Property Testing of Equivalence under a Permutation Group Action

László Babai and Sourav Chakraborty
University of Chicago
{laci, sourav}@cs.uchicago.edu

Abstract

For a permutation group G acting on the set Ω we say that two strings $x, y : \Omega \to \{0,1\}$ are G-*isomorphic* if they are equivalent under the action of G, i. e., if for some $\pi \in G$ we have $x(i^\pi) = y(i)$ for all $i \in \Omega$. Cyclic Shift, Graph Isomorphism and Hypergraph Isomorphism are special cases, and subcases corresponding to certain classes of groups have been central to the design of efficient isomorphism testing for subclasses of graphs (Luks 1982).

We study the complexity of G-isomorphism in the context of property testing: we want to find the randomized decision tree complexity of distinguishing the cases when x and y are G-isomorphic from the cases when they are at least δ-far from being G-isomorphic (in normalized Hamming distance). Error can be 1-sided or 2-sided. In each case we consider two models. In the query-1 model we assume y is known and only x needs to be queried. In the query-2 model we have to query both x and y.

We give various upper and lower bounds for the four combinations of models considered in terms of $n = |\Omega|$ and $|G|$. In most cases, substantial gaps remain between the upper and lower bounds. However, for *primitive permutation groups*, we obtain a tight (up to polylog(n) factors) bound of $\widetilde{\Theta}(\log |G|)$ for the 1-sided error query complexity in the query-1 model and a tight $\widetilde{\Theta}(\sqrt{n \log |G|})$ bound for the 1-sided error query complexity in the query-2 model. These results extend results of Fischer and Matsliah (2006) on Graph Isomorphism to a surprisingly general class of groups which also includes isomorphism of uniform hypergraphs of any rank. Our results imply a remarkable divergence between the query-1 and the query-2 complexities of k-hypergraph isomorphism for large k in the 1-sided error model.

Besides the fact that they include Graph Isomorphism and k-hypergraph isomorphism, primitive permutation groups are significant because they form the "building blocks" of all permutations groups, providing the base cases of a natural divide-and-conquer approach successfully exploited in algorithm design (Luks, 1982; Babai, Luks, Seress 1987).

While all our bounds are in terms of the order and the degree of G, it seems likely that tighter bounds will depend on the finer structure of G. Our results on primitive groups are the first steps in this direction.

1 Introduction

"Property testing" is a branch of decision tree complexity (query complexity) theory: with a small number of randomized queries to the unknown input string, we want to have a good chance of distinguishing the cases when the input has a given property from the cases when the input is "far" from any string having the property.

This concept was introduced in the context of program checking by Blum, Luby and Rubinfeld [14] who showed that *linearity* of a function over a vector space can be tested with a *constant* number of queries. A central ingredient in the proof of the MIP=NEXP theorem [11] was the proof that *mulitinearity* can be tested with a *polylogarithmic* number of queries. These two papers were among the roots of the technical developments culminating in the PCP Theorem [8, 7].

Rubinfeld and Sudan [24] formally defined property testing in the context of algebraic properties. Subsequently, the interest in property testing was extended to graph properties, with applications to learning and approximation [20]. In recent years the field of combinatorial property testing has enjoyed a rapid growth (see, e. g., [1, 2, 3, 5, 4], cf. [23, 16]).

Notably, Alon and Shapira [4] show that graph properties that are invariant under vertex removal (i. e., are inherited by induced subgraphs) are testable by a constant number of queries. Isomorphism to a given graph is an important example of a graph property that is not hereditary.

The immediate motivation of our work comes from papers by Fischer [17] and Fischer and Matsliah [18] who consider the Graph Isomorphism problem in the property testing model. Here two graphs are given as inputs and we have to test whether they are isomorphic or "far" from being isomorphic.

In this paper we consider a generalization of graph isomorphism. Let us fix a permutation group G acting on the set Ω. Given two input strings $x, y : \Omega \to \{0, 1\}$, we say x is "G-isomorphic" to y if y is a a π-shift of x for some $\pi \in G$. We want to test the property "x is G-isomorphic to y," that is, we want to distinguish the case when x and y are G-isomorphic from the case when every string that is G-isomorphic to y is far from x. [Formal definitions are given in Section 2.]

Graph Isomorphism is a special case of G-isomorphism: let Ω be the set of unordered pairs of the set V of vertices; and $G = \mathrm{Sym}^{(2)}(V)$ the induced action on Ω of $\mathrm{Sym}(V)$, the symmetric group acting on V (so $n = \binom{|V|}{2}$). We note that the induced symmetric group action on pairs is primitive (does not admit nontrivial invariant partitions of the permutation domain). This fact defines the direction in which we extend results on Graph Isomorphism. We note that by considering the induced symmetric group action on k-tuples, another primitive action, we also cover the case of k-uniform hypergraphs (k-hypergraphs). Here k is a variable, $2 \le k \le |V| - 2$. Various finite geometries also correspond to primitive groups, so G-isomorphism includes equivalence under geometric transformations (projective, orthogonal, symplectic, etc.).

Thus Graph Isomorphism and its immediate generalizations (k-hypergraphs) can be viewed as isomorphism under some primitive group action. In addition to this, primitive permutation groups are significant because they form the "building blocks" of all permutations groups in the sense that a "structure tree" can be built of which the leaves constitute the permutation domain and the action of G extends to the tree in such a way that the the stabilizer of any node in the tree acts as a primitive group on the children of the node (cf. [13]). (The stabilizer of a node is the subgroup that sends the node to itself. The structure tree formalizes the natural divide-and-conquer approach successfully exploited in algorithm design [12, 13, 22].)

In "property testing" we want to accept if the inputs are G-isomorphic and reject if they are "far" from being G-isomorphic. The complexity is the number of queries made to the input. We consider two models depending on whether we have to query both x and y or we have to query only one of them (the other is known). We call the models "query-2" and "query-1," respectively. A property test can have 1-sided or 2-sided-error.

In this paper we focus mainly on "property testing" of G-isomorphism when the group is primitive. Our main results are the tight bounds on the query complexity when we are allowed only 1-sided error, that is, the algorithm has to accept with probability 1 when the two inputs are G-isomorphic and reject with high probability when the inputs are "far" from being isomorphic. The main results are the following.

Theorem 1.1 (Tight bounds for primitive groups)**.** *If G is a primitive group then*

(a) the 1-sided-error query-1 complexity of G-isomorphism is $\widetilde{\Theta}(\log |G|)$.

(b) the 1-sided-error query-2 complexity of G-isomorphism is $\widetilde{\Theta}(\sqrt{n \log |G|})$;

The tilde in the asymptotic notation indicates polylog(n) factors.

Theorem 1.1 generalizes a result of Fischer and Matsliah [18] on Graph Isomorphism. The lower bound parts of this result constitute the main technical contributions of this paper and are proved in Section 3.

A permutation group is *transitive* if all elements of the permutation domain are equivalent under the group action. (Primitive groups are transitive by definition.) In the query-2 model, a lower bound of $\sqrt{n}$ can easily be proved for any transitive group G. In this paper we improve this lower bound to $\widetilde{\Theta}(\sqrt{n \log |G|})$ under the condition that the group is primitive. For the lower bound proofs we use a classification of primitive groups based on the O'Nan–Scott Theorem ([25], see [15, 21]) and the Classification of Finite Simple Groups. However, the consequence we use is elementary and can be understood with an undergraduate abstract algebra background.

