A NOTE ON THE INEFFECTIVENESS OF THE REGULARITY LEMMA FOR BOUNDED DEGREE GRAPHS

CLARK LYONS, GRIGORY TERLOV, AND ZOLTÁN VIDNYÁNSZKY

ABSTRACT. We show that for any $\Delta \geq 3$, there is no bound computable from (ε, r) on the size of a graph required to approximate a graph of maximum degree at most Δ up to ε error in r-neighborhood statistics. This provides a negative answer to a question posed by Lovász. Our result is a direct consequence of the recent celebrated work of Bowen, Chapman, Lubotzky, and Vidick, which refutes the Aldous–Lyons conjecture.

1. Introduction and main results

Szemerédi's regularity lemma and its variants provide a foundational framework for understanding dense graphs. Roughly speaking, the lemma asserts that every graph can be approximated, in terms of homomorphism densities, by a bounded-size graph, where the size bound is explicit and depends on the chosen error parameter.

To be more precise, for finite graphs F and G define the **homomorphism density of F in G** by

$$t(F, G) = \frac{|\mathrm{Hom}(F, G)|}{|V(G)|^{|V(F)|}},$$

that is, the probability that a uniformly random function from the vertices of F to the vertices of G defines a graph homomorphism. The following is a well-known consequence of the Frieze–Kannan (weak) regularity lemma with bounds given in [FK99].

Theorem 1 ([FK99]). *There is an absolute constant C such that for any K, any $\varepsilon > 0$, and any graph G, there is a graph G' with $|V(G')| \leq C^{K^2/\varepsilon^2}$ such that*

$$|t(F, G) - t(F, G')| < \varepsilon$$

for all graphs F with $|E(F)| \leq K$.

We note that there are variants of this theorem for notions of sampling based on injective homomorphisms and graph embeddings. The theorem can also be adapted to the setting of graphs equipped with edge and vertex labelings (see [Lov12, Chapter 9] and references therein).

For sparse graphs, any homomorphism density is close to zero and hence such an approximation is not useful. Instead, one approximates them by neighborhood statistics, defined as follows. Suppose that G is a finite graph with degrees bounded by Δ. A **rooted graph of radius** $\leq r$ is a graph $F_\bullet$ with a distinguished root vertex such that all vertices of $F_\bullet$ are of distance $\leq r$ from the root. Rooted graphs will be considered up to root-preserving isomorphism $\cong_\bullet$ of graphs. For a rooted graph $F_\bullet$ of radius r and a (finite) graph G define the corresponding r-**neighborhood statistic** in G as

$$u_r(F_\bullet, G) = \frac{\left|\{v \in V(G) \mid B_r(v) \cong_\bullet F_\bullet\}\right|}{|V(G)|}, \tag{1}$$

which is equal to the probability that the radius r ball of a uniform random vertex of G is isomorphic as a rooted graph to $F_\bullet$.

Lovász has asked for an analogue of Theorem 1 in the case of bounded degree graphs, and the following observation was soon made by Alon [Alo].

Theorem 2 (Alon [Alo], as in [Zha23, Theorem 4.8.4])**.** *For a fixed $\Delta > 0$ there is a function $N_\Delta(\varepsilon, r)$ such that for every graph G with maximum degree at most Δ, there is a graph G' of maximum degree at most Δ with $|V(G')| \leq N_\Delta(\varepsilon, r)$ such that*

$$\left|u_r(F_\bullet, G) - u_r(F_\bullet, G')\right| < \varepsilon$$

for all rooted graphs $F_\bullet$ of radius r.

This statement follows from the totally boundedness of the space of possible r-neighborhood statistics for graphs (see Section 2 for the proof). Thus, unlike the case of dense graphs, the function $N_\Delta(\varepsilon, r)$ is not constructive. It was then asked by Lovász [Lov12, page 358] (see also [Zha23, Open Problem 4.8.5]) whether it is possible to find an "effective" bound for $N_\Delta(\varepsilon, r)$. He also suggested that an efficient, algorithmic way of constructing an approximating graph could be used also on graphings, and hence to find a finitary approximation, in turn, resolving positively the famous Aldous–Lyons conjecture [AL07].

In a recent breakthrough, Bowen, Chapman, Lubotzky, and Vidick [BCLV24, BCV24] gave a negative solution to the latter conjecture. In this note we show that their main undecidability result directly implies a negative answer to Lovász's question:

Theorem 3 (Regularity is ineffective for bounded degree graphs)**.** *For any $\Delta \geq 3$, any regularity bound $N_\Delta(\varepsilon, r)$ is not computable.*

Bowen, Chapman, Lubotzky, and Vidick used so-called subgroup tests to encode non-local games and then build on the ideas from the earlier deep results of Ji, Natarajan, Vidick, Wright, and Yuen [JNV+22] to deduce undecidability. We note that although the subgroup test approach might not be necessary for their original goal [Man25] it is indeed essential for our argument. Let us also mention that solely a counterexample to the Aldous–Lyons conjecture in itself would have been likely insufficient to show Theorem 3, but, the insight provided by the undecidability results suffices for this purpose.

