbol{y} \varphi(\boldsymbol{x}, \boldsymbol{y}) \land \psi(\boldsymbol{y}, \boldsymbol{z}) \in \mathsf{M}^t$ is such that $G \models \chi(\boldsymbol{r}, \boldsymbol{s})$ and $H \not\models \chi(\boldsymbol{u}, \boldsymbol{v})$. This yields Item 1. □

Finally, Theorem 9.3.8 is derived from Theorem 9.3.9.

Proof of Theorem 9.3.8. Supposing Item 1, let $\pi \colon V(G)^t \times V(G)^t \to V(H)^t \times V(H)^t$ be a bijection such that $\mathrm{mwl}^\infty_G(\boldsymbol{r}, \boldsymbol{s}) = \mathrm{mwl}^\infty_H(\pi(\boldsymbol{r}, \boldsymbol{s}))$ for all $\boldsymbol{r}, \boldsymbol{s} \in V(G)^t$. By

Theorem 9.3.9, for $\boldsymbol{F} = (F, \boldsymbol{u}, \boldsymbol{v}) \in \mathcal{L}^+(t, t)$,

$$\hom(F, G) = \operatorname{soe}(\boldsymbol{F}_G) = \sum_{\boldsymbol{r},\boldsymbol{s} \in V(G)^t} \boldsymbol{F}_G(\boldsymbol{r}, \boldsymbol{s}) = \sum_{\boldsymbol{r},\boldsymbol{s} \in V(G)^t} \boldsymbol{F}_H(\pi(\boldsymbol{r}, \boldsymbol{s})) = \hom(F, H),$$

so Item 2 holds.

Assuming Item 2 holds, let $\Phi = \exists^{\geq \ell} \boldsymbol{x} \varphi(\boldsymbol{x})$ be an M^t-sentence where φ is an M^t-formula, $\ell \in \mathbb{N}$ and $\boldsymbol{x}$ is a tuple of $2t$ distinct variables. Let $\boldsymbol{q} \in \mathbb{R}\mathcal{L}^+(t, t)$ be for φ and $n := \max\{|V(G)|, |V(H)|\}$ as in Lemma 9.3.10. Then Item 2 implies that

$$\begin{aligned} \left|\{\boldsymbol{rs} \in V(G)^{2t} \mid G \models \varphi(\boldsymbol{r}, \boldsymbol{s})\}\right| &= \sum_{\boldsymbol{rs} \in V(G)^{2t}} \boldsymbol{q}_G(\boldsymbol{r}, \boldsymbol{s}) = \operatorname{soe}(\boldsymbol{q}_G) \\ &= \left|\{\boldsymbol{uv} \in V(H)^{2t} \mid H \models \varphi(\boldsymbol{u}, \boldsymbol{v})\}\right|. \end{aligned}$$

Hence, $G \models \Phi$ if, and only if, $H \models \Phi$. This yields Item 3.

Assuming Item 3 holds, suppose that G and H are distinguished by the mwl algorithm and let $C \subseteq V(G)^{2t}$ denote an mwl-colour class in G whose counterpart $D \subseteq V(H)^{2t}$ has different size. By Theorem 9.3.9, there exists an M^t-formula φ which is satisfied by tuples in C and D and by no other tuples. The M^t-sentence $\exists^{\geq \ell} \boldsymbol{x} \varphi(\boldsymbol{x})$ is not satisfied by both G and H for a suitable $\ell \in \mathbb{N}$. This yields Item 1. □

9.4 Lower Bounds

In this section, two hardness results for the problem HomInd are established. In both cases, we show hardness for families of minor-closed graph classes. The approaches are orthogonal in the sense that the reduction yielding coNP-hardness in Theorem 9.4.1 is from a fixed-parameter tractable problem while the reduction yielding coW[1]-hardness in Theorem 9.4.2 is not polynomial-time.

9.4.1 coNP-Hardness

The first hardness result concerns deciding whether two graphs are indistinguishable under the k-dimensional Weisfeiler–Leman algorithm wl_k when k is part of the input.[16]

Theorem 9.4.1. *The problem of deciding given two simple graphs G and H and an integer $k \in \mathbb{N}$ whether G and H are not distinguished by wl_k is* coNP*-hard under polynomial-time many-one reductions.*

The k-dimensional Weisfeiler–Leman algorithm has implementations of runtime $O(k^2 n^{k+1} \log n)$ [93], which is exponential when regarding k as part of the input. Grohe [77] showed that when k is fixed, then deciding indistinguishability under

[16] Recently, the same result was independently obtained by Lichter, Raßmann, & Schweitzer [112].

the k-dimensional Weisfeiler–Leman algorithm is PTIME-complete. Establishing lower bounds on the complexity of the problem in Theorem 9.4.1 is a challenging problem [20, 22, 82]. Theorem 9.4.1 is a first step towards resolving a question posed by Berkholz [20]: Is the decision problem in Theorem 9.4.1 EXPTIME-complete?

By Theorem 3.4.3, wl_k indistinguishability coincides with homomorphism indistinguishability over the class of graphs of treewidth at most k. Hence, the problem in Theorem 9.4.1 is clearly a special case of HomInd, i.e. with φ set to true. Thus, Theorem 9.4.1 implies that, when disregarding the parametrisation, HomInd is coNP-hard under polynomial-time many-one reductions.

Theorem 9.4.1 is obtained by reducing the NP-complete problem of deciding whether a graph of bounded degree has treewidth $\leq k$ [28, 27]. The reduction is based on the CFI construction, cf. Section 6.3.1.

Proof of Theorem 9.4.1. Recall that, for a graph G, $\Delta(G)$ denotes its maximum vertex degree. The following problem is NP-complete by [28, Theorem 11]. See also [27].

BoundedDegreeTreewidth
Input A simple graph G with $\Delta(G) \leq 9$, an integer k
Question Is the treewidth of G at most k?

By deleting isolated vertices, we may suppose that every connected component of G contains at least two vertices. If G has multiple connected components, take one vertex from each component and connect them in a path-like fashion. This increases the maximum degree potentially by one but makes the graph connected. The treewidth is invariant under this operation. Thus, we may suppose that G is connected and $\Delta(G) \leq 10$.

Given such an instance, we produce the instance (G_0, G_1, k) of WL, i.e. the problem in Theorem 9.4.1. Here, G_0 and G_1 are the even and odd CFI graphs of G, cf. Definition 6.3.6.

Then G_0 and G_1 are not distinguished by wl_k if, and only if, $\text{tw}(G) \geq k+1$. Indeed, by Theorems 6.4.1, 3.4.3, and 6.3.8, since G is connected, if $\text{tw}(G) \geq k+1$, then G_0 and G_1 are not distinguished by wl_k. Conversely, if $\text{tw}(G) < k+1$, then G_0 and G_1 are distinguished by wl_k since $\hom(G, G_0) \neq \hom(G, G_1)$ by Corollary 6.3.5 and Theorem 3.4.3.

Hence, (G,k) is a YES-instance of BoundedDegreeTreewidth if, and only if, (G_0, G_1, k) is a NO-instance of WL. The graphs G_0 and G_1 are of size $\sum_{v \in V(G)} 2^{\deg(v)-1} \leq 2^9 |V(G)|$, which is polynomial in the input. □

9.4.2 coW[1]-Hardness

The second hardness result concerns HomInd as a parametrised problem. We show that it is coW[1]-hard and that the runtime in Theorem 9.2.7 is optimal under the Exponential Time Hypothesis (ETH).

Theorem 9.4.2. HomInd *is* coW[1]*-hard under fpt-reductions. Unless* ETH *fails, there is no algorithm for* HomInd *that runs in time* $f(|\varphi| + k)n^{o(|\varphi|+k)}$ *for any computable function* $f: \mathbb{N} \to \mathbb{N}$.

We show the assertions of Theorem 9.4.2 already for the problem HomIndSize defined below. Write $\mathcal{G}_{\leq k}$ for the class of all graphs on at most k vertices.

> HomIndSize
> **Input** Simple graphs G and H, an integer $k \geq 1$.
> **Parameter** k.
> **Question** Are G and H homomorphism indistinguishable over the class $\mathcal{G}_{\leq k}$?

The problem HomIndSize fixed-parameter reduces to HomInd. To that end, consider the first-order formula $\varphi_k := \exists x_1 \dots \exists x_k \forall y \bigvee_{i=1}^{k} (y = x_i)$ for $k \in \mathbb{N}$. Then, a graph models φ_k if, and only if, it has at most k vertices. Furthermore, $|\varphi_k| = O(k)$. Hence, transforming the instance (G, H, k) of HomIndSize to the instance $(G, H, \varphi_k, k - 1)$ of HomInd gives the desired reduction. Since $|\varphi_k| + k = O(k)$, it suffices to show the assertions of Theorem 9.4.2 for HomIndSize.

Proof of Theorem 9.4.2. The proof is by reduction from the parametrised clique problem Clique, which is well-known to be W[1]-complete and which does not admit an $f(k)n^{o(k)}$-time algorithm for any computable function $f: \mathbb{N} \to \mathbb{N}$ unless ETH fails [51, Theorems 13.25, 14.21].

Let K denote the k-vertex complete graph and K_0 and K_1 the even and odd CFI graphs of K. By Theorem 6.3.8, $K_0 \equiv_{\mathcal{G}_{\leq k} \setminus \{K\}} K_1$. Indeed, no graph in $\mathcal{G}_{\leq k} \setminus \{K\}$ admits a weak oddomorphism to K since weak oddomorphisms are surjective on edges and vertices.

The reduction produces, given the instance (G, k) of Clique, the instance $(G \times K_0, G \times K_1, k)$ of HomIndSize where $\times$ denotes the categorical product of two graphs, cf. Section 2.1.1. Producing this instance is fixed-parameter tractable. Furthermore, the parameter k is not affected by this reduction. It remains to show correctness.

If $G \times K_0 \equiv_{\mathcal{G}_{\leq k}} G \times K_1$, then $\hom(K, G) = 0$. Indeed, by assumption and Equation (2.2), $\hom(K, G) \hom(K, K_0) = \hom(K, G \times K_0) = \hom(K, G \times K_1) = \hom(K, G) \hom(K, K_1)$. However, $\hom(K, K_0) \neq \hom(K, K_1)$ by Corollary 6.3.5, and thus $\hom(K, G) = 0$.

Conversely, it holds that $K_0 \equiv_{\mathcal{G}_{\leq k} \setminus \{K\}} K_1$ and hence also $G \times K_0 \equiv_{\mathcal{G}_{\leq k} \setminus \{K\}} G \times K_1$ by Lemma 7.1.1. Since $\hom(K, G) = 0$, also $G \times K_0 \equiv_{\mathcal{G}_{\leq k}} G \times K_1$. □

9.4.3 Reductions between Homomorphism Indistinguishability Problems

Few lower bounds are known for the complexity of homomorphism indistinguishability problems. The most prominent example is the class $\mathcal{P}$ of all planar graphs, for which HomInd($\mathcal{P}$) is undecidable [124, 15]. It would be desirable to extend this result to other graph classes by finding, more generally, for two graph classes $\mathcal{F}_1$ and $\mathcal{F}_2$, reductions between their homomorphism indistinguishability problems HomInd($\mathcal{F}_1$) and HomInd($\mathcal{F}_2$). In this section, the few known examples for such reductions are listed.

Taking Complements

The first reduction is by taking complements of the input graphs, cf. Theorem 7.0.1.

Lemma 9.4.3. *Let $\mathcal{C}$ denote the class of all cycles and $\mathcal{CP}$ the class of all cycles and paths. Then* HomInd($\mathcal{CP}$) *polynomial-time Turing reduces to* HomInd($\mathcal{C}$)*.*

Proof. By Theorem 3.3.2 and Corollary 3.3.3, for simple graphs G and H it holds that $G \equiv_{\mathcal{CP}} H$ if, and only if, $G \equiv_{\mathcal{C}} H$ and $\overline{G} \equiv_{\mathcal{C}} \overline{H}$. Thus HomInd($\mathcal{CP}$) on input (G, H) can be decided by calling an oracle for HomInd($\mathcal{C}$) for the instances (G, H) and $(\overline{G}, \overline{H})$. □

Since HomInd($\mathcal{C}$) and HomInd($\mathcal{CP}$) are both in polynomial time, the statement of the following lemma is in a strict sense trivial. It is stated nevertheless as it directly relates two homomorphism indistinguishability problems. The graph classes $\mathcal{C}$ and $\mathcal{CP}$ are the only graph classes we are aware of to which the above idea applies. It would be interesting to generalise this construction.