The upper bounds in Theorem 1.1 hold for any permutation group G.

Proposition 1.2 (Upper bound)**.** *Let G be a permutation group.*

(a) The 1-sided error query-1 complexity of G-isomorphism is $O(1 + \log |G|)$.

(b) The 1-sided error query-2 complexity of G-isomorphism is $O(\sqrt{n(1 + \log |G|)})$.

Consequently the same upper bounds hold for the 2-sided error models.

In Table 1, we abbreviated the expression $1 + \log |G|$ to $\log |G|$ for better typography. The only case where this makes a difference is when $|G| = 1$ so the results as stated in the Table 1 assume $|G| \geq 2$.

We also prove a lower bound for the 2-sided-error query-1 case if the group is transitive and not too large. But a significant gap remains between the upper and lower bounds even in this case.

Theorem 1.3 (Lower bound)**.** *Let G be a transitive group of order $2^{O(n^{1-\epsilon})}$. Then the 2-sided-error query-1 complexity of G-isomorphism is $\Omega(\epsilon \log n)$.*

All our lower bounds are against adaptive queries whereas all our upper bounds uses only non-adaptive queries.

In Section 2 we give the formal definitions. In Sections 3, 4 and 5 we give the proofs of the above three results. In Section 6 we state further nearly tight bounds that follow from our results (in addition to Theorem 1.1).

Table 1 summarizes our results on G-isomorphism. Table 2 gives the results of Fischer and Matsliah on Graph Isomorphism (ignoring logarithmic factors). In Table 3 we specialize our results to the case of Graph Isomorphism for comparison with the results of Fischer and Matsliah. In Table 4 we specialize our results to the case of isomorphism of k-hypergraphs. Table 4 exhibits a remarkable divergence between the query-1 and query-2 complexities for k-hypergraphs in the 1-sided error model: the query-1 complexity is $\widetilde{O}(V)$, independent of k, while the query-2 complexity grows rapidly with k (k goes in the exponent).

2 Preliminaries

2.1 Group actions

Let Ω be a set of size n. The permutations of Ω form the **symmetric group** $\mathrm{Sym}(\Omega)$ of order $n!$. We write the action of $\pi \in \mathrm{Sym}(\Omega)$ as $i \mapsto i^\pi$. For a subset $S \subseteq \Omega$ we set $S^\pi = \{i^\pi : i \in S\}$.

A subgroup G of $\mathrm{Sym}(\Omega)$ is a **permutation group**; Ω is the **permutation domain** on which G acts. G has **order** $|G|$ and **degree** $n = |\Omega|$. The **alternating group** $\mathrm{Alt}(\Omega)$ consists of the even permutations of Ω; it has order $n!/2$.

Given an (abstract) group G, a homomorphism $\varphi : G \to \mathrm{Sym}(\Omega)$ is called an **action** of G on Ω or G-action on Ω. For $\pi \in G$ and $x \in \Omega$ we write x^π instead of $x^{\varphi(\pi)}$ when the action φ is clear from the context. The action is **faithful** if $|\ker(\varphi)| = 1$. The same group can have many significantly different actions. Permutation groups will be viewed as faithful actions on their domain (φ is the identity map).

Let us fix a G-action on Ω. The action is **transitive** if $(\forall i, j \in \Omega)(\exists \pi \in G)(i^\pi = j)$. A partition $\Omega = \Omega_1 \dot\cup \dots \dot\cup \Omega_m$ of Ω into the nonempty disjoint subsets Ω_i is G**-invariant** if if $(\forall \pi \in G)(\forall i \le m)(\exists j \le m)(\Omega_i^\pi = \Omega_j)$. The trivial partitions correspond to $m = 1$ and $m = n$; these are always invariant. The G-action is **primitive** if $n \ge 2$, the action is transitive on Ω, and Ω does not admit any nontrivial G-invariant partition. All primitive permutation groups of degree n other than the symmetric group and the alternating group have order $\le \exp(O(\sqrt{n}\log^2 n))$ ([9, 10]) so except for the two classes of "giants" of orders $n!$ and $n!/2$, resp., $\log(|G|) = \widetilde{O}(\sqrt{n})$ for all primitive groups of degree n.

An important class of primitive groups arises from the **induced action of** $\mathrm{Sym}(V)$ **on the** k**-subsets of** V**;** this group is denoted by $\mathrm{Sym}_V^{(k)}$ and acts on the set $\Omega = \binom{V}{k}$ of k-subsets of V. This group is primitive for $1 \le k < |V|$, $k \ne |V|/2$.

We use the notation $[n] = \{1, 2, 3, ..., n\}$. Most often we take $\Omega = [n]$ and write S_n for $\mathrm{Sym}([n])$ and A_n for $\mathrm{Alt}([n])$.

2.2 Strings, G-isomorphism

Definition 2.1. A *partial assignment* is a function $p : S \to \{0, 1\}$ where $S \subseteq [n]$. We call S the support of this partial assignment and set $|p| = |S|$. We call x a (full) assignment if $x : [n] \to \{0, 1\}$. We say $p \subseteq x$ if x is an extension of p, *i. e.*, if $p = x|_S$ (the restriction of x to S).

$\mathrm{Ham}(x, y)$ will denote the Hamming distance of the strings (full assignments) x and y.

Definition 2.2. Let $T \subseteq [n]$ and let $\pi \in S_n$. Let G be a permutation group acting on $[n]$. Then the sets T^π, where $\pi \in G$, are called the *G-shifts* of T. If $p : T \to \{0, 1\}$ is a partial assignment then we define $p^\pi : T^\pi \to \{0, 1\}$ as $p^\pi(i) = p(i^{\pi^{-1}})$.

Given two full assignments x and y and a permutation group G we denote by $d_G(x, y)$ the minimum distance between the G-shifts of x and y. That is,

$$d_G(x, y) = \min_{\pi_1, \pi_2 \in G} \mathrm{Ham}(x^{\pi_1}, y^{\pi_2}). \tag{1}$$

Since G is a group, we have

$$d_G(x, y) = \min_{\pi \in G} \mathrm{Ham}(x, y^\pi) = \min_{\pi \in G} \mathrm{Ham}(x^\pi, y). \tag{2}$$

If $d_G(x, y) = 0$ then we say "x is G**-isomorphic** to y."

For a given constant δ a **2-sided-error** δ**-property tester** for G-isomorphism is a probabilistic decision tree, say $\mathcal{A}$, such that given $x, y \in \{0, 1\}^n$

if $d_G(x, y) = 0$ then with probability $> 2/3$ we have $\mathcal{A}(x, y) = 1$, and,

if $d_G(x, y) \ge \delta n$ then with probability $> 2/3$ we have $\mathcal{A}(x, y) = 0$.

A **1-sided-error** δ**-property tester** is one which makes no mistake if $d_G(x, y) = 0$.

The complexity of a property tester is the maximum (over all possible inputs) of the minimum number of bits that need to be queried. If neither x nor y is given (so both need to be queried) then we speak of a

	Query-1 Complexity	Query-2 Complexity
1-sided-error	$\widetilde{\Theta}(\log\lvert G\rvert)^{\ddagger}$, $\Omega(\log n)^{\dagger}$	$\widetilde{\Theta}(\sqrt{n\log\lvert G\rvert})^{\ddagger}$
2-sided-error	$O(\log\lvert G\rvert)$, $\Omega(\log n)^{\dagger}$	$O(\sqrt{n\log\lvert G\rvert})$

† The lower bound holds when G is transitive and $|G| = 2^{O(n^{1-\epsilon})}$.
‡ The lower bound is for primitive G and the upper bound has no tilde.

