

# An Upper Bound on the Weisfeiler-Leman Dimension\*

Thomas Schneider and Pascal Schweitzer  
TU Darmstadt

September 14, 2025

## Abstract

The Weisfeiler-Leman (WL) dimension is a standard measure in descriptive complexity theory for the structural complexity of a graph. We prove that the WL-dimension of a graph on  $n$  vertices is at most  $3/20 \cdot n + o(n) = 0.15 \cdot n + o(n)$ .

The proof develops various techniques to analyze the structure of coherent configurations. This includes sufficient conditions under which a fiber can be restored up to isomorphism if it is removed, a recursive proof exploiting a degree reduction and treewidth bounds, as well as an analysis of interspaces involving small fibers.

As a base case, we also analyze the dimension of coherent configurations with small fiber size and thereby graphs with small color class size.

## 1 Introduction

In recent years, the Weisfeiler-Leman (WL) dimension has evolved to become a standard measure for the structural complexity of a graph. Initially coined in the context of isomorphism questions [Bab79, Wei76, IL90], a plethora of equivalent reformulations in seemingly unrelated areas has surfaced (e.g. [CFI92, Dvo10, AM13, GO15, DGR18, AMR<sup>+</sup>19]). The concept in particular has applications in machine learning on graphs (see [MLM<sup>+</sup>23] for a survey).

In its initial formulation, the WL-dimension of a graph  $G$  is characterized as the minimum  $k$  required so that the  $k$ -dimensional Weisfeiler-Leman algorithm distinguishes  $G$  from every non-isomorphic graph. By a central result of Cai, Fürer, and Immerman [CFI92], the dimension plus one is also the least number of variables required to identify  $G$  in a particular logic (fixed-point logic with counting) and also the number of pebbles required in a particular combinatorial pebble game (the bijective pebble game).

In some sense, the Weisfeiler-Leman dimension measures how difficult it is to test isomorphism of one graph to others using a fairly general class of combinatorial algorithms. Crucially, isomorphism between graphs of bounded Weisfeiler-Leman dimension can be decided in polynomial time. More precisely, if the WL-dimension is at most  $k$ , then the problem can be solved in time  $O(n^{k+1} \log n)$ . While group theoretic techniques can circumvent the structural complexity given by high WL-dimension, the currently fastest theoretical algorithm, which runs in quasi-polynomial time [Bab16], uses a  $O(\log(n))$ -dimensional Weisfeiler-Leman algorithm as a subroutine.

Many graph classes are known to have bounded Weisfeiler-Leman dimension (e.g., all classes with a forbidden minor [Gro17] and bounded clique width graphs [GN23]), giving a polynomial time isomorphism algorithm for these classes.

---

\*The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (EngageS: grant agreement No. 820148).

Initial hopes to find a general bound on the WL-dimension of graphs, however, were dispelled by the seminal construction of Cai, Fürer, and Immerman [CFI92] which constructs graphs of order  $n$  and Weisfeiler-Leman dimension  $\Omega(n)$ .

In this paper we investigate explicit bounds for the Weisfeiler-Leman dimension. A calculation of the precise constant in [CFI92] yields a lower bound of  $0.00465 \cdot n$  as demonstrated in [PVV06, PV09]. This constitutes the current best lower bound in the literature. The best explicit upper bound is  $0.5n + 1.5$  [PVV06, PV09], which follows from upper bounds that apply in a more general context (more specifically from bounds on the “non-counting version” of the WL-algorithm).

**Results.** As main result in this paper, we prove that for all graphs  $\text{WLdim}(G) \leq 3/20 \cdot n + o(n) = 0.15 \cdot n + o(n)$ . As a side note, we also observe that for all orders there are graphs with a Weisfeiler-Leman dimension of at least  $0.0105027 \cdot n - o(n)$ . As usual, these improved bounds can both be recast in logical terms, bounding the number of variables required for graph identification in fixed point-logic with counting.

**Techniques.** On a macroscopic scale the idea for the upper bound proof is as follows. To facilitate recursion, we first generalize the problem to the realm of coherent configurations. These naturally generalize various highly regular graph families, such as strongly regular graphs and distance-regular graphs. They can be understood as the stable colorings under the 2-dimensional Weisfeiler-Leman algorithm. Increasing the dimension by at most 2, we can at any point in time, even during recursion, assume that all our objects are coherent configurations. The color classes of graphs translate into *fibers* of the coherent configuration.

The proof strategy is to reduce the number of vertices, with a focus on reducing those contained in large fibers. We repeatedly use individualizations (artificially assigning single vertices a different color) to simplify the coherent configuration. A vertex individualization decreases the dimension by at most 1. We argue that, at a sublinear cost, we can ensure that the maximum color class size is sublinear (Section 8).

With small fibers in mind, we investigate situations in which a color class can be removed without decreasing the Weisfeiler-Leman dimension. We call configurations *critical* if no fiber can be removed this way. Fibers of size at most 3, which we call *tiny*, can always be removed. For fibers that are not tiny, we develop a technique to determine from the combinatorial structure that a fiber is *restorable* and thus can be removed (Section 3). The technique is based on extendability of automorphisms of induced subconfigurations.

Our base case of the recursion is the situation in which all fibers have size at most 7. We call such fibers *small*. We analyze the structure of the configurations that small fibers can induce, as well as the possible connections between small fibers (Section 4). These connections are called *interspaces*. The *quotient graph* captures the structural information given by fibers and interspaces. Several reductions lead us to a quotient graph of maximum degree at most 3, at which point we can use bounds on the treewidth for cubic graphs to bound the Weisfeiler-Leman dimension. Overall we show that a coherent configuration with fiber size at most 7 has Weisfeiler-Leman dimension at most  $n/20 + O(1)$  (Section 7).

For the recursion we also need to understand the possible interspaces between small and large fibers (Section 5). We define a potential function (Section 9) that measures the progress we make towards the base case. It gives us the possibility to trade individualizations for a reduction of the potential. We then define a sequence of local reductions that, for various subconfigurations and types of interspace, provide a positive trade-off (Section 10). We can thus inductively assume that these subconfigurations and interspaces are not present in our configuration.

To finally reach the base case, we employ a global argument. We introduce the concept of a *t-reduced* configuration (Section 11) and show that configurations to which none of the local

reductions are applicable are  $t$ -reduced. For reduced structures, the global argument allows us to separate the graph into pieces whose underlying structure either has small treewidth or which consist only of small fibers (Section 12). Overall this recursive approach proves the main theorem (Section 13).

In contrast to the upper bounds, the lower bound simply follows straightforwardly by combining three known results on expansion, treewidth, and the CFI-construction in the evident fashion. An intermediate step in that argument is that random cubic graphs asymptotically almost surely have treewidth at least  $0.04201 \cdot n$  (Section 14).

**Related Work.** A concrete classification of graphs with Weisfeiler-Leman dimension 1 is known [AKRV17, KSS22]. Fuhlbrück, Köbler, and Verbitsky analyze the structure of graphs with Weisfeiler-Leman dimension 2 [FKV21] and bounded color class size. Some of our structural lemmas can be seen as direct generalizations of results in their paper. A recent generalization of their complexity results regarding the WL-dimension can be found in [LRS24]. We should remark that in the two papers, just like in ours, the CFI-graphs appear innately.

A survey on descriptive complexity in particular with bounds related to the Weisfeiler-Leman dimension can be found in [PV09]. We previously supervised a Bachelor's thesis by Simon Lutz [Lut20] at TU Kaiserslautern that shows an upper bound of  $\lceil \frac{n}{3} \rceil + 2$ .

The term Weisfeiler-Leman dimension was coined by Grohe in his monograph [Gro17]. The main result implies that for non-trivial minor-closed graphs classes the Weisfeiler-Leman dimension is bounded. As remarked above, this is also true for graphs of bounded rank-width [GN23]. Recently, for several graph classes explicit bounds on the dimension have been proven, including planar graphs [KPS19], distance hereditary graphs [GNP23], and permutation graphs [GGP23].

In independent work, Kiefer and Neuen [KN24] recently prove an upper bound on the Weisfeiler-Leman dimension of  $\lceil \frac{n}{4} \rceil + o(n)$ . They also observe a lower bound of  $\frac{1}{96}n$ , slightly below the one we state here using essentially the same observations.

## 2 Preliminaries

Our overall proof makes ample use of coherent configurations, for which the basic concepts and notation are introduced in this section (see also [Bab16, CP19] for the broader theory).

Let  $\Omega$  be a finite set and  $A$  be a binary relation on  $\Omega$ . We set  $n := |\Omega|$ . Throughout the paper, we call the elements in  $\Omega$  *vertices* and the elements in  $A$  *arcs*. We also write  $vw$  instead of  $(v, w)$  to denote arcs. We set  $A^* := \{wv \mid vw \in A\}$  and call it the transpose relation of  $A$ . For  $v \in \Omega$ , the set  $vA := \{w \in \Omega \mid vw \in A\}$  is called the *neighborhood* of  $v$  under  $A$  and we define  $d^A(v) := |vA|$ . Given a set  $\Delta \subseteq \Omega$ , we also use the notation  $d_A(\Delta)$  if  $d^A(v)$  is independent of the choice of  $v \in \Delta$ . We denote  $\{(v, v) \mid v \in \Omega\}$  by  $1_\Omega$ . A *coloring of a set  $A$*  is the function  $\chi: A \rightarrow C$  where  $C$  is the set of *colors*. The *color* of  $a \in A$  is  $\chi(a)$  and the set of all colors is  $\chi(A)$ . The coloring induces a *color partition*  $\pi(\chi)$  on  $A$ . For a set  $\mathcal{A}$  of binary relations, we denote the set of all unions of these relations by  $\mathcal{A}^\cup$ .

**Graphs.** Given a finite set  $\Omega$  and a binary relation  $A$ , we call the pair  $(\Omega, A)$  a (*directed*) *graph*  $G$ . We denote the set of all vertices of  $G$  by  $\Omega(G)$  and the set of all arcs of  $G$  by  $A(G)$ . We call  $G$  *undirected* if  $A = A^*$ . Two vertices  $v, w \in \Omega$  are *adjacent* if  $vw \in A$  or  $wv \in A$ . Given  $\Delta \subseteq \Omega$ , the *subgraph of  $G$  induced by  $\Delta \subseteq \Omega$*  is  $(\Delta, A \cap \Delta^2)$  and is denoted by  $G[\Delta]$ . Given the graph  $G$  and a coloring  $\chi: A(G) \rightarrow C$ , we call  $(G, \chi)$  a *colored graph*.

**Isomorphisms.** An *isomorphism* between uncolored graphs  $G$  and  $H$  is a bijection  $\varphi : V(G) \rightarrow V(H)$  which preserves adjacency and non-adjacency, that is, for all  $v, w \in \Omega(G)$  we have  $vw \in A(G)$  if and only if  $\varphi(v)\varphi(w) \in A(H)$ . Given two colored graphs  $(G, \chi_G)$  and  $(H, \chi_H)$ , an isomorphism  $\varphi$  between  $G$  and  $H$  is called *color-permuting* if for all  $v, w, v', w' \in \Omega(G)$  it satisfies

$$\chi_G(vw) = \chi_G(v'w') \iff \chi_H(\varphi(v)\varphi(w)) = \chi_H(\varphi(v')\varphi(w')),$$

and *color-preserving* if for all  $v, w \in \Omega(G)$  we have  $\chi_G(vw) = \chi_H(\varphi(v)\varphi(w))$ . If there exists an isomorphism between  $G$  and  $H$ , then we call the graphs *isomorphic* and write  $G \cong H$ . Unless otherwise stated, we require all isomorphism between colored graphs to be color-preserving.

**Weisfeiler-Leman Algorithm.** Given  $k \in \mathbb{N}_{>0}$ , the  $k$ -dimensional Weisfeiler-Leman algorithm  $\mathbf{k}$ -WL is part of family of incomplete deciders for the *isomorphism problem*, which, given two uncolored graphs  $G$  and  $H$ , asks whether  $G \cong H$  holds. Given the colored graph  $(G, \chi)$ , the algorithm  $\mathbf{k}$ -WL determines for every  $(v_1, \dots, v_k) \in \Omega(G)^k$  an initial coloring  $\chi_0^{\mathbf{k}\text{-WL}}(v_1, \dots, v_k) := (\vartheta(v_1v_1), \vartheta(v_1v_2), \dots, \vartheta(v_kv_k))$  where

$$\vartheta(vw) := \begin{cases} (0, \chi(vw)) & \text{if } v = w, \\ (1, \chi(vw)) & \text{if } vw \in A(G), \\ (2, 0) & \text{if } vw \notin A(G). \end{cases}$$

Next, it iteratively computes  $\chi_{i+1}^{\mathbf{k}\text{-WL}}(v_1, \dots, v_k)$  defined by

$$(\chi_i^{\mathbf{k}\text{-WL}}(v_1, \dots, v_k), \{(\chi_i^{\mathbf{k}\text{-WL}}(w, v_2, \dots, v_k), \dots, \chi_i^{\mathbf{k}\text{-WL}}(v_1, \dots, v_{k-1}, w)) \mid w \in \Omega(G)\})$$

for all  $(v_1, \dots, v_k) \in \Omega(G)^k$ . This process stops if  $\pi(\chi_i^{\mathbf{k}\text{-WL}})$  is *stable under  $\mathbf{k}$ -WL*, that is  $\pi(\chi_i^{\mathbf{k}\text{-WL}}) = \pi(\chi_{i+1}^{\mathbf{k}\text{-WL}})$ . For the  $i$  at which the process stops, we define  $\chi_\infty^{\mathbf{k}\text{-WL}} := \chi_i^{\mathbf{k}\text{-WL}}$ .

