Homomorphism-Distinguishing Closedness for Graphs of Bounded Tree-Width

Daniel Neuen
University of Bremen, Germany

Abstract

Two graphs are *homomorphism indistinguishable* over a graph class $\mathcal{F}$, denoted by $G \equiv_{\mathcal{F}} H$, if $\hom(F,G) = \hom(F,H)$ for all $F \in \mathcal{F}$ where $\hom(F,G)$ denotes the number of homomorphisms from F to G. A classical result of Lovász shows that isomorphism between graphs is equivalent to homomorphism indistinguishability over the class of all graphs. More recently, there has been a series of works giving natural algebraic and/or logical characterizations for homomorphism indistinguishability over certain restricted graph classes.

A class of graphs $\mathcal{F}$ is *homomorphism-distinguishing closed* if, for every $F \notin \mathcal{F}$, there are graphs G and H such that $G \equiv_{\mathcal{F}} H$ and $\hom(F,G) \neq \hom(F,H)$. Roberson conjectured that every class closed under taking minors and disjoint unions is homomorphism-distinguishing closed which implies that every such class defines a distinct equivalence relation between graphs. In this work, we confirm this conjecture for the classes $\mathcal{T}_k$, $k \geq 1$, containing all graphs of tree-width at most k.

As an application of this result, we also characterize which subgraph counts are detected by the k-dimensional Weisfeiler-Leman algorithm. This answers an open question from [Arvind et al., J. Comput. Syst. Sci., 2020].

2012 ACM Subject Classification Mathematics of computing → Graph theory; Mathematics of computing → Combinatorial algorithms; Theory of computation → Graph algorithms analysis

Keywords and phrases homomorphism indistinguishability, tree-width, Weisfeiler-Leman algorithm, subgraph counts

Digital Object Identifier 10.4230/LIPIcs.STACS.2024.53

Related Version *arXiv*: `https://arxiv.org/abs/2304.07011`

Acknowledgements I want to thank Tim Seppelt for pointing me to [24, Lemma 4].

1 Introduction

In 1967, Lovász [16] proved that two graphs G and H are isomorphic if and only if $\hom(F,G) = \hom(F,H)$ for every graph F where $\hom(F,G)$ denotes the number of homomorphisms from F to G. A natural follow-up question is to ask whether it is necessary to take the class of all graphs F to obtain the above result, and which kind of other equivalence relations can be obtained by restricting F to come from a proper subclass of all graphs. For a graph class $\mathcal{F}$, we say that two graphs G and H are $\mathcal{F}$*-equivalent*, denoted by $G \equiv_{\mathcal{F}} H$, if $\hom(F,G) = \hom(F,H)$ for all $F \in \mathcal{F}$. Hence, Lovász's [16] result says that $\equiv_{\mathcal{A}}$ is identical to the isomorphism relation where $\mathcal{A}$ denotes the class of all graphs.

In recent years, there has been a series of works giving natural algebraic and/or logical characterizations for homomorphism indistinguishability over certain restricted classes of graphs. For example, this includes graphs of bounded tree-width [8], graphs of bounded path-width [13], graphs of bounded tree-depth [12, 13] and the class of planar graphs [17]. In particular, those results imply that the equivalence relations $\equiv_{\mathcal{F}}$ obtained from the mentioned graph classes $\mathcal{F}$ do not correspond to isomorphism, and moreover, these equivalence relations are pairwise distinct.

In [21], Roberson initiated a more systematic study of the question which types of graph classes $\mathcal{F}$ lead to different equivalence relations $\equiv_{\mathcal{F}}$. A class of graphs $\mathcal{F}$ is called *homomorphism-distinguishing closed* if, for every $F \notin \mathcal{F}$, there are graphs G and H such that $G \equiv_{\mathcal{F}} H$ and $\hom(F, G) \neq \hom(F, H)$.

▶ **Conjecture 1** (Roberson [21]). *Let $\mathcal{F}$ be a graph class closed under taking disjoint unions and minors. Then $\mathcal{F}$ is homomorphism-distinguishing closed.*

In particular, this conjecture implies that every graph class closed under taking disjoint unions and minors defines a distinct equivalence relation $\equiv_{\mathcal{F}}$. Note that not every graph class is homomorphism-distinguishing closed. For example, the class $\mathcal{D}_2$ of 2-degenerate graphs (which is not closed under taking minors) is not homomorphism-distinguishing closed since the corresponding equivalence relation defines the isomorphism relation between graphs [8].

For $k \geq 1$ let $\mathcal{T}_k$ denote the class of all graphs of tree-width at most k. Roberson [21] showed that $\mathcal{T}_k$ is homomorphism-distinguishing closed for $k \in \{1, 2\}$. In this work, we generalize this to all $k \geq 1$.

▶ **Theorem 2.** *The class $\mathcal{T}_k$ is homomorphism-distinguishing closed for all $k \geq 1$.*

For the proof, we rely on known characterizations of homomorphism indistinguishability over the class $\mathcal{T}_k$ [6, 8, 13] and existing constructions of non-isomorphic pairs of graphs that are difficult to distinguish (see, e.g., [2, 5, 21]).

We remark that, since the first publication of the result, it has already been used in [22] to analyse the Lasserre semidefinite programming hierarchy for graph isomorphism via a characterization in terms of homomorphism counts. Also, the results have been used in [9] which in particular uses a similar strategy to prove that the class $\mathcal{TD}_q$ of all graphs of tree-depth at most q is homomorphism-distinguishing closed for all $q \geq 1$.

As an application of this result, we are able to characterize which subgraph counts are detected by the Weisfeiler-Leman algorithm (see also [1]). The Weisfeiler-Leman algorithm (WL) is a standard heuristic in the context of graph isomorphism testing (see, e.g., [2]) which recently also gained attention in a machine learning context [19, 20, 26, 28]. For $k \geq 1$, the k-dimensional Weisfeiler-Leman algorithm (k-WL) computes an isomorphism-invariant coloring of the k-tuples of vertices of a graph G. If the color patterns computed for two graphs G and H do not match, the graphs are non-isomorphic. In this case, we say that k-WL *distinguishes* G and H. It is known that two graphs G and H are distinguished by k-WL if and only if $G \not\equiv_{\mathcal{T}_k} H$, i.e., indistinguishability by k-WL can be characterized by homomorphism indistinguishability over the class of graphs of tree-width at most k [6, 8, 13].

