On the Weisfeiler-Leman Dimension of Fractional Packing

Vikraman Arvind[1], Frank Fuhlbrück[2(✉)], Johannes Köbler[2], and Oleg Verbitsky[2]

[1] The Institute of Mathematical Sciences (HBNI), Chennai, India
arvind@imsc.res.in

[2] Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
{fuhlbfra,koebler,verbitsky}@informatik.hu-berlin.de

Abstract. The k-dimensional Weisfeiler-Leman procedure (k-WL) has proven to be immensely fruitful in the algorithmic study of Graph Isomorphism. More generally, it is of fundamental importance in understanding and exploiting symmetries in graphs in various settings. Two graphs are k-WL-equivalent if dimention k does not suffice to distinguish them. 1-WL-equivalence is known as fractional isomorphism of graphs, and the k-WL-equivalence relation becomes finer as k increases.

We investigate to what extent standard graph parameters are preserved by k-WL-equivalence, focusing on fractional graph packing numbers. The integral packing numbers are typically NP-hard to compute, and we discuss applicability of k-WL-invariance for estimating the integrality gap of the LP relaxation provided by their fractional counterparts.

Keywords: Computational complexity · The Weisfeiler-Leman algorithm · Fractional packing

1 Introduction

The 1-dimensional version of the Weisfeiler-Leman procedure is the classical *color refinement* applied to an input graph G. Each vertex of G is initially colored by its degree. The procedure refines the color of each vertex $v \in V(G)$ in rounds, using the multiset of colors of vertices u in the neighborhood $N(v)$ of the vertex v. In the 2-dimensional version [25], all vertex pairs $xy \in V(G) \times V(G)$ are classified by a similar procedure of coloring them in rounds. The extension of this procedure to a classification of all k-tuples of G is due to Babai (see historical overview in [4,5]) and is known as the *k-dimensional Weisfeiler-Leman procedure*, abbreviated as k-WL. Graphs G and H are said to be k-WL *-equivalent* (denoted $G \equiv_{k\text{-WL}} H$) if they are indistinguishable by k-WL.

The WL Invariance of Graph Parameters. Let π be a *graph parameter*. By definition, $\pi(G) = \pi(H)$ whenever G and H are isomorphic (denoted $G \cong H$).

O. Verbitsky was supported by DFG grant KO 1053/8–1. He is on leave from the IAPMM, Lviv, Ukraine.

A. Leporati et al. (Eds.): LATA 2020, LNCS 12038, pp. 357–368, 2020.
https://doi.org/10.1007/978-3-030-40608-0_25

We say that π is a k-WL *-invariant* graph parameter if the equality $\pi(G) = \pi(H)$ is implied even by the weaker condition $G \equiv_{k\text{-WL}} H$. The smallest such k will be called the *Weisfeiler-Leman (WL) dimension* of π.

If no such k exists, we say that the WL dimension of π is *unbounded.* Knowing that a parameter π has unbounded WL dimension is important because this implies that π cannot be computed by any algorithm expressible in fixed-point logic with counting (FPC), a robust framework for study of *encoding-invariant (*or *"choiceless") computations*; see the survey [7].

The focus of our paper is on graph parameters with *bounded* WL dimension. If π is the indicator function of a graph property $\mathcal{P}$, then k-WL-invariance of π precisely means that $\mathcal{P}$ is definable in the infinitary $(k+1)$-variable counting logic $C^{k+1}_{\infty\omega}$. While minimizing the number of variables is a recurring theme in descriptive complexity, our interest in the study of k-WL-invariance has an additional motivation: If we know that a graph parameter π is k-WL-invariant, this gives us information not only about π but also *about* k-WL. For example, the largest eigenvalue of the adjacency matrix has WL dimension 1 (see [24]), and the whole spectrum of a graph has WL dimension 2 (see [8,13]), which implies that 2-WL subsumes distinguishing non-isomorphic graphs by spectral methods.

Fractional Graph Parameters. In this paper, we mainly consider *fractional* graph parameters. Algorithmically, a well-known approach to tackling intractable optimization problems is to consider an appropriate linear programming (LP) relaxation. Many standard integer-valued graph parameters have fractional real-valued analogues, obtained by LP-relaxation of the corresponding 0–1 linear program; see, e.g., the monograph [24]. The fractional counterpart of a graph parameter π is denoted by π_f. While π is often hard to compute, π_f provides, sometimes quite satisfactory, a polynomial-time computable approximation of π.

The WL dimension of a natural fractional parameter π_f is a priori bounded, where *natural* means that π_f is determined by an LP which is logically interpretable in terms of an input graph G. A striking result of Anderson, Dawar, Holm [1] says that the optimum value of an interpretable LP is expressible in FPC. It follows from the known immersion of FPC into the finite-variable infinitary counting logic $C^{\omega}_{\infty\omega} = \bigcup_{k=2}^{\infty} C^{k}_{\infty\omega}$ (see [21]), that each such π_f is k-WL-invariant for some k. While this general theorem is applicable to many graph parameters of interest, it is not easy to extract an explicit value of k from this argument, and in any case such value would hardly be optimal.

We are interested in *explicit* and, possibly, *exact* bounds for the WL dimension. A first question here would be to pinpoint which fractional parameters π_f are 1-WL-invariant. This natural question, using the concept of fractional isomorphism [24], can be recast as follows: Which *fractional* graph parameters are invariant under *fractional* isomorphism? It appears that this question has not received adequate attention in the literature. The only earlier result we could find is the 1-WL-invariance of the fractional domination number γ_f shown in the Ph.D. thesis of Rubalcaba [23].