If the initially given graph is uncolored, we start the algorithm on the monochromatic version of it. The algorithm  $\mathbf{k}$ -WL *distinguishes* graphs  $(G, \chi_G)$  and  $(H, \chi_H)$  if  $\{\{\chi_\infty^{\mathbf{k}\text{-WL}}(\bar{v}) \mid \bar{v} \in \Omega(G)^k\} \neq \{\{\chi_\infty^{\mathbf{k}\text{-WL}}(\bar{v}) \mid \bar{v} \in \Omega(H)^k\}\}$ . The notation  $(G, \chi_G) \simeq_k (H, \chi_H)$  indicates that the graphs are not distinguished by  $\mathbf{k}$ -WL. The algorithm  $\mathbf{k}$ -WL *identifies*  $(G, \chi_G)$  if it distinguishes the graph  $(G, \chi_G)$  from all non-isomorphic graphs. The *Weisfeiler-Leman dimension*  $\text{WLdim}((G, \chi_G))$  of a graph  $(G, \chi_G)$  is the minimal  $k \in \mathbb{N}$  such that  $\mathbf{k}$ -WL identifies  $G$ . Note that for a colored graph  $(G, \chi_G)$ , the WL-dimension only depends on  $\pi(\chi_G)$  and not on  $\chi_G$  itself.

Intuitively 2-WL works as follows: in each iteration, the algorithm counts for each arc  $(vw)$  in  $G$  the number of colored 2-paths from  $v$  to  $w$ , taken the colors into account. Many invariants are computed by 2-WL. It in particular it measures distances between vertices in  $G$  and is able to identify cycles and cliques. It is known that 2-WL identifies all uncolored graphs with order smaller than 16.

**Coherent configurations.** Let  $\Omega$  be a finite set and  $\mathcal{A}$  a partition of  $\Omega^2$ . Thus each  $A \in \mathcal{A}$  is a binary relation on  $\Omega$ . The pair  $(\Omega, \mathcal{A})$  is called a *coherent configuration*  $\mathfrak{X}$  on  $\Omega$  if the following conditions are satisfied:

$$(\text{CC1}) \quad 1_\Omega \in \mathcal{A}^\cup,$$

$$(\text{CC2}) \quad A^* \in \mathcal{A} \text{ for all } A \in \mathcal{A}, \text{ and}$$

$$(\text{CC3}) \quad \text{given } A, B, T \in \mathcal{A}, \text{ there is a constant } c_{A,B}^T \text{ such that } c_{A,B}^T = |vA \cap wB^*| \text{ is independent of the choice of } vw \in T.$$

We denote the set of all vertices of  $\mathfrak{X}$  by  $\Omega(\mathfrak{X})$ , and the set of all relations of  $\mathfrak{X}$  by  $\mathcal{A}(\mathfrak{X})$ . We call each  $A \in \mathcal{A}$  a *basis relation*,  $|\Omega|$  the *order* of  $\mathfrak{X}$ , and  $|\mathcal{A}|$  the *rank* of  $\mathfrak{X}$ . We call  $\mathfrak{X}$  *homogeneous* if  $1_\Omega \in \mathcal{A}$ .

We may interpret a coherent configuration  $\mathfrak{X}$  as a complete, colored graph  $(G_{\mathfrak{X}}, \chi_{\mathfrak{X}})$  where

$$\Omega(G_{\mathfrak{X}}) = \Omega(\mathfrak{X}), \quad A(G_{\mathfrak{X}}) = \Omega(\mathfrak{X})^2, \quad \text{and} \quad \pi_{G_{\mathfrak{X}}}(\chi) = \mathcal{A}.$$

We define  $\text{WLdim}(\mathfrak{X}) := \text{WLdim}((G_{\mathfrak{X}}, \chi_{\mathfrak{X}}))$ . Note that, the dimension only depends on the partition induced by  $\chi_{\mathfrak{X}}$  and not on the actual colors<sup>1</sup>.

**Lemma 2.1** (see [Bab16]). *Given a coherent configuration  $\mathfrak{X}$ , the color partition  $\pi(\chi_{\mathfrak{X}})$  is stable under 2-WL.*

**Fibers and interspaces.** Let  $\mathfrak{X}$  be a coherent configuration. A set  $\Delta \subseteq \Omega$  is called a *fiber* if  $1_\Delta \in \mathcal{A}$ . We denote the set of all fibers of  $\mathfrak{X}$  by  $F(\mathfrak{X})$ . By Property (CC1) the vertex set  $\Omega(\mathfrak{X})$  is partitioned by the collection of fibers  $F(\mathfrak{X})$ . A fiber  $\Delta$  is called a *singleton* if  $|\Delta| = 1$ .

Given  $\mathcal{R}, \mathcal{B} \in F(\mathfrak{X})^\cup$ , there is a unique subset  $\mathcal{A}' \subseteq \mathcal{A}(\mathfrak{X})$  which partitions  $\mathcal{R} \times \mathcal{B}$ . We call this partition the *interspace* between  $\mathcal{R}$  and  $\mathcal{B}$  and denote it by  $\mathfrak{X}[\mathcal{R}, \mathcal{B}]$ . Observe that  $A \in \mathfrak{X}[\mathcal{R}, \mathcal{B}]$  if and only if  $A^* \in \mathfrak{X}[\mathcal{B}, \mathcal{R}]$ . If  $|\mathfrak{X}[\mathcal{R}, \mathcal{B}]| = 1$  we call  $\mathfrak{X}[\mathcal{R}, \mathcal{B}]$  and  $\mathfrak{X}[\mathcal{B}, \mathcal{R}]$  *homogeneous*. If  $\mathcal{R} = \mathcal{B}$ , we shorten  $\mathfrak{X}[\mathcal{R}, \mathcal{R}]$  to  $\mathfrak{X}[\mathcal{R}]$ . For a union of fibers  $\mathcal{R}$ , we define  $\mathfrak{X} - \mathcal{R} = (\Omega(\mathfrak{X}) - \mathcal{R}, \mathfrak{X}[\Omega(\mathfrak{X}) - \mathcal{R}])$ .

Given two fibers  $R, B \in F(\mathfrak{X})$  and a basis relation  $A \in \mathfrak{X}[R, B]$ , we set  $d(A) := d^A(R)$  and  $d(R, B) := \min_{A \in \mathfrak{X}[R, B]} d(A)$ .

**Constituents.** For basis relation  $A$  of a coherent configuration  $\mathfrak{X}$ , we call the subgraph induced by  $A$  a *constituent* of  $\mathfrak{X}$ . To declare that the constituent  $G$  is *contained* in the interspace (or cell) between fibers  $R$  and  $B$ , we will abuse notation and write  $G \in \mathfrak{X}[R, B]$  instead of  $A(G) \in \mathfrak{X}[R, B]$ . If clear from context, abusing notation further, by  $G \in \mathfrak{X}[R, B]$  for explicitly given graphs  $G$  we will mean that  $\mathfrak{X}[R, B]$  contains a constituent that is isomorphic to  $G$ .

**Coherent closure and individualizations.** Given two coherent configurations  $\mathfrak{X}$  and  $\mathfrak{X}'$ , we say that  $\mathfrak{X}$  is at least as fine as  $\mathfrak{X}'$ , denoted by  $\mathfrak{X} \preceq \mathfrak{X}'$  if for all  $A \in \mathcal{A}(\mathfrak{X}')$  we have  $A \in \mathcal{A}(\mathfrak{X})^\cup$ .

Let  $\mathcal{A}$  be a collection of relations on  $\Omega$  which does not necessarily satisfy Property (CC3). The *coherent closure* of  $\mathcal{A}$  is the coarsest coherent configuration whose relations are each contained in some relation in  $\mathcal{A}$ . For vertices  $v_1, \dots, v_\ell \in \Omega$ , we set  $\mathfrak{X}_{v_1, \dots, v_\ell} := 2\text{-WL}(\mathcal{A}(\mathfrak{X}) \cup \{1_{v_1}, \dots, 1_{v_\ell}\})$ . Intuitively we individualize  $v_1, \dots, v_\ell$ , that is, force  $v_i$  to form its own color class (in terms of colored graphs) or equivalently to be its own fiber (in terms of coherent configuration).

**Theorem 2.2.** *Let  $\mathfrak{X}$  be a coherent configuration, and let  $v_1, \dots, v_\ell \in \Omega(\mathfrak{X})$ . Then  $\text{WLdim}(\mathfrak{X}) \leq \ell + \max\{2, \text{WLdim}(\mathfrak{X}_{v_1, \dots, v_\ell})\}$ .*

**Quotient graph.** The *quotient graph* of the coherent configuration  $\mathfrak{X}$  is the uncolored, undirected graph  $(F(\mathfrak{X}), \{RB \mid |\mathfrak{X}[R, B]| > 1\})$ . We denote the quotient graph of  $\mathfrak{X}$  by  $Q(\mathfrak{X})$ . Given  $R \in F(\mathfrak{X})$ , we define the *color degree* of  $R$  to be the number of fibers adjacent to  $R$  in  $Q(\mathfrak{X})$ .

We call a set  $\mathcal{S}$  of fibers *dominating* if every fiber of  $\mathfrak{X}$  is in  $\mathcal{S}$  or adjacent to at least one fiber of  $\mathcal{S}$  in  $Q(\mathfrak{X})$ .

<sup>1</sup>Treating configurations as uncolored objects, one may define a different notion of Weisfeiler-Leman dimension. However, the notion we use is the one suitable for application to graphs.

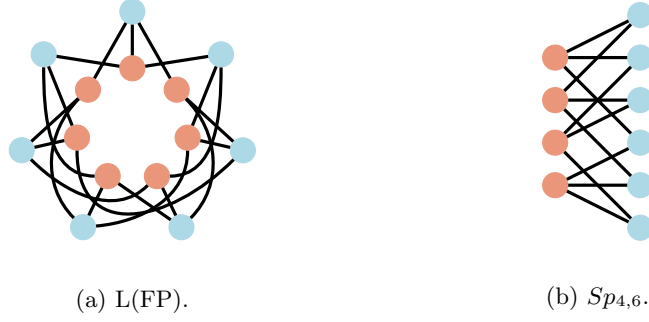


Figure 1: Special graphs.

**Special graphs.** We denote the disjoint union of  $s$  copies of a graph  $G$  by  $sG$  and write  $K_n$  for the complete, undirected graph of order  $n$  we write. We refer to the cycle of order  $n$  by  $C_n$ , and use  $\overrightarrow{C_n}$  for the directed version. We should remark that we use  $2K_3$  and  $2C_3$  both to refer to the same isomorphic graph. We refer to the rook graph on  $n \times n$  vertices by  $R_n$  and denote the Paley tournament on 7 vertices by  $PTr(7)$ .

Given two vertex sets  $R$  and  $B$  of size  $n_1$  and  $n_2$  respectively, the graph  $K_{n_1, n_2}$  is the complete, directed, bipartite graph from  $R$  to  $B$ . We call  $n_1 K_{1, n_2/n_1}$  a *star*, and  $n_1 K_{1, 1}$  a *matching*. Given disjoint fibers  $\{r_1, \dots, r_n\}$  and  $\{b_1, \dots, b_n\}$ , the set of arcs  $\{r_i b_i, r_i b_{i+1} \mid i \in \{1, \dots, n-1\}\} \cup \{r_n b_n, r_n b_1\}$  introduce a direction-alternating cycle of length  $2n$  in the interspace between the fibers. Abusing notation we denote it by  $C_{2n}$ .