In [10], Fürer initiated research on the question of which subgraph counts are detected by k-WL. Let F and G be two graphs. We write $\mathrm{sub}(F, G)$ to denote the number of subgraphs of G isomorphic to F. We say the function $\mathrm{sub}(F, \cdot)$ is *k-WL invariant* if $\mathrm{sub}(F, G) = \mathrm{sub}(F, H)$ for all graphs G, H that are indistinguishable by k-WL. For example, Fürer [10] shows that $\mathrm{sub}(C_\ell, \cdot)$ is 2-WL invariant for all $\ell \leq 6$ (where C_ℓ denotes the cycle on ℓ vertices), but $\mathrm{sub}(K_4, \cdot)$ is not 2-WL invariant. In [1], Arvind, Fuhlbrück, Köbler and Verbitsky further extended this line of research by showing $\mathrm{sub}(F, \cdot)$ is k-WL invariant for all graphs F that have hereditary tree-width at most k. For a graph F we define its *hereditary tree-width*, denoted by $\mathrm{hdtw}(F)$, to be the maximum tree-width of a homomorphic image of F. Arvind et al. [1] also provide some isolated negative results, but could not obtain a complete classification of which subgraph counts are detected by k-WL even for the special case $k = 2$.

Building on Theorem 2, we provide a complete classification of which subgraph counts are detected by k-WL for all $k \geq 1$. This answers an open question from [1].

▶ **Theorem 3.** *Let F be a graph and $k \geq 1$. Then* $\text{sub}(F, \cdot)$ *is k-WL invariant if and only if* $\text{hdtw}(F) \leq k$.

Observe that the backward direction is already proved in [1], i.e., the main contribution of this work is to show that for every graph F with $\text{hdtw}(F) > k$, the k-WL algorithm fails to detect subgraph counts from F.

For the proof, we use a well-known result [4] that allows us to formulate subgraph counts as a linear combination of certain homomorphism counts, and then combine Theorem 2 with an auxiliary lemma from [24].

We stress that Theorem 3 is also relevant in a machine-learning context. Indeed, it is known that the expressive power of graph neural networks (GNNs), which are a common tool for processing graph-structured data, is closely related to the expressive power of k-WL (see, e.g., [19, 20]). On the other hand, counting small subgraph patterns, also called network motifs [18], is a common technique in the study of large networks (see, e.g. [7, 14, 23, 27] for the use of network motifs in computational biology). Hence, it is natural to ask which subgraph counts can be detected by certain GNNs. This question has been studied in [3], but similar to [10, 1] only limited results have been obtained. Exploiting the connections between GNNs and k-WL (see, e.g., [19, 20]), Theorem 3 can provide a much more complete picture of which subgraph counts can be detected by GNNs. In fact, in a recent work, Lanzinger and Barceló [15] extend Theorem 3 to so-called knowledge graphs which are typically considered by GNNs.

We also remark that another extension of Theorem 3 has been obtained by Göbel, Goldberg and Roth [11] who determine the WL-dimension of counting the number of answers to an existential conjunctive query.

2 Preliminaries

A *graph* is a pair $G = (V, E)$ with vertex set $V = V(G)$ and edge relation $E = E(G)$. In this paper all graphs are finite, simple (no loops or multiple edges), and undirected. We denote edges by $vw \in E(G)$ where $v, w \in V(G)$. The *neighborhood* of $v \in V(G)$ is denoted by $N_G(v)$. Moreover, we write $E_G(v)$ to denote the set of edges incident to v. If the graph is clear from context, we usually omit the index G and simply write $N(v)$ and $E(v)$. For $A \subseteq V(G)$ we denote by $G[A]$ the *induced subgraph* of G on A. Also, we denote by $G \setminus A$ the induced subgraph on the complement of A, that is $G \setminus A \coloneqq G[V(G) \setminus A]$.

An *isomorphism* from a graph G to another graph H is a bijective mapping $\varphi\colon V(G) \to V(H)$ which preserves the edge relation, that is, $vw \in E(G)$ if and only if $\varphi(v)\varphi(w) \in E(H)$ for all $v, w \in V(G)$. Two graphs G and H are *isomorphic* ($G \cong H$) if there is an isomorphism from G to H. We write $\varphi\colon G \cong H$ to denote that φ is an isomorphism from G to H.

Let F and G be two graphs. A *homomorphism* from F to G is a mapping $\varphi\colon V(F) \to V(G)$ such that $\varphi(v)\varphi(w) \in E(G)$ for all $vw \in E(F)$. We write $\text{hom}(F, G)$ to denote the number of homomorphisms from F to G.

Let G be a graph. A graph H is a *minor* of G if H can be obtained from G by deleting vertices and edges of G as well as contracting edges of G. More formally, let $\mathcal{B} = \{B_1, \dots, B_h\}$ be a partition of $V(G)$ such that $G[B_i]$ is connected for all $i \in [h]$. We define $G/\mathcal{B}$ to be the graph with vertex set $V(G/\mathcal{B}) \coloneqq \mathcal{B}$ and

$$E(G/\mathcal{B}) \coloneqq \{BB' \mid \exists v \in B, v' \in B' \colon vv' \in E(G)\}.$$

A graph H is a minor of G if there is a partition $\mathcal{B} = \{B_1, \dots, B_h\}$ of connected subsets $B_i \subseteq V(G)$ such that H is isomorphic to a subgraph of $G/\mathcal{B}$. A graph G *excludes H as a minor* if H is not a minor of G.