We show that the fractional matching number ν_f is also a fractional parameter preserved by fractional isomorphism. Indeed, the matching number is an instance of the *F-packing number* π^F of a graph, corresponding to $F = K_2$. Here and throughout, we use the standard notation K_n for the complete graphs, P_n for the path graphs, and C_n for the cycle graph on n vertices. In general, $\pi^F(G)$ is the maximum number of vertex-disjoint subgraphs F' of G that are isomorphic to the fixed pattern graph F. While the matching number is computable in polynomial time, computing π^F is NP-hard whenever F has a connected component with at least 3 vertices [19], in particular, for $F \in \{P_3, K_3\}$. Note that K_3-packing is the optimization version of the archetypal NP-complete problem Partition Into Triangles [14, GT11]. We show that the fractional P_3-packing number $\nu_f^{P_3}$, like $\nu_f = \pi_f^{K_2}$, is 1-WL-invariant, whereas the WL dimension of the fractional triangle packing is 2.

In fact, we present a general treatment of fractional F-packing numbers π_f^F. We begin in Sect. 2 with introducing a concept of equivalence between two linear programs L_1 and L_2 ensuring that equivalent L_1 and L_2 have equal optimum values. Next, in Sect. 3, we consider the standard optimization versions of Set Packing and Hitting Set [14, SP4 and SP8], two of Karp's 21 NP-complete problems. These two generic problems generalize F-Packing and Dominating Set respectively. Their fractional versions have thoroughly been studied in hypergraph theory [12,20]. We observe that the LP relaxations of Set Packing (or Hitting Set) are equivalent whenever the incidence graphs of the input set systems are 1-WL-equivalent. This general fact readily implies Rubalcaba's result [23] on the 1-WL-invariance of the fractional domination number and also shows that, if the pattern graph F has ℓ vertices, then the fractional F-packing number π_f^F is k-WL-invariant for some $k < 2\,\ell$. This bound for k comes from a logical definition of the instance of Set Packing corresponding to F-Packing in terms of an input graph G (see Corollary 6). Though the bound is quite decent, it does not need to be optimal. We elaborate on a more precise bound, where we need to use additional combinatorial arguments even in the case of the fractional matching. We present a detailed treatment of the fractional matching in this exposition (Theorem 4), while the proof of our general result on the fractional F-packing numbers (Theorem 5), which includes the aforementioned cases of $F = K_3, P_3$, is postponed to the full version of the paper [2].

The *edge-disjoint* version of F-Packing is another problem that has intensively been studied in combinatorics and optimization. Since it is known to be NP-hard for any pattern F containing a connected component with at least 3 edges [10], fractional relaxations have received much attention in the literature [17,26]. We show that our techniques work well also in this case. In particular, the WL dimension of the fractional edge-disjoint triangle packing number $\rho_f^{K_3}$ is 2 (Theorem 7).

Integrality Gap via Invariance Ratio. Furthermore, we discuss the *approximate* invariance of *integral* graph parameters expressible by integer linear programs. For a first example, recall Lovász's inequality [12, Theorem 5.21] $\nu_f(G) \le \frac{3}{2}\,\nu(G)$. As ν_f is 1-WL-invariant, it follows that $\nu(G)/\nu(H) \le 3/2$

for any pair of nonempty 1-WL-equivalent graphs G and H. This bound is tight, as seen for the 1-WL-equivalent graphs $G = C_{6s}$ and $H = 2s\,C_3$. Consequently, the above relation between $\nu(G)$ and $\nu_f(G)$ is also tight. This simple example demonstrates that knowing, first, the exact value k of the WL dimension of a fractional parameter π_f and, second, the discrepancy of the integral parameter π over k-WL-invariant graphs implies a lower bound for the precision of approximating π by π_f.

Specifically, recall that the maximum $\max_G \frac{\pi_f(G)}{\pi(G)}$, (respectively $\max_G \frac{\pi(G)}{\pi_f(G)}$ for minimization problems) is known as the *integrality gap* of π_f. The integrality gap is important for a computationally hard graph parameter π, as it bounds how well the polynomial-time computable parameter π_f approximates π.

On the other hand, we define the k-WL-*invariance ratio* for the parameter π as $\max_{G,H} \frac{\pi(G)}{\pi(H)}$, where the quotient is maximized over all k-WL-equivalent graph pairs (G, H). If π is k-WL-invariant, then the k-WL-invariance ratio bounds the integrality gap from below. The following question suggests itself: How tight is this lower bound? In this context, we now consider the fractional domination number γ_f.

A general bound by Lovász [20] on the integrality gap of the fractional covering number for hypergraphs implies that the integrality gap for the domination number is at most logarithmic, specifically, $\frac{\gamma(G)}{\gamma_f(G)} \leq 1 + \ln n$ for a non-empty graph G with n vertices. This results in an LP-based algorithm for approximation of $\gamma(G)$ within a logarithmic factor, which is essentially optimal as $\gamma(G)$ is hard to approximate within a sublogarithmic factor assuming NP $\neq$ P [22]. As shown by Rubalcaba [23], γ_f is 1-WL-invariant. Along with the Lovász bound, this implies that the 1-WL-invariance ratio of γ is at most logarithmic. On the other hand, Chappell et al. [6] have shown that the logarithmic upper bound for the integrality gap of γ_f is tight up to a constant factor. In Sect. 6 we prove an $\Omega(\log n)$ lower bound even for the 1-WL-invariance ratio of γ over n-vertex graphs. This implies the integrality gap lower bound [6], reproving it from a different perspective. In Sect. 6 we also discuss the *additive* integrality gap of the fractional edge-disjoint triangle packing.