The incidence graph (or Levi graph) of the Fano plane  $L(\text{FP})$  is a bipartite graph on two parts  $R$  and  $B$  each of size 7. The set of arcs  $A(L(\text{FP}))$  is, up to isomorphism, determined by following property: for each  $X \in \{R, B\}$  and  $v, v' \in X$  there is exactly one  $w \in (R \cup B) \setminus X$  for which  $vw, v'w \in A(L(\text{FP}))$ . The graph  $L(\text{FP})$  is shown in Figure 1(a). The graph  $Sp_{4,6}$  is the graph obtained from  $K_4$  by subdividing every edge. It is a bipartite graph with vertex set  $R \dot{\cup} B$  where  $|R| = 4$  and  $|B| = 6$ . For every  $\{r, r'\} \subsetneq R$  there is exactly one  $b \in B$  such that  $rb, r'b \in A(Sp_{4,6})$ . Figure 1(b) depicts  $Sp_{4,6}$ .

### 3 Critical Configurations

**Definition 3.1.** We call a coherent configuration  $\mathfrak{X}$  *critical* if  $\text{WLdim}(\mathfrak{X}) \geq 2$  and there is no set of fibers  $\mathcal{R} \subsetneq \mathcal{F}(\mathfrak{X})$  such that  $\text{WLdim}(\mathfrak{X} - \bigcup_{R \in \mathcal{R}} R) = \text{WLdim}(\mathfrak{X})$ .

**Lemma 3.2.** Let  $\mathfrak{X}$  be a coherent configuration. If  $\mathfrak{X}$  is critical, then  $\mathcal{Q}(\mathfrak{X})$  is connected.

*Proof.* TOPROVE 0 □

**Lemma 3.3.** Let  $\mathfrak{X}$  be a coherent configuration. If  $\mathfrak{X}$  is critical, then there are no distinct fibers  $R, B \in \mathcal{F}(\mathfrak{X})$  with  $d(B, R) = 1$ , that is, fibers with  $|R|K_{1, |B|} \in \mathfrak{X}[R, B]$ .

*Proof.* TOPROVE 1 □

**Lemma 3.4.** Let  $\mathfrak{X}$  be a coherent configuration. If  $\mathfrak{X}$  is critical, then there are no fibers  $R, B \in \mathcal{F}(\mathfrak{X})$  with  $|R| = |B|$  odd such that  $d(R, B) = 2$ . In particular, if  $\mathfrak{X}$  is critical, then there is no interspace  $tC_{2x} \in \mathfrak{X}[R, B]$  for positive integers  $t, x$  with  $x$  odd.

*Proof.* TOPROVE 2 □

A fiber  $R$  of a coherent configuration is called *large* if  $8 \leq |R|$ , *small* if  $4 \leq |R| \leq 7$ , and *tiny* if  $|R| \leq 3$ . Thinking of red, blue and yellow, we will typically use the letters  $R$ ,  $B$  and  $Y$  for fibers in general. We will use  $L$  or  $S$  to indicate that the fiber in question is large or small, respectively.

**Lemma 3.5.** *Let  $\mathfrak{X}$  be a coherent configuration. If  $\mathfrak{X}$  is critical, then there is no tiny fiber in  $F(\mathfrak{X})$ .*

*Proof.* **TOPROVE 3** □

Let  $\mathcal{R}$  be a union of fibers of  $\mathfrak{X}$ , and let  $\mathcal{B}$  be the union of fibers that are not in  $\mathcal{R}$  but adjacent (in  $Q(\mathfrak{X})$ ) to some fiber in  $\mathcal{R}$ . We call  $\mathcal{R}$  *restorable* if every automorphism of  $\mathfrak{X}[\mathcal{B}]$  that extends to an automorphism of  $\mathfrak{X} - \mathcal{R}$  also extends to an automorphism of  $\mathfrak{X}[\mathcal{R} \cup \mathcal{B}]$ .

**Lemma 3.6.** *Let  $\mathfrak{X}$  be a coherent configuration. If  $\mathfrak{X}$  is critical, then there is no restorable union of fibers that is non-dominating.*

*Proof.* **TOPROVE 4** □

Let  $R$  be a fiber of a coherent configuration. We say that a fiber  $Y$  distinct from  $R$  is *taken care of* (regarding restorability of  $R$ ) if for every  $y \in Y$  and  $U \in \mathfrak{X}[Y, R]$  there is a fiber  $B$  different from  $R$  and  $Y$ , a point  $b \in B$ , and an interspace  $U_B \in \mathfrak{X}[B, R]$  such that  $bU_B \subseteq yU$ .

**Lemma 3.7.** *Let  $\mathfrak{X}$  be a coherent configuration with distinct fibers  $R$  and  $Y$ . If  $Y$  is taken care of (regarding restorability of  $R$ ) and  $R$  is restorable in  $\mathfrak{X} - Y$ , then  $R$  is restorable in  $\mathfrak{X}$ .*

*Proof.* **TOPROVE 5** □

We call a vertex set  $R \subseteq \Omega(\mathfrak{X})$  a *module* if for all  $b \in \Omega(\mathfrak{X}) \setminus R$  and all  $A \in \mathcal{A}(\mathfrak{X})$  we have  $bA \cap R \in \{\emptyset, R\}$ .

**Lemma 3.8.** *Let  $\mathfrak{X}$  be a coherent configuration that is not homogeneous. If  $\mathfrak{X}$  is critical, then there is no small fiber that can be properly partitioned into at most 3 modules.*

*Proof.* **TOPROVE 6** □

**Theorem 3.9.** *Let  $\mathfrak{X}$  be a coherent configuration. Suppose  $(R, B, Y)$  is an induced path in  $Q(\mathfrak{X})$ ,  $r \in R$ ,  $y \in Y$ ,  $U \in \mathfrak{X}[R, B]$ , and  $U' \in \mathfrak{X}[Y, B]$ . Then  $d(U) \cdot d(U') = |B| \cdot |rU \cap yU'|$ .*

*In particular  $\text{lcm}\{|B|, d(U)\} > 1$  and  $\text{lcm}\{|B|, d(U')\} > 1$ . In particular  $|B|$  is not prime.*

*Proof.* **TOPROVE 7** □

## 4 Small fibers and their interspaces

Given an interspace  $\mathfrak{X}[R, B]$  between the fibers  $R$  and  $B$ , define  $T(\mathfrak{X}[R, B])$  to be the collection of basis relations  $A$  in a  $\mathfrak{X}[R, B]$  up to isomorphism while omitting transpose basis relation  $A^*$ ,  $1_R$ , and  $1_B$ .

**Lemma 4.1.** *Every small fiber  $R$  in a coherent configuration induces, up to isomorphism, one of 22 coherent configurations. These 22 options are given in Table 1.*

*Proof.* **TOPROVE 8** □

$ R $	$T(\mathfrak{X}[R])$
1	$(K_1)$
2	$(K_2)$
3	$(K_3)$
4	$(K_4), (C_4, 2K_2), (2K_2, 2K_2, 2K_2), (\vec{C_4}, 2K_2)$
5	$(K_5), (C_5, C_5), (\vec{C_5}, \vec{C_5}),$
6	$(K_6), (2K_3, K_{3,3}), (2\vec{C_3}, K_{3,3}), (3K_2, K_{2,2,2}), (3K_2, \vec{C_3}[K_2]),$ $(C_6, 3K_2, 2K_3), (\vec{C_6}, 2\vec{C_3}, 3K_2), (3K_2, 3K_2, 2\vec{C_3}, 3K_2)$
7	$(K_7), (C_7, C_7, C_7), (PT\mathfrak{r}(7)), (\vec{C_7}, \vec{C_7}, \vec{C_7})$

Table 1: Classification of all homogeneous coherent configuration of order  $n \leq 7$ .

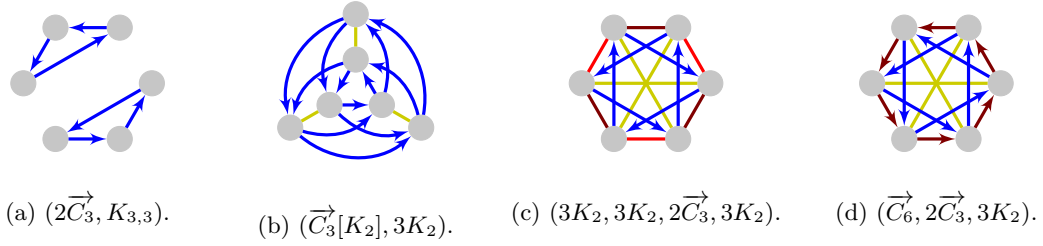


Figure 2: Coherent configurations of order 6 with directed edges.



$ R $	$ B $	$T(\mathfrak{X}[R, B])$
4	4	$(C_8, C_8), (2K_{2,2}, 2K_{2,2})$
4	6	$(Sp_{4,6}, Sp_{4,6}), (2K_{2,3}, 2K_{2,3})$
6	6	$(C_{12}, C_{12}, 3K_{2,2}), (2K_{3,3}, 2K_{3,3}), (3K_{2,2}, 3K_{2,2}, 3K_{2,2}), (3K_{2,2}, R \times B - 3K_{2,2})$
7	7	$(L(FP), R \times B - L(FP))$

Table 2: Classification of all  $\mathfrak{X}[R, B]$  with  $2 \leq |R| \leq |B| \leq 7$  in critical coherent configurations.

Note that for small fibers, the constituents of the induced homogeneous coherent configuration contained within a fiber determine the isomorphism type of said homogeneous coherent configuration. (This is not necessarily the case for larger fibers.)

**Lemma 4.2.** *Let  $\mathfrak{X}$  be a coherent configuration, and let  $(R, B)$  be an edge in  $Q(\mathfrak{X})$  such that  $2 \leq |R| \leq |B| \leq 7$ .*

- (1) *If  $|R| = 2$ , then  $|B| \in \{2, 4, 6\}$  and  $2K_{1, \frac{|B|}{2}} \in \mathfrak{X}[R, B]$ .*
- (2) *If  $|R| = 3$ , then  $|B| \in \{3, 6\}$  and  $3K_{1, \frac{|B|}{2}} \in \mathfrak{X}[R, B]$ .*
- (3) *If  $|R| = 4$ , then  $|B| \in \{4, 6\}$ . Furthermore, if  $|R| = 4$  and  $|B| = 4$ , then at least one of the constituents  $C_8$ ,  $4K_{1,1}$ , and  $2K_{2,2}$  is contained in  $\mathfrak{X}[R, B]$ , and if  $|R| = 4$  and  $|B| = 6$ , then at least one of constituents  $Sp_{4,6}$  and  $2K_{2,3}$  is contained in  $\mathfrak{X}[R, B]$ .*
- (4) *If  $|R| = 5$ , then  $|B| = 5$  and  $5K_{1,1} \in \mathfrak{X}[R, B]$ .*
- (5) *If  $|R| = 6$ , then  $|B| = 6$  and at least one of constituents  $6K_{1,1}$ ,  $C_{12}$ ,  $3K_{2,2}$ , or  $2K_{3,3}$  is contained in  $\mathfrak{X}[R, B]$ .*
- (6) *If  $|R| = 7$ , then  $|B| = 7$  and at least one of constituents  $7K_{1,1}$  and  $L(FP)$  is contained in  $\mathfrak{X}[R, B]$ .*

*Proof.* **TOPROVE 9** □

To classify all possible isomorphism types of interspaces between small fibers, we can examine the coarsest coherent configurations containing one of the graphs guaranteed to exist by Lemma 4.2 as constituent. Table 2 lists all isomorphism types between two small fibers in a critical coherent configuration. Note that the isomorphism type is determined by the isomorphism types of its constituents. (This is not necessarily the case for interspaces between larger fibers). For fibers of size 4 and 6 the classification can alternatively be derived from the upcoming Theorem 5.1.

**Lemma 4.3.** *Let  $\mathfrak{X}$  be a coherent configuration, and let  $(R, B)$  be an edge in  $Q(\mathfrak{X})$  with  $|R| \leq |B|$ , and suppose  $x \in \{4, 6\}$  and  $y \in \{2, 3\}$ .*

- (1) *If  $C_{2x} \in \mathfrak{X}[R, B]$ , then  $C_x \in \mathfrak{X}[R]$ .*
- (2) *If  $yK_{2,2} \in \mathfrak{X}[R, B]$ , then  $yK_2 \in \mathfrak{X}[R]$ .*
- (3) *If  $2K_{2,3} \in \mathfrak{X}[R, B]$ , then  $2K_2 \in \mathfrak{X}[R]$  and either  $2K_3 \in \mathfrak{X}[B]$  or  $2\vec{C}_3 \in \mathfrak{X}[B]$ .*
- (4) *If  $2K_{3,3} \in \mathfrak{X}[R, B]$ , then either  $2K_3 \in \mathfrak{X}[R]$  or  $2\vec{C}_3 \in \mathfrak{X}[B]$ .*

(5) If  $L(\text{FP}) \in \mathfrak{X}[R, B]$ , then either  $K_7 \in \mathfrak{X}[R]$  or  $\text{PTr}(7) \in \mathfrak{X}[R]$ .