Furthermore, there seems to be an unnoticed implication in the converse the direction. Namely, we provide an easy argument showing that a positive answer to the Aldous–Lyons conjecture would have implied a computable bound on $N_\Delta(\varepsilon, r)$. This observation may still be of interest, as one could attempt to directly show that such a bound does not exist, thereby providing an alternative proof of the failure of the conjecture.

Organization. In Section 2 we state the undecidability result (as a combination of the statements from [BCLV24, BCV24]) and we prove the analogous result to Theorem 3 stated for Schreier graphs (Theorem 6), which is equivalent to to Theorem 3 by Lemma 7. Section 3 is dedicated to the upper approximation of r-neighborhood statistics of graphings, and showing the implication mentioned above. In Section 4 we state several open problems. Finally, in Appendix A we present standard encodings between graphs and Schreier graphs and prove Lemma 7.

2. Proof of Theorem 3

Both sampling regularity for dense graphs and neighborhood statistic regularity are well known to extend to decorated graphs. We will work with graphs equipped with a Schreier decoration of an action of a free group on finitely many generators.

Definition 4. For a natural number d, an $\mathbb{F}_d$ **Schreier graph** G consists of a vertex set $V(G)$, a symmetric (but not necessarily reflexive) edge set $E(G) \subseteq V(G)^2$, and an edge labeling $c_G : E(G) \to \mathcal{P}(\{a_1, \dots, a_d, a_1^{-1}, \dots, a_d^{-1}\})$ such that:

- we have $a_i \in c_G(x,y) \iff a_i^{-1} \in c_G(y,x)$
- for any $x \in V(G)$ and any $\ell \in \{a_1, \ldots, a_d, a_1^{-1}, \ldots, a_d^{-1}\}$ there is a unique $y \in V(G)$ such that $(x,y) \in E(G)$ and $\ell \in c_G(x,y)$.

The edges and edge labels specify the same information as an action of the free group on n generators $a_G : \mathbb{F}_d \curvearrowright V(G)$.

For any $\mathbb{F}_d$ Schreier graph G, around any vertex $v \in V(G)$ we can consider a ball $B_r(G)$ of radius r as a rooted edge-labeled (non-simple) graph. We consider these graphs up to root-preserving and label-preserving graph isomorphism $\cong^*_\bullet$. This allows us to formulate neighborhood statistics for $\mathbb{F}_d$ Schreier graphs exactly as before but with this new notion of isomorphism. Just as in the undecorated case, there is a regularity bound obtained by a compactness (or, more precisely, totally boundedness) argument, see Section 3.

Definition 5. For a rooted edge-labeled graph $F_\bullet$ of radius r and an $\mathbb{F}_d$ Schreier graph G, define the **Schreier neighborhood statistic**

$$u_r^*(F_\bullet, G) = \frac{\left|\{v \in V(G) \mid B_r(v) \cong^*_\bullet F_\bullet\}\right|}{|V(G)|}.$$

An $\mathbb{F}_d$ **Schreier regularity bound** is a function $N_d^*(\varepsilon, r)$ such that for any $\mathbb{F}_d$ Schreier graph G there is an $\mathbb{F}_d$ Schreier graph G' with $|V(G')| \le N^*(\varepsilon, r)$ such that

$$\left|u_r^*(F_\bullet, G) - u_r^*(F_\bullet, G')\right| < \varepsilon$$

for all rooted edge-labeled graphs $F_\bullet$ of radius r.

We are now ready state a version of Theorem 3 for Schreier graphs.

Theorem 6 (Regularity is Ineffective in Schreier Graphs)**.** *For $d \ge 2$ any Schreier regularity bound $N_d^*(\varepsilon, r)$ is not a computable function.*

This version is equivalent to the original theorem for graphs by the following lemma. The proof of this equivalence relies on a well known fact that finite graphs can be encoded by Schreier graphs and visa versa, which we present it in Appendix A in full detail.

Lemma 7. *Let $\Delta \ge 3$ and $d \ge 2$. A sparse regularity bound $N_\Delta(\varepsilon, r)$ can be computed from any Schreier regularity bound $N_d^*(\varepsilon, r)$ as an oracle and vice versa.*

The powerset $\mathcal{P}(\mathbb{F}_d)$ carries compact Hausdorff topology from viewing it as a product of discrete spaces $2^{\mathbb{F}_d}$. Let $\mathscr{T}_d$ denote the set of all continuous rational-valued functions $\mathcal{T}$ on $\mathcal{P}(\mathbb{F}_d)$. By continuity, for any such function $\mathcal{T}$ there is a finite set $K \subseteq \mathbb{F}_d$ such that for any $S \subseteq \mathbb{F}_d$, then output $\mathcal{T}(S)$ is determined by $S \cap K$.

