



# The Weisfeiler-Leman Algorithm and Recognition of Graph Properties

Frank Fuhlbrück<sup>1</sup>, Johannes Köbler<sup>1</sup>, Iliia Ponomarenko<sup>2,3</sup>,  
and Oleg Verbitsky<sup>1</sup>(✉)

<sup>1</sup> Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany  
{fuhlbfra,koebler,verbitsk}@informatik.hu-berlin.de

<sup>2</sup> Central China Normal University, Wuhan, China

<sup>3</sup> Steklov Institute of Mathematics at St. Petersburg, St. Petersburg, Russia  
inp@pdmi.ras.ru

**Abstract.** The  $k$ -dimensional Weisfeiler-Leman algorithm ( $k$ -WL) is a very useful combinatorial tool in graph isomorphism testing. We address the applicability of  $k$ -WL to recognition of graph properties. Let  $G$  be an input graph with  $n$  vertices. We show that, if  $n$  is prime, then vertex-transitivity of  $G$  can be seen in a straightforward way from the output of 2-WL on  $G$  and on the vertex-individualized copies of  $G$ . This is perhaps the first non-trivial example of using the Weisfeiler-Leman algorithm for recognition of a natural graph property rather than for isomorphism testing. On the other hand, we show that, if  $n$  is divisible by 16, then  $k$ -WL is unable to distinguish between vertex-transitive and non-vertex-transitive graphs with  $n$  vertices unless  $k = \Omega(\sqrt{n})$ .

## 1 Introduction

The  $k$ -dimensional Weisfeiler-Leman algorithm ( $k$ -WL), whose original, 2-dimensional version [20] appeared in 1968, has played a prominent role in isomorphism testing already for a half century. Given a graph  $G$  with vertex set  $V$ ,  $k$ -WL computes a canonical coloring  $\text{WL}_k(G)$  of the Cartesian power  $V^k$ . Let  $\widehat{\text{WL}}_k(G)$  denote the multiset of colors appearing in  $\text{WL}_k(G)$ . The algorithm decides that two graphs  $G$  and  $H$  are isomorphic if  $\widehat{\text{WL}}_k(G) = \widehat{\text{WL}}_k(H)$ , and that they are non-isomorphic otherwise. While a negative decision is always correct, Cai, Fürer, and Immerman [5] constructed examples of non-isomorphic graphs  $G$  and  $H$  with  $n$  vertices such that  $\widehat{\text{WL}}_k(G) = \widehat{\text{WL}}_k(H)$  unless  $k = \Omega(n)$ . Nevertheless, a constant dimension  $k$  suffices to correctly decide isomorphism for many special classes of graphs (when  $G$  is in the class under consideration and  $H$  is arbitrary). For example,  $k = 2$  is enough if  $G$  is an interval graph [10],  $k = 3$  is enough for planar graphs [15], and there is a constant  $k = k(M)$  sufficient for all graphs not containing a given graph  $M$  as a minor [13]. Last but not least,  $k$ -WL is an

---

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

© Springer Nature Switzerland AG 2021

T. Calamoneri and F. Corò (Eds.): CIAC 2021, LNCS 12701, pp. 245–257, 2021.

[https://doi.org/10.1007/978-3-030-75242-2\\_17](https://doi.org/10.1007/978-3-030-75242-2_17)

important component in Babai's quasipolynomial-time algorithm [3] for general graph isomorphism.

In the present paper, we initiate a discussion of the applicability of  $k$ -WL to recognition of graph properties rather than to testing isomorphism. That is, given a single graph  $G$  as input, we are interested in knowing which properties of  $G$  can be *easily* detected by looking at  $\text{WL}_k(G)$  or, in other words, for which decision problems the execution of  $k$ -WL on an input graph is a reasonable preprocessing step. Of course, some regularity properties are recognized in a trivial way. For example,  $G$  is strongly regular if and only if 2-WL splits  $V^2$  just in the diagonal  $\{(u, u) : u \in V\}$ , the adjacency relation of  $G$ , and the complement.

For a graph property  $\mathcal{P}$ , we use the same character  $\mathcal{P}$  to denote also the class of all graphs possessing this property. While the multiset of canonical colors  $\widehat{\text{WL}}_k(G)$  retains the isomorphism type of the original graph  $G$  only if  $k$  is sufficiently large, the coloring  $\text{WL}_k(G)$  of  $V^k$  does this for every  $k$ . This means that, at least implicitly,  $\text{WL}_k(G)$  contains the information about all properties  $\mathcal{P}$  of  $G$ . It is, however, a subtle question whether any certificate of the membership of  $G$  in  $\mathcal{P}$  can be extracted from  $\text{WL}_k(G)$  efficiently. Even when the isomorphism type of every graph in  $\mathcal{P}$  is known to be identifiable by  $k$ -WL for some  $k$ , we can only be sure that  $k$ -WL distinguishes  $\mathcal{P}$  from its complement, in the following sense: If  $G \in \mathcal{P}$  and  $H \notin \mathcal{P}$ , then  $\widehat{\text{WL}}_k(G) \neq \widehat{\text{WL}}_k(H)$ . However, given the last inequality, we might never know whether  $G \in \mathcal{P}$  and  $H \notin \mathcal{P}$  or whether  $H \in \mathcal{P}$  and  $G \notin \mathcal{P}$ . As a particular example, the fact that 2-WL decides isomorphism of interval graphs or that 3-WL decides isomorphism of planar graphs does not seem to imply, on its own, any efficient recognition algorithm for these classes.

We address the applicability of  $k$ -WL to recognition of properties saying that a graph is highly symmetric.

