2013 IEEE 54th Annual Symposium on Foundations of Computer Science

Faster Canonical Forms For Strongly Regular Graphs (Extended Abstract)

László Babai*, Xi Chen†, Xiaorui Sun†, Shang-Hua Teng‡, and John Wilmes*
*University of Chicago
Email: laci@cs.uchicago.edu *and* wilmesj@math.uchicago.edu
†Columbia University
Email: xichen@cs.columbia.edu *and* xiaoruisun@cs.columbia.edu
‡University of Southern California
Email: shanghua@usc.edu

***Abstract*—We show that a canonical form for strongly regular (s. r.) graphs can be found in time $\exp(\widetilde{O}(n^{1/5}))$ and therefore isomorphism of s. r. graphs can be tested within the same time bound, where n is the number of vertices and the tilde hides a polylogarithmic factor.**

The best previous bound for testing isomorphism of s. r. graphs was $\exp(\widetilde{O}(n^{1/3}))$ (Spielman, STOC 1996) while the bound for GI in general has been standing firmly at $\exp(\widetilde{O}(n^{1/2}))$ for three decades. (These results, too, provided canonical forms.)

The previous bounds on isomorphism of s. r. graphs (Babai 1980 and Spielman 1996) were based on the analysis of the classical individualization/refinement (I/R) heuristic. The present bound depends on a combination of a deeper analysis of the I/R heuristic with Luks's group theoretic divide-and-conquer methods following Babai-Luks (STOC 1983) and Miller (1983).

Our analysis builds on Spielman's work that brought Neumaier's 1979 classification of s. r. graphs to bear on the problem. One of Neumaier's classes, the line-graphs of Steiner 2-designs, has been eliminated as a bottleneck in recent work by the present authors (STOC'13). In the remaining hard cases, we have the benefit of "Neumaier's claw bound" and its asymptotic consequences derived by Spielman, some of which we improve via a new "clique geometry."

We also prove, by an analysis of the I/R heuristic, that, with known (trivial) exceptions, s. r. graphs have $\exp(\widetilde{O}(n^{9/37}))$ automorphisms, improving Spielman's $\exp(\widetilde{O}(n^{1/3}))$ bound.

No knowledge of group theory is required for this paper. The group theoretic method is only used through an easily stated combinatorial consequence (Babai–Luks, 1983 combined with Miller, 1983).

While the bulk of this paper is joint work by the five authors, it also includes two contributions by subsets of the authors: the clique geometry [BW] and the automorphism bound [CST].

***Keywords*-Algorithms, graph isomorphism, strongly regular graphs**

I. Introduction

A. History and statement of the main result

It follows from the early theory of interactive proofs that the Graph Isomorphism problem (GI) is not NP-complete unless the polynomial-time hierarchy collapses [1] (see [2] for a self-contained proof and references). On the other hand, for three decades now, the best known upper bound on the complexity of GI has been $\exp(\widetilde{O}(\sqrt{n}))$ where n is the number of vertices and the tilde hides a polylog factor [3], [4], [5].

A *strongly regular graph* with parameters (n, k, λ, μ) is a regular graph with n vertices and degree k such that each pair of adjacent vertices has λ common neighbors and each pair of non-adjacent vertices has μ common neighbors.

The class of strongly regular graphs, while not believed to be GI-complete, has long been identified as a hard case for GI (cf. [6]).

A function F from a class $\mathcal{F}$ of finite structures to itself is called a *canonical form* if for every $X, Y \in \mathcal{F}$ we have $X \cong F(X)$ and $X \cong Y$ if and only if $F(X) = F(Y)$. Isomorphism of members of $\mathcal{F}$ can be decided by two applications of a canonical form function and comparison of the outputs.

The set $\mathrm{Iso}(G, H)$ of $G \to H$ isomorphisms is a coset of the automorphism group $\mathrm{Aut}(G)$ and can be concisely represented by a set of $O(n)$ generators of $\mathrm{Aut}(G)$ together with a single $G \to H$ isomorphism.

Our main result is the following.

Main Theorem I.1. *Let G, H be s. r. graphs. A canonical form for G can be computed, and consequently, isomorphism of G and H can be decided, in time $\exp(\widetilde{O}(n^{1/5}))$. Moreover, the set $\mathrm{Iso}(G, H)$ can be computed within the same time bound.*

Disconnected s. r. graphs are disjoint unions of

0272-5428/13 $26.00 © 2013 IEEE
DOI 10.1109/FOCS.2013.25

cliques of equal size; we refer to them and to their complements as the **"trivial s. r. graphs."** Nontrivial s. r. graphs have diameter 2 and therefore have degree $k \geq \sqrt{n-1}$. Since the complement of a s. r. graph is s. r., we shall assume throughout the paper that

$$\sqrt{n-1} \leq k \leq (n-1)/2. \tag{1}$$

Babai showed in 1980 [7] (cf. [8]) that isomorphism of s. r. graphs can be tested in time

$$\exp(\widetilde{O}(n/k)). \tag{2}$$

In the light of inequality (1) this gives an overall bound of $\exp(\widetilde{O}(\sqrt{n}))$. This was improved by Spielman [9] (STOC 1996) to $\exp(\widetilde{O}(n^{1/3}))$; no further improvement was obtained until the present paper.

B. Graphic s. r. graphs and vertex coloring

Let $G = (V, E)$ be an (undirected) graph; V is the set of vertices and E the set of edges. The *line-graph* $L(G)$ has the vertex set E with two edges of G being adjacent in $L(G)$ if they share a vertex in G.

$L(G)$ is a nontrivial s. r. graph precisely if $G = K_v$ ($n = \binom{v}{2}$) or $K_{v,v}$ ($n = v^2$) ($v \geq 3$); we refer to these graphs and their complements as the **"graphic s. r. graphs."**

We shall work with *vertex-colored graphs* (colors are preserved by isomorphisms by definition). While canonical forms for the graphic s. r. graphs can easily be found in linear time, the vertex-colored versions of each of the two classes of graphic s. r. graphs are *GI-complete* via a quadratic reduction. So an $\exp(\widetilde{O}(n^{1/4-\epsilon}))$ GI bound for either of these colored classes would improve the general bound for GI to $\exp(\widetilde{O}(n^{1/2-2\epsilon}))$ and would therefore in itself be a major result. However, our main results remain valid for the vertex-colored versions of all s. r. graphs except the graphic ones.

This comment also underlines the significance of being able to separate graphic s. r. graphs from the remaining ones. A spectral separation was first achieved by Seidel [10] who showed that for $n \geq 29$, the s. r. graphs with least eigenvalue -2 are precisely the s. r. line graphs. This was powerfully generalized by Neumaier to separate additional classes of s. r. graphs which we call *geometric* (below); following Spielman, we build on Neumaier's classification.

C. Geometric s. r. graphs and conference graphs

For the purposes of this paper, a *finite geometry* is a set $\mathcal{P}$ of v "points" and set $\mathcal{L}$ of subsets of $\mathcal{P}$ called "lines." The "length" of a line is the number of its points; we assume (i) all lines have the same length $\ell \geq 3$; (ii) every pair of lines shares at most one point. The *line-graph* of a finite geometry $(\mathcal{P}, \mathcal{L})$ has $\mathcal{L}$ for its set of vertices; adjacency corresponds to intersection.

