Counting Homomorphisms from Hypergraphs of Bounded Generalised Hypertree Width: A Logical Characterisation

Benjamin Scheidt
Humboldt-Universität zu Berlin, Germany

Nicole Schweikardt
Humboldt-Universität zu Berlin, Germany

Abstract

We introduce the 2-sorted counting logic GC^k and its restriction RGC^k that express properties of hypergraphs. These logics have available k variables to address hyperedges, an unbounded number of variables to address vertices of a hypergraph, and atomic formulas $E(e,v)$ to express that a vertex v is contained in a hyperedge e. We show that two hypergraphs H, H' satisfy the same sentences of the logic RGC^k if, and only if, they are homomorphism indistinguishable over the class of hypergraphs of generalised hypertree width at most k. Here, H, H' are called homomorphism indistinguishable over a class $\mathcal{C}$ if for every hypergraph $G \in \mathcal{C}$ the number of homomorphisms from G to H equals the number of homomorphisms from G to H'. This result can be viewed as a lifting (from graphs to hypergraphs) of a result by Dvořák (2010) stating that any two (undirected, simple, finite) graphs H, H' are indistinguishable by the $k{+}1$-variable counting logic C^{k+1} if, and only if, they are homomorphism indistinguishable over the class of graphs of tree-width at most k.

2012 ACM Subject Classification Theory of computation → Finite Model Theory; Mathematics of computing → Hypergraphs

Keywords and phrases counting logics, guarded logics, homomorphism counting, hypertree decompositions, hypergraphs, incidence graphs, quantum graphs

Digital Object Identifier 10.4230/LIPIcs.MFCS.2023.79

Related Version *Extended Version*: `https://arxiv.org/abs/2303.10980` [22]

Funding *Nicole Schweikardt*: Partially supported by the ANR project EQUUS ANR-19-CE48- 0019; funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 431183758 (gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 431183758).

Acknowledgements We thank Isolde Adler for pointing us to the results in [2, 1, 3], which led to Theorem 2.3.

1 Introduction

Counting homomorphisms from a given class $\mathcal{C}$ of graphs induces a similarity measure between graphs: Consider an arbitrary graph H. The results of the homomorphism counts for all $G \in \mathcal{C}$ in H can be represented by a mapping (or, "vector") $\mathsf{HOM}_{\mathcal{C}}(H)$ that associates with every $G \in \mathcal{C}$ the number $\hom(G,H)$ of homomorphisms from G to H. A similarity measure for the mappings $\mathsf{HOM}_{\mathcal{C}}(H)$ and $\mathsf{HOM}_{\mathcal{C}}(H')$ can then be viewed as a similarity measure of two given graphs H, H'. An overview of this approach, its relations to *graph neural networks*, and its usability as a similarity measure of graphs can be found in Grohe's survey [17].

Two graphs H, H' are viewed as "equivalent" (or, *indistinguishable*) *over* $\mathcal{C}$ if $\mathsf{HOM}_{\mathcal{C}}(H) = \mathsf{HOM}_{\mathcal{C}}(H')$, i.e., for every graph G in $\mathcal{C}$ the number of homomorphisms from G to H equals the number of homomorphisms from G to H'. A classical result by Lovász [19] shows that two graphs H, H' are indistinguishable over the class of *all* graphs if, and only if, they are isomorphic. This inspired a lot of research in recent years, examining the notion of *homomorphism indistinguishability over a class* $\mathcal{C}$ for various classes $\mathcal{C}$ [11, 10, 6, 5, 7, 16, 21]. In particular, Grohe [16] proved that two graphs are homomorphism indistinguishable over the class of graphs of *tree-depth* $\leqslant k$ if, and only if, they are indistinguishable by sentences of first-order counting logic C of quantifier-rank $\leqslant k$ (C is the extension of first-order logic with counting quantifiers of the form $\exists^{\geqslant n}x$ meaning "there exist at least n elements x"). A decade earlier, Dvořák [11] proved that two graphs are homomorphism indistinguishable over the class of graphs of *tree-width* $\leqslant k$ if, and only if, they are indistinguishable by sentences of the k+1-variable fragment C^{k+1} of C. From Cai, Fürer, Immerman [8] we know that this precisely coincides with indistinguishability by the k-dimensional Weisfeiler-Leman algorithm.

An obvious question is if and how these results can be lifted from graphs to hypergraphs. A first answer was given by Böker in [7]: He introduces a new version of a *color refinement* algorithm on hypergraphs and proves that two hypergraphs H, H' cannot be distinguished by this algorithm if, and only if, they are homomorphism indistinguishable over the class of *Berge-acyclic* hypergraphs. This is a lifting – from graphs to hypergraphs – of the result of [11, 8] for the case $k = 1$ (i.e., trees) to "tree-like" hypergraphs. Note that there are different concepts of "tree-likeness" for hypergraphs. Berge-acyclicity is a rather restricted one; it is subsumed by the more general concept of α-acyclic hypergraphs, which coincides with the hypergraphs of *generalised hypertree width* 1 (cf., [13, 14, 12]).

This paper gives a further answer to the above question: For arbitrary $k \geqslant 1$ let GHW_k be the class of hypergraphs of generalised hypertree width $\leqslant k$. Our main result provides a logical characterisation of homomorphism indistinguishability over the class GHW_k. We introduce a new logic called GC^k and a restriction RGC^k of GC^k and show that two hypergraphs are homomorphism indistinguishable over GHW_k if, and only if, they are indistinguishable by sentences of the logic RGC^k.

GC^k is a 2-sorted counting logic for expressing properties of hypergraphs. It has available k "blue" variables to address edges, and an unbounded number of "red" variables to address vertices of a hypergraph, and atomic formulas $E(e, v)$ to express that vertex v is contained in edge e, as well as atomic formulas $e = e'$ and $v = v'$ for expressing equality of edge or vertex variables. Counting quantifiers are of the form $\exists^{\geqslant n}\overline{z}$ where $\overline{z} = (z_1, \dots, z_\ell)$ is either a tuple of edge variables or a tuple of vertex variables; and their meaning is "there exist at least n tuples $\overline{z}$". In the logic GC^k, each vertex variable v has to be *guarded* by an edge variable e and an atomic statement $E(e, v)$ (meaning that vertex v is included in edge e); the use of quantifiers is restricted in a way to ensure that guards are always present. Our design of the logic GC^k is somewhat inspired by the *guarded fragment of first-order logic* (cf., [4, 15, 14]). RGC^k imposes certain restrictions on the way guards of red variables can change between quantifications in a formula.

Our main result can be viewed as a lifting – from graphs to hypergraphs – of Dvořák's [11] result: Dvořák proves that two graphs are homomorphism indistinguishable over the class TW_k of graphs of tree-width $\leqslant k$ iff they are indistinguishable by the logic C^{k+1}. We prove that two hypergraphs are homomorphism indistinguishable over the class GHW_k of hypergraphs of generalised hypertree width $\leqslant k$ iff they are indistinguishable by the logic RGC^k. This is analogous (although not tightly related) to the following classical results: Kolaitis and Vardi [18] proved that the conjunctive queries of tree-width $\leqslant k$ are precisely the queries expressible in the k+1-variable fragment of a certain subclass L of first-order logic. Gottlob et al. [14] proved that the conjunctive queries of hypertree width $\leqslant k$ are precisely the ones expressible in the k-guarded fragment of L. This is somehow parallel to our result lifting Dvořák's characterisation; it is what initially gave us the confidence to work on our hypothesis.

The proof of our theorem is at its core very similar to Dvořák's proof – but it is far from straightforward. Before being able to follow along the lines of Dvořák's proof, we first have to perform a number of reduction steps and build the necessary machinery. The first step is to move over from homomorphisms on hypergraphs to homomorphisms on incidence graphs. Fortunately, Böker [7] already implicitly achieved what is needed in our setting. The result is: Two hypergraphs H, H' are homomorphism indistinguishable over the class GHW_k iff their incidence graphs I, I' are homomorphism indistinguishable over the class IGHW_k of incidence graphs of generalised hypertree width $\leqslant k$; see Section 3.

Next, for an inductive proof in the spirit of Dvořák, we would need an inductive characterisation of the class IGHW_k in the spirit of [9]. Unfortunately, generalised hypertree decompositions seem to be unsuitable for such a characterisation. That is why we work with severely restricted decompositions that we call *entangled hypertree decompositions* (ehds). In Section 4 we prove that homomorphism indistinguishability over the class IGHW_k coincides with homomorphism indistinguishability over the class IEHW_k of incidence graphs of *entangled hypertree width* $\leqslant k$. In our opinion this is interesting on its own, since the requirements of ehds are quite harsh and $\mathit{IEHW}_k \subsetneq \mathit{IGHW}_k$ for arbitrarily large k.

In Section 6 we introduce the logic GC^k and its restriction RGC^k. The inductive characterisation of IEHW_k follows in Section 7, where we also provide the machinery of *quantum* incidence graphs as an analogue of the quantum graphs used in Dvořák's proof, tailored towards our setting. In Section 8 we prove that two incidence graphs I, I' are indistinguishable by the logic RGC^k if, and only if, they are homomorphism indistinguishable over the class IEHW_k. This is achieved by two inductive proofs: We use the inductive characterisation of IEHW_k to show that for every incidence graph J in IEHW_k and every $m \in \mathbb{N}$ there exists an RGC^k-sentence that is satisfied by an incidence graph I iff there are precisely m homomorphisms from J to I. For the opposite direction, we proceed by induction on the definition of RGC^k and construct for every sentence χ in RGC^k and certain size parameters $m, d \in \mathbb{N}$ a quantum incidence graph Q in IEHW_k satisfying the following: for all incidence graphs I that match the size parameters m, d, the number $\hom(Q, I)$ of homomorphisms from Q to I is either 0 or 1, and it is 1 if and only if I satisfies the sentence χ. Both proofs are quite intricate, and the details of the syntax definition of RGC^k had to be tweaked right in order to enable proving *both* directions.