Question 9.4.4. *For which graph classes $\mathcal{F}$ is the equivalence relation $G \approx H$ defined via $G \equiv_{\mathcal{F}} H$ and $\overline{G} \equiv_{\mathcal{F}} \overline{H}$ a homomorphism indistinguishability relation?*

Taking Categorical Products

The second reduction is by taking the categorical product of the input graphs with a fixed graph K, i.e. by transforming an input graph G to $G \times K$. Recall that, for a graph class $\mathcal{F}$ and a graph K, $\mathcal{F}_K$ denotes the class of all K-colourable graphs in $\mathcal{F}$. As observed in Section 7.3, $G \equiv_{\mathcal{F}_K} H$ if, and only if, $G \times K \equiv_{\mathcal{F}} H \times K$ for all simple graphs G, H, and K. This yields the following Lemma 9.4.5.

Lemma 9.4.5. *Let $\mathcal{F}$ be a graph class and K be a simple graph. Then* HomInd($\mathcal{F}_K$) *polynomial-time many-one reduces to* HomInd($\mathcal{F}$)*.*

By Theorem 7.3.5, if $\mathcal{F}$ is closed under subdivision and K is non-bipartite, then $\equiv_{\mathcal{F}}$ and $\equiv_{\mathcal{F}_K}$ coincide. Thus, one typically needs to assume K to be bipartite in order for Lemma 9.4.5 to yield a reduction between distinct homomorphism indistinguishability problems. For example, by Theorem 6.1.1, homomorphism indistinguishability over $\mathcal{G}_{K_2}$, the class of all bipartite graphs, is not isomorphism. Nevertheless, HomInd($\mathcal{G}_{K_2}$) is equivalent to the graph isomorphism problem under polynomial-time many-one reductions.

Theorem 9.4.6. HomInd($\mathcal{G}_{K_2}$) *and* HomInd($\mathcal{G}$) *are polynomial-time many-one interreducible.*

Proof. The lemma is implied by [144, Corollary 4.4] and Corollary 3.1.3. Alternatively, consider the following direct proof. By Lemma 9.4.5, HomInd($\mathcal{G}_{K_2}$) polynomial-time many-one reduces to HomInd($\mathcal{G}$), i.e. the graph isomorphism problem.

By [179, p. 1474], graph isomorphism polynomial-time many-one reduces to the problem of deciding whether two bipartite graphs are isomorphic. Indeed, given an instance (G, H) for graph isomorphism, we may construct (G', H') by subdividing every edge of G and H once. The resulting graphs are bipartite and it holds that $G' \cong H'$ if, and only if, $G \cong H$.

We further reduce to HomInd($\mathcal{G}_{K_2}$). By Lemma 7.3.4 and Corollary 3.1.3, the graphs G' and H' are isomorphic if, and only if, they are homomorphism indistinguishable over all bipartite graphs. Thus, graph isomorphism for bipartite graphs polynomial-time many-one reduces to HomInd($\mathcal{G}_{K_2}$). □

9.5 Essentially Finite and Profinite Graph Classes

In this section, we show that, for all essentially finite classes $\mathcal{F}$, HomInd$(\mathcal{F})$ is in polynomial time. For essentially profinite $\mathcal{F}$, the problem can be arbitrarily hard. In particular, as demonstrated by Böker, Chen, Grohe, & Rattan [33], HomInd($\mathcal{K}$) is $C_=P$-complete for the class $\mathcal{K}$ of all complete graphs.

Theorem 9.5.1. *For a graph class $\mathcal{F}$, the following are equivalent:*

1. *$\mathcal{F}$ is essentially finite.*
2. *$\mathcal{F}$ has a constant witness function, i.e. there exists a constant $C \in \mathbb{N}$ such that, for all simple graphs G and H, it holds that $G \equiv_{\mathcal{F}} H$ if, and only if, $G \equiv_{\mathcal{F}_{\leq C}} H$.*

In particular, for every essentially finite graph class $\mathcal{F}$, HomInd$(\mathcal{F})$ *is in polynomial time.*

Proof. First assume that $\mathcal{F}$ is essentially finite. Recall the notation from Section 6.5. We construct a finite subclass $\mathcal{F}' \subseteq \mathcal{F}$ such that $\mathrm{cl}(\mathcal{F}') = \mathrm{cl}(\mathcal{F})$. Then every upper bound C on the size of the graphs in $\mathcal{F}'$ is as desired.

To that end, choose, for every $\Lambda \subseteq \Gamma(\mathcal{F})$, a finite set $F_1^\Lambda, \dots, F_\ell^\Lambda \in \mathcal{F}_L$ such that the vectors $\vec{F}_1^\Lambda, \dots, \vec{F}_\ell^\Lambda \in \mathbb{R}^{\Gamma(\mathcal{F})}$ span the finite-dimensional space $\mathrm{span}\{\vec{F} \in \mathbb{R}^{\Gamma(\mathcal{F})} \mid F \in \mathcal{F}_L\}$ where $L := \coprod_{C \in \Lambda} C$. Define $\mathcal{F}'$ by taking the union over all graphs $F_1^\Lambda, \dots, F_\ell^\Lambda \in \mathcal{F}_L$ constructed in this way. By construction, $\mathcal{F}' \subseteq \mathcal{F}$. Thus, it suffices to show that $\mathcal{F} \subseteq \mathrm{cl}(\mathcal{F}')$.

To that end, let $K \in \mathcal{F}$ be a simple graph. By Lemma 6.5.5, $K \in \mathrm{cl}(\mathcal{F}')$ if, and only if, $K \in \mathrm{cl}(\mathcal{F}'_K)$. Let $L := \coprod_{C \in \Lambda} C$ where $\Lambda := \Gamma(K)$. Since L and K are homomorphically equivalent, it holds that $\mathcal{F}'_K = \mathcal{F}'_L$. By construction, $\vec{K} \in \mathrm{span}\{\vec{F}_1^\Lambda, \dots, \vec{F}_\ell^\Lambda\} \subseteq \mathbb{R}^{\Gamma(\mathcal{F})}$ and, by projection, the analogous statement holds in $\mathbb{R}^\Delta$ for $\Delta := \Gamma(\{K, F_1^\Lambda, \dots, F_\ell^\Lambda\})$. By Theorem 6.5.6, $K \in \mathrm{cl}(\{F_1^\Lambda, \dots, F_\ell^\Lambda\}) \subseteq \mathrm{cl}(\mathcal{F}')$.

Conversely, suppose that $C \in \mathbb{N}$ is such that, for all simple graphs G and H, $G \equiv_{\mathcal{F}} H$ if, and only if, $G \equiv_{\mathcal{F}_{\leq C}} H$. Then $\mathrm{cl}(\mathcal{F}) = \mathrm{cl}(\mathcal{F}_{\leq C})$. Containing only finitely many graphs, $\mathcal{F}_{\leq C}$ is essentially finite. By Corollary 6.5.9, $\mathrm{cl}(\mathcal{F}_{\leq C})$ is essentially finite. Since $\mathcal{F} \subseteq \mathrm{cl}(\mathcal{F}) = \mathrm{cl}(\mathcal{F}_{\leq C})$, the graph class $\mathcal{F}$ is essentially finite as well. □

For an essentially profinite graph class $\mathcal{F}$, the problem HomInd($\mathcal{F}$) can be arbitrarily hard. For a set $S \subseteq \mathbb{N}$ of positive integers, write Mem(S) for the problem of deciding given an integer $n \geq 1$ whether $n \notin S$. Here, n is encoded unarily, i.e. the size of an instance is n. Note that (the complement of) any decision problem $L \subseteq \{0,1\}^*$ can be reduced to Mem(S) for some S at the expense of an exponential increase in the size of the instances. Recall the definition of the essentially profinite graph class $\mathcal{K}^S$ from Equation (6.4).

Theorem 9.5.2. *For every $S \subseteq \mathbb{N}$, the problem* Mem(S) *polynomial-time many-one reduces to* HomInd($\mathcal{K}^S$).

The proof of Theorem 9.5.2 is built on the following construction from [33].

Theorem 9.5.3 ([33, Corollary 14])**.** *For every $n \geq 1$, there exist graphs G_n and H_n on at most* $\max\{1, 2(n-1)\}$ *many vertices such that, for all $\ell \geq 1$,*

$$\hom(K_\ell, G_n) \neq \hom(K_\ell, H_n) \iff \ell = n.$$

Furthermore, G_n and H_n can be constructed in polynomial time in n.

Before we prove Theorem 9.5.2, we note that Theorem 9.5.3 implies a linear lower bound on the witness functions of the class $\mathcal{K}$ of all complete graphs. Hence, the trivial upper bound from Lemma 9.1.3 is tight up to a constant.

Corollary 9.5.4. *Let $f \colon \mathbb{N} \to \mathbb{N}$ be a non-decreasing witness function for $\mathcal{K}$. Then $f(n) \geq \frac{n}{2} + 1$ for infinitely many numbers $n \in \mathbb{N}$.*

Proof. Let G_n and H_n be as in Theorem 9.5.3 for $n \geq 2$. Let $m := \max\{|V(G_n)|, |V(H_n)|\} \leq 2(n-1)$. By Theorem 9.5.3, $f(m) \geq n$. Since f is non-decreasing, $f(2(n-1)) \geq f(m) \geq n$. □

Proof of Theorem 9.5.2. Consider the following reduction. Given an instance $n \geq 1$ of Mem(S), construct G_n and H_n via Theorem 9.5.3. Then $G_n \equiv_{\mathcal{K}^S} H_n$ if, and only if, $n \notin S$. □

Theorem 9.5.2 has the fault that not only HomInd($\mathcal{K}^S$) is hard but already deciding membership in $\mathcal{K}^S$ is hard. Overcoming this deficiency, Böker, Chen, Grohe, & Rattan [33, Theorem 1] showed that there exists a polynomial-time decidable graph class $\mathcal{F}$ of bounded treewidth for which HomInd($\mathcal{F}$) is undecidable. However, this graph class is unnatural in the sense that it does not have a concise definition in logics, i.e. it is not CMSO_2-definable by Theorem 9.0.1. Addressing homomorphism indistinguishability over natural graph classes, Theorem 9.5.3 was used to show the following.

Theorem 9.5.5 ([33, Theorem 18]). HomInd($\mathcal{K}$) *is* $C_=P$*-complete under polynomial-time Turing reductions.*

Thus, HomInd($\mathcal{K}$) is not in the polynomial hierarchy unless it collapses. The complexity class $C_=P$ should be thought of as the decision version of the functional complexity class #P.

The class $\mathcal{K}$ can be defined in first-order logic and has constant clique-width, cf. [48, Definition 2.89]. Thus, from a model checking perspective, it is a rather simple graph class. Theorem 9.5.5 indicates that one cannot replace Courcelle's graph algebras for bounded treewidth, which were used for Theorem 9.0.1, by their graph algebras for bounded clique-width, which Courcelle uses to prove that MSO_1 model checking is fixed-parameter tractable on classes of bounded clique-width.

Nevertheless, we show that the requirement of bounded treewidth can be relaxed when one does not aim at a polynomial-time algorithm. In the following Theorem 9.5.6, $\chi(F)$ denotes the chromatic number of F. A typical graph class to which the theorem applies is the class $\mathcal{K}$ of all complete graphs.

Theorem 9.5.6. *Let φ be a* CMSO_2*-sentence and $\mathcal{F}$ the class of all graphs satisfying φ. Suppose there exists a computable non-decreasing function $g \colon \mathbb{N} \to \mathbb{N}$ such that $\operatorname{tw}(F) \leq g(\chi(F))$ for all $F \in \mathcal{F}$. Then* $\textsc{HomInd}(\mathcal{F})$ *is decidable.*

Proof. Let $n := \max\{|V(G)|, |V(H)|\}$. If $F \in \mathcal{F}$ has treewidth greater than $g(n)$, then $\chi(F) > n$. Since G and H both have chromatic number at most n, there is neither a homomorphism $F \to G$ nor $F \to H$. Hence, $G \equiv_{\mathcal{F}} H$ if, and only if, $G \equiv_{\mathcal{F} \cap \mathcal{TW}_{g(n)}} H$. Now, invoke Theorem 9.2.7 with input φ and $g(n)$. □

9.6 A Trichotomy for Homomorphism Indistinguishability?

Theorem 9.0.1, the central result of this chapter, asserts that deciding homomorphism indistinguishability is tractable over every recognisable graph class of bounded treewidth. In particular, Theorem 9.0.1 shows that HomInd($\mathcal{F}$) is tractable for every minor-closed graph class of bounded treewidth. Notably, this result does not rely on reformulations of homomorphism indistinguishability relations in terms of logical or equational graph isomorphism relaxations but operates with the homomorphism counts themselves.

