Rapid social connectivity [*]

Itai Benjamini[†]   Jonathan Hermon[‡]

Abstract

Given a graph $G = (V, E)$, consider Poisson($|V|$) walkers performing independent lazy simple random walks on G simultaneously, where the initial position of each walker is chosen independently with probability proportional to the degrees. When two walkers visit the same vertex at the same time they are declared to be acquainted. The social connectivity time $\mathrm{SC}(G)$ is defined as the first time in which there is a path of acquaintances between every pair of walkers. It is shown that when the average degree of G is d, with high probability

$$c \log |V| \leqslant \mathrm{SC}(G) \leqslant Cd^{1+5 \cdot 1_{G \text{ is not regular}}} \log^3 |V|.$$

When G is regular the lower bound is improved to $\mathrm{SC}(G) \geqslant \log |V| - 6 \log \log |V|$, with high probability. We determine $\mathrm{SC}(G)$ up to a constant factor in the cases that G is an expander and when it is the n-cycle.

Keywords: social network; random walks; giant component; coalescence process.
AMS MSC 2010: 82C41; 60K35; 82B43; 60J10; 05C81; 91D30.
Submitted to EJP on May 28, 2018, final version accepted on March 17, 2019.
Supersedes arXiv:arXiv:1608.07621.

1 Introduction

Consider the following model for a social network which we call the **random walks social network model** or for short, the *social network model*. We write **SN** as a shorthand for social network. Let $G = (V, E)$ be a finite connected graph, which is the underlying graph of the SN model. In the model there are walkers performing independent lazy simple random walks (**LSRW**) on G, defined as follows. If a walker's current position is v, then the walker either stays put w.p. $1/2$, or moves to one of the neighbors of v w.p. $\frac{1}{2d_v}$, where d_v is the degree of v. The walkers perform their LSRWs

simultaneously (i.e. at each time unit they all perform one step, which may be a lazy step).

At time 0 the number of walkers at vertex v is set to be N_v, where $(N_v)_{v \in V}$ are independent random variables, where for each vertex v, we have that $N_v \sim \mathrm{Poisson}(\bar{\pi}_v)$ where $\bar{\pi} := |V|\pi$ and $\pi_v := d_v/(2|E|)$ is the stationary distribution of the random walk on G.

A directed graph G is called *Eulerian* if for every vertex v the indegree of v equals its outdegree, denoted (unambiguously) by d_v. One can define LSRW on $G = (V, E)$ in an analogous manner to the undirected case (e.g., [6]). Its stationary distribution is $\pi_v := d_v/|E|$ (here E is the set of directed edges). We say that a directed graph $G = (V, E)$ is *connected* if for all $u, v \in V$ there is a directed path in G from u to v.[1] We say that an Eulerian graph $G = (V, E)$ is regular when $d_v = d_u$ for all $u, v \in V$. We denote a directed edge from u to v by (u, v) and an undirected edge by $\{u, v\}$.

We shall prove our main results in the setup that G is a connected Eulerian digraph. Note that any graph can be thought of as an Eulerian digraph in which each edge appears with both orientations. However, the examples we consider shall all be graphs. Namely, the n-cycle and regular expanders.

We say that two walkers w, w' have *met by time t*, which we denote by $w \overset{t}{\leftrightarrow} w'$, if there exists $t_0 \leqslant t$ such that they have the same position at time t_0. After two walkers meet they continue their walks independently without coalescing. "Meeting by time t" is a symmetric relation. It induces the equivalence relation of *having a path of acquaintances by time t*, denoted by $\overset{t}{\sim}$, defined as follows. Two walkers a and b have a path of acquaintances by time t iff there exist $k \in \mathbb{N}$ and walkers $a = c_0, c_1, \ldots, c_k, c_{k+1} = b$ such that $c_i \overset{t}{\leftrightarrow} c_{i+1}$, for all $0 \leqslant i \leqslant k$. Note that we are not requiring the sequence of times in which the walkers met to be non-decreasing, which is the main difference between the SN model and some existing models for spread of rumor/infection (e.g. the $A + B \to 2B$ model [14] and the Frog model [18, 3, 4, 17, 12], see §2.6 for a description of these models and a discussion about their relationship to the SN model). Consequently, the SN model often evolves faster than such models. However, Conjecture 2.3 suggests one sense in which these models might also evolve rapidly.

We are interested in the coalescence process of the equivalence classes of $\overset{t}{\sim}$ (as t varies). In particular, we are interested in the **social connectivity time** $\mathrm{SC}(G)$, defined as the minimal time in which there is only one class (i.e. every pair of walkers have a path of acquaintances between them).

2 Our results

2.1 General graphs

Our main result is

Theorem 1. Let $G = (V, E)$ be a connected n-vertex Eulerian digraph of average degree $d = \frac{|E|}{n}$. Denote $r_* := \frac{\min_{u \in V} d_u}{d}$. Then there exist absolute constants[2] $C, M, c > 0$ such that

$$\mathrm{P}\left[\mathrm{SC}(G) > Cd^{1+5\xi}(\log n)^3\right] \leqslant M/n, \quad \text{where } \xi := 1_{G \text{ is not regular}}. \qquad (2.1)$$

$$\forall \alpha \in (0, 1), \exists c_\alpha > 0, \quad \mathrm{P}[\mathrm{SC}(G) \leqslant c_\alpha \log n] \leqslant 3e^{-cr_*^2 n^\alpha}. \qquad (2.2)$$

---

[1]This is sometimes referred to as being strongly connected.

[2]We use $C, C', C_0, C_1, K, M \ldots$ (resp. $\delta, \varepsilon, c, c', c_0, c_1, \ldots$) to denote positive absolute constants which are sufficiently large (resp. small) to ensure that a certain inequality holds. Similarly, we use C_d (resp. c_d) to refer to sufficiently large (resp. small) positive constants, whose value depends on the parameters appearing in subscript. Different appearances of the same constant at different places may refer to different numeric values.

Note that (2.2) is meaningful as long as $r_* \geqslant |V|^{\beta-0.5}$ for some $\beta > 0$. Theorem 5.5 gives an extension covering in particular the case $r_* \geqslant |V|^{\beta-1}$ for some $\beta > 0$. When G is a regular graph we have the following improved lower bounds on $\mathrm{SC}(G)$.

Theorem 2. There exist $c, C > 0$ so that for every finite connected regular graph $G = (V, E)$,

$$\mathrm{P}[\mathrm{SC}(G) < \log|V| - 6\log\log|V|] \leqslant C/\log|V|. \tag{2.3}$$

$$\mathrm{P}[\mathrm{SC}(G) < s_*(G)] \leqslant C\exp\left(-c\sqrt{|V|}(\log|V|)^{-4}\right), \tag{2.4}$$

where P is the transition matrix of the LSRW,

$$\kappa_t := \min_{v \in V} \sum_{i=0}^{t} P^{2i}(v, v) \quad \text{and} \quad s_*(G) := \min\left\{t : \frac{t+1}{\kappa_{t+1}} > \frac{1}{32}\log|V|\right\}.$$

Note that $s_*(G) > \frac{1}{32}\log|V|$. We include (2.3), despite this, as it is sharp, up to the value of the sub-leading term, and the failure probability, see the discussion below about extremality of the complete graph. In some cases $s_*(G) \geqslant C(\log|V|)^2$. For instance, for the n-cycle C_n we have that $P^{2i}(v, v) = \Theta(\frac{1}{\sqrt{i+1}})$ and so $s_*(\mathrm{C}_n) = \Theta((\log n)^2)$.[3] Hence by (2.4) $\mathrm{P}[\mathrm{SC}(\mathrm{C}_n) < c(\log n)^2] = o(1)$. This is not surprising as initially there will be intervals of length $\Omega(\log n)$ which will not be occupied by a single particle. This is why in terms of the lower bound on $\mathrm{SC}(\mathrm{C}_n)$ it is more interesting to study the case that initially there is one walker per site (as done in Theorem 3). Conversely, there exists some $C > 0$ such that the probability that $\mathrm{SC}(\mathrm{C}_n) > Cs_*(\mathrm{C}_n) \geqslant C'(\log n)^2$ is at most C''/n (see §2.3). We conjecture that there exists some universal constant $C > 0$ such that for every connected finite vertex transitive graph G (Definition 8.1) $\frac{1}{C} \leqslant \frac{\mathbb{E}[\mathrm{SC}(G)]}{s_*(G)+1} \leqslant C$.

2.2 Some remarks concerning our main results

Degree dependence: We now discuss the degree dependence in (2.1) and also a refinement of (2.1). Example 8.6 (resp. 8.7) describes for some absolute constants $c_0, c_1, C_0 > 0$ for all sufficiently large n and all $C_0 \leqslant d_n \leqslant c_0 n$ a family of graphs $G_n = (V_n, E_n)$ such that $n \leqslant |V_n| \leqslant C_0 n$ which are d_n-regular (resp. of average degree d_n) for which $\mathrm{P}[\mathrm{SC}(G_n) \geqslant c_1 d_n \log^2(|V_n|/d_n)] > c_1$ (resp. $\mathrm{P}[\mathrm{SC}(G_n) \geqslant c_1 d_n^2 \log^2(|V_n|/d_n)] > c_1$). This shows that the linear dependence on d in (2.1) is optimal for regular G and that we can have $\mathbb{E}[\mathrm{SC}(G)] \geqslant n$ for an n-vertex regular G and $\mathbb{E}[\mathrm{SC}(G)] \geqslant n^2$ for non-regular G.

If G is d-regular and in addition for every set $A \subset V$ such that $|A| \leqslant \frac{2}{3}(|V| + |V|^{2/3})$, we have that

$$|\{(u, v) \in E : u \in A, v \notin A \text{ or } u \notin A, v \in A\}| \geqslant c_0 d,$$

then the degree dependence in (2.1) can be eliminated and (2.1) can be improved to

$$\mathrm{P}[\mathrm{SC}(G) > Cc_0^{-2}\log^3|V|] \leqslant M/|V|,$$

for some absolute constants C, M (see Proposition 4.12, part (iii)).

Now consider the case that G is a connected n-vertex Eulerian digraph of average degree d. Let $\xi := 1_{G \text{ is not regular}}$. From the proof of Theorem 1 it follows that we can replace $Cd^{1+5\xi}(\log n)^3$ in (2.1) by $C't_*(G)d^{1+3\xi}\log n$, for some $t_*(G) \leqslant \bar{C}d^{2\xi}(\log n)^2$. Namely,

$$t_*(G) := \inf\{t : \max_{x,y}|P^t(x, y) - \pi_y| \leqslant d_y/(3 \cdot 2^9 d\log n)\},$$

---

[3]We write $o(1)$ for terms which vanish as $n \to \infty$ (or some other index, which is clear from context). We write $f_n = o(g_n)$ or $f_n \ll g_n$ if $f_n/g_n = o(1)$. We write $f_n = O(g_n)$ and $g_n = \Omega(f_n)$ if there exists a constant $C > 0$ such that $|f_n| \leqslant C|g_n|$ for all n. We write $f_n = \Theta(g_n)$ if $f_n = O(g_n)$ and $g_n = O(f_n)$. We write $f_n \leqslant O_d(g_n)$ to indicate that the constant C above may depend on d.

where P is the transition matrix of the LSRW on G. By Fact 4.8 indeed we have that $t_*(G) \leqslant \bar{C}d^{2\xi}(\log n)^2$.

For the cycle $t_* = \Theta((\log n)^2)$. This shows that (2.1) is sharp up to a factor $O(\log n)$, when $d = O(1)$. We believe that the cycle is in some sense extremal for the SN model (see Remark 8.6, Theorem 3 and Conjecture 8.2 for more details on this point).

The next remark discusses the case in which the particle density is considered as a parameter of the model.

Allowing the particle density to vary: One may consider the case in which the density of walkers at each site is multiplied by some $\lambda > 0$ (i.e. for all $v \in V$ the number of particles initially occupying it N_v has a Poisson$(\lambda \bar{\pi}_v)$ distribution). Let $\xi := 1_{G \text{ is not regular}}$. For $C_0 \frac{\log |V|}{|V|} d^\xi \leqslant \lambda \leqslant 1$ only minor adaptations to the arguments from §4 are necessary in order to show that

$$\text{w.h.p.} \quad \text{SC}(G) \leqslant C_d \lambda^{-4} (\log |V|)^3,$$

(here we allow λ to depend on $|V|$, provided that $\lambda \geqslant C \frac{\log |V|}{|V|} d^\xi$). With additional work this can be improved to w.h.p. $\text{SC}(G) \leqslant Cd^{1+5\xi} \lambda^{-2} (\log |V|)^3$. When $\lambda \in [|V|^{-1+\varepsilon}, 1]$ for some $\varepsilon \in (0, 1)$, this is optimal up to an $O_{d,\varepsilon}(\log |V|)$ factor as described in Remark 2.1.

When G is vertex-transitive (see Definition 8.1) we conjecture that

$$\mathbb{E}_\lambda[\text{SC}(G) \mid \text{there are at least two walkers}]$$

is monotone decreasing in λ (see Conjecture 8.3 for a slightly different formulation).

It is not hard to extend the proofs of (2.2) and (2.3) to the case that the particle density at each site is multiplied by some $\lambda \in [|V|^{-\varepsilon}, 1]$ for some $\varepsilon \in (0, 1/2)$. In this case, the lower bound on $\text{SC}(G)$ is increased by a factor of λ^{-1}. Similarly, one can extend (2.4) as follows: if $G = (V, E)$ is a connected n-vertex regular graph and $\lambda \in [n^{-\varepsilon}, 1]$ for some $\varepsilon \in (0, 1/16)$ then w.h.p. $\text{SC}(G) \geqslant s_*(G, \lambda) := \min\{t : \frac{t+1}{\kappa_{t+1}} > \frac{1}{32\lambda} \log n\}$.

Other initial distributions: We may consider the case that initially each vertex is occupied by a single particle. Consider the case that G is a d-regular Eulerian graph. We outline in §4.7 the necessary adaptations required in order to show that also in this setup we have

$$\text{w.h.p.}^4 \quad \text{SC}(G) \leqslant Cd(\log |V|)^3. \tag{2.5}$$

Extremality of the complete graph: We will prove Theorem 2 for an arbitrary holding probability $p \in [0, 1)$ for the walks (i.e. the probability of a walker staying put at each given step is p). Let K_n be the complete graph on n vertices. Consider the SN model on K_n with some holding probability $p_n \in [0, 1/n]$. It is not hard[5] to verify that $\text{SC}(K_n)$ is concentrated around $\log n$ in a window whose width is of order 1. Thus, by (2.3) the complete graph is the regular graph with (asymptotically) the fastest social connectivity time (at least when the holding probability is between 0 and $1/(d+1)$).

Extensions: In §8 we discuss some extensions of our results (e.g. allowing the walkers to belong to different communities, walking on different graphs with the same vertex set).

---

[4]We say that a sequence of events A_n defined with respect to some probabilistic model on a sequence of graphs $G_n := (V_n, E_n)$ with $|V_n| \to \infty$ holds w.h.p. ("with high probability") if the probability of A_n tends to 1 as $n \to \infty$.

[5]The proof is similar to the proof that the connectivity threshold for $G(n, p)$, the Erdős and Rényi random graph, occurs in a window of order $1/n$ around $p_n = (\log n)/n$ (for the SN model the proof is still elementary, but it is somewhat tedious and thus omitted).

2.3 The n-cycle and higher dimensional tori

Theorem 3. Let C_n be the n-cycle. Consider the setup in which at time 0 there is exactly one walker at each vertex and that the walkers perform continuous-time independent SRWs (with jump rate 1). Then there exist some absolute constants $c_1, c_2, \alpha, C_1, C_2 > 0$ such that

$$P[SC(C_n) > C_1 \log^2 n] \leqslant C_2/n. \tag{2.6}$$

$$P[SC(C_n) < c_1 \log^2 n] \leqslant e^{-c_2 n^\alpha}. \tag{2.7}$$

The proof of (2.6) is due to Gady Kozma. We thank him for allowing us to present his argument. Working in continuous-time simplifies the analysis of (2.6). The proof in the usual setup of $\mathrm{Pois}(1)$ walkers per site is in fact simpler (see Remark 6.1). As mentioned after Theorem 2, in the usual Poisson, discrete-time setup (2.7) follows from (2.4).

Remark 2.1. When initially the number of particles per site is $\mathrm{Pois}(\lambda)$ distributed (independently for different sites) it is possible to extend the analysis of Theorem 3 in order to show that w.h.p. $c < \frac{SC(C_n)}{\max\{[\lambda^{-1}\log(\lambda n)]^2, 1\}} \leqslant C$ as long as $\lambda \gg \frac{1}{n}$ (cf. [5] for the analysis of the frog model on the n-cycle for $\lambda \gg n^{-1}$).

Remark 2.2. Denote the d-dimensional torus of side length n by $\mathbb{T}_d(n)$. This is the Cayley graph of $(\mathbb{Z}/n\mathbb{Z})^d$ obtained by connecting by an edge each $x, y \in (\mathbb{Z}/n\mathbb{Z})^d$ which disagree only in one coordinate, by $\pm 1 \mod n$. Then $\mathbb{T}_1(n)$ is the n-cycle. We comment that using similar ideas as in [5, Theorem 2] one can show that w.h.p. $c < \frac{SC(\mathbb{T}_2(n))}{\log n \log \log n} < C$ and $c < \frac{SC(\mathbb{T}_d(n))}{d \log n} < C$ for $d \geqslant 3$. This result appeared in an earlier version of this work. Note that $SC(\mathbb{T}_2(n)) \geqslant c \log n \log \log n$ and $SC(\mathbb{T}_d(n)) \geqslant (1 - o(1))d \log n$ w.h.p. follow from (2.4)-(2.3).

2.4 Expanders

We denote by $\gamma(G)$ the **spectral-gap** of LSRW on G, defined as the smallest non-zero eigenvalue of $I - P$, where P is the transition matrix of LSRW on G and I is the identity matrix. We say that a sequence of graphs G_n is an expander family if $\inf_n \gamma(G_n) > 0$. We say that G is a γ-**expander** if $\gamma(G) \geqslant \gamma$ (we think of γ as being uniformly bounded away from 0, independently of the size of G, although our analysis remains valid even if this fails). The following result demonstrates that Theorem 2 is sharp.

Theorem 4. Let G be a connected d-regular n-vertex λ-expander. Then there exist absolute constants C, C', C_0 such that if $\lambda \log_d n \geqslant C_0$, then $SC(G) \leqslant C\lambda^{-1} \log n$ w.p. at least $1 - C'/n$.

Theorem 5. Let G be a connected d-regular n-vertex λ-expander. Then there exist some constants $c, C_1 > 0$ such that with probability at least $1 - \exp\left(-cnd^{-4(s+1)}\right)$, after $s := \lceil C_1/\lambda \rceil$ steps there exists a class of walkers of size at least $n/6$.

We refer to a class of walkers (in the SN model on some graph $G = (V, E)$ at a certain time) of size $\Omega(|V|)$ as a **giant class**. We shall deduce Theorem 4 from Theorem 5 by bounding SC by the time it takes the giant class from Theorem 5 to "capture" the rest of the walkers. The requirement $\lambda \log_d n \geqslant C_0$ in Theorem 4 is imposed in order to ensure that the error term $\exp\left(-cnd^{-4(s+1)}\right)$ from Theorem 5 is small.

