On the Expressive Power
of Homomorphism Counts

Albert Atserias[*]   Phokion G. Kolaitis[†]   Wei-Lin Wu[‡]

February 7, 2021

Abstract

A classical result by Lovász asserts that two graphs G and H are isomorphic if and only if they have the same left-homomorphism vector, that is, for every graph F, the number of homomorphisms from F to G coincides with the number of homomorphisms from F to H. Dell, Grohe, and Rattan showed that restrictions of the left-homomorphism vector to a class of graphs can capture several different relaxations of isomorphism, including co-spectrality (i.e., two graphs having the same characteristic polynomial), fractional isomorphism and, more broadly, equivalence in counting logics with a fixed number of variables. On the other side, a result by Chaudhuri and Vardi asserts that isomorphism is also captured by the right-homomorphism vector, that is, two graphs G and H are isomorphic if and only if for every graph F, the number of homomorphisms from G to F coincides with the number of homomorphisms from H to F. In this paper, we embark on a study of the restrictions of the right-homomorphism vector by investigating relaxations of isomorphism that can or cannot be captured by restricting the right-homomorphism vector to a fixed class of graphs. Our results unveil striking differences between the expressive power of the left-homomorphism vector and the right-homomorphism vector. We show that co-spectrality, fractional isomorphism, and equivalence in counting logics with a fixed number of variables cannot be captured by restricting the right-homomorphism vector to a class of graphs. In the opposite direction, we show that chromatic equivalence cannot be captured by restricting the left-homomorphism vector to a class of graphs, while, clearly, it can be captured by restricting the right-homomorphism vector to the class of all cliques.

1

1   Introduction

Even though research on the graph isomorphism problem has spanned several decades [RC77] and in spite of some recent progress [Bab16], the exact complexity of the graph isomorphism problem remains unknown. This state of affairs has motivated the study of *relaxations* of graph isomorphism, that is, equivalence relations that are coarser than graph isomorphism. Some well known such relaxations are based on indistinguishability of two graphs via a heuristic for graph isomorphism, others are based on indistinguishability of two graphs in some logical formalism, and others are based on indistinguishability of two graphs via some graph polynomial.

The k-dimensional Weisfeiler-Lehman method, $k \geq 1$, is a prominent example of a heuristic for graph isomorphism that distinguishes some, but not all, non-isomorphic graphs. This method is an iterative algorithm that assigns colors to k-tuples of vertices of a graph until a stable coloring is achieved [Wei06]. The 1-dimensional Weisfeiler-Lehman method is also known as the vertex refinement algorithm and coincides with the fractional isomorphism test, which asks for the existence of a rational solution to a 0-1 integer linear program encoding the existence of an isomorphism between two graphs. On the side of logic, the counting logics C^k, $k \geq 1$, have found applications to the study of graph isomorphism; each logic C^k augments the syntax of first-order logic with counting quantifiers $\exists i x$ asserting that there are at least i distinct elements x, but the C^k-formulas are required to have at most k distinct variables. Cai, Fürer, and Immerman [CFI92] have shown a tight connection between the Weisfeiler-Lehman method and the counting logics, namely, for every $k \geq 2$, two graphs are indistinguishable by the $(k-1)$-Weisfeiler-Lehman method if and only if they satisfy the same C^k-sentences.

Graph polynomials are polynomials that encapsulate one or more important invariants of the graph they are associated with. Different graph polynomials give rise to different relaxations of graph isomorphism. For example, the chromatic polynomial $\chi(G, k)$ of a graph G returns the number of all k-colorings of G (see [Rea68]); it also gives rise to the *chromatic equivalence* relation, which holds between two graph G and H precisely when $\chi(G, k) = \chi(H, k)$.

Lovász [Lov67] showed that graph isomorphism can be characterized in terms of left-homomorphism counts. If G and H are two graphs, then $\hom(G, H)$ denotes the number of homomorphisms from G to H. Let $\mathscr{G}$ be the class of all graphs (graphs are assumed to be finite, undirected and to have no loops and no multiples edges) and let G be a graph. The *left-homomorphism vector* of G is the infinite vector $\hom(\mathscr{G}, G) = (\hom(F, G) \mid F \in \mathscr{G})$. Lovász's result asserts that two graphs G and H are isomorphic if and only $\hom(\mathscr{G}, G) = \hom(\mathscr{G}, H)$. Dell, Grohe, and Rattan [DGR18] investigated relaxations of graph isomorphism obtained by restricting the left-homomorphism vector to a fixed class $\mathscr{F}$ of graphs, i.e., they considered vectors of the form $\hom(\mathscr{F}, G) = (\hom(F, G) \mid F \in \mathscr{F})$ and the associated equivalence relation between graphs G and H defined by the condition $\hom(\mathscr{F}, G) = \hom(\mathscr{F}, H)$. They showed that the left-homomorphism vector restricted to the class $\mathscr{T}$ of all trees captures fractional isomorphism, i.e., two graphs G and H are fractionally isomorphic if and only if $\hom(\mathscr{T}, G) = \hom(\mathscr{T}, H)$. They also showed that the left-homomorphism vector restricted to the class $\mathscr{T}_{k-1}$ of all graphs of treewidth at most k captures indistinguishabil-

2

ity in the counting logic C^k and, hence, indistinguishability using the $(k-1)$-dimensional Weisfeiler-Lehman method, $k \geq 3$. The study of restrictions of the left-homomorphism vector was further pursued by Böker et al. [BCGR19] and by Grohe et al. [Gro20].

On the other side and in the context of database theory, Chaudhuri and Vardi [CV93] showed that graph isomorphism can also be characterized in terms of right-homomorphism counts. The *right-homomorphism vector* of a graph G is the infinite vector $\hom(G, \mathscr{G}) = (\hom(G, F) \mid F \in \mathscr{G})$. Chaudhuri and Vardi showed that two graphs G and H are isomorphic if and only if and only if $\hom(G, \mathscr{G}) = \hom(H, \mathscr{G})$.

In this paper, we embark on a study of relaxations of graph isomorphism arising by restricting the right-homomorphism vector to a fixed class $\mathscr{F}$ of graphs (vectors of the form $\hom(G, \mathscr{F}) = (\hom(G, F) \mid F \in \mathscr{F})$), that is, we study equivalence relations on two graphs G and H defined by the condition $\hom(G, \mathscr{F}) = \hom(H, \mathscr{F})$. Our main aim is to compare the expressive power of restrictions of the right-homomorphism vector vs. the expressive power of restrictions of the left-homomorphism vector. While Dell, Grohe, and Rattan [DGR18] identified relaxations of graph isomorphism that *can* be captured by restrictions of the left-homomorphism vector, here we identify relaxations of graph isomorphism that *cannot* be captured by *any* restriction of the right-homomorphism vector or by *any* restriction of the left-homomorphism vector. It should be noted that Garijo, Goodall, and Nešetřil [GGN11] considered restrictions of both the left-homomorphism vector and the right-homomorphism vector, but used these restrictions to study a different problem, namely, which graphs are uniquely determined by certain graph polynomials, such as the aforementioned chromatic polynomial.

We now present an overview of our technical results. First, we show from first principles that there is no class $\mathscr{F}$ of graphs such that the right-homomorphism vector restricted to $\mathscr{F}$ captures fractional isomorphism. After this, we use sophisticated machinery to show that for every $k \geq 3$, there is no class $\mathscr{F}$ of graphs such that the right-homomorphism vector restricted to $\mathscr{F}$ captures indistinguishability of graphs in the counting logic C^k (which, as discussed earlier, is the same as indistinguishability of graphs using the $(k-1)$-dimensional Weisfeiler-Lehman method). As a matter of fact, we show a much stronger inexpressibility result to the effect that if $\equiv$ is an equivalence relation on graphs that is finer than C^1-equivalence but coarser than C^k-equivalence for some $k \geq 2$, then there is no class $\mathscr{F}$ of graphs such that the right-homomorphism vector restricted to $\mathscr{F}$ captures the equivalence relation $\equiv$. The proof of this result uses a combination of tools from constraint satisfaction and finite model theory. The key technical tool is a definable version of the H-coloring dichotomy theorem of Hell and Nešetřil [HN90], which asserts that, for every graph H, if H is not 2-colorable, then the H-coloring problem is NP-complete, while if H is 2-colorable, then the H-coloring problem is solvable in polynomial time. The definable version of this dichotomy theorem asserts that if H is not 2-colorable, then the H-coloring problem in not definable in the infinitary counting logic $C^\omega_{\infty\omega}$, while if H is 2-colorable, then the H-coloring problem is definable by the negation of a Datalog sentence.

As a byproduct of our main inexpressibility result, we establish that there is no class $\mathscr{F}$ of graphs such that the right-homomorphism vector restricted to $\mathscr{F}$ captures co-spectrality

of graphs; by definition, two graphs are *co-spectral* if they have the same characteristic polynomial or, equivalently, if their adjacency matrices have the same multiset of eigenvalues. Note that, as pointed out by Dell, Grohe, and Rattan [DGR18], co-spectrality of graphs is captured by the left-homomorphism vector restricted to the class $\mathscr{C}$ of all cycles.

The results discussed so far concern limitations of the expressive power of restrictions of the right-homomorphism vector. Switching to the other side, we show that there is no class $\mathscr{F}$ of graphs such that the left-homomorphism vector restricted to $\mathscr{F}$ captures chromatic equivalence. In contrast, chromatic equivalence is captured by the right-homomorphism vector restricted to the class $\mathscr{K}$ of all cliques. Note that the study of chromatically equivalent graphs has a long history that starts with the work of Birkhoff [Bir12] in the early 20th Century. Furthermore, chromatically equivalent graphs have several invariants in common, including the same number of vertices, edges, triangles, components, and girth (see [Noy03]); it is worth pointing out that the first three of these invariants are left-homomorphism counts.

In this paper, we also obtain a common generalization of the characterizations of graph isomorphism by Lovász [Lov67] and by Chaudhuri and Vardi [CV93]. Finally, at the conceptual level, we discuss extensions of the framework studied here to graphs that may have loops and weights on their vertices and edges, where the weights are real numbers or, more generally, are elements of an arbitrary, but fixed, semiring. This extension of the framework makes it possible to capture several other relaxations of graph isomorphism arising from fundamental graph polynomials, such as the cluster expansion polynomial and the independence polynomial.

2   Preliminaries and Basic Concepts

Unless otherwise specified, all graphs are finite, undirected, and simple, i.e., without self-loops or multi-edges. For a graph G, we write $V(G)$ for the set of *vertices* (or *nodes*) of G and $E(G)$ for the set of *edges* of G. The vertices are *labelled*, i.e., they are elements of some fixed countable set of labels, say $\mathbb{N}$. If u and v are distinct vertices, we write (u, v) to denote the edge with endpoints u and v. Since edges are undirected, the edges (u, v) and (v, u) are the same.

Definition 1. Let G and H be two graphs. A function $g : V(G) \to V(H)$ is an *isomorphism* from G to H if g is a bijection, and for every two vertices u and v in $V(G)$, we have that $(u, v) \in E(G)$ if and only if $(g(u), g(v)) \in E(H)$. We say that G *is isomorphic to* H and we write $G \cong H$ if there exists an isomorphism g from G to H. A function $h : V(G) \to V(H)$ is a *homomorphism* from G to H if for every two vertices u and v in $V(G)$, we have that $(u, v) \in E(G)$ implies $(h(u), h(v)) \in E(H)$. We write $h : G \to H$ to denote that h is a homomorphism from G to H.