**Deciding Vertex-Transitivity.** A graph  $G$  is *vertex-transitive* if every vertex can be taken to any other vertex by an automorphism of  $G$ . It is unknown whether the class of vertex-transitive graphs is recognizable in polynomial time. The isomorphism problem for vertex-transitive graphs reduces to their recognition problem, and its complexity status is also open. In the case of graphs with a prime number  $p$  of vertices, a polynomial-time recognition algorithm is known due to Muzychuk and Tinhofer [18]. Their algorithm uses 2-WL as preprocessing and then involves a series of algebraic-combinatorial operations to find a Cayley presentation of the input graph. It is known [19] that, if  $p$  is prime, then every vertex-transitive graph with  $p$  vertices is *circulant*, i.e., a Cayley graph of the cyclic group of order  $p$ . Our first result, Theorem 1, shows a very simple, purely combinatorial way to recognize vertex-transitivity of a graph  $G$  with  $p$  vertices. Indeed, vertex-transitivity can immediately be detected by looking at the outdegrees of the monochromatic digraphs in  $\text{WL}_2(G)$  and  $\text{WL}_2(G_u)$  for all copies of  $G$  with an individualized vertex  $u$ . Our algorithm takes time  $O(p^4 \log p)$ , which is somewhat better than the running time  $O(p^5 \log^2 p)$  of the algorithm presented in [18]. However, we believe that the main beneficial factor of our approach is its conceptual and technical simplicity.

Note that the research on circulant graphs has a long history; see, e.g., [2, 17]. This class of graphs can be recognized in polynomial time [7], but whether or not this can be done by means of  $k$ -WL is widely open. The dimension  $k = 2$  would clearly suffice if the algorithm could identify a cyclic order of the vertices in an input graph corresponding to its cyclic automorphism. However, it would be too naive to hope for this because such an order is, in general, not unique, not preserved by automorphisms and, hence, not canonical, even when the number of vertices is prime.

The analysis of our algorithm is based on the theory of coherent configurations (we provide a digest of main concepts in Sect. 3.1). In fact, our exposition, apart from the well-known facts on circulants of prime order, uses only several results about the *schurity property* of certain coherent configurations.

**Lower Bounds for the WL Dimension.** Since the work of Muzychuk and Tinhofer [18] the polynomial-time recognizability of vertex-transitive graphs with a prime number of vertices remains state-of-the-art in the sense that, to the best of our knowledge, no polynomial-time algorithm is currently known that recognizes vertex-transitivity on all  $n$ -vertex input graphs for infinitely many composite numbers  $n$ . Motivated by this fact, we complement our algorithmic result by exploring the limitations of the  $k$ -WL-based combinatorial approach to vertex-transitivity. We prove that, if  $n$  is divisible by 16, then  $k$ -WL is unable to distinguish between vertex-transitive and non-vertex-transitive graphs with  $n$  vertices unless  $k = \Omega(\sqrt{n})$ ; see Theorem 7. This excludes extension of our positive result to graphs with an arbitrary number of vertices. Indeed, since the combination of 2-WL with vertex individualization is subsumed by 3-WL, such an extension would readily imply that 3-WL distinguishes any vertex-transitive graph from any non-vertex-transitive graph, contradicting our lower bound  $k = \Omega(\sqrt{n})$ . This bound as well excludes any other combinatorial approach to recognizing vertex-transitivity as long as it is based solely on  $k$ -WL for a fixed dimension  $k$ . It shows that, if such an algorithm succeeds on the  $n$ -vertex input graphs for  $n$  in a set  $S$ , then  $S$  can contain only finitely many multiples of 16.

Our lower bound is based on the Cai-Fürer-Immerman construction [5], which converts a template graph  $F$  into a pair of non-isomorphic graphs  $G$  and  $H$  indistinguishable by  $k$ -WL. To prove our lower bound for the WL dimension, we have to ensure that  $G$  is vertex-transitive and  $H$  is not. This is faced with two technical complications. First, the original CFI gadget [5, Fig. 3] involves vertices of different degrees and, hence, destroys vertex-transitivity even when the template graph  $F$  is vertex-transitive. This can be overcome by using a modified version of the CFI gadget with all vertex degrees equal, which apparently first appeared in [9]; see also the survey in [12]. Note that this approach has already been used to analyze vertex-transitivity of coherent configurations; see Evdokimov's thesis [8].

The second point is more subtle. The CFI construction replaces each vertex of the template graph  $F$  with a cell of new vertices, and vertices in different cells receive different colors. In many contexts the vertex coloring can be removed by

using additional gadgets, but this is hardly possible without losing the vertex-transitivity. The vertex colors constrain the automorphisms of the CFI graphs  $G$  and  $H$  and ensure that these graphs are non-isomorphic. We establish rather general conditions on a template graph  $F$  under which the CFI graphs retain their functionality even without colors. This result of independent interest provides a very straight way of making the CFI graphs colorless, which can be used in any of their numerous applications.

The analysis of the regularized and discolored version of the CFI construction and the proof of our lower bound (Theorem 7) can be found in a long version of this paper [11].

## 2 Notation and Definitions

We denote the vertex set of a graph  $G$  by  $V(G)$ . The notation  $\text{Aut}(G)$  stands for the automorphism group of  $G$ .

**Cayley Graphs.** Let  $\Gamma$  be a group and  $Z$  be a set of non-identity elements of  $\Gamma$  such that  $Z^{-1} = Z$ , that is, any element belongs to  $Z$  only together with its inverse. The *Cayley graph*  $\text{Cay}(\Gamma, Z)$  has the elements of  $\Gamma$  as vertices, where  $x$  and  $y$  are adjacent if  $x^{-1}y \in Z$ . This graph is connected if and only if the *connection set*  $Z$  is a generating set of  $\Gamma$ . Every Cayley graph is obviously vertex-transitive.