The **value** of a continuous rational function $\mathcal{T}$ with respect to an $\mathbb{F}_d$ Schreier graph is defined by

$$\mathrm{val}(\mathcal{T}, G) = \mathbb{E}_{v \in V(G)}\left[\mathcal{T}(\mathrm{Stab}(v))\right],$$

where $\mathrm{Stab}(v)$ is the stabilizer of v in the induced action of $\mathbb{F}_d$ on $V(G)$ and the expectation is over a uniform random choice of vertex $v \in V(G)$. The **sofic value** of a continuous rational function is the supremum of its values with respect to all (finite) $\mathbb{F}_d$ Schreier graphs

$$\mathrm{val}_{sof}(\mathcal{T}) = \sup_G \mathrm{val}(\mathcal{T}, G).$$

Then for any n the set of rational lower bounds for sofic values

$$\{(\mathcal{T}, q) \in \mathscr{T}_d \times \mathbb{Q} \mid \mathrm{val}_{sof}(\mathcal{T}) > q\}$$

is a computably enumerable set, through a brute-force search of all $\mathbb{F}_d$ Schreier graphs.

The following theorem summarizes a key step of the argument of Bowen, Chapman, Lubotzky, and Vidick in their resolution of the Aldous–Lyons conjecture. See [BCLV24, Theorem 1.13 and 7.3] and [BCV24, Theorem 1.1] for the general result.

Theorem 8 ([BCLV24, BCV24])**.** *For any $d \geq 2$ there is a computable map which takes as input a Turing machine M and outputs a continuous rational function $\mathcal{T}^M \in \mathscr{T}_d$ and rational number $\lambda^M > 0$ such that*

- *if M halts then $val_{sof}(\mathcal{T}^M) = 1$,*
- *if M does not halt then $val_{sof}(\mathcal{T}^M) \leq 1 - \lambda^M$.*

We briefly note how the result of Bowen, Chapman, Lubotzky, and Vidick is stronger than what is quoted above. For them, the continuous rational function $\mathcal{T}^M$ is restricted to be a subgroup test, which is a certain kind of continuous rational function with outputs in $[0,1]$ which mimics an interactive proof system between "prover" and "verifier" entities. Additionally, in the case that $\mathrm{val}_{sof}(\mathcal{T}^M) = 1$, there is a (finite) $\mathbb{F}_d$ Schreier graph G which achieves this supremum so $\mathrm{val}_{sof}(\mathcal{T}^M, G) = 1$. The proof of Theorem 6 only uses that the sofic values of continuous rational functions cannot be computably approximated from above uniformly.

Proof of Theorem 6. We show how to computably enumerate rational upper bounds for sofic values of a continuous rational function from a Schreier regularity bound $N_d^*(\varepsilon, r)$. Consider a continuous rational function $\mathcal{T}$ whose output is determined by intersection with the finite subset $K \subseteq \mathbb{F}_d$. Let r be a radius large enough that in the Cayley graph of $\mathbb{F}_d$, the ball of radius r around the identity contains K. Note that for a vertex v in an $\mathbb{F}_d$ Schreier graph G, the value $\mathcal{T}(\mathrm{Stab}(v))$ depends only on the edge-labeled rooted subgraph $F_\bullet$ induced by $B_r(v)$, considered up to a root and label preserving isomorphism. We denote this value $\mathcal{T}(\mathrm{Stab}(F_\bullet))$. Let $\mathcal{F}_d^*(r)$ be the finite set that consists of all rooted graphs of radius $\leq r$ with edges labeled by the standard generating set of $\mathbb{F}_d$, considered up to a rooted and label-preserving isomorphism. Let $m(\mathcal{T}) = \max_{F_\bullet \in \mathcal{F}_d^*(r)} \mathcal{T}(\mathrm{Stab}(F_\bullet))$. For any $\Theta > 0$ if

$$\left|u_r^*(F_\bullet, G) - u_r^*(F_\bullet, G')\right| < \varepsilon = \frac{\Theta}{|\mathcal{F}_d^*(r)| \cdot m(\mathcal{T})}$$

for all $F_\bullet \in \mathcal{F}_d^*(r)$, then

$$\left|\mathrm{val}(\mathcal{T}, G) - \mathrm{val}(\mathcal{T}, G')\right| \leq \sum_{F_\bullet \in \mathcal{F}_d^*(r)} |u_r^*(F_\bullet, G) - u_r^*(F_\bullet, G')| \cdot \mathcal{T}(\mathrm{Stab}(F_\bullet)) < \Theta.$$

Thus for each $\Theta > 0$ we can compute all $\mathbb{F}_d$ Schreier graphs with at most $N_d^*(\varepsilon, r)$ vertices for this value of ε, and we can compute the maximum value

$$\beta_{\mathcal{T},\Theta} = \max_{|V(G')| \leq N_d^*(\varepsilon, r)} \mathrm{val}(\mathcal{T}, G')$$

among these graphs. Then by the Schreier regularity bound we have

$$\beta_{\mathcal{T},\Theta} \leq \mathrm{val}_{sof}(\mathcal{T}) \leq \beta_{\mathcal{T},\Theta} + \Theta. \tag{2}$$

But this contradicts Theorem 8: indeed, given a Turing-machine M, we could use Theorem 8 to calculate $\mathcal{T}^M$ and λ^M. Then, using $\mathcal{T}^M$ we could find an r as above and calculate $\beta_{\mathcal{T},\Theta}$ for $\Theta = \lambda^M/2$. Finally, (2) would guarantee that M halts iff $\beta_{\mathcal{T},\Theta} > 1 - \lambda^M$.