2.5 Organization of this work

In §3 we present some preliminaries. In §4 we prove (2.1). In §5 we prove (2.2) and Theorem 2. In §6 we prove Theorem 3. In §7 we study the SN model on expanders and prove Theorems 4-5. We conclude with some open problems, conjectures and concluding remarks in §8.

2.6 Related work and connections to other models

The social network model on infinite graphs is investigated in [10]. The setup studied in [10] is as follows. Given an infinite connected regular graph $G = (V, E)$, place at each vertex $\text{Pois}(\lambda)$ walkers performing independent lazy simple random walks on G simultaneously. When two walkers visit the same vertex at the same time they are declared to be acquainted. It is shown in [10] that when G is vertex-transitive and amenable, for all $\lambda > 0$ a.s. every pair of walkers will eventually have a path of acquaintances between them. In contrast, it is shown that when G is non-amenable (not necessarily transitive) there is always a phase transition at some $\lambda_c(G) > 0$. General bounds on $\lambda_c(G)$ are given and the case that G is the d-regular tree $\mathcal{T}_d$ is studied in more detail (it is shown that $c \leqslant \lambda_c(\mathcal{T}_d)/\sqrt{d} \leqslant C$). Finally, it is shown that in the non-amenable setup there exists a finite time t such that a.s. there exists an infinite set of walkers having a path of acquaintances between them by time t. We note that the last result is the infinite setup counterpart of our Theorem 5.

We believe that, in spirit, the results in this paper should be true also for some other particle systems involving independent random walks. The $A + B \to 2B$ family of models [14, 15], often interpreted as models for a spread of an epidemic/rumor, are defined by the following rule: there are type A and B particles occupying a graph G, say with densities $\lambda_A, \lambda_B > 0$. They perform independent SRW with holding probabilities $p_A \in [0, 1]$ and $p_B \in [0, 1)$. When a type B particle collides with a type A particle, the latter transforms into a type B particle. The *frog model* is the particular case of the above dynamics in which the type A particles are immobile ($p_A = 1$) and $\lambda_A = \lambda_B = \lambda$.

We now consider a variant of the above model which is intimately related to the SN model. Consider the case in which initially only the particles at some vertex o are of type B and an additional B particle is planted at o (this is done to ensure that there is at least one B particle). Consider the case that each B particle performs only t steps of its walk before vanishing (after which it cannot become reactivated again). We call t the lifetime of the B particles. We say that the process dies out when there are no B particles left. One then defines the *susceptibility*, $\mathcal{S}(G)$ as the minimal lifetime which is sufficient so that all particles are transformed into type B particles before the process dies out (to make this definition formal, we think of the particles as first picking the infinite trajectory of their random walks, but then, once transforming into a B particle, performing only t additional steps of their walk). In the interpretation of the model as a model for a spread of a rumor/epidemic the susceptibility is meant to capture "how long should a virus live in order to infect the entire population" or "how interesting should a rumor be, so that eventually everybody will hear it". Similarly, one can consider the minimal lifetime required so that every vertex is visited by at least one B particle.

In [9, 5] parallel results to Theorems 2,3,4 and 5 are proved for the above variant in the case of the frog model, where the susceptibility replaces the social connectivity time. We strongly believe that (certain variants of) Theorems 2,3,4 and 5 should hold even when the A particles are mobile. Moreover, we have the following conjecture concerning an analog of Theorem 1.

Conjecture 2.3. *Consider the aforementioned variant of the $A + B \to 2B$ model with $\lambda_A = \lambda_B = \lambda$ and some fixed $p_A \in [0, 1]$ and $p_B \in (0, 1)$. Then there exist some $C_{d,\lambda}, \ell > 0$, such that for every sequence of finite connected graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$ of maximal degree at most d, we have that*

$$\lim_{n \to \infty} \mathbb{P}_\lambda \left[\mathcal{S}(G_n) \leqslant C_{d,\lambda} \log^\ell |V_n| \right] = 1.$$

3 Preliminaries and additional notation

3.1 Reversibility, stationarity of the occupation measure and independence of the number of walkers performing different walks.

Let $G = (V, E)$ be an Eulerian digraph. Let P be the transition kernel of LSRW on G. We denote by $P^t(u, v)$ the t-steps transition probability from u to v. We now establish a certain independence property for walks in G, which in particular implies a certain stationarity property of the SN model. Recall that $\pi_v := d_v/2|E|$ and $\bar{\pi} = |V|\pi$. For LSRW on a graph G **reversibility** is the property that for every $v, u \in V$ and $t \geqslant 0$,

$$\pi_v P^t(v, u) = \pi_u P^t(u, v).$$

A **walk** of length k in G is a sequence of $k + 1$ vertices $(v_0, v_1, \ldots, v_k)$ such that for all $0 \leqslant i < k$ either $v_i = v_{i+1}$ or $(v_i, v_{i+1}) \in E$ (or for a graph $\{v_i, v_{i+1}\} \in E$). Let Γ_k be the collection of all walks of length k in G. For a walk $\gamma = (\gamma_0, \ldots, \gamma_k) \in \Gamma_k$ for some $k \geqslant 1$, let

$$p(\gamma) := \prod_{i=0}^{k-1} P(\gamma_i, \gamma_{i+1}) \quad \text{and} \quad q(\gamma) := \bar{\pi}_{\gamma_0} p(\gamma).$$

Note that $p(\gamma)$ is precisely the probability that a walker whose starting position is γ_0 performed the walk γ (i.e. her position at time i is γ_i for all $i \leqslant k$). Let γ' be the reversal of γ (i.e. $\gamma'_i = \gamma_{|\gamma|-i}$ for all $i \leqslant |\gamma|$). Reversibility is equivalent to the following condition:

$$q(\gamma) = q(\gamma'), \quad \text{for every walk } \gamma.$$

We denote the number of walkers who performed a walk γ by X_γ. We denote the number of walkers whose position at time t is v by $Y_v(t)$. Since π is the stationary distribution of the walks, $\mathbb{E}[Y_v(t)] = \sum_{u \in V} \bar{\pi}_u P^t(u, v) = \bar{\pi}_v = \mathbb{E}[Y_v(0)]$, for every vertex v and time $t > 0$. Thus, the following fact follows easily from Poisson thinning.

Fact 3.1. *Let G be an Eulerian digraph. Then for all $t > 0$ and $\gamma \in \Gamma_t$, $X_\gamma \sim \text{Pois}(q(\gamma))$. Moreover, $(X_\gamma)_{\gamma \in \Gamma_t}$ are independent for each fixed $t > 0$. Consequently, for each fixed $t > 0$ we have that $(Y_v(t))_{v \in V}$ are independent and that $Y_v(t) \sim \text{Pois}(\bar{\pi}_v)$ for all $v \in V$.*

Next we note that if the average degree of $G = (V, E)$ is d, then for all $u, v \in V$

$$\bar{\pi}_v = |V| \cdot \frac{d_v}{|E|} = d_v/d, \qquad \bar{\pi}_v/\bar{\pi}_u = d_v/d_u. \tag{3.1}$$

We shall use the following concentration estimate. Let $Y \sim \text{Pois}(\mu)$. Then

$$\forall a \geqslant 0, \quad \max\{\mathrm{P}[Y \geqslant \mu + a], \mathrm{P}[Y \leqslant \mu - a]\} \leqslant \exp\left(-\frac{a^2}{2(\mu + a/3)}\right). \tag{3.2}$$

By taking an appropriate limit, this follows from Bernstein's inequality which says that if $\xi_1, \ldots, \xi_m$ are independent Bernoulli r.v.'s, $Y = \sum_{i \in [m]} \xi_i$, $\mu = \mathbb{E}[Y]$ and $\sigma^2 = \text{Var}(Y)$, then

$$\forall a \geqslant 0, \quad \max\{\mathrm{P}[Y \geqslant \mu + a], \mathrm{P}[Y \leqslant \mu - a]\} \leqslant \exp\left(-\frac{a^2}{2(\sigma^2 + a/3)}\right). \tag{3.3}$$

Another concentration estimate which we shall use in the proof of some of the lower bounds on SC is the McDiarmid's inequality, which we now state. Let $f : \mathcal{X}^n \to \mathbb{R}$. Let $c_i := \sup|f(\mathbf{x}) - f(\mathbf{y})|$, where the supremum is taken over all $\mathbf{x} := (x_1, \ldots, x_n), \mathbf{y} := (y_1, \ldots, y_n) \in \mathcal{X}^n$ such that $x_j = y_j$ for all $j \neq i$. Let $X_1, \ldots, X_n$ be i.i.d. $\mathcal{X}$ valued r.v.'s. Then

$$\mathrm{P}[f(X_1, \ldots, X_n) < \mathbb{E}[f(X_1, \ldots, X_n)] - \varepsilon] \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i \in [n]} c_i^2}\right). \tag{3.4}$$

3.2 Negative association

Let $G = (V, E)$ be a finite Eulerian digraph. The goal of this subsection is to formulate the intuitive notion that given a certain deterministic initial configuration of walkers, a time $t > 0$ and a vertex x, the more walkers there are at x at time t, the less walkers we expect to find at the other vertices of G.

The partial order $\preceq$ on $\mathbb{Z}_+^V$ is defined as follows. For $a, b \in \mathbb{Z}_+^V$ we have that $a \preceq b$ if $a(v) \leqslant b(v)$ for all v. We say that a function $f : \mathbb{Z}_+^V \to \mathbb{R}$ is *increasing* (resp. *decreasing*) if $f(a) \leqslant f(b)$ whenever $a \preceq b$ (resp. whenever $b \preceq a$, equivalently, $-f$ is increasing).

Let $\mathbf{X}_1 := (X_1(v))_{v \in V}$ and $\mathbf{X}_2 := (X_2(v))_{v \in V}$ be random elements in $\mathbb{Z}_+^V$. Denote their laws by μ_1 and μ_2, respectively. We say that (the law of) $\mathbf{X}_1$ *stochastically dominates* (the law of) $\mathbf{X}_2$ if $\int f d\mu_2 \leqslant \int f d\mu_1$ for all increasing f. We say that $f : \mathbb{Z}_+^V \to \mathbb{R}$ depends only on the co-ordinates in A if $f(z) = f(z')$ whenever $z(a) = z'(a)$ for all $a \in A$. We say that $\mathbf{Y} := (Y(v))_{v \in V}$ are *negatively associated* if for every two disjoint sets $A, B \subset V$ and any two increasing functions f, g such that the former depends only on the co-ordinates in A and the latter depends only on the co-ordinates in B we have that $\mathbb{E}[f(\mathbf{Y})g(\mathbf{Y})] \leqslant \mathbb{E}[f(\mathbf{Y})]\mathbb{E}[g(\mathbf{Y})]$. It is immediate from the definition that if $(Y(v))_{v \in V}$ are negatively associated then:

(1) The law of $(Y_u)_{u:u \neq v}$ stochastically dominates its conditional law given $Y_v > 0$ (take above $A = \{v\}, B = V \setminus \{v\}$ and $f = \mathbf{1}_{\{Y_v \geqslant 1\}}$).

(2) $(Y(v))_{v \in V}$ are negatively correlated, i.e. $\mathbb{E}[Y_v Y_u] \leqslant \mathbb{E}[Y_v]\mathbb{E}[Y_u]$ for all $u \neq v$.

(3) If $V = [n]$ then for all increasing (resp. decreasing) $f_1, \ldots f_n : \mathbb{Z}_+ \to \mathbb{R}$ we have that

$$\mathbb{E}[\prod_{i \in [n]} f_i(Y_i)] \leqslant \prod_{i \in [n]} \mathbb{E}[f_i(Y_i)]. \tag{3.5}$$

The following result is due to Dubhashi and Ranjan [7, Theorem 13].

Fact 3.2. *Let $X_1, \ldots, X_m$ be independent $[n]$ valued random variables. For $i \in [n]$ let $Y_i := \sum_{j \in [m]} \mathbf{1}_{\{X_j = i\}}$ be the number of X_ℓ's which take the value i. Then $(Y_i)_{i \in [n]}$ are negatively associated.*

4 A general upper bound on SC - proof of (2.1)

In this section we prove (2.1), which is the harder and more interesting part of our main result, Theo

rem 1. Throughout the section we fix some connected n-vertex Eulerian digraph $G = (V, E)$ of average degree d and set

$$t := t_{d,|V|} := \lceil Cd^{1_{G \text{ is not regular}}} \log n \rceil^2 \tag{4.1}$$

for some sufficiently large absolute constant C to be determined later.

4.1 An overview of the proof

We first describe an easy auxiliary baby model (§4.2) and analyze it (Lemma 4.2). We then (§4.3) define the notions of a marginally (α, t)-merging configuration, a globally (α, p, t)-merging configuration and a (δ, t)-balanced configuration (Definition 4.3) and explain how the analysis of the SN model can be reduced to that of the auxiliary baby model, assuming one of the first two properties hold at each given time w.h.p. (Proposition 4.7). Finally, in §4.4 we show that at each given time w.h.p. the configuration is $(1/8, t)$-balanced, and that this implies that it is marginally (α, t)-merging for some $\alpha = \alpha_d$.

4.2 Auxiliary baby models

Consider an arbitrary discrete-time coalescent process starting with m distinct classes. Define the coalescence time CT as the minimal time in which there is only one class (if there is no such time it is defined to be ∞). We now describe several simple coalescence models, and derive upper bounds on the coalescence time in each model. We intentionally start with two simple models, each involving a deterministic condition, and then work our way towards a slightly more complicated model which is more faithful to the actual details of our argument for the SN model.

Clearly, if prior to CT in every time unit every class merges with at least one other class, then deterministically, it must be the case that $CT \leqslant \lceil \log_2 m \rceil$. We now generalize this simple model, while introducing some terminology which will be used throughout the section.

Definition 4.1. *Fix some $t > 0$. Let $t_k := kt$. We call the time interval $(t_k, t_{k+1}]$ the $(k+1)$-**th round**. We say that the $(k+1)$-th round is a p-**merging round** if either $t_k \geqslant CT$ (this case is taken for technical reasons) or $t_k < CT$ and at least a p-fraction of the classes (w.r.t. time t_k) have merged with at least one other class by the end of time t_{k+1}.*

The last definition makes sense for the SN model, as well as for a general coalescence model. Note that if the number of classes at the beginning of a p-merging round is $\ell > 1$, then the number of classes at the end of the round is at most $(1 - p/2)\ell$. Assume that initially there are at most m classes. If all rounds are p-merging then, deterministically, $CT \leqslant \lceil C_p t \log m \rceil$ (for concreteness, $C_p := 1/\log(\frac{2}{2-p})$). More generally, even if not all rounds are p-merging, there can be at most $\lceil C_p \log m \rceil$ p-merging rounds prior to CT. This motivates the following simple model in which each round is p-merging with probability at least α, regardless of the history. We intentionally keep the description and analysis of this model somewhat informal, deferring a more rigorous statement to Lemma 4.2.

Simplified baby model: Start with m classes. At each round flip a coin with heads probability at least α (i.e. the heads probability in each round may depend on the history up to that time, but is guaranteed to be at least α). If it lands on heads then the round is p-merging. The list of pairs of classes which merge with one another in each round (given the evolution of the model up to the beginning of the round and the result of the coin flip corresponding to the round) is determined according to some arbitrary predetermined rule (satisfying that whenever the coin lands on heads the requirement of being p-merging is satisfied by the list of pairs of merged classes), perhaps using some additional independent source of randomness. The predetermined rule can be thought of as one decided by an adversary.

It is intuitively clear that there exists $C_{\alpha,p}$ of the form $C_{\alpha,p} = K_1 C_p / \alpha \leqslant \frac{K_2}{\alpha p}$ (for some sufficiently large absolute constants K_1, K_2, where C_p is as above) such that the probability that the total number of rounds exceeds $L = L_{m,\alpha,p} := \lceil C_{\alpha,p} \log m \rceil$ tends to 0 as $m \to \infty$. Informally, one can argue that the number of p-merging rounds in the first L rounds, stochastically dominates the $\mathrm{Bin}(L, \alpha)$ distribution and then apply an appropriate large deviation estimate in order to bound $P[\mathrm{Bin}(L, \alpha) < C_p \log m]$.

The auxiliary baby model: We now consider another small variation of the last model by introducing an error term ε_m. We refer to this variation as the *auxiliary baby model*. Consider the case that at each round with probability we only have probability $\geqslant 1 - \varepsilon_m$, for some $\varepsilon_m = o(1/\log m)$, that the model evolves according to the above description (i.e. like the "simplified baby model"), while with the complement probability it stops evolving (i.e. the coalescence process is terminated). It is intuitively clear that the conclusion from the previous paragraph remains valid, with an additional error term

of $\varepsilon_m L = o(1)$ for the probability that the number of rounds exceeds L. We will show that the analysis of the SN model on an Eulerian graph $G = (V, E)$ of average degree d can be reduced to that of the auxiliary baby model (for a suitable choice of parameters).

Observe that above it was sufficient to consider the number of classes at the end of each round in order to bound CT from above. This motivates the following lemma, which formulates rigorously the assertions made in the previous paragraphs. One should think of $J_k + 1$ below as the number of classes at the end of the k-th round, and so $T_0 := \inf\{k : J_k = 0\}$ corresponds to the coalescence time CT.

Lemma 4.2. *Let J_k be a non-increasing $\mathbb{Z}_+$-valued random process, measurable w.r.t. filtration $(\mathcal{F}_k)_{k \geqslant 0}$, with $J_0 \leqslant m$, satisfying that for some $0 < \alpha, p, \varepsilon_m < 1$, for each k there exists some events $A_k \in \mathcal{F}_k$ with $\mathrm{P}(A_k) \geqslant 1 - \varepsilon_m$ satisfying that*

$$\mathbb{E}[1_{\{J_{k+1} \leqslant (1-p/2)J_k\}} 1_{A_k} \mid \mathcal{F}_k] \geqslant \alpha 1_{A_k}. \tag{4.2}$$

Let $T_0 := \inf\{k : J_k = 0\}$. Then there exists some constant $C_{\alpha,p} \leqslant K/(\alpha p)$ such that

$$\mathrm{P}[T_0 > \lceil C_{\alpha,p} \log m \rceil] \leqslant m^{-2} + \varepsilon_m(\lceil C_{\alpha,p} \log m \rceil + 1).$$

The proof of Lemma 4.2 follows by noting that by (4.2) $M_k := (1 - \frac{1}{2}\alpha p)^{-k} J_k 1_{\cap_{i \in [k]} A_i}$ is a supermartingale and $\mathrm{P}[T_0 > k] = \mathbb{E}[J_k 1_{\cap_{i \in [k]} A_i}] + \mathrm{P}[\cup_{i \in [k]} A_i^c] \leqslant (1 - \frac{1}{2}\alpha p)^k \mathbb{E}[M_0] + \varepsilon_m k$. The details are given in the appendix in §A.2.

4.3 Reducing the SN model to the baby model via merging configurations

We are interested in reducing the SN model to the auxiliary baby model. Recall that $Y_v(s)$ is the number of walkers at vertex v at time s. It is beneficial to consider a notion of "independence of the history" which depends only on $(Y_v(s))_{v \in V}$, whose distribution is stationary and hence convenient to work with, rather than also on the identity of the classes of walkers at time s, which might have a complex dependency structure. This motivates the following key definitions.