A reasonable next step is to combine Theorem 9.0.1 with a hardness result. To that end, we propose the following working hypothesis:

Conjecture 9.6.1. *Let $\mathcal{F}$ be a minor-closed graph class.*

1. *If $\mathcal{F}$ is the class of all graphs, then* HomInd($\mathcal{F}$) *is graph isomorphism.*
2. *If $\mathcal{F}$ has bounded treewidth, then* HomInd($\mathcal{F}$) *is in polynomial time.*
3. *If $\mathcal{F}$ is proper and has unbounded treewidth, then* HomInd($\mathcal{F}$) *is undecidable.*

The first assertion is implied by Theorem 3.1.1. The second assertions amounts to derandomising our Theorem 9.0.1 and is predicted by the complexity-theoretic hypothesis that P = coRP. The third assertion is wide open: The only minor-closed graph class $\mathcal{F}$ for which HomInd($\mathcal{F}$) is known to be undecidable is the class $\mathcal{P}$ of planar graphs, as shown in [124, 15]. Conjecture 9.6.1 is inspired by this example and Theorem 7.2.7 from graph minor theory [153], which asserts that every minor-closed graph class is either of bounded treewidth or contains all planar graphs. Intuitively, HomInd($\mathcal{P}$) is undecidable since the problem amounts to solving an infinite-dimensional system of equations. Roughly speaking, the dimension corresponds to the number of labels needed to generate all planar graphs under operations like series composition. Theorem 9.0.1 makes the other direction of this vague argument precise: We show that if the number of labels is bounded (e.g. the graph class has bounded treewidth), then considering finite-dimensional spaces suffices, rendering the problem tractable. That treewidth might be the right parameter in Conjecture 9.6.1 is also suggested by the complexity dichotomy for counting homomorphisms [52].

Conjecture 9.6.1 implies Conjecture 6.0.3 asserting that $\equiv_{\mathcal{F}}$ *is not the isomorphism relation* $\cong$ *for every proper minor-closed graph class* $\mathcal{F}$. Towards Conjecture 9.6.1, one could devise reductions between HomInd($\mathcal{F}_1$) and HomInd($\mathcal{F}_2$) for distinct minor-closed graph classes $\mathcal{F}_1$ and $\mathcal{F}_2$. The reductions we are aware of are limited, cf. Section 9.4.3.

Another pathway to Conjecture 9.6.1 is suggested by Definition 9.1.1: If a decidable graph class $\mathcal{F}$ admits a computable witness function, then HomInd($\mathcal{F}$) is decidable. On the other hand, proving lower bounds on witness functions is a purely combinatorial problem and avoids the intricacies of computation. Conjecture 9.6.1 implies that no witness function is computable for any minor-closed graph class of unbounded treewidth. A first step towards Question 9.6.2 was taken in Theorem 9.1.4.

Question 9.6.2. *Given a proper minor-closed graph class of unbounded treewidth, what are lower bounds on its witness functions?*

As a generalisation of Conjecture 9.6.1, we propose to study the complexity of the problems HomInd($\mathcal{F}$) for graph classes $\mathcal{F}$ which are monotone, i.e. closed under taking subgraphs, or hereditary, i.e. closed under taking induced subgraphs. Despite of Conjecture 6.0.2 and Theorem 7.0.1, it is not clear what exactly the role of minor-closed graph classes is in the theory of homomorphism indistinguishability. We expect an analogue of Conjecture 9.6.1 for hereditary or monotone graph classes to look much more complicated. One reason for this is that there exists many monotone graph classes which are not homomorphism distinguishing closed, e.g. the class of 2-degenerate graphs (Theorem 6.1.2) and the class of 3-colourable planar graphs (Theorem 7.3.5). Furthermore, there exist graph classes such as the class $\mathcal{K}$ of all complete graphs for which HomInd($\mathcal{K}$) is decidable yet hard (Theorem 9.5.5). On the tractability side, the realm of recognisable graph classes of bounded treewidth, to which Theorem 9.0.1 applies, contains many monotone or hereditary graph classes which are not minor-closed.

Question 9.6.3. *What does the complexity-theoretic landscape of the problems* HomInd($\mathcal{F}$) *for monotone or hereditary graph classes $\mathcal{F}$ look like?*

Recognisability is a key ingredient of Theorem 9.0.1 and a fairly general property of graph classes. As a last question, we ask what the role of recognisability is in the theory of homomorphism indistinguishability. Note that there exist recognisable graph classes which are homomorphism distinguishing closed (e.g. the class of treewidth at most k, cf. Theorem 6.4.1) and recognisable graph classes which are not (e.g. the class of 3-colourable graphs of treewidth at most k, cf. Theorem 7.3.5).

Question 9.6.4. *Is the homomorphism distinguishing closure* $\mathrm{cl}(\mathcal{F})$ *of every recognisable graph class $\mathcal{F}$ recognisable?*

10 Conclusion

Originating in Lovász's Theorem 3.1.1 and facilitated by connections to fields as diverse as quantum information theory, finite model theory, and category theory, homomorphism indistinguishability has developed into a beautiful coherent theory. This thesis demonstrates that homomorphism indistinguishability not only allows to recast graph isomorphism relaxations (characterisations) but also to study their distinguishing power (closure) and their computational complexity from a unified perspective. For example, as shown in this thesis, the homomorphism indistinguishability characterisation of the Lasserre semidefinite programming hierarchy made it possible to ascertain the hierarchy's precise distinguishing power and led to novel (randomised) polynomial-time algorithms for deciding its feasibility. This conclusion gives an overview over this thesis' contributions and open problems in each of the three areas (characterisations, closure, and complexity) which constitute the cornerstones of a coherent theory of homomorphism indistinguishability.

10.1 Characterisations

In Chapters 3 to 5, we demonstrated that many graph isomorphism relaxations from diverse fields can be characterised as homomorphism indistinguishability relations. In particular, we focused on equational graph isomorphism relaxations such as the Sherali–Adams and Lasserre relaxations of the graph isomorphism quadratic program. We took two points of view:

In Chapter 4, we constructed, given a graph class $\mathcal{F}$, a system of equations whose feasibility characterises homomorphism indistinguishability over $\mathcal{F}$. Our techniques applied, for example, to the classes of graphs of bounded treewidth, pathwidth, and treedepth, and to the class of graphs admitting a pebble forest cover of bounded depth. However, these methods do not apply to all graph classes. One limitation of the linear-algebraic Theorems 4.1.4 and 4.1.10 is that the considered vector spaces must be finite-dimensional. Since the dimension of a vector space of homomorphism vectors of k-labelled graphs with respect to some n-vertex graph is at most n^k, this amounts to assuming that the number of labels required to construct the graphs in $\mathcal{F}$ is finite. Given the operations from Section 3.2, this means that $\mathcal{F}$ must be of bounded treewidth. Even under this assumption, it is not clear how to characterise homomorphism indistinguishability over $\mathcal{F}$ as a concise system of equations. Addressing this problem, we characterised homomorphism indistinguishability over every recognisable graph class of bounded treewidth as feasibility of some collection of equations when proving Theorem 9.1.6. Perhaps, Theorem 9.1.6 is the most general result to be expected in this direction.

In the case of graph classes of unbounded treewidth, it appears that the application of more sophisticated algebraic techniques, which make use of infinite-dimensional vector spaces, is necessary. Here, a source of inspiration is the result of Mančinska & Roberson [124], which ultimately relies on the Tannaka–Krein duality due to Woronowicz [176] between easy quantum groups and partition categories, cf. the monograph [121]. Orthogonal easy quantum groups were exhaustively classified in [147]. It is conceivable that their partition categories yield further graph classes to which the same machinery that yielded the characterisation of quantum isomorphism as homomorphism indistinguishability over planar graphs applies, cf. [123, Section 8.1].

In Chapter 5, we conversely considered known equational graph isomorphism relaxations and constructed graph classes whose homomorphism indistinguishability relations characterise them. Concretely, we showed that feasibility of each level of the Lasserre semidefinite programming hierarchy for the graph isomorphism quadratic program is characterised as a homomorphism indistinguishability relation. Beyond the Lasserre and Sherali–Adams hierarchies, a reasonable next step is to study the Lovász–Schrijver hierarchy [117] or, more generally, hyperbolic programming [149] using homomorphism indistinguishability. Another interesting direction is indicated by Question 8.4.2, which asks to characterise the feasibility of Sherali–Adams relaxations over the integers or finite fields.

Abstracting from Chapters 3 to 5, the most fundamental (and vaguest) question regarding homomorphism indistinguishability characterisations is the following Question 7.4.4, which asks for an axiomatisation of homomorphism indistinguishability relations. In Chapter 7, we explored some surprising necessary conditions for a graph isomorphism relaxation to be a homomorphism indistinguishability relation over a graph class with certain closure properties.

Question 7.4.4. *What are sufficient and necessary conditions for a graph isomorphism relaxation to be a homomorphism indistinguishability relation?*

10.2 Closure

The homomorphism distinguishing closure is the central notion for studying the distinguishing power of homomorphism indistinguishability relations. This thesis outlined the currently available techniques and partial results in the context of Roberson's conjecture, the central open problem regarding the homomorphism distinguishing closure.

Conjecture 6.0.2 ([150, Conjecture 1])**.** *Every minor-closed and union-closed graph class is homomorphism distinguishing closed.*

Roberson [150, p. 4] remarked that it is not clear whether minor-closed graph classes play a distinct role in a theory of homomorphism indistinguishability. In Chapter 7, we gave evidence for this hypothesis by showing that the homomorphism distinguishing closure of a minor-closed graph class is homomorphism distinguishing closed.

In Chapter 6, we made progress towards Conjecture 6.0.2 by confirming it for all essentially finite graph classes, the classes of graphs of bounded pathwidth, for every $h \geq 3$, the class of $K_{2,h}$-minor-free graphs of treewidth at most two. Moreover, we showed that the class of disjoint unions of cycles is homomorphism distinguishing closed.

As an application, the result by Neuen [134] that the classes of graphs of bounded treewidth are homomorphism distinguishing closed allowed us to precisely determine the distinguishing power of the Lasserre hierarchy compared to the Sherali–Adams hierarchy. While we precisely determined the Sherali–Adams level whose feasibility guarantees feasibility of any given Lasserre level, the converse direction remains open, cf. Question 6.4.3. To that end, one would need to show that the graph classes $\mathcal{L}_t$ from Theorem 5.1.1 are homomorphism distinguishing closed, as predicted by Conjecture 6.0.2. Besides this concrete instance, subsequent problems are Conjecture 6.0.3 and those laid out in Section 6.6.

The CFI construction is the central tool for proving that a graph class is homomorphism distinguishing closed. Its generalisations and variants [24, 82] have not yet been considered from the perspective of homomorphism indistinguishability. They could facilitate progress towards Questions 8.4.2 and 9.6.2.

10.3 Complexity

The fundamental complexity-theoretic questions regarding homomorphism indistinguishability concern the decision problem HomInd($\mathcal{F}$) for a fixed graph class $\mathcal{F}$. Motivated by Conjecture 6.0.2 and Theorem 7.0.1, we considered the problems HomInd($\mathcal{F}$) for minor-closed graph classes $\mathcal{F}$ in Chapter 9. We conjectured that, as for homomorphism counting problems [52], the complexity of HomInd($\mathcal{F}$) is determined by the treewidth of $\mathcal{F}$. Partially resolving the following Conjecture 9.6.1, we showed that HomInd($\mathcal{F}$) is in randomised polynomial time for every recognisable graph class $\mathcal{F}$ of bounded treewidth. Overcoming limitations of semidefinite programming techniques, this result yields a novel randomised polynomial-time algorithm for deciding the exact feasibility of the Lasserre relaxations for graph isomorphism.

Conjecture 9.6.1. *Let $\mathcal{F}$ be a minor-closed graph class.*

1. *If $\mathcal{F}$ is the class of all graphs, then* HomInd($\mathcal{F}$) *is graph isomorphism.*
2. *If $\mathcal{F}$ has bounded treewidth, then* HomInd($\mathcal{F}$) *is in polynomial time.*
3. *If $\mathcal{F}$ is proper and has unbounded treewidth, then* HomInd($\mathcal{F}$) *is undecidable.*

Previous Publications

This section contains the declaration in accordance with §5 (5) and (6) of the Regulations Governing Doctoral Studies of the Faculty of Mathematics, Computer Science and Natural Sciences of RWTH Aachen University (2018/174). This thesis is based on the following previously published works, from which material is included verbatim without citations for individual results.

Grohe, Rattan, & Seppelt [86] Results were obtained collaboratively with the exception of [86, Theorem 3], which is due to Gaurav Rattan and omitted here. I formulated most of the theorems and proofs. In particular, I developed the representation-theoretic foundations (Section 2.5 and Lemmas 4.1.5 and 4.1.11), the bounds in Theorems 4.1.6 and 4.1.14, and the notion of inner-product compatibility (Definition 4.3.23).

Chapter 4 (except for Section 4.5), Sections 2.1.3 and 2.5, and parts of Section 2.4 are based on the full version [87].