□

3. Local statistics of graphings

Let $\mathcal{F}_\Delta(r)$ denote the (finite) set of rooted graphs of maximal degree $\leq \Delta$ and radius $\leq r$ around the root, considered up to a rooted isomorphism. Similarly, let $\mathcal{G}_\Delta$ denote the (infinite) set of finite graphs of maximum degree $\leq \Delta$, also considered up to an isomorphism. There is a natural map

$U_r : \mathcal{G}_\Delta \to [0,1]^{\mathcal{F}_\Delta(r)}$ that encodes r-neighborhood statistics of $G \in \mathcal{G}_\Delta$ in terms of all possible neighborhood statistics $\{u_r(F_\bullet, G)\}_{F_\bullet \in \mathcal{F}_\Delta(r)}$, namely

$$U_r(G) := \{F_\bullet \mapsto u_r(F_\bullet, G)\}.$$

Since $[0,1]^{\mathcal{F}_\Delta(r)}$ is compact in the metric d_∞, the image $U_r(\mathcal{G}_\Delta)$ is totally bounded.

So for any $\varepsilon > 0$ there is a finite subset $\mathcal{N} \subseteq \mathcal{G}_\Delta$ such that for all $G \in \mathcal{G}_\Delta$ there exists $G' \in \mathcal{N}$ such that $d_\infty(U_r(G), U_r(G')) < \varepsilon$ which implies

$$\left|u_r(F_\bullet, G) - u_r(F_\bullet, G')\right| < \varepsilon$$

for all rooted graphs $F_\bullet$ of radius r.

Remark 9. In fact, taking $N_\Delta(\varepsilon, r)$ to be the maximum number of vertices of a graph in $\mathcal{N}$ yields a proof of Theorem 2 as presented in [Zha23, Theorem 4.8.4] (with the only difference that it is stated in terms of ℓ^∞ distance instead of the total variation distance, but since $\mathcal{F}_\Delta$ is finite, these distances are equivalent).

The same argument proves the version of Theorem 2 for $\mathbb{F}_d$ Schreier graphs and the map $U^*_{d,r} : \mathcal{G}^*_d \to [0,1]^{\mathcal{F}^*_d(r)}$ from finite $\mathbb{F}_d$ Schreier graphs to r-neighborhood statistics.

This map extends to the set of Schreier graphings obtained by measure preserving actions of $\mathbb{F}_d$ on a standard probability space. It is easy to see that the value only depends on the invariant random subgroup obtained as the stabilizer of a random vertex for such an action.

Consider the space of probability measures on the power set of $\mathbb{F}_d$, denoted as $\mathrm{Prob}(\mathcal{P}(\mathbb{F}_d))$. An **invariant random subgroup** of $\mathbb{F}_d$ is a Borel probability measure $\mu \in \mathrm{Prob}(\mathcal{P}(\mathbb{F}_d))$ supported on subgroups of $\mathbb{F}_d$ which is invariant under conjugation. We denote the set of invariant random subgroups of $\mathbb{F}_d$ by IRS_d and define the corresponding r-neighborhood statistic as $U^*_{d,r} : \mathrm{IRS}_d \to [0,1]^{\mathcal{F}^*_d(r)}$ as

$$U^*_{d,r}(\theta) := \{F_\bullet \mapsto u^*_{d,r}(F_\bullet, \theta)\},$$

where $u^*_{d,r}(F_\bullet, H)$ is the probability that the random subgroup θ encodes $F_\bullet$, in the sense that the radius r ball around a vertex with stabilizer θ is isomorphic to $F_\bullet$ as a rooted edge-labeled graph. Notice that the statement of the Aldous–Lyons conjecture is equivalent to $U^*_{d,r}(\mathcal{G}_d)$ being dense in $U^*_{d,r}(\mathrm{IRS}_d)$.

Given $k > 0$ let $W_d(k)$ be the ball of radius k around the identity in the Cayley graph of $\mathbb{F}_d$ with respect to the standard generating set. A k-**pseudo-subgroup** of $\mathbb{F}_d$ is a subset of $W_d(k)$ which is the intersection of $W_d(k)$ with a subgroup of $\mathbb{F}_d$. Then a k-**pseudo-IRS** is defined as a Borel probability measure on $\mathcal{P}(W_d(k)) \subseteq \mathcal{P}(\mathbb{F}_d)$ which is invariant under conjugation by the standard generators (when this conjugation is defined). Similarly to the above, we denote the set of k-pseudo-IRSs of $\mathbb{F}_d$ by $\text{P-IRS}_d(k)$. In fact, it follows that

$$\mathrm{IRS}_d = \bigcap_k \text{P-IRS}_d(k)$$

as a decreasing intersection.