> **Definition 4.3.** *Let $\mathbf{Y} := (Y_v)_{v \in V} \in \mathbb{Z}_+^V$. We identify $\mathbf{Y}$ with a configuration of walkers in which for each $v \in V$ there are initially Y_v particles at v. The walkers then perform a LSRW on G of length t, where t is as in (4.1).*
>
> - *Let $\mathcal{P}$ be a partition of the walkers into (disjoint) sets $\mathcal{A}_1, \ldots, \mathcal{A}_m$ with $m \geqslant 2$. We say that $\mathcal{P}$ **respects vertices** if walkers which initially occupy the same vertex all belong to the same set in the partition (i.e. for all v, each $\mathcal{A}_i$ contains either all Y_v walkers initially occupying v or none of them).*
>
> - *We say that the configuration $\mathbf{Y}$ is **marginally** (α, t)-**merging**, if the following holds. For every vertex-respecting partition of the walkers into sets, $\mathcal{A}_1, \ldots, \mathcal{A}_m$ with $m \geqslant 2$, for every $\mathcal{A}_i$ which contains less than half of the walkers, the probability that at least one walker from $\mathcal{A}_i$ will meet some walker not belonging to $\mathcal{A}_i$ by time t is at least α. We say such a configuration is **globally** (α, p, t)-**merging** if for every vertex-respecting partition of the walkers into sets, $\mathcal{A}_1, \ldots, \mathcal{A}_m$ with $m \geqslant 2$, we have that with probability at least p, there are at least $\alpha(m-1)$ i's such that at least one walker from $\mathcal{A}_i$ met by time t some walker not belonging to $\mathcal{A}_i$.*
>
> - *We say that the configuration $\mathbf{Y}$ is (δ, t)-**balanced**, if for all $v \in V$ we have that*
>
> $$\sum_{u \in V} Y_u P^t(u, v) \geqslant (1-\delta)\bar{\pi}_v \quad (\text{recall that } \bar{\pi}_v := |V|\pi_v), \tag{4.3}$$
>
> *and*
>
> $$\max_{u \in V} Y_u/d_u \leqslant \log |V|. \tag{4.4}$$

- *We say that the configuration* $\mathbf{Y}$ *is **fully** (δ, t)-**balanced**, if it is (δ, t)-balanced and in addition* $\sum_{u \in V} Y_u P^t(u, v) < (1 + 2\delta)\bar{\pi}_v$ *for all* $v \in V$.

We believe that some of the ideas from the proof of (2.1) and in particular the notion of balanced configurations can be useful for other particle systems. For instance, the approach taken in [11], where the mixing time of the exclusion process on regular graphs is bounded from above in terms of the mixing of independent random walks, was greatly inspired by the proof of (2.1).

Throughout, we take δ to be some constant in $(0, 1/8]$. We now record the fact that requirement (4.4) holds w.h.p. when $(Y_u)_{u \in V}$ are independent and $Y_u \sim \mathrm{Pois}(\bar{\pi}_u)$.

Lemma 4.4. *Let* $(Y_u)_{u \in V}$ *be independent and* $Y_u \sim \mathrm{Pois}(\bar{\pi}_u)$. *Then*

$$\mathrm{P}[\max_{u \in V} Y_u/d_u > \log |V|] \leqslant \sum_{u \in V} C \left(\frac{e}{\log |V|} \right)^{d_u \log |V|}.$$

Proof: $\mathrm{P}[\mathrm{Pois}(\bar{\pi}_u) > d_u \log |V|] \leqslant \mathrm{P}[\mathrm{Pois}(d_u) > d_u \log |V|] \leqslant C\mathrm{P}[\mathrm{Pois}(d_u) = \lceil d_u \log |V| \rceil]$. Finally, apply a union bound over all $u \in V$. $\square$

Consider the SN model viewed only at integer multiples of t. Think of the time interval $(kt, (k+1)t]$ as the k-th round (for $k = 0$ we take $[0, t]$ instead of $(0, t]$) Assume that the configuration at time kt is globally (α, p, t)-merging. Then the k-th round is $\alpha/2$ merging in the sense of Definition 4.1 w.p. at least p. In light of Lemma 4.2, in order to prove (2.1), we need to show that for some absolute constant c_1, at each given time the configuration is globally $(c_1/d^{1+1_{G \text{ is not regular}}}, c_1, t)$-merging, with probability at least $1 - |V|^{-2}$, where here t is as in (4.1). This observation is recorded in Proposition 4.7 below.

A priori, it may appear difficult to determine that a configuration is either marginally (α, t)-merging or globally (α, p, t)-merging (for some $\alpha, p \in (0, 1)$), as the sets $\mathcal{A}_1, \ldots, \mathcal{A}_m$ from Definition 4.3 form an arbitrary vertex-respecting partition. On the other hand, the property of being (δ, t)-balanced is more tractable as it is stated in terms of a (lack of) a large deviation behavior for the expected number of walkers occupying each site t steps into the future, given the current configuration. We now explain the relations between the four properties of being (δ, t)-balanced, fully (δ, t)-balanced, marginally (α, t)-merging and globally (α, p, t)-merging.

To prove (2.1) we shall later show that (i) if C from the definition of t in (4.1) is taken to be sufficiently large then at each given time the configuration of walkers is fully $(\frac{1}{8}, t)$-balanced w.p. at least $1 - 2|V|^{-2} + C'|V|^{2 - \log \log |V|}$ and (ii) if G is regular then a fully $(\frac{1}{8}, t)$-balanced configuration is globally $(c_1/d, c_1, t)$-merging for some absolute constant $c_1 > 0$.

Before proving (ii) we show that (iii) a $(\frac{1}{8}, t)$-balanced configuration is marginally $(c_1/d^{1+1_{G \text{ is not regular}}}, t)$-merging for some absolute constant $c_1 > 0$ and (iv) a marginally (α, t)-merging configuration is globally $(\alpha/2, \alpha/4, t)$-merging (this is done in Corollary 4.6).

To prove (iv) we use the following simple lemma which asserts that a uniform lower bound on the marginal probabilities of certain m events, implies a corresponding lower bound on the probability that a certain fraction of them occur. The first equation from the lemma follows by a straightforward application of the Paley–Zygmund inequality (e.g. [13, Lemma 4.1]). The second equation is the one-sided Chebyshev's inequality.

Lemma 4.5. *Let* $\xi_1, \ldots, \xi_m$ *be indicator random variables. Let* $p_i := \mathbb{E}[\xi_i]$, $p := \min_i p_i$, $S := \sum_{i=1}^{m} \xi_i$, $\mu = \mathbb{E}[S]$ *and* $\sigma^2 := \mathrm{Var}(S) = \sum_i p_i(1 - p_i) + 2\sum_{i,j:i<j} \mathrm{Cov}(\xi_i, \xi_j)$. *Then*

$$\mathrm{P}[S \geqslant pm/2] \geqslant \mathrm{P}[S \geqslant \mu/2] \geqslant (\mu/2)^2/\mathbb{E}[S^2] \geqslant \mu/(4m) \geqslant p/4. \tag{4.5}$$

$$\forall c \geqslant 0, \quad \mathrm{P}[S \geqslant \mu - c\sigma] \geqslant \frac{c^2}{1 + c^2}. \tag{4.6}$$

In particular, if $\mathrm{Cov}(\xi_i, \xi_j) \leqslant 0$ *for all* $i \neq j$ *then* $\sigma^2 \leqslant \mu$ *and so* $\mathrm{P}[S < \frac{1}{2}\mu] \leqslant \frac{4}{4+\mu}$.

In our setup, (4.5) implies the following. Assuming that the configuration at time $t_k = kt$ is marginally α-merging and $\mathcal{A}_1, \ldots, \mathcal{A}_m$ are the classes of walkers at time t_k (the classes clearly form a vertex-respecting partition of the walkers), then with probability at least $\alpha/4$, there exists an index set I of size at least $\lceil \alpha(m-1)/2 \rceil$ such that for all $i \in I$ there is some walker from $\mathcal{A}_i$ who met in the time interval $(t_k, t_k + t]$ some walker not belonging to $\mathcal{A}_i$. We now record this observation.

Corollary 4.6. *A marginally* (α, t)-*merging configuration is globally* $(\alpha/2, \alpha/4, t)$-*merging.*

Derivation of (2.1) **assuming that at each given time the configuration is marginally or globally merging** w.h.p.**:**

Proposition 4.7. *Let* $G = (V, E)$ *be a connected Eulerian digraph. Fix some* $t > 0$ *and* $\alpha \in (0, 1)$. *Assume that for each time* s *w.p. at least* $1 - \varepsilon_n$ *the configuration of walkers at time* s *is globally* (α, p, t)-*merging. Then there exists an absolute constant* $C > 0$ *such that*

$$\mathrm{P}[\mathrm{SC}(G) \geqslant C(\alpha p)^{-1} t \log |V|] \leqslant |V|^{-2} + \varepsilon_n C(\alpha p)^{-1} \log |V|.$$

In particular, if at each given time w.p. at least $1 - \varepsilon_n$ *the configuration of walkers is marginally* (α, t)-*merging then*

$$\mathrm{P}[\mathrm{SC}(G) \geqslant C'\alpha^{-2} t \log |V|] \leqslant |V|^{-2} + \varepsilon_n C'\alpha^{-2} \log |V|.$$

Proof. The second assertion follows from the first via Corollary 4.6. For the first assertion, if we think of each interval of length t as a round, we have that if at time kt the configuration of walkers is globally (α, p, t)-merging, then the $(k+1)$-th round (corresponding to the time interval $(kt, (k+1)t]$) is $\frac{\alpha}{4}$-merging, in the sense of Definition 4.1, with probability at least p (this holds regardless of the outcomes of the previous rounds). Since at the beginning of each round w.p. at least $1 - \varepsilon_n$ the configuration of walkers is globally (α, p, t)-merging, we are precisely at the setup of the auxiliary baby model and the assertion of the proposition now follows from Lemma 4.2. $\qquad\square$

4.4 Proof of the validity of the marginally (α, t)-merging property via a reduction to the (δ, t)-balance condition

Note that if $(Y_v(s))_{v \in V}$ is distributed like the configuration of walkers in the SN model on G at time s, then (by stationarity of the occupation measure, which allows one to consider only the case that $s = 0$) for every $v \in V$ the sum $\bar{Z}_v(s) := \sum_u P^t(u, v)Y_u(s)$ in (4.3) (whose mean is $\bar{\pi}_v$) is a linear combination (with non-negative coefficients) of independent Poisson r.v.'s. As demonstrated in Lemma 4.10 below, in order to obtain a large deviation estimate for such a linear combination (namely, for the event $\bar{Z}_v(s) < (1 - \delta)\bar{\pi}_v = (1 - \delta)\mathbb{E}[\bar{Z}_v(s)]$), it suffices to control the maximal coefficient $P^t(u, v)$ appearing in the sum. We obtain such control using the following decay estimate for $\max_{x,y} P^t(x, y)$.

Recall that d_y is the degree of vertex y.

Fact 4.8 ([6] Lemma 2.4). *There exists an absolute constant* M *such that for every finite connected Eulerian digraph* $G = (V, E)$, *LSRW on* G *satisfies*

$$\forall t \geqslant 1, \quad \max_{x,y} |P^t(x, y) - \pi_y| \leqslant M\sqrt{1/t}[d_y 1_{G \text{ is not regular}} + 1_{G \text{ is regular}}]. \tag{4.7}$$

In what comes, we shall assume that (*) $12d^{1_{G \text{ is not regular}}} \log |V| \leqslant \delta^2 |E|$. Note that if (*) fails, then $|V|$ has to be bounded (recall that δ is a constant and that $|V|d = 2|E|$, as d is

the average degree). The following proposition, whose proof is deferred to §4.5 implies (2.1) when (*) fails.

Proposition 4.9. *Let $G = (V, E)$ be a connected Eulerian digraph. Then for some $C > 0$,*

$$\mathrm{P}[\mathrm{SC}(G) > C|E||V|\log|V|] \leqslant C/|V|.$$

By Fact 4.8, if we set $C = 12M\delta^{-2}$ in the definition of t ((4.1)) then

$$\max_{x,y} P^t(x,y) \leqslant \pi_y + \frac{\delta^2 d_y}{12d\log|V|}. \tag{4.8}$$

Assume (*) holds for some fixed $0 < \delta \leqslant 1/8$. By (4.8) we have that

$$\max_{x,y} P^t(x,y) \leqslant \delta^2 d_y/(6d\log|V|). \tag{4.9}$$

In light of Lemma 4.4 and the paragraph following (4.3), by combining (4.9) with the following lemma, it follows by a union bound over the vertices of G that the probability that the configuration is not (δ, t)-balanced at some fixed time s is at most $C_1/|V|^2$ (assuming (*), when the constants are chosen as indicated above).

Lemma 4.10. *Let $\xi_1, \ldots, \xi_m$ be independent Poisson random variables. Let $p_1, \ldots, p_m \in (0,1)$. Let $p_* := \max p_i$, $S := \sum_{i=1}^m p_i \xi_i$ and $\mu := \mathbb{E}[S]$. Then for all $\varepsilon \in (0,1]$*

$$\begin{aligned}
\mathrm{P}[S \leqslant (1-\varepsilon)\mu] &\leqslant e^{-\frac{1}{2}\mu\varepsilon^2/p_*}. \\
\mathrm{P}[S \geqslant (1+\varepsilon)\mu] &\leqslant e^{-\frac{1}{4}\mu\varepsilon^2/p_*}.
\end{aligned} \tag{4.10}$$

The proof of Lemma 4.10 is given in the appendix in §A.1.

Corollary 4.11. *Let $Y_v(s)$ be the number of walkers at vertex v at time s. Assume that $Y_v(0) \sim \mathrm{Pois}(\bar{\pi}_v)$ independently for different vertices. Let t be as above and assume that (*) holds. Let $n := |V|$ and $\bar{Z}_v(s) := \sum_{u \in V} P^t(u,v) Y_u(s)$. Then for all $s \geqslant 0$ and $v \in V$ we have*

$$\mathrm{P}[\bar{Z}_v(s) \leqslant (1-\delta)\bar{\pi}_v] \leqslant \exp\left[-\frac{\bar{\pi}_v \delta^2}{2(\delta^2 d_v/(6d\log n))}\right] \leqslant n^{-3}.$$

$$\mathrm{P}[\bar{Z}_v(s) \geqslant (1+2\delta)\bar{\pi}_v] \leqslant \exp\left[-\frac{\bar{\pi}_v(2\delta)^2}{4(\delta^2 d_v/(6d\log n))}\right] \leqslant n^{-6}.$$

The following proposition concludes the proof of (2.1) when G is not regular, and provides a version of (2.1) also for a regular G which has a worse degree dependence than (2.1). Below we assume that the total number of walkers is at most $|V| + |V|^{2/3}$. By (3.2), the probability that this fails is at most $\exp(-\frac{1}{3}|V|^{1/3})$ and hence there is no harm in assuming this holds.

Proposition 4.12. *Let $G = (V, E)$ be an n-vertex connected Eulerian digraph of average degree d. Let $\xi := \mathbf{1}\{G \text{ is not regular}\}$. Let $0 < \delta \leqslant 1/8$. Assume that (*) holds. Assume that a configuration of walkers is (δ, t)-balanced for $t := \lceil 12d^\xi M\delta^{-2}(\log n)\rceil^2$ and that the total number of walkers it contains is smaller than $n + n^{2/3}$. Then there exists some absolute constant $\bar{c} > 0$ such that the following hold.*

   (i) *The configuration is marginally $(\bar{c}/d^{1+\xi}, t)$-merging.*

   (ii) *If G is regular and $|\{(u,v) \in E : u \in A, v \notin A \text{ or } u \notin A, v \in A\}| \geqslant c_0 d$ for every set $A \subset V$ such that $|A| \leqslant \frac{2}{3}(n + n^{2/3})$ then the configuration is marginally $(\bar{c}c_0, t)$-merging.*

Proof. Consider some configuration $(X_v)_{v \in V}$ of walkers (i.e. for all v the number of walkers at vertex v is X_v) with at most $n + n^{2/3}$ walkers (i.e. $\sum_v X_v \leqslant n + n^{2/3}$) and a vertex-respecting partition of the walkers into disjoint sets $\mathcal{A}_1, \ldots, \mathcal{A}_m$. Since both the properties of a configuration which we consider (being marginally (α, t)-merging and (δ, t)-balanced) are independent of the time, by stationarity of the occupation measure w.l.o.g. we may assume that this configuration corresponds to time 0. We start with part (i).

Let $\alpha_d := \bar{c}/d^{1+\xi}$, where $\bar{c} > 0$ shall be determined soon. We now fix some $1 \leqslant i \leqslant m$ such that $|\mathcal{A}_i| \leqslant (n + n^{2/3})/2$ and show that the probability that at least one walker from $\mathcal{A}_i$ meets some walker not from $\mathcal{A}_i$ by time t is at least α_d. Denote the collection of vertices occupied by the walkers from $\mathcal{A}_j$ (at time 0) by V_j (where $1 \leqslant j \leqslant m$). Denote the expected number of walkers from $\mathcal{A}_j$ (resp. not from $\mathcal{A}_j$, i.e. from $\cup_{i: i \neq j} \mathcal{A}_i$) at vertex v at time t by

$$\mu_v(j) := \sum_{u \in V_j} P^t(u, v) X_u, \quad (\text{resp. } a_v(j) := \sum_{i: i \neq j} \mu_v(i) = \sum_{i: i \neq j} \sum_{u \in V_i} P^t(u, v) X_u).$$

As $\sum_{v \in V} \mu_v(j) = |\mathcal{A}_j|$, there can be at most one set $\mathcal{A}_j$ with $\mu_v(j) > (1 - 2\delta)\bar{\pi}_v$ for all v (assuming $n \geqslant 8$; where we have used $\delta \leqslant 1/8$ and $\sum_{v \in V} X_v \leqslant n + n^{2/3}$). Hence we only have to treat the case that $\mu_v(i) \leqslant (1 - 2\delta)\bar{\pi}_v$ for all v and the case that $\mu_v(i) \leqslant (1 - 2\delta)\bar{\pi}_v$ holds for some of the vertices, but not for all.

Assume that $\mu_v(i) \leqslant (1 - 2\delta)\bar{\pi}_v$ for all v. Then since the configuration is δ-balanced

$$\forall v \in V, \quad a_v(i) \geqslant (1 - \delta)\bar{\pi}_v - \mu_v(i) \geqslant \delta \bar{\pi}_v = \frac{\delta d_v}{d}. \tag{4.11}$$

An elementary calculation (see Lemma 4.13) now shows that for all v the probability that there is a walker not from $\mathcal{A}_i$ at v at time t is at least $1 - e^{-a_v(i)}$. Hence trivially every walker from the set $\mathcal{A}_i$ has a chance of at least $1 - e^{-a_v(i)}$ of meeting some walker not from $\mathcal{A}_i$ at time t. In this case, the proof is concluded using (4.11).