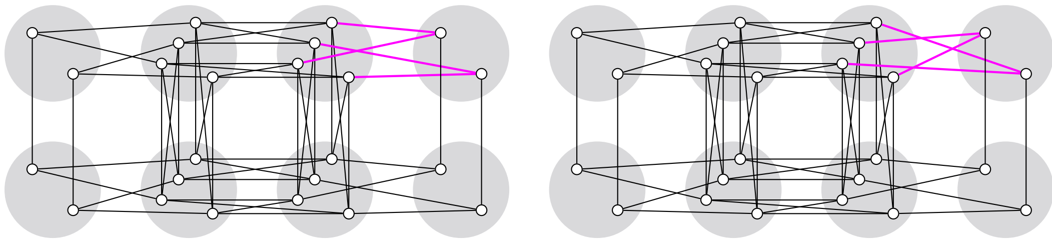

Figure 1 The figure shows the graphs $\mathrm{CFI}(G)$ and $\mathrm{CFI}^{\times}(G)$ where G is the 2×4 grid. The sets $M_{G,\emptyset}(v)$ and $M_{G,\{u_0\}}(v)$ are highlighted in gray. The vertex u_0 is located in the top-right corner of the grid. The marked edges show the difference between the two graphs.

3 Homomorphism Indistinguishability and Oddomorphisms

Toward the proof of Theorem 2, we need to cover several tools introduced in [21].

▶ **Definition 4** (Roberson [21]). *Let F and G be graphs and suppose φ is a homomorphism from F to G. We say a vertex $a \in V(F)$ is* odd (with respect to φ) *if $|N_F(a) \cap \varphi^{-1}(v)|$ is odd for every $v \in N_G(\varphi(a))$. Similarly, we say a vertex $a \in V(F)$ is* even with respect to φ *if $|N_F(a) \cap \varphi^{-1}(v)|$ is even for every $v \in N_G(\varphi(a))$.*

An oddomorphism *from F to G is a homomorphism φ from F to G such that*
(I) *every vertex $a \in V(F)$ is odd or even (with respect to φ), and*
(II) *$\varphi^{-1}(v)$ contains an odd number of odd vertices for every $v \in V(G)$.*

A weak oddomorphism *from F to G is a homomorphism φ from F to G such that there is a subgraph F' of F for which $\varphi|_{V(F')}$ is an oddomorphism from F' to G.*

Next, we introduce a construction for pairs of similar graphs from a base graph G that has also been used in [21]. Actually, variants of this construction have already been used in several earlier works (see, e.g., [2, 5]).

Let G be a graph and let $U \subseteq V(G)$. For $v \in V(G)$ we define $\delta_{v,U} := |\{v\} \cap U|$. We define the graph $\mathrm{CFI}(G, U)$ (the name refers to the authors of [2] where a very similar construction was first used in a related context) with vertex set

$$V(\mathrm{CFI}(G, U)) := \{(v, S) \mid v \in V(G), S \subseteq E(v), |S| \equiv \delta_{v,U} \bmod 2\}$$

and edge set

$$E(\mathrm{CFI}(G, U)) := \{(v, S)(u, T) \mid uv \in E(G), uv \notin S \triangle T\}$$

(here, $S \triangle T$ denotes the symmetric difference of S and T, i.e., $S \triangle T := (S \setminus T) \cup (T \setminus S)$). For $v \in V(G)$ we also write $M_{G,U}(v) := \{(v, S) \mid S \subseteq E(v), |S| \equiv \delta_{v,U} \bmod 2\}$ for the vertices in $\mathrm{CFI}(G, U)$ associated with v.

The following lemma is well-known (see, e.g., [2, 21])

▶ **Lemma 5.** *Let G be a connected graph and let $U, U' \subseteq V(G)$. Then $\mathrm{CFI}(G, U) \cong \mathrm{CFI}(G, U')$ if and only if $|U| \equiv |U'| \bmod 2$.*

We define $\mathrm{CFI}(G) := \mathrm{CFI}(G, \emptyset)$ and $\mathrm{CFI}^{\times}(G) := \mathrm{CFI}(G, \{u_0\})$ for some $u_0 \in V(G)$. A visualization can also be found in Figure 1.

▶ **Theorem 6** (Roberson [21, Theorem 3.13]). *Let F, G be graphs and suppose G is connected. Then $\hom(F, \mathrm{CFI}(G)) \geq \hom(F, \mathrm{CFI}^{\times}(G))$. Moreover, $\hom(F, \mathrm{CFI}(G)) > \hom(F, \mathrm{CFI}^{\times}(G))$ if and only if there exists a weak oddomorphism from F to G.*

We require two additional tools from [21] stated below.

▶ **Lemma 7** ([21, Lemma 5.6]). *Let F and G be graphs such that there is a weak oddomorphism from F to G. Also suppose G' is a minor of G. Then there is a minor F' of F such that there is an oddomorphism from F' to G'.*

▶ **Lemma 8** ([21, Theorem 6.2]). *Let $\mathcal{F}$ be a class of graphs such that*
(1) *if $F \in \mathcal{F}$ and there is a weak oddomorphism from F to G, then $G \in \mathcal{F}$, and*
(2) *$\mathcal{F}$ is closed under disjoint unions and restrictions to connected components.*
Then $\mathcal{F}$ is homomorphism-distinguishing closed.

4 Graphs of Bounded Tree-Width

In this section, we present the proof of Theorem 2. We rely on game characterizations for graphs of bounded tree-width as well as homomorphism indistinguishability over graphs of tree-width at most k.