Related Work. Atserias and Dawar [3] have shown that the 1-WL-invariance ratio for the vertex cover number τ is at most 2. Alternatively, this bound also follows from the 1-WL-invariance of ν_f (which implies the 1-WL-invariance of τ_f as $\tau_f = \nu_f$ by LP duality) combined with a standard rounding argument. The approach of [3] uses a different argument, which alone does not yield 1-WL-invariance of the fractional vertex cover τ_f.

The bound of 2 for the 1-WL-invariance ratio of τ is optimal. Atserias and Dawar [3] also show that the k-WL-invariance ratio for τ is at least 7/6 for each k. This implies an unconditional inapproximability result for Vertex Cover in the model of encoding-invariant computations expressible in FPC.

Notation and Formal Definitions. For $\bar{x} = (x_1, \ldots, x_k)$ in $V(G)^k$, let $\mathrm{WL}_k^0(G, \bar{x})$ be the $k \times k$ matrix $(m_{i,j})$ with $m_{i,j} = 1$ if $x_i x_j \in E(G)$, $m_{i,j} = 2$ if $x_i = x_j$ and $m_{i,j} = 0$ otherwise. We also augment $\mathrm{WL}_k^0(G, \bar{x})$

by the vector of the colors of $x_1, \ldots, x_k$ if the graph G is vertex-colored. $\mathrm{WL}_k^0(G, \bar{x})$ encodes the ordered isomorphism type of $\bar{x}$ in G and serves as an initial coloring of $V(G)^k$ for k-WL. In each refinement round, 1-WL computes $\mathrm{WL}_1^{r+1}(G, x) = (\mathrm{WL}_1^r(G, x), \{\!\{\mathrm{WL}_1^r(G, y) : y \in N(x)\}\!\})$, where $N(x)$ is the neighborhood of x and $\{\!\{\ \}\!\}$ denotes a multiset. If $k \geq 2$, k-WL refines the coloring by $\mathrm{WL}_k^{r+1}(G, x) = (\mathrm{WL}_k^r(G, \bar{x}), \{\!\{(\mathrm{WL}_k^r(G, \bar{x}_1^u), \ldots, \mathrm{WL}_k^r(G, \bar{x}_k^u) : u \in V(G)\}\!\})$, where $\bar{x}_i^u$ is the tuple $(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_k)$. If G has n vertices, the color partition stabilizes in at most n^k rounds. We define $\mathrm{WL}_k(G, \bar{x}) = \mathrm{WL}_k^{n^k}(G, \bar{x})$ and $\mathrm{WL}_k(G) = \{\!\{\mathrm{WL}_k(G, \bar{x}) : \bar{x} \in V(G)^k\}\!\}$. Now, $G \equiv_{k\text{-WL}} H$ if $\mathrm{WL}_k(G) = \mathrm{WL}_k(H)$.

The color partition of $V(G)$ according to $\mathrm{WL}_1(G, x)$ is *equitable*: for any color classes C and C', each vertex in C has the same number of neighbors in C'. Moreover, if G is vertex-colored, then the original colors of all vertices in each C are the same. If $V(G) = V(H)$, then $G \equiv_{k\text{-WL}} 1H$ exactly when G and H have a common equitable partition [24, Theorem 6.5.1].

Let G and H be graphs with vertex set $\{1, \ldots, n\}$, and let A and B be the adjacency matrices of G and H, respectively. Then G and H are isomorphic if and only if $AX = XB$ for some $n \times n$ permutation matrix X. The linear programming relaxation allows X to be a doubly stochastic matrix. If such an X exists, G and H are said to be *fractionally isomorphic*. If G and H are colored graphs with the same partition of the vertex set into color classes, then it is additionally required that $X_{u,v} = 0$ whenever u and v are of different colors. It turns out that two graphs are indistinguishable by color refinement if and only if they are fractionally isomorphic [24, Theorem 6.5.1].

2 Reductions Between Linear Programs

A *linear program (LP)* is an optimization problem of the form "maximize (or minimize) $a^t x$ subject to $Mx \leq b$", where $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, M is an $m \times n$ matrix $M \in \mathbb{R}^{m \times n}$, and x varies over all vectors in $\mathbb{R}^n$ with nonnegative entries (which we denote by $x \geq 0$). Any vector x satisfying the constraints $Mx \leq b$, $x \geq 0$ is called a *feasible solution* and the function $x \mapsto a^t x$ is called the *objective function*. We denote an LP with parameters a, M, b by $LP(a, M, b, opt)$, where $opt = \min$, if the goal is to minimize the value of the objective function, and $opt = \max$, if this value has to be maximized. The optimum of the objective function over all feasible solutions is called the *value* of the program $L = LP(a, M, b, opt)$ and denoted by $val(L)$. Our goal now is to introduce an equivalence relation between LPs ensuring equality of their values.

Equivalence of LPs. Let $L_1 = LP(a, M, b, opt)$ and $L_2 = LP(c, N, d, opt)$ be linear programs (in general form), where $a, c \in \mathbb{R}^n$, $b, d \in \mathbb{R}^m$, $M, N \in \mathbb{R}^{m \times n}$ and $opt \in \{\min, \max\}$. We say that L_1 *reduces* to L_2 ($L_1 \leq L_2$ for short), if there are matrices $Y \in \mathbb{R}^{m \times m}$ and $Z \in \mathbb{R}^{n \times n}$ such that

- $Y, Z \geq 0$
- $a^t Z \diamondsuit c^t$, where $\diamondsuit = \begin{cases} \leq, & opt = \min \\ \geq, & opt = \max \end{cases}$

– $MZ \le YN$
– $Yd \le b$

L_1 and L_2 are said to be *equivalent* ($L_1 \equiv L_2$ for short) if $L_1 \le L_2$ and $L_2 \le L_1$.