**The Weisfeiler-Leman Algorithm.** The original version of the Weisfeiler-Leman algorithm, 2-WL, operates on the Cartesian square  $V^2$  of the vertex set of an input graph  $G$ . Below it is supposed that  $G$  is undirected. We also suppose that  $G$  is endowed with a vertex coloring  $c$ , that is, each vertex  $u \in V$  is assigned a color denoted by  $c(u)$ . The case of uncolored graphs is covered by assuming that  $c(u)$  is the same for all  $u$ . 2-WL starts by assigning each pair  $(u, v) \in V^2$  the initial color  $\text{WL}_2^0(u, v) = (\text{type}, c(u), c(v))$ , where *type* takes on one of three values, namely *edge* if  $u$  and  $v$  are adjacent, *nonedge* if distinct  $u$  and  $v$  are non-adjacent, and *loop* if  $u = v$ . The coloring of  $V^2$  is then modified step by step. The  $(r + 1)$ -th coloring is computed as

$$\text{WL}_2^{r+1}(u, v) = \{ \{ (\text{WL}_2^r(u, w), \text{WL}_2^r(w, v)) \}_{w \in V}, \tag{1}$$

where  $\{ \{ \}$  denotes the multiset. In words, the new color of a pair  $uv$  is a “superposition” of all old color pairs observable along the extensions of  $uv$  to a triangle  $uvw$ . Let  $\mathcal{S}^r$  denote the partition of  $V^2$  determined by the coloring  $\text{WL}_2^r(\cdot, \cdot)$ . It is easy to notice that  $\text{WL}_2^{r+1}(u, v) = \text{WL}_2^{r+1}(u', v')$  implies  $\text{WL}_2^r(u, v) = \text{WL}_2^r(u', v')$ , which means that  $\mathcal{S}^{r+1}$  is finer than or equal to  $\mathcal{S}^r$ . It follows that the partition stabilizes starting from some step  $t \leq n^2$ , where  $n = |V|$ , that is,  $\mathcal{S}^{\ell+1} = \mathcal{S}^\ell$ , which implies that  $\mathcal{S}^r = \mathcal{S}^t$  for all  $r \geq t$ . As the stabilization is reached, 2-WL terminates and outputs the coloring  $\text{WL}_2^t(\cdot, \cdot)$ , which will be denoted by  $\text{WL}_2(\cdot, \cdot)$ .

Note that the length of  $\text{WL}_2^r(u, v)$  grows exponentially as  $r$  increases. The exponential blow-up is remedied by renaming the colors after each step.

Let  $\phi$  be an automorphism of  $G$ . A simple induction on  $r$  shows that  $WL_2^r(\phi(u), \phi(v)) = WL_2^r(u, v)$  for all  $r$  and, hence

$$WL_2(\phi(u), \phi(v)) = WL_2(u, v). \tag{2}$$

In particular, if  $G$  is vertex-transitive, then the color  $WL_2(u, u)$  is the same for all  $u \in V$ . If the last condition is fulfilled, we say that 2-WL *does not split the diagonal on  $G$* , where by the diagonal we mean the set of all loops  $(u, u)$ .

In general, the automorphism group  $\text{Aut}(G)$  of the graph  $G$  acts on the Cartesian square  $V(G)^2$ , and the orbits of this action are called *2-orbits* of  $\text{Aut}(G)$ . Thus, the partition of  $V(G)^2$  into 2-orbits is finer than or equal to the stable partition  $\mathcal{S} = \mathcal{S}^t$  produced by 2-WL.

### 3 Vertex-Transitivity on a Prime Number of Vertices

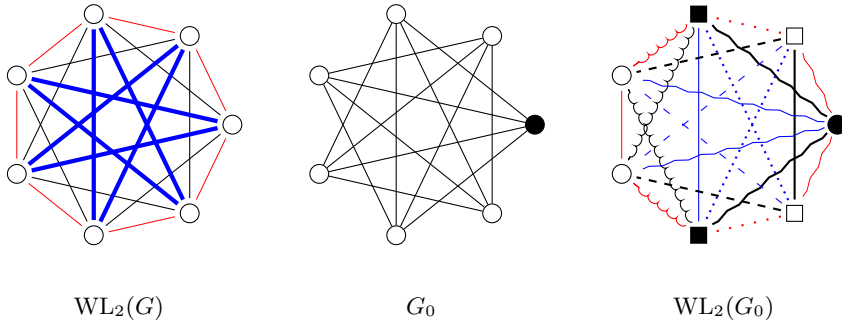
We begin with a few simple observations about the output produced by 2-WL on an input graph  $G$ . Recall that in this paper we restrict our attention to undirected graphs. Even though  $G$  is undirected, the equality  $WL_2(u, v) = WL_2(v, u)$  need not be true in general. Thus, the output of 2-WL on  $G$  can naturally be seen as a complete colored directed graph on the vertex set  $V(G)$ , which we denote by  $WL_2(G)$ . That is,  $WL_2(G)$  contains every pair  $(u, v) \in V(G)^2$  as an arc, i.e., a directed edge, and this arc has the color  $WL_2(u, v)$  returned by 2-WL. We will see  $WL_2(G)$  as containing no loops, but instead we assign each vertex  $u$  the color  $WL_2(u, u)$ . Any directed subgraph of  $WL_2(G)$  formed by all arcs of the same color is called a *constituent digraph*.

Let  $(u, v)$  and  $(u', v')$  be arcs of a constituent digraph  $C$  of  $WL_2(G)$ . Note that the vertices  $u$  and  $u'$  must be equally colored in  $WL_2(G)$ . Indeed, since the color partition of  $WL_2(G)$  is stable, there must exist  $w$  such that  $(WL_2(u', w), WL_2(w, v')) = (WL_2(u, u), WL_2(u, v))$ . The equality  $WL_2(u', w) = WL_2(u, u)$  can be fulfilled only by  $w = u'$  because any non-loop  $(u', w)$  is initially colored differently from the loop  $(u, u)$  and, hence, they are colored differently after all refinements.

Note also that, if  $u$  and  $v$  are equally colored in  $WL_2(G)$ , then they have the same outdegree in every constituent digraph  $C$ ; in particular, they simultaneously belong or do not belong to  $V(C)$ . Otherwise, contrary to the assumption that the color partition of  $WL_2(G)$  is stable, the loops  $(u, u)$  and  $(v, v)$  would receive different colors in another refinement round of 2-WL. It follows that for each constituent digraph  $C$  there is an integer  $d \geq 1$  such that all vertices in  $C$  with non-zero outdegree have outdegree  $d$ . We call  $d$  the *outdegree of  $C$* .

