A Combinatorial Characterization of Constant Mixing Time

Lap Chi Lau
University of Waterloo, Canada

Raymond Liu
University of Waterloo, Canada

Abstract

Classical spectral graph theory characterizes graphs with logarithmic mixing time. In this work, we present a combinatorial characterization of graphs with constant mixing time. The combinatorial characterization is based on the small-set bipartite density condition, which is weaker than having near-optimal spectral radius and is stronger than having near-optimal small-set vertex expansion.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains

Keywords and phrases Random walks, mixing time, bipartite density, spectral graph theory

Digital Object Identifier 10.4230/LIPIcs.ITCS.2026.92

Funding *Lap Chi Lau*: Supported by an NSERC Discovery Grant.
Raymond Liu: Supported by an NSERC Discovery Grant.

Acknowledgements We thank Thatchaphol Saranurak for discussions on length-constrained expanders that inspired the study in this work.

1 Introduction

We start with a brief review of some background in random walks and spectral graph theory. Let $G = (V, E)$ be a d-regular graph and $n := |V|$. Let P be the random walk matrix of G, with stationary distribution $\pi = \vec{1}/n$ as G is regular. The ϵ-mixing time of the random walks is defined as

$$\tau_\epsilon(P) := \min\left\{t \;\middle|\; \frac{1}{2}\left\|P^t p_0 - \pi\right\|_1 \leq \epsilon \text{ for any initial distribution } p_0\right\}.$$

Let $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$ be the eigenvalues of P and $\lambda := \max\{\lambda_2, |\lambda_n|\}$ be the spectral radius of P. The graph G is called a spectral expander if λ is a constant strictly smaller than one. Standard spectral analysis [8] shows that

$$\frac{1}{2}\left\|P^t p_0 - \pi\right\|_1 \leq \lambda^t \cdot \|p_0\|_2 \cdot \sqrt{n}. \tag{1}$$

This implies that the mixing time of random walks on a spectral expander is $O(\log n)$, and this upper bound is optimal as the diameter of the graph is $\Omega(\log n)$ when d is a constant. Cheeger's inequality states that G is a spectral expander if and only if G is a combinatorial expander (i.e., with constant edge conductance); see Section 2. This gives a combinatorial characterization of graphs with $O(\log n)$ mixing time. The relations between eigenvalues, combinatorial expansion, and mixing time are fundamental results in spectral graph theory.

Inspired by the recent development in constant-hop expander graphs (see [7, 6] and the references therein), where the focus is on sending multicommodity flows using paths of *constant* length, we are interested in characterizing graphs with constant mixing time, as these form a nice class of constant-hop expander graphs. For a d-regular graph to have constant diameter, a necessary condition is that $d \geq n^\xi$ for some small constant $\xi > 0$, so

17th Innovations in Theoretical Computer Science Conference (ITCS 2026).
Editor: Shubhangi Saraf; Article No. 92; pp. 92:1–92:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

we focus on graphs in this moderately-dense regime as in [7, 6]. Even in this regime, it is not difficult to construct spectral expanders with $\Omega(\log n)$ diameter, so we need to look for stronger conditions to guarantee constant mixing time. We say a d-regular graph has an *inverse-polynomial spectral radius* if $\lambda \lesssim 1/d^c$ for some constant $c \in (0, \frac{1}{2}]$. From (1), observe that n^ξ-regular graphs with $\lambda \lesssim 1/d^c$ have constant mixing time $O(1/(c\xi))$. Graphs with inverse-polynomial spectral radius also exhibit stronger combinatorial expansion properties: Tanner's theorem implies that such graphs have vertex expansion $\Omega(d^{2c})$ for sets of size $O(n/d^{2c})$; see Section 2. Graphs in these regimes satisfy interesting properties but were not explored much before.

These lead us to study the relations between inverse-polynomial spectral radius, small-set vertex expansion, and constant mixing time. One natural question is whether an n^ξ-regular graph with near-optimal small-set vertex-expansion has constant mixing time. In Subsection 4.1, we provide a negative example to this question, which suggests that an even stronger combinatorial condition is required to guarantee constant mixing time.

Another natural question is whether inverse-polynomial spectral radius is necessary to guarantee constant mixing time. To answer this question, we consider a combinatorial characterization of the spectral radius through the expander mixing lemma [2]: If a d-regular graph $G = (V, E)$ has spectral radius λ, then

$$\left| \big|E(S,T)\big| - \frac{d|S||T|}{n} \right| \leq \lambda d \sqrt{|S||T|} \quad \text{for any } S, T \subseteq V, \tag{2}$$

where $E(S,T) := \{(u,v) \mid u \in S, v \in T, uv \in E\}$ is the set of ordered edges where $u \in S$ and $v \in T$. The converse of the expander mixing lemma by Bilu and Linial [4] shows that if (2) is satisfied for all disjoint $S, T \subseteq V$, then the graph has spectral radius $O(\lambda \cdot \log(1 + \frac{1}{\lambda}))$. Thus, if (2) is satisfied for $\lambda \leq O(1/d^c)$ for some constant $c \in (0, \frac{1}{2}]$ and $d \geq n^\xi$ for some constant $\xi > 0$, then the graph has constant mixing time. This provides a combinatorial sufficient condition for constant mixing time, but we will show that it is not a necessary condition.

Our Results

We show that only the upper bounds in (2) are needed for constant mixing time, simplifying (2) to a condition about the bipartite density between two sets.

▶ **Definition 1** (α-Bipartite Density). *Let $G = (V, E)$ be a d-regular graph. For any $\alpha \in (\sqrt{d}, d]$, we say that G satisfies the α-bipartite density condition if*

$$|E(S,T)| \leq \frac{d|S||T|}{n} + \alpha\sqrt{|S||T|} \quad \text{for all } S, T \subseteq V. \tag{3}$$

If (3) is only satisfied for sets S, T with $|S|, |T| \leq \delta n$ for some $\delta \leq 1$, then we say that G satisfies the δ-small-set α-bipartite density condition.

We note that the α-bipartite density condition is weaker than having spectral radius α/d by (2), but it is stronger than having near-optimal small-set vertex expansion (see Subsection 4.1).

