ARTICLE

The diameter of the uniform spanning tree of dense graphs

Noga Alon[1,2], Asaf Nachmias[3] and Matan Shalev[3, *]

[1]Department of Mathematics, Princeton University, Princeton, NJ 08544, USA, [2]Schools of Mathematics and Computer Science, Tel Aviv University, Tel Aviv 69978, Israel, and [3]School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel
*Corresponding author. Email: matanshalev@mail.tau.ac.il

Abstract

We show that the diameter of a uniformly drawn spanning tree of a simple connected graph on n vertices with minimal degree linear in n is typically of order $\sqrt{n}$. A byproduct of our proof, which is of independent interest, is that on such graphs the Cheeger constant and the spectral gap are comparable.

1. Introduction

The uniform spanning tree of a finite connected graph G, denoted by $\mathsf{UST}(G)$, is a uniformly chosen random spanning tree of G. The main result of this paper is that the diameter of the UST, i.e., the largest distance between two vertices of the UST, on graphs with linear minimal degree grows like the square root of the number of vertices with high probability.

When G is the complete graph on n vertices, K_n, much more is known. A classical result of Szekeres [18] (see also [9]) explicitly provides the limiting distribution of the diameter of $\mathsf{UST}(K_n)$ scaled by $n^{-1/2}$. This was greatly extended by the influential work of Aldous [1–3] and Le Gall [12,13] who proved that $\mathsf{UST}(K_n)$, viewed as a random metric space and scaled by $n^{-1/2}$, converges in distribution with respect to the Gromov-Hausdorff distance to a canonical random compact metric space known as the Continuum Random Tree [1].

The UST is a critical statistical physics model, hence it is expected that as long as the base graph G is 'high dimensional', $\mathsf{UST}(G)$ should have a similar geometry to that of $\mathsf{UST}(K_n)$. This high dimensionality condition is typically some good isoperimetric condition. This has been pursued in [16] where the authors show that the diameter of $\mathsf{UST}(G)$ is of order $\sqrt{n}$ for a large class of high dimensional graphs including, for example, $\mathbb{Z}_n^5$, the hypercube $\{0, 1\}^n$ and regular expanders. The dense graphs we study in this paper, however, can be very far from being high dimensional. For instance, two cliques on $n/2$ vertices connected by an edge will have the worst isoperimetric inequality, yet the diameter of its UST is still of order $\sqrt{n}$. We now state our main result. For a connected graph H we write $\mathrm{diam}(H)$ for the maximal graph distance in H between any two vertices.

Theorem 1.1. *For any $\varepsilon, \delta \in (0, 1)$ there exists $C = C(\delta, \varepsilon) \in (1, \infty)$ such that if G is a connected simple graph on n vertices with minimal degree at least δn, then,*

$$\mathbb{P}\left(C^{-1}\sqrt{n} \leq \operatorname{diam}(\mathsf{UST}(G)) \leq C\sqrt{n}\right) \geq 1 - \varepsilon,$$

The main tool we use is a decomposition theorem (Lemma 2.2) which can be thought of as Szemerédi-type Regularity Lemma allowing to partition the vertices of G into $O(1)$ sets such that the induced graph on each set satisfies a sufficiently strong isoperimetric inequality and such that the number of edges connecting two such sets is sufficiently small. This partition is then used to study the behaviour of the loop-erased random walk on G which in turn provides estimates on the UST via Wilson's algorithm (Section 1.1).

It turns out that one can get significant mileage in the study of the random walk using such a decomposition theorem. One such estimate, which we believe is of independent interest, is an improvement to Cheeger's inequality on graphs of linear minimal degree. This improved inequality shows that on such graphs the Cheeger constant and the spectral gap are comparable. Denote by P the transition matrix of the simple random walk on G, and let $\pi(v) = \deg(v)/2|E(G)|$ denote its stationary distribution. Since P is self-adjoint in $L^2(\pi)$ it has n real eigenvalues in $[-1, 1]$ denoted by

$$1 = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq -1.$$

A classical highly useful inequality proved by Alon-Milman [4, 5], Lawler-Sokal [11] and Jerrum-Sinclair [7] known as Cheeger's inequality relates the spectral gap $\gamma(G) := 1 - \lambda_2$ of P with its isoperimetric constant (also known as Cheeger's constant). More precisely, for a set of vertices S of G we denote its volume by $\operatorname{Vol}(S) = \sum_{v \in S} \deg(v)$ and its edge boundary by $\partial S = \{(u, v) \in E(G) \mid u \in S, v \notin S\}$. We define the Cheeger constant as

$$\Phi(G) := \min_{S, \pi(S) \leq 1/2} \frac{|\partial S|}{\operatorname{Vol}(S)}$$

Cheeger's inequality states that

$$\Phi(G)^2/2 \leq \gamma(G) \leq 2\Phi(G). \tag{1}$$

When G is a simple graph of linear minimal degree we can improve the lower bound in Cheeger's inequality to match the order of the upper bound.

Theorem 1.2. *For any $\delta \in (0, 1)$ there exists a constant $c(\delta) > 0$ such that the following holds. Let $G = (V, E)$ be a simple graph on n vertices with minimal degree at least δn and Cheeger constant $\Phi(G)$, then*

$$\gamma(G) \geq c(\delta)\Phi(G).$$

Remark 1.3. Our proof gives $c(\delta) = \delta^{19}/2^{34}$ but we have not tried to optimize this constant.

Remark 1.4. After posting this paper we learned from Majid Farhadi, Suprovat Ghoshal, Anand Louis, and Prasad Tetali of an alternate proof of Theorem 1.2, which gives $c(\delta) = \Omega(\delta)$. Since the proofs are completely different we believe there is value in presenting both. We emphasize that our proof of Theorem 1.2 is just a byproduct of the tools we develop to prove our main result Theorem 1.1 and is also quite short given these tools. It is presented in Section 2.3.

The alternate proof follows from Theorem 1 of [10]. In our notation, it states that there exists a universal constant $C > 0$ such that for all $k \leq n$,

$$\Phi(G) \leq \frac{Ck\gamma(G)}{\sqrt{1 - \lambda_k}}. \tag{2}$$

To obtain Theorem 1.2 from (2), observe that when the minimal degree of G is at least δn, each diagonal term of P^2 is bounded from above by $1/(\delta n)$ and therefore the trace of P^2 is bounded from above by $1/\delta$. Since the trace of P^2 equals the sum of squares of the eigenvalues of P, we have

that at most $4/\delta$ eigenvalues of P are larger than $1/2$. Hence, $\lambda_{\lceil 4/\delta \rceil} \leq 1/2$. Plugging this into (2) and rearranging gives

$$\gamma(G) \geq c'\delta\Phi(G)$$

for some $c' > 0$.

1.1 Preliminaries

For a finite graph $G = (V, E)$ and a vertex $v \in V$ we denote by $\deg_G(v)$ its degree. When $U \subseteq V$ we will write $\deg_G(v, U)$ for the number of edges between v and U and $G[U]$ for the subgraph of G induced by U. We will sometimes omit the subscript when it is obvious to which G we refer. A network (G, w) is a connected graph $G = (V, E)$ endowed with a non-negative function $w : E \to [0, \infty)$ on its edges. The simple random walk $\{X_t\}_{t=0}^{\infty}$ on G is the Markov chain on the state space V that at each step moves along a uniformly chosen edge incident to it. Similarly, the simple random walk on a network (G, w) is the Markov chain such that the transition probability from v to u is proportional to $w(\{v, u\})$. In order to avoid issues of parity, we will sometimes consider the lazy random walk. Formally, at each step, with probability $1/2$ the walker stays put and otherwise chooses a neighbour uniformly (or proportionally to $w(\{v, \cdot\})$). We will often consider random walks with different starting distributions. When μ is a probability measure on V, we will use the notation $\mathbb{P}_\mu$ for the probability measure conditioned on $X_0 \sim \mu$. We will also use $\mathbb{P}_v$ for the walk conditioned on $X_0 = v$. Also, for a non-negative integer $t \geq 0$ and two vertices $v, u \in V$ we write $\mathbf{p}^t(v, u)$ for $\mathbb{P}_v(X_t = u)$. For any $a, b \in [0, \infty)$, we write $X[a, b]$ for $\langle X_i \rangle_{i \in I}$ where $I = \left[\lceil a \rceil, \lfloor b \rfloor\right] \cap \mathbb{N}$. Similarly, we write $X[a, b)$ if we wish to exclude b from I.

We will frequently use some facts about the *mixing time* of the random walk on G which we now define. The total variation distance between two probability measures μ, ν on V is

$$d_{\mathrm{TV}}(\mu, \nu) := \frac{1}{2} \sum_{u \in V} |\mu(v) - \nu(v)|.$$

For every $\varepsilon \in (0, 1/2)$, the ε-mixing time of G is defined by

$$t_{\mathrm{mix}}^G(\varepsilon) := \max_{v \in V} \min\left\{t \geq 0 : \|\mathbf{p}^t(v, \cdot) - \pi(\cdot)\|_{\mathrm{TV}} < \varepsilon\right\},$$

where $\pi(v) = \deg(V)/2|E|$, the stationary distribution of the random walk on G. To avoid issues of periodicity we emphasize that in this paper the quantity $t_{\mathrm{mix}}^G(\varepsilon)$ is defined *only* for the lazy random walk. We liberally vary the choice of ε throughout the proof; this changes the mixing time by at most a multiplicative constant. Indeed, for every $\varepsilon < 1/4$ and any integer, we have (see, [14, equations (4.34) and (4.32)])

$$t_{\mathrm{mix}}^G(\varepsilon) \leq \log_2(\varepsilon^{-1}) t_{\mathrm{mix}}(1/4) \qquad \text{and} \qquad t_{\mathrm{mix}}^G(\varepsilon^k) \leq k t_{\mathrm{mix}}(\varepsilon/2). \tag{3}$$

The uniform spanning tree (UST) of G is the uniform measure over the set of all spanning trees of G. More generally, when (G, w) is a finite network, we denote by $\mathsf{UST}(G)$ the weighted uniform spanning tree. That is, the probability measure supported on spanning trees of G that assigns to each such tree T a measure proportional to $\prod_{e \in T} w(e)$. We briefly describe here some useful properties of the UST involving sampling, conditioning and stochastic domination and refer the reader to [15, Chapter 4] for a comprehensive overview.

Our analysis of the UST will rely on Wilson's algorithm [19] for efficiently sampling the UST. This popular algorithm is frequently used not just to sample but rather to prove theorems about the UST, see [15]. Let $G = (V, E)$ be a finite connected graph. A walk of length L on G is a sequence of vertices $(X_0, \ldots, X_L)$ such that $(X_i, X_{i+1}) \in E$ for every $0 \leq i < L$. Given such a walk X, its loop erasure $\mathsf{LE}(X)$ is a sequence of vertices defined as follows. We put $\mathsf{LE}(X)_0 = X_0$ and inductively, for every $i > 0$ and given $\mathsf{LE}(X)[0, i-1]$, define $s_i := \max\{t \leq L \mid X_t = \mathsf{LE}(X)_{i-1}\}$. If $s_i = L$, the loop-erased random walk of X is $(\mathsf{LE}(X)_0, \ldots, \mathsf{LE}(X)_{i-1})$. Otherwise, let $\mathsf{LE}(X)_i = X_{s_i+1}$. In words, we walk along $(X_0, \ldots, X_L)$ and erase the loops as they are formed. Given two vertices v, u the

loop-erased random walk from v to u is defined to be $\mathsf{LE}(X)$ where X is the simple random walk started at v and terminated upon when hitting u. In a similar fashion we define the loop-erased random walk from a vertex v to a subset of vertices U. Wilson's algorithm works as follows. Choose any ordering $(v_1, \ldots, v_n)$ of the vertices of G and let T_1 be the empty tree containing v_1 and no edges. At each step $i > 1$, run a loop-erased random walk from v_i to T_{i-1} let T_i be the union of this loop-erased random walk and T_{i-1}. This process terminates after going through all vertices and results a spanning tree T_n. A remarkable theorem of Wilson [19], that we use throughout this paper, states that T_n is distributed as $\mathsf{UST}(G)$.