(6) If  $\text{Sp}_{4,6} \in \mathfrak{X}[R, B]$ , then  $K_4 \in \mathfrak{X}[R]$  and either  $3K_2, K_{2,2,2} \in \mathfrak{X}[B]$  or  $3K_2, \overrightarrow{C_3}[K_2] \in \mathfrak{X}[B]$ .

*Proof.* **TOPROVE 10** □

## 5 Interspaces between large and small fibers

To study interspaces between large and small fibers we will introduce a specialized notation that concisely captures the attachment structure that vertices in the large fiber have in the small fiber. The challenge is to capture the entire information for all constituents simultaneously.

For a relation  $A$ , we define  $\text{ul}(A)$  to be  $A \cup A^*$ . We extend this notation to graphs and coherent configurations to obtain their *underlying undirected structure*. For a coherent configuration  $\mathfrak{X}$ , we set  $\text{ul}(\mathfrak{X})$  to be  $(\Omega(\mathfrak{X}), \{\text{ul}(A) \mid A \in \mathcal{A}(\mathfrak{X})\})$ . Observe that  $\text{ul}(\mathfrak{X})$  is not necessarily coherent. For a graph  $G$ , we set  $\text{ul}(G)$  to be  $(\Omega(G), \text{ul}(A(G)))$  and call it the *underlying undirected graph* of  $G$ .

Let  $L, S$  be distinct fibers of a coherent configuration  $\mathfrak{X}$ . We say the interspace  $\mathfrak{X}[L, S]$  contains the *interspace pattern*  $(G^1, d_1^1, d_2^1, \dots, d_{t_1}^1; \dots; G^k, d_1^k, d_2^k, \dots, d_{t_k}^k)$  if  $\sum_{j=1}^k t_j = |\mathfrak{X}[L, S]| - 1$ , there are distinct  $A^1, \dots, A^k \in \mathfrak{X}[S]$  such that for each  $i \in \{1, \dots, k\}$  we have  $\text{ul}(S, A^i) \cong G^i$ , and there are distinct  $U_1^i, \dots, U_{t_i}^i \in \mathfrak{X}[L, S]$  such that for all  $\ell \in L$  and  $j \in \{1, \dots, t_i\}$  the set  $\ell U_j^i$  induces a  $d_j^i$ -clique in  $\text{ul}(S, A^i)$ . We should emphasize that the sum of the  $t_i$  is  $|\mathfrak{X}[L, S]| - 1$  and thus we omit exactly one of the interspaces in the pattern. In fact, in general the missing interspace will not satisfy a clique condition. Also, it is a priori not clear that an interspace always has a pattern of this form we just defined.

Considering  $d_j^i = 3$ , if for all  $S' \subseteq S$  that induce a 3-clique in  $\text{ul}(S, A^1)$  the common neighborhood  $\bigcap_{s \in S'} sU_j^{i*}$  is not empty, then we mark the entry in the interspace pattern by  $^\dagger$ . This means there is a vertex in  $\ell \in L$  such that  $\ell U_j^i = S'$ . On the other hand, if for all  $S' \subseteq S$  inducing a 3-clique in  $\text{ul}(S, A^1)$  we have that  $S \setminus S'$  is also a 3-clique and exactly one of  $\bigcap_{s \in S'} sU_j^{i*}$  or  $\bigcap_{s \in S \setminus S'} sU_j^{i*}$  is not empty, then we mark the entry by  $^\ddagger$ . For example, the interspace pattern  $(K_{2,2,2}, 3)$  has two subpatterns, namely  $(K_{2,2,2}, 3^\dagger)$  and  $(K_{2,2,2}, 3^\ddagger)$ . In principle there could be other subcases, but it will follow from our classification that only these two subcases arise.

**Theorem 5.1.** *Let  $\mathfrak{X}$  be a critical coherent configuration and  $(L, S)$  be an edge in  $\mathcal{Q}(\mathfrak{X})$ . If  $|S| \in \{4, 6\}$ , then  $\mathfrak{X}[L, S]$  contains one of the following interspace patterns:*

- |                   |                                |                          |                                |
|-------------------|--------------------------------|--------------------------|--------------------------------|
| (1) $(K_4, 2)$    | (6) $(3K_2, 2)$                | (11) $(K_{3,3}, 2)$      | (16) $(K_{2,2,2}, 3^\dagger)$  |
| (2) $(2K_2, 2)$   | (7) $(3K_2, 2, 2)$             | (12) $(K_{3,3}, 2, 2)$   | (17) $(K_{2,2,2}, 3^\ddagger)$ |
| (3) $(C_4, 2)$    | (8) $(C_6, 2; 3K_2, 2)$        | (13) $(K_6, 3^\dagger)$  |                                |
| (4) $(K_6, 2)$    | (9) $(3K_2, 2; 3K_2, 2)$       | (14) $(K_6, 3^\ddagger)$ |                                |
| (5) $(K_6, 2, 2)$ | (10) $(K_{2,2,2}, 2; 3K_2, 2)$ | (15) $(2K_3, 3^\dagger)$ |                                |

*Proof.* **TOPROVE 11** □

We can observe that the interspace patterns in the theorem are mutually exclusive, and we conclude the following.

**Corollary 5.2.** *Let  $\mathfrak{X}$  be a critical coherent configuration, and let  $(L, S)$  be an edge in  $\mathcal{Q}(\mathfrak{X})$ . If  $|S| \in \{4, 6\}$ , then  $\mathfrak{X}[L, S]$  contains exactly one of interspace patterns listed in Theorem 5.1.*

*Proof.* **TOPROVE 12** □

We should stress that this classification of the interspaces between fibers  $L$  and  $S$  on interspace pattern is only possible because we consider small fibers with  $S$  of size 4 or 6.

For an interspace pattern  $\mathfrak{P}$  of  $\mathfrak{X}[L, S]$  we write  $A^i(\mathfrak{X}[L, S])$  and  $U_j^i(\mathfrak{X}[L, S])$  to denote  $A^i$  and  $U_j^i$  as given in the definition of interspace pattern. While the interspace pattern  $\mathfrak{P}$  is unique for interspace  $\mathfrak{X}[L, S]$ , the choice for  $A^i$  is only determined up to isomorphism. For a given coherent configuration we will always choose an arbitrary but fixed  $A^i$  and matching  $U_j^i$ .

**Lemma 5.3.** *Let  $\mathfrak{X}$  be a critical coherent configuration and  $(R, Y, B)$  a path in  $\mathcal{Q}(\mathfrak{X})$ . If either  $K_4 \in \mathfrak{X}[Y]$  or  $K_6 \in \mathfrak{X}[Y]$ , then  $\mathfrak{X}[R, B]$  is not homogeneous.*

*Proof.* **TOPROVE 13** □

Let  $(G^1, d_1^1, d_2^1, \dots, d_{t_1}^1; \dots; G^k, d_1^k, d_2^k, \dots, d_{t_k}^k)$  be the interspace pattern of  $\mathfrak{X}[L, S]$ , and let  $i \in \{1, \dots, k\}$  and  $j \in \{1, \dots, t_i\}$ . For  $U = U_j^i(\mathfrak{X}[L, S])$ , two vertices  $\ell, \ell' \in L$  are called *equivalent* with respect to  $S$  and  $U$  if  $\ell U = \ell' U$ . We denote the set of all equivalence classes with respect to  $S$  and  $U_j^i(\mathfrak{X}[L, S])$  by  $\mathcal{P}_j^i(L, S)$ . Further, we define

$$\mathcal{P}(L, S) := \bigwedge_{\substack{i \in \{1, \dots, k\} \\ j \in \{1, \dots, t_i\}}} \mathcal{P}_j^i(L, S),$$

that is,  $\mathcal{P}(L, S)$  is the meet of all partitions  $\mathcal{P}_j^i(L, S)$ . In other words,  $\mathcal{P}(L, S)$  is the coarsest partition which still finer than all  $\mathcal{P}_j^i(L, S)$ . Also note that  $\mathcal{P}(L, S) = \mathcal{P}_1^1(L, S)$  if  $|\mathfrak{X}[L, S]| = 2$ . For a union of small fibers  $\mathcal{S}$ , we define  $\mathcal{P}(L, \mathcal{S}) := \bigwedge_{S \in \mathcal{S}} \mathcal{P}(L, S)$ .

**Lemma 5.4.** *Let  $\mathfrak{X}$  be a critical coherent configuration,  $(L, S)$  an edge in  $\mathcal{Q}(\mathfrak{X})$  with  $|S| \in \{4, 6\}$  and suppose that  $\mathfrak{X}[L, S]$  has interspace pattern  $(G^1, d_1^1, \dots, d_{t_1}^1; \dots; G^k, d_1^k, \dots, d_{t_k}^k)$ . Let  $x$  be the number of 3-cliques in  $G^1$ .*

*Then  $|\mathcal{P}_1^1(L, S)| = |A^1(\mathfrak{X}[L, S])|$  if  $d_1^1 = 2$ ,  $|\mathcal{P}(L, S)| = x$  if  $d_1^1 = 3^\ddagger$ , and  $|\mathcal{P}(L, S)| = \frac{x}{2}$  if  $d_1^1 = 3^\ddagger$ .*

*Proof.* **TOPROVE 14** □

Table 3 gives an explicit overview of how the partition size  $|\mathcal{P}_1^1(L, S)|$  depends on the interspace patterns of Theorem 5.1.

**Lemma 5.5.** *Let  $\mathfrak{X}$  be a coherent configuration, and  $(R_1, L, R_2)$  a path in  $\mathcal{Q}(\mathfrak{X})$ . If  $|\mathcal{P}_1^1(L, R_1)|$  and  $|\mathcal{P}_1^1(L, R_2)|$  are coprime, then partitions  $\mathcal{P}_1^1(L, R_1)$  and  $\mathcal{P}_1^1(L, R_2)$  are fully intersecting, that is,  $P_1 \cap P_2 \neq \emptyset$  for all  $P_1 \in \mathcal{P}_1^1(L, R_1), P_2 \in \mathcal{P}_1^1(L, R_2)$ .*

*Proof.* **TOPROVE 15** □

Let  $\mathcal{S}$  be a union of small fibers each adjacent to a large fiber  $L$  in the quotient graph of a coherent configuration  $\mathfrak{X}$ . For  $U, U' \in \mathfrak{X}[L, \mathcal{S}]$ , we define a function  $\eta_{U, U'}: \mathcal{P}(L, \mathcal{S})^2 \rightarrow \mathbb{N}$  with  $(P, P') \mapsto |pU \cap p'U'|$  where  $p \in P, p' \in P'$  are arbitrary. We define  $\mathcal{Q}$  to be the set of equivalence classes on  $\mathcal{P}(L, \mathcal{S})^2$  such that for all  $Q \in \mathcal{Q}$  we have

$$(P_1, P_2), (P'_1, P'_2) \in Q \text{ if and only if } \eta_{U, U'}(P_1, P_2) = \eta_{U, U'}(P'_1, P'_2) \text{ for all } U, U' \in \mathfrak{X}[L, \mathcal{S}].$$

We call the coherent configuration  $(\mathcal{P}(L, \mathcal{S}), \mathcal{Q})$  the *partition structure*, denoted by  $\mathfrak{S}(L, \mathcal{S})$ . Intuitively speaking, the vertices of the partition structure correspond to the parts of  $\mathcal{P}(L, \mathcal{S})$ . For the arcs between a pair of parts, the relations represent what the arcs in  $L$  “know” about the intersection of common neighborhoods of their end vertices in  $\mathcal{S}$  (with respect to basis relations in  $\mathfrak{X}[L, \mathcal{S}]$ ). Observe that  $\mathfrak{S}(L, \mathcal{S})$  is indeed a coherent configuration itself.

$ \mathcal{P}_1^1(L, S) $	interspace pattern contained in $\mathfrak{X}[L, S]$
2	$(2K_2, 2), (2K_3, 3^\dagger)$
3	$(3K_2, 2), (3K_2, 2, 2), (3K_2, 2; 3K_2, 2)$
4	$(C_4, 2), (K_{2,2,2}, 3^\dagger)$
6	$(C_6, 2; 3K_2, 2), (K_4, 2)$
8	$(K_{2,2,2}, 3^\dagger)$
9	$(K_{3,3}, 2), (K_{3,3}, 2, 2)$
10	$(K_6, 3^\dagger)$
12	$(K_{2,2,2}, 2; 3K_2, 2)$
15	$(K_6, 2), (K_6, 2, 2)$
20	$(K_6, 3^\dagger)$

Table 3: Overview of  $|\mathcal{P}_1^1(L, S)|$  depending on the interspace pattern in  $\mathfrak{X}[L, S]$ .

