

# Cutoff for Random Walk on Dynamical Erdős-Rényi Graph

Perla Sousi   Sam Thomas

## Abstract

We consider dynamical percolation on the complete graph  $K_n$ , where each edge refreshes its state at rate  $\mu \ll 1/n$ , and is then declared open with probability  $p = \lambda/n$  where  $\lambda > 1$ . We study a random walk on this dynamical environment which jumps at rate  $1/n$  along every open edge. We show that the mixing time of the full system exhibits cutoff at  $\log n/\mu$ . We do this by showing that the random walk component mixes faster than the environment process; along the way, we control the time it takes for the walk to become isolated.

*Keywords.* Dynamical percolation, Erdős-Rényi, random walk, mixing times, coupling  
*MSC 2010 subject classifications.* 05C81, 60J27, 60K35, 60K37

## 1 Introduction

In this paper we consider a random walk on a dynamically evolving random graph. Fix an underlying graph  $G = (V, E)$ . The dynamics of the graph are that of *dynamical percolation*: edges refresh at rate  $\mu$ ; upon refreshing, the edge is declared open with probability  $p$  and closed with probability  $1 - p$ . We denote the state of the graph at time  $t$  by  $\eta_t \in \{0, 1\}^E$ : 0 corresponds to a closed edge and 1 to an open edge. The location of the random walker at time  $t$  is denoted  $X_t \in V$ : it moves at rate 1; when its exponential clock rings, it chooses uniformly at random a neighbour  $v$  of its current location,  $x$  say, and jumps from  $x$  to  $v$  if and only if the edge connecting  $x$  and  $v$  is open (at this time), otherwise it remains in place.

Write  $\pi_{\text{RW}}$  for the invariant distribution of the nearest-neighbour simple random walk on  $G$  (ie the degree-biased distribution) and  $\pi_p$  for the product measure on  $\{0, 1\}^E$  with density  $p$ . This process is reversible with invariant distribution  $\pi_{\text{RW}} \times \pi_p$ .

We emphasise that the pair  $(X_t, \eta_t)_{t \geq 0}$  is Markovian, as is just the graph process  $(\eta_t)_{t \geq 0}$ , while the location of the walker alone  $(X_t)_{t \geq 0}$  is not: indeed, its transitions depend on the current graph. (Note that  $(\eta_t)_{t \geq 0}$  is a biased simple random walk on the hypercube  $\{0, 1\}^E$ .) For all our results, we take  $p = \lambda/n$ , and emphasise that  $\lambda$  is a fixed constant, while  $n$  and  $\mu = \mu_n$  vary.

This model was introduced by Peres, Stauffer and Steif in [23]. They used the torus  $\mathbb{Z}_n^d$  (with  $d$  fixed) as their underlying graph; in this paper we use the complete graph  $K_n$  as our underlying graph. Hence from now on we take

$$V = \{1, \dots, n\} \quad \text{and} \quad E = \{(i, j) \mid i, j \in \{1, \dots, n\}, i \neq j\}.$$

Percolation on the complete graph gives precisely the Erdős-Rényi graph, and hence the name ‘dynamical Erdős-Rényi’; we denote the measure of an Erdős-Rényi graph by  $\pi_{\text{ER}}$ , and note that in this case  $\pi_p = \pi_{\text{ER}}$ . Also, we denote the uniform measure on  $\{1, \dots, n\}$  by  $\pi_U$ , and note that in this case  $\pi_{\text{RW}} = \pi_U$ .

Recall that in the non-dense regime for percolation the expected degree of a vertex is order 1 (ie independent of  $n$ ). Hence, in our model, the walker takes steps on the (approximately) timescale  $n$ . We consider  $\mu \ll 1/n$  so that walk takes a large number of steps before seeing a local update to the graph; for bounded degree graphs, we could take  $\mu \ll 1$ . Our proofs actually require very slightly more, namely a polylogarithmic factor: we consider  $\mu \ll (\log n)^{-\alpha}/n$  for a fixed  $\alpha > 0$ ; no attempt has been made to optimise this parameter.

---

Statistical Laboratory, University of Cambridge, UK  
Email: {p.sousi, s.m.thomas}@statslab.cam.ac.uk

As in [23], we look at the  $\varepsilon$ -mixing time of the *full system*  $(X, \eta)$ :

$$t_{\text{mix}}(\varepsilon) = \inf\{t \geq 0 \mid \max_{x_0, \eta_0} \|\mathbb{P}_{x_0, \eta_0}((X_t, \eta_t) \in \cdot) - \pi\|_{\text{TV}} \leq \varepsilon\}.$$

We consider the *supercritical regime*, ie take  $p = \lambda/n$  where  $\lambda > 1$  is a constant.

In order to state our mixing result, we must first define some notation for iterated logarithm:

$$\text{set } \log_{(1)} n = \log n \quad \text{and define inductively } \log_{(m+1)} n = \log(\log_{(m)} n) \quad \text{for } m \geq 1.$$

**Theorem 1.1** (Cutoff for Full System). *For all  $\lambda > 1$ , all  $\varepsilon \in (0, 1)$ , all  $M \in \mathbb{N}$  and all  $n$  sufficiently large, for  $p = \lambda/n$  and  $\mu \leq (\log n)^{-20}/n$ , we have*

$$t_{\text{mix}}(\varepsilon) = \frac{1}{\mu} \log n \cdot (1 + o(1)) \quad \text{and} \quad |t_{\text{mix}}(\varepsilon) - t_{\text{mix}}(1 - \varepsilon)| \leq \frac{1}{\mu} \log_{(M)} n.$$

We also consider the ‘mixing’ of the *random walk component*:

$$t_{\text{mix}}^{\text{RW}}(\varepsilon, \eta_0) = \sup\{t \geq 0 \mid \max_{x_0} \|\mathbb{P}_{x_0, \eta_0}(X_t \in \cdot) - \pi_U\|_{\text{TV}} \geq \varepsilon\} \quad \text{for } \eta_0 \in \{0, 1\}^E.$$

Since  $X$  is not a Markov chain, we do not have a priori that the total variation distance from uniform is decreasing, hence we do not define the mixing time to be ‘the first time the total variation distance is below  $\varepsilon$ ’, but rather ‘the last time the total variation distance is above  $\varepsilon$ ’. (Of course, for a Markov chain, these notions are the same.) Note that, trivially by projection,  $t_{\text{mix}}^{\text{RW}}(\varepsilon, \eta_0) \leq t_{\text{mix}}(\varepsilon)$  for all  $\varepsilon$  and all  $\eta_0$ . We show that ‘the walk mixes faster than the environment’ when the initial environment is ‘typical’, in the following precise sense.

**Theorem 1.2** (Mixing of Random Walk). *For all  $\lambda > 1$ , all  $\varepsilon \in (0, 1)$ , all  $M \in \mathbb{N}$  and all  $n$  sufficiently large, for  $p = \lambda/n$  there exists a subset  $H \subseteq \{0, 1\}^E$  with  $\pi_{\text{ER}}(H) = 1 - o(1)$  so that, for all  $\eta_0 \in H$ , for  $\mu \leq (\log n)^{-20}/n$ , we have*

$$t_{\text{mix}}^{\text{RW}}(\varepsilon, \eta_0) \leq \log_{(M)} n / \mu;$$

ie, if  $\eta_0 \sim \pi_{\text{ER}}$  then, for all  $\varepsilon \in (0, 1)$ , all  $M \in \mathbb{N}$  and all  $n$  sufficiently large, we have

$$t_{\text{mix}}^{\text{RW}}(\varepsilon, \eta_0) \leq \log_{(M)} n / \mu \quad \text{with probability } 1 - o(1).$$

**Remark 1.3.** From this intuitively it is clear why we get cutoff. First, let the environment mix; this is just a (biased) random walk on the hypercube  $\{0, 1\}^N$  where  $N = \binom{n}{2}$ , and is known to have cutoff at time  $\log n / \mu$ . At this time, the graph is ‘approximately’ Erdős-Rényi, and so likely to be in the set  $H$ . Finally we let the walk mix; this takes time little- $o$  of the mixing of the environment. This is indeed the method that we use, but made fully rigorous.  $\triangle$

**Remark 1.4.** Above we are allowed to choose  $X_0$  dependent on  $\eta_0$  (and vice versa). In §6, we consider drawing  $\eta_0$  according to  $\pi_{\text{ER}}$ , and then choosing  $X_0$  *independently* of  $\eta_0$ . By symmetry, we may assume  $X_0 = 1$ . We then look at the mixing time of the walk on this (evolving) graph:

$$t_{\text{mix}}^{\text{RW}}(\varepsilon) = \sup\{t \geq 0 \mid \|\mathbb{P}_{1, \text{ER}}(X_t \in \cdot) - \pi_U\|_{\text{TV}} \geq \varepsilon\},$$

where  $\mathbb{P}_{x_0, \text{ER}}(\cdot) = \sum_{\eta_0} \mathbb{P}_{x_0, \eta_0}(\cdot) \pi_{\text{ER}}(\eta_0)$  averages the initial environment with respect to the Erdős-Rényi measure. We prove a sharp, up to constants, result on this mixing time. The result does not require us to consider the supercritical regime, ie  $\lambda > 1$ , but allows any  $\lambda \in (0, \infty)$ , including the critical case  $\lambda = 1$ . In §6, we prove the following.

For all  $\lambda > 0$ , there exist positive constants  $\mu_0$  and  $C$  so that, for all  $\varepsilon$  sufficiently small and all  $n$  sufficiently large, for  $p = \lambda/n$  and  $\mu \leq \mu_0/n$ , we have

$$\frac{1}{\mu} \cdot \frac{1}{2\lambda} \log(1/\varepsilon) \leq t_{\text{mix}}^{\text{RW}}(\varepsilon) \leq \frac{1}{\mu} \cdot C \log(1/\varepsilon).$$

Note that once all edges of the graph have been refreshed, the graph has the Erdős-Rényi measure, and is independent of  $\eta_0$ . This is the coupon-collector problem, and takes time concentrated at  $\log N \approx 2 \log n$ . At first glance, then, it appears that the idea of Remark 1.3 along with the above statement can be applied to easily give pre-cutoff at  $2 \log n$ , with an order  $1/\mu$  window. However, this is not the case: to prove the above statement, it is crucial that  $X_0$  is chosen *independently* of  $\eta_0$ ; we then exploit symmetry. At time  $2 \log n$ , the walk and environment are correlated, and so the argument does not apply; hence the need for Theorem 1.2.  $\triangle$

The above statement *suggests* that the upper bound in Theorem 1.2 is probably not sharp. However, the main interest in Theorem 1.2 compared with Theorem 1.1 is not the specific upper bound, but the fact that ‘most’  $\eta_0$  have  $t_{\text{mix}}^{\text{RW}}(\varepsilon, \eta_0) \ll t_{\text{mix}}(\varepsilon)$ , ie that the walk mixes faster than the full system. This is what allows us to show cutoff.

**Remark.** For the rest of the paper, with the exception of §6, we assume that  $\mu \leq (\log n)^{-20}/n$ ; we shall not repeat this in the statement of every theorem. In §6, we drop the  $\log n$  factor: we assume that  $\mu \leq \mu_0/n$ , with  $\mu_0$  a constant as small as desired. Similarly, for the entire paper we assume that  $n$  is as large as we desire.  $\triangle$

The model of dynamical percolation (without the random walker) was introduced by Häggström, Peres and Steif in [14]. The model with the random walk was then introduced by Peres, Stauffer and Steif in [23], with underlying graph the torus  $\mathbb{Z}_n^d$ . They considered the *subcritical regime*, ie  $p < p_c(\mathbb{Z}^d)$ , the critical probability for bond percolation on  $\mathbb{Z}^d$ , and obtained the correct order for the mixing of the full system, showing further that the order of the mixing of the walk is the same order as the mixing of the full system (in contrast to our model). The *supercritical regime*, ie  $p > p_c(\mathbb{Z}^d)$ , was then considered by Peres, Sousi and Steif in [22]. They considered the ‘quenched’ case, where a ‘typical’ environment process  $\{\eta_t\}_{t \geq 0}$  is fixed in advance, and the walker walks on this. They obtained the correct order for the mixing of the walk, up to polylogarithmic factors, but only in the regime where  $\theta(p) > \frac{1}{2}$ , ie the probability that the component at 0 in  $\mathbb{Z}^d$  is infinite is greater than  $\frac{1}{2}$ . The case  $\theta(p) \leq \frac{1}{2}$  remains open.

Avena, Güldaş, van der Hofstad and den Hollander in [2, 3] studied the mixing time of the non-backtracking random walk on a *dynamical configuration model*. The *configuration model* generates a graph uniformly at random among all graphs with a prescribed degree sequence, and the dynamics at every time step ‘rewire’ uniformly at random a given proportion of the edges.

It is straightforward to see that in our model when the walker first crosses a refreshed edge (which is a randomised stopping time) it is then (almost) uniform. The authors of [2, 3] considered an analogous time for the nearest-neighbour simple random walk, namely the first time the walk crosses a rewired edge, and showed that the distribution of the walk at this time is (almost) its invariant distribution. In both cases, however, these times are not sufficient to show mixing: it is not the case that the walk ‘remains close to invariant’; for example, there is significant probability that the walk will cross back over the same edge to its previous location. To resolve this issue in [2, 3], rather than considering the usual nearest-neighbour random walk, the authors consider the non-backtracking random walk. They show sharp asymptotics for the mixing time of this, and do this using the first crossing of a rewired edge: the non-backtracking removes the above issue.

Fountoulakis and Reed in [11] and Benjamini, Kozma and Wormald in [5] studied the mixing time of the nearest-neighbour simple random walk on the giant component of a supercritical Erdős-Rényi random graph, without any graph dynamics: they prove that the mixing time is order exactly  $(\log n)^2$ . Fountoulakis and Reed carefully studied the ratio between the size of the edge boundary of a set and the set itself, using a variation of the Lovász-Kannan integral, which they developed in [10]. Benjamini, Kozma and Wormald used a more geometric approach, defining a stripping process to analyse the (2-)core and the kernel of the graph; they show the kernel is a (type of) expander, and describe the *decorations* attached to the kernel.

The two works above consider mixing from the worst-case starting point. Berestycki, Lubetzky, Peres and Sly in [6] consider mixing when the starting point is chosen according to the invariant distribution. They show then that the mixing time is actually order  $\log n$ , obtaining the correct constant and also showing cutoff; contrast this with order  $(\log n)^2$  for the worst-case.

## 2 Outline of Proof and Preliminaries

### 2.1 Outline of Proofs

We now give a brief, informal outline of the proofs of the main results. First consider the following scenario: suppose a single walker is isolated at vertex  $u$ , and suppose it becomes non-isolated by the edge  $(u, v)$  opening, where  $v$  was isolated immediately before  $(u, v)$  opened. Now the pair  $\{u, v\}$  is a component of the graph. Because  $\mu \ll 1/n$ , we see that the walker takes a large number of steps before this edge closes. If it closes before any other edge incident to  $\{u, v\}$  opens (which has order 1 probability, by counting open/closed edges), then the walker is approximately uniformly distributed on  $\{u, v\}$ . So it has approximately ‘done a lazy simple random walker step’.

This motivates the following coupling. First wait for the two walkers to be simultaneously isolated *in the same environment* of two full systems. Then to couple we want to imitate the standard coupling of the lazy simple random walk on the complete graph; we do this by considering the event that when the walkers become non-isolated they connect to a vertex that was isolated immediately prior. We give a precise definition of the coupling that we use in §5.

In order to find the time it takes for two walkers to be simultaneously isolated in the same environment, we first consider how long it takes one walker to become isolated. To find this time, we observe that a walker can only become isolated if it is at a degree 1 vertex and this vertex becomes isolated prior to the walk’s leaving it. This motivates looking at the rate at which degree 1 vertices are hit. To do this, we compare the number of degree 1 vertices hit by a walker on the dynamic graph and the same quantity for a walker on a static graph; we then apply a Chernoff-style bound due to Gillman [13]. We give the precise details of this in §4.

A key element in studying the isolation time is to control how long the walk remains in the giant once it has entered. We show that since the graph updates slowly, as  $\mu n \ll (\log n)^{-19}$ , the walk does not see updates to the graph for some while; this time is long enough for the walk to become approximately uniform on the giant prior to seeing a change. We can then use structure results on the Erdős-Rényi graph to see how ‘near the core’ of the giant the walk is.

We then use the fact that an Erdős-Rényi graph with one vertex conditioned to be isolated has the distribution of an Erdős-Rényi graph on  $n - 1$  vertices (with edge-probability  $p$ ) union an isolated vertex: this allows us to say that ‘conditioning on one walker’s being isolated has almost no affect on the other walker’, which will allow us to treat the walkers as almost independent. We give the precise details of this in §4.2.

### 2.2 Notation and Terminology

For functions  $f$  and  $g$ , we write  $f(n) \lesssim g(n)$ , or  $f(n) = \mathcal{O}(g(n))$ , if there exists a positive constant  $C$  so that  $f(n) \leq Cg(n)$  for all  $n$ ; write  $f(n) \gtrsim g(n)$ , or  $f(n) = \Omega(g(n))$ , if  $g(n) \lesssim f(n)$ . We write  $f(n) \asymp g(n)$ , or  $f(n) = \Theta(g(n))$ , if we have both  $f(n) \lesssim g(n)$  and  $g(n) \lesssim f(n)$ . Write  $f(n) \ll g(n)$ , or  $f(n) = o(g(n))$ , if  $f(n)/g(n) \rightarrow 0$  as  $n \rightarrow \infty$ ; write  $f(n) \gg g(n)$ , or  $f(n) = \omega(g(n))$ , if  $g(n) \ll f(n)$ .

For random variables  $X$  and  $Y$ , we write  $X \preceq Y$  if  $Y$  stochastically dominates  $X$  (from above), ie if  $\mathbb{P}(X \geq z) \leq \mathbb{P}(Y \geq z)$  for all  $z$ ; write  $X \succ Y$  if  $Y \preceq X$ .

For a real number  $\zeta > 0$ , we write  $\mathcal{E}(\zeta)$  for the exponential random variable with rate  $\zeta$ .

For real numbers  $\alpha$  and  $\beta$ , we write  $\alpha \wedge \beta = \min\{\alpha, \beta\}$ .

### 2.3 Good Graphs and Erdős-Rényi Structure Results

In this section we state some results on Erdős-Rényi graphs. Since a realisation can be *any* graph, we want to describe explicitly what we shall mean by a *good* graph. We now define some notation for graph properties. For the moment, let  $c_*$  and  $C_*$  be any two constants.

**Notation 2.1.** Let  $n \in \mathbb{N}$ ; in the following, we suppress the  $n$ -dependence. For a graph  $G = (V, E)$  with  $V = \{1, \dots, n\}$ , we use the following notation.

- (i) For  $x \in V$ , write  $d(x)$  for the degree of  $x$  (in  $G$ ).

- (ii) Write  $\mathcal{G}$  for the (set of vertices in the) largest component, and call it the *giant*; if there is a tie, choose the option that includes the smallest labelled vertex.
- (iii) For  $x \in V$ , call an edge a *removal edge* for  $x$  if its removal breaks the component of  $x$  in two, leaving  $x$  in the smaller component (breaking ties as above). Write  $\mathcal{R}(x)$  for the set of removal edges for  $x$ , and write  $R(x) = |\mathcal{R}(x)|$ .
- (iv) For  $M \in \mathbb{N}$ , write  $\mathcal{W}^M$  for the vertices of the giant with at most  $C_* \log_{(M)} n$  removal edges, ie

$$\mathcal{W}^M = \{x \in \mathcal{G} \mid R(x) \leq C_* \log_{(M)} n\}.$$

- (v) Write  $\gamma$  for the spectral gap and  $\Phi$  for the isoperimetric constant, ie

$$\Phi = \min \left\{ \frac{|\partial S|}{d(S)} \mid S \neq \emptyset, d(S) \leq |E| \right\} \quad \text{and} \quad \gamma = 1 - \lambda_G,$$

where  $d(S) = \sum_{x \in S} d(x)$ ,  $\partial S = \{(x, y) \in E \mid x \in S, y \notin S\}$  and  $\lambda_G$  is the second largest eigenvalue of the transition matrix of the nearest-neighbour simple random walk on  $G$ .

If the graph  $G = \zeta$ , then we use subscript  $\zeta$ , eg writing  $\mathcal{G}_\zeta$  for the giant of  $\zeta$ . If the graph  $G = \eta_t$ , then we add a subscript  $t$ , eg writing  $\mathcal{G}_t$ ,  $R_t(x)$  or  $d_t(x)$ .  $\triangle$

Let  $\omega_*$  be any function of  $n$ .

**Definition 2.2** (Good Graph). Let  $n \in \mathbb{N}$ ; in the following, we suppress the  $n$ -dependence. We say that a graph  $G = (V, E)$  with  $V = \{1, \dots, n\}$  is *good*, and write  $G \in \mathcal{G}$ , if it has a unique component  $\mathcal{G}$  with  $|\mathcal{G}| \geq C_* \log n$ , which we call the *giant*, which satisfies the following properties.

- (i) *Size*. We have  $|\mathcal{G}| \geq c_* n$ .
- (ii) *Maximum degree*. The maximum degree of  $\mathcal{G}$  is at most  $C_* \log n$ .
- (iii) *Number of edges*. There are at most  $C_* n$  edges in  $\mathcal{G}$ .
- (iv) *Number of degree 1 vertices*. The number of degree 1 vertices in  $\mathcal{G}$  is at least  $c_* n$ .
- (v) *Removal edges*. We have  $R(x) \leq C_* \log n$  for all  $x \in \mathcal{G}$ .
- (vi) *Vertices far from the core*. For all  $2 \leq M \leq \omega_*(n)$ , the proportion of vertices  $x$  of  $\mathcal{G}$  with  $R(x) \geq C_* \log_{(M)} n$  is at most  $(\log_{(M-1)} n)^{-4}$ , ie  $|\mathcal{G} \setminus \mathcal{W}^M|/|\mathcal{G}| \leq (\log_{(M-1)} n)^{-4}$ .
- (vii) *Expansion properties*. We have  $\Phi_{\mathcal{G}} \geq c_*(\log n)^{-2}$  and  $\gamma_{\mathcal{G}} \geq c_*(\log n)^{-4}$ .  $\triangle$

Write  $\mathcal{G}_t = \{\eta_t \in \mathcal{G}\}$ , and also

$$\mathcal{G}[s, t] = \{\eta_u \in \mathcal{G} \forall u \in [s, t]\} = \cap_{s \leq u \leq t} \mathcal{G}_u.$$

**Proposition 2.3.** *There exist positive constants  $c_*$  and  $C_*$  and a function  $\omega_*(n) \rightarrow \infty$  so that, for  $G \sim \pi_{\text{ER}}$ , we have*

$$\mathbb{P}(G \notin \mathcal{G}) = \mathcal{O}(n^{-9}) \quad \text{and} \quad \mathbb{P}\left(\frac{1}{n} |\{x \in V \mid d(x) = 0\}| \leq c_*\right) = \mathcal{O}(n^{-9}).$$

The claims in this proposition are fairly standard, but usually in the literature the proved decay rate is only  $o(1)$ , whereas we desire  $\mathcal{O}(n^{-9})$ . As such we give the proof of this proposition, but defer it until the appendix.

For the remainder of the paper, we select  $(\omega_*, c_*, C_*)$  as guaranteed by this proposition and fix them permanently; whenever  $\omega_*$ ,  $c_*$  or  $C_*$  is written below, it will refer to these constants.

We now consider our graph dynamics. We want not only the starting graph to be good, but we want it to remain good for a long time.

**Definition 2.4.** We make the following definitions:

$$\begin{aligned}\mathcal{H} &= \{\eta_0 \in \{0, 1\}^E \mid \mathbb{P}_{\eta_0}(\mathcal{G}[0, 1/\mu]^c) \leq n^{-1}\}; \\ \mathcal{H}[s, t] &= \{\eta_u \in \mathcal{H} \forall u \in [s, t]\}; \\ H &= \{\eta_0 \in \{0, 1\}^E \mid \mathbb{P}_{\eta_0}(\mathcal{H}[0, n/\mu]^c) \leq n^{-1}\}.\end{aligned}$$

Further, if we are considering two environment processes,  $\eta$  and  $\xi$  say, then we (abuse notation slightly and) use the same notation, eg

$$\mathcal{H}[s, t] = \{\eta_u, \xi_u \in \mathcal{H} \forall u \in [s, t]\}. \quad \triangle$$

We have the following result on ‘how good’ an Erdős-Rényi graph is.

**Proposition 2.5.** *For all  $t \leq n/\mu$ , if  $\eta_0 \sim \pi_{\text{ER}}$ , then we have*

$$\mathbb{P}(\mathcal{H}[0, t]^c) = \mathcal{O}(n^{-3}), \quad \text{and hence } \pi_{\text{ER}}(H) = 1 - o(1).$$

Moreover, these still hold if we add the condition that at least a proportion  $c_*$  of the vertices are isolated to the definition of a good graph.

We state a large deviations result on the Poisson distribution, which we use on a number of occasions throughout the paper.

**Lemma 2.6.** *We have the following bounds, valid for all  $\lambda > 0$  and all  $\varepsilon \in (0, 1)$ :*

$$\begin{aligned}\mathbb{P}(\text{Po}(\lambda) \geq (1 + \varepsilon)\lambda) &\leq \exp(-\tfrac{1}{2}\lambda\varepsilon^2(1 - \tfrac{1}{3}\varepsilon)); \\ \mathbb{P}(\text{Po}(\lambda) \leq (1 - \varepsilon)\lambda) &\leq \exp(-\tfrac{1}{2}\lambda\varepsilon^2).\end{aligned}$$

**Proof of Proposition 2.5.** Since  $\pi_{\text{ER}}$  is the invariant distribution for the environment, using the concentration of the Poisson distribution and Proposition 2.3 we find that

$$\mathbb{P}(\mathcal{G}[0, 1/\mu]^c) = \mathbb{P}(\exists t \leq 1/\mu \text{ s.t. } \eta_t \notin \mathcal{G}) \leq n^2 \cdot \mathcal{O}(n^{-9}) + \exp(-\tfrac{1}{12}n^2) = \mathcal{O}(n^{-7}).$$

We now restrict the implicit sum in  $\mathbb{P}$  from  $\eta_0 \in \{0, 1\}^E$  to  $\eta_0 \in \mathcal{H}^c$ :

$$\begin{aligned}\mathbb{P}(\mathcal{G}[0, 1/\mu]^c) &= \sum_{\eta_0 \in \{0, 1\}^E} \mathbb{P}_{\eta_0}(\mathcal{G}[0, 1/\mu]^c) \pi_{\text{ER}}(\eta_0) \\ &\geq \sum_{\eta_0 \in \mathcal{H}^c} \mathbb{P}_{\eta_0}(\mathcal{G}[0, 1/\mu]^c) \pi_{\text{ER}}(\eta_0) \geq n^{-1} \cdot \pi_{\text{ER}}(\mathcal{H}^c).\end{aligned}$$

Hence we deduce that  $\pi_{\text{ER}}(\mathcal{H}^c) = \mathcal{O}(n^{-6})$ . Repeating the same argument, we find that

$$\pi_{\text{ER}}(H^c) = \mathcal{O}(n^{-2}). \quad \square$$

### 3 Hitting and Exit Times of the Giant

In this section we study the hitting time of the giant, and how long the walk remains in the giant given that it starts there. Write the following for the hitting and exit times of the giant:

$$\tau_{\mathcal{G}} = \inf\{t \geq 0 \mid X_t \in \mathcal{G}_t\} \quad \text{and} \quad \tau'_{\mathcal{G}} = \inf\{t \geq 0 \mid X_t \notin \mathcal{G}_t\}.$$

**Lemma 3.1** (Hitting the Giant). *There exists a positive constant  $c$  so that, for all  $n$  sufficiently large and all  $(x_0, \eta_0)$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(\tau_{\mathcal{G}} \leq 1/\mu) \geq c - \mathbb{P}_{\eta_0}(\mathcal{G}[0, 1/\mu]^c).$$

**Proof.** Write  $\mathcal{U}$  for the first time an edge incident to the walker refreshes and opens. By the memoryless property,  $\mathcal{U} \sim \mathcal{E}(\lambda\mu(1 - 1/n))$ . When such an edge opens, it connects to the giant with probability  $|\mathcal{G}_{\mathcal{U}}|/(n - 1)$ . Write  $\theta_t = |\mathcal{G}_t|/n$ . We then have

$$\mathbb{P}_{x_0, \eta_0}(\tau_{\mathcal{G}} \leq 1/\mu) \geq \mathbb{P}_{x_0, \eta_0}(\tau_{\mathcal{G}} = \mathcal{U}, \mathcal{U} \leq 1/\mu, \theta_{\mathcal{U}-} \geq c_*)$$

$$\begin{aligned} &\geq \mathbb{P}_{x_0, \eta_0}(\tau_{\mathcal{G}} = \mathcal{U} \mid \mathcal{U} \leq 1/\mu, \theta_{\mathcal{U}^-} \geq c_*) \cdot \mathbb{P}_{x_0, \eta_0}(\mathcal{U} \leq 1/\mu, \theta_t \geq c_* \forall t \leq 1/\mu) \\ &\geq c_* (\mathbb{P}_{x_0, \eta_0}(\mathcal{U} \leq 1/\mu) - \mathbb{P}_{\eta_0}(\{\theta_t \geq c_* \forall t \leq 1/\mu\}^c)). \end{aligned}$$

The proof is completed by noting that  $c_* \leq 1$  and  $U \sim \mathcal{E}(\lambda\mu(1-1/n))$  so  $\mathbb{P}_{x_0, \eta_0}(\mathcal{U} \leq 1/\mu) \asymp 1$ .  $\square$

We now consider how long the walker remains in the giant once it enters. Recall the definition of  $R$  and  $\mathcal{W}^M$  from §2.3. Since the number of removal edges satisfies  $R(x) \leq C_* \log n$  for all  $x \in \mathcal{G}$  when the graph is good, while the graph is good a trivial bound is  $\tau'_{\mathcal{G}} \succcurlyeq \mathcal{E}(C_* \log n)$ . We do a more careful analysis which shows that, for all  $M \in \mathbb{N}$ , ‘most of the time’  $X_t \in \mathcal{W}_t^M$ , ie satisfies  $R(X_t) \leq C_* \log_{(M)} n$ ; this is because  $|\mathcal{W}^M|/|\mathcal{G}| = o(1)$  for a good graph. The precise statement that we prove is as follows.