Next, it is very simple to prove that conditioning on the existence or absence of edges in the UST results in a UST on the graph obtained from G by contracting or erasing those edges, respectively.

Lemma 1.5. [15, Section 4.2] *Let (G, w) be a network and let e be an edge of G. The UST of (G, w) conditioned to contain e is distributed as the union of e with the UST of the network obtained from G by contracting the edge e to a single vertex.*

Lastly, we recall a few highly useful corollaries to a result of Feder and Mihail [6]. Given two probability measures μ_1 and μ_2 on 2^E, we say that μ_1 is stochastically dominated by μ_2 if there exists a probability measure μ on $2^E \times 2^E$ with marginals μ_1 and μ_2 which is supported on

$$\{(T_1, T_2) \in 2^E \times 2^E \mid T_1 \subseteq T_2\}.$$

Lemma 1.6. [15, Lemma 10.3] *Let G be a connected subgraph of a finite connected graph H. Then, $\mathsf{UST}(H) \cap E(G)$ is stochastically dominated by $\mathsf{UST}(G)$ when both are viewed as probability measures on $2^{E(G)}$.*

This lemma can be further generalized to our needs. We say that a network (G, w) is a sub-network of (H, w') if $V(G) \subseteq V(H)$ and for every edge (v, u) with $w(v, u) \neq 0$ we have $w(v, u) = w'(v, u)$. The same proof of [15, Lemma 10.3] yields a more general statement.

Lemma 1.7. *Let (G,w) be a subnetwork of a finite network (H, w'). Then, $\mathsf{UST}(H) \cap E(G)$ is stochastically dominated by $\mathsf{UST}(G)$ when both are viewed as probability measures on $2^{E(G)}$.*

Given a network G and a subset of vertices A we write G/A for the network obtained from G by contracting the vertices of A to a single vertex and keeping all edges. The following is well known and can be obtained by a similar argument to the proof of [15, Lemma 10.3].

Lemma 1.8. *Let (G, w) be a finite network and let $A \subseteq B$ be two sets of vertices of G. Then, $\mathsf{UST}(G/A)$ stochastically dominates $\mathsf{UST}(G/B)$.*

1.2 Proof outline and organization

It is easier to bound the diameter of the UST after conditioning on a long path in it. Indeed, a key lemma from [16] (see Lemma 4.1) roughly states that if a vertex set $W \subset V$ is sufficiently spread out in the sense that the random walk is unlikely to avoid it, then one can upper bound the probability that the diameter of $\mathsf{UST}(G/W)$ is much larger than $|W|$, where G/W is the graph obtained from G by identifying W to a single vertex. When G, say, is a regular expander (or any other 'high dimensional' graph) the approach in [16] is to take W to be the vertices on the unique path in $\mathsf{UST}(G)$ between two independently drawn uniform vertices of G. The expansion property is then used to show that this set has size $\Theta(\sqrt{n})$ and is sufficiently spread out, so Lemma 4.1 implies that $\mathsf{UST}(G/W)$ has diameter $\Theta(\sqrt{n})$.

The high level approach in this paper is to use our decomposition theorem (Lemma 2.2, proved in Section 2) and partition the graph into $O(1)$ sets so that with high probability the UST path between two random vertices in each set remains within the set, is of size $\Theta(\sqrt{n})$ and is sufficiently spread out within the set. Formalizing and proving this is performed in Section 3. In Section 4 we take the union of these $O(1)$ paths to be our set W and apply Lemma 4.1 from [16] to obtain that

UST(G/W) has diameter roughly of order $\sqrt{n}$ from which we deduce the desired upper bound on the diameter of UST(G).

2. Decomposition of linear minimal degree graphs

In this section we prove that any finite connected graph $G = (V, E)$ on n vertices with linear minimal degree can be decomposed into k sets, each of them linear in the number of vertices, such that a random walk typically mixes inside every such set before leaving it. Hence, when considering short times, roughly $\sqrt{n}$ steps of the walk, a random walk on the graph can be approximated well by a random walk on one of its sets in this decomposition. We will denote a partition of V by $\mathcal{P}$ and sometimes more explicitly by $V = V_1 \sqcup \ldots \sqcup V_k$. We write $[k]$ for the set $\{1, \ldots, k\}$.

Definition 2.1. Let $\varepsilon \in (0, 1)$, $\delta \in (0, 1]$ and $\beta > 0$ be fixed and let $G = (V, E)$ be a graph on n vertices with minimal degree at least δn. We say that a partition $V = V_1 \sqcup \ldots \sqcup V_k$ is an $(\varepsilon, \delta, \beta)$-good decomposition if there exists some $\theta \in [\varepsilon^{11 \cdot 2^{2/\delta}}, \varepsilon]$ such that the following conditions are satisfied.

1. The number of sets in the decomposition, denoted by k, satisfies $k \leq 2/\delta$.
2. For every $i \in [k]$ we have $|V_i| \geq \frac{\delta n}{2}$.
3. For every $i \in [k]$, the spectral gap of $G[V_i]$ is at least $\frac{\delta^{15}\theta\beta}{2^{31}n^2}$.
4. For every $i \in [k]$ and each $v \in V_i$ we have that $\deg(v, V_i) \geq \frac{\delta^4 n}{40}$.
5. For every $i \in [k]$ we have $|E(V_i, V \setminus V_i)| \leq \varepsilon^9 \theta^2 \beta$.

The majority of this section is devoted to proving the following decomposition lemma which will be key in the proof of Theorem 1.1.

Lemma 2.2. *For every $\delta > 0$, there exists a constant $c = c(\delta) > 0$ such that the following holds. For every $\varepsilon \in (0, c)$, every $\beta \in (0, 240n^2/(\varepsilon\delta^4))$ and any simple graph $G = (V, E)$ on n vertices with minimal degree at least δn there exists an $(\varepsilon, \delta, \beta)$-good decomposition of G.*

We remark that the proof of Theorem 1.2 does not use this lemma, rather a simpler decomposition lemma, Lemma 2.11, which is also the first step in the proof of Lemma 2.2.

2.1 Preliminary estimates on the spectral gap

In this subsection we prove the following lemma allowing us to lower bound the spectral gap of a decomposable graph.

Definition 2.3. Given a partition $\mathcal{P}$ of V, denoted by $V = V_1 \sqcup \ldots \sqcup V_k$, and some $c > 0$, we define the graph $H(\mathcal{P}, c)$ as follows. The vertices of $H(\mathcal{P}, c)$ are $[k]$ where each vertex $i \in [k]$ represents a set V_i of $\mathcal{P}$ and we join an edge (i, j) if $|E(V_i, V_j)| > c$.

Lemma 2.4. *Let $G = (V, E)$ be a simple graph on n vertices. Let $\mathcal{P}$ be a partition of V denoted by $V = V_1 \sqcup \ldots \sqcup V_k$. Assume that there exists a, b, c > 0 such that the following conditions hold.*

- *For every $i \in [k]$, the spectral gap of $G[V_i]$ is larger than a,*
- *For every $i \in [k]$ and every $v \in V_i$ we have $\deg(v, V_i) \geq b$,*
- *The graph $H(\mathcal{P}, c)$ is connected.*

Then, the spectral gap of G is at least $\min\left\{a, \frac{abc}{6kn^3}\right\}$.

Lemma 2.4 is an application of the main result of [8] together with some quick estimates involving the Dirichlet form (see [14, Chapter 13] for further reading on the Dirichlet form). In the rest of this subsection we cite and prove these necessary background results, then prove the lemma. Since the proof digresses from the main ideas of this paper, the reader may want to take this lemma as a 'black box' and skip reading its proof. We will typically use this Lemma when a is roughly a constant, b is of order n and c is of order $\Phi(G)n^2$.

Let (G, w) be a network where $G = (V, E)$ with a partition of its vertex set $V = V_1 \sqcup \ldots \sqcup V_k$. Let π be the stationary measure of the simple (or lazy) random walk on (G, w). For such a network with a partition of its vertex set to k sets, we define the distribution $\overline{\pi}$ on $[k]$ by setting $\overline{\pi}(i) = \sum_{v \in V_i} \pi(v)$. The projection chain is a Markov chain on $[k]$ with the following transition probabilities

$$\overline{P}_{i,j} = \frac{1}{\overline{\pi}(i)} \sum_{v \in V_i, u \in V_j} \pi(v) P(v, u). \tag{4}$$

Note that $\overline{\pi}$ is the stationary distribution of this chain. Furthermore, we define k restriction chains, to which we will also refer as the restriction walks. For every $i \in [k]$, this restriction walk is a Markov chain on V_i with transition probabilities

$$P_i(x, y) = \begin{cases} P(x, y) & x \neq y, \\ 1 - \sum_{w \in V_i \setminus \{x\}} P(x, w) & x = y. \end{cases}$$

We are now ready to state a weaker version of [8, Theorem 1] which will be useful later. We remark that our version follows easily from [8, Theorem 1] by applying trivial upper bounds to the spectral gap of the projection chain and to the probability to move from one set in the decomposition to another one.

Theorem 2.5. ([8, Theorem 1]) *Let (G, w) be a network where $G = (V, E)$ and let $V = V_1 \sqcup \ldots \sqcup V_k$ be a partition of its vertex set. Denote by $\bar{\gamma}$ the spectral gap of the projection chain associated with it and for every $i \in [k]$ let γ_i be the spectral gap of the restriction walk on V_i. Then, the spectral gap γ of the random walk on (G, w) satisfies*

$$\gamma \geq \min_i \frac{\bar{\gamma} \gamma_i}{6}.$$

For an irreducible Markov chain on a finite state space Ω with transition matrix P, stationary distribution π and a function $f : \Omega \to \mathbb{R}$, we denote

$$\mathcal{E}_\pi^P(f) := \frac{1}{2} \sum_{x,y \in \Omega} \pi(x) P(x, y) (f(x) - f(y))^2.$$

The following lemma which we will not prove is helpful in estimating spectral gaps of networks which are obtained by a small perturbation of another network.