Let  $u \in V(G)$ . A *vertex-individualized graph*  $G_u$  is obtained from  $G$  by assigning the vertex  $u$  a special color, which does not occur in  $G$ . If  $G$  is vertex-transitive, then all vertex-individualized copies of  $G$  are obviously isomorphic.

Consider now a simple and still instructive example. Let  $G = \overline{C}_7$  be the complement of the cycle graph on seven vertices  $0, 1, \dots, 6$  passed in this order. It is not hard to see that 2-WL splits  $V(G)^2$  into the four 2-orbits of  $\text{Aut}(G)$ ; the diagonal  $\{(u, u) : u \in V(G)\}$  is one of them. Note that the three constituent digraphs



**Fig. 1.** The output of 2-WL on input  $G = \overline{C_7}$  and on its vertex-individualized copy  $G_0$ . In this example, the arcs between two equally colored vertices  $u$  and  $v$  (those with  $WL_2(u, u) = WL_2(v, v)$ ) have equal colors in both directions, that is,  $WL_2(u, v) = WL_2(v, u)$ . If  $u$  and  $v$  are colored distinctly, then clearly  $WL_2(u, v) \neq WL_2(v, u)$ . As a general fact,  $WL_2(u, v) = WL_2(u', v')$  exactly when  $WL_2(v, u) = WL_2(v', u')$ . This allows us to improve the visualization by showing only one of the two mutually reversed colors. (Color figure online)

of  $WL_2(G)$  are of the same degree 2; see Fig. 1. Applying 2-WL to the vertex-individualized graph  $G_0$ , it can easily be seen that 2-WL again splits  $V(G)^2$  into the 2-orbits of  $\text{Aut}(G_0)$ . Note that  $WL_2(G_0)$  also has exactly three constituent digraphs of outdegree 2, while all other constituent digraphs of  $WL_2(G_0)$  have outdegree 1. We see that the outdegrees of the constituent digraphs for  $G$  and its vertex-individualized copies are distributed similarly. This similarity proves to be a characterizing property of vertex-transitive graphs on a prime number of vertices.

**Theorem 1.** *Let  $p$  be a prime, and  $G$  be a graph with  $p$  vertices. Suppose that  $G$  is neither complete nor empty. Then  $G$  is vertex-transitive if and only if the following conditions are true:*

1. *If run on  $G$ , 2-WL does not split the diagonal, that is, all vertices in  $WL_2(G)$  are equally colored.*
2. *All constituent digraphs of  $WL_2(G)$  have the same outdegree  $d > 1$  and, hence, there are  $\frac{p-1}{d}$  constituent digraphs.*
3. *For every  $u \in V(G)$ , exactly  $\frac{p-1}{d}$  constituent digraphs in  $WL_2(G_u)$  have outdegree  $d$ , and all others have outdegree 1.*

Since the color partition of  $WL_2(G)$  for a  $p$ -vertex graph  $G$  can be computed in time  $O(p^3 \log p)$  [14], Conditions 1–3 can be verified in time  $O(p^4 \log p)$ , which yields an algorithm of this time complexity for recognition of vertex-transitivity of graphs with a prime number of vertices.

As it will be discussed in Remark 6 below, there are graphs  $G$  and  $H$  with a prime number of vertices such that  $G$  is vertex-transitive,  $H$  is not, and still they are indistinguishable by 2-WL. This implies that Theorem 1 is optimal in that

it uses as small WL dimension as possible and, also, that the condition involving the vertex individualization cannot be dropped. Note that 1-WL, which stands for the classical degree refinement, does not suffice even when run on  $G$  and all  $G_u$  because the output of 1-WL on these inputs is subsumed by the output of 2-WL on  $G$  alone.

Theorem 1 is proved in Subsect. 3.2. The next subsection provides the necessary preliminaries.

### 3.1 Coherent Configurations

A detailed treatment of the material presented below can be found in [6]. The stable partition  $\mathcal{S}_G$  of  $V(G)^2$  produced by 2-WL on an input graph  $G$  has certain regularity properties, which are equivalent to saying that the pair  $(V, \mathcal{S})$  forms a coherent configuration. This concept is defined as follows.

A *coherent configuration*  $\mathcal{X} = (V, \mathcal{S})$  is formed by a set  $V$ , whose elements are called *points*, and a partition  $\mathcal{S} = \{S_1, \dots, S_m\}$  of the Cartesian square  $V^2$ , that is,  $\bigcup_{i=1}^m S_i = V^2$  and any two  $S_i$  and  $S_j$  are disjoint. An element  $S_i$  of  $\mathcal{S}$  is referred to as a *basis relation* of  $\mathcal{X}$ . The partition  $\mathcal{S}$  has to satisfy the following three conditions:

- (A) If a basis relation  $S \in \mathcal{S}$  contains a loop  $(u, u)$ , then all pairs in  $S$  are loops.
- (B) For every  $S \in \mathcal{S}$ , the *transpose relation*  $S^* = \{(v, u) : (u, v) \in S\}$  is also in  $\mathcal{S}$ .
- (C) For each triple  $R, S, T \in \mathcal{S}$ , the number

$$p(u, v) = |\{w : (u, w) \in R, (w, v) \in S\}|$$

for a pair  $(u, v) \in T$  does not depend on the choice of this pair in  $T$ .

In other words, if  $\mathcal{S}$  is seen as a color partition of  $V^2$ , then such a coloring is stable under 2-WL refinement.

We describe two important sources of coherent configurations. Let  $\mathcal{T}$  be an arbitrary family of subsets of the Cartesian square  $V^2$ . There exists a unique coarsest partition  $\mathcal{S}$  of  $V^2$  such that every  $T \in \mathcal{T}$  is a union of elements of  $\mathcal{S}$  and  $\mathcal{X} = (V, \mathcal{S})$  is a coherent configuration; see [6, Section 2.6.1]. We call  $\mathcal{X} = (V, \mathcal{S})$  the *coherent closure* of  $\mathcal{T}$  and denote it by  $Cl(\mathcal{T})$ .