It is not hard to see that Theorem 3 implies that one cannot algorithmically confirm whether a given open set covers all statistics of IRSs coming from an action on a finite space (i.e., there is no Turing machine which halts exactly when the containment holds, see also [BCLV24]). By contrast, using pseudo-subgroups, it can be algorithmically confirmed whether such an open set covers all the IRSs.

Theorem 10. *There is a Turing machine $M(d, r, S)$ which takes as input a number of generators d, radius r, and $S \subseteq [0,1]^{\mathcal{F}^*_d(r)}$ a union of finitely many open d_∞-neighborhoods (centered at rational points) such that $M(d, r, S)$ halts if and only if $U^*_{d,r}(IRS_d) \subseteq S$.*

Proof. First we note that the value $U^*_{d,r}(\theta)$ is determined by its restriction to $W_d(2r)$ (which is a pseudo-IRS). In particular we can extend the map to $U^*_{d,r} : \text{P-IRS}_d(k) \to [0,1]^{\mathcal{F}^*_d(r)}$ for any $k \geq 2r$. We have

$$U^*_{d,r}(\text{IRS}_d) = \bigcap_{k \geq 2r} U^*_{d,r}(\text{P-IRS}_d(k)).$$

Since the map $U^*_{d,r}$ is continuous and $U^*_{d,r}(\text{IRS}_d)$ is a closed set, by compactness of $U^*_{d,r}(\text{IRS}_d)$, if $U^*_{d,r}(\text{IRS}_d) \subseteq S$ then there is some $k \geq 2r$ such that $U^*_{d,r}(\text{P-IRS}_d(k)) \subseteq S$.

Since the linear inequalities defining $U^*_{d,r}(\text{P-IRS}_d(k))$ are computable uniformly in d and k [BCLV24, Lemma 2.16], and since checking the consistency of finite families of rational inequalities is computable, the condition $U^*_{d,r}(\text{P-IRS}_d(k)) \subseteq S$ is also computable. Thus, the Turing machine $M(d,r,S)$ can search through all $k \geq 2r$ and halt when it determines $U^*_{d,r}(\text{P-IRS}_d(k)) \subseteq S$. □

Remark 11. It follows that if for all d and r the $U^*_{d,r}(\mathcal{G}_d)$ were dense in $U^*_{d,r}(\text{IRS}_d)$ (i.e. the Aldous–Lyons conjecture holds) then, by Theorem 10, there would have been a computable upper bound to $N^*_d(\varepsilon, r)$. Indeed, a brute force search through finite $\mathbb{F}_d$ Schreier graphs would eventually enumerate a finite collection whose r-neighborhood statistics form an ε-cover of $U^*_{d,r}(\mathcal{G}_d)$, equivalently of $U^*_{d,r}(\text{IRS}_d)$, and will detect this in finite time.

4. Open Problems

We now present several related open questions. First, the present note is concerned with the bounded degree graphs, and, as mentioned in the introduction, the analogous statement is not true in the other end of the spectrum, i.e., for dense graphs. It is possible to build a continuous spectrum connecting bounded degree graphs to dense graphs, parametrized by a real, which measures their density, and to ask appropriate versions of approximation theorems (see, e.g., [BS22, Fre18] for different variants). This naturally leads to the following question.

Question 12. Can one describe the phase transition from approximable to inapproximable, as the density decreases?

It would be extremely interesting for example, if, as the density decreases, the bound in Theorem 1 would increase, eventually yielding a non-computable growth rate.

Another natural direction could be the following. Let $\mathcal{A}_r \subseteq \mathbb{R}^{\mathcal{F}_\Delta(r)}$ be the set of all probability distributions of r-neighborhoods around the root in G, where G ranges over all finite uniformly-rooted graphs. Similarly, let $\mathcal{A}'_r \subseteq \mathbb{R}^{\mathcal{F}_\Delta(r)}$ be the same, but where G ranges through all graphings. Clearly, the closure of $\mathcal{A}_r$, denoted by $\overline{\mathcal{A}_r}$, is a contained in $\mathcal{A}'_r$. The Aldous–Lyons conjecture can be stated as $\overline{\mathcal{A}_r} = \mathcal{A}'_r$ for every r. The following problems were mentioned by Lovász [Lov12, page 358] as potential ways of refuting the Aldous–Lyons conjecture. In light of Theorem 3 we still find them interesting.

Open Problem 13. Find an explicit $r > 0$ such that $\overline{\mathcal{A}_r} \neq \mathcal{A}'_r$.

