Electronic Colloquium on Computational Complexity, Report No. 81 (2011)

Erdös-Rényi Sequences and Deterministic construction of Expanding Cayley Graphs

V. Arvind [*]   Partha Mukhopadhyay[†]   Prajakta Nimbhorkar [†]

May 15, 2011

Abstract

Given a finite group G by its multiplication table as input, we give a deterministic polynomial-time construction of a directed Cayley graph on G with $O(\log|G|)$ generators, which has a rapid mixing property and a constant spectral expansion.

We prove a similar result in the undirected case, and give a new deterministic polynomial-time construction of an expanding Cayley graph with $O(\log|G|)$ generators, for any group G given by its multiplication table. This gives a completely different and elementary proof of a result of Wigderson and Xiao [10].

For any finite group G given by a multiplication table, we give a deterministic polynomial-time construction of a cube generating sequence that gives a distribution on G which is arbitrarily close to the uniform distribution. This derandomizes the well-known construction of Erdös-Rényi sequences [2].

1 Introduction

Let G be a finite group with n elements, and let $J = \langle g_1, g_2, \ldots, g_k\rangle$ be a *generating* set for the group G.

The *directed Cayley graph* $\mathrm{Cay}(G, J)$ is a directed graph with vertex set G with directed edges of the form (x, xg_i) for each $x \in G$ and $g_i \in J$. Clearly, since J is a generating set for G, $\mathrm{Cay}(G, J)$ is a strongly connected graph with every vertex of out-degree k.

The *undirected Cayley graph* $\mathrm{Cay}(G, J \cup J^{-1})$ is an undirected graph on the vertex set G with undirected edges of the form $\{x, xg_i\}$ for each $x \in G$ and $g_i \in J$. Again, since J is a generating set for G, $\mathrm{Cay}(G, J \cup J^{-1})$ is a connected regular graph of degree $|J \cup J^{-1}|$.

Let $X = (V, E)$ be an undirected regular n-vertex graph of degree D. Consider the *normalized* adjacency matrix A_X of the graph X. It is a symmetric matrix with largest eigenvalue 1. For $0 < \lambda < 1$, the graph X is an (n, D, λ)-*spectral expander* if the second largest eigenvalue of A_X, in absolute value, is bounded by λ.

The study of expander graphs and its properties is of fundamental importance in theoretical computer science; the Hoory-Linial-Wigderson monograph is an excellent source [4] for current

[*] The Institute of Mathematical Sciences, Chennai, India.Email: arvind@imsc.res.in

[†] Chennai Mathematical Institute, Siruseri, India. Emails: {partham,prajakta}@cmi.ac.in

ISSN 1433-8092

developments and applications. A central problem is the explicit construction of expander graph families [4, 5]. By explicit it is meant that the family of graphs has efficient deterministic constructions, where the notion of efficiency is often tailored to a specific application, e.g. [9]. Explicit constructions with the best known (and near optimal) expansion and degree parameters are Cayley expander families (the so-called Ramanujan graphs) [5].

Does every finite group have an expanding generator set? Alon and Roichman, in [1], answered this in the positive using the probabilistic method. Let G be any finite group with n elements. Given any constant $\lambda > 0$, they showed that a random multiset J of size $O(\log n)$ picked uniformly at random from G is, with high probability, a spectral expander with second largest eigenvalue bounded by λ. In other words, $\mathrm{Cay}(G, J \cup J^{-1})$ is an $O(\log n)$ degree, λ-spectral expander with high probability. The theorem also gives a polynomial (in n) time randomized algorithm for construction of a Cayley expander on G: pick the elements of J independently and uniformly at random and check that $\mathrm{Cay}(G, J \cup J^{-1})$ is a spectral expander. There is a brute-force deterministic simulation of this that runs in $n^{O(\log n)}$ time by cycling through all candidate sets J. Wigderson and Xiao in [10], give a very interesting $n^{O(1)}$ time derandomized construction based on Chernoff bounds for matrix-valued random variables (and pessimistic estimators). Their result is the starting point of the study presented in this paper.

In this paper, we give an entirely different and elementary $n^{O(1)}$ time derandomized construction that is based on analyzing mixing times of random walks on expanders rather than on its spectral properties. Our construction is conceptually somewhat simpler and also works for directed Cayley graphs.

The connection between mixing times of random walks on a graph and its spectral expansion is well studied. For undirected graphs we have the following.

Theorem 1.1 [8, Theorem 1] *Let A be the normalized adjacency matrix of an undirected graph. For every initial distribution, suppose the distribution obtained after t steps of the random walk following A is ϵ-close to the uniform distribution in the L_1 norm. Then the spectral gap $(1 - |\lambda_1|)$ of A is $\Omega(\frac{1}{t}\log\left(\frac{1}{\epsilon}\right))$.*

In particular, if the graph is $\mathrm{Cay}(G, J \cup J^{-1})$ for any n element group G, such that a $C \log n$ step random walk is $\frac{1}{n^c}$-close to the uniform distribution in L_1 norm, then the spectral gap is a constant $\frac{c}{C}$.

Even for directed graphs a connection between mixing times of random walks and the spectral properties of the underlying Markov chain is known.

Theorem 1.2 [6, Theorem 5.9] *Let λ_{max} denote the second largest magnitude (complex valued) eigenvalue of the normalized adjacency matrix P of a strongly connected aperiodic Markov Chain. Then the mixing time is lower bounded by $\tau(\epsilon) \geq \frac{\log(1/2\epsilon)}{\log(1/|\lambda_{max}|)}$, where ϵ is the difference between the resulting distribution and the uniform distribution in the L_1 norm.*

In [7], Pak uses this connection to prove an analogue of the Alon-Roichman theorem for directed Cayley graphs: Let G be an n element group and $J = \langle g_1, \ldots, g_k \rangle$ consist of $k = O(\log n)$ group elements picked independently and uniformly at random from G. Pak shows that for any initial distribution on G, the distribution obtained by an $O(\log n)$ steps *lazy random walk* on the directed graph $\mathrm{Cay}(G, J)$ is $\frac{1}{\mathrm{poly}(n)}$- close to the uniform distribution. Then, by Theorem 1.2, it follows that the directed Cayley graph $\mathrm{Cay}(G, J)$ has a constant spectral expansion. Crucially, we note that Pak considers lazy random walks, since his main technical tool is based on *cube generating sequences* for finite groups introduced by Erdös and Rényi in [2].

Definition 1.3 *Let G be a finite group and $J = \langle g_1, \ldots, g_k \rangle$ be a sequence of group elements. For any $\delta > 0$, J is said to be a* cube generating sequence *for G with closeness parameter δ, if the probability distribution D_J on G given by $g_1^{\epsilon_1} \ldots g_k^{\epsilon_k}$, where each ϵ_i is independently and uniformly distributed in $\{0, 1\}$, is δ-close to the uniform distribution in the L_2-norm.*

Erdös and Rényi [2] proved the following theorem.

Theorem 1.4 *Let G be a finite group and $J = \langle g_1, \ldots, g_k \rangle$ be a sequence of k elements of G picked uniformly and independently at random. Let D_J be the distribution on G generated by J, i.e $D_J(x) = \mathsf{Pr}_{\{\epsilon_i \in_R \{0,1\}\ :\ 1 \le i \le k\}}[g_1^{\epsilon_1} \ldots g_k^{\epsilon_k} = x]$ for $x \in G$, and U be the uniform distribution on G. Then the expected value $\mathbb{E}_J \|D_J - U\|_2^2 = 1/2^k(1 - 1/n)$.*

In particular if we choose $k = O(\log n)$, the resulting distribution D_J is $\frac{1}{\text{poly}(n)}$-close to the uniform distribution in L_2 norm.

Our Results

Let G be a finite group with n elements given by its multiplication table. Our first result is a derandomization of a result of Pak [7]. We show a deterministic polynomial-time construction of a generating set J of size $O(\log |G|)$ such that a lazy random walk on $\text{Cay}(G, J)$ mixes fast. Throughout the paper, we measure the distance between two distributions in L_2 norm.

Theorem 1.5 *For any constant $c > 1$, there is a deterministic* $\text{poly}(n)$ *time algorithm that computes a generating set J of size $O(\log n)$ for the given group G, such that given any initial distribution on G the lazy random walk of $O(\log n)$ steps on the directed Cayley graph* $\text{Cay}(G, J)$ *yields a distribution that is $\frac{1}{n^c}$-close (in L_2 norm) to the uniform distribution.*

Theorem 1.5 and Theorem 1.2 together yield the following corollary.

Corollary 1.6 *Given a finite group G and any $\epsilon > 0$, there is a deterministic polynomial-time algorithm to construct an $O(\log n)$ size generating set J such that* $\text{Cay}(G, J)$ *is a spectral expander (i.e. its second largest eigenvalue in absolute value is bounded by ϵ).*

Our next result yields an alternative proof of the Wigderson-Xiao result [10]. In order to carry out a similar approach as the proof of Theorem 1.5 for undirected Cayley graphs, we need a suitable generalization of cube generating sequences, and in particular, a generalization of [2]. Using this generalization, we can give a deterministic $\text{poly}(n)$ time algorithm to compute $J = \langle g_1, g_2, \ldots, g_k \rangle$ where $k = O(\log n)$ such that a lazy random walk of length $O(\log n)$ on $\text{Cay}(G, J \cup J^{-1})$ is $\frac{1}{\text{poly}(n)}$-close to the uniform distribution. Here the lazy random walk is described by the symmetric transition matrix $A_J = \frac{1}{3}I + \frac{1}{3k}(P_J + P_{J^{-1}})$ where P_J and $P_{J^{-1}}$ are the adjacency matrices of the Cayley graphs $\text{Cay}(G, J)$ and $\text{Cay}(G, J^{-1})$ respectively.

Theorem 1.7 *Let G be a finite group of order n and $c > 1$ be any constant. There is a deterministic* $\text{poly}(n)$ *time algorithm that computes a generating set J of size $O(\log n)$ for G, such that an $O(\log n)$ step lazy random walk on G, governed by the transition matrix A_J described above, is $\frac{1}{n^c}$-close to the uniform distribution, for any given initial distribution on G.*

Theorem 1.7 and the connection between mixing time and spectral expansion for undirected graphs given by Theorem 1.1 yields the following.