Our main results are that this condition implies fast mixing time.

▶ **Theorem 2** (Upper Bounding Mixing Time by Bipartite Density). *Let $G = (V, E)$ be a d-regular graph with $d = n^\xi$ for some constant $\xi > 0$. If G satisfies the α-bipartite density condition for $\alpha \lesssim d/(\log d)^2$, then*

$$\tau_{1/n}(P) \lesssim \left(\frac{\log n}{\log(d/\alpha)}\right)^2.$$

In particular, if $\alpha = d^{1-c}$ *for some constant* $c \in (0, \frac{1}{2}]$, *this implies constant mixing time such that*

$$\tau_{1/n}(P) \lesssim \frac{1}{c^2\xi^2}.$$

The standard definition of mixing time is $\tau_{1/3}(P)$, and a well-known fact [8] is that $\tau_{1/n}(P) \lesssim \tau_{1/3}(P) \cdot \log n$. In Theorem 2, we bound $\tau_{1/n}(P)$ directly without losing the logarithmic factor, matching the constant mixing time result for graphs with inverse-polynomial spectral radius.

Furthermore, we show an improved upper bound on standard mixing time using only the small-set bipartite density condition.

▶ **Theorem 3** (Upper Bounding Mixing Time by Small-Set Bipartite Density). *Let* $G = (V, E)$ *be a* d*-regular graph with* $d = n^\xi$ *for some constant* $\xi > 0$. *If* G *satisfies the* δ*-small-set* α*-bipartite density condition for some* $\alpha \lesssim d/(\log d)^2$ *and* $\delta \gtrsim \alpha/d$, *then*

$$\tau_{1/3}(P) \lesssim \frac{\log n}{\log(d/\alpha)}.$$

In particular, if $\alpha = d^{1-c}$ *for some constant* $c \in (0, \frac{1}{2}]$ *and* (3) *holds for all sets* $S, T \subseteq V$ *with* $|S|, |T| \lesssim n/d^c$, *then this implies constant standard mixing time such that*

$$\tau_{1/3}(P) \lesssim \frac{1}{c\xi}.$$

We remark that constant standard mixing time implies the graph is a constant-hop expander graph. It is our hope that the small-set bipartite density condition can lead to a simpler cut-matching game for constructing constant-hop expanders.

We also establish a lower bound on the mixing time using δ-small-set α-bipartite density. The following theorem states that the existence of a dense bipartite structure between two small sets implies slow mixing time.

▶ **Theorem 4** (Lower Bounding Mixing Time by Bipartite Density). *Let* G *be a* d*-regular graph. If there exist* $S, T \subseteq V$ *such that*

$$|E(S,T)| \geq \frac{d|S||T|}{n} + \alpha\sqrt{|S||T|} \quad \text{and} \quad |S|, |T| \leq \delta n,$$

then

$$\tau_{1/n}(P) \gtrsim \frac{\log(1/\delta)}{\log(d/\alpha)}.$$

In particular, if there are two small sets with high bipartite density such that $|S|, |T| \leq n^{1-\epsilon}$ *for some constant* ϵ *and* $\alpha = \Omega(d/(\text{poly}\log d))$, *then the graph has non-constant mixing time such that* $\tau_{1/n}(P) \gtrsim (\log n)/\log\log d$.

To summarize, we can view the δ-small-set α-bipartite density condition as a loose characterization of constant mixing time: If all sets of large enough size have low bipartite density, then the graph has constant mixing time; if some sets of small enough size have high bipartite density, then the graph has non-constant mixing time.

We think the proof approach in Theorem 2 and Theorem 3 is also interesting that it provides a clean and direct way to upper bound the mixing time using a combinatorial condition, without going through a spectral argument as usual.

2 Preliminaries

We write $f \lesssim g$ if $f = O(g)$, $f \gtrsim g$ if $f = \Omega(g)$, and $f \asymp g$ if $f = \Theta(g)$.

We assume the given graph $G = (V, E)$ is a d-regular graph throughout this paper, with $n := |V|$ vertices and $m := |E|$ edges. Let A be the adjacency matrix of G, and let $P := A/d$ be the normalized adjacency matrix, which is also the transition matrix of random walks, as G is d-regular.

Let $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$ be the eigenvalues of P. We call $\lambda := \max\{\lambda_2, |\lambda_n|\}$ the spectral radius of G. A well-known result by Alon and Boppana [10] establishes that $\lambda \geq 2\sqrt{d-1}/d$ as $n \to \infty$. A graph is called Ramanujan [9] if $\lambda \leq 2\sqrt{d-1}/d$.

Variation Distance and Mixing Time

The stationary distribution π of the transition matrix $P = A/d$ is the uniform distribution $\vec{1}/n$. The variation distance between any two probability distributions p, q is defined as $d_{\mathrm{TV}}(p, q) = \frac{1}{2}\|p - q\|_1$. The variation distance at step t of the random walk is defined as

$$d_{\mathrm{TV}}(t) = \max_p d_{\mathrm{TV}}(P^t p, \pi),$$

where the maximum is over all initial probability distributions p on V. Given $\varepsilon > 0$, the ε-mixing time of the random walk is defined as

$$\tau_\varepsilon(P) = \min\{t \mid d_{\mathrm{TV}}(t) \leq \varepsilon\}.$$

Standard spectral analysis in (1) yields $d_{TV}(P^t p, \pi) \lesssim \lambda^t \cdot \|p\|_2 \cdot \sqrt{n}$. It follows that the mixing time is upper bounded by $O(\log(n)/(1-\lambda))$, and so when G is a spectral expander, the mixing time is bounded by $O(\log n)$.

A standard fact in Markov chain [8] shows that for any $k \in \mathbb{N}$,

$$d_{\mathrm{TV}}(kt) \leq (2d_{\mathrm{TV}}(t))^k. \tag{4}$$

In particular, any graph with mixing time $\tau_{1/3}(P) \lesssim \log n$ implies $\tau_{1/n}(P) \lesssim (\log n)^2$.