One can also restrict the attention to some finite set of rooted graphs $F_1, F_2, \ldots, F_m \in \mathcal{F}_\Delta(r)$ and to each finite graph G one can associate a vector $(u_r(F_1, G), \ldots, u_r(F_m, G)) \subset \mathbb{R}^m$. Let $U_r(F_1, F_2, \ldots, F_m)$ denote the set of all such vectors as G ranges over all finite uniformly-rooted graphs. Again, let $U'_r(F_1, F_2, \ldots, F_m)$ be the same, but where G ranges through all graphings.

Open Problem 14. Given $r > 0$ and finite graphs $\{F_i\}_{i=1}^m$, determine the sets $\overline{U}_r(F_1, F_2, \ldots, F_m)$ and $U'_r(F_1, F_2, \ldots, F_m)$.

Harangi [Har13] considered this question with $m = 2$, and F_1, F_2 are cycles of length 3 and 4. For any given r his result gives an explicit description of the corresponding sets and yields that they do coincide. We also would like to point out another question of Abért stated in [Har13, Question

1.3] about the approximation of the expected spectral measures of graphings by the eigenvalue distribution of finite graphs.

Appendix A. Encoding graphs and Schreier graphs

The goal of this appendix is to prove the equivalence of Theorem 3 and Theorem 6 by encoding free group Schreier graphs with bounded degree graphs and vice versa. It is a well known fact that these objects encode each other (see e.g. [Gro77, AL07, Gri11, Can14, Tót21] and references therein) and is often mentioned without further explanation. Although the particular encoding presented below is also known to the specialists we were not able to find where it is written explicitly.

First, there is a straightforward way of encoding any locally finite graph as an $\mathbb{F}_2$ Schreier graph. For a (simple, connected) graph $G = (V, E)$, we build an $\mathbb{F}_2$ Schreier graph G^* with vertex set equal to the set of directed edges $V(G^*) = E(G)$. For each $v \in V(G)$ fix a cyclic ordering of the edges with source v. This defines an action on $a : \mathbb{Z} \curvearrowright V(G^*)$. There is also an action $b : \mathbb{Z} \curvearrowright V(G^*)$ which exchanges the source and target of the directed edge. Together these define an $\mathbb{F}_2$ action.

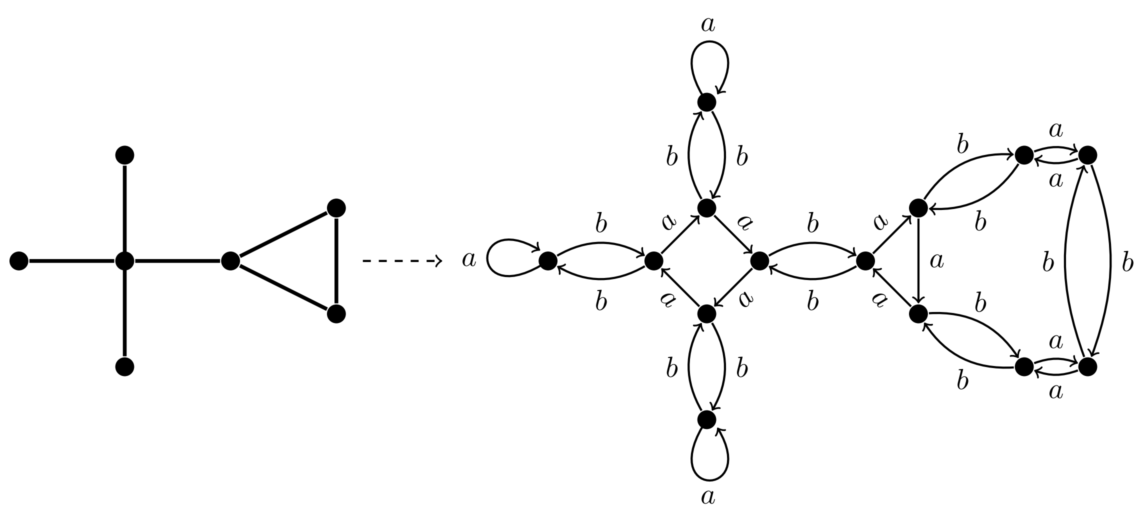

Figure 1. Example of encoding a graph as an $\mathbb{F}_2$ Schreier graph.

For the converse direction let G be an $\mathbb{F}_d$ Schreier graph. We encode this into a simple graph $\hat{G}$ of maximum degree 3. For each vertex v of G and each $\ell \in \{a_1, \ldots, a_d, a_1^{-1}, \ldots, a_d^{-1}\}$ we have a vertex $v_\ell \in \hat{G}$ and these vertices are arranged in a cycle. Additionally, for each $i \in \{1, \ldots, d\}$ and edge (x, y) labeled by a_i, there are vertices and edges forming a gadget that connects x_{a_i} and $y_{a_i^{-1}}$ as in Figure 2. Therefore the only cycles of length $2d$ in $\hat{G}$ are are those corresponding to vertices of G, and each vertex of such a cycle is the start or end of a distinct gadget, one corresponding to each $\ell \in \{a_1, \ldots, a_d, a_1^{-1}, \ldots, a_d^{-1}\}$.