Corollary 1.8 (Wigderson-Xiao) [10] *Given a finite group G by its multiplication table, there is a deterministic polynomial (in $|G|$) time algorithm to construct a generating set J such that* $\mathrm{Cay}(G, J \cup J^{-1})$ *is a spectral expander.*

Finally, we show that the construction of cube generating sequences can also be done in deterministic polynomial time.

Theorem 1.9 *For any constant $c > 1$, there is a deterministic polynomial (in n) time algorithm that outputs a cube generating sequence J of size $O(\log n)$ such that the distribution D_J on G, defined by the cube generating sequence J, is $\frac{1}{n^c}$-close to the uniform distribution.*

1.1 Organization of the paper

The paper is organized as follows. We prove Theorem 1.5 and Corollary 1.6 in Section 2. The proof of Theorem 1.7 and Corollary 1.8 are given in Section 3. We prove Theorem 1.9 in Section 4. Finally, we summarize in Section 5.

2 Expanding Directed Cayley Graphs

Let D_1 and D_2 be two probability distributions over the finite set $\{1, 2, \ldots, n\}$. We use the L_2 norm to measure the distance between the two distributions: $\|D_1 - D_2\|_2 = \left[\sum_{x\in[n]} |D_1(x) - D_2(x)|^2\right]^{\frac{1}{2}}$.

Let U denote the uniform distribution on $[n]$. We say that a distribution D is δ-close to the uniform distribution if $\|D - U\|_2 \leq \delta$.

Definition 2.1 *The* collision probability *of a distribution D on $[n]$ is defined as* $\mathrm{Coll}(D) = \sum_{i\in[n]} D(i)^2$. *It is easy to see that* $\mathrm{Coll}(D) \leq 1/n + \delta$ *if and only if* $\|D - U\|_2^2 \leq \delta$ *and* $\mathrm{Coll}(D)$ *attains its minimum value $1/n$ only for the uniform distribution.*

We prove Theorem 1.5 by giving a deterministic construction of a cube generating sequence J such that a random walk on $\mathrm{Cay}(G, J)$ mixes in $O(\log n)$ steps. We first describe a randomized construction in Section 2.1, which shows the existence of such a sequence. The construction is based on analysis of [7]. This is then derandomized in Section 2.2.

2.1 Randomized construction

For a sequence of group elements $J = \langle g_1, \ldots, g_k\rangle$, we consider the Cayley graph $\mathrm{Cay}(G, J)$, which is, in general, a directed multigraph in which both in-degree and out-degree of every vertex is k. Let A denote the adjacency matrix of $\mathrm{Cay}(G, J)$. The lazy random walk is defined by the probability transition matrix $(A+I)/2$ where I is the identity matrix. Let Q_J denote the probability distribution obtained after m steps of the lazy random walk. Pak [7] has analyzed the distribution Q_J and shown that for a random J of $O(\log n)$ size and $m = O(\log n)$, Q_J is $1/n^{O(1)}$-close to the uniform distribution. We note that Pak works with the L_∞ norm. Our aim is to give an efficient deterministic construction of J. It turns out for us that the L_2 norm and the collision probability are the right tools to work with since we can compute these quantities exactly as we fix elements of J one by one.

Consider any length-m sequence $I = \langle i_1, \ldots, i_m \rangle \in [k]^m$, where i_js are indices that refer to elements in the set J. Let R_I^J denote the following probability distribution on G. For each $x \in G$: $R_I^J(x) = \mathsf{Pr}_{\bar{\epsilon}}[g_{i_1}^{\epsilon_1} \cdot \ldots \cdot g_{i_m}^{\epsilon_m} = x]$, where $\bar{\epsilon} = (\epsilon_1, \ldots, \epsilon_m)$ and each $\epsilon_i \in \{0,1\}$ is picked independently and uniformly at random. Notice that for each $x \in G$ we have: $Q_J(x) = \frac{1}{k^m} \sum_{I \in [k]^m} R_I^J(x)$.

Further, notice that R_I^J is precisely the probability distribution defined by the cube generating sequence $\langle g_{i_1}, g_{i_2}, \ldots, g_{i_m} \rangle$, and the above equation states that the distribution Q_J is the average over all $I \in [k]^m$ of the R_I^J.

In general, the indices in $I \in [k]^m$ are not distinct. Let $L(I)$ denote the sequence of distinct indices occurring in I, in the order of their *first occurrence* in I, from left to right. We refer to $L(I)$ as the L-subsequence of I. Clearly, the sequence $L(I)$ will itself define a probability distribution $R_{L(I)}^J$ on the group G.

Suppose the elements of J are independently, randomly picked from G. The following lemma shows for any $I \in [k]^m$ that if $R_{L(I)}^J$ is δ-close to uniform distribution (in L_2 norm), in expectation, then so is R_I^J. We state it in terms of collision probabilities.

Lemma 2.2 *For a fixed I, If $\mathbb{E}_J[\mathrm{Coll}(R_{L(I)}^J)] = \mathbb{E}_J[\sum_{g \in G} R_{L(I)}^J(g)^2] \leq 1/n + \delta$ then $\mathbb{E}_J[\mathrm{Coll}(R_I^J)] = \mathbb{E}_J[\sum_{g \in G} R_I^J(g)^2] \leq 1/n + \delta$.*

A proof of Lemma 2.2 is in the appendix to keep our presentation self-contained. A similar lemma for the L_∞ norm is shown in [7, Lemma 1] (though it is not stated there in terms of the expectation).

When elements of J are picked uniformly and independently from G, by Theorem 1.4, $\mathbb{E}_J[\mathrm{Coll}(R_{L(I)}^J)] = \mathbb{E}_J[\sum_{g \in G} R_{L(I)}^J(g)^2] = \frac{1}{n} + \frac{1}{2^\ell}(1 - \frac{1}{n})$, where ℓ is the length of the L-subsequence. Thus the expectation is small provided ℓ is large enough. It turns out that most $I \in [k]^m$ have sufficiently long L-subsequences (Lemma 2.3). A similar result appears in [7]. We give a proof of Lemma 2.3 in the appendix.

Lemma 2.3 [7] *Let $a = \frac{k}{\ell - 1}$. The probability that a sequence of length m over $[k]$ does not have an L-subsequence of length ℓ is at most $\frac{(ae)^{\frac{k}{a}}}{a^m}$.*

To ensure the above probability is bounded by $\frac{1}{2^m}$, it suffices to choose $m > \frac{(k/a)\log(ae)}{\log(a/2)}$.

The following lemma (which is again an L_2 norm version of a similar statement from [7]), we observe that the expected distance from the uniform distribution is small, when $I \in [k]^m$ is picked uniformly at random. The proof of the lemma is given in the appendix.

Lemma 2.4 $\mathbb{E}_J[\mathrm{Coll}(Q_J)] = \mathbb{E}_J[\sum_{g \in G} Q_J(g)^2] \leq \frac{1}{n} + \frac{1}{2^{\Theta(m)}}$.

We can make $\frac{1}{2^{\Theta(m)}} < \frac{1}{n^c}$ for some $c > 0$, by choosing $m = O(\log n)$. That also fixes k to be $O(\log n)$ suitably.

2.2 Deterministic construction

Our goal is to compute, for any given constant $c > 0$, a multiset J of k group elements of G such that $\mathrm{Coll}(Q_J) = \sum_{g \in G} Q_J(g)^2 \ \leq \ 1/n + 1/n^c$, where both k and m are $O(\log n)$. For each J observe, by Cauchy-Schwarz inequality, that

$$\mathrm{Coll}(Q_J) = \sum_{g \in G} Q_J(g)^2 \ \leq \ \sum_{g \in G} \frac{1}{k^m} \sum_{I \in [k]^m} R_I^J(g)^2 = \frac{1}{k^m} \sum_{I \in [k]^m} \mathrm{Coll}(R_I^J). \tag{1}$$

Our goal can now be restated: it suffices to construct in deterministic polynomial time a multiset J of group elements such that the average collision probability $\frac{1}{k^m} \sum_{I \in [k]^m} \mathrm{Coll}(R_I^J) \leq 1/n + 1/n^c$.

Consider the random set $J = \{X_1, \ldots, X_k\}$ with each X_i a uniformly and independently distributed random variable over G. Combined with the proof of Lemma 2.4 (in particular from Equation 17), we observe that for any constant $c > 1$ there are k and m, both $O(\log n)$ such that

$$\mathbb{E}_J[\mathrm{Coll}(Q_J)] \ \leq \ = \ \mathbb{E}_J[\mathbb{E}_{I \in [k]^m} \mathrm{Coll}(R_I^J)] \ \leq \ \frac{1}{n} \ + \ \frac{1}{n^c}. \tag{2}$$

Our deterministic algorithm will fix the elements in J in stages. At stage 0 the set $J = J_0 = \{X_1, X_2, \ldots, X_k\}$ consists of independent random elements X_i drawn from the group G. Suppose at the j^{th} stage, for $j < k$, the set we have is $J = J_j = \{x_1, x_2, \ldots, x_j, X_{j+1}, \ldots, X_k\}$, where each $x_r (1 \leq r \leq j)$ is a fixed element of G and the $X_s (j + 1 \leq s \leq k)$ are independent random elements of G such that

$$\mathbb{E}_J[\mathbb{E}_{I \in [k]^m} \mathrm{Coll}(R_I^J)] \leq 1/n + 1/n^c.$$

Remark.

1. *In the above expression, the expectation is over the random elements of J.*
2. *If we can compute in* $\mathrm{poly}(n)$ *time a choice x_{j+1} for X_{j+1} such that $\mathbb{E}_J[\mathbb{E}_{I \in [k]^m} \mathrm{Coll}(R_I^J)] \leq 1/n + 1/n^c$ then we can compute the desired generating set J in polynomial (in n) time.*

Given $J = J_j = \{x_1, \ldots, x_j, X_{j+1}, \ldots, X_k\}$ with j fixed elements and $k - j$ random elements, it is useful to partition the set of sequences $[k]^m$ into subsets $S_{r,\ell}$ where $I \in S_{r,\ell}$ if and only if there are exactly r indices in I from $\{1, \ldots, j\}$, and of the remaining $m - r$ indices of I there are exactly ℓ distinct indices. We now define a suitable generalization of L-subsequences.

