MIXING TIME OF NEAR-CRITICAL RANDOM GRAPHS

JIAN DING, EYAL LUBETZKY AND YUVAL PERES

ABSTRACT. Let $\mathcal{C}_1$ be the largest component of the Erdős-Rényi random graph $\mathcal{G}(n, p)$. The mixing time of random walk on $\mathcal{C}_1$ in the strictly supercritical regime, $p = c/n$ with fixed $c > 1$, was shown to have order $\log^2 n$ by Fountoulakis and Reed, and independently by Benjamini, Kozma and Wormald. In the critical window, $p = (1 + \varepsilon)/n$ where $\lambda = \varepsilon^3 n$ is bounded, Nachmias and Peres proved that the mixing time on $\mathcal{C}_1$ is of order n. However, it was unclear how to interpolate between these results, and estimate the mixing time as the giant component emerges from the critical window. Indeed, even the asymptotics of the diameter of $\mathcal{C}_1$ in this regime were only recently obtained by Riordan and Wormald, as well as the present authors and Kim.

In this paper we show that for $p = (1 + \varepsilon)/n$ with $\lambda = \varepsilon^3 n \to \infty$ and $\lambda = o(n)$, the mixing time on $\mathcal{C}_1$ is with high probability of order $(n/\lambda) \log^2 \lambda$. In addition, we show that this is the order of the largest mixing time over all components, both in the slightly supercritical and in the slightly subcritical regime (i.e., $p = (1 - \varepsilon)/n$ with λ as above).

1. INTRODUCTION

There is a rich interplay between geometric properties of a graph and the behavior of a random walk on it (see, e.g., [1]). A particularly important parameter is the mixing time, which measures the rate of convergence to stationarity. In this paper we focus on random walks on the classical Erdős-Rényi random graph $\mathcal{G}(n, p)$.

The geometry of $\mathcal{G}(n, p)$ has been studied extensively since its introduction in 1959 by Erdős and Rényi [10]. An intriguing phenomenon exhibited by this model is the *double jump*: For $p = c/n$ with c fixed, the largest component $\mathcal{C}_1$ has size $O(\log n)$ with high probability (w.h.p.), when $c < 1$, it is w.h.p. linear in n for $c > 1$, and for $c = 1$ its size has order $n^{2/3}$ (the latter was proved by Bollobás [5] and Łuczak [21]). Bollobás discovered that the critical behavior extends throughout $p = (1 \pm \varepsilon)/n$ for $\varepsilon = O(n^{-1/3})$, a regime known as the *critical window*.

Only in recent years were the tools of Markov chain analysis and the understanding of the random graph sufficiently developed to enable estimating mixing times on $\mathcal{C}_1$. Fountoulakis and Reed [12] showed that, in the strictly supercritical regime ($p = c/n$ with fixed $c > 1$), the mixing time of random

walk on $\mathcal{C}_1$ w.h.p. has order $\log^2 n$. Their proof exploited fairly simple geometric properties of $\mathcal{G}(n, p)$, while the key to their analysis was a refined bound [11] on the mixing time of a general Markov chain. The same result was obtained independently by Benjamini, Kozma and Wormald [3]. There, the main innovation was a decomposition theorem for the giant component. However, the methods of these two papers do not yield the right order of the mixing time when c is allowed to tend to 1.

Nachmias and Peres [26] proved that throughout the critical window the mixing time on $\mathcal{C}_1$ is of order n. The proof there used branching process arguments, which were effective since the critical $\mathcal{C}_1$ is close to a tree.

It was unclear how to interpolate between these results, and estimate the mixing time as the giant component emerges from the critical window, since the methods used for the supercritical and the critical case were so different. The focus of this paper is primarily on the emerging supercritical regime, where $p = (1 + \varepsilon)/n$ with $\varepsilon^3 n \to \infty$ and $\varepsilon = o(1)$. In this regime, the largest component is significantly larger than the others, yet its size is still sublinear. Understanding the geometry of $\mathcal{C}_1$ in this regime has been challenging: Indeed, even the asymptotics of its diameter were only recently obtained by Riordan and Wormald [29], as well as in [9].

Our main result determines the order of the mixing time throughout the emerging supercritical regime (see Subsection 2.3 for a formal definition of mixing time).

Theorem 1 (supercritical regime). *Let $\mathcal{C}_1$ be the largest component of $\mathcal{G}(n, p)$ for $p = \frac{1+\varepsilon}{n}$, where $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$. With high probability, the mixing time of the lazy random walk on $\mathcal{C}_1$ is of order $\varepsilon^{-3} \log^2(\varepsilon^3 n)$.*

While the second largest component $\mathcal{C}_2$ has a mixing time of smaller order (it is w.h.p. a tree, and given that event, it is a uniform tree on its vertices, and as such has mixing time of order $|\mathcal{C}_2|^{3/2} \asymp \varepsilon^{-3} \log^{3/2}(\varepsilon^3 n)$), it turns out that w.h.p. there exists an even smaller component, whose mixing time is of the same order as on $\mathcal{C}_1$. This is captured by our second theorem, which also handles the subcritical regime.

Theorem 2 (controlling all components). *Let $G \sim \mathcal{G}(n, p)$ for $p = (1 \pm \varepsilon)/n$, where $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$. Let $\mathcal{C}^\star$ be the component of G that maximizes the mixing time of the lazy random walk on it, denoted by $t^\star_{\mathrm{MIX}}$. Then with high probability, $t^\star_{\mathrm{MIX}}$ has order $\varepsilon^{-3} \log^2(\varepsilon^3 n)$. This also holds when maximizing only over tree components.*

In the area of random walk on random graphs, the following two regimes have been analyzed extensively.

- The supercritical regime, where $t_{\mathrm{MIX}} \asymp (\mathrm{diam})^2$ with diam denoting the intrinsic diameter in the percolation cluster. Besides $\mathcal{G}(n, \frac{c}{n})$ for $c > 1$, this also holds in the torus $\mathbb{Z}_n^d$ by [6] and [24].
- The critical regime on a high dimensional torus, where $t_{\mathrm{MIX}} \asymp (\mathrm{diam})^3$. As mentioned above, for critical percolation on the complete graph, this was shown in [26]. For high dimensional tori, this is a consequence of [13].

To the best of our knowledge, our result is the first interpolation for the mixing time between these two different powers of the diameter.

2. Preliminaries

2.1. **Cores and kernels.** The *k-core* of a graph G, denoted by $G^{(k)}$, is the maximum subgraph $H \subset G$ where every vertex has degree at least k. It is well known (and easy to see) that this subgraph is unique, and can be obtained by repeatedly deleting any vertex whose degree is smaller than k (at an arbitrary order).

We call a path $\mathcal{P} = v_0, v_1, \ldots, v_k$ for $k > 1$ (i.e., a sequence of vertices with $v_i v_{i+1}$ an edge for each i) a *2-path* if and only if v_i has degree 2 for all $i = 1, \ldots, k-1$ (while the endpoints v_0, v_k may have degree larger than 2, and possibly $v_0 = v_k$).

The *kernel* $\mathcal{K}$ of G is obtained by taking its 2-core $G^{(2)}$ minus its disjoint cycles, then repeatedly contracting all 2-paths (replacing each by a single edge). Note that, by definition, the degree of every vertex in $\mathcal{K}$ is at least 3.

2.2. **Structure of the supercritical giant component.** The key to our analysis of the random walk on the giant component $\mathcal{C}_1$ is the following result from our companion paper [8]. This theorem completely characterizes the structure of $\mathcal{C}_1$, by reducing it to a tractable contiguous model $\tilde{\mathcal{C}}_1$.

Theorem 2.1. [8] *Let $\mathcal{C}_1$ be the largest component of $\mathcal{G}(n, p)$ for $p = \frac{1+\varepsilon}{n}$, where $\varepsilon^3 n \to \infty$ and $\varepsilon \to 0$. Let $\mu < 1$ denote the conjugate of $1 + \varepsilon$, that is, $\mu e^{-\mu} = (1 + \varepsilon) e^{-(1+\varepsilon)}$. Then $\mathcal{C}_1$ is contiguous to the following model $\tilde{\mathcal{C}}_1$:*

1. *Let $\Lambda \sim \mathcal{N}\left(1 + \varepsilon - \mu, \frac{1}{\varepsilon n}\right)$ and assign i.i.d. variables $D_u \sim \mathrm{Poisson}(\Lambda)$ ($u \in [n]$) to the vertices, conditioned that $\sum D_u \mathbb{1}_{\{D_u \geq 3\}}$ is even. Let*

$$N_k = \#\{u : D_u = k\} \quad and \quad N = \textstyle\sum_{k \geq 3} N_k \,.$$

   Select a random multigraph $\mathcal{K}$ on N vertices, uniformly among all multigraphs with N_k vertices of degree k for $k \geq 3$.
2. *Replace the edges of $\mathcal{K}$ by paths of lengths i.i.d. $\mathrm{Geom}(1 - \mu)$.*
3. *Attach an independent $\mathrm{Poisson}(\mu)$-Galton-Watson tree to each vertex.*

That is, $\mathbb{P}(\tilde{\mathcal{C}}_1 \in \mathcal{A}) \to 0$ implies $\mathbb{P}(\mathcal{C}_1 \in \mathcal{A}) \to 0$ for any set of graphs $\mathcal{A}$.

In the above, a Poisson(μ)-Galton-Watson tree is the family tree of a Galton-Watson branching process with offspring distribution Poisson(μ). We will use the abbreviation PGW(μ)-tree for this object.

Note that conditioning on $\sum D_u \mathbb{1}_{\{D_u \geq 3\}}$ being even does not pose a problem, as one can easily use rejection sampling. The 3 steps in the description of $\tilde{\mathcal{C}}_1$ correspond to constructing its kernel $\mathcal{K}$ (Step 1), expanding $\mathcal{K}$ into the 2-core $\tilde{\mathcal{C}}_1^{(2)}$ (Step 2), and finally attaching trees to it to obtain $\tilde{\mathcal{C}}_1$ (Step 3).

Further observe that $N_k = \Theta(\varepsilon^k n)$ for any fixed $k \geq 2$, and so in the special case where $\varepsilon = o(n^{-1/4})$ w.h.p. we have $D_u \in \{0, 1, 2, 3\}$ for all $u \in [n]$, and the kernel $\mathcal{K}$ is simply a uniform 3-regular multigraph.

Combining the above description of the giant component with standard tools in the study of random graphs with given degree-sequences, one can easily read off useful geometric properties of the kernel. This is demonstrated by the following lemma of [8], for which we require a few definitions: For a subset S of the vertices of a graph G, we let

$$d_G(S) \triangleq \sum_{v \in S} d_G(v)$$

denote the sum of the degrees of its vertices (also referred to as the *volume* of S in G). The *isoperimetric number* of a graph G is defined to be

$$i(G) \triangleq \min \left\{ \frac{e(S, S^c)}{d_G(S)} \; : \; S \subset V(G) \, , \, d_G(S) \leq e(G) \right\} \, .$$

Lemma 2.2 ([8, Lemma 3.5])**.** *Let $\mathcal{K}$ be the kernel of the largest component $\mathcal{C}_1$ of $\mathcal{G}(n, p)$ for $p = \frac{1+\varepsilon}{n}$, where $\varepsilon^3 n \to \infty$ and $\varepsilon \to 0$. Then w.h.p.,*

$$|\mathcal{K}| = \left(\tfrac{4}{3} + o(1) \right) \varepsilon^3 n \, , \quad e(\mathcal{K}) = (2 + o(1)) \varepsilon^3 n \, ,$$

and $i(\mathcal{K}) \geq \alpha$ for some absolute constant $\alpha > 0$.

2.3. **Notions of mixing of the random walk.** Let $G = (V, E)$ be a connected graph. For any vertex $v \in V$, let $d_G(v)$ denote its degree, and similarly put $d_G(S) = \sum_{v \in V} d_G(v)$ for $S \subset V$.

For any two distributions φ, ψ on V, the *total-variation distance* of φ and ψ is defined as

$$\|\varphi - \psi\|_{\mathrm{TV}} \triangleq \sup_{S \subset V} |\varphi(S) - \psi(S)| = \frac{1}{2} \sum_{v \in V} |\varphi(v) - \psi(v)| \, .$$

Let (S_t) denote the lazy random walk on G, i.e., the Markov chain which at each step holds its position with probability $\frac{1}{2}$ and otherwise moves to a uniformly chosen neighbor. This is an aperiodic and irreducible Markov chain, whose stationary distribution π is given by

$$\pi(x) = d_G(x)/2|E| \, .$$

We next define two notions of measuring the distance of an ergodic Markov chain (S_t), defined on a state-set V, from its stationary distribution π.

Let $0 < \delta < 1$. The (worst-case) *total-variation mixing time* of (S_t) with parameter δ, denoted by $t_{\mathrm{MIX}}(\delta)$, is defined to be

$$t_{\mathrm{MIX}}(\delta) \triangleq \min\left\{ t : \max_{v \in V} \|\mathbb{P}_v(S_t \in \cdot) - \pi\|_{\mathrm{TV}} \leq \delta \right\},$$

where $\mathbb{P}_v$ denotes the probability given that $S_0 = v$.