Both of these encodings are defined uniquely up to the choice of cyclic ordering at each vertex. Additionally, both encodings are injective maps and therefore do not lose information. It will be important in the following proof that both of these encodings are defined locally, and that they can both be decoded with small errors in the space of neighborhood statistics.

Proof of Lemma 7. First fix $\Delta \geq 3$ and suppose we have oracle access to a Schreier regularity bound $N_2^*(\varepsilon, r)$. We show how to compute a sparse regularity bound $N_\Delta(\varepsilon, r)$.

First let G be graph of maximum degree at most Δ, and let G^* be its encoding as a Schreier graph as defined above. We know that for any ε_0 and r_0 there is an $\mathbb{F}_2$ Schreier graph $G^{*\prime}$ with

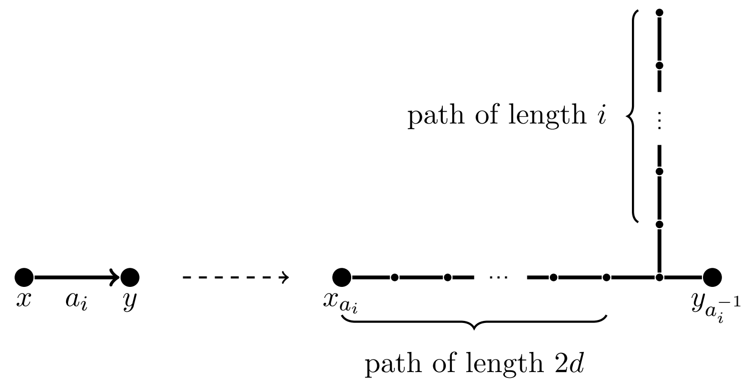

FIGURE 2. Example of encoding of a directed edge

$|V(G^{*\prime})| \le N_2^*(\varepsilon_0, r_0)$ such that

$$|u_{r_0}(F_\bullet, G^*) - u_{r_0}(F_\bullet, G^{*\prime})| < \varepsilon_0$$

for any rooted edge-labeled graph $F_\bullet \in \mathcal{F}_2^*(r_0)$. The graph $G^{*\prime}$ may not be in the image of the encoding map, but we can still decode some simple graph of degree at most Δ from the incidence of a-cycles and b-cycles in the action of $\mathbb{F}_2$ on $V(G^{*\prime})$.

We can say that a cycle for the action of a on $V(G^{*\prime})$ is 'good' if it is of length at most Δ, and if each vertex of the a-cycle is part of a b-cycle of length 2 with a distinct vertex outside of the a-cycle. We then form a graph G' of degree at most Δ with 'good' cycles as vertices and edges between 'good' cycles if there is a b-cycle of length 2 between them. By choosing ε_0 small enough and r_0 large enough we can ensure that at most ε fraction of the vertices of $G^{*\prime}$ are not part of good cycles, since every vertex of G^* is part of a good cycle and this is a locally property (i.e. determined by a neighborhood with uniformly bounded radius).

In fact, by choosing ε_0 small and r_0 large enough, we can ensure

$$|u_r(F_\bullet, G) - u_r(F_\bullet, G')| < \varepsilon$$

for any rooted graph $F_\bullet \in \mathcal{F}_\Delta(r)$. Therefore we can let $N_\Delta(\varepsilon, r)$ be the maximum number of vertices in one of these decoded graphs G'.

Conversely fix $d \ge 2$ and suppose we have oracle access to a sparse regularity bound $N_3(\varepsilon, r)$. We show how to compute a Schreier regularity bound $N_d^*(\varepsilon, r)$.

Let G be an $\mathbb{F}_d$ Schreier graph, and let $\hat{G}$ be its encoding as a simple rooted graph of degree at most 3. For any ε_0 and r_0 there is a graph $\hat{G}'$ of degree at most 3 with $|V(\hat{G}')| \le N_3(\varepsilon_0, r_0)$ such that

$$|u_{r_0}(F_\bullet, \hat{G}) - u_{r_0}(F_\bullet, \hat{G}')| < \varepsilon_0$$

for any rooted graph $F_\bullet$ of radius r_0. As before, the graph $\hat{G}'$ may not be in the image of the encoding map, but we can decode some $\mathbb{F}_d$ Schreier graph from the small cycles and the gadgets connecting them.