**Proposition 3.2** (Exit Time from the Giant). *There exists a constant  $C$  so that, for all  $M \in \mathbb{N}$ , all  $n$  sufficiently large, all  $t$  with  $(\log n)^{-5} \leq \mu t \leq \frac{1}{10}$  and all  $(x_0, \eta_0)$  with  $\eta_0 \in \mathcal{H}$  and  $x_0 \in \mathcal{G}_0$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(\tau'_{\mathcal{G}} \leq t) \leq C\mu t \log_{(M)} n.$$

Since  $\mu n \ll 1$ , the ‘majority of the time’ the walker takes a step before any edge incident to its location changes state. This motivates looking at a random walker moving on a *static* graph, ie one without graph-dynamics. We call such a walk the *static walk*, and the original walk (on the dynamic graph) the *dynamic walk*; we denote them by  $\tilde{X}$  and  $X$ , respectively.

Consider starting the two walks together. Observe that until the static walk encounters an edge that is in a different state to its original, the two walks have the same distribution, and hence we can couple them to be the same (until this time) as follows. Give  $X$  and  $\tilde{X}$  the same jump clock. When this clock rings, at time  $t$  say, both walks choose the same vertex;  $\tilde{X}$  performs the jump if and only if the connecting edge is present in  $\eta_0$ , while  $X$  performs the jump if and only if the connecting edge is present in  $\eta_t$ . We call this the *static-dynamic coupling*.

We define the *set of edges seen by the walker* (in the interval  $[s, t]$ ) as the set of all edges (open or closed) that are incident to the walker at some time (in the interval  $[s, t]$ ). Until such a seen edge changes state, we can couple  $X$  and  $\tilde{X}$  (as above).

Label the edges of the (complete) graph  $e_1, \dots, e_N$ , where  $N = \binom{n}{2}$ , in an arbitrary ordering (eg lexicographically). At time  $t$ , write  $\mathcal{O}_t = \{o_t^1, o_t^2, \dots\}$  for the (ordered) set of open edges in  $\eta_t$  and  $\mathcal{C}_t = \{c_t^1, c_t^2, \dots\}$  for the set of closed ones. We say that an edge is a *bridge* for a component if its removal splits the component into two (disconnected) parts. Recall also from Notation 2.1 the definition of a removal edge, and of  $\mathcal{R}$ .

Fix  $\rho = (\log n)^{11}$ . We now define a set process  $\mathcal{P} = (\mathcal{P}_t)_{t \geq 0}$ .

**Definition 3.3** (Set Processes). We define the set process  $\mathcal{P}$  inductively. We assume throughout the definition that the graph is good; that is, we shall only define  $(\mathcal{P}_s)_{s \leq t}$  on the event  $\mathcal{G}[0, t]$ .

Suppose we have defined the set process  $\mathcal{P}$  up until time  $t$ , ie have defined  $(\mathcal{P}_s)_{s \in [0, t]}$ . Let  $U$  be the most recent ‘update time’ for the process  $(\mathcal{P}_s)_{s \in [0, t]}$  in the following sense:

$$V' = \sup\{s \in [0, t] \mid \text{an edge of } \mathcal{P}_s \text{ changes state at time } s\}.$$

(If  $t = 0$ , then we take  $V' = 0$ . What follows is also used for the base case of the induction.) For  $r \geq 0$ , let  $\mathcal{A}'_{r+t}$  (respectively  $\mathcal{B}'_{r+t}$ ) be the set of open (respectively closed) edges seen by  $\tilde{X}$  since time  $U$ . For  $s \geq 0$ , write  $\mathcal{R}_s = \mathcal{R}_s(X_s)$ . If  $|\mathcal{A}'_{r+t} \cup \mathcal{R}_{r+t}| \leq \rho$  and  $|\mathcal{B}'_{r+t}| \leq \rho n$ , then set

$$\mathcal{A}_{r+t} = \mathcal{A}'_{r+t} \cup \mathcal{R}_{r+t} \cup \{o_{r+t}^1, \dots, o_{r+t}^{a_{r+t}}\} \quad \text{and} \quad \mathcal{B}_{r+t} = \mathcal{B}'_{r+t} \cup \{c_{r+t}^1, \dots, c_{r+t}^{b_{r+t}}\}, \quad (3.1)$$

where  $a_{r+t}$  and  $b_{r+t}$  are so that  $|\mathcal{A}_{r+t}| = \rho$  and  $|\mathcal{B}_{r+t}| = \rho n$ ; otherwise, set

$$\mathcal{A}_{r+t} = \{o_{r+t}^1, \dots, o_{r+t}^{\rho}\} \quad \text{and} \quad \mathcal{B}_{r+t} = \{c_{r+t}^1, \dots, c_{r+t}^{\rho n}\}. \quad (3.2)$$

Let  $V$  be the first update time for the process  $(\mathcal{A}_{r+t} \cup \mathcal{B}_{r+t})_{r > 0}$ :

$$V = \inf\{r \geq 0 \mid \text{an edge of } \mathcal{A}_{r+t} \cup \mathcal{B}_{r+t} \text{ changes state at time } r+t\} + t.$$



Define  $\mathcal{P}_{r+t} = \mathcal{A}_{r+t} \cup \mathcal{B}_{r+t}$  for  $r \in [0, V)$ .

If at time  $s$  we use (3.1) to define  $\mathcal{A}_s$  and  $\mathcal{B}_s$ , then we say that the set definitions *succeeded* at time  $s$ , and write  $\mathcal{S}_s$  for this event; if we used (3.2), then we say they have *failed*.

Finally, for  $t > 0$  define the event  $\mathcal{S}[0, t) = (\cap_{s < t} \mathcal{S}_s) \cap \mathcal{G}[0, t)$ .  $\triangle$

**Definition 3.4.** Set  $S = n(\log n)^8$  and  $U_0 = 0$ . We say that  $\mathcal{P}$  *updates* (at time  $t$ ) when an edge in the set  $\mathcal{P}$  refreshes and changes state (at time  $t$ ). We inductively define the sequence  $U_1, U_2, \dots$ : for all  $k \geq 1$ , let  $V_k$  be the first time after  $U_{k-1}$  that  $\mathcal{P}$  updates; set  $U_k = V_k \wedge (U_{k-1} + S)$ .  $\triangle$

Work on the event  $\mathcal{S}[0, t)$ , and fix  $s < t$ . Then  $\mathcal{A}_s$  is a collection of open edges of size  $\rho$  and  $\mathcal{B}_s$  is a collection of closed edges of size  $\rho n$ . Hence the set  $\mathcal{P}_s = \mathcal{A}_s \cup \mathcal{B}_s$  updates at rate  $\kappa\mu$  where

$$\kappa = (1-p)\rho + p\rho n = (1-\lambda/n)\rho + \lambda\rho = (1+\lambda-\lambda/n)\rho; \quad \text{note that } \kappa \asymp \rho = (\log n)^{11}.$$

By the memoryless property,  $U_k - U_{k-1} \sim^{\text{iid}} \mathcal{E}(\kappa\mu) \wedge S$ . Observe also that the walk may only leave the giant when the set  $\mathcal{P}$  updates, in particular only at one of the times  $U_1, U_2, \dots$ , but note that not all of these times are caused by updates: some are caused because of the threshold  $S$ .

We first look at the probability of the event  $\{\tau'_G \leq U_k\} \cap \mathcal{S}[0, U_k)$ .

**Lemma 3.5.** *For all  $n$ , all  $k$  and all  $(x_0, \eta_0)$  with  $x_0 \in \mathcal{G}_0$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(\tau'_G \leq U_k, \mathcal{S}[0, U_k)) \leq k \max_{\eta_0, x_0 \in \mathcal{G}_0} \mathbb{P}_{x_0, \eta_0}(\tau'_G = U_1, \mathcal{S}[0, U_1)).$$

**Proof.** Consider any  $(x_0, \eta_0)$  satisfying  $x_0 \in \mathcal{G}_0$ . By the union bound and the strong Markov property (applied at time  $U_{j-1}$  for the  $j$ -th term of the sum), we find that

$$\begin{aligned} \mathbb{P}_{x_0, \eta_0}(\tau'_G \leq U_k, \mathcal{S}[0, U_k)) &\leq \sum_{j=1}^k \mathbb{P}_{x_0, \eta_0}(\tau'_G = U_j, \mathcal{S}[0, U_j)) \\ &\leq \max_{\eta_0, x_0 \in \mathcal{G}_0} \mathbb{P}(\tau'_G = U_1, \mathcal{S}[0, U_1)) \cdot \sum_{j=1}^k \mathbb{P}_{x_0, \eta_0}(\tau'_G > U_{j-1}, \mathcal{S}[0, U_{j-1})). \end{aligned}$$

Upper bounding the sum by  $k$  completes the proof.  $\square$

We now determine the *uniform* mixing time of the static walk on a good giant. For a Markov chain  $Z$  with transition matrix  $P$  and invariant distribution  $\pi$ , the *uniform mixing time* is

$$t_{\text{unif}}(\varepsilon, Z) = \inf\{t \geq 0 \mid \max_{x, y} |1 - p_t(x, y)/\pi(y)| \leq \varepsilon\}.$$

**Lemma 3.6.** *Let  $G$  be a good graph, and let  $\mathcal{G}$  be its (unique) giant. Consider the static walk, denoted  $\tilde{X}$ , on the giant. Write  $t_{\text{unif}}(\varepsilon, \tilde{X})$  for the  $\varepsilon$ -uniform mixing time of  $\tilde{X}$  (on  $\mathcal{G}$ ). Then*

$$t_{\text{unif}}(\frac{1}{8}, \tilde{X}) \lesssim n(\log n)^6.$$

**Proof.** To prove this lemma, we compare  $\tilde{X}$  with a ‘sped-up’ version. Consider a walk  $Z$  on the giant  $\mathcal{G}$  of a good graph  $\zeta$ . Write  $m = |\mathcal{G}|$ ; so  $m \asymp n$ . Write  $d(z)$  for the degree of  $z$  in  $\zeta$ , and  $d^*$  for the maximum degree; note that  $d^* \leq C_* \log n$ . Associate to a vertex  $z \in \mathcal{G}$  the following set:

$$\mathcal{V}_z = \mathcal{N}_z \cup \{1, \dots, k_z\} \setminus \{z\} \quad \text{where } k_z \text{ is such that } V_z = |\mathcal{V}_z| = 2C_* \log n,$$

where  $\mathcal{N}_z = \{z' \in \mathcal{G} \mid \zeta(z, z') = 1\}$  is the (open) neighbourhood of  $z$  (in  $\zeta$ ). (This is possible since  $d^* \leq C_* \log n$ .) Give  $Z$  a rate 1 jump clock: when this clock rings, if  $Z$  is at  $z$  then a vertex  $z'$  is chosen uniformly at random from  $\mathcal{V}_z$  and  $Z$  moves (from  $z$ ) to  $z'$  if and only if the edge  $(z, z')$  is open, ie  $\zeta(z, z') = 1$ . Observe that  $Z$  is the same as the static walk  $\tilde{X}$ , except that it is sped-up by a factor  $n/(2C_* \log n)$ . Hence the mixing times are in ratio  $n/(2C_* \log n)$ , for both total variation and uniform mixing. We now calculate the uniform mixing time  $t_{\text{mix}}(\frac{1}{8}, Z)$ .

Since  $V_z = |\mathcal{V}_z| = 2C_* \log n \geq 2d^*$  for all  $z \in \mathcal{G}$ , we see that the chain  $Z$  is *lazy* in the sense that if we discretise by its rate 1 jump clock then the resulting discrete-time chain is *lazy*, ie  $p(z, z) \geq \frac{1}{2}$  for all  $z \in \mathcal{G}$ . Moreover,

$$\pi_Z(z) = 1/|\mathcal{G}| = 1/m \quad \text{and} \quad p(z, z') = \frac{1}{2C_* \log n} \mathbf{1}(\zeta(z, z') = 1).$$



Hence  $Z$  is reversible. It is then known that

$$t_{\text{unif}}(\varepsilon, Z) \lesssim \Phi_*^{-2}(\log(1/\pi_{\min}) + \log(1/\varepsilon)),$$

where  $\Phi_* = \inf\{\Phi_S \mid \pi_Z(S) \leq \frac{1}{2}\}$  and  $\Phi_S = \sum_{x \in A, y \in B} \pi_Z(x)p_Z(x, y)/\pi_Z(S)$ ; for a proof of this, see [17]. For any set  $S \subseteq \mathcal{G}$ , we have

$$\Phi_S = \frac{1}{2C_* \log n} \cdot \frac{|\partial S|}{|S|} \geq \frac{1}{2C_* \log n} \cdot \frac{|\partial S|}{d(S)} = \frac{1}{2C_* \log n} \cdot \Phi'_S,$$

where the prime ( $\prime$ ) denotes that we are considering the corresponding quantity for the nearest-neighbour discrete-time random walk. But we know that  $\Phi'_* \gtrsim (\log n)^{-2}$  since the graph is good, and hence  $\Phi_* \gtrsim (\log n)^{-3}$ . Hence

$$t_{\text{unif}}(\frac{1}{8}, Z) \lesssim (\log n)^7, \quad \text{and hence} \quad t_{\text{unif}}(\frac{1}{8}, \tilde{X}) \lesssim n(\log n)^6. \quad \square$$

We now use this mixing of the static walk along with our static-dynamic coupling to determine where the dynamic walk is at the update times of  $\mathcal{P}$ .

**Lemma 3.7.** *There exists a constant  $C$ , so that, for all  $M$ , all  $n$  sufficiently large, we have*

$$\max_{\eta_0, x_0 \in \mathcal{G}_0} \mathbb{P}_{x_0, \eta_0}(\tau'_G = U_1, \mathcal{S}[0, U_1]) \leq C\mu S \log_{(M)} n.$$

**Proof.** For this whole proof, we only consider the first update time  $U_1$ ; as such, we drop the 1 from the subscript, just writing  $U$ . Also, we write  $\tilde{X}$  for the static walk on  $\eta_0$  (as above).

For the walk to leave the giant, we need the time  $U$  to be triggered by an update to  $\mathcal{P}$ , ie we need  $U < S$ . If this is the case, then the walk leaves the giant if and only if the update was caused by the closing of one of the removal edges which, given  $R_{U^-}(X_U)$ , is  $R_{U^-}(X_U)/\kappa$ .

Once the event that the set definitions succeed, the static and dynamic walks can be coupled on  $[0, U)$ , so  $R_{U^-}(X_U) = R_0(\tilde{X}_U)$ ; write  $\tilde{R}_U = R_0(\tilde{X}_U)$ . Hence

$$\mathbb{P}_{x_0, \eta_0}(\tau'_G = U, \mathcal{S}[0, U]) = \frac{1}{\kappa} \mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(U < S) \mathbf{1}(\mathcal{G}[0, U])). \quad (3.3)$$

We now set  $T = n(\log n)^7$ ; so  $S = T \log n \gg T$ . We decompose according to  $\{U < T\}$  or  $\{U \geq T\}$ . When  $U < T$  we use the trivial bound  $\tilde{R}_U \leq C_* \log n$  (which holds whenever the graph is good):

$$\mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(U < T) \mathbf{1}(\mathcal{G}[0, U])) \leq C_* \log n \cdot \mathbb{P}_{x_0, \eta_0}(U < T, \mathcal{G}[0, U]).$$

Since  $\mathcal{W}_0^{k+1} \subseteq \mathcal{W}_0^k$ , for all  $M$  and all  $n$  sufficiently large, we have

$$\begin{aligned} \mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(T \leq U < S) \mathbf{1}(\mathcal{G}[0, U])) &= \mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(T \leq U < S) \mathbf{1}(\tilde{X}_U \in \mathcal{W}_0^M) \mathbf{1}(\mathcal{G}[0, U])) \\ &+ \sum_{k=1}^{M-1} \mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(T \leq U < S) \mathbf{1}(\tilde{X}_U \in \mathcal{W}_0^k \setminus \mathcal{W}_0^{k+1}) \mathbf{1}(\mathcal{G}[0, U])). \end{aligned} \quad (3.4)$$

When  $\tilde{X}_U \in \mathcal{W}_0^k$ , we have (by definition)  $\tilde{R}_U \leq C_* \log_{(k)} n$ . Hence we have

$$\begin{aligned} &\mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(T \leq U < S) \mathbf{1}(\mathcal{G}[0, U])) \\ &\leq C_* \mathbb{P}(U < S, \mathcal{G}[0, U]) \cdot (\log_{(M)} n + \sum_{k=1}^{M-1} \log_{(k)} n \cdot \mathbb{P}(\tilde{X}_U \notin \mathcal{W}_0^{k+1} \mid T \leq U < S, \mathcal{G}[0, U])). \end{aligned}$$

What is crucial is that, on the event that the graph is good, the update *times* are independent of the evolution of the walk: since  $\mathcal{P}$  always, regardless of the number of edges seen by the walker, contains precisely  $\rho$  open edges and  $\rho n$  closed edges, the update rate is always  $\kappa\mu$ . Thus an equivalent way of realising  $(\mathcal{P}_t)_{t \in [0, U]}$  is the following. Define the processes  $(\mathcal{A}_r)_{r \geq 0}$  and  $(\mathcal{B}_r)_{r \geq 0}$  as in (3.1, 3.2), taking  $t = 0$ . Then sample independently  $V \sim \mathcal{E}(\kappa\mu)$ . At time  $V$  with probability  $q = (1-p)/(1-p+\lambda)$  choose an edge uniformly at random from  $\mathcal{A}_V$  and change its state from open to closed, and with probability  $1-q$  choose an edge uniformly at random from  $\mathcal{B}_V$  and change its state from closed to open. Then set  $\mathcal{P}_r = \mathcal{A}_r \cup \mathcal{B}_r$  for all  $r \in [0, V)$ . Finally, set  $U = V \wedge S$ .

Since  $T \gg n(\log n)^6$ , which is the uniform mixing time of the static walk on a good giant (Lemma 3.6), if  $U \geq T$  then  $\tilde{X}_U$  has (uniformly) mixed and so, since the invariant distribution of the static walk is uniform (on the giant), for all  $k \leq M$ , we have

$$\mathbb{P}_{x_0, \eta_0}(\tilde{X}_U \notin \mathcal{W}_0^{k+1} \mid T \leq U < S, \mathcal{G}[0, U]) \leq 2 |\mathcal{G}_0 \setminus \mathcal{W}_0^{k+1}| / |\mathcal{G}_0| \leq 2(\log_{(k)} n)^{-4}, \quad (3.5)$$

with the final inequality holding by definition of a good graph, and noting that conditioning on  $\mathcal{G}[0, U]$  has no effect on the static walk (since  $\eta_0 \in \mathcal{H} \subseteq \mathcal{G}$ ); here we have used crucially that the path  $(\tilde{X}_t)_{t \geq 0}$  is independent of  $U$ .

Also, for any  $s \geq 0$ , we have

$$\mathbb{P}_{x_0, \eta_0}(U < s, \mathcal{G}[0, U]) \leq \mathbb{P}(\mathcal{E}(\kappa\mu) \leq s) = 1 - e^{-\kappa\mu s} \leq \kappa\mu s.$$

Hence combining these inequalities, for all  $M$  and all  $n$  for sufficiently large, we have

$$\mathbb{P}_{x_0, \eta_0}(\tau'_G = U_1, \mathcal{S}[0, U_1]) \leq 2C_*\mu S \log_{(M)} n. \quad \square$$

Let  $K$  be the (random) index given by  $U_K \leq t < U_{K+1}$ . By monotonicity of  $\{\tau'_G \leq t\}$ , we have

$$\begin{aligned} \mathbb{P}_{x_0, \eta_0}(\tau'_G \leq t) &\leq \mathbb{P}_{x_0, \eta_0}(\tau'_G \leq U_k, \mathcal{S}[0, U_k]) + \mathbb{P}_{x_0, \eta_0}(K \geq k, \mathcal{S}[0, U_k]) \\ &\quad + \mathbb{P}_{x_0, \eta_0}(\mathcal{S}[0, U_k]^c, \mathcal{G}[0, U_k]) + \mathbb{P}_{x_0, \eta_0}(\mathcal{G}[0, U_k]^c). \end{aligned}$$

We have already dealt with the first term in the previous lemmas; we now just need to show that the two ‘remainder’ terms are sufficiently small. We do this now.

**Lemma 3.8.** *For all  $n$  sufficiently large, all  $t$  with  $\mu t \geq (\log n)^{-5}$  and all  $(x_0, \eta_0)$  with  $x_0 \in \mathcal{G}_0$ , for  $k = \lceil 5t/S \rceil$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(K \geq k, \mathcal{S}[0, U_k]) \leq n^{-5}.$$

**Proof.** If we were to not have the thresholding by  $S$ , then we would have  $K \sim \text{Po}(\kappa\mu t)$ . However, we do have the thresholding. Set  $\tilde{U}_0 = 0$ , and inductively define  $\tilde{U}_j$ , for  $j = 1, 2, \dots$ , by

$$\tilde{U}_j - \tilde{U}_{j-1} = S \cdot \mathbf{1}(U_j - U_{j-1} = S), \quad \text{and set } \tilde{K} = \inf\{k \geq 0 \mid \tilde{U}_k \leq t < \tilde{U}_{k+1}\}.$$

We have  $\tilde{U}_j \leq U_j$  for all  $j \geq 0$ , and thus  $\tilde{K} \geq K$ .

Recall that when the set definitions succeed,  $\{U_j - U_{j-1}\}_{j \geq 1}$  is a collection of iid random variables, and are independent of the starting point  $(x_0, \eta_0)$ . Note that we have

$$\mathbb{P}(\mathcal{E}(\kappa\mu) \geq S) = e^{-\kappa\mu S} = 1 - o(1) \geq \frac{1}{2},$$

by the assumption  $\mu n \ll (\log n)^{-19}$ . Also let us write  $k' = \lceil t/S \rceil$ ; since  $\mu t \geq (\log n)^{-5}$  and  $\mu n \ll (\log n)^{-14}$ , we have  $t/S \gg \log n$ , and so  $k' \gg 1$  and  $k' \leq 2t/S$ . Then, on the event that the set definitions succeed, we have  $\tilde{K} \preceq \text{Po}(4t/S)$ , since  $\mathbb{P}(\mathcal{E}(\kappa\mu) \geq S) \geq \frac{1}{2}$ . Hence

$$\mathbb{P}_{x_0, \eta_0}(K \geq k, \mathcal{S}[0, U_k]) \leq \mathbb{P}(\text{Po}(4t/S) \geq 5t/S) \leq \exp(-\frac{1}{10}t/S),$$

by Poisson concentration. Since  $t/S \gg \log n$ , we deduce our lemma.  $\square$

**Lemma 3.9.** *For all  $n$  sufficiently large, all  $k$  and all  $(x_0, \eta_0)$  with  $x_0 \in \mathcal{G}_0$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(\mathcal{S}[0, U_k]^c, \mathcal{G}[0, U_k]) \leq k \cdot \exp(-\frac{1}{3}C_*(\log n)^9).$$

**Proof.** By the union bound, we have

$$\mathbb{P}_{x_0, \eta_0}(\mathcal{S}[0, U_k]^c, \mathcal{G}[0, U_k]) \leq k \cdot \max_{x_0, \eta_0} \mathbb{P}_{x_0, \eta_0}(\mathcal{S}[0, U_1]^c, \mathcal{G}[0, U_1]).$$

Since we work on the event that the graph is good, we have at most  $C_* \log n$  removal edges for each vertex; we also have that there are  $\Omega(n)$  open edges and  $\Omega(n^2)$  closed edges. Hence the only part that can ‘go wrong’ in the definitions is if the number of open or closed edges seen since the

last update is too high. However, the maximum degree is at most  $C_* \log n$  and we walk for a time at most  $S = n(\log n)^8$ , so Poisson concentration will tell us that we do not see too many.

Consider a static walk  $\tilde{X}$  on a good graph  $\eta_0$ , starting from  $x_0 \in \mathcal{G}_0$  and run for a time  $S = n(\log n)^8$ . Let  $\alpha$  be the number of open edges seen in this time, and  $\beta$  the number of closed. Write  $N$  for the number of steps taken; by Poisson thinning, we have  $N \preceq \text{Po}(C_* S \log n/n)$ , and  $S \log n/n = (\log n)^9$ . On the event  $N \leq 2C_*(\log n)^9$ , we have  $\alpha \leq 2C_*^2(\log n)^{10} \ll \rho$  and  $\beta \leq 2C_* n(\log n)^9 \leq \rho n$ , as required for the set definitions to succeed. Hence

$$\mathbb{P}_{x_0, \eta_0}(\mathcal{S}[0, U_1]^c, \mathcal{G}[0, U_1]) \leq \mathbb{P}(\text{Po}(C_*(\log n)^9) > 2C_*(\log n)^9) \leq \exp(-\frac{1}{3}C_*(\log n)^9),$$

by Poisson concentration. The result now follows from the union bound given above.  $\square$

**Corollary 3.10.** *For all  $n$  sufficiently large, all  $t$  with  $\mu t \geq (\log n)^{-5}$  and all  $(x_0, \eta_0)$  with  $x_0 \in \mathcal{G}_0$ , for  $\mu \geq n^{-8}$  and  $k = \lceil 5t/S \rceil$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(\mathcal{S}[0, U_k]^c, \mathcal{G}[0, U_k]) \leq \mu t \cdot n^{-5}.$$

**Proof.** As in Lemma 3.8 for  $\lceil t/S \rceil$ , we have  $k \leq 6t/S$ . Hence

$$k \cdot \exp(-\frac{1}{3}C_*(\log n)^9) \leq 6S^{-1} \cdot \mu t \cdot n^8 \exp(-\frac{1}{3}C_*(\log n)^9) \leq \mu t \cdot n^{-5}. \quad \square$$

We now have all the ingredients to prove Proposition 3.2 for the case  $\mu \geq n^{-8}$ .

**Proof of Proposition 3.2** (when  $\mu \geq n^{-8}$ ). Fix  $M$ . Combining the above results, we have, for  $k = \lceil 5t/S \rceil$ , recalling that  $U_k \leq kS \leq 6t$  for the times  $t$  we are considering, that

$$\begin{aligned} \mathbb{P}_{x_0, \eta_0}(\tau'_G \leq t) &\leq \mathbb{P}_{x_0, \eta_0}(\tau'_G \leq U_k, \mathcal{G}[0, U_k]) + \mathbb{P}_{x_0, \eta_0}(K \geq \lceil 5t/S \rceil, \mathcal{S}[0, U_k]) \\ &\quad + \mathbb{P}_{x_0, \eta_0}(\mathcal{S}[0, U_k]^c, \mathcal{G}[0, U_k]) + \mathbb{P}_{x_0, \eta_0}(\mathcal{G}[0, 6t]^c) \\ &\leq C\mu t \log_{(M)} n + n^{-5} + \mu t n^{-5} + n^{-1} \leq 2C\mu t \log_{(M)} n \end{aligned} \quad (3.6)$$

since  $\eta_0 \in \mathcal{H}$ ,  $\mu \geq n^{-8}$  and  $(\log n)^{-5} \leq \mu t \leq \frac{1}{10}$ .  $\square$

It remains to prove the proposition in the case  $\mu \leq n^{-8}$ . In this case, ‘almost always’ the static walk mixes on the entire giant before any of the graph even refreshes; this will make this proof easier. The general idea will be very similar, particularly to Lemma 3.7.