Table 1: Bounds on the query complexity of G-isomorphism.

	Query-1 Complexity	Query-2 Complexity
1-sided-error	$\widetilde{\Theta}(\lvert V\rvert)$	$\widetilde{\Theta}(\lvert V\rvert^{3/2})$
2-sided-error	$\widetilde{\Theta}(\sqrt{\lvert V\rvert})$	$\Omega(\lvert V\rvert)$, $\widetilde{O}(\lvert V\rvert^{5/4})$

Table 2: The results of Fischer and Matsliah for Graph Isomorphism.

query-2 δ-tester and correspondingly of **query-2 complexity**. If one of them is given (we always assume y is given) and only the other (that is x) needs to be queried then we speak of a *query-1 δ-tester* and **query-1 complexity**.

But usually we will drop the δ when talking about the testers and the complexities. Unless otherwise stated, in the rest of the paper we will assume that δ is some constant and will affect the constants in the asymptotic notations of the query complexities.

The trivial upper bound on the complexity of query-2 testers is $2n$ and of query-1 testers is n.

All our upper bound results hold for any permutation group G. But for our lower bound results we need some more structure on G. In Theorem 1.3 we assume that the group is transitive while Theorem 1.1 holds for primitive groups. Our main tool for primitive groups is the O'Nan–Scott Theorem (see Section 3). The following definition will help describe the structure of large primitive groups.

Definition 2.3. *Let $T_1, T_2, \ldots, T_s$ be disjoint sets and $r_1, r_2, \ldots, r_s$ be positive integers satisfying $\sum_{i=1}^{s} r_i = R$. Then by $\binom{T_1,T_2,\ldots,T_s}{r_1,r_2,\ldots,r_s}$ we mean the set of R-tuples formed by r_i distinct elements from the set T_i for all $1 \le i \le s$. That is,*

$$\binom{T_1, T_2, \ldots, T_s}{r_1, r_2, \ldots, r_s} = \left\{ \bigcup_{i=1}^{s} S_i \;\middle|\; S_i \subseteq T_i, |S_i| = r_i \right\}.$$

	Query-1 Complexity	Query-2 Complexity
1-sided-error	$\widetilde{\Theta}(\|V\|)^\dagger$	$\widetilde{\Theta}(\|V\|^{3/2})^\dagger$
2-sided-error	$\widetilde{O}(\|V\|), \Omega(\log(\|V\|))$	$\widetilde{O}(\|V\|^{3/2})$

† Matches the Fischer–Matsliah bounds.

Table 3: Corollaries of our results to Graph Isomorphism.

	Query-1 Complexity	Query-2 Complexity
1-sided-error	$\widetilde{\Theta}(\|V\|)$	$\widetilde{\Theta}\left(\left(\|V\|\binom{\|V\|}{k}\right)^{1/2}\right)$
2-sided-error	$\widetilde{O}(\|V\|), \Omega\left(\log\binom{\|V\|}{k}\right)$	$\widetilde{O}\left(\left(\|V\|\binom{\|V\|}{k}\right)^{1/2}\right)$

Table 4: Corollaries of our results to isomorphism of k-hypergraphs. Explanation of the *tilde* notation: In Table 1, the suppressed term is polylog(n); in Tables 2 and 3, it is polylog$(|V|)$; and in Table 4, it is polylog$\binom{|V|}{k}$.

2.3 Prior work

The query complexity of the property testing version of graph isomorphism has been studied by Fischer and Matsliah [18]. They gave tight bounds for 1-sided error and nontrivial bounds for 2-sided error (see Table 2).

Graph isomorphism is identical with G-isomorphism for the group $G = S_V^{(2)}$, where V is the vertex set of the graph. Thus our results specialize to this case (Table 3).

Isomorphism of k-hypergraphs corresponds to G-isomorphism for $G = \mathrm{Sym}_V^{(k)}$ (Table 4).

2.4 Chernoff bounds

We shall repeatedly use the following version of the Chernoff bounds, as presented by Alon and Spencer [6, Corollary A.14].

Let $X_1, X_2, \ldots, X_k$ be mutually independent indicator random variables and $Y = \sum_{i=1}^k X_i$. Let the expected value of Y be $\mu = \mathbb{E}[Y]$. For all $\alpha > 0$,

$$\Pr[|Y - \mu| > \alpha\mu] < 2e^{-c_\alpha \mu},$$

where $c_\alpha > 0$ depends only on α.

3 Primitive groups, 1-sided-error: tight bounds

3.1 Structure of primitive groups

The wreath product of permutation groups G and H, acting on the sets A and B, resp., is a permutation group generated by $|B|$ copies of G, acting independently on $|B|$ copies of A, and a copy of H which permutes the $|B|$ copies of A. Here is the formal definition.

Definition 3.1. Let G be a permutation group acting on a set A and H a permutation group acting on a set B. We define the *wreath product* $G \wr H$ as a permutation group acting on $A \times B$ as follows. $G \wr H$ contains the "base" subgroup G^B (the Cartesian product of $|B|$ copies of G), with each copy acting independently on the corresponding copy $A \times \{b\}$ of A ($b \in B$), i. e., if $(a, b) \in A \times B$ and $f \in G^B$ (f is a $B \to G$ function) then $(a, b)^f = (a^{f(b)}, b)$. $G \wr H$ also contains a subgroup H^* isomorphic to H acting only on the second components: for $(a, b) \in A \times B$ and $h \in H$ we define $(a, b)^h = (a, b^h)$. The group $G \wr H$ is defined as the subgroup of $\mathrm{Sym}(A \times B)$ generated by G^B and H^*.

It is easy to see that G^B is a normal subgroup of $G \wr H$ and its quotient by G^B is H:

$$H \cong (G \wr H)/G^B.$$

In particular, $|G \wr H| = |G|^{|B|}|H|$.

Note that if $|A|, |B| > 1$ then $G \wr H$ as defined above is an imprimitive group: the partition $(A \times \{b\} : b \in B)$ is $G \wr H$-invariant. Therefore this definition describes what is called the **imprimitive action** of the wreath product.

An important, often primitive, permutation group arises by the faithful action of $G \wr H$ on the set A^B of $B \to A$ functions, defined as follows: the base group acts coordinatewise, i. e., $p \in A^B$ and $f \in G^B$ then $(p^f)(b) = p(b)^{f(b)}$); and H acts by permuting the coordinates, i. e., $(p^h)(b) = b^h$.

It is easy to see that this defines a permutation group isomorphic to $G \wr H$; it is referred to as the **product action** of $G \wr H$.

The structure of primitive permutation groups is described by the O'Nan–Scott Theorem [25] (cf. [21]); the product action of the wreath product plays a central role in that description.

We only need a consequence of the O'Nan–Scott theorem, derived by Cameron [15].