Two classes of finite geometries are of particular interest. In a *Steiner 2-design* there is a line through every pair of points. The points of a *transversal design* are partitioned into ℓ classes of equal size; there is a line through a pair of points if and only if the points do not belong to the same class.

The line graphs of these two classes of geometries are s. r.; we refer to them and to their complements as **geometric s. r. graphs**.

A **conference graph** is a s. r. graph of degree $k = (n-1)/2$ with $\mu = (n-1)/4$ and $\lambda = \mu - 1$.

D. The Neumaier classification

The following result was central to Spielman's work and remains central to ours.

Theorem I.2 (Neumaier, 1979)**.** *A s. r. graph is trivial, graphic, geometric, or a conference graph, or satisfies "Neumaier's claw bound."*

Neumaier's claw bound [11], [12] is an inequality involving the parameter μ and the eigenvalues of the adjacency matrix (see e. g. [9]). We shall only need the following asymptotic consequences of the claw bound.

Theorem I.3. *Suppose G is a s. r. graph satisfying Neumaier's claw bound. Assum $k = o(n)$. Then*

(a) $\mu \sim k^2/n$ *(Spielman)*
(b) $\lambda = O(k^{2/3}\mu^{1/3}) = O(k^{4/3}n^{-1/3})$ *(Spielman)*
(c) *If* $k = \Omega(n^{2/3})$ *then* $\lambda = O((k\mu)^{1/2}) = O(k^{3/2}n^{-1/2})$ *([13], see Section V)*

Parts (a) and (b) are implicit in [9].

We point out the philosophical significance of these results. Written as $\mu/n \sim (k/n)^2$, item (a) can be interpreted as saying that the neighborhoods of nonadjacent pairs of vertices are "asymptotically independent." Since most pairs of vertices are not adjacent, this is the typical behavior. While the neighborhoods of adjacent vertices can be heavily positively correlated, items (b) and (c) limit this correlation. We found this intuition helpful.

In the rest of this section we discuss how we dispose of the cases that do not satisfy Neumaier's claw bound. The rest of the paper will assume the claw bound.

The following results appear in Miller [14] (1978) for transversal designs and in [15] and [16] (STOC'13) for Steiner 2-designs.

Theorem I.4. *Canonical forms for transversal designs and for Steiner 2-designs can be computed in $n^{O(\log n)}$ time where n is the number of lines.*

To apply Theorem I.4 to geometric graphs, we first need to reconstruct the underlying geometry from the line-graph. This is not always possible. However, up to a degree threshold $\sim n^{2/3}$ the geometry can be uniquely reconstructed in polynomial time, as shown by Miller [14] for transversal designs and by Spielman [9] for Steiner designs. All this is implicit in Neumaier's work, along with the following.

Proposition I.5. *Let G be a geometric s. r. graph. Then either the underlying geometry can be uniquely reconstructed in polynomial time, or G satisfies Neumaier's claw bound.*

The assumption $k = o(n)$ in Theorem I.3 is justified by (2) [7]. Indeed, for $k \geq n/\log n$, we can construct canonical forms in quasipolynomial time ($\exp(\widetilde{O}(1))$) by (2). This observation takes care in particular of conference graphs (a case in Neumaier's classification).

In the remaining cases we shall assume $k \leq n/\log n$ and that Neumaier's claw bound holds, and therefore items (a), (b), (c) of Theorem I.3 hold as well.

E. The individualization/refinement (I/R) heuristic and the automorphism bound

A classical heuristic approach to the GI problem is the "individualization/refinement" (I/R) method: we individualize t vertices (assign unique colors to them), then apply a canonical color refinement process (cf. Sec. II). If we are lucky enough that the refinement process *completely splits the graph* (assigns a unique color to each vertex after refinement) then we obtain a canonical form for the graph in time $n^{t+O(1)}$ by repeating the process for all possible choices of the list of t vertices to be individualized and picking the lexicographically first among the resulting labeled graphs.

The results of Babai [7] and Spielman [9] as well as [15] and [16] were based solely on the analysis of the I/R heuristic.

This heuristic, however, has an inherent mathematical limitation: if the individualization of t vertices completely splits the graph G then $|\operatorname{Aut}(G)| \leq n^t$. Babai has **conjectured** for three decades that for nontrivial, non-graphic s. r. graphs $|\operatorname{Aut}(G)| = \exp(n^{o(1)})$, but the best result to date in this direction is

$$|\operatorname{Aut}(G)| \leq \exp(\widetilde{O}(n^{9/37})) \tag{3}$$

by [17] (see Section VIII), improving Babai's $\exp(\widetilde{O}(n^{1/2}))$ and Spielman's $\exp(\widetilde{O}(n^{1/3}))$ bounds and inspiring the present work.

Our algorithmic result, however, goes beyond the $\exp(\widetilde{O}(n^{9/37}))$ bound and therefore necessarily involves more than an intricate analysis of I/R. Indeed, part of our work is based on Luks's group-theoretic divide-and-conquer through a simply-stated combinatorial consequence, Theorem III.1, implicit in Babai–Luks (1983) [3] combined with a result of Miller (1983) [18]. Theorem III.1 reduces the problem of finding canonical forms to a relaxed analysis of I/R, where instead of targeting a complete split, our goal is a "color-t-bounded graph" (see the definition before Theorem III.1).

F. Detailed results

Our main result follows from a combination of the following results with Theorem I.3.

Theorem I.6. *Let G be a s. r. graph satisfying Neumaier's claw bound. Assume $k = o(n)$. Then a canonical form of G can be found in time*

(a) $n^{O(\mu+\log n)}$ *(Sec. IV)*
(b) $\exp(\widetilde{O}((n/k)^{1/2}))$ *for $k = o(n^{2/3})$ (Spielman)*
(c) $\exp(\widetilde{O}(1+\lambda/\mu))$ *for $k = \omega((n\log n)^{1/2})$ (Sec. VI)*

Parts (b) and (c) of this result combined with parts (a) and (c) of Theorem I.3 yield
(b′) $\exp(\widetilde{O}((n/k)^{1/2}))$ for all k.

Remark I.7. Part (a) is the only part where the proof uses the group theory method. In all other cases, the time bounds stated suffice to *list all isomorphisms*. (This is true for the $n^{O(\log n)}$ bound in the geometric cases as well [15], [16].)

Deriving the Main Theorem I.1 from Theorems I.6, I.3, Prop. I.5, and Eq. (2)*:* We settle the trivial and the graphic cases in a straightforward manner in polynomial (in fact, linear) time. The following cases are done in quasipolynomial time: the cases $k \geq n/\log n$ by (2); and the geometric cases where we don't have the claw bound, using Theorem I.4 and Proposition I.5. Finally in the remaining cases we use part (a) of Theorem I.6 in combination with part (a) of Theorem I.3 for $k \leq n^{3/5}$ noting that $\mu \sim k^2/n \lesssim n^{1/5}$ in this range; and statement (b′) after Thm. I.6 for $n^{3/5} \leq k \leq n/\log n$. The bottleneck timing arises at $k \sim n^{3/5}$. ∎

We now state our bounds in terms of k and n, along with the history for comparison.