Martin Grohe, Gaurav Rattan, & Tim Seppelt. 'Homomorphism Tensors and Linear Equations'. In: *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*. Ed. by Mikołaj Bojańczyk, Emanuela Merelli, & David P. Woodruff. Vol. 229. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 70:1–70:20. DOI: `10.4230/LIPIcs.ICALP.2022.70`

Rattan & Seppelt [145] Results were obtained collaboratively. I formulated most of the theorems and proofs. In particular, I developed the notions and results in [145, Section 3] with the exception of the counterexample in [145, Theorem 3.2], which is due to Gaurav Rattan. The augmented homomorphism representation [145, Definition 4.2] was conceived by Gaurav Rattan. I carried out the details in the proofs of [145, Section 4].

Section 4.5 and parts of Section 2.1 are based on the full version [146].

Gaurav Rattan & Tim Seppelt. 'Weisfeiler–Leman and Graph Spectra'. In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Society for Industrial and Applied Mathematics, 2023, pp. 2268–2285. DOI: `10.1137/1.9781611977554.ch87`

Roberson & Seppelt [151, 152] The idea to study the Lasserre hierarchy using homomorphism indistinguishability was proposed by David E. Roberson during my research visit at Danmarks Tekniske Universitet in October 2022. During that visit, we collaboratively convinced ourselves of the main lemmas and theorems and developed the overall structure of their proofs. I contributed techniques to treat tree decompositions in the framework of (bi)labelled graphs (Section 5.2). Subsequently, I carried out the details of the entire paper. I developed [151, Section 5].

Chapter 5, Corollary 6.4.2 and Question 6.4.3, Section 9.3.2, and parts of Sections 2.4 and 2.6 are based on the full version [152].

David E. Roberson & Tim Seppelt. 'Lasserre Hierarchy for Graph Isomorphism and Homomorphism Indistinguishability'. In: *50th International Colloquium on Automata, Languages, and Programming (ICALP 2023)*. Ed. by Kousha Etessami, Uriel Feige, & Gabriele Puppis. Vol. 261. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 101:1–101:18. DOI: `10.4230/LIPIcs.ICALP.2023.101`

David E. Roberson & Tim Seppelt. 'Lasserre Hierarchy for Graph Isomorphism and Homomorphism Indistinguishability'. In: *TheoretiCS* 3 (Sept. 2024). DOI: `10.46298/theoretics.24.20`

Seppelt [162, 165] Chapter 7, Section 6.5, and parts of Sections 2.1 and 9.5 are based on the full version.

Tim Seppelt. 'Logical Equivalences, Homomorphism Indistinguishability, and Forbidden Minors'. In: *48th International Symposium on Mathematical Foundations of Computer Science (MFCS 2023)*. Ed. by Jérôme Leroux, Sylvain Lombardy, & David Peleg. Vol. 272. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 82:1–82:15. DOI: `10.4230/LIPIcs.MFCS.2023.82`

Tim Seppelt. 'Logical equivalences, homomorphism indistinguishability, and forbidden minors'. In: *Information and Computation* 301 (2024): *Special issue on the 48th International Symposium on Mathematical Foundations of Computer Science (MFCS 2023)*, p. 105224. DOI: `10.1016/j.ic.2024.105224`

Lichter, Pago, & Seppelt [111] I contributed [111, Sections 3, 5, and 6]. The other parts of the paper were provided by Moritz Lichter and Benedikt Pago.

Section 6.3.1 and Chapter 8 are based on the full version [110].

Moritz Lichter, Benedikt Pago, & Tim Seppelt. 'Limitations of Game Comonads for Invertible-Map Equivalence via Homomorphism Indistinguishability'. In: *32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)*. Ed. by Aniello Murano & Alexandra Silva. Vol. 288. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 36:1–36:19. DOI: `10.4230/LIPIcs.CSL.2024.36`

Seppelt [164] Chapter 9 (except for Section 9.3.2) and parts of Section 9.5 are based on the full version [163].

Tim Seppelt. 'An Algorithmic Meta Theorem for Homomorphism Indistinguishability'. In: *49th International Symposium on Mathematical Foundations of Computer Science (MFCS 2024)*. Ed. by Rastislav Královič & Antonín Kučera. Vol. 306. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 82:1–82:19. DOI: 10.4230/LIPIcs.MFCS.2024.82

Neuen & Seppelt (unpublished) Section 6.3.4 is joint work with Daniel Neuen. Daniel Neuen proved Theorem 6.3.15 for $h = 3$. I generalised their argument to all $h \geq 3$ by proving Lemma 6.3.20.

The following two publications were made in preparation of this thesis. However, results from these publications are individually cited and not reproduced verbatim.

Fluck, Seppelt, & Spitzer [68] I contributed statements and proofs in [68, Section 4.3]. Some of them are stated with individual references in Section 6.4.

Eva Fluck, Tim Seppelt, & Gian Luca Spitzer. 'Going Deep and Going Wide: Counting Logic and Homomorphism Indistinguishability over Graphs of Bounded Treedepth and Treewidth'. In: *32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)*. Ed. by Aniello Murano & Alexandra Silva. Vol. 288. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 27:1–27:17. DOI: 10.4230/LIPIcs.CSL.2024.27

Böker, Härtel, Runde, Seppelt, & Standke [34] I contributed [34, Section 6]. A result from this paper is discussed in Section 7.4.

Jan Böker, Louis Härtel, Nina Runde, Tim Seppelt, & Christoph Standke. 'The Complexity of Homomorphism Reconstructibility'. In: *41st International Symposium on Theoretical Aspects of Computer Science (STACS 2024)*. Ed. by Olaf Beyersdorff, Mamadou Moustapha Kanté, Orna Kupferman, & Daniel Lokshtanov. Vol. 289. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 19:1–19:20. DOI: 10.4230/LIPIcs.STACS.2024.19

Eidesstattliche Erklärung

Ich, Tim Frederik Seppelt, erkläre hiermit, dass diese Dissertation und die darin dargelegten Inhalte meine eigenen sind und selbstständig, als Ergebnis meiner eigenen originären Forschung, generiert wurden. Hiermit erkläre ich an Eides statt:

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;
2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;
3. Wenn immer andere eigene oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;
4. Wenn aus anderen eigenen oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;
5. Alle wesentlichen Quellen von Unterstützung wurden benannt;
6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;
7. Ein Teil oder Teile dieser Arbeit wurden zuvor veröffentlicht und zwar in Grohe, Rattan & Seppelt [86, 87], Rattan & Seppelt [145, 146], Roberson & Seppelt [151, 152], Seppelt [162, 165], Lichter, Pago & Seppelt [111, 110] und Seppelt [164, 163].

Aachen, am 9. Dezember 2024

Tim Frederik Seppelt

List of Tables

List of Figures

Bibliography

[1] Milton Abramowitz & Irene A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. 9th Revised ed. Edition. New York: Dover Publications Inc., 1965. ISBN: 0486612724 (cit. on p. 208).

[2] Samson Abramsky, Anuj Dawar, & Pengming Wang. 'The Pebbling Comonad in Finite Model Theory'. In: *Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. Reykjavík, Iceland: IEEE, 2017. DOI: 10.1109/LICS.2017.8005129 (cit. on pp. 3, 15, 90).

[3] Samson Abramsky, Tomáš Jakl, & Thomas Paine. 'Discrete Density Comonads and Graph Parameters'. In: *Coalgebraic Methods in Computer Science*. Ed. by Helle Hvid Hansen & Fabio Zanasi. Cham: Springer International Publishing, 2022, pp. 23–44. DOI: 10.1007/978-3-031-10736-8_2 (cit. on pp. 1, 4, 34, 90, 162).

[4] Samson Abramsky & Nihil Shah. 'Relating structure and power: Comonadic semantics for computational resources'. In: *Journal of Logic and Computation* 31.6 (2021), pp. 1390–1428. DOI: 10.1093/logcom/exab048 (cit. on p. 90).

[5] Isolde Adler & Eva Fluck. 'Monotonicity of the Cops and Robber Game for Bounded Depth Treewidth'. In: *49th International Symposium on Mathematical Foundations of Computer Science (MFCS 2024)*. Ed. by Rastislav Královič & Antonín Kučera. Vol. 306. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 6:1–6:18. DOI: 10.4230/LIPIcs.MFCS.2024.6 (cit. on p. 146).

[6] Manindra Agrawal, Neeraj Kayal, & Nitin Saxena. 'PRIMES is in P'. In: *Annals of Mathematics* 160.2 (2004), pp. 781–793. DOI: 10.4007/annals.2004.160.781 (cit. on p. 208).

[7] Noga Alon, Phuong Dao, Iman Hajirasouliha, Fereydoun Hormozdiari, & Süleyman Cenk Sahinalp. 'Biomolecular network motif counting and discovery by color coding'. In: *Proceedings of the 16th International Conference on Intelligent Systems for Molecular Biology (ISMB), Toronto, Canada, July 19-23, 2008*. 2008, pp. 241–249. DOI: 10.1093/bioinformatics/btn163 (cit. on p. 151).

[8] Alex Arkhipov. *Extending and Characterizing Quantum Magic Games*. 2012. arXiv: 1209.3819v1 [quant-ph] (cit. on p. 127).

[9] Sanjeev Arora & Boaz Barak. *Computational complexity: a modern approach*. Cambridge; New York: Cambridge University Press, 2009. 579 pp. ISBN: 978-0-521-42426-4 (cit. on p. 191).

[10] V. Arvind, Frank Fuhlbrück, Johannes Köbler, & Oleg Verbitsky. 'On Weisfeiler–Leman invariance: Subgraph counts and related graph properties'. In: *Journal of Computer and System Sciences* 113 (2020), pp. 42–59. DOI: 10.1016/j.jcss.2020.04.003 (cit. on p. 9).

[11] Albert Atserias, Andrei Bulatov, & Víctor Dalmau. 'On the Power of k-Consistency'. In: *Automata, Languages and Programming (ICALP 2007)*. Ed. by Lars Arge, Christian Cachin, Tomasz Jurdziński, & Andrzej Tarlecki. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 279–290. DOI: 10.1007/978-3-540-73420-8_26 (cit. on p. 145).

[12] Albert Atserias, Andrei Bulatov, & Anuj Dawar. 'Affine systems of equations and counting infinitary logic'. In: *Theoretical Computer Science* 410.18 (2009): *Automata, Languages and Programming (ICALP 2007)*, pp. 1666–1683. DOI: 10.1016/j.tcs.2008.12.049 (cit. on p. 145).

[13] Albert Atserias & Joanna Fijalkow. 'Definable Ellipsoid Method, Sums-of-Squares Proofs, and the Graph Isomorphism Problem'. In: *SIAM Journal on Computing* 52.5 (2023), pp. 1193–1229. DOI: 10.1137/20M1338435 (cit. on pp. 5, 7, 29, 95, 96, 212).

[14] Albert Atserias, Phokion G. Kolaitis, & Wei-Lin Wu. 'On the Expressive Power of Homomorphism Counts'. In: *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)* (Rome, Italy). IEEE, 2021, pp. 1–13. DOI: 10.1109/LICS52264.2021.9470543 (cit. on pp. 1, 32, 33, 47, 158, 175, 177, 190).

[15] Albert Atserias, Laura Mančinska, David E. Roberson, Robert Šámal, Simone Severini, & Antonios Varvitsiotis. 'Quantum and non-signalling graph isomorphisms'. In: *Journal of Combinatorial Theory, Series B* 136 (2019), pp. 289–328. DOI: 10.1016/j.jctb.2018.11.002 (cit. on pp. 1, 6, 10, 127, 192, 222, 227).

[16] Albert Atserias & Elitza N. Maneva. 'Sherali–Adams Relaxations and Indistinguishability in Counting Logics'. In: *SIAM Journal on Computing* 42.1 (2013), pp. 112–137. DOI: 10.1137/120867834 (cit. on pp. 5, 6, 27).

[17] László Babai. 'Graph Isomorphism in Quasipolynomial Time [Extended Abstract]'. In: *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC 2016)*. New York, NY, USA: ACM, 2016, pp. 684–697. DOI: 10.1145/2897518.2897542 (cit. on pp. 10, 192).

[18] Libor Barto, Andrei Krokhin, & Ross Willard. 'Polymorphisms, and How to Use Them'. In: *The Constraint Satisfaction Problem: Complexity and Approximability*. Ed. by Andrei Krokhin & Stanislav Živný. Vol. 7. Dagstuhl Follow-Ups. Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017, pp. 1–44. DOI: 10.4230/DFU.Vol7.15301.1 (cit. on p. 181).