Theorem 3.2 (O'Nan-Scott, Cameron)**.** *There is a (computable) constant c with the property that, if G is a primitive permutation group of degree n, then at least one of the following holds:*

1. $|G| \leq n^{c \log n}$.
2. *G is a subgroup of $\mathrm{Aut}(A_m^{(k)}) \wr S_\ell$ (product action) containing $(A_m^{(k)})^\ell$, where $A_m^{(k)}$ is the alternating group A_m acting on k-element subsets, where $m \geq 5$ and $1 \leq k < m/2$.*

Remark 3.3. While the O'Nan–Scott Theorem is elementary, Cameron has to invoke the power of the Classification of Finite Simple Groups to derive Theorem 3.2.

We need the following fact.

Fact 3.4. $\mathrm{Aut}(A_m) = S_m$ for all $m \geq 3$, $m \neq 6$. In particular, for all m,

$$|\mathrm{Aut}(A_m)| \leq 2m!$$

We now study the relation between the parameters n, m, k, ℓ in the case $|G| > n^{c\log n}$. In this case, the degree of G is given by

$$n = \binom{m}{k}^{\ell} \text{ and therefore } n \geq m^{\ell}. \tag{3}$$

It follows that $\ell \leq \log_2 n$. Now since $k < m/2$, we have

$$\binom{m}{k} \geq \left(\frac{m}{k}\right)^{k} > 2^k \text{ and therefore } k < \log_2 n. \tag{4}$$

A bound on the order of G follows, using Fact 3.4 and noting that A_m is isomorphic to $A_m^{(k)}$.

$$|G| \leq (2m!)^{\ell}(\ell!) < m^{m\ell}\ell^{\ell} \leq n^m \ell^{\ell} \text{ [From Equation (3)]} \tag{5}$$

Since $\ell \leq \log_2 n$, we have from Equation (5),

$$c(\log n)^2 < \log(|G|) < (m\log n + \ell \log \ell) \sim m \log n. \tag{6}$$

The last asymptotic equality holds because $\ell < \log n$ and therefore $\ell \log \ell = o(\log^2 n)$.

Therefore,

$$\log |G| \lesssim m \log n \text{ and } m \gtrsim c \log n. \tag{7}$$

Observation 3.5. *If $k = O(\sqrt{m})$ then* $\binom{m}{k} = \Theta\left(\frac{m^k}{k!}\right)$.

Corollary 3.6. *If G is a primitive permutation group of degree n and n is sufficiently large then either $|G| < (\log n)^3$ or we are in the second case of Theorem 3.2, $k \leq \sqrt{m}$, and*

$$\frac{m}{\sqrt{k}}\sqrt{\binom{m(1-\frac{1}{k})}{k-1}\binom{m}{k}^{\ell-1}} = \widetilde{\Omega}(\sqrt{n \log |G|}). \tag{8}$$

Proof. If we are in the first case of Theorem 3.2 then $\log |G| < c(\log n)^2 < (\log n)^3$ (for large n).

Now assume we are in the second case of Theorem 3.2. Let $k > \sqrt{m}$. Then

$$n = \binom{m}{k} > \left(\frac{m}{k}\right)^k > 2^k > 2^{\sqrt{m}}.$$

Therefore $m < (\log n)^2$ which implies by Equation (6) that $\log |G| < (\log n)^3$.

Now if $k \leq \sqrt{m}$ then the Corollary follows from the facts that $n = \binom{m}{k}^{\ell}$ (Equation (3)); $\log |G| \lesssim m \log n$ (Equation (7)); and Observation 3.5. $\square$

3.2 G-Agreeability

In this section we build our tools for the lower bounds.

Definition 3.7. Let $A, B \subseteq [n]$ and $p : A \to \{0, 1\}$ and $q : B \to \{0, 1\}$ be two partial assignments. We say that p and q are **compatible** if there exists a full assignment x on $[n]$ which is an extension of both p and q.

Let G be a permutation group on $[n]$. We say that p and q are G-**agreeable** if there exist $\pi_1, \pi_2 \in G$ such that p^{π_1} and q^{π_2} are compatible. Since G is a group this is same as saying that there exists an element $\pi \in G$ such that p^{π} and q are compatible. We say that p and q are agreeable through π.

Definition 3.8. Let G be a permutation group on $[n]$. Let x and y be two full assignments on $[n]$. We say that x and y are **k-G-agreeable** if for any sets $A, B \subseteq [n]$ with $|A|, |B| \leq k$, the partial assignments $x|_A$ and $y|_B$ are G-agreeable.

Observation 3.9. *If there exist two full assignments x, y on $[n]$ which are k-G-agreeable but satisfy $d_G(x, y) > \delta$ then the 1-sided-error query-2 complexity of G-isomorphism is greater than k.*

Proof. Let x and y be the two assignments in question. Now a 1-sided-error algorithm is forced to accept if it does not find a proof that x and y are not G-isomorphic. Let Q_x and Q_y be the sets of positions in x and y, respectively, that our 1-sided-error query-2 algorithm $\mathcal{A}$ queries. If x and y are k-G-agreeable and $|Q_x|, |Q_y| \leq k$ then this means there exists a permutation $\pi \in G$ such that $x^\pi|_{\pi(Q_x)}$ and $y|_{Q_y}$ are compatible. Hence $\mathcal{A}$ will have to accept, which is the wrong answer. □

The following is folklore.

Proposition 3.10. *Let G be a transitive group on $[n]$. Let us fix $A, B \subseteq [n]$ and let us select $\pi \in G$ uniformly at random. Then*

$$\mathbb{E}(|A^\pi \cap B|) = \frac{|A||B|}{n}. \tag{9}$$

Proof. By G-symmetry, for each $b \in B$ we have $\Pr(b \in A^\pi) = |A|/n$. Now the linearity of expectation yields the result. □

Corollary 3.11. *Let G be a transitive group on $[n]$. Let $A, B \subseteq [n]$ with $|A|, |B| \leq \epsilon\sqrt{n}$. Then,*

$$\Pr_{\pi \in G}[A^\pi \cap B = \emptyset] \geq (1 - \epsilon^2).$$

It follows that if A and B are the supports of the partial functions p and q, respectively, and $\epsilon \leq 1$ then p and q are G-agreeable. In particular, for a transitive group G, any two full assignments x and y on $[n]$ are $\lfloor\sqrt{n}\rfloor$-G-agreeable.

Proof. Immediate from Proposition 3.10 by Markov's inequality. □

The following is now immediate.

Corollary 3.12. *Let G be a transitive permutation group. The 1-sided-error query-2 complexity of G-isomorphism is at least $\lfloor\sqrt{n}\rfloor$.*

The following lemma has the most technical proof in this paper (see Section 3.4).

Lemma 3.13 (G-Agreeability Lemma for primitive groups)**.** *Let G be a primitive group. Then there exist two full assignments x and y on $[n]$ such that $d_G(x, y) \geq n/6$ and x and y are $\widetilde{\Omega}(\sqrt{n \log |G|})$-$G$-agreeable.*

3.3 Proof of the lower bounds for primitive groups

The lower bound in Theorem 1.1(b) is immediate by combining Observation 3.9 and Lemma 3.13.

Next we prove the lower bound for query-1 complexity stated in Theorem 1.1(a).

We recall the example for lower bound of 1-sided query-1 complexity of graph isomorphism given by Fischer and Matsliah [18]. In their case, the inputs are graphs. The unknown graph is the complete graph on n vertices while the known graph is the union of $n/2$ isolated vertices and a complete graph on $n/2$

vertices. Note that without querying more than $n/4$ pairs of vertices it is impossible to give a certificate of non-isomorphism. This gives the lower bound of $n/4$ for the graph isomorphism case.