Plugging together the results achieved in the previous sections yields our main theorem, provided in Section 9: Two hypergraphs are homomorphism indistinguishable over the class GHW_k of hypergraphs of generalised hypertree width $\leqslant k$ iff they are indistinguishable by the logic RGC^k.

Due to space limitations, many proof details had to be deferred to the paper's extended version [22].

2 Preliminaries

This section provides basic notions concerning hypergraphs, incidence graphs, hypertree decompositions, and homomorphisms. We write $\mathbb{R}$ for the set of reals, $\mathbb{N}$ for the set of non-negative integers, and we let $\mathbb{N}_{\geqslant 1} \coloneqq \mathbb{N} \setminus \{0\}$ and $[n] \coloneqq \{1, \dots, n\}$ for all $n \in \mathbb{N}_{\geqslant 1}$.

Hypergraphs. The hypergraphs considered in this paper are generalisations of ordinary undirected graphs, where each edge can consist of an arbitrary number of vertices. For our proofs it will be necessary to deal with hypergraphs in which the same edge can have multiple occurrences. Furthermore, it will be convenient to assume that every vertex belongs to at least one edge. This is provided by the following definition that is basically taken from [7].

A *hypergraph* $H \coloneqq (V(H), E(H), f_H)$ consists of disjoint finite sets $V(H)$ of *vertices* and $E(H)$ of edges, and an *incidence function* f_H associating with every $e \in E(H)$ the set $f_H(e) \subseteq V(H)$ of vertices incident with edge e, such that $V(H) = \bigcup_{e\in E(H)} f_H(e)$. A *simple hypergraph* is a hypergraph H where the function f_H is injective. We can identify the edges of a hypergraph H with the multiset $M_H \coloneqq \{\!\{ f_H(e) : e \in E(H) \}\!\}$; the number of occurrences of a set $s \subseteq V(H)$ in this multiset then is the number of occurrences of "edge s" in H. The *simple* hypergraphs are the hypergraphs in which every "edge s" has only one occurrence.

Every hypergraph $H = (V(H), E(H), f_H)$ can be represented by an ordinary, bipartite graph I_H in the following way: The vertices $v \in V(H)$ occur as *red* nodes of I_H, i.e., $R(I_H) \coloneqq V(H)$. The edges $e \in E(H)$ occur as *blue* nodes of I_H, i.e., $B(I_H) \coloneqq E(H)$. And there is an edge from each blue node e to all red nodes $v \in f_H(e)$. I.e., $E(I_H) \coloneqq \{(e, v) \in B(I_H) \times R(I_H) : v \in f_H(e)\}$. The condition $V(H) = \bigcup_{e\in E(H)} f_H(e)$ implies that every red node is adjacent to at least one blue node. It is straighforward to see that the mapping $H \mapsto I_H$ provides a bijection between the class of all hypergraphs and the class of all *incidence graphs*, where the notion of incidence graphs is as follows.

An *incidence graph* $I = (R(I), B(I), E(I))$ consists of disjoint finite sets $R(I)$ and $B(I)$ of *red* nodes and *blue* nodes, resp., and a set of edges $E(I) \subseteq B(I) \times R(I)$, such that each red node is adjacent to at least one blue node. As usual for graphs, the *neighbourhood* of a node v is the set $N_I(v)$ of all nodes adjacent to v. Thus, if I is the incidence graph I_H of a hypergraph H, the neighbourhood of every blue node e is $N_I(e) = f_H(e)$, i.e., the set of all vertices of H that are incident with edge e. The neighbourhood of every red node v is $N_I(v) = \{e \in E(H) : v \in f_H(e)\}$, i.e., the set of all edges of H that are incident with vertex v. Two incidence graphs I, I' are *isomorphic* ($I \cong I'$, for short) if there exists an *isomorphism* $\pi = (\pi_R, \pi_B)$ *from I to I'*, i.e, bijections $\pi_R : R(I) \to R(I')$ and $\pi_B : B(I) \to B(I')$ such that for all $(e, v) \in B(I) \times R(I)$ we have: $(e, v) \in E(I) \iff (\pi_B(e), \pi_R(v)) \in E(I')$. We sometimes drop the subscript and write $\pi(e)$ and $\pi(v)$ instead of $\pi_B(e)$ and $\pi_R(v)$.

Generalised Hypertree Decompositions. We use the same notation as [12] for decompositions of hypergraphs, but we write *bag*(t) and *cover*(t) instead of $\chi(t)$ and $\lambda(t)$, respectively, and we formalise them with respect to incidence graphs rather than hypergraphs.

▶ **Definition 2.1.** A *complete generalised hypertree decomposition* (*ghd*, for short) of an incidence graph I is a tuple $D \coloneqq (T, \mathit{bag}, \mathit{cover})$, where $T \coloneqq (V(T), E(T))$ is a finite undirected tree, and *bag* and *cover* are mappings that associate with every tree-node $t \in V(T)$ a set $\mathit{bag}(t) \subseteq R(I)$ of red nodes of I and a set $\mathit{cover}(t) \subseteq B(I)$ of blue nodes of I, having the following properties:

1. *Completeness*: For each $e \in B(I)$ there is a $t \in V(T)$ with $N_I(e) \subseteq \mathit{bag}(t)$ and $e \in \mathit{cover}(t)$.
2. *Connectedness for red nodes*: For every $v \in R(I)$ the subgraph T_v of T induced on $V_v \coloneqq \{t \in V(T) : v \in \mathit{bag}(t)\}$ is a tree.
3. *Covering of Bags*: For every $t \in V(T)$ we have $\mathit{bag}(t) \subseteq \bigcup_{e\in \mathit{cover}(t)} N_I(e)$.

It is straightforward to see that this notion of a ghd of an incidence graph I coincides with the classical notion (cf., [13, 12]) of a complete generalised hypertree decomposition of a hypergraph H where $I_H = I$. The *width* $\mathrm{w}(D)$ *of a ghd* D is defined as the maximum number of blue nodes in the cover of a tree-node, i.e., $\mathrm{w}(D) \coloneqq \max\{|\mathit{cover}(t)| : t \in V(T)\}$. We write $\mathit{ghds}(I)$ to denote the class of all ghds of an incidence graph I. The *generalised hypertree width* of an incidence graph I is $\mathrm{ghw}(I) \coloneqq \min\{\mathrm{w}(D) : D \in \mathit{ghds}(I)\}$. By IGHW_k we denote the class of all incidence graphs of generalised hypertree width $\leqslant k$. It is straightforward to see that $\mathrm{ghw}(I_H)$ coincides with the classical notion (cf., [12]) of generalised hypertree width of a hypergraph H, and IGHW_k is the class of incidence graphs I_H of all hypergraphs H of generalised hypertree width $\leqslant k$.

For our proofs we need ghds with specific further properties, defined as follows; we are not aware of any related work that studies this particular kind of decompositions. Hypertree decompositions satisfying condition 4 (but not necessarily condition 5) of Definition 2.2 are known as *strong* decompositions [14].

▶ **Definition 2.2.** An *entangled hypertree decomposition* (*ehd*, for short) of an incidence graph I is a ghd D of I that additionally satisfies the following requirements:

4. *Precise coverage of bags*: For all tree-nodes $t \in V(T)$ we have $\bigcup_{e \in cover(t)} N_I(e) = bag(t)$.
5. *Connectedness for blue nodes*: For every $e \in B(I)$ the subgraph T_e of T induced on $V_e := \{t \in V(T) : e \in cover(t)\}$ is a tree.

We write *ehds*(I) to denote the class of all ehds of an incidence graph I.

The *entangled hypertree width* of an incidence graph I is $\mathrm{ehw}(I) := \min\{\mathrm{w}(D) : D \in ehds(I)\}$. For a hypergraph H we let $\mathrm{ehw}(H) := \mathrm{ehw}(I_H)$. By $IEHW_k$ we denote the class of all incidence graphs of entangled hypertree width $\leqslant k$.

Applying results from [2, 1, 3] shows that there exist arbitrarily large k such that $IEHW_k$ is a strict subclass of $IGHW_k$. More precisely:

▶ **Theorem 2.3.** *$IEHW_k \subseteq IGHW_k$, for every $k \in \mathbb{N}_{\geqslant 1}$. Furthermore, $IEHW_1 = IGHW_1$, but $IEHW_k \subsetneq IGHW_k$ for each $k \in \{2, 3\}$. Moreover, for every $n \in \mathbb{N}$ there exists a $k \in \mathbb{N}_{\geqslant 1}$ such that $IGHW_k \nsubseteq IEHW_{k+n}$ (and hence, $IEHW_{k+n} \subsetneq IGHW_{k+n}$).*

Proof. $IEHW_k \subseteq IGHW_k$ holds because every ehd also is a ghd. $IEHW_1 = IGHW_1$ holds because ghds of width 1 are known to be equivalent to so-called *join trees*, and these can easily be translated into ehds of width 1. For the remaining statements, we use elaborate results from [2, 1, 3] that relate the *hypertree width* $\mathrm{hw}(H)$ (cf., [13, 14]) of a hypergraph to its *generalised hypertree width* $\mathrm{ghw}(H)$:

From [2, Proposition 3.3.2] (cf. also [3, Example 3]) and [1, Claim 6.1] we obtain for each $k \in \{2, 3\}$ a simple hypergraph H_k such that $\mathrm{ghw}(H_k) = k$ and $\mathrm{hw}(H_k) = k{+}1$. Furthermore, [2, Fact 3.3.1] and [1, Theorem 4.1] provide for every $n \in \mathbb{N}_{\geqslant 1}$ a simple hypergraph H^n such that $\mathrm{hw}(H^n) = \mathrm{ghw}(H^n){+}n$.[1]

It is straightforward to verify that every ehd also is a complete *hypertree decomposition* in the sense of [13, 14]. Consequently, for every hypergraph H we have $\mathrm{hw}(H) \leqslant \mathrm{ehw}(H)$. Therefore, for each $k \in \{2, 3\}$, the incidence graph of H_k witnesses that $IEHW_k \subsetneq IGHW_k$.