Finally, assume that $\mu_v(i) > (1 - 2\delta)\bar{\pi}_v$ for some v but not for all v. We argue that in this case, there must be some vertex u such that for some absolute constant $c_1 > 0$

$$c_1/d \leqslant \mu_u(i) \leqslant (1 - 2\delta)\bar{\pi}_u. \tag{4.12}$$

Once this is establish, the proof is concluded as follows. Let u be as above. Since the configuration is (δ, t)-balanced we have that

$$a_u(i) \geqslant (1 - \delta)\bar{\pi}_u - \mu_u(i) \geqslant \delta \bar{\pi}_u.$$

By (4.12) together with the elementary Lemma 4.13 below, the probability that u is occupied at time t by both a walker from $\mathcal{A}_i$ and also by a walker not from $\mathcal{A}_i$ is at least

$$(1 - e^{-\mu_u(i)})(1 - e^{-a_u(i)}) \geqslant \frac{1}{4} \mu_u(i) a_u(i) \geqslant \frac{c_1}{4d} \delta \bar{\pi}_u \geqslant \bar{c}/d^{1+\xi} \tag{4.13}$$

as desired (where we have used $1 - e^{-x} \geqslant x - \frac{1}{2}x^2 \geqslant \frac{x}{2}$ for $x \in [0, 1]$). We now verify the existence of such u satisfying (4.12). Let

$$D_i := \{v : \mu_v(i) \geqslant (1 - 2\delta)\bar{\pi}_v\}. \tag{4.14}$$

We argue that if $u \notin D_i$ and $(v, u) \in E$ for some $v \in D_i$, then u satisfies (4.12) (by our assumption on $\mathcal{A}_i$ the set D_i is non-empty and is a proper subset of V). This follows from Lemma 4.14 below (whose proof is deferred to the appendix §A.3; See Remark 4.15 for the intuition behind this lemma), applied to the collection of walkers from $\mathcal{A}_i$ (i.e. take

$Y_y(0)$ in Lemma 4.14 to be 0 if the walkers which are at time 0 at vertex y do not belong to $\mathcal{A}_i$, otherwise take it to be the number of such walkers; Note that condition (4.4) from the definition of a (δ, t)-balanced configuration implies that the term $Y \exp(-c_1 t)$ from Lemma 4.14 is much smaller than the term $\mathbb{E}[Y_v(t)]$).

We now prove part (ii). We now assume G is regular, and so $\bar{\pi}_v = 1$ for all v. We fix some i. The case that $\mu_v(i) \leqslant 1 - 2\delta$ for all v is the same as before. We now treat the case that $\mu_v(i) > 1 - 2\delta$ for some v but not for all v. We may assume that $|\mathcal{A}_i| \leqslant \frac{1}{2}(n + n^{2/3})$, as there can be at most one class which contains more than half of the walkers. Let D_i be as in (4.14). Observe that $|D_i| \leqslant \frac{|\mathcal{A}_i|}{(1-2\delta)} \leqslant \frac{2}{3}(n + n^{2/3})$ (as $\delta \leqslant 1/8$) and so $\max\{|\partial_E^{\text{out}} D_i|, |\partial_E^{\text{in}} D_i|\} \geqslant \frac{c_0 d}{2}$, where $\partial_E^{\text{out}} A := \{(u,v) \in E : u \in A, v \notin A\}$ and $\partial_E^{\text{in}} A := \{(v,u) \in E : u \in A, v \notin A\}$. We first treat the case that $|\partial_E^{\text{out}} D_i| \geqslant \frac{c_0 d}{2}$.

For each $e = (v, u)$ let Z_e be the indicator of the event that there is some walker from $\mathcal{A}_i$ which was at v at time $t-1$ and then moved to u at time t. Let $Z := \sum_{e \in \partial_E^{\text{out}} D_i} Z_e$. As in the proof of part (i), by Lemma 4.14, there exists some $c > 0$ such that $\mathbb{E}[Z_e] > c/d$ for all $e \in \partial_E^{\text{out}} D_i$. Hence $\mathbb{E}[Z] \geqslant c_1 c_0$. By Fact 3.2 we have that $(Z_e)_{e \in E}$ are negatively correlated, and thus $\text{Var}(Z) \leqslant \mathbb{E}[Z]$. Thus, using the one-sided Chebyshev's inequality (4.6) we get that $P[Z > 0] \geqslant \frac{\mathbb{E}[Z]}{1 + \mathbb{E}[Z]} \geqslant c_2 c_0$. Conditioned on $Z_e = 1$ for $e = (v, u) \in \partial_E^{\text{out}} D_i$, as in the proof of part (i), the probability that there is a walker not from $\mathcal{A}_i$ at u at time t is bounded from below (as $u \notin D_i$ and the configuration is (δ, t)-balanced). This concludes the proof in the case that $|\partial_E^{\text{out}} D_i| \geqslant \frac{c_0 d}{2}$.

Now consider the case that $|\partial_E^{\text{in}} D_i| \geqslant \frac{c_0 d}{2}$. For each $e = (v, u)$ let $\widehat{Z}_e$ be the indicator of the event that there is some walker from $\cup_{j:j \neq i} \mathcal{A}_j$ which was at v at time $t-1$ and then moved to u at time t. Let $\widehat{Z} := \sum_{e \in \partial_E^{\text{in}} D_i} \widehat{Z}_e$. As in the proof of part (i), by Lemma 4.14, there exists some $c > 0$ such that $\mathbb{E}[\widehat{Z}_e] > c/d$ for all $e \in \partial_E^{\text{in}} D_i$. Hence $\mathbb{E}[\widehat{Z}] \geqslant c_1 c_0$. By Fact 3.2 we have that $(\widehat{Z}_e)_{e \in E}$ are negatively correlated. Thus as before, using the one-sided Chebyshev's inequality (4.6), we get that $P[\widehat{Z} > 0] > c_2 c_0 > 0$. Conditioned on $\widehat{Z}_e = 1$ for $e = (v, u) \in \partial_E^{\text{in}} D_i$, as in the proof of part (i), the probability that there is a walker from $\mathcal{A}_i$ at u at time t is bounded from below (as $u \in D_i$). $\qquad\square$

Lemma 4.13. *Let $\xi_1, \ldots, \xi_m$ be independent Bernoulli random variables. Denote $p_i := \mathbb{E}[\xi_i]$, $S := \sum_i \xi_i$ and $\mu = \mathbb{E}[S]$. Then $P[S = 0] = \prod_{i=1}^m (1 - p_i) \leqslant \exp[\sum_{i=1}^m -p_i] = e^{-\mu}$.*

Lemma 4.14. *Let $G = (V, E)$ be a finite Eulerian graph. Let $Y_v(t)$ be the number of walkers which occupy vertex v at time t, where $(Y_v(0))_{v \in V}$ is some arbitrary deterministic initial configuration of walkers and different walkers perform independent LSRWs. Let $(v, u) \in E$. Let $Y := \max_{x \in V} \frac{d_v}{d_x} Y_x(0)$. Then*

$$2d_v \mathbb{E}[Y_u(t)] \geqslant \mathbb{E}[Y_v(t-1)] \geqslant \frac{1}{3}(\mathbb{E}[Y_v(t)] - Y \exp(-c_1 t)). \qquad (4.15)$$

Remark 4.15. The proof of Lemma 4.14 relies on laziness in a crucial manner (in fact, this is the only use of laziness in the proof of (2.1)). The idea is that typically a lazy walk of length t makes at least $t/3$ lazy steps (the error term $Y \exp(-c_1 t)$ corresponds to the total contribution to $\mathbb{E}[Y_v(t)]$ of the paths that make less than $t/3$ lazy steps). Such a walk could be transformed into many walks of length $t - 1$ with the same end-point by deleting one of the lazy steps. The "cost" of this operation is at most a constant factor (this uses the fact that there are at least $\Omega(t)$ lazy steps). In fact, some algebra yields that it is at most $\sup_{k \in \mathbb{N}: k \in [t/3, t]} \binom{t}{k} / \binom{t}{k-1} \leqslant 3$.

4.5 Proof of proposition 4.9

Proof: By [6, Lemma 2.4] there exists some $C > 0$ such that for $t := \lceil C|E||V| \rceil$

$$P^t(x, y) \geqslant \pi_y/2, \quad \text{for all } x, y \in V.$$

Hence, by independence, for every pair of walkers with some arbitrary initial positions we have that the probability that they are at the same position at time t is at least $\sum_{x \in V} (\pi_x/2)^2 \geqslant \frac{1}{4|V|}(\sum_x \pi_x)^2 = \frac{1}{4|V|}$. We may assume that the total number of walkers Y is greater than $3|V|/4$, as the probability that this fails is exponentially small in $|V|$. Under this assumption, it follows from Lemma 4.5 that there exists some constant $\alpha > 0$ such that for every collection of $1 + \frac{1}{2}\lceil 3|V|/4 \rceil$ walkers with some arbitrary initial conditions, the probability that the first walker met at time t one of the other walkers in the collection is at least α. It follows that (if $Y > \frac{3|V|}{4}$) the configuration of walkers at each fixed time s, deterministically, has the (α, t)-merging property. The proof is now concluded using Proposition 4.7. □

4.6 Proof of (2.1) for regular G

In this subsection we prove (2.1) when G is regular. This is done by combining Proposition 4.16 below, which refines Proposition 4.12, with Proposition 4.7.

Let $G = (V, E)$ be a connected d-regular n-vertex Eulerian digraph. Let

$$t_*(G) := \max\left\{\hat{t}_*, \left\lceil \frac{2}{c_1} \log\log n \right\rceil\right\},$$

$$\text{where } \hat{t}_* := \inf\left\{t : \max_{x,y} |P^t(x, y) - \frac{1}{n}| \leqslant \frac{1}{3 \cdot 2^8 \log n}\right\} \tag{4.16}$$

where $c_1 > 0$ is as in Lemma 4.14. Note that as the walk is lazy, the probability it stays put for $\lfloor \log_2(c \log n) \rfloor$ steps is at least $\frac{1}{c \log n}$ and thus $t_*(G) \leqslant C\hat{t}_*$. Recall that by Fact 4.8

$$t_*(G) \leqslant C'(\log n)^2.$$

The choice of the constant $3 \cdot 2^8$ was made so that by Corollary 4.11, provided that (*) holds for $\delta = 1/8$, we have that the configuration of walkers is $(\frac{1}{8}, t_*(G))$-balanced w.p. at least $1 - 2n^{-2}$. The choice of the constant $\frac{2}{c_1}$ was made to ensure that for a $(\frac{1}{8}, t_*(G))$-balanced configuration we have in Lemma 4.14 that $\mathbb{E}[Y_v(t)] - Y \exp(-c_1 t) \geqslant \frac{1}{2}\mathbb{E}[Y_v(t)]$.

In this subsection we improve Proposition 4.12 as follows.

Proposition 4.16. *Let $G = (V, E)$ be a connected d-regular n-vertex Eulerian digraph satisfying (*) for $\delta = \frac{1}{8}$. Assume that a configuration of walkers is fully $(\frac{1}{8}, t_*(G))$-balanced and that the total number of walkers it contains is smaller than $n + n^{2/3}$. Then there exists some absolute constant $\bar{c} > 0$ such that the following hold.*

(i) The configuration is globally $(\bar{c}/d, \bar{c}, t_*(G))$-merging.

(ii) If $|\{(u, v) \in E : u \in A, v \notin A \text{ or } u \notin A, v \in A\}| \geqslant c_0 d$ for every set $A \subset V$ such that $|A| \leqslant \frac{2}{3}(n + n^{2/3})$ then the configuration is marginally $(\bar{c}c_0, t_*(G))$-merging.

Proof. We first prove part (i). Consider some configuration $(X_v)_{v \in V}$ of walkers with at most $n + n^{2/3}$ walkers and a vertex-respecting partition of the walkers into disjoint sets $\mathcal{A}_1, \ldots, \mathcal{A}_m$. Again, we may think of the current time as being time 0. Let $a_u(i)$ (resp. $\mu_u(i)$) be the expected number of walkers not from (resp. from) $\mathcal{A}_i$ which occupy u at time $t_*(G)$. First consider the case that there is a set $I \subseteq [m]$ of size at least $(m-1)/2$ such that for all $i \in I$ we have that $\mu_u(i) \leqslant \frac{3}{4}$ for all u. Then from the proof of Proposition 4.12 we see that for each $i \in I$ the probability that some walker from $\mathcal{A}_i$ has met by time $t_*(G)$ some walker not from $\mathcal{A}_i$ is at least $c_3 > 0$. By Lemma 4.5 we have that with probability at least $c_3/4$, there are at least $c_3|I|$ sets $\mathcal{A}_i$ with $i \in I$ such that some walker from $\mathcal{A}_i$ has met by time $t_*(G)$ some walker not from $\mathcal{A}_i$.

Now consider the case that there is a set $J \subseteq [m]$ of size at least $(m-1)/2$ such that for all $j \in J$ we have that $\mu_u(j) \geqslant \frac{3}{4}$ for some $u \in V$ but not for all $u \in V$. For each $j \in J$,

as before, let $D_j := \{v : \mu_u(j) > \frac{3}{4}\}$. Let

$$F_j := \{v \in D_j : \text{there exists some } u \notin D_j \text{ such that } (u, v) \in E\}.$$

For every $j \in J$ fix some $v_j \in F_j$. Observe that since the configuration is fully $(\frac{1}{8}, t_*(G))$-balanced we have that the sets $(F_j)_{j \in J}$ are disjoint, and so $(v_j)_{j \in J}$ are distinct.

Let Z_j be the indicator of the event that some walker from $\mathcal{A}_j$ and also some walker not from $\mathcal{A}_j$ occupying v_j at time $t_*(G)$. As in the proof of Proposition 4.12 we have that $\mathbb{E}[Z_j] \geqslant c_3/d$. By Fact 3.2 we have that $(Z_j)_{j \in J}$ are negatively correlated (indeed, for $i \neq j$ knowing that $Z_i = 1$ decreases the chances of both the requirements for $Z_j = 1$). Hence by the one sided Chebyshev's inequality (4.6) $\mathrm{P}[\sum_{j \in J} Z_j > \frac{c_3}{2d}|J|] \geqslant c_4 > 0$ as desired.

The proof of part (ii) is identical to the proof of part (ii) of Proposition (4.16). $\qquad\square$

4.7 Deterministic initial configuration

Let $G = (V, E)$ be a connected d-regular n-vertex Eulerian digraph. Consider the case that initially there is exactly one walker at each site and that the walkers perform independent LSRWs. Let $Y_v(s)$ be the number of particles occupying v at time s. Let $t_*(G)$ be as in (4.16). Assume n is sufficiently large so that $\frac{1}{n} \leqslant \frac{1}{3 \cdot 2^9 \log n}$ (which by the definition of $t_*(G)$ implies that $\max_{x,y} P^{t_*(G)}(x, y) \leqslant \frac{1}{2^9 \log n}$). We show that at each given time the configuration is fully $(\frac{1}{8}, t_*(G))$-balanced w.p. at least $2n^{-4} + n^{-\log \log n}$. From this, as in the proof of (2.1), we get that w.h.p. $\mathrm{SC}(G) \leqslant Cdt_*(G) \log n \leqslant C'd(\log n)^3$.

The proof of the following proposition is similar to that of Lemma 4.10. The main difference is that now the number of walkers at distinct vertices are not independent and thus one has to rely on (3.5) in order to bound $\mathbb{E}[\exp(\lambda Z_v(s))]$ from above.

Proposition 4.17. *Let $Z_v(s) := \sum_{u \in V} Y_u(s) P^{t_*(G)}(u, v)$. In the above setup we have that*

$$\forall s \geqslant 0, \quad \mathrm{P}\left[Z_v(s) \notin \left(\frac{7}{8}, \frac{10}{8}\right)\right] \leqslant 2n^{-4}. \tag{4.17}$$

$$\forall s \geqslant 0, \quad \mathrm{P}[Y_v(s) \geqslant \log n] \leqslant n^{-\log \log n}. \tag{4.18}$$

Proof: Let $v \in V$ and $s \geqslant 0$ We first prove (4.17). Let $\lambda \in [-2^9 \log n, 2^9 \log n]$ to be determined later. By Fact 3.2 and (3.5) we have that

$$\mathbb{E}[\exp(\lambda Z_v(s))] = \mathbb{E} \prod_{u \in V} \exp(\lambda Y_u(s) P^{t_*(G)}(u, v))$$
$$\leqslant \prod_{u \in V} \mathbb{E}[\exp(\lambda Y_u(s) P^{t_*(G)}(u, v))]. \tag{4.19}$$

Note that $\exp(\lambda Y_u(s) P^{t_*(G)}(u, v)) = \prod_{y \in V} \exp(\lambda \xi_y(u, s) P^{t_*(G)}(u, v))]$, where $\xi_y(u, s)$ is the indicator of the event that the walker which initially occupied y is at u at time s. Then

$$\mathbb{E}[\exp(\lambda Y_u(s) P^{t_*(G)}(u, v))] = \prod_{y \in V} \mathbb{E}[\exp(\lambda \xi_y(u, s) P^{t_*(G)}(u, v))]. \tag{4.20}$$

Let $p_* := \max\{P^{t_*(G)}(x, y) : x, y \in V\} \leqslant \frac{1}{2^9 \log n}$. As $|\lambda|p_* \leqslant 1$ and $e^x - 1 \leqslant x + \frac{x^2}{1 + 1_{x \leqslant 0}}$ for $x \in [-1, 1]$ we get that for all $y, u \in V$ we have that

$$\mathbb{E}[\exp(\lambda \xi_y(u, s) P^{t_*(G)}(u, v))] - 1 \leqslant P^s(y, u)\left(e^{\lambda P^{t_*(G)}(u,v)} - 1\right)$$
$$\leqslant P^s(y, u)\left[\lambda P^{t_*(G)}(u, v) + \frac{\left(\lambda P^{t_*(G)}(u, v)\right)^2}{1 + 1_{\lambda \leqslant 0}}\right]$$
$$\leqslant P^s(y, u)\lambda P^{t_*(G)}(u, v)\left[1 + \frac{\lambda p_*}{1 + 1_{\lambda \leqslant 0}}\right].$$

Using $1 + x \leqslant e^x$, substituting the above in (4.20) and using (4.19), together with the fact that the uniform distribution is stationary (and thus $\sum_y P^s(y, u) = 1 = \sum_u P^{t_*(G)}(u, v)$) we get that

$$\mathbb{E}[\exp(\lambda Z_v(s))] \leqslant \exp\left(\lambda + \frac{\lambda^2 p_*}{1 + 1_{\lambda \leqslant 0}}\right).$$

Taking $\lambda = -\frac{1}{8 p_*}$ and noting that $\mathrm{P}[Z_v(s) \leqslant \frac{7}{8}] = \mathrm{P}\left[\exp(\lambda Z_v(s)) \geqslant e^{\frac{7\lambda}{8}}\right]$ we see that

$$\mathrm{P}[Z_v(s) \leqslant \frac{7}{8}] = \ \leqslant \mathbb{E}[\exp(\lambda Z_v(s))] e^{-\frac{7\lambda}{8}} \leqslant \exp\left(\lambda/8 + \frac{\lambda^2 p_*}{2}\right) = \exp\left(-\frac{1}{2^7 p_*}\right) \leqslant n^{-4}.$$

Taking $\lambda = \frac{1}{8 p_*}$ and noting that $\mathrm{P}\left[Z_v(s) \geqslant \frac{10}{8}\right] = \mathrm{P}\left[\exp(\lambda Z_v(s)) \geqslant e^{\frac{10\lambda}{8}}\right]$ yields

$$\mathrm{P}\left[Z_v(s) \geqslant \frac{10}{8}\right] \leqslant \mathbb{E}[\exp(\lambda Z_v(s))] e^{-\frac{10\lambda}{8}} \leqslant \exp(-\lambda/4 + \lambda^2 p_*) = \exp\left(-\frac{1}{2^6 p_*}\right) \leqslant n^{-8}.$$

We now prove (4.18). It suffices to prove the following. Let $\xi_1, \xi_2, \ldots, \xi_m$ be independent Bernoulli r.v.'s. Let $p_i := \mathbb{E}[\xi_i]$. Assume that $\sum_i p_i = 1$. Let $L > 0$ and $\lambda = \log L$. Then

$$\mathrm{P}[S > 1 + L] \leqslant L^{-(L+1)},$$

where $S := \sum_i \xi_i$. Indeed $\mathbb{E}[e^{\lambda \xi_i}] = 1 + p_i(e^\lambda - 1) \leqslant \exp(p_i(e^\lambda - 1))$. As $\mathbb{E}[e^{\lambda S}] = \prod_i \mathbb{E}[e^{\lambda \xi_i}]$,

$$\mathrm{P}[S > 1 + L] \leqslant \mathbb{E}[e^{\lambda S}] e^{-\lambda(1+L)} = \exp(e^\lambda - 1) e^{-\lambda(1+L)} \leqslant \exp(-(L+1)\log L). \qquad \square$$

5 General lower bounds on the social connectivity time

Clearly we can bound SC from below by the minimal time by which every walker has met at least one other walker. This motivates the following definition.