**Proof of Proposition 3.2** (when  $\mu \leq n^{-8}$ ). Fix  $M$ . For this proof, let  $U_1, U_2, \dots$  be the refresh times of the graph; let  $U_0 = 0$ . Note then that  $U_j - U_{j-1} \sim^{\text{iid}} \mathcal{E}(\mu N)$  where  $N = \binom{n}{2} \leq n^2$ .

By our static-dynamic coupling, we may couple so that  $X_t = \tilde{X}_t$  for all  $t \leq U_1$ , where  $\tilde{X}$  is the static walk. We wish to see how many removal edges there are for  $X$  at time  $U_1$ . Dropping the subscript 1 (as previously), we have  $U \sim \mathcal{E}(\mu N)$ , independent of  $\tilde{X}$ .

Suppose  $\eta_0 \in \mathcal{G}$  and  $x_0 \in \mathcal{G}_0$ . Then, similarly to in (3.3), we have

$$\mathbb{P}_{x_0, \eta_0}(\tau'_G = U, \mathcal{G}[0, U]) = \frac{1}{N} \mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(U \geq n^2) \mathbf{1}(\mathcal{G}[0, U])) + \mathbb{P}_{x_0, \eta_0}(U \leq n^{-2}).$$

We know that  $\mathbb{P}(U \leq n^2) = \mathbb{P}(\mathcal{E}(\mu N) \leq n^2) \leq \mu n^2 N \leq n^{-4}$ . Similarly to in (3.4), we have

$$\begin{aligned} \mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(U \geq n^2) \mathbf{1}(\mathcal{G}[0, U])) &= \mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(U \geq n^2) \mathbf{1}(\tilde{X}_U \in \mathcal{W}_0^M) \mathbf{1}(\mathcal{G}[0, U])) \\ &\quad + \sum_{k=1}^{M-1} \mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(U \geq n^2) \mathbf{1}(\tilde{X}_U \in \mathcal{W}_0^k \setminus \mathcal{W}_0^{k+1}) \mathbf{1}(\mathcal{G}[0, U])). \end{aligned}$$

When  $\tilde{X}_U \in \mathcal{W}_0^{k-1}$ , we have (by definition)  $\tilde{R}_U \leq C_* \log_{(k-1)} n$ . Hence we have

$$\mathbb{E}_{x_0, \eta_0}(\tilde{R}_U \mathbf{1}(U \geq n^2) \mathbf{1}(\mathcal{G}[0, U])) \leq \log_{(M)} n + \sum_{k=1}^{M-1} \log_{(k)} n \cdot \mathbb{P}(\tilde{X}_U \notin \mathcal{W}_0^{k+1} \mid U \geq n^2, \mathcal{G}[0, U]).$$

Exactly as in (3.5), we use the (uniform) mixing of the static walk to obtain

$$\mathbb{P}_{x_0, \eta_0}(\tilde{X}_U \notin \mathcal{W}_0^{k+1} \mid U \geq n^2, \mathcal{G}[0, U]) \leq 2 |\mathcal{G}_0 \setminus \mathcal{W}_0^{k+1}| / |\mathcal{G}_0| \leq 2(\log_{(k)} n)^{-4}.$$

Hence combining these inequalities, for all  $M$  and all  $n$  sufficiently large, we have

$$\max_{x_0, \eta_0} \mathbb{P}_{x_0, \eta_0}(\tau'_G = U_1, \mathcal{G}[0, U_1]) \leq 2C_* \frac{1}{N} \log_{(M)} n.$$

As in Lemma 3.5 (except replacing  $\mathcal{G}$  by  $\mathcal{S}$ ), for all  $k \in \mathbb{N}$ , we have

$$\mathbb{P}_{x_0, \eta_0}(\tau'_G \leq U_k, \mathcal{G}[0, U_k]) \leq k \max_{x_0, \eta_0} \mathbb{P}_{x_0, \eta_0}(\tau'_G = U_1, \mathcal{G}[0, U_1]) \leq 2C_* k \frac{1}{N} \log_{(M)} n.$$

Observe that  $U_k \sim \Gamma(k, \mu N)$ ; let  $K$  be the (random) index given by  $U_K \leq t < U_{K+1}$ , and observe then that  $K \sim \text{Po}(\mu t N)$ . Set  $k = \lceil 2\mu t N \rceil$ . By the same arguments as used in (3.6) (except without the ‘set-definitions’ term) we have

$$\begin{aligned} \mathbb{P}_{x_0, \eta_0}(\tau'_G \leq t) &\leq \mathbb{P}_{x_0, \eta_0}(\tau'_G \leq U_k, \mathcal{G}[0, U_k]) + \mathbb{P}(K \geq 2\mu t N) + \mathbb{P}(U_k \geq 3k/(\mu N)) + \mathbb{P}_{\eta_0}(\mathcal{G}[0, 7t]^c) \\ &\leq 2C_* \mu t \log_{(M)} n + \exp(-\frac{1}{3}\mu t N) + \exp(-2\mu t N) + \mathbb{P}_{x_0, \eta_0}(\mathcal{G}[0, 7t]^c). \end{aligned}$$

We now recall that we restricted consideration of  $t$  to satisfy  $(\log n)^{-5} \leq \mu t \leq \frac{1}{10}$ ; note then that  $7t \leq 1/\mu$ . We also consider only  $\eta_0 \in \mathcal{H}$ ; as such, the graph remainder term in the final line above is at most  $1/n$ . Since  $N \asymp n^2$ , we see that the first term dominates, leaving us with

$$\mathbb{P}_{x_0, \eta_0}(\tau'_G \leq t) \leq C \mu t \log_{(M)} n \quad \text{for a constant } C. \quad \square$$

## 4 Isolation Times

In this section we prove two main results on isolation times. They will involve, respectively, a single random walker on a dynamical environment and two independent random walkers on the same dynamical environment. When considering just one walk, we write  $\tau_{\text{isol}}$  for the isolation time; eg for a dynamical percolation system  $(Z, \zeta)$  we write

$$\tau_{\text{isol}}^Z = \inf\{t \geq 0 \mid d_t^\zeta(Z_t) = 0\}.$$

When the context is clear, we omit the superscript, just writing  $\tau_{\text{isol}}$ ; similarly, when the context is clear we write  $d$  for the degree, rather than  $d^\zeta$ . When we consider two walks,  $X$  and  $Y$ , on the same system,  $\eta$ , we use superscript  $X$  or  $Y$  to indicate which walk we are referring to: define

$$\tau_{\text{isol}}^X = \inf\{t \geq 0 \mid d_t(X_t) = 0\} \quad \text{and} \quad \tau_{\text{isol}}^Y = \inf\{t \geq 0 \mid d_t(Y_t) = 0\}.$$

**Theorem 4.1** (Single-Walker Isolation Time). *For all  $M \in \mathbb{N}$ , all  $n$  sufficiently large and all pairs  $(x_0, \eta_0)$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > t, \mathcal{H}[0, t]) \leq 2 \exp(-\mu t (\log_{(M)} n)^{-1}).$$

Moreover, if  $\mu t \geq 3$ , then we may remove the pre-factor of 2.

Once we have proved this, we shall be able to use a type of concentration result to prove a bound on the isolation time of two independent random walkers on the same environment. When we are considering this, we write  $\mathbb{P}_{x_0, y_0, \eta_0}$  for the measure. For two walks  $X$  and  $Y$  on the same (dynamical) environment  $\eta$ , write

$$\tau = \inf\{t \geq 0 \mid d_t(X_t) = 0 = d_t(Y_t)\}.$$

**Theorem 4.2** (Dual-Walker Isolation Time). *For all  $M \in \mathbb{N}$ , all  $n$  sufficiently large and all triples  $(x_0, y_0, \eta_0)$ , we have*

$$\mathbb{P}_{x_0, y_0, \eta_0}(\tau > t, \mathcal{H}[0, t]) \leq 2 \exp(-\mu t (\log_{(M)} n)^{-1}).$$

## 4.1 Single-Walker Isolation Time

In this section we prove Theorem 4.1 on the isolation time of a walk  $X$  on a dynamical environment  $\eta$ . In order to find the isolation time, we wait until the walk joins the giant and then look at becoming isolated from there. It is easier to consider the giant, rather than subcritical components, because we are able to use concentration results on the structure of the giant.

We first state the proposition on isolation from the giant, and then show how to conclude Theorem 4.1 from it; we then prove the proposition to finish. Throughout,  $M$  is a positive integer.

**Proposition 4.3** (Isolation from the Giant). *There exists a positive constant  $c$  so that, for all  $M$ , all  $n$  sufficiently large and all  $(x_0, \eta_0)$  with  $\eta_0 \in \mathcal{H}$  and  $x_0 \in \mathcal{G}_0$ , we have*

$$\mathbb{P}_{x_0, \eta_0} \left( \tau_{\text{isol}} \leq \frac{1}{\mu \log_{(M)} n} \right) \geq c \cdot \frac{1}{\log_{(M)} n}.$$

**Proof of Theorem 4.1.** Observe that this trivially holds (for all  $n$  large enough) if  $\mu t \leq 3$ . By monotonicity, replacing  $M$  by  $M - 1$ , it suffices to prove an upper bound of  $\exp(-c\mu t / \log_{(M)} n)$  for a positive constant  $c$  when  $\mu t \geq 3$ .

Fix  $M$ . For this proof, rescale time so that  $\mu = 1$ . We prove this theorem by performing independent experiments. Note that if  $x_0 \in \mathcal{G}_0$  then  $\tau_{\mathcal{G}} = 0$ , and otherwise we apply Lemma 3.1. By direct calculation, we have

$$\begin{aligned} \mathbb{P}_{x_0, \eta_0} (\tau_{\text{isol}} \leq 2, \mathcal{H}[0, 1]) &\geq \mathbb{P}_{x_0, \eta_0} (\tau_{\text{isol}} \leq 2, \tau_{\mathcal{G}} \leq 1, \mathcal{H}[0, 1]) \\ &\geq \mathbb{P}_{x_0, \eta_0} (\tau_{\text{isol}} - \tau_{\mathcal{G}} \leq 1 / \log_{(M)} n, \tau_{\mathcal{G}} \leq 1, \mathcal{H}[0, 1]) \\ &= \sum_{x'_0, \eta'_0} \mathbb{P}_{x'_0, \eta'_0} (\tau_{\text{isol}} \leq 1 / \log_{(M)} n) \cdot \mathbb{P}_{x_0, \eta_0} (X(\tau_{\mathcal{G}}) = x'_0, \eta(\tau_{\mathcal{G}}) = \eta'_0, \mathcal{H}[0, 1]) \\ &\geq c(\log_{(M)} n)^{-1} \cdot \left( \mathbb{P}_{x_0, \eta_0} (\tau_{\mathcal{G}} \leq 1) - \mathbb{P}_{x_0, \eta_0} (\mathcal{H}[0, 1]^c) \right), \end{aligned}$$

for a positive constant  $c$ , where for the final inequality we used that on the event  $\mathcal{H}[0, 1]$  we have  $\eta'_0 \in \mathcal{H}$ , and hence we may apply Proposition 4.3. Now applying Lemma 3.1, we obtain

$$\mathbb{P}_{x_0, \eta_0} (\tau_{\text{isol}} \leq 2, \mathcal{H}[0, 1]) \geq \frac{1}{2} c c_1 / \log_{(M)} n - \mathbb{P}_{x_0, \eta_0} (\mathcal{H}[0, 1]^c),$$

with the positive constant  $c_1$  coming from Lemma 3.1, noting that  $\mathbb{P}_{x_0, \eta_0} (\mathcal{G}[0, 1]^c) = o(1)$  since  $\eta_0 \in \mathcal{H}$ . Rearranging this, we obtain, for a positive constant  $c$ , that

$$\mathbb{P}_{x_0, \eta_0} (\tau_{\text{isol}} > 2, \mathcal{H}[0, 2]) \leq \mathbb{P}_{x_0, \eta_0} (\tau_{\text{isol}} > 2, \mathcal{H}[0, 1]) \leq \exp(-c / \log_{(M)} n).$$

Hence, for any  $k \in \mathbb{N}$ , applying the strong Markov property ( $k - 1$  times), we obtain

$$\mathbb{P}_{x_0, \eta_0} (\tau_{\text{isol}} > 2k, \mathcal{H}[0, 2k]) \leq \max_{x'_0, \eta'_0} \mathbb{P}_{x'_0, \eta'_0} (\tau_{\text{isol}} > 2, \mathcal{H}[0, 2])^k \leq \exp(-ck / \log_{(M)} n).$$

This completes the proof.  $\square$

It remains to prove Proposition 4.3. We do this via a sequence of lemmas, using the following rough methodology. Observe that the *only* way for the walk to become isolated is to be at a degree 1 vertex and for the one open edge to close before any closed incident edges open or the walker leaves the vertex. This motivates looking at the rate at which the walk hits degree 1 vertices.

Since the walk is on a dynamically evolving graph, even though when we require the graph to be good this includes that the giant has a lot of degree 1 vertices, the location of these degree 1 vertices is changing. This makes using averaging properties (like a law of large numbers) difficult. However, since we take steps at rate at least  $1/n$  (when non-isolated) and  $\mu n \ll 1 / \log n$  (the order of the maximum degree), we see that the vast majority of the time the walker takes a step before any edge incident to its location changes state. This motivates looking at the rate at which a walk (with the same walk-dynamics) hits degree 1 vertices *on a static (good) graph*, and then relating this quantity to the relevant quantity for the walk on the dynamic graph. In §3 we referred to this as the *static walk* and the original as the *dynamic walk*, denoting them by  $\tilde{X}$  and  $X$ , respectively.

With this motivation in mind, we first collect some results regarding a walk with our dynamics on a static graph. To do this, we use a Chernoff-style bound on the number of visits to a set, which is due to Gillman [13]. It applies to *discrete-time* random walks. We do not apply it to a discretisation of our continuous chain, but to the jump chain of the walk (on a static graph). We state it in a general form; an even more general form is given in [13, Theorem 2.1].

**Theorem 4.4** (Gillman [13]). *Consider the discrete-time random walk on a weighted, connected graph  $G = (V, E)$  with any initial distribution. Let  $\pi$  be the unique invariant distribution. Let  $A \subseteq V$ , and let  $N_m$  be the number of visits to  $A$  in  $m$  steps. Write  $\gamma$  for the spectral gap. Then*

$$\mathbb{P}(|N_m - m\pi(A)| \geq \varepsilon) \leq 3\pi_{\min}^{-1/2} \exp(-\frac{1}{20}\gamma\varepsilon^2/m) \quad \text{for any } \varepsilon \in [0, m].$$

We now apply this to a walk on a good (static) giant.

**Lemma 4.5.** *Consider the discrete-time nearest-neighbour simple random walk on a graph  $G$ , and write  $N_m$  for the number of visits to the set of degree 1 vertices in  $m$  steps. There exists a positive constant  $c$  so that, for all  $n$  sufficiently large and all  $m \geq (\log n)^6$ , if the graph is good, ie  $G \in \mathcal{G}$ , and the walk starts from its giant, then we have*

$$\mathbb{P}(N_m \leq cm) \leq n^{-1}.$$

**Proof.** Note that the invariant measure of this walk, which we denote  $\pi'$ , is given by  $\pi_i = d_i/d_G$ , where  $d_G = \sum_{i \in \mathcal{G}} d_i$ . Since  $d_i \geq 1$  for all  $i \in \mathcal{G}$ , we have  $\pi'_{\min} \geq 1/d_G$ . Now, trivially we have that  $d_G \leq d_{\mathcal{G}}$ , where  $d_{\mathcal{G}} = \sum_{i \in G} d_i$ , and  $d_{\mathcal{G}} \leq 2C_*n$  by Definition 2.2(iii). Hence  $1/\pi'_{\min} \leq d_G \leq 2C_*n$ .

Let  $A = \{x \in \mathcal{G} \mid d(x) = 1\}$  be the set of degree 1 vertices in the giant. Definition 2.2(iv) tells us that  $|A| \geq c_*n$ . Thus, since  $d(x) \geq 1$  for all  $x \in \mathcal{G}$  and  $d_G \leq 2C_*n$ , we have that  $\pi'(A) \geq c_*/(2C_*)$ ; let  $c = c_*/(4C_*)$  so that  $\pi'(A) \geq 2c$ .

Recall from Definition 2.2(vii) that the spectral gap  $\gamma$  of a good giant satisfies  $\gamma \geq c_*(\log n)^{-4}$ . We now take  $\varepsilon = \frac{1}{2}\pi'(A)m \leq m$  in Theorem 4.4 to obtain

$$\mathbb{P}(|N_m - m\pi'(A)| \geq \frac{1}{2}\pi'(A)m) \leq 3\sqrt{2c_*} \cdot n^{-1/2} \exp(-\frac{1}{20}c_*(\log n)^{-4} \cdot \frac{1}{4}\pi'(A)^2m).$$

Since  $\pi'(A) \geq 2c$ , taking  $m \geq (\log n)^6$  gives super-polynomial decay, completing the proof.  $\square$

We now make rigorous the motivation given at the start of this section in the following lemma.

**Lemma 4.6.** *There exists a positive constant  $q$  so that, for all  $n$  sufficiently large and all  $(x_0, \eta_0)$  with  $\eta_0 \in \mathcal{G}$  and  $x_0 \in \mathcal{G}_0$ , for  $s = n(\log n)^6$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} \leq s) \geq q\mu s.$$

**Remark.** Observe that, as a best-case scenario, if the walker were always at a degree 1 vertex until it becomes isolated, then the isolation time would simply be the time it takes for that one edge to close, which is  $\mathcal{E}(\mu(1-p))$ . Thus, for any  $(x_0, \eta_0)$  with  $d_0(x_0) \neq 0$ , we have

$$\mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} \leq s) \leq \mu s / (1-p).$$

Hence, for this  $s$ , such a result as Lemma 4.6 is best-possible up to constants.  $\triangle$

**Proof of Lemma 4.6.** Fix a pair  $(x_0, \eta_0)$  with  $\eta_0 \in \mathcal{G}$  and  $x_0 \in \mathcal{G}_0$ ; for this proof, drop it from the notation, writing  $\mathbb{P}(\cdot)$  in place of  $\mathbb{P}_{x_0, \eta_0}(\cdot)$ .

Lemma 4.5 tells us the rate at which the static walk hits degree 1 vertices (with high probability). In order to transfer this result to our dynamic walk, we define a coupling between the two walks; this was given in §3, but we recall it precisely here. Write  $\tilde{X}$  for the static walk, walking on the static graph  $\eta_0$ . Set  $X_0 = \tilde{X}_0 = x_0$ . Give  $X$  and  $\tilde{X}$  the same jump clock. When the clock rings, at time  $t$  say, both walks choose the same vertex;  $\tilde{X}$  performs the jump if and only if the edge is present in  $\eta_0$ , while  $X$  performs the jump if and only if the edge is present in  $\eta_t$ . We call this the *static-dynamic coupling*.

We now define  $\tilde{T}_i$  to be the  $i$ -th time that (the static walk)  $\tilde{X}$  hits a degree 1 vertex: set  $\tilde{T}_0 = \tilde{T}'_0 = 0$  and define inductively, for  $i \geq 1$ ,

$$\tilde{T}_i = \inf\{t \geq \tilde{T}'_{i-1} \mid d_t(\tilde{X}_t) = 1\} \quad \text{and} \quad \tilde{T}'_i = \inf\{t \geq \tilde{T}_i \mid \tilde{X}(t) \neq \tilde{X}(\tilde{T}_i)\}.$$

Since the jump rate of  $\tilde{X}$  is always at least  $1/n$ , by standard Poisson concentration it takes at least  $s/(2n)$  steps in time  $s - n$  with probability  $1 - o(1)$ . Along with Lemma 4.5 this says that

$$\mathbb{P}(\tilde{T}_k \leq s - n) = 1 - o(1) \quad \text{for} \quad k = s/(8n). \quad (4.1)$$

For  $i \geq 1$  define the event that all the (open or closed) edges incident to a vertex that the static walk visited remain in the same state between visits to degree 1 vertices:

$$\mathcal{E}_i = \{\text{neighbourhood of path of static walk did not change in } [\tilde{T}_{i-1}, \tilde{T}_i]\}.$$

Similarly, for  $u \geq v \geq 0$  define

$$\mathcal{E}_{u,v} = \{\text{neighbourhood of path of static walk did not change in } [u, v]\}.$$

Note that, by definition, on the event  $\{\tilde{T}_k \leq s\}$  we have  $\mathcal{E}_1 \cap \dots \cap \mathcal{E}_k \supseteq \mathcal{E}_{0,s}$ .

Write  $N$  for the number of steps taken by  $\tilde{X}$  in time  $s$ . Since  $\eta_0 \in \mathcal{G}$ , the maximum degree is at most  $C_* \log n$ . Hence, by Poisson thinning,  $N \preceq L \sim \text{Po}(C_* s \log n/n)$ . Let  $\alpha$  and  $\beta$  be the total number of open and closed edges, respectively, that are adjacent to the path of the (static) walk by time  $s$ . When  $N \leq 2C_* s \log n/n$ , we have  $\alpha \leq 2C_*^2 s (\log n)^2/n$  and  $\beta \leq 2C_* s \log n$ . Hence

$$\begin{aligned} \mathbb{P}(\mathcal{E}_{0,s}^c) &\leq \mathbb{P}(\mathcal{E}(\alpha(1-p)\mu + \beta p\mu) \leq s, N \leq 2C_* s \log n/n) + \mathbb{P}(L > 2C_* s \log n/n) \\ &\leq Cs^2(\log n)^2\mu/n + \exp(-cs \log n/n), \end{aligned}$$

for positive constants  $c$  and  $C$ , by Poisson concentration. Hence, since  $\mu n \ll (\log n)^{-14}$ , we have

$$\mathbb{P}(\mathcal{E}_{0,s}) = 1 - o(1) \quad \text{when} \quad s = n(\log n)^6. \quad (4.2)$$

Note that  $X$  can only become isolated when it is at a degree 1 vertex immediately prior. We use our static-dynamic coupling to lower bound:

$$\begin{aligned} \mathbb{P}(\tau_{\text{isol}} \leq s) &\geq \mathbb{P}\left(\bigcup_{i=1}^k \{\tau_{\text{isol}} \in [\tilde{T}_i, \tilde{T}'_i), \tilde{T}_i \leq s - n, \tau_{\text{isol}} - \tilde{T}_i \leq n, \cap_{j \leq i} \mathcal{E}_j\}\right) \\ &= \sum_{i=1}^k \mathbb{P}(\tau_{\text{isol}} \in [\tilde{T}_i, \tilde{T}'_i), \tau_{\text{isol}} - \tilde{T}_i \leq n \mid \tilde{T}_i \leq s - n, \cap_{j \leq i} \mathcal{E}_j) \cdot \mathbb{P}(\tilde{T}_i \leq s - n, \cap_{j \leq i} \mathcal{E}_j). \end{aligned} \quad (4.3)$$

Using the static-dynamic coupling on the event  $\mathcal{E}_1 \cap \dots \cap \mathcal{E}_i$ , we see that if the unique open edge adjacent to  $X$  at time  $\tilde{T}_i$  closes before anything else opens or  $X$  jumps, then  $X$  becomes isolated during  $[\tilde{T}_i, \tilde{T}'_i)$ . Writing  $\mathcal{E}_1, \mathcal{E}_2$  and  $\mathcal{E}_3$  for independent exponential random variables, we have

$$\begin{aligned} \mathbb{P}(\tau_{\text{isol}} \in [\tilde{T}_i, \tilde{T}'_i), \tau_{\text{isol}} - \tilde{T}_i \leq n \mid \tilde{T}_i \leq s - n, \cap_{j \leq i} \mathcal{E}_j) \\ \geq \mathbb{P}(\mathcal{E}_1((1-p)\mu) < \min\{\mathcal{E}_2(p(n-1)\mu), \mathcal{E}_3(1/n)\}, \mathcal{E}_1((1-p)\mu) \leq n) \asymp \mu n, \end{aligned}$$

since  $\mu n \ll 1$ , by comparing rates. Using this in (4.3) along with (4.1) and (4.2) we obtain

$$\begin{aligned} \mathbb{P}(\tau_{\text{isol}} \leq s) &\gtrsim \mu n k \mathbb{P}(\tilde{T}_k \leq s - n, \mathcal{E}_1 \cap \dots \cap \mathcal{E}_k) \\ &\geq \mu n k (1 - \mathbb{P}(\mathcal{E}_{0,s}^c) - \mathbb{P}(\tilde{T}_k > s - n)) \asymp \mu n k. \end{aligned}$$

Since  $k = s/(8n)$ , this concludes the proof.  $\square$

We now use this to prove our isolation result Proposition 4.3.