Definition 2.5 *An (r, ℓ)-*normal sequence *for J is a sequence $\{n_1, n_2, \ldots, n_r, \ldots, n_{r+\ell}\} \in [k]^{r+\ell}$ such that the indices $n_s, 1 \leq s \leq r$ are in $\{1, 2, \ldots, j\}$ and the indices $n_s, s > \ell$ are* all distinct *and in $\{j + 1, \ldots, k\}$. I.e. the first r indices (possibly with repetition) are from the fixed part of J and the last ℓ are all distinct elements from the random part of J.*

Transforming $S_{r,\ell}$ to (r, ℓ)-normal sequences

We use the simple fact that if $y \in G$ is picked uniformly at random and $x \in G$ be any element independent of y, then the distribution of xyx^{-1} is uniform in G.

Let $I = \langle i_1, \ldots, i_m \rangle \in S_{r,\ell}$ be a sequence. Let $F = \langle i_{f_1}, \ldots, i_{f_r} \rangle$ be the subsequence of indices for the fixed elements in I. Let $R = \langle i_{s_1}, \ldots, i_{s_{m-r}} \rangle$ be the subsequence of indices for the random elements in I, and $L = \langle i_{e_1}, \ldots, i_{e_\ell} \rangle$ be the L-subsequence in R. More precisely, notice that R is a sequence in $\{j+1,\ldots,k\}^{m-r}$ and L is the L-subsequence for R. The (r,ℓ) normal sequence $\widehat{I}$ of $I \in S_{r,\ell}$ is the sequence $\langle i_{f_1},\ldots,i_{f_r},i_{e_1},\ldots,i_{e_\ell}\rangle$.

We recall here that the multiset $J=\{x_1,\ldots,x_j,X_{j+1}\ldots,X_k\}$ is defined as before. For ease of notation we denote the list of elements of J by $g_t, 1\le t\le k$. I.e. $g_t=x_t$ for $t\le j$ and $g_t=X_t$ for $t>j$. Consider the distribution of the products $g_{i_1}^{\epsilon_1}\ldots g_{i_m}^{\epsilon_m}$ where $\epsilon_i\in\{0,1\}$ are independent and uniformly picked at random. Then we can write

$$\begin{aligned} g_{i_1}^{\epsilon_1}\ldots g_{i_m}^{\epsilon_m} &= z_0 g_{i_{f_1}}^{\epsilon_{f_1}} z_1 g_{i_{f_2}}^{\epsilon_{f_2}} z_2\ldots z_{r-1} g_{i_{f_r}}^{\epsilon_{f_r}} z_r, \text{ where} \\ z_0 z_1\ldots z_r &= g_{i_{s_1}}^{\epsilon_{s_1}} g_{i_{s_2}}^{\epsilon_{s_2}}\ldots g_{i_{s_{m-r}}}^{\epsilon_{s_{m-r}}}. \end{aligned}$$

By conjugation, we can rewrite the above expression as $g_{i_{f_1}}^{\epsilon_{f_1}} z z_1 g_{i_{f_2}}^{\epsilon_{f_2}} z_2\ldots g_{i_{f_r}}^{\epsilon_{f_r}} z_r$, where $z = g_{i_{f_1}}^{-\epsilon_{f_1}} z_0 g_{i_{f_1}}^{\epsilon_{f_1}}$.

We refer to this transformation as moving $g_{i_{f_1}}^{\epsilon_{f_1}}$ to the left. Successively moving the elements $g_{i_{f_1}}^{\epsilon_{f_1}}, g_{i_{f_2}}^{\epsilon_{f_2}},\ldots,g_{i_{f_r}}^{\epsilon_{f_r}}$ to the left we can write

$$g_{i_1}^{\epsilon_1}\ldots g_{i_m}^{\epsilon_m} = g_{i_{f_1}}^{\epsilon_{f_1}}\ldots g_{i_{f_r}}^{\epsilon_{f_r}} z_0' z_1'\ldots z_r',$$

where each $z_t' = u_t z_t u_t^{-1}$, and u_t is a product of elements from the fixed element set $\{x_1,\ldots,x_j\}$. Notice that each z_t is a product of some consecutive sequence of elements from $\langle g_{i_{s_1}}^{\epsilon_{s_1}}, g_{i_{s_2}}^{\epsilon_{s_2}},\ldots,g_{i_{s_{m-r}}}^{\epsilon_{s_{m-r}}}\rangle$. If $z_t=\prod_{a=b}^{c} g_{i_{s_a}}^{\epsilon_{s_a}}$ then $z_t'=\prod_{a=b}^{c} u_t g_{i_{s_a}}^{\epsilon_{s_a}} u_t^{-1}$. Thus, the product $z_0'z_1'\ldots z_r'$, is of the form

$$z_0' z_1'\ldots z_r' = \prod_{a=1}^{m-r} h_{s_a}^{\epsilon_{s_a}},$$

where each $h_{s_a} = y_a g_{i_{s_a}}^{\epsilon_{s_a}} y_a^{-1}$, for some elements $y_a\in G$. In this expression, observe that for distinct indices a and b, we may have $i_{s_a}=i_{s_b}$ and $y_a\ne y_b$ and hence, in general, $h_{s_a}\ne h_{s_b}$.

Recall that the L-subsequence $L=\langle i_{e_1},\ldots,i_{e_\ell}\rangle$ is a subsequence of $R=\langle i_{s_1},\ldots,i_{s_{m-r}}\rangle$. Consequently, let $(h_{e_1},h_{e_2},\ldots,h_{e_\ell})$ be the sequence of all *independent* random elements in the above product $\prod_{a=1}^{m-r} h_{s_a}^{\epsilon_{s_a}}$ that correspond to the L-subsequence. To this product, we again apply the transformation of moving to the left, the elements $h_{e_1}^{\epsilon_{e_1}}, h_{e_2}^{\epsilon_{e_2}},\ldots,h_{e_\ell}^{\epsilon_{e_\ell}}$, in that order. Putting it all together we have

$$g_{i_1}^{\epsilon_1}\ldots g_{i_m}^{\epsilon_m} = g_{i_{f_1}}^{\epsilon_{f_1}}\ldots g_{i_{f_r}}^{\epsilon_{f_r}} h_{e_1}^{\epsilon_{e_1}}\ldots h_{e_\ell}^{\epsilon_{e_\ell}} y(\bar{\epsilon}),$$

where $y(\bar{\epsilon})$ is an element in G that depends on J,I and $\bar{\epsilon}$, where $\bar{\epsilon}$ consists of all the ϵ_j for $i_j\in I\setminus(F\cup L)$. Let $J(I)$ denote the multiset of group elements obtained from J by replacing the subset $\{g_{i_{e_1}}, g_{i_{e_2}},\ldots,g_{i_{e_\ell}}\}$ with $\{h_{e_1},h_{e_2},\ldots,h_{e_\ell}\}$. It follows from our discussion that $J(I)$ has exactly j fixed elements $x_1,x_2,\ldots,x_j$ and $k-j$ uniformly distributed independent random elements. Recall that $\widehat{I}=\langle i_{f_1},i_{f_2},\ldots,i_{f_r},i_{e_1},i_{e_2},\ldots,i_{e_\ell}\rangle$ is the (r,ℓ)-normal sequence for I. Analogous to Lemma 2.2, we now compare the probability distributions R_I^J and $R_{\widehat{I}}^{J(I)}$. The proof of the lemma is in the appendix.

Lemma 2.6 *For each $j \leq k$ and $J = \{x_1, \ldots, x_j, X_{j+1}, \ldots, X_k\}$ (where $x_1, \ldots, x_j \in G$ are fixed elements and $X_{j+1}, \ldots, X_k$ are independent uniformly distributed in G), and for each $I \in [k]^m$, $\mathbb{E}_J[\mathrm{Coll}(R_I^J)] \ \leq \ \mathbb{E}_J[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$.*

Remark 2.7 *Here it is important to note that the expectation $\mathbb{E}_J[\mathrm{Coll}(R_I^J)]$ is over the random elements in J. On the other hand, the expectation $\mathbb{E}_J[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$* is over the *the random elements in $J(I)$ (which are conjugates of the random elements in J). In the rest of this section, we need to keep this meaning clear when we use $\mathbb{E}_J[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$ for different $I \in [k]^m$.*

By averaging the above inequality over all I sequences and using Equation 1, we get

$$\mathbb{E}_J[\mathrm{Coll}(Q_J)] \leq \mathbb{E}_J\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_I^J)] \leq \mathbb{E}_J\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]. \tag{3}$$

Now, by Equation 2 and following the proof of Lemma 2.4, when all k elements in J are random then we have $\mathbb{E}_J\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})] \leq 1/n + 1/n^c$. Suppose for any $J = \{x_1, \ldots, x_j, X_{j+1}, \ldots, X_k\}$ we can compute $\mathbb{E}_J\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$ in deterministic polynomial (in n) time. Then, given the bound $\mathbb{E}_J\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})] \leq 1/n + 1/n^c$ for $J = \{x_1, \ldots, x_j, X_{j+1}, \ldots, X_k\}$, we can clearly fix the $(j+1)^{st}$ element of J by choosing $X_{j+1} := x_{j+1}$ which minimizes the expectation $\mathbb{E}_J\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$. Also, it follows easily from Equation 3 and the above lemma that $\mathbb{E}_J\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})] \leq \delta$ implies $\mathbb{E}_J\mathrm{Coll}(Q_J) \leq \mathbb{E}_J\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_I^J)] \leq \delta$. In particular, when J is completely fixed after k stages, and if $\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})] \leq \delta$ then $\mathrm{Coll}(Q_J) \leq \delta$.

Remark 2.8 *In fact, the quantity $\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$ plays the role of a pessimistic estimator for $\mathbb{E}_{I\in[k]^m}[\mathrm{Coll}(R_I^J)]$.*

We now proceed to explain the algorithm that fixes X_{j+1}. To this end, it is useful to rewrite this as

$$\begin{aligned}\mathbb{E}_J\mathbb{E}_I[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})] &= \frac{1}{k^m}\left[\sum_{r,\ell}\sum_{I\in S_{r,\ell}}\mathbb{E}_J[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]\right] \\ &= \sum_{r,\ell}\frac{|S_{r,\ell}|}{k^m}\mathbb{E}_{I\in S_{r,\ell}}\mathbb{E}_J[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})] \end{aligned} \tag{4}$$

For any r, ℓ the size of $S_{r,\ell}$ is computable in polynomial time (Lemma 2.9). We include a proof in the appendix.

Lemma 2.9 *For each r and ℓ, $|S_{r,\ell}|$ can be computed in time polynomial in n.*

Since r, ℓ is of $O(\log n)$, it is clear from Equation 4 that it suffices to compute $\mathbb{E}_{I\in S_{r,\ell}}\mathbb{E}_J[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$ in polynomial time for any given r and ℓ. We reduce this computation to counting number of paths in weighted directed acyclic graphs. To make the reduction clear, we simply the expression $\mathbb{E}_{I\in S_{r,\ell}}\mathbb{E}_J[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$ as follows.

Let $\bar{u}$ be a sequence of length r from the fixed elements $x_1, x_2, \ldots, x_j$. We identify $\bar{u}$ as an element in $[j]^r$. The number of I sequences in $S_{r,\ell}$ that have $\bar{u}$ as the prefix in the (r,ℓ) normal sequence $\widehat{I}$ is $\frac{|S_{r,\ell}|}{j^r}$. Recall that $R_{\widehat{I}}^{J(I)}(g) = \text{Prob}_{\bar{\epsilon}}[g_{i_{f_1}}^{\epsilon_1} \ldots g_{i_{f_r}}^{\epsilon_r} h_{e_1}^{\epsilon_{r+1}} \ldots h_{e_\ell}^{\epsilon_{r+\ell}} = g]$. Let $\bar{u} = (g_{i_{f_1}}, \ldots, g_{i_{f_r}})$. It is convenient to denote the element $g_{i_{f_1}}^{\epsilon_1} \ldots g_{i_{f_r}}^{\epsilon_r} h_{e_1}^{\epsilon_{r+1}} \ldots h_{e_\ell}^{\epsilon_{r+\ell}}$ by $M(\bar{u}, \bar{\epsilon}, \widehat{I}, J)$.

Let $\bar{\epsilon} = (\epsilon_1, \ldots, \epsilon_{r+\ell})$ and $\bar{\epsilon}' = (\epsilon_1', \ldots, \epsilon_{r+\ell}')$ be random uniformly picked from $\{0,1\}^{r+\ell}$. Then

$$\begin{aligned} \text{Coll}(R_{\widehat{I}}^{J(I)}) &= \sum_{g\in G} (R_{\widehat{I}}^{J(I)}(g))^2 \\ &= \text{Prob}_{\bar{\epsilon},\bar{\epsilon}'}[M(\bar{u}, \bar{\epsilon}, \widehat{I}, J) = M(\bar{u}, \bar{\epsilon}', \widehat{I}, J)]. \end{aligned} \tag{5}$$

For fixed $\bar{\epsilon}, \bar{\epsilon}'$ and $\bar{u} \in [j]^r$, let $S_{r,\ell}^{\bar{u}}$ be the set of all $I \in S_{r,\ell}$ such that the subsequence of indices of I for the fixed elements $\{x_1, x_2, \ldots, x_j\}$ is precisely $\bar{u}$. Notice that $|S_{r,\ell}^{\bar{u}}| = \frac{|S_{r,\ell}|}{j^r}$.

Then we have the following.

$$\mathbb{E}_{I\in S_{r,\ell}} \mathbb{E}_J[\sum_{g\in G} (R_{\widehat{I}}^{J(I)}(g))^2] = \frac{1}{2^{2(\ell+r)}} \left[\sum_{\bar{\epsilon},\bar{\epsilon}'\in\{0,1\}^{\ell+r}} \frac{1}{|S_{r,\ell}|} \sum_{\bar{u}\in[j]^r} \sum_{I\in S_{r,\ell}^{\bar{u}}} \mathbb{E}_J[\chi_{M(\bar{u},\bar{\epsilon},\widehat{I},J)=M(\bar{u},\bar{\epsilon}',\widehat{I},J)}] \right] \tag{6}$$

where $\chi_{M(\bar{u},\bar{\epsilon},\widehat{I},J)=M(\bar{u},\bar{\epsilon}',\widehat{I},J)}$ is a $0-1$ indicator random variable that gets 1 when $M(\bar{u}, \bar{\epsilon}, \widehat{I}, J) = M(\bar{u}, \bar{\epsilon}', \widehat{I}, J)$ and 0 otherwise. Crucially, we note the following:

Claim 2.10 *For each $I \in S_{r,\ell}^{\bar{u}}$ and for fixed $\bar{\epsilon}, \bar{\epsilon}'$, the random variables $\chi_{M(\bar{u},\bar{\epsilon},\widehat{I},J)=M(\bar{u},\bar{\epsilon}',\widehat{I},J)}$ are identically distributed.*

The claim follows from the fact that for each $I \in S_{r,\ell}^{\bar{u}}$, the fixed part in $\widehat{I}$ is $\bar{u}$ and elements in the unfixed part are identically and uniformly distributed in G. We simplify the expression in Equation 6 further.

$$\begin{aligned} \frac{1}{|S_{r,\ell}|} \left[\sum_{\bar{u}\in[j]^r} \sum_{I\in S_{r,\ell}^{\bar{u}}} \mathbb{E}_J[\chi_{M(\bar{u},\bar{\epsilon},\widehat{I},J)=M(\bar{u},\bar{\epsilon}',\widehat{I},J)}] \right] &= \frac{1}{|S_{r,\ell}|} \left[\sum_{\bar{u}\in[j]^r} \frac{|S_{r,\ell}|}{j^r} \mathbb{E}_J[\chi_{M(\bar{u},\bar{\epsilon},\widehat{I},J)=M(\bar{u},\bar{\epsilon}',\widehat{I},J)}] \right] &(7) \\ &= \sum_{\bar{u}\in[j]^r} \frac{1}{j^r} \mathbb{E}_J[\chi_{M(\bar{u},\bar{\epsilon},\widehat{I},J)=M(\bar{u},\bar{\epsilon}',\widehat{I},J)}] &(8) \end{aligned}$$

where Equation 7 follows from Claim 2.10. Let $p_{\bar{u}}(\bar{\epsilon}, \bar{\epsilon}')$ be the number of different assignments of ℓ random elements in J such that $M(\bar{u}, \bar{\epsilon}, \widehat{I}, J) = M(\bar{u}, \bar{\epsilon}', \widehat{I}, J)$. Then it is easy to see that

$$\sum_{\bar{u}\in[j]^r} \frac{1}{j^r} \mathbb{E}_J[\chi_{M(\bar{u},\bar{\epsilon},\widehat{I},J)=M(\bar{u},\bar{\epsilon}',\widehat{I},J)}] = \sum_{\bar{u}} \frac{1}{j^r}\, p_{\bar{u}}(\bar{\epsilon}, \bar{\epsilon}')\, \frac{1}{n^\ell}, \tag{9}$$

where the factor $\frac{1}{n^\ell}$ accounts for the fact that ℓ unfixed elements of J are picked uniformly and independently at random from the group G.

Notice that $2^{r+\ell} \leq 2^m = n^{O(1)}$ for $m = O(\log n)$ and $\bar{\epsilon}, \bar{\epsilon}' \in \{0,1\}^{r+\ell}$. Then, combining the Equation 4 and Equation 9, it is clear that to compute $\mathbb{E}_J \mathbb{E}_I[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})]$ in polynomial time, it suffices to compute $\left[\sum_{\bar{u}\in[j]^r} \frac{1}{j^r}\ p_{\bar{u}}(\bar{\epsilon}, \bar{\epsilon}')\ \frac{1}{n^\ell}\right]$ (for fixed $r, \ell, \bar{\epsilon}, \bar{\epsilon}'$) in polynomial time. We now turn to this problem.

2.3 Reduction to counting paths in weighted DAGs

We will interpret the quantity $\left[\sum_{\bar{u}\in[j]^r} \frac{1}{j^r}\ p_{\bar{u}}(\bar{\epsilon}, \bar{\epsilon}')\ \frac{1}{n^\ell}\right]$ as the sum of weights of paths between a source vertex s and sink vertex t in a layered weighted directed acyclic graph $H = (V, E)$. The vertex set V is $G \times G \times [r+\ell+1] \cup \{s,t\}$, and $s = (e,e,0)$, where e is the identity element in G. The source vertex s is at 0-th layer and the sink t is at the $r+\ell+2$-th layer. Let $S = \{x_1, x_2, \ldots, x_j\}$. The edge set is the union $E = E_s \cup E_S \cup E_{G\setminus S} \cup E_t$, where

$$\begin{array}{rcl} E_s & = & \{(s, (g,h,1)) \mid g,h \in G\} \\ E_S & = & \{((g,h,t), (gx^{\epsilon_t}, hx^{\epsilon'_t}, t+1)) \mid g,h \in G, x \in S, 1 \leq t \leq r\}, \\ E_{G\setminus S} & = & \{((g,h,t), (gx^{\epsilon_t}, hx^{\epsilon'_t}, t+1)) \mid g,h \in G, x \in G, r < t \leq r+\ell\}, \text{ and} \\ E_t & = & \{((g,g,r+\ell+1), t) \mid g \in G\}. \end{array}$$

All edges in E_s and E_t have weights 1 each. Each edge in E_S has weight $\frac{1}{j}$. Each edge in $E_{G\setminus S}$ has weight $\frac{1}{n}$.

Each s-to-t directed path in the graph G corresponds to an (r,ℓ)-normal sequence $\widehat{I}$ (corresponding to some $I \in S_{r,\ell}$), along with an assignment of group elements to the ℓ distinct independent random elements that occur in it. For a random $I \in S_{r,\ell}$, the group element corresponding to each of the r "fixed" positions is from $\{x_1, x_2 \ldots, x_j\}$ with probability $1/j$ each. Hence each edge in E_S has weight $1/j$. Similarly, the ℓ distinct indices in I (from $\{X_{j+1}, \ldots, X_k\}$) are assigned group elements independently and uniformly at random. Hence edges in $E_{G\setminus S}$ has weight $\frac{1}{n}$.

The weight of an s-to-t path is a product of the weights of edges on the path. The graph depends on $j, \bar{\epsilon}$, and $\bar{\epsilon}'$. So for fixed r, ℓ, we denote it as $H_{r,\ell}(j, \bar{\epsilon}, \bar{\epsilon}')$. The following claim is immediate from the Equation 9.

Claim 2.11 *The sum of weights of all s to t paths in $H_{j,\bar{\epsilon},\bar{\epsilon}'}$ is $\sum_{\bar{u}\in[j]^r} \frac{1}{j^r}\ p_{\bar{u}}(\bar{\epsilon}, \bar{\epsilon}')\ \frac{1}{n^\ell}$.*

In the following lemma we observe that $\left[\sum_{\bar{u}\in[j]^r} \frac{1}{j^r}\ p_{\bar{u}}(\bar{\epsilon}, \bar{\epsilon}')\ \frac{1}{n^\ell}\right]$ can be computed in polynomial time. The proof is easy.

Lemma 2.12 *For each $j, \bar{\epsilon}, \bar{\epsilon}', r, \ell$, the quantity $\left[\sum_{\bar{u}\in[j]^r} \frac{1}{j^r}\ p_{\bar{u}}(\bar{\epsilon}, \bar{\epsilon}')\ \frac{1}{n^\ell}\right]$ can be computed in time polynomial in n.*

Proof: The graph $H_{r,\ell}(j, \bar{\epsilon}, \bar{\epsilon}')$ has n^2 vertices in each intermediate layer. For each $1 \leq t \leq r+\ell+2$, we define a matrix M_{t-1} whose rows are indexed by the vertices of layer $t-1$ and columns by vertices of layer t, and the $(a,b)^{th}$ entry of M_{t-1} is the weight of the edge (a,b) in the graph $H_{j,\bar{\epsilon},\bar{\epsilon}'}$. Their product $M = \prod_{t=0}^{r+\ell+1} M_t$ is a scalar which is precisely $\left[\sum_{\bar{u}\in[j]^r} \frac{1}{j^r}\ p_{\bar{u}}(\bar{\epsilon}, \bar{\epsilon}')\ \frac{1}{n^\ell}\right]$ As the product of the matrices M_t can be computed in time polynomial in n, the lemma follows. □

To summarize, we describe the $(j+1)^{st}$ stage of the algorithm, where a group element x_{j+1} is chosen for X_{j+1}. The algorithm cycles through all n choices for x_{j+1}. For each choice of x_{j+1}, and for each $\bar{\epsilon}, \bar{\epsilon}'$, and r, ℓ, the graph $H_{r,\ell}(j+1, \bar{\epsilon}, \bar{\epsilon}')$ is constructed. Using Lemma 2.12, the expression in 4 is computed for each choice of x_{j+1} and the algorithm fixes the choice that minimizes this expression. This completes the proof of Theorem 1.5.

By Theorem 1.2 we can bound the absolute value of the second largest eigenvalue of the matrix for $\mathrm{Cay}(G, J)$. Theorem 1.5 yields that the resulting distribution after an $O(\log n)$ step random walk on $\mathrm{Cay}(G, J)$ is $\frac{1}{\mathrm{poly}(n)}$ close to the uniform distribution in the L_2 norm. Theorem 1.2 is in terms of the L_1 norm. However, since $|L_1| \leq n|L_\infty| \leq n|L_2|$, Theorem 1.5 guarantees that the resulting distribution is $\frac{1}{poly(n)}$ close to the uniform distribution also in L_1 norm. Choose $\tau = m = c' \log n$ and $\epsilon = \frac{1}{n^c}$ in Theorem 1.2, where c, c' are fixed from Theorem 1.5. Then $|\lambda_{max}| \leq \frac{1}{2^{O(c/c')}} < 1$. This completes the proof of Corollary 1.6.□

3 Undirected Expanding Cayley Graphs

In this section, we show a deterministic polynomial-time construction of a generating set J for any group G (given by table) such that a lazy random walk on the *undirected* Cayley graph $\mathrm{Cay}(G, J \cup J^{-1})$ mixes well. As a consequence, we get Cayley graphs which have a constant spectral gap (an alternative proof of a result in [10]). Our construction is based on a simple adaptation of techniques used in Section 2.

The key point in the undirected case is that we will consider a generalization of Erdös-Renyi sequences. We consider the distribution on G defined by $g_1^{\epsilon_1} \dots g_k^{\epsilon_k}$ where $\epsilon_i \in_R \{-1, 0, 1\}$. The following lemma is an easy generalization the Erdös-Renyi result (Theorem 1.4). A similar theorem appears in [3, Theorem 14]. Our main focus in the current paper is the derandomized construction of Cayley expanders. Towards that, to make our paper self-contained, we include a short proof of Lemma 3.1 in the appendix.

Lemma 3.1 *Let G be a finite group and $J = \langle g_1, \dots, g_k \rangle$ be a sequence of k elements of G picked uniformly and independently at random. Let D_J be the following distribution: $D_J(x) = \mathsf{Pr}_{\{\epsilon_i \in_R \{-1,0,1\} \,:\, 1 \leq i \leq k\}}[g_1^{\epsilon_1} \dots g_k^{\epsilon_k} = x]$ for $x \in G$, and U be the uniform distribution on G. Then $\mathbb{E}_J\left[\sum_{x \in G} (D_J(x))^2\right] = \mathbb{E}_J[\mathrm{Coll}(D_J)] \leq \left(\frac{8}{9}\right)^k + \frac{1}{n}$.*

Deterministic construction

First, we note that analogues of Lemma 2.2, 2.3, and 2.4 hold in the undirected case too. In particular, When elements of J are picked uniformly and independently from G, by Lemma 3.1, we have $\mathbb{E}_J[\mathrm{Coll}(R^J_{L(I)})] = \mathbb{E}_J\left[\sum_{g \in G}\left(R^J_{L(I)}(g)\right)^2\right] \leq (\frac{8}{9})^\ell + \frac{1}{n}$, where ℓ is the length of the L-subsequence $L(I)$ of I. Now we state Lemma 3.2 below, which is a restatement of Lemma 2.4 for the undirected case. The proof is exactly similar to the proof of Lemma 2.4. As before, we again consider the probability that an I sequence of length m does not have an L sequence of length ℓ. Also, we fix ℓ, m to $O(\log n)$ appropriately.

Lemma 3.2 *Let $Q_J(g) = \frac{1}{k^m} \sum_{I \in [k]^m} R_I(g)$. Then $\mathbb{E}_J[\mathrm{Coll}(Q_J)] = \mathbb{E}_J[\sum_{g \in G} Q_J(g)^2] \leq 1/n + 2\left(\frac{8}{9}\right)^{\Theta(m)}$.*

Building on this, we can extend the results in Section 2.2 to the undirected case too in a straightforward manner. In particular, we can use essentially the same algorithm as described in Lemma 2.12 to compute the quantity in Equation 5 in polynomial time also in the undirected setting. The only difference we need to incorporate is that now $\bar{\epsilon}, \bar{\epsilon}' \in \{-1,0,1\}^{r+\ell}$. This essentially completes the proof of Theorem 1.7. We do not repeat all the details here.

Finally, we derive Corollary 1.8. The normalized adjacency matrix of the undirected Cayley graph (corresponding to the lazy walk we consider) is given by $A = \frac{1}{3}I + \frac{1}{3k}(P_J + P_{J^{-1}})$ where P_J and $P_{J^{-1}}$ are the corresponding permutation matrices defined by the sets J and J^{-1}. As in the proof of Corollary 1.8, we bound the distance of the resulting distribution from the uniform distribution in L_1 norm. Let $m = c' \log n$ be suitably fixed from the analysis and $|A^m \bar{v} - \bar{u}|_1 \leq \frac{1}{n^c}$. Then by Theorem 1.1, the spectral gap $1 - |\lambda_1| \geq \frac{c}{c'}$. Hence the Cayley graph is a spectral expander. It follows easily that the standard undirected Cayley graph with adjacency matrix $\frac{1}{2k}(P_J + P_{J^{-1}})$ is also a spectral expander.

4 Deterministic construction of Erdös-Rényi sequences

In this section, we prove Theorem 1.9. We use the method of conditional expectations as follows: From Theorem 1.4, we know that $E_J \|D_J - U\|_2^2 = \frac{1}{2^k}\left(1 - \frac{1}{n}\right)$. Therefore there exists a setting of J, say $J = \langle x_1, \ldots, x_k \rangle$, such that $\|D_J - U\|_2^2 \leq \frac{1}{2^k}\left(1 - \frac{1}{n}\right)$. We find such a setting of J by fixing its elements one by one. Let $\delta = \frac{1}{n^c}, c > 1$ be the required closeness parameter. Thus we need k such that $\frac{1}{2^k} \leq \delta$. It suffices to take $k > c \log n$. We denote the expression $X_{i_1}^{\epsilon_1} \ldots X_{i_t}^{\epsilon_t}$ by $\bar{X}^{\bar{\epsilon}}$ when the length t of the sequence is clear from the context.

Let after ith step, $x_1, \ldots, x_i$ be fixed and $X_{i+1}, \ldots, X_k$ are to be picked. At this stage, by our choice of $x_1, \ldots, x_i$, we have $E_{J=(X_{i+1},\ldots,X_k)}(\|D_J - U\|_2^2 \mid X_1 = x_1, \ldots, X_i = x_i) \leq \frac{1}{2^k}(1 - \frac{1}{n})$. Now we cycle through all the group elements for X_{i+1} and fix $X_{i+1} = x_{i+1}$ such that the $E_{J=(X_{i+2},\ldots,X_k)}(\|D_J - U\|_2^2 \mid X_1 = x_1, \ldots, X_{i+1} = x_{i+1}) \leq \frac{1}{2^k}(1 - \frac{1}{n})$. Such an x_{i+1} always exists by a standard averaging argument. In the next theorem, we show that the conditional expectations are efficiently computable at every stage. Theorem 1.9 is an immediate corollary.

Assume that we have picked $x_1, \ldots, x_i$ from G, and $X_{i+1}, \ldots, X_k$ are to be picked from G. Let the choice of $x_1, \ldots, x_i$ be such that $E_{J=(X_{i+1},\ldots,X_k)}(\|D_J - U\|_2^2 \mid X_1 = x_1, \ldots, X_i = x_i) \leq \frac{1}{2^k}(1 - \frac{1}{n})$.

Let, for $x \in G$ and $J = \langle X_1, \ldots, X_k \rangle$

$$Q_J(x) = \mathsf{Pr}_{\bar{\epsilon} \in \{0,1\}^k}\left[\bar{X}^{\bar{\epsilon}} = x\right]$$

When J is partly fixed,

$$\begin{aligned}
\widehat{Q}_J(x) &= \mathsf{Pr}_{\bar{\epsilon_1} \in \{0,1\}^i, \bar{\epsilon_2} \in \{0,1\}^{k-i}}\left[\bar{x}^{\bar{\epsilon_1}} \cdot \bar{X}^{\bar{\epsilon_2}} = x\right] \\
&= \sum_{y \in G} \mathsf{Pr}_{\bar{\epsilon_1}}\left[\bar{x}^{\bar{\epsilon_1}} = y\right] \mathsf{Pr}_{\bar{\epsilon_2}}\left[\bar{X}^{\bar{\epsilon_2}} = y^{-1}x\right] \\
&= \sum_{y \in G} \mu(y) \mathsf{Pr}_{\bar{\epsilon_2}}\left[\bar{X}^{\bar{\epsilon_2}} = y^{-1}x\right] \\
&= \sum_{y \in G} \mu(y) \widehat{Q}_{\bar{X}}(y^{-1}x)
\end{aligned}$$

where $\mu(y) = \mathsf{Pr}_{\bar{\epsilon}_1}\big[\bar{x}^{\bar{\epsilon}_1} = y\big]$. Then $\mathbb{E}_J[\mathrm{Coll}(D_J)] = \mathbb{E}_J\|D_J - U\|_2^2 + \frac{1}{n}$, and $\mathbb{E}_J[\mathrm{Coll}(\widehat{Q}_J)] = (\mathbb{E}_J\|D_J - U\|_2^2 \; |X_1 = x_1, X_2 = x_2, \ldots, X_i = x_i) + \frac{1}{n}$.

Next theorem completes the proof.

Theorem 4.1 *For any finite group G of order n given as multiplication table, $\mathbb{E}_J[\mathrm{Coll}(\widehat{Q}_J)]$ can be computed in time polynomial in n.*

Proof:

$$\mathbb{E}_J[\mathrm{Coll}(\widehat{Q}_J)] = \mathbb{E}_J \sum_{x \in G} \widehat{Q}_J^2(x). \tag{10}$$

Now we compute $\mathbb{E}_J \sum_{x \in G} \widehat{Q}_J^2(x)$.

$$\begin{aligned}
\mathbb{E}_J \sum_{x \in G} \widehat{Q}_J^2(x) &= \mathbb{E}_J \sum_{x \in G} \big(\sum_{y \in G} \mu(y)\widehat{Q}_{\bar{X}}(y^{-1}x)\big)\big(\sum_{z \in G} \mu(z)\widehat{Q}_{\bar{X}}(z^{-1}x)\big) \\
&= \sum_{y,z \in G} \mu(y)\mu(z)\mathbb{E}_J \sum_{x \in G} \big[\widehat{Q}_{\bar{X}}(y^{-1}x)\widehat{Q}_{\bar{X}}(z^{-1}x)\big].
\end{aligned} \tag{11}$$

Now,

$$\begin{aligned}
\sum_{x \in G} \big[\widehat{Q}_{\bar{X}}(y^{-1}x)\widehat{Q}_{\bar{X}}(z^{-1}x)\big] &= \sum_{x \in G} \mathsf{Pr}_{\bar{\epsilon}}\big[\bar{X}^{\bar{\epsilon}} = y^{-1}x\big]\mathsf{Pr}_{\bar{\epsilon}'}\big[\bar{X}^{\bar{\epsilon}'} = z^{-1}x\big] \\
&= \frac{1}{2^{2k}} \sum_{x,\bar{\epsilon},\bar{\epsilon}'} \chi_{y^{-1}x}(\bar{\epsilon})\chi_{z^{-1}x}(\bar{\epsilon}') \\
&= \frac{1}{2^{2k}} \Bigg(\sum_{\bar{\epsilon}=\bar{\epsilon}'} \sum_{x \in G} \chi_{y^{-1}x}(\bar{\epsilon})\chi_{z^{-1}x}(\bar{\epsilon}') + \sum_{\bar{\epsilon}\neq\bar{\epsilon}'} \sum_{x \in G} \chi_{y^{-1}x}(\bar{\epsilon})\chi_{z^{-1}x}(\bar{\epsilon}') \Bigg)
\end{aligned} \tag{12}$$

where $\chi_a(\bar{\epsilon})$ is an indicator variable which is 1 if $\bar{X}^{\bar{\epsilon}} = a$ and 0 otherwise. If $\bar{\epsilon} = \bar{\epsilon}'$ then $\chi_{y^{-1}x}(\bar{\epsilon}) \cdot \chi_{z^{-1}x}(\bar{\epsilon}') = \delta_{y,z}$, where $\delta_{a,b} = 1$ whenever $a = b$ and 0 otherwise.

For $\bar{\epsilon} \neq \bar{\epsilon}'$, $\chi_{y^{-1}x}(\bar{\epsilon}) \cdot \chi_{z^{-1}x}(\bar{\epsilon}') = 1$ only if $y\bar{X}^{\bar{\epsilon}} = z\bar{X}^{\bar{\epsilon}'} = x$. Therefore for $\bar{\epsilon} \neq \bar{\epsilon}'$, we have

$$\frac{1}{2^{2k}} \sum_{\bar{\epsilon}\neq\bar{\epsilon}'} \sum_{x \in G} \chi_{y^{-1}x}(\bar{\epsilon}) \cdot \chi_{z^{-1}x}(\bar{\epsilon}') \;=\; \mathbb{E}_{\bar{\epsilon},\bar{\epsilon}'}\delta_{y\bar{X}^{\bar{\epsilon}},z\bar{X}^{\bar{\epsilon}'}}(1 - \delta_{\bar{\epsilon},\bar{\epsilon}'}).$$

Putting this in Equation 12, we get

$$\frac{1}{2^{2k}} \Bigg(\sum_{\bar{\epsilon}=\bar{\epsilon}'} \sum_{x \in G} \chi_{y^{-1}x}(\bar{\epsilon})\chi_{z^{-1}x}(\bar{\epsilon}') + \sum_{\bar{\epsilon}\neq\bar{\epsilon}'} \sum_{x \in G} \chi_{y^{-1}x}(\bar{\epsilon})\chi_{z^{-1}x}(\bar{\epsilon}') \Bigg) \;=\; \frac{n}{2^k}\delta_{y,z} + \mathbb{E}_{\bar{\epsilon},\bar{\epsilon}'}\delta_{y\bar{X}^{\bar{\epsilon}},z\bar{X}^{\bar{\epsilon}'}}(1 - \delta_{\bar{\epsilon},\bar{\epsilon}'}).$$

Therefore we get

$$\begin{aligned}
\mathbb{E}_J \sum_{x \in G} \widehat{Q}_{\bar{X}}(y^{-1}x) \cdot \widehat{Q}_{\bar{X}}(z^{-1}x) &= \frac{n}{2^k}\delta_{y,z} + \mathbb{E}_J\big[\mathbb{E}_{\bar{\epsilon},\bar{\epsilon}'}\big[\delta_{y\bar{X}^{\bar{\epsilon}},z\bar{X}^{\bar{\epsilon}'}}(1 - \delta_{\bar{\epsilon},\bar{\epsilon}'})\big]\big] \\
&= \frac{n}{2^k}\delta_{y,z} + \mathbb{E}_{\bar{\epsilon},\bar{\epsilon}'}\big[(1 - \delta_{\bar{\epsilon},\bar{\epsilon}'})\mathbb{E}_J\big[\delta_{y\bar{X}^{\bar{\epsilon}},z\bar{X}^{\bar{\epsilon}'}}\big]\big] \\
&= \frac{n}{2^k}\delta_{y,z} + \mathbb{E}_{\bar{\epsilon},\bar{\epsilon}'}\big[(1 - \delta_{\bar{\epsilon},\bar{\epsilon}'})\mathsf{Pr}_{\bar{X}}(y\bar{X}^{\bar{\epsilon}} = z\bar{X}^{\bar{\epsilon}'})\big]
\end{aligned} \tag{13}$$

Claim 4.2 *For $\bar{\epsilon} \neq \bar{\epsilon}'$, $\mathsf{Pr}_{\bar{X}}(y\bar{X}^{\bar{\epsilon}} = z\bar{X}^{\bar{\epsilon}'}) = \frac{1}{n}$.*

Proof: Let j be the smallest index from left such that $\epsilon_j \neq \epsilon'_j$. Let $X_{i+1}^{\epsilon_1} \cdot \ldots \cdot X_{i+j-1}^{\epsilon_{j-1}} = a$. Let $X_{i+j+1}^{\epsilon_{i+1}} \cdot \ldots \cdot X_k^{\epsilon_{k-i}} = b$ and $X_{i+j+1}^{\epsilon'_{i+1}} \cdot \ldots \cdot X_k^{\epsilon'_{k-i}} = b'$. Also, without loss of generality, let $\epsilon_j = 1$ and $\epsilon'_j = 0$. Then we have $\mathsf{Pr}_{\bar{X}}(y\bar{X}^{\bar{\epsilon}} = z\bar{X}^{\bar{\epsilon}'}) = \mathsf{Pr}_{X_{i+j}}(yaX_{i+j}b = zab') = \frac{1}{n}$. □

Thus Equation 13 becomes

$$\mathbb{E}_J \sum_{x \in G} \widehat{Q}_{\bar{X}}(y^{-1}x) \cdot \widehat{Q}_{\bar{X}}(z^{-1}x) = \frac{n}{2^k}\delta_{y,z} + \frac{2^{2k} - 2^k}{n2^{2k}}$$

Putting this in Equation 11, we get

$$\mathbb{E}_J[\mathrm{Coll}(\widehat{Q}_J)] = \mathbb{E}_J \sum_{x \in G} \widehat{Q}_J^2(x) = \sum_{y,z \in G} \frac{1}{2^{2k}} \left[2^k \cdot n \cdot \delta_{y,z} + (2^{2k} - 2^k) \cdot \frac{1}{n}\right] \mu(y)\mu(z) \tag{14}$$

Clearly, for any $y \in G$, $\mu(y)$ can be computed in time $O(2^i)$ which is a polynomial in n since $i \leq k = O(\log n)$. Also from Equation 14, it is clear that $\mathbb{E}_J[\mathrm{Coll}(\widehat{Q}_J)]$ is computable in polynomial (in n) time. □□

5 Summary

Constructing explicit Cayley expanders on finite groups is an important problem. In this paper, we give simple deterministic construction of Cayley expanders that have a constant spectral gap. Our method is completely different and elementary than the existing techniques [10].

The main idea behind our work is a deterministic polynomial-time construction of a cube generating sequence J of size $O(\log |G|)$ such that $\mathrm{Cay}(G, J)$ has a rapid mixing property. In randomized setting, Pak [7] has used similar ideas to construct Cayley expanders. In particular, we also give a derandomization of an well-known result of Erdös and Rényi [2].

References

[1] Noga Alon and Yuval Roichman. Random cayley graphs and expanders. *Random Struct. Algorithms*, 5(2):271–285, 1994.

[2] Paul Erdös and Alfréd Rényi. Probabilistic methods in group theory. *Journal D'analyse Mathematique*, 14(1):127–138, 1965.

[3] Martin Hildebrand. A survey of results on random random walks on finite groups. *Probability Surveys*, 2:33–63, 2005.

[4] Shlomo Hoory, Nati Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. AMS*, 43(4):439–561, 2006.

[5] Alex Lubotzky, R. Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[6] Ravi Montenegro and Prasad Tetali. Mathematical aspects of mixing times in markov chains. *Foundations and Trends in Theoretical Computer Science*, 1(3), 2005.

[7] Igor Pak. Random cayley graphs with $o(\log[g])$ generators are expanders. In *Proceedings of the 7th Annual European Symposium on Algorithms*, ESA '99, pages 521–526. Springer-Verlag, 1999.

[8] Dana Randall. Rapidly mixing markov chains with applications in computer science and physics. *Computing in Science and Engg.*, 8(2):30–41, 2006.

[9] Omer Reingold. Undirected connectivity in log-space. *J. ACM*, 55(4), 2008.

[10] Avi Wigderson and David Xiao. Derandomizing the ahlswede-winter matrix-valued chernoff bound using pessimistic estimators, and applications. *Theory of Computing*, 4(1):53–76, 2008.

Appendix

We include a proof of Lemma 2.2.

Proof of Lemma 2.2

Proof: We use the simple fact that if $y \in G$ is picked uniformly at random and $x \in G$ be any element independent of y, then the distribution of xyx^{-1} is uniform in G.

Let $I = \langle i_1, \ldots, i_m \rangle$, and $L = \langle i_{r_1}, \ldots, i_{r_\ell} \rangle$ be the corresponding L-subsequence (clearly, $r_1 = 1$). Let $J = \langle g_1, g_2, \ldots, g_k \rangle$ be uniform and independent random elements from G. Consider the distribution of the products $g_{i_1}^{\epsilon_1} \ldots g_{i_m}^{\epsilon_m}$ where $\epsilon_i \in \{0, 1\}$ are independent and uniformly picked at random. Then we can write

$$g_{i_1}^{\epsilon_1} \ldots g_{i_m}^{\epsilon_m} \quad = \quad g_{i_{r_1}}^{\epsilon_{r_1}} x_1 g_{i_{r_2}}^{\epsilon_{r_2}} x_2 \ldots x_{\ell-1} g_{i_{r_\ell}}^{\epsilon_{r_\ell}} x_\ell,$$

where, by definition of L-subsequence, notice that x_j is a product of elements from $\{g_{i_{r_1}}, g_{i_{r_2}}, \ldots, g_{i_{r_{j-1}}}\}$ for each j. By conjugation, we can rewrite the above expression as $g_{i_{r_1}}^{\epsilon_{r_1}} x_1 g_{i_{r_2}}^{\epsilon_{r_2}} x_2 \ldots h^{\epsilon_{r_\ell}} x_{\ell-1} x_\ell$, where

$$h^{\epsilon_{r_\ell}} = x_{\ell-1} g_{i_{r_\ell}}^{\epsilon_{r_\ell}} x_{\ell-1}^{-1}.$$

We refer to this transformation as moving $x_{\ell-1}$ to the right. Successively applying this transformation to $x_{\ell-2}, x_{\ell-3} \ldots, x_1$ we can write

$$g_{i_1}^{\epsilon_1} \ldots g_{i_m}^{\epsilon_m} \quad = \quad h_{i_{r_1}}^{\epsilon_{r_1}} h_{i_{r_2}}^{\epsilon_{r_2}} \ldots h_{i_{r_\ell}}^{\epsilon_{r_\ell}} x_1 x_2 \ldots x_{\ell-1} x_\ell,$$

where each $h_{i_{r_j}}$ is a conjugate $z_j g_{i_{r_j}} z_j^{-1}$. Crucially, notice that the group element z_j is a product of elements from $\{g_{i_{r_1}}, g_{i_{r_2}}, \ldots, g_{i_{r_{j-1}}}\}$ for each j. As a consequence of this and the fact that $\{g_{i_{r_1}}, g_{i_{r_2}}, \ldots, g_{i_{r_\ell}}\}$ are all independent uniformly distributed elements of G, it follows that $\{h_{i_{r_1}}, h_{i_{r_2}}, \ldots, h_{i_{r_\ell}}\}$ are all independent uniformly distributed elements of G. Let J' denote the set of k group elements obtained from J by replacing the subset $\{g_{i_{r_1}}, g_{i_{r_2}}, \ldots, g_{i_{r_\ell}}\}$ with $\{h_{i_{r_1}}, h_{i_{r_2}}, \ldots, h_{i_{r_\ell}}\}$. Clearly, J' is a set of k independent, uniformly distributed random group elements from G.

Thus, we have

$$g_{i_1}^{\epsilon_1} \ldots g_{i_m}^{\epsilon_m} = h_{i_{r_1}}^{\epsilon_{r_1}} \ldots h_{i_{r_\ell}}^{\epsilon_{r_\ell}} x(\bar{\epsilon}),$$

where $x(\bar{\epsilon}) = x_1 x_2 \ldots x_r$ is an element in G that depends on J, I and $\bar{\epsilon}$, where $\bar{\epsilon}$ consists of all the ϵ_j for $i_j \in I \setminus L$. Hence, for each $g \in G$, observe that we can write

$$\begin{aligned} R_I^J(g) \quad &= \quad \mathrm{Prob}_{\epsilon_1, \ldots, \epsilon_m}[\prod_{j=1}^{m} g_{i_j}^{\epsilon_j} = g] \\ &= \quad \mathrm{Prob}_{\epsilon_1, \ldots, \epsilon_m}[h_{i_{r_1}}^{\epsilon_{r_1}} \ldots h_{i_{r_\ell}}^{\epsilon_{r_\ell}} = g x(\bar{\epsilon})^{-1}] \\ &= \quad E_{\bar{\epsilon}}[R_{L(I)}^{J'}(g x(\bar{\epsilon})^{-1})]. \end{aligned}$$

Therefore we have the following:

$$\begin{aligned}
\mathbb{E}_J[\mathrm{Coll}(R_I^J)] &= \mathbb{E}_J[\sum_g (R_I^J(g))^2] \\
&= \mathbb{E}_J[\sum_g (\mathbb{E}_{\bar{\epsilon}} R_{L(I)}^J (gx(\bar{\epsilon})^{-1}))^2] \\
&\leq \mathbb{E}_J[\sum_g \mathbb{E}_{\bar{\epsilon}} (R_{L(I)}^J (gx(\bar{\epsilon})^{-1}))^2] \qquad (15) \\
&= \mathbb{E}_{\bar{\epsilon}}[\mathbb{E}_J[\sum_g (R_{L(I)}^J (gx(\bar{\epsilon})^{-1}))^2]] \\
&= \mathbb{E}_{\bar{\epsilon}}[\mathbb{E}_J[\sum_h (R_{L(I)}^J (h))^2]] \\
&= \mathbb{E}_J[\sum_h (R_{L(I)}^J (h))^2] \\
&= \mathbb{E}_J[\mathrm{Coll}(R_{L(I)}^J)] \leq \frac{1}{n} + \delta \qquad (16)
\end{aligned}$$

where the inequality in 15 follows from Cauchy-Schwarz inequality and the last step follows from the assumption of the lemma. □□

We use simple counting argument to prove Lemma 2.3. A similar lemma appears in [7].

Proof of Lemma 2.3

Proof: Consider the event that a sequence X of length m does not have an L-subsequence of length ℓ. Thus it has at most $\ell - 1$ distinct elements, which can be chosen in at most $\binom{k}{\ell-1}$ ways. The m length sequence can be formed from them in at most $[\ell-1]^m$ ways. Therefore

$$\begin{aligned}
\mathsf{Pr}[X \text{ has L-subsequence of length } < \ell] &\leq \frac{\binom{k}{\ell-1}[\ell-1]^m}{k^m} \\
&\leq \left(\frac{ke}{\ell-1}\right)^{\ell-1} \cdot \left(\frac{\ell-1}{k}\right)^m \\
&= e^{\ell-1}\left(\frac{\ell-1}{k}\right)^{m-\ell+1} \\
&= \frac{e^{\ell-1}}{a^{m-(k/a)}} = \frac{(ae)^{k/a}}{a^m}.
\end{aligned}$$

□□

Next we prove Lemma 2.4.

Proof of Lemma 2.4

Proof:

We call $I \in [k]^m$ *good* if it has an L-subsequence of length at least ℓ, else we call it *bad*.

$$
\begin{aligned}
\mathbb{E}_J[\mathrm{Coll}(Q_J)] &= \mathbb{E}_J[\sum_{g\in G} Q_J^2(g)] \\
&= \mathbb{E}_J[\sum_{g\in G} (\mathbb{E}_I(R_I(g))^2] \\
&\leq \mathbb{E}_J[\sum_{g\in G} \mathbb{E}_I(R_I^2(g))] \text{ By Cauchy-Schwarz inequality} \qquad (17) \\
&= \mathbb{E}_I[\mathbb{E}_J[\mathrm{Coll}(R_I)]] \\
&\leq \frac{1}{k^m}\mathbb{E}_J[\sum_{\substack{I\in[k]^m \\ I \text{ is good}}} \sum_{g\in G} (R_L^J(g))^2 + \sum_{\substack{I\in[k]^m \\ I \text{ is bad}}} 1] \\
&\leq \mathsf{Pr}_I[I \text{ is good}]\left(\frac{1}{n}+\frac{1}{2^\ell}\right) + \mathsf{Pr}_I[I \text{ is bad}] \qquad (18)
\end{aligned}
$$

Here the last step follows from Lemma 2.2 and Theorem 1.4. Now we fix m from Lemma 2.3 appropriately to $O(\log n)$ such that $\mathsf{Pr}_I[I \text{ is bad}] \leq \frac{1}{2^m}$ and choose $\ell = \Theta(m)$. Hence we get that $\mathbb{E}_J[\mathrm{Coll}(Q_J)] \leq \frac{1}{n} + \frac{1}{2^{\Theta(m)}}$. □□

Next, we give the proof of Lemma 2.6

6 Proof of Lemma 2.6

Proof: For each $g \in G$, we can write

$$
\begin{aligned}
R_I^J(g) &= \mathrm{Prob}_{\epsilon_1,\ldots,\epsilon_m}[\prod_{j=1}^m g_{i_j}^{\epsilon_j} = g] = \mathrm{Prob}_{\epsilon_1,\ldots,\epsilon_m}[g_{i_{f_1}}^{\epsilon_{f_1}} \ldots g_{i_{f_r}}^{\epsilon_{f_r}} h_{e_1}^{\epsilon_{e_1}} \ldots h_{e_\ell}^{\epsilon_{e_\ell}} = gy(\bar{\epsilon})^{-1}] \\
&= E_{\bar{\epsilon}}[R_{\widehat{I}}^{J(I)}(gy(\bar{\epsilon})^{-1})].
\end{aligned}
$$

Therefore we have the following:

$$
\begin{aligned}
\mathbb{E}_J[\mathrm{Coll}(R_I^J)] &= \mathbb{E}_J[\sum_g (R_I^J(g))^2] \\
&= \mathbb{E}_J[\sum_g (\mathbb{E}_{\bar{\epsilon}} R_{\widehat{I}}^{J(I)}(gy(\bar{\epsilon})^{-1}))^2] \\
&\leq \mathbb{E}_J[\sum_g \mathbb{E}_{\bar{\epsilon}} (R_{\widehat{I}}^{J(I)}(gy(\bar{\epsilon})^{-1}))^2] \qquad (19) \\
&= \mathbb{E}_{\bar{\epsilon}}[\mathbb{E}_J[\sum_g (R_{\widehat{I}}^{J(I)}(gy(\bar{\epsilon})^{-1}))^2]] \\
&= \mathbb{E}_{\bar{\epsilon}}[\mathbb{E}_J[\sum_h (R_{\widehat{I}}^{J(I)}(h))^2]] \\
&= \mathbb{E}_J[\mathrm{Coll}(R_{\widehat{I}}^{J(I)})],
\end{aligned}
$$

where the inequality 19 follows from Cauchy-Schwarz inequality. □□

We include a short proof of Lemma 2.9.

Proof of Lemma 2.9

Proof: There are $\binom{m}{r}$ ways of picking r positions for the fixed elements in I. Each such index can be chosen in j ways. From the $(k-j)$ random elements of J, ℓ distinct elements can be picked in $\binom{k-j}{\ell}$ ways. Let $n_{m-r,\ell}$ be the number of sequences of length $m-r$ that can be constructed out of ℓ distinct integers such that every integer appears at least once. Clearly, $|S_{r,\ell}| = \binom{m}{r} j^r \binom{k-j}{\ell} n_{m-r,\ell}$. It is well known that $n_{m-r,\ell}$ is the coefficient of $x^{m-r}/(m-r)!$ in $(e^x-1)^\ell$. Thus, by the binomial theorem $n_{m-r,\ell} = \sum_{i=0}^{\ell}(-1)^i\binom{\ell}{i}(\ell-i)^{m-r}$. Since $m = O(\log n)$ and $\ell \le m$, $n_{m-r,\ell}$ can be computed in time polynomial in n. □□

Next, we give a proof of Lemma 3.1.

Proof of Lemma 3.1

Proof: The proof closely follows the proof of Erdös-Rényi for the case $\bar{\epsilon} \in \{0,1\}^k$. We briefly sketch the argument below for the sake of completeness.

We denote the expression $g_1^{\epsilon_1}\dots g_k^{\epsilon_k}$ by $\bar{g}^{\bar{\epsilon}}$. For a given J, $\chi_x(\bar{\epsilon}) = 1$ if $\bar{g}^{\bar{\epsilon}} = x$ and 0 otherwise. Let $S_1 = \{(\bar{\epsilon},\bar{\epsilon}')|\bar{\epsilon} \neq \bar{\epsilon}'; \exists i \text{ such that } \bar{\epsilon}_i \neq \bar{\epsilon}'_i \text{ and } \bar{\epsilon}_i\bar{\epsilon}'_i = 0\}$. Let $S_2 = \{(\bar{\epsilon},\bar{\epsilon}')|\bar{\epsilon} \neq \bar{\epsilon}'; \bar{\epsilon}_i \neq \bar{\epsilon}'_i \Rightarrow \bar{\epsilon}_i\bar{\epsilon}'_i = -1\}$

$$\begin{aligned}
\mathbb{E}_J[\mathrm{Coll}(D_J)] &= \mathbb{E}_J[\sum_{x\in G}(D_J(x))^2] \\
&= \mathbb{E}_J\left[\sum_{x\in G}\left(\mathsf{Pr}_{\bar{\epsilon}}[\bar{g}^{\bar{\epsilon}} = x]\right)^2\right] \\
&= \frac{1}{3^{2k}}\mathbb{E}_J\left[\sum_{x\in G}\left(\sum_{\bar{\epsilon}}\chi_x(\bar{\epsilon})\right)\left(\sum_{\bar{\epsilon}'}\chi_x(\bar{\epsilon}')\right)\right] \\
&= \frac{1}{3^{2k}}\left[\sum_{\bar{\epsilon}=\bar{\epsilon}'}\mathbb{E}_J\left[\sum_{x\in G}\chi_x(\bar{\epsilon})\chi_x(\bar{\epsilon}')\right] + \sum_{\bar{\epsilon}\neq\bar{\epsilon}'}\mathbb{E}_J\left[\sum_{x\in G}\chi_x(\bar{\epsilon})\chi_x(\bar{\epsilon}')\right]\right] \\
&= \frac{1}{3^{2k}}\left(3^k + \sum_{(\bar{\epsilon},\bar{\epsilon}')\in S_1}\mathbb{E}_J\left[\sum_{x\in G}\chi_x(\bar{\epsilon})\chi_x(\bar{\epsilon}')\right] + \sum_{(\bar{\epsilon},\bar{\epsilon}')\in S_2}\mathbb{E}_J\left[\sum_{x\in G}\chi_x(\bar{\epsilon})\chi_x(\bar{\epsilon}')\right]\right) \\
&= \frac{1}{3^{2k}}\left[3^k + \sum_{(\bar{\epsilon},\bar{\epsilon}')\in S_1}\mathsf{Pr}_{\bar{g}}(\bar{g}^{\bar{\epsilon}} = \bar{g}^{\bar{\epsilon}'}) + \sum_{(\bar{\epsilon},\bar{\epsilon}')\in S_2}\mathsf{Pr}_{\bar{g}}(\bar{g}^{\bar{\epsilon}} = \bar{g}^{\bar{\epsilon}'})\right] \\
&\le \frac{1}{3^k} + \left(1 - \frac{1}{3^k} - \frac{5^k}{9^k}\right)\frac{1}{n} + \frac{5^k}{9^k} \\
&= \left(1 - \frac{1}{n}\right)\left(\frac{1}{3^k} + \frac{5^k}{9^k}\right) + \frac{1}{n} \\
&< \left(\frac{8}{9}\right)^k + \frac{1}{n}
\end{aligned}$$

To see the last step, first notice that if $\bar{\epsilon} = \bar{\epsilon}'$ then $\sum_{x\in G}\chi_x(\bar{\epsilon})\chi_x(\bar{\epsilon}') = 1$. A simple counting argument shows that $|S_2| = \sum_{i=0}^{k}\binom{k}{i}2^i3^{k-i} = 5^k$. So $\sum_{(\bar{\epsilon},\bar{\epsilon}')\in S_2}\mathsf{Pr}_{\bar{g}}(\bar{g}^{\bar{\epsilon}} = \bar{g}^{\bar{\epsilon}'}) \le 5^k$. Now consider

$(\bar{\epsilon}, \bar{\epsilon}') \in S_1$, let j be the first position from left such that $\bar{\epsilon}_j \neq \bar{\epsilon}'_j$. W.l.o.g assume that $\bar{\epsilon}_j = 1$ (or $\bar{\epsilon}_j = -1$) and $\bar{\epsilon}'_j = 0$. In that case write $\bar{g}^{\bar{\epsilon}} = a g_j b$ and $\bar{g}^{\bar{\epsilon}'} = ab'$. Then $\mathsf{Pr}_{g_j}[g_j = b' b^{-1}] = \frac{1}{n}$. Hence $\sum_{(\bar{\epsilon}, \bar{\epsilon}') \in S_1} \mathsf{Pr}_{\bar{g}}(\bar{g}^{\bar{\epsilon}} = \bar{g}^{\bar{\epsilon}'}) = \frac{9^k - 3^k - 5^k}{n}$. □□

ECCC ISSN 1433-8092

http://eccc.hpi-web.de