Theorem 1. *If $L_1 \equiv L_2$, then $val(L_1) = val(L_2)$.*

Proof. Let $L_1 = LP(a, M, b, opt)$ and $L_2 = LP(c, N, d, opt)$ and assume $L_1 \le L_2$ via (Y, Z). We show that for any feasible solution x of L_2 we get a feasible solution $x' = Zx$ of L_1 with $a^t x' \diamondsuit c^t x$, where $\diamondsuit$ is as in the definition:

$$Mx' = \underbrace{MZ}_{\le YN} x \le Y \underbrace{Nx}_{\le d} \le Yd \le b \text{ and } a^t x' = \underbrace{a^t Z}_{\diamondsuit\ c^t} x \diamondsuit c^t x.$$

Thus, $L_1 \le L_2$ implies $val(L_1) \diamondsuit val(L_2)$ and the theorem follows. □

LPs with Fractionally Isomorphic Matrices. Recall that a square matrix $X \ge 0$ is *doubly stochastic* if its entries in each row and column sum up to 1. We call two $m \times n$ matrices M and N *fractionally isomorphic* if there are doubly stochastic matrices $Y \in \mathbb{R}^{m\times m}$ and $Z \in \mathbb{R}^{n\times n}$ such that

$$MZ = YN \text{ and } NZ^t = Y^t M. \tag{1}$$

Grohe et al. [16, Eqs. (5.1)–(5.2) in arXiv version] discuss similar definitions. They use fractional isomorphism fractional isomorphism to reduce the dimension of linear equations and LPs. The meaning of (1) will be clear from the proof of Theorem 3 below.

Lemma 2. *If M and N are fractionally isomorphic $m \times n$ matrices, then*

$$LP(\mathbb{1}_n, M, \mathbb{1}_m, opt) \equiv LP(\mathbb{1}_n, N, \mathbb{1}_m, opt),$$

where $\mathbb{1}_n$ denotes the n-dimensional all-ones vector.

Proof. Since the matrices Y and Z in (1) are doubly stochastic, $Y\mathbb{1}_m = \mathbb{1}_m$ and $\mathbb{1}_n^t Z = \mathbb{1}_n^t$. Along with the first equality in (1), these equalities imply that $L_1 \le L_2$. The reduction $L_2 \le L_1$ follows similarly from the second equality in (1) as Y^t and Z^t are doubly stochastic. □

3 Fractional Set Packing

The *Set Packing* problem is, given a family of sets $\mathcal{S} = \{S_1, \ldots, S_n\}$, where $S_j \subset \{1, \ldots, m\}$, to maximize the number of pairwise disjoint sets in this family. The maximum is called in combinatorics the *matching number of hypergraph* $\mathcal{S}$ and denoted by $\nu(\mathcal{S})$. The fractional version is given by $LP(\mathcal{S}) = LP(\mathbb{1}_n, M, \mathbb{1}_m, \max)$ where M is the $m \times n$ incidence matrix of $\mathcal{S}$, namely

$$\max \sum_{i=1}^{n} x_i \quad \text{under}$$

$$x_i \ge 0 \text{ for every } i \le n,$$

$$\sum_{i\,:\,S_i \ni j} x_i \le 1 \text{ for every } j \le m.$$

The optimum value $\nu_f(\mathcal{S}) = val(LP(\mathcal{S}))$ is called the *fractional matching number of* $\mathcal{S}$.

Let $I(\mathcal{S})$ denote the incidence graph of $\mathcal{S}$. Specifically, this is the vertex-colored bipartite graph with biadjacency matrix M on two classes of vertices; m vertices are colored red, n vertices are colored blue, and a red vertex j is adjacent to a blue vertex i if $j \in S_i$.

Theorem 3. *Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two families each consisting of n subsets of the set $\{1, \ldots, m\}$. If $I(\mathcal{S}_1) \equiv_{\text{1-WL}} I(\mathcal{S}_2)$, then $\nu_f(\mathcal{S}_1) = \nu_f(\mathcal{S}_2)$.*

Proof. Denote the incidence matrices of $\mathcal{S}_1$ and $\mathcal{S}_2$ by M and N respectively. Let

$$A_1 = \begin{pmatrix} 0 & M \\ M^t & 0 \end{pmatrix} \text{ and } A_1 = \begin{pmatrix} 0 & N \\ N^t & 0 \end{pmatrix}$$

be the adjacency matrices of $I(\mathcal{S}_1)$ and $I(\mathcal{S}_2)$ respectively. Since $I(\mathcal{S}_1)$ and $I(\mathcal{S}_2)$ are indistinguishable by color refinement, by [24, Theorem 6.5.1] we conclude that these graphs are fractionally isomorphic, that is, there is a doubly stochastic matrix X such that

$$A_1 X = X A_2 \tag{2}$$

and $X_{uv} = 0$ whenever u and v are from different vertex color classes. The latter condition means that X is the direct sum of an $n \times n$ doubly stochastic matrix Y and an $n \times n$ doubly stochastic matrix Z, that is, Equality (2) reads

$$\begin{pmatrix} 0 & M \\ M^t & 0 \end{pmatrix} \begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix} = \begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix} \begin{pmatrix} 0 & N \\ N^t & 0 \end{pmatrix},$$

yielding $MZ = YN$ and $M^t Y = ZN^t$. Thus, M and N are fractionally isomorphic. Lemma 2 implies that $LP(\mathcal{S}_1) \equiv LP(\mathcal{S}_2)$. Therefore, these LPs have equal values by Theorem 1. □