**Proof of Proposition 4.3.** In this proof, we use the following shorthand:

$$\mathbb{P}_{\mathcal{G}}(\cdot) = \min_{\eta_0 \in \mathcal{G}, x_0 \in \mathcal{G}_0} \mathbb{P}_{x_0, \eta_0}(\cdot) \quad \text{and} \quad \mathbb{P}^{\mathcal{G}}(\cdot) = \max_{\eta_0 \in \mathcal{G}, x_0 \in \mathcal{G}_0} \mathbb{P}_{x_0, \eta_0}(\cdot).$$



Consider an initial pair  $(x_0, \eta_0)$  with  $\eta_0 \in \mathcal{H}$  and  $x_0 \in \mathcal{G}_0$ . For any  $s \in \mathbb{R}$  and  $r \in \mathbb{N}$ , using the Markov property we have

$$\begin{aligned} & \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} \in (sr, s(r+1)], \mathcal{G}[0, sr]) \\ & \geq \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} \leq s(r+1) \mid \tau_{\text{isol}} > sr, X_{sr} \in \mathcal{G}_{sr}, \mathcal{G}[0, sr]) \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > sr, X_{sr} \in \mathcal{G}_{sr}, \mathcal{G}[0, sr]) \\ & \geq \mathbb{P}_{\mathcal{G}}(\tau_{\text{isol}} \leq s) \cdot \left( \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > sr, \mathcal{G}[0, sr]) - \mathbb{P}_{x_0, \eta_0}(\exists u \leq sr \text{ s.t. } X_u \notin \mathcal{G}_u) \right). \end{aligned}$$

Hence we have

$$\begin{aligned} & \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > s(r+1), \mathcal{G}[0, sr]) \\ & = \mathbb{P}_{x_0, \eta_0}(\mathcal{G}[0, sr]) - \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} \leq sr, \mathcal{G}[0, sr]) - \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} \in (sr, s(r+1)], \mathcal{G}[0, sr]) \\ & \leq \mathbb{P}_{x_0, \eta_0}(\mathcal{G}[0, sr]) - \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} \leq sr, \mathcal{G}[0, sr]) \\ & \quad - \mathbb{P}_{\mathcal{G}}(\tau_{\text{isol}} \leq s) \cdot \left( \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > sr, \mathcal{G}[0, sr]) - \mathbb{P}_{x_0, \eta_0}(\exists u \leq sr \text{ s.t. } X_u \notin \mathcal{G}_u) \right) \\ & = \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > sr, \mathcal{G}[0, sr]) \cdot \mathbb{P}^{\mathcal{G}}(\tau_{\text{isol}} > s) \\ & \quad + \mathbb{P}_{\mathcal{G}}(\tau_{\text{isol}} \leq s) \cdot \mathbb{P}_{x_0, \eta_0}(\exists u \leq sr \text{ s.t. } X_u \notin \mathcal{G}_u). \end{aligned}$$

Hence, upon iterating, we obtain

$$\begin{aligned} \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > sr, \mathcal{G}[0, sr]) & \leq \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > sr, \mathcal{G}[0, s(r-1)]) \\ & \leq \mathbb{P}^{\mathcal{G}}(\tau_{\text{isol}} > s)^r + r \cdot \mathbb{P}_{\mathcal{G}}(\tau_{\text{isol}} \leq s) \mathbb{P}_{x_0, \eta_0}(\exists u \leq sr \text{ s.t. } X_u \notin \mathcal{G}_u). \end{aligned} \quad (4.4)$$

Observe that, by the memoryless property, we have

$$\mathbb{P}_{\mathcal{G}}(\tau_{\text{isol}} \leq s) \leq \mu s. \quad (4.5)$$

We now set  $s = n(\log n)^6$ ,  $t = \gamma(\log_{(M)} n)^{-1}/\mu$  for a constant  $\gamma$ , to be chosen later, and  $r = \lfloor t/s \rfloor$ ; note then that  $\frac{2}{3}t \leq rs \leq t$  as  $\mu n \ll (\log n)^{-5}$ . Since  $\eta_0 \in \mathcal{H}$ , we may apply Proposition 3.2 for this  $t$  to obtain a constant  $C$  so that

$$\mathbb{P}_{x_0, \eta_0}(\exists u \leq sr \text{ s.t. } X_u \notin \mathcal{G}_u) \leq C\mu t \log_{(M)} n. \quad (4.6)$$

Using (4.5) and (4.6) along with Lemma 4.6 in (4.4), we find that

$$\mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > t, \mathcal{G}[0, t]) \leq (1 - q\mu s)^r + C(\mu t)^2 \log_{(M)} n \leq 1 - \frac{1}{3}q\mu t + C(\mu t)^2 \log_{(M)} n,$$

valid for any  $(x_0, \eta_0)$  with  $\eta_0 \in \mathcal{H}$  and  $x_0 \in \mathcal{G}_0$ . We then take  $\gamma = q/(6C)$  and obtain

$$\mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > t, \mathcal{G}[0, t]) \leq 1 - \frac{q^2}{36C} \cdot \frac{1}{\log_{(M)} n}.$$

Since we can take  $C \geq 1$  and  $q \leq 1$ , we then have

$$\begin{aligned} & \mathbb{P}_{x_0, \eta_0} \left( \tau_{\text{isol}} > \frac{1}{C\mu \log_{(M)} n}, \mathcal{G} \left[ 0, \frac{1}{C\mu \log_{(M)} n} \right] \right) \\ & \leq \mathbb{P}_{x_0, \eta_0} \left( \tau_{\text{isol}} > \frac{q}{3C\mu \log_{(M)} n}, \mathcal{G} \left[ 0, \frac{q}{3C\mu \log_{(M)} n} \right] \right). \end{aligned}$$

Hence there exists a positive constant  $c$  so that

$$\mathbb{P}_{x_0, \eta_0} \left( \tau_{\text{isol}} > \frac{1}{\mu \log_{(M)} n}, \mathcal{G} \left[ 0, \frac{1}{\mu \log_{(M)} n} \right] \right) \leq 1 - c \cdot \frac{1}{\log_{(M)} n}.$$

Finally,  $\eta_0 \in \mathcal{H}$ , so  $\mathbb{P}_{\eta_0}(\mathcal{G}[0, (\log_{(M)} n)^{-1}/\mu]^c) \leq n^{-1}$ , and the result follows.  $\square$

**Remark.** Observe that, as in the remark after Lemma 4.6, for this time-scale the result of Proposition 4.3 is best-possible, up to constants.  $\triangle$

## 4.2 Dual-Walker (Joint) Isolation Time

In this section we prove Theorem 4.2 on the joint isolation time of two walkers on a single dynamical environment. We start by introducing some more notation. Consider two walks  $X$  and  $Y$ , which start from  $x_0$  and  $y_0$  respectively, walking *independently* on the same environment  $\eta$ . Let

$$\tau = \inf\{t \geq 0 \mid d_t(X_t) = 0 = d_t(Y_t)\}.$$

Let  $\tau_0^X = \tau_0^Y = \hat{\tau}_0^X = \hat{\tau}_0^Y = 0$ , and for  $k \geq 1$  define inductively

$$\begin{aligned} \tau_k^X &= \inf\{t \geq \hat{\tau}_{k-1}^X \mid d_t(X_t) = 0\}, & \tau_k^Y &= \inf\{t \geq \tau_k^X \mid d_t(Y_t) = 0\}, \\ \hat{\tau}_k^X &= \inf\{t \geq \tau_k^X \mid d_t(X_t) \neq 0\}, & \hat{\tau}_k^Y &= \inf\{t \geq \tau_k^Y \mid d_t(Y_t) \neq 0\}. \end{aligned}$$

We prove a result on the joint-isolation time of two walks,  $X$  and  $Y$ , walking independently on the same (dynamic) environment  $\eta$ . For the probability measure associated to this system  $(X, Y, \eta)$ , when it is started from  $(x_0, y_0, \eta_0)$ , we write  $\mathbb{P}_{x_0, y_0, \eta_0}$ .

In order to prove the dual-walker isolation result, we first state two lemmas that we use. We prove the theorem using the lemmas, then prove the lemmas. Throughout,  $M$  is a positive integer.

**Lemma 4.7.** *There exists a positive constant  $c_1$  so that, for all  $M$ , all  $n$  sufficiently large and all  $(x_0, y_0, \eta_0)$ , we have*

$$\mathbb{P}_{x_0, y_0, \eta_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[\tau_1^X, \hat{\tau}_1^X]) \leq \exp(-c_1 / \log_{(M)} n).$$

**Lemma 4.8.** *There exists a positive constant  $c_2$  so that, for all  $M$ , all  $n$  sufficiently large and all  $(x_0, \eta_0)$ , we have*

$$\mathbb{P}_{x_0, \eta_0}(\tau_{K+1}^X > t, \mathcal{H}[0, t]) \leq \exp(-\frac{2}{3}K) \quad \text{when } K = \lfloor c_2 \mu t / \log_{(M)} n \rfloor.$$

**Proof of Theorem 4.2.** By monotonicity, replacing  $M$  by  $M - 1$ , it suffices to find a positive constant  $c$  so that the probability is upper bounded by  $2 \exp(-c \mu t / (\log_{(M)} n)^2)$  for a positive constant  $c$ . Hence we may assume that  $\mu t \geq (\log_{(M)} n)^2$ , as otherwise the result trivially holds.

Fix  $M$ . Recall that we may assume  $\mu t \geq (\log_{(M)} n)^2$  for any constant  $C$  to be chosen later, otherwise the result is trivial.

For any  $t \geq 0$  and for  $K = \lfloor c_2 \mu t / \log_{(M)} n \rfloor$ , using Lemma 4.8 we have

$$\mathbb{P}_{x_0, y_0, \eta_0}(\tau > t, \mathcal{H}[0, t]) \leq \mathbb{P}_{x_0, y_0, \eta_0}(\tau > t, \tau_{K+1}^X \leq t, \mathcal{H}[0, t]) + \exp(-\frac{2}{3}K). \quad (4.7)$$

Since  $\tau_{K+1}^X \leq t$  implies  $\hat{\tau}_K^X \leq t$ , on the event  $\{\tau_{K+1}^X \leq t\}$  we have  $\mathcal{H}[0, t] \subseteq \mathcal{H}[0, \hat{\tau}_K^X]$ . We then use the strong Markov property at time  $\hat{\tau}_1^X$  to iterate:

$$\begin{aligned} \mathbb{P}_{x_0, y_0, \eta_0}(\tau > t, \tau_{K+1}^X \leq t, \mathcal{H}[0, t]) &\leq \mathbb{P}_{x_0, y_0, \eta_0}(\cap_{k=1}^K \{\tau_k^Y > \hat{\tau}_k^X\}, \mathcal{H}[0, \hat{\tau}_K^X]) \\ &\leq \mathbb{P}_{x_0, y_0, \eta_0}(\cap_{k=2}^K \{\tau_k^Y > \hat{\tau}_k^X\}, \mathcal{H}[\hat{\tau}_1^X, \hat{\tau}_K^X] \mid \tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[0, \hat{\tau}_1^X]) \\ &\quad \cdot \mathbb{P}_{x_0, y_0, \eta_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[\tau_1^X, \hat{\tau}_1^X]) \\ &\leq \max_{x'_0, y'_0, \eta'_0} \mathbb{P}_{x'_0, y'_0, \eta'_0}(\cap_{k=1}^{K-1} \{\tau_k^Y > \hat{\tau}_k^X\}, \mathcal{H}[0, \hat{\tau}_{K-1}^X]) \cdot \mathbb{P}_{x_0, y_0, \eta_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[\tau_1^X, \hat{\tau}_1^X]) \\ &\leq \dots \leq \max_{x'_0, y'_0, \eta'_0} \mathbb{P}_{x'_0, y'_0, \eta'_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[\tau_1^X, \hat{\tau}_1^X])^K. \end{aligned} \quad (4.8)$$

Since  $\mu t \geq C \log_{(M)} n$ , we have  $K \geq \frac{1}{2} c_2 \mu t / \log_{(M)} n$ . Using (4.8) and Lemma 4.7 in (4.7), we have

$$\begin{aligned} \mathbb{P}_{x_0, y_0, \eta_0}(\tau > t, \mathcal{H}[0, t]) &\leq \max_{x'_0, y'_0, \eta'_0} \mathbb{P}_{x'_0, y'_0, \eta'_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[\tau_1^X, \hat{\tau}_1^X])^K + \exp(-\frac{2}{3}K) \\ &\leq \exp(-c_1 K / \log_{(M)} n) + \exp(-\frac{2}{3}K) \leq 2 \exp(-\frac{1}{2} c_1 c_2 \mu t / (\log_{(M)} n)^2). \quad \square \end{aligned}$$

It remains to prove Lemmas 4.7 and 4.8.

**Proof of Lemma 4.7.** By the strong Markov property, used in the same way as above, and recalling that  $\tau_1^Y$  is the first time after  $\tau_1^X$  that  $Y$  becomes isolated, we have

$$\max_{x_0, y_0, \eta_0} \mathbb{P}_{x_0, y_0, \eta_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[\tau_1^X, \hat{\tau}_1^X]) \leq \max_{\substack{x_0, y_0, \eta_0 \\ d_0(x_0)=0}} \mathbb{P}_{x_0, y_0, \eta_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[0, \hat{\tau}_1^X]). \quad (4.9)$$

For the moment, we emphasise that our underlying graph has  $n$  vertices: we do this by using super- and subscript  $n$ , eg  $\mathbb{P}^n$  and  $\mathcal{G}_n$ . Recall Theorem 4.1, which says that

$$\mathbb{P}_{y_0, \eta_0}^n(\tau_{\text{isol}}^Y > t, \mathcal{H}[0, t]) \leq \exp(-\mu t / \log_{(M)} n) \quad \text{when } \mu t \geq 3.$$

We wish to bound (the related quantity)

$$\mathbb{P}_{x_0, y_0, \eta_0}^n(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[0, \hat{\tau}_1^X] \mid \hat{\tau}_1^X).$$

This is trivially 0 for  $x_0 = y_0$ ; consider  $x_0 \neq y_0$ . To bound this, we observe that, conditional on the value of  $\hat{\tau}_1^X = T$ , this is conditioning the vertex  $x_0$  to be isolated until time  $\hat{\tau}_1^X = T$ ; the rest of the graph is unaffected.

Let  $\eta^n$  be a dynamical environment on  $n$  vertices, and let  $x_0 \in \{1, \dots, n\}$  be a vertex. Define  $\tilde{\eta}^n$  by conditioning on the event that the vertex  $x_0$  is isolated until time  $T$ . Write  $\hat{\eta}^n$  for the restriction of  $\tilde{\eta}^n$  to  $\{1, \dots, n\} \setminus \{x_0\}$ . Observe then that  $\hat{\eta}^n \sim \eta^{n-1}$  (where  $\eta^{n-1}$  is a dynamical environment on  $n-1$  vertices), up to relabelling of vertices. (In words, this says that if a vertex is conditioned to be isolated, then the rest of the graph behaves as a dynamical environment on  $n-1$  vertices.) Note also that two edges do not update at the same time, so we cannot have  $\tau_1^Y = \hat{\tau}_1^X$  (since the first requires an edge to close and the second an edge to open).

Hence,  $(Y_t \mid t \leq \hat{\tau}_1^X)$  is a walk on the environment  $\hat{\eta}^n$ , which has the distribution of  $\eta^{n-1}$ . Note that  $Y$  may still pick the (conditioned to be isolated) vertex  $x_0$  (with probability  $1/(n-1)$ ), in which case it does not move; thus, under this conditioning,  $(Y, \hat{\eta}^n)$  is simply a realisation of dynamical percolation on  $n-1$  vertices, but with added laziness: when  $Y$ 's  $\mathcal{E}(1)$  clock rings, with probability  $1/(n-1)$  it does nothing; with the remaining probability, it performs the usual step.

Note that we can rescale  $\mu$  to get rid of the laziness of  $Y$ . Indeed, the laziness has the effect of changing the walker's clock from rate 1 to rate  $1 - 1/(n-1)$ . As such, if we replace  $\mu$  by  $\mu' = \mu(1 - 1/(n-1))$ , then the ratio of the rate edge-clocks to the rate of the walker-clock is  $\mu$ : we have simply slowed both down. We then speed up everything by a factor  $1 - 1/(n-1)$ . We apply previous results with  $\mu$  replaced by  $\mu'$ . The restrictions on  $\mu$  are satisfied by  $\mu'$  also, since  $\mu' \leq \mu$ .

For all  $m$ , define  $\mathcal{G}'_m$  by replacing  $c_*$  and  $C_*$  in Definition 2.2 by  $\frac{1}{2}c_*$  and  $2C_*$ , respectively; define  $\mathcal{H}'$  in terms of  $\mathcal{G}'$  as in Definition 2.4. We then have that if  $\eta_0^n \in \mathcal{G}_n$  and the vertex  $x_0$  is isolated (in  $\eta_0^n$ ), then  $\tilde{\eta}_0^n = \eta_0^n - \{x_0\}$  defined by removing the vertex  $x_0$  satisfies  $\tilde{\eta}_0^n \in \mathcal{G}'_{n-1}$  (for  $n$  sufficiently large). Hence we have the following inequality: let  $(Z, \zeta)$  be a full system, independent of  $X$  and  $Y$ , on  $n-1$  vertices, and start it from  $(Z_0, \zeta_0) = (y_0, \hat{\eta}_0^n)$ ; we then have

$$\mathbb{P}_{x_0, y_0, \eta_0^n}^n(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}_n[0, \hat{\tau}_1^X] \mid \hat{\tau}_1^X) \leq \mathbb{P}_{y_0, \hat{\eta}_0^n}^{n-1}(\tau_1^Z > \hat{\tau}_1^X, \mathcal{H}'_{n-1}[0, \hat{\tau}_1^X] \mid \hat{\tau}_1^X).$$

Note that Theorem 4.1 still holds if we replace  $\mathcal{H}$  by  $\mathcal{H}'$  in its statement. Combining all the above considerations, applying Theorem 4.1, on the event  $\{\mu \hat{\tau}_1^X \geq 3\}$  we have

$$\begin{aligned} \mathbb{P}_{x_0, y_0, \eta_0^n}^n(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}_n[0, \hat{\tau}_1^X] \mid \hat{\tau}_1^X) &\leq \mathbb{P}_{y_0, \hat{\eta}_0^n}^{n-1}(\tau_1^Z > \hat{\tau}_1^X, \mathcal{H}'_{n-1}[0, \hat{\tau}_1^X] \mid \hat{\tau}_1^X) \\ &\leq \exp(-\mu \hat{\tau}_1^X (1 - \frac{1}{n}) / \log_{(M)}(n-1)) \leq \exp(-\frac{1}{2} \mu \hat{\tau}_1^X / \log_{(M)} n). \end{aligned}$$

We now calculate the unconditioned value. Fix  $(x_0, y_0, \eta_0)$  with  $d_0(x_0) = 0$ . We have

$$\begin{aligned} &\mathbb{P}_{x_0, y_0, \eta_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[0, \hat{\tau}_1^X] \mid \hat{\tau}_1^X) \\ &\leq \mathbb{E}_{x_0, y_0, \eta_0}(\mathbb{P}_{x_0, y_0, \eta_0}(\tau_1^Y > \hat{\tau}_1^X, \mathcal{H}[0, \hat{\tau}_1^X] \mid \hat{\tau}_1^X) \cdot \mathbf{1}(\hat{\tau}_1^X \geq 3/\mu)) + \mathbb{P}_{x_0, \eta_0}(\hat{\tau}_1^X < 3/\mu) \\ &\leq \mathbb{P}_{x_0, \eta_0}(\hat{\tau}_1^X \geq 3/\mu) \cdot \exp(-\frac{1}{2} \mu (3/\mu) / \log_{(M)} n) + \mathbb{P}_{x_0, \eta_0}(\hat{\tau}_1^X < 3/\mu) \\ &\leq 1 - c \mathbb{P}_{x_0, \eta_0}(\hat{\tau}_1^X < 3/\mu) / \log_{(M)} n \leq \exp(-c \mathbb{P}_{x_0, \eta_0}(\hat{\tau}_1^X < 3/\mu) / \log_{(M)} n) \end{aligned}$$

for a positive constant  $c$ . Since  $d_0(x_0) = 0$ , we have  $\hat{\tau}_1^X \sim \mathcal{E}(\lambda \mu (1 - 1/n))$ , and hence we have  $\mathbb{P}_{x_0, \eta_0}(\hat{\tau}_1^X \geq 3/\mu) \asymp 1$ . Substituting this into (4.9) gives the required bound.  $\square$

**Proof of Lemma 4.8.** We may assume that  $K \geq 1$ , otherwise the result is trivial.

For this lemma we only consider one walker,  $X$ ; as such, we drop the  $X$  superscripts. We define

$$\tau_{\text{isol}}(s) = \inf\{t \geq s \mid d_t(X_t) = 0\} \quad \text{and} \quad \hat{\tau}_{\text{isol}}(s) = \inf\{t \geq \tau_{\text{isol}}(s) \mid d_t(X_t) > 0\};$$

also write  $\tau_{\text{isol}} = \tau_{\text{isol}}(0)$  and  $\hat{\tau}_{\text{isol}} = \hat{\tau}_{\text{isol}}(0)$ .

For  $k = 0, \dots, 3K$ , set  $t_k = \frac{t}{3K}$  and  $t'_k = t_k + \frac{1}{2}t/(3K)$ ; also, for  $k = 1, \dots, 3K$ , write

$$\mathcal{H}_k = \mathcal{H}[t_{k-1}, t'_{k-1}] \quad \text{and} \quad \mathcal{J}_k = \{\hat{\tau}_{\text{isol}}(t_{k-1}) \leq t_k\}.$$

If  $\mathcal{J}_k$  occurs then at some point in the interval  $[t_{k-1}, t_k]$  the walk is isolated and at a later point (in the same interval) is not. Observe that we have

$$\begin{aligned} \{\tau_{K+1} > t\} \cap \mathcal{H}[0, t] &\subseteq \left\{ \sum_{k=1}^{3K} \mathbf{1}(\mathcal{J}_k) \leq K \right\} \cap \mathcal{H}[0, t] \\ &= \left\{ \sum_{k=1}^{3K} \mathbf{1}(\mathcal{J}_k^c) \geq 2K \right\} \cap \mathcal{H}[0, t] \subseteq \left\{ \sum_{k=1}^{3K} \mathbf{1}(\mathcal{J}_k^c \cap \mathcal{H}_k) \geq 2K \right\}. \end{aligned}$$

Write  $J = \sum_{k=1}^{3K} \mathbf{1}(\mathcal{J}_k^c \cap \mathcal{H}_k)$ . Note that by the Markov property we have  $J \preceq \text{Bin}(3K, q)$  where

$$q = \max_{x_0, \eta_0} \mathbb{P}_{x_0, \eta_0}(\mathcal{J}_1^c \cap \mathcal{H}_1).$$

We shall show, for a suitable constant  $c_2$  in the definition of  $K$ , that  $q \leq \frac{1}{3}$ , and then deduce that

$$\mathbb{P}_{x_0, \eta_0}(\tau_{K+1} > t, \mathcal{H}[0, t]) \leq \mathbb{P}(\text{Bin}(3K, \frac{1}{3}) \geq 2K) \leq \exp(-\frac{2}{3}K).$$

Observe that we have

$$\left\{ \hat{\tau}_{\text{isol}} > \frac{t}{3K}, \tau_{\text{isol}} \leq \frac{t}{6K} \right\} \subseteq \left\{ \hat{\tau}_{\text{isol}} - \tau_{\text{isol}} > \frac{t}{6K} \right\}.$$

Thus we have, for any  $(x_0, \eta_0)$ , that

$$\mathbb{P}_{x_0, \eta_0}(\mathcal{J}_1^c \cap \mathcal{H}_1) \leq \mathbb{P}_{x_0, \eta_0}(\hat{\tau}_{\text{isol}} - \tau_{\text{isol}} > \frac{t}{6K}) + \mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > \frac{t}{6K}, \mathcal{H}_1).$$

The first term is simply

$$\mathbb{P}(\mathcal{E}((n-1)\lambda\mu/n) > \frac{t}{6K}) \leq \mathbb{P}(\mathcal{E}(\mu) > \frac{t}{6K}) = \exp(-\frac{1}{6}\mu t/K)$$

since there are  $n-1$  edges that can open, and  $\lambda > 1$ . Applying Theorem 4.1, we have

$$\mathbb{P}_{x_0, \eta_0}(\tau_{\text{isol}} > \frac{t}{6K}, \mathcal{H}_1) \leq 2 \exp(-\frac{1}{6}\mu t/(K \log_{(M)} n)).$$

Combining these two bounds, we then find that

$$\begin{aligned} q = \max_{x_0, \eta_0} \mathbb{P}_{x_0, \eta_0}(\mathcal{J}_1^c \cap \mathcal{H}_1) &\leq \exp(-\frac{1}{6}\mu t/K) + 2 \exp(-\frac{1}{6}c_1^* \mu t/(K \log_{(M)} n)) \\ &\leq 3 \exp(-\frac{1}{6}c_1^* \mu t/(K \log_{(M)} n)). \end{aligned}$$

Hence there exists a positive constant  $c_2$  so that if  $K = \lfloor c_2 \mu t / \log_{(M)} n \rfloor$  then  $q \leq \frac{1}{3}$ .  $\square$

## 5 Coupling

### 5.1 Statement and Application of Coupling to Mixing

For this section only, we call a graph *good* if it satisfies the conditions of Definition 2.2 and in addition the condition that at least a proportion  $c_*$  of its vertices are isolated. Since this is an additional condition, the probability that a graph is good decreases, and hence all our isolation results (from §4) still hold with this extra condition. Recall also the definition of  $H$  from Definition 2.4, and in particular that  $\pi_{\text{ER}}(H) = 1 - o(1)$ .

In this section,  $(X, \eta)$  and  $(Y, \xi)$  are two realisations of the dynamical percolation system; we shall define a coupling of the two systems, and find a tail bound on the coupling time. The coupling will be *coalescent*, ie will have the property that

if  $(X_t, \eta_t) = (Y_t, \xi_t)$  then  $(X_s, \eta_s) = (Y_s, \xi_s)$  for all  $s \geq t$ , ie if  $(X_t, \eta_t) \neq (Y_t, \xi_t)$  then  $\tau_c > t$ ,

where  $\tau_c = \inf\{t \geq 0 \mid (X_t, \eta_t) = (Y_t, \xi_t)\}$  is the *coalescence time*. We look first at the case when the environments  $\eta$  and  $\xi$  start with  $\eta_0 = \xi_0$ .

**Proposition 5.1** (Coupling Tail Bound). *There exists a Markovian coupling, which we denote by  $\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}$  when  $(X, \eta)$  and  $(Y, \xi)$  start from  $(x_0, \eta_0)$  and  $(y_0, \xi_0)$  respectively, so that, for all  $M \in \mathbb{N}$ , all  $n$  sufficiently large, all  $t$  with  $\mu t \leq \log n$  and all  $(x_0, \eta_0)$  and all  $\eta_0 \in H$ , we have*

$$\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}(\tau_c > t) \leq 3 \exp(-\mu t / \log_{(M)} n).$$

From Proposition 5.1 we are able to deduce the upper bounds in Theorems 1.1 and 1.2. As before, it suffices to prove these upper bounds up to a constant multiple.

**Proof of Theorem 1.2.** Observe that we have

$$\|\mathbb{P}_{x_0, \eta_0}(X_t = \cdot) - \mathbb{P}_{y_0, \eta_0}(Y_t = \cdot)\|_{\text{TV}} \leq \|\mathbb{P}_{x_0, \eta_0}(X_t = \cdot, \eta_t = \cdot) - \mathbb{P}_{y_0, \eta_0}(Y_t = \cdot, \xi_t = \cdot)\|_{\text{TV}}.$$

While the walk component alone is not a Markov chain, the full system is. It is then standard to upper bound the total variation distance by the tail probability of the coalescence time:

$$\|\mathbb{P}_{x_0, \eta_0}(X_t = \cdot, \eta_t = \cdot) - \mathbb{P}_{y_0, \eta_0}(Y_t = \cdot, \xi_t = \cdot)\|_{\text{TV}} \leq \mathbb{P}_{(x_0, \eta_0), (y_0, \eta_0)}(\tau_c > t).$$

The coupling is coalescent, and so this tail bound is monotone in  $t$ . The theorem now follows immediately from Proposition 5.1, since  $\eta_0 \in H$ .  $\square$

**Proof of Theorem 1.1.** We first consider an upper bound on  $t_{\text{mix}}(\varepsilon)$ . Fix  $(x_0, \eta_0)$  and  $(y_0, \xi_0)$ . By Chapman-Kolmogorov, we have

$$\mathbb{P}_{x_0, \eta_0}((X_{s+t}, \eta_{s+t}) \in \cdot) = \mathbb{E}_{x_0, \eta_0}(\mathbb{P}_{X_s, \eta_s}((X_t, \eta_t) \in \cdot)).$$

Hence for any coupling  $\mathbb{Q}$  of  $\mathbb{P}_{x_0, \eta_0}((X_s, \eta_s) = \cdot)$  and  $\mathbb{P}_{y_0, \xi_0}((Y_s, \xi_s) = \cdot)$  we have

$$\begin{aligned} & \|\mathbb{P}_{x_0, \eta_0}((X_{s+t}, \eta_{s+t}) \in \cdot) - \mathbb{P}_{y_0, \xi_0}((Y_{s+t}, \xi_{s+t}) \in \cdot)\|_{\text{TV}} \\ & \leq \mathbb{Q}(\eta_s \neq \xi_s) + \mathbb{Q}(\eta_s \notin H) + \max_{x'_0, y'_0, \eta'_0 \in H} \|\mathbb{P}_{x'_0, \eta'_0}((X_t, \eta_t) \in \cdot) - \mathbb{P}_{y'_0, \eta'_0}((Y_t, \xi_t) \in \cdot)\|_{\text{TV}}. \end{aligned}$$

In particular, consider the following such coupling  $\mathbb{Q}$ : fix  $s \geq 0$ , and couple  $(\eta_s, \xi_s)$  using the optimal coupling when started from  $(\eta_0, \xi_0)$ ; given  $\eta_s$  (and the fixed  $\eta_0$ ), sample  $X_s$  conditional on  $\eta_s$  (and  $\eta_0$ ); do similarly (and independently) for  $Y_s$  with  $\xi_s$  (and  $\xi_0$ ). This then has

$$\mathbb{Q}(\eta_s \neq \xi_s) = \|\mathbb{P}_{\eta_0}(\eta_s \in \cdot) - \mathbb{P}_{\xi_0}(\xi_s \in \cdot)\|_{\text{TV}}.$$

To determine this probability, we just look at the mixing of the environment, which is simply a biased walk on the hypercube  $\{0, 1\}^N$ , where  $N = \binom{n}{2}$ , where each coordinate refreshes at rate  $\mu$ . We claim that for any  $\gamma$ , which may depend on  $n$ , we have

$$\|\mathbb{P}_{\eta_0}(\eta_r \in \cdot) - \mathbb{P}_{\xi_0}(\xi_r \in \cdot)\|_{\text{TV}}^2 \leq \exp(e^{-2\gamma}) - 1 \quad \text{when} \quad \mu r = \frac{1}{2} \log N + \gamma.$$

The eigenvalues may be found exactly: see [19, Example 12.16], replacing  $\{-1, 1\}$  with  $\{-p, 1-p\}$ . From this it is easy to deduce the claim: see [19, Lemma 12.18 and Example 12.19].

Fix  $M \in \mathbb{N}$ . Choose  $s$  so that  $\mu s = \log n + \log_{(M+1)} n$ , and so  $\mu s \geq \frac{1}{2} \log N + \log_{(M+1)} n$ , to get

$$\|\mathbb{P}_{\eta_0}(\eta_s \in \cdot) - \mathbb{P}_{\xi_0}(\xi_s \in \cdot)\|_{\text{TV}}^2 \leq \exp(n^{-2 \log_{(M+1)} n}) - 1 \leq 2n^{-2 \log_{(M+1)} n} = o(1).$$

This shows cutoff for the mixing of the environment, and hence we deduce, for this  $s$ , that

$$\mathbb{P}(\eta_s \neq \xi_s) = o(1) \quad \text{and} \quad \mathbb{P}(\eta_s \notin H) \leq \pi_{\text{ER}}(H^c) + o(1) = o(1).$$

Since our coupling  $\mathbb{P}_\cdot$ , from Proposition 5.1 is Markovian and coalescent, we have

$$\left\| \mathbb{P}_{x_0, \eta_0}((X_t, \eta_t) \in \cdot) - \mathbb{P}_{y_0, \eta_0}((Y_t, \xi_t) \in \cdot) \right\|_{\text{TV}} \leq \mathbb{P}_{(x_0, \eta_0), (y_0, \eta_0)}(\tau_c > t).$$

Noting the conditions of Proposition 5.1, this implies that

$$\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}(\tau_c > t) \leq \varepsilon^2 \quad \text{when} \quad t = \frac{1}{\mu} \log(3/\varepsilon^2) \log_{(M)} n.$$

Combining these three bounds we obtain, for these  $s$  and  $t$ , that

$$\left\| \mathbb{P}_{x_0, \eta_0}((X_{s+t}, \eta_{s+t}) \in \cdot) - \mathbb{P}_{y_0, \xi_0}((Y_{s+t}, \xi_{s+t}) \in \cdot) \right\|_{\text{TV}} \leq \varepsilon^2 + o(1) + o(1) \leq \varepsilon.$$

Hence for all  $\varepsilon \in (0, 1)$  there exists a constant  $C$  (depending on  $\varepsilon$ ) so that

$$\mu \cdot t_{\text{mix}}(\varepsilon) \leq \log n + C \log_{(M)} n.$$

This completes the proof of the upper bound.