Definition 5.1. *We say that a walker remained **isolated** for time t if this walker has not met any other walkers up to and including time t. We say that a walker is **lazy** by time t if this walker has not left her initial position by time t. We denote the event that there is an isolated lazy walker at vertex v by time t by $\mathrm{IL}_v(t)$.*

Recall that by elementary properties of the Poisson distribution, conditioned on the total number of walkers being m, each of them independently starts at a random initial position chosen according to the stationary distribution. We denote the corresponding probability and expectation for the aforementioned initial condition by $\mathbb{P}^{(m)}$ and $\mathbb{E}^{(m)}$. In what comes, it will sometimes be convenient to condition on the total number of walkers.

Proposition 5.2. *Let $G = (V, E)$ be a finite Eulerian digraph of average degree d and minimal degree $d_{\min}$. Let $m \in \mathbb{N}$. Let $Z(t) := \sum_{v \in V} \mathrm{IL}_v(t)$ be the number of isolated lazy walkers at time t. Let $Y(t)$ be the number of walkers which remained isolated for time t. Then*

$$\mathbb{P}^{(m)}[Y(t) = 0] \leqslant \exp\left(-\frac{2(\mathbb{E}^{(m)}[Y(t)])^2}{m(t+2)^2}\right). \tag{5.1}$$

$$\mathbb{E}^{(m)}[Z(t)] \geqslant m\pi\{v : \bar{\pi}_v \leqslant 2\} 2^{-t}\left(1 - \frac{2(t+1)}{|V|}\right)^{m-1} \geqslant \frac{d_{\min}}{2d} m 2^{-t}\left(1 - \frac{2(t+1)}{|V|}\right)^{m-1}. \tag{5.2}$$

In additional, if G is a regular graph then

$$\mathbb{E}^{(m)}[Y(t)] \geqslant m\left(1 - \frac{(2t+1)/|V|}{\min_{v \in V} \sum_{i=0}^{t} P^{2i}(v, v)}\right)^{m-1}. \tag{5.3}$$

Proof. Let $f(\mathbf{x}_1, \ldots, \mathbf{x}_m)$ be the number of walkers which remained isolated for time t when for all $i \in [m]$ the i-th walker performs the walk $\mathbf{x}_i$. Let $\mathbf{x} := (\mathbf{x}_1, \ldots, \mathbf{x}_m)$ and let $\mathbf{x}'$ be $\mathbf{x}$ with its ith co-ordinate replaced by some other walk $\mathbf{x}'_i$. It is easy to see that $|f(\mathbf{x}) - f(\mathbf{x}')| \leqslant t + 2$. In other words, given the walks performed by the rest of the walkers, the walk performed by the i-th walker can change the number of walkers which remained isolated for time t by an additive term whose absolute value is at most $t + 2$. Thus (5.1) follows by the McDiarmid's inequality (3.4).

Let $A := \{v : \bar{\pi}_v \leqslant 2\}$. The second inequality in (5.2) follows by noting that $|A| \geqslant |V|/2$ and so $\pi(A) \geqslant |A| \min_{v \in V} \pi_v \geqslant \frac{d_{\min}}{2d}$. We now prove the first inequality in (5.2). As each of the m walkers w.p. $\pi\{v : \bar{\pi}_v \leqslant 2\}$ has its initial position in A, it suffices to show that for each $v \in A$ the conditional probability that $\mathrm{IL}_v(t) = 1$ given that the first walker's initial position is v is at least $2^{-t}(1 - \frac{2(t+1)}{|V|})^{m-1}$. Indeed this holds as the probability of the first walker staying put up to time t is 2^{-t}, while the probability that no other walker visits v by time t is

$$(1 - \mathbb{P}_\pi[T_v \leqslant t])^{m-1} \geqslant (1 - \mathbb{E}_\pi[|\{i \leqslant t : X_i = v\}|])^{m-1} = \left(1 - \frac{2(t+1)}{|V|}\right)^{m-1},$$

where $(X_i)_{i \geqslant 0}$ is a LSRW with $X_0 \sim \pi$ and $\mathbb{P}_\pi$ and $\mathbb{E}_\pi$ are the corresponding probability and expectation (where we have used $\pi_v \leqslant 2/|V|$ since $v \in A$).

We now prove (5.3). Let $\mathbf{X} := (X_i)_{i \geqslant 0}$ and $(\widehat{X}_i)_{i \geqslant 0}$ be independent LSRWs with $X_0 \sim \pi$ and $\widehat{X}_0 \sim \pi$. Let $N(s) := |\{i \leqslant s : X_i = \widehat{X}_i\}|$ be the number of times the two walks collide by time s. Since conditioned on the walk that a certain walker performs, the events that each of the other $m - 1$ walkers does not collide with it by time t are independent with the same probability (used in the first equality in (5.4)), by Jensen's inequality we get that

$$\begin{aligned} \mathbb{E}^{(m)}[Y(t)] &= m\mathbb{E}[\mathbb{E}[1_{N(t)=0} \mid \mathbf{X}]^{m-1}] \\ &\geqslant m(\mathbb{E}[\mathbb{E}[1_{N(t)=0} \mid \mathbf{X}]])^{m-1} = m\mathbb{P}[N(t) = 0]^{m-1}. \end{aligned} \quad (5.4)$$

Now

$$\begin{aligned} \mathbb{P}[N(t) \geqslant 1] &\leqslant \frac{\mathbb{E}[N(2t)]}{\mathbb{E}[N(2t) \mid N(t) \geqslant 1]} \\ &\leqslant \frac{(2t+1)/|V|}{\min_{v \in V, i \leqslant t} \mathbb{E}[N(2t) - N(i-1) \mid X_i = \widehat{X}_i = v]} \leqslant \frac{(2t+1)/|V|}{\min_{v \in V} \sum_{j=0}^{t} P^{2j}(v,v)}, \end{aligned}$$

where we have used regularity to argue that π is the uniform distribution and hence $\mathbb{E}[N(2t)] = \sum_{i=0}^{2t} \mathbb{P}[X_i = \widehat{X}_i] = (2t+1) \sum_{v \in V} \pi_v^2 = (2t+1)/|V|$, and reversibility to argue that $\mathbb{E}[N(2t) - N(i-1) \mid X_i = \widehat{X}_i = v] = \sum_{j=0}^{2t-i} \sum_u P^j(v,u)P^j(v,u) = \sum_{j=0}^{2t-i} P^{2j}(v,v)$, as $\sum_u P^j(v,u)P^j(v,u) = \sum_u P^j(v,u)P^j(u,v) = P^{2j}(v,v)$. $\qquad \square$

Equations (2.2) and (2.4) follow from the following corollary, in conjunction with the estimate $\mathbb{P}[\mathrm{Pois}(n) \notin [\frac{n}{2}, 2n]] \leqslant 2e^{-n/16}$, which follows from (3.2).

Corollary 5.3. *Let $G = (V, E)$ be a connected n-vertex Eulerian digraph of average degree d and minimal degree $d_{\min}$. Let $m \in [\frac{n}{2}, 2n]$. Let*

$$\kappa_t := \min_{v \in V} \sum_{i=0}^{t} P^{2i}(v,v) \quad \text{and} \quad s_\alpha := \min\left\{t : \frac{t+1}{\kappa_{t+1}} > \frac{\alpha}{16} \log n\right\}.$$

Let $Y(t)$ be the number of walkers which remained isolated for time t. If $t \leqslant \frac{1}{36} \log n$ then

$$\mathbb{P}^{(m)}[Y(t) = 0] \leqslant \exp\left[-cn\left(\frac{d_{\min}}{d}\right)^2 e^{-20t}\right]. \quad (5.5)$$

In additional, if G is a regular graph then for all $\alpha \in (0,1)$ we have that

$$\mathbb{P}^{(m)}[Y(s_\alpha) = 0] \geqslant \exp\left(-cn^{1-\alpha}(\log n)^{-4}\right). \tag{5.6}$$

Finally, for the usual Poisson setup we have that

$$\mathrm{P}[Y(t) = 0] \leqslant \max_{m \in [\frac{n}{2}, 2n]} \mathbb{P}^{(m)}[Y(t) = 0] + \mathbb{P}\left[\mathrm{Pois}(n) \notin \left[\frac{n}{2}, 2n\right]\right]. \tag{5.7}$$

Proof. Equation (5.5) follows from (5.1) and (5.2) together with a little algebra. We leave the details as an exercise. Similarly, equation (5.6) follows from (5.1) and (5.3) together with a little algebra, after noting that by Fact 4.8 $s_\alpha \leqslant s_1 \leqslant C(\log n)^2$, for all $\alpha \in (0,1]$. Finally (5.7) follows by conditioning on the total number of walkers in the l.h.s.. $\qquad \square$

Recall the notation from §3 (in particular that $\bar{\pi} := |V|\pi$ that $p((\gamma_0, \cdots, \gamma_t)) := \prod_{i=1}^{t} P(\gamma_{i-1}, \gamma_i)$ and that $q(\gamma) := \bar{\pi}_{\gamma_0} p(\gamma)$ and that Γ_t is the collection of all walks of length t). The following proposition follows straightforwardly from Fact 3.1 (about Poisson thinning).

Proposition 5.4. *Let $G = (V, E)$ be a finite connected Eulerian digraph. Let $\gamma = (\gamma_0, \ldots, \gamma_t) \in \Gamma_t$. Then, given that $X_\gamma = 1$, the conditional distribution of the number of walkers that the walker who performed the path γ met by time t is $\mathrm{Poisson}(a_\gamma)$, where*

$$a_\gamma := \sum_{\gamma' \in \Gamma_t : \gamma' \neq \gamma, \exists i, \gamma_i = \gamma_i'} q(\gamma') \leqslant -q(\gamma) + \sum_{i=0}^{t} \bar{\pi}_{\gamma_i}.$$

In particular,

$$\begin{aligned} \forall v, \quad &\mathrm{P}[\mathrm{IL}_v(t)] \geqslant 2^{-t} \bar{\pi}_v e^{-(t+1)\bar{\pi}_v}, \\ \forall v \neq u, \quad &\mathrm{P}[\mathrm{IL}_v(t) \cap \mathrm{IL}_u(t)] = \mathrm{P}[\mathrm{IL}_v(t)]\mathrm{P}[\mathrm{IL}_u(t)]e^{m_{u,v}(t)}, \end{aligned} \tag{5.8}$$

where

$$m_{u,v}(t+1) := \sum_{\gamma \in \Gamma_{t+1} : u,v \in \gamma, \gamma_0 \notin \{v,u\}} q(\gamma) \leqslant t\bar{\pi}_v \mathbb{P}_v[T_u \leqslant t] + t\bar{\pi}_u \mathbb{P}_u[T_v \leqslant t].$$

.

Proof: For the inequality, note that the probability that there is one walker at v at time 0 and that this walker is lazy by time t is $\bar{\pi}_v 2^{-t} e^{-\bar{\pi}_v}$. The conditional probability that no other walker visited v by time t is $\mathbb{P}[\mathrm{Pois}(\sum_{\gamma \in \Gamma_t : v \in \gamma, \gamma_0 \neq v} q(\gamma)) = 0] \geqslant \mathbb{P}[\mathrm{Pois}(t\bar{\pi}_v) = 0] \geqslant e^{-t\bar{\pi}_v}$.

We now prove the second line of (5.8). Let

$$\Gamma_1 := \{\gamma \in \Gamma_t : v, u \in \gamma, \gamma_0 \notin \{v, u\}\},$$

$$\Gamma_2 := \{\gamma \in \Gamma_t : v \in \gamma, u \notin \gamma, \gamma_0 \neq v\} \quad \text{and}$$

$$\Gamma_3 := \{\gamma \in \Gamma_t : u \in \gamma, v \notin \gamma, \gamma_0 \neq u\}.$$

Denote $b_i := \sum_{\gamma \in \Gamma_i} q(\gamma)$ (where $i = 1, 2, 3$). Then $b_1 = m_{u,v}(t)$ and by the above reasoning

$$\mathrm{P}[\mathrm{IL}_v(t)] = 2^{-t} \bar{\pi}_v e^{-\bar{\pi}_v} e^{-(b_1+b_2)}, \quad \mathrm{P}[\mathrm{IL}_u(t)] = 2^{-t} \bar{\pi}_u e^{-\bar{\pi}_u} e^{-(b_1+b_3)}.$$

$$\mathrm{P}[\mathrm{IL}_v(t) \cap \mathrm{IL}_u(t)] = (2^{-t} \bar{\pi}_v e^{-\bar{\pi}_v})(2^{-t} \bar{\pi}_u e^{-\bar{\pi}_u}) e^{-(b_1+b_2+b_3)} = \mathrm{P}[\mathrm{IL}_v(t)]\mathrm{P}[\mathrm{IL}_u(t)]e^{b_1}. \qquad \square$$

Theorem 5.5. *Let G be an n-vertex connected graph. Let $\mu := \sum_{v:\bar\pi_v \leqslant 2} \bar\pi_v$ (recall that $\bar\pi = n\pi$) and $a_t := \mu 2^{-t} e^{-2(t+1)}$. Then for all $t \geqslant 0$ we have that*

$$P[\cup_v \mathrm{IL}_v(t)] \geqslant 1 - 7t/\sqrt{a_t}. \tag{5.9}$$

Hence, if there exists some $\delta > 0$ such that $\mu \geqslant c_1 n^\beta$ for some $\beta > 0$, then for $s_\beta := \max\{\lfloor \tilde c_\beta \log n - 2\log\log n\rfloor - 1, 0\}$, where $\tilde c_\beta := \frac{\beta}{2\log(2e^2)}$, we have that

$$P[\cup_v \mathrm{IL}_v(s_\beta)] \geqslant 1 - 12n^{-\beta/4}.$$

Proof. We may assume that $7t \leqslant \sqrt{a_t}$. Let $A := \{v : \bar\pi_v \leqslant 2\}$. Let $\nu_i(x,y) := \sum_{j=1}^i P^j(x,y)$. Consider

$$B_a := \{b \in A : \bar\pi_a \nu_t(a,b) \geqslant \bar\pi_b/(2\sqrt{a_t})\}.$$

For every $a \in A$ we have that $\sum_{b\in A} \nu_t(a,b) \leqslant t$ and so

$$\sum_{b\in B_a} \bar\pi_b \leqslant 2\bar\pi_a\sqrt{a_t} \sum_{b\in B_a} \nu_t(a,b) \leqslant 2\bar\pi_a t\sqrt{a_t}.$$

This (together with $7t \leqslant \sqrt{a_t}$, which implies that $t\sqrt{a_t} > 1$) implies that there exists some $D \subseteq A$ such that

$$\sum_{a\in D} \bar\pi_a \geqslant \mu/(1 + 2t\sqrt{a_t}) \geqslant \mu/(3t\sqrt{a_t}) \quad\text{and} \tag{5.10}$$

$$\bar\pi_a \nu_t(a,b) < \max\{\bar\pi_a, \bar\pi_b\}/(2\sqrt{a_t}), \text{ for all distinct } a, b \in D. \tag{5.11}$$

By (5.8) and (5.10), $\mathbb{E}[Z] \geqslant (a_t/\mu)\sum_{a\in D}\bar\pi_a \geqslant \sqrt{a_t}/(3t)$, where $Z := \sum_{a\in D} Z_a$ and $Z_a := 1_{\mathrm{IL}_a(t)}$. For all distinct $a, b \in D$, we have that $2\bar\pi_a t\nu_t(a,b) \leqslant 4t/\sqrt{a_t} \leqslant 4/7$. Thus by (5.11)

$$\exp(2\bar\pi_a t\mathbb{P}_a[T_b \leqslant t]) - 1 \leqslant 4\bar\pi_a t\mathbb{P}_a[T_b \leqslant t] \leqslant 2\max\{\bar\pi_a, \bar\pi_b\}t/\sqrt{a_t} \leqslant 4t/\sqrt{a_t},$$

and so by (5.8)

$$\mathrm{Cov}(Z_a, Z_b)/(\mathbb{E}[Z_a]\mathbb{E}[Z_b]) \leqslant \exp(2\bar\pi_a t\mathbb{P}_a[T_b \leqslant t]) - 1 \leqslant 4t/\sqrt{a_t}.$$

Hence $L := \sum_{a\neq b, a,b\in D} \mathrm{Cov}(Z_a, Z_b) \leqslant (\mathbb{E}[Z])^2(4t/\sqrt{a_t})$. Finally, since $\mathrm{Var} Z = \mathbb{E}[Z] + L$ by Chebyshev's inequality we have that $P[Z = 0] \leqslant \frac{\mathrm{Var} Z}{(\mathbb{E}[Z])^2} \leqslant \frac{1}{\mathbb{E}[Z]} + 4t/\sqrt{a_t} \leqslant 7t/\sqrt{a_t}$. $\square$

The next example demonstrates that the assertion of Theorem 5.5 is quite sharp.

Example 5.6. Denote $L_n := \lceil \log^{10} n\rceil$. Consider a $\lceil n/L_n\rceil$-clique, such that each vertex in the clique is the center of a star of size L_n (where all stars are disjoint). Then one can show that $\mathrm{SC} \leqslant C\log\log n$ w.h.p. (and if the walkers perform non-lazy SRWs we would get that $\mathrm{SC} \leqslant C$ w.h.p.). The reason for this is that the expected number of walkers whose initial position is not in a center of a star is at most L_n^2, while at each vertex of the clique there are $\mathrm{Pois}(\mu_n)$ walkers for some $\mu_n = L_n(1 \pm o(1))$. An easy calculation shows that w.h.p. after one step all of the walkers whose initial positions are in the clique are in the same class.