We generalize this example to primitive groups of order $> n^{c\log n}$ where c is the constant from Part 1 of Theorem 3.2, the structure theorem for primitive groups. If $|G| \le n^{c\log n}$ then $\log|G| = \widetilde{\Theta}(1)$ and the lower bound $\widetilde{\Omega}(\log|G|)$ holds vacuously.

Now assume $|G| > n^{c\log n}$. By part 2 of Theorem 3.2, combined with Fact 3.4 and inequality (7) (which implies $m \ge 7$),

$$(A_m^{(k)})^\ell \le G \le S_m^{(k)} \wr S_\ell \tag{10}$$

where the wreath product acts in its product action ("$\le$" denotes "subgroup"). Thus G is isomorphic to a subgroup of $S_m \wr S_\ell$, acting in its imprimitive action on $\mathcal{V} = \cup_{i=1}^{\ell} V_i$, where $|V_i| = m$ and the V_i are all disjoint. This conversion translates any full assignment to a function from the set $\binom{V_1,\ldots,V_\ell}{k,k,\ldots,k}$ to $\{0,1\}$ and the group G can be thought of as acting on $\mathcal{V}$.

Now we define the known and the unknown parts of the input. We partition V_1 into three disjoint parts, namely V_a, V_b, and V_c, where $|V_a| = |V_b| = |V_c| = m/3$. The known input is

$$y(w) = 1 \text{ iff } w \in \binom{V_a, V_c, V_2 \ldots, V_\ell}{1, k-1, k, \ldots, k}.$$

The unknown input is

$$x(w) = 1 \text{ iff } w \in \binom{(V_a \cup V_b), V_c, V_2, \ldots, V_\ell}{1, k-1, k, \ldots, k}.$$

Note that one needs to make at least $m/6$ queries to give a certificate of non-isomorphism between the two inputs. Now from inequality (7) we obtain a lower bound of $\Omega\left(\frac{\log(|G|)}{\log n}\right)$.

3.4 Proof of the G-agreeability lemma for primitive groups

Proof of Lemma 3.13. We may assume n is large. If $\log|G| < (\log n)^3$ then $\sqrt{n\log|G|} = \widetilde{O}(\sqrt{n})$ and the result follows from the last sentence of Corollary 3.11.

Now assume $\log|G| \ge (\log n)^3$. Then, by Corollary 3.6, we have $k < \sqrt{n}$; we are in the second case of Theorem 3.2; and equation (8) holds. It follows, as before, that equation (10) holds and we can perform the translation from the product action of the wreath product to its imprimitive action as described above after equation (10).

If $\ell = 1$ and $G = S_m^{(2)}$ then G is the group of automorphisms of the complete graph on m vertices. This case was settled by Fischer and Matsliah [18]. We generalize their technique.

The rest of our proof has the following two parts:

- Define the full assignments x and y and prove that $d_G(x,y) > \delta n$ for some (absolute) constant $\delta > 0$.
- Let Q_x and Q_y be the query sets for x and y, respectively, such that both $|Q_x|$ and $|Q_y|$ is $\widetilde{O}(\sqrt{n\log|G|})$. Then we prove that there exist a permutation $\pi = \pi_1 \times \pi_2 \times \cdots \times \pi_\ell \in (A_m^{(k)})^\ell$ such that Q_x^π and Q_y are compatible.

We start with defining x.

Definition of the full assignments x and y

Let ϵ be a small constant that will be specified later. Let $U \subseteq V_1$ such that

$$|U| = m\left(1 - \frac{1}{k}\right).$$

Now we define x. Consider a partition of $V_1 \backslash U$ into two parts, U_1 and U_2, such that

$$|U_1| = m\left(\frac{1}{2k} + \epsilon\right) \text{ and } |U_2| = m\left(\frac{1}{2k} - \epsilon\right).$$

Set

$$x(w) = 1 \text{ iff } w \in \binom{U_1, U, V_2, \ldots, V_\ell}{1, k-1, k, k, \ldots, k}.$$

Now let us define y. Consider a partition of $V_1 \backslash U$ into two parts, W_1 and W_2, such that

$$|W_1| = m\left(\frac{1}{2k} - \epsilon\right) \text{ and } |W_2| = m\left(\frac{1}{2k} + \epsilon\right).$$

Set

$$y(w) = 1 \text{ iff } w \in \binom{W_1, U, V_2, \ldots, V_\ell}{1, k-1, k, k, \ldots, k}.$$

We will consider a map $\pi : \mathcal{V} \to \mathcal{V}$ that preserves V_i for all i, that is, for all i we have $\pi(V_i) = V_i$. Such a map gives a rearranges the positions in x. Note that weight of x (number of 1s) is $m\binom{m(1-\frac{1}{k})}{k-1}(\frac{1}{2k}+\epsilon)\binom{m}{k}^{\ell-1}$ and the weight of y is $m\binom{m(1-\frac{1}{k})}{k-1}(\frac{1}{2k}-\epsilon)\binom{m}{k}^{\ell-1}$. So from the difference in weights we see that

$$d_G(x,y) \geq 2\epsilon m \binom{m(1-\frac{1}{k})}{k-1}\binom{m}{k}^{\ell-1}.$$

For $k = 1$, the right-hand side is $2\epsilon m^\ell = 2\epsilon n$. If $k \neq 1$ then from Lemma 3.5 and the fact that $\left(1-\frac{1}{k}\right)^{k-1} \geq \frac{1}{e}$ we obtain, using Observation 3.5, that

$$2\epsilon m \binom{m(1-\frac{1}{k})}{k-1} \sim 2\epsilon \frac{m^k(1-\frac{1}{k}))^{k-1}}{(k-1)!} \geq \epsilon k \frac{2m^k}{ek!} = \Theta\left(\epsilon k \binom{m}{k}\right).$$

So if we choose $\epsilon = \frac{1}{12Ck}$ where C is the constant implied in the Θ notation, we have that

$$d_G(x,y) \geq \frac{1}{6}\binom{m}{k}^\ell = \frac{1}{6}n.$$

Now we move to the second part of the proof. Note that since G contains $(A_m^{(k)})^\ell$ it can permute the elements in V_i for each i in a huge number of ways. Since x and y differ only in their association with elements in V_1, in order to make the life of the tester difficult the only tricky part is how to permute the elements in V_1 so that the tester will not find a certificate of non-isomorphism. Now consider the case that $a \in U_1$ and $b \in W_2$ and for any w' if (a, w') is queried then (b, w') is not queried. If this is the case for a pair (a, b) then one can map a to b and then all the answers to queries associated with a and b are compatible. If there are a lot of such pairs (a, b), one can map each of those $a \in U_1$ to its corresponding $b \in W_2$. And then the rest of the elements in U_1 can be mapped to W_1 arbitrarily and the element of U_2 to the rest of elements of W_2 arbitrarily. All we need to show is that if the number of queries is small, one can find such a map.

Now we formalize this idea. Let Q_x and Q_y be query sets for x and y, respectively, such that $|Q_x|, |Q_y| \leq M$ where $M = \frac{m}{18\sqrt{k}}\sqrt{\binom{m(1-\frac{1}{k})}{k-1}\binom{m}{k}^{\ell-1}}$.