We now show a lower bound on  $t_{\text{mix}}(1 - \varepsilon)$ , which will follow easily by just considering the mixing of environment. First observe that, trivially,

$$\left\| \mathbb{P}_{x_0, \eta_0}((X_t, \eta_t) \in \cdot) - \pi_U \times \pi_{\text{ER}} \right\|_{\text{TV}} \geq \left\| \mathbb{P}_{\eta_0}(\eta_t \in \cdot) - \pi_{\text{ER}} \right\|_{\text{TV}}.$$

Thus it suffices to only show that the environment has not mixed by time  $t$ . Again write  $N = \binom{n}{2}$ .

Once an edge refreshes, it is open with probability  $p = \lambda/n$  and closed with probability  $1 - p$ . Suppose we start with all edges open, ie  $\eta_0(e) = 1$  for all  $e \in E$ ; we denote this state  $\mathbf{1}$ . Write  $N_t$  for the number of open edges in  $\eta_t$ , ie

$$N_t = |\{e \in E \mid \eta_t(e) = 1\}|.$$

For an edge  $e \in E$ , let  $E_1(e)$  be the first time the edge refreshes, and let us write

$$q_t = \mathbb{P}(E_1(e) > t) = \mathbb{P}(\mathcal{E}(\mu) > t) = e^{-\mu t}.$$

Then, since all edges start open and  $p = o(1)$ , the probability a given edge is open at time  $t$  is

$$(1 - q_t)p + q_t = p + q_t(1 - p) \geq p + \frac{1}{2}q_t, \quad \text{and hence} \quad N_t \succcurlyeq \text{Bin}(N, p + \frac{1}{2}q_t).$$

Now fix  $t = (\frac{1}{2} \log N - \alpha)/\mu$ . By Hoeffding, we have

$$\mathbb{P}_{\mathbf{1}}(N_t \leq N(p + \frac{1}{3}q_t)) \leq \mathbb{P}(\text{Bin}(N, p + \frac{1}{2}q_t) \leq (p + \frac{1}{3}q_t)N) \leq \exp(-\frac{1}{18}q_t^2 N).$$

In equilibrium (ie under  $\pi_{\text{ER}}$ ), the number of open edges is  $\text{Bin}(N, p)$ , and again by Hoeffding

$$\mathbb{P}(\text{Bin}(N, p) \geq N(p + \frac{1}{6}q_t)) \leq \exp(-\frac{1}{18}q_t^2 N).$$

Hence, writing  $A$  for the set of all graphs with at least  $N(p + \frac{1}{4}q_t)$  edges, we have

$$\left\| \mathbb{P}_{\mathbf{1}}(\eta_t \in \cdot) - \pi_{\text{ER}} \right\|_{\text{TV}} \geq \mathbb{P}_{\mathbf{1}}(\eta_t \in A) - \pi_{\text{ER}}(A) \geq 1 - 2 \exp(-\frac{1}{18}q_t^2 N).$$

Observe that  $q_t^2 N = e^{2\alpha}$  when  $\mu t = \frac{1}{2} \log N - \alpha$ ; also  $N \geq n^2$ . Hence we have

$$\left\| \mathbb{P}_{\mathbf{1}}(\eta_t \in \cdot) - \pi_{\text{ER}} \right\|_{\text{TV}} \geq 1 - \varepsilon \quad \text{when} \quad t \geq \frac{1}{\mu} (\log n - \frac{1}{2} \log(18 \log(2/\varepsilon))).$$

Hence for all  $\varepsilon \in (0, 1)$  there exists a constant  $C$  (depending on  $\varepsilon$ ) so that

$$\mu \cdot t_{\text{mix}}(\varepsilon) \geq \log n - C.$$

This completes the proof of the lower bound, and hence of the theorem.  $\square$

It remains to prove Proposition 5.1. To prove this, we carefully define a coupling, and use the result on dual-walker isolation, Theorem 4.2, that we proved in the previous section.

## 5.2 Coupling Description and Proof of Tail Bound

Below, we write  $(x, y)$  for the *undirected* edge with endpoints  $x$  and  $y$ ; in particular,  $(x, y) = (y, x)$ . We also talk of “the pair of edges  $(x, y)$  in  $\eta$  and  $(u, v)$  in  $\xi$ ” to mean the edge  $(x, y)$  in the  $\eta$ -configuration and  $(u, v)$  in the  $\xi$ -configuration.

The coupling that we define will be valid whatever the state of the system is, but we shall be interested in using it when the environments are ‘close’, in the following sense.

**Definition 5.2** (Environments  $H_c$ ). Define

$$H_c = \{(\eta, \xi) \in \{0, 1\}^E \times \{0, 1\}^E \mid \exists \text{ distinct } x, y, z \in V \text{ s.t. } \eta(e) = \xi(e) \forall e \notin \{(x, z), (y, z)\}\}. \quad \triangle$$

We now define the coupling.

**Definition 5.3** (Coupling for Two Full Systems). We define four steps for the coupling.

- (i) Run the environments together (ie using the same source of randomness) and the walks independently until the environments agree. Call the runtime for this step  $\sigma_1$ .
- (ii) Run the environments together and the walks independently until the walks are simultaneously isolated (not necessarily at the same vertex); write  $x'_0$  for the location of  $X$  at this time and  $y'_0$  for  $Y$ . If  $x'_0 = y'_0$  then we have coupled; otherwise continue to the next step.
- (iii) Continue using the same edge-refresh clocks on all edges other than those incident to  $x'_0$  or  $y'_0$ . Give the  $2n - 3$  (closed) edges incident to the pair  $\{x'_0, y'_0\}$  a joint change-clock, with rate  $(2n - 3)p\mu = (2 - 3/n)\lambda\mu$ : when this clock rings we need to choose an edge to open. Call the runtime for this step (ie this clock)  $\sigma_2$ .

With probability  $1/(2n - 3)$  we open the edge  $(x'_0, y'_0)$ ; if this happens, we say the coupling has failed and return to Step (ii). Otherwise choose a vertex  $z$  uniformly among  $V \setminus \{x'_0, y'_0\}$  and proceed as follows.

In  $\eta$ , open the edge  $(x'_0, z)$  or  $(y'_0, z)$  each with probability  $\frac{1}{2}$ . Do the ‘opposite’ in  $\xi$ : if  $(x'_0, z)$  is opened in  $\eta$ , then open  $(y'_0, z)$  in  $\xi$ ; if  $(y'_0, z)$  is opened in  $\eta$ , then open  $(x'_0, z)$  in  $\xi$ . [See Figure 5.1.]

If  $z$  were isolated (when chosen) and we opened the edges  $(x'_0, z)$  in  $\eta$  and  $(y'_0, z)$  in  $\xi$ , then we continue to the next step; otherwise we say the coupling has failed and return to Step (ii). (Note that if we continue then the environments have decoupled.)

- (iv) Use the same edge-refresh clocks on all edges other than  $(x'_0, z)$  and  $(y'_0, z)$ . Give the pair of edges  $(x'_0, z)$  in  $\eta$  and  $(y'_0, z)$  in  $\xi$  the same edge-refresh clock; independently do the same for the pair  $(y'_0, z)$  in  $\eta$  and  $(x'_0, z)$  in  $\xi$ . Couple the walks  $X$  and  $Y$  similarly so that they move to  $z$  together and away from  $z$  together.

Run until an edge incident to any vertex of  $\{x'_0, y'_0, z\}$  changes state. Call the runtime for this step  $\sigma_3$ . In particular, this tells us that for any time  $t$  in this interval we have either  $\{X_t = x'_0, Y_t = y'_0\}$  or  $\{X_t = z = Y_t\}$ , since the walkers move together. [See Figure 5.2.]

If the change is caused by the open edges  $(x'_0, z)$  in  $\eta$  and  $(y'_0, z)$  in  $\xi$  both closing (recall that they have the same clocks), then we continue; otherwise we say the coupling has failed, and return to Step (i). (Note that if we continue then the environments have now recoupled.)

Write  $\sigma = \sigma_1 + \sigma_2 + \sigma_3$  (with time initialised to 0 at the start of Step (i)). If we do not have  $X_\sigma = z = Y_\sigma$ , ie we have  $X_\sigma = x'_0$  and  $Y_\sigma = y'_0$ , then we say the coupling has failed, and return to Step (ii); otherwise we have  $X_\sigma = z = Y_\sigma$  and also  $\eta_\sigma = \xi_\sigma$ . In the latter case, the full systems have coalesced, and from then on we run the systems together.

Observe that this defines a genuine, coalescent, Markovian coupling. When the systems  $(X, \eta)$  and  $(Y, \xi)$  start from  $(x_0, \eta_0)$  and  $(y_0, \xi_0)$ , respectively, we denote this coupling  $\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}$ .  $\triangle$

Observe that  $\sigma_2 + \sigma_3 \preceq \Gamma(2, \mu)$  since  $\lambda > 1$ . Suppose that the environments start in  $H_c$ . Observe that if the coupling fails at some point, then we do not necessarily have that the environments



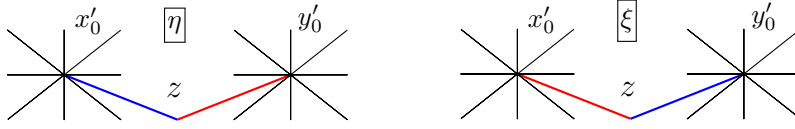


Figure 5.1. The black edges represent the closed edges incident to  $x'_0$  and  $y'_0$  (recall that they are both isolated at  $\tau$ ). The **BLUE** colour indicates opening an edge to  $z$ ; **RED** indicates leaving it closed. The other case is with the **BLUE** and **RED** swapped

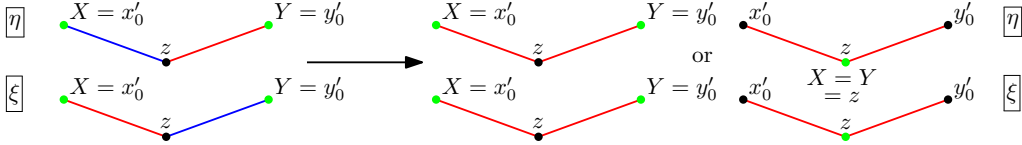


Figure 5.2. The **BLUE** colour indicates an open edge; **RED** indicates a closed edge. The walkers move up and down the **BLUE** edges, moving together. On the left-hand side the **GREEN** dots represent where the walkers start (and the black are empty). On the right-hand side the **GREEN** dots represent where the walkers end (and the black are empty)

are the same, but (as a pair) they are in  $H_c$ . Since there are then at most two mismatched edges,  $\sigma_1 \preccurlyeq \Gamma(2, \mu)$ , independent of the previous  $\Gamma$ . Hence, when  $(\eta_0, \xi_0) \in H_c$ , we have  $\sigma \preccurlyeq \Gamma(4, \mu)$ .

Before proving Proposition 5.1, we make an important observation, and define some notation. In our coupling, once the clock in Step (iii) rings, it *cannot* be the case that Step (iii) or (iv) are still running *and* both walks are isolated and the environments are the same. With this observation in mind, we now set up the following notation. Recall that we defined

$$\tau = \inf \{ t \geq 0 \mid \eta_t = \xi_t, d_t(X_t) = 0 = d_t(Y_t) \}.$$

We consider multiple attempts at coupling. Let  $\tau_0 = 0$ , and for  $k \geq 1$  define inductively

$$\begin{aligned} \hat{\tau}_k &= \inf \{ t \geq \tau_k \mid d_t^\eta(X_t) \neq 0 \text{ or } d_t^\xi(Y_t) \neq 0 \text{ or } \eta_t \neq \xi_t \}, \\ \tau_{k+1} &= \inf \{ t \geq \hat{\tau}_k \mid \eta_t = \xi_t, d_t(X_t) = 0 = d_t(Y_t) \}. \end{aligned}$$

We now state a lemma (very similar in character to Lemma 4.8) that we use to prove Proposition 5.1. We then prove the proposition with this lemma, and finally prove the lemma.

**Lemma 5.4.** *There exists a positive constant  $c$  so that, for all  $M$ , all  $n$  sufficiently large and all  $(x_0, \eta_0)$  and  $(y_0, \xi_0)$  with  $(\eta_0, \xi_0) \in H_c$ , we have (using the coupling  $\mathbb{P}_{\cdot, \cdot}$  given by Definition 5.3)*

$$\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}(\tau_c > t, \tau_{K+1} > t, \mathcal{H}[0, t]) \leq \exp(-\frac{1}{2}K) \quad \text{when } K = \lfloor c\mu t / \log_{(M)} n \rfloor.$$

**Proof of Proposition 5.1.** As before, it suffices to prove an upper bound of  $3 \exp(-c\mu t / \log_{(M)} n)$  for a positive constant  $c$ . We actually prove a slightly stronger result. We show first that there exists a positive constant  $c'$  so that, for all  $M$ , all  $n$  sufficiently large and any  $(\eta_0, \xi_0) \in H_c$ , we have

$$\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}(\tau_c > t, \mathcal{H}[0, t]) \leq 2 \exp(-c' \mu t / \log_{(M)} n).$$

Note that  $\{(\eta_0, \xi_0) \mid \eta_0 \in \{0, 1\}^E\} \subseteq H_c$ . Recall that we defined  $H$  in Definition 2.4 as

$$H = \{ \eta_0 \in \{0, 1\}^E \mid \mathbb{P}_{\eta_0}(\mathcal{H}[0, n/\mu]^c) \leq n^{-1} \}.$$

By Proposition 2.5 we have  $\pi_{\text{ER}}(H) = 1 - o(1)$ . Hence, for  $\eta_0 \in H$  and  $\mu t \leq \log n$ , we have

$$\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}(\tau_c > t) \leq 2 \exp(-c' \mu t / \log_{(M)} n) + 2n^{-1} \leq 3 \exp(-c' \mu t / \log_{(M)} n).$$

Hence we have deduced Proposition 5.1. To show that such a  $c'$  exists, we assume that  $\mu t \geq C \log_{(M)} n$  for some sufficiently large constant  $C$  to be chosen later, otherwise the result is trivial.

Fix  $M$ . By inspection there exists a positive constant  $q$  so that the probability that the first attempt at Steps (iii)–(iv) of the coupling, started from  $(x_0, \eta_0)$  and  $(y_0, \xi_0)$ , succeeds is at least

$$q \frac{1}{n} \mathbb{E}_{(x_0, \eta_0), (y_0, \xi_0)} (I - 2),$$

where  $I$  is the number of isolated vertices in the graph at the time the clock in Step (iii) rings. (Note that at this time the environments are the same, so it makes sense to talk about the number of isolated vertices at this time.) The *idea* is that ‘most of the time’ we have order  $n$  isolated vertices, and so this probability will be order 1.

We now set  $K = \lfloor c\mu t / \log_{(M)} n \rfloor$  using Lemma 5.4. We then have

$$\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)} (\tau_c > t, \mathcal{H}[0, t]) \leq \mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)} (\tau_c > t, \tau_{K+1} \leq t, \mathcal{H}[0, t]) + \exp(-\frac{1}{2}K). \quad (5.1)$$

Consider repeatedly running Steps (i)–(iv). Write  $A_k$  for the event that the coupling occurs (not necessarily for the first time) between  $\tau_k$  and  $\tau_{k+1}$ ; informally, this says that the coupling attempt Steps (iii)–(iv) started from  $\tau_k$  succeeds. Let us also write  $\mathcal{H}_k = \{\eta_{\tau_k}, \xi_{\tau_k} \in \mathcal{H}\}$ . (Note that  $\eta_{\tau_k} = \xi_{\tau_k}$  by definition of  $\tau_k$ .) Then

$$\begin{aligned} \mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)} (\tau_c > t, \tau_{K+1} \leq t, \mathcal{H}[0, t]) &\leq \mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)} (\cap_{k=1}^K (A_k^c \cap \mathcal{H}_k)) \\ &= \mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)} (\cap_{k=1}^K (A_k^c \cap \mathcal{H}_k) \mid A_1^c \cap \mathcal{H}_1) \cdot \mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)} (A_1^c \cap \mathcal{H}_1) \\ &\leq \left( \max_{x'_0, y'_0, \eta'_0} \mathbb{P}_{(x'_0, \eta'_0), (y'_0, \eta'_0)} (A_1^c) \cdot \mathbf{1}(\eta'_0 \in \mathcal{H}) \cdot \mathbf{1}(d'_0(x'_0) = 0 = d'_0(y'_0)) \right)^K, \end{aligned} \quad (5.2)$$

by the strong Markov property, conditioning on the location at the end of the attempt. From the definition of the coupling, for any  $(x'_0, y'_0)$  and  $\eta'_0$  with  $\eta'_0 \in \mathcal{H}$  and  $d'_0(x'_0) = 0 = d'_0(y'_0)$ , we have

$$\begin{aligned} \mathbb{P}_{(x'_0, \eta'_0), (y'_0, \eta'_0)} (A_1) &\geq q \frac{1}{n} \mathbb{E}_{\eta'_0} (I - 2) \geq q \frac{1}{n} \mathbb{E}_{\eta'_0} ((I - 2) \cdot \mathbf{1}(\mathcal{G}[0, \sigma_2^-])) \\ &\geq \frac{1}{2} c_* q \cdot \left( 1 - \mathbb{P}_{\eta'_0} (\mathcal{G}[0, \frac{1}{\mu}]^c) - \mathbb{P}_{\eta'_0} (\sigma_2 > \frac{1}{\mu}) \right) \geq r \end{aligned} \quad (5.3)$$

for a positive constant  $r$ , since  $\eta'_0 \in \mathcal{H}$  and  $\sigma_2 \preceq \mathcal{E}(\mu)$ , where  $c_*$  is from Definition 2.2. Also note that, for such  $(x'_0, y'_0)$  and  $\eta'_0$ , the environments are coupled until  $\sigma_2$ . Hence we have  $\mathbb{P}_{(x'_0, \eta'_0), (y'_0, \eta'_0)} (A_1^c) \leq e^{-r}$  for such  $(x'_0, y'_0)$  and  $\eta'_0$ ; note that  $r \leq 1$ . Since  $\mu t \geq C \log_{(M)} n$ , we can choose  $C$  large enough so that  $K \geq \frac{1}{2} c\mu t / \log_{(M)} n$ . Combining (5.1–5.3) we obtain

$$\begin{aligned} &\max_{(x_0, \eta_0), (y_0, \xi_0)} \mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)} (\tau_c > t, \mathcal{H}[0, t]) \cdot \mathbf{1}((\eta_0, \xi_0) \in H_c) \\ &\leq \exp(-rK) + \exp(-\frac{1}{2}K) \leq 2 \exp(-\frac{1}{2}rK) \\ &\leq 2 \exp(-\frac{1}{4}c\mu t / \log_{(M)} n). \end{aligned} \quad \square$$

It remains to prove the Hoeffding-type inequality that is Lemma 5.4. The bulk of this proof will be very similar to the proof of Lemma 4.8; with the exception of the first few paragraphs of text and some manipulation of the  $\Gamma$  distribution, the ideas will be exactly the same.

**Proof of Lemma 5.4.** We may assume that  $K \geq 1$ , otherwise the result is trivial. Also, throughout this proof we implicitly work on the event  $\{\tau_c > t\}$ , which means that we are running Steps (i)–(iv) of Definition 5.3.

We first explain the intuition behind the proof. Recall that in Step (ii) we run the walks independently and the environments together. Observe that Steps (i) and (iii)–(iv) are ‘short’ compared with Step (ii), so ‘most’ of the time of the algorithm is spent with the two environments the same and waiting for the two walks to become simultaneously isolated. Recall also that the time taken to process Steps (i), (iii) and (iv) is at most  $\Gamma(4, \mu)$ .

This motivates looking at the case where we always keep the environments the same and run the walks independently, but every time the walks become simultaneously isolated we add a holding period of  $\Gamma(4, \mu)$ . This is not quite sufficient, however: Steps (iii) and (iv) are different to (ii), and so we need to take into account from where Step (ii) is started each time.

To this end, for the moment we work with two walks,  $X$  and  $Y$ , walking independently on the *same* environment,  $\eta$ . We define

$$\tau(s) = \inf\{t \geq s \mid d_t(X_t) = 0 = d_t(Y_t)\}, \quad \text{and} \quad \hat{\tau}(s) = \inf\{t \geq \tau(s) \mid d_t(X_t) \neq 0 \text{ or } d_t(Y_t) \neq 0\};$$

also write  $\tau = \tau(0)$  and  $\hat{\tau} = \hat{\tau}(0)$ .

For  $k = 0, \dots, 3K$ , set  $t_k = \frac{t}{3K}$  and  $t'_k = t_k + \frac{1}{2}t/(3K)$ ; also, for  $k = 1, \dots, 3K$ , write

$$\mathcal{H}_k = \mathcal{H}[t_{k-1}, t'_k] \quad \text{and} \quad \mathcal{J}_k = \{\hat{\tau}(t_{k-1}) \leq t_k\}.$$

If  $\mathcal{J}_k$  occurs then at some point in the interval  $[t_{k-1}, t_k]$  the walks are jointly isolated and at a later point (in the same interval) are not. Observe that we have

$$\left\{ \sum_{k=1}^{3K} \mathbf{1}(\mathcal{J}_k) \leq K \right\} \cap \mathcal{H}[0, t] = \left\{ \sum_{k=1}^{3K} \mathbf{1}(\mathcal{J}_k^c) \geq 2K \right\} \cap \mathcal{H}[0, t] \subseteq \left\{ \sum_{k=1}^{3K} \mathbf{1}(\mathcal{J}_k^c \cap \mathcal{H}_k) \geq 2K \right\}.$$

Write  $J = \sum_{k=1}^{3K} \mathbf{1}(\mathcal{J}_k)$ . If  $\tau_{K+1} > t$  then  $J \leq K$ . Note that by the Markov property we have  $J \preceq \text{Bin}(3K, q)$  where

$$q = \max_{x_0, y_0, \eta_0} \mathbb{P}_{x_0, y_0, \eta_0}(\mathcal{J}_1^c \cap \mathcal{H}_1).$$

It is important that  $\mathbb{P}_{x_0, y_0, \eta_0}(\mathcal{J}_1^c \cap \mathcal{H}_1) \leq q$  for all  $(x_0, y_0, \eta_0)$ , as we use the strong Markov property to condition on the state of the system at the start of Step (ii). We must also take into account the  $\Gamma(4, \mu)$  ‘holding times’, of which there are at most  $K + 1$  if  $\tau_{K+1} > t$ . This gives

$$\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}(\tau_c > 2t, \tau_{K+1} > 2t, \mathcal{H}[0, t]) \leq \mathbb{P}(\text{Bin}(3K, q) \geq 2K) + \mathbb{P}(\Gamma(4(K+1), \mu) > t).$$

We have  $K = \lfloor c\mu t / \log_{(M)} n \rfloor$  with  $K \geq 1$ , and hence  $K + 1 \leq 2K$ . Hence

$$\mathbb{P}(\Gamma(4(K+1), \mu) > t) \leq \mathbb{P}(\Gamma(8c\mu t / \log_{(M)} n, 1) > t) \leq \exp(-\frac{1}{3}c\mu t) \leq \exp(-\frac{1}{6}K \log_{(M)} n),$$

with the middle inequality being an easy application of Chernoff, and the final inequality holding since  $K \geq 1$ , and so  $K \geq \frac{1}{2}c\mu t / \log_{(M)} n$ . We shall show, for a suitable constant  $c$  in the definition of  $K$ , that  $q \leq \frac{1}{3}$ , and then deduce that

$$\mathbb{P}_{(x_0, \eta_0), (y_0, \xi_0)}(\tau_c > 2t, \tau_{K+1} > 2t, \mathcal{H}[0, t]) \leq \exp(-\frac{2}{3}K) + \exp(-\frac{1}{6}K \log_{(M)} n) \leq \exp(-\frac{1}{2}K).$$

Observe that we have

$$\left\{ \hat{\tau} > \frac{t}{3K}, \tau \leq \frac{t}{6K} \right\} \subseteq \left\{ \hat{\tau} - \tau > \frac{t}{6K} \right\}.$$

Thus we have, for any  $(x_0, y_0)$  and  $\eta_0$ , that

$$\mathbb{P}_{x_0, y_0, \eta_0}(\mathcal{J}_1^c \cap \mathcal{H}_1) \leq \mathbb{P}_{x_0, y_0, \eta_0}(\hat{\tau} - \tau > \frac{t}{6K}) + \mathbb{P}_{x_0, y_0, \eta_0}(\tau > \frac{t}{6K}, \mathcal{H}_1).$$

The first term is simply

$$\mathbb{P}(\mathcal{E}((2n-3)\lambda\mu/n) > \frac{t}{6K}) \leq \mathbb{P}(\mathcal{E}(\mu) > \frac{t}{6K}) = \exp(-\frac{1}{6}\mu t / K),$$

since there are  $2n - 3$  edges that can open, and  $\lambda > 1$ . Applying Theorem 4.2, we have

$$\mathbb{P}_{x_0, y_0, \eta_0}(\tau > \frac{t}{6K}, \mathcal{H}_1) \leq 2 \exp(-\frac{1}{6}c_2^* \mu t / \log_{(M)} n / K)$$

Combining these two bounds, we then find that

$$\begin{aligned} q &= \max_{x_0, y_0, \eta_0} \mathbb{P}_{x_0, y_0, \eta_0}(\mathcal{J}_1^c \cap \mathcal{H}_1) \leq \exp(-\frac{1}{6}\mu t / K) + 2 \exp(-\frac{1}{6}c_2^* \mu t / (K \log_{(M)} n)) \\ &\leq 3 \exp(-\frac{1}{6}c_2^* \mu t / (K \log_{(M)} n)). \end{aligned}$$

Hence there exists a positive constant  $c$  so that if  $K = \lfloor c\mu t / \log_{(M)} n \rfloor$  then  $q \leq \frac{1}{3}$ .  $\square$

## 6 Invariant Initial Environment

In this section we prove the result stated in Remark 1.4, which concerns the case where we draw  $\eta_0$  according to  $\pi_{\text{ER}}$  and set  $X_0 = 1$ ; we state the theorem precisely below. Throughout this entire section we consider the measure  $\mathbb{P}_{1, \text{ER}}(\cdot)$ ; for ease of notation, we drop the subscript and just write  $\mathbb{P}(\cdot)$ .