**Lemma 5.6.** *Let  $\mathfrak{X}$  be a critical coherent configuration and  $L, S \in F(\mathfrak{X})$ . For  $|\mathcal{P}_1^1(L, S)| \leq 8$ , the coherent configuration  $\mathfrak{X}[S]$  together with the interspace pattern of  $\mathfrak{X}[L, S]$  uniquely determines the partition structure  $\mathfrak{S}(L, S)$ .*

*Proof.* **TOPROVE 16** □

We should comment that the interspace pattern and the isomorphism type of the small fiber  $S$  does not uniquely determine the partition structure  $\mathfrak{S}(L, S)$  if  $|\mathcal{P}_1^1(L, S)| \geq 9$ . Further, we remark that in a critical coherent configuration the interspace pattern might restrict the possible constituents in the small fiber, as we have shown in the proof of Theorem 5.2: For example, if an interspace  $\mathfrak{X}[L, S]$  has interspace pattern  $(C_4, 2)$  or  $(C_6, 2; 3K_2, 2)$ , then by coherence  $\overrightarrow{C_4} \notin \mathfrak{X}[S]$  (respectively  $\overrightarrow{C_6} \notin \mathfrak{X}[S]$ ) since  $U_1^1(\mathfrak{X}[L, S])$  would not be a basis relation. For similar reasons  $2\overrightarrow{C_3} \notin \mathfrak{X}[S]$  if  $\mathfrak{X}[L, S]$  has interspace pattern  $(K_{3,3}, 2)$ . For each interspace pattern and isomorphism type of configuration  $\mathfrak{X}[S]$ , the possible partition structures in a critical coherent configuration  $\mathfrak{X}$  are listed in Table 4.

## 6 Restorable fibers

**Lemma 6.1.** *Let  $\mathfrak{X}$  be a critical coherent configuration, and  $(R, S, Y)$  be a path in  $Q(\mathfrak{X})$  with  $S$  being small.*

- (1) *If both  $\mathfrak{X}[R, S]$  and  $\mathfrak{X}[Y, S]$  have interspace pattern  $(C_4, 2)$ , then there is a union  $A$  of basis relations in  $\mathfrak{X}[R, Y]$  such that  $(R \dot{\cup} Y, A) \cong 4K_{\frac{|R|}{4}, \frac{|Y|}{4}}$ .*
- (2) *If both  $\mathfrak{X}[R, S]$  and  $\mathfrak{X}[Y, S]$  have interspace pattern  $(C_6, 2; 3K_2, 2)$ , then there is a union  $A$  of basis relations in  $\mathfrak{X}[R, Y]$  such that  $(R \dot{\cup} Y, A) \cong 6K_{\frac{|R|}{6}, \frac{|Y|}{6}}$ .*

*In particular, fibers  $R$  and  $Y$  are large.*

*Proof.* **TOPROVE 17** □

**Lemma 6.2.** *Let  $\mathfrak{X}$  be a critical coherent configuration and  $S \in F(\mathfrak{X})$  be a size-4 fiber. If  $\{S\}$  is not dominating and  $|\text{ul}(\mathfrak{X}[S])| = 3$ , then there are fibers  $R, R'$  adjacent to  $S$  in  $Q(\mathfrak{X})$  such that  $\mathfrak{X}[R, S]$  has interspace pattern  $(C_4, 2)$  and  $\mathfrak{X}[R', S]$  has interspace pattern  $(2K_2, 2)$ .*

Interspace pattern of $\mathfrak{X}[L, S]$	$T(\mathfrak{X}[S])$	$T(\mathfrak{S}(L, S))$
$(K_4, 2)$	$(K_4)$	$(3K_2, K_{2,2,2})$
$(2K_2, 2)$	$(2K_2, 2K_2, 2K_2), (C_4, 2K_2), (\vec{C}_4, 2K_2)$	$(K_2)$
$(C_4, 2)$	$(2K_2, C_4)$	$(2K_2, C_4)$
$(3K_2, 2)$	$(3K_2, K_{2,2}), (C_6, 2C_3, 3K_2),$	$(K_3)$
$(3K_2, 2, 2)$	$(3K_2, K_{2,2,2}), (C_6, 2C_3, 3K_2),$	$(3K_2, 3K_2, 2\vec{C}_3, 3K_2)$
	$(3K_2, \vec{C}_3[K_2]), (\vec{C}_6, 2\vec{C}_3, 3K_2)$	$(\vec{C}_3)$
$(C_6, 2; 3K_2, 2)$	$(C_6, 2C_3, 3K_2)$	$(C_6, 2C_3, 3K_2)$
$(K_{3,3}, 2)$	$(2K_3, K_{3,3})$	$(R_3, R_3)$
$(2K_3, 3^\dagger)$	$(2C_3, K_{3,3}), (2\vec{C}_3, K_{3,3}), (C_6, 2C_3, 3K_2),$ $(\vec{C}_6, 2\vec{C}_3, 3K_2), (3K_2, 3K_2, 2\vec{C}_3, 3K_2)$	$(K_2)$
$(3K_2, 2; 3K_2, 2)$	$(3K_2, 3K_2, 2\vec{C}_3, 3K_2)$	$(C_3)$
$(K_{2,2,2}, 3^\dagger)$	$(3K_2, K_{2,2,2}), (3K_2, \vec{C}_3[K_2])$	$(2K_4, 4K_2, K_{4,4} - 4K_2)$
$(K_{2,2,2}, 3^\ddagger)$	$(3K_2, K_{2,2,2}), (3K_2, \vec{C}_3[K_2])$	$(K_4)$

Table 4: Partition structure  $\mathfrak{S}(L, S)$  determined  $\mathfrak{X}[S]$  and  $\mathfrak{X}[L, S]$  in a critical coherent configuration  $\mathfrak{X}$ .

*Proof.* **TOPROVE 18** □

The *color-disjoint union*  $\mathfrak{X}$  of two coherent configurations  $\mathfrak{X}_1$  and  $\mathfrak{X}_2$  is the coherent configuration  $(\Omega(\mathfrak{X}_1) \dot{\cup} \Omega(\mathfrak{X}_2), A(\mathfrak{X}_1) \cup A(\mathfrak{X}_2) \cup \{\Omega(\mathfrak{X}_1) \times \Omega(\mathfrak{X}_2)\})$ . Observe that  $\Omega(\mathfrak{X}_i) \in F(\mathfrak{X})^\cup$  for all  $i \in \{1, 2\}$ . Thus there is no automorphism of  $\mathfrak{X}$  which maps  $\Omega(\mathfrak{X}_1)$  to  $\Omega(\mathfrak{X}_2)$  and vice versa.

**Lemma 6.3.** *Let  $\mathfrak{X}$  be critical coherent configuration, and let  $R_0, R_1 \in F(\mathfrak{X})$  be distinct. If  $\{R_0, R_1\}$  is not dominating, then  $C_8 \notin \mathfrak{X}[R_0, R_1]$ .*

*Proof.* **TOPROVE 19** □

**Lemma 6.4.** *Let  $\mathfrak{X}$  be a critical coherent configuration and  $S \in F(\mathfrak{X})$  a fiber that is not dominating. For every  $A \in \mathfrak{X}[S]$  with  $(S, A) \cong 2K_2$  there are  $R \in F(\mathfrak{X})$  and  $U \in \mathfrak{X}[R, S]$  such that for all  $r \in R$  the set  $rU$  induces a 2-clique in  $(S, A)$ .*

*In particular, if  $\mathfrak{X}$  is critical and  $|\mathfrak{X}[S]| = 4$ , then  $\text{cdeg}(S) \geq 3$ .*

*Proof.* **TOPROVE 20** □

**Lemma 6.5.** *Let  $\mathfrak{X}$  be a coherent configuration, and let  $(R, B, Y)$  be path in  $Q(\mathfrak{X})$  with  $|B| = 6$ .*

- (1) *If  $3K_{\frac{|R|}{3}, 2} \in \mathfrak{X}[R, B]$  and  $3K_{\frac{|Y|}{3}, 2} \in \mathfrak{X}[Y, B]$ , then there is a union  $U$  of relations in  $\mathfrak{X}$  such that  $(R \dot{\cup} Y, U) \cong 3K_{\frac{|R|}{3}, \frac{|Y|}{3}}$ .*
- (2) *If both  $\mathfrak{X}[R, B]$  and  $\mathfrak{X}[Y, B]$  have interspace pattern  $(2K_3, 3^\dagger)$ , then there is a union  $U$  of relations in  $\mathfrak{X}$  such that  $(R \dot{\cup} Y, U) \cong 2K_{\frac{|R|}{2}, \frac{|Y|}{2}}$ .*

		$\mathfrak{X}[B, Y]$		
		$C_{12}$	$3K_{2,2}$	$2K_{3,3}$
$\mathfrak{X}[R, B]$	$C_{12}$	$6K_{1,1}$	$3K_{2,2}$	$R \times B$
	$3K_{2,2}$	$3K_{2,2}$	$3K_{2,2}$	$R \times B$
	$2K_{3,3}$	$R \times B$	$R \times B$	$2K_{3,3}$

Table 5: For a path  $(R, B, Y)$  of size 6 fibers, the isomorphism type of a constituent in  $\mathfrak{X}[R, Y]$  depending on the isomorphism types of the constituents contained in  $\mathfrak{X}[R, B]$  and  $\mathfrak{X}[B, Y]$ .

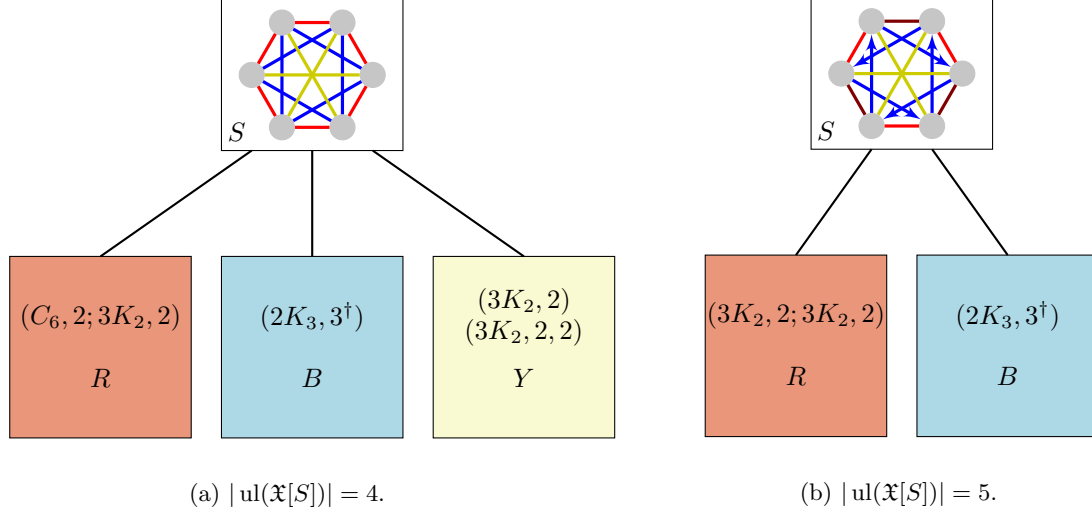


Figure 3: Visualisation of the neighbor subset partition in Lemma 6.7.

*Proof.* **TOPROVE 21** □

For the convenience of the reader, Table 5 gives an overview of the results of Lemmas 6.1 and 6.5 for size 6 fibers.

**Lemma 6.6.** *Let  $\mathfrak{X}$  be a critical coherent configuration and let  $S \in \mathcal{F}(\mathfrak{X})$  be distinct with  $|S| = 6$ . If  $\{S\}$  is not dominating, then there is no interspace pattern  $\mathfrak{P}$  in the following list such that for all fibers  $R$  adjacent to  $S$  in  $\mathcal{Q}(\mathfrak{X})$  the interspace  $\mathfrak{X}[R, S]$  has this interspace pattern  $\mathfrak{P}$ :*

- |                         |                          |                           |
|-------------------------|--------------------------|---------------------------|
| (1) $(2K_3, 3^\dagger)$ | (3) $(3K_2, 2, 2)$       | (5) $(3K_2, 2; 3K_{2,2})$ |
| (2) $(3K_2, 2)$         | (4) $(C_6, 2; 3K_{2,2})$ | (6) $(K_{3,3}, 2)$        |

*In particular, if  $\{S\}$  is not dominating, then there is at least one fiber  $R$  adjacent to  $S$  in  $\mathcal{Q}(\mathfrak{X})$  such that  $\mathfrak{X}[R, S]$  has neither interspace pattern  $(3K_2, 2)$  nor  $(3K_2, 2, 2)$ .*

*Proof.* **TOPROVE 22** □

**Lemma 6.7.** *Let  $\mathfrak{X}$  be a critical coherent configuration and  $S \in \mathcal{F}(\mathfrak{X})$  be a fiber such that  $|S| = 6$ ,  $|\text{ul}(\mathfrak{X}[S])| > 3$ , and  $\{S\}$  is not dominating. There exist disjoint subsets  $\mathcal{R}, \mathcal{B}, \mathcal{Y}$  partitioning the neighborhood of  $S$  in  $\mathcal{Q}(\mathfrak{X})$  with  $\mathcal{R}, \mathcal{B}$  non-empty (and  $\mathcal{Y}$  possibly empty) such that*

- (1) for all  $R \in \mathcal{R}$  the interspace  $\mathfrak{X}[R, S]$  has either interspace pattern  $(C_6, 2; 3K_2, 2)$  or interspace pattern  $(3K_2, 2; 3K_2, 2)$ ,
- (2) for all  $B \in \mathcal{B}$  the interspace  $\mathfrak{X}[B, S]$  has interspace pattern  $(2K_3, 3^\dagger)$ ,
- (3) for all  $Y \in \mathcal{Y}$  the interspace  $\mathfrak{X}[Y, S]$  has either interspace pattern  $(3K_2, 2)$  or interspace pattern  $(3K_2, 2, 2)$ .

