

# A COMPARISON PRINCIPLE FOR RANDOM WALK ON DYNAMICAL PERCOLATION

BY JONATHAN HERMON<sup>1</sup> AND PERLA SOUSI<sup>2</sup>

<sup>1</sup>Mathematics Department, University of British Columbia, [jhermon@math.ubc.ca](mailto:jhermon@math.ubc.ca)

<sup>2</sup>Statistical Laboratory, Cambridge University, [p.sousi@statslab.cam.ac.uk](mailto:p.sousi@statslab.cam.ac.uk)

We consider the model of random walk on dynamical percolation introduced by Peres, Stauffer and Steif in (*Probab. Theory Related Fields* **162** (2015) 487–530). We obtain comparison results for this model for hitting and mixing times and for the spectral gap and log-Sobolev constant with the corresponding quantities for simple random walk on the underlying graph  $G$ , for general graphs. When  $G$  is the torus  $\mathbb{Z}_n^d$ , we recover the results of Peres et al., and we also extend them to the critical case. We also obtain bounds in the cases where  $G$  is a transitive graph of moderate growth and also when it is the hypercube.

**1. Introduction.** In this paper we consider the model of random walk on a dynamically evolving environment introduced in [28]. Fix a base graph  $G = (V, E)$ , and let each edge  $e$  refresh at rate  $\mu$  to open with probability  $p$  and close with probability  $1 - p$  independently of other edges and previous states of the same edge. Let  $X$  be a continuous time random walk that moves as follows: at rate 1 it chooses one of its neighbours uniformly at random and only jumps there if the edge connecting the neighbour to its current location is open. Otherwise, it stays in place. We denote the state of the full system at time  $t$  by  $(X_t, \eta_t)$ , where  $X_t \in V$  and  $\eta_t \in \{0, 1\}^E$  with 0 representing a closed edge and 1 an open one. We refer to  $\eta_t$  as the environment at time  $t$ . We emphasise that  $(X_t, \eta_t)_{t \geq 0}$  is Markovian, while the location of the walker  $(X_t)_{t \geq 0}$  is not. One readily checks that  $\pi_{\text{full}, p} = \pi \times \pi_p$  is the unique stationary distribution and that the process is reversible; here,  $\pi$  is the degree biased distribution on  $V$ , that is,  $\pi(x) = \deg(x)/(2|E|)$  for all  $x$  and  $\pi_p$  is product measure of  $\text{Ber}(p)$  on the edges. Moreover, even if the environment process  $\{\eta_t\}_{t \geq 0}$  is fixed,  $\pi$  is a stationary distribution for the resulting time inhomogeneous Markov process  $(X_t)_{t \geq 0}$ .

There has been a lot of interest recently in studying these processes. The case where  $G = \mathbb{Z}_n^d$  was studied in [27, 28]. The subcritical regime seems to be fully understood [28], while the supercritical case still poses challenges. In [3], the authors established precise mixing time results for the nonbacktracking random walk on a dynamic configuration model, and in [30] the authors studied the case where  $G$  is the complete graph. In this paper we study mixing and hitting times for random walk on dynamical percolation on general graphs. We develop general machinery that allows for a comparison of these quantities with the corresponding ones for simple random walk (SRW) on the base graph  $G$ . Recall that the SRW on  $G = (V, E)$  is a Markov chain on  $V$  with transition probabilities given by  $P(x, y) = \mathbf{1}(\{x, y\} \in E) / \deg(x)$ , where  $\deg(x)$  is the degree of  $x$  (i.e., at each step the walk picks a neighbour uniformly at random and jumps to it). We emphasise that, from now on, whenever we write SRW, we refer to simple random walk on the (static) graph  $G$ , that is, where all edges in  $E$  are present. Below we shall consider its continuous-time version with jump rate 1. We note that our upper bounds on the hitting and mixing times hold for all  $\mu$  and all  $p \in (0, 1]$  with

Received April 2019; revised January 2020.

*MSC2020 subject classifications.* 60F05, 60G50.

*Key words and phrases.* Dynamical percolation, mixing times, hitting times, spectral profile.

no difference between the subcritical and supercritical regimes. This is in sharp contrast to previous works.

Let  $P$  be a transition matrix with stationary distribution  $\pi$ . We define the total variation and  $L_\infty$  mixing times as follows:

$$t_{\text{mix}}(\varepsilon) = \min \left\{ t \geq 0 : \max_x \|P^t(x, \cdot) - \pi\|_{\text{TV}} \leq \varepsilon \right\},$$

$$t_{\text{mix}}^{(\infty)}(\varepsilon) = \min \left\{ t \geq 0 : \max_x \|P^t(x, \cdot) - \pi\|_{\infty, \pi} \leq \varepsilon \right\},$$

where the total variation and  $L_\infty$  norms of a signed measure  $a$  are given by  $\|a\|_{\text{TV}} = \frac{1}{2} \sum_x |a(x)|$  and  $\|a\|_{\infty, c} = \max_x |a(x)/c(x)|$ . We are primarily interested in the total variation and  $L_\infty$  mixing times of the full system and we denote them by

$$t_{\text{mix}}^{\text{full}}(\varepsilon) = t_{\text{mix}}^{\text{full}, (\mu, p)}(\varepsilon) \quad \text{and} \quad t_{\text{mix}}^{\text{full}, (\infty)}(\varepsilon) = t_{\text{mix}}^{\text{full}, (\mu, p), (\infty)}(\varepsilon).$$

We denote the corresponding mixing times for the SRW on  $G$  by  $t_{\text{mix}}^{\text{SRW}}(G, \varepsilon)$  and  $t_{\text{mix}}^{\text{SRW}, (\infty)}(G, \varepsilon)$ . We omit  $G$  when clear from context, and  $\varepsilon$  when  $\varepsilon = 1/4$ .

Let  $\mathbb{P}_{x, \eta}$  be the law of the full process, started from initial environment  $\eta$  and initial location  $x$  for the walk. We denote the corresponding expectation by  $\mathbb{E}_{x, \eta}$ . When  $p$  and  $\mu$  are not clear from context, we write  $\mathbb{P}_{x, \eta}^{(\mu, p)}$  and  $\mathbb{E}_{x, \eta}^{(\mu, p)}$ . We write  $\mathbb{P}_{x, \eta}^t$  as a shorthand for the law of the full process at time  $t$ ,  $\mathbb{P}_{x, \eta}((X_t, \eta_t) = (\cdot, \cdot))$ .

We are also interested in *hitting times* by the full process of the form

$$t_{\text{hit}}^{\text{full}} = t_{\text{hit}}^{\text{full}, (\mu, p)} = \max_{x, y \in V, \eta \in \{0, 1\}^E} \mathbb{E}_{x, \eta}[T_{\{y\} \times \{0, 1\}^E}],$$

where for a set  $A \subset V \times \{0, 1\}^E$  its *hitting time*  $T_A = \inf\{t : (X_t, \eta_t) \in A\}$  is defined to be the first time the process visits the set  $A$ . Our Theorem 1.1 bounds  $t_{\text{hit}}^{\text{full}}$  in terms of the hitting times for the SRW which are denoted by  $t_{\text{hit}}^{\text{SRW}}(G) = \max_{x, y \in V} \mathbb{E}_x^{\text{SRW}}[T_y]$ .

For functions  $f, g$  we will write  $f(n) \lesssim g(n)$  if there exists a constant  $c > 0$  such that  $f(n) \leq cg(n)$  for all  $n$ . We write  $f(n) \gtrsim g(n)$  if  $g(n) \lesssim f(n)$ . Finally, we write  $f(n) \asymp g(n)$  if both  $f(n) \lesssim g(n)$  and  $f(n) \gtrsim g(n)$ . We write  $\asymp_a, \lesssim_a$  and  $\gtrsim_a$  when the implied constant depends on  $a$ .

Peres and Steif [29] asked whether  $t_{\text{hit}}^{\text{full}, (\mu, p)} \lesssim_{\mu, p} |V|^3$  which is the natural analog of the classical bound  $t_{\text{hit}}^{\text{SRW}} \lesssim |V|^3$  (see, e.g., [1, 22]). In the following theorem we give an upper bound on  $t_{\text{hit}}^{\text{full}, (\mu, p)}$  in terms of  $t_{\text{hit}}^{\text{SRW}}$ . Using  $t_{\text{hit}}^{\text{SRW}} \lesssim |V|^3$ , this answers affirmatively their question.

**THEOREM 1.1 (Hitting time comparison with SRW).** *For every  $\mu$  there exists a positive constant  $c_1$  such that for all graphs  $G$  and all  $p$  we have that*

$$(1.1) \quad t_{\text{hit}}^{\text{full}, (\mu, p)} \leq \frac{c_1}{p} \cdot t_{\text{hit}}^{\text{SRW}}.$$

*Moreover, there exists a constant  $c_2$  so that for all graphs  $G$  and all  $(\mu, p) \in (0, 1]^2$  we have that*

$$(1.2) \quad t_{\text{hit}}^{\text{full}, (\mu, p)} \leq c_2(\mu^{-1} t_{\text{hit}}^{\text{full}, (1, p)} + t_{\text{mix}}^{\text{full}, (\mu, p)}).$$

We believe the  $t_{\text{mix}}^{\text{full}, (\mu, p)}$  term in (1.2) in most cases satisfies  $t_{\text{mix}}^{\text{full}, (\mu, p)} \lesssim \mu^{-1} t_{\text{hit}}^{\text{full}, (1, p)}$  and can thus be removed from (1.2); see item (2) in Section 1.1.

The Dirichlet form associated to the transition matrix  $P$  (resp., generator  $\mathcal{L}$ ) is defined to be

$$\mathcal{E}_P(f, f) = \frac{1}{2} \sum_{x,y} \pi(x) P(x, y) (f(x) - f(y))^2 = \pi((I - P)f \cdot f)$$

for all  $f : \Omega \rightarrow \mathbb{R}$  (resp.,  $\mathcal{E}_{\mathcal{L}}(f, f) = \frac{1}{2} \sum_{x,y} \pi(x) \mathcal{L}(x, y) (f(x) - f(y))^2 = \pi(-\mathcal{L}f \cdot f)$ ).

For  $\varepsilon > 0$  we denote the spectral profile

$$(1.3) \quad \Lambda(\varepsilon) = \min\{\mathcal{E}(h, h) : h \in \mathbb{R}^\Omega, \text{Var}_\pi(h) = 1, \pi(\text{supp}(h)) \leq \varepsilon\},$$

where  $\text{supp}(h) = \{x \in \Omega : h(x) \neq 0\}$  is the *support* of  $h$  and  $\text{Var}_\pi(h) = \mathbb{E}_\pi[(h - \mathbb{E}_\pi h)^2]$  is the variance w.r.t.  $\pi$ .

For  $\varepsilon > 0$ , the  $\varepsilon$ -spectral-profile time is given by

$$(1.4) \quad t_{\text{spectral-profile}}(\varepsilon) = \int_{4\pi_*}^{4/\varepsilon} \frac{2d\delta}{\delta \Lambda(\delta)},$$

where  $\pi_* = \min_x \pi(x)$ .

Goel et al. [13] showed that  $t_{\text{mix}}^{(\infty)}(\varepsilon) \leq t_{\text{spectral-profile}}(\varepsilon)$  (this refines the evolving sets bound of Morris and Peres [25]). Let  $t_{\text{spectral-profile}}^{\text{SRW}}$  be  $t_{\text{spectral-profile}}$  defined w.r.t. simple random walk. We recall a result of Kozma [21] that for all (simple) finite graphs  $t_{\text{spectral-profile}}^{\text{SRW}} \lesssim t_{\text{mix}}^{\text{SRW},(\infty)} \log \log |V|$ . However, there are many families of graphs for which  $t_{\text{spectral-profile}}^{\text{SRW}} \asymp t_{\text{mix}}^{\text{SRW},(\infty)} \asymp t_{\text{mix}}^{\text{SRW}}$ . For instance, this is the case for the hypercube and for vertex-transitive graphs of moderate growth.

**THEOREM 1.2 (Mixing time comparison with SRW).** *There exists a positive constant  $c_1$  such that for all graphs  $G$  and all  $(\mu, p) \in (0, 1]^2$  we have for all  $\varepsilon \in (0, 1)$*

$$(1.5) \quad t_{\text{mix}}^{\text{full},(\mu,p),(\infty)}(\varepsilon) \leq \frac{c_1}{\mu p} \cdot t_{\text{spectral-profile}}^{\text{SRW}}(\varepsilon) + \frac{c_1}{\mu} |\log(1 - p)|.$$

**REMARK 1.3.** We note that the term  $|\log(1 - p)|/\mu$  is necessary, because the  $L_\infty$  mixing time of the environment is at least of this order. However, if we were considering the total variation mixing time of the full system, then we would get rid of this extra term. We explain this at the end of the proof of this theorem in Section 6.

We recall that the *spectral gap* of a reversible transition matrix  $P$  (resp., generator  $\mathcal{L}$ ) is defined as the smallest positive eigenvalue of  $I - P$  (resp.,  $-\mathcal{L}$ ). Its inverse is called the *relaxation time*. We denote the relaxation time of the full process by  $t_{\text{rel}}^{\text{full}} = t_{\text{rel}}^{\text{full},(\mu,p)}$  and that of the SRW by  $t_{\text{rel}}^{\text{SRW}}$ . We recall that, generally, for a continuous-time reversible Markov chain on a finite state space  $\Omega$  with stationary distribution  $\pi$  the relaxation time determines the asymptotic exponential rate of convergence to equilibrium in the following precise sense (cf. [22], Lemmas 20.5 and 20.11) for all  $\varepsilon \in (0, 1)$ :

$$t_{\text{rel}} |\log \varepsilon| \leq t_{\text{mix}}(\varepsilon/2) \leq t_{\text{mix}}^{(\infty)}(\varepsilon) \leq t_{\text{rel}} |\log(\pi_* \varepsilon)| \quad \text{where } \pi_* := \min_{x \in \Omega} \pi(x).$$

The inequality above together with Theorem 1.2 and the fact that  $t_{\text{spectral-profile}}^{\text{SRW}}(\varepsilon) \asymp |\log(\varepsilon)| t_{\text{rel}}^{\text{SRW}}$  for  $\varepsilon < 1/|V|$  imply the following corollary.

COROLLARY 1.4. *Uniformly, in  $(\mu, p) \in (0, 1]^2$  and in  $G = (V, E)$  we have that for all  $k \in \mathbb{N}$*

$$\begin{aligned} \frac{1}{k} t_{\text{mix}}^{\text{full},(\mu,p),(\infty)} \left( \frac{1}{|V|^k} \right) &\lesssim (\mu p)^{-1} t_{\text{rel}}^{\text{SRW}} \log |V| + \frac{1}{\mu} |\log(1-p)| \\ &\asymp (\mu p)^{-1} t_{\text{mix}}^{\text{SRW}} \left( \frac{1}{|V|} \right) + \frac{1}{\mu} |\log(1-p)|. \end{aligned}$$

REMARK 1.5. The following quenched statement follows easily from Corollary 1.4:

Let  $t = t(\mu, p, C) := C(\mu p)^{-1} t_{\text{rel}}^{\text{SRW}} \log |V|$ . Define  $d(t, x, \{\eta_s\}_{s \geq 0})$  to be the total variation distance from  $\pi$  of the walk coordinate  $X_t$  at time  $t$  of the full process, given the environment  $\{\eta_s\}_{s \geq 0}$  and given  $X_0 = x$ . Then, for some choice of universal constant  $C > 0$  for all  $(\mu, p) \in (0, 1]$  we have that

$$\max_{\eta} \mathbb{P}_{\eta}^{(\mu,p)} \left( \max_x d(t, x, \{\eta_s\}_{s \geq 0}) > 1/|V|^2 \right) < 1/|V|.$$

Indeed, this follows by Markov’s inequality and a union bound over  $x$ .

Recall that the *log-Sobolev constant* of a continuous-time Markov chain on a finite state space  $\Omega$ , with generator  $\mathcal{L}$  and stationary distribution  $\pi$  is given by

$$c_{\text{LS}} = \inf \left\{ \frac{\mathcal{E}_{\mathcal{L}}(h, h)}{\text{Ent}_{\pi}(h^2)} : h^2 \in [0, \infty)^{\Omega} \text{ and nonconstant} \right\},$$

where  $\text{Ent}_{\pi}(f) = \mathbb{E}_{\pi}[f \log(f/\mathbb{E}_{\pi}[f])]$  (with the convention that  $0 \log 0 = 0$ ).

THEOREM 1.6. *There exist positive constants  $c_1$  and  $c_2$  so that for every graph  $G$  and all values of  $(\mu, p)$  we have that*

$$c_{\text{LS}}^{\text{full},(\mu,p)} \geq c_1 \mu \min \left\{ p c_{\text{LS}}^{\text{SRW}}, \frac{1}{\log(1/\pi_*) \log(\frac{1}{p(1-p)})} \right\},$$

where  $\pi_*$  above is defined to be  $\min_{v \in V} \pi(v)$ . Moreover, we have that

$$t_{\text{rel}}^{\text{full},(\mu,p)} \leq c_2 \frac{1}{\mu p} t_{\text{rel}}^{\text{SRW}}.$$

REMARK 1.7. We note that for a simple random walk if  $h_v(u) := \frac{1_{u=v}}{\sqrt{\pi(v)}}$ , then

$$c_{\text{LS}}^{\text{SRW}} \leq \min_v \frac{\mathcal{E}_{\mathcal{L}}(h_v, h_v)}{\text{Ent}_{\pi}(h_v^2)} = \frac{1}{\log(1/\pi_*)},$$

and, hence, if  $p$  is bounded away from 1, then the lower bound above for the log-Sobolev constant of the full process becomes  $c_3 \mu p c_{\text{LS}}^{\text{SRW}}$ .

Let  $G = (V, E)$  be an  $n$ -vertex connected graph. We say that  $G$  is *vertex transitive* if the action of its automorphism group on its vertices is transitive. Denote the volume of a ball of radius  $r$  in  $G$  by  $V(r)$ . Denote the diameter of  $G$  by  $\gamma = \inf\{r : V(r) \geq n\}$ . Following Diaconis and Saloff-Coste [7], we say that  $G$  has  $(c, a)$ -*moderate growth* if  $V(r) \geq cn(r/\gamma)^a$  for all  $r$ . Breuillard and Tointon [6] proved that for Cayley graphs of fixed degree, this condition is equivalent in some quantitative sense to the simpler condition that  $n \leq \beta \gamma^\alpha$  for some  $\alpha, \beta > 0$ . Tessera and Tointon recently extended this result to vertex-transitive graphs [31]. We note that Diaconis and Saloff-Coste [7] proved that for vertex-transitive graphs of  $(c, a)$ -moderate growth and of degree  $d$

$$d^{-1} t_{\text{mix}}^{\text{SRW},(\infty)} \lesssim_{a,c} \gamma^2 \lesssim_{a,c} t_{\text{rel}}^{\text{SRW}} \leq t_{\text{mix}}^{\text{SRW},(\infty)}.$$

Denote the percolation cluster of vertex  $x$  by  $K_x$  and its edge boundary by  $\partial K_x$ . We identify the cluster with the vertices lying in it and denote the cardinalities of  $K_x$  and  $\partial K_x$  by  $|K_x|$  and  $|\partial K_x|$ , respectively. We also denote  $M_p = \pi_p(|\partial K_x||K_x|^2) \leq \pi_p(d|K_x|^3)$  and  $N_p = \pi_p(|K_x|)$  (by transitivity these quantities are independent of  $x$ ), where  $d$  is the degree and where  $\pi_p(f) = \sum_{\eta} \pi_p(\eta) f(\eta)$  denotes expectation w.r.t.  $\pi_p$  of  $f : \{0, 1\}^E \rightarrow \mathbb{R}$ .

**THEOREM 1.8** (Moderate growth vertex-transitive graphs, subcritical regime). *Let  $a, b, c \in \mathbb{R}_+$ . Let  $G = (V, E)$  be a connected vertex transitive graph of degree  $d$ ,  $(c, a)$ -moderate growth and diameter  $\gamma$ . Suppose that  $|\log(1 - p)| \leq \gamma^2$ ,  $M_p \leq b$  and  $N_p \leq \gamma/8$ . Then,*

$$t_{\text{mix}}^{\text{full},(\mu,p),(\infty)} \asymp_{a,b,c,d} (\mu p)^{-1} t_{\text{rel}}^{\text{SRW}} \asymp_{a,b,c} t_{\text{rel}}^{\text{full},(\mu,p)}.$$

Moreover, even if  $p$  is not subcritical, we still have that  $t_{\text{mix}}^{\text{full},(\mu,p),(\infty)} \lesssim_{a,c,d} (\mu p)^{-1} t_{\text{rel}}^{\text{SRW}}$ .

We note that the condition  $|\log(1 - p)| \leq \gamma^2$  is very mild and, typically, follows from either of the two other conditions (provided  $b$  is much smaller than  $n$ ). In applications  $\gamma \gg 1$ , and we think of  $b$  as a constant. In this case the conditions  $N_p \leq \gamma/8$  and  $|\log(1 - p)| \leq \gamma^2$  follow from the condition  $M_p \leq b$ . A more refined statement, with an explicit dependence on  $b$ , is given in Lemma 7.1.