Given a vertex-colored undirected graph  $G$  on the vertex set  $V$ , let  $\mathcal{T}$  consist of the set of the pairs  $(u, v) \in V^2$  such that  $\{u, v\}$  is an edge of  $G$  and the sets of loops  $(u, u)$  for all vertices  $u$  of the same color in  $G$ . Then  $Cl(\mathcal{T})$  is exactly the stable partition produced by 2-WL on input  $G$ . We denote this coherent configuration by  $Cl(G)$ .

Given a coherent configuration  $\mathcal{X} = (V, \mathcal{S})$  and a point  $u \in V$ , the coherent configuration  $\mathcal{X}_u = Cl(\mathcal{S} \cup \{(u, u)\})$  is called a *one-point extension* of  $\mathcal{X}$ . This concept is naturally related to the notion of a vertex-individualized graph, in that

$$Cl(G_u) = Cl(G)_u. \tag{3}$$

Another source of coherent configurations is as follows. Let  $K$  be a permutation group on a set  $V$ . Denote the set of 2-orbits of  $K$  by  $\mathcal{S}$ . Then  $\mathcal{X} = (V, \mathcal{S})$  is a coherent configuration, which we denote by  $\text{Inv}(K)$ . Coherent configurations obtained in this way are said to be *schurian*.

We define an *automorphism* of a coherent configuration  $\mathcal{X} = (V, \mathcal{S})$  as a bijection  $\alpha$  from  $V$  onto itself such that, for every  $S \in \mathcal{S}$  and every  $(u, v) \in S$ , the pair  $(\alpha(u), \alpha(v))$  also belongs to  $S$ . The group of all automorphisms of  $\mathcal{X}$  is denoted by  $\text{Aut}(\mathcal{X})$ . A coherent configuration  $\mathcal{X}$  is *schurian* if and only if

$$\mathcal{X} = \text{Inv}(\text{Aut}(\mathcal{X})). \tag{4}$$

Note also that the connection between the coherent closure of a graph and 2-WL implies that

$$\text{Aut}(Cl(G)) = \text{Aut}(G). \tag{5}$$

A set of points  $X \subseteq V$  is called a *fiber* of  $\mathcal{X}$  if the set of loops  $\{(x, x) : x \in X\}$  is a basis relation of  $\mathcal{X}$ . Denote the set of all fibers of  $\mathcal{X}$  by  $F(\mathcal{X})$ . By Property **A**,  $F(\mathcal{X})$  is a partition of  $V$ . Property **C** implies that for every basis relation  $S$  of  $\mathcal{X}$  there are, not necessarily distinct, fibers  $X$  and  $Y$  such that  $S \subseteq X \times Y$ . We use the notation  $N_S(x) = \{y : (x, y) \in S\}$  for the set of all points in  $Y$  that are in relation  $S$  with  $x$ . Note that  $|N_S(x)| = |N_S(x')|$  for any  $x, x' \in X$ . We call this number the *valency* of  $S$ . If every basis relation  $S$  of  $\mathcal{X}$  has valency 1, then  $\mathcal{X}$  is called *semiregular*.

**Proposition 2** (see [6, Exercise 2.7.35]). *A semiregular coherent configuration is schurian.*

Given a set of points  $U \subseteq V$  that is a union of fibers, let  $\mathcal{S}_U$  denote the set of all basis relations  $S \in \mathcal{S}$  such that  $S \subseteq X \times Y$  for some, not necessarily distinct, fibers  $X \subseteq U$  and  $Y \subseteq U$ . As easily seen,  $\mathcal{X}_U = (U, \mathcal{S}_U)$  is a coherent configuration.

If a coherent configuration has a single fiber, it is called *association scheme*.

### 3.2 Proof of Theorem 1

**Necessity.** Given a vertex-transitive graph  $G$  with  $p$  vertices, where  $p$  is prime, we have to check Conditions 1–3. Condition 1 follows immediately from vertex-transitivity; see the discussion in the end of Sect. 2. For Condition 2, we use two basic results on vertex-transitive graphs with a prime number of vertices. First, every such graph is isomorphic to a *circulant graph*, i.e., a Cayley graph of a cyclic group, because every transitive group of permutations of a set of prime cardinality  $p$  contains a  $p$ -cycle (Turner [19]). Let  $\mathbb{F}_p$  denote the  $p$ -element field,  $\mathbb{F}_p^+$  its additive group, i.e., the cyclic group of order  $p$ , and  $\mathbb{F}_p^\times$  its multiplicative group, which is isomorphic to the cyclic group of order  $p - 1$ . Another useful fact (Alspach [1]) is that, if a set  $Z \subset \mathbb{F}_p$  is non-empty and  $Z \neq \mathbb{F}_p^\times$ , then the automorphism group of the circulant graph  $\text{Cay}(\mathbb{F}_p^+, Z)$  consists of the permutations

$$x \mapsto ax + b, \quad x \in \mathbb{F}_p, \quad \text{for all } a \in M, b \in \mathbb{F}_p^+, \tag{6}$$



where  $M = M(Z)$  is the largest subgroup of  $\mathbb{F}_p^\times$  of even order such that  $Z$  is a union of cosets of  $M$ . This subgroup is well defined because the condition  $Z = -Z$  implies that  $Z$  is split into pairs  $\{z, -z\}$  and, hence, is a union of cosets of the multiplicative subgroup  $\{1, -1\}$ . For example,  $\overline{C_7} = \text{Cay}(\mathbb{F}_7^+, \{2, 3, 4, 5\})$  and  $M(\{2, 3, 4, 5\}) = \{1, -1\}$ . Without loss of generality we assume that  $G = \text{Cay}(\mathbb{F}_p^+, Z)$  and denote  $K = \text{Aut}(G)$ .

Let  $\mathcal{X} = (\mathbb{F}_p^+, \mathcal{S})$  be the coherent closure of  $G$ . Recall that  $\mathcal{S}$  is exactly the stable partition of  $V(G)^2$  produced by 2-WL on input  $G$ . The irreflexive basis relations of  $\mathcal{X}$  are exactly the constituent digraphs of  $\text{WL}_2(G)$ , and we have to prove that all of them have the same valency.