[19] Paul Beaujean, Florian Sikora, & Florian Yger. 'Graph Homomorphism Features: Why Not Sample?' In: *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*. Ed. by Michael Kamp, Irena Koprinska, Adrien Bibal, Tassadit Bouadi, Benoît Frénay, Luis Galárraga, José Oramas, Linara Adilova, Yamuna Krishnamurthy, Bo Kang, Christine Largeron, Jefrey Lijffijt, Tiphaine Viard, Pascal Welke, Massimiliano Ruocco, Erlend Aune, Claudio Gallicchio, Gregor Schiele, Franz Pernkopf, Michaela Blott, Holger Fröning, Günther Schindler, Riccardo Guidotti, Anna Monreale, Salvatore Rinzivillo, Przemyslaw Biecek, Eirini Ntoutsi, Mykola Pechenizkiy, Bodo Rosenhahn, Christopher Buckley, Daniela Cialfi, Pablo Lanillos, Maxwell Ramstead, Tim Verbelen, Pedro M. Ferreira, Giuseppina Andresini, Donato Malerba, Ibéria Medeiros, Philippe Fournier-Viger, M. Saqib Nawaz, Sebastian Ventura, Meng Sun, Min Zhou, Valerio Bitetta, Ilaria Bordino, Andrea Ferretti, Francesco Gullo, Giovanni Ponti, Lorenzo Severini, Rita Ribeiro, João Gama, Ricard Gavaldà, Lee Cooper, Naghmeh Ghazaleh, Jonas Richiardi, Damian Roqueiro, Diego Saldana Miranda, Konstantinos Sechidis, & Guilherme Graça. Vol. 1524. Series Title: Communications in Computer and Information Science. Cham: Springer International Publishing, 2021, pp. 216–222. DOI: 10.1007/978-3-030-93736-2_17 (cit. on p. 151).

[20] Christoph Berkholz. 'Lower Bounds for Existential Pebble Games and k-Consistency Tests'. In: *Proceedings of the 27th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. Dubrovnik, Croatia: IEEE, 2012, pp. 25–34. DOI: 10.1109/LICS.2012.14 (cit. on pp. 10, 221).

[21] Christoph Berkholz. 'The Relation between Polynomial Calculus, Sherali–Adams, and Sum-of-Squares Proofs'. In: *35th Symposium on Theoretical Aspects of Computer Science (STACS 2018)*. Ed. by Rolf Niedermeier & Brigitte Vallée. Vol. 96. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018, 11:1–11:14. DOI: 10.4230/LIPIcs.STACS.2018.11 (cit. on p. 5).

[22] Christoph Berkholz, Paul S. Bonsma, & Martin Grohe. 'Tight Lower and Upper Bounds for the Complexity of Canonical Colour Refinement'. In: *Theory of Computing Systems* 60.4 (2017), pp. 581–614. DOI: 10.1007/s00224-016-9686-0 (cit. on pp. 10, 221).

[23] Christoph Berkholz & Martin Grohe. 'Limitations of Algebraic Approaches to Graph Isomorphism Testing'. In: *Automata, Languages, and Programming*.

Ed. by Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, & Bettina Speckmann. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 155–166. DOI: 10.1007/978-3-662-47672-7_13 (cit. on pp. 6, 95, 96, 189).

[24] Christoph Berkholz & Martin Grohe. 'Linear Diophantine Equations, Group CSPs, and Graph Isomorphism'. In: *Proceedings of the 2017 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Ed. by Philip N. Klein. 2017, pp. 327–339. DOI: 10.1137/1.9781611974782.21 (cit. on pp. 189, 231).

[25] Dan Bienstock, Neil Robertson, Paul D. Seymour, & Robin Thomas. 'Quickly excluding a forest'. In: *Journal of Combinatorial Theory, Series B* 52.2 (1991), pp. 274–283. DOI: 10.1016/0095-8956(91)90068-U (cit. on pp. 147, 149).

[26] Hans L. Bodlaender. 'A partial k-arboretum of graphs with bounded treewidth'. In: *Theoretical Computer Science* 209.1 (1998), pp. 1–45. DOI: 10.1016/S0304-3975(97)00228-4 (cit. on pp. 14, 137, 138).

[27] Hans L. Bodlaender, Édouard Bonnet, Lars Jaffke, Dušan Knop, Paloma T. Lima, Martin Milanič, Sebastian Ordyniak, Sukanya Pandey, & Ondřej Suchý. 'Treewidth Is NP-Complete on Cubic Graphs'. In: *18th International Symposium on Parameterized and Exact Computation (IPEC 2023)*. Ed. by Neeldhara Misra & Magnus Wahlström. Vol. 285. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 7:1–7:13. DOI: 10.4230/LIPIcs.IPEC.2023.7 (cit. on p. 221).

[28] Hans L. Bodlaender & Dimitrios M. Thilikos. 'Treewidth for graphs with small chordality'. In: *Discrete Applied Mathematics* 79.1-3 (1997), pp. 45–61. DOI: 10.1016/S0166-218X(97)00031-0 (cit. on p. 221).

[29] Mikołaj Bojańczyk & Michał Pilipczuk. 'Definability equals recognizability for graphs of bounded treewidth'. In: *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science*. New York NY USA: ACM, 2016, pp. 407–416. DOI: 10.1145/2933575.2934508 (cit. on pp. 196, 197).

[30] Mikołaj Bojańczyk & Igor Walukiewicz. 'Forest algebras'. In: *Logic and Automata*. Ed. by Jörg Flum, Erich Grädel, & Thomas Wilke. History and Perspectives. Amsterdam University Press, 2008, pp. 107–132. ISBN: 978-90-5356-576-6. URL: http://www.jstor.org/stable/j.ctt46mv83.8 (cit. on p. 58).

[31] Jan Böker. 'Structural Similarity and Homomorphism Counts'. MA thesis. Aachen: RWTH Aachen University, 2018 (cit. on p. 34).

[32] Jan Böker. 'Color Refinement, Homomorphisms, and Hypergraphs'. In: *Graph-Theoretic Concepts in Computer Science - 45th International Workshop, WG 2019, Vall de Núria, Spain, June 19-21, 2019, Revised Papers*. Ed. by Ignasi Sau & Dimitrios M. Thilikos. Vol. 11789. Lecture Notes in Computer Science. Springer, 2019, pp. 338–350. DOI: 10.1007/978-3-030-30786-8_26 (cit. on p. 32).

[33] Jan Böker, Yijia Chen, Martin Grohe, & Gaurav Rattan. 'The Complexity of Homomorphism Indistinguishability'. In: *44th International Symposium on Mathematical Foundations of Computer Science (MFCS 2019)*. Ed. by Peter Rossmanith, Pinar Heggernes, & Joost-Pieter Katoen. Vol. 138. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019, 54:1–54:13. DOI: 10.4230/LIPIcs.MFCS.2019.54 (cit. on pp. 10, 191, 224–226).

[34] Jan Böker, Louis Härtel, Nina Runde, Tim Seppelt, & Christoph Standke. 'The Complexity of Homomorphism Reconstructibility'. In: *41st International Symposium on Theoretical Aspects of Computer Science (STACS 2024)*. Ed. by Olaf Beyersdorff, Mamadou Moustapha Kanté, Orna Kupferman, & Daniel Lokshtanov. Vol. 289. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 19:1–19:20. DOI: 10.4230/LIPIcs.STACS.2024.19 (cit. on pp. 182, 235).

[35] Andrei A. Bulatov & Amirhossein Kazeminia. 'Complexity classification of counting graph homomorphisms modulo a prime number'. In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2022)*. Rome, Italy: ACM, 2022, pp. 1024–1037. DOI: 10.1145/3519935.3520075 (cit. on p. 189).

[36] Silvia Butti & Víctor Dalmau. 'Fractional Homomorphism, Weisfeiler-Leman Invariance, and the Sherali-Adams Hierarchy for the Constraint Satisfaction Problem'. In: *46th International Symposium on Mathematical Foundations of Computer Science (MFCS 2021)*. Ed. by Filippo Bonchi & Simon J. Puglisi. Vol. 202. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021, 27:1–27:19. DOI: 10.4230/LIPIcs.MFCS.2021.27 (cit. on p. 32).

[37] Jin-Yi Cai, Martin Fürer, & Neil Immerman. 'An optimal lower bound on the number of variables for graph identification'. In: *Combinatorica* 12.4 (1992), pp. 389–410. DOI: 10.1007/BF01305232 (cit. on pp. 3, 8, 125, 126, 131, 145, 217, 219).

[38] Balder ten Cate, Víctor Dalmau, Phokion G. Kolaitis, & Wei-Lin Wu. 'When Do Homomorphism Counts Help in Query Algorithms?' In: *27th International Conference on Database Theory, ICDT 2024, March 25-28, 2024, Paestum, Italy*. Ed. by Graham Cormode & Michael Shekelyan. Vol. 290. LIPIcs. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 8:1–8:20. DOI: 10.4230/LIPICS.ICDT.2024.8 (cit. on pp. 32, 151, 190).

[39] Surajit Chaudhuri & Moshe Y. Vardi. 'Optimization of Real Conjunctive Queries'. In: *Proceedings of the Twelfth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, May 25-28, 1993, Washington, DC, USA*. Ed.

by Catriel Beeri. ACM Press, 1993, pp. 59–70. DOI: 10.1145/153850.153856 (cit. on p. 151).

[40] Gang Chen & Ilia Ponomarenko. *Lectures on Coherent Configurations*. Wuhan: Central China Normal University Press, 2018. URL: https://web.archive.org/web/20230202201123/https://www.pdmi.ras.ru/~inp/ccNOTES.pdf (cit. on pp. 105, 106).

[41] Yijia Chen, Jörg Flum, Mingjun Liu, & Zhiyang Xun. 'On Algorithms Based on Finitely Many Homomorphism Counts'. In: *47th International Symposium on Mathematical Foundations of Computer Science, MFCS 2022, August 22-26, 2022, Vienna, Austria*. Ed. by Stefan Szeider, Robert Ganian, & Alexandra Silva. Vol. 241. LIPIcs. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 32:1–32:15. DOI: 10.4230/LIPIcs.MFCS.2022.32 (cit. on p. 151).

[42] Man-Duen Choi. 'Completely positive linear maps on complex matrices'. In: *Linear Algebra and its Applications* 10.3 (1975), pp. 285–290. DOI: 10.1016/0024-3795(75)90075-0 (cit. on p. 24).

[43] Paolo Codenotti, Grant Schoenebeck, & Aaron Snook. *Graph Isomorphism and the Lasserre Hierarchy*. 2014. arXiv: 1401.0758v1 [cs.CC] (cit. on pp. 5, 29, 95).

[44] Lothar Collatz & Ulrich Sinogowitz. 'Spektren endlicher Grafen. Wilhelm Blaschke zum 70. Geburtstag gewidmet'. In: *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 21.1 (1957), pp. 63–77. DOI: 10.1007/BF02941924 (cit. on p. 7).

[45] Jesse Comer. *Lovász Theorems for Modal Languages*. 2024. arXiv: 2404.15421v1 [cs.LO] (cit. on p. 190).

[46] Bruno Courcelle. 'The monadic second-order logic of graphs. I. Recognizable sets of finite graphs'. In: *Information and Computation* 85.1 (1990), pp. 12–75. DOI: 10.1016/0890-5401(90)90043-H (cit. on pp. 192, 197, 203, 209, 210).

[47] Bruno Courcelle. 'The monadic second order logic of graphs VI: on several representations of graphs by relational structures'. In: *Discrete Applied Mathematics* 54.2 (1994), pp. 117–149. DOI: 10.1016/0166-218X(94)90019-1 (cit. on p. 177).

[48] Bruno Courcelle & Joost Engelfriet. *Graph Structure and Monadic Second-Order Logic: A Language-Theoretic Approach*. 1st. USA: Cambridge University Press, 2012. ISBN: 0-521-89833-1 (cit. on pp. 17, 192, 211, 226).

[49] Radu Curticapean, Holger Dell, & Dániel Marx. 'Homomorphisms Are a Good Basis for Counting Small Subgraphs'. In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2017)*. New York, NY, USA: ACM, 2017, pp. 210–223. DOI: 10.1145/3055399.3055502 (cit. on pp. 8, 161).

[50] D. M. Cvetković & I. Gutman. 'On Spectral Structure Of Graphs Having The Maximal Eigenvalue Not Grater Than Two'. In: *Publications de l'Institut Mathématique* 18.32 (1975), pp. 39–45. URL: http://eudml.org/doc/254702 (cit. on p. 165).

[51] Marek Cygan, Fedor V. Fomin, Łukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michał Pilipczuk, & Saket Saurabh. *Parameterized Algorithms*. Cham: Springer International Publishing, 2015. DOI: 10.1007/978-3-319-21275-3 (cit. on p. 222).

[52] Víctor Dalmau & Peter Jonsson. 'The Complexity of Counting Homomorphisms Seen from the Other Side'. In: *Theoretical Computer Science* 329.1 (2004), pp. 315–323. DOI: 10.1016/j.tcs.2004.08.008 (cit. on pp. 32, 227, 231).

[53] E.R. van Dam, W.H. Haemers, & J.H. Koolen. 'Cospectral graphs and the generalized adjacency matrix'. In: *Special Issue devoted to papers presented at the Aveiro Workshop on Graph Spectra* 423.1 (2007), pp. 33–41. DOI: 10.1016/j.laa.2006.07.017 (cit. on pp. 5, 45, 47).