In particular,  $\text{cdeg}(S) \geq 2$ . (See Figure 3)

*Proof.* **TOPROVE 23** □

**Lemma 6.8.** *Let  $\mathfrak{X}$  be a critical coherent configuration and  $R, S \in F(\mathfrak{X})$  be distinct fibers of size 6 such that  $C_{12} \in \mathfrak{X}[R, S]$  and  $\{R, S\}$  is not dominating. There exist disjoint, non-empty subsets  $\mathcal{B}, \mathcal{Y}$  of the neighborhood of  $S$  without  $R$  in  $Q(\mathfrak{X})$  such that*

- (1) for all  $B \in \mathcal{B}$  the interspace  $\mathfrak{X}[B, S]$  has interspace pattern  $(2K_3, 3^\dagger)$  and
- (2) for all  $Y \in \mathcal{Y}$  the interspace  $\mathfrak{X}[Y, S]$  has either interspace pattern  $(3K_2, 2)$  or interspace pattern  $(3K_2, 2, 2)$ .

In particular,  $\text{cdeg}(S) \geq 3$ .

*Proof.* **TOPROVE 24** □

Next we deal with interspaces isomorphic to  $Sp_{4,6} \in \mathfrak{X}[R, B]$  for fibers  $R, B$  with  $|R| < |B|$ . Such an interspace has a different pattern depending on the perspective. Interspace  $\mathfrak{X}[R, B]$  has interspace pattern  $(K_{2,2,2}, 3^\ddagger)$ , while interspace  $\mathfrak{X}[B, R]$  has interspace pattern  $(K_4, 2)$ .

**Lemma 6.9.** *Let  $\mathfrak{X}$  be critical coherent configuration, and let  $R, S \in F(\mathfrak{X})$  fibers such that  $\mathfrak{X}[R, S]$  has interspace pattern  $(K_{2,2,2}, 3^\ddagger)$ .*

- (1) If  $\text{cdeg}(R) = 1$ , then  $\{S\}$  is dominating and for all fibers  $Y$  adjacent to  $S$  in  $Q(\mathfrak{X})$  other than  $R$  the interspace  $\mathfrak{X}[Y, S]$  has either interspace pattern  $(3K_2, 2)$  or interspace pattern  $(3K_2, 2, 2)$ .
- (2) If  $Sp_{4,6} \in \mathfrak{X}[R, S]$  and all fibers of  $\mathfrak{X}$  are small, then  $\text{WLdim}(\mathfrak{X}) = 2$ .

*Proof.* **TOPROVE 25** □

**Lemma 6.10.** *If  $\mathfrak{X}$  is a critical coherent configuration, then there is no path  $(R, B, Y)$  in  $Q(\mathfrak{X})$  such that  $L(\text{FP}) \in \mathfrak{X}[R, B]$  and  $L(\text{FP}) \in \mathfrak{X}[B, Y]$ .*

*Proof.* **TOPROVE 26** □

**Lemma 6.11.** *Let  $\mathfrak{X}$  be a critical coherent configuration and  $(S, L, S')$  be a path in  $Q(\mathfrak{X})$  such that  $\mathfrak{X}[L, S]$  and  $\mathfrak{X}[L, S']$  both have interspace pattern  $(3K_2, 2)$ . If  $|L| = 9$ ,  $\text{cdeg}(L) = 2$ ,  $|\text{ul}(\mathfrak{X}[L])| = 4$ , and  $\mathcal{P}(L, S)$  and  $\mathcal{P}(L, S')$  are fully intersecting, then  $\{L\}$  is dominating.*

*Proof.* **TOPROVE 27** □

## 7 Weisfeiler-Leman dimension of graphs with small fibers

**Theorem 7.1.** *Let  $\mathfrak{X}$  be a critical coherent configuration. If all fibers of  $\mathfrak{X}$  are small, then either*

- (1) *all fibers have the same size, or*
- (2) *for all  $R \in F(\mathfrak{X})$  we have  $|R| \in \{4, 6\}$ .*

*Proof.* **TOPROVE 28** □

**Lemma 7.2.** *Let  $\mathfrak{X}$  be a critical coherent configuration such that all fibers of  $\mathfrak{X}$  have size 4. A fiber  $R$  of  $\mathfrak{X}$  and all fibers adjacent to it in  $Q(\mathfrak{X})$  split entirely into tiny fibers in  $\mathfrak{X}_r$  where  $r \in R$ .*

*In particular, if  $\text{cdeg}(R) \geq 4$ , then  $\text{WLdim}(\mathfrak{X}) = 2$  or there is a critical coherent configuration  $\mathfrak{X}'$  for which  $\text{WLdim}(\mathfrak{X}) \leq 1 + \text{WLdim}(\mathfrak{X}_r) = 1 + \text{WLdim}(\mathfrak{X}')$  with  $|\Omega(\mathfrak{X})| \leq |\Omega(\mathfrak{X}')| - 20$ .*

*Proof.* **TOPROVE 29** □

The lemma essentially allows us to reduce the degree of the quotient graph. For graphs with maximum degree at most 3, we can make use of a bound by Fomin and Høie on the pathwidth of the graph.

**Theorem 7.3** (see [FH06]). *For every  $\epsilon > 0$ , there exists  $n_0 \in \mathbb{N}$  such that every graph of maximum degree at most 3 with at least  $n$  vertices satisfies  $\text{tw}(G) \leq \text{pw}(G) \leq \frac{n}{6} + \epsilon$ .*

We use the theorem to bound the Weisfeiler-Leman dimension for certain graphs as follows.

**Theorem 7.4.** *Let  $\mathfrak{X}$  be a critical coherent configuration on  $n$  vertices such that all fibers of  $\mathfrak{X}$  have size 4 and there is no interspace containing a constituent isomorphic to  $C_8$ . If all fibers of  $\mathfrak{X}$  have at most 3 neighbors in  $Q(\mathfrak{X})$ , then  $\text{WLdim}(\mathfrak{X}) \leq \frac{n}{24} + \mathcal{O}(1)$ .*

*Proof.* **TOPROVE 30** □

**Corollary 7.5.** *Let  $\mathfrak{X}$  be a critical coherent configuration on  $n$  vertices such that all fibers of  $\mathfrak{X}$  have size 4. Then  $\text{WLdim}(\mathfrak{X}) \leq \frac{n}{20} + \mathcal{O}(1)$ .*

*Proof.* **TOPROVE 31** □

**Lemma 7.6.** *Let  $\mathfrak{X}$  be a critical coherent configuration such that all fibers of  $\mathfrak{X}$  have size 5. Then  $|F(\mathfrak{X})| = 1$  and  $\text{WLdim}(\mathfrak{X}) \leq 2$ .*

*Proof.* **TOPROVE 32** □

**Lemma 7.7.** *If  $\mathfrak{X}$  is a critical coherent configuration whose fibers each have size 4 or 6, then  $\text{WLdim}(\mathfrak{X}) \leq \frac{n}{20} + \mathcal{O}(1)$ .*

*Proof.* **TOPROVE 33** □

**Lemma 7.8.** *Let  $\mathfrak{X}$  be a critical coherent configuration such that all fibers of  $\mathfrak{X}$  have size 7. Then  $|F(\mathfrak{X})| \leq 2$  and  $\text{WLdim}(\mathfrak{X}) \leq 3$ .*

*Proof.* **TOPROVE 34** □

**Theorem 7.9.** *Let  $\mathfrak{X}$  be a critical coherent configuration such that all fibers of  $\mathfrak{X}$  are small. Then  $\text{WLdim}(\mathfrak{X}) \leq \frac{n}{20} + \mathcal{O}(1)$ .*

*Proof.* **TOPROVE 35** □



## 8 Limiting fiber sizes

For each interspace  $\mathfrak{X}[R, B]$  we arbitrarily choose a basis relation whose degree is maximal. A *non-maximal* basis relation is basis relation different than the chosen one. Similarly for each fiber  $R$  we chose a constituent with maximal degree. A non-maximal constituent is one that is not the chosen one. Define  $\overline{d(B, R)}$  to be the maximum degree among all degrees of non-maximal basis relations in  $\mathfrak{X}[R, B]$ . Set  $\overline{d(B, R)} = 0$  if  $|\mathfrak{X}[R, B]| = 1$ .

Define the *max-modules* of a coherent configuration to be the connected components of the graph in which the edges are formed by all non-maximal basis relations. (The max-modules are sections in the sense of [Gol83] and colored modules in the sense of [Sch17].) It is not difficult to see that the Weisfeiler-Leman dimension of a coherent configuration is at most the maximum of 2 and the maximum Weisfeiler-Leman dimension of its max-modules.

**Lemma 8.1** (Color valence limit, (Zemlyachenko, see [Bab81])). *Let  $\mathfrak{X}$  be a critical coherent configuration, let  $d \in \mathbb{N}$  be a positive integer. There is a set of vertices  $S$  of size at most  $\frac{2n}{d}$  so that in the components of  $\mathfrak{X}_S$  we have  $\overline{d(B, R)} \leq d$  for all (not necessarily distinct) fibers  $R, B \in F(\mathfrak{X}_S)$ .*

*Proof.* **TOPROVE 36** □

**Lemma 8.2** (Fiber size limit). *Let  $\mathfrak{X}$  be a critical coherent configuration, and let  $c, d \in \mathbb{N}$  be positive integers. There is a set  $S$  of vertices of size at most  $\frac{2n}{d} + \frac{dn}{c}$  such that for each max-module  $\overline{\mathfrak{X}}$  of  $\mathfrak{X}_S$  we have  $|R| \leq c$  and  $\overline{d(R, B)} \leq d$  for all (not necessarily distinct) fibers  $R, B \in F(\overline{\mathfrak{X}})$ .*

*Proof.* **TOPROVE 37** □

**Lemma 8.3.** *If all fibers of a coherent configuration  $\mathfrak{X}$  have size at most  $t$ , then  $\text{WLdim}(\mathfrak{X}) \leq t \cdot \text{tw}(\text{Q}(\mathfrak{X}))$ , where  $\text{tw}(\text{Q}(\mathfrak{X}))$  is the treewidth of the quotient graph.*

*Proof.* **TOPROVE 38** □

## 9 A potential function

For a coherent configuration  $\mathfrak{X}$  let  $\text{Par}(\mathfrak{X})$  be the triple  $(n_\ell, k_\ell, n_s)$ , where  $n_\ell$  is the number of vertices in large fibers,  $k_\ell$  is the number of large fibers and  $n_s$  is the number of small fibers.

Let  $\hat{f}: \mathbb{N}^3 \rightarrow \mathbb{N}$  be the function that assigns to the triple  $(n_\ell, k_\ell, n_s)$  the maximum Weisfeiler-Leman dimension of a coherent configuration  $\mathfrak{X}$  with  $\text{Par}(\mathfrak{X}) = (n_\ell, k_\ell, n_s)$ .

Define the function  $\tau: \mathbb{N}^3 \rightarrow \mathbb{R}$  with  $(n_\ell, k_\ell, n_s) \mapsto \frac{3n_\ell + n_s - 8k_\ell}{20}$ .

Note that the parameters and function  $\tau$  are additive in the sense that we can compute the parameters separately for a configuration induced by an arbitrary partition of the fibers and then add them up.

We now define  $f$  to be the function that assigns to a coherent configuration  $\mathfrak{X}$  the maximum Weisfeiler-Leman dimension among all coherent configurations  $\mathfrak{X}'$  for which  $\tau(\text{Par}(\mathfrak{X}')) \leq \tau(\text{Par}(\mathfrak{X}))$ . (In some sense  $f$  is a monotonization of  $\hat{f}$ .)