**Theorem 6.1.** *For all  $\lambda \in (0, \infty)$ , there exists a constant  $C$  so that, all  $\varepsilon \in (0, 1)$  and all  $n$  sufficiently large,  $p = \lambda/n$ , we have the following bounds on the mixing time:*

$$\begin{aligned} t_{\text{mix}}^{\text{RW}}(\varepsilon) &\geq \frac{1}{\mu} \cdot \frac{1}{2\lambda} \log(1/\varepsilon) \quad \text{if } \varepsilon \in (0, e^{-3\lambda} \wedge 1) \quad \text{for any } \mu; \\ t_{\text{mix}}^{\text{RW}}(\varepsilon) &\leq \frac{1}{\mu} \cdot C \log(1/\varepsilon) \quad \text{if } \varepsilon \in (0, \frac{1}{4}) \quad \text{when } \mu \leq \frac{2}{3}(1 + \lambda)^{-1}/n. \end{aligned}$$

**Proof of Theorem 6.1 (lower bound).** Suppose the walk starts from an isolated vertex: it cannot have mixed before an incident edge opens. We make this idea precise and rigorous. We have

$$\mathbb{P}(d_0(1) = 0) = (1 - p)^n = e^{-\lambda}(1 - o(1)).$$

Let  $\tau$  be the first time an edge incident to  $X_0 = 1$  opens. By counting edges and their respective rates, we see that  $\tau \sim \mathcal{E}(\lambda\mu(1 - 1/n)) \asymp \mathcal{E}(\lambda\mu)$ . Let  $T = \lambda\mu t$ , and observe that

$$\mathbb{P}(\tau > t \mid d_0(1) = 0) \geq \mathbb{P}(\mathcal{E}(\lambda\mu) > T/(\lambda\mu)) = e^{-T}.$$

On the event  $\{d_0(1) = 0\} \cap \{\tau > t\}$ , we have  $X_t = 1$  (in fact  $X_s = 1$  for all  $s \leq t$ ), and hence

$$\|\mathbb{P}(X_t \in \cdot) - \pi_U\|_{\text{TV}} \geq \mathbb{P}(X_t = 1) - \pi_U(1) \geq \mathbb{P}(d_0(1) = 0, \tau > t) - \pi_U(1) \geq e^{-(T+\lambda)} - \frac{1}{n}.$$

We desire  $T$  so that  $\|\mathbb{P}(X_t \in \cdot) - \pi_U\|_{\text{TV}} \geq \varepsilon$ . By the above, we may take  $T = -\log(\varepsilon + 1/n) - \lambda$ . If  $\varepsilon < e^{-3\lambda}$ , then  $T \geq \frac{1}{2} \log(1/\varepsilon)$ . This proves the lower bound, as  $t = T/(\lambda\mu)$ .  $\square$

The aim of the remainder of this section is to prove the upper bound in Theorem 6.1; herein we assume that  $\mu \leq \frac{2}{3}(1 + \lambda)^{-1}/n$ . To this end, let  $\tau$  be the first time our initial vertex is isolated and the walk is not there, ie

$$\tau = \inf\{t \geq 0 \mid d_t(X_0) = 0, X_t \neq X_0\}.$$

The *idea* is that at time  $\tau$  we are nearly uniform *and* have lost information about where we started, and so our total variation does not become large in the future. We show this rigorously.

**Proposition 6.2.** *For all  $n$  and all  $t$ , we have*

$$\|\mathbb{P}(X_t \in \cdot) - \pi_U\|_{\text{TV}} \leq \mathbb{P}(X_t = 1, \tau \leq t \wedge (n/\mu)) + \mathbb{P}(\tau > t \wedge (n/\mu)) + \frac{1}{n}.$$

**Proof.** Note that by construction and the symmetry of the graph, at all times  $t \geq 0$  we must have that  $\mathbb{P}(X_t = x)$  is constant over  $x \in V \setminus \{1\} = \{2, \dots, n\}$ : define  $\rho_t = \mathbb{P}(X_t = 1)$ ; then  $\mathbb{P}(X_t = x) = (1 - \rho_t)/(n - 1)$  for all  $x \in \{2, \dots, n\}$ . We then have

$$2 \|\mathbb{P}(X_t \in \cdot) - \pi_U\|_{\text{TV}} = (n - 1) \left| \frac{1 - \rho_t}{n - 1} - \frac{1}{n} \right| + \left| \rho_t - \frac{1}{n} \right| = 2 \left| \rho_t - \frac{1}{n} \right| \leq 2(\rho_t + \frac{1}{n}).$$

Decomposing according to the event  $\{\tau > t \wedge (n/\mu)\}$  completes the proof.  $\square$

**Proposition 6.3.** *There exists a constant  $C$  so that, for all  $K \geq 2$  and all  $n$  sufficiently large, we have*

$$\mathbb{P}(\tau > CK/\mu) \leq e^{-K}.$$

Before we prove this, we define some preliminary notation, and then state three claims. First, let  $\sigma_0 = \sigma'_1 = 0$ , and let  $\sigma_1$  be the first time the initial vertex, 1, becomes isolated, ie

$$\sigma_1 = \inf\{t \geq 0 \mid d_t(1) = 0\},$$

and for  $i \geq 1$  define inductively

$$\begin{aligned}\sigma'_{i+1} &= \inf\{t \geq \sigma_i \mid d_t(1) > 0\}, & \sigma_{i+1} &= \inf\{t \geq \sigma'_{i+1} \mid d_t(1) = 0\} \\ \text{and } \sigma''_i &= \inf\{t \leq \sigma_i \mid d_s(1) = 1 \forall s \in [t, \sigma_i]\}.\end{aligned}$$

In words,  $\sigma_i$  is the  $i$ -th time the vertex 1 becomes isolated,  $\sigma'_i$  is the first time after this that it becomes non-isolated and  $[\sigma''_i, \sigma_i)$  is the interval in which it is degree 1 immediately before becoming isolated for the  $i$ -th time. By the memoryless property,  $\sigma'_i - \sigma_{i-1} \sim^{\text{iid}} \mathcal{E}(\lambda\mu(1 - 1/n))$ .

Define  $\tau_i := \sigma_i - \sigma_{i-1}$  for  $i \geq 1$ ; then  $\tau_1$  is the time it takes to become isolated initially, and, for  $i \geq 2$ ,  $\tau_i$  is the time between the  $(i-1)$ -st and  $i$ -th times we become isolated. Note that the random variables  $\{\tau_i\}_{i \geq 2}$  are all independent and identically distributed.

We now state three lemmas which we use to deduce Proposition 6.3.

**Lemma 6.4.** *There exists a constant  $C$  so that, for all  $K \geq 1$  and all  $n$  sufficiently large, we have*

$$\mathbb{P}(\tau_1 > CK/\mu) \leq e^{-K}.$$

**Lemma 6.5.** *There exists a constant  $C$  so that, for all  $K \geq 2$  and all  $n$  sufficiently large, we have*

$$\mathbb{P}\left(\sum_{i=2}^K \tau_i > CK/\mu\right) \leq e^{-K}.$$

**Lemma 6.6.** *For all  $n$  sufficiently large and all  $i \geq 1$ , we have*

$$\mathbb{P}(X_{\sigma_i} = 1 \mid X_{\sigma_j} = 1 \forall j < i) \leq \frac{2}{3}.$$

We now show how to conclude our tail bounds on  $\tau$  from these three lemmas.

**Proof of Proposition 6.3.** Consider an integer  $K \geq 2$ . Lemma 6.6 tells us that

$$\mathbb{P}(X_{\sigma_i} = 1 \forall i \leq K) = \prod_{i=1}^K \mathbb{P}(X_{\sigma_i} = 1 \mid X_{\sigma_j} = 1 \forall j < i) \leq (2/3)^K.$$

Combining this with Lemma 6.4 and Lemma 6.5 tells us that

$$\mathbb{P}(\sigma_K \leq C'K/\mu, \exists k \leq K \text{ s.t. } X_{\sigma_k} \neq 1) \geq 1 - e^{-K} - e^{-K} - (3/2)^{-K}$$

for a suitably large constant  $C'$ . From this we deduce our claim.  $\square$

To complete the proof of our tail bound, it remains only to prove our three lemmas; we do this at the end of the section. For now, we turn to upper bounding  $\mathbb{P}(X_t = 1, \tau \leq t \wedge (n/\mu))$ .

**Lemma 6.7.** *There exists a constant  $C$  so that, for all  $n$  sufficiently large and all  $t$ , we have*

$$\mathbb{P}(X_t = 1, \tau \leq t \wedge (n/\mu)) \leq C/n.$$

**Proof.** Write  $\mathcal{I}_t$  for the set of isolated vertices at time  $t$ . Write  $s = t \wedge (n/\mu)$ . First we lower bound the number of isolated vertices at time  $\tau$  on the event  $\{\tau \leq s\}$ . From Proposition 2.5,

$$\mathbb{P}(|\mathcal{I}_\tau \setminus \{X_\tau\}| \leq \frac{1}{2}c_*n, \tau \leq s) = \mathcal{O}(n^{-2}).$$

By the symmetry of the complete graph, we must have that  $\mathbb{P}(X_t = x \mid \mathcal{F}_\tau)$  is constant over  $x \in \mathcal{I}_\tau \setminus \{X_\tau\}$  on the event  $\{\tau \leq s\}$ ; let  $\xi_t$  be this (random) value. ( $\xi_t$  is an  $\mathcal{F}_\tau$ -measurable random variable.) Now, by construction of  $\tau$ , we have that  $X_0 = 1 \in \mathcal{I}_\tau \setminus \{X_\tau\}$ . This says that

$$\xi_t = \mathbb{P}(X_t = 1 \mid \mathcal{F}_\tau) \mathbf{1}(\tau \leq s) \quad \text{and hence} \quad \mathbb{P}(X_t = 1, \tau \leq s) = \mathbb{E}(\xi_t).$$

It remains to bound  $\mathbb{E}(\xi_t)$ , which we now do. Note that we have

$$1 \geq \mathbb{P}(\tau \leq s) \geq \mathbb{E}(|\mathcal{I}_\tau \setminus \{X_\tau\}| \cdot \mathbb{P}(X_t = 1 \mid \mathcal{F}_\tau) \mathbf{1}(\tau \leq s)).$$

Letting  $A = \{|\mathcal{I}_\tau \setminus \{X_\tau\}| \geq \frac{1}{2}c_*n\}$ , we have  $\mathbb{P}(A^c, \tau \leq s) = \mathcal{O}(n^{-2})$ , as above. Hence

$$\begin{aligned} 1 &\geq \mathbb{E}(|\mathcal{I}_\tau \setminus \{X_\tau\}| \cdot \mathbb{P}(X_t = 1 \mid \mathcal{F}_\tau) \mathbf{1}(\tau \leq s) \mathbf{1}(A)) \\ &\geq \frac{1}{2}c_*n\mathbb{E}(\xi_t \mathbf{1}(A)) \geq \frac{1}{2}c_*n(\mathbb{E}(\xi_t) - \mathbb{P}(A^c, \tau \leq s)). \end{aligned}$$

Rearranging completes the proof:

$$\mathbb{P}(X_t = 1, \tau \leq s) = \mathbb{E}(\xi_t) \leq (\frac{1}{2}c_*n)^{-1} + \mathcal{O}(n^{-2}) \leq 3c_*^{-1}/n. \quad \square$$

We can now give the proof of the upper bound in Theorem 6.1.

**Proof of Theorem 6.1 (upper bound).** Lemma 6.7 says that, for all  $t$ , we have

$$\mathbb{P}(X_t = 1, \tau \leq t \wedge (n/\mu)) \leq C'/n,$$

for a constant  $C'$ . Hence we have

$$\|\mathbb{P}(X_t \in \cdot) - \pi_U\|_{\text{TV}} \leq \mathbb{P}(\tau > t \wedge (n/\mu)) + (C' + 1)/n.$$

Observe that this upper bound is (weakly) monotone-decreasing in  $t$ , and Proposition 6.3 gives us a constant  $C$  so that

$$\|\mathbb{P}(X_t \in \cdot) - \pi_U\|_{\text{TV}} \leq \varepsilon^2 + (C' + 1)/n \leq \varepsilon \quad \text{when} \quad t = 2C \log(1/\varepsilon)/\mu.$$

Hence we deduce that  $t_{\text{mix}}(\varepsilon) \leq 2C \log(1/\varepsilon)/\mu$ . □

**Proof of Lemma 6.4.** Write  $d_t = d_t(1)$ . Also rescale time by  $\mu$ , so as to remove the  $\mu$  factors from the workings. Observe that the jump-rates of  $d$  are as follows:

$$\begin{aligned} k \rightarrow k+1 &\quad \text{at rate} \quad q_+(k) = (n-1-k)p = (\lambda - \lambda(1+k)/n); \\ k \rightarrow k-1 &\quad \text{at rate} \quad q_-(k) = k(1-p) = (k - \lambda k/n). \end{aligned}$$

Let  $q(k) = q_+(k) + q_-(k)$ , and observe that  $q(k) \geq q(0) \geq 1$  for all  $k \geq 0$ . We now couple  $d$  with an auxiliary process  $d'$ , which has rate-1 jumps. Above  $3\lambda$ ,  $d'$  has probability  $\frac{2}{3}$  of going up and  $\frac{1}{3}$  of going down; below  $3\lambda$ , it has the same probabilities as  $d$ , ie  $q_+(k)/q(k)$  for up and  $q_-(k)/q(k)$  for down. Set  $d'_0 = d_0$ , and write  $\tau'_1$  for the hitting time of 0 by  $d'$ . We then have  $\tau_1 \preceq \tau'_1$ .

Note that once  $d'$  reaches  $\lceil 3\lambda \rceil$ , it moves directly to 0 (in  $\lceil 3\lambda \rceil$  steps) with probability bounded away from 0. The hitting time of  $\lceil 3\lambda \rceil$  is that of a biased simple random walk. Since  $d_0 \sim \text{Bin}(n-1, \lambda/n)$ , we may assume that  $d_0 \leq CK$  for some sufficiently large constant  $C$  with a penalty  $e^{-K}$  to the probability. Given this, we see that the hitting time of  $\lceil 3\lambda \rceil$  has mean  $\Theta(1)$  and an exponential tail. Once  $d'$  hits  $\lceil 3\lambda \rceil$ , we perform a geometric number of excursions, the length of which have an exponential tail. Hence  $\tau'_1$  has mean  $\Theta(1)$  and an exponential tail. □

**Proof of Lemma 6.5.** Again, drop the  $\mu$  factors. Note that  $\tau_i \preceq \tau_1$ , and  $\tau_1$  has mean  $\Theta(1)$  with an exponential tail. Since the  $\tau_i$  are independent, we then apply the Chernoff bound to a sum of  $K$  independent  $\tau_1$  random variables to deduce the lemma. □

**Proof of Lemma 6.6.** Fix  $i \geq 1$ . For  $t \in (\sigma''_i, \sigma_i)$  we have that  $d_t(1) = 1$ ; write  $x_i$  for the neighbour of 1 in the interval  $(\sigma''_i, \sigma_i)$ . Note that all the  $\sigma$ -times depend only on the environment, not also on the walk. We describe a coupling between  $X$  and an auxiliary walk  $X'$  which is confined to the pair  $\{1, x_i\}$ . The coupling will have the property that

$$\mathbb{P}(X_{\sigma_i} = 1 \mid X_{\sigma_j} = 1 \forall j < i) \leq \mathbb{P}(X'_{\sigma_i} = 1 \mid X_{\sigma_j} = 1 \forall j < i).$$

In particular,  $X'$  will be the usual simple random walk on  $\{1, x_i\}$ , jumping at rate  $1/(n-1)$ . Thus we shall see that the probability on the right-hand side is ‘approximately’  $\frac{1}{2}$ .

We now explicitly define the coupling. Start  $X'$  from 1. If both  $X$  and  $X'$  are at 1, then move them together; if both  $X$  and  $X'$  are at  $x_i$  and  $X$  chooses vertex 1 to jump to, then move them together; otherwise let them evolve independently. Observe that, wherever  $X$  is at time  $\sigma''_i$ , we always have for  $t \in [\sigma''_i, \sigma_i]$  that  $X_t = 1$  implies  $X'_t = 1$ . Hence our desired inequality holds.

Observe that  $X'$  is at 1 if it has taken an even number of steps (and at  $x_i$  if odd). Hence  $\mathbb{P}(X'_{\sigma_i} = 1 \mid X_{\sigma_j} = 1 \forall j < i) = \mathbb{P}(\text{Po}(\zeta) \text{ is even}) = \frac{1}{2}(1 + \mathbb{E}(e^{-2\zeta}))$  where  $\zeta = (\sigma_i - \sigma_i'')/(n-1)$ .

We have  $\sigma_i - \sigma_i'' \sim \mathcal{E}((\lambda + 1 - 3p)\mu)$  by counting edges and rates, and so the lemma follows since

$$\mathbb{E}(e^{-2\zeta}) = \frac{(\lambda + 1 - 3p)\mu}{(\lambda + 1 - 3p)\mu + 2/(n-1)} \leq \frac{1}{2}(\lambda + 1)\mu n \leq \frac{1}{3}. \quad \square$$

**Acknowledgements.** The authors would like to thank Andrew Swan for many helpful discussions at the start of this project.

## References

- [1] M. Abdullah, C. Cooper and A. Frieze (2012). Cover Time of a Random Graph with Given Degree Sequence. *Discrete Math.* **312**.21, (3146–3163) [MR2957935](#) [DOI](#)
- [2] L. Avena, H. Güldaş, R. van der Hofstad and F. den Hollander (2016). Mixing Times of Random Walks on Dynamic Configuration Models. *Available at arXiv:1606.07639*
- [3] L. Avena, H. Güldaş, R. van der Hofstad and F. den Hollander (2018). Random Walks on Dynamic Configuration Models: a Trichotomy. *Available at arXiv:1803.04824*
- [4] E. A. Bender and E. R. Canfield (1978). The Asymptotic Number of Labeled Graphs with Given Degree Sequences. *J. Combinatorial Theory Ser. A.* **24**.3, (296–307) [MR0505796](#) [DOI](#)
- [5] I. Benjamini, G. Kozma and N. Wormald (2014). The Mixing Time of the Giant Component of a Random Graph. *Random Structures Algorithms.* **45**.3, (383–407) [MR3252922](#) [DOI](#)
- [6] N. Berestycki, E. Lubetzky, Y. Peres and A. Sly (2018). Random Walks on the Random Graph. *Ann. Probab.* **46**.1, (456–490) [MR3758735](#) [DOI](#)
- [7] B. Bollobás (1980). A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs. *European J. Combin.* **1**.4, (311–316) [MR595929](#) [DOI](#)
- [8] B. Bollobás (2001). *Random Graphs*. Second ed., Cambridge University Press, Cambridge [MR1864966](#) [DOI](#)
- [9] J. Ding, E. Lubetzky and Y. Peres (2014). Anatomy of the Giant Component: the Strictly Supercritical Regime. *European J. Combin.* **35**, (155–168) [MR3090494](#) [DOI](#)
- [10] N. Fountoulakis and B. A. Reed (2007). Faster Mixing and Small Bottlenecks. *Probab. Theory Related Fields.* **137**.3-4, (475–486) [MR2278465](#) [DOI](#)
- [11] N. Fountoulakis and B. A. Reed (2008). The Evolution of the Mixing Rate of a Simple Random Walk on the Giant Component of a Random Graph. *Random Structures Algorithms.* **33**.1, (68–86) [MR2428978](#) [DOI](#)
- [12] A. Frieze and M. Karoński (2016). *Introduction to Random Graphs*. Cambridge University Press, Cambridge [MR3675279](#) [DOI](#)
- [13] D. Gillman (1998). A Chernoff Bound for Random Walks on Expander Graphs. *SIAM J. Comput.* **27**.4, (1203–1220) [MR1621958](#) [DOI](#)
- [14] O. Häggström, Y. Peres and J. E. Steif (1997). Dynamical Percolation. *Ann. Inst. H. Poincaré Probab. Statist.* **33**.4, (497–528) [MR1465800](#) [DOI](#)
- [15] R. van der Hofstad (2017). *Random Graphs and Complex Networks*. Cambridge University Press, Cambridge [MR3617364](#) [DOI](#)
- [16] M. Jerrum and A. Sinclair (1989). Approximate Counting, Uniform Generation and Rapidly Mixing Markov Chains. *Inform. and Comput.* **82**.1, (93–133) [MR1003059](#) [DOI](#)
- [17] M. Jerrum and A. Sinclair (1989). Approximating the Permanent. *SIAM J. Comput.* **18**.6, (1149–1178) [MR1025467](#) [DOI](#)



- [18] G. F. Lawler and A. D. Sokal (1988). Bounds on the  $L^2$  Spectrum for Markov Chains and Markov Processes: a Generalization of Cheeger’s Inequality. *Trans. Amer. Math. Soc.* **309**.2, (557–580) [MR930082](#) [DOI](#)
- [19] D. A. Levin, Y. Peres and E. L. Wilmer (2017). *Markov Chains and Mixing Times*. Second ed., American Mathematical Society, Providence, RI [MR3726904](#)
- [20] C. McDiarmid (1989). On the Method of Bounded Differences. *Surveys in Combinatorics, 1989 (Norwich, 1989)*, London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, (148–188) [MR1036755](#) [DOI](#)
- [21] B. D. McKay and N. C. Wormald (1991). Asymptotic enumeration by degree sequence of graphs with degrees  $o(n^{1/2})$ . *Combinatorica.* **11**.4, (369–382) [MR1137769](#) [DOI](#)
- [22] Y. Peres, P. Sousi and J. E. Steif (2017). Mixing Time for Random Walk on Supercritical Dynamical Percolation. Available at [arXiv:1707.07632](#)
- [23] Y. Peres, A. Stauffer and J. E. Steif (2015). Random Walks on Dynamical Percolation: Mixing Times, Mean Squared Displacement and Hitting Times. *Probab. Theory Related Fields.* **162**.3-4, (487–530) [MR3383336](#) [DOI](#)
- [24] B. Pittel, J. Spencer and N. Wormald (1996). Sudden Emergence of a Giant  $k$ -Core in a Random Graph. *J. Combin. Theory Ser. B.* **67**.1, (111–151) [MR1385386](#) [DOI](#)
- [25] L. Warnke (2016). On the Method of Typical Bounded Differences. *Combin. Probab. Comput.* **25**.2, (269–299) [MR3455677](#) [DOI](#)

## A Proofs of Erdős-Rényi Structure Results

In this section we give the proofs of the Erdős-Rényi structure results given in Proposition 2.3. The results will all be for supercritical Erdős-Rényi random graphs. As mentioned before, such results are usually shown to hold with probability  $1 - o(1)$  (see [8]); here we quantify this  $o(1)$ .

Properties (ii) and (iii) follow easily from properties of the Binomial distribution. We give these proofs immediately. Applying Chernoff’s bound, one can easily show, for  $\lambda > 1$ , that

$$\mathbb{P}(\text{Bin}(n - 1, \lambda/n) \geq 3\lambda K) \leq e^{-K}.$$

Taking  $K = 10 \log n$  and applying the union bound, we deduce (ii). The number of edges in the graph is precisely  $\text{Bin}(\binom{n}{2}, \lambda/n)$ , and so we deduce (iii) from concentration of the Binomial.

The remaining results take more to prove. A method we shall use repeatedly is to find expectation of various quantities and then show concentration. In particular, we shall use a result on *typical bounded difference* due to Warnke [25], which builds on work by McDiarmid [20].

Property (i) is standard, but the proof that we shall give is a nice, straightforward example of the above typical bounded differences approach, and so we give it as a type of ‘warm-up proof’.

In order to prove (iv), (v) and (vi), we use (vii). As such, we first prove (vii) directly (not using concentration, with one exception); we then formulate the concentration inequality of Warnke [25], and apply it in various situations.

Write  $G \sim \text{ER}(n, p)$  if  $G$  has the Erdős-Rényi distribution on  $n$  vertices with edge probability  $p$ . Direct calculation as at the start of the proof of [12, Theorem 2.14] tells us that if  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed then there exist positive constants  $c$  and  $C$  so that the probability there is a component with size in  $[C \log n, cn]$  is  $\mathcal{O}(n^{-100})$ .

### A.1 Expansion Properties of the Giant

We now consider (vii). We first state precisely what it is that we prove.

**Proposition A.1** (Expansion Properties of the Giant, (vii)). *Suppose  $G \sim \text{ER}(n, \lambda/n)$ , with  $\lambda > 1$  fixed, and write  $\mathcal{G}$  for its largest component (breaking ties arbitrarily). Then there exists a positive constant  $c$  so that  $\Phi_{\mathcal{G}} \geq c(\log n)^{-2}$  and  $\gamma_{\mathcal{G}} \geq c(\log n)^{-4}$  with probability  $1 - \mathcal{O}(n^{-9})$ .*

We prove this via a sequence of lemmas, the proof of which are deferred until after the proof of Proposition A.1. First, though, we make introductory observations and definitions.

Given a graph  $G$ , write  $\delta(G)$  for its minimum degree. The *core* of  $G$ , denoted  $\mathcal{C}_G = \mathcal{C}(G)$ , is obtained by removing all isolated vertices and then recursively removing all degree 1 vertices; note that  $\delta(\mathcal{C}_G) \geq 2$ . The *kernel* of  $G$ , denoted  $\mathcal{K}_G = \mathcal{K}(G)$ , is obtained from the core by replacing each maximal 2-path that joins vertices  $u$  and  $v$  of degree at least 3 by an edge  $(u, v)$ , and deleting each isolated cycle. (Note that the kernel may have self-loops and/or multiple edges, and may not be connected.) A loop contributes 2 to the degree of its incident vertex, so  $\delta(\mathcal{K}_G) \geq 3$ . Note that if  $G$  is connected, then so are  $\mathcal{C}_G$  and  $\mathcal{K}_G$ . Write  $\mathcal{G}_G = \mathcal{G}(G)$  for the largest component of  $G$ .

Given a degree sequence  $\mathbf{d}$ , write  $G(\mathbf{d})$  for the set of all (multi-)graphs with degree sequence  $\mathbf{d}$ , and write  $U(\mathbf{d})$  for  $\text{Unif}(G(\mathbf{d}))$ ; also write  $\mathbf{d}^*$  for the degree sequence  $\mathbf{d}$  with the degree 2 vertices removed. Observe that if  $G \sim U(\mathbf{d})$ , then  $\mathcal{K}_G \sim U(\mathbf{d}^*)$ .

Also write  $G'(\mathbf{d})$  for the set of all simple graphs with degree sequence  $\mathbf{d}$ , and write  $U'(\mathbf{d})$  for  $\text{Unif}(G'(\mathbf{d}))$ . Pittel, Spencer and Wormald showed in [24, Proposition 1(b)], for  $G \sim \text{ER}(n, p)$ , that  $\mathcal{C}_G \sim U'(\mathbf{d})$  conditional on having degree sequence  $\mathbf{d}$ . Since this includes conditioning on being simple, we cannot deduce that  $\mathcal{K}_G \sim U'(\mathbf{d}^*)$ ; however, it's known that a realisation of  $U(\mathbf{d})$  is simple with order 1 probability under certain conditions (we state and reference this precisely below), and this will be sufficient for us.

The *configuration model* was introduced by Bollobás in [7], and provides a method for constructing a graph with a given degree sequence with distribution that is uniform over all graphs with the given degree sequence. (Bender and Canfield in [4] also introduced a highly related object; see also Bollobás [8], or van der Hofstad [15] for a modern description.)

As some notation, for a graph  $H$ , write  $H = (V_H, E_H)$  and  $d_H(x)$  for the degree of a vertex  $x$  in  $H$ . For  $S \subseteq H$ , write  $d_H(S) = \sum_{x \in S} d_H(x)$  for the total degree of  $S$ ,  $\Phi_H(S) = e_H(S)/d_H(S)$ ,  $I_H(S)$  for the number of induced total degree (ie the sum of the degrees in the subgraph induced by  $S$ , ie twice the number of edges in the subgraph induced by  $S$ ) and  $e_H(S)$  for the number of edges in  $E_H$  between  $S$  and  $S^c$ . Then  $\Phi_H(S) = 1 - I_H(S)/d(S)$  and

$$\Phi_H = \min\{\Phi_H(S) \mid 0 < d_H(S) \leq |E_H|\}.$$

Also, for a degree sequence  $\mathbf{d}$ , write  $\delta_{\mathbf{d}}$  for its minimal degree and  $\theta_{\mathbf{d}}$  for its average degree.

**Lemma A.2.** *Let  $\mathbf{d} = (d_1, \dots, d_m)$  be a degree sequence with  $\delta_{\mathbf{d}} \geq 3$  and  $\theta_{\mathbf{d}} \leq m^{1/5}$ . Suppose that  $G \sim U(\mathbf{d})$ , and write  $\mathcal{G}$  for its largest component (breaking ties arbitrarily). We have*

$$\mathbb{P}(\Phi(\mathcal{G}) < 10^{-5}) = \mathcal{O}(m^{-50}).$$

The proof of this lemma will closely follow [1, Lemma 12], but with a few slight changes to strengthen the probability. We want to convert this from a result about  $U(\mathbf{d}^*)$  into a result about the kernel of the giant of an Erdős-Rényi graph. Recall that if  $F \sim \text{ER}(n, p)$  then  $\mathcal{C}_F \sim U'(\mathbf{d})$  conditional on having degree sequence  $\mathbf{d}$ , and that  $\mathcal{K}(U(\mathbf{d})) \sim U(\mathbf{d}^*)$ .

**Lemma A.3.** *Suppose  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  a constant, and write  $\mathcal{G}$  for its largest component (breaking ties arbitrarily). There exists a positive constant  $c$  so that  $|\mathcal{K}_G| \geq cn$  with probability at least  $1 - \mathcal{O}(n^{-10})$ .*

Lemma A.3 is proved using a concentration result; its proof is deferred until the next section.

We also need to know that if  $G \sim U(\mathbf{d})$  then  $\mathbb{P}(G \text{ simple})$  is not too small.