**THEOREM 1.9** (Hypercube). *Let  $G$  be the hypercube  $\{0, 1\}^d$ . Then, uniformly in  $(\mu, p) \in (0, 1)^2$ , we have that*

$$t_{\text{mix}}^{\text{full},(\mu,p),(\infty)} \lesssim \frac{1}{\mu p} d \log d + \frac{1}{\mu} |\log(1 - p)|.$$

It is conjectured ([32], page 59) that, when  $p = c/d$  for  $c > 1$ , the mixing time of the SRW on the giant component is w.h.p.  $\Theta(d^2)$ ; however, no polynomial in  $d$  upper bound is known for the static case.

1.1. *Remarks and open problems.*

(1) *Comparison with previous results:* We are able to extend the results of Peres, Stauffer and Steif [28] about  $\mathbb{Z}_n^d$  to the more general setup of vertex-transitive graphs of moderate growth via a simpler proof, while also eliminating the requirement that  $p$  is subcritical. Our results for the supercritical case complement the results on Peres, Sousi and Steif [26, 27]. When  $\mu$  is of order 1, our Theorem 1.8 provides better bounds, but their results are quenched and, hence, not directly comparable to ours (also, the main result from [27] required  $p$  to be supercritical in a certain quantitative way, namely, that the expected size of the giant component is at least  $|V|/2$ ). We note that we give the first bounds about the mixing time in the critical case.

Biskup and Rodriguez [5] study continuous-time random walk on  $\mathbb{Z}^d$  for  $d \geq 2$  with random, symmetric jump rates which are time dependent, stationary, ergodic and bounded from above, but are not assumed to be positive. Under mild assumptions on the environment (meant to eliminate the possibility that a certain edge has rate zero for a long period of time), they prove a quenched invariance principle.

We consider a particular case of this model in remark (9) below but do so on general vertex-transitive graphs (we also do not require the weights to be bounded). We consider the mixing time of the walk coordinate in this model. The main difference is that, while we consider the less general (and simpler) case in which the rates of different edges are independent, we provide an upper bound on the mixing time with an explicit dependence on the update rate of the rates.

(2) *Refining (1.2)*: It is quite possible that for all regular and all bounded degree graphs one has that  $t_{\text{spectral-profile}}^{\text{SRW}} \lesssim t_{\text{hit}}^{\text{SRW}}$ . We strongly believe this to be the case when  $G$  is a Cayley graph. Whenever this is the case, (1.2) reads as

$$t_{\text{hit}}^{\text{full},(\mu,p)} \lesssim (\mu p)^{-1} t_{\text{hit}}^{\text{SRW}}.$$

This follows from Theorem 1.2 and Remark 1.3. We note that this always holds up to a  $\log \log |V|$  factor, as by Kozma [21] for all graphs  $t_{\text{spectral-profile}}^{\text{SRW}} \lesssim t_{\text{mix}}^{\text{SRW},(\infty)} \log \log |V|$ , while  $t_{\text{mix}}^{\text{SRW},(\infty)} \lesssim t_{\text{hit}}^{\text{SRW}}$  ([22], Theorem 10.22).

(3) *Hitting times lower bound for subcritical  $p$* : It is reasonable to expect that for  $p$  which is subcritical, say in the sense that the second moment  $M$  of the size of a cluster is bounded, we have that  $t_{\text{hit}}^{\text{full},(\mu,p)} \gtrsim_M (\mu p)^{-1} t_{\text{hit}}^{\text{SRW}}$ . (Note that  $M$  is increasing in  $p$ . Hence, the  $p^{-1}$  term on the r.h.s. is meant to capture the dependence on  $p$  for  $p$  close to 0. Namely, for such  $p$  we expect that  $t_{\text{hit}}^{\text{full},(\mu,p)} \gtrsim (\mu p)^{-1} t_{\text{hit}}^{\text{SRW}}$ .)

(4) *Heuristics for the  $(\mu, p)$  dependence*: The following heuristics explain the appearance of the terms  $\mu$  and  $p$  in our main results when the maximal degree is bounded. Informally, when  $p$  is subcritical, viewing the walk when it moves from one cluster to another gives a new random process with small increments, which (as a discrete-time process) we expect in many cases to have the same order mixing time as SRW. However, as we now explain, the time between consecutive steps of this process is typically of order  $\frac{1}{\mu p}$ . Indeed, the walk coordinate of the full process is typically at a cluster  $\mathcal{C}$  of size  $O(1)$ . Using the fact that the degree is bounded, it requires order  $\frac{1}{|\mathcal{C}|\mu p} = \Omega(\frac{1}{\mu p})$  time units until an edge adjacent to  $\mathcal{C}$  becomes open.

(5) *Direct comparison of mixing times*: It is natural to wonder whether stronger results of the form  $t_{\text{mix}}^{\text{full},(\mu,p)} \lesssim (\mu p)^{-1} t_{\text{mix}}^{\text{SRW}} + \mu^{-1} \log |V|$  and  $t_{\text{mix}}^{\text{full},(\mu,p),(\infty)} \lesssim (\mu p)^{-1} t_{\text{mix}}^{\text{SRW},(\infty)} + \mu^{-1} |\log(1-p)|$  hold. Alas, it appears that small variants of the examples from [10, 15, 18] can be used to show that this does not hold in general. However, if  $G$  is vertex transitive, we believe this should indeed be the case with the implicit constant depending on the degree.

(6) *Commute times*: It follows from our argument that, for every  $x, y \in V$ , the expected time it takes the walk component of the full process to reach  $y$  from  $x$  and then return to  $x$  is at most  $O_\mu(p^{-1} \mathbb{E}_x^{\text{SRW}}[T_y] + \mathbb{E}_y^{\text{SRW}}[T_x])$  and also at most  $O(\frac{1}{\mu p} [\mathbb{E}_x^{\text{SRW}}[T_y] + \mathbb{E}_y^{\text{SRW}}[T_x] + t_{\text{spectral-profile}}^{\text{SRW}}])$  for  $p$  bounded away from 1.

(7) *Extending Theorem 1.9 to selfproducts*: The Cartesian product  $G_1 \times G_2 = (V', E')$  of two graphs  $G_i = (V_i, E_i)$  is defined via  $V' := V_1 \times V_2$  and

$$E' := \{(v_1, v_2), (u_1, u_2)\} : v_1 = u_1 \in V_1 \text{ and } v_2 u_2 \in E_2, \text{ or vice versa}\}.$$

For a graph  $G = (V, E)$ , we denote the  $n$ -fold self (Cartesian) product of  $G$  with itself by  $G_{\otimes n} = (V^n, E(G_{\otimes n}))$ . That is,  $G_{\otimes n} = G_{\otimes(n-1)} \times G = G \times \dots \times G$ . Note that the  $n$ -dim hypercube is the  $n$ -fold selfproduct of the complete graph on two vertices with itself. The proof of Theorem 1.9 can easily be extended to show that, uniformly, in  $(\mu, p) \in (0, 1]^2$ , for  $G_{\otimes n}$  with  $G = (V, E)$  we have that

$$t_{\text{mix}}^{\text{full},(\mu,p),(\infty)} \lesssim_{|V|} \frac{1}{\mu p} n \log n + \frac{1}{\mu} |\log(1-p)|.$$

In fact, this can be derived as an immediate corollary from either (1.5) (in a similar fashion to the derivation of Theorem 1.9) or from Theorem 1.6.

(8) *Allowing  $\mu > 1$* : We assume throughout that  $\mu \in (0, 1]$ . Our analysis of the case  $\mu = 1$  can be used almost verbatim to treat  $\mu > 1$ , in which case terms of the form  $\frac{1}{\mu}$  should be replaced by  $\frac{1+\mu}{\mu}$ .

(9) *Random rates model:* We now discuss a certain extension of our results to a *random walk on dynamical random rates model*. Let  $\nu$  be some law supported on  $\mathbb{R}_+$ . Consider the case that each edge  $e$  is updated at rate  $\mu$ , and, when it is updated, it is assigned a random rate  $r_e$  with law  $\nu$  independently. Given that the current location of the walk coordinate is  $x$  and that the current environment is  $\eta$ , where  $\eta(e)$  denotes the rate of the edge  $e$ , the walk coordinate moves to vertex  $y$  at rate  $\eta(xy)$  for all  $xy \in E$ . Observe that the case that  $\nu$  is Bernoulli( $p$ ) gives rise to random walk on dynamical percolation. Let  $X \sim \nu$ .

We consider the case that the base graph  $G = (V, E)$  is vertex transitive of degree  $d$ , that  $\nu(0) = 1 - p$ , that  $\mathbb{E}[X] < \infty$  and that  $\mathbb{E}[e^{a(X/\mathbb{E}[X])}] \leq b$  for some  $a, b > 0$ . Let  $\kappa(\mu)$  be the expected time until the walk coordinate leaves the origin for the first time in the variant of the model in which for each vertex  $v$ , at rate  $\mu$  all of the rates of the edges incident to  $v$  are refreshed simultaneously and independently, according to the law  $\nu$ . Let  $M := C'(a, b) \frac{d\mathbb{E}[X^2|X>0]}{\mathbb{E}[X]}$ , where  $C'(a, b)$  is a constant depending only on  $a, b$ , to be determined later. (Note that  $M$  implicitly depends on  $\nu(0)$ .)

Our analysis of Theorem 1.2 can be extended to this model. With some additional effort one can show that (for some appropriate  $C'(a, b)$  above) the total variation mixing time of the walk coordinate is at most

$$\frac{C(a, b)d(M + \mu)\kappa(\mu \vee M)}{\mu} t_{\text{spectral-profile}}^{\text{SRW}}$$

where  $C(a, b)$  depends only on  $(a, b)$  and where  $A \vee B := \max\{A, B\}$ .

Moreover, if  $\nu$  has a finite support, this is also an upper bound on the total variation mixing time of the full process, whereas the  $L_2$  mixing time of the full process is at most

$$(1.6) \quad \frac{C(a, b)d(M + \mu)\kappa(\mu \vee M)}{\mu} t_{\text{spectral-profile}}^{\text{SRW}} + \frac{Cd|\log \min_{x:x \neq 0} \nu(x)|}{\mu}.$$

This extension of Theorem 1.2 to the random rates model is not at all straightforward. We provide the details in [20], Section 6.1.

QUESTION 1.10. *Let  $G$  be an infinite connected vertex-transitive graph. Assume that initially the environment is stationary. Let  $P_t(o, o)$  and  $Q_t(o, o)$  be the return probability to the origin by the (continuous-time) SRW and by the walk coordinate of the full process, respectively. Is it the case that, for some  $C = C(G, p, \mu) \geq 1$ , we have*

$$\forall t \geq 0, \quad \frac{1}{C} P_{Ct}(o, o) \leq Q_t(o, o) \leq C P_{t/C}(o, o)?$$

QUESTION 1.11. *What is the order of  $t_{\text{mix}}^{\text{full}, (\mu, p)}$  for  $G = \mathbb{Z}_n^d$  when  $p$  is the critical probability  $p_c$  for Bernoulli bond percolation on  $\mathbb{Z}_n^d$ ?*

We believe that for all  $d$  the order of the mixing time when  $p = p_c$  and  $\mu = o(1)$  lies strictly between its values in the subcritical and supercritical regimes (for the same  $\mu$ ) and that it has a complicated dependence on  $\mu$  and  $d$ . (It is possible that the dependence on  $d$  becomes simple once mean-field behavior kicks in.) This might seem surprising at first glance due to the lack of exceptional times (i.e., times at which an infinite cluster exists) for critical dynamical percolation on  $\mathbb{Z}^d$  in high dimension [14]. We wish to express our immense gratitude to Gabor Pete for relevant discussions.

The *cover time*  $\tau_{\text{cov}}^{\text{SRW}}(G)$  of a graph  $G = (V, E)$  is the first time by which every vertex  $v \in V$  has been visited by SRW on  $G$ . Let  $t_{\text{cov}}^{\text{SRW}}(G) := \max_{v \in V} \mathbb{E}_v[\tau_{\text{cov}}^{\text{SRW}}(G)]$  be its worst case expectation. We can similarly define  $\tau_{\text{cov}}^{(\mu, p)}(G)$  to be the first time by which every vertex  $v \in V$  has been visited by walk coordinate of the full process with parameters  $(\mu, p)$  (i.e.,  $(v, \eta)$  has been visited for some  $\eta$ ) and  $t_{\text{cov}}^{(\mu, p)}(G) := \max_{(v, \eta)} \mathbb{E}_{(v, \eta)}^{(\mu, p)} t_{\text{cov}}^{(\mu, p)}(G)$ .

QUESTION 1.12. *Is it the case that there exist constants  $C(\mu) > 0$  such that for all finite graphs  $G$  we have that*

$$\forall \mu \in (0, 1], p \in (0, 1], \quad t_{\text{cov}}^{(\mu,p)}(G) \leq \frac{C(\mu)}{p} t_{\text{cov}}^{\text{SRW}}(G)?$$

We comment that, if  $G = (H, E)$  is a Cayley graph of an Abelian group  $H$ , then the auxiliary process from Section 3 is reversible (see Remark 3.7). In [9] the authors give a general comparison principle for cover times (of two reversible Markov chains on the same state space, provided the effective resistances of one chain are pairwise smaller than some constant multiple of those of the other). Thus, using our comparison of the transition probabilities between the auxiliary chain and the SRW from Lemma 4.4 (together with Theorem 4.1 which relates this to a comparison of effective resistances), it follows from [9] that  $t_{\text{cov}}^{(\mu,p)}(G) \leq \frac{C(\mu)}{p} t_{\text{cov}}(G)$ , by considering the cover time of  $G$  for the auxiliary walk. We omit the details.

**2. Overview of our approach.** We believe that our approach may be relevant for other models of random walks on evolving graphs, in situations in which the mixing time of the environment is smaller than that of the walk coordinate. We first explain why we may concentrate on the case that  $\mu = 1$ . Let  $\mathcal{L}_{(\mu,p)}$  be the infinitesimal generator of the full process with edge probability  $p$  and update rate  $\mu$ . One can readily see if  $\mu_1 < \mu_2$ , then the corresponding rates satisfy

$$\forall x, x' \in V, \eta, \eta' \in \{0, 1\}^E, \quad \mathcal{L}_{(\mu_2,p)}((x, \eta), (x', \eta')) \leq \frac{\mu_2}{\mu_1} \mathcal{L}_{(\mu_1,p)}((x, \eta), (x', \eta')).$$

Using this and some general theory, we are able to transfer hitting time estimates from  $\mu = 1$  to ones for  $\mu < 1$ , at a cost of a  $\mu^{-1}$  multiplicative term and an additive term of order of  $t_{\text{mix}}^{\text{full},(\mu,p)}$ . As for the mixing time, we obtain an upper bound on the spectral-profile bound when  $\mu = 1$  by a certain comparison with SRW. Again, general theory allows us to then translate this into a bound for general  $\mu < 1$  at a cost of a  $\mu^{-1}$  multiplicative term.

Now, consider the case that  $\mu = 1$ . As we now explain, we may also consider the case that  $\eta_0 \sim \pi_p$ . It is tempting to argue that it suffices to wait until every edge is updated once, which takes order  $\log |E|$  time units, and then the environment is stationary. However, since the walk co-ordinate is dependent on the environment, some difficulties arise when trying to formulate this. Using the fact that  $\mu = 1$ , we show that at some stopping time slightly larger than the first time at which each edge is updated at least once, the environment is stationary and is independent of the position of the walk. This explains why when bounding hitting times we may assume that  $\eta_0 \sim \pi_p$ .

We define a sequence of increasing stopping times (w.r.t. an enlarged filtration)  $(\tau_i : i \in \mathbb{Z}_+)$  such that the following hold:

- (1)  $\eta_{\tau_i} \sim \pi_p$  and is independent of  $X_{\tau_i}$  for all  $i$ .
- (2)  $\tau_0 = 0$  and  $\tau_{i+1} - \tau_i$  are i.i.d. such that  $\mathbb{E}[e^{\delta \tau_1}] < \infty$  for some  $\delta > 0$ .
- (3)  $Y_i = X_{\tau_i}$  is a Markov chain with the same stationary distribution as SRW. We call each such  $\tau_i$  a regeneration time.

By construction we will have for all  $i \in \mathbb{N}$  that all of the edges examined by the walk coordinate (i.e., that the walk coordinate attempted to cross) at some time during  $(\tau_{i-1}, \tau_i)$  have been refreshed since the last time in  $(\tau_{i-1}, \tau_i)$  at which they were examined. This, along with the assumption that  $\eta_0 \sim \pi_p$ , imply the crucial property (1) above.

While we do not believe the auxiliary chain  $(Y_i)_{i \in \mathbb{Z}_+}$  to be reversible in general, we show that  $\pi$  is its stationary distribution, and that the transition matrix  $Q$  of its additive symmetrization satisfies that  $Q(u, v) \gtrsim \frac{1}{p \deg u}$  which turns out to be sufficient in order to obtain



a comparison of its hitting times (more precisely, of its commute times) and of its spectral-profile with that of SRW (at a price of an  $O(\frac{1}{p})$  factor). To turn these into upper bounds on expected hitting times for  $(Y_i)_{i \in \mathbb{Z}_+}$  (rather than for its additive symmetrisation), we use the fact that the commute times for the symmetrisation are always as large as they are for the original chain. As the chain  $(Y_i)_{i \in \mathbb{Z}_+}$  is constructed by viewing the chain  $(X_t)_{t \in \mathbb{R}_+}$  at a nice sequence of stopping times, we then easily translate hitting time bounds for  $Y$  into ones for  $X$  (for hitting times of the form  $T_{Y \times \{0,1\}^E}$ ).

For the mixing time we derive a comparison of the spectral profile bound  $t_{\text{spectral-profile}}^{\text{full},(1,p)}$  for the full process with parameters  $(1, p)$  with that of SRW  $t_{\text{spectral-profile}}^{\text{SRW}}$ . The comparison method is a standard method for comparing analytic quantities that have an extremal characterization involving the Dirichlet form, between two Markov chains. Despite the fact that the auxiliary chain may be nonreversible, using the comparison method we are able to effortlessly compare the spectral-profiles of SRW with that of the auxiliary chain. We wish to use the auxiliary chain as a link to relate the SRW with the full process.

A major obstacle in doing so is that the former has state space  $V$ , while the latter  $V \times \{0, 1\}^V$ . Unfortunately, the comparison method requires the two considered chains to have the same state space or at least similar state spaces (e.g., random walks on two quasi-isometric graphs might have different state spaces but, with more effort, can still be compared). We are unaware of any previous works establishing a comparison argument between Markov chains with very different state spaces.

Our solution to this difficulty hinges on a probabilistic interpretation of the spectral profile in terms of the rate of exponential decay of the tail of hitting times starting from the stationary distribution of the chain, which we review in Section 5. This interpretation allows us to use the the auxiliary chain as a link between the full process and the SRW.

For each set  $A \subseteq V \times \{0, 1\}^E$ , we construct a certain set  $B \subseteq V$  such that  $\pi(B^c) \lesssim \pi_{\text{full},p}(A^c)$  and show that the rate of exponential decay of the tail of  $T_A$  for the full process can be controlled via that of  $T_B$  for the aforementioned auxiliary chain  $Y$ . The set  $B$  is defined to be the set of all vertices  $v \in V$  such that  $\{w \in \{0, 1\}^E : (v, w) \in A\}$  has  $\pi_p$  probability at least  $1/4$ . Let  $C \geq 1$  be some absolute constant. Let

$$\gamma_A := \sup\{a : \mathbb{E}_\pi[e^{aT_D}] < \infty \text{ for all } D \subset V \text{ such that } \pi(D^c) \leq C\pi_{\text{full},p}(A^c)\},$$

where the expectation is taken w.r.t. SRW. The connection of this quantity to the spectral profile of SRW is explained in Section 5. It follows from this connection that, to prove Theorem 1.2, it suffices to show that for some absolute constant  $c > 0$  we have that  $\mathbb{E}_{\pi_{\text{full},p}}[\exp(cp\gamma_A T_A)] < \infty$  (recall that we consider  $\mu = 1$ ).

Let  $t = \frac{C'}{\gamma_A p}$ . Loosely speaking, we show that for some choice of  $(C, C')$ , for all  $i \in \mathbb{N}$ , the full process started from its stationary distribution has a regeneration time during the time interval  $[2it, (2i + 1)t]$ , with probability bounded from below. Moreover, at each such regeneration time, with probability bounded from below, the full process is in  $A$ . Furthermore, this holds even conditioned on the information gathered at the previous time intervals  $[2jt, (2j + 1)t]$ , where for each such time interval we expose if there was a regeneration time during it, and, if there was one, we take the first one and expose whether the walk is in  $A$  at that time.

(In the actual proof we do not restrict to one regeneration time per time interval, although we could have done so. For technical reasons we make several other small modifications in the proof to what is written above.)

This is achieved by treating the walk coordinate and the environment separately. For the walk coordinate we use the comparison method to compare the auxiliary chain with the SRW. By comparing their spectral profiles and applying a certain inductive argument (which we

sketch below), we are able to show that at each time interval  $[2it, (2i + 1)t]$  the chain has probability bounded from below to visit  $B$ , even when conditioning on the information exposed during the previous time intervals. This gives us a sequence of visits to  $B$  at regeneration times by the walk coordinate which are separated by at least  $t$  time units.