Theorem I.8. *Canonical forms for a s. r. graph G can be computed in time*

(a) $\exp(\widetilde{O}(n/k))$ *for all k [7]*
(b) $\exp(\widetilde{O}(\sqrt{n/k}))$ *for $k = o(n^{2/3})$ [9] unless G is geometric*
(c) $n^{O(\log n)}$ *if G is geometric and does not satisfy Neumaier's claw bound ([15], [16] plus Prop. I.5)*
(d) *In the present paper we prove*
  (d1) $\exp(\widetilde{O}(k^2/n))$ *for $k = O(n/\log n)$, improving Spielman's bound in the range $k \leq n^{3/5}$*

(d2) $\exp(\widetilde{O}(\sqrt{n/k}))$ *for* $k = o(n)$*, extending the range of Spielman's bound.*

We still use (b) for $n^{3/5} \leq k = o(n^{2/3})$ *and (a) for* $k = \Omega(n)$ *(and (c) for the geometric case).*

G. Outline of the paper

The two basic techniques, I/R and the group theoretic method, are sketched in Sections II and III, resp. Section IV gives the full proof of part (a) of Theorem I.6 via a color-μ-bound (see def. before Thm. III.1). This is the only part of the paper that depends on the group theory method. For lack of space, all the other proofs are only sketched. Section V describes the clique geometry and indicates the proof of the improved bound on λ (part (c) of Theorem I.3). Sections VI and VII outline two proofs of the case of large degree (part (c) of Theorem I.6) based on different strategies. Finally, Section VIII outlines the proof of the $\exp(\widetilde{O}(n^{9/37}))$ bound on the number of automorphisms.

Three follow-up papers will contain the complete proofs. The bulk of the work will be covered by the "full version" of this paper; the clique structure will be presented in [13], and the automorphism bound in [17].

II. Individualization and refinement

We consider vertex-colored graphs $G = (V, E, f)$ where $f : V \to \{1, \ldots, s\}$ is a map for some s (the number of colors). The coloring f defines a partition of the vertices into color classes. A coloring g of the graph (V, E) is a *refinement* of the coloring f of (V, E) if $(\forall v, w \in V)(g(v) = g(w) \Rightarrow f(v) = f(w))$.

A *color-refinement* operator $\mathcal{R}$ over a class $\mathcal{C}$ of colored graphs assigns to every member of $\mathcal{C}$ another member with the same underlying graph and a refined coloring. We call the refinement operator $\mathcal{R}$ *canonical* if for all $G, H \in \mathcal{C}$ we have $\mathrm{Iso}(G, H) = \mathrm{Iso}(\mathcal{R}(G), \mathcal{R}(H))$.

The coloring of G is *stable* if the coloring of $\mathcal{R}(G)$ defines the same partition as G. Repeated application of a (canonical) refinement operator to a colored graph $G = (V, E, f)$ leads to a (canonical) stable refinement $G^* = (V, E, f^*)$.

We say that a vertex v has a *unique color* if no other vertex shares the color of v.

We say that a color refinement operator *completely splits* the colored graph $G = (V, E, f)$ if the stable refinement G^* of G assigns a unique color to each vertex. We note that if this is the case and $\mathcal{R}$ is canonical then $|\mathrm{Aut}(G)| = 1$ and sorting the vertices of G in the order of their f^*-values puts G into a canonical form.

Individualizing a vertex v means the assignment of a unique color to v. **Depth-d stabilization** is the process of individualizing d vertices and refining to a stable coloring. (It seems this was first introduced in [19] in the context of the Weisfeiler-Leman refinement.) We say that *depth-d stabilization achieves a certain type of coloring* if there exists a set S of size $|S| \leq d$ such that after individualizing S, the stable refinement is of the stated type.

Proposition II.1. *Let $\mathcal{C}$ be a class of colored graphs, closed under isomorphisms. Let $\mathcal{R}$ be a canonical refinement operator over $\mathcal{C}$. Suppose depth-ℓ stabilization completely splits $G \in \mathcal{C}$. Then $|\mathrm{Aut}(G)| \leq n^\ell$ and we can compute a canonical form of G and list $\mathrm{Iso}(G, H)$ for any colored graph H in time $T(\mathcal{R})n^{\ell+O(1)}$, where $T(\mathcal{R})$ denotes the cost of one execution of $\mathcal{R}$ on a colored graph on n vertices.* ∎

The *naive color refinement* process first replaces the coloring f by the coloring f' where $f'(v) = (f(v); m_i(v) : i \in [s])$ where $m_i(v)$ denotes the number of neighbors of v of color i. Next we lexicographically order the set $\{f'(v) : v \in V\}$ of strings, and define $\widehat{f}(v)$ to be the rank of $f'(v)$ in this ordering. (The $\widehat{f}$ notation hides the fact that $\widehat{f}$ depends not only on f but on (V, E, f).) The correspondence $(V, E, f) \to (V, E, \widehat{f})$ is clearly a canonical refinement; this is one round of the naive refinement process. The *naive refinement operator* replaces f by the stable coloring f^*.

Except for Section VII, the only refinement operator we consider is the naive refinement. The same is true for papers [7], [9], [15]. Paper [16] and Section VII of the present paper use a slight extension of naive refinement, subsumed by the Weisfeiler-Leman refinement [19].

III. The group theory method

The proof of part (a) of Theorem I.6 is based on a combination of Luks's group theoretic divide-and-conquer methods [20] and the I/R technique. This combination was developed by Babai and Luks in [3], see esp. Section 4.5 of that paper. Below we state the general principle implicit in [3].

Let G be a colored graph with color classes $C_1, \ldots, C_m$ (in this order). Let $B_k = \bigcup_{i=1}^{k} C_i$. We say that G *is color-t-bounded* if for $k = 1, \ldots, m$, no set of $t+1$ vertices in C_k has the same set of neighbors in B_{k-1}. (Note that $B_0 = \emptyset$, so for $k = 1$ this condition means $|C_1| \leq t$.)

Theorem III.1. *Let G be a color-t-bounded colored graph. Then a canonical form for G can be computed in $n^{O(t)}$ time.*

A weaker form of this result is implicit in [3]; this weaker form would suffice for our purposes. We stated

the result in this stronger form for its simplicity. To prove this stronger form, we need to adapt the proof of [18, Theorem 2] (G. L. Miller, 1983) to the [3] technique. (This route is also indicated in [3].)

Proposition III.2. *Let $\mathcal{C}$ be a class of colored graphs, closed under isomorphisms, and $\mathcal{R}$ a canonical refinement operator over $\mathcal{C}$. Let $G \in \mathcal{C}$ have n vertices. Suppose depth-ℓ stabilization results in a color-t-bounded graph. Then we can compute a canonical form of G in time $T(\mathcal{R})n^{\ell+O(t)}$, where $T(\mathcal{R})$ denotes the cost of one execution of $\mathcal{R}$ on a colored graph on n vertices.*

The proof is analogous to the proof of Prop. II.1.