**Lemma A.4.** *Suppose  $F \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed. Let  $\mathbf{d}$  be a random degree sequence with the distribution of the degree sequence of  $\mathcal{C}_F$ . Let  $G \sim U(\mathbf{d})$ . Then there exists a positive constant  $c$  so that, for all  $n$  sufficiently large, we have*

$$\mathbb{P}(G \text{ simple}) \geq c.$$

The above results have been about the kernel, or the core. To convert this into a result about the giant, we consider the ‘decorations’ to the kernel. Consider a graph  $G$  and its kernel  $\mathcal{K}_G$ . An

edge of  $\mathcal{K}_G$  is first expanded into a path in the core, and then trees are hung from each vertex of the path. For an edge  $e$  of  $\mathcal{K}_G$ , write  $D_e$  for the number of vertices added in this ‘decorating’ process; write  $D = \max_e D_e$ .

Suppose the edge  $(x, y)$  in  $\mathcal{K}_G$  has decoration with vertex set  $S$ , where we do not include  $x$  or  $y$  in  $S$ . If  $S = \emptyset$ , then  $(x, y)$  is an edge in  $G$ ; suppose  $S \neq \emptyset$ . Then  $G[S]$ , the subgraph of  $G$  induced by  $S$ , is a tree. Moreover, all the (potential) edges between  $S$  and  $S^c$  are closed in  $G$ , with the exception of one edge between  $S$  and  $x$  and one between  $S$  and  $y$ .

**Lemma A.5.** *Suppose that  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed. Write  $D$  for the size of the largest decoration (as above). There exists a constant  $C$  so that*

$$\mathbb{P}(D > C \log n) = \mathcal{O}(n^{-100}).$$

Given these two lemmas, we can now prove Proposition A.1.

**Proof of Proposition A.1.** We first note that the inequality  $\frac{1}{2}\Phi^2 \leq \gamma \leq 2\Phi$  holds; this is due to Jerrum and Sinclair [16] and Lawler and Sokal [18]. Hence given  $\Phi \geq c(\log n)^{-2}$  it follows that  $\gamma \geq c'(\log n)^{-4}$  with  $c' = \frac{1}{2}c^2$ . So it suffices to just prove the statement about  $\Phi_{\mathcal{G}}$ .

Let  $F \sim \text{ER}(n, \lambda/n)$ , with  $\lambda > 1$  fixed, and let  $\mathbf{d}$  be a random degree sequence with the distribution of the degree sequence of  $\mathcal{C}_F$  (ie the core of an Erdős-Rényi graph). Let  $G' \sim U'(\mathbf{d})$ ,  $G \sim U(\mathbf{d})$  and  $H \sim U(\mathbf{d}^*)$ . We first show that

$$\mathbb{P}(\Phi_{\mathcal{G}(\mathcal{K}(F))} < 10^{-5}) = \mathcal{O}(n^{-40}).$$

Let  $c_1$  be the constant from Lemma A.2,  $c_2$  from Lemma A.3 and  $c_3$  from Lemma A.4. Then

$$\begin{aligned} \mathbb{P}(\Phi_{\mathcal{G}(\mathcal{K}(F))} < 10^{-3}) &= \mathbb{P}(\Phi_{\mathcal{G}(\mathcal{K}(G'))} < 10^{-3}) \\ &\leq \mathbb{P}(\Phi_{\mathcal{G}(\mathcal{K}(G))} < 10^{-3}, |\mathcal{K}(G)| \geq c_2 n \mid G \text{ simple}) + \mathbb{P}(|\mathcal{K}(G')| < c_2 n) \\ &\leq \mathbb{P}(\Phi_{\mathcal{G}(\mathcal{K}(G))} < 10^{-3}, |\mathcal{K}(G)| \geq c_2 n) / \mathbb{P}(G \text{ simple}) + \mathbb{P}(|\mathcal{K}(G')| < c_2 n) \\ &\leq (c_1 n)^{-50} / c_3 + n^{-100} = \mathcal{O}(n^{-50}). \end{aligned} \tag{A.1}$$

We now show how to move from the isoperimetric constant of the kernel to that of the graph. To do this, we first set up some notation. Let  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed; write  $\mathcal{G}$  for the giant,  $\mathcal{C}$  for its core and  $\mathcal{K}$  for its kernel. For an edge  $(x, y) \in E_{\mathcal{K}}$ , write  $\mathcal{E}_{x,y}$  for the set of vertices in  $\mathcal{G} \setminus \mathcal{K}$  that are the decoration of the edge  $(x, y)$ ; For a vertex  $z \in \mathcal{C}$ , write  $\mathcal{T}_z$  for the dangling tree hung from  $z$  and  $\mathcal{T}'_z = \mathcal{T}_z \cup \{z\}$ . For  $x \in \mathcal{K}$ , write  $\mathcal{D}_x = \mathcal{T}_x \cup (\cup_{y \sim_{\mathcal{K}} x} \mathcal{E}_{x,y})$ ; note that  $x \notin \mathcal{D}_x$ . Define the following properties for sets  $S \subseteq \mathcal{G}$ , which we denote  $(\dagger)$  and  $(\ddagger)$ , respectively:

$$\begin{aligned} \forall x \in S \cap \mathcal{K} \text{ we have } \mathcal{D}_x \subseteq S. & \tag{\dagger} \\ \forall x \in S \setminus \mathcal{K} \exists y \in S \cap \mathcal{K} \text{ with } x \in \mathcal{D}_y; & \tag{\ddagger} \end{aligned}$$

We do not consider  $\mathcal{G}$  to be a random graph, but rather consider each realisation of it. We assume that it has the following properties: there exists a constant  $c$  so that  $|\mathcal{G}| \geq cn$ ;  $|\mathcal{K}| > 0$ ; writing  $\Delta$  as an upper bound for the maximum degree and for the largest decoration, there exists a constant  $C$  so that we may take  $\Delta \leq C \log n$ .

First we show how to restrict attention to sets  $S$  with  $\mathcal{K} \not\subseteq S$ . Indeed, suppose  $S$  has  $\mathcal{K} \subseteq S$  and  $d(S) \leq |E_{\mathcal{G}}|$ . We claim that there exist (at least)  $3\Delta^2$  vertices  $z \in \mathcal{C}$  such that  $\mathcal{T}_z \not\subseteq S$  but  $\mathcal{T}_z \subseteq \mathcal{D}_x$  for some  $x \in \mathcal{K} \subseteq S$ . Indeed,  $\cup_{x \in \mathcal{K}} \mathcal{D}_x = \mathcal{G}$  and, and so by the definition of  $\mathcal{D}_x$  were this claim not the case then  $S = \mathcal{G} \setminus T$  where  $T$  is a union of subsets of at most  $2\Delta^2$  trees  $\mathcal{T}_z$ , and hence has size at most  $2\Delta^3$ , and so  $d(T) \leq 6\Delta^3$  (since the components are trees). But  $d(S) \leq |E_{\mathcal{G}}|$ , and so  $d(T) \geq |E_{\mathcal{G}}| \geq cn$ , contradicting  $d(T) \leq 6\Delta^3 \asymp (\log n)^3$ .

Given this property, define a new set,  $S'$ , from  $S$  in the following way. First choose an element  $x \in \mathcal{K}$  and remove it from our set. Since  $d_{\mathcal{G}}(x) \leq \Delta$  and the  $|\mathcal{D}_x| \leq \Delta$ , this removes at most  $\Delta^2$  vertices  $z \in \mathcal{C}$  with the above property; hence at least  $2\Delta^2$  remain. Choose  $2\Delta$  such  $z \in \mathcal{C}$ , say  $\{z_1, \dots, z_{2\Delta}\}$ . To each  $z_i$ , associate a vertex  $x_i \in \mathcal{K}$  with  $z_i \in \mathcal{D}_{x_i}$ ; note that the  $x_i$  need not be distinct. Now add to our set all the vertices of  $\mathcal{D}_{x_i}$  for each  $i$ . So  $S' = S \cup (\cup_{i=1}^{2\Delta} \mathcal{D}_{x_i}) \setminus x$ .

We now compare the outer boundary and edge-count of  $S$  and  $S'$ . Removing the first vertex can remove at most  $d_G(x) \leq \Delta$  from the number of edges in  $S$ , but adding in the decorations adds at least  $2\Delta$  (as  $2\Delta$  vertices are added). Hence  $d(S') - d(S) \geq 2\Delta$  (since  $d(S)$  is twice the number of edges in  $S$ ). Removing the first vertex can add at most  $d_G(x) \leq \Delta$  edges to the outer boundary, but adding the decorations removes at least  $2\Delta$  (since filling such a tree  $\mathcal{T}_z$  removes at least 1). Hence  $|\partial_G S| - |\partial_G S'| \geq \Delta$ . In particular,  $\Phi_G(S') \leq \Phi_G(S)$ .

Of course, it may be the case that  $d(S') > |E_G|$ . However, we note that  $d(S') \leq d(S) + 2\Delta^2 \leq \frac{3}{2}|E_G|$ . To allow such  $S'$ , we define  $\Phi'_G$  as follows:

$$\Phi'_G = \min\{S \subseteq \mathcal{G} \mid (0 < d(S) \leq |E_G|) \text{ OR } (|E_G| \leq d(S) \leq \frac{3}{2}|E_G| \text{ and } \mathcal{K} \not\subseteq S)\}.$$

Clearly  $\Phi_G \geq \Phi'_G$ . The above considerations show that we can then exclude  $S$  with  $\mathcal{K} \subset S$ :

$$\Phi'_G = \min\{S \subseteq \mathcal{G} \mid 0 < d(S) \leq |E_G|, \mathcal{K} \not\subseteq S\}.$$

From now on we shall assume that sets  $S$  in question do not have  $\mathcal{K} \subseteq S$ , and hence  $S^c \cap \mathcal{K} \neq \emptyset$ .

Now consider restricting attention to sets satisfying  $(\dagger)$  or both  $(\dagger)$  and  $(\ddagger)$ :

$$\begin{aligned} \Phi_G^\dagger &= \min\{\Phi_G(S) \mid 0 < d(S) \leq \frac{3}{2}|E_G|, \mathcal{K} \not\subseteq S, S \text{ satisfies } (\dagger)\}; \\ \Phi_G^{\dagger\ddagger} &= \min\{\Phi_G(S) \mid 0 < d(S) \leq \frac{3}{2}|E_G|, \mathcal{K} \not\subseteq S, S \text{ satisfies } (\dagger, \ddagger)\}. \end{aligned}$$

Since decorations are trees, since they are a (core-)path with dangling trees hung from each vertex of the path, we have  $\Phi_G = \Phi_G^\dagger$ . (Note that a set  $S$  with  $S \cap \mathcal{K} = \emptyset$  satisfies  $(\dagger)$  vacuously, and a set  $S$  satisfying  $(\ddagger)$  must have  $S \cap \mathcal{K} \neq \emptyset$ .)

We now show that it suffices to consider ‘connected’  $S$ , ie  $S \subseteq \mathcal{G}$  so that the induced graph  $G[S]$  is connected. Indeed, suppose that  $S = S_1 \cup S_2$  with  $S_1 \cap S_2 = \emptyset$  and no edges between  $S_1$  and  $S_2$ . Then  $d(S) = d(S_1) + d(S_2)$  and  $|\partial S| = |\partial S_1| + |\partial S_2|$ . Now, for any  $a, b, c, d > 0$ , we have

$$\frac{a+b}{c+d} \geq \min\{a/c, b/d\}, \quad \text{and hence} \quad \Phi(S) \geq \min\{\Phi(S_1), \Phi(S_2)\},$$

from which it follows that we may assume  $S$  is connected. Indeed, suppose  $a/c < b/d$ ; then

$$\frac{a+b}{c+d} \geq \frac{a+ad/c}{c+d} = \frac{a(c+d)}{c(c+d)} = \frac{a}{c}.$$

Recall that the largest decoration is of size at most  $\Delta$ . We claim that

$$\Phi_G = \Phi_G^\dagger = \min\{\Phi_G^{\dagger\ddagger}, 1/\Delta\}. \quad (\text{A.2})$$

Indeed, if  $S$  is connected and does not satisfy  $(\ddagger)$ , then, since decorations are connected at kernel-vertices, it can contain vertices of at most one decoration and  $S \cap \mathcal{K} = \emptyset$ . Since  $|\partial S| \geq 1$  for all  $S \notin \{\emptyset, \mathcal{G}\}$  and decorations are trees, such  $S$  have  $\Phi_G(S) \geq 1/\Delta$ . This proves the claim.

Observe that for  $S$  satisfying  $(\dagger, \ddagger)$  we have

$$|\partial_G S^c| = |\partial_G S| = |\partial_{\mathcal{K}}(S \cap \mathcal{K})| = |\partial_{\mathcal{K}}(S^c \cap \mathcal{K})|. \quad (\text{A.3})$$

We claim now that a set  $S$  satisfying  $(\ddagger)$  satisfies

$$d_G(S) \leq \Delta d_{\mathcal{K}}(S \cap \mathcal{K}) + \Delta |S \cap \mathcal{K}| \leq 2\Delta d_{\mathcal{K}}(S \cap \mathcal{K}). \quad (\text{A.4a})$$

Indeed, by associating vertices  $z \in \mathcal{E}_{x,y}$  with  $(x,y) \in E_{\mathcal{K}}$ , we get the first term; with the second term coming from the dangling tree attached to vertices of the kernel. We also claim that if  $S$  satisfies both  $(\dagger, \ddagger)$ , then  $S^c$  satisfies  $(\ddagger)$  unless  $\mathcal{K} \subseteq S$ , and hence the same inequality, ie

$$d_G(S^c) \leq \Delta d_{\mathcal{K}}(S^c \cap \mathcal{K}) + \Delta |S^c \cap \mathcal{K}| \leq 2\Delta d_{\mathcal{K}}(S^c \cap \mathcal{K}). \quad (\text{A.4b})$$

Indeed, take  $x \in S^c \setminus \mathcal{K}$ . First suppose that  $x \in \mathcal{D}_{y,z}$  for some  $(y,z) \in E_{\mathcal{K}}$ . We must then have  $y, z \in S^c$ , else by  $(\dagger)$  we would have  $x \in S$ . Hence we have  $x \in \mathcal{D}_y$  with  $y \in S^c \cap \mathcal{K}$ . The other

case is to suppose that  $x \in \mathcal{T}_y$  for some  $y \in \mathcal{K}$ . Again, we must have  $y \in S^c$ , else by  $(\dagger)$  we would have  $x \in S$ . Hence we have  $x \in \mathcal{T}_y \subseteq \mathcal{D}_y$  with  $y \in S^c \cap \mathcal{K}$ . This proves the claim.

Instead of considering  $\Phi_{\mathcal{G}}^\ddagger$ , for a graph  $H$  we consider

$$\Psi_H = \min\{\Psi_H(S) \mid S \notin \{\emptyset, H\}\} \quad \text{where} \quad \Psi_H(S) = \frac{2|\partial_H S||E_H|}{d_H(S)d_H(S^c)}.$$

Since any  $S \subseteq H$  has  $d_H(S^c) \leq 2|E_H|$ , we have  $\Psi_H \geq \Phi_H$ . If we define  $\Phi_H$  by considering sets  $S$  with  $0 < d(S) \leq 2(1-c)|E_H|$ , then  $c\Psi_H \leq \Phi_H$ . Using the above expression for  $|\partial_{\mathcal{G}} S|$  and inequalities for  $d_{\mathcal{G}}(S)$  and  $d_{\mathcal{G}}(S^c)$ , namely (A.3), (A.4a) and (A.4b), for  $S$  satisfying  $(\dagger, \ddagger)$ , we have

$$\Psi_{\mathcal{G}}(S) = \frac{2|\partial_{\mathcal{G}} S||E_{\mathcal{G}}|}{d_{\mathcal{G}}(S)d_{\mathcal{G}}(S^c)} \geq \frac{2|\partial_{\mathcal{K}}(S \cap \mathcal{K})||E_{\mathcal{K}}|}{4\Delta^2 d_{\mathcal{K}}(S \cap \mathcal{K})d_{\mathcal{G}}(S^c \cap \mathcal{K})} \geq 2\Psi_{\mathcal{K}}(S \cap \mathcal{K})/\Delta^2.$$

We have  $S \cap \mathcal{K}, S^c \cap \mathcal{K} \notin \{\emptyset, \mathcal{K}\}$  by recalling the following: we are assuming  $\mathcal{K} \not\subseteq S$ , which implies  $S \cap \mathcal{K} \neq \mathcal{K}$ , which in turn implies  $S^c \cap \mathcal{K} \neq \emptyset$ ; also  $S$  satisfies  $(\ddagger)$ , which implies  $S \cap \mathcal{K} \neq \emptyset$ , which in turn implies  $S^c \cap \mathcal{K} \neq \mathcal{K}$ . Hence we deduce that

$$\Psi_{\mathcal{G}}^\ddagger = \min\{\Psi_H(S) \mid S \notin \{\emptyset, \mathcal{G}\}, S \text{ satisfies } (\dagger, \ddagger)\} \geq 2\Psi_{\mathcal{K}}/\Delta^2 \geq 2\Phi_{\mathcal{K}}/\Delta^2.$$

In our calculation of  $\Phi_{\mathcal{G}}$ , we are using the condition  $0 < d(S) \leq |E_{\mathcal{G}}|$ , and hence

$$\Phi_{\mathcal{G}}^\ddagger \geq \frac{1}{2}\Psi_{\mathcal{G}}^\ddagger \geq \Phi_{\mathcal{K}}/\Delta^2.$$

We have shown that  $\mathbb{P}(\Phi_{\mathcal{K}} < 10^{-3}) = \mathcal{O}(n^{-500})$ , and hence the proposition follows.  $\square$

It remains to give the deferred proofs of Lemmas A.2, A.4 and A.5. (Recall that the proof of Lemma A.3 is deferred further, until the next section.)

**Proof of Lemma A.2.** For any connected graph  $H$ , any  $S \notin \{\emptyset, H\}$  has  $e_H(S) \geq 1$ , and hence  $\Phi_H(S) \geq 1/d_H(S)$ . Thus we have

$$\Phi_H \geq \min\{\Phi'_H, 10^{-5}\} \quad \text{where} \quad \Phi'_H = \min\{\Phi_H(S) \mid S \subseteq H, 10^5 \leq d(S) \leq |E_H|\}.$$

Since  $\mathcal{G}$  is a connected component of  $G \sim U(\mathbf{d})$ , we can lower bound

$$\Phi_{\mathcal{G}} \geq \min\{\Phi'_{\mathcal{G}}, 10^{-5}\} \quad \text{where} \quad \Phi'_{\mathcal{G}} = \min\{\Phi_{\mathcal{G}}(S) \mid S \subseteq \mathcal{G}, 10^5 \leq d(S) \leq |E_{\mathcal{G}}|\},$$

where we define  $\Phi_{\mathcal{G}}(S) = 0$  if  $S \supseteq \mathcal{G}$ . The advantage of doing this is that the set over which we are minimising is no longer a random set. It will be easier to work with  $G$ , rather than  $\mathcal{G}$ , since  $G$  has a nice form. Fix  $\beta = 99/100$ .

*Consider 'small' sets.* For  $s \in \mathbb{N}$ , define the collection of sets

$$\mathcal{S}_0(s) = \{S \mid |S| = s, 10^5 \leq d(S) \leq (\theta m)^{1/4}\}.$$

Let  $N_0(s)$  be the expected number of sets  $S \in \mathcal{S}_0(s)$  with  $I_G(S) \geq \beta d_G(S)$ . As in [1, Lemma 12],

$$\begin{aligned} N_S(s) &= \sum_{S \in \mathcal{S}_0(s)} \binom{d(S)}{\beta d(S)} \binom{\theta m/2}{\beta d(S)/2} / \binom{\theta m}{\beta d(S)} \\ &\lesssim \sum_{S \in \mathcal{S}_0(s)} \binom{d(S)}{\beta d(S)} \left(\frac{\beta d(S)}{\theta m}\right)^{\beta d(S)/2} \left(1 - \frac{\beta d(S)}{\theta m}\right)^{(\theta m - \beta d(S))/2} \\ &\lesssim \sum_{S \in \mathcal{S}_0(s)} \binom{d(S)e}{\beta d(S)}^{\beta d(S)} \left(\frac{\beta d(S)}{\theta m}\right)^{\beta d(S)/2} \\ &= \sum_{S \in \mathcal{S}_0(s)} \left(\frac{e^2 d(S)}{\beta \theta m}\right)^{\beta d(S)/2} \lesssim \sum_{S \in \mathcal{S}_0(s)} (5m^{-1})^{(1.1)d(S)/3}, \end{aligned}$$

since  $\theta \geq 3$  and  $\beta = 99/100$ . Note that there are at most  $m^s$  terms in the sum.

Consider  $s \geq 10^4$ . We use the bound  $d(S) \geq 3|S| = 3s$ , which is valid for all sets. Hence

$$N_0(s) \lesssim m^s \left(\frac{1}{4}m\right)^{-(1.1)s} \lesssim m^{-(0.05)s} \lesssim m^{-50}.$$

Consider  $s \leq 10^4$ . We use the bound  $d(S) \geq 10^5$  for  $S \in \mathcal{S}_G(s)$ . Hence

$$N_S(s) \lesssim m^s m^{-(1.1)10^5/3} \lesssim m^{-10^4}.$$

Combining these two then tells us that

$$\sum_{s \geq 1} N_S(s) = \mathcal{O}(m^{-50}).$$

Consider ‘large’ sets. For  $s \in \mathbb{N}$ , define the collection of sets

$$\mathcal{S}_1(s) = \{S \subseteq \mathcal{S} \mid |S| = s, (\theta m)^{1/4} < d(S) \leq \theta m/2\}.$$

Let  $N_1(s)$  be the expected number of sets  $S \in \mathcal{S}_1(s)$  with  $I_G(S) \geq \beta d_G(S)$ . Following precisely the ‘large sets’ part of the proof of [1, Lemma 12], it is shown that

$$\sum_{s \geq 1} N_1(s) = \mathcal{O}((9/10)^{m^{1/16}}) = \mathcal{O}(m^{-50}).$$

Combing ‘small’ and ‘large’ sets. Recall that  $\Phi_H(S) = 1 - I_H(S)/d_H(S)$ , and so if  $I_H(S) \leq \beta d_H(S)$  then  $\Phi_H(S) \geq 1/100$ . Hence, by Markov’s inequality, we have

$$\mathbb{P}(\Phi'_G < 1/100) \leq \sum_{s \geq 1} N(s) = \mathcal{O}(m^{-50}).$$

Since  $\Phi_G \geq \min\{\Phi'_G, 10^{-5}\}$ , the lemma follows.  $\square$

To prove Lemma A.4, we reference a result due to McKay and Wormald [21] on the probability of simplicity of a graph chosen uniformly at random given its degree sequences. The type of statement is well-known now, and similar statements can be found in [15, Theorem 7.12], [7] and [8, §2.4]. The statement below is not as general as that given in [21, Lemma 5.1], but rather slightly weaker result that will be easier to apply in our situation.

**Theorem A.6** ([21, Lemma 5.1]). *Let  $\mathbf{d} = (d_1, \dots, d_m)$  be a degree sequence, with  $M = \sum_i d_i$  even and  $\max_i d_i \leq M^{1/4}$ . Let  $N = \sum_i d_i(d_i - 1)$ . Let  $G \sim U(\mathbf{d})$ . Then*

$$\mathbb{P}(G \text{ simple}) \geq \exp\left(-\frac{1}{2}N/M - \frac{1}{4}N^2/M^2 - \frac{1}{2}N^2/M^3 + \mathcal{O}(1/M^{1/4})\right).$$

We show that the degree sequence of the core, with sufficient probability, satisfies  $N \asymp M \asymp n$ .

**Proof of Lemma A.4.** Write  $\mathcal{C} = \mathcal{C}_F$ , and let  $\mathbf{d} = (d_1, \dots, d_m)$  be as in the statement. Lemma A.3 says that the kernel at least  $cn$  vertices with probability at least  $1 - n^{-10}$ , for a positive constant  $c$ , and thus the same holds true for the core, and hence  $\mathbb{P}(M \leq cn) \leq n^{-10}$ . Observe that  $d_{\mathcal{C}}(x) \leq d_F(x)$  for all  $x \in \mathcal{C}$ . Hence, in the notation of Theorem A.6, we have  $N \leq \sum_{i=1}^n d_F(i)^2$ . This formulation has the advantage that we know the distribution of  $d_F(i)$ , namely  $\text{Bin}(n-1, \lambda/n)$ .

Moreover, we know the correlations between degrees: for an edge  $e$  of the complete graph  $K_n$ , write  $\eta_e = 1$  if the edge is open in  $F$  and  $\eta_e = 0$  if it is closed; then  $d_F(i) = \sum_{y \neq x} \eta_{x,y}$ , and the random variables  $(\eta_e)_{e \in E(K_n)}$  are independent  $\text{Bern}(\lambda/n)$  random variables. It is then straightforward to show that  $\mathbb{E}(N) \asymp n$ . By Markov’s inequality, we may then choose  $C$  sufficiently large so that  $\mathbb{P}(N \geq Cn) \leq \frac{1}{2}$ . The result now follows since

$$\mathbb{P}(G \text{ simple}) \geq \mathbb{P}(G \text{ simple} \mid N \leq Cn, M \geq cn)\mathbb{P}(N \leq Cn, M \geq cn). \quad \square$$

We now give the proof of Lemma A.5.

**Proof of Lemma A.5.** This proof uses similar ideas to that of [11, Lemma 4.2]. Write  $X_k$  for the number of sets  $S$  of size  $k$  such that  $G[S]$  is a tree and  $e_G(S) \leq 2$ . Then we have

$$\mathbb{E}(X_k) \leq \binom{n}{k} \left(\frac{\lambda}{n}\right)^{k-1} k^{k-2} \left(1 - \frac{\lambda}{n}\right)^{k(n-k) + \binom{k}{2} - k - 1}. \quad (\text{A.5})$$

For  $k \leq cn$ , with positive constant  $c$  sufficiently small, it is straightforward to see, using  $\binom{n}{k} \leq (ne/k)^k$ , that there exists a positive constants  $c'$  and  $C'$  so that

$$\mathbb{E}(X_k) \leq C' n e^{-c'k}.$$

For  $k \geq cn$ , we are slightly more careful. Note that (due to positive correlation) the number of isolated vertices is stochastically dominated from below by  $\text{Bin}(n, \frac{1}{2}e^{-\lambda})$ , and hence, using concentration of the Binomial (Hoeffding) we have

$$\mathbb{P}(|\mathcal{G}| \geq (1-c)n) = \mathcal{O}(n^{-30}) \quad \text{whenever } c \leq \frac{1}{4}e^{-\lambda}.$$

Since a decoration is a subset of the giant, we need only consider  $cn \leq k \leq (1-c)n$  with  $c$  as small as we desire. Write  $k = \alpha n$  with  $c \leq \alpha \leq 1-c$ . Using Stirling's formula in (A.5), we have

$$\begin{aligned} \mathbb{E}(X_k) &\lesssim n^{-5/2} (\alpha^\alpha (1-\alpha)^{1-\alpha})^{-n} (\alpha n)^{\alpha n} (\lambda/n)^{\alpha n} \exp(-\lambda \alpha (1-\alpha)n - \frac{1}{2} \lambda \alpha^2 n + \frac{3}{2} \lambda \alpha) \\ &\lesssim \exp(-nh(\alpha, \lambda)) \quad \text{where } h(\alpha, \lambda) = (1-\alpha) \log(1-\alpha) - \alpha \log \lambda + \lambda \alpha (1-\alpha) + \frac{1}{2} \lambda \alpha^2. \end{aligned}$$

It is elementary to check that there exists a constant  $\gamma$  so that  $h(\alpha, \lambda) > \gamma$  for all  $\alpha \in [c, 1-c]$ . Hence in this case also there exist positive constants  $c'$  and  $C'$  so that

$$\mathbb{E}(X_k) \leq C' n e^{-c'k}.$$

Taking  $k = 101 \log n / c'$  we deduce the lemma by Markov's inequality.  $\square$

## A.2 Typical Bounded Difference

McDiarmid [20] in 1989 used martingale methods to show concentration results for functions with bounded differences. Warnke [25] in 2016 extended this to only consider differences on a high probability set (hence the name ‘typical’ bounded differences). We first give a general formulation, and then show how to apply it to Erdős-Rényi graphs. We shall only formulate this for  $\{0, 1\}$ -valued random variables, as this is all that we shall need; [25] covers more general functions.