4.1 Games

First, we cover the cops-and-robber game that characterizes tree-width of graphs. Fix some integer $k \geq 1$. For a graph G, we define the cops-and-robber game $\text{CopRob}_k(G)$ as follows:

- The game has two players called *Cops* and *Robber*.
- The game proceeds in rounds, each of which is associated with a pair of positions $(\bar{v}, u)$ with $\bar{v} \in \big(V(G)\big)^k$ and $u \in V(G)$.
- To determine the initial position, the Cops first choose a tuple $\bar{v} = (v_1, \dots, v_k) \in \big(V(G)\big)^k$ and then the Robber chooses some vertex $u \in V(G) \setminus \{v_1, \dots, v_k\}$ (if no such u exists, the Cops win the play). The initial position of the game is then set to $(\bar{v}, u)$.
- Each round consists of the following steps. Suppose the current position of the game is $(\bar{v}, u) = ((v_1, \dots, v_k), u)$.
  (C) The Cops choose some $i \in [k]$ and $v' \in V(G)$.
  (R) The Robber chooses a vertex $u' \in V(G)$ such that there exists a path from u to u' in $G \setminus \{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k\}$. After that, the game moves to position $\big((v_1, \dots, v_{i-1}, v', v_{i+1}, \dots, v_k), u'\big)$.
  If $u \in \{v_1, \dots, v_k\}$ the Cops win. If there is no position of the play such that the Cops win, then the Robber wins.

We say that the Cops (and the Robber, respectively) *win* $\text{CopRob}_k(G)$ if the Cops (and the Robber, respectively) have a winning strategy for the game. We also say that k *cops can catch a robber on* G if the Cops have a winning strategy in this game.

▶ **Theorem 9** ([25]). *A graph G has tree-width at most k if and only if $k + 1$ cops can catch a robber on G.*

Next, we discuss a game-theoretic characterization of two graphs being indistinguishable via homomorphism counts from graphs of tree-width at most k.

Let $k \geq 1$. For graphs G and H on the same number of vertices, we define the *bijective k-pebble game* $\text{BP}_k(G, H)$ as follows:

- The game has two players called *Spoiler* and *Duplicator*.
- The game proceeds in rounds, each of which is associated with a pair of positions $(\bar{v}, \bar{w})$ with $\bar{v} \in \big(V(G)\big)^k$ and $\bar{w} \in \big(V(H)\big)^k$.

- To determine the initial position, Duplicator plays a bijection $f\colon \big(V(G)\big)^k \to \big(V(H)\big)^k$ and Spoiler chooses some $\bar{v} \in \big(V(G)\big)^k$. The initial position of the game is then set to $(\bar{v}, f(\bar{v}))$.
- Each round consists of the following steps. Suppose the current position of the game is $(\bar{v}, \bar{w}) = ((v_1, \dots, v_k), (w_1, \dots, w_k))$.
  (S) Spoiler chooses some $i \in [k]$.
  (D) Duplicator picks a bijection $f\colon V(G) \to V(H)$.
  (S) Spoiler chooses $v \in V(G)$ and sets $w := f(v)$. Then the game moves to position $\big(\bar{v}[i/v], \bar{w}[i/w]\big)$ where $\bar{v}[i/v] := (v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k)$ is the tuple obtained from $\bar{v}$ by replacing the i-th entry by v.

  If mapping each v_i to w_i does not define an isomorphism of the induced subgraphs of G and H, Spoiler wins the play. More precisely, Spoiler wins if there are $i, j \in [k]$ such that $v_i = v_j \not\Leftrightarrow w_i = w_j$ or $v_iv_j \in E(G) \not\Leftrightarrow w_iw_j \in E(H)$. If there is no position of the play such that Spoiler wins, then Duplicator wins.

We say that Spoiler (and Duplicator, respectively) *wins* $\mathrm{BP}_k(G, H)$ if Spoiler (and Duplicator, respectively) has a winning strategy for the game. Also, for a position $(\bar{v}, \bar{w})$ with $\bar{v} \in \big(V(G)\big)^k$ and $\bar{w} \in \big(V(H)\big)^k$, we say that Spoiler (and Duplicator, respectively) *wins* $\mathrm{BP}_k(G, H)$ *from position* $(\bar{v}, \bar{w})$ if Spoiler (and Duplicator, respectively) has a winning strategy for the game started at position $(\bar{v}, \bar{w})$.

The following theorem follows from [2] and [6, 8, 13].

▶ **Theorem 10.** *Suppose $k \geq 1$. Let G and H be two graphs. Then $\hom(F, G) = \hom(F, H)$ for every $F \in \mathcal{T}_k$ if and only if Duplicator wins the game $\mathrm{BP}_{k+1}(G, H)$.*

4.2 Indistinguishable Graphs

The main step in the proof of Theorem 2 is to show that $\mathrm{CFI}(G)$ and $\mathrm{CFI}^{\times}(G)$ can not be distinguished via homomorphism counts from graphs of tree-width at most k for all connected graphs G of tree-width strictly greater than k. The proof follows similar arguments from [5] used to prove a closely related statement. Toward this end, the next lemma provides certain useful isomorphisms between CFI-graphs.

▶ **Lemma 11.** *Let G be a connected graph and suppose $u, v \in V(G)$. Let P be a path from u to v. Then there is an isomorphism $\varphi\colon \mathrm{CFI}(G, \{u\}) \cong \mathrm{CFI}(G, \{v\})$ such that*
(1) *$\varphi(M_{G,\{u\}}(w)) = M_{G,\{v\}}(w)$ for all $w \in V(G)$, and*
(2) *$\varphi(w, S) = (w, S)$ for all $w \in V(G) \setminus V(P)$ and $(w, S) \in M_{G,\{u\}}(w)$.*

Proof. Let $E(P)$ denote the set of edges on the path P. Clearly,
- $|E(P) \cap E(u)| = 1$ and $|E(P) \cap E(v)| = 1$,
- $|E(P) \cap E(w)| = 2$ for all $w \in V(P) \setminus \{u, v\}$, and
- $|E(P) \cap E(w)| = 0$ for all $w \in V(G) \setminus V(P)$.

We define $\varphi(w, S) := (w, S \triangle (E(P) \cap E(w)))$ for all $(w, S) \in \mathrm{CFI}(G, \{u\})$. It is easy to check that $\varphi\colon \mathrm{CFI}(G, \{u\}) \cong \mathrm{CFI}(G, \{v\})$ and the desired properties are satisfied. ◀

The next lemma forms the key technical step in the proof of Theorem 2.