Our next result implies Part (a) of Theorem I.6.

Theorem III.3. *After individualizing $O(\log n)$ vertices of G, naive refinement yields a color-μ-bounded graph.*

Remark III.4. We note a structural obstruction to this strategy. This comment, ending with the paragraph after Proposition III.5, requires the basics of group theory and is not needed for the rest of the paper.

Following Luks [20], let Γ_r denote the class of those groups L which have a subgroup chain $L = L_0 \geq L_1 \geq \cdots \geq L_m = \{1\}$ (the "$\leq$" sign placed between groups means "subgroup") such that for all $i \geq 1$ we have $|L_{i-1} : L_i| \leq r$. Equivalently, a group L belongs to Γ_r if all composition factors of L are isomorphic to subgroups of the symmetric group S_r.

Proposition III.5. *If depth-ℓ stabilization (with respect to any canonical refinement operator) of the graph G results in a color-t-bounded graph then $\mathrm{Aut}(G)$ has a Γ_t-subgroup of index $\leq n^\ell$.*

Ideally we would wish the conclusion (a non-algorithmic mathematical fact) to hold for some small value of $\ell + t$. It is open whether it holds with $\ell + t \leq n^{1/5-\epsilon}$ for all notrivial, non-graphic s. r. graphs. But a slightly weaker parameter that might still potentially permit the application of some version of the group theory method *is* polylogarithmically bounded, see Theorem IX.1.

IV. Generating the graph μ vertices at a time

In this section we prove Theorem III.3 and thereby part (a) of Theorem I.6.

For $a \in V(G)$, let $N(a)$ denote the set of neighbors of a (so $a \notin N(a)$). For $A \subseteq V(G)$, let $N(A) = \bigcup_{a\in A} N(a)$. Given a set A of vertices, we define $\widehat{A} = A \cup G(A)$, where

$$G(A) = \{u \in V : \exists a, b \in A \text{ s.t. } a \neq b,\ a \not\sim b, \text{ and } u \in N(a) \cap N(b)\}.$$

We say a set A is *closed* if $A = \widehat{A}$. The *closure* $\overline{A}$ of A is the smallest closed set containing A. We say a set A *generates* a set B if $B \subseteq \overline{A}$. We show that there exists a small set A of vertices that generates all of V.

Lemma IV.1. *There exists a set $A \subset V$ with $|A| = O(\log n)$ such that $\overline{A} = V$.*

Proof of Theorem III.3: Observe that by construction, for any set $B \subseteq V$, each vertex in $G(B)$ has a pair of nonadjacent neighbors in B. Therefore, no set of $\mu + 1$ vertices in $G(B)$ can have the same set of neighbors in B.

By Lemma IV.1, there is some set $A \subseteq V$ with $|A| = O(\log n)$ such that $\overline{A} = V$. Individualize A and refine the coloring until it is stable. Order the color-classes: start with the members of A (a single vertex in each class), and then repeatedly apply the $B \mapsto \widehat{B}$ operator to the union of the color-classes already listed; add the list of color classes in $\widehat{B} \setminus B$ in any order. By the foregoing, our colored graph now is color-μ-bounded. ∎

We will prove Lemma IV.1 in three stages. The first and largest step is to generate a set $\overline{A}$ containing a positive fraction of each of the neighborhoods of two vertices, starting from a set A of logarithmic size (Lemma IV.4). This is proved inductively; in the inductive step (Lemma IV.5), we show that a subset of a neighborhood of a vertex will generate a significantly larger subset of a neighborhood of a vertex after individualizing two additional vertices.

To complete the proof of Lemma IV.1, we prove that once we have a set A of the sort guaranteed in Lemma IV.4, we already have that $\overline{A} = V$. First, we show in Lemma IV.7 that $\widehat{A}$ already covers a constant fraction of the space, and then, in Lemma IV.8, we complete the proof. Both steps follow by counting the number of edges leaving the set we have already generated.

Two preliminary estimates and some additional notation are required. For $x \in V$, we write $N^+(x) = N(x) \cup \{x\}$. For $A \subseteq V$ and $y \in V$, consider the set

$$A + y := N(A \setminus N^+(y)) \cap N(y). \tag{4}$$

Clearly, $|A+y| \leq \mu|A|$. Furthermore, we note $A+y \subseteq \widehat{A \cup \{y\}} \cap N(y)$.

Lemma IV.2. *Let $L \subset N(A)$. Then*

$$\mathbb{E}_y(|(A + y) \cap L|) \geq |L|(k - \lambda - 1)/n \sim |L|(\mu/k).$$

Proof: For each $u \in L$, designate a neighbor $u' \in A$. For $y \in V$, let $\vartheta_u(y)$ denote the indicator of the event that $y \in N(u)\setminus N^+(u')$. Since $|N(u)\cap N^+(u')| = \lambda+1$, we have $\mathbb{E}_y(\vartheta_u) = (k - \lambda - 1)/n$.

Now, if $u \in L$ and $\vartheta_u = 1$ then $u \in (A+y) \cap L$. So $|(A+y) \cap L| \geq \sum_{u \in L} \vartheta_u$ and $\mathbb{E}_y(|(A+y) \cap L|) \geq \sum_{u \in L} \mathbb{E}_y(\vartheta_u) = |L|(k-\lambda-1)/n$. ∎

Fact IV.3. *Let $m \geq 1$ be real and let X be a random variable such that $0 \leq X \leq m$ and $\mathbb{E}(X) = 1$. Then $P(X > 1-\epsilon) \geq \epsilon/(m-1+\epsilon)$. In particular, $P(X > 1/2) \geq 1/(2m-1)$.* ∎

Lemma IV.4. *There exists a pair of distinct vertices $x, y \in V$ and a set A with $|A| = O(\log n)$ such that $|\overline{A} \cap N(x)| \geq k/100$ and $|\overline{A} \cap N(y)| \geq k/100$.*

To prove the lemma, we will show that given a set $A \subset N(x)$ with $|A| \leq k/100$, we can grow A into a new set $A' \subseteq N(x')$ whose size exceeds $|A|$ by a constant factor, and which is generated by A and two additional vertices; moreover, we have many choices for x'. Here is the precise statement.

Lemma IV.5 (Growth lemma)**.** *Let $x \in V$ and $A \subset N(x)$. Suppose $|A| < k/100$. Then there are $\Omega(n/\mu)$ vertices y such that there are $\Omega(n/\mu)$ vertices z such that $|\overline{A'} \cap N(z)| \gtrsim (9/8)|A|$ where $A' = A \cup \{y, z\}$.*

Lemma IV.4 will then follow by induction.