For the environment note that, by definition, at every visit of the walk coordinate to  $B$ , if the environment is distributed according to  $\pi_p$  and independent of the walk, the full process would have probability at least  $1/4$  of being in  $A$ . By property (1) above, the distribution of the environment is indeed stationary at all regeneration times and, in particular, at the regeneration times at which the walk visits  $B$ . However, (generally) if  $(Y_i)$  is a stationary Markov chain, and for each  $j$ ,  $D_j$  is an event which is determined by  $Y_j$ , then, given the values of the indicators of  $D_1, \dots, D_i$ , we no longer have that  $Y_{i+1}$  has the same law as its unconditional law (i.e., its law might differ from the invariant distribution). In particular, if at some of the previous visits to  $B$  at regeneration times we expose if the full process is in  $A$ , this affects the law of the environment at the current regeneration time, making it no longer stationary. It turns out that the fact that we consider a sequence of such visits which are separated by an amount of time units which is larger than the relaxation time of the environment allows us to control such dependencies. We now review the general principle from which this follows.

For a Markov chain  $(W_t)_{t \in \mathbb{R}_+}$  on a finite state space, reversible w.r.t. a distribution  $\hat{\pi}$ , starting from its invariant distribution, if  $s$  is a large multiple of its relaxation time, then samples of the chain at times  $0, s, 2s, \dots$  are uncorrelated (e.g., [22], Section 12.7). In fact, if  $D_0, D_1, D_2, \dots$  is a sequence of sets such that  $\inf\{\min\{\hat{\pi}(D_i), \hat{\pi}(D_i^c)\} : i \geq 0\} =: \delta > 0$ , then, as we explain below, using  $L_2$  considerations and an inductive argument, one can show that if  $s \geq C_1 t_{\text{rel}} \log(1/\delta)$  for some absolute constant  $C_1$ , then a.s.

$$\mathbb{E}[\mathbf{1}_{W_{is} \in D_i} \mid (\mathbf{1}_{W_{js} \in D_j^c} : j < i)] \geq \delta/2.$$

Note that this statement is obvious if we replace  $C_1 t_{\text{rel}} \log(1/\delta)$  by the  $\delta/2$  mixing time. As two more complicated variants of this inductive argument will be used in the proof of Theorem 1.2, we now give a sketch of the proof of this fact in order to emphasise the main idea behind the proof.

Let  $\mu$  be a distribution with  $\|\mu - \hat{\pi}\|_{2, \hat{\pi}} \leq \delta$ . Let  $\xi_i$  be the indicator of  $W_{is} \in D_i$ . Then, for all  $i$  and all  $a_0, a_1, \dots \in \{0, 1\}$ , such that  $\mathbb{P}_\mu[E_\ell] > 0$  for all  $\ell$ , where  $E_\ell = \{\xi_j = a_j \text{ for all } j \leq \ell\}$ , we have that

$$\mathbb{P}_\mu[W_{is} \in D_i \mid E_{i-1}] \in \left[ \hat{\pi}(D_i) - \frac{\delta}{2}, \hat{\pi}(D_i) + \frac{\delta}{2} \right].$$

This is proven by proving by induction that for all  $i \in \mathbb{N}$

$$\Upsilon_i := \|\mathbb{P}_\mu(W_{is} \in \cdot \mid E_i)\|_{2, \hat{\pi}}^2 \leq 4 \frac{\delta^2 + 1}{\min\{\hat{\pi}(D_i), \hat{\pi}(D_i^c)\}^2},$$

and that

$$|\mathbb{P}_\mu(W_{(i+1)s} \in D_{i+1} \mid E_i) - \hat{\pi}(D_{i+1})| \leq \|\mathbb{P}_\mu(W_{(i+1)s} \in \cdot \mid E_i) - \hat{\pi}\|_{2, \hat{\pi}} =: \varrho_i \leq \delta.$$

The induction step follows by combining the following simple observations (in proving the induction step, first use (3), then (2) and then (1)):

(1) By the definition of the total-variation distance ( $\|v - v'\|_{\text{TV}} = \max_D |v(D) - v'(D)|$ ) and the fact that  $2\|v - v'\|_{\text{TV}} \leq \|v - v'\|_{2, \pi}$ , we have

$$|\mathbb{P}_\mu(W_{(i+1)s} \in D_{i+1} \mid E_i) - \hat{\pi}(D_{i+1})| \leq \frac{1}{2} \varrho_i.$$

(2) By the Poincaré inequality

$$\rho_i \leq \Upsilon_i e^{-s/t_{\text{rel}}} = \delta^C \Upsilon_i.$$

(3) For any distribution  $\nu$  and any set  $D$  we have  $\|\nu_D - \hat{\pi}\|_{2, \hat{\pi}}^2 + 1 \leq \frac{\|\nu - \hat{\pi}\|_{2, \hat{\pi}}^2 + 1}{\nu(D)^2}$ , where  $\nu_D(b) := \frac{\nu(b)\mathbf{1}_{\{b \in D\}}}{\nu(D)}$  is  $\nu$  conditioned on  $D$  (see (6.10)). Hence, by (1)  $\Upsilon_i^2 \leq \frac{\varrho_{i-1}^2 + 1}{\min\{\hat{\pi}(D_i) - \frac{1}{2}\varrho_{i-1}, \hat{\pi}(D_i^c) - \frac{1}{2}\varrho_{i-1}\}^2}$ .

To argue that at each time interval  $[2it, (2i + 1)t]$  the chain has probability bounded from below to visit  $B$ , even when conditioning on the information exposed during the previous time intervals, we use a more complicated version of this inductive argument sketched below, in which  $\frac{1}{p\gamma_A}$  and the spectral-profile bound on the decay of the  $L_2$  distance from equilibrium play the roles of the relaxation time and the Poincaré inequality above, respectively.

Consider a sequence of visits to the set  $B$  by the walk coordinate of the full process at regeneration times, say to vertices  $b_1, b_2, \dots$ , at times that are separated apart by  $t \geq C_2$  time units, for some large absolute constant  $C_2$ . At the  $i$ th visit, if the environment is in  $D_i := \{w \in \{0, 1\}^V : (b_i, w) \in A\}$ , then the full process visits  $A$  at that time. Let  $\xi_i$  be the indicator of the event that the environment is in  $D_i$  at this time.

We start with a stationary environment, since we start from the stationary distribution on the full process, as we are seeking to show that  $\mathbb{E}_{\pi_{\text{full}, p}}[\exp(cp\gamma_A T_A)] < \infty$ . Since  $t$  is large in terms of the relaxation time of the environment, which is  $O(1)$ , in light of the above discussion it is intuitive that  $\xi_1, \xi_2, \dots$  is approximately a sequence of independent Bernoulli trials. Making this intuition precise turns out to be quite subtle and is among the main technical challenges in the proof of Theorem 1.2. Indeed, a major obstacle is the fact that the sets  $D_1, D_2, \dots$  are random and depend on the walk coordinate and, thus, also on the environment.

**3. Auxiliary chain.** Recall that  $G = (V, E)$  with  $V$  the set of vertices and  $E$  the set of edges. We fix an ordering of the edges  $E = \{e_1, e_2, \dots, e_n\}$ .

For every edge  $e_i$  we create an infinite number of copies denoted  $e_{i,1}, e_{i,2}, \dots$ . We emphasize that the copies will not be considered as edges of  $E$ . We start  $(X_0, \eta_0) \sim \pi \times \pi_p$ . For every time  $t$  we now define a set of “infected edges”  $R_t$  as follows:  $R_0 = \emptyset$  and if  $R_{t-} = A$  and the exponential clock of  $X$  rings at time  $t$ , we then add to  $R_t$  the edge that  $X$  examines to cross at time  $t$ . If this edge already exists in  $A$ , then we add its lowest numbered copy that is not in  $A$ . We next assign an ordering to the edges in  $R_t$  using the ordering of the edges of  $E$  in the following way: assign label 1 to the edge (or copy) contained in  $R_t$  with the lowest label in the ordering of  $E$ . If both the edge and some of its copies are contained in  $R_t$ , then we assign label 1 to the edge and then we give the next label to the lowest numbered copy and so on. If there are only copies of that edge, then assign label 1 to the lowest numbered copy of this edge and label 2 to the second lowest numbered copy and so on, until we exhaust all the copies of the edge. Then, we continue in the same way by finding the second edge (or copy of it) from  $E$  with the second lowest label and assign to it the next label and so on. (We note that any ordering of  $R_t$  would work, but we choose to specify one to make the construction clearer.)

When  $R_{t-} = A$ , then assign exponential clocks of rate  $\mu$  to the edges of  $E$  that are not in  $A$ , and we also generate an exponential clock of rate  $|A|\mu$ . When a clock of an edge in  $E \setminus A$  rings, refresh the state of the edge to open with probability  $p$  and closed with probability  $1 - p$ . If the exponential clock of rate  $|A|\mu$  rings at time  $t$ , choose an index from  $\{1, \dots, |A|\}$  uniformly at random and remove the edge with this label from the set  $R_{t-}$  in order to obtain the set  $R_t$ . If the edge that was chosen to be removed was an edge of  $E$ , then we also refresh its state.

With this construction  $(X, \eta)$  has the correct transition rates. This construction also enables us to couple different systems by keeping their infected sets of the same size. More specifically, suppose that  $(X^1, \eta^1), \dots, (X^m, \eta^m)$  are such that at time 0 they are i.i.d. and distributed according to  $\pi \times \pi_p$  and take  $R_0^1 = \emptyset, \dots, R_0^m = \emptyset$ . Then, we couple them all together by using the same exponential clocks of rate 1 for the walk components, and, in order to remove edges from the sets  $R_{t-}^i$ , we use the same exponential clocks of rate  $|R_{t-}^i| \mu$  and choose the same uniform number from  $\{1, \dots, R_{t-}^i\}$ . For the edges not in  $R_{t-}^i$ , we assign independent exponential clocks to the different systems. Finally, when a clock rings (either of the walk or of an edge of the set  $R$ ), the vertex to which the walk components jump or the new states of the edges chosen are independent for the different systems. With this coupling, indeed, the sizes of the sets  $R_t^i$  remain equal throughout for different  $i$ 's.

We now want to consider the jump chain of the full process. This is defined to be the full process  $(X, \eta)$  observed at the sequence of jump times which are the points of a Poisson process of parameter  $1 + \mu n$ .

The following lemma (whose proof follows immediately by induction on  $k$ ) shows that, with the coupling described above, the jump chains of the full processes are independent at every discrete time step.

LEMMA 3.1. *Suppose that  $(X_k^1, \eta_k^1)_{k \in \mathbb{N}}, \dots, (X_k^m, \eta_k^m)_{k \in \mathbb{N}}$  start independently according to  $\pi \times \pi_p$ , and suppose that an adversary prescribes in advance at which discrete time steps the  $X$  or the  $\eta$  coordinate will make a jump (i.e., the jumps of all of the systems occur simultaneously, and either in all of them, the walk coordinate attempts a move or in all of them the environment is updated). When the adversary chooses the  $\eta$  coordinate to update, he also prescribes which edge is going to be updated in each system (not necessarily the same edge!). At times at which the  $X$  coordinate is chosen, where the  $X$  coordinate jumps to is independent for the different chains. At times when the environment is chosen, the states of the chosen edges are also i.i.d. and become open with probability  $p$  and closed with probability  $1 - p$ . Then, for all  $k$  we have that  $(X_k^1, \eta_k^1), \dots, (X_k^m, \eta_k^m)$  are independent.*

LEMMA 3.2. *Consider a continuous time Markov chain with generator  $Q = e^{\lambda(P-I)}$  for  $\lambda > 0$ . Let  $X^1, \dots, X^m$  be continuous time Markov chains with generator  $Q$  which start from i.i.d. states distributed according to the invariant distribution  $\pi$ . Suppose that their exponential clocks are coupled in some way and at every discrete time step the jump chains (that evolve according to the transition matrix  $P$ ) are independent. Then, for every  $t$  we have*

$$\mathbb{P}(X_t^1 = x_1, \dots, X_t^m = x_m) = \prod_{i=1}^m \pi(x_i).$$

PROOF. Let  $N^1, \dots, N^m$  be the Poisson processes of rate  $\lambda$  associated to each Markov chain  $X^1, \dots, X^m$  that are coupled as in the statement of the lemma. Let  $Y^1, \dots, Y^m$  be the independent jump chains that start at time 0, independently, according to  $\pi$ . Then, we have

$$\begin{aligned} &\mathbb{P}(X_t^1 = x_1, \dots, X_t^m = x_m) \\ &= \sum_{k_1, \dots, k_m} \mathbb{P}(Y_{k_1}^1 = x_1, \dots, Y_{k_m}^m = x_m, N_t^1 = k_1, \dots, N_t^m = k_m) \\ &= \sum_{k_1, \dots, k_m} \mathbb{P}(Y_{k_1}^1 = x_1, \dots, Y_{k_m}^m = x_m) \mathbb{P}(N_t^1 = k_1, \dots, N_t^m = k_m) \\ &= \sum_{k_1, \dots, k_m} \pi(x_1) \cdots \pi(x_m) \mathbb{P}(N_t^1 = k_1, \dots, N_t^m = k_m) = \pi(x_1) \cdots \pi(x_m) \end{aligned}$$

and this completes the proof.  $\square$

DEFINITION 3.3. We define the auxiliary chain  $Y$  starting from  $x_0$  as follows: start  $\eta_0 \sim \pi_p$  and  $X_0 = x_0$ . Set  $R_0 = \emptyset$ , and for every  $t$  consider the set of infected edges  $R_t$ , as defined above. We define the regeneration times by letting  $\tau_0 = 0$ , and for every  $i \geq 0$  we set

$$\tau_{i+1} = \inf\{t \geq \tau_i + S_i : R_t = \emptyset\},$$

where  $\tau_i + S_i$  is the first time after time  $\tau_i$  that  $R_t$  becomes nonempty. The auxiliary chain is defined to be the discrete time chain given by  $Y_i = X_{\tau_i}$  for all  $i$ .

REMARK 3.4. Note that the  $(S_i)$ 's are i.i.d. having the exponential distribution with parameter 1 and are independent of  $(X_{\tau_i})_i$ .

LEMMA 3.5. We have that  $(\tau_i - \tau_{i-1})_{i \geq 1}$  are i.i.d., have mean  $e^{1/\mu}$  and have exponential tails. Moreover, the process  $(X_s, \eta_s, R_s)_s$  is positive recurrent.

PROOF. The first claim follows from the fact that  $|R_t|$  evolves as a birth and death chain with transition rates  $q(i, i + 1) = 1$  and  $q(i, i - 1) = \mu i$ .

We now prove that the chain  $(X, \eta, R)$  is positive recurrent. It suffices to prove that the state  $(x, \eta, \emptyset)$  with  $x \in V$  and  $\eta \in \{0, 1\}^E$  is positive recurrent. Let  $T$  be the first return time to  $(x, \eta, \emptyset)$ . Consider the chain  $Z_i = (X_{\tau_i}, \eta_{\tau_i})$ . This is clearly irreducible and positive recurrent, since it takes values in a finite state space. Let  $T_Z$  be the time it takes for  $Z$  to return to  $(x, \eta)$ . Then,  $T_Z$  has exponential tails, and we have

$$\mathbb{E}[T] \leq \sum_{i=1}^{\infty} \mathbb{E}[(\tau_i - \tau_{i-1})\mathbf{1}(T_Z \geq i)].$$

Using Cauchy–Schwarz and the exponential tails of both  $T_Z$  and  $(\tau_i - \tau_{i-1})$  proves the result. □

LEMMA 3.6. The invariant distribution of  $Y$  is  $\pi$ .

PROOF. In order to prove this result, we let  $(X^1, \eta^1), \dots, (X^m, \eta^m)$  be  $m$  systems started independently from stationarity and coupled in the way described above using the infected sets, so that at every time  $t$  they are independent and their infected sets all have the same size  $|R_t|$ . Then, by Lemmas 3.1 and 3.2 we get that  $(X_t^1, \eta_t^1), \dots, (X_t^m, \eta_t^m)$  are independent at every time  $t$  and are distributed according to  $\pi \times \pi_p$ .

Fix  $x \in V$ , and define

$$S_m(t) = \frac{1}{m} \sum_{i=1}^m \mathbf{1}(X_t^i = x).$$

Then,  $\mathbb{E}[S_m(t)] = \pi(x)$ , and by the strong law of large numbers we get that as  $m \rightarrow \infty$

$$S_m(t) \rightarrow \pi(x) \quad \text{almost surely.}$$

Using this and dominated convergence gives that as  $m \rightarrow \infty$

$$\mathbb{E}[S_m(t) \mid |R_t| = 0] = \frac{\mathbb{E}[S_m(t)\mathbf{1}(|R_t| = 0)]}{\mathbb{P}(|R_t| = 0)} \rightarrow \frac{\mathbb{E}[\pi(x)\mathbf{1}(|R_t| = 0)]}{\mathbb{P}(|R_t| = 0)} = \pi(x).$$

Using that

$$\mathbb{E}[S_m(t)\mathbf{1}(|R_t| = 0)] = \mathbb{P}(|R_t| = 0)\mathbb{E}[S_m(t) \mid |R_t| = 0]$$

and taking the limit as  $m \rightarrow \infty$ , we now obtain

$$(3.1) \quad \mathbb{P}(X_t^1 = x, |R_t| = 0) = \mathbb{P}(|R_t| = 0)\pi(x).$$

If we now take the limit as  $t \rightarrow \infty$ , then  $(X_t^1, \eta_t^1, R_t)$  will converge to its invariant distribution  $\mu$  which exists by Lemma 3.5. So we obtain

$$\mu(x, \{0, 1\}^E, \emptyset) = \nu(\emptyset)\pi(x),$$

where  $\nu$  is the marginal for  $R$  of the invariant distribution  $\mu$ .

In order to find the invariant distribution of  $Y$ , we use the ergodic theorem. First of all, it is easy to see that  $Y$  is an irreducible Markov chain. Let  $b$  be its invariant distribution. Then, by the ergodic theorem we will have that almost surely as  $n \rightarrow \infty$

$$\frac{1}{n} \sum_{i=1}^n \mathbf{1}(Y_i = x) \rightarrow b(x).$$

Using the ergodic theorem for  $(X, \eta, R)$ , we get that almost surely as  $t \rightarrow \infty$

$$\frac{1}{t} \int_0^t \mathbf{1}(X_s = x, R_s = \emptyset) ds \rightarrow \mu(x, \{0, 1\}^E, \emptyset) = \nu(\emptyset)\pi(x).$$

Let  $N_t = \sup\{i : \tau_i \leq t\}$ , and let  $(S_i)$  be i.i.d. exponential random variables of parameter 1. Then,

$$\frac{1}{t} \sum_{i=1}^{N_t-1} \mathbf{1}(X_{\tau_i} = x) S_i \leq \frac{1}{t} \int_0^t \mathbf{1}(X_s = x, R_s = \emptyset) ds \leq \frac{1}{t} \sum_{i=1}^{N_t} \mathbf{1}(X_{\tau_i} = x) S_i.$$

Using the above convergence, this now implies that almost surely as  $t \rightarrow \infty$

$$\frac{N_t}{t} \cdot \frac{1}{N_t} \sum_{i=1}^{N_t} \mathbf{1}(X_{\tau_i} = x) S_i \rightarrow \nu(\emptyset)\pi(x).$$

Since  $N_t$  is a renewal process with interarrival times having mean  $1/\nu(\emptyset)$ , we get almost surely

$$\frac{1}{N_t} \sum_{i=1}^{N_t} \mathbf{1}(X_{\tau_i} = x) S_i \rightarrow \pi(x) \quad \text{as } t \rightarrow \infty.$$

Now, we notice that, since  $N_t$  can increase by 1 at every jump, we get that the sequence  $a_n = n$  is a subsequence of  $(N_t)_t$ , and, hence, almost surely

$$\frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_{\tau_i} = x) S_i \rightarrow \pi(x) \quad \text{as } n \rightarrow \infty.$$

Now, to conclude we write  $L(n) = \sum_{i=1}^n \mathbf{1}(X_{\tau_i} = x)$ , and by relabelling  $S_i$  using the independence between  $(X_{\tau_i})_i$  and  $(S_i)_i$ , we obtain almost surely

$$\frac{1}{n} \sum_{i=1}^{L(n)} S_i \rightarrow \pi(x) \quad \text{as } n \rightarrow \infty.$$

Using the strong law of large numbers for the i.i.d. sequence  $(S_i)$  and that  $L(n) \rightarrow \infty$  as  $n \rightarrow \infty$ , we finally get almost surely

$$\frac{L(n)}{n} \rightarrow \pi(x) \quad \text{as } n \rightarrow \infty.$$

This shows that  $b(x) = \pi(x)$  for all  $x$  and concludes the proof.  $\square$

REMARK 3.7. Consider a graph  $G = (V, E)$  with the property that, for all  $x, y \in V$ , there is an automorphism  $\varphi$  of  $G$  such that  $\varphi(x) = y$  and  $\varphi(y) = x$ . Any Cayley graph of an Abelian group has this property. It is not hard to verify that graphs with this property satisfy that the transition matrix of the auxiliary chain is symmetric and, hence, reversible w.r.t. the uniform distribution.

We end this section with a lemma on the holding probabilities for the auxiliary chain that will be used later in the paper.