Condition 1 says that  $\mathcal{X}$  is an association scheme. In general, not all association schemes with a prime number of points are schurian (see, e.g., [6, Section 4.5]). Nevertheless, the theorem by Leung and Man on the structure of Schur rings over cyclic groups implies the following fact.

**Proposition 3** (see [6, Theorem 4.5.1]). *Let  $\mathcal{X} = (V, \mathcal{S})$  be an association scheme with a prime number of points. If  $\text{Aut}(\mathcal{X})$  acts transitively on  $V$ , then  $\mathcal{X}$  is schurian.*

By Equality (5),  $\text{Aut}(\mathcal{X}) = K$ . Since the group  $K$  is transitive, Proposition 3 implies that  $\mathcal{X}$  is schurian, and we have  $\mathcal{X} = \text{Inv}(K)$  by Equality (4). This yields Condition 2 for  $d = |M|$ . Indeed, every irreflexive basis relation  $S \in \mathcal{S}$  has valency  $|M|$ . To see this, it is enough to count the number of pairs  $(0, y)$  in  $S$ . Fix an arbitrary pair  $(0, y) \in S$ . A pair  $(0, y')$  is in the 2-orbit containing  $(0, y)$  if and only if  $y' = ay$  for  $a \in M$ , for which we have  $|M|$  possibilities.

It remains to prove Condition 3. By vertex-transitivity, all vertex-individualized copies of  $G$  are isomorphic and, therefore, it is enough to consider  $G_0$ . We have to count the frequencies of valencies in  $\mathcal{X}_0$ . Note that  $\mathcal{X}_0 = \text{Cl}(G_0)$  by Equality (3).

It is generally not true that a one-point extension of a schurian coherent configuration is schurian; see [6, Section 3.3.1]. Luckily, this is the case in our setting.

**Proposition 4** (see [6, Theorem 4.4.14]). *If  $\mathcal{X} = \text{Inv}(K)$ , where  $K$  is the group of permutations of the form (6) for a subgroup  $M$  of  $\mathbb{F}_p^\times$ , then the one-point extension  $\mathcal{X}_0$  is schurian.*

Taking into account Equality (5), we have

$$\text{Aut}(\mathcal{X}_0) = \text{Aut}(\text{Cl}(G_0)) = \text{Aut}(G_0) = \text{Aut}(G)_0 = K_0,$$

where  $K_0$  is the one-point stabilizer of 0 in  $K$ , that is, the subgroup of  $K$  consisting of all permutations  $\alpha \in K$  such that  $\alpha(0) = 0$ . Obviously,  $K_0 = \{x \mapsto ax, x \in \mathbb{F}_p\}_{a \in M}$ .

Let  $S$  be a 2-orbit of  $K_0$ . If  $S$  contains a pair  $(0, y)$ , then it consists of all pairs  $(0, y')$  for  $y' \in My$  and, hence, has valency  $|M|$ . If  $S$  contains a pair  $(z, y)$  with  $z \neq 0$  and  $y \neq z$ , then  $(z, y)$  is the only element of  $S$  with the first coordinate  $z$ , and  $S$  has valency 1. The proof of Condition 3 is complete.

**Sufficiency.** Let  $G$  be a graph satisfying Conditions 1–3 stated in the theorem. Let  $\mathcal{X} = Cl(G)$ . Condition 1 says that  $\mathcal{X}$  is an association scheme. By Equality (5), it suffices to prove that the group  $\text{Aut}(\mathcal{X})$  is transitive. The proof is based on the following lemma.

**Lemma 5** (see [16, Theorem 7.1]). *Let  $\mathcal{X} = (V, \mathcal{S})$  be an association scheme. Suppose that the following two conditions are true for every point  $u \in V$ :*

- (I) *the coherent configuration  $(\mathcal{X}_u)_{V \setminus \{u\}}$  is semiregular, and*
- (II)  *$F(\mathcal{X}_u) = \{N_S(u) : S \in \mathcal{S}\}$ .*

*Then the group  $\text{Aut}(\mathcal{X})$  acts transitively on  $V$ .*

Let  $u$  be an arbitrary vertex of  $G$ . By Equality (3),  $\mathcal{X}_u = Cl(G_u)$ . Now, it suffices to derive Conditions I–II in the lemma from Conditions 1–3 in the theorem.

For a fiber  $X \in F(\mathcal{X}_u)$ , note that  $\{u\} \times X$  must be a basis relation of  $\mathcal{X}_u$ . Since this relation has valency  $|X|$ , Condition 3 implies that every fiber in  $F(\mathcal{X}_u)$  is either a singleton or consists of  $d \geq 2$  points. Denote the number of singletons in  $F(\mathcal{X}_u)$  by  $a$ . Besides of them,  $F(\mathcal{X}_u)$  contains  $(p - a)/d$  fibers of size  $d$ .

For every  $X, Y \in F(\mathcal{X}_u)$  with  $|X| = 1$  and  $|Y| = d$ ,  $X \times Y$  is a basis relation of  $\mathcal{X}_u$  of valency  $d$ . It follows from Condition 3 that

$$\frac{p - 1}{d} \geq \frac{a(p - a)}{d}.$$

Therefore,  $p - 1 \geq a(p - a)$  or, equivalently,  $p(a - 1) \leq (a - 1)(a + 1)$ . Assume for a while that  $a > 1$ . It immediately follows that  $a \geq p - 1$ . Since the equality  $a = p - 1$  is impossible, we conclude that  $a = p$ . However, this implies that  $d = 1$ , a contradiction. Thus,  $a = 1$ . Consequently, every fiber of the coherent configuration  $\mathcal{X}' = (\mathcal{X}_u)_{V \setminus \{u\}}$  is of cardinality  $d$ , and  $|F(\mathcal{X}')| = (p - 1)/d$ .