Edge Conductance

Given an undirected graph $G = (V, E)$ and $S, T \subseteq V$, define

$$E(S, T) := \{(u, v) \mid u \in S, v \in T, uv \in E\},$$

where an edge with $u, v \in S \cap T$ is counted twice, as both (u, v) and (v, u) are in $E(S, T)$. The edge boundary of S is defined as $\delta(S) := E(S, V \setminus S)$.

The second eigenvalue λ_2 is closely related to the edge conductance of the graph, defined as

$$\phi(G) = \min_{S \subseteq V : |S| \leq n/2} \frac{|\delta(S)|}{d|S|}.$$

Cheeger's inequality [5, 3, 1] states that

$$\frac{1}{2}(1 - \lambda_2) \leq \phi(G) \leq \sqrt{2(1 - \lambda_2)}. \tag{5}$$

The edge conductance characterizes the mixing time of constant degree graphs:

$$\frac{1}{\phi(G)} \lesssim \tau_{1/3}(W) \lesssim \frac{\log n}{\phi(G)^2},$$

where $W = \frac{1}{2}(I + P)$ is the lazy random walk. We only consider non-lazy random walk P in this paper. Since for any initial distribution of the form $p = \chi_v$, where χ_v is the indicator vector of a vertex $v \in V$, a staying probability of $\frac{1}{2}$ automatically implies $\tau_{1/n}(P)$ is non-constant.

Graphs with Inverse-Polynomial Spectral Radius

A graph G exhibits stronger probabilistic and combinatorial properties when it has an inverse-polynomial spectral radius such that $\lambda \lesssim 1/d^c$ for some constant $c \in (0, 1/2]$.

For random walks, it follows from (1) that if $d = n^\xi$ for some constant $\xi > 0$, then the mixing time of a graph with inverse-polynomial spectral radius is a constant, upper bounded by $O(1/(c\xi))$.

A graph with inverse-polynomial spectral radius also has large small-set vertex expansion. Define the vertex expansion of a set $S \subseteq V$ as

$$\psi(S) := \frac{|\partial(S)|}{|S|} \quad \text{where } \partial(S) := \{v \notin S \mid \exists u \in S \text{ with } uv \in E\}.$$

Tanner [11] proved that

$$\psi(S) \geq \left(\frac{|S|}{n}(1 - \lambda^2) + \lambda^2\right)^{-1} - 1. \tag{6}$$

In particular, when G is Ramanujan, it has near-optimal small-set vertex expansion such that $\psi(S) = \Omega(d)$ for sets of size up to $\Omega(n/d)$.

We remark that the small-set vertex expansion condition can also be derived from the α-bipartite density condition with $\alpha = \lambda d$ (instead of the spectral radius λ).

3 Upper Bounds on Mixing Time

In this section, we prove Theorem 2 and Theorem 3.

We show that for a graph that satisfies the small-set bipartite density condition, the variation distance to the stationary distribution after $O(1/(c\xi))$ steps of random walks is essentially upper bounded by how close δ is to 1.

▶ **Theorem 5** (Upper Bounding Variation Distance by Small-Set Bipartite Density). *Let G be a d-regular graph with $d = n^\xi$ for some constant $\xi > 0$. If G satisfies the δ-small-set α-bipartite density condition for some $\alpha \lesssim d/(\log d)^2$, then*

$$d_{\mathrm{TV}}(t) \leq \frac{1}{4} \cdot \sqrt{1 - \delta} + O\left(\frac{1}{n^{c\xi/8}}\right) \quad \text{for} \quad t \geq \frac{4}{c\xi} + 1.$$

where $c := 1 - \log_d \alpha$ such that $\alpha = d^{1-c}$.

Assuming Theorem 5, the proofs of Theorem 2 and Theorem 3 follow easily.

Proof of Theorem 3. Since $c = 1 - \log_d \alpha$ and $n = d^{1/\xi}$, it follows that

$$c\xi = \xi - \xi \cdot \frac{\log \alpha}{\log d} = \xi - \frac{\log \alpha}{\log n} \quad \Longrightarrow \quad \frac{1}{n^{c\xi}} = \frac{1}{n^{(\xi - \log_n \alpha)}} = \frac{\alpha}{d}. \tag{7}$$

By (7) and Theorem 5, for $t \geq 4/c\xi + 1$,

$$d_{\mathrm{TV}}(t) \leq \frac{1}{4} \cdot \sqrt{1 - \delta} + O\left(\left(\frac{\alpha}{d}\right)^{\frac{1}{8}}\right) \leq \frac{1}{4} + O\left((\log d)^{-\frac{1}{4}}\right) \leq \frac{1}{3},$$

where the second inequality follows by the assumption that $\alpha \lesssim d/(\log d)^2$, and the last inequality holds for sufficiently large d. This implies that

$$\tau_{1/3}(P) \lesssim \frac{1}{c\xi} = \frac{\log d}{\xi \log(d/\alpha)} = \frac{\log n}{\log(d/\alpha)}$$ ◀

To prove Theorem 2, we apply the standard fact in (4).

Proof of Theorem 2. Here we assume α-bipartite density so $\delta = 1$. For $t \geq 4/c\xi + 1$, as in (7),

$$d_{\mathrm{TV}}(t) \leq \frac{1}{4} \cdot \sqrt{1-\delta} + O\Big(\frac{1}{n^{c\xi/8}}\Big) = O\Big(\frac{\alpha}{d}\Big)^{\frac{1}{8}} := \beta,$$

By (4), for $k \geq (\log \frac{1}{n})/\log(2\beta)$,

$$d_{\mathrm{TV}}(kt) \leq (2 \cdot d_{\mathrm{TV}}(t))^k \leq (2\beta)^k \leq \frac{1}{n}.$$

Since $c = 1 - \log_d \alpha = \log(d/\alpha)/\log d$, we conclude that

$$\tau_{1/n}(P) \lesssim \frac{\log n}{\log(d/\alpha)} \cdot \Big(\frac{4}{c\xi} + 1\Big) \lesssim \frac{\log n}{c\xi \log(d/\alpha)} = \frac{\log d}{\log(d/\alpha)} \cdot \frac{\log n}{\xi \log(d/\alpha)} = \Big(\frac{\log n}{\log(d/\alpha)}\Big)^2.$$

◀

3.1 Bounding Mixing Time via 2-Norm

The main idea is to measure the mixing progress by the ℓ_2-norm of the random walk distribution. For any distribution p, $\|p\|_2^2$ is lower bounded by that of the stationary distribution, which is $1/n$. The following proposition measures the progress of $\|p\|_2^2$ approaching C_δ/n, where C_δ is a constant depending on how close δ is to 1.