In the sequel, whenever we say that $\mathscr{F}$ is a *class* of graphs, we mean that $\mathscr{F}$ is a non-empty collection of graphs that is closed under isomorphisms, i.e., if $G \in \mathscr{F}$ and G is isomorphic to H, then $H \in \mathscr{F}$.

We now introduce notation for various classes of graphs and for particular graphs that will be used in the sequel. For $n \geq 1$, we write C_n, P_n, K_n and I_n to denote the cycle with n vertices (and length n), the path with n vertices (and length $n-1$), the clique with n vertices, and the independent set with n vertices, respectively. The cycles C_1 and C_2 are called *degenerate*. Clearly, $I_1 = K_1 = P_1 = C_1$ is the single-vertex graph and $K_2 = P_2 = C_2$ is the single-edge graph (on two vertices). We write $\mathscr{G}$, $\mathscr{T}$, $\mathscr{C}$, $\mathscr{P}$, $\mathscr{K}$, and $\mathscr{I}$ for the classes of all graphs, all (unrooted) trees, all cycles, all paths, all cliques, and all independent sets, respectively.

The central notion in this paper is that of *homomorphism count*, which we now introduce.

Definition 2. Let G and H be two graphs. We write $\mathrm{Hom}(G, H)$ to denote the set of all homomorphisms from G to H. In symbols,

$$\mathrm{Hom}(G, H) := \{h \mid h : G \to H\}.$$

We write $\hom(G, H)$ to denote the number of homomorphisms from G to H. In symbols,

$$\hom(G, H) := |\mathrm{Hom}(G, H)|.$$

The homomorphism count can be viewed as a function $\hom : \mathscr{G} \times \mathscr{G} \to \mathbb{N}$ from pairs of graphs to natural numbers. If we pick only one graph from each isomorphism type, then the homomorphism count can be visualized as the following infinite 2-dimensional matrix:

$\hom(\cdot, \cdot)$	G_1	$\cdots$	G_j	$\cdots$
G_1	$\hom(G_1, G_1)$	$\cdots$	$\hom(G_1, G_j)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$
G_i	$\hom(G_i, G_1)$	$\cdots$	$\hom(G_i, G_j)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

For a graph G, the column of this matrix indexed by G is referred to as the left-homomorphism vector of G, and the row of this matrix indexed by G is referred to as the right-homomorphism vector of G. If $\mathscr{F}$ is a class of graphs, the *left-homomorphism vector of G restricted to $\mathscr{F}$* and the *right-homomorphism vector of G restricted to $\mathscr{F}$* are defined as:

$$\hom(\mathscr{F}, G) := (\hom(F, G) \mid F \in \mathscr{F}),$$
$$\hom(G, \mathscr{F}) := (\hom(G, F) \mid F \in \mathscr{F}).$$

When $\mathscr{F}$ is the class $\mathscr{G}$ of all graphs, the reference to the restriction to the class $\mathscr{F}$ is omitted from the terminology.

We now state two well-known theorems by Lovász [Lov67] and by Chaudhuri and Vardi [CV93], which characterize graph isomorphism in terms of left-homomorphism vectors and right-homomorphism vectors, respectively.

Theorem 1. *[Lov67] For every two graphs G and H, we have that $G \cong H$ if and only if $\hom(\mathscr{G}, G) = \hom(\mathscr{G}, H)$.* □

Theorem 2. *[CV93] For every two graphs G and H, we have that $G \cong H$ if and only if $\mathrm{hom}(G,\mathscr{G}) = \mathrm{hom}(H,\mathscr{G})$.* □

Since we have picked only one graph from each isomorphism type in constructing the matrix, Theorem 1 asserts that *no* two columns of the matrix are the same, while Theorem 2 asserts that *no* two rows of the matrix are the same. The 'only if' direction of both these theorems is trivial.

3 Generalizing Lovász and Chaudhuri-Vardi

The preceding theorems by Lovász and by Chaudhuri and Vardi characterize graph isomorphism in terms of the left-homomorphism vector and the right-homomorphism vector on the class $\mathscr{G}$ of all graphs. In this section, we generalize these result to classes $\mathscr{F}$ of graphs that satisfy certain conditions.

Definition 3. Let G and H be two graphs. A homomorphism $h : G \to H$ is *injective* if the mapping $h : V(G) \to V(H)$ is injective. We write $\mathrm{inj}(G,H)$ to denote the number of injective homomorphisms from G to H. A homomorphism $h : G \to H$ is *surjective* if the image $h(G)$ of G under h coincides with H, i.e., $V(h(G)) = V(H)$ and $E(h(G)) = E(H)$. We write $\mathrm{sur}(G,H)$ to denote the number of surjective homomorphisms from G onto H. An isomorphism from G to G is called an *automorphism* of G. We write $\mathrm{aut}(G)$ for the number of automorphisms of G.

Definition 4. Let $\mathscr{F}$ be a class of graphs. We write $\mathrm{Inj}(\mathscr{F})$ to denote the class of all graphs G such that there is an injective homomorphism $h : G \to F$ to some graph $F \in \mathscr{F}$. We write $\mathrm{Sur}(\mathscr{F})$ to denote the class of all graphs G such that there is a surjective homomorphism $h : F \to G$ from some graph $F \in \mathscr{F}$ onto G. The *extension class of $\mathscr{F}$*, denoted by $\mathrm{Ext}(\mathscr{F})$, is the intersection of $\mathrm{Inj}(\mathscr{F})$ and $\mathrm{Sur}(\mathscr{F})$. In symbols,

$$\mathrm{Ext}(\mathscr{F}) := \mathrm{Sur}(\mathscr{F}) \cap \mathrm{Inj}(\mathscr{F}).$$

It is obvious that $\mathscr{F} \subseteq \mathrm{Inj}(\mathscr{F})$ and $\mathscr{F} \subseteq \mathrm{Sur}(\mathscr{F})$. Therefore, for every class $\mathscr{F}$ of graphs, we have that $\mathscr{F} \subseteq \mathrm{Ext}(\mathscr{F})$.

The next result is a simultaneous generalization of Theorem 1 and Theorem 2.

Theorem 3. *Let $\mathscr{F}$ be a non-empty class of graphs. For every two graphs G and H in $\mathscr{F}$, the following statements are equivalent:*

(1) G and H are isomorphic.

(2) $\mathrm{hom}(\mathrm{Ext}(\mathscr{F}),G) = \mathrm{hom}(\mathrm{Ext}(\mathscr{F}),H)$.

(3) $\mathrm{hom}(G,\mathrm{Ext}(\mathscr{F})) = \mathrm{hom}(H,\mathrm{Ext}(\mathscr{F}))$.

In particular, if $\mathrm{Ext}(\mathscr{F}) = \mathscr{F}$, then for every two graphs G and H in $\mathscr{F}$, we have that G and H are isomorphic if and only if their left-homomorphism vectors restricted to $\mathscr{F}$ are equal, and if and only if their right-homomorphism vectors restricted to $\mathscr{F}$ are equal.

Proof. In the sequel, we assume that the class $\mathscr{G}$ of graphs is linearly ordered first in increasing order of $|V(G)|$, then in increasing order of $|E(G)|$, and, finally, arbitrarily in case of a tie; e.g., in lexicographical order of the rows of the adjacency matrix. We write $G < H$ to denote that G precedes H in this ordering. Moreover, we assume that in summations over a class $\mathscr{F}$ of graphs, the summand $F \in \mathscr{F}$ runs over the isomorphism types so that isomorphic graphs contribute only once to the summation.

The directions $(1) \Rightarrow (2)$ and $(1) \Rightarrow (3)$ are trivial. To prove $(2) \Rightarrow (1)$, let $\mathscr{F}$ be a non-empty class, and let G and H be two graphs in $\mathscr{F}$ such that $\hom(\text{Ext}(\mathscr{F}), G) = \hom(\text{Ext}(\mathscr{F}), H)$. Our goal is to show that $G \cong H$, and we do so by arguing that $\text{inj}(G, H) > 0$ and $\text{inj}(H, G) > 0$.

By induction on the position of $D \in \text{Ext}(\mathscr{F})$ in the linear order $<$ on $\mathscr{G}$, we show that $\text{inj}(D, G) = \text{inj}(D, H)$. From this the goal will follow by setting $D = H$, i.e., $\text{inj}(H, G) = \text{inj}(H, H) > 0$, and $D = G$, i.e., $\text{inj}(G, H) = \text{inj}(G, G) > 0$. We start by noting that

$$\hom(D, G) = \sum_{E \in \mathscr{G}} \text{sur}(D, E) \cdot \text{inj}(E, G)/\text{aut}(E). \tag{1}$$

Next observe that, for $E \in \mathscr{G}$, if $\text{sur}(D, E) > 0$ and $\text{inj}(E, G) > 0$, then $E \in \text{Sur}(\mathscr{F})$ since $D \in \text{Ext}(\mathscr{F}) \subseteq \text{Sur}(\mathscr{F})$, and also $E \in \text{Inj}(\mathscr{F})$ since $G \in \mathscr{F}$; that is, $E \in \text{Ext}(\mathscr{F})$. This means that the sum in (1) can be restricted to $E \in \text{Ext}(\mathscr{F})$. Moreover, if E has more vertices or edges than D, or if E and D have the same number of vertices and edges but are not isomorphic, then $\text{sur}(D, E) = 0$. Thus, the sum in (1) can be restricted further to $E < D$ or $E \cong D$. Since, by convention, sums over classes of graphs are restricted to isomorphism types, we get

$$\text{inj}(D, G) + \sum_{\substack{E \in \text{Ext}(\mathscr{F}): \\ E < D}} \text{sur}(D, E) \cdot \text{inj}(E, G)/\text{aut}(E). \tag{2}$$

The same can be argued for H in place G. By induction hypothesis we have $\text{inj}(E, G) = \text{inj}(E, H)$ for every E in the sum in (2), and by assumption we have $\hom(D, G) = \hom(D, H)$. The conclusion is that $\text{inj}(D, G) = \text{inj}(D, H)$, as was to be proved.

The direction $(3) \Rightarrow (1)$ can be argued analogously, so we highlight only the key differences. This time we want to argue that $\text{sur}(G, H) > 0$ and $\text{sur}(H, G) > 0$, and we do so by arguing that $\text{sur}(G, D) = \text{sur}(H, D)$ for every $D \in \text{Ext}(\mathscr{F})$. The analogue of Equation (1) is

$$\hom(G, D) = \sum_{E \in \mathscr{G}} \text{sur}(G, E) \cdot \text{inj}(E, D)/\text{aut}(E). \tag{3}$$

Next, for $E \in \mathscr{G}$, if $\text{sur}(G, E) > 0$ and $\text{inj}(E, D) > 0$, then $E \in \text{Sur}(\mathscr{F})$ since $G \in \mathscr{F}$, and also $E \in \text{Inj}(\mathscr{F})$ since $D \in \text{Ext}(\mathscr{F}) \subseteq \text{Inj}(\mathscr{F})$; that is, $E \in \text{Ext}(\mathscr{F})$. Thus, the sum in (3) can be restricted to $\text{Ext}(\mathscr{F})$, and by the same type of argument as before, further down to

$$\text{sur}(G, D) + \sum_{\substack{E \in \text{Ext}(\mathscr{F}): \\ E < D}} \text{sur}(G, E) \cdot \text{inj}(E, D)/\text{aut}(E). \tag{4}$$

7

The same can be argued for H in place G. By induction hypothesis we have $\mathrm{sur}(E, G) = \mathrm{sur}(E, H)$ for every E in the sum in (4), and by assumption we have $\hom(G, D) = \hom(H, D)$. The conclusion is that $\mathrm{sur}(G, D) = \mathrm{sur}(H, D)$, as was to be proved. $\qquad \Box$

Several remarks about Theorem 3 are now in order.