Proof of Lemma IV.4 (based on the Growth lemma (Lemma IV.5))*:* Fix adjacent vertices y_0 and z_0, and let $A_0 = \{y_0, z_0\}$. When A_i, y_i, and z_i have been defined, and $|\overline{A_i} \cap N(z_i)| \leq k/100$, by Lemma IV.5, there exists a pair (y_{i+1}, z_{i+1}) of vertices such that, setting $A_{i+1} = A_i \cup \{y_{i+1}, z_{i+1}\}$, we have

$$|\overline{A_{i+1}} \cap N(z_{i+1})| \gtrsim (9/8)|\overline{A_i} \cap N(z_i)|.$$

Thus, for some $t = O(\log n)$, we have a set A_t of vertices such that $|A_t| \leq 2t$ and $|\overline{A_t} \cap N(z_t)| > k/100$. Furthermore, Lemma IV.5 ensures that we have $\Omega(n/\mu) = \omega(\log n)$ choices for each z_i. We repeat the construction to obtain sets A'_i for $i \leq t'$ where $t' = O(\log n)$ using a sequence of pairs (y'_i, z'_i) where the set $\{z_1, \ldots, z_t\}$ is disjoint from the set $\{z'_1, \ldots, z'_{t'}\}$. We then have $|\overline{A'_{t'}} \cap N(z'_t)| \geq k/100$ and $z'_{t'} \neq z_t$. So $A = A_t \cup A'_{t'}$ has the desired property. ∎

In this section we only prove the Growth lemma (Lemma IV.5) for $\mu \geq 8$.

The case of small μ presents an added technical difficulty because in those cases, the sets as given in the proof may not grow at all, or not at a constant rate. The proof for $\mu \leq 7$ will use similar ideas as the proof below for larger μ except we shall need to use two additional vertices y, z per round rather than just one to nudge our sets to grow. We defer the proof of the case $\mu \leq 7$ to the full version of the paper.

For $\mu \geq 8$, the Growth lemma follows from

Lemma IV.6. *Suppose $\mu \geq 8$. Let $A \subset N(x)$ and suppose $|A| < k/100$. Then there are $\Omega(n/\mu)$ vertices $y \in V$ such that $|\overline{A \cup \{y\}} \cap N(y)| \gtrsim (9/8)|A|$.*

Proof: Define R as the set of triples (a, v, b) satisfying $a \in A, b \in N(x) \setminus A$, and $v \in N(a) \cap N(b)$. Let $Q = \{(a, v, b) \in R : v \notin N^+(x)\}$. We claim that

$$|Q| \gtrsim (99/200)|A|k(\mu - 3). \tag{5}$$

The lemma then follows. Indeed, if v is such that $(a, v, b) \in Q$ for some a, b, then clearly $v \in N(A)$. Furthermore, such a vertex v can appear at most

$$|N(v) \cap N(x) \cap A| \cdot |N(v) \cap N(x) \setminus A| \leq \frac{\mu^2}{4}$$

times in Q. Therefore,

$$\begin{aligned} |N(A)| &\geq 4|Q|/\mu^2 \gtrsim (99/50)((\mu-3)/\mu)(k/\mu)|A| \\ &\geq (99/80)(k/\mu)|A|. \end{aligned}$$

It then follows by Lemma IV.2 that $\mathbb{E}_y(|A+y|) \gtrsim (99/80)|A|$. Let $Z = |A+y|$ and let us apply Fact IV.3 to the random variable $X = Z/\mathbb{E}_y(Z)$ with $\epsilon = 1/11$. Since $Z \leq \mu|A|$ for all y, we obtain that

$$P(Z > (9/8)|A|) \gtrsim \frac{\epsilon}{(80\mu/99) - 1 + \epsilon} = \frac{9}{80\mu - 90},$$

proving the lemma.

We now prove inequality (5). Define $Q' = \{(a, v, b) \in R : v \notin N(x)\}$ and $W = \{(a, v, b) \in R : a \not\sim b \text{ and } v \neq x\}$. Thus, $Q \subset Q'$, and $Q' \setminus Q$ is the set of triples of the form (a, x, b) where $a \in A$ and $b \in N(x) \setminus A$. In particular, $|Q'| = |Q| + |A|(k - |A|)$.

For every $a \in A$, there are $\geq k - \lambda - |A|$ vertices $b \in N(x) \setminus A$ with $a \not\sim b$. For every such pair (a, b), there are $\mu - 1$ vertices v such that $(a, v, b) \in W$. Thus, since $\lambda = o(k)$, we have $|W| \gtrsim |A|(k - |A|)(\mu - 1)$.

The set $W \setminus Q'$ is the collection of triples (a, v, b) of distinct vertices in $N(x)$ such that $a \in A$, $b \notin A$, and (a, v, b) induces a path (i. e., $a \sim v \sim b$ and $a \not\sim b$).

Let F be the set of edges $\{a, b\}$ such that $a \in A$ and $b \in N(x) \setminus A$. For any $(a, v, b) \in W \setminus Q'$, exactly one of the edges $\{a, v\}$ and $\{v, b\}$ is in F. Fix $\{a, b\} \in F$ with $a \in A$ and $b \notin A$, and let $K = K(a, b) = N(a) \cap N(b) \cap N(x)$. There are $\leq |N(a) \cap N(x) \setminus K| = \lambda - |K|$ vertices w such that $(w, a, b) \in W \setminus Q'$. Similarly, there are $\leq |N(b) \cap N(x) \setminus K| = \lambda - |K|$ vertices w such that $(a, b, w) \in W \setminus Q'$. Thus,

$$|W \setminus Q'| \leq \sum_{\{a,b\} \in F} 2(\lambda - |K(a, b)|)$$

On the other hand, $Q' \setminus W$ contains all triples (a, v, b) such that $\{a, b\} \in F$ with $a \in A$ and $v \in N(a) \cap$

$N(b) \setminus N(x)$. Thus, for fixed $\{a,b\} \in F$, there are exactly $|N(a) \cap N(b) \setminus K| = \lambda - |K|$ vertices v such that $(a,v,b) \in Q' \setminus W$. Thus,

$$|Q' \setminus W| \geq \sum_{\{a,b\} \in F} (\lambda - |K(a,b)|)$$

so that $|Q' \setminus W| \geq |W \setminus Q'|/2$. It follows that

$$|Q'| = |Q' \cap W| + |Q' \setminus W| \geq |Q' \cap W| + |W \setminus Q'|/2$$
$$\geq |W|/2 \gtrsim (1/2)|A|(k - |A|)(\mu - 1)\,.$$

Since $|Q| = |Q'| - |A|(k - |A|)$, we have $|Q| \gtrsim (1/2)|A|(k-|A|)(\mu - 3)$. Inequality (5) now follows, since $|A| < k/100$. ∎

Lemma IV.7. *Let $A \subseteq V$. Suppose there exist distinct vertices x, y such that $|A \cap N(x)| \geq ck$ and $|A \cap N(y)| \geq ck$ for some constant $c > 0$. Then $|\widehat{A}| \gtrsim c^2 n$.*

Proof: Let $X = A \cap N(x) \setminus N(y)$ and $Y = A \cap N(y) \setminus N(x)$. Without loss of generality, assume $|Y| \leq |X|$. We have $|N(x) \cap N(y)| \leq \lambda$ and therefore $|Y| \geq ck - \lambda \sim ck$.

Define $Q = \{(a,b,v) : a \in X, b \in Y, a \not\sim b, v \in N(a) \cap N(b)\}$. Note that if $(a,b,v) \in Q$ then $v \in \widehat{A}$.