Using this notation, note that the choice of potential function guarantees that if  $\mathfrak{X}'$  is finer than  $\mathfrak{X}$  then  $f(\mathfrak{X}') \leq f(\mathfrak{X})$ .

We define  $\tilde{f}: \mathbb{R} \rightarrow \mathbb{R}$  with  $r \mapsto \max\{f(\mathfrak{X}) \mid \tau(\text{Par}(\mathfrak{X})) \leq r\}$ .

Our goal in the upcoming sections will be to show that  $\text{WLdim}(\mathfrak{X}) \leq \tau(\text{Par}(\mathfrak{X})) + o(n)$ . A key technique will be reductions that show that  $\text{WLdim}(\mathfrak{X}) \leq t + f(\tau(\text{Par}(\mathfrak{X}')))$  for some coherent configuration  $\mathfrak{X}'$  obtained from  $\mathfrak{X}$  in a controlled manner, which will give a recursive bound of the form  $\text{WLdim}(\mathfrak{X}) \leq t + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - t)$ .

## 10 Local reductions

In this section we provide local reductions, trading progress in the potential with individualizations. This eventually allows us to restrict the structure of the coherent configuration in three ways: First, it limits the number of neighbors a large fibers can have in the quotient graph. Second, we eliminate certain interspace patterns. For some other interspace patterns we show that there can be at most one small neighboring fiber attached with this pattern. Finally we reduce the quotient graph degree of fibers of size 4 or 6.

Each reduction proceeds as follows. We examine the local structure of neighbors of a fiber. In particular, we describe the interspace patterns and the color degree. We individualize one or two carefully chosen vertices, take the coherent closure, and restore criticality. Afterwards we weigh the number of individualized vertices against the progress in the terms of the potential function  $\tau$  we made due to fibers that split.

To write the proofs of this section more concisely, we introduce an auxiliary result as follows. For triples  $(a, b, c), (a', b', c') \in \mathbb{N}^3$ , we will use the notation  $(a, b, c) \preceq (a', b', c')$  to denote that  $\tau(a, b, c) \leq \tau(a', b', c')$ . Define a function  $h: \mathbb{R}^+ \rightarrow \mathbb{R}$  by setting  $h(a) = \max\{-0.4, -3 \cdot \frac{\lceil 8a \rceil}{20}\}$ .

**Lemma 10.1.** *Let  $\mathfrak{X}, \mathfrak{X}'$  be a coherent configurations so that  $\mathfrak{X}$  consists of a single large fiber and  $\mathfrak{X}' \preceq \mathfrak{X}$ . Let  $R_1, \dots, R_u$  be the fibers of  $\mathfrak{X}'$  and set  $\{t_i \mid i \in \{1, \dots, u\}, |R_i| = t_i \cdot |\mathfrak{X}|\}$ . Then  $\tau(\text{Par}(\mathfrak{X}')) \leq \tau(\text{Par}(\mathfrak{X})) + \sum_{i=1}^u h(t_i) + 0.4$ .*

*Proof.* **TOPROVE 39** □

We should remark that while Lemma 10.1 is sufficient to calculate the progress made in  $L$  in many cases, in some other case a closer examination produces better results on the progress. This is especially true if we know conditions on the size of  $L$  or sizes of the vertex sets into which  $L$  splits.

**Lemma 10.2.** *Let  $\mathfrak{X}$  be a critical coherent configuration,  $R \in F(\mathfrak{X})$ , and  $\{L_1, L_2, L_3\}$  a set of large fibers adjacent to  $R$  in  $Q(\mathfrak{X})$ . If  $R$  is a large fiber or  $|\text{ul}(\mathfrak{X}[R])| \geq 3$ , then  $\text{WLdim}(\mathfrak{X}) \leq 1 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 1)$ .*

*Proof.* **TOPROVE 40** □

For a coherent configuration  $\mathfrak{X}$ , we define  $Q^L(\mathfrak{X})$  and  $Q^S(\mathfrak{X})$  to be the subgraph of  $Q(\mathfrak{X})$  induced by the set of all large fibers and small fibers, respectively.

**Lemma 10.3.** *If  $\mathfrak{X}$  is a coherent configuration in which no large fiber has at least 3 large neighbors. Then the connected components of  $Q^L(\mathfrak{X})$  have treewidth at most 2.*

*Proof.* **TOPROVE 41** □

**Theorem 10.4.** *Let  $\mathfrak{X}$  be a critical coherent configuration and suppose  $L, S \in F(\mathfrak{X})$  with  $L$  large and  $S$  small. If the set  $\{S\}$  is not dominating and the interspace  $\mathfrak{X}[L, S]$  has one of the interspace patterns  $(C_6, 2; 3K_2, 2)$ ,  $(3K_2, 2; 3K_2, 2)$ ,  $(3K_2, 2, 2)$ ,  $(K_{2,2,2}, 2; 3K_2, 2)$ ,  $(K_{2,2,2}, 3^\dagger)$ ,  $(K_{3,3}, 2)$ , or  $(K_{3,3}, 2, 2)$ , then  $\text{WLdim}(\mathfrak{X}) \leq 1 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 1)$ .*

*Proof.* **TOPROVE 42** □

**Theorem 10.5.** *Let  $\mathfrak{X}$  be a critical coherent configuration and let  $(S, L, S')$  be a path in  $Q(\mathfrak{X})$  with  $L$  large and  $S, S'$  small. If  $\mathfrak{X}[L, S]$  has interspace pattern  $(K_{2,2,2}, 3^\dagger)$  and  $\mathfrak{X}[L, S']$  has one of the interspace patterns  $(K_{2,2,2}, 3^\dagger)$ ,  $(C_4, 2)$ ,  $(3K_2, 2)$ ,  $(2K_2, 2)$ , or  $(2K_3, 3^\dagger)$ , then  $\text{WLdim}(\mathfrak{X}) \leq 1 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 1.1)$ .*

*Proof.* **TOPROVE 43** □

**Theorem 10.6.** *Let  $\mathfrak{X}$  be a critical coherent configuration, and let  $(S, L, S')$  be a path in  $Q(\mathfrak{X})$  with  $L$  large and  $S, S'$  small. If  $\{S\}$  and  $\{S'\}$  are not dominating and both  $\mathfrak{X}[L, S]$  and  $\mathfrak{X}[L, S']$  have one of the interspace patterns  $(C_4, 2)$ ,  $(3K_2, 2)$ ,  $(2K_2, 2)$ , or  $(2K_3, 3^\dagger)$ , then  $\text{WLdim}(\mathfrak{X}) \leq 1 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 1)$ .*

*Proof.* **TOPROVE 44** □

Given  $F \in F(\mathfrak{X})$ , we denote the number of large (respectively small) fibers adjacent to  $F$  in  $Q(\mathfrak{X})$  by  $\text{cdeg}_L(F)$  (respectively  $\text{cdeg}_S(F)$ ).

**Lemma 10.7.** *Let  $\mathfrak{X}$  be a critical coherent configuration, and let  $S$  be a size-4 fiber of  $\mathfrak{X}$  such that  $K_4 \notin \mathfrak{X}[S]$ . If*

- $\text{cdeg}(S) = 3$  and  $\text{cdeg}_L(S) \geq 1$  or
- $\text{cdeg}(S) \geq 4$ ,

*then  $\text{WLdim}(\mathfrak{X}) \leq 1 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 1)$ .*

*Proof.* **TOPROVE 45** □

**Lemma 10.8.** *Let  $\mathfrak{X}$  be a critical coherent configuration, and let  $(L, S)$  be an edge in  $Q(\mathfrak{X})$  such that  $\mathfrak{X}[L, S]$  has interspace pattern  $(K_{2,2,2}, 3^\dagger)$ . If  $\text{cdeg}(S) \geq 3$  and  $\{L, S\}$  is not dominating, then*

- $\text{WLdim}(\mathfrak{X}) \leq 1 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 1)$  or
- $\text{WLdim}(\mathfrak{X}) \leq 2 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 2.1)$

*Proof.* **TOPROVE 46** □

**Lemma 10.9.** *Let  $\mathfrak{X}$  be a critical coherent configuration, and let  $(L, S, S')$  be a path in  $Q(\mathfrak{X})$  such that interspace  $\mathfrak{X}[L, S]$  has interspace pattern  $(K_{2,2,2}, 3^\dagger)$  and interspace  $\mathfrak{X}[S, S']$  has interspace pattern  $(3K_2, 2; 3K_2, 2)$ . If  $\{L, S\}$  is not dominating, then*

- $\text{WLdim}(\mathfrak{X}) \leq 1 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 1.05)$  or
- $\text{WLdim}(\mathfrak{X}) \leq 2 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 2)$ .

*Proof.* **TOPROVE 47** □

**Lemma 10.10.** *Let  $\mathfrak{X}$  be a critical coherent configuration, and let  $S$  be a size 6 fiber of  $\mathfrak{X}$  such that  $K_6 \notin \mathfrak{X}[S]$ . If  $\text{cdeg}(S) \geq 3$ , then  $\text{WLdim}(\mathfrak{X}) \leq 1 + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - 1)$ .*

*Proof.* **TOPROVE 48** □

## 11 The structure of reduced coherent configurations

Let  $\mathfrak{X}$  be a coherent configuration. We call a small fiber  $S$  *relevant* if  $|S| \in \{4, 6\}$  and  $|\text{ul}(\mathfrak{X}[S])| > 2$ , and *irrelevant* otherwise. For a fiber  $R$  of  $\mathfrak{X}$ , we define  $\text{cdeg}_{\text{rS}}(R)$  to be number of relevant small fibers adjacent to  $R$  in  $\text{Q}(\mathfrak{X})$ .

**Lemma 11.1.** *Let  $\mathfrak{X}$  be a critical coherent configuration. If  $S$  is an irrelevant small fiber of  $\mathfrak{X}$ , then  $\text{cdeg}_S(S) \leq 1$  and the set of fibers adjacent to  $S$  in  $\text{Q}(\mathfrak{X})$  induces a clique in  $\text{Q}(\mathfrak{X})$ .*

*Furthermore, if  $\text{cdeg}_L(S) \geq 1$ , then  $S$  is not adjacent to an relevant in  $\text{Q}(\mathfrak{X})$ .*

*Proof.* **TOPROVE 49** □

**Lemma 11.2.** *Let  $\mathfrak{X}$  be a critical coherent configuration and there is  $t \in \mathbb{N}$  such that all fibers of  $\mathfrak{X}$  have size at most  $t$ . If there is a set of fibers  $\mathcal{R}$  which is dominating and there are at most  $u \in \mathbb{N}$  large fibers adjacent to a fiber of  $\mathcal{R}$ , then  $\text{WLdim}(\mathfrak{X}) \leq 2 + \sum_{R \in \mathcal{R}} |R| + u \cdot t$ .*

*Proof.* **TOPROVE 50** □

**Definition 11.3.** For  $t \in \mathbb{N}$ , we call a coherent configuration  $\mathfrak{X}$  *t-reduced*, if it satisfies all of the following properties.

- (1)  $\mathfrak{X}$  is critical.
- (2) Every fiber  $R$  has at most  $t$  vertices, i.e.,  $|R| \leq t$ .
- (3) All large fibers  $L$  have at most two large neighbors, that is  $\text{cdeg}_L(L) \leq 2$ .
- (4) A large fiber has at most one relevant small neighbor, that is  $\text{cdeg}_{\text{rS}}(L) \leq 1$ .
- (5) If a relevant small fiber  $S$  has a large neighbor (i.e., if  $\text{cdeg}_L(S) \geq 1$ ), then  $|\text{ul}(\mathfrak{X}[S])| = 3$  and its color degree is at most 2 (i.e.,  $\text{cdeg}(S) \leq 2$ ).
- (6) For every  $s \in S$  in a relevant small fiber  $S$  with three small neighbors (i.e.,  $\text{cdeg}_S(S) \geq 3$ ), we have that  $S$  is discrete in  $\mathfrak{X}_s$ .

**Lemma 11.4.** *Let  $\mathfrak{X}$  be a critical coherent configuration whose largest fibers have size  $t$ . Then*

- $\text{WLdim}(\mathfrak{X}) \leq z + \tilde{f}(\tau(\text{Par}(\mathfrak{X})) - z)$  for positive integer  $z$ ,
- $\text{WLdim}(\mathfrak{X}) \leq 10 + 2 \cdot t$ , or
- $\mathfrak{X}$  is *t-reduced* (i.e., satisfies Definition 11.3).