[54] Edwin R. van Dam & Willem H. Haemers. 'Which graphs are determined by their spectrum?' In: *Linear Algebra and its Applications* 373 (2003), pp. 241–272. DOI: 10.1016/S0024-3795(03)00483-X (cit. on pp. 7, 41, 44, 45).

[55] Anuj Dawar, Erich Grädel, & Moritz Lichter. 'Limitations of the invertible-map equivalences'. In: *Journal of Logic and Computation* (2022), exac058. DOI: 10.1093/logcom/exac058 (cit. on p. 4).

[56] Anuj Dawar, Erich Grädel, & Wied Pakusa. 'Approximations of Isomorphism and Logics with Linear-Algebraic Operators'. In: *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*. Ed. by Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, & Stefano Leonardi. Vol. 132. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019, 112:1–112:14. DOI: 10.4230/LIPIcs.ICALP.2019.112 (cit. on pp. 4, 178).

[57] Anuj Dawar, Tomáš Jakl, & Luca Reggio. 'Lovász-Type Theorems and Game Comonads'. In: *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)* (Rome, Italy). IEEE, 2021, pp. 1–13. DOI: 10.1109/LICS52264.2021.9470609 (cit. on pp. 1, 4, 15, 32, 33, 49, 52, 78, 84, 90).

[58] Anuj Dawar, Simone Severini, & Octavio Zapata. 'Descriptive complexity of graph spectra'. In: *Annals of Pure and Applied Logic* 170.9 (2019), pp. 993–1007. DOI: 10.1016/j.apal.2019.04.005 (cit. on p. 45).

[59] Anuj Dawar & Pengming Wang. 'Definability of Semidefinite Programming and Lasserre Lower Bounds for CSPs'. In: *Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. Reykjavík, Iceland: IEEE, 2017, pp. 1–12. DOI: 10.1109/LICS.2017.8005108 (cit. on p. 9).

[60] Anuj Dawar & Gregory Wilsenach. 'Symmetric Arithmetic Circuits'. In: *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*. Ed. by Artur Czumaj, Anuj Dawar, & Emanuela Merelli. Vol. 168. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020, 36:1–36:18. DOI: 10.4230/LIPIcs.ICALP.2020.36 (cit. on p. 9).

[61] Holger Dell, Martin Grohe, & Gaurav Rattan. 'Lovász Meets Weisfeiler and Leman'. In: *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*. Ed. by Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, & Donald Sannella. Vol. 107. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018, 40:1–40:14. DOI: 10.4230/LIPIcs.ICALP.2018.40 (cit. on pp. 1, 3, 5, 7, 29, 45, 47, 49, 51, 66–68, 84, 86, 136).

[62] Rodney G. Downey & Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. London: Springer London, 2013. DOI: 10.1007/978-1-4471-5559-1 (cit. on p. 197).

[63] Zdeněk Dvořák. 'On recognizing graphs by numbers of homomorphisms'. In: *Journal of Graph Theory* 64.4 (2010), pp. 330–342. DOI: 10.1002/jgt.20461 (cit. on pp. 1–4, 8, 32, 41, 47, 49, 128, 129, 180, 183, 216–218).

[64] Martin Dyer & Catherine Greenhill. 'The complexity of counting graph homomorphisms'. In: *Random Structures and Algorithms* 17.3 (2000), pp. 260–289. DOI: 10.1002/1098-2418(200010/12)17:3/4<260::AID-RSA5>3.0.CO;2-W (cit. on p. 32).

[65] H.-D. Ebbinghaus. 'Extended Logics: The General Framework'. In: *Model-Theoretic Logics*. Ed. by J. Barwise & S. Feferman. Perspectives in Logic. Cambridge: Cambridge University Press, 2017, pp. 25–76. DOI: 10.1017/9781316717158.005 (cit. on p. 16).

[66] Thomas Eiter, Georg Gottlob, & Thomas Schwentick. 'The Model Checking Problem for Prefix Classes of Second-Order Logic: A Survey'. In: *Fields of Logic and Computation*. Ed. by Andreas Blass, Nachum Dershowitz, & Wolfgang Reisig. Vol. 6300. Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 227–250. DOI: 10.1007/978-3-642-15025-8_13 (cit. on p. 177).

[67] John Faben & Mark Jerrum. 'The Complexity of Parity Graph Homomorphism: An Initial Investigation'. In: *Theory of Computing* 11.2 (2015), pp. 35–57. DOI: 10.4086/toc.2015.v011a002 (cit. on pp. 183, 184).

[68] Eva Fluck, Tim Seppelt, & Gian Luca Spitzer. 'Going Deep and Going Wide: Counting Logic and Homomorphism Indistinguishability over Graphs of Bounded Treedepth and Treewidth'. In: *32nd EACSL Annual Conference on*

Computer Science Logic (CSL 2024). Ed. by Aniello Murano & Alexandra Silva. Vol. 288. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 27:1–27:17. DOI: 10.4230/LIPIcs.CSL.2024.27 (cit. on pp. 1, 4, 9, 16, 49, 125, 146, 148, 235).

[69] Michael Freedman, László Lovász, & Alexander Schrijver. 'Reflection positivity, rank connectivity, and homomorphism of graphs'. In: *Journal of the American Mathematical Society* 20 (2007), pp. 37–51. DOI: 10.1090/S0894-0347-06-00529-7 (cit. on p. 182).

[70] G. Frobenius & I. Schur. 'Über die Äquivalenz der Gruppen linearer Substitutionen'. In: *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* 1 (1906), pp. 209–217. URL: https://www.biodiversitylibrary.org/page/29500785 (cit. on p. 25).

[71] Martin Fürer. 'Weisfeiler–Lehman Refinement Requires at Least a Linear Number of Iterations'. In: *Automata, Languages and Programming*. Ed. by Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, Fernando Orejas, Paul G. Spirakis, & Jan van Leeuwen. Vol. 2076. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 322–333. DOI: 10.1007/3-540-48224-5_27 (cit. on pp. 68, 131).

[72] Vyacheslav Futorny, Roger A. Horn, & Vladimir V. Sergeichuk. 'Specht's criterion for systems of linear mappings'. In: *Linear Algebra and its Applications* 519 (2017), pp. 278–295. DOI: 10.1016/j.laa.2017.01.006 (cit. on p. 53).

[73] Joseph A. Gallian. 'Graph Labeling'. In: *The Electronic Journal of Combinatorics* 1000 (1998). DOI: 10.37236/27 (cit. on p. 35).

[74] Dennis Geller & Saul Stahl. 'The chromatic number and other functions of the lexicographic product'. In: *Journal of Combinatorial Theory, Series B* 19.1 (1975), pp. 87–95. DOI: 10.1016/0095-8956(75)90076-3 (cit. on p. 176).

[75] Timo Gervens & Martin Grohe. 'Graph Similarity Based on Matrix Norms'. In: *47th International Symposium on Mathematical Foundations of Computer Science (MFCS 2022)*. Ed. by Stefan Szeider, Robert Ganian, & Alexandra Silva. Vol. 241. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 52:1–52:15. DOI: 10.4230/LIPIcs.MFCS.2022.52 (cit. on p. 1).

[76] Martin Grohe. 'Finite Variable Logics in Descriptive Complexity Theory'. In: *Bulletin of Symbolic Logic* 4.4 (1998), pp. 345–398. DOI: 10.2307/420954 (cit. on p. 3).

[77] Martin Grohe. 'Equivalence in Finite-Variable Logics is Complete for Polynomial Time'. In: *Combinatorica* 19.4 (1999), pp. 507–532. DOI: 10.1007/s004939970004 (cit. on pp. 192, 220).

[78] Martin Grohe. *Descriptive Complexity, Canonisation, and Definable Graph Structure Theory*. Cambridge: Cambridge University Press, 2017. DOI: 10.1017/9781139028868 (cit. on pp. 16, 158, 177).

[79] Martin Grohe. 'Counting Bounded Tree Depth Homomorphisms'. In: *Proceedings of the 35th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. New York, NY, USA: ACM, 2020, pp. 507–520. DOI: 10.1145/3373718.3394739 (cit. on pp. 1, 3, 4, 32, 49).

[80] Martin Grohe. 'word2vec, node2vec, graph2vec, X2vec: Towards a Theory of Vector Embeddings of Structured Data'. In: *Proceedings of the 39th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2020, Portland, OR, USA, June 14-19, 2020*. Ed. by Dan Suciu, Yufei Tao, & Zhewei Wei. ACM, 2020, pp. 1–16. DOI: 10.1145/3375395.3387641 (cit. on pp. 1, 33, 44, 151).

[81] Martin Grohe. 'The Logic of Graph Neural Networks'. In: *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)* (Rome, Italy). IEEE, 2021, pp. 1–17. DOI: 10.1109/LICS52264.2021.9470677 (cit. on pp. 1, 3, 18, 19, 138).

[82] Martin Grohe, Moritz Lichter, Daniel Neuen, & Pascal Schweitzer. 'Compressing CFI Graphs and Lower Bounds for the Weisfeiler–Leman Refinements'. In: *64th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2023, Santa Cruz, CA, USA, November 6-9, 2023*. IEEE, 2023, pp. 798–809. DOI: 10.1109/FOCS57990.2023.00052 (cit. on pp. 221, 231).

[83] Martin Grohe & Daniel Neuen. 'Recent advances on the graph isomorphism problem'. In: *Surveys in Combinatorics, 2021: Invited lectures from the 28th British Combinatorial Conference, Durham, UK, July 5-9, 2021*. Ed. by Konrad K. Dabrowski, Maximilien Gadouleau, Nicholas Georgiou, Matthew Johnson, George B. Mertzios, & Daniël Paulusma. Cambridge University Press, 2021, pp. 187–234. DOI: 10.1017/9781009036214.006 (cit. on pp. 1, 10).

[84] Martin Grohe & Martin Otto. 'Pebble Games and Linear Equations'. In: *The Journal of Symbolic Logic* 80.3 (2015), pp. 797–844. DOI: 10.1017/jsl.2015.28 (cit. on pp. 6, 27, 29).

[85] Martin Grohe & Wied Pakusa. 'Descriptive complexity of linear equation systems and applications to propositional proof complexity'. In: *Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. Reykjavík, Iceland: IEEE, 2017, pp. 1–12. DOI: 10.1109/LICS.2017.8005081 (cit. on p. 6).

[86] Martin Grohe, Gaurav Rattan, & Tim Seppelt. 'Homomorphism Tensors and Linear Equations'. In: *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*. Ed. by Mikołaj Bojańczyk, Emanuela Merelli,

& David P. Woodruff. Vol. 229. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 70:1–70:20. DOI: 10.4230/LIPIcs.ICALP.2022.70 (cit. on pp. 1, 53, 136, 233, 236).

[87] Martin Grohe, Gaurav Rattan, & Tim Seppelt. *Homomorphism Tensors and Linear Equations*. 2024. arXiv: 2111.11313v3 [math.CO] (cit. on pp. 53, 233, 236).

[88] Martin Grohe & Pascal Schweitzer. 'The Graph Isomorphism Problem'. In: *Communications of the ACM* 63.11 (2020), pp. 128–134. DOI: 10.1145/3372123 (cit. on pp. 1, 125).

[89] Martin Grotschel, László Lovász, & Alexander Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Berlin Heidelberg: Springer, 1998. DOI: 10.1007/978-3-642-97881-4 (cit. on p. 212).

[90] Hs. H. Günthard & H. Primas. 'Zusammenhang von Graphentheorie und MO-Theorie von Molekeln mit Systemen konjugierter Bindungen'. In: *Helvetica Chimica Acta* 39.6 (1956), pp. 1645–1653. DOI: 10.1002/hlca.19560390623 (cit. on p. 7).

[91] Lauri Hella. 'Logical Hierarchies in PTIME'. In: *Information and Computation* 129.1 (1996), pp. 1–19. DOI: 10.1006/inco.1996.0070 (cit. on pp. 3, 171).

[92] Roger A. Horn & Charles R. Johnson. *Matrix analysis*. 23rd ed. Cambridge: Cambridge University Press, 2010. 561 pp. ISBN: 978-0-521-38632-6 (cit. on p. 54).

[93] Neil Immerman & Eric Lander. 'Describing Graphs: A First-Order Approach to Graph Canonization'. In: *Complexity Theory Retrospective: In Honor of Juris Hartmanis on the Occasion of His Sixtieth Birthday, July 5, 1988*. Ed. by Alan L. Selman. New York, NY: Springer New York, 1990, pp. 59–81. DOI: 10.1007/978-1-4612-4478-3_5 (cit. on pp. 3, 10, 18, 19, 217, 220).

[94] John Isbell. 'Some inequalities in hom sets'. In: *Journal of Pure and Applied Algebra* 76.1 (1991), pp. 87–110. DOI: 10.1016/0022-4049(91)90099-N (cit. on p. 32).