We say that a cycle in $\hat{G}'$ is 'good' if it has length $2d$ and each vertex is the start or end of a distinct gadget, one corresponding to each $\ell \in \{a_1, \dots, a_d, a_1^{-1}, \dots, a_d^{-1}\}$, as in the encoding construction. We can form an $\mathbb{F}_d$ Schreier graph G' with 'good' cycles as vertices and edges between 'good' cycles if there is a gadget between them. The symbol a_i (or a_i^{-1}) will be part of the label of an edge if that edge (or its reverse) came from the corresponding gadget. So far, this only defines a partial $\mathbb{F}_d$ Schreier graph. For each vertex v and $\ell \in \{a_1, \dots, a_d, a_1^{-1}, \dots, a_d^{-1}\}$ there is at most one edge out of v with ℓ as part of its label. But for each $i \in \{1, \dots, n\}$ we have an equal number of vertices missing a_i as part of an edge coming out as those missing a_i as part of an edge coming in. This can be seen by analyzing the components of the subgraph of edges labeled a_i or a_i^{-1}, which has degree at most 2 and therefore consists of cycles and paths. Therefore we can add an additional edge from each vertex with a_i as a missing out-label to a vertex with a_i as a missing in-label. This new edge will have a_i as part of its label (and a_i^{-1} in the other direction). This ensures G' is an $\mathbb{F}_d$ Schreier graph. By choosing ε_0 small enough and r_0 large enough we can ensure that at most ε fraction of the vertices of $\hat{G}'$ are not part of good cycles or the gadgets connected to them, since this is true of all vertices in G^* and this property is determined locally. Finally, as before, by choosing ε_0 small and r_0 large enough, we can ensure

$$|u_r^*(F_\bullet, G) - u_r^*(F_\bullet, G')| < \varepsilon$$

for any labeled rooted graph $F_\bullet$ of radius r. Therefore we can let $N_\Delta^*(\varepsilon, r)$ be the maximum number of vertices in one of these decoded graphs G'. □

Acknowledgements

We would like to thank Michael Chapman, László Lovász, and all the participants of the PSQC reading seminar at Renyi for many insightful conversations. The second author was supported by the ERC Synergy Grant No. 810115 - DYNASNET. The 1st and 3rd authors were partially supported by HAS Momentum Grant 2022-58. The 3rd author was also supported by National Research, Development and Innovation Office (NKFIH) grant no. 146922.

References

[AL07] David Aldous and Russell Lyons. Processes on unimodular random networks. *Electronic Journal of Probability*, 12:1454 – 1508, 2007.

[Alo] Noga Alon. *(unpublished)*.

[BCLV24] Lewis Bowen, Michael Chapman, Alexander Lubotzky, and Thomas Vidick. The aldous–lyons conjecture i: Subgroup tests, 2024. https://arxiv.org/abs/2408.00110.

[BCV24] Lewis Bowen, Michael Chapman, and Thomas Vidick. The aldous–lyons conjecture ii: Undecidability, 2024. https://arxiv.org/abs/2501.00173.

[BS22] Ágnes Backhausz and Balázs Szegedy. Action convergence of operators and graphs. *Canadian Journal of Mathematics*, 74(1):72–121, 2022.

[Can14] Jan Cannizzo. On invariant Schreier structures. *Enseign. Math.*, 60(3-4):397–415, 2014.

[FK99] Alan Frieze and Ravindran Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19:175–220, 02 1999.

[Fre18] Péter Frenkel. Convergence of graphs with intermediate density. *Transactions of the American Mathematical Society*, 370(5):3363–3404, 2018.

[Gri11] Rostislav I Grigorchuk. Some topics in the dynamics of group actions on rooted trees. *Proceedings of the Steklov Institute of Mathematics*, 273(1):64–175, 2011.

[Gro77] Jonathan L. Gross. Every connected regular graph of even degree is a Schreier coset graph. *J. Combinatorial Theory Ser. B*, 22(3):227–232, 1977.

[Har13] Viktor Harangi. On the density of triangles and squares in regular finite and unimodular random graphs. *Combinatorica*, 33(5):531–548, 2013.

[JNV+22] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Mip*=re, 2022. https://arxiv.org/abs/2001.04383.

[Lov12] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[Man25] Aareyan Manzoor. There is an equivalence relation whose von neumann algebra is not connes embeddable, 2025. https://arxiv.org/abs/2502.06697.

[Tót21] László Márton Tóth. Invariant Schreier decorations of unimodular random networks. *Ann. H. Lebesgue*, 4:1705–1726, 2021.

[Zha23] Yufei Zhao. *Graph Theory and Additive Combinatorics: Exploring Structure and Randomness*. Cambridge University Press, 2023.

Clark Lyons, HUN-REN Alfréd Rényi Institute of Mathematics, Reáltanoda u. 13-15, Budapest 1053, Hungary.
Email address: clyons@renyi.hu

Grigory Terlov, Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC, USA, and HUN-REN Alfréd Rényi Institute of Mathematics, Reáltanoda u. 13-15, Budapest 1053, Hungary.
Email address: gterlov@unc.edu

Zoltán Vidnyánszky, Eötvös Loránd University, Institute of Mathematics, Pázmány Péter stny. 1/C, 1117 Budapest, Hungary
Email address: zoltan.vidnyanszky@ttk.elte.hu