*Proof.* **TOPROVE 51** □

## 12 A global argument

**Lemma 12.1.** *Let  $\mathfrak{X}$  be a critical coherent configuration that is *t-reduced* (Definition 11.3), and let  $(L, S, L')$  be an induced path in  $\text{Q}(\mathfrak{X})$  such that  $L, L'$  are large and  $S$  is small and relevant. There are  $s_1, \dots, s_{|S|/2-1} \in S$  such that either vertex sets  $L \cup S$  and  $L'$  or the vertex sets  $L$  and  $L' \cup S$  are homogeneously connected in  $(\mathfrak{X}[L \cup S \cup L'])_{s_1, \dots, s_{|S|/2-1}}$ . Furthermore  $L$  or  $L'$  has at least size 9 if  $|S| = 6$ .*

*Proof.* **TOPROVE 52** □

**Lemma 12.2.** *Let  $\mathfrak{X}$  be a critical coherent configuration that is  $t$ -reduced (Definition 11.3), and let  $(L_0, S_0, S_1, L_1)$  be an induced path in  $Q(\mathfrak{X})$  such that  $L_0, L_1$  are large and  $S_0, S_1$  are small and relevant. There are  $s_1, \dots, s_{|S_0|/2-1} \in S_0$  such that vertex sets  $L_0 \cup S_0$  and  $L_1 \cup S_1$  are homogeneously connected in  $(\mathfrak{X}[L_0 \cup S_0 \cup S_1 \cup L_1])_{s_1, \dots, s_{|S_0|/2-1}}$ .*

*Proof.* **TOPROVE 53** □

**Lemma 12.3.** *Let  $\mathfrak{X}$  be a  $t$ -reduced critical coherent configuration with parameters  $Par(\mathfrak{X}) = (n_\ell, k_\ell, n_s)$ . There is a set  $M$  of size  $q \leq k_\ell$  such that each connected component of  $Q(\mathfrak{X}_M)$*

- *induces a configuration of Weisfeiler-Leman dimension at most  $3t$  or*
- *has only small fibers and order at most  $n_s - r_1 \cdot 3 - r_2 \cdot 4$ , where  $r_1 \cdot 8.5 + r_2 \cdot 8 \leq n_\ell$ .*

*Proof.* **TOPROVE 54** □

**Corollary 12.4.** *Let  $\mathfrak{X}$  be a critical coherent configuration that is  $t$ -reduced. Suppose  $Par(\mathfrak{X}) = (n_\ell, k_\ell, n_s)$  are the parameters of  $\mathfrak{X}$ .*

*Then  $WLdim(\mathfrak{X}) \leq \frac{2}{20}n_\ell + \frac{1}{20}n_s + 3t \leq \tau(n_\ell, k_\ell, n_s) + 3t$ .*

*Proof.* **TOPROVE 55** □

## 13 Proof of the main theorem

**Theorem 13.1.** *Let  $\mathfrak{X}$  be a coherent configuration in which every fiber has size at most  $t$  with parameters  $Par(\mathfrak{X}) = (n_\ell, k_\ell, n_s)$ , then  $WLdim(\mathfrak{X}) \leq \tau(n_\ell, k_\ell, n_s) + 3 \cdot t + 6$ .*

*Proof.* **TOPROVE 56** □

**Theorem 13.2.** *Let  $\mathfrak{X}$  be a coherent configuration on  $n$  vertices then  $WLdim(\mathfrak{X}) \leq 3/20n + o(n)$ .*

*Proof.* **TOPROVE 57** □

## 14 Lower bound

In this section we assemble several known results to obtain an improved lower bound for the maximum Weisfeiler-Leman dimension of graphs in terms of their order.

**Theorem 14.1.** *A random cubic graph asymptotically almost surely has a treewidth of at least  $0.042011151 \cdot n$ .*

*Proof.* **TOPROVE 58** □

It is well known that cubic graphs of high treewidth yield graphs with high Weisfeiler-Leman dimension via the CFI-construction. Specifically, we have the following relation.

**Theorem 14.2** (Consequence of [DR07, Theorem 3]). *For a graph  $G$ , the CFI-graph  $CFI(G)$  with base graph  $G$  satisfies  $WLdim(G) \geq tw(G)$ .*

We remark that there are two versions of the CFI-constructions used in the literature. One where each vertex is replaced by a gadget of order 10 and one where each vertex is replaced by a gadget of order 4. (See [Für01, NS17, Lic23] for more information.) These versions are very similar and the theorem, as well as many other theorems, hold for either of them. The difference between the constructions is that the former produces CFI-graphs of order  $10|G|$ , while the latter produces graphs of order  $4|G|$ . In the terminology of our current paper, the first version produces coherent configurations with fibers of size 2, so a non-critical configuration. Removal of the tiny fibers yields the other construction.

The two theorems combine as follows.

**Corollary 14.3.** *The maximum Weisfeiler-Leman dimension for cubic graphs of order  $n$  is at least  $0.0105027 \cdot n - o(n)$ .*

In the light of our discussions in Section 4 we remark that, due to the nature of the CFI-construction, the statement is also true for graphs of color class size 4.

## References

- [AKRV17] Vikraman Arvind, Johannes Köbler, Gaurav Rattan, and Oleg Verbitsky. Graph isomorphism, color refinement, and compactness. *Comput. Complex.*, 26(3):627–685, 2017.
- [AM13] Albert Atserias and Elitza N. Maneva. Sherali-adams relaxations and indistinguishability in counting logics. *SIAM J. Comput.*, 42(1):112–137, 2013.
- [AMR<sup>+</sup>19] Albert Atserias, Laura Mancinska, David E. Roberson, Robert Sámal, Simone Severini, and Antonios Varvitsiotis. Quantum and non-signalling graph isomorphisms. *J. Comb. Theory, Ser. B*, 136:289–328, 2019.
- [Bab79] László Babai. Lectures on graph isomorphism. Mimeographed lecture notes, Department of Computer Science, University of Toronto, October 1979. Lecture notes.
- [Bab81] László Babai. Moderately exponential bound for graph isomorphism. In *Fundamentals of Computation Theory, FCT’81, Proceedings of the 1981 International FCT-Conference, Szeged, Hungary, August 24-28, 1981*, volume 117 of *Lecture Notes in Computer Science*, pages 34–50. Springer, 1981.
- [Bab16] László Babai. Graph isomorphism in quasipolynomial time [extended abstract]. In *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.
- [CFI92] Jin-yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Comb.*, 12(4):389–410, 1992.
- [CP19] G. Chen and I. Ponomarenko. *Lectures on Coherent Configurations*. Central China Normal University Press, 2019.
- [DGR18] Holger Dell, Martin Grohe, and Gaurav Rattan. Lovász meets weisfeiler and leman. In *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 40:1–40:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

- [DN16] Zdenek Dvorák and Sergey Norin. Strongly sublinear separators and polynomial expansion. *SIAM J. Discret. Math.*, 30(2):1095–1101, 2016.
- [DR07] Anuj Dawar and David Richerby. The power of counting logics on restricted classes of finite structures. In *Computer Science Logic, 21st International Workshop, CSL 2007, 16th Annual Conference of the EACSL, Lausanne, Switzerland, September 11–15, 2007, Proceedings*, volume 4646 of *Lecture Notes in Computer Science*, pages 84–98. Springer, 2007.
- [Dvo10] Zdenek Dvorák. On recognizing graphs by numbers of homomorphisms. *J. Graph Theory*, 64(4):330–342, 2010.
- [FH06] Fedor V. Fomin and Kjartan Høie. Pathwidth of cubic graphs and exact algorithms. *Information Processing Letters*, 97(5):191 – 196, 2006.
- [FKV21] Frank Fuhlbrück, Johannes Köbler, and Oleg Verbitsky. Identifiability of graphs with small color classes by the Weisfeiler-Leman algorithm. *SIAM J. Discret. Math.*, 35(3):1792–1853, 2021.
- [Für01] Martin Fürer. Weisfeiler-lehman refinement requires at least a linear number of iterations. In *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8–12, 2001, Proceedings*, volume 2076 of *Lecture Notes in Computer Science*, pages 322–333. Springer, 2001.
- [GGP23] Jin Guo, Alexander L. Gavriluk, and Ilia Ponomarenko. On the Weisfeiler-Leman dimension of permutation graphs. *CoRR*, abs/2305.15861, 2023. arXiv.
- [GN23] Martin Grohe and Daniel Neuen. Canonisation and definability for graphs of bounded rank width. *ACM Trans. Comput. Log.*, 24(1):6:1–6:31, 2023.
- [GNP23] Alexander L. Gavriluk, Roman Nedela, and Ilia Ponomarenko. The weisfeiler-leman dimension of distance-hereditary graphs. *Graphs Comb.*, 39(4):84, 2023.
- [GO15] Martin Grohe and Martin Otto. Pebble games and linear equations. *J. Symb. Log.*, 80(3):797–844, 2015.
- [Gol83] Mark K. Goldberg. A nonfactorial algorithm for testing isomorphism of two graphs. *Discret. Appl. Math.*, 6(3):229–236, 1983.
- [Gro17] Martin Grohe. *Descriptive Complexity, Canonisation, and Definable Graph Structure Theory*, volume 47 of *Lecture Notes in Logic*. Cambridge University Press, 2017.
- [HM00] Akihide Hanaki and Izumi Miyamoto. Classification of primitive association schemes of order up to 22. *Kyushu Journal of Mathematics*, 54(1):81–86, 2000.
- [HM03] Akihide Hanaki and Izumi Miyamoto. Classification of association schemes of small order. *Discret. Math.*, 264(1-3):75–80, 2003.
- [IL90] Neil Immerman and Eric Lander. Describing graphs: a first-order approach to graph canonization. In *Complexity theory retrospective*, pages 59–81. Springer, New York, 1990.
- [KN24] Sandra Kiefer and Daniel Neuen. Bounding the weisfeiler-leman dimension via a depth analysis of i/r-trees. *CoRR*, abs/2402.03274, 2024. arXiv.

- [KPS19] Sandra Kiefer, Ilya Ponomarenko, and Pascal Schweitzer. The weisfeiler-leman dimension of planar graphs is at most 3. *J. ACM*, 66(6):44:1–44:31, 2019.
- [KSS22] Sandra Kiefer, Pascal Schweitzer, and Erkal Selman. Graphs identified by logics with counting. *ACM Trans. Comput. Log.*, 23(1):1:1–1:31, 2022.
- [KW14] Brett Kolesnik and Nick Wormald. Lower bounds for the isoperimetric numbers of random regular graphs. *SIAM J. Discret. Math.*, 28(1):553–575, 2014.
- [Lic23] Moritz Lichter. *Continuing the Quest for a Logic Capturing Polynomial Time - Potential, Limitations, and Interplay of Current Approaches*. PhD thesis, Technische Universität Darmstadt, Darmstadt, July 2023.
- [LRS24] Moritz Lichter, Simon Raßmann, and Pascal Schweitzer. Computational complexity of the weisfeiler-leman dimension. *CoRR*, abs/2402.11531, 2024. arXiv.
- [Lut20] Simon Lutz. Bounds on the Weisfeiler-Leman dimension of general graphs. Bachelor’s thesis, TU Kaiserslautern, 2020.
- [MLM<sup>+</sup>23] Christopher Morris, Yaron Lipman, Haggai Maron, Bastian Rieck, Nils M. Kriege, Martin Grohe, Matthias Fey, and Karsten M. Borgwardt. Weisfeiler and Leman go machine learning: The story so far. *Journal of Machine Learning Research*, 24:1–59, 2023.
- [NS17] Daniel Neuen and Pascal Schweitzer. Benchmark graphs for practical graph isomorphism. In *25th Annual European Symposium on Algorithms, ESA 2017, September 4-6, 2017, Vienna, Austria*, volume 87 of *LIPIcs*, pages 60:1–60:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.
- [OS14] Yota Otachi and Pascal Schweitzer. Reduction techniques for graph isomorphism in the context of width parameters. In *Algorithm Theory - SWAT 2014 - 14th Scandinavian Symposium and Workshops, Copenhagen, Denmark, July 2-4, 2014. Proceedings*, volume 8503 of *Lecture Notes in Computer Science*, pages 368–379. Springer, 2014.
- [PV09] Oleg Pikhurko and Oleg Verbitsky. Logical complexity of graphs: A survey. In *Model Theoretic Methods in Finite Combinatorics - AMS-ASL Joint Special Session, Washington, DC, USA, January 5-8, 2009*, volume 558 of *Contemporary Mathematics*, pages 129–180. American Mathematical Society, 2009.
- [PVV06] Oleg Pikhurko, Helmut Veith, and Oleg Verbitsky. The first order definability of graphs: Upper bounds for quantifier depth. *Discret. Appl. Math.*, 154(17):2511–2529, 2006.
- [Sch17] Pascal Schweitzer. Towards an isomorphism dichotomy for hereditary graph classes. *Theory Comput. Syst.*, 61(4):1084–1127, 2017.
- [Wei76] Boris Weisfeiler. *On construction and identification of graphs*. Lecture Notes in Mathematics, Vol. 558. Springer-Verlag, Berlin-New York, 1976.