LEMMA 3.8. *Let  $Y$  be the auxiliary chain with parameters  $\mu = 1$  and  $p \in (0, 1)$ . Then, for all  $x \in V$  we have that  $P_{\text{aux}}(x, x) \geq 1 - 2e^2 p$  and  $(1 - p)/2 \leq P_{\text{aux}}(x, x) \leq 1 - \frac{p}{2}$ , where  $P_{\text{aux}}$  is the transition matrix of  $Y$ .*

PROOF. The claim that  $(1 - p)/2 \leq P_{\text{aux}}(x, x) \leq 1 - p/2$  follows by considering the case that the first edge to be examined by the walk after a regeneration time is closed (resp., open), and that this edge is then refreshed before the walk attempts to make further jumps.

We now prove  $\min_x P_{\text{aux}}(x, x) \geq 1 - 2e^2 p$ . In fact, we show that  $\alpha = \min_x \alpha_x \geq 1 - 2e^2 p$ , where  $\alpha_x$  is the probability that, after a regeneration time at  $x$ , the walk does not leave  $x$  before the next regeneration time. We first observe that, after a regeneration time at  $x$  until the walk coordinate leaves  $x$  for the first time, every edge it examines is open with probability at most  $p$ . To see this, by the strong Markov property we may consider starting at time 0 from a stationary environment with the walk coordinate at  $x$ . Finally, given that the edges  $e_1, \dots, e_k$  were examined, thus far, at times  $t_1 < \dots < t_k$  and were all closed, the probability that  $e_{k+1}$  is open at time  $t_{k+1} > t_k$  is exactly  $p$  if  $e_{k+1} \notin \{e_1, \dots, e_k\}$  and is  $p(1 - e^{t_i - t_{k+1}}) < p$  if  $e_{k+1} = e_i$  and  $e_{k+1} \notin \{e_{i+1}, \dots, e_k\}$ .

We now consider a birth and death chain with an additional point added to it. Its state space is  $\mathbb{Z}_+ \cup \{\infty\}$ . Its transition matrix  $Q$  is given by  $Q(0, 1) = 1 - p = 1 - Q(0, \infty)$ , while for  $i \in \mathbb{N}$  we have  $Q(i, i - 1) = \frac{i}{i+1}$ ,  $Q(i, i + 1) = \frac{(1-p)}{i+1}$  and  $Q(i, \infty) = \frac{p}{i+1}$ . Finally, we set  $Q(\infty, \infty) = 1$ . It is easy to see that

$$\alpha \geq \mathbb{P}_0^Q(T_0^+ < T_\infty),$$

where  $T_0^+$  stands for the first return to 0 and  $T_\infty$  for the first hitting time of  $\infty$  for the chain with matrix  $Q$ . We have that

$$(3.2) \quad \mathbb{P}_0^Q(T_0^+ < T_\infty) = (1 - p)\mathbb{P}_1^Q(T_0 < T_\infty).$$

Consider now a birth and death chain on  $\mathbb{Z}_+$  with transition matrix  $W(x, y) = \frac{Q(x, y)}{1 - Q(x, \infty)}$ . Let  $N$  be a geometric random variable of success probability  $p$  independent of the Markov chain  $W$ . Since the probability of jumping to  $\infty$  for the Markov chain  $Q$  is at most  $p$  at every step, we can lower bound  $\mathbb{P}_1^Q(T_0 < T_\infty)$  by the probability that the Markov chain  $W$  hits 0 starting from 1 before time  $N$ , that is,

$$\mathbb{P}_1^Q(T_0 < T_\infty) \geq \mathbb{P}_1^W(T_0 < N).$$

We now turn to lower bound  $\mathbb{P}_1^W(T_0 < N)$ . Since from now on we will only be working with the chain  $W$ , we drop the letter  $W$  from the notation. Let  $X$  be the birth and death chain with matrix  $W$ . We let  $Z = \sum_{i=0}^N \mathbf{1}(X_i = 0)$ . Then,

$$(3.3) \quad \mathbb{P}_1(T_0 < N) = \mathbb{P}_1(Z > 0) = \frac{\mathbb{E}_1[Z]}{\mathbb{E}_1[Z \mid Z > 0]} = \frac{\mathbb{E}_1[Z]}{\mathbb{E}_0[Z]},$$

where for the last equality we used the memoryless property of  $N$ . Using the independence of  $N$  and  $X$  and that  $W(0, 1) = 1$ , we get

$$(3.4) \quad \mathbb{E}_1[Z] = \sum_{i=1}^{\infty} \mathbb{P}_1(X_i = 0)(1 - p)^{i-1} \quad \text{and} \quad \mathbb{E}_0[Z] = 1 + \mathbb{E}_1[Z].$$

We notice that  $W(1, 0) = \frac{1}{2-p}$ , and so  $\mathbb{P}_1(X_i = 0) = \mathbb{P}_1(X_{i-1} = 1)/(2 - p)$ . Therefore, this gives

$$\mathbb{E}_1[Z] = \frac{1}{2 - p} \sum_{i=0}^{\infty} \mathbb{P}_1(X_i = 1)(1 - p)^i = \frac{1}{2 - p} \sum_{i=0}^{\infty} \mathbb{P}_1(X_{2i} = 1)(1 - p)^{2i}.$$

Solving the detailed balance equations it is straightforward to see that  $X$  has an invariant distribution satisfying  $\nu(1) \geq e^{-2}$ . Using that  $\mathbb{P}_1(X_{2i} = 1)$  is decreasing as a function of  $i$  and converges to  $2\nu(1)$  as  $i \rightarrow \infty$ , we obtain

$$\mathbb{E}_1[Z] \geq \frac{2\nu(1)}{2 - p} \sum_{i=0}^{\infty} (1 - p)^{2i} = \frac{2\nu(1)}{p(2 - p)^2} \geq \frac{1}{2pe^2}.$$

Substituting this bound and (3.4) into (3.3), we deduce

$$\mathbb{P}_1(T_0 < N) \geq \frac{1}{2e^2 p + 1 - p} \geq 1 - (2e^2 - 1)p.$$

Substituting this into (3.2), we finally get  $\alpha \geq 1 - 2e^2 p$ , and this concludes the proof.  $\square$

**4. Comparison of hitting times.** In this section we prove Theorem 1.1. Before doing so, we review some results about hitting and commute times.

Let  $(Y_k)_{k \in \mathbb{N}}$  be an irreducible discrete time Markov chain on a finite state space  $\Omega$  with transition matrix  $P$  and stationary distribution  $\pi$ . For  $a \in \Omega$  we write

$$T_a = \inf\{t \geq 0 : Y_t = a\} \quad \text{and} \quad T_a^+ = \inf\{t \geq 1 : Y_t = a\}.$$

For distinct states  $a, b \in \Omega$ , we write  $T_{ab}$  for the commute time between  $a$  and  $b$  (for the chain starting at  $a$ ), that is,

$$T_{ba} = \inf\{t > T_b : Y_t = a\}.$$

We write  $P^*$  for the time-reversal of  $P$ , that is, it is the transition matrix given by

$$\pi(x)P^*(x, y) = \pi(y)P(y, x).$$

Let  $\lambda(A)$  be the smallest eigenvalue of  $Q_A$ , the generator of the chain killed upon exiting  $A$ . Writing  $Q = \gamma(P - I)$  for some transition matrix  $P$  and  $\gamma > 0$  and using the Perron Frobenius Theorem, we have that ([1], Chapter 3)

$$(4.1) \quad \lambda(A) = \min\{\mathcal{E}_Q(h, h) : h \in \mathbb{R}^\Omega, \|h\|_2 = 1, \text{supp}(h) \subseteq A\},$$

where  $\|h\|_2 = (\mathbb{E}_\pi[|h|^2])^{1/2}$ .

We next recall Dirichlet’s principle for effective resistance. For a proof, see [1], Theorem 3.36.

**THEOREM 4.1.** *Let  $Y$  be a reversible Markov chain on a finite state space  $\Omega$  with transition matrix  $P$  and stationary distribution  $\pi$ . Then, for all  $a, b \in \Omega$  we have*

$$\frac{1}{\mathbb{E}_a[T_{ba}]} = \inf_{f \in \mathbb{R}^\Omega : f(a)=1, f(b)=0, 0 \leq f \leq 1} \mathcal{E}_P(f, f).$$



The following result is due to Doyle and Steiner [11] and was also rediscovered by Gaudilière and Landim [12]. We include its short proof for the reader’s convenience.

LEMMA 4.2. *Let  $Y$  be a Markov chain on a finite state space  $\Omega$  with transition matrix  $P$  and stationary distribution  $\pi$ . Let  $P^*$  be its time reversal and  $S = (P + P^*)/2$ . Then, for all  $a, b \in \Omega$  we have*

$$\mathbb{E}_a^S[T_{ba}] \leq \mathbb{E}_a^P[T_{ba}].$$

PROOF. Let  $v(x) = \mathbb{P}_x(T_a < T_b)$ . Then, using that  $v$  is harmonic off  $\{a, b\}$  (and that  $\frac{1}{\mathbb{E}_a^P[T_{ba}]} = \pi(a)\mathbb{P}_a^P[T_b < T_a^+]$ , which can be derived using Wald’s equation, by considering the number of returns of  $a$  before the first excursion that visits  $b$ ), we get

$$\frac{1}{\mathbb{E}_a^P[T_{ba}]} = \mathcal{E}_P(v, v) = \mathcal{E}_S(v, v) \geq \frac{1}{\mathbb{E}_a^S[T_{ba}]},$$

where the inequality follows from Theorem 4.1 applied to the reversible matrix  $S$  using that  $v(b) = 0$  and  $v(a) = 1$ .  $\square$

The following is similar to [16], page 12. We give the proof for the sake of completeness.

LEMMA 4.3. *There exists a positive constant  $c$  so that for every reversible Markov chain on a finite state space  $\Omega$  with stationary distribution  $\pi$ , for all  $A \subset \Omega$  we have that*

$$\max_{x \in \Omega} \mathbb{E}_x[T_A] \leq \mathbb{E}_\pi[T_A] + ct_{\text{mix}}.$$

PROOF. We start by recalling another notion of mixing introduced by Aldous [2] in the continuous time case and later studied in discrete time by Lovász and Winkler [23]. We let

$$(4.2) \quad \begin{aligned} t_{\text{stop}} &= \max_x \inf\{\mathbb{E}_x[T] : T \text{ randomised stopping time} \\ &\text{s.t. } \mathbb{P}_x(X_T = y) = \pi(y) \text{ for all } y\}. \end{aligned}$$

Note that, for every  $x$ , there is a randomised stopping time that achieves the infimum above (see [2] and [23]). Let  $x \in \Omega$ , and let  $T$  be a stopping time such that  $\mathbb{P}_x(X_T \in \cdot) = \pi(\cdot)$  and  $\mathbb{E}_x[T] \leq t_{\text{stop}}$ . Writing  $a_+ = \max\{0, a\}$  we then get

$$\mathbb{E}_x[T_A] \leq \mathbb{E}_x[T] + \mathbb{E}_x[(T_A - T)_+] \leq t_{\text{stop}} + \mathbb{E}_\pi[T_A],$$

where we have used  $\mathbb{E}_x[(T_A - T)_+] \leq \mathbb{E}_x[\inf\{t \geq T : X_t \in A\} - T] = \mathbb{E}_\pi[T_A]$ . The proof is now concluded using the result of Aldous [2] that there exists a positive constant  $c$  so that for all reversible chains  $t_{\text{stop}} \leq ct_{\text{mix}}$ .  $\square$

LEMMA 4.4. *Consider the auxiliary chain  $Y$  from Definition 3.3 with parameters  $\mu$  and  $p$ . Denote its transition matrix by  $P_{\text{aux},(\mu,p)}$ . Then, for all  $x, y$  such that  $\{x, y\} \in E$  we have that*

$$\begin{aligned} P_{\text{aux},(\mu,p)}(x, y) &\geq P_{\text{SRW}}(x, y)p \frac{\mu}{1 + \mu} \quad \text{and} \\ \frac{P_{\text{aux},(\mu,p)}(x, y) + P_{\text{aux},(\mu,p)}^*(x, y)}{2} &\geq P_{\text{SRW}}(x, y)p \frac{\mu}{1 + \mu}. \end{aligned}$$

PROOF. Start the environment at time 0 from its stationary distribution, and the walk coordinate of the full process from  $x$ . Now, consider the event that the first attempt of a jump of the walk coordinate is to  $y$ , that the edge  $\{x, y\}$  is open at that time and that, before the walk attempts to make its second jump the edge,  $\{x, y\}$  is refreshed. The probability of this event is at least  $P_{\text{SRW}}(x, y)p \frac{\mu}{1+\mu}$  and at most  $P_{\text{aux},(\mu,p)}(x, y)$ .

Using the first inequality, the fact that the invariant distribution for the auxiliary chain is the same as for the simple random walk on  $G$  by Lemma 3.6 and that the simple random walk is a reversible chain, we get that for all  $\{x, y\} \in E$

$$\frac{P_{\text{aux},(\mu,p)}(x, y) + P_{\text{aux},(\mu,p)}^*(x, y)}{2} \geq P_{\text{SRW}}(x, y)p \frac{\mu}{1 + \mu},$$

and this concludes the proof.  $\square$

We are now ready to prove Theorem 1.1.

PROOF OF THEOREM 1.1. Let  $X_0 = x$ ,  $\eta_0 = \eta$ , and suppose that the infection process  $R$  starts with all edges of  $G$  infected, that is,  $|R_0| = |E|$ . Let  $\tau$  be the first regeneration time, that is, the first time that the size of  $R$  becomes 0. Then, for all  $x, \eta$ , using that  $|R_t|$  evolves as a birth and death chain with  $q(i, i + 1) = 1$  and  $q(i, i - 1) = \mu i$ , we get that

$$(4.3) \quad \mathbb{E}[\tau] \asymp_{\mu} \log |E|.$$

We now claim that  $\eta_{\tau}$  is distributed according to  $\pi_p$  and is independent of  $X_{\tau}$ . Indeed, considering for every edge examined by the walk the last time before  $\tau$  that this happened proves the claim, since edges that were not examined are distributed according to  $\text{Ber}(p)$ . Therefore, we obtain that

$$(4.4) \quad t_{\text{hit}}^{\text{full},(\mu,p)} \leq \mathbb{E}[\tau] + \max_{x,y \in V} \mathbb{E}_{x,\pi_p}^{(\mu,p)}[T_{y \times \{0,1\}^E}].$$

Consider the auxiliary chain with parameters  $\mu$  and  $p$ , denote its transition matrix by  $P_{\text{aux},(\mu,p)}$  and let  $S = (P_{\text{aux},(\mu,p)} + P_{\text{aux},(\mu,p)}^*)/2$  be its additive symmetrisation. Let  $P_{\text{SRW}}$  denote the transition matrix of the simple random walk on  $G$ .

Lemma 4.4 immediately implies that  $\mathcal{E}_S(f, f) \geq \frac{p\mu}{1+\mu} \mathcal{E}_{P_{\text{SRW}}}(f, f)$  for all  $f \in \mathbb{R}^V$ . This together with Theorem 4.1 gives that for all  $x, y \in V$

$$p \frac{\mu}{1 + \mu} \mathbb{E}_x^{S,(\mu,p)}[T_{yx}] \leq \mathbb{E}_x^{\text{SRW}}[T_{yx}].$$

Applying Lemma 4.2, we obtain that

$$\mathbb{E}_x^{\text{aux},(\mu,p)}[T_{yx}] \leq \mathbb{E}_x^{S,(\mu,p)}[T_{yx}].$$

We next let

$$\tilde{T}_{yx} = \inf\{t > T_{y \times \{0,1\}^E} : X_t = x\}.$$

Let  $\tau_i$  be the  $i$ th regeneration time. Then, by Lemma 3.5 the variables  $\sigma_i = \tau_i - \tau_{i-1}$  are i.i.d. of mean  $e^{1/\mu}$ . Using the natural coupling of the full process started from  $(\delta_x, \pi_p)$  with the auxiliary chain started from  $x$ , we obtain

$$\tilde{T}_{yx} \leq \sum_{i=1}^{T_{yx}} \sigma_i$$

and, hence, using Wald’s identity

$$p \frac{\mu}{1 + \mu} \mathbb{E}_x^{\pi_p,(\mu,p)}[\tilde{T}_{xy}] \leq e^{1/\mu} \mathbb{E}_x^{\text{SRW}}[T_{xy}].$$

This together with (4.3) and (4.4) now implies

$$t_{\text{hit}}^{\text{full},(\mu,p)} \lesssim_{\mu} \frac{1}{p} t_{\text{hit}}^{\text{SRW}} + \log |E|.$$

Using the fact that  $t_{\text{hit}}^{\text{SRW}} \geq |V| - 1$ , we can absorb the  $\log |E|$  term inside the first term and this proves (1.1).

Next, we prove (1.2). Let  $x, y \in V$  and  $\eta \in \{0, 1\}^E$ . By Lemma 4.3,

$$(4.5) \quad \mathbb{E}_{x,\eta}^{(\mu_i,p)} [T_{\{y\} \times \{0,1\}^E}] \leq c t_{\text{mix}}^{\text{full},(\mu_i,p)} + \mathbb{E}_{\pi_{\text{full},p}^{(\mu_i,p)}} [T_{\{y\} \times \{0,1\}^E}].$$

Recall ([1], Chapter 3) that, for a general reversible Markov chain on a finite state space with stationary distribution  $\pi$ , for every set  $A$  there exists a distribution  $\mu_A$ , known as the quasi-stationary distribution of  $A$ , such that

$$(4.6) \quad \mathbb{E}_{\pi} [T_A] \leq \mathbb{E}_{\mu_A} [T_A] = 1/\lambda(A) \leq \max_x \mathbb{E}_x [T_A],$$

where  $\lambda(A)$  is the smallest eigenvalue of the generator of the chain killed upon exiting  $A$ . Applying this to  $A = \{y\} \times \{0, 1\}^E$ , we have that

$$\begin{aligned} \mathbb{E}_{\pi_{\text{full},p}^{(\mu_1,p)}} [T_{\{y\} \times \{0,1\}^E}] &\leq \frac{1}{\lambda_{(\mu_1,p)}(\{y\} \times \{0, 1\}^E)} \\ &\leq \frac{\mu_2}{\mu_1 \lambda_{(\mu_2,p)}(\{y\} \times \{0, 1\}^E)} \leq \frac{\mu_2}{\mu_1} t_{\text{hit}}^{\text{full},(\mu_2,p)}, \end{aligned}$$

where for the second inequality we used (4.1) together with the fact that if  $\mu_1 < \mu_2$ , then

$$\mathcal{L}_1((x, \eta), (x', \eta')) \leq \frac{\mu_2}{\mu_1} \mathcal{L}_2((x, \eta), (x', \eta'))$$

with  $\mathcal{L}_i$  denoting the generator of the full process when the rate at which the edges update is  $\mu_i$ . This together with (4.5) concludes the proof.  $\square$

**5. The spectral profile.** We now recall a couple of results from [13]. While some of the results below were originally stated in the case where the infinitesimal generator  $\mathcal{L}$  is of the form  $K - I$ , where  $I$  is the identity matrix and  $K$  is a transition matrix of a discrete-time Markov chain (possibly with nonzero diagonal entries), they hold for general  $\mathcal{L}$  when the state space is finite, as we can always write  $\mathcal{L} := c(K - I)$  for some  $c > 0$  and some transition matrix  $K$ , possibly with positive diagonal entries. (Namely,  $c = \max(-\mathcal{L}(x, x))$  and  $K = \mathcal{L}/c + I$ .) (All the quantities considered below scale linearly in  $c$ .)

Consider an irreducible continuous-time Markov chain on a finite state space  $\Omega$  with infinitesimal generator  $\mathcal{L}$  and stationary distribution  $\pi$ . Let  $P_t := e^{t\mathcal{L}}$ . Recall that the time  $t$  transition probabilities are given by  $P_t(\cdot, \cdot)$  and that the corresponding rates are given by  $\mathcal{L}(x, y) := \lim_{h \rightarrow 0} h^{-1}(P_h(x, y) - \mathbf{1}(x = y))$ . Denote the infinitesimal generators of the time reversal and additive symmetrization by  $\mathcal{L}^*$  and  $\mathcal{L}^s := \frac{1}{2}(\mathcal{L} + \mathcal{L}^*)$ , respectively, where  $\mathcal{L}^*$  is the dual operator of  $\mathcal{L}$  in  $(L_2(\Omega), \langle \cdot, \cdot \rangle_{\pi})$ , whose rates are given explicitly by  $\pi(x)\mathcal{L}^*(x, y) := \pi(y)\mathcal{L}(y, x)$ , where  $\langle f, g \rangle_{\pi} := \mathbb{E}_{\pi}[fg]$  is the inner product on  $\mathbb{R}^{\Omega}$  induced by  $\pi$  and  $\mathbb{E}_{\pi}[h] := \sum_{x \in \Omega} \pi(x)h(x)$  the stationary expectation of  $h \in \mathbb{R}^{\Omega}$ . We say that  $\mathcal{L}$  (and the corresponding Markov chain) is reversible if  $\mathcal{L} = \mathcal{L}^*$ . One can readily check that  $\pi$  is stationary also for  $\mathcal{L}^*$  and that  $\mathcal{L}^s$  is reversible w.r.t.  $\pi$ , as  $(\mathcal{L}^*)^* = \mathcal{L}$ .