The *Cesàro mixing time* (also known as the approximate uniform mixing time) of (S_t) with parameter δ, denoted by $\tilde{t}_{\mathrm{MIX}}(\delta)$, is defined as

$$\tilde{t}_{\mathrm{MIX}}(\delta) = \min\left\{ t : \max_{v \in V} \left\| \pi - \frac{1}{t} \sum_{i=0}^{t-1} \mathbb{P}_v(S_i \in \cdot) \right\|_{\mathrm{TV}} \leq \delta \right\}.$$

When discussing the order of the mixing-time it is customary to choose $\delta = \frac{1}{4}$, in which case we will use the abbreviations $t_{\mathrm{MIX}} = t_{\mathrm{MIX}}(\frac{1}{4})$ and $\tilde{t}_{\mathrm{MIX}} = \tilde{t}_{\mathrm{MIX}}(\frac{1}{4})$.

By results of [2] and [19] (see also [20]), the mixing time and the Cesàro mixing time have the same order for lazy reversible Markov chains (i.e., discrete-time chains whose holding probability in each state is at least $\frac{1}{2}$), as formulated by the following theorem.

Theorem 2.3. *Every lazy reversible Markov chain satisfies*

$$c_1 \, \tilde{t}_{\mathrm{MIX}}(\tfrac{1}{4}) \leq t_{\mathrm{MIX}}(\tfrac{1}{4}) \leq c_2 \, \tilde{t}_{\mathrm{MIX}}(\tfrac{1}{4})$$

for some absolute constants $c_1, c_2 > 0$.

Proof. The first inequality is straightforward and does not require laziness or reversibility. We include its proof for completeness. Notice that

$$\left\| \pi - \frac{1}{t} \sum_{i=0}^{t-1} \mathbb{P}_v(S_i \in \cdot) \right\|_{\mathrm{TV}} \leq \frac{1}{8} + \frac{1}{t} \sum_{i=t/8}^{t-1} \left\| \pi - \mathbb{P}_v(S_i \in \cdot) \right\|_{\mathrm{TV}}$$

$$\leq \frac{1}{8} + \|\pi - \mathbb{P}_v(S_{t/8} \in \cdot)\|_{\mathrm{TV}},$$

where we used the fact that $\|\pi - \mathbb{P}_v(S_t \in \cdot)\|$ is decreasing in t. Taking $t = 8t_{\mathrm{MIX}}(\frac{1}{8})$, we obtain that $\tilde{t}_{\mathrm{MIX}}(\frac{1}{4}) \leq 8t_{\mathrm{MIX}}(\frac{1}{8})$ and conclude the proof of the first inequality using the well-known fact that $t_{\mathrm{MIX}}(\frac{1}{8}) \leq 4t_{\mathrm{MIX}}(\frac{1}{4})$.

The second inequality of the theorem is significantly more involved: By combining [19, Theorem 5.4] (for a stronger version, see [20, Theorem 4.22]) and [2, Theorem C], it follows that the order of the Cesàro mixing time can be bounded by that of the mixing time for the corresponding continuous-time Markov chain. Now, using a well-known fact that the mixing time for the lazy Markov chain and the continuous-time chain have the same order (see, e.g., [16, Theorem 20.3]), the proof is concluded. ∎

Let Γ be a stopping rule (a randomized stopping time) for (S_t). That is, $\Gamma : G \times \Lambda \to \mathbb{N}$ for some probability space Λ, such that $\Gamma(\cdot, \omega)$ is a stopping time for every $\omega \in \Lambda$. Let $\sigma^\Gamma \triangleq \mathbb{P}_\sigma(S_\Gamma \in \cdot)$ when σ is a distribution on V.

Let σ, ν be two distributions on V. Note that there is always a stopping rule Γ such that $\sigma^\Gamma = \nu$, e.g., draw a vertex z according to ν and stop when reaching z. The *access time* from σ to ν, denoted by $H(\sigma, \nu)$, is the minimum expected number of steps over all such stopping rules:

$$H(\sigma, \nu) \triangleq \min_{\Gamma \,:\, \sigma^\Gamma = \nu} \mathbb{E}\Gamma \,.$$

It is easy to verify that $H(\sigma, \nu) = 0$ iff $\sigma = \nu$ and that $H(\cdot, \cdot)$ satisfies the triangle-inequality, however it is not necessarily symmetric.

The *approximate forget time* of G with parameter $0 < \delta < 1$ is defined by

$$\mathcal{F}_\delta = \min_\varphi \max_\sigma \min_{\nu \,:\, \|\nu - \varphi\|_{\mathrm{TV}} \leq \delta} H(\sigma, \nu) \,.$$

Combining Theorem 3.2 and Corollary 5.4 in [20], one immediately obtains that the approximate forget time and the Cesàro mixing time have the same order, as stated in the following theorem.

Theorem 2.4. *Every reversible Markov chain satisfies*

$$c_1 \, \mathcal{F}_{1/4} \leq \tilde{t}_{\mathrm{MIX}}(\tfrac{1}{4}) \leq c_2 \, \mathcal{F}_{1/4}$$

for some absolute constants $c_1, c_2 > 0$.

2.4. **Conductance and mixing.** Let $P = (p_{x,y})_{x,y}$ be the transition kernel of an irreducible, reversible and aperiodic Markov chain on Ω with stationary distribution π. For $S \subset \Omega$, define the *conductance* of the set S to be

$$\Phi(S) \triangleq \frac{\sum_{x \in S, y \notin S} \pi(x) p_{x,y}}{\pi(S)\pi(\Omega \setminus S)} \,.$$

We define Φ, the conductance of the chain, by $\Phi \triangleq \min\{\Phi(S) \,:\, \pi(S) \leq \tfrac{1}{2}\}$ (In the special case of a lazy random walk on a connected regular graph, this quantity is similar to the isoperimetric number of the graph, defined earlier). A well-known result of Jerrum and Sinclair [14] states that t_{MIX} is of order at most $\Phi^{-2} \log \pi_{\min}^{-1}$, where $\pi_{\min} = \min_{x \in \Omega} \pi(x)$. This bound was fine-tuned by Lovász and Kannan [18] to exploit settings where the conductance of the average set S plays a dominant role (rather than the worst set). For our upper bound of the mixing time on the random walk on the 2-core, we will use an enhanced version of the latter bound (namely, Theorem 3.6) due to Fountoulakis and Reed [11].

2.5. **Edge set notations.** Throughout the paper we will use the following notations, which will be handy when moving between the kernel and 2-core.

For $S \subset G$, let $E_G(S)$ denote the set of edges in the induced subgraph of G on S, and let $\partial_G S$ denote the edges between S and its complement $S^c \triangleq V(G) \setminus S$. Let

$$\bar{E}_G(S) \triangleq E_G(S) \cup \partial_G(S) \,,$$

and define $e_G(S) \triangleq |E_G(S)|$. We omit the subscript G whenever its identity is made clear from the context.

If $\mathcal{K}$ is the kernel in the model $\tilde{\mathcal{C}}_1$ and $\mathcal{H}$ is its 2-core, let

$$E^\star_\mathcal{H} : 2^{E(\mathcal{K})} \to 2^{E(\mathcal{H})}$$

be the operator which takes a subset of edges $T \subset E(\mathcal{K})$ and outputs the edges lying on their corresponding 2-paths in $\mathcal{H}$. For $S \subset V(\mathcal{K})$, we let

$$E^\star_\mathcal{H}(S) \triangleq E^\star_\mathcal{H}(E_\mathcal{K}(S)) \,, \quad \bar{E}^\star_\mathcal{H}(S) \triangleq E^\star_\mathcal{H}\left(\bar{E}_\mathcal{K}(S)\right) \,.$$

3. RANDOM WALK ON THE 2-CORE

The In this section, we analyze the properties of the random walk on the 2-core $\tilde{\mathcal{C}}_1^{(2)}$.

3.1. **Mixing time of the 2-core.** By the definition of our new model $\tilde{\mathcal{C}}_1$, we can study the 2-core $\mathcal{C}_1^{(2)}$ via the well-known configuration model (see, e.g., [4] for further details on this method). To simplify the notation, we let $\mathcal{H}$ denote the 2-core of $\tilde{\mathcal{C}}_1$ throughout this section.

The main goal of the subsection is to establish the mixing time of the lazy random walk on $\mathcal{H}$, as stated by the following theorem.

Theorem 3.1. *With high probability, the lazy random walk on $\mathcal{H}$ has a Cesàro mixing time $\tilde{t}_{\mathrm{MIX}}$ of order $\varepsilon^{-2} \log^2(\varepsilon^3 n)$. Consequently, w.h.p. it also satisfies $t_{\mathrm{MIX}} = \Theta\left(\varepsilon^{-2} \log^2(\varepsilon^3 n)\right)$.*

We will use a result of Fountoulakis and Reed [12], which bounds the mixing time in terms of the isoperimetric profile of the graph (measuring the expansion of sets of various volumes). As a first step in obtaining this data for the supercritical 2-core $\mathcal{H}$, the next lemma will show that a small subset of the kernel, $S \subset \mathcal{K}$, cannot have too many edges in $\bar{E}_\mathcal{H}(S)$.

Lemma 3.2. *For $v \in \mathcal{K}$, define*

$$\mathfrak{C}_{v,K} \triangleq \{S \ni v : \ |S| = K \text{ and } S \text{ is a connected subgraph of } \mathcal{K} \ \} \,.$$

The following holds w.h.p. for every $v \in \mathcal{K}$, integer K and $S \in \mathfrak{C}_{v,K}$:

(1) $|\mathfrak{C}_{v,K}| \leq \exp\left[5(K \vee \log(\varepsilon^3 n))\right]$.
(2) $d_\mathcal{K}(S) \leq 30(K \vee \log(\varepsilon^3 n))$.

Proof. By definition, $\Lambda = (2 + o(1))\varepsilon$ w.h.p., thus standard concentration arguments imply that the following holds w.h.p.:

$$N_3 = \left(\tfrac{4}{3} + o(1)\right)\varepsilon^3 n \quad \text{and} \quad N_k \leq \frac{(3\varepsilon)^k \log(1/\varepsilon)}{k!} n \text{ for } k \geq 4. \qquad (3.1)$$

Assume that the above indeed holds, and notice that the lemma trivially holds when $K \geq \varepsilon^3 n$. We may therefore further assume that $K \leq \varepsilon^3 n$.

Consider the following exploration process, starting from the vertex v. Initialize S to be $\{v\}$, and mark $v_1 = v$. At time $i \geq 1$, we explore the neighborhood of v_i (unless $|S| < i$), and for each its neighbors that does not already belong to S, we toss a fair coin to decide whether or not to insert it to S. Newly inserted vertices are labeled according to the order of their arrival; that is, if $|S| = k$ prior to the insertion, we give the new vertex the label v_{k+1}. Finally, if $|S| < i$ at time i then we stop the exploration process.

Let X_i denote the degree of the vertex v_i in the above defined process. In order to stochastically dominate X_i from above, observe that the worst case occurs when each of the vertices in $v_1, \ldots, v_{i-1}$ has degree 3. With this observation in mind, let A be a set consisting of $N_3 - K$ vertices of degree 3 and N_k vertices k (for $k \geq 4$). Sample a vertex proportional to the degree from A and let Y denote its degree. Clearly, $X_i \preceq Y_i$, where Y_i are independent variables distributed as Y, and so

$$d_{\mathcal{K}}(S) \preceq \sum_{i=1}^{K} Y_i. \qquad (3.2)$$

By the definition of our exploration process,

$$|\mathfrak{C}_{v,K}| \preceq \sum_{\ell_1 + \cdots + \ell_K = K} \prod_{i=1}^{K} \binom{Y_i}{\ell_i}.$$

We can now deduce that

$$\mathbb{E}|\mathfrak{C}_{v,K}| \leq \mathbb{E}\left[\sum_{\ell_1 + \cdots + \ell_K = K} \prod_{i=1}^{K} \binom{Y_i}{\ell_i}\right] = \sum_{\ell_1 + \cdots + \ell_K = K} \prod_{i=1}^{K} \mathbb{E}\left[\binom{Y_i}{\ell_i}\right]. \qquad (3.3)$$

For all $i \geq 4$, we have

$$\mathbb{P}(Y = i) \leq 27i\frac{(3\varepsilon)^{i-3} \log(1/\varepsilon)}{i!} = 27\frac{(3\varepsilon)^{i-3} \log(1/\varepsilon)}{(i-1)!}$$

and therefore, for sufficiently large n (recall that $\varepsilon = o(1)$),

$$\mathbb{E}\left[\binom{Y}{k}\right] \leq \binom{3}{k} + \sum_{i \geq 4} \binom{i}{k} \cdot 27\frac{(3\varepsilon)^{i-3} \log(1/\varepsilon)}{(i-1)!} \leq \frac{7}{k!} \text{ for all } k,$$

Altogether,

$$\mathbb{E}|\mathfrak{C}_{v,K}| \leq 7^K \sum_{\ell_1 + \cdots + \ell_K = K} \prod_{i=1}^{K} \frac{1}{\ell_i!}. \tag{3.4}$$

The next simple claim will provide a bound on the sum in the last expression.

Claim 3.3. *The function* $f(n) = \sum_{\ell_1 + \cdots + \ell_n = n} \prod_{k=1}^{n} \frac{1}{\ell_k!}$ *satisfies* $f(n) \leq \mathrm{e}^n$.