Lemma 2.6. ([14, Lemma 13.8]) *Let P_0 and P_1 be transition matrices with stationary distributions π_0 and π_1 over the same finite state space Ω. Let γ^0 and γ^1 be their spectral gaps, respectively. If there exists $\alpha > 0$ such that for all functions $f : \Omega \to \mathbb{R}$ we have $\mathcal{E}_{\pi^1}^{P_1}(f) \leq \alpha \mathcal{E}_{\pi^0}^{P_0}(f)$, then*

$$\gamma^1 \leq \left(\max_{\omega \in \Omega} \frac{\pi^0(\omega)}{\pi^1(\omega)} \right) \alpha \gamma^0.$$

Claim 2.7. *Let $G = (V, E)$ and let $W_0 = (G, w_0)$ and $W_1 = (G, w_1)$ be two networks such that there exists a vertex $v \in V$ such that $w_{v,v}^0 < w_{v,v}^1$ and for every other edge $(u, w) \neq (v, v)$ we have*

$w^0_{u,w} = w^1_{u,w}$. *Denote by γ^0 and γ^1 the spectral gaps corresponding to W_0 and W_1, respectively. Then, $\gamma^1 \leq \gamma^0$.*

Proof For $i \in \{0, 1\}$, we let P_i be the transition matrix corresponding to W_i. We denote

$$w^i_u = \sum_{e \ni u} w^i_e, \qquad Z_i = \sum_{u \in V} w^i_u.$$

We also denote by π^i the stationary distribution of W_i and recall that $\pi^i(u) = w^i_u / Z_i$. We will now use Lemma 2.6 to show that $\gamma^1 \leq \gamma^0$. A simple calculation shows that for every f

$$\mathcal{E}^{P_1}_{\pi^1}(f) = \frac{1}{2} \sum_{u,v \in V} \frac{w^1_{u,v}}{Z_1}(f(u) - f(v))^2 = \frac{Z_0}{Z_1}\mathcal{E}^{P_0}_{\pi^0}(f). \tag{5}$$

Also, we write $\varepsilon = w^1_{vv} - w^0_{vv} > 0$. Then, for every $u \in V$ we have

$$\frac{\pi^0(u)}{\pi^1(u)} = \frac{\frac{w^0_u}{Z_0}}{\frac{w^1_u}{Z_1}} = \begin{cases} \dfrac{Z_1}{Z_0} & u \neq v \\[2ex] \dfrac{Z_1}{Z_0} \cdot \dfrac{w^0_v}{w^0_v + \varepsilon} & u = v \end{cases}.$$

Hence, in any case $\pi^0(u)/\pi^1(u) \leq Z_1/Z_0$. By (5), we can use Lemma 2.6 with $\alpha = Z_0/Z_1$. We thus get that $\gamma^1 \leq \gamma^0$. $\qquad \square$

As mentioned in the last subsection, if P is a transition matrix of some random walk and Q is the transition matrix of its lazy version, then $Q = \frac{1}{2}(I + P)$ and hence $\gamma(Q) = \frac{1}{2}\gamma(P)$. Similarly, if X is a random walk with transition P and Y is an α-lazy random version of X, that is, at each step the walker stays put with probability α and otherwise chooses its next step according to P, then $Q = \alpha I + (1 - \alpha)P$. Hence, $\gamma(Q) = (1 - \alpha)\gamma(P)$.

We can further generalize this. For every $v \in V$, let $p_v \in [0, 1)$. We call $(p_v)_{v \in V}$ the lazy vector. Let X be some random walk on V with transition matrix P and let Y be the following random walk on V. At each step, if the walker is at some $u \in V$, it stays put with probability p_u and otherwise chooses its next step according to P. The next claim shows that we can lower bound the spectral gap associated with this random walk.

Claim 2.8. *Let P be a transition matrix of a random walk X on a finite connected graph G with spectral gap $\gamma(P)$. Let $\alpha \in (0, 1)$ and let $(p_v)_{v \in V} \in [0, \alpha)^V$ be a vector which we call the lazy vector. Let Y be the following random walk on G. If $Y_t = v$, stay put with probability p_v. Otherwise, choose X_{t+1} according to P. Let Q be the transition matrix of Y. Then, $\gamma(Q) \geq (1 - \alpha)\gamma(P)$.*

Proof Let (G, w^α) be the network associated with the graph G such that $w^\alpha_{uv} = 1$ if $(u, v) \in E$ and $w^\alpha_{vv} = \beta_v$ where β_v satisfies $\frac{\beta_v}{\deg(v) + \beta_v} = \alpha$. If P is the transition matrix of the simple random walk on G, then $Q := \alpha I + (1 - \alpha)P$ is the transition matrix corresponding to (G, w^α). Note that the spectral gap of Q satisfies $\gamma(Q) = (1 - \alpha)\gamma(P)$. Let $(p_v)_{v \in V}$ be the lazy vector with all values non-negative and smaller than α. The random walk that stays put at some $u \in V$ with probability p_u and jumps according to P otherwise can be seen as a random walk on a network (G, w^p) which can be constructed from (G, w^α) by going iteratively over all vertices $v \in V$ and decreasing the weight of w^α_{vv} at each step according to p_v. By Claim 2.7, at each step the spectral gap can be only increased. Hence, the spectral gap of the walk corresponding to (G, w^p) is larger than $\gamma(Q) = (1 - \alpha)\gamma(P)$, as required. $\qquad \square$

Another canonical method of bounding the spectral gap from below is *the path method*.

Claim 2.9. *(The path method, see [14, Corollary 13.21]) Let P be a transition matrix of a Markov chain on a finite state space Ω with stationary distribution π and spectral gap γ. For every $x, y \in \Omega$,*

denote $Q(x, y) = \pi(x)P(x, y)$. *Let* $\{\varphi_{x,y}\}_{x,y \in \Omega}$ *be a collection of paths in* Ω *from x to y such that every path* $\varphi_{x,y}$ *is a path from x to y with $Q(e) > 0$ for every $e \in \varphi_{x,y}$. Denote*

$$B := \max_e \frac{1}{Q(e)} \sum_{\varphi_{x,y} \ni e} \pi(x)\pi(y)|\varphi_{x,y}|.$$

Then, $\gamma \geq B^{-1}$.

Proof of Lemma 2.4. When $k = 1$, this is trivial and the spectral gap of G is at least a. We assume henceforth that $k \geq 2$ and we consider the projection and restriction chains associated with the decomposition of V to sets of $\mathcal{P}$ as described earlier in this section. Let $i \in [k]$ and consider the restriction chain associated with V_i, a set of $\mathcal{P}$. By our assumption, we have that the spectral gap of $G[V_i]$ is at least a. However, the restriction walk on V_i is different from the simple random walk on $G[V_i]$, as it is obtained from it by adding self loops for every edge that exits V_i. For every $v \in V$ we denote by p_v the probability to move from v to itself in the restriction chain. Since every $v \in V_i$ has $\deg(v, V_i) \geq b$, we have that $p_v \leq (1 - \frac{b}{n})$. We thus have by Claim 2.8 that γ_i, the spectral gap of the restriction chain, satisfies $\gamma_i \geq \frac{ab}{n}$.

We turn to the projection chain, which is a Markov chain on the state space $[k]$ with stationary distribution $\overline{\pi}$ and transition matrix $\overline{P}$ as in (4). We will use Claim 2.9 to bound the spectral gap associated with it from below. For every edge $e = (l, m) \in H(\mathcal{P}, c)$, we denote $Q(e) = \overline{\pi}(l)\overline{P}(l, m)$. We denote by P the transition matrix of the original simple random walk. By the definition in (4), we have

$$Q(e) = \overline{\pi}(l) \cdot \frac{1}{\overline{\pi}(l)} \sum_{u \in V_l, v \in V_m} \pi(u)P(u, v) = \sum_{u \in V_l} \frac{|E(u, V_m)|}{2|E|} \geq \frac{|E(V_l, V_m)|}{n^2} \geq \frac{c}{n^2}.$$

Since the graph $H(\mathcal{P}, c)$ is connected, for every $x, y \in [k]$ we can choose a path $\varphi_{x,y}$ connecting x and y such that every edge in this path has $Q(e) \geq c/n^2$. We choose such paths for every x, y arbitrarily. We obtain that for every edge e which belongs to any path in this choice $(\varphi_{x,y})_{x,y \in [k]}$ we have

$$\frac{1}{Q(e)} \sum_{\varphi_{x,y} \ni e} \overline{\pi}(x)\overline{\pi}(y)|\varphi_{x,y}| \leq \frac{kn^2}{c} \sum_{\varphi_{x,y} \ni e} \overline{\pi}(x)\overline{\pi}(y) \leq \frac{kn^2}{c}.$$

Hence, by Claim 2.9, the spectral gap of the projection chain $\overline{\gamma}$ is at least c/kn^2. Using Theorem 2.5, we conclude

$$\gamma \geq \min_{i \in [k]} \frac{\overline{\gamma}\gamma_i}{6} \geq \frac{abc}{6kn^3}. \qquad \square$$

2.2 Primary decomposition

Definition 2.10. Let $G = (V, E)$ be a graph on n vertices with minimal degree at least δn. A δ-primary decomposition of G is a partition of its vertices $V = V_1 \sqcup \ldots \sqcup V_k$ which has the following properties.

1. The number of sets in the decomposition, denoted by k, satisfies $k \leq 2/\delta$.

2. For every $i \in [k]$, we have $|V_i| \geq \frac{\delta n}{2}$.

3. For every $i \in [k]$ and each $v \in V_i$ we have that $\deg(v, V_i) \geq \frac{\delta^4 n}{40}$.

4. For every $i \in [k]$, the spectral gap of $G[V_i]$ is at least $\frac{\delta^{10}}{2^{22}}$.

Lemma 2.11. *Let $G = (V, E)$ be a simple graph on n vertices with minimal degree at least δn. Then, there exists a δ-primary decomposition of G.*

Proof We build the decomposition inductively each time refining the partition of V. We begin with the trivial partition $\{V\}$. At each step, if there is a subset W in the partition that can be further partitioned $W = W_1 \sqcup W_2$ such that $|E(W_1, W_2)| \leq \frac{\delta^3}{20}|W_1||W_2|$, then we refine the partition by replacing W with W_1, W_2. We call the edges $E(W_1, W_2)$ in each such refinement step negligible. The choice of W and W_1, W_2 is not necessarily unique and at each step we choose arbitrarily among all possibilities. Since the graph is finite this process must stop and we denote the final decomposition by $V = U_1 \sqcup \ldots \sqcup U_\ell$, where edges between any pair U_i and U_j are negligible. The sum of $|W_1||W_2|$, described above, over each of the ℓ refinement steps is no more than the cardinality of pairs of vertices, hence, the number of negligible edges is at most $\frac{\delta^3}{20}\binom{n}{2}$.

We call a vertex bad if the number of negligible edges touching it is larger than $\delta n/2$. Our bound on the number of negligible edges implies that there are no more than $\delta^2 n/10$ bad vertices. Every vertex which is not bad is called good. If a set U_i in the partition contains a good vertex we call it a good set, otherwise, a bad set. Since the minimal degree in the graph is larger than δn, every good vertex v touches at least $\delta n/2$ edges that are not negligible; the corresponding neighbours must be in the same set of the partition as v. Hence each good set is of size at least $\delta n/2$ and so their number is at most $2/\delta$. We call bad vertices belonging to bad sets evil. To obtain our primary decomposition, we remove all bad sets from the partition and redistribute the evil vertices among the good sets as follows. Assume without loss of generality that the good sets of the partition are $U_1, \ldots, U_k$ where $k \leq \ell$. Let $v \in V \setminus \cup_{i=1}^{k} U_i$ be an evil vertex. Since the number of good neighbours of v is at least $\delta n - \delta^2 n/10$ and $k \leq 2/\delta$, there exists some $i \in [k]$ for which $d(v, U_i) \geq \delta^2 n/3$. We add v to one such set chosen arbitrarily.

We denote the resulting decomposition by $V = V_1 \sqcup \ldots \sqcup V_k$ (with $U_i \subset V_i$ for all $i \in [k]$) and argue that it satisfies the desired conditions of Definition 2.10. Conditions (1) and (2) are immediate. To see that condition (3) is satisfied, let $i \in [k]$ and let $v \in V_i$. If v is good, then $\deg(v, V_i) \geq \delta n/2$. If v is bad but not evil, then $|E(\{v\}, U_i)| \geq \frac{\delta^3}{20}|U_i| \geq \delta^4 n/40$ since otherwise we would have partitioned U_i to $\{v\}$ and $U_i \setminus \{v\}$. If v is evil, then it was added to V_i since $d(v, U_i) \geq \delta^2 n/3$.