Recall that for  $\varepsilon > 0$ , we denote the spectral profile

$$(5.1) \quad \Lambda(\varepsilon) = \min\{\mathcal{E}(h, h) : h \in \mathbb{R}^{\Omega}, \text{Var}_{\pi}(h) = 1, \pi(\text{supp}(h)) \leq \varepsilon\},$$

where  $\text{supp}(h) = \{x \in \Omega : h(x) \neq 0\}$  is the *support* of  $h$ . We write  $\lambda = \Lambda(1)$  for the *Poincaré constant*. We recall that in the reversible setting  $\lambda$  is equal to the spectral gap. A related notion is

$$(5.2) \quad \Lambda_0(\varepsilon) = \min\{\mathcal{E}(h, h) : h \in \mathbb{R}^\Omega, \|h\|_2 = 1, \pi(\text{supp}(h)) \leq \varepsilon\},$$

where for  $h \in \mathbb{R}^\Omega$ ,  $\|h\|_\infty = \max_{x \in \Omega} |h(x)|$  and  $\|h\|_p = (\mathbb{E}_\pi[|h|^p])^{1/p}$  are the  $L_\infty$  and  $L_p$  ( $p \in [1, \infty)$ ) norms respectively. By Cauchy–Schwarz

$$\|h\|_1^2 = \|h\mathbf{1}(\text{supp}(h))\|_1^2 \leq \|h\|_2^2 \pi(\text{supp}(h)),$$

and so  $1 - \pi(\text{supp}(h)) \leq \frac{\text{Var}_\pi(h)}{\|h\|_2^2} \leq 1$ . Hence, we get that

$$(5.3) \quad (1 - \varepsilon)\Lambda(\varepsilon) \leq \Lambda_0(\varepsilon) \leq \Lambda(\varepsilon).$$

As we now explain,  $\Lambda_0$  has a probabilistic interpretation, which we shall exploit later on. For each  $A \subsetneq \Omega$ , we define  $\mathcal{L}_A$  to be the generator of the chain that is killed upon exiting  $A$ , whose rates are given explicitly by  $\mathcal{L}_A(u, v) = \mathcal{L}(u, v)\mathbf{1}(u, v \in A)$ , and define  $\lambda(A)$  to be the smallest eigenvalue of  $-\mathcal{L}_A^s = -\frac{1}{2}(\mathcal{L}_A + \mathcal{L}_A^*)$ . Under reversibility

$$(5.4) \quad \lambda(A) = \sup\left\{c : \max_{a \in A} \mathbb{E}_a[\exp(cT_{A^c})] < \infty\right\} = \sup\{c : \mathbb{E}_\pi[\exp(cT_{A^c})] < \infty\}.$$

Indeed, if  $B = \sup\{c : \max_{a \in A} \mathbb{E}_a[\exp(cT_{A^c})] < \infty\}$  and  $C = \sup\{c : \mathbb{E}_\pi[\exp(cT_{A^c})] < \infty\}$ , then, by definition,  $C \geq B$ , but as  $\pi$  has full support we actually have that  $C = B$ . Now, it is classical ([1], Chapter 4, or [4], Section 3) that under reversibility the law of  $T_{A^c}$  under initial distribution  $\pi$  is a mixture of an atom at 0 of mass  $\pi(A^c)$  and of exponential distributions whose minimal parameter is exactly  $\lambda(A)$ . From which we easily get that  $\lambda(A) = C$ .

Writing  $\mathcal{L} = c(P - I)$  for some transition matrix  $P$  and  $c > 0$  and using the Perron–Frobenius Theorem (to argue that the minimum can be attained by some nonnegative  $h$ ), we also have that (e.g., [1], Chapter 3)

$$(5.5) \quad \begin{aligned} \lambda(A) &= \min\{\mathcal{E}(h, h) : h \in \mathbb{R}^\Omega, \|h\|_2 = 1, \text{supp}(h) \subseteq A\} \\ &= \min\{\mathcal{E}(h, h) : h \in \mathbb{R}_+^\Omega, \|h\|_2 = 1, \text{supp}(h) \subseteq A\}, \end{aligned}$$

and so

$$(5.6) \quad \Lambda_0(\varepsilon) = \min\{\lambda(A) : \pi(A) \leq \varepsilon\}.$$

5.1. *Decay of  $L_2$  distances via the spectral-profile and the Poincaré constant.* Recall that the  $L_p$  norm (w.r.t.  $\pi$ ) of a signed measure  $\sigma$  is defined as

$$\|\sigma\|_{p,\pi} = \|\sigma/\pi\|_p \quad \text{where } (\sigma/\pi)(x) = \sigma(x)/\pi(x).$$

In particular, for a distribution  $\nu$  its  $L_2$  distance from  $\pi$  satisfies

$$\|\nu - \pi\|_{2,\pi}^2 = \|\nu/\pi - 1\|_2^2 = \text{Var}_\pi(\nu/\pi).$$

Let  $\nu_t = P_t^\nu := \nu e^{\mathcal{L}t}$  and  $u_t = \nu_t/\pi = e^{t\mathcal{L}^*}(\frac{\nu}{\pi})$  (where  $\nu e^{\mathcal{L}t}(x) = \sum_y \nu(y)P_t(y, x)$ , while  $e^{t\mathcal{L}^*}f(x) = \sum_y e^{t\mathcal{L}^*}(x, y)f(y)$ ). For  $f \in \mathbb{R}^\Omega$ , writing  $f_t = e^{t\mathcal{L}^*}f$ , we have  $\frac{d}{dt} \text{Var}_\pi(f_t) = \langle \mathcal{L}^*f_t, f_t \rangle_\pi + \langle f_t, \mathcal{L}^*f_t \rangle_\pi = 2\langle \mathcal{L}f_t, f_t \rangle_\pi = -2\mathcal{E}(f_t, f_t)$  (cf. [22], page 284) and so

$$(5.7) \quad \frac{d}{dt} \text{Var}_\pi(u_t) = \frac{d}{dt} \text{Var}_\pi\left(e^{t\mathcal{L}^*}\left(\frac{\nu}{\pi}\right)\right) = -2\mathcal{E}(u_t, u_t).$$

By (5.1) and the fact that the spectral gap  $\lambda$  equals  $\Lambda(1)$  we get  $\mathcal{E}(u_t, u_t) \geq \lambda \text{Var}_\pi(u_t)$  from which it follows that  $\frac{d}{dt} \text{Var}_\pi(u_t) \leq -2\lambda \text{Var}_\pi(u_t)$ , and so by Grönwall’s lemma

$$(5.8) \quad \|\nu_t - \pi\|_{2,\pi}^2 \leq \|\nu - \pi\|_{2,\pi}^2 \exp(-2\lambda t).$$

This is the well-known Poincaré inequality. The  $\varepsilon$   $L_p$ -mixing time is defined as

$$t_{\text{mix}}^{(p)}(\varepsilon) = \inf \left\{ t : \max_x \|P^t(x, \cdot) - \pi\|_{p,\pi} \leq \varepsilon \right\}.$$

It is standard (e.g., [13] or [22], Proposition 4.15) that for reversible Markov chains, for all  $x \in \Omega$  and  $t$  we have

$$(5.9) \quad \max_{y,z \in \Omega} \left| \frac{P_t(y, z)}{\pi(z)} - 1 \right| = \max_y \frac{P_t(y, y)}{\pi(y)} - 1 \quad \text{and} \quad \|P_x^t - \pi\|_{2,\pi}^2 = \frac{P_{2t}(x, x)}{\pi(x)} - 1.$$

Thus,  $t_{\text{mix}}^{(\infty)}(\varepsilon^2) = 2t_{\text{mix}}^{(2)}(\varepsilon)$  for all  $\varepsilon \leq (\max_x \frac{1-\pi(x)}{\pi(x)})^{1/2}$ .

Theorem 1.1 in [13] asserts that

$$(5.10) \quad \forall \varepsilon \in (0, 1/\pi_*], \quad t_{\text{mix}}^{(\infty)}(\varepsilon) \leq t_{\text{spectral-profile}}(\varepsilon).$$

We shall use a variant of this where we want to bound the  $L_2$  mixing-time starting from some initial distribution  $\nu$  for which we have a decent upper bound on  $\|\nu - \pi\|_{2,\pi}$ .

PROPOSITION 5.1 ([13], Lemma 2.1). *For any (nonzero)  $u \in \mathbb{R}_+^\Omega$ , we have that*

$$\frac{\mathcal{E}(u, u)}{\text{Var}_\pi(u)} \geq \frac{1}{2} \Lambda(4\|u\|_1^2 / \text{Var}_\pi(u)).$$

We recall the proof from [13] for the reader’s convenience.

PROOF. Let  $M = \text{Var}_\pi(\frac{u}{\sqrt{4\|u\|_1}})$  and  $B = \{x : u(x) \geq M\}$ . The set  $B$  is nonempty by Hölder’s inequality. Considering  $\hat{u} := (u - M)1_B = (u - M)_+$ , we now have  $\mathcal{E}(u, u) \geq \mathcal{E}(\hat{u}, \hat{u})$ . Also,

$$\begin{aligned} \text{Var}_\pi(\hat{u}) &\geq \mathbb{E}_\pi[(u - M)_+^2] - (\mathbb{E}_\pi[u])^2 \geq \|u\|_2^2 - 2M\|u\|_1 - \|u\|_1^2 \\ &= \text{Var}_\pi(u) - 2M\|u\|_1 = \frac{1}{2} \text{Var}_\pi(u). \end{aligned}$$

Finally,  $2\frac{\mathcal{E}(u,u)}{\text{Var}_\pi(u)} \geq \frac{\mathcal{E}(\hat{u},\hat{u})}{\text{Var}_\pi(\hat{u})} \geq \Lambda(\pi(\text{supp}(\hat{u}))) \geq \Lambda(\frac{4\|u\|_1^2}{\text{Var}_\pi(u)})$ , since by Markov’s inequality  $\pi(\text{supp}(\hat{u})) \leq \frac{\|u\|_1}{M} = \frac{4\|u\|_1^2}{\text{Var}_\pi(u)}$ .  $\square$

Recall that  $u_t := \nu_t / \pi$  and  $\nu_t := P_\nu^t$ . Using Proposition 5.1 and (5.7), it is not difficult to verify the assertion of the following proposition, which refines (5.8) and (5.10).

PROPOSITION 5.2 ([13], Theorem 1.1). *For any initial distribution  $\nu$  we have that*

$$(5.11) \quad \|\nu_t - \pi\|_{2,\pi}^2 \leq M \quad \text{if } t \geq \int_{4/\|\nu - \pi\|_{2,\pi}^2}^{4/M} \frac{d\delta}{\delta \Lambda(\delta)}.$$

In particular, for all  $0 < \beta < 1$  we have that

$$(5.12) \quad \|\nu_t - \pi\|_{2,\pi}^2 \leq \beta \|\nu - \pi\|_{2,\pi}^2 \quad \text{if } t \geq \frac{\log(1/\beta)}{\Lambda(4/(\beta \|\nu - \pi\|_{2,\pi}^2))}.$$

We shall sometimes use the following simple variant of (5.12).

LEMMA 5.3. *If  $\|\nu - \pi\|_{2,\pi}^2 \leq M$ , then for all  $0 < \beta < 1$  we have that*

$$(5.13) \quad \|\nu_t - \pi\|_{2,\pi}^2 \leq \beta M \quad \text{if } t \geq \frac{\log(1/\beta)}{\Lambda(4/(\beta M))}.$$

PROOF. If  $\|v - \pi\|_{2,\pi}^2 \leq \beta M$ , this follows from the fact that  $P_t$  is a contraction in  $L_2$ , that is,  $\|v_t - \pi\|_{2,\pi}^2 \leq \|v - \pi\|_{2,\pi}^2 \leq \beta M$ . If  $\|v - \pi\|_{2,\pi}^2 \in (\beta M, M]$ , then by (5.12) for  $\beta' = \frac{\beta M}{\|v - \pi\|_{2,\pi}^2} \geq \beta$ ,  $s = \frac{\log(1/\beta')}{\Lambda(4/(\beta'\|v - \pi\|_{2,\pi}^2))} = \frac{\log(1/\beta')}{\Lambda(4/(\beta M))} \leq \frac{\log(1/\beta)}{\Lambda(4/(\beta M))}$  we have  $\|v_t - \pi\|_{2,\pi}^2 \leq \beta' \|v - \pi\|_{2,\pi}^2 = \beta M$ .  $\square$

5.2. *A lower bound on  $L_2$  distances in terms of small sets probabilities.* Let  $\mathcal{P}(\Omega)$  be the collection of all distributions on  $\Omega$ . Let  $A \subsetneq \Omega$  and  $\delta \in (0, 1)$ . Let

$$\mathcal{P}_{A,\delta} := \{v \in \mathcal{P}(\Omega) : v(A) \geq \pi(A) + \delta\pi(A^c)\}.$$

Note that

$$v_{A,\delta} := \delta\pi_A + (1 - \delta)\pi \in \mathcal{P}_{A,\delta},$$

where  $\pi_A(a) := \frac{1_{\{a \in A\}}\pi(a)}{\pi(A)}$  is  $\pi$  conditioned on  $A$ . Moreover,  $\min\{\delta' : v_{A,\delta'} \in \mathcal{P}_{A,\delta}\} = \delta$ . It is thus intuitive that, for a convex distance function between distributions,  $v_{A,\delta}$  is the closest distribution to  $\pi$  in  $\mathcal{P}_{A,\delta}$ .

PROPOSITION 5.4 ([17], Proposition 4.1). *Let  $A \subsetneq V$ . Denote  $v_{A,\delta} = \delta\pi_A + (1 - \delta)\pi$ . Then,*

$$(5.14) \quad \forall \delta \in (0, 1) \quad \min_{v \in \mathcal{P}_{A,\delta}} \|v - \pi\|_{2,\pi}^2 = \|v_{A,\delta} - \pi\|_{2,\pi}^2 = \delta^2\pi(A^c)/\pi(A).$$

The proof is an exercise in Lagrange multipliers (see [17], Proposition 4.1).

5.3. *A mixing time bound for the  $p$ -tilted hypercube.* We call the hypercube  $\{0, 1\}^d$  equipped with the product measure  $\pi_p$  the  $p$ -tilted hypercube. The natural dynamics associated with it is the one at which each coordinate is updated at rate  $\mu$  and takes the value 1 with probability  $p$  and 0 with probability  $1 - p$ . This is precisely the evolution of the environment  $\eta$ .

LEMMA 5.5. *Denote by  $P_t$  the time  $t$  transition kernel of the  $p$ -tilted hypercube with update rate  $\mu$  for each coordinate. Let  $\alpha := p \wedge (1 - p)$ . Let  $t(\delta) = t_{p,\mu}(\delta) := \frac{1}{\mu} \log(d \frac{1-\alpha}{\alpha \log(1+\delta)})$ . Then,*

$$(5.15) \quad \max_{x,y \in \{0,1\}^d} \left| \frac{P_{t(\delta)}(x,y)}{\pi_p(y)} - 1 \right| \leq \delta.$$

PROOF. By scaling we can assume  $\mu = 1$ . Denote the transition kernel of a single coordinate by  $Q_t$ . Let  $v_p(1) := p =: 1 - v_p(0)$ . Then,  $Q_t(a, a) = e^{-t} + (1 - e^{-t})v_p(a)$  for  $a \in \{0, 1\}$ . Now,

$$\frac{P_t(x, x)}{\pi_p(x)} = \prod_{i \in [d]} \frac{Q_t(x_i, x_i)}{v_p(x_i)} \leq \prod_{i \in [d]} \left( 1 + \frac{1 - \alpha}{\alpha e^t} \right) \leq \exp\left(d \frac{1 - \alpha}{\alpha e^t}\right)$$

(using  $1 + x \leq e^x$ ). The proof is concluded using (5.9) by substituting  $t = t(\delta)$  above.  $\square$

**6. Proof of Theorem 1.2.**

DEFINITION 6.1. Let  $A \subset V \times \{0, 1\}^E$ . For every  $a \in V$ , we define

$$\text{Env}(a, A) = \{\eta \in \{0, 1\}^E : (a, \eta) \in A\}.$$

For  $\alpha \in [0, 1]$  we say that  $a \in V$  is  $(A, \alpha)$ -environmentally friendly if  $\pi_p(\text{Env}(a, A)) \geq \alpha$ . We denote the collection of  $(A, \alpha)$ -environmentally friendly vertices by  $A(\alpha)$ .

From now on we fix a set  $A \subset V \times \{0, 1\}^E$  with  $\pi_{\text{full},p}(A) \geq \frac{1}{2}$  and set  $B = A(\frac{1}{4})$ . For every  $b \in B$ , let  $\widehat{\text{Env}}(b) \subseteq \text{Env}(b, A)$  be some set of environments such that  $\pi_p(\widehat{\text{Env}}(b)) \in [\frac{1}{4}, \frac{1}{2}]$ . Suppose that  $\eta_0 \sim \pi_p$ . Recall from Definition 3.3 that  $(\tau_i)$  is the sequence of regeneration times with  $\tau_0 = 0$ . Let  $\widehat{\tau}_i := \tau_{\inf\{\ell: \sum_{j=1}^{\ell} 1_{\{X_{\tau_j} \in B\}} = i\}}$  be the  $i$ th regeneration time at which the walk coordinate is in  $B$ . We now take a subsequence defined as follows:  $\sigma_1 = \widehat{\tau}_1$  and inductively

$$\sigma_{i+1} = \inf\{\widehat{\tau}_j : \widehat{\tau}_j \geq \sigma_i + \kappa\},$$

where  $\kappa$  is a constant to be determined later. Finally, we let

$$T = \inf\{j : \eta_{\sigma_j} \in \widehat{\text{Env}}(X_{\sigma_j})\}.$$

LEMMA 6.2. Let  $\eta_0 \sim \pi_p$ . Then,  $\eta_{\sigma_1}$  is independent of  $X_{\sigma_1}$  and distributed according to  $\pi_p$ .

PROOF. If an edge has not been examined by the walk during  $[0, \sigma_1]$ , then at time  $\sigma_1$  it is distributed according to  $\text{Ber}(p)$ . For the edges that were examined by the walk, considering the last time before time  $\sigma_1$  that this happened, we get that at time  $\sigma_1$  they are also distributed according to  $\text{Ber}(p)$ , independently, over different edges and of the location of the walk.  $\square$

LEMMA 6.3. If  $\kappa$  in the definition of  $(\sigma_i)$  is taken sufficiently large, then

$$\forall i \in \mathbb{N}, \quad \mathbb{P}(T = i \mid T \geq i) \geq \frac{1}{8}.$$

Let  $Y$  be the auxiliary chain with parameters  $\mu = 1$  and  $p \in (0, 1)$ . We take its continuous time version, that is, we consider the continuous time chain which stays at every vertex for an exponential time of parameter 1 and then makes a jump according to  $P_{\text{aux}}$ . We write  $(P_t^{\text{aux},\text{cts}})_{t \in \mathbb{R}_+}$  for its transition semigroup, that is,  $P_s^{\text{aux},\text{cts}} = e^{s(P_{\text{aux}} - I)}$ , and define for  $\delta > 0$

$$(6.1) \quad r(\delta) = \inf\left\{s \in \mathbb{R}_+ : \max_{x,y \in V} P_t^{\text{aux},\text{cts}}(x, y) \leq 1 - \delta \text{ for all } t \geq s\right\}.$$

LEMMA 6.4. There exists  $\delta_0 \in (0, 1)$  such that for all  $p \in (0, 1)$  we have

$$(6.2) \quad r(\delta_0) \leq \frac{1}{p}.$$

PROOF. We write  $J_t = P_t^{\text{aux},\text{cts}}$ . We first note that  $\max_{x,y \in V} J_t(x, y)$  is nonincreasing in  $t$ , and so  $r(\delta) = \inf\{s \in \mathbb{R}_+ : \max_{x,y \in V} J_s(x, y) \leq 1 - \delta\}$ . Now, let  $s = 1/p$ , and let  $T_1$  be the first time the continuous time auxiliary chain jumps out of  $x$ . Then,  $T_1$  is exponential of parameter  $1 - P_{\text{aux}}(x, x) \asymp p$  by Lemma 3.8. We then get

$$J_s(x, x) \geq \mathbb{P}(T_1 > s) = e^{-s(1 - P_{\text{aux}}(x,x))} \geq c_1.$$

Let  $T_2$  be an exponential random variable independent of  $T_1$  of parameter  $\max_y(1 - P_{\text{aux}}(y, y)) \asymp p$ . We then get

$$1 - J_s(x, x) \geq \mathbb{P}_x(T_1 < s, T_2 > s) \geq c_2$$

for a positive constant  $c_2$ . Noting that  $J_s(x, y) \leq 1 - J_s(x, x)$  for  $x \neq y$  concludes the proof.  $\square$

Below we write  $\Lambda_{\text{full},(\mu,p)}$ ,  $\Lambda_{\text{aux},(\mu,p)}$  and  $\Lambda_{\text{SRW}}$  for the spectral profile of the full chain, the auxiliary chain with parameters  $(\mu, p)$  and the simple random walk, respectively. Note that  $\Lambda_{\text{aux},(\mu,p)}$  is defined with respect to the generator  $P_{\text{aux},(\mu,p)} - I$ . We write  $\mathcal{L}_{\mu,p}$  for the generator of the full process with parameters  $(\mu, p)$ . Finally, for  $A \subset V \times \{0, 1\}^E$  we write  $\lambda_{(\mu,p)}(A)$  for the smallest eigenvalue of the restriction of  $-\mathcal{L}_{\mu,p}$  to  $A$ , as in the paragraph preceding (5.4).