Since no vertex in Y is adjacent to x, each vertex in Y has at most μ neighbors in X. Therefore

$$|Q| \geq (|X| - \mu)|Y|\mu \sim c^2 k^2 \mu.$$

If $v \in N(x)$ and $(a,b,v) \in Q$ then $v \in N(b) \cap N(x)$ and $b \notin N(x)$, while $a \in N(v) \cap N(x)$. It follows that there are $\leq \mu\lambda|Y|$ such triples (a,b,v). Similarly, there are $\leq \mu\lambda|X|$ triples $(a,b,v) \in Q$ such that $v \in N(y)$. For every other vertex v, we have $|N(v) \cap X| \leq \mu$ and $|N(v) \cap Y| \leq \mu$, so there are $\leq \mu^2$ pairs a, b such that $(a,b,v) \in Q$. Thus, the number of distinct vertices $v \in V \setminus (X \cup Y)$ such that $(a,b,v) \in Q$ for some a, b is at least

$$\frac{|Q| - \mu\lambda|X| - \mu\lambda|Y|}{\mu^2} \gtrsim \frac{c^2k^2}{\mu} \sim c^2 n.$$

In particular, $|\widehat{A}| \gtrsim c^2 n$. ∎

Lemma IV.8. *Let $A \subseteq V$ is such that $|A| = \Omega(n)$. Then for n sufficiently large, $\overline{A} = V$.*

Proof: Let B be the collection of vertices with at least $\lambda + 2$ neighbors in A. Clearly $B \subset \widehat{A}$. Thus, it suffices to show that every vertex $x \notin A \cup B$ has at least $\lambda + 2$ neighbors in $A \cup B$ for n sufficiently large. Indeed, if $x \notin A \cup B$, then $|A \setminus N(x)| \geq \Omega(n - \lambda) = \Omega(n)$. Every vertex in $A \setminus N(x)$ has μ neighbors in $N(x)$, so the number of edges between $N(x)$ and $A \setminus N(x)$ is $\Omega(\mu n) = \Omega(k^2)$. Thus, $\mathbb{E}_{y \in N(x)}(|N(y) \cap A|) = \Omega(k)$, and since each vertex in $N(x)$ has at most k neighbors in A, it follows by Fact IV.3 that at least $\Omega(k)$ neighbors of x each have at least $\Omega(k)$ neighbors in A. Since $\lambda = o(k)$, for n sufficiently large, x has more than $\lambda + 2$ neighbors in B. ∎

Finally, we complete the proof of Lemma IV.1 by using Lemmas IV.4, IV.7, and IV.8 in this order. ∎

V. Clique geometry and the λ bound

Spielman found the following remarkable lower bound on the density of the subgraph induced by the set of common neighbors of a pair of adjacent vertices.

Lemma V.1 (Spielman)**.** *Let u, v be adjacent vertices. Then there are at most $(k - \lambda - 1)(\mu - 1)$ ordered pairs of nonadjacent vertices in $N(u) \cap N(v)$.*

Using this bound we can prove that for large λ, almost all vertices in $N(u) \cap N(v)$ induce a clique. Here is the exact statement.

Corollary V.2. *If $k\mu = o(\lambda^2)$ then we have the following structure.*

(i) *For every pair $\{u, v\}$ of adjacent vertices there is a unique maximal clique $C(u,v)$ of order $\sim \lambda$ such that $u, v \in C(u,v)$. Let us call these cliques* special.

(ii) *Let $D(u,v) = N(u) \cap N(v) \setminus C(u,v)$. Then the vertices of $D(u,v)$ each have at most μ neighbors in $C(u,v)$, and hence have degree $o(\lambda)$ in the subgraph of G induced on $N(u) \cap N(v)$. The special cliques are therefore easily identified, since the vertices of $C(u,v)$ have degree $\sim \lambda$ in $N(u) \cap N(v)$.*

(iii) *If $x \neq y$ are vertices in $C(u,v)$ then $C(x,y) = C(u,v)$. In other words, two distinct special cliques have at most one vertex in common.*

We defer the proof to [13].

This structure can be viewed as an approximate version of the well-studied class of "partial geometries" (where all lines must have equal length and the number of lines from x to ℓ must always be the same). The special cliques are the "lines." They are of almost equal length $\sim \lambda$; every pair of adjacent vertices belongs to a unique line; two lines intersect in at most one point. Furthermore, if a point x is not on a line ℓ then there are at most μ lines connecting x to points of ℓ.

We now use this structure to derive our bound on λ.

A *hypergraph* is a pair $\mathcal{H} = (V, \mathcal{E})$ where $\mathcal{E} \subseteq 2^V$ is the set of *edges*. Let $n = |V|$. The *degree* $\deg(v)$ of the vertex $v \in V$ is the number edges containing v. We say

that $\mathcal{H}$ is *nearly d-regular* if $\deg(v) \sim d$ for all $v \in V$, and $\mathcal{H}$ is *nearly λ-uniform* if $|A| \sim \lambda$ for all $A \in \mathcal{E}$.

Lemma V.3. *Let $\mathcal{H} = (V, \mathcal{E})$ be a nearly d-regular, nearly λ-uniform hypergraph with n vertices, such that every pair of edges shares at most one vertex. Assume $d \to \infty$ with n. Then $\lambda \lesssim \sqrt{n}$.*

Proof: Count the triples (v, A, B) where $v \in V$, $A, B \in \mathcal{E}$, $A \neq B$, and $v \in A \cap B$ in two ways. ∎

Proof of Theorem I.3 (c): Assume $k = \Omega(n^{2/3})$. Note that $\mu \sim k^2/n = \Omega(n^{1/3})$, so $k\mu = \Omega(n)$. We need to show that $\lambda = O((k\mu)^{1/2})$. Assume instead that $k\mu = o(\lambda^2)$. Let $\mathcal{H}$ be the hypergraph on V whose edges are the "special cliques" $C(u, v)$. Then $\mathcal{H}$ is nearly λ-uniform and nearly k/λ-regular. It follows by Lemma V.3 that $\lambda^2 \lesssim n = O(k\mu)$, contradicting the assumption that $k\mu = o(\lambda^2)$. ∎

VI. The $\exp(\widetilde{O}(1 + \lambda/\mu))$ bound

In this section, we will outline a proof of Theorem I.6 (c). An alternative proof will be outlined in Section VII.

In view of Prop. II.1, the theorem follows from the following lemma. We set $\nu = \max\{\lambda, \mu\}$.

Lemma VI.1. *Assume $\nu = \omega(\log n)$. For some $d = O((\nu/\mu)\log^3 n)$, depth-d stabilization completely splits G.*

The overall strategy of the proof of Lemma VI.1 is similar to that of [15]. By individualizing vertices in three stages, we gradually refine the coloring of G. In the first stage, we ensure that for a positive fraction of vertices x, the neighborhood $N(x)$ intersects many color classes. Here is a precise statement. We say that a subset T of a colored set is *stable* if T is a union of color-classes.