To address the theorem's next statement, consider an arbitrary $n \in \mathbb{N}$. Let $H := H^{n+1}$ and let $k := \mathrm{ghw}(H)$. Then, $\mathrm{ehw}(H) \geqslant \mathrm{hw}(H) = k{+}n{+}1$. Thus, the incidence graph of H belongs to $IGHW_k$ but not to $IEHW_{k+n}$. ◀

Homomorphisms. We use the classical notions for hypergraphs and incidence graphs: A *homomorphism* from a hypergraph F to a hypergraph H is a pair (h_V, h_E) of mappings $h_V : V(F) \to V(H)$ and $h_E : E(F) \to E(H)$ such that for all $e \in E(F)$ we have $f_H(h_E(e)) = \{h_V(v) : v \in f_F(e)\}$. We write $\mathrm{Hom}(F, H)$ for the set of all homomorphisms from F to H, and $\mathrm{hom}(F, H) := |\mathrm{Hom}(F, H)|$ is the number of homomorphisms from F to H.

A *homomorphism* from an incidence graph J to an incidence graph I is a pair $h = (h_R, h_B)$ of mappings $h_R : R(J) \to R(I)$ and $h_B : B(J) \to B(I)$ such that for all $(e, v) \in E(J)$ we have $(h_B(e), h_R(v)) \in E(I)$. We sometimes drop the subscript and write $h(e)$ and $h(v)$ instead of $h_B(e)$ and $h_R(v)$. By $\mathrm{Hom}(J, I)$ we denote the set of all homomorphisms from J to I, and we let $\mathrm{hom}(J, I) := |\mathrm{Hom}(J, I)|$ be the number of homomorphisms from J to I.

[1] Note that the notions c_H-$\mathrm{hw}(H)$ and c_H-$\mathrm{ghw}(H)$ in [2] correspond to $\mathrm{hw}(H)$ and $\mathrm{ghw}(H)$ for all hypergraphs H according to [2, Example 2.1.10].

As pointed out in [7], every homomorphism from a hypergraph F to a hypergraph H also is a homomorphism from the incidence graph I_F to the incidence graph I_H; but there exist homomorphisms from I_F to I_H that do not correspond to any homomorphism from F to H. In fact, every homomorphism (h_R, h_B) from I_F to I_H is a pair of mappings $(h_V, h_E) := (h_R, h_B)$ with $h_V : V(F) \to V(H)$ and $h_E : E(F) \to E(H)$ such that for every $e \in E(F)$ we have $f_H(h_E(e)) \supseteq \{h_V(v) : v \in f_F(e)\}$ – i.e., the condition "=" of the definition of hypergraph-homomorphisms is relaxed into the condition "⊇".

3 Homomorphism Indistinguishability

Let $(B, B', \mathcal{C})$ be either two incidence graphs and a class of incidence graphs or two hypergraphs and a class of hypergraphs. By $\mathsf{HOM}_{\mathcal{C}}(B)$ we denote the function $\alpha : \mathcal{C} \to \mathbb{N}$ that associates with every $A \in \mathcal{C}$ the number $\hom(A, B)$ of homomorphisms from A to B. We say that B and B' are *homomorphism indistinguishable over* $\mathcal{C}$ if $\mathsf{HOM}_{\mathcal{C}}(B) = \mathsf{HOM}_{\mathcal{C}}(B')$. Note that $\mathsf{HOM}_{\mathcal{C}}(B) \neq \mathsf{HOM}_{\mathcal{C}}(B')$ means that there exists an $A \in \mathcal{C}$ that *distinguishes* between B and B' in the sense that $\hom(A, B) \neq \hom(A, B')$.

Recall from Section 2 that $IGHW_k$ is the class of incidence graphs of generalised hypertree width $\leqslant k$. We write GHW_k for the class of all hypergraphs of generalised hypertree width $\leqslant k$ (i.e., all hypergraphs H for which $I_H \in IGHW_k$), and $sGHW_k$ for the subclass consisting of all *simple* hypergraphs (i.e., hypergraphs where each edge has multiplicity 1) in GHW_k.

▶ **Theorem 3.1** (implicit in [7]). *Let H, H' be hypergraphs.*

(a) *If H and H' are simple hypergraphs, then*
$\mathsf{HOM}_{GHW_k}(H) = \mathsf{HOM}_{GHW_k}(H') \iff \mathsf{HOM}_{sGHW_k}(H) = \mathsf{HOM}_{sGHW_k}(H')$.

(b) $\mathsf{HOM}_{GHW_k}(H) = \mathsf{HOM}_{GHW_k}(H') \iff \mathsf{HOM}_{IGHW_k}(I_H) = \mathsf{HOM}_{IGHW_k}(I_{H'})$.

Böker [7] proved the analogous statement for BA, IBA instead of $GHW_k, IGHW_k$, where BA is the class of all Berge-acyclic hypergraphs and IBA is the class of all incidence graphs of hypergraphs in BA. Böker's proof, however, works for all classes C of hypergraphs and the associated class IC of all incidence graphs of hypergraphs in C, provided that C satisfies some mild closure properties, which GHW_k satisfies.

4 Relating $IGHW_k$ to $IEHW_k$

Recall from Section 2 that $IEHW_k \subseteq IGHW_k$, for the class $IEHW_k$ of incidence graphs of entangled hypertree width $\leqslant k$. By Theorem 2.3 there exist arbitrarily large k such that $IEHW_k$ is a strict subclass of $IGHW_k$. This section's main result is that, nevertheless:

▶ **Theorem 4.1.** *For all incidence graphs I and I' we have*
$\mathsf{HOM}_{IGHW_k}(I) = \mathsf{HOM}_{IGHW_k}(I') \iff \mathsf{HOM}_{IEHW_k}(I) = \mathsf{HOM}_{IEHW_k}(I')$.

Proof sketch. The proof heavily relies on our following technical main lemma, which uses the following notation: For an arbitrary incidence graph J, for $s \subseteq R(J)$, and for $n \in \mathbb{N}$ we write $J + n{\cdot}s$ to denote the incidence graph J' obtained from J by inserting n new blue nodes $\hat{e}_1, \ldots, \hat{e}_n$ and edges $(\hat{e}_i, v)$ for all $i \in [n]$ and all $v \in s$ – i.e., $N_{J'}(\hat{e}_i) = s$.

▶ **Lemma 4.2.** *Let J, I, I' be incidence graphs with* $\hom(J, I) \neq \hom(J, I')$*, let $e \in B(J)$, and let $s \subseteq N_J(e)$. For every $m \in \mathbb{N}$ there exists an $n \in \mathbb{N}$ with $n \geqslant m$ such that $J_n := J + n{\cdot}s$ satisfies* $\hom(J_n, I) \neq \hom(J_n, I')$.

The (combinatorially quite involved) proof of Lemma 4.2 can be found in the paper's extended version [22].

The direction "$\Longrightarrow$" of Theorem 4.1 is trivial. For the direction "$\Longleftarrow$" it suffices to prove the following: If there is a $J \in \mathit{IGHW}_k$ with $\hom(J, I) \neq \hom(J, I')$, then there also exists a $J' \in \mathit{IEHW}_k$ with $\hom(J', I) \neq \hom(J', I')$. We construct such a J' in a 2-step process. We start with a ghd $D = (T, \mathit{bag}, \mathit{cover})$ of J with $\mathrm{w}(D) \leqslant k$. First, we transform D into a ghd D^1 of an incidence graph J^1 such that $\mathrm{w}(D^1) \leqslant \mathrm{w}(D)$ and $\hom(J^1, I) \neq \hom(J^1, I')$ and D^1 satisfies condition 4 of Definition 2.2 (but condition 5 might still be violated). Afterwards, we transform D^1 into a ghd D^2 of an incidence graph J^2 such that $\mathrm{w}(D^2) = \mathrm{w}(D^1)$ and $\hom(J^2, I) \neq \hom(J^2, I')$ and D^2 satisfies conditions 4 and 5 of Definition 2.2 and hence is an ehd. Letting $J' := J^2$ then completes the proof.

For the construction of D^1, J^1 we consider all those $t \in V(T)$ and $e \in \mathit{cover}(t)$ where $N_J(e) \nsubseteq \mathit{bag}(t)$ and let $s := N_J(e) \cap \mathit{bag}(t)$. We use Lemma 4.2 to choose a suitable number $n_s \geqslant 1$ and replace J by $J + n_s \cdot s$ (let us write $e'_1, \dots e'_{n_s}$ for the n_s newly inserted blue nodes). In D we replace e with e'_1 in $\mathit{cover}(t)$, and we add new leaves t_j for $j \in \{2, \dots, n_s\}$ adjacent to t with $\mathit{cover}(t_j) = \{e'_j\}$ and $\mathit{bag}(t_j) = s$. After having done this for all combinations of t and e, we end up with the desired incidence graph J^1 and ghd $D^1 = (T^1, \mathit{bag}^1, \mathit{cover}^1)$.