▶ **Proposition 6** (2-Norm Decrease). *Let G be a d-regular graph with $d = n^\xi$ for some constant $\xi > 0$. If G satisfies the δ-small set α-bipartite density condition with $\alpha \log d/(\xi d) \leq 1$ and $\delta \geq \frac{\sqrt{5}-2}{2} \cdot \alpha/d$, then*

$$\|Pp\|_2^2 \leq \frac{C_\delta}{n} + O\Bigg(\sqrt{\frac{\alpha \log d}{\xi d}}\Bigg) \cdot \|p\|_2^2$$

for any probability distribution p, where $C_\delta = \delta + 5(1-\delta)/4$.

Proof. Let p be an arbitrary probability distribution and Pp be the distribution after one step of random walks. Assume $p(1) \geq \cdots \geq p(n)$ without loss of generality, and let σ be a permutation of $[n]$ such that $Pp(\sigma(1)) \geq Pp(\sigma(2)) \geq \cdots \geq Pp(\sigma(n))$. Let $T = [\sigma(1), \cdots, \sigma(k)]$ be the largest k entries in Pp for some $k \in [n]$, and let $d_T(i) = |\{j \in T \mid ij \in E\}|$ denotes the number of vertices in T adjacent to some vertex $i \in V$. Note that

$$\sum_{1\leq j\leq k} Pp(\sigma(j)) = \sum_{1\leq j\leq k} \sum_{i:(i,\sigma(j))\in E} \frac{p(i)}{d} = \sum_{i=1}^{n} \frac{p(i)}{d} \cdot d_T(i) \leq \sum_{i=1}^{n} \frac{p(i)}{d} \cdot d'_T(i), \tag{8}$$

where $d'_T(1) \geq \cdots \geq d'_T(n)$ is a permutation of $d_T(1), \cdots, d_T(n)$ sorted from largest to smallest, with the last inequality follows from rearrangement inequality.

Consider first the case that $|T| = k \leq \delta n$. For any $r \in \{1, \cdots, \delta n\}$, it follows from the δ-small-set α-bipartite density condition and an averaging argument that

$$\sum_{i \leq r} d'_T(i) \leq \frac{dr|T|}{n} + \alpha\sqrt{r|T|} \quad \implies \quad d'_T(r) \leq \frac{d|T|}{n} + \alpha\sqrt{\frac{|T|}{r}}. \tag{9}$$

And, for any $r \geq \delta n$, it follows from the sortedness of $d'_T(i)$ that

$$d'_T(r) \leq d'_T(\delta n) \leq \frac{d|T|}{n} + \alpha\sqrt{\frac{|T|}{\delta n}}. \tag{10}$$

Denote $\bar{\alpha} := \alpha/d$. Combining (9), (10) with (8), it follows that

$$\begin{aligned}
\sum_{1 \leq j \leq k} Pp(\sigma(j)) &\leq \sum_{i=1}^{n} \frac{p(i)}{d} \cdot d'_T(i) \\
&\leq \sum_{i=1}^{\delta n} \frac{p(i)}{d} \cdot \left(\frac{d|T|}{n} + \alpha\sqrt{\frac{|T|}{i}}\right) + \sum_{i=\delta n+1}^{n} \frac{p(i)}{d} \cdot \left(\frac{d|T|}{n} + \alpha\sqrt{\frac{|T|}{\delta n}}\right) \\
&= \frac{|T|}{n} \cdot \sum_{i=1}^{n} p(i) + \bar{\alpha}\sqrt{|T|} \cdot \sum_{i=1}^{\delta n} \frac{p(i)}{\sqrt{i}} + \frac{\bar{\alpha}\sqrt{|T|}}{\sqrt{\delta n}} \cdot \sum_{i=\delta n+1}^{n} p(i) \\
&\leq \frac{|T|}{n} + \frac{\bar{\alpha}\sqrt{|T|}}{\sqrt{\delta n}} + \bar{\alpha}\sqrt{|T|} \cdot \sum_{i=1}^{n} \frac{p(i)}{\sqrt{i}},
\end{aligned}$$

where the last inequality follows from $\sum_{i=1}^{n} p(i) = 1$. Recall that $Pp(\sigma(1)) \geq \cdots \geq Pp(\sigma(n))$ and $|T| = k$. Hence, by an averaging argument,

$$Pp(\sigma(k)) \leq \left(\frac{1}{n} + \frac{1}{\sqrt{k}} \cdot \frac{\bar{\alpha}}{\sqrt{\delta n}}\right) + \frac{\bar{\alpha}}{\sqrt{k}} \cdot \sum_{i=1}^{n} \frac{p(i)}{\sqrt{i}}. \tag{11}$$

For the case that $k > \delta n$, we simply use the sortedness of Pp to obtain that

$$Pp(\sigma(k)) \leq Pp(\sigma(\delta n)) \leq \left(\frac{1}{n} + \frac{\bar{\alpha}}{\delta n}\right) + \frac{\bar{\alpha}}{\sqrt{\delta n}} \cdot \sum_{i=1}^{n} \frac{p(i)}{\sqrt{i}}. \tag{12}$$