Proof of Theorem 2: Equation (2.4) follows from Corollary 5.3. We now prove (2.3). Let $G = (V, E)$ be a connected n-vertex d-regular graph. Consider the SN model on G in which the walks have some fixed holding probability $p \in [0, 1)$. Denote the transition matrix of the corresponding walks by P. Recall that Γ_t is the collection of all walks on G of length t and that for $\gamma = (\gamma_0, \ldots, \gamma_t) \in \Gamma_t$, $p(\gamma) := \prod_{i=0}^{t-1} P^i(\gamma_i, \gamma_{i+1})$. Fix

$t = t_n = \lfloor \log n - 6 \log \log n \rfloor - 1$. Assume $t \geqslant 0$ as otherwise there is nothing to prove. For every $\Gamma' \subseteq \Gamma_t$ let $m(\Gamma') := \sum_{\gamma \in \Gamma'} p(\gamma)$. For every $\gamma \in \Gamma_t$ and $0 \leqslant i \leqslant t$ let

$$B_{\gamma,i} := \left\{ v \in V : \sum_{j=0}^{t} P^{|i-j|}(\gamma_j, v) \geqslant (t+1)^{-3} \right\},$$

$$\mathrm{Bad}_\gamma := \{ \tilde{\gamma} \in \Gamma_t : \exists i \in \{0, 1, \ldots, t\} \text{ such that } \tilde{\gamma}_i \in B_{\gamma,i} \}.$$

Clearly $|B_{\gamma,i}| \leqslant (t+1)^4$ for all $\gamma \in \Gamma_t$ and $i \in \{0, 1, \ldots, t\}$, and thus

$$m(\mathrm{Bad}_\gamma) \leqslant \sum_{i \leqslant t} \sum_{v \in B_{\gamma,i}} m(\{\tilde{\gamma} \in \Gamma_t : \tilde{\gamma}_i = v\}) \leqslant \sum_{i \leqslant t} |B_{\gamma,i}| \leqslant (t+1)^5. \tag{5.12}$$

Fix some order $\prec$ on Γ_t. Recall that X_γ denotes the number of walkers who performed the walk γ. We sequentially expose X_γ for some paths $\gamma \in \Gamma_t$ according to the following procedure. Assuming that we have already exposed $A \subset \Gamma_t$ so that the following holds: for every $\gamma \in A$ with $X_\gamma > 0$ we have that $\mathrm{Bad}_\gamma \cap \{\gamma' \in A : \gamma \prec \gamma'\} = \{\gamma\}$ (where $\gamma \prec \gamma'$ indicates that γ' is larger in the ordering than γ). In the next stage we expose $X_{\gamma'}$ for the minimal $\gamma' \in \Gamma_t \setminus B(A)$, where $B(A) := \cup_{\gamma \in A : X_\gamma > 0} \mathrm{Bad}_\gamma$. In the following stage we apply the same rule, with the set A replaced by the set $A \cup \{\gamma'\}$. At the end of this procedure we obtain a collection $\mathcal{W} = \{\gamma^1, \ldots, \gamma^{|\mathcal{W}|}\}$ (resp. $\mathcal{N}$) of all γ's for which we exposed that $X_\gamma > 0$ (resp. $= 0$), where the indices are taken so that $i < j$ iff γ^i is before γ^j in the ordering. The collection of all $\gamma \in \Gamma_t$ for which X_γ was not exposed is precisely $\cup_{\gamma \in \mathcal{W} : X_\gamma > 0} \mathrm{Bad}_\gamma \setminus \{\gamma\} = \Gamma_t \setminus (\mathcal{W} \cup \mathcal{N})$.

Let $\mathcal{N}_i$ be the collection of all $\gamma \in \mathcal{N}$ which were exposed in between γ^{i-1} and γ^i (and $\mathcal{N}_1$ is the collection of all γ's exposed prior to γ^1). For every $1 \leqslant i \leqslant |\mathcal{W}|$, the probability that $m(\mathcal{N}_i) > (t+1)^5$ is at most $\mathrm{P}[\mathrm{Pois}((t+1)^5) = 0] = e^{-(t+1)^5}$. Hence by (5.12)

$$\mathrm{P}[|\mathcal{W}| < n/(2(t+1)^5)] \leqslant n e^{-(t+1)^5} = o(1). \tag{5.13}$$

We now condition on $\mathcal{W} = W$ and $\mathcal{N} = N$ for some arbitrary $W, N \subset \Gamma_t$ so that (i) $|\mathcal{W}| \geqslant n/[2(t+1)^5]$ and (ii) $\mathrm{Bad}_\gamma \cap \{\gamma' \in W : \gamma \prec \gamma'\} = \{\gamma\}$, for all $\gamma \in W$. Observe that, given $(\mathcal{W}, \mathcal{N}) = (W, N)$, the joint distribution of $(X_\gamma)_{\gamma \in \Gamma_t \setminus (W \cup N)}$ is the same as their unconditional joint distribution. For $\gamma = (\gamma_0, \ldots, \gamma_t), \gamma' = (\gamma'_0, \ldots, \gamma'_t) \in W$ let

$$\Gamma_\gamma := \{\tilde{\gamma} \in \Gamma_t \setminus (N \cup \{\gamma\}) : \exists i \in \{0, 1, \ldots, t\} \text{ such that } \tilde{\gamma}_i = \gamma_i\} \subseteq \mathrm{Bad}_\gamma,$$

$$\Gamma_{\gamma,\gamma'} := \{\tilde{\gamma} \in \Gamma_t : \exists i, j \in \{0, 1, \ldots, t\} \text{ such that } \tilde{\gamma}_i = \gamma_i \text{ and } \tilde{\gamma}_j = \gamma'_j\} \subseteq \mathrm{Bad}_\gamma \cap \mathrm{Bad}_{\gamma'}.$$

Note that $m_\gamma := m(\Gamma_\gamma) \leqslant \sum_{i=0}^{t} m(\{\tilde{\gamma} \in \Gamma_t : \tilde{\gamma}_i = \gamma_i\}) = t + 1$. Denote

$$D_{i,j}(\gamma, \gamma') := \{\tilde{\gamma} \in \Gamma_t : \tilde{\gamma}_i = \gamma_i, \tilde{\gamma}_j = \gamma'_j, \forall k < i \ \tilde{\gamma}_k \neq \gamma_k, \forall r < j \ \tilde{\gamma}_r \neq \gamma'_r\}\}$$

By construction $\gamma' \notin \mathrm{Bad}_\gamma$, for all $\gamma, \gamma' \in W$ such that $\gamma \prec \gamma'$ and so we have that $\sum_{j=0}^{t} P^{|i-j|}(\gamma_j, \gamma'_i) < (t+1)^{-3}$ for all $0 \leqslant i \leqslant t$. Hence by reversibility

$$m_{\gamma,\gamma'} := m(\Gamma_{\gamma,\gamma'}) = \sum_{0 \leqslant i < j \leqslant t} m(D_{i,j}(\gamma, \gamma')) + m(D_{i,j}(\gamma', \gamma))$$

$$\leqslant \sum_{0 \leqslant i,j \leqslant t} P^{|i-j|}(\gamma_i, \gamma'_j) \leqslant (t+1)^{-2}.$$

Finally, for each $\gamma \in W$ set Z_γ to be the indicator of the event that $\sum_{\gamma \in \Gamma_\gamma} X_\gamma = 0$. Set $Z = \sum_{\gamma \in W} Z_\gamma$. Observe that (using $m_\gamma \leqslant t + 1$ for all $\gamma \in W$)

$$\mathbb{E}[Z] \geqslant |W| e^{-(t+1)} > \frac{n}{2(t+1)^5} \left(n^{-1} \log^6 n \right) > \tfrac{1}{2} \log n$$

while similarly to (5.8) for all $\gamma, \gamma' \in W$ we have that

$$\mathrm{Cov}(Z_\gamma, Z_{\gamma'})/(\mathbb{E}[Z_\gamma]\mathbb{E}[Z_{\gamma'}]) = e^{m_{\gamma,\gamma'}} - 1 \leqslant e^{1/(t+1)^2} - 1 =: \delta_n.$$

Hence

$$\mathrm{Var}(Z) \leqslant \mathbb{E}[Z] + \delta_n(\mathbb{E}[Z])^2,$$

which by Chebyshev implies that $\mathrm{P}[Z = 0] < \frac{1}{\mathbb{E}[Z]} + \delta_n$. $\qquad\square$

6 The cycle

In this section we consider the case that G is the n-cycle C_n. In this section we assume that at time 0 there is precisely one walker at each site and that the walkers perform independent continuous-time random walks with jump rate 1.

We first prove (2.6). We fix $s = s_n := C_1 \log^2 n$. By a union bound over the n edges of the cycle, in order to deduce that $\mathrm{P}[\mathrm{SC}(\mathrm{C}_n) \geqslant s] \leqslant C/n$ it suffices to show that for every pair of neighboring vertices u, v, the probability that the walkers who started at those vertices do not have a path of acquaintances by time s is at most Cn^{-2}.

Consider some fixed edge $e = \{u, v\}$. It partitions the vertices into 2 sides (segments) A_u, A_v of size roughly $n/2$, according to the identity of the end-point to which they are closer to (break a tie arbitrarily). Let e' be the other edge, apart from e, which connects A_u and A_v. A walker starting from either u or v is extremely unlikely to cross e' in s time units.

Let $\mathrm{CROSS}_s(e, w)$ be the event that the walker who started at vertex w did not cross e' by time s and that her position at time s is at the opposite side of e compared to w.

We now argue that for every $(w, w') \in A_u \times A_v$ the event $\mathrm{CROSS}_s(e, w) \cap \mathrm{CROSS}_s(e, w')$ "forces" the walkers who started at u and v to have a path of acquaintances by time s (which uses only walkers whose starting positions are in $\{u, v, w, w'\}$, possibly a subset of this set). Indeed, the continuous-time setup eliminates the possibility of two walkers swapping positions without meeting. This implies that on the event $\mathrm{CROSS}_s(e, w) \cap \mathrm{CROSS}_s(e, w')$ it must be the case that by time s the walkers whose starting positions are w and w', resp., have met, and that each of the walkers whose starting positions are u and v, resp., must have met at least one of the previous pair of walkers (in fact, this is the case for any pair of vertices u', v' lying in the segment between w and w' which contains e).

This reduces the proof of (2.6) into the following simple calculation: If C_1 is taken to be sufficiently large then $\mathrm{P}[\mathrm{CROSS}_s(e, w)] \geqslant 0.4$ for every edge $e = \{u, v\}$ and vertex w of distance at most $10 \log n$ from e. Indeed, starting from such a vertex w, if C_1 is sufficiently large then the probability of crossing e by time s is arbitrarily close to $1/2$ (in fact it is arbitrarily close to 1), and for all sufficiently large n the probability that e' is crossed by time s is at most $1/20$ (this probability tends to 0 as n to infinity). Hence (using $\left(\frac{12}{20}\right)^{10 \log n} \leqslant n^{-2}$)

$$\mathrm{P}[\mathrm{CROSS}_s(e)] \geqslant 1 - Cn^{-2},$$
$$\text{where } \mathrm{CROSS}_s(e) := \bigcup_{w \in A_u, w' \in A_v} \mathrm{CROSS}_s(e, w) \cap \mathrm{CROSS}_s(e, w').$$

This concludes the proof of (2.6). We now prove (2.7). We fix $t = t_n = c_1 \log^2 n$ for some constant c_1 to be determined later. Observe that if there are at least 2 edges which no walker crossed by time t, then deterministically $\mathrm{SC}(\mathrm{C}_n) > t$. Let $\mathrm{NC}_t(e)$ be the event that the edge $e \in E(\mathrm{C}_n)$ was not crossed by any walker by time t. We argue that if c_1 is taken to be sufficiently small, then for all sufficiently large n and every edge e it holds that

$$\mathrm{P}[\mathrm{NC}_t(e) \mid \mathrm{F}] \geqslant c_3 n^{-1/4}, \tag{6.1}$$

where F is the event that no walker performed at least $(\sqrt{n}/2) - 2$ steps by time t. We now explain how (2.7) can be derived from (6.1). Denote $m := \lfloor \sqrt{n} \rfloor$. Let $(e_i)_{i=1}^m$ be a collection of edges which are all at distance at least $\sqrt{n}$ from one another. Then (by a union bound over the n walkers) $1 - \mathrm{P}[\mathrm{F}] \leqslant C_2 n e^{-cn^{1/4}} \leqslant C_2' e^{-c'n^{1/5}}$. Finally, note that conditioned on F, the events $(\mathrm{NC}_t(e_i))_{i=1}^m$ become independent, and so by (6.1) the conditional probability that none of them occur is at most $(1 - c_3 n^{-1/4})^m \leqslant e^{-c_3' n^{1/4}}$.

We now prove (6.1). It is not hard to see that in order to prove prove (6.1) it suffices to show that for every k if we put one particle at each site of $\mathbb{N} := \{1, 2 \ldots\}$ and the particles perform independent continuous-time SRW on $\mathbb{Z}$ with jump rate 1, then the probability that no walker reached the origin by time k is at least $e^{-M\sqrt{k}}$, for some absolute constant M. Denote the probability of the previous event by p_k. Let $(S_r^i)_{r \geqslant 0}$ be a SRW on $\mathbb{Z}$ starting at $i \in \mathbb{N}$. Then by the reflection principle (see e.g. [16, §2.7]; the proof in continuous-time is analogous)

$$a_i(k) := \mathbb{P}\left[S_r^i > 0, \, \forall r \in [0, k]\right] = \mathbb{P}\left[S_k^0 \in \{-i+1, \ldots, i\}\right], \quad \text{for all } i \in \mathbb{N}. \tag{6.2}$$

Write $b := \lfloor 4\sqrt{k} \rfloor$. We argue that there exist constants $c_5, c_6, C_3 > 0$ such that

$$a_i(k) \geqslant c_5 i/b, \quad \text{for all } 1 \leqslant i \leqslant b. \tag{6.3}$$

$$
\begin{aligned}
a_i(k) &\geqslant 1 - 2\exp\left[-\frac{i^2}{k} + \frac{i^4}{4k^3}\right] \geqslant 1 - 2\exp\left(-\frac{i^2}{8k}\right) \\
&\geqslant \exp\left(-C_3 e^{-i^2/(8k)}\right), \text{ for } b < i \leqslant 1.2k. \\
a_i(k) &\geqslant 1 - e^{-c_6 i} \geqslant \exp\left(-C_3 e^{-c_6 i}\right), \quad \text{for all } i > 1.2k.
\end{aligned}
\tag{6.4}
$$

The first line follows from the local CLT. The first inequality in the third line follows from the fact that $1 - a_i(k) \leqslant \mathbb{P}[\mathrm{Pois}(k) \geqslant i]$. The last inequalities in the second and third lines follows from the fact that for every $0 < c < 1$ there exists $C > 0$ such that $1 - x \geqslant e^{-Cx}$, for all $x \leqslant c$ (applied to $x = 2e^{-i^2/(8k)}$ for $i > b$). The first inequality in the second line is obtained by noting that for discrete-time SRW $(\tilde{S}_r)_{r \in \mathbb{Z}_+}$ starting at the origin we have that $\mathbb{E}[e^{\lambda \tilde{S}_r}] = \left(\frac{1}{2}e^\lambda + \frac{1}{2}e^{-\lambda}\right)^r \leqslant e^{\lambda^2 r/2}$ (by comparing Taylor expansion coefficients). Hence,

$$\mathbb{E}[e^{\lambda S_k}] \leqslant \sum_r \mathrm{P}[\mathrm{Pois}(k) = r]e^{\lambda^2 r/2} = \exp[k(e^{\lambda^2/2} - 1)] \leqslant \exp\left[k\left(\lambda^2/2 + (\lambda^2/2)^2\right)\right],$$

as long as $\lambda^2/2 \leqslant 1$ (using $e^b - 1 \leqslant b + b^2$ for all $b \in [-1, 1]$). We set $\lambda = i/k$, so that indeed $\lambda^2/2 \leqslant 1$ for $i \leqslant 1.2k$. Finally, by (6.2), Markov's inequality and our choice of λ

$$1 - a_i(k) \leqslant 2\mathbb{P}[S_k \geqslant i] \leqslant 2\mathbb{E}[e^{\lambda S_k}]e^{-\lambda i} \leqslant 2\exp[-(i^2/2k) + i^4/(4k^3)].$$

We are now in a position to conclude the proof. By (6.3) and Stirling's approximation

$$\prod_{i \leqslant b} a_i(k) \geqslant c_5^b b!/b^b \geqslant \sqrt{b}(c_5/e)^b \geqslant e^{-C_4\sqrt{k}}.$$

Denote $b' := \lfloor \sqrt{k} \rfloor$. Finally, using (6.4) it is not hard to show that

$$\prod_{i > b} a_i(k) \geqslant \prod_{\ell:4 \leqslant \ell \leqslant 1.2b'} a_{\ell b'}^{b'}(k) \prod_{\ell > 1.2b'} a_\ell(k) \geqslant e^{-C_5\sqrt{k}}$$

(we leave the details to the reader; Alternatively, since $a_i(k)$ are uniformly bounded away from 0 for $i > b$, we get that $\prod_{i > b} a_i(k) \geqslant e^{-C\sum_{i>b}(1 - a_i(k))}$ and the reasoning in the following remark yields that $\sum_{i>0}(1 - a_i(k)) = \sum_{i>0} \mathrm{P}_i[T_0 \leqslant k] \leqslant C'\sqrt{k}$). We are done as $p_k := \prod_{i>0} a_i(k)$. $\qquad\square$

Remark 6.1. Consider the usual Poisson setup. The proof of the upper bound on $\mathrm{SC}(\mathrm{C}_n)$ is similar to the case of one walker per site at time 0. However, as we now explain, the proof of the lower bound on $\mathrm{SC}(\mathrm{C}_n)$ becomes much simpler. While this was already noted in the introduction after Theorem 2, we present below an alternative argument, as it will be used in the analysis of Examples 8.6 and 8.7.

For simplicity, we present the argument in the discrete-time setup. Let v be a vertex in C_n. Denote $\nu_t := \sum_{i=0}^t P^i(v,v)$ (this sum is independent of v). For $t \leqslant n^2$ we have that $\nu_t = \Theta(\sqrt{t+1})$. By Fact 3.1, the total number of walkers to reach vertex v by time t has a Poisson distribution with mean (by reversibility)

$$\mu_t := \sum_u \mathbb{P}_u[T_v \leqslant t] = \sum_u \sum_{j=0}^t \mathbb{P}_u[T_v = j] \leqslant \frac{1}{\nu_t} \sum_u \sum_{j=0}^t \mathbb{P}_u[T_v = j] \sum_{i=0}^{2t-j} P^i(v,v)$$

$$\leqslant \frac{C}{\sqrt{t+1}} \sum_u \sum_{i=0}^{2t} P^i(u,v) = \frac{C(2t+1)}{\sqrt{t+1}}.$$

Thus if $t + 1 < (\frac{\log n}{8C})^2$ we get that the probability that no walker reached v by time t is at least $e^{-\mu_t} \geqslant n^{-1/4}$. This implies the lower bound on $\mathrm{SC}(\mathrm{C}_n)$ as in the proof above.