[95] Tomáš Jakl, Dan Marsden, & Nihil Shah. 'A categorical account of composition methods in logic'. In: *Proceedings of the 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)* (Boston, USA). IEEE, 2023, pp. 1–14. DOI: 10.1109/LICS56636.2023.10175751 (cit. on p. 181).

[96] Naihuan Jing. 'Unitary and orthogonal equivalence of sets of matrices'. In: *Linear Algebra and its Applications* 481 (2015), pp. 235–242. DOI: 10.1016/j.laa.2015.04.036 (cit. on p. 53).

[97] Charles R. Johnson & Morris Newman. 'A note on cospectral graphs'. In: *Journal of Combinatorial Theory, Series B* 28.1 (1980), pp. 96–103. DOI: 10.1016/0095-8956(80)90058-1 (cit. on p. 5).

[98] Ravindran Kannan & Achim Bachem. 'Polynomial Algorithms for Computing the Smith and Hermite Normal Forms of an Integer Matrix'. In: *SIAM Journal on Computing* 8.4 (1979), pp. 499–507. DOI: 10.1137/0208040 (cit. on p. 189).

[99] Richard M. Karp. 'Reducibility among Combinatorial Problems'. In: *Complexity of Computer Computations*. Ed. by Raymond E. Miller, James W. Thatcher, & Jean D. Bohlinger. Boston, MA: Springer US, 1972, pp. 85–103. DOI: 10.1007/978-1-4684-2001-2_9 (cit. on pp. 9, 10).

[100] Sandra Kiefer. 'Power and limits of the Weisfeiler-Leman algorithm'. PhD thesis. RWTH Aachen University, 2020. DOI: 10.18154/RWTH-2020-03508 (cit. on pp. 3, 18, 158).

[101] Sandra Kiefer. 'The Weisfeiler-Leman algorithm: an exploration of its power'. In: *ACM SIGLOG News* 7.3 (2020), pp. 5–27. DOI: 10.1145/3436980.3436982 (cit. on p. 18).

[102] Nils M. Kriege, Fredrik D. Johansson, & Christopher Morris. 'A survey on graph kernels'. In: *Appl. Netw. Sci.* 5.1 (2020), p. 6. DOI: 10.1007/s41109-019-0195-3 (cit. on p. 151).

[103] Jarosław Kwiecień, Jerzy Marcinkowski, & Piotr Ostropolski-Nalewaja. 'Determinacy of Real Conjunctive Queries. The Boolean Case'. In: *Proceedings of the 41st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*. Philadelphia, PA, USA: ACM, 2022, pp. 347–358. DOI: 10.1145/3517804.3524168 (cit. on pp. 32, 151–153).

[104] Tsit-Yuen Lam. *A First Course in Noncommutative Rings*. 2. ed. Graduate texts in mathematics 131. New York, NY: Springer, 2001. DOI: 10.1007/978-1-4419-8616-0 (cit. on pp. 19, 25).

[105] Serge Lang. *Linear Algebra*. Undergraduate Texts in Mathematics. New York, NY: Springer New York, 1987. DOI: 10.1007/978-1-4757-1949-9 (cit. on pp. 19–23, 44, 45, 47).

[106] Jean B. Lasserre. 'Global Optimization with Polynomials and the Problem of Moments'. In: *SIAM Journal on Optimization* 11.3 (2001), pp. 796–817. DOI: 10.1137/S1052623400366802 (cit. on pp. 5, 7, 29).

[107] Monique Laurent. 'A Comparison of the Sherali–Adams, Lovász–Schrijver, and Lasserre Relaxations for 0–1 Programming'. In: *Mathematics of Operations Research* 28.3 (2003), pp. 470–496. URL: http://www.jstor.org/stable/4126981 (cit. on pp. 5, 29, 95).

[108] Moritz Lichter. 'Continuing the Quest for a Logic Capturing Polynomial Time – Potential, Limitations, and Interplay of Current Approaches'. PhD thesis. Darmstadt: Technische Universität Darmstadt, 2023, xii, 320 pages. DOI: 10.26083/tuprints-00024244 (cit. on pp. 4, 131).

[109] Moritz Lichter. 'Separating rank logic from polynomial time'. In: *Journal of the ACM* 70.2 (2023), pp. 1–53. DOI: 10.1145/3572918 (cit. on pp. 131, 178).

[110] Moritz Lichter, Benedikt Pago, & Tim Seppelt. *Limitations of Game Comonads via Homomorphism Indistinguishability*. 2023. arXiv: 2308.05693v2 [cs.LO] (cit. on pp. 128, 183, 234, 236).

[111] Moritz Lichter, Benedikt Pago, & Tim Seppelt. 'Limitations of Game Comonads for Invertible-Map Equivalence via Homomorphism Indistinguishability'. In: *32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)*. Ed. by Aniello Murano & Alexandra Silva. Vol. 288. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 36:1–36:19. DOI: 10.4230/LIPIcs.CSL.2024.36 (cit. on pp. 4, 128, 131, 178, 183, 234, 236).

[112] Moritz Lichter, Simon Raßmann, & Pascal Schweitzer. *Computational complexity of the Weisfeiler–Leman dimension*. 2024. arXiv: 2402.11531v1 [cs.CC] (cit. on p. 220).

[113] Fenjin Liu & Wei Wang. 'A Note on Non-ℝ-Cospectral Graphs'. In: *The Electronic Journal of Combinatorics* 24.1 (2017), P1.48. DOI: 10.37236/6002 (cit. on p. 66).

[114] László Lovász. 'Operations with structures'. In: *Acta Mathematica Academiae Scientiarum Hungarica* 18.3 (1967), pp. 321–328. DOI: 10.1007/BF02280291 (cit. on pp. iii, v, 1, 2, 10, 31–33, 125, 183, 229).

[115] László Lovász. 'On the cancellation law among finite relational structures'. In: *Periodica Mathematica Hungarica* 1.2 (1971), pp. 145–156. DOI: 10.1007/BF02029172 (cit. on pp. 32, 179, 180).

[116] László Lovász. *Large networks and graph limits*. American Mathematical Society colloquium publications volume 60. Providence, Rhode Island: American Mathematical Society, 2012. DOI: 10.1090/coll/060 (cit. on pp. 13, 34, 41, 129, 166).

[117] Lászlo Lovász & Alexander Schrijver. 'Cones of Matrices and Set-Functions and 0–1 Optimization'. In: *SIAM Journal on Optimization* 1.2 (1991), pp. 166–190. DOI: 10.1137/0801013 (cit. on pp. 5, 230).

[118] László Lovász & Alexander Schrijver. 'Semidefinite Functions on Categories'. In: *The Electronic Journal of Combinatorics* 16.2 (2009). DOI: 10.37236/80 (cit. on p. 182).

[119] László Lovász & Balázs Szegedy. 'Contractors and connectors of graph algebras'. In: *Journal of Graph Theory* 60.1 (2009), pp. 11–30. DOI: 10.1002/jgt.20343 (cit. on pp. 41, 128, 129).

[120] Martino Lupini, Laura Mančinska, & David E. Roberson. 'Nonlocal games and quantum permutation groups'. In: *Journal of Functional Analysis* 279.5 (2020), p. 108592. DOI: 10.1016/j.jfa.2020.108592 (cit. on pp. 5, 6).

[121] Laura Maaßen. 'Representation categories of compact matrix quantum groups'. PhD thesis. Aachen: RWTH Aachen University, 2021. DOI: 10.18154/RWTH-2021-06610 (cit. on p. 230).

[122] Peter N. Malkin. 'Sherali–Adams relaxations of graph isomorphism polytopes'. In: *Discrete Optimization* 12 (2014), pp. 73–97. DOI: 10.1016/j.disopt.2014.01.004 (cit. on pp. 5–7, 27–29).

[123] Laura Mančinska & David E. Roberson. *Quantum isomorphism is equivalent to equality of homomorphism counts from planar graphs*. 2019. arXiv: 1910.06958v2 [quant-ph] (cit. on p. 230).

[124] Laura Mančinska & David E. Roberson. 'Quantum isomorphism is equivalent to equality of homomorphism counts from planar graphs'. In: *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*. IEEE, 2020, pp. 661–672. DOI: 10.1109/FOCS46700.2020.00067 (cit. on pp. 1, 2, 5, 6, 10, 35, 127, 160, 178, 180, 192, 194, 222, 227, 230).

[125] Laura Mančinska, David E. Roberson, Robert Šámal, Simone Severini, & Antonios Varvitsiotis. 'Relaxations of Graph Isomorphism'. In: *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*. Ed. by Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, & Anca Muscholl. Vol. 80. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017, 76:1–76:14. DOI: 10.4230/LIPIcs.ICALP.2017.76 (cit. on p. 96).

[126] Laura Mančinska, David E. Roberson, & Antonios Varvitsiotis. 'Graph isomorphism: physical resources, optimization models, and algebraic characterizations'. In: *Mathematical Programming* (2023). DOI: 10.1007/s10107-023-01989-7 (cit. on pp. 24, 29, 96–98, 100, 102, 105–107, 109, 117).

[127] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, & Uri Alon. 'Network motifs: simple building blocks of complex networks'. In: *Science* 298.5594 (2002), pp. 824–827. DOI: 10.1126/science.298.5594.824 (cit. on p. 151).

[128] L. Mirsky. 'Results and problems in the theory of doubly-stochastic matrices'. In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 1.4 (1963), pp. 319–334. DOI: 10.1007/BF00533407 (cit. on p. 23).

[129] Yoàv Montacute & Nihil Shah. 'The Pebble-Relation Comonad in Finite Model Theory'. In: *Logical Methods in Computer Science* 20.2 (2024). DOI: 10.46298/lmcs-20(2:9)2024 (cit. on pp. 1, 4, 32, 90, 147).

[130] Christopher Morris, Martin Ritzert, Matthias Fey, William L. Hamilton, Jan Eric Lenssen, Gaurav Rattan, & Martin Grohe. 'Weisfeiler and Leman Go Neural: Higher-Order Graph Neural Networks'. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (2019), pp. 4602–4609. DOI: 10.1609/aaai.v33i01.33014602 (cit. on pp. 1, 3).

[131] Melvyn B. Nathanson. *Elementary Methods in Number Theory*. Vol. 195. Graduate Texts in Mathematics. New York, NY: Springer New York, 2000. DOI: 10.1007/b98870 (cit. on p. 185).

[132] Jaroslav Nešetřil & Patrice Ossona de Mendez. 'Tree-depth, subgraph coloring and homomorphism bounds'. In: *European Journal of Combinatorics* 27.6 (2006), pp. 1022–1041. DOI: 10.1016/j.ejc.2005.01.010 (cit. on p. 15).

[133] Jaroslav Nešetřil & Patrice Ossona de Mendez. *Sparsity: Graphs, Structures, and Algorithms*. Algorithms and combinatorics 28. Heidelberg; New York: Springer, 2012. DOI: 10.1007/978-3-642-27875-4 (cit. on pp. 15, 178).

[134] Daniel Neuen. 'Homomorphism-Distinguishing Closedness for Graphs of Bounded Tree-Width'. In: *41st International Symposium on Theoretical Aspects of Computer Science (STACS 2024)*. Ed. by Olaf Beyersdorff, Mamadou Moustapha Kanté, Orna Kupferman, & Daniel Lokshtanov. Vol. 289. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 53:1–53:12. DOI: 10.4230/LIPIcs.STACS.2024.53 (cit. on pp. 8, 145, 149, 150, 231).

[135] Daniel Neuen & Pascal Schweitzer. 'Benchmark Graphs for Practical Graph Isomorphism'. In: *25th Annual European Symposium on Algorithms, ESA 2017, September 4-6, 2017, Vienna, Austria*. Ed. by Kirk Pruhs & Christian Sohler. Vol. 87. LIPIcs. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017, 60:1–60:14. DOI: 10.4230/LIPIcs.ESA.2017.60 (cit. on p. 131).

[136] Hoang Nguyen & Takanori Maehara. 'Graph Homomorphism Convolution'. In: *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*. Vol. 119. Proceedings of Machine Learning Research. PMLR, 2020, pp. 7306–7316. URL: http://proceedings.mlr.press/v119/nguyen20c.html (cit. on p. 151).

[137] Adam Ó Conghaile & Anuj Dawar. 'Game Comonads & Generalised Quantifiers'. In: *29th EACSL Annual Conference on Computer Science Logic (CSL 2021)*. Ed. by Christel Baier & Jean Goubault-Larrecq. Vol. 183. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021, 16:1–16:17. DOI: 10.4230/LIPIcs.CSL.2021.16 (cit. on p. 4).

[138] Ryan O'Donnell, John Wright, Chenggang Wu, & Yuan Zhou. 'Hardness of Robust Graph Isomorphism, Lasserre Gaps, and Asymmetry of Random Graphs'. In: *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*. Ed. by Chandra Chekuri. SIAM, 2014, pp. 1659–1677. DOI: 10.1137/1.9781611973402.120 (cit. on pp. 5, 29, 95).