To prove that x and y are M-G-agreeable, we have to construct $\pi \in (A_m^{(k)})^\ell \subseteq G$ that maps $\mathcal{V}$ to $\mathcal{V}$ such that Q_x^π and Q_y are compatible.

If $a \in U_1$ then we define

$$q_x(a) = \left\{ w \in \binom{U_1, U, V_2, \ldots, V_\ell}{1, k-1, k, \ldots, k} \,\Big|\, w \in Q_x \text{ and } a \in w \right\}.$$

Similarly if $b \in W_2$, let

$$q_y(b) = \left\{ w \in \binom{W_2, U, V_2, \ldots, V_\ell}{1, k-1, k, \ldots, k} \,\Big|\, w \in Q_y \text{ and } b \in w \right\}.$$

Now by an averaging argument there exist sets $A \subseteq U_1$ and $B \subseteq W_2$ such that $|A| = |B| > \frac{2m}{9k}$ and for all $a \in A$ and $b \in B$ we have

$$|q_x(a)|, |q_y(b)| \leq \frac{9\sqrt{k}}{m} M.$$

Let $H = A^{(k-1)}_{m(1-\frac{1}{k})} \times (A_m^{(k)})^{\ell-1}$ acting on the set $\binom{U, V_2, \ldots, V_\ell}{k-1, k, k, \ldots, k}$. That is, every element in H maps U to U and V_i to V_i for all $i \geq 2$. Note that H acts transitively on the set $\binom{U, V_2, \ldots, V_\ell}{k-1, k, k, \ldots, k}$. Fix an arbitrary even bijection from A to B, i.e., an even permutation of $[n]$ which maps A to B. Let $a \in A$ be mapped to $b \in B$. For a $\pi' \in H$ we call a pair (a, b) π'-acceptable if $q_x(a)^{\pi'} \cap q_y(b) = \emptyset$. Using a simple probabilistic argument we show that there exists at least one $\pi' \in H$ such that at least ϵm of the pairs are π'-acceptable.

Pick a random element $\pi' \in H$. We want to calculate the probability of a pair (a, b) being π'-acceptable.

Consider the set q_a defined as

$$q_a = \{w \backslash \{a\} \mid w \in q_x(a)\}.$$

Similarly we define q_b. Note that q_a and q_b are two subsets of $\binom{U, V_2, \ldots, V_\ell}{k-1, k, \ldots, k}$. Note that M is picked such that the size of the set $\binom{U, V_2, \ldots, V_\ell}{k-1, k, \ldots, k}$ is $(18\sqrt{k}M/m)^2$. Since the sizes of both the sets q_a and q_b are less than $9\sqrt{k}M/m$, from Corollary 3.11 we see that

$$\Pr_{\pi' \in H}[(a, b) \text{ is } \pi'\text{-acceptable}] \geq \frac{3}{4}.$$

So by linearity of expectation the expected number of pairs that are π'-acceptable is $\geq \frac{3}{4}\frac{2m}{9k} = \frac{m}{6k} > 2\epsilon m$. So there exists a permutation $\pi' \in H$ such that ϵm of the (a, b) pairs are acceptable. These acceptable pairs along with the permutation π' give a map from a set $A' \subseteq A \subseteq U_1$ to a set $B' \subseteq B \subseteq W_2$ such that $|A'| = |B'| = 2\epsilon m$. Now we have

$$|U_1 \backslash A'| = |W_1|.$$

Hence π' and the map from the acceptable pairs can be extended to a mapping $\pi : \mathcal{V} \to \mathcal{V}$ by mapping $U_1 \backslash A'$ to W_1 and U_2 in x to $W_2 \backslash B'$. Thus Q_x^π and Q_y are compatible.

Finally from Corollary 3.6 we have $M = \widetilde{\Omega}(\sqrt{n \log |G|})$. □

4 Upper bounds for transitive groups

In this section we prove Proposition 1.2.

Definition 4.1. We define a **query sequence** as the sequence positions (elements of $[n]$) of the input string that are queried. If Q is a query sequence then $|Q|$ is the number of elements in the sequence.

The proofs of both parts of Proposition 1.2 are rather simple applications of the Chernoff bound; we describe the proofs for completeness.

Proof of Part (a) of Proposition 1.2. In this part we only have to query x. Let us choose a real number p, $0 < p < 1$, appropriately (see below). The length of the query sequence Q will be less than $2pn$. The following is the test:

1. Construct the query sequence Q by choosing elements of $[n]$ independently at random with probability p.
2. If m is more than $2pn$ then accept.
3. If $|Q| \leq 2pn$ then query x at Q. So we obtain the partial function $x|_Q$.
4. If for some $\pi \in G$ the partial function $x|_Q^\pi$ and y are compatible then accept. Otherwise reject.

Note that the query complexity for the test is less that $2pn$. Also note that if x and y are G-isomorphic then the test always accepts. So all we need to show is that if x and y are ϵ-far from being G-isomorphic then the test accepts with probability less than $1/3$.

By Chernoff Bound we have that

$$\Pr\left[|Q| > 2pn\right] < \exp\left(-cnp\right),$$

where c is some constant. So the probability that the test accepts in Step 2 less than $\exp(-cnp)$.

Now if x and y are ϵ-far from being G-isomorphic then for any permutation $\pi \in G$, we know that x^π and y differ in at least ϵn bits. For any fixed $\pi \in G$, $x|_Q^\pi$ and y are compatible if and only if none of those ϵn bit positions are not in Q. Since the bits are chosen at random the probability that $x|_Q^\pi$ and y are compatible is at most $(1-p)^{\epsilon n}$. By union bound the probability that the tester accepts in Step 4 is less than $|G|(1-p)^{\epsilon n}$. Thus if x and y are ϵ-far from being G-isomorphic then

$$\Pr\left[\text{The tester accepts}\right] < \left(\exp\left(-cnp\right) + |G|(1-p)^{\epsilon n}\right).$$

If $p = O((1 + \log|G|)/\epsilon n)$ then the right hand side of the above equation and hence the probability of error of the tester is less than $1/3$. Thus the query complexity for this test is less than $O((1+\log|G|)/\epsilon)$. $\square$

Proof of Part (b) of Proposition 1.2. In this part we have to query both x and y. Again we choose a real number p, $0 < p < 1$, appropriately (see below). The total length of the query sequence will be $4pn$. The following is the test:

1. Construct two query sequences Q_1 and Q_2, by choosing the elements of $[n]$ independently at random with probability p for each query sequence.

2. If $|Q_1|$ or $|Q_2|$ is more than $2pn$ then accept.
3. Query the bits of x and y corresponding to Q_1 and Q_2 respectively. So we obtain the partial functions $x|_{Q_1}$ and $y|_{Q_2}$.
4. If for some group element $\pi \in G$, the partial function $x|^{\pi}_{Q_1}$ and the partial function $y|_{Q_2}$ are compatible then accept. Otherwise reject.

Note that the query complexity for the test is less that $4pn$. Also note that if x and y are G-isomorphic then the test always accepts. So all we need to show is that if x and y are ϵ-far from being G-isomorphic then the test accepts with probability less than $1/3$.

By Chernoff Bound and union bound we have that

$$\Pr\left[\, |Q_1| \text{ or } |Q_2| > 2pn\right] < 2\exp\left(-cnp\right),$$

where c is some constant. So the probability that the test accepts in Step 2 less than $2\exp\left(-cnp\right)$.