It remains to prove condition (4). Due to Cheeger's inequality (see equation (1)), it is enough to show that for every $i \in [k]$

$$\Phi(G[V_i]) \geq \frac{\delta^5}{1200}. \qquad (6)$$

Fix $i \in [k]$. We slightly abuse notation and write Vol and π for the volume and stationary measures on $G[V_i]$ respectively, that is, for any $S \subset V_i$ we have $\mathrm{Vol}(S) = \sum_{s \in S} \deg(s, V_i)$ and $\pi(S) = \mathrm{Vol}(S)/\mathrm{Vol}(V_i)$. Let $X \subset V_i$ be a subset with $\pi(X) \leq 1/2$. If at least half of the vertices of X are evil, then its size is at most twice the number of bad vertices, i.e. $|X| \leq \delta^2 n/5$. Each evil vertex of X has at least $\delta^2 n/3 - \delta^2 n/5$ of its neighbours in V_i outside X. Hence,

$$\frac{|\partial X|}{\mathrm{Vol}(X)} \geq \frac{\frac{2\delta^2 n}{15} \cdot \frac{|X|}{2}}{n|X|} = \frac{\delta^2}{15}.$$

Suppose otherwise that at least half of the vertices of X are non-evil. Denote the set of non-evil vertices of X by R and let T be the other non-evil vertices of V_i. Note that $R \sqcup T = U_i$. The number of edges between them is at least $\frac{\delta^3}{20}|R||T|$, since otherwise U_i would have been partitioned further. Thus,

$$\frac{|\partial X|}{\mathrm{Vol}(X)} \geq \frac{|E(R,T)|}{\mathrm{Vol}(X)} \geq \frac{\frac{\delta^3}{20}|R||T|}{\mathrm{Vol}(X)} \geq \frac{\frac{\delta^3}{20}|T||X|}{2n|X|} \geq \frac{\delta^3|T|}{40n}. \tag{7}$$

It remains to lower bound $|T|$. Denote by G_i and B_i the sets of good and bad vertices of V_i, respectively. We have that $|G_i| \geq \delta n/2 - \delta^2 n/10$. Each vertex $v \in G_i$ has $\deg(v, V_i) \geq \delta n/2$ hence $\mathrm{Vol}(G_i) \geq \delta^2 n^2/5$. On the other hand, since the total number of bad vertices is at most $\delta^2 n/10$ we have $\mathrm{Vol}(B_i) \leq \delta^2 n^2/10$. We deduce that $\pi(G_i) \geq 2/3$. Since $\pi(V_i \setminus X) \geq 1/2$, we have that $\pi(T) \geq \pi(G_i \cap (V_i \setminus X)) \geq 1/6$. We bound $\mathrm{Vol}(T) \leq |T|n$ and $\mathrm{Vol}(V_i) \geq \mathrm{Vol}(G_i) \geq \delta^2 n^2/5$ which with the last estimate gives $|T| \geq \delta^2 n/30$. We plug this into (7) to obtain that $\frac{|\partial X|}{\mathrm{Vol}(X)} \geq \frac{\delta^5}{1200}$, as required. $\qquad\square$

2.3 The Cheeger constant and spectral gap are comparable on graphs with linear degree

Proof of Theorem 1.2. For brevity we denote $\Phi(G) = r$. By Lemma 2.11, there exists a δ-primary decomposition $\mathcal{P}$ of G, also denoted by $V = V_1 \sqcup \ldots \sqcup V_k$. We claim that the graph $H(\mathcal{P}, r\delta^2 n^2/2k^2)$ is connected. Indeed, assume to the contrary that it is not connected and let $S \subseteq [k], S \neq [k]$, be a connected component in this graph. Let $V_S := \cup_{i \in S} V_i$ be the set corresponding to S in G. We may assume that $\pi(V_S) \leq 1/2$ (otherwise, we will take one of the connected components of $V_{[k]\setminus S}$). Since $\mathcal{P}$ is δ-primary and $\Phi(G) = r$, we have that

$$|\partial V_S| \geq r\mathrm{Vol}(V_S) \geq r\delta^2 n^2/2.$$

Hence, we can find a component $V_j \subseteq (V \setminus V_S)$ and a component $V_i \subseteq V_S$ with $|E(V_i, V_j)| \geq r\delta^2 n^2/2k^2$, contradicting the assumption that S is a connected component of $H(\mathcal{P}, r\delta^2 n^2/2k^2)$. We can therefore apply Lemma 2.4 and obtain that Theorem 1.2 holds with the constant $\frac{\delta^{19}}{2^{34}}$ (note that $\delta \leq 1$ so the minimum in the conclusion of Lemma 2.4 is attained in the second item). $\qquad\square$

2.4 Coarsening

Given two partitions $\mathcal{P}$ and $\mathcal{P}'$, we say that $\mathcal{P}$ is a coarsening of $\mathcal{P}'$ if every set in $\mathcal{P}$ is a union of sets in $\mathcal{P}'$. Suppose that $G = (V, E)$ has minimal degree at least δn. Let $\mathcal{P}$ be a partition denoted by $V = V_1 \sqcup \ldots \sqcup V_k$, and let $\mathcal{P}'$ be a partition $V = V'_1 \sqcup \ldots \sqcup V'_\ell$ such that $\mathcal{P}$ is a coarsening of $\mathcal{P}'$. For each $i \in [k]$ we write $\mathcal{P}_i$ for the partition of V_i into sets of $\mathcal{P}'$.

Definition 2.12. For $\varepsilon, \alpha \in (0, 1)$ and $\beta > 0$ we say that $\mathcal{P}$ is an $(\varepsilon, \alpha, \beta)$-good coarsening of a partition containing ℓ sets $\mathcal{P}'$, if there exists some $\theta \in \left[\varepsilon\left(\frac{\varepsilon\alpha}{\ell^2}\right)^{2^\ell}, \varepsilon\right]$ for which the following conditions are satisfied.

1. For every $i \in [k]$ we have that $H(\mathcal{P}_i, \theta\beta)$ is connected (H is defined in Definition 2.3).
2. For every $i \in [k]$ we have $|E(V_i, V \setminus V_i)| \leq \theta^2 \beta\alpha$.

Lemma 2.13. *For any $\varepsilon, \alpha \in (0, 1)$ and any $\beta > 0$, if $G = (V, E)$ is a finite graph and $\mathcal{P}'$ is a partition of V containing ℓ sets, then there exists an $(\varepsilon, \alpha, \beta)$-good coarsening of $\mathcal{P}'$.*

Proof. Let $\varepsilon > 0$. We will construct $\mathcal{P}$, an $(\varepsilon, \alpha, \beta)$-good coarsening of $\mathcal{P}'$, which will satisfy conditions (1) and (2) of Definition 2.12 with some parameter θ. At first, if $\mathcal{P}'$ satisfies condition (2) with ε playing the role of θ, then $\mathcal{P}'$ is an $(\varepsilon, \alpha, \beta)$ good coarsening of itself. Otherwise, we build recursively a finite sequence of length at most ℓ of coarsenings $\langle \mathcal{P}_i \rangle$ of $\mathcal{P}'$ and parameters $\langle \theta_i \rangle$, such that $\mathcal{P}_i$ satisfies condition (1) with θ_i. We set $\mathcal{P}_1 := \mathcal{P}'$ and $\theta_1 = \varepsilon$.

At step $m > 1$, we are given with $\mathcal{P}_{m-1}$ and θ_{m-1} such that $\mathcal{P}_{m-1}$ satisfies condition (1) with parameter θ_{m-1}. If $\mathcal{P}_{m-1}$ also satisfies condition (2) with θ_{m-1}, then $\mathcal{P}_{m-1}$ is an $(\varepsilon, \alpha, \beta)$ good coarsening and we halt the process, denoting $\mathcal{P} := \mathcal{P}_{m-1}$ and $\theta := \theta_{m-1}$. Otherwise, there exists a set U in the partition $\mathcal{P}_{m-1}$ which has $|E(U, V \setminus U)| \geq \theta_{m-1}^2 \alpha \beta$. Since there are ℓ sets in $\mathcal{P}'$, there exists at least one pair of sets $W_1 \subseteq U$ and $W_2 \subseteq V \setminus U$, both are sets of the partition $\mathcal{P}'$, such that $|E(W_1, W_2)| \geq \theta_{m-1}^2 \alpha \beta / \ell^2$. We denote

$$\theta_m = \frac{\alpha}{\ell^2} \theta_{m-1}^2 \tag{8}$$

and form $\mathcal{P}_m$ by replacing U and the set containing W_2 in $\mathcal{P}_{m-1}$ with their union. We note that since we assumed that $\mathcal{P}_{m-1}$ is a coarsening of $\mathcal{P}'$ and satisfies condition (1) with θ_{m-1}, then $\mathcal{P}_m$ is also a coarsening of $\mathcal{P}'$ which satisfies condition (1) with θ_m.

Eventually, since the number of sets in $\mathcal{P}'$ is ℓ, this process halts within at most ℓ steps and we obtain an $(\varepsilon, \alpha, \beta)$ good coarsening $\mathcal{P}$ and θ, a parameter satisfying $\theta \geq \theta_\ell$, with which conditions (1) and (2) are satisfied. Solving (8) with initial condition $\theta_1 = \varepsilon$, we obtain

$$\theta_m = \varepsilon \left(\frac{\varepsilon \alpha}{\ell^2}\right)^{2^{m-1}-1}.$$

Therefore, we have

$$\theta \geq \theta_\ell \geq \varepsilon \left(\frac{\varepsilon \alpha}{\ell^2}\right)^{2^{\ell-1}-1} \geq \varepsilon \left(\frac{\varepsilon \alpha}{\ell^2}\right)^{2^\ell},$$

as required. □

2.5 Proof of Lemma 2.2

To prove Lemma 2.2 we will show that with the right choice of α an $(\varepsilon, \alpha, \beta)$-good coarsening of a δ-primary decomposition is in fact a $(\varepsilon, \delta, \beta)$-good decomposition.

Proof of Lemma 2.2. Let $G = (V, E)$ be a graph on n vertices with minimal degree at least δn, let $\varepsilon > 0$ and let $0 < \beta \leq 240n^2/(\varepsilon \delta^4)$. By Lemma 2.11, there exists a δ-primary decomposition of V. We denote this decomposition by $\mathcal{P}'$. By Lemma 2.13, we obtain that there exists an $(\varepsilon, \varepsilon^9, \beta)$-good coarsening of $\mathcal{P}'$, denoted by $\mathcal{P}$. We also denote this coarsening explicitly by $V = V_1 \sqcup \ldots \sqcup V_k$ and we let θ be the parameter from Lemma 2.13 to which the coarsening corresponds.

We claim that $\mathcal{P}$ is indeed an $(\varepsilon, \delta, \beta)$-good decomposition satisfying conditions (3) and (5) of Definition 2.1 with this θ. We first note that for $\varepsilon > 0$ small enough and by the properties of a good coarsening

$$\varepsilon^{11 \cdot 2^{2/\delta}} \leq \varepsilon \left(\frac{\varepsilon^{10} \delta^2}{4}\right)^{2^{2/\delta}} \leq \theta \leq \varepsilon.$$

Conditions (1), (2) and (4) of Definition 2.1 are immediate for every coarsening of a δ-primary decomposition. Condition (5) is satisfied by condition (2) of the coarsening in Definition 2.12. We are then left with verifying that condition (3) of Definition 2.1 holds. To this end, we let V_i be some set in $\mathcal{P}$ and $V_i = V_{i,1} \sqcup \ldots \sqcup V_{i,\ell_i}$ be its partition to ℓ_i sets of $\mathcal{P}'$ which we denote by $\mathcal{P}_i$. Note that for every $j \in [\ell_i]$ and for every $v \in V_{i,j}$ we have $\deg(v, V_{i,j}) \geq \delta^4 n/40$. Also, by condition (4) of the primary decomposition, Definition 2.10, we have that the spectral gap corresponding to $G[V_{i,j}]$ is at least $\delta^{10}/2^{22}$. Finally, by the properties of the coarsening, the graph $H(\mathcal{P}_i, \theta\beta)$ is connected. Hence, denoting $\gamma(G[V_i])$ for the spectral gap of $G[V_i]$ and using Lemma 2.4 we get that

$$\gamma(G[V_i]) \geq \min\left\{\frac{\delta^{10}}{2^{22}}, \quad \frac{\delta^4 n}{40} \cdot \frac{\delta^{10}}{2^{22}} \cdot \theta\beta \cdot \frac{1}{6\ell_i n^3}\right\}. \tag{9}$$

Since $\theta \leq \varepsilon$ and $\beta \leq 240n^2/(\varepsilon\delta^4)$ and $\ell_i \geq 1$ we learn that the minimum above is attained in the second term. As $\mathcal{P}$ is a coarsening of a δ-primary decomposition, we have that $\ell_i \leq 2/\delta$, hence,

$$\gamma(G[V_i]) \geq \frac{\delta^4 n}{40} \cdot \frac{\delta^{10}}{2^{22}} \cdot \theta\beta \cdot \frac{1}{6\ell_i n^3} \geq \frac{\delta^{15}\theta\beta}{2^{31} n^2}$$

and we deduce that condition (3) of Definition 2.1 holds. □

3. Random walks and Wilson's algorithm on decomposed graphs

In the rest of this paper we take $\beta = n^{3/2}$ in the decomposition of Section 2. The main goal of this section is to prove the following estimate.