Proof. The proof is by induction. For $n = 1$, the claim trivially holds. Assuming the hypothesis is valid for $n \leq m$, we get

$$f(m+1) = \sum_{k=0}^{m+1} \frac{1}{k!} f(m-k) \leq \sum_{k=0}^{m+1} \frac{\mathrm{e}^{m-k}}{k!} \leq \mathrm{e}^m \sum_{k=0}^{m+1} \frac{1}{k!} \leq \mathrm{e}^{m+1},$$

as required. ∎

Plugging the above estimate into (3.4), we conclude that $\mathbb{E}|\mathfrak{C}_{v,K}| \leq (7\mathrm{e})^K$. Now, Markov's inequality, together with a union bound over all the vertices in the kernel $\mathcal{K}$ yield the Part (1) of the lemma.

For Part (2), notice that for any sufficiently large n,

$$\mathbb{E}\mathrm{e}^Y \leq \mathrm{e}^3 + \sum_{i \geq 4} \mathrm{e}^i 27i \frac{(3\varepsilon)^{i-3} \log(1/\varepsilon)}{i!} \leq 25,$$

Therefore, (3.2) gives that

$$\mathbb{P}\left(d_{\mathcal{K}}(S) \geq 30\left(K \vee \log(\varepsilon^3 n)\right)\right) \leq \exp\left[-5\left(K \vee \log(\varepsilon^3 n)\right)\right].$$

At this point, the proof is concluded by a union bound over $\mathfrak{C}_{v,K}$ for all $v \in \mathcal{K}$ and $K \leq \varepsilon^3 n$, using the upper bound we have already derived for $|\mathfrak{C}_{v,K}|$ in the Part (1) of the lemma. ∎

Lemma 3.4. *Let* $\mathcal{L} \subset E(\mathcal{K})$ *be the set of loops in the kernel. With high probability, every subset of vertices* $S \subset \mathcal{K}$ *forming a connected subgraph of* $\mathcal{K}$ *satisfies* $|\bar{E}_{\mathcal{H}}^{\star}(S)| \leq (100/\varepsilon)\left(|S| \vee \log(\varepsilon^3 n)\right)$, *and every subset* T *of* $\frac{1}{20}\varepsilon^3 n$ *edges in* $\mathcal{K}$ *satisfies* $|E_{\mathcal{H}}^{\star}(T) \cup E_{\mathcal{H}}^{\star}(\mathcal{L})| \leq \frac{3}{4}\varepsilon^2 n$.

Proof. Assume that the events given in Parts (1),(2) of Lemma 3.2 hold. Further note that, by definition of the model $\tilde{\mathcal{C}}_1$, a standard application of CLT yields that w.h.p.

$$|\mathcal{K}| = (\tfrac{4}{3} + o(1))\varepsilon^3 n, \quad e(\mathcal{H}) = (2 + o(1))\varepsilon^2 n, \quad e(\mathcal{K}) = (2 + o(1))\varepsilon^3 n.$$

By Part (2) of that lemma, $d_{\mathcal{K}}(S) \leq 30\left(|S| \vee \log(\varepsilon^3 n)\right)$ holds simultaneously for every connected set S, hence there are at most this many edges in $\bar{E}_{\mathcal{K}}(S)$.

Let $S \subset \mathcal{K}$ be a connected set of size $|S| = s$, and let

$$K = K(s) = s \vee \log(\varepsilon^3 n).$$

Recalling our definition of the graph $\mathcal{H}$, we deduce that

$$\left|\bar{E}_{\mathcal{H}}^{\star}(S)\right| \preceq \sum_{i=1}^{30K} Z_i\,,$$

where Z_i are i.i.d. Geometric random variables with mean $\frac{1}{1-\mu}$. It is well known that the moment-generating function of such variables is given by

$$\mathbb{E}(e^{tZ_1}) = \frac{(1-\mu)e^t}{1-\mu e^t}\,.$$

Setting $t = \varepsilon/2$ and recalling that $\mu = 1-(1+o(1))\varepsilon$, we get that $\mathbb{E}(e^{\frac{\varepsilon}{2}Z_1}) \leq e$ for sufficiently large n (recall that $\varepsilon = o(1)$). Therefore, we obtain that for the above mentioned S,

$$\mathbb{P}\left(|\bar{E}_{\mathcal{H}}^{\star}(S)| \geq (100/\varepsilon)K\right) \leq \frac{\exp(30K)}{\exp(\frac{\varepsilon}{2}(100/\varepsilon)K)} = e^{-20K}\,.$$

By Part (1) of Lemma 3.2, there are at most $(\frac{4}{3}+o(1))\varepsilon^3 n\exp(5K)$ connected sets of size s. Taking a union bound over the $(\frac{4}{3} + o(1))\varepsilon^3 n$ values of s establishes that the statement of the lemma holds except with probability

$$\left(\tfrac{4}{3} + o(1)\right)\varepsilon^3 n\sum_{s} e^{-20K(s)}e^{5K(s)} \leq \left(\tfrac{16}{9} + o(1)\right)(\varepsilon^3 n)^{-13} = o(1)\,,$$

completing the proof of the statement on all connected subsets $S \subset \mathcal{K}$.

Next, if T contains t edges in $\mathcal{K}$, then the number of corresponding edges in $\mathcal{H}$ is again stochastically dominated by a sum of i.i.d. geometric variables $\{Z_i\}$ as above. Hence, by the same argument, the probability that there exists a set $T \subset E(\mathcal{K})$ of $\alpha\varepsilon^3 n$ edges in $\mathcal{K}$, which expands to at least $\beta\varepsilon^2 n$ edges in $\mathcal{H}$ for some $0 < \alpha < \frac{1}{2}$ and $0 < \beta < 1$, is at most

$$\binom{(2 + o(1))\varepsilon^3 n}{\alpha\varepsilon^3 n}\frac{e^{\alpha\varepsilon^3 n}}{e^{(\varepsilon/2)\beta\varepsilon^2 n}} \leq \exp\left[\left(2H\left(\tfrac{\alpha}{2}\right) + \alpha - \tfrac{\beta}{2} + o(1)\right)\varepsilon^3 n\right]$$

(using the well-known fact that $\sum_{i\leq\lambda m}\binom{m}{i} \leq \exp[H(\lambda)m]$ where $H(x)$ is the entropy function $H(x) \triangleq -x\log x - (1-x)\log(1-x)$). It is now easy to verify that a choice of $\alpha = \frac{1}{20}$ and $\beta = \frac{2}{3}$ in the last expression yields a term that tends to 0 as $n \to \infty$.

It remains to bound $|\mathcal{L}|$. This will follow from a bound on the number of loops in $\mathcal{K}$. Let $u \in \mathcal{K}$ be a kernel vertex, and recall that its degree D_u is distributed as an independent $\left(\mathrm{Poisson}(\Lambda)\,\big|\,\cdot \geq 3\right)$, where $\Lambda = (2 + o(1))\varepsilon$ with high probability. The expected number of loops that u obtains in a random realization of the degree sequence (via the configuration model) is clearly at most D_u^2/D, where $D = (4 + o(1))\varepsilon^3 n$ is the total of the kernel degrees. Therefore,

$$\mathbb{E}|\mathcal{L}| \leq (\tfrac{4}{3} + o(1))\varepsilon^3 n \cdot (1/D)\mathbb{E}[D_u^2] = O(1)\,,$$

and so $\mathbb{E}|E_{\mathcal{H}}^{\star}(\mathcal{L})| = O(1/\varepsilon)$. The contribution of $|E_{\mathcal{H}}^{\star}(\mathcal{L})|$ is thus easily absorbed w.h.p. when increasing β from $\frac{2}{3}$ to $\frac{3}{4}$, completing the proof. ∎

Lemma 3.5. *There exists an absolute constant $\iota > 0$ so that w.h.p. every connected set $S \subset \mathcal{H}$ with $(200/\varepsilon)\log(\varepsilon^3 n) \leq d_{\mathcal{H}}(S) \leq e(\mathcal{H})$ satisfies that $|\partial_{\mathcal{H}}S| / d_{\mathcal{H}}(S) \geq \iota\varepsilon$.*

Proof. Let $S \subset \mathcal{H}$ be as above, and write $S_{\mathcal{K}} = S \cap \mathcal{K}$. Observe that $S_{\mathcal{K}}$ is connected (if nonempty). Furthermore, since $d_{\mathcal{H}}(S) \geq (200/\varepsilon)\log(\varepsilon^3 n)$ whereas the longest 2-path in $\mathcal{H}$ contains $(1 + o(1))(1/\varepsilon)\log(\varepsilon^3 n)$ edges w.h.p., we may assume that $S_{\mathcal{K}}$ is indeed nonempty.

Next, clearly $|\partial_{\mathcal{H}}S| \geq |\partial_{\mathcal{K}}S_{\mathcal{K}}|$ (as each edge in the boundary of $S_{\mathcal{K}}$ translates into a 2-path in $\mathcal{H}$ with precisely one endpoint in S), while $|\bar{E}_{\mathcal{H}}(S)| \leq |\bar{E}_{\mathcal{H}}^{\star}(S_{\mathcal{K}})|$ (any $e \in \bar{E}_{\mathcal{H}}(S)$ belongs to some 2-path $\mathcal{P}_e$, which is necessarily incident to some $v \in S_{\mathcal{K}}$ as, crucially, $S_{\mathcal{K}}$ is nonempty. Hence, the edge corresponding to $\mathcal{P}_e$ belongs to $\bar{E}_{\mathcal{K}}(S_{\mathcal{K}})$, and so $e \in \bar{E}_{\mathcal{H}}^{\star}(S_{\mathcal{K}})$). Therefore, using the fact that $d_{\mathcal{H}}(S) \leq 2|\bar{E}_{\mathcal{H}}(S)|$,

$$\frac{|\partial_{\mathcal{H}}S|}{d_{\mathcal{H}}(S)} \geq \frac{|\partial_{\mathcal{K}}S_{\mathcal{K}}|}{2|\bar{E}_{\mathcal{H}}^{\star}(S_{\mathcal{K}})|} = \frac{|\partial_{\mathcal{K}}S_{\mathcal{K}}|}{2|\bar{E}_{\mathcal{K}}(S_{\mathcal{K}})|} \cdot \frac{|\bar{E}_{\mathcal{K}}(S_{\mathcal{K}})|}{|\bar{E}_{\mathcal{H}}^{\star}(S_{\mathcal{K}})|} . \tag{3.5}$$

Assume that the events stated in Lemma 3.4 hold. Since the assumption on $d_{\mathcal{H}}(S_{\mathcal{K}})$ gives that $|\bar{E}_{\mathcal{H}}^{\star}(S_{\mathcal{K}})| \geq (100/\varepsilon)\log(\varepsilon^3 n)$, we deduce that necessarily

$$|S_{\mathcal{K}}| \geq (\varepsilon/100)|\bar{E}_{\mathcal{H}}^{\star}(S_{\mathcal{K}})| ,$$

and thus (since $S_{\mathcal{K}}$ is connected)

$$|\bar{E}_{\mathcal{K}}(S_{\mathcal{K}})| \geq |E_{\mathcal{K}}(S_{\mathcal{K}})| \geq (\varepsilon/100)|\bar{E}_{\mathcal{H}}^{\star}(S_{\mathcal{K}})| - 1 . \tag{3.6}$$

Now,

$$d_{\mathcal{H}}(S) \leq e(\mathcal{H}) = (2 + o(1))\varepsilon^2 n ,$$

and since $d_{\mathcal{H}}(S) = 2|E_{\mathcal{H}}(S)| + |\partial_{\mathcal{H}}S|$ we have $|E_{\mathcal{H}}(S)| \leq (1 + o(1))\varepsilon^2 n$. In particular, $|E(\mathcal{H}) \setminus E_{\mathcal{H}}(S)| \geq \frac{3}{4}\varepsilon^2 n$ for sufficiently large n.

At the same time, if $\mathcal{L}$ is the set of all loops in $\mathcal{K}$ and $T = \bar{E}_{\mathcal{K}}(\mathcal{K} \setminus S_{\mathcal{K}})$, then clearly $E_{\mathcal{H}}^{\star}(T) \cup E_{\mathcal{H}}^{\star}(\mathcal{L})$ is a superset of $E(\mathcal{H}) \setminus E_{\mathcal{H}}(S)$. Therefore, Lemma 3.4 yields that $|T| \geq \frac{1}{20}\varepsilon^3 n$. Since $d_{\mathcal{K}}(S_{\mathcal{K}}) \leq 2e(\mathcal{K}) = (4 + o(1))\varepsilon^3 n$, we get

$$d_{\mathcal{K}}(\mathcal{K} \setminus S_{\mathcal{K}}) \geq |T| \geq \frac{\varepsilon^3 n}{20} \geq \frac{1 + o(1)}{80}d_{\mathcal{K}}(S_{\mathcal{K}}) .$$

At this point, by Lemma 2.2 there exists $\alpha > 0$ such that w.h.p. for any such above mentioned subset S:

$$|\partial_{\mathcal{K}}S_{\mathcal{K}}| \geq \alpha\left(d_{\mathcal{K}}(S_{\mathcal{K}}) \wedge d_{\mathcal{K}}(\mathcal{K} \setminus S_{\mathcal{K}})\right) \geq \frac{\alpha + o(1)}{80}d_{\mathcal{K}}(S_{\mathcal{K}}) . \tag{3.7}$$

Plugging (3.6),(3.7) into (3.5), we conclude that the lemma holds for any sufficiently large n with, say, $\iota = \frac{1}{2} \cdot 10^{-4}\alpha$. ∎

We are now ready to establish the upper bound on the mixing time for the random walk on $\mathcal{H}$.