For any group element π, let D_π be the set of positions of the bits of x^π that differ from y. By definition if x and y are ϵ-far from being G-isomorphic then for any permutation $\pi \in G$, we know that $|D_\pi| > \epsilon n$. Now $x|^{\pi}_{Q_1}$ and $y|_{Q_2}$ are compatible if only if for every $D_\pi \cap Q_1 \cap Q_2 = \emptyset$. Since the bits are chosen at random the probability that this happens is at most $(1-p^2)^{\epsilon n}$.

By union bound the probability that the tester accepts in Step 4 is less than $|G|(1-p^2)^{\epsilon n}$. Thus if x and y are ϵ-far from being G-isomorphic then

$$\Pr\left[\text{The tester accepts}\right] < \left(2\exp\left(-cnp\right) + |G|(1-p^2)^{\epsilon n}\right).$$

If we take $p = O(\sqrt{(1+\log|G|)/\epsilon n})$ the error is less than $1/3$. Thus the query complexity for this test is less than $O(\sqrt{n(1+\log|G|)/\epsilon})$ □

5 Lower bounds for Transitive Groups

In the section we prove Theorem 1.3.

We begin with two easy observations. Recall that a *k-hypergraph* is a family of k-subsets ("edges") of the vertex set V. The *degree* of vertex $v \in V$ is the number of edges containing v. The hypergraph is *regular* if every vertex has the same degree. A subset of the vertices is a *cover* it it hits every edge.

Observation 5.1. *If a k-uniform regular hypergraph has n vertices then every cover has size $\geq n/k$.*

Proof. By straightforward counting: let the hypergraph have m edges and be r-regular. Let T be a cover. Then $km = rn$ and $|T|r \geq m$. □

Corollary 5.2. *If G is a transitive permutation group and S is a k-subset of $[n]$ then there exist at least n/k^2 pairwise disjoint G-shifts of S.*

Proof. Let T be a maximal union of disjoint G-shifts of S. So T hits all G-shifts of S. The set of G-shifts of S is a regular k-uniform hypergraph (regular because G is transitive); therefore, by Observation 5.1, we have $|T| \geq n/k$. This requires at least n/k^2 shifted copies of S. □

We shall need the following result.

Definition 5.3. *Let D be any distribution on $\{0,1\}^n$. Then for any set $Q \subseteq [n]$ we define $D|Q$ to be the distribution on $\{0,1\}^{|Q|}$ obtained by picking $x \in \{0,1\}^n$ according to the distribution D and then projecting to the indices in Q.*

Theorem 5.4 ([16, 19])**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a function. Let $Y \subseteq \{0,1\}^n$ be the set of x such that $f(x) = 1$ and let $N \subseteq \{0,1\}^n$ be the set of x which are ϵ-far from satisfying $f(x) = 1$. Let D_Y and D_N be distributions on Y and N, respectively. If for all $Q \subseteq [n]$ of size q, and all $g \in \{0,1\}^{|Q|}$, we have $\frac{2}{3}\Pr_{D_Y|Q}[g] < \Pr_{D_N|Q}[g]$ then any 2-sided-error property test for f requires at least q queries.*

Proof of Theorem 1.3. Let x be a full assignment. For any subset $P \subseteq [n]$ of size k and any $z \in \{0,1\}^k$ let $n_z^P(x) = |\{\pi \in G : x^\pi|_P = z\}|$. We call x "almost k-universal" if for all $z \in \{0,1\}^k$ and for all $P \subseteq [n]$ of size k, we have $|n_z^P(x) - \frac{|G|}{2^k}| \leq \frac{|G|}{5(2^k)}$. Note that this means that if we pick $\pi \in G$ at random then for any $z \in \{0,1\}^k$ and for any subset P we have

$$\left|\Pr_{\pi \in G}[x^\pi|_P = z] - \mu\right| \leq \mu/5$$

where $\mu = 1/2^k$.

We prove the existence of an almost k-universal string using the probabilistic method. Pick a random full assignment x. For any subset $P \subseteq [n]$ of size k and $z \in \{0,1\}^k$ we will estimate $n_z^P(x)$. By Corollary 5.2 we can place $\frac{n}{k^2}$ disjoint G-shifts of the subset P in $[n]$. Let $\mathcal{S}$ denote the set of disjoint copies of P. Let $v_i^z(x)$ be the $(0,1)$-indicator variable indicating whether the i-th G-shift of P in $\mathcal{S}$ is same as z. Since x is chosen randomly the random variables $v_i^z(x)$ are independent. Let $v_z(x) = \sum v_i^z(x)$. That is $v_z(x)$ be the number of times z occurs in $\mathcal{S}$. The expected value of $v_z(x)$ is $\frac{n}{k^2 2^k}$. Using the Chernoff bound we obtain

$$\Pr\left[\left|v_z(x) - \frac{n}{k^2 2^k}\right| > \frac{n}{5k^2 2^k}\right] \leq 2\exp\left(-\frac{c_{1/5}n}{k^2 2^k}\right).$$

So using the union bound we have

$$\Pr\left[\forall \pi \in G, \forall z \in \{0,1\}^k, \forall P, \left|v_z(x^\pi) - \frac{n}{k^2 2^k}\right| \leq \frac{n}{5k^2 2^k}\right] \geq 1 - 2\exp\left(-\frac{c_{1/5}n}{k^2 2^k}\right)|G|\binom{n}{k}2^k.$$

If $|G| = 2^{O(n^{1-\epsilon})}$ then for any positive ϵ and $k \leq (\epsilon - \gamma)(\log n)$ (where $\gamma > 0$), this probability is non-zero. Now since we had exactly (n/k^2) number of disjoint copies of P, so there is a string x such that

$$\forall z \in \{0,1\}^k, \forall P \subseteq [n], |P| = k, \left|\Pr_{\pi \in G}[x^\pi|_P = z] - \mu\right| \leq \mu/5$$

where $\mu = 1/2^k$. Thus for any positive ϵ and $k \leq (\epsilon - \gamma)(\log n)$ (where $\gamma > 0$) there exists an "almost k-universal" string. Let it be x_Y.

Similarly one can show that existence of a full assignment such that it is $\frac{1}{3}$-far from x_Y and still "almost k-universal." Probability that a random string is $\frac{1}{3}$-close to x_Y is $\frac{1}{2^{\Omega(n)}}$. Using the same argument as above we can say that the probability that a random string is $\frac{1}{3}$-far from x_Y and is an "almost universal" string is more than $\left(1 - \frac{1}{2^{\Omega(n)}} - 2\exp\left(-\frac{c_{1/5}n}{k^2 2^k}\right)|G|\binom{n}{k}2^k\right)$. This is also positive for $k \leq \frac{\epsilon \log n}{2}$ (since $|G| = 2^{O(n^{1-\epsilon})}$). Hence for $k \leq \frac{\epsilon \log n}{2}$ there exists a full assignment in $\{0,1\}^n$ which is $\frac{1}{3}$-far from x_Y and is "almost k-universal." Let it be x_N.

Now let x_Y be the string to which we have full access. The unknown string is chosen from the following two distributions.

- D_Y: Uniform random G-shift of x_Y.
- D_N: Uniform random G-shift of x_N.