Theorem 3.1. *For any $\delta \in (0, 1]$ there exists $C = C(\delta) < \infty$ such that the following holds. Let $G = (V, E)$ be a connected simple graph on n vertices with minimal degree at least δn and $\varepsilon > 0$. Denote by $V = V_1 \sqcup \ldots \sqcup V_k$ an $(\varepsilon, \delta, n^{1.5})$-good decomposition of G with parameter θ (as guaranteed to exist by Lemma 2.2). Then for every $i \in [k]$ there are two vertices $v_1^i, v_2^i \in V_i$ such that if $\mathcal{T}$ is a uniform spanning tree of G and φ is the unique path between v_1^i and v_2^i in $\mathcal{T}$, then*

$$\mathbb{P}\left(\varepsilon^8 \theta \sqrt{n} \leq |\varphi| \leq \frac{\sqrt{n}}{\theta\varepsilon^8} \text{ and } \varphi \subseteq V_i\right) \geq 1 - C\varepsilon^2.$$

In Section 3.1 we prove a couple of preliminary useful random walk estimates on graphs with linear minimal degree that do not involve the decomposition. In Section 3.2, we show that a random walk on G stays inside one set of the decomposition for at least $\sqrt{n}$ steps with high probability; since the spectral gap of each set in the decomposition is at least of order $n^{-1/2}$ and the induced graph on the set has linear minimal degree, the random walk is mixed in this set (even though the mixing time of G may be much larger than $\sqrt{n}$). We use this estimate in Section 3.3 to prove the aforementioned Theorem 3.1.

3.1 Preliminary random walk estimates

It is a classical fact that the mixing time of the random walk on a connected graph G is always $O(\gamma^{-1}\log n)$ where n is the number of vertices and $\gamma = \gamma(G)$ is the spectral gap, see for instance [14, Theorem 12.4]. This estimate is sharp as is seen on bounded degree expander graphs where the gap is $\Omega(1)$ but at least $\Omega(\log n)$ steps are needed for the walker to be able to reach the majority of the graph. However, when the minimal degree is linear, after a single step the location is already spread on a set of linear size and this estimate can be improved.

Lemma 3.2. *For any $\delta \in (0, 1]$ there exists a constant $C = C(\delta) < \infty$ such that the following holds. For any simple graph G on n vertices with minimal degree at least δn and any $\varepsilon \in (0, 1)$ we have*

$$t_{mix}^G(\varepsilon) \leq C \log(1/\varepsilon)\left(\gamma^{-1}(G) + \log(n)\right).$$

Proof Let P be the transition matrix of the lazy random walk on G. Recall that P is a self-adjoint operator $P : L^2(\pi) \to L^2(\pi)$ where $\pi(v) = \deg_G(v)/2|E(G)|$ is the stationary distribution. We denote the eigenvalues of P by $1 = \lambda_1 \geq \ldots \geq \lambda_n \geq 0$ and $\gamma(P) = 1 - \lambda_2$ and by $\mathbf{1}$ the all 1 vector which is the eigenvalue corresponding to the eigenvalue 1. Let μ be any probability measure on V and write f for the vector $f(v) = \mu(v)/\pi(v)$. We have that $f - \mathbf{1}$ is orthogonal to $\mathbf{1}$ and for any integer $t \geq 1$ we have that $\pi(v)Pf(v) = \mathbb{P}_\mu(X_t = \cdot)$ so in particular $P^t f - \mathbf{1}$ is orthogonal to $\mathbf{1}$. Hence

$$\left\|P^t f - \mathbf{1}\right\|_2 = \left\|P^t(f - \mathbf{1})\right\|_2 \leq \lambda_2^t \left\|f - \mathbf{1}\right\|_2. \tag{10}$$

We rewrite this as

$$\left\| \frac{\mathbb{P}_\mu(X_t = \cdot)}{\pi(\cdot)} - \mathbf{1} \right\|_2 \leq \lambda_2^t \left\| f - \mathbf{1} \right\|_2 .$$

We now claim that for any $v \in V$ and every $u \in V$ we have that

$$\mathbb{P}_v(X_{\lceil \log_2(n) \rceil} = u) \leq \frac{2}{\delta n}.$$

Indeed, if the random walker made a non-lazy step at some time in $\{1, \ldots, \lceil \log_2(n) \rceil\}$, then the probability to be at any vertex u at time $\lceil \log_2(n) \rceil$ is bounded by $1/(\delta n)$. On the other hand, the probability of staying put $\lceil \log_2(n) \rceil$ steps is at most $\frac{1}{n}$. Since $\frac{\delta}{n} \leq \pi(\cdot) \leq \frac{1}{\delta n}$ it follows that

$$\left\| \frac{\mathbb{P}_v(X_{\lceil \log_2(n) \rceil} = \cdot)}{\pi(\cdot)} - \mathbf{1} \right\|_2^2 \leq \frac{2}{\delta^2}. \tag{11}$$

Using (10) and (11), for $t = \lceil \log_2(n) \rceil + \log(\sqrt{2}/\varepsilon\delta)\gamma^{-1}$ and every $v \in V$ we have

$$\left\| \frac{\mathbb{P}_v(X_t = \cdot)}{\pi(\cdot)} - \mathbf{1} \right\|_2 \leq \lambda_2^{t - \lceil \log_2(n) \rceil} \left\| \frac{\mathbb{P}_v(X_{\lceil \log_2(n) \rceil} = \cdot)}{\pi(\cdot)} - \mathbf{1} \right\|_2 \leq \frac{\sqrt{2}}{\delta}(1 - \gamma)^{t - \lceil \log_2(n) \rceil} \leq \varepsilon.$$

By [14, Lemma 12.18] we have that

$$2\|\mathbf{p}_t(v, \cdot) - \pi(\cdot)\|_{d_{\mathrm{TV}}} \leq \left\| \frac{\mathbb{P}_v(X_t = \cdot)}{\pi(\cdot)} - \mathbf{1} \right\|_2 \leq \varepsilon ,$$

concluding our proof. □

Claim 3.3. *For any $\delta > 0$ there exists $c = c(\delta) > 0$ such that the following holds. For any $\varepsilon \in (0, c)$, any simple graph $G = (V, E)$ on $n \geq \varepsilon^{-2}$ vertices with minimal degree at least δn, any $U \subset V$ with $|U| \geq \varepsilon\sqrt{n}$ and any vertex $v \in G$*

$$\mathbb{P}_v(X[0, 2(t_{mix}^G(\varepsilon/2) + \lfloor \sqrt{n} \rfloor)] \cap U = \varnothing) \leq 1 - \frac{\varepsilon\delta}{4} ,$$

where X is the simple random walk on G.

Remark 3.4. We emphasize a potentially confusing point: t_{mix} is defined in terms of the lazy random walk, but in this claim, as well as the rest of this paper, we study the non-lazy random walk running for times depending on t_{mix}.

Proof. We prove this for the *lazy* simple random walk and trivially it follows for the usual random walk. Without loss of generality we may assume that $|U| = \lceil \varepsilon\sqrt{n} \rceil$, otherwise we may take a subset of U of that size. By equation (3) we have that $2(t_{\mathrm{mix}}^G(\varepsilon/2) + \lfloor \sqrt{n} \rfloor) \geq t_{\mathrm{mix}}^G(\varepsilon^2) + \lfloor \sqrt{n} \rfloor$ so it is then enough to bound from below the probability that X hits U within $t_{\mathrm{mix}}^G(\varepsilon^2) + \lfloor \sqrt{n} \rfloor$ steps. Recall that

$$\mathbb{P}(Z > 0) \geq \frac{\mathbb{E}^2[Z]}{\mathbb{E}[Z^2]} , \tag{12}$$

for any non-negative random variable Z. Let Y be a lazy random walk starting from the stationary distribution and define $Z = |\{t \in [0, \lfloor \sqrt{n} \rfloor] \mid Y_t \in U\}|$. We will use (12) to bound $\mathbb{P}(Z > 0)$ from below. Since $|U| \geq \varepsilon\sqrt{n}$ and the minimal degree is δn, we have that $\pi(U) \geq \varepsilon\delta n^{-1/2}$ and so $\mathbb{E}[Z] = (\lfloor \sqrt{n} \rfloor + 1)\pi(U) \geq \varepsilon\delta$.

To bound $\mathbb{E}[Z^2]$, let $t < r$ be two positive integers. If $Y_t \in U$, there is a probability of $2^{-(r-t)}$ that the walker made $r - t$ lazy steps and then $Y_r = Y_t$. Else, since the minimal degree is at least

δn, the probability that $Y_r \in U$ is bounded by $|U|/(\delta n)$. Therefore,

$$\mathbb{P}(Y_t \in U, Y_r \in U) \leq \mathbb{P}(Y_t \in U) \cdot \left(\frac{|U|}{\delta n} + \frac{1}{2^{r-t}} \right) = \pi(U) \cdot \left(\frac{|U|}{\delta n} + \frac{1}{2^{r-t}} \right).$$

Hence, by summing over all $t \leq r \leq \lfloor \sqrt{n} \rfloor$

$$\mathbb{E}[Z^2] = \sum_{t=0}^{\lfloor \sqrt{n} \rfloor} \mathbb{P}(Y_t \in U) + 2 \sum_{t=0}^{\lfloor \sqrt{n} \rfloor} \sum_{r=t+1}^{\lfloor \sqrt{n} \rfloor} \mathbb{P}(Y_t \in U, Y_r \in U)$$

$$\leq \mathbb{E}[Z] + 2 \binom{\lfloor \sqrt{n} \rfloor + 1}{2} \frac{\pi(U)|U|}{\delta n} + 2 \sum_{t=0}^{\lfloor \sqrt{n} \rfloor} \pi(U) \sum_{r=t+1}^{\lfloor \sqrt{n} \rfloor} \frac{1}{2^{r-t}}.$$

Since $\mathbb{E}[Z] = (\lfloor \sqrt{n} \rfloor + 1) \pi(U)$ we upper bound the last term of the right-hand side by $2\mathbb{E}[Z]$. For the middle term we write

$$2 \binom{\lfloor \sqrt{n} \rfloor + 1}{2} \frac{\pi(U)|U|}{\delta n} \leq (\lfloor \sqrt{n} \rfloor + 1) \cdot \frac{\pi(U)|U|}{\delta \sqrt{n}} \leq \mathbb{E}[Z] \left(\frac{\varepsilon \sqrt{n} + 1}{\delta \sqrt{n}} \right) \leq \frac{1}{2} \mathbb{E}[Z],$$

where the last inequality holds for ε small enough. Therefore, for ε small enough we have that $\mathbb{E}[Z^2] \leq \frac{7}{2} \mathbb{E}[Z]$ and thus by (12)

$$\mathbb{P}(Z > 0) \geq \frac{2\mathbb{E}[Z]}{7} \geq \frac{2\varepsilon\delta}{7}.$$

By the definition of $t_{\text{mix}}^G(\varepsilon^2)$, if X is a random walk starting from some $v \in V$, we have that $d_{\text{TV}}(X_{t_{\text{mix}}(\varepsilon^2)}, \pi) \leq \varepsilon^2$. Therefore, we can couple the walk X starting from time $t_{\text{mix}}^G(\varepsilon^2)$ with an independent random walk Y starting from the stationary distribution such that the walks coincide with probability larger than $1 - \varepsilon^2$. We thus obtain that for ε small enough

$$\mathbb{P}_v(X[t_{\text{mix}}^G(\varepsilon^2), t_{\text{mix}}^G(\varepsilon^2) + \lfloor \sqrt{n} \rfloor] \cap U \neq \varnothing) \geq \varepsilon\delta/3.5 - \varepsilon^2 \geq \varepsilon\delta/4. \qquad \square$$

3.2 Random walks of length $\sqrt{n}$ stay in the same set of the decomposition

We now show that with high probability the random walker on a graph with linear minimal degree will stay in the same set of its $(\varepsilon, \delta, n^{1.5})$-good decomposition that it walked to in its first step.