4 1-WL-invariance of the Fractional Matching Number

Recall that a set of edges $M \subseteq E(G)$ is a *matching* in a graph G if every vertex of G is incident to at most one edge from M. The *matching number* $\nu(G)$ is the maximum size of a matching in G. Note that this terminology and notation agrees with Sect. 3 when graphs are considered hypergraphs with hyperedges of size 2. Fractional Matching is defined by the LP

$$\max \sum_{uv \in E(G)} x_{uv} \quad \text{under}$$

$$x_{uv} \geq 0 \text{ for every } uv \in E(G),$$

$$\sum_{v \in N(u)} x_{uv} \leq 1 \text{ for every } u \in V(G),$$

whose value is the *fractional matching number* $\nu_f(G)$. The above LP is exactly the linear program $LP(\mathcal{S}_G)$ for the instance $\mathcal{S}_G = E(G)$ of Fractional Set Packing formed by the edges of G as 2-element subsets of $V(G)$, that is, $\nu_f(G) = \nu_f(\mathcal{S}_G)$.

Theorem 4. *The fractional matching number is* 1-WL-*invariant.*

Proof. Given $G \equiv_{\text{1-WL}} H$, we have to prove that $\nu_f(G) = \nu_f(H)$ or, equivalently, $\nu_f(\mathcal{S}_G) = \nu_f(\mathcal{S}_H)$ where $\mathcal{S}_G$ is as defined above. By Theorem 3, it suffices to show that $I(\mathcal{S}_G) \equiv_{\text{1-WL}} I(\mathcal{S}_H)$. To this end, we construct a common equitable partition of $I(\mathcal{S}_G)$ and $I(\mathcal{S}_H)$, appropriately identifying their vertex sets. Recall that $V(I(\mathcal{S}_G)) = V(G) \cup E(G)$ and a red vertex $x \in V(G)$ is adjacent to a blue vertex $e \in E(G)$ if $x \in e$.

For $x \in V(G)$, let $c_G(x) = \text{WL}_1(G, x)$ and define c_H on $V(H)$ similarly. First, we identify $V(G)$ and $V(H)$ (i.e., the red parts of the two incidence graphs) so that $c_G(x) = c_H(x)$ for every x in $V(G) = V(H)$, which is possible because 1-WL-equivalent graphs have the same color palette after color refinement. The color classes of c_G now form a common equitable partition of G and H.

Next, extend the coloring c_G to $E(G)$ (the blue part of $I(\mathcal{S}_G)$) by $c_G(\{x, y\}) = \{c_G(x), c_G(y)\}$, and similarly extend c_H to $E(H)$. Denote the color class of c_G containing $\{x, y\}$ by $C_G(\{x, y\})$, the color class containing x by $C_G(x)$ etc. Note that $|C_G(\{x, y\})|$ is equal to the number of edges in G between $C_G(x)$ and $C_G(y)$ (or the number of edges within $C_G(x)$ if $c_G(x) = c_G(y)$). Since $\{C_G(x)\}_{x \in V(G)}$ is a common equitable partition of G and H, we have $|C_G(\{x, y\})| = |C_H(\{x', y'\})|$ whenever $c_G(\{x, y\}) = c_H(\{x', y'\})$ (note that $\{x, y\}$ does not need to be an edge in H, nor $\{x', y'\}$ needs to be an edge in G). This allows us to identify $E(G)$ and $E(H)$ so that $c_G(e) = c_H(e)$ for every e in $E(G) = E(H)$.

Now, consider the partition of $V(G) \cup E(G)$ into the color classes of c_G (or the same in terms of H) and verify that this is an equitable partition for both $I(\mathcal{S}_G)$ and $I(\mathcal{S}_H)$. Indeed, let $C \subseteq V(G)$ and $D \subseteq E(G)$ be color classes of c_G such that there are $x \in C$ and $e \in D$ adjacent in $I(\mathcal{S}_G)$, that is, $e = \{x, y\}$ for some vertex y of G. Note that, if considered on $V(H) \cup E(H)$, the classes C and D also must contain $x' \in C$ and $e' = \{x', y'\} \in D$ adjacent in $I(\mathcal{S}_H)$ (take $x' = x$ and any y' adjacent to x in H such that $c_H(y') = c_G(y)$). Denote $C' = C_G(y)$ (it is not excluded that $C' = C$). The vertex x has exactly as many D-neighbors in $I(\mathcal{S}_G)$ as it has C'-neighbors in G. This number depends only on C and C' or, equivalently, only on C and D. The same number is obtained also while counting the D-neighbors of x' in $I(\mathcal{S}_H)$.