Proof of Theorem 3.1. We will apply the following recent result of [11], which bounds the mixing time of a lazy chain in terms of its isoperimetric profile (a fine-tuned version of the Lovász-Kannan [18] bound on the mixing time in terms of the average conductance).

Theorem 3.6 ([11])**.** *Let $P = (p_{x,y})$ be the transition kernel of an irreducible, reversible and aperiodic Markov chain on Ω with stationary distribution π. Let $\pi_{\min} = \min_{x \in \Omega} \pi(x)$ and for $p > \pi_{\min}$, let*

$$\Phi(p) \triangleq \min\{\Phi(S) : S \text{ is connected and } p/2 \leq \pi(S) \leq p\},$$

and $\Phi(p) = 1$ if there is no such S. Then for some absolute constant $C > 0$,

$$\tilde{t}_{\mathrm{MIX}} \leq C \sum_{j=1}^{\lceil \log \pi_{\min}^{-1} \rceil} \Phi^{-2}(2^{-j}).$$

In our case, the P is the transition kernel of the lazy random walk on $\mathcal{H}$. By definition, if $S \subset \mathcal{H}$ and $d_{\mathcal{H}}(x)$ denotes the degree of $x \in \mathcal{H}$, then

$$\pi_{\mathcal{H}}(x) = \frac{d_{\mathcal{H}}(x)}{2e(\mathcal{H})}, \quad p_{x,y} = \frac{1}{2d_{\mathcal{H}}(x)}, \quad \pi_{\mathcal{H}}(S) = \frac{d_{\mathcal{H}}(S)}{2e(\mathcal{H})},$$

and so $\Phi(S) = \frac{1}{2}|\partial_{\mathcal{H}}S|/d_{\mathcal{H}}(S)$. Recall that w.h.p. $e(\mathcal{H}) = (2 + o(1))\varepsilon^2 n$. Under this assumption, for any $p \geq 120\frac{\log(\varepsilon^3 n)}{\varepsilon^3 n}$ and connected subset $S \subset \mathcal{H}$ satisfying $\pi_{\mathcal{H}}(S) \geq p/2$,

$$d_{\mathcal{H}}(S) = 2\pi_{\mathcal{H}}(S)e(\mathcal{H}) \geq (200/\varepsilon)\log(\varepsilon^3 n).$$

Therefore, by Lemma 3.5, w.h.p.

$$\Phi(p) \geq \tfrac{1}{2}\iota\varepsilon \ \text{ for all } \ 120\frac{\log(\varepsilon^3 n)}{\varepsilon^3 n} \leq p \leq \tfrac{1}{2} \ . \tag{3.8}$$

Set

$$j^* = \max\left\{j : 2^{-j} \geq 120\frac{\log(\varepsilon^3 n)}{\varepsilon^3 n}\right\} \ .$$

It is clear that $j^* = O(\log(\varepsilon^3 n))$ and (3.8) can be translated into

$$\Phi(2^{-j}) \geq \tfrac{1}{2}\iota\varepsilon, \text{ for all } 1 \leq j \leq j^* \ . \tag{3.9}$$

On the other hand, if $\pi_{\mathcal{H}}(S) \leq p < 1$ then $d_{\mathcal{H}}(S) \leq 2pe(\mathcal{H})$ while $|\partial_{\mathcal{H}}S| \geq 1$ (as $\mathcal{H}$ is connected), and so for $j^* < j \leq \lceil \log \pi_{\min}^{-1} \rceil$ we have

$$\Phi(2^{-j}) \geq \frac{2^{j-2}}{e(\mathcal{H})} \geq \frac{2^j}{10\varepsilon^2 n} \tag{3.10}$$

(where the last inequality holds for large n). Combining (3.9) and (3.10) together, we now apply Theorem 3.6 to conclude that there exists a constant $C > 0$ such that, w.h.p.,

$$\tilde{t}_{\text{MIX}} \leq C \sum_{j=1}^{\lceil \log \pi_{\min}^{-1} \rceil} \frac{1}{\Phi^2(2^{-j})} = C \left[\sum_{j=1}^{j^*} \frac{1}{\Phi^2(2^{-j})} + \sum_{j=j^*}^{\lceil \log \pi_{\min}^{-1} \rceil} \frac{1}{\Phi^2(2^{-j})} \right]$$

$$\leq C \left(j^*(\tfrac{1}{2}\iota\varepsilon)^{-2} + 2(10\varepsilon^2 n \cdot 2^{-j^*})^2 \right) = O(\varepsilon^{-2} \log^2(\varepsilon^3 n)) \,,$$

where the last inequality follows by our choice of j^*.

The lower bound on the mixing time follows immediately from the fact that, by the definition of $\tilde{\mathcal{C}}_1$, w.h.p. there exists a 2-path in $\mathcal{H}$ whose length is $(1 - o(1))(1/\varepsilon) \log(\varepsilon^3 n)$ (see [8, Corollary 1]). ∎

3.2. **Local times for the random walk on the 2-core.** In order to extend the mixing time from the 2-core $\mathcal{H}$ to the giant component, we need to prove the following proposition.

Proposition 3.7. *Let $N_{v,s}$ be the local time induced by the lazy random walk (W_t) on $\mathcal{H}$ to the vertex v up to time s, i.e., $\#\{0 \leq t \leq s : W_t = v\}$. Then there exists some $C > 0$ such that, w.h.p., for all $s > 0$ and any $u, v \in \mathcal{H}$,*

$$\mathbb{E}_u[N_{v,s}] \leq C \frac{\varepsilon s}{\log(\varepsilon^3 n)} + (150/\varepsilon) \log(\varepsilon^3 n) \,.$$

In order to prove Proposition 3.7, we wish to show that with positive probability the random walk W_t will take an excursion in a long 2-path before returning to v. Consider some $v \in \mathcal{K}$ (we will later extend this analysis to the vertices in $\mathcal{H} \setminus \mathcal{K}$, i.e., those vertices lying on 2-paths). We point out that proving this statement is simpler in case $D_v = O(1)$, and most of the technical challenge lies in the possibility that D_v is unbounded. In order to treat this point, we first show that the neighbors of vertex v in the kernel are, in some sense, distant apart.

Lemma 3.8. *For $v \in \mathcal{K}$ let $\mathcal{N}_v$ denote the set of neighbors of v in the kernel $\mathcal{K}$. Then w.h.p., for every $v \in \mathcal{K}$ there exists a collection of disjoint connected subsets $\{B_w(v) \subset \mathcal{K} : w \in \mathcal{N}_v\}$, such that for all $w \in \mathcal{N}_v$,*

$$|B_w| = \lceil (\varepsilon^3 n)^{1/5} \rceil \quad \text{and} \quad \text{diam}(B_w) \leq \tfrac{1}{2} \log(\varepsilon^3 n) \,.$$

Proof. We may again assume (3.1) and furthermore, that

$$3 \leq D_v \leq \log(\varepsilon^3 n) \text{ for all } v \in \mathcal{K} \,.$$

Let $v \in \mathcal{K}$. We construct the connected sets B_w while we reveal the structure of the kernel $\mathcal{K}$ via the configuration model, as follows: Process the vertices $w \in \mathcal{N}_v$ sequentially according to some arbitrary order. When processing such a vertex w, we expose the ball (according to the graph metric) about

it, excluding v and any vertices that were already accounted for, until its size reaches $\lceil (\varepsilon^3 n)^{1/5} \rceil$ (or until no additional new vertices can be added).

It is clear from the definition that the B_w's are indeed disjoint and connected, and it remains to prove that each B_w satisfies $|B_w| = \lceil (\varepsilon^3 n)^{1/5} \rceil$ and $\mathrm{diam}(B_w) \leq \log(\varepsilon^3 n)$.

Let R denote the tree-excess of the (connected) subset $\{v\} \cup \bigcup_w B_w$ once the process is concluded. We claim that w.h.p. $R \leq 1$. To see this, first observe that at any point in the above process, the sum of degrees of all the vertices that were already exposed (including v and $\mathcal{N}_v$) is at most

$$\lceil (\varepsilon^3 n)^{1/5} \rceil \log^2(\varepsilon^3 n) = (\varepsilon^3 n)^{1/5 + o(1)}.$$

Hence, by the definition of the configuration model (which draws a new half-edge between w and some other vertex proportional to its degree), $R \preceq Z$ where Z is a binomial variable $\mathrm{Bin}\left((\varepsilon^3 n)^{1/5 + o(1)}, (\varepsilon^3 n)^{-4/5 + o(1)} \right)$. This gives

$$\mathbb{P}(R \geq 2) = (\varepsilon^3 n)^{-6/5 + o(1)}.$$

In particular, since $D_w \geq 3$ for any $w \in \mathcal{K}$, this implies that we never fail to grow B_w to size $(\varepsilon^3 n)^{1/5}$, and that the diameter of each B_w is at most that of a binary tree (possibly plus $R \leq 1$), i.e., for any large n,

$$\mathrm{diam}(B_w) \leq \tfrac{1}{5} \log_2(\varepsilon^3 n) + 2 \leq \tfrac{1}{2} \log(\varepsilon^3 n).$$

A simple union bound over $v \in \mathcal{K}$ now completes the proof. ∎

We distinguish the following subset of the edges of the kernel, whose paths are suitably long:

$$\mathcal{E} \triangleq \left\{ e \in E(\mathcal{K}) : |\mathcal{P}_e| \geq \frac{1}{20\varepsilon} \log(\varepsilon^3 n) \right\},$$

where $\mathcal{P}_e$ is the 2-path in $\mathcal{H}$ that corresponds to the edge $e \in E(\mathcal{K})$. Further define $\mathcal{Q} \subset 2^{\mathcal{K}}$ to be all the subsets of vertices of $\mathcal{K}$ whose induced subgraph contains an edge from $\mathcal{E}$:

$$\mathcal{Q} \triangleq \{ S \subset \mathcal{K} : E_{\mathcal{K}}(S) \cap \mathcal{E} \neq \emptyset \}.$$

For each $e \in \mathcal{K}$, we define the median of its 2-path, denoted by $\mathrm{med}(\mathcal{P}_e)$, in the obvious manner: It is the vertex $w \in \mathcal{P}_e$ whose distance from the two endpoints is the same, up to at most 1 (whenever there are two choices for this w, pick one arbitrarily). Now, for each $v \in \mathcal{H}$ let

$$\mathcal{E}_v \triangleq \{ \mathrm{med}(\mathcal{P}_e) \ : \ e \in \mathcal{E} , \ v \notin \mathcal{P}_e \}.$$

The next lemma provides a lower bound on the effective conductance between a vertex v in the 2-core and its corresponding above defined set $\mathcal{E}_v$. See, e.g., [23] for further details on conductances/resistances.

Lemma 3.9. *Let $C_{\mathrm{eff}}(v \leftrightarrow \mathcal{E}_v)$ be the effective conductance between a vertex $v \in \mathcal{H}$ and the set $\mathcal{E}_v$. With high probability, for any $v \in \mathcal{H}$,*

$$C_{\mathrm{eff}}(v \leftrightarrow \mathcal{E}_v)/D_v \geq \varepsilon/(100 \log(\varepsilon^3 n)) \ .$$

Proof. In order to bound the effective conductance, we need to prove that for any $v \in \mathcal{K}$, there exist D_v disjoint paths of length at most $(100/\varepsilon) \log(\varepsilon^3 n)$ leading to the set $\mathcal{E}_v$. By Lemmas 3.4 and 3.8, it suffices to prove that w.h.p. for any $v \in \mathcal{K}$ and $w \in \mathcal{N}_v$, we have that $E(B_w) \cap \mathcal{E} \neq \emptyset$, where $\mathcal{N}_v$ and B_w are defined as in Lemma 3.8 (in this case, the path from v to some $e \in \mathcal{E}$ within B_w will have length at most $\frac{1}{2} \log(\varepsilon^3 n)$ in $\mathcal{K}$, and its length will not be exceed $(100/\varepsilon) \log(\varepsilon^3 n)$ after being expanded in the 2-core).

Notice that if Y is the geometric variable $\mathrm{Geom}(1 - \mu)$ then

$$\mathbb{P}\big(Y \geq \tfrac{1}{10\varepsilon} \log(\varepsilon^3 n)\big) = \mu^{\frac{1}{10\varepsilon} \log(\varepsilon^3 n)} \geq (\varepsilon^3 n)^{-1/10 + o(1)} \ .$$

Therefore, by the independence of the lengths of the 2-paths and the fact that $|B_w| = \lceil (\varepsilon^3 n)^{1/5} \rceil$, we obtain that

$$\mathbb{P}\left(E(B_w) \cap \mathcal{E} = \emptyset\right) \leq \left(1 - (\varepsilon^3 n)^{-1/10 + o(1)}\right)^{(\varepsilon^3 n)^{1/5}} \leq \mathrm{e}^{-(\varepsilon^3 n)^{1/10 - o(1)}} \ .$$

At this point, a union bound shows that the probability that for some $v \in \mathcal{K}$ there exists some $w \in \mathcal{N}_v$, such that $E(B_w)$ does not intersect $\mathcal{E}$, is at most

$$\left(\tfrac{4}{3} + o(1)\right) \varepsilon^3 n \cdot \log(\varepsilon^3 n) \cdot \mathrm{e}^{-(\varepsilon^3 n)^{1/10 - o(1)}} = o(1) \, . \qquad \blacksquare$$

We are ready to prove the main result of this subsection, Proposition 3.7, which bounds the local times induced by the random walk on the 2-core.