First, Theorem 3 is indeed a common generalization of Theorem 1 and Theorem 2 because it is clear that the condition $\mathrm{Ext}(\mathscr{G}) = \mathscr{G}$ holds for the class $\mathscr{G}$ of all graphs.

Second, since $\mathrm{Ext}(\mathscr{F}) = \mathrm{Sur}(\mathscr{F}) \cap \mathrm{Inj}(\mathscr{F})$, we have that if $\mathrm{Sur}(\mathscr{F}) = \mathscr{F}$ or if $\mathrm{Inj}(\mathscr{F}) = \mathscr{F}$ hold, then $\mathrm{Ext}(\mathscr{F}) = \mathscr{F}$ holds. Therefore, Theorem 3 applies to every class $\mathscr{F}$ of graphs such that $\mathrm{Sur}(\mathscr{F}) = \mathscr{F}$ or $\mathrm{Inj}(\mathscr{F}) = \mathscr{F}$.

Third, Theorem 3 applies to each of the following classes $\mathscr{F}$ of graphs: the class $\mathscr{P}$ of all paths, the class $\mathscr{T}$ of all trees, the class $\mathscr{T}_k$ of graphs of treewidth at most k, the class $\mathscr{K}$ of all cliques, the class $\mathscr{X}_k$ of all k-colorable graphs, and the class $\mathscr{D}_k$ of all graphs of degree at most k. For the classes $\mathscr{T}_k$, $\mathscr{X}_k$, and $\mathscr{D}_k$, we actually have $\mathrm{Inj}(\mathscr{F}) = \mathscr{F}$, while for $\mathscr{K}$, we have $\mathrm{Sur}(\mathscr{K}) = \mathscr{K}$.

To see that $\mathrm{Ext}(\mathscr{T}) = \mathscr{T}$, it suffices to show $\mathrm{Ext}(\mathscr{T}) \subseteq \mathscr{T}$ since it is clear that $\mathscr{T} \subseteq \mathrm{Ext}(\mathscr{T})$. First, observe that trees are connected graphs and hence so are their homomorphic images since the homomorphic image of a connected graph is connected. In other words, the graphs in $\mathrm{Sur}(\mathscr{T})$ are connected. Next, $\mathrm{Inj}(\mathscr{T})$ contains all subgraphs of trees. Therefore, every graph in $\mathrm{Ext}(\mathscr{T}) = \mathrm{Sur}(\mathscr{T}) \cap \mathrm{Inj}(\mathscr{T})$ is a connected subgraph of a tree, which must be a tree. It follows that $\mathrm{Ext}(\mathscr{T}) \subseteq \mathscr{T}$. The same type of argument works for $\mathscr{P}$.

The final remark of this section concerns the class $\mathscr{T}_k$ of graphs of treewidth at most k. It follows from the above discussion that the left-homomorphism vector $\hom(\mathscr{T}_k, \cdot)$ characterizes isomorphism on $\mathscr{T}_k$, namely, for graphs G and H in $\mathscr{T}_k$ it holds that G and H are isomorphic if and only if $\hom(\mathscr{T}_k, G) = \hom(\mathscr{T}_k, H)$. By Theorems 1 and 3 in [DGR18], the vector $\hom(\mathscr{T}_k, \cdot)$ also characterizes C^{k+1}-equivalence; see also Theorems 4, 5, 6 and 8 in this paper. Consequently, the equivalence in C^{k+1} determines isomorphism on $\mathscr{T}_k$. This reproves and strengthens a result of Grohe and Mariño [GM99]; see Theorem 4 in that paper.

4 Right-Homomorphism Vectors

Let $k \geq 1$ be a fixed positive integer. The *k-dimensional Weisfeiler-Lehman method* [Wei06] is an iterative algorithm for graph isomorphism that assigns colors to each k-tuple of vertices. At each iteration, the algorithm refines the color classes of the k-tuples; the algorithm stops when a *stable* coloring is achieved, i.e., when no further refinement of the color classes of the k-tuples is possible. This method is a heuristic for graph isomorphism in the sense that if the k-dimensional Weisfeiler-Lehman method produces different stable colorings when applied to two graphs G and H, then the graphs G and H are not isomorphic. However, this is not a complete isomorphism test because, for each $k \geq 1$, there are non-isomorphic graphs G_k and H_k on which the k-dimensional Weisfeiler-Lehman method produces the same coloring. We write $G \equiv_{\mathrm{WL}}^k H$ to denote that the graphs G and H are indistinguishable via

the k-dimensional Weisfeiler-Lehman method. In view of the preceding discussion, $\equiv_{\mathrm{WL}}^{k}$ is an equivalence relation that is strictly coarser than isomorphism.

We describe in more detail the 1-dimensional Weisfeiler-Lehman method. Initially, all vertices have the same color. At each iteration, two vertices u and v that were in the same color class are assigned different colors if there is a color c such that the number of neighbors of u that have color c is different from the number of neighbors of v that have color c. This process continues until all vertices in the same color class have the same number of neighbors in every color class. Tinhofer [Tin86, Tin91] has shown that the 1-dimensional Weisfeiler-Lehman method amounts to the *fractional isomorphism* test. Let $G = (V, E)$ and $H = (V, E')$ be two graphs on the same set of vertices and let A and B be their adjacency matrices. Clearly, G and H are isomorphic if and only if there is permutation matrix X such that $AX = XB$, where a permutation matrix is a square matrix of 0's and 1's such that each row and each column contains exactly one 1. Consider now the system of linear equations $AX = XB$, $Xe = e$, $e^{\top}X = e^{\top}$, where e is the vector of 1's of length $|V|$. The graphs G and H are said to be *fractionally isomorphic* if this system has a non-negative solution over the rational numbers, which is called a *fractional isomorphism* between G and H.

Theorem 4. *[Tin86, Tin91] For every two graphs G and H, we have that G and H are fractionally isomorphic if and only if $G \equiv_{\mathrm{WL}}^{1} H$.*

Cai, Fürer, and Immerman [CFI92] characterized indistinguishability via the Weisfeiler-Lehman method in terms of indistinguishability in first-order logic with counting and a fixed number of variables. A *counting quantifier* is a quantifier of the form $\exists ix$, where i is a positive integer. The meaning of a formula $\exists ix\varphi(x)$ is that there are at least i distinct elements x such that $\varphi(x)$ holds. If φ is an FO-formula, then the formula $\exists ix\varphi(x)$ is clearly equivalent to an FO-formula, so counting quantifiers do not add expressive power to first-order logic. Counting quantifiers, however, become interesting when we consider logics with a fixed number of distinct variables. For every $k \geq 1$, let FO^{k} be the fragment of first-order logic FO consisting of all FO-formulas with at most k distinct variables, and let C^{k} be the logic obtained from FO^{k} by augmenting the syntax with all counting quantifiers $\exists ix$, where $i \geq 1$. For example, the formula

$$\exists ix(x = x) \wedge \neg\exists(i + 1)x(x = x) \wedge \forall x\forall y(x \neq y \rightarrow E(x, y))$$

is a C^2-formula that is satisfied by a graph G precisely when G is the clique K_i with i vertices. We say that two graphs G and H are C^{k}-*equivalent*, denoted by $G \equiv_{\mathrm{C}}^{k} H$, if G and H satisfy the same C^{k}-sentences. Detailed information about counting logics with finitely many variables can be found in the monograph [Ott17].

Theorem 5. *[CFI92] For every $k \geq 2$ and every two graphs G and H, we have that $G \equiv_{\mathrm{C}}^{k} H$ if and only if $G \equiv_{\mathrm{WL}}^{k-1} H$.*

An immediate consequence of Theorems 4 and 5 is that two graphs are fractionally isomorphic if and only if they are C^2-equivalent. The following family of pairwise fractionally isomorphic graphs will be of interest to us in the sequel.
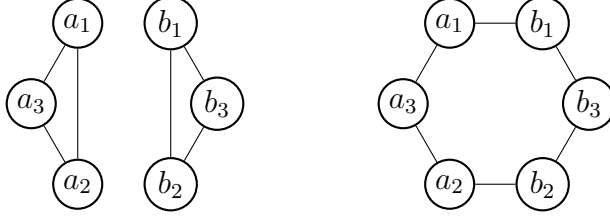
Figure 1: Two fractionally isomorphic graphs, G_3 (left) and H_3 (right).

Example 1. For every $n \geq 3$, consider the graphs G_n and H_n defined as follows: $G_n :=$ $A_n \oplus B_n$, where $\oplus$ denotes the disjoint union of two graphs, A_n and B_n are two isomorphic copies of K_n, and H_n is formed from G_n by first removing an arbitrary edge $(a_1, a_2) \in E(A_n)$ and an arbitrary edge $(b_1, b_2) \in E(B_n)$, and then adding two edges $(a_1, b_1), (a_2, b_2)$. The graphs G_3, H_3 are depicted in Figure 1.

The graphs G_n and H_n are fractionally isomorphic. The reason is that both are regular (each vertex has $n-1$ neighbors) and they have the same number of vertices and edges, thus the 1-dimensional Weisfeiler-Lehman method terminates at the first step with all vertices assigned the same color.

4.1 Fractional Isomorphism

Dell, Grohe, and Rattan [DGR18] showed that fractional isomorphism can be characterized by restricting the left-homomorphism vector.

Theorem 6. *[DGR18] For every two graphs G and H, we have that G and H are fractionally isomorphic if and only if* $\mathrm{hom}(\mathscr{T}, G) = \mathrm{hom}(\mathscr{T}, H)$, *where $\mathscr{T}$ is the class of all trees.*

In contrast, we show that fractional isomorphism cannot be characterized by restricting the right-homomorphism vector. While this result will follow from the more general Theorem 10 proved in the next subsection, the proof for the special case of fractional isomorphism is more elementary and self-contained, so we include it here.

Theorem 7. *There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that G and H are fractionally isomorphic if and only if* $\mathrm{hom}(G, \mathscr{F}) = \mathrm{hom}(H, \mathscr{F})$.

Proof. Towards a contradiction, assume that such a class $\mathscr{F}$ exists. We distinguish the following two cases.

Case 1. $\mathscr{F} \subseteq \mathscr{I}$, i.e., each graph in $\mathscr{F}$ is an independent set. This leads immediately to a contradiction. Indeed, the cliques K_2 and K_3 are not fractionally isomorphic, but $\mathrm{hom}(K_2, \mathscr{F}) = \mathrm{hom}(K_3, \mathscr{F})$ because, for every $F \in \mathscr{F}$, we have that $\mathrm{hom}(K_2, F) = 0 = \mathrm{hom}(K_3, F)$.