Lemma 3.5. *Let $\delta \in (0, 1]$, $\varepsilon > 0$ and $G = (V, E)$ be a simple graph on n vertices with minimal degree at least δn. Also let $V = V_1 \sqcup \ldots \sqcup V_k$ be an $(\varepsilon, \delta, n^{1.5})$-good decomposition of G with parameter θ (as guaranteed to exist by Lemma 2.2). Then for any $C > 0$ and any $i \in [k]$*

$$\mathbb{P}_v \left(\exists t \in [1, C\sqrt{n}] : X_t \in V_i, X_{t+1} \notin V_i \right) \leq \frac{C\theta^2\varepsilon^9}{\delta^2}, \tag{13}$$

where X is the simple random walk on G. Furthermore, for any $i \in [k]$ there exists a set $V_i' \subseteq V_i$ satisfying $|V_i'| \geq \delta^4 n/80$ such that for every $v \in V_i'$

$$\mathbb{P}_v \left(X[0, C\sqrt{n}] \subseteq V_i \right) \geq 1 - \frac{80C\theta^2\varepsilon^9}{\delta^6}. \tag{14}$$

Proof. Let $i \in [k]$ and let θ be the parameter from the $(\varepsilon, \delta, n^{1.5})$-good decomposition $\mathcal{P}$. For every $t \geq 1$, we have

$$\mathbb{P}(X_t \in V_i, X_{t+1} \notin V_i) \leq \sum_{w \in V_i} \mathbb{P}(X_t = w) \cdot \mathbf{p}(w, V_i^c) \leq \sum_{w \in V_i} \frac{1}{\delta n} \frac{|E(w, V_i^c)|}{\delta n} = \frac{|E(V_i, V_i^c)|}{\delta^2 n^2}.$$

Hence by taking the union over $t \in [1, C\sqrt{n}]$ and using condition (5) of a $(\varepsilon, \delta, n^{1.5})$-good decomposition (Definition 2.1) we immediately obtain (13). To prove (14) let $C > 0$ and fix some $v \in V_i$. By condition (4) of Definition 2.1, we have that $\deg(v, V_i) \geq \delta^4 n / 40$. By (13) we have

$$\mathbb{P}_v(X_1 \in V_i \text{ and } X\left[1, C\sqrt{n} + 1\right] \not\subseteq V_i) \leq \frac{C\theta^2 \varepsilon^9}{\delta^2}.$$

Yet on the other hand,

$$\mathbb{P}_v(X_1 \in V_i \text{ and } X[1, C\sqrt{n}] \not\subseteq V_i) \geq \frac{1}{n} \sum_{u \sim v, u \in V_i} \mathbb{P}_u\left(X[0, C\sqrt{n}] \not\subseteq V_i\right).$$

Thus the number of $u \in V_i$ with $u \sim v$ satisfying

$$\mathbb{P}_u\left(X[0, C\sqrt{n}] \not\subseteq V_i\right) \geq \frac{80 C \theta^2 \varepsilon^9}{\delta^6}.$$

cannot be larger than $\delta^4 n / 80$. Since $\deg(v, V_i) \geq \delta^4 n / 40$ we conclude the proof of (14). □

3.3 LERWs and Wilson's Algorithm on the decomposed graph

We now proceed to the proof of Theorem 3.1. In the rest of this section we assume that $\delta \in (0, 1]$ and $\varepsilon > 0$ are given, that $G = (V, E)$ is a connected simple graph on n vertices with minimal degree δn and that $V = V_1 \sqcup \ldots \sqcup V_k$ is an $(\varepsilon, \delta, n^{1.5})$-good decomposition with parameter θ as guaranteed to exist by Lemma 2.2. Lastly, note that if $\sqrt{n} \leq 1/(\theta \varepsilon^8)$, then Theorem 3.1 is trivial, so we assume the contrary.

As described in Section 1.1, the UST path between two vertices of G is distributed as the LERW between them. A recurring problem in analysing the UST is that the random walk path between two vertices may be much longer than its loop erasure, meaning that most of the random walk path is erased during the loop erasure. To overcome this obstacle we use the following idea which goes back to Wilson [19] and was used extensively by Peres and Revelle [17]. Let G^ρ be the network obtained from $G = (V, E)$ by adding a vertex ρ, connecting it to each $v \in V$ and assigning edge weights

$$w(v, \rho) = \frac{\theta \varepsilon^4 \deg_G(v)}{\sqrt{n} - \theta \varepsilon^4}.$$

for any $v \in V$. An immediate calculation shows that with these edge weights the probability that the random walk starting from any $v \in V$ moves to ρ in the first step is $\theta \varepsilon^4 n^{-1/2}$ and so τ_ρ is a geometric random variable with expectation $\sqrt{n}/\theta \varepsilon^4$. Furthermore, if X is a simple random walk on this network, then conditioned on $\tau_\rho = m$, we have that $X[0, m - 1]$ is distributed like a simple random walk on G of length $m - 1$. Thus, in G^ρ, the random walk typically takes $\sqrt{n}$ steps to hit ρ. It turns out that a positive fraction of such a walk survives the loop erasure with high probability, and is hence easier to analyse. To deduce information about the LERW in G rather than G^ρ we use Lemma 1.7 stating that $\mathsf{UST}(G)$ stochastically dominates $\mathsf{UST}(G^\rho) \cap E(G)$. Hence, if v_1 and v_2 are two distinct vertices and φ and φ' are the unique paths between them in $\mathsf{UST}(G)$ and $\mathsf{UST}(G^\rho)$, respectively, then

$$d_{\mathrm{TV}}(\varphi, \varphi') \leq \mathbb{P}(\rho \in \varphi'). \tag{15}$$

The proof strategy of Theorem 3.1 is as follows. Fix some $i \in [k]$ and two vertices v_1, v_2 in V_i which we will choose according to (14). We run Wilson's algorithm (see Section 1.1) on G^ρ where the first three vertices in the ordering of the vertices of G^ρ are (ρ, v_1, v_2). We will first show that with high probability the random walk from v_1 to ρ stays within V_i except for the last step (Claim 3.6) and that the length of its loop erasure is at least $\varepsilon \sqrt{n}$ (Claim 3.7); it is also unlikely to

contain v_2. Next (Lemma 3.8) we show that conditioned on this first LERW, with high probability the second LERW starting at v_2 hits the first LERW in a vertex different than ρ and stays in V_i until that visit. We will also show there that the second LERW is typically longer than $\varepsilon^8 \theta \sqrt{n}$. This gives a lower bound on $|\varphi'|$ and by (15) a lower bound for $|\varphi|$ is obtained.

Claim 3.6. *For any $i \in [k]$ there exists a set $V'_i \subseteq V_i$ with $|V'_i| \geq \delta^4 n/80$ such that for every $v \in V_i'$, a simple random walk on G^ρ starting from v satisfies*

$$\mathbb{P}_v(\tau_\rho < \tau_{V \setminus V_i}) \geq 1 - B\varepsilon^2,$$

for some $B = B(\delta) < \infty$.

Proof. Apply Lemma 3.5 with $C = 1/\theta\varepsilon^6$ to obtain a set V'_i such that for every $v \in V_i'$, if Y is a simple random walk on the graph G, then

$$\mathbb{P}_v(Y[0, \sqrt{n}/\theta\varepsilon^6] \subseteq V_i) \geq 1 - \frac{80\theta\varepsilon^3}{\delta^6} \geq 1 - B\varepsilon^3. \tag{16}$$

Also, for a random walk X on G^ρ we have $\mathbb{E}_v \tau_\rho = \sqrt{n}/\theta\varepsilon^4$ and thus Markov's inequality gives

$$\mathbb{P}_v(\tau_\rho \leq \sqrt{n}/\theta\varepsilon^6) \geq 1 - \varepsilon^2 \tag{17}$$

Let X be a random walk on G^ρ. Conditioned on τ_ρ the random path $X[0, \tau_\rho - 1]$ has the distribution of a random walk on G. Hence combining (16) and (17) yields the desired result. $\square$

Claim 3.7. *For any $i \in [k]$ there exists a set $V'_i \subseteq V_i$ with $|V'_i| \geq \delta^4 n/80$ such that for every $v \in V_i'$ if X is a simple random walk on G^ρ starting at v and stopped when hitting ρ, then*

$$\mathbb{P}_v(|\mathsf{LE}(X)| \geq \varepsilon\sqrt{n} \text{ and } \mathsf{LE}(X) \subseteq V_i \cup \{\rho\}) \geq 1 - C\varepsilon^2,$$

for some $C = C(\delta) < \infty$.