Let  $S$  be a basis relation of  $\mathcal{X}$ . If  $S$  is reflexive, then  $N_S(u) = \{u\}$ . If  $S$  is irreflexive, then  $N_S(u)$  must be a union of fibers in  $F(\mathcal{X}')$ . By Condition 2, the number of irreflexive basis relations in  $\mathcal{S}$  is  $(p - 1)/d$ . It follows that  $N_S(u)$  actually coincides with one of the fibers of  $\mathcal{X}'$ . This proves Condition II.

Since  $\mathcal{X}_u$  contains  $(p - 1)/d$  basis relations of the kind  $\{u\} \times X$  for  $X \in F(\mathcal{X}')$ , Condition 3 implies that every basis relation of  $\mathcal{X}'$  is of valency 1, yielding Condition I.

The proof of Theorem 1 is complete.

*Remark 6.* We now argue that there is a vertex transitive graph  $G$  and a non-vertex-transitive graph  $H$  such that  $G$  and  $H$  are indistinguishable by 2-WL. Recall that a *strongly regular graph* with parameters  $(n, d, \lambda, \mu)$  is an  $n$ -vertex  $d$ -regular graph where every two adjacent vertices have  $\lambda$  common neighbors, and every two non-adjacent vertices have  $\mu$  common neighbors. As easily seen, two strongly regular graphs with the same parameters are indistinguishable by 2-WL, and our example will be given by  $G$  and  $H$  of this kind. Let  $p$  be a prime (or a prime power) such that  $p \equiv 1 \pmod{4}$ . The *Paley graph* on  $p$  vertices is

the Cayley graph  $\text{Cay}(\mathbb{F}_p^+, Y_p)$  where  $Y_p$  is the subgroup of  $\mathbb{F}_p^\times$  formed by all quadratic residues modulo  $p$ . The assumption  $p \equiv 1 \pmod{4}$  ensures that  $-1$  is a quadratic residue modulo  $p$  and, hence,  $Y_p = -Y_p$ . The Paley graph on  $p$  vertices is strongly regular with parameters  $(p, \frac{p-1}{2}, \frac{p-5}{4}, \frac{p-1}{4})$ .

Let  $G$  be the Paley graph on 29 vertices. It is known (Bussemaker and Spence; see, e.g., [4, Section 9.9]) that there are 40 other strongly regular graphs with parameters  $(29, 14, 6, 7)$ . Let  $H$  be one of them. We have only to show that  $H$  is not vertex-transitive. Otherwise, by Turner’s theorem [19] this would be a circulant graph, that is, we would have  $H = \text{Cay}(\mathbb{F}_p^+, Z)$  for some connection set  $Z$ . In this case, the coherent closure  $Cl(H)$  must be schurian by Proposition 3. Since  $H$  is strongly regular, 2-WL colors all pairs of adjacent vertices uniformly and, therefore, they form a 2-orbit of  $\text{Aut}(H)$ . It follows that the stabilizer  $\text{Aut}(H)_0$  acts transitively on  $N(0)$ , the neighborhood of 0 in  $H$ . The aforementioned result of Alspach [1], implies that  $Z$  is the subgroup of  $\mathbb{F}_p^\times$  of order  $(p - 1)/2$ , i.e.,  $M = Z$  in (6). This means that  $Z = Y_p$  and  $H = G$ , a contradiction.

### 4 A Lower Bound for the WL Dimension

We now state a negative result on the recognizability of vertex-transitivity by  $k$ -WL. We begin with a formal definition of the  $k$ -dimensional algorithm. Let  $k \geq 2$ . Given a graph  $G$  with vertex set  $V$  as input,  $k$ -WL operates on  $V^k$ . The initial coloring of  $\bar{u} = (u_1, \dots, u_k)$  encodes the equality type of this  $k$ -tuple and the ordered isomorphism type of the subgraph of  $G$  induced by the vertices  $u_1, \dots, u_k$ . The color refinement is performed similarly to (1). Specifically,  $k$ -WL iteratively colors  $V^k$  by  $\text{WL}_k^{r+1}(\bar{u}) = \{ \{ \text{WL}_k^r(\bar{u}_1^w), \dots, \text{WL}_k^r(\bar{u}_k^w) \} \mid w \in V(G), \text{ where } \bar{u}_i^w = (u_1, \dots, u_{i-1}, w, u_{i+1}, \dots, u_k) \}$ . If  $G$  has  $n$  vertices, the color partition stabilizes in  $t \leq n^k$  rounds, and  $k$ -WL outputs the coloring  $\text{WL}_k(\cdot) = \text{WL}_k^t(\cdot)$ .

We say that  $k$ -WL *distinguishes* graphs  $G$  and  $H$  if the final color palettes are different for  $G$  and  $H$ , that is,  $\{ \{ \text{WL}_k(\bar{u}) \} \mid \bar{u} \in V(G)^k \} \neq \{ \{ \text{WL}_k(\bar{u}) \} \mid \bar{u} \in V(H)^k \}$  (note that color renaming in each refinement round must be performed on  $G$  and  $H$  synchronously).

**Theorem 7.**

1. For every  $n$  divisible by 16 there are  $n$ -vertex graphs  $G$  and  $H$  such that  $G$  is vertex-transitive,  $H$  is not, and  $G$  and  $H$  are indistinguishable by  $k$ -WL as long as  $k \leq 0.01 \sqrt{n}$ .
2. For infinitely many  $n$  there are  $n$ -vertex graphs  $G$  and  $H$  such that  $G$  is vertex-transitive,  $H$  is not, and  $G$  and  $H$  are indistinguishable by  $k$ -WL as long as  $k \leq 0.001 n$ .

### 5 Concluding Discussion

We have suggested a new, very simple combinatorial algorithm recognizing, in polynomial time, vertex-transitivity of graphs with a prime number of vertices.

The algorithm consists, in substance, in running 2-WL on an input graph and all its vertex-individualized copies. This is perhaps the first non-trivial example of using the Weisfeiler-Leman algorithm for recognition of a natural graph property rather than for isomorphism testing.