Case 2. $\mathscr{F} \not\subseteq \mathscr{I}$, which means that there is at least one graph $F \in \mathscr{F}$ such that F contains K_2 as a subgraph. We now bring into the picture the family of graphs G_n and H_n, with $n \geq 3$, considered in Example 1. Since G_n and H_n are fractionally isomorphic graphs,

10

the hypothesis for the class $\mathscr{F}$ implies that $\mathrm{hom}(G_n, F) = \mathrm{hom}(H_n, F)$, for every $n \geq 3$. Furthermore, we claim that the following properties hold for every graph D and every $n \geq 3$. We use the notation $G \subseteq H$ to denote the fact that H contains G as a subgraph.

P1. $\mathrm{hom}(G_n, D) > 0$ if and only if $K_n \subseteq D$.
P2. $\mathrm{hom}(H_n, D) > 0$ if and only if $K_{n-1} \subseteq D$.

For the first property, observe that if $h : G_n \rightarrow D$ is a homomorphism, then the homomorphic image $h(G_n)$ must contain K_n as a subgraph because K_n is a subgraph of G_n and is preserved under h; it follows that D contains K_n as a subgraph, too. Conversely, if D contains K_n as a subgraph, then there exists a homomorphism $h' : G_n \rightarrow D$ that maps each of the two disjoint copies A_n and B_n of K_n in G_n to the K_n in D.

For the second property, the 'only if' direction can be argued in the same way as that in the first part using the fact that K_{n-1} is a subgraph of H_n. For the 'if' direction, assume that D contains K_{n-1} as a subgraph; then there is a homomorphism $h' : H_n \rightarrow D$ that maps H_n onto K_{n-1} in D where $h(a_1) = h(a_2) = h(b_3)$ and $h(b_1) = h(b_2) = h(a_3)$ for some $a_3 \in V(A_n) \setminus \{a_1, a_2\}$ and some $b_3 \in V(B_n) \setminus \{b_1, b_2\}$.

Using these two properties, we will show by induction on n that F contains K_n as a subgraph, for every $n \geq 2$, which is absurd since $|V(F)|$ is finite. The base case ($n = 2$) is just the assumption that F contains K_2 as a subgraph. For the inductive case, let $n \geq 2$ and suppose that F contains K_n as a subgraph. By taking $D = F$ and applying the second property, we have that $\mathrm{hom}(H_{n+1}, F) > 0$, hence $\mathrm{hom}(G_{n+1}, F) > 0$ since $\mathrm{hom}(G_{n+1}, F) = \mathrm{hom}(H_{n+1}, F)$. By applying the first property, we conclude that F contains K_{n+1} as a subgraph. This completes the proof of the theorem. $\qquad\square$

4.2   Counting-Logic Equivalence

Dell, Grohe, and Rattan [DGR18] also showed that, for every $k \geq 3$, equivalence in C^k can be characterized by restricting the left-homomorphism vector.

Theorem 8. *[DGR18] For every integer $k \geq 3$ and every two graphs G and H, we have that $G \equiv_{\mathrm{C}}^k H$ if and only if $\mathrm{hom}(\mathscr{T}_{k-1}, G) = \mathrm{hom}(\mathscr{T}_{k-1}, H)$, where $\mathscr{T}_{k-1}$ is the class of all graphs of treewidth at most $k - 1$.*

In contrast, we show that C^k-equivalence, $k \geq 2$, cannot be characterized by restricting the right-homomorphism vector. Since fractional isomorphism and C^2-equivalence are the same, the case $k = 2$ coincides with the statement of Theorem 7. We chose to state Theorem 7 earlier and present its proof, because that proof is self-contained, unlike the proof of Theorem 9 below, which requires sophisticated tools.

Theorem 9. *For every $k \geq 2$, there is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that $G \equiv_{\mathrm{C}}^k H$ if and only if $\mathrm{hom}(G, \mathscr{F}) = \mathrm{hom}(H, \mathscr{F})$.*

This theorem will be an immediate consequence of the following much more general result, which states that if an equivalence relation on graphs interpolates between $\equiv_{\mathrm{C}}^k$ and $\equiv_{\mathrm{C}}^1$ for

some $k \geq 2$, then it cannot be characterized by restricting the right-homomorphism vector. We note that, for loopless graphs (as is our case), two graphs are $\equiv_{\mathrm{C}}^1$-equivalent if and only if they have the same number of vertices.

Theorem 10. *For every equivalence relation $\equiv$ on graphs that is finer than $\equiv_{\mathrm{C}}^1$ and coarser than $\equiv_{\mathrm{C}}^k$ for some $k \geq 2$, there is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that $G \equiv H$ if and only if $\mathrm{hom}(G, \mathscr{F}) = \mathrm{hom}(H, \mathscr{F})$.*

To prove this we need to bring in some tools from the theory of constraint satisfaction. We focus on the special case of this theory that applies to graphs since this is all we need.

Fix a graph H. A graph G is called *H-colorable* if there is a homomorphism from G to H. The H-coloring problem asks: given a graph as input, is it H-colorable? The name of the problem reflects the fact that the K_k-coloring problem is the same as that of deciding whether a given graph as input has a proper coloring with k colors. The celebrated H-Coloring Dichotomy Theorem, due to Hell and Nešetřil [HN90], asserts that the H-coloring problem is NP-complete if H is non-2-colorable, and solvable in polynomial time if H is 2-colorable. Here, we will use what we call the *Definable H-Coloring Dichotomy Theorem*. The logics that are mentioned in its statement are defined immediately following it.

Theorem 11. *Let H be a graph. The class of H-colorable graphs is not definable in the logic $\mathrm{C}_{\infty\omega}^\omega$ unless H is 2-colorable, in which case it is definable by the negation of a 3-Datalog sentence.*

The logic $\mathrm{C}_{\infty\omega}^\omega$ denotes the union of the logics $\mathrm{C}_{\infty\omega}^k$ as k ranges over the positive integers, where $\mathrm{C}_{\infty\omega}^k$ is the logic that is obtained from C^k by augmenting the syntax with infinitary disjunctions and conjunctions. It is known that $\mathrm{C}_{\infty\omega}^k$, and even its counting-free, existential, positive fragment $\exists \mathrm{L}_{\infty\omega}^k$, is at least as expressive as the k-variable fragment k-Datalog of Datalog; see Theorem 4.1 in [KV00].

We claim that Theorem 11 is implicit in the literature and follows by combining known results. Concretely, Theorem 11 can be seen to follow by combining the implication (c) to (b) of Theorem 1 in [Bul05] with Corollary 23 in [ABD09]. Unfortunately, showing that the output of Theorem 1 in the first of these references can serve as input for Corollary 23 in the second reference is not straightforward because it would require us to introduce the notions of Tame Congruence Theory that are used in the statement of Corollary 23 in [ABD09]. For this reason, and also because we were not able to find a concrete reference where Theorem 11 is stated and proved, we provide some details here.

The part of Theorem 11 that states that the class of H-colorable graphs is definable by the negation of a 3-Datalog sentence when H is 2-colorable is well-known: If H is an independent set, then the class of H-colorable graphs coincides with class $\mathscr{I}$ of independent sets, which is clearly definable as the negation of a Datalog program even with two variables. If H has at least one edge but is 2-colorable, then the class of H-colorable graphs is precisely the class of 2-colorable graphs, which is one of the canonical examples of definability in (the negation of) Datalog by stating that the graph contains an odd cycle. A Datalog program for this that uses four variables was given in Section 4.1 of [KV00], but a simple optimization shows that three variables are enough. It is also easy to see that three variables are necessary.

12

Next, we focus on the non-2-colorable case of Theorem 11. The development of the theory is facilitated by the use of finite relational structures of richer vocabularies. Graphs are seen as finite structures of a vocabulary that has a single binary relation symbol whose interpretation is symmetric and irreflexive. Let A be a finite relational structure with domain $D(A)$. If $S \subseteq D(A)$ is a subset of the domain of A, then we write A_S for the substructure of A induced by S; i.e., the domain of A_S is S, and its relations are the sets of tuples of elements in S that are tuples in the corresponding relation in A. If $E \subseteq D(A) \times D(A)$ is an equivalence relation on the domain of A, then we write A/E for the quotient structure of A; i.e., the domain of A/E is the set of E-equivalence classes $[a]$ of elements a in the domain of A, and its relations are the sets of tuples of equivalence classes $([a_1], \ldots, [a_r])$ such that $(a_1, \ldots, a_r)$ is a tuple in the corresponding relation in A. We write A^c for the singleton-expansion of A; i.e., the expansion of A with a new unary relation symbol P_a interpreted by the singleton set $\{a\}$ for each element a in the domain of A. A *primitive positive* formula, or pp-formula, is a first-order formula of the form $\exists y_1 \cdots \exists y_s \psi(x_1, \ldots, x_r, y_1, \ldots, y_s)$, where ψ is a conjunction of atomic formulas on the variables $x_1, \ldots, x_r$ and $y_1, \ldots, y_s$; among the atomic formulas we include equalities between variables. Let R be a relation on $D(A)$ of arity r. We say that a A *pp-defines* R if there exists a pp-formula $\varphi(x_1, \ldots, x_r)$ such that $R = \{(a_1, \ldots, a_r) \in D(A)^r : A \models \varphi(x_1/a_1, \ldots, x_r/a_r)\}$. We also say that R is *pp-definable* in A. The following is the algebraic version of the H-Coloring Dichotomy Theorem due to Bulatov; see the implication (c) to (b) in Theorem 1 in [Bul05].

Theorem 12. *[Bul05] If H is a non-2-colorable core graph, then there exists a subset S of vertices of H and an equivalence relation $E \subseteq S \times S$ on S such that the following three statements hold: (1) H^c pp-defines S, (2) H_S^c pp-defines E, and (3) $(H_S)/E$ is isomorphic to K_3.*

A graph G is a *core* if G does not have a retraction to a proper subgraph H, i.e., a homomorphism from $V(G)$ to $V(H)$ that is the identity on $V(H)$ (see [HN04] for detailed discussion of retractions and cores). Also, to clarify, the statement (c) in Theorem 1 in [Bul05] says that the equivalence relation E is pp-definable in H^c, but the proof actually shows that it is pp-definable in H_S^c. This is, in principle, a stronger statement, and it is what we actually need below. We note that the implications (a) to (b) and (a) to (c) in Theorem 1 in [Bul05] tacitly assume that P $\neq$ NP, but we do not use that part of the theorem.

The next ingredient in the proof is the concept of logical reducibility between classes of finite structures. We refer the reader to Definition 1 in [ABD09] for the definition of Datalog-reducibility, denoted by $\leq_{\text{datalog}}$ there and by $\leq_d$ here, for brevity. We need only one property about $\leq_d$, namely, that if $\mathscr{B}$ a class of finite structures that is $C_{\infty\omega}^\omega$-definable and $\mathscr{A}$ is another class of finite structures such that $\mathscr{A} \leq_d \mathscr{B}$, then also $\mathscr{A}$ is $C_{\infty\omega}^\omega$-definable. This follows from the aforementioned fact that Datalog is a fragment of $\exists L_{\infty\omega}^\omega \subseteq C_{\infty\omega}^\omega$. For a finite structure A, let $\text{CSP}(A)$ denote the class of structures of the same vocabulary as A that have a homomorphism to A. The following was proved in [BJK05] for polynomial-time reducibility, and in [ABD09] for $\leq_d$ reducibility.

Theorem 13. *[ABD09] Let A be a finite relational structure with at least two elements,*

let $S \subseteq D(A)$ be a subset of the domain of A, and let $E \subseteq D(A) \times D(A)$ be an equivalence relation on the domain of A. The following hold:

(1) $\mathrm{CSP}(A) \leq_\mathrm{d} \mathrm{CSP}(A^c)$.