Now we know that x_Y and x_N are $\frac{1}{3}$-far. Also since x_Y and x_N are "almost k-universal" for $k = (\epsilon/2)\log n$, so for all subsets $P \subseteq [n]$ of size $(\epsilon/2)\log n$ and all $z \in \{0,1\}^{(\epsilon/2)\log n}$, we have

$$2/3 \Pr_{\pi \in G}\left[x_Y^{\pi}|_P = z\right] \leq \Pr_{\pi}\left[x_N^{\pi}|_P = z\right].$$

Now by Theorem 3.5 we can say that it will be impossible to test G-isomorphism with less than $(\epsilon/2)\log n$ queries. So the query-1 complexity of any property tester of G-isomorphism is $\Omega(\log n)$.

□

6 Tight bounds and comparisons

Our main result, Theorem 1.1, gives tight bounds for primitive groups in the 1-sided error model.

In this section we point out that we have tight bounds for small transitive groups; "small" means their order is polynomially or quasiolynomially bounded as a function of their degree. These include a number of classes of groups of interest. In particular, all permutation represetations of all finite simple groups except the alternating groups have size $n^{O(\log n)}$. Moreover, all finite simple groups except the alternating groups and the "classical groups" of unbounded dimension (linear, symplectic, orthogonal, and unitary groups) have polynomially bounded order (cf. [15]).

Corollary 6.1. *Let G be a transitive permutation group. Assume $|G| = n^{O(1)}$. Then the query-1 complexity of 1-sided-error and 2-sided-error property testing of G-isomorphism is* $\Theta(\log n)$.

Proof. If $|G| = n^{o(1)}$ then from Proposition 1.2 we obtain the upper bound of $O(\log n)$ for both 1-sided-error and 2-sided-error cases. Theorem 1.3 gives the matching lower bound. □

Corollary 6.2. *Let G be a transitive group. Assume* $\log(|G|) = (\log n)^{O(1)}$. *Then the 1-sided-error query-2 complexity of G-isomorphism is* $\widetilde{\Theta}(\sqrt{n})$.

Proof. The lower bound follows from Corollary 3.12; the upper bound from Proposition 1.2. □

7 Open questions

We have obtained tight bounds for the 1-sided-error query complexity when the group is primitive. Obtaining tight bounds for the 2-sided error query complexity that match the Fischer–Matsliah bounds in the special case of Graph Isomorphism would be of considerable interest. We note that no tight bounds are known even for Graph Isomorphism in the 2-sided error, query-2 model.

Another natural direction would be to extend our bounds to all transitive groups in the one-sided error model. For primitive groups, we expressed our tight bounds in terms of the (approximate) values of the parameters n and $|G|$ (the degree and the order of the permutation group). It is possible that these parameters alone do not suffice in the case of transitive groups and any tight bound must depend on the finer structure of the transitive group G.

A test case would be the automorphism group of a complete binary tree in its action on the leaves. Let T_n denote the complete binary tree with $n = 2^h$ leaves. Let G be the action of the automorphism group of the tree on the leaves. (This is the h-fold iterated wreath product $S_2 \wr S_2 \wr \cdots \wr S_2$.) Let us index the positions in the string x of length n by the leaves of T_n. Then G permutes the positions of x. For this particular transitive group, the query-1 and query-2 complexities of testing G-isomorphism are wide open both in the 1-sided error and 2-sided error models.

We have reason to believe that a solution for this group would bring us close to solving the corresponding problem for all transitive groups.

References

[1] Noga Alon, Eldar Fischer, Michael Krivelevich, and Mario Szegedy. Efficient testing of large graphs. In *Combinatorica*, volume 20, pages 451–476, 2000.

[2] Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: It's all about regularity. In *Proc. 38th ACM Symp. on Theory of Computing (STOC)*, pages 251–260, 2006.

[3] Noga Alon and Asaf Shapira. Testing subgraphs in directed graphs. In *Proc. 35th ACM Symp. on Theory of Computing (STOC)*, pages 700–709, 2003.

[4] Noga Alon and Asaf Shapira. A characterization of the (natural) graph properties testabale with one-sided error. In *Proc. 46th Ann. Symp. on Foundations of Computer Science (FOCS)*, pages 429–438, 2005.

[5] Noga Alon and Asaf Shapira. Linear equations, arithmetic progressions and hypergraph property testing. In *Proc. 16th ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 708–717, 2005.

[6] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. Wiley-Interscience (John Wiley & Sons), New York, 1992.

[7] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and hardness of approximation problems. In *J. ACM*, volume 45, pages 501–555, 1998.

[8] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: a new characterization of NP. In *J. ACM*, volume 45, pages 70–122, 1998.

[9] László Babai. On the order of uniprimitive permutation groups. In *Annals of Mathematics*, volume 113, pages 553–568, 1981.

[10] László Babai. On the order of doubly transitive permutation groups. In *Inventiones Math.*, volume 65, pages 473–484, 1982.

[11] László Babai, Lance Fortnow, and Carsten Lund. Nondeterministic exponential time has two-prover interactive protocols. In *Computational Complexity*, volume 1, pages 3–40, 1991.

[12] László Babai and Eugene M. Luks. Canonical labeling of graphs. In *Proc. 15th ACM Symp. on Theory of Computing (STOC)*, pages 171–183, 1983.

[13] László Babai, Eugene M. Luks, and Ákos Seress. Permutation groups in NC. In *Proc. 19th ACM Symp. on Theory of Computing (STOC)*, pages 409–420, 1987.

[14] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. In *J. of Computer and System Sciences*, volume 47, pages 549–595, 1993.

[15] Peter J. Cameron. Finite permutation groups and finite simple groups. In *Bull. London Math. Soc.*, volume 13, pages 1–22, 1981.

[16] Eldar Fischer. The art of uninformed decisions: A primer to property testing. In G. Rozenberg G. Paun and A. Salomaa, editors, *Current Trends in Theoretical Computer Science: The Challenge of the New Century*, volume I, pages 229–264. World Scientific Pub. Co. Inc., 2004.

[17] Eldar Fischer. The difficulty of testing for isomorphism against a graph that is given in advance. In *SIAM J. on Computing*, volume 34, pages 1147–1158, 2005.

[18] Eldar Fischer and Arie Matsliah. Testing graph isomorphism. In *SIAM J. on Computing*, volume 38, pages 2007–225, 2008.

[19] Eldar Fischer, Ilan Newman, and Jiři Sgall. Functions that have read-twice constant width branching programs are not necessarily testable. In *Random Structures and Algorithms*, volume 24, pages 175–193, 2004.

[20] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. In *J. ACM*, volume 45, pages 653–750, 1998.

[21] Martin W. Liebeck, Cheryl E. Praeger, and Jan Saxl. On the O'Nan-Scott theorem for finite primitive permutation groups. In *J. Austral. Math. Soc. Ser. A*, volume 44, pages 389–396, 1988.

[22] Eugene M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. In *J. Computer and System Sciences*, volume 25, pages 42–65, 1982.

[23] Dana Ron. Property testing. In J. H. Reif S. Rajasekaran, P. M. Pardalos and J. D. P. Rolim, editors, *Handbook of Randomized Computing*, volume II, chapter 15, pages 597–650. Kluwer Academic Publishers, 2001.

[24] Ronitt Rubinfeld and Madhu Sudan. Robust characterization of polynomials with applications to program testing. In *SIAM J. on Computing*, volume 25, pages 252–271, 1996.

[25] Leonard L. Scott. Representations in characteristic p. In *Proc. Sympos. Pure Math*, volume 37, pages 319–322. A.M.S., 1980.