Proof. As in the previous proof, for every $v \in V_i$ we have that $\mathbb{P}(\tau_\rho > \sqrt{n}) \geq 1 - \varepsilon^4$. We condition on this event on τ_ρ and on the first $\tau_\rho - \varepsilon\sqrt{n}$ steps of the random walk. Let us assume first that

$$|\mathsf{LE}(X[0, \tau_\rho - \varepsilon\sqrt{n}))| \geq \varepsilon\sqrt{n}.$$

In this case we denote by U the first $\varepsilon\sqrt{n}$ vertices of $\mathsf{LE}(X[0, \tau_\rho - \varepsilon\sqrt{n}))$. As before, conditioned on τ_ρ the walk $X[0, \tau_\rho - 1]$ is distributed as an unconditional random walk. By the Markov property, the random walk at times $[\tau_\rho - \varepsilon\sqrt{n}, \tau_\rho)$ is distributed as an unconditional random walk on the graph G. Hence, since the minimal degree of G is at least δn, the probability that $X_t \in U$ for some $t \in [\tau_\rho - \varepsilon\sqrt{n}, \tau_\rho]$ is at most ε^2/δ. If this does not occur, then the set U survives the loop erasure. It follows that

$$\mathbb{P}\left(|\mathsf{LE}\left(X[0, \tau_\rho - \varepsilon\sqrt{n}]\right)| \geq \varepsilon\sqrt{n} \text{ and } |\mathsf{LE}(X)| \leq \varepsilon\sqrt{n}\right) \leq \varepsilon^4 + \varepsilon^2/\delta \leq C\varepsilon^2,$$

for some $C = 1 + \delta^{-1}$. In the second case $|\mathsf{LE}(X[0, \tau_\rho - \varepsilon\sqrt{n}))| \leq \varepsilon\sqrt{n}$. In this case the last $\varepsilon\sqrt{n}$ steps of X will survive the loop erasure high probability. Indeed, by Markov's inequality and since the degrees are at least δn we deduce that the walk $X[\tau_\rho - \varepsilon\sqrt{n}, \tau_\rho]$ has no loops with probability at least $1 - \varepsilon^2/\delta$. By the same reasoning, the walk $X[\tau_\rho - \varepsilon\sqrt{n}, \tau_\rho]$ does not visit $|\mathsf{LE}(X[0, \tau_\rho - \varepsilon\sqrt{n}))|$ with probability at least $1 - \varepsilon^2/\delta$. On these two events $\mathsf{LE}(X)$ contains $X[\tau_\rho - \varepsilon\sqrt{n}, \tau_\rho]$. It follows that

$$\mathbb{P}(|\mathsf{LE}(X[0, \tau_\rho - \varepsilon\sqrt{n}))| \leq \varepsilon\sqrt{n} \text{ and } |\mathsf{LE}(X)| \leq \varepsilon\sqrt{n}) \leq \varepsilon^4 + 2\varepsilon^2/\delta \leq C\varepsilon^2,$$

for some $C = 1 + 2\delta^{-1}$. Combining the last two inequalities and using Claim 3.6 finishes the proof. $\square$

Lemma 3.8. *For any $i \in [k]$ there exist two distinct vertices $v_1, v_2 \in V_i$ such that the following holds. Let X be a simple random walk in G^ρ starting at v_1 and stopped when hitting ρ, and conditioned*

on X, let Y be an independent simple random walk on G^ρ starting at v_2 and stopped when hitting $\mathsf{LE}(X)$. Then,

$$\mathbb{P}\left(Y_{\tau_{\mathsf{LE}(X)}} \neq \rho \text{ and } |\mathsf{LE}(Y)| \geq \varepsilon^8 \theta \sqrt{n} \text{ and } \mathsf{LE}(Y) \subseteq V_i)\right) \geq 1 - C\varepsilon^2, \tag{18}$$

for some $C = C(\delta) < \infty$.

Proof. We apply Claim 3.7 to obtain the set V_i' and take v_1, v_2 to be any two distinct vertices of V_i'. Let Z be simple random walk on $G[V_i]$ starting at v_2 independent of X (note that unlike Y, the random walk Z does not leave V_i and does not visit ρ). By Lemma 3.2 and condition (3) of an $(\varepsilon, \delta, n^{1.5})$-good decomposition (Definition 2.1) we have that for some $C = C(\delta)$

$$t_{\mathrm{mix}}^{G[V_i]}(\varepsilon/2) + \left\lfloor \sqrt{n} \right\rfloor \leq C \log(1/\varepsilon) \frac{\sqrt{n}}{\theta}.$$

Fix a constant

$$A = \left\lceil 320 \log(1/\varepsilon)/\varepsilon\delta^4 \right\rceil,$$

so that by the last estimate, and since $v_2 \in V_i'$ the assertion of Lemma 3.5 implies that

$$\mathbb{P}_{v_2}\left(Y\left[0, 2A(t_{\mathrm{mix}}^{G[V_i]}(\varepsilon/2) + \left\lfloor \sqrt{n} \right\rfloor)\right] \subseteq V_i\right) \geq \mathbb{P}_{v_2}\left(Y\left[0, \frac{2AC \log(1/\varepsilon)\sqrt{n}}{\theta}\right] \subseteq V_i\right)$$

$$\geq 1 - \frac{160AC\theta\varepsilon^8}{\delta^6} - \mathbb{P}\left(\tau_\rho \leq \frac{2AC \log(1/\varepsilon)\sqrt{n}}{\theta}\right)$$

$$\geq 1 - C\varepsilon^2, \tag{19}$$

for some $C = C(\delta)$. Thus we learn that with probability larger than $1 - C\varepsilon^2$ we can couple Y and Z such that $Y_t = Z_t$ for all $t \leq 2A(t_{\mathrm{mix}}^{G[V_i]}(\varepsilon/2) + \left\lfloor \sqrt{n} \right\rfloor)$. Now, since $v_1 \in V_i'$ the assertion of Claim 3.7 states that with probability at least $1 - C\varepsilon^2$ we have that $|\mathsf{LE}(X)| \geq \varepsilon\sqrt{n}$ and $\mathsf{LE}(X) \subseteq V_i \cup \{\rho\}$. The minimal degree in $G[V_i]$ is at least $\delta^4 n/40$ and $\delta n/2 \leq |V_i| \leq n$, allowing us to apply Claim 3.3 A times together with the previous estimate to obtain that

$$\mathbb{P}_{v_2}(Z[0, 2A(t_{\mathrm{mix}}^{G[V_i]}(\varepsilon/2) + \left\lfloor \sqrt{n} \right\rfloor)] \cap \mathsf{LE}(X) = \varnothing) \leq \left(1 - \frac{\varepsilon\delta^4}{160}\right)^A + C\varepsilon^2 \leq (C+1)\varepsilon^2,$$

by our choice of A. This together with the coupling of Y and Z and (19) gives

$$\mathbb{P}\left(Y_{\tau_{\mathsf{LE}(X)}} \neq \rho \text{ and } \mathsf{LE}(Y) \subset V_i\right) \geq 1 - C\varepsilon^2.$$

We are left with bounding $|\mathsf{LE}(Y)|$ from below. We will show that with large probability $Y[0, \varepsilon^8 \theta \sqrt{n}] \subseteq \mathsf{LE}(Y)$. Indeed, since $\mathbb{E}_{v_1} \tau_\rho = \sqrt{n}/\theta\varepsilon^4$ we have that $|\mathsf{LE}(X)| \leq \sqrt{n}/\theta\varepsilon^6$ with probability at least $1 - \varepsilon^2$. Furthermore, since $v_1 \neq v_2$ and the degree is at least δn we have that $v_2 \notin \mathsf{LE}(X)$ with probability at least $1 - \varepsilon^2 - \delta^{-1} n^{-1/2}/\theta\varepsilon^6 \geq 1 - C\varepsilon^2$. Hence

$$\mathbb{P}(Y[0, \varepsilon^8 \theta \sqrt{n}] \cap \mathsf{LE}(X) \neq \varnothing) \leq \varepsilon^8 \theta \sqrt{n} \frac{\sqrt{n}/\theta\varepsilon^6}{\delta n} + C\varepsilon^2 \leq C\varepsilon^2. \tag{20}$$

Furthermore, since the degree is at least δn, the union bound gives that

$$\mathbb{P}(Y[0, \varepsilon^8 \theta \sqrt{n}] \cap Y[\varepsilon^8 \theta \sqrt{n}, \sqrt{n}/\theta\varepsilon^6] \neq \varnothing) \leq \frac{\varepsilon^2}{\delta}. \tag{21}$$

Since $\tau_\rho \geq \sqrt{n}/\theta\varepsilon^6$ occurs with probability at most ε^2, we deduce by (20), (21) that

$$\mathbb{P}\left(\mathsf{LE}(Y[0, \varepsilon^8 \theta \sqrt{n}]) \subseteq \mathsf{LE}(Y)\right) \geq 1 - C\varepsilon^2.$$

for some $C = C(\delta) < \infty$. Lastly, again by the linear minimal degree and the union bound, the probability that there is a repeating vertex in $Y[0, \varepsilon^8 \theta \sqrt{n}]$ is at most $\varepsilon^{16}\theta^2/\delta$; when this does not occur $Y[0, \varepsilon^8 \theta \sqrt{n}] = \mathsf{LE}(Y[0, \varepsilon^8 \theta \sqrt{n}])$, concluding our proof. □

Proof of Lemma 3.1. If $\theta \varepsilon^8 \le 1/\sqrt{n}$, the claim is trivial. We assume the converse is true, and we take the vertices v_1^i and v_2^i from Lemma 3.8. We denote by φ and φ' the paths between them in $\mathsf{UST}(G)$ and $\mathsf{UST}(G^\rho)$, respectively. As mentioned in the beginning of this subsection, we couple $\mathsf{UST}(G)$ and $\mathsf{UST}(G^\rho)$ such that $\mathsf{UST}(G^\rho) \cap E(G) \subseteq \mathsf{UST}(G)$. We use (15) and recall that under this coupling, if $\rho \notin \varphi'$, then $\varphi = \varphi'$. Hence it suffices to show that

$$\mathbb{P}\left(\varepsilon^8 \theta \sqrt{n} \le |\varphi'| \le \frac{\sqrt{n}}{\theta \varepsilon^8} \text{ and } \varphi' \subseteq V_i \text{ and } \rho \notin \varphi' \right) \ge 1 - C\varepsilon^2,$$

for some $C = C(\delta)$. By Wilson's algorithm, we can sample φ' by sampling $\mathsf{LE}(X)$, a LERW from v_1^i to ρ and then sampling $\mathsf{LE}(Y)$, another LERW from v_2^i to $\mathsf{LE}(X)$. The path between v_1^i and v_2^i in $\mathsf{LE}(X) \cup \mathsf{LE}(Y)$ is distributed as the path between v_1^i and v_2^i in $\mathsf{UST}(G^\rho)$. By Lemma 3.8 and Claim 3.7, this path is contained in V_i and does not contain ρ with probability larger than $1 - C\varepsilon^2$. By construction $|\varphi'| \ge |\mathsf{LE}(Y)|$ hence Lemma 3.8 gives the required lower bound on $|\varphi'|$. Finally, as the length of $\mathsf{LE}(X)$ and $\mathsf{LE}(Y)$ is bounded by two independent random variables with the distribution of τ_ρ, by Markov's inequality

$$\mathbb{P}\left(|\mathsf{LE}(X)| + |\mathsf{LE}(Y)| \ge \frac{\sqrt{n}}{\theta \varepsilon^8} \right) \le 1 - C\varepsilon^4,$$

concluding the proof. $\qquad\square$

4. Proof of main theorem

In [16], the following strategy was used to show that the diameter grows like $\sqrt{n}$. First, a small part of the UST is sampled. This part contains roughly $\sqrt{n}$ vertices (in [16], it is simply a path between two vertices). Then, it is shown that this part of the UST is difficult to avoid in the sense that random walks starting from any vertex of the graph will hit it with positive probability within roughly $\sqrt{n}$ steps. To formalize and quantify this we first define

$$\mathbf{p}_W^t(v, v) = \mathbb{P}_v(X_t = v, \tau_W > t),$$

for any $W \subset V$. Next we define the W-bubble sum by

$$\mathcal{B}_W(G) = \sum_{t=0}^{\infty} (t + 1) \sup_{v \in G} \mathbf{p}_W^t(v, v).$$

If the set W is difficult to avoid, then the $\mathbf{p}_W^t(v, v)$ decays fast with t and thus $\mathcal{B}_W(G)$ is small. It is shown in [16] that if $\mathcal{B}_W(G)$ is small, then the diameter of $\mathsf{UST}(G/W)$ cannot be too large:

Lemma 4.1. ([16, Lemma 3.13]) *Let $G = (V, E)$ be a connected graph, let $D = \frac{\max_v \deg(v)}{\min_v \deg(v)}$ and let W be a non-empty vertex set. Let $\mathcal{T}_W$ be a UST on the graph G/W. Then*

$$\mathbb{P}(diam(\mathcal{T}_W) \ge \ell) \le \frac{C_3 |W|}{\ell},$$

for $C_3 = 138420 \cdot D^4 \mathcal{B}_W(G)^3 \log(192 D \mathcal{B}_W(G))$.

    In our context, we will take W to be the union of k paths in the UST drawn according to an $(\varepsilon, \delta, n^{1.5})$-good decomposition. In the next few claims, using the results we obtained in Section 3.3, we will show that with high probability $\mathcal{B}_W(G) = O(1)$, after which we will prove Theorem 1.1.