(2) If A pp-defines E, then $\mathrm{CSP}(A/E) \leq_\mathrm{d} \mathrm{CSP}(A)$.

(3) If A pp-defines S, then $\mathrm{CSP}(A_S) \leq_\mathrm{d} \mathrm{CSP}(A)$.

(4) If A is a core, then $\mathrm{CSP}(A^c) \leq_\mathrm{d} \mathrm{CSP}(A)$.

In the statement of part (4), the core of a relational structure is the straightforward generalization of the core of a graph to structures of arbitrary relational vocabularies. To see how Theorem 13 follows from the statements in [ABD09], note that (1) follows from a special case of Lemma 11 there, that (2) and (3) are special cases of Theorem 18 there, and that (4) follows from the second part of Lemma 19 there.

By combining Theorem 13 with Theorem 12, we get that if H is a non-2-colorable graph, then $\mathrm{CSP}(K_3) \leq_\mathrm{d} \mathrm{CSP}(H)$. Here we used also the obvious fact that if H' is the unique core graph of H, then $\mathrm{CSP}(H') = \mathrm{CSP}(H)$. The final link that yields the proof of Theorem 11 is the following well-known result of Dawar, which was obtained by adapting the main result in [CFI92]; see Theorem 4.11 and Remark 4.12 in [Daw98].

Theorem 14. *[Daw98] The class of 3-colorable graphs is not $\mathrm{C}^\omega_{\infty\omega}$-definable.*

By the aforementioned downwards preservation of $\mathrm{C}^\omega_{\infty\omega}$-definability through the reducibility $\leq_\mathrm{d}$, this completes the proof of Theorem 11. We are now ready to prove Theorem 10 and, hence, Theorem 9.

Proof of Theorem 10. Fix an equivalence relation $\equiv$ that is finer than $\equiv^1_\mathrm{C}$ and is coarser than $\equiv^k_\mathrm{C}$ for some $k \geq 2$. Towards a contradiction, assume that there exists a class of graphs $\mathscr{F}$ such that $G \equiv H$ holds if and only if $\mathrm{hom}(G, \mathscr{F}) = \mathrm{hom}(H, \mathscr{F})$ holds. We distinguish the following two cases:

Case 1. All graphs in $\mathscr{F}$ are 2-colorable. This leads immediately to a contradiction. Indeed, the cliques K_3 and K_4 have different sizes, so they are not C^1-equivalent, hence they are not $\equiv$-equivalent. But $\mathrm{hom}(K_3, F) = \mathrm{hom}(K_4, F) = 0$ for every 2-colorable graph F, so $\mathrm{hom}(K_3, \mathscr{F}) = \mathrm{hom}(K_4, \mathscr{F})$.

Case 2. Some graph in $\mathscr{F}$ is non-2-colorable. Let $H \in \mathscr{F}$ be such a graph. By Theorem 11, the class of H-colorable graphs is not $\mathrm{C}^k_{\infty\omega}$-definable. By Corollary 2.4 in [Ott17] this means that the class of H-colorable graphs is not preserved by $\equiv^k_\mathrm{C}$-equivalence, i.e., there exists a pair of graphs G_0 and G_1 such that $G_0 \equiv^k_\mathrm{C} G_1$ with $\mathrm{hom}(G_0, H) = 0$ and $\mathrm{hom}(G_1, H) \neq 0$. In particular, $G_0 \equiv G_1$ and $\mathrm{hom}(G_0, H) \neq \mathrm{hom}(G_1, H)$, against the fact that H is in $\mathscr{F}$ and the assumption on $\mathscr{F}$. $\qquad\square$

4.3 Cospectrality

The *characteristic polynomial* of a graph G is defined as $p(G, x) = \det(xI - A_G)$, where A_G denotes the adjacency matrix of G and I denotes the identity matrix of the same dimension

as A_G. Here, det denotes the determinant of a square matrix. The zeros of $p(G, x)$ are the eigenvalues of the adjacency matrix A_G. Two graphs G and H are called *cospectral* if their characteristic polynomials are the same; equivalently, G and H are cospectral if their adjacency matrices A_G and A_H have the same multisets of eigenvalues. Dell, Grohe and Rattan proved the following characterization of cospectral graphs; see Proposition 9 in [DGR18], attributed to [VDH03].

Theorem 15. *[DGR18] For every two graphs G and H, we have that G and H are cospectral if and only if $\hom(\mathscr{C}, G) = \hom(\mathscr{C}, H)$, where $\mathscr{C}$ is the class of all cycles, including the degenerate cycles with one and two vertices.*

Here we show that, in contrast, cospectrality cannot be characterized as the restriction of a right-homomorphism vector. Interestingly, as we will see, the proof of this result is an application of Theorem 10 with $k = 3$.

Theorem 16. *There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that G and H are cospectral if and only if $\hom(G, \mathscr{F}) = \hom(H, \mathscr{F})$.*

Proof. It is known that any two $\equiv_{\mathrm{C}}^3$-equivalent graphs are cospectral; since all cycles have treewidth at most two, this follows by combining Theorems 15 and 8, but see also Theorem 2.1 in [DSZ17]. Moreover, any two cospectral graphs have the same number of vertices. Therefore, cospectrality is finer than $\equiv_{\mathrm{C}}^1$ and coarser than $\equiv_{\mathrm{C}}^3$. The result follows from Theorem 10. □

5 Left-Homomorphism Vectors

For every graph G, there is a polynomial $\chi(G, x)$ in the variable x, called the *chromatic polynomial* of G, such that, for every $k \geq 1$, the value $\chi(G, k)$ is the number of k-colorings of G (see the survey [Rea68] for the history of chromatic polynomials and basic facts about them). Furthermore, if G has n vertices, then $\chi(G, x)$ has degree n and its leading coefficient is 1. For example,

- $\chi(I_n, x) = x^n$, where $n \geq 1$;
- $\chi(K_n, x) = x(x - 1) \cdots (x - n + 1)$, where $n \geq 1$;
- $\chi(C_n, x) = (x - 1)^n + (-1)^n(x - 1)$, where $n \geq 3$.

There are two useful techniques for deriving the chromatic polynomial $\chi(G, x)$ of an arbitrary graph G.

Multiplicativity. If $G = G_1 \oplus G_2$ is the disjoint union of two graphs G_1 and G_2, then $\chi(G, x) = \chi(G_1, x) \cdot \chi(G_2, x)$.

Addition-Contraction Recursion. If $u, v \in V(G)$ are distinct and furthermore if $(u, v) \notin E(G)$, then $\chi(G, x) = \chi(G_1, x) + \chi(G_2, x)$ where G_1 is obtained from G by adding the edge (u, v), while G_2 is obtained from G by contracting the two vertices u, v.

Figure 2: Two chromatically equivalent graphs, X_1 (left) and X_2 (right).

The base cases of the addition-contraction recursion are of the form $\chi(K_n, x) = x(x - 1) \cdots (x - n + 1)$. Thus, to obtain $\chi(G, x)$ for a graph G that is not a clique and has n vertices, we can use this recursion to expand $\chi(G, x)$ until every term in the expansion is of the form $\chi(K_m, x)$ with $m \leq n$.

Example 2. Consider the two graphs $X_1 = I_1 \oplus P_3$ and $X_2 = P_2 \oplus P_2$ depicted in Figure 2. Using the preceding two techniques and the chromatic polynomials of K_2 and K_3, it is easy to verify that $\chi(X_1, x) = x^2(x - 1)^2 = \chi(X_2, x)$.

We say that two graphs G and H are *chromatically equivalent* if they have the same number of n-colorings, for every $n \geq 1$. This is equivalent to saying that G and H have the same chromatic polynomial. The two graphs X_1, X_2 in Example 2 are chromatically equivalent, a fact that will be used in what follows. Furthermore, all trees with the same number of vertices are pairwise chromatically equivalent, because if T is a tree with n vertices, then $\chi(T, x) = x(x - 1)^{n-1}$.

Chromatic equivalence is characterized by restricting the right-homomorphism vector to the class $\mathscr{K}$ of all cliques, that is, two graphs G and H are chromatically equivalent if and only if $\hom(G, \mathscr{K}) = \hom(H, \mathscr{K})$. In contrast, we will show that chromatic equivalence cannot be characterized by restricting the left-homomorphism vector.

Theorem 17. *There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that G and H are chromatically equivalent if and only if $\hom(\mathscr{F}, G) = \hom(\mathscr{F}, H)$.*

Before proving Theorem 17, we present two lemmas.

Lemma 1. *For every 2-colorable connected graph G with at least three vertices, we have that $\mathrm{sur}(G, P_3) > 0$.*

Proof. Let G be a 2-colorable connected graph with $|V(G)| \geq 3$. We argue that $\mathrm{sur}(G, P_3) > 0$.

For simplicity, in what follows we assume that K_2 is a graph with vertex set $V(K_2) = \{v_1, v_2\}$ and edge set $E(K_2) = \{(v_1, v_2)\}$, and that P_3 is a graph with $V(P_3) = \{v_1, v_2, v_3\}$ and $E(P_3) = \{(v_1, v_2), (v_2, v_3)\}$; in other words, K_2 is a subgraph of P_3.

If $|V(G)| = 3$, then G is isomorphic to P_3 and obviously $\mathrm{sur}(G, P_3) > 0$.

Assume now that $|V(G)| > 3$. Then G must contain an edge because it is connected. Moreover, we have $\mathrm{sur}(G, K_2) > 0$ as G is 2-colorable. Let $h : G \to K_2$ be a surjective homomorphism and, without loss of generality, assume that $|h^{-1}(v_1)| > 1$, i.e., more than

one vertex in $V(G)$ is mapped to $v_1 \in V(K_2)$ under h (note that at least one of $h^{-1}(v_1)$ and $h^{-1}(v_2)$ has cardinality > 1). Pick two distinct vertices $a_1, a_2 \in V(G)$ such that $h(a_1) = h(a_2) = v_1$; then there are two vertices $b_1, b_2 \in V(G)$ (not necessarily distinct) such that $(a_1, b_1), (a_2, b_2) \in E(G)$ (since G is connected) and $h(b_1) = h(b_2) = v_2$. It follows that the homomorphism $h' : G \to P_3$ such that $h'(u) = h(u)$ for $u \neq a_2$ and $h'(a_2) = v_3$ is surjective. We conclude that $\mathrm{sur}(G, P_3) > 0$. $\hfill\square$

Given any graph G, we define the *n-fold disjoint union* $G^{\oplus n}$ *of* G, where $n \geq 0$, inductively as follows:
$$G^{\oplus 0} \ := \ \emptyset; \ \ G^{\oplus n+1} \ := \ G^{\oplus n} \oplus G.$$
We also set $\mathrm{hom}(\emptyset, H) := 1$ for every graph H.

The next lemma tells which graphs G distinguish the graphs X_1 and X_2 in Example 2 via $\mathrm{hom}(G, X_1)$ and $\mathrm{hom}(G, X_2)$.

Lemma 2. *If G is a graph, then $\mathrm{hom}(G, X_1) \geq \mathrm{hom}(G, X_2)$. Furthermore, $\mathrm{hom}(G, X_1) = \mathrm{hom}(G, X_2)$ holds if and only if either G is not 2-colorable or G is a disjoint union of vertices and edges (in symbols, $G = I_m \oplus P_2^{\oplus n}$ for some $m, n \geq 0$ such that $m + n \geq 1$).*

Proof. Let G be a graph. We distinguish two cases.