Lemma VI.2. *For some $t = O((\nu/\mu)\log^2 n)$, depth-t stabilization yields a stable set A with $|A| = \Omega(n)$ such that for every $x \in A$ there are $\Omega(k/\nu)$ disjoint stable sets C such that $|N(x) \cap C| = \Omega(\mu)$.*

Proof sketch: We show that with a single individualization, we can either increase the number of color classes by a factor of $(1+\Omega(\mu/(\nu\,\mathrm{polylog}(n))))$, or else obtain a collection of $\Omega(k/\nu)$ color classes of bounded size. By iterating this process $\widetilde{O}(\nu/\mu)$ times, we ensure that the latter event obtains. The neighborhoods of these bounded-size color classes will be the stable sets C in the statement of the lemma. The set of all vertices $x \in V$ such that $|N(x) \cap C| = \Omega(\mu)$ is stable, and we show that the number of such x is $\Omega(n)$. ∎

Next, we give unique colors to a substantial fraction of all vertices.

Lemma VI.3. *Suppose there is a stable set A with $|A| = \Omega(n)$ such that for every $x \in A$ there are $\Omega(k/\nu)$ disjoint stable sets C such that $|N(x) \cap C| = \Omega(\mu)$. Then for some $t = O((\nu/\mu)\log^3 n)$, depth-t stabilization produces $\Omega(n)$ vertices with unique colors.*

First, in Lemma VI.4, we show that if a pair of vertices has $\Omega(k/\nu)$ different colors appearing in the symmetric difference of their neighborhoods, then with high probability they will receive different colors in the stable refinement after individualizing $\widetilde{O}(\nu/\mu)$ random vertices.

Lemma VI.4. *For any $\delta > 0$ there is some $M > 0$ such that the following holds. Suppose $x, y \in V$ are such that $N(x) \setminus N(y)$ intersects at least $\delta k/\nu$ color classes. Individualize at least $M(\nu/\mu)\log^2 n$ random vertices. Then for n sufficiently large, x and y get different colors in the stable refinement with probability at least $1/2$.*

Proof sketch: By taking intersections with the neighborhoods of random vertices, we iteratively shrink the color classes in $N(x) \triangle N(y)$. With high probability, some color class shrinks to a single vertex after $\widetilde{O}(\nu/\mu)$ such intersections. ∎

Sketch of proof of Lemma VI.3: The hypotheses of Lemma VI.4 are satisfied by pairs of nonadjacent vertices in the set A of Lemma VI.3, and so individualizing some set of $\widetilde{O}(\nu/\mu)$ random vertices will suffice to partition the set A of Lemma VI.3 into color classes of size at most λ in the stable refinement. This coloring of A will induce a relatively fine coloring of neighborhoods for those vertices which intersect A substantially. Hence, by repeating the argument, we can give different colors to every pair of vertices whose neighborhoods have large intersection with A. ∎

Lemma VI.5. *Suppose $\Omega(n)$ vertices have unique colors. Then for n sufficiently large, G is completely split in the stable refinement.*

The proof is similar to that of Lemma IV.8. ∎

Proof of Lemma VI.1: Apply Lemmas VI.2, VI.3, and VI.5, in that order. ∎

The details of the proofs are left to the full paper.

VII. Bipartite structure

In this section we outline an alternative proof of Theorem I.6 (c). As in [16] (as well as in [7], [9]), the overall strategy is to distinguish pairs of vertices.

Given a s.r. graph, we achieve this by building a family of bipartite structures that can be utilized in a

multi-stage analysis. When aiming at distinguishing two vertices $u \neq v$, we simultaneously grow two sequences of bipartite structures, one for u and one for v, with the assistance of some random vertices (seeds). At each step, the bipartite structures have the following properties: either their interactions with a small number of random seeds can introduce the desired asymmetry with high probability, or their interactions with "not too many" random seeds likely produce another pair of bipartite structures with measurable progress. Note that in order to perform multilevel analyses of non-trivial s. r. graphs whose diameter is 2 it is usually essential to build substructures within a s. r. graph that can stretch.

For a vertex u, we focus on the induced bipartite subgraph between $N(u)$ and $V \setminus N^+(u)$. Specifically, we focus on a family of *bipartite systems*, each consisting of a sequence of induced bipartite subgraphs $((A_1, B_1), \ldots, (A_\gamma, B_\gamma))$, where the A_i are disjoint subsets of $N(u)$ and $B_i \subseteq V \setminus N^+(u)$. For our construction and analysis, in addition to demanding that every (A_i, B_i) be dense enough, we further require that (i) the sizes of all the A_i be within a factor of 2 of each other, (ii) all degrees involved by vertices in $\bigcup_i B_i$ be within a factor of 2 of each other, (iii) the numbers of B_i to which each vertex in $\bigcup_i B_i$ belongs be within a constant factor of each other.

With these strong "regularity" conditions, we have the following property: Assume $((A_i, B_i))$ and $((A'_i, B'_i))$ are a pair of bipartite structures built for u and v, respectively. If $|A_i \cap A'_i| = o(|A_i|)$ and $|A_i|$ $(= |A'_i|)$ is smaller than $k/\max(\lambda, \mu)$, then we have $|B_i \cap B'_i| = o(|B_i|)$. Thus, if $|\bigcup_i B_i|$ is very close to n, then the interaction of a small number of random seeds w with $\bigcup_i B_i$ and $\bigcup_i B'_i$ is likely to produce the asymmetry that we aim for: for some i, $w \in B_i$ but $w \notin B'_i$.

Our initial bipartite systems for u, v are simply $((N(u), V \setminus N^+(u))$ and $((N(v), V \setminus N^+(v))$, respectively, which do not meet the size condition above since $|N(u)| = |N(v)| = k$ is greater than $k/\max(\lambda, \mu)$. To make progress, we draw a small number of random seeds and use the following process to *partition* a bipartite system $((A_i, B_i))$ to $((A'_j, B'_j))$. For a seed z, if $z \in B_i$ for some i, then we extract a new induced bipartite graph (A', B'), where A' contains all the vertices of A_i that are also neighbors of z, and B' contains all neighbors of A' in $V \setminus N^+(u)$. We collect all such new induced bipartite graphs and then, "clean up" the bipartite graphs to make the new bipartite system satisfy the desired "regularity" conditions. Our two goals are to ensure that (1) the union of the B-part of the new bipartite system still contains almost all vertices in $V \setminus N^+(u)$, and (2) the A-part of the new bipartite system is smaller than the old one by a factor of $O(n^{-\Omega(1)})$. Thus, a constant number of steps is sufficient to obtain the desired structures.

We show that our construction of bipartite systems is isomorphism-invariant, and that $\widetilde{O}(1 + \lambda/\mu)$ random seeds are sufficient to distinguish all pairs of vertices with high probability. As a result, this gives an alternative algorithm for testing isomorphism of s. r. graphs that satisfy Neumaier's claw bound with running time $\exp(\widetilde{O}(1 + \lambda/\mu))$. For technical reasons, our analysis works only for s. r. graphs of $k = O(n^{1-\epsilon})$ for some arbitrary constant $\epsilon > 0$. For larger k we refer to the bound (2) [7]. We defer the details to the full paper.

VIII. The combinatorial $9/37$ bound

We now outline the proof of inequality (3). We state the algorithmic result which implies (3).