**PROPOSITION 6.5.** *There exist positive constants  $M$  and  $L$  so that the following holds. Let  $A \subset V \times \{0, 1\}$  be such that  $\pi_{\text{full},p}(A^c) \leq 1/2$ , let  $B = A(\frac{1}{4}) \subseteq V$  be as above, let  $\delta_0$  be as in Lemma 6.4 and let  $(Y_s)_{s \geq 0}$  be the continuous time chain with generator  $P_{\text{aux},(1,p)} - I$  with  $p \in (0, 1)$ . There exists a sequence of stopping times  $T_1 < T_2 < \dots$  (w.r.t. the chain  $(Y_s)_{s \geq 0}$ ) such that for all  $i$  we have  $Y_{T_i} \in B$  and  $T_i / \lceil \frac{M}{\Lambda_{\text{aux},(1,p)}(M\pi_{\text{full},p}(A^c))} + r(\delta_0) \rceil$  is stochastically dominated by the law of  $\sum_{j=1}^i Z_j$ , where  $Z_1, Z_2, \dots$  are i.i.d. Geometric random variables with mean at most  $L$ , where  $r(\delta_0)$  is as in (6.1).*

We defer the proofs of Lemma 6.3 and Proposition 6.5 until after the proof of Theorem 1.2.

**PROOF OF THEOREM 1.2.** To simplify notation we write  $r = r(\delta_0)$ .

The proof is mostly a formal exercise involving translating the assertion of Proposition 6.5 concerning the rate of exponential decay of the tail of  $T_i / (i \lceil \frac{M}{\Lambda_{\text{aux},(1,p)}(M\pi_{\text{full},p}(A^c))} + r \rceil)$  into one about the rate of exponential decay of the tail of  $\sigma_i / (i \lceil \frac{M}{\Lambda_{\text{aux},(1,p)}(M\pi_{\text{full},p}(A^c))} + r \rceil)$ . This is straightforward in light of the fact that the spacings between the regeneration times are i.i.d. with an exponentially decaying tail and that each such spacing is at least  $\kappa$  w.p. at least  $e^{-\kappa}$ . We now give the formal details.

Let  $t = \frac{1}{2\mu} \log(|E|(1 - \alpha)/(\alpha \log 2))$ , where  $\alpha = p \wedge (1 - p)$ . Then, it is not hard to see that

$$(6.3) \quad \max_{x,\eta} \|\mathbb{P}_{(x,\eta)}^t - \pi_{\text{full},p}\|_{2,\pi_{\text{full},p}}^2 \leq \frac{2}{\min_{v \in V} \pi(v)}.$$

Indeed, let  $P_t = P_t^{(\mu,p)}$  and  $Q_t = Q_t^{(\mu,p)}$  be the transition kernels for time  $t$  of the full process and of the environment, respectively. The latter is simply a continuous-time SRW on the  $p$ -tilted hypercube, and its  $1$ - $L_2$  mixing time (i.e.,  $t_{\text{mix}}^{(2)}(\varepsilon)$  with  $\varepsilon = 1$ ) is at most  $t$  by Lemma 5.5.

Thus,  $\frac{P_{2t}((x,\eta),(x,\eta))}{\pi_{\text{full},p}(x,\eta)} \leq \frac{Q_{2t}(\eta,\eta)}{\pi_p(\eta) \min_{v \in V} \pi(v)} \leq 2 / \min_{v \in V} \pi(v)$  which shows (6.3).

Using (6.3) and (5.11), we get

$$(6.4) \quad t_{\text{mix}}^{\text{full},(\mu,p),(\infty)}(\varepsilon^2) \leq 2t + 2 \int_{2 \min_{v \in V} \pi(v)}^{4/\varepsilon} \frac{d\delta}{\delta \Lambda_{\text{full},(\mu,p)}(\delta)}.$$

Let  $\mu_1 < \mu_2$ . Then,  $\mathcal{L}_{\mu_2,p}((x, \eta), (y, \eta')) \leq \frac{\mu_2}{\mu_1} \mathcal{L}_{\mu_1,p}((x, \eta), (y, \eta'))$  for all  $(x, \eta), (y, \eta') \in V \times \{0, 1\}^E$ . Using also (5.5), (5.1) and noting that both  $\mathcal{L}_{\mu_1,p}$  and  $\mathcal{L}_{\mu_2,p}$  are reversible w.r.t.  $\pi_{\text{full},p}$ , we obtain

$$(6.5) \quad \Lambda_{\text{full},(\mu_2,p)}(\delta) \leq \frac{\mu_2}{\mu_1} \Lambda_{\text{full},(\mu_1,p)}(\delta).$$



By combining (6.4) and (6.5) together with (5.3) and (5.6), we see that in order to conclude the proof of Theorem 1.2 it suffices to consider  $\mu = 1$  and prove that for a positive constant  $M$  we have that for all  $A \subset V \times \{0, 1\}^E$  with  $\pi_{\text{full},p}(A^c) \leq 1/2$

$$(6.6) \quad \frac{p}{M\lambda_{(1,p)}(A^c)} \leq \frac{1}{\Lambda_{\text{SRW}}(M\pi_{\text{full},p}(A^c))}$$

(the l.h.s. is defined w.r.t. the full process). We write  $\Lambda^{\text{aux},(\mu,p)}$  for the spectral profile of the auxiliary chain with parameters  $\mu$  and  $p$ .

By Lemma 4.4 we have that, for all  $\delta > 0$ ,

$$\frac{p}{\Lambda^{\text{aux},(1,p)}(\delta)} \leq \frac{2}{\Lambda_{\text{SRW}}(\delta)}.$$

Hence, to conclude the proof of Theorem 1.2 it suffices to show that for  $A$  as above

$$(6.7) \quad \frac{1}{M\lambda_{(1,p)}(A^c)} \leq \frac{1}{\Lambda^{\text{aux},(1,p)}(M\pi_{\text{full},p}(A^c))}.$$

Throughout the remainder of the section, we fix  $\mu = 1$ . Our strategy for proving (6.7) is to find a set  $B \subseteq V$  with  $\pi(B^c) \lesssim \pi_{\text{full},p}(A^c)$  and such that the asymptotic rate of decay of the tail of  $T_A$  can be controlled via the time spent in  $B$  by the auxiliary chain. Roughly speaking, we want to have a set  $B$  such that every visit of the auxiliary chain to  $B$  is a visit of the full process to  $A$  with some probability bounded away from 0. Here, we are using the fact that the auxiliary chain can be coupled with the full process by viewing it along regeneration times.

We next claim that it suffices to prove that for  $M$  and  $r$ , as in Proposition 6.5, there exists a positive constant  $c$  such that

$$(6.8) \quad \mathbb{E}_{\pi_{\text{full},p}}[e^{32\alpha\sigma_i}] \leq e^i \quad \text{for some } \alpha \text{ s.t. } \frac{c}{\alpha} \leq \frac{1}{\Lambda_{\text{aux},(1,p)}(M\pi_{\text{full},p}(A^c))} + r.$$

Indeed, for  $(\sigma_i)$  and  $T$  as defined earlier

$$\mathbb{P}_{\pi_{\text{full},p}}(T_A > t) \leq \mathbb{P}_{\pi_{\text{full},p}}(\sigma_{\lceil 16\alpha t \rceil} > t) + \mathbb{P}_{\pi_{\text{full},p}}(T > \lceil 16\alpha t \rceil).$$

Now, by Lemma 6.3 we have  $\mathbb{P}_{\pi_{\text{full},p}}(T > \lceil 16\alpha t \rceil) \leq (7/8)^{\lceil 16\alpha t \rceil} \leq e^{-2\alpha t}$  and  $\mathbb{P}_{\pi_{\text{full},p}}(\sigma_{\lceil 16\alpha t \rceil} > t) \leq \mathbb{E}_{\pi_{\text{full},p}}[e^{32\alpha\sigma_{\lceil 16\alpha t \rceil}}]e^{-32\alpha t}$  which, by our assumption, is at most  $e^{-32\alpha t} \times e^{\lceil 16\alpha t \rceil} \leq e^{1-16\alpha t}$ . Therefore, this would then imply that  $\mathbb{E}_{\pi_{\text{full},p}}[e^{\alpha T_A}] < \infty$  and, hence,  $\lambda_{(1,p)}(A^c) \geq \alpha$ .

So we now turn to prove (6.8). Recall that  $\tau_i$  is the  $i$ th regeneration time as in Definition 3.3. We now consider the discrete time auxiliary chain  $Y$ . Let  $\tilde{Y}$  be the continuous time version of  $Y$ , that is, we let  $N$  be an independent Poisson process of rate 1 and set  $\tilde{Y}_t = Y_{N(t)}$ . We set

$$S = 8 \left[ \frac{M}{\Lambda_{\text{aux},(1,p)}(M\pi_{\text{full},p}(A^c))} + r \right].$$

Hence, it suffices to show that, by setting  $L$  to be sufficiently large, there exists a positive constant  $c$  such that for all  $x \geq 1$  and all  $i$

$$\mathbb{P}(\sigma_i > xL^3Si) \leq ce^{-10xi}.$$

Let  $(T_i)$  be the stopping times from Proposition 6.5. Recall that  $N(t) \sim \text{Poisson}(t)$  is the number of jumps the continuous-time version of the auxiliary chain makes by time  $t$ . Hence,  $N(T_i)$  is the number of jumps it makes by the stopping time  $T_i$ . We have generated the full process, the auxiliary chain  $(Y_j)$  and its continuous-time version  $\tilde{Y}_t := Y_{N(t)}$  on the same probability space (the auxiliary chain is generated by viewing the walk coordinate of the

full process at regeneration times, and the continuous-time version of the auxiliary chain is generated from the auxiliary chain by using an independent rate one Poisson process  $(N(t) : t \geq 0)$ . Hence, we can consider  $\rho(i) := \tau_{N(T_i)}$  which is the time at which the  $N(T_i)$ -th regeneration time of the full process occurs. The walk coordinate of the full process at time  $\rho(i)$  is  $\tilde{Y}_{T_i}$ . For all  $i$  we set  $Z_i$  to be the time between  $\rho(i)$  and the first time after  $\rho(i)$  that the walk  $X$  examines an edge. (Note  $Z_i$  is an exponential variable of parameter 1.) We also define

$$\xi_j = \mathbf{1}(N(T_j) - N(T_{j-1}) \geq 1, Z_j \geq \kappa) \quad \text{and} \quad J_\ell = \sum_{j=1}^{\ell} \xi_j.$$

Note that the variables  $(\xi_j)$  are i.i.d., and, since  $T_j - T_{j-1} \geq 1$  for all  $j$  and the two events appearing in the definition of  $\xi_j$  are independent, we get that  $\mathbb{P}(\xi_j = 1) \geq (1 - e^{-1})e^{-\kappa}$ . Using the definitions above, we then have the following inclusions for all  $x$  and  $i$ :

$$\begin{aligned} \{\sigma_i \geq xL^3 Si\} &\subseteq \{\rho(\lceil xLi/4 \rceil) \geq xL^3 Si\} \cup \{J_{\lceil xLi/4 \rceil} < i\}, \\ \{\rho(\lceil xLi/4 \rceil) \geq xL^3 Si\} &\subseteq \{N(T_{\lceil xLi/4 \rceil}) \geq xL^2 Si\} \cup \{\tau_{\lceil xL^2 Si \rceil} \geq xL^3 Si\}, \\ \{N(T_{\lceil xLi/4 \rceil}) \geq xL^2 Si\} &\subseteq \{T_{\lceil xLi/4 \rceil} \geq xL^2 Si/100\} \cup \{N(xL^2 Si/100) \geq xL^2 Si\}. \end{aligned}$$

(For instance, the first inclusion follows by noting that if  $J(j) \geq \ell$  and  $\rho(j) \leq t$ , then we must have that  $\sigma_\ell \leq t$ . Indeed, if  $J(j) \geq \ell$ , then, among  $T_1, \dots, T_j$ , there are at least  $\ell$  that contribute  $+1$  toward increasing the index of  $\sigma$ , that is, toward  $\inf\{k : \sigma_k \geq \rho(j)\}$ . If moreover  $\rho(j) \leq t$ , then  $T_j$  is generated in the full process at some time which is smaller or equal to  $t$ .)

The proof is now concluded by taking  $L$  sufficiently large and using Proposition 6.5 for the tails of the stopping times  $T_i$ , the fact that  $J$  is the sum of i.i.d. indicators with probability bounded away from 0, large deviations for Poisson random variables and Lemma 3.5 for the tails of  $(\tau_i - \tau_{i-1})$ .

Using that  $t_{\text{spectral-profile}}^{\text{SRW}}(\varepsilon) \gtrsim (\log|V|/\varepsilon)$  for all  $\varepsilon \in (0, 1)$  shows that we can absorb the logarithmic terms (one of which is coming from the term  $\int_{2 \min_{v \in V} \pi(v)}^{4/\varepsilon} \frac{r}{\delta} d\delta$ , using  $\min_v \pi(v) \geq |V|^{-2}$ , while the other one from (6.4)) and this completes the proof.  $\square$

REMARK 6.6. We now explain how to get rid of the log term in the statement of Theorem 1.2 when considering total variation mixing. Let  $t = \frac{2 \log n}{\mu}$  and  $A$  be the event that all edges of  $G$  have been updated by time  $t$ . Define  $\nu_1 = \mathbb{P}_{(x,\eta)}^t(\cdot | A)$  and  $\nu_2 = \mathbb{P}_{(x,\eta)}^t(\cdot | A^c)$ . Then, we have

$$\mathbb{P}_{(x,\eta)}^{t+s} = \mathbb{P}(A)\mathbb{P}_{\nu_1}^s + \mathbb{P}(A^c)\mathbb{P}_{\nu_2}^s,$$

and, hence, by convexity and Jensen’s inequality we obtain

$$\|\mathbb{P}_{(x,\eta)}^{t+s} - \pi_{\text{full},p}\|_{\text{TV}} \leq \mathbb{P}(A^c) + \|\mathbb{P}_{\nu_1}^s - \pi_{\text{full},p}\|_{2,\pi_{\text{full},p}}.$$

Similar to (6.3), we have that

$$\|\nu_1 - \pi_{\text{full},p}\|_{2,\pi_{\text{full},p}}^2 \leq \frac{1}{\min_v \pi(v)}.$$

The rest of the proof is identical to the proof of Theorem 1.2.

LEMMA 6.7. Let  $\pi$  be a distribution of full support on a finite set  $\Omega$ . Let  $J$  be an  $\Omega$ -valued random variable and  $A$  an event. Then,

$$(6.9) \quad \|\mathbb{P}(J \in \cdot | A) - \pi\|_{2,\pi}^2 \leq \frac{\|\mathbb{P}(J \in \cdot) - \pi\|_{2,\pi}^2 + 1}{\mathbb{P}(A)^2} - 1.$$

In particular, if  $\nu$  is some distribution on  $\Omega$  and  $\widehat{\nu}$  is  $\nu$  conditioned on  $A \subseteq \Omega$ , then

$$(6.10) \quad \|\widehat{\nu} - \pi\|_{2,\pi}^2 \leq \frac{\|\nu - \pi\|_{2,\pi}^2 + 1}{\nu(A)^2} - 1.$$

PROOF. We have that  $\|\mathbb{P}(J \in \cdot) - \pi\|_{2,\pi}^2 + 1 = \sum_x \pi(x) \left(\frac{\mathbb{P}(J=x)}{\pi(x)}\right)^2 \geq \sum_x \pi(x) \times \left(\frac{\mathbb{P}(\{J=x\} \cap A)}{\pi(x)}\right)^2$ . By the same reasoning  $\|\mathbb{P}(J \in \cdot | A) - \pi\|_{2,\pi}^2 + 1 = \frac{1}{\mathbb{P}(A)^2} \sum_{x \in A} \pi(x) \times \left(\frac{\mathbb{P}(\{J=x\} \cap A)}{\pi(x)}\right)^2$ .  $\square$

PROOF OF LEMMA 6.3. We need to show that

$$\mathbb{P}(\eta_{\sigma_i} \in \widehat{\text{Env}}(X_{\sigma_i}) | T \geq i) \geq \frac{1}{8}.$$

Let  $J_i$  be the set of edges examined by the walk during the time interval  $[\sigma_{i-1}, \sigma_i]$ , and let  $E_i = E \setminus J_i$ . Crucially, given  $(J_i, E_i)$ ,  $\eta_{\sigma_{i-1}}$  and the whole history  $\sigma(X_t : t \leq \sigma_i)$ , we have that the law of  $\eta_{\sigma_i}$  can be described as follows:

- (1) The different coordinates of  $\eta_{\sigma_i}$  are independent;
- (2) For  $e \in J_i$  we have that  $\eta_{\sigma_i}(e) \sim \text{Bernoulli}(p)$ ;
- (3) For  $e \in E_i$  we have that  $\eta_{\sigma_i}(e) \sim \text{Bernoulli}(p)$  with probability  $1 - e^{-(\sigma_i - \sigma_{i-1})}$ , and, otherwise,  $\eta_{\sigma_i}(e) = \eta_{\sigma_{i-1}}(e)$ . In other words, the restriction of the environment to  $E_i$  evolves during  $[\sigma_{i-1}, \sigma_i]$  by updating each edge at rate 1 to be either open w.p.  $p$  or closed w.p.  $1 - p$ .

For any sequence  $(x_i)$  with  $x_i \in B$  for all  $i$  and all sequences  $(S_i)$  of subsets of  $E$ , define  $A_i = \{X_{\sigma_1} = x_1, \dots, X_{\sigma_i} = x_i, E_1 = S_1, \dots, E_i = S_i\}$ . For every subset of edges  $S$ , we denote by  $\pi_p^S$  the Bernoulli( $p$ ) product measure on  $S$ . By properties (1) and (2) above we then have

$$\|\mathbb{P}(\eta_{\sigma_i} \in \cdot | T \geq i, A_i) - \pi_p\|_{2,\pi_p} = \|\mathbb{P}(\eta_{\sigma_i}|_{S_i} \in \cdot | T \geq i, A_i) - \pi_p^{S_i}\|_{2,\pi_p^{S_i}},$$

where  $\eta|_S$  denotes the restriction of  $\eta$  on the set of edges  $S$ . For every  $\eta \in \{0, 1\}^{S_i}$ , we now let

$$\nu_i^\eta(\cdot) = \mathbb{P}(\sigma_i - \sigma_{i-1} \in \cdot | \eta_{\sigma_{i-1}}|_{S_i} = \eta, T \geq i, A_i).$$

We note that under the conditioning above, the distribution of  $\sigma_i - \sigma_{i-1}$  does not depend on  $\{\eta_{\sigma_{i-1}}|_{S_i} = \eta\}$ , since  $S_i$  is the set of edges that the walk does not examine during  $[\sigma_{i-1}, \sigma_i]$ . Therefore, we get for all  $\eta$

$$\nu_i^\eta(\cdot) = \mathbb{P}(\sigma_i - \sigma_{i-1} \in \cdot | T \geq i, A_i) =: \nu_i(\cdot).$$

Using again that  $S_i$  is the set of edges that the walk does not examine during  $[\sigma_{i-1}, \sigma_i]$ , we get

$$(6.11) \quad \mu_i(\eta) := \mathbb{P}(\eta_{\sigma_{i-1}}|_{S_i} = \eta | T \geq i, A_i) = \mathbb{P}(\eta_{\sigma_{i-1}}|_{S_i} = \eta | T \geq i, A_{i-1}).$$

Conditional on  $\{T \geq i\} \cap A_i$ , the distribution of  $\eta_{\sigma_i}|_{S_i}$  is that of a  $p$ -tilted random walk on the hypercube  $\{0, 1\}^{S_i}$  started from  $\eta_{\sigma_{i-1}}|_{S_i}$  and run for time  $\sigma_i - \sigma_{i-1}$ . Let  $\tilde{\eta}$  be a continuous time  $p$ -tilted random walk on  $\{0, 1\}^{S_i}$ . Hence, putting all things together (and recalling that

$\sigma_i - \sigma_{i-1} \geq \kappa$  by construction) we obtain

$$\begin{aligned}
 & \|\mathbb{P}(\eta_{\sigma_i} | S_i \in \cdot | T \geq i, A_i) - \pi_p^{S_i}\|_{2, \pi_p^{S_i}}^2 \\
 &= \sum_{\xi \in \{0,1\}^{S_i}} \pi_p^{S_i}(\xi) \left( \frac{\mathbb{P}(\eta_{\sigma_i} | S_i = \xi | T \geq i, A_i)}{\pi_p^{S_i}(\xi)} - 1 \right)^2 \\
 &= \sum_{\xi \in \{0,1\}^{S_i}} \pi_p^{S_i}(\xi) \left( \frac{1}{\pi_p^{S_i}(\xi)} \sum_{\eta \in \{0,1\}^{S_i}} \mu_i(\eta) \int_{\kappa}^{\infty} \mathbb{P}(\tilde{\eta}_t = \xi | \tilde{\eta}_0 = \eta) v_i(dt) - 1 \right)^2 \\
 (6.12) \quad &= \sum_{\xi \in \{0,1\}^{S_i}} \pi_p^{S_i}(\xi) \left( \int_{\kappa}^{\infty} \frac{\mathbb{P}_{\mu_i}(\tilde{\eta}_t = \xi)}{\pi_p^{S_i}(\xi)} v_i(dt) - 1 \right)^2 \\
 &= \sum_{\xi \in \{0,1\}^{S_i}} \pi_p^{S_i}(\xi) \left( \int_{\kappa}^{\infty} \left( \frac{\mathbb{P}_{\mu_i}(\tilde{\eta}_t = \xi)}{\pi_p^{S_i}(\xi)} - 1 \right) v_i(dt) \right)^2 \\
 &\leq \int_{\kappa}^{\infty} \sum_{\xi \in \{0,1\}^{S_i}} \pi_p^{S_i}(\xi) \left( \frac{\mathbb{P}_{\mu_i}(\tilde{\eta}_t = \xi)}{\pi_p^{S_i}(\xi)} - 1 \right)^2 v_i(dt),
 \end{aligned}$$

where we used Jensen’s inequality for the last bound. The spectral gap of the chain  $\tilde{\eta}$  is 1, and, hence, using Poincaré’s inequality yields for all  $t$

$$\begin{aligned}
 \sum_{\xi \in \{0,1\}^{S_i}} \pi_p^{S_i}(\xi) \left( \frac{\mathbb{P}_{\mu_i}(\tilde{\eta}_t = \xi)}{\pi_p^{S_i}(\xi)} - 1 \right)^2 &= \|\mathbb{P}_{\mu_i}(\tilde{\eta}_t = \cdot) - \pi_p^{S_i}\|_{2, \pi_p^{S_i}}^2 \\
 &\leq e^{-2t} \|\mu_i - \pi_p^{S_i}\|_{2, \pi_p^{S_i}}^2.
 \end{aligned}$$