Let $L_1(k) := \frac{1}{n} + \frac{1}{\sqrt{k}} \cdot \frac{\bar{\alpha}}{\sqrt{\delta n}}$ and $H_1(k) := \frac{\bar{\alpha}}{\sqrt{k}} \cdot \sum_{i=1}^{n} \frac{p(i)}{\sqrt{i}}$ denote the two terms in (11). Similarly, let $L_2 := \frac{1}{n} + \frac{\bar{\alpha}}{\delta n}$ and $H_2 := \frac{\bar{\alpha}}{\sqrt{\delta n}} \cdot \sum_{i=1}^{n} \frac{p(i)}{\sqrt{i}}$ denote the two terms in (12). It follows that

$$\begin{aligned}
\|Pp\|_2^2 &= \sum_{k=1}^{\delta n} Pp(\sigma(k))^2 + \sum_{k=\delta n+1}^{n} Pp(\sigma(k))^2 \\
&\leq \sum_{k=1}^{\delta n} \left(L_1(k)^2 + 2L_1(k)H_1(k) + H_1(k)^2\right) + \sum_{k=\delta n+1}^{n} \left(L_2^2 + 2L_2H_2 + H_2^2\right) \\
&\leq \sum_{k=1}^{\delta n} L_1(k)^2 + \sum_{k=\delta n+1}^{n} L_2^2 + \sum_{k=1}^{n} \left(2L_2H_2 + H_2^2\right) \\
&= \sum_{k=1}^{\delta n} L_1(k)^2 + (1-\delta)n \cdot L_2^2 + n \cdot \left(2L_2H_2 + H_2^2\right),
\end{aligned}$$

where the second inequality follows since $L_1(k) \leq L_2$ and $H_1(k) \leq H_2$ for all $k \leq \delta n$.

Let $\gamma := \bar{\alpha}/\delta$. It remains to analyze the three terms. For the first term,

$$\begin{aligned}\sum_{k=1}^{\delta n} L_1(k)^2 = \sum_{k=1}^{\delta n}\Big(\frac{1}{n}+\frac{1}{\sqrt{k}}\cdot\frac{\bar{\alpha}}{\sqrt{\delta n}}\Big)^2 &\le \sum_{k=1}^{\delta n}\frac{1}{n^2}+\frac{2\bar{\alpha}}{n\sqrt{\delta n}}\sum_{k=1}^{\delta n}\frac{1}{\sqrt{k}}+\frac{\bar{\alpha}^2}{\delta n}\sum_{k=1}^{n}\frac{1}{k}\\ &\le \frac{1}{n}\cdot\Big(\delta+O\big(\bar{\alpha}+\bar{\alpha}\gamma\log n\big)\Big)\\ &\le \frac{1}{n}\cdot\Big(\delta+O\Big(\frac{\bar{\alpha}\log d}{\xi}\Big)\Big), \end{aligned} \tag{13}$$

where the second inequality uses $\sum_{k=1}^{\delta n} 1/\sqrt{k} = O(\sqrt{\delta n})$ and $\sum_{k=1}^{n} 1/k = O(\log n)$, and the last inequality follows by our assumptions that $\gamma \le O(1)$ and $n = d^{1/\xi}$.

For the second term, using the assumption that $\gamma \le \frac{\sqrt{5}}{2} - 1$,

$$(1-\delta)n\cdot L_2^2 = (1-\delta)n\cdot\Big(\frac{1}{n}+\frac{\bar{\alpha}}{\delta n}\Big)^2 = (1-\delta)\cdot\frac{(1+\gamma)^2}{n} \le (1-\delta)\cdot\frac{5}{4n}. \tag{14}$$

For the last term, applying the Cauchy-Schwarz inequality,

$$\begin{aligned}2L_2H_2+H_2^2 &= \frac{2(1+\gamma)}{n}\cdot\frac{\bar{\alpha}}{\sqrt{\delta n}}\sum_{i=1}^{n}\frac{p(i)}{\sqrt{i}}+\frac{\bar{\alpha}^2}{\delta n}\Big(\sum_{i=1}^{n}\frac{p(i)}{\sqrt{i}}\Big)^2\\ &\le \frac{2\bar{\alpha}(1+\gamma)}{n\sqrt{\delta n}}\cdot\sqrt{\sum_{i=1}^{n}\frac{1}{i}}\cdot\sqrt{\sum_{i=1}^{n}p(i)^2}+\frac{\bar{\alpha}^2}{\delta n}\cdot\Big(\sum_{i=1}^{n}\frac{1}{i}\Big)\cdot\Big(\sum_{i=1}^{n}p(i)^2\Big)\\ &\lesssim \frac{2\bar{\alpha}(1+\gamma)}{n\sqrt{\delta n}}\cdot\sqrt{\log n}\cdot\|p\|_2+\frac{\bar{\alpha}^2}{\delta n}\cdot\log n\cdot\|p\|_2^2.\end{aligned}$$

Since $n = d^{1/\xi}$ and $\|p\|_2 \ge 1/\sqrt{n}$, it follows that

$$\begin{aligned}n\cdot\big(2L_2H_2+H_2^2\big) &\lesssim \frac{2\bar{\alpha}(1+\gamma)}{\sqrt{\delta}}\cdot\sqrt{\frac{\log d}{\xi}}\cdot\|p\|_2^2+\frac{\bar{\alpha}^2}{\delta}\cdot\frac{\log d}{\xi}\cdot\|p\|_2^2\\ &\lesssim \sqrt{\gamma}(1+\gamma)\cdot\sqrt{\frac{\bar{\alpha}\log d}{\xi}}+\gamma\cdot\frac{\bar{\alpha}\log d}{\xi}\cdot\|p\|_2^2 \qquad (\text{by } \gamma = \tfrac{\bar{\alpha}}{\delta})\\ &\lesssim \sqrt{\frac{\bar{\alpha}\log d}{\xi}}\cdot\|p\|_2^2,\end{aligned}$$

where the last inequality uses the assumptions that $\gamma \le O(1)$ and $(\bar{\alpha}\log d)/\xi \le 1$.

Combining the three terms and using $1/n \le \|p\|_2^2$ and $(\bar{\alpha}\log d)/\xi \le 1$, we conclude that

$$\begin{aligned}\|Pp\|_2^2 &\le \frac{1}{n}\cdot\Big(\delta+O\Big(\frac{\bar{\alpha}\log d}{\xi}\Big)\Big)+(1-\delta)\cdot\frac{5}{4n}+O\Big(\sqrt{\frac{\bar{\alpha}\log d}{\xi}}\Big)\cdot\|p\|_2^2\\ &= \frac{1}{n}\cdot\Big(\delta+(1-\delta)\cdot\frac{5}{4}\Big)+O\Big(\frac{\bar{\alpha}\log d}{\xi}\Big)\cdot\frac{1}{n}+O\Big(\sqrt{\frac{\bar{\alpha}\log d}{\xi}}\Big)\cdot\|p\|_2^2\\ &\le \frac{1}{n}\cdot\Big(\delta+(1-\delta)\cdot\frac{5}{4}\Big)+O\Big(\sqrt{\frac{\bar{\alpha}\log d}{\xi}}\Big)\cdot\|p\|_2^2.\end{aligned}$$

◀

3.2 Proof of Theorem 5

Proposition 6 allows us to derive the number of steps needed for the squared 2-norm of any initial distribution to drop to around C_δ/n.

▶ **Lemma 7.** *Let G be a d-regular graph with $d = n^\xi$ for some constant $\xi > 0$. If G satisfies the δ-small-set α-bipartite density condition for some $\alpha = d^{1-c} \le \xi^2 d/(\log d)^2$ and $\delta \gtrsim \alpha/d$, then*

$$\|P^t p_0\|_2^2 \le \frac{C_\delta}{n} + O\Big(\frac{1}{n^{1+c\xi/4}}\Big) \quad \textit{for any initial distribution } p_0 \textit{ and any } t \ge \frac{4}{c\xi} + 1.$$

Proof. Let $1/\beta := O\big(\sqrt{\alpha \log d/(\xi d)}\big)$ denote the drop rate in Proposition 6. For any probability distribution p, if $\|p\|_2^2 \lesssim C_\delta \beta/n$, then by Proposition 6,

$$\|Pp\|_2^2 \le \frac{C_\delta}{n} + \frac{1}{\beta} \cdot \|p\|_2^2 \le \frac{C_\delta}{n} + \frac{1}{\beta} \cdot O\Big(\frac{C_\delta \beta}{n}\Big) \le O\Big(\frac{C_\delta}{n}\Big).$$

This implies that after the next step of random walks,

$$\begin{aligned}\|P^2 p\|_2^2 \le \frac{C_\delta}{n} + \frac{1}{\beta} \cdot \|Pp\|_2^2 &\le \frac{C_\delta}{n} + \frac{1}{\beta} \cdot O\Big(\frac{C_\delta}{n}\Big)\\ &= \frac{C_\delta}{n} + O\bigg(\sqrt{\frac{\alpha \log d}{\xi d}} \cdot \frac{C_\delta}{n}\bigg)\\ &= \frac{C_\delta}{n} + O\bigg(\sqrt{\frac{\log d}{\xi d^c}} \cdot \frac{1}{n}\bigg).\end{aligned}$$

Note that our assumption $\alpha = d^{1-c} \le \xi^2 d/(\log d)^2$ implies that $d^{\frac{c}{2}} \ge (\log d)/\xi$, and thus

$$\|P^2 p\|_2^2 \le \frac{C_\delta}{n} + O\bigg(\sqrt{\frac{\log d}{\xi d^c}} \cdot \frac{1}{n}\bigg) \le \frac{C_\delta}{n} + O\Big(\frac{1}{d^{c/4}} \cdot \frac{1}{n}\Big) = \frac{C_\delta}{n} + O\Big(\frac{1}{n^{c\xi/4+1}}\Big),$$

where the last equality follows by the assumption $d = n^\xi$. Hence, for distribution p that is already close to stationary distribution (i.e., $\|p\|^2 \lesssim C_\delta \beta/n$), its squared 2-norm drops to $C_\delta/n + O(1/n^{c\xi/4+1})$ in two steps.

On the other hand, if $\|p\|^2 \ge C_\delta \beta/n$, there is a large drop rate such that

$$\|Pp\|_2^2 \le \frac{C_\delta}{n} + \frac{1}{\beta} \cdot \|p\|_2^2 \le \frac{2}{\beta} \cdot \|p\|_2^2.$$

This implies that for any initial distribution p_0, $\|P^t p_0\|_2^2 \lesssim C_\delta \beta/n$ for $t \ge 4/c\xi - 1$ because

$$\|P^t p_0\|_2^2 \le \Big(\frac{2}{\beta}\Big)^t \cdot \|p_0\|_2^2 \le \Big(\frac{2}{\beta}\Big)^{\frac{4}{c\xi}-1} = \frac{\beta}{2} \cdot O\bigg(\sqrt{\frac{\log d}{\xi d^c}}\bigg)^{\frac{4}{c\xi}} \lesssim \frac{\beta}{2} \cdot \Big(\frac{1}{d^{c/4}}\Big)^{\frac{4}{c\xi}} \le \frac{\beta}{n} \le \frac{C_\delta \beta}{n},$$

where the third last inequality uses $d^{\frac{c}{2}} \ge (\log d)/\xi$ that we derived above.

To summarize, $\|P^t p_0\|_2^2 \lesssim C_\delta \beta/n$ after $4/c\xi - 1$ steps of random walks, and the lemma follows after two more steps of random walks using the calculation in the first paragraph. ◀

We are ready to prove Theorem 5.

Proof of Theorem 5. Let p be an arbitrary probability distribution. Note that an upper bound on $\|p\|_2^2$ implies an upper bound on $d_{TV}(p,\pi)$ by Cauchy-Schwarz:

$$d_{\mathrm{TV}}(p,\pi) = \frac{1}{2}\|p-\pi\|_1 \le \frac{\sqrt{n}}{2}\|p-\pi\|_2 = \frac{\sqrt{n}}{2}\sqrt{\|p\|_2^2 - 2\langle p,\pi\rangle + \|\pi\|_2^2} = \frac{\sqrt{n}}{2}\sqrt{\|p\|_2^2 - \frac{1}{n}}. \tag{15}$$

By Lemma 7, for any initial distribution p_0 and any $t \ge 4/c\xi + 1$,

$$\|P^t p_0\|_2^2 \le \frac{C_\delta}{n} + O\Big(\frac{1}{n^{1+c\xi/4}}\Big).$$

Combining this with (15) and the fact that $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$, we conclude that

$$\begin{aligned} d_{\mathrm{TV}}\big(P^t p_0, \pi\big) \le \frac{\sqrt{n}}{2}\sqrt{\frac{C_\delta - 1}{n} + O\Big(\frac{1}{n^{1+c\xi/4}}\Big)} &\le \frac{\sqrt{n}}{2}\bigg(\sqrt{\frac{C_\delta - 1}{n}} + O\Big(\frac{1}{n^{1/2+c\xi/8}}\Big)\bigg) \\ &= \frac{\sqrt{1-\delta}}{4} + O\Big(\frac{1}{n^{c\xi/8}}\Big) \end{aligned}$$

where the last equality follows since $C_\delta - 1 = 5(1-\delta)/4 - (1-\delta) = (1-\delta)/4$. ◀

4 Lower Bounds on Mixing Time

We prove Theorem 4 in this section. The proof relies on the following lemma on variation distance.