Case 1: G is connected. If G is not 2-colorable, then $\mathrm{hom}(G, X_1) = 0 = \mathrm{hom}(G, X_2)$, since both X_1 and X_2 are 2-colorable. If $|V(G)| \leq 2$, then, since G is connected, we have that $G = I_1$ or $G = P_2$, from which it follows easily that $\mathrm{hom}(G, X_1) = 4 = \mathrm{hom}(G, X_2)$. So, assume that G is 2-colorable and $|V(G)| \geq 3$. We will show that, in this case, $\mathrm{hom}(G, X_1) > \mathrm{hom}(G, X_2)$. Since G is connected, it must have at least one edge. Then any homomorphic image of G must contain an edge and must be connected as well since the homomorphic image of a connected graph is connected. Thus, the calculation of $\mathrm{hom}(G, X_1)$ and of $\mathrm{hom}(G, X_2)$ can be broken down to the calculation of the number of surjective homomorphisms from G onto a connected subgraph of X_1 or X_2 that contains at least one edge; note that such a subgraph must be isomorphic to P_2 or P_3. Since X_1 has exactly two subgraphs isomorphic to P_2 and one subgraph isomorphic to P_3 whereas X_2 has exactly two subgraphs isomorphic P_2 and no subgraph isomorphic to P_3, we obtain $\mathrm{hom}(G, X_1) = 2\mathrm{sur}(G, P_2) + \mathrm{sur}(G, P_3)$ and $\mathrm{hom}(G, X_2) = 2\mathrm{sur}(G, P_2)$. By Lemma 1, we have that $\mathrm{sur}(G, P_3) > 0$. It follows that $\mathrm{hom}(G, X_1) > \mathrm{hom}(G, X_2)$.

Thus, if G is connected, then $\mathrm{hom}(G, X_1) \geq \mathrm{hom}(G, X_2)$, and the equality holds if and only if G is not 2-colorable or $G \in \{I_1, P_2\}$.

Case 2: G is not connected. Let $H_1, \ldots, H_k$, $k \geq 2$, be the connected components of G. Thus, $G = \oplus_{i=1}^{k} H_i$ and so $\mathrm{hom}(G, X_1) = \Pi_{i=1}^{k} \mathrm{hom}(H_i, X_1)$ and $\mathrm{hom}(G, X_2) = \Pi_{i=1}^{k} \mathrm{hom}(H_i, X_2)$. Since each H_i is connected, the previous case implies that $\mathrm{hom}(H_i, X_1) \geq \mathrm{hom}(H_i, X_2)$, for $1 \leq i \leq k$. Consequently, $\mathrm{hom}(G, X_1) \geq \mathrm{hom}(G, X_2)$. Furthermore, the equality holds if and only if G is not 2-colorable (i.e., at least one H_i is not 2-colorable) or every $H_i \in \{I_1, P_2\}$, which means that G is not 2-colorable or $G = I_m \oplus P_2^{\oplus k-m}$, for some m with $0 \leq m \leq k$. $\hfill\square$

We are now ready to prove Theorem 17.

Proof of Theorem 17. Towards a contradiction, suppose that there is a class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that G and H are chromatically equivalent if and only $\hom(\mathscr{F}, G) = \hom(\mathscr{F}, H)$. In particular, for the chromatically equivalent graphs X_1, X_2 in Example 2, we have that $\hom(F, X_1) = \hom(F, X_2)$, for every $F \in \mathscr{F}$. Lemma 2 implies that $\mathscr{F} \subseteq \mathscr{F}_1 \cup \mathscr{F}_2$, where

$$\begin{aligned} \mathscr{F}_1 &= \{G \mid G \text{ is not 2-colorable}\} \\ \mathscr{F}_2 &= \{I_m \oplus P_2^{\oplus n} \mid m, n \geq 0 \text{ and } m + n \geq 1\}. \end{aligned}$$

Let $G = C_8$ be the 8-cycle and let $H = C_4 \oplus C_4$ be the disjoint union of two 4-cycles. Then G and H are not chromatically equivalent because they have different chromatic polynomials. Indeed, $\chi(G, x) = (x - 1)^8 + (x - 1)$, while $\chi(H, x) = ((x - 1)^4 + (x - 1))^2$. However, no graph F in $\mathscr{F}_1 \cup \mathscr{F}_2$ distinguishes G and H via $\hom(F, G)$ and $\hom(F, H)$. The reason is that if F is not 2-colorable, then $\hom(F, G) = 0 = \hom(F, H)$, while if $F = I_m \oplus P_2^{\oplus n}$ for some $m, n \geq 0$ with $m + n \geq 1$, then $\hom(F, G) = 8^m 16^n = \hom(F, H)$. $\qquad\square$

6 Left- and Right-Homomorphism Vectors

In the previous two sections, we saw that several natural relaxations of isomorphism can be captured by restricting one of the two (the left or the right) homomorphism vectors, but not by the other. In this section, we shall see that other natural, logic-based relaxations of isomorphism cannot be captured by restricting either of the two homomorphism vectors.

Let $k \geq 2$ be a positive integer. Recall that FO^k is the fragment of first-order logic consisting of all FO-formulas with at most k distinct variables. Two graphs G and H are said to be FO^k-equivalent, denoted by $G \equiv_{\mathrm{FO}}^k H$, if G and H satisfy the same FO^k-sentences. The finite-variable logic FO^k, $k \geq 2$, have been studied extensively in finite model theory (see [Lib04, GKL$^+$07]); in particular, it is well known that FO^k-equivalence can be characterized in terms of the k-pebble game, $k \geq 2$. The next result tells that FO^k-equivalence cannot be captured by restricting the left-homomorphism vector or the right-homomorphism vector.

Proposition 1. Consider a positive integer $k \geq 2$.

1. There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that $G \equiv_{\mathrm{FO}}^k H$ if and only if $\hom(\mathscr{F}, G) = \hom(\mathscr{F}, H)$.

2. There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that $G \equiv_{\mathrm{FO}}^k H$ if and only if $\hom(G, \mathscr{F}) = \hom(H, \mathscr{F})$.

Proof. For the first part, suppose that such a class $\mathscr{F}$ exists. Pick a graph D in $\mathscr{F}$ and let c be the number of the vertices of D. Let $m = \max(c, k)$ and consider the cliques K_m and K_{m+1} with $V(K_m) = \{v_1, \ldots, v_m\}$ and $V(K_{m+1}) = \{v_1, \ldots, v_m, v_{m+1}\}$. Clearly, $K_m \equiv_{\mathrm{FO}}^m K_{m+1}$, hence $K_m \equiv_{\mathrm{FO}}^k K_{m+1}$ because $k \leq m$. By the assumption about the class $\mathscr{F}$, we have $\hom(\mathscr{F}, K_m) = \hom(\mathscr{F}, K_{m+1})$; in particular, we have $\hom(D, K_m) = \hom(D, K_{m+1})$. Furthermore, $\hom(D, K_m) > 0$, since $m \geq c = |V(D)|$ and there is an injective homomorphism

18

from D to K_m. Therefore, $\hom(D, K_m) = \hom(D, K_{m+1}) > 0$. This, however, is a contradiction, since $\hom(D, K_{m+1}) > \hom(D, K_m)$. Indeed, first note that $\mathrm{Hom}(D, K_m) \subseteq \mathrm{Hom}(D, K_{m+1})$. Next, given a homomorphism $h : D \to K_m$, we take the least-indexed vertex in the image $h(D)$, say v_r (note that $r \leq m$), and substitute v_{m+1} in h for v_r to obtain a homomorphism $h' : D \to K_{m+1}$ such that for every $u \in V(D)$, we have that $h'(u) = v_{m+1}$ if $u \in h^{-1}(v_r)$, and $h'(u) = h(u)$ if $u \in V(D) \setminus h^{-1}(v_r)$.

Obviously, $h' \notin \hom(D, K_m)$, which proves the first part.

For the second part, suppose that such a class $\mathscr{F}$ exists. It cannot be the case $\mathscr{F} = \{I_1\}$. Otherwise, we would have $\hom(I_1, \mathscr{F}) = \hom(I_2, \mathscr{F})$ which, by assumption, would imply $I_1 \equiv^k_{\mathrm{FO}} I_2$; this is absurd because $I_1 \models \forall x \forall y\, x = y$, while $I_2 \not\models \forall x \forall y\, x = y$. Therefore, $\mathscr{F}$ contains a graph D with at least two vertices. Let $c = |V(D)| \geq 2$ and let $m = \max(c, k)$. Consider two independent sets, I_m and I_{m+1}. Obviously, $I_m \equiv^m_{\mathrm{FO}} I_{m+1}$ and hence $I_m \equiv^k_{\mathrm{FO}} I_{m+1}$ since $m \geq k$. The assumption about $\mathscr{F}$ implies that $\hom(I_m, \mathscr{F}) = \hom(I_{m+1}, \mathscr{F})$. Therefore, $\hom(I_m, D) = \hom(I_{m+1}, D)$, which is a contradiction because $c = |V(D)| \geq 2$ and $\hom(I_m, D) = c^m < c^{m+1} = \hom(I_{m+1}, D)$. $\qquad\square$

The *quantifier-depth* of a first-order formula is a positive integer that measures the nesting of quantifiers in that formula; it gives rise to a parametrization of first-order logic that is different from the parametrization according to the number of distinct formulas. Let k be a positive integer and let G and H be two graphs. We write $G \equiv^k_{\mathrm{QR}} H$ to denote that G and H satisfy the same FO-sentences of quantifier rank at most k. It is well known that the equivalence relation $\equiv^k_{\mathrm{QR}}$ is characterized in terms of the k-move Ehrenheufcht-Fraïssé game (see [Lib04, GKL$^+$07]). The next result tells that $\equiv^k_{\mathrm{QR}}$ cannot be captured by restricting the left-homomorphism vector or the right-homomorphism vector. The proof is omitted because it is essentially the same as that of Proposition 1.

Proposition 2. Consider a positive integer $k \geq 2$.

1. There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that $G \equiv^k_{\mathrm{QR}} H$ if and only if $\hom(\mathscr{F}, G) = \hom(\mathscr{F}, H)$.

2. There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that $G \equiv^k_{\mathrm{QR}} H$ if and only if $\hom(G, \mathscr{F}) = \hom(H, \mathscr{F})$.

7   Extensions and Discussion

Up to this point, we have focused on graphs that have no self-loops, no multi-edges, and no weights on the vertices or the edges; furthermore, the computation of the homomorphism count uses integer arithmetic only. In this section, we discuss two extensions of the framework: the first involves graphs with loops and real numbers as weights on the vertices and the edges, while the second involves defining the homomorphism count over arbitrary semirings.

7.1 Extension to Graphs with Loops and Real Weights

There is substantial literature on extended notions of homomorphism counts that involve *weighted* graphs, that is, undirected graphs with self-loops and weights on each vertex and each edge (including self-loops), but no multi-edges. In particular, Lovász defines the notion of the homomorphism count $\hom(G, H)$, where G is an ordinary graph (no self-loops, no multi-edges, no weights) and H is a weighted graph [Lov12]. Specifically, assume that $H = (V(H), E(H), w)$ is a weighted graph, where w is a real-valued weight function defined on $V(H) \cup E(H)$. If $G = (V(G), E(G))$ is a graph, then a *homomorphism* from G to H is a homomorphism from G to the (unweighted) looped graph $H' = (V(H), E(H))$ (note that here an edge of G can be mapped to a self-loop of H'). Let $\mathrm{Hom}(G, H')$ denote the set of all homomorphisms from G to H' and define the homomorphism count $\hom(G, H)$ as

$$\hom(G, H) = \sum_{h \in \mathrm{Hom}(G, H')} \prod_{u \in V(G)} w(h(u)) \prod_{e \in E(G)} w(h(e))$$

The earlier definition of $\hom(G, H)$ between two graphs G and H is the special case of this in which H is turned into a weighted graph with weight 1 on each vertex and each edge.