On the other hand, e has exactly one neighbor x in C if $C' \neq C$ and exactly two C-neighbors x and y if $C' = C$. What is the case depends only on D and C, and is the same in $I(\mathcal{S}_G)$ and $I(\mathcal{S}_H)$. Thus, we do have a common equitable partition of $I(\mathcal{S}_G)$ and $I(\mathcal{S}_H)$. □

As was discussed in Sect. 1, we are able to generalize Theorem 4 to any fractional F-packing number π_f^F. For a graph G, let $\mathcal{S}_{F,G}$ be the family of subsets of $V(G)$ consisting of the vertex sets $V(F')$ of all subgraphs F' of G isomorphic to the pattern graph F. Now, $\pi_f^F(G) = \nu_f(\mathcal{S}_{F,G})$. Dell et al. [9] establish a close connection between homomorphism counts and k-WL equivalence, which motivates the following definition. The *homomorphism-hereditary treewidth* of a graph F, denoted by $htw(F)$, is the maximum treewidth $tw(F')$ over all homomorphic images F' of F. The proof of the following result can be found in the full version of the paper [2].

Theorem 5. *If $htw(F) \leq k$, then π_f^F is k-WL-invariant.*

First-Order Interpretability. Our approach to proving Theorem 4 was, given an instance graph G of Fractional Matching Problem, to define an instance $\mathcal{S}_G$ of Fractional Set Packing Problem having the same LP value. The following definition concerns many similar situations. Given a correspondence $G \mapsto \mathcal{S}_G$, we say that an istance $\mathcal{S}_G$ of Fractional Set Packing is definable over a graph G with *excess* e if $G \equiv_{(1+e)\text{-WL}} H$ implies $I(\mathcal{S}_G) \equiv_{1\text{-WL}} I(\mathcal{S}_H)$.

This definition includes a particular situation when $I(\mathcal{S}_G)$ is first-order *interpretable* in G in the sense of [11, Chapter 12.3], which means that for the color predicates (to be red or blue respectively) as well as for the adjacency relation of $I(\mathcal{S}_G)$ we have first order formulas defining them on $V(G)^k$ for some k in terms of the adjacency relation of G. The number k is called *width* of the interpretation. In this case, if there is a first-order sentence over s variables that is true on $I(\mathcal{S}_G)$ but false on $I(\mathcal{S}_H)$, then there is a first-order sentence over sk variables that is true on G but false on H. Cai, Fürer, and Immerman [5] showed that two structures are $\equiv_{k\text{-WL}}$-equivalent iff they are equivalent in the $(k+1)$-variable counting logic C^{k+1}. Therefore, Theorem 3 has the following consequence.

Corollary 6. *Let π_f be a fractional graph parameter such that $\pi_f(G) = \nu_f(\mathcal{S}_G)$, where $\mathcal{S}_G$ admits a first-order interpretation of width k in G (even possibly with counting quantifiers). Under these conditions, $\mathcal{S}_G$ is definable over G with excess $2(k-1)$ and, hence, π_f is $(2k-1)$-WL-invariant.*

To obtain 1-WL-invariance via Corollary 6, we would need an interpretation of width 1. This is hardly possible in the case of the fractional matching number, and an interpretation of width 2 could only give us 3-WL-invariance of ν_f. Thus, our purely combinatorial argument for Theorem 4 is preferable here.

5 Fractional Edge-Disjoint Triangle Packing

We now show that the approach we used in the proof of Theorem 4 works as well for edge-disjoint packing. Given a graph G, let $T(G)$ denote the family of all sets $\{e_1, e_2, e_3\}$ consisting of the edges of a triangle subgraph in G. We regard $T(G)$ as a family $\mathcal{S}_G$ of subsets of the edge set $E(G)$. The optimum value of Set Packing Problem on $\mathcal{S}_G$, which we denote by $\rho^{K_3}(G)$, is equal to the maximum number of edge-disjoint triangles in G. Let $\rho_f^{K_3}(G) = \nu_f(\mathcal{S}_G)$ be the corresponding fractional parameter.

Theorem 7. *The fractional packing number $\rho_f^{K_3}$ is 2-WL-invariant.*

Proof. Given a graph G, we consider the coloring c_G of $E(G) \cup T(G)$ defined by $c_G(\{x,y\}) = \{\mathrm{WL}_2(G,x,y), \mathrm{WL}_2(G,y,x)\}$ on $E(G)$ and $c_G(\{e_1,e_2,e_3\}) = \{\!\{c_G(e_1), c_G(e_2), c_G(e_3)\}\!\}$ on $T(G)$. Like in the proof of Theorem 4, the upper case notation $C_G(w)$ will be used to denote the color class of $w \in E(G) \cup T(G)$.

Suppose that $G \equiv_{2\text{-WL}} H$. This condition means that we can identify the sets $E(G)$ and $E(H)$ so that $c_G(e) = c_H(e)$ for every e in $E(G) = E(H)$.

Moreover, the 2-WL-equivalence of G and H implies that $|C_G(t)| = |C_H(t')|$ for any $t \in T(G)$ and $t' \in T(H)$ with $c_G(t) = c_H(t')$. This allows us to identify $T(G)$ and $T(H)$ so that $c_G(t) = c_H(t)$ for every t in $T(G) = T(H)$. As in the proof of Theorem 4, it suffices to argue that $\{C_G(w)\}_{w \in E(G) \cup T(G)}$ is a common equitable partition of the incidence graphs $I(\mathcal{S}_G)$ and $I(\mathcal{S}_H)$. The equality $\rho_f^{K_3}(G) = \rho_f^{K_3}(H)$ will then follow by Theorem 3.