Let  $N \in \mathbb{N}$ , and let  $Z = (Z_1, \dots, Z_N)$  be a family of independent,  $\{0, 1\}$ -valued random variables with  $\mathbb{P}(Z_k = 1) = p_k$ . Let  $\mathcal{Z} \subseteq \{0, 1\}^N$ , and suppose that the function  $f : \{0, 1\}^N \rightarrow \mathbb{R}$  satisfies the following *typical Lipschitz condition*: there exist numbers  $(\alpha_k)_{k=1}^N$  and  $(\beta_k)_{k=1}^N$  such that, for each  $k$ , we have  $\alpha_k \leq \beta_k$  and, whenever  $z, z' \in \{0, 1\}^N$  with  $z_j = z'_j$  for all  $j \neq k$ ,

$$|f(z) - f(z')| \leq \begin{cases} \alpha_k & \text{if } z, z' \in \mathcal{Z}, \\ \beta_k & \text{otherwise;} \end{cases}$$

we say that  $f : \{0, 1\}^N \rightarrow \mathbb{R}$  satisfies  $\text{TL}(\mathcal{Z}, \alpha, \beta)$ .

Warnke [25] showed in essence that, providing we choose  $\mathcal{Z}$  to be such that  $\mathbb{P}(Z \notin \mathcal{Z})$  is sufficiently small, we can restrict our attention from worst-case differences to ‘typical’ differences.

**Theorem A.7** ([25, Theorem 2 and Remark 10]). *Consider  $Z$  and  $\mathcal{Z}$  as above. Suppose  $f : \{0, 1\}^N \rightarrow \mathbb{R}$  satisfies  $\text{TL}(\mathcal{Z}, \alpha, \beta)$ . For any numbers  $(\gamma_k)_{k=1}^N \subseteq (0, 1]^N$ , we have*

$$\mathbb{P}(|f(Z) - \mathbb{E}(f(Z))| \geq R) \leq 2 \exp\left(-\frac{R^2}{2 \sum_{k=1}^N p_k (1-p_k) (\alpha_k + e_k)^2 + 2CR/3}\right) + \Gamma Q \mathbb{P}(Z \notin \mathcal{Z}),$$

where  $e_k = \gamma_k(\beta_k - \alpha_k)$ ,  $C = \max_k(\alpha_k + e_k)$ ,  $q_k \leq \min_b \mathbb{P}(Z_k = b)$ ,  $\Gamma = \sum_k \gamma_k^{-1}$  and  $Q = \sum_k q_k^{-1}$ .



We now formulate this in terms of an Erdős-Rényi graph with  $n$  vertices and edge probability  $p = \lambda/n$  with  $\lambda > 1$  fixed. We take  $N = \binom{n}{2}$  to be the number of edges in the complete graph on  $n$  vertices, and  $Z_k$  to be the state of the  $k$ -th edge, so  $q_k = \lambda/n$  for all  $k$ . Let us also restrict attention to bounded  $f$ , say  $f : \{0, 1\}^N \rightarrow [0, M]$ ;  $M$  is allowed to depend on  $n$ . We then set  $\beta_k = M$  and  $\gamma_k = (Mn)^{-1}$  for all  $k$ . Since  $\alpha_k \leq \beta_k$ , we then have  $0 \leq \alpha_k + e_k \leq \alpha_k + 1/n$ , and so  $C \leq \max_k \alpha_k + 1/n$ . From this we obtain the following corollary.

**Corollary A.8.** *Consider an  $\text{ER}(n, \lambda/n)$  graph. Let  $Z_k$  be the state of the  $k$ -th edge. Suppose  $f : \{0, 1\}^N \rightarrow [0, n]$  satisfies  $\text{TL}(\mathcal{Z}, \alpha, n)$ . We have*

$$\mathbb{P}(|f(Z) - \mathbb{E}(f(Z))| \geq R) \leq 2 \exp\left(-\frac{R^2}{2\lambda n^{-1} \sum_{k=1}^N (\alpha_k + 1/n)^2 + 2CR/3}\right) + n^6 \mathbb{P}(Z \notin \mathcal{Z}),$$

where  $C = \max_k \alpha_k + 1/n$ .

First we note some preliminary results which will go into the definition of  $\mathcal{Z}$ . As proved at the start of this section (§A) there exist positive constants  $c_0$  and  $C_0$  so that the probability there exists a component with size in  $[C_0 \log n, c_0 n]$  is  $\mathcal{O}(n^{-100})$ . A minor adaptation to the end of the proof of [12, Theorem 2.14] replacing  $c_1 = c - \log n/n$  by  $c_1 = c - (\log n)^2/n$ , shows that the probability there exists more than one component of size at least  $c_0 n$  is  $\mathcal{O}(n^{-100})$ . Fix this  $c_0$  and  $C_0$  for the rest of the paper.

**Lemma A.9.** *Suppose  $G \sim \text{ER}(n, \lambda/n)$ , with  $\lambda > 1$  fixed, and write  $\mathcal{G}$  for its largest component (breaking ties arbitrarily). There exists a positive constant  $c$  so that*

$$\mathbb{P}(|\mathcal{G}| \leq cn) = \mathcal{O}(n^{-90}).$$

**Proof.** By counting the expected number of subcritical components, we see that the expected size of the largest component is order  $n$ , say at least  $2cn$  for a positive constant  $c$ . For the details of this calculation, see the proof of [12, Theorem 2.14].

Let  $\mathcal{Z}$  be the set of all graphs with no components with size in  $[C_0 \log n, c_0 n]$  and at most one component with size at least  $c_0 n$ . Note also that  $|\mathcal{G}| \leq n$ . Now consider a graph  $G \in \mathcal{Z}$ , and consider adding/removing an edge to obtain a new graph  $G'$ , and require that  $G' \in \mathcal{Z}$ . Adding an edge can only increase  $|\mathcal{G}|$ , and by at most  $C_0 \log n$ , since no non-giant component has size larger than this. Removing an edge can only decrease  $|\mathcal{G}|$ , and by at most  $C_0 \log n$ , since the new component must have size at most this. Hence in the typical bounded difference formulation we may take  $\alpha_k = C_0 \log n$  for all  $k$ ; this gives  $C \leq 2C_0 \log n$  and  $\sum_{k=1}^N (\alpha_k + 1/n)^2 \leq C_0^2 n^2 (\log n)^2$ . Take  $R = cn$ . Hence overall we have

$$\mathbb{P}(|\mathcal{G}| \leq cn) \leq 2 \exp\left(-\frac{c^2 n^2}{2\lambda C_0^2 n (\log n)^2 + 2cC_0 n \log n}\right) + n^6 \mathcal{O}(n^{-100}) = \mathcal{O}(n^{-90}). \quad \square$$

We now prove Lemmas A.3 and A.5. The proofs have two parts: first we find the expectation; then we show concentration. To find the expectation, we use the contiguous model of Ding, Lubetzky and Peres, which was stated in Theorem A.13; concentration will be proved using the typical bounded differences method above.

One part of the proof of Lemma A.3 will rely on a short lemma, which we state and prove now.

**Lemma A.10.** *Write  $\mathcal{D}$  for the set of all edges of the giant whose removal causes the giant to split into two components (ie edges that are a bridge for the giant). For  $e \in \mathcal{D}$ , write  $R_e$  for the size of the smaller component resulting from the removal of  $e$  (breaking ties arbitrarily). Write  $R = \max_e R_e$ . There exists a constant  $C$  so that*

$$\mathbb{P}(R > C \log n) = \mathcal{O}(n^{-100}).$$

**Proof.** Let  $F$  be a graph, and let  $F \setminus e$  denote the graph  $F$  with the edge  $e$  closed. Let  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed. Assuming  $e$  is open in  $F$ , using independence of the edges we have

$$\mathbb{P}(G = F) = \mathbb{P}(G = F \setminus e) \cdot p/(1-p).$$

Now consider removing an edge  $e \in \mathcal{D}$ . Removing  $e$  creates a new component. Since the probability that more than one component of size greater than  $C_0 \log n$  exists is  $\mathcal{O}(n^{-100})$ , using the above inequality we deduce the lemma in the following way.

Let  $\mathcal{A}$  be the set of all graphs with the property that its largest component (breaking ties arbitrarily) has an edge whose removal disconnects the component leaving two components each of size greater than  $C_0 \log n$ ; associate to each graph  $F \in \mathcal{A}$  an edge with the above removal property, and call this edge  $e_F$ . Let  $\mathcal{B}$  be the set of all graphs with two components of size greater than  $C_0 \log n$ . Then we have

$$\{F \setminus e_F \mid F \in \mathcal{A}\} \subseteq \mathcal{B}.$$

Hence using the first inequality we obtain

$$\begin{aligned} \mathbb{P}(G \in \mathcal{A}) &= \sum_{F \in \mathcal{A}} \mathbb{P}(G = F) = \sum_{F \in \mathcal{A}} \mathbb{P}(G = F \setminus e_F) \cdot p/(1-p) \\ &\leq \sum_{H \in \mathcal{B}} \mathbb{P}(G = H) \cdot p/(1-p) = \mathbb{P}(G \in \mathcal{B}) \cdot p/(1-p). \end{aligned}$$

Finally, if a graph  $F$  has  $R > C_0 \log n$ , then  $F \in \mathcal{A}$ . We know that  $\mathbb{P}(G \in \mathcal{B}) = \mathcal{O}(n^{-100})$ , and hence we have proved the lemma.  $\square$

An immediate corollary of this is that any  $x \in \mathcal{G}$  has at most  $C \log n$  removal edges.

**Corollary A.11** (Maximum Removal Edges, (v)). *There exists a constant  $C$  so that*

$$\mathbb{P}(\exists x \in \mathcal{G} \text{ s.t. } R(x) > C \log n) = \mathcal{O}(n^{-100}).$$

**Proof.** Recall that we write  $\mathcal{R}(x)$  for the set of removal edges for  $x$ . Then  $\mathcal{R}(x)$  defines a ‘dangling component’ in the sense of Lemma A.10; the result follows.  $\square$

**Remark A.12.** Observe that the proof of Lemma A.10 immediately extends in the following way: instead of considering sets  $S \subseteq \mathcal{G}$  which can be disconnected from  $\mathcal{G}$  by the removal of one edge, we require two edges to be removed.  $\triangle$

To prove this lemma, instead of considering the kernel directly, we construct a ‘contiguous model’ of the giant, as given in [9, Theorem 1]; we state the precise formulation now.

**Theorem A.13** ([9, Theorem 1]). *Write  $\mathcal{G}$  for the (random) largest component of an  $\text{ER}(n, \lambda/n)$  graph (breaking ties arbitrarily), with  $\lambda > 1$  fixed. Let  $\theta$  be the unique solution in  $[0, 1]$  of  $\theta e^{-\theta} = \lambda e^{-\lambda}$ ; write  $\chi = \lambda - \theta$ . Generate  $\tilde{\mathcal{G}}$  in the following way.*

- (i) *Constructing the kernel. Let  $X \sim N(\chi, 1/n)$  and, independently of  $X$ , let  $D_i \sim \text{Po}(X)$  for  $i = 1, \dots, n$ , conditioned that  $\sum_i D_i \mathbf{1}(D_i \geq 3)$  is even (where  $\text{Po}(\xi) = 0$  if  $\xi \leq 0$ ). Let*

$$N_k = |\{i \mid D_i = k\}| \quad \text{and} \quad N = \sum_{k \geq 3} N_k.$$

*Choose  $\tilde{\mathcal{K}}$  uniformly at random among all (multi-)graphs with  $N_k$  vertices of degree  $k$  for  $k \geq 3$  (and hence with a total of  $N$  vertices).*

- (ii) *Expanding to the core. Obtain  $\tilde{\mathcal{C}}$  by replacing the edges of  $\tilde{\mathcal{K}}$  by iid paths of geometric length with parameter  $1 - \theta$  (and mean  $1/\theta$ ).*
- (iii) *Hanging dangling trees. Obtain  $\tilde{\mathcal{G}}$  by attaching an independent Galton-Watson tree with Poisson( $\theta$ ) offspring to each vertex of  $\tilde{\mathcal{C}}$ .*

Then, for any set  $\mathcal{A}$  of graphs,  $\mathbb{P}(\tilde{\mathcal{G}} \in \mathcal{A}) \rightarrow 0$  implies  $\mathbb{P}(\mathcal{G} \in \mathcal{A}) \rightarrow 0$ .

**Proof of Lemma A.3.** We start by finding the expectation via the contiguous model of Theorem A.13. In that notation, the size of the kernel is exactly  $N$ . Note that  $X \sim N(\chi, 1/n)$  and, independently of  $X$ ,  $D_i \sim \text{Po}(X)$  for  $i = 1, \dots, n$  conditioned that  $\sum_i D_i \mathbf{1}(D_i \geq 3)$  is even, and  $N = \sum_i \mathbf{1}(D_i \geq 3)$ . Hence there exists a positive constant  $c$  so that  $\mathbb{P}(\text{Po}(X) \geq 3) \geq 2c$ . Then

$$\mathbb{P}(N \leq cn \mid \sum_i D_i \mathbf{1}(D_i \geq 3) \text{ even}) \leq \mathbb{P}(\sum_i \mathbf{1}(D_i \geq 3) \leq cn) / \mathbb{P}(\sum_i D_i \mathbf{1}(D_i \geq 3) \text{ even})$$

$$\begin{aligned}
&= \mathbb{P}(\text{Bin}(n, \mathbb{P}(D_i \geq 3)) \leq cn) / \mathbb{P}(\sum_i D_i \mathbf{1}(D_i \geq 3) \text{ even}) \\
&\leq \mathbb{P}(\text{Bin}(n, 2c) \leq cn) / \mathbb{P}(\sum_i D_i \mathbf{1}(D_i \geq 3) \text{ even}) \\
&\leq \exp(-2c^2 n) / \mathbb{P}(\sum_i D_i \mathbf{1}(D_i \geq 3) \text{ even}),
\end{aligned}$$

with the final inequality holding by Hoeffding. We now show that the denominator is not too small. By conditioning that  $\sum_{i=1}^{n-1} D_i \mathbf{1}(D_i \geq 3)$  is either odd or even and taking a minimum, we see that

$$\mathbb{P}(\sum_{i=1}^n D_i \mathbf{1}(D_i \geq 3) \text{ even}) \geq \min\{\mathbb{P}(D_1 = 3), \mathbb{P}(D_1 = 4)\}.$$

Hence we see that there exists a positive constant  $c'$  so that  $\mathbb{P}(\sum_i D_i \mathbf{1}(D_i \geq 3) \text{ even}) \geq c'$ . From this we deduce that

$$\mathbb{P}(|\mathcal{K}_{\mathcal{G}}| \leq cn) = o(1), \quad \text{and hence} \quad \mathbb{E}(|\mathcal{K}_{\mathcal{G}}|) \geq \frac{1}{2}cn,$$

where  $\mathcal{G}$  is the largest component of an  $\text{ER}(n, \lambda/n)$  graph, with  $\lambda > 1$  fixed.

We now show concentration about the mean. To do this, we let  $f(Z) = |\mathcal{K}(\mathcal{G}(G))|$  where we identify the vector  $Z$  in  $\{0, 1\}^N$  with a graph  $G$ . Note that  $f$  is increasing in the sense that if an edge is added to the largest component then (assuming there was no tie for largest component) the size of the kernel can only increase. Let  $\mathcal{Z}$  be the set of all graphs with no components with size in  $[C_0 \log n, c_0 n]$  and precisely one component with size at least  $c_0 n$ .

We now consider how the size of the kernel can change when the state of an edge is changed. Write  $K$  for the original kernel and  $K'$  for the one obtained by the removal. We show that  $\|K\| - \|K'\| \lesssim \log n$ ; the rest of the proof follows exactly as in Lemma A.9.

*Consider removing the edge  $(x, y)$ .*

Removing an edge cannot increase the size of the kernel (using the uniqueness of the giant).

Suppose first that the removal of the edge does not disconnect the kernel (ie the edge is not a *bridge* for the kernel). If the removed edge lies in a dangling tree, then this has no effect on the size of the kernel. Suppose the edge lies in a core-path, and so removing it removes the corresponding kernel-edge; let  $z$  be an endpoint of the *kernel-edge* corresponding to  $(x, y)$ . If  $d_K(z) \geq 4$ , then  $z$  remains in the kernel. If  $d_K(z) = 3$ , then the other two core-paths incident to  $z$  are joined into a single core-path, and the vertex  $x$  is removed from the kernel but all other vertices remain. Hence we see that removing such an edge can remove at most 2 vertices from the kernel.

Now suppose that the edge is a bridge for the kernel. Since the graph obtained by removing  $(x, y)$  must be in  $\mathcal{Z}$ , the disconnected component must have size at most  $C_0 \log n$ . In addition to these, as in the above case, one further vertex is removed from the giant (through the joining of two core-paths at this vertex).

*Consider adding the edge  $(x, y)$ .*

Adding an edge cannot decrease the size of the kernel (using the uniqueness of the giant).

If the added edge has both ends in the same dangling tree, then this can add at most two vertices to the kernel, namely the most recent common ancestor of  $x$  and  $y$  in the dangling tree and its root.

Similarly, if  $x$  and  $y$  are in different dangling trees (allowing  $x$  and  $y$  to be the roots), then two vertices are added, namely the roots of the dangling trees in which  $x$  and  $y$  lie; that is, writing  $r_z$  for the root of the dangling tree of  $z$ , we have  $K' = K \cup \{r_x, r_y\}$ .

Connecting a subcritical component to the giant can add at most  $C_0 \log n + 1$ , in the same way as removing a bridge can remove this many.  $\square$

Having now completed the proof of the expansion properties (vii), we move onto calculating the number of degree 1 vertices in the giant. (This result is proved independently of the expansion properties, except for quoting a small calculation that we did in Lemma A.3.)

**Lemma A.14** (Degree 1 Vertices in the Giant, (iv)). *Suppose  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed, and write  $N$  for the number of degree 1 vertices in the giant. There exists a positive constant  $c$  so that*

$$\mathbb{P}(N \leq cn) = \mathcal{O}(n^{-100}).$$

**Proof.** First we lower bound the expected number of degree 1 vertices in the giant in the contiguous model; write  $\tilde{N}$  for  $N$  in the contiguous model. As in the proof of Lemma A.3, there exists a constant  $c_1$  so that  $\mathbb{P}(|\mathcal{K}(\tilde{\mathcal{G}})| \leq c_1 n) = o(1)$ . Note that PGW( $\theta$ ) trees are hung from the vertices of the kernel (as well as of the other vertices of the core). Since  $\theta$  is a constant, the probability that the PGW-tree is just the root and one offspring is a constant. Hence there exists a positive constant  $c_2$  so that  $\mathbb{P}(\tilde{N} \leq c_2 n) = o(1)$ . Hence, as in the proof of Lemma A.3, we have  $\mathbb{E}(N) \geq \frac{1}{2}c_2 n$ .

We now consider the effect on  $N$  of adding/removing an edge. By recalling that both the original and new graph must have subcritical components of size at most  $C_0 \log n$ , we see that the largest change is at most  $C_0 \log n$ . The rest of the proof follows exactly as in Lemma A.9.  $\square$

Finally we prove (vi) on the number of vertices ‘far away from the core’. Recall the definitions of  $R$  and  $\mathcal{W}$  from Notation 2.1, as well as that of  $\log_{(M)}$ .

**Proposition A.15** (Vertices Far from the Core, (vi)). *Suppose  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed. There exists a constant  $C$  so that, for all  $M$ , for all  $n$  sufficiently large, we have*

$$\mathbb{P}(|\mathcal{G} \setminus \mathcal{W}^M|/|\mathcal{G}| \geq (\log_{(M-1)} n)^{-4}) \leq Cn^{-50}.$$

**Remark A.16.** The proposition needs  $n$  to be sufficiently large depending on  $M$ , say  $n \geq n_0(M)$ . If we choose  $\omega_*$  such that  $n_0(\omega_*(n)) \leq n$ , then we can apply the above proposition simultaneously for all  $M \leq \omega_*(n)$ . Further we can require  $\omega_*(n) \leq n$ , and then deduce (vi) by the union bound.  $\triangle$

We prove this by a sequence of lemmas. In essence, we split our consideration into dangling trees and parts of the giant that can be removed by removing a kernel-edge. The latter parts we think of as ‘bad’ parts of the giant; we consider these first.

For a graph  $F$ , write  $\mathcal{P}_F$  for the set of all  $x \in \mathcal{G}(F)$  with the following property: there exists an edge of the giant whose removal disconnects the giant, leaving  $x$  in the smaller component (breaking ties arbitrarily), and this smaller component is not a tree; that is, the removed edge disconnects the kernel. We call  $\mathcal{P}_F$  the set of *peninsular vertices* of the giant of  $F$ .

**Lemma A.17.** *Suppose  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed. There exists a constant  $C$  so that*

$$\mathbb{P}(|\mathcal{P}_G| > C \log n) = \mathcal{O}(n^{-50}).$$

**Proof.** We first bound the expectation; we then use the typical bounded differences methodology.

We consider the kernel, and consider the expected number of vertices of the kernel that can be removed from the kernel by removing one kernel-edge. We showed in the proof of Proposition A.1, specifically (A.1), that the kernel is an expander, ie has isoperimetric constant  $\Phi = \Theta(1)$ , with sufficiently high probability, and so any components that are ‘broken off’ will have size in the kernel at most  $1/\Phi = \Theta(1)$ . Moreover, any component of the kernel that is broken off cannot be a tree (since the kernel has minimal degree 3). Using similar arguments as in Lemma A.10, we now show that  $\mathbb{E}(|\mathcal{P}_G|) = \mathcal{O}(\log n)$ .

For a graph  $F$ , let  $\mathcal{E}_F$  be the set of all edges  $e$  with the property that removing  $e$  disconnects the kernel of the largest component (breaking ties arbitrarily). Write  $F \setminus \mathcal{E}_F$  for the graph with all edge  $e \in \mathcal{E}_F$  closed. Then, with  $p = \lambda/n$ , we have

$$\mathbb{P}(G = F) = \mathbb{P}(G = F \setminus \mathcal{E}_F) \cdot (p/(1-p))^{|\mathcal{E}_F|} \leq \mathbb{P}(G = F \setminus \mathcal{E}_F).$$

Write  $|\mathcal{T}'_F|$  for the number of vertices in non-tree subcritical components. Then observe that  $|\mathcal{P}_F| \leq D_F \cdot |\mathcal{T}'_{F \setminus \mathcal{E}_F}|$  where  $D_F$  is the size of the largest decoration. Write  $\mathcal{A}$  for the set of all graphs  $F$  with  $D_F \leq C \log n$ , with  $C$  coming from Lemma A.5. We then have

$$\mathbb{E}(|\mathcal{P}_G| \mathbf{1}(G \in \mathcal{A})) = \sum_{F \in \mathcal{A}} |\mathcal{P}_F| \mathbb{P}(G = F)$$

$$\begin{aligned}
&\leq \sum_{F \in \mathcal{A}} |\mathcal{P}_F| \mathbb{P}(G = F \setminus \mathcal{E}_F) \\
&\leq \sum_{F \in \mathcal{A}} D_F |\mathcal{T}'_{F \setminus \mathcal{E}_F}| \mathbb{P}(G = F \setminus \mathcal{E}_F) \\
&\leq C \log n \sum_H |\mathcal{T}'_H| \mathbb{P}(G = H) \\
&= C \log n \mathbb{E}(|\mathcal{T}'_G|).
\end{aligned}$$

It is a standard calculation to show that  $\mathbb{E}(|\mathcal{T}'_G|) = \Theta(1)$  (see the proof of [12, Theorem 2.4], for example). Lemma A.5 says that  $\mathbb{P}(G \notin \mathcal{A}) = \mathcal{O}(n^{-100})$  and  $|\mathcal{P}_G| \leq n$ , hence  $\mathbb{E}(|\mathcal{P}_G|) = \mathcal{O}(\log n)$ .

We now apply the typical bounded differences methodology. Recall Lemma A.10, which says that the largest set which can be disconnected from the giant by removing one edge has size at most  $C \log n$ , with probability  $1 - \mathcal{O}(n^{-100})$ .

Let  $\mathcal{Z}$  the set of all graphs with this connectivity property, with no components with size in  $[C_0 \log n, c_0 n]$  and with precisely one with size at least  $c_0 n$ , which we call the giant. We now consider changing the state of a single edge. We show that the change is at most  $(2C + C_0) \log n$ ; the rest of the proof follows exactly as in Lemma A.9.

*Consider removing an edge.* A set which was previously connected to the giant via two edges can be reduced to only one edge, and so at most  $C \log n$  can be added. A set which was previously connected to the giant via one edge can be removed, and so at most  $C \log n$  can be removed.

*Consider adding an edge.* A subcritical component can be connected to a vertex already in  $\mathcal{P}$ , adding at most  $C_0 \log n$  vertices to  $\mathcal{P}$ . Also, the edge could make a dangling tree into a dangling non-tree; this component could then be disconnected by removing one edge, and so must have size at most  $C \log n$ . If the edge connects such a bad component to another vertex in the giant, then this can only remove vertices from  $\mathcal{P}$ . It can remove two such bad components, each of which must have size at most  $C \log n$ , and so remove at most  $2C \log n$  vertices from  $\mathcal{P}$ .  $\square$

We consider now the vertices in dangling trees that are ‘far from the core’.

**Lemma A.18.** *Suppose  $G \sim \text{ER}(n, \lambda/n)$  with  $\lambda > 1$  fixed. There exist positive constants  $c$  and  $C$  so that, writing  $T_k$  for the number of dangling trees of size  $k$ , for  $K \leq C_0 \log n$ , we have*

$$\mathbb{E}(\sum_{k \geq K} k T_k) \leq C e^{-cK}.$$

**Proof.** Recall that the largest dangling tree is size order  $\log n$  with high probability. Choose  $c$  sufficiently small so that  $c \cdot C_0 \leq 10$ ; then  $\exp(-cK) \geq n^{-10}$  for all  $K \leq C_0 \log n$ .

We upper bound  $T_k$  by the number of tree components of size  $k$  with exactly 1 edge coming out. Using standard combinatorial counting, we then see that

$$\mathbb{E}(T_k) \lesssim n(\lambda e^{1-\lambda})^k.$$

The lemma then follows.  $\square$

We now have the ingredients required to prove Proposition A.15.

**Proof of Proposition A.15.** As is often the case, we first look at expectation and then use the typical bounded differences methodology.

Write  $\mathcal{T}_k$  for the number of vertices in dangling trees of size at least  $k$ . (So  $\mathcal{T}_k = \sum_{k \geq K} k T_k$  in the notation of Lemma A.18.) Write

$$\mathcal{V}^K = \{x \in \mathcal{G} \mid R(x) \geq K\}; \quad \text{so} \quad \mathcal{V}^{\log_{(\kappa)} n} = \mathcal{W}^K.$$

For  $K \leq C_0 \log n$ , we upper bound

$$|\mathcal{V}^K| \leq \mathcal{T}_K + |\mathcal{P}|.$$

In Lemma A.17, we showed that  $|\mathcal{P}| \leq C \log n$  with probability  $1 - \mathcal{O}(n^{-50})$ . Hence this combined with Lemma A.18 says that

$$\mathbb{E}(|\mathcal{V}^K|) \leq C n e^{-cK} + C' \log n.$$

In particular, if we take  $K = 5c^{-1} \log_{(M)} n$ , then we see that

$$\mathbb{E}(|\mathcal{W}^M|) \leq \frac{1}{2} n (\log_{(M-1)} n)^{-4};$$

moreover, we may allow  $M$  to depend on  $n$  if we require  $M$  so that  $\log_{(M)} n \geq 1$  (ie  $M \leq \log^* n - 1$ ).

We now use the typical bounded differences methodology, applied to  $\mathcal{T}_k$ . Let  $\mathcal{Z}$  the set of all graphs with no components with size in  $[C_0 \log n, c_0 n]$ , with precisely one with size at least  $c_0 n$ , which we call the giant, and largest decoration to the giant of size at most  $C_0 \log n$ . We now consider changing the state of a single edge.

*Consider adding an edge.* If both endpoints are in the giant, then this can only decrease the number of vertices in dangling trees (or leave unchanged); two full dangling trees can be removed, removing at most  $2C_0 \log n$  vertices. If one endpoint is outside the giant, then a subcritical component can be added, adding at most  $C_0 \log n$  vertices.

*Consider removing an edge.* If the core is unchanged, then a dangling tree has been ‘pruned’, and hence at most  $C_0 \log n$  vertices can be removed. If the kernel is not disconnected, but a kernel-edge is removed, then all the vertices that were in this core-path then belong to dangling trees in the new graph, and hence at most  $C_0 \log n$  can be added. If the kernel is disconnected, then the vertices on the core-edge, of which there are at most  $C_0 \log n$ , can be added; all the vertices in the new subcritical component are removed, and so at most  $C_0 \log n$  are removed.

Hence  $\mathcal{T}_k$  can change by at most  $3C_0 \log n$ . Taking  $R = n/(\log n)^{-6}$  in Corollary A.8 we obtain

$$\mathbb{P}(\mathcal{T}_k \geq n(Ce^{-ck} + (\log n)^{-6})) = \mathcal{O}(n^{-90}).$$

Note that  $R \gg \log n$ . This proves the proposition. □