One can consider another, conceptually even simpler approach. If an input graph  $G$  is vertex-transitive, then  $k$ -WL colors all diagonal  $k$ -tuples  $(u, \dots, u)$ ,  $u \in V(G)$ , in the same color. Is this condition for a possibly large, but fixed  $k$  sufficient to claim vertex-transitivity? In general, a negative answer immediately follows from Theorem 7. Does there exist a fixed dimension  $k$  such that the answer is affirmative for graphs with a prime number of vertices? This is apparently a hard question; it seems that we cannot even exclude that  $k = 3$  suffices.

Another interesting question is whether  $k$ -WL is able to efficiently recognize vertex-transitivity on  $n$ -vertex input graphs for  $n$  in a larger range than the set of primes. The lower bound of Theorem 7 excludes this only for  $n$  divisible by 16, in particular, for the range of  $n$  of the form  $16p$  for a prime  $p$ . Can  $k$ -WL be successful on the inputs with  $2p$  vertices? Conversely, can the negative result of Theorem 7 be extended to a larger range of  $n$ ?<sup>1</sup>

Finally, we remark that the results similar to Theorems 1 and 7 can be obtained for recognition of arc-transitivity. We refer an interested reader to a long version of this paper [11].

## References

1. Alspach, B.: Point-symmetric graphs and digraphs of prime order and transitive permutation groups of prime degree. *J. Comb. Theory Ser. B* **15**, 12–17 (1973)
2. Babai, L.: Isomorphism problem for a class of point-symmetric structures. *Acta Math. Acad. Sci. Hungar.* **29**(3–4), 329–336 (1977). <https://doi.org/10.1007/BF01895854>
3. Babai, L.: Graph isomorphism in quasipolynomial time. In: *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC 2016)*, pp. 684–697 (2016). <https://doi.org/10.1145/2897518.2897542>
4. Brouwer, A.E., Haemers, W.H.: *Spectra of Graphs*. Springer, Berlin (2012). <https://doi.org/10.1007/978-1-4614-1939-6>
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). <https://doi.org/10.1007/BF01305232>
6. Chen, G., Ponomarenko, I.: *Coherent Configurations*. Central China Normal University Press, Wuhan (2019), a draft version is available at <http://www.pdmi.ras.ru/~inp/ccNOTES.pdf>
7. Evdokimov, S., Ponomarenko, I.: Circulant graphs: recognizing and isomorphism testing in polynomial time. *St. Petersburg. Math. J.* **15**(6), 813–835 (2004)
8. Evdokimov, S.: *Schurity and separability of association schemes*. Ph.D. thesis, St. Petersburg University, St. Petersburg (2004)

<sup>1</sup> A simple inspection of the proof shows that Theorem 7 can be extended to the range of  $n$  divisible by  $8p$  for each prime  $p$ .

9. Evdokimov, S., Ponomarenko, I.: On highly closed cellular algebras and highly closed isomorphisms. *Electr. J. Comb.* **6** (1999). [http://www.combinatorics.org/Volume\\_6/Abstracts/v6i1r18.html](http://www.combinatorics.org/Volume_6/Abstracts/v6i1r18.html)
10. Evdokimov, S., Ponomarenko, I., Tinhofer, G.: Forestal algebras and algebraic forests (on a new class of weakly compact graphs). *Discrete Math.* **225**(1–3), 149–172 (2000). [https://doi.org/10.1016/S0012-365X\(00\)00152-7](https://doi.org/10.1016/S0012-365X(00)00152-7)
11. Fuhlbrück, F., Köbler, J., Ponomarenko, I., Verbitsky, O.: The Weisfeiler-Leman algorithm and recognition of graph properties. Tech. rep. [arxiv.org/abs/2005.08887](https://arxiv.org/abs/2005.08887) (2020)
12. Fürer, M.: On the combinatorial power of the Weisfeiler-Lehman algorithm. In: Fotakis, D., Pagourtzis, A., Paschos, V.T. (eds.) *CIAC 2017. LNCS*, vol. 10236, pp. 260–271. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-57586-5\\_22](https://doi.org/10.1007/978-3-319-57586-5_22)
13. Grohe, M.: Fixed-point definability and polynomial time on graphs with excluded minors. *J. ACM* **59**(5), 27:1–27:64 (2012). <https://doi.org/10.1145/2371656.2371662>
14. Immerman, N., Lander, E.: Describing graphs: a first-order approach to graph canonization. In: Selman, A.L. (eds) *Complexity Theory Retrospective*. Springer, New York, NY (1990). [https://doi.org/10.1007/978-1-4612-4478-3\\_5](https://doi.org/10.1007/978-1-4612-4478-3_5)
15. Kiefer, S., Ponomarenko, I., Schweitzer, P.: The Weisfeiler-Leman dimension of planar graphs is at most 3. *J. ACM* **66**(6), 44:1–44:31 (2019). <https://doi.org/10.1145/3333003>
16. Muzychuk, M., Ponomarenko, I.: On pseudocyclic association schemes. *ARS Math. Contemp.* **5**(1), 1–25 (2012)
17. Muzychuk, M.E., Klin, M.H., Pöschel, R.: The isomorphism problem for circulant graphs via Schur ring theory. In: *Codes and Association Schemes. DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 56, pp. 241–264. DIMACS/AMS (1999). <https://doi.org/10.1090/dimacs/056/19>
18. Muzychuk, M.E., Tinhofer, G.: Recognizing circulant graphs of prime order in polynomial time. *Electr. J. Comb.* **5** (1998). [http://www.combinatorics.org/Volume\\_5/Abstracts/v5i1r25.html](http://www.combinatorics.org/Volume_5/Abstracts/v5i1r25.html)
19. Turner, J.: Point-symmetric graphs with a prime number of points. *J. Comb. Theory* **3**, 136–145 (1967)
20. 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). [https://www.iti.zcu.cz/wl2018/pdf/wl\\_paper\\_translation.pdf](https://www.iti.zcu.cz/wl2018/pdf/wl_paper_translation.pdf)