Let $C \subseteq E(G)$ and $D \subseteq T(G)$ be color classes of c_G such that there is an edge between them in $I(\mathcal{S}_G)$, that is, there are $e \in C$ and $t \in D$ such that $t = \{e, e_2, e_3\}$. If considered on $E(H) \cup T(H)$, the classes C and D also must contain $e' \in C$ and $t' = \{e', e'_2, e'_3\} \in D$ adjacent in $I(\mathcal{S}_H)$ (take, for example, the edge $e' = e$ of H and extend it to a triangle with other two edges e'_2 and e'_3 such that $c_H(e'_2) = c_G(e_2)$ and $c_H(e'_3) = c_G(e_3)$, which must exist in H because H and G are 2-WL-equivalent). Denote $C' = C_G(e_2)$ and $C'' = C_G(e_3)$ (it is not excluded that some of the classes C, C', and C'' coincide).

Let x, y, and z be the vertices of the triangle t in G, and suppose that $e = \{x, y\}$. The number of D-neighbors that e has in $I(\mathcal{S}_G)$ is equal to the number of vertices z' such that $(\mathrm{WL}_2(G, x, z'), \mathrm{WL}_2(G, z', y))$ is one of the 8 pairs in $(c_G(\{x, z\}) \times c_G(\{y, z\})) \cup (c_G(\{y, z\}) \times c_G(\{x, z\}))$, like $(\mathrm{WL}_2(G, z, y), \mathrm{WL}_2(G, x, z))$ (some of these pairs can coincide). Since the partition of $V(G)^2$ by the coloring $\mathrm{WL}_2(G, \cdot, \cdot)$ is not further refined by 2-WL, this number depends only on C and D. We obtain the same number also while counting the D-neighbors of e' in $I(\mathcal{S}_H)$.

On the other hand, t has exactly one neighbor e in C if C differs from both C' and C'', exactly two C-neighbors if C coincides with exactly one of C' and C'', and exactly three C-neighbors e, e_2, and e_3 if $C = C' = C''$. Which of the three possibilities occurs depends only on D and C, and is the same in $I(\mathcal{S}_G)$ and $I(\mathcal{S}_H)$. This completes our verification that we really have a common equitable partition. □

6 Invariance Ratio and Integrality Gap

Recall the discussion in the introduction about the domination number $\gamma(G)$.

Theorem 8. *For infinitely many n, there are n-vertex* 1-WL-*equivalent graphs G and H such that $\gamma(G)/\gamma(H) > \frac{1}{20} \ln n - 1$.*

Proof. It suffices to show that the variation of the domination number among n-vertex d-regular graphs is logarithmic for an appropriate choice of the degree function $d = d(n)$.

Assuming that dn is even, let $\mathbb{R}(n, d)$ denote a random d-regular graph on n vertices. Given $p \in (0, 1)$, let $\mathbb{G}(n, p)$ denote the Erdős–Rényi random graph with edge probability p. Kim and Vu [18] proved for certain degree functions $d = d(n)$ that the distribution $\mathbb{R}(n, d)$ can be approximated from below and above, with respect to the subgraph relation, by distributions $\mathbb{G}(n, p_1)$ and $\mathbb{G}(n, p_2)$ with $p_1 = (1 - o(1))\frac{d}{n}$ and $p_2 = (1 + o(1))\frac{d}{n}$. We need the part of this sandwiching result about the approximation from above.

For our purposes, we consider pairs n, d such that $n = (2d)^4$ and, thus, $d = n^{1/4}/2$. Applied to this case, the Kim-Vu theorem says that there is a joint distribution of $\mathbb{R}(n, d)$ and $\mathbb{G}(n, p)$ with $p = (1 + o(1))\frac{d}{n} = (\frac{1}{2} + o(1))n^{-3/4}$ such that $\Delta(\mathbb{R}(n, d) \setminus \mathbb{G}(n, p)) \leq 4$ with probability $1 - o(1)$ as n increases. It follows that

$$\gamma(\mathbb{G}(n, p)) \leq 5\,\gamma(\mathbb{R}(n, d))$$

with probability $1 - o(1)$. Glebov et al. [15] proved that $\gamma(\mathbb{G}(n, p)) = \frac{\ln(np)}{p}(1 + o(1))$ with probability $1-o(1)$ whenever $p \to 0$ and $pn \to \infty$. Hence $\gamma(\mathbb{R}(n, d)) \geq \frac{1}{5}\frac{n}{d}\ln d$ with probability $1-o(1)$. As a consequence, there is an n-vertex d-regular graph G with $\gamma(G) \geq \frac{1}{5}\frac{n}{d}\ln d$.

On the other hand, consider $H = \frac{n}{2d}\,K_{d,d}$, where $K_{s,t}$ stands for the complete bipartite graph with vertex classes of size s and t, and note that $\gamma(H) = \frac{n}{d}$. Therefore, $\gamma(G)/\gamma(H) \geq \frac{1}{5}\ln d$, which readily gives us the desired bound. □

We conclude with a discussion of Edge-Disjoint Triangle Packing. Haxell and Rödl [17] proved that ρ^{K_3} is well approximated by $\rho_f^{K_3}$ on dense graphs as $\rho_f^{K_3}(G) - \rho^{K_3}(G) = o(n^2)$ for n-vertex G. On the other hand, Yuster [26] showed that $\rho_f^{K_3}(G) - \rho^{K_3}(G) = \Omega(n^{1.5})$ for infinitely many G, and it is open whether this lower bound is tight. Define the *invariance discrepancy* of ρ^{K_3} as the function $D_{K_3}(n) = \max|\rho^{K_3}(G) - \rho^{K_3}(H)|$ where the maximum is taken over 2-WL-equivalent n-vertex graphs G and H. As follows from Theorem 7, this function provides a lower bound for the maximum integrality gap $\rho_f^{K_3}(G) - \rho^{K_3}(G)$ over n-vertex graphs. In this respect, it is reasonable to ask what the asymptotics of $D_{K_3}(n)$ is. The following fact is a step towards this goal.