Plugging this into (6.12) gives

$$(6.13) \quad \|\mathbb{P}(\eta_{\sigma_i} | S_i \in \cdot | T \geq i, A_i) - \pi_p^{S_i}\|_{2, \pi_p^{S_i}} \leq e^{-\kappa} \|\mu_i - \pi_p^{S_i}\|_{2, \pi_p^{S_i}}.$$

Using (6.11) and the fact that the  $L_2$  distance does not increase under projections, we get

$$\|\mu_i - \pi_p^{S_i}\|_{2, \pi_p^{S_i}} \leq \|\mathbb{P}(\eta_{\sigma_{i-1}} \in \cdot | T \geq i, A_{i-1}) - \pi_p\|_{2, \pi_p}.$$

Lemma 6.7 (with  $(\eta_{\sigma_{i-1}}, \mathbb{P}$  given  $T \geq i - 1$  and  $A_{i-1}, \{T \geq i\} \cap A_{i-1}$ ) here playing the roles of  $(J, \mathbb{P}, A)$  from (6.9), resp.) gives that

$$\begin{aligned}
 & \|\mathbb{P}(\eta_{\sigma_{i-1}} \in \cdot | T \geq i, A_{i-1}) - \pi_p\|_{2, \pi_p}^2 \\
 (6.14) \quad & \leq \frac{1}{(\mathbb{P}(T \neq i - 1 | T \geq i - 1, A_{i-1}))^2} \|\mathbb{P}(\eta_{\sigma_{i-1}} \in \cdot | T \geq i - 1, A_{i-1}) - \pi_p\|_{2, \pi_p}^2 \\
 & \quad + \frac{1}{(\mathbb{P}(T \neq i - 1 | T \geq i - 1, A_{i-1}))^2} - 1.
 \end{aligned}$$

Setting  $\theta_i(\cdot) = \mathbb{P}(\eta_{\sigma_i} \in \cdot | T \geq i, A_i)$ , so far we have shown that

$$(6.15) \quad e^{2\kappa} \|\theta_i - \pi_p\|_{2, \pi_p}^2 \leq \frac{\|\theta_{i-1} - \pi_p\|_{2, \pi_p}^2}{(\mathbb{P}(T \neq i - 1 | T \geq i - 1, A_{i-1}))^2}$$

$$(6.16) \quad + \frac{1}{(\mathbb{P}(T \neq i - 1 | T \geq i - 1, A_{i-1}))^2} - 1.$$

We next show that for all  $i$

$$(6.17) \quad \|\theta_i - \pi_p\|_{2,\pi_p} \leq \frac{1}{8}.$$

Since  $\eta_0 \sim \pi_p$ , Lemma 6.2 gives that, conditional on  $X_{\sigma_1}$ , we have that  $\eta_{\sigma_1} \sim \pi_p$ . Therefore, we get that (6.17) is true for  $i = 1$ . Suppose it holds for  $i - 1$ , we show that it also holds for  $i$ . By the definition of  $T$ , we have for all  $i$

$$\begin{aligned} \mathbb{P}(T = i \mid T \geq i, A_i) &= \mathbb{P}(\eta_{\sigma_i} \in \widehat{\text{Env}}(x_i) \mid T \geq i, A_i) \\ &= \sum_{\eta \in \widehat{\text{Env}}(x_i)} \mathbb{P}(\eta_{\sigma_i} |_{S_i} = \eta |_{S_i}, \eta_{\sigma_i} |_{S_i^c} = \eta |_{S_i^c} \mid T \geq i, A_i) \\ &= \sum_{\eta \in \widehat{\text{Env}}(x_i)} \pi_p^{S_i^c}(\eta |_{S_i^c}) \cdot \mathbb{P}(\eta_{\sigma_i} |_{S_i} = \eta |_{S_i} \mid T \geq i, A_i), \end{aligned}$$

where in the last equality we used the i.i.d. property of the different coordinates of  $\eta_{\sigma_i}$  and the fact that they are  $\text{Ber}(p)$  (properties (1) and (2) from the beginning of the proof). Writing  $\lambda_i$  for the projection map from  $\{0, 1\}^E \rightarrow \{0, 1\}^{S_i}$  and noting that

$$\pi_p(\widehat{\text{Env}}(x_i)) = \sum_{\eta \in \widehat{\text{Env}}(x_i)} \pi_p^{S_i^c}(\eta |_{S_i^c}) \cdot \pi_p^{S_i}(\eta |_{S_i})$$

we get

$$\begin{aligned} &|\mathbb{P}(T = i \mid T \geq i, A_i) - \pi_p(\widehat{\text{Env}}(x_i))| \\ &\leq \sum_{\eta \in \lambda_i(\widehat{\text{Env}}(x_i))} \pi_p^{S_i}(\eta) \left| \frac{\mathbb{P}(\eta_{\sigma_i} |_{S_i} = \eta \mid T \geq i, A_i)}{\pi_p^{S_i}(\eta)} - 1 \right| \\ &\leq \|\mathbb{P}(\eta_{\sigma_i} |_{S_i} \in \cdot \mid T \geq i, A_i) - \pi_p^{S_i}\|_{2,\pi_p^{S_i}} \\ &\leq \|\mathbb{P}(\eta_{\sigma_i} \in \cdot \mid T \geq i, A_i) - \pi_p\|_{2,\pi_p}, \end{aligned}$$

where for the second inequality we used Cauchy–Schwarz and for the third one the fact that the  $L_2$  distance does not increase under projections. From this it now follows that if

$$\|\mathbb{P}(\eta_{\sigma_{i-1}} \in \cdot \mid T \geq i - 1, A_{i-1}) - \pi_p\|_{2,\pi_p} \leq \frac{1}{8},$$

then

$$|\mathbb{P}(T = i - 1 \mid T \geq i - 1, A_{i-1}) - \pi_p(\widehat{\text{Env}}(x_{i-1}))| \leq \frac{1}{8}.$$

Since  $\pi_p(\widehat{\text{Env}}(x_{i-1})) \leq 1/2$ , the above implies that

$$\mathbb{P}(T = i - 1 \mid T \geq i - 1, A_{i-1}) \leq \frac{5}{8}.$$

We are now ready to show that if (6.17) holds for  $i - 1$ , then it also holds for  $i$ . Indeed, substituting the above bound into (6.15) and using the induction hypothesis  $\|\theta_{i-1} - \pi_p\|_{2,\pi_p} \leq \frac{1}{8}$  give

$$\|\theta_i - \pi_p\|_{2,\pi_p} \leq e^{-\kappa} \left( \frac{(1/8)^2}{(3/8)^2} + \frac{1}{(3/8)^2} - 1 \right)^{1/2} \leq \frac{\sqrt{56}}{3e^\kappa},$$

which by taking  $\kappa$  sufficiently large can be made smaller than  $1/8$  and this completes the inductive step and the proof of the lemma.  $\square$

PROOF OF PROPOSITION 6.5. First, we claim that for all  $\alpha$  we have that  $(1 - \alpha)\pi(A(\alpha)^c) \leq \pi_{\text{full},p}(A^c)$ . Indeed,  $\pi_{\text{full},p}(A) \leq \pi(A(\alpha)) + \alpha\pi(A(\alpha)^c)$ , and, hence,  $\pi_{\text{full},p}(A^c) \geq (1 - \alpha)\pi(A(\alpha)^c)$ . Since  $B = A(1/4)$ , this now gives that  $\pi(B^c) \leq \frac{4}{3} \times \pi_{\text{full},p}(A^c) \leq \frac{2}{3}$ . By increasing  $M$  by a  $\frac{4}{3}$ -factor, we may replace  $\pi_{\text{full},p}(A^c)$  in (6.7) as well as in the statement of Proposition 6.5 by  $\pi(B^c)$ .

We write  $r = r(\delta_0)$  with  $\delta_0$  from Lemma 6.4, and we let  $\alpha = \min((\delta_0/2), 1/4)$ . To simplify notation, we write  $\Lambda = \Lambda_{\text{aux},(1,p)}$ . Let  $\kappa \in \mathbb{N}$  to be chosen later. We now define a sequence of times  $(t_j)$  by setting for all  $j \geq 0$

$$t_j = j \left( \frac{\log(4\alpha^{-2\kappa})}{\Lambda(16\alpha^{-2\kappa}/\|\pi_{B^c} - \pi\|_{2,\pi}^2)} \vee r \right) = j \left( \frac{\log(4\alpha^{-2\kappa})}{\Lambda(16\alpha^{-2\kappa}\pi(B^c)/\pi(B))} \vee r \right),$$

where we write  $a \vee b$  for  $\max(a, b)$  and  $\pi_D$  for  $\pi$  conditioned on  $D$  (i.e.,  $\pi_D(x) = \frac{\pi(x)\mathbf{1}_{x \in D}}{\pi(D)}$ ). We will now construct a sequence of random sets  $D_0, D_1, \dots$  such that  $B^c \subseteq D_i$  for all  $i$  and if  $\xi_i = \mathbf{1}(Y_{t_i} \notin D_i)$ , then for all  $i$  almost surely

$$\mathbb{P}(\xi_i = 1 \mid \xi_0, \dots, \xi_{i-1}) \geq \alpha.$$

This will imply the assertion of the proposition by setting

$$T_i = \inf \left\{ t_j : \sum_{k=0}^j \xi_k = i \right\},$$

that is,  $T_i$  is the  $i$ th time  $t_j$  such that  $Y_{t_j} \notin D_j$ .

So now we turn to define the sets  $D_i$ . We do this by induction. For  $i = 0$ , we set  $D_0 = B^c$ . For  $i \geq 1$ , we will define  $D_i$  as a measurable function of  $\xi_0, \dots, \xi_{i-1}$ . Since  $\pi(B^c) \leq 2/3$  and  $Y_0 \sim \pi$ , we immediately get that

$$\mathbb{P}(T_1 = 0) = \mathbb{P}(\xi_0 = 1) = \pi(B) \geq \frac{1}{3}.$$

We note that if  $Y_0 \sim \pi$ , then, given  $Y_0 \notin B$ , we have that  $Y_0 \sim \pi_{B^c}$  and, similarly, if  $Y_0 \in B$ , then  $Y_0 \sim \pi_B$ . We now consider the measures  $\nu_i(\cdot) = \mathbb{P}_{\pi_{B^i}}(Y_{t_1} \in \cdot)$  for  $i = 0, 1$ , where we set  $B^0 = B$  and  $B^1 = B^c$ . We now argue that

$$(6.18) \quad \|\nu_0 - \pi\|_{2,\pi}^2 \vee \|\nu_1 - \pi\|_{2,\pi}^2 \leq \alpha^\kappa \frac{\pi(B)}{\pi(B^c)}.$$

Indeed, using (5.12) we obtain

$$(6.19) \quad \|\nu_1 - \pi\|_{2,\pi}^2 \leq \alpha^\kappa \|\pi_{B^c} - \pi\|_{2,\pi}^2 = \alpha^\kappa \frac{\pi(B)}{\pi(B^c)}.$$

We now verify that also  $\|\nu_0 - \pi\|_{2,\pi}^2 \leq \alpha^\kappa \frac{\pi(B)}{\pi(B^c)}$ . Clearly,  $\|\nu_0 - \pi\|_{2,\pi}^2 \leq \|\pi_B - \pi\|_{2,\pi}^2 = \frac{\pi(B^c)}{\pi(B)}$ . Hence, it suffices to consider the case that  $\frac{\pi(B^c)}{\pi(B)} \geq \alpha^\kappa \frac{\pi(B)}{\pi(B^c)}$ . In this case, using the fact that  $\Lambda(\cdot)$  is nonincreasing  $\Lambda(16\alpha^{-2\kappa}\pi(B^c)/\pi(B)) \leq \Lambda(4(4\alpha^{-\kappa})/\|\pi_B - \pi\|_{2,\pi}^2)$  and so by (5.12), we get that

$$(6.20) \quad \|\nu_0 - \pi\|_{2,\pi}^2 \leq \frac{1}{4}\alpha^\kappa \|\pi_B - \pi\|_{2,\pi}^2 = \frac{1}{4}\alpha^\kappa \frac{\pi(B^c)}{\pi(B)} \leq \alpha^\kappa \frac{\pi(B)}{\pi(B^c)},$$

as desired, where in the last inequality we have used the fact that  $\pi(B^c) \leq 2/3$ .

From (6.18) together with Proposition 5.4 we get that  $\nu_i(B^c) \leq \pi(B^c) + \alpha^{\kappa/2}$  for  $i = 0, 1$ . Fix  $i \in \{0, 1\}$ . Let  $D_i^j$  be a set in

$$\{D \supseteq B^c : \nu_i(D) \geq \delta_0/2\}$$

with minimal  $\nu_i$  probability. By the definition of  $r$  and the fact that  $\pi(B^c) \leq 2/3$ , we have that for all  $i = 0, 1$  and for  $\kappa$  sufficiently large

$$\nu_i(D_1^i) \in \left[ \frac{\delta_0}{2}, (\pi(B^c) + \alpha^{\kappa/2}) \vee \left( \frac{\delta_0}{2} + 1 - \delta_0 \right) \right] \subseteq \left[ \frac{\delta_0}{2} \wedge \frac{1}{4}, \left( 1 - \frac{\delta_0}{2} \right) \vee \frac{3}{4} \right].$$

To see this, consider the cases  $\nu_i(B^c) \leq \delta_0/2$  and  $\nu_i(B^c) \geq \delta_0/2$ . In the latter we may take  $D_1^i = B^c$ , while in the former, by minimality  $\nu_i(D_1^i) \leq \delta_0/2 + \max_z \nu_i(z)$  which is at most  $1 - \delta_0/2$  as  $t_1 \geq r$  (using the definition of  $r$ ). We set  $D_1 = D_1^0$  of  $Y_0 \in B^0$  and  $D_1 = D_1^1$  of  $Y_0 \in B^1$ .

This concludes the construction of  $D_1$ . We now proceed by induction. For  $a \in \{0, 1\}^i$  and  $j \leq i$ , let  $a(j) \in \{0, 1\}^j$  be the first  $j$  coordinates of  $a$ . Assume that for each  $a \in \{0, 1\}^i$  we have defined sets  $D_j^{a(j)} \supseteq B^c$  for all  $j \leq i$  such that  $\nu_a$ , the law of  $Y_i$  given that, for each  $j < i$ , we have that  $Y_{t_j} \in D_j^{a(j)}$  iff the  $j$ th coordinate of  $a$  is 1, satisfies that

$$\nu_a(D_i^a) \in [\alpha, 1 - \alpha] \quad \text{and} \quad \|\nu_a - \pi\|_{2,\pi}^2 \leq \alpha^\kappa \left( \frac{\pi(B)}{\pi(B^c)} + \frac{\alpha^{-2}}{1 - \alpha^{\kappa-2}} \right).$$

Note that we have already checked that this holds for our distributions  $\nu_i$  for  $i = 0, 1$ . We now want to construct for each  $b \in \{0, 1\}^{i+1}$  a distribution  $\nu_b$  and the set  $D_{i+1}^b$  such that  $\nu_b(D_{i+1}^b) \in [\alpha, 1 - \alpha]$  and  $\|\nu_b - \pi\|_{2,\pi}^2 \leq \alpha^\kappa \left( \frac{\pi(B)}{\pi(B^c)} + \frac{\alpha^{-2}}{1 - \alpha^{\kappa-2}} \right)$ . For  $a \in \{0, 1\}^i$ , let  $\nu_a^{(0)}$  (resp.,  $\nu_a^{(1)}$ ) be the measure  $\nu_a$  conditioned on  $(D_i^a)^c$  (resp.,  $D_i^a$ ). Then, by Lemma 6.7 for  $j = 0, 1$  we get

$$\begin{aligned} \|\nu_a^{(j)} - \pi\|_{2,\pi} &\leq \frac{\|\nu_a - \pi\|_{2,\pi}^2 + 1}{j\nu_a(D_i^a)^2 + (1-j)\nu_a((D_i^a)^c)^2} \leq \alpha^{-2}(\|\nu_a - \pi\|_{2,\pi}^2 + 1) \\ &\leq \alpha^{\kappa-2} \frac{\pi(B)}{\pi(B^c)} + \frac{\alpha^{-2}}{1 - \alpha^{\kappa-2}} =: M. \end{aligned}$$

Using  $\pi(B^c) \leq 2/3$ , we see that  $M \leq \frac{\pi(B)}{\pi(B^c)}(\alpha^{\kappa-2} + 2\frac{\alpha^{-2}}{1 - \alpha^{\kappa-2}}) \leq 3\frac{\pi(B)}{\pi(B^c)}\alpha^{-2}$ , provided  $\kappa \geq \kappa_0(\alpha)$ . Using this bound, we see that

$$t_1 \geq \frac{\log(\alpha^{-2\kappa+2})}{\Lambda(4\alpha^{-2\kappa+2}/M)}.$$

For  $b \in \{0, 1\}^{i+1}$ , we now define  $\nu_b(\cdot) = \mathbb{P}_{\nu_{b(i)}^{(b_{i+1})}}(Y_{t_1} \in \cdot)$  with  $b_{i+1}$  denoting the  $i + 1$ st coordinate of  $b$  (and  $b(i)$  its first  $i$  coordinates). Using (5.13), this time we obtain that for sufficiently large  $\kappa$

$$\|\nu_b - \pi\|_{2,\pi}^2 \leq \alpha^{2\kappa-2} M \leq \alpha^\kappa \left( \frac{\pi(B)}{\pi(B^c)} + \frac{\alpha^{-2}}{1 - \alpha^{\kappa-2}} \right).$$

In particular,  $\|\nu_b - \pi\|_{2,\pi}^2 \leq 3\alpha^{\kappa-2} \frac{\pi(B)}{\pi(B^c)}$  provided  $\kappa$  is sufficiently large. Proposition 5.4 gives that  $\nu_b(B^c) \leq \pi(B^c) + \sqrt{3}\alpha^{(\kappa-2)/2} \leq \pi(B^c) + \alpha^{\kappa/4}$  for  $\kappa$  sufficiently large. In the same way as above when defining the set  $D_1^i$ , we get that if  $\kappa$  is sufficiently large, there exists some set  $D_{i+1}^b \supseteq B^c$  such that  $\nu_b(D_{i+1}^b) \in [\alpha, 1 - \alpha]$ . This completes the induction and the proof of the proposition.  $\square$

**7. Applications: Transitive graphs of moderate growth and the hypercube.** In this section we prove Theorems 1.8 and 1.9. Below we identify a percolation cluster with the vertices lying in it. Recall that we denote the cluster of vertex  $x$  by  $K_x$  and its edge boundary by  $\partial K_x$ . Finally, recall that we write  $M_p = \pi_p(|\partial K_x||K_x|^2)$  and  $N_p = \pi_p(|K_x|)$ .

LEMMA 7.1. *Let  $c, a \in \mathbb{R}$ . There exists a positive constant  $c_1 = c_1(a, c)$  such that the following holds. Let  $G = (V, E)$  be a connected vertex-transitive graph of  $(c, a)$ -moderate growth and diameter  $\gamma$ . Suppose that  $N_p \leq \gamma/4$ . Then, we have*

$$t_{\text{rel}}^{\text{full},(\mu,p)} \geq c_1 \frac{(\gamma - 4N_p)^2}{\mu p M_p}.$$

PROOF. Denote the cluster of  $x \in V$  w.r.t.  $\eta$  by  $K_x(\eta)$ , and, as usual, we identify it with the set of vertices lying in it.