7 Expanders

In this section we study the case that G is a d-regular λ-expander. It is not difficult to extend the results to the case G is an expander of maximal degree d. LSRW on a regular λ-expander $G = (V, E)$ mixes rapidly in the following sense

$$\max_{x,y \in V} |P^t(x,y) - \pi_y| \leqslant (1-\lambda)^t, \quad \text{for all } t. \tag{7.1}$$

Indeed $\max_{x \in V} |P^t(x,x) - \pi_x| \leqslant (1-\lambda)^t$ follows from the spectral decomposition of $P^t(x,x)$ along with the non-negativity of the eigenvalues of P. However, for LSRW on a regular graph $\max_{x,y \in V} |P^t(x,y) - \pi_y| = \max_{x \in V} |P^t(x,x) - \pi_x|$, since in general (cf. [8, (2.2)] for even t and [16, p. 135] for moving from even t to odd t using laziness)

$$\max_{x,y} |P^t(x,y) - \pi_y| \leqslant \sqrt{\max_{u,v}(\pi_u/\pi_v)} \max_x |P^t(x,x) - \pi_x|.$$

7.1 Proof of Theorem 4

We argue that Theorem 4 follows from Theorem 5. This follows at once from Corollary 7.2 below. Indeed, after a ("giant") class of walkers $\mathcal{A}$ of size at least $n/6$ emerges, by Corollary 7.2 (using a union bound over the walkers not in $\mathcal{A}$, together with the concentration of the total number of walkers around n), w.p. at least $1 - C_2 n^{-1}$, the additional amount of time until every walker not from $\mathcal{A}$ will meet some walker from $\mathcal{A}$ is at most $\lceil C\lambda^{-1} \log n \rceil$. $\qquad \square$

Lemma 7.1. *Let G be a connected d-regular n-vertex λ-expander. Then there exists $C > 0$ such that for $t := \lceil C\lambda^{-1} \log n \rceil$ and every $v, u_1, u_2, \ldots, u_t$ sequence of vertices, the probability that a LSRW started at v visits some u_i at time i for some $1 \leqslant i \leqslant t$ is at least $\frac{1}{8}\min(\frac{\lambda t}{n}, 1)$.*

Proof. Let $(X_s)_{s \geqslant 0}$ be a LSRW on G started at v. Let $Y_i = 1_{X_i = u_i}$ and $Y := \sum_{i=1}^t Y_i$, where $t := \lceil C\lambda^{-1} \log n \rceil$, for some constant C, to be determined shortly. By (7.1), if C is sufficiently large, $\mathbb{E}[Y] \geqslant t/(2n)$, whereas

$$\mathbb{E}[Y_i Y_{i+j}] \leqslant \mathbb{E}[Y_i] P^j(u_i, u_{i+j}) \leqslant \mathbb{E}[Y_i](n^{-1} + (1-\lambda)^j)$$

and so $\mathbb{E}[Y^2] \leqslant 2\mathbb{E}[Y](\lambda^{-1} + t/n)$. Finally,

$$\mathrm{P}[Y > 0] \geqslant \frac{(\mathbb{E}[Y])^2}{\mathbb{E}[Y^2]} \geqslant \frac{\mathbb{E}[Y]}{2\left(\gamma^{-1} + \frac{t}{n}\right)} \geqslant \frac{1}{8} \min\left\{\frac{\lambda t}{n}, 1\right\}. \qquad \square$$

Corollary 7.2. *Let $G = (V, E)$ be a connected d-regular n-vertex λ-expander. Place $\ell := \lceil n/6 \rceil$ walkers arbitrarily on G along with one extra walker. Assume all walkers perform independent LSRWs on G simultaneously. Then there exist some constants C, n_0, independent of G, d, λ and the initial locations of the walks, such that if $n \geqslant n_0$ the extra walker meets at least one of the other ℓ walkers by time $t := \lceil C\lambda^{-1} \log n \rceil$ with probability at least $1 - 1/n^2$ (as usual, meeting means visiting the same vertex at the same time).*

Proof. First, condition on the walk that the extra walker performed by time t, denoted by γ. Then use Lemma 7.1 and independence to obtain the desired estimate for the conditional probability (uniformly, for all possible values of γ). Finally, average over γ. $\square$

7.2 Proof of Theorem 5

Proof: We adapt a technique from [2, Proposition 3.1] to our setup. Let $t \leqslant n$ to be determined later. Set $s = s(t, \lambda) := \lceil 8(t+2)/\lambda \rceil$. Call a class of walkers at time s **good** if the collection of walkers from this class occupy at time s at least t vertices. By Corollary 7.4 below (if t is sufficiently large), w.p. at least $1 - \exp(-c_1 n/d^{4(s+1)})$ at time s there exists a collection containing at most $n/(2t)$ good classes such that the walkers belonging to the union of these classes occupy at least $n/2$ vertices at time s. We conditioned on this event and on the identity of the good classes from this collection and on the positions at time s of the corresponding walkers. We claim, that if t is taken to be sufficiently large, then w.p. at least $1 - \exp(-c_2 n)$ there is no way of splitting this collection into 2 (disjoint) collections, $\mathcal{A}, \mathcal{B}$, such that

 (i) The walkers belonging to the union of the classes in $\mathcal{A}$ (resp. $\mathcal{B}$), occupy at time s at least $n/6$ vertices.
 (ii) No walker from (the union of the classes in) $\mathcal{A}$ met a walker from $\mathcal{B}$ by time $s + r$, where $r := \lceil C\lambda^{-1} \rceil$ for some sufficiently large constant C.

Indeed, this follows from Lemma 7.5 below which asserts that for any choice of $\mathcal{A}, \mathcal{B}$ satisfying (i), the probability that (ii) holds is at most $e^{-c_2 n}$, by a union bound over all $\leqslant 2^{n/(2t)}$ such partitions (and picking t such that $t \geqslant \lceil 1/c_2 \rceil$; Recall that in the considered collection of good classes there are at most $n/(2t)$ classes). $\square$

Lemma 7.3. *Let $G = (V, E)$ be a connected d-regular n-vertex λ-expander. Consider the SN model on G. Let $\mathcal{W}_u(s)$ denote the walkers who occupy vertex u at time s. Write $u \Leftrightarrow_s v$ if there exists a pair of walkers $(w, w') \in \mathcal{W}_u(s) \times \mathcal{W}_v(s)$ who met each other by time s. Let $A_{u,r}(s)$ be the event that $|\{v : u \Leftrightarrow_s v\}| \geqslant r$. Denote $s := \lceil 8(t+2)/\lambda \rceil$. Then there exists an absolute constant $c > 0$ such that if $s \leqslant n$ then*

$$\mathrm{P}[A_{u,t}(s)] \geqslant (1 - e^{-1}) - e^{-ct}.$$

Proof. We have that $\mathrm{P}[|\mathcal{W}_u(s)| > 0] = 1 - e^{-1}$. Throughout the proof we condition on $|\mathcal{W}_u(s)| > 0$. Moreover, we fix some walker $w \in \mathcal{W}_u(s)$ and condition on the walk $\mathbf{w} = (\mathbf{w}(0), \ldots, \mathbf{w}(s) = u)$ that she performed by time s. By averaging, it suffices to show that the conditional probability that $A_{u,t}(s)$ fails is at most e^{-ct} for some constant $c > 0$ independent of $\mathbf{w}$. Below, all probabilities and expectations are the conditional ones, given $\mathbf{w}$.

Let $(X_k)_{k \geqslant 0}$ be a LSRW on G. We denote the law of $(X_k)_{k \geqslant 0}$ given $X_0 = v$ by $\mathbb{P}_v$. Let $\tau_{\mathbf{w}} := \inf\{k \in [0, s] : X_k = \mathbf{w}(s - k)\}$. By Fact 3.1 and reversibility, the (conditional) distribution (given $\mathbf{w}$) of the number of walkers not from $\mathcal{W}_u(s)$ that w met by time s has a Poisson distribution with mean $\mu_{\mathbf{w}} := \sum_{v:v \neq u} \mathbb{P}_v[\tau_{\mathbf{w}} \leqslant s]$. We argue that $\mu_{\mathbf{w}} \geqslant 4t$. Indeed, let $b_{v,\mathbf{w}} := \sum_{i=0}^{s} P^i(v, \mathbf{w}(s - i))$. By reversibility and (7.1) $\mu'_{\mathbf{w}} := \sum_{v:v \neq u} b_{v,\mathbf{w}} = s - b_{u,\mathbf{w}} \geqslant s - \lambda^{-1} - 1$. Again by (7.1) and reversibility, for all v

$$b_{v,\mathbf{w}} = \sum_{i \leqslant s} \mathbb{P}_v[\tau_{\mathbf{w}} = i] \sum_{j=0}^{s-i} P^j(\mathbf{w}(s-i), \mathbf{w}(s-i-j)) \leqslant \sum_{j=0}^{s-1} \max_{x,y} P^j(x, y) \mathbb{P}_v[\tau_{\mathbf{w}} \leqslant s]$$

$$\leqslant \mathbb{P}_v[\tau_{\mathbf{w}} \leqslant s](\lambda^{-1} + 1) \leqslant 2\lambda^{-1} \mathbb{P}_v[\tau_{\mathbf{w}} \leqslant s].$$

Summing over all $v \in V \setminus \{u\}$ we get that $\mu_{\mathbf{w}} \geqslant (\lambda/2)\mu'_{\mathbf{w}} \geqslant (\lambda/2)(s - \lambda^{-1} - 1) \geqslant 4t$, as desired.

Let Z_v be the number of walkers in $\mathcal{W}_v(s)$ who met w by time s. Then (conditioned on $\mathbf{w}$) $(Z_v)_{v:v \neq u}$ are independent Poisson random variables such that $\mathbb{E}[Z_v] \leqslant 1$ for all v and $\sum_{v:v \neq u} \mathbb{E}[Z_v] \geqslant 4t$. Finally, by (the elementary) Lemma 7.6 below we have that $\mathbb{E}[1_{Z_v \geqslant 1}] \geqslant (1 - e^{-1})\mathbb{E}[Z_v]$ for all v, and so $\sum_{v:v \neq u} \mathbb{E}[1_{Z_v \neq 0}] > (1 - e^{-1})4t > 2t$. By Bernstein's inequality (3.3) $\mathbb{P}[\sum_{v:v \neq u} 1_{Z_v \neq 0} \leqslant t] \leqslant e^{-ct}$, for some constant $c > 0$, as desired. $\square$

Corollary 7.4. *In the setup and notation of Lemma 7.3, there exist $t_0, c_1 > 0$ such that*

$$\mathbb{P}[\sum_{u \in V} 1_{A_{u,t}(s)} < n/2] \leqslant \exp(-c_1 n/d^{4(s+1)}), \text{ for all } t \geqslant t_0 \text{ such that } s = \lceil 8(t+2)/\lambda \rceil \leqslant n.$$

Proof. Use Lemma 7.3 together with Azuma inequality on an appropriate Doob's sequence, by exposing the value of indicators on the l.h.s. sequentially, one at a time. The absolute value of the increments are bounded by the maximal size of a ball of radius $2s$ (which is at most $d^{2(s+1)}$). $\square$

Lemma 7.5. *Let $G = (V, E)$ be a connected d-regular n-vertex λ-expander. Let $0 < \delta \leqslant 1/2$. Let $A, B \subset V$ be two disjoint sets of size at least δn. Put one walker at each vertex of $A \cup B$ and let them perform independent LSRWs. There exists an absolute constant $c > 0$ such that the probability that no walker from A met a walker from B by time $r := \lceil \lambda^{-1} \log(4/\delta^2) \rceil$ is at most $\exp[-c\delta^2 n]$.*

Proof: Let $\mathcal{W}_A, \mathcal{W}_B$ be the collection of walkers occupying A and B, respectively, at time 0 (as indicated in the statement of the lemma, at each vertex there is exactly one walker at time 0). Let D_t be the set of vertices which are occupied at time t by at least one walker in $\mathcal{W}_B$. Using Lemma 7.6 below and the fact that $\mathbb{P}[u \in D_t] = P^t(u, B) \leqslant 1$ for all u, it is easy to verify that $\mathbb{E}[|D_t|] \geqslant |B|(1 - e^{-1})$ for every t. Thus by applying Azuma inequality to an appropriate Doob's sequence (by exposing the positions at time t of the walkers in $\mathcal{W}_B$ sequentially, one at a time), we get that

$$\mathbb{P}[|D_t| < \delta n/4] \leqslant e^{-c_1 \delta n} \tag{7.2}$$

(the absolute values of the increments of this martingale are bounded by 1). Denote the uniform distribution on A by π_A and by $\mathrm{P}^r_{\pi_A}$ the distribution at time r of LSRW with $X_0 \sim \pi_A$. If some walker from $\mathcal{W}_A$ is at D_r at time r then this walker meets a walker from $\mathcal{W}_B$ at time r. Given that $D_r = D$, the conditional probability that no walker from $\mathcal{W}_A$ is at D_r at time r is at most

$$\prod_{a \in A} (1 - P^r(a, D)) \leqslant \exp\left[-\sum_{a \in A} P^r(a, D)\right] = \exp\left[-|A|\mathrm{P}^r_{\pi_A}(D)\right].$$

To conclude the proof, we show that $P_{\pi_A}^r(D) \geqslant \delta/8$ for all $D \subseteq V$ such that $|D| \geqslant \delta n/4$, which implies the assertion of the Lemma using (7.2).

Let π be the uniform distribution on V. Recall that for every $p \geqslant 1$ the ℓ_p distance between a pair μ, ν of distributions on V is defined as $\|\mu - \nu\|_{p,\pi} := [\sum_{v \in V} \pi(v)(\frac{\mu(v) - \nu(v)}{\pi(v)})^p]^{1/p}$ and $\|\mu - \nu\|_{\infty,\pi} := \max_{v \in V} |\mu(v) - \nu(v)|$. Then

$$\|\pi_A - \pi\|_{2,\pi} \leqslant \|\pi_A - \pi\|_{\infty,\pi} \leqslant n/|A|.$$

By Jensen's inequality (first inequality below) and the Poincaré inequality (second inequality below) satisfied by the spectral gap (cf. [1, Lemma 3.26])

$$\|P_{\pi_A}^r - \pi\|_{1,\pi} \leqslant \|P_{\pi_A}^r - \pi\|_{2,\pi} \leqslant (1-\lambda)^r \|\pi_A - \pi\|_{2,\pi} \leqslant \frac{n}{|A|} e^{-\lambda r} \leqslant \delta/4,$$

where the last inequality follows from the choice of r. As

$$\|P_{\pi_A}^r - \pi\|_{1,\pi} = \sum_v |P_{\pi_A}^r(v) - \pi(v)| = 2\left(\max_{D \subset V} \pi(D) - P_{\pi_A}^r(D)\right),$$

it follows that if $|D| \geqslant \delta n/4$ then

$$P_{\pi_A}^r(D) \geqslant \pi(D) - \frac{1}{2}\|P_{\pi_A}^r - \pi\|_{1,\pi} = |D|/n - \delta/8 \geqslant \delta/8. \qquad \square$$

Lemma 7.6. *For all $x \in [0,1]$, $(1 - e^{-x}) \geqslant (1 - e^{-1})x$. In particular, for all $x \in [0,1]$, if $Y \sim \mathrm{Pois}(x)$ or $Y = \sum_{i=1}^m \xi_i$, where $\xi_1, \ldots, \xi_m$ are independent Bernoulli r.v.'s with $\sum_{i=1}^m \mathbb{E}[\xi_i] = x$, then (by Lemma 4.13) $P[Y \geqslant 1] \geqslant 1 - e^{-\mathbb{E}[Y]} \geqslant (1 - e^{-1})\mathbb{E}[Y]$.*

8 Concluding remarks, conjectures and open problems

8.1 Allowing different communities of walkers

It is not hard to see that the proof of (2.1) can also be extended to the following setup in which there are m "communities" of walkers. Let $m \leqslant n$. Let $k_1, \ldots, k_m > 0$ be such that $\sum_{i=1}^m k_i = n$. Let $G_1, \ldots, G_m$ be connected graphs on the same vertex set V of size n. Denote the stationary distribution of LSRW on G_i by π^i. At time 0 there are $\mathrm{Pois}(k_i \pi_v^i)$ walkers at vertex v performing independent LSRWs on G_i (for all $1 \leqslant i \leqslant m$ and $v \in V$, independently). As usual, when two walkers reach the same vertex at the same time they become acquainted, even if they walk on different graphs. The bound on SC depends in this case on $\max_{i \leqslant m} d^{(i)}$, where $d^{(i)}$ is the average degree of G_i.

8.2 Dependence on the maximal degree and the vertex-transitive setup

Definition 8.1. *We say that a bijection $\varphi : V \to V$ is an automorphism of a graph $G = (V, E)$ if $\{u, v\} \in E$ iff $\{\varphi(u), \varphi(v)\} \in E$. A graph G is said to be* vertex-transitive *if the action of its automorphisms group $\mathrm{Aut}(G)$ on its vertices is transitive (i.e. $\{\varphi(v) : \varphi \in \mathrm{Aut}(G)\} = V$ for all v).*

Conjecture 8.2. *There exists some $C > 0$ such that for every connected n-vertex graph G of average degree d, with probability at least $1 - C/n$ we have that*

$$\frac{\mathrm{SC}(G)}{(\log n)^2} \leqslant C(d^2 1_{G \text{ is non-regular}} + d 1_{G \text{ is regular but not vertex-transitive}} + 1)$$

Conjecture 8.3 (Monotonicity in the number of walkers)**.** *Denote by $\mathrm{SC}(G, k)$ the social connectivity time when we have k walkers, each starting at a vertex chosen from the stationary distribution π independently of the rest of the walkers. Then for every vertex-transitive graph G we have that $\mathbb{E}[\mathrm{SC}(G, k)]$ is monotonically non-increasing in k for $k \geqslant 2$.*

Open Problem 8.4. Let $G = (V, E)$ be a finite connected vertex-transitive graph. Let $o \in V$. Consider the SN model on G. Consider

- $t(G)$ - the minimal integer satisfying $t(G) \geqslant \log |V| \sum_{i=0}^{t(G)} P^i(o, o)$,
- τ_1 - the minimal time in which all classes are of size at least 2,
- and τ_2 - the minimal time by which every vertex has been visited by at least one walker.

Is there an absolute constant C such that

$$\max\{\mathbb{E}[\mathrm{SC}(G)], \mathbb{E}[\tau_1], \mathbb{E}[\tau_2], t(G)\} \leqslant C \min\{\mathbb{E}[\mathrm{SC}(G)], \mathbb{E}[\tau_1], \mathbb{E}[\tau_2], t(G)\} ?$$

Conjecture 8.5. *Consider a sequence G_n of vertex-transitive graphs of increasing sizes. Then in the notation of Open problem 8.4,*

$$\mathrm{SC}(G_n)/\mathbb{E}[\mathrm{SC}(G_n)] \to 1 \quad and \quad \tau_i(G_n)/\mathbb{E}[\tau_i(G_n)] \to 1 \ (i = 1, 2) \ in \ distribution, \ as \ n \to \infty.$$

We believe that in the setup of the previous conjecture it is often the case that $[\mathrm{SC}(G_n) - \tau_1(G_n)]/\mathbb{E}[\mathrm{SC}(G_n)] \to 0$ in distribution (e.g. we believe this is the case for transitive expanders). However, this might not be the case for the n-cycle. It is plausible that Conjecture 8.5 holds even without the transitivity assumption as long as the graphs G_n are of bounded degree. Conversely, if G_n is two disjoint n-cliques connected by a single edge, then $\frac{\mathrm{SC}(G_n)}{\mathbb{E}[\mathrm{SC}(G_n)]}$ converges in distribution to the Exponential distribution with parameter 1.