Proof of Proposition 3.7. For some vertex $v \in \mathcal{H}$ and subset $A \subset \mathcal{H}$, let

$$\tau_v^+ \triangleq \min\{t > 0 : W_t = v\}, \quad \tau_A \triangleq \min\{t : W_t \in A\} \, .$$

It is well-known (see, e.g., [23, equation (2.4)]) that the effective conductance has the following form:

$$\mathbb{P}_v(\tau_A < \tau_v^+) = \frac{C_{\mathrm{eff}}(v \leftrightarrow A)}{D_v} \, .$$

Combined with Lemma 3.9, it follows that

$$\mathbb{P}_v(\tau_{\mathcal{E}_v} < \tau_v^+) = \frac{C_{\mathrm{eff}}(v \leftrightarrow \mathcal{E}_v)}{D_v} \geq \varepsilon / \left(100 \log(\varepsilon^3 n)\right) \ .$$

On the other hand, for any $v \in \mathcal{H}$, by definition $w \in \mathcal{E}_v$ is the median of some 2-path, which does not contain v and has length at least $\frac{1}{20\varepsilon} \log(\varepsilon^3 n)$. Hence, by well-known properties of hitting times for the simple random walk on the integers, there exists some absolute constant $c > 0$ such that for any $v \in \mathcal{H}$ and $w \in \mathcal{E}_v$:

$$\mathbb{P}_w\left(\tau_v^+ \geq c\varepsilon^{-2} \log^2(\varepsilon^3 n)\right) \geq \mathbb{P}_w\left(\tau_{\mathcal{K}} \geq c\varepsilon^{-2} \log^2(\varepsilon^3 n)\right) \geq \tfrac{2}{3} \, ,$$

Altogether, we conclude that

$$\mathbb{P}_v\left(\tau_v^+ \geq c\varepsilon^{-2}\log^2(\varepsilon^3 n)\right) \geq \mathbb{P}_v\left(\tau_{\mathcal{E}_v} < \tau_v^+\right) \min_{w\in\mathcal{E}_v}\left\{\mathbb{P}_w(\tau_v^+ \geq c\varepsilon^{-2}\log^2(\varepsilon^3 n))\right\}$$
$$\geq \varepsilon/\left(150\log(\varepsilon^3 n)\right).$$

Setting $t_c = c\varepsilon^{-2}\log^2(\varepsilon^3 n)$, we can rewrite the above as

$$\mathbb{P}_v(N_{v,t_c} \geq 2) \leq 1 - \varepsilon/\left(150\log(\varepsilon^3 n)\right).$$

By the strong Markovian property (i.e., $(W_{\tau_v^+ + t})$ is a Markov chain with the same transition kernel of (W_t)), we deduce that

$$\mathbb{P}(N_{v,t_c} \geq k) \leq \left[1 - \varepsilon/\left(150\log(\varepsilon^3 n)\right)\right]^{k-1},$$

and hence

$$\mathbb{E}N_{v,t_c} \leq (150/\varepsilon)\log(\varepsilon^3 n).$$

The proof is completed by observing that $\mathbb{E}_v(N_{v,s}) \leq \lceil s/t_c\rceil\mathbb{E}_v N_{v,t_c}$ and that $\mathbb{E}_u N_{v,s} \leq \mathbb{E}_v N_{v,s}$ for any u. ∎

4. Mixing on the giant component

In this section, we prove Theorem 1, which establishes the order of the mixing time of the lazy random walk on the supercritical $\mathcal{C}_1$.

4.1. Controlling the attached Poisson Galton-Watson trees.
So far, we have established that w.h.p. the mixing time of the lazy random walk on the 2-core $\tilde{\mathcal{C}}_1^{(2)}$ has order $\varepsilon^{-2}\log^2(\varepsilon^3 n)$. To derive the mixing time for $\tilde{\mathcal{C}}_1$ based on that estimate, we need to consider the delays due to the excursions the random walk makes in the attached trees. As we will later see, these delays will be upper bounded by a certain a linear combination of the sizes of the trees (with weights determined by the random walk on the 2-core). The following lemma will play a role in estimating this expression.

Lemma 4.1. *Let $\{\mathcal{T}_i\}$ be independent $\mathrm{PGW}(\mu)$-trees. For any two constants $C_1, C_2 > 0$ there exists some constant $C > 0$ such that the following holds: If $\{a_i\}_{i=1}^m$ is a sequence of positive reals satisfying*

$$\sum_{i=1}^m a_i \leq C_1\varepsilon^{-2}\log^2(\varepsilon^3 n), \tag{4.1}$$

$$\max_{1\leq i\leq m} a_i \leq C_2\varepsilon^{-1}\log(\varepsilon^3 n), \tag{4.2}$$

then

$$\mathbb{P}\Big(\sum_{i=1}^m a_i|\mathcal{T}_i| \geq C\varepsilon^{-3}\log^2(\varepsilon^3 n)\Big) \leq (\varepsilon^3 n)^{-2}.$$

Proof. It is well-known (see, e.g., [28]) that the size of a Poisson(γ)-Galton-Watson tree $\mathcal{T}$ follows a Borel(γ) distribution, namely,

$$\mathbb{P}(|\mathcal{T}| = k) = \frac{k^{k-1}}{\gamma k!}(\gamma e^{-\gamma})^k . \tag{4.3}$$

The following is a well-known (and easy) estimate on the size of a PGW-tree; we include its proof for completeness.

Claim 4.2. *Let $0 < \gamma < 1$, and let $\mathcal{T}$ be a PGW(γ)-tree. Then*

$$\mathbb{E}|\mathcal{T}| = \frac{1}{1 - \gamma} , \qquad \mathrm{Var}(|\mathcal{T}|) = \frac{\gamma}{(1 - \gamma)^3} .$$

Proof. For $k = 0, 1, \ldots$, let L_k be the number of vertices in the k-th level of the tree $\mathcal{T}$. Clearly, $\mathbb{E}L_k = \gamma^k$, and so $\mathbb{E}|\mathcal{T}| = \mathbb{E}\sum_k L_k = \frac{1}{1-\gamma}$.

By the total-variance formula,

$$\mathrm{Var}(L_i) = \mathrm{Var}\left(\mathbb{E}\left(L_i \mid L_{i-1}\right)\right) + \mathbb{E}\left(\mathrm{Var}\left(L_i \mid L_{i-1}\right)\right)$$
$$= \gamma^2 \mathrm{Var}(L_{i-1}) + \gamma\mathbb{E}L_{i-1} = \gamma^2 \mathrm{Var}(L_{i-1}) + \gamma^i .$$

By induction,

$$\mathrm{Var}(L_i) = \sum_{k=i}^{2i-1} \gamma^k = \gamma^i \frac{1 - \gamma^i}{1 - \gamma} . \tag{4.4}$$

We next turn to the covariance of L_i, L_j for $i \leq j$:

$$\mathrm{Cov}(L_i, L_j) = \mathbb{E}[L_i L_j] - \mathbb{E}L_i \mathbb{E}L_j = \gamma^{j-i}\mathbb{E}L_i^2 - \gamma^{i+j}$$
$$= \gamma^{j-i} \mathrm{Var}(L_i) = \gamma^j \frac{1 - \gamma^i}{1 - \gamma} .$$

Summing over the variances and covariances of the L_i's, we deduce that

$$\mathrm{Var}(|\mathcal{T}|) = 2\sum_{i=0}^{\infty}\sum_{j=i}^{\infty} \gamma^j \frac{1 - \gamma^i}{1 - \gamma} - \sum_{i=0}^{\infty} \gamma^i \frac{1 - \gamma^i}{1 - \gamma} = \frac{\gamma}{(1 - \gamma)^3} . \qquad \blacksquare$$

We need the next lemma to bound the tail probability for $\sum a_i |\mathcal{T}_i|$.

Lemma 4.3 ([15, Corollary 4.2]). *Let $X_1, \ldots, X_m$ be independent r.v.'s with $\mathbb{E}[X_i] = \mu_i$. Suppose there are b_i, d_i and ξ_0 such that $\mathrm{Var}(X_i) \leq b_i$, and*

$$\left|\mathbb{E}\left[(X_i - \mu_i)^3 e^{\xi(X_i - \mu_i)}\right]\right| \leq d_i \quad \text{for all } 0 \leq |\xi| \leq \xi_0 .$$

If $\delta\xi_0 \sum_{i=1}^m d_i \leq \sum_{i=1}^m b_i$ for some $0 < \delta \leq 1$, then for all $\Delta > 0$,

$$\mathbb{P}\left(\left|\sum_{i=1}^m X_i - \sum_{i=1}^m \mu_i\right| \geq \Delta\right) \leq \exp\left(-\frac{1}{3}\min\left\{\delta\xi_0\Delta, \frac{\Delta^2}{\sum_{i=1}^m b_i}\right\}\right) .$$

Let $T_i = |\mathcal{T}_i|$ and $X_i = a_i T_i$ for $i \in [m]$. Claim 4.2 gives that

$$\mu_i = \mathbb{E} X_i = a_i/(1-\mu).$$

Now set

$$\xi_0 = \varepsilon^3/(10 C_2 \log(\varepsilon^3 n)).$$

For any $|\xi| \leq \xi_0$, we have $a_i|\xi| \leq \varepsilon^2/10$ by the assumption (4.2), and so

$$\left| \mathbb{E} \left[(X_i - \mu_i)^3 e^{\xi(X_i - \mu_i)} \right] \right| = a_i^3 \left| \mathbb{E} \left[\left(T_i - \frac{1}{1-\mu} \right)^3 e^{\xi a_i (T_i - \frac{1}{1-\mu})} \right] \right|$$

$$\leq a_i^3 \mathbb{E} \left[(1-\mu)^{-3} \mathbb{1}_{\{T_i < (1-\mu)^{-1}\}} \right] + a_i^3 \mathbb{E} \left[T_i^3 e^{\xi a_i T_i} \mathbb{1}_{\{T_i \geq (1-\mu)^{-1}\}} \right]$$

$$\leq a_i^3 (1-\mu)^{-3} + a_i^3 \mathbb{E} \left[T_i^3 \exp(\varepsilon^2 T_i/10) \right]. \tag{4.5}$$

Recalling the law of T_i given by (4.3), we obtain that

$$\mathbb{E} \left(T_i^3 \exp(\varepsilon^2 T_i/10) \right) = \sum_{k=1}^{\infty} \frac{k^{k-1}}{\mu k!} (\mu e^{-\mu})^k k^3 e^{\varepsilon^2 k/10}.$$

Using Stirling's formula, we obtain that for some absolute constant $c > 1$,

$$\mathbb{E} \left(T_i^3 \exp(\varepsilon^2 T_i/10) \right) \leq c \sum_{k=1}^{\infty} \frac{k^{k-1}(\mu e^{-\mu})^k}{\mu (k/e)^k \sqrt{k}} k^3 e^{\varepsilon^2 k/10}$$

$$= \frac{c}{\mu} \sum_{k=1}^{\infty} k^{3/2} (\mu e^{1-\mu})^k e^{\varepsilon^2 k/10}. \tag{4.6}$$

Recalling that $\mu = 1 - \varepsilon + \frac{2}{3}\varepsilon^2 + O(\varepsilon^3)$ and using the fact that $1-x \leq e^{-x-x^2/2}$ for $x \geq 0$, we get that for sufficiently large n (and hence small enough ε),

$$\mu e^{1-\mu} = (1 - (1-\mu)) e^{1-\mu} \leq \exp \left(-\tfrac{1}{2}\varepsilon^2 + O(\varepsilon^3) \right) \leq e^{-\varepsilon^2/3}. \tag{4.7}$$

Plugging the above estimate into (4.6), we obtain that for large n,

$$\mathbb{E} \left[T_i^3 \exp(\varepsilon^2 T_i/10) \right] \leq 2c \sum_{k=1}^{\infty} k^{3/2} e^{-\varepsilon^2 k/6} \leq 4c \int_0^{\infty} x^{3/2} e^{-\varepsilon^2 x/6} \, dx$$

$$\leq 400 c \varepsilon^{-5} \int_0^{\infty} x^{3/2} e^{-x} \, dx = 300 \sqrt{\pi} c \varepsilon^{-5}.$$

Going back to (4.5), we get that for some absolute $c' > 1$ and any large n,

$$\left| \mathbb{E} \left[(X_i - \mu_i)^3 e^{\xi(X_i - \mu_i)} \right] \right| \leq a_i^3 \left(2\varepsilon^{-3} + c'\varepsilon^{-5} \right) \leq a_i \cdot 2c' C_2^2 \varepsilon^{-7} \log^2(\varepsilon^3 n) \stackrel{\triangle}{=} d_i,$$

where the second inequality used (4.2).