▶ **Lemma 8.** *Let G be a d-regular graph. Suppose there exist $S,T \subseteq V$ such that $|E(S,T)| \ge d|S||T|/n + \alpha\sqrt{|S||T|}$ for some α. Then, for any $t \in \mathbb{N}$,*

$$d_{\mathrm{TV}}(t) \ge \frac{1}{2}\cdot\Big(\frac{\alpha}{2d}\Big)^{2t} - \frac{\min\{|S|,|T|\}}{2n}.$$

Proof. We argue that if S,T form a dense bipartite structure, then the random walks starting at $U_S := \chi_S/|S|$ (the uniform distribution on S) should bounce back and forth between S,T, causing slow mixing time.

Assume without loss of generality that S,T are minimal set that satisfies $|E(S,T)| \ge d|S||T|/n + \alpha\sqrt{|S||T|}$, and that $|S| = \min\{|S|,|T|\}$. First, we argue that a dense bipartite structure implies a lower bound on the degree for each $v \in S$ and $u \in T$. For any $v \in S$, by minimality,

$$\begin{aligned} |E(S,T)| - d_T(v) = |E(S\setminus\{v\},T)| &< \frac{d(|S|-1)|T|}{n} + \alpha\sqrt{(|S|-1)|T|} \\ &< \frac{d|S||T|}{n} + \alpha\sqrt{|S||T|} - \alpha\sqrt{|T|}(\sqrt{|S|} - \sqrt{|S|-1}) \\ &\le |E(S,T)| - \frac{\alpha}{2}\sqrt{\frac{|T|}{|S|}}, \end{aligned}$$

where the last inequality uses $\sqrt{x} - \sqrt{x-1} > \frac{1}{2\sqrt{x}}$. This implies a lower bound on $d_T(v)$ and similarly on $d_S(u)$ for $u \in T$ such that

$$d_T(v) \ge \frac{\alpha}{2}\sqrt{\frac{|T|}{|S|}} \quad \text{and} \quad d_S(v) \ge \frac{\alpha}{2}\sqrt{\frac{|S|}{|T|}}.$$

Let $d_{\min} := \min\left\{\frac{\alpha}{2}\sqrt{\frac{|T|}{|S|}}, \frac{\alpha}{2}\sqrt{\frac{|S|}{|T|}}\right\}$. Since $d_{\min} \le d$, it also follows that

$$|S| \ge \frac{\alpha^2}{4d^2}|T| \quad \text{and} \quad |T| \ge \frac{\alpha^2}{4d^2}|S| \quad \implies \quad d_{\min} \ge \frac{\alpha^2}{4d}.$$

Consider the random walk starting at U_S. After one step of random walks, since $d_S(u) \ge d_{\min}$ for all $u \in T$, $PU_S(u) \ge d_{\min}/(d|S|)$. Similarly, for any $v \in S$, after another step of random walks, $P^2U_S(v) \ge d_{\min}^2/(d^2|S|)$. By induction, after the t-th step, for all $v \in S$,

$$P^tU_S(v) \ge \left(\frac{d_{\min}}{d}\right)^t \cdot \frac{1}{|S|} \ge \left(\frac{\alpha^2}{4d^2}\right)^t \cdot \frac{1}{|S|}.$$

It follows from the definition of $d_{\mathrm{TV}}(t)$ that

$$d_{\mathrm{TV}}(t) \ge \frac{1}{2}\|P^tU_S - \pi\|_1 \ge \frac{1}{2}\sum_{v\in S}\left|P^tU_S(v) - \frac{1}{n}\right| \ge \frac{1}{2}\cdot|S|\cdot\left(\left(\frac{\alpha^2}{4d^2}\right)^t\cdot\frac{1}{|S|} - \frac{1}{n}\right) = \frac{1}{2}\cdot\left(\frac{\alpha}{2d}\right)^{2t} - \frac{|S|}{2n}.$$

◀

Theorem 4 follows immediately from Lemma 8.

Proof of Theorem 4. It follows from Lemma 8 and the assumption $|S|, |T| \le \delta n$ that for any $t \in \mathbb{N}$,

$$d_{\mathrm{TV}}(t) \ge \frac{1}{2}\cdot\left(\frac{\alpha}{2d}\right)^{2t} - \frac{\min\{|S|,|T|\}}{2n} \ge \frac{1}{2}\cdot\left(\frac{\alpha}{2d}\right)^{2t} - \frac{\delta}{2}$$

For $d_{\mathrm{TV}}(t)$ to be below $\frac{1}{n}$, we need

$$\frac{1}{2}\cdot\left(\frac{\alpha}{2d}\right)^{2t} - \frac{\delta}{2} \le \frac{1}{n} \quad \implies \quad t \ge \frac{\log(\delta + \frac{2}{n})}{2\log\alpha/(2d)} \gtrsim \frac{\log\delta}{\log(\alpha/d)} = \frac{\log(1/\delta)}{\log(d/\alpha)}.$$ ◀

4.1 Expanders and Small-Set Vertex Expanders

Some natural combinatorial conditions to consider for constant mixing time of a graph are its expansion properties, such as edge conductance and small-set vertex expansion.

We discuss why they are not strong enough to attain constant mixing time in this section. The graphs that we construct have a small dense bipartite structure embedded while having near optimal edge conductance or small-set vertex expansion. It follows from Theorem 4 that it does not have constant mixing time.

Counterexample for Expanders

Example. Let G be a d-regular graph with $d \ge 4$, where there exist two small disjoint sets $S, T \subseteq V$ such that $|S| = |T| = n/(d+2) \le n/d$. Each vertex in S has $d/2$ edges into T, and similarly each vertex in T has $d/2$ edges into S. The remaining $d/2$ edges of each $v \in S \cup T$ go to unique neighbours in $V \setminus S \cup T$. The induced subgraph $G[V \setminus S \cup T]$ forms a graph of degree $d - 1$ with edge conductance $1/2$.

▷ Claim 9. G has edge conductance at least $1/8$.

Proof. Let $W \subseteq V$. Let $W_1 = W \cap (S \cup T)$, and let $W_2 = W \cap (V \setminus S \cap T)$. By construction, $|\delta(W_1)| \geq \frac{d}{2}|W_1|$ and $|\delta(W_2)| \geq \frac{d-1}{2}|W_2|$. Assume $|W_1| \geq |W_2|$. Since $|E(W_1, W_2)| \leq |W_2|$ by construction, it follows that

$$|\delta(W)| \geq |\delta(W_1)| - |W_2| \geq \frac{d}{2} \cdot |W_1| - |W_1| \geq \frac{d}{4} \cdot |W_1| \geq \frac{d}{8} \cdot |W|,$$

where the second last inequality holds for any $d \geq 4$. Assume $|W_2| \geq |W_1|$, then

$$|\delta(W)| \geq |\delta_{V \setminus S \cup T}(W_2)| \geq \frac{d-1}{2} \cdot |W_2| \geq \frac{d}{4} \cdot |W_1| \geq \frac{d}{8} \cdot |W|.$$

We conclude that G has edge conductance at least 1/8. ◁

It remains to see that G has a small dense bipartite structure.

▷ Claim 10.

$$|E(S,T)| \geq \frac{d|S||T|}{n} + \frac{d}{4}\sqrt{|S||T|}$$

Proof. Since $|S| = |T| \leq n/d$, we have

$$\frac{d|S||T|}{n} + \frac{d}{4}\sqrt{|S||T|} \leq |S| + \frac{d}{4}|S| \leq \frac{d}{2}|S| = |E(S,T)|$$

where the last inequality follows for any $d \geq 4$. ◁

It follows from Theorem 4 that G does not have constant mixing time.

Counterexample for Small-Set Vertex Expanders

To see that near-optimal small set vertex expansion also fails to attain constant mixing time, consider the same graph construction, except now the induced subgraph $G[V \setminus S \cup T]$ is small-set vertex expander of size $dn/(d+2)$ and degree $d-1$, where each set of size at most $dn/((d+2)(d-1))$ has vertex expansion at least $d/2$. The small dense bipartite structure (S,T) still exists in G, so G does not have constant mixing time by the same argument above. It suffices to check that G has near optimal small-set vertex expansion.

▷ Claim 11. For every subset W of size at most n/d, its vertex expansion in G is at least $d/8$.

Proof. Let $W \subseteq V$ with $|W| \leq n/d$, define $N(W) = \{u \in V \mid \exists v \in W \text{ with } uv \in E\}$ as the neighbor set of W. We prove the claim by lower bounding $N(W)$. Let $W_1 = W \cap (S \cup T)$ and $W_2 = W \cap (V \setminus S \cap T)$. By construction, $|N(W_1)| \geq d|W_1|/2$ and $|N(W_2)| \geq d|W_2|/2$. Assume $|W_1| \geq |W_2|$. Since $|N(W_1)| \geq d|W_1|/2$ and $2|W_1| \geq |W_1| + |W_2| = |W|$,

$$|N(W)| \geq |N(W_1)| \geq \frac{d}{4} \cdot 2|W_1| \geq \frac{d}{4} \cdot |W|.$$

The case where $|W_2| \geq |W_1|$ follows similarly.

Therefore, we can lower bound the vertex expansion as

$$|\partial(W)| \geq |N(W)| - |W| \geq (\frac{d}{4} - 1)|W| \geq \frac{d}{8} \cdot |W|$$

where the last inequality holds for any $d \geq 8$. ◁

References

1 N. Alon. Eigenvalues and expanders. *Combinatorica*, 6:83–96, 1986. Theory of computing (Singer Island, Fla., 1984). `doi:10.1007/BF02579166`.

2 N. Alon and F. R. K. Chung. Explicit construction of linear sized tolerant networks. In *Proceedings of the First Japan Conference on Graph Theory and Applications (Hakone, 1986)*, volume 72, pages 15–19, 1988. `doi:10.1016/0012-365X(88)90189-6`.

3 N. Alon and V. D. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *J. Combin. Theory Ser. B*, 38(1):73–88, 1985. `doi:10.1016/0095-8956(85)90092-9`.

4 Yonatan Bilu and Nathan Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006. `doi:10.1007/s00493-006-0029-7`.

5 Jeff Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Sympos. in honor of Salomon Bochner, Princeton Univ., Princeton, N.J., 1969)*, pages 195–199. Princeton Univ. Press, Princeton, NJ, 1970.

6 Bernhard Haeupler, Jonas Huebotter, and Mohsen Ghaffari. A cut-matching game for constant-hop expanders. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1651–1678. SIAM, Philadelphia, PA, 2025. `doi:10.1137/1.9781611978322.51`.

7 Bernhard Haeupler, Harald Räcke, and Mohsen Ghaffari. Hop-constrained expander decompositions, oblivious routing, and distributed universal optimality. In *STOC '22—Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1325–1338. ACM, New York, [2022] ©2022.

8 David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson. `doi:10.1090/mbk/058`.

9 A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988. `doi:10.1007/BF02126799`.

10 A. Nilli. On the second eigenvalue of a graph. *Discrete Math.*, 91(2):207–210, 1991. `doi:10.1016/0012-365X(91)90112-F`.

11 R. Michael Tanner. Explicit concentrators from generalized N-gons. *SIAM J. Algebraic Discrete Methods*, 5(3):287–293, 1984. `doi:10.1137/0605030`.