For the construction of D^2, J^2, for each $e \in B(J^1)$ we let m_e be the number of connected components of the subgraph T^1_e, i.e., the subgraph of T^1 induced on $V_e := \{t \in V(T^1) : e \in \mathit{cover}^1(t)\}$. Let $V_{e,0}, \dots, V_{e,m_e-1}$ be the sets of tree-nodes (i.e., nodes in $V(T^1)$) of these connected components. We consider all those $e \in B(J^1)$ where $m_e \geqslant 2$ and let $s := N_{J^1}(e)$. We use Lemma 4.2 to choose a suitable number $n_e \geqslant m_e - 1$ and replace J with $J + n_e \cdot s$ (let us write $e'_1, \dots, e'_{n_e}$ for the n_e newly inserted blue nodes). In D^1 we consider for every $i \in \{1, \dots, m_e - 1\}$ all $t \in V_{e,i}$ and replace e with e'_i in $\mathit{cover}^1(t)$. Furthermore, we pick an arbitrary $t \in V_{e,0}$, and for each $i \in [n_e]$ with $i \geqslant m_e$, we insert into T^1 a new leaf $t_{e,i}$ adjacent to t and let $\mathit{bag}^1(t_{e,i}) := s$ and $\mathit{cover}^1(t_{e,i}) := \{e'_{e,i}\}$. After having done this for all $e \in B(J^1)$ with $m_e \geqslant 2$, we end up with the desired incidence graph J^2 and ehd D^2. This completes the proof sketch of Theorem 4.1. ◀

5 Notation for Partial Functions

We introduce some further notation that will be convenient for the remaining parts of the paper. We write $f : A \rightharpoonup B$ to indicate that f is a partial function from A to B. By $\mathrm{dom}(f)$ we denote the domain of f, i.e., the set of all $a \in A$ on which $f(a)$ is defined. By $\mathrm{img}(f)$ we denote the image of f, i.e., $\mathrm{img}(f) = \{f(a) : a \in \mathrm{dom}(f)\}$. Two partial functions $f : A \rightharpoonup B$ and $g : A \rightharpoonup B$ are called *compatible* if $f(a) = g(a)$ holds for all $a \in \mathrm{dom}(f) \cap \mathrm{dom}(g)$.

We identify a partial function f with the set $\{(a, f(a)) : a \in \mathrm{dom}(f)\}$. This allows us to compare and combine partial functions via standard notation from set theory. E.g., $f \subseteq g$ indicates that $\mathrm{dom}(f) \subseteq \mathrm{dom}(g)$ and $f(a) = g(a)$ for all $a \in \mathrm{dom}(f)$. And $f \cup g$ denotes the partial function h with $\mathrm{dom}(h) = \mathrm{dom}(f) \cup \mathrm{dom}(g)$ and $h(a) = f(a)$ for all $a \in \mathrm{dom}(f)$ and $h(a) = g(a)$ for all $a \in \mathrm{dom}(g) \setminus \mathrm{dom}(f)$; note that f has precedence over g in case that f and g are not compatible. For a set S we write $f - S$ to denote the partial function g with $g \subseteq f$ and $\mathrm{dom}(g) = \mathrm{dom}(f) \setminus S$.

6 2-Sorted Counting Logic with Guards: GC^k and RGC^k

This section provides the syntax and semantics of our 2-sorted logics GC^k and RGC^k. Formulas of these logics are evaluated on incidence graphs (cf. Section 2). We fix a $k \in \mathbb{N}_{\geqslant 1}$.

To address *blue* nodes (i.e., *edges* of a hypergraph), we have available k *blue variables* $\mathtt{e}_1, \dots, \mathtt{e}_k$. To address *red* nodes (i.e., *vertices* of a hypergraph), we have available countably many *red variables* $\mathtt{v}_1, \mathtt{v}_2, \mathtt{v}_3, \dots$. An atomic formula $E(\mathtt{e}_j, \mathtt{v}_i)$ states that a hypergraph's vertex $\mathtt{v}_i$ is included in the hypergraph's edge $\mathtt{e}_j$. Let $\mathsf{Var}_B := \{\mathtt{e}_1, \ldots, \mathtt{e}_k\}$, $\mathsf{Var}_R := \{\mathtt{v}_i : i \in \mathbb{N}_{\geqslant 1}\}$, and $\mathsf{Var} := \mathsf{Var}_B \cup \mathsf{Var}_R$. An *interpretation* $\mathcal{I} = (I, \beta)$ consists of an incidence graph $I = (R(I), B(I), E(I))$ and an *assignment* β *in* I, i.e., a mapping $\beta : \mathsf{Var} \to R(I) \cup B(I)$ with $\beta(\mathtt{e}_j) \in B(I)$ for all $\mathtt{e}_j \in \mathsf{Var}_B$ and $\beta(\mathtt{v}_i) \in R(I)$ for all $\mathtt{v}_i \in \mathsf{Var}_R$. In the formulas of our logics, red variables $\mathtt{v}_i$ have to be *guarded* by a blue variable $\mathtt{e}_j$ in the sense that $E(\mathtt{e}_j, \mathtt{v}_i)$ holds. This is formalised by a *guard function*, i.e., a *partial function* $g : \mathbb{N}_{\geqslant 1} \rightharpoonup [k]$ with *finite* domain $\mathrm{dom}(g)$. Every guard function g corresponds to the formula $\Delta_g := \bigwedge_{i \in \mathrm{dom}(g)} E(\mathtt{e}_{g(i)}, \mathtt{v}_i)$, and for the special case where $\mathrm{dom}(g) = \varnothing$ we let $\Delta_g := \top$ where $\top$ is a special atomic formula satisfied by *every* interpretation $\mathcal{I}$. We let $\mathrm{free}(\Delta_g)$ be the set of all (red or blue) variables that occur in Δ_g. An interpretation $\mathcal{I} = (I, \beta)$ *satisfies* a guard function g (in symbols: $\mathcal{I} \models \Delta_g$) if for all $i \in \mathrm{dom}(g)$ we have: $(\beta(\mathtt{e}_{g(i)}), \beta(\mathtt{v}_i)) \in E(I)$. I.e., for every $i \in \mathrm{dom}(g)$, the red variable $\mathtt{v}_i$ is guarded by the blue variable $\mathtt{e}_{g(i)}$ in the sense that it is connected to it by an edge of the incidence graph.

For any formula χ we write $\mathit{ifree}_B(\chi)$ for the set of all indices $j \in [k]$ such that the blue variable $\mathtt{e}_j$ belongs to $\mathrm{free}(\chi)$. Accordingly, $\mathit{ifree}_R(\chi) := \{i \in \mathbb{N}_{\geqslant 1} : \mathtt{v}_i \in \mathrm{free}(\chi)\}$. The definition of the syntax of GC^k is inductively given as follows.

Base cases: The atomic formulas in GC^k are of the form $\top$, $E(\mathtt{e}_j, \mathtt{v}_i)$, $\mathtt{e}_j{=}\mathtt{e}_{j'}$, and $\mathtt{v}_i{=}\mathtt{v}_{i'}$ for $j, j' \in [k]$ and $i, i' \in \mathbb{N}_{\geqslant 1}$.

Inductive cases:
1. If $\psi \in \mathsf{GC}^k$, then $\neg\psi \in \mathsf{GC}^k$.
2. If $\psi_1, \psi_2 \in \mathsf{GC}^k$, then $(\psi_1 \wedge \psi_2) \in \mathsf{GC}^k$.
3. If $\psi \in \mathsf{GC}^k$ and g is a guard function with $\mathrm{dom}(g) = \mathit{ifree}_R(\psi)$ and $n, \ell \in \mathbb{N}_{\geqslant 1}$ and, for $\chi := (\Delta_g \wedge \psi)$ and $i_1 < \cdots < i_\ell$ with
   (a) $i_1, \ldots, i_\ell \in \mathit{ifree}_R(\chi)$, then $\varphi \in \mathsf{GC}^k$ for $\varphi := \exists^{\geqslant n}(\mathtt{v}_{i_1}, \ldots, \mathtt{v}_{i_\ell}).(\Delta_g \wedge \psi)$;
   (b) $i_1, \ldots, i_\ell \in \mathit{ifree}_B(\chi)$, then $\varphi \in \mathsf{GC}^k$ for $\varphi := \exists^{\geqslant n}(\mathtt{e}_{i_1}, \ldots, \mathtt{e}_{i_\ell}).(\Delta_g \wedge \psi)$.

The semantics are defined as expected. In particular, an interpretation $\mathcal{I} = (I, \beta)$ satisfies the formula $\varphi := \exists^{\geqslant n}(\mathtt{v}_{i_1}, \ldots, \mathtt{v}_{i_\ell}).(\Delta_g \wedge \psi)$ iff there are at least n tuples $(v_{i_1}, \ldots, v_{i_\ell}) \in R(I)^\ell$ such that $\mathcal{I}' = (I, \beta')$ satisfies $(\Delta_g \wedge \psi)$, where $\beta'(\mathtt{v}_{i_j}) = v_{i_j}$ for all $j \in [\ell]$ and $\beta'(\mathtt{x}) = \beta(\mathtt{x})$ for all $\mathtt{x} \in \mathsf{Var} \setminus \{\mathtt{v}_{i_1}, \ldots, \mathtt{v}_{i_\ell}\}$. Similarly, $\mathcal{I} = (I, \beta)$ satisfies $\varphi := \exists^{\geqslant n}(\mathtt{e}_{i_1}, \ldots, \mathtt{e}_{i_\ell}).(\Delta_g \wedge \psi)$ iff there are at least n tuples $(e_{i_1}, \ldots, e_{i_\ell}) \in B(I)^\ell$ such that $\mathcal{I}' = (I, \beta')$ satisfies $(\Delta_g \wedge \psi)$, where $\beta'(\mathtt{e}_{i_j}) = e_{i_j}$ for all $j \in [\ell]$ and $\beta'(\mathtt{x}) = \beta(\mathtt{x})$ for all $\mathtt{x} \in \mathsf{Var} \setminus \{\mathtt{e}_{i_1}, \ldots, \mathtt{e}_{i_\ell}\}$. Obviously we can emulate the $\forall$-quantifier (and disjunction) using $\exists^{\geqslant 1}$ and $\neg$ (and $\wedge$ and $\neg$, respectively).

We write $\mathcal{I} \models \chi$ to indicate that $\mathcal{I}$ satisfies the formula χ; and $\mathcal{I} \not\models \chi$ indicates that $\mathcal{I}$ does not satisfy χ. *Sentences* of GC^k are formulas $\chi \in \mathsf{GC}^k$ with $\mathrm{free}(\chi) = \varnothing$. For an incidence graph I and a sentence $\chi \in \mathsf{GC}^k$ we write $I \models \chi$ to indicate that $\mathcal{I} \models \chi$ where $\mathcal{I} = (I, \beta)$ for any assignment β in I (since χ has no free variable, the assignment does not matter). For a hypergraph H and a sentence $\chi \in \mathsf{GC}^k$ we write $H \models \chi$ to indicate that $I_H \models \chi$. For two incidence graphs I and I' we write $I \equiv_{\mathsf{GC}^k} I'$ and say that I and I' are *indistinguishable by the logic* GC^k if for all sentences $\chi \in \mathsf{GC}^k$ we have: $I \models \chi \iff I' \models \chi$.

Let us now introduce a restriction of GC^k that we call RGC^k. Every formula of RGC^k will be of the form $(\Delta_g \wedge \psi)$, where g is a guard function whose domain $\mathrm{dom}(g)$ consists of all indices $i \in \mathbb{N}_{\geqslant 1}$ such that the red variable $\mathtt{v}_i$ is a free variable of ψ. We let $\mathrm{free}((\Delta_g \wedge \psi)) := \mathrm{free}(\Delta_g) \cup \mathrm{free}(\psi)$ denote the set of free variables of the formula. The definition of the syntax of RGC^k is inductively given as follows.

Base cases: $(\Delta_g \wedge \psi) \in \mathsf{RGC}^k$ for all ψ and all $g : \mathbb{N}_{\geqslant 1} \rightharpoonup [k]$ matching one of the following:
1. ψ is $E(\mathtt{e}_j, \mathtt{v}_i)$ and $\mathrm{dom}(g) = \{i\}$ and $j \in [k]$ (note that $g(i) \in [k]$ can be chosen arbitrarily);
2. ψ is $\mathtt{e}_j{=}\mathtt{e}_{j'}$ with $\mathrm{dom}(g) = \varnothing$ and $j, j' \in [k]$;
3. ψ is $\mathtt{v}_i{=}\mathtt{v}_{i'}$ with $\mathrm{dom}(g) = \{i, i'\}$.

Inductive cases:

4. If $(\Delta_g \wedge \psi) \in \mathsf{RGC}^k$, then $(\Delta_g \wedge \neg\psi) \in \mathsf{RGC}^k$;
5. If $(\Delta_{g_i} \wedge \psi_i) \in \mathsf{RGC}^k$ for $i \in [2]$ and g_1 and g_2 are compatible (i.e., they agree on $\operatorname{dom}(g_1) \cap \operatorname{dom}(g_2)$), then $(\Delta_g \wedge \varphi) \in \mathsf{RGC}^k$ for $g := g_1 \cup g_2$ and $\varphi := (\psi_1 \wedge \psi_2)$;
6. If $(\Delta_g \wedge \psi) \in \mathsf{RGC}^k$ and $n, \ell \in \mathbb{N}_{\geqslant 1}$, and $i_1, \ldots, i_\ell \in \operatorname{dom}(g)$ with $i_1 < \cdots < i_\ell$, then $(\Delta_{\tilde{g}} \wedge \varphi) \in \mathsf{RGC}^k$ for $\varphi := \exists^{\geqslant n}(\mathtt{v}_{i_1}, \ldots, \mathtt{v}_{i_\ell}).(\Delta_g \wedge \psi)$ and $\tilde{g} := g - \{i_1, \ldots, i_\ell\}$ (note that $\operatorname{free}(\varphi) = \operatorname{free}((\Delta_g \wedge \psi)) \setminus \{\mathtt{v}_{i_1}, \ldots, \mathtt{v}_{i_\ell}\}$);
7. If $(\Delta_g \wedge \psi) \in \mathsf{RGC}^k$ and $n, \ell \in \mathbb{N}_{\geqslant 1}$, and $S := \{i_1, \ldots, i_\ell\} \subseteq \mathit{ifree}_B(\chi)$ for $\chi := (\Delta_g \wedge \psi)$ with $i_1 < \cdots < i_\ell$, and if $\tilde{g} : \mathbb{N}_{\geqslant 1} \rightharpoonup [k]$ with $\operatorname{dom}(\tilde{g}) = \operatorname{dom}(g)$ such that all $i \in \operatorname{dom}(g)$ satisfy
$$\tilde{g}(i) = g(i) \quad \text{or} \quad \tilde{g}(i) \in S \quad \text{or} \quad \tilde{g}(i) \notin \operatorname{img}(g)\,, \quad \text{then} \tag{1}$$
$(\Delta_{\tilde{g}} \wedge \varphi) \in \mathsf{RGC}^k$ for $\varphi := \exists^{\geqslant n}(\mathtt{e}_{i_1}, \ldots, \mathtt{e}_{i_\ell}).(\Delta_g \wedge \psi)$ (here, $\operatorname{free}(\varphi) = \operatorname{free}(\chi) \setminus \{\mathtt{e}_{i_1}, \ldots, \mathtt{e}_{i_\ell}\}$).

Let us have a closer look at rule 7): The formula φ has exactly the same free red variables as the formula χ. But the guard of red variable $\mathtt{v}_i$ in $\tilde{\chi} := (\Delta_{\tilde{g}} \wedge \varphi)$ is $j' := \tilde{g}(i)$, whereas in χ it is $j := g(i)$. Condition (1) is equivalent to the following: the guard remains unchanged (i.e., $j'{=}j$), or the new guard j' has become "available" by the quantification (i.e., $j' \in S$) or it has not been used as a guard by g (i.e., $j' \notin \operatorname{img}(g)$).

Note that $\mathsf{RGC}^k \subseteq \mathsf{GC}^k$. Furthermore, for all $\chi := (\Delta_g \wedge \psi) \in \mathsf{RGC}^k$ we have $\operatorname{dom}(g) = \{i \in \mathbb{N}_{\geqslant 1} : \mathtt{v}_i \in \operatorname{free}(\chi)\}$. *Sentences* of RGC^k are formulas $\chi := (\Delta_g \wedge \psi)$ in RGC^k with $\operatorname{free}(\chi) = \varnothing$. Since $\operatorname{dom}(g) = \{i \in \mathbb{N}_{\geqslant 1} : \mathtt{v}_i \in \operatorname{free}(\chi)\}$, this implies that $\operatorname{dom}(g) = \varnothing$, i.e., $g = g_\varnothing$ where $g_\varnothing$ is the uniquely defined partial mapping with empty domain; recall that $\Delta_{g_\varnothing} = \top$. For two incidence graphs I and I' we write $I \equiv_{\mathsf{RGC}^k} I'$ and say that I and I' are *indistinguishable by the logic* RGC^k if for all sentences $\chi \in \mathsf{RGC}^k$ we have: $I \models \chi \iff I' \models \chi$. The subsequent sections of this paper are devoted to proving the following theorem, stating that indistinguishability by the logic RGC^k coincides with homomorphism indistinguishability over the class IEHW_k of incidence graphs of entangled hypertree width at most k.

▶ **Theorem 6.1.** *For all incidence graphs I, I' and all $k \in \mathbb{N}_{\geqslant 1}$ we have:*
$I \equiv_{\mathsf{RGC}^k} I' \iff \mathsf{HOM}_{\mathit{IEHW}_k}(I) = \mathsf{HOM}_{\mathit{IEHW}_k}(I')$.

This result can be viewed as a lifting of Dvořák's theorem [11] stating that any two graphs G, G' are indistinguishable by the $k{+}1$-variable logic C^{k+1} if, and only if, they are homomorphism indistinguishable over the class TW_k of graphs of tree-width $\leqslant k$. Our proof of Theorem 6.1 is heavily inspired by Dvořák's proof. But in order to proceed along a similar construction, we first have to provide a suitable inductive characterisation of the class IEHW_k. This is presented in Section 7, where we also provide the machinery of *quantum* incidence graphs as an analogue of the quantum graphs used in Dvořák's proof. Section 8 is devoted to the proof of Theorem 6.1.

Before we close this section, let us have a look at some examples. Let $k = 2$. Consider the following formula $\psi_1 \in \mathsf{GC}^k$:

$$\psi_1 \;:=\; \exists^{\geqslant 1}(\mathtt{v}_1).\big(E(\mathtt{e}_1, \mathtt{v}_1) \wedge E(\mathtt{e}_2, \mathtt{v}_1)\big).$$

ψ_1 expresses that the hyperedges $\mathtt{e}_1$ and $\mathtt{e}_2$ share at least one vertex $\mathtt{v}_1$, i.e. they intersect. Since we quantify over $\mathtt{v}_1$, the definition of GC^k requires us to insert a guard ranging over the set of free red variables, i.e. over $\{\mathtt{v}_1\}$. We chose $E(\mathtt{e}_1, \mathtt{v}_1)$ as the guard but note that $E(\mathtt{e}_2, \mathtt{v}_1)$ would have been a valid choice as well. Next, consider the formula $\psi_2 \in \mathsf{GC}^k$:

$$\psi_2 \;:=\; \bigwedge_{j \in \{1,2\}} \exists^{\geqslant 3}(\mathtt{v}_1).\big(E(\mathtt{e}_j, \mathtt{v}_1) \wedge E(\mathtt{e}_j, \mathtt{v}_1)\big).$$

ψ_2 expresses that each of the hyperedges $\mathtt{e}_1$, $\mathtt{e}_2$ contains at least three vertices. Again, we have to insert a guard after the quantifier, which is why $E(\mathtt{e}_j, \mathtt{v}_1)$ appears twice in ψ_2 – as a guard *and* as our "actual" subformula.

Finally, we use the formulas ψ_1, ψ_2 to construct a sentence $\varphi \in \mathsf{GC}^k$:

$$\varphi \;\; := \;\; \neg\, \exists^{\geqslant 1}(\mathtt{e}_1, \mathtt{e}_2).\big(\top \wedge \big(\, (\psi_1 \wedge \psi_2) \; \wedge \neg\; \mathtt{e}_1{=}\mathtt{e}_2 \,\big)\big).$$

φ expresses that there is no pair of non-equal hyperedges $(\mathtt{e}_1, \mathtt{e}_2)$ that intersect and that both contain at least 3 vertices. I.e., φ expresses that all hyperedges that contain at least 3 vertices are pairwise disjoint. Once again, the quantification requires us to insert a guard; since there are no free red variables, we insert $\top$ as the guard. Note that $(\top \wedge \varphi) \in \mathsf{RGC}^k$. For more examples of formulas in GC^k and RGC^k, consult the paper's extended version [22].

7 An Inductive Characterisation of *IEHW*$_k$

In this section we give an inductive definition of what we call *guarded k-labeled incidence graphs* (*GLI*$_k$), and we show that these are equivalent to the incidence graphs in *IEHW*$_k$. Throughout this section, we fix an arbitrary number $k \in \mathbb{N}_{\geqslant 1}$.

k-Labeled Incidence Graphs and the Class *GLI*$_k$. We enrich an incidence graph I by labeling some of its blue (red) nodes with labels in $[k]$ (in $\mathbb{N}_{\geqslant 1}$), and by providing, for every $i \in \mathbb{N}_{\geqslant 1}$ that is used as a label for a red node, a "blue label" $g(i) \in [k]$ that should be regarded as "the guard" of i. Each label can only be used once, not all labels have to be used, not all vertices have to be labeled, one vertex may have multiple labels, and "guards" can be chosen arbitrarily. This is formalised as follows: A *k-labeled incidence graph* $L = (I, r, b, g)$ consists of an incidence graph I and partial mappings $r : \mathbb{N}_{\geqslant 1} \rightharpoonup R(I)$, $b : [k] \rightharpoonup B(I)$, and $g : \mathbb{N}_{\geqslant 1} \rightharpoonup [k]$ such that $\mathrm{dom}(g) = \mathrm{dom}(r)$ is finite. We write I_L, r_L, b_L, g_L to address L's components I, r, b, g. Let $L = (I, r, b, g)$ be a k-labeled incidence graph. If $j \in \mathrm{dom}(b)$, the blue node $b(j)$ of I is labeled with the number j. If $i \in \mathrm{dom}(r)$, the red node $r(i)$ of I is labeled with the number i, and $g(i) = j$ indicates that the blue node labeled with the number j (if it exists) should be regarded as "the guard" of the red node labeled with the number i.

We say that L *has real guards* if for every $i \in \mathrm{dom}(r)$ the red node v labeled i is "guarded" by the blue node e labeled $j := g(i)$ in the sense that I contains an edge from e to v. This is formalised as follows: A k-labeled incidence graph $L = (I, r, b, g)$ is said to *have real guards w.r.t. f* for a partial function $f : \mathbb{N}_{\geqslant 1} \rightharpoonup [k]$ if $\mathrm{dom}(f) \subseteq \mathrm{dom}(r)$ and for all $i \in \mathrm{dom}(f)$ we have $f(i) \in \mathrm{dom}(b)$ and $(b(f(i)), r(i)) \in E(I)$. We say that L *has real guards* if it has real guards w.r.t. g. Particularly simple examples of k-labeled incidence graphs with real guards are provided by the following definition.

▶ **Definition 7.1.** Let $f : \mathbb{N}_{\geqslant 1} \rightharpoonup [k]$ with finite $\mathrm{dom}(f) \neq \varnothing$. The k-labeled incidence graph M_f *defined by* f is the k-labeled incidence graph $L = (I, r, b, g)$ with $g := f$, where I consists of a red node v_i for every $i \in \mathrm{dom}(f)$, a blue node e_j for every $j \in \mathrm{img}(f)$, and an edge $(e_{f(i)}, v_i)$ for every $i \in \mathrm{dom}(f)$, and where $\mathrm{dom}(r) = \mathrm{dom}(f)$ and $r(i) = v_i$ for all $i \in \mathrm{dom}(r)$, and $\mathrm{dom}(b) = \mathrm{img}(f)$ and $b(j) = e_j$ for all $j \in \mathrm{dom}(b)$. Note that M_f has real guards.

We introduce a number of operations on k-labeled incidence graphs. The first kind of operations provides ways to modify the labels (the latter two of these do not necessarily preserve real guards). Let $L = (I, r, b, g)$ be a k-labeled incidence graph. Let $X_r \subseteq \mathbb{N}_{\geqslant 1}$ be finite, and let $X_b \subseteq [k]$.

1. Removing from the red nodes all the labels in X_r is achieved by the operation $L[X_r{\to}\bullet] := (I, r', b, g')$ with $r' := r - X_r$ and $g' := g - X_r$.
2. Removing from the blue nodes all the labels in X_b is achieved by the operation $L\langle X_b{\to}\bullet\rangle := (I, r, b', g)$ with $b' := b - X_b$
3. Let $X_r = \{i_1, \ldots, i_\ell\}$ for $\ell := |X_r|$ and $i_1 < \cdots < i_\ell$. For every $\overline{v} = (v_1, \ldots, v_\ell) \in R(I)^\ell$ we let $L[X_r{\to}\overline{v}] := (I, r', b, g)$ with $\mathrm{dom}(r') = \mathrm{dom}(r) \cup X_r$ and $r'(i_j) = v_j$ for all $j \in [\ell]$ and $r'(i) = r(i)$ for all $i \in \mathrm{dom}(r) \setminus X_r$ (i.e., for each $j \in [\ell]$, the red label i_j is moved onto the red node v_j, and all other labels remain unchanged).
4. Let $X_b = \{i_1, \ldots, i_\ell\}$ for $\ell := |X_b|$ and $i_1 < \cdots < i_\ell$. For every $\overline{e} = (e_1, \ldots, e_\ell) \in B(I)^\ell$ we let $L\langle X_b{\to}\overline{e}\rangle := (I, r, b', g)$ with $\mathrm{dom}(b') = \mathrm{dom}(b) \cup X_b$ and $b'(i_j) = e_j$ for all $j \in [\ell]$ and $b'(i) = b(i)$ for all $i \in \mathrm{dom}(b) \setminus X_b$ (i.e., for each $j \in [\ell]$, the blue label i_j is moved onto the blue node e_j, and all other labels remain unchanged).

The next operation enables us to *glue* two k-labeled incidence graphs L_1 and L_2. This is achieved by first taking the disjoint union of L_1 and L_2 and then merging all red (blue) nodes that carry the same label into a single red (blue) node that inherits all neighbours of the merged nodes. We write $(L_1 \cdot L_2)$ to denote the resulting k-labeled incidence graph. We need one further operation on k-labeled incidence graphs, namely, one that admits us to change its guard function:

▶ **Definition 7.2** (Applying a transition)**.** Consider a partial function $g : \mathbb{N}_{\geqslant 1} \rightharpoonup [k]$.

(a) A *transition for* g is a partial function $f : \mathbb{N}_{\geqslant 1} \rightharpoonup [k]$ with $\varnothing \neq \mathrm{dom}(f) \subseteq \mathrm{dom}(g)$ satisfying the following: for every $i \in \mathrm{dom}(g)$ with $g(i) \in \mathrm{img}(f)$ we have $i \in \mathrm{dom}(f)$.

(b) Let $L = (I, r, b, g)$ be a k-labeled incidence graph, and let f be a transition for g. *Applying* the transition f to L yields the k-labeled incidence graph $L[\rightsquigarrow f] := (\, M_f \,\cdot\, L\langle X_b{\to}\bullet\rangle\,)$, where $X_b := \mathrm{img}(g) \cap \mathrm{img}(f) \cap \mathrm{dom}(b)$, and M_f is provided by Definition 7.1.

The idea of applying a transition f to a k-labeled incidence graph $L = (I, r, b, g)$ is to assign new guards to a set of labeled red vertices (i.e. the domain of f). These new guards should by newly inserted nodes, and they should be *real* guards. To this end, for every $j \in \mathrm{img}(f)$ we add a new blue node e'_j labeled j; and in case that the label j had already been used by a blue node e of L (i.e., $j \in \mathrm{dom}(b)$) *and* served as a guard according to g (i.e., $j \in \mathrm{img}(g)$), we remove this label from e. For each $i \in \mathrm{dom}(f)$ with $f(i) = j$ we add an edge from the red node of L labeled i to the new blue node e'_j.

The formal definition $L[\rightsquigarrow f] := (M_f \cdot L\langle X_b{\to}\bullet\rangle)$ achieves this as follows: By $L\langle X_b{\to}\bullet\rangle$ we release from L all blue labels j that are present in L and that we want to assign to newly created nodes. This is achieved by letting $X_b = \mathrm{img}(g) \cap \mathrm{img}(f) \cap \mathrm{dom}(b)$. Afterwards, adding the edges from the nodes of L that carry a red label $i \in \mathrm{dom}(f)$ to the new blue node $e'_{f(i)}$ is achieved by glueing M_f to $L\langle X_b{\to}\bullet\rangle$. Note that releasing from L all blue labels in X_b might be problematic: Consider a red node v labeled i that was originally guarded by the blue node e of L that carried the label $j := g(i)$. Releasing the label j from node e means that v loses its guard in case that $i \notin \mathrm{dom}(f)$. Therefore, for f to be a transition for g, we require in Definition 7.2 that it assigns a new guard to all the affected labeled red vertices, i.e. we require $i \in \mathrm{dom}(f)$, if $g(i) \in \mathrm{img}(f)$. Finally, we are ready to define the class GLI_k:

▶ **Definition 7.3** (GLI_k)**.** The class GLI_k of *guarded k-labeled incidence graphs* is inductively defined as follows:

Base case: Any k-labeled incidence graph $L = (I, r, b, g)$ with $R(I) = \mathrm{img}(r)$, $B(I) = \mathrm{img}(b)$, and with real guards belongs to GLI_k.

Inductive cases: Let $L = (I, r, b, g) \in \mathit{GLI}_k$.

1. $L[X_r{\to}\bullet] \in \mathit{GLI}_k$ for every $X_r \subseteq \mathrm{dom}(r)$.
2. $L\langle X_b{\to}\bullet\rangle \in \mathit{GLI}_k$ for every $X_b \subseteq \mathrm{dom}(b) \setminus \mathrm{img}(g)$.
3. $L[\rightsquigarrow f] \in \mathit{GLI}_k$ for every transition f for g.
4. $(L \cdot L') \in \mathit{GLI}_k$ for every $L' = (I', r', b', g') \in \mathit{GLI}_k$ such that g and g' are compatible.

An easy inductive proof shows that every $L \in \mathit{GLI}_k$ has real guards.
A k-labeled incidence graph L is called *label-free* if $\mathrm{dom}(r_L) = \mathrm{dom}(b_L) = \mathrm{dom}(g_L) = \varnothing$. We are now ready for this section's technical main result, which states that, essentially, GLI_k provides an inductive characterisaton of IEHW_k:

▶ **Theorem 7.4.**
(a) *The incidence graph I_L of every $L \in \mathit{GLI}_k$ is in IEHW_k.*
(b) *For every $I \in \mathit{IEHW}_k$ there exists a label-free $L \in \mathit{GLI}_k$ such that $I \cong I_L$.*

The proof of (a) proceeds by induction on the definition of GLI_k and explicitly constructs an ehd of width $\leqslant k$ for each $L \in \mathit{GLI}_k$; for this it utilises that L has real guards. The proof of (b) starts with an ehd of I of width $\leqslant k$, chooses a suitable root node of the ehd's tree and performs a bottom-up traversal of this tree to associate each tree-node t with a corresponding k-labeled incidence graph L_t. This construction's details have to be carried out with care to ensure that $L_t \in \mathit{GLI}_k$.

Homomorphisms on k-Labeled Incidence Graphs and their Quantum Analogues. We define the notion of *homomorphisms* of k-labeled incidence graphs in such a way that it respects labels, but ignores the guard function: Let $L = (I, r, b, g)$ and $L' = (I', r', b', g')$ be k-labeled incidence graphs. If $\mathrm{dom}(r) \not\subseteq \mathrm{dom}(r')$ or $\mathrm{dom}(b) \not\subseteq \mathrm{dom}(b')$, then there exists no homomorphism from L to L'. Otherwise, a homomorphism from L to L' is a homomorphism $h = (h_R, h_B)$ from I to I' satisfying the following condition: $h(r(i)) = r'(i)$ for all $i \in \mathrm{dom}(r)$ and $h(b(j)) = b'(j)$ for all $j \in \mathrm{dom}(b)$. By $\mathrm{Hom}(L, L')$ we denote the set of all homomorphisms from L to L', and we let $\hom(L, L') := |\mathrm{Hom}(L, L')|$ be the number of homomorphisms from L to L'. In particular, if L is *label-free*, then $\hom(L, L') = \hom(I_L, I_{L'})$.

In order to enable us to "aggregate" homomorphism counts, we proceed in a similar way as Dvořák [11]: we use a variant of the *quantum graphs* of Lovász and Szegedy [20], tailored towards our setting. We say that k-labeled incidence graphs $L_1, \dots, L_d$ are *compatible* if their labeling functions all have the same domain and they all have the same guard function, i.e., $\mathrm{dom}(r_{L_1}) = \mathrm{dom}(r_{L_i})$, $\mathrm{dom}(b_{L_1}) = \mathrm{dom}(b_{L_i})$, and $g_{L_1} = g_{L_i}$ for all $i \in [d]$.

A *k-labeled quantum incidence graph* Q is a formal finite non-empty linear combination with real coefficients of compatible k-labeled incidence graphs. We represent a k-labeled quantum incidence graph Q as $\sum_{i=1}^{d} \alpha_i L_i$, where $d \in \mathbb{N}_{\geqslant 1}$, $\alpha_i \in \mathbb{R}$, and L_i is a k-labeled incidence graph for $i \in [d]$. We let $\mathrm{dr}_Q \coloneqq \mathrm{dom}(r_{L_1}) = \cdots = \mathrm{dom}(r_{L_d})$, $\mathrm{db}_Q \coloneqq \mathrm{dom}(b_{L_1}) = \cdots = \mathrm{dom}(b_{L_d})$, and $g_Q \coloneqq g_{L_1} = \cdots = g_{L_d}$. The α_i's and L_i's are called the *coefficients* and *components*, respectively, and d is called the *degree* of Q. Note that a k-labeled incidence graph is a k-labeled quantum incidence graph with degree 1 and coefficient 1. For a k-labeled quantum incidence graph $Q = \sum_{i=1}^{d} \alpha_i L_i$ and an arbitrary k-labeled incidence graph L' we let $\hom(Q, L') \coloneqq \sum_{i=1}^{d} \alpha_i \cdot \hom(L_i, L') \in \mathbb{R}$.

We adapt the operations for k-labeled incidence graphs to their quantum equivalent in the expected way: $Q[X_r{\to}\bullet] \coloneqq \sum_{i=1}^{d} L_i[X_r{\to}\bullet]$, $Q\langle X_b{\to}\bullet\rangle \coloneqq \sum_{i=1}^{d} L_i\langle X_b{\to}\bullet\rangle$, $Q[\rightsquigarrow f] \coloneqq \sum_{i=1}^{d} \alpha_i L_i[\rightsquigarrow f]$. Glueing two k-labeled quantum incidence graphs $Q = \sum_{i=1}^{d} \alpha_i L_i$ and $Q' = \sum_{j=1}^{d'} \alpha'_j L'_j$ is achieved by pairwise glueing of their components and multiplication of their respective coefficients, i.e. $(Q \cdot Q') \coloneqq \sum_{\substack{i\in[d]\\ j\in[d']}} (\alpha_i{\cdot}\alpha'_j)\, (L_i \cdot L'_j)$.

The following can easily be proved for the case where Q, Q' have degree 1 and coefficient 1 (i.e., Q, Q' are k-labeled incidence graphs), and then be generalised to quantum incidence graphs by simple linear arguments.

▶ **Lemma 7.5.** *For all k-labeled quantum incidence graphs Q, Q' and all k-labeled incidence graphs L we have:*

1. $\hom((Q \cdot Q'), L) = \hom(Q, L) \cdot \hom(Q', L)$.
2. $\hom(Q[X_r{\to}\bullet], L) = \sum_{\overline{v} \in R(I_L)^\ell} \hom(Q, L[X_r{\to}\overline{v}])$, *for all* $X_r \subseteq \mathrm{dr}_Q$ *and* $\ell := |X_r|$.
3. $\hom(Q\langle X_b{\to}\bullet\rangle, L) = \sum_{\overline{e} \in B(I_L)^\ell} \hom(Q, L\langle X_b{\to}\overline{e}\rangle)$, *for all* $X_b \subseteq \mathrm{db}_Q$ *and* $\ell := |X_b|$.
4. $\hom(Q[\rightsquigarrow f], L) = \hom(M_f, L) \cdot \sum_{\overline{e} \in B(I_L)^\ell} \hom(Q, L\langle X_b{\to}\overline{e}\rangle)$, *for all transitions f for g_Q, for* $X_b := \mathrm{db}_Q \cap \mathrm{img}(f) \cap \mathrm{img}(g)$ *and* $\ell := |X_b|$. *Note that* $\hom(M_f, L) \in \{0, 1\}$.

The class $QGLI_k$ of *guarded k-labeled quantum incidence graphs* consists of those k-labeled quantum incidence graphs where all components belong to GLI_k. The following lemma was provided for series-parallel quantum graphs by Lovász and Szegedy [20] and for labeled quantum graphs of tree-width $\leqslant k$ by Dvořák [11]; their proof also works for $QGLI_k$.

▶ **Lemma 7.6.** *Let $X, Y \subseteq \mathbb{N}$ be disjoint and finite, and let $Q \in QGLI_k$. There exists a $Q[X, Y] \in QGLI_k$ with the same parameters $\mathrm{dr}_Q, \mathrm{db}_Q, g_Q$ as Q, such that for all k-labeled incidence graphs L with real guards w.r.t. g_Q we have:*

1. *If* $\hom(Q, L) \in X$ *then* $\hom(Q[X, Y], L) = 0$.
2. *If* $\hom(Q, L) \in Y$ *then* $\hom(Q[X, Y], L) = 1$.

8 Proof of Theorem 6.1

Finally, we have available all the machinery so that, from a high-level point of view, our proof of Theorem 6.1 can follow a similar approach as Dvořák's proof in [11]. Analogously to the two main lemmas in [11], we provide a key lemma for each of the directions "$\Longleftarrow$" and "$\Longrightarrow$" of Theorem 6.1. These lemmas use the following notion: The *interpretation* $\mathcal{I}_{L'}$ associated with a k-labeled incidence graph L' is an interpretation (I, β) with $I := I_{L'}$ and $\beta(\mathtt{v}_i) := r_{L'}(i)$ for all $i \in \mathrm{dom}(r_{L'})$ and $\beta(\mathtt{e}_j) := b_{L'}(j)$ for all $j \in \mathrm{dom}(b_{L'})$.

▶ **Lemma 8.1.** *Let $L = (I, b, r, g) \in GLI_k$. For every $m \in \mathbb{N}$ there is a formula $\varphi_{L,m}$ with $(\Delta_g \wedge \varphi_{L,m}) \in \mathsf{RGC}^k$ and $\mathrm{free}((\Delta_g \wedge \varphi_{L,m})) = \{\mathtt{v}_i : i \in \mathrm{dom}(r)\} \cup \{\mathtt{e}_j : j \in \mathrm{dom}(b)\}$ such that for every k-labeled incidence graph L' with $\mathrm{dom}(b_{L'}) \supseteq \mathrm{dom}(b)$, $\mathrm{dom}(r_{L'}) \supseteq \mathrm{dom}(r)$, and with real guards w.r.t. g we have: $\mathcal{I}_{L'} \models \Delta_g$, and $\hom(L, L') = m \iff \mathcal{I}_{L'} \models \varphi_{L,m}$.*

▶ **Lemma 8.2.** *Let $\chi := (\Delta_g \wedge \psi) \in \mathsf{RGC}^k$ and let $m, d \in \mathbb{N}$ with $m \geqslant 1$. There exists a $Q := Q_{\chi,m,d} \in QGLI_k$ with $g_Q = g$, $\mathrm{db}_Q = \mathit{ifree}_B(\chi)$, $\mathrm{dr}_Q = \mathrm{dom}(g) = \mathit{ifree}_R(\chi)$ such that for all k-labeled incidence graphs $L' = (I', b', r', g')$ with $|B(I')| = m$ and $\max\{|N_{I'}(e)| : e \in B(I')\} \leqslant d$ and $\mathrm{dom}(b') \supseteq \mathrm{db}_Q$, $\mathrm{dom}(r') \supseteq \mathrm{dr}_Q$, $g' \supseteq g$, and with real guards w.r.t. g we have: $\mathcal{I}_{L'} \models \Delta_g$, and $\hom(Q, L') = 1$ if $\mathcal{I}_{L'} \models \chi$, and $\hom(Q, L') = 0$ if $\mathcal{I}_{L'} \not\models \chi$.*

The proofs of both lemmas are technically quite intricate because the concept of generalised hypertree width (as well as the classes IEHW_k and GLI_k) is much more complicated than the concept of tree-width. For Lemma 8.1 we proceed by induction based on Definition 7.3; for Lemma 8.2 we proceed by induction on the construction of χ. Finally, the proof of Theorem 6.1 can easily be achieved by using Theorem 7.4 and the Lemmas 8.1 (for direction "$\Longleftarrow$") and 8.2 (for direction "$\Longrightarrow$").

9 Conclusion

Combining the Theorems 3.1, 4.1, 6.1 yields:

▶ **Theorem 9.1** (Main Theorem). *Let H, H' be hypergraphs.*

(a) $I_H \equiv_{\mathsf{RGC}^k} I_{H'} \iff \mathsf{HOM}_{GHW_k}(H) = \mathsf{HOM}_{GHW_k}(H')$.

(b) *If H and H' are simple, then:* $I_H \equiv_{\mathsf{RGC}^k} I_{H'} \iff \mathsf{HOM}_{sGHW_k}(H) = \mathsf{HOM}_{sGHW_k}(H')$.

An obvious question is whether RGC^k-sentences have the same expressive power as GC^k-sentences. Since the submission of this paper, we were able to prove that this is indeed the case, i.e., any sentence of the logic GC^k can be transformed into an equivalent sentence in RGC^k. This implies that $I_H \equiv_{\mathsf{RGC}^k} I_{H'} \iff I_H \equiv_{\mathsf{GC}^k} I_{H'}$. Details are provided in the paper's extended version [22].

For our proofs it was crucial to consider ehds instead of generalised hypertree decompositions. To the best of our knowledge, ehds have not been studied before. From Theorem 2.3 we know that there exist arbitrarily large k such that $IEHW_k$ is a strict subclass of $IGHW_k$; but nevertheless, according to Theorem 4.1 homomorphism indistinguishability coincides for both classes. Many other questions remain open, in particular: How hard is it, given a hypergraph H and a number k, to determine whether $\text{ehw}(H) \leqslant k$? For $\mathcal{C} := IEHW_k$: how hard is it to compute the function (or, "vector") $\mathsf{HOM}_{\mathcal{C}}(H)$ for a given hypergraph H? Which properties does it have? What is the expressive power of the logic GC^k? How does a suitable pebble game for GC^k look like? Our result lifts Dvořák's result for tree-width $\leqslant k$ [11] from graphs to hypergraphs. Does there also exist a lifting of Grohe's result for *tree-depth* $\leqslant k$ [16] from graphs to hypergraphs? Seeing that Dvořák's result lifted nicely to hypergraphs, we believe that there should also be a lifting of Cai, Fürer and Immerman's result [8], i.e., a hypergraph-variant of the Weisfeiler-Leman algorithm, whose distinguishing power matches precisely the logic GC^k. We plan to study this in future work.

References

1. Isolde Adler. Marshals, monotone marshals, and hypertree-width. *J. Graph Theory*, 47(4):275–296, 2004. doi:10.1002/jgt.20025.
2. Isolde Adler. *Width Functions for Hypertree Decompositions*. PhD thesis, Albert-Ludwigs Universität Freiburg, 2006. Available at https://d-nb.info/979896851/34.
3. Isolde Adler, Georg Gottlob, and Martin Grohe. Hypertree width and related hypergraph invariants. *European Journal of Combinatorics*, 28(8):2167–2181, 2007. doi:10.1016/j.ejc.2007.04.013.
4. Hajnal Andréka, István Németi, and Johan van Benthem. Modal languages and bounded fragments of predicate logic. *J. Philos. Log.*, 27(3):217–274, 1998. doi:10.1023/A:1004275029985.
5. Jan Böker, Yijia Chen, Martin Grohe, and Gaurav Rattan. The complexity of homomorphism indistinguishability. In *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, volume 138 of *LIPIcs*, pages 54:1–54:13. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.MFCS.2019.54.
6. Jan Böker. Structural similarity and homomorphism counts. Master's thesis, RWTH Aachen University, 2018.
7. Jan Böker. Color Refinement, Homomorphisms, and Hypergraphs. In Ignas Sau and Dimitrios M. Thilikos, editors, *Graph-Theoretic Concepts in Computer Science*, volume 11789 of *Lecture Notes in Computer Science*, pages 338–350. Springer, Cham, 2019.
8. Jin-Yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, December 1992. doi:10.1007/BF01305232.

9 Bruno Courcelle. Graph Grammars, Monadic Second-Order Logic And The Theory Of Graph Minors. In Neil Robertson and Paul Seymour, editors, *Graph Structure Theory*, volume 147 of *Contemporary Mathematics*, pages 565–590. AMS, 1993.

10 Holger Dell, Martin Grohe, and Gaurav Rattan. Lovász meets Weisfeiler and Leman. In *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 40:1–40:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.ICALP.2018.40.

11 Zdeněk Dvořák. On Recognizing Graphs by Numbers of Homomorphisms. *Journal of Graph Theory*, 64(4):330–342, 2010.

12 Georg Gottlob, Gianluigi Greco, Nicola Leone, and Francesco Scarcello. Hypertree decompositions: Questions and answers. In Tova Milo and Wang-Chiew Tan, editors, *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2016, San Francisco, CA, USA, June 26 – July 01, 2016*, pages 57–74. ACM, 2016.

13 Georg Gottlob, Nicola Leone, and Francesco Scarcello. Hypertree Decompositions and Tractable Queries. *Journal of Computer and System Sciences*, 64(3):579–627, May 2002.

14 Georg Gottlob, Nicola Leone, and Francesco Scarcello. Robbers, marshals, and guards: game theoretic and logical characterizations of hypertree width. *Journal of Computer and System Sciences*, 66(4):775–808, June 2003.

15 Erich Grädel. On the restraining power of guards. *J. Symb. Log.*, 64(4):1719–1742, 1999. doi:10.2307/2586808.

16 Martin Grohe. Counting Bounded Tree Depth Homomorphisms. In *Proceedings of the 35th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS'20, pages 507–520. ACM, July 2020. doi:10.1145/3373718.3394739.

17 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'20, pages 1–16. ACM, 2020. doi:10.1145/3375395.3387641.

18 Phokion G. Kolaitis and Moshe Y. Vardi. Conjunctive-Query Containment and Constraint Satisfaction. *Journal of Computer and System Sciences*, 61(2):302–332, October 2000.

19 László Lovász. Operations with structures. *Acta Mathematica Academiae Scientiarum Hungaricae*, 18(3):321–328, 1967.

20 László Lovász and Balázs Szegedy. Contractors and Connectors of Graph Algebras. *Journal of Graph Theory*, 60(1):11–30, 2008.

21 Laura Mancinska and 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*, pages 661–672. IEEE, 2020. doi:10.1109/FOCS46700.2020.00067.

22 Benjamin Scheidt and Nicole Schweikardt. Counting homomorphisms from hypergraphs of bounded generalised hypertree width: A logical characterisation. *CoRR*, abs/2303.10980, 2023. doi:10.48550/arXiv.2303.10980.