Several important graph invariants can be expressed as graph polynomials. In Section 5, we encountered such a graph polynomial, namely, the chromatic polynomial $\chi(G, k)$, which gives the number of the k-colorings of G. It is known that several other fundamental graph polynomials can be expressed using the preceding extended notion of homomorphism counts (for an overview, see [Lov12, Section 5.3]). Here, we discuss two such polynomials.

The *cluster expansion polynomial* $\mathrm{CEP}(G; x, y)$ of a graph G is a bivariate polynomial that, among other things, generalizes the chromatic polynomial. If G is a graph, then, by definition,

$$\mathrm{CEP}(G; x, y) = \sum_{A \subseteq E(G)} x^{c(A)} y^{|A|},$$

where $c(A)$ is the number of connected components of the graph $(V(G), A)$ and $|A|$ is the cardinality of the set A. It can be shown that $\mathrm{CEP}(G; k, -1) = \chi(G, k)$. More importantly, the cluster expansion polynomial is a version of the Tutte polynomial [Tut04], which is arguably the most fundamental graph polynomial as it encapsulates a great deal of information about the graph with which it is associated.

The cluster expansion polynomial can be expressed in terms of the homomorphism count between graphs and certain weighted graphs (see [Lov12, Section 5.3])). Specifically, for every $k \geq 1$ and every real y, let $K_{k,y}$ be the clique on k vertices with a self-loop added at every vertex and with the following weights: every vertex has weight 1, every self-loop has weight $1 + y$, and every other edge has weight 1. Then it can be proved that $\mathrm{CEP}(G; k, y) = \hom(G, K_{k,y})$. It follows that if $\mathscr{F}$ is the class of all weighted graphs of the form $K_{k,y}$, then the right-homomorphism vector restricted to $\mathscr{F}$ captures the equivalence relation "the graphs G and H have the same cluster expansion polynomial".

The *independence polynomial* $I(G; x, y)$ of a graph G is a bivariate polynomial that encapsulates information about the independent sets of G. If G is a graph, then let $\mathscr{I}(G)$

denote the collection of its independent sets, and define

$$I(G; x, y) = \sum_{U \in \mathscr{I}(G)} x^{|U|} y^{|V(G) \setminus U|}.$$

Note that

$$I(G; x, 1) = \sum_{U \in \mathscr{I}(G)} x^{|U|},$$

which is the univariate independence polynomial introduced by Gutman and Harary [GH83]. Like the cluster expansion polynomial, the independence polynomial can be expressed in terms of the homomorphism count between graphs and certain weighted graphs (see [GGN11, Section 4.2]). Specifically, let $L = (V(L), E(L))$ be the "lollipop" graph, i.e., the graph with two vertices a, b, an edge (a, b) and a self-loop (b, b). For every two real numbers x and y, let $L_{x,y}$ be the weighted graph obtained from L by putting x as the weight of the vertex a, putting y as the weight of the vertex b, and putting 1 as the weight of the edge (a, b) and the self-loop (b, b). Then it can be proved that $I(G; x, y) = \hom(G, L_{x,y})$. It follows that if $\mathscr{F}$ is the class of all weighted graphs of the form $L_{x,y}$, then the right-homomorphism vector restricted to $\mathscr{F}$ captures the equivalence relation "the graphs G and H have the same independence polynomial".

The preceding discussion suggests that the investigation we embarked on here should be expanded to an investigation of the expressive power of homomorphism vectors restricted to classes of weighted graphs, as such vectors give rise to a variety of equivalence relations arising from graph polynomials.

7.2   Extension to Arbitrary Semirings

Let G be a graph and let H be a weighted graph. The expression defining the homomorphism count $\hom(G, H)$ is a sum of products of real numbers. This sum of products is also meaningful over an arbitrary semiring $\mathbb{K} = (K, +, \times, 0, 1)$, where $+$ and $\times$ are the addition and multiplication operations on K, and 0 and 1 are the identity elements of $+$ and $\times$. Thus, we can define the homomorphism count $\hom_{\mathbb{K}}(G, H)$ for a graph G and a $\mathbb{K}$-weighted graph H, where the weight function takes values in the universe K of an arbitrary, but fixed, semiring $\mathbb{K}$. For example, the standard homomorphism count $\hom(G, H)$, where G and H are graphs, coincides with the homomorphism count $\hom_{\mathbb{N}}(G, H)$, where $\mathbb{N} = (\{0, 1, 2, \ldots\}, +, \times, 0, 1)$ is the *bag* semiring of the non-negative integers, and H is viewed as a $\mathbb{N}$-weighted graph with weight 1 on each vertex and each edge. And, of course, the discussion in the preceding section is about the homomorphism count $\hom_{\mathbb{R}}(G, H)$, where $\mathbb{R} = (R, +, \times, 0, 1)$ is the semiring (actually, the field) of the real numbers.

These considerations pave the way for a further expansion of the framework to homomorphism counts with respect to some arbitrary, but fixed, semiring. As a concrete case in point, consider the *Boolean* semiring $\mathbb{B} = (\{0, 1\}, \vee, \wedge, 0, 1)$, which has disjunction $\vee$ and conjunction $\wedge$ as operations, and 0 (false) and 1 (true) as the identity elements of $\vee$ and $\wedge$. Let G be a graph and let H be a $\mathbb{B}$-graph with weight 1 on each vertex and each edge.

Then $\hom_{\mathbb{B}}(G, H)$ is the *sign* of the standard homomorphism count $\hom(G, H)$ indicating the existence or non-existence of a homomorphism from G to H, that is, $\hom_{\mathbb{B}}(G, H) = 1$ if there is a homomorphism from G to H, and $\hom_{\mathbb{B}}(G, H) = 0$, otherwise.

Let G be a graph and let $\mathscr{F}$ be a class of graphs. Put

$$\begin{aligned}
\hom_{\mathbb{B}}(\mathscr{F}, G) &:= (\hom_{\mathbb{B}}(D, G) \mid D \in \mathscr{F}), \\
\hom_{\mathbb{B}}(G, \mathscr{F}) &:= (\hom_{\mathbb{B}}(G, D) \mid D \in \mathscr{F})
\end{aligned}$$

for the left-homomorphism and the right-homomorphism vectors of G restricted to $\mathscr{F}$, respectively.

Recall that $\mathscr{G}$ is the class of all graphs. Obviously, for every graph, there is a homomorphism from that graph to itself. Therefore, for every two graphs G and H, we have that the following statements are equivalent:

1. $\hom_{\mathbb{B}}(\mathscr{G}, G) = \hom_{\mathbb{B}}(\mathscr{G}, H)$.

2. $\hom_{\mathbb{B}}(G, \mathscr{G}) = \hom_{\mathbb{B}}(H, \mathscr{G})$.

3. G and H are homomorphically equivalent.

By definition, G and H are *homomorphically equivalent* if there are homomorphisms from G to H, and from H to G. The notion of homomorphic equivalence plays an important role in several different areas, including database theory and constraint satisfaction. Furthermore, homomorphic equivalence coincides with isomorphism on graphs that are cores (see [HN04] for detailed information about these notions).

The *chromatic number* $\chi(G)$ of a graph G is the smallest positive integer k such that G has a k-coloring. Clearly, G has a k-coloring if and only if $\hom(G, K_k) > 0$; furthermore, if $\hom(G, K_k) > 0$, then $\hom(G, K_m) > 0$, for all $m > k$. Thus, two graphs G and H have the same chromatic number if and only if $\hom_{\mathbb{B}}(G, \mathscr{K}) = \hom_{\mathbb{B}}(H, \mathscr{K})$. In contrast, the next result asserts that the equivalence relation of two graphs having the same chromatic number cannot be captured by restricting the left-homomorphism vector $\hom_{\mathbb{B}}(\mathscr{G}, G)$.

Proposition 3. There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that $\chi(G) = \chi(H)$ if and only if $\hom_{\mathbb{B}}(\mathscr{F}, G) = \hom_{\mathbb{B}}(\mathscr{F}, H)$.

Proof. Towards a contradiction, assume such a class $\mathscr{F}$ exists.

Consider the cycles C_{2k} of even length, $k \geq 2$. They have chromatic number 2, hence for all $k, m \geq 2$, we have $\hom_{\mathbb{B}}(\mathscr{F}, C_{2k}) = \hom_{\mathbb{B}}(\mathscr{F}, C_{2m})$. Next, consider the cycles C_{2k+1} of odd length, $k \geq 1$. They have chromatic number 3, hence $\hom_{\mathbb{B}}(F, C_{2k+1}) = \hom_{\mathbb{B}}(\mathscr{F}, C_{2m+1})$, for all $k, m \geq 1$. We will show that $\hom_{\mathbb{B}}(\mathscr{F}, C_{2k}) = \hom_{\mathbb{B}}(\mathscr{F}, C_{2k+1})$, for $k \geq 2$, which will be a contradiction, since C_{2k} and C_{2k+1} have different chromatic numbers.

Fix a $k \geq 2$ and let D be an arbitrary graph in $\mathscr{F}$.

Case 1. D is a 2-colorable graph. In this case, we have that $\hom_{\mathbb{B}}(D, C_{2k}) = 1 = \hom_{\mathbb{B}}(D, C_{2k+1})$.

Case 2. D is a non-2-colorable graph. It follows that D contains an odd cycle C_{2l+1}, for some $l \geq 1$. Pick a cycle C_{2n} with $2n > 2l + 1$. Then $\hom_{\mathbb{B}}(D, C_{2n}) = 0$, since

homomorphisms map odd cycles to odd cycles of smaller or equal length, but C_{2n} contains no such cycle. Therefore, $\hom_{\mathbb{B}}(D, C_{2k}) = 0$ as well. Next, pick a cycle C_{2n+1} with $n > l$. By the same reasoning, $\hom_{\mathbb{B}}(D, C_{2n+1}) = 0$, and so $\hom_{\mathbb{B}}(D, C_{2k+1}) = 0$. We now conclude that for every $k \geq 2$, we have that $\hom_{\mathbb{B}}(D, C_{2k}) = \hom_{\mathbb{B}}(D, C_{2k+1})$, for every $D \in \mathscr{F}$. Thus, $\hom_{\mathbb{B}}(\mathscr{F}, C_{2k}) = \hom_{\mathbb{B}}(\hom(\mathscr{F}, C_{2k+1})$, a contradiction. $\qquad\square$

The *clique number* $\omega(G)$ of a graph G is the largest positive integer k such that G has a k-clique as a subgraph. Clearly, two graphs G and H have the same clique number if and only if $\hom_{\mathbb{B}}(\mathscr{K}, G) = \hom_{\mathbb{B}}(\mathscr{K}, H)$. In contrast, we show that the clique number cannot be captured by any restriction of the right-homomorphism vector over the Boolean semiring. The proof uses the classical result of Erdös that there exist graphs of arbitrarily large chromatic number and, simultaneously, arbitrarily large girth; see, e.g., Corollary 3.13 in [HN04].