In the last example we have that $\mathbb{E}[\mathrm{SC}(G_n)] = \Theta(n)$. This shows $\mathbb{E}[\mathrm{SC}(G_n)]$ may increase with the maximal degree. Below are two families of graphs, generalizing the previous example, which we believe to be extremal.

Example 8.6. In general, $\mathbb{E}[\mathrm{SC}(G)]$ may grow linearly in d, even when G is regular, as the following example demonstrates. Consider two disjoint n-cliques. Delete an edge from each clique and connected the two cliques by two edges so that the obtained graph is $(n-1)$-regular. By considering the time required until some walker visits both cliques it is not hard to show that $\mathbb{E}[\mathrm{SC}] \geqslant cn$. We now describe a similar construction of d-regular graphs which satisfy $\mathbb{E}[\mathrm{SC}] \geqslant cd \log^2(n/d)$.

Fix some $2 \leqslant d, n$ such that $2d \leqslant n$. Let J_k be a graph obtained from the complete graph on k vertices by deleting a single edge. Consider $\lceil n/d \rceil$ disjoint copies of J_d: $I_0, \ldots, I_{\lceil n/d \rceil - 1}$, where for all $0 \leqslant j < \lceil n/d \rceil$, I_j is connected to I_{j+1} ($j+1$ is defined $\mod \lceil n/d \rceil$) by a single edge that connects two degree $d-1$ vertices. This can be done so that the obtained graph is d-regular. It is not hard to verify that the obtained graph satisfies $\mathbb{E}[\mathrm{SC}] \geqslant cd \log^2(n/d) =: s_{d,n} = s$ (in fact, if $d \ll n$ we have that w.h.p. $\mathrm{SC} \geqslant s$). To see this, observe that similarly to the analysis of the n-cycle it suffices to show that provided that c from the definition of s is taken to be sufficiently small, with probability bounded from below (in fact, w.h.p. when $d \ll n$) there will be multiples j's such that no walker crossed the edge connecting I_j with I_{j+1} by time s. To prove this, one can imitate the analysis from Remark 6.1. The claimed bound becomes intuitive when one considers the fact that the projection of the walk to the index of the copy of J_d to which it belongs, behaves similarly to a random walk on the $\lceil n/d \rceil$ cycle with holding probability $1 - \frac{1}{d^2}$. Thus the expected number of walkers to cross an edge connecting two copies of J_d by time s is comparable to d times the expected number of walkers which cross a certain edge by time s/d for the n-cycle (one has to multiply by d as J_d has d vertices).

We now consider a non-regular family of examples.

Example 8.7. Consider two disjoint n-cliques connected by a path of length n. Using the fact that the number of walkers which initially occupy the n-path has a Poisson distribution with a constant mean, it is not hard to show that $\mathbb{E}[\mathrm{SC}] \geqslant cn^2$.

We now describe a similar construction whose maximal degree is $d \leqslant 8|V|$ which satisfy $\mathbb{E}[\text{SC}] \geqslant c[d \log(|V|/d)]^2$. Take $\lceil n/d \rceil$ copies of J_d: $I_0, \ldots, I_{\lceil n/\ell \rceil - 1}$ and for all $0 \leqslant k < \lceil n/d \rceil - 1$ connect I_k to I_{k+1} via a path of length d.

To see that $\mathbb{E}[\text{SC}] \geqslant c[d \log(n/d)]^2 = t_{d,n}$, observe that with positive probability (in fact, w.h.p. if $d \ll n$) there will be many vertices lying in the middle of a length d path (connecting some I_k to I_{k+1}) which will not be visited by a single walker up to time $t_{d,n}$. We leave the details to the reader (hint: use the idea from Remark 6.1).

8.3 Additional Open Problems

The following conjecture concerns the existence of infinite classes of walkers after a constant number of steps in the setup of infinite graphs.

Conjecture 8.8. *Let G be a connected infinite bounded degree graph such that the critical density for independent percolation on G is strictly less than 1. If the number of walkers at time 0 at each vertex v has a $\text{Pois}(\lambda d_v)$ distribution, independently for different v's (let $\mathbb{P}_\lambda$ be the corresponding probability measure), then for every $\lambda > 0$ there exists $t_c(\lambda, G) > 0$ such that for all $t > t_c(\lambda, G)$,*

$$\mathbb{P}_\lambda[\text{there exists an infinite class of walkers at time } t] = 1.$$

Let K_n be the complete graph on n vertices. When the holding probabilities of the walks are taken to be either $q_n = 0$ or $q_n = 1/n$, one can show that after a single step the SN model (with $\text{Pois}(1)$ walkers per site) behaves in some sense like a critical Erdős Rényi random graph, $G(n, 1/n)$, while after two steps it behaves like the super-critical random graph $G(n, 2/n)$ (hence w.h.p. a giant class of walkers emerges after precisely 2 steps). We believe that also in the continuous-time setup there is a phase transition.

Open Problem 8.9. Consider the SN model on K_n in continuous-time (each walker has jump rate 1). What is the critical time t_c for the emergence of a giant class of walkers?

A Deferred proofs

A.1 Proof of Lemma 4.10

Lemma 4.10 Let $\xi_1, \ldots, \xi_m$ be independent Poisson random variables. Let $p_1, \ldots, p_m \in (0, 1)$. Denote $p_* := \max p_i$, $S := \sum_{i=1}^m p_i \xi_i$ and $\mu := \mathbb{E}[S]$. Then for all $\varepsilon \in (0, 1]$

$$P[S \leqslant (1 - \varepsilon)\mu] \leqslant e^{-\frac{1}{2}\mu\varepsilon^2/p_*}.$$
$$P[S \geqslant (1 + \varepsilon)\mu] \leqslant e^{-\frac{1}{4}\mu\varepsilon^2/p_*}.$$

Proof: Denote $\mu_i := \mathbb{E}[\xi_i]$ and $\lambda := \varepsilon/p_*$. As $\sum_k e^{-\lambda p_i k} \mu_i^k/k! = e^{\mu_i e^{-\lambda p_i}}$ and $\lambda p_i \leqslant 1$ (which implies that $e^{-\lambda p_i} - 1 \leqslant -\lambda p_i + \frac{1}{2}(\lambda p_i)^2$) we have that

$$\mathbb{E}[e^{-\lambda p_i \xi_i}] = \exp[\mu_i(e^{-\lambda p_i} - 1)] \leqslant \exp[\mu_i(-\lambda p_i + \frac{1}{2}(\lambda p_i)^2)] \leqslant \exp[\mu_i(-\lambda p_i + \frac{1}{2}\lambda^2 p_* p_i)],$$

$$\mathbb{E}[e^{-\lambda S}] = \prod_{i=1}^m \mathbb{E}[e^{-\lambda p_i \xi_i}] \leqslant \exp[\sum_{i=1}^m \mu_i(-\lambda p_i + \frac{1}{2}\lambda^2 p_* p_i)] = \exp[\mu(-\lambda + \frac{1}{2}\lambda^2 p_*)].$$

Finally, by Markov's inequality and the choice of λ

$$P[S \leqslant (1 - \varepsilon)\mu] = P[e^{-\lambda S} \geqslant e^{-\lambda(1-\varepsilon)\mu}] \leqslant \mathbb{E}[e^{-\lambda S}]e^{\lambda(1-\varepsilon)\mu} \leqslant e^{\mu(-\lambda\varepsilon + \frac{1}{2}\lambda^2 p_*)} = e^{-\frac{1}{2}\mu\varepsilon^2/p_*}.$$

Similarly, let $a := \varepsilon/(2p_*)$. Then $\sum_k e^{a p_i k} \mu_i^k/k! = e^{\mu_i e^{a p_i}}$. Since $a p_i \leqslant 1$ we have that $e^{a p_i} - 1 \leqslant a p_i + (a p_i)^2$ and so

$$\mathbb{E}[e^{a p_i \xi_i}] = \exp[\mu_i(e^{a p_i} - 1)] \leqslant \exp[\mu_i(a p_i + (a p_i)^2)] \leqslant \exp[\mu_i(a p_i + a^2 p_* p_i)],$$

$$\mathbb{E}[e^{aS}] = \prod_{i=1}^{m} \mathbb{E}[e^{ap_i\xi_i}] \leqslant \exp[\sum_{i=1}^{m} \mu_i(ap_i + a^2 p_* p_i)] = \exp[\mu(a + \lambda^2 p_*)].$$

Finally, by Markov's inequality and the choice of a we get that

$$\mathrm{P}[S \geqslant (1+\varepsilon)\mu] = \mathrm{P}[e^{aS} \geqslant e^{a(1+\varepsilon)\mu}] \leqslant \mathbb{E}[e^{aS}]e^{-a(1+\varepsilon)\mu} \leqslant e^{\mu(-a\varepsilon + a^2 p_*)} = e^{-\mu\varepsilon^2/(4p_*)}. \qquad \square$$

A.2 Proof of Lemma 4.2

Lemma 4.2 Let J_k be a non-increasing $\mathbb{Z}_+$-valued random process, measurable w.r.t. filtration $(\mathcal{F}_k)_{k \geqslant 0}$, with $J_0 \leqslant m$, satisfying that for some $0 < \alpha, p, \varepsilon_m < 1$, for each k there exists some events $A_k \in \mathcal{F}_k$ with $\mathrm{P}(A_k) \geqslant 1 - \varepsilon_m$ satisfying that

$$\mathbb{E}[1_{\{J_{k+1} \leqslant (1-p/2)J_k\}} 1_{A_k} \mid \mathcal{F}_k] \geqslant \alpha 1_{A_k}.$$

Let $T_0 := \inf\{k : J_k = 0\}$. Then there exists some constant $C_{\alpha,p} \leqslant K/(\alpha p)$ such that

$$\mathrm{P}[T_0 > \lceil C_{\alpha,p} \log m \rceil] \leqslant m^{-2} + \varepsilon_m(\lceil C_{\alpha,p} \log m \rceil + 1).$$

Proof: Denote the complement of A_k by B_k. Let $\tau := \inf\{k : 1_{B_k} = 1\}$. Define $I_k = J_k$ for $k < \tau$ and $I_k = 0$ for $k \geqslant \tau$. Denote $L := \lceil C_{\alpha,p} \log m \rceil$, where $C_{\alpha,p}$ shall be determined soon. Let $T_0' := \inf\{k : I_k = 0\}$. By a union bound over $\bigcup_{i=0}^{L} B_k$ it suffices to show that

$$\mathrm{P}[T_0' \geqslant L] \leqslant m^{-2}$$

(since $\mathrm{P}[T_0 > L] - \mathrm{P}[T_0' > L] \leqslant \mathrm{P}(\bigcup_{i=0}^{L} B_k)$). Let $a := 1/(1 - \frac{1}{2}\alpha p)$. By (4.2) $M_k := a^k I_k$ is a super-martingale (w.r.t. the filtration $(\mathcal{F}_k)_{k \geqslant 0}$). In particular, for $C_{\alpha,p} := 3/\log a$ we have

$$\mathbb{E}[I_L] \leqslant a^{-L}\mathbb{E}[M_0] \leqslant m^{-3}\mathbb{E}[J_0] \leqslant m^{-2}.$$

By Markov's inequality

$$\mathrm{P}[T_0' \geqslant L] = \mathrm{P}[I_L \geqslant 1] \leqslant \mathbb{E}[I_L] \leqslant m^{-2}. \quad \square$$

A.3 Proof of Lemma 4.14

Lemma 4.14 Let $G = (V, E)$ be a finite Eulerian graph. Let $Y_v(t)$ be the number of walkers which occupy vertex v at time t, where $(Y_v(0))_{v \in V}$ is some arbitrary deterministic initial configuration of walkers and different walkers perform independent LSRWs. Let $(v, u) \in E$. Let $Y := \max_{x \in V} \frac{d_v}{d_x} Y_x(0)$. Then

$$2d_v\mathbb{E}[Y_u(t)] \geqslant \mathbb{E}[Y_v(t-1)] \geqslant \frac{1}{3}(\mathbb{E}[Y_v(t)] - Y\exp(-c_1 t)). \tag{A.1}$$

Proof: The first inequality in (A.1) is trivial. We now prove the second inequality. Let $v \in V$ and $t \in \mathbb{N}$. We say that a path γ' is non-lazy if $\gamma'(i) \neq \gamma'(i+1)$ for all i. Let $\hat{\Gamma}_\ell$ be the collection of all non-lazy paths of length at most ℓ which terminate at v. For a walk γ let γ' be its projection to its non-lazy version (obtained by omitting from γ coordinates with the same value as that of the previous time unit).

For every $\gamma' \in \hat{\Gamma}_t$ (resp. $\gamma' \in \hat{\Gamma}_{t-1}$) let $W_{\gamma'}(t)$ (resp. $W_{\gamma'}(t-1)$) be the collection of walks of length t (resp. $t-1$) whose non-lazy version is γ'. Let $\hat{P}$ be the transition matrix of the non-lazy walk (i.e. $P = \frac{1}{2}(I + \hat{P})$, where I is the identity matrix). Let $b_k := \mathbb{P}[\mathrm{Bin}(t, \frac{1}{2}) = k]$ and $a_k := \mathbb{P}[\mathrm{Bin}(t, \frac{1}{2}) > k]$. We first note that the contribution to $\mathbb{E}[Y_v(t)]$ coming from walks in $\cup_{\gamma' \in \hat{\Gamma}_t \setminus \hat{\Gamma}_{\frac{2t}{3}}} W_{\gamma'}(t)$ (i.e. from walks γ's with less than $t/3$ lazy steps) is at most

$$\sum_{k>2t/3} b_k \sum_x Y_x(0)\hat{P}^k(x,v) \leqslant \sum_{k>2t/3} b_k \frac{Y}{d_v} \sum_x d_x \hat{P}^k(x,v) = a_{2t/3} Y \leqslant Y e^{-c_1 t}, \tag{A.2}$$

where we have used $\sum_{k>2t/3} b_k \leqslant e^{-c_1 t}$, and $\pi \hat{P}^k = \pi$ to argue that $\sum_x d_x \hat{P}^k(x,v) = d_v$.

Recall (§3) that $p(\gamma)$ is the probability that a walker starting from $\gamma(0)$ performs the walk γ. Let $\gamma' \in \hat{\Gamma}_{\frac{2t}{3}}$. We argue that the total contribution to $\mathbb{E}[Y_v(t)]$ coming from walks in $W_{\gamma'}(t)$ (which equals $\sum_{\gamma \in W_{\gamma'}(t)} p(\gamma) Y_{\gamma(0)}(0)$) is larger than the contribution to $\mathbb{E}[Y_v(t-1)]$ coming from walks in $W_{\gamma'}(t-1)$ (which equals $\sum_{\gamma \in W_{\gamma'}(t-1)} p(\gamma) Y_{\gamma(0)}(0)$) by a factor of at most 3. Since the total number of lazy steps performed is at least $t/3$, the "cost" $\frac{\sum_{\gamma \in W_{\gamma'}(t)} p(\gamma)}{\sum_{\gamma \in W_{\gamma'}(t-1)} p(\gamma)}$ of making one less lazy step is bounded from above by

$$\sup_{k \in \mathbb{N}: k \in [t/3, t]} \binom{t}{k} / \binom{t}{k-1} = \sup_{k \in [t/3, t]} \frac{t - k + 1}{k} \leqslant 3.$$

Putting everything together gives:

$$\mathbb{E}[Y_v(t-1)] \geqslant \sum_{\gamma' \in \hat{\Gamma}_{2t/3}} \sum_{\gamma \in W_{\gamma'}(t-1)} p(\gamma) Y_{\gamma(0)}(0) \geqslant \frac{1}{3} \sum_{\gamma' \in \hat{\Gamma}_{2t/3}} \sum_{\gamma \in W_{\gamma'}(t)} p(\gamma) Y_{\gamma(0)}(0).$$

By (A.2) the rightmost term above is at least $\frac{1}{3}(\mathbb{E}[Y_v(t)] - Y \exp(-c_1 t))$, as desired. $\qquad\square$

References

[1] David Aldous and Jim Fill, *Reversible Markov chains and random walks on graphs*, 2002, Unfinished manuscript. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[2] Noga Alon, Itai Benjamini, and Alan Stacey, *Percolation on finite graphs and isoperimetric inequalities*, Ann. Probab. **32** (2004), no. 3A, 1727–1745. MR-2073175

[3] O. S. M. Alves, F. P. Machado, and S. Yu. Popov, *Phase transition for the frog model*, Electron. J. Probab. **7** (2002), no. 16, 21. MR-1943889

[4] O. S. M. Alves, F. P. Machado, and S. Yu. Popov, *The shape theorem for the frog model*, Ann. Appl. Probab. **12** (2002), no. 2, 533–546. MR-1910638

[5] Itai Benjamini, Luiz Renato Fontes, Jonathan Hermon, and Fabio Pabio Machado, *On an epidemic model on finite graphs*, Arxiv preprint arXiv:1610.04301 (2016).

[6] Lucas Boczkowski, Yuval Peres, and Perla Sousi, *Sensitivity of mixing times in Eulerian digraphs*, SIAM J. Discrete Math. **32** (2018), no. 1, 624–655. MR-3775128

[7] Devdatt Dubhashi and Desh Ranjan, *Balls and bins: a study in negative dependence*, Random Structures Algorithms **13** (1998), no. 2, 99–124. MR-1642566

[8] Sharad Goel, Ravi Montenegro, and Prasad Tetali, *Mixing time bounds via the spectral profile*, Electron. J. Probab. **11** (2006), no. 1, 1–26. MR-2199053

[9] Jonathan Hermon, *Frogs on trees?*, Electron. J. Probab. **23** (2018), Paper No. 17, 40. MR-3771754

[10] Jonathan Hermon, Ben Morris, Chuan Qin, and Allan Sly, *The social network model on infinite graphs*, Arxiv preprint arXiv:1610.04293 (2016).

[11] Jonathan Hermon and Richard Pymar, *The exclusion process mixes (almost) faster than independent particles*, arXiv preprint arXiv:1808.10846 (2018).

[12] Christopher Hoffman, Tobias Johnson, and Matthew Junge, *Recurrence and transience for the frog model on trees*, Ann. Probab. **45** (2017), no. 5, 2826–2854. MR-3706732

[13] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR-1876169

[14] Harry Kesten and Vladas Sidoravicius, *The spread of a rumor or infection in a moving population*, Ann. Probab. **33** (2005), no. 6, 2402–2462. MR-2184100

[15] Harry Kesten and Vladas Sidoravicius, *A shape theorem for the spread of an infection*, Ann. of Math. **167** (2008), no. 3, 701–766. MR-2415386

[16] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2017, Second edition of [MR2466937], With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson. MR-3726904

[17] S. Yu. Popov, *Frogs in random environment*, J. Statist. Phys. **102** (2001), no. 1-2, 191–201. MR-1819703

[18] András Telcs and Nicholas C. Wormald, *Branching and tree indexed random walks on fractals*, J. Appl. Probab. **36** (1999), no. 4, 999–1011. MR-1742145

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm