Cutoff for Ramanujan graphs via degree inflation

Jonathan Hermon*

Abstract

Recently Lubetzky and Peres showed that simple random walks on a sequence of d-regular Ramanujan graphs $G_n = (V_n, E_n)$ of increasing sizes exhibit cutoff in total variation around the diameter lower bound $\frac{d}{d-2} \log_{d-1} |V_n|$. We provide a different argument under the assumption that for some $r(n) \gg 1$ the maximal number of simple cycles in a ball of radius $r(n)$ in G_n is uniformly bounded in n.

1 Introduction

Generically, we denote the stationary distribution of an ergodic Markov chain $(X_t)_{t \geq 0}$ by π, its state space by Ω and its transition matrix by P. We denote by P_x^t (resp. P_x) the distribution of X_t (resp. $(X_t)_{t \geq 0}$), given that the initial state is x. The total variation distance of two distributions on Ω is $\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2} \sum_y |\mu(y) - \nu(y)|$. The total variation ε-mixing time is $t_{\mathrm{mix}}(\varepsilon) := \inf\{t : \max_x \|\mathrm{P}_x^t - \pi\|_{\mathrm{TV}} \leq \varepsilon\}$. Next, consider a sequence of chains, $((\Omega_n, P_n, \pi_n))_{n \in \mathbb{N}}$, each with its mixing time $t_{\mathrm{mix}}^{(n)}(\cdot)$. We say that the sequence exhibits a cutoff if the following sharp transition in its convergence to stationarity occurs:

$$\forall \varepsilon \in (0, 1/2], \quad \lim_{n \to \infty} t_{\mathrm{mix}}^{(n)}(\varepsilon) / t_{\mathrm{mix}}^{(n)}(1 - \varepsilon) = 1. \tag{1.1}$$

A family of d-regular graphs G_n with $d \geq 3$ is called an *expander family*, if the second largest eigenvalues of the corresponding adjacency matrices are uniformly bounded away from d. Lubotzky, Phillips, and Sarnak [6] defined a connected finite d-regular graph G with $d \geq 3$ to be **Ramanujan** if the eigenvalues of the transition matrix of simple random walk (SRW) on G all lie in $\{\pm 1\} \cup [-\rho_d, \rho_d]$, where $\rho_d := \frac{2\sqrt{d-1}}{d}$ is the spectral radius of SRW on the infinite d-regular tree $\mathbb{T}_d$. Lubotzky, Phillips, and Sarnak [6], Margulis [8] and Morgenstern [9] constructed d-regular Ramanujan graphs for all d of the form $d = p^m + 1$, where p is a prime number. Recently, Marcus, Spielman and Srivastava [7] proved the existence of bipartite d-regular Ramanujan graphs for all $d \geq 3$. In light of the Alon-Boppana bound [10], Ramanujan graphs are "optimal expanders" as they have asymptotically the largest spectral-gap.

Let $G_n = (V_n, E_n)$ be a sequence of finite connected d_n-regular graphs. Let P_n be the transition matrix of SRW on G_n. Denote the eigenvalues of P_n by $1 = \lambda_1(n) > \lambda_2(n) \geq$

*Faculty of mathematics and computer science, Weizmann Institute of Science, Rehovot, Israel.
E-mail: jonathah@weizmann.ac.il.

$\cdots \geq \lambda_{|V_n|}(n) \geq -1$. We say that the sequence is asymptotically Ramanujan if $|V_n| \to \infty$ and

$$\max\{|\lambda_i(n)| : |\lambda_i(n)| \neq 1\} \leq \rho_{d_n}^{1-o(1)}.$$

We say that the sequence is asymptotically one-sided Ramanujan if $|V_n| \to \infty$, $\lambda_2(n) \leq \rho_{d_n}^{1-o(1)}$ and $\liminf_{n\to\infty} \min\{\lambda_i(n) : \lambda_i(n) \neq -1\} > -1$. Friedman [3] showed that a sequence of d-regular random graphs of increasing sizes is w.h.p. asymptotically Ramanujan.

Remark 1.1. Our definition of asymptotically Ramanujan graphs is not the standard one. The more standard definition is that $\max\{|\lambda_i(n)| : |\lambda_i(n)| \neq 1\} \leq \rho_{d_n} + o(1)$.

It is elementary to show that for every n-vertex d-regular graph, the $1 - \varepsilon$ total variation mixing time for the SRW is at least $t_{d,\varepsilon,n} := \frac{d}{d-2} \log_{d-1} n - C\sqrt{n|\log \varepsilon|/d}$, for some constant $C > 0$.[1] The following precise formulation of this fact is due to Lubeztky and Peres [4].

Lemma 1.2 (Trivial diameter lower bound - c.f. [4] (2.2)–(2.3) pg. 9)**.** *Let $G = (V, E)$ be an n-vertex d-regular graph with $d \geq 3$. Let $c_d := \frac{2\sqrt{d(d-1)}}{(d-2)^{3/2}}$ and Φ^{-1} be the inverse function of the CDF of the standard Normal distribution. Then SRW on G satisfies*

$$\forall \varepsilon \in (0,1), \quad t_{\mathrm{mix}}(1 - \varepsilon - o(1)) \geq \frac{d}{d-2} \log_{d-1} n + c_d \Phi^{-1}(\varepsilon)\sqrt{\log_{d-1} n}.$$

Recently, Lubetzky and Peres [4] showed that simple random walks on a sequence of non-bipartite d_n-regular Ramanujan graphs $G_n = (V_n, E_n)$ of increasing sizes exhibit cutoff around the diameter lower bound $\frac{d_n}{d_n-2} \log_{d_n-1} |V_n|$. In this work we present an alternative argument and prove the same result under the following assumption:

Assumption 1. There exists a diverging sequence r_n such that the maximal number of simple cycles in a ball of radius r_n in G_n is uniformly bounded in n.

Theorem 1.3. *Let $G_n = (V_n, E_n)$ be a sequence of non-bipartite, finite, connected, d_n-regular asymptotically one-sided Ramanujan graphs.*

  (i) If $d_n = d$ for all n and Assumption 1 holds then the corresponding sequence of simple random walks exhibits cutoff around time $\frac{d}{d-2} \log_{d-1} |V_n|$.

  (ii) If d_n diverges and $\log d_n = o(\log_{d_n} |V_n|)$ then the corresponding sequence of simple random walks exhibits cutoff around time $\log_{d_n} |V_n|$.

Remark 1.4. If there is no cutoff, then cutoff must fail on some subsequence (n_k) such that either $\lim_{k\to\infty} d_{n_k} = \infty$ or $d_{n_k} = d$ for all k for some fixed $d \geq 3$. Thus there is no loss of generality in assuming that either $\lim_{n\to\infty} d_n = \infty$ or $d_n = d$ for all n.

Assumption 1 is rather mild as it is quite difficult to construct a family of asymptotically one-sided Ramanujan graphs violating this assumption. In particular, it is satisfied w.h.p. by a sequence of random d-regular graphs of increasing sizes [5]. It follows from [1, Theorem 1] that if G_n is a sequence of d-regular transitive asymptotically Ramanujan graphs of increasing sizes then $\lim_{n\to\infty} \mathrm{girth}(G_n) = \infty$, where for a graph G, $\mathrm{girth}(G)$ denotes its girth[2] (and so Assumption 1 holds).

The argument of Lubetzky and Peres [4] does not require Assumption 1 (nor the assumption $\log d_n = o(\log_{d_n} |V_n|)$). They studied the Jordan decomposition of the transition matrix of the non-backtracking walk[3] and used it to derive cutoff for the non-backtracking

---

[1]This can be derived from the fact that C can be chosen so that the probability that the probability that the distance of the walk at time $t_{d,\varepsilon,n}$ from its starting point is at least $\lfloor \log_{d-1}(\frac{1}{4}\varepsilon n) \rfloor$ with probability at most $\frac{\varepsilon}{2}$ (together with the fact that a ball of radius $\lfloor \log_{d-1}(\frac{1}{4}\varepsilon n) \rfloor$ contains at most $\frac{1}{2}\varepsilon n$ vertices).

[2]The girth of a graph G is the length of the shortest cycle in G.

[3]This is a random walk on the directed edges of the graph, with transition matrix $P_{\mathrm{NB}}((x,y)(z,w)) = \frac{1_{z=y, w\neq x}}{\deg(y)-1}$.

walk, which for a regular graph implies cutoff also for the SRW. In this note we study the SRW by looking at it only when it crosses distance k from its previous position, for some large k.

1.1 Organization of this note

In § 2, as a warm up, we present an extremely simple and short proof for the occurrence of cutoff for SRW on a sequence of asymptotically Ramanujan graphs of diverging degree. In § 3 we present some machinery for bounding mixing times using hitting times. We then apply this machinery to prove Part (ii) of Theorem 1.3. In § 4 we give an overview of the proof of Part (i) of Theorem 1.3. In § 5 we prove two auxiliary results. Finally, in § 6 we conclude the proof of Theorem 1.3.

2 A warm up

It turns out that for a sequence of asymptotically Ramanujan graphs of diverging degree the trivial diameter lower bound (of Lemma 1.2) is matched by the trivial spectral-gap upper bound on the L_2 mixing time obtained via the Poincaré inequality. As a warm up and motivation for what comes we now prove the following theorem.

Theorem 2.1. *Let $G_n = (V_n, E_n)$ be a sequence of non-bipartite, finite, connected, d_n-regular asymptotically Ramanujan graphs with $d_n \to \infty$. Then the corresponding sequence of simple random walks exhibits cutoff around time $\log_{d_n} |V_n|$.*

Note that in Part (ii) of Theorem 1.3 the graphs are assumed to be only asymptotically one-sided Ramanujan. Before proving Theorem 2.1 we need a few basic definitions and facts. Let

$$\lambda := \max\{|a| : a \neq 1, a \text{ is an eigenvalue of } P\} \quad \text{and} \quad t_{\text{rel}} := \frac{1}{1 - \lambda}.$$

The L_2 distance of P_x^t from π is defined as

$$\|P_x^t - \pi\|_{2,\pi}^2 = \sum_y \pi(y)(P^t(x,y)/\pi(y))^2 - 1.$$

By Jensen's and the Poincaré inequalities, for all t and x we have that

$$4\|P_x^t - \pi\|_{\text{TV}}^2 \leq \|P_x^t - \pi\|_{2,\pi}^2 \leq \lambda^{2t}\|P_x^0 - \pi\|_{2,\pi}^2 \leq \lambda^{2t}/\pi(x).$$

Hence for SRW on an n-vertex regular graph we have for all t and x that

$$4\|P_x^t - \pi\|_{\text{TV}}^2 \leq n\lambda^{2t} \implies t_{\text{mix}}(\varepsilon) \leq \frac{1}{2}\log_{\frac{1}{\lambda}}(n\varepsilon^{-2}). \tag{2.1}$$

Proof of Theorem 2.1: By assumption $\lambda = \rho_{d_n}^{1-o(1)} = d_n^{-\frac{1}{2}(1-o(1))}$. Thus $\frac{1}{2}\log_{\frac{1}{\lambda}}|V_n| = (1 + o(1))\log_{d_n}|V_n|$. The proof is concluded by combining (2.1) with Lemma 1.2. □

3 Replacing the Poincaré inequality by its hitting time analog

In the proof of Theorem 1.3 we exploit the general connection between mixing times and escape times from small sets, established in [2] (Corollary 3.1 eq. (3.2)): There exists some absolute constant $C > 0$ such that for every reversible chain (with a finite state space),

$$\forall \alpha, \varepsilon \in (0,1), \quad t_{\text{mix}}(\varepsilon + \alpha) \leq \text{hit}_{1-\alpha}(\varepsilon) + Ct_{\text{rel}}\log(1/\alpha), \tag{3.1}$$

where $\text{hit}_{1-\alpha}(\varepsilon) := \inf\{t : \max_{x,A:\pi(A)\leq\alpha} P_x[T_{A^c} > t] \leq \varepsilon\}$ and $T_B := \inf\{t : X_t \in B\}$ is the hitting time of the set B. In the proof of Theorem 1.3 we replace the naive L_2 bound

used in the proof of Theorem 2.1 by its hitting time counterpart: Under reversibility, for all $A \subsetneq \Omega$, $a \in A$ and $t \geq 0$

$$\pi_A(a)(\mathrm{P}_a[T_{A^c} > t])^2 \leq \sum_{b \in A} \pi_A(b)(\mathrm{P}_b[T_{A^c} > t])^2 = \|P_A^t 1_A\|_{2,A}^2 \leq [\lambda(A)]^{2t}, \qquad (3.2)$$

where π_A is π conditioned on A, P_A is the restriction of the transition matrix P to A (this is the transition matrix of the chain which is "killed" upon escaping A), $\|f\|_{2,A}^2 := \sum_{b \in A} \pi_A(b) f^2(b)$ for $f \in \mathbb{R}^A$ and $\lambda(A)$ is the largest eigenvalue of P_A.

The following proposition relates $\lambda(A)$ to λ_2, the second largest eigenvalue of P.

Proposition 3.1 (e.g. [2] Lemma 3.8). *For every reversible Markov chain and any set A,*

$$\lambda(A) \leq \lambda_2 + \pi(A), \qquad (3.3)$$

Similarly to (2.1), by (3.1)–(3.3) we have for every reversible chain on a finite state space with $\lambda_2 < 1/2$ and every $\alpha \in (0, \lambda_2]$ that

$$\begin{aligned} \mathrm{hit}_{1-\alpha}(\sqrt{\alpha}) &\leq \frac{1}{2} |\log_{\frac{1}{2\lambda_2}}(\min_v \pi(v))|, \\ t_{\mathrm{mix}}(2\sqrt{\alpha}) &\leq \frac{1}{2} |\log_{\frac{1}{2\lambda_2}}(\min_v \pi(v))| + C t_{\mathrm{rel}} \log(1/\alpha). \end{aligned} \qquad (3.4)$$

We are now in a position to give a short proof for Part (ii) of Theorem 1.3.

Proof. Let $G_n = (V_n, E_n)$ be a sequence of non-bipartite, finite, connected, d_n-regular asymptotically one-sided Ramanujan graphs. Assume that d_n diverges and $\log d_n = o(\log_{d_n} |V_n|)$. Let $\alpha = \alpha_n = d_n^{-1/2} = o(1)$. Let $\lambda_2 = \lambda_2(n)$ be the second largest eigenvalue of the transition matrix of SRW on G_n. By our assumptions $2\lambda_2 = d_n^{-\frac{1}{2}+o(1)}$ and so by (3.4) we have that

$$t_{\mathrm{mix}}(2\sqrt{\alpha}) \leq \frac{1}{2} \log_{\frac{1}{2\lambda_2}} |V_n| + C' \log(1/\alpha) = (1 + o(1)) \log_{d_n} |V_n|.$$

The proof is concluded using Lemma 1.2. $\qquad \square$

4 Degree inflation

The simple proof of Part (ii) of Theorem 1.3 motivates looking at the following graph.

Definition 4.1. *Given a graph $G = (V, E)$, we define $G(k) = (V, E(k))$ via*

$$E(k) := \{\{u, v\} : \mathrm{dist}_G(u, v) = k, \ u, v \in V\},$$

where $\mathrm{dist}_G(u, v)$ denotes the graph distance of u and v w.r.t. G. Denote the transition matrix of SRW on $G(k)$ by K.

Definition 4.2. *Consider SRW on G, $(X_t)_{t=0}^\infty$. Let $T_0 := 0$ and inductively set $T_{i+1} := \inf\{t \geq T_i : \mathrm{dist}_G(X_{T_{i+1}}, X_{T_i}) = k\}$. Consider the chain $\mathbf{Y} := (Y_j)_{j=0}^\infty$ defined via $Y_i := X_{T_i}$ for all i, and denote its transition matrix by W.*

Remark 4.3. It is possible that $G(k) := (V, E(k))$ is not connected. This could be rectified, say by connecting every vertex to its entire k-neighborhood. However, below we only use the fact that the SRW on $G(k)$ is reversible w.r.t. $\pi_{G(k)}(x) := \deg_{G(k)}(x)/(2|E(k)|)$.

Let $G = (V, E)$ be a d-regular finite Ramanujan graph. Assume that Assumption 1 holds. Let $r = r_n$ be as in Assumption 1. Fix some $k = k_n$ such that $1 \ll k \ll \sqrt{r}$.

Remark 4.4. Let K, W and T_i be as in Definitions 4.1 and 4.2. By Assumption 1, for every $x, y \in V$ of distance k from one another $1 \leq K(x, y) d(d-1)^{k-1} \leq C_1(d)$. In Lemma 5.2 we show that for such x, y also $1 \leq W(x, y) d(d-1)^{k-1} \leq C_2(d)$. In fact, Assumption 1 could have been replaced by the assumption that $\max\{W(x, y), K(x, y)\} \leq (d-1)^{-k(1-o(1))}$ and that T_1 is concentrated around $\frac{dk}{d-2}$ (uniformly for all initial states).

4.1 An overview of the proof of Part (i) of Theorem 1.3

Let G, k and r be as above. Intuitively, if either the SRW on $G(k)$ or the chain $\mathbf{Y}$ (from Definitions 4.1 and 4.2) exhibit an abrupt convergence to stationarity around time $t = t_n$, then also the SRW on G should exhibit an abrupt convergence to stationarity around time $t \cdot \frac{d}{d-2}k$. The term $\frac{d}{d-2}k$ comes from the fact that (by Assumption 1) the expected time it takes the walk on G to get within distance k from its current position is $\frac{d}{d-2}k(1 + o(1))$.

While the chain $\mathbf{Y}$ is more directly related to the SRW on G, it is harder to analyze it directly since it need not be reversible and a-priori it is not clear that its stationary distribution is close to the uniform distribution. Instead we analyze the walk on $G(k)$ and use it to learn about $\mathbf{Y}$ and then in turn about the walk on G.

In light of Part (ii) of Theorem 1.3 (which has already been proven) a natural strategy for proving Part (i) of Theorem 1.3 is to show that $\lambda_2(K) = \rho_D^{1-o(1)} = (d-1)^{-\frac{k}{2}(1-o(1))}$, where D is the maximal degree in $G(k)$, K is the transition matrix of SRW on $G(k)$ and $\lambda_2(K)$ is its second largest eigenvalue. Unfortunately, we do not know how to show this (see the first paragraph of § 5). Instead, we obtain such an estimate for $\lambda_K(A)$, the largest eigenvalue of K_A, the restriction of K to A, for any "small" set A. By small we mean that its stationary probability is at most $\alpha := (d-1)^{-3k^2}$. Indeed, the key to the proof of Part (i) of Theorem 1.3 is to show that $\lambda_K(A) \le (d-1)^{-\frac{k}{2}(1-o(1))}$ for every small set A. Using (3.2) we get for the walk on $G(k)$ that $\mathrm{P}_a[T_{A^c} > (1 + o(1))\frac{1}{k}\log_{d-1}|V|] = (d-1)^{-\frac{k}{2}(1-o(1))}$. We then show that the same holds for $\mathbf{Y}$ (this is obvious when $2k < \mathrm{girth}(G)$; The general case is derived using the fact that, as mentioned in Remark 4.4, $cW(x,y) \le K(x,y) \le CW(x,y)$ for all x,y). Finally, using an obvious coupling between $\mathbf{Y}$ and the SRW on G, after multiplying by $\frac{d}{d-2}k(1 + o(1))$ the last bound is transformed into a bound on $\mathrm{hit}_{1-\alpha}(o(1))$ for SRW on G (for some $o(1)$ terms).

5 Auxiliary results

In order to control $\lambda_K(A)$ (for small A), apart from Proposition 3.1 we need the following comparison result. While there are similar comparison techniques for the spectral-gap, we are not aware of a comparison technique which allows one to argue that λ_2 (the second largest eigenvalue of the transition matrix) of one chain is close to 0 (say, that $\lambda_2 = o(1)$) if that of another chain is close to 0.

Proposition 5.1. *Let $P^{(1)}$ and $P^{(2)}$ be two transition matrices on the same finite state space Ω, both reversible w.r.t. $\pi^{(1)}$ and $\pi^{(2)}$, respectively. Assume that $P^{(1)}(x,y) \le C_1 P^{(2)}(x,y)$ and $1/C_2 \le \pi^{(1)}(x)/\pi^{(2)}(x) \le C_2$ for all x, y. Let $A \subsetneq \Omega$ and let $\lambda_{P^{(i)}}(A)$ be the largest eigenvalue of $P_A^{(i)}$, the restriction of $P^{(i)}$ to A ($i = 1, 2$). Then*

$$\lambda_{P^{(1)}}(A) \le C_1 C_2^2 \lambda_{P^{(2)}}(A).$$

Proof: Denote $\langle f, g \rangle_{\pi_A^{(i)}} := \sum_{x \in A} \pi_A^{(i)}(x)g(x)f(x)$. By the Perron-Frobenius Theorem

$$\lambda_{P^{(1)}}(A) = \max_{f \in \mathbb{R}_+^A, f \ne 0} \frac{\langle P_A^{(1)}f, f \rangle_{\pi_A^{(1)}}}{\langle f, f \rangle_{\pi_A^{(1)}}} \le C_1 C_2^2 \max_{f \in \mathbb{R}_+^A, f \ne 0} \frac{\langle P_A^{(2)}f, f \rangle_{\pi_A^{(2)}}}{\langle f, f \rangle_{\pi_A^{(2)}}} = C_1 C_2^2 \lambda_{P^{(2)}}(A). \quad \square$$

Before proving Theorem 1.3 we need one more lemma.

Lemma 5.2. *Let $G = (V, E)$ be a d-regular graph ($d \ge 3$). Let $v \in V$. For $i, k \in \mathbb{N}$ let $D_i := \{u \in V : \mathrm{dist}_G(u, v) = i\}$, $B_i := \cup_{j=0}^{i} D_j$ (the ball of radius i around v) and*

$$\mathrm{t}(B_k) := |\{\{x, y\} \in E : y \in B_{k-1}, x \in B_k\}| - |B_k|.$$

For any $s \geq 0$ there exist some constant $C(s,d) > 0$ and k_s such that if $k \geq k_s$, $\mathrm{t}(B_k) \leq s$ and $D_k \neq \varnothing$ then

$$\frac{1}{d(d-1)^{k-1}} \leq \min_{u \in D_k} \mathrm{P}_v[T_{D_k} = T_u] \leq \max_{u \in D_k} \mathrm{P}_v[T_{D_k} = T_u] \leq \frac{C(s,d)}{d(d-1)^{k-1}}. \tag{5.1}$$

Proof. Let $u \in D_k$. We first prove that $\mathrm{P}_v[T_{D_k} = T_u] \geq \frac{1}{d(d-1)^{k-1}}$. This follows from a standard argument involving the covering tree of G. A non-backtracking path of length ℓ is a sequence of vertices $(v_0, v_1, \ldots, v_\ell)$ such that $\{v_i, v_{i-1}\} \in E$ and $v_{i+2} \neq v_i$ for all i. Let $\mathcal{P}_\ell$ be the collection of all non-backing paths of length ℓ starting from v. Let $\mathbb{T}_d$ be the (infinite) d-regular tree. We may label the ℓth level of $\mathbb{T}_d$ by the set $\mathcal{P}_\ell$ (in a bijective manner) such that the children of $(v, v_1, \ldots, v_\ell)$ are $\{(v, v_1, \ldots, v_\ell, v') : (v, v_1, \ldots, v_\ell, v') \in \mathcal{P}_{\ell+1}\}$. For $\gamma = (v, v_1, \ldots, v_\ell)$ let $\phi(\gamma) := v_\ell$. Note that if $(S_n)_{n=0}^\infty$ is a SRW on $\mathbb{T}_d$ (labeled as above) started from (v) (which is the root) then $(\phi(S_n))_{n=0}^\infty$ is a SRW on G started from v. Denote the law of $(S_n)_{n=0}^\infty$ by $\mathbb{P}_v$.

Fix some $\gamma := (v, v_1, \ldots, v_k) \in \mathcal{P}_k$ such that $v_k = u$. Finally, observe that

$$\mathrm{P}_v[T_{D_k} = T_u] \geq \mathbb{P}_v[T_{\mathcal{P}_k} = T_\gamma] = \frac{1}{|\mathcal{P}_k|} = \frac{1}{d(d-1)^{k-1}}.$$

We now prove that $\mathrm{P}_v[T_{D_k} = T_u] \leq \frac{C(s,d)}{d(d-1)^{k-1}}$. We prove this by induction on s. The base case $\mathrm{t}(B_k) = 0$ is trivial (it holds with $C(1,d) = 1$). Now consider the case that $\mathrm{t}(B_k) = s > 0$. Let $z \in D_k$ be such that $\mathrm{P}_v[T_{D_k} = T_z] = \max_{u \in D_k} \mathrm{P}_v[T_{D_k} = T_u]$. For an edge $e := \{x, y\} \in E$ let $G_e := (V, E \setminus \{e\})$ be the graph obtained by deleting e from G. Let $H_e := (V_e, E_e)$ be the graph obtained from G_e by connecting x (resp. y) to the root of a d-ary tree[4] $\mathcal{T}_x$ (resp. $\mathcal{T}_y$). Denote the law of SRW on H_e by $\mathrm{P}^{(e)}$. Let $D_i^{(e)} := \{u \in V_e : \mathrm{dist}_{H_e}(u, v) = i\}$ and $B_k^{(e)} := \cup_{i=0}^k D_i^{(e)}$. We now show that there is some constant $K(s,d)$ and an edge $e = \{x, y\} \in E$ belonging to some cycle in B_k such that $x \in B_k, y \in B_{k-1}$ and

$$\mathrm{P}_v[T_{D_k} = T_z] \leq K(s,d)\mathrm{P}_v^{(e)}[T_{D_k^{(e)}} = T_z]. \tag{5.2}$$

Once this is established, invoking the induction hypothesis concludes the induction step.

Consider an arbitrary cycle in B_k with at most one vertex in D_k. Let x be the vertex of the cycle which maximizes $\mathrm{P}_x[T_{D_k} = T_z]$. Let $e = \{x, y\}, e' = \{x, y'\}$ be the two edges of the cycle which are incident to x. Without loss of generality, let e be the one through which x is less likely to be reached. More precisely, assume that

$$\mathrm{P}_v[X_{T_x-1} = y, T_x \leq T_{D_k}] \leq \mathrm{P}_v[X_{T_x-1} = y', T_x \leq T_{D_k}]. \tag{5.3}$$

Also, by the choice of x we have that

$$\mathrm{P}_x[T_{D_k} = T_z] \geq \mathrm{P}_y[T_{D_k} = T_z]. \tag{5.4}$$

Note that if $x \in D_k$ and $x \neq z$ then $\mathrm{P}_v[T_{D_k} = T_z] = \mathrm{P}_v^{(e)}[T_{D_k^{(e)}} = T_z]$. If $x = z$ then by (5.3) $\mathrm{P}_v[T_{D_k} = T_z] \leq 2\mathrm{P}_v^{(e)}[T_{D_k^{(e)}} = T_z]$. Now consider the case that $x \notin D_k$. Denote $T_{x,y} := \min\{T_x, T_y\}$ and $T_x^+ := \inf\{t > 0 : X_t = x\}$. Observe that

$$\mathrm{P}_v[T_{D_k} = T_z < T_{x,y}] = \mathrm{P}_v^{(e)}[T_{D_k^{(e)}} = T_z < T_{x,y}]. \tag{5.5}$$

Thus in order to conclude the proof of (5.2) it remains only to show that

$$\mathrm{P}_v[T_{D_k} = T_z > T_{x,y}] \leq \tilde{C}(s,d)\mathrm{P}_v^{(e)}[T_{D_k^{(e)}} = T_z > T_{x,y}].$$

---

[4]The root of a d-ary tree is of degree $d - 1$.

By (5.3) we have that

$$\mathrm{P}_v[T_x < \min\{T_{D_k}, T_y\}] \geq \mathrm{P}_v[T_y < T_x < T_{D_k}] \geq \frac{1}{d}\mathrm{P}_v[T_y < T_{D_k}].$$

Thus $\mathrm{P}_v[T_x < T_{D_k}] \geq \frac{2}{d}\mathrm{P}_v[T_y < T_{D_k}]$. By (5.4) we get that

$$\mathrm{P}_v[T_x < T_{D_k} = T_z] = \mathrm{P}_v[T_x < T_{D_k}]\mathrm{P}_x[T_{D_k} = T_z] \geq \frac{2}{d}\mathrm{P}_v[T_y < T_{D_k}]\mathrm{P}_y[T_{D_k} = T_z]$$

$= \frac{2}{d}\mathrm{P}_v[T_y < T_{D_k} = T_z]$. Hence, there exists some constant $M(s, d)$ such that

$$\begin{aligned}
\mathrm{P}_v[T_{D_k} = T_z > T_{x,y}] &\leq \mathrm{P}_v[T_{D_k} = T_z > T_x] + \mathrm{P}_v[T_{D_k} = T_z > T_y] \\
&\leq (1 + \frac{d}{2})\mathrm{P}_v[T_{D_k} = T_z > T_x] \leq (d+2)\mathrm{P}_v[T_{D_k} = T_z, T_x < \min\{T_{D_k}, T_y\}] \\
&\leq M(s,d)\mathrm{P}_v[T_x < \min\{T_{D_k}, T_y\}]\mathrm{P}_x[T_{D_k} = T_z, \min\{T_x^+, T_y\} > T_{D_k}] \\
&\leq M(s,d)\mathrm{P}_v^{(e)}[T_x < \min\{T_{D_k^{(e)}}, T_y\}]\mathrm{P}_x^{(e)}[T_{D_k^{(e)}} = T_z < T_x^+] \\
&\leq M(s,d)\mathrm{P}_v^{(e)}[T_{D_k^{(e)}} = T_z > T_{x,y}],
\end{aligned} \tag{5.6}$$

where in the second inequality we have used the fact that $\mathrm{P}_x[\min\{T_x^+, T_y\} > T_{D_k}] \geq c(s,d)$ for some constant $c(s,d) > 0$[5] and that by the choice of x (namely, by (5.4)) we have that $\mathrm{P}_y[T_{D_k} = T_z \mid T_x > T_{D_k}] \leq \mathrm{P}_x[T_{D_k} = T_z] = \mathrm{P}_x[T_{D_k} = T_z \mid T_x^+ > T_{D_k}]$ and so

$$\mathrm{P}_x[T_{D_k} = T_z \mid \min\{T_x^+, T_y\} > T_{D_k}] \geq \mathrm{P}_x[T_{D_k} = T_z \mid T_x^+ > T_{D_k}] = \mathrm{P}_x[T_{D_k} = T_z].$$

We leave the missing details as an exercise. Finally, combining (5.5) and (5.6) yields (5.2). □

6 Proof of Theorem 1.3

Part (ii) was proven in § 3. Let $G_n = (V_n, E_n)$ be a sequence of non-bipartite, finite, connected, d-regular asymptotically one-sided Ramanujan graphs satisfying Assumption 1. Let $r_n \to \infty$ be as in Assumption 1. Pick some $k = k_n \to \infty$ such that $k_n^2 = o(r_n)$. From this point on we often suppress the dependence on n from our notation. Denote the transition matrix of SRW on G (resp. $G(k)$) by P (resp. K) and its stationary distribution by π (resp. $\pi_{G(k)}$). Let A be an arbitrary set such that $\pi(A) \leq \alpha = \alpha_n := d^{-3k^2}$. Denote $Q := P^{k+2k^2}$.

Before proceeding with the proof, we explain the choice of $k + 2k^2$ in the definition of Q. In order to obtain an upper bound on $\lambda_K(A)$ we shall apply Proposition 5.1 with P^t (for some t) and K in the roles of $P^{(2)}$ and $P^{(1)}$ (respectively) from Proposition 5.1. The obtained estimate is useful only when $t \geq ck^2$. Heuristically, this is related to the fact that a SRW on a d-regular tree is much more likely to be at time t at some given vertex of distance $O(\sqrt{t})$ from its starting point, than at some other given vertex at distance $\gg \sqrt{t}$ from its starting point (and we want $k = O(\sqrt{t})$).

Recall that $\rho_d := \frac{2\sqrt{d-1}}{d}$. Let λ_2 and λ_2' be the second largest eigenvalues of P and Q, respectively. Since $\lambda_2 = \rho_d^{1-o(1)}$, by decreasing k if necessary, we may assume that $\lambda_2 \leq \rho_d^{1-\frac{1}{3k^2}\frac{1}{\log d}}$. By Proposition 3.1 (using the notation from there) and our choice of α,

$$\lambda_Q(A) \leq \lambda_2' + \alpha = \lambda_2^{k+2k^2} + \alpha \leq C_1\rho_d^{k+2k^2}. \tag{6.1}$$

---

[5]This could be proved by induction on s.

Let $(S_t)_{t=0}^{\infty}$ be SRW on $\mathbb{T}_d$, the infinite d-regular tree rooted at o. Denote its transition kernel by $P_{\mathbb{T}_d}$. Denote the ith level of $\mathbb{T}_d$ by $\mathcal{L}_i$. Let $\tilde{S}_t$ be the level S_t belongs to. Let $v \in \mathcal{L}_k$. Let $T_0^+ := \inf\{t > 0 : \tilde{S}_t = 0\}$. Then by Lemma 6.1 (second inequality)

$$
\begin{aligned}
|\mathcal{L}_k| P_{\mathbb{T}_d}^{k+2k^2}(o,v) = \mathrm{P}_0[\tilde{S}_{k+2k^2} = k] &\geq \mathrm{P}_0[\tilde{S}_{k+2k^2} = k, T_0^+ > k + 2k^2] \\
&\geq c_0 k^{-2} 2^{k+2k^2}(d-1)^{k^2+k-1} d^{-(k+2k^2)+1} \geq c_1 k^{-2}(d-1)^{\frac{k}{2}} \rho_d^{2k^2+k}
\end{aligned}
\tag{6.2}
$$

Let x, y be a pair of adjacent vertices in $G(k)$. It is standard that $P^t(x,y) \geq P_{\mathbb{T}_d}^t(o,v)$ for all t (where v is as above), and so by (6.2)

$$
Q(x,y) = P^{k+2k^2}(x,y) \geq P_{\mathbb{T}_d}^{k+2k^2}(o,v) \geq (d-1)^{\frac{k}{2}(1-o(1))} \rho_d^{2k^2+k} =: C_k. \tag{6.3}
$$

By Proposition 5.1 (and borrowing the notation from there) in conjunction with (6.1), (6.3) and Assumption 1 (which implies that there exists some constant $C_0 = C_0(d) > 0$ such that $L := \frac{\max_x \deg_{G(k)}(x)}{\min_y \deg_{G(k)}(y)} \leq C_0$ and that if x, y are of distance k in G then $K(x,y) \leq C_0(d-1)^{-k}$), we have that

$$
\lambda_K(A) \leq \lambda_Q(A) C_0^3 (d-1)^{-k} / C_k = (d-1)^{-\frac{k}{2}(1-o(1))}.
$$

Denote the probability w.r.t. SRW on $G(k)$ by $\mathbb{P}$. By (3.2) we have for all t (uniformly) that

$$
\max_{(a,A):a \in A, \pi(A) \leq \alpha} \mathbb{P}_a[T_{A^c} > t] \leq \sqrt{C_0 \alpha |V|}(d-1)^{-\frac{tk}{2}(1-o(1))} = \sqrt{\alpha |V|}(d-1)^{-\frac{tk}{2}(1-o(1))},
\tag{6.4}
$$

where we have used the fact that $\max_{x \in V} \pi_{G(k)}(x)/\pi(x) \leq C_0$, where C_0 is as above.

Consider SRW on G, $(X_t)_{t=0}^{\infty}$. Let $T_0 := 0$ and inductively, $T_{i+1} := \inf\{t \geq T_i : \mathrm{dist}_G(X_{T_{i+1}}, X_{T_i}) = k\}$. As in Definition 4.2, consider the chain $\mathbf{Y} = (Y_i)_{i=0}^{\infty}$, where $Y_i := X_{T_i}$ for all i. Let W be its transition matrix. By Assumption 1 and Lemma 5.2 there exists some constant $C = C(d)$ such that for all $x, y \in V$ of distance k from one another (in G),

$$
1/C \leq W(x,y)/K(x,y) \leq C. \tag{6.5}
$$

Denote the probability w.r.t. $\mathbf{Y}$ by $\mathbf{P}$. Then by (6.4) and (6.5)

$$
\max_{(a,A):a \in A, \pi(A) \leq \alpha} \mathbf{P}_a[T_{A^c} > t] \leq C^t \max_{(a,A):a \in A, \pi(A) \leq \alpha} \mathbb{P}_a[T_{A^c} > t] \leq \sqrt{\alpha |V|}(d-1)^{-\frac{tk}{2}(1-o(1))},
\tag{6.6}
$$

uniformly for all t. Denote the distribution of SRW on G by P. Observe that for all $s, t \geq 0$

$$
\max_{(a,A):a \in A, \pi(A) \leq \alpha} \mathrm{P}_a[T_{A^c} > t + s] \leq \max_{(a,A):a \in A, \pi(A) \leq \alpha} \mathbf{P}_a[T_{A^c} > \tau(t)] + \max_{a \in V} \mathrm{P}_a[T_{\tau(t)} > t + s],
$$

where

$$
\tau(t) := \lceil \frac{(d-2)t}{dk} \rceil.
$$

To conclude the proof (using (3.1) in conjunction with Lemma 1.2), we now show that (for some $o(1)$ terms) substituting above $t = \lceil (1 + o(1)) \frac{d}{d-2} \log_{d-1} |V| \rceil$ and $s = t/\sqrt{k} + t^{2/3}$ (the value $2/3$ in the exponent can be replaced by any number in $(1/2, 1)$) yields $\max_{(a,A):a \in A, \pi(A) \leq \alpha} \mathrm{P}_a[T_{A^c} > t + s] = o(1)$. By (6.6) it suffices to show that for this choice of s and t we have that $\max_{a \in V} \mathrm{P}_a[T_{\tau(t)} > t + s] = o(1)$.

Fix s and t as above. We say that time j is *good* if X_j has $d - 1$ neighbors of greater distance from $X_{T_{i(j)}}$, where $i(j)$ is the index for which $j \in [T_{i(j)}, T_{i(j)+1})$. Let

$$
U_i := |\{t \in [T_i, T_{i+1}) : t \text{ is not good}\}| \quad \text{and} \quad U := \sum_{i=0}^{\tau(t)} U_i.
$$

By Assumption 1 we have that $\max_v \mathrm{P}_v[U_0 > \ell] \le C' e^{-c\ell}$ for all ℓ, for some constant $c, C' > 0$ (this is left as an exercise). By the Markov property, it follows that

$$\max_v \mathrm{P}_v[U > \frac{t}{\sqrt{k}}] = o(1).$$

Consider a coupling of the SRW on G $(X_j)_{j=0}^{\infty}$ with the SRW on $\mathbb{T}_d$ started from its root o $(S_j)_{j=0}^{\infty}$ in which if j is the ℓth good time, then $\mathrm{dist}_G(X_{j+1}, X_{T_{i(j)}}) < \mathrm{dist}_G(X_j, X_{T_{i(j)}})$ iff $\mathrm{dist}_{\mathbb{T}_d}(S_{\ell+1}, o) < \mathrm{dist}_{\mathbb{T}_d}(S_\ell, o)$ (unless $S_\ell = o$, but there is no harm in neglecting this possibility, as the number of returns to o has a Geometric distribution). Using this coupling we get that for all $a \in V$ we have that

$$\mathrm{P}_a[T_{\tau(t)} > t + s] \le \mathrm{P}_a[U > \frac{t}{\sqrt{k}}] + \max_{0 \le j \le \lceil \frac{t}{\sqrt{k}} \rceil} \mathrm{P}_o[S_{t+s-j} \in \cup_{i=0}^{\tau(t)+j} \mathcal{L}_i] = o(1).$$

To see that $\max_{0 \le j \le \lceil \frac{t}{\sqrt{k}} \rceil} \mathrm{P}_o[S_{t+s-j} \in \cup_{i=0}^{\tau(t)+j} \mathcal{L}_i] = o(1)$ use the fact that the distance of S_{t+s-j} from o is concentrated around $\frac{d-2}{d}(t + s - j)$ within a window whose length is of order $\sqrt{t}$ (c.f. [4] (2.2)–(2.3) pg. 9) and that by our choice of s we have that $\frac{d-2}{d}(t + s - j) - (\tau(t) + j) \gg \sqrt{t}$, for all $0 \le j \le \lceil \frac{t}{\sqrt{k}} \rceil$. $\qquad \square$

Lemma 6.1. *Let M be the number of paths of length $k + 2k^2$ in $\mathbb{Z}$, starting from 0, which end at k and do not return to 0. Then $M \ge c_0 2^{k+2k^2}/k^2$.*

Proof. Let $(Z_i)_{i=0}^{\infty}$ be a SRW on $\mathbb{Z}$. Let $T_0^+ := \inf\{t > 0 : Z_t = 0\}$. Then

$$\mathrm{P}_0[Z_{k+2k^2} = k, T_0^+ > k + 2k^2] \ge \mathrm{P}_0[T_0^+ > k + 2k^2 \ge T_k] \min_{0 \le i \le k^2} \mathrm{P}_k[T_0 > 2i, Z_{2i} = k] \ge c_0 k^{-2},$$

where we have used the fact that $\mathrm{P}_0[T_0^+ > k + 2k^2 \ge T_k] \ge c_1 \mathrm{P}_0[T_0^+ > T_k] = c_1/(2k)$ and that $\mathrm{P}_k[T_0 > 2i, Z_{2i} = k] \ge \mathrm{P}_k[T_{\{0,2k\}} > 2i] \mathrm{P}_k[Z_{2i} = k \mid T_{\{0,2k\}} > 2i] \ge c_2 \cdot \frac{1}{2k}$ for all $i \le k^2$. $\qquad \square$

References

[1] Abért, M., Glasner, Y. and Virág, B.: The measurable Kesten theorem. *Ann. Probab.*, 44(3):1601–1646, 2016. MR-3502591

[2] Riddhipratim. B., Hermon, J. and Peres Y.: Characterization of cutoff for reversible Markov chains. *Ann. Probab.*, 45(3):1448–1487, 2017. MR-3650406

[3] Friedman, J.: A proof of Alon's second eigenvalue conjecture. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, pages 720–724. ACM, New York, 2003. MR-2120428

[4] Lubetzky, E. and Peres, Y.: Cutoff on all Ramanujan graphs. *Geom. Funct. Anal.*, 26(4):1190–1216, 2016. MR-3558308

[5] Lubetzky, E. and Sly, A: Cutoff phenomena for random walks on random regular graphs. *Duke Math. J.*, 153(3):475–510, 2010. MR-2667423

[6] Lubotzky, A., Phillips, R. and Sarnak, P.: Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988. MR-963118

[7] Marcus, A. W., Spielman, D. A., and Srivastava, N.: Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. of Math. (2)*, 182(1):307–325, 2015. MR-3374962

[8] Margulis, G. A.: Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988. MR-939574

[9] Morgenstern, M.: Existence and explicit constructions of $q + 1$ regular Ramanujan graphs for every prime power q. *J. Combin. Theory Ser. B*, 62(1):44–62, 1994. MR-1290630

[10] Nilli, A.: On the second eigenvalue of a graph. *Discrete Math.*, 91(2):207–210, 1991. MR-1124768