Proposition 9. *$D_{K_3}(n) = \Omega(n)$.*

Proof. Consider $G = tS$ and $H = tR$, where S and R are the Shrikhande and 4×4 rook's graphs respectively. Both have vertex set $\mathbb{Z}_4 \times \mathbb{Z}_4$, and (i, j) and (i', j') are adjacent in S if ($i = i'$ and $j' = j + 1$) or ($j = j'$ and $i' = i + 1$) or ($i' = i+1$ and $j' = j+1$), where equality is in $\mathbb{Z}_4$, while they are adjacent in R if $i = i'$ (row 4-clique) or $j = j'$ (column 4-clique). S is completely decomposable into edge-triangles $\{(i, j), (i + 1, j), (i + 1, j + 1)\}$ and, hence, $\rho^{K_3}(S) = 16$. On the other hand, in R the edges of each K_3 all belong to the same row or column 4-clique. Since a packing can take at most one K_3 from each row/column K_4, we have $\rho^{K_3}(R) = 8$. This yields $\rho^{K_3}(G) - \rho^{K_3}(H) = 8t$ as desired. □

References

1. Anderson, M., Dawar, A., Holm, B.: Solving linear programs without breaking abstractions. J. ACM **62**(6), 48:1–48:26 (2015)
2. Arvind, V., Fuhlbrück, F., Köbler, J., Verbitsky, O.: On the Weisfeiler-Leman dimension of Fractional Packing. Technical report, arxiv.org/abs/1910.11325 (2019)
3. Atserias, A., Dawar, A.: Definable inapproximability: new challenges for duplicator. In: Proceedings of CSL 2018. LIPIcs, vol. 119, pp. 7:1–7:21 (2018)

4. Babai, L.: Graph isomorphism in quasipolynomial time. In: Proceedings of STOC 2016, pp. 684–697 (2016)
5. Cai, J., Fürer, M., Immerman, N.: An optimal lower bound on the number of variables for graph identifications. Combinatorica **12**(4), 389–410 (1992)
6. Chappell, G., Gimbel, J., Hartman, C.: Approximations of the domination number of a graph. J. Combin. Math. Combin. Comput. **104**, 287–297 (2018)
7. Dawar, A.: The nature and power of fixed-point logic with counting. SIGLOG News **2**(1), 8–21 (2015)
8. Dawar, A., Severini, S., Zapata, O.: Pebble games and cospectral graphs. Electron. Notes Discrete Math. **61**, 323–329 (2017)
9. Dell, H., Grohe, M., Rattan, G.: Lovász meets Weisfeiler and Leman. In: Proceedings of ICALP 2018. LIPIcs, vol. 107, pp. 40:1–40:14 (2018)
10. Dor, D., Tarsi, M.: Graph decomposition is NP-complete: a complete proof of Holyer's conjecture. SIAM J. Comput. **26**(4), 1166–1187 (1997)
11. Ebbinghaus, H.D., Flum, J.: Finite Model Theory. Springer Monographs in Mathematics. Springer, Berlin (2006)
12. Füredi, Z.: Matchings and covers in hypergraphs. Graphs Comb. **4**(1), 115–206 (1988)
13. Fürer, M.: On the power of combinatorial and spectral invariants. Linear Algebra Appl. **432**(9), 2373–2380 (2010)
14. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman, San Francisco (1979)
15. Glebov, R., Liebenau, A., Szabó, T.: On the concentration of the domination number of the random graph. SIAM J. Discrete Math. **29**(3), 1186–1206 (2015)
16. Grohe, M., Kersting, K., Mladenov, M., Selman, E.: Dimension reduction via colour refinement. In: Schulz, A.S., Wagner, D. (eds.) ESA 2014. LNCS, vol. 8737, pp. 505–516. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44777-2_42. arXiv version: arxiv.org/abs/1307.5697
17. Haxell, P.E., Rödl, V.: Integer and fractional packings in dense graphs. Combinatorica **21**(1), 13–38 (2001)
18. Kim, J., Vu, V.: Sandwiching random graphs: universality between random graph models. Adv. Math. **188**(2), 444–469 (2004)
19. Kirkpatrick, D.G., Hell, P.: On the complexity of general graph factor problems. SIAM J. Comput. **12**(3), 601–609 (1983)
20. Lovász, L.: On the ratio of optimal integral and fractional covers. Discrete Math. **13**(4), 383–390 (1975)
21. Otto, M.: Bounded Variable Logics and Counting: A Study in Finite Models. Lecture Notes in Logic, vol. 9. Cambridge University Press, Cambridge (2017)
22. Raz, R., Safra, S.: A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In: Proceedings of STOC 1997, pp. 475–484. ACM (1997)
23. Rubalcaba, R.R.: Fractional domination, fractional packings, and fractional isomorphisms of graphs. Ph.D. thesis, Auburn University (2005)
24. Scheinerman, E.R., Ullman, D.H.: Fractional Graph Theory. A Rational Approach to the Theory of Graphs. Wiley, Hoboken (1997)
25. Weisfeiler, B., Leman, A.: The reduction of a graph to canonical form and the algebra which appears therein. NTI Ser. 2 **9**, 12–16 (1968). English translation is available at https://www.iti.zcu.cz/wl2018/pdf/wl_paper_translation.pdf
26. Yuster, R.: Integer and fractional packing of families of graphs. Random Struct. Algorithms **26**(1–2), 110–118 (2005)