▶ **Lemma 12.** *Let G be a connected graph of tree-width $\mathrm{tw}(G) \geq k$. Then Duplicator wins the k-bijective pebble game played on $\mathrm{CFI}(G)$ and $\mathrm{CFI}^{\times}(G)$.*

Proof. Let us fix some vertex $u_0 \in V(G)$ so that $\mathrm{CFI}^{\times}(G) = \mathrm{CFI}(G, \{u_0\})$. Since $\mathrm{tw}(G) \geq k$, the Robber has a winning strategy in the cops-and-robber game $\mathrm{CopRob}_k(G)$ by Theorem 9. We translate the winning strategy for the Robber in $\mathrm{CopRob}_k(G)$ into a winning strategy for Duplicator in the k-bijective pebble game played on $\mathrm{CFI}(G)$ and $\mathrm{CFI}^{\times}(G)$.

We first construct the bijection f for the initialization round. Suppose $\bar{x} = (x_1, \dots, x_k) \in (V(\mathrm{CFI}(G)))^k$. We define $A(\bar{x}) := (v_1, \dots, v_k)$ where $v_i \in V(G)$ is the unique vertex such that $x_i \in M_{G,\emptyset}(v_i)$.

Now let u be the vertex chosen by the Robber if the Cops initially place themselves on $A(\bar{x})$. Let P be a shortest path from u to u_0 (recall that G is connected), and let φ denote the isomorphism from $\mathrm{CFI}(G, \{u\})$ to $\mathrm{CFI}(G, \{u_0\})$ constructed in Lemma 11. We set $f(\bar{x}) := (\varphi(x_1), \dots, \varphi(x_k))$. It is easy to see that this gives a bijection f (we use the same isomorphism φ for all tuples $\bar{x}$ having the same associated tuple $A(\bar{x})$).

Now, throughout the game, Duplicator maintains the following invariant. Let $(\bar{x}, \bar{y})$ denote the current position. Then there is a vertex $u \in V(G)$ and an isomorphism $\varphi\colon \mathrm{CFI}(G, \{u\}) \cong \mathrm{CFI}(G, \{u_0\})$ such that

- $\varphi(M_{G,\{u\}}(w)) = M_{G,\{u_0\}}(w)$ for all $w \in V(G)$,
- $\varphi(\bar{x}) = \bar{y}$,
- u does not appear in the tuple $A(\bar{x})$, and
- the Robber wins from the position $(A(\bar{x}), u)$, i.e., if the Cops are placed on $A(\bar{x})$ and the Robber is on u.

Note that this condition is satisfied by construction after the initialization round.

Also observe that Duplicator never looses the game in such a position. Indeed, the mapping φ restricts to an isomorphism between $\mathrm{CFI}(G, \emptyset) - M_{G,\emptyset}(u) = \mathrm{CFI}(G, \{u\}) - M_{G,\{u\}}(u)$ and $\mathrm{CFI}(G, \{u_0\}) - M_{G,\{u_0\}}(u)$. Hence, since no vertex associated with u is pebbled in either graph, the pair $(\bar{x}, \bar{y})$ induces a local isomorphism.

So it remains to show that Duplicator can maintain the above invariant in each round of the k-bijective pebble game. Suppose $(\bar{x}, \bar{y})$ is the current position. Also let $(A(\bar{x}), u)$ be the associated position in the cops-and-robber game. Suppose that $A(\bar{x}) = (v_1, \dots, v_k)$.

Let $i \in [k]$ denote the index chosen by Spoiler. We describe the bijection f chosen by Duplicator. Let $v \in V(G)$. Let u' be the vertex the Robber moves to if the Cops choose i and v (i.e., the i-th cop changes its position to v) in the position $(A(\bar{x}), u)$. Let P denote a path from u to u' that avoids $\{v_1, \dots, v_k\} \setminus \{v_i\}$. Let ψ denote the isomorphism from $\mathrm{CFI}(G, \{u'\})$ to $\mathrm{CFI}(G, \{u\})$ constructed in Lemma 11. We set $f(x) := \varphi(\psi(x))$ for all $x \in M_{G,\emptyset}(v)$.

It is easy to see that f is a bijection. Let x denote the vertex chosen by Spoiler and let $y := f(x)$. Let $\bar{x}' := \bar{x}[i/x]$ and $\bar{y}' := \bar{y}[i/y]$, i.e., the pair $(\bar{x}', \bar{y}')$ is the new position of the game. Also, we set $\varphi' := \psi \circ \varphi$ (i.e., $\varphi'(z) = \varphi(\psi(z))$) where ψ denotes the isomorphism from $\mathrm{CFI}(G, \{u'\})$ to $\mathrm{CFI}(G, \{u\})$ used in the definition of $f(x)$.

Clearly, $\varphi'(M_{G,\{u'\}}(w)) = M_{G,\{u_0\}}(w)$ for all $w \in V(G)$, since the corresponding conditions are satisfied for the mappings ψ and φ. We have $\varphi'(x) = y$ by definition. All the other entries of $\bar{x}'$ are fixed by the mapping ψ (see Lemma 11, Part (2)) which overall implies that $\varphi'(\bar{x}') = \bar{y}'$. Also, u' does not appear in the tuple $A(\bar{x}')$ by construction, and the Robber wins from the position $(A(\bar{x}'), u')$.

So overall, this means that Duplicator can maintain the above invariant which provides the desired winning strategy. ◀

With this, we are almost ready to prove Theorem 2. The next corollary states the key consequence of Lemma 12 that allows us to apply Lemma 8.

▶ **Corollary 13.** *Let $k \geq 1$ and let F be a graph of tree-width* $\mathrm{tw}(F) \leq k$*. Also let G be a graph and suppose there is a weak oddomorphism from F to G. Then* $\mathrm{tw}(G) \leq k$*.*

Proof. Suppose towards a contradiction that $\mathrm{tw}(G) > k$. Then there is a connected subgraph G' of G such that $\mathrm{tw}(G') > k$. By Lemma 7, we conclude that there is a minor F' of F such that there is an oddomorphism from F' to G'. In particular, $\mathrm{tw}(F') \leq \mathrm{tw}(F) \leq k$. By Theorem 6, we conclude that $\hom(F', \mathrm{CFI}(G')) > \hom(F', \mathrm{CFI}^{\times}(G'))$. Using Theorem 10 it follows that Spoiler wins the $(k+1)$-bijective pebble game $\mathrm{BP}_{k+1}(\mathrm{CFI}(G'), \mathrm{CFI}^{\times}(G'))$. But this contradicts Lemma 12 since $\mathrm{tw}(G') \geq k+1$. ◀

Proof of Theorem 2. Let $k \geq 1$ be fixed. By Corollary 13, the class $\mathcal{T}_k$ satisfies Condition 1 from Lemma 8. Also, the class $\mathcal{T}_k$ clearly satisfies Condition 2 from Lemma 8. So $\mathcal{T}_k$ is homomorphism-distinguishing closed by Lemma 8. ◀

5 Weisfeiler-Leman and Subgraph Counts

In this section, we prove Theorem 3. Towards this end, we first need to formally introduce the WL algorithm.

5.1 The Weisfeiler-Leman Algorithm

Let $\chi_1, \chi_2 \colon V^k \to C$ be colorings of the k-tuples of vertices, where C is some finite set of colors. We say χ_1 *refines* χ_2, denoted $\chi_1 \preceq \chi_2$, if $\chi_1(\bar{v}) = \chi_1(\bar{w})$ implies $\chi_2(\bar{v}) = \chi_2(\bar{w})$ for all $\bar{v}, \bar{w} \in V^k$. The colorings χ_1 and χ_2 are *equivalent*, denoted $\chi_1 \equiv \chi_2$, if $\chi_1 \preceq \chi_2$ and $\chi_2 \preceq \chi_1$.

We describe the *k-dimensional Weisfeiler-Leman algorithm* (k-WL) for all $k \geq 1$. For an input graph G let $\chi^{k,G}_{(0)} \colon (V(G))^k \to C$ be the coloring where each tuple is colored with the isomorphism type of its underlying ordered subgraph. More precisely, $\chi^{k,G}_{(0)}(v_1, \dots, v_k) = \chi^{k,G}_{(0)}(v'_1, \dots, v'_k)$ if and only if, for all $i, j \in [k]$, it holds that $v_i = v_j \Leftrightarrow v'_i = v'_j$ and $v_i v_j \in E(G) \Leftrightarrow v'_i v'_j \in E(G)$.

We then recursively define the coloring $\chi^{k,G}_{(i+1)}$ obtained after $i+1$ rounds of the algorithm (for $i \geq 0$). For $k \geq 2$ and $\bar{v} = (v_1, \dots, v_k) \in (V(G))^k$ we define

$$\chi^{k,G}_{(i+1)}(\bar{v}) \coloneqq \left(\chi^{k,G}_{(i)}(\bar{v}), \mathcal{M}_i(\bar{v})\right)$$

where

$$\mathcal{M}_i(\bar{v}) \coloneqq \left\{\!\!\left\{ \left(\chi^{k,G}_{(i)}(\bar{v}[1/w]), \dots, \chi^{k,G}_{(i)}(\bar{v}[k/w])\right) \;\middle|\; w \in V(G) \right\}\!\!\right\}$$

and $\bar{v}[i/w] \coloneqq (v_1, \dots, v_{i-1}, w, v_{i+1}, \dots, v_k)$ is the tuple obtained from $\bar{v}$ by replacing the i-th entry by w. For $k = 1$, the definition is similar, but we only iterate over neighbors of v_1, i.e.,

$$\mathcal{M}_i(v_1) \coloneqq \left\{\!\!\left\{ \chi^{k,G}_{(i)}(w) \;\middle|\; w \in N_G(v_1) \right\}\!\!\right\}.$$

There is a minimal $i_\infty \geq 0$ such that $\chi^{k,G}_{(i_\infty)} \equiv \chi^{k,G}_{(i_\infty+1)}$ and for this i_∞ we define $\chi^{k,G} \coloneqq \chi^{k,G}_{(i_\infty)}$.

Let G and H be two graphs. We say that k-WL *distinguishes* G and H if there exists a color c such that

$$\left|\left\{\bar{v} \in \left(V(G)\right)^k \;\middle|\; \chi^{k,G}(\bar{v}) = c\right\}\right| \neq \left|\left\{\bar{w} \in \left(V(H)\right)^k \;\middle|\; \chi^{k,H}(\bar{w}) = c\right\}\right|.$$

We write $G \simeq_k H$ if k-WL does not distinguish G and H.

Recall that $\mathcal{T}_k$ denotes the class of graphs of tree-width at most k. The following characterization follows from [6, 8, 13] (see also Theorem 10).

▶ **Theorem 14.** *Suppose $k \geq 1$. Let G and H be two graphs. Then $G \simeq_k H$ if and only if $G \equiv_{\mathcal{T}_k} H$.*

Recall that we write $\text{sub}(F, G)$ to denote the number of subgraphs of G isomorphic to F. We write $\text{sub}(F, \cdot)$ to denote the function that maps each graph G to the corresponding subgraph count $\text{sub}(F, G)$.

▶ **Definition 15.** *Let F be a graph. The function* $\text{sub}(F, \cdot)$ *is* k-WL invariant *if*

$$\text{sub}(F, G) = \text{sub}(F, H) \tag{1}$$

for all graphs G, H such that $G \simeq_k H$.

5.2 Subgraph Counts

Using the framework from [4], it is possible to describe the subgraph count $\text{sub}(F, G)$ as a linear combination

$$\text{sub}(F, G) = \sum_{i \in [\ell]} \alpha_i \cdot \text{hom}(F_i, G)$$

for certain graphs $F_1, \dots, F_\ell$ and coefficients $\alpha_1, \dots, \alpha_\ell \in \mathbb{R}$ that only depend on F. More precisely, the graphs $F_1, \dots, F_\ell$ are exactly the homomorphic images of F.

▶ **Definition 16.** *Let F and H be two graphs. We say that H is a* homomorphic image *of F if there is a surjective homomorphism $\varphi \colon V(F) \to V(H)$ such that*

$$E(H) = \{\varphi(v)\varphi(w) \mid vw \in E(F)\}.$$

We write $\text{spasm}(F)$ *to denote the set of homomorphic images of F. The* hereditary tree-width *of F, denoted by* $\text{hdtw}(F)$*, is the maximum tree-width of a graph in* $\text{spasm}(F)$*, i.e.,*

$$\text{hdtw}(F) := \max_{H \in \text{spasm}(F)} \text{tw}(H).$$

In the following, we assume that $\text{spasm}(F)$ contains only one representative from each isomorphism class, i.e., for every homomorphic image H of F there is exactly one graph $H' \in \text{spasm}(F)$ that is isomorphic to H. In particular, the set $\text{spasm}(F)$ is finite.

The backward direction of Theorem 3 has already been proved in [1].

▶ **Lemma 17** ([1, Corollary 4.3])**.** *Let F be a graph such that* $\text{hdtw}(F) \leq k$*. Then* $\text{sub}(F, \cdot)$ *is k-WL invariant.*

For the sake of completeness, we still include the simple proof.

Proof. Let F be a graph such that $\text{hdtw}(F) \leq k$ and let $\mathcal{L} := \text{spasm}(F)$. By [4] there is a unique function $\alpha \colon \mathcal{L} \to \mathbb{R} \setminus \{0\}$ such that

$$\text{sub}(F, G) = \sum_{L \in \mathcal{L}} \alpha(L) \cdot \text{hom}(L, G)$$

for all graphs G.

Now let G, H be two graphs such that $G \simeq_k H$. Then $\text{hom}(L, G) = \text{hom}(L, H)$ for all graph $L \in \mathcal{T}_k$ by Theorem 14. Since $\text{hdtw}(F) \leq k$, we get that $\mathcal{L} \subseteq \mathcal{T}_k$. So, in particular, $\text{hom}(L, G) = \text{hom}(L, H)$ for all graph $L \in \mathcal{L}$. It follows that $\text{sub}(F, G) = \text{sub}(F, H)$. ◀

For the other direction, we combine Theorem 2 and the following lemma from [24].

▶ **Lemma 18** ([24, Lemma 4]). *Let $\mathcal{F}$ be a class of graphs that is homomorphism-distinguishing closed. Let $\mathcal{L}$ be a finite set of pairwise non-isomorphic graphs and $\alpha\colon \mathcal{L} \to \mathbb{R} \setminus \{0\}$. Also suppose that for all graphs G, H it holds that*

$$G \equiv_{\mathcal{F}} H \quad \implies \quad \sum_{L \in \mathcal{L}} \alpha(L) \cdot \hom(L, G) = \sum_{L \in \mathcal{L}} \alpha(L) \cdot \hom(L, H). \tag{2}$$

Then $\mathcal{L} \subseteq \mathcal{F}$.

▶ **Lemma 19.** *Let F be a graph such that $\mathrm{sub}(F, \cdot)$ is k-WL invariant. Then $\mathrm{hdtw}(F) \leq k$.*

Proof. Let $\mathcal{F}$ denote the class of graphs of tree-width at most k. By Theorem 2 the class $\mathcal{F}$ is homomorphism-distinguishing closed. Let $\mathcal{L} \coloneqq \mathrm{spasm}(F)$. By [4] there is a unique function $\alpha\colon \mathcal{L} \to \mathbb{R} \setminus \{0\}$ such that

$$\mathrm{sub}(F, G) = \sum_{L \in \mathcal{L}} \alpha(L) \cdot \hom(L, G)$$

for all graphs G. Since $\mathrm{sub}(F, \cdot)$ is k-WL invariant it follows that Equation (2) is satisfied for all graphs G, H using Theorem 14. So $\mathrm{spasm}(F) = \mathcal{L} \subseteq \mathcal{F}$ by Lemma 18. This implies that $\mathrm{hdtw}(F) \leq k$. ◀

Proof of Theorem 3. The theorem follows directly from Lemmas 17 and 19. ◀

6 Conclusion

We proved that for every $k \geq 1$ the class $\mathcal{T}_k$ of all graphs of tree-width at most k is homomorphism-distinguishing closed. As a consequence, we could answer an open question from [1] and precisely classify the subgraph counts detected by k-WL.

Still, Conjecture 1 remains wide open. As an intermediate step, it may be interesting to consider minor- and union-closed classes of bounded tree-width. More precisely, let $\mathcal{F}$ be a graph class closed under taking disjoint unions and minors, and there is some $k \geq 1$ such that every $F \in \mathcal{F}$ has tree-width at most k. Can we show that $\mathcal{F}$ is homomorphism-distinguishing closed? Towards this end, it may also be interesting to obtain a direct proof of Corollary 13 that does not rely on the characterization from Theorem 10.

References

1 Vikraman Arvind, Frank Fuhlbrück, Johannes Köbler, and Oleg Verbitsky. On Weisfeiler-Leman invariance: Subgraph counts and related graph properties. *J. Comput. Syst. Sci.*, 113:42–59, 2020. doi:10.1016/j.jcss.2020.04.003.

2 Jin-yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Comb.*, 12(4):389–410, 1992. doi:10.1007/BF01305232.

3 Zhengdao Chen, Lei Chen, Soledad Villar, and Joan Bruna. Can graph neural networks count substructures? In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors, *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. URL: https://proceedings.neurips.cc/paper/2020/hash/75877cb75154206c4e65e76b88a12712-Abstract.html.

4 Radu Curticapean, Holger Dell, and Dániel Marx. Homomorphisms are a good basis for counting small subgraphs. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 210–223. ACM, 2017. doi:10.1145/3055399.3055502.

5 Anuj Dawar and David Richerby. The power of counting logics on restricted classes of finite structures. In Jacques Duparc and Thomas A. Henzinger, editors, *Computer Science Logic, 21st International Workshop, CSL 2007, 16th Annual Conference of the EACSL, Lausanne, Switzerland, September 11-15, 2007, Proceedings*, volume 4646 of *Lecture Notes in Computer Science*, pages 84–98. Springer, 2007. doi:10.1007/978-3-540-74915-8_10.

6 Holger Dell, Martin Grohe, and Gaurav Rattan. Lovász meets Weisfeiler and Leman. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *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.

7 Banu Dost, Tomer Shlomi, Nitin Gupta, Eytan Ruppin, Vineet Bafna, and Roded Sharan. QNet: a tool for querying protein interaction networks. *J. Comput. Biol.*, 15(7):913–925, 2008. doi:10.1089/cmb.2007.0172.

8 Zdeněk Dvořák. On recognizing graphs by numbers of homomorphisms. *J. Graph Theory*, 64(4):330–342, 2010. doi:10.1002/jgt.20461.

9 Eva Fluck, Tim Seppelt, and Gian Luca Spitzer. Going deep and going wide: Counting logic and homomorphism indistinguishability over graphs of bounded treedepth and treewidth. *CoRR*, abs/2308.06044, 2023. doi:10.48550/arXiv.2308.06044.

10 Martin Fürer. On the combinatorial power of the Weisfeiler-Lehman algorithm. In Dimitris Fotakis, Aris Pagourtzis, and Vangelis Th. Paschos, editors, *Algorithms and Complexity - 10th International Conference, CIAC 2017, Athens, Greece, May 24-26, 2017, Proceedings*, volume 10236 of *Lecture Notes in Computer Science*, pages 260–271, 2017. doi:10.1007/978-3-319-57586-5_22.

11 Andreas Göbel, Leslie Ann Goldberg, and Marc Roth. The Weisfeiler-Leman dimension of existential conjunctive queries. *CoRR*, abs/2310.19006, 2023. doi:10.48550/arXiv.2310.19006.

12 Martin Grohe. Counting bounded tree depth homomorphisms. In Holger Hermanns, Lijun Zhang, Naoki Kobayashi, and Dale Miller, editors, *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020*, pages 507–520. ACM, 2020. doi:10.1145/3373718.3394739.

13 Martin Grohe, Gaurav Rattan, and Tim Seppelt. Homomorphism tensors and linear equations. In Mikolaj Bojanczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France*, volume 229 of *LIPIcs*, pages 70:1–70:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.ICALP.2022.70.

14 Falk Hüffner, Sebastian Wernicke, and Thomas Zichner. Algorithm engineering for color-coding with applications to signaling pathway detection. *Algorithmica*, 52(2):114–132, 2008. doi:10.1007/s00453-007-9008-7.

15 Matthias Lanzinger and Pablo Barceló. On the power of the Weisfeiler-Leman test for graph motif parameters. *CoRR*, abs/2309.17053, 2023. doi:10.48550/arXiv.2309.17053.

16 László Lovász. Operations with structures. *Acta Math. Acad. Sci. Hungar.*, 18:321–328, 1967. doi:10.1007/BF02280291.

17 Laura Mancinska and David E. Roberson. Quantum isomorphism is equivalent to equality of homomorphism counts from planar graphs. In Sandy Irani, editor, *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.

18 Ron Milo, Shai S. Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri B. Chklovskii, and Uri Alon. Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002. doi:10.1126/science.298.5594.824.

19 Christopher Morris, Yaron Lipman, Haggai Maron, Bastian Rieck, Nils M. Kriege, Martin Grohe, Matthias Fey, and Karsten M. Borgwardt. Weisfeiler and Leman go machine learning: The story so far. *CoRR*, abs/2112.09992, 2021. arXiv:2112.09992.

20 Christopher Morris, Martin Ritzert, Matthias Fey, William L. Hamilton, Jan Eric Lenssen, Gaurav Rattan, and Martin Grohe. Weisfeiler and Leman go neural: Higher-order graph neural networks. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, pages 4602–4609. AAAI Press, 2019. doi:10.1609/aaai.v33i01.33014602.

21 David E. Roberson. Oddomorphisms and homomorphism indistinguishability over graphs of bounded degree. *CoRR*, abs/2206.10321, 2022. doi:10.48550/arXiv.2206.10321.

22 David E. Roberson and Tim Seppelt. Lasserre hierarchy for graph isomorphism and homomorphism indistinguishability. In Kousha Etessami, Uriel Feige, and Gabriele Puppis, editors, *50th International Colloquium on Automata, Languages, and Programming, ICALP 2023, July 10-14, 2023, Paderborn, Germany*, volume 261 of *LIPIcs*, pages 101:1–101:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.ICALP.2023.101.

23 Jacob Scott, Trey Ideker, Richard M. Karp, and Roded Sharan. Efficient algorithms for detecting signaling pathways in protein interaction networks. *J. Comput. Biol.*, 13(2):133–144, 2006. doi:10.1089/cmb.2006.13.133.

24 Tim Seppelt. Logical equivalences, homomorphism indistinguishability, and forbidden minors. In Jérôme Leroux, Sylvain Lombardy, and David Peleg, editors, *48th International Symposium on Mathematical Foundations of Computer Science, MFCS 2023, August 28 to September 1, 2023, Bordeaux, France*, volume 272 of *LIPIcs*, pages 82:1–82:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.MFCS.2023.82.

25 Paul D. Seymour and Robin Thomas. Graph searching and a min-max theorem for tree-width. *J. Comb. Theory, Ser. B*, 58(1):22–33, 1993. doi:10.1006/jctb.1993.1027.

26 Nino Shervashidze, Pascal Schweitzer, Erik Jan van Leeuwen, Kurt Mehlhorn, and Karsten M. Borgwardt. Weisfeiler-Lehman graph kernels. *J. Mach. Learn. Res.*, 12:2539–2561, 2011. doi:10.5555/1953048.2078187.

27 Tomer Shlomi, Daniel Segal, Eytan Ruppin, and Roded Sharan. Qpath: a method for querying pathways in a protein-protein interaction network. *BMC Bioinformatics*, 7:199, 2006. doi:10.1186/1471-2105-7-199.

28 Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. URL: https://openreview.net/forum?id=ryGs6iA5Km.