Proposition 4. There is no class $\mathscr{F}$ of graphs such that for every two graphs G and H, we have that $\omega(G) = \omega(H)$ if and only if $\hom_{\mathbb{B}}(G, \mathscr{F}) = \hom_{\mathbb{B}}(H, \mathscr{F})$.

Proof. Towards a contradiction, assume such a class $\mathscr{F}$ exists. Pick any graph D in $\mathscr{F}$ and let $m = |V(D)|$. The graph D cannot be an independent set since otherwise $\hom_{\mathbb{B}}(K_2, D) = \hom_{\mathbb{B}}(K_3, D) = 0$ but $\omega(K_2) \neq \omega(K_3)$. Let G be any graph of chromatic number larger than m and girth larger than 3. In particular $\hom_{\mathbb{B}}(G, D) = 0$ and $\omega(G) = 2$. Now take $F = D \times K_2$, i.e., $V(F) = V(D) \times \{0, 1\}$ and there is an edge between (u, a) and (v, b) in F (where $u, v \in V(D)$ and $a, b \in \{0, 1\}$) if and only if (u, v) is an edge in F and $a \neq b$. The projections into each component are homomorphisms into D and K_2, respectively. Hence $\hom_{\mathbb{B}}(F, D) \neq 0$ and F is 2-colorable. In particular F is triangle-free and, since D is not an independent set, $\omega(F) = 2$. We have shown that $\omega(F) = \omega(G) = 2$ but $\hom_{\mathbb{B}}(F, D) \neq \hom_{\mathbb{B}}(G, D)$. $\qquad\square$

It is obvious that $\omega(G) \leq \chi(G)$. The inequality is strict for the odd cycles (since $\omega(C_{2n+1}) = 2$ and $\chi(C_{2n+1}) = 3$) and for many other prominent graphs, such as the Petersen graph. There is a vast literature on graph parameters that interpolate between the clique number and the chromatic number. A well-known such parameter is the *fractional chromatic number* of a graph G, denoted by $\chi_f(G)$, which arises in the study of fractional colorings [Gel76, Sta76]. By definition, the quantity $\chi_f(G)$ is the optimum of a linear program that has one real variable x_U for each independent set $U \in \mathscr{I}(G)$ of G and is defined as follows:

$$
\begin{aligned}
\min \quad & \textstyle\sum_{U \in \mathscr{I}(G)} x_U \\
\text{s.t.} \quad & \textstyle\sum_{U \in \mathscr{I}(G):v \in U} x_U \geq 1 \quad && \text{for all } v \in V(G), \\
& x_U \geq 0 \quad && \text{for all } U \in \mathscr{I}(G).
\end{aligned}
$$

It is known that $\omega(G) \leq \chi_f(G) \leq \chi(G)$ and, again, the inequalities can be strict. In particular, the inequalities are strict for odd cycles of length at least 5, since $\chi_f(C_{2n+1}) = 2 + \frac{1}{n}$.

The fractional chromatic number of a graph can also be characterized combinatorially. We write $K_{a:b}$ to denote the *Kneser* graph with parameters a and b, where $a \geq 2b$. This is the graph whose vertices are the b-element subsets of $\{1, \ldots, a\}$, and where two b-element subsets

are joined by an edge if the sets are disjoint. It is known that $\chi_f(G)$ is the smallest rational number a/b for which there is a homomorphism from G to the Kneser graph $K_{a:b}$. It follows that two graphs G and H satisfy $\chi_f(G) = \chi_f(H)$ if and only if $\hom_\mathbb{B}(G, \mathscr{K}_f) = \hom_\mathbb{B}(G, \mathscr{K}_f)$, where $\mathscr{K}_f$ denotes the class of all Kneser graphs. We refer the reader to Chapter 6 of [HN04] for the fascinating interplay between graph coloring theory and the Kneser graphs.

Interestingly, the fractional chromatic number of a graph G has a dual, called the *fractional clique number* of G, denoted by $\omega_f(G)$. By definition, this is the optimum of the *dual* of the linear program that defines $\chi_f(G)$. Concretely, $\omega_f(G)$ is the optimum of

$$
\begin{array}{ll}
\max & \sum_{v \in V(G)} y_v \\
\text{s.t.} & \sum_{v \in U} y_v \leq 1 \quad \text{for all } U \in \mathscr{I}(G), \\
& y_v \geq 0 \qquad\qquad \text{for all } v \in V(G).
\end{array}
$$

By the Duality Theorem of Linear Programming, it holds that $\omega_f(G) = \chi_f(G)$. In view of this, and of the characterization of χ_f in terms of the sign of the right-homomorphism vector $\hom_\mathbb{B}(\cdot, \mathscr{K}_f)$ restricted to Kneser graphs, it would be interesting to know whether the equivalence relation of having the same fractional clique number can be captured as the restriction of a left-homomorphism vector. We leave this as an open problem.

8   Concluding Remarks

In this paper, we investigated relaxations of graph isomorphism obtained by restricting the left-homomorphism vector and the right-homomorphism vector to a class of graphs. We showed that these two types of restrictions have incomparable expressive power. In particular, we established a number of results to the effect that certain natural relaxations of graph isomorphism *cannot* be obtained by restricting one of these two homomorphism vectors.

The work reported here motivates several different directions for future research, including investigating extensions of this framework to weighted graphs over semirings, as discussed in Section 7. One of the ultimate goals is to characterize the relaxations of graph isomorphism that are captured by restrictions of the left-homomorphism vector or by restrictions of the right-homomorphism vector. An important result in this vein is the Freedman-Lovász-Schrijver Theorem [FLS07], which characterizes the *graph parameters* (invariant numbers) of a graph G that are equal to $\hom(G, F)$ for some fixed weighted graph F. Much more remains to be done to obtain such characterizations for restrictions of the left-homomorphism vector or the right-homomorphism vector to an arbitrary class $\mathscr{F}$ of graphs (not just to a single graph F).

References

[ABD09]   Albert Atserias, Andrei A. Bulatov, and Anuj Dawar. Affine systems of equations and counting infinitary logic. *Theor. Comput. Sci.*, 410(18):1666–1683, 2009.

[Bab16]   László Babai. Graph isomorphism in quasipolynomial time [extended abstract]. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 684–697. ACM, 2016.

[BCGR19] Jan Böker, Yijia Chen, Martin Grohe, and Gaurav Rattan. The complexity of homomorphism indistinguishability. In Peter Rossmanith, Pinar Heggernes, and Joost-Pieter Katoen, editors, *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.

[Bir12]   George D Birkhoff. A determinant formula for the number of ways of coloring a map. *The Annals of Mathematics*, 14(1/4):42–46, 1912.

[BJK05]   Andrei A. Bulatov, Peter Jeavons, and Andrei A. Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM J. Comput.*, 34(3):720–742, 2005.

[Bul05]   Andrei A. Bulatov. H-coloring dichotomy revisited. *Theor. Comput. Sci.*, 349(1):31–39, 2005.

[CFI92]   Jin-yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identifications. *Comb.*, 12(4):389–410, 1992.

[CV93]   Surajit Chaudhuri and Moshe Y. Vardi. Optimization of *Real* conjunctive queries. In Catriel Beeri, editor, *Proceedings of the Twelfth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, May 25-28, 1993, Washington, DC, USA*, pages 59–70. ACM Press, 1993.

[Daw98]   Anuj Dawar. A restricted second order logic for finite structures. *Inf. Comput.*, 143(2):154–174, 1998.

[DGR18]   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.

[DSZ17]   Anuj Dawar, Simone Severini, and Octavio Zapata. Pebble games and cospectral graphs. *Electronic Notes in Discrete Mathematics*, 61:323–329, 2017.

[FLS07]   Michael Freedman, László Lovász, and Alexander Schrijver. Reflection positivity, rank connectivity, and homomorphism of graphs. *Journal of the American Mathematical Society*, 20(1):37–51, 2007.

[Gel76]   Dennis P Geller. r-tuple colorings of uniquely colorable graphs. *Discrete Mathematics*, 16(1):9–12, 1976.

[GGN11]   Delia Garijo, Andrew J. Goodall, and Jaroslav Nešetřil. Distinguishing graphs by their left and right homomorphism profiles. *Eur. J. Comb.*, 32(7):1025–1053, 2011.

[GH83]   Ivan Gutman and Frank Harary. Generalizations of the matching polynomial. *Utilitas Mathematica*, 24(1):97–106, 1983.

[GKL$^+$07]   Erich Grädel, Phokion G. Kolaitis, Leonid Libkin, Maarten Marx, Joel Spencer, Moshe Y. Vardi, Yde Venema, and Scott Weinstein. *Finite Model Theory and Its Applications*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2007.

[GM99]   Martin Grohe and Julian Mariño. Definability and descriptive complexity on databases of bounded tree-width. In Catriel Beeri and Peter Buneman, editors, *Database Theory - ICDT '99, 7th International Conference, Jerusalem, Israel, January 10-12, 1999, Proceedings*, volume 1540 of *Lecture Notes in Computer Science*, pages 70–82. Springer, 1999.

[Gro20]   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.

[HN90]   Pavol Hell and Jaroslav Nešetřil. On the complexity of H-coloring. *J. Comb. Theory, Ser. B*, 48(1):92–110, 1990.

[HN04]   Pavol Hell and Jaroslav Nešetřil. *Graphs and homomorphisms*, volume 28 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, 2004.

[KV00]   Phokion G. Kolaitis and Moshe Y. Vardi. Conjunctive-query containment and constraint satisfaction. *J. Comput. Syst. Sci.*, 61(2):302–332, 2000.

[Lib04]   Leonid Libkin. *Elements of Finite Model Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004.

[Lov67]   László Lovász. Operations with structures. *Acta Mathematica Academiae Scientiarum Hungarica*, 18(3-4):321–328, 1967.

[Lov12]   László Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.

[Noy03]   Marc Noy. Graphs determined by polynomial invariants. *Theor. Comput. Sci.*, 307(2):365–384, 2003.

[Ott17]   Martin Otto. *Bounded Variable Logics and Counting: A Study in Finite Models*, volume 9 of *Lecture Notes in Logic*. Cambridge University Press, 2017.

[RC77]   Ronald C Read and Derek G Corneil. The graph isomorphism disease. *Journal of Graph Theory*, 1(4):339–363, 1977.

[Rea68]   Ronald C Read. An introduction to chromatic polynomials. *Journal of Combinatorial Theory*, 4(1):52–71, 1968.

[Sta76]   Saul Stahl. n-tuple colorings and associated graphs. *Journal of Combinatorial Theory, Series B*, 20(2):185–203, 1976.

[Tin86]   Gottfried Tinhofer. Graph isomorphism and theorems of birkhoff type. *Computing*, 36(4):285–300, 1986.

[Tin91]   Gottfried Tinhofer. A note on compact graphs. *Discret. Appl. Math.*, 30(2-3):253–264, 1991.

[Tut04]   WT Tutte. Graph-polynomials. *Advances in Applied Mathematics*, 32(1-2):5–9, 2004.

[VDH03]   Edwin R Van Dam and Willem H Haemers. Which graphs are determined by their spectrum? *Linear Algebra and its Applications*, 373:241–272, 2003.

[Wei06]   Boris Weisfeiler. *On construction and identification of graphs*, volume 558. Springer, 2006.