Fix some  $o \in V$ . Let  $f(x, \eta) = \frac{1}{|K_x(\eta)|} \sum_{v \in K_x(\eta)} d_G(v, o)$ , where  $d_G$  is the graph distance w.r.t.  $G$ . For  $\eta \in \{0, 1\}^E$ , denote by  $\eta^e$  the environment obtained from  $\eta$  by setting  $e$  to be open if it was not already open (i.e.,  $\eta^e(e) = 1$  and  $\eta^e(e') = \eta(e')$  for all  $e' \in E \setminus \{e\}$ ). Observe that the value of  $f$  cannot change as a result of a jump in the random walk coordinate (as such a jump leaves the walk in the same percolation cluster). Thus,

$$\begin{aligned} \mathcal{E}(f, f) &= \sum_{x \in V, \eta \in \{0,1\}^E, e \in \partial K_x(\eta)} \pi(x) \pi_p(\eta) \mu p (f(x, \eta) - f(x, \eta^e))^2 \\ &\leq \mu p \sum_{x, \eta, e \in \partial K_x(\eta)} \pi(x) \pi_p(\eta) |K_x(\eta^e)|^2 \\ (7.1) \quad &= \mu p \sum_{\eta, e \in \partial K_o(\eta)} \pi_p(\eta) |K_o(\eta^e)|^2 \\ &= \mu p \sum_{x, y: \{x, y\} \in E} \pi_p[\mathbf{1}(y \notin K_o \ni x) (|K_o| + |K_y|)^2], \end{aligned}$$

where the penultimate equality follows from transitivity and the last one from the fact that on  $\{y \notin K_o \ni x\}$  we have that  $|K_o^{\{x,y\}}| = |K_o| + |K_y|$ .

Observe that for all  $A \subset V$  with  $o, x \in A$  and  $y \notin A$ , given that  $K_o = A$ , we have that  $|K_y|$  is distributed as the size of the percolation cluster of  $y$  (with parameter  $p$ ) on the induced graph on  $V \setminus A$  which by an obvious coupling with  $\pi_p$  is stochastically dominated by the (unconditional) law of  $|K_y|$  under  $\pi_p$ . Thus, for all such  $A$  we have that  $\mathbb{E}\pi_p[|K_y|^a \mid K_o = A] \leq \mathbb{E}\pi_p[|K_y|^a]$ , and, hence, for all  $a > 0$  we have that

$$(7.2) \quad \mathbb{E}\pi_p[\mathbf{1}(y \notin K_o \ni x) |K_y|^a \mid K_o] \leq \mathbb{E}\pi_p[|K_y|^a] = \mathbb{E}\pi_p[|K_o|^a].$$

Plugging (7.2) in (7.1) and summing over  $\{x, y\} \in E$  yield

$$(7.3) \quad \mathcal{E}(f, f) \leq 4\mu p M_p.$$

We conclude the proof by showing that  $\text{Var}_{\pi_{\text{full},p}}(f) \geq c_{a,c}(\gamma - 4N_p)^2$ . It is not hard to verify that

$$(7.4) \quad \forall x, \quad \left| d_G(x, o) - \sum_{\eta} \pi_p(\eta) f(x, \eta) \right| \leq \pi_p(|K_x|) = N_p.$$

Let  $(X, \eta) \sim \pi \times \pi_p$  and  $(Y, \eta') \sim \pi \times \pi_p$  be independent. Define

$$A = \{d_G(Y, o) \leq \gamma/4, d_G(X, o) \geq 3\gamma/4\}.$$

Since  $G$  is of moderate growth, there exists a positive constant  $b$  such that  $\mathbb{P}(A) \geq b$ . Finally, by the independence between  $X$  and  $\eta$  together with (7.4)  $\mathbb{E}[f(X, \eta) \mid A] \geq 3\gamma/4 - N_p$  and, similarly, we also have that  $\mathbb{E}[f(Y, \eta) \mid A] \leq \gamma/4 + N_p$ , which together yield that

$$\begin{aligned} 2 \text{Var}_{\pi_{\text{full},p}}(f) &= \mathbb{E}[(f(X, \eta) - f(Y, \eta'))^2] \geq \mathbb{E}[(f(X, \eta) - f(Y, \eta'))^2 \mathbf{1}(A)] \\ &\geq (\mathbb{P}(A))^2 (\mathbb{E}[f(X, \eta) \mid A] - \mathbb{E}[f(Y, \eta') \mid A])^2 \geq b^2(\gamma - 4N_p)^2/4, \end{aligned}$$



where for the second inequality we used Jensen’s inequality and for the last one we used the assumption that  $\gamma \geq 4N_p$ .  $\square$

**PROOF OF THEOREM 1.8.** Let  $P$  be the transition matrix of simple random walk (SRW) on  $G$ . Diaconis and Saloff-Coste [7] showed that for a Cayley graph  $G$  of  $(c, a)$ -moderate growth we have

$$c^2\gamma^24^{-2a-1} \leq t_{\text{rel}} \lesssim t_{\text{mix}}^{(\infty)} \lesssim_{c,a} \gamma^2.$$

Using this, Lemma 7.1 and the assumptions on  $M_p$  and  $N_p$  we have

$$t_{\text{rel}}^{\text{full},(\mu,p)} \gtrsim_{a,b,c} \frac{1}{\mu p} \gamma^2 \asymp_{a,c} \frac{1}{\mu p} t_{\text{rel}}^{\text{SRW}}.$$

As proved in [19], Proposition 8.1, and previously noted in [24], for vertex-transitive graphs of degree  $d$  and  $(c, a)$ -moderate growth one has that  $t_{\text{spectral-profile}}^{\text{SRW}} \lesssim_{a,c,d} \gamma^2$ . This together with Theorem 1.2 yields

$$t_{\text{rel}}^{\text{full},(\mu,p)} \lesssim t_{\text{mix}}^{\text{full},(\mu,p),(\infty)} \lesssim_{a,c,d} \frac{1}{\mu p} \gamma^2 + \frac{1}{\mu} |\log(1-p)| \lesssim \frac{1}{\mu p} \gamma^2 \asymp \frac{1}{\mu p} t_{\text{rel}}^{\text{SRW}},$$

where for the last inequality we used the assumption  $|\log(1-p)| \leq \gamma^2$ . This completes the proof.  $\square$

**PROOF OF THEOREM 1.9.** The proof follows from Theorem 1.2 together with the upper bound on the spectral profile of simple random walk on the hypercube (see [19], Section 7).  $\square$

**8. Log-Sobolev constant.** Theorem 1.2 in [18] asserts that for every Markov chain on a finite state space

$$(8.1) \quad \frac{1}{17} \leq \inf_{\varepsilon \in (0,1/2]} \frac{\log(1/\varepsilon)c_{\text{LS}}}{\Lambda_0(\varepsilon)} \leq 1.$$

We note that the result in [18] is stated for the case that  $\mathcal{L} = P - I$  for some transition matrix  $P$ , but, as noted several times before, the general case can be reduced to the case  $\mathcal{L} = c(P - I)$  for some  $c > 0$ , and the relevant quantities scale linearly in  $c$ . We also note that the equivalence between (8.1) and Theorem 1.2 in [18] relies on the general fact that if  $A = \bigcup_{i=1}^r A_i$  and the sets  $A_i$  are disjoint and satisfy  $\mathcal{L}(x, y) = 0$  for all  $x \in A_i$  and  $y \in A_j$  for  $i \neq j$ , then (e.g., by (5.5))

$$(8.2) \quad \lambda(A) = \min_{i \in [r]} \lambda(A_i).$$

As for every singleton  $x$  by (5.5), we have that  $\lambda(\{x\}) = -\mathcal{L}(x, x)$  it follows that

$$(8.3) \quad c_{\text{LS}} \leq \min_x \frac{-\mathcal{L}(x, x)}{\log(1/\pi(x))}.$$

**PROOF OF THEOREM 1.6.** Let  $M$  be as in (6.6). Recall that  $\Lambda_0(\varepsilon)$  and  $\Lambda(\varepsilon)$  are non-increasing in  $\varepsilon$ . Thus, by increasing  $M$  if necessary, we may assume that  $M \geq 2$ . Using this again, as well as  $\Lambda_0(1/2) \leq \Lambda(1) \leq 2\Lambda_0(1/2)$  (this is used to treat  $\varepsilon \in (\pi_*/M, \pi_*] \cup (\frac{1}{2M}, \frac{1}{2}]$ ; for the first inequality see the proof of [13], Lemma 2.2, the second inequality follows from (5.3) by monotonicity), by (6.6) together with (8.1) we have that

$$\mu p \min_{\varepsilon \in [\pi_*/M, 1/2]} \frac{\log(1/\varepsilon)}{\Lambda_0^{\text{full},(\mu,p)}(\varepsilon)} \lesssim \min_{\varepsilon \in [\pi_*, 1/2]} \frac{\log(1/\varepsilon)}{\Lambda_0^{\text{SRW}}(\varepsilon)} \asymp 1/c_{\text{LS}}^{\text{SRW}}.$$

It is left to prove  $\min_{\varepsilon \in (0, \pi_*/M]} \frac{\log(1/\varepsilon)}{\Lambda_0^{\text{full}, (\mu, p)}(\varepsilon)} \lesssim \frac{\log(1/\pi_*) \log(\frac{1}{p(1-p)})}{\mu}$ .

Let  $\delta \leq 1/M$ , and let  $B$  be such that  $\pi_{\text{full}, p}(B) = \delta\pi_*$ . Then, for all  $v \in V$  we have that

$$\pi_p(\text{Env}(v, B)) \leq \frac{\sum_u \pi(u) \pi_p(\text{Env}(u, B))}{\pi_*} = \frac{\pi_{\text{full}, p}(B)}{\pi_*} = \delta.$$

Consider now the (reducible) Markov chain on  $\Omega := V \times \{0, 1\}^E$  in which the walk coordinate cannot change, and the environment evolves in the usual fashion by refreshing each edge at rate  $\mu$  (and declaring it to be open with probability  $p$  and closed w.p.  $1 - p$ ). Note that it is reversible w.r.t.  $\pi_{\text{full}, p}$ . Denote the corresponding generator by  $\mathcal{L}'$  while that of the (usual) full process with parameters  $(\mu, p)$  by  $\mathcal{L}$ . Let  $\lambda'(B)$  and  $\lambda(B)$  be the minimal Dirichlet eigenvalue of  $B$  w.r.t.  $\mathcal{L}'$  (i.e., the minimal positive eigenvalue of  $\mathcal{L}'_B$ , given by  $\mathcal{L}'_B(a, b) := \mathcal{L}'(a, b)\mathbf{1}(a, b \in B)$ ) and  $\mathcal{L}$ , respectively. Then, by (5.5) we have that

$$(8.4) \quad \begin{aligned} \lambda'(B) &= \min\{\mathcal{E}_{\mathcal{L}'}(h, h) : h \in \mathbb{R}_+^\Omega, \|h\|_2 = 1, \text{supp}(h) \subseteq B\} \\ &\leq \min\{\mathcal{E}_{\mathcal{L}}(h, h) : h \in \mathbb{R}_+^\Omega, \|h\|_2 = 1, \text{supp}(h) \subseteq B\} = \lambda(B). \end{aligned}$$

(The first equality holds even without irreducibility.) Since w.r.t.  $\mathcal{L}'$  the sets  $(\{v\} \times \text{Env}(v, B) : v \in V)$  are disconnected (i.e.,  $\mathcal{L}'(a, b) = 0$  for all  $a \in \{v\} \times \text{Env}(v, B)$  and  $b \in \{u\} \times \text{Env}(v, B)$ , for all  $v \neq u$ ) and their union is  $B$ , by (8.2) we have that

$$(8.5) \quad \lambda'(B) = \min_v \lambda'(\{v\} \times \text{Env}(v, B)).$$

As the evolution of the chain corresponding to  $\mathcal{L}'$  on each set of the form  $\{v\} \times \{0, 1\}^E$  is simply that of the rate  $\mu$   $p$ -tilted hypercube, for every  $v$  we have that

$$(8.6) \quad \lambda'(\{v\} \times \text{Env}(v, B)) = \lambda_{p\text{-tilted hypercube with rate } \mu}(\text{Env}(v, B)).$$

As  $\pi_p(\text{Env}(v, B)) \leq \delta$ , by (8.1) we have that for all  $v$  (uniformly in  $p$  and  $\mu$ )

$$(8.7) \quad \lambda_{p\text{-tilted hypercube with rate } \mu}(\text{Env}(v, B)) \gtrsim \hat{c} \log(1/\delta),$$

where  $\hat{c} = \hat{c}(\mu, p)$  is the log-Sobolev constant of the rate  $\mu$   $p$ -tilted hypercube (this is a dimension free quantity [8], but here the coordinates of the hypercube are labeled by the set  $E$ ). Finally, combining (8.4)–(8.7) and using the facts that  $\pi_{\text{full}, p}(B) = \delta\pi_*$  and  $\delta \leq 1/M \leq 1/2$ , we see that

$$\begin{aligned} \frac{\lambda(B)}{\log(1/\pi_{\text{full}, p}(B))} &\gtrsim \frac{\hat{c} \log(1/\delta)}{\log(1/(\delta\pi_*))} \gtrsim \hat{c} / \log(1/\pi_*) \\ &\gtrsim \frac{\mu(1 - 2 \min(p, 1 - p))}{\log(\max(p, 1 - p) / \min(p, 1 - p)) \log(\frac{1}{\pi_*})}, \end{aligned}$$

where in the last inequality we have used

$$\hat{c} \asymp \mu(1 - 2 \min(p, 1 - p)) / \log\left(\frac{\max(p, 1 - p)}{\min(p, 1 - p)}\right)$$

(e.g., [8], alternatively, this can be seen using the facts that: (i) the log-Sobolev constant scales linearly in  $\mu$ , (ii) for product chains it is the same as the minimal log-Sobolev constant of a single coordinate [8], and (iii) the log-Sobolev constant of a single coordinate can be approximated using (8.1)). We note that for  $p = 1/2$  this should be interpreted as  $(1 - 2 \min(p, 1 - p)) / \log(\frac{\max(p, 1 - p)}{\min(p, 1 - p)}) = 1$ .

We now prove the result about  $t_{\text{rel}}$ . Combining (6.5) and (6.6), we have that

$$(\mu p)^{-1} \Lambda_{(\mu, p)}(1/2) \gtrsim \min_a \Lambda_{\text{SRW}}(a) = (t_{\text{rel}}^{\text{SRW}})^{-1}.$$

Finally, using the fact that for every reversible chain  $\Lambda(1/2) \leq 2/t_{\text{rel}}$  [13], Lemma 2.2, completes the proof.  $\square$

**Acknowledgments.** We are grateful to Tom Hutchcroft for several useful suggestions. This work was supported by the Engineering and Physical Sciences Research Council: JH by EP/L018896/1 and PS by EP/R022615/1.

## REFERENCES

- [1] ALDOUS, D. and FILL, J. (2002). Reversible Markov chains and random walks on graphs. Unfinished manuscript. Available at <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- [2] ALDOUS, D. J. (1982). Some inequalities for reversible Markov chains. *J. Lond. Math. Soc.* (2) **25** 564–576. MR0657512 <https://doi.org/10.1112/jlms/s2-25.3.564>
- [3] AVENA, L., GÜLDAŞ, H., VAN DER HOFSTAD, R. and DEN HOLLANDER, F. (2018). Mixing times of random walks on dynamic configuration models. *Ann. Appl. Probab.* **28** 1977–2002. MR3843821 <https://doi.org/10.1214/17-AAP1289>
- [4] BASU, R., HERMON, J. and PERES, Y. (2017). Characterization of cutoff for reversible Markov chains. *Ann. Probab.* **45** 1448–1487. MR3650406 <https://doi.org/10.1214/16-AOP1090>
- [5] BISKUP, M. and RODRIGUEZ, P.-F. (2018). Limit theory for random walks in degenerate time-dependent random environments. *J. Funct. Anal.* **274** 985–1046. MR3743189 <https://doi.org/10.1016/j.jfa.2017.12.002>
- [6] BREUILLARD, E. and TOINTON, M. C. H. (2016). Nilprogressions and groups with moderate growth. *Adv. Math.* **289** 1008–1055. MR3439705 <https://doi.org/10.1016/j.aim.2015.11.025>
- [7] DIACONIS, P. and SALOFF-COSTE, L. (1994). Moderate growth and random walk on finite groups. *Geom. Funct. Anal.* **4** 1–36. MR1254308 <https://doi.org/10.1007/BF01898359>
- [8] DIACONIS, P. and SALOFF-COSTE, L. (1996). Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** 695–750. MR1410112 <https://doi.org/10.1214/aoap/1034968224>
- [9] DING, J., LEE, J. R. and PERES, Y. (2012). Cover times, blanket times, and majorizing measures. *Ann. of Math.* (2) **175** 1409–1471. MR2912708 <https://doi.org/10.4007/annals.2012.175.3.8>
- [10] DING, J. and PERES, Y. (2013). Sensitivity of mixing times. *Electron. Commun. Probab.* **18** no. 88. MR3141797 <https://doi.org/10.1214/ECP.v18-2765>
- [11] DOYLE, P. G. and STEINER, J. (2011). Commuting time geometry of ergodic markov chains. Preprint. Available at [arXiv:1107.2612](https://arxiv.org/abs/1107.2612).
- [12] GAUDILLIÈRE, A. and LANDIM, C. (2014). A Dirichlet principle for non reversible Markov chains and some recurrence theorems. *Probab. Theory Related Fields* **158** 55–89. MR3152780 <https://doi.org/10.1007/s00440-012-0477-5>
- [13] GOEL, S., MONTENEGRO, R. and TETALI, P. (2006). Mixing time bounds via the spectral profile. *Electron. J. Probab.* **11** 1–26. MR2199053 <https://doi.org/10.1214/EJP.v11-300>
- [14] HÄGGSTRÖM, O., PERES, Y. and STEIF, J. E. (1997). Dynamical percolation. *Ann. Inst. Henri Poincaré Probab. Stat.* **33** 497–528. MR1465800 [https://doi.org/10.1016/S0246-0203\(97\)80103-3](https://doi.org/10.1016/S0246-0203(97)80103-3)
- [15] HERMON, J. (2018). On sensitivity of uniform mixing times. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 234–248. MR3765888 <https://doi.org/10.1214/16-AIHP802>
- [16] HERMON, J. (2018). A spectral characterization for concentration of the cover time. *J. Theoret. Probab.* To appear. Preprint. Available at [arXiv:1809.00145](https://arxiv.org/abs/1809.00145).
- [17] HERMON, J. and PERES, Y. (2018). A characterization of  $L_2$  mixing and hypercontractivity via hitting times and maximal inequalities. *Probab. Theory Related Fields* **170** 769–800. MR3773799 <https://doi.org/10.1007/s00440-017-0769-x>
- [18] HERMON, J. and PERES, Y. (2018). On sensitivity of mixing times and cutoff. *Electron. J. Probab.* **23** Paper No. 25. MR3779818 <https://doi.org/10.1214/18-EJP154>
- [19] HERMON, J. and PYMAR, R. (2018). The exclusion process mixes (almost) faster than independent particles. Preprint. Available at [arXiv:1808.10846](https://arxiv.org/abs/1808.10846).
- [20] HERMON, J. and SOUSI, P. (2019). A comparison principle for random walk on dynamical percolation. Preprint. Available at [arXiv:1902.02770](https://arxiv.org/abs/1902.02770).
- [21] KOZMA, G. (2007). On the precision of the spectral profile. *ALEA Lat. Am. J. Probab. Math. Stat.* **3** 321–329. MR2372888
- [22] LEVIN, D. A. and PERES, Y. (2017). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. Second edition of [MR2466937]. MR3726904
- [23] LOVÁSZ, L. and WINKLER, P. (1998). Mixing times. In *Microsurveys in Discrete Probability* (Princeton, NJ, 1997). DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **41** 85–133. Amer. Math. Soc., Providence, RI. MR1630411
- [24] LYONS, R. and OVEIS GHARAN, S. (2018). Sharp bounds on random walk eigenvalues via spectral embedding. *Int. Math. Res. Not. IMRN* **2018** 7555–7605. MR3892273 <https://doi.org/10.1093/imrn/rnx082>

- [25] MORRIS, B. and PERES, Y. (2005). Evolving sets, mixing and heat kernel bounds. *Probab. Theory Related Fields* **133** 245–266. MR2198701 <https://doi.org/10.1007/s00440-005-0434-7>
- [26] PERES, Y., SOUSI, P. and STEIF, J. E. (2019). Quenched exit times for random walk on dynamical percolation. *Markov Process. Related Fields*. To appear.
- [27] PERES, Y., SOUSI, P. and STEIF, J. E. (2020). Mixing time for random walk on supercritical dynamical percolation. *Probab. Theory Related Fields* **176** 809–849. MR4087484 <https://doi.org/10.1007/s00440-019-00927-z>
- [28] PERES, Y., STAUFFER, A. and STEIF, J. E. (2015). Random walks on dynamical percolation: Mixing times, mean squared displacement and hitting times. *Probab. Theory Related Fields* **162** 487–530. MR3383336 <https://doi.org/10.1007/s00440-014-0578-4>
- [29] PERES, Y. and STEIF, J. E. Private communication.
- [30] SOUSI, P. and THOMAS, S. (2018). Cutoff for random walk on dynamical Erdős Rényi graph. Preprint. Available at [arXiv:1807.04719](https://arxiv.org/abs/1807.04719).
- [31] TESSERA, R. and TOINTON, M. (2019). A finitary structure theorem for vertex-transitive graphs of polynomial growth. Preprint. Available at [arXiv:1908.06044](https://arxiv.org/abs/1908.06044).
- [32] VAN DER HOFSTAD, R. and NACHMIAS, A. (2017). Hypercube percolation. *J. Eur. Math. Soc. (JEMS)* **19** 725–814. MR3612867 <https://doi.org/10.4171/JEMS/679>