Theorem VIII.1. *If G and H are non-trivial, non-graphic s. r. graphs then* $\mathrm{Iso}(G, H)$ *can be listed in time* $\exp(\widetilde{O}(n^{9/37}))$.

This section is not required for our main algorithmic result.

The proof follows from a combination of results stated in the Introduction and the following lemma that gives a strong combinatorial bound for small k.

Lemma VIII.2. *Assume G satifies the claw bound. Suppose $k \leq n^{7/13}/\log n$. Then for some $d = O((k^{17/3}/n^{8/3}) \log^5 n)$, depth-$d$ stabilization completely splits G.*

Proof of Theorem VIII.1 from the Lemma: As in the derivation of Theorem I.1 near the end of Section I-F, we quickly reduce to the case when G satisfies Neumaier's claw bound. We then use Lemma VIII.2 (as per Prop. II.1) for the case when $k \leq n^{19/37}$, then Theorem I.6 (b′) up to $k \leq n/\log n$, and finally (2) for $k \geq n/\log n$ (see Remark I.7). The bottleneck for the automorphism bound arises at $k \sim n^{19/37}$. ■

Outline of the proof of Lemma VIII.2: The Lemma follows the high-level strategy of [7], [9], and [16]. We show that, for each pair of vertices $u, v \in V$, if d (as given in the statement of Lemma VIII.2) vertices are sampled uniformly at random and individualized, then with high probability, u and v will receive distinct colors after three steps of refinement. Lemma VIII.2 then follows by a union bound.

To this end, we show that, with high probability, there are three individualized vertices w_1, w_2 and w_3 such that the following event holds for u but does not hold for v: There exist two vertices p, q such that w_2 and w_3

are both neighbors of q; w_1 and q are both neighbors of p; p is a neighbor of u (but no such pair of vertices exists for v). It is then clear that, after three steps of refinement, u and v receive distinct colors. ■

We defer the details of the proof to [17].

IX. Open problems

The main open problem in the area is to reduce the complexity of GI to $\exp(n^{0.49})$. The target for isomorphism of s. r. graphs is subexponential ($\exp(n^{o(1)})$) or even quasipolynomial ($\exp((\log n)^{O(1)})$). This may not be entirely out of reach using the group theory method in the light of the following new result.

Theorem IX.1 ([21]). *Suppose the alternating group A_t is involved in $\mathrm{Aut}(G)$ (as a quotient of a subgroup) for a nontrivial, non-graphic s. r. graph G with n vertices. Then $t = O((\log n)^2/\log\log n)$.*

This implies that any primitive permutation group involved in $\mathrm{Aut}(G)$ has quasipolynomially bounded order, a key condition for the quasipolynomial-time efficiency of Luks's divide-and-conquer. However, the obstacles to applying Luks's method are still considerable, given that s. r. graphs lack an evident recursive structure (substructures that satisfy the same constraint on their automorphism groups).

A major open problem, related to an obstruction to the I/R method without group theory, is a subexponential bound on $|\mathrm{Aut}(G)|$ (see the bound (3) and the paragraph preceding it in Sec. I-E).

Among the open cases of isomorphism of s. r. graphs we should highlight the line-graphs of *partial geometries;* nothing better than our $\exp(\widetilde{O}(n^{1/5}))$ is known for them in spite of their geometric structure.

Acknowledgments

Partial support by NSF grants is acknowledged as follows: L. Babai CCF-1017781, X. Chen and X. Sun CCF-1149257, S-H. Teng 1111270 and 0964481, J. Wilmes DGE-1144082. X. Chen also acknowledges a Sloan Fellowship.

References

[1] O. Goldreich, S. Micali, and A. Wigderson, "Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof system," *J. ACM*, vol. 38, no. 1, pp. 691–729, 1991.

[2] L. Babai and S. Moran, "Arthur-Merlin games: a randomized proof system, and a hierarchy of complexity classes," *J. Comp. Syst. Sci.*, vol. 36, pp. 254–276, 1988.

[3] L. Babai and E. M. Luks, "Canonical labeling of graphs," In: *15th STOC*, pp. 171–183, 1983.

[4] L. Babai, W. M. Kantor, and E. M. Luks, "Computational complexity and the classification of finite simple groups," In: *24th FOCS*, pp. 162–171, 1983.

[5] V. N. Zemlyachenko, N. M. Korneenko, and R. I. Tyshkevich, "Graph isomorphism problem," *Zapiski Nauchn. Semin. LOMI*, vol. 118, pp. 83–158, 215, 1982.

[6] R. C. Read and D. G. Corneil, "The graph isomorphism disease," *J. Graph Th.*, vol. 1, no. 4, pp. 339–363, 1977.

[7] L. Babai, "On the complexity of canonical labeling of strongly regular graphs," *SIAM J. Comput.*, vol. 9, no. 1, pp. 212–216, 1980.

[8] ——, "On the order of uniprimitive permutation groups," *Annals of Math.*, vol. 113, no. 3, pp. 553–568, 1981.

[9] D. A. Spielman, "Faster isomorphism testing of strongly regular graphs," In: *28th STOC*, pp. 576–584, 1996.

[10] J. J. Seidel, "Strongly regular graphs with $(-1, 1, 0)$-adjacency matrix having eigenvalue 3," *Linear Algebra Appl.*, vol. 1, pp. 281–298, 1968.

[11] A. Neumaier, "Strongly regular graphs with smallest eigenvalue $-m$," *Arch. Math.*, vol. 33, no. 4, pp. 392–400, 1979.

[12] ——, "Quasiresidual 2-designs, $1\frac{1}{2}$-designs, and strongly regular multigraphs," *Geom. Dedicata*, vol. 12, no. 4, pp. 351–366, 1982.

[13] L. Babai and J. Wilmes, "Asymptotic Delsarte cliques in strongly regular graphs," in preparation.

[14] G. L. Miller, "On the $n^{\log n}$ isomorphism technique: A preliminary report," In: *10th STOC*, pp. 51–58, 1978.

[15] L. Babai and J. Wilmes, "Quasipolynomial-time canonical form for Steiner designs," In: *45th STOC*, pp. 261–270, 2013.

[16] X. Chen, X. Sun, and S.-H. Teng, "Multi-stage design for quasipolynomial-time isomorphism testing of Steiner 2-systems," In: *45th STOC*, pp. 271–280, 2013.

[17] ——, "A new bound on the order of the automorphism group of strongly regular graphs," in preparation.

[18] G. L. Miller, "Isomorphism of graphs which are pairwise k-separable," *Information and Control*, vol. 56, no. 1-2, pp. 21–33, 1983.

[19] B. Weisfeiler, Ed., *On construction and identification of graphs*, ser. Lecture Notes in Mathematics. Springer-Verlag, 1976, vol. 558.

[20] E. M. Luks, "Isomorphism of graphs of bounded valence can be tested in polynomial time," *J. Comp. Syst. Sci.*, vol. 25, no. 1, pp. 42–65, 1982.

[21] L. Babai, "Alternating groups involved in the automorphism groups of strongly regular graphs," 2013, in preparation.