By Claim 4.2, it follows that for large enough n,

$$\mathrm{Var}(X_i) = a_i^2 \, \mathrm{Var}(T_i) = a_i^2 \frac{\mu}{(1-\mu)^3} \leq 2a_i^2 \varepsilon^{-3} \leq a_i \cdot 2C_2 \varepsilon^{-4} \log(\varepsilon^3 n) \stackrel{\triangle}{=} b_i.$$

Since $\sum_i d_i = \left(c'C_2\varepsilon^{-3}\log(\varepsilon^3 n)\right)\sum_i b_i$, by setting $\delta = 1$ (and recalling our choice of ξ_0) we get

$$\delta\xi_0 \sum_{i=1}^m d_i = \frac{\delta c'}{10}\sum_i b_i \leq \sum_{i=1}^m b_i\,.$$

We have thus established the conditions for Lemma 4.3, and it remains to select Δ. For a choice of $\Delta = (60C_2 \vee \sqrt{12C_1 C_2})\varepsilon^{-3}\log^2(\varepsilon^3 n)$, by definition of ξ_0 and the b_i's we have

$$\xi_0\Delta \geq 6\log(\varepsilon^3 n)\,,$$
$$\Delta^2/\textstyle\sum_i b_i \geq 6C_1\varepsilon^{-2}\log^3(\varepsilon^3 n)/\textstyle\sum_i a_i \geq 6\log(\varepsilon^3 n)\,,$$

where the last inequality relied on (4.1). Hence, an application of Lemma 4.3 gives that for large enough n,

$$\mathbb{P}\Big(\textstyle\sum_i a_i T_i - \sum_i \mu_i \geq \Delta\Big) \leq (\varepsilon^3 n)^{-2}\,.$$

Finally, by (4.1) and using the fact that $1 - \mu \geq \varepsilon/2$ for any large n, we have $\sum_i \mu_i = (1-\mu)^{-1}\sum_i a_i \leq 2C_1\varepsilon^{-3}\log^2(\varepsilon^3 n)$. The proof of Lemma 4.1 is thus concluded by choosing $C = 2C_1 + (60C_2 \vee \sqrt{12C_1 C_2})$. ∎

To bound the time it takes the random walk to exit from an attached PGW-tree (and enter the 2-core), we will need to control the diameter and volume of such a tree. The following simple lemma of [9] gives an estimate on the diameter of a PGW-tree:

Lemma 4.4 ([9, Lemma 3.2]). *Let $\mathcal{T}$ be a $\mathrm{PGW}(\mu)$-tree and L_k be its k-th level of vertices. Then $\mathbb{P}(L_k \neq \emptyset) = \Theta\left(\varepsilon e^{-k(\varepsilon + O(\varepsilon^2))}\right)$ for any $k \geq 1/\varepsilon$.*

The next lemma gives a bound on the volume of a PGW-tree:

Lemma 4.5. *Let $\mathcal{T}$ be a $\mathrm{PGW}(\mu)$-tree. Then*

$$\mathbb{P}(|\mathcal{T}| \geq 6\varepsilon^{-2}\log(\varepsilon^3 n)) = o(\varepsilon(\varepsilon^3 n)^{-2})\,.$$

Proof. Recalling (4.3) and applying Stirling's formula, we obtain that for any $s > 0$,

$$\mathbb{P}(|\mathcal{T}| \geq s) = \sum_{k \geq s} \frac{k^{k-1}}{\mu k!}(\mu e^{-\mu})^k = \Theta\left(\sum_{k \geq s} \frac{(\mu e^{1-\mu})^k}{k^{3/2}}\right). \tag{4.8}$$

Write $r = \log(\varepsilon^3 n)$. By estimate (4.7), we now get that for large enough n,

$$\sum_{k \geq 6\varepsilon^{-2}r} k^{-3/2}(\mu e^{1-\mu})^k \leq \sum_{k \geq 6\varepsilon^{-2}r} k^{-3/2}e^{-\varepsilon^2 k/3} = O(e^{-2r}\varepsilon/\sqrt{r})\,,$$

and combined with (4.8) this concludes the proof. ∎

Finally, for the lower bound, we will need to show that w.h.p. one of the attached PGW-trees in $\tilde{\mathcal{C}}_1$ is suitably large, as we next describe. For a rooted tree $\mathcal{T}$, let L_k be its k-th level of vertices and $\mathcal{T}_v$ be its entire subtree rooted at v. Define the event

$$A_{r,s}(\mathcal{T}) \triangleq \{\exists v \in L_r \text{ such that } |\mathcal{T}_v| \geq s \}.$$

The next lemma gives a bound on the probability of this event when $\mathcal{T}$ is a PGW(μ)-tree.

Lemma 4.6. *Let $\mathcal{T}$ be a PGW(μ)-tree and take $r = \lceil \frac{1}{8}\varepsilon^{-1}\log(\varepsilon^3 n) \rceil$ and $s = \frac{1}{8}\varepsilon^{-2}\log(\varepsilon^3 n)$. Then for any sufficiently large n,*

$$\mathbb{P}(A_{r,s}(\mathcal{T})) \geq \varepsilon(\varepsilon^3 n)^{-2/3}.$$

Proof. We first give a lower bound on the probability that $|\mathcal{T}| \geq s$. By (4.8), we have $\mathbb{P}(|\mathcal{T}| \geq s) \geq c\sum_{k \geq s} k^{-3/2}(\mu e^{1-\mu})^k$ for some absolute $c > 0$. Recalling that $\mu = 1 - \varepsilon + \frac{2}{3}\varepsilon^2 + O(\varepsilon^3)$, we have that for n large enough,

$$\mu e^{1-\mu} \geq e^{-(\varepsilon+\varepsilon^2)}e^{\varepsilon-\varepsilon^2} \geq e^{-2\varepsilon^2}.$$

Therefore, for $s = \frac{1}{8}\varepsilon^{-2}\log(\varepsilon^3 n)$ this gives that

$$\mathbb{P}(|\mathcal{T}| \geq s) \geq c\sum_{s \leq k \leq 2s} k^{-3/2}e^{-2\varepsilon^2 k} \geq cs(2s)^{-3/2}e^{-4\varepsilon^2 s} \geq \varepsilon(\varepsilon^3 n)^{-1/2+o(1)}.$$

Combining this with the fact that $\{\mathcal{T}_v : v \in L_r\}$ are i.i.d. PGW(μ)-trees given L_r, we get

$$\mathbb{P}\left(A_{r,s}(\mathcal{T}) \mid L_r\right) = 1 - (1 - \mathbb{P}(|\mathcal{T}| \geq s))^{|L_r|} \geq 1 - (1 - \varepsilon(\varepsilon^3 n)^{-1/2+o(1)})^{|L_r|}.$$

Taking expectation over L_r, we conclude that

$$\begin{aligned}
\mathbb{P}\left(A_{r,s}(\mathcal{T})\right) &\geq 1 - \mathbb{E}\left((1 - \varepsilon(\varepsilon^3 n)^{-1/2+o(1)})^{|L_r|}\right) \\
&\geq \varepsilon(\varepsilon^3 n)^{-1/2+o(1)}\mathbb{E}|L_r| - \varepsilon^2(\varepsilon^3 n)^{-1+o(1)}\mathbb{E}|L_r|^2 .
\end{aligned} \tag{4.9}$$

For $r = \lceil \frac{1}{8}\varepsilon^{-1}\log(\varepsilon^3 n) \rceil$ we have

$$\mathbb{E}(|L_r|) = \mu^r \geq e^{-(\varepsilon+O(\varepsilon^2))r} \geq (\varepsilon^3 n)^{-1/8+o(1)},$$

and by (4.4),

$$\text{Var}\,|L_r| = \mu^r \frac{1-\mu^r}{1-\mu} \leq e^{-\varepsilon r}2\varepsilon^{-1} \leq 2\varepsilon^{-1}(\varepsilon^3 n)^{-1/8}.$$

Plugging these estimates into (4.9), we obtain that

$$\mathbb{P}\left(A_{r,s}(\mathcal{T})\right) \geq \varepsilon(\varepsilon^3 n)^{-5/8+o(1)} \geq \varepsilon(\varepsilon^3 n)^{-2/3},$$

where the last inequality holds for large enough n, as required. $\blacksquare$

4.2. **Proof of Theorem 1: Upper bound on the mixing time.** By Theorem 2.1, it suffices to consider $\tilde{\mathcal{C}}_1$ instead of $\mathcal{C}_1$. As in the previous section, we abbreviate $\tilde{\mathcal{C}}_1^{(2)}$ by $\mathcal{H}$.

For each vertex v in the 2-core $\mathcal{H}$, let $\mathcal{T}_v$ be the PGW-tree attached to v in $\tilde{\mathcal{C}}_1$. Let (S_t) be the lazy random walk on $\tilde{\mathcal{C}}_1$, define $\tau_0 = 0$ and for $j \geq 0$,

$$\tau_{j+1} = \begin{cases} \tau_j + 1 & \text{if } S_{\tau_j+1} = S_{\tau_j}, \\ \min\{t > \tau_j : S_t \in \mathcal{H}, S_t \neq S_{\tau_j}\} & \text{otherwise.} \end{cases}$$

Defining $W_j \triangleq S_{\tau_j}$, we observe that (W_j) is a lazy random walk on $\mathcal{H}$. Furthermore, started from any $w \in \mathcal{H}$, there are two options:

(i) Do a step in the 2-core (either stay in w via the lazy rule, which has probability $\frac{1}{2}$, or jump to one of the neighbor of w in $\mathcal{H}$, an event that has probability $d_{\mathcal{H}}(w)/2d_{\tilde{\mathcal{C}}_1}(w)$).

(ii) Enter the PGW-tree attached to w (this happens with probability $d_{\mathcal{T}_w}(w)/2d_{\tilde{\mathcal{C}}_1}(w)$).

It is the latter case that incurs a delay for the random walk on $\tilde{\mathcal{C}}_1$. Since the expected return time to w once entering the tree $\mathcal{T}_w$ is $2(|\mathcal{T}_w| - 1)/d_{\mathcal{T}_w}(w)$, and as the number of excursions to the tree follows a geometric distribution with success probability $1 - d_{\mathcal{T}_w}(w)/2d_{\tilde{\mathcal{C}}_1}(w)$, we infer that

$$\mathbb{E}_w \tau_1 = 1 + \frac{2(|\mathcal{T}_w| - 1)}{d_{\mathcal{T}_w}(w)} \cdot \frac{2d_{\tilde{\mathcal{C}}_1}(w)}{2d_{\tilde{\mathcal{C}}_1}(w) - d_{\mathcal{T}_w}(w)} \leq 4|\mathcal{T}_w|.$$

For some constant $C_1 > 0$ to be specified later, let

$$\ell = C_1 \varepsilon^{-2} \log^2(\varepsilon^3 n), \quad \text{and} \quad a_{v,w}(\ell) = \sum_{j=0}^{\ell-1} \mathbb{P}_v(W_j = w).$$

It follows that

$$\mathbb{E}_v(\tau_\ell) = \sum_{j=0}^{\ell-1} \sum_{w \in \mathcal{H}} \mathbb{P}_v(S_{\tau_j} = w)\mathbb{E}_w \tau_1$$

$$= \sum_{w \in \mathcal{H}} \sum_{j=0}^{\ell-1} \mathbb{P}_v(W_j = w)\mathbb{E}_w \tau_1 \leq 4 \sum_{w \in \mathcal{H}} a_{v,w}(\ell)|\mathcal{T}_w|. \tag{4.10}$$

We now wish to bound the last expression via Lemma 4.1. Let $v \in \mathcal{K}$. Note that, by definition,

$$\sum_{w \in \mathcal{H}} a_{v,w}(\ell) = \ell = C_1 \varepsilon^{-2} \log^2(\varepsilon^3 n).$$

Moreover, by Proposition 3.7, there exists some constant $C_2 > 0$ (which depends on C_1) such that w.h.p.

$$\max_{w \in \mathcal{H}} a_{v,w}(\ell) \leq C_2 \varepsilon^{-1} \log(\varepsilon^3 n).$$

Hence, Lemma 4.1 (applied on the sequence $\{a_{v,w}(\ell) : w \in \mathcal{H}\}$) gives that there exists some constant $C > 0$ (depending only on C_1, C_2) such that

$$\sum_{w \in \mathcal{H}} a_{v,w}(\ell)|\mathcal{T}_v| \leq C\varepsilon^{-3} \log^2(\varepsilon^3 n) \quad \text{except with probability } (\varepsilon^3 n)^{-2} \,.$$

Since $|\mathcal{K}| = (\frac{4}{3} + o(1))\varepsilon^3 n$ w.h.p., taking a union bound over the vertices of the kernel while recalling (4.10) implies that w.h.p.,

$$\mathbb{E}_v(\tau_\ell) \leq C\varepsilon^{-3} \log^2(\varepsilon^3 n) \quad \text{for all } v \in \mathcal{K} \,. \tag{4.11}$$

We next wish to bound the hitting time to the kernel $\mathcal{K}$, defined next:

$$\tau_\mathcal{K} = \min\{t : S_t \in \mathcal{K}\} \,.$$

Recall that from any $v \in \tilde{\mathcal{C}}_1$, after time τ_1 we will have hit a vertex in the 2-core, hence for any $v \in \tilde{\mathcal{C}}_1$ we have

$$\mathbb{E}_v \tau_\mathcal{K} \leq \mathbb{E}_v \tau_1 + \max_{w \in \mathcal{H}} \mathbb{E}_w \tau_\mathcal{K} \,. \tag{4.12}$$

To bound the first summand, since

$$\max_{v \in \tilde{\mathcal{C}}_1} \mathbb{E}_v \tau_1 = \max_{w \in \mathcal{H}} \max_{v \in \mathcal{T}_w} \mathbb{E}_v \tau_1 \,,$$

it clearly suffices to bound $\mathbb{E}_v \tau_w$ for all $w \in \mathcal{H}$ and $v \in \mathcal{T}_w$. To this end, let $w \in \mathcal{H}$, and let $\tilde{S}_t$ be the lazy random walk on $\mathcal{T}_w$. As usual, define $\tilde{\tau}_v = \min\{t : \tilde{S}_t = v\}$. Clearly, for all $v \in \mathcal{T}_w$ we have $\mathbb{E}_v \tau_w = \mathbb{E}_v \tilde{\tau}_w$. We bound $\mathbb{E}_v \tilde{\tau}_w$ by $\mathbb{E}_v \tilde{\tau}_w + \mathbb{E}_w \tilde{\tau}_v$, i.e., the commute time between v and w. Denote by $\mathrm{R}_{\mathrm{eff}}(v, w)$ the effective resistance between v and w when each edge has unit resistance. The commute time identity of [7] (see also [30]) yields that

$$\mathbb{E}_v \tilde{\tau}_w + \mathbb{E}_w \tilde{\tau}_v \leq 4|\mathcal{T}_w| R_{\mathrm{eff}(v \leftrightarrow w)} \leq 4|\mathcal{T}_w| \operatorname{diam}(\mathcal{T}_w) \,, \tag{4.13}$$

Now, Lemmas 4.4 and 4.5 give that for any $w \in \mathcal{H}$, with probability at least $1 - O(\varepsilon(\varepsilon^3 n)^{-2})$,

$$|\mathcal{T}_w| \leq 6\varepsilon^{-2} \log(\varepsilon^3 n) \quad \text{and} \quad \operatorname{diam}(\mathcal{T}_w) \leq 2\varepsilon^{-1} \log(\varepsilon^3 n) \,. \tag{4.14}$$

Since w.h.p. $|\mathcal{H}| = (2 + o(1))\varepsilon^2 n$, we can sum the above over the vertices of $\mathcal{H}$ and conclude that w.h.p., (4.14) holds simultaneously for all $w \in \mathcal{H}$. Plugging this in (4.13), we deduce that

$$\mathbb{E}_v \tilde{\tau}_w + \mathbb{E}_w \tilde{\tau}_v \leq 48\varepsilon^{-3} \log^2(\varepsilon^3 n) \,,$$

and altogether, as the above holds for every $w \in \mathcal{H}$,

$$\max_{v \in \tilde{\mathcal{C}}_1} \mathbb{E}_v \tau_1 \leq 48\varepsilon^{-3} \log^2(\varepsilon^3 n) \,. \tag{4.15}$$

For the second summand in (4.12), consider $e \in \mathcal{K}$ and let $\mathcal{P}_e$ be the 2-path corresponding to e in the 2-core $\mathcal{H}$. Recall that w.h.p. the longest

such 2-path in the 2-core has length $(1 + o(1))\varepsilon^{-1} \log(\varepsilon^3 n)$. Since from each point $v \in \mathcal{P}_e$ we have probability at least $2/|\mathcal{P}_e|$ to hit one of the endpoints of the 2-path (belonging to $\mathcal{K}$) before returning to v, it follows that w.h.p., for every $e \in \mathcal{K}$ and $v \in \mathcal{P}_e$ we have

$$\max_{w \in \mathcal{P}_e} \mathbb{E}_w \#\{t \leq \tau_{\mathcal{K}} : W_t = v\} \leq (\tfrac{1}{2} + o(1))\varepsilon^{-1} \log(\varepsilon^3 n) \,. \tag{4.16}$$

We now wish to apply Lemma 4.1 to the sequence $a_v = \max_{w \in \mathcal{P}_e} \mathbb{E}_w \#\{t \leq \tau_{\mathcal{K}} : W_t = v\}$. Since this sequence satisfies

$$\max_{v \in \mathcal{P}_e} a_v \leq (\tfrac{1}{2} + o(1))\varepsilon^{-1} \log(\varepsilon^3 n) \,, \quad \sum_{v \in \mathcal{P}_e} a_v \leq (\tfrac{1}{2} + o(1))\varepsilon^{-2} \log^2(\varepsilon^3 n) \,,$$

we deduce that there exists some absolute constant $C' > 0$ such that, except with probability $O((\varepsilon^3 n)^{-2})$, every $w \in \mathcal{P}_e$ satisfies

$$\mathbb{E}_w \tau_{\mathcal{K}} \leq C' \varepsilon^{-3} \log^2 \varepsilon^3 n \,. \tag{4.17}$$

Recalling that $e(\mathcal{K}) = (2 + o(1))\varepsilon^3 n$ w.h.p., we deduce that w.h.p. this statement holds simultaneously for all $w \in \mathcal{H}$. Plugging (4.15) and (4.17) into (4.12) we conclude that w.h.p.

$$\mathbb{E}_v \tau_{\mathcal{K}} \leq (C' + 48)\varepsilon^{-3} \log^2 \varepsilon^3 n \text{ for all } v \in \tilde{\mathcal{C}}_1 \,.$$

Finally, we will now translate these hitting time bounds into an upper bound on the forget-time for S_t. Let $\pi_{\mathcal{H}}$ denote the stationary measure on the walk restricted to $\mathcal{H}$:

$$\pi_{\mathcal{H}}(w) = d_{\mathcal{H}}(w)/2e(\mathcal{H}) \quad \text{for } w \in \mathcal{H} \,.$$

Theorem 3.1 enables us to choose some absolute constant $C_1 > 0$ so that $\ell = \ell(C_1)$ would w.h.p. satisfy

$$\max_{w \in \mathcal{H}} \left\| \frac{1}{\ell} \sum_{j=1}^{\ell} \mathbb{P}_w(W_j \in \cdot) - \pi_{\mathcal{H}} \right\|_{\mathrm{TV}} \leq \frac{1}{4} \,. \tag{4.18}$$

Define $\bar{\tau}_0 = \tau_{\mathcal{K}}$ and for $j \geq 0$, define $\bar{\tau}_{j+1}$ as we did for τ_j's, that is,

$$\bar{\tau}_{j+1} = \begin{cases} \bar{\tau}_j + 1 & \text{if } S_{\bar{\tau}_j + 1} = S_{\bar{\tau}_j} \,, \\ \min\{t > \bar{\tau}_j : S_t \in \mathcal{H}, \ S_t \neq S_{\bar{\tau}_j}\} & \text{otherwise.} \end{cases} \,.$$

Let Γ be the stopping rule that selects $j \in \{0, \ldots, \ell - 1\}$ uniformly and then stops at $\bar{\tau}_j$. By (4.18), w.h.p.

$$\max_{v \in \tilde{\mathcal{C}}_1} \left\| \mathbb{P}_v(S_\Gamma \in \cdot) - \pi_{\mathcal{H}} \right\|_{\mathrm{TV}} \leq \frac{1}{4} \,.$$

Furthermore, combining (4.11) and (4.17), we get that w.h.p. for any $v \in \tilde{\mathcal{C}}_1$:

$$\mathbb{E}_v \bar{\tau}_\ell \leq (C + C' + 48)\varepsilon^{-3} \log^2(\varepsilon^3 n) \,.$$

Altogether, we can conclude that the forget-time for S_t w.h.p. satisfies that

$$\mathcal{F}_{1/4} \leq \max_{v \in \tilde{\mathcal{C}}_1} \mathbb{E}_v \bar{\tau}_\ell \leq (C + C' + 48)\varepsilon^{-3} \log^2(\varepsilon^3 n) \,.$$

This translates into the required upper bound on t_{MIX} via an application of Theorems 2.3 and 2.4. ∎

4.3. **Proof of Theorem 1: Lower bound on the mixing time.** As before, by Theorem 2.1 it suffices to prove the analogous statement for $\tilde{\mathcal{C}}_1$.

Let r, s be as in Lemma 4.6, i.e.,

$$r = \lceil \tfrac{1}{8}\varepsilon^{-1} \log(\varepsilon^3 n) \rceil \quad \text{and} \quad s = \tfrac{1}{8}\varepsilon^{-2} \log(\varepsilon^3 n) \,.$$

Let $\mathcal{T}_v$ for $v \in \mathcal{H}$ be the $\mathrm{PGW}(\mu)$-tree that is attached to the vertex v. Lemma 4.6 gives that when n is sufficiently large, every $v \in \mathcal{H}$ satisfies

$$\mathbb{P}(A_{r,s}(\mathcal{T}_v)) \geq \varepsilon(\varepsilon^3 n)^{-2/3} \,.$$

Since $|\mathcal{H}| = (2+o(1))\varepsilon^2 n$ w.h.p. (recall Theorem 2.1), and since $\{\mathcal{T}_v : v \in \mathcal{H}\}$ are i.i.d. given $\mathcal{H}$, we can conclude that w.h.p. there exists some $\rho \in \mathcal{H}$ such that $A_{r,s}(\mathcal{T}_\rho)$ holds. Let $\rho \in \mathcal{H}$ therefore be such a vertex.

Let (S_t) be a lazy random walk on $\tilde{\mathcal{C}}_1$ and π be its stationary distribution. As usual, let $\tau_v \overset{\triangle}{=} \min\{t : S_t = v\}$. We wish to prove that

$$\max_{w \in \mathcal{T}_\rho} \mathbb{P}_w \left(\tau_\rho \geq \tfrac{2}{3} rs \right) \geq \tfrac{1}{3} \,. \tag{4.19}$$

For $w \in \mathcal{T}_\rho$, let $\mathcal{T}_w$ be the entire subtree rooted at w. Further let L_r be the vertices of the r-th level of $\mathcal{T}_\rho$. By our assumption on $\mathcal{T}_\rho$, there is some $\xi \in L_r$ such that $|\mathcal{T}_\xi| \geq s$.

We will derive a lower bound on $\mathbb{E}_\xi \tau_\rho$ from the following well-known connection between hitting-times of random walks and flows on electrical networks (see [30] and also [23, Proposition 2.19]).

Lemma 4.7 ([30])**.** *Given a graph $G = (V, E)$ with a vertex z and a disjoint subset of vertices Z, let $v(\cdot)$ be the voltage when a unit current flows from z to Z and the voltage is 0 on Z. Then $\mathbb{E}_z \tau_Z = \sum_{x \in V} d(x)v(x)$.*

In our setting, we consider the graph $\tilde{\mathcal{C}}_1$. Clearly, the effective resistance between ρ and ξ satisfies $R_{\mathrm{eff}}(\rho \leftrightarrow \xi) = r$. If a unit current flows from ξ to ρ and $v(\rho) = 0$, it follows from Ohm's law that $v(\xi) = r$. Notice that for any $w \in \mathcal{T}_\xi$, the flow between w and ξ is 0. Altogether, we deduce that

$$v(w) = r \quad \text{for all } w \in \mathcal{T}_\xi \,.$$

Therefore, Lemma 4.7 implies that

$$\mathbb{E}_\xi \tau_\rho \geq r|T_\xi| \geq rs \,.$$

Clearly, if $w^\star \in \mathcal{T}_\rho$ attains $\max\{\mathbb{E}_w \tau_\rho : w \in \mathcal{T}_\rho\}$ then clearly

$$\mathbb{E}_{w^\star} \tau_\rho \leq \tfrac{2}{3} rs + \mathbb{P}_{w^\star}\left(\tau_\rho \geq \tfrac{2}{3} rs\right) \mathbb{E}_{w^\star} \tau_\rho \, .$$

On the other hand,

$$\mathbb{E}_{w^\star} \tau_\rho \geq \mathbb{E}_\xi \tau_\rho \geq rs \, ,$$

hence we obtain (4.19).

Recall that w.h.p. $|\tilde{\mathcal{C}}_1| = (2 + o(1))\varepsilon n$. Together with Lemma 4.5, we deduce that w.h.p. every $v \in \mathcal{H}$ satisfies

$$|\mathcal{T}_v| \leq 6\varepsilon^{-2} \log(\varepsilon^3 n) = o(|\tilde{\mathcal{C}}_1|) \, .$$

In particular, $|\mathcal{T}_\rho| = o(|\tilde{\mathcal{C}}_1|)$, and so (as it is a tree) $\pi(\mathcal{T}_\rho) = o(1)$. However, (4.19) states that with probability at least $\frac{1}{3}$, the random walk started at some $w \in \mathcal{T}_\rho$ does not escape from $\mathcal{T}_\rho$, hence

$$\max_{w \in \tilde{\mathcal{C}}_1} \|\mathbb{P}_w(S_{2rs/3} \in \cdot) - \pi\|_{\mathrm{TV}} \geq \tfrac{1}{4} \, .$$

where π is the stationary measure for the random walk S_t on $\tilde{\mathcal{C}}_1$. In other words, we have that

$$t_{\mathrm{MIX}}\left(\tfrac{1}{4}\right) \geq \tfrac{2}{3} rs = \tfrac{1}{96}\varepsilon^{-3} \log^2(\varepsilon^3 n) \, ,$$

as required. ∎

5. Mixing in the subcritical regime

In this section, we give the proof of Theorem 2. By Theorem 1 and the well known duality between the subcritical and supercritical regimes (see [21]), it suffices to establish the statement for the subcritical regime of $\mathcal{G}(n, p)$.

For the upper bound, by results of [5] and [21] (see also [25]), we know that the largest component has size $O(\varepsilon^{-2} \log(\varepsilon^3 n))$ w.h.p., and by results of [22], the largest diameter of a component is w.h.p. $O(\varepsilon^{-1} \log(\varepsilon^3 n))$. Therefore, the maximal hitting time to a vertex is $O(\varepsilon^{-3} \log^2(\varepsilon^3 n))$ uniformly for all components, yielding the desired upper bound on the mixing time.

In order to establish the lower bound, we will demonstrate the existence of a component with a certain structure, and show that the order of the mixing time on this particular component matches the above upper bound.

To find this component, we apply the usual exploration process until εn vertices are exposed. By definition, each component revealed is a Galton-Watson tree (the exploration process does not expose the tree-excess) where the offspring distribution is stochastically dominated by $\mathrm{Bin}(n, \frac{1-\varepsilon}{n})$ and stochastically dominates $\mathrm{Bin}(n, \frac{1-2\varepsilon}{n})$.

It is well known (see, e.g., [17, equation (1.12)]) that for any $\lambda > 0$,

$$\left\|\mathrm{Bin}(n, \tfrac{\lambda}{n}) - \mathrm{Po}(\lambda)\right\|_{\mathrm{TV}} \leq \lambda^2/n \, .$$

It follows that when discovering the first εn vertices, we can approximate the binomial variables by Poisson variables, at the cost of a total error of at most $\varepsilon n(1/n) = \varepsilon = o(1)$.

Lemma 5.1. *With high probability, once εn vertices are exposed in the exploration process, we will have discovered at least $\varepsilon^2 n/2$ components.*

Proof. Notice that each discovered component is stochastically dominated (with respect to containment) by a Poisson$(1-\varepsilon)$-Galton-Watson tree. Thus, the probability that the first $\varepsilon^2 n/2$ components contain more than εn vertices is bounded by the probability that the total size of $\varepsilon^2 n/2$ independent PGW$(1 - \varepsilon)$-trees is larger than εn. The latter can be estimated (using Chebyshev's inequality and Claim 4.2) by

$$\mathbb{P}\Big(\sum_{i=1}^{\varepsilon^2 n/2}|\mathcal{T}_i| \geq \varepsilon n\Big) \leq \frac{\varepsilon^2 n \varepsilon^{-3}}{(\varepsilon n/2)^2} = 4(\varepsilon^3 n)^{-1} = o(1)\,. \qquad \blacksquare$$

For a rooted tree $\mathcal{T}$, we define the following event, analogous to the event $A_{r,s}(\mathcal{T})$ from Subsection 4.1:

$$B_{r,s}(\mathcal{T}) \triangleq \{\exists v, w \in \mathcal{T} \text{ such that } |\mathcal{T}_v| \geq s\,, |\mathcal{T}_w| \geq s \text{ and } \operatorname{dist}(v,w) = r\}\,.$$

The next lemma estimates the probability that the above defined event occurs in a PGW-tree.

Lemma 5.2. *Let $\mathcal{T}$ be a PGW$(1 - 2\varepsilon)$-tree and set $r = \lceil \frac{1}{20}\varepsilon^{-1}\log(\varepsilon^3 n)\rceil$ and $s = \frac{1}{64}\varepsilon^{-2}\log(\varepsilon^3 n)$. Then for some $c > 0$ and any sufficiently large n,*

$$\mathbb{P}(B_{r,s}(\mathcal{T})) \geq c\varepsilon(\varepsilon^3 n)^{-1/2}\,.$$

Proof. The proof follows the general argument of Lemma 4.6. By Lemma 4.4,

$$\mathbb{P}(L_{1/\varepsilon} \neq \emptyset) = \Theta(\varepsilon)\,.$$

Combined with the proof of Claim 4.2 (see (4.4) in particular), we get that

$$\mathbb{E}\big(|L_{1/\varepsilon}| \mid L_{1/\varepsilon} \neq \emptyset\big) = \Theta\big(\varepsilon^{-1}\big)\,, \text{ and } \operatorname{Var}\big(|L_{1/\varepsilon}| \mid L_{1/\varepsilon} \neq \emptyset\big) = \Theta\big(\varepsilon^{-2}\big)\,.$$

Applying Chebyshev's inequality, we get that for some constants $c_1, c_2 > 0$

$$\mathbb{P}\big(|L_{1/\varepsilon}| > c_1\varepsilon^{-1} \mid L_{1/\varepsilon} \neq \emptyset\big) \geq c_2\,.$$

Repeating the arguments for the proof of Lemma 4.6, we conclude that for a PGW$(1-2\varepsilon)$-tree $\mathcal{T}$, the probability that the event $A_{r,s}(\mathcal{T})$ occurs (using r, s as defined in the current lemma) is at least $\varepsilon(\varepsilon^3 n)^{-1/4}$ for n large enough. Thus (by the independence of the subtrees rooted in the $(1/\varepsilon)$-th level),

$$\mathbb{P}\left(\bigcup\Big\{A_{r,s}(\mathcal{T}_u) \cap A_{r,s}(\mathcal{T}_{u'}) : {}^{u,u'\in L_{1/\varepsilon}}_{u \neq u'}\Big\} \,\Big|\, |L_{1/\varepsilon}| > c_1\varepsilon^{-1}\right) \geq c(\varepsilon^3 n)^{-1/2}$$

for some $c > 0$. Altogether, we conclude that for some $c' > 0$,

$$\mathbb{P}\bigg(\bigcup\Big\{A_{r,s}(\mathcal{T}_u) \cap A_{r,s}(\mathcal{T}_{u'}) : \substack{u,u'\in L_{1/\varepsilon} \\ u\neq u'}\Big\}\bigg) \geq c'\varepsilon(\varepsilon^3 n)^{-1/2}\,,$$

which immediately implies that required bound on $\mathbb{P}(B_{r,s}(\mathcal{T}))$. ∎

Combining Lemmas 5.1 and 5.2, we conclude that w.h.p., during our exploration process we will find a tree $\mathcal{T}$ which satisfies the event $B_{r,s}(\mathcal{T})$ for r,s as defined in Lemma 5.2. Next, we will show that the component of $\mathcal{T}$ is indeed a tree, namely, it has no tree-excess. Clearly, edges belonging to the tree-excess can only appear between vertices that belong either to the same level or to successive levels (the root of the tree $\mathcal{T}$ is defined to be the vertex in $\mathcal{T}$ that is first exposed). Therefore, the total number of candidates for such edges can be bounded by $4\sum_i |L_i|^2$ where L_i is the i-th level of vertices in the tree. The next claim provides an upper bound for this sum.

Claim 5.3. *Let r,s be defined as in Lemma 5.2. Then the $\mathrm{PGW}(1-\varepsilon)$-tree $\mathcal{T}$ satisfies $\mathbb{E}\big[\sum_i |L_i|^2 \mid B_{r,s}(\mathcal{T})\big] = O\big(\varepsilon^{-3}\sqrt{\varepsilon^3 n}\big)$.*

Proof. Recalling Claim 4.2 and in particular equation (4.4), it follows that $\mathbb{E}\big(\sum_i |L_i|^2\big) \leq \varepsilon^{-2}$. Lemma 5.2 now implies the required upper bound. ∎

By the above claim and Markov's inequality, we deduce that w.h.p. there are, say, $O\big(\varepsilon^{-3}(\varepsilon^3 n)^{2/3}\big)$ candidates for edges in the tree-excess of the component of $\mathcal{T}$. Crucially, whether or not these edges appear is independent of the exploration process, hence the probability that any of them appears is at most $O\big((\varepsilon^3 n)^{-1/3}\big) = o(1)$. Altogether, we may assume that the component of $\mathcal{T}$ is indeed a tree which satisfies the event $B_{r,s}(\mathcal{T})$.

It remains to establish the lower bound on the mixing time of the random walk on the tree $\mathcal{T}$. Let v,w be two distinct vertices in the r-th level satisfying $|\mathcal{T}_v| \geq s$ and $|\mathcal{T}_w| \geq s$. By the same arguments used to prove (4.19), we have that

$$\max_{u\in\mathcal{T}_v}\mathbb{P}_u(\tau_w \geq 10^{-3}rs) \geq 1 - 10^{-3}\,.$$

Recall that w.h.p. $|\mathcal{T}| \leq 6\varepsilon^{-2}\log(\varepsilon^3 n) = 384s$. It now follows that w.h.p. the mixing time of the random walk on this components satisfies

$$t_{\mathrm{MIX}}(\delta) \geq 10^{-3}rs\,, \text{ for } \delta = \tfrac{1}{384} - 10^{-3} \geq 10^{-3}\,.$$

The lower bound on $t_{\mathrm{MIX}}(\tfrac{1}{4})$ now follows from the definition of r,s. ∎

ACKNOWLEDGMENTS

References

[1] D. Aldous and J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs*. In preparation, `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[2] D. Aldous, L. Lovász, and P. Winkler, *Mixing times for uniformly ergodic Markov chains*, Stochastic Process. Appl. **71** (1997), no. 2, 165–185.

[3] I. Benjamini, G. Kozma, and N. C. Wormald, *The mixing time of the giant component of a random graph*. preprint, available at `http://arxiv.org/abs/math/0610459`.

[4] B. Bollobás, *Random graphs*, 2nd ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.

[5] B. Bollobás, *The evolution of random graphs*, Trans. Amer. Math. Soc. **286** (1984), no. 1, 257–274.

[6] I. Benjamini and E. Mossel, *On the mixing time of a simple random walk on the super critical percolation cluster*, Probab. Theory Related Fields **125** (2003), no. 3, 408–420.

[7] A. K. Chandra, P. Raghavan, W. L. Ruzzo, R. Smolensky, and P. Tiwari, *The electrical resistance of a graph captures its commute and cover times*, Comput. Complexity **6** (1996/97), no. 4, 312–340.

[8] J. Ding, J. H. Kim, E. Lubetzky, and Y. Peres, *Anatomy of a young giant component in the random graph*. preprint, available at `http://arxiv.org/abs/0906.1839`.

[9] J. Ding, J. H. Kim, E. Lubetzky, and Y. Peres, *Diameters in supercritical random graphs via first passage percolation*. available at `http://arxiv.org/abs/0906.1840`.

[10] P. Erdős and A. Rényi, *On random graphs. I*, Publ. Math. Debrecen **6** (1959), 290–297.

[11] N. Fountoulakis and B. A. Reed, *Faster mixing and small bottlenecks*, Probab. Theory Related Fields **137** (2007), no. 3-4, 475–486.

[12] N. Fountoulakis and B. A. Reed, *The evolution of the mixing rate of a simple random walk on the giant component of a random graph*, Random Structures Algorithms **33** (2008), no. 1, 68–86.

[13] M. Heydenreich and R. van der Hofstad, *Random graph asymptotics on high-dimensional tori II. Volume, diameter and mixing time*. preprint, available at `http://arxiv.org/abs/0903.4279`.

[14] M. Jerrum and A. Sinclair, *Approximate counting, uniform generation and rapidly mixing Markov chains*, Inform. and Comput. **82** (1989), no. 1, 93–133.

[15] J. H. Kim, *Poisson cloning model for random graphs*, International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 873–897.

[16] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[17] T. Lindvall, *Lectures on the coupling method*, Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1992 original.

[18] L. Lovász and R. Kannan, *Faster mixing via average conductance*, Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999), ACM, New York, 1999, pp. 282–287 (electronic).

[19] L. Lovász and P. Winkler, *Mixing of random walks and other diffusions on a graph*, Surveys in combinatorics, 1995 (Stirling), London Math. Soc. Lecture Note Ser., vol. 218, Cambridge Univ. Press, Cambridge, 1995, pp. 119–154.

[20] L. Lovász and P. Winkler, *Mixing times*, Microsurveys in discrete probability (Princeton, NJ, 1997), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 41, Amer. Math. Soc., Providence, RI, 1998, pp. 85–133.

[21] T. Łuczak, *Component behavior near the critical point of the random graph process*, Random Structures Algorithms **1** (1990), no. 3, 287–310.

[22] T. Łuczak, *Random trees and random graphs*, Proceedings of the Eighth International Conference "Random Structures and Algorithms" (Poznan, 1997), 1998, pp. 485–500.

[23] R. Lyons with Y. Peres, *Probability on Trees and Networks*, Cambridge University Press, 2008. In preparation. Current version available at `http://mypage.iu.edu/~rdlyons/prbtree/book.pdf`.

[24] P. Mathieu and E. Remy, *Isoperimetry and heat kernel decay on percolation clusters*, Ann. Probab. **32** (2004), no. 1A, 100–128.

[25] A. Nachmias and Y. Peres, *Component sizes of the random graph outside the scaling window*, ALEA Lat. Am. J. Probab. Math. Stat. **3** (2007), 133–142 (electronic).

[26] A. Nachmias and Y. Peres, *Critical random graphs: diameter and mixing time*, Ann. Probab. **36** (2008), no. 4, 1267–1286.

[27] Y. Peres, *Probability on trees: an introductory climb*, Lectures on probability theory and statistics (Saint-Flour, 1997), Lecture Notes in Math., vol. 1717, Springer, Berlin, 1999, pp. 193–280.

[28] J. Pitman, *Enumerations of trees and forests related to branching processes and random walks*, Microsurveys in discrete probability (Princeton, NJ, 1997), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 41, Amer. Math. Soc., Providence, RI, 1998, pp. 163–180.

[29] O. Riordan and N. C. Wormald, *The diameter of sparse random graphs.* preprint, available at `http://arxiv.org/abs/0808.4067`.

[30] P. Tetali, *Random walks and the effective resistance of networks*, J. Theoret. Probab. **4** (1991), no. 1, 101–109.

Jian Ding

Department of Statistics, UC Berkeley, Berkeley, CA 94720, USA.
  E-mail address: `jding@stat.berkeley.edu`

Eyal Lubetzky

Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA.
  E-mail address: `eyal@microsoft.com`

Yuval Peres

Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA.
  E-mail address: `peres@microsoft.com`