Lemma 4.2. *For any $\delta > 0$ there exists $b(\delta) > 0$ such that for any $\varepsilon \in (0, b)$ there exists $C = C(\varepsilon, \delta) < \infty$ and $c = c(\varepsilon, \delta) > 0$ such that the following holds. Let $G = (V, E)$ be a connected simple graph on n vertices with minimal degree at least δn and $\varepsilon > 0$. Denote by $V = V_1 \sqcup \ldots \sqcup V_k$ an*

$(\varepsilon, \delta, n^{1.5})$-*good decomposition of G with parameter θ (as guaranteed to exist by Lemma 2.2). Then,* *for every set W that satisfies $|W \cap V_i| \geq \varepsilon^8 \theta \sqrt{n}$ for every $i \in [k]$, we have that*

$$\sum_{t=1}^{\infty} (t+1) \sup_{v \in G} \boldsymbol{p}_W^t(v, v) \leq C.$$

Proof. Let W be such a set and fix $v \in V$. We will first show that

$$\mathbb{P}_v \left(X \left[0, C\sqrt{n} \right] \cap W \neq \varnothing \right) \geq c. \tag{22}$$

for some C,c depending on ε and δ. There exists at least one component V_i in the decomposition such that $\mathbb{P}_v(X_1 \in V_i) \geq 1/k$ (note that v does not necessarily belong to V_i). Let Y be a random walk on $G[V_i]$ starting from some $u \in V_i$. By condition (4) of an $(\varepsilon, \delta, n^{1.5})$-good decomposition (Definition 2.1), the minimal degree of $G[V_i]$ is at least $\delta^4 n/40$. We apply Claim 3.3 to the graph $G[V_i]$, with $\varepsilon^8 \theta$ playing the role of ε in the claim, to obtain that for every $u \in V_i$

$$\mathbb{P}_u \left(Y \left[0, 2 \left(t_{\text{mix}}^{G[V_i]}(\varepsilon^8 \theta) + \lfloor \sqrt{n} \rfloor \right) \right] \cap W \neq \varnothing \right) \geq \frac{\varepsilon^8 \theta \delta^4}{160}. \tag{23}$$

By definition of an $(\varepsilon, \delta, n^{1.5})$-good decomposition (Definition 2.1) we have $\theta \geq \varepsilon^{11 \cdot 2^{2/\delta}}$. Hence by (3), Lemma 3.2 and condition (3) of Definition 2.1 we get

$$t_{\text{mix}}^{G[V_i]}(\varepsilon^8 \theta) + \lfloor \sqrt{n} \rfloor \leq (8 + 11 \cdot 2^{2/\delta})(t_{\text{mix}}^{G[V_i]}(\varepsilon/2)) + \lfloor \sqrt{n} \rfloor \leq \frac{B \log(1/\varepsilon) \sqrt{n}}{\theta}, \tag{24}$$

for some $B = B(\delta)$. By Lemma 3.5

$$\mathbb{P} \left(X_1 \in V_i \text{ and } \exists t \leq 2(t_{\text{mix}}^{G[V_i]}(\varepsilon^8 \theta) + \lfloor \sqrt{n} \rfloor) \text{ with } X_t \notin V_i \right) \leq \frac{2B\theta \log(1/\varepsilon)\varepsilon^9}{\delta^2}. \tag{25}$$

Hence, conditioned on $X_1 \in V_i$, if we set $Y_0 = X_1$ then we can couple these two walks such that $Y_t = X_{t+1}$ for all $t \leq 2(t_{\text{mix}}^{G[V_i]}(\varepsilon^8 \theta) + \lfloor \sqrt{n} \rfloor)$ with failure probability bounded by the right-hand side of (25). This and (23) imply that

$$\mathbb{P}_v \left(X[0, 2(t_{\text{mix}}^{G[V_i]}(\varepsilon^8 \theta) + \lfloor \sqrt{n} \rfloor)] \cap W \neq \varnothing \right) \geq \frac{1}{k} \left(\frac{\varepsilon^8 \theta \delta^4}{160} - \frac{2B\theta \log(1/\varepsilon)\varepsilon^9}{\delta^2} \right).$$

Plugging in (23) and (25) we obtain that the right-hand side is bounded from below by

$$\frac{1}{k} \left(\frac{\varepsilon^8 \theta \delta^4}{160} - \frac{2B\theta \log(1/\varepsilon)\varepsilon^9}{\delta^2} \right),$$

which is lower bounded by some $c = c(\varepsilon, \delta)$. Now (22) follows by (24) and taking $C = \frac{B \log(1/\varepsilon)}{\theta}$. Now by (22) and the Markov property, for any positive integer m and any $t \in [mC\sqrt{n}, (m+1)C\sqrt{n}]$ we have

$$\boldsymbol{p}_W^t(v, v) \leq \frac{(1-c)^m}{\delta n},$$

where the denominator accounts for the last step returning to v. We conclude that

$$\mathcal{B}_W(G) = \sum_{t=0}^{\infty} (t+1) \boldsymbol{p}_W^t(v, v) \leq \sum_{m=0}^{\infty} \sum_{t=mC\sqrt{n}}^{(m+1)C\sqrt{n}} ((m+1)C\sqrt{n}) \frac{(1-c)^m}{\delta n}$$

$$\leq \frac{C^2}{\delta} \sum_{m=0}^{\infty} (m+1)(1-c)^m,$$

and this concludes our proof since the infinite sum above converges. $\qquad\square$

Proof of Theorem 1.1. Let $\varepsilon > 0$ and let $G = (V, E)$ be a connected simple graph on n vertices with minimal degree at least δn. By Lemma 2.2, there exists an $(\varepsilon, \delta, n^{1.5})$ good decomposition of G, denoted by $V = V_1 \sqcup \ldots \sqcup V_k$. By Theorem 3.1 there exist some $C' = C'(\delta)$ and k pairs of distinct vertices $v_1^i, v_2^i \in V_i$, such that if φ_i is the random path between v_1^i and v_2^i in $\mathsf{UST}(G)$, then

$$\mathbb{P}\left(\forall i \in [k]: \varepsilon^8 \theta \sqrt{n} \le |\varphi_i| \le \frac{\sqrt{n}}{\theta \varepsilon^8} \text{ and } \varphi_i \subseteq V_i\right) \ge 1 - C' k \varepsilon^2. \tag{26}$$

We condition on this event and on the collection of paths $(\varphi_i)_{i \in [k]}$ and denote by W the set of vertices of $(\varphi_i)_{i \in [k]}$. Let H be the graph obtained from G by contracting each φ_i into a single vertex. Then Lemma 1.5 implies that $\mathsf{UST}(H) \cup \{\varphi_i\}_{i \in [k]}$ has the distribution of $\mathsf{UST}(G)$. Hence $\mathrm{diam}(\mathsf{UST}(G)) \le \mathrm{diam}(\mathsf{UST}(H)) + |W|$. Denote also by $\mathcal{T}_W$ the UST on G/W. By Lemma 1.8, we have that $\mathsf{UST}(H)$ stochastically dominates $\mathcal{T}_W := \mathsf{UST}(G/W)$ (when viewed as random subsets of $E(G)$) hence there is a coupling such that $\mathcal{T}_W \subseteq \mathsf{UST}(H)$ and since H has $k - 1$ vertices more than G/W we deduce that $\mathsf{UST}(H)$ is a union of $\mathcal{T}_W$ and at most $k - 1$ more edges. Hence the diameter of $\mathsf{UST}(H))$ is at most $k - 1$ times the diameter of $\mathcal{T}_W$. We conclude that

$$\mathrm{diam}(\mathsf{UST}(G)) \le (k - 1)\mathrm{diam}(\mathcal{T}_W) + |W|,$$

Now, Lemma 4.2 implies that the set W has $\mathcal{B}_W(G) \le C$ for some $C = C(\varepsilon, \delta)$ so that Lemma 4.1 gives

$$\mathbb{P}(\mathrm{diam}(\mathcal{T}_W) \ge \ell) \le \frac{C_3 |W|}{\ell}, \tag{27}$$

for $C_3 = C_3(\varepsilon, \delta)$ and any $\ell \ge 1$. Hence under our conditioning we have that for any $A > k/\theta \varepsilon^8$

$$\mathbb{P}(\mathsf{UST}(G) \ge A\sqrt{n}) \le \mathbb{P}\left(\mathrm{diam}(\mathcal{T}_W) \ge (A - k/\theta \varepsilon^8)\sqrt{n}/(k - 1)\right) \le \frac{C_3 k(k - 1)}{\theta \varepsilon^8 (A - k/\theta \varepsilon^8)},$$

which together with (26) concludes the proof. □

References

[1] Aldous, D. J. (1991) The continuum random tree. I. *Ann. Prob.* **19**(1) 1–28.

[2] Aldous, D. J. (1991) The continuum random tree. II. An overview. In *Stochastic Analysis (Durham, 1990)*, pp. 23–70.

[3] Aldous, D. J. (1993) The continuum random tree. III. *Ann. Prob.* **21**(1) 248–289.

[4] Alon, N. (1986) Eigenvalues and expanders. *Combinatorica* **6** 83–96.

[5] Alon, N. and Milman, V. D. (1985) λ1; isoperimetric inequalities for graphs, and superconcentrators. *J. Comb. Theory Ser. B* **38**(1) 73–88.

[6] Feder, T. and Mihail, M. (1992) Balanced matroids. In *Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing*, pp. 26–38.

[7] Jerrum, M. and Sinclair, A. (1989) Approximating the permanent. *SIAM J. Comput.* **18**(6) 1149–1178.

[8] Jerrum, M., Son, J.-B., Tetali, P. and Vigoda, E. (2004) Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains. *Ann. Appl. Prob.* **14**(4) 1741–1765 (English).

[9] Kolchin, V. F. (1986) *Random Mappings, Translation Series in Mathematics and Engineering*, Optimization Software, Inc., Publications Division, New York,. Translated from the Russian, With a foreword by S. R. S. Varadhan.

[10] Kwok, T. C., Lau, L. C., Lee, Y. T., Gharan, S. O. and Trevisan, L. (2013) Improved Cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap. In *STOC'13| Proceedings of the 2013 ACM Symposium on Theory of Computing*, pp. 11–20.

[11] Lawler, G. F. and Sokal, A. D. (1988) Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Trans. Am. Math. Soc.* **309**(2) 557–580.

[12] Le Gall, J.-F. (2005) Random trees and applications. *Probab. Surv.* **2** 245–311.

[13] Le Gall, J.-F. (2006) Random real trees. *Ann. Fac. Sci. Toulouse Math. (6)* **15**(1) 35–62.

[14] Levin, D. A. and Peres, Y. (2017) *Markov Chains and Mixing Times*, American Mathematical Society, Providence, RI. Second edition of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson. Available at https://pages.uoregon.edu/dlevin/MARKOV/mcmt2e.pdf.

[15] Lyons, R. and Peres, Y. (2016) *Probability on Trees and Networks*, Vol. 42 of Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, New York.

[16] Michaeli, P., Nachmias, A. and Shalev, M. (2021) The diameter of uniform spanning trees in high dimensions. *Probab. Theory Related Fields* **179**(1–2) 261–294.

[17] Peres, Y. and Revelle, D. (2004) Scaling limits of the uniform spanning tree and loop-erased random walk onfinite graphs, arXiv Mathematics e-prints, available at math/0410430.

[18] Szekeres, G. (1983) Distribution of labelled trees by diameter. In *Combinatorial Mathematics X (Adelaide, 1982), Lecture Notes in Mathematics,* pp. 392–397.

[19] Wilson, D. B. (1996) Generating random spanning trees more quickly than the cover time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)*, pp. 296–303.