[139] Martin Otto. *Bounded Variable Logics and Counting: A Study in Finite Models*. Vol. 9. Lecture Notes in Logic. Cambridge University Press, 2017. DOI: 10.1017/9781316716878 (cit. on pp. 3, 16).

[140] Christopher J. Pappacena. 'An Upper Bound for the Length of a Finite-Dimensional Algebra'. In: *Journal of Algebra* 197.2 (1997), pp. 535–545. DOI: 10.1006/jabr.1997.7140 (cit. on p. 54).

[141] Azaria Paz. 'An application of the Cayley–Hamilton theorem to matrix polynomials in several variables'. In: *Linear and Multilinear Algebra* 15.2 (1984), pp. 161–170. DOI: 10.1080/03081088408817585 (cit. on p. 54).

[142] Carl Pearcy. 'A complete set of unitary invariants for operators generating finite W^*-algebras of type I'. In: *Pacific Journal of Mathematics* 12.4 (1962), pp. 1405–1416. DOI: 10.2140/pjm.1962.12.1405 (cit. on pp. 54, 55, 63).

[143] A. Pultr. 'Isomorphism types of objects in categories determined by numbers of morphisms'. In: *Acta Scientiarum Mathematicarum* 35 (1973), pp. 155–160 (cit. on p. 32).

[144] Motakuri V. Ramana, Edward R. Scheinerman, & Daniel Ullman. 'Fractional isomorphism of graphs'. In: *Discrete Mathematics* 132.1 (1994), pp. 247–265. DOI: 10.1016/0012-365X(94)90241-0 (cit. on pp. 26, 34, 224).

[145] Gaurav Rattan & Tim Seppelt. 'Weisfeiler–Leman and Graph Spectra'. In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Society for Industrial and Applied Mathematics, 2023, pp. 2268–2285. DOI: 10.1137/1.9781611977554.ch87 (cit. on pp. 53, 90, 91, 233, 236).

[146] Gaurav Rattan & Tim Seppelt. *Weisfeiler–Leman and Graph Spectra*. 2023. arXiv: 2103.02972v4 [cs.DS] (cit. on pp. 53, 90, 233, 236).

[147] Sven Raum & Moritz Weber. 'The Full Classification of Orthogonal Easy Quantum Groups'. In: *Communications in Mathematical Physics* 341.3 (2016), pp. 751–779. DOI: 10.1007/s00220-015-2537-z (cit. on p. 230).

[148] Luca Reggio. 'Polyadic sets and homomorphism counting'. In: *Advances in Mathematics* 410 (2022), p. 108712. DOI: 10.1016/j.aim.2022.108712 (cit. on pp. 4, 32).

[149] James Renegar. 'Accelerated first-order methods for hyperbolic programming'. In: *Mathematical Programming* 173.1 (2019), pp. 1–35. DOI: 10.1007/s10107-017-1203-y (cit. on p. 230).

[150] David E. Roberson. *Oddomorphisms and homomorphism indistinguishability over graphs of bounded degree*. 2022. arXiv: 2206.10321v1 [math.CO] (cit. on pp. iii, v, 1, 8, 126–128, 130–137, 151, 152, 156, 159, 162, 165, 191, 230, 231).

[151] David E. Roberson & Tim Seppelt. 'Lasserre Hierarchy for Graph Isomorphism and Homomorphism Indistinguishability'. In: *50th International Colloquium on Automata, Languages, and Programming (ICALP 2023)*. Ed. by Kousha Etessami, Uriel Feige, & Gabriele Puppis. Vol. 261. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 101:1–101:18. DOI: 10.4230/LIPIcs.ICALP.2023.101 (cit. on pp. 1, 96, 128, 193, 233, 234, 236).

[152] David E. Roberson & Tim Seppelt. 'Lasserre Hierarchy for Graph Isomorphism and Homomorphism Indistinguishability'. In: *TheoretiCS* 3 (2024). DOI: 10.46298/theoretics.24.20 (cit. on pp. 29, 95, 96, 128, 193, 233, 234, 236).

[153] Neil Robertson & Paul D. Seymour. 'Graph minors. V. Excluding a planar graph'. In: *Journal of Combinatorial Theory, Series B* 41.1 (1986), pp. 92–114. DOI: 10.1016/0095-8956(86)90030-4 (cit. on pp. 178, 227).

[154] Neil Robertson & Paul D. Seymour. 'Graph Minors. XIII. The Disjoint Paths Problem'. In: *Journal of Combinatorial Theory, Series B* 63.1 (1995), pp. 65–110. DOI: 10.1006/jctb.1995.1006 (cit. on p. 113).

[155] Neil Robertson & Paul D. Seymour. 'Graph Minors. XX. Wagner's conjecture'. In: *Journal of Combinatorial Theory, Series B* 92.2 (2004): *Special Issue Dedicated to Professor W.T. Tutte*, pp. 325–357. DOI: 10.1016/j.jctb.2004.08.001 (cit. on pp. 113, 136, 192).

[156] Barkley Rosser. 'Explicit Bounds for Some Functions of Prime Numbers'. In: *American Journal of Mathematics* 63.1 (1941), p. 211. DOI: 10.2307/2371291 (cit. on p. 207).

[157] Benjamin Scheidt. 'On Homomorphism Indistinguishability and Hypertree Depth'. In: *51st International Colloquium on Automata, Languages, and Programming, ICALP 2024, July 8-12, 2024, Tallinn, Estonia*. Ed. by Karl Bringmann, Martin Grohe, Gabriele Puppis, & Ola Svensson. Vol. 297. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024, 152:1–152:18. DOI: 10.4230/LIPICS.ICALP.2024.152 (cit. on p. 32).

[158] Benjamin Scheidt & Nicole Schweikardt. 'Counting Homomorphisms from Hypergraphs of Bounded Generalised Hypertree Width: A Logical Characterisation'. In: *48th International Symposium on Mathematical Foundations of Computer Science, MFCS 2023, August 28 to September 1, 2023, Bordeaux, France*. Ed. by Jérôme Leroux, Sylvain Lombardy, & David Peleg. Vol. 272. LIPIcs. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 79:1–79:15. DOI: 10.4230/LIPICS.MFCS.2023.79 (cit. on p. 32).

[159] Edward R. Scheinerman & Daniel H. Ullman. *Fractional graph theory. A rational approach to the theory of graphs*. In collab. with Claude Berge. Wiley-Interscience series in discrete mathematics and optimization. New York: Wiley, 1997. ISBN: 978-0-471-17864-4 (cit. on pp. 5, 26).

[160] Alexander Schrijver. *Theory of Linear and Integer Programming*. Wiley-Interscience Series in Discrete Mathematics & Optimization. Chichester: Wiley, 1998. 471 pp. ISBN: 978-0-471-98232-6 (cit. on pp. 134, 189).

[161] Allen J. Schwenk. 'Almost all trees are cospectral'. In: *New directions in the theory of graphs, Proc. third Ann Arbor Conf.* University of Michigan, 1973, pp. 275–307 (cit. on p. 158).

[162] Tim Seppelt. 'Logical Equivalences, Homomorphism Indistinguishability, and Forbidden Minors'. In: *48th International Symposium on Mathematical Foundations of Computer Science (MFCS 2023)*. Ed. by Jérôme Leroux, Sylvain Lombardy, & David Peleg. Vol. 272. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 82:1–82:15. DOI: 10.4230/LIPIcs.MFCS.2023.82 (cit. on pp. 128, 160, 193, 234, 236).

[163] Tim Seppelt. *An Algorithmic Meta Theorem for Homomorphism Indistinguishability*. 2024. arXiv: 2402.08989v1 [cs.LO] (cit. on pp. 193, 235, 236).

[164] Tim Seppelt. 'An Algorithmic Meta Theorem for Homomorphism Indistinguishability'. In: *49th International Symposium on Mathematical Foundations of Computer Science (MFCS 2024)*. Ed. by Rastislav Královič & Antonín Kučera. Vol. 306. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 82:1–82:19. DOI: 10.4230/LIPIcs.MFCS.2024.82 (cit. on pp. 193, 235, 236).

[165] Tim Seppelt. 'Logical equivalences, homomorphism indistinguishability, and forbidden minors'. In: *Information and Computation* 301 (2024): *Special issue on the 48th International Symposium on Mathematical Foundations of Computer Science (MFCS 2023)*, p. 105224. DOI: 10.1016/j.ic.2024.105224 (cit. on pp. 128, 160, 193, 234, 236).

[166] Luke Sernau. 'Graph operations and upper bounds on graph homomorphism counts'. In: *Journal of Graph Theory* 87.2 (2018), pp. 149–163. DOI: 10.1002/jgt.22148 (cit. on p. 182).

[167] Paul D. Seymour. 'Hadwiger's Conjecture'. In: *Open Problems in Mathematics*. Ed. by Jr. Nash John Forbes & Michael Th. Rassias. Cham: Springer International Publishing, 2016, pp. 417–437. DOI: 10.1007/978-3-319-32162-2_13 (cit. on p. 157).

[168] Paul D. Seymour & Robin Thomas. 'Graph Searching and a Min-Max Theorem for Tree-Width'. In: *Journal of Combinatorial Theory, Series B* 58.1 (1993), pp. 22–33. DOI: 10.1006/jctb.1993.1027 (cit. on p. 145).

[169] Hanif D. Sherali & Warren P. Adams. 'A Hierarchy of Relaxations between the Continuous and Convex Hull Representations for Zero-One Programming Problems'. In: *SIAM Journal on Discrete Mathematics* 3.3 (1990), pp. 411–430. DOI: 10.1137/0403036 (cit. on pp. 5, 27).

[170] Wilhelm Specht. 'Zur Theorie der Matrizen. II.' In: *Jahresbericht der Deutschen Mathematiker-Vereinigung* 50 (1940), pp. 19–23. URL: http://resolver.sub.uni-goettingen.de/purl?GDZPPN002132982 (cit. on pp. 7, 53, 54).

[171] Maciej M. Sysło. 'Characterizations of outerplanar graphs'. In: *Discrete Mathematics* 26.1 (1979), pp. 47–53. DOI: 10.1016/0012-365X(79)90060-8 (cit. on pp. 118, 138).

[172] Gottfried Tinhofer. 'Graph isomorphism and theorems of Birkhoff type'. In: *Computing* 36.4 (1986), pp. 285–300. DOI: 10.1007/BF02240204 (cit. on pp. 5, 23, 26, 61).

[173] K. Wagner. 'Beweis einer Abschwächung der Hadwiger-Vermutung'. In: *Mathematische Annalen* 153.2 (1964), pp. 139–141. DOI: 10.1007/BF01361181 (cit. on p. 157).

[174] B. YU. Weisfeiler & A. A. Leman. 'A reduction of a graph to canonical form and an algebra arising during this reduction'. English. Trans. Russian by Grigory Ryabov. In: *NTI*. 2nd ser. (1968). URL: https://web.archive.org/web/20240406105827/https://www.iti.zcu.cz/wl2018/pdf/wl_paper_translation.pdf (cit. on pp. 3, 18).

[175] N. A. Wiegmann. 'Necessary and sufficient conditions for unitary similarity'. In: *Journal of the Australian Mathematical Society* 2.1 (1961), pp. 122–126. DOI: 10.1017/S1446788700026422 (cit. on pp. 7, 53, 54).

[176] S. L. Woronowicz. 'Tannaka–Krein duality for compact matrix pseudogroups. Twisted SU(N) groups'. In: *Inventiones mathematicae* 93.1 (1988), pp. 35–76. DOI: 10.1007/BF01393687 (cit. on p. 230).

[177] Keyulu Xu, Weihua Hu, Jure Leskovec, & Stefanie Jegelka. 'How Powerful are Graph Neural Networks?' In: *International Conference on Learning Representations*. 2018. URL: https://openreview.net/forum?id=ryGs6iA5Km (cit. on pp. 1, 3).

[178] Doron Zeilberger. 'A combinatorial proof of Newton's identities'. In: *Discrete Mathematics* 49.3 (1984), p. 319. DOI: 10.1016/0012-365X(84)90171-7 (cit. on p. 45).

[179] V. N. Zemlyachenko, N. M. Korneenko, & R. I. Tyshkevich. 'Graph isomorphism problem'. In: *Journal of Soviet Mathematics* 29.4 (1985), pp. 1426–1481. DOI: 10.1007/BF02104746 (cit. on p. 224).

[180] Bohang Zhang, Jingchu Gai, Yiheng Du, Qiwei Ye, Di He, & Liwei Wang. 'Beyond Weisfeiler–Lehman: A Quantitative Framework for GNN Expressiveness'. In: *The Twelfth International Conference on Learning Representations*. 2024. URL: https://openreview.net/forum?id=HSKaGOi7Ar (cit. on pp. 1, 125).